

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**

**Σχολή Χρηματοοικονομικής και Στατιστικής**



**Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ**

**ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΜΕΛΕΤΗ ΠΑΡΑΓΟΝΤΩΝ ΠΟΥ ΕΠΙΔΡΟΥΝ**

**ΣΤΗΝ ΕΠΙΛΟΓΗ ΑΘΛΗΤΙΚΗΣ**

**ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ ΜΕΣΩ**

**ΔΙΑΔΙΚΤΥΑΚΗΣ ΠΛΑΤΦΟΡΜΑΣ**

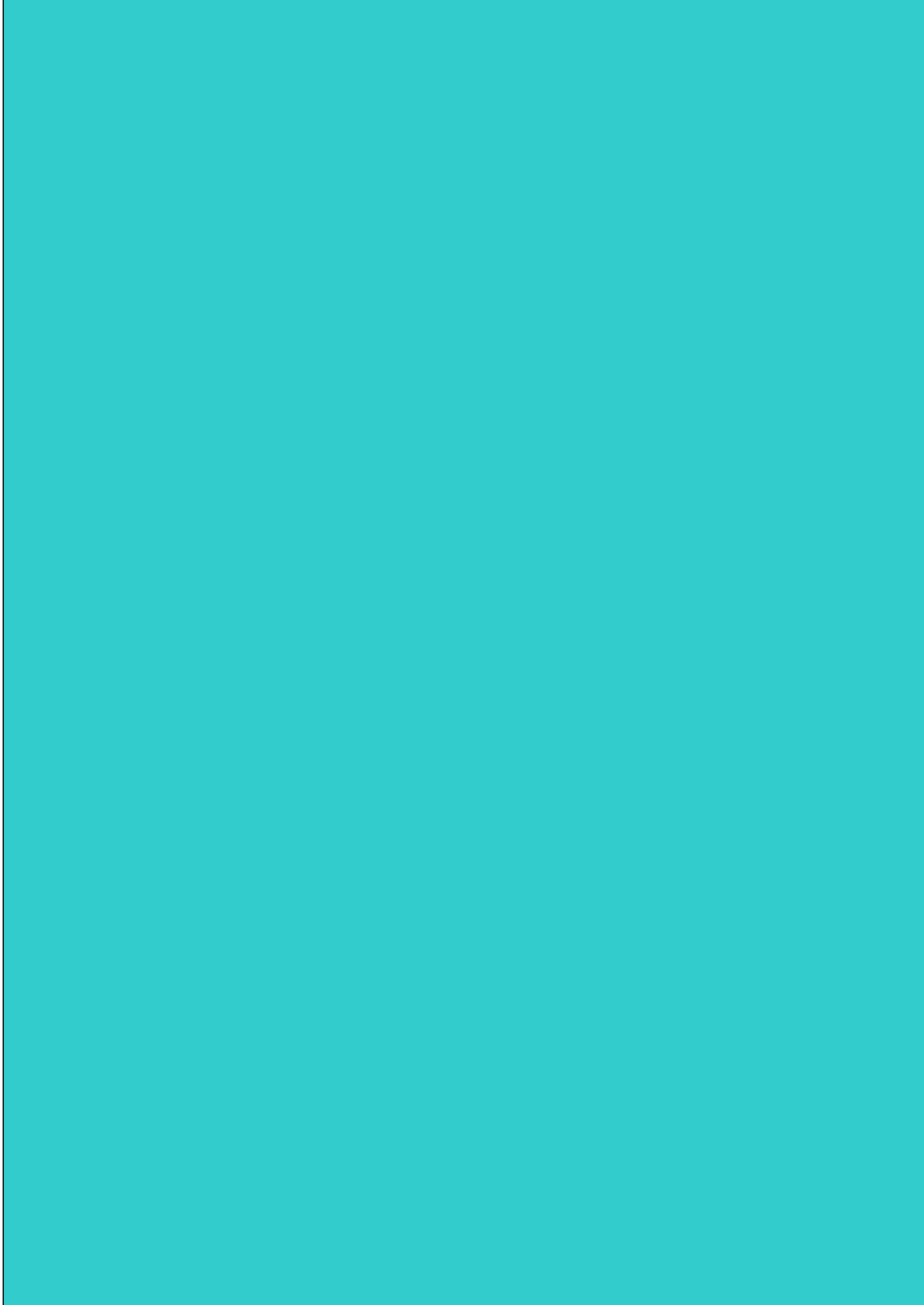
**Γιάννης Κρητικός**

**Διπλωματική Εργασία**

που υποβλήθηκε στο Τμήμα Στατιστικής και  
Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς  
ως μέρος των απαιτήσεων για την απόκτηση του  
Μεταπτυχιακού Διπλώματος Ειδίκευσης στην  
*Εφαρμοσμένη Στατιστική*

Πειραιάς

ΜΑΪΟΣ 2024



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**

**Σχολή Χρηματοοικονομικής και Στατιστικής**



**Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ**

**ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΜΕΛΕΤΗ ΠΑΡΑΓΟΝΤΩΝ ΠΟΥ ΕΠΙΔΡΟΥΝ**

**ΣΤΗΝ ΕΠΙΛΟΓΗ ΑΘΛΗΤΙΚΗΣ**

**ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ ΜΕΣΩ**

**ΔΙΑΔΙΚΤΥΑΚΗΣ ΠΛΑΤΦΟΡΜΑΣ**

**Γιάννης Κρητικός**

**Διπλωματική Εργασία**

που υποβλήθηκε στο Τμήμα Στατιστικής και  
Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς  
ως μέρος των απαιτήσεων για την απόκτηση του  
Μεταπτυχιακού Διπλώματος Ειδίκευσης στην  
*Εφαρμοσμένη Στατιστική*

Πειραιάς

ΜΑΪΟΣ 2024

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμό ..... συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Ιωάννης Τριανταφύλλου (Επιβλέπων), Επίκουρος Καθηγητής
- Χαράλαμπος Ευαγγελάρας, Αναπληρωτής Καθηγητής
- Αθανάσιος Ρακιτζής, Επίκουρος Καθηγητής

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνωμών του συγγραφέα.

**UNIVERSITY OF PIRAEUS**  
**School of Finance and Statistics**



**Department of Statistics and Insurance Science**

**POSTGRADUATE PROGRAM IN**  
**APPLIED STATISTICS**

**STUDY OF FACTORS INFLUENCING THE**  
**SELECTION OF SPORTS ACTIVITIES**  
**THROUGH ONLINE PLATFORMS**

By

**Ioannis Kritikos**

MSc Dissertation

submitted to the Department of Statistics and  
Insurance Science of the University of Piraeus in partial  
fulfilment of the requirements for the degree of  
Master of Science in Applied Statistics

Piraeus, Greece

May 2024



*Στους γονείς μου  
Γιώργο και Μαρίνα*





## Περίληψη

Η παρούσα μελέτη εστίασε στη διερεύνηση των παραγόντων που επηρεάζουν τη συχνότητα συμμετοχής και το συνολικό χρηματικό ποσό που συνεισφέρουν οι χρήστες σε μια πλατφόρμα αθλητικών εγκαταστάσεων. Μέσω της χρήσης ενός πλήρους συνόλου δεδομένων, διενεργήθηκε μια σειρά στατιστικών αξιολογήσεων, που περιλαμβάνουν ANOVA, T-tests και συντελεστές συσχέτισης Spearman, για τον εντοπισμό υποκείμενων μοτίβων και συνδέσεων. Τα αποτελέσματα αποκαλύπτουν σημαντικές διακυμάνσεις στη δραστηριότητα και τις οικονομικές δαπάνες των χρηστών, βασισμένες σε δημογραφικά κριτήρια όπως η ηλικία και το φύλο, καθώς και σε μεταβλητές όπως η διάρκεια της συνδρομής και η επιλογή δραστηριοτήτων.

Διαπιστώθηκε ότι μια εκτεταμένη διάρκεια συνδρομής συμβαδίζει με αυξημένες χρηματικές συνεισφορές, υπονοώντας μια σημαντική δέσμευση μεταξύ των χρηστών και της πλατφόρμας. Επίσης, φάνηκαν διακρίσεις με βάση το φύλο, με τους άνδρες συμμετέχοντες να εμφανίζουν συχνότερα και υψηλότερες οικονομικές δαπάνες. Η ανάλυση επισήμανε επιπλέον την επίδραση ορισμένων δραστηριοτήτων στη συμπεριφορά των χρηστών. Οι δραστηριότητες που σχετίζονται με το νερό και οι ειδικές προγράμματα αν ανακαλύφθηκαν να προσελκύουν μεγαλύτερη συμμετοχή και χρηματική επένδυση. Η παρούσα ανάλυση προσφέρει πληροφορίες σχετικά με τις προτιμήσεις και τη συμπεριφορά των χρηστών, θέτοντας ένα στέρεο θεμέλιο για τη διαμόρφωση ακριβών στρατηγικών μάρκετινγκ και τη βελτίωση των υπηρεσιών. Αυτές στοχεύουν στην ενίσχυση της εμπειρίας του χρήστη και στην αύξηση της οικονομικής απόδοσης της πλατφόρμας.

Λέξεις-Κλειδιά: Χρήση Αθλητικών Εγκαταστάσεων, Εμπλοκή Χρηστών, Προτιμήσεις Δραστηριότητας, Στρατηγικές Μάρκετινγκ, Εμπειρία Πελατών, Διάρκεια Συνδρομής.

## **Abstract**

The investigation conducted in this study provides an extensive examination of the determinants that influence the frequency of participation and the total monetary contributions made by users on a sports facility platform. Through the utilization of a comprehensive dataset, a sequence of statistical evaluations was executed, encompassing ANOVA, T-tests, and Spearman's correlation coefficients, to identify underlying patterns and connections. The results elucidate notable variances in user activity and financial outlay, predicated on demographic criteria such as age and gender, in addition to variables like the length of subscription and the selection of activities.

It was discerned that an elongated duration of subscription is congruent with increased financial contributions, insinuating a substantial commitment between the users and the platform. Gender-based disparities were also manifest, with male participants typically exhibiting greater frequency in participation and higher financial expenditure. The analysis further highlighted the impact of particular activities on user conduct. Activities associated with water and specialized programs were discovered to attract greater engagement and fiscal investment.

The analysis presented herein offers pivotal insights into the predilections and conduct of users, laying a robust groundwork for the formulation of precise marketing stratagems and service enhancements. These are designed to augment the user experience and elevate the economic performance of the platform.

**Keywords:** Sports Facility Usage, User Engagement, Activity Preferences, Marketing Strategies, Customer Experience, Subscription Duration

## Περιεχόμενα

Λίστα Πινάκων .....	12
Λίστα Γραφημάτων .....	14
Εισαγωγή .....	15
1.2 Επαγωγική Στατιστική Συμπερασματολογία .....	26
1.3 Ανάλυση συσχέτισης.....	30
1.4 Ανάλυση παλινδρόμησης.....	32
1.5 Μέθοδοι Γραφικής Απεικόνισης .....	33
1.6 Δεοντολογικά ζητήματα στην ανάλυση δεδομένων .....	35
Κεφάλαιο 4: Αποτελέσματα.....	48
4.1 Περιγραφικά Στατιστικά.....	48
4.2 Παλινδρόμηση .....	78
4.3 ANOVA .....	86
4.4 T-test.....	88
4.5 Συντελεστής Spearman.....	89

## Λίστα Πινάκων

<a href="#">Πίνακας 1: Σύνοψη των μεταβλητών που περιλαμβάνουν ημερομηνίες</a>	48
<a href="#">Πίνακας 2: Στοιχεία Περιγραφικής Στατιστικής</a>	41
<a href="#">Πίνακας 3: Στοιχεία Περιγραφικής Στατιστικής</a>	42
<a href="#">Πίνακας 4: Κατανομή Συχνότητας Ηλικιακών Ομάδων Χρηστών</a>	42
<a href="#">Πίνακας 5: Συχνότητα Φύλου Χρηστών</a>	49
<a href="#">Πίνακας 6: Συχνότητα Διάρκειας Εγγραφής Hughes</a>	50
<a href="#">Πίνακας 7: Συχνότητα Διάρκειας Εγγραφής (2 μήνες έως 12+ μήνες)</a>	51
<a href="#">Πίνακας 8: Συχνότητα Αποχής Χρηστών Πριν τη Λήξη Εγγραφής Hughes</a>	51
<a href="#">Πίνακας 9: Συχνότητα Ημερών Αδράνειας Πριν τη Λήξη Εγγραφής (2)</a>	52
<a href="#">Πίνακας 10: Συχνότητα Συνολικού Ποσού Πληρωμής Hughes</a>	53
<a href="#">Πίνακας 11: Συχνότητα Μέσης Εβδομαδιαίας Συχνότητας Κατά τη Διάρκεια Εγγραφής Hughes</a>	64
<a href="#">Πίνακας 12: Συχνότητα Εγγεγραμμένων Χρηστών στην Περίοδο</a>	64
<a href="#">Πίνακας 13: Συχνότητα Συμμετοχής σε Στίβο</a>	65
<a href="#">Πίνακας 14: Συχνότητα Συμμετοχής σε Δραστηριότητες Νερού</a>	66
<a href="#">Πίνακας 15: Συχνότητα Συμμετοχής στο Κολύμπι</a>	67
<a href="#">Πίνακας 16: Συχνότητα Συμμετοχής στον Χορό</a>	68
<a href="#">Πίνακας 17: Συχνότητα Συμμετοχής σε Ομαδικά Αθλήματα</a>	69
<a href="#">Πίνακας 18: Συχνότητα Συμμετοχής σε Αθλήματα Ρακέτας</a>	60
<a href="#">Πίνακας 19: Συχνότητα Συμμετοχής σε Πολεμικές Τέχνες</a>	60
<a href="#">Πίνακας 20: Συχνότητα Συμμετοχής σε Δραστηριότητες στη Φύση</a>	61
<a href="#">Πίνακας 21: Συχνότητα Συμμετοχής σε Ειδικές Δραστηριότητες</a>	71
<a href="#">Πίνακας 22: Συχνότητα Συμμετοχής σε Άλλες Δραστηριότητες</a>	63
<a href="#">Πίνακας 23: Συχνότητα Επισκέψεων στην Αθλητική Πλατφόρμα από Έναρξη Έως Διακοπή Εγγραφής Hughes</a>	63
<a href="#">Πίνακας 24: Συχνότητα Συμμετοχής σε Δραστηριότητες από Έναρξη Έως Διακοπή Εγγραφής Hughes</a>	74
<a href="#">Πίνακας 25: Συχνότητα Μέσου Αριθμού Συμμετοχής σε Δραστηριότητες από Έναρξη Έως Διακοπή Εγγραφής Hughes</a>	75
<a href="#">Πίνακας 26: Συχνότητα Σχέσης Πραγματικών Επισκέψεων προς Πιθανές Επισκέψεις (2)</a>	75
<a href="#">Πίνακας 27: Συχνότητα Αριθμού Ατόμων με Οικογενειακή ή Φιλική Σχέση με τον Χρήστη</a>	76

<u>Πίνακας 28: Συχνότητα Κατάστασης Χρήστη στην Υποδεικνυόμενη Ημερομηνία</u> .....	77
<u>Πίνακας 29: Περίληψη Μοντέλου Παλινδρόμησης Συνολικού ποσού πληρωμής</u> .....	78
<u>Πίνακας 30: Πίνακας ANOVA Συνολικού ποσού πληρωμής</u> .....	80
<u>Πίνακας 31: Πίνακας Συντελεστών Συνολικού ποσού πληρωμής</u> .....	70
<u>Πίνακας 32: Περίληψη Μοντέλου Παλινδρόμησης Μέσης εβδομαδιαίας συχνότητας σε όλη την περίοδο εγγραφής</u> .....	71
<u>Πίνακας 33: Πίνακας ANOVA Μέσης εβδομαδιαίας συχνότητας σε όλη την περίοδο εγγραφής</u> .....	72
<u>Πίνακας 34: Πίνακας Συντελεστών Μέσης εβδομαδιαίας συχνότητας σε όλη την περίοδο εγγραφής</u> .....	73
<u>Πίνακας 35: Περίληψη Μοντέλου Παλινδρόμησης Σχέσης μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει</u> .....	84
<u>Πίνακας 36: Πίνακας ANOVA Σχέσης μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει</u> .....	85
<u>Πίνακας 37: Πίνακας Συντελεστών Σχέσης μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει</u> .....	855
<u>Πίνακας 38: Επίδραση Ηλικιακής Ομάδας στο Συνολικό Ποσό Πληρωμής</u> .....	86
<u>Πίνακας 39: Επίδραση Διάρκειας Συνδρομής στο Συνολικό Ποσό Πληρωμής</u> .....	87
<u>Πίνακας 40: Επίδραση Ηλικιακής Ομάδας στη Μέση Εβδομαδιαία Συχνότητα Συμμετοχής</u> .....	87
<u>Πίνακας 41: Επίδραση Διάρκειας Συνδρομής στη Μέση Εβδομαδιαία Συχνότητα Συμμετοχής</u> .....	88
<u>Πίνακας 42: Στατιστικά Στοιχεία ανά Φύλο για Μέση Εβδομαδιαία Συχνότητα</u> .....	88
<u>Πίνακας 43: Στατιστικά Στοιχεία Φύλου για Συνολικό Ποσό Πληρωμής</u> .....	89
<u>Πίνακας 44: Συσχετίσεις Spearman</u> .....	80

## Λίστα Γραφημάτων

<a href="#">Γράφημα 1: Διάγραμμα κουτιού της διάρκειας εγγραφής χρηστών</a>	42
<a href="#">Γράφημα 2: Θηκόγραμμα Ηλικίας</a>	43
<a href="#">Γράφημα 3: Θηκόγραμμα Αριθμού Ημερών Πριν Θεωρηθεί Λήξη Εγγραφής</a>	43
<a href="#">Γράφημα 4: Θηκόγραμμα Συνολικού Ποσού Πληρωμής</a>	44
<a href="#">Γράφημα 5: Θηκόγραμμα Μέσης Εβδομαδιαίας Συχνότητας σε Όλη την Περίοδο Εγγραφής</a>	44
<a href="#">Γράφημα 6: Θηκόγραμμα Αριθμού Δραστηριοτήτων</a>	45
<a href="#">Γράφημα 7: Θηκόγραμμα Αριθμού Επισκέψεων Από Έναρξη Έως Διακοπή Εγγραφής</a>	45
<a href="#">Γράφημα 8: Θηκόγραμμα Μέσου Αριθμού Συμμετοχής σε Δραστηριότητες</a>	46
<a href="#">Γράφημα 9: Θηκόγραμμα Εβδομαδιαίων Επισκέψεων Τελευταίων 2 Μηνών Εγγραφής</a>	55
<a href="#">Γράφημα 10: Θηκόγραμμα Αριθμού Επισκέψεων στις Εγκαταστάσεις τους Τελευταίους 2 Μήνες Εγγραφής</a>	56
<a href="#">Γράφημα 11: Θηκόγραμμα Πραγματικού Αριθμού Επισκέψεων</a>	56
<a href="#">Γράφημα 12: Θηκόγραμμα Σχέσης Πραγματικών και Θεωρητικών Επισκέψεων</a>	57
<a href="#">Γράφημα 13: Θηκόγραμμα Αριθμού Ανανεώσεων Κατά την Περίοδο Εγγραφή</a>	57
<a href="#">Γράφημα 14: Κατανομή Συχνότητας Ηλικιακών Ομάδων Χρηστών</a>	58
<a href="#">Γράφημα 15: Συχνότητα Φύλου Χρηστών</a>	50
<a href="#">Γράφημα 16: Συχνότητα Διάρκειας Εγγραφής (2 μήνες έως 12+ μήνες)</a>	52
<a href="#">Γράφημα 17: Συχνότητα Ημερών Αδράνειας Πριν τη Λήξη Εγγραφής (2)</a>	53
<a href="#">Γράφημα 18: Συχνότητα Εγγεγραμμένων Χρηστών στην Περίοδο</a>	65
<a href="#">Γράφημα 19: Συχνότητα Συμμετοχής σε Στίβο</a>	66
<a href="#">Γράφημα 20: Συχνότητα Συμμετοχής σε Δραστηριότητες Νερού</a>	67
<a href="#">Γράφημα 21: Συχνότητα Συμμετοχής στο Κολύμπι</a>	68
<a href="#">Γράφημα 22: Συχνότητα Συμμετοχής σε Ομαδικά Αθλήματα</a>	60
<a href="#">Γράφημα 23: Συχνότητα Συμμετοχής σε Αθλήματα Ρακέτας</a>	61
<a href="#">Γράφημα 24: Συχνότητα Συμμετοχής σε Πολεμικές Τέχνες</a>	62
<a href="#">Γράφημα 25: Συχνότητα Συμμετοχής σε Ειδικές Δραστηριότητες</a>	63
<a href="#">Γράφημα 26: Συχνότητα Συμμετοχής σε Άλλες Δραστηριότητες</a>	64
<a href="#">Γράφημα 27: Συχνότητα Σχέσης Πραγματικών Επισκέψεων προς Πιθανές Επισκέψεις (2)</a>	67
<a href="#">Γράφημα 28: Συχνότητα Αριθμού Ατόμων με Οικογενειακή ή Φιλική Σχέση με τον Χρήστη</a>	68
<a href="#">Γράφημα 29: Συχνότητα Κατάστασης Χρήστη στην Υποδεικνυόμενη Ημερομηνία</a>	69

## Εισαγωγή

Στη σύγχρονη εποχή, το ψηφιακό πλαίσιο επικοινωνίας και κοινωνικοποίησης έχει αναδιαμορφώσει τον τρόπο με τον οποίο τα άτομα μπορούν να αλληλεπιδρούν με τους τομείς του αθλητισμού και διαφόρων άλλων ψυχαγωγικών δραστηριοτήτων. Η αυξανόμενη επιρροή των διαδικτυακών πηγών έχει εισαγάγει μια μεταρρύθμιση στη δυναμική της αθλητικής ενασχόλησης, παρέχοντας στους χρήστες πρόσβαση σε μια ποικιλία από αθλητικές δραστηριότητες. Η κατανόηση των παραγόντων που επηρεάζουν την επιλογή των αθλητικών δραστηριοτήτων ενός ατόμου στη σφαίρα αυτών των διαδικτυακών πηγών είναι καίριας σημασίας για τις επιχειρήσεις, τους οργανισμούς και τους ενδιαφερόμενους φορείς που επιδιώκουν να προσαρμόσουν τις προσφορές τους στις προτιμήσεις και τις κλίσεις του κοινού-στόχου τους.

Ο πολλαπλασιασμός των δεδομένων που δημιουργούνται από αυτές τις διαδικτυακές πλατφόρμες αποτελεί ευκαιρία να γίνει ανάλυση των υποκείμενων τάσεων, των προτιμήσεων και των προτύπων συμπεριφοράς που διέπουν την επιλογή αθλητικών δραστηριοτήτων. Μέσω της ανάλυσης δεδομένων, καθίσταται δυνατός ο εντοπισμός των περιπλοκών της συμπεριφοράς των χρηστών, η διάκριση των παραγόντων που οδηγούν στις προτιμήσεις αθλητικών δραστηριοτήτων και ο εντοπισμός πιθανών τομέων για στρατηγική παρέμβαση και βελτίωση. Αξιοποιώντας τη δύναμη της στατιστικής ανάλυσης, μπορεί να επιτευχθεί μια ολοκληρωμένη κατανόηση των κλίσεων των χρηστών, θέτοντας τα θεμέλια για τεκμηριωμένη λήψη αποφάσεων και εφαρμογή στοχευμένων στρατηγικών για τη βελτιστοποίηση της αφοσίωσης και της ικανοποίησης των χρηστών.

Η παρούσα εργασία στοχεύει να διερευνήσει τη σύνθετη αλληλεπίδραση παραγόντων που επηρεάζουν την επιλογή αθλητικών δραστηριοτήτων μέσω διαδικτυακών πλατφορμών. Αξιοποιώντας την ισχύ των στατιστικών μεθοδολογιών, η μελέτη προσπαθεί να αναλύσει τις προτιμήσεις των χρηστών, τις δημογραφικές παραλλαγές και τους ψυχολογικούς καθοριστικούς παράγοντες που καθοδηγούν την πορεία της ενασχόλησης με τον αθλητισμό στην ψηφιακή σφαίρα. Επιπλέον, επιδιώκεται να γεφυρωθεί το χάσμα μεταξύ της θεωρητικής προσέγγισης και των πρακτικών επιπτώσεων, διευκολύνοντας την ανάπτυξη εφαρμοσμένων γνώσεων που μπορούν να συμβάλουν στη βελτίωση των στρατηγικών μάρκετινγκ, στη

βελτιστοποίηση της εμπειρίας των χρηστών και στην προώθηση ενός πιο προσαρμοσμένου πλαισίου αθλητικής ενασχόλησης.

Τα επόμενα κεφάλαια θα εστιάσουν στις θεμελιώδεις αρχές της στατιστικής ανάλυσης, παρουσιάζοντας τα απαραίτητα εργαλεία και τις τεχνικές που συμβάλλουν στην ανάλυση δεδομένων και στην εξαγωγή συμπερασμάτων. Μέσα από το πρίσμα των περιγραφικών στατιστικών και της επαγωγικής στατιστικής συμπερασματολογίας, γίνεται προσπάθεια ανάδειξης υποκείμενων τάσεων και συσχετίσεων μεταξύ διαφόρων μεταβλητών, προσφέροντας μια ολοκληρωμένη κατανόηση των βασικών παραγόντων που επηρεάζουν τις επιλογές αθλητικών δραστηριοτήτων στη σφαίρα των διαδικτυακών πλατφορμών.

Υπό το πρίσμα των εξελισσόμενων ηθικών κριτηρίων που αφορούν το απόρρητο και την ασφάλεια των δεδομένων, αυτή η έρευνα θα διατηρήσει τα πρότυπα ακεραιότητας και εμπιστευτικότητας των δεδομένων, διασφαλίζοντας την προστασία των πληροφοριών των χρηστών και τηρώντας τις αρχές της δεοντολογικής έρευνας. Με την προώθηση μιας διαφανούς και υπεύθυνης προσέγγισης στην ανάλυση δεδομένων, η μελέτη στοχεύει να δημιουργήσει κλίμα εμπιστοσύνης μεταξύ των ερωτηθέντων, ενισχύοντας έτσι ένα ευνοϊκό περιβάλλον για διαρκή συμμετοχή.

