



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Πτυχιακή εργασία

From digital footprints to facts: mining social data for marketing practices

Λιναράκης Δημήτριος - Ιωάννης

2022201600096

Βλάχος Στέφανος

2022201600019

Επιβλέπουσα:

Παρασκευή Ραυτοπούλου

ΕΔΙΠ

Τρίπολη, Απρίλιος 2021

Περίληψη

Στη σημερινή εποχή, είναι κοινώς αποδεκτό πως τα μέσα κοινωνικής δικτύωσης έχουν καταλάβει το μεγαλύτερο φάσμα της καθημερινότητας του ανθρώπου. Το πλήθος των δυνατοτήτων που προσφέρουν στους χρήστες έχει διαμορφώσει σε μεγάλο βαθμό τις διαπροσωπικές σχέσεις αλλά και πληθώρα καθημερινών δραστηριοτήτων τόσο σε προσωπικό, όσο και σε επαγγελματικό επίπεδο. Όλο και περισσότερες γίνονται οι επιχειρήσεις που χρησιμοποιούν το διαδίκτυο για να διαφυλάξουν την ομαλή λειτουργία τους και, παράλληλα, να διαφημίσουν τα προϊόντα τους σε ένα ευρύτερο κοινό απ' ό,τι θα τους επέτρεπε κάποιο φυσικό μέσο. Επιπρόσθετα, ένας μέσος χρήστης έρχεται συχνά σε επαφή με πολύ μεγάλο όγκο περιεχομένου (*content*). Μπορεί κανείς, λοιπόν, να αντιληφθεί πως οι πλατφόρμες δικτύωσης αποτελούν ίσως το ιδανικότερο μέσο διαφημιστικής δραστηριότητας κατά τη τρέχουσα χρονική περίοδο. Η δραστηριότητα αυτή μπορεί να ταξινομηθεί σε δύο βασικές κατηγορίες. Η πρώτη, είναι εκείνη όπου οι διαφημίσεις (*advertisements*) γίνονται απευθείας από την επιχείρηση, η οποία επιθυμεί να προβάλλει τα προϊόντα της η ίδια μέσω των προφίλ της, με χρήση φωτογραφιών και βίντεο. Η δεύτερη κατηγορία αφορά στη διαφήμιση επιχειρήσεων μέσω ατόμων, τα οποία ασκούν υψηλή επιρροή στο κοινό που τα ακολουθεί. Τα άτομα αυτά (γνωστά ως *Influencers*) συνήθως είναι διάσημα πρόσωπα, όπως καλλιτέχνες, αθλητές ή μοντέλα, κι έχουν έναν αξιοσημείωτο αριθμό ακολούθων στα κοινωνικά τους προφίλ. Αυτό το είδος προβολής είναι γνωστό στην παγκόσμια γλώσσα ως *Influencer Marketing* και αποτελεί μεγάλη τάση στον σύγχρονο διαδικτυακό κόσμο.

Το ενδιαφέρον που παρουσιάζουν τα παραπάνω δεδομένα, σε συνδυασμό με το γεγονός ότι μέχρι σήμερα δεν υπάρχουν αρκετά εμπειριστατωμένες έρευνες όσον αφορά στο ελληνικό κοινό, συνέβαλαν στο να αποτελέσει κύριο αντικείμενο της παρούσας μελέτης η συμπεριφορά των Ελλήνων χρηστών στα μέσα κοινωνικής δικτύωσης. Συγκεκριμένα, τέθηκε ως στόχος να αποκωδικοποιηθούν οι προτιμήσεις του κοινού και το ποσοστό επιρροής του από τις διάφορες μορφές των αναρτημένων δημοσιεύσεων, όπως επίσης και να σκιαγραφηθούν οι άξονες, πάνω στους οποίους θα μπορούσαν οι εταιρείες και οι *Influencers* να κατευθύνουν με μεγαλύτερη ευστοχία το περιεχόμενο τους. Αρχικά, χρειάστηκε να επιλεγθούν οι πλατφόρμες δικτύωσης που θα αποτελέσουν τη βάση της έρευνας, κάτι το οποίο δημιούργησε προβληματισμούς, καθώς αυτές πολλαπλασιάζονται καθημερινά με ταχείς ρυθμούς. Το κριτήριο που καθόρισε την επιλογή αυτή, ήταν η απήχηση της εκάστοτε πλατφόρμας κατά τη χρονική περίοδο 2020-2021. Με γνώμονα το παραπάνω, αποφασίστηκε πως τα Instagram και YouTube πληρούσαν όλες τις προδιαγραφές, έτσι ώστε να προκύψει μια ολοκληρωμένη και πολύπλευρη έρευνα γύρω από την ψηφιακή δραστηριοποίηση της ελληνικής κοινότητας.

Στα πλαίσια της παρούσας εργασίας, στήθηκε και χρησιμοποιήθηκε ένας μηχανισμός εξαγω-

γής δεδομένων από το Instagram και το YouTube. Τα δεδομένα αποθηκεύτηκαν στη MongoDB NoSQL βάση δεδομένων και στη συνέχεια, αφού μελετήθηκαν τα ποσοτικά και τα ποιοτικά χαρακτηριστικά των δεδομένων, αναπτύχθηκαν διαγράμματα, τα οποία βοήθησαν στην εξαγωγή συμπερασμάτων γύρω από το θέμα της έρευνας. Επιπλέον, διεξήχθη ερωτηματολόγιο, το οποίο είχε ως σκοπό να συμπληρώσει και να εμπλουτίσει τα αποτελέσματα, με γνώση που δεν ήταν δυνατό να εξαχθεί μέσω των δεδομένων που αντλήθηκαν από τα προφίλ στα κοινωνικά δίκτυα.

Abstract

Nowadays, it is commonly accepted that social media have conquered the widest spectrum of humans' daily life. The multitude of abilities that are provided to the users has greatly shaped the interpersonal relationships as well as a large variety of everyday activities, both on a personal and a business level. More and more companies use the internet to support their activities and, in parallel, to advertise their products to a larger audience than any physical means would allow them. Furthermore, an average user experiences a high volume of content daily. Therefore, it is well perceived that social networks have turned to the most ideal mean for advertising in the current period. The advertising activity through social networks can be divided into two main categories. The first one consists of businesses that conduct their advertisements using their social profiles, through images and videos. On the opposite side, there are businesses that choose to promote their products by collaborating with content creators, called Influencers, as they have a remarkable number of followers in their social network profiles and seem to have influence on their audience. These creators are usually celebrities, such as artists, athletes or models. This type of promotion is known worldwide as Influencers Marketing and serves as a huge trend in the modern digital era.

The shift in the marketing practices, in combination with the fact that there are not detailed enough researches concerning the behaviour of the Greek public in the social media, contributed to the subject of research of this study. In particular, the principal objective was, on one hand, to decode the users' preferences and the degree of the influence that they receive from the various types of online publications and, on the other hand, to outline the axes on which both the companies and the Influencers could direct their content more accurately. Initially, the platforms that would form the basis of the study should be chosen. However, the rapid growth of newly established platforms was a serious challenge. The basic criterion which determined this selection was the impact of each platform on the audience during the period 2020-2021. Based on the above, it was decided that Instagram and YouTube satisfied all the requirements that would lead to comprehensive research on the activity of the Greek community on social media.

For the work presented in this thesis, tools and techniques were applied, parametrised and used to scrape Instagram and YouTube publicly available profiles. The data upon retrieved was stored in MongoDB NoSQL database and, consequently, based on the qualitative and the quantitative properties of the data, charts were developed to draw conclusions on the topic of this research. Finally, a questionnaire was also prepared and made available to collect the user preferences concerning their shopping practices and the way they receive advertisements through social media.

Ευχαριστίες

Πρώτα απ' όλα, θα θέλαμε να ευχαριστήσουμε την επιβλέπουσα της πτυχικής μας, Παρασκευή Ραυτοπούλου, για την εμπιστοσύνη που μας έδειξε και την καθοδήγησή της όλη αυτή τη χρονική περίοδο. Είναι αξιοσημείωτο το γεγονός ότι σε κάθε μας δυσκολία ήταν πάντα διαθέσιμη και πρόσχαρη στο να μας βοηθήσει. Την ευχαριστούμε, επίσης, για τις πρωτοβουλίες που μας επέτρεψε να πάρουμε και για τις ευκαιρίες που μας έχει προσφέρει μέχρι σήμερα.

Θα θέλαμε, επίσης, να ευχαριστήσουμε ο ένας τον άλλον για την υπομονή και την αφοσίωση απέναντι στον κοινό μας στόχο, αλλά και για την αμοιβαία εμπιστοσύνη όλων αυτόν τον καιρό. Η συνεργασία ήταν άψογη και υπήρχε κατανόηση και χώρος για όλες τις απόψεις και τις ιδέες.

Ένα μεγάλο ευχαριστώ στις οικογένειές μας για την υπομονή και τις θυσίες τους καθ' όλη τη διάρκεια της φοιτητικής μας ζωής, που παρά τις όλες τις δυσκολίες που παρουσιάστηκαν μας προσέφεραν την ευκαιρία να σπουδάσουμε και να διευρύνουμε τους ορίζοντές μας. Χωρίς αυτές δεν θα είχαμε φτάσει ως εδώ και σίγουρα τους οφείλουμε ένα μεγάλο μερίδιο των επιτυχιών μας.

Δεν θα μπορούσαμε να παραλείψουμε τους φίλους και συμφοιτητές μας, με τους οποίους ζήσαμε ανεκτίμητες στιγμές. Σε κάθε δυσκολία ήταν εκεί για να δώσουν μια χιουμοριστική χροιά στις καταστάσεις. Είμαστε ευγνώμονες για όλα τα συναισθήματα που ανταλλάξαμε και είμαστε σίγουροι πως έχουμε πολλά ακόμα να ζήσουμε μαζί.

Περιεχόμενα

1	Εισαγωγή	1
1.1	Ανάλυση προβλήματος	1
1.2	Η συνεισφορά μας στο πρόβλημα	2
1.3	Οργάνωση της πτυχιακής εργασίας	3
2	Σχετικό Υλικό	5
2.1	Web Crawling και Web Scraping	5
2.1.1	Web Crawling έναντι Web Scraping	6
2.1.2	Η λειτουργία των Web Crawlers και Web Scrapers	7
2.1.3	Τεχνικές Web Scraping	8
2.1.4	Διαθέσιμα Εργαλεία	12
2.1.5	Προκλήσεις κατά το Web Scraping	13
2.1.6	Πολιτική προστασίας δεδομένων ενάντια στο Web Scraping	15
2.2	Βάσεις Δεδομένων	20
2.2.1	SQL έναντι NoSQL	21
2.2.2	Localhost έναντι Cloud-based	22
2.3	Μελέτη δεδομένων	22
2.3.1	Τα στάδια της ερευνητικής διαδικασίας	23
2.3.2	Προκλήσεις κατά την ερευνητική διαδικασία	24
2.3.3	Διεξαγωγή ερωτηματολογίου	25
3	Εργαλεία που χρησιμοποιήθηκαν	27
3.1	Scrapy Framework	27
3.1.1	Επισκόπηση των δυνατοτήτων του Scrapy	27
3.1.2	Αρχιτεκτονική του Scrapy	28
3.1.3	Βασική δομή ενός Scrapy πρότζεκτ	30
3.2	YouTube Data API	31
3.3	MongoDB	35
3.3.1	MongoDB Compass	36
3.3.2	MongoDB Atlas	37
3.3.3	MongoDB Charts	37
3.4	Google Forms	37
4	Συγκέντρωση και αποθήκευση των δεδομένων της έρευνας	39
4.1	Συγκομιδή στοιχείων του Instagram	39

4.1.1	Δομή του Instagram και δεδομένα	39
4.1.2	Κώδικας Instagram	46
4.2	Συγκομιδή στοιχείων του YouTube	60
4.2.1	Δομή του YouTube και δεδομένα	61
4.2.2	Κώδικας YouTube	67
4.3	Συγκέντρωση δεδομένων και προκλήσεις	82
5	Μελέτη δεδομένων	85
5.1	Διερεύνηση του Instagram	85
5.1.1	Ανάλυση των συγκεντρωμένων δεδομένων	85
5.2	Διερεύνηση του YouTube	138
5.2.1	Ανάλυση των συγκεντρωμένων δεδομένων	139
6	Μελέτη ερωτηματολογίου	197
6.1	Ερωτήσεις Instagram	200
6.1.1	Πόσο καιρό χρησιμοποιείτε το Instagram;	200
6.1.2	Πόσο συχνά χρησιμοποιείτε το Instagram;	202
6.1.3	Ακολουθείτε Influencers;	204
6.1.4	Αν ναι, τι πιστεύετε ότι σας προσφέρει το περιεχόμενο των Influencers;	206
6.1.5	Τι θεματικό περιεχόμενο σας αρέσει να παρακολουθείτε στο Instagram;	207
6.1.6	Προτιμάτε να επισκέπτεστε εξωτερικούς συνδέσμους (links) μέσω swipe up στα stories ή μέσω link που βρίσκεται στις πληροφορίες ενός προφίλ (Bio);	208
6.1.7	Έχετε αναζητήσει/αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση;	209
6.1.8	Έστω ότι δύο Instagram προφίλ, ένα με 500 followers και ένα με 10.000 followers παρουσιάζουν το ίδιο προϊόν. Την άποψη ποιου από τα δύο προφίλ θα εμπιστευτείτε;	212
6.1.9	Σας ενοχλεί η ύπαρξη χορηγούμενων δημοσιεύσεων;	213
6.1.10	Επισκέπτεστε τα προφίλ χορηγούμενων δημοσιεύσεων;	214
6.1.11	Κάνετε Like/Comment σε δημοσιεύσεις λογαριασμών που δεν ακολουθείτε;	215
6.1.12	Έχετε συμμετάσχει ποτέ σε κάποιο giveaway;	216
6.1.13	Αν ναι, πώς ανακαλύψατε αυτό το giveaway;	218
6.1.14	Αναζητείτε πληροφορίες, για προϊόντα που ενδιαφέρεστε να αγοράσετε, στο Instagram;	219
6.1.15	Ποια θεωρείτε ότι είναι η ιδανική συχνότητα αναρτήσεων;	220
6.1.16	Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο (timeline) ή μέσω των instagram ιστοριών (stories);	221
6.1.17	Έχετε ανακαλύψει ποτέ νέες δημοσιεύσεις αναζητώντας κάποιο hashtag στο Instagram;	223
6.1.18	Κάνετε repost (επαναδημοσίευση) δημοσιεύσεις που σας αρέσουν;	224
6.1.19	Τι κοιτάτε πρώτα σε μία δημοσίευση (timeline post);	225
6.1.20	Ποια μορφή δημοσίευσης (timeline post) προτιμάτε;	226
6.1.21	Ακολουθείτε τους/τις αγαπημένους/νες σας YouTubers στο Instagram;	227
6.2	Ερωτήσεις YouTube	229

6.2.1	Πόσο καιρό χρησιμοποιείτε το YouTube;	229
6.2.2	Γνωρίζετε ότι το YouTube αποτελεί ένα ανερχόμενο μέσο κοινωνικής δικτύωσης, κάνοντας διαθέσιμη τη δημιουργία προσωπικών post και stories; 231	
6.2.3	Πόσο συχνά παρακολουθείτε βίντεο στο YouTube;	233
6.2.4	Τι περιεχόμενο προτιμάτε να παρακολουθείτε στο YouTube;	235
6.2.5	Με ποιον τρόπο ανακαλύπτετε νέα βίντεο;	236
6.2.6	Τι κοιτάτε πρώτα σε ένα βίντεο;	239
6.2.7	Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών;	240
6.2.8	Κάνετε Like/Comment στα βίντεο που σας αρέσουν;	243
6.2.9	Κοινοποιείτε βίντεο που σας αρέσουν σε κάποιο άλλο μέσο κοινωνικής δικτύωσης (πχ Facebook, Instagram);	244
6.2.10	Παρακολουθείτε βίντεο-reviews (κριτικές) στο YouTube για προϊόντα που σας ενδιαφέρουν να αγοράσετε;	246
6.2.11	Έχετε αναζητήσει/αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube;	247
6.2.12	Έστω ότι δύο YouTube κανάλια, ένα με 500 subscribers και ένα με 10.000 subscribers διαφημίζουν το ίδιο προϊόν. Ποιο από τα δύο κανάλια θα εμπιστευόσασταν;	250
6.2.13	Έχετε συμμετάσχει ποτέ σε κάποιο YouTube διαγωνισμό;	250
6.2.14	Αν ναι, πώς ανακαλύψατε αυτό το Giveaway;	251
6.2.15	Έχετε αναζητήσει/επισκεφθεί ποτέ κάποιο προϊόν/βίντεο που εμφανίζεται σε διαφήμιση στο YouTube;	252
6.2.16	Έχετε αποχωρήσει ποτέ από βίντεο λόγω αναδυόμενων διαφημίσεων;	253
6.2.17	Κάθε πότε προτιμάτε να αναρτά βίντεο κάποιος δημιουργός;	254
6.2.18	Έχετε επισκεφθεί ποτέ κάποιο προφίλ του Instagram που ανακαλύψατε μέσω του YouTube;	255
7	Συμπεράσματα και Μελλοντική Έρευνα	259
7.1	Συμπεράσματα	259
7.1.1	Συμπεράσματα σχετικά με το Instagram	259
7.1.2	Συμπεράσματα σχετικά με το YouTube	261
7.1.3	Γενικότερες παρατηρήσεις	263
7.1.4	Επιχειρηματικός κλάδος και Influencer Marketing στα μέσα κοινωνικής δικτύωσης	264
7.1.5	COVID-19 και μέσα κοινωνικής δικτύωσης	265
7.2	Μελλοντική Έρευνα	266
	Βιβλιογραφία	267

Κατάλογος πινάκων

2.1	Web Crawling έναντι Web Scraping	6
2.2	Manual Web Scraping	9
2.3	Automated Web Scraping	9
2.4	Κύριες διαφορές μεταξύ SQL και NoSQL βάσεων δεδομένων	22
2.5	Πλεονεκτήματα και μειονεκτήματα της τοπικής αποθήκευσης δεδομένων	23
2.6	Πλεονεκτήματα και μειονεκτήματα της αποθήκευσης δεδομένων στο “νέφος” .	23

Κατάλογος σχημάτων

2.1	Βασική δομή ενός HTML αρχείου	10
2.2	Βασική δομή του δέντρου DOM	10
2.3	Βασική δομή του XPath μοντέλου	11
2.4	Το “robots.txt” αρχείο του YouTube	16
2.5	Διαφορετικά είδη CAPTCHAs	17
2.6	Το πεδίο User Agent που αντιστοιχεί στον περιηγητή Mozilla Firefox	18
2.7	Κεφαλίδες HTTP αιτήματος προς ένα προφίλ του Instagram	19
3.1	Scrapy αρχιτεκτονική και ροή δεδομένων μεταξύ των δομικών στοιχείων του	29
3.2	Η εξωτερική δομή ενός Scrapy πρότζεκτ	30
3.3	Η εσωτερική δομή ενός Scrapy πρότζεκτ	31
3.4	Δημιουργία νέου πρότζεκτ	32
3.5	Υποβολή ονόματος του πρότζεκτ	33
3.6	Αναδυόμενη λίστα με όλα τα διαθέσιμα πρότζεκτ του λογαριασμού	33
3.7	Ενεργοποίηση του εργαλείου YouTube Data API	34
3.8	Επιλογή της καρτέλας Credentials	34
3.9	Δημιουργία ενός κλειδιού API	35
3.10	Μορφή ενός MongoDB εγγράφου	35
3.11	Μορφή ενός ερωτήματος σε Aggregation Pipeline	36
4.1	Διαστρωμάτωση ενός λογαριασμού Instagram	40
4.2	Βασική δομή μίας δημοσίευσης του Instagram	42
4.3	Επιθεώρηση του πηγαίου κώδικα της σελίδας	43
4.4	Εύρεση του HTML στοιχείου script	44
4.5	Η JSON δομή ενός λογαριασμού Instagram	44
4.6	Επιθεώρηση του πηγαίου κώδικα της σελίδας	56
4.7	Επιλογή της καρτέλας Δίκτυο	57
4.8	Επιλογή διαδικτυακού στοιχείου	57
4.9	Εύρεση του αιτήματος στη λίστα των απεσταλμένων αιτημάτων	57
4.10	Κεφαλίδες του αιτήματος	58
4.11	Διαστρωμάτωση ενός λογαριασμού YouTube	62
4.12	Βασική δομή ενός βίντεο του YouTube	63
4.13	Ενδεχόμενη διάρθρωση μιας απόκρισης του YouTube Data API	64
5.1	Κατηγορίες χρηστών του Instagram	86
5.2	Κορυφαίες κατηγορίες λογαριασμών (Άνδρες)	87

5.3	Κορυφαίες κατηγορίες λογαριασμών (Γυναίκες)	87
5.4	Κορυφαίες κατηγορίες λογαριασμών (Επιχειρήσεις)	88
5.5	Κορυφαίες κατηγορίες λογαριασμών (Άλλο)	88
5.6	Κατανομή του αριθμού των ακολούθων (κάτω των 500K ακολούθων)	89
5.7	Κατανομή του αριθμού των ακολούθων (άνω των 500K ακολούθων)	90
5.8	Κατηγορίες χρηστών με τους περισσότερους ακόλουθους	91
5.9	Κατηγορίες περιεχομένου με τους περισσότερους ακόλουθους	91
5.10	Αριθμός λογαριασμών με βάση τον αριθμό των ακολούθων	92
5.11	Σχέση μεταξύ αριθμού ακολούθων και ακολουθούμενων	92
5.12	Μέσος αριθμός ακολουθούμενων ανά κατηγορία χρηστών	93
5.13	Κατανομή του ER Post	94
5.14	Κατανομή του ER Comments Post	94
5.15	Κατανομή του ER View	95
5.16	Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (0K-20K ακόλουθοι)	96
5.17	Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (20K-100K ακόλουθοι)	97
5.18	Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (100K-280K ακόλουθοι)	97
5.19	Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (ακόλουθοι > 280K)	98
5.20	Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (γενικό σύνολο)	99
5.21	Ποσοστό αλληλεπίδρασης ανά κατηγορία χρηστών	99
5.22	Ποσοστό αλληλεπίδρασης ανά αριθμό ακολούθων (0K-20K ακόλουθοι)	100
5.23	Ποσοστό αλληλεπίδρασης ανά αριθμό ακολούθων (20K-100K ακόλουθοι)	101
5.24	Ποσοστό αλληλεπίδρασης ανά αριθμό ακολούθων (100K-280K ακόλουθοι)	101
5.25	Ποσοστό αλληλεπίδρασης ανά αριθμό ακολούθων (ακόλουθοι > 280K)	102
5.26	Ποσοστό αλληλεπίδρασης ανά θεματική κατηγορία λογαριασμού	103
5.27	Αριθμός δημοσιεύσεων ανά κατηγορία χρηστών	104
5.28	Αριθμός δημοσιεύσεων ανά θεματική κατηγορία λογαριασμού	105
5.29	Σχέση μεταξύ αριθμού ακολούθων-συχνότητα αναρτήσεων	105
5.30	Συχνότητα ανάρτησης δημοσιεύσεων ανά κατηγορία χρηστών	106
5.31	Σχέση μεταξύ συχνότητας αναρτήσεων-ποσοστού αλληλεπίδρασης	107
5.32	Αριθμός δημοσιεύσεων και μέσου ποσοστού αλληλεπίδρασης ανά εβδομάδα του 2020	108
5.33	Αριθμός δημοσιεύσεων ανά μήνα και κατηγορία χρηστών	109
5.34	Αριθμός δημοσιεύσεων ανά μέρα/ώρα (Άνδρες, Γυναίκες)	110
5.35	Αριθμός δημοσιεύσεων ανά μέρα/ώρα (Επιχειρήσεις)	111
5.36	Αριθμός δημοσιεύσεων ανά μέρα/ώρα (Άλλο)	111
5.37	Είδη δημοσίευσης ανά κατηγορία χρηστών	112
5.38	Ποσοστό αλληλεπίδρασης ανά κατηγορία δημοσιεύσεων	113
5.39	Σχέση ποσοστού ER Views και θέσης του βίντεο σε μία δημοσίευση τύπου slideshow 114	
5.40	Μετρικές απόδοσης ανά θεματικό περιεχόμενο δημοσίευσης (Giveaway, Sponsorship, Other)	115

5.41	Αριθμός διαγωνισμών (Giveaways) ανά κατηγορία χρηστών	116
5.42	Αριθμός χορηγιών (Sponsorships) ανά κατηγορία χρηστών	117
5.43	Αριθμός διαγωνισμών και χορηγιών ανά θεματική κατηγορία λογαριασμού . .	117
5.44	Αριθμός διαγωνισμών (Giveaways) ανά μήνα του 2020	118
5.45	Αριθμός χορηγιών (Sponsorships) ανά μήνα του 2020	119
5.46	Σχέση μεταξύ αριθμού επισυναπτόμενων χρηστών και ποσοστού αλληλεπίδρασης	120
5.47	Δημοφιλέστεροι επισυναπτόμενοι χρήστες (Άνδρες)	121
5.48	Δημοφιλέστεροι επισυναπτόμενοι χρήστες (Γυναίκες)	122
5.49	Δημοφιλέστεροι επισυναπτόμενοι χρήστες (Επιχειρήσεις)	123
5.50	Δημοφιλέστεροι επισυναπτόμενοι χρήστες (Άλλο)	123
5.51	Αριθμός hashtags ανά αριθμό ακολούθων	124
5.52	Σχέση μεταξύ ποσοστού αλληλεπίδρασης και αριθμού hashtags (0K-20K ακό- λουθοι)	125
5.53	Σχέση μεταξύ ποσοστού αλληλεπίδρασης και αριθμού hashtags (20K-100K ακό- λουθοι)	125
5.54	Σχέση μεταξύ ποσοστού αλληλεπίδρασης και αριθμού hashtags (100K-280K ακό- λουθοι)	126
5.55	Σχέση μεταξύ ποσοστού αλληλεπίδρασης και αριθμού hashtags (ακόλουθοι > 280K)	126
5.56	Δημοφιλέστερα hashtags ανά ανδρικούς λογαριασμούς	128
5.57	Δημοφιλέστερα hashtags ανά δημοσιεύσεις ανδρικών λογαριασμών	128
5.58	Δημοφιλέστερα hashtags ανά γυναικίους λογαριασμούς	129
5.59	Δημοφιλέστερα hashtags ανά δημοσιεύσεις γυναικείων λογαριασμών	130
5.60	Δημοφιλέστερα hashtags ανά επιχειρησιακούς λογαριασμούς	131
5.61	Δημοφιλέστερα hashtags ανά δημοσιεύσεις επιχειρησιακών λογαριασμών . . .	132
5.62	Δημοφιλέστερα hashtags ανά λογαριασμούς τύπου Άλλο	133
5.63	Δημοφιλέστερα hashtags ανά δημοσιεύσεις λογαριασμών τύπου Άλλο	133
5.64	Δημοφιλέστερα hashtags χρονικής περιόδου Ιανουάριος-Φεβρουάριος	134
5.65	Δημοφιλέστερα hashtags χρονικής περιόδου Μάρτιος-Μάιος	135
5.66	Δημοφιλέστερα hashtags χρονικής περιόδου Ιούνιος-Αύγουστος	136
5.67	Δημοφιλέστερα hashtags χρονικής περιόδου Σεπτέμβριος-Οκτώβριος	137
5.68	Δημοφιλέστερα hashtags χρονικής περιόδου Νοέμβριος-Δεκέμβριος	138
5.69	Κατανομή του αριθμού εγγεγραμμένων χρηστών (κάτω των 500K εγγεγραμμένων)	140
5.70	Κατανομή του αριθμού εγγεγραμμένων χρηστών (άνω των 500K εγγεγραμμένων)	140
5.71	Κατανομή του αριθμού προβολών των βίντεο (κάτω των 1M προβολών)	141
5.72	Κατανομή του αριθμού προβολών των βίντεο (άνω των 1M προβολών)	142
5.73	Κατανομή του αριθμού θετικών αντιδράσεων των βίντεο (κάτω των 100K θετι- κών αντιδράσεων)	143
5.74	Κατανομή του αριθμού θετικών αντιδράσεων των βίντεο (άνω των 100K θετικών αντιδράσεων)	143
5.75	Κατανομή του αριθμού σχολίων των βίντεο (κάτω των 10K σχολίων)	144
5.76	Κατανομή του αριθμού σχολίων των βίντεο (άνω των 10K σχολίων)	144
5.77	Κατανομή του ποσοστού αλληλεπίδρασης	145
5.78	Κατηγορίες χρηστών του YouTube	146
5.79	Θεματικές κατηγορίες βίντεο ανά κατηγορία χρηστών	146

5.80	Κατηγορίες χρηστών ανά αριθμό εγγεγραμμένων χρηστών	147
5.81	Μέσος αριθμός εγγεγραμμένων χρηστών ανά κατηγορία χρηστών	148
5.82	Μέσος αριθμός εγγεγραμμένων χρηστών ανά θεματική κατηγορία περιεχομένου	149
5.83	Θεματικές κατηγορίες περιεχομένου σε μεγάλα κανάλια (άνω των 400K εγγε- γραμμένων χρηστών)	150
5.84	Βαθμός αλληλεπίδρασης ανά έτος δημιουργία των καναλιού	151
5.85	Σχέση αριθμού θετικών αντιδράσεων και αριθμού εγγεγραμμένων χρηστών . .	152
5.86	Σχέση σχολίων και αριθμού εγγεγραμμένων χρηστών	152
5.87	Σχέση αριθμού προβολών και αριθμού εγγεγραμμένων χρηστών	153
5.88	Σχέση μέσου ποσοστού αλληλεπίδρασης και αριθμού εγγεγραμμένων χρηστών	154
5.89	Μέσος αριθμός θετικών και αρνητικών αντιδράσεων ανά κατηγορία χρήστη .	155
5.90	Σχέση μεταξύ θετικών και αρνητικών αντιδράσεων ανά κατηγορία χρήστη . .	155
5.91	Μέσος αριθμός σχολίων ανά κατηγορία χρήστη	156
5.92	Μέσος αριθμός προβολών ανά κατηγορία χρήστη (κάτω των 250.000 προβολών)	157
5.93	Μέσος αριθμός προβολών ανά κατηγορία χρήστη (γενικά)	157
5.94	Μέσο ποσοστό αλληλεπίδρασης ανά κατηγορία χρήστη	158
5.95	Μέσος αριθμός θετικών αντιδράσεων ανά θεματική κατηγορία βίντεο (κάτω των 15.000 θετικών αντιδράσεων)	159
5.96	Μέσος αριθμός θετικών αντιδράσεων ανά θεματική κατηγορία βίντεο (άνω των 15.000 θετικών αντιδράσεων)	160
5.97	Σχέση μεταξύ θετικών και αρνητικών αντιδράσεων ανά θεματική κατηγορία βίντεο	160
5.98	Μέσος αριθμός σχολίων ανά θεματική κατηγορία βίντεο (κάτω των 1.000 σχολίων)	161
5.99	Μέσος αριθμός σχολίων ανά θεματική κατηγορία βίντεο (άνω των 1.000 σχολίων)	162
5.100	Μέσος αριθμός προβολών ανά θεματική κατηγορία βίντεο (κάτω των 250.000 προβολών)	163
5.101	Μέσος αριθμός προβολών ανά θεματική κατηγορία βίντεο (άνω των 250.000 προβολών)	163
5.102	Μέσο ποσοστό αλληλεπίδρασης ανά θεματική κατηγορία βίντεο	164
5.103	Μέσο ποσοστό αλληλεπίδρασης ανά τύπο βίντεο	167
5.104	Σχέση μεταξύ εγγεγραμμένων χρηστών και προβολών ανά κατηγορία χρηστών	168
5.105	Σχέση μεταξύ εγγεγραμμένων χρηστών και προβολών ανά θεματική κατηγορία βίντεο	169
5.106	Σχέση μεταξύ εγγεγραμμένων χρηστών και προβολών ανά τύπο βίντεο	170
5.107	Αριθμός βίντεο ανά κατηγορία χρήστη	171
5.108	Αριθμός βίντεο ανά θεματική κατηγορία βίντεο	172
5.109	Κορυφαίες θεματικές κατηγορίες βίντεο σε κανάλια δημιουργημένα εντός του 2020	172
5.110	Αριθμός βίντεο ανά μήνα	173
5.111	Αριθμός διαγωνισμών (Giveaway) ανά μήνα	174
5.112	Αριθμός διαγωνισμών (Giveaway) ανά κατηγορία χρήστη	175
5.113	Αριθμός διαγωνισμών (Giveaway) ανά θεματική κατηγορία βίντεο	175
5.114	Αριθμός βίντεο τύπου Unboxing και βίντεο αξιολόγησης (Review) ανά μήνα . .	176
5.115	Αριθμός βίντεο ψυχαγωγίας (Entertainment) και κωμωδίας (Comedy) ανά μήνα	177
5.116	Αριθμός βίντεο νέων και πολιτικής (News & Politics) ανά μήνα	177
5.117	Μέσο ποσοστό αλληλεπίδρασης ανά μήνα	178

5.118	Αριθμός βίντεο ανά ημέρα και κατηγορία χρηστών	179
5.119	Αριθμός βίντεο ανά ημέρα και ώρα (Άνδρες, Γυναίκες)	180
5.120	Αριθμός βίντεο ανά ημέρα και ώρα (Επιχειρήσεις, Άλλο)	180
5.121	Σχέση μεταξύ συχνότητας ανάρτησης βίντεο και αριθμού εγγεγραμμένων χρηστών	181
5.122	Μέση διάρκεια βίντεο ανά θεματική κατηγορία βίντεο	182
5.123	Μέση διάρκεια βίντεο ανά τύπο βίντεο	182
5.124	Σχέση μεταξύ προβολών και μέση διάρκειας βίντεο	183
5.125	Σχέση μεταξύ αριθμού λέξεων-κλειδιών καναλιών και εγγεγραμμένων χρηστών	184
5.126	Δημοφιλέστερες λέξεις-κλειδιά καναλιών (Άνδρες)	185
5.127	Δημοφιλέστερες λέξεις-κλειδιά καναλιών (Γυναίκες)	186
5.128	Δημοφιλέστερες λέξεις-κλειδιά καναλιών (Επιχειρήσεις)	186
5.129	Δημοφιλέστερες λέξεις-κλειδιά καναλιών (Άλλο)	187
5.130	Σχέση μεταξύ αριθμού λέξεων-κλειδιών βίντεο και αριθμού εγγεγραμμένων χρηστών	188
5.131	Σχέση μεταξύ αριθμού λέξεων-κλειδιών βίντεο και αριθμού προβολών	188
5.132	Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Άνδρες)	189
5.133	Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Γυναίκες)	190
5.134	Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Επιχειρήσεις)	191
5.135	Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Άλλο)	192
5.136	Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Ιανουάριος-Φεβρουάριος)	193
5.137	Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Μάρτιος-Μάιος)	193
5.138	Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Ιούνιος-Αύγουστος)	194
5.139	Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Σεπτέμβριος-Οκτώβριος)	195
5.140	Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Νοέμβριος-Δεκέμβριος)	196
6.1	Προσδιορισμός φύλου	198
6.2	Προσδιορισμός ηλικίας	198
6.3	Προτιμάτε να αναζητάτε προϊόντα σε ηλεκτρονικά καταστήματα ή σε μέσα κοινωνικής δικτύωσης;	199
6.4	Ποια από τα παρακάτω μέσα κοινωνικής δικτύωσης χρησιμοποιείτε;	200
6.5	Πόσο καιρό χρησιμοποιείτε το Instagram;	201
6.6	Πόσο καιρό χρησιμοποιείτε το Instagram; (18-34)	201
6.7	Πόσο καιρό χρησιμοποιείτε το Instagram; (35+)	202
6.8	Πόσο συχνά χρησιμοποιείτε το Instagram;	203
6.9	Πόσο συχνά χρησιμοποιείτε το Instagram; (18-34)	203
6.10	Πόσο συχνά χρησιμοποιείτε το Instagram; (35+)	204
6.11	Ακολουθείτε Influencers;	205
6.12	Ακολουθείτε Influencers; (18-44)	205
6.13	Ακολουθείτε Influencers; (45+)	206
6.14	Αν ναι, τι πιστεύετε ότι σας προσφέρει το περιεχόμενο των Influencers;	207
6.15	Τι θεματικό περιεχόμενο σας αρέσει να παρακολουθείτε στο Instagram; (Άνδρες)	208
6.16	Τι θεματικό περιεχόμενο σας αρέσει να παρακολουθείτε στο Instagram; (Γυναίκες)	208
6.17	Προτιμάτε να επισκέπτεστε εξωτερικούς συνδέσμους (links) μέσω swipe up στα stories ή μέσω link που βρίσκεται στις πληροφορίες ενός προφίλ (Bio);	209

6.18 Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (Άνδρες)	210
6.19 Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (Γυναίκες)	210
6.20 Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (18-24)	211
6.21 Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (25-44)	211
6.22 Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (45+)	212
6.23 Έστω ότι δύο Instagram προφίλ, ένα με 500 followers και ένα με 10.000 followers παρουσιάζουν το ίδιο προϊόν. Την άποψη ποιου από τα δύο προφίλ θα εμπιστευτείτε;	213
6.24 Σας ενοχλεί η ύπαρξη χορηγούμενων δημοσιεύσεων; (Άνδρες)	214
6.25 Σας ενοχλεί η ύπαρξη χορηγούμενων δημοσιεύσεων; (Γυναίκες)	214
6.26 Επισκέπτεστε τα προφίλ χορηγούμενων δημοσιεύσεων; (Άνδρες)	215
6.27 Επισκέπτεστε τα προφίλ χορηγούμενων δημοσιεύσεων; (Γυναίκες)	215
6.28 Κάνετε Like/Comment σε δημοσιεύσεις λογαριασμών που δεν ακολουθείτε; . .	216
6.29 Έχετε συμμετάσχει ποτέ σε κάποιο giveaway;	217
6.30 Έχετε συμμετάσχει ποτέ σε κάποιο giveaway; (18-34)	217
6.31 Έχετε συμμετάσχει ποτέ σε κάποιο giveaway; (35+)	218
6.32 Αν ναι, πώς ανακαλύψατε αυτό το giveaway;	219
6.33 Αναζητείτε πληροφορίες, για προϊόντα που ενδιαφέρεστε να αγοράσετε, στο Instagram;	220
6.34 Ποια θεωρείτε ότι είναι η ιδανική συχνότητα αναρτήσεων;	221
6.35 Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο ή μέσω των instagram ιστοριών;	222
6.36 Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο (timeline) ή μέσω των instagram ιστοριών (stories); (18-34)	222
6.37 Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο (timeline) ή μέσω των instagram ιστοριών (stories); (35-44)	223
6.38 Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο (timeline) ή μέσω των instagram ιστοριών (stories); (45+)	223
6.39 Έχετε ανακαλύψει ποτέ νέες δημοσιεύσεις αναζητώντας κάποιο hashtag στο Instagram;	224
6.40 Κάνετε repost (επαναδημοσίευση) δημοσιεύσεις που σας αρέσουν;	225
6.41 Τι κοιτάτε πρώτα σε μία δημοσίευση (timeline post);	226
6.42 Ποια μορφή δημοσίευσης (timeline post) προτιμάτε;	227
6.43 Ακολουθείτε τους/τις αγαπημένους/νες σας YouTubers στο Instagram;	228
6.44 Ακολουθείτε τους/τις αγαπημένους/νες σας YouTubers στο Instagram; (18-34) .	228
6.45 Ακολουθείτε τους/τις αγαπημένους/νες σας YouTubers στο Instagram; (35+) . .	229
6.46 Πόσο καιρό χρησιμοποιείτε το YouTube;	230
6.47 Πόσο καιρό χρησιμοποιείτε το YouTube; (18-34)	230
6.48 Πόσο καιρό χρησιμοποιείτε το YouTube; (35+)	231

6.49	Γνωρίζετε ότι το YouTube αποτελεί ένα ανερχόμενο μέσο κοινωνικής δικτύωσης, κάνοντας διαθέσιμη τη δημιουργία προσωπικών post και stories;	232
6.50	Γνωρίζετε ότι το YouTube αποτελεί ένα ανερχόμενο μέσο κοινωνικής δικτύωσης, κάνοντας διαθέσιμη τη δημιουργία προσωπικών post και stories; (18-34)	232
6.51	Γνωρίζετε ότι το YouTube αποτελεί ένα ανερχόμενο μέσο κοινωνικής δικτύωσης, κάνοντας διαθέσιμη τη δημιουργία προσωπικών post και stories; (35+)	233
6.52	Πόσο συχνά παρακολουθείτε βίντεο στο YouTube;	234
6.53	Πόσο συχνά παρακολουθείτε βίντεο στο YouTube; (18-34)	234
6.54	Πόσο συχνά παρακολουθείτε βίντεο στο YouTube; (35+)	235
6.55	Τι περιεχόμενο προτιμάτε να παρακολουθείτε στο YouTube; (Άνδρες)	236
6.56	Τι περιεχόμενο προτιμάτε να παρακολουθείτε στο YouTube; (Γυναίκες)	236
6.57	Με ποιον τρόπο ανακαλύπτετε νέα βίντεο; (Άνδρες)	237
6.58	Με ποιον τρόπο ανακαλύπτετε νέα βίντεο; (Γυναίκες)	238
6.59	Με ποιον τρόπο ανακαλύπτετε νέα βίντεο; (18-34)	239
6.60	Με ποιον τρόπο ανακαλύπτετε νέα βίντεο; (35+)	239
6.61	Τι κοιτάτε πρώτα σε ένα βίντεο;	240
6.62	Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (Άνδρες)	241
6.63	Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (Γυναίκες)	241
6.64	Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (18-24)	242
6.65	Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (25-44)	242
6.66	Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (45+)	243
6.67	Κάνετε Like ή Comment στα βίντεο που σας αρέσουν;	244
6.68	Κοινοποιείτε βίντεο που σας αρέσουν σε κάποιο άλλο μέσο κοινωνικής δικτύωσης (πχ Facebook, Instagram);	245
6.69	Κοινοποιείτε βίντεο που σας αρέσουν σε κάποιο άλλο μέσο κοινωνικής δικτύωσης (πχ Facebook, Instagram); (18-34)	245
6.70	Κοινοποιείτε βίντεο που σας αρέσουν σε κάποιο άλλο μέσο κοινωνικής δικτύωσης (πχ Facebook, Instagram); (35+)	246
6.71	Παρακολουθείτε βίντεο-reviews (κριτικές) στο YouTube για προϊόντα που σας ενδιαφέρουν να αγοράσετε; (Άνδρες)	247
6.72	Παρακολουθείτε βίντεο-reviews (κριτικές) στο YouTube για προϊόντα που σας ενδιαφέρουν να αγοράσετε; (Γυναίκες)	247
6.73	Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube; (Άνδρες)	248
6.74	Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube; (Γυναίκες)	248
6.75	Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube; (18-34)	249
6.76	Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube; (35+)	249

6.77 Έστω ότι δύο YouTube κανάλια, ένα με 500 subscribers και ένα με 10.000 subscribers διαφημίζουν το ίδιο προϊόν. Ποιο από τα δύο κανάλια θα εμπιστευόσασταν; . . .	250
6.78 Έχετε συμμετάσχει ποτέ σε κάποιο YouTube διαγωνισμό;	251
6.79 Αν ναι, πώς ανακαλύψατε αυτό το Giveaway;	252
6.80 Έχετε αναζητήσει ή επισκεφτεί ποτέ κάποιο προϊόν ή βίντεο που εμφανίζεται σε διαφήμιση στο YouTube;	253
6.81 Έχετε αποχωρήσει ποτέ από βίντεο λόγω αναδυόμενων διαφημίσεων;	254
6.82 Κάθε πότε προτιμάτε να αναρτά βίντεο κάποιος δημιουργός;	255
6.83 Έχετε επισκεφθεί ποτέ κάποιο προφίλ του Instagram που ανακαλύψατε μέσω του YouTube;	256
6.84 Έχετε επισκεφθεί ποτέ κάποιο προφίλ του Instagram που ανακαλύψατε μέσω του YouTube; (18-34)	257
6.85 Έχετε επισκεφθεί ποτέ κάποιο προφίλ του Instagram που ανακαλύψατε μέσω του YouTube; (35+)	257

Κεφάλαιο 1

Εισαγωγή

Στην παρούσα πτυχιακή εργασία παρουσιάζεται ο μηχανισμός που αναπτύχθηκε για την εξόρυξη δημόσιων δεδομένων από προφίλ χρηστών στα μέσα κοινωνικής δικτύωσης, Instagram και YouTube. Επιπλέον, περιγράφεται η συνολική διαδικασία εξαγωγής συμπερασμάτων σχετικά με τη συμπεριφορά των Ελλήνων χρηστών στα κοινωνικά δίκτυα και, τέλος, εξετάζεται η ουσιαστικότητα τους. Σε αυτό το κεφάλαιο, αναλύεται η φύση του αρχικού προβλήματος και περιγράφεται ο τρόπος, με τον οποίο αυτή η εργασία συμβάλλει στην επίλυση του.

1.1 Ανάλυση προβλήματος

Σε μια εποχή που ο ψηφιακός κόσμος αναπτύσσεται με ραγδαίους ρυθμούς, το ενδιαφέρον όλων στρέφεται προς το διαδίκτυο. Η αμεσότητα και η ελευθερία που προσφέρει στους χρήστες έχουν οδηγήσει στη ψηφιοποίηση πολλών πτυχών της καθημερινότητας, όπως η ανάπτυξη διαπροσωπικών σχέσεων μέσω των κοινωνικών δικτύων ή ακόμα και η διαδικτυακή παρουσία όλο και περισσότερων εταιρειών. Τη στροφή αυτή προς το διαδίκτυο έχει ακολουθήσει και η διαφημιστική δραστηριότητα. Το γεγονός ότι η μεγαλύτερη μερίδα του πληθυσμού περιηγείται καθημερινά στο διαδίκτυο και αλληλεπιδρά με πληθώρα αναρτήσεων, έχει οδηγήσει πολλές επιχειρήσεις στο να εγκαταλείψουν τους παραδοσιακούς τρόπους διαφήμισης, οι οποίοι περιόριζαν σημαντικά το κοινό που έρχεται σε επαφή με αυτές, και να διεξάγουν τις διαφημιστικές τους εκστρατείες διαδικτυακά. Η έντονη δραστηριοποίηση στα μέσα κοινωνικής δικτύωσης αποτελεί το σημαντικότερο μέρος της προσπάθειας αυτής.

Με την πάροδο του χρόνου, όλο και περισσότερα είναι τα άτομα που χρησιμοποιούν τα μέσα κοινωνικής δικτύωσης τόσο σε προσωπικό, όσο και σε επαγγελματικό επίπεδο. Οι επιχειρήσεις αντιλαμβανόμενες την τάση της μάζας προς τα κοινωνικά δίκτυα, ξεκίνησαν να χιτίζουν την διαδικτυακή τους παρουσία σε αυτά, προβάλλοντας και προωθώντας τα προϊόντα τους στο ήδη εδραιωμένο κοινό τους, αλλά και σε μελλοντικά υποψήφιους πελάτες. Ωστόσο, για να επιτευχθεί κάτι τέτοιο, οι επιχειρήσεις έρχονται καθημερινά αντιμέτωπες με ερωτήματα που αφορούν την συμπεριφορά των καταναλωτών στα κοινωνικά δίκτυα, όπως με ποιους τρόπους μπορούν να προσεγγίσουν αποτελεσματικότερα το στοχευμένο κοινό τους, τι είναι αυτό που διατηρεί ζωντανό το ενδιαφέρον του κοινού ή με τι είδους περιεχόμενο προτιμάει το κοινό να αλληλεπιδρά. Οι επιχειρήσεις, μέσα από την ανάλυση της συμπεριφοράς του κοινού, αποσκοπούν στο να αυξήσουν την απήχηση των προφίλ τους στα κοινωνικά δίκτυα, κάτι το οποίο

συνεπάγεται και με την αύξηση των κερδών τους.

Η είσοδος των επιχειρήσεων και η ανάπτυξη εμπορικών σχέσεων στα μέσα κοινωνικής δικτύωσης προσέδωσαν εμπορικό ενδιαφέρον και σε χρήστες με μεγάλο κοινό, καθώς φαίνεται ότι μπορούν να ασκήσουν επιρροή στο κοινό που τους ακολουθεί. Χρήστες ευρέως γνωστοί, όπως καλλιτέχνες ή αθλητές, αποτελούν στόχο για εταιρείες που αναζητούν να παρουσιάσουν τα προϊόντα τους σε μία ήδη εδραιωμένη αγοραστική κοινότητα. Το ενδιαφέρον αυτό οφείλεται στο γεγονός, ότι οι χρήστες αυτοί, γνωστοί ως *Influencers* έχουν καταφέρει, μέσα σε μικρό χρονικό διάστημα, κάτι που οι επιχειρήσεις προσπαθούσαν και συνεχίζουν να προσπαθούν να πετύχουν μέχρι και σήμερα, δηλαδή την ανάπτυξη μίας ειλικρινούς επικοινωνίας σε συνδυασμό με την παροχή περιεχομένου σε σταθερή συχνότητα. Έτσι, δημιουργήθηκε ο κλάδος του *Influencer Marketing* [16]. Εταιρείες που αναζητούν να προωθήσουν τα προϊόντα τους με μεγάλο ποσοστό επιτυχίας και στον λιγότερο δυνατό χρόνο, στρέφονται σε *Influencers* με χαρακτήρα με προφίλ αντίστοιχο με αυτόν της επιχείρησης. Επομένως, άλλη μια πρόκληση για τις επιχειρήσεις είναι να εντοπίσουν προφίλ που διατηρούν υψηλό ποσοστό αλληλεπίδρασης με το κοινό τους και είναι συμβατά με την διαδικτυακή τους εικόνα.

Συμπερασματικά, η είσοδος του εμπορίου στα μέσα κοινωνικής δικτύωσης έχει δημιουργήσει αρκετά σημεία που χρήζουν μελέτης. Στην εργασία αυτή, το πρόβλημα έχει προσεγγιστεί ως προς τρεις βασικούς άξονες, (i) τη συμπεριφορά των απλών χρηστών, (ii) τη συμπεριφορά των *Influencers*, και (iii) το ενδιαφέρον των επιχειρήσεων ως προς αυτές τις συμπεριφορές. Η κατηγοριοποίηση αυτή, ανταποκρίνεται στην πραγματική εικόνα των κοινωνικών δικτύων όπου, το μεγαλύτερο μέρος του κοινού είναι απλοί χρήστες που επιλέγουν τα μέσα με μοναδικό σκοπό την ψυχαγωγία τους, ενώ σεβαστό είναι και το πλήθος των *Influencers* και των επιχειρήσεων που επιλέγουν τα μέσα με σκοπό την αποκόμιση εσόδων, μέσω της προώθησης υπηρεσιών και προϊόντων. Εντούτοις, ο κλάδος των κοινωνικών δικτύων είναι τεράστιος, γεγονός που δημιουργεί την ανάγκη για εστίαση σε συγκεκριμένες πλατφόρμες. Καθοριστικό ρόλο για την επιλογή σε αυτή την εργασία αποτέλεσε η απήχηση της εκάστοτε πλατφόρμας, σε συνδυασμό με το εμπορικό ενδιαφέρον που παρουσιάζει η καθεμία από αυτές κατά το τελευταίο χρονικό διάστημα, και συγκεκριμένα κατά την περίοδο 2020-2021. Συμπληρωματικά, ο τεράστιος όγκος των δεδομένων που αφορούν στα κοινωνικά δίκτυα σε παγκόσμιο επίπεδο, απαιτούσε την επιλογή ενός υποσυνόλου του συνολικού πλήθους των χρηστών, με αποτέλεσμα η έρευνα να έχει ως θεματικό άξονα Έλληνες χρήστες με περισσότερους από 1.000 ακόλουθους. Στην απόφαση αυτή συνέβαλε το γεγονός ότι τα διαθέσιμα στατιστικά στοιχεία, όσον αφορά το ελληνικό κοινό, είναι απαρχαιωμένα και μη εμπειριστατωμένα, καθώς και ότι οι χρήστες με τουλάχιστον χίλιους ακόλουθους εμφανίζουν αρκετά υψηλή δραστηριότητα. Λαμβάνοντας υπόψη όλα τα παραπάνω, η μελέτη επικεντρώθηκε στα Ελληνικά δεδομένα των κοινωνικών δικτύων Instagram και YouTube, με το πρώτο να αποτελεί το δημοφιλέστερο μέσο κοινωνικής δικτύωσης τον τελευταίο χρόνο, και το δεύτερο να κυριαρχεί στην κατηγορία των βίντεο, σε μία εποχή που η ανάγκη του κοινού για οπτικοακουστικό υλικό βρίσκεται στην κορύφωση της.

1.2 Η συνεισφορά μας στο πρόβλημα

Στην παρούσα έρευνα, τέθηκε ως κύριος στόχος η ανάλυση των παραπάνω προβληματισμών και η εξαγωγή γνώσης που θα μπορούσε να βοηθήσει στην αντιμετώπιση τους. Η διαδικασία αυτή εκτυλίχθηκε σε τρία στάδια: (i) τη συγκέντρωση των δεδομένων, (ii) την εξαγωγή συμπε-

ρασμάτων με βάση τα δεδομένα αυτά και (iii) τη διεξαγωγή ερωτηματολογίου με σκοπό την ενίσχυση και συμπλήρωση των συμπερασμάτων.

Στο στάδιο της συγκέντρωσης των δεδομένων, αναπτύχθηκαν δύο Web Spiders με χρήση του Scrapy [8], ένα για το Instagram και ένα για το YouTube αντίστοιχα, τα οποία λάμβαναν ως είσοδο μία λίστα με ονόματα χρηστών, επισκέπτονταν τα προφίλ αυτά και συγκέντρωναν όλες τις διαθέσιμες πληροφορίες. Στην συνέχεια, οι πληροφορίες για τα προφίλ αποθηκεύονταν ως έγγραφα (documents) στην MongoDB [25] NoSQL βάση δεδομένων. Τα έγγραφα που δημιουργήθηκαν, μοιράστηκαν σε δύο συλλογές δεδομένων (collections), μία για κάθε μέσο κοινωνικής δικτύωσης. Συνολικά, συγκεντρώθηκαν δεδομένα για 2.727 προφίλ του Instagram και για 2.057 προφίλ του YouTube, τα οποία αποτέλεσαν την βάση για την μελέτη της εργασίας και τη δημιουργία στατιστικών στοιχείων.

Ως επόμενο βήμα, ακολούθησε η οπτική αναπαράσταση των συλλεχθέντων δεδομένων με την χρήση των δυνατοτήτων που προσφέρει η πλατφόρμα του MongoDB Atlas [24]. Σε αυτό το στάδιο, δημιουργήθηκαν γραφήματα με σκοπό την απάντηση ερωτημάτων, όπως εκείνα που τέθηκαν στην Ενότητα 1.1 και, γενικότερα, την ανάλυση της δραστηριότητας των Ελλήνων χρηστών στα κοινωνικά δίκτυα. Με την χρήση των γραφημάτων ανιχνεύθηκαν και μελετήθηκαν τα πιθανά μοτίβα συμπεριφοράς τόσο των απλών χρηστών, όσο και των επαγγελματικών προφίλ, δηλαδή των *Influencers* και των επιχειρήσεων. Τα αποτελέσματα που προέκυψαν από την προκειμένη ανάλυση βοήθησαν στην χαρτογράφηση της Ελληνικής ψηφιακής κοινότητας και διευκρίνισαν τις σχέσεις αλληλεπίδρασης μεταξύ των διαφόρων κατηγοριών των χρηστών.

Παρά το γεγονός ότι τα δημόσια δεδομένα των προφίλ που συγκεντρώθηκαν ήταν αρκετά για να οδηγηθεί η μελέτη σε επαρκή συμπεράσματα, υπήρχαν πεδία που δεν ήταν δυνατό να παρατηρηθούν, λόγω του σεβασμού της ιδιωτικότητας αρκετών προσωπικών στοιχείων. Ένα μεγάλο ποσοστό των προσωπικών δεδομένων, όπως για παράδειγμα, το φύλλο, η ηλικία και τα στατιστικά απόδοσης της κάθε δημοσίευσης, των χρηστών διαφυλάσσεται από την εκάστοτε διαδικτυακή κοινότητα και χρήζουν άδειας, ώστε να διατεθούν για δημόσιες έρευνες. Καθώς το τελευταίο δεν ήταν δυνατό στα πλαίσια αυτής της πτυχιακής εργασίας, αποφασίστηκε να διαμοιραστεί ένα ερωτηματολόγιο, το οποίο καλούνταν να καλύψει την έλλειψη δημόσιων δημογραφικών στοιχείων και δεδομένων που αφορούν τις συνήθειες του κοινού στα μέσα κοινωνικής δικτύωσης. Το ερωτηματολόγιο αυτό απαντήθηκε από 270 άτομα που κατατάσσονται σε ολόκληρο το ηλικιακό φάσμα και, παράλληλα, εμπλούτισε τα ήδη καταγεγραμμένα αποτελέσματα.

1.3 Οργάνωση της πτυχιακής εργασίας

Τα ακόλουθα κεφάλαια της πτυχιακής εργασίας οργανώνονται ως εξής. Στο Κεφάλαιο 2 παρουσιάζονται οι ήδη υπάρχουσες εργασίες που είναι σχετικές με το αντικείμενο της μελέτης μας. Στο Κεφάλαιο 3 περιγράφονται με λεπτομέρεια όλα τα εργαλεία που χρησιμοποιήθηκαν για την ολοκλήρωση της εργασίας (Scrapy, MongoDB, MongoCompass, κ.κ.). Στο Κεφάλαιο 4 αναλύεται η αρχιτεκτονική των Web Spiders που δημιουργήθηκαν για την συγκέντρωση δεδομένων από τα Instagram και YouTube. Στη συνέχεια, στο Κεφάλαιο 5 αναπτύσσεται η διαδικασία που ακολουθήθηκε με σκοπό την εξαγωγή συμπερασμάτων από τα συγκεντρωμένα δεδομένα και παρουσιάζονται οι γραφικές αναπαραστάσεις που οδηγούν τα συμπεράσματα αυτά. Στο Κεφάλαιο 6 παρουσιάζεται το σκεπτικό πίσω από τον σχεδιασμό του ερωτηματολογίου

που διεξήχθη και γίνεται αναφορά στα ποιοτικά και δημογραφικά χαρακτηριστικά των απαντήσεων που συλλέχθηκαν. Τέλος, στο Κεφάλαιο 7 γίνεται η διασταύρωση των στοιχείων από τα Κεφάλαια 5 και 6 και προκύπτουν τα τελικά συμπεράσματα της παρούσας μελέτης, καθώς επίσης ορίζονται και οι προοπτικές για την επέκταση και την βελτίωση της λειτουργικότητας της παρούσας εργασίας.

Κεφάλαιο 2

Σχετικό Υλικό

Η συλλογή δεδομένων από το διαδίκτυο και η μελέτη τους αποτελεί μία πολύπλοκη διαδικασία, η οποία για να έρθει εις πέρας, απαιτεί την συνύπαρξη διαφορετικών εργαλείων και τεχνικών. Τα βήματα που πρέπει να ακολουθηθούν για την επίτευξη των παραπάνω είναι: (i) η συγκομιδή διαδικτυακών δεδομένων με την χρήση web crawlers/scrapers, (ii) η αποθήκευση των δεδομένων σε μία, κατάλληλη για το πρόβλημα, βάση δεδομένων και (iii) η εξαγωγή γνώσης από τα συλλεχθέντα δεδομένα.

Στις επόμενες ενότητες περιγράφονται αναλυτικότερα όλα τα παραπάνω στάδια και γίνεται αναφορά στα διαθέσιμα εργαλεία, όπως και σε πιθανούς προβληματισμούς που μπορεί να προκύψουν κατά τη διαδικασία της συλλογής και της επεξεργασίας δεδομένων.

2.1 Web Crawling και Web Scraping

Τα Web Scrapers και Web Crawlers, ή αλλιώς Web Spiders [19], [33], αποτελούν ένα πρόγραμμα/ μηχανισμό που περιηγείται στον Παγκόσμιο Ιστό (World Wide Web) με αυτοματοποιημένο τρόπο, ο οποίος καθορίζεται από τον δημιουργό του, με σκοπό τη στοχευμένη αναζήτηση και εξαγωγή πληροφορίας από μία πληθώρα διαδικτυακών πηγών. Το διαδίκτυο, το οποίο αποτελεί το εικονικό περιβάλλον αναζήτησης των δεδομένων, χαρακτηρίζεται ως ένα κατευθυνόμενο διάγραμμα, όπου οι ιστοσελίδες παρουσιάζονται σαν κόμβοι και οι υπερσύνδεσμοι (hyper-links) αναπαρίστανται ως ακμές, που συνδέουν τις διάφορες ιστοσελίδες. Έτσι, ξεκινώντας από μία μόνο ιστοσελίδα, ο μηχανισμός αυτός είναι δυνατό να επισκεφθεί αρκετές νέες ιστοσελίδες, σχετικές με τον θεματικό άξονα της αρχικής αναζήτησης [53]. Με την πάροδο των χρόνων, ο Παγκόσμιος Ιστός επεκτάθηκε ραγδαία σε μέγεθος, γεγονός το οποίο προσέθεσε ιδιαίτερη πολυπλοκότητα στην οργάνωση του περιεχομένου της εκάστοτε ιστοσελίδας. Από το 1994, που δημιουργήθηκε το πρώτο web crawler [10], έως και σήμερα, η βάση για την δημιουργία ενός τέτοιου μηχανισμού παραμένει αναλλοίωτη. Ωστόσο, η εκθετική αύξηση των ιστοσελίδων, σε συνδυασμό με την όλο και πολυπλοκότερη δομή τους, απαιτούν εκσυγχρονισμένες και σύνθετες τεχνικές για την εξαγωγή πληροφοριών από αυτές.

2.1.1 Web Crawling έναντι Web Scraping

Είναι σημαντικό να γίνει η διάκριση μεταξύ των όρων Web Crawling και Web Scraping. Το γεγονός ότι υπάρχουν μεγάλες ομοιότητες μεταξύ των δύο διαδικασιών αποτελεί τον λόγο για τον οποίο οι δύο όροι συχνά χρησιμοποιούνται απαράλλαχτα. Παρ' όλα αυτά, υπάρχουν σημαντικές διαφορές μεταξύ τους.

Ο όρος Web Crawling απευθύνεται στην διαδικασία χρήσης αυτοματοποιημένου λογισμικού (γνωστό ως *bot* ή *spider*) για την ανάγνωση και την αποθήκευση ολόκληρου του περιεχομένου μίας ιστοσελίδας, με σκοπό την αρχειοθέτηση και την ευρετηρίαση των πληροφοριών της. Αρκετοί περιηγητές (web browsers), όπως το Bing και το Google, χρησιμοποιούν Web Crawlers για να εξάγουν όλες τις διαθέσιμες πληροφορίες από την κάθε ιστοσελίδα και, στη συνέχεια, να δημιουργήσουν ένα ευρετήριο για αυτές στις μηχανές αναζήτησης τους [33]. Η λειτουργία ενός web crawler συνηθίζεται να παρομοιάζεται με την διαδικασίας ύφανσης του ιστού μίας αράχνης, εξού και η εναλλακτική του ονομασία Web Spider. Όπως προαναφέρθηκε, το Web Crawling εφαρμόζεται σε όλες τις σελίδες ενός ιστότοπου και όχι σε ένα υποσύνολο των περιεχομένων του, γεγονός που ενισχύει την καταλληλότητα του για εφαρμογές μεγαλύτερης κλίμακας. Το μέγεθος των εφαρμογών αυτών απαιτεί, επιπλέον, την αποτροπή συλλογής διπλότυπων πληροφοριών (deduplication), με σκοπό την αποφυγή επιβάρυνσης του συστήματος και, παράλληλα, την εξοικονόμηση αποθηκευτικού χώρου [33], [52].

Ο όρος Web Scraping απευθύνεται, επίσης, στην διαδικασία χρήσης αυτοματοποιημένου λογισμικού για την ανάγνωση και την αποθήκευση περιεχομένου μίας σελίδας. Ωστόσο, στην προκειμένη περίπτωση, συγκεντρώνεται ένας στοχευμένο υποσύνολο δεδομένων και όχι όλο το περιεχόμενο της ιστοσελίδας. Σε αντίθεση με το Web Crawling, το Web Scraping, δεν περιορίζεται στα όρια του διαδικτύου, αλλά μπορεί να εφαρμοστεί σε οποιαδήποτε δομή δεδομένων όπως καθορίζεται από τον χρήστη του μηχανισμού. Η χρήση του Web Scraping, συνηθίζεται σε εφαρμογές μικρότερης κλίμακας, με αποτέλεσμα ο έλεγχος για την ύπαρξη διπλότυπων να είναι λιγότερο απαιτητικός, όμως αυτό δεν αποτρέπει την χρήση του σε μεγαλύτερου μεγέθους δεδομένα. Εν ολίγοις, παρατηρείται πως ένα Web Scraper λειτουργεί αρκετά πιο επιλεκτικά, σε σχέση με ένα Web Crawler [33], [52]. Στον παρακάτω πίνακα αντιπαραθέτονται εν συντομία οι προαναφερθείσες διαφορές.

Web Crawling	Web Scraping
<ul style="list-style-type: none"> • Συγκεντρώνει όλα τα περιεχόμενα μίας ιστοσελίδας • Αρχειοθετεί και ευρετηριάζει όλα τα δεδομένα • Κατάλληλο για μεγάλης κλίμακας εφαρμογές • Απαιτείται αφαίρεση διπλότυπων πληροφοριών 	<ul style="list-style-type: none"> • Συγκεντρώνει ένα στοχευμένο υποσύνολο των περιεχομένων μίας ιστοσελίδας • Συλλέγει τα επιθυμητά δεδομένα • Κατάλληλο για οποιασδήποτε κλίμακας εφαρμογές • Δεν απαιτείται αφαίρεση διπλότυπων πληροφοριών

Πίνακας 2.1: Web Crawling έναντι Web Scraping

2.1.2 Η λειτουργία των Web Crawlers και Web Scrapers

Εφόσον, στην Ενότητα 2.1.1, πραγματοποιήθηκε η διάκριση μεταξύ των βασικών χαρακτηριστικών των Web Crawlers και Web Scrapers, στο σημείο αυτό θα παρουσιαστούν αναλυτικότερα τα επιμέρους βήματα που ακολουθεί ο κάθε μηχανισμός για τη συγκομιδή δεδομένων.

Η διαδικασία του Web Crawling μπορεί να αναλυθεί στα εξής βήματα [33], [19]:

1. *Επιλογή του αρχικού ή των αρχικών URLs*
2. *Προσθήκη του αρχικού ή των αρχικών URLs στο frontier. Το frontier αποτελεί το μέρος του συστήματος συγκομιδής δεδομένων (crawling system), το οποίο αποφασίζει τη λογική και την πολιτική που ακολουθούνται κατά την επίσκεψη ιστοσελίδων [38]*
3. *Εξαγωγή του URL ή των URLs από το frontier*
4. *Συγκέντρωση των δεδομένων της ιστοσελίδας που ανταποκρίνεται σε αυτό το URL*
5. *Ανάλυση των περιεχομένων της ιστοσελίδας με σκοπό την εύρεση νέων URLs*
6. *Προσθήκη όλων των νέων URL στο frontier*
7. *Επανάληψη της διαδικασίας από το βήμα 3, ωστόσο το frontier μείνει άδειο*

Αντίστοιχα, η διαδικασία του Web Scraping αναλύεται σε τρία βασικά βήματα [33]:

1. *Αίτημα-Απάντηση (Request-Response)*
 - Ως πρώτο βήμα ορίζεται το αίτημα για τα περιεχόμενα ενός συγκεκριμένου URL μιας ιστοσελίδας.
 - Ως απάντηση στο αίτημα αυτό, το Web Scraper λαμβάνει τις ζητούμενες πληροφορίες σε μορφή HTML κώδικα.
2. *Ανάλυση των δεδομένων και εξαγωγή πληροφοριών*
 - Σε αυτό το σημείο, αναλύεται ο κώδικας HTML της σελίδας και συλλέγονται οι επιθυμητές πληροφορίες, όπως τίτλοι παραγράφων, ολόκληρες παράγραφοι ή υπερσύνδεσμοι (links) για άλλες σελίδες.
3. *Αποθήκευση των δεδομένων*
 - Η αποθήκευση των δεδομένων αποτελεί το τελικό στάδιο της διαδικασίας του Web Scraping. Κατά το βήμα αυτό, οι πληροφορίες που εξήχθησαν από τον HTML κώδικα της ιστοσελίδας αποθηκεύονται σε αρχεία JSON, CSV ή ακόμα και σε βάσεις δεδομένων, με σκοπό τη μελλοντική τους χρήση.

Στα πλαίσια αυτής της εργασίας, επιλέχθηκε, ως καταλληλότερη για το τη λύση του προβλήματος, η διαδικασία του Web Scraping. Η επιλογή αυτή καθορίστηκε από το γεγονός, ότι τα προς συλλογή δεδομένα ήταν συγκεκριμένα και δεν υπήρχε η ανάγκη για την αποθήκευση ολόκληρου του περιεχομένου της κάθε ιστοσελίδας. Επιπλέον, το σύνολο των ιστοσελίδων που

έπρεπε να προσπελαστούν ήταν προκαθορισμένο, με αποτέλεσμα να μην πραγματοποιηθεί η δυναμική επίσκεψη νεοανακαληφθέντων URLs. Ως εκ τούτου, η επόμενη ενότητα εστιάζουν σε πληροφορίες που αφορούν αποκλειστικά το Web Scraping.

2.1.3 Τεχνικές Web Scraping

Η συγκομιδή δεδομένων από το διαδίκτυο αποτελεί μια δραστηριότητα για την οποία έχουν αναπτυχθεί διαφορετικές τεχνικές, που ποικίλουν ανάλογα με τη φύση του προβλήματος. Το μέγεθος των δεδομένων που προορίζονται για συλλογή, η οργάνωση αυτών των δεδομένων, οι δικλίδες ασφαλείας της κάθε ιστοσελίδας, καθώς και οι δυνατότητες της εκάστοτε γλώσσας προγραμματισμού σχηματίζουν την κατευθυντήρια γραμμή για τον σχεδιασμό της μεθοδολογίας που θα εφαρμοστεί.

Οι τεχνικές για την συλλογή των δεδομένων μπορούν να χωριστούν σε δύο κατηγορίες με κεντρικό κριτήριο να αποτελεί η ύπαρξη ή όχι αυτοματοποιημένου λογισμικού. Η χειροκίνητη συγκομιδή δεδομένων, γνωστή και ως Manual Scraping, αφορά στην περίπτωση κατά την οποία ο χρήστης επιλέγει με τον κέρσορα του τα επιθυμητά πεδία μιας σελίδας και με Αντιγραφή/Επικόλληση (Copy/Paste) αποθηκεύει τα δεδομένα σε μία δομή δεδομένων, όπως μία φόρμα Excel ή μία βάση δεδομένων. Στον αντίποδα, το αυτοματοποιημένο λογισμικό (Automated Scraping) παρέχει την δυνατότητα στον χρήστη να εκτελεί παράλληλες διαδικασίες αναζήτησης και εξαγωγής δεδομένων από πολλαπλές διαδικτυακές πηγές, χωρίς να είναι αναγκαία η διαρκής ενσχόληση του με τις τρέχουσες διεργασίες.

Σε αρκετές περιπτώσεις, το Manual Scraping συνιστά την πρακτικότερη δίοδο. Συχνά, τα δεδομένα που προορίζονται για συλλογή έχουν μικρό μέγεθος και δεν απαιτούν σύνθετη και επαναλαμβανόμενη επεξεργασία. Η απλότητα που διέπει τέτοιου είδους προβλήματα, σε αντιπαράθεση με τον χρόνο που πρέπει να αφιερωθεί για την εγκατάσταση ενός αυτοματοποιημένου μηχανισμού, οδηγεί στην υιοθέτηση χειροκίνητων πρακτικών. Επιπλέον, τα μέτρα διασφάλισης των δεδομένων που εφαρμόζονται από τις ιστοσελίδες είναι, πολλές φορές, αποτρεπτικά για την δυναμική προσπέλαση τους [9]. Παρ' όλα αυτά, στις περιπτώσεις που ο όγκος των δεδομένων είναι αυξημένος, η εφαρμογή αυτοματοποιημένων τεχνικών μπορεί να εξοικονομήσει αρκετό χρόνο και κόπο από τη πλευρά του χρήστη. Τα μοντέρνα Web Scraping συστήματα παρέχουν φιλικές διεπαφές χρήστη (User Interface) που επιτρέπουν σε οποιονδήποτε να διεξάγει διαδικασίες εξαγωγής δεδομένων χωρίς οι εξεζητημένες γνώσεις προγραμματισμού να είναι απαραίτητη προϋπόθεση. Στους πίνακες 2.2 και 2.3 παρουσιάζονται συνοπτικά τα πλεονεκτήματα και τα μειονεκτήματα της κάθε μιας από τις παραπάνω προσεγγίσεις.

Όπως αναφέρθηκε στο Κεφάλαιο 1, και συγκεκριμένα στην Ενότητα 1.1, το διαδίκτυο αναπτύσσεται με ραγδαίους ρυθμούς. Η ανάγκη των χρηστών του διαδικτύου για υψηλής αισθητικής και πλούσιο περιεχόμενο έχει αυξήσει τον φόρτο των πληροφοριών που στεγάζουν οι ιστοσελίδες, με αποτέλεσμα η οργάνωση τους να καθίσταται όλο και πολυπλοκότερη. Το γεγονός αυτό αποτελεί τροχοπέδη για τις σύγχρονες μεθόδους αυτοματοποιημένης συγκομιδής δεδομένων, οι οποίες καλούνται να αντεπεξέλθουν στη συγκομιδή μιας πληθώρας, διαφορετικού τύπου, δεδομένων. Ως εκ τούτου, ανάλογα με τα δεδομένα που πρέπει να συλλεχθούν από μια ιστοσελίδα, ακολουθείται και η αντίστοιχη προσέγγιση.

Πλεονεκτήματα	Μειονεκτήματα
<ul style="list-style-type: none"> • Εύκολη μέθοδος, δεν απαιτούνται προγραμματιστικές γνώσεις • Αποφυγή σφαλμάτων λόγω του ποιοτικότερου ελέγχου στα υπό επεξεργασία δεδομένα • Αποφυγή αποκλεισμού από τα συστήματα ασφαλείας των ιστοσελίδων, λόγω αργής συλλογής δεδομένων 	<ul style="list-style-type: none"> • Αργή μέθοδος • Πιο επιρρεπές σε ανθρώπινα σφάλματα

Πίνακας 2.2: Manual Web Scraping

Πλεονεκτήματα	Μειονεκτήματα
<ul style="list-style-type: none"> • Πολύ γρήγορη συλλογή δεδομένων • Μεγάλη εξοικονόμηση χρόνου • Τα μοντέρνα Web Scrapers είναι φιλικά προς τον χρήστη 	<ul style="list-style-type: none"> • Απαιτεί γνώσεις προγραμματισμού • Πολλές ιστοσελίδες απορρίπτουν τα αιτήματα από αυτοματοποιημένα συστήματα • Μικρότερα βαθμός ελέγχου κατά την εξαγωγή δεδομένων

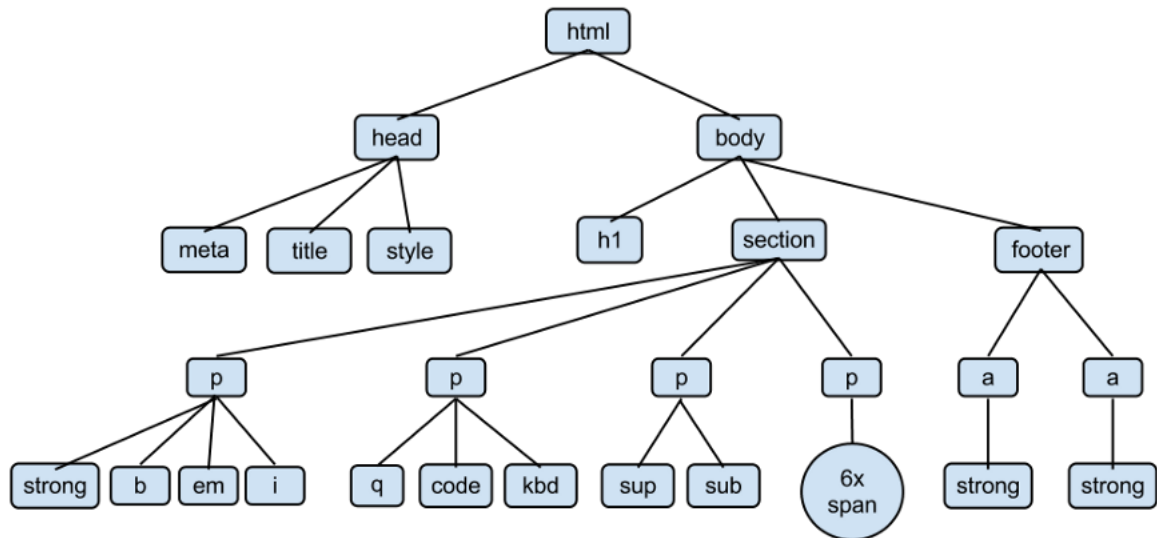
Πίνακας 2.3: Automated Web Scraping

HTML Parsing

Η HTML είναι μια γλώσσα σήμανσης υπερκειμένου (HyperText Markup Language) που χρησιμοποιείται ως βάση για τη δημιουργία ιστοσελίδων. Ο HTML κώδικας συγκροτείται από επιμέρους στοιχεία (HTML Elements), τα οποία διαβιβάζουν στον περιηγητή (Browser) πληροφορίες για τη διάταξη των δεδομένων κατά την παρουσίαση τους. Η ανάλυση της HTML δομής μιας ιστοσελίδας και η εξαγωγή των επιθυμητών δεδομένων από το εσωτερικό HTML στοιχείων, ονομάζεται “HTML Parsing” και συνιστά μία από τις συχνότερες τεχνικές για εξαγωγή πληροφοριών από διαδικτυακές πηγές. Με τις κατάλληλες γλώσσες σεναρίων (scripting languages) ή τα κατάλληλα εργαλεία συγκέντρωσης πληροφοριών, όπως τα CSS Selectors [7], από τον Παγκόσμιο Ιστό, κάθε σελίδα με παρόμοιο προγραμματιστικό μοτίβο σχεδίασης με αυτό που παραθέτεται στο Σχήμα 2.1 [9], μπορεί να χρησιμοποιηθεί ως πηγή δεδομένων.

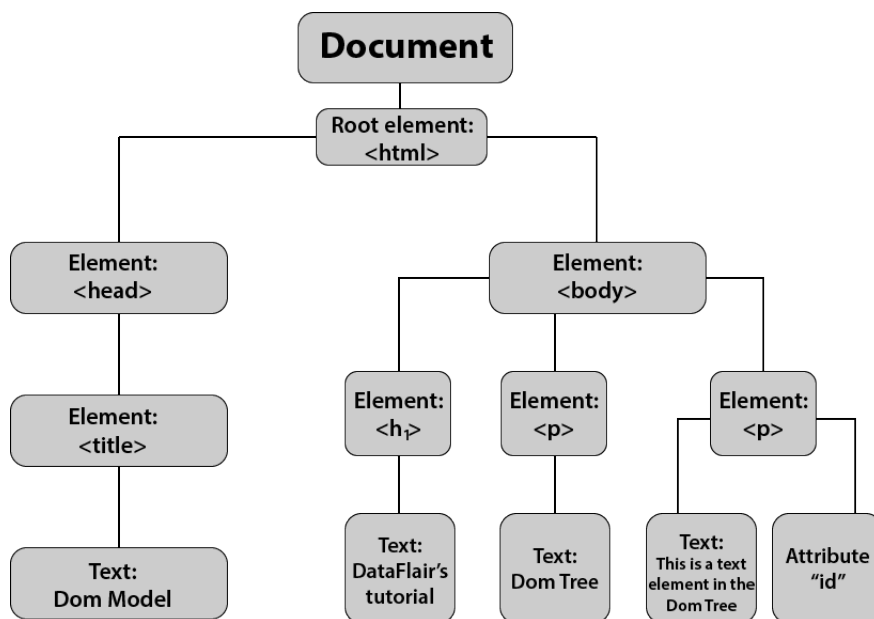
DOM Parsing

Η ανάπτυξη των γλωσσών προγραμματισμού και των περιηγητών, σε συνδυασμό με την ανάγκη για βαθύτερη εξέταση της δομής ενός HTML αρχείου, οδήγησαν στην εξέλιξη της μεθόδου HTML Parsing σε DOM Parsing. Το DOM (Document Object Model) αναπαριστά τα περιεχόμενα ενός HTML ή XML αρχείου σε μορφή δέντρου, όπου κάθε κόμβος αντιπροσωπεύει ένα HTML ή XML στοιχείο (Element) του αρχείου. Ένα δέντρο DOM μπορεί να παρομοιαστεί με ένα γενεαλογικό δέντρο, όπου οι κόμβοι με μικρότερο βάθος αποτελούν τους προγόνους των κόμβων με



Σχήμα 2.1: Βασική δομή ενός HTML αρχείου

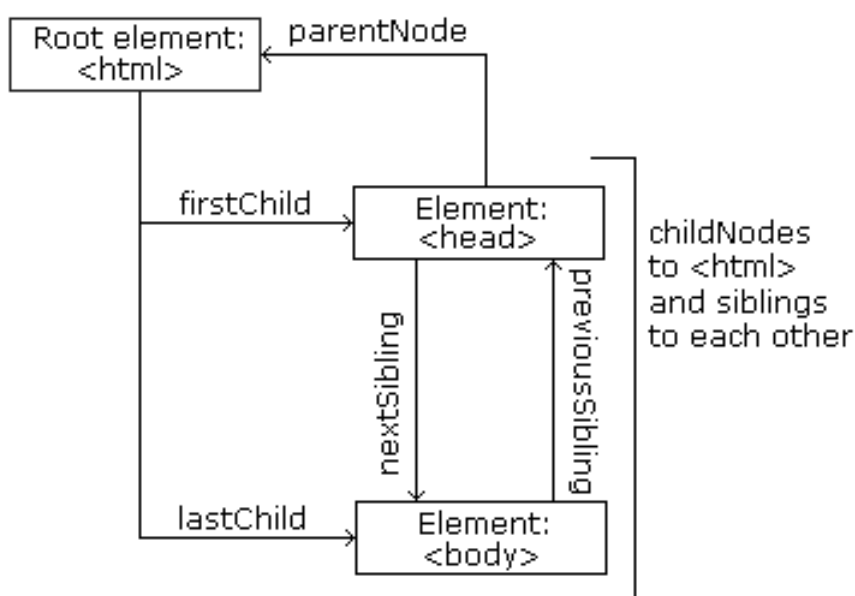
μεγαλύτερο βάθος. Το βάθος του κάθε κόμβου στο δέντρο καθορίζει τον βαθμό συγγενείας του με τα υπόλοιπα στοιχεία του εγγράφου. Η τεχνική DOM Parsing μπορεί να χρησιμοποιηθεί από τους scrapers, με σκοπό την ανάκτηση των δεδομένων που βρίσκονται εμφωλευμένα σε κόμβους και, στη συνέχεια, με χρήση εργαλείων (για παράδειγμα XPath), να εξαχθούν οι απαραίτητες πληροφορίες από αυτά. Στο σχήμα 2.2 παρατηρείται η δενδροειδής δομή του DOM μιας τυπικής ιστοσελίδας [9], [2], [18].



Σχήμα 2.2: Βασική δομή του δέντρου DOM

XPath

Η XPath (XML Path Language) είναι μια γλώσσα ερωτημάτων (query language), η οποία χρησιμοποιείται για την εξαγωγή πληροφοριών από XML αρχεία. Λόγω του γεγονότος ότι τα XML, όπως και τα HTML, αρχεία έχουν δένδροειδή δομή, η XPath υποστηρίζει την εκτέλεση ερωτημάτων και σε HTML έγγραφα. Η συγκεκριμένη γλώσσα αντιλαμβάνεται τα δεδομένα με πιο σύνθετο τρόπο, σε σχέση με το DOM, καθώς οι σχέσεις που αναπτύσσονται μεταξύ των κόμβων του δέντρου δεν είναι μονομερείς, αλλά σχέσεις γονέα-παιδιού, όπως είναι εμφανές και από το Σχήμα 2.3. Τα XPath ερωτήματα, γνωστά και ως εκφράσεις διαδρομής, συντάσσονται με τη χρήση καθέτων (/), έτσι ώστε να διαχωρίζονται τα διαδοχικά βήματα προς τον κόμβο που περιλαμβάνει τα επιθυμητά δεδομένα, όπως τα URLs και οι διευθύνσεις διαδρομής στα συστήματα αρχείων [9], [2].



Σχήμα 2.3: Βασική δομή του XPath μοντέλου

Αντιστοίχιση Μοτίβου-Κειμένου

Στο σύνολο των τεχνικών συγκέντρωσης δεδομένων ανήκει και η *Αντιστοίχιση Μοτίβου-Κειμένου* (Text Pattern Matching). Η λειτουργία της βασίζεται, αρχικά, στον καθορισμό της ακολουθίας χαρακτήρων, δηλαδή στο μοτίβο της πληροφορίας που πρέπει να συλλεχθεί και, στη συνέχεια, στην ανάλυση του κειμένου και στην εξαγωγή των τμημάτων που είναι ταυτοτικά ίδια με το μοτίβο σύγκρισης. Ένα από τα βασικά εργαλεία *Αντιστοίχισης Μοτίβου-Κειμένου* είναι οι τυπικές εκφράσεις (regex ή regexp). Ως τυπική έκφραση ορίζεται το σύνολο των κανόνων, με βάση τους οποίους θα πραγματοποιηθεί ο εντοπισμός των συμβολοσειρών που αντιστοιχίζονται με αυτούς. Η μέθοδος της *Αντιστοίχισης Μοτίβου-Κειμένου* είναι κατάλληλη σε περιπτώσεις όπου το ζητούμενο μοτίβο, για παράδειγμα λέξεις ή φράσεις κλειδιά, είναι επαναλαμβανόμενο και αυστηρώς καθορισμένο [2].

APIs

Πολλές φορές, η συλλογή δεδομένων από διαδικτυακές πηγές με τη χρήση των προαναφερθέντων τεχνικών δεν είναι δυνατή, είτε επειδή εφαρμόζονται μέτρα για τη διαφύλαξη των πληροφοριών ή επειδή τα δεδομένα είναι δομημένα με αρκετά περίπλοκο τρόπο, με αποτέλεσμα να καθίσταται δύσκολη η συγκέντρωσή τους. Σε αυτές τις περιπτώσεις συστήνεται η χρήση των APIs (Application Programming Interface). Ένα API συγκροτείται από ένα σύνολο διαδικασιών και πρωτοκόλλων επικοινωνίας, τα οποία διευθύνουν την πρόσβαση χρηστών στα δεδομένα μιας εφαρμογής, ενός λειτουργικού συστήματος ή άλλων υπηρεσιών. Η λειτουργία ενός API αντιστοιχεί σε αυτή του μεσολαβητή μεταξύ ενός χρήστη και μιας πηγής δεδομένων, με αποτέλεσμα πολλές εφαρμογές να μπορούν να αξιοποιήσουν δεδομένα από την ίδια πηγή, επικοινωνώντας πρώτα με το καθορισμένο API [32]. Ωστόσο, οι υπηρεσίες που παρέχει το κάθε API ορίζονται από τον δημιουργό τους και συνήθως τα δεδομένα για τα οποία επιτρέπεται η πρόσβαση είναι περιορισμένα ή απαιτείται χρηματική αμοιβή για τη συλλογή τους, τακτική την οποία εφαρμόζουν στα API τους αρκετές γνωστές πλατφόρμες, όπως το Instagram και το YouTube.

2.1.4 Διαθέσιμα Εργαλεία

Στα πλαίσια της συγκέντρωσης δεδομένων από το διαδίκτυο διατίθεται, από αρκετές γλώσσες προγραμματισμού, μία μεγάλη ποικιλία εργαλείων, με τις πιο συχνά χρησιμοποιούμενες να είναι η Java, Python και JavaScript. Παρακάτω, παρατίθενται μία συνοπτική περιγραφή των σχετικών βιβλιοθηκών και frameworks, που συναντήθηκαν κατά τη διεκπεραίωση της εργασίας.

Scrapy Framework

Το Scrapy είναι ένα framework συγκομιδής διαδικτυακών πόρων βασισμένο στη γλώσσα προγραμματισμού Python, το οποίο επιτρέπει την δημιουργία διαφορετικών τύπων *spiders* (βλ. Υποενότητα 2.1.1), με σκοπό την ανάκτηση του πηγαίου κώδικα της επιθυμητής ιστοσελίδας. Η διαδικασία συγκομιδής δεδομένων ξεκινά με τον ορισμό του συνόλου των διευθύνσεων (URLs) που πρέπει να επισκεφθούν και, στη συνέχεια, αναπτύσσεται η μέθοδος “parse”, κατά την οποία εξάγονται αυτομάτως πληροφορίες από τον κάθε ιστότοπο με την χρήση εκφράσεων XPath (βλ. Υποενότητα 2.1.3). Το Scrapy είναι βασισμένο στο Twisted, μια ασύγχρονη βιβλιοθήκη της Python που υποβάλλει τα αιτήματα με ασύγχρονο τρόπο, ενισχύοντας έτσι την απόδοση του *spider* [12], [44].

Requests

Η Requests είναι μία βιβλιοθήκη της Python, που επιτρέπει την αποστολή HTTP αιτημάτων με εύκολο τρόπο, καθώς η πολυπλοκότητα που χαρακτηρίζει την σύνταξη HTTP αιτημάτων αντικαθίσταται από ένα, φιλικό προς τον χρήστη, API. Οι δυνατότητες της βιβλιοθήκης αυτής περιορίζονται μόνο στην ανάκτηση του πηγαίου κώδικα ενός ιστότοπου και για αυτό τον λόγο η χρήση της συνοδεύεται από επιπλέον εργαλεία, με σκοπό την εξαγωγή πληροφοριών από τον κώδικα [36].

Beautiful Soup

Η Beautiful Soup είναι μία βιβλιοθήκη της Python, η οποία, σε αντίθεση με την Requests, χρησιμοποιείται για την ανάλυση/εξαγωγή δεδομένων από ήδη ανακτημένα HTML ή XML αρχεία. Η χρήση της είναι εύκολη προς τον χρήστη και μπορεί να συνδυαστεί με ένα ευρύ φάσμα από APIs, που θα βοηθήσουν στην εξαγωγή πληροφοριών από ετερογενείς πηγές. Ωστόσο, υπενθυμίζεται ότι πρακτικά η Beautiful Soup δεν είναι ένας αυτοτελής εξαγωγέας πληροφοριών, αλλά στην πραγματικότητα χρησιμοποιεί στο παρασκήνιο άλλα εργαλεία για να εξυπηρετήσουν τον σκοπό αυτό [49]. Παρέχει όλους τους απαραίτητους μηχανισμούς για τον εντοπισμό πληροφοριών κατά μήκος της δένδροειδούς απεικόνισης μίας σελίδας, για παράδειγμα CSS Selectors, όπως και για την απόσπαση των δεδομένων από αυτή.

Selenium

Το Selenium αποτελεί ένα σύστημα εργαλείων που ειδικεύονται στην αυτοματοποιημένη διαδικτυακή πλοήγηση και είναι συμβατό με αρκετές γλώσσες προγραμματισμού. Τα εργαλεία αυτά χρησιμοποιούν τεχνικές για τον απομακρυσμένο έλεγχο ενός προγράμματος περιήγησης και την προσομοίωση της ανθρώπινης αλληλεπίδρασης με αυτό. Το συγκεκριμένο σύστημα επιτρέπει στους χρήστες να αυτοματοποιούν συχνές δραστηριότητες τους, όπως τη συμπλήρωση ψηφιακών φορμών, την επιλογή συνδέσμων ή ακόμα και την κίνηση του κέρσορα σε διάφορα σημεία της σελίδας [48]. Συμπερασματικά, το Selenium μπορεί να συμβάλλει στη συλλογή δεδομένων από το διαδίκτυο, δρομολογώντας μια ακολουθία διαδοχικών βημάτων, που τελικά οδηγούν στην επιλογή των επιθυμητών στοιχείων μιας ιστοσελίδας.

2.1.5 Προκλήσεις κατά το Web Scraping

Η εκμείωση διαδικτυακών δεδομένων, ιδιαιτέρως όταν πραγματοποιείται με αυτοματοποιημένο τρόπο και όταν ο όγκος των επιθυμητών δεδομένων είναι μεγάλος, εγκυμονεί αρκετές προκλήσεις, η οποίες στέκονται ως εμπόδια στην διεκπεραίωση της διαδικασίας. Κατά της διάρκειας της παρούσας έρευνας, παρατηρήθηκαν οι παρακάτω εστίες προβληματισμού.

Μεταβολές στα περιεχόμενα των ιστοσελίδων

Ένα από τα συχνότερα προβλήματα κατά το Web Scraping αποτελεί η συντήρηση και η εξέλιξη των ιστοσελίδων-στόχων, κάτι το οποίο συνεπάγεται και τη μεταβολή της δομής του πηγαίου τους κώδικα. Μια μόνο αλλαγή σε ένα από τα επιθυμητά πεδία μιας ιστοσελίδας ενδέχεται να προκαλέσει σφάλματα ή ακόμα και να θέσει εκτός λειτουργίας ολόκληρο τον μηχανισμό εξαγωγής. Γίνεται αντιληπτό, λοιπόν, πως η τακτική ενημέρωση των λειτουργιών ενός Web Scraper, έτσι ώστε να ανταποκρίνεται στα συχνά τροποποιημένα δεδομένα του εκάστοτε ιστότοπου, καθίσταται επιτακτική ανάγκη για τη διασφάλιση της λειτουργικότητάς του.

Σελιδοποίηση περιεχομένου ιστοσελίδων

Η σελιδοποίηση (Pagination) είναι μία ευρέως χρησιμοποιούμενη τεχνική σχεδιασμού ιστοσελίδων, η οποία διαμοιράζει το περιεχόμενο διαδικτυακών τόπων σε επιμέρους σελίδες, καθιστώντας έτσι πιο ομαλή την αλληλεπίδραση ενός χρήστη με μεγάλα σύνολα δεδομένων [23]. Η μορφή της σελιδοποίησης που εφαρμόζει η κάθε ιστοσελίδα στα περιεχόμενα της, επηρεάζει την λειτουργία των Web Scrapers. Πρακτικές, όπως η αρίθμηση των σελίδων με διαδοχικούς αριθμούς,

μπορούν να βοηθήσουν την ορθή προσπέλασή τους από έναν Web Scraper, ο οποίος θα μπορεί να αναγνωρίσει το μοτίβο της σελιδοποίησης και να κινηθεί αναμεταξύ τους. Επομένως, προτείνεται, πριν την έναρξη της συλλογής δεδομένων από μία διαδικτυακή πηγή, να εξετάζεται η παρουσία ή μη σελιδοποίησης των δεδομένων και, κατ' επέκταση, η προσαρμογή της λειτουργίας του μηχανισμού συγκομιδής στα χαρακτηριστικά που τη διέπουν.

Δυναμικό περιεχόμενο ιστοσελίδων

Οι ιστοσελίδες, των οποίων ο σχεδιασμός βασίζεται στην δυναμική φόρτωση του περιεχομένου τους με τη χρήση των γλωσσών JavaScript και AJAX, ενισχύουν την πολυπλοκότητα της εξαγωγής δεδομένων από αυτές. Παρά το γεγονός ότι η δυναμική φόρτωση στοιχείων διευκολύνει την προσπέλαση τους από τους χρήστες, δεν ισχύει το ίδιο για τους Web Scrapers, καθώς αρκετοί από αυτούς έχουν τη δυνατότητα να εντοπίζουν αποκλειστικά στατικές πληροφορίες από το αρχείο HTML της κάθε ιστοσελίδας. Η απόσπαση δυναμικά εμφανιζόμενων πληροφοριών απαιτεί χρήση εξειδικευμένων εργαλείων και τεχνικών, όπως τα PhantomJS¹ και Octoparse², τα οποία δεν αναλύονται στο πλαίσιο της παρούσας πτυχιακής εργασίας.

Απαίτηση εγγραφής/σύνδεσης

Ένα ακόμα συχνό φαινόμενο, κατά τη διαδικασία του Web Scraping, είναι η απαίτηση για εγγραφή ή σύνδεση του χρήστη στο σύστημα της ιστοσελίδας, με σκοπό την απόκτηση πρόσβασης στα υπό προστασία περιεχόμενα της. Εφόσον, ένας Web Scraper πραγματοποιήσει με αυτόματο τρόπο τη διαδικασία της αυθεντικοποίησης, το πρόγραμμα περιήγησης παρέχει στον μηχανισμό συγκομιδής δεδομένων τα *cookies* της σύνδεσης, τα οποία ενσωματώνονται σε όλα τα ακόλουθα αιτήματα και αποδεικνύουν ότι αυτά προέρχονται από έναν εξουσιοδοτημένο χρήστη.

Μέτρα αποτροπής του Web Scraping από ιστοσελίδες

Παρά το γεγονός ότι υπάρχει μία πληθώρα εργαλείων με σκοπό την εξαγωγή πληροφορίας από διαδικτυακές πηγές, οι κάτοχοι των ιστοσελίδων δεν ενθαρρύνουν τους χρήστες του διαδικτύου να πράξουν κάτι ανάλογο στα δεδομένα τους. Για τον λόγο αυτό, οι διαχειριστές αρκετών διαδικτυακών τόπων εφαρμόζουν τεχνικές, έτσι ώστε να αποτραπεί η εκμετάλλευση των δεδομένων τους ή, τουλάχιστον, να ελαχιστοποιηθεί ο αντίκτυπος τέτοιου είδους πρακτικών. Στην Υποενοότητα 2.1.6 περιγράφονται με μεγαλύτερη ακρίβεια τα διάφορα μέτρα προστασίας δεδομένων που λαμβάνουν οι ιστοσελίδες, καθώς και τεχνικές για την παράκαμψη τους.

Διαχείριση μεγάλου όγκου πληροφοριών

Η εξαγωγή δεδομένων σε μεγάλη κλίμακα απαιτεί και την διαχείριση μεγάλου όγκου πληροφοριών, όπως είναι η αποθήκευση και η επεξεργασία τους. Απαραίτητα χαρακτηριστικά γνωρίσματα που πρέπει να φέρει ένα σύστημα διαχείρισης μεγάλης κλίμακας δεδομένων είναι: (i) η ανεκτικότητα σε σφάλματα, (ii) η επεκτασιμότητα, (iii) η ασφάλεια και (iv) η υψηλή διαθεσιμότητα. Εάν η βάση δεδομένων δεν έχει κατασκευαστεί με τις απαραίτητες προδιαγραφές, αν είναι για παράδειγμα ασταθής ή μη προσβάσιμη, τότε η οργάνωση, η αναζήτηση και το

¹<https://phantomjs.org>

²<https://www.octoparse.com>

φιλτράρισμα των δεδομένων χαρακτηρίζονται ως δύσκολες και χρονοβόρες διαδικασίες ή και αδύνατες σε κάποιες περιπτώσεις. Για να αντιμετωπιστεί το προαναφερθέν σενάριο, δίνεται η δυνατότητα στους ενδιαφερόμενους, αντί να στήσουν τη δική τους βάση δεδομένων από την αρχή, να χρησιμοποιούν ήδη υπάρχουσες, όπως την MongoDB³, οι οποίες ενσωματώνουν όλες τις απαραίτητες προδιαγραφές για μία σωστά δομημένη βάση δεδομένων.

Χρονική διάρκεια διαδικασίας

Ο χρόνος που καταναλώνεται για την ολοκλήρωση μιας διαδικασίας ανάκτησης πληροφοριών από ιστοσελίδες είναι επίσης ένας δεσμευτικός παράγοντας. Τα δεδομένα απαιτούν χρόνο για να αποσπαστούν από τις πηγές τους και να μεταφερθούν στη δομή αποθήκευσής τους. Γίνεται αντιληπτό, λοιπόν, πως η χρονική διάρκεια μιας διαδικασίας Web Scraping είναι ανάλογη με τον όγκο των δεδομένων που προορίζονται για συλλογή. Παράμετροι όπως οι δυνατότητες του συστήματος, η ταχύτητα της σύνδεσης στο διαδίκτυο και η τεχνική επεξεργασίας των δεδομένων παίζουν επίσης σημαντικό ρόλο στην απόδοση του μηχανισμού συγκομιδής δεδομένων. Τεχνικές όπως η ασύγχρονη εξαγωγή πληροφοριών από πολλαπλές διαδικτυακές πηγές ή η παράλληλη λειτουργία πολλαπλών Web Scrapers μπορούν να επιταχύνουν σε μεγάλο βαθμό την εξαγωγή δεδομένων, περιορίζοντας έτσι την ανεξέλεγκτη αύξηση του χρόνου που θα προέκυπτε ακολουθώντας μια απλούστερη τακτική.

2.1.6 Πολιτική προστασίας δεδομένων ενάντια στο Web Scraping

Οι ιστοσελίδες παράγουν και στεγάζουν, καθημερινά, μεγάλο όγκο δεδομένων, τα οποία διατίθενται δημόσια προς το ευρύ κοινό. Όλοι οι επισκέπτες μία ιστοσελίδας έχουν το δικαίωμα προσπέλασης και ανάγνωσης των κοινοποιημένων πληροφοριών, ωστόσο αρκετές είναι οι περιπτώσεις κατά τις οποίες θεωρείται αναγκαία η απόσπαση πληροφοριών από διαδικτυακές πηγές, όπως συμβαίνει κατά τη συγκέντρωση δεδομένων με σκοπό τη διεξαγωγή ερευνών. Αρκετοί ιστότοποι παρέχουν στο κοινό τη δυνατότητα συλλογής των δεδομένων τους, μέσω της χρήσης εξουσιοδοτημένων APIs. Παρ' όλα αυτά, όταν δεν διατίθεται κάποιο επίσημο API, η μοναδική διέξοδος για τη εξαγωγή δεδομένων από μία ιστοσελίδα αποτελεί το Web Scraping. Το γεγονός ότι πληροφορίες μία διαδικτυακής πηγής μπορούν να γίνουν αντικείμενο εκμετάλλευσης, πέραν των απλών επισκεπτών, από κακόβουλους χρήστες ή ανταγωνιστικές επιχειρήσεις, οι οποίες επιθυμούν να υποκλέψουν πληροφορίες για το επιχειρησιακό μοντέλο μιας άλλης επιχείρησης, έχει οδηγήσει πολλές ιστοσελίδες στο να λάβουν μέτρα αποτροπής του Web Scraping [45].

Η πολιτική προστασίας των δεδομένων έχει δημιουργήσει αρκετούς προβληματισμούς σε όσους επιλέγουν τη χρήση των Web Scrapers, καθώς πολλές απόπειρες εξαγωγής στοιχείων από ιστοσελίδες καταλήγουν στην απόρριψη. Οι anti-scraping τεχνολογίες είναι αρκετά ανεπτυγμένες, με αποτέλεσμα να ανιχνεύουν με ευκολία ενέργειες που προέρχονται από αυτοματοποιημένους μηχανισμούς συγκομιδής δεδομένων και να επιβάλλουν κυρώσεις, με αυτές να ποικίλουν από πιο ήπιες, όπως η απόρριψη ενός αιτήματος, σε πιο σκληρές, όπως η απαγόρευση πρόσβασης συγκεκριμένων διευθύνσεων IP σε μια ιστοσελίδα. Στις επόμενες παραγράφους παρουσιάζονται περιληπτικά συγκεκριμένες στρατηγικές που μπορούν να ακολουθηθούν, με σκοπό την αποφυγή τέτοιου είδους κυρώσεων, οι οποίες είναι απαγορευτικές στα πλαίσια της διαδικασίας συλλογής δεδομένων.

³<https://www.mongodb.com>

Σεβασμός προς το αρχείο robots.txt

Οι περισσότερες ιστοσελίδες που μερμινούν για την διαφύλαξη των περιεχομένων τους, διαθέτουν ένα αρχείο που φέρει το όνομα “robots.txt” και μπορεί να προβληθεί, αναζητώντας το URL της ιστοσελίδας ακολουθούμενο από την κατάληξη “robots.txt” (δηλαδή, https://www.<website_name>.com/robots.txt). Σε αυτό το αρχείο οριοθετείται η επιτρεπόμενη συμπεριφορά το κάθε Web Scraper απέναντι στην εκάστοτε ιστοσελίδα. Συγκεκριμένα, ορίζονται τα πεδία της ιστοσελίδας των οποίων η συλλογή επιτρέπεται ή απαγορεύεται, καθώς και η αποδεκτή συχνότητα αποστολής αιτημάτων. Συνήθως, τα περισσότερα “robots.txt” αρχεία απορρίπτουν οποιασδήποτε μορφής δραστηριότητα που σχετίζεται με το Web Scraping, ωστόσο η εξαγωγή δεδομένων παραμένει εφικτή ωστόσο εντοπίζεται από τους μηχανισμούς ασφάλειας. Συμπερασματικά, συστήνεται να μελετάται πρώτα το συγκεκριμένο αρχείο και, στη συνέχεια, να λαμβάνονται αποφάσεις σχετικά με τις ενέργειες εξαγωγής δεδομένων. Στην ακόλουθη εικόνα παρατηρείται η δομή του “robots.txt” αρχείου του YouTube.

```
# robots.txt file for YouTube
# Created in the distant future (the year 2000) after
# the robotic uprising of the mid 90's which wiped out all humans.

User-agent: Mediapartners-Google*
Disallow:

User-agent: *
Disallow: /channel/*/community
Disallow: /comment
Disallow: /get_video
Disallow: /get_video_info
Disallow: /get_midroll_info
Disallow: /live_chat
Disallow: /login
Disallow: /results
Disallow: /signup
Disallow: /t/terms
Disallow: /timedtext_video
Disallow: /user/*/community
Disallow: /verify_age
Disallow: /watch_ajax
Disallow: /watch_fragments_ajax
Disallow: /watch_popup
Disallow: /watch_queue_ajax

Sitemap: https://www.youtube.com/sitemaps/sitemap.xml
```

Σχήμα 2.4: Το “robots.txt” αρχείο του YouTube

Για παράδειγμα, με τις εντολές “User-agent: Mediapartners-Google*” και “Disallow:” που εμφανίζονται στο συγκεκριμένο αρχείο, απορρίπτεται η ευρετηρίαση της σελίδας από το πρόγραμμα ανίχνευσης του AdSense.

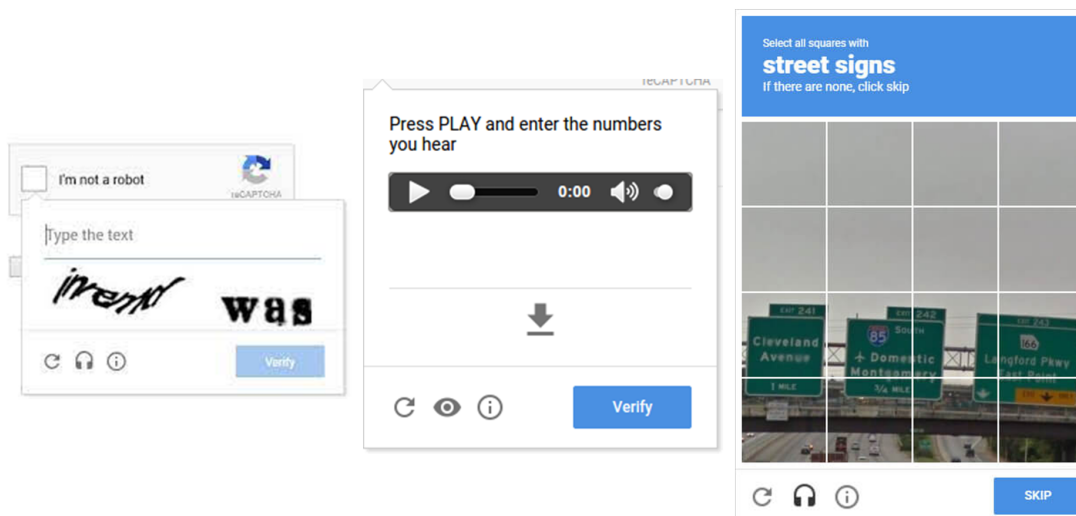
Προσομοίωση ανθρώπινης συμπεριφοράς

Η δράση ενός μέσου ανθρώπου στα πλαίσια της επίσκεψής του σε μια ιστοσελίδα θα μπορούσε να χαρακτηριστεί ακαθόριστη και απρόβλεπτη. Συνήθως, ένας επισκέπτης προβαίνει σε μια τυχαία αλληλουχία από ενέργειες, οι οποίες δεν εμφανίζουν κάποιο συγκεκριμένο μοτίβο και απέχουν μεταξύ τους τυχαία και ανθρωπίνως δυνατά χρονικά διαστήματα. Εφόσον μια τέτοια συμπεριφορά γίνεται αποδεκτή από τους εξυπηρετητές των ιστοσελίδων, προκύπτει η ιδέα για προσομοίωση της ανθρώπινης συμπεριφοράς στη λειτουργία ενός Web Scraper. Πράγματι, η ιδέα αυτή αποτελεί άλλη μια από τις μεθόδους παράκαμψης των μέτρων ασφάλειας της κάθε ιστοσελίδας, καθώς εφαρμόζοντας τυχαίες χρονικές καθυστερήσεις μεταξύ των αιτημάτων που πραγματοποιεί ο αυτόματος μηχανισμός συγκομιδής δεδομένων, οργανώνοντας με μη προβλέψιμο τρόπο τα αιτήματα ή ακόμα και προγραμματίζοντας την αυτόματη και τυχαία κίνηση του

κέρσορα στο περιεχόμενο της, είναι δυνατό να ξεγελάσουν τα ευφυή συστήματα ασφαλείας των ιστοσελίδων και να προσδοθεί στον μηχανισμό ανθρώπινος χαρακτήρας.

Χρήση μηχανισμών συμπλήρωσης CAPTCHA

Το CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) είναι μία δημοφιλής μορφή αντίστροφων δοκιμών Turing (reverse Turing tests), η οποία χρησιμοποιείται για να διατηρήσει έναν ιστότοπο ασφαλή από μηχανισμούς συγκομιδής δεδομένων. Συγκεκριμένα, το εργαλείο αυτό έχει τη δυνατότητα να αναγνωρίζει αυτοματοποιημένους Web Scrapers, μέσω της υποβολής σύντομων ερωτήσεων, των οποίων η απάντηση προϋποθέτει την ύπαρξη ανθρώπινη νοημοσύνη, και μπορεί να διακριθεί σε τρεις κατηγορίες αναγνώρισης περιεχομένου: (i) τα CAPTCHAs στατικού κειμένου, τα οποία απαιτούν την αναγνώριση παραμορφωμένου και δυσανάγνωστου κειμένου, (ii) τα CAPTCHAs ήχου, τα οποία χρησιμοποιούν παραμορφωμένη φωνητική ομιλία με σκοπό την αποτύπωσή της και (iii) τα CAPTCHAs βίντεο/εικόνας, που αντιπροσωπεύουν μεγαλύτερο εμπόδιο για την αυτοματοποιημένη πρόσβαση σε σύγκριση με το στατικό κείμενο, διότι η αναγνώριση αντικειμένων σε εικόνες με τη χρήση πληροφοριακών μέσων συνιστά ένα δύσκολο έργο. Παρά το γεγονός ότι η παρουσία των CAPTCHAs αποτελεί τροχοπέδη για τους μηχανισμούς Web Scraping, έχουν αναπτυχθεί τεχνικές, όπως για παράδειγμα το AntiCAPTCHA⁴, οι οποίες μπορούν να τα παρακάμψουν. Παρ' όλα αυτά, οι συγκεκριμένες τεχνικές μπορεί να είναι ιδιαίτερος κοστοβόρος ή, ακόμα, και να επιβραδύνουν τη συνολική διαδικασία εξαγωγής δεδομένων [23].



Σχήμα 2.5: Διαφορετικά είδη CAPTCHAs

Αποφυγή Honeypots


Η εξέλιξη των μεθόδων παραβίασης των μέτρων ασφαλείας ενάντια στο Web Scraping έχει οδηγήσει τις ιστοσελίδες στο να αναπτύσσουν ευφυέστερους τρόπους ανίχνευσης και αποκλεισμού των παραβατών. Μέρος τις προσπάθειας αυτής είναι και ο σχεδιασμός των "Honeypots".

⁴<https://anti-captcha.com/mainpage>

Τα Honeypots αντιπροσωπεύουν ένα σύνολο από κρυφούς συνδέσμους, οι οποίοι δεν είναι ορατοί στη διεπαφή χρήστη-ιστοσελίδας, αλλά ορίζονται στον πηγαίο κώδικα της. Ένας αυτοματοποιημένος μηχανισμός, ο οποίος έχει προγραμματιστεί να συλλέγει όλους τους συνδέσμους από έναν ιστότοπο, είναι αναπόφευκτο να μην υποπέσει στην συλλογή των κρυφών συνδέσμων-παγίδων. Με αυτό τον τρόπο, οι εξυπηρετητές των ιστοσελίδων είναι ικανοί να εντοπίζουν τις δραστηριότητες αλληλεπίδρασης με Honeypots και να απαγορεύουν την πρόσβαση στον μηχανισμό που κρύβεται πίσω από αυτές. Με σκοπό την αποφυγή των Honeypots συστήνεται ο έλεγχος του πηγαίου κώδικα για συνδέσμους που περιέχουν την παράμετρο "display: none" ή "visibility: hidden", καθώς και για συνδέσμους που τους έχει δοθεί ίδιο χρώμα με αυτό του φόντου της ιστοσελίδας.

Χρήση User Agents

Κάθε φορά που ένας χρήστης επιθυμεί να επισκεφθεί μία σελίδα στο διαδίκτυο, αποστέλλεται ένα αίτημα από το πρόγραμμα περιήγησης που χρησιμοποιεί, έτσι ώστε να του επιστραφεί η συγκεκριμένη ιστοσελίδα ως απάντηση. Το αίτημα απαρτίζεται από αρκετά πεδία, μέρος των οποίων είναι οι HTTP κεφαλίδες του αιτήματος, οι οποίες δημιουργούνται με σκοπό η ιστοσελίδα να μπορεί να συγκεντρώνει και να αναγνωρίζει ορισμένες σημαντικές λεπτομέρειες για τον επισκέπτη. Μια βασική HTTP κεφαλίδα, για κάθε αίτημα προς μια ιστοσελίδα, είναι και το εν ονόματι πεδίο User Agent. Το πεδίο User Agent είναι μία χαρακτηριστική συμβολοσειρά, η οποία επιτρέπει στους διακομιστές και τους ομότιμους χρήστες του δικτύου (network peers) να αναγνωρίζουν την εφαρμογή, το λειτουργικό σύστημα και την έκδοση του προγράμματος περιήγησης που χρησιμοποιείται για την αποστολή του αιτήματος από τον χρήστη. Η προσθήκη στο εκάστοτε αίτημα ενός πεδίου User Agent, που αντιστοιχεί σε κάποιο σύγχρονο περιηγητή, μπορεί να ενισχύσει την αληθοφάνεια του και να το βοηθήσει να προσπεράσει με επιτυχία αρκετούς μηχανισμούς ασφαλείας, καθώς τα HTTP αιτήματα που παράγονται από τους Web Scrapers δεν διαθέτουν προσχεδιασμένες κεφαλίδες. Ωστόσο, αξίζει να επισημανθεί πως είναι αρκετά εύκολο ο ιστότοπος-στόχος να ανιχνεύσει και, στη συνέχεια, να αποκλείσει έναν User Agent, εάν χρησιμοποιείται ο ίδιος σε κάθε αίτημα που στέλνεται από τον αυτοματοποιημένο μηχανισμό. Για να αποφευχθεί κάτι τέτοιο, ο χρήστης που χειρίζεται τον τρέχοντα μηχανισμό θα πρέπει να μεριμνήσει για την συνεχή εναλλαγή του User Agent μεταξύ των αιτημάτων, με χρήση μιας λίστα από δημοφιλή και εκσυγχρονισμένα User Agents [44].

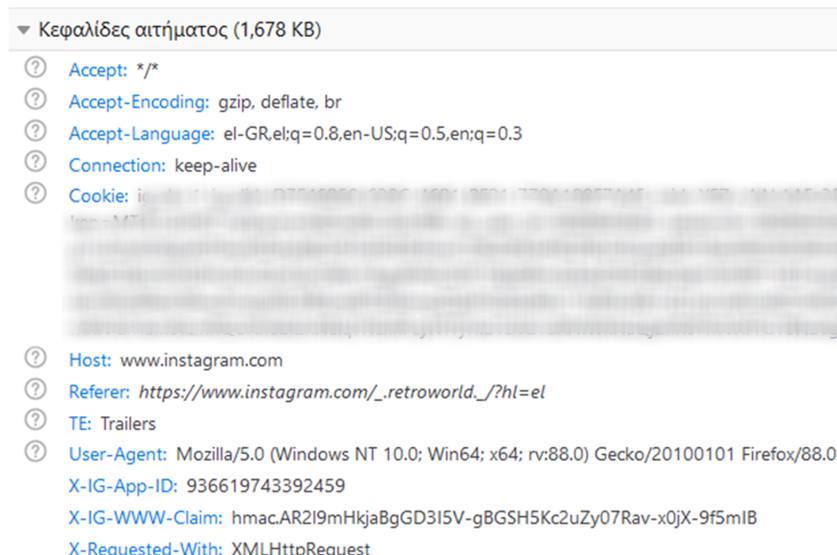
 **User-Agent:** Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:88.0) Gecko/20100101 Firefox/88.0

Σχήμα 2.6: Το πεδίο User Agent που αντιστοιχεί στον περιηγητή Mozilla Firefox

Ανάλυση των HTTP αιτημάτων

Όταν η προσθήκη ενός User Agent στα αιτήματα του μηχανισμού συγκομιδής δεδομένων δεν αρκεί για την αποφυγή του αποκλεισμού, τότε εισέρχεται στο προσκήνιο η ανάλυση επιπλέον HTTP κεφαλίδων. Κάθε HTTP αίτημα που παράγεται από έναν κοινό χρήστη στο διαδίκτυο, περιέχει ένα σύνολο από πεδία κεφαλίδων, τα οποία είναι σημαντικά για την επικοινωνία με τον εκάστοτε εξυπηρετητή και την εξακρίβωση της ψηφιακής ταυτότητας του αποστολέα. Εκτός του πεδίου User Agent που προαναφέρθηκε, υπάρχουν εξίσου σημαντικά πεδία, όπως είναι το "Referrer", που υποδηλώνει την ιστοσελίδα προς την οποία γίνεται η ανακατεύθυνση του χρήστη, και το Cookies, το οποίο ανανεώνεται με την εγκαθίδρυση νέας συνεδρίας και είναι ανα-

γκαίο για την επικοινωνία με διαδικτυακούς τόπους που απαιτούν την ύπαρξη του. Επομένως, όταν ένας χρήστης επιθυμεί να συγκεντρώσει δεδομένα από μία ιστοσελίδα, θα μπορούσε να αναλύσει τα πεδία-κεφαλίδες που παράγονται, υπό κανονικές συνθήκες, κατά την αποστολή αιτημάτων στη συγκεκριμένη ιστοσελίδα και να τα προσαρμόσει στα αιτήματα που θα αποσταλούν από τον μηχανισμό Web Scraping [15].



Σχήμα 2.7: Κεφαλίδες HTTP αιτήματος προς ένα προφίλ του Instagram

Δυναμική εναλλαγή διεύθυνσης IP

Κατά την εξαγωγή δεδομένων από μία ιστοσελίδα, η διεύθυνση IP του χρήστη που επιτελεί αυτή την διαδικασία, είναι ορατή προς τον εξυπηρετητή (server) της ιστοσελίδας. Το συγκεκριμένο γεγονός οφείλεται στο ότι ο ιστότοπος στον οποίο πραγματοποιείται η επίσκεψη, έχει την δυνατότητα να συλλέγει πληροφορίες σχετικά με τους χρήστες και να ελέγχει αν κάποιος από αυτούς συμπεριφέρεται ύποπτα, δηλαδή αν συλλέγει μεγάλο όγκο δεδομένων, αν χρησιμοποιεί κάποιο σταθερό μοτίβο κατά τη διάρκεια της συλλογής δεδομένων ή αν στέλνει πολλά και αλληλέγγυα αιτήματα, ενέργειες που κάποιος άνθρωπος δεν θα ήταν δυνατό να πραγματοποιήσει χωρίς τη χρήση αυτοματοποιημένου μηχανισμού. Σε περίπτωση που παρατηρηθεί κάποιο από τα προαναφερθέντα, γίνεται αντιληπτό ότι ο επισκέπτης υπέπεσε σε παραβίαση των όρων προστασίας των δεδομένων και, ως εκ τούτου, αποκλείεται η διεύθυνση IP του από την τρέχουσα ιστοσελίδα με συνοπτικές διαδικασίες [4]. Μια από τις μεθόδους για την αποτροπή του αποκλεισμού της διεύθυνσης IP είναι η δημιουργία ενός ηλεκτρονικού αποθετηρίου, το οποίο θα απαρτίζεται από μία λίστα με διαφορετικά IPs (proxy servers). Κάθε φορά που ο άνωθεν χρήστης θα συγκεντρώνει ένα πλήθος δεδομένων, ανάλογα με τους κανονισμούς της εκάστοτε ιστοσελίδας, θα μεριμνά για την έγκαιρη και αυτόματη εναλλαγή της διεύθυνσης IP, η οποία θα “περιβάλλει” την διεύθυνση IP της συσκευής του. Παρόμοιες λύσεις στο παραπάνω πρόβλημα προσφέρει και χρήση των VPN⁵ και TOR⁶ [15]. Ωστόσο, η σύνδεση σε απομακρυσμένο εξυπηρετητή/μεσολαβητή επιβαρύνει την ταχύτητα αποστολής και λήψης αιτημάτων, κάτι το οποίο μπορεί να αποβεί μοιραίο για τη συνολική διαδικασία, προκαλώντας σφάλματα

⁵<https://us.norton.com/internetsecurity-privacy-what-is-a-vpn.html>

⁶<https://www.torproject.org>

εξάντλησης του χρόνου αναμονής (timeout). Η διαδικασία που μόλις περιγράφηκε, ονομάζεται δυναμική εναλλαγή διεύθυνσης IP.

Χρήση Headless Browsers

Τα εργαλεία που χρησιμοποιούν οι ιστοσελίδες ενάντια στο Web Scraping γίνονται καθημερινά εξυπνότερα, καθώς η τεχνητή νοημοσύνη που βρίσκεται πίσω αυτά τροφοδοτείται συνεχώς με νέα δεδομένα, καθιστώντας, έτσι, όλο και πιο περίπλοκες τις τεχνικές αποφυγής τους. Πλέον, είναι δυνατό να ανιχνευθεί ένας μηχανισμός συγκομιδής δεδομένων ακόμα και από τον κώδικα της JavaScript που εκτελείται κατά την αλληλεπίδραση του χρήστη με το παραθυρικό περιβάλλον του κάθε περιηγητή. Λύση στο πρόβλημα αυτό μπορεί να παρέχει η χρήση των “Headless” περιηγητών (Headless browsers). Ένας headless browser είναι ένας πλήρως λειτουργικός περιηγητής ιστοσελίδων, ο οποίος δεν διαθέτει παραθυρικό περιβάλλον, αλλά η διαχείρισή του πραγματοποιείται με τη χρήση της γραμμής εντολών. Η έλλειψη της διεπαφής χρήστη-ιστοσελίδας ευνοεί τη λειτουργία ενός Web Scraper, καθώς αποφεύγεται η επαφή με πιθανές παγίδες, που μπορεί να ελλοχεύουν στα κομμάτια κώδικα που αφορούν την εμφάνιση της ιστοσελίδας, και επιταχύνεται η διαδικασία συλλογής δεδομένων, λόγω του χαμηλότερου φόρτου κώδικα. Εντούτοις, η πολυπλοκότητα που εγκυμονεί το στήσιμο ενός Headless περιηγητή, αντισταθμίζει τα πλεονεκτήματα της χρήσης του. Υπάρχουν αρκετές διαθέσιμες βιβλιοθήκες που μπορούν να υποστηρίξουν την λειτουργία των “Headless” περιηγητών, με τις πιο γνωστές να είναι το Selenium⁷, το Puppeteer⁸ και το Playwright⁹ [15], [42].

Χρήση κρυφής μνήμης Google

Σε περιπτώσεις όπου καμία από τις παραπάνω μεθόδους δεν είναι δυνατό να εφαρμοστεί, διατίθεται η λιγότερο διαδεδομένη διέξοδος της χρήσης της κρυφής μνήμης Google (Google Cache). Η Google διατηρεί αποθηκευμένα αντίγραφα μιας πληθώρας διαδικτυακών τόπων, με σκοπό την χρήση αυτών όταν ο αντίστοιχος ιστότοπος δεν είναι διαθέσιμος. Η προκειμένη πρακτική καθιστά εφικτή τη συλλογή των επιθυμητών πληροφοριών από την κρυφή μνήμη της Google. Παρά το γεγονός ότι μέσω της χρήσης της κρυφής μνήμης αποφεύγεται η αλληλεπίδραση με τους μηχανισμούς ασφαλείας της εκάστοτε ιστοσελίδας, είναι αρκετά πιθανό να μην υπάρχει κάποιο αντίγραφο των δεδομένων που αναζητούνται. Η παραπάνω διαδικασία μπορεί να επιτευχθεί αναζητώντας των σύνδεσμο της ιστοσελίδας που στεγάζει τα απαραίτητα δεδομένα συνοδευόμενο από το πρόθεμα “http://webcache.googleusercontent.com/search?q=cache:” [50].

2.2 Βάσεις Δεδομένων

Επόμενο στάδιο, μετά την συλλογή δεδομένων με την χρήση του κατάλληλου μηχανισμού, αποτελεί η αποθήκευσή τους σε μία βάση δεδομένων που ικανοποιεί τις λειτουργικές απαιτήσεις του αντίστοιχου έργου. Η επιλογή της βέλτιστης βάσης δεδομένων χαρακτηρίζεται ως μία πολύπλοκη διαδικασία, κατά την οποία ο ενδιαφερόμενος θα πρέπει να αξιολογήσει σωστά την κατάσταση, ώστε να μη χρειαστεί να βρεθεί αντιμέτωπος με μελλοντικά προβλήματα, όπως ο ελλιπής χώρος αποθήκευσης δεδομένων ή η τροποποίηση των δεδομένων πριν την αποθήκευση, και δυσλειτουργίες, όπως μεγάλες καθυστερήσεις διεξαγωγής αιτημάτων στη βάση. Με

⁷<https://www.selenium.dev>

⁸<https://github.com/puppeteer/puppeteer>

⁹<https://github.com/microsoft/playwright>

σκοπό να αποφευχθούν τα παραπάνω, θα πρέπει να ληφθούν υπόψη ορισμένοι σημαντικοί παράγοντες, οι οποίοι θα καθορίσουν και την τελική επιλογή του χρήστη. Συγκεκριμένα, πρώτο κριτήριο επιλογής συνιστά (i) ο όγκος των δεδομένων προς συλλογή, καθώς κάθε σύστημα διαχείρισης δεδομένων παρέχει περιορισμένο χώρο αποθήκευσης, στον οποίο μπορούν να αποθηκευτούν και να υποβληθούν σε επεξεργασία οι καταχωρήσεις. Δεύτερο κριτήριο, είναι (ii) η υποδομή και το κόστος της βάσης, λαμβάνοντας υπόψη ότι τα δεδομένα που είναι διαθέσιμα στο διαδίκτυο αυξάνονται καθημερινά με ραγδαίους ρυθμούς. Έτσι, ο χρήστης καλείται να αποφασίσει αν το λογισμικό που θα στεγάζονται τα δεδομένα του θα φιλοξενηθεί σε τοπικό επίπεδο (localhost) ή στο λεγόμενο νέφος (cloud-based), επιλογή η οποία μπορεί να επηρεάσει και το χρηματικό κόστος της αποθήκευσης των δεδομένων. Επόμενος παράγοντας είναι (iii) η ταχύτητα διεξαγωγής αιτημάτων στην βάση, η οποία θα πρέπει να πληρεί τις ανάγκες και τις απαιτήσεις που σχετίζονται με τον στόχο της εκάστοτε διαδικασίας. Το τέταρτο κριτήριο αντιστοιχεί σε εκείνο (iv) της επεκτασιμότητας μίας βάσης δεδομένων. Η επεκτασιμότητα αντιπροσωπεύει την ικανότητα της εκάστοτε βάσης να διατηρεί την λειτουργικότητα της, παρά τον αυξανόμενο φόρτο των αποθηκευμένων δεδομένων, αλλά και ενδεχόμενες αλλαγές στη δομή τους, με τους δύο βασικούς άξονες να είναι η κατακόρυφη (αύξηση της υπολογιστικής ισχύος του συστήματος αποθήκευσης) και η οριζόντια (αύξηση των συστημάτων αποθήκευσης) επεκτασιμότητα. Τελευταίο, αλλά εξίσου σημαντικό, είναι (v) το κριτήριο της ασφάλειας των δεδομένων. Θα πρέπει να διασφαλιστεί ότι τηρούνται όλα τα απαραίτητα μέτρα ασφαλείας των δεδομένων, όπως ο έλεγχος πρόσβασης των χρηστών, η ακεραιότητα των δεδομένων, καθώς και η παροχή λειτουργιών δημιουργίας αντιγράφων ασφαλείας και ανάκτησης δεδομένων.

2.2.1 SQL έναντι NoSQL

Όπως προαναφέρθηκε, ένα από τα σημαντικότερα διλήμματα κατά την επιλογή μιας βάσης δεδομένων είναι η ανάπτυξη μιας SQL/Σχεσιακής ή NoSQL/Μη-σχεσιακής δομής. Τα πλεονεκτήματα που προσφέρει κάθε μία από τις παραπάνω προσεγγίσεις είναι πολυάριθμα, ωστόσο οι διαφορές μεταξύ τους είναι αυτές που καθορίζουν την καταλληλότητα της κάθε στρατηγικής για το εκάστοτε πρόβλημα [13].

Οι SQL βάσεις δεδομένων είναι σχεσιακές βάσεις και χρησιμοποιούν τη γλώσσα ερωτημάτων SQL (Structured Query language) για τον καθορισμό της μορφής των δεδομένων και την επεξεργασία τους. Η SQL είναι η τυπική γλώσσα που χρησιμοποιείται για την αλληλεπίδραση και την διαχείριση σχεσιακών βάσεων δεδομένων (RDBMS), οι οποίες καθορίζουν τις σχέσεις/εξαρτήσεις μεταξύ των αποθηκευμένων οντοτήτων με την μορφή πίνακα (table). Η γλώσσα αυτή είναι ιδιαίτερα χρήσιμη στο χειρισμό δομημένων δεδομένων, δηλαδή δεδομένων που κατέχουν μία αυστηρά καθορισμένη δομή. Χαρακτηριστικά παραδείγματα σχεσιακών βάσεων δεδομένων αποτελούν οι PostgreSQL¹⁰ και MySQL¹¹.

Από την άλλη μεριά, οι NoSQL βάσεις δεδομένων είναι μη-σχεσιακές βάσεις, καθώς δεν απαιτούν την ύπαρξη σχήματος, στο οποίο θα αποδίδεται η προκαθορισμένη δομή των αποθηκευμένων δεδομένων. Λόγω του γεγονότος ότι στην περίπτωση των NoSQL βάσεων τα δεδομένα είναι μη-δομημένα, η αποθήκευσή τους δεν πραγματοποιείται σε μορφή πίνακα, αλλά είναι προσανατολισμένη σε δομές στηλών (column oriented), εγγράφων (document oriented), γράφων (graph based) ή και ζευγών κλειδιού-τιμής (key-value stores). Κάποιες από τις πιο δημοφιλείς

¹⁰<https://www.postgresql.org>

¹¹<https://www.mysql.com>

μη-σχεσιακές βάσεις δεδομένων είναι οι MongoDB και Apache Cassandra¹².

Στον ακόλουθο πίνακα παρουσιάζονται οι κύριες διαφορές μεταξύ των SQL και NoSQL βάσεων δεδομένων [43].

SQL	NoSQL
<ul style="list-style-type: none"> • Είναι σχεσιακές βάσεις δεδομένων • Έχουν προκαθορισμένο σχήμα για δομημένα δεδομένα • Είναι κατακόρυφα επεκτάσιμες • Έχουν δομή πίνακα 	<ul style="list-style-type: none"> • Είναι μη σχεσιακές βάσεις δεδομένων • Έχουν δυναμικό σχήμα για μη δομημένα δεδομένα • Είναι οριζόντια επεκτάσιμες • Η δομή του ποικίλει ανάλογα με την περίπτωση (δομή στηλών, εγγράφων, γράφων ή ζευγών κλειδιού-τιμής)

Πίνακας 2.4: Κύριες διαφορές μεταξύ SQL και NoSQL βάσεων δεδομένων

2.2.2 Localhost έναντι Cloud-based

Η συλλογή δεδομένων είναι ένα από τα βασικότερα σημεία για την διεκπεραίωση μίας έρευνας, καθώς βοηθάει στην εξαγωγή στατιστικών στοιχείων και αριθμητικών δεδομένων. Ωστόσο, στο παρασκήνιο αυτής της διαδικασίας δημιουργείται ένας ιδιαίτερος προβληματισμός που αφορά το μέρος στο οποίο θα αποθηκευτούν όλες οι σχετικές πληροφορίες της έρευνας. Η αποθήκευση δεδομένων αποτελεί παράγοντα ζωτικής σημασίας είτε πρόκειται για τον απλό χρήστη είτε για επιχειρήσεις. Οι δύο επικρατέστερες επιλογές, την παρούσα χρονική στιγμή, είναι η τοπική αποθήκευση (Localhost) και η αποθήκευση στο “νέφος” (Cloud-based), οι οποίες εξυπηρετούν διαφορετικούς σκοπούς και, ως εκ τούτου, προτιμούνται σε διαφορετικές καταστάσεις.

Ο όρος “Localhost” απευθύνεται στην διαδικασία αποθήκευσης και συντήρησης δεδομένων σε υλικό εξοπλισμό που στεγάζεται στο περιβάλλον του χρήστη, όπως εξωτερικές μονάδες αποθήκευσης ή προσωπικά υπολογιστικά συστήματα. Η τοπική αποθήκευση δεδομένων προτιμάται, συνήθως, από μεμονωμένους χρήστες που διαχειρίζονται μικρό μέγεθος δεδομένων. Αντιθέτως, κατά την αποθήκευση δεδομένων στο “νέφος”, τα ψηφιακά δεδομένα αποστέλλονται και φιλοξενούνται σε απομακρυσμένες εγκαταστάσεις, οι οποίες είναι ιδιόκτητες από εξωτερικές επιχειρήσεις. Η “Cloud-based” τακτικές αποθήκευσης διευκολύνουν την παράλληλη πρόσβαση των χρηστών σε ένα κοινό σύνολο δεδομένων, γεγονός που τις καθιστά ιδανικές για επιχειρήσεις και, γενικότερα, ομάδες χρηστών [6].

Παρακάτω, στους Πίνακες 2.5 και 2.6, παρατίθενται τα πλεονεκτήματα και τα μειονεκτήματα που συνοδεύουν την εκάστοτε στρατηγική αποθήκευσης δεδομένων [5], [30], [22], [6].

2.3 Μελέτη δεδομένων

Στο τελικό στάδιο μιας εμπειριστατωμένης έρευνας συναντάται η μελέτη των συλλεχθέντων δεδομένων και η εξαγωγή γνώσης από αυτά. Ωστόσο, για να στεφθεί με επιτυχία η εξαγωγή

¹²<https://cassandra.apache.org>

Πλεονεκτήματα	Μειονεκτήματα
<ul style="list-style-type: none"> • Πλήρης έλεγχος των δεδομένων • Γρήγορη πρόσβαση • Δεν απαιτείται σύνδεση στο διαδίκτυο για την πρόσβαση στα δεδομένα 	<ul style="list-style-type: none"> • Ευάλωτη σε απώλεια δεδομένων λόγω φυσικών φθορών του εξοπλισμού • Μεγαλύτερο κόστος για την ασφάλεια των δεδομένων • Δυσκολότερη επεκτασιμότητα • Περισσότερα έξοδα για εγκατάσταση και συντήρηση της βάσης δεδομένων • Δυσκολότερη προσβασιμότητα από εξωτερικούς χρήστες

Πίνακας 2.5: Πλεονεκτήματα και μειονεκτήματα της τοπικής αποθήκευσης δεδομένων

Πλεονεκτήματα	Μειονεκτήματα
<ul style="list-style-type: none"> • Ευκολότερη επεκτασιμότητα • Προσπελάσιμη από οπουδήποτε υπάρχει σύνδεση στο διαδίκτυο • Λιγότερο δαπανηρή σε γενικές γραμμές • Υψηλότερη και ευκολότερα επιτεύξιμη η ασφάλεια των δεδομένων • Υψηλός βαθμός συνέπειας των δεδομένων • Ευκολότερη πρόσβαση από παράλληλους χρήστες 	<ul style="list-style-type: none"> • Μεγαλύτερες καθυστερήσεις κατά την πρόσβαση στα δεδομένα • Περιορισμένος έλεγχος σε σχέση με την τοπική αποθήκευση • Αδυναμία πρόσβασης στα δεδομένων χωρίς σύνδεση στο διαδίκτυο • Εξάρτηση της απόδοσης του διακομιστή από τη συμφόρηση του

Πίνακας 2.6: Πλεονεκτήματα και μειονεκτήματα της αποθήκευσης δεδομένων στο “νέφος”

συμπερασμάτων, θα πρέπει να έχει τηρηθεί ευλαβικά μία ακολουθία προτεινόμενων βημάτων, που προηγείται του τελικού αυτού σταδίου της έρευνας. Στην συνέχεια της προκειμένης ενότητας, περιγράφονται τα διαδοχικά βήματα προς μία επιτυχημένη διεξαγωγή έρευνας, καθώς και προκλήσεις που ενδέχεται να προκύψουν [47], [3], [34].

2.3.1 Τα στάδια της ερευνητικής διαδικασίας

Η ερευνητική διαδικασία μπορεί να διαμοιραστεί σε έξι επιμέρους στάδια, με σκοπό την μεθοδική και εμπειριστατωμένη διεκπεραίωση της. Τα στάδια αυτά διακρίνονται ως εξής:

1. Διατύπωση του προβλήματος

Το πρώτο βήμα της διαδικασίας είναι η διατύπωση του προβλήματος, κατά το οποίο επιλέγεται ο θεματικός άξονας που θα ερευνηθεί, και διατυπώνονται ερωτήματα γύρω από αυτόν, που χρίζουν απάντησης.

2. Αξιολόγηση του διαθέσιμου υλικού προς διερεύνηση

Στη συνέχεια, αξιολογείται η ικανοποιησιμότητα του διαθέσιμου πλήθους πληροφοριών

που αφορούν το θέμα της έρευνας.

3. **Καθορισμός των στόχων της έρευνας και της μεθοδολογίας που θα ακολουθηθεί**
Έπειτα, καθορίζονται οι βασικοί στόχοι, τους οποίους θα κληθεί να επιτύχει η έρευνα, και σχεδιάζεται το πλάνο που θα ακολουθηθεί για την ολοκλήρωση της.
4. **Συλλογή των δεδομένων**
Στο παρόν στάδιο, λαμβάνει χώρα η συγκομιδή των απαραίτητων δεδομένων από ένα σύνολο πηγών, τα οποία θα αποτελέσουν τη βάση της εξαγωγής των τελικών συμπερασμάτων.
5. **Επεξεργασία και ανάλυση των δεδομένων**
Ακολουθώς, τα συλλεχθέντα δεδομένα υποβάλλονται σε επεξεργασία, με σκοπό την μετάφραση τους σε γνώση.
6. **Εξαγωγή συμπερασμάτων και ολοκλήρωση της έρευνας**
Τέλος, διατυπώνονται και αναπαρίστανται τα τελικά συμπεράσματα με τρόπο κατανοητό προς το ευρύ κοινό.

2.3.2 Προκλήσεις κατά την ερευνητική διαδικασία

Η διαδικασία της ανάλυσης των δεδομένων, λόγω των χρονικών περιορισμών, του όγκου των δεδομένων και της εμπειρίας που απαιτεί ώστε να εκπονηθεί, συνοδεύεται από αρκετές προκλήσεις που απασχολούν τους ερευνητές. Με αφορμή το γεγονός αυτό, στην παρούσα υποενότητα έχουν επιλεγεί ορισμένες από τις σημαντικότερες και τις συνηθέστερες προκλήσεις που μπορεί να συναντήσει ένας εν δυνάμει ερευνητής [37].

