



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ  
ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

## Ανάλυση της επιρροής της τιμής της βενζίνης στις κυκλοφοριακές συνθήκες της Αθήνας

Στυλιανός Πειθής  
Επιβλέπων: Γιώργος Γιαννής, Καθηγητής Ε.Μ.Π  
Αθήνα, Νοέμβριος 2022

## Περιεχόμενα

Περίληψη .....	9
1. Εισαγωγή.....	13
1.2 Στόχος Διπλωματικής εργασίας .....	16
1.3 Δομή διπλωματικής εργασίας .....	17
2. Βιβλιογραφική ανασκόπηση.....	17
2.1 Εισαγωγή.....	17
2.2 Συναφείς έρευνες και μεθοδολογίες.....	18
2.3 Συμπεράσματα βιβλιογραφίας.....	21
3. Θεωρητικό Υπόβαθρο.....	23
3.1 Εισαγωγή.....	23
3.2 Βασικές έννοιες στατιστικής.....	23
3.3 Συντελεστής συσχέτισης μεταβλητών .....	25
3.4 Κανονική Κατανομή .....	26
3.5 Μαθηματικά Πρότυπα.....	26
3.5.1 Γραμμική Παλινδρόμηση .....	26
3.5.2 Λογαριθμοκανονική Παλινδρόμηση.....	28
3.5.3 Γενικό γραμμικό μοντέλο.....	28
3.5.4 Γραμμικό μικτό μοντέλο (Linear mixed model) .....	29
3.5.5 Συντελεστής προσδιορισμού (Coefficient of determination).....	30
4. Συλλογή και επεξεργασία στοιχείων .....	31
4.1 Εισαγωγή.....	31
4.2 Συλλογή δεδομένων .....	31
4.2.1 Δεδομένα Κυκλοφοριακού φόρτου.....	31
4.2.2 Δεδομένα για την τιμή της βενζίνης.....	36
4.2.3 Δεδομένα θερμοκρασίας και βροχόπτωσης .....	36
4.2.4 Δεδομένα ενεργών κρουσμάτων Covid-19.....	36
4.3 Επεξεργασία δεδομένων .....	36
4.4 Περιγραφική στατιστική .....	38
5. Ανάπτυξη μαθηματικών προτύπων.....	42
5.1 Εισαγωγή.....	42
5.1.2 Λογισμικό R και βιβλιοθήκες .....	42

5.1.3 Εισαγωγή βάσης δεδομένων στο προγραμματιστικό περιβάλλον R .....	43
5.1.4 Συσχέτιση μεταβλητών (Pearson Correlation) .....	44
5.2 Εφαρμογή στατιστικών μοντέλων .....	46
Μοντέλο A: Επιρροή της τιμής της βενζίνης στην κυκλοφορία της ευρύτερης περιοχής της Αθήνας ..	47
Μαθηματική σχέση Μοντέλου A.....	48
Ποιότητα Μοντέλου A .....	48
Σχολιασμός αποτελεσμάτων Μοντέλου A.....	48
Σχετική επιρροή Μοντέλου A .....	49
Ανάλυση Ευαισθησίας.....	51
Εφαρμογή μοντέλου για τη βάση δεδομένων με μετρητές εκτός Δακτυλίου .....	54
Μαθηματική σχέση μοντέλου για μετρητές εκτός Δακτυλίου.....	55
Ποιότητα μοντέλου για μετρητές εκτός Δακτυλίου .....	55
Σχολιασμός αποτελεσμάτων μοντέλου για μετρητές εκτός Δακτυλίου .....	55
Σχετική επιρροή των ανεξάρτητων μεταβλητών για μετρητές εκτός Δακτυλίου .....	56
Ανάλυση ευαισθησίας των ανεξάρτητων μεταβλητών για μετρητές εκτός Δακτυλίου .....	56
Συνολικός σχολιασμός των διαγραμμάτων .....	58
Εφαρμογή μοντέλου για βάση δεδομένων με μετρητές εντός του Δακτυλίου .....	59
Μαθηματική σχέση μοντέλου για μετρητές εντός Δακτυλίου.....	59
Ποιότητα μοντέλου για μετρητές εντός Δακτυλίου .....	60
Σχολιασμός αποτελεσμάτων μοντέλου για μετρητές εντός του Δακτυλίου .....	60
Σχετική επιρροή των ανεξάρτητων μεταβλητών για μετρητές εντός Δακτυλίου .....	60
6. Συμπεράσματα.....	62
6.1 Σύνοψη αποτελεσμάτων .....	62
6.2 Συμπεράσματα.....	63
6.3 Προτάσεις .....	65
6.4 Προτάσεις για περαιτέρω έρευνα .....	65
Βιβλιογραφία .....	67



## Ευχαριστίες

Με την ολοκλήρωση της παρούσας Διπλωματικής Εργασίας ολοκληρώνεται ο κύκλος των σπουδών μου στη Σχολή Πολιτικών Μηχανικών του Εθνικού Μετσοβίου Πολυτεχνείου.

Θα ήθελα πρωτίστως να ευχαριστήσω τον κ.Γ.Γιαννή, Καθηγητή της Σχολής Πολιτικών Μηχανικών ΕΜΠ, για την εμπιστοσύνη που μου έδειξε με την ανάθεση του ερευνητικού αντικειμένου, καθώς επίσης και για την πολύ σημαντική υποστήριξη και καθοδήγηση σε όλα τα στάδια της εκπόνησης της Διπλωματικής Εργασίας.

Επίσης θα ήθελα να ευχαριστήσω την υποψήφια διδάκτορα Βιργινία Πετράκη για την καθοριστική της συνεισφορά της στη διεκπεραίωση της παρούσας εργασίας, μέσω των πολύτιμων συμβουλών και υποδείξεων της, την υποστήριξη της και το εξαιρετικό κλίμα συνεργασίας και επικοινωνίας που διαμόρφωσε.

Τέλος θα ήθελα να ευχαριστήσω μέσα από την καρδιά μου κυρίως την οικογένεια μου που με στήριξε καθ' όλη τη διάρκεια των σπουδών μου και τους φίλους μου.

Αθήνα, Νοέμβριος 2022  
Στυλιανός Πειθής

# Ανάλυση της επιρροής της τιμής της βενζίνης στις κυκλοφοριακές συνθήκες της Αθήνας

Στυλιανός Πειθής

Επιβλέπων: Γιώργος Γιαννής, Καθηγητής Ε.Μ.Π.

## Σύνοψη

Στόχος της συγκεκριμένης Διπλωματικής Εργασίας αποτελεί η ανάλυση της επιρροής της τιμής της βενζίνης στον κυκλοφοριακό φόρτο της Αθήνας. Για τον σκοπό αυτό συλλέχθηκαν δεδομένα σχετικά με τις κυκλοφοριακές συνθήκες της πόλης, τιμολόγησης της βενζίνης καθώς και άλλων παραμέτρων οι οποίες επηρεάζουν τον κυκλοφοριακό φόρτο όπως είναι τα κλιματικά δεδομένα. Η ανάλυση έγινε με χρήση γραμμικών μικτών μοντέλων (Linear Mixed Models) με πλήθος ανεξάρτητων μεταβλητών που ερμηνεύουν τις μεταβολές του κυκλοφοριακού φόρτου. Από την εφαρμογή της μεθοδολογίας για τη χρονική περίοδο Νοέμβριος 2021-Ιούλιος 2022 προέκυψαν τρία στατιστικά μοντέλα με κοινές μεταβλητές για διαφορετικές βάσεις δεδομένων, ανάλογα με την τοποθεσία των μετρητών κυκλοφοριακού φόρτου. Η αυξημένη τιμή της βενζίνης έχει αρνητική επίδραση στο μέγεθος του κυκλοφοριακού φόρτου ενώ σημαντική επίδραση έχουν και οι παράγοντες του κλίματος όπως η βροχόπτωση, η θερμοκρασία αλλά και η χρονική περίοδος δηλαδή αργίες, σαββατοκύριακα και σχολικές διακοπές. Ύστερα από την ανάλυση ελαστικότητας για τις παραπάνω ανεξάρτητες μεταβλητές προκύπτουν διαφορές ανάμεσα στην επιρροή αυτών για τον κυκλοφοριακό φόρτο εκτός και εντός του Δακτυλίου της Αθήνας.

Λέξεις κλειδιά : Κόστος μετακίνησης, ακρίβεια , κυκλοφοριακός φόρτος , κλιματικά δεδομένα , γραμμικά μικτά μοντέλα παλινδρόμησης, ανάλυση ελαστικότητας, ανάλυση ευαισθησίας , βενζίνη, επαγωγικοί βρόγχοι

## Impact analysis of gasoline price on traffic conditions in Athens

Stylios Peithis

Supervisor: George Yannis, Professor NTUA

### **Abstract**

The objective of this thesis is the impact analysis of gasoline price on traffic conditions in Athens. For this purpose, data concerning traffic conditions, gasoline prices and a wide variety of traffic shaping factors such as weather data was collected. In order to accomplish this thesis' objective Linear Mixed Effects models with random effect intercept were deployed using as independent variables various metrics with explanative power about traffic volume's fluctuations. During the application of the methods mentioned above three statistical models were constructed for separate databases, depending on the geographical position of the traffic flow measuring devices but with the same independent variables across all models. The rise of gasoline prices has a negative impact on traffic volume while the impact of other factors such weather conditions like temperature and rainfall are also substantial. Schools' vacations, weekends and public holidays also have a decreasing impact in traffic conditions. After conducting an elasticity estimation for the independent variables mentioned, differences in the weight of impact between the traffic zones.

Key word : Trip costs, expensiveness, traffic volume , weather data, linear mixed models regression, elasticity estimation , sensitivity estimation, gasoline , measuring devices, conductive loops





## Περίληψη

Αντικείμενο της διπλωματικής εργασίας αποτέλεσε η **ανάλυση της επιρροής της τιμής της βενζίνης στις κυκλοφοριακές συνθήκες της Αθήνας** και συγκεκριμένα στον κυκλοφοριακό φόρτο της πόλης.

Ύστερα από τον καθορισμό του στόχου της διπλωματικής εργασίας, ακολούθησε η **βιβλιογραφική ανασκόπηση** ώστε να μελετηθούν υφιστάμενες έρευνες και μεθοδολογίες με συναφή αντικείμενο. Στη συνέχεια γίνεται η **συλλογή** και **επεξεργασία** όλων των απαραίτητων δεδομένων από διάφορες πηγές συνθέτοντας τις βάσεις με τις οποίες θα εκπονηθεί η στατιστική ανάλυση. Εν συνεχεία γίνεται η επιλογή της κατάλληλης **μεθοδολογίας**, η εφαρμογή των μαθηματικών προτύπων που αυτή περιλαμβάνει, ελέγχεται η στατιστική σημαντικότητα των μοντέλων και ερμηνεύονται τα αποτελέσματα. Η στατιστική ανάλυση πραγματοποιήθηκε με χρήση του **Γραμμικού Μικτού Μοντέλου** (Linear Mixed Effects Model). Ως εξαρτημένη μεταβλητή ορίστηκε ο **ημερήσιος κυκλοφοριακός φόρτος** με μονάδες διελύσεις οχημάτων ανά ημέρα και ως ανεξάρτητες μεταβλητές ορίστηκαν η διαφορά της ημερήσιας τιμής της **βενζίνης** από τον μέσο όρο των επτά προηγούμενων ημερών με μονάδες ευρώ ανά λίτρο, η θερμοκρασία ημέρας σε βαθμούς κελσίου και μια σειρά από ανεξάρτητες ποιοτικές μεταβλητές (dummy variables) οι οποίες αποτελούν σημαντικούς παράγοντες διαμόρφωσης των κυκλοφοριακών συνθηκών. Αυτές είναι η ημέρα της εβδομάδας, αν δηλαδή πρόκειται για εργάσιμες ή σαββατοκύριακα, οι αργίες και οι σχολικές διακοπές. Η εφαρμογή ενιαίου μοντέλου γίνεται σε τρεις διαφορετικές **βάσεις δεδομένων**, η πρώτη αφορά όλους τους μετρητές για κάθε οδό του δείγματος συνολικά, η δεύτερη μόνο τους μετρητές εκτός δακτυλίου ενώ η τελευταία αφορά μόνο τους μετρητές εντός του δακτυλίου.

Σημειώνεται πώς η επιλογή του στατιστικού μοντέλου προκύπτει ύστερα από αρκετές δοκιμές συνδυασμών ανεξάρτητων μεταβλητών και εξαρτημένων μεταβλητών ενώ αρκετές ήταν και οι δοκιμές σχετικά με τον παράγοντα που θα καθορίσει την τυχαία επίδραση του σταθερού όρου (random intercept). Η ομαδοποίηση έγινε δηλαδή με βέλτιστο τρόπο προκειμένου οι ομάδες των παρατηρήσεων να περιέχουν μετρητές με κοινά χαρακτηριστικά κυκλοφοριακών συνθηκών μειώνοντας έτσι τις διαφορές κυκλοφοριακών συνθηκών που μετράει ο κάθε μετρητής.

