



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ –
ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ
ΤΟΜΕΑΣ ΕΡΓΩΝ ΥΠΟΔΟΜΗΣ ΚΑΙ ΑΓΡΟΤΙΚΗΣ ΑΝΑΠΤΥΞΗΣ

Διπλωματική Εργασία

**«Εφαρμογή αλγορίθμων συσταδοποίησης στην
αναγνώριση χώρο-χρονικών προτύπων επιλογής
μέσου μετακίνησης.
Η περίπτωση της Αθήνας. »**

Ειρήνη Μαρία Ανδρινοπούλου

*Επιβλέπων: Κωνσταντίνος Κεπαπτσόγλου,
Αναπληρωτής Καθηγητής ΕΜΠ*

ΕΡΓΑΣΤΗΡΙΟ
ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΤΕΧΝΙΚΗΣ



Αθήνα, Μάρτιος 2024



NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF RURAL, SURVEYING AND GEOMATICS
ENGINEERING
DEP. OF INFRASTRUCTURE AND RURAL DEVELOPMENT

Diploma Thesis

**«Application of clustering algorithms in identifying
spatiotemporal patterns of transportation mode
choice.
The case of Athens. »**

Eirini Maria Andrinopoulou

*Supervisor: Konstantinos Kepaptsoglou,
Associate Professor NTUA*

LABORATORY OF
TRANSPORTATION ENGINEERING



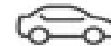
Athens, March 2024



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ –
ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ
ΤΟΜΕΑΣ ΕΡΓΩΝ ΥΠΟΔΟΜΗΣ ΚΑΙ ΑΓΡΟΤΙΚΗΣ ΑΝΑΠΤΥΞΗΣ

Διπλωματική Εργασία

**«Εφαρμογή αλγορίθμων συσταδοποίησης στην
αναγνώριση χώρο-χρονικών προτύπων επιλογής
μέσου μετακίνησης.
Η περίπτωση της Αθήνας.»**



Ειρήνη Μαρία Ανδρινοπούλου

*Επιβλέπων: Κωνσταντίνος Κεπαπτσόγλου,
Αναπληρωτής Καθηγητής ΕΜΠ*

ΕΡΓΑΣΤΗΡΙΟ
ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΤΕΧΝΙΚΗΣ



Αθήνα, Μάρτιος 2024

ΔΗΛΩΣΗ ΑΝΑΓΝΩΡΙΣΗΣ ΠΝΕΥΜΑΤΙΚΩΝ ΔΙΚΑΙΩΜΑΤΩΝ

Δηλώνω ότι η διπλωματική αυτή εργασία αποτελεί στο σύνολο της δική μου εργασία, και κανένα τμήμα της δεν έχει χρησιμοποιηθεί για την κτήση άλλου τίτλου σπουδών. Όπου έχει χρησιμοποιηθεί υλικό από άλλες πηγές, αυτές έχουν αναφερθεί με ακρίβεια και πληρότητα.

Ειρήνη Μαρία Ανδρινοπούλου

Υπογραφή Φοιτητή/τριας

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον κ. Κεπαπτσόγλου Κωνσταντίνο, Αναπληρωτή Καθηγητή της Σχολής Αγρονόμων και Τοπογράφων Μηχανικών-Μηχανικών Γεωπληροφορικής του Εθνικού Μετσόβιου Πολυτεχνείου, για την ανάθεση της παρούσας Διπλωματικής Εργασίας και την πολύτιμη βοήθειά του.

Επίσης θα ήθελα να ευχαριστήσω τον κ. Τζούρα Παναγιώτη, Μεταδιδακτορικό Ερευνητή του Εθνικού Μετσόβιου Πολυτεχνείου, για την βοήθεια, την υποστήριξη και την άψογη συνεργασία καθ' όλη την διάρκεια εκπόνησης της Διπλωματικής Εργασίας.

Τέλος, θέλω να ευχαριστήσω την οικογένεια και τους φίλους μου, οι οποίοι με υποστήριξαν κατά την διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Η παρούσα Διπλωματική Εργασία επικεντρώνεται στην ανάλυση των αστικών μετακινήσεων όσον αφορά την επιλογή μεταφορικού μέσου, στην Μητροπολιτική περιοχή της Αθήνας. Αντλώντας δεδομένα από τις καθημερινές μετακινήσεις των πολιτών, η έρευνα αποσκοπεί στην κατανόηση και βελτίωση των συστημάτων μεταφοράς, μέσω της εφαρμογής αλγορίθμων συσταδοποίησης για την ανίχνευση χώρο-χρονικών προτύπων. Παρά την σημαντικότητα των αστικών μετακινήσεων στην καθημερινότητα της πόλης, η έλλειψη πρόσφατων και ευρέως διαθέσιμων δεδομένων καθιστά δύσκολη την ανάλυση και τον σχεδιασμό βελτιωτικών παρεμβάσεων. Η έρευνα αυτή επιχειρεί να γεφυρώσει αυτό το κενό, αναδεικνύοντας τη σημασία των κοινωνικό-οικονομικών και δημογραφικών παραγόντων στις μετακινήσεις και προτείνοντας νέες μεθοδολογίες για την αποτελεσματική ανάλυση και πρόγνωση των συνθηκών που εκδηλώνουν οι κάτοικοι της Αθήνας σχετικά με τις μετακινήσεις τους. Μέσω της συλλογής και ανάλυσης δεδομένων, καθώς και της εφαρμογής καινοτόμων αλγορίθμων, η εργασία στοχεύει στην ενίσχυση της κατανόησης των αστικών μετακινήσεων και στη διαμόρφωση στρατηγικών για τη βελτίωση της ποιότητας ζωής στις πόλεις. Η δομή της εργασίας περιλαμβάνει την εισαγωγή στο θέμα, την παρουσίαση των στόχων, τη βιβλιογραφική ανασκόπηση, τη μεθοδολογία της έρευνας, τη συλλογή και περιγραφή δεδομένων, την επεξεργασία και ανάλυση των δεδομένων, καθώς και τα συμπεράσματα και προτάσεις για μελλοντική έρευνα.

ABSTRACT

This thesis focuses on the analysis of urban mobility in terms of transportation mode choice in the Metropolitan area of Athens. By drawing data from the daily movements of citizens, the research aims to understand and improve transportation systems through the application of clustering algorithms to detect spatiotemporal patterns. Despite the significance of urban mobility to the city's daily life, the lack of recent and widely available data makes it difficult to analyze and plan for improvements. This study attempts to bridge this gap, highlighting the importance of socio-economic and demographic factors in mobility and proposing new methodologies for the effective analysis and forecasting of the mobility habits of Athens' residents. Through the collection and analysis of data, as well as the application of innovative algorithms, the work aims to enhance the understanding of urban movements and to formulate strategies for improving the quality of life in cities. The structure of the thesis includes an introduction to the topic, the presentation of objectives, a literature review, research methodology, data collection and description, data processing and analysis, as well as conclusions and suggestions for future research.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1	ΕΙΣΑΓΩΓΗ.....	1
1.1	Γενικά	1
1.2	Σκοπός	2
1.3	Δομή.....	2
2	ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ	5
2.1	Ορισμοί.....	5
2.2	Αντικείμενο	6
2.3	Παράμετροι Επιλογής Μέσου	6
2.4	Μοντέλα Επιλογής Μέσου	8
2.5	Αλγόριθμοι Συσταδοποίησης στην Προτυποποίηση των Μετακινήσεων	9
2.6	Δημιουργία Συνθετικής Ζήτησης	10
2.7	Σύνοψη.....	12
3	ΜΕΘΟΔΟΛΟΓΙΑ.....	13
3.1	Περιγραφή Διαδικασίας.....	13
3.2	Μέθοδοι Δειγματοληψίας	14
3.3	Μέθοδοι Ανάλυσης Δεδομένων.....	19
3.4	Μέθοδοι Συσταδοποίησης	23
4	ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ.....	35
4.1	Γενικά	35
4.2	Περιγραφή Ερωτηματολογίων	35
4.3	Μεταβλητές.....	36
4.4	Περιγραφική Στατιστική.....	42
4.5	Επεξεργασία Δεδομένων	56
5	ΕΠΕΞΕΡΓΑΣΙΑ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ	75
5.1	Γενικά	75
5.2	Ανάλυση Δεδομένων	75
5.3	Διαδικασία Συσταδοποίησης.....	77
5.4	Αποτελέσματα Συσταδοποίησης.....	80
5.5	Αναγνώριση Συστάδων.....	105
6	ΣΥΜΠΕΡΑΣΜΑΤΑ.....	111
6.1	Σύνοψη.....	111
6.2	Κύρια ευρήματα	111
6.3	Περιορισμοί	112

6.4	Προτάσεις.....	113
7	ΑΝΑΦΟΡΕΣ	115
7.1	Διεθνής Βιβλιογραφία	115
7.2	Ελληνική Βιβλιογραφία	121
7.3	Ιστοσελίδες.....	121

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 4.1 Κωδικοποίηση μεταβλητών	36
Πίνακας 4.2 Πλήθος συμμετεχόντων ανά ζώνη κατοικίας	47
Πίνακας 4.3 Συσχετίσεις Μεταβλητών	55
Πίνακας 4.4 Κατανομή πλήθους μετακινήσεων ανά περιοχή κατοικίας.....	67
Πίνακας 5.1 Κατανομή πλήθους μετακινήσεων ανά συστάδα.....	80
Πίνακας 5.2 Μέση τιμή μεταβλητών ανά συστάδα	82
Πίνακας 5.3 Διάμεσος μεταβλητών ανά συστάδα	82
Πίνακας 5.4 Συνοπτική παρουσίαση χαρακτηριστικών κάθε συστάδας	109
Πίνακας 5.5 Δημογραφικά χαρακτηριστικά συστάδων.....	110

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 4.1 Ποσοστιαία κατανομή φύλου	42
Διάγραμμα 4.2 Κατανομή ηλικιών	43
Διάγραμμα 4.3 Ποσοστιαία κατανομή ηλικιακών ομάδων	44
Διάγραμμα 4.4 Ποσοστιαία κατανομή εκπαιδευτικών βαθμίδων	45
Διάγραμμα 4.5 Ποσοστιαία κατανομή εργασιακής κατάστασης	45
Διάγραμμα 4.6 Ποσοστιαία κατανομή εισοδηματικών κατηγοριών.....	46
Διάγραμμα 4.7 Ποσοστιαία κατανομή κατοχής ιδιωτικού αυτοκινήτου	47
Διάγραμμα 4.8 Ποσοστιαία κατανομή πλήθους διαδοχικών μετακινήσεων	50
Διάγραμμα 4.9 Κατανομή πλήθους διαδοχικών μετακινήσεων ανά ηλικιακή ομάδα	51
Διάγραμμα 4.10 Ποσοστιαία κατανομή αποστάσεων των μετακινήσεων	52
Διάγραμμα 4.11 Ποσοστιαία κατανομή σκοπών μετακινήσεων	53
Διάγραμμα 4.12 Ποσοστιαία κατανομή μέσων μετακίνησης	53
Διάγραμμα 4.13 Κατανομή ωρών έναρξης μετακινήσεων.....	54
Διάγραμμα 4.14 Ποσοστιαία κατανομή μέσων μετακίνησης ανά φύλο.....	56
Διάγραμμα 4.15 Ποσοστιαία κατανομή μέσων μετακίνησης ανά ηλικιακή ομάδα.....	57
Διάγραμμα 4.16 Ποσοστιαία κατανομή μέσων μετακίνησης ανά κατηγορία εισοδήματος	59
Διάγραμμα 4.17 Ποσοστιαία κατανομή μέσων μετακίνησης αναλόγως την κατοχή αυτοκινήτου.....	60
Διάγραμμα 4.18 Ποσοστιαία κατανομή μέσων μετακίνησης ανά επίπεδο εκπαίδευσης.....	62
Διάγραμμα 4.19 Ποσοστιαία κατανομή μέσων μετακίνησης ανά εργασιακή κατάσταση	63
Διάγραμμα 4.20 Ποσοστιαία κατανομή μέσων μετακίνησης ανά σκοπό μετακίνησης	66

Διάγραμμα 4.21 Ποσοστιαία κατανομή μέσων μετακίνησης ανά ζώνη κατοικίας	67
Διάγραμμα 4.22 Ποσοστιαία κατανομή μέσων μετακίνησης αναλόγως την απόσταση	69
Διάγραμμα 4.23 Ποσοστιαία κατανομή σκοπών μετακινήσης ανά ηλικιακή ομάδα	70
Διάγραμμα 4.24 Ποσοστιαία κατανομή σκοπών μετακινήσης ανά κατηγορία εισοδήματος.....	72
Διάγραμμα 4.25 Ποσοστιαία κατανομή σκοπών μετακινήσης ανά εργασιακή κατάσταση	73
Διάγραμμα 5.1 Ποσοστιαία κατανομή συστάδων.....	81
Διάγραμμα 5.2 Στατιστικά μέτρα της ώρας έναρξης ταξιδιού ανά συστάδα.....	84
Διάγραμμα 5.3 Στατιστικά μέτρα της απόστασης ταξιδιού ανά συστάδα.....	85
Διάγραμμα 5.4 Στατιστικά μέτρα του μέσου μετακίνησης ανά συστάδα	86
Διάγραμμα 5.5 Στατιστικά μέτρα του σκοπού μετακίνησης ανά συστάδα	86
Διάγραμμα 5.6 Θηκογράμματα μεταβλητών ανά συστάδα.....	87
Διάγραμμα 5.7 Διασπορά μεταβλητών ανά συστάδα.....	88
Διάγραμμα 5.8 Ποσοστιαία κατανομή μέσων μετακίνησης ανά συστάδα	89
Διάγραμμα 5.9 Ποσοστιαία κατανομή σκοπών μετακίνησης ανά συστάδα	91
Διάγραμμα 5.10 Κατανομή ώρας έναρξης των μετακινήσεων ανά συστάδα	93
Διάγραμμα 5.11 Κατανομή ώρας έναρξης του 75% των μετακινήσεων ανά συστάδα.....	94
Διάγραμμα 5.12 Κατανομή ώρας έναρξης του 50% των μετακινήσεων ανά συστάδα.....	95
Διάγραμμα 5.13 Κατανομή απόστασης των μετακινήσεων ανά συστάδα	96
Διάγραμμα 5.14 Κατανομή απόστασης του 75% των μετακινήσεων ανά συστάδα	98
Διάγραμμα 5.15 Κατανομή απόστασης του 50% των μετακινήσεων ανά συστάδα	99
Διάγραμμα 5.16 Διάγραμμα συσχετίσεων μεταβλητών πρώτης συστάδας	100
Διάγραμμα 5.17 Διάγραμμα συσχετίσεων μεταβλητών δεύτερης συστάδας	101
Διάγραμμα 5.18 Διάγραμμα συσχετίσεων μεταβλητών τρίτης συστάδας	102
Διάγραμμα 5.19 Διάγραμμα συσχετίσεων μεταβλητών τέταρτης συστάδας	103
Διάγραμμα 5.20 Διάγραμμα συσχετίσεων μεταβλητών πέμπτης συστάδας	104
Διάγραμμα 5.21 Διάγραμμα συσχετίσεων μεταβλητών έκτης συστάδας.....	105
Διάγραμμα 5.22 Τρισδιάστατη απεικόνιση των παρατηρήσεων των συστάδων	106

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1.1 Διάρθρωση διπλωματικής εργασίας	4
Εικόνα 3.1 Στάδια έρευνας	13
Εικόνα 3.2 Μεθοδολογία περιγραφικής στατιστικής.....	20
Εικόνα 3.3 Τεταρτημόρια σε διάφορες κατανομές	22
Εικόνα 3.4 Παράδειγμα Elbow Method.....	27
Εικόνα 3.5 Παράδειγμα Silhouette Method.....	28
Εικόνα 3.6 Παράδειγμα Gap Method.....	28
Εικόνα 5.1 Απόσπασμα κώδικα αποκλιμάκωσης δεδομένων	77
Εικόνα 5.2 Απόσπασμα κώδικα συνάρτησης k-means.....	78
Εικόνα 5.3 Απόσπασμα κώδικα δημιουργίας μήτρας ομοιότητας	78
Εικόνα 5.4 Απόσπασμα κώδικα υπολογισμού ιδιοδιανυσμάτων	79
Εικόνα 5.5 Απόσπασμα κώδικα εφαρμογής συσταδοποίησης	80

1 ΕΙΣΑΓΩΓΗ

1.1 Γενικά

Η αστική μετακίνηση και η επιλογή των μέσων μεταφοράς αποτελούν κεντρικά θέματα μελέτης στον τομέα της συγκοινωνιολογίας, με ιδιαίτερη έμφαση στη βελτιστοποίηση των υποδομών και στην αύξηση της αποδοτικότητας των μεταφορών.

Διαχρονικά, οι μετακινήσεις και οι μεταφορές αποτελούν καθοριστικό παράγοντα για την ευημερία των πόλεων. Συνιστούν κρίσιμο συντελεστή στην ανάπτυξη της κοινωνίας αλλά και στη δημιουργία μίας ομαλής καθημερινότητας για τους πολίτες της. Σε καθημερινή βάση, οι κάτοικοι της Αθήνας καλούνται να μετακινηθούν μέσα και έξω από την πόλη προκειμένου να καλύψουν τις καθημερινές τους ανάγκες σε θέματα εργασίας, ψυχαγωγίας, εκπαίδευσης, αγορών κ.α. Ταυτόχρονα όμως, έρχονται αντιμέτωποι με την επιλογή του μέσου μετακίνησης, της διαδρομής και της κατάλληλης ώρας έναρξης των μετακινήσεων. Οι επιλογές αυτές επηρεάζονται άμεσα από πολλούς σχετικούς με τα κοινωνικό-οικονομικά και δημογραφικά χαρακτηριστικά, καθώς επίσης και από τα χαρακτηριστικά της ίδιας της μετακίνησης και του δικτύου των μεταφορών. Για τη βελτίωση των μετακινήσεων και κατά επέκταση της ποιότητας ζωής των κατοίκων στη Μητροπολιτική περιοχή της Αθήνας, θα πρέπει να κατανοηθεί πλήρως η επιρροή που ασκείται από διάφορους παράγοντες στη διαμόρφωση της γενικότερης συμπεριφοράς μετακίνησης.

Η αδυναμία εφαρμογής καινοτόμων εργαλείων και η ελλιπής διαθεσιμότητα επικαιροποιημένων δεδομένων σχετικά με τις καθημερινές επιλογές μετακίνησης συμβάλουν στην αδυναμία εύρεσης βέλτιστων λύσεων και πρακτικών στο σχεδιασμό αστικών συστημάτων μεταφορών στην Ελλάδα. Πράγματι, η τελευταία έρευνα Προέλευσης – Προορισμού (ΠΠ) στην Μητροπολιτική περιοχή της Αθήνας, πραγματοποιήθηκε από τον ΟΑΣΑ το 2006 και έκτοτε δεν έχουν καταγραφεί ξανά συστηματικά δεδομένα για την κινητικότητα από κάποιο επίσημο φορέα. Μεμονωμένες προσπάθειες συλλογής ερωτηματολογίων έχουν διενεργηθεί στο πλαίσιο διαφόρων μελετών, όμως τις περισσότερες φορές, τα δεδομένα είτε δεν είναι ανοικτά ή δεν μπορούν να αξιοποιηθούν για διαφορετικό σκοπό εκτός αυτού για τον οποίο αρχικά συλλέχθηκαν. Επίσης, λόγω του παραπάνω προβλήματος, εκλείπουν τα μοντέλα και εργαλεία πρόβλεψης που θα μπορούσαν να δημιουργήσουν πειραματικά δεδομένα αξιοποιήσιμα στο πλαίσιο αξιολογήσεων μελλοντικών σεναρίων. Έτσι, οι λόγοι που οδηγούν καθημερινά τους κατοίκους της Μητροπολιτικής περιοχής της Αθήνας να επιλέγουν συγκεκριμένα μέσα και τύπους μετακίνησης παραμένουν μερικώς άγνωστοι, δεν έχουν ποσοτικοποιηθεί και δύσκολα μπορούν να παρατηρηθούν αλλά και να προβλεφθούν. Επιπλέον, δεν έχει αναλυθεί ποτέ στην Αθήνα η χωρική και η χρονική διάσταση των προτύπων μετακίνησης.

1.2 Σκοπός

Η παρούσα Διπλωματική Εργασία εστιάζει στην εφαρμογή αλγορίθμων συσταδοποίησης με απώτερο σκοπό την ανάγνωση διάφορων χώρο-χρονικών προτύπων που επηρεάζουν την επιλογή μέσου μετακίνησης στην Αθήνα. Η μελέτη αποσκοπεί στην κατανόηση των μετακινήσεων των ατόμων μέσα στην πόλη, με επίκεντρο την ανάλυση των κοινωνικό-οικονομικών παραμέτρων, των δημογραφικών χαρακτηριστικών και των προτιμήσεων μετακίνησης. Συνεπώς, ο σκοπός της εργασίας είναι διττός: αφενός, να αναλυθούν τα δεδομένα της έρευνας με στόχο τη στατιστική περιγραφή και την εξαγωγή συμπερασμάτων σχετικά με τις προτιμήσεις μετακίνησης των κατοίκων της Αθήνας και αφετέρου, να εξεταστεί η εφαρμογή και η αποδοτικότητα των αλγορίθμων συσταδοποίησης στην επεξεργασία και ανάλυση χώρο-χρονικών δεδομένων. Η εφαρμογή των αλγορίθμων K-μέσου και Φασματικής Συσταδοποίησης προσφέρει μια νέα διάσταση στην επεξεργασία και ανάλυση των μετακινήσεων, ενισχύοντας την κατανόηση των παραγόντων που επηρεάζουν τις επιλογές μετακίνησης.

Η προσέγγιση αυτή όχι μόνο αναδεικνύει την πρωτοτυπία και τη σημασία της εργασίας στον τομέα των αστικών μετακινήσεων αλλά παρέχει επίσης ένα ισχυρό θεμέλιο για περαιτέρω ερευνητική εργασία στην ανάλυση δεδομένων μετακινήσεων. Μέσα από την κατανόηση των χώρο-χρονικών προτύπων μετακίνησης και την **εφαρμογή σύγχρονων αλγορίθμων συσταδοποίησης**, η εργασία στοχεύει και στην αποτελεσματικότερη σχεδίαση των μεταφορικών υποδομών μέσα από τα ποιοτικά συμπεράσματα που εξάγονται στο τέλος.

Ειδικότεροι στόχοι της διπλωματικής εργασίας είναι:

- Η διερεύνηση των προτύπων μετακίνησης ως προς την επιλογή μέσου, μέσα από τη λεπτομερή ανάλυση των δεδομένων ερωτηματολογίου,
- Η εξέταση και ανάδειξη των κοινωνικών και οικονομικών παραγόντων που επηρεάζουν την επιλογή του μέσου μετακίνησης,
- Η μελέτη των χωρικών και χρονικών διαστάσεων του προβλήματος επιλογής μέσου μετακίνησης πέρα από τις τυπικές μεταβλητές που χρησιμοποιούνται,
- Η ανάπτυξη μοντέλων και εργαλείων για την προτυποποίηση των μετακινήσεων και κατ' επέκταση για την πρόβλεψη των επιλογών και την αναπαραγωγή των προτύπων σε συνθετικό πληθυσμό.

1.3 Δομή

Η διπλωματική εργασία οργανώνεται ως εξής:

Το **πρώτο κεφάλαιο** περιλαμβάνει την εισαγωγή στο θέμα, το σκοπό και τη δομή της εργασίας. Επίσης, παρουσιάζεται το πρόβλημα που πραγματεύεται η παρούσα Διπλωματική Εργασία καθώς η εισαγωγή θέτει το πλαίσιο και τους στόχους της έρευνας.

Το **δεύτερο κεφάλαιο** αφορά τη βιβλιογραφική ανασκόπηση, η οποία περιγράφει τα θεωρητικά θεμέλια, εστιάζοντας στους ορισμούς, στο αντικείμενο της μελέτης, καθώς και σε προηγούμενες έρευνες. Αναλυτικά, σε αυτό το κεφάλαιο παρουσιάζονται μελέτες σχετικά με την επιλογή μέσου μετακίνησης, τα μοντέλα που εφαρμόζονται στην προτυποποίηση των μετακινήσεων, τη συσταδοποίηση των μετακινήσεων και διάφορα μοντέλα συνθετικής ζήτησης, με εφαρμογές στον αστικό χώρο.

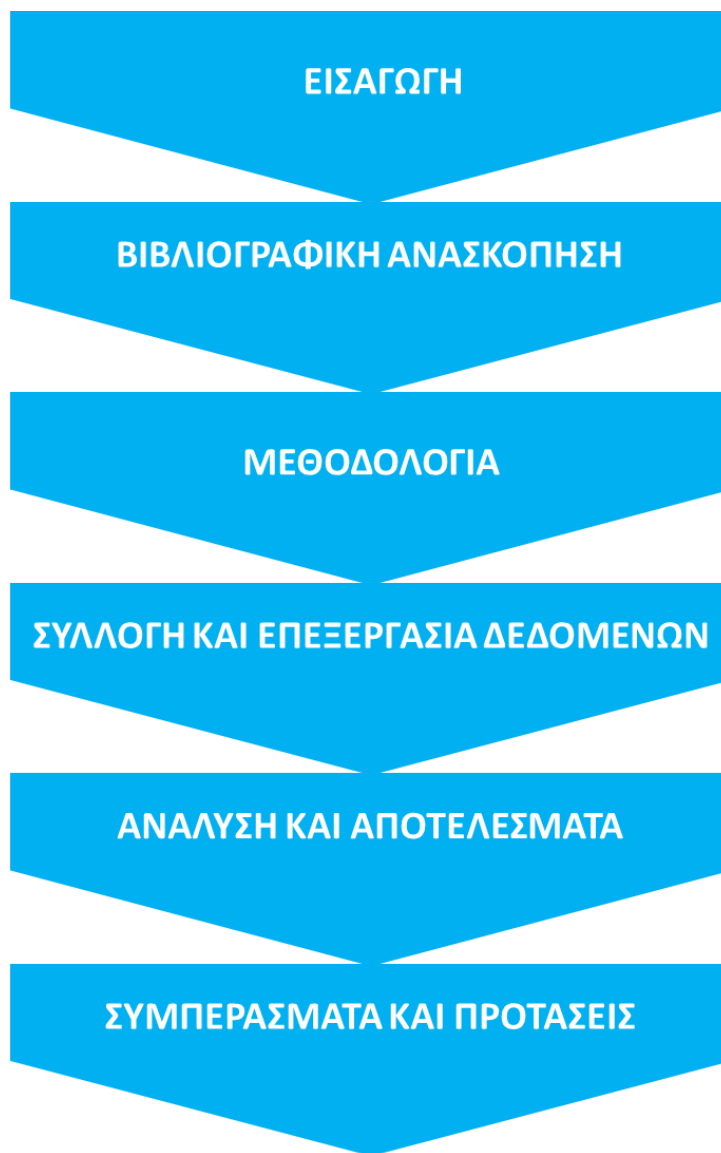
Το **τρίτο κεφάλαιο** εξετάζει τη μεθοδολογία που ακολουθήθηκε για τη συλλογή, στατιστική περιγραφή και επεξεργασία των δεδομένων. Το κεφάλαιο της μεθοδολογίας περιγράφει την προσέγγιση και τις μεθόδους της περιγραφής δεδομένων, καθώς και τις μεθόδους ανάλυσης των δεδομένων, δηλαδή των αλγορίθμων που εφαρμόστηκαν μαζί με τη διαχείριση των προκλήσεων που προέκυψαν.

Το **τέταρτο κεφάλαιο** αναφέρεται στη συλλογή και στατιστική περιγραφή των δεδομένων. Σε αυτό το κεφάλαιο, παρουσιάζεται μία περιεκτική περιγραφή των ερωτηματολογίων από όπου προέκυψαν τα δεδομένα της έρευνας καθώς και των μεταβλητών της. Επίσης, διεξάγεται διεξοδική περιγραφική στατιστική του δείγματος και περεταίρω ανάλυση του εμβαθύνοντας σε συγκεκριμένες μεταβλητές, ιδίως αυτές που αφορούν το μέσο πραγματοποίησης των μετακινήσεων.

Το **πέμπτο κεφάλαιο** αφορά την επεξεργασία και ανάλυση των δεδομένων καθώς επίσης, περιλαμβάνει τα αποτελέσματα της συσταδοποίησης. Παρουσιάζονται η διαδικασία και τα αποτελέσματα της συσταδοποίησης, αναγνωρίζονται τα χαρακτηριστικά των συστάδων και ταξινομούνται οι μετακινήσεις. Η εκάστοτε συστάδα, περιγράφεται στατιστικά όσον αφορά, τόσο τις μεταβλητές που συμμετείχαν στη συσταδοποίηση, όσο και τα κοινωνικό-οικονομικά χαρακτηριστικά της.

Τέλος, στο **έκτο κεφάλαιο** παρουσιάζονται τα συμπεράσματα και προτάσεις για μελλοντική έρευνα, παρέχοντας μια συνοπτική ανασκόπηση των κυριότερων ευρημάτων και των πιθανών εφαρμογών τους στη βελτίωση των αστικών μετακινήσεων. Επιπλέον, παρουσιάζονται οι περιορισμοί και οι προκλήσεις που αντιμετωπίστηκαν κατά τη διάρκεια διεξαγωγής της μελέτης.

Αυτή η δομή σχεδιάστηκε για να προσφέρει μια συνεκτική και λογική ανάλυση των χώρο-χρονικών προτύπων μετακίνησης, με έμφαση στην περίπτωση της Αθήνας, και να παρέχει ένα στέρεο θεμέλιο για μελλοντικές έρευνες στον τομέα των συγκοινωνιακών μελετών.



Εικόνα 1.1 Διάρθρωση διπλωματικής εργασίας

2 ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

Στο κεφάλαιο αυτό, παρουσιάζεται η βιβλιογραφική ανασκόπηση της παρούσας Διπλωματικής Εργασίας. Συγκεκριμένα, παρουσιάζονται διεθνείς μελέτες σχετικά με τα πρότυπα επιλογής μέσου μετακίνησης και μελέτες αλγορίθμων συσταδοποίησης. Στο τέλος του κεφαλαίου, συνοψίζονται οι παράγοντες που επηρεάζουν την επιλογή μέσου μετακίνησης και ο τρόπος που προκύπτουν και διερευνώνται τα χώρο-χρονικά πρότυπα επιλογής μέσου μετακίνησης.

2.1 Ορισμοί

Κάτωθι, παρουσιάζονται συνοπτικά, οι ορισμοί κάποιων απαραίτητων εννοιών για την ομαλή κατανόηση και διεξαγωγή της έρευνας.

Μοντέλα Μετακινήσεων

Τα Μοντέλα Μετακινήσεων είναι ένα σύνολο μαθηματικών σχέσεων για την αναπαράσταση (μοντελοποίηση) των επιλογών που κάνουν οι άνθρωποι κατά τη μετακίνησή τους. Αυτές οι επιλογές περιλαμβάνουν το πλήθος των μετακινήσεων τους, την προέλευση, τον προορισμό και με το μέσο μετακίνησης. Η ζήτηση μίας μετακίνησης είναι το συνδυασμένο αποτέλεσμα χιλιάδων ατόμων που κάνουν αυτές τις επιλογές. Επίσης τα μοντέλα αυτά είναι ένα εργαλείο που βοηθά τους σχεδιαστές να μελετήσουν τις επιπτώσεις εναλλακτικών σεναρίων μεταφορών, όπως νέοι αυτοκινητόδρομοι, αλλαγές στις διαδρομές λεωφορείων ή περιορισμοί στο πάρκινγκ, στη μελλοντική ζήτηση ταξιδιών, προκειμένου να ληφθούν επικαιροποιημένες πολιτικές αποφάσεις.

Μοντέλα Βάσει Συνθετικής Ζήτησης

«Το σημερινό σύστημα μεταφορών στην κοινωνία μας αποτελείται από σύνθετες αλληλεπιδράσεις μεγάλης κλίμακας που κατευθύνονται μερικώς από την απρόβλεπτη συμπεριφορά των ταξιδιωτών καθώς αλληλεπιδρούν μεταξύ τους και με το δυναμικό τους περιβάλλον, σε μια προσπάθεια να μεταβούν από το ένα μέρος στο άλλο. Τα συστήματα βασισμένα σε πράκτορες προσφέρουν μια κατάλληλη μέθοδο για τη μοντελοποίηση τέτοιας πολυπλοκότητας. Προερχόμενα από τον τομέα των υπολογιστικών συστημάτων, είναι συστήματα όπου οι πράκτορες, αντιπροσωπεύοντας οντότητες μέσα στο περιβάλλον τους, εκτελούν ανεξάρτητα τα καθήκοντά τους ενώ αλληλεπιδρούν με άλλους πράκτορες.» (Grace O. Kaghoa et al., 2020).

Η μοντελοποίηση βασισμένη σε πράκτορες (AgBM) είναι μια υπολογιστική μέθοδος που προσομοιώνει τη συμπεριφορά ατομικών "πρακτόρων" και τις αλληλεπιδράσεις τους εντός ενός συστήματος. Ένας πράκτορας μπορεί να είναι ένας άνθρωπος, ένα όχημα ή οποιαδήποτε άλλη

οντότητα που αλληλεπιδρά με άλλους πράκτορες. Τα τελευταία χρόνια, το AgBM έχει γίνει ένα όλο και πιο δημοφιλές εργαλείο για τη μοντελοποίηση σύνθετων συστημάτων, συμπεριλαμβανομένων εκείνων στον τομέα της κυκλοφορίας και των μεταφορών. Αναπαριστώντας οδηγούς, οχήματα και τη ροή της κυκλοφορίας, ένα AgBM μπορεί να χρησιμοποιηθεί για να κατανοήσουμε καλύτερα τη συμπεριφορά των συστημάτων κυκλοφορίας και να αναπτύξουμε πιο αποτελεσματικές πολιτικές μεταφορών. (Luis Willumsen, 2023)

2.2 Αντικείμενο

Η βιομηχανοποίηση έχει οδηγήσει σε υψηλότερους ρυθμούς ανάπτυξης, αυξημένο εισόδημα και μεγάλη ζήτηση για κινητικότητα. Έτσι, ο αυξανόμενος αριθμός οχημάτων στην πόλη προκαλεί συμφόρηση και περιβαλλοντικά προβλήματα που οδηγούν σε δυσμενείς συνθήκες κυκλοφορίας, όπως καθυστερήσεις και ατυχήματα, προκαλώντας τεράστια οικονομική ζημία κάθε χρόνο. Η προσέλκυση των χρηστών ιδιωτικών μέσων σε μαζικά μεταφορικά μέσα φαίνεται να είναι μια λύση, αλλά δεν είναι εύκολα εφικτή λόγω του παράγοντα της άνεσης των μέσων μαζικής μεταφοράς.

Για την ανακούφιση από τις επιδεινούμενες συνθήκες μεταφοράς, έχουν πραγματοποιηθεί μελέτες για την κατανόηση της σχέσης μεταξύ της επιλογής μεταφορικού μέσου και των διάφορων παραγόντων που την επηρεάζουν. Η ανάλυση της επιλογής μεταφορικού μέσου ωφελεί τους μηχανικούς και τους σχεδιαστές μεταφορών ώστε να μελετήσουν το υφιστάμενο σύστημα μεταφορών και να προβλέψουν τις μελλοντικές ανάγκες του συστήματος, καθώς δημιουργείται μια εικόνα για τις προτιμήσεις και τις απαιτήσεις των επιβατών.

Τα βήματα του σχεδιασμού των μεταφορών, κατά σειρά, είναι η γένεση των μετακινήσεων, η κατανομή των μετακινήσεων, η επιλογή μέσου μετακίνησης και ο καταμερισμός στο δίκτυο. Η ανάλυση της επιλογής μέσου μετακίνησης αποσκοπεί αρχικά στην κατανομή των μετακινούμενων στα διάφορα μέσα και έπειτα στην κατανομή των μετακινήσεων σε ζώνες. Σαφώς, είναι υψίστης σημασίας στη διαδικασία σχεδιασμού των μεταφορών και έχει άμεσο αντίκτυπο στη λήψη αποφάσεων, καθώς μέσω αυτής, επιτυγχάνεται η διαχείριση της ζήτησης για μεταφορές και επισημαίνονται οι αλλαγές που απαιτούνται ώστε αυτή να εξυπηρετηθεί. (Minal & Sekhar, 2014).

2.3 Παράμετροι Επιλογής Μέσου

Η επιλογή του μεταφορικού μέσου είναι μία πολυδιάστατη διαδικασία, η οποία επηρεάζεται από πληθώρα παραγόντων. Σύμφωνα με τους Minal και Sekhar (2014), η επιλογή μεταφορικού μέσου αποτελεί ένα από τα πιο κρίσιμα στάδια στη διαδικασία σχεδιασμού των μεταφορών και έχει άμεσο αντίκτυπο στην πολιτική λήψης αποφάσεων. Στην ίδια έρευνα επισημαίνεται ότι, η επιλογή

μέσου μετακίνησης επηρεάζεται από διάφορους κοινωνικούς, οικονομικούς, πολιτιστικούς και περιβαλλοντικούς παράγοντες, όπως είναι ο χρόνος ταξιδιού, το κόστος της μετακίνησης, η άνεση, ο χρόνος αναμονής κ.λπ.. Οι παράγοντες αυτοί κυμαίνονται από κοινωνικούς και οικονομικούς έως πολιτιστικούς και περιβαλλοντικούς, με το χρόνο ταξιδιού, το κόστος, την άνεση και το χρόνο αναμονής να αποτελούν καθοριστικούς παράγοντες (Racca & Ratledge, 2004).

Οι παράμετροι μπορούν να κατηγοριοποιηθούν σε δημογραφικές, οικονομικές, ψυχολογικές, κοινωνικό-πολιτισμικές και περιβαλλοντικές.

Από δημογραφικής πλευράς, η ηλικία, το φύλο και η εκπαίδευση έχουν αναδειχθεί ως κρίσιμοι παράγοντες επιλογής. Μελέτες έχουν δείξει ότι νεότεροι επιβάτες και φοιτητές είναι πιο πιθανό να επιλέξουν δημόσια συγκοινωνία, ενώ η προτίμηση για το αυτοκίνητο αυξάνεται με την ηλικία και το επίπεδο εισοδήματος (Schwanen et al., 2005).

Όσον αφορά τους οικονομικούς παράγοντες όπως το κόστος ταξιδιού, η διαθεσιμότητα και η προσβασιμότητα των μεταφορικών μέσων επίσης παίζουν σημαντικό ρόλο. Η σχετική μελέτη των Ben-Akiva και Lerman (1985) υπογραμμίζει τη σημασία της σύγκρισης κόστους και χρόνου ταξιδιού μεταξύ διαφόρων επιλογών μεταφοράς.

Ψυχολογικοί και κοινωνικοί παράγοντες, όπως η αντίληψη για την άνεση, την ασφάλεια και την προσωπική προτίμηση, επίσης επηρεάζουν την επιλογή μεταφορικού μέσου. Μια θεωρία που χρησιμοποιείται είναι η θεωρία της προσχεδιασμένης συμπεριφοράς (Ajzen, 1991). Σε άλλες έρευνες μελετάται η επιρροή της ατομικής ταυτότητας και προσωπικότητας στις επιλογές για τις μετακινήσεις. Μάλιστα, σε σχετικές έρευνες η μελέτη της ατομικής ταυτότητας αναδείχθηκε ως ένας σημαντικός παράγοντας πρόβλεψης συμπεριφοράς (Fekadu & Kraft, 2001). Επίσης, ο αυτοπροσδιορισμός ως χρήστης ενός συγκεκριμένου μέσου, βρέθηκε ότι αυξάνει την πιθανότητα χρήσης του συγκεκριμένου μέσου και ότι η ατομική «ταυτότητα» (self identity) δεν εξηγεί μόνο την επιλογή μέσου για ένα μόνο ταξίδι, αλλά και τη συχνότητα χρήσης του συγκεκριμένου μέσου (Heinen, 2016).

Η τεχνολογική πρόοδος και η εισαγωγή νέων μορφών μεταφοράς, όπως το car-sharing και το bike-sharing, έχουν αλλάξει το τοπίο της μετακίνησης, προσφέροντας περισσότερες και πιο ευέλικτες επιλογές στους χρήστες, ενισχύοντας την ποικιλομορφία στην επιλογή μεταφορικού μέσου (Shaheen et al., 2010).

Τέλος, περιβαλλοντικοί παράγοντες και η αυξανόμενη ευαισθητοποίηση για την κλιματική αλλαγή έχουν οδηγήσει σε μια στροφή προς πιο βιώσιμες μορφές μετακίνησης. Η έρευνα των

Gatersleben και Uzzell (2007) επισημαίνει την ανάγκη για πολιτικές που ενθαρρύνουν τη χρήση δημόσιας συγκοινωνίας και μη μηχανοκίνητων μέσων μεταφοράς ως μέσο προώθησης της βιωσιμότητας. Η μελέτη των Litman και Burwell (2006) συνδέει την επιλογή μεταφορικού μέσου με την ύπαρξη περιβαλλοντικών συνειδητήσεων και την άσκηση κοινωνικών πιέσεων.

Συνοψίζοντας, η επιλογή του μεταφορικού μέσου είναι μια πολύπλοκη διαδικασία που επηρεάζεται από διάφορους παράγοντες. Η προσεκτική μελέτη των παραμέτρων αυτών και η εφαρμογή των σχετικών ευρημάτων στο σχεδιασμό και την υλοποίηση πολιτικών μπορεί να βελτιώσει σημαντικά την αποτελεσματικότητα των συστημάτων μεταφορών.

2.4 Μοντέλα Επιλογής Μέσου

Η επιλογή του μεταφορικού μέσου είναι πιθανότατα ένα από τα πιο σημαντικά κλασσικά μοντέλα στο σχεδιασμό των μεταφορών. Αυτό οφείλεται στον κρίσιμο ρόλο που διαδραματίζει η δημόσια συγκοινωνία στη διαμόρφωση της πολιτικής (Ortuzar και Willumsen, 2002).

Ένα σημαντικό εργαλείο στην ανάλυση της επιλογής μεταφορικού μέσου είναι τα διακριτά μοντέλα επιλογής, με το πολυωνυμικό λογιστικό μοντέλο (Multinomial Logit Model - MNL) να αποτελεί μια από τις πιο διαδεδομένες προσεγγίσεις (Ben-Akiva & Lerman, 1985; McFadden, 1973). Αυτά τα μοντέλα βασίζονται στη θεωρία χρησιμότητας και προσπαθούν να προβλέψουν την επιλογή ενός ατόμου μεταξύ πολλαπλών εναλλακτικών με βάση τη μέγιστη προσδοκώμενη χρησιμότητα.

Πέρα από το MNL, έχουν αναπτυχθεί και άλλα μοντέλα για να αντιμετωπίσουν ορισμένους από τους περιορισμούς του, όπως το Nested Logit Model, το οποίο επιτρέπει την υποδιαίρεση των επιλογών σε ομάδες με παρόμοια χαρακτηριστικά, προσφέροντας μια πιο λεπτομερή ανάλυση (Ben-Akiva & Lerman, 1985). Επιπλέον, το Probit Model προσφέρει μια εναλλακτική προσέγγιση, λαμβάνοντας υπόψη τη συσχέτιση μεταξύ των εναλλακτικών και την αντιμετώπιση της υπόθεσης ανεξαρτησίας από τις άσχετες εναλλακτικές (IIA) που έχει το MNL (Train, 2009).

Το Mixed Logit Model, από την άλλη πλευρά, επιτρέπει την εισαγωγή ατομικής ετερογένειας στις παραμέτρους του μοντέλου, προσφέροντας μια πιο ευέλικτη και ακριβή προσέγγιση στην πρόβλεψη της επιλογής μεταφορικού μέσου (Train, 2009). Αυτή η προσέγγιση είναι ιδιαίτερα χρήσιμη στην αντιμετώπιση της ποικιλομορφίας των προτιμήσεων μεταξύ διαφορετικών ατόμων.

Πρόσφατες τεχνολογικές εξελίξεις και η αυξανόμενη διαθεσιμότητα μεγάλων συνόλων δεδομένων έχουν επίσης επιτρέψει την εφαρμογή προηγμένων τεχνικών μηχανικής μάθησης και τεχνητής νοημοσύνης στην πρόβλεψη της επιλογής μεταφορικού μέσου. Μέθοδοι όπως η συσταδοποίηση (clustering), οι αλγόριθμοι κατηγοριοποίησης και τα νευρωνικά δίκτυα, επιτρέπουν την ανάλυση

και την ερμηνεία πολύπλοκων συνόλων δεδομένων, αποκαλύπτοντας νέες προοπτικές και κατανοήσεις στις προτιμήσεις των επιβατών και τις δυναμικές της επιλογής μεταφορικού μέσου.

Η συσταδοποίηση, για παράδειγμα, μπορεί να χρησιμοποιηθεί για να ομαδοποιήσει επιβάτες με παρόμοιες προτιμήσεις ή συμπεριφορές μετακίνησης, παρέχοντας έτσι στους σχεδιαστές των μεταφορών τη δυνατότητα να προσαρμόσουν και να βελτιστοποιήσουν τα συστήματα μεταφοράς σύμφωνα με τις συγκεκριμένες ανάγκες και τις προτιμήσεις των διαφόρων υποομάδων (Xu & Wunsch, 2005).

Οι αλγόριθμοι συσταδοποίησης, όπως τα Δέντρα Αποφάσεων, ο Random Forest και οι Υποστηρικτικές Μηχανές Διανυσμάτων (SVM), έχουν επίσης εφαρμοστεί με επιτυχία στην πρόβλεψη της επιλογής μεταφορικού μέσου, επιτρέποντας την ακριβή κατηγοριοποίηση των επιβατών βάσει των προτιμήσεών τους (Breiman, 2001).

Τέλος, η χρήση Τεχνητών Νευρωνικών Δικτύων (ANN) στην ανάλυση της επιλογής μεταφορικού μέσου προσφέρει μια ιδιαίτερα δυναμική και ευέλικτη μέθοδο για την εξέταση πολύπλοκων μη γραμμικών σχέσεων μεταξύ των παραμέτρων που επηρεάζουν τις επιλογές των επιβατών (Kroese & Smagt, 1996). Η δυνατότητα των ANN να "μαθαίνουν" από τα δεδομένα και να προσαρμόζονται σε νέες πληροφορίες καθιστά αυτή την προσέγγιση ιδιαίτερα υποσχόμενη για την ανάπτυξη πιο εξατομικευμένων και αποτελεσματικών συστημάτων μεταφοράς.

Η συνεχής πρόοδος στις τεχνολογίες μηχανικής μάθησης και τεχνητής νοημοσύνης ανοίγει νέους δρόμους για την ανάλυση και τη βελτίωση των μεταφορικών συστημάτων, επιτρέποντας την πιο λεπτομερή κατανόηση των παραγόντων που επηρεάζουν τις επιλογές μεταφορικού μέσου και την ανάπτυξη πιο προσαρμοσμένων λύσεων για την εξυπηρέτηση των αναγκών των επιβατών.

2.5 Αλγόριθμοι Συσταδοποίησης στην Προτυποποίηση των Μετακινήσεων

Η έρευνα των Ester, M. et al., (1996) σχετικά με τους αλγόριθμους εύρεσης συστάδων σε μεγάλα σύνολα χωρικών δεδομένων με θόρυβο, παρουσιάζει τον αλγόριθμο DBSCAN, έναν κλασικό αλγόριθμο συσταδοποίησης που έχει εφαρμοστεί ευρέως σε δεδομένα μετακινήσεων. Επίσης, στην έρευνα των MacQueen, J. (1967), για τις μεθόδους ομαδοποίησης και ανάλυσης πολύ-παραμετρικών παρατηρήσεων, περιγράφεται ο αλγόριθμος K-means, ένας από τους πιο διαδεδομένους αλγόριθμους συσταδοποίησης, κατάλληλος για την ανάλυση δεδομένων μετακινήσεων.

Η έρευνα των Shidong Liang et al. (2022), προτείνει μια μέθοδο για την ταξινόμηση της κατάστασης της κυκλοφορίας χρησιμοποιώντας μια συνδυασμένη εφαρμογή του αλγορίθμου k-

medoids και του αλγορίθμου φασματικής συσταδοποίησης. Οι συγγραφείς χρησιμοποίησαν το συντελεστή σιλουέτας, το δείκτη Davies-Bouldin (DB) και το δείκτη Krzanowski-Lai (KL) για να καθορίσουν τον αριθμό των συστάδων στον αλγόριθμο k-medoids. Η εργασία συζητά επίσης τη χρήση του αλγορίθμου αυτό-ρυθμιζόμενης φασματικής συσταδοποίησης για τη διάκριση καταστάσεων της κυκλοφορίας βάσει παραμέτρων κυκλοφορίας. Τα αποτελέσματα του αλγορίθμου αυτόματης ρύθμισης φασματικής συσταδοποίησης συγκρίνονται με αυτά άλλων αλγορίθμων συσταδοποίησης, όπως ο παραδοσιακός αλγόριθμος φασματικής συσταδοποίησης, ο αλγόριθμος k-means και ο αλγόριθμος FCM, δείχνοντας ότι ο αλγόριθμος αυτόματης ρύθμισης φασματικής συσταδοποίησης παράγαγε καλύτερα αποτελέσματα.

2.6 Δημιουργία Συνθετικής Ζήτησης

Οι Bastariento, FF et al. (2021) παρουσίασαν μια εκτενή επισκόπηση της τρέχουσας κατάστασης των μοντέλων βασισμένων σε πράκτορες (συνθετική ζήτηση) στην έρευνα αστικών μεταφορών, περιλαμβάνοντας κενά ερευνητικού ενδιαφέροντος, προκλήσεις, και δυνητικές μελλοντικές κατευθύνσεις έρευνας. Η μεθοδολογία που χρησιμοποιήθηκε στη μελέτη περιλαμβάνει τη συλλογή δεδομένων, τη βιβλιογραφική ανάλυση, την ανάλυση περιεχομένου, και την οπτικοποίηση για να τονίσει την ποικιλομορφία των εφαρμογών των μοντέλων βασισμένων σε πράκτορες στις αστικές μεταφορές.

Η εργασία των Hörl et al. (2021), παρουσιάζει μια διαδικασία για τη δημιουργία συνθετικής ζήτησης μετακίνησης στο Παρίσι και στην περιοχή Île-de-France, βασισμένη εξ ολοκλήρου σε ανοικτά δεδομένα και ανοικτό λογισμικό, η οποία μπορεί να αναπαραχθεί από οποιονδήποτε ερευνητή και προορίζεται να χρησιμοποιηθεί ως μια πλήρης βάση δεδομένων για προσομοιώσεις μεταφορών βασισμένες σε πράκτορες και ως ένα πεδίο δοκιμών για αλγόριθμους σύνθεσης πληθυσμού και ζήτησης. Η εργασία περιγράφει έναν πλήρη κύκλο εργασιών από τα αρχικά σύνολα δεδομένων έως ένα τελικό σύνολο δεδομένων ζήτησης μετακίνησης, σχεδιασμένο να είναι αναπαραγώγιμος και προσβάσιμος σε οποιονδήποτε ερευνητή. Η εργασία παρέχει μια επικύρωση για τη ζήτηση μετακίνησης του Παρισιού και συζητά τις δομές συσχέτισης που δημιουργούνται από οποιαδήποτε διαδικασία συνθετικής ζήτησης.

Σύμφωνα με τους Ziemke et al. (2019), το MATSim Open Berlin, είναι ένα σενάριο προσομοίωσης μεταφορών για τη μητροπολιτική περιοχή του Βερολίνου που βασίζεται αποκλειστικά σε ανοικτά δεδομένα και μια πλήρως συνθετική διαδικασία, αποδεικνύοντας τη δυνατότητα δημιουργίας ενός ρεαλιστικού σεναρίου προσομοίωσης μεταφορών χωρίς την ανάγκη για πληροφορίες από έρευνα ημερολογίου μετακινήσεων. Σε αντίθεση με τα περισσότερα σενάρια προσομοίωσης μεταφορών, δεν απαιτεί πληροφορίες από έρευνα ημερολογίου μετακινήσεων ως

είσοδο. Η διαδικασία δημιουργίας σεναρίου που περιγράφεται στη μελέτη είναι χωρικά μεταφερόμενη και διευκολύνει τη δημιουργία σεναρίων προσομοίωσης μεταφορών βασισμένων σε πράκτορες για τυχαίες περιοχές. Το άρθρο τονίζει την ανάγκη για εκφραστικά μοντέλα μεταφορών δηλώνοντας ότι με την εμφάνιση νέων υπηρεσιών και τεχνολογιών μεταφορών, το σύστημα μεταφορών γίνεται πιο ατομικοποιημένο σε πολλές πτυχές του. Επίσης, υπογραμμίζει ότι τα μοντέλα μεταφορών, ως το πιο σημαντικό εργαλείο για την αξιολόγηση πολιτικών και σχεδιασμών, πρέπει να είναι αρκετά εκφραστικά για να αντιμετωπίσουν αυτές τις εξελίξεις. Το άρθρο συγκεκριμένα αναφέρει ότι τα μοντέλα μεταφορών βασισμένα σε πράκτορες, τα οποία διευθετούν τους ταξιδιώτες με ατομικές ιδιότητες, που έχουν την ικανότητα να δρουν και να αποφασίζουν αυτόνομα, επιτρέπουν την κατάλληλη μοντελοποίηση και ανάλυση της ατομικοποιημένης φύσης του εξελισσόμενου συστήματος μεταφορών.

Στο πλαίσιο του eqasim για αναπαραγώγιμες προσομοιώσεις κινητικότητας βασισμένες σε πράκτορες, σύμφωνα με τους Hörl et al. (2021), χρησιμοποιήθηκε το eqasim ως ένα κανάλι για παροχή προσομοιώσεων των μεταφορών, από αρχικά δεδομένα. Στην έρευνα συμπεριλαμβάνονται ενότητες για την αξιολόγηση της προσομοίωσης και των μοντέλων που βασίζονται σε πράκτορες. Επιπλέον, παρέχονται πληροφορίες σχετικά με την προσέγγιση του eqasim και παρουσιάζεται μια υλοποίηση του καναλιού σε ένα μοντέλο για τη Ζυρίχη, Ελβετία.

Σύμφωνα με την εργασία των Eftekhar, (2023), διερευνώνται τα χρονικά πρότυπα στην παραγωγή συνθετικών μετακινήσεων. Συγκεκριμένα, χρησιμοποιήθηκαν δεδομένα μετακινήσεων ενός ολόκληρου μήνα από ζώνες κυκλοφορίας στην Ολλανδία. Η ανάλυση συσταδοποίησης έδειξε ότι τα συνολικά πρότυπα παραγωγής ζήτησης είναι αναγνωρίσιμα βάσει των χαρακτηριστικών χρήσης γης. Η εκτελεσμένη χρονική ανάλυση των υποκείμενων προτύπων είναι πολύτιμη για την προσαρμογή των μοντέλων ζήτησης και πρόβλεψης, και τα ευρήματα υποδεικνύουν ότι, δεδομένων των χαρακτηριστικών χρήσης γης κάθε ζώνης, το πιο πιθανό χρονικό πρότυπο παραγωγής μετακινήσεων τους είναι ανιχνεύσιμο, το οποίο είναι ωφέλιμο για τα δυναμικά μοντέλα πρόβλεψης ζήτησης.

Στην συνέχεια, οι Hagenauer και Helbich (2012), στα πλαίσια της έρευνάς τους σχετικά με την χρήση ταξινομητών μηχανικής μάθησης στη μοντελοποίηση μέσου μετακίνησης, σύγκριναν την απόδοση επτά διαφορετικών ταξινομητών μηχανικής μάθησης. Χρησιμοποιήθηκαν δεδομένα ημερολογίων μετακινήσεων από την Ολλανδία από το 2010 έως το 2012 και τονίστηκε η σημασία της ακριβούς μοντελοποίησης για το σχεδιασμό μεταφορών και τη λήψη αποφάσεων. Βρέθηκε ότι ο Random Forest έχει σημαντικά καλύτερη απόδοση από οποιαδήποτε άλλη μέθοδο ταξινόμησης που μελετήθηκε.

Οι Gang Shen και Dongmei YeH (2017), ανέπτυξαν και προσάρμοσαν έναν καινοτόμο αλγόριθμο φασματικής συσταδοποίησης, χρησιμοποιώντας τα ιδιοδιανύσματα του πίνακα αποστάσεων για τη διαμέριση κόμβων, και υλοποιώντας έναν κατανεμημένο αλγόριθμο για φασματική ανάλυση χωρίς την ανταλλαγή πληροφοριών απόστασης μεταξύ κόμβων. Οι προσομοιώσεις απέδειξαν την αποτελεσματικότητα της προτεινόμενης προσέγγισης τόσο σε συνθετικά όσο και σε ελεγχόμενα σύνολα δεδομένων

2.7 Σύνοψη

Η παρούσα βιβλιογραφική ανασκόπηση εστιάζει στα πρότυπα επιλογής μέσου μετακίνησης και στην ανάλυση αλγορίθμων συσταδοποίησης, με βάση διεθνείς μελέτες. Αναλύονται οι παράγοντες που επηρεάζουν την επιλογή μέσου μετακίνησης και πώς αυτοί συντελούν στη διαμόρφωση χώρο-χρονικών προτύπων. Συγκεκριμένα, η εργασία αναφέρεται στα Μοντέλα Μετακινήσεων, που περιγράφουν μαθηματικά τις επιλογές ταξιδιού των ανθρώπων, και στα Μοντέλα Βάσει Συνθετικής Ζήτησης, αναδεικνύοντας τη σημασία τους στο σχεδιασμό των μεταφορών. Παρουσιάζονται επίσης διάφοροι αλγόριθμοι συσταδοποίησης, όπως ο DBSCAN και ο K-means, και η εφαρμογή τους σε δεδομένα μετακίνησης. Ειδική αναφορά γίνεται στην προσέγγιση του MATSim Open Berlin για τη δημιουργία σεναρίων προσομοίωσης μεταφορών βασισμένων αποκλειστικά σε ανοιχτά δεδομένα, καθώς και στη σημασία της αναπαραγωγιμότητας και της διαφάνειας στις μελέτες προσομοίωσης μεταφορών βασισμένων σε πράκτορες.

3 ΜΕΘΟΔΟΛΟΓΙΑ

Στο παρόν κεφάλαιο, θα γίνει περιγραφή της μεθοδολογίας που ακολουθήθηκε για την επίτευξη του στόχου της Διπλωματικής Εργασίας. Θα αναλυθούν τα στάδια της έρευνας και το θεωρητικό υπόβαθρο των μεθόδων και των αλγορίθμων που χρησιμοποιήθηκαν.

3.1 Περιγραφή Διαδικασίας

Τα στάδια της έρευνας της Διπλωματικής Εργασίας διαρθρώνονται ως εξής:



Εικόνα 3.1 Στάδια έρευνας

Μελέτη βιβλιογραφίας

Η μελέτη της διεθνούς βιβλιογραφίας αποτέλεσε το πρώτο στάδιο της έρευνας. Αρχικά, μελετήθηκαν οι μεταβλητές που συμπεριλαμβάνονται στις έρευνες γύρω από την κατανομή των μετακινήσεων και στη συνέχεια, μελετήθηκαν έρευνες που χρησιμοποιούν με παρόμοιο τρόπο μεθόδους συσταδοποίησης. Το περιεχόμενο της μελέτης της βιβλιογραφίας παρουσιάστηκε λεπτομερώς στο δεύτερο κεφάλαιο της Διπλωματικής Εργασίας.

Μετά τη βιβλιογραφική ανασκόπηση, επιλέχθηκαν οι μεταβλητές που θα διερευνηθούν στα πλαίσια της παρούσας Διπλωματικής Εργασίας. Με στόχο την αναγνώριση των χώρο-χρονικών προτύπων επιλογής μέσου, διακρίθηκαν και επιλέχθηκαν συγκεκριμένα, οι μεταβλητές της απόστασης, του χρόνου, του μέσου και του σκοπού της μετακίνησης. Επιπλέον, επιλέχθηκαν για να διερευνηθούν τα οικονομικά, κοινωνικά και δημογραφικά χαρακτηριστικά του δείγματος

Ανάλυση δεδομένων

Σε αυτό το στάδιο της μελέτης, έγινε η ανάλυση των δεδομένων που συλλέχθηκαν από τα ερωτηματολόγια. Όλα τα δεδομένα περιεγράφηκαν στατιστικά, ώστε να αποκαλυφθούν τα χαρακτηριστικά των συμμετεχόντων της έρευνας. Η μεθοδολογία που ακολουθήθηκε για τη στατιστική περιγραφή των μεταβλητών αναλύεται στη συνέχεια.

Έπειτα, το σύνολο των δεδομένων οπτικοποιήθηκε και χαρακτηρίστηκε από ποσοστά προκειμένου να ερμηνευτεί καταλλήλως. Ακολούθησε ο υπολογισμός των κεντρικών τάσεων και

των συσχετίσεων των βασικών μεταβλητών. Η οπτικοποίηση των στατιστικών μέτρων των μεταβλητών πραγματοποιήθηκε με χρήση του περιβάλλοντος της R-studio.

Επεξεργασία και αποτελέσματα

Στη συνέχεια, πραγματοποιήθηκε η επεξεργασία των δεδομένων με σκοπό τη συσταδοποίηση τους και απώτερο στόχο τη διερεύνηση των προτύπων μετακίνησης στη Μητροπολιτική περιοχή της Αθήνας. Για να επιτευχθεί η συσταδοποίηση έπρεπε τα δεδομένα των μεταβλητών να τροποποιηθούν προς στην κατάλληλη μορφή. Αυτή η επεξεργασία των δεδομένων έγινε με χρήση του R-studio καθώς και η διαδικασία της συσταδοποίησης. Στη συσταδοποίηση χρησιμοποιήθηκαν διάφορες μεθοδολογίες και αλγόριθμοι όπως αυτός των κ-μέσων και της φασματικής συσταδοποίησης, οι οποίοι θα περιγραφούν διεξοδικά.