1. **Σωστή διαχείριση χρόνου**
Η διεξαγωγή μίας έρευνας αποτελεί μία χρονοβόρα διαδικασία και, ως εκ τούτου, ο σχεδιασμός ενός χρονοδιαγράμματος των διαδοχικών βημάτων προς την ολοκλήρωση του τελικού στόχου φαίνεται να είναι αδήριτος. Είναι σημαντικό να αφιερωθεί ο απαραίτητος χρόνος στην οργάνωση του έργου, ώστε να αποφευχθούν ζημιογόνες παραπομπές στις απαρχές της διαδικασίας, κατά τα μεταγενέστερα στάδια της έρευνας.
2. **Επιβλητικός όγκος δεδομένων**
Το διαδίκτυο επεκτείνεται καθημερινά με ιλιγγιώδεις ρυθμούς, φαινόμενο που επιβάλλει στον ερευνητή να επιλέξει με σύνεση ένα υποσύνολο των διαθέσιμων δεδομένων και όχι να τα καταχραστεί. Αρκετές είναι οι φορές, κατά τις οποίες δίνεται μεγαλύτερη σημασία στην ποσότητα των πληροφοριών μιας έρευνας, παρά στην ποιότητα τους, με αποτέλεσμα να ευδοκιμούν οι πιθανότητες αποπροσανατολισμού από τον θεματικό της άξονα. Τακτικές όπως η διαλογή και ο ορισμός ενός ανώτατου ορίου των υπό συλλογή δεδομένων, καθώς και η αξιοποίηση ήδη υπάρχουσας γνώσης, μπορούν να συμβάλλουν στην ευκολότερη διαχείριση των αστείρευτων πηγών πληροφορίας.
3. **Διαχείριση ακραίων περιπτώσεων**
Συχνό φαινόμενο αποτελεί η ύπαρξη ακραίων περιπτώσεων στο σύνολο των συγκεντρωμένων δεδομένων, οι οποίες μπορούν εύκολα να κατευθύνουν τα συμπεράσματα μακριά από την αντιπροσωπευτική τους εικόνα. Η εξαίρεση αυτών των περιπτώσεων από τις στατιστικές μετρήσεις είναι δυνατό να προσδώσει μεγαλύτερη αξιοπιστία στα εξαγο-

μένα συμπεράσματα. Ωστόσο, αν κρίνεται αναγκαία η συμπερίληψη όλων των διαθέσιμων δειγμάτων, ο εκάστοτε ερευνητής θα πρέπει να προσμένει πως τα αποτελέσματα της έρευνας θα απέχουν ως ένα βαθμό από την πραγματικότητα, ή ακόμα και να υπολογίσει το μέγεθος της απόκλισης που δημιουργούν οι ακραίες περιπτώσεις.

4. *Αντικειμενική στάση απέναντι στα δεδομένα*

Κάθε έρευνα θα πρέπει να αναπαριστά μία όψη της αντικειμενικής πραγματικότητας και όχι να επηρεάζεται από υποκειμενικά κριτήρια. Συνεπώς, είναι αναγκαίο να διατηρείται μία αμερόληπτη στάση απέναντι στα δεδομένα και ο ερευνητής να αφήνει τις πληροφορίες να τον οδηγήσουν στα αντίστοιχα συμπεράσματα, αντί να προσπαθεί να κατευθύνει τα αποτελέσματα, με απώτερο σκοπό να ταυτιστούν με τις προσωπικές του προσδοκίες.

5. *Παγίωση των προκύπτων συμπερασμάτων*

Σε περιπτώσεις που τα αποτελέσματα μιας έρευνας δεν απαντούν πλήρως σε όλα τα αρχικά ερωτήματα που είχαν τεθεί ως στόχοι, τότε δίνεται η επιλογή στον ερευνητή να εφαρμόσει επιπλέον τεχνικές για την συγκέντρωση νέων δεδομένων, τα οποία θα επαυξήσουν και θα ενισχύσουν την ήδη εξαχθείσα γνώση. Πρακτικές όπως η διεξαγωγή ενός ερωτηματολογίου μπορούν να εστιάσουν σε πληροφορίες που δεν θα ήταν εφικτό να αποκομιστούν μέσω των αρχικά εφαρμοσμένων μεθόδων.

2.3.3 Διεξαγωγή ερωτηματολογίου

Όπως αναφέρθηκε στην Υποενότητα 2.3.2, η διεξαγωγή ερωτηματολογίου συνιστά ένα επικοινωνικό ερευνητικό εργαλείο για τη συγκομιδή δεδομένων, που αφορούν ένα συγκεκριμένο θέμα. Ένα ερωτηματολόγιο απαρτίζεται από μία σειρά ερωτήσεων, που απευθύνονται σε ένα στοχευμένο υποσύνολο του κοινού και έχουν ως στόχο να σκιαγραφήσουν τη άποψη αυτού γύρω από τον θεματικό άξονα της έρευνας. Με τη χρήση του ερωτηματολογίου διευκολύνεται η ταχεία συγκέντρωση δεδομένων και δίνεται η δυνατότητα για εστίαση σε ιδιαίτερης φύσεως πληροφορίες, όπως δημογραφικά στοιχεία ή προσωπικές απόψεις συγκεκριμένων κοινωνικών ομάδων.

Όσον αφορά τον χαρακτήρα των ερωτήσεων ενός ερωτηματολογίου διακρίνονται δύο βασικές κατηγορίες ερευνητικής διαδικασίας, η ποσοτική και η ποιοτική, οι οποίες είναι συμπληρωματικές μέθοδοι που μπορούν να συνδυαστούν, με σκοπό τον σχηματισμό πληρέστερων συμπερασμάτων. Κατά την ποιοτική μέθοδο, οι ερωτήσεις είναι περισσότερο αφηρημένες και επιδιώκουν να αποσπάσουν διάφορες απόψεις αναφορικά με ένα συγκεκριμένο θέμα. Η διατύπωση αυτής της πληθώρας απόψεων επιτυγχάνεται μέσω ανοιχτού τύπου ερωτήσεων. Τέτοιου τύπου ερωτήσεις είναι δύσκολο να αναλυθούν από τους ερευνητές, αλλά μπορούν να προσδώσουν βαθύτερο νόημα στη φύση του υπό διερεύνηση προβλήματος [14]. Στην περίπτωση της ποσοτικής μεθόδου συλλέγονται τυποποιημένες απαντήσεις, οι οποίες μπορούν να μεταφραστούν άμεσα σε στατιστικά στοιχεία γύρω από το αντικείμενο της μελέτης. Η μέθοδος αυτή μπορεί να οδηγήσει ευκολότερα σε συμπεράσματα στατιστικής φύσεως χωρίς, ωστόσο, αυτά να παρέχουν μία βαθύτερη αναπαράσταση του προβλήματος, αλλά μία γενικότερη εικόνα του [46].

Κεφάλαιο 3

Εργαλεία που χρησιμοποιήθηκαν

Στο τρέχον κεφάλαιο γίνεται αναφορά στα εργαλεία που χρησιμοποιήθηκαν στην παρούσα πτυχιακή εργασία, με σκοπό την συγκέντρωση, την αποθήκευση και την αναπαράσταση των δεδομένων. Στο Κεφάλαιο 2, στην Υποενότητα 2.1.4, έχει παρουσιαστεί ένα σύνολο εργαλείων που διατίθενται προς τους χρήστες που επιθυμούν τη συγκομιδή δεδομένων από διαδικτυακές πηγές, ωστόσο παρατέθηκαν μόνο τα στοιχειώδη χαρακτηριστικά τους, χωρίς να αναλυθεί εκτενέστερα ο τρόπος λειτουργίας και τα πλεονεκτήματα που αυτά προσφέρουν. Στα πλαίσια της προκείμενης έρευνας, κατά το στάδιο της συγκομιδής δεδομένων από τα μέσα κοινωνικής δικτύωσης Instagram και YouTube, αξιοποιήθηκε η παρεχόμενη λειτουργικότητα των **Scrapy Framework**[8] και **YouTube Data API**[55]. Η αποθήκευση και η διαχείριση των συγκεντρωμένων δεδομένων πραγματοποιήθηκε με τη βοήθεια της **MongoDB**[25] βάσης δεδομένων και των υπηρεσιών που αυτή διαθέτει, **MongoDB Compass**[29] και **MongoDB Atlas**[24] και **MongoDB Charts**[26]. Τέλος, με σκοπό τη σύνταξη ερωτηματολογίου, το οποίο θα εμπλούτιζε τα αποτελέσματα της έρευνας, έγινε χρήση της πλατφόρμας **Google Forms**[21]. Στις επόμενες ενότητες παρουσιάζονται διεξοδικά τα παραπάνω εργαλεία.

3.1 Scrapy Framework

3.1.1 Επισκόπηση των δυνατοτήτων του Scrapy

Το Scrapy είναι ένα framework ανοιχτού κώδικα, βασισμένο στη γλώσσα προγραμματισμού Python, το οποίο χρησιμοποιείται για την εξαγωγή δομημένων δεδομένων από διάφορους διαδικτυακούς τόπους. Αν και βασική αρμοδιότητα του αποτελεί το Web Scraping, μπορεί να χρησιμοποιηθεί επίσης για την συγκομιδή δεδομένων μέσω APIs, καθώς και ως ένας Web Crawler γενικού σκοπού.

Το Scrapy Framework διέπεται από ένα πλήθος χαρακτηριστικών, τα οποία καθιστούν ευκολότερη την διαδικασία του Web Scraping και παρέχουν στον χρήστη όλα τα απαραίτητα εφόδια για την επιτυχή διεκπεραίωση της. Τα σημαντικότερα από αυτά είναι τα εξής [8], [39], [29]:

- Ενσωματωμένη υποστήριξη για την επιλογή και εξαγωγή δεδομένων από HTML/ XML μέσω της χρήσης CCS Selectors και XPath εκφράσεων
- Δυνατότητα αποθήκευσης των εξαγόμενων δεδομένων σε JSON, CSV, XML αρχεία
- Παροχή του Scrapy Shell, μίας διεπαφής η οποία επιτρέπει στον χρήστη να εκτελεί διάφορες εργασίες, όπως η αποσφαλμάτωση (debugging) του κώδικα του, η αποστολή αιτημάτων και η διαχείριση αντίστοιχων απαντήσεων, και η επεξεργασία των ρυθμίσεων που θα εφαρμοστούν κατά την εξαγωγή δεδομένων
- Ασύγχρονη διαχείριση των αιτημάτων βασισμένη στην βιβλιοθήκη Twisted, γεγονός που επιταχύνει σε μεγάλο βαθμό τη διαδικασία συλλογής δεδομένων
- Μεγάλο πλήθος ρυθμίσεων και ενσωματωμένων επεκτάσεων, τα οποία μπορούν να φανούν χρήσιμα σε ποικίλες περιπτώσεις, όπως η αποφυγή του αποκλεισμού από ιστοσελίδες (διαχείριση cookies, χρήση δυναμικά εναλλασσόμενων IPs/ User Agents, ελεγχόμενο χρονικό περιθώριο μεταξύ των αιτημάτων) ή ο χειρισμός των χειρισμός των επιμέρους HTTP αιτημάτων
- Ισχυρή επεκτασιμότητα, λόγω της δυνατότητας ενισχύσεις της λειτουργικότητας με νέες εξατομικευμένες λειτουργίες

3.1.2 Αρχιτεκτονική του Scrapy

Η αρχιτεκτονική του Scrapy (βλ. Σχήμα 3.1) μπορεί να διασπαστεί σε επτά επιμέρους δομικά στοιχεία [8]:

- **Scrapy Engine**

Το *Scrapy Engine* συνιστά την κεντρική μονάδα του μηχανισμού συγκομιδής δεδομένων, καθώς είναι υπεύθυνο για τον έλεγχο της ροής δεδομένων και για την λήψη αποφάσεων, ως αντίδραση σε συγκεκριμένα συμβάντα.

- **Scheduler**

Το *Scheduler* λαμβάνει αιτήματα από το *Engine*, τα οποία ταξινομούνται και επιστρέφονται στο *Engine* όταν αυτά ζητηθούν.

- **Downloader**

Το *Downloader* είναι αρμόδιο για την ανάκτηση των επιθυμητών ιστοσελίδων και την προώθηση τους στο *Engine*, μέσω του οποίου, στη συνέχεια, οδηγούνται στα *Spiders*.

- **Spiders**

Τα *Spiders* είναι συντασσόμενες, από τον χρήστη, κλάσεις, μέσω των οποίων εξάγονται τα επιθυμητά δεδομένα από τις ανακτημένες ιστοσελίδες και διαμορφώνονται τα αιτήματα που θα αποσταλούν στη συνέχεια.

- **Item Pipeline**

Το *Item Pipeline* επεξεργάζεται τα δεδομένα που έχουν μόλις εξαχθεί από τα *Spiders*· οι αρμοδιότητες του ποικίλουν από το φιλτράρισμα και την επικύρωση των δεδομένων,

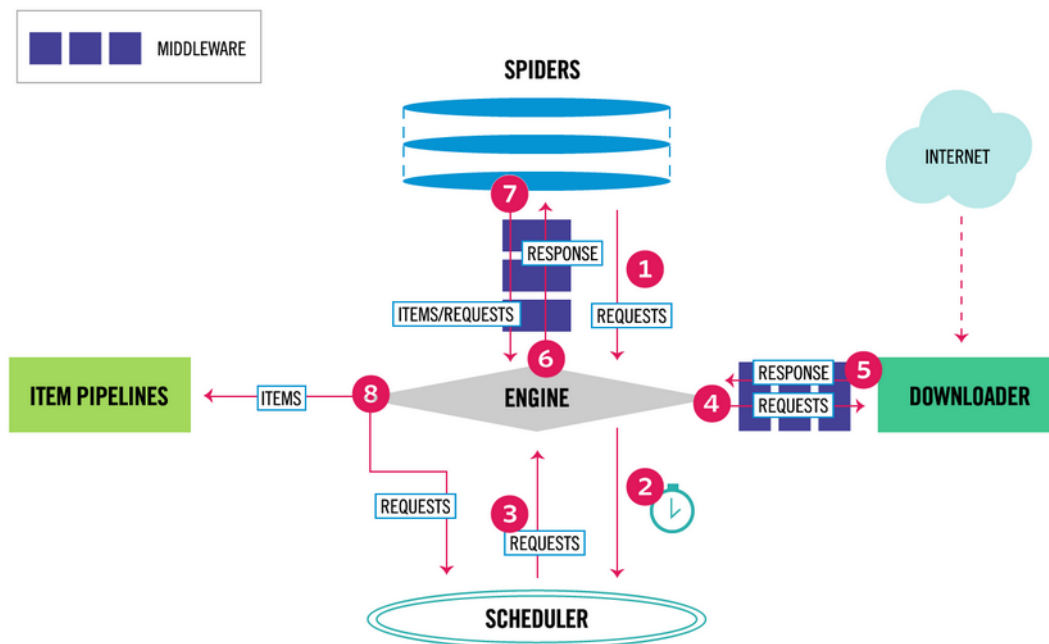
μέχρι την αποθήκευσή τους σε ετερογενείς δομές δεδομένων (βάσεις δεδομένων, αρχεία).

• Downloader Middlewares

Τα *Downloader Middlewares* είναι τμήματα του μηχανισμού, τα οποία συναντούνται μεταξύ του *Downloader* και του *Engine* και μεταχειρίζονται τα αιτήματα που αυτά τα δύο ανταλλάσσουν. Οι περιπτώσεις κατά τις οποίες τα *Downloader Middlewares* μπορούν να φανούν χρήσιμα είναι: (i) η επεξεργασία ενός αιτήματος πριν αυτό σταλεί στο *Downloader*, (ii) η τροποποίηση μιας ληφθείσας απόκρισης πριν αυτή μεταβιβαστεί στα *Spiders*, (iii) η αποστολή ενός νέου αιτήματος αντί για την μεταφορά μιας απόκρισης στα *Spiders*, και (iv) η απόρριψη συγκεκριμένων αιτημάτων.

• Spider Middlewares

Τα *Spider Middlewares* είναι τμήματα του μηχανισμού τα οποία βρίσκονται μεταξύ του *Engine* και των *Spiders* και είναι σε θέση να επεξεργάζονται τις εισόδους (ληφθείσες αποκρίσεις) και τις εξόδους (συλλεγμένα δεδομένα από ιστοσελίδες) των *Spiders*. Οι περιπτώσεις κατά τις οποίες τα *Downloader Middlewares* μπορούν να φανούν χρήσιμα είναι: (i) η τροποποίηση των αιτημάτων ή των δεδομένων που εξάγονται από τα *Spiders* και (ii) η αντιμετώπιση σφαλμάτων και εξαιρέσεων που προκύπτουν κατά τη λειτουργία των *Spiders*.



Σχήμα 3.1: Scrapy αρχιτεκτονική και ροή δεδομένων μεταξύ των δομικών στοιχείων του

Σύμφωνα με το παραπάνω διάγραμμα, μπορεί να παρατηρηθεί η εξής ροή δεδομένων κατά την εξαγωγή πληροφοριών με τη χρήση του Scrapy Framework:

1. Το *Engine* λαμβάνει από τα *Spiders* τα αρχικά αιτήματα που προορίζονται για αποστολή.
2. Το *Engine* προγραμματίζει τη χρονική σειρά των αιτημάτων στο *Scheduler* και ζητά τα αιτήματα που έχουν προτεραιότητα να αποσταλούν.

3. Το *Scheduler* επιστρέφει τα προς αποστολή αιτήματα στο *Engine*.
4. Το *Engine* στέλνει τα αιτήματα στο *Downloader*, περνώντας τα διαμέσου του *Downloader Middlewares*.
5. Μόλις ολοκληρωθεί η ανάκτηση της επιθυμητής ιστοσελίδας, το *Downloader* παράγει μία απόκριση η οποία περιέχει τα δεδομένα της ιστοσελίδας και την αποστέλλει στο *Engine*, διαμέσου του *Downloader Middlewares*.
6. Το *Engine* λαμβάνει την απόκριση απο το *Downloader* και την αποστέλλει στα *Spiders* για επεξεργασία, διαμέσου των *Spider Middlewares*.
7. Τα *Spiders* επεξεργάζονται την απόκριση και επιστρέφουν τις εξαχθείσες πληροφορίες, καθώς και τα νέα αιτήματα προς αποστολή, στο *Engine*, διαμέσου του *Spider Middlewares*.
8. Το *Engine* προσκομίζει τα επεξεργασμένα δεδομένα στο *Item Pipeline* και, ακολούθως, τα νέα αιτήματα στο *Scheduler*, ζητώντας του να τα αποστείλει.
9. Η διαδικασία επαναλαμβάνεται από το Βήμα 1, έως ότου δεν υπάρχουν πλέον αιτήματα προς αποστολή από την μεριά του *Scheduler*.

3.1.3 Βασική δομή ενός Scrapy πρότζεκτ



Στη συγκεκριμένη υποενότητα θα περιγραφεί ο τρόπος με τον οποίο μπορεί να δημιουργηθεί μία διαδικασία (αλλιώς πρότζεκτ) συλλογής δεδομένων στο Scrapy, καθώς και η βασική δομή των *spiders*, μέσω των οποίου θα λάβει μέρος η διαδικασία [8], [17], [29].

Δεδομένου ότι ο μηχανισμός του Scrapy έχει εγκατασταθεί επιτυχώς στο υποψήφιο σύστημα, ένα πρότζεκτ είναι δυνατό να δημιουργηθεί επιλέγοντας, αρχικά, μέσω της γραμμής εντολών (CLI) τον φάκελο στον οποίο θα στεγαστεί και, στη συνέχεια, πληκτρολογώντας την εντολή:

```
scrapy startproject first_scrapy
```

όπου το *first_scrapy* είναι το επιλεγμένο όνομα για το συγκεκριμένο πρότζεκτ.

Η εκτέλεση της παραπάνω εντολής θα οδηγήσει στην αυτόματη δημιουργία ενός φακέλου, ο οποίος θα περιέχει όλα τα απαραίτητα για την ομαλή λειτουργία του *spider* αρχεία. Η δομή του νεοσύστατου φακέλου θα είναι η εξής:

 first_scrapy	11/5/2021 6:08 μμ	Φάκελος αρχείων	
 scrapy.cfg	11/5/2021 6:08 μμ	Αρχείο CFG	1 KB

Σχήμα 3.2: Η εξωτερική δομή ενός Scrapy πρότζεκτ

Επίσης, ο φάκελος *first_scrapy* παρουσιάζει τη μορφή που φαίνεται στο Σχήμα 3.3:

Πιο συγκεκριμένα, ένα Scrapy πρότζεκτ απαρτίζεται από τα παρακάτω αρχεία/φακέλους:

- **scrapy.cfg**: αρχείο ρύθμισης παραμέτρων
- **__init__.py**: αρχείο αρχικοποίησης

spiders	19/10/2020 7:48 μμ	Φάκελος αρχείων	
__init__.py	5/8/2020 7:16 μμ	Python File	0 KB
items.py	11/5/2021 5:36 μμ	Python File	1 KB
middlewares.py	11/5/2021 5:36 μμ	Python File	4 KB
pipelines.py	11/5/2021 5:36 μμ	Python File	1 KB
settings.py	11/5/2021 5:36 μμ	Python File	3 KB

Σχήμα 3.3: Η εσωτερική δομή ενός Scrapy πρότζεκτ

- **items.py**: αρχείο που περιλαμβάνει καθορισμένα από τον χρήστη αντικείμενα, τα οποία περιγράφουν τη δομή των προς συλλογή δεδομένων
- **middlewares.py**: αρχείο στο οποίο ορίζονται τα *Middlewares* του πρότζεκτ (βλ. Υποενότητα 3.1.2)
- **pipelines.py**: αρχείο στο οποίο ορίζονται τα *Pipelines* του πρότζεκτ (βλ. Υποενότητα 3.1.2)
- **settings.py**: αρχείο ρυθμίσεων
- **spiders**: φάκελος στον οποίο στεγάζονται τα αρχεία των *Spiders* του εκάστοτε χρήστη

Κάθε αρχείο που αποθηκεύεται στον φάκελο *spiders* αντιπροσωπεύει έναν αυτοσχέδιο μηχανισμό συγκομιδής δεδομένων ή αλλιώς ένα *spider*. Όλα τα *spiders* διαθέτουν μια υποδειγματική δομή, την οποία ο χρήστης καλείται να επεκτείνει με βάση τις απαιτήσεις του. Τα βασικά στοιχεία, των οποίων η ύπαρξη είναι απαραίτητη για την ομαλή λειτουργία του μηχανισμού, είναι τα εξής:

- **name**: το αναγνωριστικό όνομα του *spider*
- **allowed_domains**: λίστα που περιλαμβάνει του δικτυακούς τόπους, τους οποίους θα επισκεφθεί ο μηχανισμός κατά τη συλλογή δεδομένων
- **start_urls**: λίστα διευθύνσεων URL, από τις οποίες θα ξεκινήσει αυτόματα η αναζήτηση δεδομένων
- **start_requests()**: μέθοδος που καλείται κατά την εκκίνηση της διαδικασίας και αποστέλλει αιτήματα προς τις διευθύνσεις URL από τις οποίες θα ξεκινήσει η αναζήτηση των δεδομένων· η λειτουργία της μεθόδου *start_requests()* είναι ταυτοτική με εκείνη της λίστας *start_urls*, επομένως η επιλογή χρήσης μίας από τις δύο ανατίθεται στον χρήστη
- **parse()**: μέθοδος που θα κληθεί να χειριστεί την απόκριση, η οποία έχει ληφθεί για καθένα από τα αιτήματα που έχουν υποβληθεί, και να εξάγει τις επιθυμητές πληροφορίες από αυτή

3.2 YouTube Data API

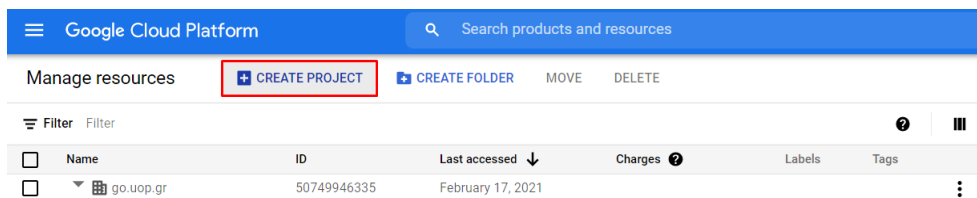
Όπως προαναφέρθηκε στο Κεφάλαιο 2, στην Υποενότητα 2.1.3, αρκετά συχνά, η συλλογή δεδομένων από διαδικτυακές πηγές με τη χρήση μηχανισμών συγκομιδής δεν είναι δυνατή, είτε επειδή εφαρμόζονται μέτρα για τη διαφύλαξη των πληροφοριών ή επειδή τα δεδομένα είναι

δομημένα με αρκετά περίπλοκο τρόπο, με αποτέλεσμα να καθίσταται δύσκολη η συγκέντρωση τους. Το τελευταίο ενδεχόμενο παρουσιάστηκε κατά τη συλλογή δεδομένων από το YouTube, γεγονός που συνέβαλε στην απόφαση για τη χρήση του YouTube Data API [55].

Το YouTube Data API αποτελεί ένα από τα επίσημα APIs του YouTube και επιτρέπει την ενσωμάτωση στην εφαρμογή του χρήστη μιας πληθώρας λειτουργιών που παρέχονται από την πλατφόρμα. Με τη χρήση του συγκεκριμένου εργαλείου, αποκτάται πρόσβαση σε ένα ευρύ σύνολο μεθόδων και προτύπων επικοινωνίας, μέσω των οποίων μπορούν να ανακτηθούν όλες οι διαθέσιμες πληροφορίες που αφορούν ένα ενδεχόμενο κανάλι ή βίντεο. Ένας χρήστης που επιθυμεί να αξιοποιήσει τις λειτουργίες του YouTube Data API, θα πρέπει να προβεί στην έκδοση ενός ψηφιακού κλειδιού, το οποίο θα χρησιμοποιείται ως αναγνωριστικό της εφαρμογής του και, παράλληλα, θα καθορίζει το ανώτατο όριο μονάδων που επιτρέπεται να εξαργυρώνει καθημερινά. Οι μονάδες αυτές αντιπροσωπεύουν την βαρύτητα του εκάστοτε απεσταλμένου αιτήματος. Κατά το χρονικό διάστημα που πραγματοποιήθηκε η παρούσα πτυχιακή εργασία, το προκαθορισμένο ανώτατο όριο μονάδων ανερχόταν στις 10.000/ημέρα.

Οι διαδοχικές ενέργειες που απαιτούνται για την απόκτηση ενός ψηφιακού κλειδιού αναφέρονται αριθμημένες στην λίστα που ακολουθεί:

1. Εφόσον έχει προηγηθεί σύνδεση του χρήστη στον Google λογαριασμό του, για την απόκτηση ενός κλειδιού API του YouTube πρέπει, ως πρώτο βήμα, να πλοηγηθεί στην διεύθυνση <https://console.developers.google.com/project>.
2. Στη συνέχεια της διαδικασίας, θα πρέπει να πατήσει το κουμπί που φέρει την ονομασία *CREATE PROJECT* στο πάνω μέρος της σελίδα, όπως φαίνεται στο Σχήμα 3.4, με σκοπό να δημιουργηθεί ένα νέο πρότζεκτ.



Σχήμα 3.4: Δημιουργία νέου πρότζεκτ

3. Ύστερα, όπως δείχνει και το Σχήμα 3.5, ο χρήστης θα μεταφερθεί σε σχετική καρτέλα, στην οποία θα κληθεί να δώσει όνομα στο συγκεκριμένο πρότζεκτ και, στη συνέχεια, να πατήσει το μπλε κουμπί που βρίσκεται στο κάτω και αριστερά μέρος της ιστοσελίδας για την δημιουργία του πρότζεκτ.

Google Cloud Platform

New Project

You have 11 projects remaining in your quota. Request an increase or delete projects. [Learn more](#)

[MANAGE QUOTAS](#)

Project name *
My Project 59706

Project ID: nomadic-proton-314615. It cannot be changed later. [EDIT](#)

Organization *
go.uop.gr

Select an organization to attach it to a project. This selection can't be changed later.

Location *
go.uop.gr [BROWSE](#)

Parent organization or folder

[CREATE](#) CANCEL

Σχήμα 3.5: Υποβολή ονόματος του πρότζεκτ

4. Μόλις δημιουργηθεί το πρότζεκτ, ο χρήστης θα πρέπει, επιλέγοντας το μενού που εμφανίζεται πάνω και αριστερά της σελίδας (τρεις οριζόντιες γραμμές), να επιλέξει την καρτέλα *Home*. Έχοντας λοιπόν, μεταφερθεί στην αρχική σελίδα, καλείται να επιλέξει το πρότζεκτ για το οποίο θέλει να εκδώσει ψηφιακό κλειδί. Αυτό μπορεί να επιτευχθεί πατώντας, αρχικά, το αναδυόμενο μενού που βρίσκεται στην επικεφαλίδα της ιστοσελίδας και φέρει το όνομα *Select a project* ή το όνομα ενός από τα πρότζεκτ του και, στη συνέχεια, επιλέγοντας το όνομα του επιθυμητού πρότζεκτ από την εμφανιζόμενη λίστα (βλ. Σχήμα 3.6).

Select from GO.UOPGR NEW PROJECT

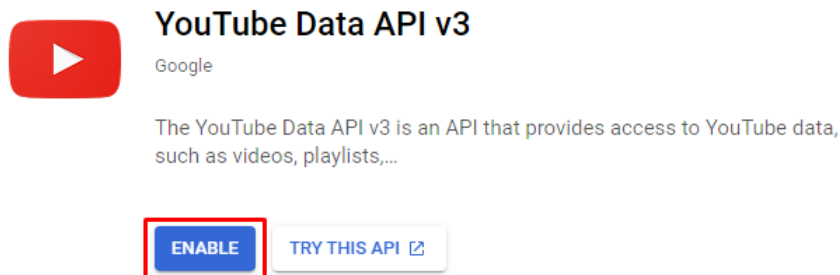
Search projects and folders

RECENT STARRED ALL

Name	ID
✓ ☆ My Project 59706	nomadic-proton-314615
go.uop.gr	50749946335

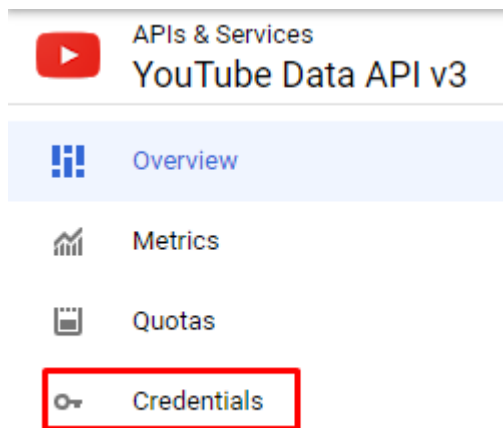
Σχήμα 3.6: Αναδυόμενη λίστα με όλα τα διαθέσιμα πρότζεκτ του λογαριασμού

5. Ως επόμενη κίνηση ορίζεται η ενεργοποίηση του API που σχετίζεται με το συγκεκριμένο πρότζεκτ. Έτσι, ο χρήστης καλείται να επιλέξει ξανά το μενού που εμφανίζεται πάνω και αριστερά της σελίδας (τρεις οριζόντιες γραμμές), να οδηγήσει τον κέρσορα του στην επιλογή *APIs & Services* και, στη συνέχεια, να επισκεφθεί την υποκατηγορία *Library*, με σκοπό να πλοηγηθεί σε μία ψηφιακή βιβλιοθήκη, στην οποία εμπεριέχεται και το YouTube Data API. Ακολούθως, ο χρήστης θα πρέπει να αναζητήσει το προαναφερθέν API και να πατήσει πάνω του με την βοήθεια του κέρσορα. Μόλις μεταφερθεί στην αντίστοιχη σελίδα, θα πρέπει, τελικά, να επιλέξει το μπλε κουμπί με την ονομασία *Enable*, όπως παρουσιάζεται στο Σχήμα 3.7, ώστε να ενεργοποιηθεί το API.



Σχήμα 3.7: Ενεργοποίηση του εργαλείου YouTube Data API

6. Στο σημείο αυτό, εφόσον έχει ενεργοποιηθεί το API, ο χρήστης θα μεταβεί αυτομάτως στον πίνακα ελέγχου του YouTube Data API για το συγκεκριμένο πρότζεκτ, από όπου θα πρέπει να ανακατευθυνθεί στην καρτέλα *Credentials*, επιλέγοντας την από το κατακόρυφο μενού στα αριστερά της σελίδας (βλ. Σχήμα 3.8).



Σχήμα 3.8: Επιλογή της καρτέλας Credentials

7. Μόλις ο χρήστης φτάσει στην καρτέλα *Credentials*, τότε στο επάνω μέρος της σελίδας θα εμφανίζεται ένα κουμπί με την επιγραφή *CREATE CREDENTIALS*. Όταν επιλεγθεί το κουμπί αυτό, θα αναπτυχθεί μια λίστα με επιλογές, από τις οποίες ο χρήστης θα πρέπει να επιλέξει εκείνη που αναγράφεται ως *API key*. Ως αποτέλεσμα της ενέργειας αυτής, θα εμφανιστεί ένα αναδυόμενο παράθυρο με το νεοσύστατο ψηφιακό κλειδί API, όπως φαίνεται στο σχήμα 3.9, και ο χρήστης θα έχει την δυνατότητα να το αντιγράψει και να το χρησιμοποιήσει σε αιτήματα που, ενδεχομένως, θα χρειαστεί να συντάξει μελλοντικά, ώστε να αποσπάσει βασικές πληροφορίες για ένα κανάλι του YouTube.

API key created

Use this key in your application by passing it with the `key=API_KEY` parameter.

Your API key

⚠ Restrict your key to prevent unauthorized use in production.

[CLOSE](#) [RESTRICT KEY](#)

Σχήμα 3.9: Δημιουργία ενός κλειδιού API

3.3 MongoDB

Με σκοπό την αποθήκευση των δεδομένων που συλλέχθηκαν με τη χρήση των παραπάνω εργαλείων, χρησιμοποιήθηκε η MongoDB βάση δεδομένων [25]. Η MongoDB αποτελεί μία NoSQL (βλ. Υποενότητα 2.2.1) βάση δεδομένων εγγράφων (document database) και προσφέρει μία μεγάλη ποικιλία υπηρεσιών, που εξυπηρετούν τόσο την αποθήκευση και τη διαχείριση των δεδομένων όσο και την αναπαράστασή τους.

Ο όρος “βάση δεδομένων εγγράφων” αντικατοπτρίζει την ιδιότητα της βάσης να αποθηκεύει τα δεδομένα σε μορφή που μοιάζει με ένα JSON έγγραφο. Όσον αφορά στη MongoDB, κάθε εγγραφή αντιπροσωπεύει ένα έγγραφο, δηλαδή μια δομή δεδομένων που απαρτίζεται από ζεύγη κλειδιού-τιμής. Τα έγγραφα της MongoDB είναι πανομοιότυπα με τα αντικείμενα JSON. Οι τιμές των πεδίων μπορεί να περιλαμβάνουν άλλα έγγραφα, πίνακες ή ακόμα και πίνακες εγγράφων. Οι εγγραφές δεδομένων της MongoDB, που αποθηκεύονται ως έγγραφα συγκροτούν ένα ευρύτερο σύνολο δεδομένων, το οποίο ονομάζεται συλλογή (Collection). Μία βάση δεδομένων συνίσταται από μία ή περισσότερες συλλογές εγγράφων. Στην ακόλουθη εικόνα παρουσιάζεται η τυπική μορφή ενός εγγράφου MongoDB [25].

```

{
  name: "sue",
  age: 26,
  status: "A",
  groups: [ "news", "sports" ]
}

```

← field: value
 ← field: value
 ← field: value
 ← field: value

Σχήμα 3.10: Μορφή ενός MongoDB εγγράφου

Σε συνδυασμό με τα προτερήματα που συνοδεύουν μία NoSQL βάσης δεδομένων (βλ. Υποενότητα 2.2.1), ορισμένα από τα χαρακτηριστικά που ενισχύουν την καταλληλότητα της MongoDB για ένα ευρύ φάσμα εφαρμογών, συμπεριλαμβανομένης και της δικής μας, είναι τα εξής [28]:

- **Υψηλή ταχύτητα ενημέρωσης των δεδομένων**, λόγω του αυτόματου ευρετηριασμού τους (Indexing) και της ύπαρξης εσωτερικής μνήμης. Χωρίς την ύπαρξη ευρετηρίων, η MongoDB θα έπρεπε να εκτελέσει σάρωση συλλογής, δηλαδή σάρωση για κάθε αποθηκευμένο έγγραφο σε μια συλλογή, για να επιλέξει τα έγγραφα αυτά που ταιριάζουν με τη δήλωση του συντασσόμενου από τον χρήστη ερωτήματος.
- **Πλούσια γλώσσα ερωτημάτων**, χάρη στη δυνατότητα εκτέλεσης τυπικών ή συγκεντρωτικών ερωτημάτων (Aggregations). Ένα παράδειγμα τυπικού ερωτήματος είναι το `db.collection.find({ age: { $gt: 20 } })`, κατά το οποίο ζητείται από την συλλογή να επιστραφούν τα έγγραφα των οποίων το πεδίο “age” έχει τιμή μεγαλύτερη (`$gt`: greater than) από 20. Από την άλλη, τα συγκεντρωτικά ερωτήματα επεξεργάζονται τα δεδομένα των εγγραφών, εκτελώντας πολλαπλά σειριακά τυπικά ερωτήματα και επιστρέφουν τα αντίστοιχα αποτελέσματα. Η MongoDB παρέχει τρεις τρόπους για την εκτέλεση τέτοιου είδους ερωτημάτων: (i) τη λειτουργία *Aggregation Pipeline* (βλ. Σχήμα 3.11), (ii) την διαδικασία *Map-Reduce* και (iii) τις λιγότερο σύνθετες μεθόδους συγκέντρωσης απλού σκοπού, όπως οι μέθοδοι `count()` και `distinct()`.

```
db.orders.aggregate([
  { $match: { status: "A" } },
  { $group: { _id: "$cust_id", total: { $sum: "$amount" } } }
])
```

Σχήμα 3.11: Μορφή ενός ερωτήματος σε Aggregation Pipeline

- **Διασφάλιση της οριζόντιας επεκτασιμότητας**, με τον διαμοιρασμό του φόρτου εκτέλεσης σε πολλαπλούς μηχανισμούς (Sharding). Τα συστήματα βάσεων δεδομένων με μεγάλη σύνολα δεδομένων ή εφαρμογές υψηλής απόδοσης μπορούν να προκαλέσουν την εξάντληση της χωρητικότητας ή της επεξεργαστικής ισχύος ενός διακομιστή. Για αυτόν τον λόγο εφαρμόζονται τεχνικές Sharding, κατά τις οποίες τα δεδομένα διασπώνται σε επιμέρους τμήματα (Shards), δέχονται επεξεργασία από πολυάριθμες επεξεργαστικές μονάδες και, τέλος, τα αποτελέσματα συνδυάζονται, συνθέτοντας έτσι την τελική έξοδο της διαδικασίας.
- **Υψηλή διαθεσιμότητα των δεδομένων**, μέσω της δημιουργίας αντιγράφων σε πολλαπλούς διακομιστές (Replication). Η διαδικασία αυτή διασφαλίζει την προστασία των δεδομένων από περιπτώσεις διακοπής λειτουργίας ενός διακομιστή, διακοπής υπηρεσίας παροχής διαδικτύου ή βλάβης του υλικού εξοπλισμού.
- **Παροχή πρόσθετων υπηρεσιών**, οι οποίες διευκολύνουν την αποθήκευση, την διαχείριση και την οπτικοποίηση των δεδομένων. Χαρακτηριστικά αναφέρονται οι MongoDB Compass, MongoDB Atlas και MongoDB Charts, κάθε μία από τις οποίες παρουσιάζεται παρακάτω με περισσότερη λεπτομέρεια.

3.3.1 MongoDB Compass

Το MongoDB Compass [27] είναι το γραφικό περιβάλλον διεπαφής χρήστη (GUI) που παρέχεται από την MongoDB. Το συγκεκριμένο εργαλείο επιτρέπει στον χρήστη να επεξεργάζεται, να

αναλύει και να εκτελεί στα δεδομένα της βάσης του, χωρίς να κρίνεται απαραίτητη η γνώση της σύνταξης MongoDB ερωτημάτων. Το MongoDB Compass ενσωματώνει λειτουργίες μέσω των οποίων καθίστανται δυνατά η βελτιστοποίηση της απόδοσης των ερωτημάτων, η διαχείριση των ευρετηρίων, καθώς και ο ορισμός ενός σχήματος (schema) για τα υπό αποθήκευση έγγραφα.

3.3.2 MongoDB Atlas

Κατά τη διεκπεραίωση της προκείμενης έρευνας, η ανάγκη για συνεργασία των συντελεστών και για απόκτηση πρόσβασης σε μία κοινή βάση δεδομένων από απομακρυσμένα σημεία οδήγησε στην αξιοποίηση του MongoDB Atlas [24]. Το MongoDB Atlas είναι μία διαδικτυακή υπηρεσία της MongoDB, η οποία προσφέρει όλα τα απαραίτητα εφόδια για την γρήγορη και πλήρη ανάπτυξη ενός Cluster σε περιβάλλον νέφους (Cloud). Με τη χρήση της παρούσας υπηρεσίας επιτυγχάνεται με ευκολία η πρόσβαση στη βάση δεδομένων από πολλαπλούς παράλληλους χρήστες, καθώς δίνεται η δυνατότητα καθορισμού των χρηστών που μπορούν να επισκεφτούν τη βάση και των δικαιωμάτων που τους αναλογούν.

3.3.3 MongoDB Charts

Το MongoDB Charts [26] συνιστά ένα εργαλείο μέσω του οποίου επιτυγχάνεται η δημιουργία γραφικών παραστάσεων, με βάση τα αποθηκευμένα δεδομένα μιας MongoDB βάσης δεδομένων. Η οπτικοποίηση των δεδομένων είναι ζωτικής σημασίας για τον εντοπισμό συσχετίσεων που μπορεί να υπάρχουν μεταξύ των στοιχείων μιας έρευνας, γεγονός που, σε συνδυασμό με τους πολυάριθμους τύπους γραφημάτων που διατίθενται και τη δυνατότητα δημιουργίας σύνθετων μεταβλητών από επιμέρους πεδία της βάσης, καθιστά το MongoDB Charts μια ασφαλή δίοδο προς τη μετουσίωση των συγκεντρωμένων πληροφοριών μιας μελέτης σε πολύτιμα συμπεράσματα.

3.4 Google Forms

Το Google Forms [21] είναι μια διαδικτυακή εφαρμογή που χρησιμοποιείται για τη δημιουργία φορμών, με σκοπό την συλλογή δεδομένων. Κάθε υποψήφιος ερευνητής έχει την δυνατότητα να αξιοποιήσει την εφαρμογή αυτή για να διεξάγει μία έρευνα, δημιουργώντας και διαμοιράζοντας ερωτηματολόγια στο ευρύτερο κοινό. Αξίζει να σημειωθεί πως το Google Forms επιτρέπει σε πολλαπλούς χρήστες να διαχειρίζονται από κοινού και σε πραγματικό χρόνο το ερευνητικό υλικό. Τα δεδομένα που συλλέγονται με τη χρήση των φορμών, μπορούν να προβληθούν είτε σε μορφή υπολογιστικού φύλλου, είτε μέσω ενός αρχείου CSV (Comma-Separated Values) και, τελικά, να μεταφραστούν σε διαγράμματα, χάρη στην αυτοματοποιημένη απεικόνιση των αποτελεσμάτων που προσφέρει η πλατφόρμα.

Κεφάλαιο 4

Συγκέντρωση και αποθήκευση των δεδομένων της έρευνας

Εφόσον παρουσιάστηκαν οι σημαντικότερες πτυχές της προκειμένης έρευνας και παρατέθηκαν οι απαραίτητες, για την κατανόηση της, πληροφορίες, καθίσταται πλέον εφικτή η εμβάθυνση στα διαδοχικά βήματα που ακολουθήθηκαν, τηρώντας το τρίπτυχο συγκέντρωση, αποθήκευση και μελέτη των δεδομένων, με σκοπό την εξαγωγή εμπειριστατωμένων και αξιόπιστων συμπερασμάτων.

Στο κεφάλαιο αυτό θα γίνει αναλυτική αναφορά στη διαδικασία συγκέντρωσης και αποθήκευσης δεδομένων από τις πλατφόρμες Instagram και YouTube με τη χρήση των εργαλείων που αναφέρθηκαν εκτενώς στο Κεφάλαιο 3.

4.1 Συγκομιδή στοιχείων του Instagram

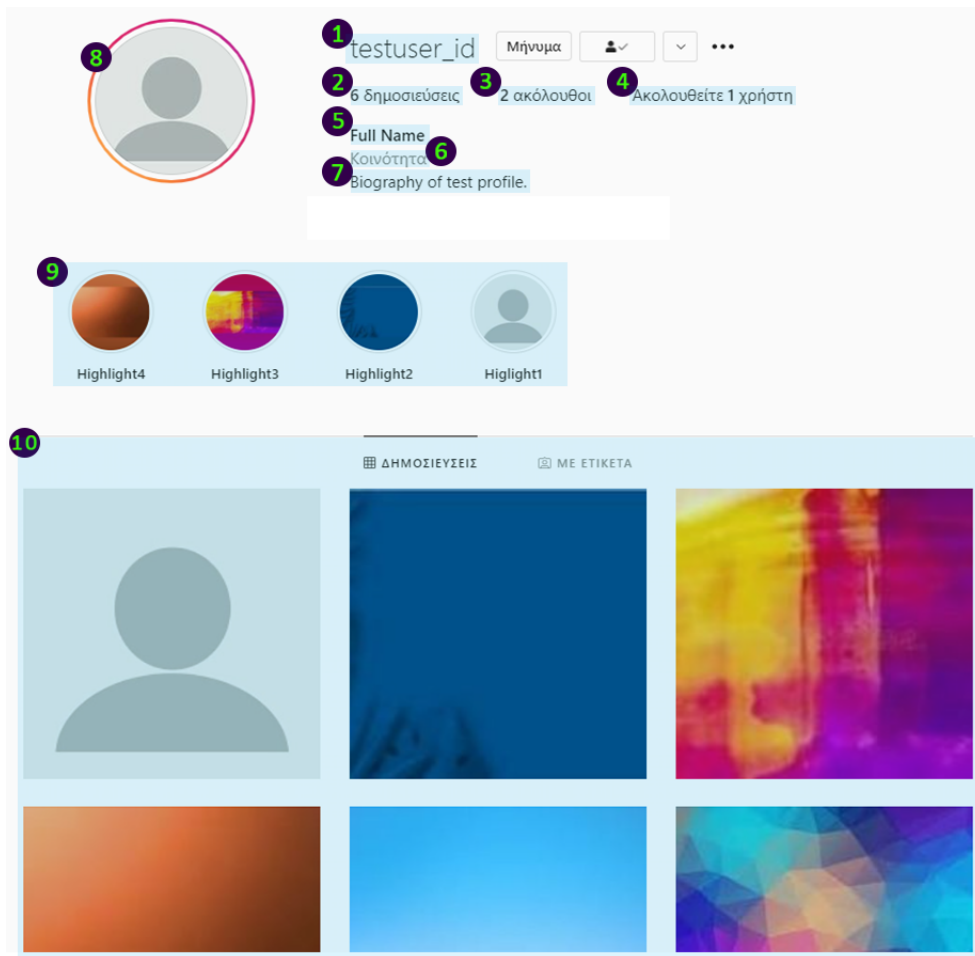
Προκειμένου να περιγραφούν τα πρωτογενή (τα οποία συλλέχθηκαν και αποθηκεύτηκαν αυτούσια, χωρίς να υποστούν φιλτράρισμα ή τροποποίηση) και δευτερογενή (τα οποία προέκυψαν από την επεξεργασία πρωτογενών στοιχείων) δεδομένα που εξήχθησαν από το Instagram, με σκοπό τη δημιουργία στατιστικών στοιχείων, κρίνεται απαραίτητη η ανάλυση της βασικής δομής ενός Instagram λογαριασμού και, κατ' επέκταση, μιας Instagram δημοσίευσης, καθώς και η επισήμανση των προκλήσεων και των δυσκολιών που έπρεπε να αντιμετωπιστούν, ώστε να στεφθεί με επιτυχία η συγκομιδή του ερευνητικού υλικού από αυτή την πλατφόρμα.

4.1.1 Δομή του Instagram και δεδομένα

Η τυπική διαστρωμάτωση ενός λογαριασμού Instagram παρουσιάζεται στο Σχήμα 4.1 και είναι η ακόλουθη:

1. **Κωδικό όνομα χρήστη (User name ID):** η κωδική ονομασία του χρήστη, η οποία χρησιμοποιείται ως μοναδικό αναγνωριστικό του
2. **Αριθμός δημοσιεύσεων (Posts):** ο αριθμός των συνολικών δημοσιεύσεων του χρήστη
3. **Αριθμός ακολούθων (Followers):** ο αριθμός των χρηστών που έχουν επιλέξει να ενημερώνονται για τις δημοσιεύσεις του χρήστη, ακολουθώντας τον

4. **Αριθμός ακολουθούμενων (Following):** ο αριθμός των χρηστών, για των οποίων τις δημοσιεύσεις ο χρήστης έχει επιλέξει να ενημερώνεται, ακολουθώντας τους
5. **Πλήρες όνομα χρήστη (Full name):** το κείμενο που έχει επιλέξει ο χρήστης να αναγράφεται ως το πλήρες όνομα του
6. **Κατηγορία λογαριασμού (Account category):** πεδίο που εμφανίζεται αν ο χρήστης έχει κατατάξει τον λογαριασμό του σε κάποια επαγγελματική κατηγορία (π.χ καλλιτέχνης, δημιουργός βίντεο ή ένδυση)
7. **Σύντομο βιογραφικό χρήστη (Biography):** μικρής έκτασης κείμενο, στο οποίο ο χρήστης αναφέρει λεπτομέρειες για το προφίλ του
8. **Ιστορία Instagram (Instagram story):** στιγμιότυπο το οποίο έχει επιλέξει ο χρήστης να εμφανίζεται στον λογαριασμό του για 24 ώρες
9. **Αποθηκευμένα στιγμιότυπα (Highlights reel):** το σύνολο των Instagram ιστοριών (Instagram stories) τις οποίες ο χρήστης έχει επιλέξει να παραμένουν διαρκώς ορατές προς το κοινό
10. **Σύνολο δημοσιεύσεων χρήστη (Profile feed):** η λίστα των δημοσιεύσεων του χρήστη



Σχήμα 4.1: Διαστρωμάτωση ενός λογαριασμού Instagram

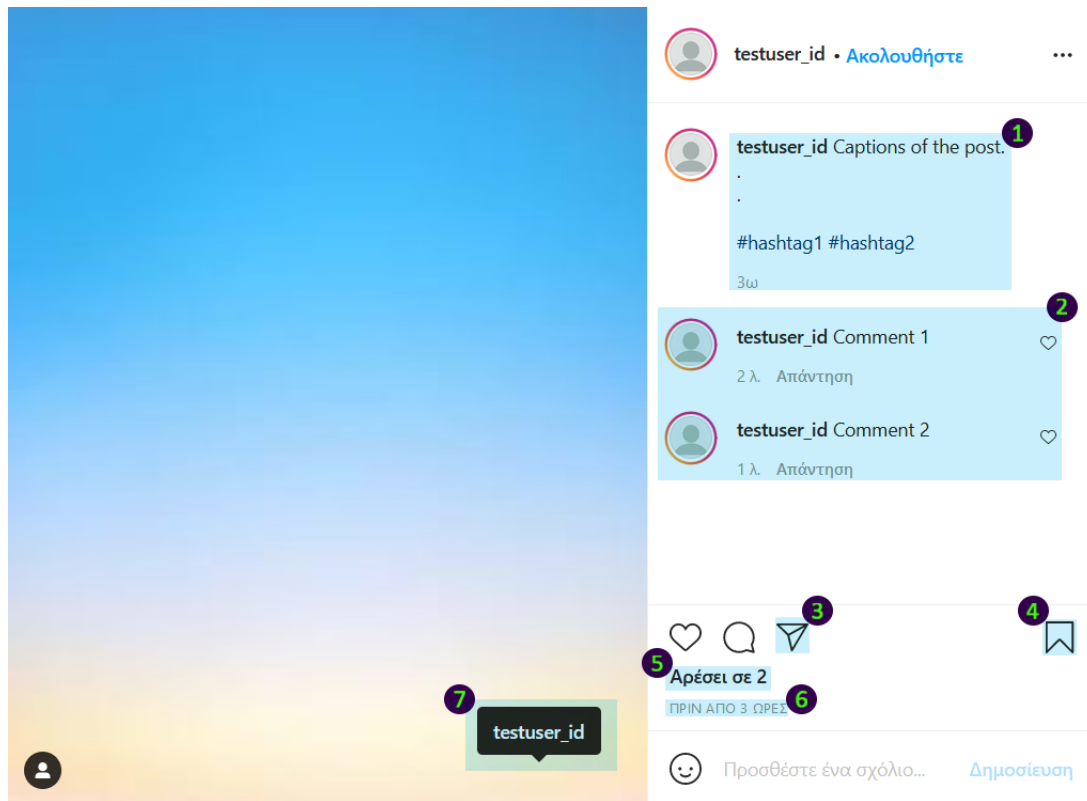
Κάθε χρήστης έχει τη δυνατότητα να δημοσιεύει διαφορετικού τύπου περιεχόμενο, το οποίο μπορεί να παραμένει ορατό για περιορισμένο χρονικό διάστημα, δηλαδή οι ιστορίες Instagram που προαναφέρθηκαν, ή ωστόσο αποφασιστεί η διαγραφή του. Στην τελευταία κατηγορία αναρτήσεων ανήκουν τα εξής:

- **Φωτογραφία:** ανάρτηση μιας απλής φωτογραφίας
- **Βίντεο μικρής διάρκειας:** ανάρτηση ενός βίντεο με διάρκεια μικρότερη του ενός λεπτού
- **Βίντεο μεγάλης διάρκειας:** ανάρτηση ενός βίντεο με διάρκεια μεγαλύτερη του ενός λεπτού· τα βίντεο αυτά κατατάσσονται στην θεματική κατηγορία IGTV
- **Παρουσίαση πολλαπλού περιεχομένου (Slideshow):** ανάρτηση μιας συλλογής φωτογραφιών και βίντεο

Όλες οι μορφές δημοσιευμένου περιεχομένου, εκτός των των ιστοριών Instagram, μοιράζονται ένα κοινό πρότυπο αναπαράστασης με τα βασικά του δομικά στοιχεία, τα οποία φαίνονται στο Σχήμα 4.2, να είναι τα ακόλουθα:

1. **Περιγραφή δημοσίευσης (Captions):** σύντομη περιγραφή του περιεχομένου της δημοσίευσης· συχνό φαινόμενο αποτελεί η επισήμανση χρηστών, υπό τη μορφή “@user_name”, ή δημοφιλών hashtags, γράφοντας #hashtag1, στο συγκεκριμένο πεδίο
2. **Σχόλια (Comments):** το σύνολο των σχολιασμών που απέσπασε η δημοσίευση από τον ίδιο το δημιουργό της ή άλλους χρήστες
3. **Κοινοποίηση (Share):** αντίδραση που επιτρέπει στους χρήστες να προωθήσουν τη δημοσίευση σε άλλους χρήστες της πλατφόρμας
4. **Αποθήκευση (Save):** αντίδραση που επιτρέπει στους χρήστες να αποθηκεύσουν τη δημοσίευση στη λίστα των αγαπημένων τους
5. **Αριθμός αντιδράσεων “Μου αρέσει” (Likes)/Προβολές βίντεο (Video views):** το σύνολο των χρηστών που αντέδρασαν “Μου αρέσει” στη δημοσίευση· στην περίπτωση που η δημοσίευση είναι βίντεο, τότε στο πεδίο αυτό αναγράφεται, επιπλέον, ο αριθμός των χρηστών που το παρακολούθησαν
6. **Ημερομηνία κοινοποίησης (Upload date):** ημερομηνία κατά την οποία κοινοποιήθηκε η δημοσίευση
7. **Χρήστες που έχουν επισυναφθεί (Mentioned users):** το σύνολο των λογαριασμών των οποίων το όνομα παρουσιάζεται στο εσωτερικό του περιεχομένου της δημοσίευσης

Βασικό άξονα στην επιλογή των κατάλληλων, για την έρευνα, δεδομένων αποτέλεσε η μελέτη της απήχησης των αναρτήσεων από λογαριασμούς ποικίλου βεληνεκούς, όσον αφορά το μέγεθος του κοινού που τους ακολουθεί. Για να εξυπηρετηθεί ο σκοπός αυτός, κρίθηκε πως η ανταπόκριση του κοινού απέναντι σε μία δημοσίευση μπορεί να αντικατοπτρισθεί στον αριθμό των αντιδράσεων που απέσπασε η τελευταία. Επομένως, η συγκέντρωση του πλήθους των αντιδράσεων “Μου αρέσει”, των συνολικών σχολίων και των προβολών, σε περιπτώσεις βίντεο, μιας δημοσίευσης συνέβαλε στον σχηματισμό μιας πρωταρχικής εικόνας για την απόδοση της ανάρτησης. Τα παραπάνω αριθμητικά δεδομένα, σε συνδυασμό με το είδος της δημοσίευσης, το



Σχήμα 4.2: Βασική δομή μίας δημοσίευσης του Instagram

σύνολο των επισυνημμένων λογαριασμών και των αναγραφόμενων hashtags, αλλά και το πλήθος των ακολούθων του λογαριασμού, είναι δυνατό να παρέχουν μία βαθύτερη οπτική των παραγόντων που καθορίζουν το ποσοστό επιρροής ενός δημιουργού περιεχομένου στο Instagram.

Λαμβάνοντας υπόψιν τα προαναφερθέντα, τα στοιχεία που συλλέχθηκαν από τον κάθε λογαριασμό του Instagram παρουσιάζονται παρακάτω:

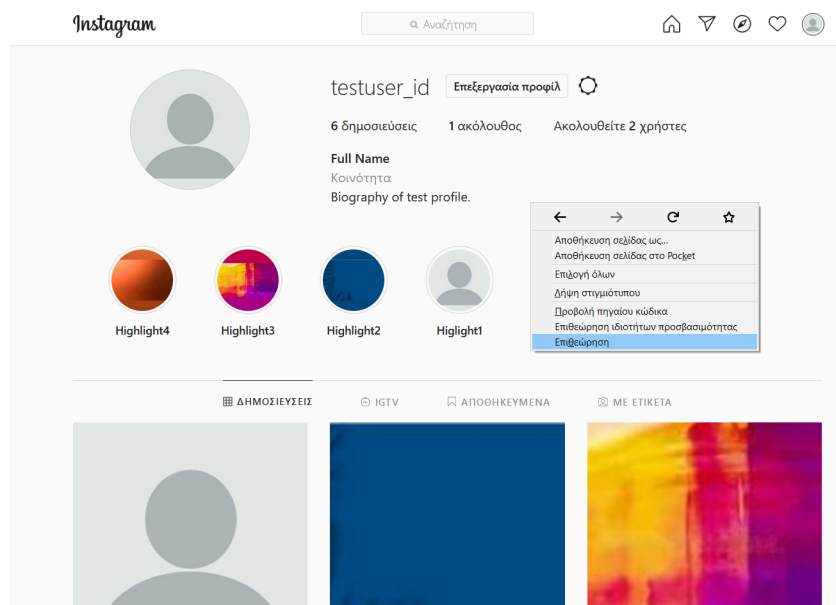
- **Κωδικό όνομα χρήστη (User name ID):** πεδίο που χρησιμοποιήθηκε ως το αναγνωριστικό του εκάστοτε χρήστη· αξίζει να σημειωθεί πως οποιαδήποτε αναφορά σε έναν λογαριασμό γίνεται μέσω του κωδικού ονόματος του
- **Αριθμός ακολούθων (Followers)/ακολουθούμενων (Following):** πεδία που βοήθησαν στην ανάλυση της επιρροής ενός λογαριασμού, με βάση το πλήθος των ακολούθων του και των ακολουθούμενων του
- **Κατηγορία λογαριασμού (Account category):** πεδίο που συνέβαλε στην ανάλυση της ψηφιακής παρουσίας ενός λογαριασμού, με βάση την θεματική κατηγορία του περιεχομένου του
- **Σύντομο βιογραφικό χρήστη (Biography):** πεδίο από το οποίο αποσπάστηκαν πιθανά hashtags και κωδικά ονόματα χρηστών, που σχετίζονται με τον συγκεκριμένο λογαριασμό
- **Περιγραφή δημοσίευσης (Captions):** πεδίο από το οποίο αποσπάστηκαν πιθανά hashtags και κωδικά ονόματα χρηστών, που σχετίζονται με τη συγκεκριμένη ανάρτηση
- **Αριθμός σχολίων δημοσίευσης (Comments):** πεδίο που βοήθησε στην ανάλυση της

απόδοσης μιας δημοσίευσης, με βάση τον αριθμό σχολίων που απέσπασε

- **Αριθμός αντιδράσεων “Μου αρέσει” δημοσίευσης (Likes):** πεδίο που βοήθησε στην ανάλυση της απόδοσης μιας δημοσίευσης, με βάση τον αριθμό των αντιδράσεων “Μου αρέσει” που απέσπασε
- **Προβολές βίντεο (Video views):** πεδίο που συνέβαλε στην αποτύπωση της ακροαματικότητας μιας δημοσίευσης-βίντεο, με βάση το πλήθος των ατόμων που το παρακολούθησαν
- **Ημερομηνία κοινοποίησης δημοσίευσης (Upload date):** πεδίο που βοήθησε στην ανάλυση της απόδοσης των δημοσιεύσεων, με βάση τη χρονική περίοδο κατά την οποία αναρτήθηκαν
- **Χρήστες που έχουν επισυναφθεί στη δημοσίευση (Mentioned users)** πεδίο το οποίο συντέλεσε στη χαρτογράφηση των σχέσεων μεταξύ χρηστών, αλλά και στη μελέτη της επιρροής των επισυνημμένων λογαριασμών στην απόδοση της ανάρτησης

Η συγκομιδή των πρωτογενών στοιχείων της έρευνας, αναφορικά με το Instagram, πραγματοποιήθηκε μέσω της επεξεργασίας του πηγαίου κώδικα της πλατφόρμας και, συγκεκριμένα, ενός HTML στοιχείου *script*, το οποίο είναι προσβάσιμο από τον κάθε φυλλομετρητή. Για την εύρεση του τμήματος κώδικα που περιλαμβάνει τις επιθυμητές πληροφορίες απαιτούνται οι εξής ενέργειες:

1. Επίσκεψη στον λογαριασμό, από τον οποίο θα συλλεχθούν τα δεδομένα
2. Επιθεώρηση του πηγαίου κώδικα της σελίδας, πατώντας δεξί κλικ και επιλέγοντας *Επιθεώρηση* ή κάποια αντίστοιχη λειτουργία που παρέχεται από το πρόγραμμα περιήγησης (βλ. Σχήμα 4.3)



Σχήμα 4.3: Επιθεώρηση του πηγαίου κώδικα της σελίδας

3. Εύρεση του HTML στοιχείου *script*, το οποίο περικλείεται στο στοιχείο *body* του HTML κώδικα της σελίδας και περιέχει τη JavaScript μεταβλητή *window._sharedData* (βλ. Σχήμα

4.4) η μεταβλητή αυτή ενθυλακώνει μια δομή JSON με όλα τα διαθέσιμα προς συλλογή στοιχεία του λογαριασμού

```
<!DOCTYPE html>
<html class="js logged-in client-root js-focus-visible sDNSV" lang="el" >
<head>
<body class="" style="" >
<div id="react-root" >
<link rel="stylesheet" href="/static/bundles/es6/ConsumerUICommons.css/59a1ecc2764.css" type="text/css" crossorigin="anonymous">
<script type="text/javascript">
window._sharedData = {
  "config": {
    "csrf_token": "D6G3DKXU833ofqU53UUhW674C8n1M0Qq",
    "viewer": {
      "biography": "Biography of test profile.",
      "business_contact_method": null,
      "business_email": null,
      "business_phone_number": null,
      "can_see_organic_insights": true,
      "category_name": "u039a\u03bf\u03b9\u03bd\u03cc",
      "external_url": null,
      "fbid": "17841446368038548",
      "full_name": "Full Name",
      "has_phone_number": false,
      "has_profile_pic": true,
      "has_tabbed_inbox": true,
      "hide_like_and_view_counts": false,
      "id": "46324385601",
      "is_joined_recently": false,
      "is_private": false,
      "is_professional_account": true,
      "profile_pic_url": "https://instagram.fath3-4.fna.fbcdn.net/v/t51.2885-19/5150x150/37405323_667260650?count=0",
      "comments_disabled": false,
      "taken_at_timestamp": "1620494086",
      "edge_liked_by": {
        "count": 36,
        "location": "4s08480/182858838_24029674551603_357810823898753172_n.jpg?tp1\u0026_nc_ht=instagram.fath3-4.fna.fbcdn.net\u0026_nc_cat=107\u0026_nc_ohc=46836"
      }
    }
  },
  "node": {
    "typename": "GraphImage",
    "id": "2551108940571056510",
    "shortcode": "CnNXCqD5L-",
    "dimensions": {
      "height": 1080,
      "width": 1080
    },
    "display_url": "https://instagram.fath3-4.fna.fbcdn.net/v/t51.2885-19/5150x150/37405323_667260650?count=0",
    "edge_media_to_tagged_user": {
      "edges": [
        {
          "owner": {
            "id": "5689845662",
            "username": "_saintforth"
          },
          "is_video": false
        }
      ]
    },
    "fact_check_overall_rating": null,
    "fact_check_information": null,
    "gating_info": null,
    "sharing_friction_info": {
      "should_have_sharing_friction": false,
      "blocs_app_url": null
    },
    "media_overlay_info": null,
    "media_preview": "ACoqzRS030fu1KUYDj5T0NVZCzMsQVx0eM8909S3RiJ/"
  }
}
```

Σχήμα 4.4: Εύρεση του HTML στοιχείου script

Στο Σχήμα 4.5 παρουσιάζονται διαδοχικά τα τμήματα της JSON δομής, τα οποία περιγράφουν τις γενικές πληροφορίες ενός λογαριασμού και τα στοιχεία μιας ανάρτησης, αντίστοιχα.

```
{
  "config": {
    "csrf_token": "D6G3DKXU833ofqU53UUhW674C8n1M0Qq",
    "viewer": {
      "biography": "Biography of test profile.",
      "business_contact_method": null,
      "business_email": null,
      "business_phone_number": null,
      "can_see_organic_insights": true,
      "category_name": "u039a\u03bf\u03b9\u03bd\u03cc",
      "external_url": null,
      "fbid": "17841446368038548",
      "full_name": "Full Name",
      "has_phone_number": false,
      "has_profile_pic": true,
      "has_tabbed_inbox": true,
      "hide_like_and_view_counts": false,
      "id": "46324385601",
      "is_joined_recently": false,
      "is_private": false,
      "is_professional_account": true,
      "profile_pic_url": "https://instagram.fath3-4.fna.fbcdn.net/v/t51.2885-19/5150x150/37405323_667260650?count=0",
      "comments_disabled": false,
      "taken_at_timestamp": "1620494086",
      "edge_liked_by": {
        "count": 36,
        "location": "4s08480/182858838_24029674551603_357810823898753172_n.jpg?tp1\u0026_nc_ht=instagram.fath3-4.fna.fbcdn.net\u0026_nc_cat=107\u0026_nc_ohc=46836"
      }
    }
  },
  "node": {
    "typename": "GraphImage",
    "id": "2551108940571056510",
    "shortcode": "CnNXCqD5L-",
    "dimensions": {
      "height": 1080,
      "width": 1080
    },
    "display_url": "https://instagram.fath3-4.fna.fbcdn.net/v/t51.2885-19/5150x150/37405323_667260650?count=0",
    "edge_media_to_tagged_user": {
      "edges": [
        {
          "owner": {
            "id": "5689845662",
            "username": "_saintforth"
          },
          "is_video": false
        }
      ]
    },
    "fact_check_overall_rating": null,
    "fact_check_information": null,
    "gating_info": null,
    "sharing_friction_info": {
      "should_have_sharing_friction": false,
      "blocs_app_url": null
    },
    "media_overlay_info": null,
    "media_preview": "ACoqzRS030fu1KUYDj5T0NVZCzMsQVx0eM8909S3RiJ/"
  }
}
```

Σχήμα 4.5: Η JSON δομή ενός λογαριασμού Instagram

Ωστόσο, αρκετά ήταν τα δεδομένα, τα οποία παρουσίαζαν ιδιαίτερο ενδιαφέρον, αλλά ήταν ανέφικτη η συλλογή τους, είτε λόγω της σύντομης δημόσιας παρουσίας τους στην πλατφόρμα (ιστορίες Instagram), είτε λόγω των μέτρων διαφύλαξης των προσωπικών δεδομένων που εφαρμόζει το Instagram (αριθμός κοινοποιήσεων και αποθηκεύσεων μιας ανάρτησης, στοιχεία σημαντικά για τον αλγόριθμο προώθησης του Instagram [51]).

Η αντιμετώπιση των πρωτογενών στοιχείων της έρευνας ως μεμονωμένα αριθμητικά δεδομένα

ων αναδείχθηκε κρίσιμη για την εξαγωγή πρώιμων συμπερασμάτων, τα οποία αποτέλεσαν τον θεμέλιο λίθο για την εξιχνίαση σύνθετων συσχετίσεων μεταξύ των συγκεντρωμένων πληροφοριών, μέσω του σχηματισμού των δευτερογενών στοιχείων. Τα δευτερογενή στοιχεία της μελέτης προσέδωσαν έναν πιο ακριβή και αντιπροσωπευτικό χαρακτήρα στην τεκμηρίωση της έρευνας, σχετικά με την αποτύπωση των κοινωνικών και επαγγελματικών σχέσεων που αναπτύσσονται μεταξύ των λογαριασμών του Instagram. Πιο συγκεκριμένα, το σύνολο των δευτερογενών μετρικών που δημιουργήθηκαν κατά την εξέταση της προκείμενης πλατφόρμας παρουσιάζεται στην ακόλουθη λίστα (υπενθυμίζεται πως το σύνολο των δημοσιεύσεων αφορά στη χρονική περίοδο Ιανουάριος 2020-Ιανουάριος 2021) [41]:

- **Engagement Rate Post:**

$$\frac{\text{Αντιδράσεις "Μου αρέσει" μιας δημοσίευσης}}{\text{Αριθμός ακολούθων του λογαριασμού}} \times 100$$

Μεταβλητή που βοηθάει στην ανάλυση της απόδοσης μιας δημοσίευσης, συγκρίνοντας τον αριθμό των θετικών αντιδράσεων της με τον αριθμό των ακολούθων του λογαριασμού που τη κοινοποίησε. Η συγκεκριμένη μεταβλητή μπορεί να καταλάβει τιμές από μηδέν, στις περιπτώσεις δημοσιεύσεων με μηδενικές θετικές αντιδράσεις, ενώ δεν είναι δυνατό να καθοριστεί το ανώτατο όριο τιμών, καθώς αρκετές είναι οι δημοσιεύσεις που συγκεντρώνουν αριθμό θετικών αντιδράσεων μεγαλύτερο από εκείνον των ακολούθων του αντίστοιχου λογαριασμού, όπως δημοσιεύσεις στις οποίες επισυνάπτονται επιπλέον λογαριασμοί ή χορηγούμενες αναρτήσεις.

- **Engagement Rate Comments Post:**

$$\frac{\text{Αντιδράσεις "Μου αρέσει" μιας δημοσίευσης} + \text{Σχόλια μιας δημοσίευσης}}{\text{Αριθμός ακολούθων του λογαριασμού}} \times 100$$

Μεταβλητή που παρέχει μια βαθύτερη εικόνα για την απόδοση μιας δημοσίευσης, σε σχέση με τη μεταβλητή *Engagement Rate Post*, καθώς συνυπολογίζει τον αριθμό των συγκεντρωμένων σχολιασμών της ανάρτησης. Αντίστοιχα με τη μεταβλητή *Engagement Rate Post*, η συγκεκριμένη μεταβλητή μπορεί να καταλάβει τιμές από μηδέν, στις περιπτώσεις δημοσιεύσεων με μηδενικές θετικές αντιδράσεις και σχολιασμούς, ενώ δεν είναι δυνατό να καθοριστεί το ανώτατο όριο τιμών, καθώς αρκετές είναι οι δημοσιεύσεις που συγκεντρώνουν αριθμό θετικών αντιδράσεων και σχολιασμών μεγαλύτερο από εκείνον των ακολούθων του αντίστοιχου λογαριασμού, όπως δημοσιεύσεις στις οποίες επισυνάπτονται επιπλέον λογαριασμοί ή αναρτήσεις διαγωνισμών.

- **Engagement Rate Views:**

$$\frac{\text{Αντιδράσεις "Μου αρέσει" μιας δημοσίευσης βίντεο}}{\text{Προβολές μιας δημοσίευσης βίντεο}} \times 100$$

. Μεταβλητή που βοηθάει στην ανάλυση της απόδοσης μιας δημοσίευσης τύπου βίντεο, συγκρίνοντας τον αριθμό των θετικών αντιδράσεων της με τον συνολικό αριθμό των προβολών που απέσπασε. Η συγκεκριμένη μεταβλητή μπορεί να καταλάβει τιμές από μηδέν, στις περιπτώσεις βίντεο με μηδενικές θετικές αντιδράσεις, ενώ δεν είναι δυνατό να καθοριστεί το ανώτατο όριο τιμών, καθώς αρκετά είναι τα βίντεο που συγκεντρώνουν αριθμό

θετικών αντιδράσεων μεγαλύτερο από εκείνον των προβολών τους, όπως αναρτήσεις πολλαπλού περιεχομένου (*slideshow*), στις οποίες βίντεο που βρίσκονται σε μεγαλύτερο βάθος της συλλογής δε παρακολουθούνται από όλα τα άτομα που αντιδρούν με αυτή.

- **Average Days Between Posts:**

$$\frac{\text{Άθροισμα των ημερών που μεσολάβησαν μεταξύ όλων των διαδοχικών δημοσιεύσεων}}{\text{Πλήθος των δημοσιεύσεων ενός λογαριασμού} - 1}$$

Μεταβλητή που συμβάλει στην αποτύπωση της μέσης συχνότητας δημοσίευσης αναρτήσεων του κάθε λογαριασμού.

4.1.2 Κώδικας Instagram

Στο σημείο αυτό, αφού προηγήθηκε η υψηλού επιπέδου περιγραφή της διαδικασίας συγκομιδής δεδομένων, το ενδιαφέρον πλέον στρέφεται προς την εσωτερική λειτουργία και τα δομικά στοιχεία του αυτοματοποιημένου μηχανισμού που σχεδιάστηκε. Στην παρούσα υποενότητα εξετάζονται τα κυριότερα τμήματα κώδικα που συνείσφεραν στην υλοποίηση των διάφορων λειτουργιών, τόσο στο επίπεδο της παραμετροποίησης του μηχανισμού όσο και της συλλογής και αποθήκευσης των δεδομένων. Καθοριστική για το σχεδιασμό του μηχανισμού συγκομιδής δεδομένων από το Instagram ήταν η συμβολή του άρθρου *The Easy Way to Scrape Instagram Using Python Scrapy & GraphQL* [20].

Αρχείο `settings.py`

Προκειμένου να αρχικοποιηθούν κάποιες σημαντικές, για τον σκοπό λειτουργίας του μηχανισμού, παράμετροι και για να διευθυνθούν οι επιμέρους διαδικασίες με τον επιθυμητό τρόπο, κρίθηκε απαραίτητη η μερική τροποποίηση του αρχείου `settings.py`, το οποίο φιλοξενεί το σύνολο όλων των διαθέσιμων ρυθμίσεων του μηχανισμού.

Οι παράμετροι που συναντούνται πρώτες στο αρχείο `settings.py` ενός *Scrapy Spider* παρουσιάζονται με την ακόλουθη αλληλουχία:

```
BOT_NAME = 'instascraper'
```

```
SPIDER_MODULES = ['instascraper.spiders']
```

```
NEWSPIDER_MODULE = 'instascraper.spiders'
```

Σύμφωνα με τις παραπάνω ρυθμίσεις ορίζεται η ονομασία που φέρει το πρότζεκτ (`BOT_NAME`), καθώς και τα τμήματα του, στα οποία το Scrapy θα αναζητά υπάρχοντα *spiders*, καθώς και θα τοποθετεί τα νεοσύστατα από αυτά αρχεία (`SPIDER_MODULES` και `NEWSPIDER_MODULE` αντίστοιχα). Στην προκειμένη περίπτωση και οι δύο τελευταίες λειτουργίες πραγματοποιούνται στον φάκελο *spiders* του πρότζεκτ *instascraper* (`instascraper.spiders`).

Στη συνέχεια, το σύνολο των ρυθμίσεων που ακολουθούν έχουν ως στόχο να ελαχιστοποιήσουν την “επιθετικότητα” του *spider* κατά την συλλογή δεδομένων και, ως εκ τούτου, να αποτρέψουν τον εντοπισμό του μηχανισμού συγκομιδής δεδομένων από τα μέτρα ασφαλείας που λαμβάνει η πλατφόρμα του Instagram.

```

ROBOTSTXT_OBEY = False
COOKIES_ENABLED = False
LOG_LEVEL = "INFO"

DOWNLOAD_DELAY = 5
DOWNLOAD_TIMEOUT = 110
CONCURRENT_REQUESTS_PER_IP = 1

AUTOTHROTTLE_ENABLED = True
AUTOTHROTTLE_START_DELAY = 5
AUTOTHROTTLE_MAX_DELAY = 110
AUTOTHROTTLE_TARGET_CONCURRENCY = 1

```

Ειδικότερα, ως πρώτο βήμα για την αποφυγή της απαγόρευσης πρόσβασης στο Instagram τέθηκε η κατάργηση της υπακοής προς το αρχείο *robots.txt* της πλατφόρμας (βλ. Υποενότητα 2.1.6) και η απενεργοποίηση της αυτόματης λειτουργίας πρόσαψης των *Cookies* στα απεσταλμένα αιτήματα από το Scrapy. Το εργαλείο Scrapy, με βάση τις προκαθορισμένες ρυθμίσεις του, αποστέλλει τα αιτήματα διαδοχικά και σε ομάδες των 16, ωστόσο αυτό θα προκαλούσε συμφόρηση στους διακομιστές τους Instagram, με αποτέλεσμα να οριστεί ως ανώτατο όριο αιτημάτων ανα διεύθυνση IP το ένα αίτημα και να επιβληθεί μεταξύ των αιτημάτων ένα διάστημα αδράνειας των πέντε δευτερολέπτων. Για λόγους συνέπειας οριοθετήθηκε η μέγιστη αναμονή ανάκτηση μίας σελίδας στα 110 δευτερόλεπτα, μετά το πέρας της οποίας θα επέλθει η διακοπή της λειτουργίας. Με σκοπό την βελτιστοποίηση της εφαρμογής διαστημάτων αδράνειας μεταξύ των αιτημάτων, αξιοποιήθηκε η λειτουργία του Scrapy *AutoThrottle*, η οποία προσαρμόζει αυτόματα τον χρόνο καθυστέρησης, ανάλογα με τα δεδομένα της εκάστοτε διαδικασίας. Στα πλαίσια της λειτουργίας αυτής, καθορίστηκε ως αρχική διάρκεια αναμονής τα πέντε δευτερόλεπτα, η οποία θα προσαρμοστεί στη συνέχεια στις ανάγκες της διαδικασίας, ως μέγιστη διάρκεια αναμονής τα 110 δευτερόλεπτα, ενώ ως μέσος αριθμός διαδοχικών αιτημάτων το ένα, καθιστώντας με αυτόν τον τρόπο τη συγκομιδή δεδομένων ιδιαίτερα ήπια. Αξίζει να σημειωθεί, πως τα χρονικά διαστήματα που ορίστηκαν παραπάνω εξυπηρέτησαν της ανάγκες της παρούσας έρευνας, γεγονός που επιτρέπει στον εκάστοτε χρήστη να τις μεταβάλλει.

Ως τελευταία κατηγορία των απαραίτητων ρυθμίσεων μπορεί να θεωρηθεί η οργάνωση της ροής δεδομένων μεταξύ των δομικών στοιχείων του Scrapy, Pipelines και Middlewares (βλ. Υποενότητα 3.1.2).

```

ITEM_PIPELINES = {
    'instascraper.pipelines.DatabasePipeline': 300,
}

#SPIDER_MIDDLEWARES = {
#    'instascraper.middlewares.InstascraperSpiderMiddleware': 543,
#}

#DOWNLOADER_MIDDLEWARES = {
#    'instascraper.middlewares.InstascraperDownloaderMiddleware': 543

```

```
#}
```

Συγκεκριμένα, υπαγορεύτηκε τα αποτελέσματα της επεξεργασίας δεδομένων να οδηγούνται στην βάση δεδομένων διαμέσου του *DatabasePipeline*, του οποίου η δομή θα αναλυθεί παρακάτω. Η χρήση των Middlewares δεν κρίθηκε απαραίτητη, με αποτέλεσμα αυτά να μην συμπεριληφθούν στην υλοποίηση. Οι αριθμοί που παρατηρούνται δίπλα από κάθε Pipeline/Middleware αντιπροσωπεύουν το βαθμό προτεραιότητας των δομικών στοιχείων, όσον αφορά στη σειρά με την οποία προσπελούνται από τα εξαγόμενα δεδομένα. Για παράδειγμα, αν ορίζονταν δύο Item Pipelines, το *instascraper.pipelines.Pipeline1:300* και το *instascraper.pipelines.Pipeline2:305*, τότε τα δεδομένα που θα συλλέγονταν από την εκάστοτε ιστοσελίδα θα δρομολογούνταν από το *Engine* πρώτα προς το Pipelines1 και διαδοχικά προς το Pipeline2.

Αρχείο *items.py*

Με αφορμή το γεγονός ότι ο όγκος των δεδομένων που προορίζονταν για συλλογή ήταν αρκετά μεγάλος και ότι τα στοιχεία των λογαριασμών του Instagram, όπως αναφέρθηκε παραπάνω, μοιράζονται ένα κοινό δομικό πρότυπο, αποφασίστηκε να δημιουργηθεί ένα πρότυπο για όλα τα δεδομένα τα οποία θα αποθηκευτούν στη βάση δεδομένων, αξιοποιώντας το αρχείο *items.py* του πρότζεκτ, έτσι ώστε να ομαδοποιηθούν οι σχετικές μεταξύ τους πληροφορίες και, παράλληλα, να επιτευχθεί η μέγιστη δυνατή οργάνωση.

```
class UserData(scrapy.Item):
    personal_info = scrapy.Field()
    tagged_users = scrapy.Field()
    hashtags = scrapy.Field()
    user_posts = scrapy.Field()

class PersonalInfo(scrapy.Item):
    user_name = scrapy.Field()
    user_id = scrapy.Field()
    account_type = scrapy.Field()
    account_category = scrapy.Field()
    owned_by = scrapy.Field()
    followers = scrapy.Field()
    following = scrapy.Field()
    posts = scrapy.Field()
    videos = scrapy.Field()
    avg_erpost = scrapy.Field()
    avg_erview = scrapy.Field()
    avg_likes = scrapy.Field()
    avg_comments = scrapy.Field()
    avg_engagement = scrapy.Field()
    avg_days_between_posts = scrapy.Field()

class RegularPost(scrapy.Item):
    post_id = scrapy.Field()
    post_type = scrapy.Field()
    post_date_timestamp = scrapy.Field()
```

```

likes = scrapy.Field()
comments = scrapy.Field()
views = scrapy.Field()
post_tagged_users = scrapy.Field()
post_hashtags = scrapy.Field()
er_view = scrapy.Field()
er_post = scrapy.Field()
er_comments_post = scrapy.Field()

```

```

class SlideShow(scrapy.Item):
    post_id = scrapy.Field()
    post_type = scrapy.Field()
    post_date_timestamp = scrapy.Field()
    likes = scrapy.Field()
    comments = scrapy.Field()
    post_tagged_users = scrapy.Field()
    post_hashtags = scrapy.Field()
    er_post = scrapy.Field()
    er_comments_post = scrapy.Field()
    slidePosts = scrapy.Field()
    slideshow_erview = scrapy.Field()

```

```

class SlidePost(scrapy.Item):
    slide_id = scrapy.Field()
    slide_type = scrapy.Field()
    slide_views = scrapy.Field()
    er_view = scrapy.Field()

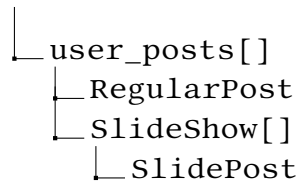
```

Όπως γίνεται αντιληπτό από τα περιεχόμενα του αρχείου *items.py*, τα στοιχεία του εκάστοτε λογαριασμού έχουν τμηματοποιηθεί σε επιμέρους δομές δεδομένων, με κύρια υποδομή να αποτελεί η κλάση *UserData*, η οποία ενθυλακώνει το σύνολο των ομαδοποιημένων πληροφοριών του λογαριασμού. Αυτό περιλαμβάνει τις γενικές πληροφορίες του λογαριασμού (*personal_info*), το σύνολο των επισυναπτόμενων χρηστών (*tagged_users*), το σύνολο των αναφερόμενων *hashtags* (*hashtags*) και τον συνολικό αριθμό των δημοσιεύσεων (*user_posts*), καθ' όλη την περίοδο του 2020. Κατά τη διάρκεια της συλλογής δεδομένων από ένα προφίλ του Instagram, παράγονται δομές για κάθε πιθανό τύπο δημοσιευμένης ανάρτησης του χρήστη (Regular post: φωτογραφία ή βίντεο, Slideshow: Συλλογή φωτογραφιών ή βίντεο), οι οποίες σε συνεργασία με τις γενικότερες πληροφορίες του προφίλ, συμπληρώνουν τα παραπάνω πεδία και συγκροτούν τελικά μια ολοκληρωμένη και έτοιμη για αποθήκευση συλλογή δεδομένων. Αξίζει να επισημανθεί πως υπάρχει πλήρης αντιστοιχία μεταξύ των πεδίων των δομών δεδομένων και των πληροφοριών ενός λογαριασμού που επιλέχθηκε να εξαχθούν στην Υποενότητα 4.1.1. Παρακάτω παρουσιάζεται η διαστρωμάτωση μιας δομής δεδομένων που αντιστοιχεί σε έναν λογαριασμό του Instagram.

```

UserData
├── personal_info
│   └── PersonalInfo
├── tagged_users[]
└── hashtags[]

```



Αρχείο `pipelines.py`

Με σκοπό την αποθήκευση των συγκεντρωμένων δεδομένων από την πλατφόρμα του Instagram στη βάση δεδομένων, τροποποιήθηκε καταλλήλως και, στη συνέχεια, χρησιμοποιήθηκε το αρχείο `pipelines.py`.

```

#db_address = address of database

from pymongo import MongoClient

class DatabasePipeline:

    myclient = MongoClient(<db_address>)
    db = myclient["Thesis"]
    collection = db["InstagramAccounts"]

    def process_item(self, item, spider):
        user_name=item['personal_info']['user_name']
        self.collection.replace_one(
            {'personal_info.user_name':user_name},
            item, upsert=True)

```

Όπως περιγράφηκε στην Υποενότητα 3.1.2, οι πληροφορίες της κάθε ιστοσελίδας, αφού εξαχθούν από το *Engine* του μηχανισμού συγκομιδής δεδομένων, αποστέλλονται στα *Item Pipelines*, απ' όπου και αποθηκεύονται στην επιθυμητή δομή δεδομένων. Σεβόμενοι την τακτική αυτή του Scrapy, ορίστηκε οι συλλογές πληροφοριών για τον κάθε λογαριασμό, να οδηγούνται στο Pipeline με όνομα *DatabasePipeline*, όπου εγκαθιδρύεται η σύνδεση με την βάση και οι συλλογές αποθηκεύονται σε αυτή με την μορφή εγγράφων. Όλα τα έγγραφα που αφορούν στη πλατφόρμα του Instagram στεγάζονται στη συλλογή με το όνομα *InstagramAccounts*, η οποία με τη σειρά της είναι μέρος της βάσης δεδομένων *Thesis*. Η βάση *Thesis* φιλοξενεί όλα τα στοιχεία της παρούσας έρευνας. Στην περίπτωση κατά την οποία τα στοιχεία για κάποιο προφίλ είναι ήδη αποθηκευμένα στην βάση, τότε αυτά υπερεγγράφονται μέσω της μεθόδου `self.Collection.replace_one()`.

Φάκελος `resources`

Κατά τη συγγραφή του κώδικα του μηχανισμού συγκομιδής δεδομένων, θεωρήθηκε σκόπιμο λίστες δεδομένων, οι οποίες καταλάμβαναν μεγάλο όγκο στο αρχείο του *spider* και αποτελούσαν σημαντικό πόρο για την λειτουργία του, να αποσπαστούν από το συγκεκριμένο αρχείο και να καταλάβουν αυτόνομη θέση σε έναν νέο φάκελο με το όνομα *resources*. Μοναδικό παράδειγμα στην περίπτωση του Instagram αποτελεί το JSON αρχείο *BusinessProfiles.json*, το οποίο εμπεριέχει έναν πίνακα με ζεύγη ονόματος-αναγνωριστικού που αντιστοιχούν σε λογαριασμούς Instagram ελληνικών επιχειρήσεων, των οποίων το περιεχόμενο αξιοποιήθηκε στα

πλαίσια της έρευνας. Πηγή των δεδομένων αποτέλεσε η ιστοσελίδα *socialbakers.com*¹⁵, από την οποία αποσπάστηκαν χειροκίνητα τα ονόματα 336 επιχειρησιακών λογαριασμών και, στη συνέχεια, αποθηκεύτηκαν στο αρχείο *BusinessProfiles.json* συνοδευόμενα από το αναγνωριστικό τους. Η μορφή που φέρει το συγκεκριμένο αρχείο παρουσιάζεται παρακάτω:

```
{
  "Profiles": [
    {
      "Name": "jumbostores",
      "Id": "jumbostoresgreece"
    },
    {
      "Name": "efood",
      "Id": "efoodgr"
    },
    .
    .
    .
    {
      "Name": "WIND Hellas",
      "Id": "wind_hellas"
    },
    {
      "Name": "Kotsovolos",
      "Id": "kotsovolos_greece"
    }
  ]
}
```

Φάκελος tools

Παρόμοιο σκοπό με εκείνον του φακέλου *resources* κλήθηκε να εξυπηρετήσει ο φάκελος *tools*. Τμήματα κώδικα, τα οποία θα μπορούσαν να επαναχρησιμοποιηθούν σε αρκετά σημεία της έρευνας, κρίθηκε προτιμότερο να συγκεντρωθούν σε έναν ξεχωριστό φάκελο, έτσι ώστε να εξοικονομηθεί χώρος στο βασικό κώδικα του μηχανισμού συγκομιδής δεδομένων και να επιτευχθεί η μέγιστη δυνατή ομαδοποίηση των διάφορων λειτουργιών. Η κύρια λειτουργία που υποστηρίχθηκε, κατά τη διαδικασία συλλογής στοιχείων, από τον φάκελο *tools*, ήταν η επικοινωνία με τη βάση δεδομένων και η διεξαγωγή των επιθυμητών αιτημάτων προς αυτήν με αντιπροσωπευτικό παράδειγμα να είναι το αρχείο *documents_exporter.py*.

```
#db_address = address of database
```

```
from pymongo import MongoClient
```

```
class DocumentsExporter:
```

```
    myclient = MongoClient(<db_address>)
```

¹⁵<https://www.socialbakers.com/statistics/youtube/channels/greece>

```
db = myclient["Thesis"]
collection = db["InstagramAccounts"]
```

Το συγκεκριμένο αρχείο φιλοξενεί μια πληθώρα μεθόδων που αφορούν στην εύρεση και συλλογή επιλεγμένων πεδίων της βάσης δεδομένων. Μια τέτοια μέθοδος είναι και η *findDocuments()*, η οποία επιτρέπει τη δημιουργία και την αποστολή παντός τύπου ερωτημάτων *find* στη βάση. Η συγκεκριμένη μέθοδος δέχεται ως ορίσματα: (i) την παράμετρο φιλτραρίσματος (*filter*), (ii) την παράμετρο προβολής (*project*), (iii) τον αριθμό παραλειπομένων εγγράφων (*skip_num*) και (iv) τον αριθμό των επιστρεφόμενων εγγράφων (*limit_num*), τα οποία συναρμολογούνται σε ένα κοινό ερώτημα *find* και αποστέλλονται στη βάση δεδομένων. Τελικά, τα αποτελέσματα της αναζήτησης επιστρέφονται μέσω της εντολής *return*.

```
def findDocuments(self, filter, project, skip_num, limit_num):
    if skip_num != None:
        if limit_num != None:
            if project != None:
                return self.collection.find(filter,
                    project).skip(skip_num).limit(limit_num)
            return
                self.collection.find(filter).skip(skip_num).limit(limit_num)
        else:
            if project != None:
                return self.collection.find(filter,
                    project).skip(skip_num)
            return self.collection.find(filter).skip(skip_num)
    else:
        if limit_num != None:
            if project != None:
                return self.collection.find(filter,
                    project).limit(limit_num)
            return self.collection.find(filter).limit(limit_num)
        return self.collection.find(filter)
```

Συγκέντρωση ονομάτων λογαριασμών

Επόμενο βήμα για τη συγκέντρωση του ερευνητικού υλικού, αναφορικά με το Instagram, υπήρξε η αναζήτηση των λογαριασμών που πληρούσαν τις προδιαγραφές της έρευνας και οι οποίοι, τελικά, θα εγκαθίδρυναν τη βάση της συλλογής πληροφοριών. Με αφορμή την πρόκληση που προέκυψε κατά τη διαδικασία ανακάλυψης νέων λογαριασμών (βλ. Ενότητα 4.3), το δείγμα της μελέτης στηρίχθηκε, κατά το πλείστον, στον ιστότοπο *starngage.com*¹³, ο οποίος στεγάζει λίστες με τους δημοφιλέστερους ελληνικούς λογαριασμούς του Instagram, ανάμεσα σε μία πληθώρα θεματικών κατηγοριών περιεχομένου. Η συγκεκριμένη ιστοσελίδα για κάθε θεματική κατηγορία παρέχει έναν πίνακα με ονόματα χρηστών καταταγμένα με βάση τον αριθμό των ακολούθων τους, περιεχόμενα τα οποία είναι προσβάσιμος μέσω ενός συνδέσμου της εξής μορφής:

¹³<https://starngage.com/app/global/influencer/ranking/greece>

```
https://starngage.com/app/global/influencer/ranking/greece
/<category_name>
```

Το γεγονός ότι το μεγάλο πλήθος λογαριασμών για κάθε κατηγορία δεν είναι δυνατό να συγχω-
νευτεί στα όρια της μιας σελίδας, οδήγησε την ιστοσελίδα *stangage.com* στη σελιδοποίηση του
εκάστοτε πίνακα περιεχομένων, με αποτέλεσμα η πρόσβαση στις επιμέρους σελίδες να απαιτεί
τη χρήση ενός συνδέσμου αντίστοιχης μορφής με τον παρακάτω:

```
https://starngage.com/app/global/influencer/ranking/greece
/<category_name>?<number_of_page>
```

Ο κώδικας για την υλοποίησης του μηχανισμού εύρεσης ονομάτων από την ιστοσελίδα *starngage.com*
ξεκινάει με την δήλωση των απαραίτητων μεταβλητών στιγμίουτυπου.

```
import scrapy
from ..items import UserData
from ..items import PersonalInfo

class StarngageSpider(scrapy.Spider):

    name = 'StarngageSpider'
    website_url="https://starngage.com/app/global/influencer/ranking/greece"
    allowed_domains = ['starngage.com']
    categories = ["All", "fashion", "celebrities", "art",
                 "health-fitness", "animals-pets", "architecture",
                 "cars-motorcycles", "cooking",
                 "design", "diy-crafts", "education",
                 "film-music-books", "food-drink", "game",
                 "gardening", "hair-beauty", "humor",
                 "kids-parenting", "nature-outdoors",
                 "photography", "technology", "travel"]
    visited_links=set()
```

Στο παραπάνω απόσπασμα του κώδικα δηλώνονται το όνομα του μηχανισμού (*name*), η διεύ-
θυνση (*website_url*) και το domain (*allowed_domains*) της ιστοσελίδας από την οποία θα συλλε-
χθούν τα ονόματα, μια λίστα με όλες τις διαθέσιμες, από την ιστοσελίδα, θεματικές κατηγορίες
περιεχομένου (*categories[]*), με σκοπό την αναζήτηση λογαριασμών με βάση αυτές και ένα βοη-
θητικό *set()* (*visited_links*), στο οποίο θα αποθηκεύονται οι επισκεπτόμενοι σύνδεσμοι, έτσι ώστε
να αποφευχθεί η επαναλαμβανόμενη προσπέλαση τους.

Στη συνέχεια, εκκινείται η διαδικασία αναζήτησης μέσω της μεθόδου *start_requests()*. Για κάθε
θεματική κατηγορία της λίστας *categories[]* καλείται η μέθοδος *getUrl()*, η οποία λαμβάνει ως
είσοδο την εκάστοτε κατηγορία και επιστρέφει τον σύνδεσμο του ιστοτόπου που της αντιστοι-
χεί, σε αλφαριθμητική μορφή. Τελικά, η μέθοδος *start_requests()* αποστέλλει ένα αίτημα για
κάθε έναν από τους προκύπτοντες συνδέσμους και τα αποτελέσματα της ανάκτησης οδηγού-
νται στην μέθοδο *parse()*.

```

def start_requests(self):
    for category in self.categories:
        yield
            scrapy.Request(self.getUrl(category), callback=self.parse)

def getUrl(self, string):
    if string=="All":
        return self.website_url
    return f"{self.website_url}/{string}"

```

Η μέθοδος *parse()* με την σειρά της, λαμβάνει την απόκριση για κάθε αίτημα της *start_requests* και αφότου αποσπάσει με την χρήση *xpath* τα ονόματα των εμφανιζόμενων λογαριασμών (*user_accounts*) και τους συνδέσμους που παραπέμπουν σε γειτονικές σελίδες (*page_links*), τότε συνθέτει μια δομή δεδομένων για κάθε έναν από τους λογαριασμούς (βλ. παράγραφο της Υποενότητας 4.1.2) και τις δρομολογεί προς αποθήκευση στη βάση δεδομένων με τη χρήση της εντολής *yield dict(user_data)*, η οποία μεταφέρει τις συλλογές δεδομένων σε μορφή *dictionary* από το *Engine* του μηχανισμού στα *Pipelines* (βλ. Παράγραφο της Υποενότητας Section: Αρχείο *pipelines.py*). Αφού ολοκληρωθούν τα παραπάνω, η διαδικασία επαναλαμβάνεται διαδοχικά για όλους τους συνδέσμους που εντοπίστηκαν, ωστόσο αυτοί εξαντληθούν.

```

def parse(self, r):
    user_accounts =
        r.xpath("//table/tbody//tr//a[@target='_blank']/text()").extract()
    page_links = r.xpath("//ul[@class='pagination
        justify-content-center']//li[@class='page-item
        ']/a/@href").extract()

    for user in user_accounts:
        user_data = UserData()
        personal_info = PersonalInfo()

        personal_info["user_name"] = user[1:]
        user_data["personal_info"] = personal_info

        yield dict(user_data)

    for link in page_links:
        if not link in self.visited_links:
            yield scrapy.Request(link, callback=self.parse)
        self.visited_links.add(link)

```

Αρχείο *instagram_spider*

Κατόπιν της αναζήτησης και αποθήκευσης ελληνικών λογαριασμών στη βάση δεδομένων, έλαβε δράση το αρχείο *instagram_spider.py*, το οποίο ενσωματώνει όλες τις απαραίτητες λειτουργίες για τη συγκομιδή στοιχείων από το Instagram. Πιο συγκεκριμένα, οι λειτουργίες αυτές αφορούν την αντιστοίχιση των ονομάτων που είναι καταγεγραμμένα στην βάση, με τους ομώνυμους λογαριασμούς της πλατφόρμας και, έπειτα, την συλλογή και αποθήκευση των ζητούμενων πλη-

ροφοριών τους, ακολουθώντας τα κριτήρια που τέθηκαν στην Υποενότητα 4.1.1.

Η συγγραφή του αρχείου *instagram_spider.py* ξεκινά με τη συμπερίληψη χρήσιμων βιβλιοθηκών και εργαλείων, καθώς με και την αρχικοποίηση των μεταβλητών στιγμιοτύπου της κύριας κλάσης του μηχανισμού *InstagramSpider()*.

```
import scrapy
import json
import re
from urllib.parse import urlencode
from datetime import datetime
from ..tools.documents_exporter import DocumentsExporter
from ..items import PersonalInfo
from ..items import SlideShow
from ..items import SlidePost
from ..items import UserData
from ..items import RegularPost

class InstagramSpider(scrapy.Spider):
    name = 'InstagramSpider'
    allowed_domains = ['instagram.com']
    users_to_scrape = []
    for item in
        list(DocumentsExporter.findDocuments(DocumentsExporter, {},
        None, 0, 1)):
            users_to_scrape.append(item["personal_info"]["user_name"])
scraped_users = []
followers_limit = 1000
mozilla_cookies = <cookies_session>
request_header = {
    "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64;
        rv:84.0) Gecko/20100101 Firefox/84.0",
    "Accept": "*/*",
    "Accept-Encoding": "gzip, deflate, br",
    "Accept-Language": "el-GR, el;q=0.8, en-US;q=0.5, en;q=0.3",
    "Connection": "keep-alive",
    "Cookie": mozilla_cookies,
    "Host": "www.instagram.com",
    "Referer": "https://www.instagram.com/?hl=el",
    "TE": "Trailers",
    "X-IG-App-ID": "936619743392459",
    "X-IG-WWW-Claim":
        "hmac.AR2l9mHkjaBgGD3I5V-gBGSH5Kc2uZy07Rav-x0jX-9f5h-j",
    "X-Requested-With": "XMLHttpRequest"
}
```

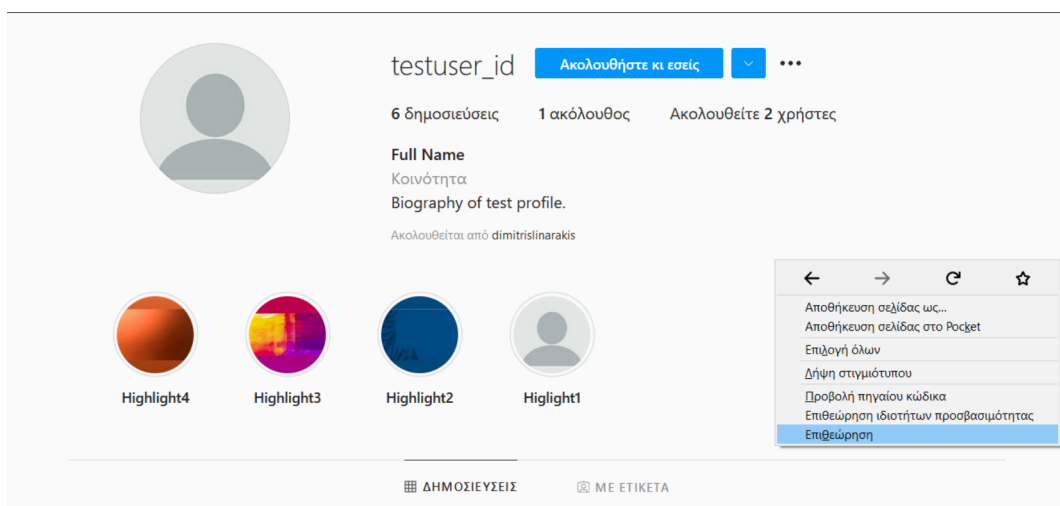
Εξετάζοντας τον παραπάνω κώδικα, αρχικά, παρατηρείται η απόδοση ονόματος στον μηχανισμό συγκομιδής δεδομένων (*name*) και η δήλωση της ιστοσελίδας από την οποία θα αντληθούν δεδομένα (*allowed_domains*). Στη συνέχεια, με τη βοήθεια του εργαλείου *DocumentsExporter* (βλ. Παράγραφο στην Υποενότητα 4.1.2) εξάγονται από την βάση δεδομένων τα ονόματα που

συλλέχθηκαν κατά τη διαδικασία αναζήτησης και αποθηκεύονται στη λίστα (`users_to_scrape[]`), η οποία θα περιλαμβάνει το σύνολο των λογαριασμών για τους οποίους θα συγκεντρωθούν δεδομένα. Ο ασύγχρονος τρόπος λειτουργίας του μηχανισμού, κατέστησε αναγκαία την καταγραφή των λογαριασμών, για τους οποίους η διαδικασία συλλογής είχε ολοκληρωθεί, έτσι ώστε να προληφθεί η σύγχυση μεταξύ των εξαγόμενων πληροφοριών. Η ανάγκη αυτή ικανοποιήθηκε μέσω της δημιουργίας της υποστηρικτικής λίστα `scraped_users[]`. Επιπλέον, η απαίτηση για την αναζήτηση λογαριασμών με περισσότερους των 1.000 ακολούθων οδήγησε στην οριοθέτηση της διαδικασίας, μέσω της μεταβλητής `followers_limit`.

Κομβικό σημείο για την ορθή λειτουργία του μηχανισμού αποτέλεσε η προσπάθεια αποφυγής του αποκλεισμού πρόσβασης στην πλατφόρμα του Instagram, όπως αναφέρθηκε στο Κεφάλαιο 2, στην Υποενότητα 2.1.6, ένας από τους προτεινόμενους τρόπους αποτροπής του μπλοκαρίσματος από μία ιστοσελίδα είναι η εφαρμογή τυχαίων και συχνών διακοπών μεταξύ των διαδοχικών αιτημάτων. Κάτι τέτοιο, ωστόσο, δε στάθηκε αρκετό στην περίπτωση του Instagram, καθώς απαιτούνταν περαιτέρω πρακτικές εξακρίβωσης της ψηφιακής ταυτότητας των αποστέλλομενων αιτημάτων. Λαμβάνοντας υπόψιν το παραπάνω, ήταν αναπόφευκτη η εκμείευση των χρησιμοποιούμενων, από τα αιτήματα του Instagram, κεφαλίδων (`request_header`) κατά την αποστολή HTTP αιτημάτων και προσαρμογή τους στα αιτήματα του μηχανισμού συγκομιδής δεδομένων, με σκοπό την ενίσχυση της εμπιστευτικότητας τους. Επισημαίνεται πως το πεδίο `Cookie` των κεφαλίδων μεταβάλλεται ανά τακτά χρονικά διαστήματα, επομένως ο χρήστης θα πρέπει να μεριμνά για τη συστηματική ενημέρωση του πεδίου αυτού.

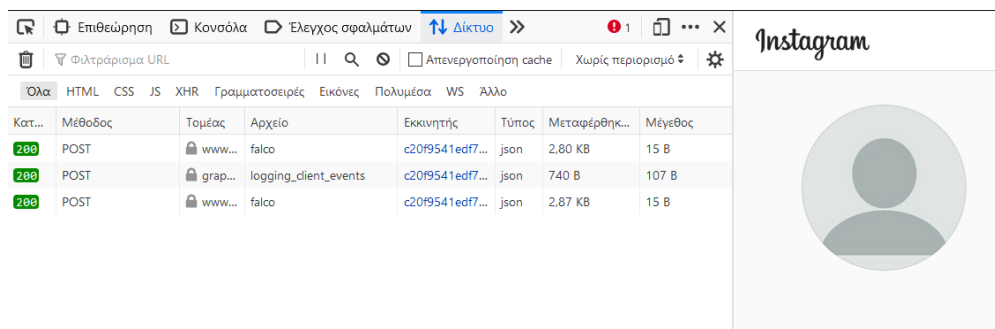
Οι κεφαλίδες ενός HTTP αιτήματος στο Instagram, και συγκεκριμένα για την περίπτωση χρήσης της μηχανής αναζήτησης Mozilla Firefox, η οποία και χρησιμοποιήθηκε στα πλαίσια της παρούσας πτυχιακής εργασίας, μπορούν να εντοπιστούν ακολουθώντας τα εξής βήματα:

- **Επιθεώρηση του πηγαίου κώδικα της σελίδας, πατώντας δεξί κλικ και επιλέγοντας *Επιθεώρηση* (βλ. Σχήμα 4.6)**



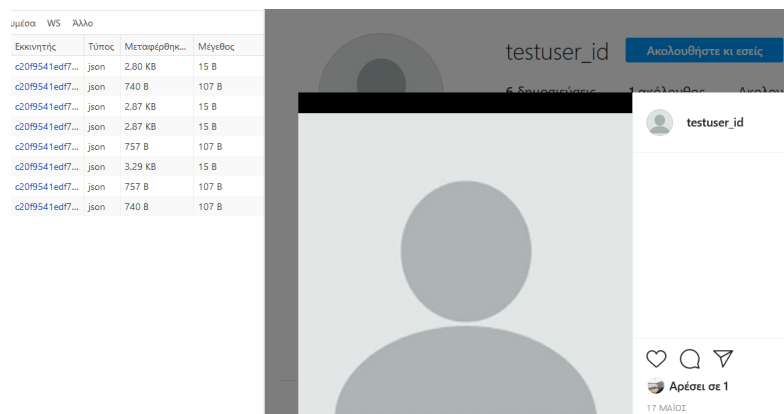
Σχήμα 4.6: Επιθεώρηση του πηγαίου κώδικα της σελίδας

- **Επιλογή της καρτέλας Δίκτυο (βλ. Σχήμα 4.7)**



Σχήμα 4.7: Επιλογή της καρτέλας Δίκτυο

- **Επιλογή ενός από τα διαδικτυακά στοιχεία της σελίδας του Instagram με σκοπό την αποστολή ενός στοχευμένου αιτήματος (βλ. Σχήμα 4.8)**



Σχήμα 4.8: Επιλογή διαδικτυακού στοιχείου

- **Εύρεση του στοχευμένου αιτήματος στη λίστα απεσταλμένων αιτημάτων της καρτέλας Δίκτυο· το αίτημα μπορεί να εντοπιστεί ευκολότερα αναζητώντας την ένδειξη `/graphql/query` (βλ. Σχήμα 4.9)**

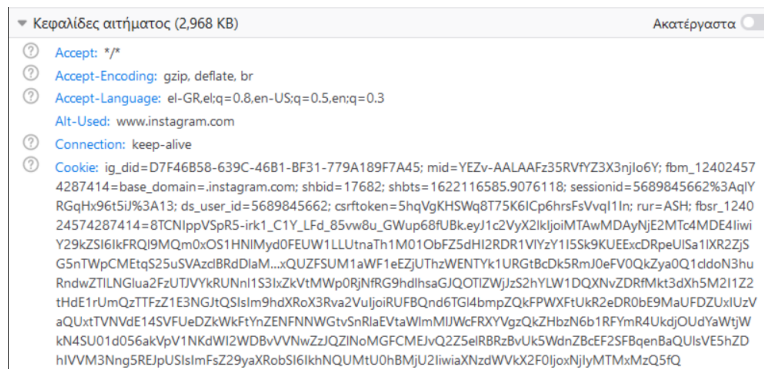
Κατ...	Μέθοδος	Τομέας	Αρχείο	Εκκινήτης	Τύπος	Μεταφέρθηκ...	Μέγεθος
200	GET	www...	/testuser_id/?_a=1	c20f9541edf7...	json	6,53 KB	22,52 KB
304	GET	insta...	187273296_1902822086	48e0f28aa47...	jpeg	στη μνήμη ca...	3,38 KB
304	GET	insta...	176565000_9304070775	48e0f28aa47...	jpeg	στη μνήμη ca...	3,82 KB
200	POST	www...	/qp/batch_fetch_web/?hl=	c20f9541edf7...	json	3,01 KB	159 B
200	GET	www...	/graphql/query?que...	c20f9541edf7...	json	7,09 KB	23,76 KB
304	GET	insta...	176565000_9304070775	c20f9541edf7...	jpeg	στη μνήμη ca...	3,82 KB
304	GET	insta...	187273296_1902822086	c20f9541edf7...	jpeg	στη μνήμη ca...	3,38 KB
304	GET	insta...	186297500_1215191358	imageset	jpeg	στη μνήμη ca...	3,60 KB

78 αιτήματα | 1,13 MB / 218,73 KB μεταφέρθηκαν | Ολοκλήρωση: 24,03 δ | DOMContentLoaded: 2,28 δ | load: 3,0 δ

Κεφαλίδες Cookies Αίτημα Απάντηση Χρονισμοί Ίχνος στοιβας Ασφάλεια

Σχήμα 4.9: Εύρεση του αιτήματος στη λίστα των απεσταλμένων αιτημάτων

- Εντοπισμός στο πεδίο *Κεφαλίδες αιτήματος* (βλ. Σχήμα 4.10)



Σχήμα 4.10: Κεφαλίδες του αιτήματος

Κατόπιν της δήλωσης των απαραίτητων μεταβλητών στιγμιοτύπου, το ενδιαφέρον στρέφεται στη μέθοδο `start_requests()`. Η μέθοδος `start_requests()`, όπως προαναφέρθηκε, είναι υπεύθυνη για την αυτόματη έναρξη της διαδικασίας συγκομιδής στοιχείων και είναι εκείνη η οποία, στην προκειμένη περίπτωση, αποστέλλει αιτήματα για τους ζητούμενους λογαριασμούς του Instagram, προσπελώνοντας τη λίστα `users_to_scrape`, και μεταβιβάζει τις αποκρίσεις τους στη μέθοδο `parse()`.

```
def start_requests(self):
    for user in self.users_to_scrape:
        profile_url = f"https://www.instagram.com/{user}/?hl=el"
        yield scrapy.Request(profile_url, callback=self.parse,
                             headers=self.request_header)
```

Οι παραπάνω αποκρίσεις διοχετεύονται στην μέθοδο `parse()`, με σκοπό να υποστούν επεξεργασία και να αντληθούν οι επιθυμητές πληροφορίες από αυτές. Όπως αναφέρθηκε στην Υποενότητα 4.1.1, κάθε απόκριση του Instagram που αφορά έναν συγκεκριμένο λογαριασμό εμπεριέχει μια JSON δομή, η οποία ενθυλακώνει όλα τα διαθέσιμα στοιχεία για αυτόν. Συμπερασματικά, ως πρώτο βήμα της επεξεργασίας της απόκρισης τίθεται η αποκοπή αυτού του τμήματος κώδικα μέσω των ακόλουθων εντολών:

```
def parse(self, response):
    response_script =
        response.xpath("//script[starts-with(., 'window._sharedData')]/text()").extract_first()
```

Εφόσον αποκτηθεί πρόσβαση στην δομή JSON, τότε διατίθενται όλοι οι απαιτούμενοι πόροι για τη σύσταση μιας συλλογής πληροφοριών για τον εκάστοτε χρήστη, με αυτές να απευθύνονται τόσο στις γενικές πτυχές του λογαριασμού του, όπως το όνομα χρήστη ή ο αριθμός ακολούθων του, όσο και σε συγκεκριμένα στοιχεία των αναρτήσεων του, όπως ο αριθμός αντιδράσεων ή οι διάφορες μετρικές απόδοσης. Όλα τα συγκεντρωμένα δεδομένα ενός λογαριασμού, τελικά, συμπεριλαμβάνονται σε μια κοινή δομή δεδομένων και αποθηκεύονται στη βάση δεδομένων

υπό τη μορφή εγγράφου. Λόγω του γεγονότος ότι η μέθοδος *parse()* ανάγεται σε έναν μεγάλο αριθμό γραμμών κώδικα, αποφασίστηκε, σε αυτό το σημείο, να μην συμπεριληφθεί αυτούσιο το τμήμα κώδικα που εξυπηρέτησε την απόσπαση στοιχείων από την δομή JSON, αλλά να περιγραφεί ο τρόπος λειτουργίας υπό τη μορφή απλοποιημένων προτάσεων, μέσω της ανάπτυξης ψευδοκώδικα σε αφηρημένο επίπεδο.

```

1. Μέθοδος parse (απόκριση_αιτήματος):
2.
3. Αν απόκριση_αιτήματος είναι έγκυρη:
4.   συλλεγμένες_πληροφορίες = {}
5.   γενικές_πληροφορίες_λογαριασμού = []
6.   αναρτήσεις_λογαριασμού = []
7.
8.   script_πληροφοριών = {}
9.   Ενημέρωσε το script_πληροφοριών με τα δεδομένα από την απόκριση_αιτήματος
10.
11. Αν πληροφορίες_λογαριασμού.ιδιωτικός_λογαριασμός δεν είναι αληθές:
12.   Αν πληροφορίες_λογαριασμού.αριθμός_ακολουθών είναι μεγαλύτερο από ή ίσο με 1000
13.
14.     Αποθήκευσε στις γενικές_πληροφορίες_λογαριασμού τις γενικές πληροφορίες του λογαριασμού
15.
16.   Για κάθε ανάρτηση από script_πληροφοριών.αναρτήσεις_λογαριασμού:
17.     Αν ανάρτηση.ημερομηνία_δημιουργίας είναι μεγαλύτερο από ή ίσο με 1/1/2020:
18.       πληροφορίες_ανάρτησης = []
19.       Αν ανάρτηση.τύπος_ανάρτησης είναι εικόνα ή βίντεο:
20.         Ενημέρωσε κατάλληλα τις πληροφορίες_ανάρτησης
21.         Αλλιώς αν ανάρτηση.τύπος_ανάρτησης είναι συλλογή εικόνων ή βίντεο:
22.           σύνολο_slides = []
23.           Ενημέρωσε το σύνολο_slides της ανάρτησης
24.           Για κάθε slide από σύνολο_slides:
25.             πληροφορίες_slide = []
26.             Ενημέρωσε κατάλληλα τις πληροφορίες_slide
27.             Πρόσθεσε στις πληροφορίες_ανάρτησης τις πληροφορίες_slide
28.
29.           Αποθήκευσε στις αναρτήσεις_λογαριασμού τις πληροφορίες_ανάρτησης
30.         Αλλιώς:
31.           Σταμάτα την αναζήτηση
32.
33.     Αποθήκευσε τις γενικές_πληροφορίες_λογαριασμού στις συλλεγμένες_πληροφορίες
34.     Αποθήκευσε τις αναρτήσεις_λογαριασμού στις συλλεγμένες_πληροφορίες
35.
36.   Αν (Υπάρχουν_επόμενες_12_δημοσιεύσεις = Αληθές) ΚΑΙ (Σταμάτα_την_αναζήτηση = Ψευδές):
37.     Στείλε αίτημα για τις επόμενες 12 δημοσιεύσεις
38.     Για κάθε απόκριση κάλεσε τη συνάρτηση parse_pages (απόκριση, συλλεγμένες_πληροφορίες)
39.   Αλλιώς:
40.     Στείλε στη βάση δεδομένων τις συλλεγμένες_πληροφορίες

```

Οι μέθοδοι *parse()* και *parse_pages()* οφείλουν την ύπαρξη τους στην σελιδοποίηση που εφαρμόζει το Instagram στις αναρτήσεις των λογαριασμών, ομαδοποιώντας τες σε σύνολα των 12. Το γεγονός ότι η απόκριση για την πρώτη δωδεκάδα αναρτήσεων ενός λογαριασμού εμφανίζει διαφορετική δομή από αυτή των επόμενων, ανέθεσε στην μέθοδο *parse()* την αρμοδιότητα της επεξεργασίας των πρώτων 12 δημοσιεύσεων του κάθε λογαριασμού, αν υπάρχουν τόσες, και, στη συνέχεια, της μεταβίβασης της μεταχείρισης των μεταγενέστερων αναρτήσεων στη μέθοδο *parse_pages()*.

Κατά συνέπεια, η διαδικασία συλλογής στοιχείων ξεκινά από την πιο πρόσφατα δημοσιευμένη ανάρτηση (Βήμα 16 της μεθόδου *parse*) και αν αυτή βρίσκεται εντός του θεμιτού χρονικού διαστήματος (Βήμα 17 της μεθόδου *parse*), τότε συλλέγονται οι επιθυμητές πληροφορίες για αυτή (Βήμα 18 έως Βήμα 29 της μεθόδου *parse*) και η διαδικασία επαναλαμβάνεται για τις υπολειπόμενες αναρτήσεις της πρώτης σελίδας. Στη περίπτωση που κάποια από τις προσπελάσιμες αναρτήσεις είναι δημοσιευμένες σε προγενέστερο χρονικό διάστημα από το αποδεκτό, η διαδικασία εξαγωγής πληροφοριών τερματίζεται (Βήμα 30 και Βήμα 31 της μεθόδου *parse*), καθώς υπενθυμίζεται πως οι δημοσιεύσεις είναι ταξινομημένες με βάση την ημερομηνία δημιουργίας τους. Με την προϋπόθεση ότι ο εξεταζόμενος λογαριασμός διαθέτει περισσότερες από 12 δημοσιεύσεις εντός του παραπάνω διαστήματος, τα μέχρι τότε συγκεντρωμένα στοιχεία του λογαριασμού ενοποιούνται σε μία κοινή δομή δεδομένων, η οποία μεταβιβάζεται στη μέθοδο *parse_pages()*, έτσι ώστε να συμπληρωθεί με τα στοιχεία των επόμενων δημοσιεύσεων (Βήμα 33 έως Βήμα 38 της μεθόδου *parse*). Εφόσον εξαντληθούν οι αναζητούμενες αναρτήσεις, τα συγκεντρωμένα δεδομένα αποθηκεύονται στη βάση δεδομένων και η διαδικασία ολοκληρώνεται (Βήμα 40 της μεθόδου *parse*). Επισημαίνεται πως η λειτουργία της μεθόδου *parse_pages()* είναι αντίστοιχη με εκείνη της *parse*.

1. Μέθοδος *parse_pages* (απόκριση, συλλεγμένες_πληροφορίες):
- 2.
3. Για κάθε ανάρτηση στις απόκριση.αναρτήσεις_λογαριασμού:
4. Αν ανάρτηση.ημερομηνία_δημιουργίας είναι μεγαλύτερο από ή ίσο με 1/1/2020:
5. Αν ανάρτηση.τύπος_ανάρτησης είναι εικόνα ή βίντεο:
6. πληροφορίες_ανάρτησης = []
7. **Ενημέρωσε** κατάλληλα τις πληροφορίες_ανάρτησης
8. **Αλλιώς αν** ανάρτηση.τύπος_ανάρτησης είναι συλλογή εικόνων ή βίντεο:
9. σύνολο_slides = []
10. **Ενημέρωσε** το σύνολο_slides της ανάρτησης
11. **Για κάθε** slide από σύνολο_slides:
12. πληροφορίες_slide = []
13. **Ενημέρωσε** κατάλληλα τις πληροφορίες_slide
14. **Πρόσθεσε** στις πληροφορίες_ανάρτησης τις πληροφορίες_slide
- 15.
16. **Αποθήκευσε** στις συλλεγμένες_πληροφορίες.αναρτήσεις_λογαριασμού τις πληροφορίες_ανάρτησης
- 17.
18. **Αλλιώς:**
19. Σταμάτα την αναζήτηση
- 20.
21. Αν (υπάρχουν_επόμενες_12_δημοσιεύσεις = Αληθές) **ΚΑΙ** (σταμάτα_την_αναζήτηση = Ψευδές):
22. **Στείλε** αίτημα για τις επόμενες 12 δημοσιεύσεις
23. **Για κάθε** απόκριση κάλεσε τη συνάρτηση *parse_pages* (νέα_απόκριση, συλλεγμένες_πληροφορίες)
24. **Αλλιώς:**
25. **Στείλε** στη βάση δεδομένων τις συλλεγμένες_πληροφορίες

4.2 Συγκομιδή στοιχείων του YouTube

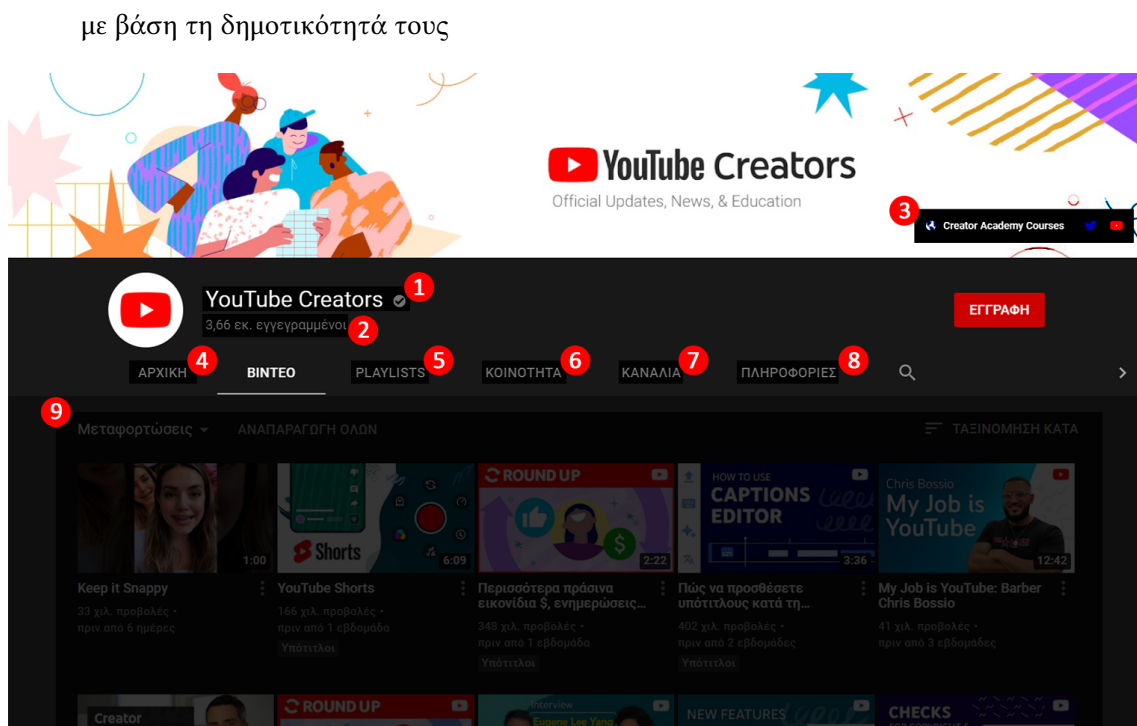
Αντίστοιχα με το Instagram, η διαδικασία συλλογής των απαραίτητων δεδομένων για το YouTube διακρίνεται σε δύο επίπεδα, εκείνο της συγκομιδής των πρωτογενών στοιχείων και εκείνο της σύνθεσης των δευτερογενών, με σκοπό τη μετάφραση των συγκεντρωμένων αριθμητικών δεδομένων σε στατιστικά στοιχεία και, κατ'επέκταση τη διαμόρφωση συμπερασμάτων για το ποσοστό επιρροής που ασκούν οι διάφοροι δημιουργοί περιεχομένου στο κοινό, αλλά και για τη σημασία των σχέσεων που αναπτύσσονται στα πλαίσια της ελληνικής κοινότητας του YouTube.

Στις επόμενες υποενότητες, αρχικά, αναλύεται η δομή ενός τυπικού λογαριασμού του YouTube, έτσι ώστε να σκιαγραφηθεί η πηγή των δεδομένων, έπειτα περιγράφονται τα στοιχεία που φάνηκαν χρήσιμα για την πραγματοποίηση της μελέτης και, καταληκτικά, γίνεται αναφορά σε προκλήσεις που προέκυψαν κατά τη διαδικασία συγκομιδής τους.

4.2.1 Δομή του YouTube και δεδομένα

Η τυπική διαστρωμάτωση ενός λογαριασμού YouTube παρουσιάζεται στο Σχήμα 4.11 και είναι η ακόλουθη:

1. **Όνομα καναλιού (Channel title):** το όνομα με το οποίο εμφανίζεται το κανάλι στο YouTube και οποιοσδήποτε χρήστης της πλατφόρμας μπορεί να το αναζητήσει με αυτό
2. **Αριθμός συνδρομητών (Subscriber count):** το σύνολο των χρηστών που έχουν επιλέξει να ενημερώνονται για τη δραστηριότητα του καναλιού, κάνοντας εγγραφή στο κανάλι
3. **Εξωτερικοί σύνδεσμοι καναλιού (External links):** επισυναπτόμενοι σύνδεσμοι που παραπέμπουν σε εξωτερικές πλατφόρμες ή ιστοσελίδες, με τις οποίες είναι άμεσα συνδεδεμένο το παρόν κανάλι
4. **Καρτέλα “Αρχική”:** καρτέλα στην οποία ο διαχειριστής του καναλιού καθορίζει το περιεχόμενο που θα συναντήσει αρχικά ένας χρήστης, κατά την επίσκεψή του στο κανάλι. Σε αυτή την ενότητα του καναλιού, ο διαχειριστής έχει τη δυνατότητα να προβάλλει επιλεγμένα βίντεο με την σειρά που εκείνος επιθυμεί, ή ακόμα και κανάλια άλλων χρηστών με τα οποία συσχετίζεται
5. **Καρτέλα “Playlists”:** καρτέλα στην οποία ο διαχειριστής του καναλιού καλείται να δημιουργήσει λίστες αναπαραγωγής, κατατάσσοντας τα βίντεο του σε θεματικές κατηγορίες της επιλογής του
6. **Καρτέλα “Κοινότητα”:** προαιρετική καρτέλα στην οποία ενσωματώνεται η υπό-ανάπτυξη λειτουργικότητα κοινωνικής δικτύωσης του YouTube, επιτρέποντας στον διαχειριστή του καναλιού να κοινοποιεί αναρτήσεις αντίστοιχης φιλοσοφίας με εκείνες των αρχέτυπων μέσω δικτύωσης, όπως φωτογραφίες, στιγμιότυπα ή δημοσκοπήσεις, και να αλληλεπιδρά σε πραγματικό χρόνο με το κοινό του
7. **Καρτέλα “Κανάλια”:** στην καρτέλα αυτή παρουσιάζονται τα κανάλια στα οποία είναι εγγεγραμμένο το παρόν κανάλι. ο διαχειριστής του καναλιού έχει τη δυνατότητα να απενεργοποιήσει αυτή την λειτουργία, ώστε να μην είναι δημοσίως ορατές οι εγγραφές του στο ευρύ κοινό
8. **Καρτέλα “Πληροφορίες”:** καρτέλα στην οποία αναγράφονται πληροφορίες σχετικές με το κανάλι, όπως η περιγραφή του καναλιού, η τοποθεσία του, εξωτερικοί σύνδεσμοι και στατιστικά στοιχεία που σχετίζονται με αυτό. στο πεδίο των στατιστικών ανήκουν η ημερομηνία δημιουργίας του καναλιού και ο συνολικός αριθμός προβολών των βίντεο του σε πραγματικό χρόνο
9. **Μεταφορτώσεις καναλιού (Channel videos):** οι μεταφορτώσεις του καναλιού βρίσκονται στην καρτέλα “Βίντεο” και αποτελούν το σύνολο όλων των κοινοποιημένων βίντεο του καναλιού, ταξινομημένα είτε με βάση την χρονολογική σειρά ανάρτησης τους, είτε



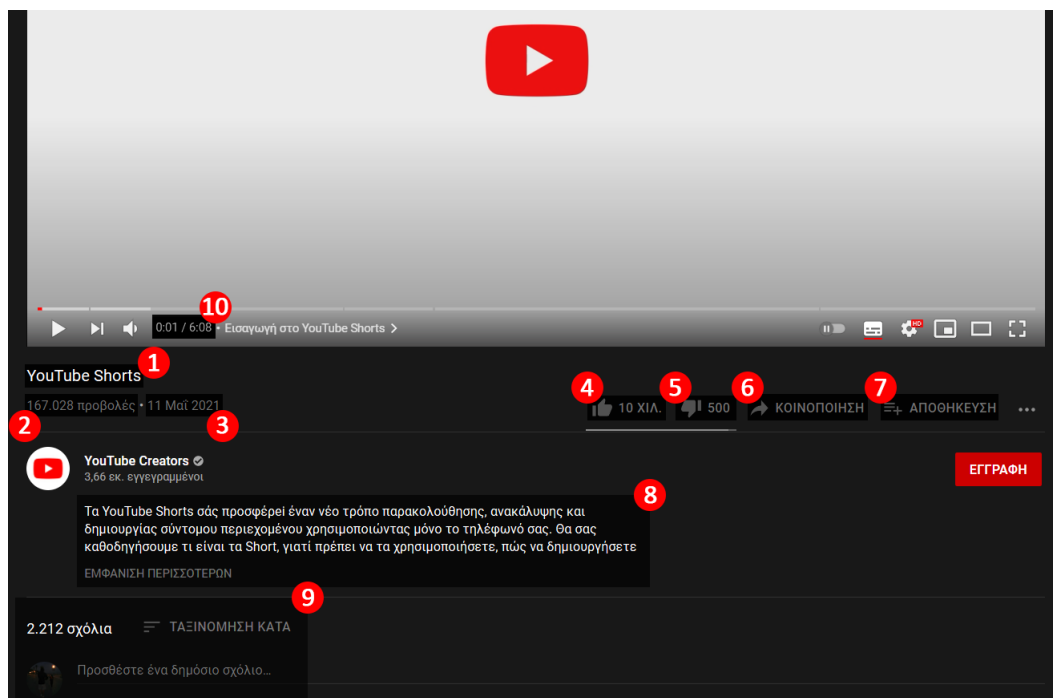
Σχήμα 4.11: Διαστρωμάτωση ενός λογαριασμού YouTube

Σχετικά με τη δομή των μεταφορτώσεων ενός καναλιού, τηρείται μία πρότυπη μορφή αναπαράστασης, με τις τελευταίες να μοιράζονται ένα σύνολο κοινών χαρακτηριστικών που αντικατοπτρίζουν την ανταπόκριση τους από το κοινό. Ωστόσο, αρκετά είναι τα κανάλια τα οποία, θέλοντας να διαφυλάξουν την δημόσια εικόνα τους, επιλέγουν να αποκρύψουν αρκετά από αυτά. Τα δομικά στοιχεία τα οποία συγκροτούν την πρότυπη μορφή ενός βίντεο φαίνονται στο Σχήμα 4.12 και αναλύονται ακολούθως:

1. **Τίτλος βίντεο (Video title):** η ονομασία που φέρει το βίντεο, με σκοπό την εύρεση του από τους χρήστες της πλατφόρμας
2. **Αριθμός προβολών βίντεο (View count):** ο συνολικός αριθμός προβολών που έχει συγκεντρώσει το βίντεο, με την προϋπόθεση ο εκάστοτε θεατής έχει ξεπεράσει το ελάχιστο χρονικό διάστημα παρακολούθησης (view threshold), μετρική που καθορίζεται από την διάρκεια του βίντεο
3. **Ημερομηνία μεταφόρτωσης βίντεο (Upload date):** η ημερομηνία κατά την οποία μεταφορτώθηκε το βίντεο στην πλατφόρμα
4. **Αριθμός αντιδράσεων “Μου αρέσει” του βίντεο (Like count):** ο συνολικός αριθμός των θετικών αντιδράσεων που έλαβε το βίντεο· η κάθε ψήφος αντιστοιχεί και σε έναν μοναδικό λογαριασμό
5. **Αριθμός αντιδράσεων “Δεν μου αρέσει” του βίντεο (Dislike count):** ο συνολικός αριθμός των αρνητικών αντιδράσεων που έλαβε το βίντεο· η κάθε ψήφος αντιστοιχεί και σε έναν μοναδικό λογαριασμό
6. **Κουμπί κοινοποίησης του βίντεο (Share button):** κουμπί που δίνει την δυνατότητα

στον χρήστη να κοινοποιήσει το βίντεο σε εξωτερικές πλατφόρμες, με αυτές να ποικίλλουν από μέσα κοινωνικής δικτύωσης μέχρι ιστοσελίδες· ο συνολικός αριθμός των κοινοποιήσεων του βίντεο δεν παρατίθεται από το YouTube προς τους επισκέπτες, παρά μόνο στον ιδιοκτήτη του καναλιού

7. **Κουμπί αποθήκευσης του βίντεο (Save button):** κουμπί που δίνει την δυνατότητα στον χρήστη να αποθηκεύσει το βίντεο στην λίστα των αγαπημένων του· ο συνολικός αριθμός των αποθηκεύσεων του βίντεο δεν παρατίθεται από το YouTube προς τους επισκέπτες, παρά μόνο στον ιδιοκτήτη του καναλιού
8. **Περιγραφή του βίντεο (Video description):** σύντομο κείμενο που μπορεί να περιγράφει το περιεχόμενο του βίντεο, να παρουσιάζει τους συντελεστές και τους συνεργάτες που συνέβαλαν στο έργο ή ακόμα και να παραθέτει τα στοιχεία επικοινωνίας τους
9. **Αριθμός σχολίων του βίντεο (Comment count):** ο συνολικός αριθμός σχολίων που απέσπασε το βίντεο
10. **Διάρκεια του βίντεο (Video duration):** η συνολική χρονική διάρκεια του βίντεο, χωρίς να συμπεριλαμβάνεται σε αυτή η διάρκεια των ενδεχόμενων αναδυόμενων διαφημίσεων



Σχήμα 4.12: Βασική δομή ενός βίντεο του YouTube

Καθαριστικό ρόλο στην επιλογή των κατάλληλων για την έρευνα στοιχείων του YouTube έπαιξε, πρωτίστως, η δυνατότητα συλλογής τους μέσω της χρήσης του YouTube Data API και, δευτερευόντως, το ποσοστό της ενδεχόμενης συνεισφοράς τους στην εκμείωση χρησμών συμπερασμάτων. Στα πλαίσια του YouTube υιοθετήθηκε αντίστοιχη φιλοσοφία με αυτή του Instagram, θέτοντας ως τον κεντρικό πόλο ενδιαφέροντος το σύνολο των αντιδράσεων που συγκεντρώνει το κάθε βίντεο. Λαμβάνοντας υπόψιν τον αριθμό των θετικών και αρνητικών αντιδράσεων ενός βίντεο, των σχολίων και των προβολών που αυτό απέσπασε είναι δυνατό να αντληθούν πληροφορίες για την ανταπόκριση του κοινού απέναντι σε αυτό, κάτι το οποίο, σε συνδυα-

σμό με το θεματικό περιεχόμενο του βίντεο αλλά και τη γενικότερη απόδοση των βίντεο ενός καναλιού, μπορεί να παράξει γνώση, τόσο όσον αφορά στο χαρακτήρα του εκάστοτε δημιουργού περιεχομένου του YouTube όσο και στη συμπεριφορά και στις προτιμήσεις του ελληνικού κοινού.