Συνολική βάση δεδομένων					Βάση δεδομένων μετρητές εκτός δακτυλίου								Βάση δεδομένων μετρητές εντός δακτυλίου					
	Fixed Effects				Σχετική Επιρροή		Fixed Effects				Σχετική Επιρροή		Fixed Effects				Σχετική επιρροή	
	Estimate	Std. Error	t-value	p-value	e <sub>i</sub>	e <sub>i</sub> <sup>*</sup>	Estimate	Std. Error	t-value	p-value	e <sub>i</sub>	e <sub>i</sub> <sup>*</sup>	Estimate	Std. Error	t-value	p-value	e <sub>i</sub>	e <sub>i</sub> <sup>*</sup>
Intercept	4.442	0.0169	262.62	0.000			4.496	0.0247	181.72	0.000			4.372	0.0199	219.49	0.000		
Weekend1	-0.059	0.0008	-71.00	0.000	-0.1276	3.376	-0.077	0.0011	-69.00	0.000	-0.1626	3.651	-0.037	0.0012	-29.32	0.000	-0.0807	2.803
Public_Holidays1	-0.132	0.0020	-65.31	0.000	-0.2623	6.940	-0.137	0.0027	-50.43	0.000	-0.2697	6.055	-0.127	0.0030	-42.05	0.000	-0.2534	8.799
Schools_Vacations1	-0.017	0.0011	-15.88	0.000	-0.0378	1.000	-0.020	0.0014	-14.04	0.000	-0.0445	1.000	-0.013	0.0016	-8.07	0.000	-0.0288	1.000
Termperature	0.001	0.0001	10.59	0.000	0.0003	-4.852	0.001	0.0001	9.69	0.000	0.0004	-5.947	0.0004	0.0001	4.97	0.000	0.0002	-3.697
GasDiff_Today-Weekbf	-0.333	0.0178	-18.73	0.000	-0.0001	1.000	-0.354	0.0238	-14.89	0.000	-0.0001	1.000	-0.306	0.0265	-11.53	0.000	-0.0001	1.000
Random Effects					Random Effects								Random Effects					
Groups	Variance	Std.Dev	R <sup>2</sup>		0.91		Variance	Std.Dev	R <sup>2</sup>		0.92		Variance	Std.Dev	R <sup>2</sup>		0.86	
id	0.04499	0.21211					0.054251	0.23292					0.02637	0.1624				
Residual	0.004613	0.06792					0.004654	0.06822					0.00436	0.06603				

Σύνοψη 1 : Αποτελέσματα στατιστικών μοντέλων

Βάσεις των αποτελεσμάτων που προέκυψαν κατά την εφαρμογή της μεθοδολογίας, προκύπτουν ορισμένα συμπεράσματα άμεσα σχετιζόμενα με τον στόχος της διπλωματικής εργασίας.

1. Για το ερευνητικό αντικείμενο της συγκεκριμένης διπλωματικής εργασίας παρατηρείται **ενιαία** αρνητική επιρροή της τιμής της **βενζίνης** στον λογάριθμο του κυκλοφοριακού φόρτου. Σε κάθε βάση δεδομένων η διαφορά της ημερήσιας τιμής της βενζίνης από το μέσο όρο των προηγούμενων επτά ημερών, δηλαδή η ακρίβεια στο κόστος ανεφοδιασμού μειώνει τον κυκλοφοριακό φόρτο, άρα το κόστος μετακινήσεις και πιο συγκεκριμένα η τιμή της βενζίνης όσο αυξάνει ο λογάριθμος του κυκλοφοριακού φόρτου μειώνει. Το παραπάνω ισχύει και αντίστροφα, όσο δηλαδή η τιμή της βενζίνης μειώνεται άρα η τιμή της βενζίνης είναι μειωμένη συγκριτικά με το μέσο όρο των επτά προηγούμενων ημερών τότε ο λογάριθμος του κυκλοφοριακού φόρτου αυξάνει.
2. Εκτός του δακτυλίου η τιμή της βενζίνης έχει ελαφρώς μεγαλύτερη σχετική επιρροή στον κυκλοφοριακό φόρτο. Η διαφορά αυτή ενδεχομένως να οφείλεται στις κυκλοφοριακές συνθήκες που επικρατούν στις οδούς εντός του δακτυλίου. Η κυκλοφορία σε οδούς στο κέντρο της πόλης χαρακτηρίζεται συχνά από συμφορημένη ροή και η πυκνότητα κυκλοφορίας έχει φτάσει στο μέγιστο. Τότε ακόμα και αν η ακρίβεια της βενζίνης μειώνει τις μετακινήσεις η μεταβολή δεν μετακυλά στον κυκλοφοριακό φόρτο καθώς στις οδούς αυτούς στην πράξη ο κυκλοφοριακός φόρτος ορισμένες φορές αυξάνει μια και από την περιοχή συμφορημένης ροής η οδός περνά στην περιοχή μη συμφορημένης ροής.
3. Ύστερα από αρκετές δοκιμές σχετικά με την ανεξάρτητη μεταβλητή που θα εκφράσει το κόστος της βενζίνης είναι φανερό πώς οι οδηγοί αντιλαμβάνονται και μειώνουν τις μετακινήσεις με ιδιωτικά οχήματα ύστερα από μία χρονική περίοδο και όχι άμεσα. Αυτό συμβαίνει καθώς ο ανεφοδιασμός των οχημάτων τους δεν προκύπτει καθημερινά

ενώ παράλληλα από τη στιγμή που παρατηρούν την αύξηση, χρειάζονται και κάποιο χρόνο για να τροποποιήσουν το πρόγραμμα μετακινήσεων τους.

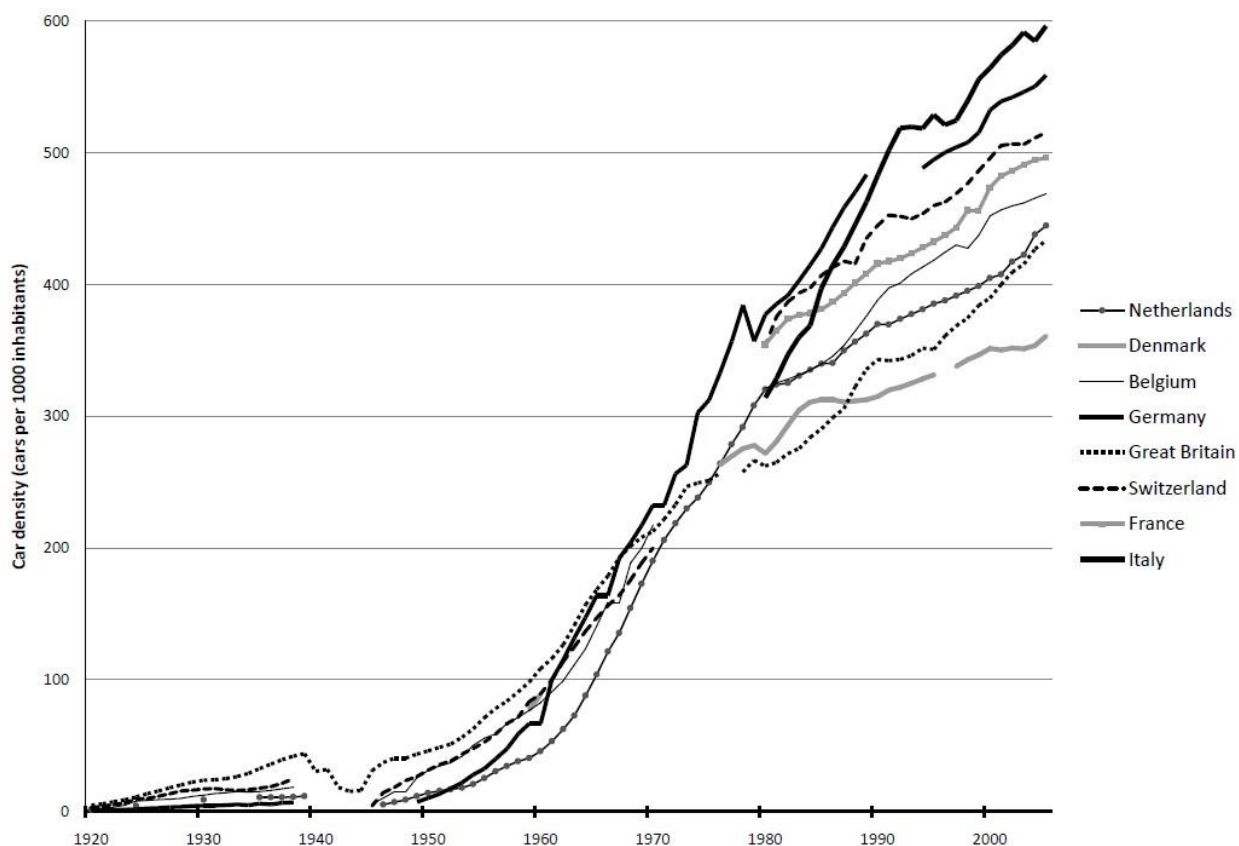
4. Η άνοδος της θερμοκρασίας σε όλα τα στατιστικά μοντέλα αυξάνει τον κυκλοφοριακό φόρτο. Όταν ο καιρός είναι αίθριος οι άνθρωποι τείνουν να κάνουν παραπάνω μετακινήσεις μιας και το κλίμα τους επιτρέπει πέρα από τις μετακινήσεις που αφορούν βιοποριστικές ανάγκες όπως η εργασία, να πραγματοποιούν και μετακινήσεις αναψυχής. Επίσης εκτός δακτυλίου η θερμοκρασία φαίνεται να έχει μεγαλύτερη επιρροή συγκριτικά με εντός, αυτό συμβαίνει καθώς ακριβώς οι μετακινήσεις εκτός δακτυλίου με προορισμό από και προς τα προάστια είναι παραπάνω για τις ημέρες με θερμό κλίμα αλλά και επειδή οι μετακινήσεις εντός δακτυλίου είναι συνήθως μετακινήσεις για την εξυπηρέτηση αναγκών και επαγγελματικών υποχρεώσεων άρα το κλίμα δεν επηρεάζει τόσο την πραγματοποίησή τους.
5. Οι αργίες για όλα τα στατιστικά μοντέλα έχουν τη μεγαλύτερη σχετική επιρροή από τις ποιοτικές ανεξάρτητες μεταβλητές. Οι μετακινήσεις από και προς την εργασία δεν υφίστανται άρα ο κυκλοφοριακός φόρτος μειώνεται αρκετά
6. Τα σαββατοκύριακα επηρεάζουν καθοδικά τον κυκλοφοριακό φόρτο για όλα τα στατιστικά μοντέλα. Σημαντική είναι η διαφορά που προκύπτει ανάμεσα στις μετακινήσεις εκτός και εντός δακτυλίου. Η διαφορά αυτή προκύπτει από τον εμπορικό χαρακτήρα που έχει το κέντρο της πόλης, πολλές φορές οι οδηγοί πραγματοποιούν μετακινήσεις για αγορές στο κέντρο με αφορμή ακριβώς το σαββατοκύριακο προκειμένου να αποφύγουν τη συμφόρηση των εργασιμων ημερών.
7. Οι σχολικές διακοπές για κάθε μοντέλο αποτελούν παράγοντα που μειώνει τον κυκλοφοριακό φόρτο σε κάθε βάση δεδομένων. Οι σχολικές μετακινήσεις αποτελούν μέρος της κυκλοφορίας ασχέτως της τοποθεσίας με αποτέλεσμα να μην υπάρχει μεγάλη διαφορά στην επιρροή τους στον κυκλοφοριακό φόρτο για εκτός και εντός δακτυλίου.
8. Τα γραμμικά μικτά μοντέλα είναι κατάλληλα για την ανάλυση κυκλοφοριακού φόρτου σε σχέση με οικονομικούς παράγοντες ενώ ιδιαίτερα σημαντική είναι η δυνατότητα προσθήκης τυχαίων επιδράσεων ανά ομάδα παρατηρήσεων μειώνοντας έτσι τη συνολική διακύμανση.
9. Ο κυκλοφοριακός φόρτος είναι πιο ευαίσθητος σε εξωγενείς παράγοντες εκτός του δακτυλίου των Αθηνών σε σύγκριση με εντός. Αυτό συμβαίνει καθώς όπως έχει αναλυθεί το κέντρο της πόλης είναι πάντοτε πόλος έλξης μετακινήσεων για τους κατοίκους όλου του λεκανοπεδίου με αποτέλεσμα να εξυπηρετεί παραπάνω ανάγκες για μεγαλύτερο πληθυσμό. Σημαντική λεπτομέρεια στο παραπάνω φαινόμενο αποτελεί πώς οι περιοχές εντός δακτυλίου διαθέτουν καλύτερη υποδομή μέσων μαζικής μετακίνησης από εκτός αλλά παρόλα αυτά σε περιόδους ακρίβειας οι μετακινήσεις μειώνονται λιγότερο από ότι εκτός του δακτυλίου, αυτό υποδεικνύει ότι οι οδηγοί προτιμούν να μειώσουν ίσως άσκοπες μετακινήσεις παρά να αλλάξουν μέσο μετακίνησης για την καθημερινότητά τους.



## 1. Εισαγωγή

Οι **μεταφορές** και πιο συγκεκριμένα οι μετακινήσεις ανθρώπων και αγαθών με αυτοκίνητα αποτελούν θεμελιώδη στοιχείο των σύγχρονων κοινωνιών. Ο άνθρωπος βασίζεται στην μετακίνηση με αυτοκίνητο για να καλύψει πληθώρα αναγκών, για παράδειγμα βιοποριστικούς σκοπούς όπως είναι η μετακίνηση από και προς το χώρο εργασίας, προμήθεια αγαθών. Ακριβώς επειδή οι μετακινήσεις με αυτοκίνητο αποτέλεσαν ικανή και αναγκαία συνθήκη για την ανάπτυξη των παγκόσμιων οικονομιών όπως τις βιώνουμε σήμερα, συχνά αποτελούν θέμα έρευνας μιας και είναι έντονα συσχετισμένες με διάφορους οικονομικούς δείκτες.

Ενδεικτικά ανατρέχοντας στην ιστορία της Ευρώπης παρατηρείται πως έπειτα την περίοδο του Β Παγκοσμίου πολέμου η **ιδιοκτησία των αυτοκινήτων** για τις πιο ανεπτυγμένες χώρες της Ευρώπης αυξάνεται με εκθετικό ρυθμό, επιβεβαιώνοντας πως η χρήση των αυτοκινήτων αδιαμφισβήτητα έχει σημαντική θέση στη δομή των σύγχρονων πόλεων και οικονομιών.