Ερμηνεία-Συμπεράσματα

Τελικό στάδιο της παρούσας έρευνας αποτέλεσε η παρουσίαση και η ερμηνεία των αποτελεσμάτων που προέκυψαν από την ανάλυση και τη συσταδοποίηση των δεδομένων. Φυσικά σαν προέκταση των συμπερασμάτων, παρουσιάστηκαν οι περιορισμοί της έρευνας και προτάσεις για μελλοντικές έρευνες.

3.2 Μέθοδοι Δειγματοληψίας

Για τον προσδιορισμό της επίδρασης και τον υπολογισμό της στατιστικής σημαντικότητας των μελετηθέντων μεταβλητών σε σχέση με την επιλογή μέσου μετακίνησης, αναπτύχθηκε ένα ερωτηματολόγιο. Μέσω αυτού, πραγματοποιήθηκε η έρευνα για τα πρότυπα επιλογής μέσου μετακίνησης των κατοίκων της Αθήνας.

Το ερωτηματολόγιο αποτελεί μια γλωσσική τεχνική για τη διενέργεια ερευνών και καθίσταται το κύριο εργαλείο αλληλεπίδρασης μεταξύ του ερευνητή και του ατόμου που υποβάλλεται σε έρευνα. Επομένως, η διαμόρφωση του ερωτηματολογίου στα πλαίσια μιας έρευνας είναι το στάδιο με τη μεγαλύτερη σημασία για την εξέλιξη της ερευνητικής διαδικασίας. Αυτή η αναγκαιότητα επιβάλλει στον ερευνητή να καθορίσει με ακρίβεια τους στόχους της έρευνας, τη μεθοδολογία και το είδος της, τις πληροφορίες που πρέπει να συγκεντρωθούν, τις ερωτήσεις που θα διευκολύνουν τη συλλογή δεδομένων και το δείγμα που είναι απαραίτητο.

Διάφορες τεχνικές και εργαλεία είναι διαθέσιμα για τη συγκέντρωση δεδομένων στο πλαίσιο μιας έρευνας. Ανάμεσά τους, δεδομένα μπορούν να συλλεγούν μέσω εστιασμένων ομαδικών συνεντεύξεων, ημερολογίων μετακινήσεων και ερωτηματολογίων. Οι εστιασμένες ομαδικές συνεντεύξεις παρέχουν ποιοτικά δεδομένα μέσα από διαδραστικές συζητήσεις καθοδηγούμενες

από τον ερευνητή, ενώ τα ημερολόγια μετακινήσεων συλλέγουν σημαντικές πληροφορίες για το σχεδιασμό των μεταφορικών δικτύων, καταγράφοντας τα δημογραφικά, κοινωνικοοικονομικά και ταξιδιωτικά χαρακτηριστικά. Τέλος, τα ερωτηματολόγια, ως εργαλεία ποσοτικής έρευνας, επιτρέπουν την καταγραφή είτε των εκδηλωμένων προτιμήσεων, που αντανακλούν την πραγματική συμπεριφορά, είτε των δηλωμένων προτιμήσεων για υποθετικές καταστάσεις, είτε ενός συνδυασμού τους.

Η διαδικασία συλλογής δεδομένων σε μια έρευνα επηρεάζεται σημαντικά από τη μέθοδο διανομής. Έτσι, διακρίνουμε τις εξής πρακτικές για τη διανομή και συλλογή ερωτηματολογίων:

- Ταχυδρομική αποστολή, με την οποία τα ερωτηματολόγια στέλνονται και επιστρέφονται ταχυδρομικά. Παρά το οικονομικό της συμφέρον και την ευκολία της, η ταχυδρομική αποστολή είναι ωστόσο μια διαδικασία που επισύρει μακροχρόνια αναμονή και συχνά χαμηλό ποσοστό ανταπόκρισης.
- Τηλεφωνική έρευνα, επιτρέπει την άμεση συλλογή πληροφοριών μέσω διαλόγου. Μπορεί να αντιμετωπιστεί με διστακτικότητα από τους συμμετέχοντες, περιορίζοντας την αποτελεσματικότητα της μεθόδου.
- Προσωπικές συνεντεύξεις, όπου η συμπλήρωση γίνεται δια ζώσης, εξασφαλίζοντας μεγάλη ευελιξία και ακρίβεια στις απαντήσεις, αλλά με το μειονέκτημα του υψηλού κόστους.
- Άμεση παράδοση και συλλογή, όπου ο ερευνητής ή ένας αντιπρόσωπος παραδίδει και συλλέγει τα ερωτηματολόγια αυτοπροσώπως, προσφέροντας έλεγχο στη διαδικασία αλλά ενέχει τον κίνδυνο της πιθανής χαμηλής ανταπόκρισης λόγω έλλειψης χρόνου των ερωτώμενων.
- Διαδικτυακή διάθεση, με τα ερωτηματολόγια να είναι διαθέσιμα ηλεκτρονικά σε ιστοσελίδα, επιτρέπει ευρεία πρόσβαση και χαμηλό κόστος, αλλά με περιορισμούς στην πληρότητα και ορθότητα των απαντήσεων και στη δυνατότητα διευκρίνισης ερωτημάτων.

Η επιλογή της κατάλληλης μεθόδου εξαρτάται από τους στόχους της έρευνας, την απαιτούμενη ακρίβεια στις απαντήσεις, το επιθυμητό μέγεθος δείγματος, την έκταση και την πολυπλοκότητα του ερωτηματολογίου, το διαθέσιμο χρόνο, καθώς και από τη χρηματοδότηση και τη διαχείριση της έρευνας.

Στη διαδικασία διαμόρφωσης ενός ερωτηματολογίου, η επιλογή των ερωτήσεων πρέπει να ανταποκρίνεται στους στόχους της έρευνας, να είναι διατυπωμένες απλά και κατανοητά, καθώς και να τοποθετούνται σε μια λογική σειρά και μορφή. Επιπρόσθετα, ζωτικής σημασίας είναι οι έννοιες και οι όροι, που ενδέχεται να μην είναι γνωστοί στους ερωτώμενους, να εξηγούνται επαρκώς, ειδικά όταν το ερωτηματολόγιο δεν διενεργείται τηλεφωνικώς ή δια ζώσης.

Υπάρχουν κυρίως δύο είδη ερωτήσεων: ανοικτές και κλειστές. Οι ανοικτές ερωτήσεις, παρόλο που είναι σχετικά απλές στο σχεδιασμό τους, φέρουν το μειονέκτημα της δυσκολίας στατιστικής επεξεργασίας των απαντήσεων, καθώς παρέχουν την ελευθερία στους ερωτώμενους να εκφράζονται χωρίς περιορισμούς. Αντίθετα, οι κλειστές ερωτήσεις, συχνά πολλαπλών επιλογών ή βαθμολογούμενες κλίμακες, προσφέρουν τη δυνατότητα για ταχύτερη και πιο οργανωμένη απάντηση. Είναι σημαντικό, ωστόσο, να διαμορφώνονται με σαφήνεια, παρέχοντας διακριτές επιλογές στους ερωτώμενους.

Η συνολική δομή ενός ερωτηματολογίου επηρεάζει την αποτελεσματικότητα της έρευνας. Επισημαίνεται ότι, η τοποθέτηση των ερωτήσεων σε μία συγκεκριμένη σειρά μπορεί να βελτιώσει την ανταπόκριση των συμμετεχόντων.

Συνοψίζοντας, η επιτυχία ενός ερωτηματολογίου εξαρτάται από την πληρότητα, τη σαφήνεια, τη συνοχή, και την οργανωμένη δομή του. Επιπλέον, η συντομία, η περιληπτικότητα και η παροχή κατευθυντήριων οδηγιών και επεξηγήσεων εννοιολογικών όρων είναι ζωτικής σημασίας.

Πέρα από τις ερωτήσεις που στοχεύουν στην κατανόηση των χαρακτηριστικών του επιλεγμένου δείγματος, στα ερωτηματολόγια περιλαμβάνονται επίσης ερωτήσεις που συλλέγουν τις προτιμήσεις των ατόμων σχετικά με το αντικείμενο μελέτης. Από τις απαντήσεις σε αυτές τις ερωτήσεις, σε συνδυασμό με τα δεδομένα για τα χαρακτηριστικά του δείγματος, προκύπτουν τα ευρήματα της συγκεκριμένης έρευνας.

Η ανάλυση των προτιμήσεων των αποκρινόμενων μπορεί να εκτελεστεί μέσω δύο βασικών προσεγγίσεων:

A) Μέθοδος των εκδηλωμένων προτιμήσεων (RP – Revealed Preference)

B) Μέθοδος των δηλωμένων προτιμήσεων (SP – Stated Preference)

Στο πλαίσιο της εκπόνησης της παρούσας Διπλωματικής Εργασίας, υλοποιήθηκε έρευνα με τη χρήση της μεθόδου των εκδηλωμένων (αποκαλυπτόμενων) προτιμήσεων.

A) Η θεωρία των εκδηλωμένων προτιμήσεων (Revealed Preference), μια οικονομική έννοια που εισήγαγε ο Πολ Άντονι Σάμουελσον, Αμερικανός οικονομολόγος, το 1938, υποστηρίζει ότι οι επιλογές των καταναλωτών, υπό σταθερές συνθήκες εισοδήματος και τιμών προϊόντων, αποτελούν τον πιο ακριβή δείκτη των προτιμήσεων τους. Αυτή η προσέγγιση προτείνει ότι η ανάλυση της αγοραστικής συμπεριφοράς των καταναλωτών είναι η καλύτερη μέθοδος για τον προσδιορισμό των προτιμήσεων τους.

Η μέθοδος των εκδηλωμένων προτιμήσεων βρίσκει εφαρμογή σε διάφορους τομείς, όπως είναι ο τομέας των μεταφορών. Τα δεδομένα για τις έρευνες αυτού του τύπου συλλέγονται από την καταγραφή πραγματικών αποφάσεων και επιλογών των ατόμων που έχουν γίνει στο παρελθόν.

Όσον αφορά τα πλεονεκτήματα και τα μειονεκτήματα των ερευνών εκδηλωμένων προτιμήσεων σημειώνεται σαν θετικό στοιχείο το ότι οι προτιμήσεις που καταγράφονται αντικατοπτρίζουν πραγματικές επιλογές, εξαλείφοντας έτσι την αντίθεση μεταξύ απαντήσεων και πραγματικής συμπεριφοράς. Η σχεδίαση της έρευνας είναι συνήθως πιο απλή με αυτή τη μέθοδο. Ωστόσο, η χαμηλή διακύμανση των παρατηρηθέντων χαρακτηριστικών μπορεί να δυσχεραίνει τον εντοπισμό συσχετίσεων και τη διαμόρφωση σχέσεων. Επίσης, δεν είναι εφικτή η πρόβλεψη της επιλογής ενός νέου προϊόντος ή συστήματος με χαρακτηριστικά διαφορετικά από αυτά των υφιστάμενων.

B) Η τεχνική των δεδηλωμένων προτιμήσεων (Stated Preference), αποτελεί μια στατιστική μέθοδο για την καταγραφή των απόψεων και των επιθυμιών ενός δείγματος πληθυσμού σε σχέση με ένα συγκεκριμένο θέμα. Κατά τους Kroes και Sheldon (1988), οι δηλωμένες προτιμήσεις άρχισαν να χρησιμοποιούνται στη δεκαετία του '70 κυρίως στο πεδίο του marketing, επεκτεινόμενες αργότερα και στις μελέτες πάνω στις μεταφορές, όπως σημειώνει ο Hensser (1994).

Ειδικότερα στον τομέα των μεταφορών, οι μελέτες που βασίζονται στη μέθοδο των δηλωμένων προτιμήσεων αναλύουν τις επιλογές των ατόμων μέσα σε υποθετικές καταστάσεις. Παρουσιάζοντας διάφορα σενάρια μετακίνησης, αυτές οι έρευνες εξερευνούν ένα ευρύ φάσμα υποθετικών συνθηκών και αποτιμήσεων, διασφαλίζοντας την αναγκαία διακύμανση για την ακριβή εκτίμηση των μοντελικών παραμέτρων. Ωστόσο, όπως επισημαίνουν οι Kroes και Sheldon (1988), υπάρχει πάντα ο κίνδυνος οι ερωτώμενοι να μην επιλέγουν ό,τι θα επέλεγαν στην πραγματικότητα, ενώ ο αριθμός των προσφερόμενων σεναρίων μπορεί να περιορίζεται για να μην καταστεί η διαδικασία υπερβολικά περίπλοκη για τους συμμετέχοντες.

Τα πλεονεκτήματα των μεθόδων δηλωμένων προτιμήσεων περιλαμβάνουν τη δυνατότητα δημιουργίας στατιστικά ικανών μοντέλων και τη μείωση του αριθμού των απαιτούμενων συνεντεύξεων και του κόστους έρευνας σε σύγκριση με τις έρευνες εκδηλωμένων προτιμήσεων. Ωστόσο, η διαδικασία σχεδιασμού μπορεί να είναι πιο περίπλοκη, ενώ υφίστανται ζητήματα ακρίβειας λόγω της ελλείψεως πλήρους ελέγχου στις απαντήσεις των ερωτώμενων και της δυσκολίας στην επιβεβαίωση της αντιστοιχίας των δηλώσεων με τις πραγματικές τους επιλογές.

Ο όρος δειγματοληψία αναφέρεται στη διαδικασία συλλογής και ανάλυσης συγκεκριμένων χαρακτηριστικών από μια επιλεγμένη ομάδα ατόμων ενός πληθυσμού, γνωστή ως δείγμα. Η δειγματοληψία στοχεύει στο να αντλήσει πληροφορίες για τον ευρύτερο πληθυσμό μέσω της εξέτασης του δείγματος. Η εγκυρότητα των συμπερασμάτων που αφορούν τον πληθυσμό εξαρτάται από την τεχνική δειγματοληψίας και την ποιότητα του δείγματος. Οι μέθοδοι δειγματοληψίας διακρίνονται σε πιθανοτικές (αντιπροσωπευτικές) και μη πιθανοτικές (δειγματοληψία κρίσης) κατηγορίες.

Στην παρούσα μελέτη εφαρμόστηκε πιθανοτική δειγματοληψία, η οποία βασίζεται στους κανόνες της πιθανότητας, επιτρέποντας τη γενίκευση των ευρημάτων από το δείγμα στο συνολικό πληθυσμό και τον υπολογισμό του σφάλματος της εκτίμησης. Περιλαμβάνει τεχνικές όπως:

- Απλή τυχαία δειγματοληψία, που παρέχει ίσες πιθανότητες συμμετοχής σε κάθε μέλος του πληθυσμού. Απαιτεί την καταγραφή του πληθυσμού σε έναν αριθμημένο κατάλογο και την τυχαία επιλογή δειγμάτων.
- Συστηματική δειγματοληψία, που συνεπάγεται την επιλογή δειγμάτων από τον κατάλογο του πληθυσμού σε τακτικά διαστήματα, εφαρμόζοντας έναν κανόνα επιλογής βάσει της σχέσης μεγέθους πληθυσμού προς μέγεθος δείγματος.
- Στρωματοποιημένη δειγματοληψία, η οποία διασφαλίζει την αντιπροσώπευση όλων των τμημάτων του πληθυσμού, μειώνοντας το σφάλμα εκτίμησης και περιλαμβάνοντας αρκετά μέλη από κάθε υποομάδα του πληθυσμού.
- Δειγματοληψία κατά συστάδες, χρησιμοποιείται για έρευνες σε μεγάλες και γεωγραφικά διασπαρμένες περιοχές, όπου το δείγμα προκύπτει από την επιλογή ολόκληρων ομάδων του πληθυσμού.

Αντίθετα, η μη πιθανοτική δειγματοληψία εφαρμόζεται όταν η πιθανοτική δειγματοληψία δεν είναι εφικτή ή όταν απαιτείται γρήγορη εκτέλεση, όπως σε πιλοτικές έρευνες. Τα αποτελέσματα από

μια τέτοια έρευνα δεν είναι γενικεύσιμα και δεν επιτρέπουν τον υπολογισμό του σφάλματος εκτίμησης, περιορίζοντας τη χρησιμότητα και την εφαρμογή των συμπερασμάτων.

3.3 Μέθοδοι Ανάλυσης Δεδομένων

Η επιλογή της μεθόδου ανάλυσης δεδομένων εξαρτάται πρωτίστως από τη φύση των δεδομένων που έχουν συλλεχθεί. Τα δεδομένα χωρίζονται σε δύο βασικές κατηγορίες: ποσοτικά και ποιοτικά. Ποσοτικά είναι εκείνα τα δεδομένα που αφορούν μετρήσιμες τιμές, όπως αριθμοί και ποσότητες, ενώ ποιοτικά είναι τα δεδομένα που βασίζονται σε υποκειμενικές ερμηνείες και δεν μπορούν να μετρηθούν άμεσα. Οι ποσοτικές μεταβλητές μπορούν να καταταχθούν περαιτέρω ως διακριτές ή συνεχείς.

Υπάρχουν διάφορες προσεγγίσεις για την ανάλυση δεδομένων, με τις κύριες να περιλαμβάνουν:

- Περιγραφική Ανάλυση, που αποτελεί την αφετηρία για κάθε αναλυτική εργασία, καθώς παρέχει μια αρχική επισκόπηση των δεδομένων μέσω του υπολογισμού βασικών στατιστικών μετρήσεων και της παρουσιάσής τους, ενίοτε μέσω γραφημάτων.
- Διερευνητική Ανάλυση, που εξερευνά τα κύρια χαρακτηριστικά και τις τάσεις των δεδομένων, βοηθώντας στη διαμόρφωση υποθέσεων και στην επιλογή της κατάλληλης μεθοδολογίας για περαιτέρω ανάλυση.
- Προγνωστική Ανάλυση, που αναλύει τις τάσεις και τα δεδομένα για την πρόβλεψη μελλοντικών καταστάσεων.

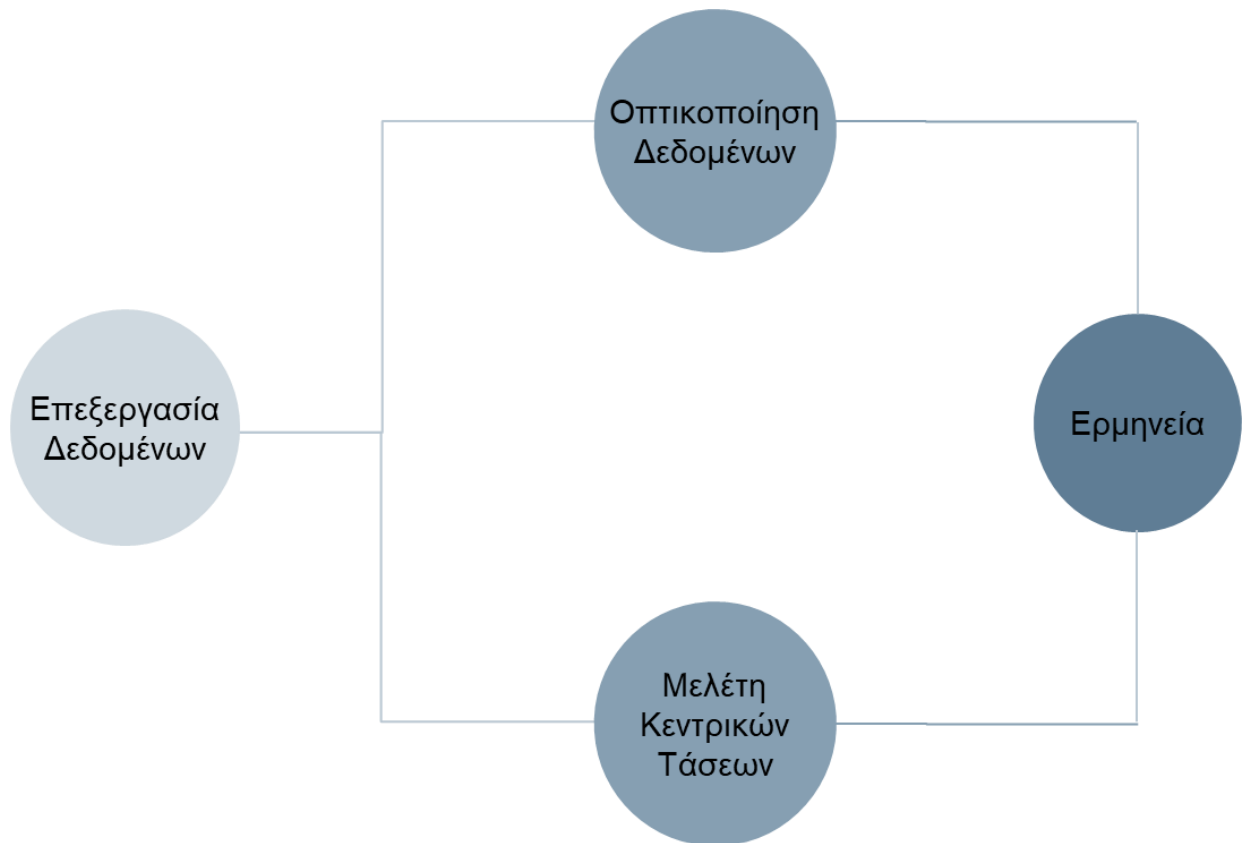
Η Περιγραφική Στατιστική στοχεύει στην οργάνωση και στην παρουσίαση πληροφοριών μέσω γραφημάτων και περιγραφικών τεχνικών.

Στο σχήμα που ακολουθεί, παρουσιάζεται συνοπτικά η διαδικασία της περιγραφικής στατιστικής που ακολουθείται για την ερμηνεία των συλλεγμένων δεδομένων. Όπως φαίνεται, μετά τη συλλογή των δεδομένων, ακολουθεί η επεξεργασία τους. Στη συνέχεια, διεξάγονται δύο διαφορετικές διαδικασίες:

A) Πραγματοποιείται η οπτικοποίηση των επεξεργασμένων δεδομένων.

B) Υπολογίζονται οι κεντρικές τάσεις, οι συχνότητες, οι συσχετίσεις και άλλα στατιστικά μεγέθη.

Εφόσον πραγματοποιήθηκαν οι παραπάνω διαδικασίες, συνδυαστικά μπορεί να προκύψει η ερμηνεία του συνόλου των δεδομένων που συλλέχθηκαν. Η ερμηνεία των συλλεγμένων δεδομένων αφορά την περιγραφή των βασικών χαρακτηριστικών ενός συνόλου δεδομένων χωρίς να εξάγονται συμπεράσματα για τον πληθυσμό αυτό.



Εικόνα 3.2 Μεθοδολογία περιγραφικής στατιστικής

Για τη διεξαγωγή της περιγραφικής στατιστικής ανάλυσης του δείγματος της παρούσας εργασίας, υπολογίστηκαν διάφορα βασικά στατιστικά μεγέθη όπως:

- Η μέση τιμή: στατιστικό μέτρο που υπολογίζεται προσθέτοντας όλες τις τιμές μιας μεταβλητής και διαιρώντας το αποτέλεσμα με τον αριθμό των παρατηρήσεων.
- Η διάμεσος: η τιμή που διαιρεί το σύνολο των δεδομένων σε δύο ίσα μέρη, έτσι ώστε το ήμισυ των τιμών να είναι μικρότερο και το άλλο ήμισυ μεγαλύτερο από αυτήν. Έχει χρησιμότητα σαν δείκτης, ειδικά σε δεδομένα με ακραίες τιμές.
- Η τυπική απόκλιση: μετρά το πόσο διασκορπισμένες είναι οι τιμές γύρω από τον μέσο όρο. Όσο μεγαλύτερη είναι η τυπική απόκλιση, τόσο μεγαλύτερη είναι η διακύμανση στο σύνολο των δεδομένων.
- Το εύρος: η διαφορά μεταξύ της μεγαλύτερης και της μικρότερης τιμής σε ένα σύνολο δεδομένων. Αντικατοπτρίζει την κατανομή των τιμών στο σύνολο.

- Η συχνότητα: ο αριθμός των φορών που εμφανίζεται μια συγκεκριμένη τιμή ή κατηγορία δεδομένων. Οι πίνακες συχνοτήτων και οι γραφικές παραστάσεις (όπως τα ιστογράμματα) χρησιμοποιούνται για την απεικόνιση των συχνοτήτων.

Μέσω αυτών των μεγεθών παρέχονται τα βασικά εργαλεία για την ανάλυση και την ερμηνεία των δεδομένων, βοηθώντας στην κατανόηση των τάσεων, των παραλλαγών, και των ανωμαλιών μέσα στο σύνολο των δεδομένων.

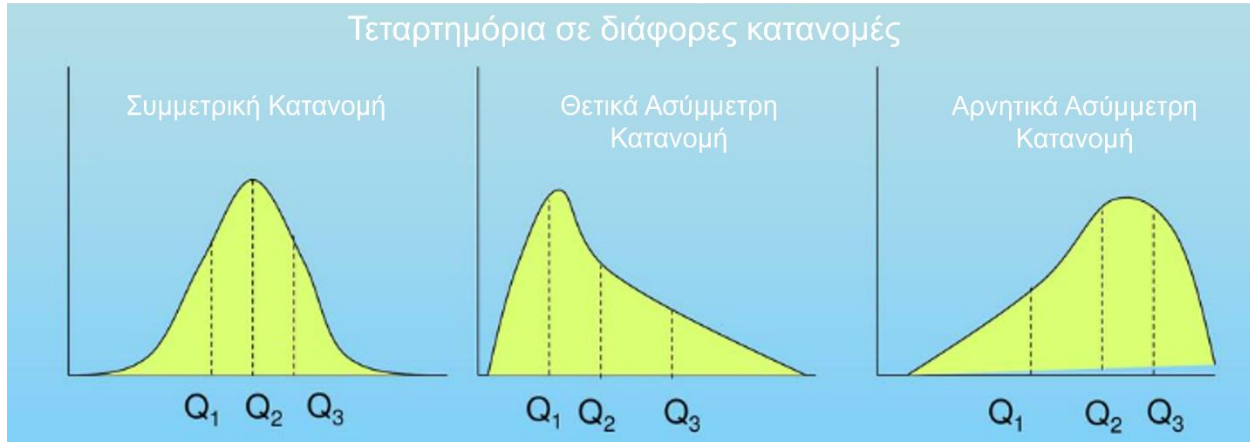
Για την οπτικοποίηση των αποτελεσμάτων της περιγραφικής στατιστικής χρησιμοποιούνται κάποιοι βασικοί τύποι διαγραμμάτων:

- ραβδογράμματα,
- κυκλικά διαγράμματα,
- ιστογράμματα,
- θηκογράμματα,
- χάρτες συσχέτισης
- χάρτες

Επιπλέον, για την επιτυχή ολοκλήρωση της περιγραφικής στατιστικής, το σύνολο των τιμών μπορεί να καταταχθεί αναλόγως το μέγεθος. Η συγκεκριμένη διαδικασία κατάταξης διευκολύνει τον εντοπισμό των οριακών τιμών στο εξεταζόμενο σύνολο δεδομένων και τον προσδιορισμό των θέσεων των στοιχείων του συνόλου με βάση τα εκατοστημόρια. Σημαντικότερος φαίνεται να είναι ο ρόλος των τεταρτημόριων στην ανάλυση και ερμηνεία της διανομής των δεδομένων.

Στην κανονική κατανομή, η τιμή που βρίσκεται στο πρώτο τεταρτημόριο αντιπροσωπεύει το σημείο κάτω από το οποίο βρίσκεται το 25% των παρατηρήσεων και πάνω από το οποίο βρίσκεται το 75% των παρατηρήσεων. Από την άλλη πλευρά, το δεύτερο τεταρτημόριο, που τοποθετείται στο κέντρο της διανομής, λειτουργεί ως διάμεσος, δηλαδή το σημείο όπου το 50% των παρατηρήσεων είναι είτε πριν είτε μετά από αυτό, διαχωρίζοντας ισομερώς το σύνολο των δεδομένων.

Όπως φαίνεται στο σχήμα που ακολουθεί, όσον αφορά τη συμμετρική κατανομή το Q1 απέχει από το Q2, όσο απέχει και το Q3 από το Q2. Σχετικά με τη θετικά ασύμμετρη κατανομή παρατηρείται ότι το Q1 είναι πιο κοντά στο Q2, ενώ το Q3 είναι πιο μακριά από το Q2. Τέλος, στην αρνητικά ασύμμετρη κατανομή το Q1 είναι μακριά από το Q2, ενώ το Q3 είναι κοντά στο Q2.



Εικόνα 3.3 Τεταρτημόρια σε διάφορες κατανομές

Για την κανονική κατανομή ισχύει ότι συγκεντρώνονται πολλές τιμές στη μέση, λίγες μεγάλες τιμές και λίγες μικρές τιμές στα άκρα. Η θετική ασυμμετρία συγκεντρώνει πολλές μικρές τιμές, κάποιες τιμές στη μέση και λίγες μεγάλες τιμές. Επίσης η μέση τιμή είναι πιο δεξιά από τη διάμεσο. Σχετικά με την αρνητική ασυμμετρία παρατηρείται ότι υπάρχουν πολλές μεγάλες τιμές, κάποιες τιμές στη μέση και λίγες μικρές τιμές. Σε αυτή την κατανομή ισχύει ότι η μέση τιμή είναι πιο μικρή σε σχέση με τη διάμεσο.

Να σημειωθεί ότι όλες οι αναλύσεις και οι οπτικοποιήσεις του συνόλου των δεδομένων της παρούσας Διπλωματικής Εργασίας διεξήχθησαν μέσω του λογισμικού της R-studio, με χρήση της γλώσσας προγραμματισμού R.

Με την ολοκλήρωση της περιγραφικής στατιστικής του συνόλου των δεδομένων, ακολουθεί η ανάλυση των δεδομένων. Υπάρχουν αρκετές μέθοδοι για την ανάλυση των δεδομένων. Μεταξύ των βασικών μεθόδων ανάλυσης δεδομένων συγκαταλέγονται:

- Ανάλυση Συστάδων, η οποία ομαδοποιεί τα δεδομένα σε ομάδες με βάση τις ομοιότητές τους.
- Ανάλυση Παλινδρόμησης, για την εκτίμηση των σχέσεων μεταξύ μεταβλητών.
- Ανάλυση Κόορτης, για την παρακολούθηση συγκεκριμένων ομάδων ατόμων στο χρόνο.
- Νευρωνικά Δίκτυα, για την επεξεργασία και ανάλυση πολύπλοκων δεδομένων.
- Παραγοντική Ανάλυση, για την αναγνώριση υποκείμενων διαστάσεων στα δεδομένα.
- Data Mining, για την ανακάλυψη προτύπων και σχέσεων σε μεγάλα σύνολα δεδομένων.

Στην περίπτωση της παρούσας Διπλωματικής Εργασίας, χρησιμοποιήθηκε η Ανάλυση Συστάδων για την επεξεργασία των συλλεγμένων δεδομένων από τα ερωτηματολόγια.

3.4 Μέθοδοι Συσταδοποίησης

Η ανάλυση συστάδων είναι μία διαδικασία αυτόματης αναγνώρισης ομάδων αντικειμένων βασιζόμενη στην ομοιότητά τους. Η βασική πρόκληση είναι η ενίσχυση της ομοιότητας εντός των αντικειμένων που υποτίθεται ότι ανήκουν στην ίδια συστάδα, ενώ παράλληλα η μείωση της ομοιότητας μεταξύ των αντικειμένων διαφορετικών συστάδων. Αυτό επιτυγχάνεται μέσω μιας μεθόδου ταξινόμησης χωρίς επίβλεψη, όπου η κατηγοριοποίηση των αντικειμένων στηρίζεται κυρίως σε πληροφορίες που παρέχονται από τα ίδια τα αντικείμενα. Αυτή η διαδικασία είναι σημαντική για την εξόρυξη δεδομένων.

Στο πεδίο της εξόρυξης δεδομένων, κεντρικά ερωτήματα περιλαμβάνουν τον καθορισμό του τρόπου συσταδοποίησης των δεδομένων και την αναγνώριση των αντικειμένων που αποκλίνουν έντονα από τις συστάδες. Αν και η ανάλυση συστάδων έχει μελετηθεί εκτενώς, το ενδιαφέρον για την ανακάλυψη νέων μεθόδων ομαδοποίησης έχει αναζωπυρωθεί, με τις πρόσφατες ερευνητικές προσπάθειες να μην περιορίζονται μόνο στην ανάπτυξη νέων στρατηγικών αλλά και στην εξέταση της επεκτασιμότητας και της διαχείρισης υψηλών διαστάσεων. Η κατανόηση των νέων προσεγγίσεων, των πλεονεκτημάτων και των περιορισμών τους αποτελεί μια πρόκληση. Ωστόσο, για μια αποτελεσματική τεχνική ομαδοποίησης στην εξόρυξη δεδομένων, υπάρχουν κάποιες βασικές απαιτήσεις που πρέπει να πληρούνται.

Οι βασικές απαιτήσεις για μια αποδοτική τεχνική συσταδοποίησης περιλαμβάνουν την ικανότητα να επεκτείνεται σε μεγάλες βάσεις δεδομένων, να προσαρμόζεται σε διαφορετικούς τύπους δεδομένων, να ανιχνεύει ομάδες με ποικίλα μορφολογικά χαρακτηριστικά, να λειτουργεί με ελάχιστες απαιτήσεις εισόδου, να είναι ανθεκτική στο θόρυβο, να μην επηρεάζεται από τη σειρά εισαγωγής των δεδομένων και να μπορεί να διαχειρίζεται δεδομένα υψηλών διαστάσεων. Αναλυτικά:

- **Επεκτασιμότητα:** Η τεχνική οφείλει να είναι εφικτή στην εφαρμογή σε μεγάλους όγκους δεδομένων με την απόδοση να μειώνεται γραμμικά με την αύξηση του όγκου των δεδομένων.
- **Ευελιξία:** Θα πρέπει να υποστηρίζει διάφορους τύπους δεδομένων, όπως αριθμητικά, boolean ή κατηγορικά.
- **Ανακάλυψη διαφορετικών σχημάτων συστάδων:** Κρίσιμη απαίτηση, ιδιαίτερα για την ανάλυση χωρικών δεδομένων, επιτρέποντας την ανίχνευση ομάδων πέρα από τα συνήθως σφαιρικά σχήματα.

- Ελάχιστη εισαγωγική παράμετρος: Απαιτείται λιγοστή προκαταρκτική γνώση για την αποτελεσματική ομαδοποίηση, αντιμετωπίζοντας το πρόβλημα πολλών αλγορίθμων που απαιτούν περίπλοκες παραμέτρους.
- Ανθεκτικότητα στο θόρυβο: Είναι ζωτικής σημασίας να μπορεί ο αλγόριθμος να διαχειρίζεται επιτυχώς σενάρια με υψηλό επίπεδο θορύβου.
- Αδιαφορία στη σειρά των δεδομένων: Η ομαδοποίηση πρέπει να παρέχει σταθερά αποτελέσματα, ανεξαρτήτως της σειράς παρουσίασης των δεδομένων.
- Κλιμάκωση σε υψηλές διαστάσεις: Η δυνατότητα επεξεργασίας δεδομένων με υψηλές διαστάσεις είναι απαραίτητη για την αντιμετώπιση πραγματικών σεναρίων.

Έως σήμερα, δεν έχει βρεθεί κάποιος αλγόριθμος που να έχει την ικανότητα να ικανοποιεί απόλυτα όλες τις προηγουμένως αναφερθείσες απαιτήσεις. Σημαντική είναι η γνώση των δυνατοτήτων και των χαρακτηριστικών των διαθέσιμων αλγορίθμων, ώστε να επιλεγεί εκείνος που ταιριάζει καλύτερα στις ανάγκες ενός συγκεκριμένου προβλήματος συσταδοποίησης. Παρ' όλα αυτά, υπάρχουν σύγχρονες τεχνικές συσταδοποίησης που δείχνουν εξαιρετική βελτίωση

Η διαδικασία της συσταδοποίησης συνίσταται στην ομαδοποίηση σημείων δεδομένων σε συγκεκριμένο αριθμό ομάδων βάσει της ομοιότητάς τους, ώστε τα δεδομένα εντός μιας ομάδας να είναι περισσότερο παρόμοια μεταξύ τους σε σύγκριση με δεδομένα από άλλες ομάδες.

Η βασική διάκριση της συσταδοποίησης από άλλες μεθόδους μηχανικής μάθησης, όπως η ταξινόμηση και η παλινδρόμηση, είναι πως εστιάζει στο είδος των δεδομένων που χρησιμοποιούνται για την εκπαίδευση των μοντέλων. Αυτή η διαφορά καθορίζει εάν η τεχνική είναι υπό επίβλεψη ή χωρίς επίβλεψη. Ενώ η ταξινόμηση και η παλινδρόμηση χρησιμοποιούν δεδομένα με συγκεκριμένες ετικέτες για την εκπαίδευση (εποπτευόμενη μάθηση), η συσταδοποίηση λειτουργεί με δεδομένα χωρίς προκαθορισμένες ετικέτες (μη εποπτευόμενη μάθηση).

Για την επιλογή της πιο κατάλληλης τεχνικής συσταδοποίησης, συνιστάται συνήθως η σύγκριση ανάμεσα σε πέντε διάφορους αλγορίθμους: k-means, φασματική συσταδοποίηση, ιεραρχική συσταδοποίηση, χωρική συσταδοποίηση βάσει πυκνότητας (DBSCAN), και Birch. Στην παρούσα Διπλωματική Εργασία διεξάχθηκε φασματική συσταδοποίηση. Από αυτή τη σύγκριση προκύπτουν τρεις κύριες κατηγορίες αλγορίθμων συσταδοποίησης:

- Κατάτμηση

- Ιεραρχικές διαδικασίες
- Βασισμένες στην πυκνότητα

Οι αλγόριθμοι κατάτμησης κατανέμουν τα δεδομένα σε έναν προκαθορισμένο αριθμό ομάδων, αξιοποιώντας την ομοιότητα ή την απόσταση ανάμεσα στα επιμέρους δείγματα. Από την άλλη πλευρά, οι ιεραρχικοί αλγόριθμοι δομούν τις ομάδες σε έναν ιεραρχικό σχηματισμό. Αλγόριθμοι που εστιάζουν στην πυκνότητα εντοπίζουν περιοχές μεγάλης πυκνότητας ανάμεσα στα δείγματα για τη δημιουργία συστάδων, ενώ οι περιοχές με χαμηλή πυκνότητα λειτουργούν ως διαχωριστικές γραμμές μεταξύ αυτών των ομάδων.

Υπάρχει μια πληθώρα αλγορίθμων για τη συσταδοποίηση δεδομένων, οι οποίοι γενικά μπορούν να κατηγοριοποιηθούν σε τέσσερις βασικές κατηγορίες: τεχνικές κατάτμησης, ιεραρχικές τεχνικές, τεχνικές βασισμένες στην πυκνότητα, και τεχνικές βασισμένες σε πλέγμα. Η αξιολόγηση και η ταξινόμηση αυτών των τεχνικών συσταδοποίησης διενεργείται μέσω εμπειρικών μελετών που εξετάζουν εκπροσωπητικές μεθόδους από κάθε κατηγορία.

Η αποτελεσματικότητα των αλγορίθμων συσταδοποίησης αξιολογείται εμπειρικά μέσω της χρήσης διαφόρων μεθόδων όπως ο αλγόριθμος k-means, χρησιμοποιώντας δισδιάστατα χωρικά δεδομένα για την ευκολία της οπτικής αξιολόγησης. Η επιλογή δισδιάστατων δεδομένων δικαιολογείται από την ανθρώπινη ικανότητα της αναγνώρισης ομάδων διαισθητικά, κάτι που γίνεται πιο δύσκολο σε περιπτώσεις δεδομένων με υψηλότερες διαστάσεις.

Όσον αφορά τις τεχνικές κατάτμησης, οι μέθοδοι αυτοί χωρίζουν τα δεδομένα σε k ομάδες ξεκινώντας από ένα αρχικό σετ n αντικειμένων. Η λογική πίσω από την κατάτμηση είναι άμεσα κατανοητή, με τη διαδικασία διαχωρισμού να ακολουθεί μια σταθερή προσέγγιση για την επίτευξη των ιδανικών κριτηρίων. Οι πιο γνωστές και ευρέως χρησιμοποιούμενες μέθοδοι κατάτμησης είναι οι k-means και k-medoids, όπου στην περίπτωση του k-means κάθε συστάδα αναπαρίσταται από το κεντροειδές της, ενώ στην περίπτωση του k-medoids από ένα από τα κεντρικά αντικείμενα της συστάδας.

Αφού καθοριστούν οι εκπρόσωποι των συστάδων, τα δεδομένα ανατίθενται σε αυτούς και στη συνέχεια διενεργείται μια επαναληπτική διαδικασία για την επιλογή νέων, βελτιωμένων εκπροσώπων, με τα δεδομένα να ανατίθενται εκ νέου μέχρι να μην παρατηρείται πλέον καμία αλλαγή. Αυτό ισχύει και για τη μέθοδο k-means (k-μέσων), η οποία χρησιμοποιήθηκε στην παρούσα Διπλωματική Εργασία, συνδυαστικά με τη φασματική συσταδοποίηση.

Οι μέθοδοι κατάτμησης παρουσιάζουν συνήθως ποιότητα στη συσταδοποίηση, αντιμετωπίζοντας ωστόσο κάποιες κοινές προκλήσεις, όπως η ανάγκη προκαταρκτικής γνώσης του αριθμού των

συστάδων πριν την εφαρμογή της συσταδοποίησης, η δυσκολία στην εύρεση συστάδων με διαφορετικά μεγέθη, και η περιορισμένη εφαρμογή της μεθόδου σε συστάδες που παρουσιάζουν κοίλες και σφαιρικές δομές.

Στη συσταδοποίηση με k-means, οι αποστάσεις μεταξύ ομοιογενών ομάδων μπορούν να υπολογιστούν προσδιορίζοντας την απόσταση κάθε εξεταζόμενου από το κέντρο κάθε συστάδας. Το κέντρο κάθε συστάδας αντιπροσωπεύεται από ένα διάνυσμα μέσων -που ονομάζεται κέντρο συστάδας- και αντιστοιχεί στις μεταβλητές που χρησιμοποιούνται για την ομαδοποίηση των υποψηφίων. Κάθε εξεταζόμενος εκχωρείται σε μία συστάδα, υπολογίζοντας την απόσταση μεταξύ του εξεταζόμενου και κάθε συστάδας και αναθέτοντας τον/την στην πλησιέστερη συστάδα (Anderberg, 1973).

Ο αλγόριθμος K-means είναι μια δημοφιλής μέθοδος συσταδοποίησης στη μηχανική μάθηση, χρησιμοποιούμενη για την ομαδοποίηση παρόμοιων δεδομένων σε k προκαθορισμένο αριθμό συστάδων. Ακολουθούν τα βασικά βήματα της διαδικασίας:

1. Επιλογή του αριθμού συστάδων k: Καθορίζεται ο αριθμός των συστάδων (k) στις οποίες θα ομαδοποιηθούν τα δεδομένα.
2. Αρχικοποίηση των κέντρων των συστάδων: Επιλέγονται τυχαία k σημεία από τα δεδομένα ως αρχικά κέντρα των συστάδων.
3. Ανάθεση κάθε σημείου στην πλησιέστερη συστάδα: Υπολογίζεται η απόσταση από κάθε σημείο δεδομένων προς τα κέντρα όλων των συστάδων και το σημείο ανατίθεται στη συστάδα με το πλησιέστερο κέντρο.
4. Επανυπολογισμός των κέντρων των συστάδων: Μετά την ανάθεση των σημείων, τα κέντρα των συστάδων ενημερώνονται υπολογίζοντας το μέσο όρο όλων των σημείων που ανήκουν στην εκάστοτε συστάδα.
5. Επανάληψη: Οι βήματα 3 και 4 επαναλαμβάνονται μέχρι να μην υπάρχει περαιτέρω μεταβολή στα κέντρα των συστάδων (ή μέχρι να συμβεί πολύ μικρή αλλαγή που δεν ξεπερνά ένα καθορισμένο κατώφλι) και έτσι ο αλγόριθμος συγκλίνει.

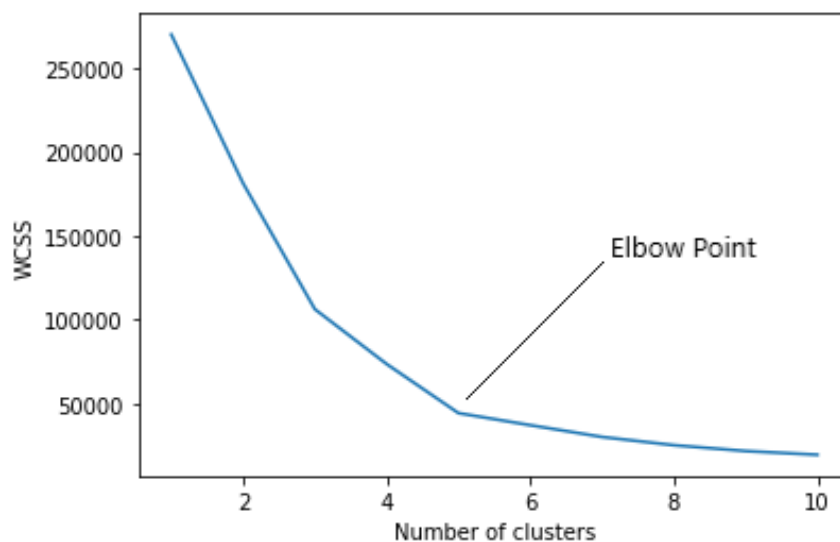
Αυτή η διαδικασία απλοποιεί το πρόβλημα της συσταδοποίησης διανέμοντας τα δεδομένα σε ομάδες με βάση την εγγύτητα τους προς τα κέντρα των συστάδων, με στόχο τη μείωση της διακύμανσης εντός των συστάδων και την αύξηση της διακύμανσης μεταξύ διαφορετικών συστάδων.

Προτού εφαρμοστεί ο αλγόριθμος k-means, απαιτείται ο καθορισμός του κατάλληλου αριθμού συστάδων (k) για την ανάλυση του συνόλου δεδομένων. Για την εύρεση του ιδανικού αριθμού συστάδων, χρησιμοποιούνται κυρίως τρεις μέθοδοι:

- Μέθοδος του Αγκώνα (Elbow method)
- Μέθοδος της Σιλουέτας (Silhouette method)
- Μέθοδος Gap (Gap static method)

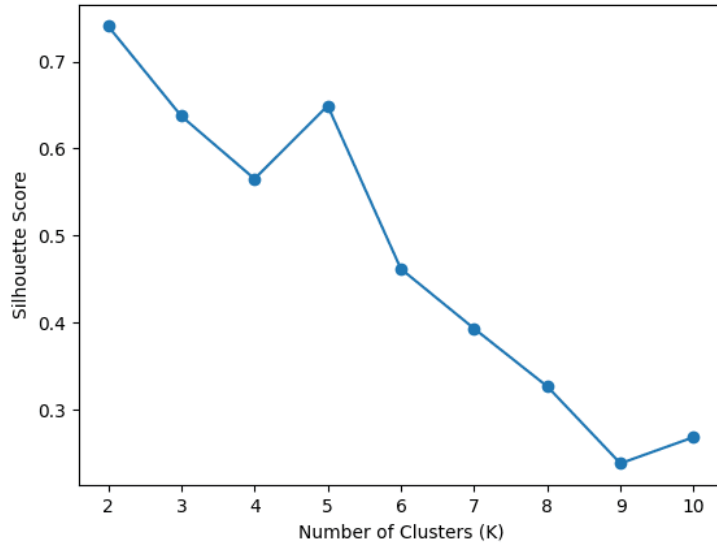
Οι μέθοδοι του Αγκώνα και της Σιλουέτας ακολουθούν μια άμεση προσέγγιση, ενώ η μέθοδος Gap βασίζεται στην εφαρμογή στατιστικών αναλύσεων.

Η μέθοδος του Αγκώνα υπολογίζει τη συνολική εντός-συστάδας αθροιστική απόσταση (within-cluster sum of squares, WCSS) για διάφορους αριθμούς συστάδων και επιλέγει το σημείο όπου η προσθήκη επιπλέον συστάδων δεν βελτιώνει σημαντικά την WCSS. Αυτό συνήθως παρατηρείται σαν ένας "αγκώνας" στο γράφημα της WCSS σε σχέση με τον αριθμό των συστάδων.



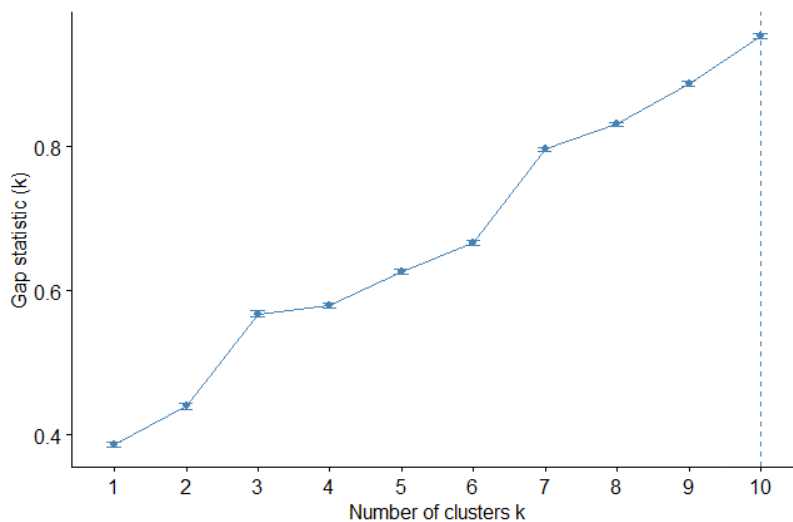
Εικόνα 3.4 Παράδειγμα Elbow Method

Η μέθοδος Silhouette μετρά πόσο καλά έχει γίνει η ανάθεση κάθε σημείου σε μια συστάδα. Ο δείκτης silhouette ποικίλλει από -1 έως 1, όπου ένας υψηλός δείκτης υποδεικνύει ότι το σημείο είναι καλά τοποθετημένο στη συστάδα του. Για την εκτίμηση των k κέντρων, εφαρμόζεται η μέθοδος Silhouette και ο υπολογισμός του δείκτη για ένα σύνολο k κέντρων. Συγκριτικά, επιλέγεται ο δείκτης που βρίσκεται πιο κοντά στο 1.



Εικόνα 3.5 Παράδειγμα Silhouette Method

Η μέθοδος Gap Statistic συγκρίνει την εσωτερική συνοχή των συστάδων σε σχέση με την αναμενόμενη τυχαία συνοχή. Σε αυτή τη μέθοδο, για κάθε τιμή του k , υπολογίζεται το λογάριθμο άθροισμα των τετραγώνων αποστάσεων ($\log(SSD)$) και συγκρίνεται με το αναμενόμενο $\log(SSD)$ που θα προέκυπτε από τα τυχαία δεδομένα. Η διαφορά μεταξύ των δύο τιμών ονομάζεται gap statistic.



Εικόνα 3.6 Παράδειγμα Gap Method

Με την αύξηση των διαστάσεων, ο όγκος του χώρου των δεδομένων αυξάνεται εκθετικά, κάνοντας τα δεδομένα να διασπαρθούν. Αυτή η διασπορά καθιστά πολύ περίπλοκη την διαχείριση και κατηγοριοποίηση των δεδομένων. Επομένως, η εφαρμογή της τεχνικής συσταδοποίησης με τον

αλγόριθμο k-means, ο οποίος εξαρτάται από τον υπολογισμό των αποστάσεων μεταξύ των σημείων, δεν είναι αποτελεσματική σε περιπτώσεις μεγάλων διαστάσεων, λόγω της αυξημένης αραίωσης των δεδομένων, που φαίνεται να απέχουν πολύ μεταξύ τους. Έτσι, πριν την εφαρμογή του αλγορίθμου k-means στα δεδομένα, είναι ουσιαστικό να προηγηθεί μια διαδικασία μείωσης των διαστάσεων, όπως η τεχνική της φασματικής συσταδοποίησης.

Στη φασματική συσταδοποίηση, διανύονται τρία βασικά στάδια: η κατασκευή ενός γράφου που απεικονίζει τις ομοιότητες, η μετάθεση των δεδομένων σε έναν χώρο με λιγότερες διαστάσεις, και η τελική συσταδοποίηση των δεδομένων. Αναλυτικά, για ένα σύνολο σημείων σε έναν χώρο μεγαλύτερων διαστάσεων, η διαδικασία περιλαμβάνει τα εξής βήματα:

1. Δημιουργία ενός πίνακα που απεικονίζει τις αποστάσεις μεταξύ των σημείων
2. Μετατροπή του πίνακα αποστάσεων σε έναν πίνακα ομοιότητας S (similarity matrix)
3. Υπολογισμός του πίνακα βαθμού D και δημιουργία του Λαπλασιανού πίνακα $L = D - S$
4. Εύρεση των ιδιοτιμών και των αντίστοιχων ιδιοδιανυσμάτων του L
5. Συγκρότηση ενός πίνακα από τα ιδιοδιανύσματα που αντιστοιχούν στις k μεγαλύτερες ιδιοτιμές
6. Κανονικοποίηση των διανυσμάτων που προκύπτουν
7. Οργάνωση των δεδομένων σε συστάδες μέσα στο χώρο k διαστάσεων.

Πριν γίνει έναρξη της φασματικής συσταδοποίησης είναι απαραίτητη η εξασφάλιση αναπαραγωγιμότητας των αποτελεσμάτων. Χρησιμοποιείται η συνάρτηση `set.seed()`, ώστε να διασφαλιστούν συνεπή αποτελέσματα σε διαφορετικές εκτελέσεις. Αυτή η πρακτική αποτελεί σημαντικό βήμα για στατιστικές αναλύσεις που περιλαμβάνουν την παραγωγή τυχαίων αριθμών, όπως η αρχικοποίηση των κέντρων σε μεθόδους συσταδοποίησης μέσω του αλγορίθμου k-μέσων, εξασφαλίζοντας ίδια αποτελέσματα σε κάθε εκτέλεση. Με αυτόν τον τρόπο ορίζεται ο αριθμός έναρξης για τη γεννήτρια τυχαίων αριθμών. Όταν ορίζεται ένα συγκεκριμένο `seed` πριν από τυχαίες επιλογές ή δειγματοληψίες, εξασφαλίζεται ότι οι τυχαίες επιλογές θα είναι οι ίδιες κάθε φορά που εκτελείται ο κώδικας και με την ίδια ακριβώς σειρά. Αυτό συνεπάγεται ότι αν επαναληφθεί η ίδια ανάλυση με το ίδιο `seed`, θα εξαχθούν τα ίδια ακριβώς αποτελέσματα.

Στη συνέχεια, για τη δημιουργία του πίνακα αποστάσεων μεταξύ των σημείων, μπορούν να χρησιμοποιηθούν διάφορα μέτρα υπολογισμού των αποστάσεων. Στην επεξεργασία δεδομένων και την ανάλυση συστάδων, ο υπολογισμός του πίνακα αποστάσεων είναι θεμελιώδης για την

αξιολόγηση της ομοιότητας ή διαφορετικότητας μεταξύ των δεδομένων. Ανάλογα με τη φύση και τα χαρακτηριστικά των δεδομένων, μπορεί να χρησιμοποιηθεί μια πληθώρα μετρήσεων απόστασης. Εδώ παρουσιάζονται κάποιες από τις συνηθέστερες μετρήσεις απόστασης και πότε χρησιμοποιούνται:

- **Ευκλείδεια Απόσταση (Euclidean Distance):** Είναι η πιο γνωστή και ευρέως χρησιμοποιούμενη μέτρηση απόστασης. Αντιστοιχεί στην "κανονική" απόσταση μεταξύ δύο σημείων σε έναν ευκλείδειο χώρο. Χρησιμοποιείται συνήθως όταν τα δεδομένα είναι συνεχή και οι διαστάσεις έχουν την ίδια σημασία.
- **Απόσταση Μανχάταν (Manhattan Distance):** Επίσης γνωστή ως απόσταση L1, είναι η άθροιση των απόλυτων διαφορών των συντεταγμένων. Χρησιμοποιείται συχνά σε περιπτώσεις όπου οι διαφορές μεταξύ των διαστάσεων είναι εξίσου σημαντικές και όταν τα δεδομένα διαφέρουν σε τάξη μεγέθους.
- **Απόσταση Chebyshev:** Η μέγιστη διαφορά μεταξύ δύο σημείων σε όλες τις διαστάσεις. Χρησιμοποιείται όταν η μέγιστη διαφορά είναι πιο σημαντική από τις μικρότερες διαφορές.
- **Ομοιότητα Cosine:** Αντί για απόσταση, μετρά την ομοιότητα μεταξύ δύο διανυσμάτων με βάση το συνημίτονο της γωνίας μεταξύ τους. Είναι ιδιαίτερα χρήσιμη σε δεδομένα με χαρακτήρες, καθώς λαμβάνει υπόψη την κατεύθυνση των διανυσμάτων αλλά όχι το μέγεθος τους.
- **Απόσταση Χάμμινγκ (Hamming Distance):** Μετρά τον αριθμό των θέσεων στις οποίες τα αντίστοιχα σύμβολα σε δύο συμβολοσειρές ή δυαδικά διανύσματα διαφέρουν. Χρησιμοποιείται συχνά στην επεξεργασία και ανάλυση δυαδικών δεδομένων ή σε σενάρια κωδικοποίησης.
- **Απόσταση Μινκόφσκι (Minkowski Distance):** Γενίκευση των αποστάσεων Ευκλείδειας και Μανχάταν, όπου μία παράμετρος p καθορίζει το είδος της απόστασης. Για $p=2$, αντιστοιχεί στην Ευκλείδεια απόσταση, ενώ για $p=1$, στην απόσταση Μανχάταν. Επιτρέπει την προσαρμογή της μέτρησης απόστασης ανάλογα με τις ανάγκες της εφαρμογής.

Κάθε μέτρηση απόστασης έχει τις ιδιαιτερότητές της και επιλέγεται με βάση τα χαρακτηριστικά του συνόλου δεδομένων και τους στόχους της ανάλυσης. Η επιλογή της κατάλληλης απόστασης μπορεί να επηρεάσει σημαντικά την αποτελεσματικότητα των αποτελεσμάτων και την ανάλυση συστάδων.

Στην παρούσα Διπλωματική Εργασία, ο υπολογισμός των αποστάσεων μεταξύ των παρατηρήσεων πραγματοποιήθηκε με χρήση της απόστασης Μανχάταν: Η απόσταση Μανχάταν μεταξύ δύο σημείων είναι το άθροισμα των απόλυτων διαφορών των συντεταγμένων τους.

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (1)$$

Υπάρχουν πολλοί λόγοι για τους οποίους επιλέχθηκε η απόσταση Manhattan αντί για άλλες μετρήσεις, όπως η Ευκλείδεια απόσταση, για τους παρακάτω λόγους:

- Ρεαλιστική Μοντελοποίηση σε πραγματικούς χώρους: Σε ορισμένες περιπτώσεις, η απόσταση Manhattan μπορεί να αντιπροσωπεύει πιο ακριβώς τον τρόπο που η απόσταση "λειτουργεί" στον πραγματικό κόσμο. Για παράδειγμα, σε αστικές περιοχές όπου οι δρόμοι τέμνονται κάθετα, η απόσταση που πρέπει να διανύσει κανείς περπατώντας ή οδηγώντας αντικατοπτρίζεται καλύτερα με τη μέτρηση Manhattan παρά με την Ευκλείδεια.
- Κατανομή Σπανιότητας (Sparsity): Η απόσταση Manhattan λειτουργεί καλύτερα από την ευκλείδεια όταν τα δεδομένα παρουσιάζουν σπανιότητα. Δηλαδή, σε σενάρια όπου τα δεδομένα περιέχουν πολλά μηδενικά ή μικρές τιμές, η Manhattan απόσταση μπορεί να δώσει πιο ακριβείς μετρήσεις ομοιότητας μεταξύ των δειγμάτων.
- Ανθεκτικότητα στις Ακραίες Τιμές (Outliers): Η απόσταση Manhattan είναι λιγότερο ευαίσθητη στις ακραίες τιμές σε σύγκριση με την ευκλείδεια απόσταση. Αυτό σημαίνει ότι μπορεί να παρέχει πιο σταθερά αποτελέσματα σε περιπτώσεις όπου τα δεδομένα περιέχουν θορύβο ή ακραίες τιμές.
- Ευαισθησία στην Κλίμακα και την Παραμόρφωση: Η απόσταση Manhattan μπορεί να είναι λιγότερο ευαίσθητη σε διαφορετικές κλίμακες διαστάσεων ή σε ανωμαλίες στα δεδομένα, κάνοντάς την κατάλληλη για εφαρμογές όπου οι διαστάσεις δεν έχουν ομοιόμορφη σημασία.
- Διαφορετικές Γεωμετρίες στα Δεδομένα: Σε κάποιες περιπτώσεις, η γεωμετρική δομή των δεδομένων μπορεί να αντιστοιχίζεται καλύτερα με την απόσταση Manhattan από ό,τι με την ευκλείδεια. Αυτό συμβαίνει ιδίως σε προβλήματα όπου οι διαστάσεις είναι ορθογώνια διακριτές και ανεξάρτητες.

- Υπολογιστική Απλότητα: Η απόσταση Manhattan μπορεί να είναι υπολογιστικά λιγότερο απαιτητική από την ευκλείδεια απόσταση, καθώς δεν απαιτεί την υπολογιστικά επίπονη διαδικασία της ρίζας. Αυτό την καθιστά πιο αποτελεσματική σε μεγάλα σετ δεδομένων.