Όπως αναφέρθηκε και νωρίτερα, με σκοπό την συγκομιδή δεδομένων από το YouTube, αξιοποιήθηκε το εργαλείο YouTube Data API (βλ. Ενότητα 3.2). Η επικοινωνία με το συγκεκριμένο API επιτυγχάνεται μέσω της σύνταξης προτυποποιημένων αιτημάτων, τα οποία φέρουν παραμέτρους που υποδηλώνουν τα ζητούμενα από τον χρήστη στοιχεία. Έστω, λοιπόν, πως ένας χρήστης επιθυμεί να συλλέξει τις βασικές πληροφορίες για ένα κανάλι του YouTube, λόγου χάριν για το κανάλι *YouTube Creators*. Σε αυτή την περίπτωση θα πρέπει ο χρήστης να συντάξει και να αποστείλει στο API το κάτωθι αίτημα πληκτρολογώντας το στη γραμμή αναζήτησης του φυλλομετρητή που χρησιμοποιεί:

```
https://www.googleapis.com/youtube/v3/channels?
key=AIzaSyDZlmyWH7bFqlxxZHHvatMkdacjgi9mp8
&part=snippet,statistics
&fields=items(snippet/title,snippet/publishedAt,statistics)
&id=UCkRfArvrzheW2E7b6SVT7vQ
```

Το API, ως απάντηση στο παραπάνω αίτημα, επιστρέφει στον χρήστη τον τίτλο και την ημερομηνία δημιουργίας του καναλιού, καθώς και πληροφορίες στατιστικής φύσεως που αφορούν στο κανάλι, πληροφορίες οι οποίες κατατάσσονται στις, προκαθορισμένες από το ίδιο, υποκατηγορίες *snippet* και *statistics*. Στην Εικόνα 4.13 αναπαρίσταται η διάρθρωση της απόκρισης αυτής.

```

▼ items:
  ▼ 0:
    ▼ snippet:
      title: "YouTube Creators"
      publishedAt: "2013-05-06T21:25:34Z"
    ▼ statistics:
      viewCount: "200337635"
      subscriberCount: "3670000"
      hiddenSubscriberCount: false
      videoCount: "397"

```

Σχήμα 4.13: Ενδεχόμενη διάρθρωση μιας απόκρισης του YouTube Data API

Το YouTube Data API χάρη στο μεγάλο σύνολο πόρων που παρέχει στους χρήστες, κατάφερε να εμπλουτίσει το εύρος των συλλεχθέντων πληροφοριών, δίνοντας πρόσβαση σε δυσεύρετα, στα πλαίσια του πηγαίου κώδικα, στοιχεία, όπως η θεματική κατηγορία και οι λέξεις-κλειδιά ενός βίντεο. Συγκεντρωτικά, τα πρωτογενή στοιχεία του YouTube που φάνηκαν χρήσιμα στο σχηματισμό μιας γενικής εικόνας του ποσοστού επιρροής των χρηστών από διαφορετικού τύπου βίντεο παρατίθενται στην επόμενη λίστα:

- **Αναγνωριστικό καναλιού (Channel ID):** πεδίο που χρησιμοποιήθηκε ως το αναγνωριστικό ενός καναλιού· αξίζει να σημειωθεί πως οποιαδήποτε αναφορά σε ένα συγκεκριμένο κανάλι μπορεί να πραγματοποιηθεί με τη χρήση του αναγνωριστικού του
- **Όνομα καναλιού (Channel title):** πεδίο που βοήθησε στην αναγνώριση ενός καναλιού στην βάση δεδομένων μέσω του επίσημου ονόματός του, καθώς το αναγνωριστικό του, αν και μοναδικό, δεν είναι ευανάγνωστο από τον χρήστη
- **Λέξεις-κλειδιά καναλιού (Channel keywords):** πεδίο υπό τη μορφή λίστας, το οποίο περιέχει λέξεις-κλειδιά που μπορούν να βοηθήσουν το κανάλι να εμφανιστεί στα αποτελέσματα αναζήτησης κάποιας από τις καθορισμένες λέξεις· συνέβαλε στην ανάλυση της επιρροής που ασκούν οι λέξεις-κλειδιά στην απήχηση ενός καναλιού
- **Ημερομηνία δημιουργίας καναλιού (Channel creation date):** πεδίο που χρησιμοποιήθηκε για την εξέταση του ποσοστού επιρροής ενός καναλιού, σε σχέση με την ημερομηνία δημιουργίας του
- **Αριθμός συνολικών προβολών (Total views count):** χρησιμοποιήθηκε ως βασικό πεδίο για την ανάλυση της γενικότερης απήχησης και, κατ' επέκταση, της επιρροής που ασκεί ένα κανάλι στο κοινό που το παρακολουθεί, σε σχέση με τις αντιδράσεις και τα σχόλια που αυτό έχει λάβει συνολικά
- **Αριθμός συνδρομητών (Subscriber count):** πεδίο που βοήθησε στην εξέταση των συντελεστών απόδοσης των καναλιών με βάση τη κλίμακα εγγεγραμμένων χρηστών στην οποία κατατάσσονται
- **Συνολικός αριθμός βίντεο καναλιού (Video count):** πεδίο που συνέβαλε στην ανάλυση της δραστηριότητας ενός καναλιού, με βάση το συνολικό αριθμό βίντεο που έχει μεταφορτώσει από την ημερομηνία δημιουργίας του
- **Λέξεις-κλειδιά βίντεο (Video tags):** πεδίο υπό την μορφή λίστας, το οποίο φέρει λέξεις-κλειδιά που μπορούν να αξιοποιηθούν με σκοπό την εμφάνιση του βίντεο στα αποτελέσματα αναζήτησης κάποιας από τις καθορισμένες λέξεις· συνέβαλε στην ανάλυση της επιρροής που έχουν οι λέξεις-κλειδιά στην απόδοση ενός βίντεο
- **Κατηγορία βίντεο (Video category):** πεδίο που χρησιμοποιήθηκε για την ανάλυση της επιρροής ενός βίντεο, με βάση τη θεματική κατηγορία στην οποία ανήκει
- **Αναγνωριστικό βίντεο (Video ID):** πεδίο που χρησιμοποιήθηκε ως το αναγνωριστικό ενός βίντεο· αξίζει να σημειωθεί πως οποιαδήποτε αναφορά σε ένα συγκεκριμένο βίντεο μπορεί να πραγματοποιηθεί με τη χρήση του αναγνωριστικού του
- **Ημερομηνία μεταφόρτωσης βίντεο (Upload date):** πεδίο το οποίο χρησιμοποιήθηκε για τη συλλογή βίντεο που εντάσσονται εντός το χρονικού διαστήματος Ιανουάριος 2020-Ιανουάριος 2021 και βοήθησε στην ανάλυση της απόδοσής τους, με βάση τη χρονική περίοδο κατά την οποία αναρτήθηκαν
- **Αριθμός προβολών βίντεο (View count):** πεδίο που συνέβαλε στην ανάλυση της επιρροής ενός βίντεο, με βάση τον αριθμό προβολών που έχει συγκεντρώσει
- **Διάρκεια βίντεο (View duration):** πεδίο που χρησιμοποιήθηκε για την μελέτη της απή-

χησης που έχει ένα βίντεο στο κοινό, με γνώμονα τη χρονική του διάρκεια

- **Αριθμός αντιδράσεων “Μου αρέσει” του βίντεο (Like count):** πεδίο που βοήθησε στην ανάλυση της απόδοσης ενός βίντεο, με βάση των αριθμό των θετικών αντιδράσεων που απέσπασε
- **Αριθμός αντιδράσεων “Δεν μου αρέσει” του βίντεο (Dislike count):** πεδίο που βοήθησε στην ανάλυση της απόδοσης ενός βίντεο, με βάση των αριθμό των αρνητικών αντιδράσεων που απέσπασε
- **Αριθμός σχολίων του βίντεο (Comment count):** πεδίο που βοήθησε στην ανάλυση της απόδοσης ενός βίντεο, με βάση τον αριθμό σχολίων που απέσπασε
- **Τίτλος του βίντεο (Video title):** πεδίο που βοήθησε στην αναγνώριση ενός βίντεο στη βάση δεδομένων με βάση το επίσημο όνομά του, καθώς το αναγνωριστικό του, αν και μοναδικό, δεν είναι ευανάγνωστο από τον χρήστη

Ωστόσο, παρόμοια με το Instagram, αρκετά ήταν τα δεδομένα τα οποία παρουσίαζαν ιδιαίτερο ενδιαφέρον αλλά ήταν ανέφικτη η συλλογή τους λόγω των μέτρων διαφύλαξης των προσωπικών δεδομένων που εφαρμόζει το YouTube, όπως ο αριθμός κοινοποιήσεων ή ο μέσος χρόνος παρακολούθησης ενός βίντεο, τα οποία είναι σημαντικά στοιχεία για τον αλγόριθμο προώθησης του YouTube [40].

Εφόσον συλλέχθηκαν και μελετήθηκαν τα απαραίτητα πρωτογενή στοιχεία από το YouTube, συγκροτήθηκε η βάση για τη δημιουργία νέων, σύνθετων μετρικών που θα μπορούσαν να επεκτείνουν το σύνολο των τελικών αποτελεσμάτων με πολυδιάστατα και λιγότερο εμφανή συμπεράσματα για τη φύση της ελληνικής YouTube κοινότητας. Ο λόγος γίνεται για τα δευτερογενή στοιχεία της μελέτης του YouTube, των οποίων η προέλευση αναλύεται παρακάτω (υπενθυμίζεται πως το σύνολο των βίντεο αφορά τη χρονική περίοδο Ιανουάριος 2020-Ιανουάριος 2021):

- **Video Dislikes Ratio:**

$$\frac{\text{Αρνητικές αντιδράσεις ενός βίντεο}}{\text{Θετικές} + \text{Αρνητικές αντιδράσεις ενός βίντεο}}$$

Μεταβλητή η οποία βοηθάει στην αποτύπωση του αρνητικού αντίκτυπου του περιεχομένου ενός βίντεο προς το κοινό, μετρώντας το ποσοστό των αρνητικών αντιδράσεων σε σχέση με τον συνολικό αριθμό των αντιδράσεων. Η συγκεκριμένη μεταβλητή μπορεί να καταλάβει τιμές από μηδέν, στις περιπτώσεις βίντεο με μηδενικές αρνητικές αντιδράσεις, έως 100, αντιπροσωπεύοντας βίντεο με 100% αρνητικές αντιδράσεις.

- **Video Engagement:**

$$\frac{\text{Αριθμός θετικών αντιδράσεων ενός βίντεο} + \text{Αριθμός σχολίων ενός βίντεο}}{\text{Προβολές ενός βίντεο}}$$

Μεταβλητή η οποία συμβάλει στην αποτύπωση του ποσοστού αλληλεπίδρασης των θεατών ενός βίντεο με αυτό, συγκρίνοντας το σύνολο των θετικών αντιδράσεων και σχολιασμών με τον αριθμό των προβολών του βίντεο. Η συγκεκριμένη μεταβλητή μπορεί να καταλάβει τιμές από μηδέν, στις περιπτώσεις βίντεο με μηδενικές θετικές αντιδράσεις και σχολιασμούς, ενώ δεν είναι δυνατό να καθοριστεί το ανώτατο όριο τιμών, καθώς αρκετές

είναι είναι τα βίντεο που συγκεντρώνουν αριθμό θετικών αντιδράσεων και σχολιασμών μεγαλύτερο από εκείνον των προβολών τους, όπως συμβαίνει με τα βίντεο διαγωνισμών.

- **Views Subscribers Ratio:**

$$\frac{\text{Συνολικές προβολές ενός βίντεο}}{\text{Αριθμός συνδρομητών του καναλιού που δημοσίευσε το βίντεο}}$$

Μεταβλητή η οποία συμβάλλει στο σχηματισμό μιας εικόνας για το ποσοστό των συνδρομητών ενός καναλιού που παρακολουθούν τα βίντεο του κατά μέσο όρο. Η συγκεκριμένη μεταβλητή μπορεί να καταλάβει τιμές από μηδέν, στις περιπτώσεις βίντεο με μηδενικό αριθμό προβολών, ενώ δεν είναι δυνατό να καθοριστεί το ανώτατο όριο τιμών, καθώς αρκετά είναι τα βίντεο που συγκεντρώνουν αριθμό προβολών μεγαλύτερο από εκείνον των προβολών τους, όπως συμβαίνει με βίντεο που γίνονται ευρέως γνωστά.

4.2.2 Κώδικας YouTube

Μέσω της μελέτης και της ανάλυσης των διαθέσιμων από την πλατφόρμα του YouTube δεδομένων δημιουργήθηκε το υπόβαθρο πάνω στο οποίο θα στηριχθεί η λεπτομερέστερη περιγραφή της εσωτερικής δομής του μηχανισμού συλλογής δεδομένων που σχεδιάστηκε. Στην παρούσα υποενότητα αναπτύσσονται τα τμήματα κώδικα τα οποία συνέβαλαν τόσο στην υποστήριξη όσο και στη υλοποίηση των βασικών λειτουργιών του μηχανισμού που απαιτούνταν για τη συγκομιδή πληροφοριών του YouTube.

Αρχείο `settings.py`

Πρωταρχικό βήμα για τη διασφάλιση της ορθής λειτουργίας του μηχανισμού αποτέλεσε η διαδικασία παραμετροποίησης του, μέσω του αρχείου `settings.py`. Σε αντίθεση με την περίπτωση του Instagram, το γεγονός ότι κατά τη συλλογή των δεδομένων από το YouTube χρησιμοποιήθηκε το εξουσιοδοτημένο εργαλείο YouTube Data API δε δημιούργησε την ανάγκη για την αποφυγή εντοπισμού του μηχανισμού από τις δικλίδες ασφαλείας της πλατφόρμας, με αποτέλεσμα το σύνολο των ρυθμίσεων να περιοριστεί στις απολύτως απαραίτητες.

Τηρώντας την τυπική αλληλουχία ρυθμίσεων των Scrapy μηχανισμών, πρώτες συναντούνται οι εξής παράμετροι:

```
BOT_NAME = 'channel'

SPIDER_MODULES = ['yt.spiders.channel']
NEWSPIDER_MODULE = 'yt.spiders'
```

Βάσει των παραπάνω ρυθμίσεων, ορίζεται η ονομασία του πρότζεκτ (`BOT_NAME`), καθώς και τα τμήματα του, στα οποία το Scrapy θα αναζητά υπάρχοντα `spiders`, καθώς και θα τοποθετεί τα νεοσύστατα από αυτά αρχεία (`SPIDER_MODULES` και `NEWSPIDER_MODULE` αντίστοιχα). Στην προκειμένη περίπτωση, υπαγορεύεται στο Scrapy να αναγνωρίζει μόνο τον μηχανισμό συγκομιδής `channel.py` (`yt.spiders.channel`) και να αποθηκεύει τους νεοσύστατους μηχανισμούς στον φάκελο `spiders` (`yt.spiders`).

Στη συνέχεια, το σύνολο των ρυθμίσεων που ακολουθούν αφορούν στην υπακοή του μηχανι-

σμού στο αρχείο *robots.txt* του YouTube (*ROBOTSTXT_OBEY*) και στον καθορισμό του αριθμού των διαδοχικών αιτημάτων που θα αποστέλλονται μαζικά προς το YouTube Data API (*CONCURRENT_REQUESTS*).

```
ROBOTSTXT_OBEY = True
CONCURRENT_REQUESTS = 28
```

Τέλος, διευθετείται η πορεία την οποία θα ακολουθήσουν τα συλλεγμένα δεδομένα, ενεργοποιώντας τη δρομολόγηση τους προς το *DatabasePipeline*, ενός από τα Item Pipelines του μηχανισμού (βλ. Υποενότητα 3.1.2). Η χρήση των Middlewares δεν κρίθηκε απαραίτητη, με αποτέλεσμα αυτά να μην συμπεριληφθούν στην υλοποίηση. Οι αριθμοί που παρατηρούνται δίπλα από κάθε Pipeline/Middleware αντιπροσωπεύουν τον βαθμό προτεραιότητας των δομικών στοιχείων, όσον αφορά στη σειρά με την οποία προσπελούνται από τα εξαγόμενα δεδομένα. Για παράδειγμα, αν ορίζονταν δύο Item Pipelines, το *yt.pipelines.Pipeline1:300* και το *yt.pipelines.Pipeline2:305*, τότε τα δεδομένα που θα συλλέγονταν από την εκάστοτε ιστοσελίδα θα δρομολογούνταν από το *Engine* πρώτα προς το Pipelines1 και διαδοχικά προς το Pipeline2.

```
ITEM_PIPELINES = {
    'yt.pipelines.DatabasePipeline': 300
}

# SPIDER_MIDDLEWARES = {
#     'yt.middlewares.YtSpiderMiddleware': 543,
# }

# DOWNLOADER_MIDDLEWARES = {
#     'yt.middlewares.YtDownloaderMiddleware': 543,
# }
```

Αρχείο *items.py*

Καθώς ο όγκος των δεδομένων που προορίζονταν για συλλογή από το YouTube ήταν αρκετά μεγάλος, υπήρχε η ανάγκη για την οργάνωσή τους, με σκοπό την ευκολότερη κατανόηση, διαχείριση και αποθήκευσή τους. Στο έργο αυτό πρωταγωνιστικό ρόλο κατείχε το αρχείο *items.py*, μέσω του οποίου προσδιορίστηκε η ιεραρχική δομή των εξαγόμενων στοιχείων.

```
class ChannelItem(scrapy.Item):
    channel_data = scrapy.Field()
    videos = scrapy.Field()

class ChannelData(scrapy.Item):
    title = scrapy.Field()
    id = scrapy.Field()
    creation_date = scrapy.Field()
    view_count = scrapy.Field()
    subscriber_count = scrapy.Field()
```

```

video_count = scrapy.Field()
keywords = scrapy.Field()
owned_by = scrapy.Field()

class VideoData(scrapy.Item):
    video_id = scrapy.Field()
    video_title = scrapy.Field()
    upload_date = scrapy.Field()
    duration = scrapy.Field()
    view_count = scrapy.Field()
    like_count = scrapy.Field()
    dislike_count = scrapy.Field()
    comment_count = scrapy.Field()
    video_category = scrapy.Field()
    tags = scrapy.Field()

```

Αναλυτικότερα, κύρια μονάδα αποθήκευσης αποτέλεσε η δομή *ChannelItem*, στην οποία περικλείεται ένα συνονθύλευμα όλων των επιθυμητών πληροφοριών ενός λογαριασμού του YouTube. Τα περικλειόμενα δεδομένα με τη σειρά τους διαχωρίζονται σε δύο επιμέρους κατηγορίες, το σύνολο των γενικών πληροφοριών του λογαριασμού (*ChannelData()*) και το σύνολο των πληροφοριών του εκάστοτε αναρτημένου βίντεο (*VideoData()*). Ως εκ τούτου, κατά τη συγκομιδή δεδομένων ενός καναλιού, εφόσον συμπληρωθούν όλα τα πεδία των παραπάνω δομικών στοιχείων, αυτά ενσωματώνονται στη κοινή δομή *ChannelItem* και αποθηκεύονται στη βάση δεδομένων υπό τη μορφή εγγράφου. Αξίζει να επισημανθεί πως υπάρχει πλήρης αντιστοιχία μεταξύ των πεδίων των δομών δεδομένων και των πληροφοριών ενός λογαριασμού, που επιλέχθηκε να εξαχθούν στην Υποενότητα 4.2.1 Παρακάτω παρατηρείται η διαστρωμάτωση μιας δομής δεδομένων, η οποία αντιστοιχεί σε έναν λογαριασμό του YouTube.

```

ChannelItem
├── channel_data
│   └── ChannelData
├── videos[]
│   └── VideoData

```

Αρχείο *pipelines.py*

Η μεταφορά των συγκεντρωμένων δεδομένων στη βάση δεδομένων, με σκοπό την αποθήκευσή τους, επιτεύχθηκε μέσω των Item Pipelines του μηχανισμού και, συγκεκριμένα, με τη χρήση του αρχείου *pipelines.py*.

```

#db_address = address of database

from pymongo import MongoClient

class DatabasePipeline:

    myClient=MongoClient(<db_adress>)
    db=myClient['Thesis']

```

```

collection=db[ 'YouTubeChannels' ]

def process_item(self, item, spider):
    channel_id = item[ 'channel_data' ][ 'id' ]

    self.collection.replace_one(
        { 'channel_data.id': channel_id },
        item, upsert = True)

```

Όπως περιγράφηκε στην Υποενότητα 3.1.2, οι πληροφορίες της κάθε ιστοσελίδας, αφού εξαχθούν από το *Engine* του μηχανισμού συγκομιδής δεδομένων, αποστέλλονται στα *Item Pipelines*, απ' όπου και αποθηκεύονται στην επιθυμητή δομή δεδομένων. Σεβόμενοι την τακτική αυτή του Scrapy, ορίστηκαν οι συλλεγμένες πληροφορίες για το κάθε κανάλι να οδηγούνται στο Pipeline με όνομα DatabasePipeline, όπου εγκαθιδρύεται η σύνδεση με την βάση και οι συλλογές αποθηκεύονται σε αυτή υπό την μορφή εγγράφων. Όλα τα έγγραφα που αφορούν στην πλατφόρμα του YouTube στεγάζονται στη συλλογή με το όνομα *YouTubeChannels*, η οποία με τη σειρά της είναι μέρος της βάσης δεδομένων *Thesis*. Η βάση *Thesis* φιλοξενεί όλα τα στοιχεία της παρούσας έρευνας. Στην περίπτωση κατά την οποία τα στοιχεία για κάποιο κανάλι είναι ήδη αποθηκευμένα στην βάση, τότε αυτά υπερεγγράφονται μέσω της μεθόδου `self.Collection.replace_one()`.

Φάκελος resources

Αντίστοιχα με την περίπτωση του Instagram, θεωρήθηκε σκόπιμο λίστες δεδομένων, οι οποίες καταλάμβαναν μεγάλο όγκο στο αρχείο του spider και αποτελούσαν σημαντικό πόρο για την λειτουργία του, να αποσπαστούν από το συγκεκριμένο αρχείο και να καταλάβουν αυτόνομη θέση σε έναν νέο φάκελο με το όνομα *resources*. Πιο συγκεκριμένα, τα τρία βασικά αρχεία που απαρτίζουν τον φάκελο αυτό είναι τα: (i) *APIkeys.json*, (ii) *CategoryIDs.json* και (iii) *BusinessChannels.json*.

Κατά τη περιγραφή των χαρακτηριστικών που διέπουν το YouTube Data API (βλ. Ενότητα 3.2) επισημάνθηκε πως ο μέγιστος αριθμός εκμεταλλεύσιμων μονάδων ανερχόταν στις 10.000 ημερησίως, γεγονός το οποίο περιόριζε σημαντικά την αποδοτικότητα του μηχανισμού, καθώς ο όγκος των, προς συγκέντρωση, δεδομένων ήταν αρκετά μεγάλος, σε αντίθεση με τη ποσότητα των στοιχείων που ήταν δυνατό να αποκτηθούν αξιοποιώντας τις παραπάνω μονάδες. Η ανάγκη για τη ελαχιστοποίηση του χρονικού κόστους της διαδικασίας σε συνδυασμό με τα προαναφερθέντα, παρότρυνε τη χρήση πολλαπλών κλειδιών API και, κατά συνέπεια, τη δημιουργία πολλαπλών λογαριασμών Google, διότι στον κάθε λογαριασμό παρέχεται το δικαίωμα χρήσης ενός μόνο κλειδιού API. Με σκοπό τη δημιουργία ενός ψηφιακού αποθετηρίου κλειδιών API και τη διευκόλυνση της δυναμικής εναλλαγής κλειδιών κατά τη διαδικασία συλλογής στοιχείων, συντάχθηκε το αρχείο *API keys.json*. Το συγκεκριμένο αρχείο περιλαμβάνει ένα σύνολο από ζεύγη κλειδιού API-ονόματος λογαριασμού σε μορφή JSON, με την υποδειγματική μορφή του να είναι η εξής:

```

{
  "Keys": [
    {
      "Name": "Profile1",

```

```

    "Id" : "AIzaSyDZl1myWH7bFq1xxZHHvatMkdacjgi9mp8"
  },
  {
    "Name" : "Profile2",
    "Id" : "AIzaSyCVBhNVKm2yuIonEAOrOzJyC9nmaDJVqa0"
  },
  .
  .
  .
  {
    "Name" : "Profile(N-1)",
    "Id" : "AIzaSyDXBmHwOfxOGINj-uknDHZA74rz1ECzvBQ"
  },
  {
    "Name" : "Profile(N)",
    "Id" : "AIzaSyC2lZln_iYsabulLEVk5Y_x8o07ymTv560"
  }
]
}

```

Σημαντική πτυχή της έρευνας αποτέλεσε η μελέτη της ανταπόκρισης του κοινού απέναντι σε συγκεκριμένες κατηγορίες βίντεο του YouTube και για αυτό το λόγο αποφασίστηκε η συγκομιδή της θεματικής κατηγορίας του κάθε βίντεο. Παρόλα αυτά, το YouTube Data API εφαρμόζει κωδικοποίηση στο συγκεκριμένο πεδίο, αντιστοιχίζοντας τα ονόματα των κατηγοριών με ακέραιους αριθμούς. Η επιθυμία για αποκωδικοποίηση των κωδικών αριθμών και την αποθήκευση στη βάση δεδομένων της θεματικής κατηγορίας κάθε βίντεο σε αλφαριθμητική μορφή, οδήγησε στη δημιουργία του αρχείου *CategoryIDs.json*. Το συγκεκριμένο αρχείο περιέχει ένα σύνολο με όλα τα ζεύγη θεματικών κατηγοριών-κωδικών αριθμών που χρησιμοποιεί το YouTube Data API, έτσι ώστε να επιτευχθεί η αντιστοίχιση μεταξύ αυτών των δύο κατά τη διαδικασία της συγκέντρωσης δεδομένων, και φέρει την εξής δομή:

```

{
  "1" : "Film & Animation",
  "2" : "Autos & Vehicles",
  .
  .
  .
  "43" : "Shows",
  "44" : "Trailers"
}

```

Τέλος, το αρχείο *BusinessChannels.json* εμπεριέχει έναν πίνακα με ζεύγη ονόματος-αναγνωριστικού που αντιστοιχούν σε κανάλια του YouTube ελληνικών επιχειρήσεων, των οποίων το περιεχόμενο αξιοποιήθηκε στα πλαίσια της έρευνας. Παρόμοια με το Instagram, πηγή των δεδομένων αποτέλεσε η ιστοσελίδα *socialbakers.com*¹⁵, από την οποία αποσπάστηκαν χειροκίνητα τα ονόματα 336 επιχειρησιακών καναλιών και, στη συνέχεια, αποθηκεύτηκαν στο αρχείο *Business*

¹⁵<https://www.socialbakers.com/statistics/youtube/channels/greece>

Channels.json συνοδευόμενα από το αναγνωριστικό τους. Η δομή του προκείμενου αρχείου παρουσιάζεται παρακάτω:

```
{
  "Channels": [
    {
      "Name": "jumbostores",
      "Id": "UCmkhcBmnTelmvUCKzy5VFJg"
    },
    {
      "Name": "efood",
      "Id": "UCU7HP8QwDDM85NADD-exRsa"
    },
    .
    .
    .
    {
      "Name": "aktaia-sea-crystals-deli-foods",
      "Id": "UC_UJLy2eBF27R1qmCQWhNUw"
    },
    {
      "Name": "avra-travel",
      "Id": "UCYZOhwwqTc6tqOESfq-XhDA"
    }
  ]
}
```

Φάκελος *tools*

Παρόμοιο σκοπό με εκείνον του φακέλου *resources* κλήθηκε να εξυπηρετήσει ο φάκελος *tools*. Τμήματα κώδικα, τα οποία θα μπορούσαν να επαναχρησιμοποιηθούν σε αρκετά σημεία της έρευνας, κρίθηκε προτιμότερο να συγκεντρωθούν σε έναν ξεχωριστό φάκελο, έτσι ώστε να εξοικονομηθεί χώρος στο βασικό κώδικα του μηχανισμού συγκομιδής δεδομένων και να επιτευχθεί η μέγιστη δυνατή ομαδοποίηση των διάφορων λειτουργιών. Η κύρια λειτουργία που υποστηρίχθηκε, κατά τη διαδικασία της συλλογής δεδομένων, από τον φάκελο *tools* ήταν η επικοινωνία με τη βάση δεδομένων και η διεξαγωγή των επιθυμητών αιτημάτων προς αυτήν με αντιπροσωπευτικά παραδείγματα να είναι το αρχεία *documents_finder.py* και *documents_deleter.py*.

Δεδομένου ότι ο ρόλος των παραπάνω αρχείων αφορά στη διεκπεραίωση διαδικασιών στη βάση δεδομένων, το αρχικό βήμα που εφαρμόστηκε και στα δύο αρχεία ήταν η εγκαθίδρυση σύνδεσης με τη βάση, μέσω του εξής τμήματος κώδικα:

```
#db_address = address of database
#collection_name = name of the database collection

myClient = MongoClient(<db_address>)
db = myClient['Thesis']
collection=db[<collection_name>]
```

Το αρχείο *documents_finder.py* φιλοξενεί τη μέθοδο *findDocuments*, η οποία αφορά στην εύρεση και συλλογή επιλεγμένων πεδίων της βάσης δεδομένων και, ειδικότερα, επιτρέπει τη σύνθεση και την αποστολή παντός τύπου ερωτημάτων *find* στη βάση. Η συγκεκριμένη μέθοδος δέχεται ως ορίσματα: (i) την παράμετρο φιλτραρίσματος (*filter*), (ii) την παράμετρο προβολής (*project*), (iii) τον αριθμό παραλειπομένων εγγράφων (*skip_num*) και (iv) τον αριθμό των επιστρεφόμενων εγγράφων (*limit_num*), τα οποία συναρμολογούνται σε ένα κοινό ερώτημα *find* και αποστέλλονται στη βάση δεδομένων. Τελικά, τα αποτελέσματα της αναζήτησης επιστρέφονται μέσω της εντολής *return*.

```
def findDocuments(self, filter, project, skip_num, limit_num):
    if skip_num != None:
        if limit_num != None:
            if project != None:
                return self.collection.find(filter,
                    project).skip(skip_num).limit(limit_num)
            return
                self.collection.find(filter).skip(skip_num).limit(limit_num)
        else:
            if project != None:
                return self.collection.find(filter,
                    project).skip(skip_num)
            return self.collection.find(filter).skip(skip_num)
    else:
        if limit_num != None:
            if project != None:
                return self.collection.find(filter,
                    project).limit(limit_num)
            return self.collection.find(filter).limit(limit_num)
        return self.collection.find(filter)
```

Παρομοίως, το αρχείο *documents_deleter.py* σχεδιάστηκε με σκοπό τη στέγαση της μεθόδου *deleteDocument*, η οποία είναι υπεύθυνη για τη διαγραφή επιλεγμένων εγγράφων από τη βάση δεδομένων. Η συγκεκριμένη μέθοδος δέχεται ως μοναδικό όρισμα το αναγνωριστικό του προς διαγραφή καναλιού (*id*), το οποίο εντοπίζεται και, εν τέλει, διαγράφεται οριστικά από τη βάση.

```
def deleteDocument(self, id):
    self.collection.delete_one({"channel_data.id": id})
```

Συγκέντρωση ονομάτων καναλιών

Στη συνέχεια της διαδικασίας συγκέντρωσης του ερευνητικού υλικού, αναφορικά με το YouTube, ακολούθησε η αναζήτηση των λογαριασμών που πληρούσαν τις απαιτήσεις της έρευνας και οι οποίοι, τελικά, θα δημιουργούσαν μία ολοκληρωμένη βάση για τη συλλογή πληροφοριών. Με αφορμή την πρόκληση που προέκυψε κατά τη διαδικασία ανακάλυψης νέων λογαριασμών (βλ. Ενότητα 4.3), το δείγμα της μελέτης στηρίχθηκε, κατά το πλείστον, στον ιστότοπο *stats.video*¹⁴, ο

¹⁴<https://stats.video/top/most-subscribed/youtube-channels/greece/of-all-time>

οποίος φιλοξενεί λίστες με τα δημοφιλέστερα ελληνικά κανάλια της πλατφόρμας του YouTube, ανάμεσα σε μία πληθώρα θεματικών κατηγοριών περιεχομένου. Η συγκεκριμένη ιστοσελίδα για κάθε θεματική κατηγορία παρέχει έναν πίνακα με ονόματα καναλιών καταταγμένα με βάση τον αριθμό των εγγεγραμμένων χρηστών που κατέχουν. Παρά την ποικιλία θεματικού περιεχομένου, η επιλογή των λογαριασμών βασίστηκε στον συγκεντρωτικό πίνακα καναλιών από όλο το φάσμα θεματικών κατηγοριών που διατίθεται, τα περιεχόμενα του οποίου είναι προσβάσιμα μέσω του συνδέσμου που φαίνεται ακολούθως:

```
https://stats.video/top/most-subscribed/youtube-channels/greece/of-all-time
```

Το γεγονός ότι το μεγάλο πλήθος λογαριασμών δεν είναι δυνατό να αναπαρασταθεί στα πλαίσια της μιας σελίδας, οδήγησε τον ιστότοπο *stats.video* στη σελιδοποίηση των περιεχομένων του παραπάνω πίνακα ανά 48 κανάλια, με την κάθε σελίδα να είναι προσβάσιμη μέσω του εξής συνδέσμου:

```
https://stats.video/top/most-subscribed/youtube-channels/greece/of-all-time/page/<page_number>
```

Η σύνταξη του κώδικα για το μηχανισμό συλλογής ονομάτων από την ιστοσελίδα *stats.video* ξεκινάει με την ανάθεση τιμών στις απαραίτητες μεταβλητές στιγμιοτύπου.

```
from scrapy import Spider, Request
import re
from ..items import ChannelItem

class Statsvideo(Spider):

    name = 'StatsvideoSpider'
    allowed_domains = ['stats.video']
    website_url: str =
        'https://stats.video/top/most-subscribed/youtube-channels/
        greece/of-all-time'
    visited_links = {"of-all-time"}
```

Όπως γίνεται αντιληπτό από το παραπάνω τμήμα κώδικα, σε πρώτο επίπεδο δίνεται όνομα στον μηχανισμό συγκομιδής δεδομένων (*name*), δηλώνεται ο ιστότοπος από τον οποίο θα συλλεχθούν τα δεδομένα (*allowed_domains*) και παρατίθεται ο σύνδεσμος με τον οποίο επιτυγχάνεται η επίσκεψη σε αυτόν (*website_url*). Έπειτα, αρχικοποιείται η λίστα *visited_links* με την τιμή *of-all-time*, η οποία αντιπροσωπεύει το αναγνωριστικό της αρχικής σελίδας του συγκεντρωτικού πίνακα καναλιών. Η χρήση της λίστας *visited_links* αποσκοπεί στην αποτροπή πολλαπλών επισκέψεων στις σελίδες του πίνακα, μέσω της διατήρησης ενός ιστορικού προσπέλασης αυτών.

Αφού ολοκληρωθεί η διαδικασία αρχικοποίησης, η αναζήτηση λογαριασμών ξεκινά από την αρχική σελίδα του συγκεντρωτικού πίνακα καναλιών με τη βοήθεια της μεθόδου *start_requests*.

```
def start_requests(self):
    yield Request(self.website_url, callback=self.parse)
```

Όπως είναι φανερό από τη μέθοδο *start_requests*, η απόκριση για τη πρώτη σελίδα του πίνακα μεταβιβάζεται στη μέθοδο *parse()*, με σκοπό την επεξεργασία της (*callback=self.parse*). Στο σημείο αυτό, εκκινείται η διαδοχική προσπέλαση όλων των σελίδων του πίνακα και η απόσπαση των χρήσιμων πληροφοριών από αυτές. Εξαιτίας του γεγονότος ότι οι καταχωρήσεις του πίνακα περιλαμβάνουν μόνο το όνομα του καναλιού και των αριθμών των συνδρομητών του, αλλά όχι το αναγνωριστικό του, το οποίο αποτελεί αναγκαίο πεδίο για τον εντοπισμό του περιεχομένου του καναλιού στη πλατφόρμα του YouTube, η διαδικασία αναζήτησης διακρίθηκε σε δύο στάδια.

Το πρώτο στάδιο είναι εκείνο της προσπέλασης του πίνακα, μέσω της μεθόδου *parse()*, και της διαλογής των ονομάτων των λογαριασμών με αριθμό συνδρομητών μεγαλύτερο του 5.000, ενώ το δεύτερο είναι εκείνο της αναζήτησης του αναγνωριστικού του εκάστοτε επιλεγμένου καναλιού στο πλαίσιο της εξατομικευμένης σελίδας, η οποία του παραχωρείται από την ιστοσελίδα <https://stats.video>, με τη χρήση της μεθόδου *parse_profile()*.

```
def parse(self, response):
    array_fields=response.xpath("//tbody//tr").extract()
    for field in array_fields:
        profile_link=re.findall(r'(?<=<a class="channelLink"
            href=").*(?=<">)', field)[0]

        pattern = "<a class="channelLink"
            href="\""+profile_link+"\">(.*?)</a>"
        channel_name = re.findall(pattern, field)[-1]
        subscriber_count =
            int(re.findall(r'(?<=title="Subscribers"></i>\>\n
                ).*(?=</button>)', field)[0].replace(',',' '))
        if subscriber_count>5000:
            yield Request(f"https://stats.video{profile_link}",
                callback=self.parse_profile,meta={"channel_name":channel_name})

    page_links=response.xpath("//a[@class='btn btn-danger purple
        m-1']//@href").extract()
    for link in page_links:
        pattern="^[^/]+$"
        page_index=re.findall(pattern, str(link))[0]
        if not page_index in self.visited_links:
            self.visited_links.add(page_index)
            yield Request(f"{self.website_url}/page/{page_index}",
                callback=self.parse)

def parse_profile(self, response):
    channel_id=response.xpath("//div[@data-original-title="YouTube
        Channel's ID"]>text()").extract()[0].strip()
    channel_item=ChannelItem()
    channel_item["channel_id"]=channel_id
```

```
channel_item["channel_name"]=response.meta["channel_name"]

yield channel_item
```

Αναλυτικότερα, κατά την επίσκεψη στην κάθε σελίδα, πρωταρχικό μέλημα αποτελεί η απόσπαση της δομής πίνακα που περικλείει τα στοιχεία των λογαριασμών του YouTube (*array_fields*). Έχοντας αποκτήσει πρόσβαση στον πίνακα στοιχείων, ξεκινά η προσπέλαση των πεδίων του, με στόχο την απόσπαση του ονόματος του εκάστοτε καναλιού (*channel_name*) και του αριθμού των συνδρομητών του (*subscriber_count*). Σε περίπτωση που το υπό εξέταση κανάλι κατέχει περισσότερους από 5000 εγγεγραμμένους χρήστες, το ενδιαφέρον στρέφεται στην εύρεση του αναγνωριστικού του, το οποίο μπορεί να εντοπιστεί στην εξατομικευμένη σελίδα του λογαριασμού, ακολουθώντας τον επισυναπτόμενο, στο αντίστοιχο πεδίο του πίνακα, σύνδεσμο (*profile_link*). Η διαδικασία αυτή υλοποιείται στη μέθοδο *parse_profile()*, κατά την οποία, αφότου συλλεχθεί το αναγνωριστικό του καναλιού (*channel_id*), δημιουργείται μια δομή δεδομένων με όλα τα μέχρι τότε συγκεντρωμένα στοιχεία του καναλιού (*channel_item*) και αποστέλλεται στην βάση δεδομένων προς αποθήκευση (*yield channel_item*). Παράλληλα, η δραστηριότητα της μεθόδου *parse()* παραμένει ενεργή, καθώς ακολουθεί η εύρεση όλων των συνδέσμων προς γειτονικές σελίδες και η αποστολή αιτήματος πρόσβασης σε αυτές που δεν έχουν εξεταστεί ακόμα, επαναλαμβάνοντας έτσι όλη την διαδικασία για κάθε μία από αυτές.

Αρχείο `channel.py`

Έχοντας συμπληρώσει τη βάση δεδομένων με τα ζεύγη ονόματος-αναγνωριστικού ενός ικανοποιητικού συνόλου λογαριασμών του YouTube, διατίθονταν όλοι οι απαιτούμενοι πόροι για την έναρξη της συγκομιδής των πληροφοριών που κρίθηκαν καθοριστικές για την επιτυχή διεκπεραίωση της έρευνας, όσον αφορά στην ελληνική κοινότητα του YouTube (βλ. Ενότητα 4.2.1). Το σύνολο των λειτουργιών που κλήθηκαν να φέρουν εις πέρας την συγκεκριμένη διαδικασία συγκεντρώνονται στο αρχείο *channel.py*, το οποίο συνιστά την υλοποίηση του μηχανισμού συγκομιδής δεδομένων από την πλατφόρμα του YouTube.

Στην ορθή λειτουργία των διαδικασιών εξαγωγής δεδομένων συνέβαλε η χρήση βιβλιοθηκών και μεταβλητών, στις οποίες βρίσκονταν καταχωρημένες θεμελιώδεις πληροφορίες. Συνεπώς, η σύνταξη του μηχανισμού ξεκινάει με τη συμπερίληψη των απαραίτητων βιβλιοθηκών και την αρχικοποίηση των μεταβλητών στιγμιοτύπου.

```
from scrapy import Spider, Request
from scrapy.exceptions import CloseSpider
from ..items import ChannelItem
from ..items import ChannelData
from ..items import VideoData
from ..tools.documents_finder import DocumentsFinder
from ..tools.documents_deleter import DocumentsDeleter
import isodate
import json

class YoutubeSpider(Spider):
```

```

name = 'youtubescraper'
allowed_domains = ['googleapis.com']
youtube_url = 'https://www.googleapis.com/youtube/v3'

handle_httpstatus_list = [403]

API_keys = {}
API_keys_file = ``<file_directory>'
with open(API_keys_file, encoding='utf8') as json_file:
    data = json.load(json_file)
    for key in data["Keys"]:
        API_keys[key["Id"]]="active"

subs_limit = 5000
channel_items = {}
channels_to_scrape = []
for channel in
    list(DocumentsFinder.findDocuments(DocumentsFinder, {},
    None, 0, 1)):
        channels_to_scrape.append(channel["channel_data"]["id"])

video_categories = {}
video_categories_file = ``<file_directory>'
with open(video_categories_file, encoding='utf8') as json_file:
    video_categories = json.load(json_file)

```

Εξετάζοντας το άνωθεν κομμάτι του κώδικα, αρχικά παρατηρείται η τυπική διαδικασία απόδοσης ονόματος στο μηχανισμό συγκομιδής δεδομένων (*name*) και καθορισμού του διαδικτυακού τύπου, από τον οποίο θα αντληθούν πληροφορίες (*allowed_domains*, *youtube_url*).

Όπως αναφέρθηκε στην Παράγραφο 4.2.2, μονόδρομο αποτέλεσε η χρήση πολλαπλών κλειδιών API για την επικοινωνία με το YouTube Data API. Ωστόσο, η αξιοποίηση τους προϋπέθετε επιπλέον τεχνικές για τη διαχείριση τους και, παράλληλα, επιφύλασσε απειλές για την ομαλή λειτουργία του μηχανισμού. Ειδικότερα, αναπτύχθηκε ένα λεξικό (*API_keys dictionary*), στο οποίο εναποτέθηκαν όλα τα κλειδιά API που βρίσκονταν αποθηκευμένα στο αρχείο *APIkeys.json* (*API_keys_file*), συνοδευόμενα από την τιμή *active*, έτσι ώστε να διατηρηθεί μία λίστα με όλα τα ενεργά και τα ανενεργά κλειδιά API· ενεργά ορίζονται τα κλειδιά για τα οποία υπάρχουν διαθέσιμες μονάδες αναζήτησης (βλ. Κεφάλαιο 3, Ενότητα 3.2). Ο πεπερασμένος αριθμός διαθέσιμων αιτημάτων για κάθε κλειδί API αποτέλεσε πηγή σφαλμάτων, καθώς κάθε φορά που εξαντλούνταν τα αιτήματα για κάποιο κλειδί, με αποτέλεσμα αυτό να οδηγείται σε ανενεργή κατάσταση, το YouTube Data API επέστρεφε, ως απάντηση στα αιτήματα, HTTP αποκρίσεις με κωδικό σφάλματος 403, καταστέλλοντας τη λειτουργία του μηχανισμού. Η ανάγκη για την ειδική διαχείριση τέτοιου είδους σφαλμάτων, σε συνδυασμό με το γεγονός ότι το Scrapy είναι προκαθορισμένο να τα διαχειρίζεται αυτόματα, οδήγησε στη χρήση της μεταβλητής *handle_httpstatus_list*, η οποία απαλλάσσει το Scrapy από την ευθύνη της αυτόματης διαχείρισης σφαλμάτων με κωδικό 403.

Έχοντας θέσει ως στόχο τη διερεύνηση καναλιών με περισσότερους των 5.000 συνδρομητών, δημιουργήθηκε η μεταβλητή *subs_limit*, η οποία απέκτησε το ρόλο του φιλτραρίσματος των

προς συγκέντρωση λογαριασμών. Το σύνολο των ονομάτων των λογαριασμών, για τους οποίους θα πραγματοποιούνταν αναζήτηση, αποθηκεύτηκε στη λίστα *channels_to_scrape*, διοχετεύοντας από τη βάση δεδομένων σε αυτή τα χαρακτηριστικά τους. Ο ασύγχρονος τρόπος της συλλογής δεδομένων για κάθε κανάλι, έθετε σε κίνδυνο την ακεραιότητα τους, αυξάνοντας την πιθανότητα σύγχυσης των πληροφοριών. Επομένως, κρίθηκε αναγκαία η αξιοποίηση μιας μεταβλητής στιγμιοτύπου υπό τη μορφή λεξικού (*channel_items dictionary*), η οποία θα διατηρούσε αποθηκευμένες όλες τις υπό συγκέντρωση δομές δεδομένων για κάθε κανάλι έως ότου αυτές να συμπληρωθούν πλήρως, αποτελώντας έτσι ένα κοινό σημείο αναφοράς για όλες τις ανεξάρτητες μεθόδους της κλάσης.

Τελευταία, αλλά εξίσου σημαντική, ήταν η μεταβλητή στιγμιοτύπου *video_categories*, η οποία συνέβαλε στην αντιστοίχιση μεταξύ των κωδικών αριθμών και των ονομάτων των θεματικών κατηγοριών των βίντεο, φορτώνοντας όλα τα ζεύγη κωδικού αριθμού-θεματικής κατηγορίας που διακρίνει το YouTube Data API μέσω του αρχείου *CategoryIDs.json* (βλ. Παράγραφο 4.2.2).

Τη διαδικασία αρχικοποίησης των μεταβλητών στιγμιοτύπου διαδέχεται η έναρξη της αναζήτησης στοιχείων για κάθε ένα από τα επιθυμητά κανάλια με τη χρήση της μεθόδου *start_requests()*.

```
def start_requests(self):
    for channel_id in self.channels_to_scrape:
        API_key = self.get_active_API()
        yield Request(self.get_url(channel_id, "channel",
            API_key), callback=self.parse, meta={"channel_id":
            channel_id, "usedAPI": API_key})
```

Στα πλαίσια της λειτουργίας της μεθόδου *start_requests()* συμμετέχουν οι υποστηρικτικές συναρτήσεις *get_active_API()* και *get_url()*, κάθε μία από τις οποίες επιτελεί ένα ξεχωριστό μερίδιο της διαδικασίας. Αναλυτικότερα, η μέθοδος *get_active_API()* έχει την αρμοδιότητα της εύρεσης και της επιστροφής ενός ενεργού κλειδιού API από τη λίστα *API_keys*, έτσι ώστε αυτό να προσαρτηθεί στο απεσταλμένο προς το YouTube Data API αίτημα. Από την άλλη μεριά, η μέθοδος *get_url()* αναλαμβάνει τη σύνθεση των αιτημάτων, παράγοντας ένα ξεχωριστό αίτημα για κάθε τύπο απαιτούμενου πόρου, ο οποίος εισάγεται σε αυτή υπό τη μορφή παραμέτρου (*return_type*).

```
def get_active_API(self):
    for key in self.API_keys:
        if self.API_keys[key] == "active":
            return key
    raise CloseSpider('All API keys exceeded.')
```

```
def get_url(self, id, return_type, API_key):
    if return_type == "videos_list":
        return <request_for_videos_list>
    elif return_type == "video":
        return <request_for_videos>
    elif return_type == "next_page":
        return <request_for_next_page>
    elif return_type == "channel":
```

```
return <request_for_channel>
```

Επόμενες στην αλυσίδα συγκομιδής συναντούνται μέθοδοι, οι οποίες φέρουν την ευθύνη της μεταχείρισης των αποκρίσεων για κάθε κανάλι, αποσπώντας από αυτές τα ζητούμενα στοιχεία και αποθηκεύοντας τα στη βάση δεδομένων. Ο σχεδιασμός των μεθόδων εκμείωσης δεδομένων επηρεάστηκε από την τακτική οργάνωσης των πληροφοριών εκ μέρους του YouTube Data API, σύμφωνα με την οποία οι διαθέσιμοι πόροι για κάθε λογαριασμό διαμοιράζονται σε τρεις επιμέρους κατηγορίες: (i) τα γενικά στοιχεία ενός λογαριασμού, (ii) τη λίστα των βίντεο του λογαριασμού και (iii) τις αναλυτικές πληροφορίες ενός βίντεο. Με αφορμή το παραπάνω γεγονός, η υλοποίηση της επεξεργασίας των αποκρίσεων από το YouTube Data API διαχωρίστηκε σε τρία επίπεδα: (i) τη μέθοδο *parse()*, (ii) τη μέθοδο *parse_videos()* και (iii) τη μέθοδο *parse_video_stats()*. Παρακάτω, περιγράφεται η λειτουργία για κάθε μία από τις προαναφερθείσες μεθόδους, υπό τη μορφή ψευδοκώδικα, καθώς η έκταση του κώδικα ήταν αποτρεπτική για τη συνεκτικότητα της αναφοράς.

Τα αιτήματα που αποστέλλονται από τη μέθοδο *start_requests()* αφορούν στην πρόσληψη των γενικών πληροφοριών του κάθε καναλιού. Επομένως, ως πρώτο βήμα η μέθοδος *parse()* καλείται να διαχειριστεί αυτές τις πληροφορίες, διαλέγοντας εκείνες που παρουσιάζουν ερευνητικό ενδιαφέρον (βλ. Υποενότητα 4.2.1), και, στη συνέχεια, να αποκτήσει πρόσβαση στη λίστα μεταφορτώσεων του καναλιού, μεταβιβάζοντας την αντίστοιχη απόκριση στη μέθοδο *parse_videos()* για επεξεργασία.

1. Μέθοδος *parse* (απόκριση_αιτήματος):
- 2.
3. Αν απόκριση_αιτήματος περιέχει σφάλμα_403:
4. **Κάλεσε** τη μέθοδο *handle_API_error()*
5. **Αλλιώς:**
6. Αν απόκριση_αιτήματος.συνδρομητές_καναλιού είναι δημόσιο:
7. Αν απόκριση_αιτήματος.συνδρομητές_καναλιού είναι μεγαλύτερο από 5000:
- 8.
9. πληροφορίες_καναλιού=[]
10. γενικές_πληροφορίες_καναλιού=[]
11. βίντεο_καναλιού=[]
- 12.
13. **Πρόσθεσε** τις γενικές_πληροφορίες_καναλιού στις πληροφορίες_καναλιού=[]
14. **Πρόσθεσε** τα βίντεο_καναλιού στις πληροφορίες_καναλιού
15. **Πρόσθεσε** τις πληροφορίες_καναλιού στη μεταβλητή στιγμιότυπου *channel_items*
- 16.
17. **Ενημέρωσε** τις *channel_items.πληροφορίες_καναλιού.γενικές_πληροφορίες_καναλιού* με απόκριση_αιτήματος.γενικές_πληροφορίες_καναλιού
18. **Στείλε** αίτημα για τη λίστα βίντεο του καναλιού
19. **Κάλεσε** τη μέθοδο *parse_videos(απόκριση_σελίδας)*
20. **Αλλιώς:**
21. **Διέγραψε** το κανάλι από τη **βάση δεδομένων**
22. **Αλλιώς:**
23. **Διέγραψε** το κανάλι από τη **βάση δεδομένων**

Τα, απευθυνόμενα προς το YouTube Data API, αιτήματα πρόσβασης στις γενικές πληροφορίες των καναλιών του YouTube συντάχθηκαν ως εξής:

```
https://www.googleapis.com/youtube/v3/channels?key={API_key}
&part=snippet,id,statistics,brandingSettings
&fields=items(brandingSettings/channel/keywords,
snippet/title, snippet/publishedAt,statistics,id)
&id={id}
```

Το YouTube Data API κατακερματίζει το σύνολο των μεταφορτώσεων του εκάστοτε λογαριασμού σε μικρότερα υποσύνολα βίντεο, τις λεγόμενες σελίδες. Κάθε σελίδα περιλαμβάνει μία λίστα με τα αναγνωριστικά ενός πεπερασμένου αριθμού μεταφορτώσεων του καναλιού, χωρίς να παρέχονται εκτενείς πληροφορίες για αυτές, καθώς και δείκτες προς τις επόμενες και τις προηγούμενες σελίδες, αν αυτές υπάρχουν. Συνεπώς, ο ρόλος της μεθόδου *parse_videos()* οριοθετείται στο να προσπελάσει όλες τις διαδοχικές σελίδες μεταφορτώσεων και, για κάθε ένα από τα βίντεο, να αναζητήσει τα αναλυτικά του στοιχεία. Η απόκριση με το σύνολο των δεδομένων του κάθε βίντεο διοχετεύεται, εν τέλει, στη μέθοδο *parse_video_stats()*, με σκοπό την επεξεργασία της.

1. Μέθοδος **parse_videos(απόκριση_σελίδας)**:
- 2.
3. Αν απόκριση περιέχει σφάλμα_403:
4. **Κάλεσε** τη μέθοδο **handle_API_error()**
5. **Αλλιώς**:
6. Αν το κανάλι δεν έχει κανένα βίντεο:
7. **Διέγραψε** το κανάλι από τη **βάση δεδομένων**
8. **Αλλιώς**:
9. **Για κάθε** βίντεο από απόκριση_σελίδας:
10. Αν το βίντεο δεν έχει ήδη συλλεχθεί:
11. **Στείλε** αίτημα για τις πληροφορίες του βίντεο
12. **Κάλεσε** τη μέθοδο **parse_video_stats(απόκριση_βίντεο)**
13. Αν υπάρχει απόκριση_σελίδας.επόμενη_σελίδα:
14. **Στείλε** αίτημα για την επόμενη σελίδα βίντεο
15. **Κάλεσε** τη μέθοδο **parse_videos(απόκριση_επόμενης_σελίδας)**

Τα, απευθυνόμενα προς το YouTube Data API, αιτήματα πρόσβασης στις σελίδες μεταφορτώσεων ενός καναλιού του YouTube συντάχθηκαν ως εξής:

```
https://www.googleapis.com/search?key={API_key}
&channelId={id}
&part=id
&order=date
&type=video
&publishedAfter=2020-01-01T00%3A00%3A00Z
&maxResults=50
```

Στο κατώτερο επίπεδο της αλυσίδας συγκομιδής δεδομένων βρίσκεται η μέθοδος *parse_video_stats()*, στο πλαίσιο της οποίας πραγματοποιείται η απόσπαση των επιθυμητών πληροφοριών για κάθε βίντεο, από τα περιεχόμενα της απόκρισης που το αντιπροσωπεύει. Στο τέλος της προκειμένης διαδικασίας, τα μέχρι στιγμής συλλεγμένα δεδομένα συγκεντρώνονται για το κάθε κανάλι σε

μία κοινή δομή δεδομένων και ενημερώνουν την τρέχουσα κατάσταση του εγγράφου της βάσης δεδομένων, το οποίο αντιστοιχεί στο κανάλι αυτό.

1. Μέθοδος `parse_video_stats`(απόκριση_βίντεο):
- 2.
3. Αν απόκριση περιέχει σφάλμα_403:
4. **Κάλεσε** τη μέθοδο `handle_API_error()`
5. **Αλλιώς:**
6. `πληροφορίες_βίντεο = []`
7. **Ενημέρωσε** τις `πληροφορίες_βίντεο` με `απόκριση_βίντεο.πληροφορίες_βίντεο`
8. **Πρόσθεσε** τις `πληροφορίες_βίντεο` στη `μεταβλητή` `στιγμιότυπου` `channel_items.πληροφορίες_καναλιού.βίντεο_καναλιού`
9. **Ενημέρωσε** το έγγραφο του καναλιού στη **βάση δεδομένων** με τις `channel_items.πληροφορίες_καναλιού`

Τα, απευθυνόμενα προς το YouTube Data API, αιτήματα πρόσβασης στις πληροφορίες μιας μεταφόρτωσης ενός καναλιού του YouTube συντάχθηκαν ως εξής:

```
https://www.googleapis.com/videos?key={API_key}
&id={id}
&part=statistics,id,snippet,contentDetails
&fields=items(id,snippet/publishedAt,snippet/title,
snippet/categoryId,snippet/tags,contentDetails/duration,
contentDetails/licensedContent,contentDetails/caption,
statistics)
```

Στην ανάλυση των παραπάνω μεθόδων παρατηρείται η συμμετοχή της συνάρτησης `handle_API_error()`. Όπως προαναφέρθηκε, η χρήση πολλαπλών κλειδιών API εξέθεσε τον μηχανισμό σε σφάλματα που προκαλούνταν κατά την εξάντληση των αιτημάτων κάποιου από τα διαθέσιμα κλειδιά. Λύση στο συγκεκριμένο πρόβλημα αποτέλεσε η σύνταξη μιας μεθόδου, η οποία κάθε φορά που τα αιτήματα ενός κλειδιού API εξαντλούνταν, θα ανανέωνε τη κατάσταση του συγκεκριμένου κλειδιού σε ανενεργή και θα επαναλάμβανε τη διαδικασία, κατά την οποία είχε προκύψει το σφάλμα, χρησιμοποιώντας αυτή τη φορά ένα ενεργό κλειδί από το ψηφιακό αποθετήριο κλειδιών API. Η λειτουργικότητα αυτή υλοποιήθηκε μέσω της μεθόδου `handle_API_error()`.

```
def handle_API_error(self, channel_id, video_id, used_API,
callback_function):
    self.API_keys[used_API] = "inactive"
    new_API = self.get_active_API()

    if callback_function == "parse":
        yield Request(self.get_url(channel_id, "channel",
            new_API), callback=self.parse, meta={"usedAPI":
            new_API, "channel_id": channel_id})
        return
    elif callback_function == "parse_videos":
        yield Request(self.get_url(channel_id, "videosList",
            new_API), callback=self.parse_videos,
            meta={"channel_id": channel_id, "usedAPI": new_API})
        return
```

```

yield Request(self.get_url(channel_id, "video", new_API),
              callback=self.parse_video_stats,
              meta={"channel_id":channel_id, "video_id":video_id,
                  "usedAPI":new_API})
return

```

4.3 Συγκέντρωση δεδομένων και προκλήσεις

Κατά τη συγκέντρωση δεδομένων από το Instagram και το YouTube, πέραν της επιλεκτικότητας που απαιτούσε η συγκομιδή τους, ιδιαίτερη προσοχή έπρεπε να δοθεί και σε διάφορες προκλήσεις που η διαδικασία εγκυμονούσε.

Πρωτίστως, κύριο μέλημα της ερευνητικής διαδικασίας υπήρξε ο εντοπισμός ενός ικανοποιητικού αριθμού ελληνικών λογαριασμών, οι οποίοι θα συγκροτούσαν το βασικό δείγμα της μελέτης. Η αρχική προσέγγιση για την επίτευξη αυτού ήταν η επιλογή ενός λογαριασμού, ο οποίος θα αποτελούσε την αφετηρία για την αναζήτηση νέων χρηστών. Την έναρξη της διαδικασίας αναζήτησης θα σηματοδοτούσε η εύρεση πιθανών επισυναπτόμενων προφίλ στα πλαίσια των αναρτήσεων του λογαριασμού-αφετηρία και, στη συνέχεια, η επανάληψη του ίδιου βήματος για κάθε ένα από αυτά. Ωστόσο, μια τέτοια μεθοδολογία θα αποδεικνύονταν μη καρποφόρα, καθώς ο τεράστιος αριθμός λογαριασμών με μικρό αριθμό ακολούθων/εγγεγραμμένων ελαχιστοποιούσε τις πιθανότητες εύρεσης λογαριασμών μεγαλύτερης δημοτικότητας σε συντομότερο χρονικό διάστημα, αυξάνοντας, έτσι, κατακόρυφα το χρονικό και επεξεργαστικό κόστος της αναζήτησης. Παράλληλα, το Instagram δεν παραθέτει δημόσια την εθνικότητα του εκάστοτε χρήστη, με αποτέλεσμα η τυχαία αναζήτηση ελληνικών λογαριασμών να καθίστανται ένα πολύ δύσκολο έργο. Προκειμένου να αποφευχθούν τα παραπάνω εμπόδια, αποφασίστηκε πως η αξιοποίηση των δημοσιοποιημένων λιστών δημοφιλών Ελλήνων χρηστών του Instagram και του YouTube από διαθέσιμες διαδικτυακές πηγές^{13,14} θα μπορούσε να επικεντρώσει την προσοχή αποκλειστικά σε ελληνικά προφίλ, χωρίς να υποπέσει στην αδυναμία εύρεσης λογαριασμών με μεγάλο αριθμό ακολούθων/εγγεγραμμένων.

Το σύνολο των λογαριασμών που συλλέχθηκαν υπόκειντο σε περαιτέρω φιλτράρισμα, με σκοπό να πληρούν όλες τις απαραίτητες, για την έρευνα, προδιαγραφές. Ειδικότερα, η ανάγκη για μια έρευνα που θα ανταποκρίνεται στα σύγχρονα ψηφιακά δρώμενα επέβαλε τη διατήρηση μόνο εκείνων των αναρτήσεων/ μεταφορτώσεων που είχαν κοινοποιηθεί από τον Ιανουάριο του 2020 και ύστερα, γεγονός που είχε ως αποτέλεσμα να απορριφθούν οι λογαριασμοί που δεν εμφάνιζαν δραστηριότητα κατά την συγκεκριμένη χρονική περίοδο. Εξίσου σημαντική παράμετρος που καθόρισε το τελικό δείγμα της μελέτης ήταν και το κατώτατο όριο ακολούθων/εγγεγραμμένων έτσι ώστε ένας λογαριασμός να γίνει αποδεκτός. Παρατηρήθηκε πως προφίλ με αριθμό ακολούθων και εγγεγραμμένων μικρότερο των 1000 και 5000, αντίστοιχα, παρουσίαζαν, αρκετά συχνά, μεγάλες περιόδους αδράνειας ή διατηρούσαν το περιεχόμενο τους ιδιωτικό, στην περίπτωση του Instagram, κάτι το οποίο αντέβαινε τις προκαθορισμένες απαιτήσεις και οδήγησε στην απόφαση για παράληψη τέτοιου είδους λογαριασμών. Συμπερασματικά, μετά το πέρας όλων των διαδικασιών διαλογής και φιλτραρίσματος των συγκεντρωμένων δεδομένων, το σύνολο λογαριασμών που αποτέλεσαν το καταληκτικό δείγμα της έρευνας ανερχόταν στα 2.727

¹³<https://starngage.com/app/global/influencer/ranking/greece>

¹⁴<https://stats.video/top/most-subscribed/youtube-channels/greece/of-all-time>

ελληνικά προφίλ του Instagram και στους 2.057 λογαριασμούς του YouTube.

Εφόσον τα ονόματα των επιθυμητών λογαριασμών είχαν αποθηκευτεί στην βάση δεδομένων, έπρεπε, στη συνέχεια, να γίνει αντιστοίχιση της κάθε εγγραφής της βάσης με τον αντίστοιχο λογαριασμό του Instagram και του YouTube σε πραγματικό χρόνο, με σκοπό να συλλεχθούν όλα τα επιθυμητά στοιχεία για τον κάθε χρήστη. Κάτι τέτοιο αποτέλεσε ένα απλό έργο για το YouTube, καθώς η χρήση του YouTube Data API εξασφάλισε ένα σταθερό μέσο πρόσβασης στα δεδομένα των λογαριασμών του YouTube. Στην αντίπερα όχθη, το γεγονός ότι τα δεδομένα του Instagram έπρεπε να συλλεχθούν από τον πηγαίο κώδικα της ιστοσελίδας κατέστησε αναγκαία τη συνεχή επίβλεψη και συντήρηση του αυτοματοποιημένου μηχανισμού της έρευνας, αφού μία μικρή τροποποίηση του πηγαίου κώδικα από τη μεριά του Instagram θα ήταν αρκετή για να τον θέσει εκτός λειτουργίας.

Ακόμη ένα εμπόδιο που παρουσιάστηκε την ολοκλήρωση της συγκομιδής δεδομένων ήταν η κατηγοριοποίηση των συγκεντρωμένων προφίλ με βάση τον διαχειριστή τους. Στην προσπάθεια για χαρτογράφηση της συμπεριφοράς διακριτών κατηγοριών χρηστών των Instagram και YouTube, έγινε αντιληπτό πως πληροφορίες σχετικές με τη φύση του διαχειριστή δεν ήταν δημοσίως διαθέσιμες, με αποτέλεσμα η διάκριση των χρηστών να στηριχθεί σε προσωπικά κριτήρια. Ως βάση της διάκρισης αυτής στάθηκε η αναζήτηση του επιχειρηματικού ενδιαφέροντος που πιθανώς να διέπει της σχέσεις μεταξύ των χρηστών, επομένως ορίστηκαν οι εξής τέσσερις στοιχειώδεις κατηγορίες λογαριασμών:

- **Ανδρας:** κατηγορία στην οποία κατατάσσονται λογαριασμοί των οποίων το περιεχόμενο στηρίζεται στη διαδικτυακή παρουσία του αποκλειστικού διαχειριστή τους, ανδρικού φύλου
- **Γυναίκα:** κατηγορία στην οποία κατατάσσονται λογαριασμοί των οποίων το περιεχόμενο στηρίζεται στη διαδικτυακή παρουσία του αποκλειστικού διαχειριστή τους, γυναικείου φύλου
- **Επιχείρηση:** κατηγορία στην οποία κατατάσσονται λογαριασμοί που προσφέρουν υπηρεσίες ή προϊόντα και είναι πιθανό να αναζητήσουν την προώθηση τους μέσω άλλων χρηστών του Instagram/YouTube· τονίζεται πως η διάκριση βασίστηκε στην προσωπική μας κρίση
- **Άλλο:** κατηγορία στην οποία κατατάσσονται λογαριασμοί των οποίων το περιεχόμενο δεν στηρίζεται στην εικονική παρουσία του αποκλειστικού διαχειριστή τους και δεν αναζητούν την προώθηση του περιεχομένου τους· στην κατηγορία αυτή ανήκουν λογαριασμοί όπως δισκογραφικές εταιρίες ή απρόσωποι/ομαδικοί λογαριασμοί αμιγώς ψυχαγωγικού περιεχομένου

Αξίζει να σημειωθεί, επίσης, η ιδιαιτερότητα που παρατηρήθηκε κατά την εξαγωγή πληροφοριών από τις αναρτήσεις ενός λογαριασμού του Instagram. Έχοντας ως σκοπό την ανάκτηση όλων των δημοσιεύσεων που κοινοποιήθηκαν κατά τη χρονική περίοδο Ιανουάριος 2020 - Ιανουάριος 2021, σε κάθε λογαριασμό η προσπέλαση ξεκινούσε από την πιο πρόσφατη ανάρτηση και, αναπηδώντας στις διαδοχικές δημοσιεύσεις, κατέληγε στην παλαιότερη, εντός του 2020, ανάρτηση. Η διαδικασία αναπήδησης αν και αρχικά φάνταζε απλή, έκρυβε αρκετή πολυπλοκότητα, λόγω σελιδοποίησης που εφαρμόζει η πλατφόρμα στις δημοσιεύσεις. Πιο συγκεκριμένα, οι αναρτήσεις του κάθε λογαριασμού ομαδοποιούνται σε σύνολα των 12 και η αναπήδηση από

την μία ομάδα αναρτήσεων στην επόμενη απαιτεί ειδική μεταχείριση, η οποία περιγράφεται αναλυτικότερα στην Υποενότητα 4.1.2.

Επιπλέον πολυπλοκότητα στη διαδικασία συλλογής δεδομένων προσέθεσε το γεγονός ότι δίνεται η δυνατότητα από το YouTube στους δημιουργούς να διατηρούν ιδιωτικό των αριθμό των αντιδράσεων του εκάστοτε βίντεο, με αποτέλεσμα τα συγκεκριμένα στοιχεία να μην είναι διαθέσιμα για ένα μέρος των υπό συλλογή βίντεο. Κάτι τέτοιο απαιτούσε την ειδική μεταχείριση των αντίστοιχων μεταφορτώσεων, μέσω της πραγματοποίησης ελέγχων για την ιδιωτικότητα ή μη των πληροφοριών αυτών και, κατ' επέκταση, της διαμόρφωσης των εγγράφων της βάσης δεδομένων, με αυτά να συμπεριλαμβάνουν τον αριθμό των αντιδράσεων, όταν αυτές διατίθενται, ενώ σε αντίθετη περίπτωση, να μην στεγάζουν τα αντίστοιχα πεδία.

Τέλος, δεσμευτική για την έρευνα κατέστη η τακτική που εφαρμόζει η πλατφόρμα του YouTube, κατά την οποία δεν επιτρέπεται η πρόσβαση σε ολόκληρο το σύνολο των βίντεο καναλιών με αριθμό μεταφορτώσεων μεγαλύτερο του 500, προσεγγιστικά, αλλά σε ένα υποσύνολο αυτών. Αυτό είχε ως αποτέλεσμα, να μην συλλεχθούν όλες οι διαθέσιμες πληροφορίες για τα κανάλια που εμφάνιζαν πολύ υψηλή καθημερινή δραστηριότητα. Παρόλα αυτά, τέτοιου είδους κανάλια αποτέλεσαν μειονότητα, κάτι που σε συνδυασμό με το ότι η γενική απόδοση τους ήταν δυνατό να σκιαγραφηθεί ακόμα και μέσω των περιορισμένων στοιχείων που κατέχονταν για αυτά, λόγω του μεγάλου πλήθους μεταφορτώσεων, προσέδωσε μη αποτρεπτικό, για τη μελέτη, χαρακτήρα στον παραπάνω περιορισμό του YouTube.

Κεφάλαιο 5

Μελέτη δεδομένων

5.1 Διερεύνηση του Instagram

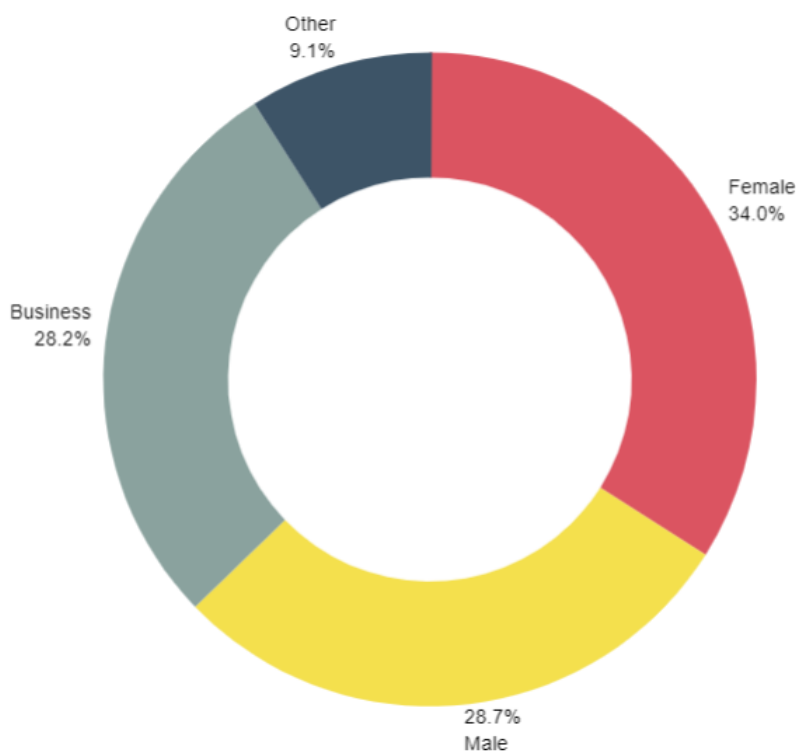
Το πρώτο μέσο κοινωνικής δικτύωσης που αποφασίστηκε να εξεταστεί για τις ανάγκες της έρευνας είναι το Instagram. Τον βασικό πυλώνα της απόφασης αυτής αποτέλεσε η συνεχώς αυξανόμενη δημοτικότητα που παρουσιάζει τα τελευταία χρόνια η συγκεκριμένη πλατφόρμα, με τους μηνιαίους ενεργούς χρήστες της να ανέρχονται στους 1 δισεκατομμύριο κατά το χρονολογικό έτος 2020-2021 [35]. Το Instagram παρέχει ένα ψηφιακό περιβάλλον στο οποίο οι χρήστες μπορούν να κοινοποιούν φωτογραφίες, βίντεο ή καθημερινά στιγμιότυπα (Instagram stories), καθώς και να αλληλεπιδρούν με δημοσιεύσεις άλλων χρηστών, δημιουργώντας έτσι ένα προσωπικό δίκτυο διαδικτυακών γνωριμιών. Το τακτικά ανανεώσιμο περιεχόμενο και η αμεσότητα που διέπει την αλληλεπίδραση με το αναρτημένο υλικό έχει καταφέρει να διατηρήσει το ενδιαφέρον του κοινού σε υψηλό επίπεδο, με χρήστες από ολόκληρο το ηλικιακό φάσμα να αφιερώνουν αρκετό χρόνο καθημερινά περιηγούμενοι στο συγκεκριμένο κοινωνικό δίκτυο. Η συγκέντρωση ενός μεγάλου συνόλου του πληθυσμού στο Instagram, σε συνδυασμό με τη συνεχή αλληλεπίδρασή του με πληθώρα δημοσιεύσεων, έχει ωθήσει πολλές επιχειρήσεις στο να εισέλθουν στον κόσμο του Instagram, εγκαθιδρύοντας την ψηφιακή τους παρουσία και προωθώντας τα προϊόντα τους, αλλά και χρήστες με μεγάλο αριθμό ακολούθων να αναλαμβάνουν τον ρόλο των διαφημιστών προϊόντων και υπηρεσιών στο διαδίκτυο. Φαίνεται, λοιπόν, πως το Instagram συγκεντρώνει αρκετό ερευνητικό ενδιαφέρον, τόσο όσον αφορά στη μελέτη της συμπεριφοράς των καθημερινών χρηστών όσο και τη διερεύνηση της εμπορικής δραστηριοποίησης στα κοινωνικά δίκτυα.

Στην ακόλουθη υποενότητα αναλύονται τα δεδομένα τα οποία συγκεντρώθηκαν από το Instagram και παρουσιάζονται διεξοδικά τα εξαχθέντα συμπεράσματα, τα οποία αφορούν στη πλατφόρμα αυτή και, ειδικότερα, στη διαδικτυακή συμπεριφορά του ελληνικού κοινού στα πλαίσια αυτής κατά το χρονολογικό έτος 2020.

5.1.1 Ανάλυση των συγκεντρωμένων δεδομένων

Η μελέτη των δεδομένων του Instagram αποφασίστηκε να διεξαχθεί μεταβαίνοντας από το γενικότερο στο ειδικότερο επίπεδο, ξεκινώντας με την περιγραφή των θεμελιωδών στοιχείων για την ομαλή ανάπτυξη της έρευνας και συνεχίζοντας με την εστίαση σε στοχευμένα σημεία ενδιαφέροντος της πλατφόρμας.

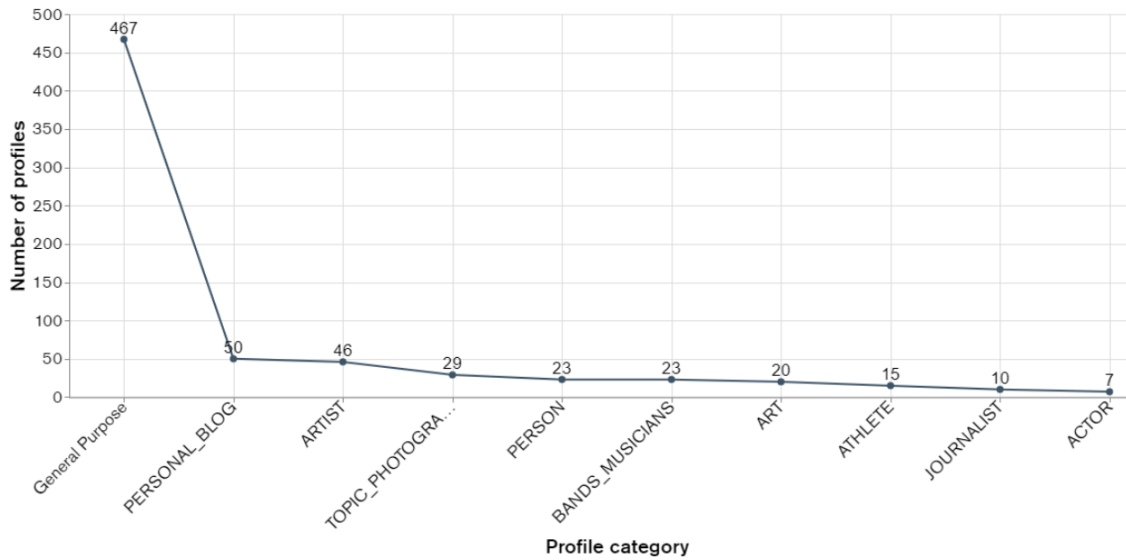
Σημείο εκκίνησης της έρευνας αποτέλεσε η μέτρηση του ποσοστού το οποίο καταλαμβάνει η εκάστοτε κατηγορία χρηστών από το σύνολο των ελληνικών λογαριασμών που συλλέχθηκαν, με βάση τη διάκριση των προφίλ που έλαβε χώρα στο Κεφάλαιο 4 και συγκεκριμένα στην Ενότητα 4.3.



Σχήμα 5.1: Κατηγορίες χρηστών του Instagram

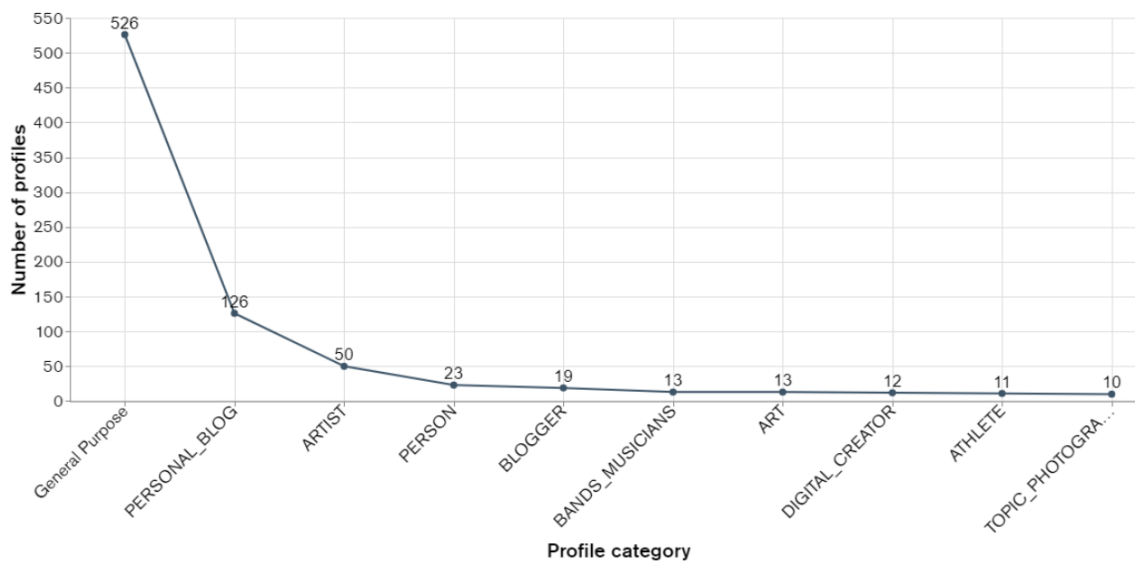
Όπως παρατηρείται στο Σχήμα 5.1, δύο στα τρία προφίλ του συνολικού δείγματος αντιστοιχούν σε γυναίκες ή άνδρες, με τις πρώτες να εμφανίζουν υπεροχή ως προς τον αριθμό λογαριασμών στον Instagram (34%). Αισθητή, επίσης, είναι η παρουσία λογαριασμών οι οποίοι κατέχονται από επιχειρήσεις, με αυτούς να καταλαμβάνουν το 28,2% του συνόλου, ποσοστό το οποίο τείνει στο 1/3 των συγκεντρωμένων προφίλ. Το μικρότερο ποσοστό λογαριασμών κατέχεται από τη κατηγορία Άλλο (όπως κωμικές/ψυχαγωγικές, ενημερωτικές σελίδες κ.α.), με αυτό να ανέρχεται μόλις στο 9,1%.

Λαμβάνοντας υπόψη την κατηγοριοποίηση που εφαρμόζει η πλατφόρμα του Instagram στους λογαριασμούς, ως επόμενο βήμα επιλέχθηκε η ανάδειξη των κορυφαίων θεματικών κατηγοριών ανά κατηγορία χρηστών (Άνδρας, Γυναίκα, Επιχείρηση, Άλλο).



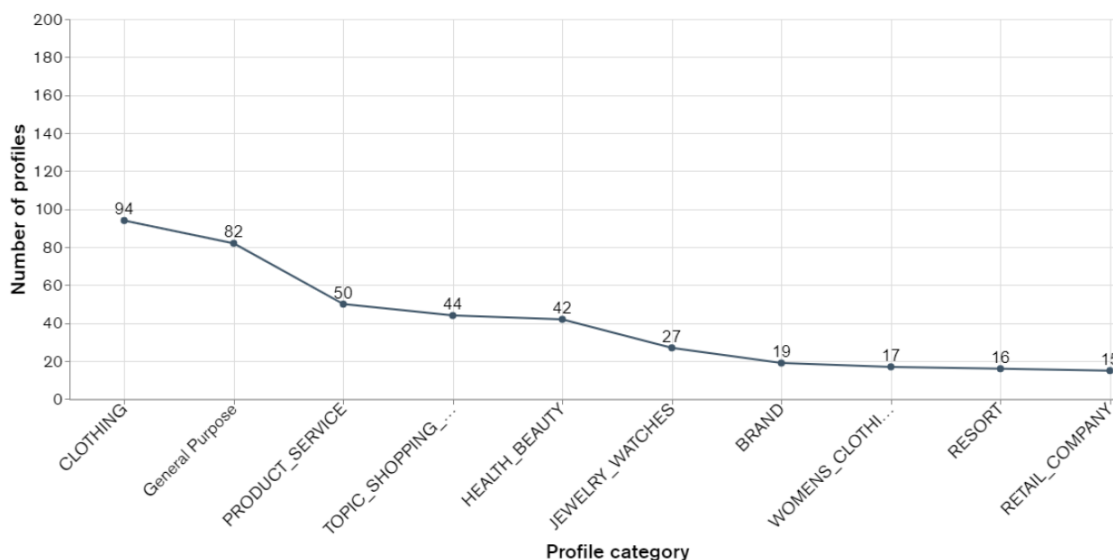
Σχήμα 5.2: Κορυφαίες κατηγορίες λογαριασμών (Άνδρες)

Ξεκινώντας με τους άνδρες, φαίνεται στο Σχήμα 5.2, εξαιρώντας τους λογαριασμούς τύπου *General Purpose*, οι οποίοι δεν έχουν κατατάξει το περιεχόμενό τους σε κάποια συγκεκριμένη θεματική κατηγορία, πιο συχνά συναντούνται λογαριασμοί που αναρτούν περιεχόμενο τεχνών (ARTIST, ART, ACTOR, BANDS_MUSICIANS), αθλητισμού (ATHLETE), φωτογραφίας (TOPIC_PHOTOGRAPHER) και προσωπικής καθημερινότητας (PERSONAL_BLOG, PERSON).



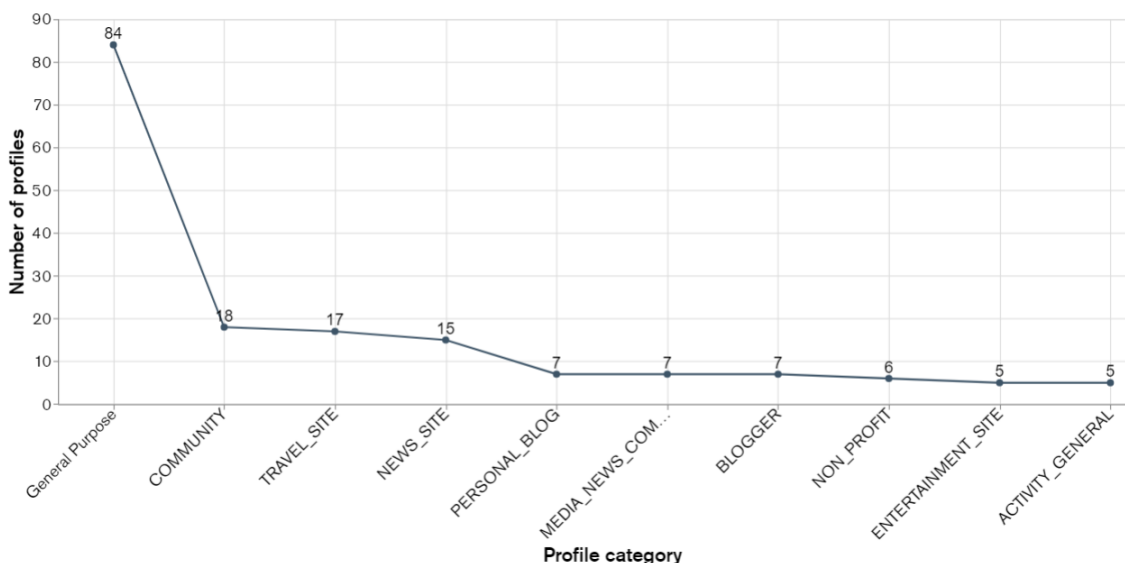
Σχήμα 5.3: Κορυφαίες κατηγορίες λογαριασμών (Γυναίκες)

Αντίστοιχη εικόνα με εκείνη των ανδρών, παρουσιάζουν οι γυναίκες (βλ. Σχήμα 5.3). Ωστόσο, στη περίπτωση των γυναικείων λογαριασμών εμφανίζεται μια αύξηση στον αριθμό των προφίλ με περιεχόμενο προσωπικής καθημερινότητας (PERSONAL_BLOG, PERSON, BLOGGER).



Σχήμα 5.4: Κορυφαίες κατηγορίες λογαριασμών (Επιχειρήσεις)

Όσον αφορά στην κατηγορία των επιχειρησιακών προφίλ, είναι ορατό στο Σχήμα 5.4 ότι ένα μεγάλο ποσοστό των επιχειρήσεων οι οποίες δραστηριοποιούνται στο Instagram είναι επικεντρωμένες στον κλάδο της ένδυσης (CLOTHING, JEWELRY_WATCHES, WOMENS_CLOTHING_STORE, BRAND), ενώ την εμφάνισή τους κάνουν και λογαριασμοί περιεχομένου ευεξίας/ομορφιάς (HEALTH_BEAUTY), τουρισμού (RESORT) και, γενικότερα, προώθησης προϊόντων.



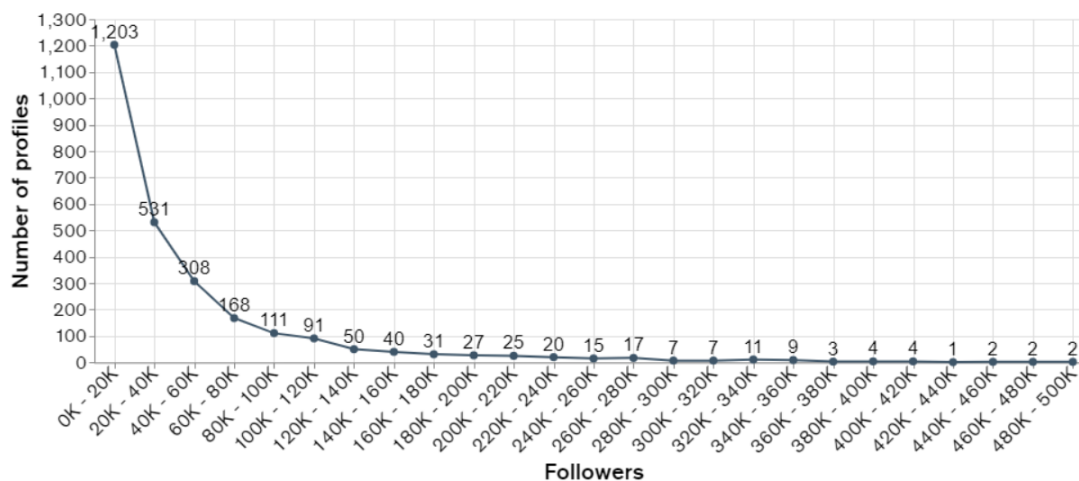
Σχήμα 5.5: Κορυφαίες κατηγορίες λογαριασμών (Άλλο)

Τέλος, αναλύοντας τις θεματικές κατηγορίες που υπερισχύουν στους λογαριασμούς τύπου Άλλο (βλ. Σχήμα 5.5), σκιαγραφείται ένας κοινός άξονας με όλες τις προηγούμενες κατηγορίες προφίλ χρηστών, με το περιεχόμενο ταξιδιωτικού ενδιαφέροντος (TRAVEL_SITE) και προσωπικής καθημερινότητας (PERSONAL_BLOG, BLOGGER) να συναντάται σε αρκετούς λογαριασμούς. Ση-

μαντικό μερίδιο καταλαμβάνουν, επίσης, προφίλ ενημερωτικού (NEWS_SITE, MEDIA_NEWS_COMPANY), κοινωνικού (COMMUNITY, NON_PROFIT) και ψυχαγωγικού χαρακτήρα (ENTERTAINMENT_SITE).

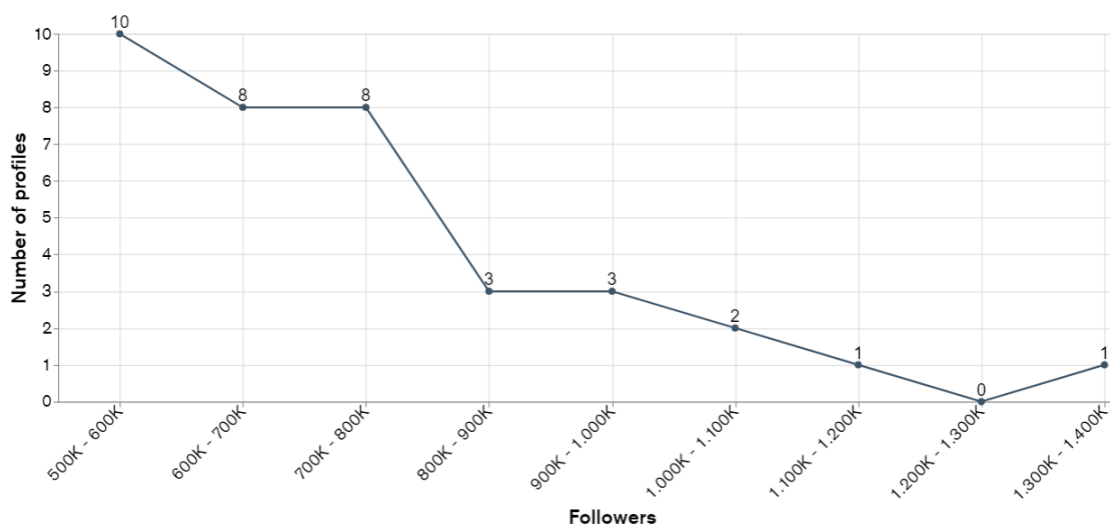
Ως ένα πρωταρχικό συμπέρασμα, με βάση τα παραπάνω διαγράμματα, μπορεί να εξαχθεί ότι η ελληνική κοινότητα στρέφεται προς το Instagram, με σκοπό τη ψυχαγωγία (τέχνες, αθλητισμός, καθημερινή ζωή), τη διάδραση με ταξιδιωτικό/φωτογραφικό περιεχόμενο και την αναζήτηση προϊόντων, ιδιαίτερα προϊόντων ρουχισμού, ενώ ενδιαφέρουσα είναι και η παρουσία ενημερωτικών λογαριασμών, η οποία πιθανώς μαρτυρά την ανάγκη του κοινού για ενημέρωση, μέσω των κοινωνικών δικτύων.

Συνεχίζοντας την ανάλυση των συλλεγμένων λογαριασμών του Instagram, θεωρήθηκε σημαντικό, πριν από τη μελέτη των αριθμητικών στοιχείων, να πραγματοποιηθεί η απεικόνιση της κατανομής αυτών, έτσι ώστε να σχηματιστεί μία ακριβής εικόνα για τις τιμές, στις οποίες τείνουν οι διάφορες μεταβλητές και, ως εκ τούτου, να εντοπισθούν οι ακραίες τιμές για κάθε μία από αυτές. Οι μεταβλητές για των οποίων τη κατανομή κρίθηκε απαραίτητη η εξέταση είναι: (i) ο αριθμός ακολούθων (*followers*), (ii) το *ER Post* της κάθε δημοσίευσης, (iii) το *ER Comments Post* της κάθε δημοσίευσης και (iii) το *ER View* της κάθε δημοσίευσης-βίντεο.



Σχήμα 5.6: Κατανομή του αριθμού των ακολούθων (κάτω των 500K ακολούθων)

Ξεκινώντας με τον αριθμό των ακολούθων των Ελλήνων χρηστών γίνεται αντιληπτό από τα Σχήματα 5.6 και 5.7, πως το ποσοστό των λογαριασμών οι οποίοι καταφέρνουν να ανελιχθούν σε μεγάλο πλήθος ακολούθων είναι συντριπτικά μικρότερο σε σχέση με αυτούς που διατηρούν λιγότερο ευρύ κοινό, με την πλειοψηφία να κυμαίνεται στην κατηγορία των μηδέν έως 100.000 ακολούθων. Είναι απαραίτητο να επισημανθεί πως, για να εξασφαλιστεί η ομοιομορφία των διαγραμμάτων κατανομής ακολούθων, εξαιρέθηκε από το δείγμα της έρευνας το προφίλ του *giannis_an34*, ο οποίος κατέλαβε την πρώτη θέση του δείγματος σε αριθμό ακολούθων, με αυτούς να ξεπερνούν τους 8.000.000.

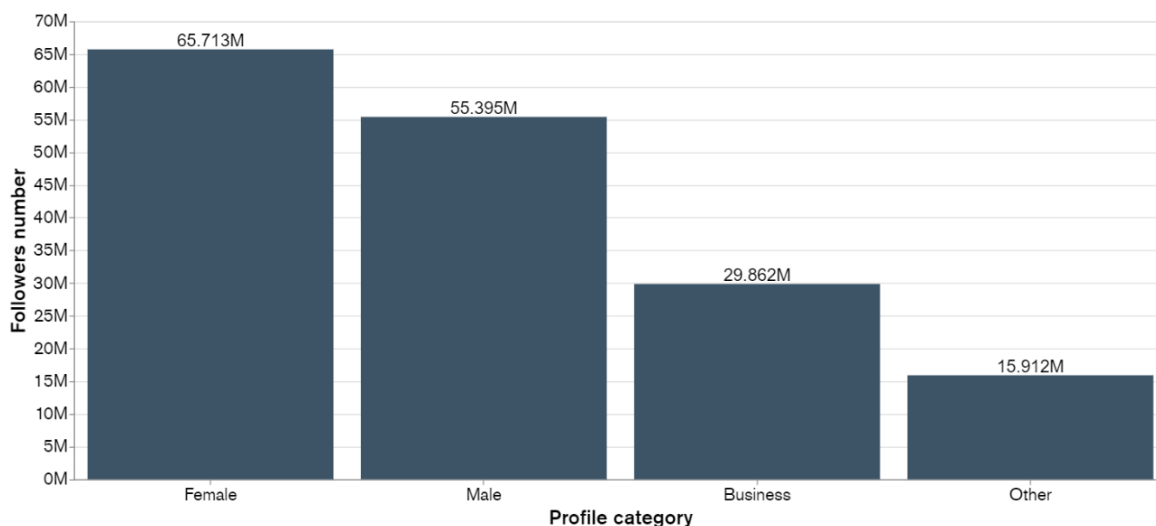


Σχήμα 5.7: Κατανομή του αριθμού των ακολούθων (άνω των 500K ακολούθων)

Παρατηρώντας την κατανομή του αριθμού των ακολούθων, είναι δυνατό να διακριθούν τέσσερις επιμέρους ομάδες: (i) 0-20.000 ακόλουθοι, (ii) 20.001-100.000 ακόλουθοι, (iii) 100.001-280.000 και (iv) ακόλουθοι > 280.000, οι οποίες ορίστηκαν με βάση το πλήθος των λογαριασμών οι οποίοι βρίσκονται στις αντίστοιχες βαθμίδες και χρησιμοποιήθηκαν για την ακριβέστερη ανάλυση των λογαριασμών διαφορετικής εμβέλειας.

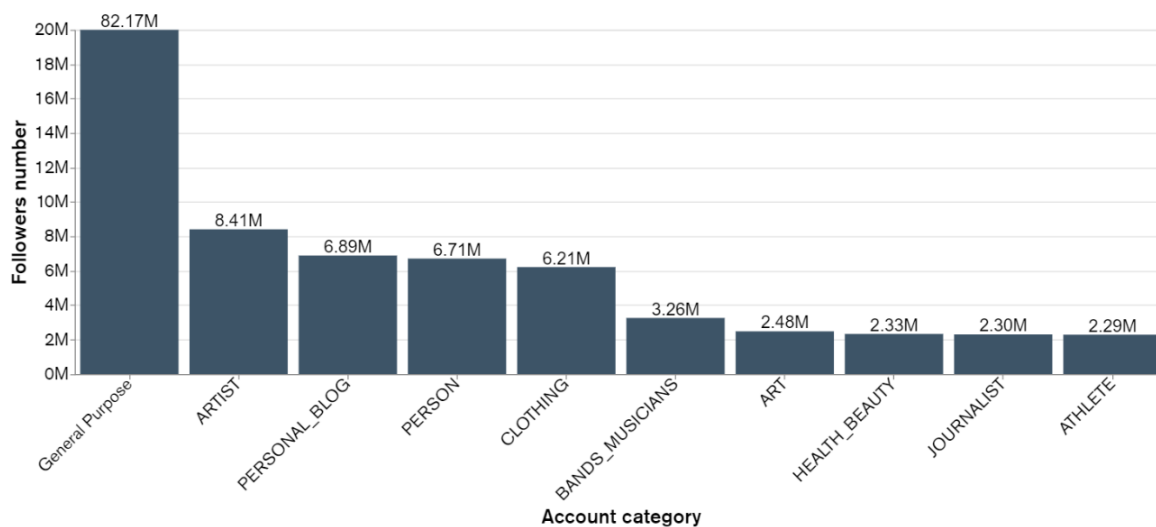
Την αναπαράσταση της κατανομής του αριθμού των ακολούθων, δύναται να διαδεχθεί η εξέταση του διαμοιρασμού τους στις διάφορες κατηγορίες χρηστών και θεματικές κατηγορίες περιεχομένου, υπολογίζοντας τον συνολικό αριθμό ακολούθων τον οποίο συγκεντρώνει κάθε μία από αυτές.

Όσον αφορά στο σύνολο ακολούθων ανά κατηγορία λογαριασμών, είναι φανερό στο Σχήμα 5.7 πως οι γυναίκες και οι άνδρες συγκεντρώνουν αθροιστικά τους περισσότερους ακολούθους, με τις πρώτες να διατηρούν τα ηνία. Τις δύο τελευταίες θέσεις της κατάταξης καταλαμβάνουν, με τη σειρά, οι επιχειρησιακοί λογαριασμοί και οι λογαριασμοί τύπου Άλλα. Το συγκεκριμένο αποτέλεσμα ήταν αναμενόμενο, καθώς ο αριθμός λογαριασμών ανά κατηγορία χρηστών, ο οποίος αναπαρίσταται στο Σχήμα 5.8, έχει άμεση επιρροή στο αντίστοιχο άθροισμα ακολούθων. Ωστόσο, παρά το γεγονός ότι το σύνολο των επιχειρησιακών λογαριασμών είναι σχεδόν ίσο με αυτό των ανδρών, το πρώτο φαίνεται να συγκεντρώνει αναλογικά αρκετά λιγότερους ακόλουθους, φαινόμενο το οποίο μπορεί να υποδηλώνει τη δυσκολία εγκαθίδρυσης ενός ευρέως κοινού εκ μέρους των επιχειρήσεων.



Σχήμα 5.8: Κατηγορίες χρηστών με τους περισσότερους ακόλουθους

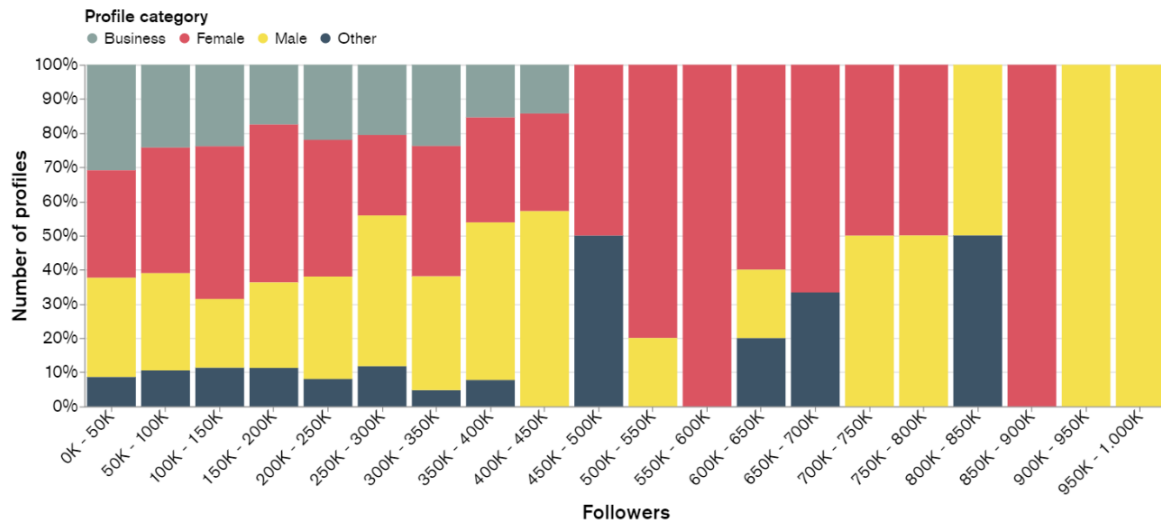
Η κατανομή των ακολούθων στις ποικίλες θεματικές κατηγορίες συνιστά ακόμα μία ένδειξη της προτίμησης του ελληνικού κοινού απέναντι σε περιεχόμενο (i) τεχνών, (ii) προσωπικής καθημερινότητας, (iii) ένδυσης/ομορφιάς και (iv) άθλησης (βλ. Σχήμα 5.9).



Σχήμα 5.9: Κατηγορίες περιεχομένου με τους περισσότερους ακόλουθους

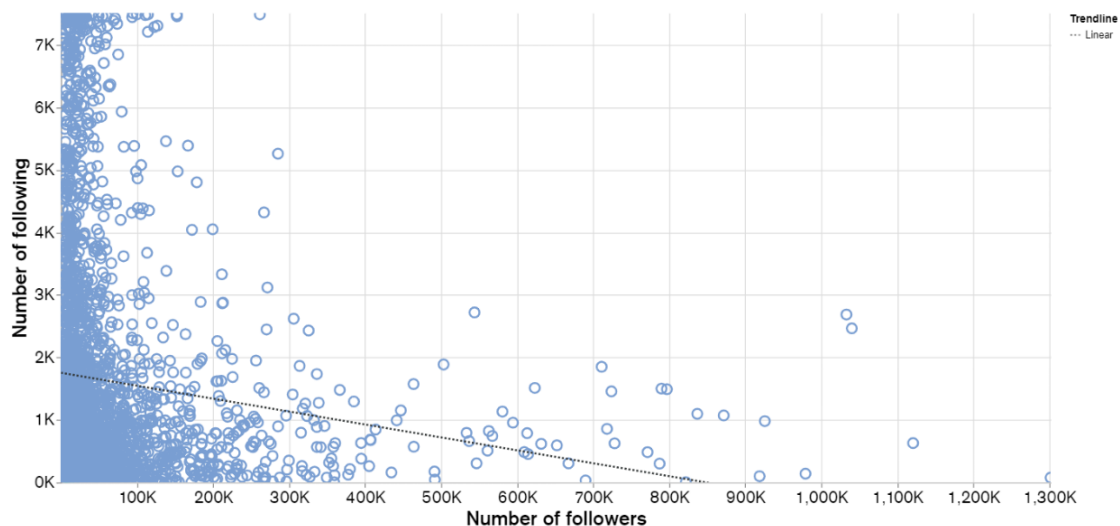
Ο υπολογισμός του πλήθους των λογαριασμών με βάση τον αριθμό των ακολούθων τους είναι δυνατό να παρέχει μια αναλυτικότερη εικόνα για την έκταση του κοινού κάθε μίας από τις κατηγορίες λογαριασμών. Ειδικότερα, βάσει του Σχήματος 5.10, στα υψηλότερα στρώματα αριθμού ακολούθων φαίνεται να κυριαρχούν λογαριασμοί διαχειριζόμενοι από άνδρες και γυναίκες, με ένα μικρό υποσύνολο των περιπτώσεων να αντιστοιχεί σε λογαριασμούς τύπου Άλλο. Στην αντίπερα όχθη, η επιχειρησιακοί λογαριασμοί, αλλά και το μεγαλύτερο μέρος των λογαριασμών τύπου Άλλο, περιορίζονται στην κλίμακα 0-450.000 ακολούθων, επαναφέροντας στο

προσκήνιο τη δυσκολία προσέλκυσης κοινού, η οποία διέπει τις δύο τελευταίες κατηγορίες λογαριασμών. Υπενθυμίζεται πως, όπως φάνηκε στο Σχήμα 5.7, τα ανώτερα στρώματα ακολούθων καταλαμβάνονται από μονοψήφιο αριθμό χρηστών.



Σχήμα 5.10: Αριθμός λογαριασμών με βάση τον αριθμό των ακολούθων

Στα πλαίσια της μελέτης των αριθμητικών στοιχείων γύρω από το πλήθος των ακολούθων ενός λογαριασμού, παρατηρήθηκε μια ενδιαφέρουσα συμπεριφορά των χρηστών του Instagram, σύμφωνα με την οποία λογαριασμοί με μεγαλύτερο πλήθος ακολούθων τείνουν να ακολουθούν λιγότερους χρήστες σε σχέση με λογαριασμούς με μικρότερο κοινό, όπως φαίνεται στο παρακάτω διάγραμμα (βλ. Σχήμα 5.11).

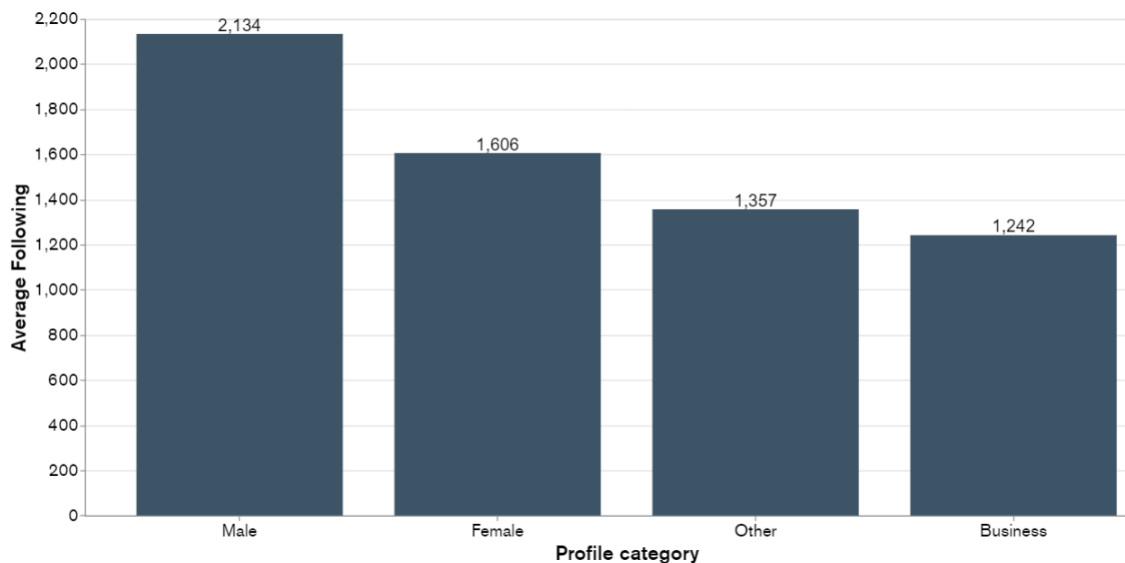


Σχήμα 5.11: Σχέση μεταξύ αριθμού ακολούθων και ακολουθούμενων

Οι λόγοι στους οποίους ενδέχεται να οφείλεται η παραπάνω διαπίστωση είναι, είτε η προσπάθεια των μικρότερων λογαριασμών να αποκτήσουν ακολούθους, ακολουθώντας περισσότερα

άτομα ή η διαφύλαξη της δημόσιας εικόνας των μεγαλύτερων λογαριασμών, ακολουθώντας στοχευμένους χρήστες.

Στη βάση της προηγούμενης παρατήρησης, αναλύθηκε ο μέσος αριθμός ακολουθούμενων χρηστών από την κάθε κατηγορία χρηστών, από την οποία προέκυψε ότι, κατά μέσο όρο, οι επιχειρήσεις και οι λογαριασμοί τύπου Άλλο ακολουθούν του λιγότερους χρήστες, με τις γυναίκες να τους διαδέχονται στην κατάταξη. Ως εκ τούτου, το υψηλότερο ποσοστό ακολουθούμενων χρηστών κατέχουν οι άνδρες. Είναι φανερό, λοιπόν, πως οι επιχειρήσεις, οι λογαριασμοί τύπου Άλλο και οι γυναίκες υιοθετούν μία περισσότερο συγκρατημένη συμπεριφορά ως αναφορά στη δημιουργία ψηφιακών δεσμών, σε σχέση με τους άνδρες (βλ. Σχήμα 5.12).

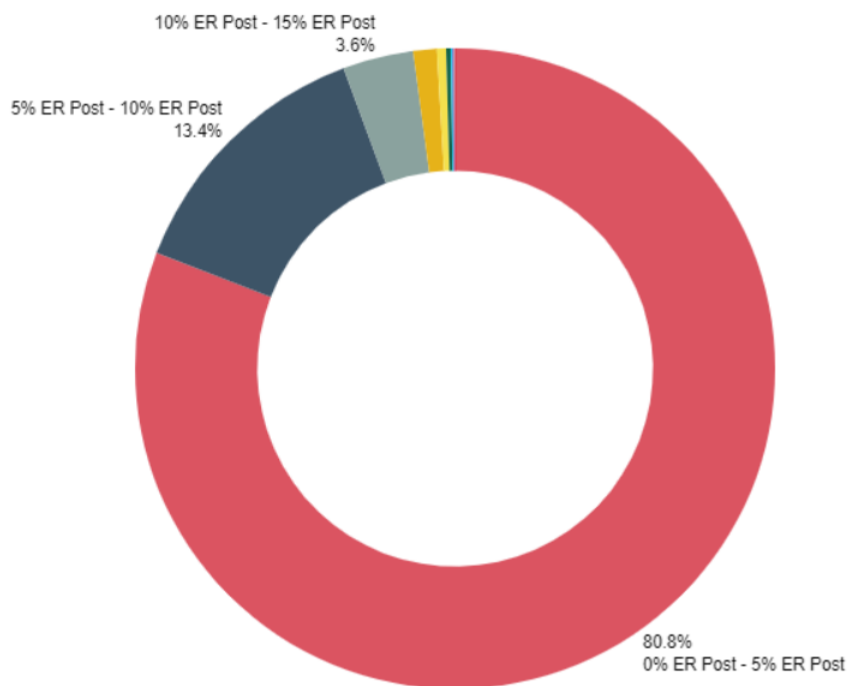


Σχήμα 5.12: Μέσος αριθμός ακολουθούμενων ανά κατηγορία χρηστών

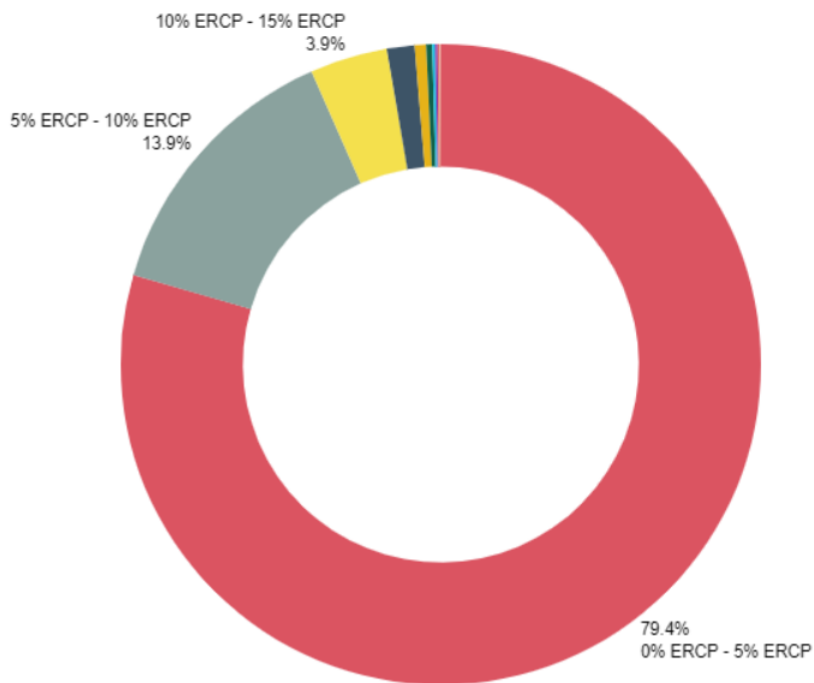
Σειρά έχουν οι δευτερογενείς μεταβλητές (i) *ER Post*, (ii) *ER Comments Post* και (iii) *ER View*, η σημασία των οποίων αναλύθηκε εκτενέστερα στο Κεφάλαιο 4 και συγκεκριμένα στην Υποενότητα 4.1.1. Υπενθυμίζεται πως βασικά συστατικά στοιχεία των παραπάνω μεταβλητών αποτέλεσαν (i) οι θετικές αντιδράσεις μίας δημοσίευσης (*likes*), (ii) οι σχολιασμοί μίας δημοσίευσης (*comments*), (iii) οι ακόλουθοι ενός λογαριασμού (*followers*) και (iv) οι προβολές μίας δημοσίευσης-βίντεο (*views*). Το γεγονός ότι οι τιμές των συγκεκριμένων στοιχείων ποικίλουν και είναι δυνατό να αυξηθούν απεριόριστα επηρέασε τις παραπάνω δευτερογενείς μεταβλητές, καθιστώντας δυσδιάκριτες τις ανώτατες τιμές τις οποίες αυτές μπορούν να λάβουν.

Παρατηρώντας τις κατανομές των μεταβλητών *ER Post* (βλ. Σχήμα 5.13) και *ER Comments Post* (βλ. Σχήμα 5.14), είναι εμφανές ότι οι δύο αυτές μετρικές παρουσιάζουν κοινή συμπεριφορά ως προς το πλήθος των αναρτήσεων που απαρτίζουν τις διαφορετικές βαθμίδες, κάτι το οποίο οφείλεται στο ότι και οι δύο υπολογίζουν τον βαθμό αλληλεπίδρασης του κοινού με μία δημοσίευση, με την ειδοποίησή τους διαφορά να είναι ο συμψηφισμός του αριθμού των σχολιασμών στη μεταβλητή *ER Comments Post*. Επομένως, το *ER Comments Post* φαίνεται να παρέχει μια πληρέστερη εικόνα της απήχησης μίας δημοσίευσης του Instagram, με τον κίνδυνο, ωστόσο, της απρόβλεπτης αύξησης της τιμής του να ελλοχεύει, λόγω της ενδεχόμενης παρουσίας μεγάλ-

λου αριθμού σχολιασμών σε κάποια ανάρτηση, όπως συμβαίνει συχνά στις περιπτώσεις των διαγωνισμών.



Σχήμα 5.13: Κατανομή του ER Post



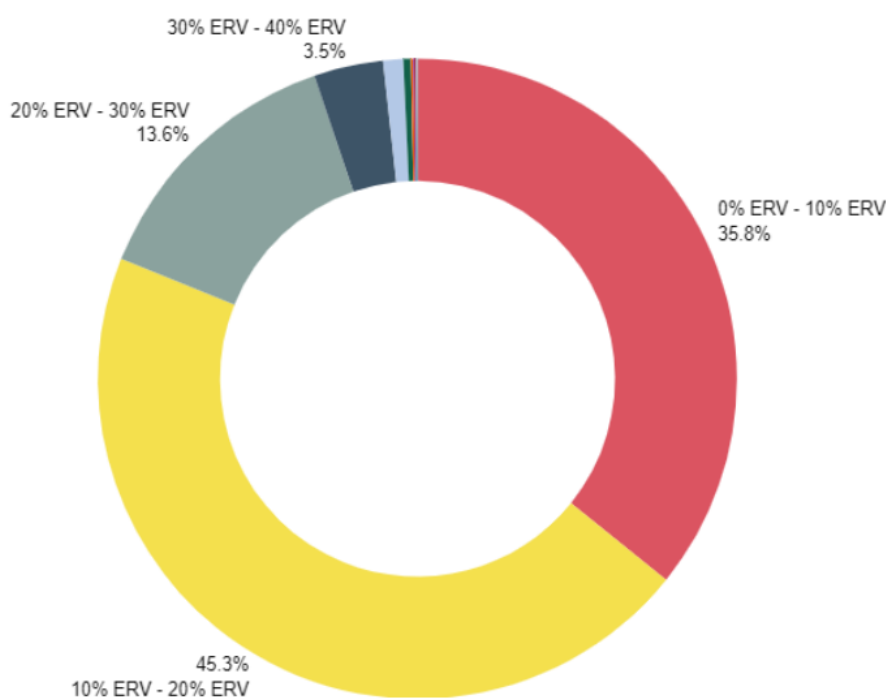
Σχήμα 5.14: Κατανομή του ER Comments Post

Με βάση τα δύο παραπάνω διαγράμματα εξάγεται το συμπέρασμα ότι, κατά μέσο όρο, η πλειο-

νότητα των ελληνικών δημοσιεύσεων, κατά το 2020, συγκέντρωσαν ποσοστό αλληλεπίδρασης της τάξης του 0% έως 5%, ενώ μια σημαντική μερίδα αυτών ανήλθε στο 15%. Αναλυτικότερα, οκτώ στις δέκα δημοσιεύσεις συγκέντρωσαν θετικές αντιδράσεις μόνο από το 0%-5% των ακολούθων του λογαριασμού που τις κοινοποίησε, ενώ μόλις 14 στις 100 δημοσιεύσεις απέσπασαν θετικές αντιδράσεις, στη καλύτερη των περιπτώσεων, 10% των ακολούθων του λογαριασμού. Ακόμα μικρότερο είναι το ποσοστό οι οποίες συναντούνται στη κλίμακα 10%-15% αλληλεπίδρασης, σχεδόν το 4% των συνολικών δημοσιεύσεων.

Με σκοπό την ακριβέστερη μελέτη της δραστηριότητας του γενικότερου συνόλου των λογαριασμών, αποφασίστηκε ως βάση να χρησιμοποιηθούν μόνο δημοσιεύσεις με αλληλεπίδραση μικρότερη του 20%, έτσι ώστε να ληφθεί υπόψιν ένα αντιπροσωπευτικό σύνολο του δείγματος, χωρίς αυτό να επηρεαστεί από τη μειονότητα των περιπτώσεων. Παρατηρείται πως το ποσοστό αλληλεπίδρασης του ελληνικού κοινού με τις δημοσιεύσεις του Instagram διατηρήθηκε σε χαμηλά επίπεδα, κάτι το οποίο, πιθανώς, να δικαιολογείται από το γεγονός ότι ένα μεγάλο μέρος των δημοσιεύσεων, οι οποίες συγκαταλέχθηκαν στην έρευνα, προέρχονται από λογαριασμούς με υψηλό αριθμό ακολούθων και, ενδεχομένως, με περιορισμένο ενεργό κοινό.

Τελευταίο στη λίστα των δευτερογενών μεταβλητών συναντάται το *ER View*, μέσω του οποίου είναι δυνατή η παρακολούθηση της αλληλεπίδρασης του κοινού με μία ανάρτηση βίντεο, συγκρίνοντας τον αριθμό προβολών βίντεο με το πλήθος των συγκεντρωμένων θετικών αντιδράσεων της δημοσίευσης.



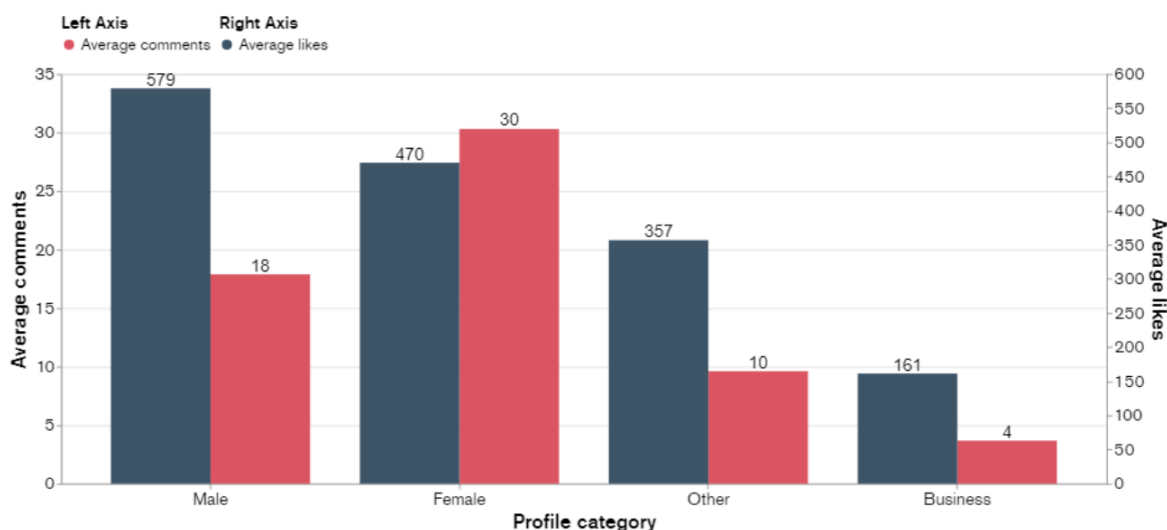
Σχήμα 5.15: Κατανομή του ER View

Σύμφωνα με το Σχήμα 5.15, διαπιστώνεται πως η μεγαλύτερη μερίδα δημοσιευμένων ελληνικών βίντεο στο Instagram απέσπασε, κατά το έτος 2020, ER View της τάξης 0% έως 20%, με τα περισσότερα από αυτά να συγκεντρώνονται στην κλίμακα 10%-20% αλληλεπίδρασης. Πιο συγκεκριμένα, φαίνεται πως το 35,8% των βίντεο αντέδρασε θετικά, κατά μέσο όρο, μηδέν έως ένας

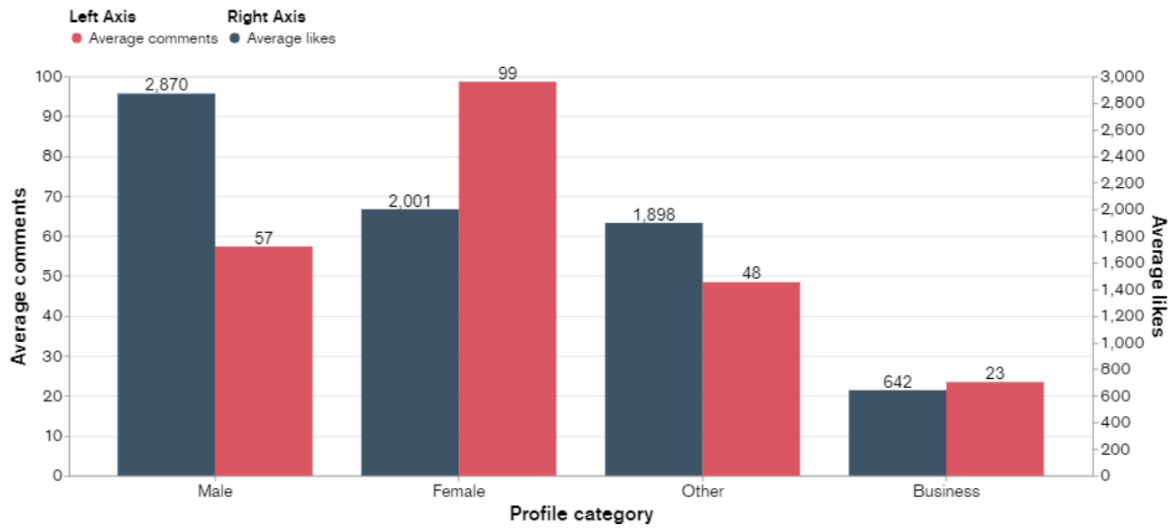
από τους 10 χρήστες που παρακολούθησαν το βίντεο, ενώ το 45,3% αυτών απέσπασε θετική αντίδραση από έναν έως δύο από τους 10 χρήστες. Σεβαστός επίσης είναι ο αριθμός των βίντεο που σημείωσαν ποσοστό αλληλεπίδρασης, το οποίο άγγιξε το 30% (το 13,6% των συλλεγμένων βίντεο). Είναι σημαντικό να αναφερθεί πως ως προβολή προσμετράται η παρακολούθηση ενός βίντεο για τουλάχιστον τρία δευτερόλεπτα. Επιπλέον, κάθε προβολή ανήκει σε έναν μοναδικό χρήστη, χωρίς να συμψηφίζονται οι πολλαπλές προβολές, οι οποίες προέρχονται από τον ίδιο λογαριασμό [11].

Με σκοπό το σχηματισμό μιας γενικής εικόνα γύρω από την απόδοση των ελληνικών βίντεο του Instagram, αποφασίστηκε να ληφθούν υπόψη μόνο τα βίντεο με ποσοστό αλληλεπίδρασης μικρότερο του 40%, έτσι ώστε να συμπεριληφθεί στην έρευνα ένα ικανοποιητικό μέρος του δείγματος και, παράλληλα, να αποφευχθούν οι ακραίες τιμές.

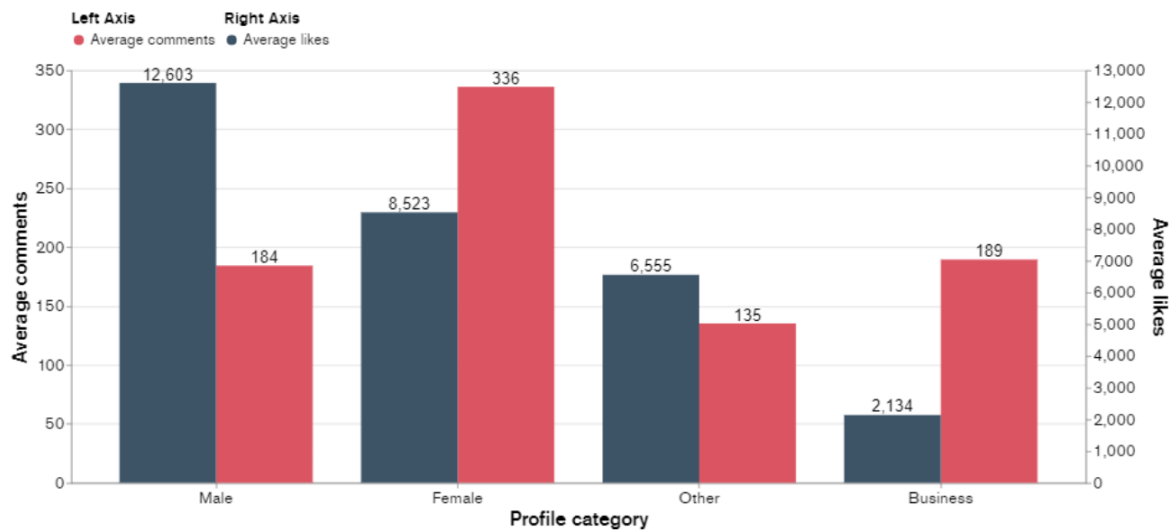
Αξιοποιώντας τη κατηγοριοποίηση των προφίλ με βάση τον αριθμό των ακολούθων τους, η οποία υλοποιήθηκε νωρίτερα, καθίσταται εφικτός ο έλεγχος της μεταβολής του αριθμού των αντιδράσεων, τις οποίες συγκεντρώνει κατά μέσο όρο η εκάστοτε κατηγορία λογαριασμών, ανά τις διαφορετικές ομάδες ακολούθων. Με βάση τα παραπάνω προκύπτουν τα ακόλουθα διαγράμματα για τη δημιουργία των οποίων χρησιμοποιήθηκαν μόνο δημοσιεύσεις με *ER Comments Post* μικρότερο ή ίσο του 20%.



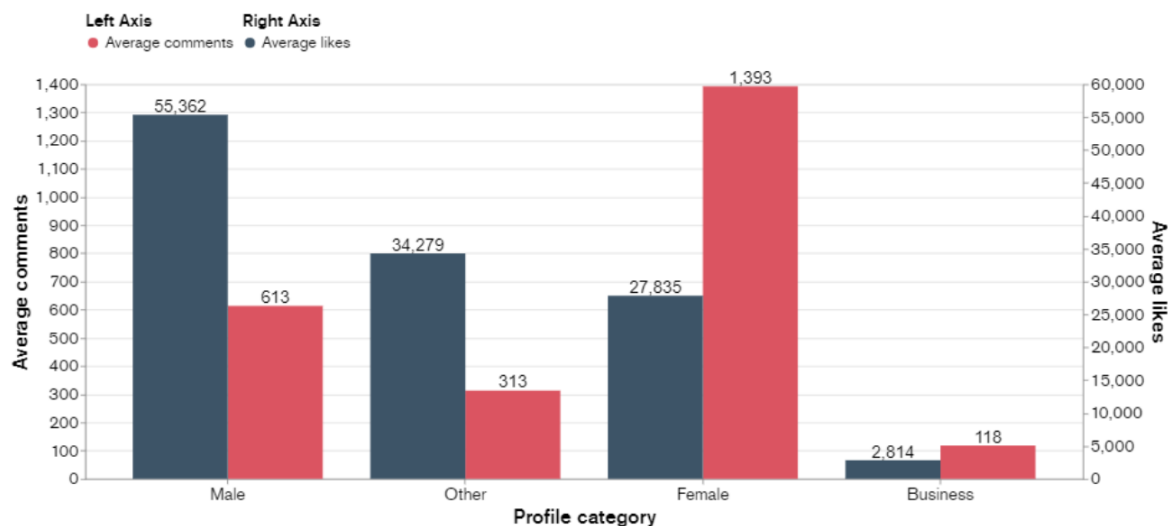
Σχήμα 5.16: Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (0K-20K ακόλουθοι)



Σχήμα 5.17: Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (20K-100K ακόλουθοι)



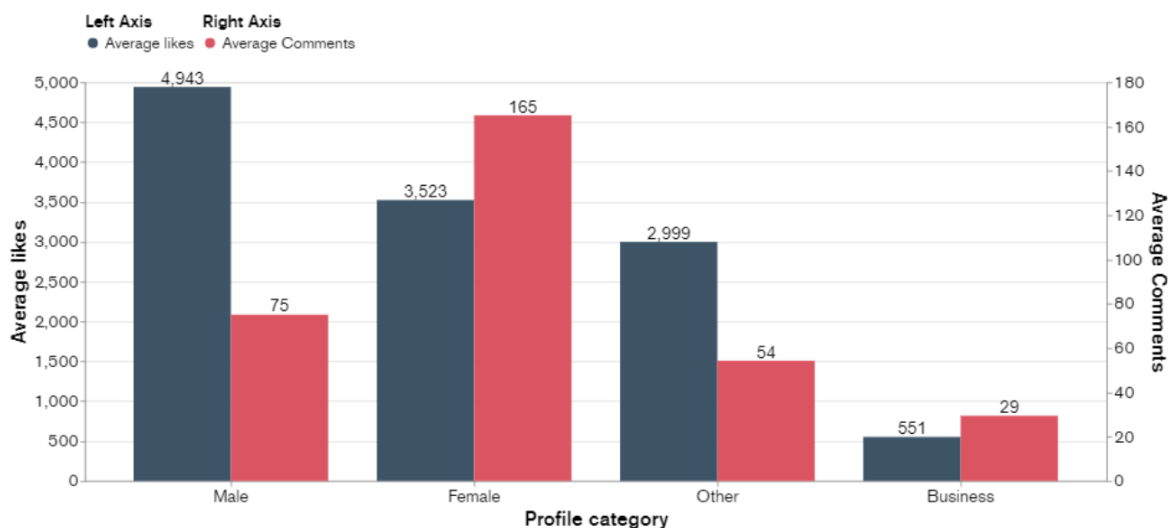
Σχήμα 5.18: Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (100K-280K ακόλουθοι)



Σχήμα 5.19: Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (ακόλουθοι > 280K)

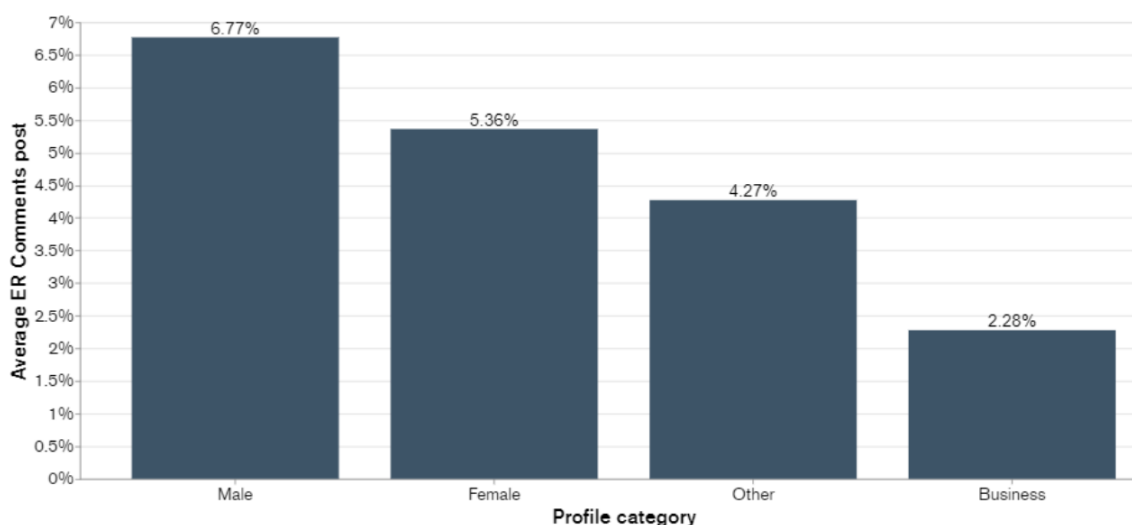
Πέραν του μέσου αριθμού αντιδράσεων που παρουσιάζονται στα παραπάνω διαγράμματα, είναι δυνατό να εξαχθούν ορισμένα γενικότερα συμπεράσματα, τόσο όσον αφορά στη σχέση μεταξύ του αριθμού ακολούθων-αριθμού αντιδράσεων, όσο και στις επιμέρους κατηγορίες χρηστών.

Αρχικά, η αύξηση του αριθμού των ακολούθων ενός λογαριασμού φαίνεται να συνεπάγεται την αύξηση του αριθμού των θετικών αντιδράσεων, ανεξαρτήτου κατηγορίας, κάτι το οποίο ισχύει και για τον αριθμό των σχολιασμών, οι οποίοι, ωστόσο, αυξάνονται με μικρότερο και λιγότερο προβλέψιμο ρυθμό. Συνολικά, οι άνδρες, οι γυναίκες και οι λογαριασμοί τύπου Άλλο του Instagram φαίνεται να συγκέντρωσαν, κατά το 2020, μεγαλύτερο αριθμό αντιδράσεων από τους επιχειρησιακούς λογαριασμούς, γεγονός το οποίο είναι πιθανό να οφείλεται στο ότι οι επιχειρήσεις δεν έχουν καταφέρει, μέχρι στιγμής, να εδραιώσουν το κοινό τους στην παρούσα πλατφόρμα. Επιπλέον, είναι εμφανές πως, κατά μέσο όρο, οι γυναίκες συγκέντρωσαν περισσότερους σχολιασμούς σε ολόκληρο το φάσμα του αριθμού ακολούθων, φαινόμενο που αναδεικνύει το υψηλό ποσοστό αλληλεπίδρασης που διατήρησαν με το κοινό τους. Από την άλλη μεριά, οι άνδρες φαίνεται να διατήρησαν τον υψηλότερο μέσο αριθμό θετικών αντιδράσεων, γεγονός το οποίο σε συνδυασμό με τα προηγούμενα εκφράζει το, επίσης, μεγάλο ποσοστό αλληλεπίδρασης τους με το κοινό, το οποίο ίσως να συνοδεύεται από ένα χαμηλότερο βαθμό εξοικείωσης με αυτό, αναλογικά με τις γυναίκες. Όλα όσα προαναφέρθηκαν μπορούν να επιβεβαιωθούν, μελετώντας το διάγραμμα τιμών για το γενικό σύνολο (βλ. Σχήμα 5.20).



Σχήμα 5.20: Μέσος αριθμός θετικών αντιδράσεων, σχολίων ανά κατηγορία λογαριασμών (γενικό σύνολο)

Γνωρίζοντας ότι ο υπολογισμός του ποσοστού αλληλεπίδρασης ενός χρήστη με τους ακολούθους του είναι μία συνάρτηση η οποία περιλαμβάνει τα αριθμητικά στοιχεία που εξετάστηκαν παραπάνω, αναμένεται τα ποσοστά αλληλεπίδρασης να παρουσιάζει ανάλογα πορεία με εκείνη του μέσου αριθμού αντιδράσεων ανά κατηγορία χρηστών.

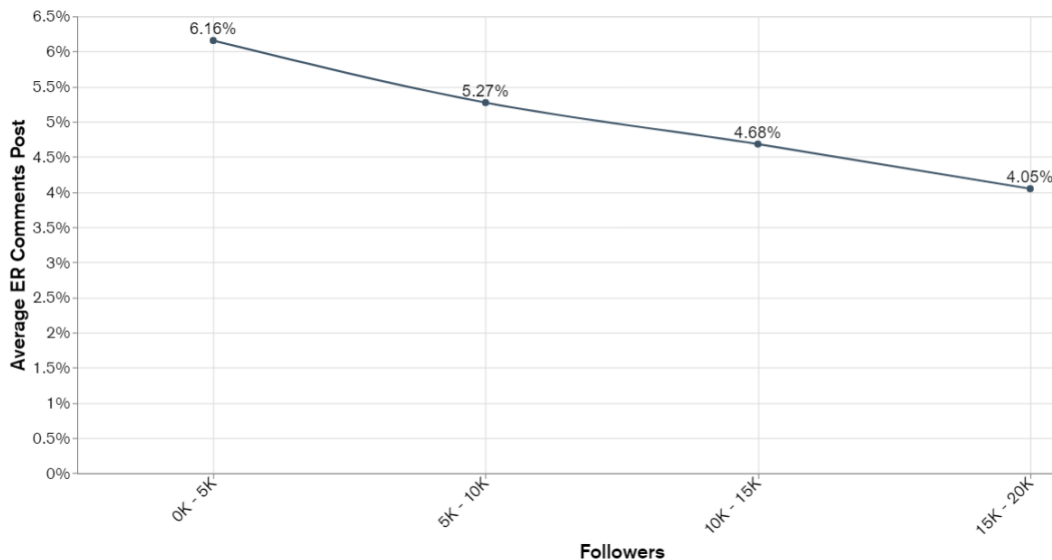


Σχήμα 5.21: Ποσοστό αλληλεπίδρασης ανά κατηγορία χρηστών

Πράγματι, η παραπάνω πρόβλεψη επαληθεύεται στο Σχήμα 5.21, με τους άνδρες και τις γυναίκες να διατηρούν το υψηλότερο ποσοστό *ER Comments Post*, ενώ οι επιχειρησιακοί λογαριασμοί, για το έτος 2020, παρέμειναν στη τελευταία βαθμίδα κατάταξης. Παρατηρείται ότι λογαριασμοί διαχειριζόμενοι από άνδρες και γυναίκες επέτυχαν καλύτερη επικοινωνία με τους ακολούθους, σε σχέση με τις άλλες δύο κατηγορίες λογαριασμών, κάτι το οποίο μπορεί να υπο-

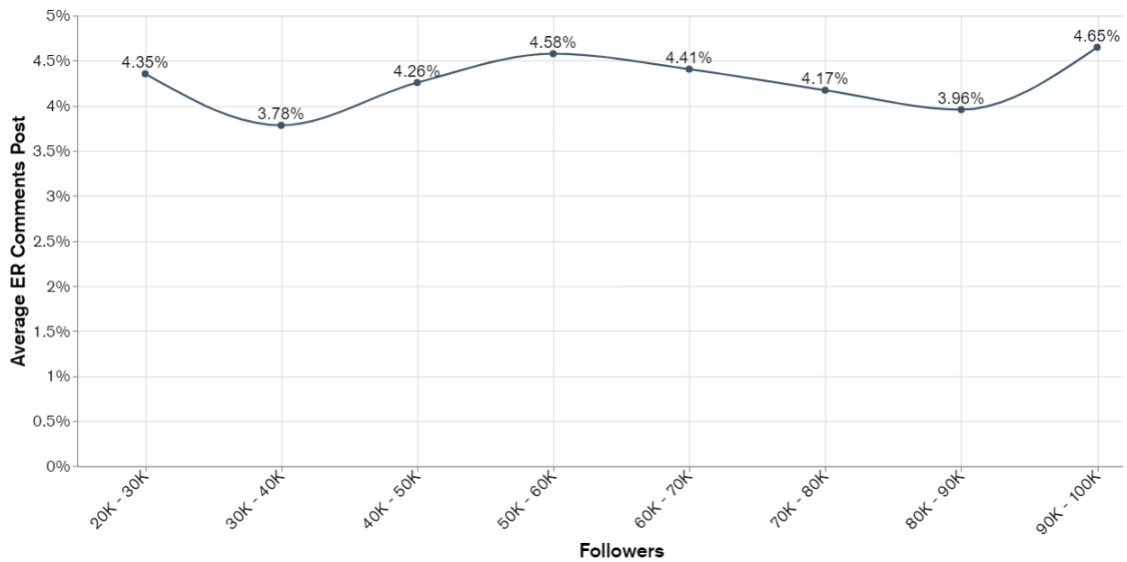
δηλώνει την μεγαλύτερη οικειότητα των χρηστών απέναντι σε λογαριασμούς οι οποίοι αναρτούν πιο προσωπικό περιεχόμενο, σε αντίθεση με τους περισσότερο απρόσωπους επιχειρησιακούς λογαριασμούς ή προφίλ τύπου Άλλο.

Προκειμένου να εντοπισθούν επιπλέον παράγοντες οι οποίοι είναι πιθανόν να επηρεάζουν το ποσοστό αλληλεπίδρασης των χρηστών του Instagram με τους ακολούθους τους, θεωρήθηκε απαραίτητη η αναζήτηση κάποιας συσχέτισης μεταξύ του αριθμού των ακολούθων ενός λογαριασμού και της μετρικής *ER Comments Post*. Η ανάλυση αυτή στηρίχθηκε στην κατηγοριοποίηση των χρηστών με βάση τον αριθμό των ακολούθων τους, η οποία, όπως αναφέρθηκε και σε προγενέστερα σημεία της ενότητας, είναι (i) 0-20.000 ακόλουθοι, (ii) 20.001-100.000 ακόλουθοι, (iii) 100.001-280.000 και (iv) ακόλουθοι > 280.000.