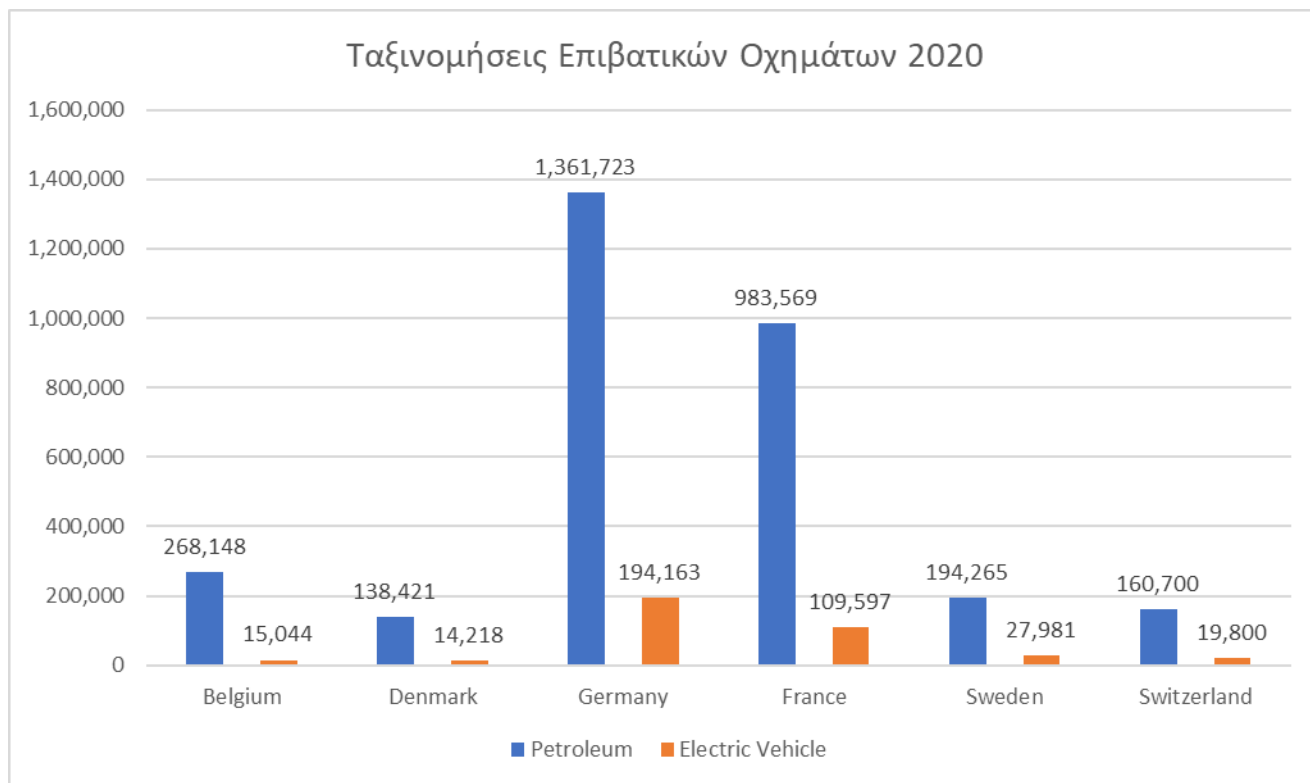
Αυτή η εκθετική ανάπτυξη αποτέλεσε εναρκτήριο λάκτισμα για αρκετούς κρατικούς και όχι φορείς να ερευνήσουν αρχικώς πώς θα μπορέσουν να ανταποκριθούν σε αυτή την υψηλή ζήτηση μιας και οι υποδομές των πόλεων εξελίσσονταν παράλληλα με την παραπάνω αύξηση. Αρχικά, οι πόλεις αφοσιώθηκαν στην αύξηση των λωρίδων και στην όδευση νέων οδών προκειμένου να καλύψουν τις ανάγκες της κυκλοφορίας όμως γρήγορα έγινε αντιληπτό πως τα οχήματα θα πρέπει με κάποιο τρόπο να περιοριστούν μίας και οι γειτονιές των κατοίκων άρχισαν να αλλοιώνονται. Ενδεικτικά, ένα εκ των πρώτων σχεδίων διαχείρισης της κυκλοφορίας εφαρμόστηκε στην Ολλανδία το 1960 γνωστό και ως Wooners Design το οποίο έθετε τον πεζό ως κύριο χρήστη της οδού και μέσω διαφόρων τεχνικών όπως φυσικά εμπόδια η πεζοδρομήσεις ώθησε τους οδηγούς στην επιλογή άλλων οδών ή ακόμα και στη χρήση άλλων μέσων όπως είναι το ποδήλατο. Τα παραπάνω αφορούν εξωγενείς φυσικούς παράγοντες όπως η διαμόρφωση της οδού που μπορούν να επηρεάσουν την κυκλοφορία.

Στα επόμενα χρόνια και ειδικότερα κατά την περίοδο της Πετρελαϊκής κρίσης του 1973 έγινε πλέον πλήρως αντιληπτό πώς η **κίνηση των οχημάτων δεν είναι ανεξάρτητη της οικονομίας** αλλά τότε άρχισε να ποσοτικοποιείται και η επίδραση οικονομικών δεικτών στη χρήση των οχημάτων ενώ παράλληλα φάνηκε πόσο εξαρτημένη είναι η καθημερινότητα των ανθρώπων από τη χρήση των οχημάτων. Από τη μία οι τιμές ρεκόρ του αργού πετρελαίου περιόρισαν τις μετακινήσεις ενώ από την άλλη τα κράτη σε μια προσπάθεια να καλύψουν την αυξημένη ζήτηση ενέργειας επέλεξαν μέτρα όπως η απαγόρευση χρήσης ιδιωτικών οχημάτων τις Κυριακές.

Η ευαισθησία της κυκλοφορίας των οχημάτων ως προς τα **κόστη μετακίνησης** αποτελεί διαχρονικά ερευνητικό αντικείμενο αφενός διότι είναι δύσκολο να προβλέψει κανείς και πόσο μάλλον να ποσοτικοποιήσει όλους τους δυνατούς παράγοντες που ενδεχομένως να μεταβάλλουν τα κυκλοφοριακά μεγέθη, αφετέρου επειδή όπως όλα τα μακροοικονομικά μεγέθη θα χρειαστεί ένα αρκετά μεγάλο χρονικό δείγμα, στο οποίο πολλοί παράγοντες μεταβάλλονται, προκειμένου να δοθεί κάποιο σαφές συμπέρασμα. Παρόλα αυτά μιας και η διαχείριση της κυκλοφορίας από τις αρχές της εμφάνισης των οχημάτων αποτελεί καίριο ενδιαφέρον των κρατικών φορέων αλλά και των εταιρειών που αναλαμβάνουν τη διαχείριση των οδών, γίνονται συνεχώς προσπάθειες για τη βελτίωση των υφιστάμενων μοντέλων.

Διαχρονικά ένας από τους πιο άρρηκτα συνδεδεμένους οικονομικούς δείκτες με τη χρήση των οχημάτων είναι η **τιμή του βαρελιού αργού πετρελαίου**. Επειδή η συντριπτική πλειοψηφία των οχημάτων κινείται με κινητήρες εσωτερικής καύσης είναι φυσικό επόμενο η τιμή του αργού πετρελαίου να επηρεάζει άμεσα το κόστος χρήσης ενός οχήματος μιας και από την απόσταση του αργού πετρελαίου προκύπτουν τα συμβατικά καύσιμα όπως η αμόλυβδη βενζίνη και το πετρέλαιο κίνησης. Αν και με την εξέλιξη της τεχνολογίας των μπαταριών λιθίου καθώς και την βελτίωση της τεχνολογίας αλλά και της υποδομής φόρτισης παρατηρείται αύξηση στα ποσοστά ηλεκτρικών αυτοκινήτων αυτά ακόμα παραμένουν χαμηλά για τον Ευρωπαϊκό στόλο. Ενδεικτικά παρακάτω δίνονται οι ταξινομήσεις για το έτος 2020 για ηλεκτρικά οχήματα σε σύγκριση με

αυτές των οχημάτων εσωτερικής καύσης σε έξι (6) χώρες με τα υψηλότερα ποσοστά ηλεκτροκίνησης.



Γράφημα 1 : Ταξινομήσεις επιβατικών οχημάτων 2020. Πηγή Δεδομένων : Eurostat

Με αιτία την **παγκόσμια ενεργειακή κρίση** που καλπάζει στην Ευρώπη ύστερα της εμπόλεμης κατάστασης στην Ουκρανία προέκυψαν σημαντικές και πρωτοφανείς ανατιμήσεις στα καύσιμα με αποτέλεσμα οι οδηγοί να πρέπει να καλύψουν ένα αρκετά επαυξημένο έξοδο προκειμένου να πραγματοποιήσουν τις καθιερωμένες μετακινήσεις τους. Η παγκόσμια οικονομία ήδη πριν από την έναρξη του πολέμου γνωρίζοντας πως η Ρωσία αποτελεί την τρίτη μεγαλύτερη χώρα σε εξόρυξη αργού πετρελαίου σε μια προσπάθεια να προλάβει οποιοδήποτε εμπόδιο τροφοδοσίας θα πρόκυπτε από την εμπόλεμη ζώνη προσπάθησε να εξασφαλίσει μεγαλύτερα αποθέματα αργού πετρελαίου με αποτέλεσμα η τιμή του αργού πετρελαίου ανά βαρέλι να ξεκινήσει μια άνοδο. Εν συνεχεία ύστερα από το Ευρωπαϊκό εμπάργκο στο ρωσικό πετρέλαιο αλλά και τις κυρώσεις στη Ρωσία ακολούθησαν σημαντικές ελλείψεις στο Ευρωπαϊκό ενεργειακό μείγμα. Χώρες όπως η Γερμανία ήταν ιδιαίτερα ενεργειακά εξαρτημένες όχι μόνο από το ρωσικό αργό πετρέλαιο αλλά κυρίως από το φυσικό αέριο του οποίου η τροφοδοσία πλέον με τις διακοπές στη λειτουργία του αγωγού Nord Stream 1 δεν καλύπτει την Ευρωπαϊκή ζήτηση.



Γράφημα 2 : Μερίδιο Παγκόσμιας Παραγωγής Πετρελαίου. Πηγή Δεδομένων : U.S.Energy Information Administration

## 1.2 Στόχος Διπλωματικής εργασίας

Στόχος της παρούσας Διπλωματικής εργασίας είναι η **διερεύνηση της επιρροής της τιμής της αμόλυβδης βενζίνης στις κυκλοφοριακές συνθήκες της Αθήνας.**

Συγκεκριμένα, επιδιώκεται να εξεταστεί ο βαθμός στον οποίο η τιμή της βενζίνης καθορίζει τον ημερήσιο κυκλοφοριακό φόρτο στην Αθήνα κατά τη διάρκεια της παγκόσμιας ενεργειακής κρίσης λόγω της εμπόλεμης κατάστασης στην Ουκρανία (2/2022-). Προκειμένου να επιτευχθεί ο εν λόγω στόχος συλλέχθηκε ο κυκλοφοριακός φόρτος από 158 μετρητές στην Αθήνα σε 24 επιλεγμένες οδούς καθώς και η μέση ημερήσια τιμή της αμόλυβδης βενζίνης στην Αττική. Η βάση δεδομένων συμπληρώθηκε με τις ημερήσιες καιρικές συνθήκες όπως ύψος βροχής και θερμοκρασία, τον χαρακτηρισμό της ημέρας σε καθημερινή ή σαββατοκύριακο και τις δημόσιες αργίες. Στην συνέχεια, μέσω της ανάπτυξης κατάλληλων στατιστικών μοντέλων περιγράφεται και ποσοτικοποιείται επαρκώς η επιρροή των παραμέτρων που επηρεάζουν τον κυκλοφοριακό φόρτο στην Αθήνα.

Η παρούσα Διπλωματική εργασία στοχεύει στο να προσεγγίσει με μεθοδικότητα και σφαιρικότητα το ερευνητικό αντικείμενο προκειμένου να αντληθούν σαφή συμπεράσματα που με τη σειρά τους θα συνδράμουν στην κατανόηση και εν δυνάμει πρόβλεψη της μεταβλητότητας του κυκλοφοριακού φόρτου της πόλης με γνώμονα την τιμή του καυσίμου κίνησης.



### 1.3 Δομή διπλωματικής εργασίας

Το **πρώτο κεφάλαιο** αποτελεί μια ευρύτερη εισαγωγή στο ερευνητικό αντικείμενο της εργασίας. Παρουσιάζονται κάποια γενικά στοιχεία σχετικά με τις μετακινήσεις με οχήματα στην Ευρώπη και τη σημασία αυτών στην οικονομία. Γίνεται μια σύντομη ιστορική αναδρομή στην πορεία ανάπτυξης του ευρωπαϊκού στόλου οχημάτων καθώς και στην εδραίωση της κυκλοφορίας όπως την γνωρίζουμε σήμερα. Εν συνεχεία, παρουσιάζονται δεδομένα σχετικά με τις πρόσφατες ανατιμήσεις στα καύσιμα και γίνεται αναφορά στον απώτερο σκοπό της εργασίας.

Το **δεύτερο κεφάλαιο** περιλαμβάνει τη βιβλιογραφική ανασκόπηση σχετικά με το αντικείμενο της εργασίας όπου παρουσιάζονται οι παγκόσμιες αναλύσεις και επιστημονικές θέσεις επί του αντικειμένου. Στόχος του κεφαλαίου είναι να ερευνηθεί η υφιστάμενη προσέγγιση και να αντληθούν χρήσιμες πληροφορίες επί του ερευνητικού ζητήματος. Τέλος, εντοπίζονται σημεία όπου η παρούσα Διπλωματική θα συμβάλλει στο ερευνητικό φάσμα της ήδη υπάρχουσας βιβλιογραφίας.