Ο πρωταρχικός σκοπός της μελέτης είναι να εξετάσει τους πολύπλευρους παράγοντες που επηρεάζουν τις επιλογές αθλητικών δραστηριοτήτων και τα μοτίβα αφοσίωσης των χρηστών στο πλαίσιο της διαδικτυακής πλατφόρμας. Αξιοποιώντας προηγμένες στατιστικές μεθοδολογίες και τεχνικές ανάλυσης δεδομένων, η μελέτη στοχεύει να διερευνήσει τις υποκείμενες τάσεις και τη δυναμική συμπεριφοράς που διαμορφώνουν τις προτιμήσεις των χρηστών και τη συμμετοχή σε διάφορες αθλητικές δραστηριότητες. Επιπλέον, επιδιώκει να παρέχει αξιόπιστες πληροφορίες που μπορούν να συμβάλλουν στην ανάπτυξη στοχευμένων στρατηγικών, παρεμβάσεων και πρωτοβουλιών με επίκεντρο τον χρήστη που στοχεύουν στην ενίσχυση της ικανοποίησης των χρηστών, της διατήρησης και της συνολικής εμπλοκής στο περιβάλλον των ψηφιακών αθλημάτων. Τα βασικά ερευνητικά ερωτήματα συνοψίζονται ακολούθως:



Ερευνητικά ερωτήματα:

- ✚ Ποιος είναι ο αντίκτυπος δημογραφικών στοιχείων όπως η ηλικία, το φύλο και η διάρκεια εγγραφής στην αφοσίωση των χρηστών και στις προτιμήσεις αθλητικών δραστηριοτήτων εντός της διαδικτυακής πλατφόρμας;
- ✚ Πώς διαφορετικές κατηγορίες δραστηριοτήτων, συμπεριλαμβανομένων των αθλημάτων, των δραστηριοτήτων νερού, των προγραμμάτων φυσικής κατάστασης, του χορού, των ομαδικών αθλημάτων και άλλων, επηρεάζουν τη συμμετοχή των χρηστών και τα μοτίβα αφοσίωσης στην διαδικτυακή αθλητική πλατφόρμα;
- ✚ Ποια είναι η σχέση μεταξύ της συχνότητας των επισκέψεων και των ποσοστών εγκατάλειψης μεταξύ των χρηστών και πώς συσχετίζονται αυτοί οι παράγοντες με τους τύπους δραστηριοτήτων και τα επίπεδα αφοσίωσης που εμφανίζονται από τη βάση χρηστών;
- ✚ Πώς επηρεάζουν μεταβλητές, όπως η μέση συχνότητα, τα μαθήματα και οι επιτρεπόμενες εβδομαδιαίες συμμετοχές, τη διατήρηση των χρηστών και τα ποσοστά εγκατάλειψης στο πλαίσιο της αθλητικής δέσμευσης στο διαδίκτυο;
- ✚ Τι ρόλο παίζει η παρουσία αναφορών, ο αριθμός των ανανεώσεων και οι συγκεκριμένες προτιμήσεις δραστηριότητας στην πρόβλεψη της μακροπρόθεσμης αφοσίωσης και συμμετοχής των χρηστών στην πλατφόρμα ψηφιακών αθλημάτων;



# ΚΕΦΑΛΑΙΟ 1

## Στατιστικές Μέθοδοι Ανάλυσης Δεδομένων

### 1.1 Περιγραφικά στατιστικά στοιχεία

Στον τομέα της ανάλυσης αθλητικών δραστηριοτήτων μέσω διαδικτυακών πλατφορμών, η εφαρμογή περιγραφικών στατιστικών χρησιμεύει ως θεμελιώδης πυλώνας για την κατανόηση και τη σύνοψη των χαρακτηριστικών που υπάρχουν στα δεδομένα. Με την αποσαφήνιση των κεντρικών τάσεων και μεταβλητών εντός του συνόλου δεδομένων, τα περιγραφικά στατιστικά στοιχεία διαδραματίζουν καθοριστικό ρόλο στην αποκάλυψη της υποκείμενης κατανομής και δομής των προτιμήσεων αθλητικών δραστηριοτήτων, διευκολύνοντας έτσι την κατανόηση της συμπεριφοράς και των προτιμήσεων των χρηστών.

Επίκεντρο της περιγραφικής στατιστικής είναι η εξέταση των μέτρων κεντρικής τάσης. Αυτά τα μέτρα, συμπεριλαμβανομένου του μέσου όρου, χρησιμεύουν ως κύριοι δείκτες των κεντρικών τιμών γύρω από τις οποίες συγκεντρώνονται τα δεδομένα (Nick (2007)). Ο μέσος όρος, που υπολογίζεται ως το άθροισμα όλων των τιμών διαιρούμενο με τον συνολικό αριθμό των παρατηρήσεων, προσφέρει πληροφορίες για τη μέση τιμή που αντιπροσωπεύει καλύτερα το σύνολο δεδομένων (Fisher & Marshall (2009)). Ο μέσος όρος αποτελεί μία βασική έννοια στην περιγραφική στατιστική, παρέχοντας πληροφορίες σχετικά με την κεντρική τάση ενός συνόλου δεδομένων. Η μαθηματική αναπαράσταση του μέσου όρου ( $\mu$ ) για ένα σύνολο δεδομένων με  $N$  παρατηρήσεις ( $x_1, x_2, \dots, x_N$ ) δίνεται από την εξίσωση:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i, \quad (1.1)$$

όπου

- $x_i$  αντιπροσωπεύει κάθε τιμή στο σύνολο δεδομένων,
- $N$  είναι ο συνολικός αριθμός των παρατηρήσεων.

Η εξίσωση αυτή αναδεικνύει τη σημασία του μέσου όρου ως μέτρου κεντρικής τάσης, δείχνοντας τη μέση τιμή γύρω από την οποία κατανέμονται οι τιμές των δεδομένων.

Ο μέσος όρος είναι ιδιαίτερα χρήσιμος για την κατανόηση του γενικού επιπέδου ενός χαρακτηριστικού στον πληθυσμό που μελετάται. Στο πλαίσιο της διατριβής σας για τις επιλογές αθλητικών δραστηριοτήτων μέσω διαδικτυακής πλατφόρμας, ο μέσος όρος μπορεί να εφαρμοστεί για την ανάλυση διάφορων ποσοτικών πτυχών, όπως η μέση ηλικία των συμμετεχόντων, η μέση διάρκεια αφοσίωσης σε μια συγκεκριμένη δραστηριότητα ή η μέση συχνότητα επισκέψεων των χρηστών στην πλατφόρμα.

Ωστόσο, είναι σημαντικό να αναγνωριστεί η ευαισθησία του μέσου όρου σε ακραίες τιμές (outliers), δηλαδή σε τιμές που διαφέρουν σημαντικά από τις υπόλοιπες. Σε τέτοιες περιπτώσεις, ο μέσος όρος μπορεί να μην αντικατοπτρίζει με ακρίβεια την κεντρική τάση του συνόλου δεδομένων. Αυτός ο περιορισμός τονίζει τη σημασία συμπληρωματικών μέτρων, όπως η διάμεσος και η επικρατούσα, για την παροχή μιας πιο αξιόπιστης κατανόησης της κατανομής των δεδομένων.

Αποτελεί ένα ορόσημο που έθεσε τις βάσεις για τις σύγχρονες στατιστικές τεχνικές, συμπεριλαμβανομένου του υπολογισμού και της ερμηνείας του μέσου όρου (Fisher (1925)). Επιπλέον, παρέχει μια εκτενή επισκόπηση των στατιστικών βασικών αρχών, συμπεριλαμβανομένου του μέσου όρου, και είναι κατάλληλο τόσο για αρχάριους όσο και για προχωρημένους ερευνητές που επιδιώκουν να εμβαθύνουν την κατανόησή τους στις στατιστικές μεθοδολογίες (Bulmer (1979)).

Η διάμεσος, που βρίσκεται στη μέση των δεδομένων όταν είναι διατεταγμένα σε αύξουσα ή φθίνουσα σειρά, παρέχει ένα ισχυρό μέτρο που είναι ανθεκτικό σε ακραίες τιμές, προσφέροντας έτσι μια πιο αντιπροσωπευτική κεντρική τάση παρά την παρουσία ακραίων τιμών (Kaur et al. (2018)). Είναι ένα στατιστικό μέτρο που χρησιμοποιείται για να καθορίσει την τιμή στο κέντρο ενός δεδομένου συνόλου παρατηρήσεων. Όταν οι τιμές του συνόλου δεδομένων ταξινομούνται σε αύξουσα ή φθίνουσα σειρά, η διάμεσος είναι η τιμή που βρίσκεται στην ακριβή μέση, έτσι ώστε ο μισός αριθμός των παρατηρήσεων να είναι μικρότερος και ο μισός μεγαλύτερος από αυτήν.

Για ένα σύνολο δεδομένων  $X = (x_1, x_2, \dots, x_N)$  με  $n$  ταξινομημένες τιμές, η διάμεσος ( $M$ ) υπολογίζεται ως εξής:

1. Αν  $n$  είναι περιττός αριθμός, η διάμεσος είναι η τιμή που βρίσκεται στη θέση  $\frac{n+1}{2}$  μετά την ταξινόμηση των δεδομένων, δηλαδή υπολογίζεται σύμφωνα με τον τύπο

$$M = x_{\frac{n+1}{2}} \quad (1.2)$$

2. Αν  $n$  είναι άρτιος αριθμός, η διάμεσος είναι ο μέσος όρος των δύο μεσαίων τιμών, δηλαδή των τιμών που βρίσκονται στις θέσεις  $\frac{n}{2}$  και  $\frac{n}{2} + 1$ , δηλαδή υπολογίζεται σύμφωνα με τον τύπο

$$M = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2} + 1}}{2} \quad (1.3)$$

Η διάμεσος παρέχει ένα ανθεκτικό μέτρο κεντρικής τάσης, που δεν επηρεάζεται εύκολα από ακραίες τιμές σε αντίθεση με τον μέσο όρο. Αυτό την καθιστά ιδιαίτερα χρήσιμη σε σύνολα δεδομένων με σημαντική ασυμμετρία ή όταν υπάρχουν ακραίες παρατηρήσεις που μπορεί να παραμορφώσουν την πραγματική εικόνα της κεντρικής τάσης του συνόλου.

Παρουσιάζεται μια προσβάσιμη εισαγωγή στις μη παραμετρικές στατιστικές μεθόδους, συμπεριλαμβανομένης της διάμεσου, και την εφαρμογή τους στην ανάλυση δεδομένων (Corder, & Foreman (2009)).

Η επικρατούσα τιμή, που υποδηλώνει την πιο συχνά εμφανιζόμενη τιμή στο σύνολο δεδομένων, παρέχει πληροφορίες σχετικά με τις κυρίαρχες επιλογές ή προτιμήσεις που επιδεικνύει η βάση χρηστών, προσφέροντας έτσι πολύτιμες πληροφορίες για τις πιο δημοφιλείς αθλητικές δραστηριότητες που ευνοούνται στην κοινότητα της διαδικτυακής πλατφόρμας. Για την επικρατούσα τιμή, δεν υπάρχει μια ενιαία μαθηματική εξίσωση για τον υπολογισμό της, όπως συμβαίνει με τον μέσο όρο ή τη διάμεσο. Αυτό οφείλεται στο ότι η τάση βασίζεται στη συχνότητα εμφάνισης των τιμών και όχι στην αριθμητική τους τιμή. Η εύρεση της τάσης περιλαμβάνει την καταμέτρηση του πλήθους εμφάνισης κάθε διαφορετικής τιμής στο σύνολο δεδομένων και την αναγνώριση της τιμής με τη μεγαλύτερη συχνότητα (Bulmer, M.G. (1979)).

Πέρα από τα μέτρα κεντρικής τάσης, τα μέτρα ασυμμετρίας και κύρτωσης παρέχουν επιπλέον πληροφορίες σχετικά με τη σχήμα και την κατανομή των δεδομένων. Τα μέτρα ασυμμετρίας αναφέρονται στο βαθμό με τον οποίο η κατανομή των δεδομένων αποκλίνει από τη συμμετρία γύρω από τον μέσο όρο (Cramer (2003)). Μια θετική

ασυμμετρία δηλώνει μια κατανομή με μακριά ουρά προς τα δεξιά, ενώ μια αρνητική ασυμμετρία υποδηλώνει μια κατανομή με μακριά ουρά προς τα αριστερά (Cramer (2003)). Αυτό το μέτρο είναι κρίσιμο για την κατανόηση της πιθανής διασποράς στην κατανομή των δεδομένων και μπορεί να επηρεάσει την ερμηνεία των άλλων στατιστικών μέτρων. Ο συντελεστής ασυμμετρίας, ο οποίος συχνά συμβολίζεται με  $\gamma_1$ , υπολογίζεται ως εξής:

$$\gamma_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left( \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \right)^3}, \quad (1.3)$$

όπου:

- $x_i$  είναι η τιμή της  $i$ -οστής παρατήρησης,
- $\bar{x}$  είναι ο μέσος όρος των παρατηρήσεων,
- $n$  είναι ο αριθμός των παρατηρήσεων.

Όταν  $\gamma_1 > 0$ , η κατανομή έχει θετική ασυμμετρία και υποδηλώνει μια κατανομή με μακριά ουρά προς τα δεξιά. Αντίστοιχα, όταν  $\gamma_1 < 0$ , η κατανομή έχει αρνητική ασυμμετρία και υποδηλώνει μια κατανομή με μακριά ουρά προς τα αριστερά. Μια τιμή του  $\gamma_1$  κοντά στο 0 υποδηλώνει μια σχετικά συμμετρική κατανομή. Η κατανόηση της ασυμμετρίας είναι κρίσιμη, καθώς μπορεί να επηρεάσει την ερμηνεία άλλων στατιστικών μετρήσεων, όπως ο μέσος όρος και η διάμεσος, και να υποδηλώσει την παρουσία πιθανών ακραίων τιμών ή της ύπαρξης δύο διακριτών υποομάδων εντός του συνόλου δεδομένων (Cramer (2003)).

Τα μέτρα κύρτωσης, περιγράφουν το βαθμό κορεσμού της κατανομής των δεδομένων σε σύγκριση με μια κανονική κατανομή. Υψηλές τιμές κύρτωσης υποδηλώνουν μια κατανομή με παχιές ουρές και μια πιο αμβλεία κορυφή, ενώ χαμηλές τιμές κύρτωσης υποδηλώνουν μια πιο επίπεδη κατανομή (Westfall (2014)). Η κύρτωση είναι ιδιαίτερα χρήσιμη για την ανίχνευση της παρουσίας ακραίων τιμών στα δεδομένα και του πόσο συχνά εμφανίζονται σε σύγκριση με μια κανονική κατανομή (Westfall (2014)). Ο συντελεστής κύρτωσης, συνήθως συμβολίζεται ως  $\beta_2$  ή κάποιες φορές ως  $\gamma_2$ , υπολογίζεται με την επόμενη έκφραση:

$$\beta_2 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{(\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2})^4} \quad (1.4)$$

όπου:

- $x_i$  είναι η τιμή της  $i$ -οστής παρατήρησης,
- $\bar{x}$  είναι ο μέσος όρος των παρατηρήσεων,
- $n$  είναι ο αριθμός των παρατηρήσεων.

Η τιμή της κύρτωσης συγκρίνεται συνήθως με την τιμή 3, που αντιπροσωπεύει την κύρτωση μιας τυπικής κανονικής κατανομής. Τιμές κύρτωσης μεγαλύτερες από 3 υποδηλώνουν κατανομή με "παχιές" ουρές και μια πιο αμβλεία κορυφή (λεπτοκορυφή κατανομή), ενώ τιμές μικρότερες από 3 υποδηλώνουν μια πιο επίπεδη κατανομή με λιγότερο πυκνές ουρές (πλατυκορυφή κατανομή).

Η κατανόηση της κύρτωσης είναι ουσιαστική για την εκτίμηση του κινδύνου και της πιθανότητας ακραίων τιμών σε ένα σύνολο δεδομένων, καθώς κατανομές με υψηλή κύρτωση τείνουν να παρουσιάζουν περισσότερες ακραίες τιμές σε σύγκριση με μια κανονική κατανομή.

Σε συνδυασμό με μέτρα κεντρικής τάσης, η περιγραφική στατιστική περιλαμβάνει την εξέταση των μέτρων μεταβλητότητας. Αυτά τα μέτρα, συμπεριλαμβανομένου του εύρους, της διακύμανσης και της τυπικής απόκλισης, φανερώνουν τη διασπορά ή την εξάπλωση των δεδομένων γύρω από τις κεντρικές τάσεις (Fisher & Marshall (2009)). Το εύρος, που υποδηλώνει τη διαφορά μεταξύ των μέγιστων και ελάχιστων τιμών εντός του συνόλου δεδομένων, παρέχει μια συνοπτική κατανόηση της έκτασης της μεταβλητότητας που εμφανίζεται στα δεδομένα (Nick (2007)). Υπολογίζεται ως η διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής του συνόλου δεδομένων. Μαθηματικά, το εύρος ( $R$ ) περιγράφεται από την εξίσωση:

$$R = x_{\max} - x_{\min} \quad (1.5)$$

όπου:

- $x_{\max}$  είναι η μέγιστη τιμή του συνόλου δεδομένων,
- $x_{\min}$  είναι η ελάχιστη τιμή του συνόλου δεδομένων.

Παρόλο που το εύρος είναι ένα πολύ απλό μέτρο και παρέχει μια γρήγορη ένδειξη της μεταβλητότητας, δεν λαμβάνει υπόψη την κατανομή των τιμών εντός του συνόλου δεδομένων και μπορεί να επηρεαστεί σημαντικά από ακραίες τιμές (DeGroot & Schervish (2012)). Ως εκ τούτου, συχνά χρησιμοποιούνται επιπλέον μέτρα διακύμανσης, όπως η τυπική απόκλιση και η ενδιάμεση εμβέλεια, για μια πιο ολοκληρωμένη αξιολόγηση της μεταβλητότητας.

Η διακύμανση, ποσοτικοποιώντας τη μέση τετραγωνική απόκλιση κάθε σημείου δεδομένων από τη μέση τιμή, προσφέρει μια ολοκληρωμένη αξιολόγηση της συνολικής μεταβλητότητας, διευκολύνοντας έτσι κατανόηση της κατανομής δεδομένων (Kaur et al. (2018)). Η διακύμανση (συμβολίζεται συνήθως ως  $\sigma^2$  για τον πληθυσμό και  $s^2$  για το δείγμα) υπολογίζεται σύμφωνα με τις επόμενες εκφράσεις:

- Για τον πληθυσμό:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{N} \quad (1.6)$$

- Για το δείγμα:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (1.7)$$

όπου:

- $x_i$  είναι η τιμή της  $i$ -οστής παρατήρησης,
- $\mu$  είναι ο μέσος όρος του πληθυσμού,
- $\bar{x}$  είναι ο μέσος όρος του δείγματος,
- $N$  είναι το μέγεθος του πληθυσμού,
- $n$  είναι το μέγεθος του δείγματος.

Η διακύμανση παρέχει μια ακριβή ένδειξη της συνολικής μεταβλητότητας ενός συνόλου δεδομένων, δίνοντας έμφαση στην ποσότητα της διακύμανσης που υπάρχει. Ωστόσο, επειδή οι μονάδες της διακύμανσης είναι τετραγωνικές σε σχέση με τις μονάδες των δεδομένων, συχνά χρησιμοποιείται η τυπική απόκλιση (η οποία είναι η τετραγωνική ρίζα της διακύμανσης) για μια πιο άμεσα κατανοητή μέτρηση της μεταβλητότητας (Moore & McCabe (2003)).

Η τυπική απόκλιση, η τετραγωνική ρίζα της διακύμανσης, χρησιμεύει ως μέτρηση για την οριοθέτηση του βαθμού διασποράς γύρω από τον μέσο όρο, βοηθώντας έτσι στην αξιολόγηση της συνέπειας ή της μεταβλητότητας των προτιμήσεων των χρηστών



για διάφορες αθλητικές δραστηριότητες (George & Mallery (2018)). Η τυπική απόκλιση για ένα δείγμα υπολογίζεται σύμφωνα με τις επόμενες εκφράσεις:

- Για τον πληθυσμό:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{N}} \quad (1.8)$$

- Για το δείγμα:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (1.9)$$

όπου:

- $x_i$  είναι η τιμή της  $i$ -οστής παρατήρησης,
- $\mu$  είναι ο μέσος όρος του πληθυσμού,
- $\bar{x}$  είναι ο μέσος όρος του δείγματος,
- $N$  είναι το μέγεθος του πληθυσμού,
- $n$  είναι το μέγεθος του δείγματος.

Συμπληρώνοντας αυτές τις αριθμητικές περιλήψεις, οι γραφικές αναπαραστάσεις των δεδομένων χρησιμεύουν ως απαραίτητα εργαλεία για την απεικόνιση και την ερμηνεία των υποκείμενων τάσεων και προτύπων (George & Mallery (2018)). Τα ιστογράμματα, για παράδειγμα, επιτρέπουν την απεικόνιση της κατανομής συχνότητας των αθλητικών δραστηριοτήτων, παρέχοντας μια ολοκληρωμένη επισκόπηση της σχετικής δημοτικότητας ή επικράτησης συγκεκριμένων αθλημάτων στην κοινότητα της διαδικτυακής πλατφόρμας. Οι γραφικές παραστάσεις, από την άλλη, διευκολύνουν μια διαφοροποιημένη κατανόηση της κατανομής, της διασποράς και των ακραίων τιμών εντός του συνόλου δεδομένων, βοηθώντας έτσι στον εντοπισμό πιθανών ασαφειών ή ιδιαιτεροτήτων στις προτιμήσεις αθλητικών δραστηριοτήτων.

Αξιοποιώντας τα περιγραφικά στατιστικά, αυτή η μελέτη στοχεύει να αναδείξει τις προτιμήσεις των χρηστών και τη δυναμική συμπεριφοράς στη σφαίρα της διαδικτυακής αθλητικής δέσμευσης. Μέσω μιας ολοκληρωμένης ανάλυσης των κεντρικών τάσεων, μεταβλητοτήτων και γραφικών αναπαραστάσεων, τίθενται οι βάσεις για την κατανόηση των υποκείμενων τάσεων και προτύπων που διέπουν τις επιλογές αθλητικών δραστηριοτήτων, διευκολύνοντας έτσι τη λήψη ενημερωμένων αποφάσεων

και την εφαρμογή στοχευμένων στρατηγικών στην ενίσχυση της αφοσίωσης και της ικανοποίησης των χρηστών των ψηφιακών αθλημάτων.

## 1.2 Επαγωγική Στατιστική Συμπερασματολογία

Υπερβαίνοντας την απλή περιγραφή των δεδομένων και εμβαθύνοντας στην ερμηνεία και στην πρόβλεψη δεδομένων, τα στοιχεία που προκύπτουν από την επαγωγική στατιστική συμπερασματολογία διαδραματίζουν κρίσιμο ρόλο στην εξαγωγή τεκμηριωμένων συμπερασμάτων σχετικά με τον μεγαλύτερο πληθυσμό με βάση ένα δείγμα δεδομένων (Marshall & Jonker (2011)).

Για να διεξαχθούν στατιστικά συμπεράσματα είναι σημαντική η κατανόηση των τεχνικών δειγματοληψίας, οι οποίες επιτρέπουν την επιλογή αντιπροσωπευτικών υποσυνόλων από τον μεγαλύτερο πληθυσμό (Asadoorian & Kantarelis (2005)). Μέσω μεθοδολογιών τυχαίας δειγματοληψίας, όπως η απλή τυχαία δειγματοληψία, η στρωματοποιημένη δειγματοληψία ή η δειγματοληψία σε ομάδες, οι ερευνητές μπορούν να εξαγάγουν συμπεράσματα σχετικά με την ευρύτερη βάση χρηστών, μετριάζοντας έτσι τους κινδύνους μεροληψίας και διασφαλίζοντας την αξιοπιστία και την εγκυρότητα των αποτελεσμάτων (Marshall & Jonker (2011)). Επιλέγοντας και αναλύοντας προσεκτικά ένα αντιπροσωπευτικό δείγμα, οι ερευνητές μπορούν να βγάλουν ακριβή και αξιόπιστα συμπεράσματα σχετικά με τις προτιμήσεις, τις τάσεις και τις κλίσεις ολόκληρου του πληθυσμού χρηστών που συμμετέχει σε αθλητικές δραστηριότητες μέσω διαδικτυακών πλατφορμών (Sahu et al. (2015)).

Βασικό στοιχείο της επαγωγικής στατιστικής συμπερασματολογίας είναι ο έλεγχος υποθέσεων, ο οποίος διευκολύνει τη συστηματική αξιολόγηση της εγκυρότητας των εικασιών ή των υποθέσεων που γίνονται για τον πληθυσμό με βάση τα δεδομένα του δείγματος (Asadoorian & Kantarelis (2005)). Διατυπώνοντας μηδενικές και εναλλακτικές υποθέσεις και υποβάλλοντας τα δεδομένα του δείγματος σε ακριβείς ελέγχους, οι ερευνητές μπορούν να διακρίνουν τη στατιστική σημασία των σχέσεων και των προτύπων που παρατηρούνται στο σύνολο δεδομένων (Sahu et al. (2015)). Μέσω της εφαρμογής ελέγχων και γενικότερα επαγωγικών μεθόδων στατιστικής συμπερασματολογίας, όπως t-test ή ANOVA, οι ερευνητές μπορούν να διερευνήσουν τη σημασία των συσχετίσεων μεταξύ διαφορετικών μεταβλητών, διευκολύνοντας έτσι

την ολοκληρωμένη κατανόηση των παραγόντων που επηρεάζουν σημαντικά τις επιλογές αθλητικών δραστηριοτήτων στην κοινότητα της διαδικτυακής πλατφόρμας.

Ο έλεγχος t-test είναι μια στατιστική δοκιμασία που χρησιμοποιείται για να συγκρίνει τα μέσα δύο δειγμάτων, για να διαπιστώσει εάν υπάρχει στατιστικά σημαντική διαφορά μεταξύ τους (Gosset (1908)). Οι βασικότεροι τύποι t-test περιλαμβάνουν τον ανεξάρτητο t-test (για δύο ανεξάρτητα δείγματα) και τον εξαρτημένο ή ζευγαρωμένο t-test (για δύο μη ανεξάρτητα δείγματα).

Προϋποθέσεις για τον έλεγχο t-test για ανεξάρτητα δείγματα:

- Ανεξαρτησία παρατηρήσεων: Οι παρατηρήσεις στις δύο ομάδες πρέπει να είναι ανεξάρτητες μεταξύ τους.
- Κανονικότητα: Οι παρατηρήσεις σε κάθε ομάδα θα πρέπει να προέρχονται από έναν κανονικά κατανομημένο πληθυσμό. Αυτή η προϋπόθεση μπορεί να μην είναι τόσο αυστηρή με τη χρήση του κεντρικού οριακού θεωρήματος για μεγάλα δείγματα.
- Ομοσκεδαστικότητα (Homoscedasticity): Οι διακυμάνσεις των δύο ομάδων πρέπει να είναι στατιστικά ίσες.

Οι δύο υποθέσεις που ελέγχονται στον συγκεκριμένο έλεγχο διατυπώνονται ως εξής:

- Μηδενική υπόθεση ( $H_0$ ): Δεν υπάρχει διαφορά μεταξύ των μέσων των δύο δειγμάτων.
- Εναλλακτική υπόθεση ( $H_1$ ): Υπάρχει μια στατιστικά σημαντική διαφορά μεταξύ των μέσων των δύο δειγμάτων.

Στατιστική συνάρτηση για έλεγχο t-test για ανεξάρτητα δείγματα:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}}, \quad (1.10)$$

όπου  $\bar{x}_1$  και  $\bar{x}_2$  είναι τα δειγματικά μέσα των δύο ομάδων, και  $s_{\bar{x}_1 - \bar{x}_2}$  είναι το τυπικό σφάλμα της διαφοράς μεταξύ των δύο μέσων.

Κριτήρια αποδοχής/απόρριψης: Εάν η τιμή  $t$  που υπολογίζεται υπερβαίνει την κρίσιμη τιμή  $t$  από τους πίνακες κατανομής  $t$  για το επιθυμητό επίπεδο σημαντικότητας (συνήθως 0.05 ή 0.01), τότε η μηδενική υπόθεση απορρίπτεται.

Προϋποθέσεις για τον εξαρτημένο ή ζευγαρωμένο t-test:

- Ζεύγη συνδεδεμένων δειγμάτων: Κάθε παρατήρηση στην μία ομάδα είναι συνδεδεμένη ή αντιστοιχεί σε μία παρατήρηση στην άλλη ομάδα.
- Κανονικότητα των διαφορών: Οι διαφορές μεταξύ των ζευγών των παρατηρήσεων πρέπει να ακολουθούν μια κανονική κατανομή.

Η ανάλυση διακύμανσης (ANOVA) χρησιμοποιείται για να συγκρίνει τα μέσα από δύο, τρεις ή περισσότερες δειγματικές ομάδες, για να διαπιστώσει αν υπάρχει τουλάχιστον μία στατιστικά σημαντική διαφορά μεταξύ των μέσων των ομάδων (Fisher (1918)).

Προϋποθέσεις:

- Ανεξαρτησία των παρατηρήσεων: Οι παρατηρήσεις μέσα σε κάθε ομάδα και μεταξύ των ομάδων πρέπει να είναι ανεξάρτητες.
- Κανονικότητα: Οι παρατηρήσεις σε κάθε ομάδα θα πρέπει να προέρχονται από κανονικά κατανεμημένους πληθυσμούς. Όπως και στον t-test, αυτή η προϋπόθεση μπορεί να χαλαρώσει για μεγάλα δείγματα λόγω του κεντρικού οριακού θεωρήματος.
- Ισότητα των διακυμάνσεων (Homogeneity of Variances): Οι διακυμάνσεις μεταξύ των ομάδων πρέπει να είναι στατιστικά ίσες, όπως επιβεβαιώνεται από τεστ όπως το Levene's test.