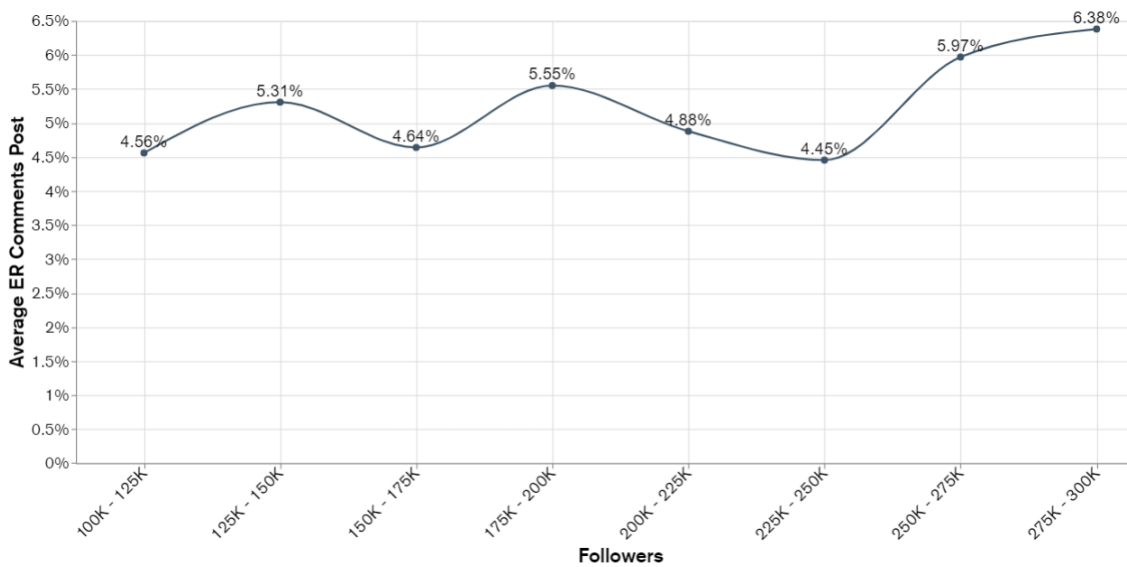


Σχήμα 5.22: Ποσοστό αλληλεπίδρασης ανά αριθμό ακολούθων (0K-20K ακόλουθοι)

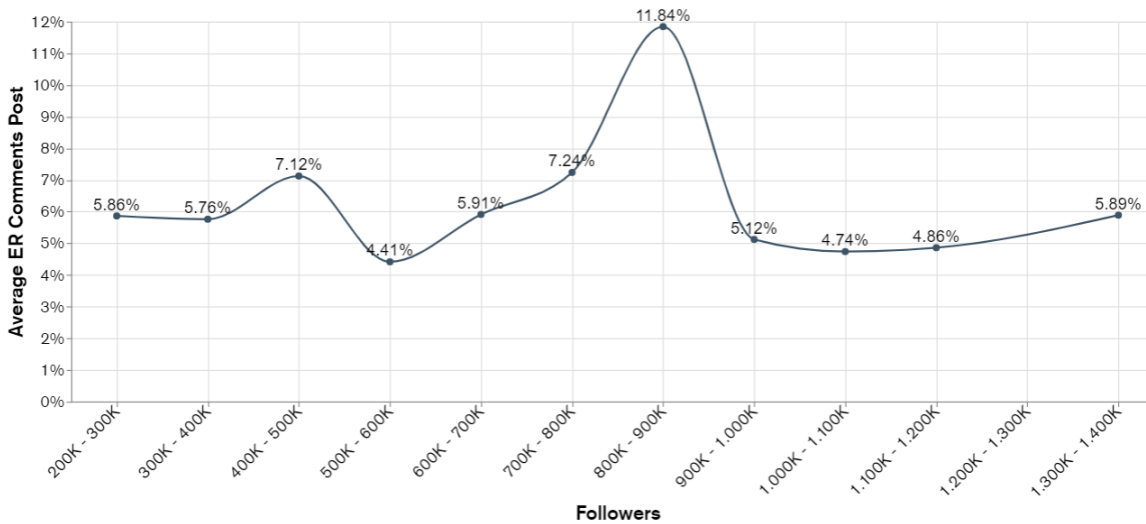
Παρακολουθώντας τη πορεία της μεταβλητής *ER Comments Post* στα χαμηλότερα στρώματα ακολούθων, είναι ορατό στο Σχήμα 5.22 πως η αύξηση του αριθμού ακολούθων ενός προφίλ συνδέεται με τη μείωση του ποσοστού αλληλεπίδρασης. Ωστόσο, κάτι τέτοιο δικαιολογείται αν αναλογιστεί κανείς ότι χρήστες με μικρότερο αριθμό ακολούθων είναι πιθανότερο να κατέχουν υψηλότερο ποσοστό ενεργού κοινού, διότι ένα μεγάλο μέρος αυτού συγκαταλέγεται στις στενές επαφές τους (οικογένεια, φίλοι).



Σχήμα 5.23: Ποσοστό αλληλεπίδρασης ανά αριθμό ακολούθων (20K-100K ακόλουθοι)



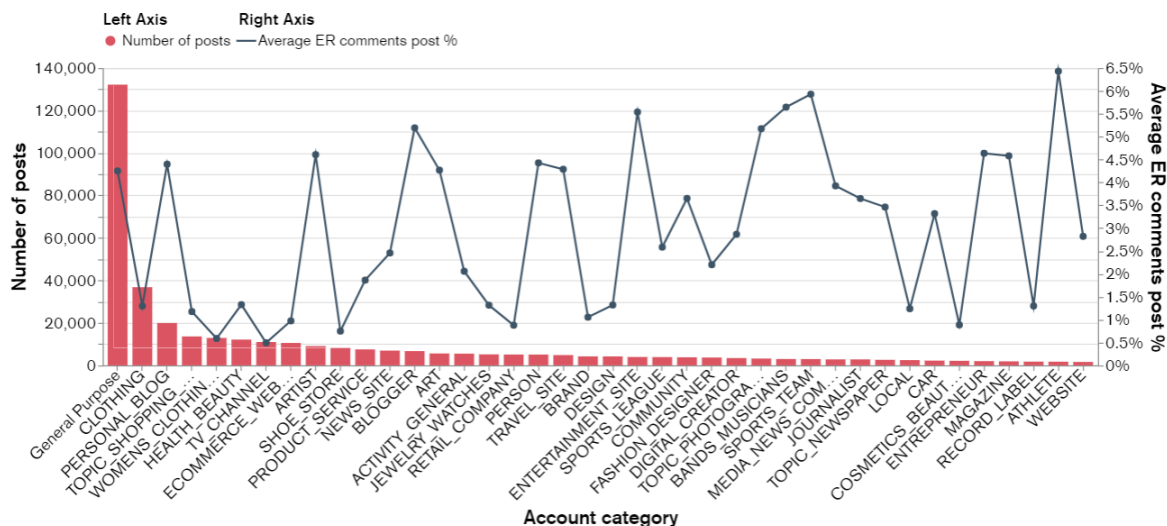
Σχήμα 5.24: Ποσοστό αλληλεπίδρασης ανά αριθμό ακολούθων (100K-280K ακόλουθοι)



Σχήμα 5.25: Ποσοστό αλληλεπίδρασης ανά αριθμό ακολούθων (ακόλουθοι > 280K)

Η παρατήρηση της εξέλιξης των τιμών τις οποίες λαμβάνει το *ER Comments Post* κατά τα υψηλότερα στρώματα ακολούθων, οδηγεί στο συμπέρασμα ότι η αύξηση του αριθμού των ακολούθων ενός λογαριασμού προκαλεί μια μικρή αύξηση του ποσοστού αλληλεπίδρασης, της τάξης του 1%-2%, η οποία συνοδεύεται από αρκετές αυξομειώσεις. Παρόλα αυτά, διαφαίνεται πως οι υψηλότερες τιμές της μεταβλητής *ER Comments Post* καταγράφονται από χρήστες οι οποίοι ανήκουν σε ανώτερες βαθμίδες, όσον αφορά στον αριθμό ακολούθων τους, χωρίς αυτό να αποτελεί τον κανόνα.

Σε συνέχεια της αναζήτησης παραγόντων, οι οποίοι καθορίζουν το ποσοστό αλληλεπίδρασης ενός χρήστη με το κοινό του, ερευνήθηκαν οι θεματικές κατηγορίες λογαριασμών του Instagram που σημειώνουν τις μεγαλύτερες τιμές *ER Comments Post* κατά μέσο όρο, στη διάρκεια του 2020. Με στόχο τη μεγιστοποίηση της εγκυρότητας των συμπερασμάτων, λήφθηκαν υπόψη οι θεματικές κατηγορίες με τις περισσότερες αναρτημένες δημοσιεύσεις, έτσι ώστε κάθε μία από αυτές να διαθέτει ένα αξιόλογο δείγμα προς μελέτη.

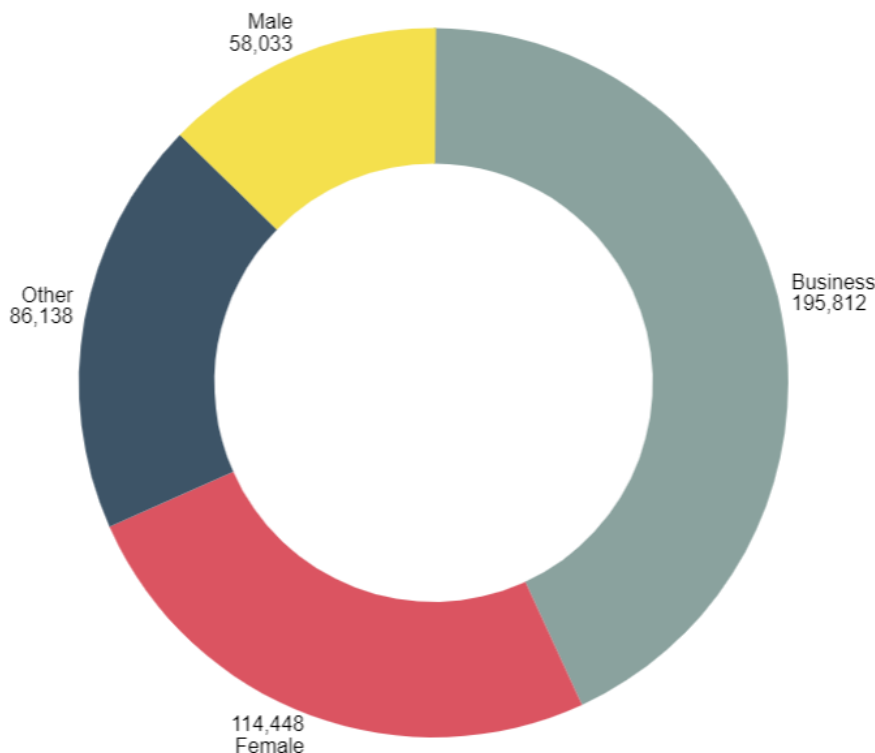


Σχήμα 5.26: Ποσοστό αλληλεπίδρασης ανά θεματική κατηγορία λογαριασμού

Στο παραπάνω σχεδιάγραμμα (βλ. Σχήμα 5.26) παρουσιάζονται ταξινομημένες, με βάση το πλήθος των αναρτήσεων, οι θεματικές κατηγορίες λογαριασμών του Instagram, συναρτήσεως του μέσου ποσοστού αλληλεπίδρασης (*ER Comments Post*) της εκάστοτε κατηγορίας. Ορίζοντας ως κατώτατη τιμή της μεταβλητής *ER Comments Post* το 4%, καθίσταται δυνατό να αναδειχθούν οι κατηγορίες, οι οποίες σημείωσαν την υψηλότερη διάδραση με το κοινό.

Με γνώμονα το πλήθος των δημοσιεύσεων, λογαριασμοί οι οποίοι είναι επικεντρωμένοι στην ένδυση και τον καλλωπισμό εμφάνισαν αισθητή δραστηριότητα, σε σχέση με τις λοιπές θεματικές κατηγορίες. Αναφορικά με το ποσοστό αλληλεπίδρασης, λογαριασμοί, των οποίων το περιεχόμενο σχετίζεται με (i) προσωπική καθημερινότητα, (ii) τέχνες, (iii) ταξιδιωτικό περιεχόμενο, (iv) άθληση και (v) ψυχαγωγία είναι αυτοί που κέντρισαν περισσότερο το ενδιαφέρον των Ελλήνων χρηστών του Instagram, κάτι το οποίο μπορεί να επαληθευτεί συγκρίνοντας τα προηγούμενα αποτελέσματα με εκείνα των Σχημάτων 5.2, 5.3, 5.4, 5.5, αλλά και 5.9.

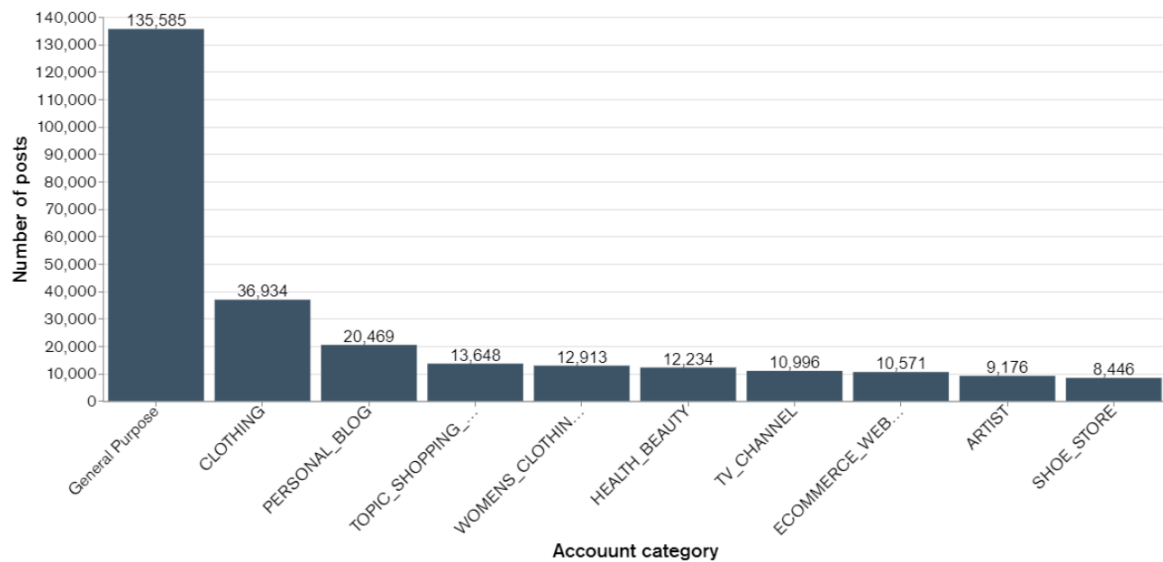
Έναν από τους κεντρικούς θεματικούς άξονες της μελέτης του Instagram αποτέλεσε, επίσης, η εξέταση του αριθμού δημοσιεύσεων και της συχνότητας αναρτήσεως τους από τις διαφορετικές κατηγορίες λογαριασμών, καθώς και η αναζήτηση κάποιας ενδεχόμενης σχέσης εξάρτησης μεταξύ αυτών και του ποσοστού αλληλεπίδρασης.



Σχήμα 5.27: Αριθμός δημοσιεύσεων ανά κατηγορία χρηστών

Ως πρώτο βήμα της προκειμένης διαδικασίας, μελετήθηκε ο συνολικός αριθμός δημοσιεύσεων, κατά τη χρονική περίοδο 2020-2021, ανά κατηγορία χρηστών και ανά κατηγορία θεματικού περιεχομένου. Σύμφωνα με το Σχήμα 5.27, παρατηρείται πως οι επιχειρησιακοί λογαριασμοί ανέπτυξαν τη ψηφιακή τους παρουσία στη πλατφόρμα του Instagram, διατηρώντας αρκετά υψηλότερη δραστηριότητα σε σχέση με τις υπόλοιπες κατηγορίες χρηστών. Από την άλλη μεριά, παρά το γεγονός ότι οι λογαριασμοί οι οποίοι διαχειρίζονται από τους άνδρες είναι σχεδόν ισάριθμοι με αυτών των επιχειρήσεων, κατέλαβαν την τελευταία θέση της κατάταξης, γεγονός το οποίο υποδηλώνει την μεγαλύτερη αδράνεια που τους διακατέχει ως ομάδα χρηστών. Σε αντίθεση με τους άνδρες, οι γυναίκες συναντούνται στη θέση με τον δεύτερο υψηλότερο αριθμό δημοσιεύσεων, για το 2020, κάνοντας έτσι ιδιαίτερα αισθητή τη παρουσία τους στο Instagram.

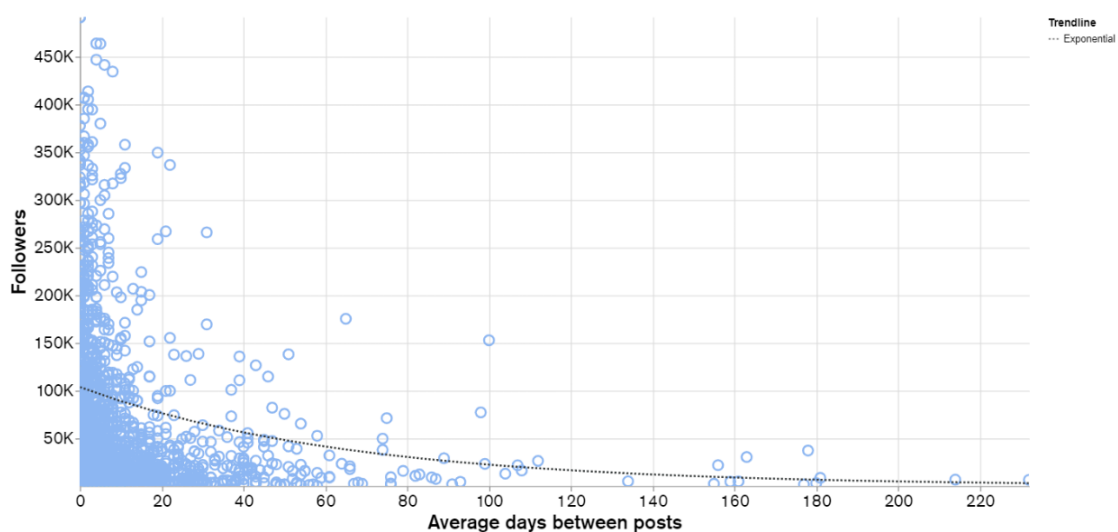
Στη συνέχεια, εφαρμόζοντας το ίδιο κριτήριο ταξινόμησης για τις ποικίλες θεματικές κατηγορίες λογαριασμών, είναι εμφανές στο Σχήμα 5.28 πως, αν εξαιρεθούν οι λογαριασμοί γενικού περιεχομένου (General Purpose), τη πλειοψηφία των ανώτερων θέσεων της κατάταξης απαρτίζουν χρήστες των οποίων το περιεχόμενο σχετίζεται με την παρουσία προϊόντων, και ειδικότερα προϊόντων ένδυσης (CLOTHING, TOPIC_SHOPPING_RETAIL, WOMEN_CLOTHING_STORE, ECOMMERCE_WEBSITE, SHOE_STORE), ενώ, για ακόμη μία φορά, σε υψηλά θέσεις διατηρούνται λογαριασμοί με θεματικό περιεχόμενο προσωπικής καθημερινότητας, υγείας/ομορφιάς και τεχνών, μαρτυρώντας έτσι την ανεπτυγμένη δραστηριοποίηση τους στη πλατφόρμα του Instagram.



Σχήμα 5.28: Αριθμός δημοσιεύσεων ανά θεματική κατηγορία λογαριασμού

Με σκοπό τη μελέτη της συχνότητας αναρτήσεων και του αντίκτυπου που πιθανώς αυτή έχει στις μετρικές της έρευνας, υπολογίστηκε για τον εκάστοτε λογαριασμό του δείγματος το μέσο διάστημα το οποίο μεσολαβεί μεταξύ δύο διαδοχικών δημοσιεύσεων του. Πρέπει να επισημανθεί πως η προαναφερθείσα τεχνική αποτύπωσης της συχνότητας αναρτήσεων δε λαμβάνει υπόψη το πλήθος των συνολικών δημοσιεύσεων, αλλά μόνο το μέσο διάστημα μεταξύ αυτών. Για παράδειγμα, έστω ότι ένας λογαριασμός είχε κοινοποιήσει συνολικά δύο αναρτήσεις κατά τη διάρκεια του 2020 με διαφορά μίας ημέρας μεταξύ τους, τότε προκύπτει ότι ο συγκεκριμένος ο χρήστης δημιουργεί περιεχόμενο, κατά μέσο όρο, σε καθημερινή βάση, διατηρώντας όμως στην πραγματικότητα πολύ χαμηλή δραστηριότητα.

Η πρώτη προς εξέταση σχέση επιλέχθηκε να είναι εκείνη, η οποία σκιαγραφείται μεταξύ του αριθμού των ακολούθων ενός λογαριασμού και της συχνότητας αναρτήσεων του.

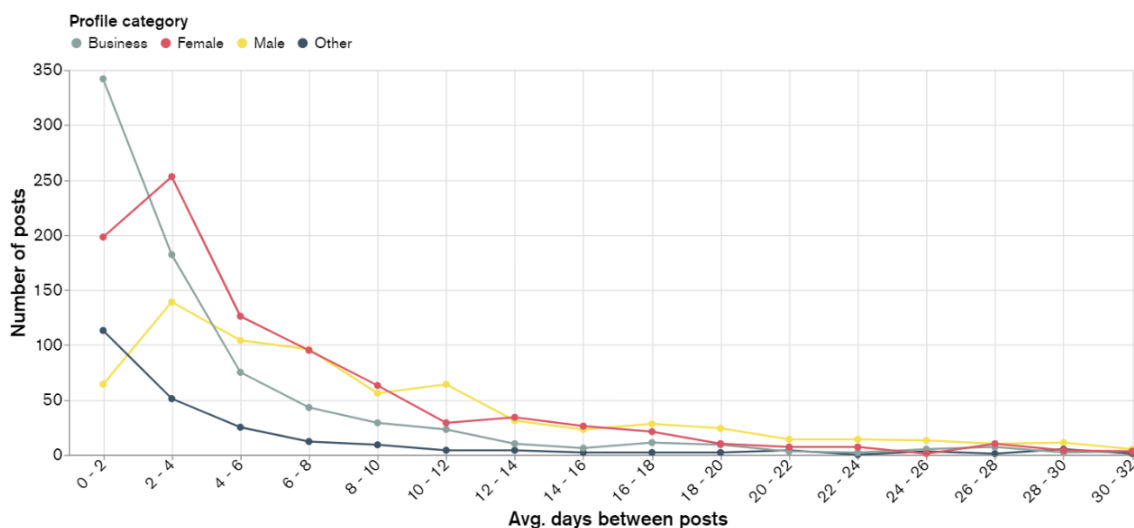


Σχήμα 5.29: Σχέση μεταξύ αριθμού ακολούθων-συχνότητα αναρτήσεων

Παρακολουθώντας τη γραμμή τάσης του Σχήματος 5.29, γίνεται αντιληπτό πως οι λογαριασμοί οι οποίοι συγκεντρώνουν υψηλό αριθμό ακολούθων δημοσιοποίησαν περιεχόμενο αρκετά συχνότερα, κατά τη διάρκεια του 2020, σε σχέση με λογαριασμούς μικρότερης εμβέλειας. Πιο συγκεκριμένα, μεταβαίνοντας στα ανώτερα στρώματα ακολούθων φαίνεται το μέσο διάστημα μεταξύ των αναρτήσεων να τείνει στον κατακόρυφο άξονα, δηλαδή στη καθημερινή κοινοποίηση περιεχομένου. Το γεγονός αυτό, είναι πιθανό να συνδέεται με την προσπάθεια των μεγαλύτερων προφίλ να συντηρήσουν την έντονη δραστηριότητα τους και, συμπερασματικά, να διαφυλάξουν την παρουσία τους στο προσκήνιο των ψηφιακών δρώμενων.

Με στόχο τη βέλτιστη αναπαράσταση των παραπάνω δεδομένων, αποφασίστηκε το δείγμα να περιοριστεί στους λογαριασμούς με αριθμό ακολούθων μικρότερο ή ίσο των 500.000, καλύπτοντας με αυτό τον τρόπο ένα σεβαστό ποσοστό λογαριασμών και αποφεύγοντας την μεγαλύτερη διασπορά των τιμών στα υψηλότερα στρώματα ακολούθων.

Προς επαύξηση του παραπάνω συμπεράσματος, είναι δυνατό να διερευνηθεί η συχνότητα ανάρτησης δημοσιεύσεων ανά τις διαφορετικές κατηγορίες χρηστών.



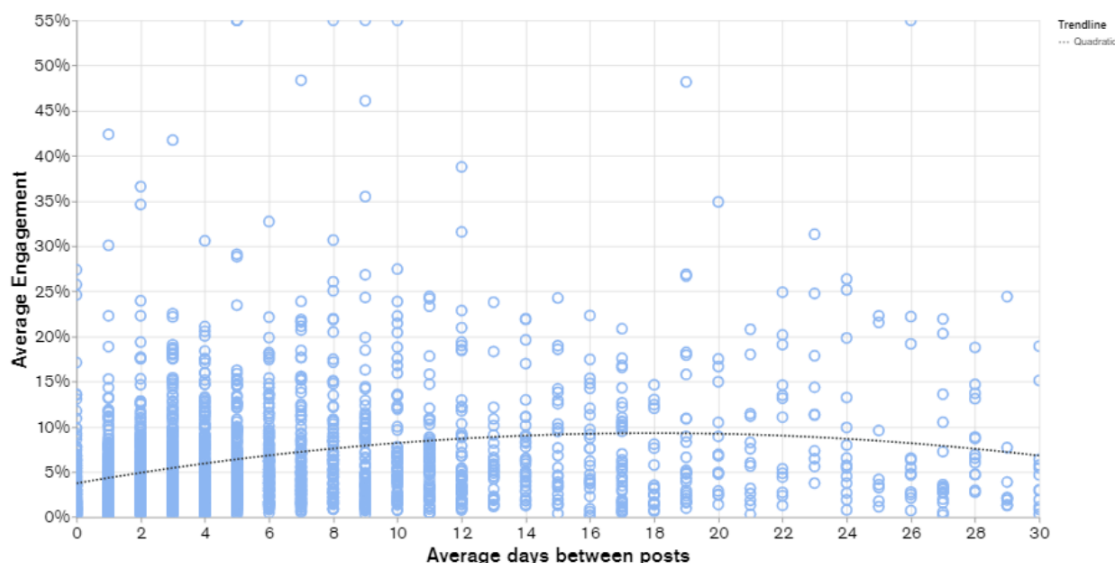
Σχήμα 5.30: Συχνότητα ανάρτησης δημοσιεύσεων ανά κατηγορία χρηστών

Βάσει του Σχήματος 5.30, διακρίνεται πως οι επιχειρησιακοί λογαριασμοί και τα προφίλ τύπου Άλλο προτίμησαν, κατά το 2020, να κοινοποιούν περιεχόμενο αρκετά συχνά (καθημερινά έως και κάθε δύο ημέρες), ενώ οι άνδρες και οι γυναίκες υιοθέτησαν μία λιγότερη έντονη δραστηριότητα, με το μέσο διάστημα μεταξύ των δημοσιεύσεων τους να κυμαίνεται από δύο έως τέσσερις ημέρες. Παρόλα αυτά, κοινή συμπεριφορά όλων των κατηγοριών χρηστών αποτέλεσε η αποφυγή μεγαλύτερων περιόδων αδράνειας ενδιάμεσα των διαδοχικών αναρτήσεων.

Εφόσον πραγματοποιήθηκε μια στοιχειώδης διερεύνηση του πλήθους των δημοσιεύσεων και της συχνότητας ανάρτησης τους, τέθηκαν τα θεμέλια για την αναζήτηση πιθανών συσχετίσεων μεταξύ των παραπάνω μεγεθών του ποσοστού αλληλεπίδρασης ενός χρήστη του Instagram με τους ακολούθους του.

Ξεκινώντας με τη συχνότητα κοινοποίησης περιεχομένου, η ζητούμενη σχέση εξετάστηκε με

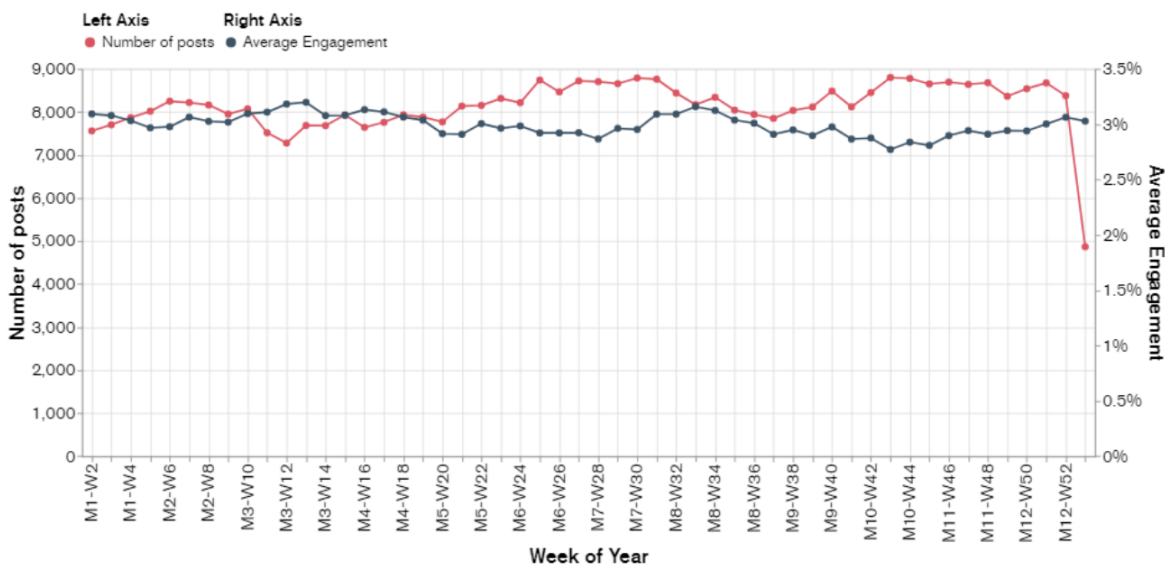
βάση τους λογαριασμούς, οι οποίοι αναρτούν υλικό κατά μέσο όρο εντός 30 ημερών από την προηγούμενη τους δημοσίευση. Από το συγκεκριμένο υποσύνολο λογαριασμών αξιοποιήθηκαν όλα τα προφίλ, ανεξαρτήτως του μέσου ποσοστού αλληλεπίδρασης τους. Ωστόσο, για λόγους παρουσίασης, το αντίστοιχο σχεδιάγραμμα εστίασε στην κλίμακα 0%-55% ποσοστού αλληλεπίδρασης.



Σχήμα 5.31: Σχέση μεταξύ συχνότητας αναρτήσεων-ποσοστού αλληλεπίδρασης

Μέσω της ανάλυσης της γραμμής τάσεως του Σχήματος 5.31, σχηματίζεται ένα πρώιμο συμπέρασμα, κατά το οποίο λογαριασμοί οι οποίοι επέλεξαν να ακολουθήσουν έναν πιο ήπιο ρυθμό δημοσίευσης περιεχομένου (≈ 1 δημοσίευση/14 ημέρες) συγκέντρωσαν μεγαλύτερο ποσοστό αλληλεπίδρασης συγκριτικά με προφίλ τα οποία κατέκλυσαν την πλατφόρμα με δημοσιεύσεις καθημερινά. Αντίστοιχο αρνητικό αντίκτυπο στο ποσοστό αλληλεπίδρασης φαίνεται να είχε και το πέρας εκτενούς χρονικού διαστήματος μεταξύ των διαδοχικών δημοσιεύσεων ενός χρήστη.

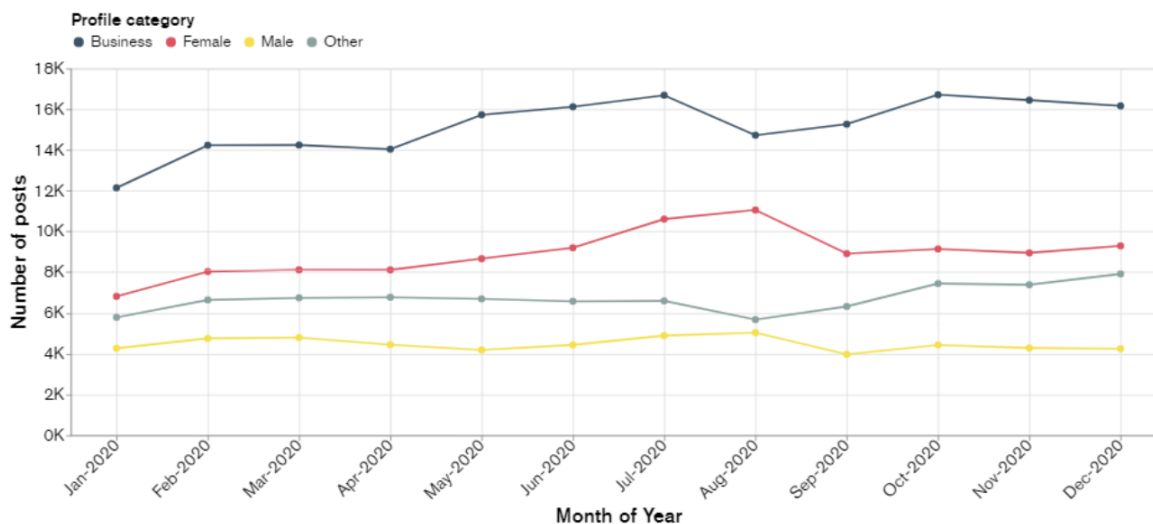
Μια εναλλακτική προσέγγιση της συσχέτισης μεταξύ ποσοστού αλληλεπίδρασης και ρυθμού ανάρτησης περιεχομένου είναι δυνατό να παρέχει η απεικόνιση του αριθμού δημοσιεύσεων ανά εβδομάδα του έτους 2020, σε συνδυασμό με τη μέση τιμή της μεταβλητής *ER Comments Post*, η οποία σημειώθηκε κατά τις αντίστοιχες χρονικές περιόδους. Η σχέση αυτή αναπαρίστανται στο Σχήμα 5.31, όπου η κάθε εβδομάδα αναγράφεται με ένα αναγνωριστικό της μορφής *MonthX-WeekX*.



Σχήμα 5.32: Αριθμός δημοσιεύσεων και μέσου ποσοστού αλληλεπίδρασης ανά εβδομάδα του 2020

Ξεκινώντας την ανάλυση του διαγράμματος με βάση τον αριθμό δημοσιεύσεων ανά εβδομάδα, διαπιστώνεται πως την ομαλή αύξηση των αναρτήσεων κατά την περίοδο του Ιανουαρίου-Φεβρουαρίου διαδέχθηκε μια απότομη πτώση στον αριθμό τους με τη χαμηλότερη τιμή να αγγίζεται κατά τη δεύτερη εβδομάδα του Μαρτίου. Η συγκεκριμένη πτώση είναι πιθανό να οφείλεται στην έναρξη του πρώτου κύματος της πανδημίας του *COVID-19* στην Ελλάδα, τον Μάρτιο του 2020 και, ίσως, μεταφράζεται στην επίπτωση της πρώτης καραντίνας στη ψυχολογία των Ελλήνων χρηστών του Instagram. Η ακόλουθη περίοδος, από τα μέσα Μαρτίου έως τα μέσα Μαΐου, χαρακτηρίστηκε από μία μικρή επάνοδο και σταθεροποίηση του αριθμού αναρτήσεων, η οποία ωστόσο διατηρήθηκε σε σχετικά χαμηλά επίπεδα.

Σε αντίθεση με τις προγενέστερες χρονικές περιόδους, σημειώθηκε μεγάλη αύξηση των κοινοποιημένων αναρτήσεων κατά το χρονικό διάστημα τέλη Μαΐου-τέλη Ιουλίου, φαινόμενο το οποίο ενδεχομένως να απορρέει από την έναρξη της καλοκαιρινής περιόδου και την παράλληλη σταδιακή άρση των περιοριστικών μέτρων. Παρόλα αυτά, ο Αύγουστος χαρακτηρίστηκε στο γενικό σύνολο των λογαριασμών από ακόμη μία πτώση στον αριθμό των δημοσιεύσεων. Με σκοπό τη διαλεύκανση της προαναφερθείσας παρατήρησης, στάθηκε χρήσιμος ο έλεγχος του πλήθους αναρτήσεων ανά μήνα του 2020 για τις διάφορες κατηγορίες χρηστών.



Σχήμα 5.33: Αριθμός δημοσιεύσεων ανά μήνα και κατηγορία χρηστών

Όπως γίνεται αντιληπτό από το Σχήματα 5.33, η πτώση του γενικού αριθμού των αναρτήσεων επηρεάστηκε από τα προφίλ τύπου Άλλο και τους επιχειρησιακούς λογαριασμούς, οι οποίοι φαίνεται να διατήρησαν μικρότερη δραστηριότητα στην καρδιά της θερινής τουριστικής περιόδου του έτους 2020. Αντιθέτως, ο αριθμός αναρτήσεων των γυναικών και των ανδρών του Instagram αυξήθηκε σταθερά καθ' όλη τη διάρκεια του καλοκαιριού, με την κορύφωσή του να επέρχεται τον Αύγουστο, υποδηλώνοντας κατά αυτόν τον τρόπο την υψηλότερη διάθεση και ευχέρεια χρόνου του συγκεκριμένου υποσυνόλου χρηστών για περιήγηση στη πλατφόρμα του Instagram κατά τη περίοδο αυτή.

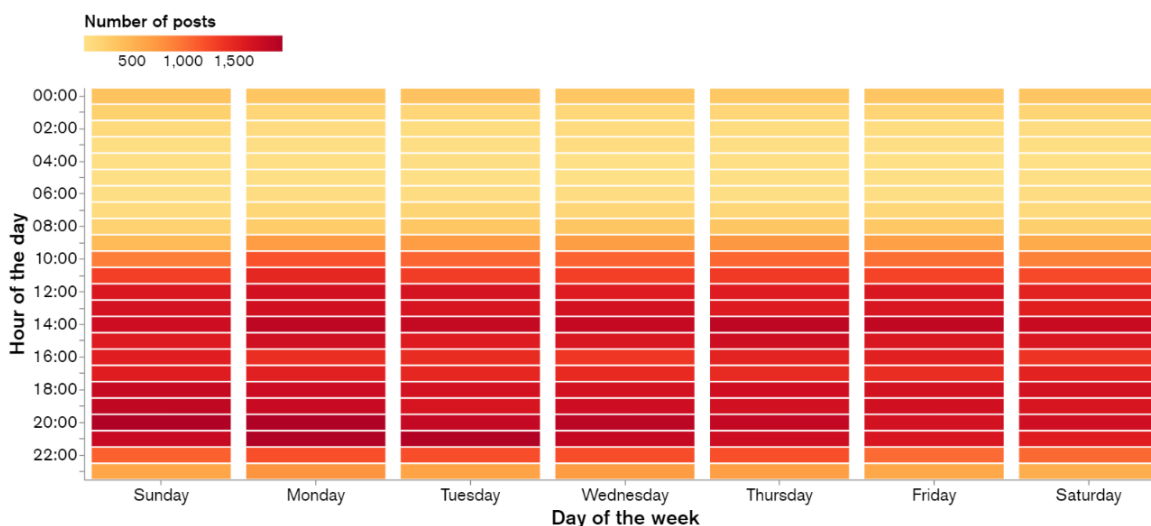
Η μετάβαση προς τη φθινοπωρινή-χειμερινή περίοδο συνοδεύτηκε από την επιστροφή του γενικού αριθμού αναρτήσεων σε υψηλά επίπεδα, διατηρώντας τον ίδιο ρυθμό μέχρι την εκπνοή του 2020. Η οπτική αυτή, όμως, καθορίστηκε από τη συμπεριφορά των επιχειρησιακών λογαριασμών και των προφίλ τύπου Άλλο, τα οποία επανήλθαν δυναμικά στην ενεργό δράση μετά το τέλος του καλοκαιριού, ενώ οι άνδρες και οι γυναίκες επέστρεψαν στους φυσιολογικούς, για τα δεδομένα τους, ρυθμούς δραστηριότητας στο Instagram. Συνολικά, λαμβάνοντας υπόψη πως κατά τη διάρκεια του φθινοπώρου του 2020 επανατέθηκαν σε ισχύ τα περιοριστικά μέτρα αντιμετώπισης του *COVID-19*, παρά το γεγονός αυτό, φαίνεται το κοινό να παρέμεινε περισσότερο δραστήριο αναλογικά με την περίοδο της πρώτης καραντίνας, μαρτυρώντας κατά ένα ποσοστό τη προσφυγή του κοινού στα μέσα κοινωνικής δικτύωσης εν καιρώ εγκλεισμού.

Συνεχίζοντας με την ανάλυση του ποσοστού αλληλεπίδρασης ανά εβδομάδα, είναι εμφανές πως αυτό δε μεταβλήθηκε ραγδαία κατά τη διάρκεια του 2020, αλλά εμφάνισε μία ελαφρώς πτωτική πορεία, η οποία αντικρούστηκε από συγκεκριμένες περιόδους ανάκαμψης. Πιο συγκεκριμένα, οι μήνες Μάρτιος, Αύγουστος και Δεκέμβριος, περίοδοι οι οποίοι ενείχαν την πρώτη καραντίνα, την κορύφωση της θερινής τουριστικής περιόδου και τις γιορτές των Χριστουγέννων, αποτέλεσαν τα διαστήματα των μεγίστων τιμών του ποσοστού αλληλεπίδρασης του κοινού με το αναρτημένο περιεχόμενο. Γενικότερα, η σύγκριση της πορείας των δύο μεταβλητών του Σχήματος 5.32 οδηγεί στην παρατήρηση πως περίοδοι αύξησης του αριθμού κοινοποιήσεων συνοδεύονται από σχετική μείωση του μέσου ποσοστού αλληλεπίδρασης και το αντίστροφο. Διασταυ-

ρώνοντας τη διαπίστωση αυτή με τα συμπεράσματα των Σχημάτων 5.30 και 5.31 ενισχύεται η ένδειξη πως οι Έλληνες χρήστες του Instagram ανταποκρίνονται καλύτερα σε ηπιότερους ρυθμούς ανάρτησης δημοσιεύσεων, με την υπερβολικά υψηλή δραστηριότητα να δημιουργεί κορεσμό στο κοινό, ενώ με την αρκετά χαμηλή δραστηριότητα να εγκυμονεί την εξασθένηση του ενδιαφέροντος των ακολούθων ενός λογαριασμού.

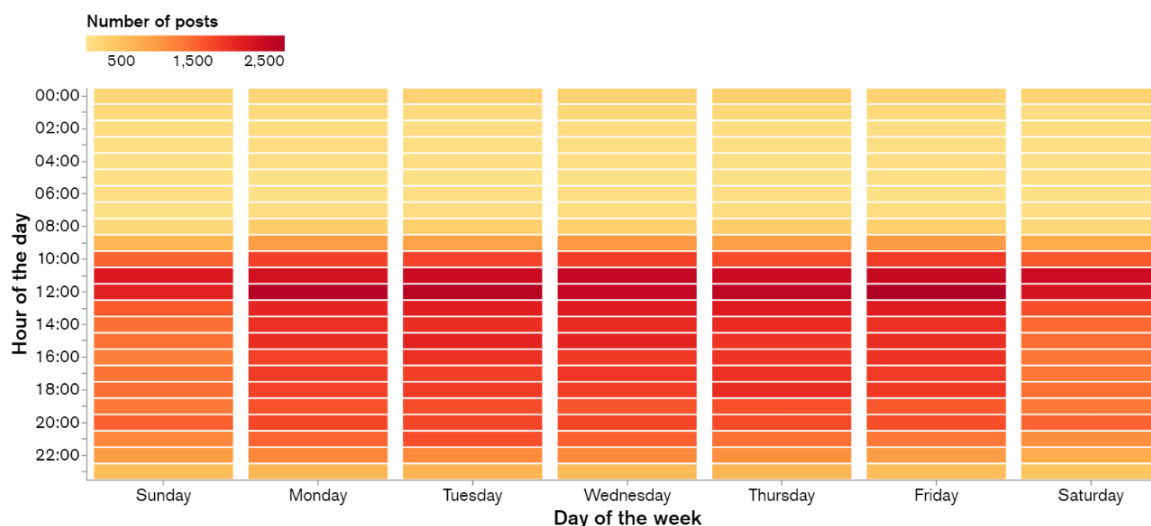
Η ανάγκη για μία βαθύτερη οπτική της συμπεριφοράς του ελληνικού κοινού, μετέφερε το ερευνητικό ενδιαφέρον από το επίπεδο μηνών σε επίπεδο ημερών και συγκεκριμένα στη διαλεύκανση των ημερών και ωρών του 2020, κατά τις οποίες καταγράφηκε η υψηλότερη δραστηριότητα ανά κατηγορία χρηστών.

Με αφετηρία τους άνδρες και τις γυναίκες, οι οποίοι παρουσίασαν κοινές προτιμήσεις όσον αφορά στις καθημερινές ώρες ανάρτησης δημοσιεύσεων, είναι φανερό από το Σχήμα 5.32 πως αυτοί διατήρησαν σταθερή δραστηριότητα καθ' όλη τη διάρκεια της εβδομάδας, με το χρονικό διάστημα το οποίο συγκέντρωσε την υψηλότερη κίνηση καθημερινά να βρίσκεται μεταξύ 10:00 π.μ. και 10:00 μ.μ.. Εστιάζοντας, παρατηρούνται δύο ζώνες αιχμής κατά τη διάρκεια της ημέρας, εκείνες των 12:00 μ.μ.-02:00 μ.μ. και των 06:00 μ.μ.-09:00 μ.μ..



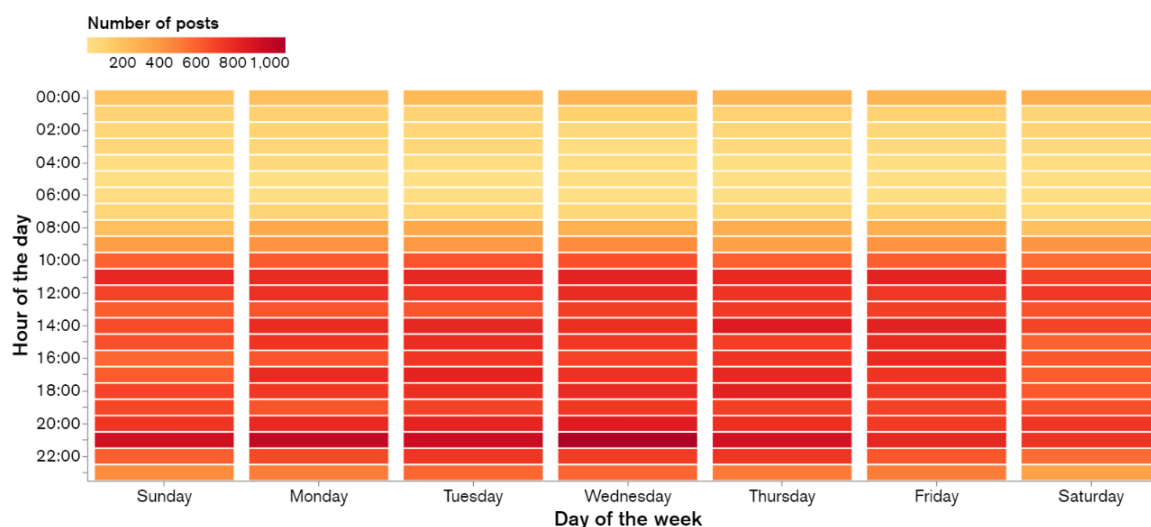
Σχήμα 5.34: Αριθμός δημοσιεύσεων ανά μέρα/ώρα (Άνδρες, Γυναίκες)

Αντίστοιχα με τους άνδρες και τις γυναίκες, η πλειοψηφία των δημοσιεύσεων, οι οποίες προήλθαν από επιχειρησιακούς λογαριασμούς, αναπτύχθηκαν κατά το χρονικό διάστημα 10:00 π.μ. έως 10:00 μ.μ. (βλ. Σχήμα 5.35). Ωστόσο, στη προκειμένη περίπτωση οι ώρες αιχμής περιορίστηκαν στις πρωινές και μεσημεριανές ώρες και συγκεκριμένα από τις 11:00 π.μ. έως τις 12:00 μ.μ., με τις απογευματινές ώρες να συγκεντρώνουν μειωμένη αλλά σεβαστή δραστηριότητα. Αξιοσημείωτο είναι το γεγονός ότι η ψηφιακή παρουσία των επιχειρήσεων στο Instagram ελαττώθηκε τα Σαββατοκύριακα, φαινόμενο που πιθανώς είναι απόρροια του καθημερινού ωραρίου λειτουργίας τους.



Σχήμα 5.35: Αριθμός δημοσιεύσεων ανά μέρα/ώρα (Επιχειρήσεις)

Καταλήγοντας στους λογαριασμούς τύπου Άλλο, διαπιστώνεται, μέσω του Σχήματος 5.36, πως εμφάνισαν παρόμοια συμπεριφορά με εκείνη των ανδρών και των γυναικών, δηλαδή η δραστηριότητα τους κυμάνθηκε σε σταθερά επίπεδα κατά τη διάρκεια ολόκληρης της εβδομάδας με μία μικρή αύξηση των δημοσιεύσεων να συναντάται κατά τις εργάσιμες μέρες και προς τις βραδινές ώρες (08:00 μ.μ-09:00 μ.μ.).



Σχήμα 5.36: Αριθμός δημοσιεύσεων ανά μέρα/ώρα (Άλλο)

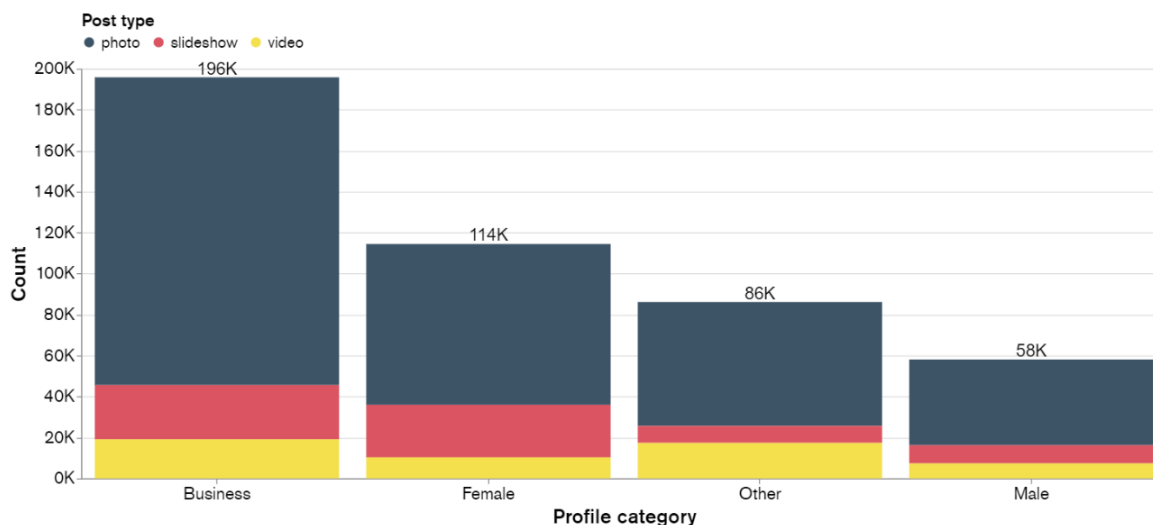
Εξίσου σημαντικό πόλο ερευνητικού ενδιαφέροντος αποτέλεσε ο έλεγχος της επιρροής την οποία ασκούν τα ποικίλα είδη δημοσιεύσεων, τεχνικά και θεματικά, στους Έλληνες χρήστες του Instagram.

Με τον όρο *τεχνικά* πραγματοποιείται η διάκριση των δημοσιεύσεων, με βάση τη τεχνική αναπαράστασης του περιεχομένου τους. Όπως προαναφέρθηκε στο Κεφάλαιο 4, στην Υποενότητα

4.1.1, τα είδη των αναρτήσεων, τα οποία μελετήθηκαν στα πλαίσια της παρούσας έρευνας, είναι τα: (i) φωτογραφία, (ii) βίντεο μικρής διάρκειας, (iii) βίντεο μεγάλης διάρκειας και (iv) παρουσίαση πολλαπλού περιεχομένου (slideshow). Διευκρινίζεται πως τα βίντεο μικρής και μεγάλης διάρκειας λήφθηκαν υπόψη από κοινού, ως μία ενιαία κατηγορία αναρτήσεων.

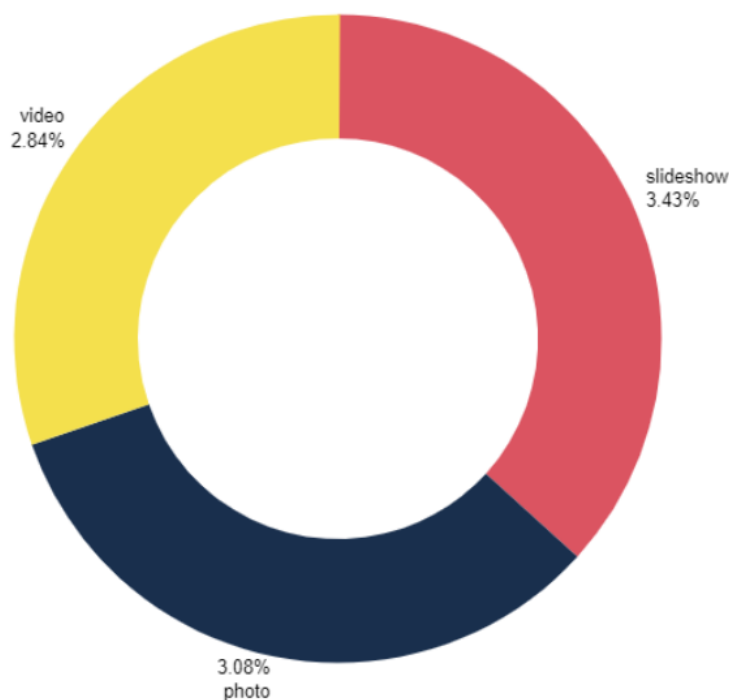
Πρωτίστως, θεωρήθηκε απαραίτητος ο υπολογισμός του ποσοστού χρήσης του εκάστοτε τύπου δημοσίευσης από τις διαφορετικές κατηγορίες χρηστών. Σύμφωνα με το Σχήμα 5.37, προκύπτει το συμπέρασμα ότι, στο συνολικό φάσμα των κατηγοριών χρηστών, περισσότερο διαδεδομένες είναι η δημοσιεύσεις τύπου *φωτογραφία* με μεγάλη διαφορά από τις υπόλοιπες, ενώ αμέσως επόμενες στη κατάταξη των προτιμήσεων των χρηστών βρίσκονται οι δημοσιεύσεις τύπου *slideshow*. Με μία πιο προσεκτική ματιά, παρατηρείται ότι οι επιχειρησιακοί λογαριασμοί και οι γυναίκες καταλαμβάνουν το υψηλότερο ποσοστό χρήσης αναρτήσεων τύπου *slideshow*, ενώ ιδιαίτερη προτίμηση προς τα βίντεο δείχνουν οι επιχειρήσεις και τα προφίλ τύπου Άλλο.

Η υπεροχή των φωτογραφιών, πιθανώς, οφείλεται στον απλούστερο και λιγότερο χρονοβόρο τρόπο δημιουργίας και ανάρτησης ο οποίος τις διέπει. Από την άλλη μεριά, η αισθητή παρουσία των βίντεο σε ορισμένες κατηγορίες χρηστών, ενδεχομένως, μαρτυρούν το όφελος το οποίο αποσπούν οι τελευταίες από την αυξημένη παραστατικότητα, η οποία χαρακτηρίζει τέτοιου είδους αναρτήσεις.



Σχήμα 5.37: Είδη δημοσίευσης ανά κατηγορία χρηστών

Ακολούθως, είναι δυνατό να προσδιοριστεί η απήχηση των παραπάνω τύπων δημοσιεύσεων στο κοινό, μετρώντας τη μέση τιμή του *ER Comments Post* για κάθε μία από αυτές.



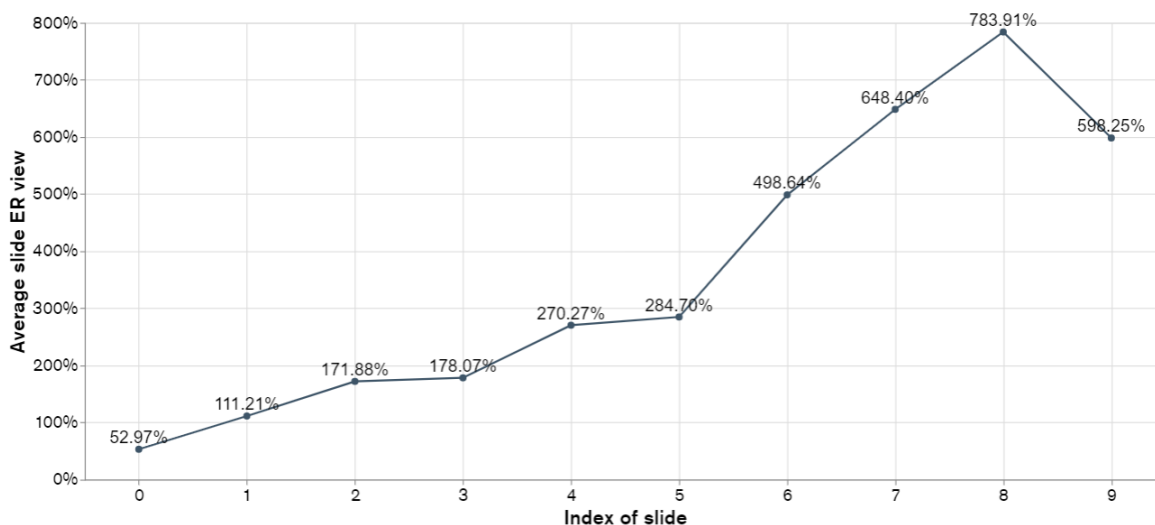
Σχήμα 5.38: Ποσοστό αλληλεπίδρασης ανά κατηγορία δημοσιεύσεων

Παρά το γεγονός ότι το μεγαλύτερο ποσοστό χρήσης αφορά σε φωτογραφικές δημοσιεύσεις, φαίνεται από το Σχήμα 5.38 ότι το υψηλότερο ποσοστό αλληλεπίδρασης σημειώθηκε, κατά μέσο όρο, από δημοσιεύσεις τύπου *slideshow*, φανερώνοντας έτσι την αποδοτικότερη προσέγκυση αντιδράσεων από τέτοιου είδους δημοσιεύσεις. Ωστόσο, η τιμή του ποσοστού αλληλεπίδρασης δε μεταβλήθηκε σημαντικά μεταξύ των τριών κατηγοριών αναρτήσεων, καθώς παρέμεινε κοντά στο 3% για την εκάστοτε κατηγορία.

Επικεντρώνοντας τη προσοχή στις δημοσιεύσεις τύπου *slideshow*, το ενδιαφέρον στράφηκε στην αναζήτηση κάποιας σχέσης μεταξύ της θέσης μιας εικόνας ή ενός βίντεο στη συλλογή περιεχομένων και της επιρροής την οποία αυτή ή αυτό, αντίστοιχα, ασκεί στο κοινό. Η τακτική του Instagram, κατά την οποία οι θετικές αντιδράσεις μιας δημοσίευσης τύπου *slideshow* συλλέγονται μαζικά για όλες τις φωτογραφίες ή βίντεο που συμπεριλαμβάνονται στη συλλογή, ανεξαρτήτως του πλήθους χρηστών το οποίο προσέπελασε τη κάθε μία από αυτές, επιβάρυνε την επίτευξη του παραπάνω στόχου. Καθοριστική υπήρξε η χρήση της μετρικής *ER Views*, βάσει της οποίας προσμετράται το ποσοστό θετικών αντιδράσεων αναλογικά με τον συνολικό αριθμό προβολών ενός βίντεο του Instagram (βλ. Κεφάλαιο 4, Υποενότητα 4.1.1). Το γεγονός ότι κάθε βίντεο το οποίο ανήκει σε μία δημοσίευση τύπου *slideshow* διατηρεί τον αποκλειστικό αριθμό των προβολών του παρά το μαζικό συνυπολογισμό των θετικών αντιδράσεων, σε συνδυασμό με τη χρήση της μεταβλητής *ER Views*, οδήγησε στην έμμεση εκτίμηση του ποσοστού παρακολούθησης ενός βίντεο σε σχέση με τη θέση του στην αντίστοιχη συλλογή.

Αναλυτικότερα, έστω ότι μια συλλογή *slideshow* εμπεριέχει δύο βίντεο και έχει αποσπάσει 100 θετικές αντιδράσεις συνολικά. Μετρώντας τη τιμή του *ER Views* για κάθε ένα από αυτά, προέκυψε ότι, παραδείγματος χάριν, το πρώτο συγκέντρωσε μικρότερο ποσοστό αλληλεπίδρασης από το δεύτερο. Στο σημείο αυτό υπενθυμίζεται ότι η μετρική *ER Views* υπολογίζεται ως συνολι-

κός αριθμός θετικών αντιδράσεων προς τον αριθμό προβολών βίντεο και ότι τα δύο βίντεο της συγκεκριμένης συλλογής διαθέτουν από κοινού τον ίδιο αριθμό θετικών αντιδράσεων, μέσω των οποίων εξάγεται το συμπέρασμα ότι το δεύτερο βίντεο της συλλογής παρακολούθηθηκε από λιγότερους χρήστες, συγκριτικά με το πρώτο.



Σχήμα 5.39: Σχέση ποσοστού ER Views και θέσης του βίντεο σε μία δημοσίευση τύπου *slideshow*

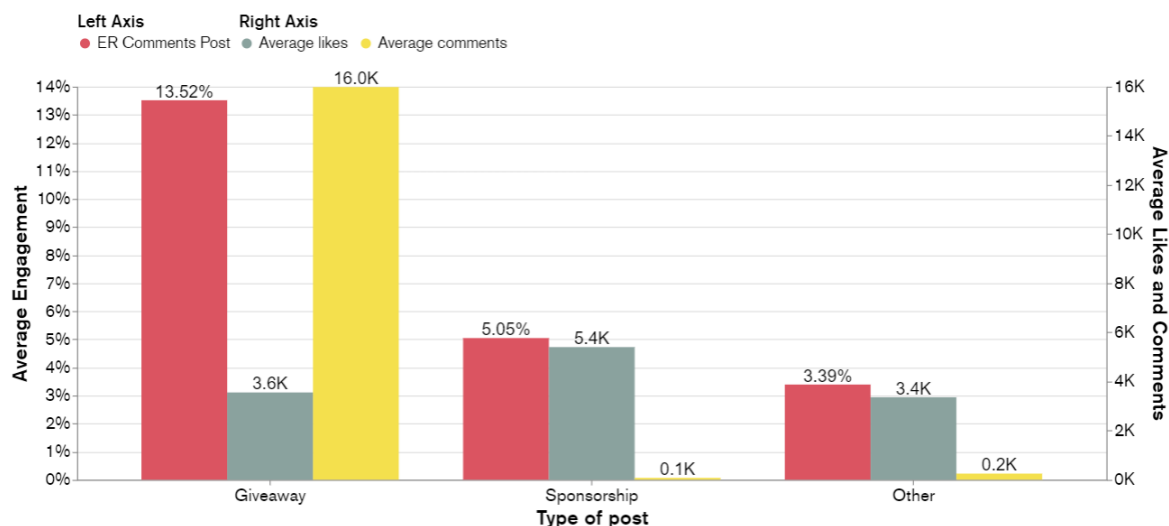
Η εφαρμογή αντίστοιχου συλλογισμού με εκείνον του παραπάνω παραδείγματος, κατά τη μελέτη του Σχήματος 5.39, συνέβαλε στη διαπίστωση ότι βίντεο τα οποία βρίσκονταν σε μεγαλύτερο βάθος μίας δημοσίευσης τύπου *slideshow* έλαβαν λιγότερες προβολές σε σχέση με βίντεο που βρίσκονταν στη κορυφή της συλλογής. Επεκτείνοντας αυτή τη παρατήρηση, συνάγεται το συμπέρασμα ότι φωτογραφίες ή βίντεο μιας συλλογής τα οποία βρίσκονται σε μεγαλύτερο βάθος λαμβάνουν αναλογικά λιγότερη προσοχή από το κοινό, με το χάσμα να αυξάνεται σημαντικά σε δημοσιεύσεις τύπου *slideshow* μεγέθους μεγαλύτερου των πέντε διαφανειών.

Η θεματική διάκριση των δημοσιεύσεων πραγματοποιήθηκε στη βάση των δημοφιλών θεματικών κατηγοριών περιεχομένου, οι οποίες στηρίζονται στη τοποθέτηση προϊόντων και, κατά συνέπεια, επιτρέπουν διερεύνηση της επιρροής των επιχειρήσεων στο ελληνικό κοινό του Instagram. Οι δημοσιεύσεις κατηγοριοποιήθηκαν σε τρεις επιμέρους ομάδες, οι οποίες είναι οι εξής:

- **Διαγωνισμός (Giveaway):** τύπος ανάρτησης στα πλαίσια της οποίας ο διαχειριστής του λογαριασμού διοργανώνει έναν διαγωνισμό, καλώντας τους χρήστες του Instagram να αντιδράσουν στη συγκεκριμένη δημοσίευση ώστε να διεκδικήσουν το διαφημιζόμενο προϊόν
- **Χορηγία (Sponsorship):** τύπος ανάρτησης στα πλαίσια της οποίας ο διαχειριστής του λογαριασμού συνεργάζεται με μία επιχείρηση, με σκοπό την προώθηση κάποιου προϊόντος ή υπηρεσίας της
- **Άλλο (Other):** κατηγορία η οποία περιλαμβάνει όλες τις δημοσιεύσεις που δεν κατατάσσονται σε κάποιον από τους παραπάνω τύπους αναρτήσεων

Η αναζήτηση και κατηγοριοποίηση των δημοσιεύσεων επετεύχθη με τη χρήση των *hashtags*,

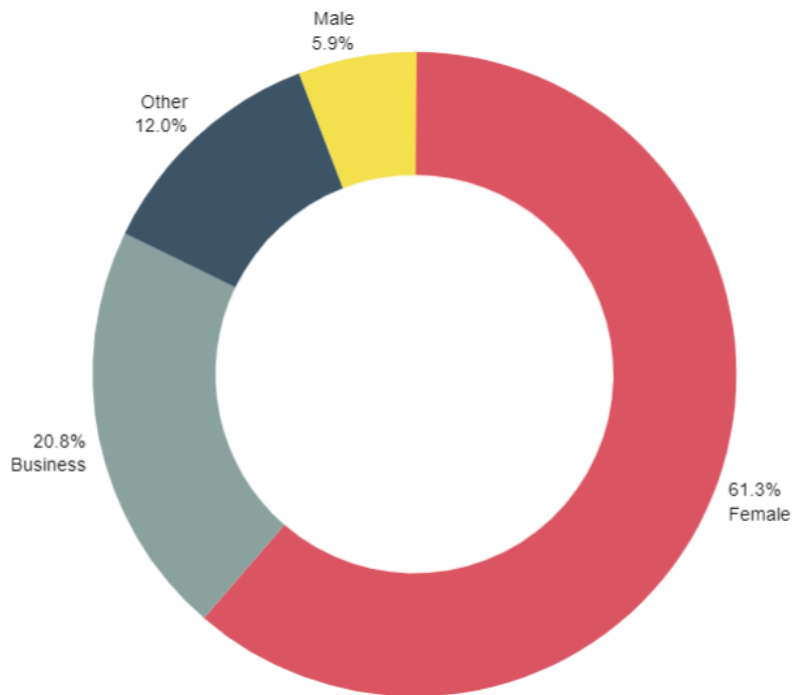
όπως για παράδειγμα #giveaway και #ad. Ωστόσο, το γεγονός ότι μία μερίδα διαγωνισμών και χορηγούμενων αναρτήσεων δεν χρησιμοποιούν *hashtags* καθιστά αναπόφευκτη την διαστρέβλωση της πραγματικής εικόνας των στατιστικών, με αρκετές δημοσιεύσεις να συγκαταλέγονται εσφαλμένα στη κατηγορία Άλλο. Η ύπαρξη της κατηγορίας Άλλο εξυπηρετεί τη σύγκριση της απόδοσης μεταξύ δημοσιεύσεων τοποθέτησης προϊόντων και αναρτήσεων γενικότερου ενδιαφέροντος. Η σύγκριση αυτή διεξάγεται στο Σχήμα 5.40, όπου αντιπαρατίθενται οι μετρικές απόδοσης (ER Comments Post, μέσος αριθμός αντιδράσεων, μέσος αριθμός σχολιασμών) για κάθε μία από τις παραπάνω κατηγορίες αναρτήσεων



Σχήμα 5.40: Μετρικές απόδοσης ανά θεματικό περιεχόμενο δημοσίευσης (Giveaway, Sponsorship, Other)

Με γνώμονα τις τιμές του Σχήματος 5.40, διακρίνεται πως το κοινό αλληλεπίδρασε, κατά το 2020, ευκολότερα με δημοσιεύσεις οι οποίες περιέλαβαν τοποθέτηση προϊόντος και ιδιαίτερος με διαγωνισμούς, καθώς εκείνοι φαίνεται να διατήρησαν σημαντικά μεγαλύτερο ποσοστό αλληλεπίδρασης από τις υπόλοιπες κατηγορίες δημοσιεύσεων. Στη διαμόρφωση του ποσοστού αυτού καθοριστικό ρόλο έπαιξε ο συντριπτικά μεγαλύτερος αριθμός σχολιασμών, ο οποίος σημειώθηκε σε τέτοιου είδους αναρτήσεις. Παρά το γεγονός ότι οι χορηγούμενες δημοσιεύσεις κατείχαν τον μικρότερο, κατά μέσο όρο, αριθμό σχολιασμών, κατέλαβαν τη δεύτερη θέση της κατάταξης, ως αναφορά στο ποσοστό αλληλεπίδρασης, κάτι το οποίο οφείλεται στον αισθητά αυξημένο μέσο αριθμό θετικών αντιδράσεων. Λαμβάνοντας υπόψη, λοιπόν, πως τα στατιστικά της κατηγορίας Άλλο είναι ενισχυμένα από τη μοιραία συμπερίληψη διαγωνισμών και χορηγιών, ενώ, παράλληλα, αποσπούν το χαμηλότερο ποσοστό αλληλεπίδρασης, ενισχύεται το αρχικό συμπέρασμα πως το ελληνικό κοινό του Instagram εξέφρασε ιδιαίτερο ενδιαφέρον προς τις δημοσιεύσεις οι οποίες αναλαμβάνουν την παρουσίαση κάποιου προϊόντος.

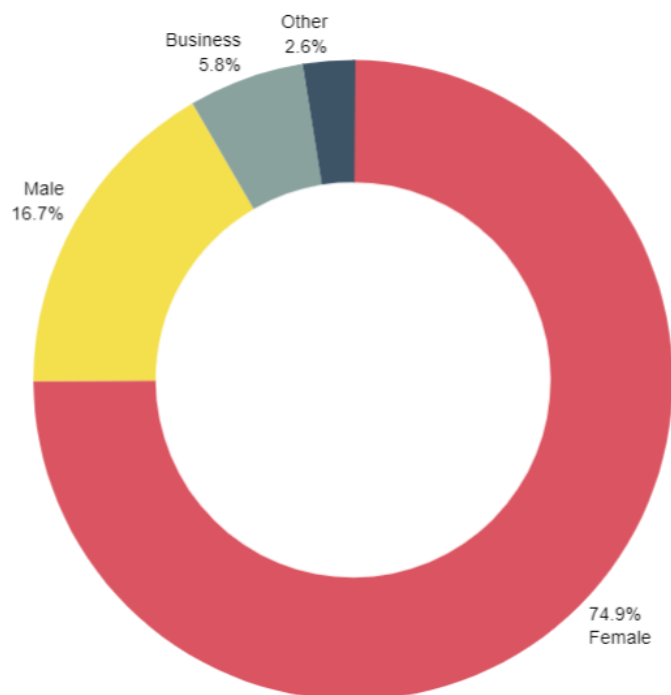
Προς βαθύτερη διερεύνηση του παραπάνω ενδιαφέροντος, αποφασίστηκε η εκτενέστερη ανάλυση συγκεκριμένων πτυχών των δημοσιεύσεων τύπου *διαγωνισμός* και *χορηγία*. Αρχικό βήμα στη παρούσα ανάλυση αποτέλεσε η μέτρηση του ποσοστού ανάρτησης τέτοιου είδους δημοσιεύσεων από τις τέσσερις κατηγορίες χρηστών (Άνδρες, Γυναίκες, Επιχειρήσεις, Άλλο).



Σχήμα 5.41: Αριθμός διαγωνισμών (Giveaways) ανά κατηγορία χρηστών

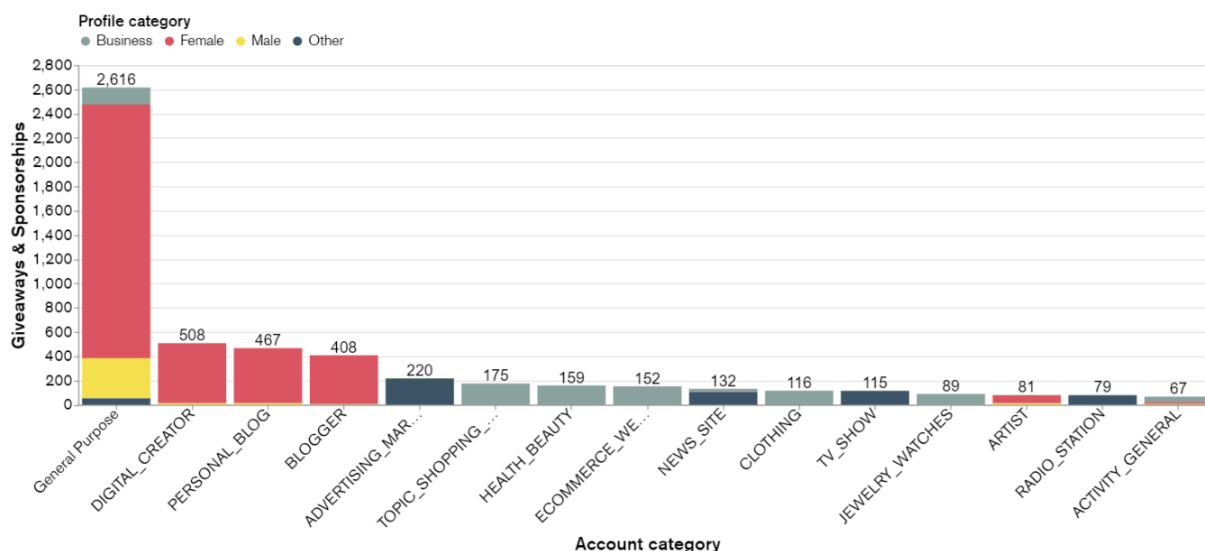
Παρατηρώντας το Σχήμα 5.41, είναι εμφανές πως η πλειοψηφία των διαγωνισμών, κατά τη διάρκεια του 2020, διεξήχθησαν από γυναικείους λογαριασμούς, ενώ τη δεύτερη θέση σε αριθμό διαγωνισμών κατέλαβαν οι επιχειρησιακοί λογαριασμοί. Παρόμοια εικόνα παρουσιάζεται στο Σχήμα 5.42, όπου οι γυναίκες, για ακόμα μία φορά, διατήρησαν τα πρωτεία στον αριθμό χορηγούμενων δημοσιεύσεων. Την εμφάνιση τους στις πρώτες θέσεις έκαναν επίσης οι άνδρες, σε αντίθεση με τη περίπτωση των διαγωνισμών κατά την οποία αποτέλεσαν τον ουραγό της κατάταξης. Από τη μεριά τους, τα προφίλ τύπου Άλλο διοργάνωσαν έναν σεβαστό αριθμό διαγωνισμών, κατά το ίδιο χρονικό διάστημα, ενώ αναφορικά με τις χορηγούμενες δημοσιεύσεις η δραστηριότητα τους παρέμεινε σε αρκετά χαμηλά επίπεδα.

Συμπερασματικά, εξέχουσα θέση στη δημιουργία περιεχομένου, το οποίο περιλαμβάνει τοποθέτηση προϊόντων, κατέλαβαν οι γυναικείοι λογαριασμοί, αναδεικνύοντας έτσι το εμπορικό ενδιαφέρον πίσω από τις γυναίκες του Instagram. Στον τομέα των χορηγιών οι γυναίκες φαίνεται να θεωρήθηκαν οι περισσότερο κατάλληλες για τη προώθηση προϊόντων, φαινόμενο το οποίο, σε συνδυασμό, με την υψηλή τους παρουσία στη διεκπεραίωση διαγωνισμών αποδεικνύει τη στροφή των επιχειρήσεων προς το γυναικείο καταναλωτικό πληθυσμό. Επιπλέον, η σχετικά υψηλή δραστηριότητα των επιχειρησιακών προφίλ στον κλάδο των διαγωνισμών συνιστά ισχυρή ένδειξη ότι τέτοιου είδους αναρτήσεις παρέχουν σε αυτά την ευκαιρία να ενισχύσουν τη ψηφιακή τους παρουσία και να διευρύνουν την αναγνωρισιμότητα τους. Τέλος, η σχεδόν αποκλειστική εμφάνιση γυναικών και ανδρών στη διενέργηση χορηγούμενων αναρτήσεων, πιθανώς, μαρτυρά την τάση των επιχειρήσεων για προώθηση των προϊόντων ή των υπηρεσιών τους μέσω τρίτων και, ενδεχομένως, δημοφιλών χρηστών, τακτική η οποία είναι γνωστή και ως *Influencer Marketing*.



Σχήμα 5.42: Αριθμός χορηγιών (Sponsorships) ανά κατηγορία χρηστών

Έχοντας εντοπίσει τις κατηγορίες χρηστών από τις οποίες προήλθαν οι περισσότερες δημοσιεύσεις τύπου *διαγωνισμός* και *χορηγία* κατά το έτος 2020, η ερευνητική διαδικασία εστιάστηκε στη χαρτογράφηση του θεματικού περιεχομένου, το οποίο αναρτάται από αυτές, μέσω της μέτρησης των διαγωνισμών και των χορηγιών οι οποίες πραγματοποιήθηκαν από την εκάστοτε θεματική κατηγορία λογαριασμών.

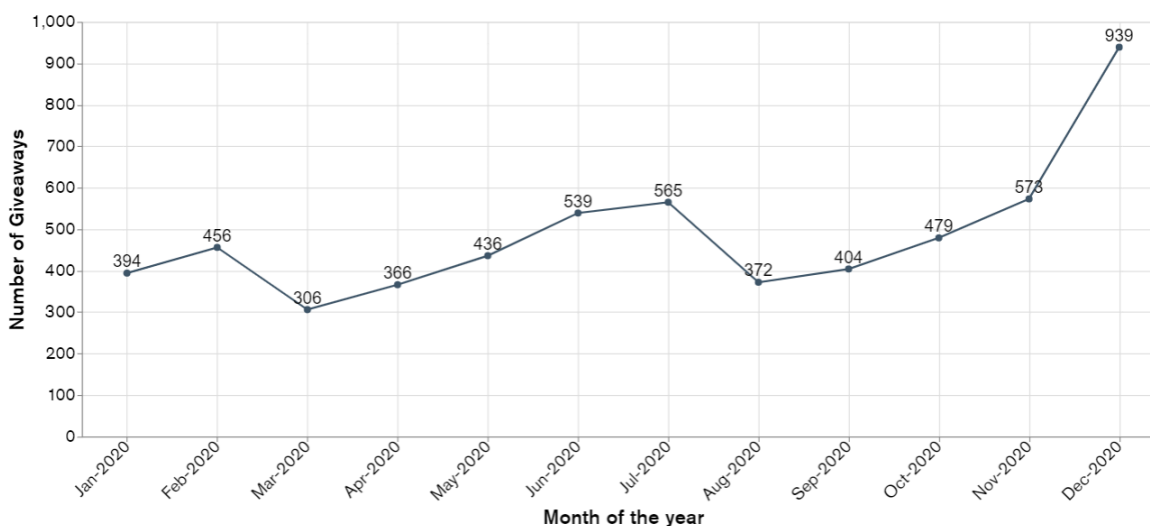


Σχήμα 5.43: Αριθμός διαγωνισμών και χορηγιών ανά θεματική κατηγορία λογαριασμού

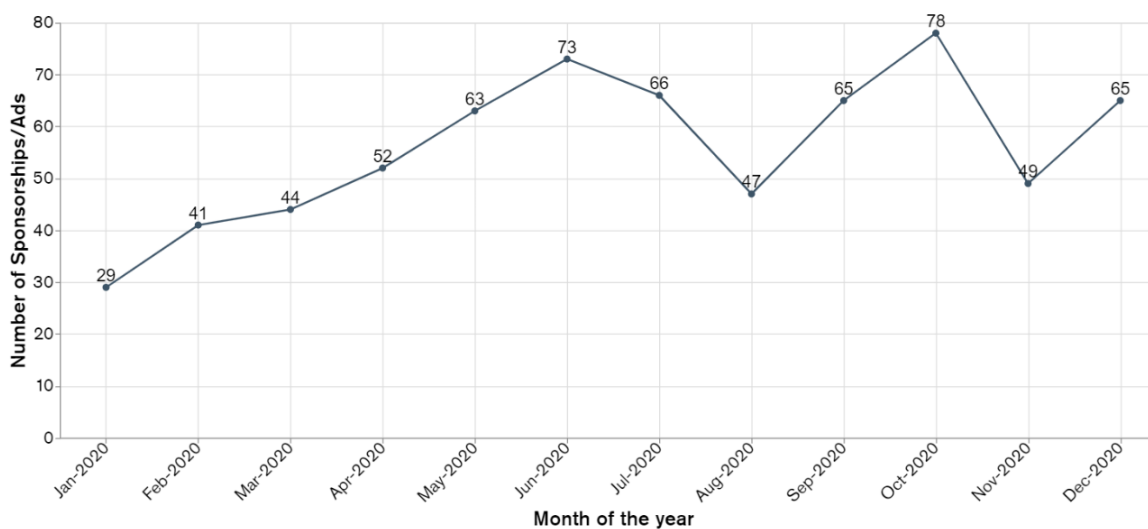
Συγκρίνοντας τα δεδομένα του Σχήματος 5.43 με αυτά των Σχημάτων 5.41 και 5.42 επαληθεύεται

το γεγονός ότι οι γυναίκες, οι οποίες ήταν περισσότερο δραστήριες στον χώρο του Instagram (βλ. Σχήμα 5.27), και ιδιαίτερα εκείνες οι οποίες δημοσίευσαν περιεχόμενο που ανταποκρινόταν στην καθημερινή τους ζωή, αποτέλεσαν την κύρια προτίμηση των επιχειρήσεων για την προώθηση των προϊόντων και των υπηρεσιών τους. Οι άνδρες εμφανίστηκαν στις ίδιες κατηγορίες με τις γυναίκες, όμως, σε πολύ χαμηλότερο ποσοστό. Η περιορισμένη παρουσία ανδρικών λογαριασμών στο κομμάτι των διαγωνισμών και των χορηγιών, ίσως είναι αποτέλεσμα της αρκετά χαμηλής δραστηριότητας την οποία εκείνοι διατήρησαν, σε σχέση με τις υπόλοιπες κατηγορίες χρηστών (βλ. Σχήμα 5.27). Καταληκτικά, με τη βοήθεια των Σχημάτων 5.4 και 5.28 ενισχύεται το πόρισμα ότι οι επιχειρήσεις οι οποίες άσκησαν τη μεγαλύτερη επιρροή στο ελληνικό κοινό του Instagram ήταν αυτές που ασχολούνταν με περιεχόμενο ένδυσης και ομορφιάς.

Εφόσον σκιαγραφήθηκε η συμπεριφορά των διαφόρων κατηγοριών χρηστών και θεματικών κατηγοριών λογαριασμών σχετικά με τις δημοσιεύσεις τοποθέτησης προϊόντων, ακολούθησε το τελικό στάδιο της διερεύνησης των αναρτήσεων τύπου *διαγωνισμός* και *χορηγία*, κατά το οποίο παρακολούθηθηκε η διακύμανση του πλήθους τους ανά τους μήνες του έτους 2020.



Σχήμα 5.44: Αριθμός διαγωνισμών (Giveaways) ανά μήνα του 2020

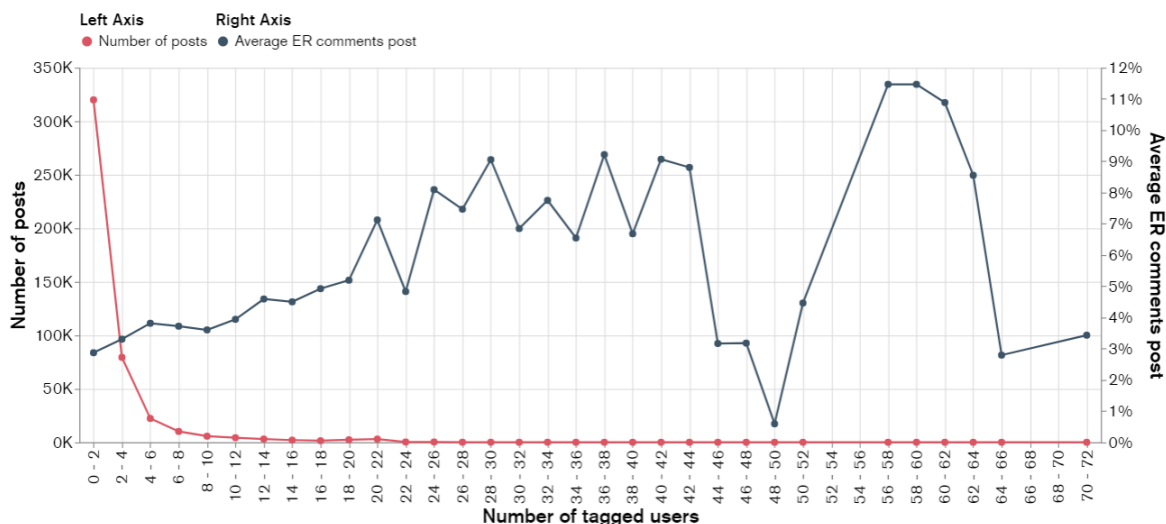


Σχήμα 5.45: Αριθμός χορηγιών (Sponsorships) ανά μήνα του 2020

Μελετώντας από κοινού τα δύο παραπάνω διαγράμματα, διαπιστώνεται πως δημοσιεύσεις οι οποίες περιλάμβαναν προώθηση προϊόντων σημείωσαν ανοδική πορεία προς τους καλοκαιρινούς μήνες του 2020, με την κορύφωση να επέρχεται τον Ιούλιο και προς τα τέλη του έτους. Αμφότεροι οι δύο τύποι δημοσιεύσεων παρουσίασαν μια πτώση στον αριθμό τους κατά τον Αύγουστο, ακολουθώντας την ίδια πορεία με εκείνη του συνολικού αριθμού δημοσιεύσεων του 2020, η οποία αποτυπώθηκε στο Σχήμα 5.33. Στη συνέχεια της λεπτομερέστερης ανάλυσης των διαγραμμάτων, διακρίνεται ένα τοπικό μέγιστο στον αριθμό των διαγωνισμών κατά τον Φεβρουάριο, διάστημα στο οποίο μεσολάβησε η γιορτή του Αγίου Βαλεντίνου και ακολουθήθηκε από μια απότομη πτώση του αριθμού τους, παράλληλα με την έναρξη της πρώτης καραντίνας ενάντια στον COVID-19. Από την άλλη, ο αριθμός των χορηγιών εμφάνισε μια μεγάλη αύξηση κατά την περίοδο Σεπτεμβρίου-Οκτωβρίου, γεγονός για την εξήγηση του οποίου δεν διατίθενται επαρκή στοιχεία, πέραν της μικρής αύξησης των αριθμών των δημοσιεύσεων από επιχειρήσεις και προφίλ τύπου Άλλο κατά το ίδιο χρονικό διάστημα. Στη γενικότερη εικόνα, είναι ορατό πως δημοσιεύσεις οι οποίες αφορούσαν στη προώθηση προϊόντων κοινοποιήθηκαν περισσότερο κατά τη διάρκεια εορτών και του καλοκαιριού, γεγονός το οποίο, πιθανότατα, εκφράζει την τακτική ωφέλεια των επιχειρήσεων σε περιόδους χαλαρότητας και έντονης καταναλωτικής δραστηριότητας.

Η ανάγκη για μία περισσότερο εμπειριστατωμένη οπτική των προτιμήσεων του ελληνικού κοινού του Instagram, υπέδειξε την αξιοποίηση όλων των διαθέσιμων στοιχείων γύρω από τους επισυναπτόμενους χρήστες, οι οποίοι έκαναν την εμφάνισή τους σε αναρτήσεις ελληνικών λογαριασμών.

Αρχικά, το ενδιαφέρον επικεντρώθηκε στην εξέταση της επιρροής που μπορεί να ασκεί η ύπαρξη ενός ή περισσότερων επισυναπτόμενων χρηστών στα πλαίσια μιας δημοσίευσης, ως αναφορά στο ποσοστό αλληλεπίδρασης του κοινού με το κοινοποιημένο περιεχόμενο.



Σχήμα 5.46: Σχέση μεταξύ αριθμού επισυναπτόμενων χρηστών και ποσοστού αλληλεπίδρασης

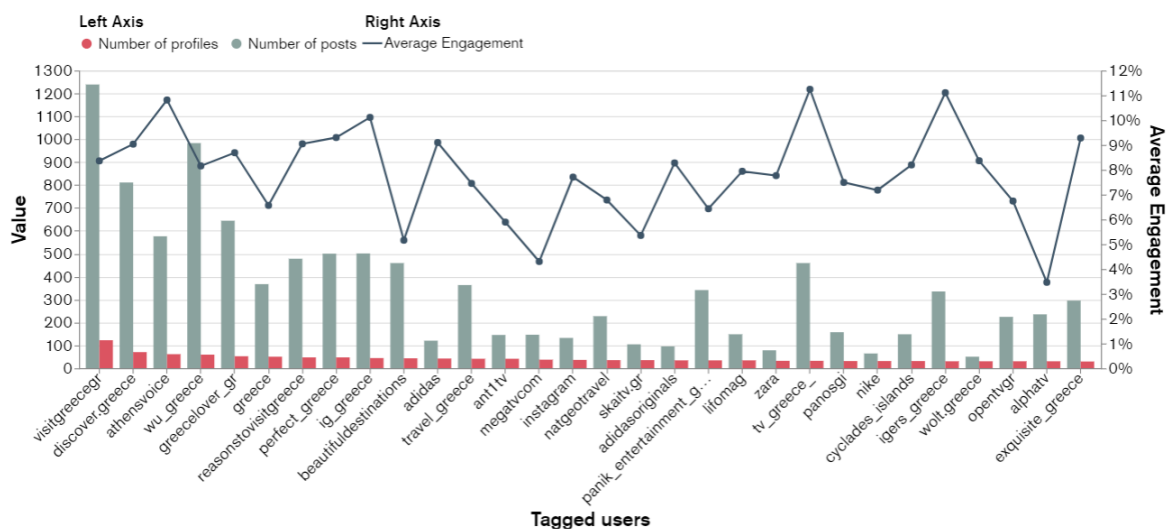
Με μια γρήγορη ματιά στο Σχήμα 5.47 είναι δυνατό κανείς να αντιληφθεί πως η συνύπαρξη διψήφιου αριθμού χρηστών στην ίδια δημοσίευση αποτελεί λιγότερο συχνό, έως και σπάνιο φαινόμενο, καθώς η πλειοψηφία των αναρτήσεων ενθυλακώνει κανέναν έως και 10 χρήστες. Ωστόσο, παρατηρώντας τη πορεία της μεταβλητής *ER Comments Post*, διαπιστώνεται πως η συγκέντρωση υψηλότερων ποσοστώ αλληλεπίδρασης επιτεύχθηκε από τη μειοψηφία των δημοσιεύσεων. Αναλυτικότερα, φαίνεται πως στην κλίμακα των 0-22 επισυναπτόμενων χρηστών, αναλογιζόμενοι τη μεγαλύτερη ακρίβεια των αριθμητικών στοιχείων στη παρούσα κλίμακα λόγω του μεγαλύτερου δείγματος αναρτήσεων, η αύξηση των ατόμων που εμφανίζονται στα πλαίσια μίας ανάρτησης επέφερε την άνοδο του ποσοστού αλληλεπίδρασης. Το γεγονός ότι οι υψηλότερες κλίμακες τους πλήθους επισυναπτόμενων χρηστών βασίζονται σε αρκετά περιορισμένο αριθμό δημοσιεύσεων καθιστά την εξέλιξη των τιμών της μεταβλητής ασταθή. Εντούτοις, οι υψηλότερες τιμές του ποσοστού αλληλεπίδρασης σημειώθηκαν σε μεγαλύτερους αριθμούς επισυναπτόμενων χρηστών, χωρίς αυτό να αντιπροσωπεύει ολόκληρο το φάσμα των ανώτερων στρωμάτων.

Η γενικότερη αύξηση ποσοστού αλληλεπίδρασης, η οποία ακολουθεί τον μεγαλύτερο αριθμό των χρηστών που συμπεριλαμβάνονται σε μία δημοσίευση, ενδεχομένως, οφείλεται στην έκθεση του κοινοποιημένου περιεχομένου σε ένα ευρύτερο κοινό από αυτό του δημιουργού της ανάρτησης.

Διαδοχικό βήμα της ανάλυσης συνέστησε η εύρεση των χρηστών με τις περισσότερες εμφανίσεις σε διαφορετικές αναρτήσεις και λογαριασμούς ανά τις επιμέρους κατηγορίες χρηστών. Η συγκεκριμένη τακτική αποσκοπούσε στη διασαφήνιση των ενδιαφερόντων του κοινού και στη σκιαγράφηση των πιθανών ψηφιακών σχέσεων, οι οποίες αναπτύσσονται μεταξύ τους.

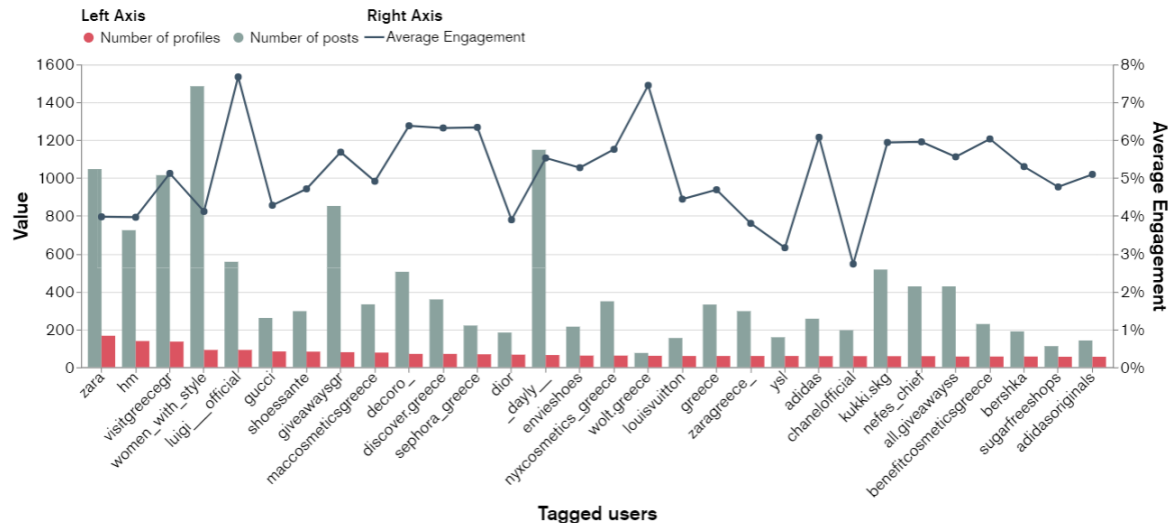
Στα παρακάτω διαγράμματα, η κατάταξη των χρηστών πραγματοποιήθηκε με βάση τον αριθμό των μοναδικών εμφανίσεων τους σε ξεχωριστούς λογαριασμούς, έτσι ώστε να μην επηρεαστεί η τελική κατάταξη από την πολλαπλή επισήμανση ενός χρήστη από έναν μεμονωμένο λογαριασμό. Η αναπαράσταση του πλήθους των εμφανίσεων του εκάστοτε χρήστη συνοδεύτηκε από το μέσο ποσοστό αλληλεπίδρασης των δημοσιεύσεων στις οποίες εκείνος συμπεριλήφθηκε, έτσι

ώστε να αναδειχθούν οι περισσότεροι προσοδοφόρες συνεργασίες.



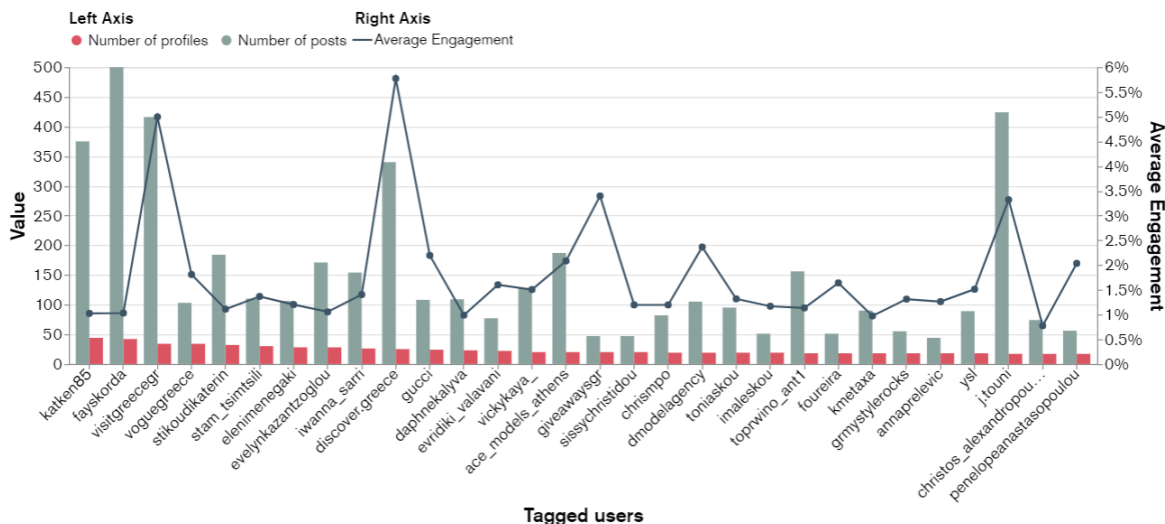
Σχήμα 5.47: Δημοφιλέστεροι επισυναπτόμενοι χρήστες (Άνδρες)

Εφόσον προηγήθηκε μία σύντομη περιήγηση στις κοινοποιημένες δημοσιεύσεις των λογαριασμών, οι οποίοι αναφέρθηκαν περισσότερο από ανδρικά προφίλ, έγινε αντιληπτό πως η μεγαλύτερη μερίδα κατάταξης του Σχήματος 5.47 απαρτίζεται από λογαριασμούς ταξιδιωτικού περιεχομένου και προβολής τουριστικών προορισμών, δείχνοντας έτσι το ενδιαφέρον τους και τον υψηλό βαθμό συναναστροφής τους με τέτοιου είδους λογαριασμούς. Το συγκεκριμένο συμπέρασμα θα μπορούσε να προοικονομηθεί από το Σχήμα 5.2, όπου φανεωνόταν μια μερίδα ανδρών χρηστών οι οποίοι ασχολούνταν με τη θεματική φωτογράφιση. Στις κορυφαίες θέσεις της κατάταξης παρευρίσκονται, επίσης, προφίλ τηλεοπτικών καναλιών και περιοδικών προσαρτώντας τη ψυχαγωγία και την ενημέρωση στο σύνολο των ανδρικών ενδιαφερόντων. Κλείνοντας την κατηγορία των ανδρών, δεν θα μπορούσαν να απέχουν οι επιχειρήσεις, λαμβάνοντας υπόψη όλα όσα έχουν προαναφερθεί, με έναν μικρό αριθμό εταιρειών ρουχισμού να δίνουν το παρόν στη παραπάνω κατάταξη.



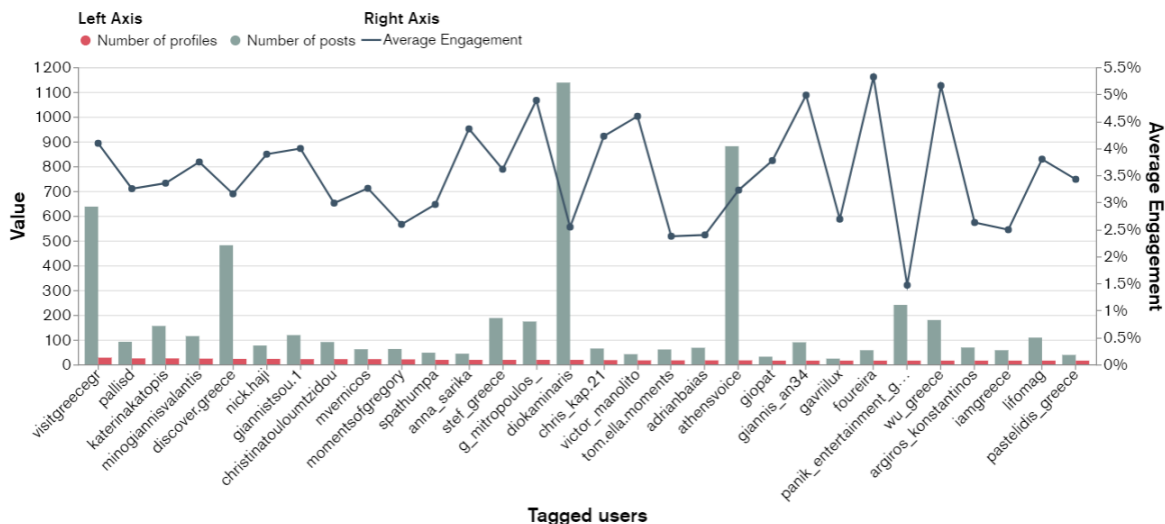
Σχήμα 5.48: Δημοφιλέστεροι επισυναπτόμενοι χρήστες (Γυναίκες)

Αναφορικά με το γυναικείο υποσύνολο των συγκεντρωμένων λογαριασμών, η κατάταξη του Σχήματος 5.48 κατακλύζεται από επιχειρησιακούς λογαριασμούς, με τη συντριπτική τους πλειοψηφία αυτών να εντάσσεται στον κλάδο του ρουχισμού και του καλλωπισμού. Αποδεικνύεται, λοιπόν, πως ένα σημαντικό μέρος των γυναικών του Instagram έθρεψε ενδιαφέρον για την ανάπτυξη επιχειρηματικών συνεργασιών, κυρίως στον κλάδο της ένδυσης και ομορφιάς, και, παράλληλα, επιχειρήσεις, οι οποίες δραστηριοποιούνταν στη πλατφόρμα, στράφηκαν κατά ένα μεγάλο ποσοστό στον γυναικείο πληθυσμό, αναπτύσσοντας κατ' αυτόν το τρόπο μία αμφίδρομη σχέση αλληλεπίδρασης. Ισχυρή απόδειξη της συγκεκριμένης διατύπωσης αποτελεί το Σχήμα 5.49, μέσω του οποίου επιβεβαιώνεται πως οι εταιρείες, κατά κύριο λόγο, συνεργάστηκαν κατά το έτος 2020 με γυναικίους λογαριασμούς, οι οποίοι μάλιστα διατηρούσαν μεγάλο αριθμό ακολούθων, επαναφέροντας έτσι στο προσκήνιο την τακτική του *Influencer Marketing*. Πέραν των επιχειρησιακών λογαριασμών, υψηλές θέσεις της κατάταξης καταλαμβάνουν και λογαριασμοί ταξιδιωτικού περιεχομένου υποδηλώνοντας το γενικότερο ενδιαφέρον των ανδρών και των γυναικών προς αυτή τη κατηγορία περιεχομένου.



Σχήμα 5.49: Δημοφιλέστεροι επισυναπτόμενοι χρήστες (Επιχειρήσεις)

Καταλήγοντας στους λογαριασμούς τύπου Άλλο φαίνεται πως το περιεχόμενο των προφίλ, τα οποία επισυνάφθηκαν σε αυτούς, ακολούθησαν, κατά βάση, ένα κοινό μοτίβο τουριστικών/ταξιδιωτικών αναρτήσεων. Η δραστηριότητα αυτή παρουσιάζει αρκετές ομοιότητες με εκείνη των ανδρικών λογαριασμών, προβάλλοντας τη σύμπλευση των ανδρών με τους λογαριασμούς τύπου Άλλο ως προς το θεματικό περιεχόμενο το οποίο αναπαράγουν. Παρόλα αυτά, δε παρατηρείται κάποια σημαντική υπεροχή μεταξύ επισυναπτόμενων ανδρών και γυναικών, γεγονός το οποίο σε συνδυασμό με την ποικιλομορφία των λογαριασμών οι οποίοι παρατηρούνται στις χαμηλότερες θέσεις κατάταξης (λογαριασμοί άθλησης, τέχνης και ψυχαγωγίας), ίσως, αποδεικνύει τον πολυδιάστατο χαρακτήρα των λογαριασμών τύπου Άλλο.

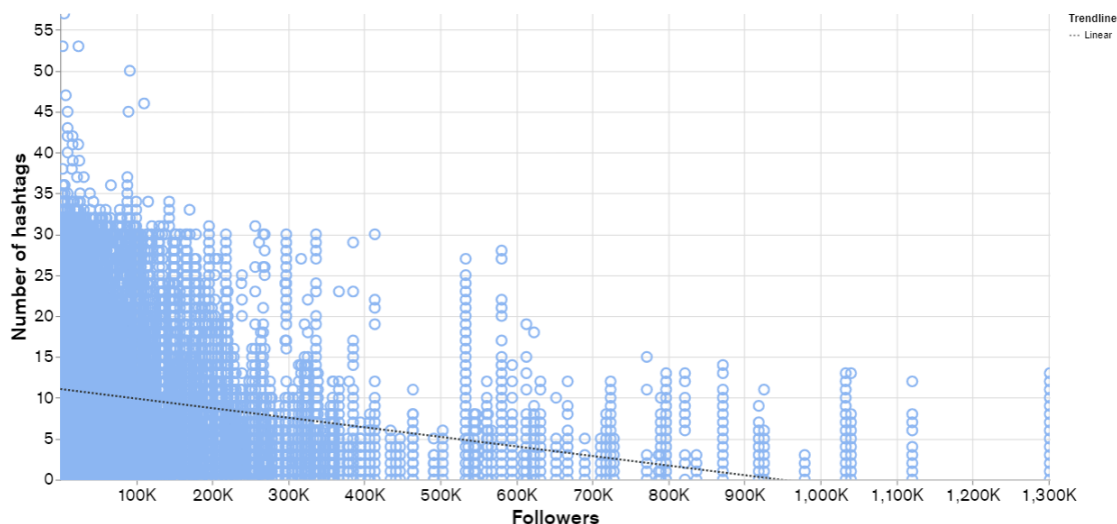


Σχήμα 5.50: Δημοφιλέστεροι επισυναπτόμενοι χρήστες (Άλλο)

Την ολοκλήρωση της ανάλυσης των συγκεντρωμένων δεδομένων για το Instagram σηματο-

δότησε η εκμετάλλευση των *hashtags*, με σκοπό τη μελέτη του τρόπου με τον οποίο η ύπαρξη τους επηρεάζει την απόδοση μιας δημοσίευσης και τον σχηματισμό μιας πλήρους εικόνας των ενδιαφερόντων του ελληνικού κοινού του Instagram.

Πρωταρχικό μέλημα του σταδίου αυτού αποτέλεσε ο υπολογισμός του βαθμού χρήσης των *hashtags* αναλογικά με το πλήθος των ακολούθων ενός λογαριασμού.



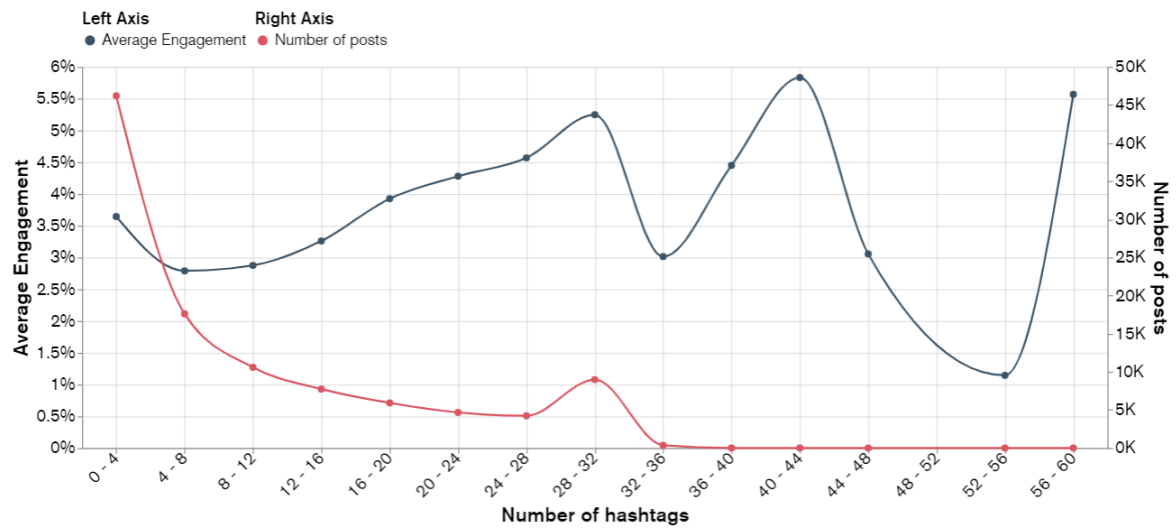
Σχήμα 5.51: Αριθμός *hashtags* ανά αριθμό ακολούθων

Παρατηρώντας το Σχήμα 5.51 φαίνεται να υπάρχει σύνδεση μεταξύ των δύο παραπάνω μεγεθών. Συγκεκριμένα, λογαριασμοί με μεγαλύτερο αριθμό ακολούθων είχαν την τάση να χρησιμοποιούν αρκετά λιγότερα *hashtags* σε σχέση με τους μικρότερους. Το συγκεκριμένο φαινόμενο είναι δυνατό να ερμηνευτεί με δύο πιθανές εξηγήσεις. Κατά πρώτων, αν αναλογιστεί κανείς πως η χρησιμότητα των *hashtags* αφορά στη προβολή μιας ανάρτησης σε ένα ευρύτερο κοινό, τότε είναι εύκολα αντιληπτό ότι η χρήση τους δεν ωφελεί ιδιαίτερα λογαριασμούς με ήδη ανεπτυγμένο κοινό, σε αντίθεση με τα λιγότερο γνωστά προφίλ. Ως δεύτερο πιθανό σενάριο θεωρήθηκε η φειδώ στη χρήση των *hashtags* από τους δημοφιλέστερους λογαριασμούς, με στόχο τη προστασία της δημόσιας εικόνας.

Προς έλεγχο της ορθότητας των δύο παραπάνω σεναρίων, εξετάστηκε η σχέση μεταξύ του ποσοστού αλληλεπίδρασης και του αριθμού των *hashtags* σε μία δημοσίευση ανά τις επιμέρους ομάδες ακολούθων: (i) 0-20.000 ακόλουθοι, (ii) 20.001-100.000 ακόλουθοι, (iii) 100.001-280.000 ακόλουθοι και (iv) ακόλουθοι > 280.000. Όπως γίνεται φανερό στα παρακάτω σχήματα, οι περιπτώσεις κατά τις οποίες μια δημοσίευση περιέχει περισσότερα από 32 *hashtags* είναι ολιγάριθμες, με αποτέλεσμα η πορεία των τιμών της μεταβλητής *ER Comments post* να είναι λιγότερο αντιπροσωπευτική σε αυτήν την κλίμακα. Επομένως, το ενδιαφέρον επικεντρώθηκε στις αναρτήσεις με αριθμό *hashtags* μικρότερο του 32.

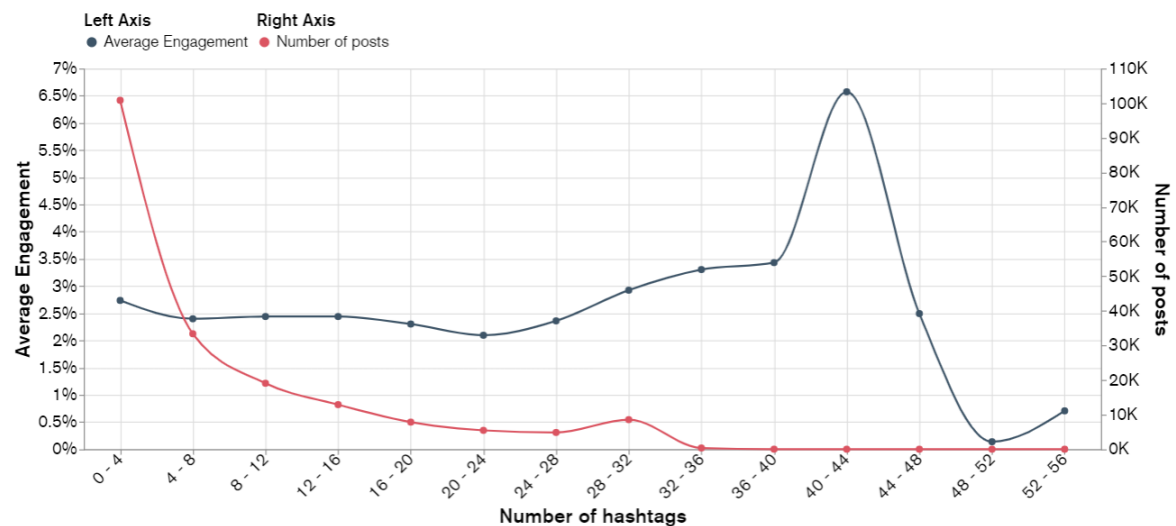
Ξεκινώντας με τους λογαριασμούς οι οποίοι κατέχουν λιγότερους από 20.000 ακόλουθους, είναι εμφανές πως η αύξηση του αριθμού των *hashtags* επέφερε μεγαλύτερα ποσοστά αλληλεπίδρασης στις δημοσιεύσεις, κατά μέσο όρο. Η πτώση του *ER Comments post*, η οποία παρατηρείται στο 0-4 έως 4-8, δε συνεπάγεται το συμπέρασμα ότι η εντονότερη χρήση *hashtags* έχει επίπτωση στο ποσοστό αλληλεπίδρασης, αλλά, πιθανώς, δικαιολογείται από το γεγονός ότι λογαριασμοί

οι οποίοι αποσπούν υψηλά ποσοστά δεν επιζητούν την ευρύτερη προώθηση του περιεχομένου τους, μέσω της χρήσης αυτών.



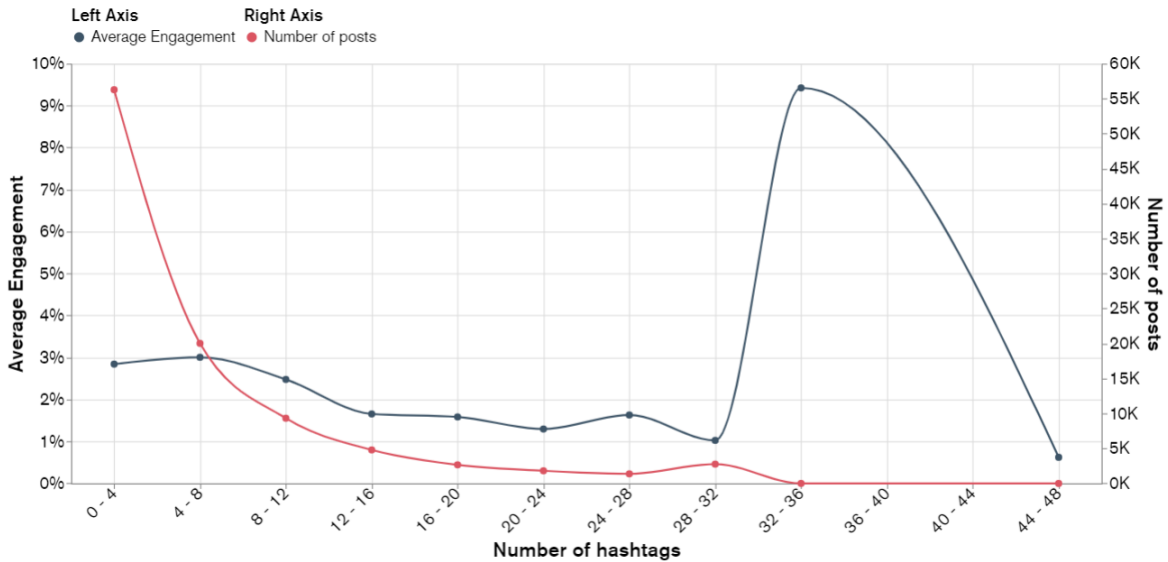
Σχήμα 5.52: Σχέση μεταξύ ποσοστού αλληλεπίδρασης και αριθμού hashtags (0K-20K ακόλουθοι)

Η μετάβαση στην αμέσως επόμενη κλίμακα ακολούθων (20K-100K ακόλουθοι) συνοδεύεται από έναν πολύ χαμηλότερο ρυθμό μεταβολής του ποσοστού αλληλεπίδρασης, καθώς οι τιμές της μεταβλητής *ER Comments post* εμφανίζουν μικρές αυξομειώσεις της τάξης του 0,5%, με την τάση προς τον μεγαλύτερο αριθμό *hashtags* να είναι ελαφρώς ανοδική. Η πορεία αυτή των τιμών του ποσοστού αλληλεπίδρασης είναι δυνατό να μεταφραστεί στην μετρίαση του θετικού αντίκτυπου των *hashtags* στην απόδοση μιας ανάρτησης προς τα μεγαλύτερα στρώματα ακολούθων.

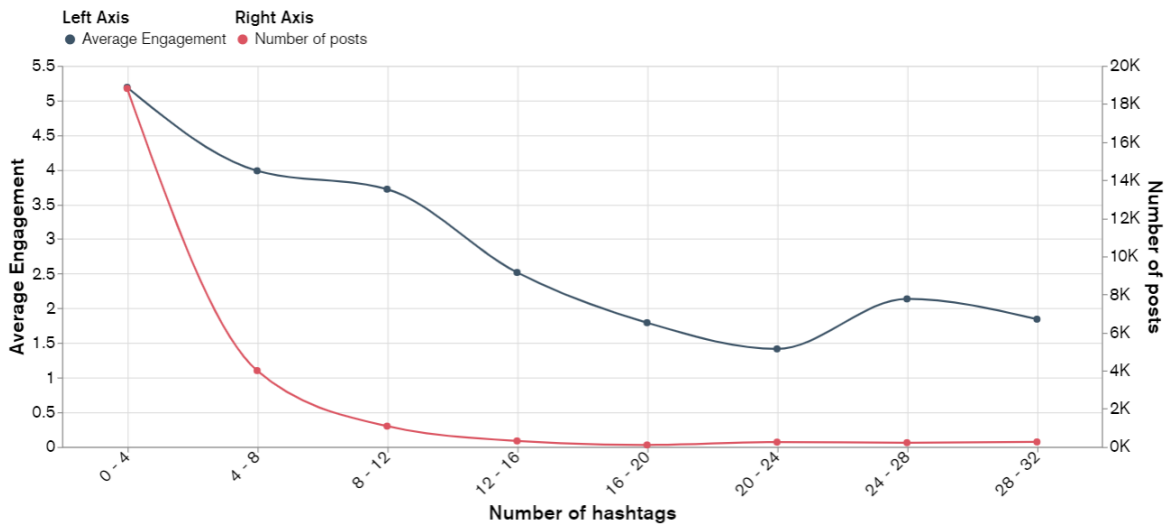


Σχήμα 5.53: Σχέση μεταξύ ποσοστού αλληλεπίδρασης και αριθμού hashtags (20K-100K ακόλουθοι)

Την μετρίαση της συνεισφοράς των *hashtags*, η οποία παρουσιάστηκε κατά τη προηγούμενη βαθμίδα, διαδέχεται η αναποτελεσματικότητα της χρήσης τους από λογαριασμούς με αριθμό ακολούθων μεγαλύτερο των 100.000, καθώς η μεταβλητή *ER Comments post* σημειώνει καθοδική πορεία σε αυτά τα επίπεδα ακολούθων. Όπως προαναφέρθηκε, η συγκεκριμένη πτώση δεν αντικατοπτρίζει τη μείωση του ποσοστού αλληλεπίδρασης λόγω της χρήσης μεγαλύτερου αριθμού *hashtags*, αλλά τη προσπάθεια των λογαριασμών χαμηλότερης απόδοσης να διαδώσουν το περιεχόμενο τους σε περισσότερους χρήστες της πλατφόρμας, μέσω της χρήσης αυτών.



Σχήμα 5.54: Σχέση μεταξύ ποσοστού αλληλεπίδρασης και αριθμού *hashtags* (100K-280K ακόλουθοι)



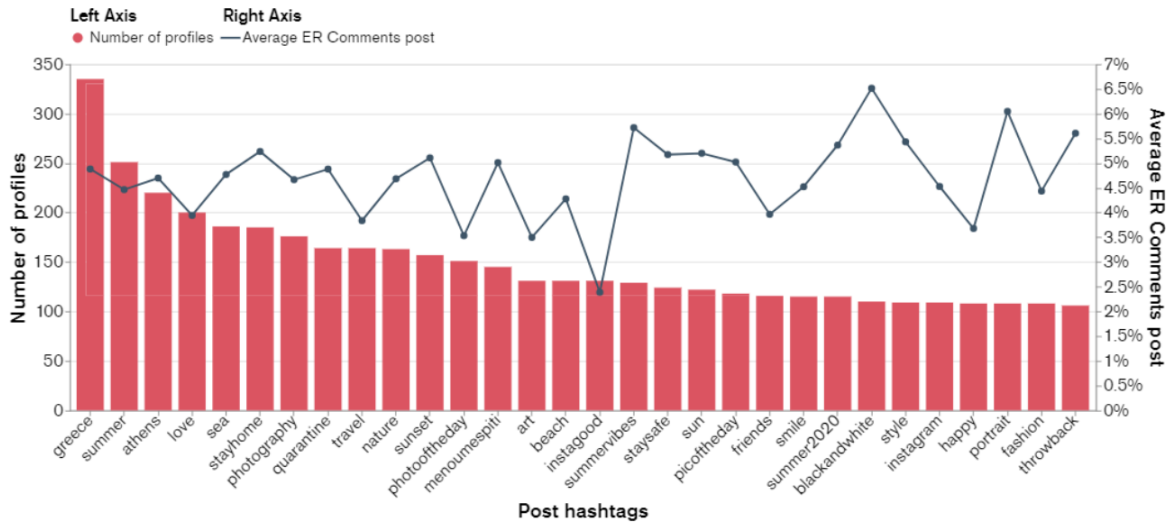
Σχήμα 5.55: Σχέση μεταξύ ποσοστού αλληλεπίδρασης και αριθμού *hashtags* (ακόλουθοι > 280K)

Συνοψίζοντας όλα τα παραπάνω, προκύπτει το συμπέρασμα πως εκείνοι οι οποίοι ευνοήθηκαν

περισσότερο από τη χρήση πολλαπλών *hashtags*, όσον αφορά στη βελτίωση της απόδοσης του περιεχομένου τους, ήταν οι λογαριασμοί οι οποίοι ακολουθούνταν από περιορισμένο σύνολο χρηστών, διότι μέσω αυτής της τακτικής αυξάνονταν οι πιθανότητες προσέγγισης νέων υποψήφιων ακολούθων. Αντιθέτως, η κατοχή υψηλού αριθμού ακολούθων από ένα προφίλ κατέστησε τη χρήση των *hashtags* λιγότερο αποδοτική και, κατά συνέπεια, λιγότερο απαραίτητη [54].

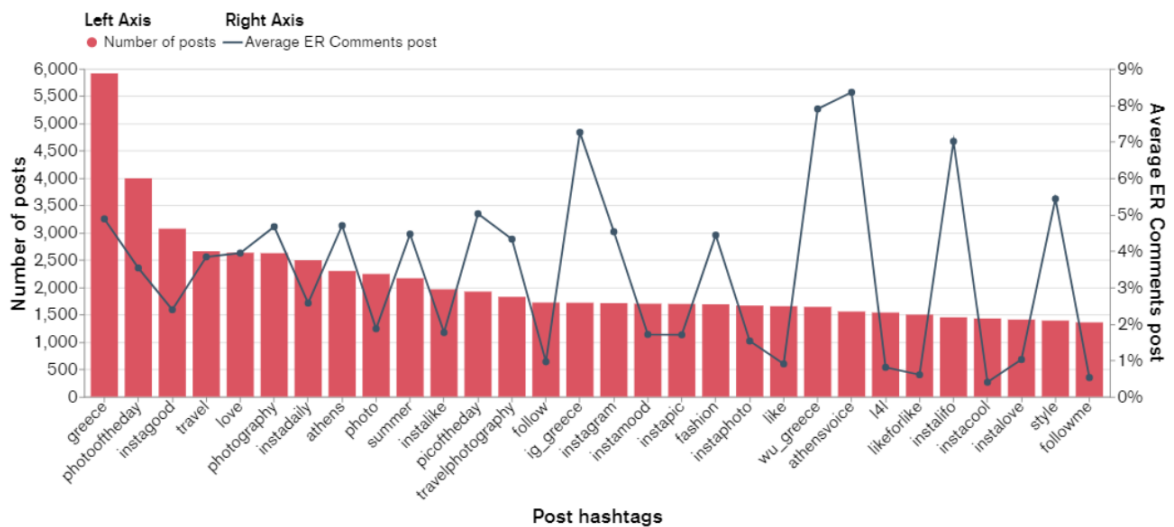
Η αποκωδικοποίηση των ενδιαφερόντων των Ελλήνων χρηστών του Instagram, μέσω της αξιοποίησης των *hashtags*, στηρίχθηκε σε δύο θεμελιώδεις άξονες: (i) το πλήθος των λογαριασμών από τους οποίους χρησιμοποιήθηκε ένα *hashtag* και (ii) το πλήθος των δημοσιεύσεων από τις οποίες χρησιμοποιήθηκε ένα *hashtag*. Βάση της προκείμενης διάκρισης αποτέλεσε το γεγονός ότι τα δύο προαναφερθέντα μεγέθη παρέχουν μια διαφορετική οπτική του ίδιου αντικειμένου μελέτης. Από τη μία, ο αριθμός των λογαριασμών από τους οποίους χρησιμοποιήθηκε ένα *hashtag* προσδιορίζει τους θεματικούς πόλους, οι οποίοι συγκεντρώνουν το ενδιαφέρον της μεγαλύτερης μερίδας χρηστών, ενώ, από την άλλη, το πλήθος των δημοσιεύσεων από τις οποίες χρησιμοποιήθηκε ένα *hashtag* αναδεικνύει τις θεματικές κατηγορίες περιεχομένου με την υψηλότερη ψηφιακή δραστηριότητα. Κατ' επέκταση, αποφασίστηκε η εξέταση των δύο παραπάνω μεγεθών να υλοποιηθεί ξεχωριστά για κάθε μία από τις τέσσερις κατηγορίες χρηστών: (i) άνδρες, (ii) γυναίκες, (iii) επιχειρήσεις και (iv) άλλο.

Η έναρξη της διερεύνησης των δημοφιλέστερων *hashtags* έλαβε χώρα στα πλαίσια των ανδρών χρηστών του Instagram. Ύστερα από την παρατήρηση και τη σύγκριση των Σχημάτων 5.56 και 5.57 διαπιστώνεται πως οι θεματικές κατηγορίες, με βάση τα αντίστοιχα *hashtags*, οι οποίες προσέλκυσαν το ενδιαφέρον του μεγαλύτερου ποσοστού του ανδρικού κοινού κατά το 2020, ταυτίζονται, κατά το πλείστον, με εκείνες οι οποίες σημείωσαν υψηλή δραστηριότητα. Συγκεκριμένα, οι άνδρες έδειξαν ιδιαίτερο ενδιαφέρον προς περιεχόμενο το οποίο απεικόνιζε φυσικά τοπία (*#nature*, *#sunset*, *#sun*) και ταξιδιωτικούς προορισμούς (*#travel*, *#ig_greece*, *#wu_greece*), ενώ την εμφάνιση τους κάνουν και θεματικές, όπως η απαθανάτιση καθημερινών στιγμών ευτυχίας (*#happy*, *#friends*, *#photooftheday*) και η αναπαράσταση περιεχομένου καλλιτεχνικής φύσεως (*#portrait*, *#blackandwhite*). Παράλληλα, ένα μεγάλο μέρος των δημοφιλέστερων *hashtags* αναφερόταν στη καλοκαιρινή περίοδο (*#summer*, *#summervibes*, *#sea*), ενισχύοντας, κατ' αυτόν τον τρόπο, το συμπέρασμα των Σχήματος 5.33 ότι οι καλοκαιρινοί μήνες αποτέλεσαν περίοδο υψηλής διάθεσης και δραστηριότητα του ελληνικού κοινού.



Σχήμα 5.56: Δημοφιλέστερα hashtags ανά ανδρικούς λογαριασμούς

Λαμβάνοντας υπόψη τα *hashtags* τα οποία χρησιμοποιήθηκαν από τους περισσότερους ανδρικούς λογαριασμούς, γίνεται εμφανής η επιρροή του ελληνικού κοινού από την έξαρση του ιού *COVID-19* και την εφαρμογή περιοριστικών μέτρων κατά το έτος 2020, μέσω της χρήσης *hashtags* όπως *#stayhome*, *#quarantine*, *#menoumespiti* και *#staysafe*. Το ίδιο φαινόμενο ισχύει και τα τις υπόλοιπες κατηγορίες χρηστών.

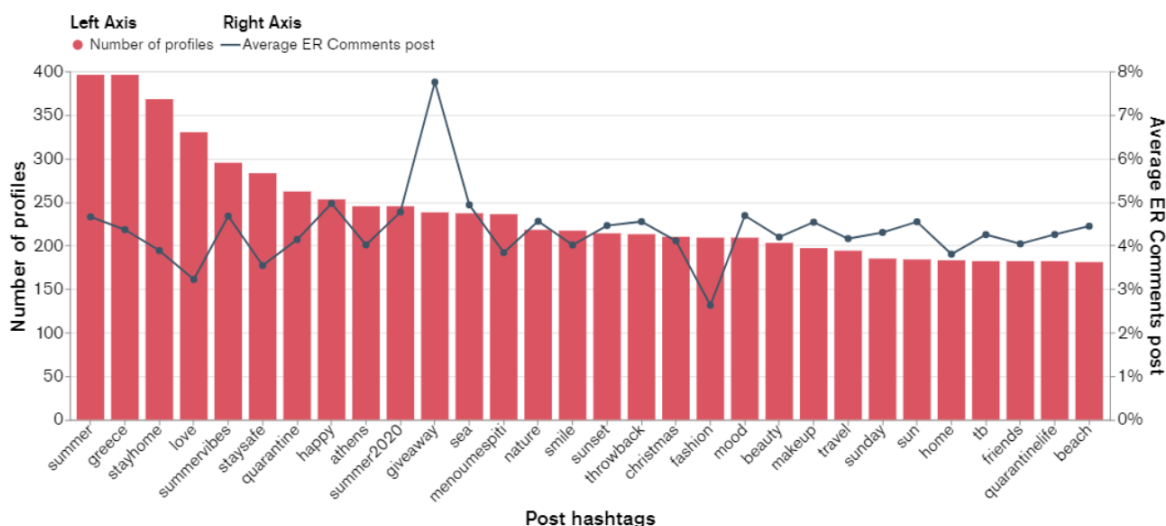


Σχήμα 5.57: Δημοφιλέστερα hashtags ανά δημοσιεύσεις ανδρικών λογαριασμών

Με βάση τα περισσότερο χρησιμοποιημένα *hashtags*, όσον αφορά στο πλήθος των δημοσιεύσεων στις οποίες αναφέρθηκαν, συχνή είναι η παρουσία εκείνων με αρκετά χαμηλό ποσοστό αλληλεπίδρασης, όπως για παράδειγμα τα *#followme*, *#instacool* και *#likeforlike*. Το φαινόμενο αυτό, πιθανώς, συνδέεται με την τακτική των λογαριασμών χαμηλής απόδοσης, η οποία βασίζεται στη χρήση δημοφιλών *hashtags*, με σκοπό την προσέλκυση ευρύτερου κοινού. Ωστόσο, κάτι

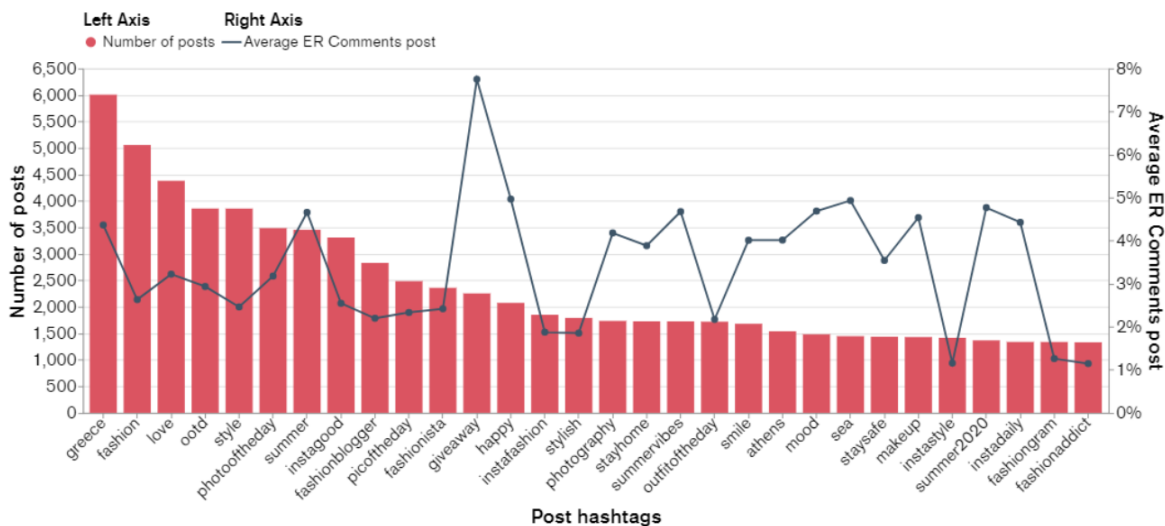
τέτοιο δε λειτουργεί πάντα ευεργετικά για μια δημοσίευση, καθώς η χρήση ευρέως γνωστών *hashtags* ενδέχεται να τη συγκαταλέξει σε ένα τεράστιο σύνολο αναρτήσεων, με αποτέλεσμα αυτή να μηδαμινό όφελος με τη χρήση τέτοιου είδους *hashtags*.

Σε συνέχεια της διερεύνησης των *hashtags*, το ενδιαφέρον στράφηκε στον γυναικείο πληθυσμό του Instagram. Με γνώμονα την κατάταξη του Σχήματος 5.58, γίνεται αντιληπτό πως η πλειοψηφία των γυναικών εμφανίζει αρκετά κοινά σημεία ενδιαφέροντος με εκείνα των ανδρών, καθώς το σύνολο των δημοφιλέστερων θεματικών κατηγοριών περιεχομένου, με βάση τα αντίστοιχα *hashtags*, απαρτίζεται από φυσικά τοπία (*#nature*, *#sunset*, *#sun*), ταξιδιωτικούς προορισμούς (*#travel*, *#greece*) και καθημερινές στιγμές ευτυχίας (*#mood*, *#sunday*, *#smile*). Αντίστοιχα με τους άνδρες, το ενδιαφέρον των γυναικών φαίνεται να ενισχύθηκε κατά τους καλοκαιρινούς μήνες (*#summervibes*, *#summer2020*, *#sea*), αλλά και κατά την περίοδο των Χριστουγέννων (*#christmas*), κάτι το οποίο, επίσης, συνεπάγεται από το Σχήμα 5.33.



Σχήμα 5.58: Δημοφιλέστερα *hashtags* ανά γυναικείους λογαριασμούς

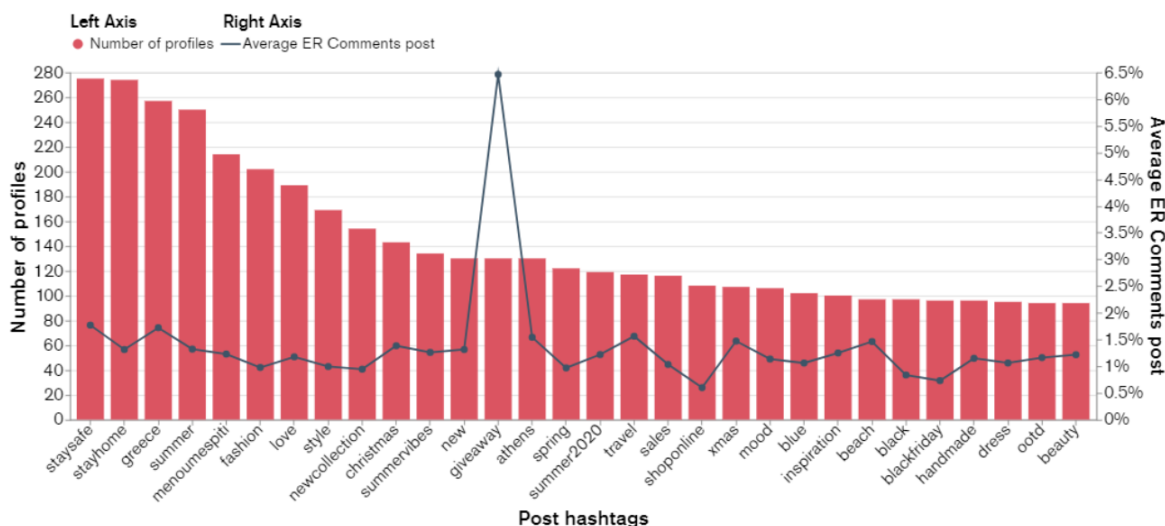
Η εξάπλωση του ιού *COVID-19* δεν άφησε ανεπηρέαστο τον γυναικείο πληθυσμό, γεγονός το οποίο γίνεται αντιληπτό από την αυξημένη χρήση ποικίλων *hashtags* σχετικά με την πανδημία, όπως τα *#stayhome*, *#staysafe*, *#menoumespiti*.



Σχήμα 5.59: Δημοφιλέστερα hashtags ανά δημοσιεύσεις γυναικείων λογαριασμών

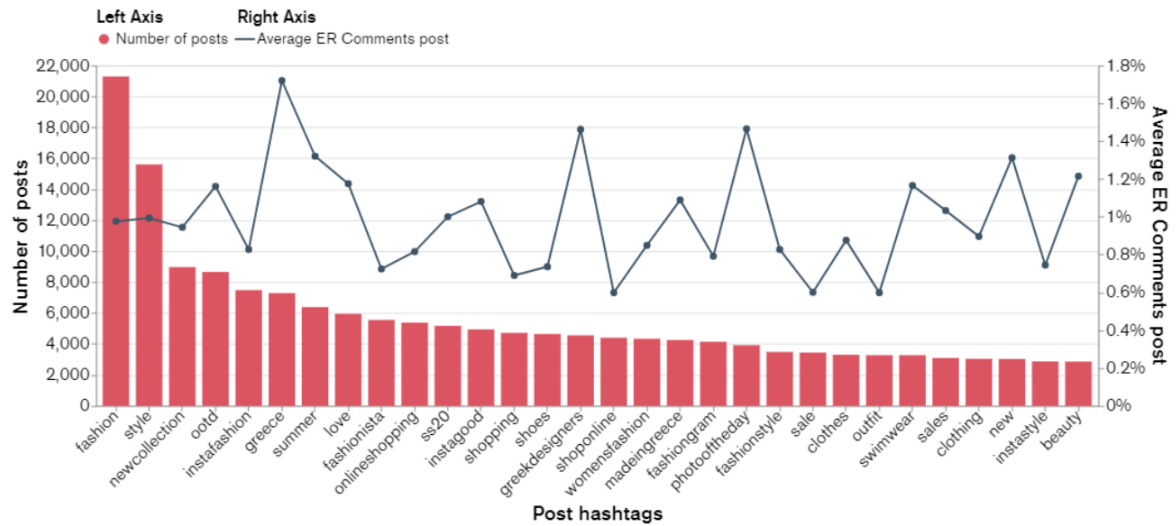
Μολονότι μια σημαντική μερίδα των γυναικών φάνηκε να ταυτίζεται με τις παραπάνω θεματικές κατηγορίες περιεχομένου, αποτυπώθηκε έντονη διαφοροποίηση στους τομείς με την υψηλότερη δραστηριότητα. Αναλυτικότερα, σύμφωνα με το Σχήμα 5.59 η πλειονότητα των *hashtags*, τα οποία επισυνάφθηκαν σε μεγαλύτερο αριθμό δημοσιεύσεων κατά τη διάρκεια του 2020, ήταν επικεντρωμένα στη μόδα/ένδυση (*#fashion*, *#outfitoftheday*, *#fashionblogger*, *#instafashion*) και στον καλλωπισμό (*#makeup*, *#beauty*), αποδεικνύοντας για ακόμη μία φορά τη παραγωγή μεγάλου όγκου περιεχομένου από λογαριασμούς αντίστοιχων κλάδων του Instagram. Τέλος, αξιοσημείωτη είναι η παρουσία των διαγωνισμών (*#giveaway*), οι οποίοι κυριάρχησαν στις τιμές του ποσοστού αλληλεπίδρασης, και, έτσι, επαλήθευσαν την αποδοτικότητα τους, όσον αφορά το ποσοστό επιρροής του κοινού και την υψηλή προτίμηση τους από τις γυναίκες χρήστες (βλ. επίσης Σχήμα 5.40).