Στο **τρίτο κεφάλαιο** δίνεται το απαραίτητο θεωρητικό υπόβαθρο το οποίο συνοδεύει τις στατιστικές αναλύσεις. Γίνεται αναλυτική περιγραφή της διαδικασίας που οδήγησε στο τελικό μοντέλο καθώς και ποιοι έλεγχοι, ποια μεγέθη εξασφαλίζουν τη στατιστική σημαντικότητα αυτού.

Το **τέταρτο κεφάλαιο** αναφέρεται στην διαδικασία συλλογής και επεξεργασίας των δεδομένων. Επεξηγείται η διαδικασία που έδωσε το τελικό αρχείο (Masterfile) πάνω στο οποίο θα εφαρμοστούν τα στατιστικά μοντέλα ενώ παράλληλα ενώ παράλληλα εξηγείται η κάθε μεταβλητή που εμφανίζεται σε αυτή. Στην συνέχεια, παρατίθενται διάφορα περιγραφικά στοιχεία των συλλεγόμενων δεδομένων.

Στο **πέμπτο κεφάλαιο** παρουσιάζεται ολόκληρη η διαδικασία ανάπτυξης και εφαρμογής της μεθοδολογίας για την παραγωγή των τελικών μαθηματικών μοντέλων. Αναλύονται οι συσχετίσεις των μεταβλητών μεταξύ τους και οι επιλεγόμενες προς τελική χρήση. Κλείνοντας, παρουσιάζονται τα τελικά μοντέλα.

Το **έκτο κεφάλαιο** περιλαμβάνονται τα συνολικά συμπεράσματα, όπως αυτά προέκυψαν από την ερμηνεία των παραχθέντων μαθηματικών προτύπων. Καταγράφεται οργανωμένα η γνώση που απορρέει από το υπολογιστικό τμήμα της παρούσας Διπλωματικής Εργασίας και διατυπώνονται προτάσεις για περαιτέρω έρευνα πάνω στο συγκεκριμένο αντικείμενο.

Το **έβδομο** και τελευταίο κεφάλαιο αποτελεί τη βιβλιογραφία που χρησιμοποιήθηκε κατά την εκπόνηση της Διπλωματικής Εργασίας. Παρουσιάζεται, με μορφή καταλόγου, το σύνολο των ερευνών που έχουν αναφερθεί στα προηγούμενα κεφάλαια.

## 2. Βιβλιογραφική ανασκόπηση

### 2.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζεται η παγκόσμια έρευνα και μεθοδολογίες όπως αυτές έχουν εφαρμοστεί ανά περιόδους σχετικά με το αντικείμενο της Διπλωματικής εργασίας. Όπως

αναφέρθηκε και παραπάνω η συσχέτιση του κυκλοφοριακού φόρτου συναρτήσει της τιμής της βενζίνης και γενικότερα του κόστους μετακίνησης διαχρονικά αποτελεί επίκαιρο αντικείμενο έρευνας είτε για ακαδημαϊκά ιδρύματα είτε από κρατικούς φορείς στα πλαίσια μελέτης για ενδεχόμενες πολιτικές επί του ζητήματος.

Ανατρέχοντας στη βιβλιογραφία προκύπτει ο στόχος της συγκεκριμένης Διπλωματικής ενώ εμφανίζονται οι όποιες δυσκολίες προκύπτουν κατά την προσπάθεια ανάλυσης των οικονομικών παραμέτρων που επηρεάζουν την κυκλοφορία. Ταυτόχρονα καταγράφονται οι απαραίτητες τεχνικές και θεωρήσεις οι οποίες επιτρέπουν στον ερευνητή να βγάλει σαφή συμπεράσματα. Η σφαιρικότητα του ερευνητικού αντικειμένου επιτρέπει πληθώρα αναλύσεων ενώ εξάγει εξίσου πολλά συμπεράσματα. Κοινό στοιχείο των περισσότερων ερευνών αποτελεί το μακροοικονομικό πρίσμα υπό το οποίο προσεγγίζεται το ερευνητικό αντικείμενο.

## 2.2 Συναφείς έρευνες και μεθοδολογίες

Η ανάλυση της επιρροής της τιμής της βενζίνης στην κυκλοφορία και πιο συγκεκριμένα εντός αστικού ιστού είναι από τα πιο χρήσιμα εργαλεία προκειμένου να αντιληφθεί κανείς πώς αντιδρά ο οδηγός σε αλλαγές επί του κόστους μετακίνησης (Zhang and Burke, 2020).

Σε έρευνα που έγινε σε στοιχεία από την Νέα Νότια Ουαλία στην Αυστραλία τα ευρήματα παρουσιάζουν ενδιαφέρον. Η έρευνα έγινε πάνω σε 11,9 εκατομμύρια ωριαίες παρατηρήσεις από 118 σταθμούς μέτρησης της κυκλοφορίας για τις χρονικές περιόδους 2010-2017. Συγκεκριμένα, η ανάλυση έγινε επί του κυκλοφοριακού φόρτου όπου δίνεται ιδιαίτερη έμφαση στον ορισμό προκειμένου να ερμηνευτούν ορθολογικά τα ευρήματα της ανάλυσης. Η πλειοψηφία των σταθμών μέτρησης κυκλοφοριακού φόρτου βρίσκεται γύρω από την περιοχή του Σίδνεϋ και του Νιουκάστλ. Σημειώνεται πως όπως και στην παρούσα Διπλωματική έτσι και σε αυτή την ανάλυση η πόλη της οποίας τα κυκλοφοριακά μεγέθη αναλύθηκαν είναι η πολυπληθέστερη και ταυτόχρονα και η πρωτεύουσα της πολιτείας, όπως ακριβώς συμβαίνει και με την Αθήνα για την Ελλάδα. Εντός της ανάλυσης υπάρχουν δύο (2) οδοί με διόδια και εκατόν δεκαέξι (116) οδοί χωρίς διόδια εκ των οποίων οι είκοσι εννέα (29) είναι αυτοκινητόδρομοι. Οι μετρητές κυκλοφοριακού φόρτου ορισμένες μέρες ή ώρες του χρόνου ενδεχομένως να υπολειτουργούν με αποτέλεσμα ορισμένες μετρήσεις να απέχουν από τις πραγματικές διελεύσεις ή ακόμα και να μην λειτουργούν καθόλου δημιουργώντας έτσι κενά. Επιπρόσθετα, οι μετρητές δεν έχουν την δυνατότητα να διαχωρίσουν τις διελεύσεις οχημάτων ανά όχημα οπότε δεν προκύπτει κατηγοριοποίηση των οχημάτων ανάμεσα σε ελαφρά και βαριά. Ο μεγάλος αριθμός παρατηρήσεων που είχαν στη διάθεση τους οι ερευνητές επέτρεψε σε αυτούς για κάθε ώρα όπου ο εκάστοτε μετρητής δεν λειτουργούσε ορθώς να διαγράφεται η συγκεκριμένη μέρα από τα δεδομένα εξασφαλίζοντας έτσι ομοιογένεια ανάμεσα στις μετρήσεις. Αναφορικά με τα δεδομένα για την τιμή της βενζίνης η έρευνα χρησιμοποιεί τον δείκτη TGP ο οποίος δεν εμπεριέχει τα κόστη ανεφοδιασμού και το κέρδος κάθε πρατηρίου.

Σύμφωνα με την παραπάνω έρευνα ο κυκλοφοριακός φόρτος δεν επηρεάζεται αποκλειστικά από την τιμή της βενζίνης αλλά και από άλλους παράγοντες οι οποίοι θα πρέπει να

συμπεριληφθούν στο στατιστικό μοντέλο προκειμένου αυτό να αποκτήσει επεξηγηματική ισχύ. Οι μεταβλητές που επιλέχθηκαν παρουσιάζονται παρακάτω με την αναλυτική περιγραφή τους:

ΟΝΟΜΑ ΜΕΤΑΒΛΗΤΗΣ	ΤΥΠΟΣ ΜΕΤΑΒΛΗΤΗΣ	ΕΠΕΞΗΓΗΣΗ
<b>Ln Traffic Flow</b>	Εξαρτημένη	Αριθμός διελεύσεων οχημάτων για κάθε μετρητή εντός μίας ώρας.
<b>Ln Real Gasoline Price</b>	Ανεξάρτητη	Μέση ημερήσια τιμή της βενζίνης για την πολιτεία της Νέας Νότιας Ουαλίας.
<b>Ln S&amp;P/ ASX 200 Price</b>	Ανεξάρτητη	Χρηματιστηριακός δείκτης για τις 200 μεγαλύτερες με βάση την κεφαλαιοποίηση τους εταιρείες στο Αυστραλιανό χρηματιστήριο.
<b>Rainfall(mm)</b>	Ανεξάρτητη	Ύψος βροχής σε χιλιοστά για την πόλη του Σύδνεϋ.
<b>Maximum Temperature</b>	Ανεξάρτητη	Μέγιστη ημερήσια θερμοκρασία στην περιοχή του Σύδνεϋ.
<b>Public Holiday</b>	Ανεξάρτητη	Αργίες
<b>School Holiday</b>	Ανεξάρτητη	Σχολικές διακοπές
<b>Hour of day dummy Day of week dummy</b>	Ανεξάρτητη	Μεταβλητές για την ώρα της ημέρας και τη μέρα της εβδομάδας

Πίνακας 1 : Μεταβλητές Ανάλυσης(Zhang and Burke, 2020)

Σύμφωνα με το μοντέλο στο οποίο καταλήγουν προκύπτουν χρήσιμα συμπεράσματα. Αρχικά, η τιμή της βενζίνης όταν αυξάνεται κατά 1% τότε ο κυκλοφοριακός φόρτος μειώνεται κατά 0,04%. Επιπρόσθετα, από τον συντελεστή του οικονομικού δείκτη (Ln S%P / ASX 200) υπογραμμίζεται η σχέση ανάμεσα στα κυκλοφοριακά μεγέθη και την εύρυθμη λειτουργία της οικονομίας. Όπως είναι αναμενόμενο όταν η οικονομία μιας χώρας καλπάζει τότε αυξάνει η αγοραστική ικανότητα του πληθυσμού της με αποτέλεσμα να πραγματοποιούνται παραπάνω μετακινήσεις. Σχετικά με το ύψος βροχόπτωσης οι συγγραφείς δικαιολογούν τον αρνητικό συντελεστή καθώς σε περιόδους βροχόπτωσης οι κάτοικοι ενδεχομένως να αποφεύγουν την οδήγηση γενικότερα μιας και οι συνθήκες στους δρόμους τείνουν να είναι πιο επικίνδυνες εξαιτίας της ολισθηρότητας του οδοστρώματος αλλά και της περιορισμένης ορατότητας. Όσον αφορά τη μέγιστη θερμοκρασία είναι αρκετά εύλογο όταν ο καιρός είναι αίθριος να παρατηρούνται παραπάνω μετακινήσεις και άρα ο κυκλοφοριακός φόρτος να αυξάνει.

Στα παραπάνω μεγέθη τη μεγαλύτερη κατά απόλυτη τιμή επιρροή στον κυκλοφοριακό φόρτο δείχνει να έχουν οι αργίες και οι διακοπές ενώ μικρότερη αλλά επίσης στατιστικά σημαντική φαίνεται να είναι και η επιρροή των σχολικών διακοπών. Αμφότερα μεγέθη έχουν αρνητική

επιρροή καθώς σε περιόδους διακοπών είναι αποδεκτό ότι οι κυκλοφοριακοί φόρτοι μειώνονται σημαντικά.

Επιπρόσθετα οι ερευνητές εξετάζουν αν υπάρχει διαφορά στο πώς αντιδρούν οι οδηγοί στην αλλαγή της τιμής βενζίνης ανάλογα με το εάν χρησιμοποιούν οδό με διόδια ή χωρίς και καταλήγουν πως δεν υπάρχει διαφορά ανάμεσα στην επιρροή της τιμής της βενζίνης για οδούς με ή χωρίς διόδια.

Η έρευνα των (Bento et al., 2013) με ερευνητικό αντικείμενο την αντίδραση των οδηγών σε μεταβολές του κόστους καυσίμου εν συναρτήσει με τη χρήση λωρίδων οχημάτων υψηλής πλήρωσης (HOV) επιβεβαιώνει την θεωρία πως ο κυκλοφοριακός φόρτος καθώς και οι επιλογές των οδηγών είναι άρρηκτα συνδεδεμένες με το κόστος μετακίνησης. Η έρευνα δεν περιορίζεται αποκλειστικά στη μελέτη οδών όπου υπάρχει η δυνατότητα χρήσης λωρίδας για οχήματα υψηλής πλήρωσης (HOV Lanes) αλλά αντιθέτως οι παρατηρήσεις αφορούν οδούς με και χωρίς λωρίδες HOV. Αν και οι Ηνωμένες Πολιτείες της Αμερικής (ΗΠΑ) έχουν επενδύσει αρκετά είτε άμεσα με τη χάραξη τέτοιων λωρίδων είτε έμμεσα με την προώθηση γενικότερα της χρήσης ενός κοινόχρηστου οχήματος που εξυπηρετεί παραπάνω από ένα άτομο (carpooling) με ευνοϊκές πολιτικές, ελάχιστες είναι οι έρευνες που μελετάνε πώς το κόστος μετακίνησης επηρεάζει τους οδηγούς. Στην συγκεκριμένη μελέτη εξετάζεται κατά πόσο μια αύξηση στην τιμή της βενζίνης μπορεί να μειώσει τον συνολικό κυκλοφοριακό φόρτο της πόλης με αποτέλεσμα να υπάρχουν λιγότερα οχήματα στις οδούς και ενδεχομένως λιγότερα ατυχήματα. Τα δεδομένα για την ανάλυση προκύπτουν από το California Freeway Performance Measurement System (PeMS) το οποίο συλλέγει σε πραγματικό χρόνο δεδομένα κυκλοφορίας.