Για τον υπολογισμό του πίνακα ομοιότητας χρειάζεται και ο ορισμός της παραμέτρου σ . Μία συνηθισμένη πρακτική για τον υπολογισμό του σ είναι η χρήση της διάμεσης των αποστάσεων μεταξύ όλων των ζευγών των σημείων στο σύνολο δεδομένων. Αυτός ο τρόπος προσφέρει μια καλή ισορροπία, καθώς η διάμεση αποτελεί μια κεντρική τάση των αποστάσεων, αποφεύγοντας τις ακραίες τιμές που μπορεί να παραμορφώσουν την κλίμακα ομοιότητας. Η τροποποίηση αυτής της τιμής επηρεάζει την κλίμακα ομοιότητας μεταξύ των σημείων και καθορίζει την επιρροή που έχει ένα σημείο στα άλλα στο πλαίσιο της ομοιότητας. Αυτό επιτρέπει τη μετατροπή των αποστάσεων σε τιμές ομοιότητας, όπου υψηλότερες τιμές σημαίνουν μεγαλύτερη ομοιότητα. Η χρήση του σ είναι κρίσιμη για την κατάλληλη κλιμάκωση των τιμών ομοιότητας.

Ο πίνακας ομοιότητας υπολογίζεται με την εφαρμογή του τύπου που ακολουθεί:

$$S_{ij} = e^{\left(-\frac{D_{ij}^2}{2\sigma^2}\right)} \quad (2)$$

Συνεπώς, ο πίνακας αποστάσεων μετασχηματίζεται σε έναν πίνακα ομοιότητας, όπου τα σημεία που βρίσκονται πιο κοντά έχουν μεγαλύτερη ομοιότητα. Η χρήση της εκθετικής συνάρτησης διασφαλίζει ότι η ομοιότητα μειώνεται εκθετικά με την αύξηση της απόστασης.

Ακολούθησε η κατασκευή του Λαπλασιανού πίνακα, με χρήση του πίνακα ομοιότητας. Αρχικά, δημιουργήθηκε ένας πίνακας βαθμίδων (degree matrix), υπολογίζοντας το άθροισμα κάθε γραμμής του πίνακα ομοιότητας και τοποθετώντας αυτά τα αθροίσματα στη διαγώνιο ενός διαγώνιου πίνακα. Ο πίνακας βαθμίδων είναι ένας διαγώνιος πίνακας που κάθε στοιχείο του διαγώνιου D_{ii} αντιπροσωπεύει το βαθμό της κορυφής i , δηλαδή το άθροισμα των βαρών όλων των ακμών που συνδέονται με αυτή την κορυφή.

$$d_i = \sum_{j|(i,j) \in E} w_{ij} \quad (3)$$

Ο Λαπλασιανός πίνακας υπολογίζεται αφαιρώντας τον πίνακα ομοιότητας από τον πίνακα βαθμίδων, όπως φαίνεται παρακάτω:

$$L = D - S \quad (4)$$

Ο μη κανονικοποιημένος Λαπλασιανός πίνακας δημιουργείται από την αφαίρεση του πίνακα συγγένειας από τον πίνακα βαθμίδων. Αυτός ο πίνακας έχει ιδιαίτερη σημασία στην ανάλυση γραφημάτων και στη φασματική συσταδοποίηση, καθώς αντικατοπτρίζει πώς η δομή των συνδέσεων μεταξύ των κορυφών επηρεάζει την ομοιότητα και την ικανότητα διαχωρισμού των κορυφών σε συστάδες.

Μέσω του Λαπλασιανού πίνακα, υπολογίστηκαν οι ιδιοτιμές και τα ιδιοδιανύσματα της μήτρας. Ο Λαπλασιανός πίνακας είναι συμμετρικός, έτσι διασφαλίζεται η βέλτιστη απόδοση των υπολογισμών. Από το σύνολο των ιδιοδιανυσμάτων επιλέχθηκαν τα πρώτα (μικρότερα) k ιδιοδιανύσματα, όπου k είναι ο αριθμός των επιθυμητών συστάδων. Τα ιδιοδιανύσματα επιλέχθηκαν βάσει των ιδιοτιμών τους και χρησιμοποιήθηκαν για να μετασχηματίσουν τα δεδομένα σε έναν νέο χώρο, με λιγότερες διαστάσεις.

Στη συνέχεια, κανονικοποιήθηκαν οι γραμμές του πίνακα που περιείχε τα επιλεγμένα ιδιοδιανύσματα, ώστε κάθε γραμμή να έχει μοναδιαίο μήκος. Αυτό βοηθά στην προετοιμασία των δεδομένων για τη συσταδοποίηση.

Τέλος, τα κανονικοποιημένα δεδομένα συσταδοποιήθηκαν με χρήση της μεθόδου k -means, όπως αυτή παρουσιάστηκε παραπάνω. Ο αλγόριθμος εκτελέστηκε για τα κανονικοποιημένα δεδομένα με κέντρα τα επιλεγμένα k μικρότερα ιδιοδιανύσματα του Λαπλασιανού πίνακα. Επίσης, ο αλγόριθμος εκτελέστηκε αρκετές φορές με μία δομή επανάληψης και με διαφορετικά αρχικά κέντρα, καταλήγοντας στην καλύτερη λύση.

Μέσω της υλοποίησης της Φασματικής Συσταδοποίησης σε συνδυασμό με τον αλγόριθμο k -means, η απόσταση μεταξύ των υποψηφίων προς συσταδοποίηση, δεν προσδιορίζεται από έναν απλό μέσο όρο αλλά από έναν σύνολο μέσων τιμών. Επίσης, αντί να κατατάσσονται οι εξεταζόμενοι σε μια συνεχή κλίμακα, κατανέμονται σε διάφορες διακριτές ομάδες ή συστάδες. Αυτές οι συστάδες μπορούν να χρησιμεύσουν ως εργαλεία για τη διαδικασία ορισμού των προτύπων, ενώνοντας τις συστάδες με τις κατηγορίες που καθορίζονται από τις διαδικασίες καθορισμού προτύπων και ανάπτυξης δοκιμών. Η λογική πίσω από τη χρήση της ανάλυσης συστάδων, αντιμετωπίζει δύο βασικούς περιορισμούς.

Αρχικά, η διαδικασία επικεντρώνεται στην ανάλυση των απαντήσεων των συμμετεχόντων στις έρευνες, περιορίζοντας την ικανότητα να θέτει πρότυπα που ξεπερνούν τις πραγματικές επιδόσεις των συμμετεχόντων. Αυτό σημαίνει ότι τα πρότυπα προκύπτουν από τις ενέργειες συγκεκριμένων ομάδων εξεταζόμενων και όχι από προκαθορισμένες προσδοκίες. Παρόλο που αυτός ο περιορισμός μπορεί να φαίνεται σημαντικός, είναι απίθανο μια δοκιμή να σχεδιαστεί τόσο

απομακρυσμένα από τα επίπεδα απόδοσης των εξεταζόμενων ώστε να μην ανταποκρίνεται σε πραγματικά πρότυπα απόδοσης.

Δεύτερον, οι τεχνικές συσταδοποίησης δημιουργούν ομάδες ανεξάρτητα από την πραγματική ύπαρξη διακριτών ομάδων εξεταζόμενων. Επομένως, απαιτείται ένα εξωτερικό κριτήριο, δηλαδή μια μεταβλητή που δεν εμπλέκεται στη συσταδοποίηση αλλά σχετίζεται με την πραγματική δομή των συστάδων, για να επιβεβαιωθεί η εγκυρότητα των αποτελεσμάτων και να διασφαλιστεί ότι οι συστάδες που προκύπτουν είναι πραγματικά διακριτές μεταξύ τους.

Σε αντίθεση με την παραδοσιακή χρήση της ανάλυσης συστάδων, ο στόχος όταν χρησιμοποιείται για τη διαδικασία ορισμού προτύπων δεν είναι η αποκάλυψη της "πραγματικής" δομής των δεδομένων αλλά η εύρεση της βέλτιστης διανομής του εξεταζόμενου πληθυσμού που ανταποκρίνεται στις εκ των προτέρων ομαδοποιήσεις επάρκειας. Όσον αφορά την επιλογή των μεταβλητών για τη συσταδοποίηση, υπάρχουν τουλάχιστον τρία σενάρια: η χρήση όλων των μεμονωμένων στοιχείων της δοκιμής, η χρήση βαθμολογιών παραγόντων που προέρχονται από την ανάλυση παραγόντων σε επίπεδο στοιχείου, και η χρήση υποβαθμολογιών που προκύπτουν από στοιχεία εντός των βασικών τομέων περιεχομένου της δοκιμής.

4 ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

4.1 Γενικά

Σε αυτό το κεφάλαιο θα περιγραφεί συνοπτικά το ερωτηματολόγιο και το πως διεξάχθηκε η έρευνα συλλογής των δεδομένων. Στη συνέχεια, θα παρουσιαστούν αναλυτικά οι μεταβλητές καθώς και η κωδικοποίηση τους όπως αυτή χρησιμοποιήθηκε για την επεξεργασία των δεδομένων στο λογισμικό της R-Studio. Θα γίνει λεπτομερής περιγραφική στατιστική του δείγματος και όλων των μεταβλητών που το αφορούν μέσα από διαγράμματα. Εν συνεχεία, θα αναλυθούν τα δεδομένα συνδυαστικά και με επίκεντρο το μέσο και το σκοπό της μετακίνησης. Θα παρουσιαστούν διαγράμματα και η διεξοδική περιγραφή τους.

4.2 Περιγραφή Ερωτηματολογίων

Τα δεδομένα που χρησιμοποιήθηκαν για τη διεξαγωγή της παρούσας έρευνας συλλέχθηκαν από πρόσφατη έρευνα αποκαλυπτόμενων προτιμήσεων που προωθήθηκε από την ΕΡΤ. Συγκεκριμένα, μοιράστηκαν 513 ερωτηματολόγια έρευνας ταξιδιών. Να αναφερθεί ότι η τελευταία έρευνα ταξιδιών στη μητροπολιτική περιοχή της Αθήνας πραγματοποιήθηκε το 2006 από τον ΟΑΣΑ.

Αρχικά, οι συμμετέχοντες απάντησαν σε ερωτήσεις δημογραφικού και κοινωνικοοικονομικού περιεχομένου. Αναλυτικότερα, ζητήθηκαν το φύλο, η ηλικία, το επίπεδο εκπαίδευσης, η εργασιακή κατάσταση, η εισοδηματική κατηγορία, η ζώνη κατοικίας και εάν υπάρχει στην κατοχή του εκάστοτε ατόμου ιδιωτικό αυτοκίνητο.

Στη συνέχεια, οι ερωτηθέντες κλήθηκαν να περιγράψουν πέντε ταξίδια που πραγματοποίησαν σε μία τυπική εργάσιμη ημέρα. Σε κάθε ταξίδι, θα έπρεπε να αναφερθεί ο σκοπός του ταξιδιού, ο χρόνος έναρξης του ταξιδιού, το μέσο μετακίνησης και η ζώνη προορισμού. Τα μέσα μετακίνησης των ταξιδιών μπορεί να είναι: 1: αυτοκίνητο, 2: ταξί, 3: λεωφορείο, 4: τραίνο (Μέσα Σταθερής Τροχιάς), 5: μηχανή, 6: ποδήλατο, 7: περπάτημα, 8: ηλεκτρικό πατίνι (e-scooter). Οι σκοποί των ταξιδιών μπορεί να είναι: 1: εργασία, 2: επιστροφή στο σπίτι, 3: εκπαίδευση, 4: αγορά, 5: αναψυχή, 6: υπηρεσία και 7: άλλο. Ο χρόνος ταξιδιού πρέπει να δίνεται σε χρονικό πλαίσιο 24 ωρών.

Επιπλέον, η μητροπολιτική περιοχή της Αθήνας χωρίστηκε σε 36 ζώνες, με μία από αυτές τις ζώνες να αναφέρεται στη μετάβαση σε μια ζώνη εκτός της περιοχής μελέτης. Μέσω αυτού του συστήματος διαχωρισμού σε ζώνες, οι αποκρινόμενοι ανέφεραν τη ζώνη κατοικίας τους χωρίς να καθορίζουν περαιτέρω την τοποθεσία τους. Επιπλέον, μια βασική υπόθεση που έγινε είναι ότι το πρώτο ταξίδι της ημέρας ξεκινά από το σπίτι. Ακόμα, γίνεται η υπόθεση ότι η ζώνη προορισμού

του ταξιδιού i είναι η ζώνη προέλευσης του ταξιδιού $i + 1$, δημιουργώντας μια αλυσίδα δραστηριοτήτων.

Η καινοτομία της διαδικασίας συλλογής των δεδομένων σχετίζεται και με τη στρατηγική διανομής. Χρησιμοποιήθηκαν οι ιστοσελίδες και οι ραδιοσυχνότητες της ΕΡΤ για να φτάσουν τα ερωτηματολόγια σε ένα ευρύτερο ακροατήριο. Η παρούσα έρευνα ήταν ανώνυμη και οι αποκρινόμενοι δεν υποχρεώθηκαν να παρέχουν όλα τα κοινωνικό-δημογραφικά τους στοιχεία. Λήφθηκε συγκατάθεση από όλους τους συμμετέχοντες στην έρευνα. (Chatziioannou, Tzouras, Tsigdinos & Keraptsoglou n.d.)

4.3 Μεταβλητές

Στον παρακάτω πίνακα παρουσιάζονται οι μεταβλητές της βάσης δεδομένων, με την αντίστοιχη περιγραφή τους, την κωδικοποίηση τους, το εύρος και τον τύπο των τιμών τους. Ο τύπος της μεταβλητής διακρίνει αν ανήκει στην κατηγορία των ποιοτικών ή ποσοτικών. Οι ποιοτικές μεταβλητές δεν είναι μετρήσιμες, ενώ οι ποσοτικές μπορούν να μετρηθούν με μια κατάλληλη κλίμακα. Οι ποιοτικές μεταβλητές χωρίζονται σε δύο κατηγορίες: ονομαστικές, όπου οι τιμές δεν υπόκεινται σε σύγκριση ή διαβάθμιση, και τακτικές, όπου οι τιμές ταξινομούνται με βάση μια κλίμακα που δείχνει αξίες μεγαλύτερες ή μικρότερες από άλλες. Τα κατηγορικά δεδομένα αντιστοιχούν σε μεταβλητές όπου οι τιμές ανήκουν σε κατηγορίες που δεν μπορούν να συγκριθούν ή να διαβαθμιστούν. Τα τακτικά δεδομένα, από την άλλη, ταξινομούνται με βάση μια κλίμακα, αποδεικνύοντας ότι ορισμένα έχουν αξία μεγαλύτερη ή μικρότερη από άλλα.

Πίνακας 4.1 Κωδικοποίηση μεταβλητών

Περιγραφή	Κωδικός	Εύρος	Τύπος
Ταυτότητα	pid	101-613	Scale
Φύλο	gender	0: female	Nominal (2 κλάσεις)
		1: male	
Ηλικία	age	16-84	Scale
Μορφωτικό Επίπεδο	education	1: primary school	Ordinal
		2: high school	
		3: bachelor	
		4: master or phd	
Επαγγελματική Κατάσταση	employment	1: inactive	Nominal (4 κλάσεις)
		2: unemployed	
		3: student	

		4: active	
Μηνιαίο Εισόδημα	income	0: no income	Ordinal
		1: 750 or less	
		2: 750-1500	
		3: 1500-2500	
		4: 2500 or more	
Ιδιοκτησία Αυτοκινήτου	car_own	0: no	Nominal (2 κλάσεις)
		1: yes	
Ζώνη Κατοικίας	home	1-36	Ordinal
Ζώνη Πρώτου Προορισμού	dest1	1-36	Ordinal
Σκοπός Πρώτης Μετακίνησης	purp1	1: work	Nominal (7 κλάσεις)
		2: return home	
		3: education	
		4: market	
		5: recreation	
		6: service	
		7: other	
Πρώτο Μέσο Μετακίνησης	mode1	1: car	Nominal (8 κλάσεις)
		2: taxi	
		3: bus	
		4: train	
		5: motorcycle	
		6: bicycle	
		7: walk	
		8: escooter	
Χρονική Διάρκεια Πρώτης Μετακίνησης	time1	0-22	Ordinal
Ζώνη Δεύτερου Προορισμού	dest2	1-36	Ordinal
Σκοπός Δεύτερης Μετακίνησης	purp2	1: work	Nominal (7 κλάσεις)
		2: return home	
		3: education	
		4: market	
		5: recreation	
		6: service	

		7: other	
Δεύτερο Μέσο Μετακίνησης	mode2	1: car	Nominal (8 κλάσεις)
		2: taxi	
		3: bus	
		4: train	
		5: motorcycle	
		6: bicycle	
		7: walk	
		8: scooter	
Χρονική Διάρκεια Δεύτερης Μετακίνησης	time2	0-22	Ordinal
Ζώνη Τρίτου Προορισμού	dest3	1-36	Ordinal
Σκοπός Τρίτης Μετακίνησης	purp3	1: work	Nominal (7 κλάσεις)
		2: return home	
		3: education	
		4: market	
		5: recreation	
		6: service	
		7: other	
Τρίτο Μέσο Μετακίνησης	mode3	1: car	Nominal (8 κλάσεις)
		2: taxi	
		3: bus	
		4: train	
		5: motorcycle	
		6: bicycle	
		7: walk	
		8: scooter	
Χρονική Διάρκεια Τρίτης Μετακίνησης	time3	0-22	Ordinal
Ζώνη Τέταρτου Προορισμού	dest4	1-36	Ordinal
Σκοπός Τέταρτης Μετακίνησης	purp4	1: work	Nominal (7 κλάσεις)
		2: return home	
		3: education	
		4: market	
		5: recreation	

		6: service	
		7: other	
Τέταρτο Μέσο Μετακίνησης	mode4	1: car	Nominal (8 κλάσεις)
		2: taxi	
		3: bus	
		4: train	
		5: motorcycle	
		6: bicycle	
		7: walk	
		8: scooter	
Χρονική Διάρκεια Τέταρτης Μετακίνησης	time4	0-22	Ordinal
Ζώνη Πέμπτου Προορισμού	dest5	1-36	Ordinal
Σκοπός Πέμπτης Μετακίνησης	purp5	1: work	Nominal (7 κλάσεις)
		2: return home	
		3: education	
		4: market	
		5: recreation	
		6: service	
		7: other	
Πέμπτο Μέσο Μετακίνησης	age	1: car	Nominal (8 κλάσεις)
		2: taxi	
		3: bus	
		4: train	
		5: motorcycle	
		6: bicycle	
		7: walk	
		8: scooter	
Χρονική Διάρκεια Πέμπτης Μετακίνησης	time5	0-22	Ordinal
Ζώνη Προέλευσης	InputID	1: athens_center	Nominal (36 κλάσεις)
		2: pagkrati	
		3: votanikos	
		4: sepolia	

		5: patisia	
		6: kipseli	
		7: goudi	
		8: chaidari	
		9: kamatero	
		10: airport	
		11: aigalew	
		12: east_att	
		13: virwnas	
		14: galatsi	
		15: zografou	
		16: dafni	
		17: west_att	
		18: hliopouli	
		19: ilion	
		20: kaisariani	
		21: kalithea	
		22: marousi	
		23: mosxato	
		24: herakleio	
		25: nea_smirni	
		26: nea_filadelfia	
		27: renti	
		28: p_faliro	
		29: papaXol	
		30: peiraias	
		31: perama	
		32: peristeri	
		33: salamina	
		34: filothei	
		35: chalandri	
		36: ektos	
Zώνη Προορισμού	TargetID	1: athens_center	

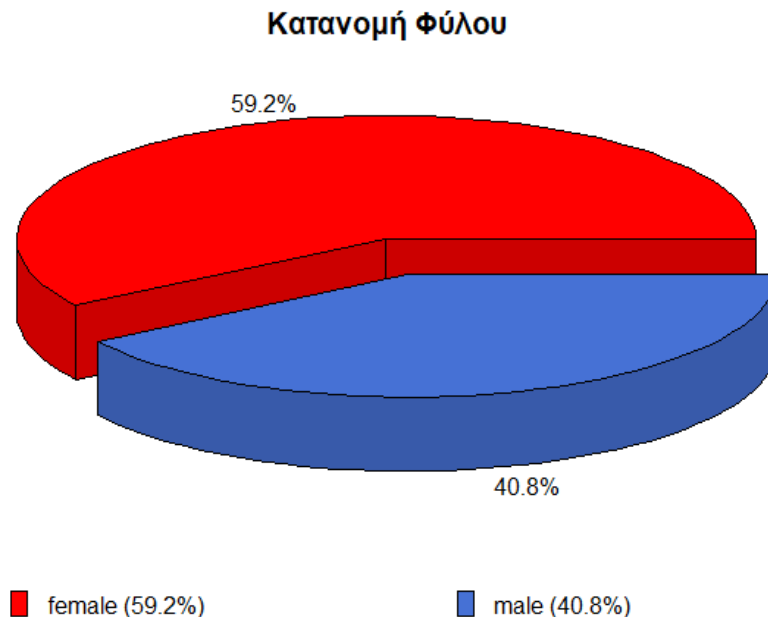
		2: pagkrati	Nominal (36 κλάσεις)
		3: votanikos	
		4: sepolia	
		5: patisia	
		6: kipseli	
		7: goudi	
		8: chaidari	
		9: kamatero	
		10: airport	
		11: aigalew	
		12: east_att	
		13: virwnas	
		14: galatsi	
		15: zografou	
		16: dafni	
		17: west_att	
		18: hliopouli	
		19: ilion	
		20: kaisariani	
		21: kalithea	
		22: marousi	
		23: mosxato	
		24: herakleio	
		25: nea_smirni	
		26: nea_filadelfia	
		27: renti	
		28: p_faliro	
		29: papaXol	
		30: peiraias	
		31: perama	
		32: peristeri	
		33: salamina	
		34: filothei	

		35: chalandri	
		36: ektos	
Απόσταση μεταξύ ζωνών Προέλευσης-Προορισμού	Distance	1405.958545 - 50517.84553	Ordinal

4.4 Περιγραφική Στατιστική

Στο παρόν υπό-κεφάλαιο, παρουσιάζεται μια λεπτομερής ανάλυση της περιγραφικής στατιστικής των μεταβλητών που συμπεριλαμβάνονται στο δείγμα. Η περιγραφική ανάλυση περιλαμβάνει γραφήματα, διαγράμματα και παρατηρήσεις που αποκαλύπτουν τη διακύμανση, τη διανομή και τη συσχέτιση μεταξύ των μεταβλητών. Αυτή η περιγραφή, αποσκοπεί στο να κατανοηθούν καλύτερα τα χαρακτηριστικά του δείγματος και οι σχέσεις μεταξύ των μεταβλητών. Τα γραφήματα και οι παρατηρήσεις που παρουσιάζονται εδώ θα αποτελέσουν τη βάση για την περαιτέρω ανάλυση και ερμηνεία των δεδομένων, καθώς και τη λήψη αποφάσεων βασισμένων στα αποτελέσματα της διερεύνησης τους.

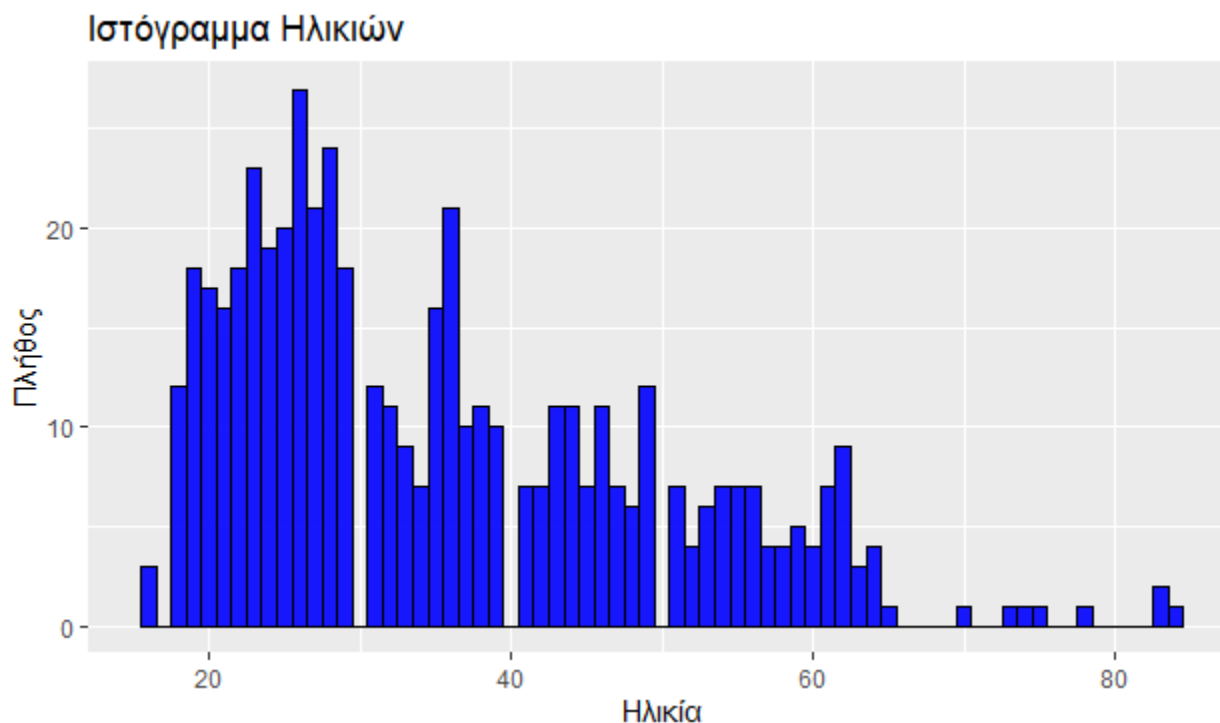
Στο διάγραμμα που ακολουθεί, φαίνεται η κατανομή του φύλου των συμμετεχόντων της έρευνας. Με κόκκινο χρώμα απεικονίζεται το κομμάτι του πληθυσμού που είναι γυναίκες και με μπλε χρώμα απεικονίζεται το πλήθος των αντρών.



Διάγραμμα 4.1 Ποσοστιαία κατανομή φύλου

Παρατηρείται ότι, το δείγμα αποτελείται κατά 59,2% από γυναίκες και κατά 40,8% από άντρες. Αναλυτικότερα, οι ερωτηθέντες αριθμητικά ήταν 300 γυναίκες και 207 άντρες. Το ποσοστό των γυναικών υπερτερεί των αντρών όμως και τα δύο ποσοστά είναι ικανοποιητικά για την παρούσα έρευνα.

Στη συνέχεια, παρουσιάζεται το ιστόγραμμα με τη διακύμανση των ηλικιών όλων των ατόμων που συμμετείχαν στο ερωτηματολόγιο.

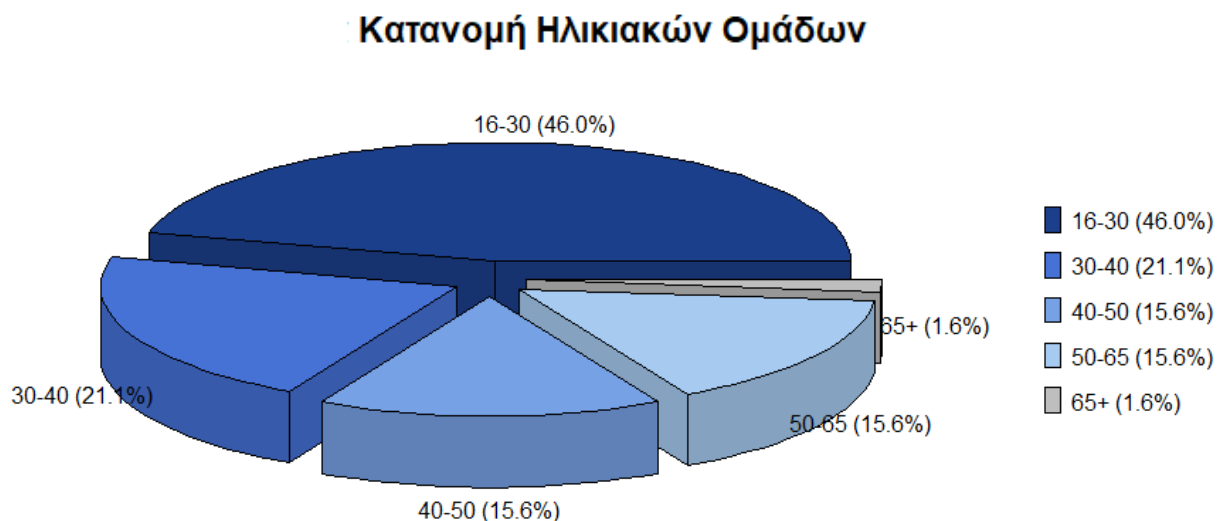


Διάγραμμα 4.2 Κατανομή ηλικιών

Μέσω του ιστογράμματος, προκύπτει ότι, το μεγαλύτερο μέρος του δείγματος συγκεντρώνεται στις ηλικίες μεταξύ 20 έως 30 χρονών. Στις ηλικίες 30 έως 60 χρονών υπάρχει αρκετή συμμετοχή στην έρευνα ενώ σε ηλικίες μεγαλύτερες από 60 χρονών οι συμμετέχοντες μειώνονται αισθητά. Παρατηρούνται, κενά σε συγκεκριμένες ηλικίες όπως 30, 40 και 50 ετών. Το ελάχιστο στη μεταβλητή «ηλικία» σημειώνεται στα 16 έτη και το μέγιστο στα 84. Τα ερωτηματολόγια χωρίς απάντηση στο ερώτημα της ηλικίας ήταν μόνο 4. Ο μέσος όρος των ηλικιών βρίσκεται στα 35,54 έτη.

Για τις ανάγκες της περιγραφικής στατιστικής του συνόλου των ερωτηματολογίων οι ηλικίες των ατόμων χωρίστηκαν σε 5 ηλικιακές ομάδες. Οι ηλικιακές ομάδες δημιουργήθηκαν ως εξής: 16 έως 30 έτη, 30 έως 40 έτη, 40 έως 50 έτη, 50 έως 65 έτη και άνω των 65 ετών.

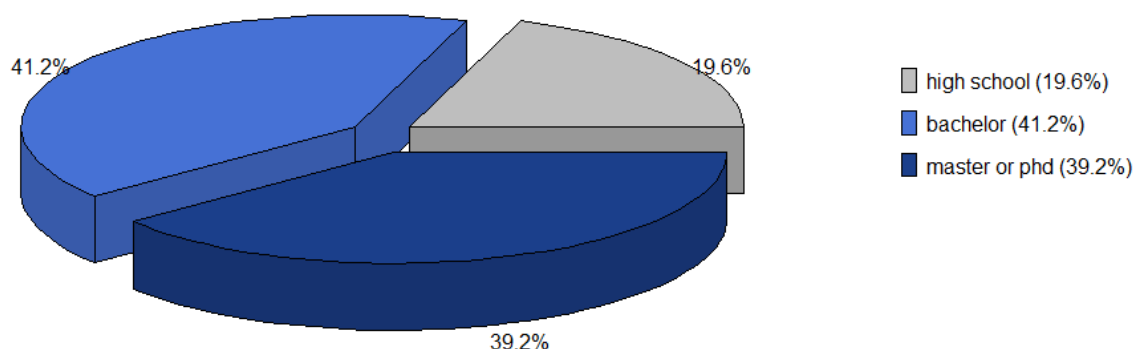
Στο παρακάτω διάγραμμα, γίνεται ποσοστιαία κατανομή των ηλικιακών ομάδων που δημιουργήθηκαν για καλύτερη στατιστική ανάλυση του δείγματος. Από το διάγραμμα επιβεβαιώνεται ότι το μεγαλύτερο ποσοστό (46%) των ερωτηθέντων συγκεντρώνεται στην ηλικιακή ομάδα 16 έως 30 χρονών. Ακολουθούν οι ηλικίες μεταξύ 30 και 40 ετών με ποσοστό 21,1%. Σε όμοια συγκέντρωση μεταξύ τους βρίσκονται οι ηλικιακές ομάδες 40 έως 50 ετών και 50 έως 65 ετών με ποσοστά 15,6% και οι δύο. Τέλος, το μικρότερο ποσοστό των συμμετεχόντων παρατηρείται στην ηλικιακή ομάδα άνω των 65 ετών με ποσοστό 1,6%.



Διάγραμμα 4.3 Ποσοστιαία κατανομή ηλικιακών ομάδων

Ακολουθεί διάγραμμα με την ποσοστιαία κατανομή του μορφωτικού επιπέδου. Οι τέσσερις κατηγορίες έτσι όπως διαμορφώθηκαν είναι απόφοιτος Δημοτικού, απόφοιτος Λυκείου, Πτυχιούχος και τέλος κάτοχος Μεταπτυχιακού-Διδακτορικού. Φαίνεται ότι, υπήρξε μόνο ένας απόφοιτος Δημοτικού σχολείου ανάμεσα στους ερωτηθέντες με ποσοστό 0,2%. Επειδή η παρουσία μόνο ενός ατόμου στην κατηγορία απόφοιτος Δημοτικού Σχολείου επηρεάζει οπτικά το αποτέλεσμα στην πίτα που ακολουθεί αποφασίστηκε να εξαιρεθεί από το διάγραμμα. Στη συνέχεια, σημειώνεται ότι, οι απόφοιτοι Λυκείου αποτελούν το 19,6% επί του συνόλου. Παρατηρείται ότι, το μεγαλύτερο μέρος των συμμετεχόντων είναι πτυχιούχοι τριτοβάθμιας εκπαίδευσης συγκεκριμένα σε ποσοστό 41,2%. Τέλος, το 39,2% του δείγματος είναι κάτοχοι Μεταπτυχιακού ή/και Διδακτορικού τίτλου σπουδών.

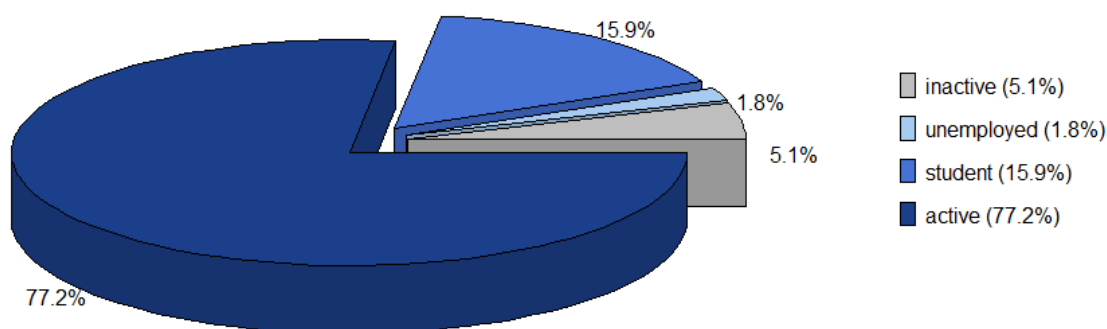
Κατανομή Εκπαιδευτικών Βαθμίδων



Διάγραμμα 4.4 Ποσοστιαία κατανομή εκπαιδευτικών βαθμίδων

Το επόμενο διάγραμμα αφορά τη μεταβλητή της εργασιακής κατάστασης. Η εργασιακή κατάσταση χωρίζεται σε τέσσερις κατηγορίες: Άεργος, Άνεργος, Μαθητής και Εργαζόμενος. Όπως φαίνεται και στο διάγραμμα, οι εργαζόμενοι αντιπροσωπεύουν την πλειοψηφία του δείγματος με ποσοστό 77,2%. Στη συνέχεια, ακολουθούν, οι μαθητές με ποσοστό 15,9%, αξιόλογο μέγεθος και το δεύτερο μεγαλύτερο του δείγματος. Αυτές οι δύο κατηγορίες αποτελούν το μεγαλύτερο ποσοστό επί του συνόλου των ερωτηθέντων, συγκεκριμένα, αποτελούν το 93,1% του δείγματος. Αυτό το 93,1% ουσιαστικά είναι ο ενεργός πληθυσμός. Ύστερα, παρατηρείται ότι, οι άεργοι εμφανίζονται με ποσοστό 5,1% και οι άνεργοι σε ποσοστό 1,8%, το μικρότερο σε σχέση με όλες υπόλοιπες κατηγορίες εργασιακής κατάστασης. Μαζί, οι δύο τελευταίες κατηγορίες αποτελούν το 6,9% του συνόλου των ερωτηθέντων του δείγματος.

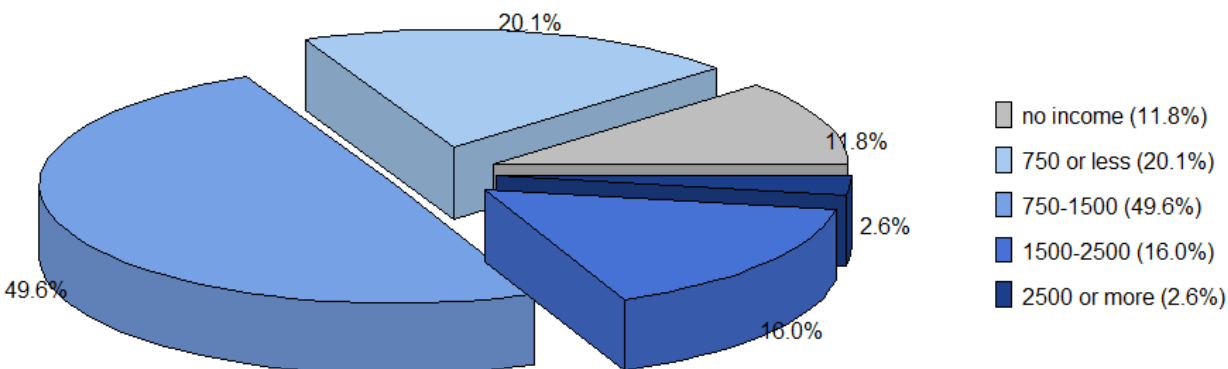
Κατανομή Εργασιακής Κατάστασης



Διάγραμμα 4.5 Ποσοστιαία κατανομή εργασιακής κατάστασης

Εν συνεχεία, παρουσιάζεται το διάγραμμα ποσοστιαίας κατανομής των εισοδηματικών κατηγοριών. Οι κατηγορίες διαμορφώθηκαν ως εξής: Κανένα Εισόδημα, έως 750€, 750€ έως 1500€, 1500€ έως 2500€, περισσότερα από 2500€. Από το διάγραμμα-πίτα προκύπτει ότι, η μεγαλύτερη συγκέντρωση του δείγματος δήλωσε ότι βρίσκεται στην εισοδηματική κατηγορία των 750€-1500€ σε ποσοστό 49,6%. Αμέσως μετά έρχεται με ποσοστό 20,1% η κατηγορία του εισοδήματος έως 750€. Το 16% των ερωτηθέντων δήλωσαν ότι βρίσκονται στην εισοδηματική κατηγορία των 1500€ έως 2500€. Προς το τέλος, έρχεται η κατηγορία των ατόμων με μηδενικό εισόδημα με ποσοστό 11,8%. Τέλος, η μειοψηφία του δείγματος με ποσοστό 2,6% συγκεντρώνεται στο εισόδημα μεγαλύτερο από 2500€. Παρατηρείται ότι, το μεγαλύτερο ποσοστό συμμετεχόντων της έρευνας λαμβάνει εισόδημα μεγαλύτερο από 750€ σε ποσοστό 68,2%. Από αυτό μόνο το 18,6% λαμβάνει εισόδημα μεγαλύτερο από τα 1500€.

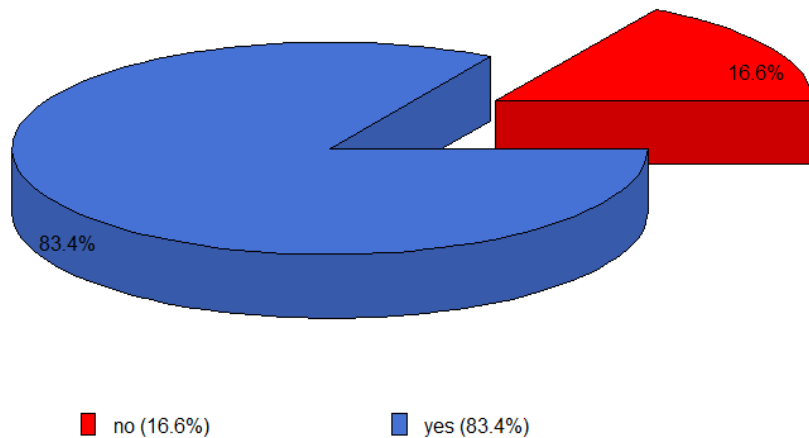
Κατανομή Εισοδηματικών Κατηγοριών



Διάγραμμα 4.6 Ποσοστιαία κατανομή εισοδηματικών κατηγοριών

Στο διάγραμμα που ακολουθεί, παρουσιάζονται τα ποσοστά των ατόμων που είτε έχουν είτε όχι στην κατοχή τους ιδιωτικής χρήσης αυτοκίνητο. Με κόκκινο χρώμα σημειώνεται το πλήθος των ατόμων που δεν έχουν στην κατοχή τους ιδιωτικό αυτοκίνητο και με μπλε χρώμα το πλήθος όσων έχουν. Όπως φαίνεται, το 16,6% των ατόμων που συμμετείχαν στην έρευνα δεν έχουν στην κατοχή τους αυτοκίνητο. Η πλειοψηφία των συμμετεχόντων με ποσοστό 83,4% έχει στην κατοχή του ιδιωτικό αυτοκίνητο.

Κατανομή Κατοχής Αυτοκινήτου



Διάγραμμα 4.7 Ποσοστιαία κατανομή κατοχής ιδιωτικού αυτοκινήτου

Σε αυτό το σημείο, αναλύεται η μεταβλητή «home» που εκφράζει τη ζώνη κατοικίας. Για τις ανάγκες της έρευνας η Μητροπολιτική Περιοχή της Αθήνας χωρίστηκε σε ζώνες. Στον πίνακα που ακολουθεί, φαίνεται το πλήθος των συμμετεχόντων της έρευνας ανά ζώνη.

Πίνακας 4.2 Πλήθος συμμετεχόντων ανά ζώνη κατοικίας

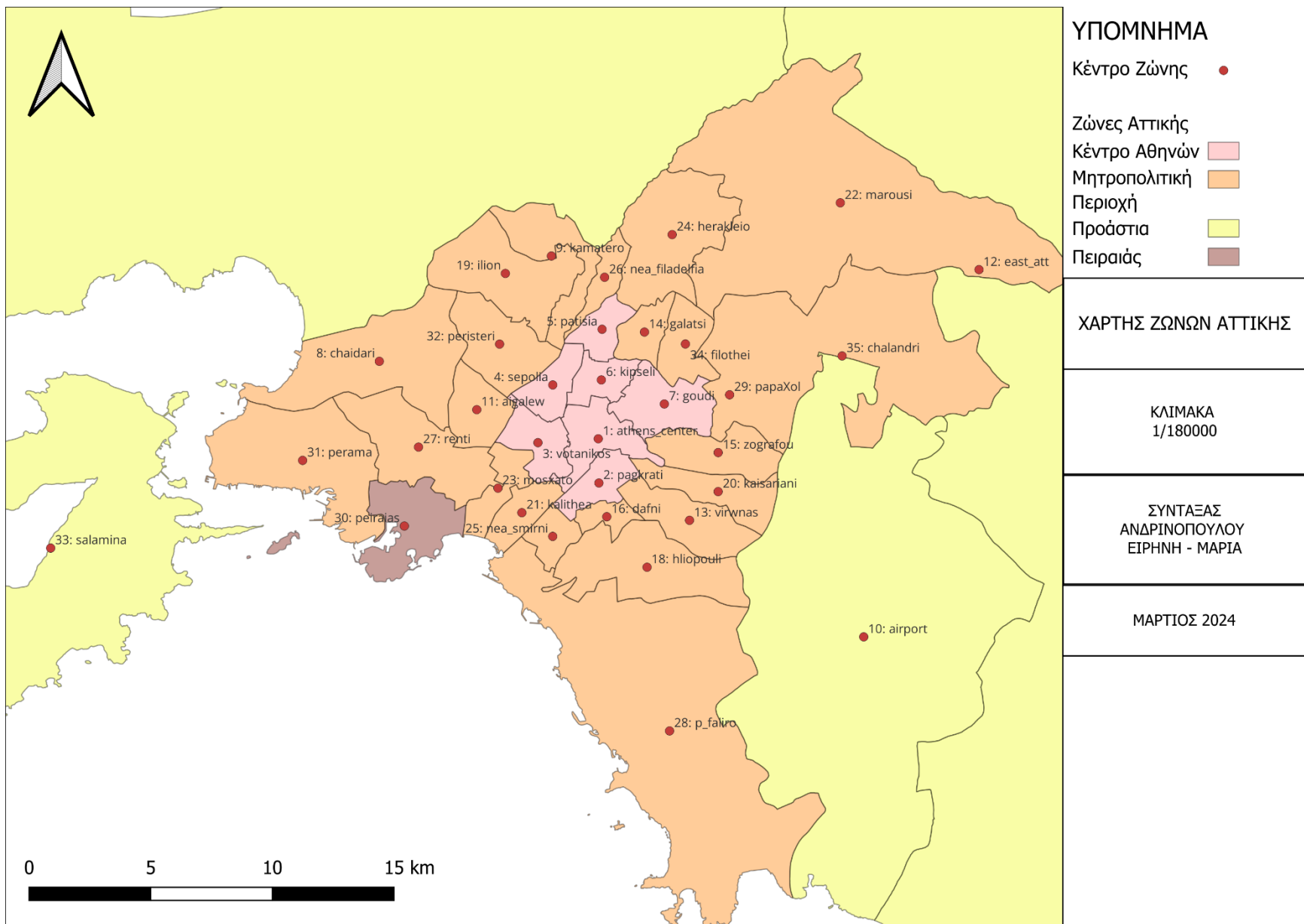
ID	Περιοχή	Πλήθος	ID	Περιοχή	Πλήθος
1	Κέντρο Αθηνών	41	19	Ίλιον	22
2	Παγκράτι	26	20	Καισαριανή	10
3	Βοτανικός	6	21	Καλλιθέα	27
4	Σεπόλια	10	22	Μαρούσι	26
5	Πατήσια	12	23	Μοσχάτο	3
6	Κυψέλη	8	24	Νέο Ηράκλειο	14
7	Γουδί	38	25	Νέα Σμύρνη	24
8	Χαϊδαρι	10	26	Νέα Φιλαδέλφεια	10
9	Καματερό	7	27	Ρέντη	1
10	Αεροδρόμιο	5	28	Παλαιό Φάληρο	12
11	Αιγάλεω	8	29	Παπάγου-Χολαργός	9
12	Ανατολική Αττική	12	30	Πειραιάς	21
13	Βύρωνα	4	31	Πέραμα	7
14	Γαλάτσι	6	32	Περιστερί	16
15	Ζωγράφου	36	33	Σαλαμίνα	2
16	Δάφνη	9	34	Φιλοθέη	8
17	Δυτική Αττική	7	35	Χαλάνδρι	28
18	Ηλιούπολη	17	36	Εκτός Αττικής	11

Στο σύνολό τους τα ερωτηματολόγια ήταν 513. Μέσω του πίνακα με τα αθροίσματα εμφάνισης των ζωνών κατοικίας στο δείγμα, προκύπτει ότι, η μεγαλύτερη συγκέντρωση του δείγματος σημειώνεται στο κέντρο της Αθήνας με πλήθος 41 άτομα, στη συνέχεια ακολουθούν το Γουδί με 38 από τους 513 ερωτηθέντες να κατοικούν σε αυτό και η περιοχή του Ζωγράφου με πλήθος 36 άτομα. Αξιοσημείωτη είναι η συγκέντρωση στο Χαλάνδρι με πλήθος 28 κατοίκων από το δείγμα και η Καλλιθέα με πλήθος 27.

Μέτριες συγκεντρώσεις εμφανίζονται στις περιοχές: Παγκράτι και Μαρούσι με 26 ερωτηθέντες-κατοίκους η κάθε μία, Νέα Σμύρνη με 24, Ίλιον με 22, Πειραιάς με 21, Ηλιούπολη με 17, Περιστέρι με 16, Νέο Ηράκλειο με 14, Ανατολική Αττική και Παλαιό Φάληρο με 12 άτομα η κάθε περιοχή. Εκτός Αττικής οι συμμετέχοντες ήταν 11 και στις περιοχές Σεπόλια, Καισαριανή και Νέα Φιλαδέλφεια ήταν 10 στην εκάστοτε περιοχή.

Η μειοψηφούσα ζώνη είναι ο Ρέντης με μόνο έναν από τους συμμετέχοντες να διαμένουν σε αυτόν. Ύστερα, ακολουθούν η Σαλαμίνα με 2, το Μοσχάτο με 3, ο Βύρωνας με 4, η περιοχή του αεροδρομίου με 5, ο Βοτανικός και το Γαλάτσι με 6 άτομα το κάθε ένα. Το Καματερό, η Δυτική Αττική και το Πέραμα συγκεντρώνουν 7 συμμετέχοντες η εκάστοτε περιοχή. Ακολουθεί, η Κυψέλη, το Αιγάλεω και η Φιλοθέη με 8 άτομα η κάθε μία και τέλος η Δάφνη και ο Χολαργός με 9 ερωτηθέντες η κάθε περιοχή.

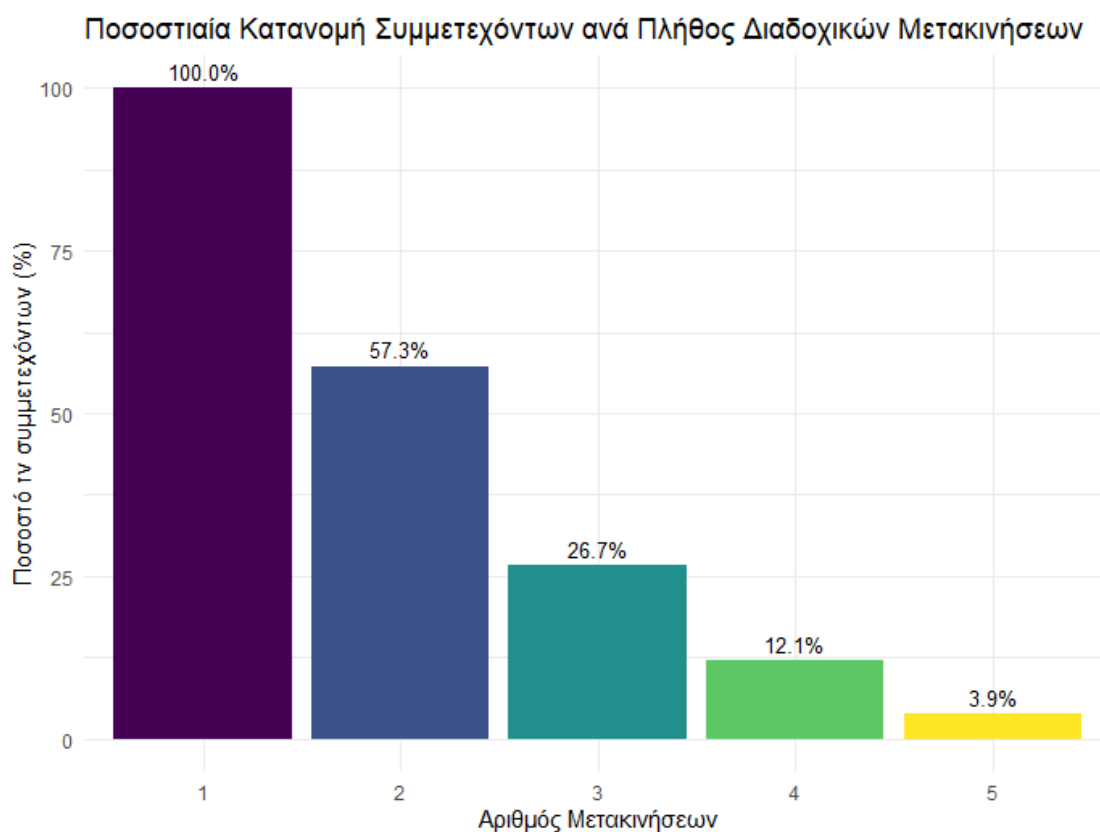
Παρακάτω, παρουσιάζεται ο χάρτης με τα όρια των ζωνών που ορίστηκαν στην Αττική καθώς και το κέντρο της κάθε ζώνης σε κλίμακα 1/180000. Με διαφορετικά χρώματα σημειώνονται οι ζώνες που απαρτίζουν το κέντρο της Αθήνας, την ευρύτερη Μητροπολιτική Περιοχή των Αθηνών, τα προάστια, και τον Πειραιά. Τα κέντρα των ζωνών χωροθετούνται ως κόκκινα σημεία πάνω στο χάρτη. Βάσει αυτών των κέντρων υπολογίστηκαν οι αποστάσεις μεταξύ των ζωνών.



Χάρτης 4.1 Χάρτης ζωνών Αττικής με κεντροειδή

Ύστερα, για τη χωροθέτηση όλων των μετακινήσεων των συμμετεχόντων χρησιμοποιήθηκε η ίδια λογική με αυτή της μεταβλητής “home”. Οι μεταβλητές “dist1”, “dist2”, “dist3”, “dist4” και “dist5” αφορούν τους προορισμούς των πέντε διαδοχικών μετακινήσεων που μπορούσαν να συμπληρώσουν οι ερωτηθέντες της έρευνας. Οι τιμές που μπορούν λάβουν αυτές οι μεταβλητές των διαδοχικών προορισμών είναι οι ίδιες ζώνες που χρησιμοποιήθηκαν για τη χωροθέτηση της κατοικίας. Στους συμμετέχοντες της έρευνας δόθηκε η δυνατότητα να δηλώσουν έως και πέντε μετακινήσεις στις οποίες προβαίνουν μία τυπική μέρα της καθημερινότητας τους. Με αυτόν τον τρόπο προέκυψαν και ερωτηματολόγια που περιείχαν λιγότερες από πέντε μετακινήσεις.

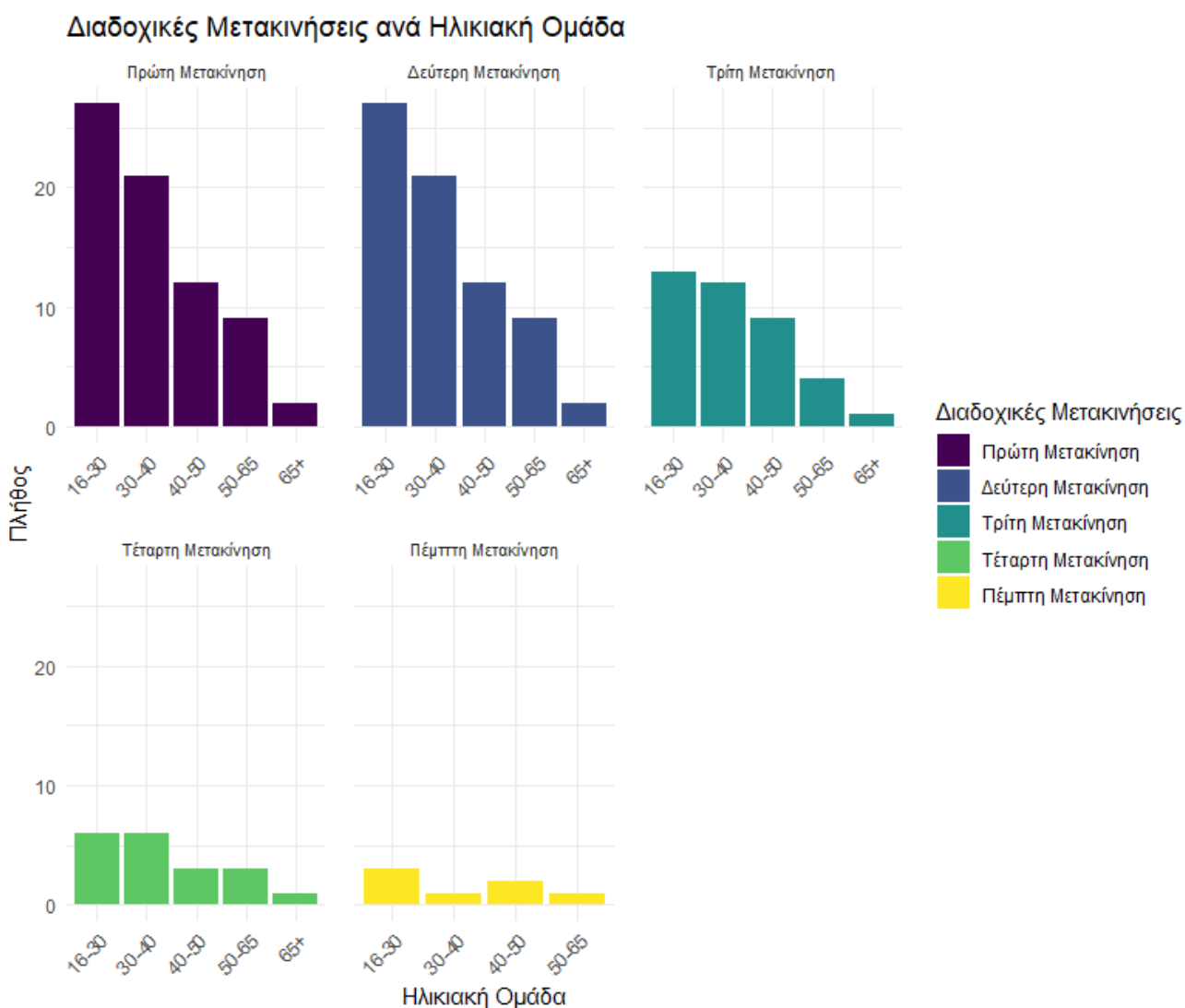
Στην συνέχεια, παρουσιάζεται το διάγραμμα κατανομής των συμμετεχόντων και των διαδοχικών τους μετακινήσεων. Παρατηρείται ότι, και οι 513 συμμετέχοντες δήλωσαν μία μετακίνηση, όμως 294 από αυτούς, δηλαδή το 57% του δείγματος συμπλήρωσε τη δεύτερη διαδοχική μετακίνηση χωρίς να μεταβεί σε τρίτη. Οι 137 από τους 513 συμμετέχοντες, σε ποσοστό το 26,7% του δείγματος, συμπλήρωσαν τρεις διαδοχικές μετακινήσεις. Έπειτα το 12,1% του συνόλου των συμμετεχόντων (62 άτομα), δήλωσαν συνολικά τέσσερις διαδοχικές μετακινήσεις και τέλος, μόνο το 3,9% του δείγματος, συγκεκριμένα 20 συμμετέχοντες, δήλωσαν πέντε διαδοχικές μετακινήσεις.



Διάγραμμα 4.8 Ποσοστιαία κατανομή πλήθους διαδοχικών μετακινήσεων

Στο διάγραμμα που ακολουθεί, εμφανίζονται πέντε ιστογράμματα, όσα δηλαδή και ο μέγιστος αριθμός των μετακινήσεων ενός ατόμου. Σε αυτά τα ιστογράμματα εμφανίζεται το πλήθος των συμμετεχόντων ανά διαδοχική μετακίνηση και ανά ηλικιακή ομάδα. Μέσω του διαγράμματος προκύπτουν συμπεράσματα που αφορούν το ποσοστό του δείγματος που προβαίνει σε παραπάνω από μία μετακινήσεις. Όπως φαίνεται, οι μικρότερες ηλικιακές ομάδες προβαίνουν σε περισσότερες διαδοχικές μετακινήσεις.

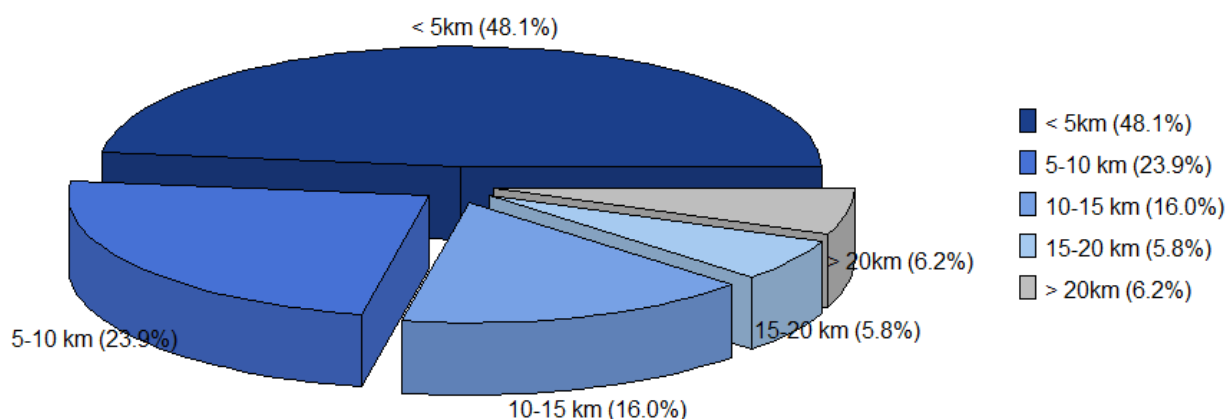
Συγκεκριμένα, η ηλικιακή ομάδα των 16 έως 30 ετών εμφανίζει το μεγαλύτερο πλήθος σε όλες τις διαδοχικές μετακινήσεις. Ύστερα, ακολουθεί, η ομάδα των 30 έως 40 ετών, με μεγάλο αριθμό μετακινήσεων στην τρίτη και τέταρτη μετακίνηση. Στην πέμπτη διαδοχική μετακίνηση δεύτερο μεγαλύτερο πλήθος σημειώνει η ηλικιακή ομάδα των 40 έως 50 ετών. Η ομάδα 50 έως 65 ετών εμφανίζει σταθερά φθίνουσα πορεία όσο εναλλάσσονται οι διαδοχικές μετακινήσεις. Τέλος, η ομάδα των ατόμων άνω των 65 ετών δεν εμφανίζει μηδενικό πλήθος στην πέμπτη διαδοχική μετακίνηση.



Διάγραμμα 4.9 Κατανομή πλήθους διαδοχικών μετακινήσεων ανά ηλικιακή ομάδα

Εν συνεχεία, παρουσιάζεται το διάγραμμα με την ποσοστιαία κατανομή των αποστάσεων των μετακινήσεων. Οι μετακινήσεις κατηγοριοποιήθηκαν σε αποστάσεις μικρότερες από 5 χιλιόμετρα, από 5 έως 10 χιλιόμετρα, από 10 έως 15 χιλιόμετρα, από 15 έως 20 χιλιόμετρα, μεγαλύτερες από 20 χιλιόμετρα.