Στη σφαίρα των επιχειρησιακών λογαριασμών, η μελέτη των *hashtags*, τα οποία χρησιμοποιήθηκαν από τις περισσότερες επιχειρήσεις (βλ. Σχήμα 5.60), μαρτυρά την στρατηγική έντονης δραστηριότητας των εταιριών κατά τη διάρκεια περιόδων υψηλού καταναλωτισμού, όπως οι καλοκαιρινοί μήνες (*#summer*, *#summervibes*), τα Χριστούγεννα (*#christmas*, *#xmas*) και το *Black Friday* (*#black*, *#blackfriday*). Επιπλέον, η κατάληψη υψηλών θέσεων της κατάταξης από *hashtags* σχετικά με τον ιό *COVID-19*, πιθανώς, υποδηλώνει την ευκαιρία η οποία δόθηκε στις επιχειρήσεις από την εφαρμογή περιοριστικών μέτρων για την εγκαθίδρυση και την εδραίωση της ψηφιακής τους παρουσίας στην ελληνική κοινότητα του Instagram.



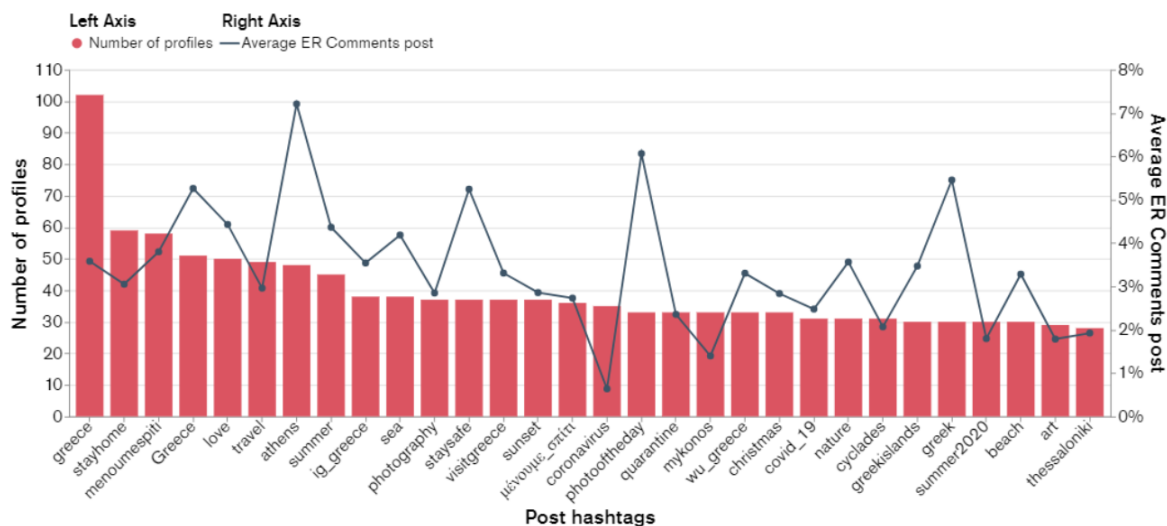
Σχήμα 5.60: Δημοφιλέστερα hashtags ανά επιχειρησιακούς λογαριασμούς

Το σύνολο των ευρύτερα χρησιμοποιημένων *hashtags* σε δημοσιεύσεις επιχειρήσεων προσδίδει ένα σαφέστερο χαρακτήρα στις εταιρείες, οι οποίες άσκησαν την εντονότερη δραστηριοποίηση στον ψηφιακό κόσμο του Instagram, κατά τη διάρκεια του έτους 2020. Είναι ορατό από το Σχήμα 5.61 ότι οι επιχειρησιακοί λογαριασμοί έδωσαν ιδιαίτερη έμφαση στη μόδα/ένδυση (*#fashion*, *#newcollection*) και στον καλλωπισμό (*#beauty*). Οι πολλαπλές ομοιότητες μεταξύ των *hashtags* τα οποία χρησιμοποιήθηκαν από τις γυναίκες και τις επιχειρήσεις, αντίστοιχα, σε συνδυασμό με την ένδειξη για ενασχόληση των επιχειρήσεων με τον γυναικείο πληθυσμό (*#womensfashion*), σκιαγραφούν έναν κοινό άξονα ανάμεσα στις δύο κατηγορίες χρηστών και, παράλληλα, θέτουν τα θεμέλια για την ανάπτυξη αμφίδρομων σχέσεων δραστηριότητας. Τέλος, δεν μπορεί να παραληφθεί η παρουσία των διαγωνισμών (*#giveaway*) ανάμεσα στα δημοφιλέστερα *hashtags* των επιχειρήσεων, καθώς το επιβλητικό ποσοστό αλληλεπίδρασης το οποίο συγκέντρωσαν τους κατέστησε ένα ισχυρό μέσο προώθησης των προϊόντων ή των υπηρεσιών, τα οποία παρέχονται από την εκάστοτε εταιρία (βλ. επίσης Σχήμα 5.40)



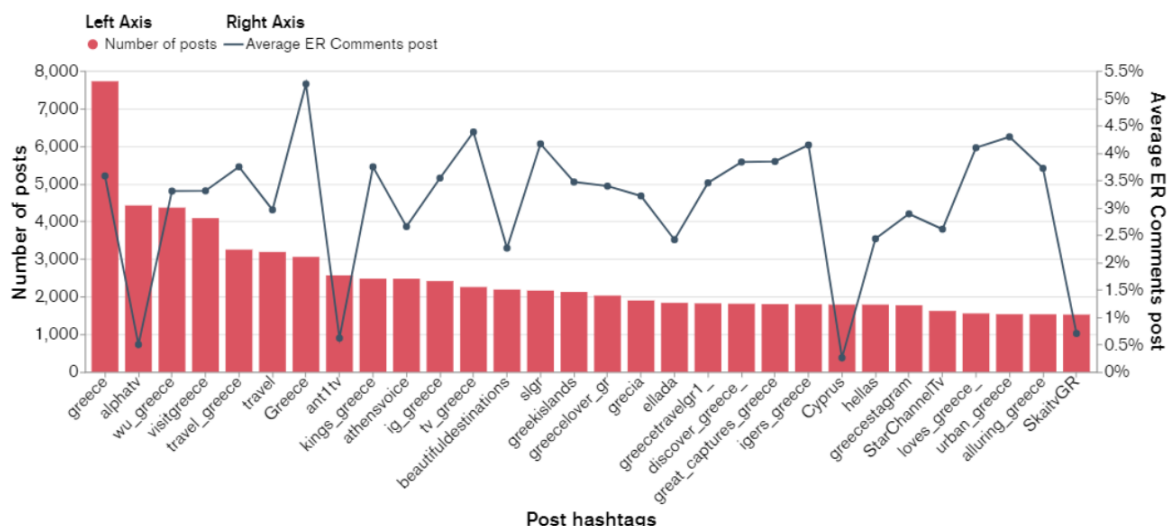
Σχήμα 5.61: Δημοφιλέστερα hashtags ανά δημοσιεύσεις επιχειρησιακών λογαριασμών

Τελευταίοι στη λίστα της μελέτης των επισυναπτόμενων *hashtags* βρίσκονται οι λογαριασμοί τύπου Άλλο. Αναλύοντας τις λέξεις-κλειδιά, η οποίες χρησιμοποιήθηκαν από το μεγαλύτερο σύνολο της συγκεκριμένης κατηγορίας χρηστών κατά το έτος 2020, διακρίνονται αρκετά κοινά σημεία ενδιαφέροντος με όλες τις προαναφερθείσες κατηγορίες. Η χρήση *hashtags*, τα οποία σχετίζονται με ταξιδιωτικούς προορισμούς (*#greece*, *#travel*, *#visitgreece*), καλοκαιρινό/γιορτινό περιεχόμενο (*#summer2020*, *#summer*, *#christmas*) και φυσικά τοπία (*#nature*) θα μπορούσε να χαρακτηρίσει το περιεχόμενο των λογαριασμών τύπου Άλλο ως μια τομή των ενδιαφερόντων των ανδρικών, των γυναικείων και των επιχειρησιακών λογαριασμών. Αντίστοιχα με τις παραπάνω κατηγορίες χρηστών, η σημαντική επιρροή του ιού *COVID-19* στη δραστηριότητα των προφίλ τύπου Άλλο είναι εμφανής, λόγω της έντονης παρουσίας ποικίλων *hashtags*, τα οποία σχετίζονται με την πανδημία στη κορυφή της κατάταξης, όπως τα *#stayhome*, *#menoumespiti*, *#staysafe* και *#μένουμε_σπίτι*.



Σχήμα 5.62: Δημοφιλέστερα hashtags ανά λογαριασμούς τύπου Άλλο

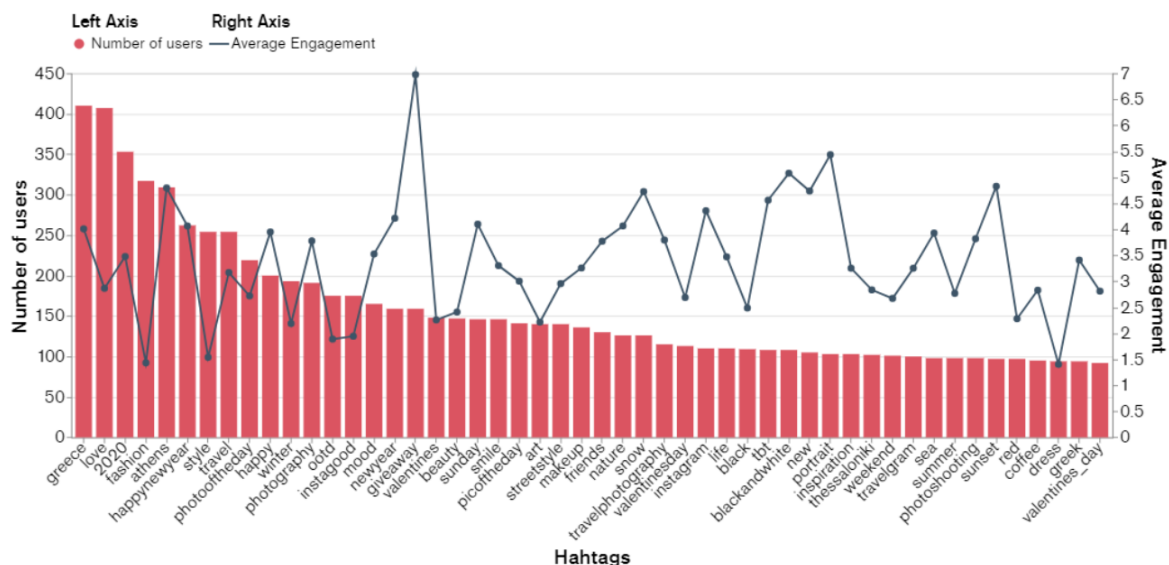
Ωστόσο, σύμφωνα με το Σχήμα 5.63, φαίνεται πως στον άξονα της υψηλότερης δραστηριότητας αναδείχθηκαν δύο επιμέρους θεματικές κατηγορίες περιεχομένου. Ο κατακλυσμός των υψηλότερων θέσεων της κατάταξης από *hashtags* ταξιδιωτικής φύσεως (*#wu_greece*, *#visitgreece*, *#travel_greece*) υποδηλώνει το αυξημένο ενδιαφέρον των λογαριασμών τύπου Άλλο για αναπαγωγή τέτοιου είδους υλικού, ενώ η ύπαρξη ενός σεβαστού αριθμού τηλεοπτικών καναλιών στη κατάταξη (*#alphatv*, *#ant1tv*, *#tv_greece*) μαρτυρά την προσπάθεια αυτής της υποκατηγορίας λογαριασμών να εισέλθουν δυναμικά στον ψηφιακό κόσμο του Instagram, δημιουργώντας ψυχαγωγικό και ενημερωτικό περιεχόμενο. Αξίζει να επισημανθεί, επίσης, πως η διακριτική παρουσία *hashtags* τα οποία αφορούσαν αθλητισμό (*#slgr*) και τέχνες (*#art*) αποτελεί ένδειξη της ποικιλομορφίας που διέπει το θεματικό περιεχόμενο λογαριασμών τύπου Άλλο.



Σχήμα 5.63: Δημοφιλέστερα hashtags ανά δημοσιεύσεις λογαριασμών τύπου Άλλο

Ως τελικός σταθμός της μελέτης των *hashtags* επιλέχθηκε η ανάδειξη εκείνων τα οποία χρησιμοποιήθηκαν από τους περισσότερους χρήστες ανά τις διαδοχικές χρονικές περιόδους του 2020, με σκοπό την εποπτεία της μεταβολής των γενικότερων ενδιαφερόντων του κοινού. Η διάκριση των χρονικών περιόδων στηρίχθηκε, κατά κύριο λόγο, στο μοτίβο των εποχών του έτους, παρόλα αυτά, λόγω του γεγονότος ότι ο Δεκέμβριος δεν ήταν ορθό να συμπεριληφθεί στο διάγραμμα των μηνών Ιανουάριος-Φεβρουάριος, εξαιτίας του χρονικού χάσματος μεταξύ τους, σε συνδυασμό με τη σταδιακή αύξηση της δραστηριότητας, η οποία παρατηρήθηκε κατά τους μήνες Νοέμβριο και Δεκέμβριο (βλ. Σχήμα 5.32), αποφασίστηκε οι δύο τελευταίοι μήνες του έτους να αποσπαστούν σε ξεχωριστό διάγραμμα.

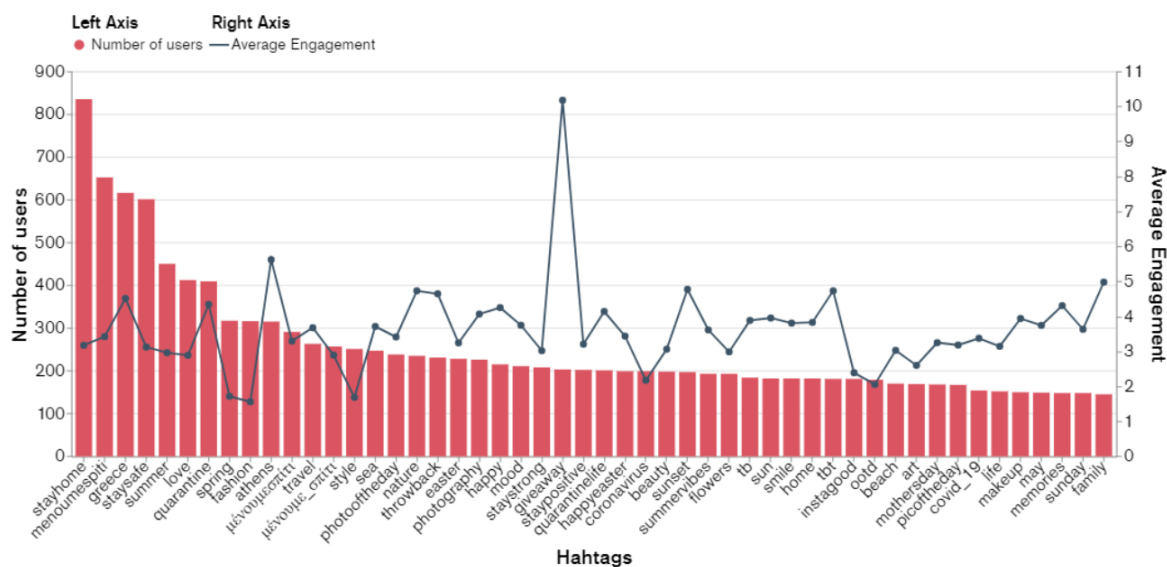
Η έναρξη του έτους 2020 σηματοδοτήθηκε από την ευρεία χρήση *hashtags* για την υποδοχή του νέου έτους, όπως τα #2020, #happynewyear και #newyear, ενώ δεν έλειψαν επίσης αναφορές στη χειμερινή ατμόσφαιρα (#winter, #snow). Ακόμη μία απόδειξη της ανταπόκρισης του ελληνικού κοινού σε εορτές και επετείους αποτέλεσε η παρουσία *hashtags* σχετικά με τη γιορτή του Αγίου Βαλεντίνου σε υψηλές θέσεις της κατάταξης (#valentines, #valentinesday). Ιδιαίτερη έκπληξη προκάλεσε η εμφάνιση καλοκαιρινών *hashtags* εν μέσω χειμερινής περιόδου (#sea, #summer), δείχνοντας έτσι τη μεγάλη συμπάθεια του κοινού προς τους καλοκαιρινούς μήνες, η οποία γίνεται αντιληπτή και από το Σχήμα 5.33. Σε γενικότερο πλαίσιο παρατηρείται πως τα ενδιαφέροντα της πλειονότητας των χρηστών του Instagram ακολούθησαν κοινή πορεία με εκείνη, η οποία αποτυπώθηκε στα προηγούμενα οκτώ διαγράμματα, και συνδέεται με θεματικές κατηγορίες περιεχομένου όπως προβολή φυσικών τοπίων και ταξιδιωτικών προορισμών, μόδα/καλλιπισμός, καλλιτεχνική φωτογραφία και καθημερινές στιγμές ευτυχίας.



Σχήμα 5.64: Δημοφιλέστερα *hashtags* χρονικής περιόδου Ιανουάριος-Φεβρουάριος

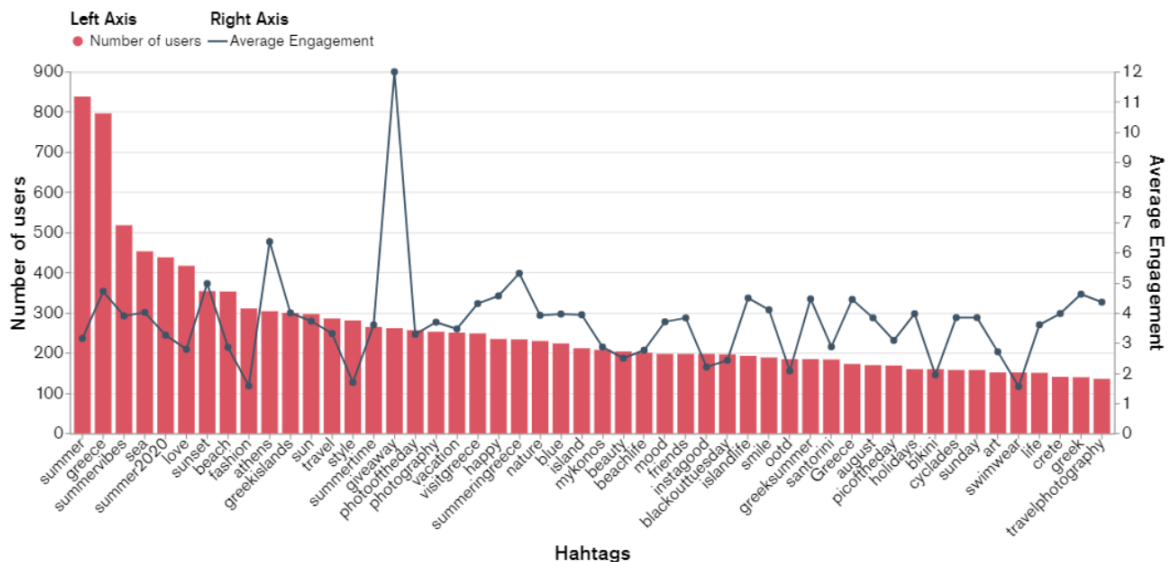
Η περίοδος μεταξύ Μαρτίου και Μαΐου στιγματίστηκε από την έξαρση της πανδημίας του *COVID-19* και την επιβολή περιοριστικών μέτρων, γεγονός το οποίο γίνεται ξεκάθαρο από τη κατάληψη της συντριπτικής πλειοψηφίας των θέσεων της κατάταξης από *hashtags* αναφερόμενα στον ιό, όπως τα #stayhome, #menoumespiti και #staysafe. Οι επιπτώσεις του ιού *COVID-19* στη ψυχολογία του ελληνικού κοινού του Instagram μεταφράστηκε σε δύο μεταβολές στη

ψηφιακή δραστηριότητα των χρηστών, με βάση το Σχήμα 5.65. Η αύξηση του αριθμού των χρηστών οι οποίοι χρησιμοποίησαν *hashtags* ανάλογα της πανδημίας, παράλληλα με τη συχνότερη εμφάνιση λέξεων-κλειδιών οι οποίες διακατέχονταν από ένα αίσθημα αναπόλησης (*#throwback*, *#tb*, *#memories*), υποδηλώνουν την έντονη ευαισθητοποίηση του κοινού η οποία υπήρξε κατά την περίοδο αυτή. Ωστόσο, για ακόμη μία φορά, δεν απείχαν από την κατάταξη οι συνήθεις πόλοι ενδιαφέροντος φυσικά τοπία/ταξιδιωτικοί προορισμοί (*#nature*, *#travel*), ανοιξιάτικο/καλοκαιρινό περιεχόμενο (*#spring*, *#summer*), προσωπικές στιγμές (*#love*, *#family*) και μόδα (*#fashion*, *#style*). Τέλος, η ανταπόκριση του κοινού σε εορταστικό/επετειακό περιεχόμενο έγινε ξανά αντιληπτή κατά τη περίοδο της άνοιξης του 2020 με χαρακτηριστικές αναφορές στις γιορτές του Πάσχα (*#happyeaster*) και στην εορτή της μητέρας (*#mothersday*).



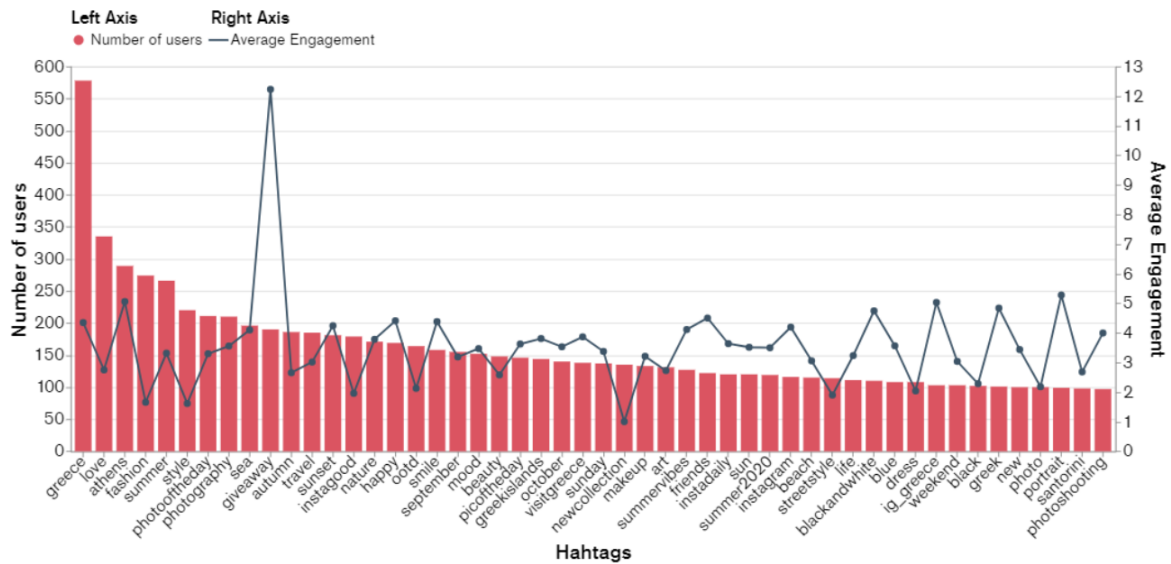
Σχήμα 5.65: Δημοφιλέστερα *hashtags* χρονικής περιόδου Μάρτιος-Μάιος

Ο ερχομός της καλοκαιρινής περιόδου σε συνδυασμό με την άρση των περιοριστικών μέτρων της πανδημίας επέφερε την επιστροφή του ελληνικού κοινού στη κανονικότητα, κάτι το οποίο επιβεβαιώθηκε και από τη ψηφιακή τους συμπεριφορά στο Instagram κατά τους καλοκαιρινούς μήνες. Πιο συγκεκριμένα, η κυριαρχία των *hashtags* τα οποία αφορούσαν σε καλοκαιρινές δραστηριότητες, όπως είναι η απόδραση σε καλοκαιρινούς προορισμούς (*#beach*, *#beachlife*, *#vacation*), η προβολή τουριστικών θερέτρων (*#greekislands*, *#summeringreece*, *#crete*), αλλά και η προώθηση καλοκαιρινών ειδών ένδυσης (*#bikini*, *#swimwear*), επαληθεύει ότι το καλοκαίρι του 2020 αποτέλεσε για τους Έλληνες χρήστες του Instagram μια περίοδο ευδαιμονίας, χαλαρότητας και αυξημένης δραστηριότητας στη πλατφόρμα (βλ. επίσης Σχήμα 5.33). Πέραν του τουριστικού ενδιαφέροντος, το οποίο αναδείχθηκε κατά τη περίοδο Ιουνίου-Αύγουστος, αξιοσημείωτη ήταν η κινητοποίηση ενός σεβαστού συνόλου του κοινού ενάντια στα επεισόδια αστυνομικής βίας, τα οποία καταγράφηκαν το αντίστοιχο χρονικό διάστημα στην Αμερική, μέσω της χρήσης του *hashtag* *#blackouttuesday*.



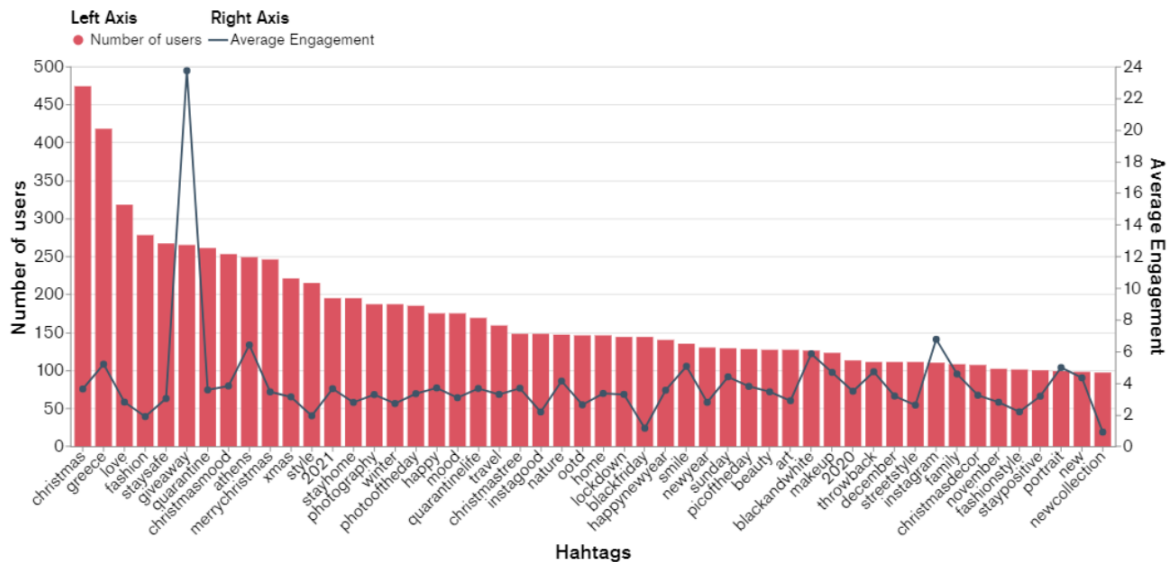
Σχήμα 5.66: Δημοφιλέστερα hashtags χρονικής περιόδου Ιούνιος-Αύγουστος

Εφόσον ολοκληρώθηκε η καλοκαιρινή περίοδος και ο καθολικός περιορισμός των μετακινήσεων, λόγω της πανδημίας, δεν τέθηκε ακόμη σε ισχύ, η μετάβαση στο φθινόπωρο ακολουθήθηκε από τη σταδιακή επιστροφή του ελληνικού κοινού στο προ-καλοκαιριού μοντέλο ψηφιακής συμπεριφοράς, ως προς τις θεματικές κατηγορίες περιεχομένου, με *hashtags* αναφερόμενα σε φυσικά τοπία (*#nature*), φθινοπωρινό περιεχόμενο (*#autumn*, *#september*), καθημερινές στιγμές ευτυχίας (*#mood*, *#sunday*, *#smile*) και μόδα (*#fashion*, *#style*) να επανέρχονται στις υψηλότερες θέσεις της κατάταξης. Παρατηρείται, επιπλέον, από το Σχήμα 5.67 πως, παρά το πέρας της καλοκαιρινής περιόδου, *hashtags* σχετικά με το καλοκαίρι, όπως η προώθηση τουριστικών θερέτρων (*#greekislands*, *#santorini*) και, γενικότερα, οι καλοκαιρινές δραστηριότητες (*#summer*, *#summervibes*), καλύπτουν ένα αξιοσέβαστο ποσοστό των θέσεων της κατάταξης, δηλώνοντας κατά αυτόν τον τρόπο την έντονη συμπάθεια των Ελλήνων χρηστών του Instagram προς τη συγκεκριμένη εποχή. Αξίζει να σημειωθεί πως, πέραν της μερικής επιστροφής των χρηστών στην κανονικότητα, το φθινόπωρο συνοδεύτηκε, επίσης, από μειωμένη δραστηριότητα, όσον αφορά στη χρήση *hashtags*, φαινόμενο το οποίο είχε παρουσιαστεί και στο Σχήμα 5.33, υποδηλώνοντας έτσι τη χαμηλότερη διάθεση του κοινού κατά τους φθινοπωρινούς μήνες.



Σχήμα 5.67: Δημοφιλέστερα hashtags χρονικής περιόδου Σεπτέμβριος-Οκτώβριος

Καταλήγοντας στα τέλη του 2020 και συγκεκριμένα στους μήνες Νοέμβριο και Δεκέμβριο, καθοριστικό ρόλο στη δραστηριότητα των χρηστών έπαιξε η επαναφορά των περιοριστικών μέτρων για τη καταπολέμηση του ιού *COVID-19*, κάτι το οποίο καθίσταται σαφές μέσω της χρήσης *hashtags* σχετικών με την πανδημία (*#staysafe*, *#quarantine*, *#quarantinelifelife*) από μία μεγάλη μερίδα των Ελλήνων χρηστών του Instagram κατά τη συγκεκριμένη χρονική περίοδο (βλ. Σχήμα 5.68). Ωστόσο, συγκρίνοντας το ποσοστό χρήσης τέτοιου είδους *hashtags* με αυτό της περιόδου Μαρτίου-Μαΐου (βλ. Σχήμα 5.65), διάστημα στο οποίο εφαρμόστηκαν τα πρώτα περιοριστικά μέτρα, παρατηρείται πως το αντίκτυπο των περιορισμών στη ψυχολογία του ελληνικού κοινού ήταν ασθενέστερο κατά το δεύτερο κύμα της πανδημίας. Εγκαταλείποντας το κλάδο του ιού *COVID-19*, πέραν των συνηθέστερων θεματικών κατηγοριών: (i) φυσικά τοπία/προορισμοί, (ii) καθημερινές στιγμές και (iii) μόδα, ιδιαίτερα αισθητό γίνεται στο παρακάτω διάγραμμα το εορταστικό κλίμα, το οποίο επικράτησε κατά τη Χριστουγεννιάτικη περίοδο του 2020 (*#christmas*, *#christmasmood*, *#merrychristmas*), αναδεικνύοντας έτσι το αντίστοιχο βεληνεκού ενδιαφέρον του κοινού για τα Χριστούγεννα, με αυτό για την εποχή του καλοκαιριού. Τέλος, με μια πιο προσεκτική ματιά στο Σχήμα 5.68 διακρίνεται το *hashtag #blackfriday*, το οποίο, για πολλοστή φορά, αποδεικνύει τον εμπορικό χαρακτήρα της πλατφόρμας του Instagram σε περιόδους υψηλού καταναλωτισμού.



Σχήμα 5.68: Δημοφιλέστερα hashtags χρονικής περιόδου Νοέμβριος-Δεκέμβριος

Συνοψίζοντας την ανάλυση των δημοφιλέστερων *hashtags* στις επιμέρους χρονικές περιόδους του έτους 2020, είναι αξιοσημείωτη η συνεχής παρουσία των διαγωνισμών (*#giveaway*) στις υψηλότερες θέσεις της κατάταξης, συνοδευόμενη από υπερμεγέθη ποσοστά αλληλεπίδρασης, καθιστώντας, κατ' αυτόν τον τρόπο, κοινώς αποδεκτό πως οι διαγωνισμοί και, κατ' επέκταση, το Instagram αποτελούν ένα κατάλληλο μέσο για τη προβολή και προώθηση προϊόντων και υπηρεσιών.

5.2 Διερεύνηση του YouTube

Ως επόμενο μέσο κοινωνικής δικτύωσης προς εξέταση επιλέχθηκε η πλατφόρμα του YouTube. Η ραγδαία ανάπτυξη των μέσων κοινωνικής δικτύωσης κατά τη διάρκεια της τελευταίας δεκαετίας ευνόησε και το YouTube, με αυτό να αποτελεί τη δεύτερη πιο δημοφιλή μηχανή αναζήτησης μετά το Google και να διατηρεί περισσότερους από 2 δισεκατομμύρια μηνιαίους χρήστες κατά το χρονολογικό έτος 2020-2021 [35]. Το μεγάλο αυτό ενδιαφέρον του κοινού προς τη συγκεκριμένη πλατφόρμα οφείλεται στη ψηφιοποίηση στην οποία υπόκειται τα τελευταία χρόνια ο τομέας της ψυχαγωγίας, καθώς και στην αυξανόμενη ζήτηση των χρηστών για οπτικοακουστικό υλικό. Το YouTube προσφέρεται να καλύψει τις παραπάνω ανάγκες, δίνοντας τη δυνατότητα στους χρήστες να κοινοποιούν βίντεο, μικρής ή μεγάλης διάρκειας, ή να αλληλεπιδρούν με το υλικό άλλων δημιουργών, και στεγάζοντας θεματικό περιεχόμενο που χαρακτηρίζεται από τεράστια ποικιλομορφία, με εκπαιδευτικό, καλλιτεχνικό, ψυχαγωγικό και ενημερωτικό υλικό να διατίθεται προς το κοινό καθημερινά. Αν και ο αρχικός σκοπός του YouTube ήταν η δημιουργία ενός μέσου για την παρακολούθηση βίντεο, η τάση του κοινού προς την ανάπτυξη διαδικτυακών κοινωνικών σχέσεων συνέβαλλε στη μετεξέλιξη της πλατφόρμας σε ένα ευρύτερο κοινωνικό δίκτυο, με χαρακτήρα παρόμοιο με αυτό των υπόλοιπων μέσων δικτύωσης. Παρόμοια με την περίπτωση του Instagram, πολλές ήταν οι εταιρείες που αποφάσισαν να εκμεταλλευτούν την απήχηση του κοινού στο YouTube, διεξάγοντας διαφημιστικές καμπάνιες η προωθώντας τα προϊόντα τους μέσω ευρέως γνωστών δημιουργών, συνθέτοντας έτσι ένα δίκτυο εμπορικών

σχέσεων, το οποίο αφήνει χώρο για βαθύτερη μελέτη.

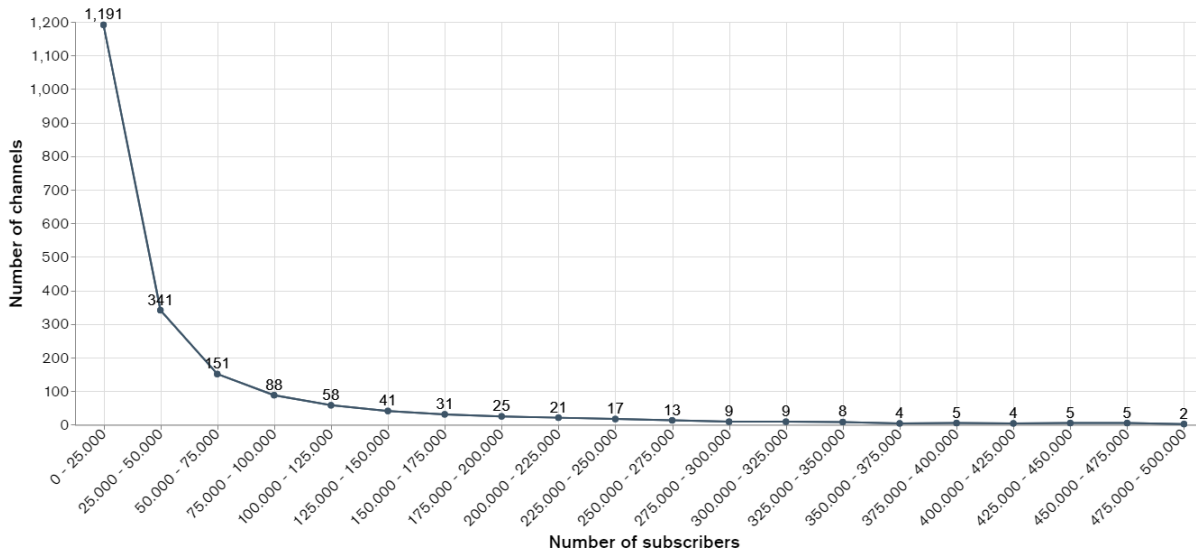
Στην ακόλουθη υποενότητα αναλύονται τα δεδομένα τα οποία συγκεντρώθηκαν από το YouTube και παρουσιάζονται διεξοδικά τα εξαχθέντα συμπεράσματα που αφορούν στη πλατφόρμα αυτή και, ειδικότερα, στη διαδικτυακή συμπεριφορά του ελληνικού κοινού στα πλαίσια αυτής κατά το χρονολογικό έτος 2020.

5.2.1 Ανάλυση των συγκεντρωμένων δεδομένων

Η μελέτη των δεδομένων του YouTube, όπως και στη περίπτωση του Instagram, αποφασίστηκε να διεξαχθεί μεταβαίνοντας από το γενικότερο στο ειδικότερο επίπεδο, ξεκινώντας με την περιγραφή των θεμελιωδών στοιχείων για την ομαλή ανάπτυξη της έρευνας και συνεχίζοντας με την εστίαση σε στοχευμένα σημεία ενδιαφέροντος της πλατφόρμας. Επισημαίνεται πως, τα αποτελέσματα της έρευνας γύρω από την ελληνική κοινότητα του YouTube είναι επηρεασμένα από την δυνατότητα την οποία παρέχει η συγκεκριμένη πλατφόρμα στους χρήστες της να παρακολουθούν βίντεο, χωρίς να απαιτείται η δημιουργία λογαριασμού εκ μέρους τους. Η προκείμενη κατηγορία χρηστών μπορεί να παρακολουθεί οποιοδήποτε βίντεο επιθυμεί, χωρίς ωστόσο να έχει τη δυνατότητα να αλληλεπιδράσει με αυτό ή με τον δημιουργό του, κάνοντας *like*, *comment* ή *subscribe* στο αντίστοιχο κανάλι. Ως εκ τούτου, σε αντίθεση με τις προβολές των βίντεο, αριθμητικά στοιχεία τα οποία βοηθούν στην αποτύπωση της συμπάθειας του κοινού προς συγκεκριμένες θεματικές κατηγορίες περιεχομένου και δημιουργούς παραμένουν αμετάβλητες, αλλοιώνοντας την ακρίβεια των τελικών συμπερασμάτων.

Σημείο εκκίνησης της έρευνας αποτέλεσε η απεικόνιση της κατανομής των τιμών τις οποίες λαμβάνουν οι θεμελιώδεις μεταβλητές του YouTube, με σκοπό την ανίχνευση των ακραίων περιπτώσεων του συλλεγμένου δείγματος και, ως εκ τούτου, την εξαγωγή αντιπροσωπευτικών συμπερασματικών για τη δραστηριότητα της ελληνικής κοινότητας της πλατφόρμας. Συγκεκριμένα, οι μεταβλητές για των οποίων η κατανομή κρίθηκε απαραίτητη η ανάλυση είναι οι εξής: (i) ο αριθμός εγγεγραμμένων χρηστών ενός καναλιού (*subscribers*), (ii) ο αριθμός προβολών ενός βίντεο (*views*), (iii) ο αριθμός των θετικών αντιδράσεων ενός βίντεο (*likes*), (iv) ο αριθμός των σχολίων ενός βίντεο (*comments*) και (v) το ποσοστό αλληλεπίδρασης του κοινού με κάποιο βίντεο.

Ξεκινώντας με την ανάλυση του αριθμού εγγεγραμμένων χρηστών, η αποτύπωση των αντίστοιχων τιμών έλαβε μέρος σε δύο επιμέρους διαγράμματα, ένα για τα κανάλια τα οποία κατέχουν λιγότερους από 500.000 εγγεγραμμένους χρήστες και ένα για εκείνα με περισσότερους εγγεγραμμένους, έτσι ώστε να διασφαλιστεί η ευκρίνεια της οπτικής αναπαράστασης τους. Μέσω της συνδυαστικής εποπτείας των δύο παρακάτω σχημάτων, είναι ορατό πως ο αριθμός των καναλιών παρουσιάζει εκθετική μείωση, μεταβαίνοντας στα υψηλότερα στρώματα εγγεγραμμένων χρηστών. Αναλυτικότερα, επιβεβαιώνεται από το Σχήμα 5.69 πως οι μεγαλύτερη μερίδα των συλλεγμένων καναλιών συναντώνται στην κλίμακα των 0-100.000 εγγεγραμμένων χρηστών.



Σχήμα 5.69: Κατανομή του αριθμού εγγεγραμμένων χρηστών (κάτω των 500K εγγεγραμμένων)

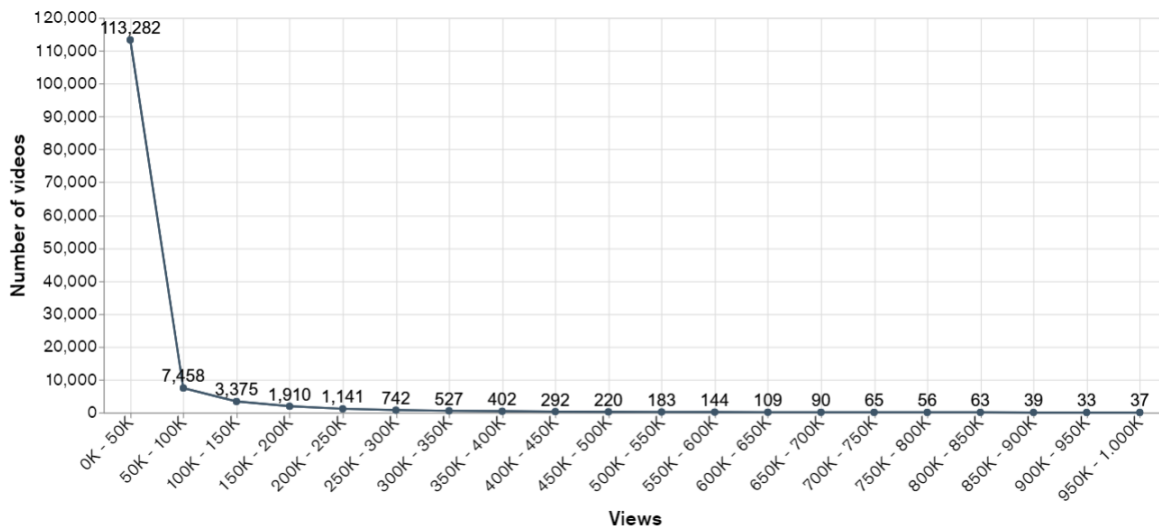
Στην αντίθετη πλευρά της κατάταξης, σύμφωνα με το Σχήμα 5.70, ο υψηλότερος αριθμός εγγεγραμμένων χρηστών, ο οποίος καταγράφηκε στα πλαίσια του δείγματος, ανέρχεται στους 2.500.000. Ωστόσο, παρατηρείται πως τα κανάλια τα οποία ξεπερνούν τους 1.000.000 εγγεγραμμένους χρήστες αγγίζουν μονοψήφιο αριθμό.



Σχήμα 5.70: Κατανομή του αριθμού εγγεγραμμένων χρηστών (άνω των 500K εγγεγραμμένων)

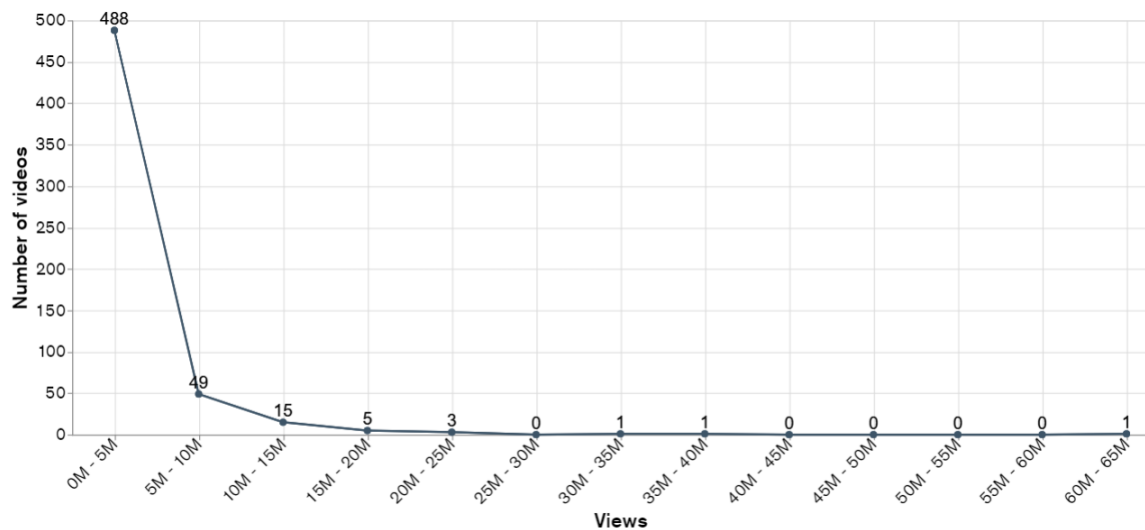
Την αναπαράσταση της κατανομής του αριθμού εγγεγραμμένων χρηστών διαδέχεται η μελέτη της κατανομής του πλήθους προβολών των βίντεο του δείγματος. Με αφορμή την διασκόρπιση των βίντεο με περισσότερες του 1.000.000 προβολές σε μία ευρύτατη κλίμακα και με σκοπό τη βέλτιστη οπτικοποίηση των τιμών, αποφασίστηκε η κατανομή των προβολών να εξετασθεί σε

δύο στάδια. Το πρώτο στάδιο αφορά στα βίντεο με λιγότερες από 1.000.000 προβολές, ενώ το δεύτερο στα περισσότερα γνωστά βίντεο.



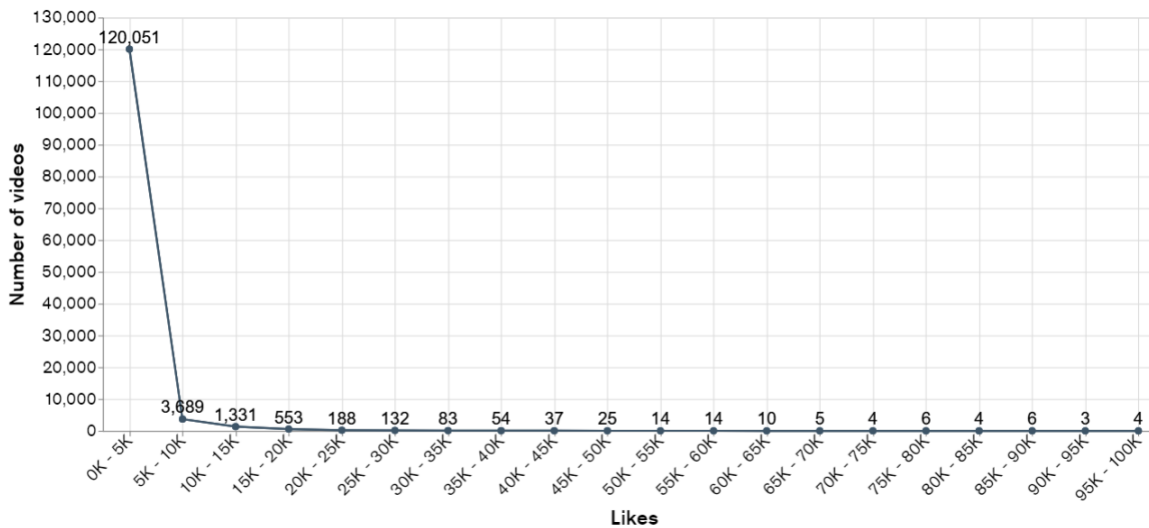
Σχήμα 5.71: Κατανομή του αριθμού προβολών των βίντεο (κάτω των 1M προβολών)

Αντίστοιχα με την κατανομή του αριθμού εγγεγραμμένων χρηστών, το πλήθος των βίντεο διαγράφει εκθετική και, παράλληλα, πτωτική πορεία σε συνάρτηση με την αύξηση των προβολών ενός βίντεο. Με βάση το Σχήμα 5.71, είναι εμφανές πως η πλειοψηφία των αναρτημένων βίντεο εντός του 2020, τα οποία συλλέχθηκαν, συγκέντρωσαν λιγότερες των 250.000 προβολές. Από την άλλη πλευρά, βλέποντας το Σχήμα 5.72, παρατηρείται εξαιρετικά υψηλός αριθμός προβολών (έως και 65.000.000 προβολές) από ένα μικρό υποσύνολο των βίντεο του δείγματος. Εντούτοις, διαπιστώνεται, όπως προαναφέρθηκε, ότι ο περιορισμένος αυτός αριθμός βίντεο είναι κατανομημένος σε διαστήματα τα οποία βρίσκονται σε μεγάλες αποστάσεις μεταξύ τους, όσον αφορά στον αριθμό των προβολών τους.



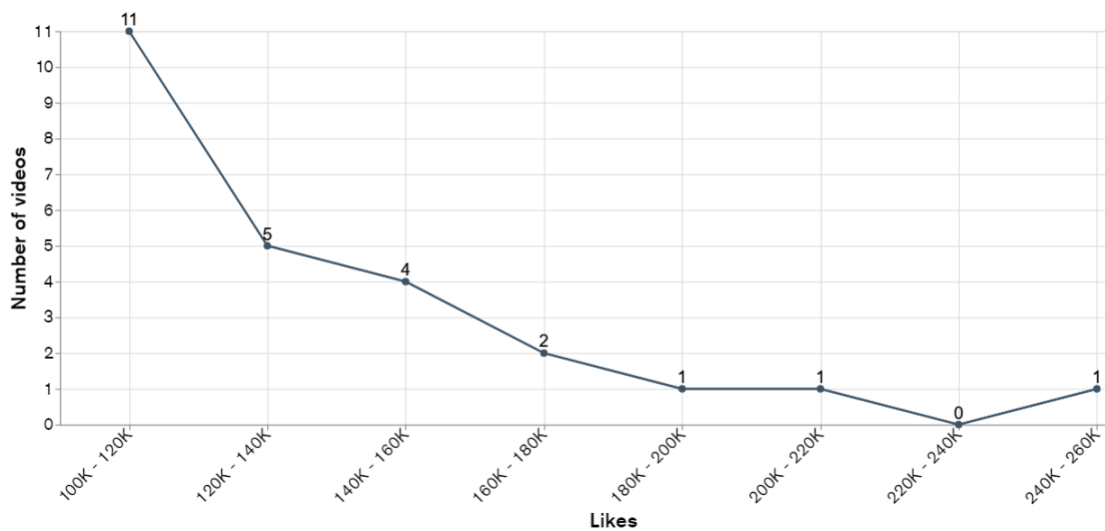
Σχήμα 5.72: Κατανομή του αριθμού προβολών των βίντεο (άνω των 1M προβολών)

Επόμενη μετρική στα πλαίσια της ανάλυσης της κατανομής αποτελεί ο αριθμός των θετικών αντιδράσεων ο οποίος συγκεντρώθηκε από τα, εντός του 2020, αναρτημένα βίντεο. Καθοριστικό ρόλο στην αναπαράσταση της κατανομής των θετικών αντιδράσεων έπαιξε ο καταμερισμός της συγκεκριμένης διαδικασίας σε δύο διαγράμματα. Για ακόμη μία φορά, το μεγάλο εύρος τιμών των θετικών αντιδράσεων, σε συνδυασμό με τη μεγάλη διασπορά των τιμών στα στρώματα άνω των 100.000 θετικών αντιδράσεων, οδήγησε στη διάσπαση του δείγματος με βάση τη τιμή αυτή και, τελικά, στη ξεχωριστή μελέτη των εκατέρωθεν τμημάτων. Παρόμοια με τις παραπάνω μεταβλητές, στο Σχήμα 5.73 παρατηρείται η εκθετική μείωση του αριθμού βίντεο, καθώς αυξάνεται το πλήθος των συγκεντρωμένων θετικών αντιδράσεων. Είναι εμφανές πως η συντριπτική πλειοψηφία των βίντεο του 2020 απέσπασαν έως 5.000 θετικές αντιδράσεις, ενώ ένα σεβαστό υποσύνολο αυτών εισέπραξε από 5.0001 έως 15.000 θετικές αντιδράσεις.



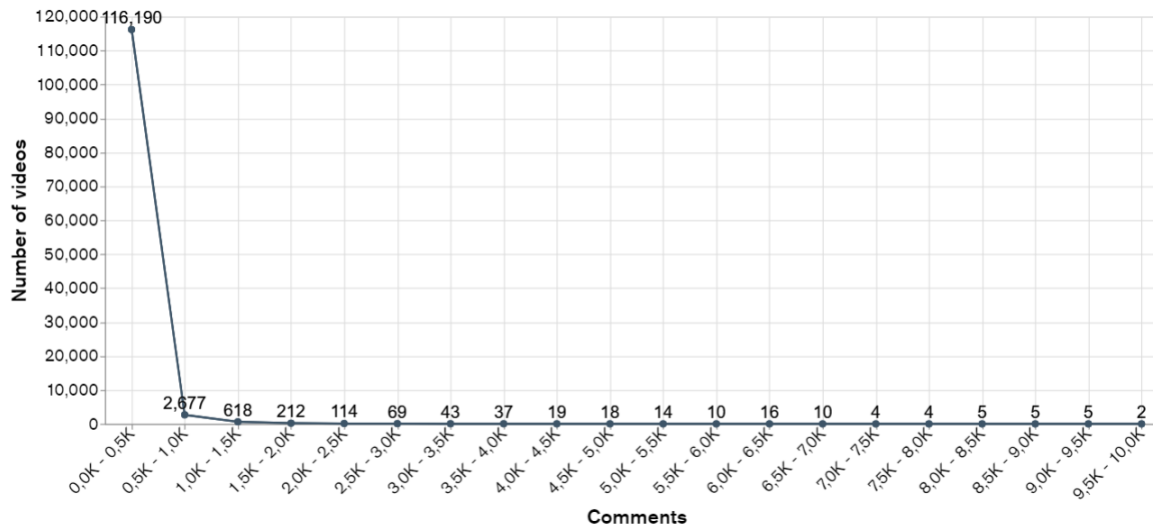
Σχήμα 5.73: Κατανομή του αριθμού θετικών αντιδράσεων των βίντεο (κάτω των 100K θετικών αντιδράσεων)

Ως αναφορά στο σύνολο των βίντεο τα οποία συγκέντρωσαν περισσότερες από 100.000 θετικές αντιδράσεις (βλ. Σχήμα 5.74), η μέγιστη τιμή θετικών αντιδράσεων ανήλθε στις 260.000. Ωστόσο, τα βίντεο τα οποία απαρτίζουν τη συγκεκριμένη κλίμακα αποτελούν τη μειοψηφία του συνολικού δείγματος.



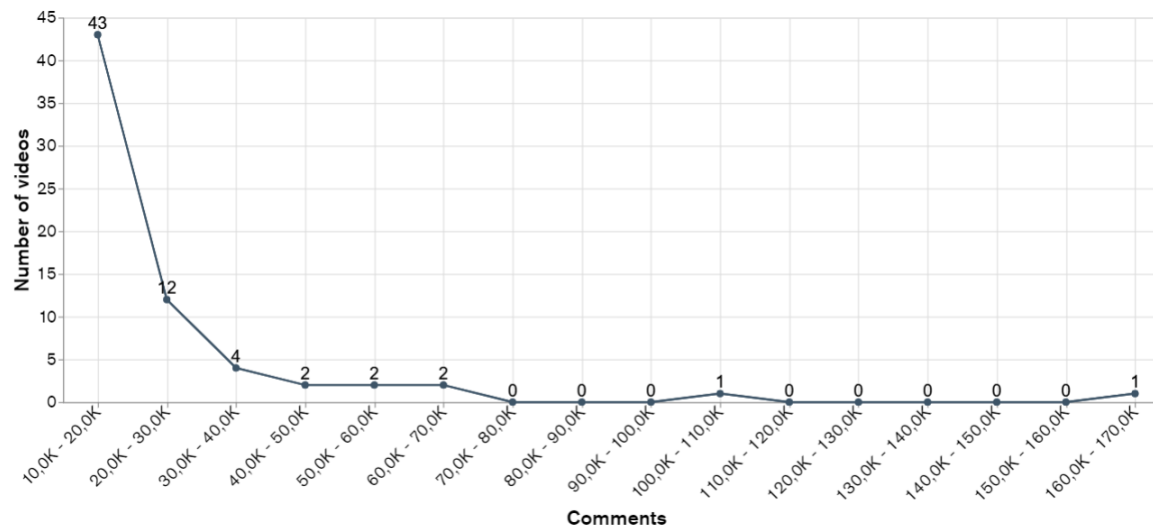
Σχήμα 5.74: Κατανομή του αριθμού θετικών αντιδράσεων των βίντεο (άνω των 100K θετικών αντιδράσεων)

Συνεχίζοντας την ανάλυση των θεμελιωδών μεταβλητών του YouTube, η προσοχή στρέφεται στο πλήθος των σχολιασμών.



Σχήμα 5.75: Κατανομή του αριθμού σχολίων των βίντεο (κάτω των 10K σχολίων)

Με γνώμονα τα Σχήματα 5.75 και 5.76 διαπιστώνεται πως η πλειονότητα των βίντεο τα οποία αναρτήθηκαν εντός του 2020 δε κατάφερε να ξεπεράσει τα 1.000 σχόλια. Παρόλα αυτά, μια αρκετά περιορισμένη μερίδα βίντεο κατόρθωσε να ανελιχθεί στην κατάταξη και να αγγίξει έως και τα 170.000 σχόλια.

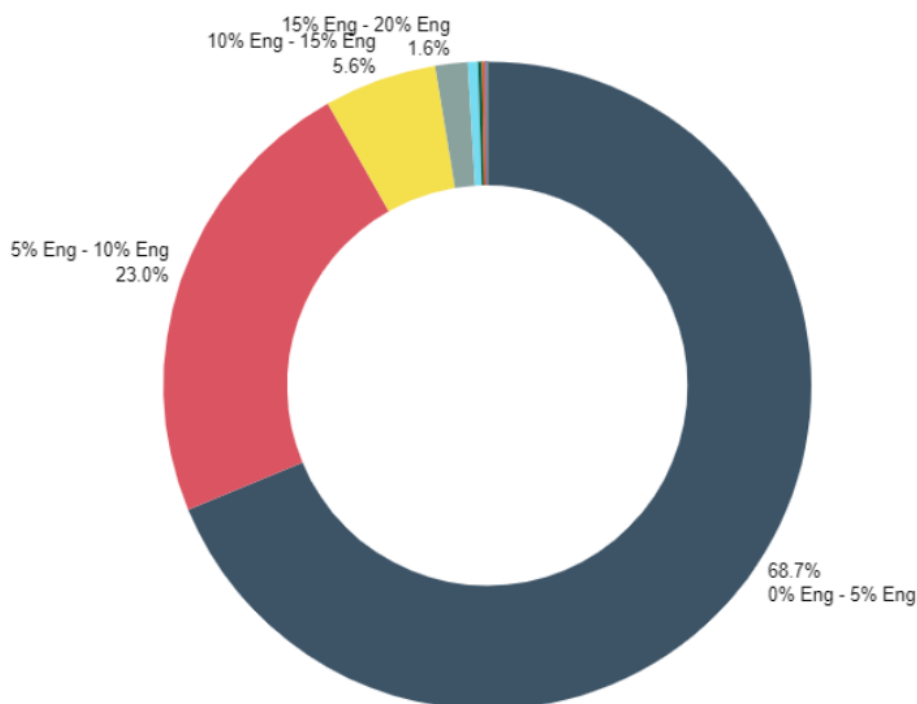


Σχήμα 5.76: Κατανομή του αριθμού σχολίων των βίντεο (άνω των 10K σχολίων)

Τελικό σταθμό της ανάλυσης της κατανομής των διαφόρων μεταβλητών του YouTube αποτελεί η εξέταση του ποσοστού αλληλεπίδρασης του ελληνικού κοινού με τα βίντεο τα οποία δημοσιεύτηκαν κατά το έτος 2020. Υπενθυμίζεται πως βασικά συστατικά στοιχεία της παρούσας μεταβλητής είναι (i) οι θετικές αντιδράσεις ενός βίντεο (*likes*), (ii) τα σχόλια ενός βίντεο (*comments*) και (iii) οι προβολές ενός βίντεο (*views*). Το γεγονός ότι οι τιμές των συγκεκριμένων στοιχείων ποικίλουν και είναι δυνατό να αυξηθούν απεριόριστα επηρέασε τη παραπάνω δευ-

τερογενή μεταβλητή, καθιστώντας δυσδιάκριτες τις ανώτατες τιμές τις οποίες αυτή μπορεί να λάβει.

Παρατηρώντας το Σχήμα 5.77 εξάγεται το συμπέρασμα ότι, κατά μέσο όρο, η πλειοψηφία των ελληνικών βίντεο συγκέντρωσαν ποσοστό αλληλεπίδρασης της τάξης του 0% έως 5%, ενώ μια σημαντική μερίδα αυτών ανήλθε στο 20%. Αναλυτικότερα, επτά στα δέκα βίντεο συγκέντρωσαν άθροισμα θετικών αντιδράσεων και σχολίων μόνο από το 0%-5% των ατόμων που τα παρακολούθησαν, ενώ μόλις 23 στα 100 βίντεο απέσπασαν άθροισμα θετικών αντιδράσεων και σχολίων, στη καλύτερη των περιπτώσεων, 10% των ατόμων που τα παρακολούθησαν. Ακόμα μικρότερο είναι το ποσοστό των αναρτήσεων οι οποίες συναντούνται στη κλίμακα 10%-20% αλληλεπίδρασης, σχεδόν το 7% των συνολικών βίντεο.

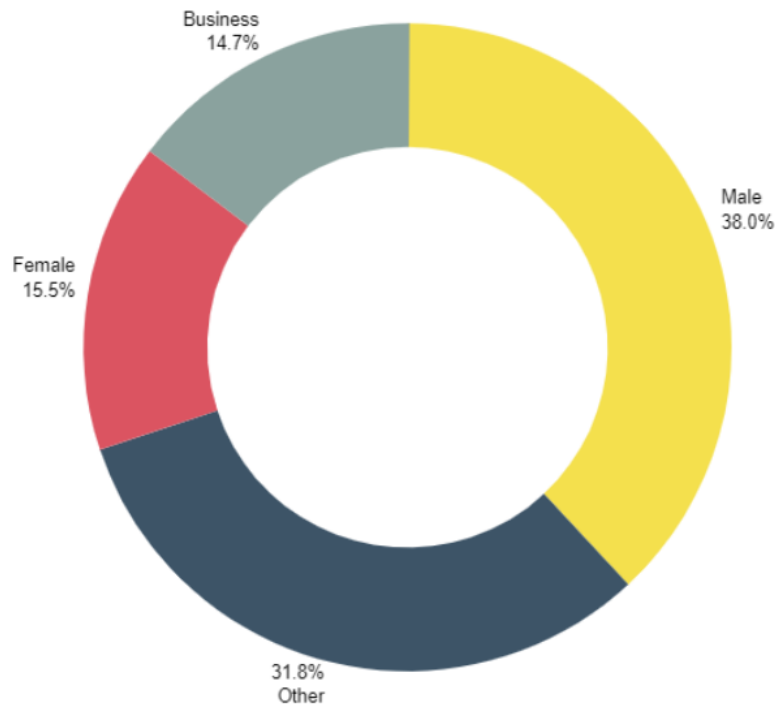


Σχήμα 5.77: Κατανομή του ποσοστού αλληλεπίδρασης

Εφόσον ολοκληρώθηκε η εξέταση των θεμελιωδών μεταβλητών του YouTube, καθίσταται δυνατή η εστίαση σε συγκεκριμένα σημεία ερευνητικού ενδιαφέροντος γύρω από την ελληνική κοινότητα της πλατφόρμας. Σημείο εκκίνησης της προκείμενης διαδικασίας αποτέλεσε η μέτρηση του ποσοστού το οποίο καταλαμβάνει η εκάστοτε κατηγορία χρηστών από το σύνολο των ελληνικών λογαριασμών που συλλέχθηκαν, με βάση τη διάκριση των προφίλ που έλαβε χώρα στο Κεφάλαιο 4 και συγκεκριμένα στην Ενότητα 4.3.

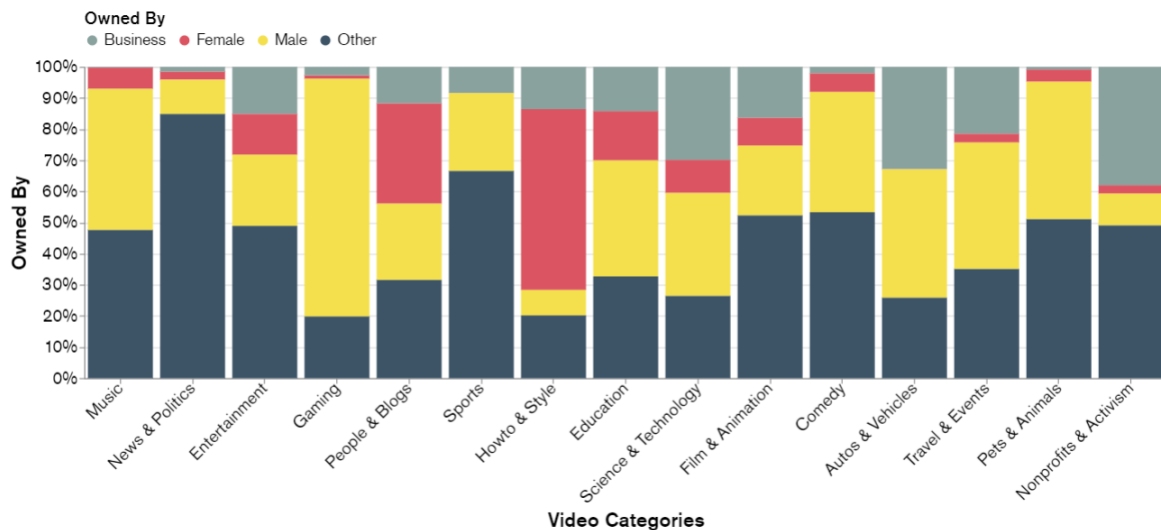
Βάσει του Σχήματος 5.78, προκύπτει ότι το μεγαλύτερο ποσοστό του συνολικού δείγματος καταλαμβάνουν οι ανδρικοί λογαριασμοί (38%), ενώ ακολουθούν οι λογαριασμοί τύπου Άλλο με μικρή διαφορά από τους πρώτους. Λιγότερο έντονη, αλλά αισθητή, είναι η παρουσία γυναικών και επιχειρησιακών λογαριασμών, καθώς αυτοί μοιράζονται, κατά προσέγγιση, το 30% των συλλεγμένων, στα πλαίσια της έρευνας, καναλιών (15,5% και 14,7% αντίστοιχα). Θα μπορούσε, λοιπόν, να εξαχθεί ως ένα πρώιμο συμπέρασμα ότι η πλατφόρμα του YouTube προσελ-

κύει σε μεγαλύτερο βαθμό άνδρες και λογαριασμούς τύπου Άλλο, όπως ομάδες δημιουργών περιεχομένου, τηλεοπτικά κανάλια ή δισκογραφικές εταιρείες.



Σχήμα 5.78: Κατηγορίες χρηστών του YouTube

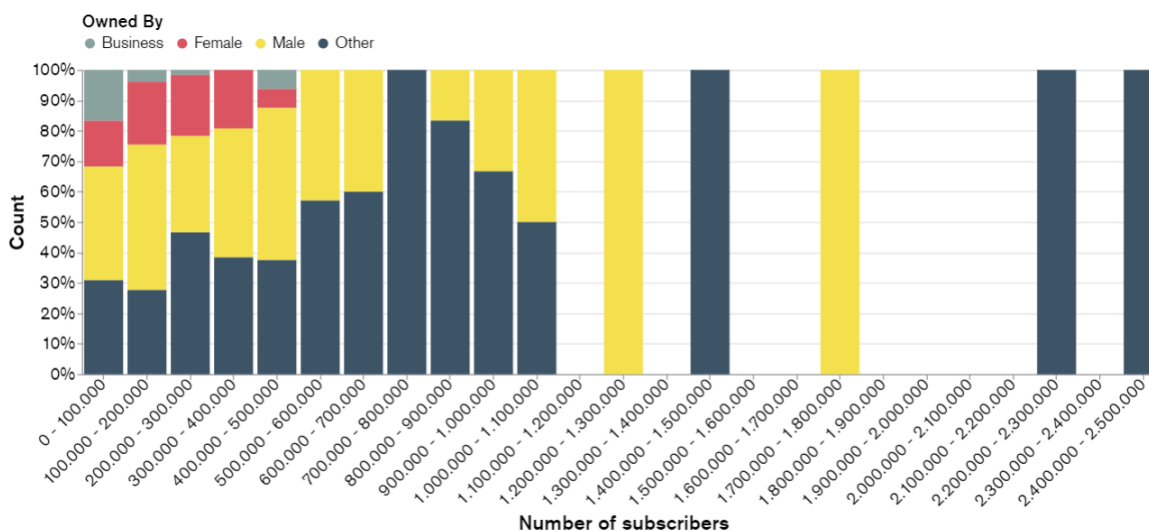
Λαμβάνοντας υπόψη την κατηγοριοποίηση την οποία εφαρμόζει η πλατφόρμα του YouTube στα βίντεο, βασισμένη στον θεματικό τους άξονα, ως επόμενο βήμα επιλέχθηκε η ανάδειξη των κορυφαίων θεματικών κατηγοριών ανά κατηγορία χρηστών (Άνδρας, Γυναίκα, Επιχείρηση, Άλλο) για το έτος 2020.



Σχήμα 5.79: Θεματικές κατηγορίες βίντεο ανά κατηγορία χρηστών

Με μια πρώτη ματιά στο Σχήμα 5.79 είναι φανερό πως, κατά κύριο λόγο, οι τέσσερις προαναφερθείσες κατηγορίες χρηστών εκδήλωσαν ενδιαφέρον προς όλες τις θεματικές κατηγορίες βίντεο. Ωστόσο, παρατηρείται πως η εκάστοτε κατηγορία χρηστών παρουσίασε μια ροπή προς συγκεκριμένες θεματικές. Ξεκινώντας με τους άνδρες, οι κατηγορίες στις οποίες έκαναν αρκετά αισθητή την παρουσία τους είναι: (i) το Ηλεκτρονικό Παιχνίδι (*Gaming*), (ii) η Μουσική (*Music*), (iii) το Εκπαιδευτικό περιεχόμενο (*Education*), (iv) το Επιστημονικό/Τεχνολογικό περιεχόμενο (*Science & Technology*) και (v) η Αυτοκίνηση (*Autos & Vehicles*). Ακολουθώντας, τα κανάλια τύπου Άλλο εμφάνισαν μία τάση προς τις κατηγορίες: (i) Ενημέρωση/Πολιτική (*News & Politics*), (ii) Άθληση (*Sports*), (iii) Ψυχαγωγία (*Entertainment*), (iv) Μουσική (*Music*) και (v) Κωμωδία (*Comedy*). Με τη σειρά τους, οι γυναίκες υπερτέρησαν στις εξής θεματικές κατηγορίες: (i) Στυλιστικό περιεχόμενο (*Howto & Style*) και (ii) Καθημερινότητα και Ανθρώπινες Σχέσεις (*People & Blogs*). Ως τελευταία ομάδα χρηστών, οι επιχειρήσεις έδειξαν ιδιαίτερη προτίμηση προς τις κατηγορίες περιεχομένου: (i) Επιστήμες/Τεχνολογία (*Science & Technology*), (ii) Αυτοκίνηση (*Autos & Vehicles*), (iii) Ταξιδιωτικές Δραστηριότητες (*Travel & Events*) και (iv) Ακτιβισμός (*Nonprofits & Activism*).

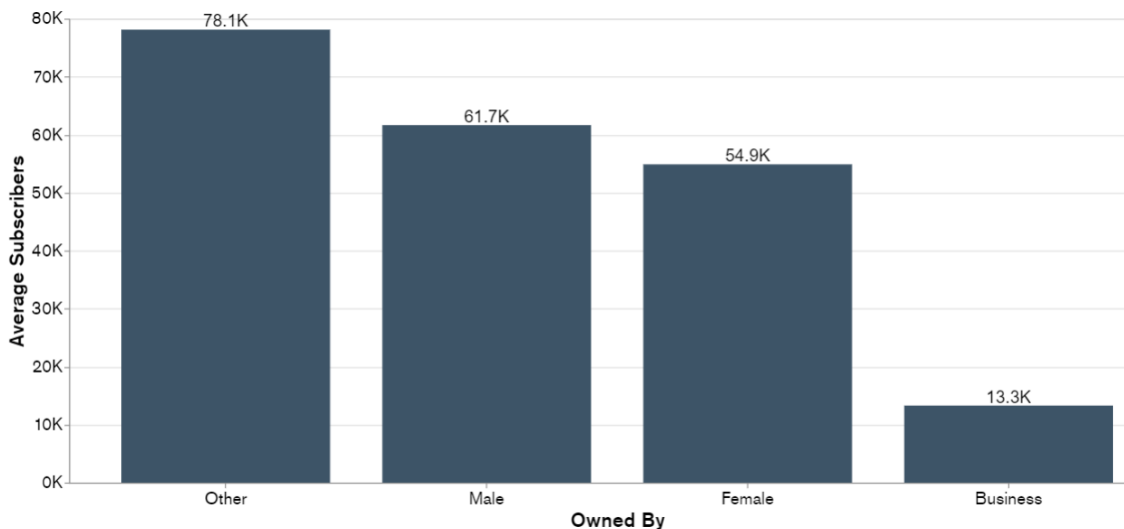
Τη πρωταρχική χαρτογράφηση της ελληνικής κοινότητας του YouTube, η οποία πραγματοποιήθηκε νωρίτερα, διαδέχεται το προπαρασκευαστικό στάδιο της διερεύνησης των προτιμήσεων των Ελλήνων χρηστών, μέσω της αξιοποίησης του αριθμού εγγεγραμμένων χρηστών των καναλιών του δείγματος. Ο αριθμός των εγγεγραμμένων χρηστών ενός καναλιού εκφράζει το πλήθος των ατόμων τα οποία έχουν επιλέξει να ενημερώνονται τακτικά για το νέο περιεχόμενο του και, κατά συνέπεια, αποτελεί μία ένδειξη του ποσοστού ενδιαφέροντος του κοινού. Με αφορμή το γεγονός αυτό, είναι δυνατό να εξετασθεί ο αριθμός εγγεγραμμένων χρηστών συναρτήσει των διαφορετικών κατηγοριών χρηστών και κατηγοριών θεματικού περιεχομένου.



Σχήμα 5.80: Κατηγορίες χρηστών ανά αριθμό εγγεγραμμένων χρηστών

Σύμφωνα με το Σχήμα 5.80, διαπιστώνεται πως στα ανώτερα στρώματα του αριθμού εγγεγραμμένων χρηστών κυριαρχούν κανάλια διαχειριζόμενα από άνδρες και λογαριασμοί τύπου Άλλο, αναδεικνύοντας, με αυτόν τον τρόπο, την υπεροχή των δύο αυτών κατηγοριών χρηστών στη πλατφόρμα του YouTube από μία διαφορετική οπτική γωνία, σε σχέση με το Σχήμα 5.78. Από

την άλλη πλευρά, οι γυναίκες και οι επιχειρήσεις διατηρούν ένα αρκετά πιο περιορισμένο κοινό, ως αναφορά στους εγγεγραμμένους χρήστες, με τους 500.000 *subscribers* να αποτελούν το ταβάνι και των δύο κατηγοριών. Ωστόσο, το ποσοστό των επιχειρησιακών λογαριασμών, οι οποίοι καταφέρνουν να ανελιχθούν στα υψηλότερα στρώματα εγγεγραμμένων, είναι μικρό, καθώς οι περισσότερες επιχειρήσεις συναντούνται στην κλίμακα 0-100.000 εγγεγραμμένων, φαινόμενο το οποίο υποδηλώνει τη δυσκολία ανάπτυξης ενός εδραιωμένου κοινού εκ μέρους των επιχειρήσεων.

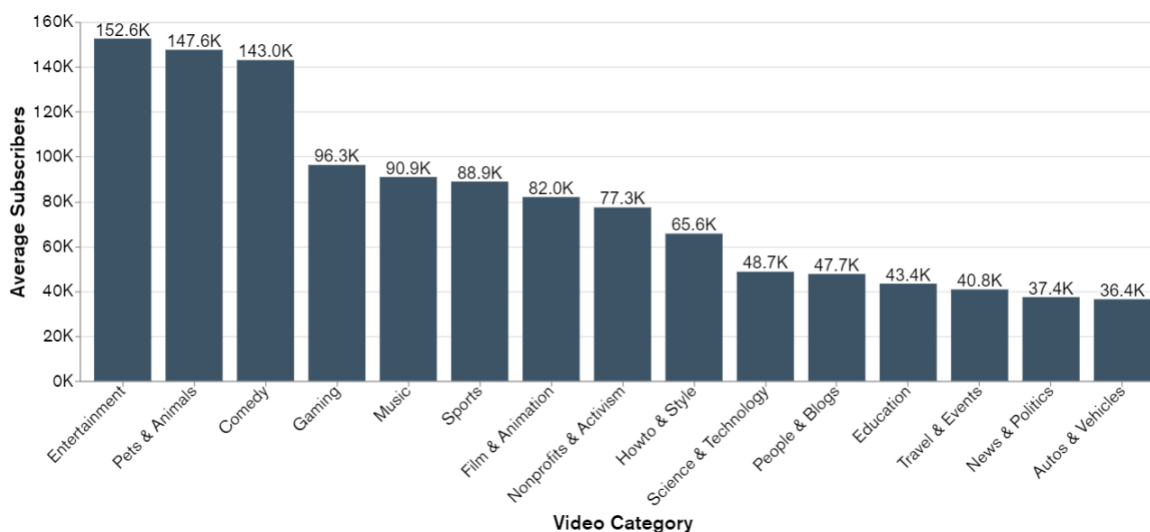


Σχήμα 5.81: Μέσος αριθμός εγγεγραμμένων χρηστών ανά κατηγορία χρηστών

Παρακολουθώντας τις μέσες τιμές εγγεγραμμένων χρηστών ανά τις διάφορες κατηγορίες χρηστών (βλ. Σχήμα 5.81), διακρίνεται πως παρά τις περιπτώσεις αντρικών καναλιών με πολύ υψηλό αριθμό εγγεγραμμένων και το περιορισμένο κοινό των γυναικών, οι δύο αυτές ομάδες χρηστών διατηρούν, κατά μέσο όρο, παρόμοιο αριθμό εγγεγραμμένων, ο οποίος κυμαίνεται στην κλίμακα των 50.000-60.000 *subscribers*. Τα δύο άκρα της κατάταξης καταλαμβάνονται από τους λογαριασμούς τύπου Άλλο και τα επιχειρησιακά κανάλια, με τους πρώτους να ανέρχονται προσεγγιστικά στους 78.100 εγγεγραμμένους και τα τελευταία στους 13.300 εγγεγραμμένους, επιβεβαιώνοντας το αυξημένο ενδιαφέρον του κοινού για το περιεχόμενο των καναλιών τύπου Άλλο και τη αδυναμία των επιχειρήσεων για την προσέλκυση αυτού.

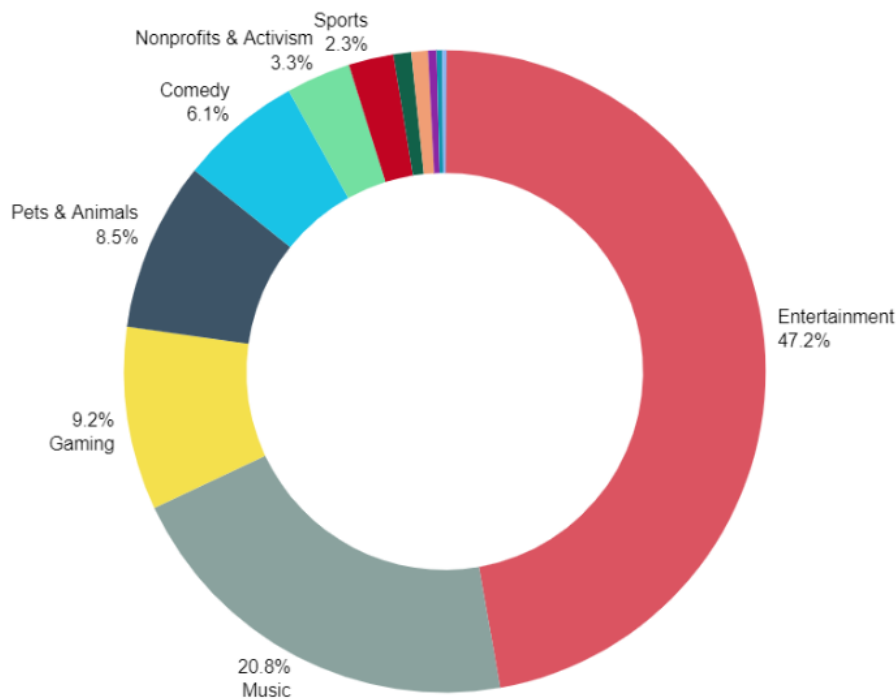
Η μετάβαση της μελέτης στις θεματικές κατηγορίες βίντεο επιτυγχάνεται μέσω του Σχήματος 5.82. Ο υπολογισμός του μέσου αριθμού εγγεγραμμένων χρηστών για κάθε μία από τις θεματικές κατηγορίες είναι δυνατό να προσδιορίσει ποιες από αυτές εμφανίζονται προσεγγιστικά σε κανάλια μεγαλύτερης ή μικρότερης εμβέλειας και, συνεπώς, αφορούν μια μεγαλύτερη ή μικρότερη μερίδα του κοινού. Η εφαρμογή αυτού του κριτηρίου στη κατάταξη του παρακάτω σχήματος συγκεντρώνει την προσοχή στις θεματικές κατηγορίες *Ψυχαγωγία (Entertainment)*, *Ζώα και Κατοικίδια (Pets & Animals)* και *Ηλεκτρονικό Παιχνίδι (Gaming)*, η οποίες κατακτούν τις κορυφαίες θέσεις της κατάταξης με εμφανή διαφορά από τις υπόλοιπες κατηγορίες, αντικατοπτρίζοντας πιθανώς την αρέσκεια της πλειοψηφίας των Ελλήνων χρηστών σε αυτές τις θεματικές βίντεο. Σε αρκετά υψηλές θέσεις της κατάταξης συναντούνται επίσης οι κατηγορίες *Ηλεκτρονικό Παιχνίδι (Gaming)*, *Μουσική (Music)* και *Άθληση (Sports)*. Στον αντίποδα, τον μικρότερο μέσο αριθμό εγγεγραμμένων χρηστών σημειώνουν οι κατηγορίες *Νέα και Πολιτική (News)*

& Politics) και Αυτοκίνηση (Autos & Vehicles), υποδεικνύοντας τη λιγότερη ταύτιση του κοινού με αυτές τις θεματικές κατηγορίες περιεχομένου.



Σχήμα 5.82: Μέσος αριθμός εγγεγραμμένων χρηστών ανά θεματική κατηγορία περιεχομένου

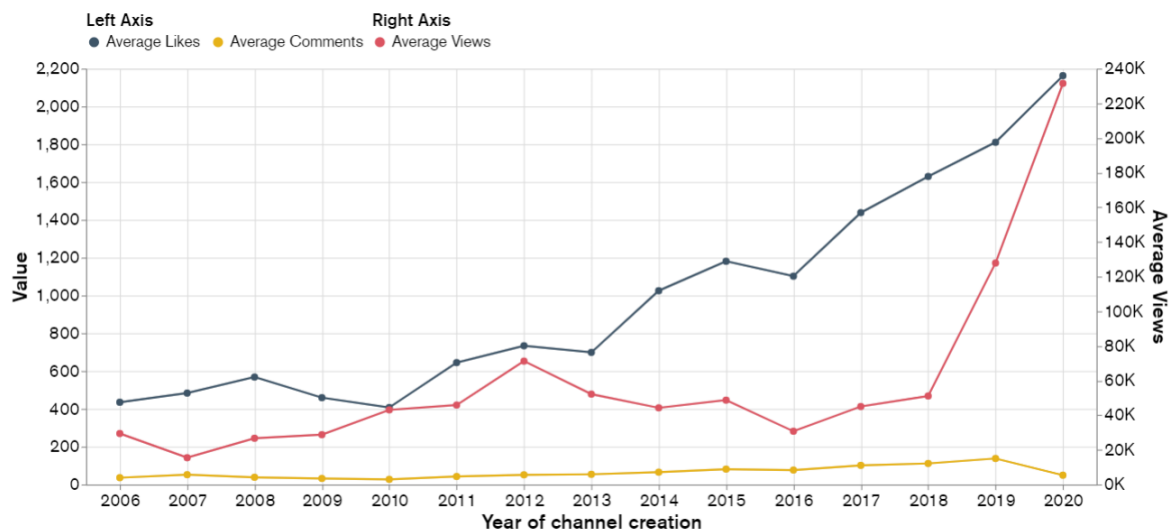
Προς ενίσχυση των παραπάνω παρατηρήσεων αναζητήθηκαν οι θεματικές κατηγορίες περιεχομένου οι οποίες εμφανίστηκαν με μεγαλύτερη συχνότητα σε κανάλια άνω των 400.000 χιλιάδων εγγεγραμμένων χρηστών, κατά το έτος 2020. Αντίστοιχα με το προηγούμενο διάγραμμα, ο στόχος του Σχήματος 5.83 είναι η ανάδειξη των δημοφιλέστερων θεματικών κατηγοριών περιεχομένου, μέσω της μελέτης των καναλιών τα οποία συγκεντρώνουν τη μεγαλύτερη μερίδα του ελληνικού κοινού. Φαίνεται, λοιπόν, από το συγκεκριμένο σχήμα πως σχεδόν τα μισά βίντεο των μεγαλύτερων καναλιών κατατάσσονται στον κλάδο της ψυχαγωγίας (*Entertainment*), ενώ αισθητή είναι και η παρουσία μουσικών βίντεο (*Music*). Την εμφάνιση κάνουν επίσης οι κατηγορίες *Ηλεκτρονικό Παιχνίδι (Gaming)*, *Ζώα και Κατοικίδια (Pets & Animals)* και *Κωμωδία (Comedy)*, επιβεβαιώνοντας τις αρχικές θέσεις της κατάταξης του Σχήματος 5.82. Τέλος, μικρότερη αλλά αξιοσημείωτη είναι η στροφή των μεγαλύτερων καναλιών προς ακτιβιστικό (*Nonprofits & Activism*) και αθλητικό (*Sports*) περιεχόμενο.



Σχήμα 5.83: Θεματικές κατηγορίες περιεχομένου σε μεγάλα κανάλια (άνω των 400K εγγεγραμμένων χρηστών)

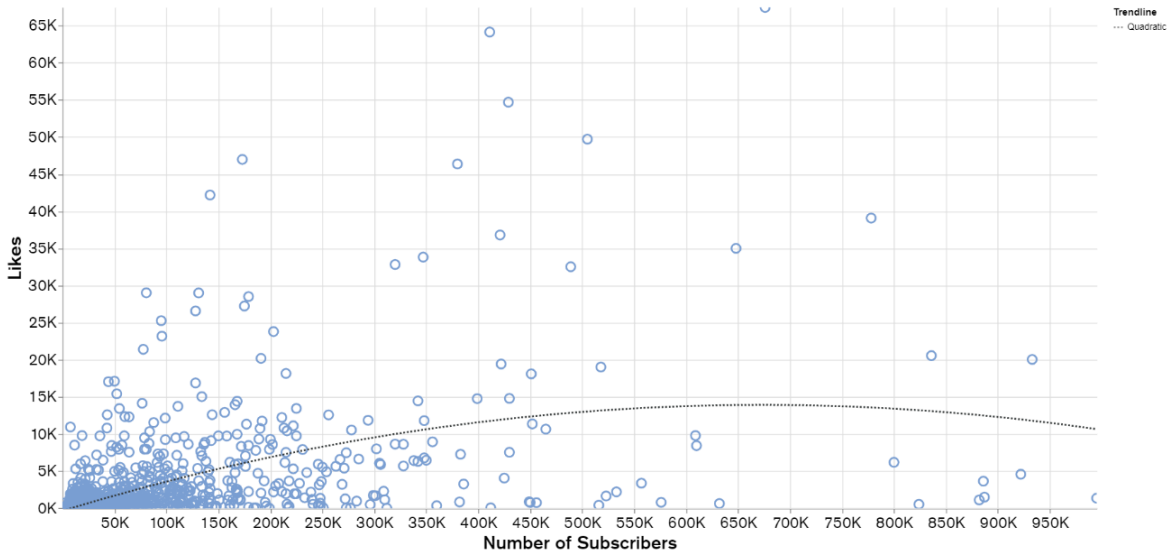
Σε συνέχεια της ερευνητικής διαδικασίας, το ενδιαφέρον επικεντρώνεται στην εξέταση των παραγόντων, οι οποίοι επηρεάζουν τις ποικίλες μετρικές απόδοσης των βίντεο, και, κατά συνέπεια, το ποσοστό αλληλεπίδρασης του κοινού με αυτά. Αναλυτικότερα, οι μεταβλητές των οποίων κρίθηκε απαραίτητη η ανάλυση είναι: (i) ο αριθμός θετικών αντιδράσεων (*Likes*), (ii) ο αριθμός σχολίων (*Comments*), (iii) ο αριθμός προβολών (*Views*) και (iv) το ποσοστό αλληλεπίδρασης (*Average Engagement*), το οποίο προκύπτει ως συνδυασμός των παραπάνω.

Ως πρώτος παράγοντας προς μελέτη επιλέχθηκε το έτος δημιουργίας του εκάστοτε καναλιού, συνοδευόμενο από τις μετρικές απόδοσης των αναρτημένων βίντεο του, κατά τη διάρκεια του έτους 2020. Εστιάζοντας, αρχικά, στον μέσο αριθμό θετικών αντιδράσεων και σχολίων του Σχήματος 5.84, γίνεται αντιληπτό πως τα πιο πρόσφατα δημιουργημένα κανάλια συγκέντρωσαν μεγαλύτερο πλήθος αντιδράσεων στα βίντεο τους, σε σχέση με εκείνα τα οποία χρονολογούνται σε παλαιότερα έτη δημιουργίας. Επιπροσθέτως, μεγάλη αύξηση σημειώνεται στον μέσο αριθμό προβολών των καναλιών, τα οποία δημιουργήθηκαν κατά τα χρονολογικά έτη 2019 και 2020. Το φαινόμενο αυτό πιθανώς να εκφράζει την καλύτερη ανταπόκριση των νεότερων καναλιών στις σύγχρονες ανάγκες του ελληνικού κοινού (ποιότητα και θεματική περιεχομένου) ή ακόμα και την είσοδο νέου κοινού στην πλατφόρμα του YouTube κατά τα τελευταία έτη.



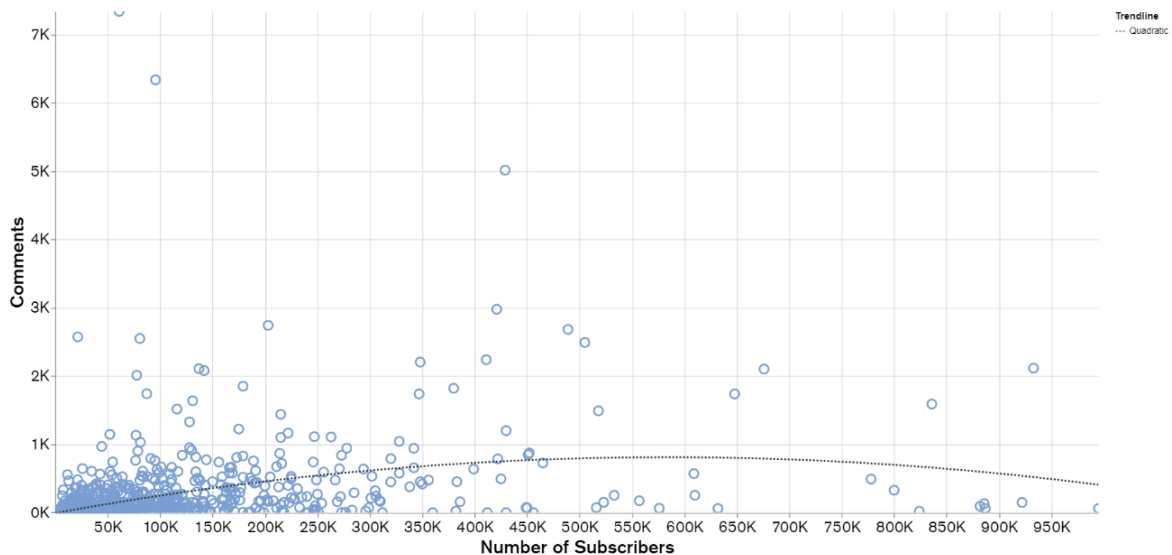
Σχήμα 5.84: Βαθμός αλληλεπίδρασης ανά έτος δημιουργία των καναλιού

Επόμενος, κατά σειρά, παράγοντας, του οποίου η επιρροή στις μετρικές απόδοσης ενός βίντεο επιλέχθηκε να εξετασθεί, είναι ο αριθμός των εγγεγραμμένων χρηστών ενός καναλιού. Αξίζει να σημειωθεί ότι για τις ανάγκες των παρακάτω διαγραμμάτων και με σκοπό την αποφυγή ακραίων περιπτώσεων, ο αριθμός των εγγεγραμμένων χρηστών περιορίστηκε στους 1.000.000, ενώ το ποσοστό αλληλεπίδρασης στο 20% (βλ. Σχήμα 5.77). Ξεκινώντας με τη σχέση μεταξύ του αριθμού θετικών αντιδράσεων και του πλήθους των εγγεγραμμένων χρηστών, διαπιστώνεται πως κανάλια μεγαλύτερης εμβέλειας τείνουν να συγκεντρώνουν, κατά μέσο όρο, υψηλότερο αριθμό θετικών αντιδράσεων, σε σχέση με αυτά των χαμηλότερων στρωμάτων εγγεγραμμένων. Εντούτοις, η ευρύτερη αύξηση του αριθμού των εγγεγραμμένων ενός καναλιού φαίνεται να μην επιφέρει ανάλογο κέρδος στον αριθμό των *likes*, με τη σταδιακή μείωση αυτών να είναι ορατή στα ανώτατα επίπεδα του ακόλουθου σχήματος. Σημειώνεται πως η συγκεκριμένη πτωτική πορεία πιθανώς δεν είναι απολύτως αντιπροσωπευτική, λόγω του μικρού αριθμού καναλιών στα υψηλότερα στρώματα εγγεγραμμένων χρηστών.



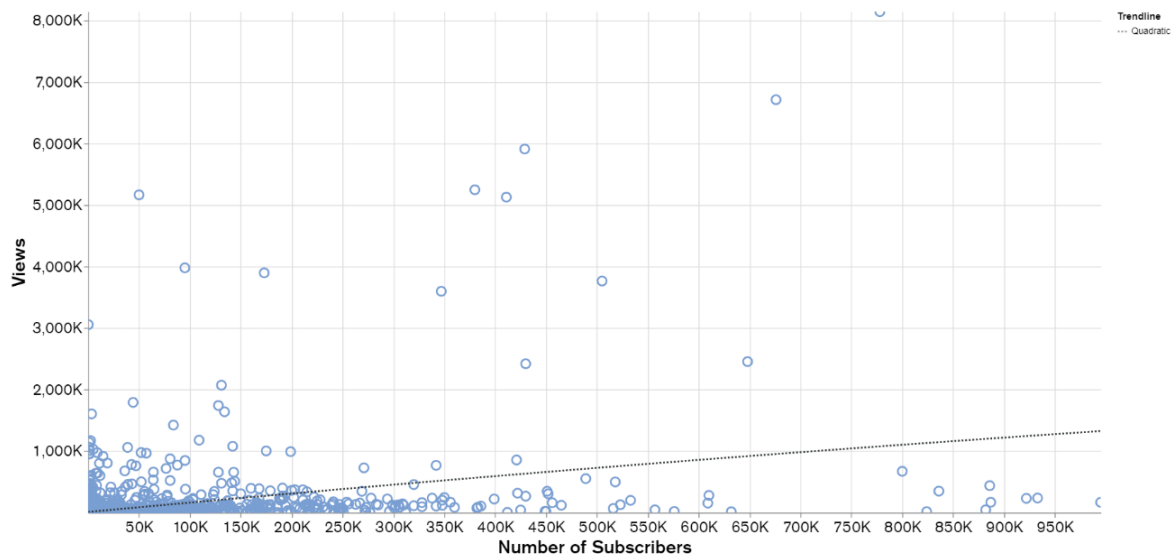
Σχήμα 5.85: Σχέση αριθμού θετικών αντιδράσεων και αριθμού εγγεγραμμένων χρηστών

Όσον αφορά στον αριθμό σχολίων τα οποία λαμβάνει ένα βίντεο, παρατηρείται πως διαγράφουν πανομοιότυπη πορεία με εκείνη των θετικών αντιδράσεων μεταβαίνοντας προς τους μεγαλύτερους αριθμούς εγγεγραμμένων χριστών, δείχνοντας, αντιστοίχως, πως κανάλια μεγαλύτερης εμβέλειας τείνουν να συγκεντρώνουν, κατά μέσο όρο, υψηλότερο αριθμό σχολίων. Ωστόσο, η ευρύτερη αύξηση του αριθμού των εγγεγραμμένων ενός καναλιού φαίνεται να μην επιφέρει ανάλογο κέρδος στον αριθμό των *comments*, με τη σταδιακή μείωση αυτών να είναι ορατή στα ανώτατα επίπεδα του παρακάτω σχήματος. Αξίζει να σημειωθεί πως η συγκεκριμένη πτωτική πορεία ενδέχεται να μην είναι απολύτως αντιπροσωπευτική, λόγω του μικρού αριθμού καναλιών στα υψηλότερα στρώματα εγγεγραμμένων χρηστών.



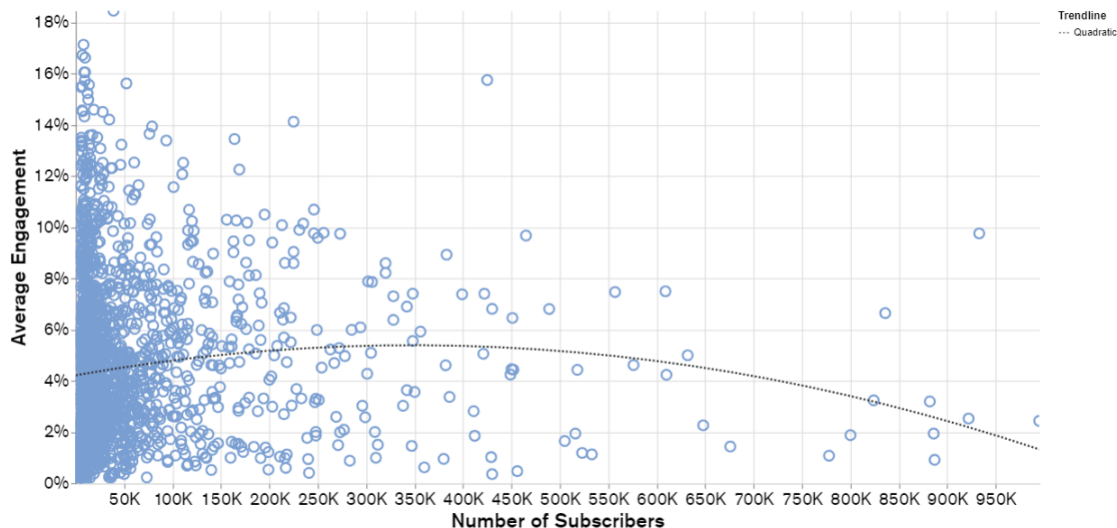
Σχήμα 5.86: Σχέση σχολίων και αριθμού εγγεγραμμένων χρηστών

Παρά τη συντηρητικά ανοδική πορεία η οποία διακατέχει τη σχέση μεταξύ των δύο παραπάνω μετρικών και του αριθμού των εγγεγραμμένων χρηστών και σύμφωνα με το Σχήμα 5.87, φαίνεται πως ο αριθμός προβολών των βίντεο παρουσιάζει μια πιο ξεκάθαρη αυξητική τάση. Παρόλα αυτά, το συγκεκριμένο γεγονός δεν αποτελεί τον κανόνα, καθώς αξιόλογος είναι ο αριθμός των περιπτώσεων κατά τις οποίες κανάλια με μικρότερο αριθμό εγγεγραμμένων χρηστών σημείωσαν υψηλό αριθμό προβολών κατά το 2020.



Σχήμα 5.87: Σχέση αριθμού προβολών και αριθμού εγγεγραμμένων χρηστών

Με γνώμονα τις πορείες των τριών παραπάνω μεταβλητών και γνωρίζοντας ότι το ποσοστό αλληλεπίδρασης του κοινού με ένα βίντεο προκύπτει από την διαίρεση του αθροίσματος των θετικών αντιδράσεων και των σχολίων με το πλήθος προβολών του βίντεο, καθίσταται δυνατή η πρόβλεψη της εξέλιξης του ποσοστού αλληλεπίδρασης σε συνάρτηση με τον αριθμό εγγεγραμμένων χρηστών ενός καναλιού, η οποία αποτυπώνεται στο Σχήμα 5.88. Με βάση το προκείμενο διάγραμμα προκύπτει ότι, κατά κύριο λόγο, κανάλια με μεγαλύτερο αριθμό εγγεγραμμένων χρηστών διατηρούν μία καλύτερη σχέση με το κοινό τους, συγκριτικά με τα κανάλια μικρότερης εμβέλειας. Ωστόσο, η παρατήρηση αυτή παύει να ισχύει για κανάλια με εκτενές εγγεγραμμένο κοινό, γεγονός το οποίο είναι δικαιολογημένο, αν λάβει κανείς υπόψη του τη δυσκολία διατήρησης ενός αναλογικά μεγάλου ενεργού κοινού από κανάλια ευρέως βεληνεκούς. Καταληκτικά, επισημαίνεται πως η ύπαρξη καναλιών με αρκετά υψηλό ποσοστό αλληλεπίδρασης στα χαμηλότερα στρώματα εγγεγραμμένων χρηστών οφείλεται στην απόσπαση ενός ποσοστιαία μεγάλου αριθμού θετικών αντιδράσεων και σχολίων από βίντεο με αρκετά περιορισμένο αριθμό προβολών.



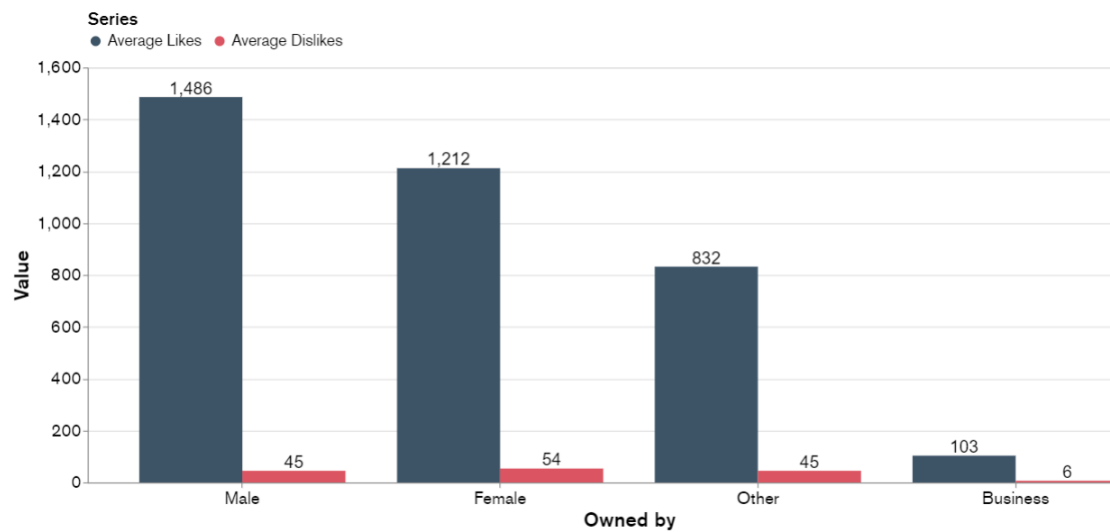
Σχήμα 5.88: Σχέση μέσου ποσοστού αλληλεπίδρασης και αριθμού εγγεγραμμένων χρηστών

Εφόσον εξετάστηκε η σχέση μεταξύ των μετρικών απόδοσης των βίντεο του YouTube και του αριθμού εγγεγραμμένων χρηστών, ως επόμενος στόχος τέθηκε η μελέτη της επιρροής της κατηγορίας χρήστη, στην οποία κατατάσσεται ο διαχειριστής του εκάστοτε καναλιού, στη πορεία των μετρικών αυτών.

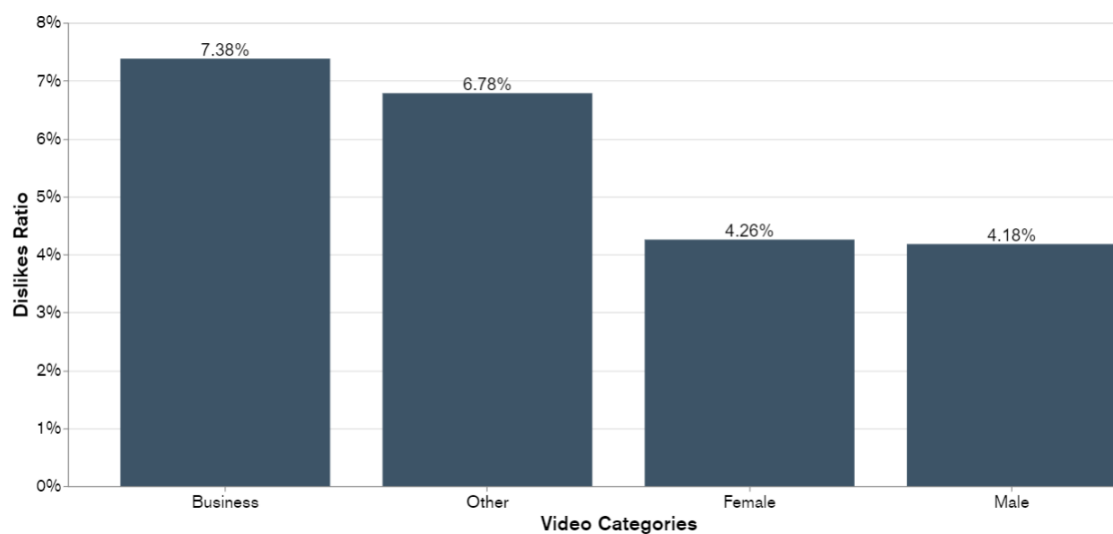
Αρχικό σταθμό της προκείμενης διαδικασίας αποτέλεσαν οι θετικές και αρνητικές αντιδράσεις των βίντεο του δείγματος. Η ανάγκη για αποφυγή της αλλοίωσης των τελικών αποτελεσμάτων από ακραίες περιπτώσεις οδήγησε στη παράλειψη των βίντεο τα οποία ξεπερνούσαν τις 15.000 θετικές αντιδράσεις. Η εφαρμογή του κριτηρίου αυτού διαμόρφωσε τη κατάταξη του Σχήματος 5.89, σύμφωνα με την οποία τη κορυφαία θέση του μέσου αριθμού θετικών αντιδράσεων κατέλαβαν οι άνδρες χρήστες του YouTube, διατηρώντας μικρή τη διαφορά τους από τις δεύτερες στη κατάταξη γυναίκες. Σε χαμηλότερα επίπεδα συναντούνται οι λογαριασμοί τύπου Άλλο και οι επιχειρησιακοί λογαριασμοί, με τους τελευταίους να κατέχουν το αρνητικό ρεκόρ του χαμηλότερου μέσου αριθμού θετικών αντιδράσεων για το έτος 2020.

Από τη μεριά των αρνητικών αντιδράσεων, στη πρώτη θέση βρίσκονται οι γυναικείοι λογαριασμοί με ελάχιστη, όμως, διαφορά από τους άνδρες και τους κανάλια τύπου Άλλο. Όπως και στη περίπτωση των θετικών αντιδράσεων, ουραγό της κατάταξης του μέσου αριθμού *dislikes* συνιστούν οι επιχειρήσεις. Με μία πρώτη ματιά, θα μπορούσε να ειπωθεί πως οι επιχειρήσεις λαμβάνουν στα βίντεο τους τον χαμηλότερο αριθμό αρνητικών αντιδράσεων, ωστόσο κάτι τέτοιο θα θεωρούνταν άστοχο, καθώς, όπως φαίνεται στο Σχήμα 5.90, οι επιχειρήσεις, κατά το έτος 2020, σημείωσαν το υψηλότερο ποσοστό αρνητικών αντιδράσεων, το οποίο άγγιξε το 7%. Το ποσοστό αυτό εκφράζει ότι, κατά μέσο όρο, αν ένα βίντεο επιχειρησιακού καναλιού αποσπούσε 100 θετικές και αρνητικές αντιδράσεις συνολικά, τότε οι επτά από αυτές θα αποτελούσαν *dislikes*. Υψηλό ποσοστό αρνητικών αντιδράσεων σημείωσαν επίσης οι λογαριασμοί τύπου Άλλο, ενώ οι άνδρες και οι γυναίκες παρέμειναν χαμηλότερα, με το ποσοστό τους να βρίσκεται κοντά στο 4%. Θα άξιζε να παρατηρηθεί από τα παραπάνω δεδομένα πως κανάλια τα οποία δε βασίζονται στη ψηφιακή παρουσία ενός συγκεκριμένου προσώπου/διαχειριστή έχουν την τάση να συγκεντρώνουν ευκολότερα αρνητικές αντιδράσεις σε σχέση τα προσωπικά κανάλια (άνδρες, γυναίκες). Πάντως, η ρίζα του φαινομένου αυτού είναι πιθανό να έγκειται σε περαιτέρω πα-

ράγοντες, όπως οι θεματική κατηγορία περιεχομένου του βίντεο ή η διαδικτυακή συμπεριφορά του εκάστοτε δημιουργού.

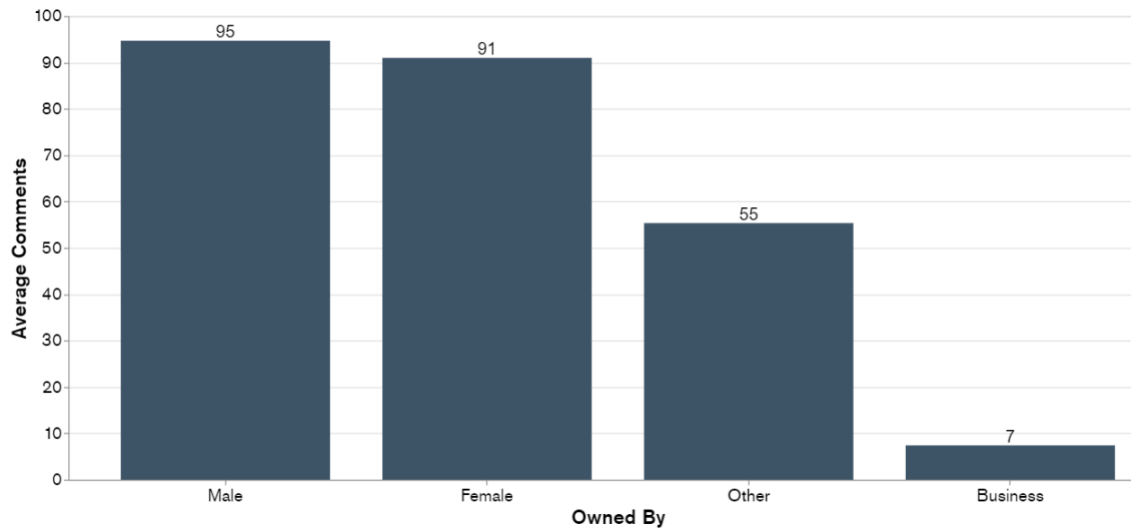


Σχήμα 5.89: Μέσος αριθμός θετικών και αρνητικών αντιδράσεων ανά κατηγορία χρήστη



Σχήμα 5.90: Σχέση μεταξύ θετικών και αρνητικών αντιδράσεων ανά κατηγορία χρήστη

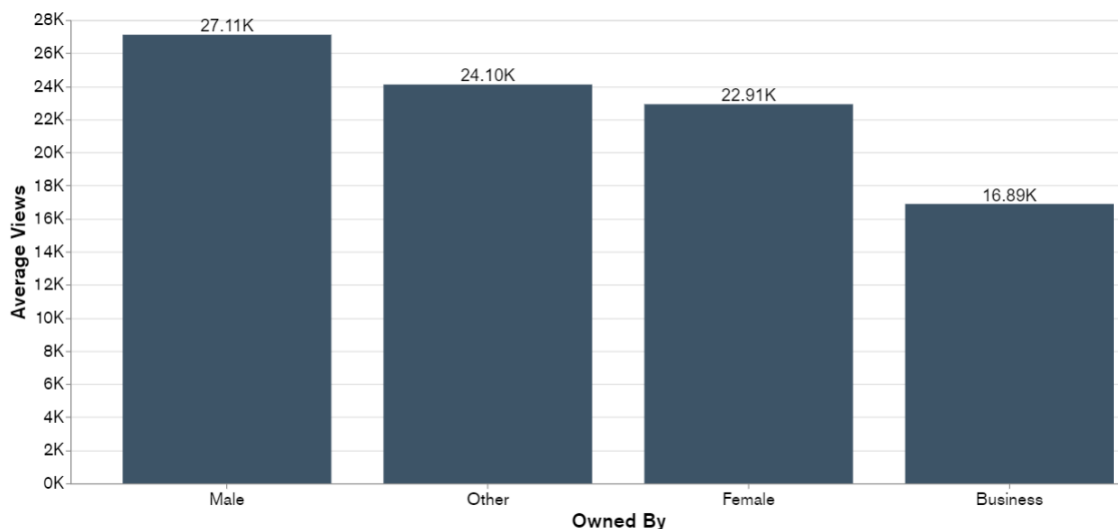
Ακολούθως, η σχέση μεταξύ των σχολίων ενός βίντεο και κατηγορίας χρήστη του δημιουργού του εμφανίζει αρκετά κοινά χαρακτηριστικά με εκείνη μεταξύ των θετικών αντιδράσεων και των διαφόρων κατηγοριών χρηστών (βλ. Σχήμα 5.89), καθώς όπως φαίνεται στο Σχήμα 5.91 οι άνδρες και οι γυναίκες βρίσκονται στις πρώτες θέσεις της κατάταξης, με μικρή διαφορά ανάμεσα τους, ενώ ακολουθούν τα κανάλια τύπου Άλλο και οι επιχειρησιακοί λογαριασμοί, οι οποίοι καταλαμβάνουν και πάλι τη θέση των λιγότερο λαμβανόμενων αντιδράσεων, απέχοντας κατά πολύ από τις υπόλοιπες κατηγορίες χρηστών.



Σχήμα 5.91: Μέσος αριθμός σχολίων ανά κατηγορία χρήστη

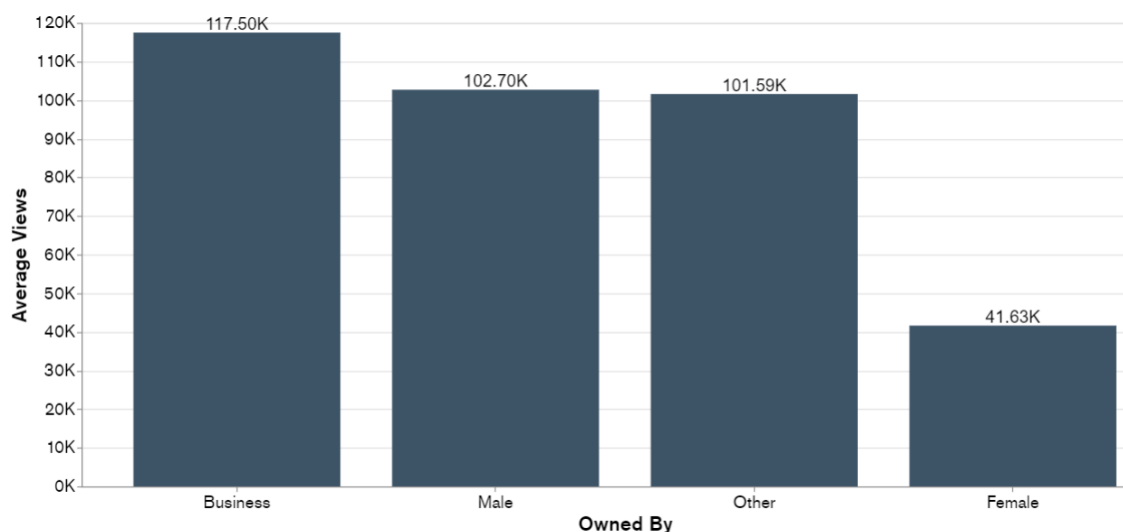
Επόμενος σταθμός της διαδικασίας ήταν η ανάλυση του μέσου αριθμού προβολών των βίντεο του δείγματος, με βάση τη κατηγορία χρήστη του δημιουργού τους. Εξαιτίας του γεγονότος ότι ένα μικρό υποσύνολο και συγκεκριμένα αυτών τα οποία ξεπερνούσαν τις 250.000 προβολές ήταν δυνατό να μεταβάλλει ριζικά τα τελικά αποτελέσματα (βλ. Σχήμα 5.71) αποφασίστηκε η ανάλυση να πραγματοποιηθεί σε δύο επιμέρους στάδια, εκείνο της μελέτης των βίντεο με μέγιστο αριθμό προβολών τις 250.000 και αυτό της μελέτης όλων των βίντεο του δείγματος.

Εστιάζοντας στα βίντεο των οποίων οι προβολές δεν ξεπέρασαν τις 250.000, από το Σχήμα 5.92 προκύπτει ότι κατά μέσο όρο τα βίντεο των ανδρών παρακολούθηθηκαν κατά το 2020 περισσότερο σε σχέση με αυτά των υπόλοιπων κατηγοριών, ενώ τα βίντεο των επιχειρήσεων κατατάσσονται στον χαμηλότερο βαθμό παρακολούθησης. Ένα πιθανό αίτιο στο οποίο θα μπορούσε να αποδοθεί το φαινόμενο αυτό είναι η σχέση μεταξύ του αριθμού εγγεγραμμένων χρηστών και του αριθμού προβολών ενός βίντεο. Αναλυτικότερα, όπως φάνηκε στο Σχήμα 5.87, κανάλια με μεγαλύτερο αριθμό εγγεγραμμένων χρηστών έχουν την τάση να συγκεντρώνουν υψηλότερο αριθμό προβολών στα βίντεο τους, κάτι το οποίο σε συνδυασμό με την κατάταξη των κατηγοριών χρηστών με βάση των μέσο αριθμό εγγεγραμμένων (βλ. Σχήμα 5.81), δικαιολογεί, εν μέρει, το αποτέλεσμα του ακόλουθου διαγράμματος. Παρόλα αυτά, ακόμη ένας λόγος στον οποίο θα μπορούσε να αποδοθεί το συγκεκριμένο φαινόμενο είναι η θεματολογία των βίντεο της εκάστοτε κατηγορίας χρήστη (βλ. Σχήμα 5.82).



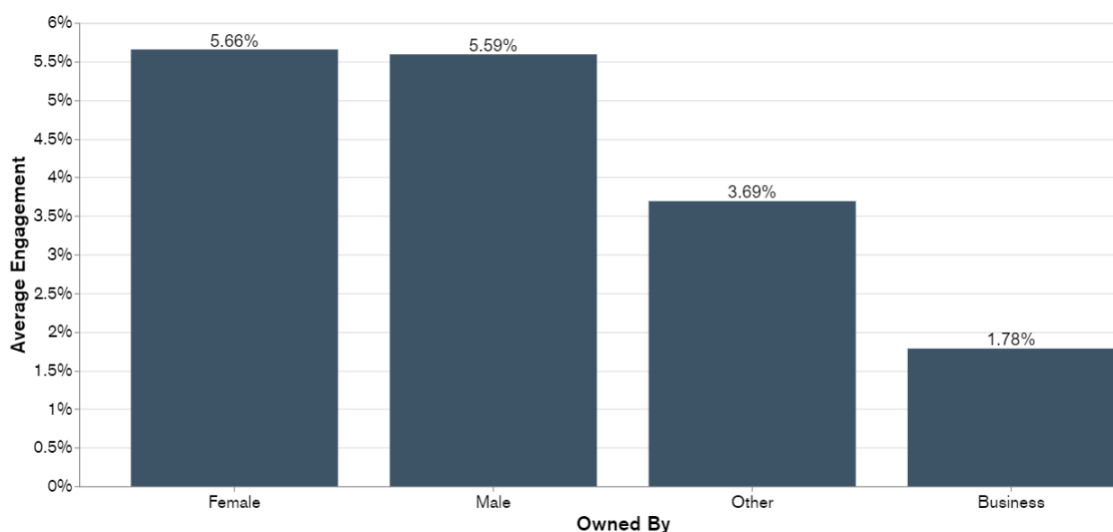
Σχήμα 5.92: Μέσος αριθμός προβολών ανά κατηγορία χρήστη (κάτω των 250.000 προβολών)

Όσον αφορά στη γενική εικόνα των βίντεο, είναι ορατό στο Σχήμα 5.93 πως αυτή διαφοροποιείται, κατά ένα μεγάλο βαθμό, από εκείνη του προηγούμενως εξεταζόμενου υποσυνόλου των βίντεο. Οι επιχειρησιακοί λογαριασμοί, παρά το περιορισμένο εγγεγραμμένο κοινό τους, φαίνεται πως κατά το έτος 2020 δημιούργησαν περιεχόμενο, το οποίο συγκέντρωσε αρκετά υψηλό αριθμό προβολών, ανεβάζοντας τους έτσι στη κορυφή της κατάταξης του μέσου αριθμού προβολών βίντεο για το ίδιο έτος. Η ριζική αυτή μεταβολή στην εικόνα των επιχειρήσεων, πιθανώς, οφείλεται στον συμψηφισμό περισσότερων καναλιών από το μικρό υποσύνολο των επιχειρήσεων, οι οποίες διατηρούν σχετικά ευρύ εγγεγραμμένο κοινό (βλ. Σχήμα 5.80). Από την άλλη μεριά, οι γυναίκες παρουσιάζουν τη μικρότερη βελτίωση του μέσου αριθμού προβολών τους, αναλογικά με τις υπόλοιπες κατηγορίες χρηστών, αποδεικνύοντας την ασθενέστερη παρουσία τους στα υψηλότερα στρώματα του αριθμού προβολών.



Σχήμα 5.93: Μέσος αριθμός προβολών ανά κατηγορία χρήστη (γενικά)

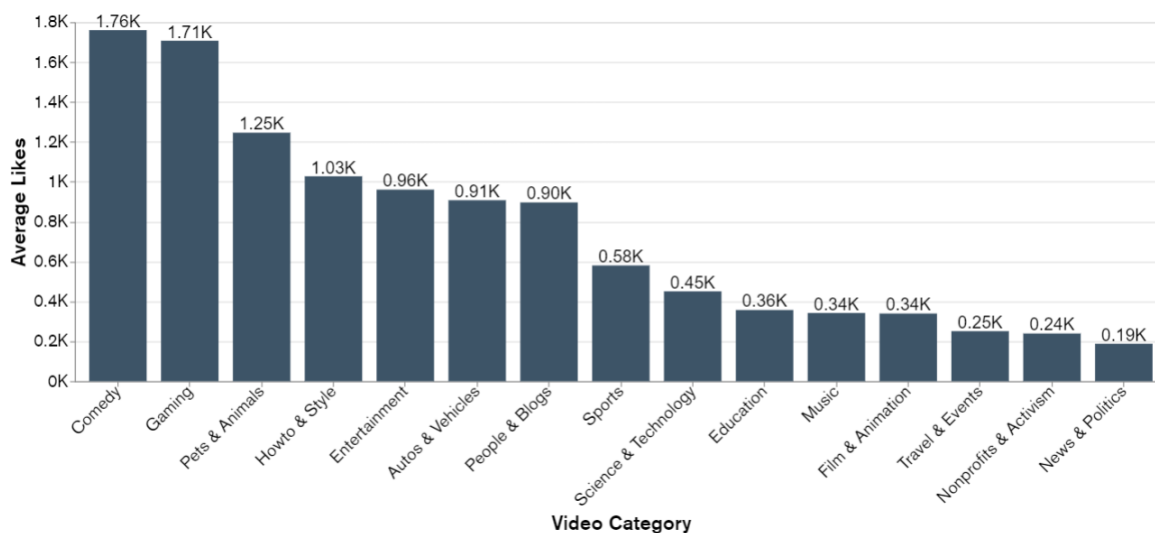
Όλα τα παραπάνω είναι δυνατό να συνοψιστούν στο μέσο ποσοστό αλληλεπίδρασης ανά κατηγορία χρήστη. Από το Σχήμα 5.94 εξάγεται το συμπέρασμα ότι οι προσωπικοί λογαριασμοί του YouTube, δηλαδή οι γυναίκες και οι άνδρες, κατάφεραν, κατά το έτος 2020, να διατηρήσουν το υψηλότερο το ποσοστό αλληλεπίδρασης με το κοινό τους, κάτι το οποίο μεταφράζεται και στον μεγαλύτερο βαθμό εξοικείωσης τον οποίον έχουν αναπτύξει με αυτό. Στον αντίποδα, τα επιχειρησιακά κανάλια παρέμειναν στη χαμηλότερη θέση της κατάταξης, υποδηλώνοντας την αδυναμία τους στην προσέλκυση της δυναμικής αλληλεπίδρασης του κοινού με το περιεχόμενό τους.



Σχήμα 5.94: Μέσο ποσοστό αλληλεπίδρασης ανά κατηγορία χρήστη

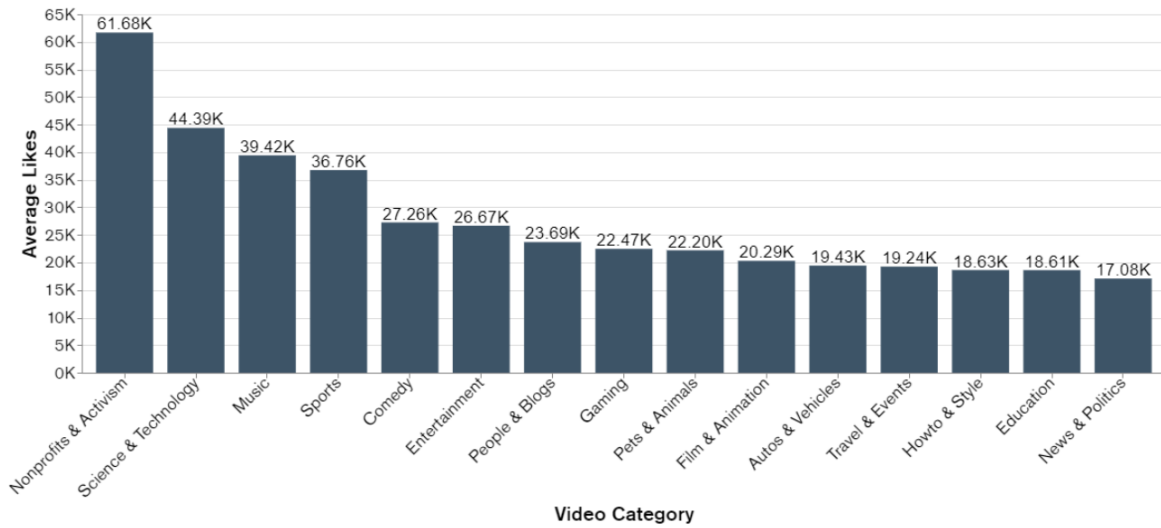
Έχοντας αναπτύξει τα παραπάνω, η ανάλυση των μετρικών απόδοσης των βίντεο του YouTube μεταφέρεται στον άξονα των θεματικών κατηγοριών βίντεο. Μέσω της μελέτης των αριθμητικών δεδομένων με κάθε μία από τις διαφορετικές θεματικές κατηγορίες περιεχομένου, είναι δυνατό να σχηματιστεί μια λεπτομερέστερη εικόνα σχετικά με ποιες από αυτές συγκέντρωσαν το μεγαλύτερο ενδιαφέρον του ελληνικού κοινού κατά τη διάρκεια του έτους 2020.

Αρχικά, στη βάση των θετικών αντιδράσεων διαπιστώνεται από το Σχήμα 5.95 ότι οι κατηγορίες *Κωμωδία (Comedy)* και *Ηλεκτρονικό Παιχνίδι (Gaming)* αναδείχθηκαν ως οι περισσότερο αρεστές κατηγορίες περιεχομένου του 2020, εισπράττοντας, κατά μέσο όρο, τον μεγαλύτερο αριθμό *likes*. Κοντά στη κορυφή της κατάταξης συναντήθηκαν, επίσης, οι κατηγορίες *Ζώα και Κατοικία (Pets & Animals)* και *Στιλιστικό περιεχόμενο (Howto & Style)*, ενώ στην αντίθετη πλευρά βρέθηκαν οι κατηγορίες *Νέα και Πολιτική (News & Politics)*, *Ακτιβισμός (Nonprofits & Activism)* και *Ταξιδιωτικές Δραστηριότητες (Travel & Events)*. Επισημαίνεται πως προς αποφυγή των ακραίων περιπτώσεων του δείγματος, στα πλαίσια του ακόλουθου διαγράμματος λήφθηκαν υπόψη μόνο βίντεο τα οποία είχαν λιγότερες από 15.000 θετικές αντιδράσεις (βλ. Σχήμα 5.73).



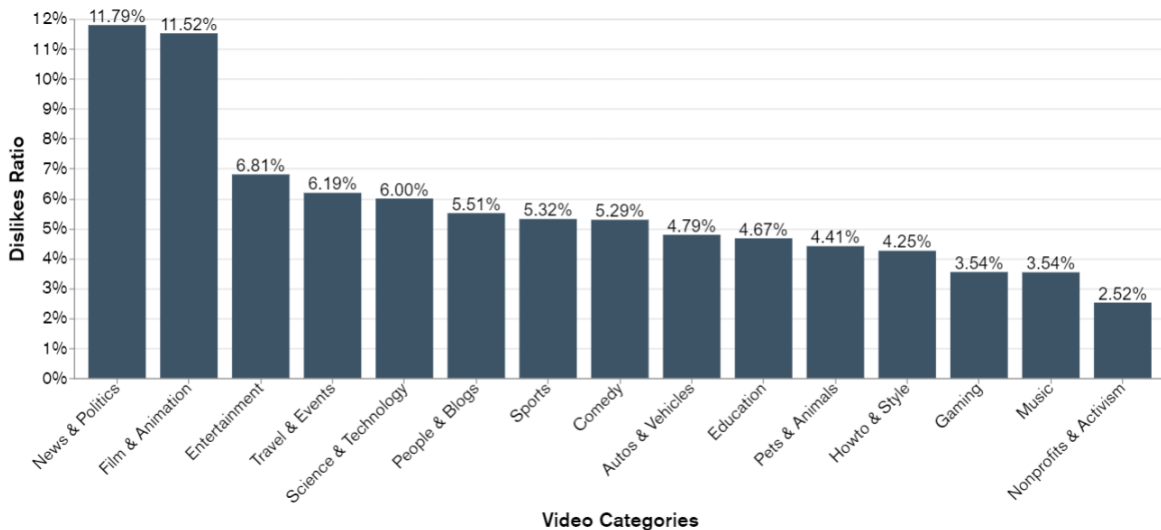
Σχήμα 5.95: Μέσος αριθμός θετικών αντιδράσεων ανά θεματική κατηγορία βίντεο (κάτω των 15.000 θετικών αντιδράσεων)

Με σκοπό την απόκτηση μιας πληρέστερης εικόνας για την κατανομή των θετικών αντιδράσεων στις θεματικές κατηγορίες περιεχομένου, κρίθηκε σημαντική η εξέταση της μειοψηφίας των βίντεο του δείγματος, τα οποία ξεπέρασαν τις 15.000 χιλιάδες θετικές αντιδράσεις. Όπως παρατηρείται στο Σχήμα 5.96, αρκετές ήταν οι θεματικές κατηγορίες οι οποίες, παρά τον χαμηλό μέσο αριθμό θετικών αντιδράσεων της πλειοψηφίας των βίντεο τους, παρήγαγαν μια μικρή ποσότητα υλικού ικανή να αποσπάσει τα μεγαλύτερα ποσοστά θετικής ανταπόκρισης του ελληνικού κοινού. Ο λόγος γίνεται κυρίως για τις κατηγορίες *Ακτιβισμός (Nonprofits & Activism)*, *Επιστήμες και Τεχνολογία (Science & Technology)*, *Μουσική (Music)* και *Αθλητισμός (Sports)*. Η μεγάλη διασπορά του αριθμού θετικών αντιδράσεων, η οποία παρατηρείται στις συγκεκριμένες κατηγορίες πιθανώς να αποτελεί ένδειξη των προοπτικών εξέλιξης τους στα πλαίσια της ελληνικής κοινότητας της πλατφόρμας.



Σχήμα 5.96: Μέσος αριθμός θετικών αντιδράσεων ανά θεματική κατηγορία βίντεο (άνω των 15.000 θετικών αντιδράσεων)

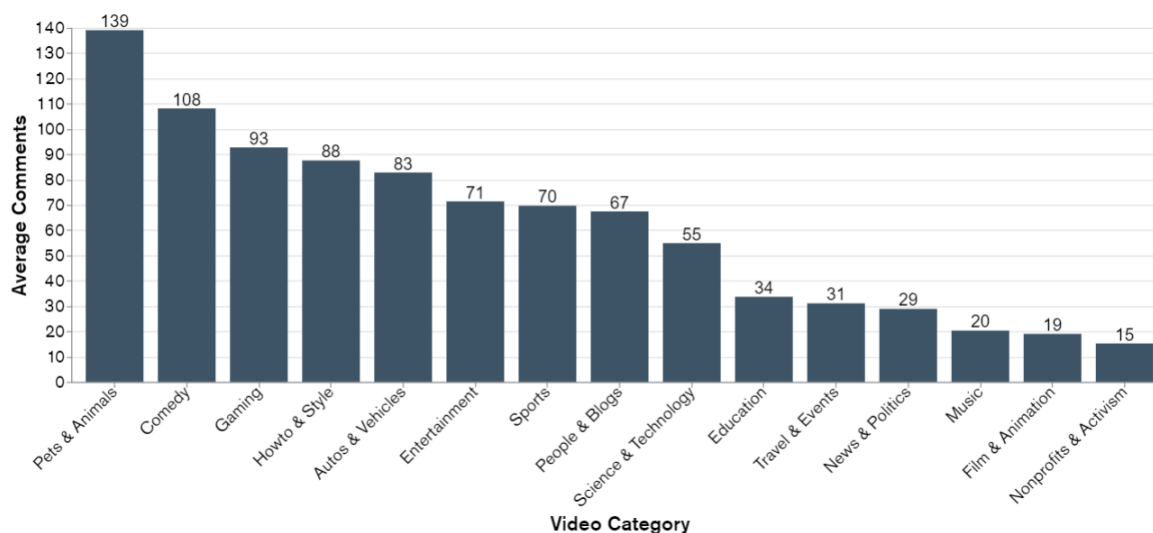
Πέραν της αποτύπωσης της θετικής ανταπόκρισης του κοινού απέναντι στις ποικίλες θεματικές κατηγορίες, το ενδιαφέρον κέντρισε η διερεύνηση του ποσοστού αρνητικών αντιδράσεων, το οποίο καθόρισε την κάθε μία από αυτές. Σύμφωνα με το Σχήμα 5.95, οι θεματικές κατηγορίες *Νέα και Πολιτική* και *Κινηματογράφος και Κινούμενα Σχέδια* αποτέλεσαν τις λιγότερο συμπαθείς κατηγορίες περιεχομένου για το έτος 2020, καθώς το ποσοστό των αρνητικών αντιδράσεων τους αναρριχήθηκε κοντά στο 12%. Ειδικότερα, το συγκεκριμένο ποσοστό μεταφράζεται ως εξής: έστω ότι κάποιο βίντεο μιας από τις δύο παραπάνω κατηγορίες συγκέντρωσε άθροισμα θετικών και αρνητικών αντιδράσεων ίσο με 100· τότε προσεγγιστικά οι 12 από αυτές θα είναι *dislikes*.



Σχήμα 5.97: Σχέση μεταξύ θετικών και αρνητικών αντιδράσεων ανά θεματική κατηγορία βίντεο

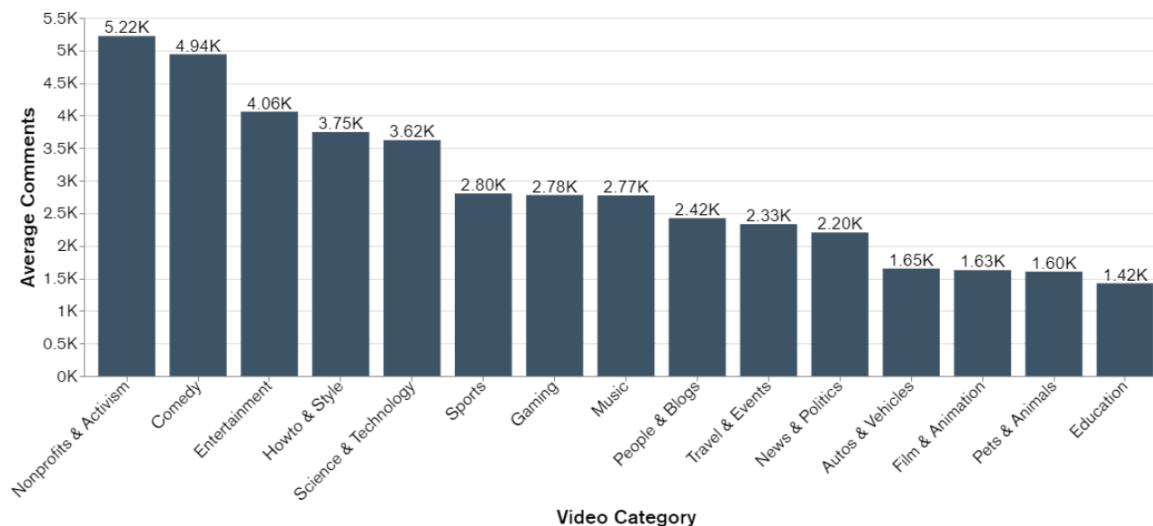
Εφόσον εξετάστηκαν οι θετικές και οι αρνητικές αντιδράσεις, σειρά έχει η μελέτη του μέσου αριθμού σχολίων, ο οποίος συγκεντρώθηκε από την εκάστοτε θεματική κατηγορία περιεχομένου, με βάση τα δημοσιευμένα εντός του 2020 βίντεο του δείγματος. Η ανάγκη για μια αντιπροσωπευτική αποτύπωση της κατανομής του αριθμού σχολίων στις διάφορες θεματικές κατηγορίες περιεχομένου οδήγησε στη διάσπαση της διαδικασίας της ανάλυσης σε δύο επιμέρους στάδια. Με γνώμονα το Σχήμα 5.75 διακρίθηκαν δύο υποσύνολα βίντεο, αυτά τα οποία συγκέντρωσαν λιγότερα από 1.000 σχόλια και εκείνα τα οποία ξεπέρασαν το συγκεκριμένο όριο κατά το 2020.

Εστιάζοντας στη πλειοψηφία των βίντεο του δείγματος (βλ. Σχήμα 5.98), διαφαίνεται πως οι θεματικές κατηγορίες περιεχομένου οι οποίες παρείχαν το γονιμότερο έδαφος για συζήτηση και ανταλλαγή απόψεων, κατά τη διάρκεια του 2020, ήταν οι εξής: (i) *Ζώα και Κατοικίδια (Pets & Animals)*, (ii) *Κωμωδία (Comedy)*, (iii) *Ηλεκτρονικό Παιχνίδι (Gaming)*, (iv) *Στιλιστικό περιεχόμενο (Howto & Style)* και (v) *Αυτοκίνηση (Autos & Vehicles)*.