Για την παραπάνω μελέτη συλλέχθηκαν δεδομένα από 1.727 σταθμούς μέτρησης για την περίοδο Ιούλιος 2000 έως Δεκέμβριος 2007. Από αυτούς οι 442 αφορούν αποκλειστικά λωρίδες HOV, οι 173 κύριες λωρίδες ενώ οι 1.112 οδούς με κύριες λωρίδες και HOV. Έτσι μπορούν να προκύψουν συμπεράσματα σχετικά με το συνολικό φόρτο των οδών αλλά και ανά κατηγορία, διαχωρίζεται έτσι ο κυκλοφοριακός φόρτος σε φόρτο οδών όπου υπάρχουν αποκλειστικά κύριες λωρίδες, σε φόρτο ειδικών HOV λωρίδων αλλά και σε φόρτο οδών όπου η υποδομή επιτρέπει στους οδηγούς είτε να κινηθούν σε κύριες λωρίδες εξυπηρετώντας ατομικές μετακινήσεις είτε σε λωρίδες για οχήματα υψηλής πλήρωσης (HOV) όπου συχνά η μετακίνηση εξυπηρετεί δύο ή και παραπάνω άτομα. Τα αποτελέσματα του στατιστικού μοντέλου επιβεβαιώνουν τον παραπάνω διαχωρισμό μίας και κατά την αύξηση της τιμής της βενζίνης σημειώνεται μείωση του κυκλοφοριακού φόρτου για αμφότερες οδούς με κύριες λωρίδες ανεξάρτητα της ύπαρξης λωρίδων HOV. Η διαφορά έγκειται στην τάξη μεγέθους καθώς όπου υπάρχει η εναλλακτική της χρήσης ειδικών λωρίδων οι οποίες αφορούν αποκλειστικά οχήματα υψηλής πλήρωσης η κοινή χρήση οχήματος (carpooling) παρατηρείται μεγαλύτερη μείωση του κυκλοφοριακού φόρτου στις κύριες λωρίδες έναντι των οδών όπου αυτή η εναλλακτική δεν δίνεται στους οδηγούς. Αντίστοιχα αποκλειστικά για τις ειδικές λωρίδες παρατηρείται αύξηση του κυκλοφοριακού φόρτου. Συμπεραίνουν λοιπόν οι ερευνητές πώς σε περιόδους αύξησης του κόστους μετακίνησης και πιο συγκεκριμένα της τιμής της βενζίνης οι οδηγοί είτε θα περιορίσουν τις

μετακινήσεις με ιδιωτικά οχήματα είτε θα επιλέξουν να μοιραστούν το κόστος της διαδρομής κάνοντας χρήση των λωρίδων HOV. Το παραπάνω φαινόμενο μάλιστα είναι πιο έντονο κατά τις περιόδους εκτός τις ώρες αιχμής συμπεραίνοντας πώς οι οδηγοί είναι πιο σκεπτικοί όταν πρόκειται για μετακινήσεις που δεν αφορούν βιοποριστικούς λόγους σε περιόδους αυξημένου κόστους μετακίνησης.

Περαιτέρω έρευνες όπως αυτή των (Delsaut, 2014) για την περίπτωση του εθνικού δικτύου της Γαλλίας ενισχύουν τη θεωρία πώς ο κυκλοφοριακός φόρτος δεν είναι ανελαστικός απέναντι σε αυξήσεις στο κόστος των καυσίμων. Πιο συγκεκριμένα λαμβάνοντας υπόψιν το εθνικό ακαθάριστο προϊόν (ΑΕΠ) της χώρας, την τιμή της βενζίνης αλλά και το μήκος του δικτύου αυτοκινητοδρόμων για κάθε έτος στην Γαλλία το στατιστικό μοντέλο των ερευνητών εκτιμά πως μια αύξηση της τάξης του 10 % στην τιμή του καυσίμου προκαλεί πτώση της 1.3% των συνολικών οχηματο-χιλιομέτρων.

### 2.3 Συμπεράσματα βιβλιογραφίας

Η παγκόσμια βιβλιογραφία εμπεριέχει αρκετές αναφορές στις μεταβολές του κυκλοφοριακού φόρτου είτε πρόκειται για έρευνες οι οποίες αφιερώνονται καθαρά στην βιβλιογραφική ανασκόπηση είτε πρόκειται για έρευνες οι οποίες με στατιστικές μεθόδους αναλύουν και επεξηγούν όλους τους πιθανούς παράγοντες οι οποίοι σχηματίζουν και τροφοδοτούν κατά κάποιο ποσοστό τα κυκλοφοριακά μεγέθη όπως αυτά τα γνωρίζουμε σήμερα. Τα πλεονεκτήματα που προκύπτουν από τέτοιες μελέτες είναι αναρίθμητα καθώς δεν είναι λίγες οι φορές όπου τέτοιες έρευνες αποτελούν θεμέλιο για την εδραίωση πολιτικών τιμολόγησης και φορολογίας στο ευρύτερο κομμάτι των μετακινήσεων είτε αυτό αφορά τις οδικές μετακινήσεις είτε γενικότερα τις συγκοινωνίες.

Στην πλειοψηφία των αναλύσεων προκύπτει πως το κόστος ανεφοδιασμού είναι αρκετά σημαντικός παράγοντας για τους οδηγούς ενώ ανάλογα την περίπτωση παρατηρείται και η αλλαγή στη συμπεριφορά αυτών. Για παράδειγμα στην έρευνα των (Bento et al., 2013) ορισμένοι οδηγοί με την αύξηση της τιμής καυσίμου κίνησης, μείωσαν τις μετακινήσεις τους ενώ άλλοι αποφάσισαν να μοιραστούν το κόστος. Επίσης, παρατηρείται ένας κοινός παρονομαστής στις περισσότερες έρευνες, οι μετακινήσεις με αυτοκίνητο αποτελούν καθημερινότητα των πολιτών ειδικά για την κάλυψη βιοποριστικών αναγκών όπως είναι η μετακίνηση από και προς το χώρο εργασίας γεγονός που αφενός δείχνει ότι εύκολα και άμεσα οι χρήστες της οδού αντιλαμβάνονται το αυξημένο κόστος που προκύπτει από τις ανατιμήσεις στις τιμές της βενζίνης και αφετέρου ακριβώς επειδή είναι αναπόσπαστο κομμάτι της καθημερινότητας τους θα προτιμήσουν να περιορίσουν μετακινήσεις που δεν προκύπτουν κατά τις ώρες αιχμής παρά τις κύριες κατά την ώρα αιχμής. Στις περισσότερες μελέτες αναφέρεται πως για να αποφανθεί κανείς εάν το αυξημένο κόστος των καυσίμων μειώνει τον κυκλοφοριακό φόρτο θα πρέπει να κατέχει κυκλοφοριακά δεδομένα για μεγάλο χρονικό διάστημα και ειδικότερα αρκετά μετέπειτα της αύξησης του κόστους διότι οι οδηγοί δεν αντιδρούν άμεσα αλλά χρειάζονται κάποιο εύλογο χρονικό διάστημα για να αναπροσαρμόσουν την καθημερινότητα τους και να αναζητήσουν πιο οικονομικές εναλλακτικές. Η συγκεκριμένη

Διπλωματική εργασία με γνώμονα τα άνωθεν θα προσπαθήσει να υπογραμμίσει τη σημασία του κόστους ανεφοδιασμού για τους οδηγούς της Αθήνας και παράλληλα να εξηγήσει περαιτέρω παράγοντες που επηρεάζουν την κυκλοφορία στην πόλη.

### 3. Θεωρητικό Υπόβαθρο

#### 3.1 Εισαγωγή

Στο παρόν κεφάλαιο περιγράφεται αναλυτικά το θεωρητικό υπόβαθρο επι του οποίου πραγματοποιήθηκε επεξεργασία και ανάλυση των δεδομένων της εργασίας. Οι στατιστικές αναλύσεις αποτελούν βασικό εργαλείο στην φαρέτρα των αναλυτών και έχουν αναρίθμητα πεδία εφαρμογής. Με την κατάλληλη εφαρμογή στατιστικών μοντέλων η ανθρωπότητα έχει ερμηνεύσει και κατανοήσει ποικίλα φαινόμενα ενώ παράλληλα η στατιστική είναι αυτή που σε κάθε έρευνα δίνει εγκυρότητα. Η ενότητα θα χωριστεί σε δύο υποενότητες, αρχικά θα δοθούν αναλυτικοί ορισμοί στατιστικών μεγεθών τα οποία διασφαλίζουν την ορθή αντίληψη, διαχείριση των δεδομένων, θα οριστούν τα μοντέλα παλινδρόμησης και τέλος θα οριστούν σημαντικές έννοιες της κυκλοφοριακής τεχνικής μέσω των οποίων καθορίζονται οι μετρήσιμες ποσότητες της εργασίας και επεξηγούνται ιδιότητες του κυκλοφοριακού φόρτου.

#### 3.2 Βασικές έννοιες στατιστικής

Με τον όρο **πληθυσμός** αναφερόμαστε σε μια ολότητα τιμών αναφερόμενη σε κάποιο ποσοτικό η ποιοτικό χαρακτηριστικό στο οποίο εστιάζει η στατιστική έρευνα. Πρόκειται για ένα καλά ορισμένο σύνολο στοιχείων το οποίο συνήθως είναι πεπερασμένο. Αρκετές φορές υπάρχουν πληθυσμοί των οποίων το μέγεθος αν και πεπερασμένο είναι μεγάλο και η ολικής απογραφή του είναι πρακτικώς αδύνατη.

Ο όρος **δείγμα** αποτελεί ένα υποσύνολο του πληθυσμού το οποίο επιλέγεται με κατάλληλο τρόπο ώστε να είναι αντιπροσωπευτικό του πληθυσμού τον οποίο η έρευνα μελετά. Είναι σημαντικό οι τιμές για τα μεγέθη που γίνεται η ανάλυση να κατανέμονται με τον ίδιο τρόπο στο δείγμα και στον πληθυσμό. Τα στοιχεία που ανήκουν σε ένα δείγμα ανήκουν και στον πληθυσμό όμως το αντίστροφο δεν συμβαίνει. Όταν το δείγμα το οποίο αναλύεται είναι αντιπροσωπευτικό του πληθυσμού τότε τα συμπεράσματα που προκύπτουν από τη στατιστική μελέτη για το δείγμα ισχύουν με ικανοποιητική ακρίβεια και για τον πληθυσμό.

Το αρχικό στάδιο κάθε στατιστική μελέτης είναι η συλλογή δεδομένων τα οποία με τη σειρά τους θα αναλυθούν από τα στατιστικά μοντέλα και θα δώσουν τα αποτελέσματα της μελέτης. Τα δεδομένα χωρίζονται σε δύο κατηγορίες, τα ποιοτικά (κατηγορικά) και τα ποσοτικά (αριθμητικά).

- **Ποιοτικά δεδομένα** : Πρόκειται για δεδομένα που αφορούν μη αριθμητικά χαρακτηριστικά για τα στοιχεία ενός δείγματος. Έστω ότι έχουμε διαφορετικές οδούς σε ένα δείγμα μετρήσεων τότε εάν αντιστοιχήσουμε σε κάθε οδό τους αριθμούς 1 ή 0 ανάλογα με το εάν η οδός είναι εντός η εκτός του δακτυλίου της Αθήνας τότε έχουμε ένα σύνολο ποιοτικών δεδομένων. Δεδομένα όπως αυτά δεν έχουν κάποια ιεράρχηση και ονομάζονται δεδομένα ονομαστικής κλιμάκωσης. Ενδέχεται όμως να υπάρχει στο δείγμα κάποια ιεράρχηση στα ποιοτικά μας δεδομένα, για παράδειγμα αν αντιστοιχήσουμε

τους αριθμούς δυο, τρία και τέσσερα σε οδούς ανάλογα με το πλήθος λωρίδων σε κάθε οδό τότε τα ποιοτικά μας δεδομένα ονομάζονται δεδομένα ιεραρχικής κλιμάκωσης.

- **Ποσοτικά δεδομένα** : Είναι τα δεδομένα με αριθμητικά χαρακτηριστικά και καταμετρήσεις. Δεδομένα μετρήσεων ονομάζονται αυτά που αντιστοιχούν σε μετρήσεις φυσικής ιδιότητας για τα στοιχεία του δείγματος. Δεδομένα καταμέτρησης ονομάζονται εκείνα που αναφέρονται σε καταμετρήσεις αριθμού των εμφανίσεων κάποιου φαινομένου ή σε καταμετρήσεις αριθμού των στοιχείων που ανήκουν σε μια κατηγορία όπως αυτή έχει οριστεί από τους ερευνητές.

Από τη συλλογή των δεδομένων κατά την εκπόνηση μιας στατιστικής μελέτης προκύπτουν και οι διάφορες μεταβλητές πάνω στις οποίες θα στηριχθεί η στατιστική ανάλυση. Με τον όρο **μεταβλητές** εννοούμε όλα τα χαρακτηριστικά που μετρούνται κατά την ανάλυση και θα ενταχθούν στο στατιστικό μοντέλο. Οι μεταβλητές με τη σειρά τους ανάλογα με το ποιο μέγεθος ή ιδιότητα για κάθε στοιχείο του δείγματος εκφράζονται διαχωρίζονται σε επιμέρους κατηγορίες οι οποίες παρουσιάζονται παρακάτω.