Η μη τήρηση αυτών των προϋποθέσεων μπορεί να οδηγήσει σε παραποιημένα αποτελέσματα και εσφαλμένα συμπεράσματα. Σε περίπτωση που οι προϋποθέσεις δεν πληρούνται, μπορεί να χρησιμοποιηθούν μη παραμετρικές μέθοδοι ως εναλλακτικές.

Υπόθεση:

- Μηδενική υπόθεση ( $H_0$ ): Όλες οι ομάδες έχουν το ίδιο μέσο.
- Εναλλακτική υπόθεση ( $H_1$ ): Τουλάχιστον μία ομάδα έχει μέσο διαφορετικό από τουλάχιστον μία άλλη.

Στατιστική συνάρτηση:

Η στατιστική συνάρτηση  $F$  του ελέγχου υπολογίζεται ως:

$$F = \frac{MSTR}{MSE}, \quad (1.11)$$

όπου *MSTR* ονομάζεται μέσο άθροισμα τετραγώνων των θεραπειών και *MSE* μέσο άθροισμα των τετραγώνων των σφαλμάτων.

Κριτήρια αποδοχής/απόρριψης: Εάν η τιμή *F* που υπολογίζεται υπερβαίνει την κρίσιμη τιμή *F* από τους πίνακες κατανομής *F* για το επιθυμητό επίπεδο σημαντικότητας, τότε η μηδενική υπόθεση απορρίπτεται.

Επιπλέον, τα διαστήματα εμπιστοσύνης χρησιμεύουν ως κρίσιμα εργαλεία στην επαγωγική στατιστική συμπερασματολογία, επιτρέποντας την εκτίμηση του εύρους εντός του οποίου είναι πιθανό να βρίσκονται οι παράμετροι του πληθυσμού με βάση τα δεδομένα του δείγματος (Sahu et al. (2015)). Καθιερώνοντας διαστήματα εμπιστοσύνης γύρω από τις εκτιμήσεις δειγμάτων, οι ερευνητές μπορούν να εξακριβώσουν την ακρίβεια και την αξιοπιστία των ευρημάτων τους, ενισχύοντας έτσι την κατανόηση των πιθανών παραλλαγών και διακυμάνσεων στον ευρύτερο πληθυσμό χρηστών (Asadoorian & Kantarelis (2005)). Αυτό βοηθά στη διαπίστωση του βαθμού βεβαιότητας που σχετίζεται με τα ευρήματα, επιτρέποντας την τεκμηριωμένη λήψη αποφάσεων και την επινόηση στενευμένων παρεμβάσεων που στοχεύουν στη βελτιστοποίηση της αφοσίωσης και της ικανοποίησης των χρηστών στη σφαίρα ενασχόλησης με τα διαδικτυακά αθλητικά.

Μέσω της εφαρμογής της επαγωγικής στατιστικής συμπερασματολογίας, η μελέτη επιδιώκει να ξεπεράσει την απλή περιγραφή δεδομένων και να επεκταθεί στην ουσιαστική ερμηνεία και πρόβλεψη. Αξιοποιώντας τις τεχνικές δειγματοληψίας, του ελέγχου υποθέσεων και των διαστημάτων εμπιστοσύνης, γίνεται προσπάθεια να αποκαλυφθεί η σύνθετη αλληλεπίδραση μεταβλητών και παραγόντων που επηρεάζουν τις προτιμήσεις αθλητικών δραστηριοτήτων στον ψηφιακό τομέα. Εξάγοντας αξιόπιστα και ουσιαστικά συμπεράσματα για τον ευρύτερο πληθυσμό χρηστών, τίθενται οι βάσεις για την εφαρμογή στοχευμένων στρατηγικών και παρεμβάσεων που στοχεύουν στη βελτίωση των εμπειριών των χρηστών και στην προώθηση ενός πιο ελκυστικού περιβάλλοντος αθλητικής ενασχόλησης στη διαδικτυακή πλατφόρμα.

### 1.3 Ανάλυση συσχέτισης

Εξετάζοντας συστηματικά τον βαθμό συσχέτισης μεταξύ διαφόρων μεταβλητών, η ανάλυση συσχέτισης δίνει τη δυνατότητα στους ερευνητές να διακρίνουν τις αλληλεξαρτήσεις και τις συσχετίσεις που στηρίζουν το περίπλοκο πλαίσιο των προτιμήσεων των χρηστών και τις κλίσεις συμπεριφοράς μέσα στο τοπίο της ενασχόλησης με τα ψηφιακά αθλήματα.

Θεμελιώδης στην ανάλυση συσχέτισης είναι η έννοια της συσχέτισης, η οποία ποσοτικοποιεί τον βαθμό στον οποίο οι αλλαγές σε μια μεταβλητή αντιστοιχούν σε αλλαγές σε μια άλλη μεταβλητή (Gogtay & Thatte (2017)). Μετρώντας το είδος και την ισχύ αυτής της σχέσης, μπορούν να διακριθούν εάν οι μεταβλητές παρουσιάζουν θετική συσχέτιση, υποδεικνύοντας ότι μια αύξηση σε μια μεταβλητή αντιστοιχεί σε αύξηση στην άλλη ή αρνητική συσχέτιση, που σημαίνει ότι μια αύξηση σε μια μεταβλητή σχετίζεται με μείωση στην άλλη (Charman et al. (2013)). Επιπλέον, η απουσία σημαντικής συσχέτισης συνεπάγεται ότι οι μεταβλητές είναι ανεξάρτητες η μία από την άλλη, υπογραμμίζοντας έτσι την απουσία οποιασδήποτε συστηματικής σχέσης μεταξύ των υπό εξέταση μεταβλητών (Charman et al. (2013)).

Κεντρική θέση στην ανάλυση συσχέτισης κατέχει ο συντελεστής συσχέτισης Pearson, μια στατιστική μέτρηση που μετράει τη γραμμική σχέση μεταξύ δύο συνεχών μεταβλητών. Κυμαίνεται μεταξύ -1 και 1, ο συντελεστής συσχέτισης Pearson παρέχει ένα μετρήσιμο μέτρο της ισχύος και της κατεύθυνσης της γραμμικής σχέσης μεταξύ των μεταβλητών. Ένας συντελεστής κοντά στο 1 ή -1 υποδηλώνει μια ισχυρή θετική ή αρνητική συσχέτιση, αντίστοιχα, υποδηλώνοντας μια ισχυρή και συνεπή συσχέτιση μεταξύ των μεταβλητών (Senthilnathan (2019)). Αντίθετα, ένας συντελεστής κοντά στο 0 υποδηλώνει ασθενή ή αμελητέα συσχέτιση, υπογραμμίζοντας την απουσία οποιασδήποτε σημαντικής σχέσης μεταξύ των μεταβλητών (Gogtay & Thatte (2017)).

Επιπλέον, η ερμηνεία της ανάλυσης συσχέτισης εκτείνεται πέρα από τις απλές αριθμητικές τιμές για να συμπεριλάβει την οπτική αναπαράσταση των σχέσεων μέσω διαγραμμάτων διασποράς. Σχεδιάζοντας τα δεδομένα για τις δύο μεταβλητές σε ένα γράφημα, οι ερευνητές μπορούν να διακρίνουν οπτικά το μοτίβο και την κατεύθυνση της σχέσης, διευκολύνοντας έτσι μια πιο διαισθητική κατανόηση των υποκείμενων συσχετισμών (Senthilnathan (2019)). Αυτό βοηθά στον εντοπισμό πιθανών τάσεων, ομάδων ή μοτίβων μέσα στο σύνολο δεδομένων, επιτρέποντας τη διάκριση των

παραγόντων που επηρεάζουν σημαντικά τις προτιμήσεις και τις επιλογές αθλητικών δραστηριοτήτων εντός της κοινότητας της διαδικτυακής πλατφόρμας.

Μέσω της ανάλυσης της συσχέτισης, αναδεικνύονται οι αλληλεξαρτήσεις και οι σχέσεις που στηρίζουν τις προτιμήσεις των χρηστών και τα πρότυπα συμπεριφοράς στη σφαίρα εμπλοκής των ψηφιακών αθλημάτων. Με τη συστηματική αξιολόγηση της κατεύθυνσης και της ισχύος των σχέσεων μεταξύ διαφορετικών μεταβλητών, αυτή η παρέχονται πολύτιμες γνώσεις σχετικά με τους παράγοντες που καθοδηγούν τις επιλογές και τις προτιμήσεις των χρηστών, διευκολύνοντας έτσι την ανάπτυξη στοχευμένων στρατηγικών και παρεμβάσεων που στοχεύουν στη βελτιστοποίηση της αφοσίωσης και της ικανοποίησης των χρηστών στα διαδικτυακά αθλήματα.

Εκτός από τον συντελεστή συσχέτισης Pearson, ένα άλλο σημαντικό εργαλείο στην ανάλυση συσχέτισης είναι ο συντελεστής Spearman. Ο συντελεστής Spearman είναι ένα μη παραμετρικό μέτρο συσχέτισης που χρησιμοποιείται για να ποσοτικοποιήσει την ισχύ και τη μορφή της σχέσης μεταξύ δύο μεταβλητών. Σε αντίθεση με τον Pearson, που μετρά τη γραμμική σχέση, ο Spearman είναι ιδανικός για την ανάλυση των σχέσεων που δεν ακολουθούν απαραίτητα μια γραμμική μορφή, καθώς και για δεδομένα που είναι ιεραρχικά (Conover (1999)).

Η χρήση του συντελεστή Spearman προτιμάται όταν τα δεδομένα περιέχουν πολλές δεσμευμένες τιμές ή όταν η κατανομή των δεδομένων δεν είναι κανονική. Αυτό τον καθιστά ευέλικτο εργαλείο για την ανάλυση σε ποικίλες περιστάσεις όπου οι παραδοχές του Pearson δεν ικανοποιούνται. Ο συντελεστής Spearman υπολογίζεται με βάση τις διαφορές των βαθμίδων των τιμών των δύο μεταβλητών, παρέχοντας μια αντικειμενική εκτίμηση της σχέσης μεταξύ τους (Hauke & Kossowski (2011)).

Ο συντελεστής συσχέτισης Spearman, συχνά συμβολιζόμενος ως  $\rho$  (rho), είναι ένας μη παραμετρικός δείκτης που χρησιμοποιείται για να καθορίσει την ισχύ και την κατεύθυνση της σχέσης μεταξύ δύο μεταβλητών (Conover (1999)). Είναι ιδιαίτερα χρήσιμος για δεδομένα που δεν είναι κανονικά κατανομημένα ή για την αξιολόγηση μονοτονικών σχέσεων που δεν είναι απαραίτητα γραμμικές.

#### Υπολογισμός του Spearman

Ο συντελεστής Spearman υπολογίζεται ως εξής:

- ❖ Αρχικά, αποδίδονται βαθμίδες στις τιμές κάθε μεταβλητής.
- ❖ Στη συνέχεια, υπολογίζεται η διαφορά βαθμίδων  $d$  για κάθε ζεύγος τιμών.
- ❖ Τέλος, χρησιμοποιείται η εξής εξίσωση για τον υπολογισμό του  $\rho$ :

$$\rho = 1 - \frac{6 \sum_{i=1}^n d^2}{n(n^2-1)}, \quad (1.12)$$

όπου:

- $d$  είναι η διαφορά των βαθμίδων των δύο μεταβλητών για κάθε ζεύγος παρατηρήσεων,
- $n$  είναι ο αριθμός των ζευγών παρατηρήσεων.

Κριτήρια Αξιολόγησης: Ο συντελεστής Spearman κυμαίνεται μεταξύ -1 και 1. Μια τιμή κοντά στο 1 υποδηλώνει μια ισχυρή θετική μονοτονική σχέση, ενώ μια τιμή κοντά στο -1 υποδηλώνει μια ισχυρή αρνητική μονοτονική σχέση. Μια τιμή κοντά στο 0 υποδηλώνει την έλλειψη μιας μονοτονικής σχέσης. Για την ερμηνεία της στατιστικής σημαντικότητας του συντελεστή Spearman, χρησιμοποιείται ένα τεστ σημαντικότητας (π.χ., z-test μετατροπής του Spearman σε έναν κανονικά κατανομημένο δείκτη) για να καθοριστεί εάν η παρατηρούμενη τιμή του  $\rho$  είναι στατιστικά σημαντική.

#### 1.4 Ανάλυση παλινδρόμησης

Εστιάζοντας στην εξερεύνηση των αιτιακών σχέσεων και επιτρέποντας την εκτίμηση του μεγέθους και της κατεύθυνσης των επιδράσεων, η ανάλυση παλινδρόμησης διαδραματίζει καθοριστικό ρόλο στην αποκάλυψη της αλληλεπίδρασης παραγόντων που επηρεάζουν τις προτιμήσεις των χρηστών και τις κλίσεις συμπεριφοράς στο διαδικτυακό τοπίο της ενασχόλησης με τον αθλητισμό.

Καίριο ρόλο έχει η έννοια του μοντέλου παλινδρόμησης, το οποίο επιτρέπει στους ερευνητές να εξακριβώσουν τη λειτουργική σχέση μεταξύ της εξαρτημένης μεταβλητής και μιας ή περισσότερων ανεξάρτητων μεταβλητών (Freund et al. (2006)). Διατυπώνοντας και προσαρμόζοντας μοντέλα παλινδρόμησης, όπως η απλή γραμμική παλινδρόμηση, η πολλαπλή γραμμική παλινδρόμηση ή η λογιστική παλινδρόμηση, μπορεί να διακριθεί ο βαθμός στον οποίο οι αλλαγές στις ανεξάρτητες μεταβλητές αντιστοιχούν σε αλλαγές στην εξαρτημένη μεταβλητή, διευκολύνοντας να γίνουν κατανοητά οι υποκείμενες τάσεις και τα μοτίβα που διέπουν τις προτιμήσεις αθλητικών δραστηριοτήτων στην κοινότητα της διαδικτυακής πλατφόρμας.



Η εφαρμογή της απλής γραμμικής παλινδρόμησης επιτρέπει την εξέταση της σχέσης μεταξύ μιας ανεξάρτητης μεταβλητής και μιας εξαρτημένης μεταβλητής, διευκολύνοντας την εκτίμηση της επίδρασης της ανεξάρτητης μεταβλητής στην εξαρτημένη μεταβλητή (Chatterjee & Hadi (2013)). Οριοθετώντας την κλίση και την τομή της ευθείας της παλινδρόμησης, διακρίνεται η κατεύθυνση και το μέγεθος της σχέσης, διευκολύνοντας την κατανόηση του βαθμού στον οποίο οι αλλαγές στην ανεξάρτητη μεταβλητή αντιστοιχούν σε αλλαγές στην εξαρτημένη μεταβλητή (Sarstedt et al. (2019)).

Αντίθετα, η πολλαπλή γραμμική παλινδρόμηση επιτρέπει τη διερεύνηση της συνδυασμένης επίδρασης πολλαπλών ανεξάρτητων μεταβλητών στην εξαρτημένη μεταβλητή, επιτρέποντας τη διάκριση της συλλογικής επιρροής διάφορων παραγόντων στις προτιμήσεις των χρηστών και στις συμπεριφορικές κλίσεις στο διαδικτυακό τοπίο της ενασχόλησης με τον αθλητισμό. Εκτιμώντας τους συντελεστές παλινδρόμησης για κάθε ανεξάρτητη μεταβλητή, προσδιορίζεται η σχετική συμβολή κάθε παράγοντα στη μεταβλητότητα της εξαρτημένης μεταβλητής, διευκολύνοντας τον προσδιορισμό των βασικών καθοριστικών παραγόντων που επηρεάζουν σημαντικά τις επιλογές αθλητικών δραστηριοτήτων στον ψηφιακό τομέα.

Επιπλέον, η λογιστική παλινδρόμηση χρησιμεύει ως εργαλείο για την ανάλυση της επίδρασης ανεξάρτητων μεταβλητών σε μια κατηγορική εξαρτημένη μεταβλητή, ώστε να γίνεται πρόβλεψη της πιθανότητας εμφάνισης ενός συγκεκριμένου γεγονότος ή αποτελέσματος (Montgomery et al. (2021)). Διακρίνοντας τη σχέση μεταξύ των ανεξάρτητων μεταβλητών και της πιθανότητας μιας συγκεκριμένης επιλογής αθλητικής δραστηριότητας, η λογιστική παλινδρόμηση επιτρέπει τον εντοπισμό των βασικών καθοριστικών παραγόντων που επηρεάζουν τις προτιμήσεις και τις επιλογές των χρηστών.

### 1.5 Μέθοδοι Γραφικής Απεικόνισης

Μεταφράζοντας σύνθετα σύνολα δεδομένων σε διαισθητικές και οπτικά ελκυστικές αναπαραστάσεις, οι μέθοδοι γραφικής απεικόνισης δεδομένων διαδραματίζουν καθοριστικό ρόλο στη διευκόλυνση της ολοκληρωμένης ερμηνείας και κατανόησης των προτιμήσεων των χρηστών, της δυναμικής συμπεριφοράς και των υποκείμενων τάσεων στο οικοσύστημα εμπλοκής ψηφιακών αθλημάτων.

Οι μέθοδοι γραφικής απεικόνισης των δεδομένων είναι η χρήση διαγραμμάτων διασποράς, τα οποία επιτρέπουν τη γραφική αναπαράσταση της σχέσης μεταξύ δύο συνεχών μεταβλητών (Post et al. (2002)). Σχεδιάζοντας τα δεδομένα σε ένα δισδιάστατο γράφημα, διακρίνεται οπτικά το μοτίβο, η κατεύθυνση και η δύναμη της σχέσης μεταξύ των μεταβλητών, διευκολύνοντας έτσι μια λεπτή κατανόηση των υποκείμενων συσχετίσεων (Brodie et al. (2012)). Τα διαγράμματα διασποράς χρησιμεύουν ως απαραίτητα εργαλεία για τον εντοπισμό πιθανών τάσεων, συστάδων ή μοτίβων εντός του συνόλου δεδομένων, φανερώνοντας τους παράγοντες που επηρεάζουν σημαντικά τις προτιμήσεις και τις επιλογές αθλητικών δραστηριοτήτων στην κοινότητα της διαδικτυακής πλατφόρμας.

Επιπλέον, τα γραφήματα ράβδων χρησιμεύουν ως εργαλεία για την οπτική απεικόνιση κατηγορικών δεδομένων, επιτρέποντας έτσι τη συγκριτική ανάλυση διαφορετικών κατηγοριών ή ομάδων εντός του συνόλου δεδομένων (Zhao et al. (2020)). Με την αναπαράσταση δεδομένων χρησιμοποιώντας ορθογώνιες ράβδους με μήκη ανάλογα με τις τιμές που αντιπροσωπεύουν, τα γραφήματα ράβδων παρέχουν μια συνοπτική επισκόπηση της σχετικής επικράτησης ή δημοτικότητας διαφόρων αθλητικών δραστηριοτήτων στην κοινότητα της διαδικτυακής πλατφόρμας. Τα διαγράμματα ράβδων χρησιμεύουν ως αποτελεσματικά εργαλεία για τη διάκριση των πιο αγαπημένων ή συχνά επιλεγμένων αθλητικών δραστηριοτήτων, διευκολύνοντας έτσι την πληρέστερη κατανόηση των προτιμήσεων και των κλίσεων της βάσης των χρηστών.

Εκτός από τα διαγράμματα διασποράς και τα γραφήματα ράβδων, οι χάρτες θερμότητας χρησιμεύουν ως ισχυρά εργαλεία γραφικής απεικόνισης δεδομένων για την απεικόνιση πολύπλοκων μοτίβων και κατανομών των δεδομένων (Ajibade & Adediran (2016)). Χρησιμοποιώντας χρωματικές διαβαθμίσεις για την αναπαράσταση του μεγέθους των τιμών σε μια δισδιάστατη απεικόνιση, οι χάρτες θερμότητας επιτρέπουν στους ερευνητές να διακρίνουν χωρικές και χρονικές τάσεις εντός του συνόλου δεδομένων, ώστε να γίνεται ευκολότερη η κατανόηση της γεωγραφικής κατανομής ή των χρονικών διακυμάνσεων στις προτιμήσεις και επιλογές αθλητικών δραστηριοτήτων. Οι χάρτες θερμότητας χρησιμεύουν ως όργανα για τον εντοπισμό hotspot, τάσεων ή συμπλεγμάτων εντός του συνόλου δεδομένων, επιτρέποντας να διακριθούν οι παράγοντες που επηρεάζουν σημαντικά την αφοσίωση και τις προτιμήσεις των χρηστών σε συγκεκριμένες περιοχές ή χρονικά πλαίσια.

Το θηκόγραμμα επίσης είναι μια μέθοδος γραφικής απεικόνισης που παρουσιάζει την κατανομή των δεδομένων μέσω ορισμένων περιγραφικών στατιστικών (ελάχιστη τιμή, πρώτο τεταρτημόριο, διάμεσος, τρίτο τεταρτημόριο και μέγιστη τιμή), καθώς και "μουστάκια" που εκτείνονται για να δείξουν την παρουσία ακραίων τιμών (Tukey (1977)). Τα θηκογράμματα είναι ιδιαίτερα χρήσιμα για την απεικόνιση της διακύμανσης, της συμμετρίας της κατανομής και την παρουσία ακραίων τιμών, παρέχοντας μια συνοπτική εικόνα της κατανομής των δεδομένων.

Το ιστόγραμμα συχνοτήτων, από την άλλη, είναι μια γραφική απεικόνιση που χρησιμοποιείται για να δείξει την κατανομή των δεδομένων σε διακριτές ομάδες ή διαστήματα, γνωστές ως "κλάσεις" (Scott (1979)). Κάθε ράβδος στο ιστόγραμμα αντιπροσωπεύει τη συχνότητα (ή τον αριθμό των παρατηρήσεων) εντός κάθε κλάσης, παρέχοντας μια εικόνα της κατανομής των τιμών. Τα ιστογράμματα συχνοτήτων είναι ιδιαίτερα χρήσιμα για την κατανόηση της κατανομής και της κεντρικής τάσης των δεδομένων, καθώς και για τον εντοπισμό πιθανών συστάδων ή ακραίων τιμών.

Τέλος, η εισαγωγή του τριδιάστατου διαγράμματος διασποράς προσθέτει μια επιπλέον διάσταση στην ανάλυση, επιτρέποντας την απεικόνιση των σχέσεων μεταξύ τριών συνεχών μεταβλητών σε έναν τρισδιάστατο χώρο (Cleveland & McGill (1988)). Αυτή η τεχνική επιτρέπει στους ερευνητές να εξετάζουν πολύπλοκες αλληλεπιδράσεις και αλληλεξαρτήσεις μεταξύ μεταβλητών, συνεισφέροντας στην κατανόηση των υποκείμενων μοτίβων και των δυναμικών εντός του συνόλου των δεδομένων.

## 1.6 Δεοντολογικά ζητήματα στην ανάλυση δεδομένων

Δεδομένης της ευαίσθητης φύσης των πληροφοριών χρήστη και των πιθανών επιπτώσεων της ανάλυσης δεδομένων στο απόρρητο και την εμπιστευτικότητα των χρηστών, οι ηθικοί παράγοντες που αφορούν την ανάλυση δεδομένων υπογραμμίζουν την ανάγκη τήρησης των υψηλότερων προτύπων ακεραιότητας, διαφάνειας και προστασίας των χρηστών (Arifin (2018)).

Θεμελιώδης για την ηθική ανάλυση δεδομένων είναι η επιτακτική ανάγκη διασφάλισης του απόρρητου και της ασφάλειας των δεδομένων. Ο σεβασμός της εμπιστευτικότητας και της ανωνυμίας των πληροφοριών των χρηστών είναι καίριας

σημασίας για τη διατήρηση της ατομικής εμπιστοσύνης και της εμπιστοσύνης της κοινότητας της διαδικτυακής πλατφόρμας. Η εφαρμογή ισχυρών πρωτοκόλλων κρυπτογράφησης δεδομένων, ασφαλών συστημάτων αποθήκευσης δεδομένων και αυστηρών ελέγχων πρόσβασης είναι σημαντική ανάγκη για την προστασία των πληροφοριών χρήστη από μη εξουσιοδοτημένη πρόσβαση ή παραβιάσεις, διασφαλίζοντας έτσι την προστασία των ευαίσθητων δεδομένων χρήστη από πιθανούς κινδύνους ή απειλές (Saltz & Dewar (2019)).

Επιπλέον, η αρχή της ενημερωμένης συναίνεσης έχει κρίσιμη σημασία για την ηθική διεξαγωγή της ανάλυσης δεδομένων. Η προτεραιότητα της συναίνεσης των χρηστών και της διαφάνειας στις διαδικασίες συλλογής δεδομένων ενισχύει μια κουλτούρα εμπιστοσύνης και αμοιβαίου σεβασμού μεταξύ των ερευνητών και της κοινότητας των χρηστών (Walker (2007)). Η σαφής διατύπωση του σκοπού, του πεδίου και των πιθανών επιπτώσεων της ανάλυσης δεδομένων στους χρήστες τους δίνει τη δυνατότητα να λαμβάνουν τεκμηριωμένες αποφάσεις σχετικά με τη χρήση και τη διάδοση των δεδομένων τους, διασφαλίζοντας έτσι την εθελοντική και ρητή συμμετοχή των χρηστών στην ερευνητική προσπάθεια (Arifin (2018)).

Επιπλέον, ο υπεύθυνος χειρισμός ευαίσθητων πληροφοριών απαιτεί την εφαρμογή τεχνικών «ανωνυμοποίησης» δεδομένων. Αφαιρώντας ή κρυπτογραφώντας προσωπικά αναγνωρίσιμες πληροφορίες από το σύνολο δεδομένων, οι ερευνητές μπορούν να διατηρήσουν την ανωνυμία και την εμπιστευτικότητα των πληροφοριών των χρηστών, μετριάζοντας έτσι τους κινδύνους παραβίασης δεδομένων ή μη εξουσιοδοτημένης πρόσβασης (Cacciattolo (2015)). Η τήρηση των αρχών της ελαχιστοποίησης των δεδομένων και της ανωνυμοποίησης δεδομένων είναι πολύ σημαντική για τη διατήρηση του απορρήτου και της εμπιστευτικότητας των δεδομένων των χρηστών, ενισχύοντας έτσι μια κουλτούρα εμπιστοσύνης και διαφάνειας στην διαδικτυακή κοινότητα αθλητικής δέσμευσης.

Πέρα από το απόρρητο και την ασφάλεια των δεδομένων, η ηθική ανάλυση δεδομένων συνεπάγεται την υπεύθυνη διάδοση των ερευνητικών ευρημάτων. Η διασφάλιση της διαφάνειας και της ακρίβειας των αποτελεσμάτων της έρευνας, η αποκάλυψη πιθανών προκαταλήψεων ή περιορισμών και η αναγνώριση της συνεισφοράς άλλων ερευνητών καλλιεργεί μια κουλτούρα ακαδημαϊκής ακεραιότητας και αμοιβαίου σεβασμού εντός της ερευνητικής κοινότητας (Cacciattolo (2015)). Η

τήρηση των αρχών της πνευματικής ειλικρίνειας και της διαφάνειας στην ανάλυση δεδομένων είναι ζωτικής σημασίας για την προώθηση ενός ευνοϊκού περιβάλλοντος για ανοιχτό διάλογο, κριτική έρευνα και συλλογική διάδοση γνώσης στον τομέα της διαχείρισης αθλημάτων και της ψηφιακής δέσμευσης.

Μέσω της ενσωμάτωσης αυτών των ηθικών θεωρήσεων, αυτή η έρευνα προσπαθεί να διατηρήσει τα υψηλότερα πρότυπα ακεραιότητας, διαφάνειας και προστασίας των χρηστών στην ανάλυση των προτιμήσεων αθλητικών δραστηριοτήτων στο οικοσύστημα της διαδικτυακής πλατφόρμας. Δίνοντας προτεραιότητα στο απόρρητο και την ασφάλεια των δεδομένων, την ενημερωμένη συγκατάθεση, την ανωνυμοποίηση δεδομένων και την υπεύθυνη διάδοση των ευρημάτων της έρευνας, αυτή η μελέτη στοχεύει να καλλιεργήσει μια κουλτούρα εμπιστοσύνης, σεβασμού και ηθικής συμπεριφοράς στο τοπίο της ενασχόλησης με τα ψηφιακά αθλήματα, θέτοντας έτσι το προηγούμενο για τους υπεύθυνους και ευσυνείδητη ενσωμάτωση της ανάλυσης δεδομένων στην επιδίωξη τεκμηριωμένης λήψης αποφάσεων και παρεμβάσεων με επίκεντρο τον χρήστη.