Σχήμα 5.98: Μέσος αριθμός σχολίων ανά θεματική κατηγορία βίντεο (κάτω των 1.000 σχολίων)

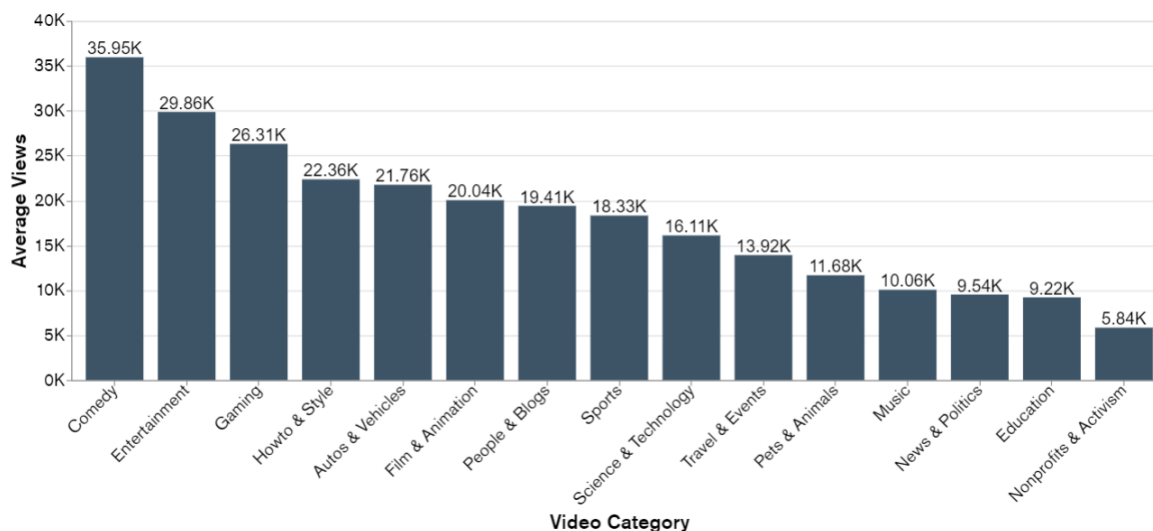
Όσον αφορά στη μικρή μερίδα βίντεο τα οποία ξεπέρασαν τα 1.000 σχόλια, φαίνεται από το Σχήμα 5.99 πως οι κορυφαίες θεματικές κατηγορίες, με βάση τον μέσο αριθμό σχολιασμών, δεν μεταβλήθηκαν ολοκληρωτικά σε σχέση με την εικόνα της πλειοψηφίας, καθώς οι κατηγορίες *Κωμωδία (Comedy)* και *Στιλιστικό περιεχόμενο (Howto & Style)* διατήρησαν την υψηλή τους θέση στη κατάταξη. Ωστόσο, την εμφάνισή τους στα ανώτερα στρώματα σχολιασμού έκαναν και θεματικές κατηγορίες περιεχομένου οι οποίες συναντήθηκαν σε χαμηλότερες θέσεις του Σχήματος 5.98, όπως ο *Ακτιβισμός (Nonprofits & Activism)*, η *Ψυχαγωγία (Entertainment)* και οι *Επιστήμες και Τεχνολογία (Science & Technology)*. Το φαινόμενο αυτό είναι πιθανό να οφείλεται, εκτός από την ύπαρξη μεμονωμένων περιπτώσεων βίντεο με υψηλό αριθμό σχολιασμών, στην παρουσία βίντεο/διαγωνισμών, τα οποία βασίζονται στην ανταπόκριση του κοινού μέσω σχολιασμών και θα αναλυθούν εκτενέστερα σε μεταγενέστερο στάδιο της ερευνητικής διαδικασίας.



Σχήμα 5.99: Μέσος αριθμός σχολίων ανά θεματική κατηγορία βίντεο (άνω των 1.000 σχολίων)

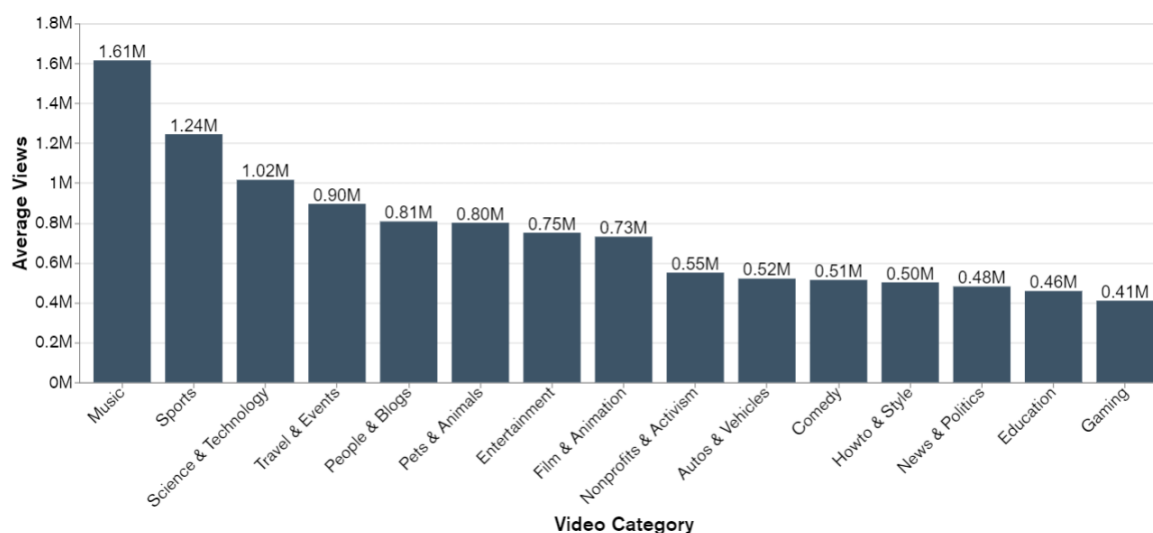
Στο επίπεδο του αριθμού προβολών των βίντεο, για λόγους αντίστοιχους με εκείνους των παραπάνω διαγραμμάτων, η μελέτη της σχέσης μεταξύ των προβολών ενός βίντεο και της θεματικής κατηγορίας, στην οποία αυτό ανήκει, πραγματοποιήθηκε ξεχωριστά για τα βίντεο του δείγματος με προβολές λιγότερες από 250.000 και για αυτά τα οποία τις ξεπέρασαν. Το όριο των 250.000 προβολών προέκυψε λαμβάνοντας υπόψη τα δεδομένα του Σχήματος 5.71.

Ξεκινώντας με την πλειονότητα των βίντεο, δηλαδή εκείνα τα οποία δεν κατάφεραν να ξεπεράσουν τις 250.000 προβολές κατά τη διάρκεια του έτους 2020, είναι ορατό στο Σχήμα 5.100 ότι οι θεματικές κατηγορίες περιεχομένου οι οποίες παρακολούθηθηκαν περισσότερο, κατά μέσο όρο, ήταν η *Κωμωδία (Comedy)*, η *Ψυχαγωγία (Entertainment)* και το *Ηλεκτρονικό Παιχνίδι (Gaming)*, ενώ σε υψηλές θέσεις της κατάταξης βρέθηκαν, επίσης, το *Στιλιστικό περιεχόμενο (Howto & Style)* και η *Αυτοκίνηση (Autos & Vehicles)*.



Σχήμα 5.100: Μέσος αριθμός προβολών ανά θεματική κατηγορία βίντεο (κάτω των 250.000 προβολών)

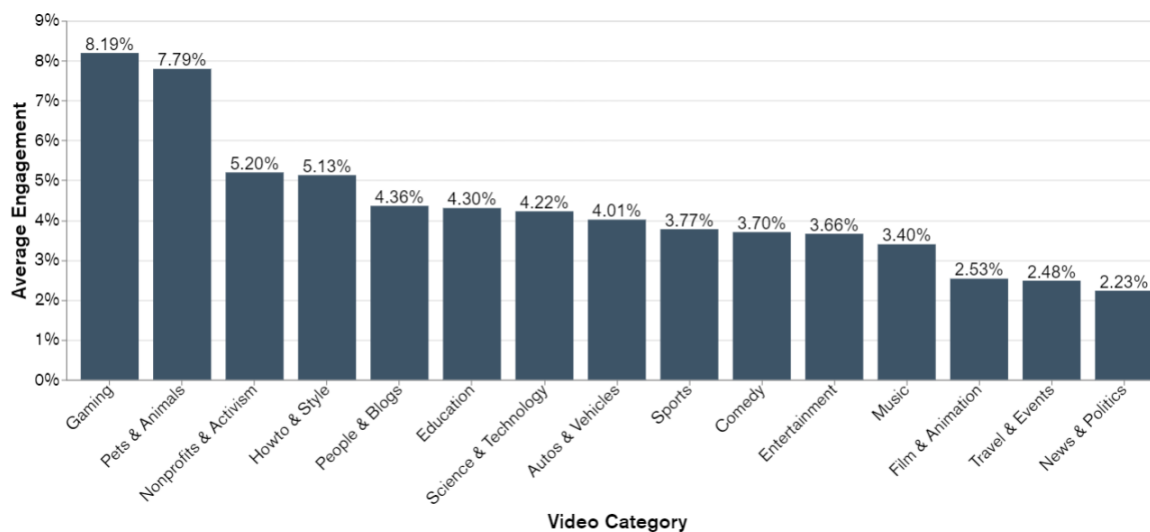
Αναφορικά με τη μικρότερη μερίδα των βίντεο, τα οποία απέσπασαν περισσότερες από 250.000 προβολές εντός του 2020, η κορυφή της κατάταξης κατακλύστηκε από θεματικές κατηγορίες περιεχομένου, οι οποίες διατηρούσαν χαμηλότερες θέσεις στην κατάταξη του Σχήματος 5.99, αναδεικνύοντας έτσι της προοπτικές εξέλιξης των κατηγοριών στα πλαίσια της ελληνικής κοινότητας του YouTube. Αναλυτικότερα, οι κατηγορίες οι οποίες διακρίθηκαν στα υψηλότερα στρώματα προβολών ήταν η *Μουσική (Music)*, ο *Αθλητισμός (Sports)*, οι *Επιστήμες και Τεχνολογία (Science & Technology)* και οι *Ταξιδιωτικές Δραστηριότητες (Travel & Events)*.



Σχήμα 5.101: Μέσος αριθμός προβολών ανά θεματική κατηγορία βίντεο (άνω των 250.000 προβολών)

Όλες οι παραπάνω μετρικές απόδοσης του YouTube είναι δυνατό να συνοψιστούν στο ποσοστό αλληλεπίδρασης μεταξύ του κοινού και των βίντεο του δείγματος. Ο υπολογισμός του μέσου ποσοστού αλληλεπίδρασης για κάθε μία από τις θεματικές κατηγορίες περιεχομένου μπορεί να αναδείξει αυτές, οι οποίες λαμβάνουν την εντονότερη ανταπόκριση του κοινού, είτε αυτή πρόκειται για θετική αντίδραση ή για σχολιασμό. Κατά τη διαδικασία μελέτης της παρούσας μεταβλητής, με σκοπό την αποφυγή των ακραίων περιπτώσεων, αποφασίστηκε να ληφθούν υπόψη μόνο τα βίντεο τα οποία δεν ξεπέρασαν το 20% ποσοστού αλληλεπίδρασης κατά το έτος 2020. Στον καθορισμό αυτής της ανώτατης τιμής συντέλεσαν τα δεδομένα του Σχήματος 5.77.

Σύμφωνα με το Σχήμα 5.102, οι θεματικές κατηγορίες περιεχομένου, οι οποίες συγκέντρωσαν το μεγαλύτερο ποσοστό αντιδράσεων αναλογικά με το πλήθος των χρηστών που τις παρακολούθησαν ήταν το Ηλεκτρονικό Παιχνίδι (*Gaming*), τα Ζώα και Κατοικίδια (*Pets & Animals*), ο Ακτιβισμός (*Nonprofits & Activism*) και το Στυλιστικό περιεχόμενο (*Howto & Style*). Συγκεκριμένα, παρατηρείται πως οι δύο πρώτες κατηγορίες σημείωσαν αρκετά υψηλότερα ποσοστά από τις υπόλοιπες, γεγονός το οποίο αποδεικνύει τη θερμή ανταπόκριση του κοινού προς αυτές. Στους πρόποδες της κατάταξης εντοπίζονται οι κατηγορίες *Κινηματογράφος και Κινούμενα Σχέδια* (*Film & Animation*), *Ταξιδιωτικές Δραστηριότητες* (*Travel & Events*) και *Νέα και Πολιτική* (*News & Politics*), αποτελώντας έτσι τις κατηγορίες με τον μικρότερο βαθμό διάδρασης μεταξύ αυτών και του κοινού για το έτος 2020.



Σχήμα 5.102: Μέσο ποσοστό αλληλεπίδρασης ανά θεματική κατηγορία βίντεο

Εν κατακλείδι, με βάση τη πλειοψηφία των βίντεο τα οποία δημοσιεύτηκαν κατά τη διάρκεια του έτους 2020, οι θεματικές κατηγορίες περιεχομένου οι οποίες προσέλκυσαν περισσότερο το ενδιαφέρον του κοινού ήταν η *Κωμωδία* (*Comedy*), το *Ηλεκτρονικό Παιχνίδι* (*Gaming*), το *Στυλιστικό περιεχόμενο* (*Howto & Style*) και η *Ψυχαγωγία* (*Entertainment*). Παρόλα αυτά, στις ακραίες περιπτώσεις θετικών αντιδράσεων, σχολιασμών και προβολών συναντήθηκαν αρκετά οι κατηγορίες *Μουσική* (*Music*), *Αθλητισμός* (*Sports*) και *Επιστήμες και Τεχνολογία* (*Science & Technology*), δείχνοντας πως οι συγκεκριμένες κατηγορίες επίσης αποτέλεσαν πόλους έλξης του ενδιαφέροντος του κοινού και πως δίνεται η δυνατότητα στους δημιουργούς περιεχομένου να ευδοκιμή-

σουν στους αντίστοιχους κλάδους της πλατφόρμας. Επιπλέον, διαπιστώθηκε πως το ελληνικό κοινό ανέπτυξε μία ιδιαίτερα στενή σχέση αλληλεπίδρασης με βίντεο *Ηλεκτρονικού Παιχνιδιού (Gaming)* και *Ζώων και Κατοικίδιων (Pets & Animals)*, φανερώνοντας τη συμπάθεια του προς αυτές. Τέλος, κατηγορίες όπως τα *Νέα και Πολιτική (News & Politics)*, το *Εκπαιδευτικό περιεχόμενο (Education)* και ο *Ακτιβισμός (Nonprofits & Activism)* συγκέντρωσαν, κατά μέσο όρο, το μικρότερο ποσοστό ενδιαφέροντος του κοινού, με την πρώτη να σημειώνει αρνητικό ρεκόρ ως προς τις λαμβανόμενες αρνητικές αντιδράσεις σε βίντεο.

Πέραν της παραπάνω διάκρισης των βίντεο, η οποία πραγματοποιήθηκε στον άξονα των θεματικών κατηγοριών περιεχομένου, τα βίντεο του δείγματος είναι επίσης δυνατό να κατηγοριοποιηθούν με βάση τη δομή του περιεχομένου τους. Πιο συγκεκριμένα, οι κατηγορίες οι οποίες ορίστηκαν είναι οι εξής:

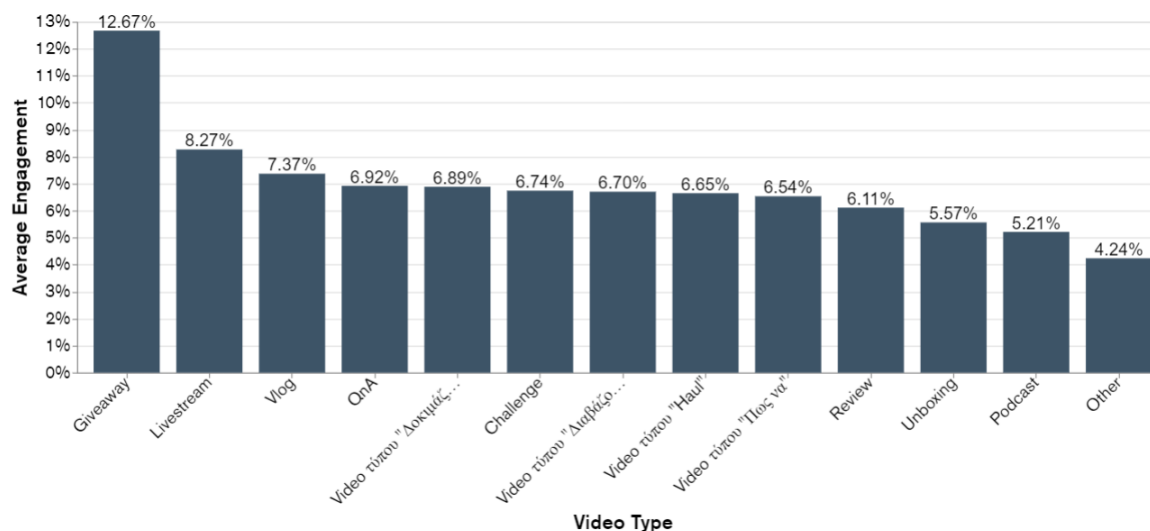
- **Διαγωνισμός (Giveaway):** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού διοργανώνει έναν διαγωνισμό, καλώντας τους χρήστες του YouTube να αντιδράσουν στο συγκεκριμένο βίντεο, ώστε να διεκδικήσουν το διαφημιζόμενο προϊόν
- **Ζωντανή Ροή (Livestream):** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού μεταδίδει ζωντανό περιεχόμενο, αλληλεπιδρώντας άμεσα με το κοινό του
- **Βλογκ (Vlog):** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού παρουσιάζει, λιγότερο επεξεργασμένες, στιγμές από τη προσωπική του καθημερινότητα
- **Ερωτήσεις και Απαντήσεις (QnA):** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού ζητάει από το κοινό να διατυπώσει ερωτήσεις, τις οποίες ο ίδιος καλείται να απαντήσει
- **Βίντεο τύπου “Δοκιμάζουμε”:** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού δοκιμάζει προϊόντα, όπως τρόφιμα, καλλυντικά ή είδη τεχνολογίας
- **Πρόκληση (Challenge):** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού ζητάει από το κοινό να θέσει κάποιες προκλήσεις, με σκοπό εκείνος να τις πραγματοποιήσει
- **Βίντεο τύπου “Διαβάζουμε”:** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού ζητάει από το κοινό να παραθέσει προσωπικές ιστορίες ή απόψεις, τις οποίες εκείνος θα σχολιάσει
- **Βίντεο τύπου “Haul”:** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού παρουσιάζει στο κοινό του προϊόντα, όπως είδη ρουχισμού και είδη οικιακής χρήσης, τα οποία ο ίδιος αγόρασε νωρίτερα
- **Βίντεο τύπου “Πως να” (How to):** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού παρουσιάζει στο κοινό του βήματα και συμβουλές, με σκοπό την εκπλήρωση ενός στόχου
- **Αξιολόγηση (Review):** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού αξιολογεί κάποιο προϊόν, όπως προϊόντα αισθητικής, μία ταινία ή ένα προϊόν τεχνολογίας, το οποίο πιθανώς έχει λάβει από κάποια εταιρεία για τον σκοπό αυτό
- **Βίντεο τύπου “Unboxing”:** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του

καναλιού λαμβάνει από κάποια εταιρεία ένα προϊόν, το οποίο καλείται να το βγάλει από την συσκευασία και να το παρουσιάσει στο κοινό του

- **Διαδικτυακή Συζήτηση (Podcast):** τύπος βίντεο στα πλαίσια του οποίου ο διαχειριστής του καναλιού αναπτύσσει συζήτηση με κάποιον καλεσμένο, ο οποίος ενδεχομένως είναι δημοφιλής
- **Άλλο (Other):** κατηγορία η οποία περιλαμβάνει όλες τα βίντεο που δεν κατατάσσονται σε κάποιον από τους παραπάνω τύπους βίντεο

Η άμεση αλληλεπίδραση του κοινού, η δημοτικότητα κατά το 2020 και το εμπορικό ενδιαφέρον, λόγω της παρουσίας επιχειρήσεων, αποτέλεσαν τους κύριους λόγους για τους οποίους οι παραπάνω τύποι βίντεο επιλέχθηκαν προς βαθύτερη ανάλυση. Είναι σημαντικό να επισημανθεί πως οι παραπάνω τύποι βίντεο εντοπίστηκαν ανάμεσα στη πληθώρα μεταφορτώσεων του δείγματος, μέσω της αναζήτησης λέξεων-κλειδιών στους τίτλους τους, όπως *Giveaway*, *Vlog* ή *QnA*.

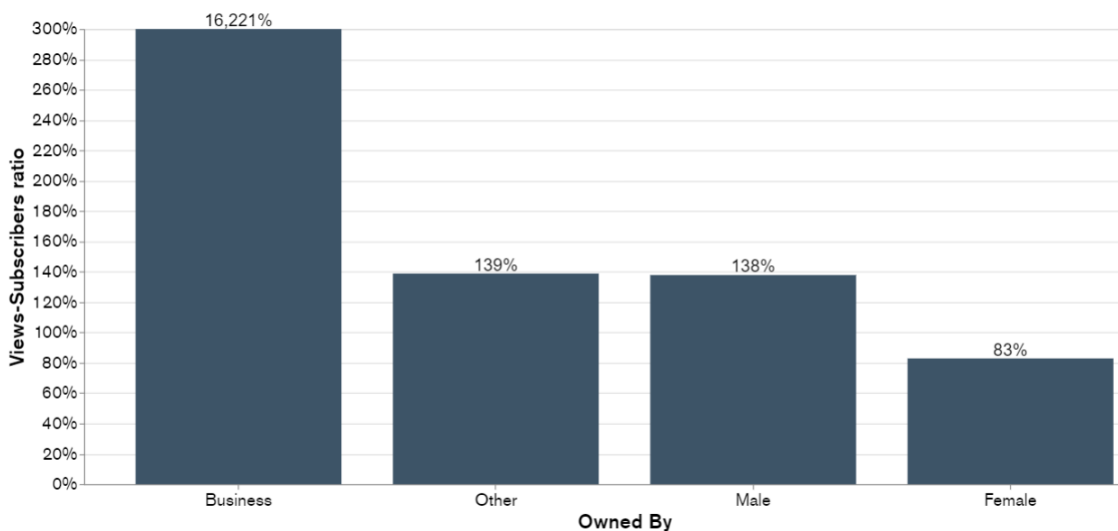
Ως πρώτο βήμα, στο Σχήμα 5.103, μελετήθηκε το μέσο ποσοστό αλληλεπίδρασης με κάθε έναν τους επιλεγμένους τύπους βίντεο. Παρατηρώντας τα δεδομένα του προκείμενου σχήματος, πράγματι, αποδεικνύεται πως όλοι οι επιλεγμένοι τύποι βίντεο απέσπασαν υψηλότερα ποσοστά αλληλεπίδρασης, συγκριτικά με κάθε άλλο τύπο βίντεο κατά τη διάρκεια του 2020. Τα βίντεο τύπου *Vlog* και οι *Ζωντανές Ροές* φαίνεται πως κατάφεραν να προσελκύσουν, σε ελαφρώς μεγαλύτερο βαθμό, τις αντιδράσεις του κοινού, ενώ τα βίντεο τα οποία αφορούσαν σε *Διαγωνισμούς (Giveaway)* στάθηκαν στη κορυφή της κατάταξης με εμφανή διαφορά στο ποσοστό αλληλεπίδρασης από τους υπόλοιπους τύπους βίντεο, εξακριβώνοντας πως τέτοιου είδους βίντεο αποτελούν την πλέον αποδοτικότερη τακτική των επιχειρήσεων για τη προώθηση των προϊόντων τους στο YouTube. Οι παρατηρήσεις αυτές, σε συνδυασμό με το αρκετά χαμηλό ποσοστό αλληλεπίδρασης που σημειώθηκε από τα βίντεο των επιχειρήσεων κατά το έτος 2020 (βλ. Σχήμα 5.94), επαναφέρουν στο προσκήνιο την έννοια του *Influencer Marketing*, κατά την οποία οι επιχειρήσεις επιλέγουν να διαφημίσουν τα προϊόντα ή τις υπηρεσίες τους σε ένα ευρύτερο κοινό, μέσω δημοφιλών δημιουργών περιεχομένου.



Σχήμα 5.103: Μέσο ποσοστό αλληλεπίδρασης ανά τύπο βίντεο

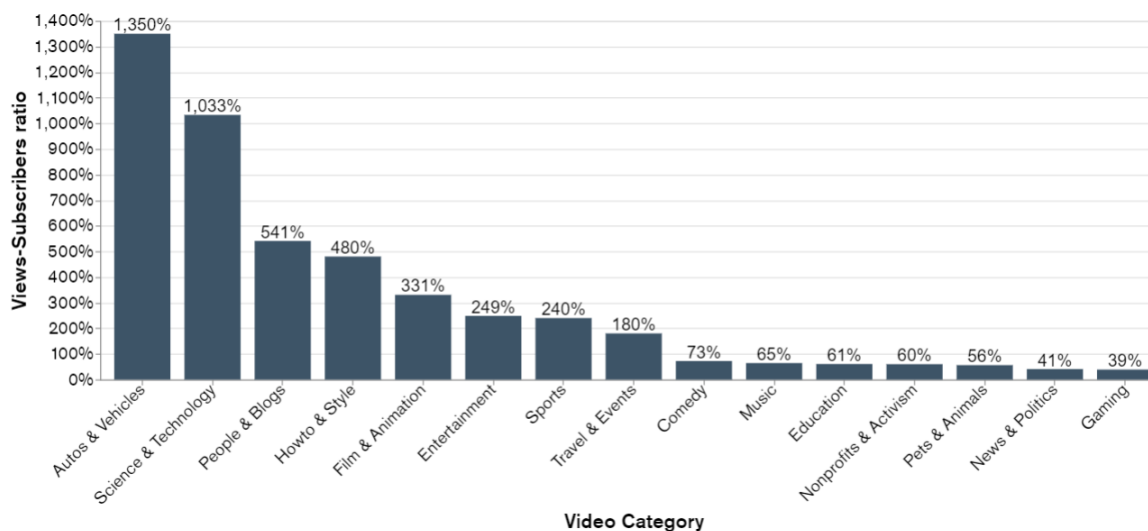
Μια εναλλακτική εικόνα της ανταπόκρισης του ελληνικού κοινού απέναντι στις διάφορες κατηγορίες χρηστών, κατηγορίες θεματικού περιεχομένου και τύπους βίντεο είναι δυνατό να παρέχει η μετρική *Views-Subscribers Ratio*. Η συγκεκριμένη μεταβλητή είναι υπεύθυνη για την αποτύπωση της σχέσης μεταξύ του μέσου αριθμού προβολών ενός καναλιού και του αριθμού των εγγεγραμμένων χρηστών του, διαιρώντας τις προβολές του εκάστοτε βίντεο με το πλήθος των εγγεγραμμένων του. Το πεδίο τιμών της μεταβλητής αυτή δεν μπορεί να οριστεί αυστηρά, καθώς δύναται είτε να ξεπεράσει κατά πολύ το 100%, σε περιπτώσεις όπου οι προβολές ενός βίντεο ξεπερνούν τον αριθμό εγγεγραμμένων χρηστών, ή να παραμείνει κοντά στο 0%, σε περιπτώσεις βίντεο με υπερβολικά χαμηλό αριθμό προβολών.

Ο υπολογισμός της μεταβλητής *Views-Subscribers Ratio*, αρχικά, επικεντρώθηκε στις κατηγορίες χρηστών του YouTube *Άνδρας*, *Γυναίκα*, *Επιχείρηση* και *Άλλο*. Σύμφωνα με τα δεδομένα του Σχήματος 5.104, διαπιστώνεται πως τα γυναικεία κανάλια σημείωσαν το μικρότερο ποσοστό μεταξύ προβολών και εγγεγραμμένων χρηστών κατά το 2020, το οποίο παρέμεινε κάτω του 100%, δείχνοντας έτσι πως το περιεχόμενο των γυναικών δημιουργών απευθύνεται σε ένα πιο στοχευμένο κοινό. Ανεβαίνοντας στη κατάταξη, συναντούνται οι άνδρες και οι λογαριασμοί τύπου *Άλλο*, οι οποίοι παρέμειναν στα ίδια επίπεδα ποσοστού *Views-Subscribers Ratio*, λίγο πάνω από το 100%, φανερώνοντας κατά αυτόν τον τρόπο την προσέλκυση προβολών από χρήστες εκτός του εύρους του πιστού τους κοινού. Φυσικά, δε θα μπορούσε να μείνει απαρατήρητο το υπέρογκο ποσοστό προβολών-εγγεγραμμένων χρηστών των επιχειρήσεων, το οποίο ξεπέρασε το 16.000%. Ο τεράστιος αυτός αριθμός σε συνδυασμό με τον χαμηλό, κατά μέσο όρο, αριθμό εγγεγραμμένων χρηστών των επιχειρήσεων (βλ. Σχήμα 5.81), ενισχύει το συμπέρασμα πως αυτές προσπαθούν να διαδώσουν ευρέως το περιεχόμενό τους, ακολουθώντας, πιθανώς, *επιθετικές* τακτικές χωρίς, ωστόσο, να καταφέρουν να αναπτύξουν μία αφοσιωμένη κοινότητα. Συμπληρώνοντας τα παραπάνω, ο αρνητικός αντίκτυπος της δραστηριότητας των επιχειρησιακών καναλιών στο ελληνικό κοινό του YouTube, μπορούσε να συναχθεί και από την αναλογικά υψηλή συγκέντρωση αρνητικών αντιδράσεων στα βίντεο τους, η οποία αναπαραστάθηκε στο Σχήμα 5.90.



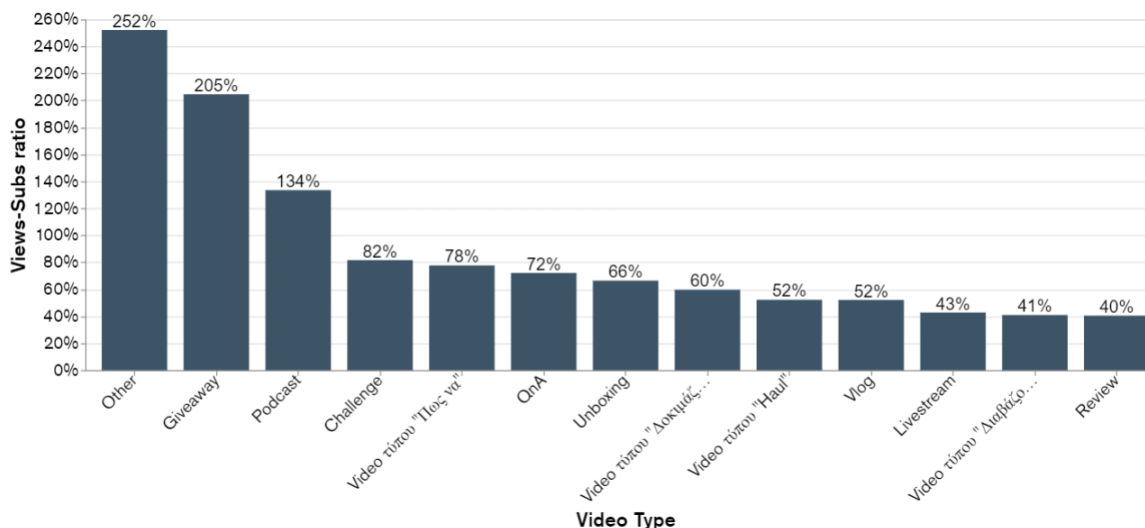
Σχήμα 5.104: Σχέση μεταξύ εγγεγραμμένων χρηστών και προβολών ανά κατηγορία χρηστών

Συνεχίζοντας στις θεματικές κατηγορίες περιεχομένου, με μία πρώτη ματιά στο Σχήμα 5.105, φαίνεται πως στις υψηλότερες θέσεις της κατάταξης, και με ποσοστό προβολών-εγγεγραμμένων χρηστών μεγαλύτερο του 100%, βρέθηκαν οι θεματικές κατηγορίες οι οποίες αφορούσαν στην *Αυτοκίνηση (Autos & Vehicles)*, στις *Επιστήμες και Τεχνολογία (Science & Technology)*, στις *Καθημερινότητα και Ανθρώπινες Σχέσεις (People & Blogs)* και στο *Στιλιστικό περιεχόμενο (Howto & Style)*, υποδηλώνοντας πως σε αυτούς τους κλάδους του YouTube δίνεται το περιθώριο στους δημιουργούς να διευρύνουν το σύνολο των πιστών ακολούθων τους. Ωστόσο, είναι πιθανό τα στατιστικά αυτά να είναι αρκετά επηρεασμένα από την παρουσία επιχειρησιακών λογαριασμών, όπως παρουσιάστηκε στο Σχήμα 5.104. Στην αντίπερα όχθη, τα βίντεο *Ηλεκτρονικού Παιχνιδιού (Gaming)* συγκέντρωσαν κατά το έτος 2020 το χαμηλότερο ποσοστό *Views-Subscribers Ratio*, γεγονός το οποίο σε συνδυασμό με τον αρκετά υψηλό μέσο αριθμό προβολών, τον οποίο αυτά απέσπασαν κατά το ίδιο χρονικό διάστημα (βλ. Σχήμα 5.100) φανερώνει την ύπαρξη ενός ευρέως εδραιωμένου κοινού, το οποίο παρακολουθεί τη συγκεκριμένη θεματική κατηγορία περιεχομένου.



Σχήμα 5.105: Σχέση μεταξύ εγγεγραμμένων χρηστών και προβολών ανά θεματική κατηγορία βίντεο

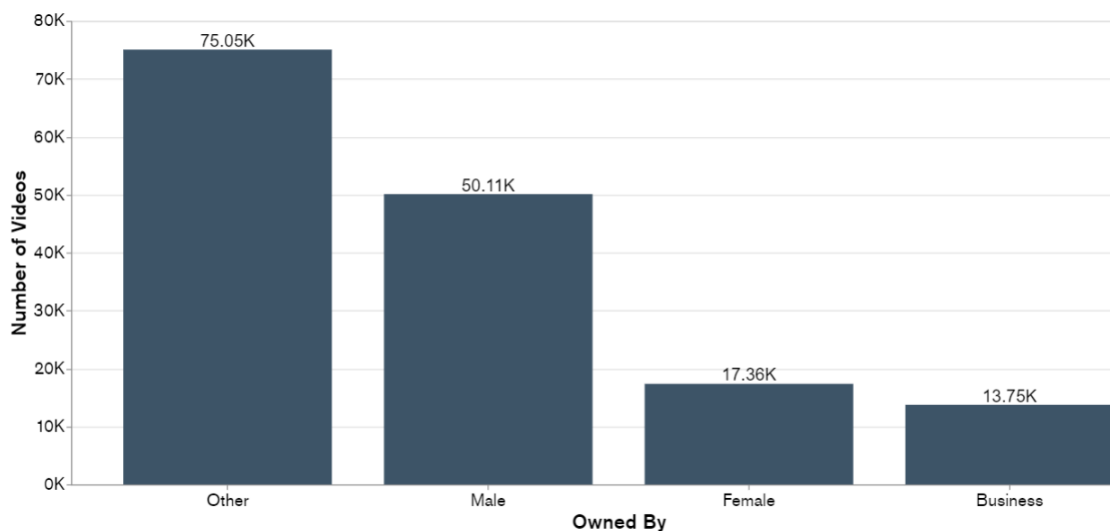
Η μελέτη της σχέσης μεταξύ των προβολών ενός βίντεο και των εγγεγραμμένων χρηστών του καναλιού από το οποίο αυτό δημοσιεύτηκε ολοκληρώθηκε με τους διάφορους τύπους βίντεο, οι οποίοι προαναφέρθηκαν στο Σχήμα 5.103. Αναλυτικότερα, βάσει του Σχήματος 5.106, προκύπτει ότι αρκετά υψηλό ποσοστό *Views-Subscribers Ratio* σημείωσαν τα βίντεο *Διαγωνισμών (Giveaway)*, υπενθυμίζοντας για ακόμη μία φορά την υψηλή αποδοτικότητα τέτοιου είδους βίντεο στη προσέγγιση νέου κοινού. Στα υψηλότερα επίπεδα της κατάταξης έκαναν, επίσης, την εμφάνισή τους τα βίντεο *Διαδικτυακής Συζήτησης (Podcast)* με ποσοστό μεγαλύτερο του 100%, κάτι το οποίο υποδεικνύει τις δυνατότητες προσέλκυσης νέων εγγεγραμμένων χρηστών στον προκειμένο κλάδο των βίντεο. Σε αντίθεση με τα παραπάνω, τα μικρότερα ποσοστά προβολών - εγγεγραμμένων χρηστών κατείχαν οι *Αξιολογήσεις (Review)*, τα βίντεο τύπου *Διαβάζουμε* και οι *Ζωντανές Ροές (Livestream)*, δείχνοντας πως τέτοιου είδους βίντεο, πιθανώς, απευθύνονται στο πιστό κοινό του δημιουργού και, ως εκ τούτου, προσελκύουν δυσκολότερα νέους χρήστες στο κανάλι. Η παρουσία των βίντεο τύπου Άλλο στην κορυφή της κατάταξης δεν είναι δυνατό να αποδοθεί σε κάποιον συγκεκριμένο παράγοντα, καθώς η συγκεκριμένη κατηγορία ενθυλακώνει μια ευρεία γκάμα ετερογενών βίντεο.



Σχήμα 5.106: Σχέση μεταξύ εγγεγραμμένων χρηστών και προβολών ανά τύπο βίντεο

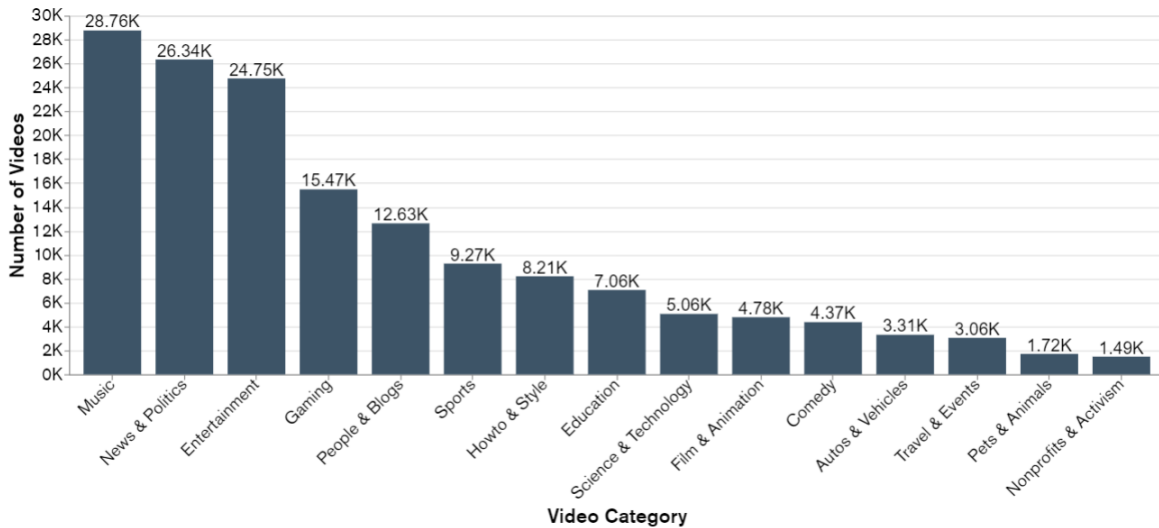
Έναν από τους βασικούς άξονες της έρευνας του YouTube αποτέλεσε, επίσης, η εξέταση του αριθμού των δημοσιευμένων βίντεο και η ανάπτυξη συγκεκριμένων τύπων βίντεο κατά τη διάρκεια του χρονολογικού έτους 2020, καθώς και η αναζήτηση κάποιας ενδεχόμενης σχέσης εξάρτησης μεταξύ του συνόλου του συλλεγμένων βίντεο του δείγματος και του ποσοστού αλληλεπίδρασης ανά τους διαφορετικούς μήνες.

Σημείο εκκίνησης της προκείμενης διαδικασίας ήταν η ανάλυση της δραστηριότητας των διαφορετικών κατηγοριών χρηστών του YouTube κατά το έτος 2020. Σύμφωνα με το Σχήμα 5.107, παρατηρείται πως οι λογαριασμοί τύπου Άλλο παρήγαγαν τα περισσότερα βίντεο, ξεπερνώντας κατά ένα μεγάλο βαθμό, σε ποσότητα υλικού, τις υπόλοιπες κατηγορίες. Η κατάληψη της δεύτερης θέσης της κατάταξης από τους άνδρες, με επίσης υψηλό αριθμό δημοσιευμένων βίντεο, επιβεβαιώνει πως εκείνοι και οι λογαριασμοί τύπου Άλλο αποτέλεσαν τα κανάλια τα οποία έκαναν εντονότερη την παρουσία τους στην ελληνική κοινότητα του YouTube. Ωστόσο, αξιοσημείωτη είναι η εμφάνιση των γυναικών στα χαμηλότερα στρώματα του αριθμού μεταφορτώσεων για το 2020, καθώς εκείνες, παρά τη χαμηλή δραστηριότητα τους, κατάφεραν να ανταγωνιστούν όλες τις υπόλοιπες κατηγορίες χρηστών, καταγράφοντας αρκετά υψηλό αριθμό προβολών και αντιδράσεων στα βίντεο τους (βλ. Σχήματα 5.89, 5.91, 5.92 και 5.94). Θα μπορούσε να εξαχθεί, λοιπόν, από τα παραπάνω πως οι γυναίκες διατηρούν πιο περιορισμένο, αλλά ιδιαίτερα δυναμικό κοινό. Στην τελευταία θέση βρέθηκαν οι επιχειρήσεις, οι οποίες δημιούργησαν τη μικρότερη ποσότητα υλικού κατά το ίδιο χρονικό διάστημα, γεγονός το οποίο επηρεάστηκε και από τον χαμηλό αριθμό επιχειρησιακών καναλιών στο δείγμα της έρευνας (βλ. Σχήμα 5.78).



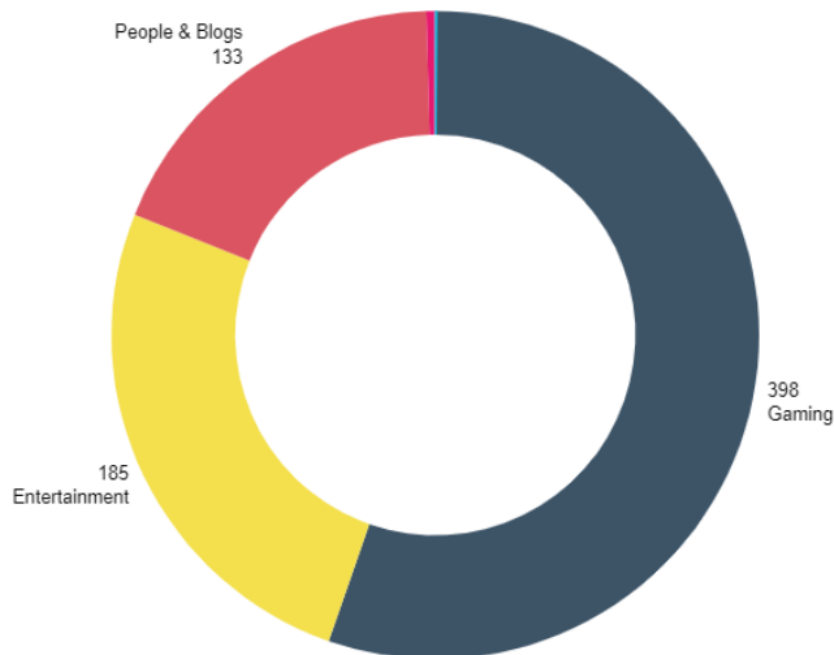
Σχήμα 5.107: Αριθμός βίντεο ανά κατηγορία χρήστη

Ακολουθώντας την ίδια τακτική για τις ποικίλες θεματικές κατηγορίες περιεχομένου, προκύπτει από το Σχήμα 5.108 πως οι κατηγορίες οι οποίες παρήγαγαν τη μεγαλύτερη ποσότητα υλικού κατά το 2020 ήταν η *Μουσική (Music)*, τα *Νέα και Πολιτική (News & Politics)* και η *Ψυχαγωγία (Entertainment)*. Εστιάζοντας στις κατηγορίες *Μουσική (Music)* και *Ψυχαγωγία (Entertainment)*, παρά τη μεγάλη τους δραστηριότητα και τον αρκετά υψηλό αριθμό προβολών τον οποίον αυτές συγκέντρωσαν (βλ. Σχήματα 5.100 και 5.101), το ποσοστό αλληλεπίδρασης τους με το κοινό παρέμεινε αρκετά χαμηλό (βλ. Σχήμα 5.102), υποδηλώνοντας ότι το κοινό καταναλώνει με γρήγορους ρυθμούς τέτοιου είδους περιεχόμενο, αλλά δεν καταφέρνει να απορροφήσει την ουσία του και, κατ' επέκταση, να ταυτιστεί ουσιαστικά με αυτό. Όσον αφορά στον κλάδο *Νέα και Πολιτική (News & Politics)*, διαπιστώνεται πως η πολύ μεγάλη προσφορά αυτών των βίντεο δεν απορροφήθηκε από το κοινό, συμπέρασμα το οποίο είναι εύκολο να εξαχθεί από τους πολύ χαμηλούς αριθμούς αντιδράσεων και προβολών, οι οποίοι σημειώθηκαν στα Σχήματα 5.95, 5.97, 5.98, 5.100 και 5.102. Αρκετά υψηλό αριθμό βίντεο παρήγαγαν, επίσης, οι κατηγορίες *Ηλεκτρονικό Παιχνίδι (Gaming)* και *Καθημερινότητα και Ανθρώπινες Σχέσεις (People & Blogs)*, κάτι το οποίο σε συνδυασμό με τις μετρικές απόδοσης τους επιβεβαιώνει την μεγάλη απήχηση των βίντεο παιχνιδιού και το αξιοσημείωτο ενδιαφέρον του κοινού για βίντεο ανθρώπινων σχέσεων κατά το έτος 2020. Από την άλλη μεριά, οι κατηγορίες *Κωμωδία (Comedy)*, *Αυτοκίνηση (Autos & Vehicles)* και *Επιστήμες και Τεχνολογία (Science & Technology)*, ενώ συγκέντρωσαν έναν σημαντικό αριθμό προβολών και αντιδράσεων, διατήρησαν έναν χαμηλό ρυθμό δημιουργίας περιεχομένου, φαινόμενο το οποίο πιθανώς υποδεικνύει τα περιθώρια παραγωγής περισσότερου υλικού στις κατηγορίες αυτές.



Σχήμα 5.108: Αριθμός βίντεο ανά θεματική κατηγορία βίντεο

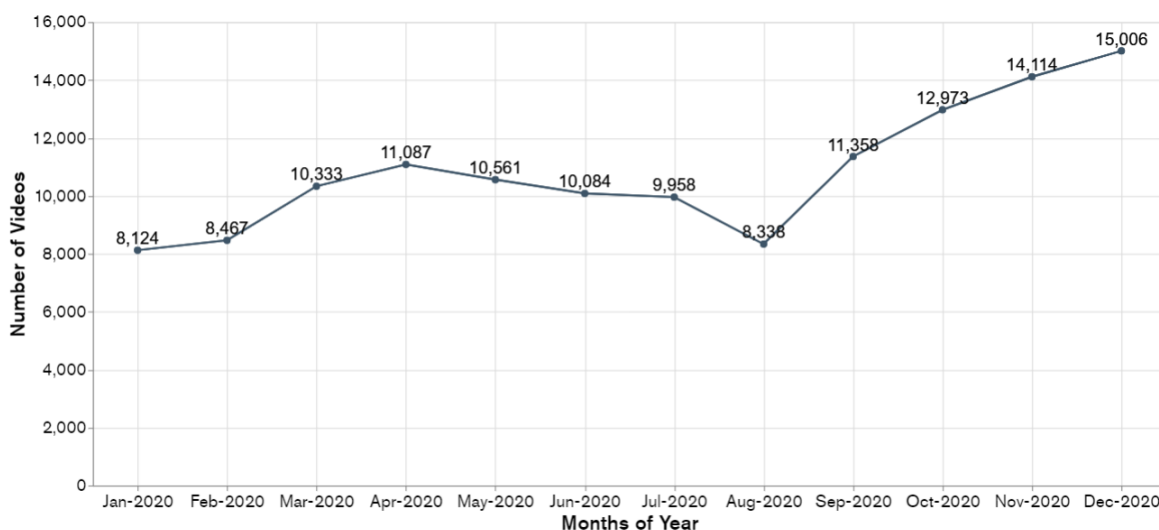
Με σκοπό την εκτενέστερη εξιχνίαση των σύγχρονων τάσεων στην ελληνική κοινότητα του YouTube, θεωρήθηκε ενδιαφέρουσα η αναζήτηση των πιο διαδεδομένων θεματικών κατηγοριών περιεχομένου με βάση τα εντός του 2020 δημοσιευμένα βίντεο, από κανάλια τα οποία δημιουργήθηκαν την ίδια χρονιά. Αποδεικνύεται, για ακόμη μια φορά, πως οι κατηγορίες *Ηλεκτρονικό Παιχνίδι (Gaming)*, *Ψυχαγωγία (Entertainment)* και *Καθημερινότητα και Ανθρώπινες Σχέσεις (People & Blogs)* αποτελούν πόλους υψηλού ενδιαφέροντος για το κοινό, καθώς, όπως φαίνεται στο Σχήμα 5.109, η συντριπτική πλειοψηφία των βίντεο νεοσύστατων καναλιών στράφηκε προς τις κατευθύνσεις αυτών των θεματικών κατηγοριών.



Σχήμα 5.109: Κορυφαίες θεματικές κατηγορίες βίντεο σε κανάλια δημιουργημένα εντός του 2020

Την εξέταση του γενικότερο αριθμού μεταφορτώσεων, με βάση τις διαφορετικές κατηγορίες χρηστών και περιεχομένου, διαδέχθηκε η παρακολούθηση της εξέλιξης του πλήθους δημοσιευμένων βίντεο ανά τους διαδοχικούς μήνες του έτους 2020.

Ξεκινώντας τη μελέτη του Σχήματος 5.110 και μεταβαίνοντας προς τους ανοιξιάτικους μήνες του 2020, γίνεται φανερή η αύξηση της δραστηριότητας των Ελλήνων δημιουργών, η οποία έγινε εντονότερη κατά τον Μάρτιο και κορυφώθηκε τον Απρίλιο. Αξίζει να σημειωθεί πως η συγκεκριμένη αύξηση στον αριθμό των συγκεκριμένων βίντεο έλαβε χώρα κατά το χρονικό διάστημα το οποίο ενείχε την εφαρμογή των πρώτων περιοριστικών μέτρων ενάντια στον ιό COVID-19, καθώς και τις πασχαλινές εορτές. Συνεχίζοντας προς το καλοκαίρι, φαίνεται πως η παραγωγικότητα των δημιουργών ακολούθησε φθίνουσα πορεία, με τον χαμηλότερο αριθμό δημοσιεύσεων για εκείνο το διάστημα να σημειώνεται κατά τον Αύγουστο, φαινόμενο το οποίο πιθανώς να οφείλεται στην χαλαρότητα που διέπει τη θερινή τουριστική περίοδο. Φτάνοντας προς τους τελευταίους μήνες του 2020, παρατηρείται πως η δραστηριότητα των δημιουργών επανήλθε σε υψηλότερα επίπεδα, σε σχέση με την προ-καλοκαιρινή περίοδο, και συνέχισε την σταθερά αύξουσα πορεία της ωσότου αυτή καταλήξει σε ολικό μέγιστο κατά τη χριστουγεννιάτικη περίοδο. Σε γενικότερα πλαίσια, παρατηρείται πως η παραγωγικότητα των ελληνικών καναλιών αυξήθηκε κατά τη διάρκεια του 2020, γεγονός στο οποίο, πιθανότατα, συνέβαλαν οι συνθήκες εγκλεισμού οι οποίες χαρακτήρισαν το συγκεκριμένο έτος.

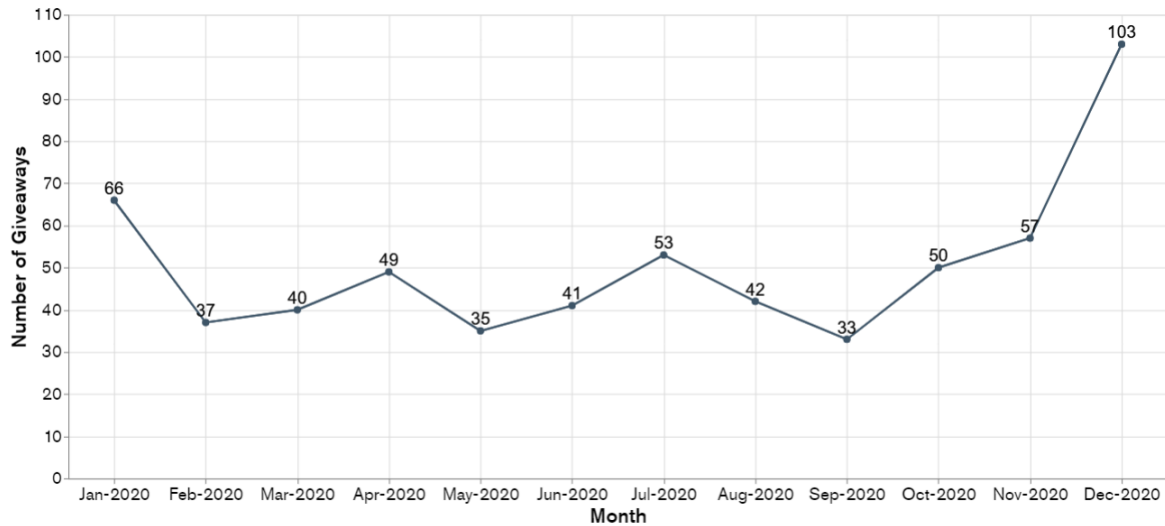


Σχήμα 5.110: Αριθμός βίντεο ανά μήνα

Τοποθετώντας στο μικροσκόπιο τον αριθμό των δημοσιευμένων βίντεο ανά μήνα, η προσοχή εστιάστηκε σε συγκεκριμένους τύπους βίντεο και θεματικές κατηγορίες περιεχομένου, τα οποία είτε εμφάνιζαν εμπορικό ενδιαφέρον, λόγω της εμπλοκής επιχειρήσεων, ή παρείχαν επιπλέον στοιχεία για τις τάσεις της ελληνικής κοινότητας του YouTube.

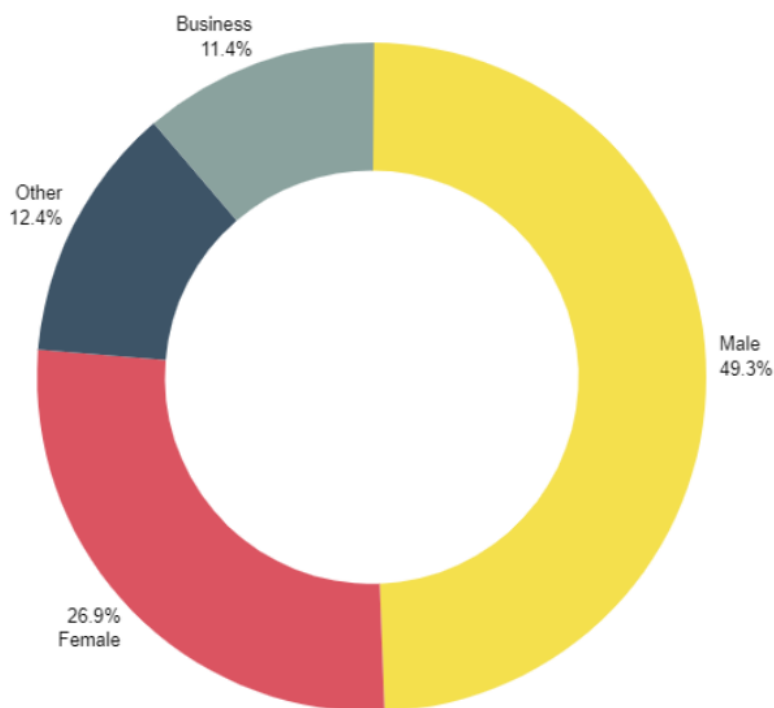
Το υψηλό ποσοστό συμμετοχής των επιχειρήσεων στα βίντεο διαγωνισμών (*Giveaway*) αποτέλεσε τον κυριότερο λόγο για τον οποίο επιλέχθηκε η λεπτομερέστερη μελέτη αυτού του είδους βίντεο. Σύμφωνα με την εξέλιξη του πλήθους των διαγωνισμών ανά τους μήνες του 2020, η οποία αποτυπώνεται στο Σχήμα 5.111, διαπιστώνεται πως οι μήνες Ιανουάριος, Απρίλιος, Ιούλιος και

Δεκέμβριος ήταν αυτοί κατά τους οποίους σημειώθηκε κορύφωση στον αριθμό των *Giveaways*.

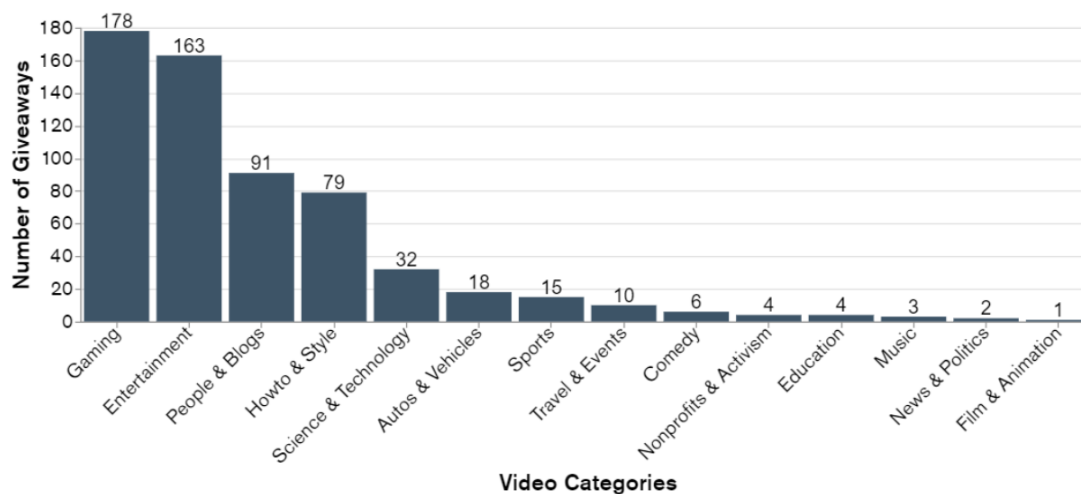


Σχήμα 5.111: Αριθμός διαγωνισμών (Giveaway) ανά μήνα

Το ενδιαφέρον για την εξιχνίαση του προφίλ των καναλιών, τα οποία επιλέγονται από τις επιχειρήσεις για τη προώθηση των προϊόντων τους, οδήγησε στη δημιουργία των Σχημάτων 5.112 και 5.113. Είναι, πλέον, δυνατό να επιβεβαιωθεί πως οι επιχειρήσεις στρέφονται προς άνδρες και γυναίκες δημιουργούς, με σκοπό τη διοργάνωση διαγωνισμών και, παράλληλα, τη προώθηση τους σε ένα ευρύτερο κοινό, τακτική γνωστή και ως *Influencer Marketing*. Εντούτοις, στη περίπτωση του YouTube, φαίνεται να υπάρχει ιδιαίτερη προτίμηση προς τους ανδρικούς λογαριασμούς για τη διεξαγωγή διαγωνισμών, καθώς σχεδόν ο μισός αριθμός διαγωνισμών του δείγματος ανήκει σε αυτούς. Η συγκεκριμένη παρατήρηση είναι δυνατό να συνδεθεί και με το Σχήμα 5.113, από όπου προκύπτει ότι το *Ηλεκτρονικό Παιχνίδι (Gaming)* και η *Ψυχαγωγία (Entertainment)*, κατηγορίες στις οποίες δραστηριοποιούνται αρκετά ανδρικά κανάλια (βλ. Σχήμα 5.79), φιλοξένησαν τον μεγαλύτερο αριθμό διαγωνισμών για τη χρονική περίοδο του 2020. Αρκετοί διαγωνισμοί, επίσης, καταγράφηκαν στις κατηγορίες *Καθημερινότητα και Ανθρώπινες Σχέσεις (People & Blogs)* και *Στιλιστικό περιεχόμενο (Howto & Style)*, τομείς οι οποίοι καταλήφθηκαν κατά ένα μεγάλο μέρος από γυναίκες δημιουργούς. Τέλος, ένας σεβαστός αριθμός διαγωνισμών διοργανώθηκαν στα πλαίσια των *Επιστημών και Τεχνολογίας (Science & Technology)*, υποδηλώνοντας το, επίσης, υπάρχον εμπορικό ενδιαφέρον στη προκείμενη θεματική κατηγορία περιεχομένου.



Σχήμα 5.112: Αριθμός διαγωνισμών (Giveaway) ανά κατηγορία χρήστη

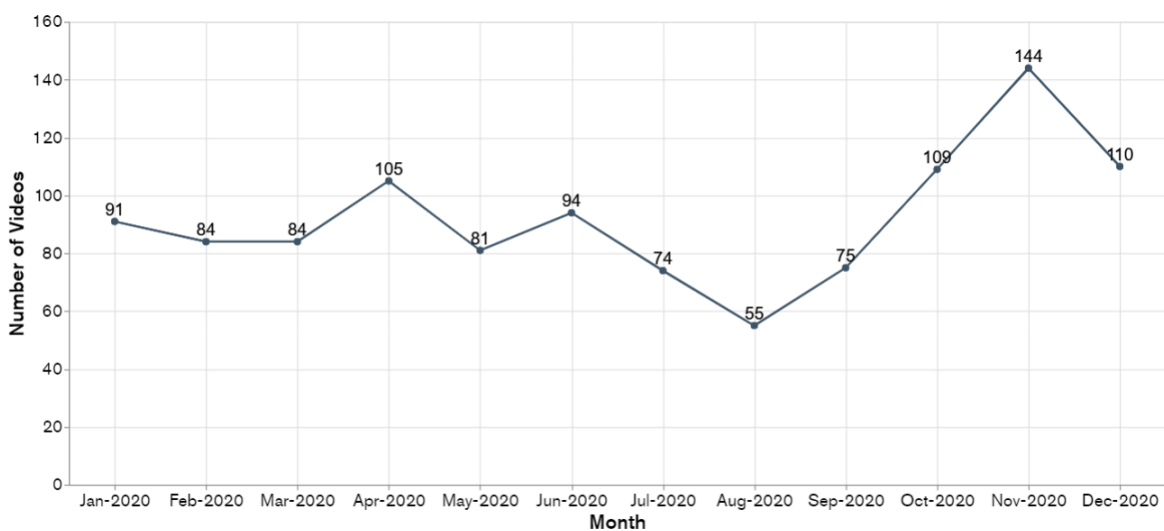


Σχήμα 5.113: Αριθμός διαγωνισμών (Giveaway) ανά θεματική κατηγορία βίντεο

Παρόμοιας φύσεως με τα βίντεο Διαγωνισμών (*Giveaways*) είναι τα βίντεο τύπου *Unboxing* και οι *Αξιολογήσεις (Reviews)*. Όπως προαναφέρθηκε, στα πλαίσια αυτών των τύπων βίντεο, ο δημιουργός περιεχομένου καλείται είτε να αποσυσκευάσει και, στη συνέχεια, να παρουσιάσει στο κοινό ένα προϊόν ή να ασκήσει κριτική στο προϊόν αυτό, το οποίο πιθανώς έχει παραλάβει για τον σκοπό αυτόν από κάποια επιχείρηση. Η ενεργή συμμετοχή των εταιρειών στη δημιουργία τέτοιου είδους περιεχομένου παρότρυνε την εξατομικευμένη μελέτη της εξέλιξης των βίντεο τύπου *Unboxing* και των *Αξιολογήσεων (Reviews)* ανά τους διαδοχικούς μήνες του 2020. Ως εκ τούτου, διαπιστώνεται από το Σχήμα 5.114 ότι το πλήθος των παραπάνω τύπων βίντεο, ακολούθησε

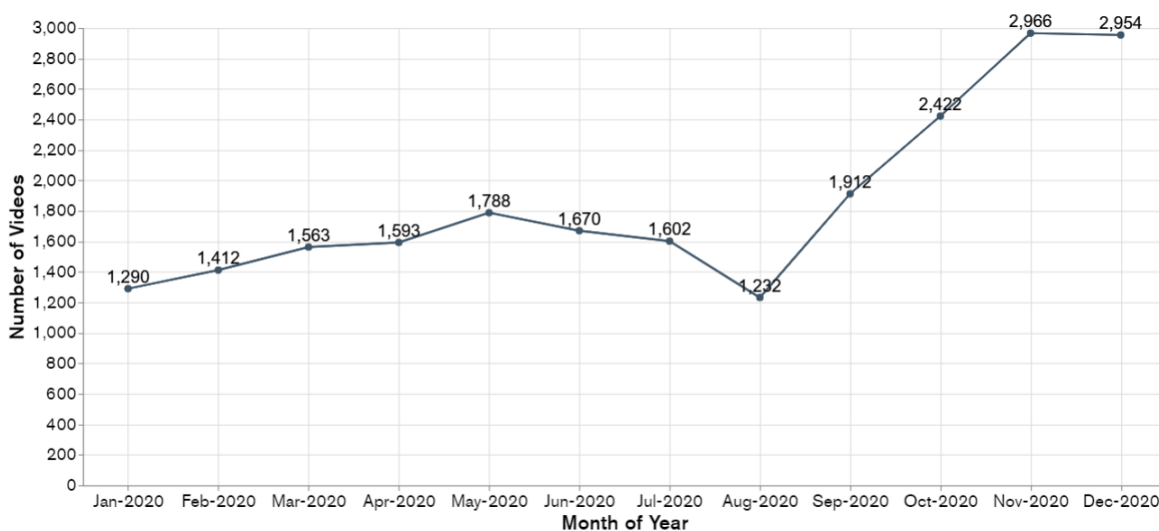
αντίστοιχο μοτίβο με εκείνο των *Διαγωνισμών (Giveaways)* (βλ. Σχήμα 5.111), παρουσιάζοντας μια μικρή άνηση κατά τους μήνες Απρίλιο (Πασχαλινές εορτές), Ιούνιο (έναρξη καλοκαιρινών διακοπών) και κατά τους τελευταίους μήνες του έτους, οδεύοντας προς τις εορτές των Χριστουγέννων.

Διασταυρώνοντας τα Σχήματα 5.111 και 5.114, γίνεται αντιληπτό πως οι επιχειρήσεις ενέτειναν τις προσπάθειες προώθησης των προϊόντων και των υπηρεσιών τους σε περιόδους αυξημένης καταναλωτικής δραστηριότητας, όπως είναι οι Πασχαλινές εορτές, η καλοκαιρινή περίοδος και οι γιορτές των Χριστουγέννων, με τις τελευταίες να χαρακτηρίζονται από την εντονότερη τοποθέτηση προϊόντων των Ελλήνων δημιουργών για το έτος 2020.

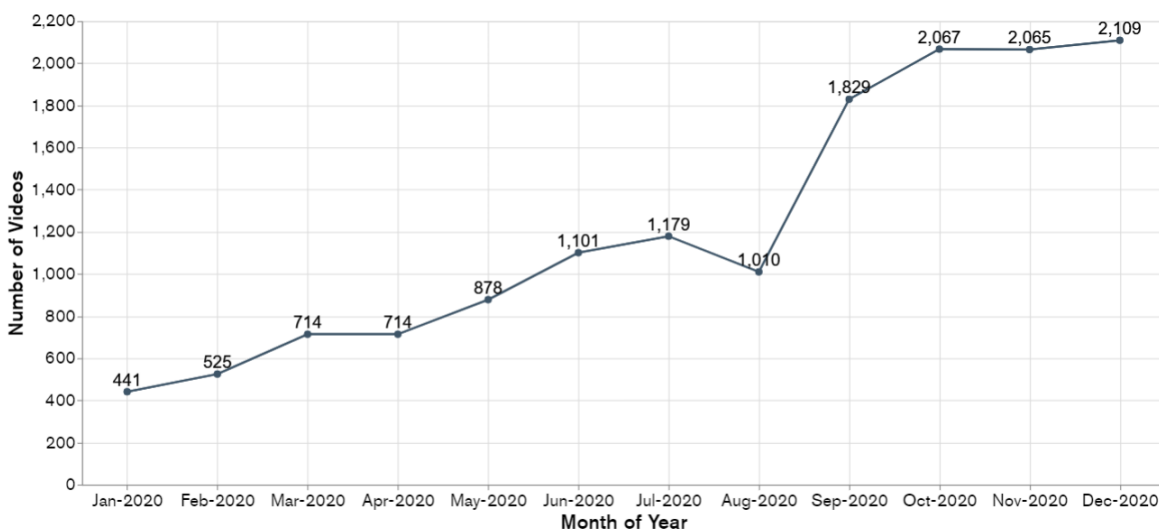


Σχήμα 5.114: Αριθμός βίντεο τύπου Unboxing και βίντεο αξιολόγησης (Review) ανά μήνα

Όσον αφορά στις θεματικές κατηγορίες περιεχομένου, ιδιαίτερα ενδιαφέρουσα θεωρήθηκε η εξέλιξη των βίντεο *Ψυχαγωγίας (Entertainment)*, *Κωμωδίας (Comedy)* και *Νέων και Πολιτικής (News & Politics)* κατά τη διάρκεια του 2020, καθώς, όπως είναι φανερό στα Σχήματα 5.115 και 5.116 ο αριθμός τους ακολούθησε ανοδική πορεία, δείχνοντας έτσι πως το συγκεκριμένο έτος, συμπεριλαμβανομένης της μεταβολής της καθημερινότητας του κοινού εξαιτίας της πανδημίας του *COVID-19*, ευνόησε την ευημερία αυτών των θεματικών κατηγοριών περιεχομένου. Παρατηρείται, ωστόσο, πως η αύξηση των τριών προηγούμενων θεματικών κατηγοριών αντιμετώπισε μια σύντομη πτώση κατά τον μήνα Αύγουστο, γεγονός το οποίο ήρθε σε συμφωνία με τα δεδομένα της καλοκαιρινής περιόδου του Σχήματος 5.110, σηματοδοτώντας τη πτώση της παραγωγικότητας των Ελλήνων δημιουργών κατά τη κορύφωση της καλοκαιρινής τουριστικής περιόδου.



Σχήμα 5.115: Αριθμός βίντεο ψυχαγωγίας (Entertainment) και κωμωδίας (Comedy) ανά μήνα

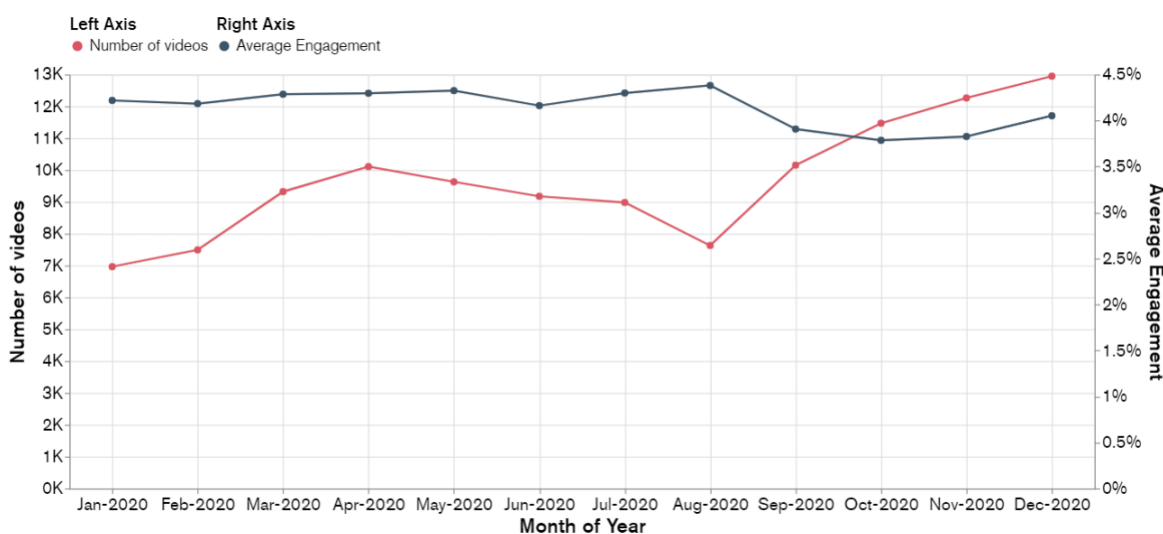


Σχήμα 5.116: Αριθμός βίντεο νέων και πολιτικής (News & Politics) ανά μήνα

Η ανταπόκριση του κοινού στη παραπάνω καταγεγραμμένη δραστηριότητα των Ελλήνων δημιουργών, εξετάστηκε μέσω του υπολογισμού του ποσοστού αλληλεπίδρασης του με τα δημοσιευμένα βίντεο ανά τους μήνες του έτους 2020. Για τις ανάγκες της συγκεκριμένης διερεύνησης και με σκοπό την αποφυγή των ακραίων περιπτώσεων, λήφθηκαν υπόψη μόνο τα βίντεο τα οποία δε ξεπέρασαν, κατά το 2020, το 20% του ποσοστού αλληλεπίδρασης.

Με βάση το Σχήμα 5.117, είναι εμφανές πως το μέσο ποσοστό αλληλεπίδρασης δεν μεταβλήθηκε δραματικά κατά τη διάρκεια του έτους. Πιο συγκεκριμένα, φαίνεται πως υπήρξε μια μικρή αύξηση τον Μάρτιο, η οποία διατηρήθηκε καθ' όλη τη διάρκεια της άνοιξης. Δεδομένης της αύξησης αυτής, θα μπορούσε να εκτιμηθεί πως η έναρξη της καραντίνας τον μήνα Μάρτιο, κατά ένα μικρό ποσοστό, ενίσχυσε το ενδιαφέρον του κοινού για τα περιεχόμενα της πλατφόρμας.

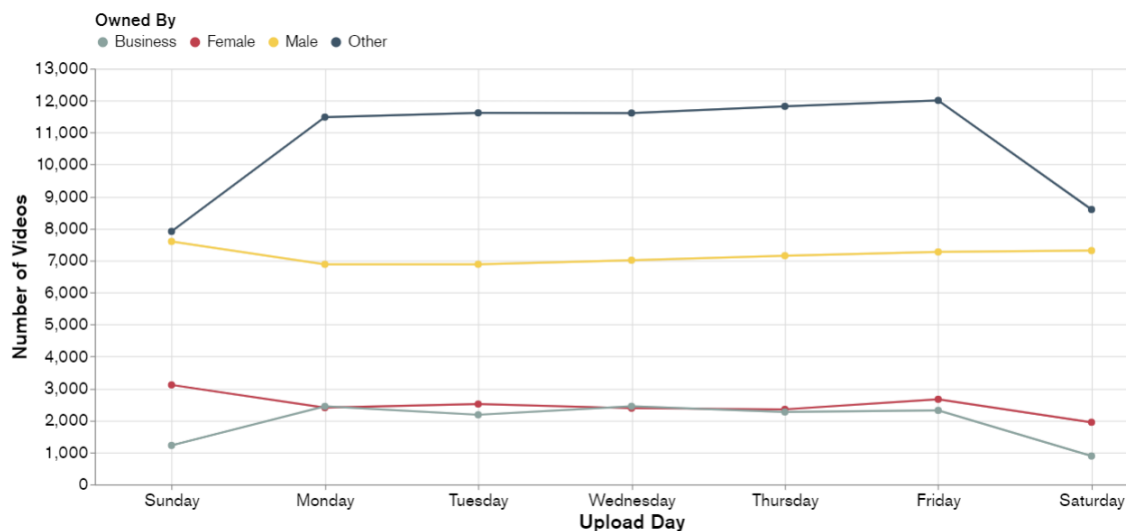
Εξαιρουμένης της μικρής πτώσης του ποσοστού αλληλεπίδρασης η οποία επήλθε τον Ιούνιο, αυτό ακολούθησε σταθερά ανοδική πορεία κατά τους καλοκαιρινούς μήνες, ωστόσο να κορυφωθεί τον Αύγουστο. Παρατηρείται, λοιπόν, πως σε αντίθεση με τους Έλληνες δημιουργούς της πλατφόρμας, ο μήνας Αύγουστος αποτέλεσε τη γονιμότερη περίοδο του έτους, όσον αφορά στην διάδραση του κοινού με τα αναρτημένα βίντεο, γεγονός στο οποίο ενδεχομένως συνέβαλαν στη χαλαρότητα της καλοκαιρινής περιόδου, το θεματικό περιεχόμενο εκείνου του διαστήματος, αλλά και ο χαμηλότερος φόρτος βίντεο. Οι φθινοπωρινοί μήνες, παράλληλα με την έντονη αύξηση των προς παρακολούθηση βίντεο, χαρακτηρίστηκαν από μία αισθητή πτώση του μέσου ποσοστού αλληλεπίδρασης, το οποίο τελικά επανήλθε κοντά στις αρχικές τιμές του έτους, μεταβαίνοντας προς τις Χριστουγεννιάτικες εορτές του Δεκεμβρίου. Αξίζει να σημειωθεί πως οι χαμηλότερες τιμές του μέσου ποσοστού αλληλεπίδρασης προς τα τέλη του έτους 2020 επηρεάστηκαν, έως ένα βαθμό, από τον χαμηλότερο βαθμό ωρίμανσης των πιο πρόσφατα αναρτημένων βίντεο, σε σχέση με βίντεο που δημοσιεύτηκαν σε προγενέστερο χρονικό διάστημα.



Σχήμα 5.117: Μέσο ποσοστό αλληλεπίδρασης ανά μήνα

Ως επόμενο στάδιο της ερευνητικής διαδικασίας επιλέχθηκε η ανάλυση των προτιμήσεων των Ελλήνων δημιουργών, όσον αφορά στη χρονική στιγμή ανάρτησης των βίντεο σε επίπεδο ημέρας και ώρας, καθώς και η αναζήτηση κάποιας σχέσης μεταξύ της συχνότητας ανάρτησης βίντεο και του αριθμού εγγεγραμμένων χρηστών ενός καναλιού.

Σύμφωνα με το Σχήμα 5.118, διακρίνεται πως τα κανάλια τύπου Άλλο και οι επιχειρησιακοί λογαριασμοί ακολούθησαν, κατά το 2020, ένα πρόγραμμα δημοσίευσης περιεχομένου βασισμένο στις εργάσιμες ημέρες Δευτέρα έως Παρασκευή. Ο συνδυασμός του πολύ μεγάλου αριθμού δημοσιεύσεων των λογαριασμών τύπου Άλλο και της τάσης αυτών να δημοσιεύουν το περιεχόμενο τους κατά τις εργάσιμες ημέρες, ίσως, αποτελεί ένδειξη ενός πιο αυστηρού και επαγγελματικού χαρακτήρα τον οποίο υιοθετούν τέτοιου είδους λογαριασμοί. Από την άλλη πλευρά, οι άνδρες και οι γυναίκες του YouTube φαίνεται πως προσαρμόσαν, έως ένα βαθμό, τη δραστηριότητά τους γύρω από τις ημέρες του Σαββατοκύριακου, με μια ιδιαίτερη προτίμηση προς την Παρασκευή και την Κυριακή να είναι εμφανής, δείχνοντας έτσι την κλίση τους προς τις ημέρες της εβδομάδας με λιγότερες εργασιακές και σχολικές υποχρεώσεις.

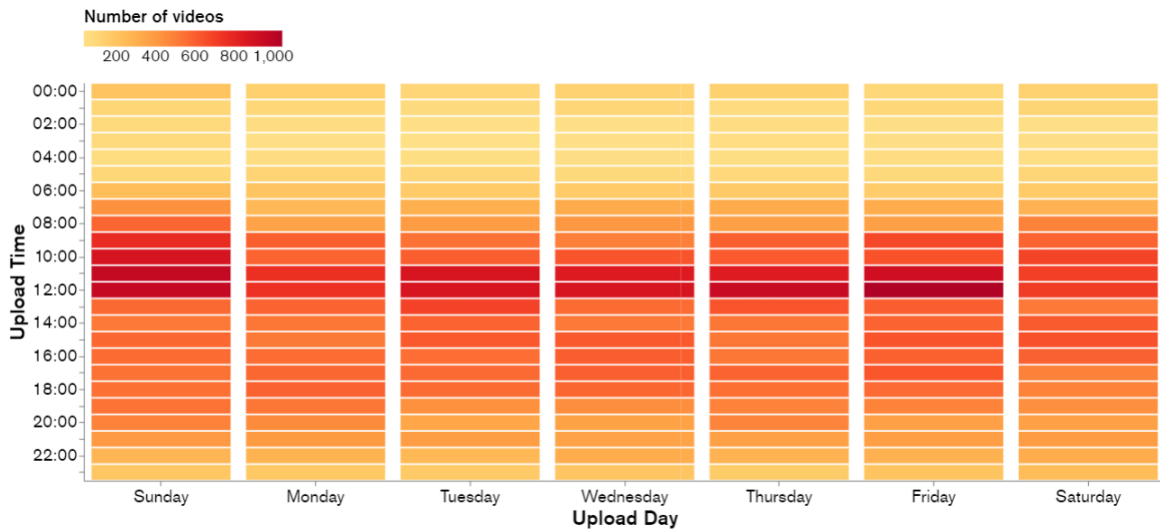


Σχήμα 5.118: Αριθμός βίντεο ανά ημέρα και κατηγορία χρηστών

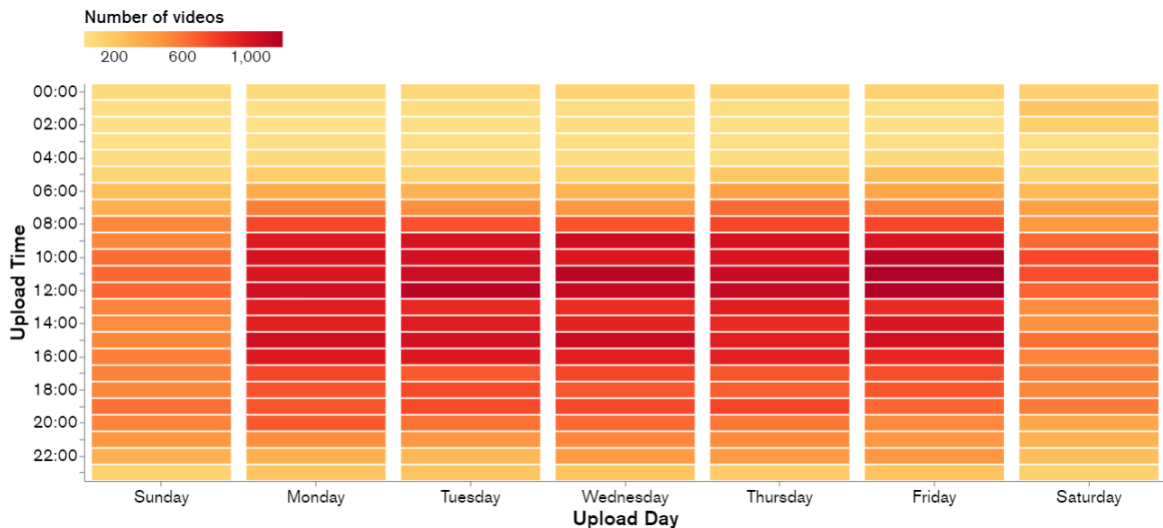
Η ανάγκη για μία βαθύτερη οπτική της συμπεριφοράς των Ελλήνων δημιουργών, μετέφερε το ερευνητικό ενδιαφέρον από το επίπεδο ημερών σε επίπεδο ωρών και συγκεκριμένα στη διαλεύκανση των ωρών της ημέρας, κατά τις οποίες καταγράφηκε η υψηλότερη δραστηριότητα ανά κατηγορία χρηστών.

Με αφετηρία τους άνδρες και τις γυναίκες, οι οποίοι παρουσίασαν κοινές προτιμήσεις όσον αφορά στις ημέρες ανάρτησης βίντεο, είναι φανερό από το Σχήμα 5.119 πως αυτοί δημοσίευσαν τη πλειοψηφία των βίντεο τους κατά τις μεσημεριανές και απογευματινές ώρες, με τις ώρες αιχμής να είναι 11:00 π.μ.-12:00 μ.μ.. Ωστόσο, φαίνεται πως η έντονη δραστηριότητα των ανδρών και των γυναικών τη Κυριακή κάλυψε ένα μεγαλύτερο χρονικό παράθυρο, μεταξύ 09:00 π.μ. και 12:00 μ.μ..

Συνεχίζοντας στους επιχειρησιακούς λογαριασμούς και τα κανάλια τύπου Άλλο, διαπιστώνεται στο Σχήμα 5.120 ότι η δραστηριότητα τους παρέμεινε σταθερή μεταξύ των εργάσιμων ημερών, με τις ώρες αιχμής να εντοπίζονται μεταξύ 08:00 π.μ. και 07:00 μ.μ.. Όπως προέκυψε και από το Σχήμα 5.118, η παραγωγικότητα των συγκεκριμένων κατηγοριών χρηστών παρέμεινε σε αρκετά χαμηλά επίπεδα κατά τη διάρκεια των Σαββατοκύριακων.

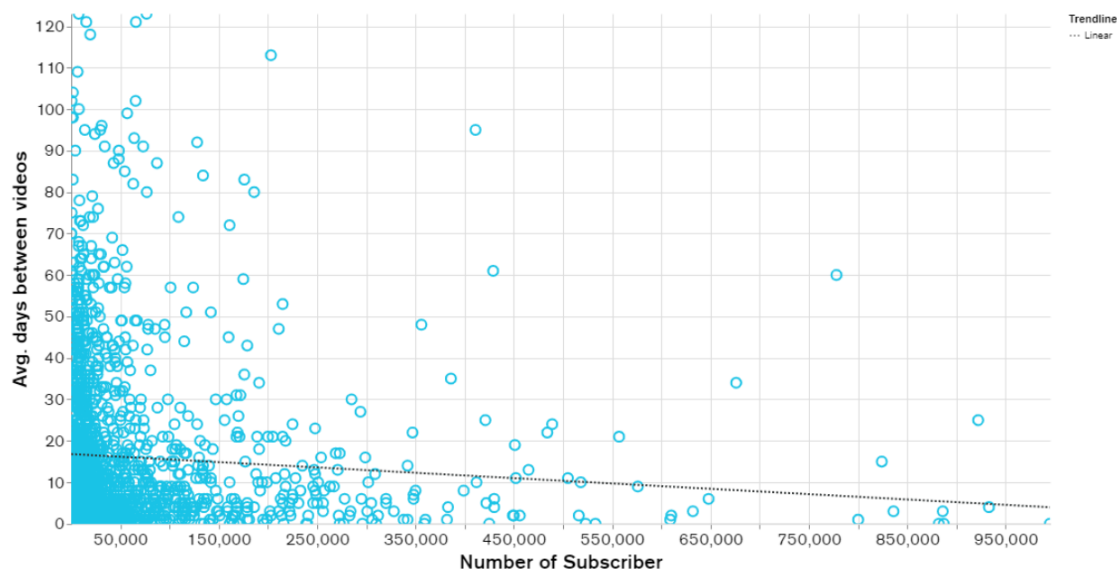


Σχήμα 5.119: Αριθμός βίντεο ανά ημέρα και ώρα (Άνδρες, Γυναίκες)



Σχήμα 5.120: Αριθμός βίντεο ανά ημέρα και ώρα (Επιχειρήσεις, Άλλο)

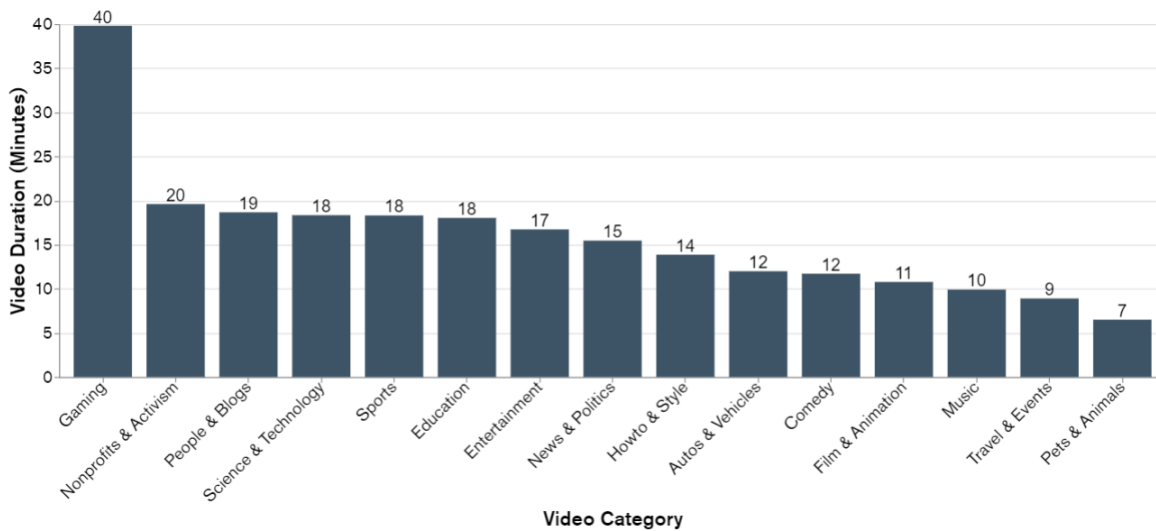
Καταλήγοντας στη σχέση μεταξύ της συχνότητας ανάρτησης βίντεο και του αριθμού εγγεγραμμένων χρηστών, μέσω της ανάλυσης τους Σχήματος 5.121, εξάγεται το συμπέρασμα ότι, κατά μέσο όρο, κανάλια τα οποία συγκεντρώνουν ένα ευρύτερο πιστό κοινό δημοσιεύουν περιεχόμενο αρκετά συχνότερα σε σχέση με τα μικρότερης έκτασης κανάλια. Ωστόσο, η διαπίστωση αυτή δεν αποτελεί τον κανόνα, καθώς αρκετά είναι τα κανάλια μικρότερης εμβέλειας τα οποία παραμένουν εξίσου δραστήρια. Το φαινόμενο αυτό πιθανότατα ανάγεται στην προσπάθεια των μεγαλύτερων καναλιών να διατηρήσουν σε υψηλά επίπεδα το ενδιαφέρον του κοινού τους και, παράλληλα, να διασφαλίσουν την ομαλή λειτουργία τους. Επισημαίνεται πως, με σκοπό τη βέλτιστη οπτική αναπαράσταση του παρακάτω διαγράμματος, η μελέτη επικεντρώθηκε στα κανάλια με αριθμό εγγεγραμμένων χρηστών μικρότερο του 1.000.000.



Σχήμα 5.121: Σχέση μεταξύ συχνότητας ανάρτησης βίντεο και αριθμού εγγεγραμμένων χρηστών

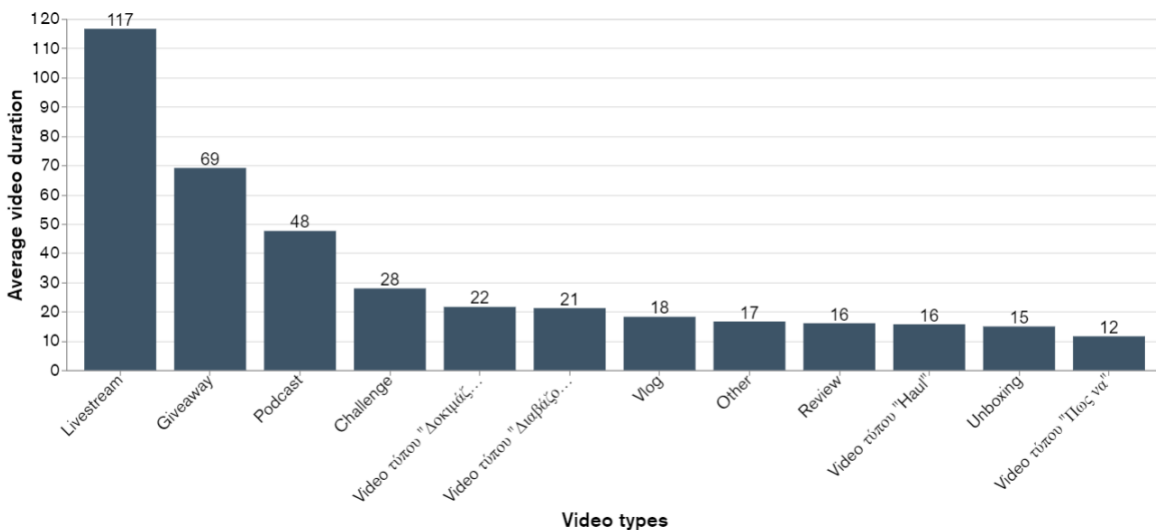
Ακόμα ένα στοιχείο των βίντεο του YouTube, το οποίο θα μπορούσε να προσφέρει ενδιαφέρουσες πληροφορίες στην ερευνητική διαδικασία, είναι η χρονική διάρκεια των βίντεο. Για τον λόγο αυτό, αποφασίστηκε να υπολογιστεί η μέση διάρκεια των ποικίλων θεματικών κατηγοριών και τύπων βίντεο και, τελικά, να αποτυπωθεί η σχέση μεταξύ των προβολών και της διάρκειας ενός βίντεο. Με σκοπό την αποφυγή των ακραίων περιπτώσεων, όσον αφορά στα βίντεο με μεγάλη χρονική διάρκεια, λήφθηκαν υπόψη μόνο οι μεταφορτώσεις με διάρκεια μικρότερη ή ίση των τριών ωρών.

Στον άξονα των θεματικών κατηγοριών βίντεο, βάσει του Σχήματος 5.122, γίνεται αντιληπτό πως η πλειοψηφία των θεματικών κατηγοριών κυμαίνεται σε χρονική διάρκεια μεταξύ 10 και 20 λεπτών, ωστόσο τα βίντεο *Ηλεκτρονικού Παιχνιδιού (Gaming)*, τα οποία διέπρεψαν κατά το 2020 όπως φάνηκε παραπάνω, φαίνεται να αποκλίνουν από τον μέσο όρο, με τη μέση διάρκειά τους να αγγίζει τα 40 λεπτά. Στην αντίθετη πλευρά, οι κατηγορίες *Μουσική (Music)*, *Ταξιδιωτικές Δραστηριότητες (Travel & Events)* και *Ζώα και Κατοικίδια (Pets & Animals)* απαρτίζονται σε μεγαλύτερο ποσοστό από βίντεο μικρότερης χρονικής διάρκειας.



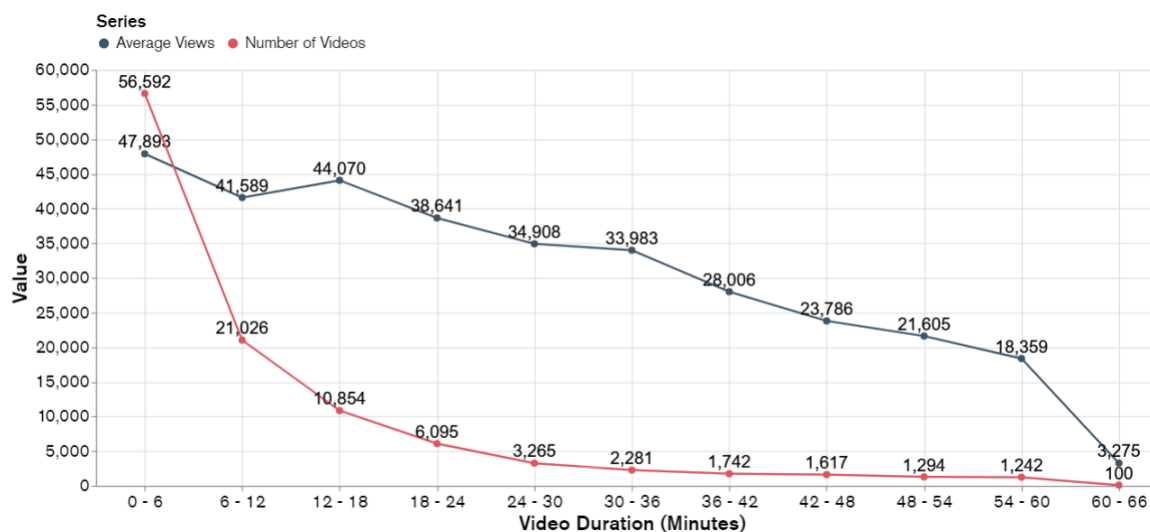
Σχήμα 5.122: Μέση διάρκεια βίντεο ανά θεματική κατηγορία βίντεο

Σχετικά με τους διαφορετικούς τύπους βίντεο, προκύπτει από το Σχήμα 5.123, όπως και στο παραπάνω διάγραμμα, πως η πλειονότητα των των βίντεο κυμαίνεται από 10 έως 20 λεπτά. Ελαφρώς μεγαλύτερη διάρκεια φαίνεται να κάλυψαν τα βίντεο *Προκλήσεων* (*Challenges*), ενώ οι *Ζωντανές Ροές* (*Livestreams*) και τα βίντεο *Διαδικτυακής Συζήτησης* (*Podcasts*) διέφεραν του μέσου όρου, με τα πρώτα, τα οποία συγκέντρωσαν αρκετά υψηλό ποσοστό αλληλεπίδρασης κατά το 2020 (βλ. Σχήμα 5.103), να ανέρχονται στη μέση χρονική διάρκεια των 117 λεπτών. Επίσης μακροσκελή παρατηρείται πως ήταν τα βίντεο *Διαγωνισμών* (*Giveaways*), τα οποία δημιουργήθηκαν κατά το 2020. Ωστόσο το συγκεκριμένο αποτέλεσμα είναι επηρεασμένο σε μεγάλο βαθμό από την ανομοιογένεια αυτού του τύπου βίντεο (βλ. Σχήμα 5.113). Τον συντομότερο τύπο βίντεο, κατά μέσο όρο, για το έτος 2020 αποτέλεσαν τα βίντεο τύπου *Πως να* (*How to*), τα οποία σημείωσαν μέση διάρκεια 12 λεπτών.



Σχήμα 5.123: Μέση διάρκεια βίντεο ανά τύπο βίντεο

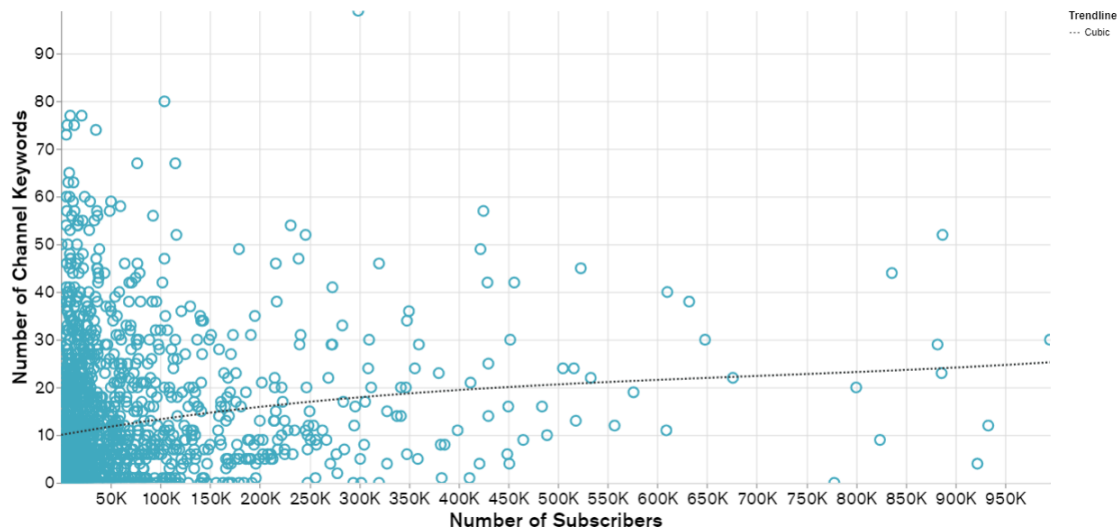
Η σχέση μεταξύ της χρονικής διάρκειας ενός βίντεο και της ανταπόκριση του κοινού σε αυτό, αναφορικά με τον αριθμό προβολών τις οποίες απέσπασε, σκιαγραφήθηκε στο Σχήμα 5.124. Πέραν του συμπεράσματος ότι τα μεγαλύτερης χρονικής διάρκειας βίντεο επιλέχθηκαν από μικρότερο αριθμό δημιουργών κατά το 2020, με τη πλειοψηφία των βίντεο να συναντάται στην κλίμακα των 0-6 λεπτών, διαπιστώνεται από το συγκεκριμένο διάγραμμα πως βίντεο μεγαλύτερης διάρκειας τείνουν να συγκεντρώνουν μικρότερο αριθμό προβολών, σε σχέση με τα συντομότερα. Με μία καλύτερη ματιά στο σχήμα, παρατηρείται πως η προαναφερθείσα διαφορά στον αριθμό προβολών ενός βίντεο αρχίζει να εκδηλώνεται εντονότερα για χρονικές διάρκειες μεγαλύτερες των 18 λεπτών, τιμή η οποία προσέγγισε το άνω όριο της μέσης διάρκειας βίντεο στα δύο προηγούμενα διαγράμματα. Στα πλαίσια του συγκεκριμένου διαγράμματος, συμπεριλήφθηκαν μόνο τα βίντεο με διάρκεια εντός της μίας ώρας και με ποσοστό αλληλεπίδρασης μικρότερο ή ίσο του 20%, έτσι ώστε η προσοχή να εστιαστεί στην πλειοψηφία των βίντεο.



Σχήμα 5.124: Σχέση μεταξύ προβολών και μέση διάρκειας βίντεο

Έχοντας ολοκληρώσει το παραπάνω κομμάτι της έρευνας γύρω από την πλατφόρμα του YouTube, είναι δυνατό να πραγματοποιηθεί η μετάβαση στη διερεύνηση των λέξεων-κλειδιών των ελληνικών καναλιών του δείγματος. Οι λέξεις-κλειδιά ενός καναλιού χρησιμοποιούνται από τους δημιουργούς, με σκοπό να αποκτήσει πληροφορίες η πλατφόρμα για το περιεχόμενο του καναλιού τους και να το προωθήσει, ως αποτέλεσμα, στις αντίστοιχες αναζητήσεις του κοινού [31].

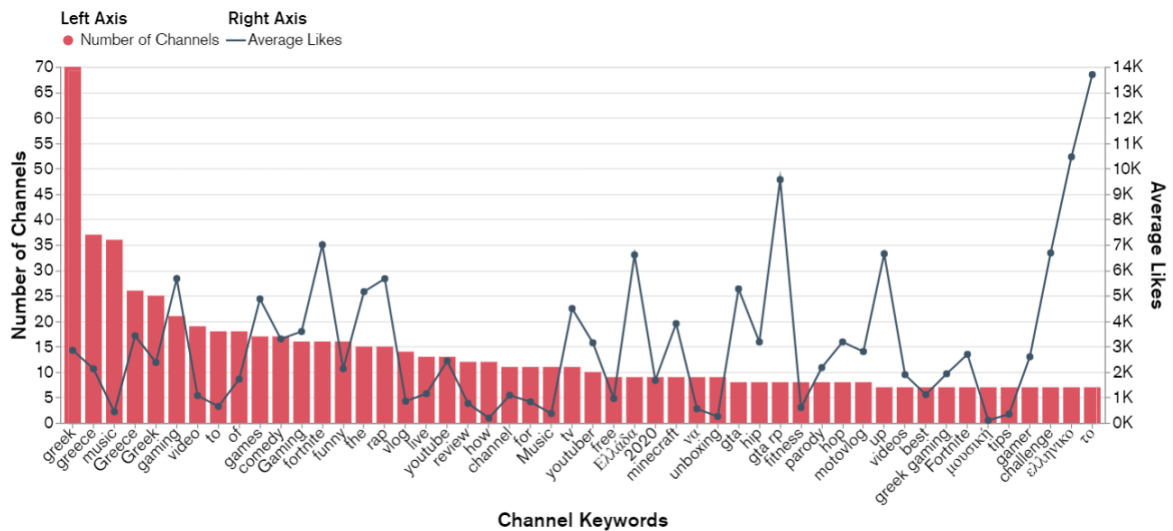
Πρωταρχικό μέλημα στο παρόν στάδιο της έρευνας ήταν η εξέταση του ποσοστού χρήσης λέξεων-κλειδιών από κανάλια ολοκλήρου του φάσματος του αριθμού εγγεγραμμένων χρηστών. Με γνώμονα τα δεδομένα του Σχήματος 5.125, διαπιστώνεται ότι κανάλια με υψηλό αριθμό εγγεγραμμένων χρηστών έχουν την τάση να χρησιμοποιούν περισσότερες λέξεις κλειδιά σε σχέση με τα μικρότερα, κάτι το οποίο επιβεβαιώνει τη σημαντικότητα της χρήσης τους. Παρόλα αυτά, δεν λείπουν οι περιπτώσεις κατά τις οποίες κανάλια μικρότερης εμβέλειας χρησιμοποιούν αρκετά μεγάλο αριθμό λέξεων-κλειδιών, φαινόμενο το οποίο πιθανώς έγκειται στη προσπάθεια ευρύτερης ανάπτυξης τους.



Σχήμα 5.125: Σχέση μεταξύ αριθμού λέξεων-κλειδιών καναλιών και εγγεγραμμένων χρηστών

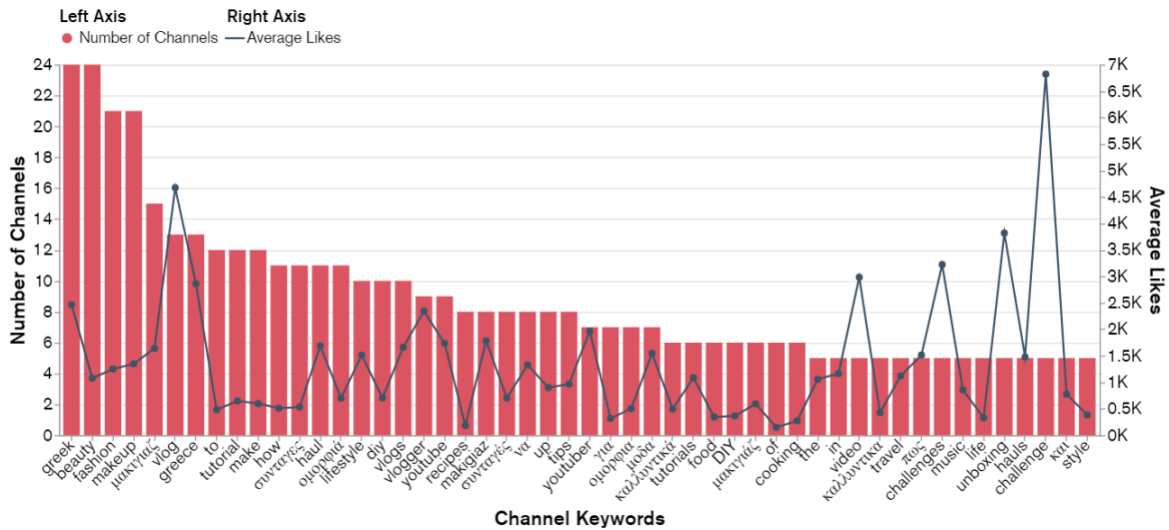
Στη βαθύτερη κατανόηση των τάσεων και των προτιμήσεων του ελληνικού κοινού του YouTube θα μπορούσε να συντελέσει ο εντοπισμός των δημοφιλέστερων λέξεων-κλειδιών στα κανάλια των διαφορετικών κατηγοριών χρηστών. Ως δημοφιλέστερες κρίθηκαν οι λέξεις-κλειδιά, οι οποίες χρησιμοποιήθηκαν από τα περισσότερα κανάλια κατά τη χρονική διάρκεια του έτους 2020. Παράλληλα στο ποσοστό χρήσης της εκάστοτε λέξης-κλειδιού τοποθετήθηκε ο μέσος αριθμός θετικών αντιδράσεων (*Likes*) για κάθε μία από αυτές, με σκοπό την απόκτηση μιας πιο ολοκληρωμένης εικόνας της ανταπόκρισης του κοινού.

Με μια πιο προσεκτική ματιά στις δημοφιλέστερες λέξεις κλειδιά των ανδρικών καναλιών (βλ. Σχήμα 5.126) είναι ορατό πως το μεγαλύτερο μέρος αυτών αφορά στο *Ηλεκτρονικό Παιχνίδι* (*#gaming*, *#fortnite*, *#gta rp*, *#minecraft*), στην *Κωμωδία* (*#comedy*, *#funny*, *#parody*), στη *Μουσική* (*#music*, *#rap*, *#hip*, *#hop*), ενώ αισθητή έκανε την παρουσία της και η *Άθληση* (*Sports*) με τη λέξη-κλειδί *fitness*. Αρκετές ήταν, επίσης, οι λέξεις-κλειδιά οι οποίες αναφέρονταν σε συγκεκριμένους τύπους βίντεο (βλ. Σχήμα 5.103), όπως τα βίντεο τύπου *Πως να* (*#how*, *#to*, *#tips*), τα βίντεο τύπου *Unboxing* (*#unboxing*), οι *Αξιολογήσεις* (*#review*), τα *Vlog* (*#vlog*, *#motovlog*), οι *Ζωντανές ροές* (*#live*) και οι *Προκλήσεις* (*#challenge*). Παρατηρείται, επιπλέον, πως οι μεγαλύτεροι αριθμοί θετικών αντιδράσεων συγκεντρώθηκαν κυρίως από τη θεματική κατηγορία περιεχομένου *Ηλεκτρονικό Παιχνίδι* (*Gaming*), αποδεικνύοντας για πολλοστή φορά την απήχηση της κατά το έτος 2020.



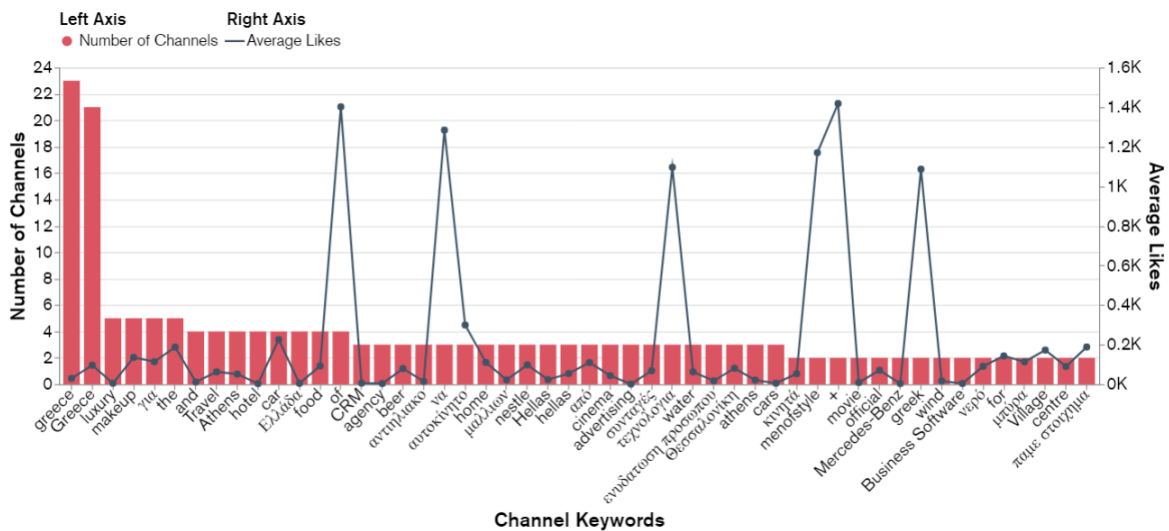
Σχήμα 5.126: Δημοφιλέστερες λέξεις-κλειδιά καναλιών (Άνδρες)

Συνεχίζοντας με τη διαδικασία ανάδειξης των δημοφιλέστερων λέξεων-κλειδιών σε γυναικεία κανάλια (βλ. Σχήμα 5.127), έντονη ήταν η παρουσία της θεματικής κατηγορίας περιεχομένου *Στυλιστικό περιεχόμενο και Ομορφιά* με τις λέξεις-κλειδιά *#beauty*, *#fashion*, *#makeup*, *#μακιγιαζ* και *#ομορφιά* να συναντούνται σε υψηλές θέσεις της κατάταξης. Αισθητή την παρουσία του έκανα, επίσης, οι θεματικές *Μαγειρική* (*#συνταγες*, *#recipes*, *#food*, *#cooking*) και *Ταξιδιωτικές Δραστηριότητες* (*#travel*). Όπως και στην περίπτωση των ανδρικών καναλιών, αρκετές ήταν οι λέξεις-κλειδιά οι οποίες αναφέρονταν σε συγκεκριμένους τύπους βίντεο (βλ. Σχήμα 5.103), όπως τα βίντεο τύπου *Vlog* (*#vlog*, *#vlogs*, *#vlogger*), τα βίντεο τύπου *Πως να* (*#how*, *#to*, *#tips*, *#diy*, *#tutorial*), τα βίντεο τύπου *Haul* (*hauls*), τα βίντεο τύπου *Unboxing* (*#unboxing*) και οι *Προκλήσεις* (*#challenge*, *#challenges*). Τέλος, στο επίπεδο των θετικών αντιδράσεων ιδιαίτερη επιτυχία φαίνεται πως σημείωσαν τα, εντός του 2020, βίντεο γυναικών τα οποία οποία αφορούσαν σε *Προκλήσεις* (*Challenges*), *Vlogs* και *Unboxing*.



Σχήμα 5.127: Δημοφιλέστερες λέξεις-κλειδιά καναλιών (Γυναίκες)

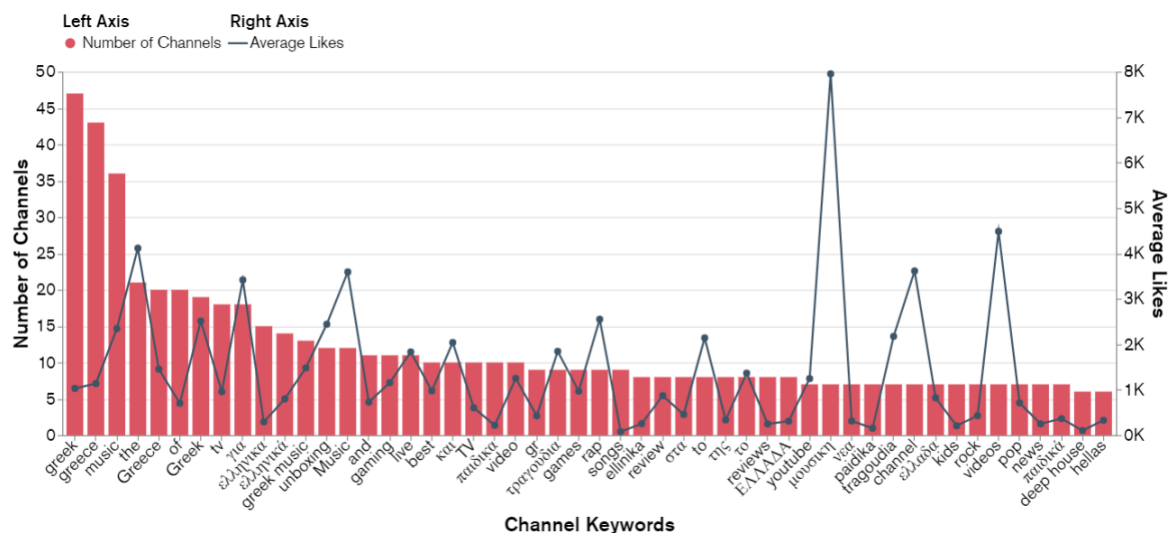
Αναφορικά με το υποσύνολο των επιχειρησιακών καναλιών, το αρκετά χαμηλό ποσοστό χρήσης κοινών λέξεων-κλειδιών δεν επιτρέπει την εξαγωγή ενός αντιπροσωπευτικού γενικού συμπεράσματος για τα κανάλια των επιχειρήσεων, μέσω του Σχήματος 5.128. Παρόλα αυτά, οι λέξεις κλειδιά οι οποίες εμφανίζονται στη κατάταξη αφορούν σε ποικίλες θεματικές κατηγορίες περιεχομένου, όπως Ταξιδιωτικές Δραστηριότητες (#Travel, #hotel), Αυτοκίνηση (#car, #αυτοκίνητο, #Mercedes-Benz), Επιστήμες και Τεχνολογία (#τεχνολογία, #wind, #κινητά) και Στυλιστικό περιεχόμενο(#makeup, #αντηλιακό, #ενυδάτωση προσώπου).



Σχήμα 5.128: Δημοφιλέστερες λέξεις-κλειδιά καναλιών (Επιχειρήσεις)

Τέλος, στο πλαίσιο των δημοφιλέστερων λέξεων-κλειδιών σε κανάλια τύπου Άλλα, είναι φανερή στο Σχήμα 5.129 η γενικότερη τάση των δημιουργών προς τη θεματική κατηγορία της Μουσικής (#Music, #greek music, #τραγουδια, #rap, #rock). Σε υψηλή θέση της κατάταξης εμφα-

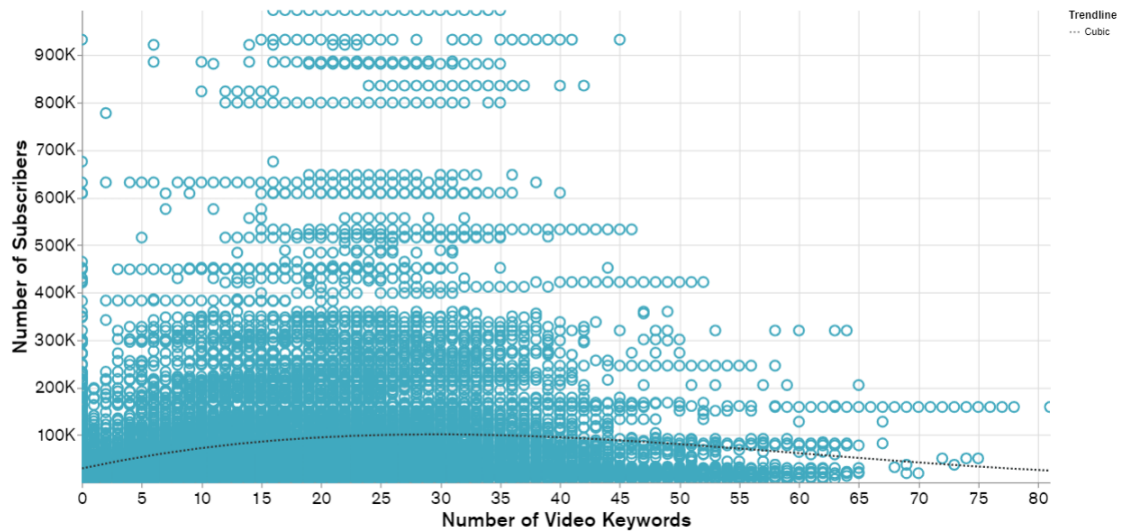
νίζεται, επίσης, το *Ηλεκτρονικό Παιχνίδι* με τις λέξεις-κλειδιά *#gaming*, *#games*, ενώ για πρώτη φορά εντοπίζονται λέξεις-κλειδιά οι οποίες εστιάζουν σε παιδικό περιεχόμενο, όπως *#παιδικα*, *#paidika* και *#kids*. Για ακόμη μία φορά, εμφανίστηκαν στην κατάταξη τα βίντεο τύπου *Unboxing* (*#unboxing*) και οι *Αξιολογήσεις* (*#review*), επαληθεύοντας τη δημοτικότητα τέτοιου είδους περιεχομένου κατά το έτος 2020.



Σχήμα 5.129: Δημοφιλέστερες λέξεις-κλειδιά καναλιών (Άλλο)

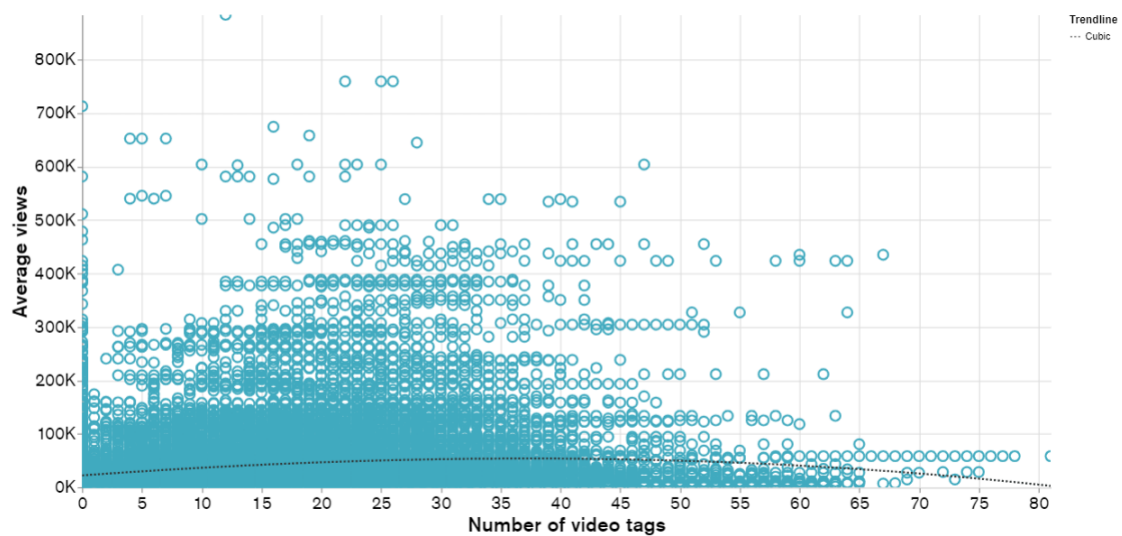
Όπως γίνεται αντιληπτό από την παραπάνω ομάδα διαγραμμάτων, η αξιοποίηση των λέξεων-κλειδιών δεν ήταν αρκετή για τον σχηματισμό μιας πλήρως αντιπροσωπευτικής εικόνας για τις τάσεις και τις προτιμήσεις του ελληνικού κοινού, λόγω της μικρότερης συνάφειας μεταξύ των λέξεων-κλειδιών, τις οποίες ενσωματώνουν τα διαφορετικά κανάλια του δείγματος. Στην αντιμετώπιση του προβλήματος αυτού κλήθηκαν να συνδράμουν οι λέξεις-κλειδιά των βίντεο του δείγματος. Οι λέξεις-κλειδιά προσαρτώνται από τους δημιουργούς στα βίντεο τους, έτσι ώστε να τα βοηθήσουν να αναδειχθούν στα αποτελέσματα των αντίστοιχων αναζητήσεων του κοινού [1]. Η μετάβαση της έρευνας από το επίπεδο των καναλιών στο επίπεδο των βίντεο, μέσω της χρήσης των λέξεων-κλειδιών βίντεο, θα ήταν δυνατό να ενισχύσει ή ακόμα και να επεκτείνει τα προηγθέντα συμπεράσματα με περισσότερη πληροφορία.

Ως πρώτα βήμα της μελέτης των λέξεων-κλειδιών βίντεο, τέθηκε η εξιχνίαση της σχέσης την οποία διατηρούν κανάλια διαφορετικής εμβέλειας με τη χρήση αυτών. Στο Σχήμα 5.130 διακρίνεται η τάση των καναλιών με μεγαλύτερο αριθμό εγγεγραμμένων χρηστών να χρησιμοποιούν περισσότερες λέξεις-κλειδιά στα βίντεο τους, με το προτιμότερο πλήθος λέξεων-κλειδιών για τα κανάλια αυτά να κυμαίνεται μεταξύ των 25 και των 35. Η ευρεία εκμετάλλευση των λέξεων-κλειδιών των βίντεο από κανάλια μεγαλύτερης εμβέλειας αποδεικνύει τη σημαντικότητα της χρήσης τους. Ωστόσο, η ενσωμάτωση αρκετά εκτεταμένου αριθμού λέξεων-κλειδιών στα βίντεο δημιουργών φαίνεται να μη προτιμάται ιδιαίτερα.



Σχήμα 5.130: Σχέση μεταξύ αριθμού λέξεων-κλειδιών βίντεο και αριθμού εγγεγραμμένων χρηστών

Με σκοπό την εξέταση της αποδοτικότητας της χρήσης των λέξεων-κλειδιών σε βίντεο του YouTube, αναζητήθηκε η σχέση μεταξύ του πλήθους των λέξεων-κλειδιών σε ένα βίντεο και των προβολών τις οποίες αυτό απέσπασε. Σημειώνεται πως, για τη καλύτερη οπτική αναπαράσταση της προκείμενης σχέσης και προς αποφυγή των ακραίων περιπτώσεων, λήφθηκαν υπόψη μόνο τα βίντεο του 2020, τα οποία δε ξεπέρασαν τις 1.000.000 προβολές. Με βάση το Σχήμα 5.131, διαπιστώνεται πως βίντεο στα οποία ενσωματώνονται περισσότερες λέξεις-κλειδιά τείνουν να συγκεντρώνουν υψηλότερο αριθμό προβολών, με τη βέλτιστη απόδοση να επιτυγχάνεται στο διάστημα 30 έως 40 λέξεων-κλειδιών. Ωστόσο, παρατηρείται πως η αφαιμάξη των λέξεων-κλειδιών δεν ανταποκρίνεται στις προσδοκίες των δημιουργών για υψηλότερο αριθμό προβολών στα βίντεο τους, με την πτώση αυτού προς τα μεγαλύτερα σύνολα λέξεων-κλειδιών, πιθανώς, να μαρτυρά την εκτεταμένη χρήση τους από κανάλια μικρότερης απόδοσης.

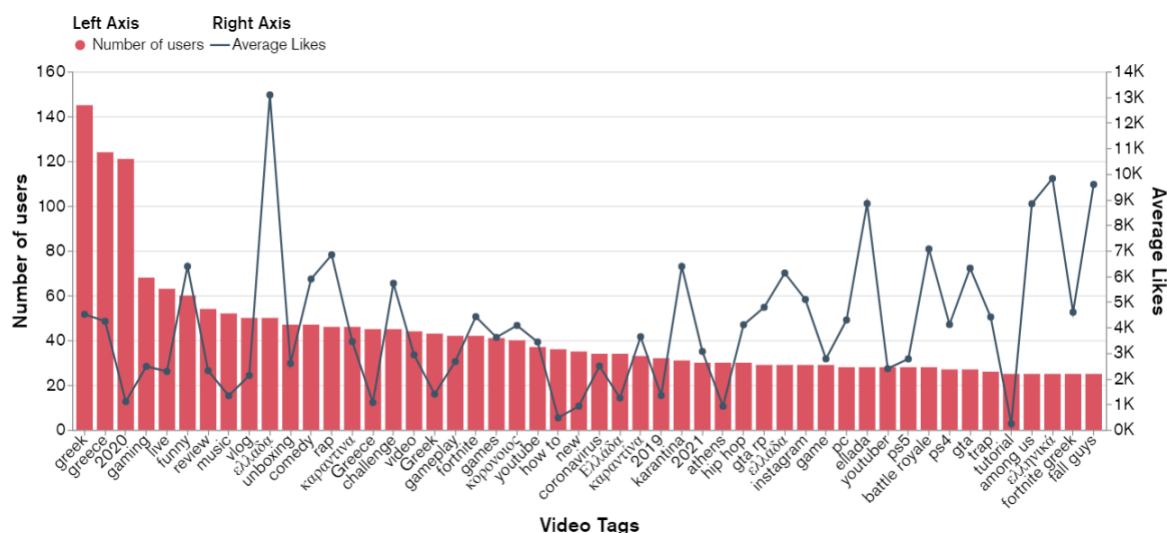


Σχήμα 5.131: Σχέση μεταξύ αριθμού λέξεων-κλειδιών βίντεο και αριθμού προβολών

Η ακριβέστερη αποκωδικοποίηση των ενδιαφερόντων των Ελλήνων δημιουργών και χρηστών του YouTube πραγματοποιήθηκε μέσω του εντοπισμού των δημοφιλέστερων λέξεων-κλειδιών σε βίντεο του 2020, ανά τις διαφορετικές κατηγορίες χρηστών. Λόγω του γεγονότος ότι αρκετές λέξεις-κλειδιά χρησιμοποιήθηκαν σε μεγάλο βαθμό από μεμονωμένα κανάλια, η κατάταξη των λέξεων-κλειδιών, με βάση το πλήθος των βίντεο στα οποία αυτά εμφανίστηκαν, επηρεάστηκε σε μεγάλο βαθμό από συγκεκριμένους δημιουργούς. Προς αποφυγή αυτού, η ανάδειξη των δημοφιλέστερων λέξεων-κλειδιών βίντεο επιτεύχθηκε στη βάση αυτών, τα οποία εμφανίστηκαν στα περισσότερα κανάλια του δείγματος. Επισημαίνεται στο σημείο αυτό, πως όπως γίνεται αντιληπτό από τα παρακάτω διαγράμματα, η χρήση κοινών λέξεων-κλειδιών βίντεο από διαφορετικά κανάλια ήταν και πάλι περιορισμένη, αλλά ήταν αρκετή για να προσδώσει έναν περισσότερο αντιπροσωπευτικό χαρακτήρα στα συμπεράσματα της παρούσας έρευνας

Η έναρξη της διερεύνησης των δημοφιλέστερων λέξεων-κλειδιών έλαβε χώρα στα πλαίσια των ανδρικών καναλιών του YouTube. Ύστερα από την παρατήρηση τους Σχήματος 5.132, διαπιστώνεται πως μια μεγάλη μερίδα των ανδρών δημιουργών, κατά το 2020, προσανατόλισε το περιεχόμενο της στο Ηλεκτρονικό Παιχνίδι με λέξεις κλειδιά, όπως #gaming, #gameplay, #fortnite και #gta rp να κατακλύζουν την κατάταξη του διαγράμματος. Επιπλέον, από την κατάταξη δεν έλειψαν η Κωμωδία (#funny, #comedy) και η Μουσική (#music, #rap, #trap), υποδηλώνοντας έτσι το επίσης ανεπτυγμένο ενδιαφέρον της συγκεκριμένης κατηγορίας χρηστών προς αυτές τις θεματικές κατηγορίες περιεχομένου. Πέραν των παραπάνω, την εμφάνιση τους έκαναν αρκετοί τύποι βίντεο, όπως η Ζωντανή Ροή (#live), η Αξιολόγηση (#review), οι Προκλήσεις (#challenge) και τα βίντεο τύπου Unboxing (#unboxing), Vlog (#vlog) και Πως να (#how to, #tutorial), υπογραμμίζοντας τη συμπάθεια των δημιουργών της πλατφόρμας προς αυτού του είδους τα βίντεο.

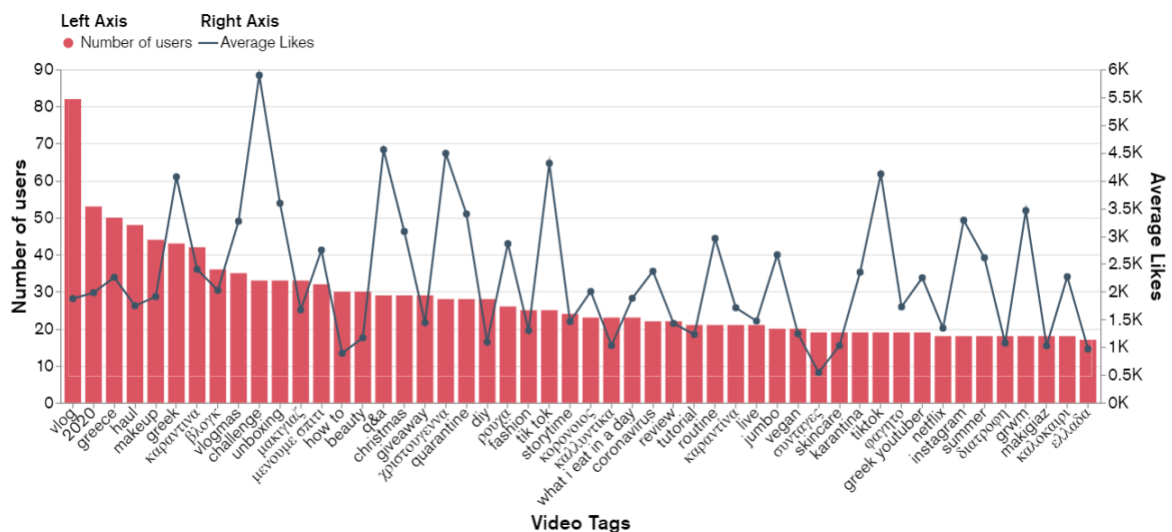
Όσον αφορά στο κοινό, είναι ορατό πως αυτό ανταποκρίθηκε στις προτιμήσεις των δημιουργών, καθώς υψηλά ποσοστά θετικών αντιδράσεων συγκεντρώθηκαν από όλες τις προαναφερθείσες θεματικές κατηγορίες περιεχομένου, με το ηλεκτρονικά παιχνίδια *gta rp*, *fortnite*, *among us* και *fall guys*, τη ραπ μουσική, τα βίντεο προκλήσεων και τα βίντεο κωμικού περιεχομένου να διακρίνονται στον άξονα της θετικής ανταπόκρισης του κοινού για το έτος 2020.



Σχήμα 5.132: Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Άνδρες)

Συνεχίζοντας στα γυναικεία κανάλια, με μια γενική ματιά στις δημοφιλέστερες λέξεις-κλειδιά (βλ. Σχήμα 5.133), των εντός του 2020 βίντεο τους, διακρίνεται η ιδιαίτερη συμπάθεια των γυναικών δημιουργών προς τη μόδα, τον καλλωπισμό και την προβολή στιγμιοτύπων από την προσωπική τους καθημερινότητα. Ειδικότερα, στο επίπεδο των θεματικών κατηγοριών περιεχομένου αρκετά έντονη στην κατάταξη είναι η παρουσία του *Στιλιστικού περιεχομένου* και της ομορφιάς, μέσω της χρήσης λέξεων-κλειδιών, όπως *#makeup*, *#beauty*, *#ρουχα*, *#fashion* και *#skincare*, ενώ, επίσης, εντοπίζεται ένας μικρότερος αριθμός λέξεων-κλειδιών ο οποίος απευθύνεται σε βίντεο διατροφής (*#vegan*, *#φαγητο*, *#συνταγες*, *#διατροφή*). Η ανάδειξη της λέξης-κλειδιού *#vlog* ως η δημοφιλέστερη μεταξύ των γυναικών δημιουργών, σε συνδυασμό με την εμφάνιση ετικετών, όπως *#βλογκ*, *#vlogmas*, *#what i eat in a day* και *#routine*, μαρτυρά την τάση των γυναικών να μοιράζονται καθημερινές στιγμές από την ζωή τους με το κοινό τους. Ακόμη, αξιοσημείωτη είναι η παρουσία αρκετών λέξεων-κλειδιών οι οποίες αφορούν σε βίντεο που περιλαμβάνουν τοποθέτηση προϊόντων, όπως *#haul*, *#unboxing*, *#review* και *#giveaway*, καθώς και η σχετικά υψηλή προτίμηση των γυναικών για βίντεο Προκλήσεων (*#challenge*), Ερωτωαπαντήσεων (*#q&a*), τύπου Πως να (*#how to*, *#tutorial*) και Ζωντανής Ροής (*#live*). Επιπρόσθετα, η εμφάνιση των ετικετών *#vlogmas*, *#christmas*, *#χριστουγεννα*, *#summer* και *#καλοκαίρι* υποδεικνύουν το ενδεχόμενο ποσοστό επιρροής των γυναικών δημιουργών από τις συγκεκριμένες χρονικές περιόδους του έτους.

Αναφορικά με την ανταπόκριση του ελληνικού κοινού στο περιεχόμενο των γυναικείων καναλιών, παρατηρείται πως κατά μέσο όρο οι περισσότερες θετικές αντιδράσεις συγκεντρώθηκαν από βίντεο προκλήσεων, ερωτωαπαντήσεων, προβολής της προσωπικής καθημερινότητας και από Χριστουγεννιάτικο περιεχόμενο. Τέλος, ενδιαφέρουσα είναι, επίσης, η συγκέντρωση ενός εξίσου υψηλού αριθμού θετικών αντιδράσεων από βίντεο σχετικά με τις πλατφόρμες *TikTok* και *Instagram*.

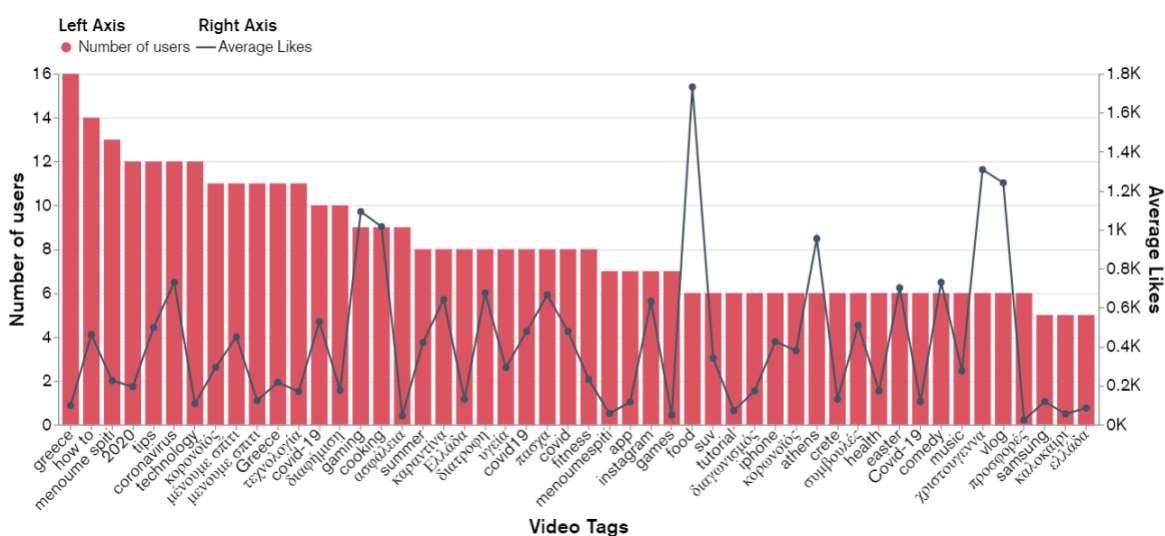


Σχήμα 5.133: Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Γυναίκες)

Από την πλευρά των επιχειρησιακών καναλιών, προκύπτει με βάση τις δημοφιλέστερες λέξεις-κλειδιά του Σχήματος 5.134 ότι δραστηριοποιήθηκαν κατά το 2020 σε ποικίλες θεματικές κατηγορίες περιεχομένου, με την *Τεχνολογία και τις Επιστήμες* (*#technology*, *#τεχνολογία*) να συνα-

ντούνται σε μεγαλύτερο βαθμό από τις υπόλοιπες κατηγορίες. Ετικέτες όπως *#gaming*, *#cooking*, *#υγεία*, *#fitness*, *#comedy* και *#music* αποδεικνύουν την ποικιλομορφία η οποία διέπει τον χαρακτήρα των επιχειρήσεων στο YouTube. Σχετικά με τους διαφορετικούς τύπους βίντεο, οι επιχειρήσεις εμφάνισαν μια κλίση προς τα βίντεο τύπου *Πως να*, δεδομένου των ετικετών *#how to*, *#tips*, *#tutorial* και *#συμβουλές*, ενώ στην κατάταξη εμφανίστηκαν επίσης, σε μικρότερο βαθμό, βίντεο *Διαγωνισμών* (*#διαγωνισμός*) και βίντεο τύπου *Blog* (*#vlog*). Τέλος, ο εντοπισμός των λέξεων-κλειδιών *#πασχα*, *#easter*, *#καλοκαίρι*, *#summer* και *#χριστουγεννα* σε βίντεο επιχειρήσεων επαναφέρει στο προσκήνιο την τακτική αυτών για προώθηση των προϊόντων και των υπηρεσιών τους κατά τη διάρκεια περιόδων υψηλής καταναλωτικής δραστηριότητας.