Ποσοστιαία Κατανομή Αποστάσεων των Μετακινήσεων



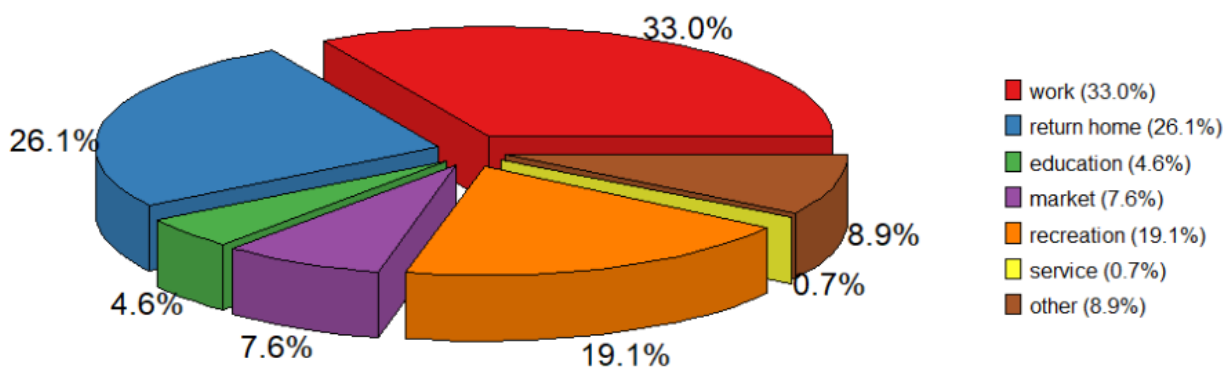
Διάγραμμα 4.10 Ποσοστιαία κατανομή αποστάσεων των μετακινήσεων

Παρατηρείται ότι, οι περισσότερες καταγεγραμμένες μετακινήσεις αφορούν αποστάσεις μικρότερες των 5 χιλιομέτρων, με ποσοστό 48%, δηλαδή σχεδόν οι μισές από τις συνολικές μετακινήσεις. Ακολουθούν οι μετακινήσεις με απόσταση μεταξύ 5-10 χιλιομέτρων (24%) και στην συνέχεια οι μετακινήσεις με απόσταση μεταξύ 10-15 χιλιομέτρων (16%). Τέλος, οι μετακινήσεις με απόσταση μεταξύ των 15-20 χιλιομέτρων αφορούν το 5,8% του δείγματος και οι μετακινήσεις με απόσταση μεγαλύτερη των 20 χιλιομέτρων αφορούν το 6,2% του δείγματος.

Ακολουθεί το διάγραμμα με την ποσοστιαία κατανομή των σκοπών μετακίνησης. Οι σκοποί μετακίνησης χωρίστηκαν σε εργασία, επιστροφή στην οικία, εκπαίδευση, αγορές, αναψυχή, υπηρεσίες και άλλο.

Από το διάγραμμα προκύπτει ότι, η πλειοψηφία των μετακινήσεων διεξάγεται για μετάβαση στην εργασία (33%). Η επιστροφή στο σπίτι συγκεντρώνει ποσοστό 26,1% των μετακινήσεων. Φυσικά η επιστροφή στο σπίτι αποτελεί απαραίτητη μετακίνηση κατά την διάρκεια της ημέρας. Ακολουθεί, η μετάβαση για εκπαιδευτικούς λόγους (4,6%) και για αγορές (7,6%). Επίσης η αναψυχή αφορά το 19,1% των μετακινήσεων και οι υπηρεσίες το 0,7%. Τέλος, οι μετακινήσεις για άλλους λόγους έχουν ποσοστό 8,9%

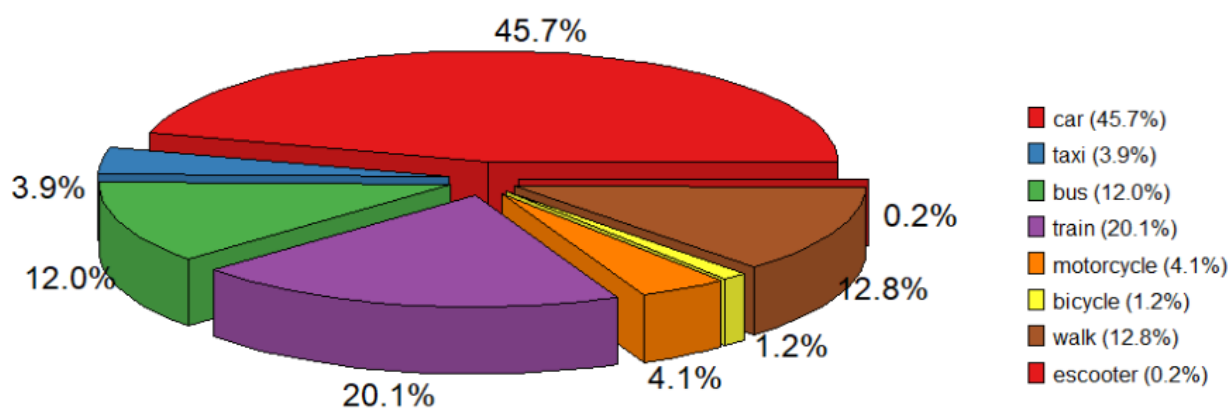
Ποσοστιαία Κατανομή Σκοπών των Μετακινήσεων



Διάγραμμα 4.11 Ποσοστιαία κατανομή σκοπών μετακινήσεων

Παρακάτω παρουσιάζεται το διάγραμμα κατανομής των μέσων μετακίνησης, τα οποία διακρίνονται σε: αυτοκίνητο, ταξί, λεωφορείο, μέσα σταθερής τροχιάς, μοτοσυκλέτα, ποδήλατο, περπάτημα, ηλεκτρικό πατίνι.

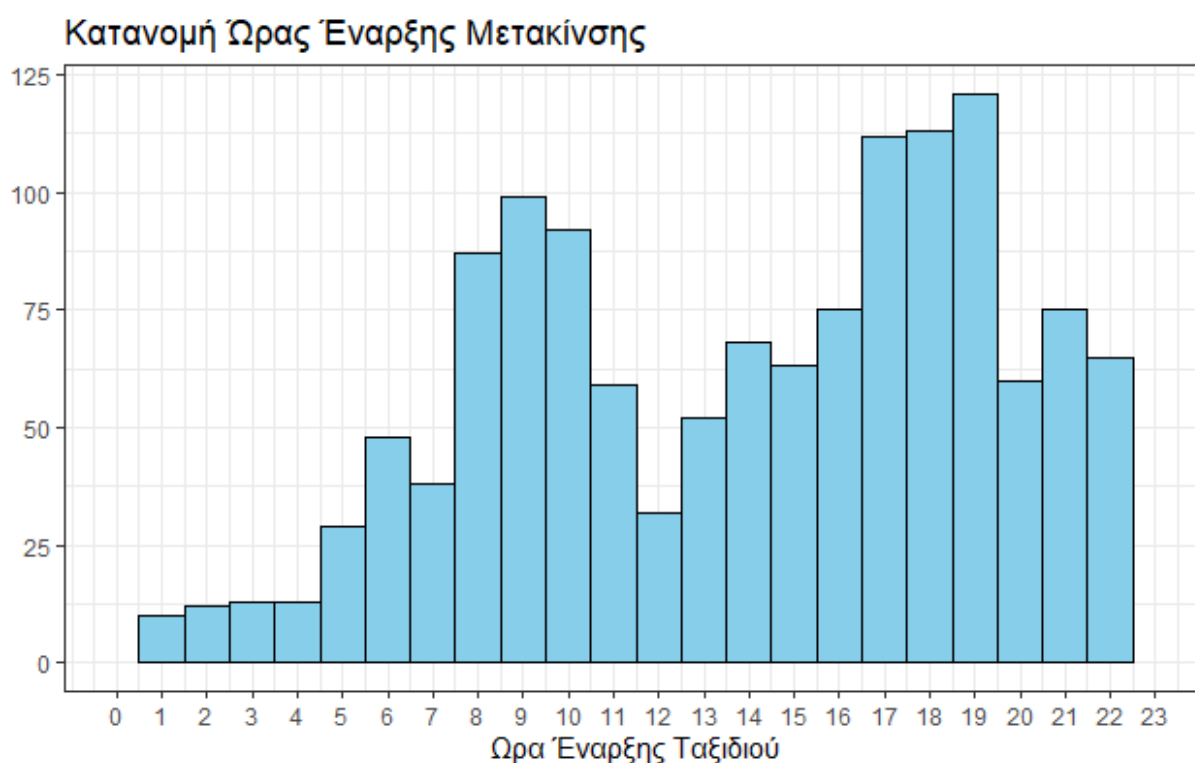
Ποσοστιαία Κατανομή Μέσων των Μετακινήσεων



Διάγραμμα 4.12 Ποσοστιαία κατανομή μέσων μετακίνησης

Όπως φαίνεται στο διάγραμμα το ΙΧ αυτοκίνητο χρησιμοποιείται για το 45,7% των μετακινήσεων. Ταυτόχρονα το ταξί χρησιμοποιείται για το 3,9% των μετακινήσεων, ενώ αθροιστικά οι μετακινήσεις που διεξάγονται με αυτοκίνητο αγγίζουν το 50%. Με Δημόσια Συγκοινωνία διεξάγεται το 32% των μετακινήσεων (λεωφορείο 12% και Μέσα Σταθερής Τροχιάς 20%). Οι μετακινήσεις με μοτοσυκλέτα συγκεντρώνουν περίπου το 4% των μετακινήσεων, ενώ με ποδήλατο το 1,2%. Έπειτα, το 12,8% των μετακινήσεων διεξάγεται με περπάτημα και με ποσοστό μικρότερο του 1% με ηλεκτρικό πατίνι (0,2%).

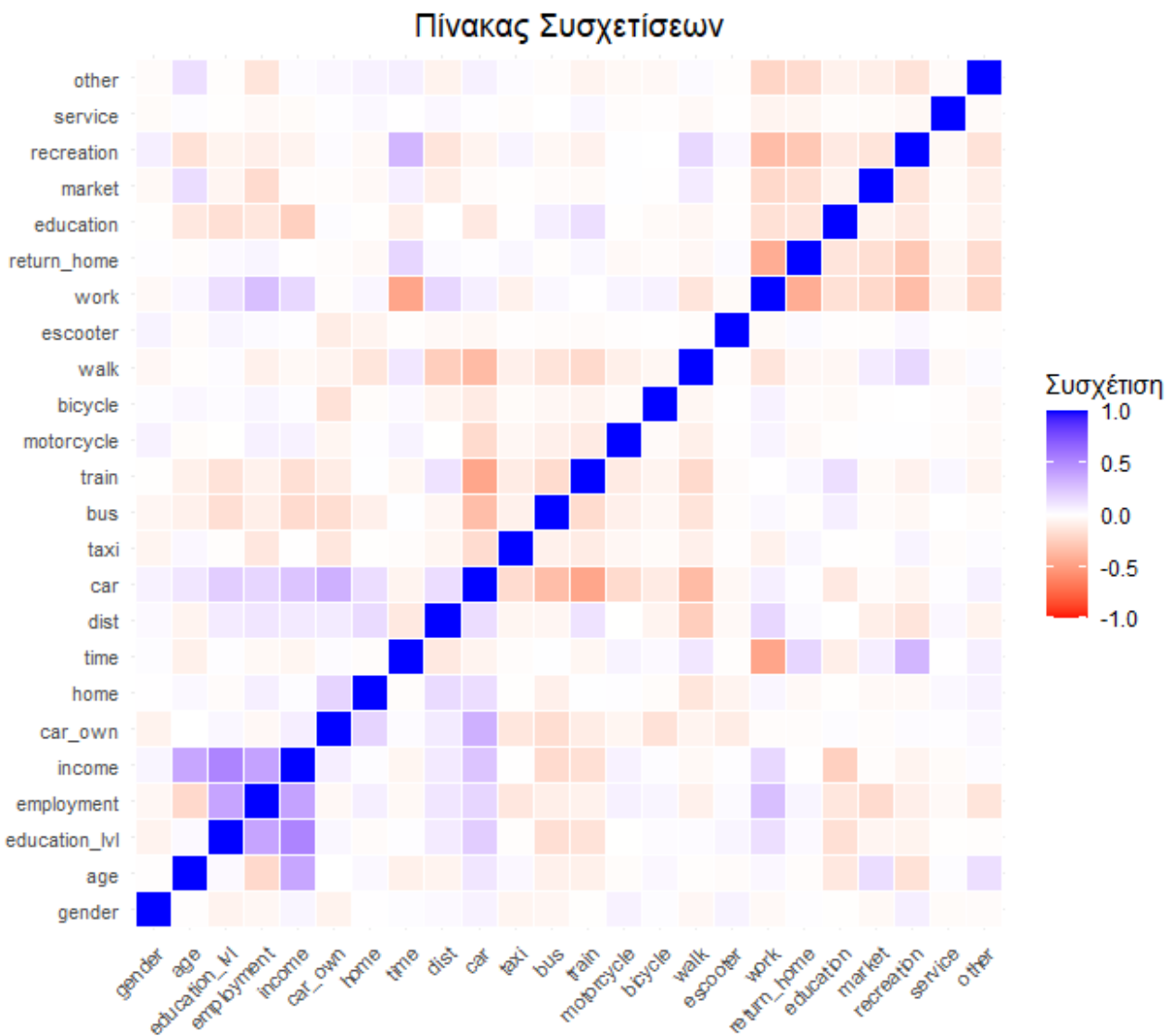
Ακολουθεί, το ιστόγραμμα κατανομής των ωρών έναρξης των μετακινήσεων. Οι ώρες έναρξης των ταξιδιών αφορούν ένα ολόκληρο εικοσιτετράωρο, δηλαδή κυμαίνονται από τις 00:00 έως τις 23:00. Παρατηρείται ότι, μεγάλο πλήθος μετακινήσεων συγκεντρώνεται στις ώρες μεταξύ 8:00 και 11:00, δηλαδή τις πρωινές ώρες καθώς και στις ώρες μεταξύ 16:00 και 19:00, δηλαδή τις απογευματινές. Γενικά οι μεγαλύτερες συγκεντρώσεις εμφανίζονται τις πρωινές και τις απογευματινές ώρες, ενώ τις βραδινές ώρες μεταξύ 20:00-00:00 καθώς και τις πρώτες πρωινές ώρες από 00:00 έως 7:00 παρατηρείται πτώση στο πλήθος των μετακινήσεων. Επιπλέον, μείωση στις μετακινήσεις εμφανίζεται τις μεσημεριανές ώρες από 11:00 έως 15:00.



Διάγραμμα 4.13 Κατανομή ωρών έναρξης μετακινήσεων

Στον πίνακα που ακολουθεί, παρουσιάζονται οι τιμές των συσχετίσεων μεταξύ των μεταβλητών του δείγματος. Οι μεταβλητές που συμμετέχουν είναι: το φύλο, η ηλικία, το επίπεδο εκπαίδευσης, η εργασιακή κατάσταση, το εισόδημα, η κτήση αυτοκινήτου, η ζώνη κατοικίας, η ώρα μετακίνησης, η απόσταση, το μέσο διεξαγωγής της μετακίνησης (αυτοκίνητο, ταξί, λεωφορείο, μέσα σταθερής τροχιάς, μοτοσυκλέτα, ποδήλατο, βάδισμα, πατίνι) και τον σκοπό της μετακίνησης (εργασία, επιστροφή στην κατοικία, εκπαίδευση, αγορές, αναψυχή, υπηρεσίες και άλλο).

Πίνακας 4.3 Συσχετίσεις Μεταβλητών

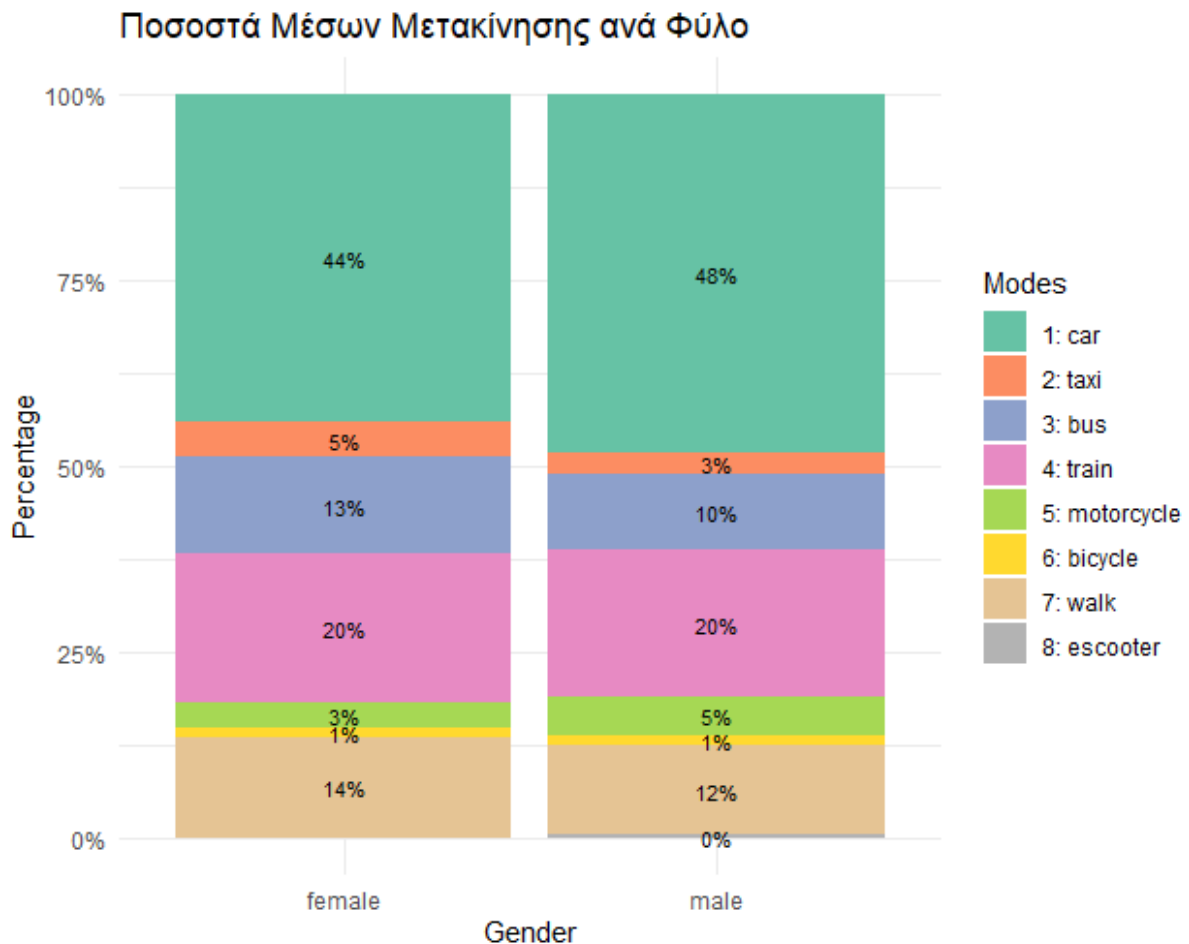


Παρατηρείται ότι, με διαβαθμίσεις του μπλε σημειώνονται οι θετικές συσχετίσεις και με διαβαθμίσεις του κόκκινου οι αρνητικές. Οι θετικές συσχετίσεις εμφανίζονται σε ανάλογες μεταβλητές, ενώ οι αρνητικές συσχετίσεις αφορούν αντιστρόφως ανάλογες μεταβλητές.

Ενδεικτικά, φαίνεται ότι, τα Μέσα Σταθερής Τροχιάς και το περπάτημα έχουν αρνητική συσχέτιση με το αυτοκίνητο. Επίσης, αρνητική συσχέτιση έχουν το περπάτημα με την απόσταση. Θετική συσχέτιση εμφανίζουν τα Μέσα Σταθερής Τροχιάς με τις μετακινήσεις για εκπαιδευτικούς λόγους. Επιπλέον, θετική συσχέτιση εμφανίζουν οι μετακινήσεις για αναψυχή και το περπάτημα. Τέλος, οι μετακινήσεις για αναψυχή και η μεταβλητή του χρόνου έναρξης παρουσιάζουν θετική συσχέτιση. Μέσω του πίνακα μπορούν να αξιολογηθούν όλες οι συσχετίσεις των μεταβλητών.

4.5 Επεξεργασία Δεδομένων

Σε αυτό το στάδιο, έγινε εστιασμένη ανάλυση των δεδομένων με επίκεντρο το μέσο μετακίνησης των ταξιδιών. Παρακάτω, παρουσιάζονται διάφορα διαγράμματα που προκύπτουν από συνδυασμό της μεταβλητής του μέσου μετακίνησης με της μεταβλητές της το φύλο, η ηλικία, το εισόδημα, η εκπαίδευση και άλλα κοινωνικοοικονομικά χαρακτηριστικά.



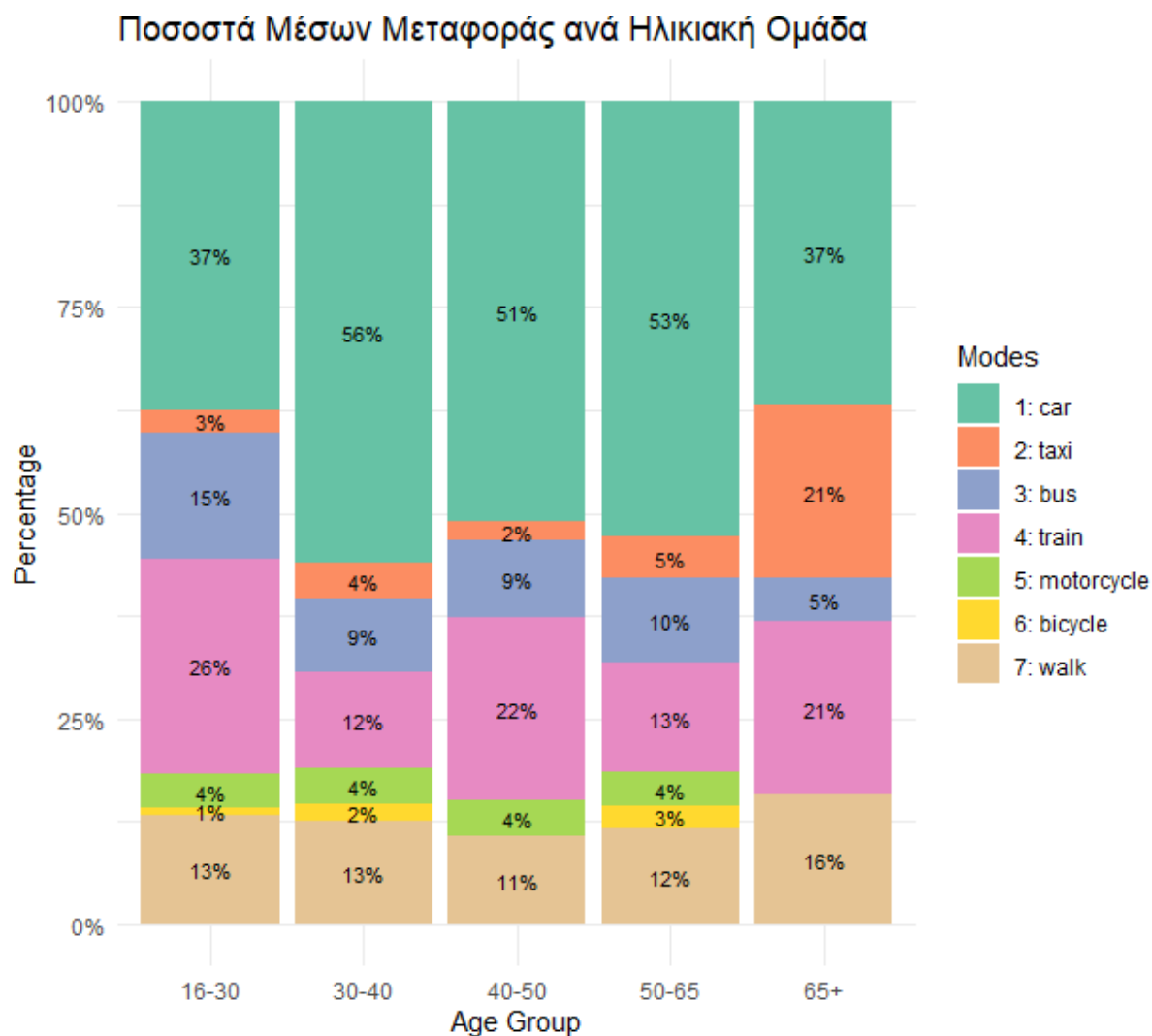
Διάγραμμα 4.14 Ποσοστιαία κατανομή μέσων μετακίνησης ανά φύλο

Στο παραπάνω διάγραμμα, φαίνεται η κατανομή των μέσων μεταφοράς σε ποσοστά ανά φύλο. Παρατηρείται ότι, όσον αφορά της γυναίκες, η πλειοψηφία της χρησιμοποιεί το αυτοκίνητο ως μέσο μεταφοράς σε ποσοστό 44%. Ακολουθούν, τα μέσα σταθερής τροχιάς με ποσοστό 20% και το περπάτημα σε ποσοστό 14%. Οι γυναίκες, χρησιμοποιούν για το 13% των μετακινήσεων της, το λεωφορείο και κατά 5% ταξί. Τέλος, με 3% εμφανίζεται η μοτοσυκλέτα και μόνο με 1% το ποδήλατο.

Παρομοίως, ο ανδρικός πληθυσμός του δείγματος, χρησιμοποιεί ως επί το πλείστον το αυτοκίνητο για της καθημερινές του μετακινήσεις με ποσοστό 48%, κατά 4% μεγαλύτερο από αυτό των γυναικών. Τα μέσα σταθερής τροχιάς, χρησιμοποιούνται από της άντρες, σε ποσοστό 20%, ακριβώς ίσο με αυτό των γυναικών. Ακολουθεί η μετακίνηση με περπάτημα, με ποσοστό 12%, κατά 2% λιγότερο από αυτό

του γυναικείου πληθυσμού. Της, οι άντρες φαίνεται να χρησιμοποιούν κατά 2% περισσότερο την μοτοσυκλέτα, συνολικά σε ποσοστό 5%. Τέλος, το 1% χρησιμοποιεί το ποδήλατο, όσο δηλαδή και οι γυναίκες. Στο διάγραμμα φαίνεται ότι το scooter σημειώνει ποσοστό μικρότερο του 1%.

Εν συνεχεία, παρουσιάζεται το διάγραμμα κατανομής των μέσων μεταφοράς ανά ηλικιακή ομάδα.



Διάγραμμα 4.15 Ποσοστιαία κατανομή μέσων μετακίνησης ανά ηλικιακή ομάδα

Παρατηρείται ότι, στην ηλικιακή ομάδα 16 έως 30 ετών επικρατεί η χρήση του αυτοκινήτου, με ποσοστό 37%, ενώ παράλληλα παρατηρείται μέγιστο της χρήσης μέσων σταθερής τροχιάς, σε σχέση με της της υπόλοιπες ηλικιακές ομάδες, με ποσοστό 26%. Της, μέγιστο σημειώνει και το λεωφορείο, με ποσοστό 15%. Φαίνεται ότι, οι νέοι 16 έως 30 ετών χρησιμοποιούν το περπάτημα σε ποσοστό 13% των μετακινήσεών της. Επιπλέον, η χρήση της μηχανής συγκεντρώνει ποσοστό 4%, η χρήση ταξί συγκεντρώνει 3%, ενώ μόνο το 1% των μετακινήσεων των ατόμων μεταξύ 16 και 30 ετών διεξάγεται με ποδήλατο.

Η δεύτερη ηλικιακή ομάδα, που αφορά της ηλικίες μεταξύ 30 και 40 ετών, παρουσιάζει την πιο αυξημένη χρήση ιδιωτικού αυτοκινήτου σε σχέση με της της υπόλοιπες ηλικιακές ομάδες, σε ποσοστό 56%. Ακολουθούν, τα μέσα σταθερής τροχιάς με ποσοστό 12%, το μικρότερο ποσοστό χρήσης μέσων σταθερής τροχιάς ανάμεσα σε της της ηλικιακές κατηγορίες. Της, το 13% των μετακινήσεων γίνεται με περπάτημα, ενώ το 9% με λεωφορείο. Το ταξί και η μοτοσυκλέτα συγκεντρώνουν ποσοστό 4% η κάθε μία. Τέλος, το ποδήλατο εξυπηρετεί το 2% των μετακινήσεων της ομάδας, ελαφρώς ανεβασμένο ποσοστό σε σχέση με της μικρότερες ηλικίες.

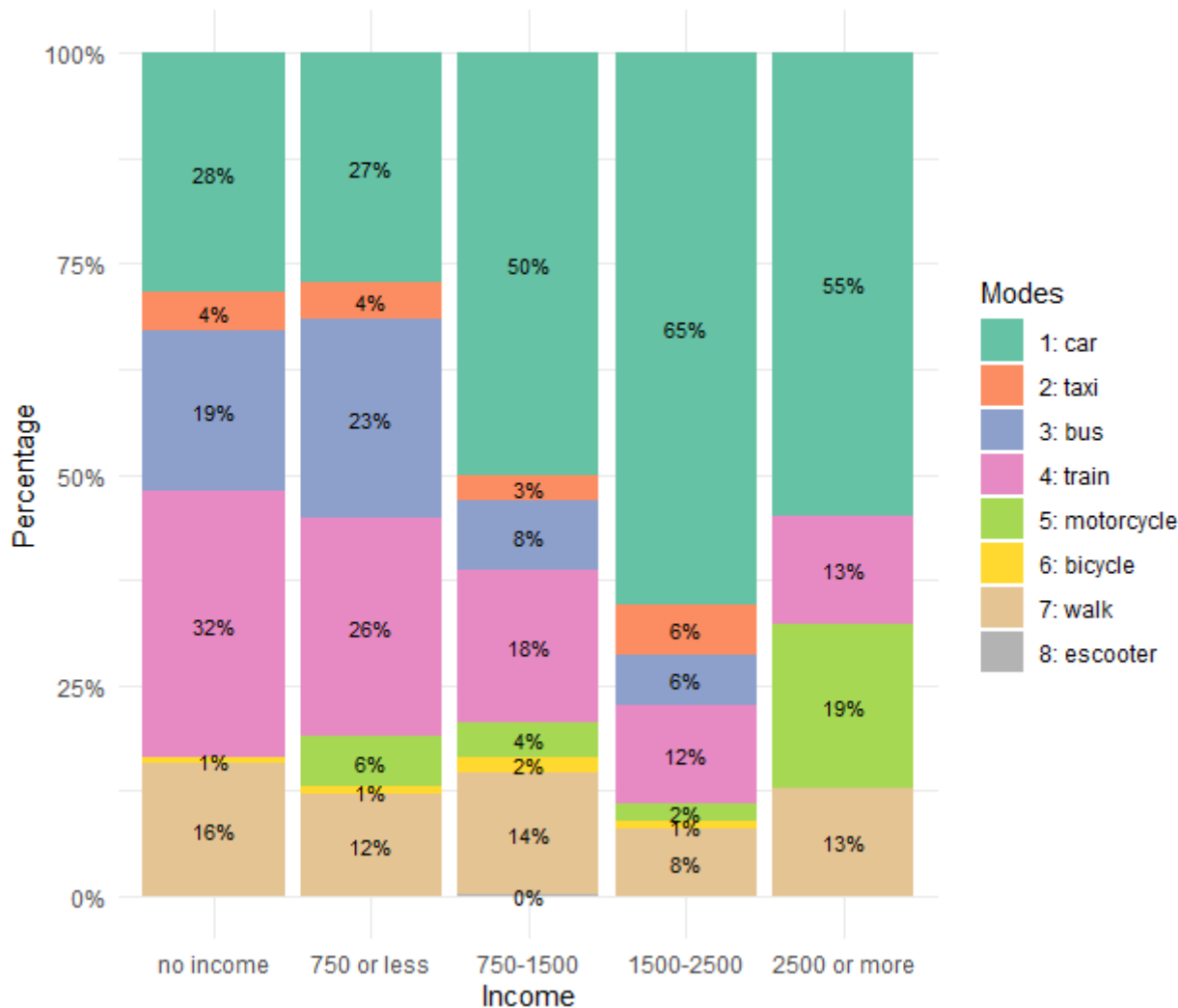
Στη συνέχεια, στην ηλικιακή ομάδα 40 έως 50 ετών, το αυτοκίνητο συγκεντρώνει το 51% και τα μέσα σταθερής τροχιάς το 22% των μετακινήσεων της ομάδας. Επιπλέον, το 11% των μετακινήσεων γίνονται με περπάτημα και το 9% γίνεται με το λεωφορείο. Σταθερό και σε αυτή την ηλικιακή ομάδα, παραμένει το ποσοστό χρήσης της μοτοσυκλέτας στο 4%, ενώ η χρήση ταξί εδώ σημειώνει ελάχιστο, με ποσοστό 2%. Αξιοσημείωτη, είναι η απουσία της χρήσης ποδηλάτου.

Στην επόμενη ηλικιακή ομάδα 50 έως 65 ετών, το αυτοκίνητο εξυπηρετεί το 53% των καθημερινών μετακινήσεων. Επιπλέον, οι μετακινήσεις με μέσα σταθερής τροχιάς έχουν ποσοστό 13% και το περπάτημα έχει 12%. Ακόμα, το λεωφορείο απαντά στο 10% των μετακινήσεων της της ομάδας. Ελαφρώς αυξημένη σε σχέση με της προηγούμενες ηλικιακές ομάδες, είναι η χρήση του ταξί (5%) και του ποδηλάτου με 3%, ποσοστό μεγαλύτερο σε σχέση με της της υπόλοιπες ομάδες. Τέλος, η μοτοσυκλέτα χρησιμοποιείται για το 4% των μετακινήσεων των ερωτηθέντων της ομάδας.

Η τελευταία ηλικιακή ομάδα, αφορά της συμμετέχοντες άνω των 65 ετών. Το αυτοκίνητο σημειώνει ποσοστό 37%, ενώ τα μέσα σταθερής τροχιάς ποσοστό 21%. Μεγάλη αύξηση, παρατηρείται στη χρήση ταξί, με της μετακινήσεις να αγγίζουν το ποσοστό 21% και στο περπάτημα με ποσοστό 16%. Ακόμα, μόνο το 5% των μετακινήσεων διεξάγεται με λεωφορείο, μικρότερο ποσοστό χρήσης λεωφορείου σε σχέση με της υπόλοιπες ομάδες. Αξιοσημείωτη, είναι η απουσία της μοτοσυκλέτας και του ποδηλάτου από την ομάδα των ατόμων που έχουν υπερβεί τα 65 έτη ηλικίας.

Παρακάτω, παρουσιάζεται το διάγραμμα κατανομής των μέσων μετακίνησης αναλόγως την εισοδηματική κατηγορία.

Ποσοστά Μέσων Μετακίνησης ανά Εισοδηματική Κατηγορία



Διάγραμμα 4.16 Ποσοστιαία κατανομή μέσων μετακίνησης ανά κατηγορία εισοδήματος

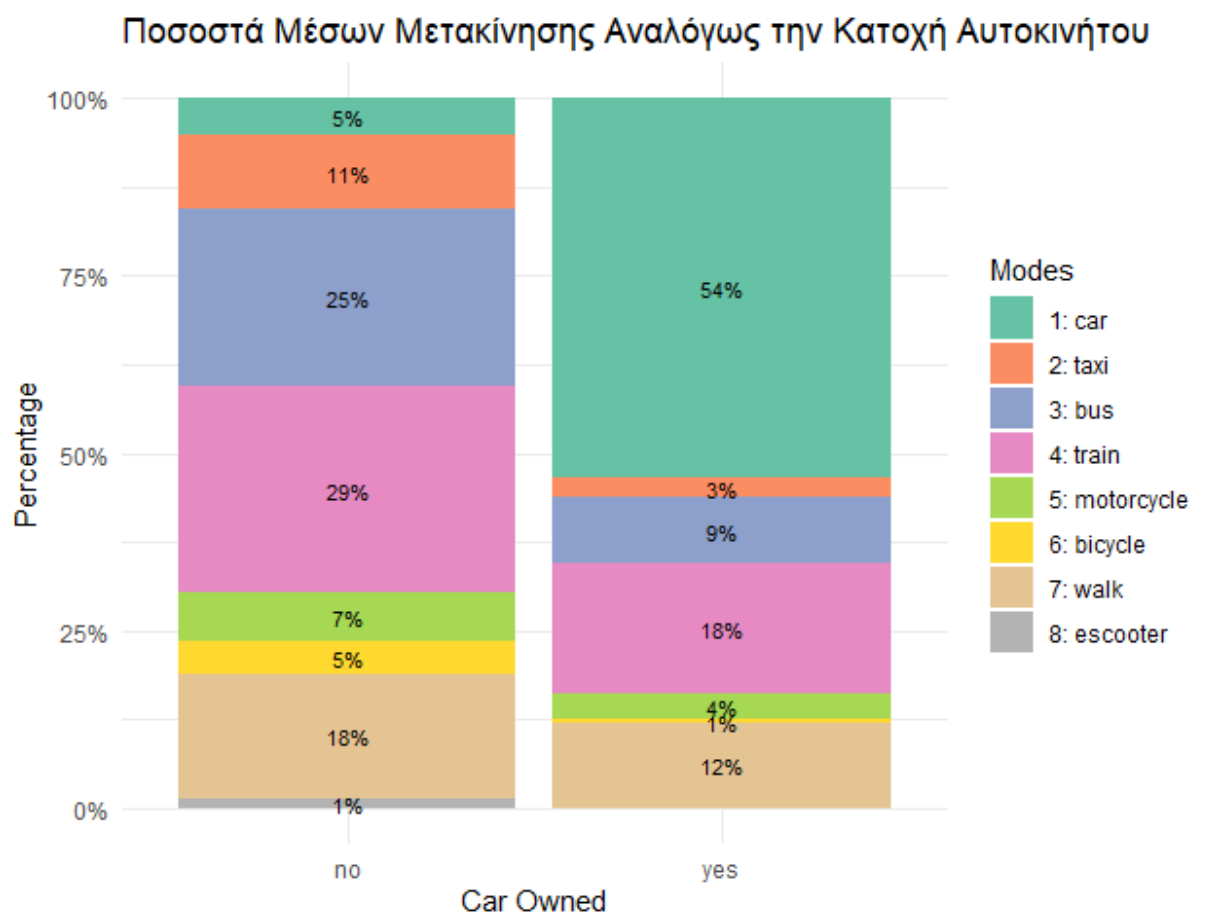
Αρχικά, το μεγαλύτερο ποσοστό μετακινήσεων με αυτοκίνητο, παρατηρείται στην εισοδηματική κατηγορία 1500€-2500€, με ποσοστό 65%. Ακολουθεί η εισοδηματική κατηγορία άνω των 2500€, με ποσοστό 55%. Της, αξιόλογο ποσοστό (50%) παρουσιάζει και η κατηγορία 750€ έως 1500€. Πολύ μικρότερα ποσοστά παρουσιάζονται της κατηγορίες λιγότερα από 750€ και μηδενικό εισόδημα, με ποσοστά 27% και 28%, αντίστοιχα.

Όσον αφορά, τα μέσα σταθερής τροχιάς, η κατηγορία του μηδενικού εισοδήματος κατέχει ποσοστό 32% των μετακινήσεων. Ακολουθεί, η κατηγορία κάτω από 750€, με ποσοστό 26%. Της, τα μέσα σταθερής τροχιάς σημείωσαν ποσοστό 18% στην κατηγορία των 750€ έως 1500€. Της εισοδηματικές κατηγορίες 1500€ έως 2500€ και άνω των 2500€, τα μέσα σταθερής τροχιάς επιλέγονται για της καθημερινές μετακινήσεις, με ποσοστό 12% και 13% αντίστοιχα.

Οι μετακινήσεις με λεωφορείο εμφανίζουν μέγιστο στην εισοδηματική κατηγορία λιγότερα από 750€, με ποσοστό 23%. Στη συνέχεια, στην κατηγορία του μηδενικού εισοδήματος, το 19% των

μετακινήσεων πραγματοποιείται με το λεωφορείου. Ακολουθεί το 8% των μετακινήσεων της κατηγορίας 750-1500€. Της, με λεωφορείο διεξάγεται και το 6% των μετακινήσεων της ομάδας 1500-2500€. Στην κατηγορία με εισόδημα άνω των 2500€ δεν εμφανίζονται μετακινήσεις με χρήση λεωφορείου.

Στη συνέχεια, ακολουθεί, διάγραμμα κατανομής των μέσων μετακινήσεων αναλόγως αν οι συμμετέχοντες κατέχουν ιδιωτικό αυτοκίνητο.



Διάγραμμα 4.17 Ποσοστιαία κατανομή μέσων μετακίνησης αναλόγως την κατοχή αυτοκινήτου

Παρατηρείται ότι, οι κάτοχοι ιδιωτικής χρήσης επιβατικού αυτοκινήτου, το χρησιμοποιούν για το 54% των μετακινήσεων τους, ενώ οι μη-κάτοχοι αυτοκινήτου, μόνο για το 5% των μετακινήσεων τους. Το ταξί χρησιμοποιείται από την ομάδα των κατόχων αυτοκινήτου στο 3% των μετακινήσεων τους, κατά 8% λιγότερο από το ποσοστό όσων δεν έχουν αυτοκίνητο, το οποίο αγγίζει το 11%.

Οι μετακινήσεις όσων δεν έχουν στην κατοχή τους ΙΧΕ αυτοκίνητο σημειώνουν μεγάλα ποσοστά στη χρήση μέσων σταθερής τροχιάς και λεωφορείου, 29% και 25% αντίστοιχα στο κάθε μέσο. Στην ομάδα των κατόχων ΙΧΕ αυτοκινήτου το 18% των μετακινήσεων γίνεται με μέσα σταθερής τροχιάς και 9% των μετακινήσεων η χρήση λεωφορείου, αισθητά μικρότερα ποσοστά σε σχέση με την άλλη ομάδα.

Η μοτοσυκλέτα χρησιμοποιείται σε ποσοστό 7% των μετακινήσεων της ομάδας χωρίς αυτοκίνητο και σε ποσοστό 4% των μετακινήσεων της ομάδας με κατοχή αυτοκινήτου.

Μεγάλη απόκλιση παρατηρείται και στη χρήση ποδηλάτου για τις καθημερινές μετακινήσεις των ομάδων, καθώς η ομάδα κατόχων αυτοκινήτου, φαίνεται να χρησιμοποιεί το ποδήλατο στο 1% των μετακινήσεων της, ενώ η ομάδα ατόμων χωρίς αυτοκίνητο στο 5% των μετακινήσεων της. Το 18% των μετακινήσεων των ατόμων που δεν κατέχουν αυτοκίνητο πραγματοποιείται με τα πόδια και το 1% των μετακινήσεων τους με ηλεκτρικό πατίνι (escooter). Το 12% των μετακινήσεων των κατόχων αυτοκινήτου πραγματοποιείται με τα πόδια, ποσοστό μικρότερο από αυτό της ομάδας χωρίς ΙΧΕ αυτοκίνητο.

Σε αυτό το σημείο, αναλύθηκε η διάκριση χρήσης του εκάστοτε μέσου μεταφοράς αναλόγως το επίπεδο εκπαίδευσης των συμμετεχόντων. Υπήρξε μόνο ένα άτομο που δήλωσε απόφοιτος Δημοτικού Σχολείου, αυτό επηρέασε το διάγραμμα, όπως φαίνεται στην πρώτη ράβδο το 100% των μετακινήσεων διαξήχθησαν με χρήση λεωφορείου. Το συγκεκριμένο ποσοστό αφορά ουσιαστικά μόνο μία μετακίνηση που δήλωσε ότι εκτελεί αυτό το άτομο.

Η χρήση του αυτοκινήτου παρουσίασε ποσοστό 26% των μετακινήσεων στην ομάδα των απόφοιτων Λυκείου, 45% των μετακινήσεων των πτυχιούχων και 56% των μετακινήσεων των κατόχων μεταπτυχιακού/διδακτορικού τίτλου σπουδών. Παρατηρείται ραγδαία αύξηση της χρήσης αυτοκινήτου με την αύξηση του επιπέδου εκπαίδευσης.

Το ταξί χρησιμοποιείται στο 6% των μετακινήσεων της ομάδας των αποφοίτων Λυκείου, στο 3% των μετακινήσεων των πτυχιούχων και στο 4% των μετακινήσεων των κατόχων μεταπτυχιακού/διδακτορικού.

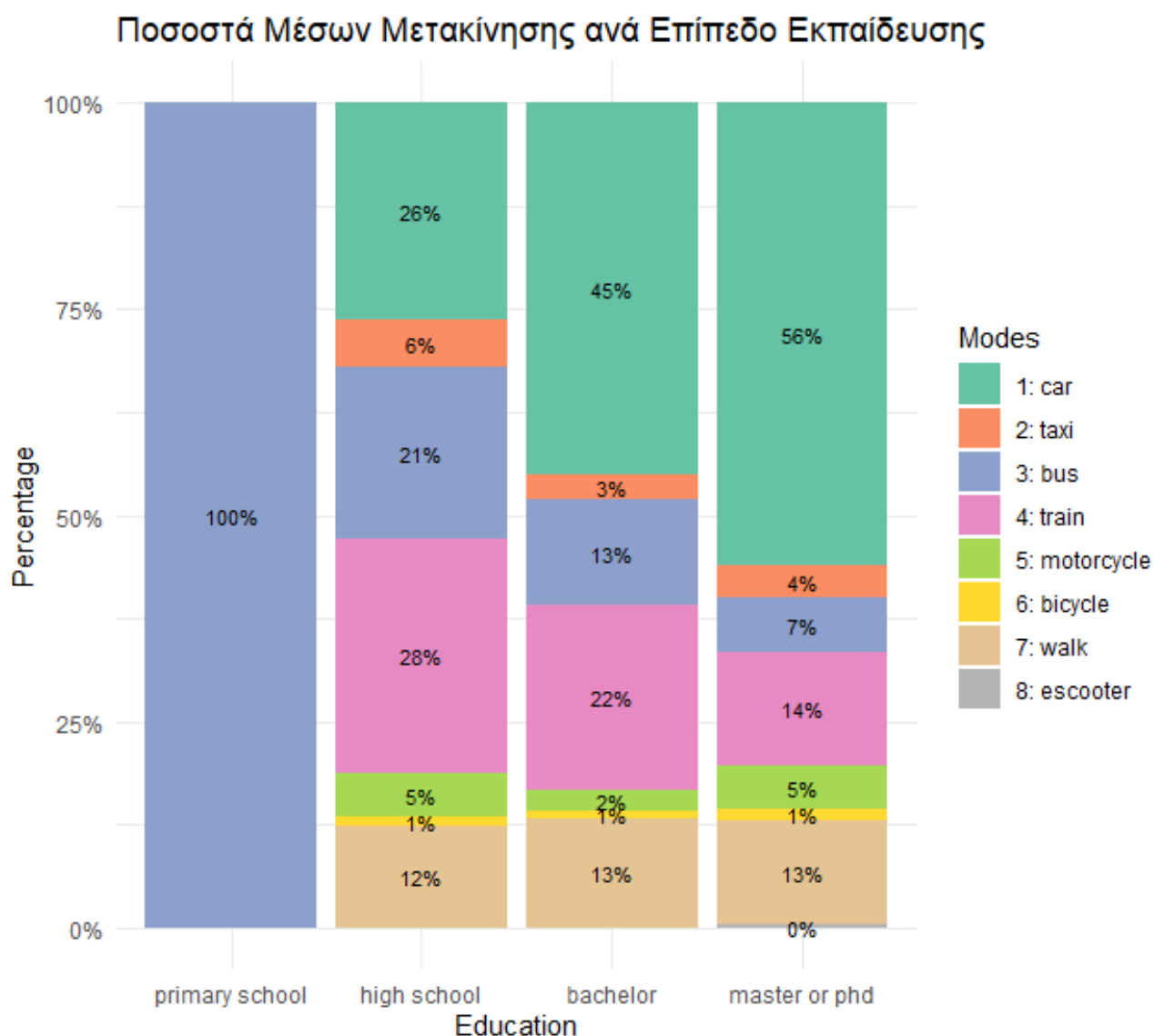
Το 21% των μετακινήσεων των συμμετεχόντων, που δήλωσαν απόφοιτοι Λυκείου, διεξάγεται με χρήση λεωφορείου, ποσοστό μεγαλύτερο από αυτό των πτυχιούχων που αγγίζει το 13% των μετακινήσεων και των κατόχων μεταπτυχιακού ή διδακτορικού που αγγίζει το 7% των μετακινήσεων τους. Παρατηρείται, μείωση της χρήσης του λεωφορείου στις καθημερινές μετακινήσεις των ατόμων με ανώτερο εκπαιδευτικό επίπεδο.

Όσον αφορά τα μέσα σταθερής τροχιάς, το μεγαλύτερο ποσοστό χρήσης τους εμφανίζεται στις μετακινήσεις της ομάδας των αποφοίτων Λυκείου, κατά 28%. Ακολουθούν οι ομάδες των πτυχιούχων με ποσοστό 22% και αυτή των κατόχων μεταπτυχιακού ή διδακτορικού τίτλου σπουδών με ποσοστό 14%.

Η μοτοσυκλέτα εξυπηρετεί το 5% των μετακινήσεων των αποφοίτων Λυκείου, το 2% των μετακινήσεων των πτυχιούχων και το 5% των μετακινήσεων των κατόχων μεταπτυχιακού ή διδακτορικού. Κατά μήκος και των τριών επιπέδων εκπαίδευσης σημειώνεται 1% χρήση ποδηλάτου για τις καθημερινές μετακινήσεις.

Με ποσοστό 12% εκφράζονται οι μετακινήσεις που διενεργούνται με περπάτημα στην ομάδα των αποφοίτων λυκείου και με ποσοστό 13% οι μετακινήσεις των πτυχιούχων, καθώς και των κατόχων μεταπτυχιακού ή διδακτορικού τίτλου σπουδών. Σημειώνεται ότι, και στις τρεις κατηγορίες η μετακίνηση με τα πόδια εμφανίζει παρόμοια ποσοστά.

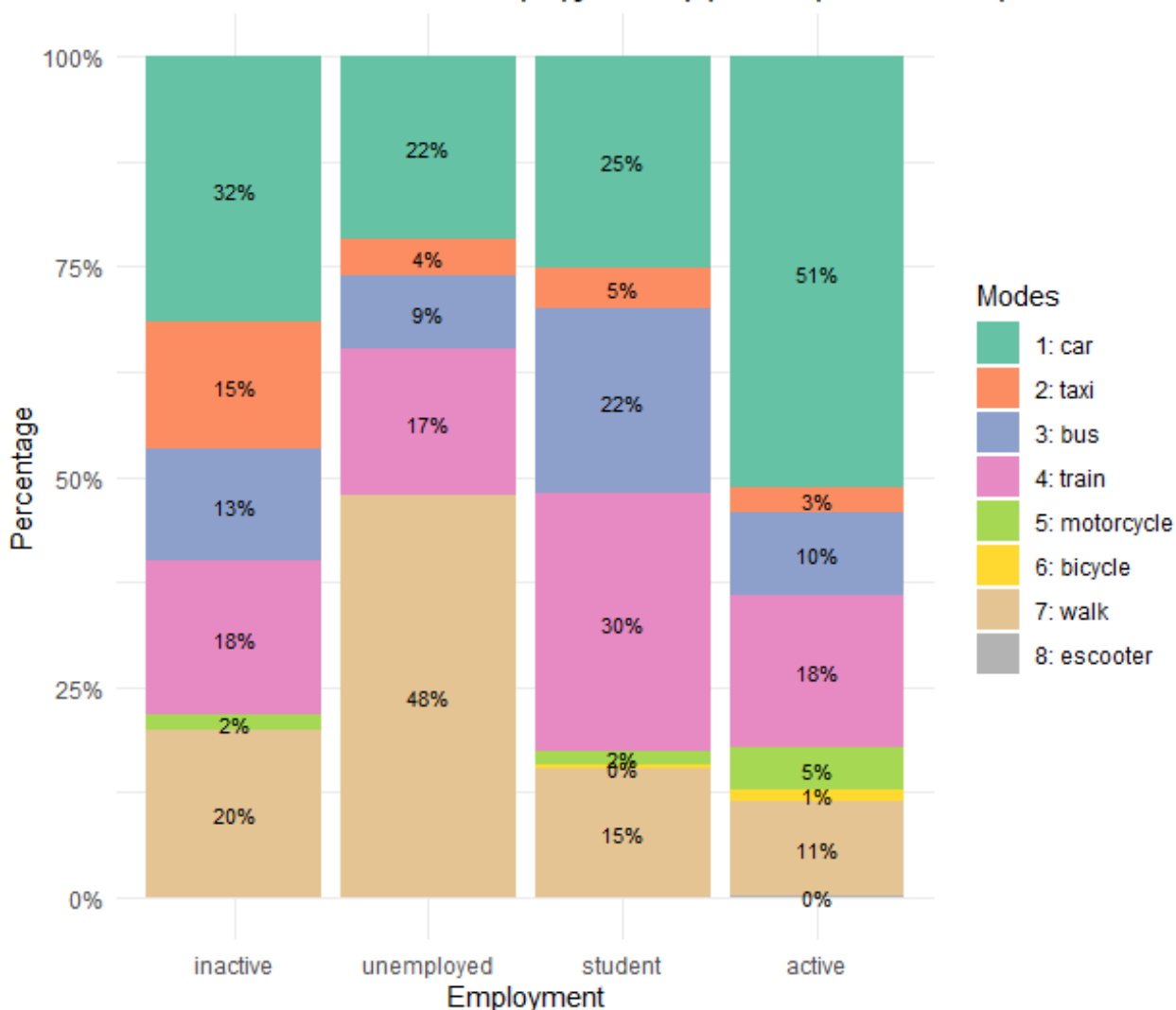
Τέλος, η χρήση του escooter για τις καθημερινές μετακινήσεις εμφανίζεται μόνο στην κατηγορία των κατόχων μεταπτυχιακού/ διδακτορικού, με ποσοστό μικρότερο από το 1%. Αυτό συμβαίνει καθώς ήταν ελάχιστοι οι αποκρινόμενοι που δήλωσαν τη χρήση escooter σε καθημερινή βάση.



Διάγραμμα 4.18 Ποσοστιαία κατανομή μέσων μετακίνησης ανά επίπεδο εκπαίδευσης

Παρακάτω, δημιουργήθηκε ένα συνδυαστικό διάγραμμα των μεταβλητών του μέσου μεταφοράς και της εργασιακής κατάστασης.

Ποσοστά Μέσων Μετακίνησης ανά Εργασιακή Κατάσταση



Διάγραμμα 4.19 Ποσοστιαία κατανομή μέσων μετακίνησης ανά εργασιακή κατάσταση

Όπως φαίνεται στην συνέχεια, η ομάδα των άεργων χρησιμοποιεί, με φθίνουσα σειρά, στο 32% των μετακινήσεων της το αυτοκίνητο, στο 20% των μετακινήσεων της το περπάτημα, στο 18% των μετακινήσεων τα μέσα σταθερής τροχιάς, στο 15% των μετακινήσεων ταξί, στο 13% των μετακινήσεων το λεωφορείο και κατά 5% την μοτοσυκλέτα. Να σημειωθεί ότι απουσιάζει η χρήση ποδηλάτου και ηλεκτρονικού πατινιού (escooter).

Η επόμενη ομάδα, που αφορά τους άνεργους, φαίνεται να σημειώνει μέγιστο στις μετακινήσεις με περπάτημα, σε ποσοστό 48%, μεγαλύτερο από όλες τις υπόλοιπες κατηγορίες. Ακολουθούν, οι μετακινήσεις με αυτοκίνητο σε ποσοστό 22%, οι μετακινήσεις με μέσα σταθερής τροχιάς σε ποσοστό 17%, οι μετακινήσεις με λεωφορείο, με ποσοστό 9% και οι μετακινήσεις με ταξί, με ποσοστό 4%. Απουσιάζουν οι μετακινήσεις με μοτοσυκλέτα, ποδήλατο και escooter.

Η ομάδα των συμμετεχόντων που καταγράφηκαν ως μαθητές-σπουδαστές, στην πλειοψηφία της, χρησιμοποιεί μέσα σταθερής τροχιάς για τις καθημερινές μετακινήσεις, με ποσοστό 30%. Σε φθίνουσα σειρά αναφέρονται οι μετακινήσεις με αυτοκίνητο σε ποσοστό 25%, με λεωφορείο σε ποσοστό 22% , με περπάτημα σε ποσοστό 15%, και με 5% εμφανίζονται οι μετακινήσεις με χρήση ταξί. Οι μετακινήσεις με μοτοσυκλέτα καταλαμβάνουν το 2% του δείγματος ενώ οι μετακινήσεις με ποδήλατο καταλαμβάνουν ποσοστό μικρότερο του 1%. Δεν σημειώνονται μετακινήσεις με ηλεκτρικό πατίνι (escooter).

Στην τελευταία κατηγορία, αυτή των κατόχων μεταπτυχιακού ή διδακτορικού τίτλου σπουδών, εμφανίζει μέγιστο η χρήση του αυτοκινήτου με το 51% των μετακινήσεων να διεξάγεται με χρήση αυτού. Το 18% των μετακινήσεων διενεργείται με τα μέσα σταθερής τροχιάς, ενώ το 10% των μετακινήσεων πραγματοποιείται με το λεωφορείο. Μόνο το 11% των μετακινήσεων γίνεται με περπάτημα. Αυξημένο εμφανίζεται το ποσοστό των μετακινήσεων με μοτοσυκλέτα, σε σχέση με τις υπόλοιπες ομάδες, στο 5%. Μείωση σημειώνεται στη χρήση ταξί με ποσοστό 3%. Το ποδήλατο χρησιμοποιείται σε ποσοστό 1% των μετακινήσεων, μεγαλύτερο από κάθε άλλη κατηγορία. Σε αυτή την ομάδα και μόνο εμφανίζεται το ηλεκτρικό πατίνι (escooter) και καλύπτει ποσοστό μικρότερο από το 1% των μετακινήσεων

Παρακάτω, παρουσιάζεται το διάγραμμα που περιλαμβάνει το μέσο μετακίνησης συναρτήσει του σκοπού μετακίνησης στην Αθήνα. Εκ πρώτης όψεως, εντοπίζεται η χρήση του αυτοκινήτου, με αυτό να κυριαρχεί σαν μέσο μετακίνησης σε όλους τους σκοπούς εκτός από τη μετάβαση για εκπαιδευτικούς λόγους.

Ειδικότερα, στην πρώτη κατηγορία που αφορά τη μετακίνηση προς την εργασία, το 50% των μεταβάσεων πραγματοποιείται με το αυτοκίνητο, ενώ στη δεύτερη θέση με ποσοστό 18%, βρίσκονται οι μετακινήσεις με μέσα σταθερής τροχιάς. Ακολουθεί, το λεωφορείο (12%) και στη συνέχεια η μετακίνηση με περπάτημα (9%). Τέλος, τα μικρότερα ποσοστά του δείγματος είναι η μετάβαση στην εργασία με μοτοσυκλέτα (5%) ενώ μόνο το 2% χρησιμοποιεί τα ταξί και τα ποδήλατα.

Ο δεύτερος σκοπός μετακίνησης είναι εκείνος της επιστροφής στον τόπο κατοικίας. Το 46% αυτών των μετακινήσεων πραγματοποιείται με αυτοκίνητο ενώ το 20% με μέσα σταθερής τροχιάς. Ακολουθεί, η μετάβαση στην κατοικία με περπάτημα (13%), με το λεωφορείο (11%), ενώ αισθητά μικρότερα είναι τα ποσοστά του ταξί, της μοτοσυκλέτας και του ποδηλάτου με ποσοστά 5%, 4% και 1% αντίστοιχα.

Ο επόμενος σκοπός μετακίνησης είναι για εκπαιδευτικούς λόγους, με το μεγαλύτερο ποσοστό των μετακινήσεων να συγκεντρώνεται στα μέσα σταθερής τροχιάς (34%). Στη δεύτερη θέση βρίσκεται το αυτοκίνητο με ποσοστό 32%, ενώ την τρίτη θέση καταλαμβάνει η μετακίνηση με το λεωφορείο (19%). Στις τρεις τελευταίες θέσεις βρίσκονται οι μετακινήσεις προς τον τόπο εκπαίδευσης με το βάδισμα, τη

μοτοσυκλέτα και το ταξί με ποσοστά 11%, 3% και 2% αντίστοιχα. Σημειώνεται ότι, για εκπαιδευτικούς λόγους δεν παρατηρούνται μετακινήσεις με χρήση escooter.

Ο τέταρτος σκοπός μετακίνησης αφορά τις αγορές. Σε αυτή την κατηγορία μετακινήσεων το μεγαλύτερο ποσοστό παρατηρείται ξανά στο αυτοκίνητο (44%). Στη συνέχεια, ακολουθεί, η μετακίνηση με περπάτημα, με ποσοστό 20%, ενώ το τρίτο μεγαλύτερο ποσοστό εντοπίζεται στις μετακινήσεις με μέσα σταθερής τροχιάς (15%). Ακολουθεί το λεωφορείο, με ποσοστό χρήσης 12%, ενώ στις τρεις τελευταίες θέσεις βρίσκονται ξανά η μοτοσυκλέτα, το ταξί και το ποδήλατο με ποσοστά 5%, 4% και 1% αντίστοιχα.