# ΚΕΦΑΛΑΙΟ 2

## Μεθοδολογία μελέτης

Στην παρούσα μελέτη θα χρησιμοποιηθούν τα εξής εργαλεία ανάλυσης:

### ❖ Γραμμική παλινδρόμηση

Μεθοδολογία: Χρησιμοποιείται για την κατανόηση της σχέσης μεταξύ μιας εξαρτημένης μεταβλητής και μιας ή περισσότερων ανεξάρτητων μεταβλητών.

Προϋποθέσεις Εφαρμογής: Η εξαρτημένη μεταβλητή πρέπει να είναι συνεχής και οι ανεξάρτητες μεταβλητές μπορεί να είναι συνεχείς ή κατηγορικές (κωδικοποιημένες).

Υποθέσεις: Ελέγχει εάν οι ανεξάρτητες μεταβλητές προβλέπουν σημαντικά την εξαρτημένη μεταβλητή.

$H_0$ : Δεν υπάρχει σχέση μεταξύ των ανεξάρτητων μεταβλητών και της εξαρτημένης μεταβλητής, δηλαδή  $\beta_1 = \beta_2 = \dots = \beta_n = 0$  (εξαιρείται ο συντελεστής  $\beta_0$ ).

$H_1$ : Υπάρχει τουλάχιστον μία ανεξάρτητη μεταβλητή που προβλέπει στατιστικά σημαντικά την εξαρτημένη μεταβλητή.

Γραμμικό μοντέλο πολλαπλής παλινδρόμησης:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon,$$

όπου  $Y$  είναι η εξαρτημένη μεταβλητή,  $X_i$  είναι ανεξάρτητες μεταβλητές,  $\beta_i$  είναι οι συντελεστές και  $\varepsilon$  είναι ο όρος σφάλματος.

Κριτήρια απόφασης: Σημαντικότητα των συντελεστών παλινδρόμησης (p-values), τιμή για την προσαρμογή του μοντέλου  $R^2$ .

### ❖ Ανάλυση Διακύμανσης (ANOVA)

Μεθοδολογία: Συγκρίνει τους μέσους σε τρεις ή περισσότερες ομάδες για να δει εάν τουλάχιστον μία μέση τιμή κάποιας ομάδας διαφέρει στατιστικά σημαντικά.

Προϋποθέσεις Εφαρμογής: Η εξαρτημένη μεταβλητή πρέπει να είναι συνεχής και η ανεξάρτητη μεταβλητή πρέπει να είναι κατηγορική με τρεις ή περισσότερες ομάδες.

Υποθέσεις: Η μηδενική υπόθεση δηλώνει ότι όλοι οι μέσοι των ομάδων μιας μεταβλητής είναι ίσοι, ενώ η εναλλακτική υπόθεση δηλώνει ότι τουλάχιστον ένας είναι διαφορετικός.

$$H_0: \mu_1 = \mu_2 = \dots = \mu_v$$

$H_1$ : Τουλάχιστον ένας μέσος μιας ομάδας διαφέρει στατιστικά σημαντικά από τους άλλους.

Κανόνας απόφασης: Η τιμή  $F$  και η τιμή  $p$  στον πίνακα ANOVA καθορίζουν εάν οι διαφορές μεταξύ των μέσων είναι στατιστικά σημαντικές.

❖ T-test

Μεθοδολογία: Ελέγχει εάν υπάρχει στατιστικά σημαντική διαφορά μεταξύ των μέσων δύο ομάδων.

Προϋποθέσεις Εφαρμογής: Μία συνεχής εξαρτημένη μεταβλητή και μία δυαδική ανεξάρτητη μεταβλητή.

Υποθέσεις: Η μηδενική υπόθεση δηλώνει ότι δεν υπάρχει διαφορά μεταξύ των μέσων των δύο ομάδων.

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Κανόνας απόφασης: Η τιμή  $t$  και  $p$ -value καθορίζουν εάν η διαφορά μεταξύ των μέσων της ομάδας είναι στατιστικά σημαντική.

❖ Συντελεστής Συσχέτισης Spearman

Μεθοδολογία: Μετρά την ισχύ και τη μορφή της συσχέτισης μεταξύ δύο ιεραρχικών μεταβλητών.

Προϋποθέσεις Εφαρμογής: Και οι δύο μεταβλητές πρέπει να είναι τουλάχιστον τακτικές.

Υποθέσεις: Ελέγχει εάν υπάρχει στατιστικά σημαντική  $\rho$  μεταξύ δύο μεταβλητών.



$H_0$ : Δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών  $\rho=0$ .

$H_1$ : Υπάρχει στατιστικά σημαντική συσχέτιση μεταξύ των μεταβλητών  $\rho \neq 0$ .

Κανόνας απόφασης: Ο συντελεστής συσχέτισης κατάταξης του Spearman ( $\rho$ ) και  $p$ -*value*. Το θετικό  $\rho$  υποδηλώνει θετική συσχέτιση και το αρνητικό  $\rho$  υποδηλώνει αρνητική συσχέτιση.



## Κεφάλαιο 3: Ερευνητικό Εργαλείο

Στις σύγχρονες αθλητικές πλατφόρμες, η ενσωμάτωση του Enterprise Resource Planning (ERP) και των συστημάτων ελέγχου πρόσβασης έχει φέρει επανάσταση στη συλλογή λεπτομερών και με ακρίβεια δεδομένων σχετικά με τη συμπεριφορά των χρηστών, ιδιαίτερα σε θέματα αφοσίωσης και συνέπειας. Αυτά τα προηγμένα συστήματα επιτρέπουν τη συλλογή περιεκτικών συνόλων δεδομένων που περιλαμβάνουν ένα ευρύ φάσμα δημογραφικών χαρακτηριστικών και ιδιοτήτων που σχετίζονται με υπηρεσίες, τα οποία μπορούν να αναλυθούν για να κατανοηθούν οι παράγοντες που επηρεάζουν την αφοσίωση και τη συνέπεια των χρηστών. Συγκεκριμένα χρησιμοποιήθηκε το εργαλείο συλλογής των Pinheiro & Cavique (2021).

Το σύνολο δεδομένων που λαμβάνεται από αυτά τα συστήματα περιέχει διάφορα χαρακτηριστικά που περιγράφουν την αλληλεπίδραση του χρήστη με τις αθλητικές πλατφόρμες, το δημογραφικό του προφίλ και τη δέσμευσή του με τις προσφερόμενες υπηρεσίες. Ακολουθεί μια σύντομη επισκόπηση ορισμένων βασικών χαρακτηριστικών που περιλαμβάνονται στο σύνολο δεδομένων.

- ✓ **Δημογραφικά χαρακτηριστικά:** Περιλαμβάνουν βασικές πληροφορίες όπως το αναγνωριστικό χρήστη (μοναδικό αναγνωριστικό), την ηλικία, την ηλικιακή ομάδα και το φύλο. Αυτά τα χαρακτηριστικά βοηθούν στην κατανόηση της δημογραφικής κατανομής των χρηστών της πλατφόρμας και στην τμηματοποίηση των δεδομένων για πιο στοχευμένη ανάλυση.
- ✓ **Χαρακτηριστικά εγγραφής:** Χαρακτηριστικά όπως οι ημερομηνίες έναρξης εγγραφής και τέλος εγγραφής παρέχουν πληροφορίες σχετικά με τη διάρκεια της αφοσίωσης ενός χρήστη με την πλατφόρμα. Ο χρήστης θεωρείται ενεργός από τη στιγμή της πρώτης εγγραφής του έως ότου είτε εκφράσει την επιθυμία να διακόψει είτε χαρακτηριστεί ως διαγραμμένος λόγω μη πληρωμής, σύμφωνα με τους κανονισμούς της επιχείρησης.
- ✓ **Χαρακτηριστικά δέσμευσης δραστηριότητας:** Αυτά περιλαμβάνουν τον αριθμό των δραστηριοτήτων στις οποίες είναι εγγεγραμμένος ένας χρήστης, τους συγκεκριμένους τύπους δραστηριοτήτων (π.χ. στίβος, δραστηριότητες στο νερό, δραστηριότητες φυσικής κατάστασης, κ.λπ.) και τον αριθμό των

δραστηριοτήτων που παρακολουθεί. Τέτοια χαρακτηριστικά είναι πολύ σημαντικά για την ανάλυση των προτιμήσεων των χρηστών και του βαθμού της ενεργής συμμετοχής τους.

- ✓ Χαρακτηριστικά συχνότητας άθλησης: Χαρακτηριστικά όπως η ημερομηνία τελευταίας συμμετοχής, συχνότητα και μέρες εκτός βοηθούν στην κατανόηση της συχνότητας των επισκέψεων ενός χρήστη στις αθλητικές πλατφόρμα. Αυτοί είναι σημαντικοί δείκτες της αφοσίωσης των χρηστών και των πιθανών κινδύνων διακοπής.
- ✓ Οικονομικά χαρακτηριστικά: Χαρακτηριστικό όπως το συνολικό ποσό που έχει ξοδέψει ένας χρήστης κατά την περίοδο εγγραφής του, προσφέροντας πληροφορίες για την οικονομική αξία που σχετίζεται με κάθε χρήστη. Αυτό, σε συνδυασμό με τον αριθμό των ανανεώσεων και τους τύπους των υπηρεσιών που έχουν εγγραφεί, παρέχει μια ολοκληρωμένη εικόνα της οικονομικής δέσμευσης του χρήστη.
- ✓ Δείκτης εγκατάλειψης: Το χαρακτηριστικό διακοπής υποδεικνύει εάν ένας χρήστης θεωρείται ότι εγκαταλείπεται τη στιγμή της εξαγωγής δεδομένων. Αυτό το χαρακτηριστικό είναι απαραίτητο για την ανάλυση των ποσοστών αφοσίωσης και την κατανόηση των παραγόντων που συμβάλλουν στην διακοπή των χρηστών.

Μεταβλητή	Περιγραφή
Ηλικία	Η τρέχουσα ηλικία του ατόμου
Ηλικιακή Ομάδα	Κατηγοριοποίηση ηλικίας (π.χ. 20-29, 30-39 κλπ.)
Φύλο	Το φύλο του ατόμου (π.χ. άνδρας, γυναίκα κλπ.)
Έναρξη Εγγραφής	Ημερομηνία έναρξης της εγγραφής
Λήξη Εγγραφής	Ημερομηνία λήξης της εγγραφής
Διάρκεια Εγγραφής	Συνολική διάρκεια της εγγραφής
Κλάση Διάρκειας Εγγραφής 1	Κατηγορία βάσει διάρκειας εγγραφής (π.χ. βραχυπρόθεσμη, μακροπρόθεσμη)
Κλάση Διάρκειας Εγγραφής 2	Μια άλλη κατηγοριοποίηση για τη διάρκεια εγγραφής πιθανώς με διαφορετικά κριτήρια
Έναρξη Τελευταίας Περιόδου	Ημερομηνία έναρξης της τελευταίας ενεργής περιόδου

Λήξη Τελευταίας Περιόδου	Ημερομηνία λήξης της τελευταίας ενεργής περιόδου
Ημερομηνία Τελευταίας Επίσκεψης	Η ημερομηνία της τελευταίας επίσκεψης του χρήστη στην πλατφόρμα
Ημέρες Χωρίς Συχνότητα	Αριθμός ημερών χωρίς συμμετοχή σε δραστηριότητα
Κλάση Χωρίς Συχνότητα 1 Ημερών	Κατηγορία των ημερών χωρίς συχνότητα σύμφωνα με ένα συγκεκριμένο σύστημα κατηγοριοποίησης
Κλάση Χωρίς Συχνότητα 2 Ημερών	Άλλη κατηγορία των ημερών χωρίς συχνότητα με διαφορετικό σύστημα κατηγοριοποίησης
Αξία Ζωής	Η συνολική οικονομική αξία που έχει προσφέρει ένας χρήστης στην πλατφόρμα μέχρι στιγμής
Κλάση Αξίας Ζωής 1	Κατηγοριοποίηση της αξίας ζωής σε κλάσεις σύμφωνα με κάποιο σύστημα
Μέση Συχνότητα	Η μέση συχνότητα συμμετοχής του χρήστη στις δραστηριότητες
Κλάση Μέσης Συχνότητας 1	Κατηγοριοποίηση της μέσης συχνότητας σε κλάσεις
Κλάση Μέσης Συχνότητας 2	Δεύτερη κατηγοριοποίηση της μέσης συχνότητας σε κλάσεις
Χρήση Χρονική Περίοδο Ανά	Πόσο συχνά ο χρήστης χρησιμοποιεί την πλατφόρμα κατά τις διάφορες ώρες της ημέρας ή της νύχτας
Δραστηριότητες Στίβου	Συμμετοχή του χρήστη σε δραστηριότητες στίβου
Δραστηριότητες στο Νερό	Συμμετοχή του χρήστη σε δραστηριότητες στο νερό
Δραστηριότητες Φυσικής Κατάστασης	Συμμετοχή του χρήστη σε προγράμματα φυσικής κατάστασης
Δραστηριότητες Χορού	Συμμετοχή του χρήστη σε δραστηριότητες χορού
Ομαδικά Αθλήματα	Συμμετοχή του χρήστη σε ομαδικά αθλήματα
Αθλήματα Ρακέτας	Συμμετοχή του χρήστη σε αθλήματα ρακέτας
Πολεμικές Τέχνες	Συμμετοχή του χρήστη σε πολεμικές τέχνες

Δραστηριότητες στη Φύση	Συμμετοχή του χρήστη σε δραστηριότητες που πραγματοποιούνται στη φύση
Ειδικές Δραστηριότητες	Συμμετοχή του χρήστη σε ειδικές δραστηριότητες (μπορεί να περιλαμβάνει π.χ. θεραπευτική γυμναστική, ειδικά προγράμματα προπόνησης κτλ.)
Άλλες Δραστηριότητες	Συμμετοχή του χρήστη σε άλλες δραστηριότητες που δεν κατατάσσονται στις προηγούμενες κατηγορίες
Αριθμός Δραστηριοτήτων	Το συνολικό αριθμό δραστηριοτήτων στις οποίες συμμετέχει ένας χρήστης
Αριθμός Συχνοτήτων	Συχνότητα συμμετοχής του χρήστη στις δραστηριότητες
Αριθμός Συχνοτήτων Κλάσης 1	Κατηγοριοποίηση του χρήστη βάσει της συχνότητας των συμμετοχών του σε δραστηριότητες σε μια συγκεκριμένη κλάση
Παρακολουθούμενα Μαθήματα	Αριθμός των μαθημάτων ή των συνεδριών που έχει παρακολουθήσει ο χρήστης
Παρακολουθούμενα Μαθήματα Κλάσης 1	Κατηγοριοποίηση του χρήστη βάσει του αριθμού των μαθημάτων που έχει παρακολουθήσει σε μια συγκεκριμένη κλάση
Μέσος Όρος Παρακολουθημένων Μαθημάτων	Μέσος όρος παρακολουθημένων μαθημάτων ανά εγγεγραμμένο χρήστη
Κατηγορία Μέσου Όρου Παρακολουθημένων Μαθημάτων 1	Κατηγοριοποίηση του μέσου όρου παρακολουθημένων μαθημάτων σε κλάση 1
Επιτρεπόμενες Εβδομαδιαίες Επισκέψεις Σύμφωνα με SLA	Επιτρεπόμενες εβδομαδιαίες επισκέψεις βάσει του επιπέδου συμφωνημένης υπηρεσίας

Επιτρεπόμενος Αριθμός Επισκέψεων Σύμφωνα με SLA	Επιτρεπόμενος αριθμός επισκέψεων βάσει του επιπέδου συμφωνημένης υπηρεσίας
Πραγματικός Αριθμός Επισκέψεων	Πραγματικός αριθμός επισκέψεων που έγιναν από τον χρήστη
Λόγος	Λόγος των πραγματικών προς τις επιτρεπόμενες επισκέψεις
Κατηγορία Λόγου 2	Κατηγοριοποίηση του λόγου των πραγματικών προς τις επιτρεπόμενες επισκέψεις σε κλάση 2
Αριθμός Ανανεώσεις	Αριθμός ανανεώσεων εγγραφής του χρήστη
Έχει Συστάσεις	Εάν ο χρήστης έχει παραπομπές (συστάσεις) μέσα στο σύστημα
Αριθμός Συστάσεων	Αριθμός παραπομπών που έχει ο χρήστης
Αποχώρηση	Εάν ο χρήστης εγκατέλειψε το πρόγραμμα ή την πλατφόρμα

Πίνακας 1

Η συλλογή τέτοιων δεδομένων μέσω συστημάτων ERP και ελέγχου πρόσβασης επιτρέπει στις αθλητικές πλατφόρμες να κατανοήσουν τη συμπεριφορά των χρηστών, των προτιμήσεων και των προτύπων αφοσίωσης. Αναλύοντας αυτά τα χαρακτηριστικά, οι πλατφόρμες μπορούν να εντοπίσουν βασικές τάσεις στην αφοσίωση και τη διατήρηση, να προβλέψουν πιθανές διακοπές και να προσαρμόσουν τις υπηρεσίες τους ώστε να ανταποκρίνονται καλύτερα στις ανάγκες και τις προτιμήσεις των χρηστών τους.

Η χρήση συστημάτων ERP και ελέγχου πρόσβασης για τη συλλογή δεδομένων σε αθλητικές πλατφόρμες αντιπροσωπεύει σημαντική πρόοδο στην κατανόηση και τη βελτίωση της αφοσίωσης και της διατήρησης των χρηστών. Το σύνολο των δεδομένων που λαμβάνεται από αυτά τα συστήματα παρέχει τη βάση για τεκμηριωμένη λήψη αποφάσεων, επιτρέποντας στις πλατφόρμες να υιοθετήσουν μια προσέγγιση βάσει δεδομένων για τη βελτίωση της αφοσίωσης των χρηστών και τη μείωση των ποσοστών εγκατάλειψης.

## Κεφάλαιο 4: Αποτελέσματα

### 4.1 Περιγραφικά Στατιστικά

Ο Πίνακας 1 περιέχει τη σύνοψη των μεταβλητών που περιλαμβάνουν ημερομηνίες.

Μεταβλητές Ημερομηνιών	Ελάχιστη	Διάμεσος	Μέγιστη
Ημερομηνία πρώτης εγγραφής	2/6/2014	23/9/2016	21/10/2019
Ημερομηνία λήξης της τελευταίας εγγραφής	11/7/2014	5/6/2018	31/10/2019
Τελευταία περίοδος έναρξης	2/6/2014	7/4/2018	30/10/2019
Τελευταία περίοδος λήξης	11/7/2014	7/4/2018	31/10/2019
Ημερομηνία και ώρα της τελευταίας συμμετοχής του χρήστη	3/7/2014	18/1/2018	31/10/2019

Πίνακας 2: Σύνοψη των μεταβλητών που περιλαμβάνουν ημερομηνίες.

Η μέση ηλικία των ερευνώμενων είναι περίπου 26 έτη με τους μισούς να είναι κάτω των 23 ετών. Οι ηλικίες κυμαίνονται από 0 έως 87 ετών, υποδηλώνοντας μια ευρεία βάση χρηστών ή πιθανές ακραίες τιμές.

Η μέση διάρκεια εγγραφής να είναι σχεδόν 14 μήνες και τα δεδομένα να δείχνουν μια λοξή κατανομή προς τα δεξιά, υποδηλώνοντας ότι οι περισσότεροι χρήστες έχουν μικρότερες περιόδους εγγραφής.

Ο αριθμός των ημερών που ο χρήστης δεν επισκέφτηκε τις πλατφόρμα πριν θεωρηθεί λήξη δείχνει μια πολύ έντονη λοξότητα και κύρτωση, υποδεικνύοντας έναν σημαντικό αριθμό ακραίων τιμών και μια εξαιρετικά οξεία κατανομή. Οι μέγιστες ημέρες χωρίς επίσκεψη πριν θεωρηθεί εγκατάλειψη είναι 1745, το οποίο είναι εξαιρετικά υψηλός αριθμός.

Η κατανομή των συνολικών πληρωμών είναι ιδιαίτερα λοξή προς τα δεξιά με μέση τιμή 302,56, υποδεικνύοντας ότι ενώ οι περισσότεροι χρήστες ξοδεύουν λιγότερα, μερικοί ξοδεύουν σημαντικά περισσότερα, έως και 6728.

Η συχνότητα των επισκέψεων ανά εβδομάδα, με μέσο όρο 0,76, δείχνει μέτρια χρήση των πλατφορμών.



Οι περισσότεροι χρήστες είναι εγγεγραμμένοι σε μία δραστηριότητα, με μικρό εύρος (1 έως 5), αλλά η κατανομή είναι πολύ λοξή προς τα δεξιά, υποδεικνύοντας ότι λίγοι χρήστες συμμετέχουν σε πολλαπλές δραστηριότητες.

Με μέσο όρο 40,16 και υψηλό μέγιστο (1031), αυτό το χαρακτηριστικό εμφανίζει επίσης μια κατανομή με λοξή δεξιά ασυμμετρία, υποδεικνύοντας ότι οι περισσότεροι χρήστες επισκέπτονται την πλατφόρμα μέτρια, αλλά λίγοι την επισκέπτονται πολύ συχνά.

Στον αριθμό των εβδομαδιαίων επισκέψεων στους τελευταίους 2 μήνες της εγγραφής του η αρνητική λοξότητα (-1,222) είναι ασυνήθιστη και υποδηλώνει ότι η κατανομή είναι λοξή προς τα αριστερά, που σημαίνει ότι περισσότεροι χρήστες πραγματοποίησαν συμμετοχές πιο κοντά στο μέγιστο επιτρεπόμενο, με τη διάμεσο στο μέγιστο (7 συμμετοχές).

Στη μεταβλητή που αφορά τη σχέση μεταξύ του πραγματικού αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει η αναλογία, με μέσο όρο 15,61, υποδηλώνει κατά μέσο όρο ότι ο πραγματικός αριθμός επισκέψεων είναι περίπου 15,61% των πιθανών επισκέψεων. Η κατανομή είναι λοξή προς τα δεξιά, υποδεικνύοντας μεταβλητότητα στον τρόπο με τον οποίο οι χρήστες αξιοποιούν τις ευκαιρίες των επισκέψεών τους.

Ο αριθμός ανανεώσεων κατά την περίοδο εγγραφής έχει μέσο όρο 1,21 και εύρος από 0 έως 6, οι περισσότεροι χρήστες ανανεώνουν την εγγραφή τους τουλάχιστον μία φορά, με μια σχετικά χαμηλή τυπική απόκλιση που υποδηλώνει περιορισμένη μεταβλητότητα.

Η ύπαρξη ατόμων με τα οποία σχετίζεται ο χρήστης από οικογενειακή σχέση ή φίλια διακρίνεται από έντονη λοξότητα και κύρτωση, αυτή η δυαδική μεταβλητή δείχνει ότι οι περισσότεροι χρήστες δεν έχουν οικογενειακές ή φιλικές συνδέσεις εντός της αθλητικής πλατφόρμας, με πολύ λίγες εξαιρέσεις.

ΜΕΛΕΤΗ ΠΑΡΑΓΟΝΤΩΝ ΠΟΥ ΕΠΙΔΡΟΥΝ ΣΤΗΝ ΕΠΙΛΟΓΗ  
ΑΘΛΗΤΙΚΗΣ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ ΜΕΣΩ ΔΙΑΔΙΚΤΥΑΚΗΣ ΠΛΑΤΦΟΡΜΑΣ

Μεταβλητές	$Q_1$	$Q_3$	Μέση τιμή	95% Διάστημα Εμπιστοσύνης για τη μέση τιμή (Κάτω όριο)	95% Διάστημα Εμπιστοσύνης για τη μέση τιμή (Άνω όριο)	Διάμεσος
Ηλικία	19	31	26,02	25,79	26,24	23
Διάρκεια Εγγραφής	4	19	13,97	13,74	14,2	9
Αριθμός ημερών που ο χρήστης δεν επισκέφτηκε τις πλατφόρμα πριν θεωρηθεί λήξη	13	84	81,22	78,91	83,54	41
Συνολικό ποσό πληρωμής	83,6	355,1	302,56	296,72	308,4	166,2
Μέση εβδομαδιαία συχνότητα σε όλη την περίοδο εγγραφής	0,29	1,01	0,76	0,75	0,77	0,57
Αριθμός δραστηριοτήτων	1	1	1,1	1,09	1,1	1
Αριθμός επισκέψεων στην αθλητική πλατφόρμα από την ημερομηνία έναρξης και την ημερομηνία διακοπής	7	46	40,16	39,11	41,21	18
Μέσος όρος αριθμού συμμετοχής σε δραστηριότητες από την ημερομηνία έναρξης και την ημερομηνία διακοπής	0	0,1	0,14	0,13	0,14	0
Αριθμός των εβδομαδιαίων επισκέψεων στις πλατφόρμα τους τελευταίους 2 μήνες της εγγραφής του	4	7	5,75	5,72	5,79	7
Αριθμός των επισκέψεων στις πλατφόρμα τους τελευταίους 2 μήνες της εγγραφής του	25,72	60,97	41,64	41,3	41,97	38,99
Πραγματικός αριθμός των επισκέψεων	1	7	5,32	5,22	5,42	4
Σχέση μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει	3,22	22,96	15,61	15,33	15,89	9,99
Αριθμός ανανεώσεων κατά την περίοδο εγγραφής	0	2	1,21	1,18	1,23	1
Υπαρξη ατόμων με τα οποία σχετίζεται ο χρήστης από οικογενειακή σχέση ή φίλια	0	0	0,02	0,02	0,02	0

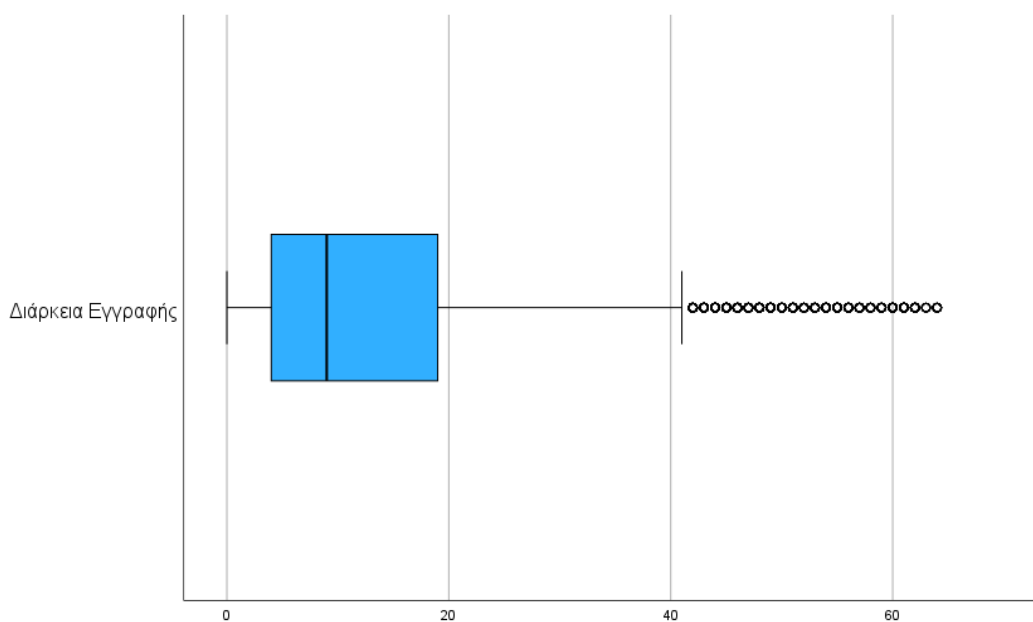
Πίνακας 3: Στοιχεία Περιγραφικής Στατιστικής