**Κατηγορικές μεταβλητές** : Ονομάζονται οι μεταβλητές των οποίων οι τιμές αναφέρονται στο διαχωρισμό των στοιχείων ενός δείγματος σε επιμέρους κατηγορίες. Διαχωρίζονται σε **ονομαστικές** (nominal) και **διάταξης** (ordinal). Οι πρώτες αφορούν χαρακτηριστικά των στοιχείων του δείγματος για τα οποία δεν υπάρχει κάποια σχέση ανάμεσα στις διαφορετικές κλάσεις. Για παράδειγμα σε μια ανάλυση μια τέτοια μεταβλητή θα μπορούσε να ήταν η τεχνολογία λειτουργίας του κινητήρα εσωτερικής καύσης των οχημάτων, αν η καύση δηλαδή γίνεται με χρήση αμόλυβδης βενζίνης ή πετρελαίου. Οι κατηγορικές μεταβλητές διάταξης δηλώνουν μια ιεράρχηση ανάμεσα στις πιθανές κατηγορίες, για παράδειγμα πάλι στους κινητήρες εσωτερικής καύσης θα μπορούσαμε να θέσουμε μεταβλητές για την κατανάλωση καυσίμου του εκάστοτε κινητήρα όπως «χαμηλή», «μέση», «υψηλή».

**Ποσοτικές μεταβλητές** : Είναι οι μεταβλητές οι οποίες εκφράζουν μια μετρήσιμη ποσότητα. Τέτοιες μεταβλητές έχουν αριθμητική τιμή η οποία αναφέρεται σε κάποια μέτρηση και συνοδεύονται πάντα από τις μονάδες μέτρησης. Οι ποσοτικές μεταβλητές διαχωρίζονται σε **συνεχείς** (continuous) και **διακριτές** (discrete). Συνεχείς ονομάζονται οι μεταβλητές των οποίων οι τιμές είναι πραγματικοί αριθμοί και αποτελούν υποσύνολο πραγματικών αριθμών άρα και δύο τιμές μπορούν να διαφέρουν κατά οποιαδήποτε μικρή ποσότητα. Τέτοια μεταβλητή είναι για παράδειγμα η πραγματική αριθμητική κατανάλωση ενός οχήματος. Διακριτές είναι οι μεταβλητές οι οποίες παίρνουν τιμές μόνο φυσικούς αριθμούς όπως είναι για παράδειγμα ο αριθμός των θέσεων ενός οχήματος υψηλής πλήρωσης.

**Μέτρα κεντρικής τάσης** : Για περιπτώσεις ανάλυσης ενός δείγματος με στοιχεία  $x_1, x_2, \dots, x_n$  η **μέση τιμή** υπολογίζεται σύμφωνα με τη σχέση:

$$\bar{x} = \frac{(x_1 + x_2 + \dots + x_n)}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$



**Μέτρα διασποράς και μεταβλητότητας :** Η διακύμανση των δεδομένων ενός δείγματος συμβολίζεται με  $s^2$  και διαιρείται με  $(n-1)$ :

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

Όπου  $\bar{x}$  ο δειγματικός μέσος, δηλαδή η **μέση τιμή** των παρατηρήσεων στο δείγμα.

Η τυπική απόκλιση του δείγματος δίνεται από τη μαθηματική σχέση :

$$s = \sqrt{s^2} = \left[ \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \right]^{1/2}$$

Για συμμετρικά κατανομημένο δείγμα δεδομένων, σύμφωνα με έναν εμπειρικό κανόνα προκύπτει ότι το διάστημα :

- $(-s, +s)$  περιέχει περίπου το 68% των δεδομένων.
- $(-2s, +2s)$  περιέχει περίπου το 95% των δεδομένων.
- $(-3s, +3s)$  περιέχει περίπου το 99% των δεδομένων.

**Συνδιακύμανση :** Αποτελεί ένα μέτρο της σχέσης μεταξύ δύο περιοχών δεδομένων.

$$cov(X, Y) = \left[ \frac{1}{n-1} \right] \sum_{i=1}^n [(X_i - \bar{X})(Y_i - \bar{Y})]$$

**Μέτρα αξιοπιστίας :**

- **Επίπεδο εμπιστοσύνης :** η αναλογία των περιπτώσεων που μια εκτίμηση θα είναι σωστή
- **Επίπεδο σημαντικότητας :** η αναλογία των περιπτώσεων που ένα συμπέρασμα είναι εσφαλμένο.

### 3.3 Συντελεστής συσχέτισης μεταβλητών

Προκειμένου να προσδιοριστεί η συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών υπολογίζεται ο συντελεστής συσχέτισης  $r$  του Pearson (Pearson Correlation) για δύο μεταβλητές και ορίζεται από την παρακάτω εξίσωση:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Οι τιμές του συντελεστή συσχέτισης κυμαίνονται μεταξύ -1 και 1, με  $r=0$  να ερμηνεύεται ως μηδενική συσχέτιση,  $r=1$  πλήρη θετική συσχέτιση και  $r=-1$  πλήρη αρνητική συσχέτιση. Θετική συσχέτιση ανάμεσα σε δύο μεταβλητές παρατηρείται όταν για αύξηση της τιμής της μίας παρατηρείται και αύξηση της τιμής της άλλης ενώ αντιστρόφως ανάλογα κατά δυο μεταβλητές έχουν αρνητική συσχέτιση όταν για αύξηση της τιμής της πρώτης έχουμε μείωση στην τιμή της άλλης μεταβλητής. Σημειώνεται πώς η συσχέτιση δεν εκφράζει αιτία άρα δεν μπορούμε να αποφανθούμε πώς για θετική συσχέτιση αύξηση κάποιας μεταβλητής έχει ως αποτέλεσμα την αύξηση κάποιας άλλης απλά παρατηρούμε πώς οι αυξήσεις συμβαίνουν ταυτόχρονα.

### 3.4 Κανονική Κατανομή

Από τις πιο σημαντικές κατανομές πιθανότητας η οποία αναφέρεται σε συνεχείς μεταβλητές είναι η κανονική κατανομή ή κατανομή Gauss. Μια συνεχής τυχαία μεταβλητή  $X$  θεωρείται ότι ακολουθεί την κανονική κατανομή με παραμέτρους  $\mu, \sigma$  ( $-\infty < \mu < +\infty, \sigma > 0$ ) και γράφεται  $X \sim N(\mu, \sigma^2)$ , όταν έχει συνάρτηση πυκνότητας πιθανότητας την :

$$f(x) = \left( \frac{1}{\sigma \sqrt{2\pi}} \right) e^{\left[ -\frac{(x-\mu)^2}{2\sigma^2} \right]}$$

Όπου  $\mu$  και  $\sigma$  είναι σταθερές ίσες με τη μέση τιμή και την τυπική απόκλιση αντίστοιχα.

### 3.5 Μαθηματικά Πρότυπα

#### 3.5.1 Γραμμική Παλινδρόμηση

Το τμήμα της στατιστικής το οποίο εξετάζει τη σχέση μεταξύ δύο η περισσότερων μεταβλητών, ώστε να είναι δυνατή η πρόβλεψη της μίας από τις υπόλοιπες, ονομάζεται **ανάλυση παλινδρόμησης (regression analysis)**. Εξαρτημένη μεταβλητή ορίζεται η μεταβλητή η οποία η μελέτη έχει ως στόχο να προβλέψει, ενώ ανεξάρτητη μεταβλητή είναι η μεταβλητή μέσω της οποίας θα γίνει η πρόβλεψη για την εξαρτημένη μεταβλητή. Η ανεξάρτητη μεταβλητή δεν θεωρείται τυχαία, αλλά παίρνει καθορισμένες τιμές ενώ η εξαρτημένη μεταβλητή θεωρείται τυχαία και προκύπτει από την ανεξάρτητη μεταβλητή. Για να προσδιοριστεί αν μια ανεξάρτητη μεταβλητή ή συνδυασμός ανεξάρτητων μεταβλητών προκάλεσε τη μεταβολή της εξαρτημένης μεταβλητής είναι απαραίτητη η ανάπτυξη μαθηματικών, στατιστικών μοντέλων,

Στόχος των στατιστικών αναλύσεων είναι η ανάπτυξη εξισώσεων που περιγράφουν τη σχέση μεταξύ των ανεξάρτητων μεταβλητών και της εξαρτημένης. Η επιλογή της μεθόδου ανάπτυξης ενός μοντέλου βασίζεται στο αν η εξαρτημένη μεταβλητή είναι συνεχής ή διακριτό μέγεθος.

Προκειμένου να αναλυθεί εξαρτημένη μεταβλητή η οποία είναι συνεχής και ακολουθεί την κανονική κατανομή, μια από τις πλέον εφαρμοσμένες στατιστικές τεχνικές είναι η γραμμική παλινδρόμηση. Η απλούστερη μορφή γραμμικής παλινδρόμησης είναι η **απλή γραμμική παλινδρόμηση (simple linear regression)**.

Στην απλή γραμμική παλινδρόμηση υπάρχει μόνο μίας ανεξάρτητη μεταβλητή X και μια εξαρτημένη μεταβλητή Y, που προσεγγίζεται ως γραμμική συνάρτηση του X. Η τιμή  $y_i$  για κάθε τιμή της  $x_i$  προκύπτει από τη σχέση :

$$y_i = a + b * x_i + \varepsilon_i$$

Στόχος της γραμμικής παλινδρόμησης είναι η εύρεση των παραμέτρων  $a$  και  $b$  που εκφράζουν καλύτερα τη γραμμική εξάρτηση της Y από τη X. Κάθε ζεύγος τιμών ( $a, b$ ) καθορίζει μια διαφορετική γραμμική σχέση που εκφράζεται γεωμετρικά από ευθεία γραμμή και οι δύο παράμετροι ορίζονται ως εξής :

- Ο σταθερός όρος  $a$  είναι η τιμή του  $y$  για  $x=0$ .
- Ο συντελεστής  $b$  του  $x$  είναι η κλίση (slope) της ευθείας ή αλλιώς ο συντελεστής παλινδρόμησης (regression coefficient) και εκφράζει τη μεταβολή της μεταβλητής Y όταν η μεταβλητή X μεταβληθεί κατά μία μονάδα.

Ο όρος  $\varepsilon_i$  είναι το σφάλμα **παλινδρόμησης** (regression error). Στην πράξη ο γραμμικός προσδιορισμός που επιτυγχάνεται μέσω της μεθόδου της γραμμικής παλινδρόμησης μπορεί μόνο να προσεγγίσει την πραγματική μαθηματική σχέση μεταξύ των δύο μεταβλητών X και Y. Έτσι είναι απαραίτητο να συμπεριληφθεί στο πρότυπο ο όρος του σφάλματος  $\varepsilon_i$ . Αυτό γίνεται προκειμένου να αντιπροσωπευθούν στο πρότυπο τυχόν παραληφθείσες μεταβλητές, όσο και να ληφθεί υπόψη κάθε σφάλμα προσέγγισης που σχετίζεται με τη γραμμική συναρτησιακή μορφή (A. Σταθόπουλος and M. Καρλαύτης, 2016).

Όταν η τυχαία εξαρτημένη μεταβλητή Y εξαρτάται από περισσότερες από μία μεταβλητές  $X_1, X_2, X_3, \dots, X_k$ , τότε πρόκειται για **πολλαπλή γραμμική παλινδρόμηση** (multiple linear regression). Η εξίσωση που περιγράφει τη σχέση μεταξύ εξαρτημένων και ανεξάρτητων μεταβλητών είναι η εξής :

$$y_i = \beta_0 + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \beta_3 * x_{3i} + \dots + \beta_k * x_{ki} + \varepsilon_i$$

Η διαφορά της πολλαπλής γραμμικής παλινδρόμησης από την απλή γραμμική έγκειται κυρίως στη μεθοδολογία επιλογής των ανεξάρτητων μεταβλητών καθώς για να έχει επεξηγηματική ισχύ το μοντέλο και κάθε ανεξάρτητη μεταβλητή να είναι στατιστικά σημαντική θα πρέπει πέρα από το να είναι σχετική με την εξαρτημένη, θα πρέπει να έχει **χαμηλή συσχέτιση με τις υπόλοιπες ανεξάρτητες μεταβλητές** στις οποίες θα βασίζεται η παλινδρόμηση.

Η εκτίμηση των παραμέτρων της πολλαπλής γραμμικής παλινδρόμησης γίνεται με τη μέθοδο των ελαχίστων τετραγώνων (method of least squares). Ο προσδιορισμός των  $\beta_i$  συντελεστών δίνει μια προσεγγιστική ευθεία η οποία συνδέει τις τιμές της μεταβλητής Y δοθέντων των τιμών των X. Σκοπός είναι το άθροισμα των τετραγώνων των κατακόρυφων αποστάσεων των σημείων (X,Y) από την ευθεία να είναι ελάχιστο.

Προκειμένου το πρότυπο να μπορεί να προσεγγίσει την επιρροή των ανεξάρτητων μεταβλητών στην εξαρτημένη με όσο το δυνατόν πιο αξιόπιστο τρόπο, θα πρέπει να πληρούνται οι παρακάτω προϋποθέσεις.

- Η υπόθεση της **γραμμικότητας**, που δηλώνει ότι η σχέση μεταξύ δύο μεταβλητών X και Y είναι κατά προσέγγιση γραμμική.
- Η υπόθεση της **ανεξαρτησίας**, που δηλώνει ότι τα υπόλοιπα (σφάλματα, αποκλίσεις) για διαφορετικές παρατηρήσεις πρέπει να είναι ανεξάρτητα μεταξύ τους.
- Η υπόθεση της **κανονικότητας**, που δηλώνει ότι η απόκλιση πρέπει να είναι (προσεγγιστικά ) κανονικά κατανοημένη.
- Η υπόθεση της **ίσης διακύμανσης**, που δηλώνει ότι η διακύμανση των σφαλμάτων πρέπει να παραμένει στο ίδιο εύρος για όλες τις παρατηρήσεις.