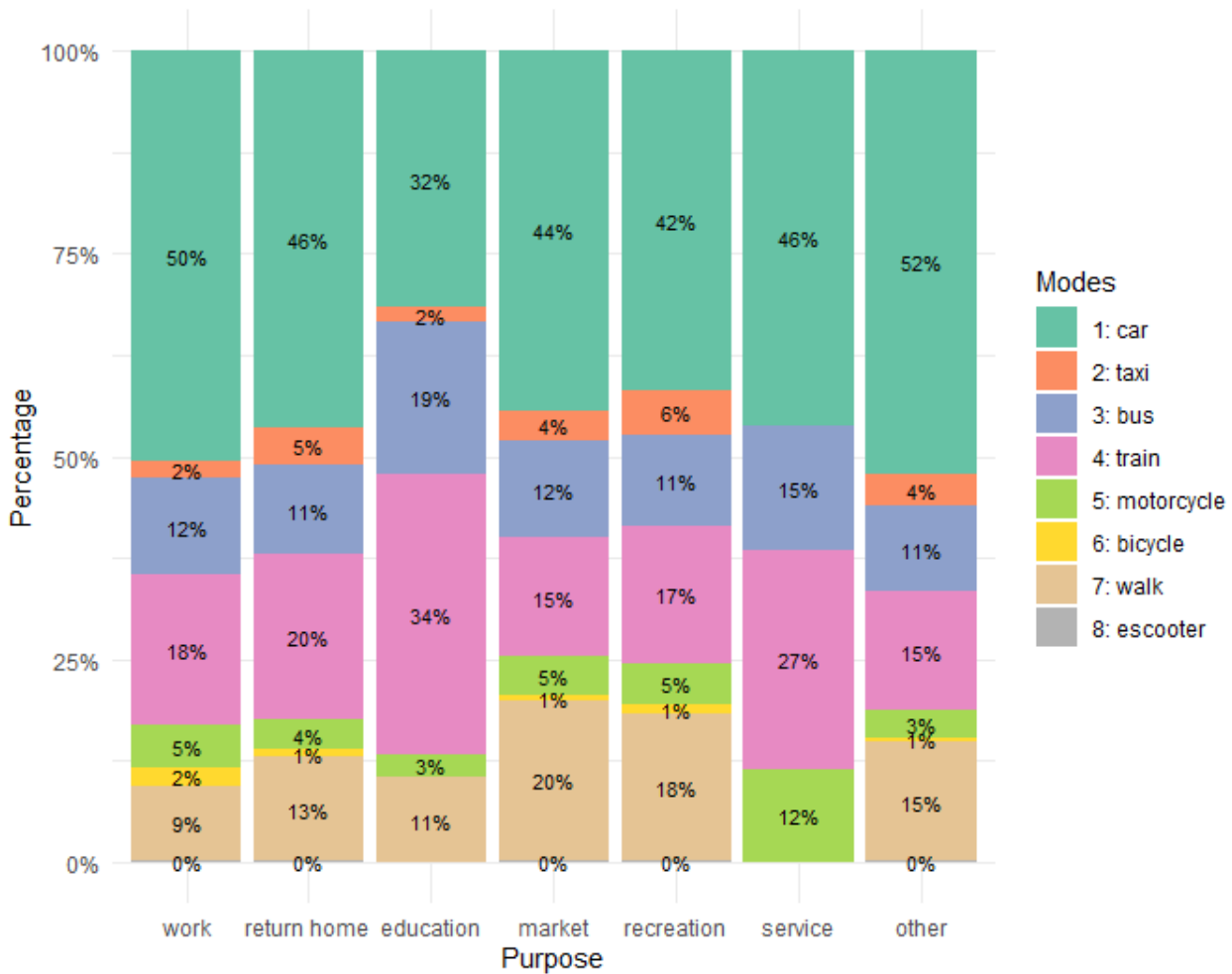
Ο πέμπτος σκοπός μετακίνησης αφορά την αναψυχή. Σε αυτή την κατηγορία το αυτοκίνητο πρωτοστατεί ξανά με ποσοστό 42%. Δεύτερη έρχεται η μετακίνηση με περπάτημα συγκεντρώνοντας ποσοστό 18%. Εν συνεχεία, τρίτη ακολουθεί η μετακίνηση με τα μέσα σταθερής τροχιάς, με ποσοστό 17%. Την τέταρτη θέση καταλαμβάνει το λεωφορείο με ποσοστό 11%. Στις τρεις τελευταίες θέσεις βρίσκεται το ταξί, η μοτοσυκλέτα και το ποδήλατο, όπως και στον προηγούμενο σκοπό μετακίνησης, με ποσοστά 6%, 5% και 1% αντίστοιχα.

Όσον αφορά τη μετάβαση σε υπηρεσίες παρατηρούνται μόνο τέσσερα διαφορετικά μέσα μετακίνησης. Στην πλειοψηφία τους οι μετακινήσεις διεξάγονται με αυτοκίνητο, σε ποσοστό 46%, ενώ ακολουθούν οι μετακινήσεις με μέσα σταθερής τροχιάς, σε ποσοστό 27%. Στις δύο τελευταίες θέσεις εμφανίζονται οι μετακινήσεις με λεωφορείο και μοτοσυκλέτα με ποσοστά 15% και 12% αντίστοιχα. Επίσης, η μοτοσυκλέτα σε αυτή την κατηγορία μετακινήσεων σημειώνει μέγιστο. Παρατηρείται ότι, σε αυτή την κατηγορία μετακινήσεων δεν εμφανίζεται καθόλου η χρήση ποδηλάτου, ταξί, escooter και η πεζή μετακίνηση.

Ο τελευταία κατηγορία αναφέρεται σε λοιπούς σκοπούς μετακίνησης. Στη συγκεκριμένη κατηγορία η μετακίνηση με αυτοκίνητο παρουσιάζει το μεγαλύτερο ποσοστό σε σχέση με όλους τις υπόλοιπες, με ποσοστό 52%. Ακολουθεί, η μετακίνηση με μέσα σταθερής τροχιάς με ποσοστό 15%. Ομοίως και οι πεζές μετακινήσεις συγκεντρώνουν ποσοστό 15%. Στη συνέχεια, το λεωφορείο χρησιμοποιείται για το 11% των συνολικών μετακινήσεων αυτής της κατηγορίας. Τελευταίες βρίσκονται ξανά οι μετακινήσεις με ταξί, μοτοσυκλέτα και ποδήλατο, συγκεντρώνοντας ποσοστά 4%, 3% και 1% αντίστοιχα.

Στο διάγραμμα διάκρισης μέσων μετακίνησης ανά σκοπό μετακίνησης φαίνεται ότι, στους σκοπούς μετακίνησης «εργασία», «επιστροφή στην οικία», «αγορές», «αναψυχή» και «άλλο» εντοπίζεται και η χρήση ηλεκτρικού πατινιού (escooter) ως μέσο στις καθημερινές μετακινήσεις, με ποσοστά μικρότερα του 1%.

Ποσοστά Μέσων Μετακίνησης ανά Σκοπό Μετακίνησης



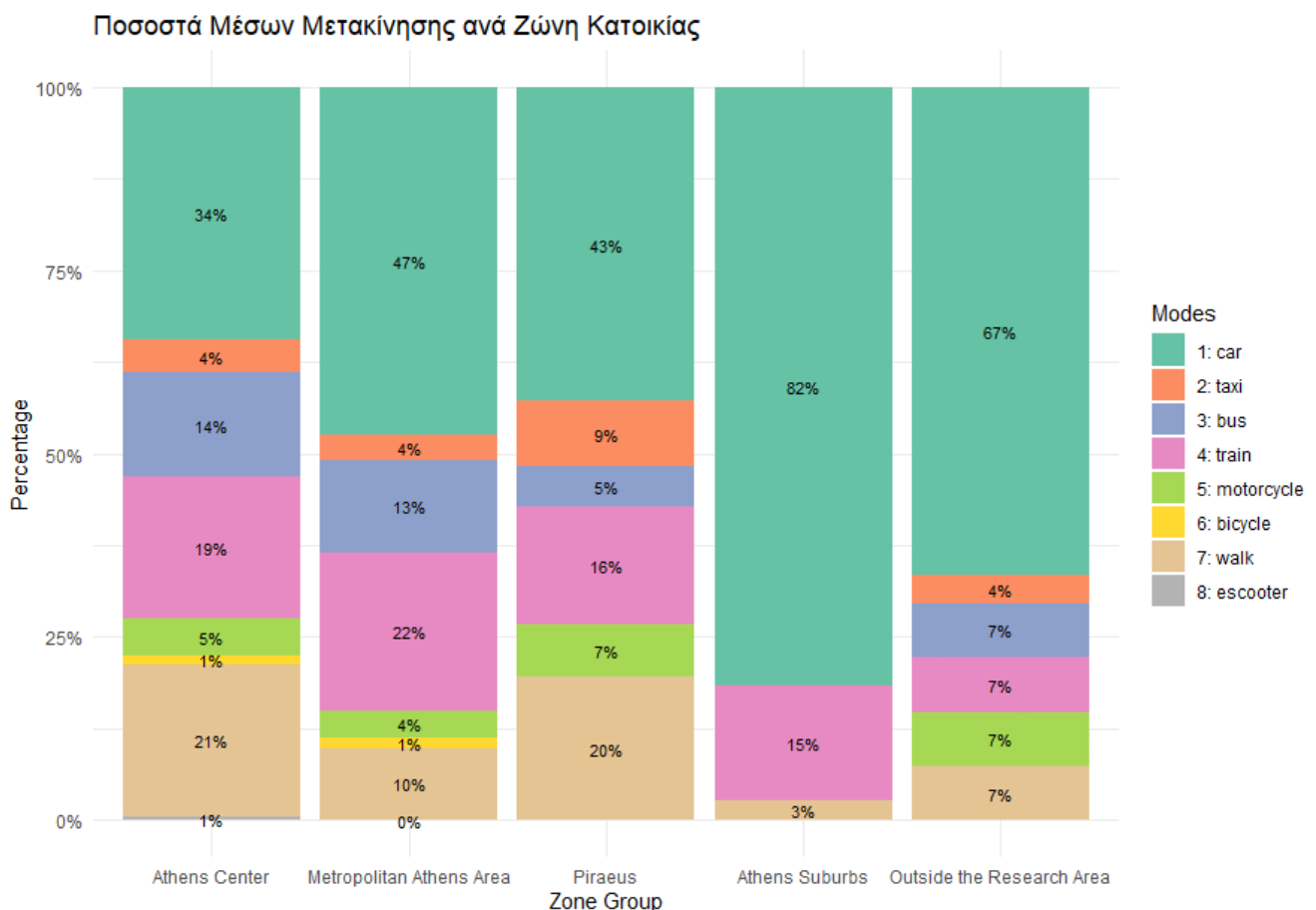
Διάγραμμα 4.20 Ποσοστιαία κατανομή μέσων μετακίνησης ανά σκοπό μετακίνησης

Όπως σχολιάστηκε και στις προηγούμενες ενότητες η ευρύτερη περιοχή της Αθήνας χωρίστηκε σε 36 διαφορετικές ζώνες, οι οποίες εμπίπτουν σε 4 κατηγορίες ομαδοποίησης. Αυτές οι ομάδες ζωνών είναι: το κέντρο της Αθήνας, η ευρύτερη μητροπολιτική περιοχή των Αθηνών, ο Πειραιάς, τα προάστια και οι περιοχές εκτός των Αθηνών. Η ομάδα του κέντρου της Αθήνας περιλαμβάνει τις ζώνες: 1,2,3,4,5,6,7. Η ευρύτερη μητροπολιτική περιοχή των Αθηνών συμπεριλαμβάνει τις ζώνες: 8, 9, 11, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 34, 35. Η ομάδα των προαστίων της Αθήνας περιέχει τις ζώνες: 10, 12, 17 και 33. Η ομάδα Πειραιάς αναφέρεται στη ζώνη 30 που αφορά τον Πειραιά. Τέλος, η ομάδα που αφορά την περιοχή εκτός της Αθήνας περιλαμβάνει τα ερωτηματολόγια των κατοίκων της ζώνης 36. Σε αυτό το σημείο παρατίθεται ο πίνακας με το πλήθος των μετακινήσεων αναλόγως τη ζώνη κατοικίας του εκάστοτε μετακινούμενου.

Πίνακας 4.4 Κατανομή πλήθους μετακινήσεων ανά περιοχή κατοικίας

Ζώνη Κατοικίας	Κέντρο Αθηνών	Μητροπολιτική Περιοχή Αθήνας	Πειραιάς	Προάστια	Εκτός Αθηνών
Πλήθος Μετακινήσεων	380	813	56	71	27
Ποσοστό	29%	60%	4%	5%	2%

Βάσει των ερωτηματολογίων της έρευνας προκύπτει ότι, 380 μετακινήσεις εκτελούνται από κατοίκους του κέντρου της Αθήνας, 813 μετακινήσεις διεξάγονται από τους κατοίκους της μητροπολιτικής περιοχής της Αθήνας, 56 ταξίδια πραγματοποιούνται από άτομα που διαμένουν στον Πειραιά, 71 ταξίδια γίνονται από κατοίκους των προαστίων της Αθήνας και μόνο 27 μετακινήσεις διεξάγονται από συμμετέχοντες που διαμένουν εκτός Αθηνών. Στο σύνολο τους οι μετακινήσεις είναι 1347. Στο παρακάτω διάγραμμα φαίνονται τα μέσα μετακίνησης που χρησιμοποιούνται για τις μετακινήσεις των ερωτηθέντων αναλόγως την ομαδοποιημένη ζώνη κατοικίας τους.



Διάγραμμα 4.21 Ποσοστιαία κατανομή μέσων μετακίνησης ανά ζώνη κατοικίας

Μέσω του διαγράμματος παρατηρείται ότι, οι κάτοικοι του κέντρου των Αθηνών χρησιμοποιούν το αυτοκίνητο λιγότερο από τους κατοίκους των υπόλοιπων ζωνών, με ποσοστό 34%. Επίσης, σε αυτή

την ομάδα παρουσιάζεται μέγιστο στις μετακινήσεις με περπάτημα, συγκεντρώνοντας ποσοστό που αγγίζει το 21%. Το 19% των μετακινήσεων των κατοίκων του κέντρου των Αθηνών, διεξάγεται με μέσα σταθερής τροχιάς και το 14% με λεωφορείο. Με χρήση ταξί πραγματοποιείται το 4% των μετακινήσεων των κατοίκων αυτής της ομάδας. Μόνο το 5% των μετακινήσεων γίνεται με χρήση μοτοσυκλέτας. Το ποδήλατο και το escooter συγκεντρώνουν το 1% των μετακινήσεων το εκάστοτε μέσο.

Στην επόμενη κατηγορία, που αφορά τις μετακινήσεις των κατοίκων της μητροπολιτικής περιοχής της Αθήνας, σημειώνει μέγιστο η χρήση αυτοκινήτου, με ποσοστό 47%. Οι μετακινήσεις με μέσα σταθερής τροχιάς συγκεντρώνουν ποσοστό 22% και με λεωφορείο 13%. Με περπάτημα διεξάγεται το 10% και με ταξί το 4% των μετακινήσεων. Το 4% συγκεντρώνουν και οι μετακινήσεις με μοτοσυκλέτα. Ακολουθούν, οι μετακινήσεις με ποδήλατο που αγγίζουν το 1% των μετακινήσεων της ομάδας και με ποσοστό μικρότερο του 1% εμφανίζονται οι μετακινήσεις με escooter.

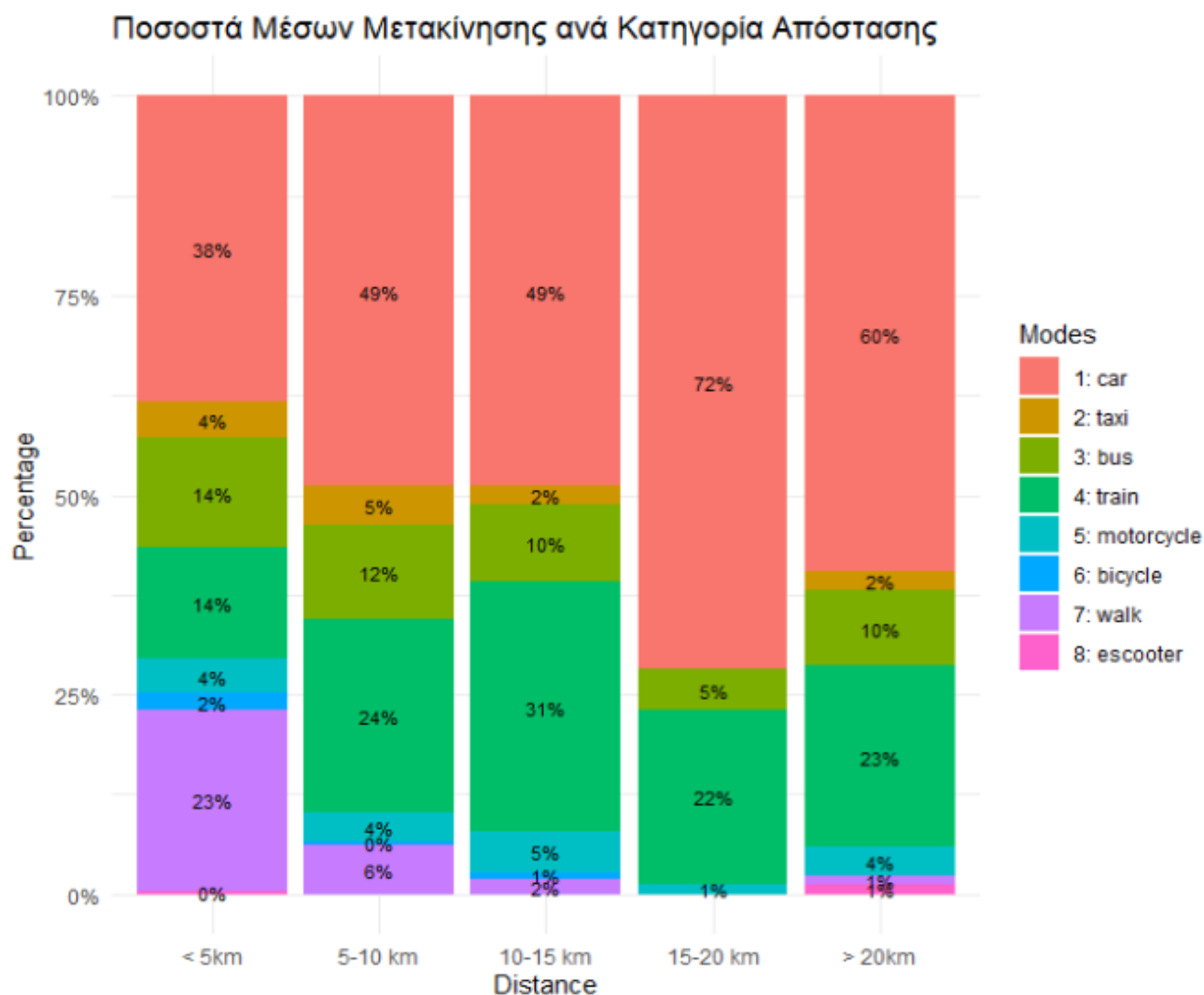
Όσον αφορά την περιοχή του Πειραιά, παρατηρείται ότι το 43% των μετακινήσεων των κατοίκων του διεξάγεται με χρήση αυτοκινήτου. Υψηλό ποσοστό συγκεντρώνουν και οι μετακινήσεις με περπάτημα στο Πειραιά (20%). Με 16% εμφανίζονται οι μετακινήσεις με μέσα σταθερής τροχιάς, ενώ μόνο το 5% των μετακινήσεων των κατοίκων του Πειραιά διεξάγεται με το λεωφορείο. Μέγιστο μεταξύ των ομάδων σημειώνεται στον Πειραιά με τη χρήση ταξί να συγκεντρώνει ποσοστό 9%. Στη συνέχεια, οι μετακινήσεις με μοτοσυκλέτα καταλαμβάνουν το 7% του συνόλου των μετακινήσεων αυτής της ομάδας. Σε αυτή την ομάδα μετακινήσεων απουσιάζουν το ποδήλατο και το escooter. Φυσικά, οι μετακινήσεις του δείγματος που αφορούν τους κατοίκους του Πειραιά (56) είναι αρκετά λιγότερες σε σχέση με το πλήθος των μετακινήσεων στο κέντρο της Αθήνας και στην ευρύτερη μητροπολιτική περιοχή της Αθήνας.

Ακολουθεί η ομάδα των προαστίων των Αθηνών, όπου η πλειοψηφία των μετακινήσεων διεξάγεται με ιδιωτικής χρήσης αυτοκίνητο, με ποσοστό 82%. Σε αυτή την ομάδα τα μέσα σταθερής τροχιάς χρησιμοποιούνται για το 15% και το περπάτημα για το 3% των μετακινήσεων. Από την ομάδα των κατοίκων των προαστίων της Αθήνας απουσιάζουν: η χρήση λεωφορείου, ταξί, μοτοσυκλέτας, ποδηλάτου και escooter.

Τελευταία κατηγορία είναι αυτή των κατοίκων περιοχών εκτός Αθηνών. Το πλήθος των μετακινήσεων των ερωτηθέντων που διαμένουν εκτός της ευρύτερης περιοχής της Αθήνας είναι μόλις 27. Όσον αφορά το μέσο μετακίνησης τους, το αυτοκίνητο συγκεντρώνει το 67% των μετακινήσεων τους. Το λεωφορείο καθώς και τα μέσα σταθερής τροχιάς, η μοτοσυκλέτα και το περπάτημα συγκεντρώνουν το καθένα ποσοστό 7%. Το ταξί εμφανίζεται σε ποσοστό 4% των μετακινήσεων της ομάδας. Όπως φαίνεται και στο διάγραμμα, απουσιάζουν οι μετακινήσεις με χρήση ηλεκτρικού πατινιού (escooter).

Στην συνέχεια, παρουσιάζεται το ραβδόγραμμα των ποσοστών των μέσων μετακίνησης αναλόγως την απόσταση του ταξιδιού. Με στόχο την ευκολότερη ανάγνωση των δεδομένων, οι αποστάσεις των

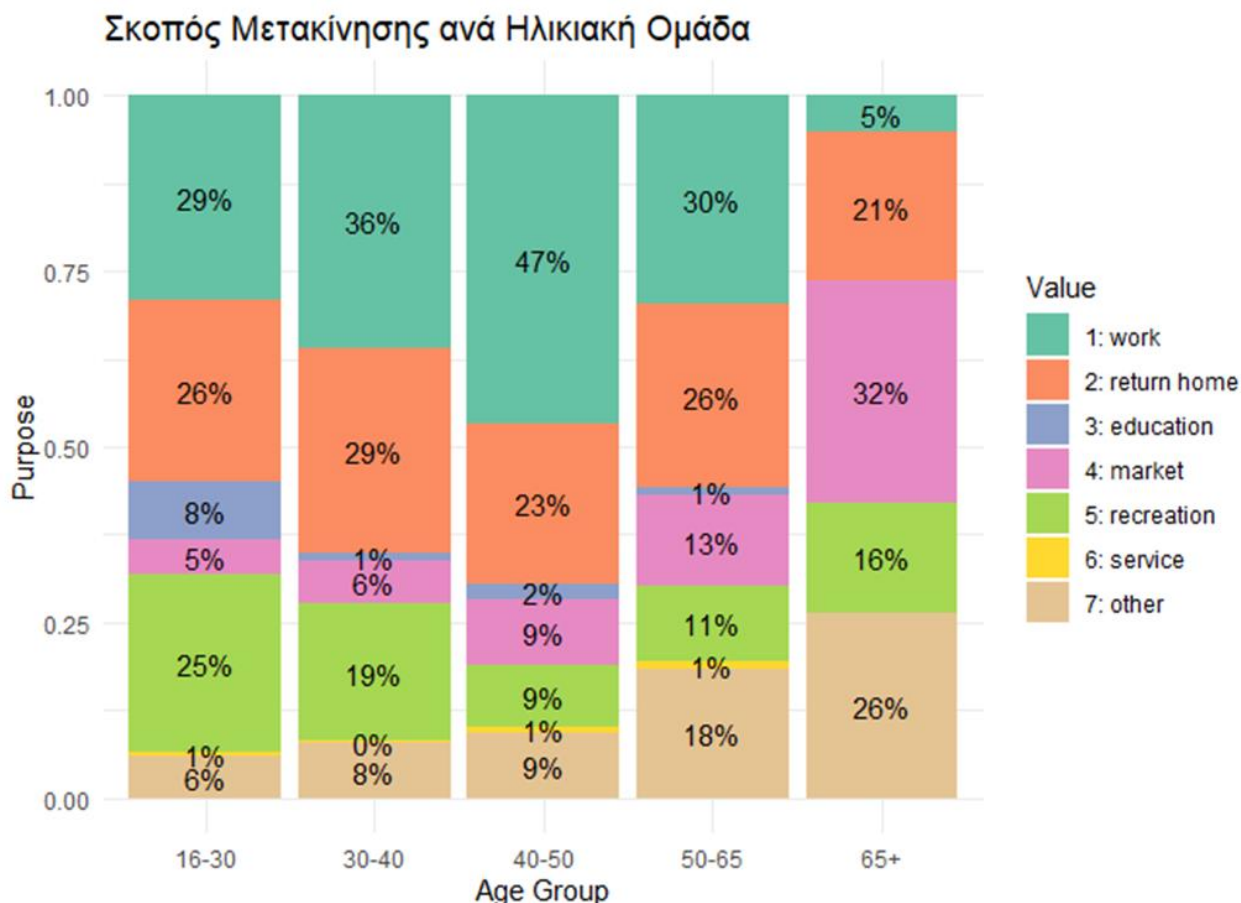
μετακινήσεων, κατηγοριοποιήθηκαν σε μετακινήσεις με απόσταση έως 5 χιλιόμετρα, από 5 έως 10 χιλιόμετρα, από 10 έως 15 χιλιόμετρα, από 15 έως 20 χιλιόμετρα και μεγαλύτερη από 20 χιλιόμετρα.



Διάγραμμα 4.22 Ποσοστιαία κατανομή μέσων μετακίνησης αναλόγως την απόσταση

Παρατηρείται ότι, στην κατηγορία των αποστάσεων μικρότερων των 5 χιλιομέτρων, οι μετακινήσεις διεξάγονται κυρίως με αυτοκίνητο στο 38%, όμως πολύ μεγάλο ποσοστό σημειώνουν και οι μετακινήσεις με περπάτημα (23%), ειδικά σε σχέση με τις υπόλοιπες αποστάσεις. Οι μετακινήσεις με απόσταση μεταξύ των 5-10 χιλιομέτρων διεξάγονται κυρίως με αυτοκίνητο (49%) και με τη Δημόσια Συγκοινωνία (36%). Στην συνέχεια, οι μετακινήσεις με απόσταση από 10 έως 15 χιλιόμετρα ως επί το πλείστον πραγματοποιούνται με το αυτοκίνητο (49%) και τα Μέσα Μαζικής Μεταφοράς (41%). Στην επόμενη κατηγορία μετακινήσεων, με απόσταση από 15 έως 20 χιλιόμετρα, αυτές γίνονται μέσω του αυτοκινήτου στην μεγάλη τους πλειοψηφία (72%). Τέλος, οι μετακινήσεις με αποστάσεις μεγαλύτερες των 20 χιλιομέτρων διεξάγονται είτε με αυτοκίνητο, με ποσοστό 60%, είτε με χρήση της Δημόσιας Συγκοινωνίας, με ποσοστό 34%.

Ακολουθεί το διάγραμμα ποσοστιαίας κατανομής των σκοπών μετακίνησης ανά ηλικιακή ομάδα. Οι σκοποί των ταξιδιών μπορεί να είναι: 1: εργασία, 2: επιστροφή στο σπίτι, 3: εκπαίδευση, 4: αγορά, 5: αναψυχή, 6: υπηρεσία και 7: άλλο.



Διάγραμμα 4.23 Ποσοστιαία κατανομή σκοπών μετακίνησης ανά ηλικιακή ομάδα

Στην πρώτη ηλικιακή ομάδα (16-30 ετών), επικρατούν οι μετακινήσεις για σκοπούς εργασίας με 29% και ακολουθούν οι μετακινήσεις για επιστροφή στο σπίτι με ποσοστό 26%. Το 25% των μετακινήσεων της πρώτης ηλικιακής ομάδας αφορούν μετακινήσεις για αναψυχή. Με 8% εμφανίζονται οι μετακινήσεις που αφορούν την εκπαίδευση, με 5% εμφανίζονται οι μετακινήσεις για αγορές και με 1% οι μετακινήσεις για πρόσβαση σε υπηρεσίες. Οι «άλλοι» σκοποί μετακίνησης καταλαμβάνουν το 6% των μετακινήσεων της ομάδας.

Όσον αφορά τη δεύτερη ηλικιακή ομάδα (30-40 ετών), επικρατούν οι μετακινήσεις προς την εργασία, με ποσοστό 36%, ακολουθούν οι μετακινήσεις για επιστροφή στο σπίτι, με ποσοστό 29% και οι μετακινήσεις για λόγους αναψυχής σημειώνουν ποσοστό 19%. Οι μετακινήσεις για αγορές συγκεντρώνουν ποσοστό 6% και οι μετακινήσεις για εκπαιδευτικούς σκοπούς ποσοστό 1%. Με ποσοστό μικρότερο του 1% εμφανίζονται οι μετακινήσεις προς υπηρεσίες σε αυτή την ηλικιακή ομάδα. Οι «άλλοι» σκοποί μετακίνησης απαντούν στο 8% των μετακινήσεων της ομάδας.

Το 47% των μετακινήσεων της τρίτης ηλικιακής ομάδας (40-50 ετών) αφορά τη μετάβαση στο χώρο εργασίας, ενώ το 23% των μετακινήσεων αφορά την επιστροφή στο σπίτι. Με ποσοστό 9% εμφανίζονται οι μετακινήσεις για αγορές και οι μετακινήσεις για αναψυχή. Οι μετακινήσεις για εκπαιδευτικούς σκοπούς αφορούν το 2% των μετακινήσεων αυτής της ομάδας, ενώ οι μετακινήσεις για μετάβαση σε υπηρεσίες μόνο το 1%. Ο σκοπός μετακίνησης «άλλο» σημειώνει ποσοστό 9% επί του συνόλου των μετακινήσεων της ηλικιακής ομάδας.

Η τέταρτη ηλικιακή ομάδα (50-65 ετών) εμφανίζει μέγιστο στις μετακινήσεις προς την εργασία, με ποσοστό 30%. Ύστερα, ακολουθούν οι μετακινήσεις για επιστροφή στο σπίτι, με ποσοστό 26%. Σε αυτή την ηλικιακή ομάδα σημειώνουν μεγάλη αύξηση οι μετακινήσεις για «άλλους» σκοπούς, με ποσοστό 18%. Οι μετακινήσεις για αγορές συγκεντρώνουν το 13% των μετακινήσεων, για αναψυχή το 11% των μετακινήσεων. Με πολύ μικρά ποσοστά της τάξεως του 1% εμφανίζονται οι μετακινήσεις για εκπαιδευτικούς σκοπούς και για πρόσβαση σε υπηρεσίες.

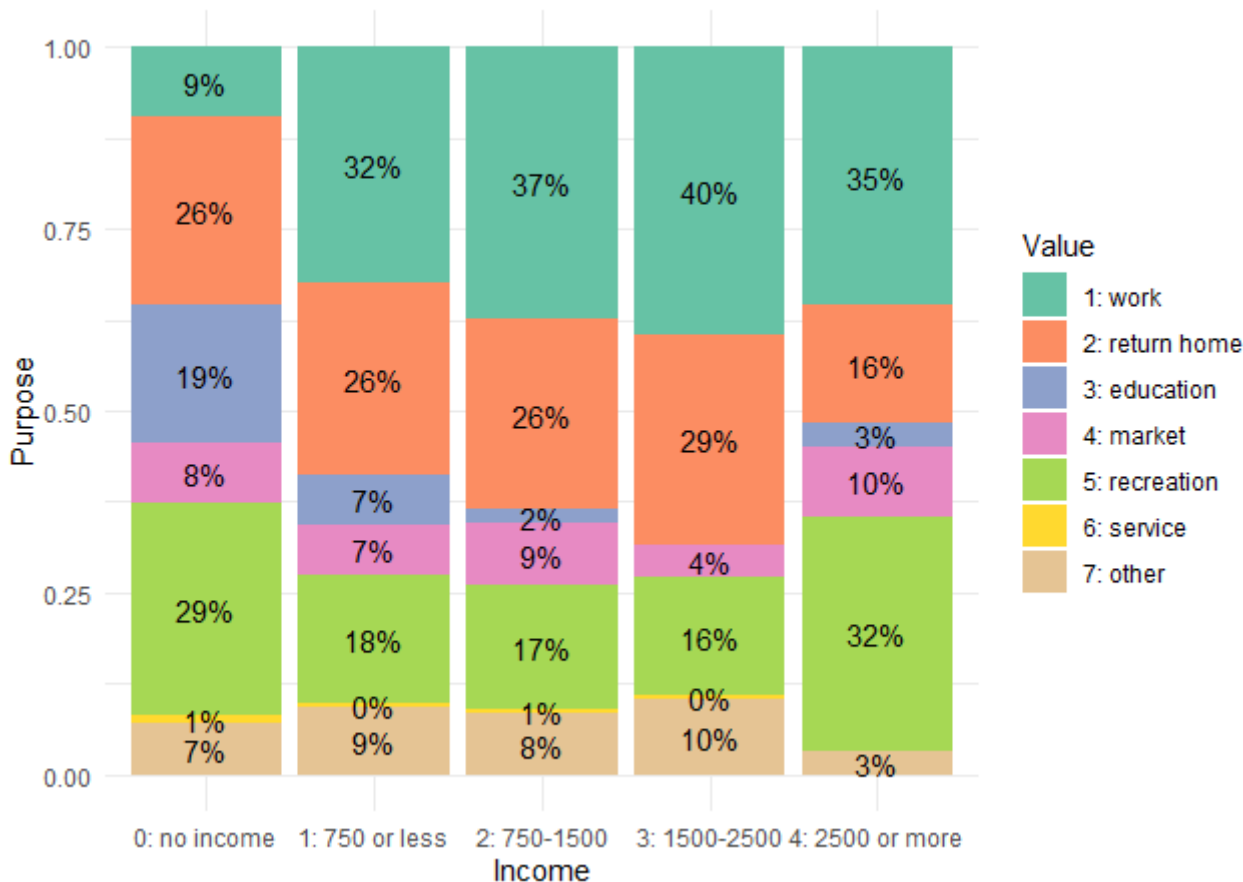
Στην τελευταία ηλικιακή ομάδα, που αφορά άτομα άνω των 65 ετών, επικρατούν οι μετακινήσεις για αγορές. Μέγιστο σε σχέση με τις υπόλοιπες ομάδες σημειώνουν σε αυτή την ομάδα οι μετακινήσεις για «άλλους» σκοπούς μετακίνησης εμφανίζοντας ποσοστό 26%. Η μετακίνηση για επιστροφή στο σπίτι συγκεντρώνει το 21% των μετακινήσεων της ομάδας και η μετακίνηση για αναψυχή συγκεντρώνει το 16% του συνόλου των μετακινήσεων αυτής της ηλικιακής ομάδας. Ελάχιστο σημειώνει η μετακίνηση για μετάβαση στην εργασία καταγράφοντας σε ποσοστό του 5% των μετακινήσεων.

Εν συνεχεία, παρατίθεται το διάγραμμα ποσοστιαίας κατανομής των σκοπών μετακίνησης ανά εισοδηματική κατηγορία.

Στην κατηγορία του μηδενικού εισοδήματος, η πλειοψηφία των μετακινήσεων γίνονται με σκοπό την αναψυχή, σε ποσοστό 29% και ακολουθεί η επιστροφή στην κατοικία, με ποσοστό 26% των μετακινήσεων. Σε αυτή την κατηγορία οι μετακινήσεις για σκοπούς εκπαίδευσης αφορούν το 19% του δείγματος. Με μικρότερα ποσοστά εμφανίζονται οι μετακινήσεις προς την εργασία (9%), οι μετακινήσεις για αγορές (8%) και οι μετακινήσεις προς υπηρεσίες (1%). Η κατηγορία μετακινήσεων «άλλο» σημείωσε ποσοστό 7% στην ομάδα του μηδενικού εισοδήματος.

Όσον αφορά την εισοδηματική κατηγορία των 750€ ή λιγότερων, επικρατούν οι μετακινήσεις με σκοπό την εργασία (32%) και μετά οι μετακινήσεις επιστροφής στην κατοικία (26%). Οι μετακινήσεις για αναψυχή αγγίζουν το 18%, ενώ οι μετακινήσεις για εκπαίδευση και αγορές σημειώνουν ποσοστό 7% η κάθε μία. Με ποσοστό μικρότερο του 1% εμφανίζονται οι μετακινήσεις προς διάφορες υπηρεσίες. Το 9% των μετακινήσεων των ατόμων που ανήκουν σε αυτή την εισοδηματική κατηγορία διεξάγονται για «άλλους» λόγους.

Ποσοστά Σκοπών Μετακίνησης ανά Κατηγορία Εισοδήματος



Διάγραμμα 4.24 Ποσοστιαία κατανομή σκοπών μετακινήσεως ανά κατηγορία εισοδήματος

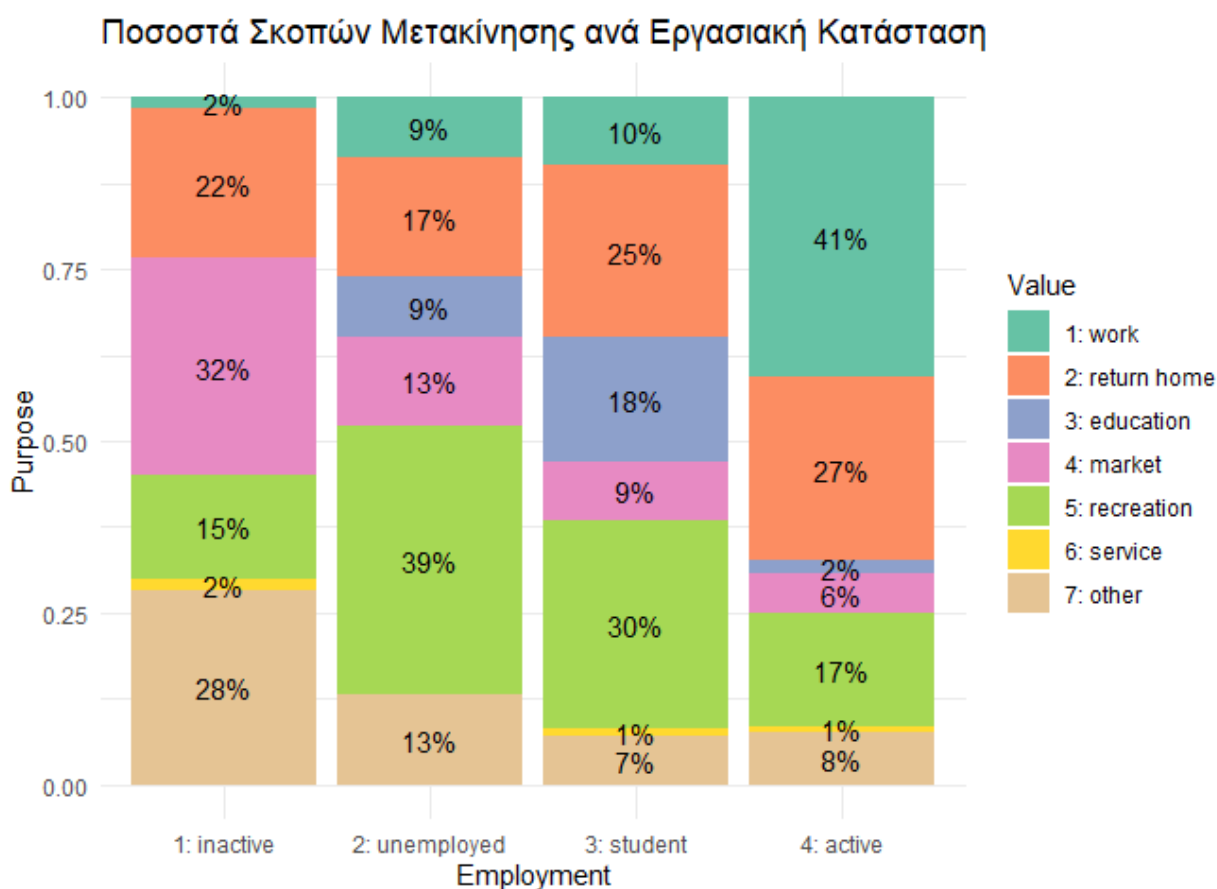
Έπειτα, γίνεται η ποσοστιαία κατανομή των σκοπών μετακίνησης για τις μετακινήσεις των ατόμων που δήλωσαν εισόδημα 750€ έως 1500€. Σε αυτή την κατηγορία πρωτοστατούν οι μετακινήσεις με σκοπό την εργασία με ποσοστό 37%. Οι μετακινήσεις για επιστροφή στο σπίτι σημειώνουν ποσοστό 26%. Ακολουθούν οι μετακινήσεις για λόγους αναψυχής με ποσοστό 17% και οι μετακινήσεις για αγορές με ποσοστό 9%. Μικρά ποσοστά εμφανίζουν οι σκοποί: «άλλο» καταλαμβάνοντας το 8% των μετακινήσεων, «εκπαίδευση» καταλαμβάνοντας το 2% και «υπηρεσίες» με ποσοστό 1%.

Στην εισοδηματική κατηγορία 1500-2500€ σημειώνουν μέγιστο οι μετακινήσεις με σκοπό την εργασία, σημειώνοντας ποσοστό ίσο με 40%. Το 29% των μετακινήσεων της κατηγορίας αφορά τη μετάβαση στην κατοικία. Οι μετακινήσεις για αγγίζουν το 16% και οι μετακινήσεις για «άλλους» λόγους αγγίζουν το 10%. Οι αγορές συγκεντρώνουν το 4% των μετακινήσεων της ομάδας, ενώ οι υπηρεσίες συγκεντρώνουν ποσοστό μικρότερο του 1%. Να σημειωθεί ότι, από αυτή την εισοδηματική κατηγορία απουσιάζουν μετακινήσεις για εκπαιδευτικούς σκοπούς.

Πέμπτη και τελευταία εισοδηματική κατηγορία είναι αυτή των εισοδημάτων άνω των 2500€. Σε αυτή την κατηγορία οι μετακινήσεις με σκοπό την εργασία καταλαμβάνουν το 35% του συνόλου, ενώ οι

μετακινήσεις για επιστροφή στην κατοικία το 16%. Μέγιστο μεταξύ των κατηγοριών σημειώνουν εδώ οι μετακινήσεις για λόγους αναψυχής συγκεντρώνοντας ποσοστό 32%. Μέγιστο σημειώνουν και οι μετακινήσεις για αγορές με ποσοστό 10%. Με 3% εμφανίζονται οι μετακινήσεις για εκπαίδευση όπως επίσης με 3% εμφανίζονται οι μετακινήσεις για «άλλους» λόγους. Παρατηρείται ότι, οι μετακινήσεις προς υπηρεσίες απουσιάζουν από αυτή την εισοδηματική κατηγορία.

Η εισοδηματική κατηγορία των συμμετεχόντων σχετίζεται άμεσα και με την εργασιακή τους κατάσταση. Στη συνέχεια, παρουσιάζεται το διάγραμμα ποσοστιαίας κατανομής των σκοπών μετακίνησης των συμμετεχόντων αναλόγως της εργασιακής κατάστασης. Οι κατηγορίες που χωρίστηκαν οι ερωτηθέντες είναι: 1: άεργος, 2: άνεργος, 3: μαθητής/φοιτητής, 4: εργαζόμενος.



Διάγραμμα 4.25 Ποσοστιαία κατανομή σκοπών μετακίνησης ανά εργασιακή κατάσταση

Στην κατηγορία των αέργων παρατηρείται μέγιστο στις μετακινήσεις για αγορές, με ποσοστό 32%. Ακολουθούν οι μετακινήσεις που χαρακτηρίστηκαν ως «άλλο», με ποσοστό 28%. Με 22% εμφανίζονται οι μετακινήσεις που αφορούν την επιστροφή στην κατοικία και με 15% οι μετακινήσεις για λόγους αναψυχής. Με μικρά ποσοστά εμφανίζονται οι μετακινήσεις για πρόσβαση σε υπηρεσίες (2%).

Η δεύτερη κατηγορία αφορά τους ανέργους. Σε αυτή την κατηγορία η πλειοψηφία των μετακινήσεων γίνεται για αναψυχή με ποσοστό 39%. Οι μετακινήσεις που αφορούν την επιστροφή στο σπίτι αγγίζουν το 17%, αυτές που αφορούν την αγορά είναι το 13% του δείγματος και όσες αφορούν την εκπαίδευση είναι στο 9%. Οι μετακινήσεις της ομάδας «άλλο» κατέχουν ποσοστό 13% των συνολικών. Να σημειωθεί ότι, σε αυτή την κατηγορία δεν υπάρχουν μετακινήσεις προς υπηρεσίες.

Στη συνέχεια, οι μετακινήσεις της κατηγορίας των μαθητών/φοιτητών είναι κατά 30% για αναψυχή και κατά 25% για επιστροφή στην κατοικία. Ακολουθούν οι μετακινήσεις για εκπαιδευτικούς σκοπούς σε ποσοστό 18%. Αυτή η ομάδα συγκεντρώνει ποσοστό 10% των μετακινήσεων στο σκοπό «μετάβαση στην εργασία» και 9% των μετακινήσεων στο σκοπό «αγορές». Με 7% εμφανίζονται οι μετακινήσεις για «άλλους» λόγους και με 1% οι μετακινήσεις για μετάβαση σε υπηρεσίες.

Στην τελευταία κατηγορία, η οποία αφορά τα εργαζόμενα άτομα, παρατηρείται ότι, το 41% των μετακινήσεων διεξάγεται για μετάβαση στην εργασία. Επιπλέον, το 27% των μετακινήσεων αφορά την επιστροφή στην κατοικία και το 17% των μετακινήσεων αφορά την αναψυχή. Σε αυτή την κατηγορία, οι μετακινήσεις προς τις αγορές αγγίζουν το 6%, οι μετακινήσεις για εκπαιδευτικούς σκοπούς αγγίζουν το 2% και οι μετακινήσεις προς υπηρεσίες μόνο το 1% του συνόλου. Ο σκοπός μετακίνησης «άλλο» συγκεντρώνει το 8% των μετακινήσεων της κατηγορίας.

5 ΕΠΕΞΕΡΓΑΣΙΑ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 Γενικά

Στο παρόν κεφάλαιο, θα παρουσιαστεί αναλυτικά, η διαδικασία με την οποία έγινε η επεξεργασία των δεδομένων της έρευνας, ώστε να έρθουν στην κατάλληλη μορφή για τη διεξαγωγή της συσταδοποίησης. Στη συνέχεια, θα αναλυθεί ο κώδικας που συντάχθηκε με χρήση της γλώσσας προγραμματισμού «R» και θα παρουσιαστούν στιγμιότυπα από το περιβάλλον της R-Studio. Επίσης, θα περιγραφεί διεξοδικά, η διαδικασία της συσταδοποίησης και του κώδικα που την αφορά. Τέλος, θα παρουσιαστούν τα αποτελέσματα της διαδικασίας της συσταδοποίησης μαζί με τα σχετικά διαγράμματα που προκύπτουν για την επίτευξη της βέλτιστης ερμηνείας τους.

5.2 Ανάλυση Δεδομένων

Τα δεδομένα που χρησιμοποιήθηκαν για την εκτέλεση της συσταδοποίησης ήταν αυτά των αποτελεσμάτων των ερωτηματολογίων και των αποστάσεων μεταξύ των ζωνών.

Απαραίτητη προϋπόθεση για τη διεξαγωγή της συσταδοποίησης ήταν ο υπολογισμός των μέσων αποστάσεων μεταξύ των ζωνών. Η υλοποίηση των ζωνών είχε γίνει ήδη για τις ανάγκες της διεξαγωγής της έρευνας, όπως αναφέρεται και στα προηγούμενα κεφάλαια. Έπειτα, στο περιβάλλον του λογισμικού QGIS και μέσω του αρχείου οπτικοποίησης των ζωνών, δημιουργήθηκε το κεντροειδές της εκάστοτε ζώνης. Στη συνέχεια, μέσω μίας σειράς ενεργειών, εξάχθηκε ο πίνακας των αποστάσεων των κεντροειδών των ζωνών. Ο πίνακας αποστάσεων περιείχε την απόσταση των κέντρων δύο ζωνών, για όλους τους συνδυασμούς από και προς κάθε μία από τις 36 ζώνες που δημιουργήθηκαν. Επίσης, ο πίνακας των αποστάσεων αποτελείται από τρεις μεταβλητές: «Ζώνη Προέλευσης», «Ζώνη Προορισμού» και «Απόσταση».

Εν συνεχεία, τέθηκε υπό επεξεργασία το σύνολο των δεδομένων που προέκυψαν από τα ερωτηματολόγια. Τα δεδομένα αυτά περιέχουν τον αριθμό ταυτοποίησης (pid), το φύλο (gender), την ηλικία (age), το επίπεδο εκπαίδευσης (education), την εργασιακή κατάσταση (employment), την εισοδηματική κατηγορία (income), τη ζώνη κατοικίας των συμμετεχόντων (home) και εάν κατέχουν ιδιωτικό αυτοκίνητο (car_own). Επίσης, περιλαμβάνουν τον προορισμό (dest), το σκοπό (purp), το μέσο (mode) και την ώρα έναρξης (time) έως και πέντε μετακινήσεων.

Αυτό το σύνολο δεδομένων, όπως αναλύθηκε παραπάνω, εισάχθηκε στο περιβάλλον της R-studio, με χρήση κατάλληλων εντολών. Εν συνεχεία, το σύνολο των δεδομένων μετασχηματίστηκε ακολουθώντας μια σειρά ενεργειών. Αρχικά, οι διαδοχικές μετακινήσεις του κάθε συμμετέχοντα χωρίστηκαν σε ξεχωριστές παρατηρήσεις. Αυτό είχε σαν αποτέλεσμα να δημιουργηθεί ένας μεγαλύτερος πίνακας δεδομένων, όπου υπήρχαν τα κοινωνικοοικονομικά δεδομένα του κάθε συμμετέχοντα σε πέντε γραμμές, μία για την κάθε διαδοχική μετακίνηση, μαζί με τις πληροφορίες της εκάστοτε μετακίνησης. Φυσικά, για τους συμμετέχοντες που δεν εκτέλεσαν συνολικά 5 διαδοχικές μετακινήσεις, προστέθηκαν τόσες γραμμές όσες οι μετακινήσεις τους (πχ. 1,2,3 ή 4). Ταυτόχρονα,

προστέθηκε μία μεταβλητή που αφορά τον αριθμό ταυτοποίησης των μετακινήσεων. Αυτή η μεταβλητή πήρε τιμές από το 1 και με αύξουσα σειρά έφτασε έως το 1347 και χαρακτήριζε όλες τις μετακινήσεις του δείγματος. Συμπερασματικά, προκύπτει ότι, οι συνολικές μετακινήσεις του δείγματος ήταν 1347.

Έπειτα, χρησιμοποιήθηκε ο πίνακας των αποστάσεων μεταξύ των ζωνών και προστέθηκε η μεταβλητή της απόστασης της κάθε μετακίνησης στο σύνολο των 1347 μετακινήσεων. Με αυτόν τον τρόπο η κάθε μετακίνηση κατέληξε να χαρακτηρίζεται από το μέσο, το σκοπό, την απόσταση και την ώρα έναρξης της.

Για την επίτευξη της συσταδοποίησης, οι μεταβλητές του μέσου και του σκοπού μετακίνησης μετασχηματίστηκαν από ονομαστικές (nominal) σε δυαδικές (binary). Για αυτόν το λόγο η μεταβλητή μέσο (mode) μετατράπηκε σε οκτώ διαφορετικές δυαδικές μεταβλητές, μία για κάθε μέσο. Συγκεκριμένα δημιουργήθηκαν οι μεταβλητές: αυτοκίνητο (car), ταξί (taxi), λεωφορείο (bus), τραίνο (train), μοτοσυκλέτα (motorcycle), ποδήλατο (bicycle), περπάτημα (walk) και e-scooter. Αυτές οι δυαδικές μεταβλητές έπαιρναν την τιμή 1 στο πεδίο του μέσου με το οποίο πραγματοποιούνταν η συγκεκριμένη μετακίνηση και την τιμή 0 σε όλα τα υπόλοιπα πεδία.

Η ίδια διαδικασία επαναλήφθηκε για τη μεταβλητή του σκοπού μετακίνησης. Συγκεκριμένα, ο σκοπός μετακίνησης (purp) μετατράπηκε σε επτά διαφορετικές δυαδικές (binary) μεταβλητές. Αναλυτικά, δημιουργήθηκαν οι μεταβλητές: εργασία (work), επιστροφή στο σπίτι (return home), εκπαίδευση (education), αγορές (market), αναψυχή (recreation), υπηρεσίες (service), άλλο (other). Αυτές οι δυαδικές μεταβλητές έπαιρναν την τιμή 1 στο πεδίο του μέσου με το οποίο πραγματοποιούνταν η συγκεκριμένη μετακίνηση και την τιμή 0 σε όλα τα υπόλοιπα πεδία.

Αφού οι μεταβλητές ήρθαν στη μορφή που απαιτείται, επιλέχθηκαν όσες από αυτές θα συμμετέχουν στη συσταδοποίηση. Αναλυτικότερα, επιλέχθηκαν οι μεταβλητές: απόσταση μετακίνησης (dist), χρόνος έναρξης μετακίνησης (time), αυτοκίνητο (car), ταξί (taxi), λεωφορείο (bus), τραίνο (train), μοτοσυκλέτα (motorcycle), ποδήλατο (bicycle), περπάτημα (walk), ηλεκτρονικό πατίνι (escooter), εργασία (work), επιστροφή στο σπίτι (return home), εκπαίδευση (education), αγορές (market), αναψυχή (recreation), υπηρεσίες (service), άλλο (other).

Η μεταβλητή της απόστασης των μετακινήσεων (dist) περιείχε την απόσταση της εκάστοτε μετακίνησης σε μέτρα με τρία δεκαδικά ψηφία. Η ίδια δημιουργήθηκε βάσει του πίνακα αποστάσεων όπως αναφέρεται και παραπάνω. Όμως ο πίνακας αποστάσεων περιέχει όλους τους δυνατούς συνδυασμούς μετακινήσεων εκτός από τις αποστάσεις των ενδοζωνικών μετακινήσεων, δηλαδή τις μετακινήσεις εντός τις ίδιας ζώνης. Αυτό συμβαίνει γιατί οι αποστάσεις μετρήθηκαν μέσω των κεντροειδών των ζωνών. Για τη διεξαγωγή της έρευνας έγινε η παραδοχή ότι οι ενδοζωνικές μετακινήσεις, θα σημειώνουν απόσταση «1000» μέτρων. Έτσι, τα πεδία της μεταβλητής της

απόστασης, τα οποία ήταν κενά επειδή αφορούσαν ενδοζωνικές μετακινήσεις, συμπληρώθηκαν με την τιμή «1000» μέτρα.

Μετά από τις ενέργειες που αναλύθηκαν σε αυτό το κεφάλαιο, το σύνολο των δεδομένων μετασχηματίστηκε στην κατάλληλη μορφή προκειμένου να πραγματοποιηθεί η διαδικασία της συσταδοποίησης.

5.3 Διαδικασία Συσταδοποίησης

Σε αυτό το υπό-κεφάλαιο θα παρουσιαστεί η συσταδοποίηση, όπως αυτή αναλύθηκε διεξοδικά στο κεφάλαιο της μεθοδολογίας, μαζί με αποσπάσματα του κώδικα που συντάχθηκε στο περιβάλλον της R-Studio.

Αρχικά, επειδή οι μεταβλητές μεταξύ τους βρίσκονται σε διαφορετική κλίμακα, ακολουθήθηκε μία διαδικασία, όπως θα παρουσιαστεί στη συνέχεια, προκειμένου να προκύψουν οι τιμές των μεταβλητών στην ίδια κλίμακα. Η μεταβλητή της απόστασης είναι της τάξεως των δεκάδων χιλιάδων και η μεταβλητή της ώρας έναρξης είναι της τάξεως των δεκάδων. Ταυτόχρονα, οι υπόλοιπες μεταβλητές του συνόλου των δεδομένων είναι δυαδικές δηλαδή παίρνουν τιμές 0 και 1. Αυτό δημιουργεί πρόβλημα στη συσταδοποίηση καθώς συνεπάγεται διαφορετική βαρύτητα μεταβλητών. Για να αποφευχθεί αυτό το πρόβλημα εκτελέστηκαν οι παρακάτω εντολές, προκειμένου να τοποθετηθούν οι τιμές των μεταβλητών της απόστασης και της ώρας έναρξης του ταξιδιού μεταξύ των τιμών 0 και 1.

```
1. min_vals <- sapply(cluster_data[, 1:2], min)
2. max_vals <- sapply(cluster_data[, 1:2], max)
3. cluster_data[, 1:2] <- mapply(function(x, min, max) (x - min) / (max - min), cluster_data[, 1:2],
min_vals, max_vals)
```

Εικόνα 5.1 Απόσπασμα κώδικα αποκλιμάκωσης δεδομένων

Εφόσον πραγματοποιήθηκε με επιτυχία η αποκλιμάκωση των μεταβλητών αυτών, ακολούθησε η συγγραφή της συνάρτησης που εκτελεί την ομαδοποίηση μέσω του αλγορίθμου k-μέσων. Αυτή η συνάρτηση λάμβανε ως ορίσματα το σύνολο των δεδομένων, τον επιθυμητό αριθμό των κέντρων και το μέγιστο αριθμό επαναλήψεων. Στη συνέχεια, εκτελεί τον αλγόριθμο k-μέσων ακολουθώντας τα εξής βήματα: αρχικά αναθέτει τα σημεία στο κοντινότερο κέντρο, μετά ενημερώνει τα κέντρα βάσει των νέων ομάδων και επαναλαμβάνει αυτά τα βήματα μέχρι τα κέντρα να συγκλίνουν ή μέχρι να ξεπεραστεί ο μέγιστος αριθμός επαναλήψεων. Τέλος, η συνάρτηση επιστρέφει μια λίστα που περιέχει τα κέντρα και τις ομάδες των δεδομένων.

```
1. k_means <- function(data, k, max_iter = 100) {
2.   n <- nrow(data)
3.   data_matrix <- as.matrix(data) # Convert data to matrix
4.   centroids <- data_matrix[sample(1:n, k), ]
5.   clusters <- numeric(n)
```

```

6.  distances <- matrix(NA, n, k)
7.
8.  for (iter in 1:max_iter) {
9.    # Step 2: Assign points to the nearest centroid
10.   for (i in 1:k) {
11.     distances[, i] <- rowSums((sweep(data_matrix, 2, centroids[i, ], "-"))^2)
12.   }
13.   clusters <- apply(distances, 1, which.min)
14.
15.   # Step 3: Update centroids
16.   old_centroids <- centroids
17.   for (i in 1:k) {
18.     centroids[i, ] <- colMeans(data_matrix[clusters == i, ], na.rm = TRUE)
19.   }
20.
21.   # Check for convergence (if centroids don't change)
22.   if (all(old_centroids == centroids)) {
23.     break
24.   }
25. }
26.
27. return(list(centroids = centroids, clusters = clusters))
28. }

```

Εικόνα 5.2 Απόσπασμα κώδικα συνάρτησης k-means

Η παραπάνω συνάρτηση θα κληθεί στη συνέχεια προκειμένου να πραγματοποιηθεί η φασματική συσταδοποίηση με το κατάλληλο σύνολο δεδομένων.

Στη συνέχεια, χρησιμοποιείται η συνάρτηση `set.seed()` για να ορίσει το `seed`, δηλαδή τον αριθμό έναρξης για τη γεννήτρια τυχαίων αριθμών. Στην προκειμένη περίπτωση επιλέγεται τυχαία το 42 ως αριθμός έναρξης της γεννήτριας τυχαίων αριθμών. Έπειτα, υπολογίζονται οι αποστάσεις μεταξύ όλων των ζευγών των σημείων των δεδομένων μέσω της απόστασης Manhattan και εκχωρεί τα αποτελέσματα σε μια μήτρα. Το αποτέλεσμα των υπολογισμών είναι ένας πίνακας στον οποίο κάθε κελί αναπαριστά την απόσταση ανάμεσα σε ένα ζευγάρι σημείων.

Όπως φαίνεται παρακάτω, ορίζεται η στατιστική παράμετρος «σίγμα», η οποία χρησιμοποιείται στην κανονικοποίηση των αποστάσεων. Η παράμετρος σ ρυθμίζει την ευαισθησία της μέτρησης ομοιότητας στις αποστάσεις. Η ίδια καταφέρνει να μετασχηματίζει την απόσταση ως τη διάμεση των προηγουμένως υπολογισμένων αποστάσεων των ζευγών. Εφόσον οριστεί το «σίγμα» δημιουργείται ένας πίνακας ομοιότητας, όπου οι αποστάσεις μετατρέπονται σε τιμές ομοιότητας μεταξύ των σημείων, χρησιμοποιώντας την αρνητική εκθετική συνάρτηση των κανονικοποιημένων αποστάσεων. Σκοπός είναι να μετατρέψει τις αποστάσεις σε μια κλίμακα από 0 έως 1, όπου οι μικρότερες αποστάσεις (δηλαδή, τα πιο κοντινά σημεία) έχουν υψηλότερες τιμές ομοιότητας.

```

1. # Set a seed for reproducibility
2. set.seed(42)
3.
4. # Calculate pairwise Manhattan distances
5. pairwise_distances <- as.matrix(dist(cluster_data, method = "manhattan"))
6.
7. # Set sigma for the similarity matrix, you can adjust this value

```

```
8. sigma <- median(pairwise_distances)
9.
10. # Create the similarity matrix using the Manhattan distance
11. similarity_matrix <- exp(-pairwise_distances / sigma)
12.
```

Εικόνα 5.3 Απόσπασμα κώδικα δημιουργίας μήτρας ομοιότητας

Ύστερα, δημιουργείται ο πίνακας βαθμών, ο οποίος είναι διαγώνιος και περιέχει το άθροισμα των τιμών ομοιότητας για κάθε σημείο. Στη συνέχεια, υπολογίζεται η Λαπλασιανή μήτρα με αφαίρεση της μήτρας ομοιότητας από τον πίνακα βαθμών. Η Λαπλασιανή μήτρα χρησιμοποιείται στη φασματική συσταδοποίηση για να αναδειχθεί η δομή των δεδομένων βασιζόμενη στις σχέσεις ομοιότητας.