Μεταβλητές	Τυπική απόκλιση	Ελάχιστη τιμή	Μέγιστη τιμή	Εύρος	Ασυμμετρία	Κύρτωση
Ηλικία	14,155	0	87	87	1,074	1,678
Διάρκεια Εγγραφής	14,164	0	64	64	1,556	1,783
Αριθμός ημερών που ο χρήστης δεν επισκέφτηκε τις πλατφόρμα πριν θεωρηθεί λήξη	144,2	0	1745	1745	5,195	36,668
Συνολικό ποσό πληρωμής	364,32	0	6728	6728	3,054	18,267
Μέση εβδομαδιαία συχνότητα σε όλη την περίοδο εγγραφής	0,668	0	5	5	1,834	4,474
Αριθμός δραστηριοτήτων	0,324	1	5	4	3,733	16,499
Αριθμός επισκέψεων στην αθλητική πλατφόρμα από την ημερομηνία έναρξης και την ημερομηνία διακοπής	65,478	1	1031	1030	4,837	37,638
Μέσος όρος αριθμού συμμετοχής σε δραστηριότητες από την ημερομηνία έναρξης και την ημερομηνία διακοπής	0,292	0	4	4	2,96	12
Αριθμός των εβδομαδιαίων επισκέψεων στις πλατφόρμα τους τελευταίους 2 μήνες της εγγραφής του	2,122	1	7	6	-1,222	-0,338

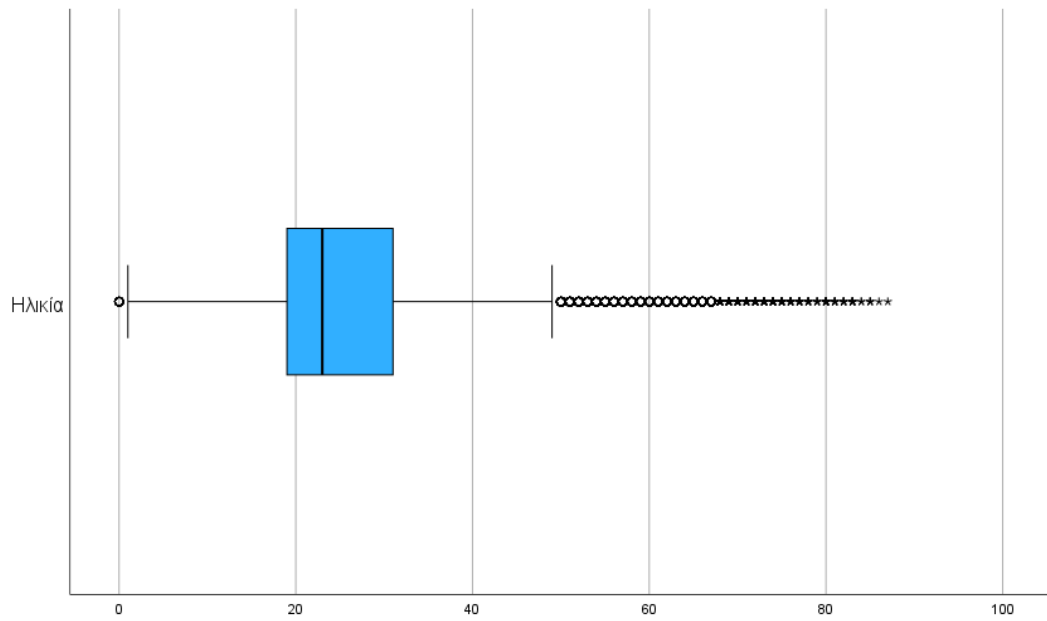
ΜΕΛΕΤΗ ΠΑΡΑΓΟΝΤΩΝ ΠΟΥ ΕΠΙΔΡΟΥΝ ΣΤΗΝ ΕΠΙΛΟΓΗ  
ΑΘΛΗΤΙΚΗΣ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ ΜΕΣΩ ΔΙΑΔΙΚΤΥΑΚΗΣ ΠΛΑΤΦΟΡΜΑΣ

Αριθμός των επισκέψεων στις πλατφόρμα τους τελευταίους 2 μήνες της εγγραφής του	21,066	1	240	239	1,09	9,753
Πραγματικός αριθμός των επισκέψεων	6,333	0	84	84	2,873	14,621
Σχέση μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει	17,539	0	100	100	1,784	3,758
Αριθμός ανανεώσεων κατά την περίοδο εγγραφής	1,381	0	6	6	1,199	0,836
Υπαρξη ατόμων με τα οποία σχετίζεται ο χρήστης από οικογενειακή σχέση ή φίλια	0,139	0	1	1	6,893	45,516

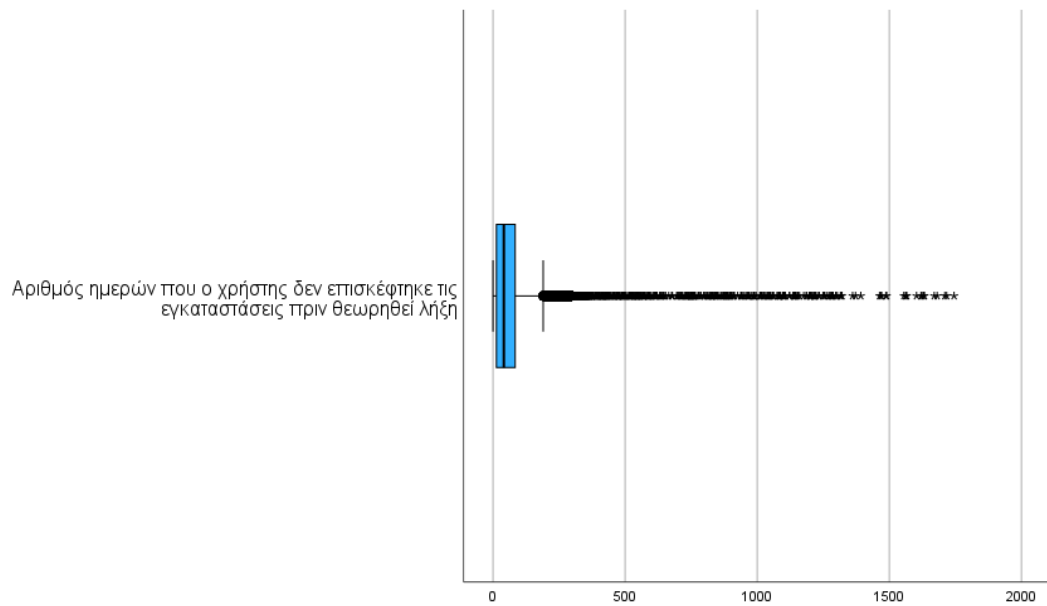
Πίνακας 3: Στοιχεία Περιγραφικής Στατιστικής



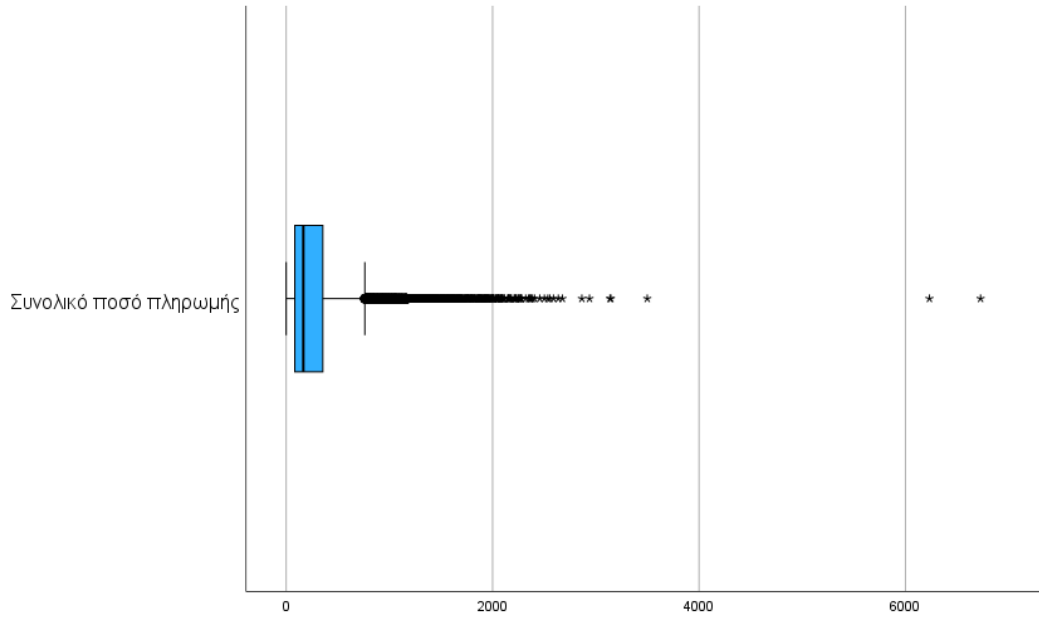
Γράφημα 1: Διάγραμμα κουτιού της διάρκειας εγγραφής χρηστών



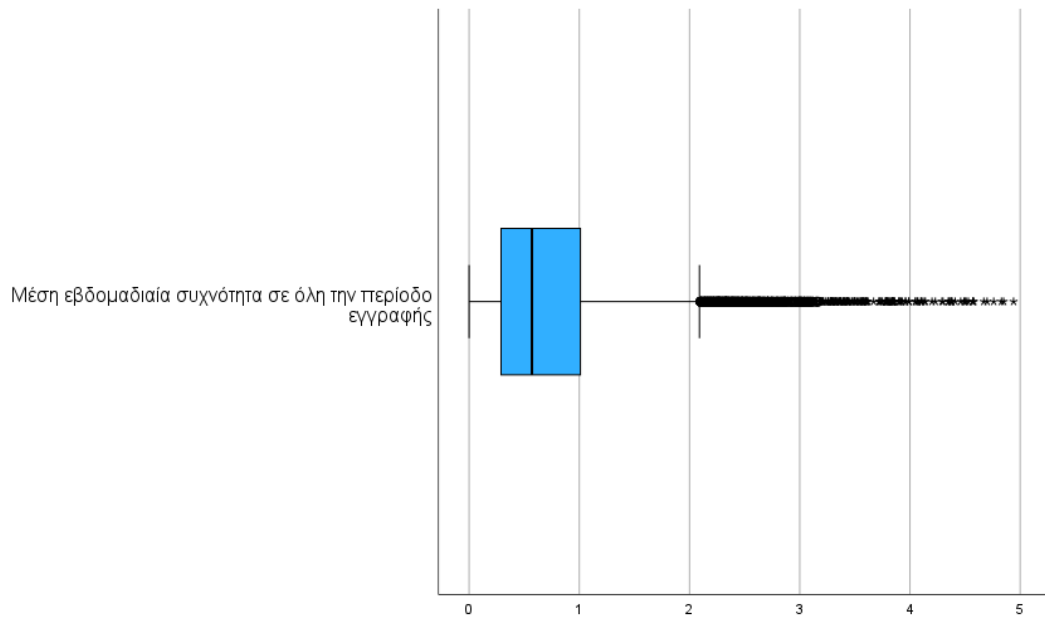
Γράφημα 2: Θηκόγραμμα Ηλικίας



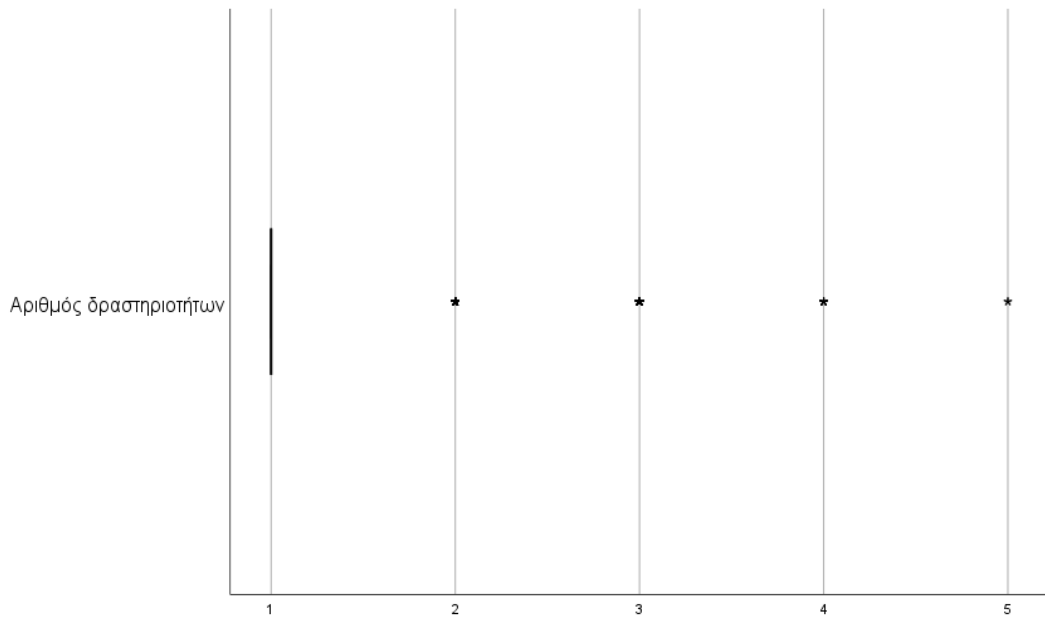
Γράφημα 3: Θηκόγραμμα Αριθμού Ημερών Πριν Θεωρηθεί Λήξη Εγγραφής



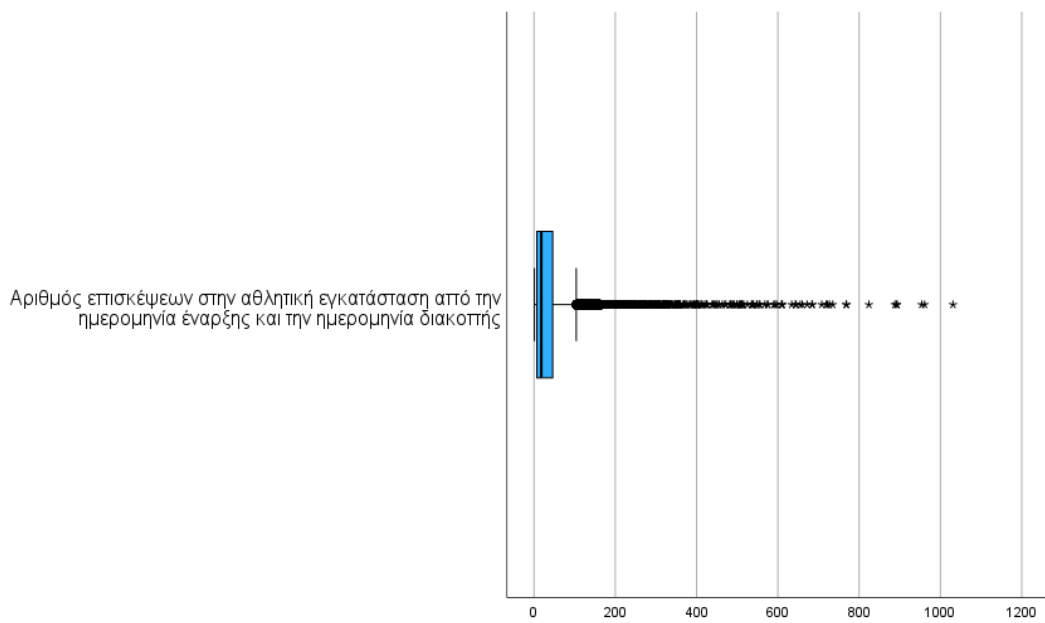
Γράφημα 4: Θηκόγραμμα Συνολικού Ποσού Πληρωμής



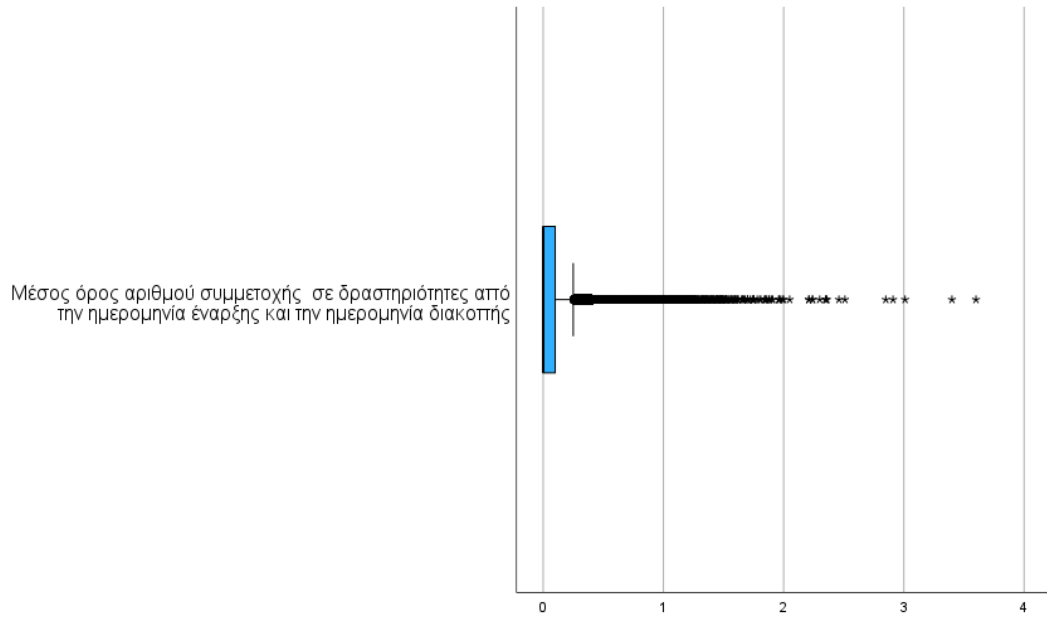
Γράφημα 5: Θηκόγραμμα Μέσης Εβδομαδιαίας Συχνότητας σε Όλη την Περίοδο Εγγραφής



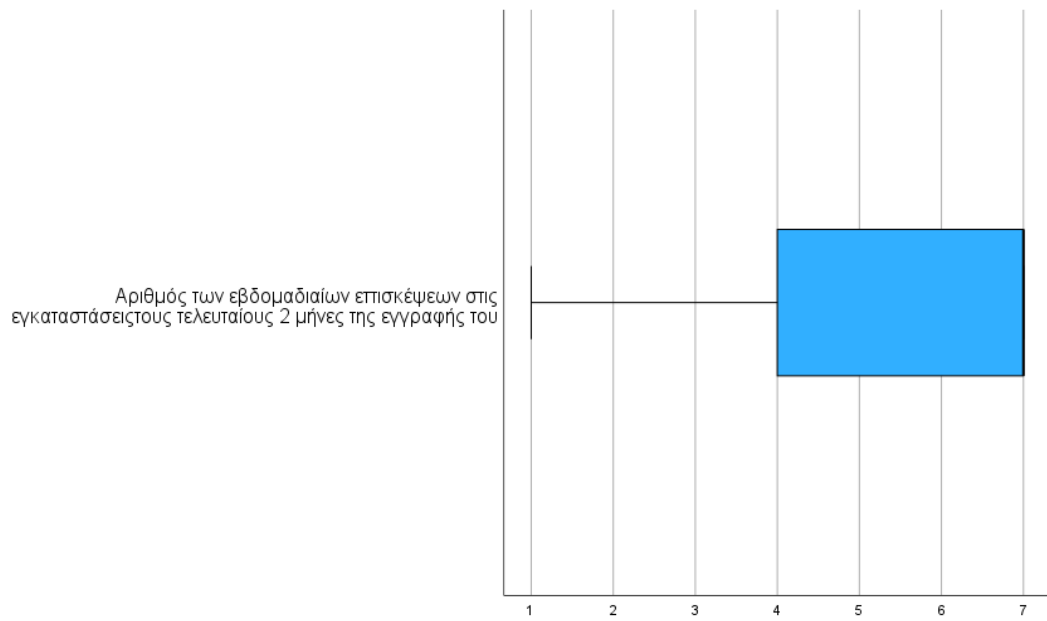
Γράφημα 6: Θηκόγραμμα Αριθμού Δραστηριοτήτων



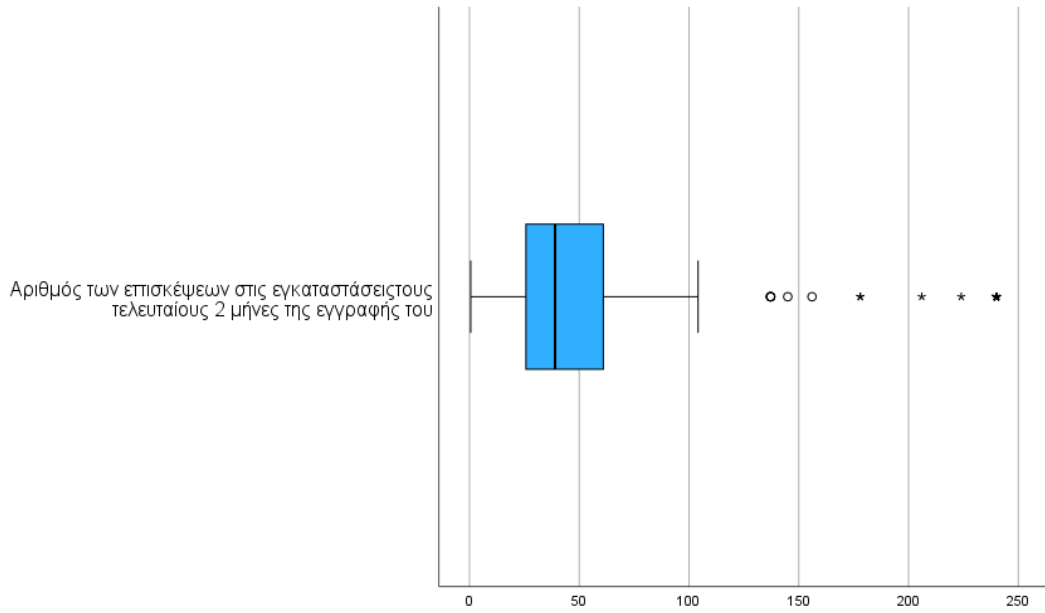
Γράφημα 7: Θηκόγραμμα Αριθμού Επισκέψεων Από Έναρξη Έως Διακοπή Εγγραφής



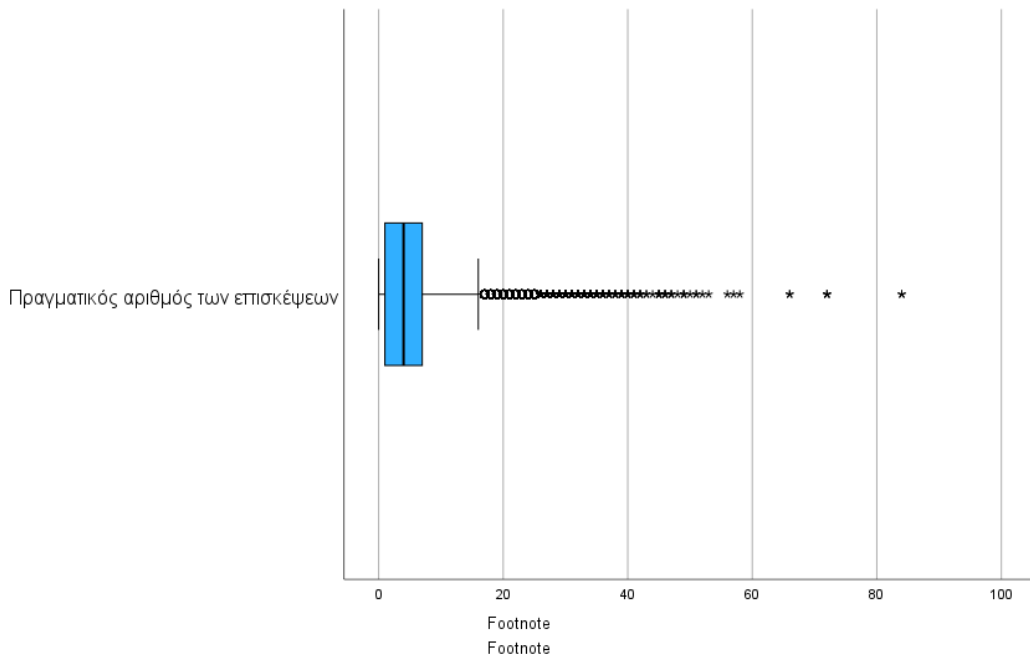
Γράφημα 8: Θηκόγραμμα Μέσου Αριθμού Συμμετοχής σε Δραστηριότητες



Γράφημα 9: Θηκόγραμμα Εβδομαδιαίων Επισκέψεων Τελευταίων 2 Μηνών Εγγραφής

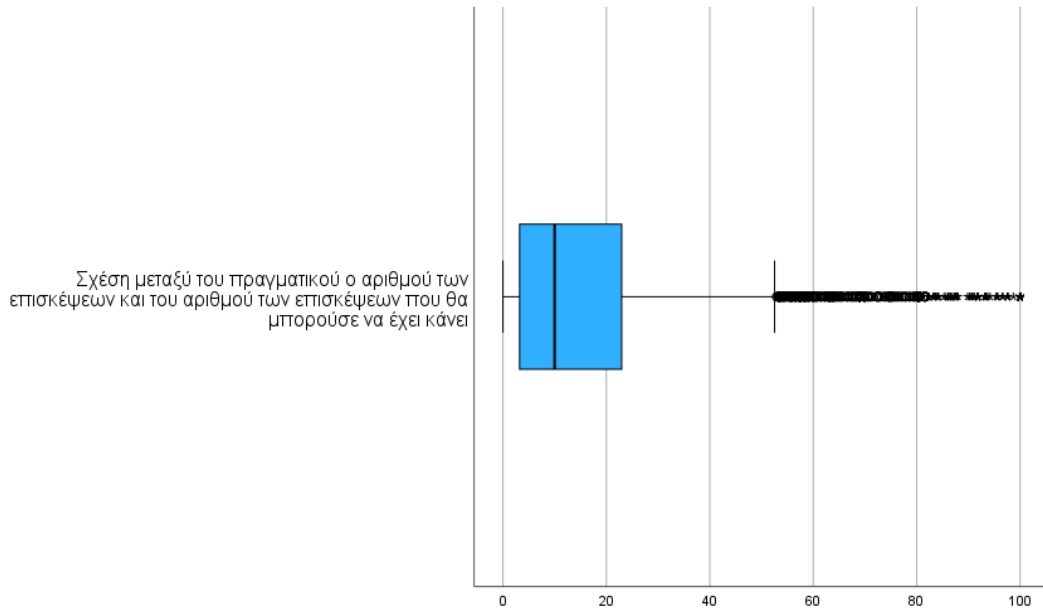


Γράφημα 10: Θηκόγραμμα Αριθμού Επισκέψεων στις Εγκαταστάσεις τους Τελευταίους 2 Μήνες Εγγραφής

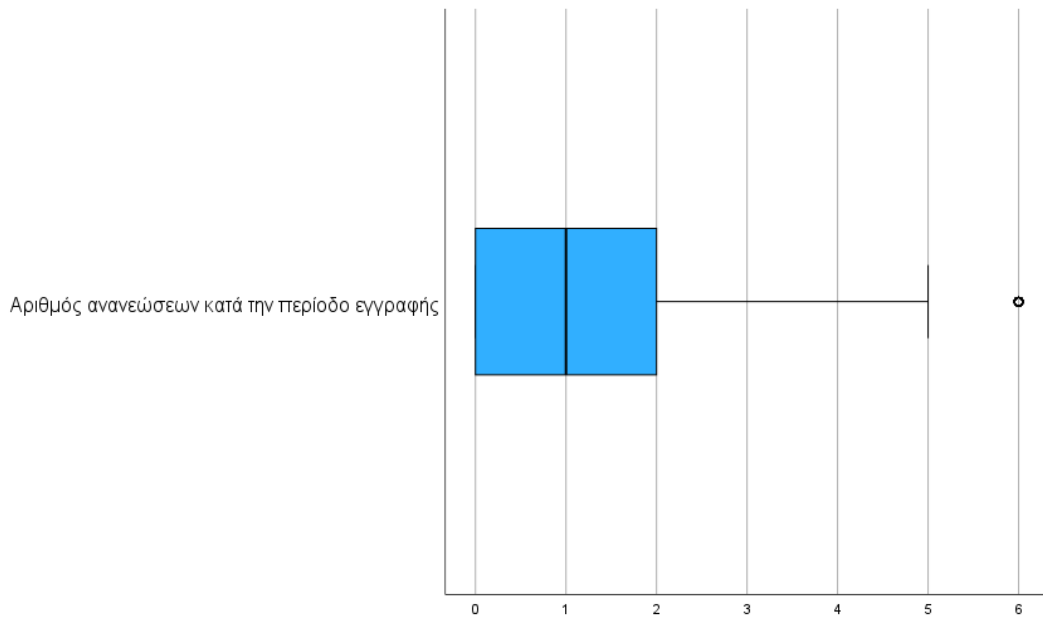


Γράφημα 11: Θηκόγραμμα Πραγματικού Αριθμού Επισκέψεων





Γράφημα 12: Θηκόγραμμα Σχέσης Πραγματικών και Θεωρητικών Επισκέψεων



Γράφημα 13: Θηκόγραμμα Αριθμού Ανανεώσεων Κατά την Περίοδο Εγγραφή

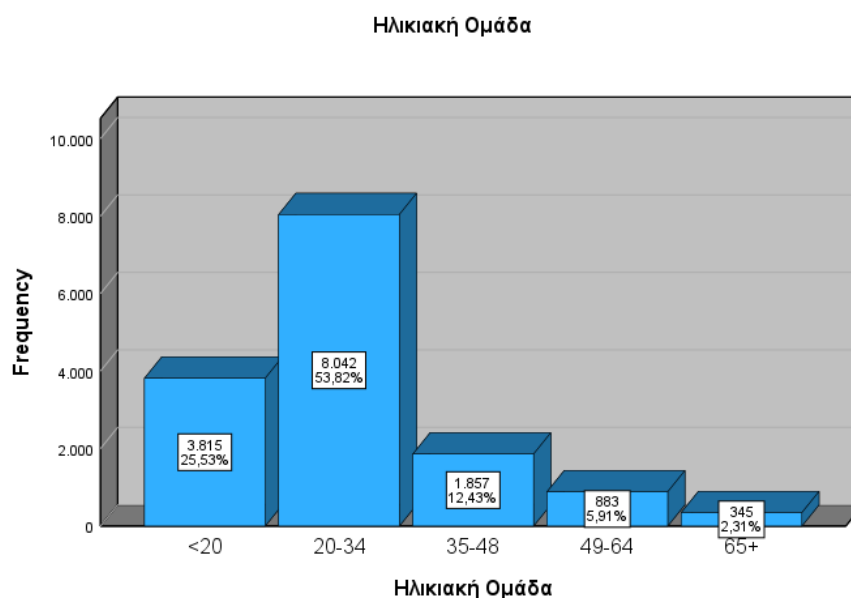
Σχετικά με την κατανομή των ηλικιακών ομάδων, 3.815 άτομα είναι κάτω των 20 ετών αποτελούν (25,5%), τονίζοντας τη μεγάλη συμμετοχή των νεότερων εγγεγραμμένων. Η ηλικιακή ομάδα 20-34 ετών είναι η επικρατέστερη 8.042 άτομα, αποτελώντας το 53,8%, δείχνοντας ότι οι νέοι ενήλικες είναι ιδιαίτερα δραστήριοι. Η ηλικιακή ομάδα 35-48 ετών, αποτελείται από 1.857 άτομα (12,4%). Η ηλικιακή ομάδα 49-64 ετών περιλαμβάνει 883 άτομα (5,9%), δείχνοντας μικρότερη συμμετοχή από

μεγαλύτερης ηλικίας άτομα, ενώ η κατηγορία 65+ έχει 345 άτομα (2,3%) είναι η ομάδα με τη μικρότερη συχνότητα, πιθανόν αντανακλώντας τη μειωμένη συμμετοχή από τα ηλικιωμένα άτομα.