### 3.5.2 Λογαριθμοκανονική Παλινδρόμηση

Μέσω της λογαριθμοκανονικής παλινδρόμησης (log-normal regression) δίνεται η δυνατότητα ανάπτυξης ενός πρότυπου που συσχετίζει δύο ή περισσότερες μεταβλητές. Η σχέση που συνδέει την εξαρτημένη με τις ανεξάρτητες μεταβλητές είναι γραμμική. Στη λογαριθμοκανονική παλινδρόμηση οι συντελεστές των μεταβλητών του μοντέλου είναι συντελεστές της γραμμικής παλινδρόμησης και υπολογίζονται από την ανάλυση της παλινδρόμησης με βάση την **αρχή των ελαχίστων τετραγώνων**.

Η λογαριθμοκανονική παλινδρόμηση βασίζεται στην υπόθεση ότι τα στοιχεία που περιέχονται στη βάση δεδομένων είναι μη αρνητικά, ο φυσικός λογάριθμος της ανεξάρτητης μεταβλητής ακολουθεί την κανονική κατανομή και ο αριθμητικός μέσος είναι σχετικά μεγάλος. Η μαθηματική σχέση που περιγράφει τη μέθοδο αυτή είναι :

$$\log y_i = \beta_0 + \beta_{1*} * x_{1i} + \beta_2 * x_{2i} + \beta_3 * x_{ei} + \dots + \beta_k * x_{ki} + \varepsilon_i$$

Όπου y είναι η εξαρτημένη μεταβλητή ,  $\beta_0, \beta_1, \dots, \beta_k$  οι συντελεστές μερικής παλινδρόμησης και  $x_{1i}, \dots, x_{ki}$  οι ανεξάρτητες μεταβλητές και  $\varepsilon_i$  το σφάλμα της παλινδρόμησης.

### 3.5.3 Γενικό γραμμικό μοντέλο

Το **γενικό γραμμικό μοντέλο** (General Linear Model) μπορεί να θεωρηθεί ως επέκταση της γραμμικής πολλαπλής παλινδρόμησης για μια μεμονωμένη εξαρτημένη μεταβλητή. Η διαφορά του από το μοντέλο πολλαπλής παλινδρόμησης έγκειται στον αριθμό των εξαρτημένων μεταβλητών που μπορεί να αναλυθεί.

Η μαθηματική σχέση, που περιγράφει τη μέθοδο, για μια εξαρτημένη μεταβλητή  $x_{ij}$  , όπου  $j=1,2, \dots, J$  ο εκάστοτε παράγοντας είναι :

$$x_{ij} = g_{i1} * \beta_{1j} + g_{i2} * \beta_{2j} + \dots + g_{ik} * \beta_{kj} + e_{ij}$$

Όπου το  $i=1,2,\dots,I$  δηλώνει την παρατήρηση.

Το γενικό γραμμικό μοντέλο βασίζεται στην υπόθεση ότι τα σφάλματα ( $e_{ij}$ ) είναι ανεξάρτητα και κατανέμονται κανονικά  $[ N(0, \sigma_j^2) ]$ . Οι συντελεστές  $g_{ik}$  είναι μεταβλητές που σχετίζονται με τις συνθήκες κάτω από τις οποίες πραγματοποιήθηκε η παρατήρηση  $i$ . Αυτοί οι συντελεστές μπορούν να είναι δύο ειδών :

- Μια συμμεταβλητή ( μεταβλητή ελέγχου – covariate ). Στην περίπτωση αυτή η παραπάνω εξίσωση είναι ένα πολυμεταβλητό μοντέλο παλινδρόμησης.
- Εικονικές μεταβλητές. Ο συγκεκριμένος τύπος μεταβλητών χρησιμοποιεί ακέραιες τιμές για να εκφράσει το επίπεδο ενός παράγοντα, δεδομένου του οποίου μετριέται η εξαρτημένη μεταβλητή.

#### 3.5.4 Γραμμικό μικτό μοντέλο (Linear mixed model)

Το γραμμικό μικτό μοντέλο επεκτείνει το γενικό γραμμικό μοντέλο, έτσι ώστε να επιτρέπεται οι όροι του σφάλματος (error terms ) και οι τυχαίες επιδράσεις (random effects ) να εμφανίζουν συσχέτιση και μη σταθερή μεταβλητότητα. Παρέχει, επομένως τη δυνατότητα να διαμορφώσει όχι μόνο τη μέση τιμή της μεταβλητής απόκρισης, αλλά και τη δομή συνδιακύμανσης του.

Επίσης οι **παράγοντες (factor)** και οι **συμμεταβλητές (covariates)** θεωρείται ότι **έχουν γραμμική σχέση με την εξαρτημένη μεταβλητή**.

Οι κατηγορικές μεταβλητές (categorical predictors) μπορούν να επιλεγθούν ως παράγοντες στο μοντέλο. Πρόκειται για μια ανεξάρτητη μεταβλητή που ορίζει μια ομάδα περιπτώσεων. Κάθε τιμή του παράγοντα μπορεί να έχει μια διαφορετική γραμμική επίδραση στην τιμή της εξαρτημένης μεταβλητής.

Οι παράγοντες διαχωρίζονται σε δύο κατηγορίες :

- **Παράγοντες σταθερών επιδράσεων ( fixed effect factors )** : Μεταβλητές των οποίων οι τιμές που ενδιαφέρουν παρουσιάζονται όλες στον πίνακα δεδομένων.
- **Παράγοντες τυχαίων επιδράσεων ( random effects factors )** : Μεταβλητές των οποίων οι τιμές στον πίνακα δεδομένων μπορούν να θεωρηθούν ως ένα τυχαίο δείγμα ενός μεγαλύτερου πληθυσμού τιμών.

Συμμεταβλητές ορίζονται οι συνεχείς μεταβλητές ( scale predictos ), όπως για παράδειγμα το εισόδημα μετρούμενο σε χιλιάδες δολάρια ή η ηλικία σε χρόνια. Σε συνδυασμούς με τις τιμές των παραγόντων, οι τιμές των συμμεταβλητών θεωρείται ότι είναι γραμμικά συσχετισμένες με τις τιμές της εξαρτημένης μεταβλητής.

Επιπλέον, το γραμμικό μικτό μοντέλο επιτρέπει τον προσδιορισμό των αλληλεπιδράσεων των παραγόντων, γεγονός που σημαίνει ότι κάθε συνδυασμός των τιμών των παραγόντων έχει διαφορετική γραμμική επίδραση στην εξαρτημένη μεταβλητή. Είναι επίσης δυνατός ο προσδιορισμός των αλληλεπιδράσεων των παραγόντων και των συμμεταβλητών, εάν υπάρχει η πεποίθηση ότι η γραμμική σχέση μεταξύ της συμμεταβλητής και της εξαρτημένης μεταβλητής αλλάζει ανάλογα με τις τιμές του παράγοντα.

Τέλος με τη διαδικασία του γραμμικού μοντέλου, όταν περιλαμβάνονται μεταβλητές επαναλαμβανόμενων επιδράσεων (repeated effects variables), επιτρέπεται ο προσδιορισμός της δομής της συνδιακύμανσης των σφαλμάτων. Για να συμβεί αυτό θα πρέπει να προσδιοριστούν τα ακόλουθα:

- Μεταβλητές επαναλαμβανόμενων επιδράσεων ορίζονται οι μεταβλητές των οποίων οι τιμές στον πίνακα δεδομένων μπορούν να θεωρηθούν ως δείκτες πολλαπλών παρατηρήσεων ενός μόνο υποκειμένου (subject).
- Οι μεταβλητές-υποκείμενα ορίζουν τα μεμονωμένα υποκείμενα των επαναλαμβανόμενων μετρήσεων. Οι όροι σφάλματος κάθε μεμονωμένου υποκειμένου είναι ανεξάρτητοι από αυτούς των άλλων μεμονωμένων υποκειμένων.
- Η δομή της συνδιακύμανσης (covariance structure) προσδιορίζει τη σχέση μεταξύ των τιμών της μεταβλητής επαναλαμβανόμενων επιδράσεων.

### 3.5.5 Συντελεστής προσδιορισμού (Coefficient of determination)

Ο συντελεστής προσδιορισμού  $R^2$  υπολογίζει το ποσοστό διακύμανσης της εξαρτημένης μεταβλητής η οποία ερμηνεύεται από τις ανεξάρτητες μεταβλητές του μαθηματικού προτύπου.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Όπου :

n: αριθμός παρατηρήσεων

$y_i$ : πραγματικές τιμές εξαρτημένης μεταβλητής Y

$\bar{y}$ : μέση τιμή της μεταβλητής Y

$\hat{y}_i$ : εκτιμημένες τιμές της μεταβλητής Y

Ο συντελεστής προσδιορισμού μετρά την ικανότητα των παραγόντων να ερμηνεύσουν ένα φαινόμενο και οι τιμές του κυμαίνονται από 0 έως 1. Το βέλτιστο μοντέλο, για το οποίο οι ανεξάρτητες μεταβλητές ερμηνεύουν 100% τη διακύμανση της εξαρτημένης μεταβλητής, έχει συντελεστή  $R^2$  ίσο με 1. Αντίθετα, όταν οι ανεξάρτητες μεταβλητές δεν μπορούν να ερμηνεύσουν καθόλου τη διακύμανση της εξαρτημένης μεταβλητής ο συντελεστής  $R^2$  είναι 0.

## 4. Συλλογή και επεξεργασία στοιχείων

### 4.1 Εισαγωγή

Κάθε μελέτη προκειμένου να εφαρμοστεί σωστά η οποιαδήποτε στατιστική μέθοδος χρειάζεται σωστά και καταλλήλως επεξεργασμένα δεδομένα. Η συλλογή των δεδομένων πρέπει να γίνει μεθοδικά και τα μεγέθη που συλλέγονται να είναι πλήρως ορισμένα, με επεξεργασία αυτών θα προκύψουν όλες οι ανεξάρτητες μεταβλητές μέσω των οποίων το μαθηματικό πρότυπο θα κάνει προσαρμογή. Στην συγκεκριμένη ενότητα παρουσιάζεται η συλλογή των δεδομένων, η επεξεργασία αυτών καθώς και περιγραφικά στατιστικά τα οποία παρέχουν μια πρώτη ιδέα για το ερευνητικό αντικείμενο ενώ συνεισφέρουν σημαντικά στην επιλογή κατάλληλου στατιστικού μοντέλου.

Η συλλογή και επεξεργασία των δεδομένων είναι ένα κομμάτι μεγάλης σημασίας της διπλωματικής εργασίας. Τα διάφορα δεδομένα πρώτα συλλέγονται από ορισμένες πηγές στη συνέχεια επεξεργάζονται είτε για να αντιμετωπιστούν προβληματικές μετρήσεις είτε για να αποκτήσουν κάποια φυσική ερμηνεία και τέλος γίνεται η σύνθεση αυτών σε ένα υπολογιστικό φύλο διαμορφώνοντας έτσι τις βάσεις δεδομένων.

### 4.2 Συλλογή δεδομένων

Αφετηρία έναρξης αποτελεί η μαζική συλλογή των δεδομένων τα οποία θα αξιοποιηθούν στη στατιστική μελέτη. Για την επίτευξη του στόχου της διπλωματικής εργασίας απαιτείται η συλλογή στοιχείων για μια συγκεκριμένη χρονική περίοδο που αφορούν τον **κυκλοφοριακό φόρτο** της Αθήνας, τη τιμή της **βενζίνης** στα πρατήρια αλλά και στοιχεία για παράγοντες που ενδεχομένως να επηρεάζουν τη διακύμανση της ανεξάρτητης μεταβλητής δηλαδή του κυκλοφοριακού φόρτου. Τέτοιοι παράγοντες για τους οποίους συλλέχθηκαν δεδομένα είναι το **ύψος βροχής, η θερμοκρασία**, τα κρούσματα της πανδημίας **Covid 19**. Προκειμένου να μελετηθεί η επιρροή της τιμής της βενζίνης χρειάζεται μια περίοδος έντονης μεταβολής της τιμής της αρκετή ώστε η επίδραση της ακρίβειας να μετακυλήσει στους οδηγούς μέσω του κόστους ανεφοδιασμού. Έτσι λοιπόν ως **περίοδος για τη μελέτη** ορίζεται από Νοέμβριο του 2021 έως Ιούλιο του 2022.

#### 4.2.1 Δεδομένα Κυκλοφοριακού φόρτου

Οι μετρήσεις του κυκλοφοριακού φόρτου συλλέχθηκαν από το Κέντρο Διαχείρισης Κυκλοφορίας (Κ.Δ.Κ.). Για το νομό της Αττικής λειτουργεί από τον Ιούλιο του 2004 αυτοματοποιημένο σύστημα διαχείρισης της κυκλοφορίας. Αυτό αποτελείται από 550 περίπου θέσεις μέτρησης κυκλοφοριακών δεδομένων, 217 κάμερες εποπτείας της κυκλοφορίας, 24 Πινακίδες Μεταβλητών Μηνυμάτων, σύστημα εποπτείας της κυκλοφορίας SITRAFFIC CONCERT και περίπου 850 ρυθμιστές κυκλοφορίας στους σηματοδοτούμενους κόμβους. Η μέτρηση

κυκλοφοριακών μεγεθών γίνεται είτε με απλούς επαγωγικούς βρόχους είτε με βρόχους «μηχανικής όρασης». Οι επαγωγικοί βρόγχοι είναι ουσιαστικά ένα πηνίο το οποίο διαρρέεται από ηλεκτρικό ρεύμα των οποίων η παρουσία οχημάτων μεταβάλλει τη συμπεριφορά και έτσι μετρούνται οι διελεύσεις των οχημάτων. Εδώ σημειώνεται πως οι βρόγχοι δεν διαχωρίζουν τα οχήματα σε ελαφρά και βαριά ούτε αντιλαμβάνονται τις διελεύσεις δίκυκλων επειδή για τα δίκυκλα απαιτείται ειδική διαμόρφωση του βρόχου. Ο κυκλοφοριακός φόρτος καταμετράται ως **διελεύσεις οχημάτων ανά ώρα** για κάθε μετρητή. Η παραπάνω πληροφορία δίνεται μαζί με την γεωγραφική θέση του κάθε μετρητή με τη μορφή γεωγραφικών συντεταγμένων και μια σύντομη περιγραφή. Παρακάτω δίνονται ενδεικτικά πίνακες με το κυκλοφοριακό μέγεθος του ημερήσιου φόρτου όπως αυτά προκύπτουν από τους μετρητές καθώς και με τις γεωγραφικές πληροφορίες του κάθε μετρητή.