Έπειτα, υπολογίζονται οι ιδιοτιμές και τα ιδιοδιανύσματα της Λαπλασιανής μήτρας. Οι ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα βοηθούν στην ανίχνευση των συστάδων μέσα στα δεδομένα. Έτσι, μέσω των ιδιοτιμών δίνεται μία ένδειξη για τον αριθμό των σημαντικών συστάδων, ενώ τα ιδιοδιανύσματα χρησιμοποιούνται για να προβάλουν τα δεδομένα στο χώρο όπου οι συστάδες είναι ευδιάκριτες. Η παράμετρος `symmetric = TRUE` χρησιμοποιείται για να βελτιστοποιήσει την απόδοση, επιτρέποντας στη συνάρτηση να εκμεταλλευτεί τη συμμετρία της μήτρας.

```
1. # Construct the Laplacian matrix
2. degree_matrix <- diag(colSums(similarity_matrix))
3. laplacian_matrix <- degree_matrix - similarity_matrix
4.
5. # Compute Eigenvalues and Eigenvectors
6. eigen_data <- eigen(laplacian_matrix, symmetric = TRUE) # Ensure symmetric is TRUE for performance
7. eigenvalues <- eigen_data$values
8. eigenvectors <- eigen_data$vectors
```

Εικόνα 5.4 Απόσπασμα κώδικα υπολογισμού ιδιοδιανυσμάτων

Τελικά, με τον υπολογισμό των απαραίτητων ιδιοδιανυσμάτων, ορίστηκε και ο επιθυμητός αριθμός των συστάδων. Συγκεκριμένα, ο αριθμός που επιλέχθηκε ήταν έξι ομάδες δεδομένων. Ύστερα, από το σύνολο ιδιοδιανυσμάτων, επιλέχθηκαν τα πρώτα k (σε αυτή την περίπτωση, 6) για χρήση στη διαδικασία της ομαδοποίησης. Τα πρώτα έξι ιδιοδιανύσματα αντιστοιχούν στις πιο μικρές (αλλά μη μηδενικές) ιδιοτιμές του πίνακα που μελετάται. Τα έξι αυτά πρώτα ιδιοδιανύσματα τοποθετήθηκαν σε έναν πίνακα που ονομάστηκε « V ».

Ακολούθησε η κανονικοποίηση των γραμμών του V , όπου για κάθε γραμμή του πίνακα, υπολογίστηκε η κανονικοποιημένη έκδοση διαιρώντας κάθε στοιχείο της γραμμής με την τετραγωνική ρίζα του αθροίσματος των τετραγώνων των στοιχείων της γραμμής. Έτσι εξασφαλίστηκε ότι κάθε γραμμή του V είχε μοναδιαίο μήκος.

Εν τέλει, έγινε η συσταδοποίηση των γραμμών του πίνακα με τα έξι μικρότερα ιδιοδιανύσματα, με χρήση του αλγορίθμου k -μέσων, όπως αυτός ορίστηκε προηγουμένως. Αυτή η συνάρτηση χρησιμοποιεί την αλγοριθμική μέθοδο k -μέσων για να ομαδοποιήσει τις κανονικοποιημένες γραμμές του πίνακα ιδιοδιανυσμάτων (V) σε k ομάδες, εδώ συγκεκριμένα σε 6 ομάδες, με 25 εκκινήσεις για να

βελτιστοποιήσει την εύρεση των καλύτερων κέντρων των ομάδων. Τα αποτελέσματα της φασματικής συσταδοποίησης εκχωρήθηκαν σε μία λίστα «spectral clusters».

```
1. # Determine the number of clusters k
2. k <- 6
3.
4. # Select the first k eigenvectors
6. V <- eigenvectors[, 1:k]
7.
8. # Normalize the rows of V
9. V_normalized <- t(apply(V, 1, function(row) row / sqrt(sum(row^2))))
10.
11. # Cluster the rows of V using K-means
12. spectral_clusters <- kmeans(V_normalized, centers = k, nstart = 25)
```

Εικόνα 5.5 Απόσπασμα κώδικα εφαρμογής συσταδοποίησης

Τα αποτελέσματα όλης της διαδικασίας της φασματικής συσταδοποίησης, όπως αυτή αναλύθηκε στο παρόν κεφάλαιο, θα παρουσιαστούν στη συνέχεια και θα αναλυθούν διεξοδικά με χρήση διάφορων οπτικοποιήσεων-διαγραμμάτων.

5.4 Αποτελέσματα Συσταδοποίησης

Σε αυτό το σημείο θα παρουσιαστούν τα αποτελέσματα της φασματικής συσταδοποίησης, θα περιγραφούν στατιστικά και θα αναγνωριστούν οι κατηγορίες των μετακινήσεων που δημιουργήθηκαν βάσει των δεδομένων που συλλέχθηκαν.

Αρχικά, στον πίνακα που ακολουθεί καταγράφεται ο αριθμός των σημείων ανά συστάδα.

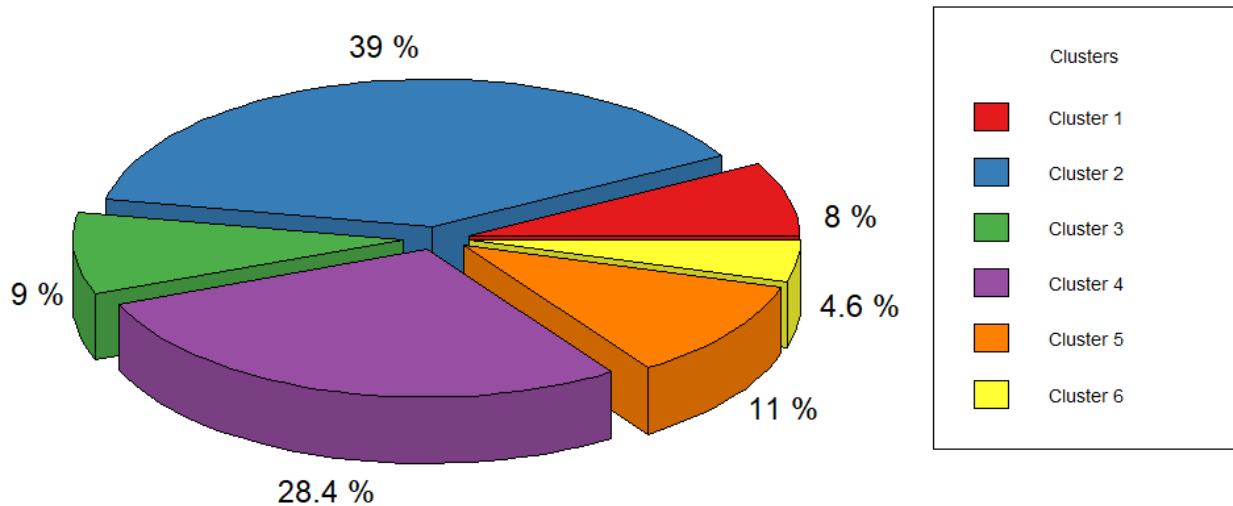
Πίνακας 5.1 Κατανομή πλήθους μετακινήσεων ανά συστάδα

Αριθμός Συστάδας	1	2	3	4	5	6	Σύνολο
Πλήθος Μετακινήσεων	108	525	121	383	148	62	1347

Στο διάγραμμα που ακολουθεί φαίνεται η ποσοστιαία κατανομή του πλήθους των παρατηρήσεων ανά συστάδα. Με κόκκινο χρώμα παρουσιάζεται η πρώτη συστάδα, που συγκεντρώνει το 8% των μετακινήσεων του δείγματος. Με μπλε παρουσιάζεται η δεύτερη συστάδα, με ποσοστό 39% των παρατηρήσεων. Η τρίτη συστάδα εμφανίζεται με πράσινο χρώμα και περιλαμβάνει το 9% των παρατηρήσεων του δείγματος. Στη συνέχεια, με μωβ απεικονίζεται το πλήθος της τέταρτης συστάδας, που συγκεντρώνει το 28,4% επί του συνόλου των παρατηρήσεων και με πορτοκαλί απεικονίζεται το ποσοστό της πέμπτης συστάδας που αγγίζει το 11%. Τέλος, με κίτρινο χρώμα παρουσιάζεται η έκτη συστάδα, στην οποία ανήκει το 4,6% του πλήθους των παρατηρήσεων. Σημειώνεται ότι, η συστάδα με τη μεγαλύτερη συγκέντρωση παρατηρήσεων είναι η δεύτερη, ενώ αυτή με τη μικρότερη συγκέντρωση είναι η έκτη. Η συστάδα με τη δεύτερη μεγαλύτερη συγκέντρωση είναι η τέταρτη, ενώ

οι υπόλοιπες τρεις συστάδες (πρώτη, τρίτη, πέμπτη) κυμαίνονται στο 8% με 11% επί του συνόλου των παρατηρήσεων.

Ποσοστιαία Κατανομή Συστάδων



Διάγραμμα 5.1 Ποσοστιαία κατανομή συστάδων

Εφόσον παρουσιάστηκε το πλήθος των παρατηρήσεων που απαρτίζουν τις συστάδες, ακολουθεί η ανάλυση της κεντρικής τάσης και της διασποράς για όλες τις μεταβλητές ανά συστάδα. Συγκεκριμένα, θα υπολογιστούν η μέση τιμή, η διάμεσος και η τυπική απόκλιση κάθε συστάδας σε σχέση με κάθε χαρακτηριστικό των δεδομένων.

Για τη δημιουργία των πινάκων της μέσης τιμής και της διαμέσου των μεταβλητών κάθε συστάδας, οι μεταβλητές «mode» και «ruipr», οι οποίες αφορούν το μέσο και το σκοπό μετακίνησης αντίστοιχα, υπέστησαν επεξεργασία. Αναλυτικότερα, από τις τιμές αυτών των μεταβλητών αφαιρέθηκε το λεκτικό και παρέμεινε το αριθμητικό τους μέρος προκειμένου να είναι ευκολότεροι οι υπολογισμοί των μεγεθών. Έτσι, για τη μεταβλητή του μέσου μετακίνησης (mode) ισχύει ότι: το 1 αντιστοιχεί στο αυτοκίνητο, το 2 αντιστοιχεί στο ταξί, το 3 αντιστοιχεί στο λεωφορείο, το 4 αντιστοιχεί στα Μέσα Σταθερής Τροχιάς, το 5 αντιστοιχεί στη μοτοσυκλέτα, το 6 αντιστοιχεί στο ποδήλατο, το 7 αντιστοιχεί στο περπάτημα και το 8 αντιστοιχεί στο πατίνι. Για τη μεταβλητή του σκοπού μετακίνησης (ruipr) ισχύει ότι: το 1 αντιστοιχεί στη μετακίνηση για εργασία, το νούμερο 2 αντιστοιχεί στην επιστροφή στο σπίτι, το 3 αντιστοιχεί στη μετακίνηση για εκπαιδευτικούς λόγους, το 4 αντιστοιχεί στις αγορές, το 5 αντιστοιχεί στην αναψυχή, το 6 αντιστοιχεί στις υπηρεσίες και το 7 αντιστοιχεί στο «άλλο».

Πίνακας 5.2 Μέση τιμή μεταβλητών ανά συστάδα

Group	mode	purp	time	dist
1	4.916667	3.435185	19.398148	1687.428
2	2.965714	3.129524	14.739048	2696.678
3	3.165289	2.661157	16.000000	14040.563
4	1.885117	2.287206	12.778068	12949.350
5	3.466216	3.000000	8.885135	10604.203
6	3.016129	4.322581	9.693548	3674.860

Από τον παραπάνω πίνακα συμπεραίνεται η μέση τιμή κάθε μεταβλητής σε κάθε ομάδα. Όσον αφορά την ώρα έναρξης των ταξιδιών παρατηρείται ότι στην πρώτη συστάδα η μέση τιμή είναι 19,3, στη δεύτερη συστάδα η μέση τιμή είναι 14,7, στην τρίτη συστάδα η μέση τιμή είναι ίση με 16, στην τέταρτη συστάδα η τιμή είναι 12,7, στην πέμπτη συστάδα η τιμή είναι 8,8 και στην έκτη συστάδα η μέση τιμή είναι περίπου 9,6.

Για τη μεταβλητή της απόστασης των ταξιδιών παρατηρείται ότι η μέση τιμή: στην πρώτη συστάδα είναι 1687 μέτρα, στη δεύτερη ομάδα είναι 2696 μέτρα, στην τρίτη είναι 14040 μέτρα, στην τέταρτη 12949 μέτρα, στην πέμπτη συστάδα είναι 10604 μέτρα και στην έκτη συστάδα είναι 3674 μέτρα.

Όσον αφορά τις μεταβλητές του μέσου και του σκοπού μετακίνησης ο πίνακας των μέσων τιμών δεν διευκολύνει την ερμηνεία των ομάδων. Αυτό συμβαίνει γιατί το κάθε μέσο και ο κάθε σκοπός μετακίνησης αντιστοιχίζονται με ακέραιες μονάδες. Η μέση τιμή σε αυτή την περίπτωση είναι δεκαδική και η στρογγυλοποίηση της δεν θα βοηθούσε στην ερμηνεία των μέσων και των σκοπών μετακίνησης των συστάδων.

Πίνακας 5.3 Διάμεσος μεταβλητών ανά συστάδα

Group	mode	purp	time	dist
1	4.0	2	20	1000.000
2	2.0	2	15	2366.143

Group	mode	purp	time	dist
3	4.0	2	17	11806.600
4	1.0	1	13	10588.343
5	4.0	2	8	10422.414
6	2.5	4	11	1817.523

Στον παραπάνω πίνακα παρουσιάζεται η διάμεσος της κάθε μεταβλητής ανά συστάδα. Η διάμεσος προκύπτει ως ο αριθμός που βρίσκεται ακριβώς στη μέση, εφόσον τοποθετηθούν σε αύξουσα σειρά οι τιμές.

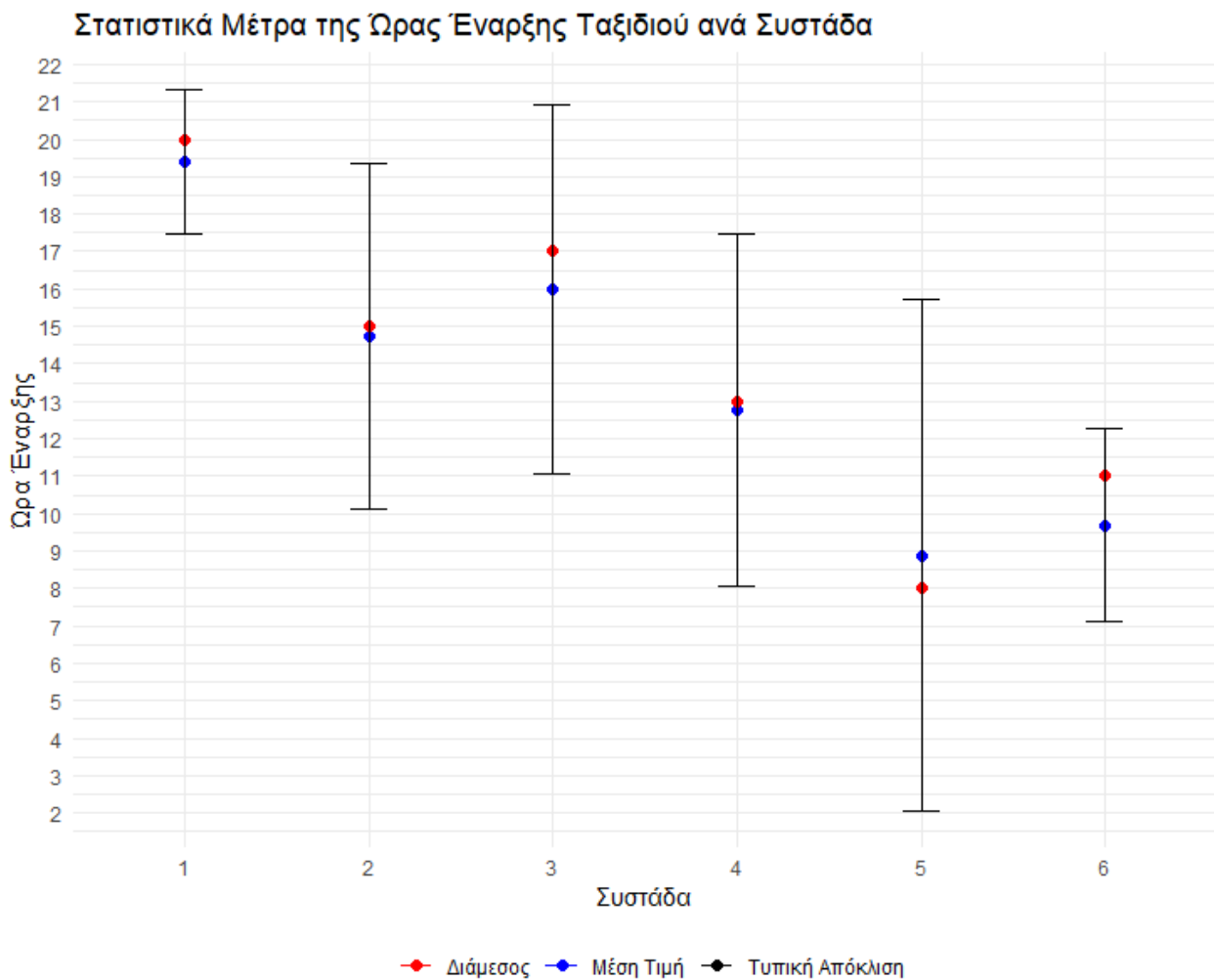
Παρατηρείται ότι, η διάμεσος της μεταβλητής του μέσου μετακίνησης ανά τις συστάδες είναι: στην πρώτη ομάδα στα μέσα σταθερής τροχιάς, στη δεύτερη στο ταξί, στην τρίτη στα μέσα σταθερής τροχιάς, στην τέταρτη στο αυτοκίνητο, στην πέμπτη στα μέσα σταθερής τροχιάς και στην έκτη μεταξύ του ταξί και του λεωφορείου.

Σε σχέση με τη μεταβλητή του σκοπού της μετακίνησης, η διάμεσος της πρώτης, της δεύτερης και της τρίτης συστάδας είναι στην επιστροφή στο σπίτι, της τέταρτης συστάδας είναι στη μετακίνηση για εργασία, στην πέμπτη συστάδα είναι στην επιστροφή στο σπίτι και στην έκτη συστάδα είναι στη μετακίνηση για αγορές.

Όσον αφορά τη μεταβλητή της ώρας έναρξης του ταξιδιού, η διάμεσος της πρώτης συστάδας είναι 20, η διάμεσος της δεύτερης συστάδας είναι 15, η διάμεσος της τρίτης συστάδας είναι 17, της τέταρτης είναι 13, της πέμπτης είναι 8 και της έκτης είναι 11.

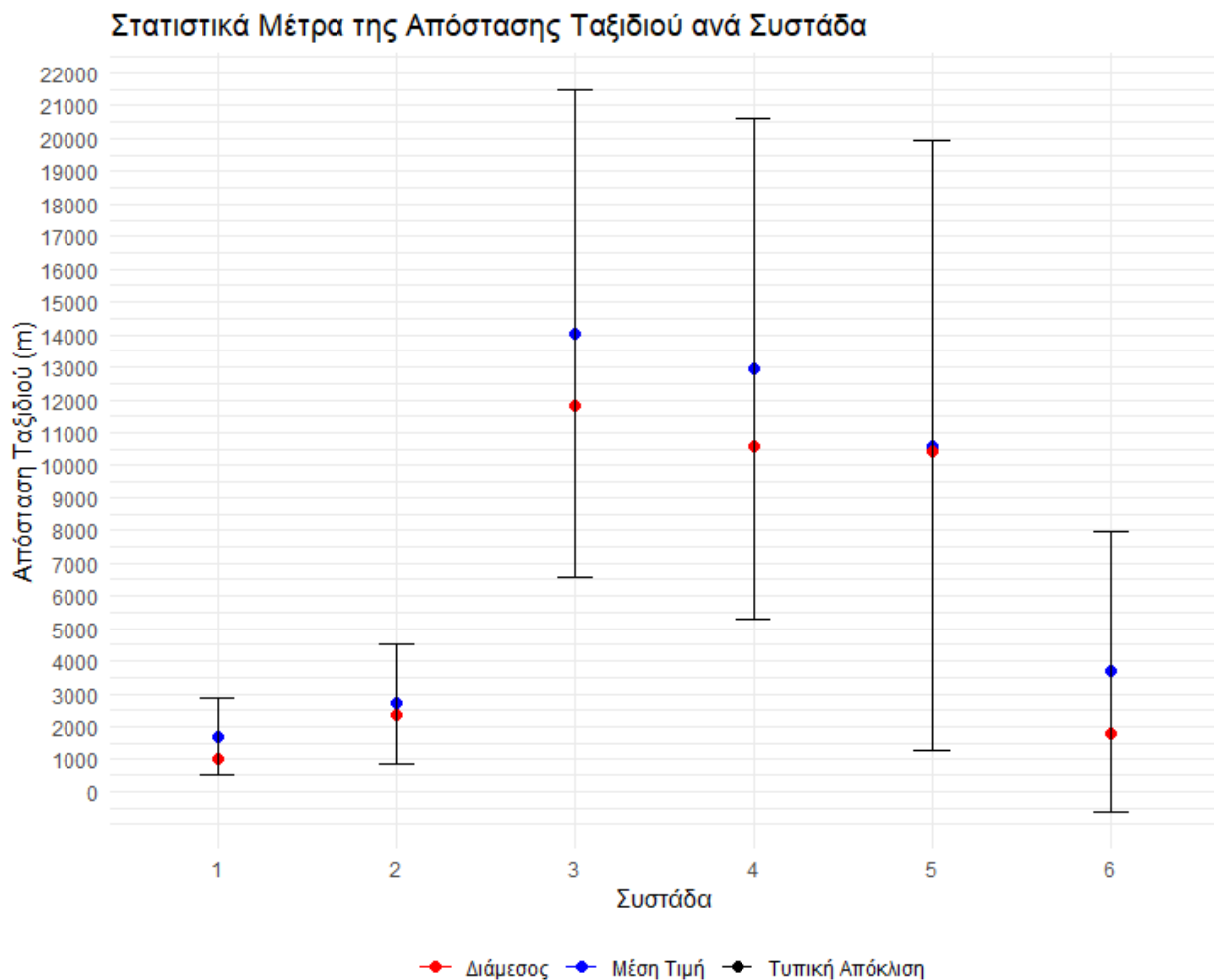
Για τη μεταβλητή της απόστασης του ταξιδιού προκύπτει ότι: η τιμή της διαμέσου στην πρώτη συστάδα είναι 1000 μέτρα, στη δεύτερη είναι 2366 μέτρα, στην τρίτη είναι 11806 μέτρα, στην τέταρτη είναι 10588 μέτρα, στην πέμπτη είναι 10422 μέτρα και στην έκτη είναι 1817 μέτρα.

Ακολουθούν τα διαγράμματα οπτικοποίησης της μέσης τιμής και των διαμέσων των μεταβλητών των ομάδων. Επίσης στα διαγράμματα που ακολουθούν φαίνεται η τυπική απόκλιση των μεταβλητών ανά συστάδα.



Διάγραμμα 5.2 Στατιστικά μέτρα της ώρας έναρξης ταξιδιού ανά συστάδα

Στο παραπάνω διάγραμμα, με τα στατιστικά μέτρα της ώρας έναρξης των μετακινήσεων, με κόκκινό χρώμα παρουσιάζεται η διάμεσος, με μπλε χρώμα η μέση τιμή και με μαύρο χρώμα η τυπική απόκλιση. Σημειώνεται ότι, η πρώτη συστάδα έχει τη μικρότερη τυπική απόκλιση, όσον αφορά την ώρα έναρξης των μετακινήσεων, αυτό συνεπάγεται ότι οι τιμές της αποκλίνουν από τη μέση τιμή λιγότερο σε σχέση με όλες τις υπόλοιπες συστάδες. Αμέσως μετά ακολουθεί η έκτη συστάδα, με την αμέσως μικρότερη τυπική απόκλιση, άρα και τις αμέσως λιγότερο διάσπαρτες παρατηρήσεις γύρω από τη μέση τιμή της, σε σχέση με την ώρα έναρξης των ταξιδιών. Μεγαλύτερες και σχεδόν ίσες τυπικές αποκλίσεις έχουν οι συστάδες 2,3 και 4. Η συστάδα με τη μεγαλύτερη τυπική απόκλιση είναι η πέμπτη, γεγονός που συνεπάγεται ότι οι παρατηρήσεις που την απαρτίζουν παρουσιάζουν μεγάλη διασπορά στη μεταβλητή του χρόνου έναρξης των μετακινήσεων.

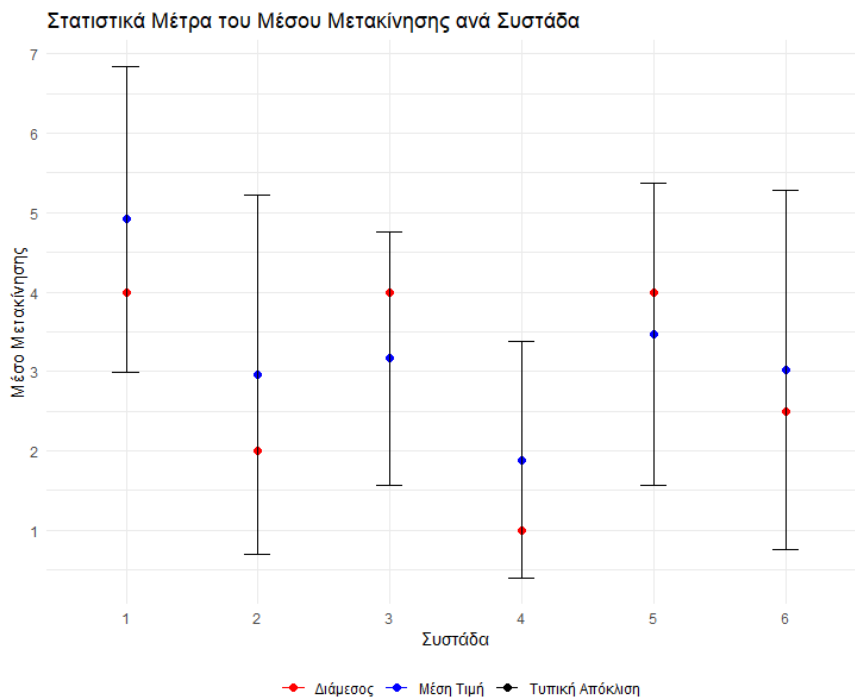


Διάγραμμα 5.3 Στατιστικά μέτρα της απόστασης ταξιδιού ανά συστάδα

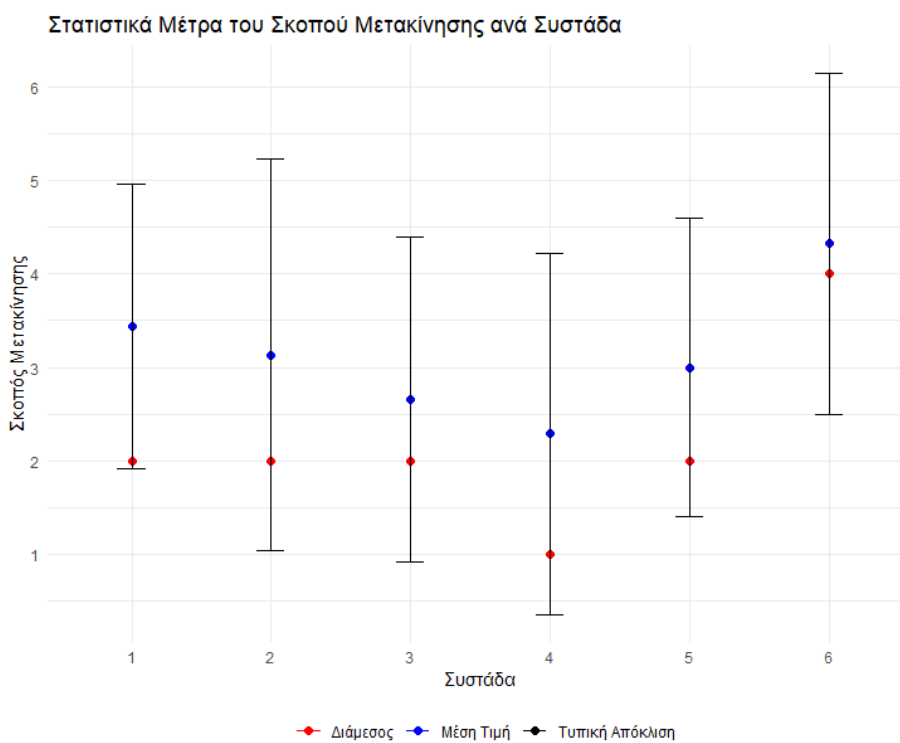
Βάσει του παραπάνω διαγράμματος συμπεραίνεται ότι, όσον αφορά τη μεταβλητή της απόστασης του ταξιδιού, η πρώτη συστάδα σημειώνει τη μικρότερη τυπική απόκλιση, αρά έχει και τη μικρότερη διασπορά. Ακολουθεί η δεύτερη συστάδα με την αμέσως μεγαλύτερη τυπική απόκλιση γύρω από τη μεταβλητή της απόστασης. Η απόσταση παρουσιάζει αρκετά μικρή τυπική απόκλιση και στην έκτη συστάδα. Τέλος, με μεγάλες και παρόμοιες τυπικές αποκλίσεις στη μεταβλητή της απόστασης του ταξιδιού, έρχονται οι συστάδες 3,4 και 5.

Τα διαγράμματα που παρουσιάζονται στη συνέχεια, αφορούν τα στατιστικά μέτρα των μεταβλητών του μέσου και του σκοπού μετακίνησης ανά συστάδα. Παρατηρώντας τα διαγράμματα, μπορεί να σχολιαστεί το μέγεθος της τυπικής απόκλισης των μέσων και των σκοπών μετακίνησης, που παρουσιάζει η κάθε συστάδα ξεχωριστά. Συγκεκριμένα, φαίνεται ότι, όλες οι συστάδες εμφανίζουν παρόμοιες τυπικές αποκλίσεις στις μεταβλητές του μέσου και του σκοπού μετακίνησης. Αυτό συνεπάγεται ότι όλες οι συστάδες παρουσιάζουν παρόμοια διασπορά γύρω από τη μέση τιμή της εκάστοτε μεταβλητής ανά συστάδα. Στα διαγράμματα που ακολουθούν, η διάμεσος θα μπορούσε να

χρησιμοποιηθεί σαν κέντρο βάρους των μεταβλητών. Αναλόγως τα συγκεντρώνονται μεγαλύτερα πλήθη παρατηρήσεων στην εκάστοτε μεταβλητή και συστάδα, μετατοπίζεται και η διάμεσος.



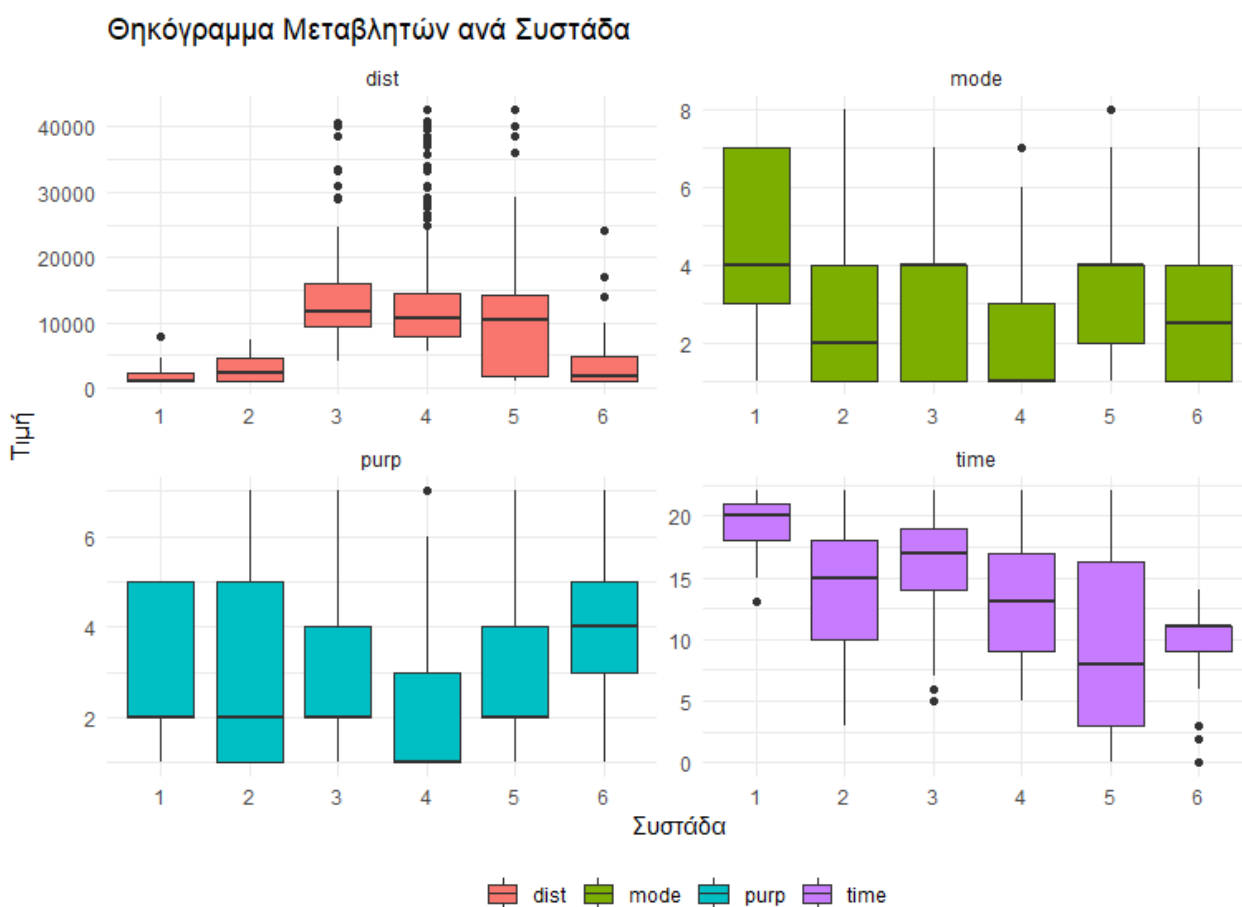
Διάγραμμα 5.4 Στατιστικά μέτρα του μέσου μετακίνησης ανά συστάδα



Διάγραμμα 5.5 Στατιστικά μέτρα του σκοπού μετακίνησης ανά συστάδα

Στη συνέχεια, παρουσιάζονται τα θηκογράμματα των μεταβλητών ανά συστάδα. Στα θηκογράμματα που ακολουθούν συμμετέχουν το πρώτο και το τρίτο τεταρτημόριο των τιμών των μεταβλητών. Κάθε υποδιάγραμμα αντιπροσωπεύει μια διαφορετική μεταβλητή (dist, mode, purp, time), και η κατανομή τους παρουσιάζεται σε έξι διαφορετικές συστάδες (1 έως 6). Το «dist» παρουσιάζει την κατανομή των αποστάσεων μετακίνησης. Οι μεσαίες τιμές (μεσαία γραμμή στο box plot) φαίνεται να διαφέρουν ανά συστάδα, με τη συστάδα 1 να έχει σημαντικά μικρότερες μεσαίες αποστάσεις σε σχέση με τις άλλες. Έπειτα για το «mode» που αφορά την κατηγορική μεταβλητή του μέσου μετακίνησης παρατηρείται ότι λαμβάνει τιμές από 1 έως 8, που αντιστοιχούν σε διαφορετικούς τρόπους μετακίνησης. Οι διαφορές μεταξύ των συστάδων είναι λιγότερο έντονες σε αυτήν τη μεταβλητή. Το «purp» είναι η μεταβλητή που περιγράφει το σκοπό της μετακίνησης με τιμές από 1 έως 7. Σημειώνεται ότι, οι κατανομές είναι πιο ομοιόμορφες μεταξύ των συστάδων σε σύγκριση με τη μεταβλητή "dist".

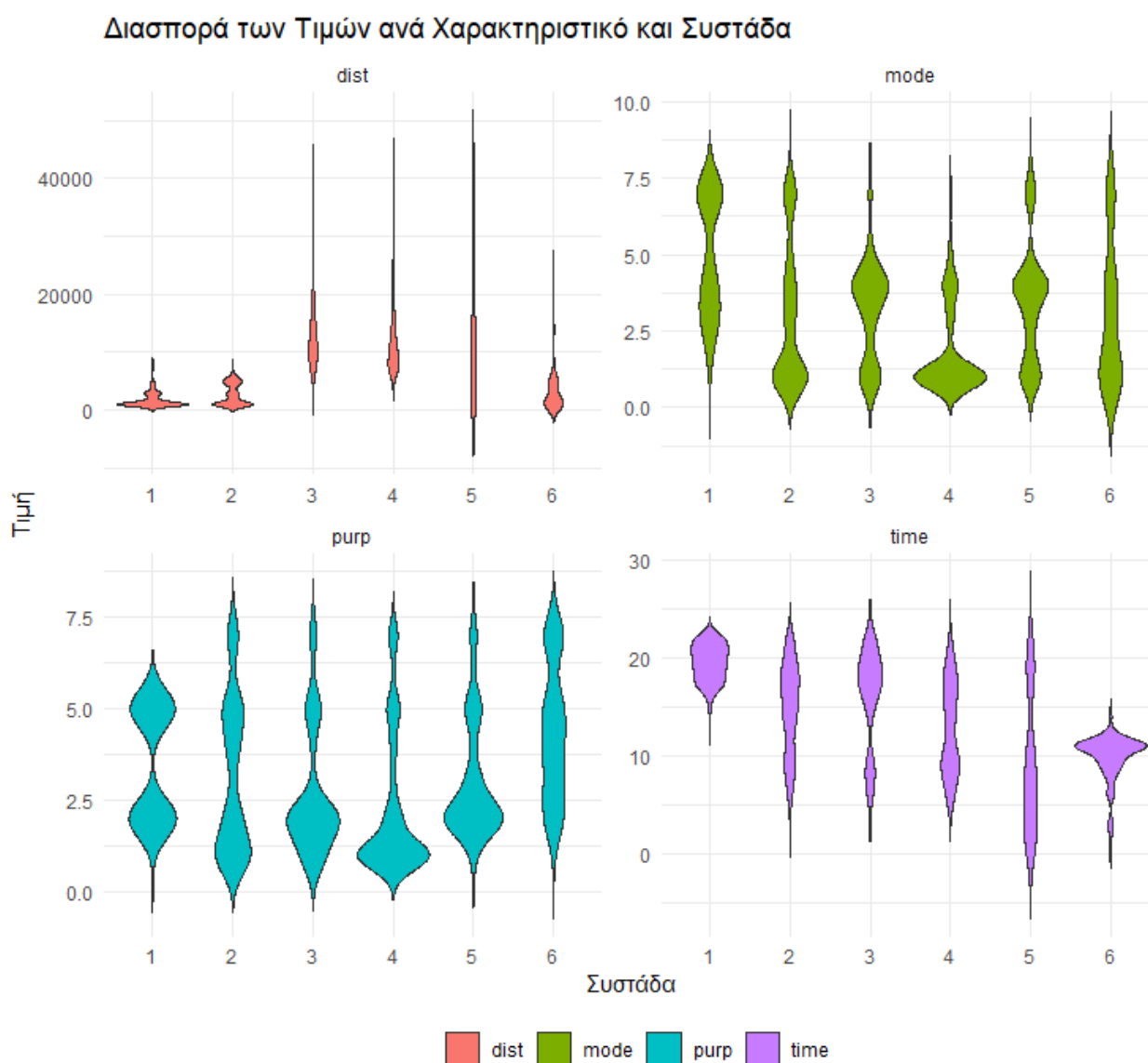
Το «time» δείχνει την κατανομή των ωρών έναρξης της μετακίνησης, σε μια κλίμακα από 0 έως 24, αναπαριστώντας το 24ωρο. Υπάρχουν σημαντικές διακυμάνσεις μεταξύ των συστάδων σε αυτή τη μεταβλητή, με κάποιες συστάδες να έχουν πιο συγκεντρωμένες ώρες μετακίνησης και άλλες πιο διασπαρμένες.



Διάγραμμα 5.6 Θηκογράμματα μεταβλητών ανά συστάδα

Στο επόμενο διάγραμμα απεικονίζεται η διασπορά των τιμών των μεταβλητών ανά ομάδα. Επιβεβαιώνεται ότι, οι τιμές των μεταβλητών του μέσου (mode) και του σκοπού (purp) των μετακινήσεων παρουσιάζουν μεγάλη διασπορά. Οι τιμές των μεταβλητών της απόστασης (dist) και της ώρας έναρξης (time) των μετακινήσεων παρουσιάζουν μικρότερη διασπορά.

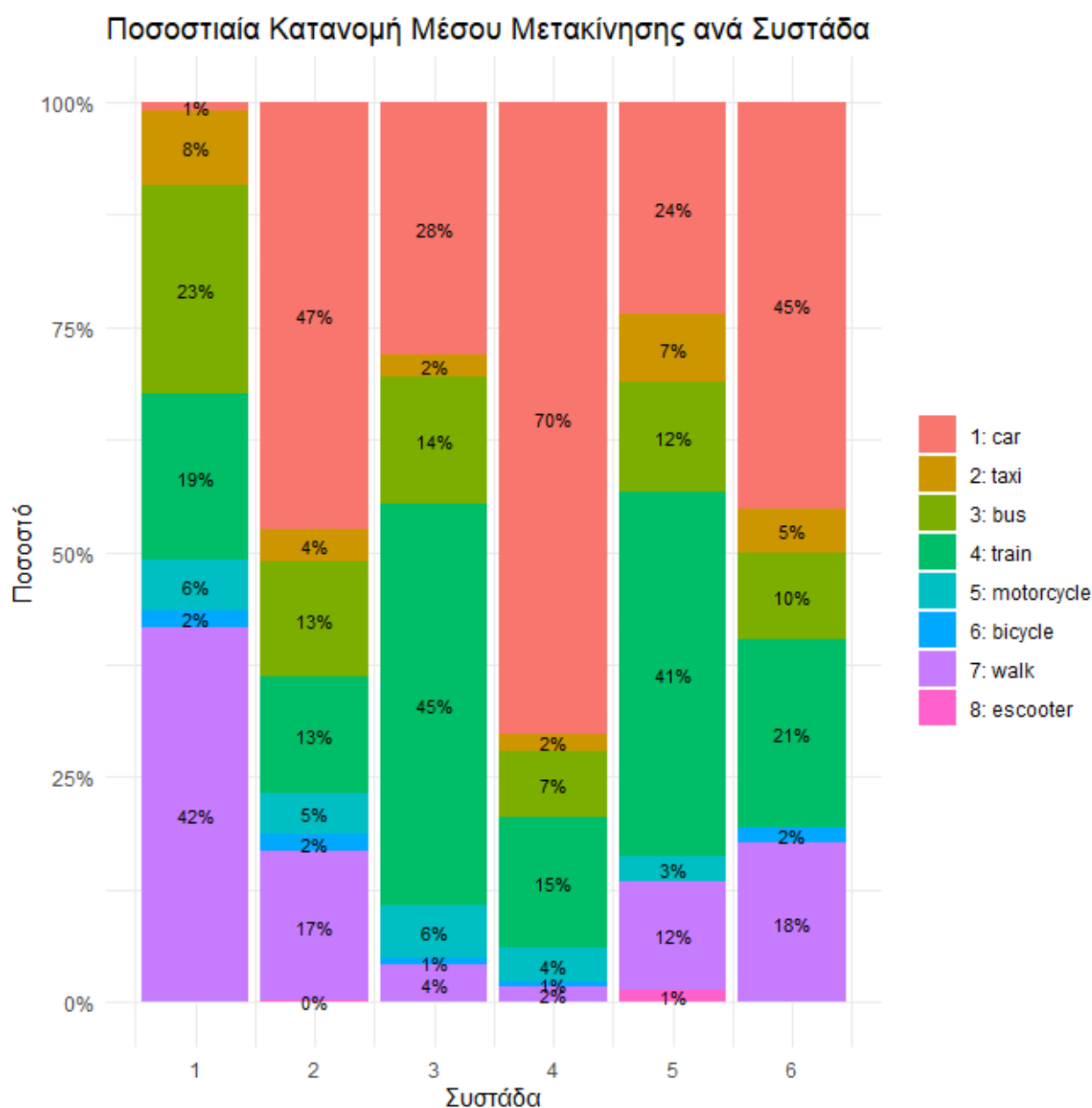
Αναλυτικότερα, όσον αφορά τη μεταβλητή της απόστασης της μετακίνησης (dist), οι συστάδες 1,2 και 6 παρουσιάζουν μικρή διασπορά, ενώ οι συστάδες 3,4 και 5 λίγο μεγαλύτερη. Επίσης, οι τιμές της μεταβλητής της ώρας έναρξης του ταξιδιού παρουσιάζουν πολύ μικρή διασπορά στις συστάδες 1 και 2, ενώ λίγο μεγαλύτερη διασπορά εμφανίζουν στις συστάδες 2,3,4 και 5. Για τη δημιουργία και αυτού του διαγράμματος εξαιρέθηκαν οι χαρακτήρες από τις τιμές των μεταβλητών του μέσου και του σκοπού μετακίνησης και διατηρήθηκε μόνο το αριθμητικό κομμάτι.



Διάγραμμα 5.7 Διασπορά μεταβλητών ανά συστάδα

Από το παραπάνω διάγραμμα προκύπτουν και διακυμάνσεις στη συγκέντρωση των παρατηρήσεων ανά μεταβλητή και ανά συστάδα, όμως αναλύονται παρακάτω με περισσότερη λεπτομέρεια και ποσοστά.

Ακολουθεί διάγραμμα ποσοστιαίας κατανομής του μέσου μετακίνησης ανά συστάδα, δηλαδή για κάθε συστάδα ξεχωριστά, το ποσοστό των μετακινήσεων που εξυπηρετείται από το εκάστοτε μέσο. Τα μέσα μετακίνησης είναι: 1: αυτοκίνητο, 2: ταξί, 3: λεωφορείο, 4: Μέσα Σταθερής Τροχιάς, 5: μοτοσυκλέτα, 6: ποδήλατο, 7: περπάτημα 8: ηλεκτρικό πατίνι. Μέσω αυτού του διαγράμματος μπορεί να χαρακτηριστεί το περιεχόμενο των συστάδων, όσον αφορά το μέσο που εξυπηρετεί κυρίως τις μετακινήσεις που απαρτίζουν την εκάστοτε συστάδα.



Διάγραμμα 5.8 Ποσοστιαία κατανομή μέσων μετακίνησης ανά συστάδα

Αρχικά, στην πρώτη συστάδα παρατηρείται ότι, τη μεγαλύτερη συγκέντρωση σημειώνει το περπάτημα, με ποσοστό 42% και αμέσως μετά ακολουθεί, το λεωφορείο, με ποσοστό 23%. Οι μετακινήσεις με χρήση μέσων σταθερής τροχιάς σημειώνουν ποσοστό 19% των μετακινήσεων της πρώτης ομάδας και οι μετακινήσεις με χρήση ταξί σημειώνουν ποσοστό 8%. Με πολύ μικρά ποσοστά εμφανίζονται η μοτοσυκλέτα (6%), το ποδήλατο (2%) και το αυτοκίνητο (1%).

Στη δεύτερη συστάδα μέγιστο σημειώνει η χρήση του αυτοκινήτου με ποσοστό 47% των μετακινήσεων. Ποσοστό 17% συγκεντρώνουν οι μετακινήσεις με περπάτημα, ενώ 13% συγκεντρώνουν η χρήση του λεωφορείου και των μέσων σταθερής τροχιάς, η κάθε μία. Ακολουθούν οι μετακινήσεις με μοτοσυκλέτα (5%), με ταξί (4%) και με ποδήλατο (2%). Με ποσοστό μικρότερο του 1% εμφανίζονται και οι μετακινήσεις με χρήση ηλεκτρικού πατινιού.

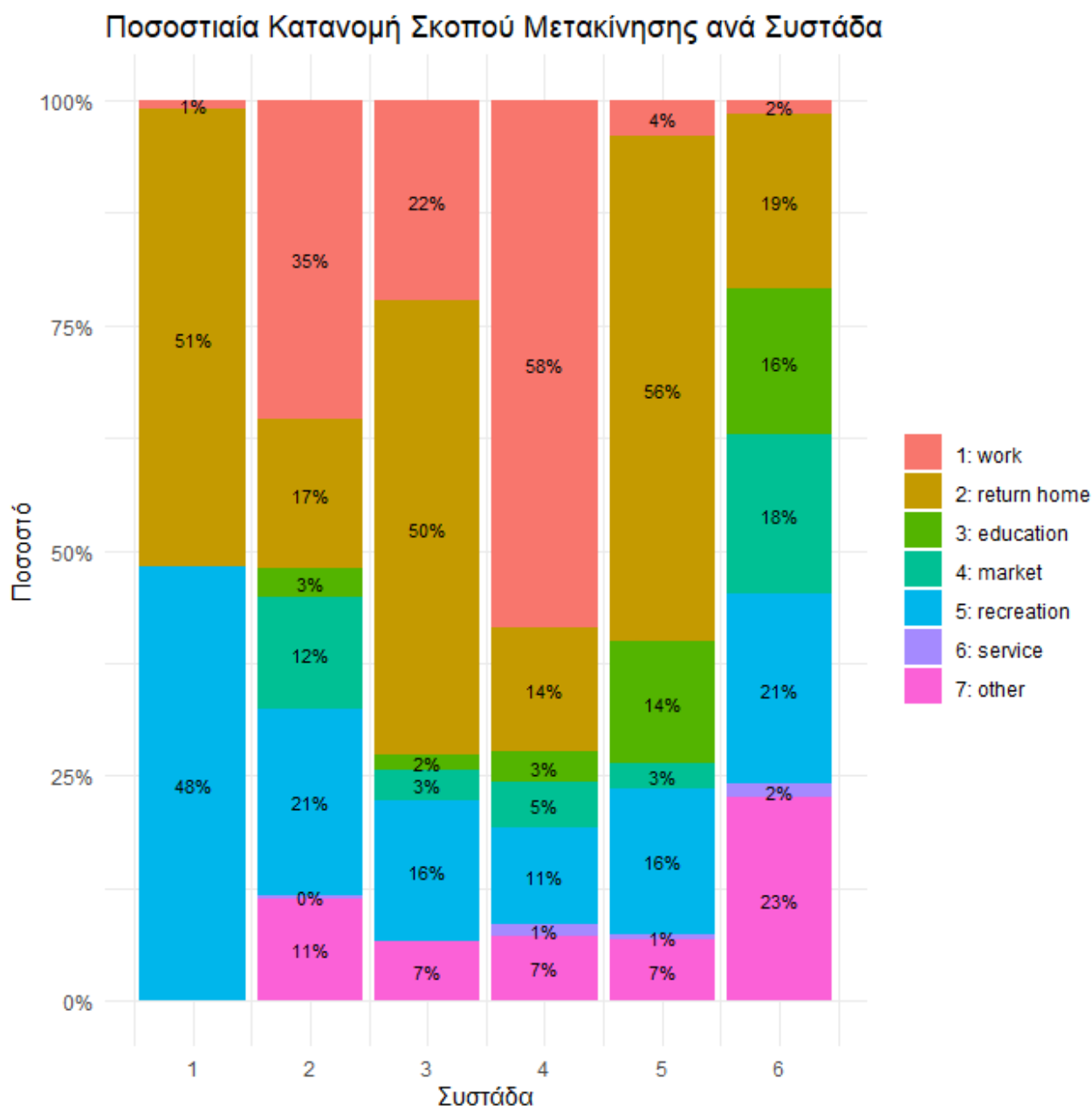
Όσον αφορά την τρίτη συστάδα, η πλειοψηφία των μετακινήσεων της διεξάγεται με τα μέσα σταθερής τροχιάς, συγκεντρώνοντας ποσοστό 45%. Αρκετά μεγάλο ποσοστό, της τάξεως του 28% εμφανίζει και η χρήση του αυτοκινήτου για την πραγματοποίηση των μετακινήσεων της τρίτης συστάδας. Η χρήση του λεωφορείου καλύπτει το 14% των μετακινήσεων και η χρήση της μοτοσυκλέτας το 6% των μετακινήσεων της συστάδας. Μικρότερα ποσοστά συγκεντρώνουν οι μετακινήσεις με περπάτημα (4%), με ταξί (2%) και με ποδήλατο (1%).

Στη συνέχεια, στην τέταρτη συστάδα επικρατεί η χρήση του αυτοκινήτου, με ποσοστό 70%, μεγαλύτερο ποσοστό χρήσης αυτοκινήτου μεταξύ όλων των συστάδων. Στη συνέχεια, τα μέσα σταθερής τροχιάς περιλαμβάνουν το 15% των μετακινήσεων και το λεωφορείο περιλαμβάνει το 7% των μετακινήσεων. Μικρές συγκεντρώσεις παρουσιάζουν η χρήση της μοτοσυκλέτας (4%), το ταξί (2%), το περπάτημα (2%) και το ποδήλατο (1%).

Το μεγαλύτερο μέρος των μετακινήσεων της πέμπτης συστάδας διεξάγονται με τα μέσα σταθερής τροχιάς, με ποσοστό 41%. Επίσης, το 24% των μετακινήσεων διεξάγεται με το αυτοκίνητο και το 12% των μετακινήσεων με λεωφορείο. Ομοίως οι μετακινήσεις με περπάτημα εμφανίζονται σε ποσοστό 12%. Επιπροσθέτως, το ταξί χρησιμοποιείται για το 7% των μετακινήσεων της πέμπτης συστάδας. Το 3% των μετακινήσεων αυτής της ομάδας διεξάγεται μέσω μοτοσυκλέτας και το 1% μέσω του πατινιού. Σημειώνεται ότι, από την πέμπτη συστάδα απουσιάζουν οι μετακινήσεις με χρήση ποδηλάτου.

Τέλος, η έκτη συστάδα αφορά μετακινήσεις που σε ποσοστό 45% εκτελούνται με το αυτοκίνητο. Τα μέσα σταθερής τροχιάς καταλαμβάνουν το 21% των μετακινήσεων της ομάδας και το βάδισμα καταλαμβάνει το 18% των μετακινήσεων της ομάδας. Επίσης, το 10% των μετακινήσεων της έκτης συστάδας πραγματοποιείται μέσω της χρήσης λεωφορείου. Με μικρότερα ποσοστά εμφανίζονται η χρήση του ταξί (5%) και η χρήση της μοτοσυκλέτας (2%). Να σημειωθεί ότι, από την έκτη συστάδα απουσιάζουν μετακινήσεις με χρήση ποδηλάτου και πατινιού.

Στη συνέχεια παρατίθεται το διάγραμμα ποσοστιαίας κατανομής των σκοπών μετακίνησης ανά συστάδα, δηλαδή για κάθε συστάδα ξεχωριστά, το ποσοστό των μετακινήσεων που αντιστοιχεί στον εκάστοτε σκοπό μετακίνησης. Οι σκοποί μετακίνησης είναι: 1: εργασία, 2: επιστροφή στην κατοικία, 3: εκπαίδευση, 4: αγορά, 5: αναψυχή, 6: υπηρεσίες, 7: άλλο. Μέσω αυτού του διαγράμματος μπορεί να χαρακτηριστεί το περιεχόμενο των συστάδων, όσον αφορά το σκοπό για τον οποίο πραγματοποιούνται οι μετακινήσεις που απαρτίζουν την εκάστοτε συστάδα.



Διάγραμμα 5.9 Ποσοστιαία κατανομή σκοπών μετακίνησης ανά συστάδα

Παρατηρείται ότι, στην πρώτη συστάδα επικρατούν κυρίως δύο σκοποί μετακινήσεων. Συγκεκριμένα, οι μετακινήσεις για επιστροφή στο σπίτι καταλαμβάνουν το 51% του συνόλου των μετακινήσεων και οι μετακινήσεις για αναψυχή καταλαμβάνουν το 48% των συνολικών μετακινήσεων της συστάδας.

Απομένει μόνο το 1% των μετακινήσεων της συστάδας το οποίο αποδίδεται σε μετακινήσεις που αφορούν τη μετάβαση στην εργασία. Να σημειωθεί ότι από τη συγκεκριμένη συστάδα απουσιάζουν οι περισσότεροι σκοποί μετακίνησης, συγκεκριμένα: εκπαίδευση, αγορά, υπηρεσία και άλλο.

Στη συνέχεια, εμφανίζονται οι μετακινήσεις της δεύτερης συστάδας. Σε αυτή τη συστάδα η πλειοψηφία των μετακινήσεων εκτελείται με σκοπό την εργασία, με ποσοστό 35%. Οι μετακινήσεις με σκοπό την αναψυχή συγκεντρώνουν το 21% των συνολικών μετακινήσεων της συστάδας. Ακολουθούν οι μετακινήσεις για επιστροφή στην κατοικία (17%), οι μετακινήσεις για αγορές (12%) και οι μετακινήσεις για «άλλους» σκοπούς (11%). Προς το τέλος, μικρότερα ποσοστά σημειώνουν οι μετακινήσεις για εκπαιδευτικούς σκοπούς, με ποσοστό 3%. Οι μετακινήσεις για μετάβαση σε υπηρεσίες καταλαμβάνουν ποσοστό μικρότερο του 1%.

Το μεγαλύτερο μέρος των μετακινήσεων της τρίτης συστάδας διεξάγεται με σκοπό την επιστροφή στην κατοικία. Οι μετακινήσεις για μετάβαση στην εργασία καταλαμβάνουν το 22% των συνολικών μετακινήσεων της ομάδας και οι μετακινήσεις για αναψυχή καταλαμβάνουν το 16%. Η μετακίνηση για άλλους λόγους σημειώνει ποσοστό 7%, για αγορές σημειώνει 3% και για εκπαίδευση μόνο 2%. Από την τρίτη συστάδα απουσιάζουν μετακινήσεις με σκοπό τη μετάβαση σε υπηρεσίες.

Στην τέταρτη συστάδα σημειώνει μέγιστο η μετακίνηση για μετάβαση στην εργασία, με ποσοστό 58% και ακολουθεί η μετακίνηση για επιστροφή στην κατοικία με ποσοστό 14%. Ποσοστό μόλις 11% συγκεντρώνουν οι μετακινήσεις με σκοπό την αναψυχή. Με πολύ μικρότερα ποσοστά εμφανίζονται οι μετακινήσεις για αγορές (5%), για εκπαίδευση (3%), για υπηρεσίες (1%) και άλλες (7%).

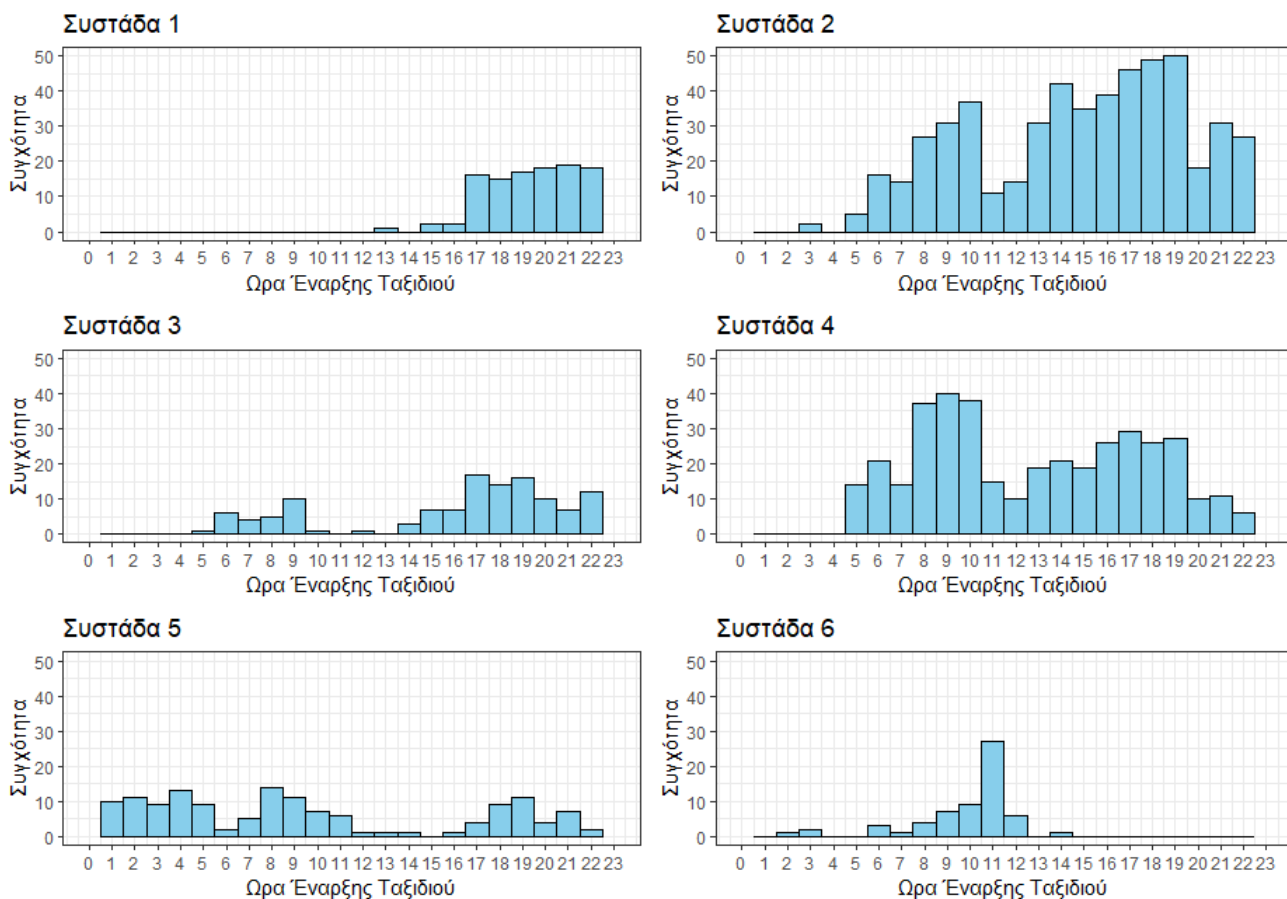
Όσον αφορά την πέμπτη συστάδα, μέγιστο σημειώνουν οι μετακινήσεις για επιστροφή στην κατοικία, συγκεντρώνοντας ποσοστό 56%. Οι μετακινήσεις με σκοπό την αναψυχή κατέχουν ποσοστό 16% και με σκοπό την εκπαίδευση κατέχουν το 14% των συνολικών μετακινήσεων της συστάδας. Το «άλλο» εμφανίζεται με ποσοστό 7%. Μικρά ποσοστά καταλαμβάνουν οι μετακινήσεις για μετάβαση στην εργασία (4%), για μετάβαση σε αγορές (3%) και για μετάβαση σε υπηρεσίες (1%).