#### Ηλικιακή Ομάδα

	N	%
<20	3815	25,5%
20-34	8042	53,8%
35-48	1857	12,4%
49-64	883	5,9%
65+	345	2,3%

Πίνακας 4: Κατανομή Συχνότητας Ηλικιακών Ομάδων Χρηστών



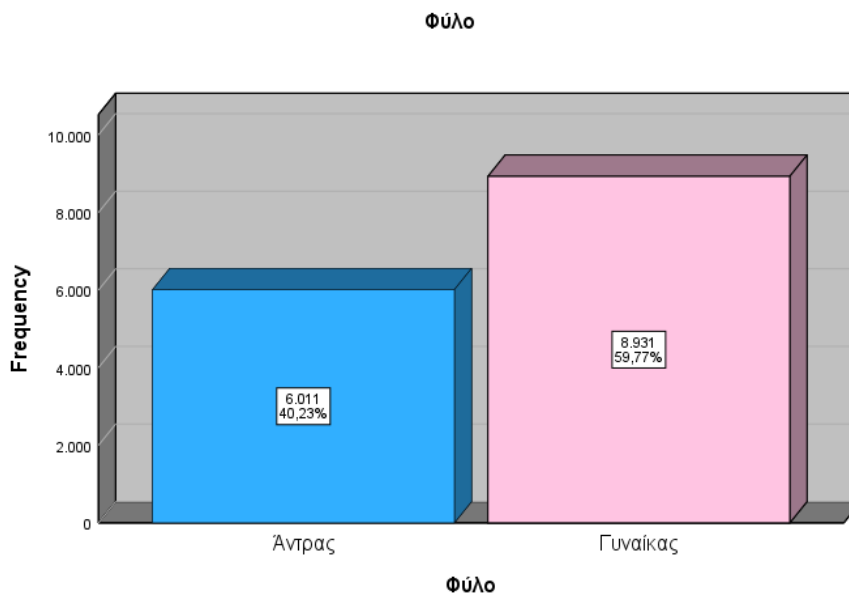
Γράφημα 14: Κατανομή Συχνότητας Ηλικιακών Ομάδων Χρηστών

Σχετικά με την κατανομή βάσει φύλου, τα δεδομένα δείχνουν ότι 6,011 άντρες αποτελούν το 40,2% του συνόλου, ενώ 8,931 γυναίκες αντιπροσωπεύουν το 59,8%. Οι γυναίκες συμμετέχουν περισσότερο σε αθλητικές δραστηριότητες, σε σύγκριση με τους άντρες.

### Φύλο

	N	%
Άντρας	6011	40,2%
Γυναίκα	8931	59,8%

Πίνακας 5: Συχνότητα Φύλου Χρηστών



Γράφημα 15: Συχνότητα Φύλου Χρηστών

Στο πλαίσιο της διάρκειας εγγραφής, σύμφωνα με τη μέθοδο Hughes, τα δεδομένα αποκαλύπτουν μια ισορροπημένη κατανομή μεταξύ των κατηγοριών. Αναλυτικότερα, 2,989 άτομα ανήκουν σε κάθε κατηγορία, αποτελώντας το 20% του συνόλου.

### Διάρκεια Εγγραφής Hughes

	N	%
1	2989	20,0%
2	2989	20,0%
3	2988	20,0%
4	2988	20,0%

5	2988	20,0%
---	------	-------

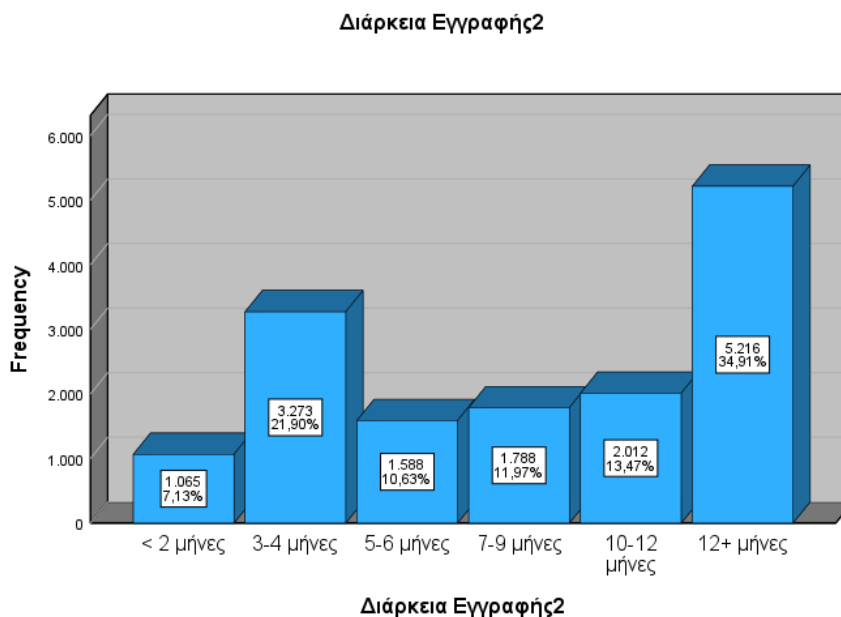
Πίνακας 6: Συχνότητα Διάρκειας Εγγραφής Hughes

Για τη διάρκεια εγγραφής, όπως καταγράφεται στη δεύτερη κατηγοριοποίηση, τα δεδομένα φανερώνουν ενδιαφέρουσες τάσεις στην εμπλοκή των συμμετεχόντων. Συγκεκριμένα, 1,065 άτομα με διάρκεια εγγραφής κάτω από 2 μήνες αντιστοιχούν στο 7,1% του συνόλου, δείχνοντας την πιο σύντομη δραστηριοποίηση. Η κατηγορία των 3-4 μηνών είναι η δεύτερη πιο μεγάλη με 3,273 άτομα, αντιπροσωπεύοντας το 21,9%, υποδηλώνοντας ότι η πλειονότητα των ατόμων παραμένει ενεργή για αυτό το χρονικό διάστημα. Η κατηγορία των 5-6 μηνών περιλαμβάνει 1,588 άτομα ή 10,6%, ενώ 1,788 άτομα με διάρκεια εγγραφής 7-9 μήνες αντιστοιχούν στο 12%. Η ομάδα 10-12 μηνών εμφανίζει 2,012 άτομα ή 13,5%. Τέλος, η μεγαλύτερη εμπλοκή παρατηρείται στην κατηγορία των 12+ μηνών με 5,216 άτομα, αντιπροσωπεύοντας το 34,9% του συνόλου, αποδεικνύοντας ότι ένα σημαντικό ποσοστό συμμετεχόντων διατηρεί μακροχρόνια συνέπεια στον αθλητισμό.

#### Διάρκεια Εγγραφής<sup>2</sup>

	N	%
< 2 μήνες	1065	7,1%
3-4 μήνες	3273	21,9%
5-6 μήνες	1588	10,6%
7-9 μήνες	1788	12,0%
10-12 μήνες	2012	13,5%
12+ μήνες	5216	34,9%

Πίνακας 7: Συχνότητα Διάρκειας Εγγραφής (2 μήνες έως 12+ μήνες)



Γράφημα 16: Συχνότητα Διάρκειας Εγγραφής (2 μήνες έως 12+ μήνες)

Αναφορικά με τον αριθμό ημερών που ένας χρήστης δεν επισκέφτηκε την πλατφόρμα πριν θεωρηθεί ως λήξη, σύμφωνα με τη μέθοδο Hughes, τα δεδομένα δείχνουν μια ομοιόμορφη κατανομή στις κατηγορίες. Κάθε μία από τις τέσσερις κατηγορίες περιλαμβάνει 2,989 άτομα, αντιπροσωπεύοντας ακριβώς το 20% του συνόλου.

**Αριθμός ημερών που ο χρήστης δεν επισκέφτηκε τις πλατφόρμα πριν θεωρηθεί λήξη Hughes**

	N	%
1	2989	20,0%
2	2989	20,0%
3	2988	20,0%
4	2988	20,0%
5	2988	20,0%

Πίνακας 8: Συχνότητα Αποχής Χρηστών Πριν τη Λήξη Εγγραφής Hughes

Σε ό,τι αφορά τον αριθμό των ημερών που ο χρήστης δεν επισκέφτηκε την πλατφόρμα πριν θεωρηθεί λήξη, 0-7 μέρες απουσίας είχαν 2,745 χρήστες 18,4%,

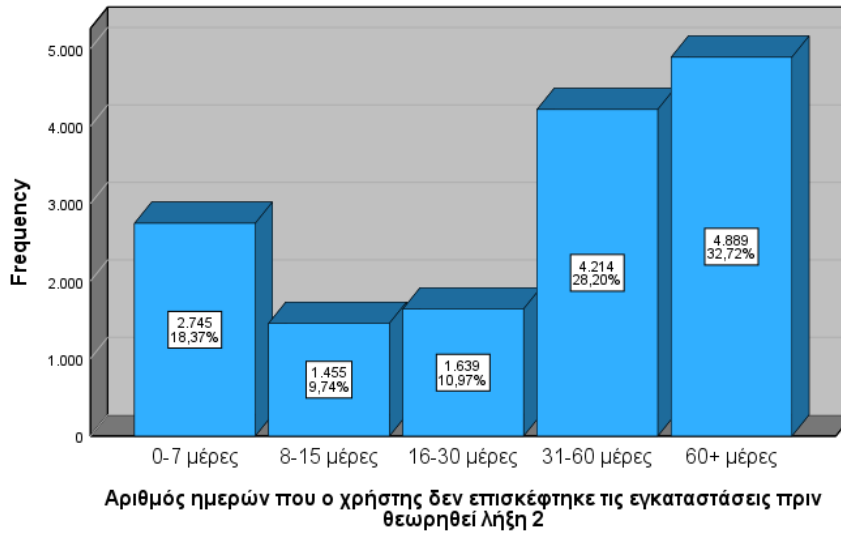
δείχνοντας ότι ένα σημαντικό ποσοστό των χρηστών διατηρεί σχετικά συχνή παρουσία. Στο εύρος των 8-15 μερών, 1,455 χρήστες 9,7% φαίνεται να απέχουν για μικρότερες περιόδους. Η κατηγορία 16-30 μερών περιλαμβάνει 1,639 χρήστες 11%, υποδηλώνοντας μέτρια συχνότητα επισκέψεων. Για το διάστημα των 31-60 μερών, ο αριθμός αυξάνεται σημαντικά στους 4,214 χρήστες, αντιπροσωπεύοντας το 28,2%, ενώ το μεγαλύτερο ποσοστό, το 32,7% δηλαδή 4,889 χρήστες, αφορά σε εκείνους που απέχουν για περισσότερες από 60 μέρες. Αυτή η κατανομή υποδηλώνει μια τάση αυξανόμενης αποχής από τις αθλητικές πλατφόρμες με το πέρασμα του χρόνου, με τους χρήστες που απέχουν περισσότερο από δύο μήνες να αποτελούν την πλειοψηφία.

**Αριθμός ημερών που ο χρήστης δεν επισκέφτηκε την πλατφόρμα πριν θεωρηθεί λήξη 2**

	N	%
0-7 μέρες	2745	18,4%
8-15 μέρες	1455	9,7%
16-30 μέρες	1639	11,0%
31-60 μέρες	4214	28,2%
60+ μέρες	4889	32,7%

Πίνακας 9: Συχνότητα Ημερών Αδράνειας Πριν τη Λήξη Εγγραφής (2)

**Αριθμός ημερών που ο χρήστης δεν επισκέφτηκε τις εγκαταστάσεις πριν θεωρηθεί λήξη 2**



Γράφημα 17: Συχνότητα Ημερών Αδράνειας Πριν τη Λήξη Εγγραφής (2)

Σχετικά με το συνολικό ποσό πληρωμής βάσει της κατηγοριοποίησης Hughes, τα δεδομένα δείχνουν έναν ομοιόμορφο διαμοιρασμό των συμμετεχόντων μεταξύ των πέντε κατηγοριών. Κάθε κατηγορία περιλαμβάνει 2,989 άτομα, με καθεμία να αντιστοιχεί σε ποσοστό 20% του συνόλου.

**Συνολικό ποσό πληρωμής Hughes**

	N	%
1	2989	20,0%
2	2989	20,0%
3	2988	20,0%
4	2988	20,0%
5	2988	20,0%

Πίνακας 10: Συχνότητα Συνολικού Ποσού Πληρωμής Hughes

Για τη μέση εβδομαδιαία συχνότητα κατά τη διάρκεια της εγγραφής, σύμφωνα με την κατηγοριοποίηση Hughes, παρατηρείται μια ακριβής ισορροπία στην κατανομή των συμμετεχόντων ανάμεσα στις διάφορες κατηγορίες. Συγκεκριμένα, 2,989 άτομα εντάσσονται στην κάθε κατηγορία και περιλαμβάνουν από 2,988 άτομα, με καθεμία από αυτές τις κατηγορίες να αποτελεί το 20% του συνόλου.

**Μέση εβδομαδιαία συχνότητα σε  
όλη την περίοδο εγγραφής Hughes**

	N	%
1	2989	20,0%
2	2989	20,0%
3	2988	20,0%
4	2988	20,0%
5	2988	20,0%

Πίνακας 11: Συχνότητα Μέσης Εβδομαδιαίας Συχνότητας Κατά τη Διάρκεια Εγγραφής Hughes

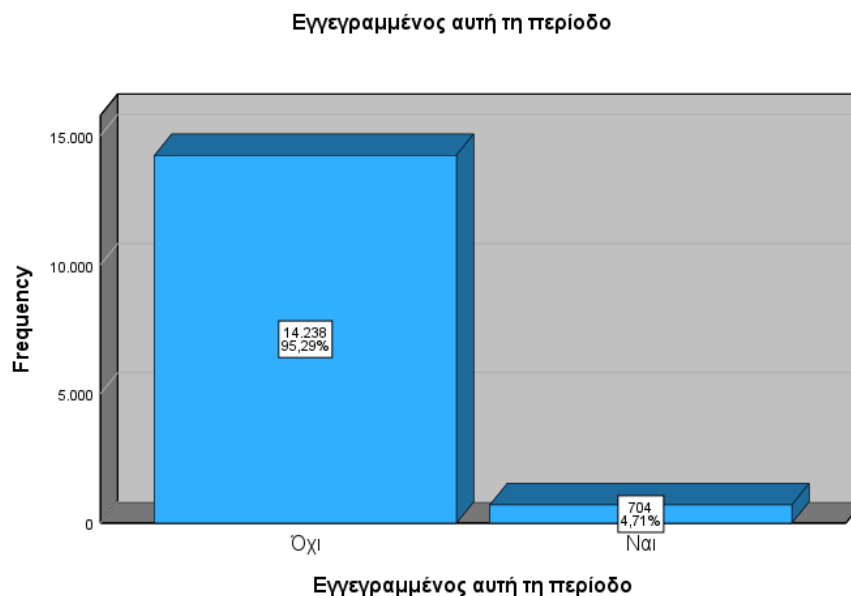
Στην κατηγορία των εγγεγραμμένων αυτή την περίοδο, τα δεδομένα δείχνουν ότι 14,238 άτομα, ή το 95,3% του συνόλου, δεν ήταν εγγεγραμμένα στις πλατφόρμες ή σε κάποια συγκεκριμένη δραστηριότητα κατά τη διάρκεια της εξεταζόμενης περιόδου. Από την άλλη πλευρά, μόλις 704 άτομα, το 4,7% του συνόλου, ήταν πράγματι εγγεγραμμένοι και ενεργοί συμμετέχοντες.

**Εγγεγραμμένος αυτή τη περίοδο**

	N	%
Όχι	14238	95,3%
Ναι	704	4,7%

Πίνακας 12: Συχνότητα Εγγεγραμμένων Χρηστών στην Περίοδο





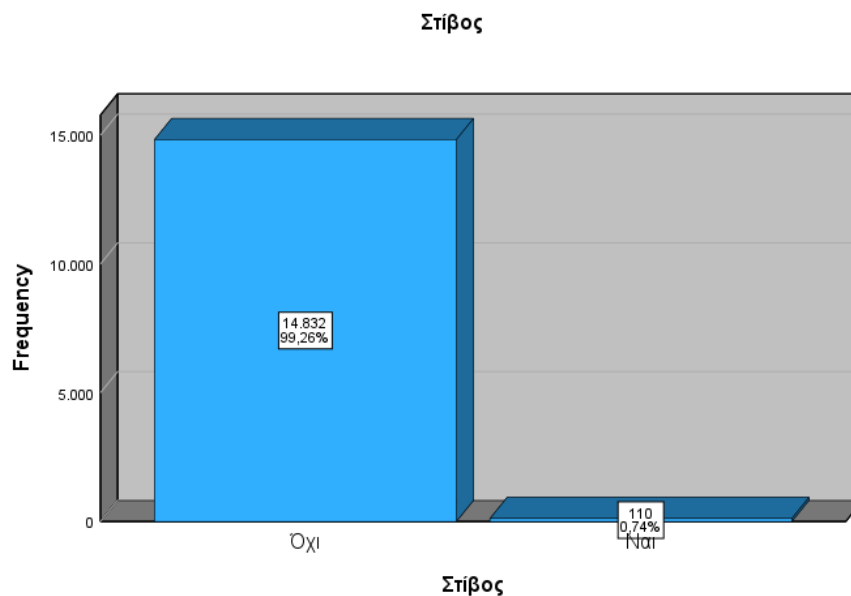
Γράφημα 18: Συχνότητα Εγγεγραμμένων Χρηστών στην Περίοδο

Όσον αφορά τη συμμετοχή σε δραστηριότητες στίβου, τα δεδομένα αποκαλύπτουν ότι 14,832 άτομα, 99,3% του συνόλου, δεν έχουν συμμετάσχει σε τέτοιου είδους δραστηριότητες. Αντιθέτως, μόνο 110 άτομα, που αντιστοιχούν το 0,7% του δείγματος, έχουν ενεργή συμμετοχή.

**Στίβος**

	N	%
Όχι	14832	99,3%
Ναι	110	0,7%

Πίνακας 13: Συχνότητα Συμμετοχής σε Στίβο



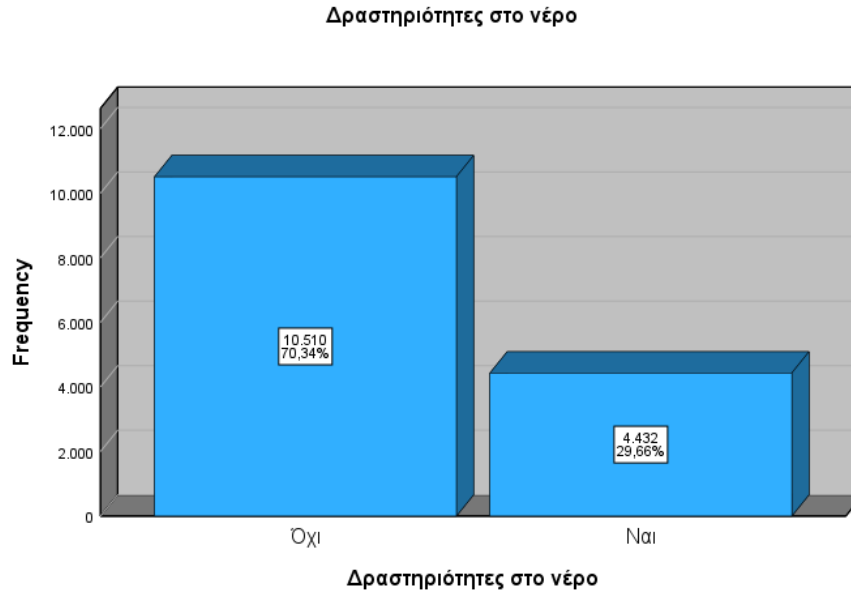
Γράφημα 19: Συχνότητα Συμμετοχής σε Στίβο

Στον τομέα των δραστηριοτήτων στο νερό, η κατανομή των συμμετεχόντων δείχνει ότι 10,510 άτομα, 70,3% του συνόλου, δεν έχουν συμμετάσχει σε τέτοιου είδους δραστηριότητες. Από την άλλη, 4,432 άτομα, το 29,7% του δείγματος, έχουν ενεργή συμμετοχή σε δραστηριότητες στο νερό.

#### Δραστηριότητες στο νερό

	N	%
Όχι	10510	70,3%
Ναι	4432	29,7%

Πίνακας 14: Συχνότητα Συμμετοχής σε Δραστηριότητες Νερού



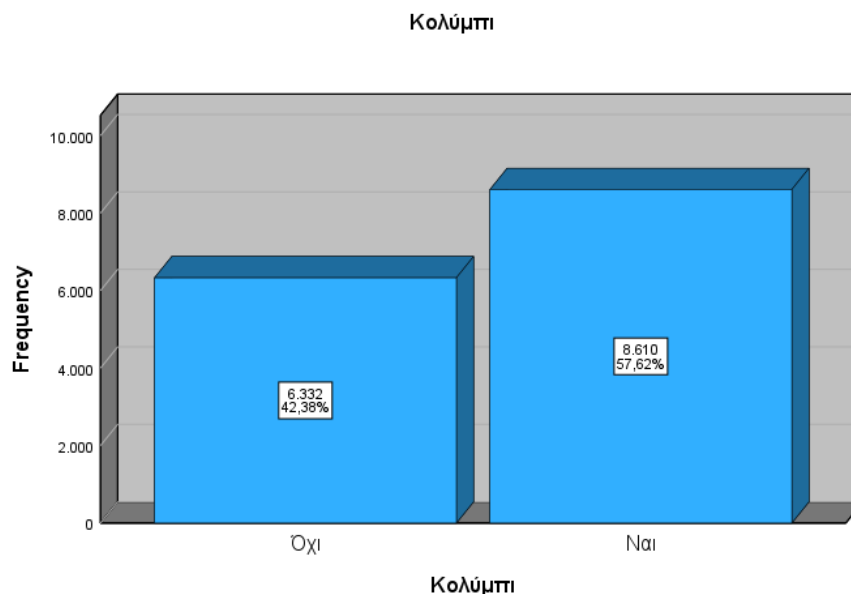
Γράφημα 20: Συχνότητα Συμμετοχής σε Δραστηριότητες Νερού

Σε ό,τι αφορά τη συμμετοχή στο κολύμπι, τα δεδομένα αναδεικνύουν ότι 6,332 άτομα, 42,4% του συνόλου, δεν έχουν εμπλακεί σε δραστηριότητες κολύμβησης. Αντιθέτως, 8,610 άτομα, δηλαδή το 57,6% των συμμετεχόντων, έχουν συμμετάσχει ενεργά στο κολύμπι.

**Κολύμπι**

	N	%
Όχι	6332	42,4%
Ναι	8610	57,6%

Πίνακας 4: Συχνότητα Συμμετοχής στο Κολύμπι



Γράφημα 21: Συχνότητα Συμμετοχής στο Κολύμπι

Όσον αφορά τη συμμετοχή στον χορό, τα δεδομένα δείχνουν ότι όλα τα 14,942 άτομα του δείγματος, δηλαδή το 100% του συνόλου, δεν έχουν συμμετάσχει σε δραστηριότητες χορού.

#### Χορός

	N	%
Όχι	14942	100,0%

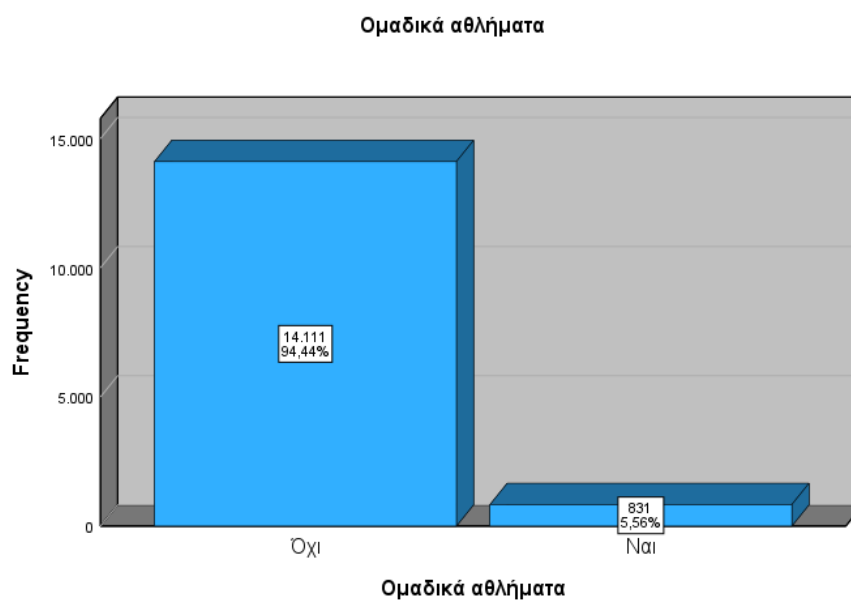
Πίνακας 16: Συχνότητα Συμμετοχής στον Χορό

Σχετικά με τη συμμετοχή σε ομαδικά αθλήματα, 14,111 άτομα, 94,4% του συνόλου, δεν έχουν συμμετάσχει σε τέτοιου είδους δραστηριότητες. Αντίθετα, μόλις 831 άτομα, το 5,6% του δείγματος, έχουν ενεργή συμμετοχή σε ομαδικά αθλήματα.