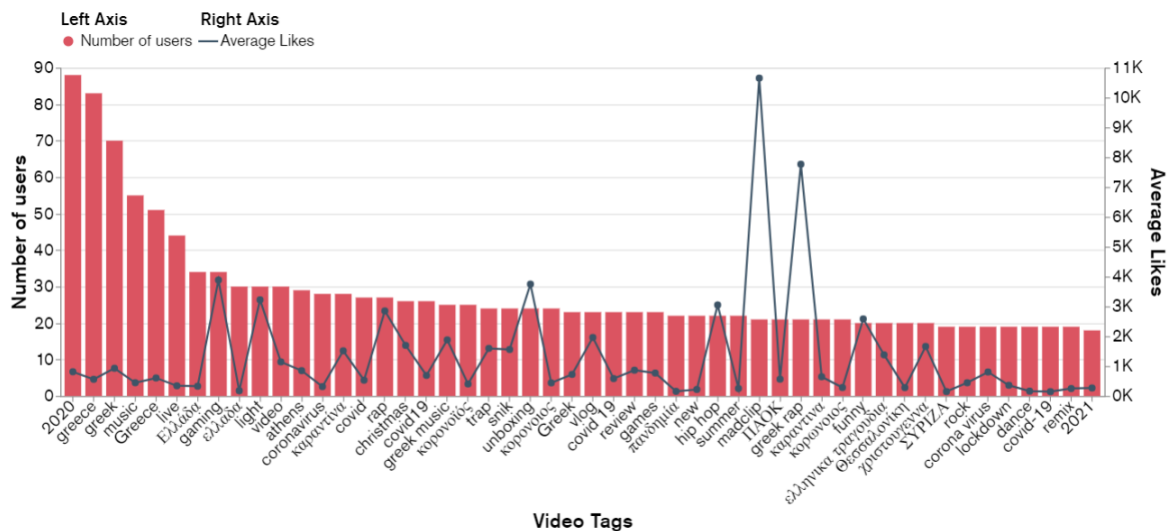
Η θετική ανταπόκριση του κοινού στα δημοσιευμένα, εντός του 2020, βίντεο των επιχειρήσεων, εκφράστηκε κατά το πλείστον σε βίντεο ηλεκτρονικού παιχνιδιού, μαγειρικής, προβολής προσωπικής καθημερινότητας και σε Χριστουγεννιάτικο περιεχόμενο.



Σχήμα 5.134: Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Επιχειρήσεις)

Τελευταία εξετάστηκε η κατηγορία χρηστών Άλλο, η οποία, σύμφωνα με το Σχήμα 5.135, εστίασε κατά έναν μεγάλο βαθμό στη *Μουσική*. Η συγκεκριμένη παρατήρηση είναι εύκολο να δικαιολογηθεί από την πληθώρα των ετικετών, οι οποίες επίσης σχετίζονται με τη μουσική, μεταξύ των οποίων βρίσκονται οι λέξεις *#music*, *#rap*, *#hip hop* και *#ελληνικά τραγούδια*. Ανάμεσα στις δημοφιλέστερες λέξεις-κλειδιά έκαναν επιπλέον την εμφάνιση τους οι ετικέτες *#gaming* και *#funny*, υποδηλώνοντας την παρουσία βίντεο *Ηλεκτρονικού Παιχνιδιού* (*Gaming*) και *Κωμωδίας* (*Comedy*), στα πλαίσια των καναλιών τύπου Άλλο. Για ακόμη μια φορά στην κατάταξη εντοπίστηκαν οι λέξεις-κλειδιά *#live*, *#unboxing*, *#vlog* και *#review*, γεγονός το οποίο, σε συνδυασμό με τα παραπάνω αποτελέσματα, επιβεβαιώνει την υψηλή δημοτικότητα αυτού του τύπου βίντεο στην ελληνική κοινότητα του YouTube.

Από τη μεριά του κοινού, διαπιστώνεται πως το υψηλότερο ενδιαφέρον για το έτος 2020 απέσπασε η ελληνική ραπ μουσική, με ιδιαίτερες περιπτώσεις να αποτελούν οι καλλιτέχνες *madclip* και *light*, το ηλεκτρονικό παιχνίδι, καθώς και τα βίντεο τύπου *Unboxing*.

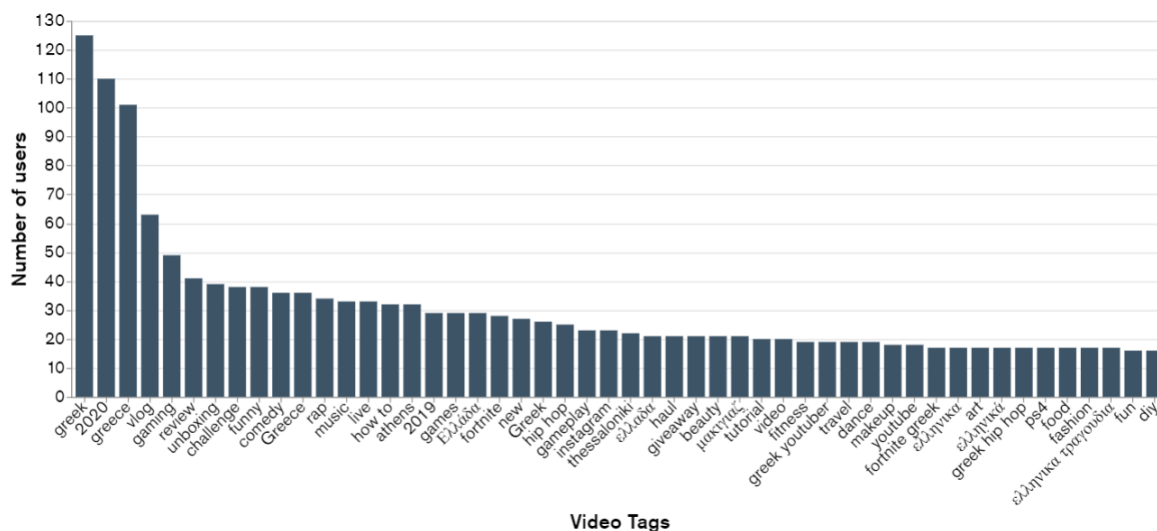


Σχήμα 5.135: Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Άλλο)

Μελετώντας τα παραπάνω διαγράμματα, δε θα μπορούσε να παραληφθεί το κοινό σημείο αναφοράς τους, το οποίο ήταν η επιρροή όλων των κατηγοριών χρηστών από την πανδημία του COVID-19. Ο εντοπισμός μιας πληθώρας λέξεων-κλειδιών σχετιζόμενων με την πανδημία, όπως *#καραντίνα*, *#κορονοϊός*, *#quarantine*, *#coronavirus* και *#μενουμε spiti*, αποδεικνύει πως η δραστηριότητα των Ελλήνων δημιουργών για το έτος 2020 διαμορφώθηκε κατά έναν σημαντικό βαθμό από τις συνθήκες εγκλεισμού, οι οποίες ακολούθησαν την εξάπλωση του ιού. Πιο συγκεκριμένα, παρατηρήθηκε πως το περιεχόμενο των επιχειρήσεων και των καναλιών τύπου Άλλο επηρεάστηκε σε μεγαλύτερο βαθμό από αυτό των υπόλοιπων κατηγοριών χρηστών, καθώς η ευαισθητοποίηση και η κινητοποίηση τους, κατά την περίοδο της καραντίνας, έγινε φανερή μέσω της εκτεταμένης χρήσης ανάλογων ετικετών στα βίντεο τους.

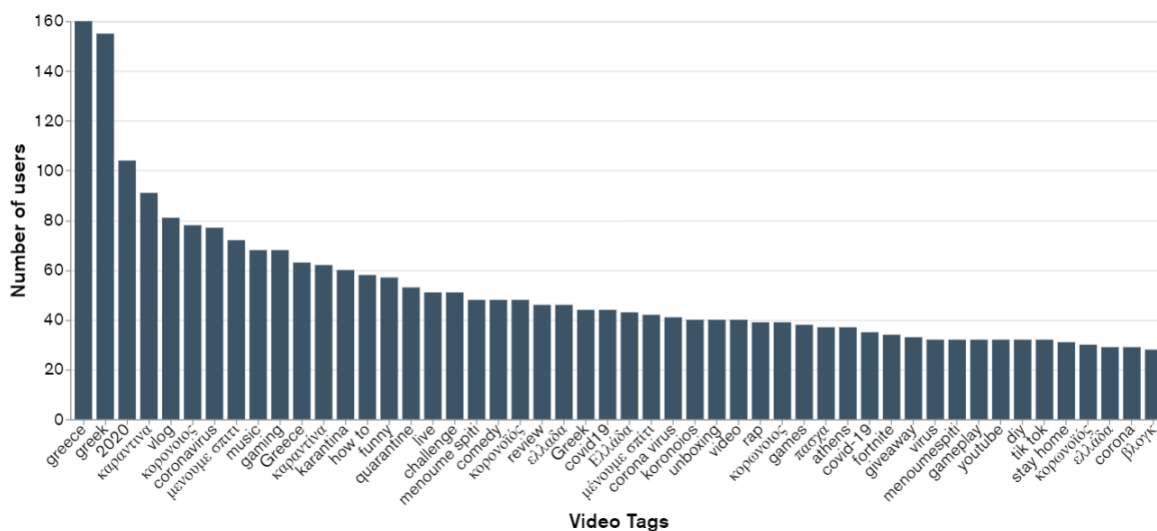
Η μεταβολή της δραστηριότητας των Ελλήνων δημιουργών, εξαιτίας του ιού COVID-19, αλλά και η γενικότερη η εξέλιξη των ενδιαφερόντων τους, είναι δυνατό να μελετηθούν μέσω της εποπτείας των λέξεων-κλειδιών των βίντεο, οι οποίες χρησιμοποιήθηκαν από τα περισσότερα κανάλια, ανά τις διαδοχικές χρονικές περιόδους του 2020. Η διάκριση των χρονικών περιόδων στηρίχθηκε, κατά κύριο λόγο, στο μοτίβο των εποχών του έτους, παρόλα αυτά, λόγω του γεγονότος ότι ο Δεκέμβριος δεν ήταν ορθό να συμπεριληφθεί στο διάγραμμα των μηνών Ιανουάριος-Φεβρουάριος, εξαιτίας του χρονικού χάσματος μεταξύ τους, σε συνδυασμό με τη σταδιακή αύξηση της δραστηριότητας, η οποία παρατηρήθηκε κατά τους μήνες Νοέμβριο και Δεκέμβριο (βλ. Σχήμα 5.110), αποφασίστηκε οι δύο τελευταίοι μήνες του έτους να αποσπαστούν σε ξεχωριστό διάγραμμα.

Η έναρξη του έτους 2020 συνοδεύτηκε από την ευρεία χρήση της λέξης-κλειδιού *#2020*, με τους δημιουργούς να υποδέχονται κατά αυτόν τον τρόπο το νέο έτος. Με βάση το Σχήμα 5.136, γίνεται αντιληπτό πως τα ενδιαφέροντα των Ελλήνων δημιουργών αποτέλεσαν ένα συνονθύλευμα των Σχημάτων 5.132, 5.133, 5.134 και 5.135, με το Ηλεκτρονικό Παιχνίδι (*Gaming*), την Κωμωδία (*Comedy*), τη Μουσική (*Music*) και το Στιλιστικό περιεχόμενο (*Howto & Style*), άλλα και του τύπου βίντεο *Vlog*, *Αξιολόγηση*, *Ζωντανή Ροή*, *Διαγωνισμό*, *Haul* και *Πως να*, να αποτελούν το επίκεντρο της προσοχής για τους μήνες Ιανουάριο και Φεβρουάριο.



Σχήμα 5.136: Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Ιανουάριος-Φεβρουάριος)

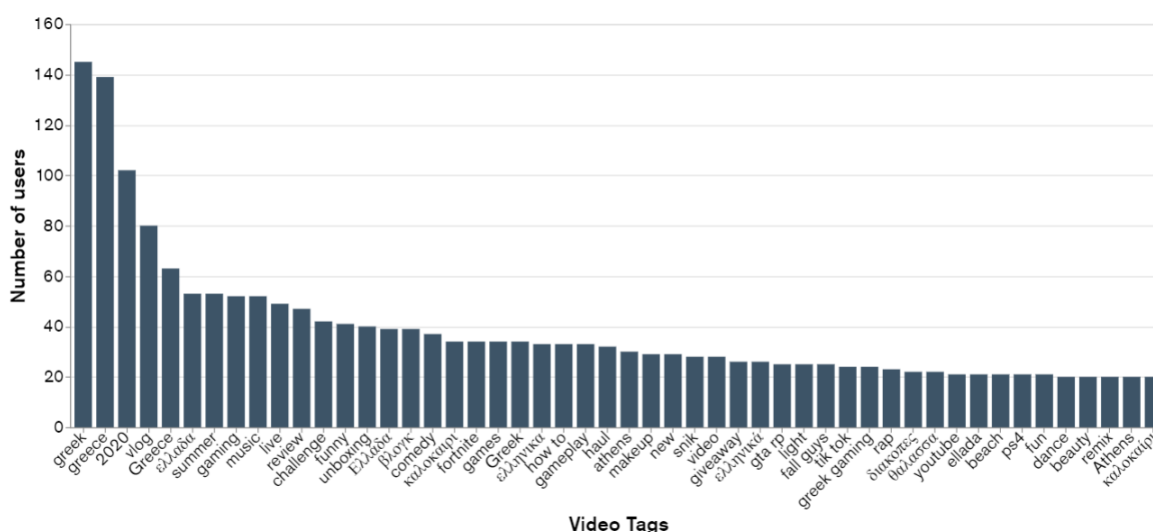
Η περίοδος μεταξύ Μαρτίου και Μαΐου στιγματίστηκε από την έξαρση της πανδημίας του COVID-19 και την επιβολή περιοριστικών μέτρων, γεγονός το οποίο γίνεται ξεκάθαρο από τη κατάληψη της συντριπτικής πλειοψηφίας των θέσεων της κατάταξης από ετικέτες αναφερόμενες στον ιό, όπως #καραντινα, #κορονοιος, #coronavirus και #μενουμε σπιτι. Ωστόσο, για ακόμη μια φορά δεν απείχαν από την κατάταξη οι συνηθείς πόλοι ενδιαφέροντος, οι οποίοι παρουσιάστηκαν και στα πλαίσια της χειμερινής περιόδου (βλ. Σχήμα 5.136). Τέλος, η ανταπόκριση των δημιουργών σε εορταστικό περιεχόμενο έγινε ξανά αντιληπτή κατά τη περίοδο της άνοιξης του 2020 με τη χαρακτηριστική αναφορά στην εορτή του Πάσχα (#πασχα).



Σχήμα 5.137: Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Μάρτιος-Μαΐος)

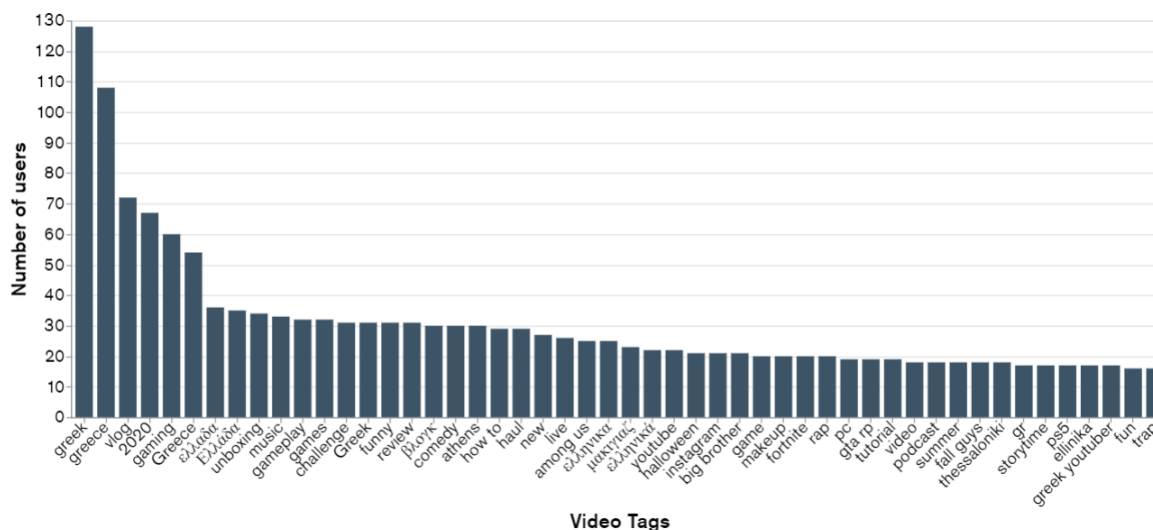
Ο ερχομός της καλοκαιρινής περιόδου σε συνδυασμό με την άρση των περιοριστικών μέτρων

της πανδημίας επέφερε την επιστροφή του Ελλήνων δημιουργών στη κανονικότητα, κάτι το οποίο επιβεβαιώθηκε και από τη ψηφιακή τους συμπεριφορά στο YouTube κατά τους καλοκαιρινούς μήνες (βλ. Σχήμα 5.117). Πιο συγκεκριμένα, η αισθητή παρουσία λέξεων-κλειδιών οι οποίες αφορούν σε καλοκαιρινές δραστηριότητες, όπως *#summer*, *#καλοκαιρι*, *#διακοπες* και *#θαλασσα*, σε συνδυασμό με τη μειωμένη δραστηριότητα των δημιουργών και, συνάμα, με το αυξημένο ποσοστό αλληλεπίδρασης του κοινού, επαληθεύει ότι το καλοκαίρι του 2020 αποτέλεσε για τους Έλληνες χρήστες του YouTube μια περίοδο χαλαρότητας και αυξημένης κίνησης στη πλατφόρμα. Η επιστροφή στην κανονικότητα δε θα μπορούσε να μη συνοδευτεί, επίσης, από την δυναμική επάνοδο των προαναφερθέντων δημοφιλών θεματικών κατηγοριών περιεχομένου και τύπων βίντεο στις θέσεις της κατάταξης



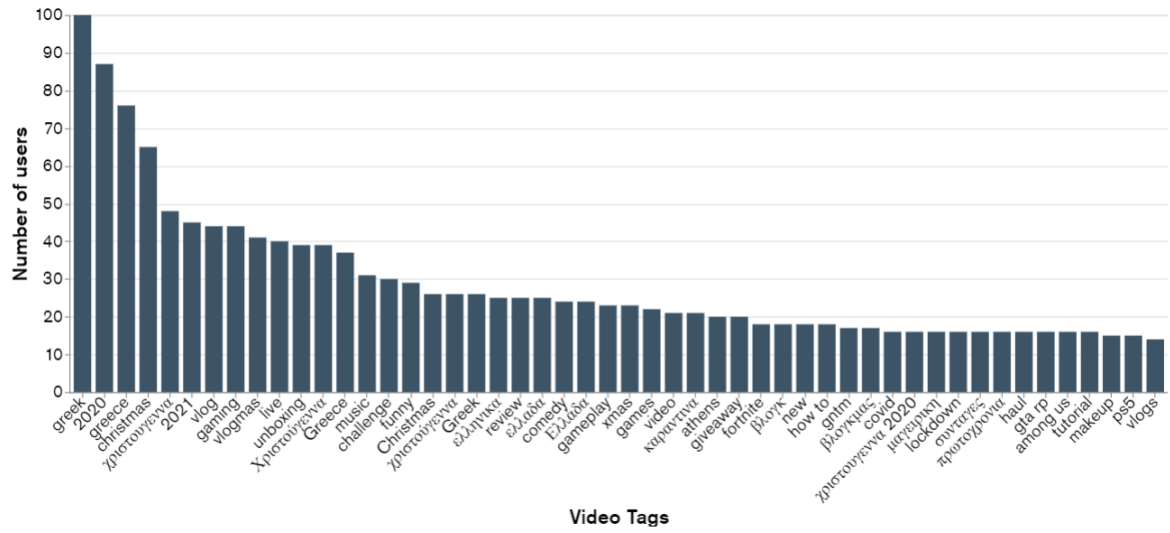
Σχήμα 5.138: Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Ιούνιος-Αύγουστος)

Εφόσον ολοκληρώθηκε η καλοκαιρινή περίοδος και ο καθολικός περιορισμός των μετακινήσεων, λόγω της πανδημίας, δεν τέθηκε ακόμη σε ισχύ, η μετάβαση στο φθινόπωρο ακολούθηθηκε από τη σταδιακή επιστροφή των Ελλήνων δημιουργών στο προ-καλοκαιριού μοντέλο ψηφιακής συμπεριφοράς, ως προς τις θεματικές κατηγορίες και τύπους βίντεο, με λέξεις-κλειδιά αναφερόμενες στο καλοκαίρι να φθίνουν σημαντικά. Αξιοσημείωτη η εμφάνιση των βίντεο *Διαδικτυακή Συζήτηση (Podcast)* μεταξύ των κορυφαίων ετικετών της συγκεκριμένης περιόδου, γεγονός το οποίο σε συνδυασμό με το υψηλό τους ποσοστό προβολών-εγγεγραμμένων χρηστών (βλ. Σχήμα 5.106), υποδηλώνει την ανάπτυξη και τις προοπτικές περαιτέρω εξέλιξης αυτού του τύπου βίντεο στην ελληνική κοινότητα της πλατφόρμας. Τέλος, η παρουσία της λέξης-κλειδιού *#halloween* στρέφει ξανά τη προσοχή στην κλίση των δημιουργών προς το εορταστικό/επειτακό περιεχόμενο



Σχήμα 5.139: Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Σεπτέμβριος-Οκτώβριος)

Καταλήγοντας στα τέλη του 2020 και συγκεκριμένα στους μήνες Νοέμβριο και Δεκέμβριο, αισθητό αντίκτυπο στη δραστηριότητα των χρηστών είχε η επαναφορά των περιοριστικών μέτρων για τη καταπολέμηση του ιού *COVID-19*, κάτι το οποίο καθίστανται σαφές μέσω της χρήσης ετικετών σχετικών με την πανδημία, όπως *#καραντινα* και *#lockdown* κατά τη συγκεκριμένη χρονική περίοδο (βλ. Σχήμα 5.140). Ωστόσο, συγκρίνοντας το ποσοστό χρήσης τέτοιου είδους λέξεων-κλειδιών με αυτό της περιόδου Μαρτίου-Μαΐου (βλ. Σχήμα 5.137), διάστημα στο οποίο εφαρμόστηκαν τα πρώτα περιοριστικά μέτρα, παρατηρείται πως το αντίκτυπο των περιορισμών στη ψυχολογία των Ελλήνων δημιουργών ήταν αρκετά ασθενέστερο κατά το δεύτερο κύμα της πανδημίας. Εγκαταλείποντας το κλάδο του ιού *COVID-19*, πέραν των συνηθέστερων θεματικών κατηγοριών και τύπων βίντεο, ιδιαίτερα αισθητό γίνεται στο παρακάτω διάγραμμα το εορταστικό κλίμα, το οποίο επικράτησε κατά τη Χριστουγεννιάτικη περίοδο του 2020 (*#christmas*, *#χριστουγεννα*, *#vlogmas*, *#xmas*), αναδεικνύοντας έτσι το έντονο ενδιαφέρον της ελληνικής κοινότητας του YouTube για τα Χριστούγεννα. Τέλος, αξίζει να παρατηρηθεί, πως η συγκεκριμένη περίοδος του έτους είναι στενά συνδεδεμένη με την αυξημένη παραγωγή τύπου *Vlog*, ενώ η γενικότερη κλίση των δημιουργών προς αυτόν τον τύπο βίντεο, καθ' όλη τη διάρκεια του έτους, αποδεικνύει τη συμπάθεια των δημιουργών και του κοινού της πλατφόρμας για κοινοποίηση προσωπικών και καθημερινών στιγμιotypών (βλ. επίσης Σχήμα 5.103).



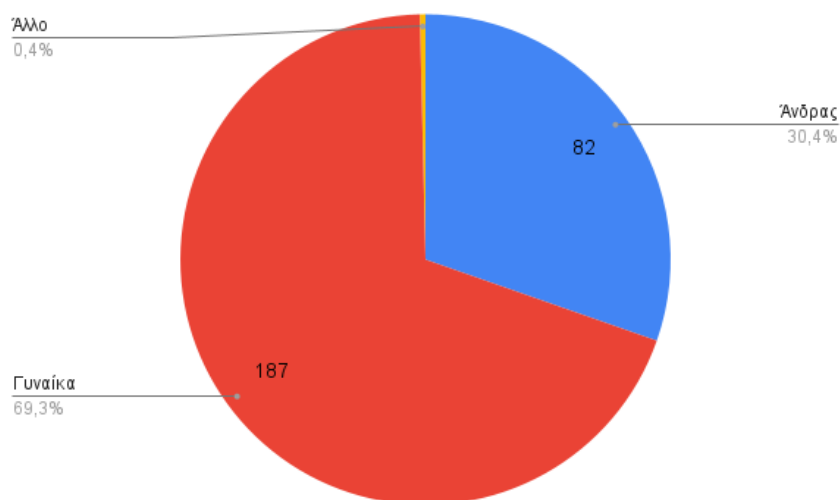
Σχήμα 5.140: Δημοφιλέστερες λέξεις-κλειδιά βίντεο (Νοέμβριος-Δεκέμβριος)

Κεφάλαιο 6

Μελέτη ερωτηματολογίου

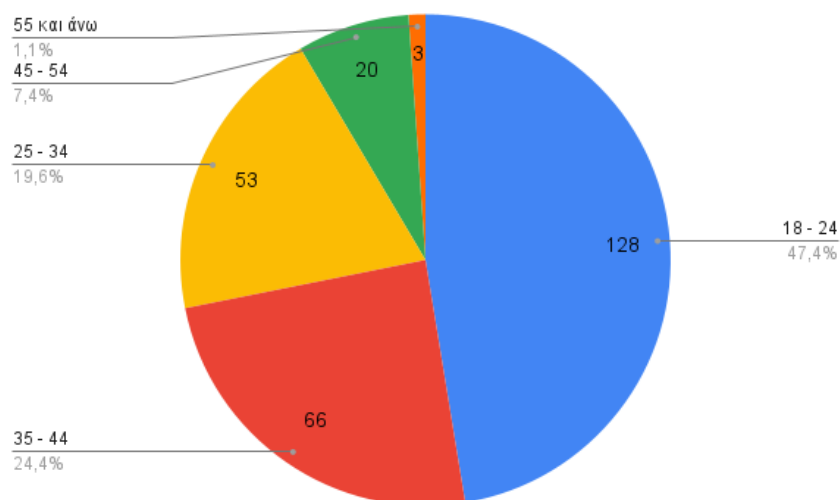
Όπως προαναφέρθηκε στα αρχικά στάδια της παρούσας έρευνας, η ανάγκη για προστασία των προσωπικών δεδομένων των χρηστών εκ μέρους των πλατφορμών Instagram και YouTube, καθώς και ο σχετικά περιορισμένος όγκος των διαθέσιμων προς συγκομιδή πληροφοριών, απέτρεψαν τη συγκέντρωση αρκετών στοιχείων, τα οποία θα μπορούσαν να προσδώσουν μεγαλύτερη πληρότητα στα συμπεράσματα της μελέτης, όσον αφορά στη συμπεριφορά των Ελλήνων χρηστών των μέσων κοινωνικής δικτύωσης. Με σκοπό την αντιμετώπιση των συγκεκριμένου προβλήματος, αποφασίστηκε η δημιουργία και ο διαμοιρασμός ενός ερωτηματολογίου, το οποίο καλούνταν να καλύψει τα παραπάνω ερευνητικά κενά.

Το ερωτηματολόγιο απευθύνθηκε σε Έλληνες χρήστες των μέσων κοινωνικής δικτύωσης Instagram και YouTube, οι οποίοι κατατάσσονταν σε ολόκληρο το ηλικιακό φάσμα και προέρχονταν από ποικίλες κοινωνικές ομάδες. Συγκεκριμένα, το ερωτηματολόγιο απαντήθηκε από 270 άτομα άνω των 18 ετών, από τα οποία 187 ήταν γυναίκες (69,3%), 82 ήταν άνδρες (30,4%), ενώ ένα άτομο κατέταξε τον εαυτό στην κατηγορία Άλλο (βλ. Σχήμα 6.1). Επισημαίνεται πως τα αποτελέσματα του ερωτηματολογίου, πιθανώς, αποκλίνουν της πραγματικότητας, είτε λόγω των περιορισμένων απαντήσεων ή λόγω της χαμηλής συνειδητότητας των χρηστών, όσων αφορά στη χρήση των μέσων κοινωνικής δικτύωσης.



Σχήμα 6.1: Προσδιορισμός φύλου

Όπως φαίνεται στο Σχήμα 6.2, στον άξονα του ηλικιακού φάσματος διακρίθηκαν πέντε επιμέρους κατηγορίες χρηστών, αυτές των 18-24 ετών, η οποία συγκέντρωσε το μεγαλύτερο μέρος των απαντήσεων (47,4%), των 25-34 ετών (19,6%), των 35-44 ετών (24,4%), των 45-54 ετών (7,4%) και τέλος των άνω των 55 ετών, στην οποία κατατάχθηκαν μόλις τρία άτομα (1,1%).

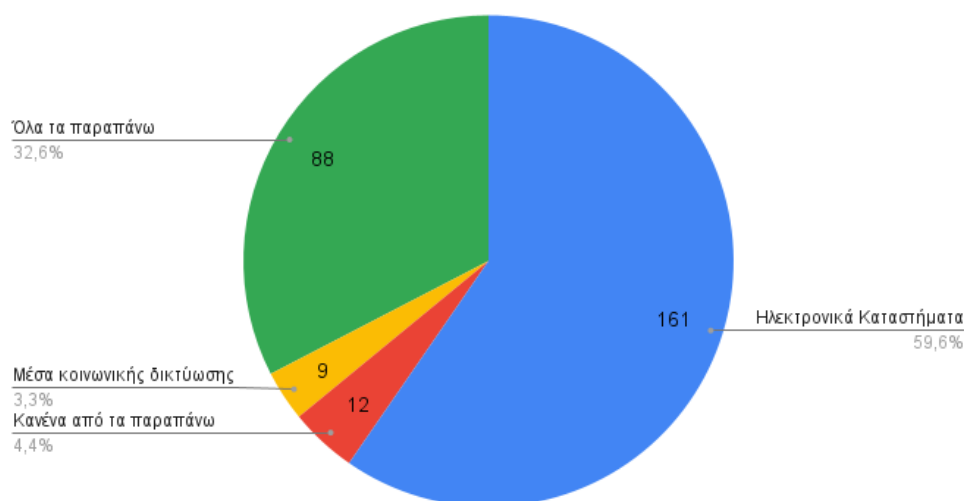


Σχήμα 6.2: Προσδιορισμός ηλικίας

Με στόχο τη διαμόρφωση μιας πρώιμης και γενικής εικόνας για το ποίον των ερωτηθέντων, υποβλήθηκαν δύο επιπλέον ερωτήσεις στο εισαγωγικό τμήμα του ερωτηματολογίου, σύμφωνα με τις οποίες οι υποψήφιοι ερωτήθηκαν αν προτιμούν να αναζητούν προϊόντα σε ηλεκτρονικά καταστήματα ή σε μέσα κοινωνικής δικτύωσης και ποιο από τα μέσα κοινωνικής δικτύωσης (Instagram, YouTube) χρησιμοποιούσαν κατά τη διάρκεια του έτους 2020.

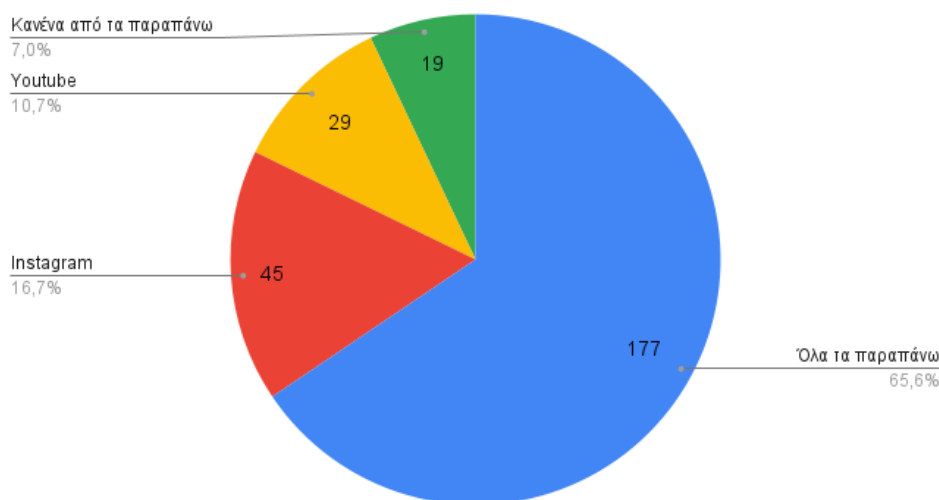
Σχετικά με την πρώτη ερώτηση, εκείνη συμπεριλήφθηκε έτσι ώστε να εξετασθεί το ποσοστό

επιρροής του εμπορικού τομέα στα μέσα κοινωνικής δικτύωσης, μέσω της αναζήτησης και αγοραπωλησίας προϊόντων σε αυτά. Σύμφωνα με το Σχήμα 6.3, γίνεται αισθητή η παρουσία των προϊόντων στα μέσα κοινωνικής δικτύωσης, καθώς το 32,6% των απαντήσεων δήλωσε πως αναζητά προϊόντα, τόσο στα μέσα κοινωνικής δικτύωσης όσο και σε ηλεκτρονικά καταστήματα, ενώ ένα μικρό ποσοστό των απαντήσεων (3,3%) στρέφεται αποκλειστικά στα μέσα κοινωνικής δικτύωσης για την αναζήτηση προϊόντων. Ωστόσο, η πλειοψηφία των ερωτηθέντων επέμεινε στην αποκλειστική χρήση των ηλεκτρονικών καταστημάτων για τον ίδιο σκοπό, με το ποσοστό αυτό να αγγίζει το 59,6%. Τέλος, ένα μικρό υποσύνολο των απαντήσεων (4,4%) απείχε από την αναζήτηση προϊόντων στα μέσα κοινωνικής δικτύωσης, αλλά και σε ηλεκτρονικά καταστήματα, αντιπροσωπεύοντας πιθανώς την μερίδα του πληθυσμού, η οποία προτιμά τις φυσικές αγορές έναντι των ψηφιακών.



Σχήμα 6.3: Προτιμάτε να αναζητάτε προϊόντα σε ηλεκτρονικά καταστήματα ή σε μέσα κοινωνικής δικτύωσης;

Η τελευταία ερώτηση την οποία κλήθηκαν να απαντήσουν οι ερωτηθέντες κατά το εισαγωγικό τμήμα του ερωτηματολογίου αφορούσε στο ποια από τις δύο πλατφόρμες, Instagram ή YouTube, χρησιμοποιούσε. Παρατηρώντας, στο Σχήμα 6.4, πως το 93% των απαντήσεων στράφηκε τουλάχιστον σε ένα από τα δύο μέσα κοινωνικής δικτύωσης, αποδεικνύεται η ένταξη των κοινωνικών δικτύων στη καθημερινή ζωή ενός μεγάλου μέρους του ελληνικού πληθυσμού. Ειδικότερα, φαίνεται πως κατά το 2020 οι δύο πλατφόρμες είχαν σχεδόν αντίστοιχη απήχηση στο ελληνικό κοινό, με το Instagram να είναι ελαφρώς περισσότερο διαδεδομένο, καθώς το ποσοστό αποκλειστικής του χρήσης άγγιξε το 16,7%, ενώ εκείνο του YouTube το 10,7%. Από την άλλη πλευρά, μια μικρή μερίδα των ερωτηθέντων δήλωσε πως δε χρησιμοποιεί καμία από τις δύο πλατφόρμες, με αποτέλεσμα εκείνοι να μην ληφθούν υπόψη στα μεταγενέστερα στάδια του ερωτηματολογίου.



Σχήμα 6.4: Ποια από τα παρακάτω μέσα κοινωνικής δικτύωσης χρησιμοποιείτε;

Εφόσον απαντήθηκαν οι εισαγωγικές ερωτήσεις του ερωτηματολογίου, η προσοχή εστιάστηκε σε συγκεκριμένους πόλους ενδιαφέροντος του Instagram και του YouTube. Αξίζει να σημειωθεί πως η μελέτη των ακόλουθων ερωτήσεων βασίστηκε στις απαντήσεις του γενικού συνόλου των ερωτηθέντων, ενώ όπου κρίθηκε απαραίτητο η εξέταση πραγματοποιήθηκε στη βάση των διαφορετικών φύλων και ηλικιακών ομάδων του δείγματος.

6.1 Ερωτήσεις Instagram

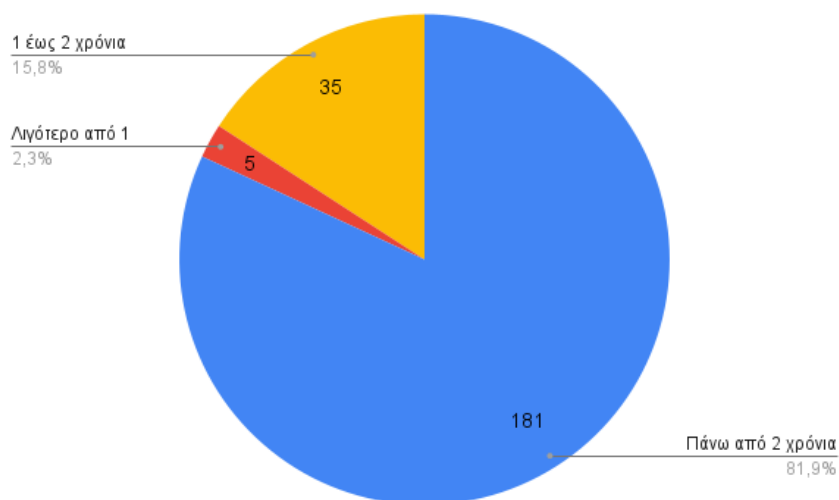
Στη συγκεκριμένη υποενότητα αναλύονται οι απαντήσεις των ερωτηθέντων στις ερωτήσεις του ερωτηματολογίου, οι οποίες αφορούσαν στη πλατφόρμα του Instagram.

6.1.1 Πόσο καιρό χρησιμοποιείτε το Instagram;

Η πρώτη σχετική με το Instagram ερώτηση, η οποία τέθηκε στους συμμετέχοντες του ερωτηματολογίου, πραγματεύεται το ιστορικό της δραστηριότητας του εκάστοτε χρήστη στην πλατφόρμα και, κατ' επέκταση, το ποσοστό εξοικείωσης του με αυτή.

Γενικά

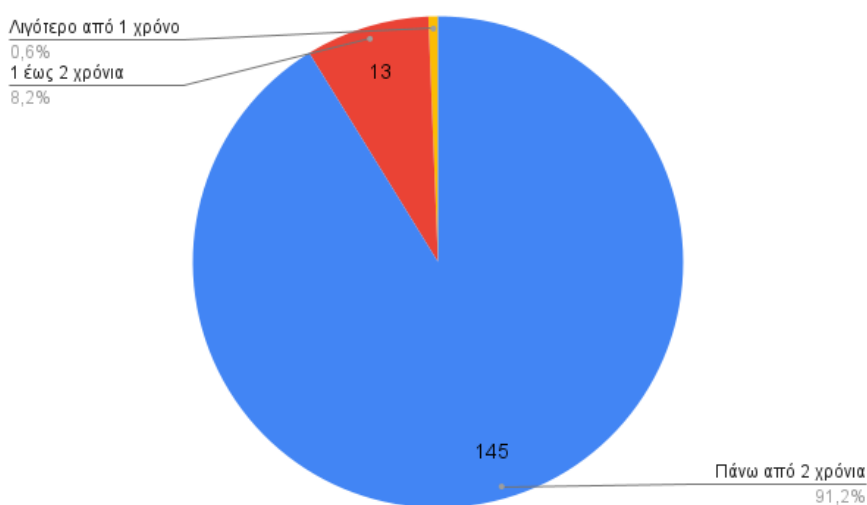
Αναφορικά με το γενικό σύνολο των απαντήσεων, προκύπτει από το Σχήμα 6.5 ότι, κατά προσέγγιση, οχτώ στους δέκα χρήστες της πλατφόρμας διέθεταν λογαριασμό σε αυτή για παραπάνω από δύο χρόνια, ενώ ένα σεβαστό υποσύνολο αυτών (15,8%) είχε εισέλθει στη διαδικτυακή κοινότητα του Instagram μεταξύ ενός και δύο ετών από την διεξαγωγή του ερωτηματολογίου. Μόλις το 2,3% των ερωτηθέντων απάντησαν πως χρησιμοποιούσαν την πλατφόρμα για λιγότερο από έναν χρόνο. Γίνεται, λοιπόν, αντιληπτό, πως η πλειοψηφία των χρηστών ανέπτυξαν τη ψηφιακή τους παρουσία στην πλατφόρμα για παραπάνω από δύο έτη, γεγονός το οποίο συνεπάγεται την εδραίωση του Instagram στον ελληνικό χώρο.



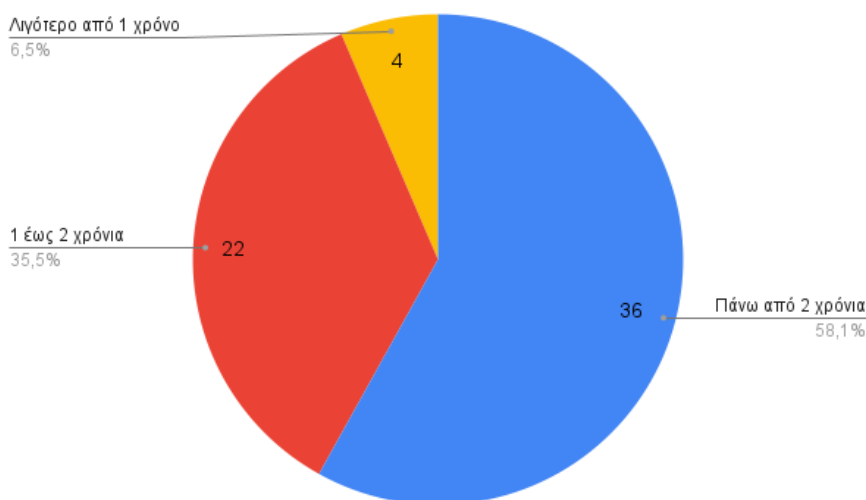
Σχήμα 6.5: Πόσο καιρό χρησιμοποιείτε το Instagram;

Ηλικιακές ομάδες

Αναλύοντας τις απαντήσεις στο επίπεδο των ηλικιών των ερωτηθέντων, προκύπτει από τα Σχήματα 6.6 και 6.7 ότι, παρά τη πλειοψηφία των χρηστών η οποία χρησιμοποιεί το Instagram για περισσότερο από δύο χρόνια, αρκετά άτομα μεγαλύτερης ηλικίας (35 και άνω) αποτελούν σχετικά νέους χρήστες της πλατφόρμας, με το 42% αυτών να δηλώνει πως ανήκει στην κοινότητα του Instagram για λιγότερο από δύο χρόνια. Γίνεται αντιληπτό, λοιπόν, πως η συγκεκριμένη πλατφόρμα δικτύωσης, αν και αποτελεί κύριο πόλο έλξης για τους νεότερους ηλικιακά χρήστες, έχει ξεκινήσει να προσελκύει άτομα μεγαλύτερης ηλικίας, επεκτείνοντας το εύρος του ενεργού του κοινού.



Σχήμα 6.6: Πόσο καιρό χρησιμοποιείτε το Instagram; (18-34)



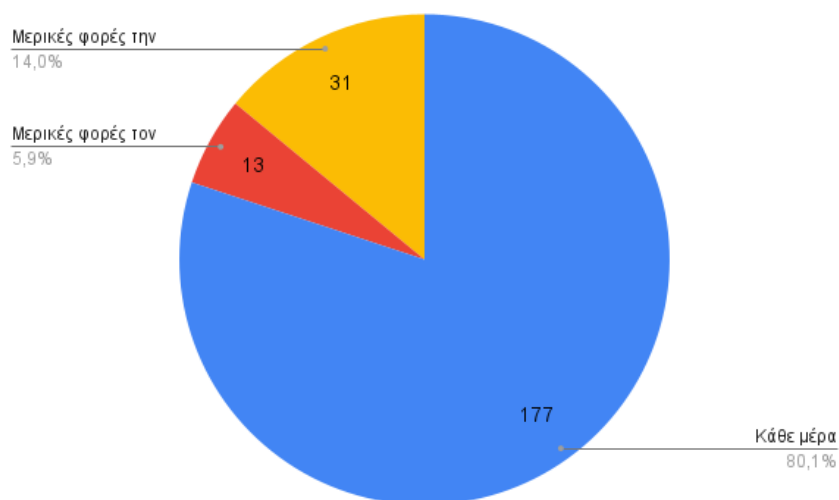
Σχήμα 6.7: Πόσο καιρό χρησιμοποιείτε το Instagram; (35+)

6.1.2 Πόσο συχνά χρησιμοποιείτε το Instagram;

Η συγκεκριμένη ερώτηση αποσκοπούσε στην εξέταση του ποσοστού χρήσης της πλατφόρμας από τους χρήστες και, κατά συνέπεια, στην αποτύπωση της επιρροής της προς αυτούς.

Γενικά

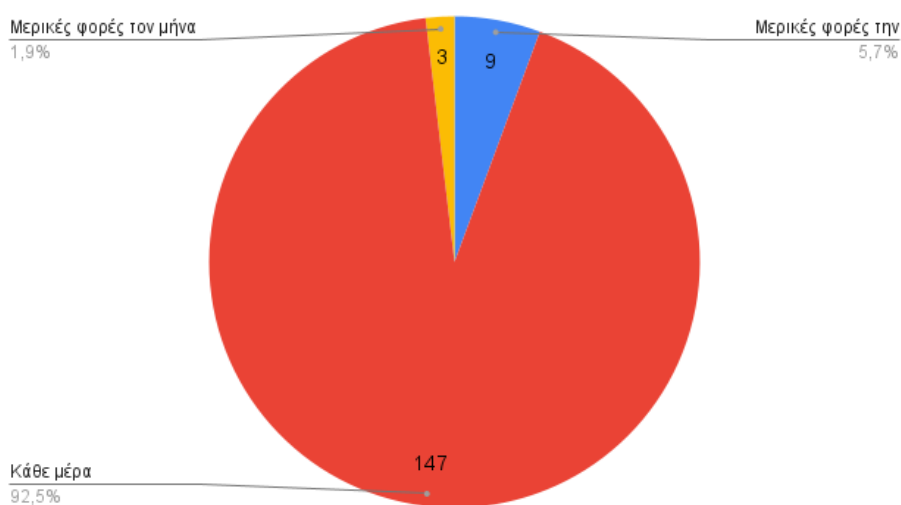
Με γνώμονα το Σχήμα 6.8, διαπιστώνεται πως η πλειονότητα των χρηστών (80,1%) πλοηγείται σε καθημερινή βάση στη πλατφόρμα, ενώ το 14% των χρηστών διατηρεί παρόμοια επίπεδα δραστηριότητας, δηλώνοντας πως χρησιμοποιεί το Instagram μερικές φορές την εβδομάδα. Μικρότερος είναι ο αριθμός των ατόμων τα οποία παραμένουν λιγότερο ενεργά στο Instagram (μερικές φορές τον μήνα), με αυτά να καταλαμβάνουν το 5,9% των δοθέντων απαντήσεων. Είναι εμφανές, επομένως, πως η πλατφόρμα του Instagram διατηρεί αμείωτο το ενδιαφέρον της πλειοψηφίας των Ελλήνων χρηστών, με την τακτική δραστηριότητα τους να αποτελεί ένδειξη της υψηλής επιρροής της σε αυτούς.



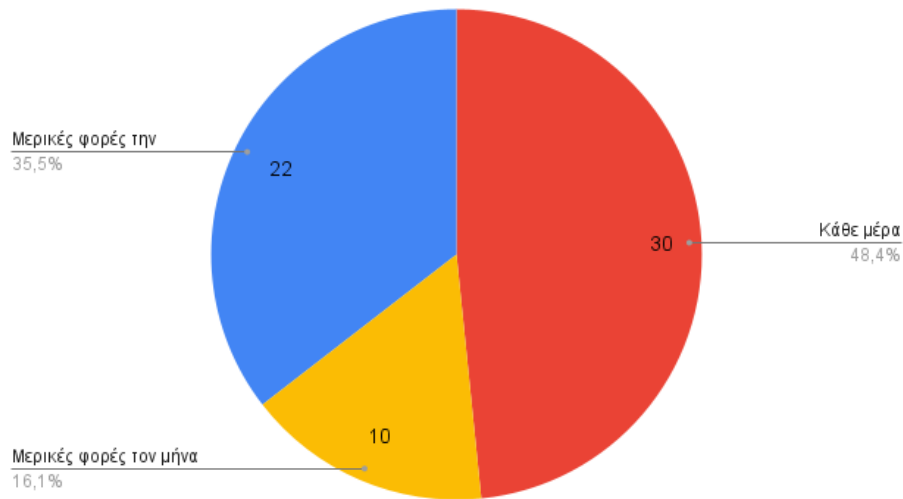
Σχήμα 6.8: Πόσο συχνά χρησιμοποιείτε το Instagram;

Ηλικιακές ομάδες

Λαμβάνοντας υπόψη τις ηλικίες των ερωτηθέντων, είναι ορατό στα Σχήματα 6.9 και 6.10 ότι η εντονότερη δραστηριότητα στο Instagram ασκείται από χρήστες μικρότερης ηλικίας (18-34), καθώς το 95,2% αυτών δήλωσε ότι περιηγείται, σε αντίθεση με τους μεγαλύτερους ηλικιακά χρήστες, ένα σημαντικό μέρος των οποίων αποκλίνει από την καθημερινή χρήση της πλατφόρμας και δραστηριοποιείται στο επίπεδο της εβδομάδας (35,5%), αλλά και του μήνα (16,1%).



Σχήμα 6.9: Πόσο συχνά χρησιμοποιείτε το Instagram; (18-34)



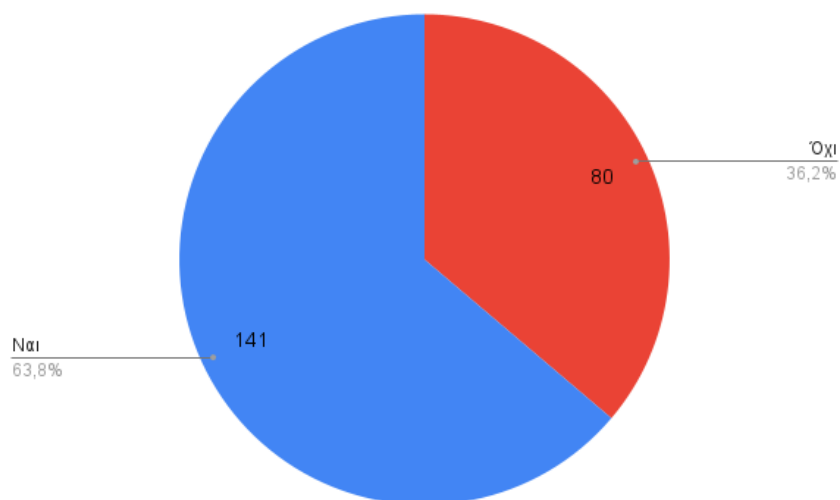
Σχήμα 6.10: Πόσο συχνά χρησιμοποιείτε το Instagram; (35+)

6.1.3 Ακολουθείτε Influencers;

Ένα σύγχρονο φαινόμενο των μέσων κοινωνικής δικτύωσης αποτελούν οι *Influencers* ή αλλιώς οι χρήστες μεγάλης δημοτικότητας, οι οποίοι ασκούν υψηλή επιρροή στο κοινό τους μέσω της ψηφιακής δραστηριότητάς τους. Το ποσοστό επιρροής του ελληνικού κοινού από τους *Influencers* είναι δυνατό να εξεταστεί σε αρχικό στάδιο, αξιοποιώντας τις ληφθείσες απαντήσεις για την παρούσα ερώτηση.

Γενικά

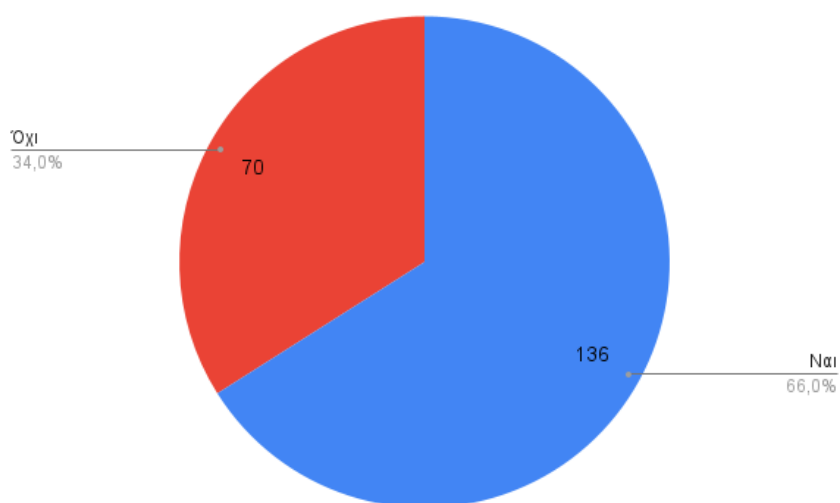
Πράγματι, παρατηρώντας το Σχήμα 6.11, είναι ορατή η τάση των Ελλήνων χρηστών να ακολουθούν *Influencers*, καθώς το 63,8% των ερωτηθέντων αποκρίθηκαν θετικά στην αντίστοιχη ερώτηση. Εντούτοις, σεβαστό είναι το ποσοστό των χρηστών, οι οποίοι δεν τάχθηκαν υπέρ της ακολούθησης *Influencers* (36,2%).



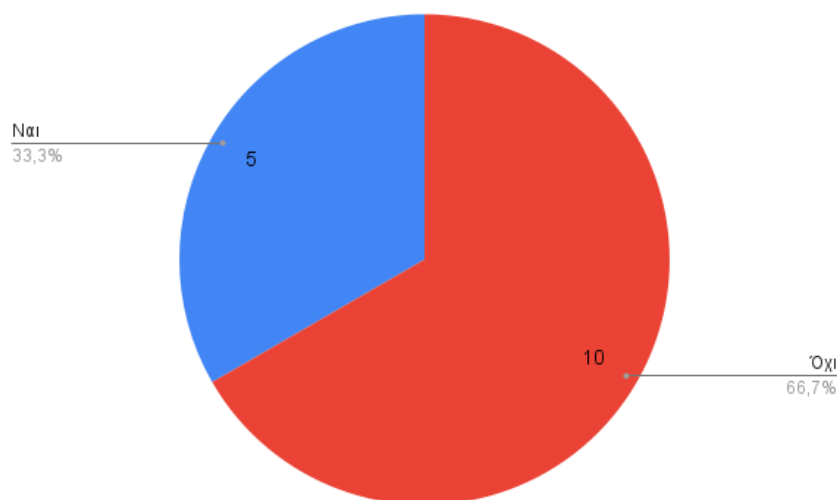
Σχήμα 6.11: Ακολουθείτε Influencers;

Ηλικιακές ομάδες

Με γνώμονα τις ηλικίες των ερωτηθέντων, διαπιστώνεται από τα Σχήματα 6.12 και 6.13 ότι το σύνολο των ακολούθων των *Influencers* απαρτίζεται, κατά βάση, από νεότερα άτομα (18-44), ενώ το 66,7% των χρηστών ηλικίας 45 ετών και άνω δήλωσαν πως δεν ακολουθούν τέτοιου είδους δημιουργούς περιεχομένου. Στο γεγονός αυτό, πιθανώς, έχει συμβάλει ο μικρότερος βαθμός εξοικείωσης των μεγαλύτερων ηλικιακά χρηστών με την πλατφόρμα (βλ. Σχήμα 6.7), αλλά και η ασυμφωνία μεταξύ του θεματικού περιεχομένου των *Influencers* και των ενδιαφερόντων των γηραιότερων ατόμων.



Σχήμα 6.12: Ακολουθείτε Influencers; (18-44)



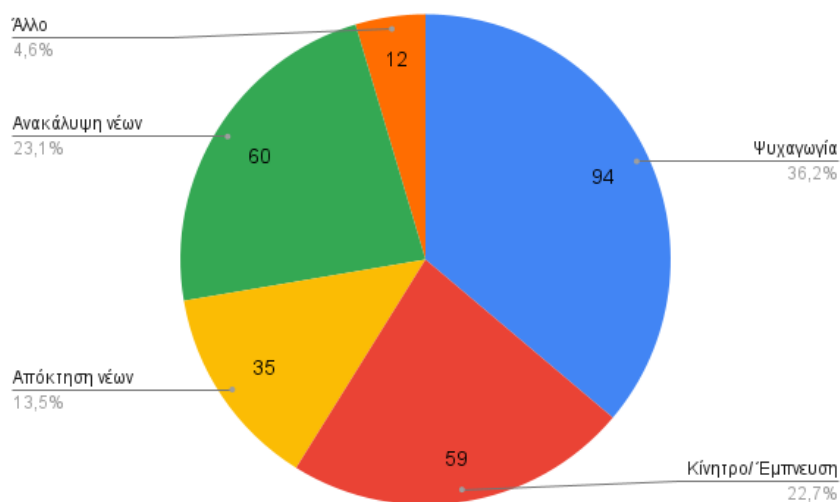
Σχήμα 6.13: Ακολουθείτε Influencers; (45+)

6.1.4 Αν ναι, τι πιστεύετε ότι σας προσφέρει το περιεχόμενο των Influencers;

Σε συνέχεια της ανάλυσης της επιρροής των *Influencers* στο ελληνικό κοινό, αναζητήθηκαν τα κίνητρα τα οποία ωθούν τους χρήστες στο να ακολουθούν τέτοιου είδους δημιουργούς περιεχομένου.

Γενικά

Σύμφωνα με το Σχήμα 6.14, εξάγεται το συμπέρασμα ότι η ψυχαγωγία αποτελεί ίσως τον σημαντικότερο λόγο για τον οποίο οι Έλληνες χρήστες ακολουθούν *Influencers*, καταλαμβάνοντας το 36,2% των απαντήσεων. Σεβαστό ήταν το πλήθος των χρηστών οι οποίοι δήλωσαν πως παρακολουθούν το περιεχόμενο των *Influencers*, με σκοπό την άντληση κινήτρου ή έμπνευσης και την ανακάλυψη νέων προϊόντων, αποδεικνύοντας για ακόμη μία φορά το εμπορικό/επιχειρηματικό ενδιαφέρον, το οποίο ενυπάρχει τα προφίλ των δημοφιλών δημιουργών περιεχομένου. Η απόκτηση νέων γνώσεων, μέσω του περιεχομένου των *Influencers*, συνέστησε τη λιγότερο διαδεδομένη απάντηση (13,5%), ενώ 12 ήταν εκείνοι οι οποίοι δήλωσαν ότι ακολουθούν *Influencers* για κάποιον άλλο σκοπό.



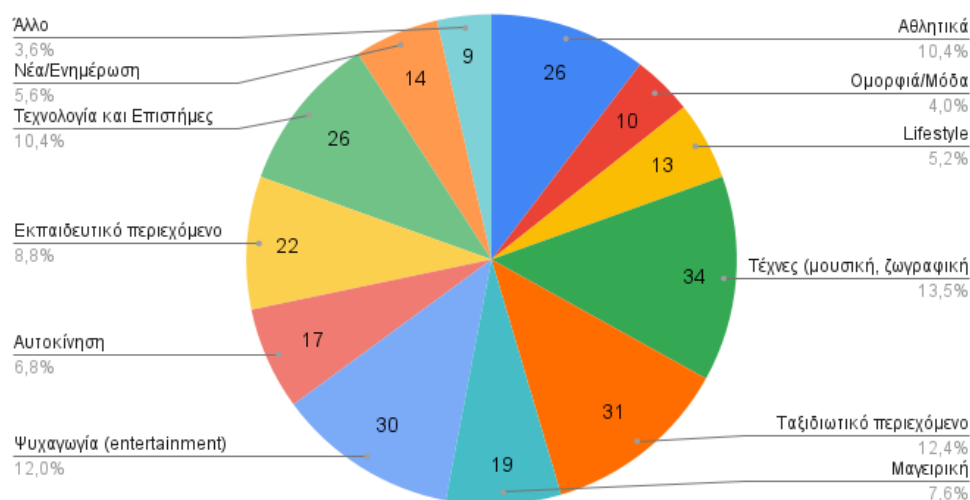
Σχήμα 6.14: Αν ναι, τι πιστεύετε ότι σας προσφέρει το περιεχόμενο των Influencers;

6.1.5 Τι θεματικό περιεχόμενο σας αρέσει να παρακολουθείτε στο Instagram;

Μια ακριβέστερη αποτύπωση των προτιμήσεων της ελληνικής κοινότητας του Instagram, σε σχέση με αυτήν η οποία προέκυψε από τα σχήματα του Κεφαλαίου 5 και συγκεκριμένα της Ενότητας 5.1, θα μπορούσε να παρέχει η άμεση ερώτηση του κοινού για το ποιες είναι οι θεματικές κατηγορίες περιεχομένου τις οποίες προτιμά να παρακολουθεί στα πλαίσια της πλατφόρμας. Καθώς, όπως φάνηκε στο Κεφάλαιο 5, οι τάσεις των ανδρών και των γυναικών εμφάνισαν αρκετές διαφορές κατά το έτος 2020, κρίθηκε σημαντική η ξεχωριστή μελέτη των απαντήσεων οι οποίες προήλθαν από τα δύο αυτά φύλα.

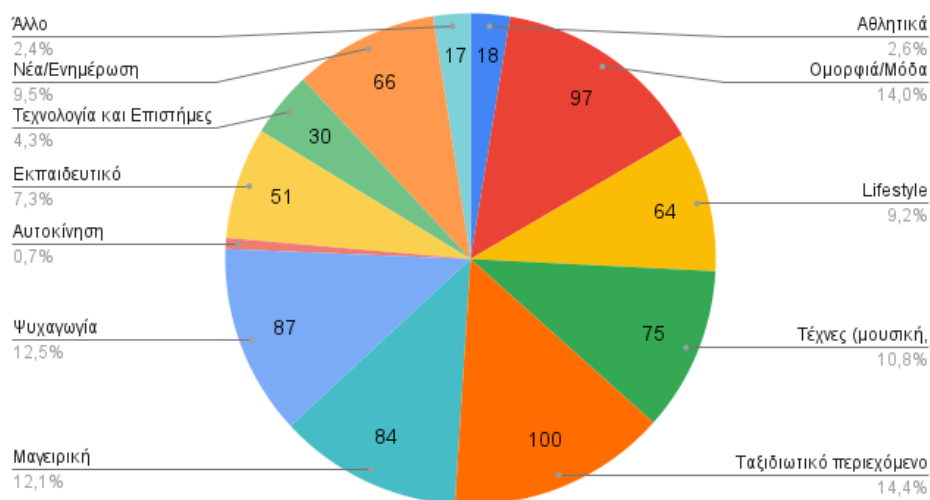
Άνδρες-Γυναίκες

Όσον αφορά στους άνδρες ερωτηθέντες, εκείνοι παρουσίασαν μία ιδιαίτερη κλίση προς τις τέχνες (13,5%), όπως καλλιτεχνική φωτογραφία, εικαστικές τέχνες ή μουσική, το ταξιδιωτικό περιεχόμενο (12,4%) και η ψυχαγωγία, αποτελέσματα τα οποία έρχονται σε συμφωνία με τα αντίστοιχα συμπεράσματα της Ενότητας 5.1. Αρκετά υψηλά στις προτιμήσεις του ανδρικού κοινού βρίσκονται, επίσης, καταλαμβάνοντας το ίδιο ποσοστό του συνόλου των απαντήσεων (10,4%), το αθλητικό και το τεχνολογικό/επιστημονικό περιεχόμενο.



Σχήμα 6.15: Τι θεματικό περιεχόμενο σας αρέσει να παρακολουθείτε στο Instagram; (Άνδρες)

Από την άλλη μεριά, οι γυναίκες της πλατφόρμας ανέδειξαν το ταξιδιωτικό περιεχόμενο (14,4%), το περιεχόμενο ομορφιάς/μόδας (14%) και την ψυχαγωγία (12,5%) ως τις περισσότερο αρεστές κατηγορίες περιεχομένου. Αποτελέσματα τα οποία, επίσης, συμβαδίζουν με τα συμπεράσματα της Ενότητας 5.1. Επιπλέον, πόλους έλξης του ενδιαφέροντος του γυναικείου κοινού φαίνεται πως αποτέλεσαν η μαγειρική (12,1%), οι τέχνες (10,8%) και το ενημερωτικό περιεχόμενο (9,5%).



Σχήμα 6.16: Τι θεματικό περιεχόμενο σας αρέσει να παρακολουθείτε στο Instagram; (Γυναίκες)

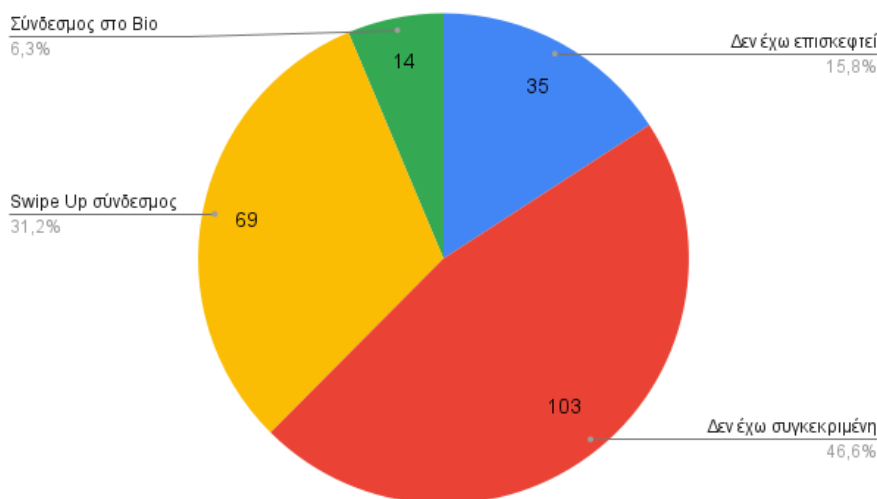
6.1.6 Προτιμάτε να επισκέπτεστε εξωτερικούς συνδέσμους (links) μέσω swipe up στα stories ή μέσω link που βρίσκεται στις πληροφορίες ενός προφίλ (Bio);

Πολλές φορές οι δημιουργοί περιεχομένου επιθυμούν να παρουσιάσουν στο κοινό τους περιεχόμενο το οποίο δεν στεγάζεται στην πλατφόρμα το Instagram, όπως για παράδειγμα η ιστοσε-

λίδα ενός επιχειρησιακού λογαριασμού. Το Instagram παρέχει τη δυνατότητα στους δημιουργούς να κατευθύνουν το κοινό προς μία εξωτερική διαδικτυακή τοποθεσία, προσαρτώντας τον αντίστοιχο σύνδεσμο, είτε στη περιγραφή του προφίλ τους (*bio*) ή σε κάποια κοινοποιημένη ιστορία (*swipe up story*). Η αποδοτικότητα της κάθε μίας από τις παραπάνω τακτικές είναι εφικτό εξεταστεί μέσω των απαντήσεων της παρούσας ερώτησης.

Γενικά

Δεδομένων των αριθμητικών στοιχείων του Σχήματος 6.17, είναι εμφανής η ουδετερότητα της πλειοψηφίας των χρηστών της πλατφόρμα, ωστόσο η προτίμηση του 31,2% των ερωτηθέντων για *swipe up* συνδέσμοι έναντι του 6,3% αυτών για συνδέσμοι στην περιγραφή λογαριασμών (*bio*), υποδηλώνει την γενικότερη κλίση του κοινού προς τους *swipe up* συνδέσμοι, γεγονός το οποίο πιθανώς απορρέει από την αμεσότητα, η οποία διέπει την επίσκεψη τέτοιου είδους εξωτερικών συνδέσμων.



Σχήμα 6.17: Προτιμάτε να επισκέπτεστε εξωτερικούς συνδέσμοι (links) μέσω *swipe up* στα stories ή μέσω link που βρίσκεται στις πληροφορίες ενός προφίλ (Bio);

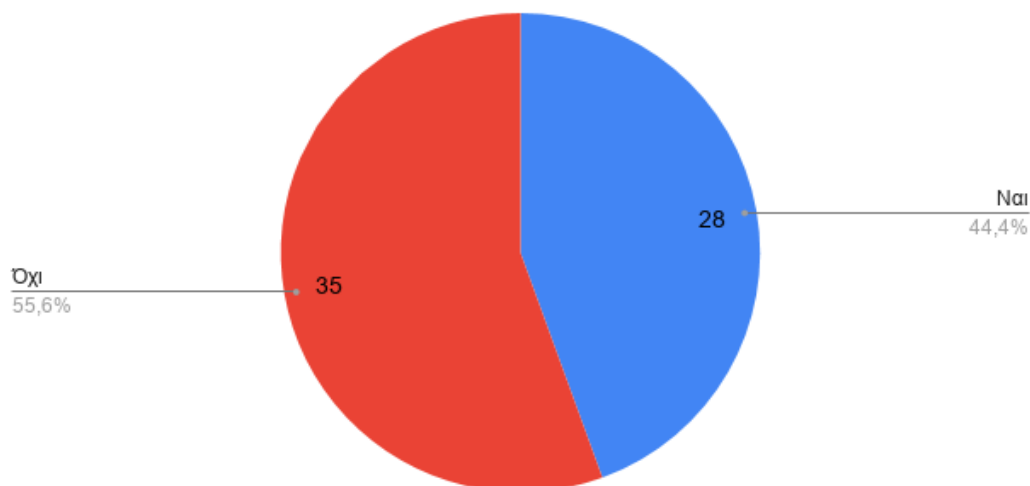
6.1.7 Έχετε αναζητήσει/αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση;

Όπως έχει ήδη ειπωθεί πολλάκις στην παρούσα έρευνα, αρκετές είναι οι εταιρείες οι οποίες, εκμεταλλευόμενες την υψηλή δραστηριότητα των χρηστών στο Instagram, αποσκοπούν στην προώθηση των προϊόντων ή των υπηρεσιών τους, είτε σε πρώτο πρόσωπο, μέσω προσωπικών δημοσιεύσεων, ή μέσω δημοφιλών δημιουργών περιεχομένου, τακτική η οποία είναι γνωστή και ως *Influencer Marketing*. Το ποσοστό επιρροής του κοινού από τέτοιου είδους δημοσιεύσεις είναι δυνατό να μελετηθεί αξιοποιώντας τη συγκεκριμένη ερώτηση.

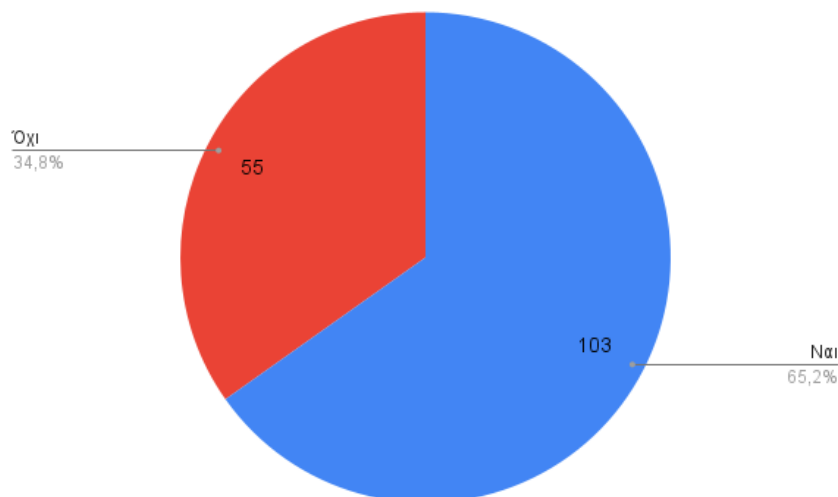
Άνδρες-Γυναίκες

Το γεγονός ότι στην Ενότητα 5.1 είχε φανεί πως οι επιχειρήσεις οι οποίες δραστηριοποιούνται στο Instagram είχαν την τάση, κατά το 2020, να απευθύνονται σε γυναίκες δημιουργούς/χρή-

στες για την προώθηση του περιεχομένου τους, έδωσε το έναυσμα για την ξεχωριστή ανάλυση των απαντήσεων, οι οποίες προήλθαν από τα δύο φύλα. Παρατηρώντας τα Σχήματα 6.18 και 6.19, διαπιστώνεται πως το γυναικείο κοινό παρουσιάζει μεγαλύτερη τάση από τους άνδρες, προς την αναζήτηση ή αγορά προϊόντων τα οποία ανακάλυψαν από κάποια δημοσίευση του Instagram, φαινόμενο το οποίο, πιθανώς, υποδηλώνει την καλύτερη ανταπόκριση των επιχειρήσεων και γενικότερα της ελληνικής κοινότητας της πλατφόρμας στις απαιτήσεις του γυναικείου κοινού.



Σχήμα 6.18: Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (Άνδρες)

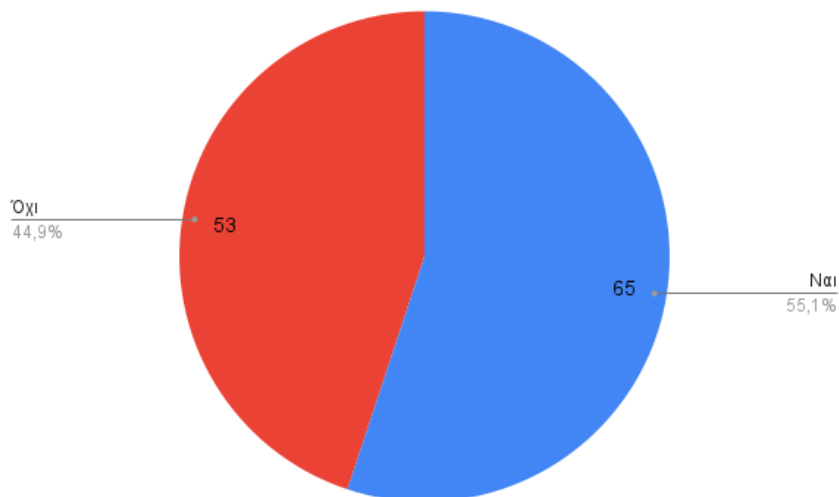


Σχήμα 6.19: Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (Γυναίκες)

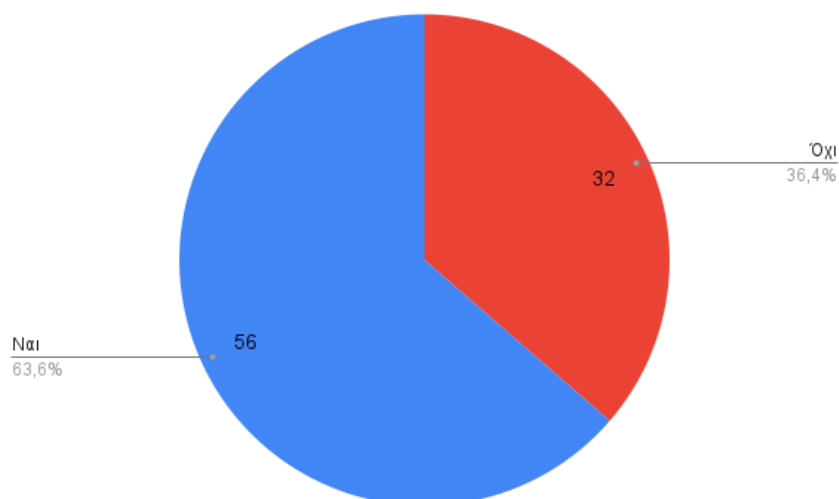
Ηλικιακές ομάδες

Στον άξονα των ηλικιών των ερωτηθέντων, παρατηρείται στα Σχήματα 6.20, 6.21 και 6.22 ότι, μεταβαίνοντας προς τις μεγαλύτερες ηλικιακές κλίμακες, οι χρήστες εμφανίζουν μεγαλύτερη

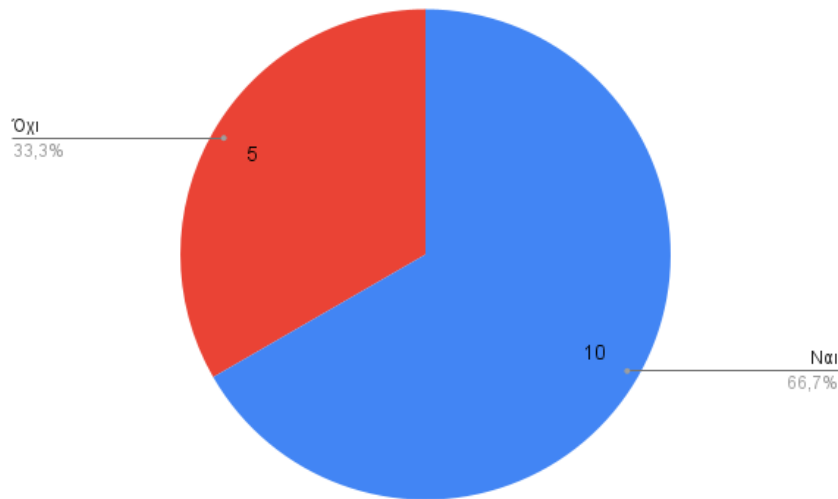
τάση στην αναζήτηση ή αγορά προϊόντων, τα οποία ανακάλυψαν σε κάποια διαφήμιση της πλατφόρμας, δείχνοντας, ενδεχομένως, τον υψηλότερο βαθμό επιρροής των γηραιότερων χρηστών από την προώθηση προϊόντων στο Instagram, σε σύγκριση με τους νεότερους.



Σχήμα 6.20: Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (18-24)



Σχήμα 6.21: Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (25-44)



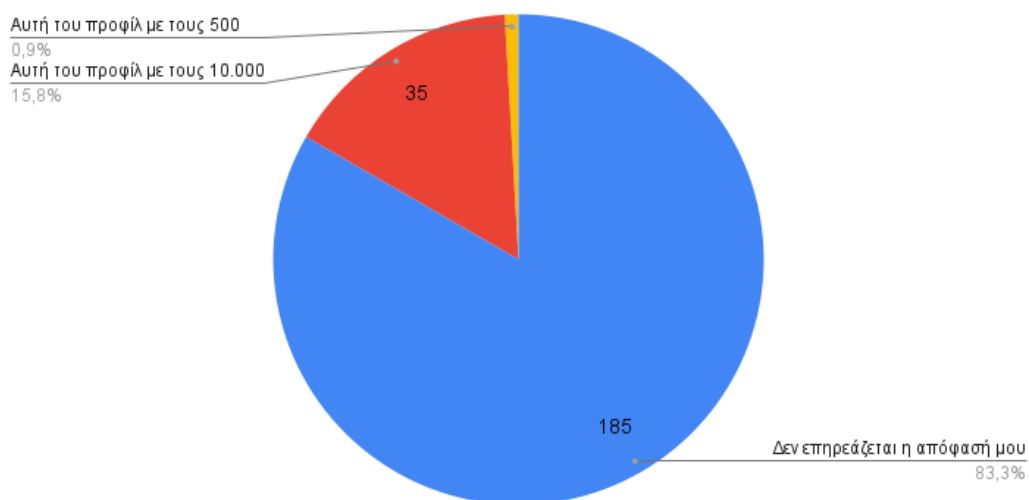
Σχήμα 6.22: Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που διαφημίστηκε σε κάποια δημοσίευση; (45+)

6.1.8 Έστω ότι δύο Instagram προφίλ, ένα με 500 followers και ένα με 10.000 followers παρουσιάζουν το ίδιο προϊόν. Την άποψη ποιου από τα δύο προφίλ θα εμπιστευτείτε;

Στο επίπεδο του *Influencer Marketing* και του ποσοστού επιρροής του κοινού από τους δημιουργούς περιεχομένου, όσον αφορά στην προώθηση ενός προϊόντος, θεωρήθηκε ενδιαφέρουσα η μελέτη της σχέσης μεταξύ του αριθμού ακολούθων ενός προφίλ και της εμπιστοσύνης την οποία δείχνει το κοινό σε αυτόν.

Γενικά

Παρά την ουδέτερη στάση του κοινού ως προς την εμπιστευτικότητα της άποψης ενός μικρού ή μεγάλου προφίλ, αντίστοιχα, για ένα προϊόν (βλ. Σχήμα 6.23), είναι φανερό πως αν έπρεπε να επιλέξει ένα από τα δύο προφίλ, αυτό πιθανότατα θα ήταν εκείνο με τον μεγαλύτερο αριθμό ακολούθων (15,8%), δείχνοντας έτσι το φαινομενικά μεγαλύτερο κύρος και, κατά συνέπεια, το μεγαλύτερο ποσοστό επιρροής, το οποίο προσδίδει ο μεγαλύτερος αριθμός ακολούθων σε έναν δημιουργό περιεχομένου.



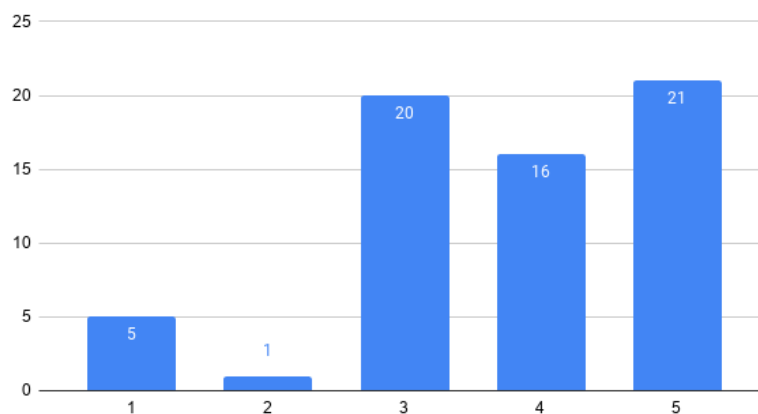
Σχήμα 6.23: Έστω ότι δύο Instagram προφίλ, ένα με 500 followers και ένα με 10.000 followers παρουσιάζουν το ίδιο προϊόν. Την άποψη ποιου από τα δύο προφίλ θα εμπιστευτείτε;

6.1.9 Σας ενοχλεί η ύπαρξη χορηγούμενων δημοσιεύσεων;

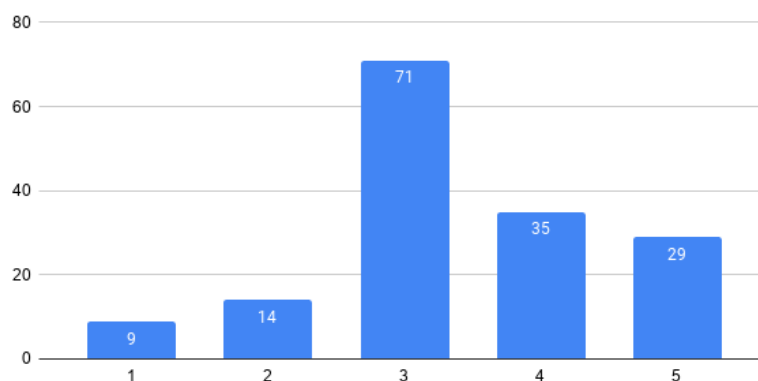
Στο κυνήγι για την ευρεία προώθηση του περιεχομένου τους, οι επιχειρήσεις, αλλά και γενικότερα οι δημιουργοί περιεχομένου, προβαίνουν στη δημιουργία διαφημίσεων. Οι δημοσιεύσεις οι οποίες επιλέγονται να διαφημιστούν προβάλλονται σε έναν μεγάλο αριθμό χρηστών, οι οποίοι πιθανότατα δεν αποτελούν ακόλουθους των διαφημιζόμενων προφίλ και, ως εκ τούτου, δεν είναι εξοικειωμένοι με το προβαλλόμενο περιεχόμενο. Η συγκεκριμένη έλλειψη εξοικείωσης, συχνά, μεταφράζεται σε αρνητικό αντίκτυπο της διαφημιζόμενης ανάρτησης στο κοινό, φαινόμενο το οποίο τέθηκε προς εξέταση μέσω της παρούσας ερώτησης.

Άνδρες-Γυναίκες

Συνοψίζοντας τα Σχήματα 6.24 και 6.25 παρατηρείται ότι το γυναικείο κοινό διατήρησε μια περισσότερο δεκτική στάση απέναντι στις διαφημίσεις του Instagram σε σχέση με τους άνδρες, οι οποίες τάχθηκαν αρκετά κατά αυτών. Η στάση αυτή των γυναικών παραπέμπει στα συμπεράσματα της ερώτησης στην Υποενότητα 6.1.7, όπου και πάλι το γυναικείο κοινό παρουσίασε μία περισσότερο θετική ανταπόκριση όσον αφορά στην αναζήτηση ή αγορά διαφημιζόμενων προϊόντων στο Instagram.



Σχήμα 6.24: Σας ενοχλεί η ύπαρξη χορηγούμενων δημοσιεύσεων; (Άνδρες)



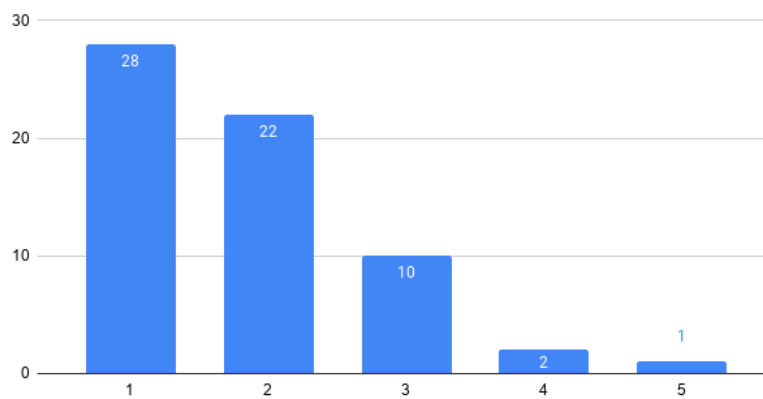
Σχήμα 6.25: Σας ενοχλεί η ύπαρξη χορηγούμενων δημοσιεύσεων; (Γυναίκες)

6.1.10 Επισκέπτεστε τα προφίλ χορηγούμενων δημοσιεύσεων;

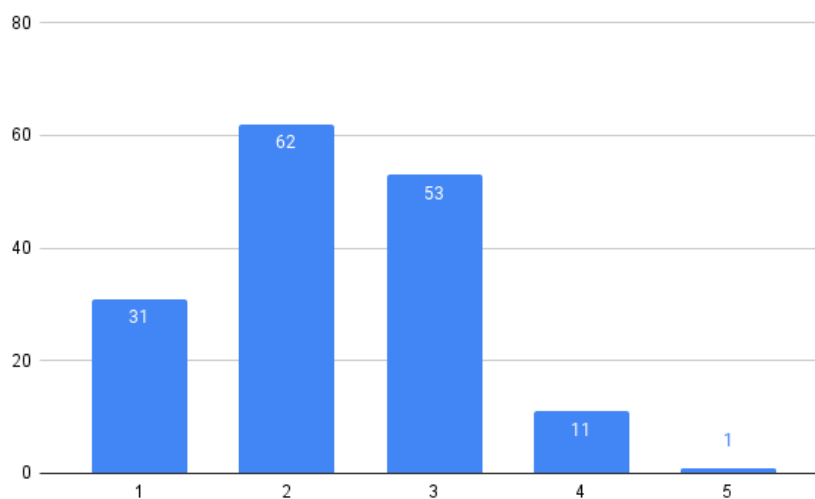
Σε συνέχεια της μελέτης των χορηγούμενων δημοσιεύσεων, το ενδιαφέρον στράφηκε στο κατά πόσο ο συγκεκριμένος τύπος αναρτήσεων είναι ικανός να προσελκύσει νέους υποψήφιους ακόλουθους στους διαφημιζόμενους λογαριασμούς, πέραν της αυξημένης προβολής, την οποία προσφέρει στη δημοσίευση.

Άνδρες-Γυναίκες

Για ακόμη μία φορά, οι γυναίκες φαίνεται πως είναι περισσότερο δεκτικές προς τις χορηγούμενες δημοσιεύσεις αναλογικά με τους άνδρες, οι οποίοι διατήρησαν και στη συγκεκριμένη, σχετική με τις διαφημίσεις, ερώτηση αρνητική στάση. Ωστόσο, λαμβάνοντας υπόψη την γενικότερη εικόνα των απαντήσεων, διαπιστώνεται πως οι χορηγούμενες δημοσιεύσεις δεν εξασφαλίζουν απαραίτητα την διεύρυνση του πιστού κοινού ενός δημιουργού, αλλά μόνο την ευρεία προβολή της διαφημιζόμενης ανάρτησης.



Σχήμα 6.26: Επισκέπτεστε τα προφίλ χορηγούμενων δημοσιεύσεων; (Άνδρες)



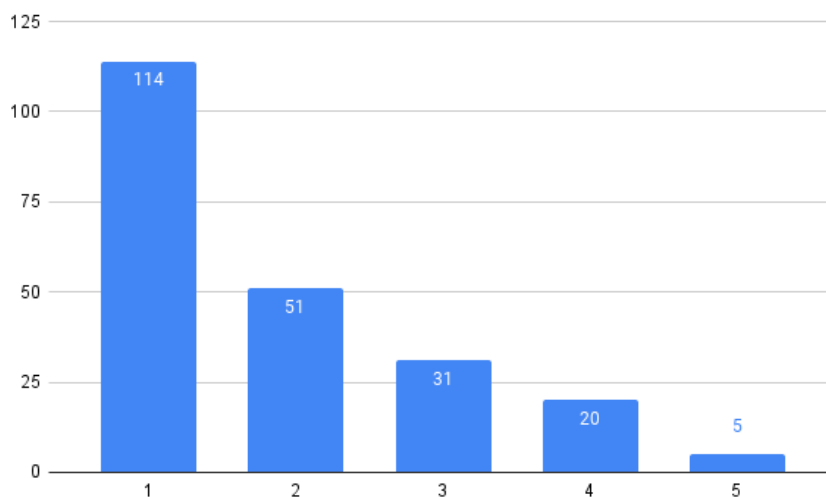
Σχήμα 6.27: Επισκέπτεστε τα προφίλ χορηγούμενων δημοσιεύσεων; (Γυναίκες)

6.1.11 Κάνετε Like/Comment σε δημοσιεύσεις λογαριασμών που δεν ακολουθείτε;

Ολοκληρώνοντας την εξέταση των χορηγούμενων απαντήσεων, κρίθηκε σημαντική η αναζήτηση του ποσοστού αλληλεπίδρασης του κοινού με αυτές και, κατ' επέκταση με αναρτήσεις λογαριασμών τους οποίους δεν ακολουθούν, μέσω των θετικών αντιδράσεων και των σχολιασμών.

Γενικά

Η ελλιπής εξοικείωση του κοινού με το περιεχόμενο των διαφημιζόμενων λογαριασμών φαίνεται πως παίζει καθοριστικό ρόλο στη διάδραση των χρηστών με τις χορηγούμενες δημοσιεύσεις, καθώς προκύπτει από το Σχήμα 6.28 ότι οι άνδρες και οι γυναίκες διατήρησαν από κοινού αρνητική στάση, αντιδρώντας στον ελάχιστο βαθμό με το προβαλλόμενο περιεχόμενο.



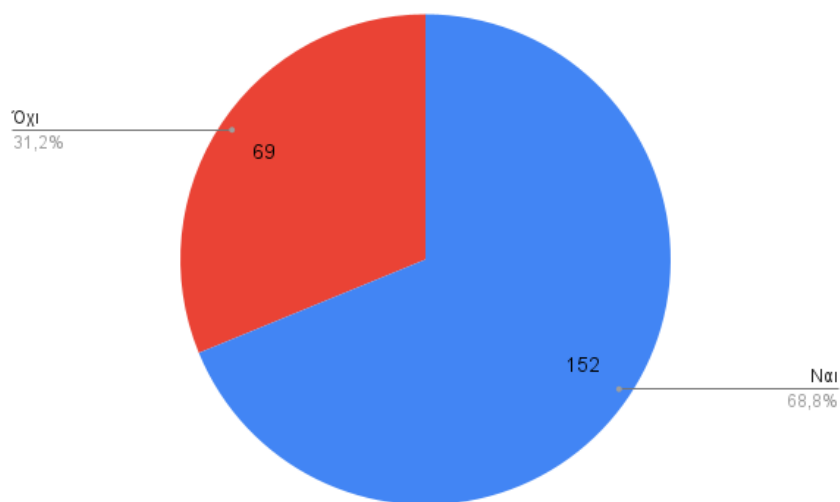
Σχήμα 6.28: Κάνετε Like/Comment σε δημοσιεύσεις λογαριασμών που δεν ακολουθείτε;

6.1.12 Έχετε συμμετάσχει ποτέ σε κάποιο *giveaway*;

Με τον όρο *giveaway* υπενθυμίζεται ότι γίνεται αναφορά σε αναρτήσεις/διαγωνισμούς, στα πλαίσια των οποίων ο διαχειριστής του λογαριασμού καλεί το κοινό να αντιδράσει στη συγκεκριμένη δημοσίευση, με σκοπό τη διεκδίκηση ενός προϊόντος. Ο πρωταγωνιστικός ρόλος των προϊόντων και η μεγάλη ανταπόκριση του κοινού σε τέτοιου είδους δημοσιεύσεις, όπως φάνηκε και στην Ενότητα 5.1, έχουν καταστήσει τα *giveaways* ως ένα ιδανικό μέσο προώθησης των επιχειρήσεων στα μέσα κοινωνικής δικτύωσης. Ο συγκεκριμένη ερώτηση τέθηκε έτσι ώστε να εξακριβωθεί η ανταπόκριση του ελληνικού κοινού σε αναρτήσεις διαγωνισμών στο Instagram.

Γενικά

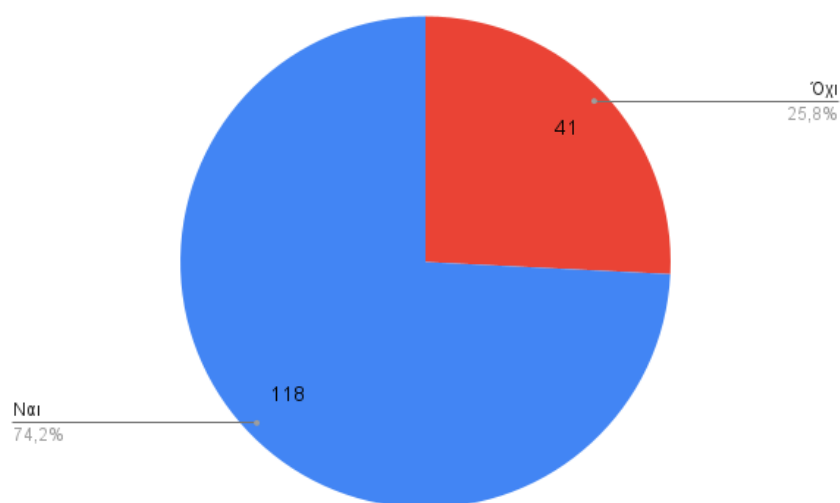
Πράγματι, μέσω του Σχήματος 6.29, επιβεβαιώνεται η μεγάλη δημοφιλία των διαγωνισμών, καθώς το 68,8% των ερωτηθέντων απάντησε πως έχει συμμετάσχει έστω και μία φορά σε κάποιο *giveaway*, αποδεικνύοντας πως τέτοιου είδους δημοσιεύσεις αποτελούν έναν από τους πιο κατάλληλους τρόπους ανάδειξης των επιχειρήσεων του Instagram.



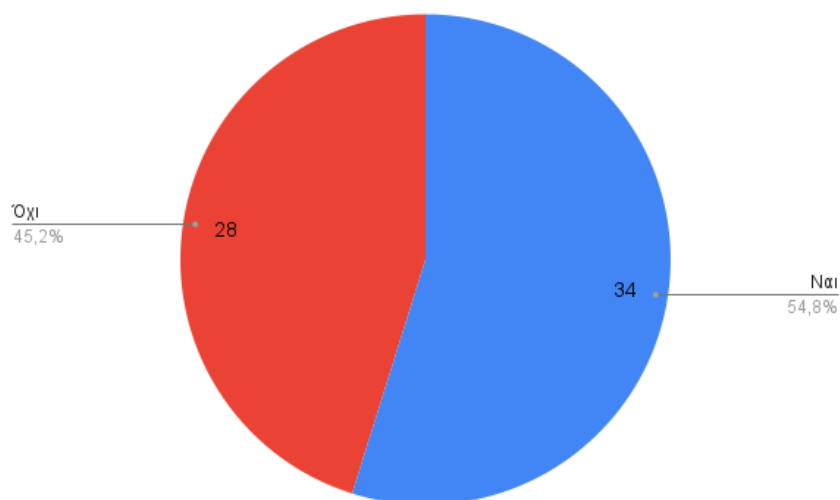
Σχήμα 6.29: Έχετε συμμετάσχει ποτέ σε κάποιο giveaway;

Ηλικιακές ομάδες

Όσον αφορά στις ηλικίες των ερωτηθέντων και με βάση τα Σχήματα 6.30 και 6.31, φαίνεται πως, παρά τη μεγάλη συμμετοχή χρηστών από ολόκληρο το ηλικιακό φάσμα στους διαγωνισμούς του Instagram, το σύνολο των συμμετεχόντων σε *giveaways* απαρτίζεται σε μεγαλύτερο ποσοστό από νεότερους χρήστες (18-34), φαινόμενο το οποίο είναι πιθανό να συνδέεται με τη μικρότερη στροφή των γηραιότερων χρηστών (45 και άνω) προς τους δημοφιλείς δημιουργούς περιεχομένου (βλ. Σχήμα 6.13), αλλά και το γενικότερα μειωμένο ενδιαφέρον αυτών για τέτοιου είδους αναρτήσεις.



Σχήμα 6.30: Έχετε συμμετάσχει ποτέ σε κάποιο giveaway; (18-34)



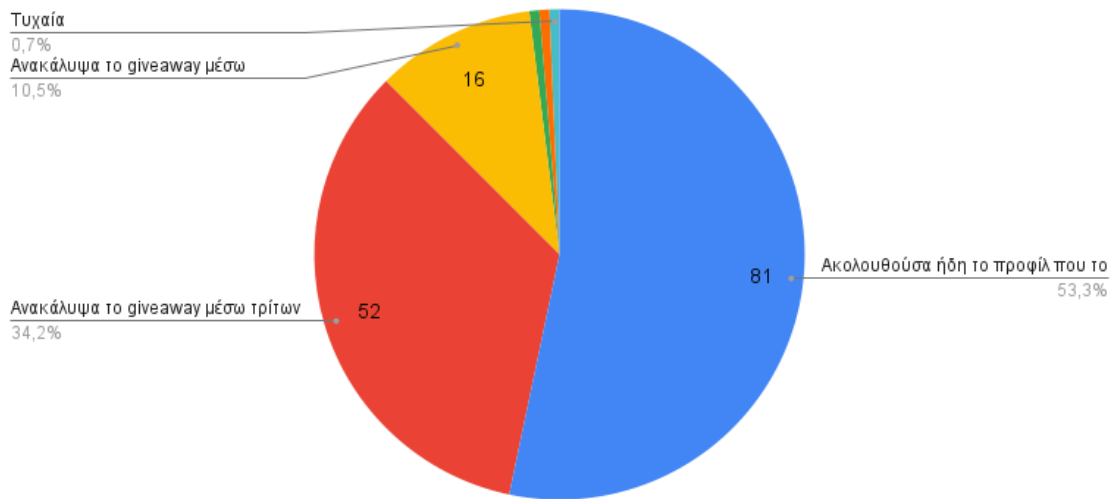
Σχήμα 6.31: Έχετε συμμετάσχει ποτέ σε κάποιο giveaway; (35+)

6.1.13 Αν ναι, πώς ανακαλύψατε αυτό το giveaway;

Τα *giveaways* απευθύνονται σε ένα ευρύτατο κοινό, το οποίο περιλαμβάνει αρκετούς χρήστες που δεν ακολουθούν τον λογαριασμό του διοργανωτή. Με αφορμή το γεγονός αυτό, κρίθηκε σημαντική η αναζήτηση των τρόπων προσέλκυσης νέων ατόμων στα πλαίσια των διαγωνισμών.

Γενικά

Οι δύο βασικοί δίαυλοι γνωστοποίησης των *giveaways*, βάσει του Σχήματος 6.32, φαίνεται πως είναι η κοινοποίηση τους στο ήδη εδραιωμένο κοινό του διοργανωτή (53,3%) και η κοινοποίηση τους μέσω άλλων χρηστών της πλατφόρμας, με σκοπό τη προσέγγιση ατόμων εκτός του φάσματος των ακολούθων του διοργανωτή (34,2%). Τέλος, λιγότερο διαδεδομένη, αλλά συγκεντρώνοντας ένα σεβαστό ποσοστό των απαντήσεων (10,5%), φαίνεται να είναι η τακτική διαφήμισης των διαγωνισμών στην πλατφόρμα του Instagram.



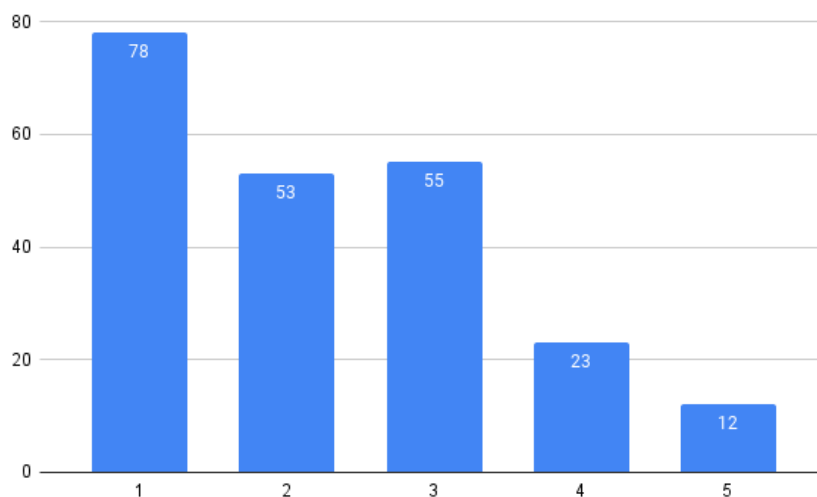
Σχήμα 6.32: Αν ναι, πώς ανακαλύψατε αυτό το giveaway;

6.1.14 Αναζητείτε πληροφορίες, για προϊόντα που ενδιαφέρεστε να αγοράσετε, στο Instagram;

Ολοκληρώνοντας τη μελέτη του εμπορικού χαρακτήρα του Instagram, αναλύεται ο βαθμός κατά τον οποίο το περιβάλλον της συγκεκριμένης πλατφόρμας ευνοεί την ανάπτυξη καταναλωτικών σχέσεων, μέσω της παρουσίασης, προώθησης και αγοραπωλησίας προϊόντων.

Γενικά

Στην ερώτηση για το εάν το ελληνικό κοινό στρέφεται στο Instagram για την αναζήτηση πληροφοριών για προϊόντα τα οποία το ενδιαφέρουν, φάνηκε πως η πλειοψηφία των χρηστών χρησιμοποιεί λιγότερο συχνά έως και καθόλου την πλατφόρμα για τον συγκεκριμένο σκοπό. Παρά το γεγονός αυτό, ένα σεβαστό υποσύνολο των ερωτηθέντων διατήρησαν ουδέτερη έως και θετική στάση στην ερώτηση αυτή. Συμπερασματικά, διαπιστώνεται πως η προώθηση και η αγοραπωλησία προϊόντων δεν αποτελεί το ένα από τα θεμελιώδη γνωρίσματα της πλατφόρμας του Instagram, όμως εκείνη διαθέτει το υπόβαθρο για την ανάπτυξη τέτοιου είδους δραστηριοτήτων.



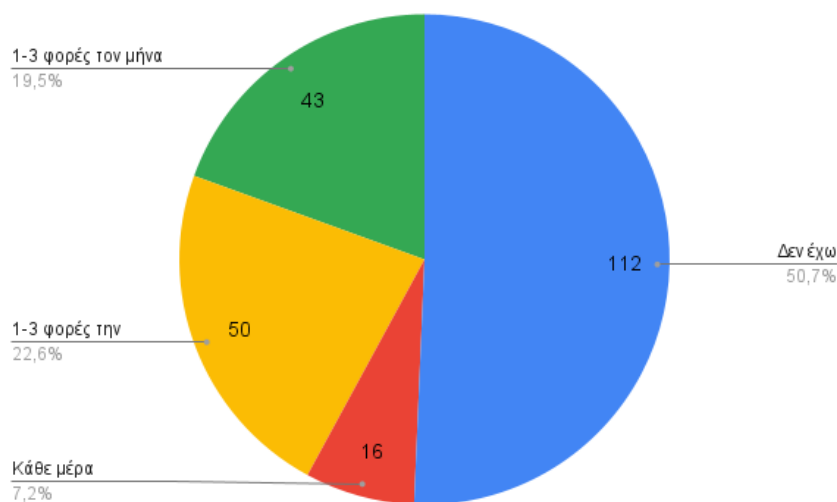
Σχήμα 6.33: Αναζητείτε πληροφορίες, για προϊόντα που ενδιαφέρεστε να αγοράσετε, στο Instagram;

6.1.15 Ποια θεωρείτε ότι είναι η ιδανική συχνότητα αναρτήσεων;

Στο σημείο αυτό, το ερευνητικό ενδιαφέρον μεταβαίνει στην ανάλυση των προτιμήσεων του κοινού, όσον αφορά στη γενικότερη χρήση της πλατφόρμας. Ως πρώτο βήμα του συγκεκριμένου στάδιο της μελέτης, αναζητήθηκε η συχνότητα ανάρτησης των δημοσιεύσεων, η οποία ανταπεξέρχεται με τον βέλτιστο τρόπο στις ανάγκες του ελληνικού κοινού.

Γενικά

Όπως παρατηρείται στο Σχήμα 6.34, ένας στους δύο Έλληνες χρήστες, κατά προσέγγιση (50,7%), δήλωσε ότι δεν έχει κάποια συγκεκριμένη προτίμηση για τη συχνότητα ανάρτησης δημοσιεύσεων, ωστόσο εστιάζοντας στο υπόλοιπο 49,3% των απαντήσεων, παρουσιάζεται η κλίση του κοινού προς τους χαμηλότερους ρυθμούς ανάρτησης των δημοσιεύσεων, με την ιδανική να φαίνεται πως είναι η 1-3 δημοσιεύσεις την εβδομάδα. Η καθημερινή κοινοποίηση περιεχομένου αποτέλεσε τη λιγότερο αρεστή επιλογή του κοινού, ρίχνοντας φως στο Σχήμα 5.31, όπου είχε αποτυπωθεί το μειωμένο ποσοστό αλληλεπίδρασης των χρηστών με το περιεχόμενο των δημιουργών, οι οποίοι διατηρούσαν ένα καθημερινό πρόγραμμα αναρτήσεων.



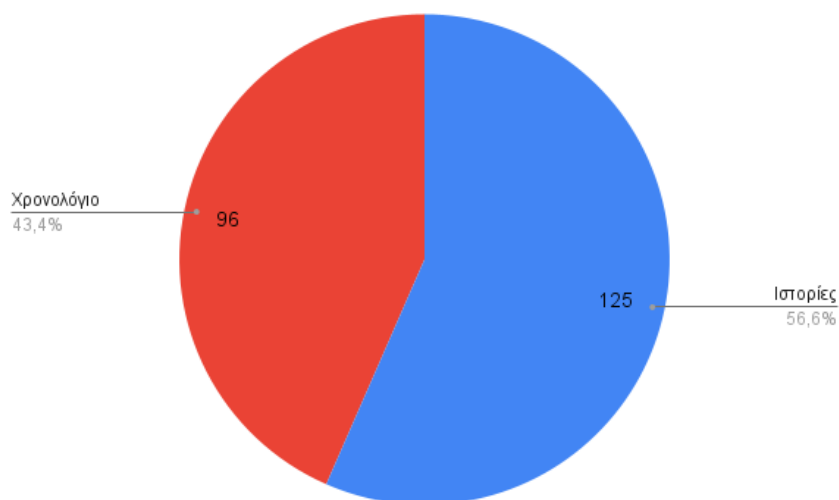
Σχήμα 6.34: Ποια θεωρείτε ότι είναι η ιδανική συχνότητα αναρτήσεων;

6.1.16 Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο (timeline) ή μέσω των instagram ιστοριών (stories);

Καθημερινά η πλατφόρμα του Instagram κατακλύζεται από νέες δημοσιεύσεις, τόσο στον χώρο του χρονολογίου (*timeline*), όσο και στον χώρο των ιστοριών (*stories*), με αποτέλεσμα να υπάρχει ο κίνδυνος αυτές να χαθούν μέσα στο πλήθος και να μην λάβουν την επιθυμητή προσοχή. Στη προσπάθεια αποφυγής του συγκεκριμένου φαινομένου, αρκετοί χρήστες επιλέγουν να κοινοποιήσουν τις δημοσιεύσεις τους και στον χώρο των ιστοριών, έτσι ώστε να άτομα τα οποία δείχνουν μεγαλύτερο ενδιαφέρον προς τις ιστορίες να το εντοπίσουν ευκολότερα. Η παρούσα ερώτηση, λαμβάνοντας υπόψη τη γενικότερη στροφή στις ιστορίες του Instagram, αποσκοπεί στην ανάλυση της αποτελεσματικότητας της προκείμενης τακτικής.

Γενικά

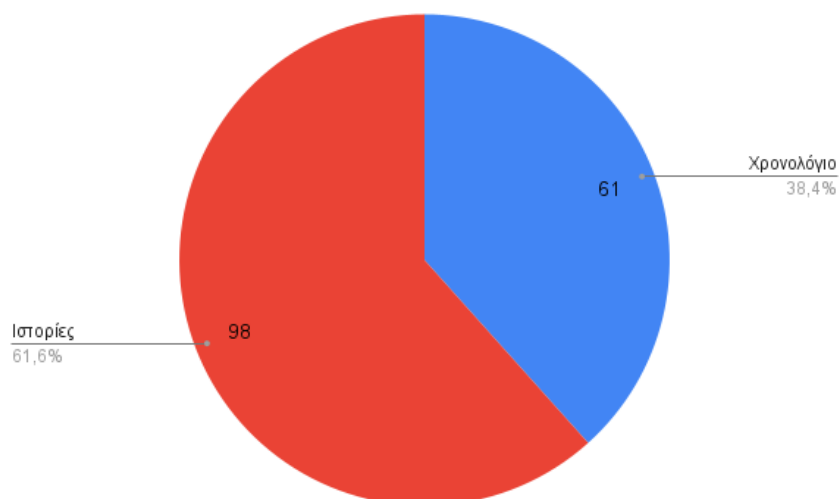
Γίνεται φανερό από το Σχήμα 6.35, ότι όντως το ελληνικό κοινό ανακαλύπτει συχνότερα νέες δημοσιεύσεις μέσω των ιστοριών του Instagram (56,6%), αποδεικνύοντας έτσι τη σημαντικότητα της προώθησης νέων αναρτήσεων με τη χρήση αυτών.



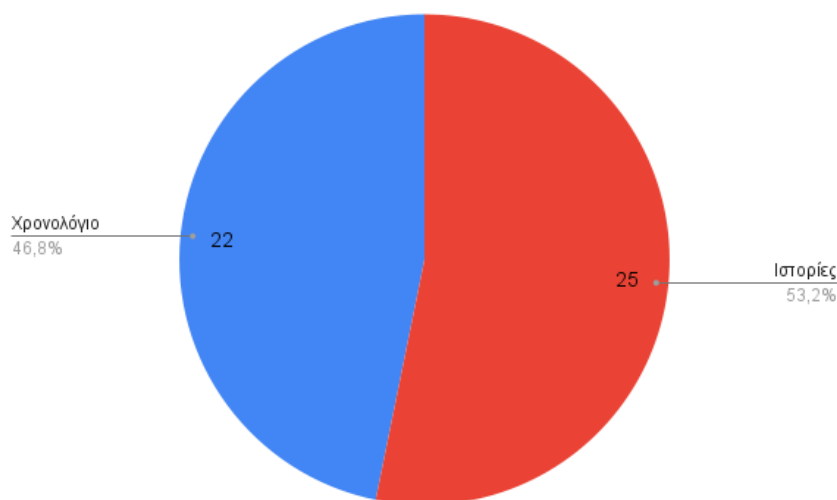
Σχήμα 6.35: Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο ή μέσω των instagram ιστοριών;

Ηλικιακές ομάδες

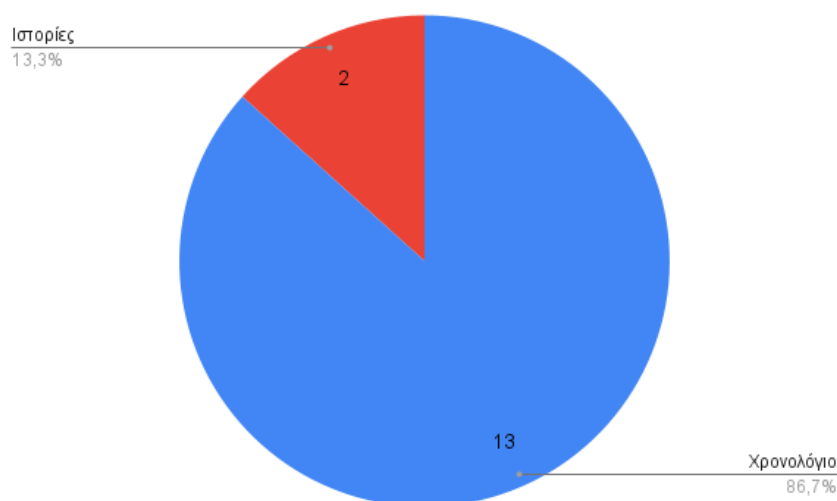
Στρέφοντας το ενδιαφέρον προς τις ηλικιακές ομάδες των ερωτηθέντων, γίνεται αντιληπτό μέσω των Σχημάτων 6.36, 6.37 και 6.38 ότι το πεδίο των ιστοριών του Instagram σημειώνει μεγαλύτερη απήχηση από νεότερους ηλικιακά χρήστες, καθώς η μετάβαση προς τις ανώτερες ηλικιακές ομάδες συνοδεύεται από τη συγκέντρωση του ενδιαφέροντος της πλειοψηφίας των χρηστών στον χώρο του χρονολογίου.



Σχήμα 6.36: Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο (timeline) ή μέσω των instagram ιστοριών (stories); (18-34)



Σχήμα 6.37: Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο (timeline) ή μέσω των instagram ιστοριών (stories); (35-44)



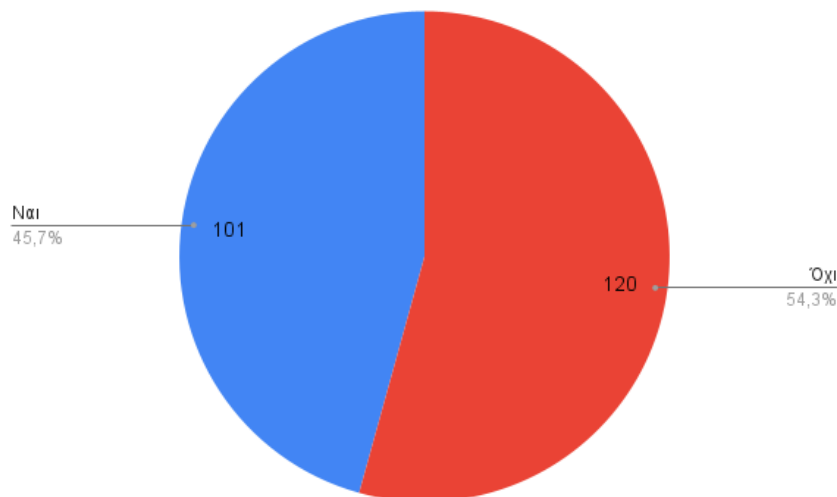
Σχήμα 6.38: Πιο συχνά εντοπίζετε μία νέα δημοσίευση από το χρονολόγιο (timeline) ή μέσω των instagram ιστοριών (stories); (45+)

6.1.17 Έχετε ανακαλύψει ποτέ νέες δημοσιεύσεις αναζητώντας κάποιο hashtag στο Instagram;

Υποστηρικτικό ρόλο στην απόδοση μιας δημοσίευσης κατέχει, επίσης, η χρήση λέξεων-κλειδιών (*hashtags*), τα οποία όπως αναφέρθηκε σε προγενέστερο στάδιο της έρευνας, βοηθηθούν την ανάρτηση να εμφανιστεί σε ένα ευρύτερο κοινό με αντίστοιχα ενδιαφέροντα και, κατά συνέπεια, των δημιουργό περιεχομένου να αναδειχθεί και να προσελκύσει νέος ακόλουθους (βλ. Σχήματα 5.51 έως 5.55). Με γνώμονα τα παραπάνω, η παρούσα ερώτηση αποσκοπούσε στην εξέταση του κατά πόσο το ελληνικό κοινό αξιοποιεί τις λέξεις-κλειδιά, έτσι ώστε να ανακαλύψει νέο περιεχόμενο.

Γενικά

Είναι ορατό στο Σχήμα 6.39, πως ένα σημαντικό μέρος των ερωτηθέντων (54,3%) αποκρίθηκε αρνητικά στην ερώτηση, δείχνοντας, κατά αυτόν τον τρόπο, πως οι Έλληνες χρήστες δεν διαθέτουν έναν αρκετά υψηλό βαθμό εξοικείωσης με τον συγκεκριμένο τρόπο ανακάλυψης δημοσιεύσεων.



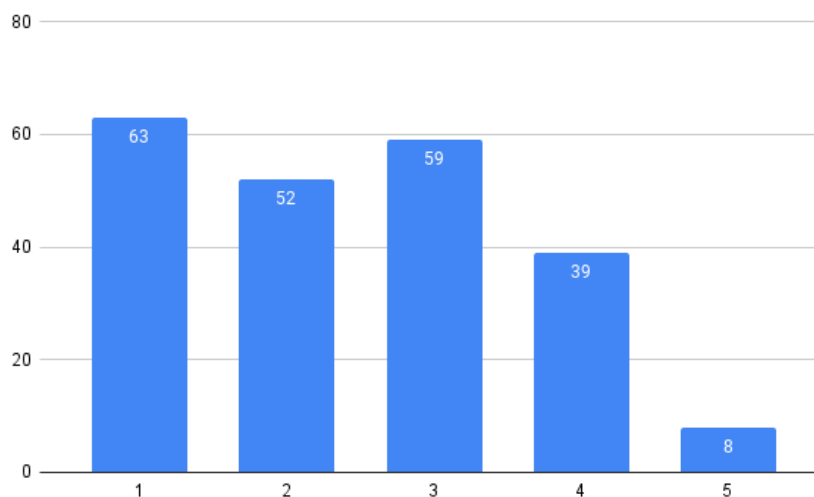
Σχήμα 6.39: Έχετε ανακαλύψει ποτέ νέες δημοσιεύσεις αναζητώντας κάποιο hashtag στο Instagram;

6.1.18 Κάνετε repost (επαναδημοσίευση) δημοσιεύσεις που σας αρέσουν;

Πέραν των ενεργειών προώθησης του δημοσιευμένου περιεχομένου, εκ μέρους των δημιουργών, εξίσου κομβικός είναι ο ρόλος του κοινού στην ανάδειξη μιας δημοσίευσης, μέσω της κοινοποίησης της στις ιστορίες της πλατφόρμας. Ως εκ τούτου, η ενεργή υποστήριξη του ελληνικού κοινού προς το αγαπημένο τους περιεχόμενο εξετάζεται στα πλαίσια της προκείμενης ερώτησης.

Γενικά

Λαμβάνοντας υπόψη τα δεδομένα του Σχήματος 6.40, παρατηρείται πως οι Έλληνες χρήστες διατήρησαν μια αμφίρροπη στάση, όσον αφορά στην επαναδημοσίευση των αγαπημένων τους αναρτήσεων. Ειδικότερα, 115 από τους 221 ερωτηθέντες χρήστες του Instagram παρουσίασαν λιγότερο υποστηρικτικοί απέναντι στο αγαπημένο τους περιεχόμενο, ενώ οι 98 από αυτούς δήλωσαν ότι κοινοποιούν δημοσιεύσεις οι οποίες τους αρέσουν, μερικές φορές έως και αρκετά συχνά. Παρόλα αυτά, πολύ μικρή είναι η μερίδα των χρηστών (8 χρήστες) η οποία διατηρεί μια σταθερά υποστηρικτική σχέση στους δημιουργούς περιεχομένου, κοινοποιώντας πάντα τις δημοσιεύσεις που προσελκύουν το ενδιαφέρον τους.



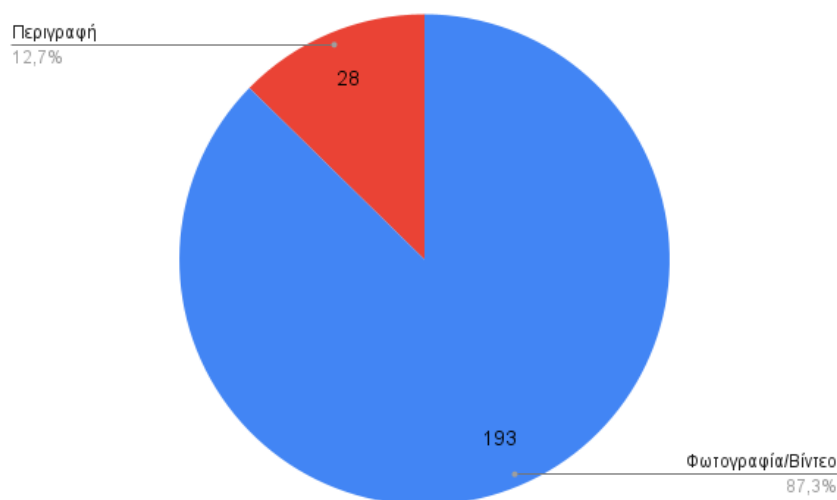
Σχήμα 6.40: Κάνετε repost (επαναδημοσίευση) δημοσιεύσεις που σας αρέσουν;

6.1.19 Τι κοιτάτε πρώτα σε μία δημοσίευση (timeline post);

Όπως αναφέρθηκε παραπάνω, η πλατφόρμα του Instagram στεγάζει καθημερινά έναν τεράστιο φόρτο νέων δημοσιεύσεων, με αποτέλεσμα οι δημιουργοί να αναζητούν τρόπους για να ξεχωρίσει το περιεχόμενό τους από το πλήθος. Με αφορμή το γεγονός αυτό, αναζητήθηκε το χαρακτηριστικό μιας δημοσίευσης το οποίο είναι δυνατό να συγκεντρώσει ευκολότερα τα βλέμματα του κοινού.

Γενικά

Είναι εμφανές από το Σχήμα 6.41 ότι η απεικόνιση του περιεχομένου μιας δημοσίευσης παίζει καθοριστικό ρόλο στην απόδοση μιας δημοσίευσης, καθώς η συντριπτική πλειοψηφία των ερωτηθέντων δήλωσε πως το χαρακτηριστικό το οποίο τραβάει πρώτα την προσοχή τους σε μία δημοσίευση είναι η οπτική αναπαράσταση της (87,3%), ενώ σε δεύτερη μοίρα έρχεται η περιγραφή της (12,7%).



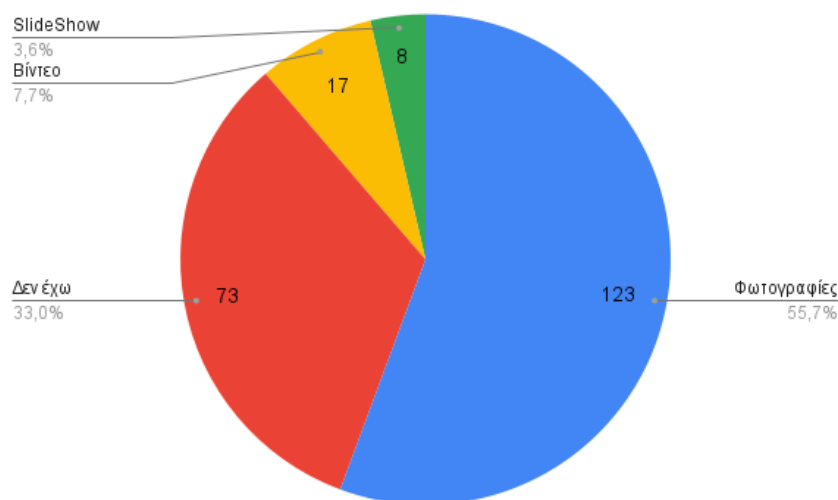
Σχήμα 6.41: Τι κοιτάτε πρώτα σε μία δημοσίευση (timeline post);

6.1.20 Ποια μορφή δημοσίευσης (timeline post) προτιμάτε;

Συνεχίζοντας την ανάλυση των προτιμήσεων του ελληνικού κοινού, το ενδιαφέρον στράφηκε στη μελέτη της συμπάθειας των χρηστών για τους διαφορετικούς τύπους αναρτήσεων στο Instagram (φωτογραφία, βίντεο μικρής και μεγάλης διάρκειας, slideshow).

Γενικά

Γίνεται αντιληπτό από το Σχήμα 6.42 πως οι Έλληνες χρήστες στην πλειονότητα τους (55,7%) προτιμούν το φωτογραφικό περιεχόμενο, ενώ λιγότεροι είναι εκείνοι οι οποίοι τάχθηκαν υπέρ των βίντεο (7,7%) και της παρουσίασης πολλαπλού περιεχομένου/slideshow (3,6%). Εντούτοις, το 33% των χρηστών δήλωσε πως δεν τρέφει κάποια ιδιαίτερη προτίμηση προς έναν συγκεκριμένο τύπο δημοσιεύσεων. Αντίστοιχο θέμα προς διερεύνηση είχε τεθεί στο Σχήμα 5.38 της Ενότητας 5.1.



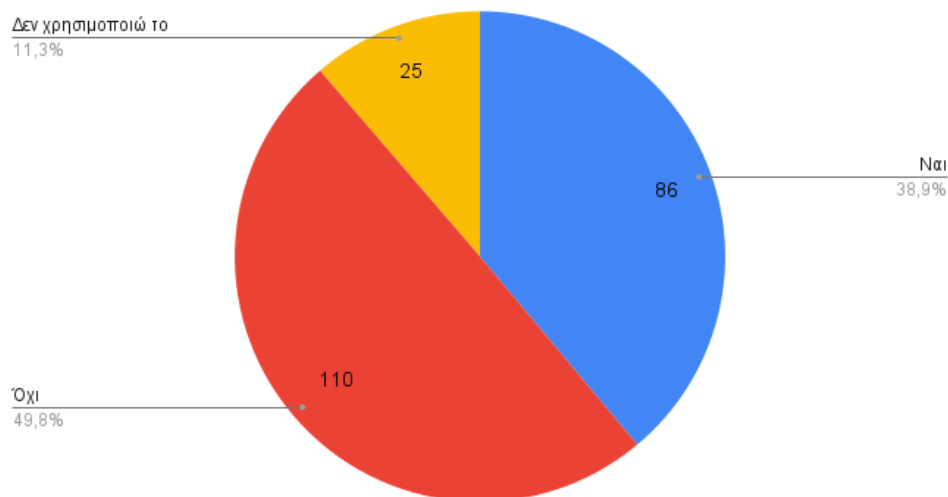
Σχήμα 6.42: Ποια μορφή δημοσίευσης (timeline post) προτιμάτε;

6.1.21 Ακολουθείτε τους/τις αγαπημένους/νες σας YouTubers στο Instagram;

Αρκετοί δημιουργοί περιεχομένου, εκμεταλλευόμενοι τη μεγάλη άνθηση των μέσων κοινωνικής δικτύωσης, αναπτύσσουν τη διαδικτυακή τους παρουσία σε πολλαπλές πλατφόρμες, έτσι ώστε να προσεγγίσουν νέο κοινό. Μεταξύ αυτών βρίσκεται και το YouTube, καθώς πολλοί είναι οι δημιουργοί βίντεο, οι οποίοι έχουν επεκτείνει τη δραστηριότητα τους στο Instagram, έτσι ώστε να διατηρήσουν ενεργό το ενδιαφέρον του κοινού τους. Η συγκεκριμένη ερώτηση αποσκοπεί στην εξέταση του βαθμού κατά τον οποίο οι Έλληνες δημιουργοί βίντεο έχουν καταφέρει να μεταφέρουν το κοινό τους στην πλατφόρμα του Instagram.

Γενικά

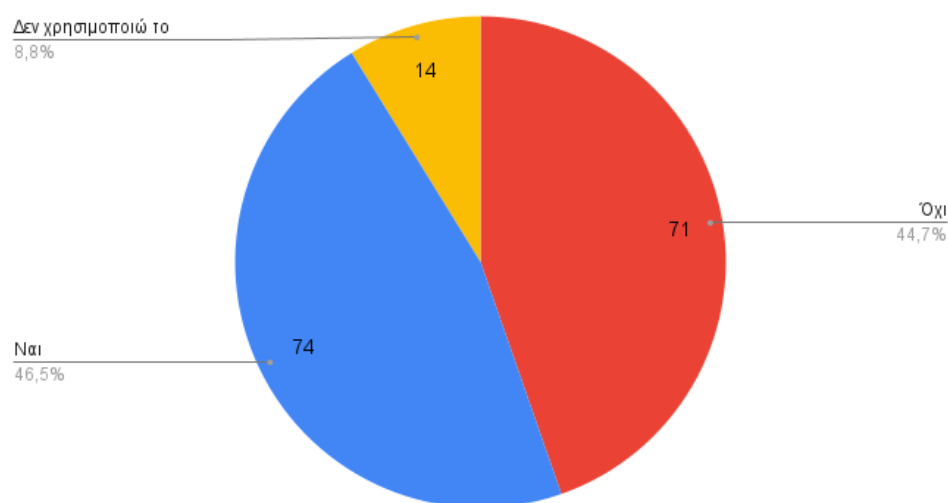
Εξαιρώντας το υποσύνολο των χρηστών του Instagram οι οποίοι δεν χρησιμοποιούν το YouTube (11,3%), διαπιστώνεται μέσω του Σχήματος 6.43 ότι, κατά προσέγγιση, ένας στους δύο χρήστες του Instagram δεν ακολουθεί τους αγαπημένους του YouTubers, γεγονός το οποίο αναδεικνύει τα περιθώρια των δημιουργών βίντεο να διευρύνουν ακόμα περισσότερο το κοινό τους στην πλατφόρμα του Instagram, ενθαρρύνοντας του εγγεγραμμένους χρήστες των καναλιών τους να τους ακολουθήσουν.



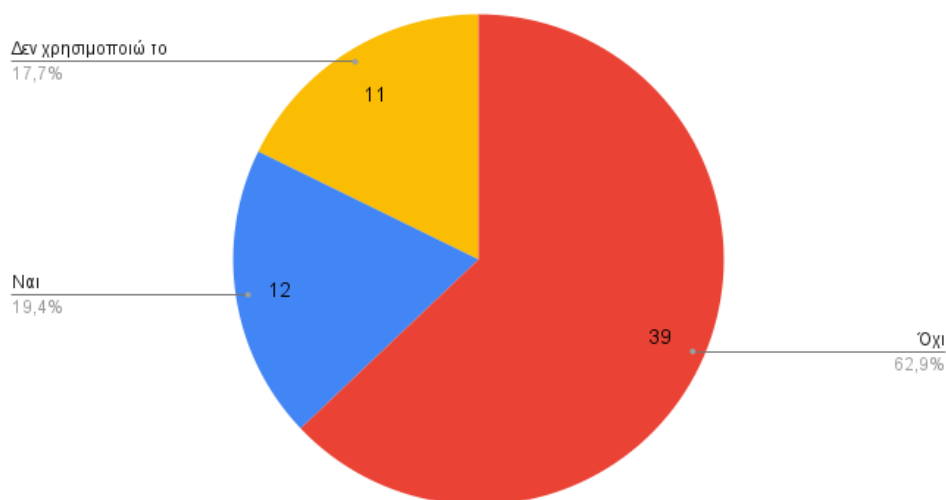
Σχήμα 6.43: Ακολουθείτε τους/τις αγαπημένους/νες σας YouTubers στο Instagram;

Ηλικιακές ομάδες

Στο επίπεδο των ηλικιών των ερωτηθέντων, εξαιρώντας και πάλι το υποσύνολο των χρηστών του Instagram οι οποίοι δεν χρησιμοποιούν το YouTube, είναι δυνατό να παρατηρηθεί στα Σχήματα 6.44 και 6.45 πως οι γηραιότεροι χρήστες της πλατφόρμας του Instagram απέχουν σε μεγαλύτερο βαθμό, σε σχέση με τους νεότερους, από την ακολούθηση των αγαπημένων τους δημιουργών βίντεο στο Instagram. Το συγκεκριμένο φαινόμενο, πιθανώς, απορρέει από τη χαμηλότερη εξοικείωση των μεγαλύτερων ηλικιακά χρηστών με τα μέσα κοινωνικής δικτύωσης, αλλά και, όπως θα φανεί παρακάτω, από τη μικρότερη τάση αυτών να εγγράφονται στα κανάλια δημιουργών και, κατά συνέπεια, να ταυτίζονται με συγκεκριμένους δημιουργούς βίντεο.



Σχήμα 6.44: Ακολουθείτε τους/τις αγαπημένους/νες σας YouTubers στο Instagram; (18-34)



Σχήμα 6.45: Ακολουθείτε τους/τις αγαπημένους/νες σας YouTubers στο Instagram; (35+)

6.2 Ερωτήσεις YouTube

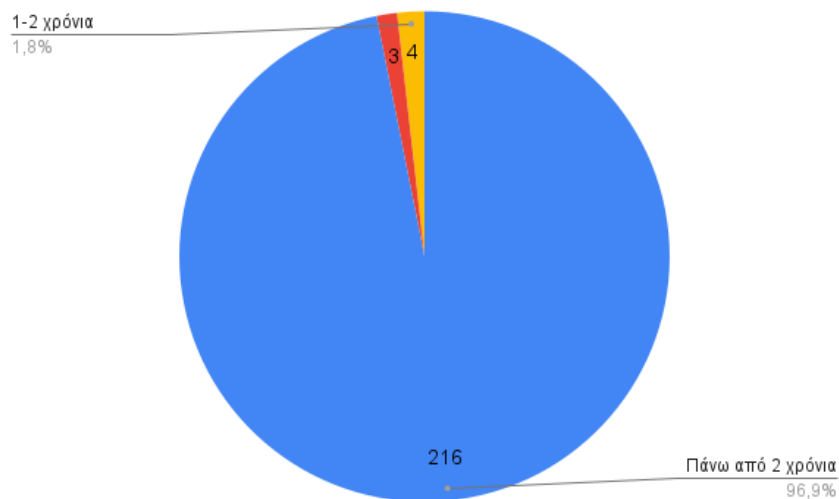
Στη συγκεκριμένη υποενότητα αναλύονται οι απαντήσεις των ερωτηθέντων στις ερωτήσεις του ερωτηματολογίου, οι οποίες αφορούσαν στη πλατφόρμα του YouTube.

6.2.1 Πόσο καιρό χρησιμοποιείτε το YouTube;

Η πρώτη σχετική με το YouTube ερώτηση, η οποία τέθηκε στους συμμετέχοντες του ερωτηματολογίου, πραγματεύεται το ιστορικό της δραστηριότητας του εκάστοτε χρήστη στην πλατφόρμα και, κατ' επέκταση, το ποσοστό εξοικείωσης του με αυτή.

Γενικά

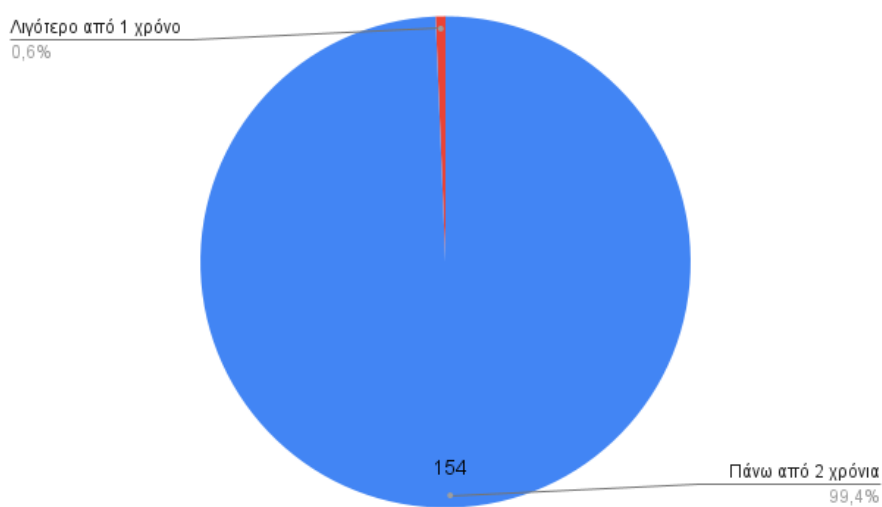
Αναφορικά με το γενικό σύνολο των απαντήσεων, προκύπτει από το Σχήμα 6.46 ότι η συντριπτική πλειοψηφία των ερωτηθέντων (96,9%) χρησιμοποιούσαν το YouTube για παραπάνω από δύο χρόνια, ενώ ένα μικρό υποσύνολο αυτών (1,8%) είχε εισέλθει στη διαδικτυακή κοινότητα της συγκεκριμένης πλατφόρμας κατά τα τελευταία δύο χρόνια (3,1%). Γίνεται, λοιπόν, αντιληπτό, πως το ελληνικό κοινό είναι εξοικειωμένο κατά ένα μεγάλο βαθμό με το τρόπο λειτουργίας της πλατφόρμας του Youtube.



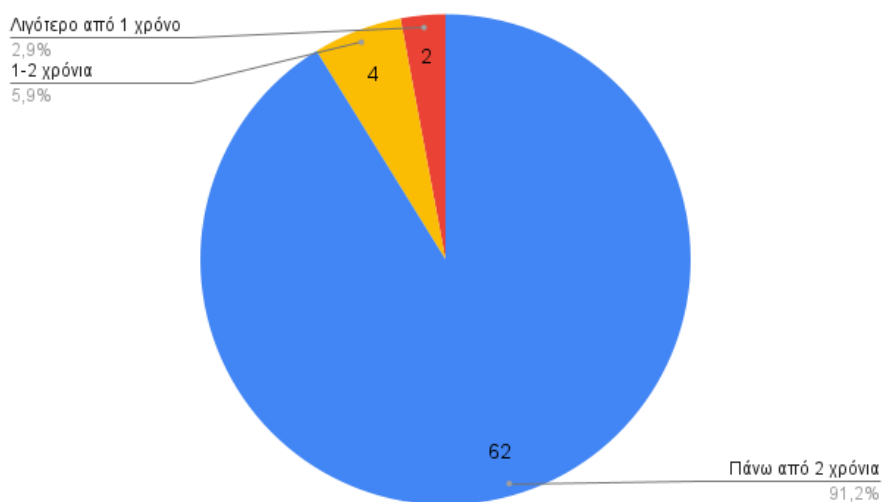
Σχήμα 6.46: Πόσο καιρό χρησιμοποιείτε το YouTube;

Ηλικιακές Ομάδες

Αναλύοντας τις απαντήσεις στο επίπεδο των ηλικιών των ερωτηθέντων, προκύπτει από τα Σχήματα 6.47 και 6.48 ότι, παρά το υψηλότερο ποσοστό χρηστών μεγαλύτερης ηλικίας οι οποίοι εισήλθαν στην πλατφόρμα του YouTube κατά τελευταία δύο έτη, η γενικότερη πλειοψηφία των ερωτηθέντων χρησιμοποιούσε το YouTube για περισσότερα από δύο χρόνια, επιβεβαιώνοντας τον μεγάλο βαθμό εξοικείωσης της ελληνικής κοινότητας με τη συγκεκριμένη πλατφόρμα.



Σχήμα 6.47: Πόσο καιρό χρησιμοποιείτε το YouTube; (18-34)



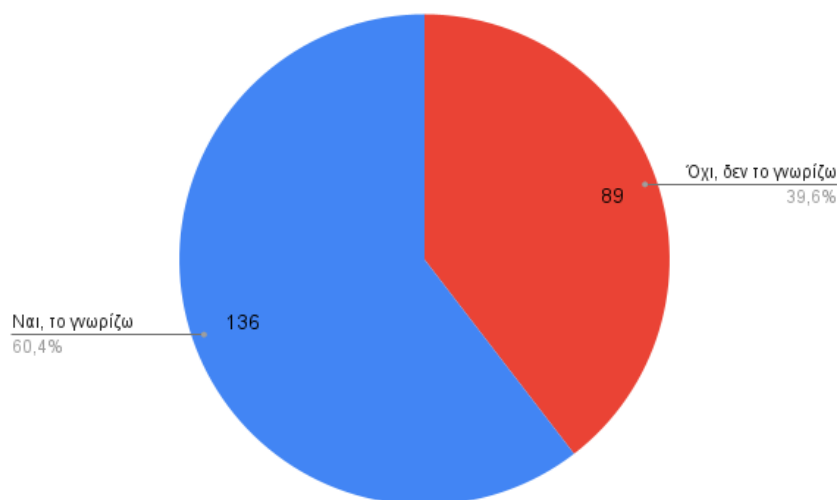
Σχήμα 6.48: Πόσο καιρό χρησιμοποιείτε το YouTube; (35+)

6.2.2 Γνωρίζετε ότι το YouTube αποτελεί ένα ανερχόμενο μέσο κοινωνικής δικτύωσης, κάνοντας διαθέσιμη τη δημιουργία προσωπικών post και stories;

Το YouTube, ακολουθώντας την τάση του κοινού προς τα μέσα κοινωνικής δικτύωσης, τα τελευταία χρόνια βρίσκεται σε μία προσπάθεια προσθήκης νέων χαρακτηριστικών στην πλατφόρμα, επηρεασμένο από δημοφιλείς πλατφόρμες δικτύωσης, όπως για παράδειγμα οι ιστορίες και οι αναρτήσεις στο χρονολόγιο. Η παρούσα ερώτηση επιδίωξε να εξετάσει το κατά πόσο η μεταβολή του παραδοσιακού χαρακτήρα του YouTube έχει γίνει αντιληπτή από το ελληνικό κοινό και να διασαφηνίσει αν αυτό αντιμετωπίζει την πλατφόρμα πλέον ως ένα μέσο κοινωνικής δικτύωσης ή απλά ως ένα μέσο αναπαραγωγής βίντεο.