Η έκτη συστάδα φαίνεται να περιέχει μετακινήσεις σε παρόμοιες συγκεντρώσεις. Σε μεγαλύτερη συγκέντρωση εμφανίζονται οι μετακινήσεις που αφορούν το σκοπό μετακίνησης «άλλο», με ποσοστό 23%. Αμέσως μετά, ακολουθούν οι μετακινήσεις για αναψυχή με ποσοστό 21% και οι μετακινήσεις για επιστροφή στην κατοικία με ποσοστό 19%. Το 18% των συνολικών μετακινήσεων καταλαμβάνουν οι μετακινήσεις με σκοπό τις αγορές και το 16% καταλαμβάνουν οι μετακινήσεις με σκοπό την εκπαίδευση. Μόνο 2% των μετακινήσεων διεξάγεται με σκοπό την πρόσβαση στην εργασία και το άλλο 2% των μετακινήσεων διεξάγεται με σκοπό τη μετάβαση σε υπηρεσίες.

Μέχρι αυτό το σημείο, οι μετακινήσεις των συστάδων έχουν αναλυθεί σχετικά με το μέσο και το σκοπό μετακίνησης. Απομένουν για ανάλυση οι μεταβλητές της ώρας έναρξης και της απόστασης των ταξιδιών. Η μεταβλητή της ώρας έναρξης των ταξιδιών λαμβάνει τις ακέραιες τιμές από 0 έως 23 και

ουσιαστικά εκφράζει τη μέρα στο εικοσιτετράωρο σύστημα. Όσον αφορά τη μεταβλητή της απόστασης, χρησιμοποιούνται ως μονάδα μέτρησης τα μέτρα και το βήμα του άξονα χ των ιστογραμμάτων που ακολουθούν είναι τα 1000 μέτρα.

Ακολουθούν τα ιστογράμματα κατανομής της ώρας έναρξης των όλων των ταξιδιών ανά συστάδα και στη συνέχεια, για καλύτερη ερμηνεία των αποτελεσμάτων, θα δημιουργηθούν τα διαγράμματα πυκνότητας του 75% και 50% των μετακινήσεων των συστάδων.



Διάγραμμα 5.10 Κατανομή ώρας έναρξης των μετακινήσεων ανά συστάδα

Μέσω του ιστογράμματος κατανομής της ώρας έναρξης όλων των ταξιδιών, παρατηρείται ότι, στην πρώτη συστάδα συμπεριλαμβάνονται μετακινήσεις που έχουν σε μέτρια συγκέντρωση ώρες έναρξης 13:00, 15:00 και 16:00. Σε μεγαλύτερη συγκέντρωση εμφανίζονται οι μετακινήσεις που έχουν ώρες έναρξης 17:00, 18:00, 19:00, 20:00, 21:00 και 22:00. Συνολικά, φαίνεται ότι, στην πρώτη συστάδα συγκεντρώνονται μετακινήσεις με ώρες έναρξης μεταξύ 13:00 και 22:00.

Στη δεύτερη συστάδα, η οποία περιέχει αισθητά μεγαλύτερο πλήθος μετακινήσεων, παρατηρείται ότι, υπάρχει μεγάλη διασπορά μεταξύ των τιμών της ώρας έναρξης. Οι ώρες έναρξης που παρουσιάζουν τις μεγαλύτερες συγκεντρώσεις είναι: 14:00, 17:00, 18:00 και 19:00. Στη συγκεκριμένη συστάδα δεν υπάρχουν μεγάλες διαφοροποιήσεις μεταξύ των συγκεντρώσεων στις ώρες έναρξης.

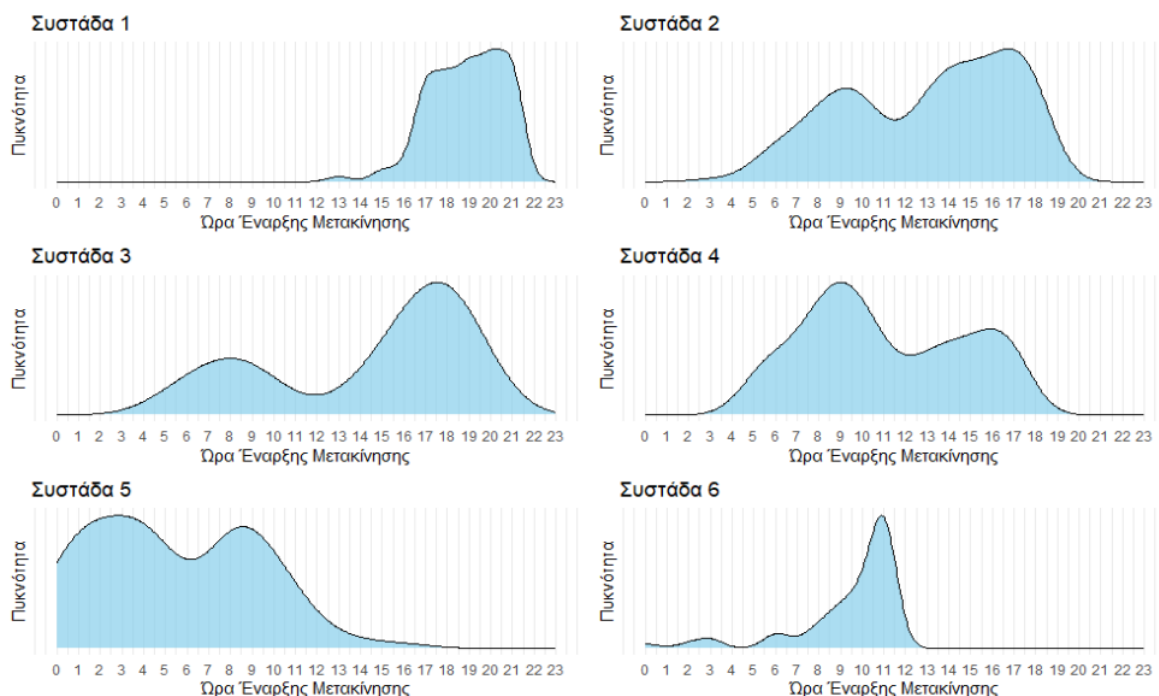
Ακολουθεί η τρίτη συστάδα όπου οι μετακινήσεις φαίνεται να συγκεντρώνονται στις ώρες μεταξύ 6:00 και 9:00 το πρωί καθώς επίσης και μεταξύ 16:00 και 22:00 το απόγευμα. Οι περισσότερες παρατηρήσεις τοποθετούνται μεταξύ 17:00 και 22:00 το απόγευμα.

Η τέταρτη συστάδα περιέχει μεγάλο αριθμό παρατηρήσεων και αυτό αντικατοπτρίζεται και στο ιστόγραμμα ώρας έναρξης των μετακινήσεων, καθώς φαίνεται να περιέχει μετακινήσεις που απαντούν σε διάφορες ώρες έναρξης. Οι μετακινήσεις της συστάδας έχουν ώρες έναρξης από τις 5:00 έως τις 22:00. Οι μεγαλύτερες συγκεντρώσεις παρατηρούνται στις ώρες από 8:00 έως 9:00 το πρωί.

Στην πέμπτη συστάδα κυρίως συμμετέχουν μετακινήσεις από τη 1:00 τα ξημερώματα έως τις 10:00 το πρωί και από τις 18:00 έως τις 21:00 το βράδυ. Πολύ λίγες παρατηρήσεις αποδίδονται στις μεσημεριανές ώρες μεταξύ 12:00 και 17:00. Οι μετακινήσεις της συστάδας αυτής σημειώνονται στις ώρες που προαναφέρθηκαν χωρίς να παρουσιάζουν μεγάλες διαφοροποιήσεις στο πλήθος συγκέντρωσης.

Όσον αφορά την τελευταία συστάδα, οι περισσότερες μετακινήσεις χρονολογούνται στις 11:00 το πρωί. Γύρω από τις 11:00 παρατηρούνται σποραδικά μετακινήσεις σε διάφορες ώρες όπως στις 8:00, 9:00, 10:00, 12:00 και 14:00. Η έκτη συστάδα εμφανώς περιλαμβάνει κυρίως πρωινές μετακινήσεις.

Κατ' επέκταση, παρατίθεται το διάγραμμα κατανομής της ώρας έναρξης των ταξιδιών του 75% των μετακινήσεων ανά συστάδα, με σκοπό την ορθότερη απόδοση χαρακτηριστικών στις συστάδες.

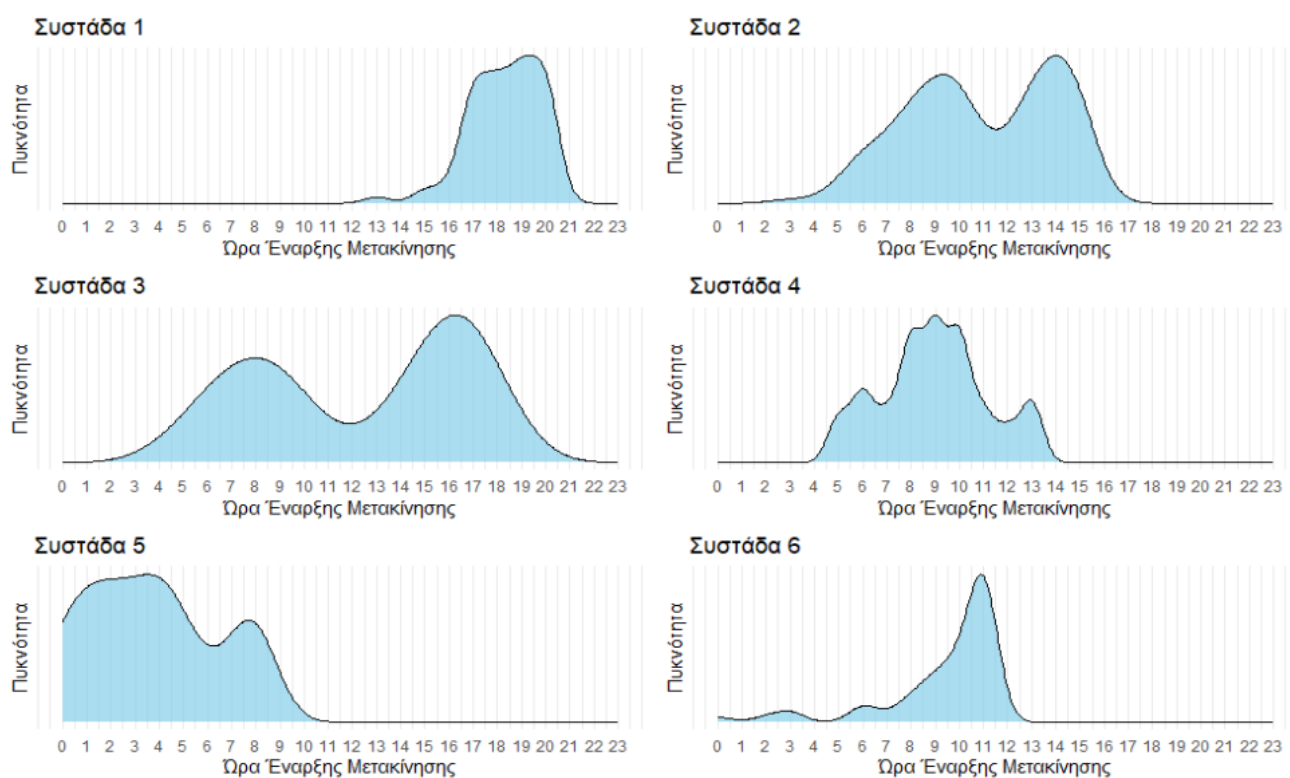


Διάγραμμα 5.11 Κατανομή ώρας έναρξης του 75% των μετακινήσεων ανά συστάδα

Μέσω του παραπάνω διαγράμματος, είναι πιο ευδιάκριτες οι συγκεντρώσεις των μετακινήσεων ανά συστάδα. Στην πρώτη συστάδα φαίνεται πως η πλειοψηφία των μετακινήσεων συγκεντρώνεται στις

ώρες μεταξύ 16:00 και 22:00. Σχετικά με τη δεύτερη συστάδα, παρατηρούνται δύο κορυφές στο διάγραμμα πυκνότητας της. Οι δύο αυτές κορυφώσεις προδίδουν ότι οι μεγαλύτερες συγκεντρώσεις στο 75% των μετακινήσεων παρατηρούνται στις ώρες από 7:00 έως 10:00 και από 13:00 έως 19:00. Ομοίως και στο διάγραμμα της τρίτης συστάδας σχηματίζονται δύο κορυφές στις ώρες μεταξύ 6:00-10:00 και μεταξύ 15:00-20:00. Στη συνέχεια, αναλύοντας το διάγραμμα που αφορά την τέταρτη συστάδα, αξίζει να σημειωθεί ότι παρουσιάζονται δύο κορυφώσεις χωρίς όμως να μεσολαβούν τόσο έντονα «χαμηλά» μεταξύ τους. Οι μέγιστες συγκεντρώσεις για το 75% των μετακινήσεων της τέταρτης συστάδας κυμαίνονται μεταξύ 6:00 και 17:00. Το διάγραμμα της πέμπτης συστάδας μοιάζει αρκετά με αυτό της τέταρτης συστάδας με τη διαφορά ότι είναι μετατοπισμένο προς τις πρώτες πρωινές ώρες. Συγκεκριμένα, φαίνεται ότι, οι μετακινήσεις της πέμπτης συστάδας αφορούν ώρες έναρξης από 0:00 έως 12:00. Τέλος, όσον αφορά το 75% των μετακινήσεων της έκτης συστάδας παρατηρείται ότι σχηματίζουν μία πολύ έντονη/απότομη κορύφωση μεταξύ των ωρών 8:00 και 12:00.

Συμπληρωματικά στα δύο προηγούμενα διαγράμματα και με απώτερο σκοπό την ερμηνεία των συστάδων σε σχέση με τη μεταβλητή της ώρας έναρξης των ταξιδιών, δημιουργήθηκε το διάγραμμα πυκνότητας για το 50% των μετακινήσεων.

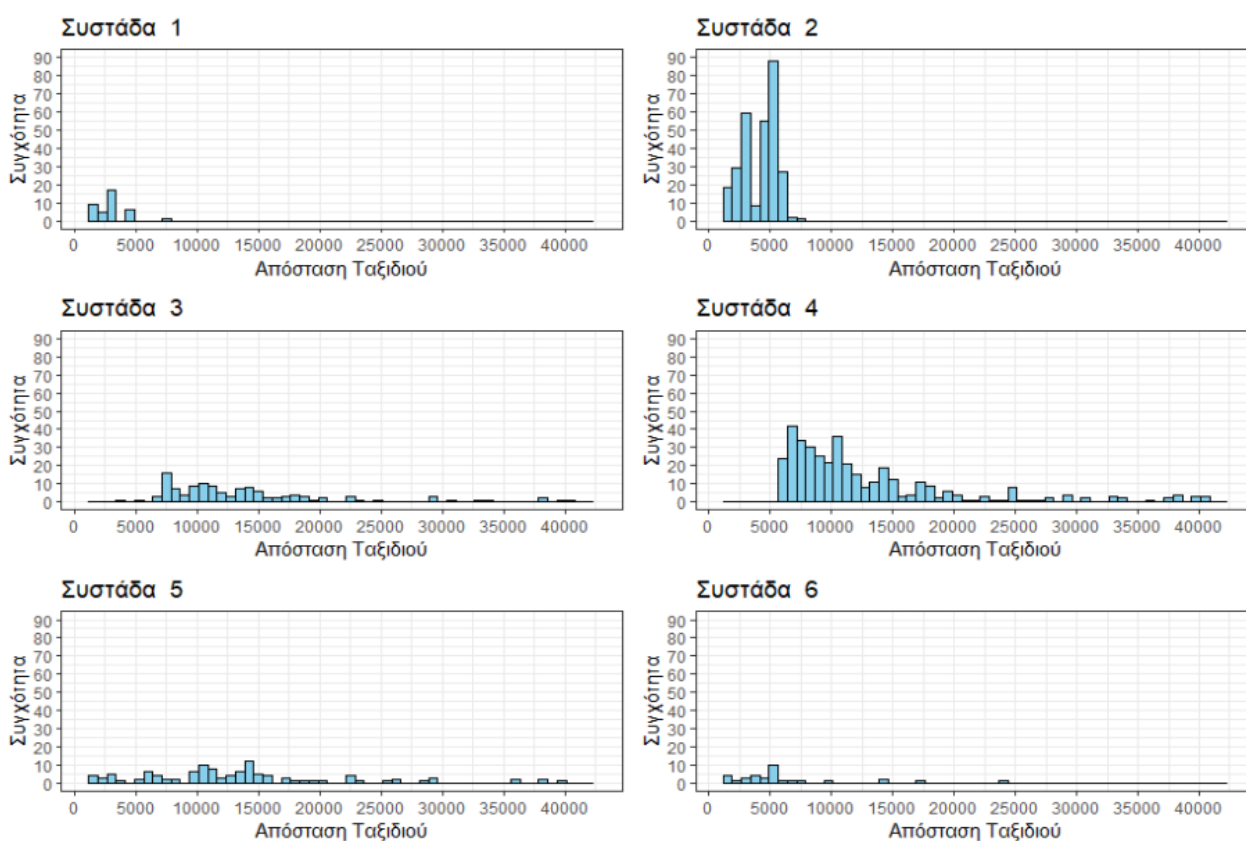


Διάγραμμα 5.12 Κατανομή ώρας έναρξης του 50% των μετακινήσεων ανά συστάδα

Το παραπάνω διάγραμμα μοιάζει κατά πολύ με το διάγραμμα της κατανομής του 75% των μετακινήσεων, με τη διαφορά ότι αυτό είναι πιο «απλοποιημένο» και «απότομο» λόγω της έλλειψης παρατηρήσεων. Στην πρώτη συστάδα παρατηρείται «στένωση» της κορυφής που δημιουργείται σε

σχέση με το διάγραμμα του 75% των παρατηρήσεων. Στη συνέχεια, το διάγραμμα της δεύτερης συστάδας διατηρεί το σχήμα του, δημιουργώντας ακόμα πιο ευδιάκριτες κορυφώσεις στις 9:00 και στις 14:00. Επιπλέον, τα όρια της στενεύουν στις ώρες από τις 7:00 έως τις 16:00. Όσο για την τρίτη συστάδα φαίνονται ξεκάθαρα οι κορυφές που σχηματίζονται στο διάγραμμά της, στις 8:00 και στις 16:00. Εξακολουθεί να καλύπτει μεγάλο μέρος του 24ώρου με ελάχιστες μειώσεις στα άκρα. Το διάγραμμα της τέταρτης συστάδας, σε αυτή την κατανομή, αποκαλύπτει διάφορες αιχμές στις 6:00, 8:00, 9:00, 10:00 και 13:00. Το 50% των μετακινήσεων έχει ώρες έναρξης από τις 4:00 έως τις 14:00, φάσμα αισθητά μειωμένο σε σχέση με την κατανομή του 75% των μετακινήσεων. Η πέμπτη συστάδα δεν εμφανίζει μεγάλες διαφορές σε αυτή την κατανομή σε σχέση με την προηγούμενη του 75% των μετακινήσεων. Μόνη διαφοροποίηση παρατηρείται και πάλι στα όρια των πιθανών ωρών έναρξης, όπου στην προκειμένη περίπτωση μετατοπίστηκαν μεταξύ των 0:00-9:00. Τέλος, στο διάγραμμα της έκτης συστάδας σχηματίζεται μία ξεκάθαρη αιχμή στις μετακινήσεις που ξεκινούν στις 11:00. Το διάγραμμα της έκτης συστάδας φαίνεται να συγκεντρώνεται μεταξύ των ωρών 9:00 και 12:00.

Στη συνέχεια θα αναλυθούν τα αποτελέσματα της συσταδοποίησης σε σχέση με τη μεταβλητή της απόστασης των ταξιδιών. Ακολουθεί, το διάγραμμα κατανομής της απόστασης όλων των μετακινήσεων της κάθε συστάδας.



Διάγραμμα 5.13 Κατανομή απόστασης των μετακινήσεων ανά συστάδα

Όπως φαίνεται, η πρώτη συστάδα περιλαμβάνει μετακινήσεις που έχουν αποστάσεις από 1000 έως και 8000 μέτρα. Υπάρχουν αρκετές μετακινήσεις της τάξεως των 1000 μέτρων, λιγότερες μετακινήσεις μεταξύ των 2000 μέτρων, πολλές μετακινήσεις μεταξύ των 3000 μέτρων, ξανά λιγότερες μετακινήσεις με απόσταση έως 5000 μέτρα και τέλος ελάχιστες μετακινήσεις με απόσταση 8000 μέτρα.

Στη δεύτερη συστάδα περιέχεται μεγάλο πλήθος μετακινήσεων που αφορούν αποστάσεις μεταξύ των 1000 μέτρων και των 9000 μέτρων. Οι περισσότερες παρατηρήσεις τοποθετούνται σε αποστάσεις της τάξεως των 6000 μέτρων. Ακολουθούν μετακινήσεις της τάξεως των 3000 μέτρων και 5000 μέτρων. Γενικά, η συστάδα αυτή περιλαμβάνει μετακινήσεις με απόσταση μικρότερη από 9000 μέτρα. Το πλήθος των αποστάσεων που αγγίζουν τα 8000 και τα 9000 μέτρα είναι πολύ μικρό.

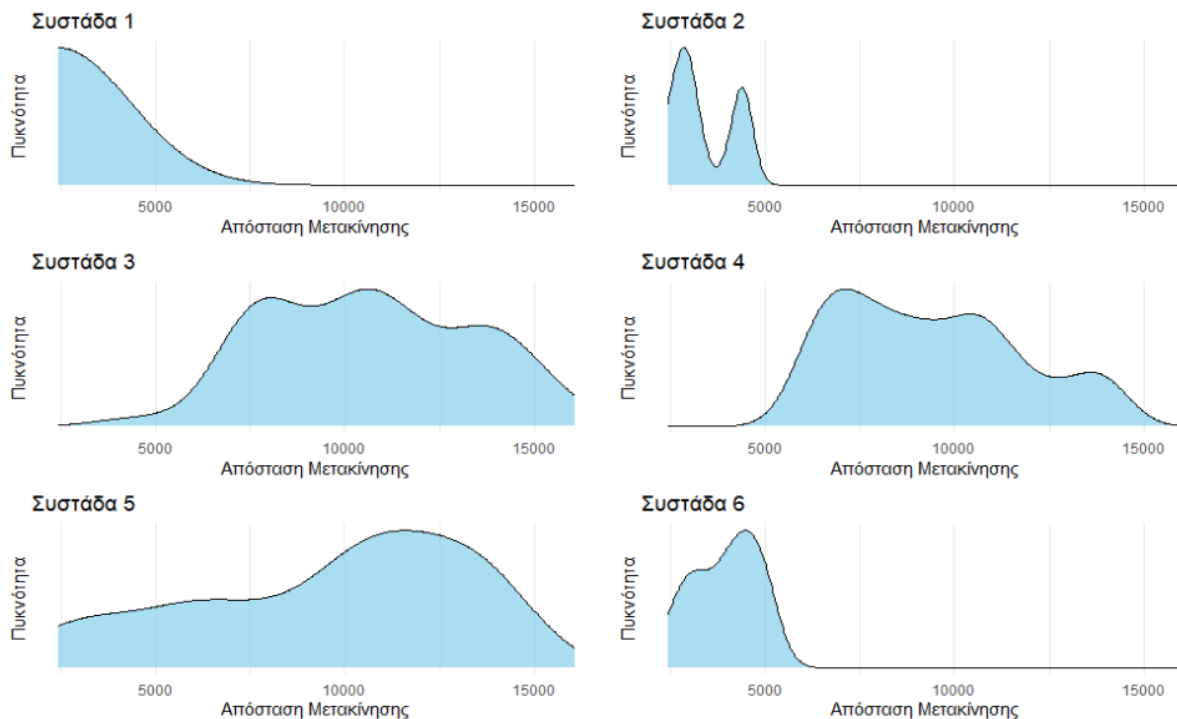
Η τρίτη συστάδα περιλαμβάνει μετακινήσεις διάφορων αποστάσεων χωρίς να παρουσιάζει μεγάλες διαφοροποιήσεις μεταξύ των συγκεντρώσεων των αποστάσεων. Το μεγαλύτερο μέρος των αποστάσεων αυτής της συστάδας είναι από 8000 έως 20000 μέτρα. Παρατηρούνται διάσπαρτες μετακινήσεις σε αρκετές αποστάσεις, οι οποίες είναι μεγαλύτερες των 20000 μέτρων έως και τα 40000 μέτρα.

Όλες οι μετακινήσεις της τέταρτης συστάδας αφορούν αποστάσεις μεγαλύτερες από τα 7000 μέτρα. Οι περισσότερες μετακινήσεις της συστάδας βρίσκονται μεταξύ των 7000 και των 15000 μέτρων. Οι αποστάσεις με τις μεγαλύτερες συγκεντρώσεις του πληθυσμού είναι της τάξεως των 8000 μέτρων και των 11000 μέτρων, ενώ και οι ενδιάμεσες αυτών αποστάσεις συγκεντρώνουν αρκετά μεγάλο κομμάτι του δείγματος.

Η πέμπτη συστάδα έχει μεγάλο εύρος τιμών. Οι αποστάσεις των μετακινήσεων της κυμαίνονται από τα 1000 έως τα 40000 μέτρα, παρουσιάζοντας ορισμένα «κενά». Οι περισσότερες παρατηρήσεις έχουν αποστάσεις από 1000 έως 20000 μέτρα, ενώ σημειώνονται κάποιες διάσπαρτα από τα 20000 έως τα 30000 μέτρα και στα 40000 μέτρα.

Στην έκτη συστάδα παρατηρούνται μετακινήσεις κυρίως στις αποστάσεις από 1000 έως 10000 μέτρα. Επίσης, περιλαμβάνονται κάποιες ελάχιστες μετακινήσεις με απόσταση 15000, 18000 και 25000 μέτρα. Η απόσταση με τις περισσότερες μετακινήσεις είναι αυτή των 6000 μέτρων.

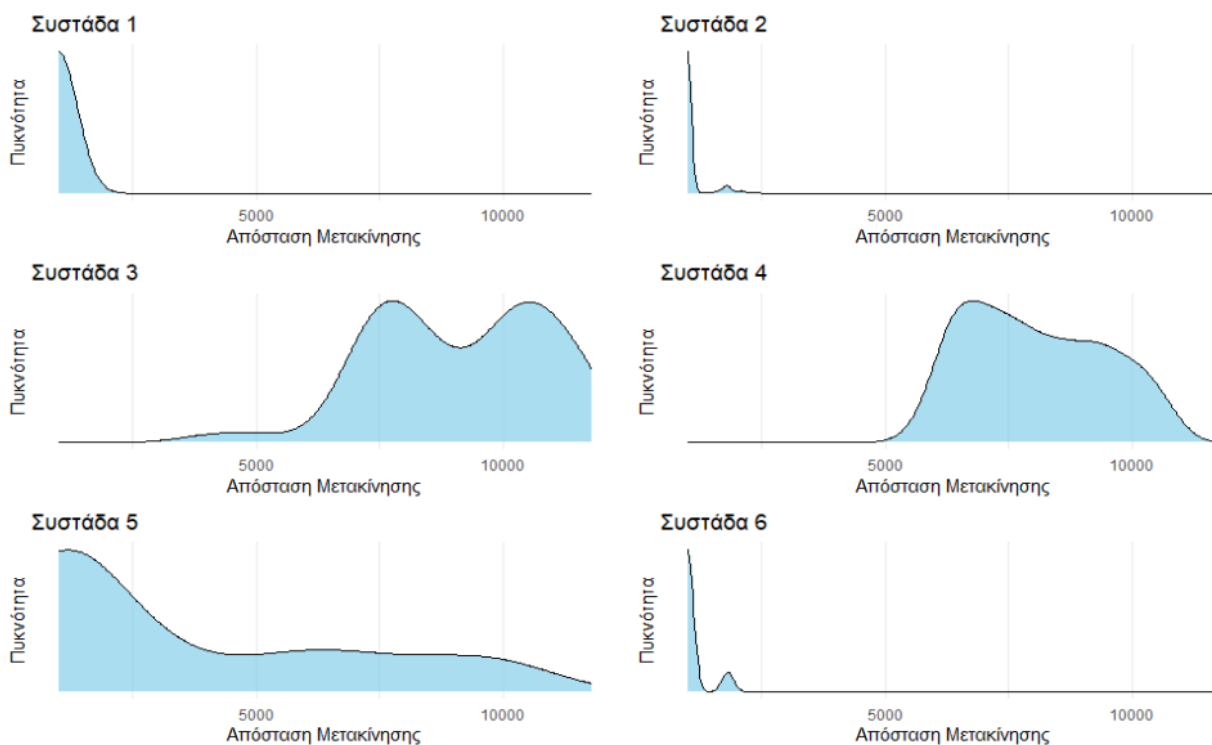
Η μεταβλητή της απόστασης, στη συνέχεια, θα προσεγγιστεί με τον ίδιο τρόπο που προσεγγίστηκε η μεταβλητή της ώρας έναρξης των μετακινήσεων. Έτσι, δημιουργήθηκαν τα διαγράμματα πυκνότητας για το 75% και 50% των μετακινήσεων, όπως παρουσιάζονται στη συνέχεια.



Διάγραμμα 5.14 Κατανομή απόστασης του 75% των μετακινήσεων ανά συστάδα

Βάσει του ανωτέρω διαγράμματος, συμπεραίνεται ότι, όσον αφορά το 75% των μετακινήσεων της πρώτης συστάδας αυτές συγκεντρώνονται μεταξύ των 1000-5000 μέτρων. Στη δεύτερη συστάδα δημιουργούνται δύο κορυφώσεις σε αποστάσεις μικρότερες από τα 5000 μέτρα. Το διάγραμμα της τρίτης συστάδας παρουσιάζει τρεις αιχμές στα 8000, 11000 και 13000 μέτρα. Το 75% των μετακινήσεων της συγκεντρώνονται μεταξύ των 5000 και 15000 μέτρων. Παρομοίως και η τέταρτη συστάδα εμφανίζει τρεις κορυφώσεις στα 7000, 11000 και 13000 μέτρα. Τα εύρος που καλύπτει το 75% των τιμών της τέταρτης συστάδας είναι παρόμοιο με το εύρος των τιμών της τρίτης συστάδας, ενώ η πτώση στα άκρα του όγκου που δημιουργείται για την τέταρτη συστάδα είναι πιο απότομη από αυτή της τρίτης συστάδας. Το διάγραμμα του 75% των τιμών της πέμπτης συστάδας έχει πολύ μεγάλο εύρος τιμών από τα 1000 έως τα 15000 μέτρα, ενώ παράλληλα εμφανίζει ένα μέγιστο στην απόσταση των 12000 μέτρων. Από το τελευταίο διάγραμμα πυκνότητας προκύπτει ότι το 75% των μετακινήσεων της έκτης συστάδας αφορά μετακινήσεις με αποστάσεις μικρότερες των 6000 μέτρων. Στο διάγραμμα δημιουργούνται δύο κορυφές στις αποστάσεις της τάξεως των 1000 και 4000 μέτρων.

Μειώνοντας λίγο ακόμα το ποσοστό των μετρήσεων της κάθε συστάδας που θα συμμετέχουν στο εκάστοτε διάγραμμα, προέκυψαν τα παρακάτω διάγραμμα πυκνότητας που παρουσιάζουν την κατανομή του 50% των μετακινήσεων της κάθε συστάδας.

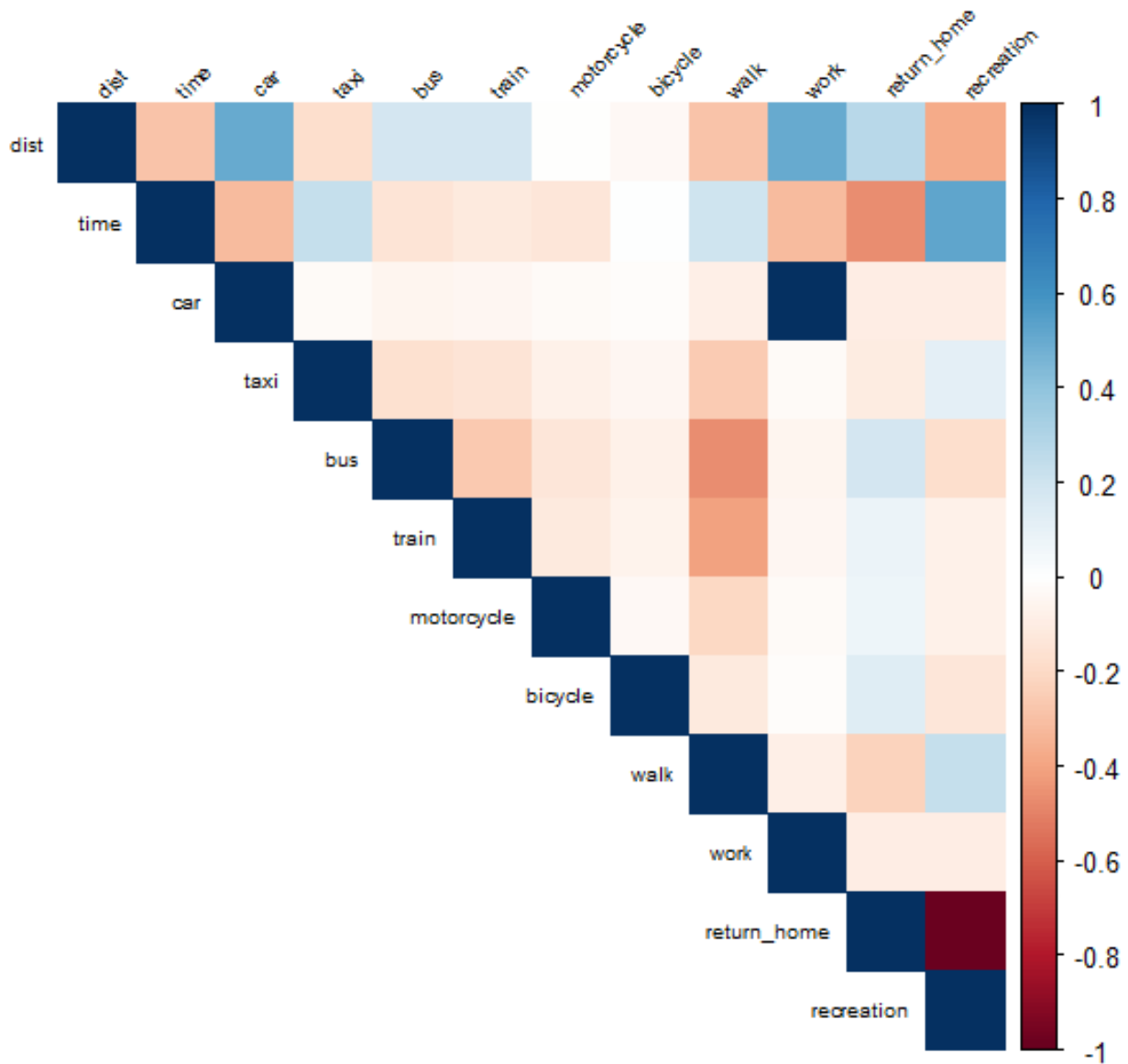


Διάγραμμα 5.15 Κατανομή απόστασης του 50% των μετακινήσεων ανά συστάδα

Λαμβάνοντας υπόψιν μόνο το 50% των μετακινήσεων ανά συστάδα, προκύπτει ότι στην πρώτη συστάδα αφορούν μετακινήσεις με αποστάσεις μικρότερες των 2500 μέτρων. Στη δεύτερη συστάδα το 50% των μετακινήσεων αντιστοιχεί σε αποστάσεις της τάξεως των 1000 μέτρων. Το 50% των μετακινήσεων της τρίτης συστάδας είναι συνδεδεμένο με αποστάσεις μεταξύ των 5000 και 12000 μέτρων. Οι κορυφές σε αυτή την κατανομή της συστάδας γίνονται δύο και σημειώνονται στις αποστάσεις των 8000 και 11000 μέτρων. Η τέταρτη συστάδα περιλαμβάνει κατά 50% μετακινήσεις που έχουν αποστάσεις από τα 6000 μέτρα έως τα 1000 μέτρα και εμφανίζει μέγιστο στις αποστάσεις της τάξεως των 7000 μέτρων. Όσον αφορά την πέμπτη συστάδα, το 50% των μετακινήσεων της καταλαμβάνει μεγάλο εύρος των τιμών και παρουσιάζει μεγάλη συγκέντρωση σε αποστάσεις μικρότερες των 2500 μέτρων. Το διάγραμμα κατανομής της απόστασης του 50% των μετακινήσεων της έκτης συστάδας μοιάζει πολύ με αυτό της δεύτερης συστάδας. Η μεγάλη πλειοψηφία του 50% των μετακινήσεων της ομάδας αφορά αποστάσεις των 1000 μέτρων.

Εν συνεχεία, παρουσιάζονται οι πίνακες συσχέτισης των μεταβλητών για κάθε συστάδα ξεχωριστά. Όσες μεταβλητές δεν συμμετείχαν σε κάποια συστάδα αφαιρέθηκαν από το διάγραμμα της εκάστοτε συστάδας. Οι συσχετίσεις μπορούν να είναι είτε αρνητικές (με διαβαθμίσεις του κόκκινου) είτε θετικές (με διαβαθμίσεις του μπλε).

Διάγραμμα Συσχέτισης Χαρακτηριστικών για τη Συστάδα 1

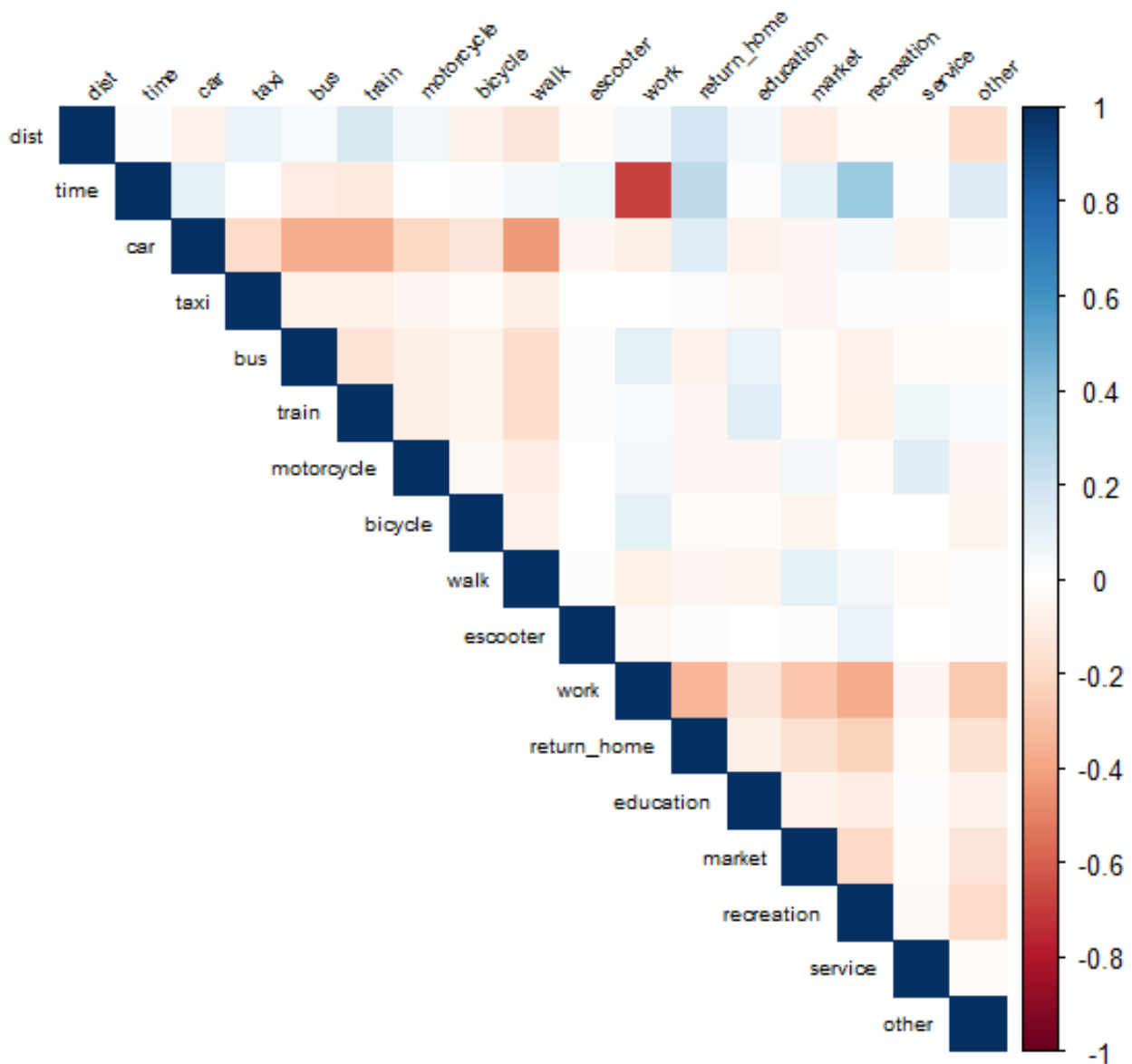


Διάγραμμα 5.16 Διάγραμμα συσχετίσεων μεταβλητών πρώτης συστάδας

Ενδεικτικά, στην πρώτη συστάδα, έντονα θετική συσχέτιση παρουσιάζουν το αυτοκίνητο με την επιστροφή στην οικία, η απόσταση με το αυτοκίνητο καθώς και η απόσταση με την εργασία. Επίσης, θετική συσχέτιση έχουν ο χρόνο έναρξης του ταξιδιού με την αναψυχή. Σε αντίθεση, αρνητικές συσχετίσεις εμφανίζουν οι μεταβλητές της επιστροφής στην κατοικία και της αναψυχής. Επιπλέον, αρνητική συσχέτιση έχουν το λεωφορείο με το περπάτημα, καθώς και τα Μέσα Σταθερής Τροχιάς με το περπάτημα.

Ακολουθεί το διάγραμμα συσχετίσεων για τη δεύτερη συστάδα.

Διάγραμμα Συσχέτισης Χαρακτηριστικών για τη Συστάδα 2

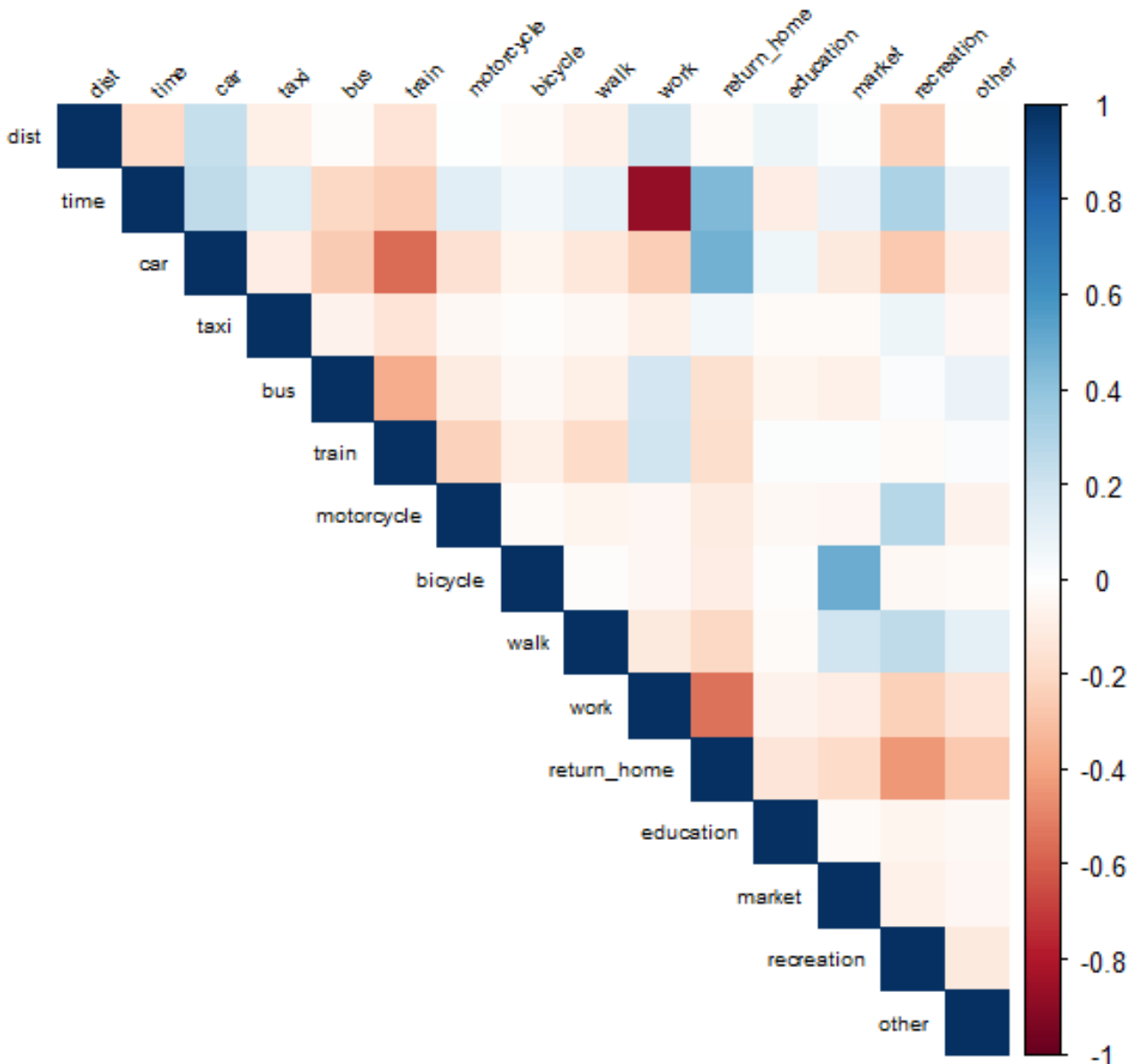


Διάγραμμα 5.17 Διάγραμμα συσχετίσεων μεταβλητών δεύτερης συστάδας

Στη δεύτερη συστάδα μετακινήσεων, παρατηρείται έντονη αρνητική συσχέτιση μεταξύ των μεταβλητών του χρόνου έναρξης της μετακίνησης και της εργασίας. Αρκετά μεγάλη θετική συσχέτιση έχουν ο χρόνος έναρξης της μετακίνησης με τη μετακίνηση για λόγους αναψυχής.

Στη συνέχεια παρουσιάζεται το διάγραμμα με τις συσχετίσεις μεταξύ των μεταβλητών για τις μετακινήσεις της τρίτης συστάδας.

Διάγραμμα Συσχέτισης Χαρακτηριστικών για τη Συστάδα 3

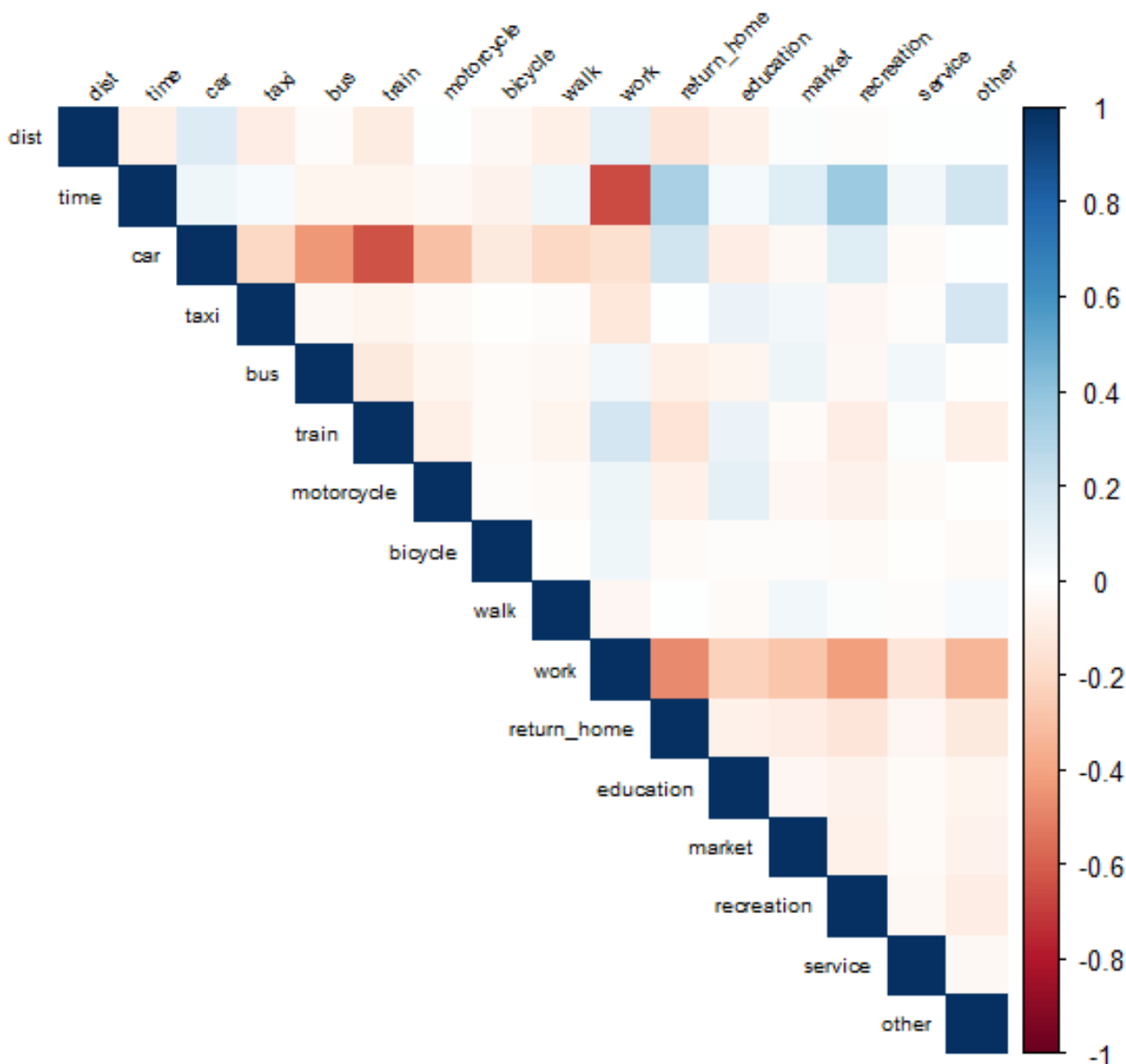


Διάγραμμα 5.18 Διάγραμμα συσχετίσεων μεταβλητών τρίτης συστάδας

Φαίνεται πως στην τρίτη συστάδα, υπάρχει έντονη αρνητική συσχέτιση μεταξύ των μεταβλητών της ώρας έναρξης της μετακίνησης και της μετακίνησης με σκοπό την εργασία, καθώς επίσης και στις μετακινήσεις με αυτοκίνητο και στις μετακινήσεις με μέσα σταθερής τροχιάς. Θετικές συσχετίσεις σημειώνονται μεταξύ των μεταβλητών του αυτοκινήτου και της επιστροφής στην κατοικία και του χρόνου έναρξης της μετακίνησης και της επιστροφής στην κατοικία. Επίσης, σε αυτή την ομάδα μετακινήσεων σημειώνουν θετική συσχέτιση η μεταβλητή της μετακίνησης για αγορές με το ποδήλατο και η μεταβλητή της μετακίνησης για εργασία με το λεωφορείο και τα μέσα σταθερής τροχιάς.

Παρακάτω, παρατίθεται το διάγραμμα των συσχετίσεων των μετακινήσεων της τέταρτης συστάδας.

Διάγραμμα Συσχέτισης Χαρακτηριστικών για τη Συστάδα 4

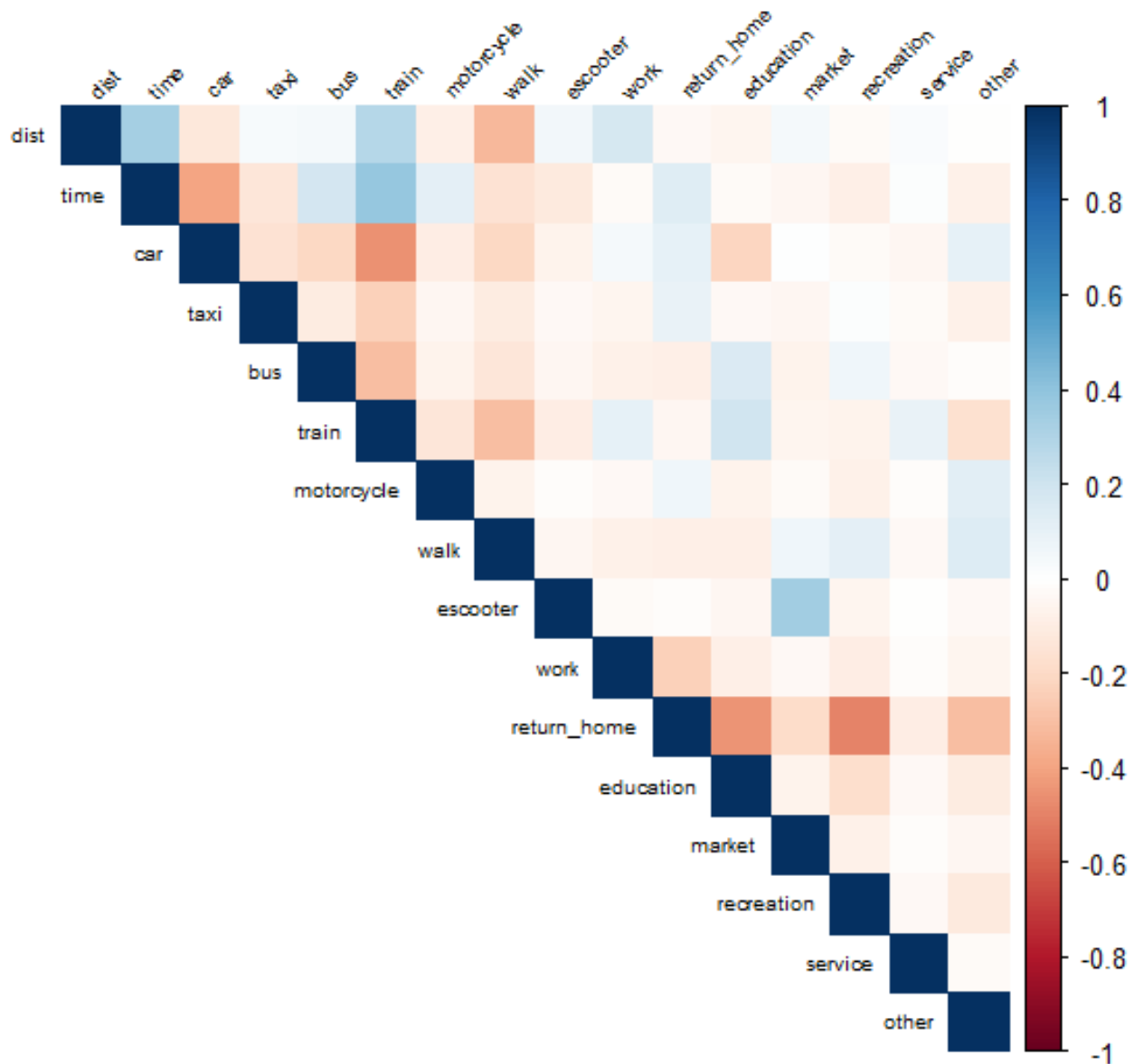


Διάγραμμα 5.19 Διάγραμμα συσχετίσεων μεταβλητών τέταρτης συστάδας

Σε αυτή τη συστάδα, αρνητική συσχέτιση έχουν μεταξύ τους οι μεταβλητές του χρόνου έναρξης της μετακίνησης και της μετάβασης στην εργασία καθώς και οι μεταβλητές της μετακίνησης με αυτοκίνητο και της μετακίνησης με μέσα σταθερής τροχιάς. Το αυτοκίνητο έχει θετική συσχέτιση με την απόσταση της μετακίνησης,

Στη συνέχεια, παρουσιάζεται, ο πίνακας συσχέτισης των μεταβλητών των μετακινήσεων της πέμπτης συστάδας.

Διάγραμμα Συσχέτισης Χαρακτηριστικών για τη Συστάδα 5

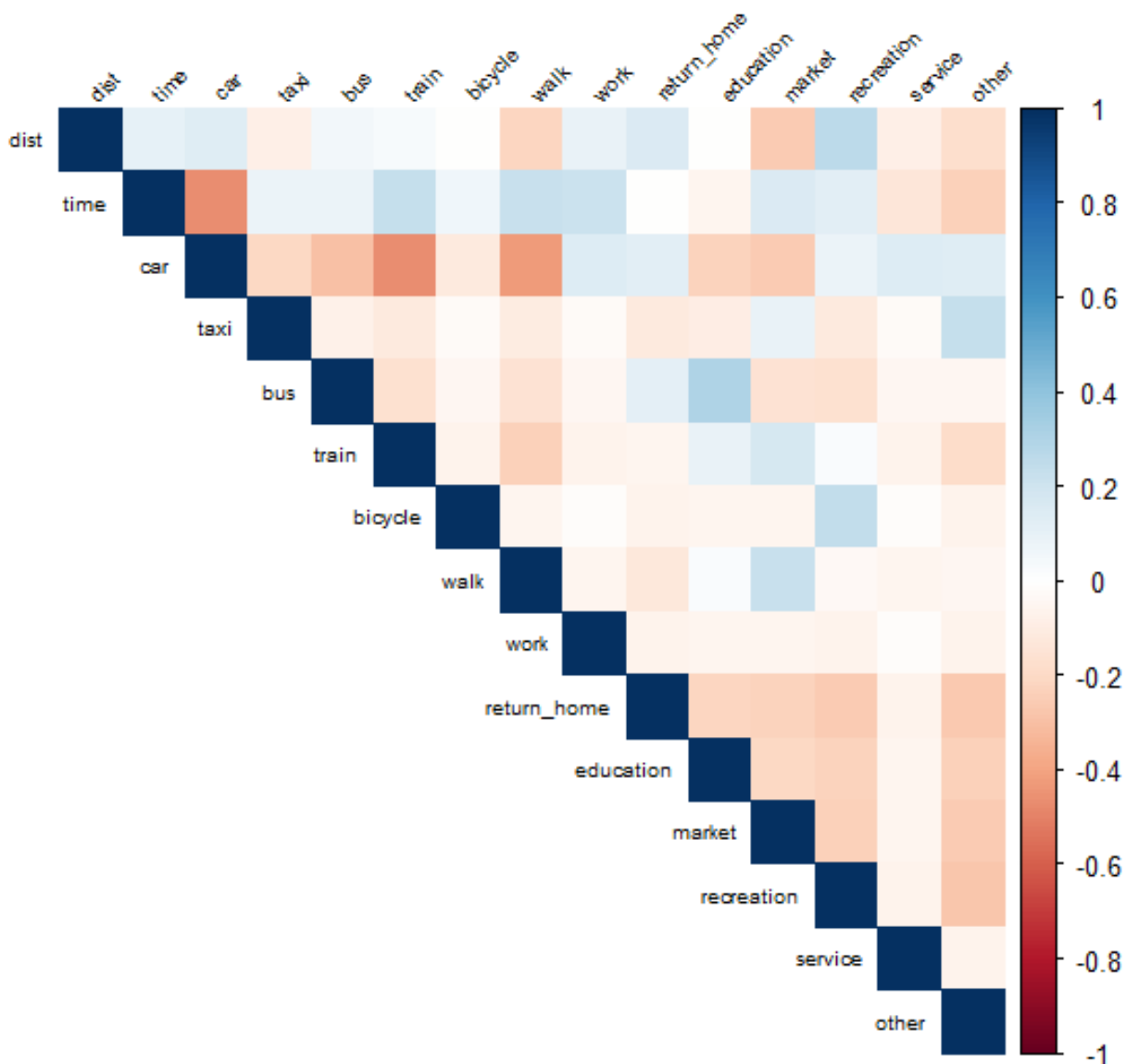


Διάγραμμα 5.20 Διάγραμμα συσχετίσεων μεταβλητών πέμπτης συστάδας

Όσον αφορά την πέμπτη συστάδα, παρατηρείται ότι, η μεταβλητή της απόστασης εμφανίζει θετική συσχέτιση με τις μεταβλητές του χρόνου έναρξης των μετακινήσεων και των μέσων σταθερής τροχιάς. Επίσης, θετική συσχέτιση υπάρχει μεταξύ των μεταβλητών της ώρας έναρξης της μετακίνησης και των μέσων σταθερής τροχιάς. Σημειώνεται ότι θετική συσχέτιση έχουν και οι μετακινήσεις με ηλεκτρικό πατίνι και οι μετακινήσεις για αγορές. Στη συνέχεια, αρνητικές συσχετίσεις έχουν οι μετακινήσεις με αμάξι και οι μετακινήσεις για εκπαιδευτικούς λόγους.