#### Ομαδικά αθλήματα

	N	%
Όχι	14111	94,4%
Ναι	831	5,6%

Πίνακας 17: Συχνότητα Συμμετοχής σε Ομαδικά Αθλήματα



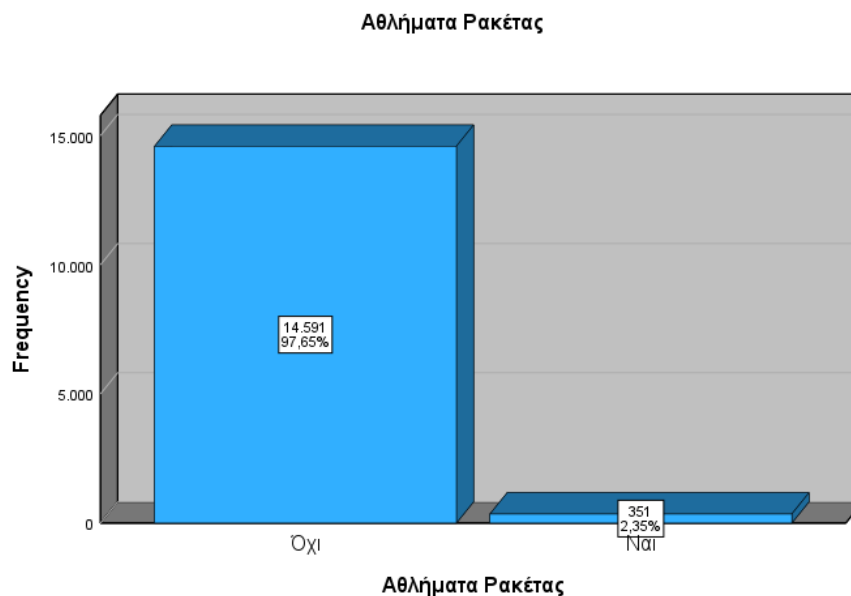
Γράφημα 22: Συχνότητα Συμμετοχής σε Ομαδικά Αθλήματα

Αναφορικά με τη συμμετοχή σε αθλήματα ρακέτας, 14,591 άτομα, 97,7%, δεν έχουν εμπλακεί σε τέτοιου είδους δραστηριότητες ενώ 351 άτομα, 2,3% έχουν συμμετάσχει σε αθλήματα ρακέτας.

#### **Αθλήματα Ρακέτας**

	N	%
Όχι	14591	97,7%
Ναι	351	2,3%

Πίνακας 18: Συχνότητα Συμμετοχής σε Αθλήματα Ρακέτας



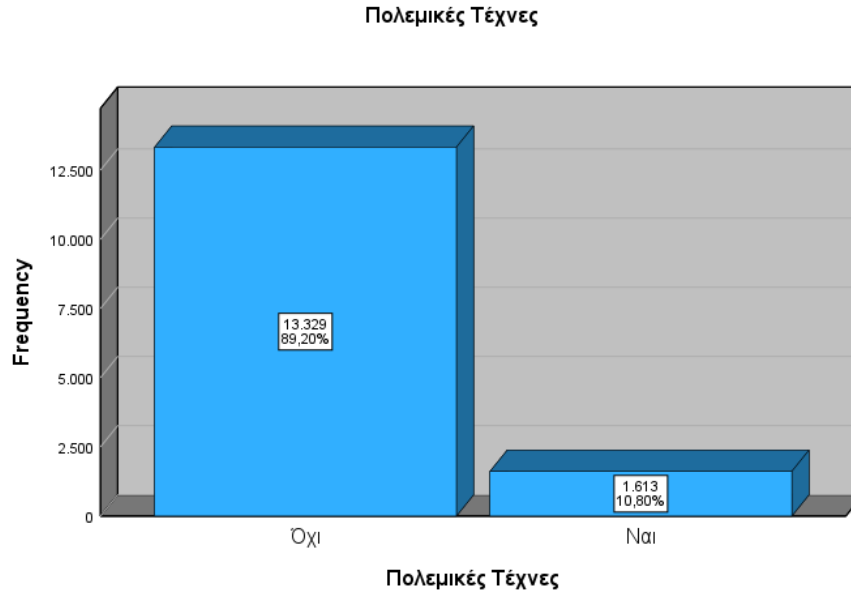
Γράφημα 23: Συχνότητα Συμμετοχής σε Αθλήματα Ρακέτας

Σε ό,τι αφορά τη συμμετοχή στις πολεμικές τέχνες, 13,329 άτομα, 89,2%, δεν έχουν συμμετάσχει σε τέτοιου είδους δραστηριότητες. Αντιθέτως, 1,613 άτομα, 10,8%, έχουν ενεργή συμμετοχή στις πολεμικές τέχνες.

#### Πολεμικές Τέχνες

	N	%
Όχι	13329	89,2%
Ναι	1613	10,8%

Πίνακας 19: Συχνότητα Συμμετοχής σε Πολεμικές Τέχνες



Γράφημα 24: Συχνότητα Συμμετοχής σε Πολεμικές Τέχνες

Σχετικά με τη συμμετοχή σε δραστηριότητες στη φύση, τα δεδομένα δείχνουν ότι κανένα από τα 14,942 άτομα του δείγματος, δηλαδή το 100% του συνόλου, δεν έχει συμμετάσχει σε τέτοιου είδους δραστηριότητες.

#### Δραστηριότητες στη φύση

	N	%
Όχι	14942	100,0%

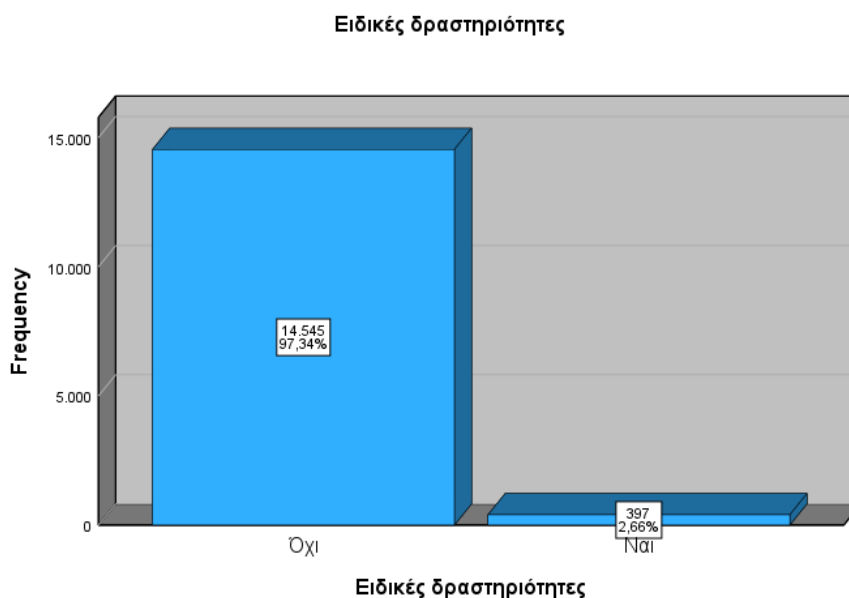
Πίνακας 20: Συχνότητα Συμμετοχής σε Δραστηριότητες στη Φύση

Σε ό,τι αφορά τη συμμετοχή σε ειδικές δραστηριότητες, 14,545 άτομα, 97,3% δεν έχουν συμμετάσχει σε τέτοιου είδους δραστηριότητες, ενώ 397 άτομα, 2,7% έχουν ενεργή συμμετοχή σε ειδικές δραστηριότητες.

#### Ειδικές δραστηριότητες

	N	%
Όχι	14545	97,3%
Ναι	397	2,7%

Πίνακας 21: Συχνότητα Συμμετοχής σε Ειδικές Δραστηριότητες



Γράφημα 25: Συχνότητα Συμμετοχής σε Ειδικές Δραστηριότητες

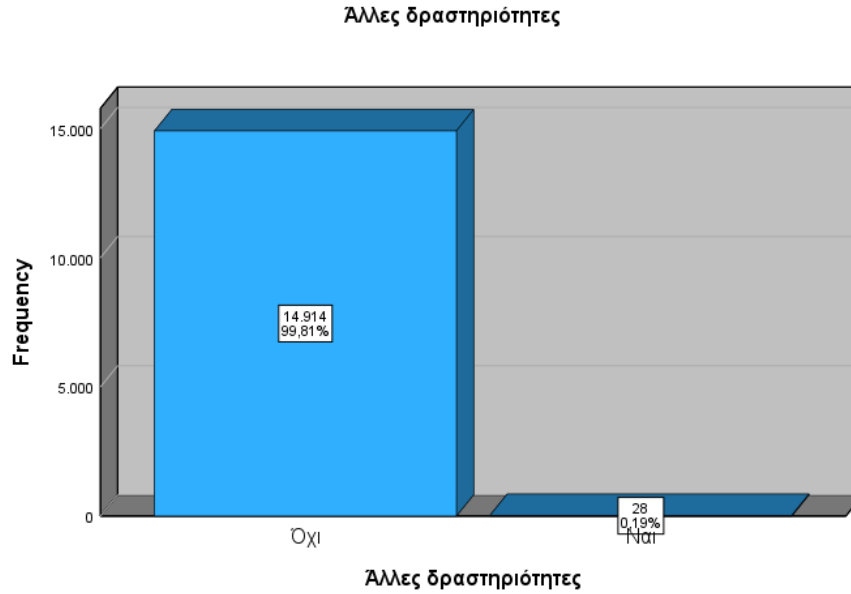
Αναφορικά με τη συμμετοχή σε άλλες δραστηριότητες, 14,914 άτομα, 99,8%, δεν έχουν εμπλακεί σε άλλες δραστηριότητες ενώ, μόλις 28 άτομα, 0,2% έχουν συμμετάσχει σε άλλες δραστηριότητες.

**Άλλες δραστηριότητες**

	N	%
Όχι	14914	99,8%
Ναι	28	0,2%

Πίνακας 22: Συχνότητα Συμμετοχής σε Άλλες Δραστηριότητες





Γράφημα 26: Συχνότητα Συμμετοχής σε Άλλες Δραστηριότητες

Σε ό,τι αφορά τον αριθμό επισκέψεων στην αθλητική πλατφόρμα από την ημερομηνία έναρξης έως την ημερομηνία διακοπής, σύμφωνα με την κατηγοριοποίηση Hughes, τα δεδομένα δείχνουν μια ομοιόμορφη κατανομή των συμμετεχόντων στις κατηγορίες. Κάθε μία από τις πέντε κατηγορίες περιλαμβάνει περίπου 2,989 άτομα, με καθεμία να αντιπροσωπεύει το 20% του συνόλου των συμμετεχόντων.

**Αριθμός επισκέψεων στην αθλητική πλατφόρμα από την ημερομηνία έναρξης και την ημερομηνία διακοπής Hughes**

	N	%
1	2989	20,0%
2	2989	20,0%
3	2988	20,0%
4	2988	20,0%
5	2988	20,0%

Πίνακας 23: Συχνότητα Επισκέψεων στην Αθλητική Πλατφόρμα από Έναρξη Έως Διακοπή Εγγραφής Hughes

Όσον αφορά τον αριθμό συμμετοχής σε δραστηριότητες από την ημερομηνία έναρξης έως την ημερομηνία διακοπής, σύμφωνα με την κατηγοριοποίηση Hughes, τα δεδομένα δείχνουν ένα ομοιόμορφο μοτίβο κατανομής. Κάθε κατηγορία, από την 1 έως την 5, περιλαμβάνει 2,989 άτομα, αντίστοιχα, και κάθε μία από αυτές αντιστοιχεί σε ποσοστό 20% του συνόλου των συμμετεχόντων.

**Αριθμός συμμετοχής σε δραστηριότητες από την ημερομηνία έναρξης και την ημερομηνία διακοπής Hughes**

	N	%
1	2989	20,0%
2	2989	20,0%
3	2988	20,0%
4	2988	20,0%
5	2988	20,0%

*Πίνακας 24: Συχνότητα Συμμετοχής σε Δραστηριότητες από Έναρξη Έως Διακοπή Εγγραφής Hughes*

Σχετικά με τον μέσο όρο του αριθμού συμμετοχής σε δραστηριότητες από την ημερομηνία έναρξης έως την ημερομηνία διακοπής, βάσει της κατηγοριοποίησης Hughes, τα δεδομένα εμφανίζουν μια πλήρως ομοιόμορφη κατανομή. Αναλυτικότερα, καθεμιά από τις πέντε κατηγορίες περιλαμβάνει 2,989 άτομα, με κάθε μία να αντιστοιχεί στο 20% του συνολικού αριθμού των συμμετεχόντων.

**Μέσος όρος αριθμού συμμετοχής σε δραστηριότητες από την ημερομηνία έναρξης και την ημερομηνία διακοπής Hughes**

	N	%
1	2989	20,0%
2	2989	20,0%
3	2988	20,0%

4	2988	20,0%
5	2988	20,0%

Πίνακας 25: Συχνότητα Μέσου Αριθμού Συμμετοχής σε Δραστηριότητες από Έναρξη Έως Διακοπή Εγγραφής Hughes

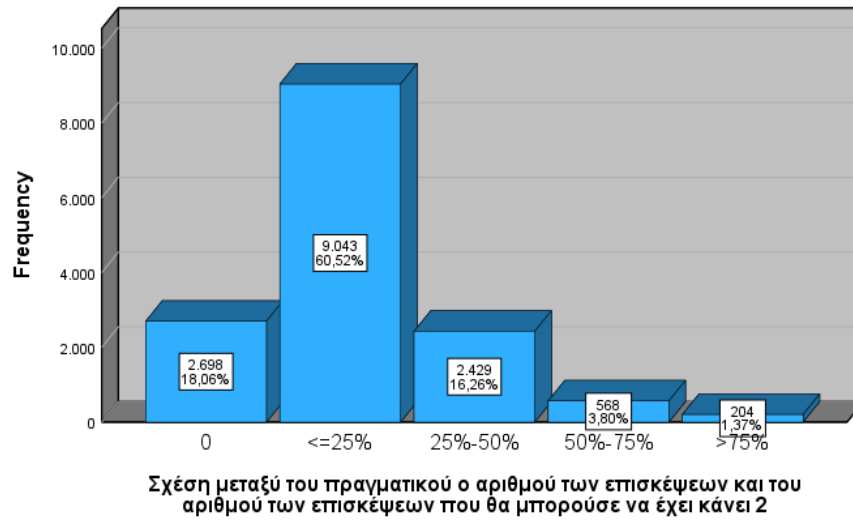
Όσον αφορά τη σχέση μεταξύ του πραγματικού αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει ένας χρήστης, 2,698 άτομα, 18,1%, δεν έχουν κάνει καμία επίσκεψη, ενώ η πλειοψηφία 9,043 άτομα 60,5%, έχει κάνει συμμετοχές που αντιστοιχούν σε ποσοστό μέχρι 25% των επισκέψεων που θα μπορούσαν να έχουν πραγματοποιήσει. Επιπλέον, 2,429 άτομα, 16,3%, έχουν πραγματοποιήσει το 25% έως 50% των δυνατών επισκέψεων. Μικρότερη αναλογία, 568 άτομα 3,8%, έχουν πραγματοποιήσει το 50% έως 75% των δυνατών επισκέψεων και τέλος, μόλις 204 άτομα, 1,4%, έχουν υπερβεί το 75% των δυνατών επισκέψεων.

**Σχέση μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει 2**

	N	%
0	2698	18,1%
<=25%	9043	60,5%
25%-50%	2429	16,3%
50%-75%	568	3,8%
>75%	204	1,4%

Πίνακας 26: Συχνότητα Σχέσης Πραγματικών Επισκέψεων προς Πιθανές Επισκέψεις (2)

Σχέση μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει 2



Γράφημα 27: Συχνότητα Σχέσης Πραγματικών Επισκέψεων προς Πιθανές Επισκέψεις (2)

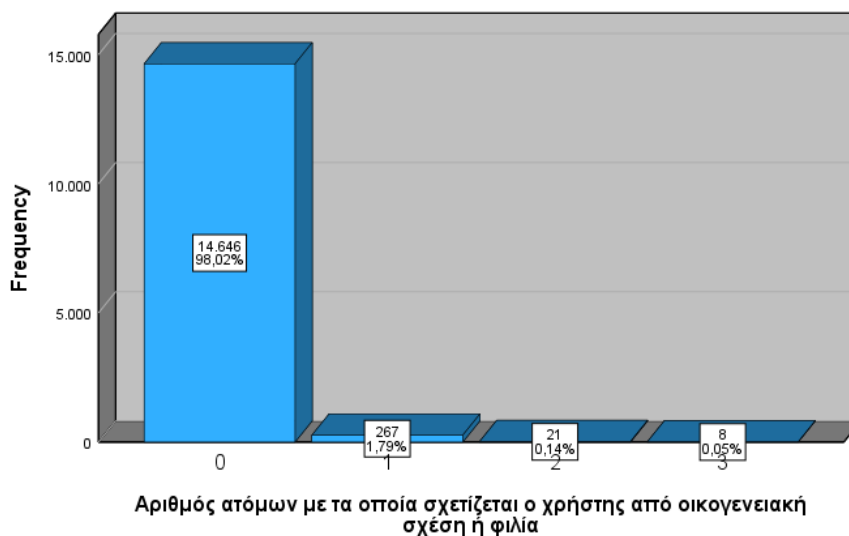
Σχετικά με τον αριθμό ατόμων με τα οποία σχετίζεται ο χρήστης από οικογενειακή σχέση ή φίλια, 14,646 άτομα, 98%, δεν έχουν κανέναν με τον οποίο να σχετίζονται από οικογενειακή σχέση ή φίλια, ενώ 267 άτομα, 1,8% έχουν ένα άτομο με το οποίο σχετίζονται, ενώ μικρότερο ποσοστό, 21 άτομα 0,1%, έχουν δύο τέτοιες σχέσεις. Επιπλέον, 8 άτομα 0,1%, έχουν τρεις σχέσεις οικογενειακές ή φιλικές εντός των αθλητικών πλατφορμών.

**Αριθμός ατόμων με τα οποία  
σχετίζεται ο χρήστης από οικογενειακή  
σχέση ή φίλια**

	N	%
0	14646	98,0%
1	267	1,8%
2	21	0,1%
3	8	0,1%

Πίνακας 27: Συχνότητα Αριθμού Ατόμων με Οικογενειακή ή Φιλική Σχέση με τον Χρήστη

Αριθμός ατόμων με τα οποία σχετίζεται ο χρήστης από οικογενειακή σχέση ή φίλια



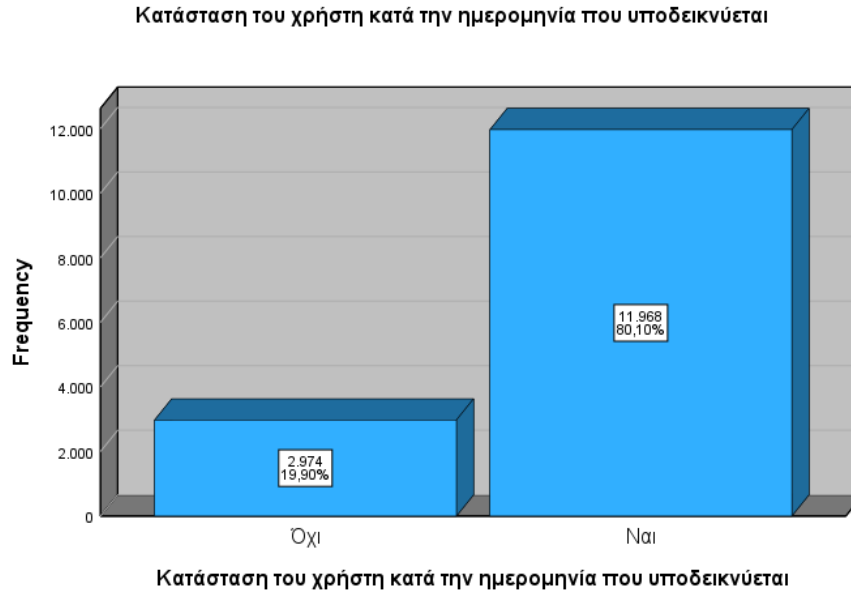
Γράφημα 28: Συχνότητα Αριθμού Ατόμων με Οικογενειακή ή Φιλική Σχέση με τον Χρήστη

Όσον αφορά την κατάσταση του χρήστη κατά την ημερομηνία που υποδεικνύεται, 2,974 άτομα 19,9%, δεν ήταν ενεργά, ενώ, 80,1%, 11,968 άτομα, ήταν.

**Κατάσταση του χρήστη κατά την  
ημερομηνία που υποδεικνύεται**

	N	%
Όχι	2974	19,9%
Ναι	11968	80,1%

Πίνακας 28: Συχνότητα Κατάστασης Χρήστη στην Υποδεικνυόμενη Ημερομηνία



Γράφημα 29: Συχνότητα Κατάστασης Χρήστη στην Υποδεικνυόμενη Ημερομηνία

## 4.2 Παλινδρόμηση

Αρχικά εφαρμόζεται γραμμικό πολλαπλό μοντέλο Παλινδρόμησης και εξαρτημένη μεταβλητή το συνολικό ποσό πληρωμής και ένα πλήθος ανεξάρτητων μεταβλητών που παρουσιάζονται στον επόμενο πίνακα.

### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,794 <sup>a</sup>	,630	,630	221,541

a. Predictors: (Constant), Άλλες δραστηριότητες, Στίβος, Ομαδικά αθλήματα, Αθλήματα Ρακέτας, Πολεμικές Τέχνες, Διάρκεια Εγγραφής, Ηλικία, Φύλο, Ειδικές δραστηριότητες, Δραστηριότητες στο νερό, Κολύμπι

Πίνακας 29: Περίληψη Μοντέλου Παλινδρόμησης Συνολικού ποσού πληρωμής

Το  $R$  είναι 0,794, δείχνοντας μια ισχυρή θετική σχέση μεταξύ των ανεξάρτητων μεταβλητών και του συνολικού ποσού πληρωμής. Το  $R^2$  είναι 0,630, υποδηλώνοντας ότι το 63% της διακύμανσης του συνολικού ποσού πληρωμής μπορεί να εξηγηθεί από τις ανεξάρτητες μεταβλητές του μοντέλου. Η στατιστική  $F$  είναι 2315,927 με  $p$ -value

< 0,001, υποδηλώνοντας ότι το μοντέλο είναι στατιστικά σημαντικό και οι ανεξάρτητες μεταβλητές έχουν σημαντική επίδραση στην εξαρτημένη μεταβλητή.

Συντελεστές

- **Διάρκεια Εγγραφής:** Έχει την ισχυρότερη θετική επίδραση ( $\beta=0,708$ ,  $t=134,767$ ) στο συνολικό ποσό πληρωμής, υποδηλώνοντας ότι όσο μεγαλύτερη είναι η διάρκεια της εγγραφής, τόσο μεγαλύτερο είναι το ποσό που πληρώνει ένα άτομο.
- **Δραστηριότητες στο Νερό και Αθλήματα Ρακέτας:** Έχουν επίσης σημαντική θετική επίδραση στο συνολικό ποσό πληρωμής.
- **Κολύμπι:** Αξιοσημείωτο είναι το γεγονός ότι το κολύμπι έχει αρνητική επίδραση ( $-\beta=0,091$ ), υποδηλώνοντας ότι οι συμμετέχοντες στο κολύμπι τείνουν να πληρώνουν λιγότερο συνολικά.
- **Φύλο:** Έχει ελαφρώς αρνητική επίδραση, υποδηλώνοντας μικρές διαφορές στο ποσό πληρωμής με βάση το φύλο.
- **Πολεμικές Τέχνες και Άλλες Δραστηριότητες:** Δεν έχουν στατιστικά σημαντική επίδραση στο συνολικό ποσό πληρωμής.

ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1250330853,798	11	113666441,254	2315,927	<,001 <sup>b</sup>
	Residual	732769338,865	14930	49080,331		
	Total	1983100192,663	14941			

a. Dependent Variable: Συνολικό ποσό πληρωμής

b. Predictors: (Constant), Άλλες δραστηριότητες, Στίβος, Ομαδικά αθλήματα, Αθλήματα Ρακέτας, Πολεμικές Τέχνες, Διάρκεια Εγγραφής, Ηλικία, Φύλο, Ειδικές δραστηριότητες, Δραστηριότητες στο νερό, Κολύμπι

Πίνακας 30: Πίνακας ANOVA Συνολικού ποσού πληρωμής

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	60,489	9,607		6,296	<,001
	Ηλικία	,006	,137	,000	,041	,968
	Φύλο	-10,723	3,760	-,014	-2,852	,004
	Διάρκεια Εγγραφής	18,210	,135	,708	134,767	<,001
	Στίβος	-71,631	21,645	-,017	-3,309	<,001
	Δραστηριότητες στο νερό	125,624	6,929	,158	18,130	<,001
	Κολύμπι	-67,242	6,727	-,091	-9,996	<,001
	Ομαδικά αθλήματα	45,252	9,429	,028	4,799	<,001
	Αθλήματα Ρακέτας	73,288	13,033	,030	5,623	<,001
	Πολεμικές Τέχνες	-2,265	7,426	-,002	-,305	,760
	Ειδικές δραστηριότητες	102,303	12,052	,045	8,489	<,001
	Άλλες δραστηριότητες	-14,311	42,459	-,002	-,337	,736

a. Dependent Variable: Συνολικό ποσό πληρωμής

Πίνακας 31: Πίνακας Συντελεστών Συνολικού ποσού πληρωμής

Η διάρκεια εγγραφής είναι ο πιο κρίσιμος παράγοντας που επηρεάζει το συνολικό ποσό πληρωμής στις αθλητικές πλατφόρμες, ενώ ορισμένες συγκεκριμένες δραστηριότητες, όπως οι δραστηριότητες στο νερό και τα αθλήματα ρακέτας, επίσης συμβάλλουν θετικά. Ενδιαφέρον είναι ότι το κολύμπι, παρά την θετική συσχέτιση με τις δραστηριότητες στο νερό, έχει αρνητική επίδραση, πιθανώς λόγω διαφορετικών τιμολογιακών πολιτικών ή προσφορών για τη συγκεκριμένη δραστηριότητα. Οι



πολεμικές τέχνες και οι άλλες δραστηριότητες δεν φαίνεται να έχουν σημαντική επίδραση στο συνολικό ποσό πληρωμής, που μπορεί να αντανακλά την λιγότερη δημοτικότητα αυτών των δραστηριοτήτων.

Ο συντελεστής συσχέτισης ( $R$ ) είναι 0,292, υποδηλώνοντας μια ασθενή θετική σχέση μεταξύ των ανεξάρτητων και της εξαρτημένης μεταβλητής. Το  $R^2$  είναι 0,085, που σημαίνει ότι μόνο το 8,5% της διακύμανσης της μέσης εβδομαδιαίας συχνότητας μπορεί να εξηγηθεί από τις επιλεγμένες ανεξάρτητες μεταβλητές.

Η στατιστική  $F$  είναι 126,525 με  $p$ -value <0,001, υποδηλώνοντας ότι το μοντέλο είναι στατιστικά σημαντικό, παρόλο που η εξήγηση της διακύμανσης είναι σχετικά χαμηλή.