Γενικά

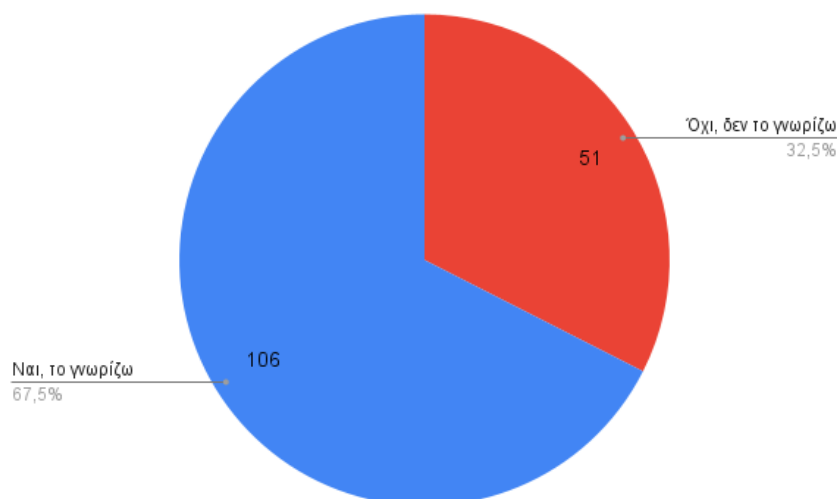
Είναι, λοιπόν, φανερό στο Σχήμα 6.49, πως μία μεγάλη μερίδα των Ελλήνων χρηστών του YouTube (60,4%) έχει αντιληφθεί τη σταδιακή μετατροπή της πλατφόρμας σε ένα μέσο κοινωνικής δικτύωσης. Από την άλλη μεριά, ένα σεβαστό μέρος των ερωτηθέντων (39,4%) δεν γνωρίζει για την αλλαγή αυτή, δείχνοντας έτσι πως το YouTube διατηρεί έναν μεικτό χαρακτήρα και, πιθανώς, βρίσκεται ακόμα σε ένα μεταβατικό στάδιο.



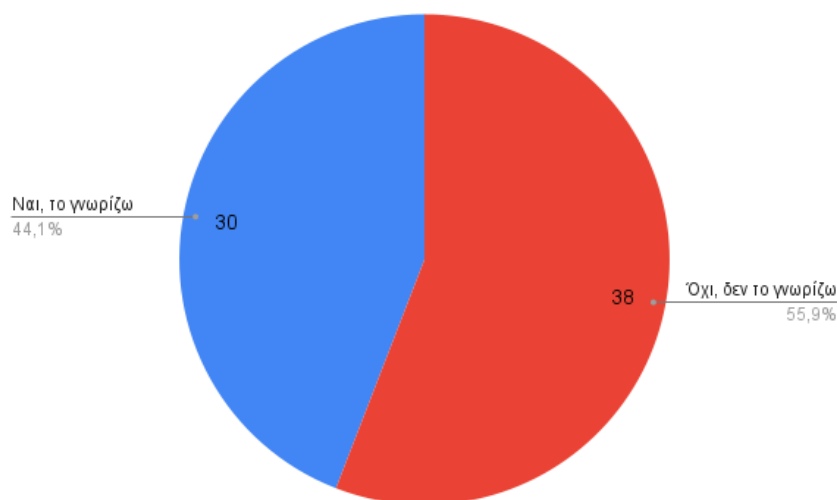
Σχήμα 6.49: Γνωρίζετε ότι το YouTube αποτελεί ένα ανερχόμενο μέσο κοινωνικής δικτύωσης, κάνοντας διαθέσιμη τη δημιουργία προσωπικών post και stories;

Ηλικιακές Ομάδες

Λαμβάνοντας υπόψη τις ηλικίες των ερωτηθέντων και με βάση τα Σχήματα 6.50 και 6.51, διαφαίνεται ότι νεότεροι σε ηλικία χρήστες του YouTube διαθέτουν μία καλύτερη αντίληψη σε σχέση με τους γηραιότερους, όσον αφορά στον εκσυγχρονισμό και τη σταδιακή εξέλιξη της πλατφόρμας σε ένα μέσο κοινωνικής δικτύωσης. Το γεγονός αυτό θεωρούταν αναμενόμενο, δεδομένης της περιορισμένης οικειότητας των ατόμων μεγαλύτερης ηλικίας με τα σύγχρονα μέσα κοινωνικής δικτύωσης.



Σχήμα 6.50: Γνωρίζετε ότι το YouTube αποτελεί ένα ανερχόμενο μέσο κοινωνικής δικτύωσης, κάνοντας διαθέσιμη τη δημιουργία προσωπικών post και stories; (18-34)



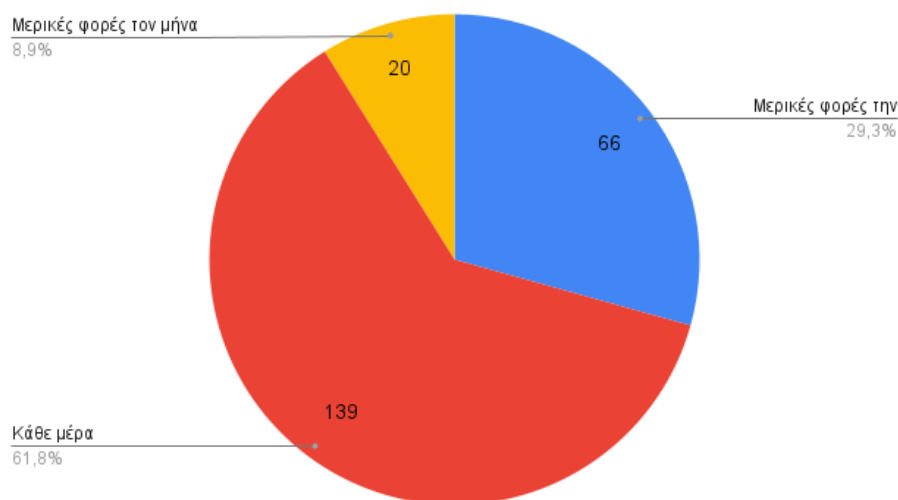
Σχήμα 6.51: Γνωρίζετε ότι το YouTube αποτελεί ένα ανερχόμενο μέσο κοινωνικής δικτύωσης, κάνοντας διαθέσιμη τη δημιουργία προσωπικών post και stories; (35+)

6.2.3 Πόσο συχνά παρακολουθείτε βίντεο στο YouTube;

Η συγκεκριμένη ερώτηση αποσκοπούσε στην εξέταση του ποσοστού χρήσης της πλατφόρμας από τους χρήστες και, κατά συνέπεια, στην αποτύπωση της επιρροής της προς αυτούς.

Γενικά

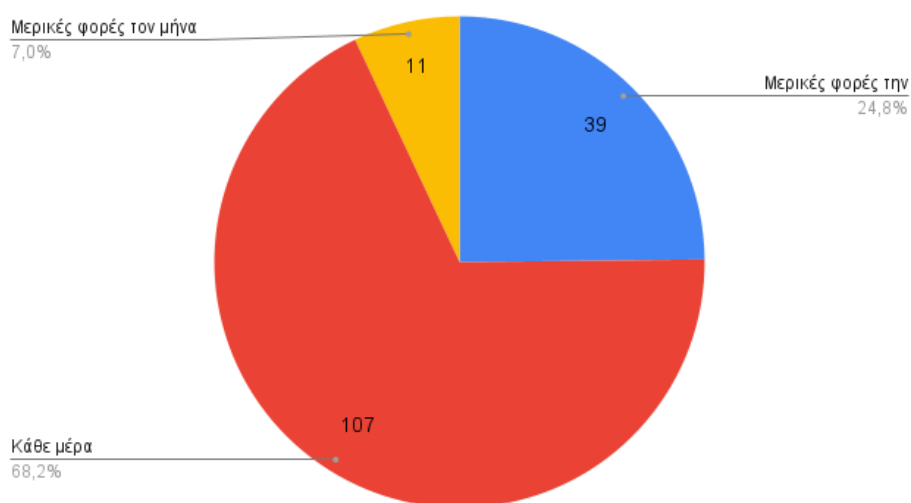
Με γνώμονα το Σχήμα 6.52, διαπιστώνεται πως η πλειονότητα των χρηστών (61,8%) πλοηγείται σε καθημερινή βάση στη πλατφόρμα, ενώ το 29,3% των χρηστών διατηρεί παρόμοια επίπεδα δραστηριότητας, δηλώνοντας πως χρησιμοποιεί το YouTube μερικές φορές την εβδομάδα. Μικρότερος είναι ο αριθμός των ατόμων τα οποία παραμένουν λιγότερο ενεργά στο Instagram (μερικές φορές τον μήνα), με αυτά να καταλαμβάνουν το 8,9% των δοθέντων απαντήσεων. Είναι εμφανές, επομένως, πως η πλατφόρμα του YouTube, όπως και του Instagram, διατηρεί αμείωτο το ενδιαφέρον της πλειοψηφίας των Ελλήνων χρηστών, με την τακτική δραστηριότητα τους να αποτελεί ένδειξη της υψηλής επιρροής της σε αυτούς.



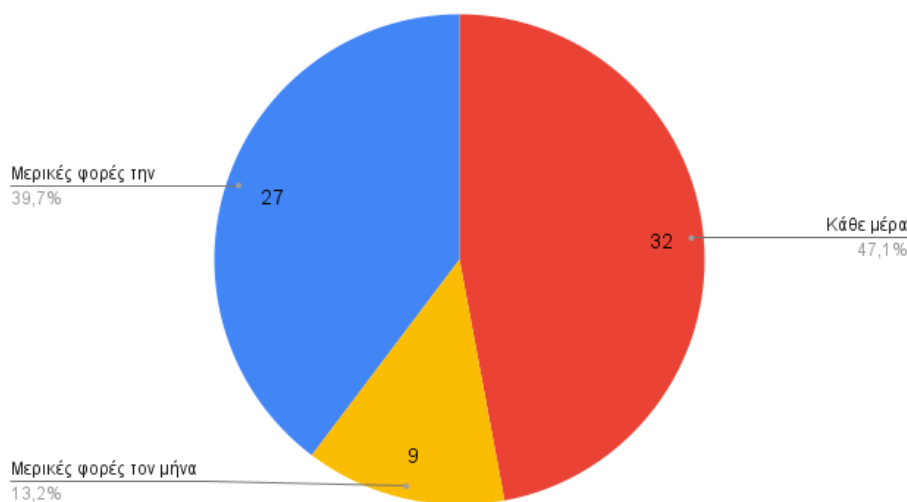
Σχήμα 6.52: Πόσο συχνά παρακολουθείτε βίντεο στο YouTube;

Ηλικιακές Ομάδες

Κατηγοριοποιώντας τις απαντήσεις με βάση τις ηλικίες των ερωτηθέντων, είναι ορατό στα Σχήματα 6.53 και 6.54 ότι η εντονότερη δραστηριότητα στο YouTube ασκείται από χρήστες μικρότερης ηλικίας (18-34), καθώς το 68,2% αυτών δήλωσε ότι περιηγείται. Από την άλλη πλευρά μεγαλύτεροι ηλικιακά χρήστες φαίνεται να διατηρούν μία περισσότερο ήπια δραστηριότητα, καθώς ένα σημαντικό μέρος αποκλίνει από την καθημερινή χρήση της πλατφόρμας και δραστηριοποιείται στο επίπεδο της εβδομάδας (39,7%), αλλά και του μήνα (13,2%).



Σχήμα 6.53: Πόσο συχνά παρακολουθείτε βίντεο στο YouTube; (18-34)



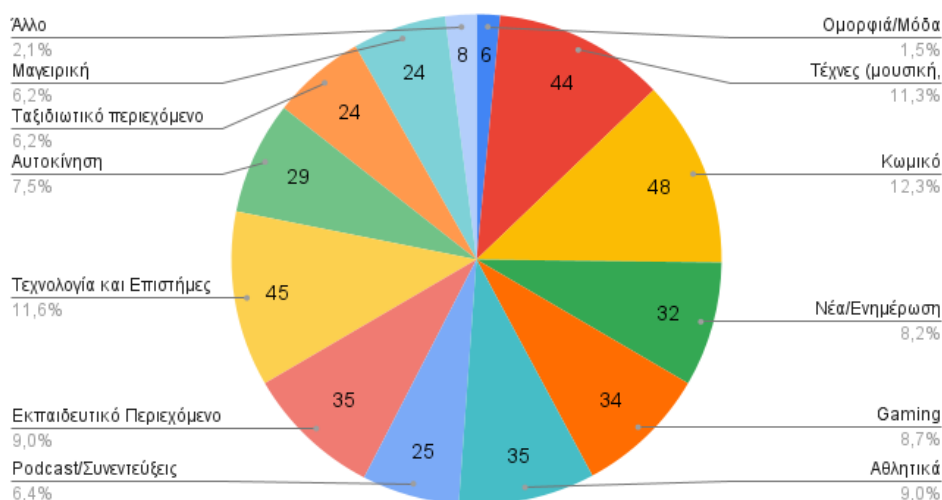
Σχήμα 6.54: Πόσο συχνά παρακολουθείτε βίντεο στο YouTube; (35+)

6.2.4 Τι περιεχόμενο προτιμάτε να παρακολουθείτε στο YouTube;

Μια ακριβέστερη αποτύπωση των προτιμήσεων της ελληνικής κοινότητας του YouTube, σε σχέση με αυτήν η οποία προέκυψε από τα σχήματα του Κεφαλαίου 5 και συγκεκριμένα της Ενότητας 5.2, θα μπορούσε να παρέχει η άμεση ερώτηση του κοινού για το ποιες είναι οι θεματικές κατηγορίες περιεχομένου τις οποίες προτιμά να παρακολουθεί στα πλαίσια της πλατφόρμας. Καθώς, όπως φάνηκε στο Κεφάλαιο 5, οι τάσεις των ανδρών και των γυναικών εμφάνισαν αρκετές διαφορές κατά το έτος 2020, κρίθηκε σημαντική η ξεχωριστή μελέτη των απαντήσεων οι οποίες προήλθαν από τα δύο αυτά φύλα.

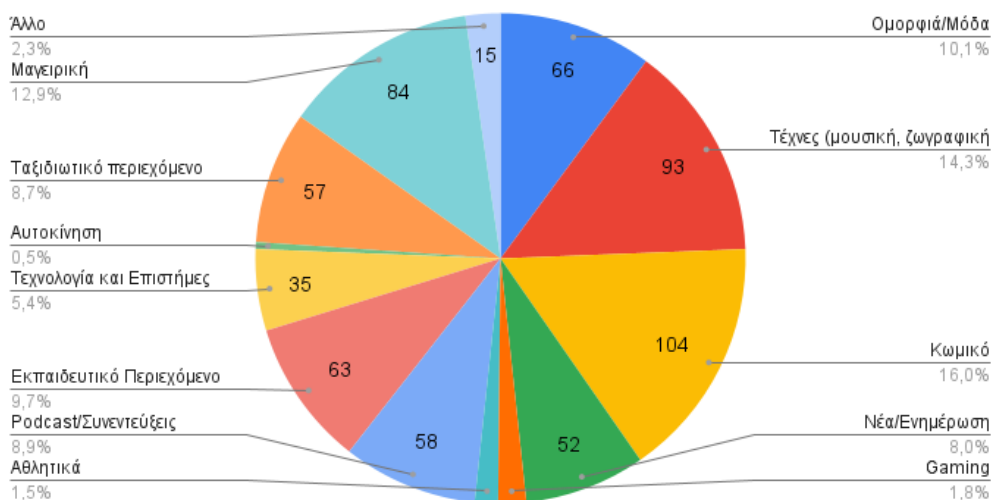
Άνδρες-Γυναίκες

Όσον αφορά στους άνδρες ερωτηθέντες, εκείνοι παρουσίασαν μία ιδιαίτερη κλίση προς το κωμικό περιεχόμενο (12,3%) και το τεχνολογικό/επιστημονικό περιεχόμενο (11,6%), ενώ αρκετά υψηλά στις προτιμήσεις τους βρίσκονται, επίσης, οι τέχνες (11,3%), το αθλητικό περιεχόμενο (9%), το εκπαιδευτικό περιεχόμενο (9%) και το ηλεκτρονικό παιχνίδι (8,7%), αποτελέσματα τα οποία παραπέμπουν στα αντίστοιχα συμπεράσματα της Ενότητας 5.2



Σχήμα 6.55: Τι περιεχόμενο προτιμάτε να παρακολουθείτε στο YouTube; (Άνδρες)

Από την άλλη μεριά, οι γυναίκες της πλατφόρμας ανέδειξαν το κωμικό περιεχόμενο (16%), τις τέχνες (14,3%) και τη μαγειρική (12,9%) ως τις περισσότερο αρεστές κατηγορίες περιεχομένου, ενώ, πόλους έλξης του ενδιαφέροντος τους φαίνεται πως αποτέλεσαν, επίσης, η μόδα/ομορφιά (10,1%) και το εκπαιδευτικό περιεχόμενο (9,7%).



Σχήμα 6.56: Τι περιεχόμενο προτιμάτε να παρακολουθείτε στο YouTube; (Γυναίκες)

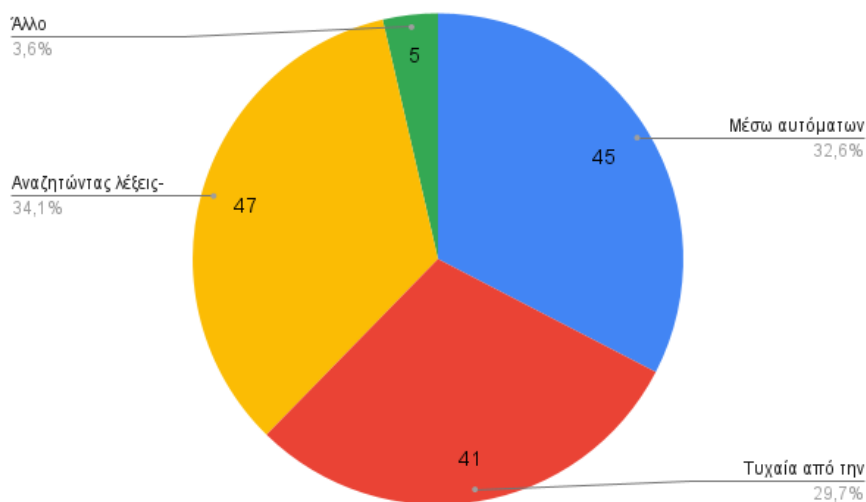
6.2.5 Με ποιον τρόπο ανακαλύπτετε νέα βίντεο;

Καθημερινά, οι δημιουργοί βίντεο του YouTube παράγουν νέο υλικό και, παράλληλα, το κοινό κατακλύζεται από μία πληθώρα νέων μεταφορτώσεων διαθέσιμων προς παρακολούθηση. Το γεγονός αυτό έχει ως αποτέλεσμα οι δημιουργοί να μάχονται με διάφορα μέσα, όπως χρήση λέξεων-κλειδιών, δημιουργία έξυπνων τίτλων ή εγκαθίδρυση ενός εγγεγραμμένου κοινού έτσι ώστε τα βίντεο τους να φτάσουν στις οθόνες των χρηστών της πλατφόρμας, ακόμα και των

ίδιων τους των πιστών ακολούθων. Με σκοπό την εξαγωγή συμπερασμάτων, όσον αφορά στους τρόπους με τους οποίους το κοινό αναζητά ή ανακαλύπτει, κρίθηκε αναγκαία η σύνταξη της προκειμένης ερώτησης. Αξίζει να σημειωθεί πως η διερεύνηση του παρόντος ερωτήματος έγινε ξεχωριστά για άνδρες, γυναίκες και για τις ποικίλες ηλικιακές ομάδες, έτσι ώστε να εξαχθεί μια πιο ξεκάθαρη εικόνα της συμπεριφορά του κάθε φύλου, αλλά και χρηστών από ολόκληρο το ηλικιακό φάσμα, στο επίπεδο της αναζήτησης.

Άνδρες-Γυναίκες

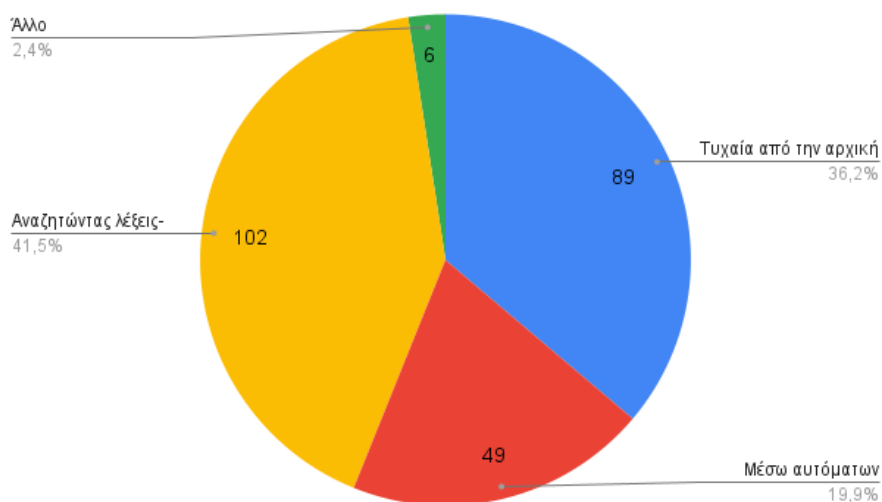
Στο επίπεδο των ανδρών χρηστών της πλατφόρμας, σύμφωνα με το Σχήμα 6.57, είναι ορατό πως ο πιο δημοφιλής τρόπος αναζήτησης νέων βίντεο αποτελεί η χρήση λέξεων-κλειδιών (34,1%), δηλαδή η αναζήτηση στοχευμένα λέξεων ή φράσεων, οι οποίες αφορούν στο επιθυμητό περιεχόμενο. Στη δεύτερη θέση της κατάταξης βρίσκεται η αναζήτηση μέσω αυτόματων ειδοποιήσεων (32,6%), κάτι το οποίο μαρτυρά πως οι άνδρες του, κατά ένα σεβαστό ποσοστό, κάνουν εγγραφή στα αγαπημένα τους κανάλια στο YouTube, με σκοπό να ενημερώνονται για νέες μεταφορτώσεις. Ως τελευταία επιλογή αναζήτησης, αλλά συγκεντρώνοντας ένα παρόμοιο πλήθος απαντήσεων με τις προαναφερθείσες επιλογές, παρουσιάζεται η τυχαία ανακάλυψη βίντεο στην αρχική σελίδα της πλατφόρμας (29,7%). Τέλος, μόλις το 3,6% των ερωτηθέντων δήλωσε πως ανακαλύπτει νέα βίντεο στο YouTube με διαφορετικό τρόπο από τους παραπάνω.



Σχήμα 6.57: Με ποιον τρόπο ανακαλύπτετε νέα βίντεο; (Άνδρες)

Από την άλλη μεριά, οι γυναίκες της πλατφόρμας φαίνεται πως λειτουργούν με διαφορετικό, εν μέρει, τρόπο στην αναζήτηση και ανακάλυψη νέων βίντεο, σε σχέση με τους άνδρες. Πιο συγκεκριμένα, το μεγαλύτερο ποσοστό στο διάγραμμα του Σχήματος 6.58 κατέχει η αναζήτηση με χρήση λέξεων-κλειδιών (41,5%), ενώ στη δεύτερη θέση βρίσκεται η εντοπισμός βίντεο μέσα από την αρχική σελίδα της πλατφόρμας (36,2%), γεγονός το οποίο φανερώνει πως οι γυναίκες προτιμούν να εξερευνούν την πλατφόρμα σε μεγαλύτερο βαθμό από τους άνδρες, οι οποίοι, όπως φάνηκε παραπάνω, έχουν την τάση να ενημερώνονται μέσω της εγγραφής τους σε επιλεγμένα κανάλια της αρεσκείας τους. Από την πλευρά των γυναικών, η ενημέρωση μέσω αυτόματων ειδοποιήσεων (εγγραφή) σε συγκεκριμένα κανάλια φαίνεται να αποτελεί λιγότερο δημοφιλή

τακτική (19,9%). Τέλος, μόλις το 2,4% των ερωτηθεισών γυναικών δήλωσε πως ανακαλύπτει νέα βίντεο στο YouTube με διαφορετικό τρόπο από τους παραπάνω.

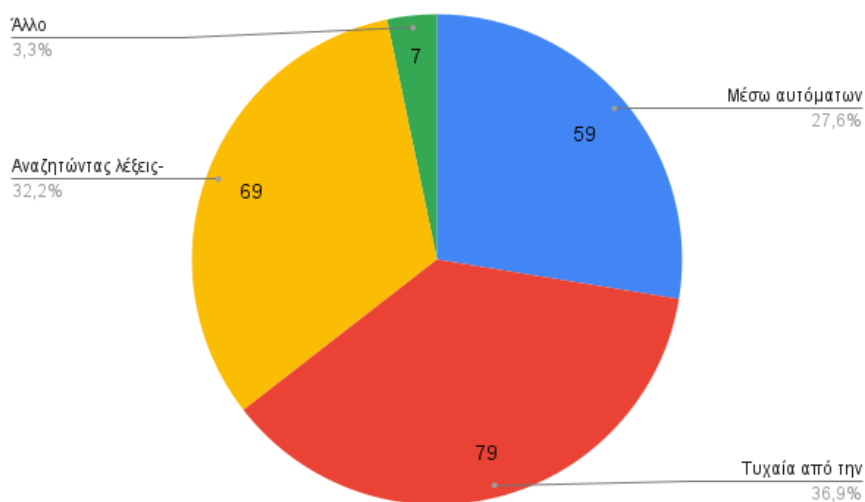


Σχήμα 6.58: Με ποιον τρόπο ανακαλύπτετε νέα βίντεο; (Γυναίκες)

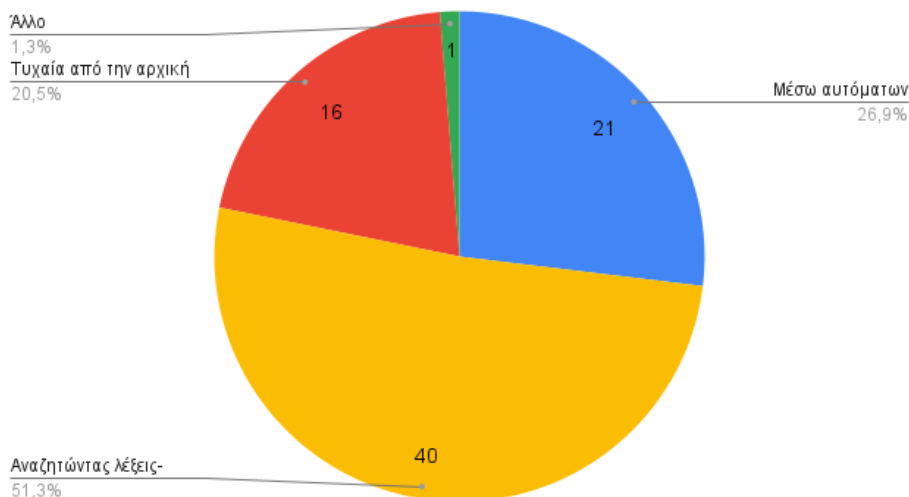
Συνοψίζοντας τα δύο παραπάνω διαγράμματα, η ευρεία αναζήτηση βίντεο μέσω της χρήσης λέξεων-κλειδιών και από τα δύο φύλλα αποδεικνύει τη σημαντικότητα της προσάρτησης των κατάλληλων ετικετών στα βίντεο εκ μέρους των δημιουργών, ενώ το σεβαστό ποσοστό των απαντήσεων οι οποίες αφορούν στη τυχαία ανακάλυψη βίντεο, μέσω της αρχικής του YouTube, υποδεικνύει την ανάγκη για χρήση έξυπνων φωτογραφιών και τίτλων σε νέα βίντεο. Το φαινόμενο αυτό είναι πιθανό να οφείλετε

Ηλικιακές Ομάδες

Με γνώμονα τα δεδομένα των Σχημάτων 6.59 και 6.60 προκύπτει ότι χρήστες μικρότερης ηλικίας συνηθίζουν να αναζητούν βίντεο τυχαία, από την αρχική σελίδα της πλατφόρμας. Η συγκεκριμένη τακτική φαίνεται να μην είναι το ίδιο διαδεδομένη στους μεγαλύτερους ηλικιακά χρήστες (35 και άνω), καθώς η πλειοψηφία εκείνων (51,3%) ανέδειξε τη στοχευμένη αναζήτηση βίντεο με χρήση λέξεων-κλειδιών ως την προτιμότερη. Το φαινόμενο αυτό είναι πιθανό να υποδεικνύει την περισσότερο αφηρημένη περιήγηση των νεαρών χρηστών στα περιεχόμενα της πλατφόρμας χωρίς κάποιον αυστηρά καθορισμένο στόχο αναζήτησης, απέναντι στη συντηρητική και στοχευμένη χρήση της πλατφόρμας εκ μέρους των μεγαλύτερων ηλικιακά χρηστών. Παρόλα αυτά, μία σεβαστή μερίδα και των δύο παραπάνω ηλικιακών ομάδων, 27,6% και 26,9% αντίστοιχα, δήλωσε πως ενημερώνεται για καινούργιες μεταφορτώσεις, κάνοντας εγγραφή σε αγαπημένα κανάλια και λαμβάνοντας αυτόματες ειδοποιήσεις για το νέο υλικό.



Σχήμα 6.59: Με ποιον τρόπο ανακαλύπτετε νέα βίντεο; (18-34)



Σχήμα 6.60: Με ποιον τρόπο ανακαλύπτετε νέα βίντεο; (35+)

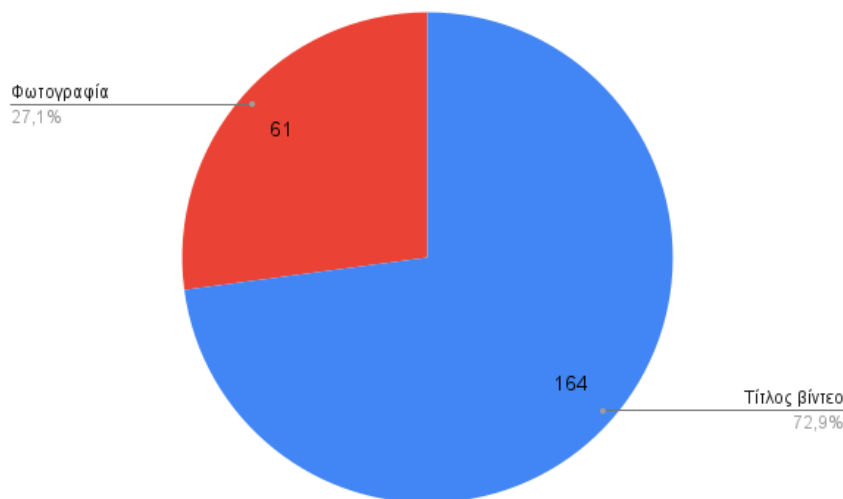
6.2.6 Τι κοιτάτε πρώτα σε ένα βίντεο;

Είναι ευρέως γνωστό, πλέον, όπως αναφέρθηκε παραπάνω, πως η πλατφόρμα του YouTube στεγάζει καθημερινά ένα τεράστιο φορτίο νέων βίντεο, με αποτέλεσμα οι δημιουργοί να αναζητούν τρόπους για να ξεχωρίσει το περιεχόμενό τους από το πλήθος. Με αφορμή το γεγονός αυτό, αναζητήθηκε το χαρακτηριστικό ενός βίντεο το οποίο είναι δυνατό να συγκεντρώσει ευκολότερα τα βλέμματα του κοινού.

Γενικά

Είναι εμφανές στο Σχήμα 6.61 ότι, σε αντίθεση με το Instagram, ο τίτλος ενός βίντεο παίζει καθοριστικό ρόλο στην απόδοσή του, καθώς η συντριπτική πλειοψηφία των ερωτηθέντων δήλωσε

πως το χαρακτηριστικό αυτό τραβάει πρώτα την προσοχή τους σε ένα βίντεο (72,9%), ενώ σε δεύτερη μοίρα έρχεται η φωτογραφία εξωφύλλου του βίντεο (27,1%).



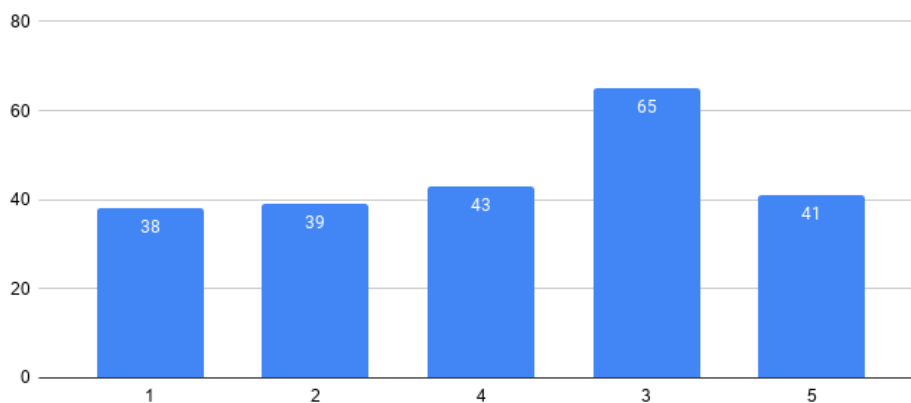
Σχήμα 6.61: Τι κοιτάτε πρώτα σε ένα βίντεο;

6.2.7 Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών;

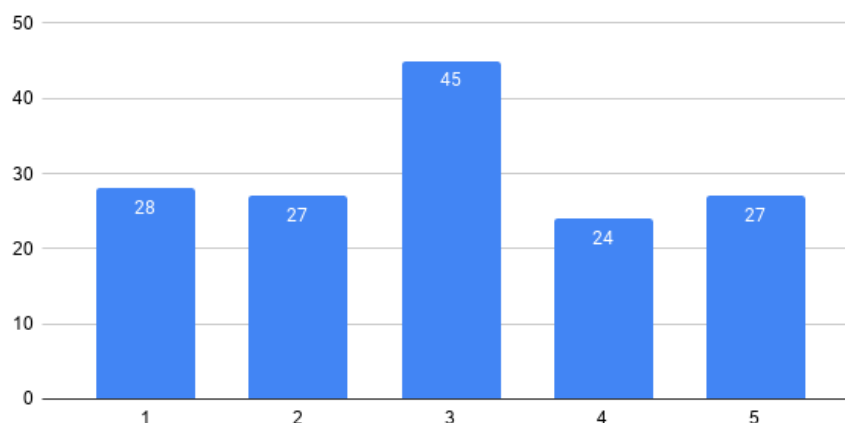
Η παρούσα ερώτηση συμπεριλήφθηκε στο ερωτηματολόγιο, με σκοπό την απόκτηση μιας ακριβέστερης εικόνας για τον βαθμό κατά τον οποίο το ελληνικό κοινό εγγράφεται στα κανάλια των αγαπημένων του δημιουργών, έτσι ώστε να ενημερώνεται άμεσα για τα νέα του βίντεο. Θεωρήθηκε ενδιαφέρον οι απαντήσεις της συγκεκριμένης ερώτησης να αναλυθούν ξεχωριστά για άνδρες, γυναίκες και για τις διαφορετικές ηλικιακές ομάδες.

Άνδρες-Γυναίκες

Σχολιάζοντας από κοινού τα Σχήματα 6.62 και 6.63, διαπιστώνεται πως οι άνδρες γίνονται ευκολότερα συνδρομητές στα αγαπημένα τους κανάλια σε σχέση με τις γυναίκες, φαινόμενο το οποίο είχε παρατηρηθεί επίσης στα Σχήματα 6.57 και 6.58 και, πιθανώς, συνδέεται με τον περιορισμένο αριθμό εγγεγραμμένων χρηστών των γυναικείων καναλιών (βλ. Σχήμα 5.80).



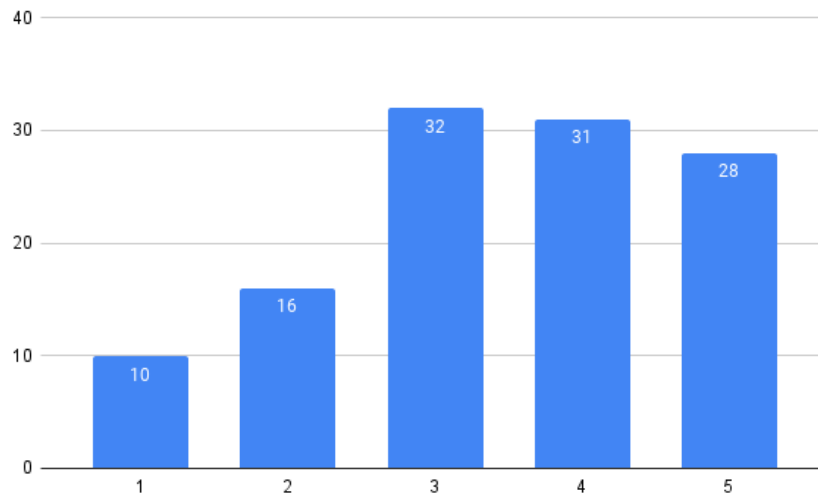
Σχήμα 6.62: Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (Άνδρες)



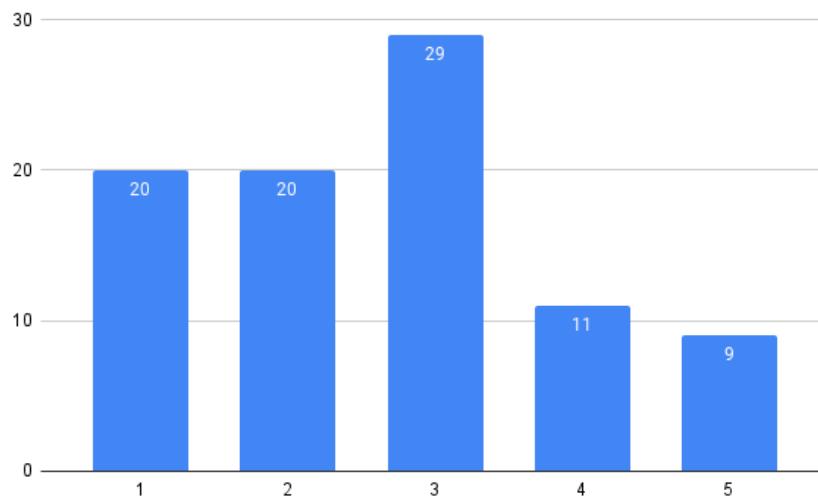
Σχήμα 6.63: Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (Γυναίκες)

Ηλικιακές Ομάδες

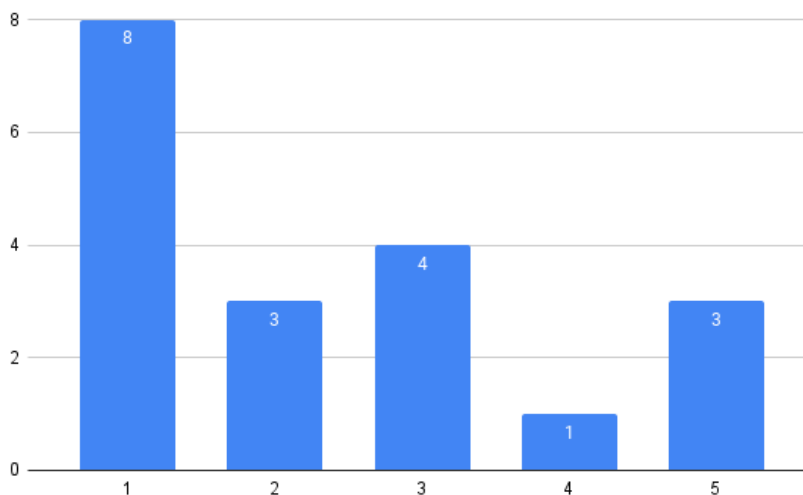
Όπως είχε προοικονομηθεί στα πλαίσια του Σχήματος 6.44, παρατηρείται στα τρία ακόλουθα σχήματος η τάση των μεγαλύτερων ηλικιακά χρηστών να μην εγγράφονται στα αγαπημένα τους κανάλια, σε αντίθεση με τους νεότερους χρήστες της πλατφόρμας, η πλειοψηφία των οποίων δήλωσε πως κάνει εγγραφή στα αγαπημένα κανάλια μερικές φορές έως και πάντα. Η προϋπόθεση της δημιουργίας λογαριασμού για την εγγραφή σε κάποιο κανάλι, σε συνδυασμό με τη δυνατότητα παρακολούθησης βίντεο χωρίς τη δημιουργία λογαριασμού και τη λιγότερη απασχόληση των γηραιότερων χρηστών στη πλατφόρμα (βλ. Σχήμα 6.54) αποτελεί, πιθανώς, έναν από τους λόγους οι οποίοι οδήγησαν στο συγκεκριμένο αποτέλεσμα. Πέραν των παραπάνω, υπάρχει το ενδεχόμενο το χαμηλότερο ποσοστό εγγραφών των μεγαλύτερων ηλικιακά χρηστών να συνδέεται με τη λιγότερο συνήθη ταύτιση αυτών με τους συγκεκριμένους δημιουργούς βίντεο.



Σχήμα 6.64: Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (18-24)



Σχήμα 6.65: Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (25-44)



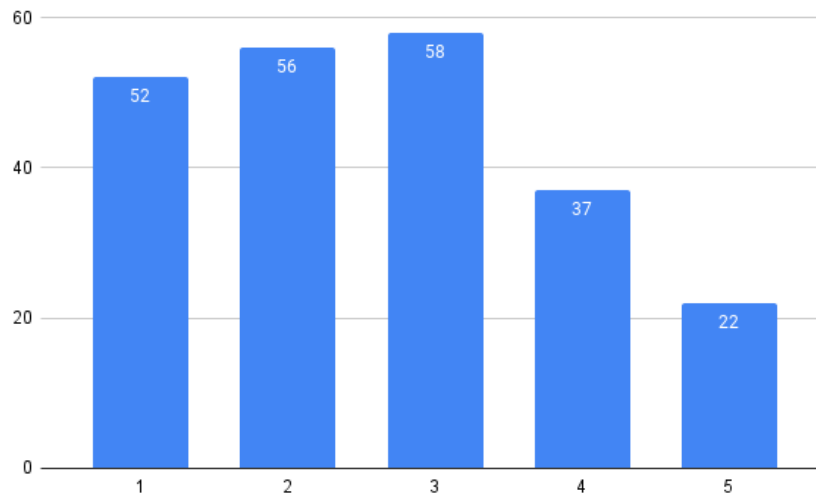
Σχήμα 6.66: Κάνετε Subscribe (συνδρομή) για να ενημερώνεστε για νέο περιεχόμενο των αγαπημένων σας καναλιών; (45+)

6.2.8 Κάνετε Like/Comment στα βίντεο που σας αρέσουν;

Στο επίπεδο της μελέτης της αλληλεπίδρασης του κοινού με τα δημοσιευμένα βίντεο, οι χρήστες του YouTube ερωτήθηκαν για το αν αντιδρούν θετικά ή σχολιάζουν σε βίντεο τα οποία τους ενδιαφέρουν.

Γενικά

Με βάση το Σχήμα 6.67, είναι φανερό ότι η ελληνική κοινότητα του YouTube διατηρεί μια σχετικά διστακτική θέση, όσον αφορά στην αλληλεπίδραση με βίντεο, καθώς η μεγαλύτερη μερίδα των ερωτηθέντων μερικές φορές έως αρκετά σπάνια αντιδρά θετικά ή σχολιάζει τα αγαπημένα της βίντεο, με μόλις 59 στους 224 να ανταποκρίνονται θετικά. Το φαινόμενο αυτό μεταφράζεται σε αρνητικό αντίκτυπο στις μετρικές απόδοσης των βίντεο των δημιουργών και χρίζει περαιτέρω παρακίνησης και καλύτερης προσέγγισης του κοινού εκ μέρους των δημιουργών, με σκοπό την θερμότερη ανταπόκριση στα βίντεο τους.



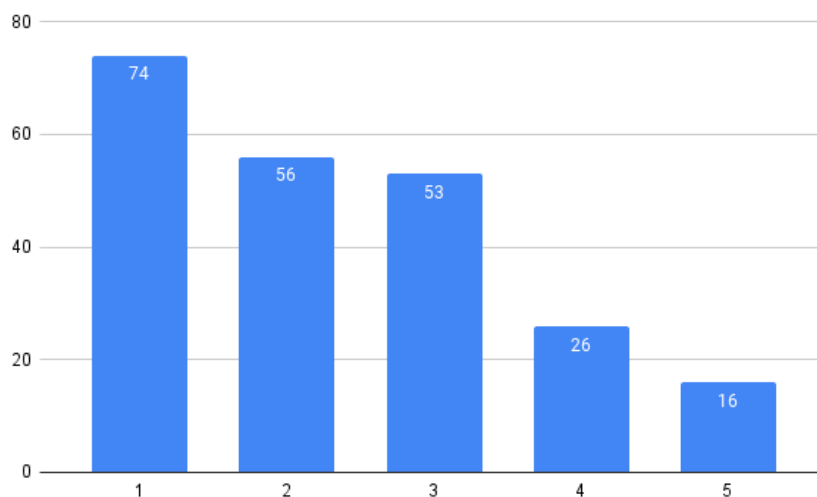
Σχήμα 6.67: Κάνετε Like ή Comment στα βίντεο που σας αρέσουν;

6.2.9 Κοινοποιείτε βίντεο που σας αρέσουν σε κάποιο άλλο μέσο κοινωνικής δικτύωσης (πχ Facebook, Instagram);

Ένας από τους σημαντικότερους τρόπους με τον οποίο ένας χρήστης μπορεί να αλληλεπιδράσει με ένα βίντεο είναι να το κοινοποιήσει στις στενές του επαφές σε κάποια άλλη πλατφόρμα κοινωνικής δικτύωσης, όπως το Facebook ή το Instagram, καθώς έτσι αυξάνεται η πιθανότητα να προσελκύσει νέους χρήστες και, ως εκ τούτου, ο δημιουργός να διευρύνει το σύνολο των πιστών του ακολούθων. Για τον λόγο αυτό αναζητήθηκε ο βαθμός κατά τον οποίο οι Έλληνες χρήστες του YouTube κοινοποιούν τα βίντεο που παρακολουθούν.

Γενικά

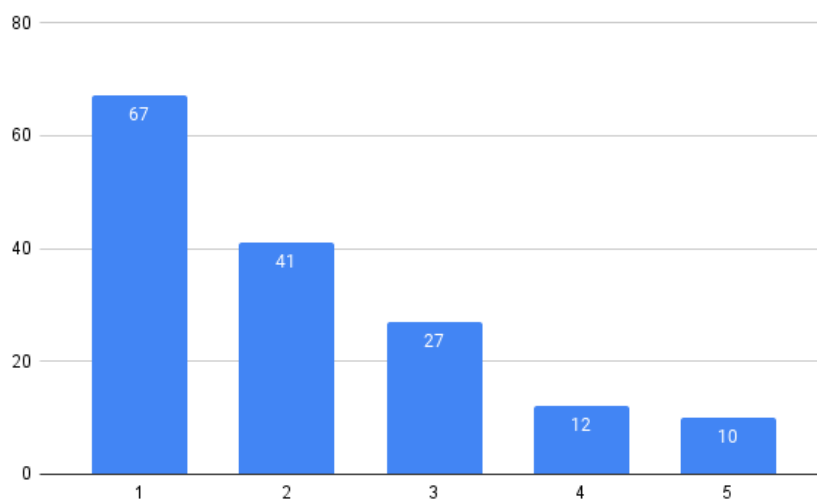
Αντίστοιχο φαινόμενο με εκείνο του Σχήματος 6.67 παρουσιάζεται και στην συγκεκριμένη περίπτωση, καθώς όπως φαίνεται στο Σχήμα 6.68, η πλειοψηφία των Ελλήνων χρηστών κοινοποιεί μερικές φορές έως και αρκετά σπάνια τα αγαπημένα της βίντεο, επιβεβαιώνοντας τη διστακτικότητα η οποία διακατέχει την ελληνική κοινότητα, αναφορικά με την αντίδραση και την προώθηση του περιεχομένου.



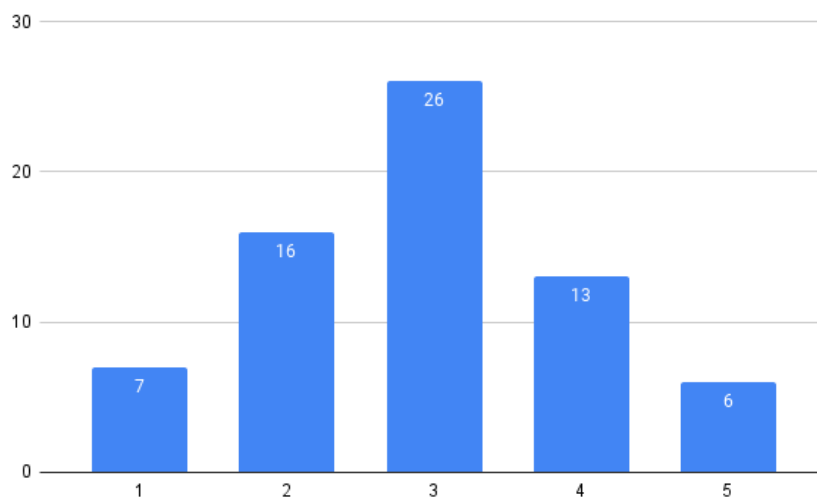
Σχήμα 6.68: Κοινοποιείτε βίντεο που σας αρέσουν σε κάποιο άλλο μέσο κοινωνικής δικτύωσης (πχ Facebook, Instagram);

Ηλικιακές Ομάδες

Με βάση την ηλικιακή διάκριση των απαντήσεων, διαπιστώνεται μέσω των Σχημάτων 6.69 και 6.70 πως οι γηραιότεροι χρήστες (35 και άνω) διατηρούν μία πιο θετική στάση απέναντι στη κοινοποίηση των αγαπημένων τους βίντεο σε άλλες πλατφόρμες δικτύωσης, σε αντίθεση με τους νεότερους, μια μεγάλη μερίδα των οποίων απάντησε πως ποτέ έως σπάνια κοινοποιεί βίντεο του YouTube.



Σχήμα 6.69: Κοινοποιείτε βίντεο που σας αρέσουν σε κάποιο άλλο μέσο κοινωνικής δικτύωσης (πχ Facebook, Instagram); (18-34)



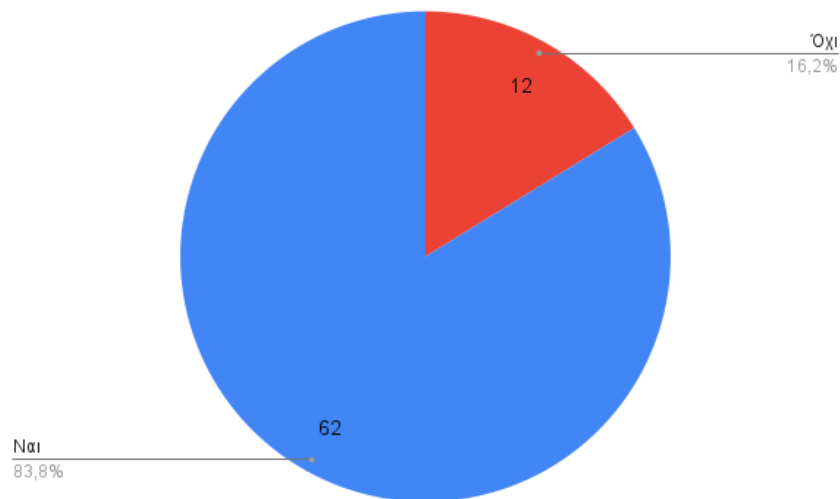
Σχήμα 6.70: Κοινοποιείτε βίντεο που σας αρέσουν σε κάποιο άλλο μέσο κοινωνικής δικτύωσης (πχ Facebook, Instagram); (35+)

6.2.10 Παρακολουθείτε βίντεο-reviews (κριτικές) στο YouTube για προϊόντα που σας ενδιαφέρουν να αγοράσετε;

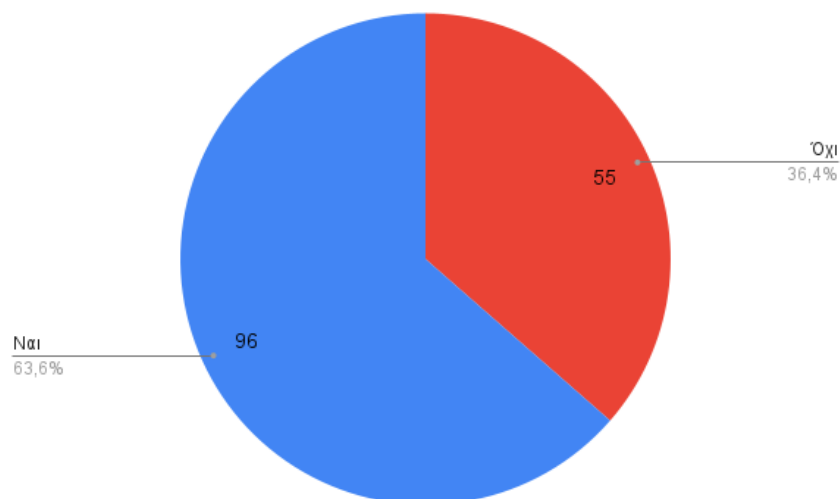
Τα βίντεο αξιολογήσεων (reviews) αποτελούν βίντεο στα πλαίσια των οποίων ο δημιουργός καλείται να αξιολογήσει ένα προϊόν, το οποίο πιθανώς έχει παραλάβει από κάποια επιχείρηση για τον σκοπό αυτόν. Η εξέταση του εμπορικού χαρακτήρα της πλατφόρμας και συγκεκριμένα της συμβολής της στη προώθηση προϊόντων είναι δυνατό να επιτευχθεί, αξιοποιώντας τον συγκεκριμένο τύπο βίντεο.

Άνδρες-Γυναίκες

Το υψηλό ποσοστό γυναικών και ανδρών, οι οποίοι δήλωσαν ότι παρακολουθούν βίντεο αξιολογήσεων για τα προϊόντα που τους ενδιαφέρουν (βλ. Σχήματα 6.71 και 6.72), μαρτυρά τον υποστηρικτικό ρόλο του YouTube και, ειδικότερα, των βίντεο αξιολόγησης στην γεφύρωση των εμπορικών σχέσεων μεταξύ του κοινού και εταιρειών. Με μία πιο προσεκτική ματιά στα παρακάτω σχήματα παρατηρείται πως οι άνδρες συνηθίζουν να παρακολουθούν περισσότερο βίντεο αξιολογήσεων (83,8%), αναλογικά με τις γυναίκες (63,6%), γεγονός το οποίο, ενδεχομένως, εκφράζει ότι η ελληνική κοινότητα του YouTube παρέχει ένα γονιμότερο έδαφος για την προώθηση προϊόντων ανδρικού ενδιαφέροντος.



Σχήμα 6.71: Παρακολουθείτε βίντεο-reviews (κριτικές) στο YouTube για προϊόντα που σας ενδιαφέρουν να αγοράσετε; (Ανδρες)



Σχήμα 6.72: Παρακολουθείτε βίντεο-reviews (κριτικές) στο YouTube για προϊόντα που σας ενδιαφέρουν να αγοράσετε; (Γυναίκες)

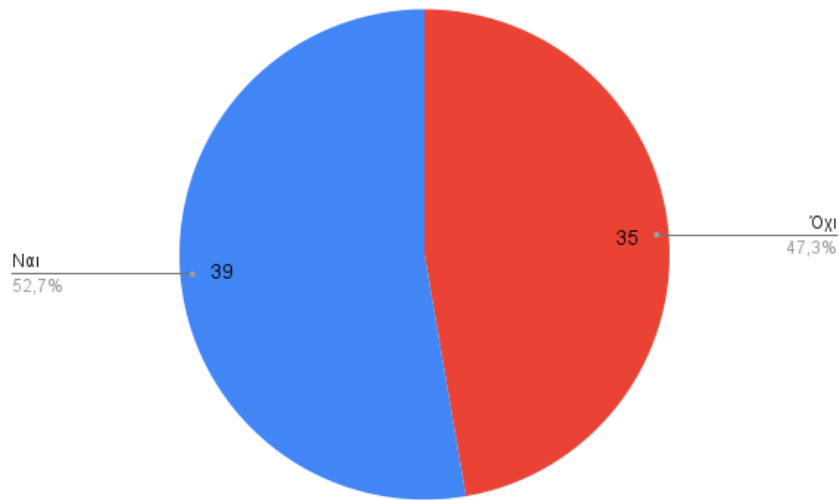
6.2.11 Έχετε αναζητήσει/αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube;

Όπως έχει προαναφερθεί αρκετές φορές, πολλές είναι οι εταιρείες οι οποίες, εκμεταλλευόμενες την υψηλή δραστηριότητα των χρηστών στα μέσα κοινωνικής δικτύωσης, αποσκοπούν στην προώθηση των προϊόντων ή των υπηρεσιών τους, είτε σε πρώτο πρόσωπο, μέσω προσωπικών δημοσιεύσεων, ή μέσω δημοφιλών δημιουργών περιεχομένου, τακτική η οποία είναι γνωστή και ως *Influencer Marketing*. Εστιάζοντας στην πλατφόρμα του YouTube, το ποσοστό επιρροής του κοινού από τέτοιου είδους δημοσιεύσεις είναι δυνατό να μελετηθεί αξιοποιώντας τη συγκεκριμένη ερώτηση. Το γεγονός ότι στην Ενότητα 5.2 είχε φανεί πως βίντεο του 2020, τα οποία

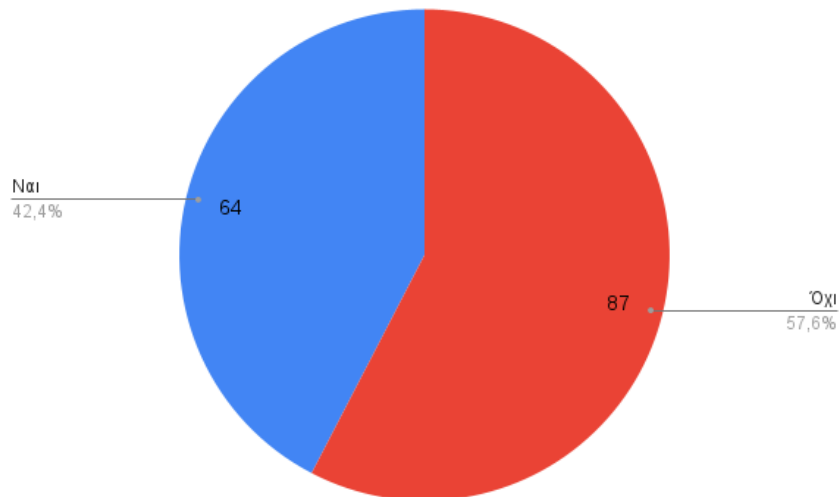
περιλάμβαναν τη παρουσία επιχειρήσεων και προϊόντων, είχαν σχετικά μεγαλύτερη απήχηση από άνδρες δημιουργούς, έδωσε το έναυσμα για την ξεχωριστή ανάλυση των απαντήσεων, οι οποίες προήλθαν από τα δύο φύλα.

Άνδρες-Γυναίκες

Παρατηρώντας τα Σχήματα 6.73 και 6.74, διαπιστώνεται πως το ανδρικό κοινό παρουσιάζει μεγαλύτερη τάση από τις γυναίκες, προς την αναζήτηση ή αγορά προϊόντων τα οποία ανακάλυψαν από κάποιο βίντεο του YouTube, φαινόμενο το οποίο, πιθανώς, υποδηλώνει την καλύτερη ανταπόκριση των επιχειρήσεων και γενικότερα της ελληνικής κοινότητας της πλατφόρμας στις απαιτήσεις του ανδρικού κοινού.



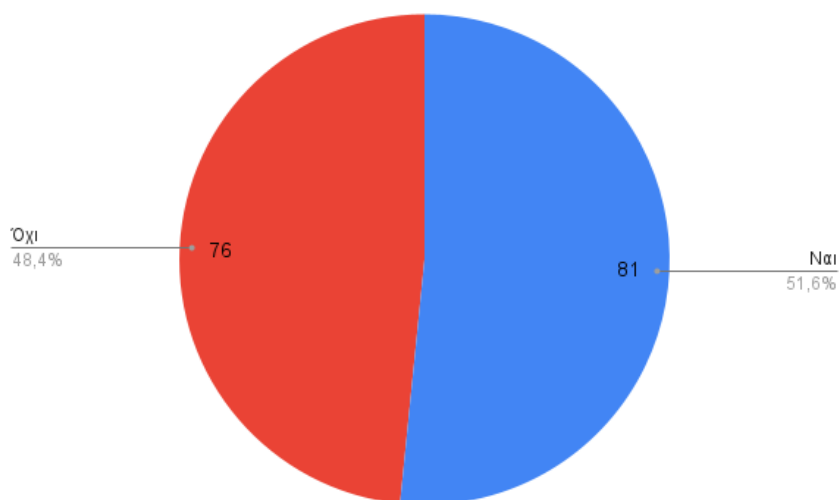
Σχήμα 6.73: Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube; (Άνδρες)



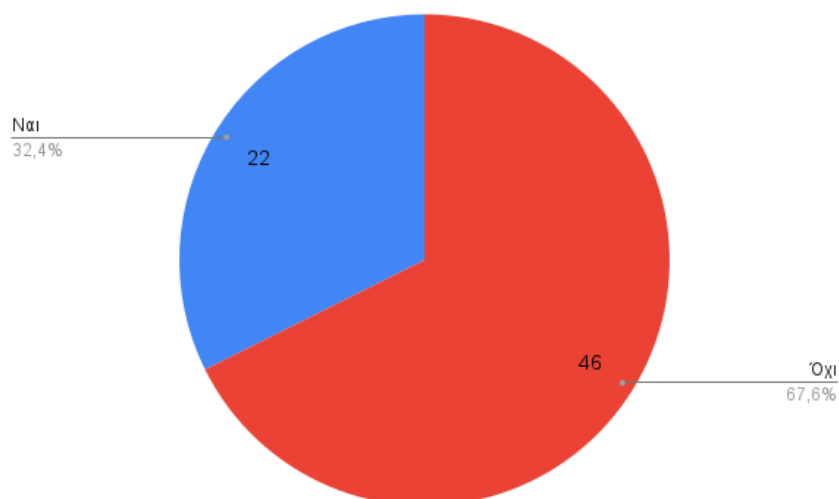
Σχήμα 6.74: Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube; (Γυναίκες)

Ηλικιακές Ομάδες

Στον άξονα των ηλικιών των ερωτηθέντων, παρατηρείται στα Σχήματα 6.75 και 6.76 ότι χρήστες 35 ετών και άνω εμφανίζουν μικρότερη τάση στην αναζήτηση ή αγορά προϊόντων τα οποία ανακάλυψαν σε κάποιο βίντεο της πλατφόρμας, σε σύγκριση με τους νεότερους και σε αντίθεση με την αντίστοιχη συμπεριφορά τους στο Instagram (βλ. Υποενότητα 6.1.7), δείχνοντας, ενδεχομένως, τον υψηλότερο βαθμό επιρροής των μικρότερων ηλικιακά χρηστών από την προώθηση προϊόντων στο YouTube. Η περιορισμένη καθημερινή χρήση της πλατφόρμας (βλ. Σχήμα 6.54) και ταύτιση των γηραιότερων χρηστών με το περιεχόμενό της (βλ. Σχήμα 6.66), αλλά και η ασυμφωνία μεταξύ των ενδιαφερόντων αυτών και των προϊόντων τα οποία τοποθετούνται σε βίντεο του YouTube, είναι δυνατό να αποτελούν αίτια του παραπάνω φαινομένου.



Σχήμα 6.75: Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube; (18-34)



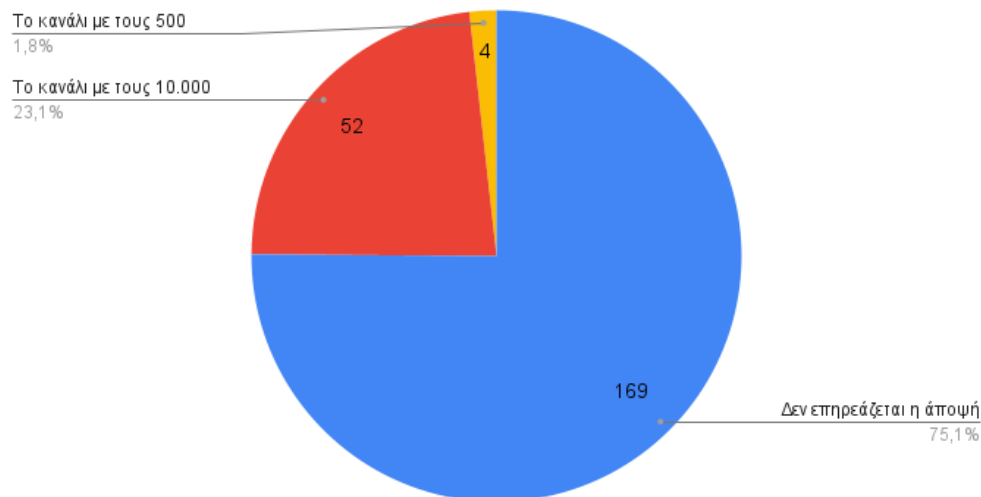
Σχήμα 6.76: Έχετε αναζητήσει ή αγοράσει ποτέ κάποιο προϊόν που είδατε για πρώτη φορά σε κάποιο βίντεο στο YouTube; (35+)

6.2.12 Έστω ότι δύο YouTube κανάλια, ένα με 500 subscribers και ένα με 10.000 subscribers διαφημίζουν το ίδιο προϊόν. Ποιο από τα δύο κανάλια θα εμπιστευόσασταν;

Στο επίπεδο του *Influencer Marketing* και του ποσοστού επιρροής του κοινού από τους δημιουργούς περιεχομένου, όσον αφορά στην προώθηση ενός προϊόντος, θεωρήθηκε ενδιαφέρουσα η μελέτη της σχέσης μεταξύ του αριθμού εγγεγραμμένων ενός καναλιού και της εμπιστοσύνης την οποία δείχνει το κοινό σε αυτόν.

Γενικά

Παρά την ουδέτερη στάση του κοινού ως προς την εμπιστευτικότητα της άποψης ενός μικρού ή μεγάλου προφίλ, αντίστοιχα, για ένα προϊόν (βλ. Σχήμα 6.77), είναι φανερό πως αν έπρεπε να επιλέξει ένα από τα δύο προφίλ, αυτό πιθανότατα θα ήταν εκείνο με τον μεγαλύτερο αριθμό εγγεγραμμένων χρηστών (23,1%), δείχνοντας έτσι το φαινομενικά μεγαλύτερο κύρος και, κατά συνέπεια, το μεγαλύτερο ποσοστό επιρροής, το οποίο προσδίδει ο μεγαλύτερος αριθμός εγγεγραμμένων σε έναν δημιουργό βίντεο.



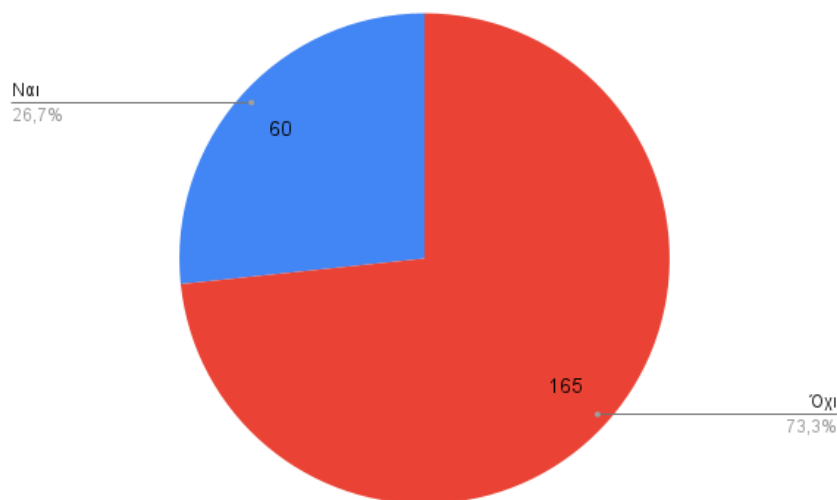
Σχήμα 6.77: Έστω ότι δύο YouTube κανάλια, ένα με 500 subscribers και ένα με 10.000 subscribers διαφημίζουν το ίδιο προϊόν. Ποιο από τα δύο κανάλια θα εμπιστευόσασταν;

6.2.13 Έχετε συμμετάσχει ποτέ σε κάποιο YouTube διαγωνισμό;

Με τον όρο *giveaway* υπενθυμίζεται ότι γίνεται αναφορά σε βίντεο διαγωνισμών, στα πλαίσια των οποίων ο διαχειριστής του καναλιού καλεί το κοινό να αντιδράσει στο συγκεκριμένο βίντεο, με σκοπό τη διεκδίκηση ενός προϊόντος. Ο πρωταγωνιστικός ρόλος των προϊόντων και η μεγάλη ανταπόκριση του κοινού σε τέτοιου είδους δημοσιεύσεις, όπως φάνηκε και στην Ενότητα 5.2, έχουν καταστήσει τα *giveaways* ως ένα ιδανικό μέσο προώθησης των επιχειρήσεων στα μέσα κοινωνικής δικτύωσης. Ο συγκεκριμένη ερώτηση τέθηκε έτσι ώστε να εξακριβωθεί η ανταπόκριση του ελληνικού κοινού σε βίντεο διαγωνισμών στο YouTube.

Γενικά

Σε αντίθεση με το Instagram, προκύπτει μέσω του Σχήματος 6.29, ότι τα βίντεο διαγωνισμών είναι λιγότερο διαδεδομένα στη πλατφόρμα του YouTube, καθώς το 73,3% των ερωτηθέντων απάντησε πως δεν έχει συμμετάσχει ποτέ σε κάποιο *giveaway*. Ωστόσο, οι μετρικές απόδοσης τέτοιου είδους βίντεο παρέμειναν σε υψηλά επίπεδα (βλ. Σχήμα 5.103), υπενθυμίζοντας πως αυτές αποτελούν ένα κατάλληλο μέσο ανάδειξης των επιχειρήσεων στο YouTube.



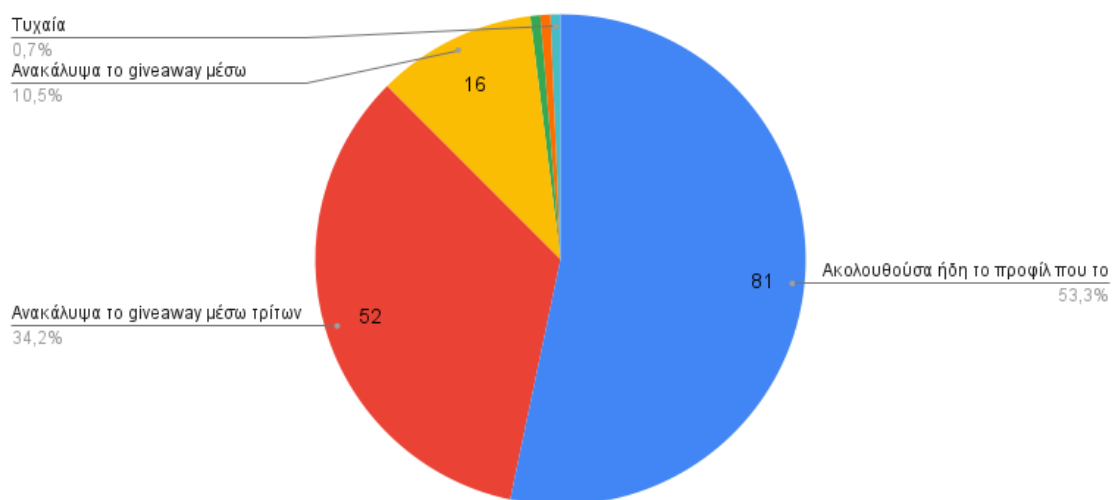
Σχήμα 6.78: Έχετε συμμετάσχει ποτέ σε κάποιο YouTube διαγωνισμό;

6.2.14 Αν ναι, πώς ανακαλύψατε αυτό το Giveaway;

Τα *giveaways* απευθύνονται σε ένα ευρύτατο κοινό, το οποίο περιλαμβάνει αρκετούς χρήστες που δεν αποτελούν εγγεγραμμένους χρήστες του διοργανωτή. Με αφορμή το γεγονός αυτό, κρίθηκε σημαντική η αναζήτηση των τρόπων προσέλκυσης νέων ατόμων στα πλαίσια των διαγωνισμών.

Γενικά

Εξαιρουμένων των χρηστών οι οποίοι ήταν ήδη εγγεγραμμένοι στο κανάλι που διοργάνωσε τον διαγωνισμό και κατέλαβαν τη πλειοψηφία των απαντήσεων (68,3%), οι δύο βασικοί διάλογοι γνωστοποίησης των *giveaways*, βάσει του Σχήμα 6.79, φαίνεται πως είναι η κοινοποίηση τους σε άλλες πλατφόρμες δικτύωσης (16,7%) και η κοινοποίηση τους μέσω άλλων χρηστών της πλατφόρμας (8,3%), ενέργειες οι οποίες είναι άμεσα συνδεδεμένες και συμβάλλουν στη προσέγγιση ατόμων εκτός του φάσματος των εγγεγραμμένων χρηστών του διοργανωτή. Τέλος, ένα μικρό υποσύνολο των ερωτηθέντων (6,7%) δήλωσε πως ανακάλυψε τον διαγωνισμό, στον οποίο συμμετείχε, τυχαία.



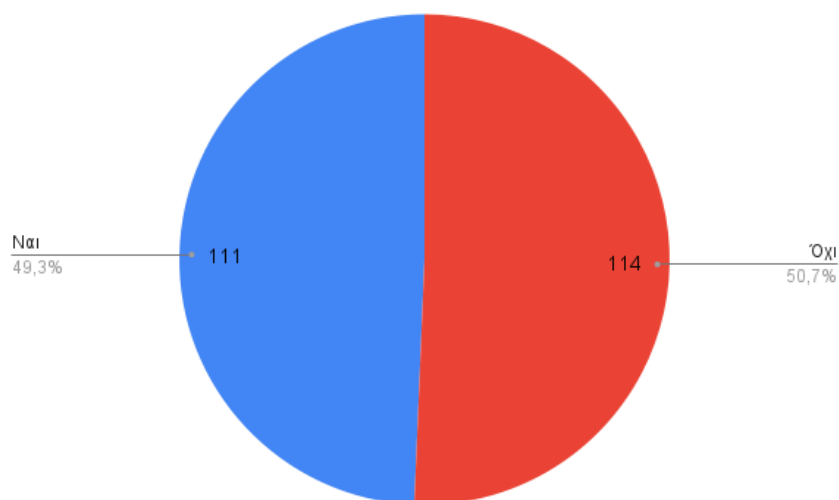
Σχήμα 6.79: Αν ναι, πώς ανακαλύψατε αυτό το Giveaway;

6.2.15 Έχετε αναζητήσει/επισκεφτεί ποτέ κάποιο προϊόν/βίντεο που εμφανίζεται σε διαφήμιση στο YouTube;

Πέραν της τακτικής του *Influencer Marketing* και της προώθησης των εταιρειών από δημοφιλείς δημιουργούς περιεχομένου, δίνεται η δυνατότητα στις επιχειρήσεις να αναπτύξουν τη δική τους ψηφιακή παρουσία στο YouTube και να παρουσιάσουν τα προϊόντα και τις υπηρεσίες τους στα πλαίσια των βίντεο τους. Αρκετές επιχειρήσεις αλλά και γενικότερα δημιουργοί βίντεο, θέλοντας να προβάλλουν το περιεχόμενό τους σε ένα ευρύ φάσμα χρηστών, επιλέγουν να προωθήσουν τα βίντεο τους μέσω διαφημίσεων. Με σκοπό τη μελέτη του ποσοστού επιρροής των διαφημίσεων στη δραστηριότητα του κοινού, τέθηκε η παρούσα ερώτηση.

Γενικά

Με γνώμονα τα δεδομένα του Σχήματος 6.80, γίνεται αντιληπτό πως η αποδοτικότητα των διαφημίσεων στη πλατφόρμα του YouTube είναι αμφιλεγόμενη, καθώς ένας στους δύο ερωτηθέντες δήλωσε πως έχει αναζητήσει ή επισκεφθεί στο παρελθόν ένα προϊόν ή βίντεο, το οποίο εμφανίστηκε σε διαφήμιση, ενώ το υπόλοιπο 50,7% αποκρίθηκε αρνητικά στην ερώτηση.



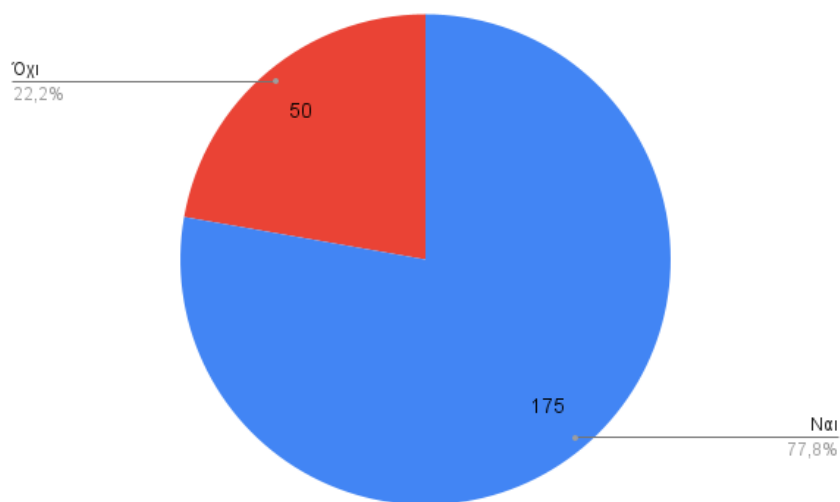
Σχήμα 6.80: Έχετε αναζητήσει ή επισκεφτεί ποτέ κάποιο προϊόν ή βίντεο που εμφανίζεται σε διαφήμιση στο YouTube;

6.2.16 Έχετε αποχωρήσει ποτέ από βίντεο λόγω αναδυόμενων διαφημίσεων;

Η παρουσίαση του διαφημιζόμενου περιεχομένου σε κοινό της πλατφόρμας, το οποίο δεν ανήκει στο σύνολο των εγγεγραμμένων χρηστών του αντίστοιχου δημιουργού, συχνά εγκυμονεί τον κίνδυνο της δυσανασχέτησης αρκετών χρηστών, λόγω της έλλειψης εξοικείωσης τους με το αντικείμενο διαφήμισης. Με αφορμή το γεγονός αυτό, το κοινό ρωτήθηκε αν έχει αποχωρήσει ποτέ από κάποιο βίντεο, λόγω αναδυόμενων διαφημίσεων.

Γενικά

Σύμφωνα με το Σχήμα 6.81, είναι εμφανής ο αρνητικός αντίκτυπος της στρατηγικής διαφημίσεων στην ελληνική κοινότητα του YouTube, καθώς το σχεδόν οχτώ στους δέκα χρήστες (77,8%) αποκρίθηκαν πως έχουν αποχωρήσει από κάποιο βίντεο, εξαιτίας των διαφημίσεων. Το φαινόμενο αυτό πιθανώς να συνδέεται, ως έναν βαθμό, με το υψηλό ποσοστό αρνητικών αντιδράσεων, το οποίο συγκέντρωσαν τα βίντεο των επιχειρήσεων κατά το έτος 2020 (βλ. Σχήμα 5.90).



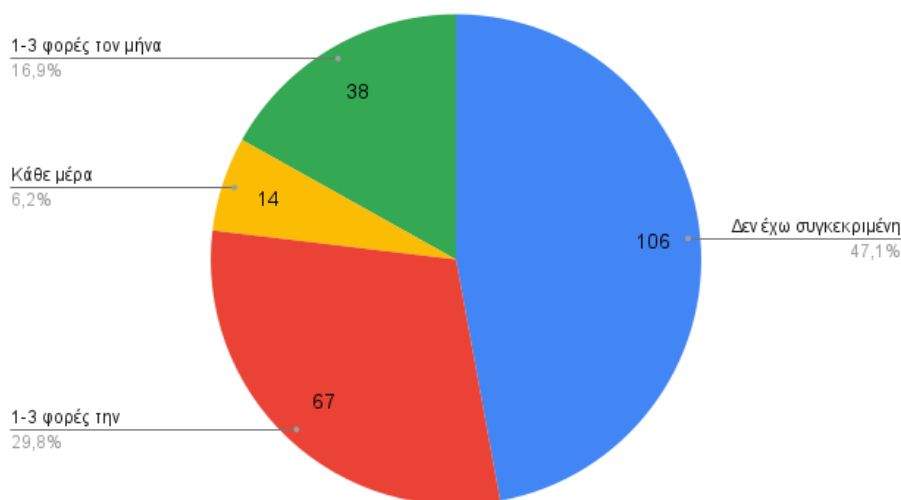
Σχήμα 6.81: Έχετε αποχωρήσει ποτέ από βίντεο λόγω αναδυόμενων διαφημίσεων;

6.2.17 Κάθε πότε προτιμάτε να αναρτά βίντεο κάποιος δημιουργός;

Επιστρέφοντας στην ανάλυση των προτιμήσεων του κοινού, όσον αφορά στη γενικότερη χρήση της πλατφόρμας, αναζητήθηκε η συχνότητα ανάρτησης των βίντεο, η οποία ανταπεξέρχεται με τον βέλτιστο τρόπο στις ανάγκες του ελληνικού κοινού.

Γενικά

Όπως παρατηρείται στο Σχήμα 6.34, ένας στους δύο Έλληνες χρήστες, κατά προσέγγιση (47,1%), δήλωσε ότι δεν έχει κάποια συγκεκριμένη προτίμηση για τη συχνότητα ανάρτησης δημοσιεύσεων, ωστόσο εστιάζοντας στο υπόλοιπο 52,9% των απαντήσεων, παρουσιάζεται η κλίση του κοινού προς τους χαμηλότερους ρυθμούς ανάρτησης των δημοσιεύσεων, με την ιδανική να φαίνεται πως είναι η 1-3 βίντεο την εβδομάδα (29,8%). Η καθημερινή κοινοποίηση περιεχομένου αποτέλεσε τη λιγότερο αρεστή επιλογή του κοινού (6,2%), συμπεριφορά η οποία ακολούθησε το ίδιο μοτίβο με εκείνη των χρηστών του Instagram απέναντι στη συχνότητα δημοσιεύσεων (βλ. Σχήμα 6.34).



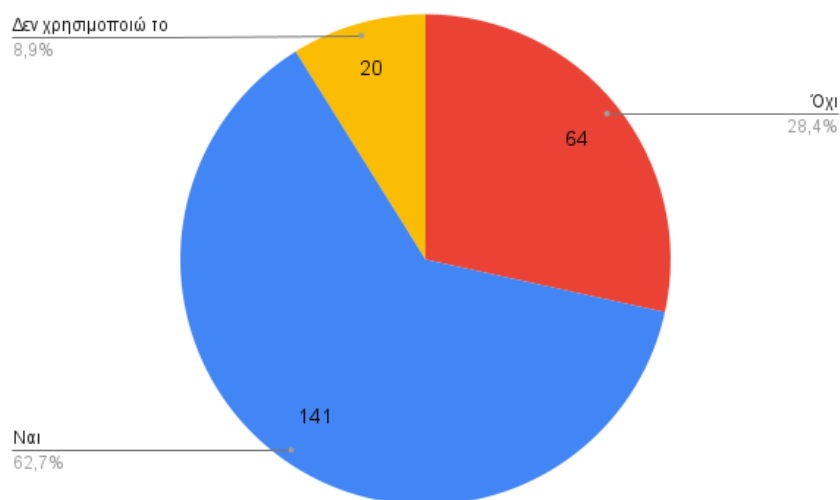
Σχήμα 6.82: Κάθε πότε προτιμάτε να αναρτά βίντεο κάποιος δημιουργός;

6.2.18 Έχετε επισκεφθεί ποτέ κάποιο προφίλ του Instagram που ανακαλύψατε μέσω του YouTube;

Αρκετοί δημιουργοί περιεχομένου, εκμεταλλευόμενοι τη μεγάλη άνθηση των μέσων κοινωνικής δικτύωσης, αναπτύσσουν τη διαδικτυακή τους παρουσία σε πολλαπλές πλατφόρμες, έτσι ώστε να προσεγγίσουν νέο κοινό. Αντίστοιχα, πολλοί είναι οι δημιουργοί βίντεο οι οποίοι επεκτείνουν τη δραστηριότητα τους στη πλατφόρμα του Instagram, αλλά και οι δημιουργοί περιεχομένου του Instagram οι οποίοι εμφανίζονται σε βίντεο του YouTube, έτσι ώστε να προσελκύσουν νέο κοινό στο προφίλ τους. Η συγκεκριμένη κινητικότητα μεταξύ των δύο πλατφορμών είναι δυνατόν να εξεταστεί, αξιοποιώντας τις απαντήσεις της προκείμενης ερώτησης.

Γενικά

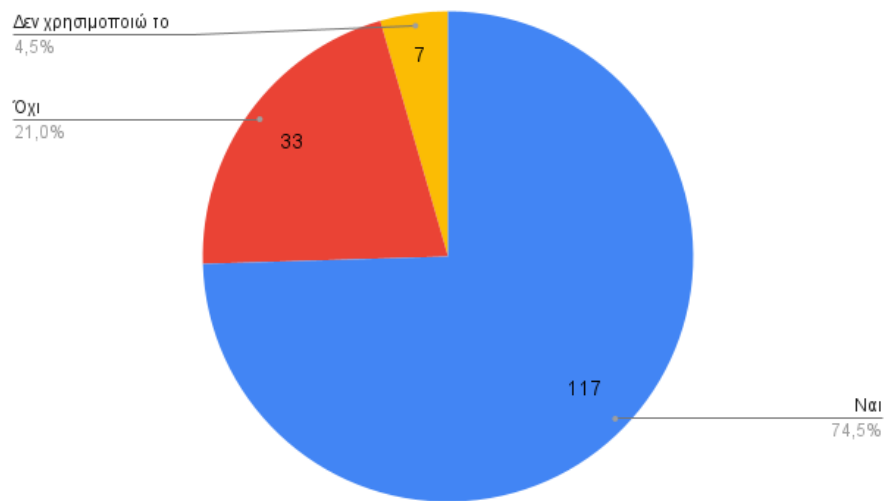
Λαμβάνοντας υπόψη μόνο το σύνολο των ερωτηθέντων οι οποίοι είναι χρήστες του Instagram (91,1%), διαπιστώνεται μέσω του Σχήματος 6.83 ότι, κατά προσέγγιση, έξι στους δέκα χρήστες του YouTube έχουν επισκεφθεί κάποιο προφίλ του Instagram που ανακάλυψαν μέσω του YouTube, γεγονός το οποίο αναδεικνύει τα περιθώρια των δημιουργών βίντεο να διευρύνουν ακόμα περισσότερο το κοινό τους στην πλατφόρμα του Instagram, ενθαρρύνοντας τους εγγεγραμμένους χρήστες των καναλιών τους να τους ακολουθήσουν, αλλά και των δημιουργών του Instagram να προσελκύσουν νέους ακολούθους, κάνοντας την εμφάνιση τους στα βίντεο του YouTube.



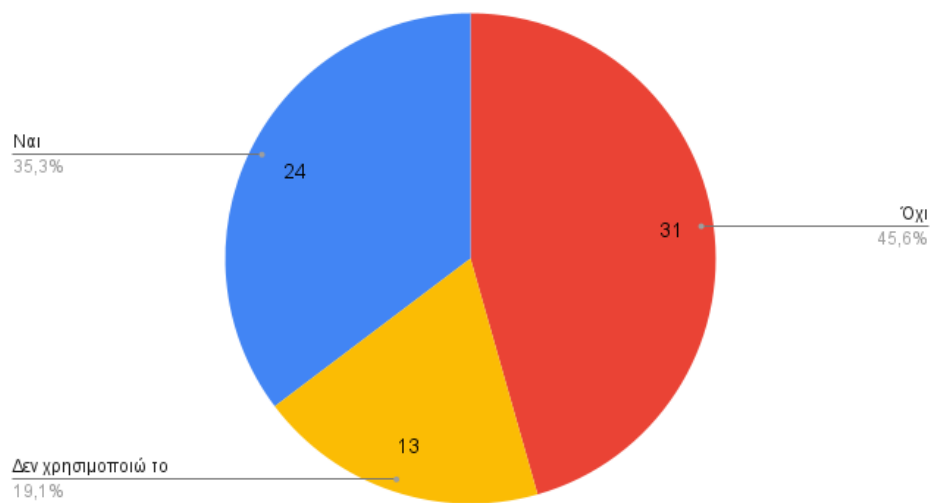
Σχήμα 6.83: Έχετε επισκεφθεί ποτέ κάποιο προφίλ του Instagram που ανακαλύψατε μέσω του YouTube;

Ηλικιακές Ομάδες

Με γνώμονα τα δεδομένα των Σχημάτων 6.84 και 6.85 και εξαιρώντας το υποσύνολο των χρηστών του YouTube οι οποίοι δε χρησιμοποιούν το Instagram, επιβεβαιώνεται η μειωμένη κινητικότητα μεταξύ των δύο πλατφορμών Instagram και YouTube από τους γηραιότερους χρήστες, καθώς είναι δυνατό να παρατηρηθεί πως εκείνοι απέχουν σε μεγαλύτερο βαθμό, σε σχέση με τους νεότερους, από την αναζήτηση δημιουργών περιεχομένου του Instagram τους οποίους ανακάλυψαν στα πλαίσια κάποιου βίντεο του YouTube. Όπως έχει προαναφερθεί και στην Ενότητα 6.1.21, το συγκεκριμένο φαινόμενο, πιθανώς, απορρέει από τη χαμηλότερη εξοικείωση των μεγαλύτερων ηλικιακά χρηστών με τα μέσα κοινωνικής δικτύωσης, αλλά και, όπως φάνηκε στις Ενότητες 6.2.7 και 6.13, από τη μικρότερη τάση αυτών να εγγράφονται στα κανάλια δημιουργών ή να ακολουθούν δημοφιλείς δημιουργούς περιεχομένου και, κατά συνέπεια, να επηρεάζονται λιγότερο από την παρουσία αυτών σε κάποιο βίντεο ή ανάρτηση.



Σχήμα 6.84: Έχετε επισκεφθεί ποτέ κάποιο προφίλ του Instagram που ανακαλύψατε μέσω του YouTube; (18-34)



Σχήμα 6.85: Έχετε επισκεφθεί ποτέ κάποιο προφίλ του Instagram που ανακαλύψατε μέσω του YouTube; (35+)

Κεφάλαιο 7

Συμπεράσματα και Μελλοντική Έρευνα

Στο σημείο αυτό, έχοντας ολοκληρώσει τη μελέτη των συγκεντρωμένων δεδομένων από τις πλατφόρμες του Instagram και του YouTube, καθώς και την εξέταση των απαντήσεων του ελληνικού κοινού στο δοθέν ερωτηματολόγιο, καθίστανται πλέον εφικτή η διατύπωση των γενικότερων συμπερασμάτων της έρευνας και η παράθεση προτάσεων, μέσω των οποίων θα μπορούσε να βελτιωθεί και να επεκταθεί το έργο της παρούσας εργασίας.

7.1 Συμπεράσματα

Είναι πλέον κοινώς αποδεκτό πως τα μέσα κοινωνικής δικτύωσης έχουν εισέλθει δυναμικά στις καθημερινές ζωές των Ελλήνων, με χρήστες από ολόκληρο το ηλικιακό και κοινωνικό φάσμα να δραστηριοποιούνται στις ψηφιακές κοινότητες των Instagram και YouTube για περισσότερο από δύο χρόνια και σε καθημερινή βάση, παράγοντας και καταναλώνοντας περιεχόμενο. Καθ' όλη τη διάρκεια της χρονιάς, οι χρήστες περιηγούνται στα μέσα κοινωνικής δικτύωσης, εκδηλώνοντας την επιρροή τους από ποικίλα δρώμενα της καθημερινότητας, όπως ασχολίες/ ενδιαφέροντα, καθημερινές οικογενειακές/προσωπικές στιγμές, εορταστικές, επετειακές και τουριστικές περιόδους ή ακόμα και κοινωνικές εξελίξεις.

Παρά τη γενικευμένη κλίση του κοινού προς τα μέσα κοινωνικής δικτύωσης, η κάθε πλατφόρμα ανταποκρίνεται καλύτερα σε συγκεκριμένες ανάγκες του κοινού. Οι διαφορές στη φύση των πλατφορμών του Instagram και του YouTube προσδίδουν στη κάθε μία από αυτές ξεχωριστά γνωρίσματα, τα οποία ευνοούν την ανάπτυξη ποικίλων θεματικών κατηγοριών και τύπων περιεχομένου και, κατά συνέπεια, προσελκύουν άτομα από διαφορετικές κατηγορίες χρηστών, είτε όσον αφορά το φύλλο, το προφίλ ή την ηλικία του εκάστοτε χρήστη.

7.1.1 Συμπεράσματα σχετικά με το Instagram

Το Instagram αποτελεί μια πλατφόρμα με χαρακτήρα αμιγούς κοινωνικού δικτύου, το οποίο βασίζεται στην ανάρτηση φωτογραφιών, βίντεο και στιγμιотύπων. Η φωτογραφία κατέχει πρωταγωνιστικό ρόλο στη συγκεκριμένη πλατφόρμα, ενώ, σε γενικότερο πλαίσιο, το αναρτημένο περιεχόμενο του Instagram διαθέτει μία σχετικά σύντομη διάρκεια ζωής, αναφορικά με τον χρόνο

κατά τον οποίο τραβούν την προσοχή της μεγαλύτερης μερίδας του κοινού. Το φαινόμενο αυτό ευνοεί τη τακτική ενασχόληση του κοινού με την πλατφόρμα και συγκεκριμένα την αρκετά συχνή παραγωγή και κατανάλωση περιεχομένου. Το παρόν μέσο κοινωνικής δικτύωσης φαίνεται πως έχει εδραιωθεί στην ελληνική κοινότητα και, ιδιαίτερα, στο σύνολο των νεότερων ηλικιακά ατόμων, με τη πλειοψηφία αυτών να χρησιμοποιούν τη πλατφόρμα για περισσότερο από δύο χρόνια. Ωστόσο, η είσοδος όλο και περισσότερων γηραιότερων χρηστών στο Instagram αποδεικνύει τον αυξανόμενο βαθμό επιρροής του στο γενικότερο σύνολο των χρηστών μέσω κοινωνικής δικτύωσης.

Η μελέτη της ψηφιακής συμπεριφοράς των δημοφιλών Ελλήνων χρηστών του Instagram πραγματοποιήθηκε μέσω της κατηγοριοποίησης αυτών σε τέσσερις βασικές κατηγορίες, εκείνες των ανδρών, των γυναικών, των επιχειρήσεων και των λογαριασμών τύπου Άλλο, κατηγορία η οποία συμπεριέλαβε λογαριασμούς που δεν κατατάσσονταν σε καμία από τις υπόλοιπες ομάδες χρηστών, και περιλαμβάνουν για παράδειγμα κωμικές/ψυχαγωγικές, ενημερωτικές σελίδες, κ.α.. Η μετρήσεις της έρευνας γύρω από την πλατφόρμα του Instagram βασίστηκαν σε προφίλ Ελλήνων χρηστών με αριθμό ακολούθων μεγαλύτερο του 1.000, τα οποία ανήλθαν στα 2.727.

Με βάση την παραπάνω διάκριση των χρηστών, παρατηρήθηκε πως στον χώρο του Instagram ιδιαίτερο ενδιαφέρον συγκεντρώνουν οι γυναικείοι λογαριασμοί, με αυτούς να κατέχουν τον μεγαλύτερο αριθμό ακολούθων, ενώ αμέσως επόμενοι στην αντίστοιχη κατάταξη βρίσκονται οι άνδρες. Το γεγονός αυτό, σε συνδυασμό με το υψηλότερο ποσοστό αλληλεπίδρασης το οποίο σημειώθηκε από τις αναρτήσεις των ανδρών και των γυναικών εντός του 2020, υποδηλώνει πως λογαριασμοί βασισμένοι στη προσωπική παρουσία του διαχειριστή τους διατηρούν μια καλύτερη επαφή με το κοινό, σε αντίθεση με τους περισσότερο απρόσωπους επιχειρησιακούς λογαριασμούς και τα προφίλ τύπου Άλλο. Από την άλλη μεριά, φάνηκε πως οι επιχειρήσεις διατηρούν μια αυστηρή στάση, όσον αφορά στο πλήθος των ατόμων τα οποία ακολουθούν, διασφαλίζοντας τον επαγγελματικό τους χαρακτήρα. Παρόμοια συμπεριφορά με εκείνη των επιχειρήσεων φαίνεται να υιοθετούν και οι γυναίκες, δρώντας περισσότερο επιλεκτικά στην ακολούθηση χρηστών σε σχέση με τους άνδρες, οι οποίοι έχουν τη τάση να ακολουθούν τους περισσότερους χρήστες.

Στον άξονα των θεματικών κατηγοριών περιεχομένου διαπιστώθηκε πως ένα σημαντικό μέρος των Ελλήνων χρηστών του Instagram ακολουθούν λογαριασμούς των οποίων το περιεχόμενο βασίζεται, κατά κύριο λόγο, στη δημοσίευση καθημερινών προσωπικών στιγμών, στην επίδειξη καλλιτεχνικών δεξιοτήτων και στην ανάδειξη της μόδας και του καλλωπισμού.

Στο κομμάτι της παραγωγικότητας των χρηστών, αποδείχθηκε πως κατά το έτος 2020 πρωταγωνιστικό ρόλο κατείχαν οι επιχειρήσεις και οι γυναίκες, οι οποίες ήταν οι κατηγορίες χρηστών με τη μεγαλύτερη παραγωγή δημοσιεύσεων και, συνεπώς, αυτές με την περισσότερο ενεργή δραστηριοποίηση στη πλατφόρμα. Στον αντίποδα, οι άνδρες διατήρησαν έναν αρκετά χαμηλό αριθμό δημιουργίας περιεχομένου, εκφράζοντας την μεγαλύτερη αδράνεια που τους διακατέχει ως χρήστες της πλατφόρμας, αναλογικά με τις υπόλοιπες ομάδες χρηστών. Παρά την αισθητή προαναφερθείσα διαφορά μεταξύ της δραστηριότητας των ανδρών και των γυναικών, οι δύο αυτές κατηγορίες χρηστών δραστηριοποιήθηκαν στο Instagram αρκετά έντονα κατά τους καλοκαιρινούς μήνες, μαρτυρώντας την υψηλή απήχηση του καλοκαιρινού περιεχομένου στη πλατφόρμα.

Οι θεματικές κατηγορίες περιεχομένου οι οποίες παρουσίασαν τη μεγαλύτερη άνθηση κατά το

έτος 2020 ήταν οι καθημερινές προσωπικές στιγμές, το ταξιδιωτικό περιεχόμενο, η μόδα και ο καλλωπισμός, υποδεικνύοντας πως το Instagram παρέχει ένα γόνιμο έδαφος για την ανάπτυξη τέτοιου είδους περιεχομένου. Ωστόσο η εκάστοτε κατηγορία χρηστών διέπρεψε στη παραγωγή ενός υψηλού αριθμού δημοσιεύσεων διαφορετικής φύσεως. Πιο συγκεκριμένα, οι γυναίκες και οι επιχειρήσεις αναδείχθηκαν στον τομέα της μόδας και του καλλωπισμού, οι άνδρες στο ταξιδιωτικό περιεχόμενο, ενώ οι λογαριασμοί τύπου Άλλο συναντήθηκαν σε ένα πιο ευρύ φάσμα κατηγοριών περιεχομένου, με το ταξιδιωτικό περιεχόμενο, τις τέχνες και τη ψυχαγωγία να βρίσκονται μεταξύ αυτών.

Λαμβάνοντας υπόψη τη τάση του γενικότερου συνόλου των δημοφιλών Ελλήνων χρηστών του Instagram, παρατηρήθηκε πως εκείνοι τρέφουν ιδιαίτερο ενδιαφέρον για τη δημοσίευση προσωπικών καθημερινών στιγμών, τις τέχνες, το ταξιδιωτικό περιεχόμενο και τη ψυχαγωγία. Παρόλα αυτά, όπως παρουσιάστηκε και παραπάνω, η κάθε κατηγορία χρηστών εμφάνισε ροπή προς συγκεκριμένες κατηγορίες περιεχομένου. Πέραν της ψυχαγωγίας, του ταξιδιωτικού περιεχομένου και των τεχνών, κατηγορίες οι οποίες αποτέλεσαν τον κοινό άξονα μεταξύ των ενδιαφερόντων των ανδρών και των γυναικών, οι άνδρες εξέφρασαν το ενδιαφέρον τους για την άθληση και την τεχνολογία, ενώ οι γυναίκες στράφηκαν προς τη μόδα, τον καλλωπισμό και την μαγειρική.

Επιπλέον, αξιοσημείωτη ήταν η ευρεία παρουσία αναρτήσεων διαγωνισμών στην ελληνική κοινότητα του Instagram. Η διοργάνωση πολυάριθμων διαγωνισμών από χρήστες διαφορετικών κατηγοριών και ποικίλων ενδιαφερόντων, σε συνδυασμό με τον μεγαλειώδη βαθμό αλληλεπίδρασης του κοινού, ανεξαρτήτως ηλικιακής ομάδας, οδήγησαν στη διαπίστωση πως οι διαγωνισμοί (*Giveaways*) του Instagram αποτέλεσαν έναν από τους τύπους δημοσιεύσεων με τη μεγαλύτερη απήχηση του κοινού για το έτος 2020.

7.1.2 Συμπεράσματα σχετικά με το YouTube

Το YouTube αποτελεί ένα μέσο κοινωνικής δικτύωσης με αρκετά διαφορετικά χαρακτηριστικά από το Instagram. Βασικό πυλώνα της συγκεκριμένης πλατφόρμας συνιστά η κοινοποίηση και η αναπαραγωγή βίντεο, ενώ σταδιακά εκείνη υιοθετεί αρκετά στοιχεία των σύγχρονων μέσων κοινωνικής δικτύωσης, όπως η δημοσίευση γραπτών αναρτήσεων και στιγμιότυπων, δίνοντας έτσι τη δυνατότητα στο κοινό και στους δημιουργούς βίντεο να αλληλεπιδρούν μέσω ποικίλων διαύλων επικοινωνίας. Δεδομένου ότι το YouTube υπήρχε αρκετά χρόνια πριν από την ίδρυση του Instagram, το ελληνικό κοινό διαθέτει έναν σημαντικό βαθμό εξοικείωσης με τον χαρακτήρα της πλατφόρμας, με χρήστες από ολόκληρο το ηλικιακό φάσμα να τη χρησιμοποιούν για περισσότερα από δύο χρόνια. Ωστόσο, παρά τη πολυετή παρουσία των γηραιότερων χρηστών στη πλατφόρμα, φαίνεται πως ο εκσυγχρονισμός του YouTube αφομοιώνεται σε μεγαλύτερο ποσοστό από τους νεότερους ηλικιακά χρήστες. Η δυναμική είσοδος της παρούσας πλατφόρμας στη καθημερινή ζωή των Ελλήνων χρηστών επιβεβαιώνεται από την τακτική παρακολούθηση βίντεο εκ μέρους του κοινού. Παρόλα αυτά, ο περισσότερος χρόνος που απαιτείται από τους χρήστες για την ενασχόληση τους με το περιεχόμενο του YouTube, σε σχέση με εκείνο του Instagram, έχει ως αποτέλεσμα τη μερική στροφή του κοινού προς την εβδομαδιαία χρήση, έναντι της καθημερινής.

Η μελέτη της ψηφιακής συμπεριφοράς των δημοφιλών Ελλήνων δημιουργών του YouTube πραγματοποιήθηκε μέσω της κατηγοριοποίησης αυτών σε τέσσερις βασικές κατηγορίες, εκείνες των ανδρών, των γυναικών, των επιχειρήσεων και των καναλιών τύπου Άλλο, κατηγορία η οποία

συμπεριέλαβε δημιουργούς που δεν κατατάσσονταν σε καμία από τις υπόλοιπες ομάδες χρηστών και περιλαμβάνει δισκογραφικές εταιρείες, ομάδες δημιουργών βίντεο, τηλεοπτικά κανάλια κ.α..

Με γνώμονα τη παραπάνω κατηγοριοποίηση των δημιουργών βίντεο, διαπιστώθηκε πως το YouTube είναι μία πλατφόρμα στην οποία δραστηριοποιούνται σε μεγάλο βαθμό άνδρες και κανάλια τύπου Άλλο, με την παρουσία των γυναικών και των επιχειρήσεων να χαρακτηρίζεται ως αισθητή, αλλά λιγότερο έντονη. Επιπροσθέτως, παρατηρήθηκε ότι μεγάλο όγκο εγγεγραμμένων χρηστών έχουν μονάχα κανάλια που ανήκουν στις κατηγορίες Άνδρες ή Άλλο, με τις γυναίκες να παραμένουν σε χαμηλότερα, αλλά αξιοπρεπή επίπεδα, δείχνοντας πως, πιθανώς, απευθύνονται σε ένα περισσότερο περιορισμένο σύνολο χρηστών. Όσον αφορά στη σχέση του ελληνικού κοινού με την εκάστοτε κατηγορία δημιουργών, έγινε αντιληπτό ότι οι επιχειρήσεις σημείωσαν τη λιγότερο καρποφόρα σχέση αλληλεπίδρασης με το κοινό τους, συγκεντρώνοντας τον μικρότερο αριθμό εγγεγραμμένων χρηστών και, παράλληλα, έναν υψηλό αριθμό αρνητικών αντιδράσεων στα εντός του 2020 αναρτημένα βίντεο τους. Σε γενικότερο επίπεδο, φάνηκε πως κανάλια διαχειριζόμενα από έναν συγκεκριμένο δημιουργό περιεχομένου ανέπτυξαν, κατά μέσο όρο, καλύτερη σχέση με το κοινό τους, σε αντίθεση με τους περισσότερο απρόσωπους επιχειρησιακούς λογαριασμούς και λογαριασμούς τύπου Άλλο. Αντίστοιχο φαινόμενο παρουσιάστηκε και στα νεοϊδρυθέντα κανάλια, τα οποία έλαβαν έναν αυξημένο αριθμό αντιδράσεων στα βίντεο τους, υποδηλώνοντας πως, ενδεχομένως, αυτά ανταπεξέρχονται καλύτερα στις σύγχρονες ανάγκες του ελληνικού κοινού της πλατφόρμας.

Αναφορικά με την δραστηριότητα των Ελλήνων δημιουργών βίντεο, οι άνδρες και τα κανάλια τύπου Άλλο αναδείχθηκαν ως τα περισσότερο ενεργά στην πλατφόρμα, παράγοντας τον μεγαλύτερο όγκο περιεχομένου για το έτος 2020. Παρόλα αυτά, το συγκεκριμένο έτος αποτέλεσε μία αρκετά γόνιμη περίοδο για το σύνολο της ελληνικής κοινότητας, καθώς κατά τη διάρκεια αυτής η γενικότερη παραγωγικότητα των δημιουργών αυξήθηκε σημαντικά. Εντούτοις, η προκείμενη αύξηση συνάντησε ένα τέλμα κατά του καλοκαιρινούς μήνες, γεγονός το οποίο, σε συνδυασμό με την άνοδο του ποσοστού αλληλεπίδρασης του κοινού κατά την ίδια περίοδο, μαρτυρά τους χαμηλούς ρυθμούς δημιουργίας περιεχομένου και, παράλληλα, την αυξημένη διάθεση του κοινού για παρακολούθηση βίντεο.

Οι θεματικές κατηγορίες περιεχομένου οι οποίες εμφάνισαν τη μεγαλύτερη άνθηση κατά το έτος 2020 ήταν το ηλεκτρονικό παιχνίδι, η μουσική, η ψυχαγωγία, η κωμωδία και το ενημερωτικό περιεχόμενο, με τα τρία τελευταία να παρουσιάζουν ιδιαίτερη αύξηση στο πλήθος των παραγόμενων βίντεο, αποδεικνύοντας πως το συγκεκριμένο έτος ευνόησε την ανάπτυξη τέτοιου είδους περιεχομένου. Αξίζει να σημειωθεί ότι η εκάστοτε κατηγορία δημιουργών διακρίθηκε στη δημιουργία βίντεο συγκεκριμένων θεματικών κατηγοριών. Αναλυτικότερα, μια μεγάλη μερίδα βίντεο ανδρών δημιουργών και καναλιών τύπου Άλλο εστίασε στην κωμωδία, το ηλεκτρονικό παιχνίδι και τη μουσική, ενώ οι γυναίκες δημιουργοί στράφηκαν, κατά ένα μεγάλο βαθμό, προς τη δημοσίευση καθημερινών προσωπικών στιγμών, το στιλιστικό περιεχόμενο, τον καλλωπισμό και τη μαγειρική. Από τη μεριά των επιχειρήσεων, αυτές παρουσίασαν ποικιλομορφία στον χαρακτήρα του περιεχομένου τους, με το δημοσιευμένο υλικό τους να αφορά σε θεματικές, όπως η τεχνολογία, η διατροφή, το ηλεκτρονικό παιχνίδι και η κωμωδία.

Στον άξονα της απήχησης του κοινού, φάνηκε πως κατά το έτος 2020 πόλους υψηλού ενδιαφέροντος αποτέλεσαν οι κατηγορίες ηλεκτρονικό παιχνίδι, κωμωδία, ψυχαγωγία, τέχνες και

εκπαιδευτικό περιεχόμενο, με τις τρεις πρώτες από αυτές να συναντώνται συχνότερα σε βίντεο καναλιών, τα οποία απευθύνονται σε ένα ευρύ σύνολο εγγεγραμμένων χρηστών και, παράλληλα, συγκεντρώνουν τον, κατά μέσο όρο, υψηλότερο αριθμό προβολών για το ίδιο έτος. Επιπλέον, ανέπτυξε μια αρκετά θερμή σχέση αλληλεπίδρασης με το ηλεκτρονικό παιχνίδι, τη μόδα/καλλωπισμό και με βίντεο σχετικά με κατοικίδια και ζώα, εκδηλώνοντας τη συμπάθεια του προς αυτές τις κατηγορίες, ενώ αντίθετη αντιμετώπιση είχαν τα βίντεο νέων και πολιτικής, τα οποία κατείχαν το ρεκόρ αρνητικών αντιδράσεων για το 2020. Το ηλεκτρονικό παιχνίδι, η ψυχαγωγία και οι ανθρώπινες σχέσεις διαπιστώθηκε πως ανταποκρίνονται στις σύγχρονες ανάγκες της ελληνικής κοινότητας του YouTube, συνιστώντας τους κύριους θεματικούς άξονες των βίντεο καναλιών που δημιουργήθηκαν εντός του έτους 2020.

Αρκετές κατηγορίες, ενώ σε μεμονωμένες περιπτώσεις προσέλκυσαν τα βλέμματα ενός ευρύτατου συνόλου χρηστών, δε κατάφεραν στη πλειοψηφία των βίντεο να συγκεντρώσουν αντίστοιχο ενδιαφέρον, μαρτυρώντας την δυνατότητα των δημιουργών βίντεο να εξελιχθούν περισσότερο στους συγκεκριμένους κλάδους. Μεταξύ των κατηγοριών αυτών εντοπίστηκαν η μουσική, οι επιστήμες/τεχνολογία και ο αθλητισμός. Παρόμοιο φαινόμενο παρατηρήθηκε και σε βίντεο σχετικά με την αυτοκίνηση, τις επιστήμες/τεχνολογία, τη μόδα/καλλωπισμό και τις ανθρώπινες σχέσεις, τα οποία, κατά μέσο όρο, συγκέντρωσαν έναν αρκετά υψηλότερο αριθμό προβολών από το πλήθος των εγγεγραμμένων χρηστών του καναλιού, φανερώνοντας τις προοπτικές διεύρυνσης του πιστού κοινού των δημιουργών σε αυτούς τους τομείς.

Εστιάζοντας στις προτιμήσεις συγκεκριμένων κατηγοριών χρηστών, έγινε αντιληπτό πως οι άνδρες και οι γυναίκες είχαν μία ιδιαίτερη ροπή προς διαφορετικές κατηγορίες περιεχομένου. Πέραν των τεχνών, της κωμωδίας και του εκπαιδευτικού περιεχομένου, τα οποία αποτέλεσαν κοινά ενδιαφέροντα των δύο φύλων, οι άνδρες εξέφρασαν το ενδιαφέρον τους για το ηλεκτρονικό παιχνίδι και τον αθλητισμό, ενώ οι γυναίκες στράφηκαν επίσης προς τη παρακολούθηση της καθημερινότητας των δημιουργών, το στιλιστικό περιεχόμενο, τον καλλωπισμό και την μαγειρική.

Γενικότερα, βίντεο τα οποία περιλαμβάνουν την ενεργή συμμετοχή του κοινού, όπως βίντεο ερωτήσεων/απαντήσεων, προκλήσεων, τύπου *Διαβάζουμε τις ιστορίες σας*, αλλά και τη παρουσίαση προϊόντων, όπως βίντεο αξιολόγησης, τύπου *Haul* ή *Unboxing*, φάνηκε πως προσελκύουν σε μεγάλο βαθμό το ενδιαφέρον του κοινού. Εξίσου δημοφιλή αναδείχθηκαν τα βίντεο παρουσί-ασης στιγμιότυπων από τη καθημερινή ζωή των δημιουργών (*Vlogs*), τα οποία άκμασαν κατά τη περίοδο των Χριστουγέννων. Τέλος, τα βίντεο διαγωνισμών (*Giveaways*) συμπεράθηκε ότι δεν είναι τόσο διαδεδομένα στο YouTube όσο στη πλατφόρμα του Instagram, εντούτοις διατηρούν σε αρκετά υψηλά επίπεδα το ποσοστό αλληλεπίδρασης μεταξύ του κοινού και των δημιουργών.

7.1.3 Γενικότερες παρατηρήσεις

Μελετώντας τα δεδομένα της έρευνας, προέκυψαν ορισμένες γενικότερες παρατηρήσεις όσον αφορά στη συμπεριφορά του ελληνικού κοινού στα μέσα κοινωνικής δικτύωσης.

Αρχικά, έγινε αντιληπτό πως τα νεότερα ηλικιακά άτομα κάνουν εντονότερη χρήση των μέσων κοινωνικής δικτύωσης, αναλογικά με τους γηραιότερους χρήστες. Το φαινόμενο αυτό είναι πιθανό να συνδέεται με την τάση των γηραιότερων, η οποία παρατηρήθηκε σε συγκεκριμένα σημεία της έρευνας, να αναζητούν περιεχόμενο στα μέσα κοινωνικής δικτύωσης με περισσότερο στοχευμένο τρόπο σε αντίθεση με τους νέους, οι οποίοι φάνηκε, ως ένα βαθμό, ότι περιηγούνται

στις πλατφόρμες κοινωνικής δικτύωσης δίχως αυστηρά καθορισμένο σκοπό.

Επιπλέον, αξιοσημείωτη είναι η ανταπόκριση του κοινού σε καθημερινά κοινωνικά, καλλιτεχνικά και πολιτικά δρώμενα, μέσω της κοινοποίησης περιεχομένου και της χρήσης λέξεων-κλειδιών σχετικών με τις τρέχουσες καταστάσεις. Μάλιστα, ιδιαίτερη συμπάθεια προς το καλοκαίρι και τα Χριστούγεννα εκδηλώθηκε από το κοινό, με την αλληλεπίδραση του να κορυφώνεται κατά τις περιόδους αυτές.

Στο επίπεδο των αναρτήσεων, διαπιστώθηκε πως η απόδοση αυτών είναι δυνατό να ευνοηθεί με ποικίλους τρόπους, όπως με τη χρήση λέξεων-κλειδιών, τη συνεργασία με άλλους χρήστες ή τη κοινοποίηση τους στα πλαίσια της ίδιας ή διαφορετικών πλατφορμών δικτύωσης. Ωστόσο, το ελληνικό κοινό φάνηκε πως διατηρεί μία συντηρητική, έως και αρνητική, στάση απέναντι στην υποστήριξη και τη προώθηση του αγαπημένου του περιεχομένου, υποδηλώνοντας την ανάγκη για γεφύρωση μιας θερμότερης σχέσης μεταξύ των δημιουργών περιεχομένου και τους κοινού τους.

Τέλος, αποτυπώθηκε η προτίμηση του κοινού προς τους ηπιότερους ρυθμούς ανάρτησης περιεχομένου, με την καθημερινή, αλλά και την αρκετά σπάνια, δημοσίευση περιεχομένου εκ μέρους των δημιουργών να εξασθενίζει την αλληλεπίδραση του κοινού με αυτό.

7.1.4 Επιχειρηματικός κλάδος και Influencer Marketing στα μέσα κοινωνικής δικτύωσης

Έναν από τους θεμελιώδεις άξονες της παρούσας έρευνας αποτέλεσε η μελέτη της δραστηριότητας των επιχειρήσεων στο χώρο των μέσων κοινωνικής δικτύωσης. Όπως αναφέρθηκε και νωρίτερα, η μαζική στροφή των ανθρώπων προς τις πλατφόρμες δικτύωσης έχει δώσει την ευκαιρία σε πολλές επιχειρήσεις να προωθήσουν τα προϊόντα τους και τις υπηρεσίες τους σε ένα εκτεταμένο κοινό, αναπτύσσοντας τη ψηφιακή τους παρουσία.

Πράγματι, έγινε φανερό μέσω της έρευνας ότι ένας μεγάλος αριθμός επιχειρήσεων δραστηριοποιείται στα μέσα κοινωνικής δικτύωσης. Αναλυτικότερα, φάνηκε πως το Instagram παρέχει ένα αρκετά έφορο έδαφος για την ανάπτυξη εμπορικών σχέσεων, καθώς έχει καταφέρει να προσελκύσει μία πληθώρα επιχειρήσεων, οι οποίες, μάλιστα, ακολουθούν μία αρκετά έντονη δραστηριότητα. Από την άλλη μεριά, στο YouTube η παρουσία των επιχειρήσεων είναι αισθητή, αλλά αρκετά πιο περιορισμένη, σε σχέση με το Instagram. Παρόλα αυτά, ανεξαρτήτου πλατφόρμας, οι επιχειρήσεις εμφάνισαν ορισμένα κοινά χαρακτηριστικά, όσον αφορά στην ψηφιακή συμπεριφορά τους. Εμφανής ήταν ο πιο αυστηρός και επαγγελματικός χαρακτήρας των εταιρειών στα μέσα κοινωνικής δικτύωσης, με αυτές να τηρούν ένα συγκεκριμένο μοτίβο δραστηριότητας καθ' όλη τη διάρκεια της εβδομάδας και να διαχειρίζονται με αρκετή φειδώ την ακολούθηση άλλων χρηστών. Ιδιαίτερα ενδιαφέρουσα θεωρήθηκε, επίσης, η τακτική των επιχειρήσεων να εντείνουν τη δραστηριότητα τους σε περιόδους διακοπών, εορτών και επετείων, όπως το Πάσχα, το καλοκαίρι, τα Χριστούγεννα ή η εορτή του Αγίου Βαλεντίνου, φανερώνοντας πως εκείνες αποτελούν περιόδους αυξημένης εμπορικής σημασίας. Εντούτοις, η δραστηριοποίηση αυτή των επιχειρήσεων δεν έχει καταφέρει να εγκαθιδρύσει, μέχρι στιγμής, ένα εδραιωμένο κοινό, καθώς διαπιστώθηκε σε αρκετά σημεία της έρευνας πως οι χρήστες διατηρούν ένα σημαντικό χαμηλό ποσοστό αλληλεπίδρασης με δημοσιευμένο περιεχόμενο των εταιριών.

Με σκοπό την αποδοτικότερη προώθηση του περιεχομένου τους, οι επιχειρήσεις στρέφονται είτε στην αυτοπρόσωπη διαφήμιση τους ή στη διαφήμιση τους μέσω δημοφιλών δημιουργών περιεχομένου, η οποία είναι γνωστή και ως *Influencer Marketing*.

Στο επίπεδο της αυτοπρόσωπης διαφήμισης προέκυψε το συμπέρασμα ότι τέτοιου είδους τακτικές δεν είναι σημαντικά προσοδοφόρες. Η παρουσίαση του διαφημιζόμενου περιεχομένου σε ένα σύνολο χρηστών οι οποίοι δεν ανήκουν στο σύνολο των πιστών ακολούθων των αντίστοιχων επιχειρήσεων φάνηκε πως οδηγεί στην αρκετά χαμηλή αλληλεπίδραση του κοινού με το περιεχόμενο, αλλά και σε ορισμένες περιπτώσεις στο αίσθημα της ενόχλησης των χρηστών. Η ελληνική κοινότητα του Instagram παρουσιάστηκε περισσότερο ανεκτική στις διαφημίσεις, σε σχέση με εκείνη του YouTube, ενώ σε γενικότερο επίπεδο παρατηρήθηκε πως υψηλότερο ποσοστό επιρροής από τέτοιου είδους τακτικές ασκείται σε μεγαλύτερους ηλικιακά χρήστες.

Στον αντίποδα, το *Influencer Marketing* αποδείχθηκε ένας περισσότερο αποδοτικός τρόπος προώθησης των προϊόντων των επιχειρήσεων. Τα υψηλότερα ποσοστά αλληλεπίδρασης του κοινού με αναρτήσεις δημιουργών, οι οποίες περιλαμβάνουν τοποθέτηση προϊόντος, πιθανότατα αποτέλεσαν την αφορμή με την οποία πολυάριθμες επιχειρήσεις στράφηκαν σε δημοφιλείς χρήστες, κατά το 2020, για τη προώθηση τους, οργανώνοντας αναρτήσεις ποικίλων τύπων, όπως διαγωνισμούς, χορηγίες, unboxings ή αξιολογήσεις. Πέραν της συνολικά θετικής ανταπόκρισης του κοινού, η τακτική του *Influencer Marketing* φάνηκε πως ασκεί μεγαλύτερη επιρροή στις μερίδες των νεότερων ηλικιακά χρηστών των μέσων κοινωνικής δικτύωσης.

Όπως έχει γίνει κατανοητό μέχρι στιγμής, η διαφορά στη φύση των Instagram και YouTube έχουν ευνοήσει την ανάπτυξη διαφορετικών θεματικών κατηγοριών και τύπων περιεχομένου στη κάθε μία πλατφόρμα από αυτές. Το φαινόμενο αυτό διαπιστώθηκε πως έχει διαμορφώσει ως ένα βαθμό τις τακτικές των επιχειρήσεων, οι οποίες θέτουν διαφορετικούς στόχους στα δύο αυτά μέσα κοινωνικής δικτύωσης. Ειδικότερα, στον ελληνικό χώρο του Instagram οι επιχειρήσεις εκδήλωσαν τη προτίμησή τους προς τις γυναίκες δημιουργούς περιεχομένου και ειδικά προς εκείνες οι οποίες δημοσιεύουν περιεχόμενο που ανταποκρίνεται σε προσωπικές τους καθημερινές στιγμές, ενώ οι άνδρες παρέμειναν δεύτεροι στις προτιμήσεις των εταιριών. Επιπλέον, το ιδιαίτερα υψηλό ενδιαφέρον των επιχειρήσεων εντοπίστηκε στον κλάδο της μόδας και του καλλωπισμού, ο οποίος αποτέλεσε τη θεματική κατηγορία με τη μεγαλύτερη δραστηριότητα των επιχειρησιακών λογαριασμών. Από την πλευρά του YouTube, προέκυψε ότι οι άνδρες δημιουργοί βίντεο βρέθηκαν στο επίκεντρο της προσοχής των επιχειρήσεων, με τις γυναίκες, παρόλα αυτά, να κάνουν αισθητή τη συμβολή τους στη δημιουργία βίντεο, τα οποία περιλαμβάνουν τοποθέτηση προϊόντος. Στο επίπεδο των θεματικών κατηγοριών περιεχομένου, οι επιχειρήσεις του YouTube φάνηκε να εστιάζουν στο ηλεκτρονικό παιχνίδι, τις επιστήμες/τεχνολογία, τη μόδα/καλλωπισμό, αλλά και τα ταξίδια. Τέλος, αξίζει να επισημανθεί ότι, όπως παρατηρήθηκε, η δραστηριοποίηση των επιχειρήσεων στο Instagram ασκεί επιρροή σε ένα ευρύ ηλικιακό φάσμα χρηστών, σε αντίθεση με το YouTube, στο οποίο οι επιχειρήσεις ανταποκρίνονται καλύτερα στις ανάγκες του νεότερου ηλικιακά κοινού.

7.1.5 COVID-19 και μέσα κοινωνικής δικτύωσης

Το 2020 ήταν ένα έτος το οποίο έφερε αρκετές αλλαγές στην καθημερινότητα των ανθρώπων, καθώς αποτέλεσε την περίοδο έξαρσης του ιού *COVID-19*. Τα μέτρα αντιμετώπισης του ιού και οι συνθήκες εγκλεισμού, οι οποίες επικράτησαν κατά το συγκεκριμένο έτος, είχαν άμεσο

αντίκτυπο στη ψυχολογία των Ελλήνων, η οποία μεταφράστηκε και σε μεταβολή της ψηφιακής τους συμπεριφοράς στα μέσα κοινωνικής δικτύωσης.

Η έντονη επιρροή του κοινού από την πανδημία έγινε φανερή μέσω της στροφής των χρηστών προς τα μέσα κοινωνικής δικτύωσης και, συγκεκριμένα, μέσω της χρήσης πληθώρας λέξεων-κλειδιών εκ μέρους μια μεγάλης μερίδας της ελληνικής κοινότητας, οι οποίες αναφέρονταν στον κορονοϊό και αρκετές φορές, μάλιστα, διέπονταν από ένα αίσθημα αναπόλησης και νοσταλγίας. Ωστόσο, διαπιστώθηκε πως μεγαλύτερο αντίκτυπο στο κοινό είχε το πρώτο κύμα της πανδημίας (Μάρτιος 2020- Μάιος 2020), με τη δεύτερη φάση (Οκτώβριος 2020-Δεκέμβριος 2020) να συνοδεύεται από την εξασθενημένη, αλλά αισθητή ευαισθητοποίηση των χρηστών.

Γενικότερα, οι επικρατούσες συνθήκες κατά το 2020 είχαν ως αποτέλεσμα την αύξηση της δραστηριότητας του κοινού στα μέσα κοινωνικής δικτύωσης, φαινόμενο από το οποίο επωφελήθηκε ιδιαίτερα το YouTube, όπου ο αριθμός των βίντεο σημείωσε σημαντική αύξηση έως τα τέλη του ίδιου έτους. Η ενίσχυση αυτή της κινητικότητας των χρηστών στα μέσα κοινωνικής δικτύωσης, σε συνδυασμό με την αναστολή λειτουργίας των επιχειρήσεων, φαίνεται πως έδωσαν την ευκαιρία στις επιχειρήσεις να ενισχύσουν τη ψηφιακή τους παρουσία στις πλατφόρμες δικτύωσης, ενώ αντίστοιχη τακτική παρατηρήθηκε πως ακολούθησαν και λογαριασμοί τύπου Άλλο, κάνοντας αρκετά έντονη χρήση λέξεων-κλειδιών σχετικών με την καραντίνα. Τέλος, αρκετές ήταν, επίσης, οι θεματικές κατηγορίες περιεχομένου οι οποίες ευνοήθηκαν κατά την περίοδο της πανδημίας, όπως το ηλεκτρονικό παιχνίδι, η ψυχαγωγία, η κωμωδία και η ενημέρωση/πολιτική, τομείς των οποίων η παραγωγικότητα αυξήθηκε σε μεγάλο βαθμό.

7.2 Μελλοντική Έρευνα

Το έργο της παρούσας έρευνας θα μπορούσε να ενισχυθεί, είτε βελτιώνοντας τον μηχανισμό συγκομιδής δεδομένων ή επεκτείνοντας το σύνολο των υπό διερεύνηση πληροφοριών.

Αναλυτικότερα, ο μηχανισμός συγκομιδής δεδομένων, ο οποίος σχεδιάστηκε στα πλαίσια της προκείμενης πτυχιακής εργασίας, δίνετε να τροποποιηθεί έτσι ώστε να γίνει ταχύτερος και αποδοτικότερος, όσον αφορά στην αποφυγή αποκλεισμού της πρόσβασης του στις ποικίλες διαδικτυακές πηγές δεδομένων, οι οποίες εφαρμόζουν τεχνικές αποτροπής του *Web Scraping*. Δεδομένων των συχνών τροποποιήσεων του πηγαίου κώδικα των μέσων κοινωνικής δικτύωσης, από τα οποία συλλέχθηκαν τα δεδομένα της μελέτης, καθίσταται αναγκαία η συντήρηση και η τακτική ενημέρωση του κώδικα του μηχανισμού συγκομιδής, με στόχο τη διαφύλαξη της λειτουργικότητας του.

Η λειτουργικότητα του μηχανισμού είναι δυνατό να επεκταθεί, με σκοπό τη συλλογή δεδομένων, τα οποία δε συλλέχθηκαν στη παρούσα μελέτη, και, συνεπώς, τον εμπλουτισμό των διαθέσιμων στοιχείων. Επιπλέον, η διεύρυνση του συνόλου των χρηστών, με βάση τους οποίους ολοκληρώθηκε η έρευνα, καθώς και η επανεκτέλεση της ίδιας ερευνητικής διαδικασίας για μεταγενέστερα έτη, με σκοπό τη σύγκριση των συμπερασμάτων γύρω από τις ελληνικές κοινότητες του Instagram και του YouTube, θα μπορούσε να ενισχύσει, αλλά και να επαυξήσει τα αποτελέσματα της παρούσας πτυχιακής εργασίας.

Βιβλιογραφία

- [1] *Add tags to videos*. URL: <https://support.google.com/youtube/answer/146402?hl=en>. (accessed: 26.08.2021).
- [2] *An Overview of Web Scraping Techniques*. URL: <https://www.radwarebotmanager.com/what-are-the-different-scraping-techniques/>. (accessed: 19.04.2021).
- [3] *Basic Steps in the Research Process*. URL: <https://www.nhcc.edu/student-resources/library/doinglibraryresearch/basic-steps-in-the-research-process>. (accessed: 05.05.2021).
- [4] Bhagyeshwari Chauhan. *How to Bypass Anti-Scraping Tools on Websites*. URL: <https://www.blog.datahut.co/post/web-scraping-how-to-bypass-anti-scraping-tools-on-websites>. (accessed: 14.12.2020).
- [5] *Cloud vs local Server – Where should you store your data?* URL: <https://www.labfolder.com/cloud-vs-local-server/>. (accessed: 04.05.2021).
- [6] *Cloud vs Local Servers: Weighing up the Pros and Cons*. URL: <https://evolvit.co.uk/it-support/server-support/cloud-vs-local-servers-weighing-up-the-pros-and-cons/>. (accessed: 04.05.2021).
- [7] *CSS Selectors*. URL: https://www.w3schools.com/css/css_selectors.asp. (accessed: 21.04.2021).
- [8] Scrapy developers. *Scrapy 2.5 documentation*. URL: <https://docs.scrapy.org/en/latest/>. (accessed: 19.10.2020).
- [9] Vojtech Draxl. «Web Scraping Data Extraction from websites». BACHELOR PAPER. University of Applied Sciences Technikum Wien, Φεβ. 2018. URL: https://www.academia.edu/35901535/BACHELOR_PAPER_Web_Scraping_Data_Extraction_from_websites..
- [10] *Evolution of Web Crawling as a Market Segment*. URL: <https://www.promptcloud.com/blog/evolution-of-web-crawling-how-crawling-emerged-as-mainstream-discipline/>. (accessed: 16.04.2021).
- [11] Kristijan Glas. *Instagram views: What Counts as a view on Instagram Videos*. URL: <https://thefutureofthings.com/13283-instagram-views-what-counts-as-a-view-on-instagram-videos/>. (accessed: 14.07.2020).
- [12] Daniel Glez-Peña, Anália Lourenço, Hugo López-Fernández, Miguel Reboiro-Jato και Florentino Fdez-Riverola. «Web scraping technologies in an API world». Στο: *BRIEFINGS IN BIOINFORMATICS* 15.5 (2013), σσ. 788–797. DOI: <http://dx.doi.org/10.1093/bib/bbt026>.

- [13] Leitao Guo. *How to Efficiently Choose the Right Database for Your Applications*. URL: <https://pingcap.com/blog/how-to-efficiently-choose-the-right-database-for-your-applications>. (accessed: 03.05.2021).
- [14] Sandy Harden. *What is a Survey and What Types of Surveys are Qualitative vs. Quantitative?* URL: <https://blog.surveyplanet.com/what-is-a-survey-and-what-types-of-surveys-are-qualitative-vs-quantitative/>. (accessed: 05.05.2021).
- [15] *How to scrape websites without getting blocked*. URL: <https://www.scrapehero.com/how-to-prevent-getting-blacklisted-while-scraping/?fbclid=IwAR3lDgMNU2ci9OwuwNXrPw2YdgzDaFxlIElK7upxnEVzHl1DhVuqYnBCBX8#unique-identifier2>. (accessed: 14.12.2020).
- [16] *Influencer Marketing: Η νέα τάση στο marketing*. URL: <https://developgreece.com/influencer-marketing/>. (accessed: 15.04.2021).
- [17] Hafsa Jabeen. *Making Web Crawlers Using Scrapy for Python*. URL: <https://www.datacamp.com/community/tutorials/making-web-crawlers-scrapy-python#overview>. (accessed: 11.05.2021).
- [18] *JavaScript DOM (Document Object Model) – Guide for Novice Learners*. URL: <https://data-flair.training/blogs/javascript-dom/>. (accessed: 20.04.2021).
- [19] Md. Abu Kausar, V. S. Dhaka και Sanjeev Kumar Singh. «Web Crawler: A Review». Στο: *International Journal of Computer Applications (0975 – 8887)* 63.2 (2013), σσ. 31–36. DOI: <http://dx.doi.org/10.5120/10440-5125>.
- [20] Ian Kerins. *The Easy Way to Scrape Instagram Using Python Scrapy & GraphQL*. URL: <https://dev.to/iankerins/the-easy-way-to-build-an-instagram-spider-using-python-scrapy-graphql-4gko>. (accessed: 19.10.2020).
- [21] Christa Love. *5 Reasons to Use Google Forms with Your Students*. URL: <https://www.technokids.com/blog/apps/reasons-to-use-google-forms-with-your-students/>. (accessed: 12.05.2021).
- [22] Sandra Melo. *Cloud storage vs local storage: What is the right for your business?* URL: <https://mydatascope.com/blog/en/cloud-storage-vs-local-storage-what-is-the-right-for-your-business/>. (accessed: 04.05.2021).
- [23] Philipp Meschenmoser, Norman Meuschke, Bela Gipp και Manuel Hotz. «Scraping Scientific Web Repositories: Challenges and Solutions for Automated Content Extraction». Στο: *D-Lib Magazine* 22.9/10 (2016). DOI: <http://dx.doi.org/10.1045/september2016-meschenmoser>.
- [24] MongoDB. *MongoDB Atlas*. URL: <https://docs.atlas.mongodb.com>. (accessed: 29.10.2020).
- [25] MongoDB. *Welcome to the MongoDB Documentation*. URL: <https://docs.mongodb.com/>. (accessed: 29.10.2020).
- [26] *MongoDB Charts*. URL: <https://docs.mongodb.com/charts/saas/>. (accessed: 13.05.2021).
- [27] *MongoDB Compass*. URL: <https://docs.mongodb.com/compass/current/>. (accessed: 13.05.2021).
- [28] *MongoDB Tutorial*. URL: <https://www.tutorialspoint.com/mongodb/>. (accessed: 13.05.2021).

- [29] Daniel Ni. *The Ultimate Guide To Building Scalable Web Scrapers With Scrapy*. URL: <https://www.smashingmagazine.com/2019/07/ultimate-guide-scalable-web-scrapers-scrapy/>. (accessed: 11.05.2021).
- [30] *On-Premise vs Cloud Database Hosting: Which is Right for You?* URL: <https://www.insightsforprofessionals.com/it/storage/onpremise-or-cloud-database-hosting>. (accessed: 04.05.2021).
- [31] *Optimize Your YouTube Videos*. URL: <https://backlinko.com/hub/youtube/channel-keywords>. (accessed: 25.08.2021).
- [32] Martin Perez. *Web Scraping vs API: What's the Difference?* URL: <https://www.parsehub.com/blog/web-scraping-vs-api/>. (accessed: 21.04.2021).
- [33] Martin Perez. *Web Scraping vs Web Crawling: What's the Difference?* URL: <https://www.parsehub.com/blog/web-scraping-vs-web-crawling/>. (accessed: 11.10.2020).
- [34] *Research Process: 8 Steps in Research Process*. URL: <https://www.iedunote.com/research-process>. (accessed: 05.05.2021).
- [35] Ryan Robinson. *The 7 Top Social Media Sites You Need to Care About in 2020*. URL: <https://www.adobe.com/express/learn/blog/top-social-media-sites>. (accessed: 16.05.2021).
- [36] Alex Ronquillo. *Python's Requests Library (Guide)*. URL: <https://realpython.com/python-requests/>. (accessed: 21.04.2021).
- [37] Deirdre Schwiesow. «7 Research Challenges (And how to overcome them)». Στο: *Walden Magazine* 5.1 (2010), σσ. 24–32.
- [38] ScrapingHub. *What is a Crawl Frontier?* URL: <https://frontera.readthedocs.io/en/v0.2.0/topics/what-is-a-crawl-frontier.html>. (accessed: 18.04.2021).
- [39] *Scrapy Tutorial*. URL: <https://www.tutorialspoint.com/scrapy/index.htm>. (accessed: 11.05.2021).
- [40] Sapir Segal. *How to Measure Video Marketing: The Metrics that Matter*. URL: <https://www.oktopost.com/blog/measure-video-marketing/>. (accessed: 20.12.2020).
- [41] Katie Sehl. *All the Different Ways to Calculate Engagement Rate (Free Calculator)*. URL: <https://blog.hootsuite.com/calculate-engagement-rate/>. (accessed: 20.12.2020).
- [42] Kartik Singh. *Web Scraping with a Headless Browser: A Puppeteer Tutorial*. URL: <https://www.blog.datahut.co/post/web-scraping-headless-browser-puppeteer>. (accessed: 26.04.2021).
- [43] Mark Smallcombe. *SQL vs NoSQL: 5 Critical Differences*. URL: <https://www.xplenty.com/blog/the-sql-vs-nosql-difference/>. (accessed: 03.05.2021).
- [44] Velotio Technologies. *Web Scraping: Introduction, Best Practices & Caveats*. URL: <https://medium.com/velotio-perspectives/web-scraping-introduction-best-practices-caveats-9cbf4acc8d0f>. (accessed: 21.04.2021).
- [45] *The Anti-Scraping Solution For Your Website, App & APIs*. URL: <https://www.radwarebotmanager.com/anti-scraping/>. (accessed: 23.04.2021).
- [46] *The difference between quantitative vs. qualitative research*. URL: <https://www.surveymonkey.com/mp/quantitative-vs-qualitative-research/>. (accessed: 05.05.2021).

- [47] *The research process*. URL: <https://www.futurelearn.com/info/courses/research-process/0/steps/71889>. (accessed: 05.05.2021).
- [48] *The Selenium project and tools*. URL: https://www.selenium.dev/documentation/en/introduction/the_selenium_project_and_tools/. (accessed: 21.04.2021).
- [49] Gabriele Tomassetti. *Parsing HTML: A Guide to Select the Right Library*. URL: <https://tomassetti.me/parsing-html/>. (accessed: 22.04.2021).
- [50] *View web pages cached in Google Search Results*. URL: <https://support.google.com/websearch/answer/1687222?hl=en>. (accessed: 25.04.2021).
- [51] Jillian Warren. *This is How the Instagram Algorithm Works in 2021*. URL: <https://later.com/blog/how-instagram-algorithm-works/>. (accessed: 16.12.2020).
- [52] *Web Data Crawling vs Web Data Scraping Python*. URL: <https://www.promptcloud.com/blog/data-scraping-vs-data-crawling/>. (accessed: 13.10.2021).
- [53] *What Is a Web Crawler? | How Web Spiders Work*. URL: <https://www.cloudflare.com/learning/bots/what-is-a-web-crawler/>. (accessed: 09.10.2020).
- [54] Patrick Whatman. *Social Mythbusters: Does Using Instagram Hashtags Mean More Engagement?* URL: <https://mention.com/en/blog/instagram-hashtags-engagement/>. (accessed: 31.07.2021).
- [55] *YouTube Data API Overview*. URL: <https://developers.google.com/youtube/v3/getting-started>. (accessed: 11.05.2021).