Ακολουθεί το διάγραμμα συσχετίσεων των μεταβλητών των μετακινήσεων που ανήκουν την έκτη συστάδα.

Διάγραμμα Συσχέτισης Χαρακτηριστικών για τη Συστάδα 6



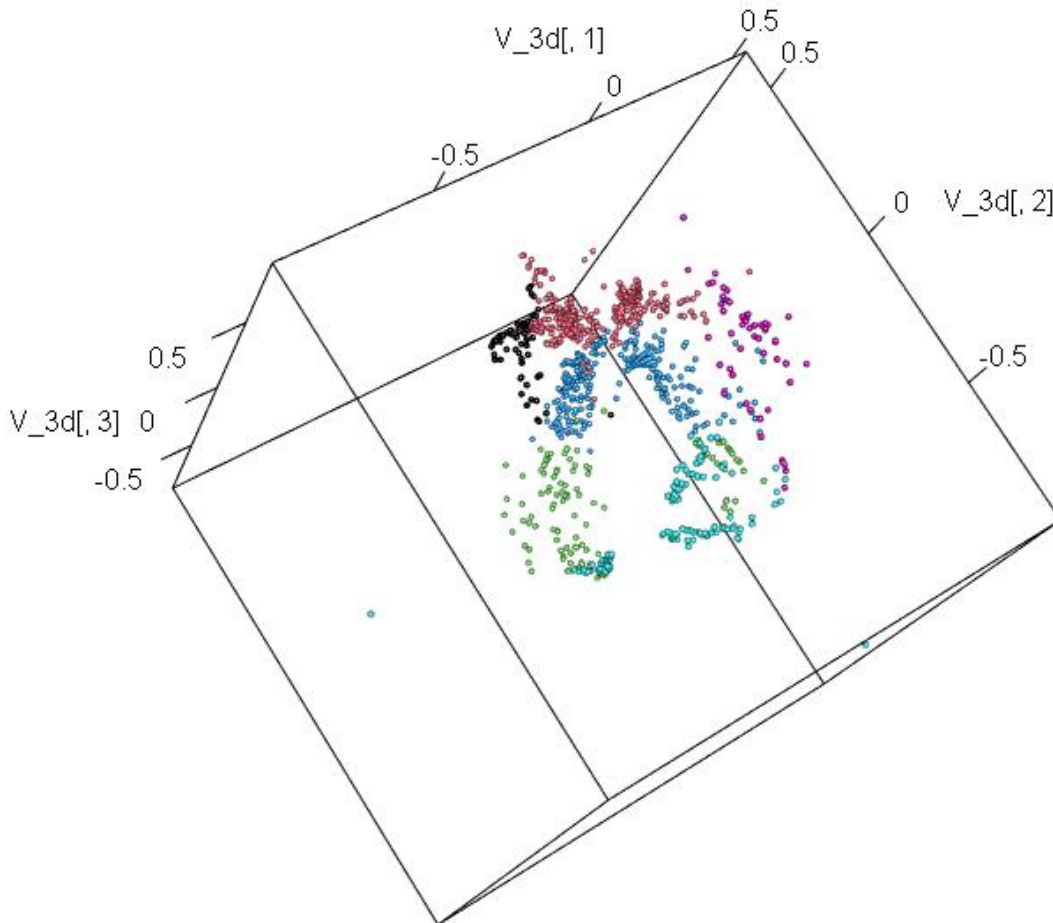
Διάγραμμα 5.21 Διάγραμμα συσχετίσεων μεταβλητών έκτης συστάδας

Σε αυτή τη συστάδα, παρατηρούνται θετικές συσχετίσεις μεταξύ της μεταβλητής του αυτοκινήτου με τις μεταβλητές της μετάβασης στην εργασία και της επιστροφής στην κατοικία, υπηρεσίες και «άλλο». Ιδιαίτερα έντονη θετική συσχέτιση έχει το ταξί με τη μετάβαση σε «άλλες» υποχρεώσεις. Το περπάτημα έχει θετική συσχέτιση με τις αγορές και το ποδήλατο με την αναψυχή. Αρνητική συσχέτιση, εμφανίζουν το αυτοκίνητο με τη μετακίνηση για εκπαιδευτικούς λόγους και η χρήση μέσω σταθερής τροχιάς για «άλλες» υποχρεώσεις.

5.5 Αναγνώριση Συστάδων

Σε αυτό το υπό-κεφάλαιο θα γίνει η αναγνώριση των χαρακτηριστικών της κάθε συστάδας, ώστε να προκύψει το είδος των μετακινήσεων που συμπεριλαμβάνονται σε κάθε μία. Τελικά, θα προκύψουν

κάποιες ομάδες μετακινήσεων, σαν ένας διαφορετικός τρόπος κατηγοριοποίησης και οργάνωσης των μετακινήσεων στην Αθήνα. Ακολουθεί μια τρισδιάστατη αναπαράσταση των σημείων των συστάδων, όπου η κάθε συστάδα απεικονίζεται με διαφορετικό χρώμα.



Διάγραμμα 5.22 Τρισδιάστατη απεικόνιση των παρατηρήσεων των συστάδων

Συστάδα 1^η

Η πρώτη συστάδα, απαρτίζεται από 108 καταγεγραμμένες μετακινήσεις, οι οποίες φαίνεται να αφορούν κατά το 42% μετακινήσεις που πραγματοποιούνται με περπάτημα. Πράγματι το 65% των μετακινήσεων αυτής της ομάδας διεξάγεται είτε με περπάτημα είτε με επιβίβαση σε λεωφορείο, ενώ το 85% των συνολικών μετακινήσεων διεξάγεται με περπάτημα, λεωφορείο ή μέσα σταθερής τροχιάς. Συνεπώς, η πλειοψηφία των μετακινήσεων αυτής της συστάδας διεξάγεται με μέσα πιο «φιλικά» προς το περιβάλλον. Σε σχέση με το σκοπό διεξαγωγής των μετακινήσεων της πρώτης συστάδας, συμπεραίνεται ότι, αυτές γίνονται κατά 48% για λόγους αναψυχής.

Επιπλέον, σχετικά με την ώρα έναρξης των μετακινήσεων της πρώτης συστάδας, συμπεραίνεται ότι, τουλάχιστον το 75% των μετακινήσεων ξεκινά τις απογευματινές και βραδινές ώρες μεταξύ 16:00 και 22:00. Ακόμα, στη συστάδα αυτή, η απόσταση του ταξιδιού εμφανίζει μέση τιμή περίπου ίση με 1687

μέτρα. Άρα η μέση τιμή των αποστάσεων των μετακινήσεων είναι περίπου 1500 μέτρα ενώ η διάμεσος, δηλαδή το κέντρο βάρους των αποστάσεων τείνει να είναι 1000 μέτρα. Από αυτές τις παραμέτρους εξάγεται το συμπέρασμα ότι οι παρατηρήσεις της πρώτης συστάδας αφορούν ενδοζωνικές μετακινήσεις είτε μετακινήσεις σε γειτονικές ή όμορες ζώνες. Όπως φαίνεται και στα προηγούμενα διαγράμματα το 75% των μετακινήσεων της συστάδας έχουν απόσταση μικρότερη ή ίση με 5000 μέτρα.

Συστάδα 2^η

Στη δεύτερη συστάδα ενσωματώνονται 525 μετακινήσεις δηλαδή το 39% του συνόλου των παρατηρήσεων του δείγματος. Η δεύτερη ομάδα είναι η μεγαλύτερη βάσει του πλήθους των παρατηρήσεων. Το ιδιωτικό αυτοκίνητο κυριαρχεί ως μέσο μετακίνησης με ποσοστό 47% των μετακινήσεων, ενώ αν προστεθεί και το 4% των μετακινήσεων που διεξάγεται με ταξί, τότε το ποσοστό των μετακινήσεων που γίνεται με αυτοκίνητο (ως όχημα) πλησιάζει το 51%. Με άλλα λόγια, οι μισές μετακινήσεις της συστάδας πραγματοποιούνται με αυτοκίνητο. Το λεωφορείο συμμετέχει με ποσοστό 13%, τα μέσα σταθερής τροχιάς με 13%, το ποδήλατο με 2% και το περπάτημα με 17%. Άρα το 40% των μετακινήσεων γίνεται με εναλλακτικά μέσα μετακίνησης. Όσον αφορά τους σκοπούς των μετακινήσεων παρατηρείται ότι, το 35% των μετακινήσεων διεξάγεται για μετάβαση στην εργασία, ενώ μεγάλα ποσοστά συγκεντρώνει και η αναψυχή (21%).

Η δεύτερη συστάδα διαφοροποιείται από τις άλλες καθώς οι ώρες έναρξης των ταξιδιών που την απαρτίζουν συγκεντρώνονται στις πρωινές και απογευματινές ώρες. Συγκεκριμένα, τουλάχιστον το 75% των μετακινήσεων της συστάδας πραγματοποιείται από τις 7:00 έως τις 11:00 το πρωί και από τις 13:00 έως τις 19:00. Προκύπτει ότι, οι ώρες έναρξης του 50% των ταξιδιών της δεύτερης συστάδας αντιστοιχούν περίπου σε όσες δεν καλύπτει η πρώτη συστάδα. Επιπροσθέτως, οι αποστάσεις που απαντούν στις μετακινήσεις της δεύτερης συστάδας είναι παρόμοιες με τις αποστάσεις της πρώτης συστάδας. Πράγματι αυτές κυμαίνονται μεταξύ των 1000 και 6000 μέτρων. Μάλιστα, το 75% των μετακινήσεων συγκεντρώνεται σε αποστάσεις μικρότερες των 5 χιλιομέτρων και το 50% σχεδόν εξ' ολοκλήρου σε αποστάσεις της τάξεως του 1 χιλιομέτρου.

Συστάδα 3^η

Το πλήθος των μετακινήσεων της τρίτης συστάδας αγγίζει τις 121. Αναφορικά με τη μεταβλητή του μέσου μετακίνησης, αυτή η συστάδα χαρακτηρίζεται από ταξίδια που κατά 45% εξυπηρετούνται από τα Μέσα Σταθερής Τροχιάς. Επίσης, το 14% των μετακινήσεων διεξάγεται με χρήση του λεωφορείου, άρα συνολικά περίπου το 60% των μετακινήσεων της συστάδας διεξάγεται με τη Δημόσια Συγκοινωνία. Από την άλλη το 30% των μετακινήσεων πραγματοποιείται με χρήση αυτοκινήτου (ΙΧ ή ταξί). Σε αυτή την ομάδα, η μετάβαση για εργασία σημειώνει ποσοστό 22% και για αναψυχή 16%. Αξιοσημείωτα ποσοστά έχουν η μετάβαση στην εργασία (22%) και η αναψυχή (17%). Συμπεραίνεται ότι, οι κυρίαρχοι

σκοποί μετακίνησης στην τρίτη συστάδα είναι η επιστροφή στην κατοικία, η μετάβαση στην εργασία και η μετακίνηση για αναψυχή.

Οι ώρες έναρξης των ταξιδιών της τρίτης συστάδας, για τουλάχιστον το 50% έως 75% των μετακινήσεων, κυμαίνονται μεταξύ 7:00 με 10:00 και κυρίως 16:00 με 20:00. Όσον αφορά την απόσταση των μετακινήσεων, το 75% των παρατηρήσεων της συστάδας εμφανίζει αποστάσεις μεταξύ των 5000 και 15000 μέτρων, ενώ το 50% των παρατηρήσεων μετατοπίζεται από τις 5000 έως τις 10000 μέτρα.

Συστάδα 4^η

Η τέταρτη συστάδα αποτελείται 383 μετακινήσεις. Η συντριπτική πλειοψηφία των μετακινήσεων της τέταρτης συστάδας πραγματοποιούνται με χρήση του αυτοκινήτου, με ποσοστό 70%. Ο κύριος σκοπός μετακίνησης είναι η μετάβαση στην εργασία που αγγίζει το 58%. Οι ώρες έναρξης του 75% των μετακινήσεων κυμαίνονται μεταξύ 6:00 και 17:00, ενώ το 50% των μετακινήσεων κυμαίνεται στις ώρες από 6:00 έως 13:00. Οι αποστάσεις που απαντούν στο 75% των μετακινήσεων της συστάδας είναι μεταξύ των 5000 και 15000 μέτρων, και το 50% των μετακινήσεων συγκεντρώνεται στις αποστάσεις από 5000 έως 10000 μέτρα.

Συστάδα 5^η

Η πέμπτη συστάδα, φαίνεται να συγκεντρώνει 148 μετακινήσεις, οι οποίες γίνονται με Μέσα Σταθερής Τροχιάς, λεωφορείο ή με περπάτημα με συνολικό ποσοστό 65% των μετακινήσεων και κάποιες μετακινήσεις με χρήση αυτοκινήτου (ΙΧ ή ταξί) σε ποσοστό 31%. Εδώ συμμετέχουν άλλοι λόγοι μετακίνησης όπως αυτός της εκπαίδευσης και της αναψυχής. Το 75% των ωρών έναρξης των ταξιδιών της συστάδας βρίσκεται μεταξύ 0:00 και 12:00 και το 50% των ωρών συγκεντρώνεται από 0:00 έως 10:00. Οι αποστάσεις των μετακινήσεων που περιλαμβάνονται στην πέμπτη συστάδα, στο 75% των ταξιδιών έχουν μεγάλο εύρος τιμών, όμως παρουσιάζουν υψηλή πυκνότητα στις τιμές μεταξύ 10000 και 15000 μέτρων. Το 50% των μετακινήσεων αφορά αποστάσεις μεταξύ των 1000 και 10000 μέτρων.

Συστάδα 6^η

Στην έκτη συστάδα περιλαμβάνονται 62 μετακινήσεις και επικρατεί η χρήση του αυτοκινήτου (ΙΧ) με ποσοστό 45% και προσθέτοντας τη χρήση ταξί, το ποσοστό χρήσης αυτοκινήτου αγγίζει το 50%. Ικανοποιητικό ποσοστό συγκεντρώνουν και οι πιο πράσινοι τρόποι μετακίνησης όπως το λεωφορείο, τα Μέσα Σταθερής Τροχιάς και το περπάτημα (49%). Σε αυτή τη συστάδα εμφανίζονται όλοι οι σκοποί μετακίνησης με παρόμοια ποσοστά, εκτός από το σκοπό της μετάβασης στην εργασία. Επιπλέον, οι ώρες έναρξης των ταξιδιών αυτών χρονολογούνται, σε ποσοστό μεγαλύτερο του 75%, στις ώρες από 8:00 έως 12:00. Η μεταβλητή της απόστασης για το 75% των μετακινήσεων φαίνεται να λαμβάνει τιμές από 1000 έως 5000 μέτρα, ενώ για το 50% των μετακινήσεων φαίνεται τεράστια συγκέντρωση στην απόσταση των 1000 μέτρων.

Στη συνέχεια, παρουσιάζεται ένας συνοπτικός πίνακας που συγκεντρώνει τα βασικά χαρακτηριστικά κάθε συστάδας, όσον αφορά τα γνωρίσματα των ίδιων των μετακινήσεων που τις απαρτίζουν.

Πίνακας 5.4 Συνοπτική παρουσίαση χαρακτηριστικών κάθε συστάδας

Συστάδα	Κύρια Μέσα Μετακίνησης	Μέση Απόσταση Μετακίνησης (Τυπική Απόκλιση)	Ώρες Μετακινήσεων (75%)	Κύριοι Σκοποί Μετακίνησης
Συστάδα 1	Περπάτημα (42%) Λεωφορείο (23%)	1.68 χλμ. (± 1.17)	16:00-22:00	Αναψυχή (48%) Εργασία (1%)
Συστάδα 2	Αυτοκίνητο (47%) Περπάτημα (17%)	2.70 χλμ. (± 1.82)	07:00-11:00	Εργασία (35%) Αναψυχή (21%)
Συστάδα 3	ΜΣΤ (45%) Αυτοκίνητο (28%)	14.04 χλμ. (± 7.46)	07:00-20:00	Εργασία (22%) Αναψυχή (16%)
Συστάδα 4	Αυτοκίνητο (70%) ΜΣΤ (15%)	12.95 χλμ. (± 7.67)	06:00-17:00	Εργασία (58%) Αναψυχή (11%)
Συστάδα 5	ΜΣΤ (41%) Αυτοκίνητο (24%)	10.60 χλμ. (± 9.35)	00:00-12:00	Αναψυχή (16%) Εκπαίδευση (14%)
Συστάδα 6	Αυτοκίνητο (45%) ΜΣΤ (21%)	3.67 χλμ. (± 4.30)	08:00-12:00	Άλλος (23%) Αναψυχή (21%)

Επιπροσθέτως, δημιουργήθηκε ένας πίνακας με τα βασικά δημογραφικά χαρακτηριστικά της κάθε συστάδας. Όπως φαίνεται παρακάτω, έγινε ποσοστιαία κατανομή των κοινωνικό-οικονομικών γνωρισμάτων του δείγματος σε όλες τις συστάδες,

Πίνακας 5.5 Δημογραφικά χαρακτηριστικά συστάδων

		Συστάδα						Σύνολο
		1	2	3	4	5	6	
Φύλο	Άντρας	3.5%	14.0%	3.4%	12.2%	4.8%	2.0%	40.0%
	Γυναίκα	4.4%	25.0%	5.7%	16.1%	6.2%	2.7%	60.0%
Ηλικιακή Ομάδα (έτη)								
	16-30	3.7%	17.3%	4.7%	12.4%	6.7%	1.7%	46.5%
	30-40	2.4%	8.6%	1.3%	6.9%	2.0%	0.8%	21.9%
	40-50	0.8%	5.3%	2.0%	5.6%	1.5%	0.4%	15.5%
	50-65	0.9%	7.5%	1.0%	3.2%	0.7%	1.4%	14.7%
	>65	0.2%	0.3%	0.1%	0.3%	0.2%	0.5%	1.4%
Μορφωτικό Επίπεδο								
	Δημοτικό	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	0.1%
	Λύκειο	1.6%	7.2%	2.0%	3.7%	3.7%	1.6%	19.7%
	Πτυχίο	3.1%	16.5%	3.8%	11.6%	3.6%	1.5%	40.1%
	Μεταπτυχιακό/Διδακτορικό	3.4%	15.4%	3.1%	13.0%	3.7%	1.5%	40.0%
Επαγγελματική Κατάσταση								
	Άεργος	0.4%	1.9%	0.1%	0.7%	0.2%	1.2%	4.5%
	Άνεργος	0.4%	0.5%	1.6%	0.3%	0.4%	0.1%	3.4%
	Φοιτητής	1.7%	6.0%	7.3%	2.9%	3.4%	1.0%	22.3%
	Εργαζόμενος	5.4%	30.6%	0.7%	24.6%	6.9%	2.4%	70.5%
Εισόδημα (€)								
	0	1.4%	4.7%	1.1%	2.5%	2.1%	0.9%	12.7%
	<750	1.6%	7.6%	1.9%	4.8%	3.1%	0.9%	19.9%
	750-1500	3.8%	20.0%	4.0%	14.5%	4.1%	2.6%	49.0%
	1500-2500	1.1%	7.1%	1.5%	5.1%	1.1%	0.3%	16.3%
	>2500	0.3%	0.4%	0.2%	1.2%	0.3%	0.0%	2.5%

6 ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 Σύνοψη

Η παρούσα Διπλωματική Εργασία ανέλαβε την εξερεύνηση των δυναμικών μοτίβων επιλογής μεταφορικού μέσου στην ευρύτερη περιοχή της Αθήνας, μελετώντας τα χώρο-χρονικά πρότυπα των μετακινήσεων. Η τεχνική κατανόησης αυτών των μοτίβων βασίστηκε στην ταξινόμηση των μετακινήσεων αναλόγως του μέσου μεταφοράς, του σκοπού, του χρόνου και της απόστασης, ενώ παράλληλα λήφθηκαν υπόψη και στοιχεία κοινωνικοοικονομικού χαρακτήρα των ατόμων που εκτελούν τις μετακινήσεις.

Στην προσπάθεια αυτή, επιστρατεύτηκε η ανάλυση της διεθνούς βιβλιογραφίας σχετικά με μοντέλα μετακίνησης και τεχνικές συσταδοποίησης. Επίσης, χρησιμοποιήθηκαν δεδομένα που συλλέχθηκαν μέσω ερωτηματολογίων. Τα ερωτηματολόγια διαμορφώθηκαν με βάση την τεχνική των αποκαλυπτόμενων προτιμήσεων, καλώντας τους συμμετέχοντες να παρέχουν δημογραφικές πληροφορίες και τις συνήθειες μετακίνησής τους κατά τη διάρκεια μιας τυπικής ημέρας, συγκεκριμένα το μέσο, τον σκοπό, την ώρα, την προέλευση και τον προορισμό της μετακίνησης.

Ακολούθησε η στατιστική επεξεργασία και ανάλυση των συλλεγμένων απαντήσεων, χρησιμοποιώντας αλγορίθμους συσταδοποίησης που ενσωμάτωσαν τις μεταβλητές του μέσου μεταφοράς, του σκοπού, της ώρας και της απόστασης των μετακινήσεων. Τα ευρήματα αυτής της διαδικασίας παρουσιάστηκαν μέσω ποσοτικής ανάλυσης, προσδιορίζοντας στατιστικά τα χαρακτηριστικά των μετακινήσεων.

Στο τελικό στάδιο, ερμηνεύτηκαν τα αποτελέσματα και προέκυψαν τα συμπεράσματα, όπου αναδείχθηκαν συγκεκριμένα μοτίβα μετακίνησης, επιτρέποντας την αναγνώριση περιορισμών και την σύνταξη προτάσεων για μελλοντικές έρευνες.

6.2 Κύρια ευρήματα

Αφού πραγματοποιήθηκε η συσταδοποίηση των μετακινήσεων, έγινε ο σχολιασμός των αποτελεσμάτων. Τα αποτελέσματα αυτά, σε συνδυασμό με την βιβλιογραφία και την γενικότερη έρευνα που πραγματοποιήθηκε στα πλαίσια της εργασίας, οδήγησαν σε μια σειρά συμπερασμάτων, τα οποία εκπληρώνουν τον αρχικό στόχο της εργασίας.

Από την διερεύνηση προέκυψαν, 6 διαφορετικά πρότυπα μετακίνησης, τα οποία εμφανίζουν ποικιλομορφία όσον αφορά τα χαρακτηριστικά τους. Αρχικά, ένα από τα πρότυπα που παρουσιάστηκαν είναι αυτό της «Βραδινής βόλτας κάπου κοντά». Συμπεραίνεται ότι, για μετακινήσεις σε μικρή απόσταση, που αφορούν την αναψυχή, κυρίως τις βραδινές ώρες, επιλέγεται ως μέσο το περπάτημα ή το λεωφορείο. Επιπλέον, προέκυψε το πρότυπο «Η δουλειά είναι κοντά, αλλά επιλέγω αυτοκίνητο». Έγινε εμφανές ότι πολλοί κάτοικοι της Αθήνας, επιλέγουν το αυτοκίνητο για μετάβαση στην εργασία, ακόμα και αν αυτή βρίσκεται σε μικρή απόσταση από την κατοικία τους. Επόμενο, είναι

το πρότυπο «Με το Μετρό και τη Δημόσια Συγκοινωνία», αφορά μετακινήσεις μεγάλων αποστάσεων κυρίως προς την εργασία, όπου παρατηρήθηκε αυξημένη χρήση των Μέσων Σταθερής Τροχιάς και των Δημόσιων Συγκοινωνιών. Ένα ακόμα πρότυπο («Η δουλειά είναι μακριά, άρα επιλέγω αυτοκίνητο») που δημιουργήθηκε, αφορά κυρίως τις πρωινές μετακινήσεις με σκοπό την εργασία, σε αρκετά μεγάλες αποστάσεις, όπου ως μέσο επιλέγεται ως επί το πλείστον το αυτοκίνητο. Οι «Μετακινήσεις νέων» όπως χαρακτηρίστηκε το επόμενο πρότυπο αφορά πρωινές μετακινήσεις για αναψυχή αλλά και εκπαίδευση, με μέτριες αποστάσεις, όπου σημειώθηκε αυξημένη χρήση των Μέσων Σταθερής Τροχιάς. Τέλος, παρατηρήθηκε ένα πρότυπο που αφορά μετακινήσεις σε σχετικά μικρές αποστάσεις προς άλλες υποχρεώσεις, εκτός εργασίας και εκπαίδευσης («Μικρές μετακινήσεις προς άλλες υποχρεώσεις»). Εκεί παρατηρήθηκε έντονη χρήση του αυτοκινήτου αλλά και των Μέσων Σταθερής Τροχιάς, ίσως λόγω του ότι οι μετακινήσεις αυτές διεξάγονται τις πρωινές ώρες.

Γενικότερα, όσον αφορά το μέσο μετακίνησης, το περπάτημα συνδέεται άμεσα με τις μετακινήσεις για λόγους αναψυχής. Επίσης, το μεγαλύτερο μέρος των μετακινήσεων για εκπαιδευτικούς σκοπούς διεξάγεται με τη Δημόσια Συγκοινωνία, δηλαδή είτε με Μέσα Σταθερής Τροχιάς, είτε με λεωφορείο. Επιπλέον, οι μετακινήσεις για αγορές στην πλειοψηφία τους πραγματοποιούνται με περπάτημα. Όπως είναι αναμενόμενο, οι μετακινήσεις με μικρότερες αποστάσεις εξυπηρετούνται αρκετά από την Δημόσια Συγκοινωνία, σε αντίθεση με τις μετακινήσεις μεγαλύτερων αποστάσεων όπου χρησιμοποιείται περισσότερο το αυτοκίνητο.

Όσον αφορά την ώρα έναρξης των μετακινήσεων, παρατηρείται ότι, οι πρωινές μετακινήσεις διεξάγονται κυρίως με αυτοκίνητο, λεωφορείο και μέσα σταθερής τροχιάς. Αντιθέτως, οι απογευματινές και βραδινές μετακινήσεις πραγματοποιούνται κατά κύριο λόγο με περπάτημα, μοτοσυκλέτα και ποδήλατο.

Σε σχέση με την ηλικία, σημειώνεται ότι, οι νέοι επιλέγουν περισσότερο την Δημόσια Συγκοινωνία σε σχέση με τις μεγαλύτερες ηλικιακές ομάδες που προτιμούν το αυτοκίνητο. Επιπλέον, η ίδια προτίμηση για το αυτοκίνητο παρατηρείται και με την αύξηση του μορφωτικού επιπέδου του δείγματος.

6.3 Περιορισμοί

Κατά τη διεξαγωγή της έρευνας παρουσιάστηκαν ορισμένοι περιορισμοί. Συγκεκριμένα, όσον αφορά το δείγμα, που αποτελείται από 513 άτομα, παρατηρείται ότι, αν και υπάρχουν απαντήσεις σε όλες τις ηλικιακές ομάδες, το μεγαλύτερο ποσοστό συγκεντρώνεται στις ηλικίες 16-30 ετών. Επίσης, σχετικά με το μορφωτικό επίπεδο, παρατηρείται ασυνήθιστα μεγάλη συγκέντρωση κατόχων μεταπτυχιακού, στοιχείο που κατά πάσα πιθανότητα δεν αντιπροσωπεύει απόλυτα τους κατοίκους των Αθηνών. Συνεπώς, με καλύτερη κατανομή του δείγματος, σε αυτά τα δημογραφικά στοιχεία ίσως επιτυγχάνονταν περαιτέρω αντιπροσωπευτικά αποτελέσματα περιγραφικής στατιστικής. Επιπλέον, όσον αφορά τον σκοπό μετακίνησης, μεγάλο μέρος των μετακινήσεων διεξάγεται για επιστροφή στην κατοικία. Όμως η επιστροφή στην οικία αποτελεί αναπόφευκτη μετακίνηση κατά την διάρκεια μίας

τυπικής ημέρας, συνεπώς δεν εκδηλώνει χωρικά ή χρονικά πρότυπα. Έτσι, η οπτικοποίηση των δεδομένων χωρίς την χρήση αυτού του σκοπού μετακίνησης, ίσως να οδηγούσε σε καλύτερη περιγραφή του δείγματος.

Επιπλέον περιορισμός της έρευνας αποτέλεσε η έλλειψη των πραγματικών αποστάσεων των μετακινήσεων, καθώς χρησιμοποιήθηκαν ζώνες προέλευσης και προορισμού των μετακινήσεων και οι αποστάσεις μετρήθηκαν από τα κέντρα τους. Συνεπώς οι αποστάσεις των μετακινήσεων εκτιμήθηκαν με αυτόν τον τρόπο. Επίσης, κατά την διεξαγωγή της έρευνας δεν χρησιμοποιήθηκαν χωρικά δεδομένα που αφορούν τα μέσα μαζικής μεταφοράς. Δεν διερευνήθηκε η προσβασιμότητα που είχαν οι συμμετέχοντες σε μέσα μαζικής μεταφοράς και σε εναλλακτικούς τρόπους μετακίνησης. Έπειτα, δεν έγινε διαχωρισμός των δημοσιών συγκοινωνιών πέρα από τις κατηγορίες του λεωφορείου και των μέσων σταθερής τροχιάς. Κατ' επέκταση η διερεύνηση έδωσε μία εικόνα του προβλήματος όσον αφορά τα πρότυπα επιλογής μέσου μετακίνησης αλλά δεν τεκμηριώνει ποσοτικά τις πραγματικές ελλείψεις και τα προβλήματα του δικτύου της Δημόσιας Συγκοινωνίας σε διάφορες γειτονίες της Αθήνας. Επίσης, στην έρευνα δεν διερευνήθηκαν καθόλου συνδυασμένες μετακινήσεις, ενώ αποτελούν πραγματικό φαινόμενο. Χρήσιμα θα ήταν και δεδομένα για το πλήθος των αυτοκινήτων ανά νοικοκυριό, καθώς επίσης και δεδομένα που αφορούν το μέσο διεξαγωγής των μετακινήσεων προς το κέντρο της Αθήνας, τις μέρες που δεν επιτρέπεται η πρόσβαση με αυτοκίνητο.

Ένας ακόμα πιθανός περιορισμός είναι η επιρροή της πανδημίας COVID-19, η οποία άλλαξε ριζικά τα πρότυπα των μετακινήσεων. Η πανδημία, οι διαδοχικοί περιορισμοί της κινητικότητας και εμφάνιση της τηλεργασίας αποτελούν φαινόμενα που ανέτρεψαν τα υφιστάμενα πρότυπα μετακίνησης, Βέβαια, η εξ αποστάσεως εργασία, με το πέρασμα των χρόνων, μειώνεται διαρκώς, συνεπώς η χρονική στιγμή που επιλέχθηκε για τη διεξαγωγή της έρευνας ερωτηματολογίου έχει ιδιαίτερη σημασία. Το αποτέλεσμα της έρευνας θα μπορούσε να είναι πολύ διαφορετικό σε 5,10 ή 15 χρόνια μετά την πανδημία.

6.4 Προτάσεις

Με βάση τους περιορισμούς της παρούσας έρευνας, θα μπορούσαν να μελετηθούν ακόμα περισσότερο οι παράγοντες που επηρεάζουν την επιλογή του μέσου μετακίνησης.

Αρχικά, θα μπορούσαν να προστεθούν ακόμα περισσότερες μεταβλητές, οι οποίες θα αφορούν άλλα χαρακτηριστικά της ίδιας της μετακίνησης, όπως το κόστος, η άνεση, η προσβασιμότητα. Επιπλέον, θα μπορούσαν να αξιοποιηθούν κυκλοφοριακά δεδομένα, όπως η ροή και η ταχύτητα κυκλοφορίας. Φυσικά, όπως προαναφέρθηκε, χρήσιμη θα ήταν και η διεξοδική χωρική ανάλυση των μετακινήσεων, με δεδομένα που αφορούν το δίκτυο δημόσιας συγκοινωνίας και το δίκτυο ποδηλατόδρομων. Επιπροσθέτως, χρήσιμη φαίνεται να είναι η αναλυτικότερη διάσπαση των ζωνών της περιοχής μελέτης, ώστε να έχουν μικρότερη έκταση άρα και μεγαλύτερη ακρίβεια στις αποστάσεις. Από το ερωτηματολόγιο εκλείπουν επίσης, ψυχολογικοί και κοινωνικοί παράγοντες.

Επιπλέον όσον αφορά τον παράγοντα της προσβασιμότητας, ένα μεγάλο μέρος των μετακινήσεων που θα έπρεπε να μελετηθεί σε βάθος είναι οι μετακινήσεις των ατόμων με ειδικές ανάγκες. Στην περίπτωση των ατόμων με ειδικές ανάγκες, η προσβασιμότητα παίρνει άλλες διαστάσεις, καθώς οι Δημόσιες Συγκοινωνίες στην πόλη των Αθηνών, απέχουν κατά πολύ από τις εκσυγχρονισμένες που μπορούν να φιλοξενήσουν αυτά τα άτομα και να εξυπηρετήσουν επαρκώς τις ανάγκες τους.

Προτείνεται επίσης, η διεξαγωγή έρευνας με καλύτερα κατανεμημένα τα δημογραφικά στοιχεία, όπως είναι η ηλικία και το μορφωτικό επίπεδο. Το δείγμα θα μπορούσε να γίνει ακόμα πιο αντιπροσωπευτικό του πληθυσμού των κατοίκων της Αθήνας, στοχεύοντας στην συγκέντρωση ακόμα περισσότερων ερωτηματολογίων.

Τέλος, προτείνεται να διερευνηθεί κατά πόσο οι κάτοικοι της Αθήνας που χρησιμοποιούν με μεγάλη συχνότητα το αυτοκίνητο στις καθημερινές τους μετακινήσεις, είναι όντως εξαρτημένοι από αυτό ή αν κάποιος άλλος παράγοντας επηρεάζει την επιλογή τους.

7 ΑΝΑΦΟΡΕΣ

7.1 Διεθνής Βιβλιογραφία

- Allahviranloo M, Regue R, Recker W (2017) Modeling the activity profiles of a population. *Transportmetrica B* 5:431–454. <https://doi.org/10.1080/21680566.2016.1241960>
- Andrew Y. Ng, Michael I. Jordan, Yair Weiss, 2001. On Spectral Clustering: Analysis and an algorithm
- Axhausen, K.W., Horni, A., Nagel, K., Axhausen, K.W., 2016. The Multi-Agent Transport Simulation MATSim. The Multi-Agent Transport Simulation MATSim. Ubiquity Press <https://doi.org/10.5334/baw>.
- Bachir, D., Khodabandelou, G., Gauthier, V., El Yacoubi, M., Puchinger, J., 2019. Inferring dynamic origin-destination flows by transport mode using mobile phone data. *Transp. Res. Part C Emerg. Technol.* 101, 254–275. <https://doi.org/10.1016/j.trc.2019.02.013> .
- Bakker, M.M., Alam, S.J., van Dijk, J., Rounsevell, M.D.A., 2014. Land-use change arising from rural land exchange: an agent-based simulation model. *Landsc. Ecol.* 30, 273–286. <https://doi.org/10.1007/s10980-014-0116-x> .
- Ballis, H., Dimitriou, L., 2019. Optimal Population of Trip Chains Synthesis from Multi-Period Origin-Destination Matrices, in: *Proceedings of Transportation Research Board 98th Annual Meeting*, Washington D.C.
- Ballis H. and Dimitriou L., 2020. “Revealing personal activities schedules from synthesizing multi-period origin-destination matrices”. *Transportation Research Part B: Methodological* (2020) 139 224-258
- Barthelemy, J., Toint, P.L., 2013. Synthetic Population Generation Without a Sample. *Transp. Sci.* 47, 266–279. <https://doi.org/10.1287/trsc.1120.0408> .
- Bastariento FF, Hancock TO, Choudhury CF, Manley E, 2023. Agent-based models in urban transportation: review, challenges, and opportunities. *European Transport Research Review* 15
- Behara KNS, Bhaskar A, Chung E ,2020. A novel approach for the structural comparison of origin-destination matrices: Levenshtein distance. *Transp Res Part C Emerg Technol* 111:513–530. <https://doi.org/10.1016/j.trc.2020.01.005>
- Ben-Akiva M, Bierlaire M ,1999. Discrete Choice Methods and their Applications to Short Term Travel Decisions. In: Hall RW (ed) *Handbook of Transportation Science*. International Series in Operations Research & Management Science. Springer US, Boston
- Ben-Akiva, M., Bottom, J., Gao, S., Koutsopoulos, H.N., Wen, Y., 2007. Towards Disaggregate Dynamic Travel Forecasting Models. *Tsinghua Sci. Technol* 12, 115–130. [https://doi.org/10.1016/S1007-0214\(07\)70019-6](https://doi.org/10.1016/S1007-0214(07)70019-6) .
- Bhat, C.R., 1996. A generalized multiple durations proportional hazard model with an application to activity behavior during the evening work-to-home commute. *Transp. Res. Part B Methodol.* 30, 465–480. [https://doi.org/10.1016/0191-2615\(96\)00070-0](https://doi.org/10.1016/0191-2615(96)00070-0) .
- Bhat, C.R., Koppelman, F.S., 1999. Activity-Based Modeling of Travel Demand 35–61. https://doi.org/10.1007/978-1-4615-5203-1_3
- Bhat, C.R., Srinivasan, S., Axhausen, K.W., 2005. An analysis of multiple interepisode durations using a unifying multivariate hazard model. *Transp. Res. Part B Methodol.* 39, 797–823. <https://doi.org/10.1016/j.trb.2004.11.002> .
- Bonnel, P. , Hombourger, E. , Olteanu-Raimond, A.M. , Smoreda, Z. , 2015. Passive Mobile Phone Dataset to Construct Origin-Destination matrix: Potentials and limitations, in: *Transportation Research Procedia*. Elsevier, pp. 381–398 .
- Bowman, J.L. , 1998. The Day Activity Schedule Approach to Travel Demand Analysis. *Metro*, p. 185.

- Bowman, J.L., Ben-Akiva, M., 2000. Activity-based disaggregate travel demand model system with activity schedules. *Transp. Res. Part A Policy Pract.* 35, 1–28. [https://doi.org/10.1016/S0965-8564\(99\)00043-9](https://doi.org/10.1016/S0965-8564(99)00043-9) .
- Bradley, M., Bowman, J.L., Griesenbeck, B., 2010. SACSIM: an applied activity-based model system with fine-level spatial and temporal resolution. *J. Choice Model.* 3, 5–31. [https://doi.org/10.1016/S1755-5345\(13\)70027-7](https://doi.org/10.1016/S1755-5345(13)70027-7) .
- Caceres, N., Romero, L.M., Benitez, F.G., 2013. Inferring origin-destination trip matrices from aggregate volumes on groups of links: a case study using volumes inferred from mobile phone data. *J. Adv. Transp.* 47, 650–666. <https://doi.org/10.1002/atr.187> .
- Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., Ratti, C., Lorenzo, G., Di, Ferreira, J., Ratti, C., 2013. Understanding individual mobility patterns from urban sensing data - A mobile phone trace example. *Transp. Res. Part C Emerg. Technol.* 26, 301–313. <https://doi.org/10.1016/j.trc.2012.09.009>.
- Cantelmo, G., Qurashi, M., Prakash, A.A., Antoniou, C., Viti, F., 2019. Incorporating trip chaining within online demand estimation. *Transp. Res. Part B Methodol.* <https://doi.org/10.1016/j.trb.2019.05.010> .
- Chatziioannou, T., Tzouras, P.G., Tsigdinos, S. & Kepaptsoglou, K. n.d. *Application of hierarchical plan clustering to identify temporal travel demand patterns in Athens Metropolitan Area*. Available from: <https://forms.gle/L6ABTptKUTfx27Rd9>.
- Chen, C., Bian, L., Ma, J., 2014. From traces to trajectories: how well can we guess activity locations from mobile phone traces? *Transp. Res. Part C Emerg. Technol.* 46, 326–337. <https://doi.org/10.1016/j.trc.2014.07.001> .
- Chen, C., Ma, J., Susilo, Y., Liu, Y., Wang, M., 2016. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transp. Res. Part C Emerg. Technol.* <https://doi.org/10.1016/j.trc.2016.04.005> .
- Cheng, E., Grossman, J.W., Lipman, M.J., 2003. Time-stamped graphs and their associated influence digraphs. *Discret. Appl. Math.* 128, 317–335. [https://doi.org/10.1016/S0166-218X\(02\)00497-3](https://doi.org/10.1016/S0166-218X(02)00497-3) .
- Cich, G., Knapen, L., Maciejewski, M., Yasar, A.U.H., Bellemans, T., Janssens, D., 2017. Modeling Demand Responsive Transport using SARL and MATSim. in: *Procedia Computer Science* 1074–1079. <https://doi.org/10.1016/j.procs.2017.05.387> .
- Çolak, S., Alexander, L.P., Alvim, B.G., Mehndiratta, S.R., Gonzalez, M.C., 2015. Analyzing Cell Phone Location Data for Urban Travel. *Transp. Res. Rec. J. Transp. Res. Board* 2526, 126–135. <https://doi.org/10.3141/2526-14> .
- Department for Transport, 2017. National Travel Survey: england 2016. National Travel Survey .
- Djavadian, S., Chow, J.Y.J., 2017. An agent-based day-to-day adjustment process for modeling 'Mobility as a Service' with a two-sided flexible transport market. *Transp. Res. Part B Methodol.* 104, 36–57. <https://doi.org/10.1016/j.trb.2017.06.015> .
- E. Ramadan, O., Sisiopiku, V., P., 2019. A Critical Review on Population Synthesis for Activity- and Agent-Based Transportation Models, in: *transportation* <https://doi.org/10.5772/intechopen.86307> .
- Eagle, N., Pentland, A.S., 2009. Eigenbehaviors: identifying structure in routine. *Behav. Ecol. Sociobiol.* 63, 1057–1066. <https://doi.org/10.1007/s00265-009-0739-0> .
- Ebadi, N., Kang, J.E., Hasan, S., 2017. Constructing activity–mobility trajectories of college students based on smart card transaction data. *Int. J. Transp. Sci. Technol.* 6, 316–329. <https://doi.org/10.1016/j.ijtst.2017.08.003> .
- Ettema, D., Borgers, A., Timmermans, H., 1995. Competing risk hazard model of activity choice, timing, sequencing, and duration. *Transp. Res. Rec.* 1493, 101–109 .

- Ferreira, A., 2004. Building a reference combinatorial model for MANETs. *IEEE Netw* 18, 24–29. <https://doi.org/10.1109/MNET.2004.1337732> .
- Flötteröd, G., Bierlaire, M., Nagel, K., 2011. Bayesian demand calibration for dynamic traffic simulations. *Transp. Sci.* 45, 541–561. <https://doi.org/10.1287/trsc.1100.0367> .
- Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.-L., 2008. Understanding individual human mobility patterns 453, 779–782. <https://doi.org/10.1038/nature06958> .
- Goodwin, P.B., 1981. The usefulness of travel budgets. *Transp. Res. Part A Gen.* 15, 97–106. [https://doi.org/10.1016/0191-2607\(83\)90019-5](https://doi.org/10.1016/0191-2607(83)90019-5) .
- Goulet Langlois, G., Koutsopoulos, H.N., Zhao, J., 2016. Inferring patterns in the multi-week activity sequences of public transport users. *Transp. Res. Part C Emerg. Technol.* 64, 1–16. <https://doi.org/10.1016/j.trc.2015.12.012> .
- Goulias, K.G. , Kitamura, R. , 1991. Recursive Model System for Trip Generation and Trip Chaining. *Transp. Res. Rec* 59–66 .
- Grace O. Kagho , Milos Balac , Kay W. Axhausen, 2020. Agent-Based Models in Transport Planning: Current State, Issues, and Expectations.
- Han, G., Sohn, K., 2016. Activity imputation for trip-chains elicited from smart-card data using a continuous hidden Markov model. *Transp. Res. Part B Methodol.* 83, 121–135. <https://doi.org/10.1016/j.trb.2015.11.015> .
- Hörl S. and Balac M., 2021. “Introducing the eqasim pipeline: From raw data to agent-based transport simulation”.
- Hörl S. and Balac M., 2021. “Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data”.
- Hörl S, Balac M (2021a) Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data. *Transp Res Part C Emerg Technol* 130:103291. <https://doi.org/10.1016/j.trc.2021.103291>
- Hörl S, Balac M (2021b) Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data. *Transp Res Part C Emerg Technol* 130:.. <https://doi.org/10.1016/j.trc.2021.103291>
- Hart, W.E., Laird, C.D., Watson, J.-P., Woodruff, D.L., Hackebeil, G.A., Nicholson, B.L., Sirola, J.D., 2017. *Pyomo — Optimization Modeling in Python*. <https://doi.org/10.1007/978-3-319-58821-6>
- IBM, 2020. CPLEX Optimizer | IBM [WWW Document]. URL <https://www.ibm.com/analytics/cplex-optimizer> (accessed 7.28.18).
- Iqbal, M.S., Choudhury, C.F., Wang, P., Gonzalez, M.C., 2014. Development of origin-destination matrices using mobile phone call data. *Transp. Res. Part C Emerg. Technol.* 40, 63–74. <https://doi.org/10.1016/j.trc.2014.01.002> .
- Jiang, S., Ferreira, J., González, M.C., 2012. Clustering daily patterns of human activities in the city, in: *Data Min Knowl Discov* 478–510. <https://doi.org/10.1007/s10618-012-0264-z> .
- Jiang, S., Yang, Y., Gupta, S., Veneziano, D., Athavale, S., González, M.C., 2016. The TimeGeo modeling framework for urban motility without travel surveys, in: *proceedings of the National Academy of Sciences of the United States of America*. pp. E5370–E5378. <https://doi.org/10.1073/pnas.1524261113>
- Joh, C.H., Arentze, T., Hofman, F., Timmermans, H., 2002. Activity pattern similarity: a multidimensional sequence alignment method. *Transp. Res. Part B Methodol.* 36, 385–403. [https://doi.org/10.1016/S0191-2615\(01\)00091-1](https://doi.org/10.1016/S0191-2615(01)00091-1) .

- Jun, C., Dongyuan, Y., 2013. Estimating smart card commuters origin-destination distribution based on APTS data. *J. Transp. Syst. Eng. Inf. Technol.* 13, 47–53. [https://doi.org/10.1016/S1570-6672\(13\)60116-6](https://doi.org/10.1016/S1570-6672(13)60116-6) .
- Konduri, K.C., You, D., Garikapati, V.M., Pendyala, R.M., 2016. Enhanced synthetic population generator that accommodates control variables at multiple geographic resolutions. *Transp. Res. Rec.* 2563, 40–50. <https://doi.org/10.3141/2563-08>
- Kostakos, V., 2009. Temporal graphs. *Phys. A Stat. Mech. its Appl* 388, 1007–1023. <https://doi.org/10.1016/j.physa.2008.11.021>
- Kumar, R., Calders, T., 2018. 2SCENT: an Efficient algorithm for enumerating all simple temporal cycles. *Proc. VLDB Endow.* 11, 1441–1453. <https://doi.org/10.14778/3236187.3236197>
- Lee, Y., Hickman, M., Washington, S., 2007. Household type and structure, time-use pattern, and trip-chaining behavior. *Transp. Res. Part A Policy Pract.* 41, 1004–1020. <https://doi.org/10.1016/j.tra.2007.06.007> .
- Lin, X., Sun, W., Veeraraghavan, M., Hu, W., 2016. Time-shifted multilayer graph: a routing framework for bulk data transfer in optical circuit-switched networks with assistive storage. *J. Opt. Commun. Netw.* 8, 162–174. <https://doi.org/10.1364/JOCN.8.000162> .
- Lin, Z. , Yin, M. , Feygin, S. , Sheehan, M. , Paiement, J.-F. , Pozdnoukhov, A. , 2017. Deep Generative Models of Urban Mobility. *ACM SIGKDD Conf.* 9 <https://doi.org/10.475/123> .
- Lindveld, C.D.R., 2003. Dynamic O-D matrix estimation: a behavioural approach.
- Liu, F., Janssens, D., Cui, J., Wang, Y., Wets, G., Cools, M., 2014. Building a validation measure for activity-based transportation models based on mobile phone data. *Expert Syst. Appl.* 41, 6174–6189. <https://doi.org/10.1016/j.eswa.2014.03.054> .
- Liu, F., Janssens, D., Cui, J., Wets, G., Cools, M., 2015. Characterizing activity sequences using profile Hidden Markov Models. *Expert Syst. Appl* 42, 5705–5722. <https://doi.org/10.1016/j.eswa.2015.02.057> .
- Liu, F., Janssens, D., Wets, G., Cools, M., 2013. Annotating mobile phone location data with activity purposes using machine learning algorithms. *Expert Syst. Appl.* 40, 3299–3311. <https://doi.org/10.1016/j.eswa.2012.12.100> .
- Ma, J., Li, H., Yuan, F., Bauer, T., 2013. Deriving Operational Origin-Destination Matrices From Large Scale Mobile Phone Data. *Int. J. Transp. Sci. Technol.* 2, 183–204. <https://doi.org/10.1260/2046-0430.2.3.183> .
- Ma, L., Srinivasan, S., 2016. An empirical assessment of factors affecting the accuracy of target-year synthetic populations. *Transp. Res. Part A Policy Pract.* 85, 247–264. <https://doi.org/10.1016/j.tra.2016.01.016> .
- Ma, L., Srinivasan, S., 2015. Synthetic population generation with multilevel controls: a fitness-based synthesis approach and validations. *Comput. Civ. Infrastruct. Eng.* 30, 135–150. <https://doi.org/10.1111/mice.12085> .
- McBride, E.C., Davis, A.W., Goulias, K.G., 2018. A Spatial Latent Profile Analysis to Classify Land Uses for Population Synthesis Methods in Travel Demand. Forecasting. *Transp. Res. Rec.* 2672, 158–170. <https://doi.org/10.1177/0361198118799168> .
- McGuckin, N., Murakami, E., 1999. Examining Trip-Chaining Behavior: comparison of Travel by Men and Women. *Transp. Res. Rec. J. Transp. Res. Board* 1693, 79–85. <https://doi.org/10.3141/1693-12>
- Morimura, T., Osogami, T., 2013. Solving inverse problem of Markov chain with partial observations.
- Ni, B., Shen, Q., Xu, J., Qu, H., 2017. Spatio-temporal flow maps for visualizing movement and contact patterns. *Vis. Informatics* 1, 57–64. <https://doi.org/10.1016/j.visinf.2017.01.007> .

- Nie, Y., Zhang, H.M., Recker, W.W., 2005. Inferring origin-destination trip matrices with a decoupled GLS path flow estimator. *Transp. Res. Part B Methodol.* 39, 497–518. <https://doi.org/10.1016/j.trb.2004.07.002> .
- Nurul Habib, K.M., 2011. A random utility maximization (RUM) based dynamic activity scheduling model: application in weekend activity scheduling. *Transportation (Amst)* 38, 123–151. <https://doi.org/10.1007/s11116-010-9294-9> .
- Ortúzar, J., de, D., Willumsen, L.G., 2011. *Modelling Transport*. Modelling Transport, 4th ed Wiley-Blackwell <https://doi.org/10.1002/9781119993308>
- Ozimek, A., Miles, D., 2011. Stata utilities for geocoding and generating travel time and travel distance information. *Stata J* 11, 106–119. <https://doi.org/10.1177/1536867x1101100107>
- Pan, C., Lu, J., Di, S., Ran, B., 2006. Cellular-based data-extracting method for trip distribution. in: *Transportation Research Record* 33–39. <https://doi.org/10.3141/1945-04>
- Pappalardo, L., Simini, F., 2018. Data-driven generation of spatio-temporal routines in human mobility. *Data Min. Knowl. Discov.* 32, 787–829. <https://doi.org/10.1007/s10618-017-0548-4> .
- Pendyala, R.M. Goulias, K.G. , 2002. Time use and activity perspectives in travel behavior research. *Transportation (Amst)* 29, 1–4 .
- Pengxiang Zhao, Dominik Bucher, Henry Martin and Martin Raubal, 2020. A Clustering-Based Framework for Understanding Individuals' Travel Mode Choice Behavior. https://doi.org/10.1007/978-3-030-14745-7_5
- Phan, D., Xiao, L., Yeh, R., Hanrahan, P., Winograd, T., 2005. Flow map layout. *Proc. - IEEE Symp. Inf. Vis. INFO VIS* 219–224. <https://doi.org/10.1109/INFVIS.2005.1532150> .
- Phithakkitnukoon, S., Horanont, T., Di Lorenzo, G., Shibasaki, R., Ratti, C., 2010. Activity-aware map: identifying human daily activity pattern using mobile phone data, in: *lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp. 14–25. https://doi.org/10.1007/978-3-642-14715-9_3
- Pinjari, A.R. , Bhat, C.R. , 2011. Activity-based Travel Demand Analysis. In: de Palma, A., Lindsey, R., Quinet, E. (Eds.), *A Handbook of Transport Economics*. Edward Elgar Publishing Ltd, pp. 213–248 <https://doi.org/10.4337/9780857930873.00017> .
- Primerano, F., Taylor, M.A.P., Pitaksringkarn, L., Tisato, P., 2008. Defining and understanding trip chaining behaviour. *Transportation (Amst)*. <https://doi.org/10.1007/s11116-007-9134-8> .
- Rasouli S, Timmermans H (2014) Applications of theories and models of choice and decision-making under conditions of uncertainty in travel behavior research. *Travel Behav Soc* 1:79–90. <https://doi.org/10.1016/j.tbs.2013.12.001>
- Raux, C., Ma, T.Y., Cornelis, E., 2016. Variability in daily activity-travel patterns: the case of a one-week travel diary. *Eur. Transp. Res. Rev* 8. <https://doi.org/10.1007/s12544-016-0213-9> .
- Recker, W.W., 2001. A bridge between travel demand modeling and activity-based travel analysis. *Transp. Res. Part B Methodol.* 35, 481–506. [https://doi.org/10.1016/S0191-2615\(00\)00060-0](https://doi.org/10.1016/S0191-2615(00)00060-0) .
- Redondo, J.L., Pelegrin, B., Fernandez, P., Garcia, I., Ortigosa, P.M., 2011. Finding multiple global optima for unconstrained discrete location problems. *Optim. Methods Softw.* 26, 207–224. <https://doi.org/10.1080/10556780903567760> .
- Ronald, N., Thompson, R., Winter, S., 2015. Simulating Demand-responsive Transportation: a Review of Agent-based Approaches. *Transp. Rev.* 35, 404–421. <https://doi.org/10.1080/01441647.2015.1017749> .
- Saadi, I., Mustafa, A., Teller, J., Cools, M., 2016. Forecasting travel behavior using Markov Chains-based approaches. *Transp. Res. Part C Emerg. Technol* 69, 402–417. <https://doi.org/10.1016/j.trc.2016.06.020> .

- Santoro, N. , Quattrociochi, W. , Flocchini, P. , Casteigts, A. , Amblard, F. , 2011. Time-varying graphs and social network analysis: temporal indicators and metrics, in: AISB 2011. Social Networks and Multiagent Systems 33–38 .
- Schneider, C.M., Belik, V., Couronné, T., Smoreda, Z., González, M.C., 2013. Unravelling daily human mobility motifs. *J. R. Soc. Interface* 10, 20130246. <https://doi.org/10.1098/rsif.2013.0246> .
- Schneider, F., Ton, D., Zomer, L.-B., Daamen, W., Duives, D., Hoogendoorn-Lanser, S., Hoogendoorn, S., 2020. Trip chain complexity: a comparison among latent classes of daily mobility patterns. *Transportation (Amst)* 1–23. <https://doi.org/10.1007/s11116-020-10084-1> .
- Schoenfelder, S., Axhausen, K.W., 2001. Analysing the rhythms of travel using survival analysis, in: transportation Research Board (TRB) Annual Meeting. <https://doi.org/10.3929/ETHZ-A-004241369>
- Sedgewick, R., 2001. Algorithms in C. Addison-Wesley Professional [https://doi.org/10.1016/0965-9978\(92\)90046-i](https://doi.org/10.1016/0965-9978(92)90046-i) .
- Tesselkin, A., Khabarov, V., 2017. Estimation of Origin-Destination Matrices Based on Markov Chains. *Procedia Eng* 178, 107–116. <https://doi.org/10.1016/j.proeng.2017.01.071> .
- Thill, J.-C, Thomas, I., 1987. Toward Conceptualizing Trip-Chaining Behavior: a Review. *Geogr. Anal.* 19, 1–17. <https://doi.org/10.1111/j.1538-4632.1987.tb00110.x> .
- Tolouei, R. , Psarras, S. , Prince, R. , 2017. Origin-Destination Trip Matrix Development: Conventional Methods Versus Mobile Phone Data, in: Transportation Research Procedia. Elsevier, pp. 39–52 .
- Toole, J.L., Colak, S., Sturt, B., Alexander, L.P., Evsukoff, A., Gonzalez, M.C., 2015. The path most traveled: travel demand estimation using big data resources. *Transp. Res. Part C Emerg. Technol.* 58, 162–177. <https://doi.org/10.1016/j.trc.2015.04.022> .
- Vlahogianni, E.I., Park, B.B., van Lint, J.W.C.W.C., 2015. Big data in transportation and traffic engineering. *Transp. Res. Part C Emerg. Technol.* 58, 161. <https://doi.org/10.1016/j.trc.2015.08.006>
- Von Landesberger, T., Brodkorb, F., Roskosch, P., Andrienko, N., Andrienko, G., Kerren, A., 2016. MobilityGraphs: visual Analysis of Mass Mobility Dynamics via Spatio-Temporal Graphs and Clustering. *IEEE Trans. Vis. Comput. Graph.* 22, 11–20. <https://doi.org/10.1109/TVCG.2015.2468111>
- Von Luxburg U. ,2007. A Tutorial on Spectral Clustering.
- Wang, W., Attanucci, J.P., Wilson, N.H.M.M., 2011. Bus Passenger Origin-Destination Estimation and Related Analyses Using Automated Data Collection Systems. *J. Public Transp.* 14, 131–150. <https://doi.org/10.5038/2375-0901.14.4.7> .
- Wang, Y., Yuan, Y., Ma, Y., Wang, G., 2019. Time-Dependent Graphs: definitions, Applications, and Algorithms. *Data Sci. Eng.* <https://doi.org/10.1007/s41019-019-00105-0> .
- Wehmuth, K. , Ziviani, A. , Fleury, E. , 2015. A Unifying Model For Representing Time-Varying graphs, in: Proceedings of the 2015 IEEE International Conference On Data Science and Advanced Analytics. DSAA 2015. Institute of Electrical and Electronics Engineers Inc .
- Wilson, W.C., 1998. Activity pattern analysis by means of sequence-alignment methods. *Environ. Plan. A* 30, 1017–1038. <https://doi.org/10.1068/a301017> .
- Wood, J., Dykes, J., Slingsby, A., 2010. Visualisation of Origins, Destinations and Flows with OD Maps. *Cartogr. J* 47, 117–129. <https://doi.org/10.1179/000870410x12658023467367> .
- Ye, P., Zhu, F., Sabri, S., Wang, F.Y., 2020. Consistent Population Synthesis with Multi-Social Relationships Based on Tensor Decomposition. *IEEE Trans. Intell. Transp. Syst.* 21, 2180–2189. <https://doi.org/10.1109/TITS.2019.2916867> .
- Yue, Y., Lan, T., Yeh, A.G.O., Li, Q.Q., 2014. Zooming into individuals to understand the collective: a review of trajectory-based travel behaviour studies. *Travel Behav. Soc.* 1, 69–78. <https://doi.org/10.1016/j.tbs.2013.12.002> .

Zhang, W., Thill, J.C., 2017. Detecting and visualizing cohesive activity-travel patterns: a network analysis approach. *Comput. Environ. Urban Syst.* 66, 117-129. <https://doi.org/10.1016/j.compenvurbsys.2017.08.004> .

Zhao, J., Rahbee, A., Wilson, N.H.M., 2007. Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. *Comput. Civ. Infrastruct. Eng.* 22, 376–387. <https://doi.org/10.1111/j.1467-8667.2007.00494.x> .

Zilske, M., Nagel, K., 2015. A simulation-based approach for constructing all-day travel chains from mobile phone data. in: *Procedia Computer Science* 468–475. <https://doi.org/10.1016/j.procs.2015.05.017> .

Ziemke D., Kaddoura I. and Nagel K., 2019. “The MATSim Open Berlin Scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data”.

7.2 Ελληνική Βιβλιογραφία

Σενικίδου Ν., 2019. “Διερεύνηση της συσχέτισης ατομικών χαρακτηριστικών και επιλογής μέσου μεταφοράς”.

Τριανταφυλλίδη Ε., 2021. “ΔΙΕΡΕΥΝΗΣΗ ΠΑΡΑΜΕΤΡΩΝ ΠΡΟΤΙΜΗΣΗΣ ΤΗΣ ΧΡΗΣΗΣ ΚΟΙΝΟΧΡΗΣΤΩΝ ΑΥΤΟΝΟΜΩΝ ΟΧΗΜΑΤΩΝ”.

7.3 Ιστοσελίδες

https://cs.brown.edu/courses/csci1820/notes/CH5_Notes.pdf

<https://eranraviv.com/understanding-spectral-clustering/>

<https://www.r-bloggers.com/2019/10/fast-adaptive-spectral-clustering-in-r-brain-cancer-rna-seq/>

<https://scalefreegan.github.io/Teaching/DataIntegration/practicals/p2.html#spectral-clustering-full>

https://mhahsler.github.io/Introduction_to_Data_Mining_R_Examples/book/clustering-analysis.html

<https://www.analyticsvidhya.com/blog/2021/05/what-why-and-how-of-spectral-clustering/>

<https://calculatedcontent.com/2012/10/09/spectral-clustering/>

<https://rpubs.com/gargeejagtap/SpectralClustering>

<https://www.di.fc.ul.pt/~jpn/r/spectralclustering/spectralclustering.html>

https://www.di.ens.fr/~fbach/nips03_cluster.pdf

<https://rpubs.com/nurakawa/spectral-clustering>