#### Συντελεστές

- **Ηλικία:** Η ηλικία έχει θετική σχέση με τη μέση εβδομαδιαία συχνότητα, αλλά η επιρροή της είναι μικρή.
- **Φύλο:** Το φύλο έχει αρνητική σχέση με τη μέση εβδομαδιαία συχνότητα, δηλαδή ένα φύλο μπορεί να είναι λιγότερο ενεργό από το άλλο.
- **Διάρκεια Εγγραφής:** Έχει ελαφρώς αρνητική επίδραση, υποδηλώνοντας ότι μεγαλύτερη διάρκεια εγγραφής μπορεί να μην οδηγεί σε υψηλότερη συχνότητα συμμετοχής.
- **Ειδικές δραστηριότητες:** Έχουν σημαντική θετική επίδραση στη μέση εβδομαδιαία συχνότητα, ενώ οι άλλες δραστηριότητες έχουν αρνητική επίδραση.
- **Κολύμπι και Δραστηριότητες στο Νερό:** Και οι δύο έχουν θετική σχέση με τη μέση εβδομαδιαία συχνότητα, υποδηλώνοντας ότι οι συμμετέχοντες σε αυτές τις δραστηριότητες τείνουν να είναι πιο ενεργοί.
- **Ομαδικά αθλήματα και Πολεμικές Τέχνες:** Έχουν αρνητική σχέση με τη μέση εβδομαδιαία συχνότητα, υποδηλώνοντας μικρότερη ενεργητικότητα από τους συμμετέχοντες σε αυτές τις δραστηριότητες.

Στη συνέχεια εφαρμόζεται γραμμικό πολλαπλό μοντέλο Παλινδρόμησης και εξαρτημένη μεταβλητή την εβδομαδιαία συχνότητα και ένα πλήθος ανεξάρτητων μεταβλητών που παρουσιάζονται στον επόμενο πίνακα.

### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,292 <sup>a</sup>	,085	,085	,639

a. Predictors: (Constant), Άλλες δραστηριότητες, Στίβος, Ομαδικά αθλήματα, Αθλήματα Ρακέτας, Πολεμικές Τέχνες, Διάρκεια Εγγραφής, Ηλικία, Φύλο, Ειδικές δραστηριότητες, Δραστηριότητες στο νερό, Κολύμπι

Πίνακας 32: Περίληψη Μοντέλου Παλινδρόμησης Μέσης εβδομαδιαίας συχνότητας σε όλη την περίοδο εγγραφής

### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	567,668	11	51,606	126,525	<,001 <sup>b</sup>
	Residual	6089,565	14930	,408		
	Total	6657,233	14941			

a. Dependent Variable: Μέση εβδομαδιαία συχνότητα σε όλη την περίοδο εγγραφής

b. Predictors: (Constant), Άλλες δραστηριότητες, Στίβος, Ομαδικά αθλήματα, Αθλήματα Ρακέτας, Πολεμικές Τέχνες, Διάρκεια Εγγραφής, Ηλικία, Φύλο, Ειδικές δραστηριότητες, Δραστηριότητες στο νερό, Κολύμπι

Πίνακας 33: Πίνακας ANOVA Μέσης εβδομαδιαίας συχνότητας σε όλη την περίοδο εγγραφής

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,735	,028		26,555	<,001
	Ηλικία	,005	,000	,116	13,892	<,001
	Φύλο	-,124	,011	-,091	-11,414	<,001

Διάρκεια Εγγραφής	-,004	,000	-,090	-10,936	<,001
Στίβος	,065	,062	,008	1,046	,296
Δραστηριότητες στο νερό	,127	,020	,087	6,354	<,001
Κολύμπι	,201	,019	,149	10,375	<,001
Ομαδικά αθλήματα	-,174	,027	-,060	-6,398	<,001
Αθλήματα Ρακέτας	-,067	,038	-,015	-1,774	,076
Πολεμικές Τέχνες	-,179	,021	-,083	-8,346	<,001
Ειδικές δραστηριότητες	,484	,035	,117	13,924	<,001
Άλλες δραστηριότητες	-,465	,122	-,030	-3,802	<,001

a. Dependent Variable: Μέση εβδομαδιαία συχνότητα σε όλη την περίοδο εγγραφής

Πίνακας 34: Πίνακας Συντελεστών Μέσης εβδομαδιαίας συχνότητας σε όλη την περίοδο εγγραφής

Το μοντέλο υποδηλώνει ότι η ηλικία, ορισμένες ειδικές δραστηριότητες, το κολύμπι και άλλες δραστηριότητες στο νερό και το φύλο έχουν σημαντική επίδραση στη μέση εβδομαδιαία συχνότητα συμμετοχής σε δραστηριότητες. Ωστόσο, οι συνολικές επιδράσεις είναι σχετικά μικρές ( $R^2 = 0,085$ ) υποδηλώνοντας ότι άλλοι παράγοντες που δεν εξετάστηκαν στο μοντέλο μπορεί να επηρεάζουν επίσης τη συχνότητα συμμετοχής.

Ο συντελεστής συσχέτισης  $R$  είναι 0,255, δείχνοντας μια σχετικά χαμηλή θετική σχέση μεταξύ των μεταβλητών. Το  $R^2$  είναι 0,065 υποδηλώνοντας ότι μόνο το 6,5% της διακύμανσης της εξαρτημένης μεταβλητής εξηγείται από το μοντέλο.

Η στατιστική  $F$  είναι 94,009 με  $p$ -value <0,001, υποδηλώνοντας ότι το μοντέλο είναι στατιστικά σημαντικό.

#### Συντελεστές

- **Ηλικία:** Έχει αρνητική επίδραση στην εξαρτημένη μεταβλητή, δηλαδή μεγαλύτερη ηλικία σχετίζεται με χαμηλότερη σχέση πραγματικών προς δυνατές συμμετοχές.
- **Φύλο:** Επίσης έχει αρνητική επίδραση, υποδηλώνοντας ότι ίσως υπάρχουν διαφορές στη συμπεριφορά συμμετοχής μεταξύ των φύλων.

- **Διάρκεια Εγγραφής:** Θετική επίδραση, υποδηλώνοντας ότι μακρύτερες περίοδοι εγγραφής συσχετίζονται με υψηλότερη σχέση πραγματικών προς τις εν δυνάμει συμμετοχές.
- **Δραστηριότητες στο Νερό και Ειδικές Δραστηριότητες:** Έχουν θετική επίδραση, υποδηλώνοντας ότι οι συμμετέχοντες σε αυτές τις δραστηριότητες είναι πιο πιθανό να εκμεταλλεύονται πλήρως τις δυνατότητες συμμετοχής.
- **Ομαδικά Αθλήματα και Πολεμικές Τέχνες:** Έχουν αρνητική επίδραση, υποδηλώνοντας ότι η συμμετοχή σε αυτές τις δραστηριότητες μπορεί να συνδέεται με χαμηλότερη σχέση πραγματικών προς τις εν δυνάμει συμμετοχές.

Στη συνέχεια εφαρμόζεται γραμμικό πολλαπλό μοντέλο Παλινδρόμησης και εξαρτημένη μεταβλητή που σχετίζεται με τον αριθμό επισκέψεων και ένα πλήθος ανεξάρτητων μεταβλητών που παρουσιάζονται στον επόμενο πίνακα.

#### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,255 <sup>a</sup>	,065	,064	16,968

a. Predictors: (Constant), Άλλες δραστηριότητες, Στίβος, Ομαδικά αθλήματα, Αθλήματα Ρακέτας, Πολεμικές Τέχνες, Διάρκεια Εγγραφής, Ηλικία, Φύλο, Ειδικές δραστηριότητες, Δραστηριότητες στο νερό, Κολύμπι

*Πίνακας 35: Περίληψη Μοντέλου Παλινδρόμησης Σχέσης μεταξύ του πραγματικού αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει*

#### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	297716,958	11	27065,178	94,009	<,001 <sup>b</sup>
	Residual	4298345,496	14930	287,900		
	Total	4596062,453	14941			

a. Dependent Variable: Σχέση μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει

b. Predictors: (Constant), Άλλες δραστηριότητες, Στίβος, Ομαδικά αθλήματα, Αθλήματα Ρακέτας, Πολεμικές Τέχνες, Διάρκεια Εγγραφής, Ηλικία, Φύλο, Ειδικές δραστηριότητες, Δραστηριότητες στο νερό, Κολύμπι

Πίνακας 36: Πίνακας ANOVA Σχέσης μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	18,904	,736		25,691	<,001
	Ηλικία	-,052	,010	-,042	-4,928	<,001
	Φύλο	-1,850	,288	-,052	-6,426	<,001
	Διάρκεια Εγγραφής	,090	,010	,072	8,657	<,001
	Στίβος	-2,610	1,658	-,013	-1,574	,115
	Δραστηριότητες στο νερό	3,874	,531	,101	7,301	<,001
	Κολύμπι	-,944	,515	-,027	-1,833	,067
	Ομαδικά αθλήματα	-6,498	,722	-,085	-8,998	<,001
	Αθλήματα Ρακέτας	6,414	,998	,055	6,426	<,001
	Πολεμικές Τέχνες	-7,940	,569	-,140	-13,961	<,001
	Ειδικές δραστηριότητες	8,959	,923	,082	9,706	<,001
	Άλλες δραστηριότητες	-,365	3,252	-,001	-,112	,911

a. Dependent Variable: Σχέση μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει

Πίνακας 37: Πίνακας Συντελεστών Σχέσης μεταξύ του πραγματικού ο αριθμού των επισκέψεων και του αριθμού των επισκέψεων που θα μπορούσε να έχει κάνει

Το μοντέλο δείχνει ότι ορισμένες δραστηριότητες και παράγοντες, όπως η διάρκεια εγγραφής, οι δραστηριότητες στο νερό και ειδικές δραστηριότητες, έχουν θετική σχέση με την αξιοποίηση των εν δυνάμει επισκέψεων από τους χρήστες. Ωστόσο, παράγοντες όπως η ηλικία και το φύλο μπορεί να επηρεάσουν αρνητικά αυτή τη σχέση. Η συνολική επεξήγηση της διακύμανσης της εξαρτημένης μεταβλητής από το μοντέλο είναι

σχετικά χαμηλή ( $R^2 = 0,065$ ), υποδηλώνοντας ότι άλλοι παράγοντες μπορεί επίσης να επηρεάζουν τη σχέση μεταξύ του πραγματικού και των εν δυνάμει συμμετοχών.

#### 4.3 ANOVA

Για την εκτίμηση της επίδρασης της ηλικιακής ομάδας στο συνολικό ποσό πληρωμής, οι μεταβλητές που συμμετέχουν στο μοντέλο είναι η Ηλικιακή Ομάδα και το Συνολικό Ποσό Πληρωμής. Η ηλικιακή ομάδα έχει σημαντική επίδραση στο συνολικό ποσό πληρωμής, με τις διαφορές μεταξύ των ομάδων να είναι στατιστικά σημαντικές. Αυτό υποδηλώνει ότι η ηλικία μπορεί να επηρεάζει την ποσότητα των χρημάτων που δαπανούνται σε αθλητικές δραστηριότητες ή υπηρεσίες.

#### ANOVA

Συνολικό ποσό πληρωμής

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	203448451,71	4	50862112,929	426,897	<,001
6					
Within Groups	1779651740,9	14937	119143,854		
46					
Total	1983100192,6	14941			
63					

Πίνακας 38: Επίδραση Ηλικιακής Ομάδας στο Συνολικό Ποσό Πληρωμής

Για την αξιολόγηση της σημασίας της Διάρκειας Συνδρομής ως παράγοντα που επηρεάζει το συνολικό ποσό πληρωμής, οι σχετικές μεταβλητές είναι η Διάρκεια Συνδρομής και το Συνολικό Ποσό Πληρωμής. Η διάρκεια συνδρομής έχει σημαντική επίδραση στο συνολικό ποσό πληρωμής, με τη διαφορά μεταξύ των ομάδων να είναι στατιστικά σημαντική. Αυτό υποδηλώνει ότι η περίοδος συνδρομής ενός χρήστη μπορεί να επηρεάζει το πόσο χρήματα ξοδεύει σε σχετικές υπηρεσίες ή δραστηριότητες.

## ANOVA

Συνολικό ποσό πληρωμής

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	13480209,264	1	13480209,264	102,250	<,001
Within Groups	1969619983,399	14940	131835,340		
Total	1983100192,663	14941			

Πίνακας 39: Επίδραση Διάρκειας Συνδρομής στο Συνολικό Ποσό Πληρωμής

Για την αξιολόγηση του ρόλου της Ηλικιακής Ομάδας στην Μέση Εβδομαδιαία Συχνότητα Συμμετοχής σε δραστηριότητες, οι μεταβλητές που περιλαμβάνονται στην ανάλυση είναι η Ηλικιακή Ομάδα και η Μέση Εβδομαδιαία Συχνότητα. Η ηλικιακή ομάδα έχει σημαντική επίδραση στη μέση εβδομαδιαία συχνότητα συμμετοχής σε δραστηριότητες κατά την περίοδο της συνδρομής, με τη διαφορά μεταξύ των ομάδων να είναι στατιστικά σημαντική.

## ANOVA

Μέση εβδομαδιαία συχνότητα σε όλη την περίοδο εγγραφής

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	189,079	4	47,270	109,161	<,001
Within Groups	6468,154	14937	,433		
Total	6657,233	14941			

Πίνακας 40: Επίδραση Ηλικιακής Ομάδας στη Μέση Εβδομαδιαία Συχνότητα Συμμετοχής

Για την αξιολόγηση του ρόλου της Ηλικιακής Ομάδας στην Μέση Εβδομαδιαία Συχνότητα Συμμετοχής σε δραστηριότητες, οι μεταβλητές που περιλαμβάνονται στην ανάλυση είναι η Ηλικιακή Ομάδα και η Μέση Εβδομαδιαία Συχνότητα. Η διάρκεια συνδρομής έχει σημαντική επίδραση στη μέση εβδομαδιαία συχνότητα συμμετοχής σε δραστηριότητες κατά την περίοδο εγγραφής, με τη διαφορά μεταξύ των τάξεων της διάρκειας συνδρομής να είναι στατιστικά σημαντική.

## ANOVA

Μέση εβδομαδιαία συχνότητα σε όλη την περίοδο εγγραφής

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	193,895	5	38,779	89,614	<,001
Within Groups	6463,338	14936	,433		
Total	6657,233	14941			

Πίνακας 41: Επίδραση Διάρκειας Συνδρομής στη Μέση Εβδομαδιαία Συχνότητα Συμμετοχής

## 4.4 T-test

Στην εφαρμογή του t-test για να εκτιμηθεί η διαφορά στη μέση εβδομαδιαία συχνότητα συμμετοχής σε δραστηριότητες μεταξύ των δύο φύλων, οι εμπλεκόμενες μεταβλητές είναι το Φύλο (άντρας, γυναίκα) και η Μέση Εβδομαδιαία Συχνότητα συμμετοχής. Υπάρχει στατιστικά σημαντική διαφορά στη μέση εβδομαδιαία συχνότητα συμμετοχής σε δραστηριότητες μεταξύ ανδρών και γυναικών. Οι άνδρες έχουν υψηλότερη μέση συχνότητα (0,80) σε σύγκριση με τις γυναίκες (0,73), με τη διαφορά αυτή να είναι στατιστικά σημαντική.

### Group Statistics

	Φύλο	N	Mean	Std. Deviation	Std. Error Mean
Μέση εβδομαδιαία συχνότητα σε όλη την περίοδο εγγραφής	Άντρας	6011	,80	,705	,009
	Γυναίκα	8931	,73	,640	,007

Πίνακας 42: Στατιστικά Στοιχεία ανά Φύλο για Μέση Εβδομαδιαία Συχνότητα

Για την ανίχνευση στατιστικά σημαντικών διαφορών στο συνολικό ποσό πληρωμής μεταξύ των φύλων, οι μεταβλητές που λαμβάνονται υπόψη στην ανάλυση είναι πάλι το Φύλο και το Συνολικό Ποσό Πληρωμής. Η ανάλυση δείχνει στατιστικά σημαντικές διαφορές στο συνολικό ποσό πληρωμής μεταξύ ανδρών και γυναικών, με τους άνδρες να ξοδεύουν κατά μέσο όρο περισσότερα (339,17) σε σύγκριση με τις γυναίκες (277,92).



### Group Statistics

	Φύλο	N	Mean	Std. Deviation	Std. Error Mean
Συνολικό ποσό πληρωμής	Άντρας	6011	339,17	392,022	5,056
	Γυναίκα	8931	277,92	342,246	3,622

Πίνακας 5: Στατιστικά Στοιχεία Φύλου για Συνολικό Ποσό Πληρωμής

### 4.5 Συντελεστής Spearman

- **Ηλικία και Διάρκεια Εγγραφής:** Η συσχέτιση είναι πολύ χαμηλή ( $\rho=,003$ ) και δεν είναι στατιστικά σημαντική ( $p=,747$ ), υποδηλώνοντας ότι δεν υπάρχει ουσιαστική σχέση μεταξύ της ηλικίας των ατόμων και της διάρκειας της συνδρομής τους.
- **Ηλικία και Συνολικό Ποσό Πληρωμής:** Υπάρχει αρνητική συσχέτιση ( $\rho = -0,022$ ) η οποία είναι στατιστικά σημαντική ( $p = 0,009$ ), υποδηλώνοντας μια ελαφρά τάση όπου μεγαλύτερη ηλικία συσχετίζεται με ελαφρώς χαμηλότερο συνολικό ποσό πληρωμής.
- **Διάρκεια Συνδρομής και Συνολικό Ποσό Πληρωμής:** Υπάρχει ισχυρή θετική συσχέτιση ( $\rho = 0,796$ ), που είναι στατιστικά σημαντική ( $p < 0,001$ ), υποδηλώνοντας ότι μεγαλύτερη διάρκεια συνδρομής συσχετίζεται με υψηλότερο συνολικό ποσό πληρωμής.
- **Αριθμός Δραστηριοτήτων και άλλες μεταβλητές:** Ο αριθμός των δραστηριοτήτων που συμμετέχει ένα άτομο έχει θετική συσχέτιση με την ηλικία ( $\rho=0,079$  και  $p<0,001$ ), τη διάρκεια εγγραφής ( $\rho=0,208$  και  $p<0,001$ ), και το συνολικό ποσό πληρωμής ( $\rho=0,203$  και  $p<0,001$ ). Αυτό υποδηλώνει ότι οι άνθρωποι που συμμετέχουν σε περισσότερες δραστηριότητες τείνουν να είναι ελαφρώς μεγαλύτεροι σε ηλικία, να έχουν μεγαλύτερη διάρκεια εγγραφής, και να ξοδεύουν περισσότερα συνολικά.

### Correlations

	Ηλικία	Διάρκεια Εγγραφής	Συνολικό ποσό πληρωμής	Αριθμός δραστηριοτήτων
Ηλικία	1			
Διάρκεια Εγγραφής		1		
Συνολικό ποσό πληρωμής			1	
Αριθμός δραστηριοτήτων				1

Spearman's rho	Ηλικία	Correlation Coefficient	1,000	,003	-,022**	,079**
		Sig. (2-tailed)	.	,747	,009	<,001
		N	14942	14942	14942	14942
Διάρκεια Εγγραφής		Correlation Coefficient	,003	1,000	,796**	,208**
		Sig. (2-tailed)	,747	.	<,001	<,001
		N	14942	14942	14942	14942
Συνολικό-ποσό πληρωμής		Correlation Coefficient	-,022**	,796**	1,000	,203**
		Sig. (2-tailed)	,009	<,001	.	<,001
		N	14942	14942	14942	14942
Αριθμός δραστηριοτήτων		Correlation Coefficient	,079**	,208**	,203**	1,000
		Sig. (2-tailed)	<,001	<,001	<,001	.
		N	14942	14942	14942	14942

\*\* . Correlation is significant at the 0.01 level (2-tailed).

Πίνακας 44: Συσχετίσεις Spearman

## Συμπεράσματα

Η παρούσα μελέτη αποσκοπούσε στην εξέταση διαφόρων παραγόντων που ενδέχεται να επηρεάζουν τη συχνότητα συμμετοχής σε δραστηριότητες και το συνολικό ποσό πληρωμής από τους χρήστες σε μία πλατφόρμα αθλητικών εγκαταστάσεων. Η ανάλυση επικεντρώθηκε σε δημογραφικά χαρακτηριστικά όπως η ηλικία και το φύλο, καθώς και σε παράγοντες όπως η διάρκεια εγγραφής και οι επιλεγμένες δραστηριότητες.

Από τα αποτελέσματα προέκυψε ότι υπάρχουν σημαντικές διαφορές στη συχνότητα συμμετοχής και το συνολικό ποσό πληρωμής με βάση την ηλικία και το φύλο των χρηστών. Συγκεκριμένα, διαπιστώθηκε ότι οι άνδρες συμμετέχουν συχνότερα και

ξοδεύουν περισσότερο σε σύγκριση με τις γυναίκες. Επιπρόσθετα, η διάρκεια εγγραφής αποδείχθηκε ένας σημαντικός παράγοντας, καθώς μακροχρόνιοι συνδρομητές τείνουν να έχουν υψηλότερη συχνότητα συμμετοχής και μεγαλύτερο συνολικό ποσό πληρωμής.

Επιπροσθέτως, οι επιλεγμένες δραστηριότητες είχαν σημαντική επίδραση, με τις δραστηριότητες στο νερό και τις ειδικές δραστηριότητες να συσχετίζονται με αυξημένη συχνότητα και συνολικό ποσό πληρωμής, ενώ ορισμένες άλλες κατηγορίες όπως τα ομαδικά αθλήματα και οι πολεμικές τέχνες φάνηκαν να έχουν αντίθετη επίδραση.

Συνοψίζοντας, τα ευρήματα της έρευνας υποδεικνύουν ότι τα δημογραφικά χαρακτηριστικά και οι προτιμήσεις δραστηριοτήτων παίζουν σημαντικό ρόλο στην προσέλκυση και διατήρηση των πελατών σε μια πλατφόρμα αθλητικών εγκαταστάσεων. Οι πληροφορίες αυτές μπορούν να αξιοποιηθούν για την ανάπτυξη στοχευμένων στρατηγικών μάρκετινγκ και υπηρεσιών που θα απευθύνονται στις ειδικές ανάγκες και προτιμήσεις διαφορετικών ηλικιακών και φυλετικών ομάδων, με στόχο τη βελτίωση της πελατειακής εμπειρίας και της οικονομικής απόδοσης των εγκαταστάσεων.

## Βιβλιογραφία

### Ελληνική

Κούτρας, Μ. & Δαμιανού, Χ. (2000). Εισαγωγή στη Στατιστική (Μέρος 1ο), Εκδόσεις Συμμετρία, Αθήνα.

Σαχλάς, Α. & Μπερσίμης (2018). Εφαρμοσμένη Στατιστική με Έμφαση στις Επιστήμες Υγείας, Εκδόσεις Τζιόλα, Αθήνα.

Γιώργος Κ. Παπαδόπουλος (2015). Εισαγωγή στις Πιθανότητες και στη Στατιστική, Εκδόσεις Gutenberg, Αθήνα.

### Ξενόγλωσση

Ajibade, S. S., & Adediran, A. (2016). An overview of big data visualization techniques in data mining. *International Journal of Computer Science and Information Technology Research*, 4(3), 105-113.

Arifin, S. R. M. (2018). Ethical considerations in qualitative study. *International journal of care scholars*, 1(2), 30-33.

Asadoorian, M. O., & Kantarelis, D. (2005). *Essentials of inferential statistics*. University Press of America.

Brodlić, K., Allendes Osorio, R., & Lopes, A. (2012). A review of uncertainty in data visualization. *Expanding the frontiers of visual analytics and visualization*, 81-109.

Bulmer, M.G. (1979). *Principles of Statistics*. Dover Publications.

Cacciattolo, M. (2015). Ethical considerations in research. In *The Praxis of English Language Teaching and Learning (PELT)* (pp. 55-73). Brill.

Chapman, C. H., Nazem-Zadeh, M., Lee, O. E., Schipper, M. J., Tsien, C. I., Lawrence, T. S., & Cao, Y. (2013). Regional variation in brain white matter diffusion index changes following chemoradiotherapy: a prospective study using tract-based spatial statistics. *PLoS One*, 8(3), e57768.

Chatterjee, S., & Hadi, A. S. (2013). *Regression analysis by example*. John Wiley & Sons.

- Cleveland, W. S., McGill, M. E., & McGill, R. (1988). The shape parameter of a two-variable graph. *Journal of the American Statistical Association*, 83(402), 289-300.
- Conover, W. J. (1999). *Practical nonparametric statistics* (Vol. 350). John Wiley & Sons.
- Conover, W.J. (1999). *Practical Nonparametric Statistics*. John Wiley & Sons.
- Corder, G.W., & Foreman, D.I. (2009). *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*. John Wiley & Sons.
- Cramer, D. (2003). *Advanced Quantitative Data Analysis*. McGraw-Hill Education.
- Cramer, D. (2003). *Fundamental statistics for social research: step-by-step calculations and computer techniques using SPSS for Windows*. Routledge.
- DeGroot, M.H., & Schervish, M.J. (2012). *Στατιστική Μεθοδολογία* (4η Έκδοση). Pearson.
- Fisher, M. J., & Marshall, A. P. (2009). Understanding descriptive statistics. *Australian critical care*, 22(2), 93-97.
- Fisher, R. A. (1925, July). Theory of statistical estimation. In *Mathematical proceedings of the Cambridge philosophical society* (Vol. 22, No. 5, pp. 700-725). Cambridge University Press.
- Fisher, R.A. (1918). "The Correlation Between Relatives on the Supposition of Mendelian Inheritance". *Transactions of the Royal Society of Edinburgh*.
- Freund, R. J., Wilson, W. J., & Sa, P. (2006). *Regression analysis*. Elsevier.
- George, D., & Mallery, P. (2018). Descriptive statistics. In *IBM SPSS Statistics 25 Step by Step* (pp. 126-134). Routledge.
- Gogtay, N. J., & Thatte, U. M. (2017). Principles of correlation analysis. *Journal of the Association of Physicians of India*, 65(3), 78-81.
- Gosset, W.S. (1908). "The Probable Error of a Mean". *Biometrika*.
- Hauke, J., & Kossowski, T. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones geographicae*, 30(2), 87-93.

- Kaur, P., Sharma, M., & Mittal, M. (2018). Big data and machine learning based secure healthcare framework. *Procedia computer science*, *132*, 1049-1059.
- Marshall, G., & Jonker, L. (2011). An introduction to inferential statistics: A review and practical guide. *Radiography*, *17*(1), e1-e6.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley & Sons.
- Moore, D.S., & McCabe, G.P. (2003). *Introduction to the Practice of Statistics*. W.H. Freeman and Company.
- Nick, T. G. (2007). Descriptive statistics. *Topics in biostatistics*, 33-52.
- Pinheiro, P., & Cavique, L. (2021). Regular sports services: Dataset of demographic, frequency and service level agreement. *Data in Brief*, *36*, 107054.
- Post, J. E., Preston, L. E., & Sachs, S. (2002). Managing the extended enterprise: The new stakeholder view. *California management review*, *45*(1), 6-28.
- Sahu, P. K., Pal, S. R., & Das, A. K. (2015). *Estimation and inferential statistics*. New Delhi: Springer.
- Saltz, J. S., & Dewar, N. (2019). Data science ethical considerations: a systematic literature review and proposed project framework. *Ethics and Information Technology*, *21*, 197-208.
- Sarstedt, M., Hair Jr, J. F., Cheah, J. H., Becker, J. M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian marketing journal*, *27*(3), 197-211.
- Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, *66*(3), 605-610.
- Senthilnathan, S. (2019). Usefulness of correlation analysis. *Available at SSRN 3416918*.
- Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2, pp. 131-160).
- Walker, M. U. (2007). *Moral understandings: A feminist study in ethics*. Oxford University Press.

Westfall, P. H. (2014). Kurtosis as peakedness, 1905–2014. RIP. *The American Statistician*, 68(3), 191-195.

Zhao, S., Lin, Q., Ran, J., Musa, S. S., Yang, G., Wang, W., ... & Wang, M. H. (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International journal of infectious diseases*, 92, 214-217.