Deviceid	Road Name	Road Info	Latitude	Longitude
MS277	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΚΗΦΙΣΙΑ ΠΡΙΝ ΑΠΟ ΤΗΝ ΕΘΝΙΚΗΣ ΑΝΤΙΣΤΑΣΕΩΣ	38.00956	23.77977
MS278	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΠΡΟΣ ΚΕΝΤΡΟ ΠΡΙΝ ΑΠΟ ΤΗΝ ΕΘΝΙΚΗΣ ΑΝΤΙΣΤΑΣΕΩΣ	38.01027	23.78014
MS280	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΚΗΦΙΣΙΑ ΠΡΙΝ ΑΠΟ ΤΗΝ ΑΓ. ΦΙΛΟΘΕΗΣ	38.01708	23.78648
MS281	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΠΡΟΣ ΚΕΝΤΡΟ ΠΡΙΝ ΑΠΟ ΤΗΝ ΑΓ. ΦΙΛΟΘΕΗΣ	38.01778	23.78692
MS290	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΚΗΦΙΣΙΑ 120 Μ. ΠΡΙΝ ΑΠΟ ΤΗΝ ΚΑΡΕΛΛΑ (ΚΟΜΒΟΣ ΥΓΕΙΑ)	38.02687	23.78980
MS291	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΚΕΝΤΡΟ 120 Μ. ΠΡΙΝ ΑΠΟ ΤΗΝ ΕΡΥΘΡΟΥ ΣΤΑΥΡΟΥ (ΚΟΜΒΟΣ ΥΓΕΙΑ)	38.02787	23.79030
MS295	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΚΗΦΙΣΙΑ ΠΡΙΝ ΑΠΟ ΤΗ ΡΑΜΠΑ ΕΙΣΟΔΟΥ ΣΤΗ Λ. ΚΗΦΙΣΙΑΣ ΤΗΣ ΣΠΥΡΟΥ ΛΟΥΗ/	38.03370	23.79559
MS300	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΚΕΝΤΡΟ 150 Μ. ΜΕΤΑ ΤΟΝ ΚΟΜΒΟ ΠΑΡΑΔΕΙΣΟΥ (ΜΕΤΑ ΤΗΝ ΟΔΟ ΧΩΜΑΤΙΑΝΟ	38.0366	23.79829
MS302	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΚΗΦΙΣΙΑ ΜΕΤΑ ΤΗ ΡΑΜΠΑ ΕΞΟΔΟΥ ΠΡΟΣ ΑΤΤΙΚΗ ΟΔΟ ΚΑΙ ΔΑΚΤΥΛΙΔΙ ΟΤΕ	38.04377	23.80487
MS308	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΚΕΝΤΡΟ ΜΕΤΑ ΤΗ ΡΑΜΠΑ ΕΞΟΔΟΥ ΤΗΣ ΠΑΡΑΔΕΙΣΟΥ/ΑΡΤΕΜΙΔΟΣ	38.0369	23.79864
MS315	Λ. ΚΗΦΙΣΙΑΣ	ΚΥΡΙΟΣ ΔΡΟΜΟΣ ΜΕ ΚΑΤΕΥΘΥΝΣΗ ΚΕΝΤΡΟ ΜΕΤΑ ΤΗ ΡΑΜΠΑ ΕΞΟΔΟΥ ΤΗΣ Λ. ΚΗΦΙΣΙΑΣ ΠΡΟΣ ΔΑΚΤΥΛΙΔΙ ΟΤΕ ΚΑ	38.04373	23.80469

Πίνακας 2: Πληροφορίες μετρητών.



Deviceid	Hour of Appprocesstime	17	18	19	20	21	22	23	24
MS277	0	3000	2600	1512	1712	1891	1498	1458	826
MS277	1	2850	1111	904	1237	1241	1207	952	1056
MS277	2	2056	593	505	694	694	694	504	975
MS277	3	1414	429	255	372	391	537	343	557
MS277	4	1173	354	269	317	367	432	286	348
MS277	5	1105	667	690	655	646	632	439	489
MS277	6	1153	2286	2234	2229	2140	998	852	631
MS277	7	1259	4281	4908	4417	3972	1255	994	450
MS277	8	1381	4900	4909	4823	4475	1531	1341	641
MS277	9	1845	4544	4429	4459	4409	1866	1763	945
MS277	10	2787	4624	4866	4519	4344	2506	2124	1369
MS277	11	3910	4353	4538	4309	4390	3075	2510	2193
MS277	12	4358	4302	4538	3597	4278	3510	2675	3009
MS277	13	4307	4372	4361	4249	4211	3477	2562	2550
MS277	14	4394	4424	4467	4301	3776	2999	2557	1660
MS277	15	3704	4129	4072	4346	3801	2128	1986	851
MS277	16	3317	4269	4382	4258	3908	2247	1691	873
MS277	17	3521	4578	4391	4149	3887	2233	1496	1187
MS277	18	3553	4777	4495	4380	3659	2388	1502	1320
MS277	19	3329	4873	4615	4153	3294	2098	1530	1537
MS277	20	3327	4334	4414	3974	3034	1648	1525	1631
MS277	21	3164	4142	4338	4301	2790	1615	1624	1823
MS277	22	2912	3071	3060	3013	2264	1842	1406	1444
MS277	23	2452	1972	2087	2502	1760	1477	799	1003

Πίνακας 3: Δεδομένα ωριαίων κυκλοφοριακών φόρτων.

Οι μετρητές ενώ γενικότερα συλλέγουν αξιόπιστα δεδομένα υπάρχουν ώρες είτε ολόκληρες ημέρες κατά τις οποίες υπολειπονται ή δεν λειτουργούν καθόλου με αποτέλεσμα να εμφανίζονται σημαντικά μειωμένες μετρήσεις οι οποίες απέχουν από το μέσο όρο αρκετά ή δεν καταγράφουν καμία διέλευση αφήνοντας έτσι την ώρα ή ημέρα κενή. Η ανασκόπηση στη βιβλιογραφία έδειξε πως το παραπάνω είναι αρκετά συνηθισμένο και υπάρχουν δύο σχολές σκέψης για την αντιμετώπιση του προβλήματος. Οι ερευνητές είτε διαγράφουν εντελώς από το δείγμα τις ημέρες όπου υπάρχουν προβληματικές μετρήσεις είτε χρησιμοποιούν τεχνικές προκειμένου να συμπληρώσουν τις ελλειπείς τιμές του δείγματος. Στη διπλωματική εργασία επιλέχθηκε η πρώτη λύση σύμφωνα με την οποία για κάθε προβληματικό μετρητή η προβληματική ημέρα έγινε αφαίρεση των μετρήσεων από το δείγμα. Έτσι λοιπόν τα δεδομένα που κρατάμε αφορούν **158 διαφορετικούς μετρητές για 209 ημέρες** εντός της περιόδου που αναφέρεται παραπάνω. Οι μετρητές αυτοί με τη σειρά τους ανήκουν σε **24 διαφορετικές οδούς της Αθήνας** εκ των οποίων κάποιες βρίσκονται εκτός δακτυλίου και κάποιες **εντός δακτυλίου**. Δίνεται αναλυτικός πίνακας με τους μετρητές της κάθε οδού.

Οδός	road_id
ΑΘΗΝΩΝ	1
ΑΜΑΛΙΑΣ	2
ΑΜΦΙΘΕΑΣ	3
ΑΧΑΡΝΩΝ	4
ΒΑΣ. ΚΩΝΣΤΑΝΤΙΝΟΥ	5
ΒΑΣ. ΣΟΦΙΑΣ	6
ΕΡΑΤΟΣΘΕΝΟΥΣ	7
ΗΠΕΙΡΟΥ	8
ΙΕΡΑ ΟΔΟΣ	9
ΚΑΛΛΙΡΡΟΗΣ	10
ΚΩΝΣΤΑΝΤΙΝΟΥΠΟΛΕΩΣ	11
Λ. ΑΘΗΝΩΝ	12
Λ. ΑΛΕΞΑΝΔΡΑΣ	13
Λ. ΚΗΦΙΣΙΑΣ	14
Λ. ΚΗΦΙΣΟΥ	15
Λ. ΚΩΝΣΤΑΝΤΙΝΟΥΠΟΛΕΩΣ	16
Λ. ΜΕΣΟΓΕΙΩΝ	17
Λ. ΣΥΓΓΡΟΥ	18
ΛΕΝΟΡΜΑΝ	19
ΜΙΧΑΛΑΚΟΠΟΥΛΟΥ	20
ΠΕΙΡΑΙΩΣ	21
ΠΕΤΡΟΥ ΡΑΛΛΗ	22
ΣΤΑΔΙΟΥ	23
ΧΑΜΟΣΤΕΡΝΑΣ	24

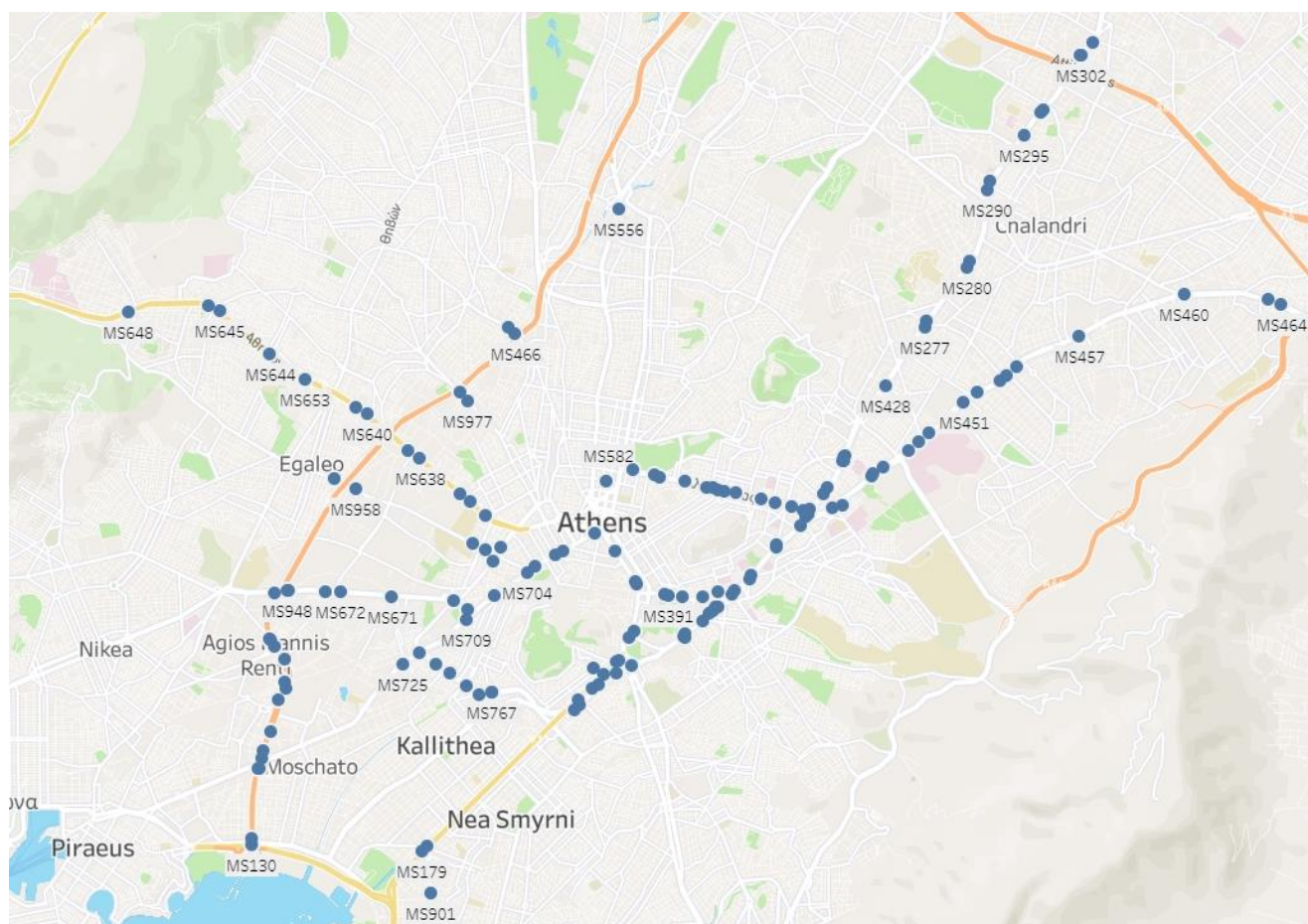
Πίνακας 4 : Αντιστοιχία αριθμών σε οδούς.

Καθώς στο προγραμματιστικό περιβάλλον R δεν είναι βολική η χρήση χαρακτήρων για τις διάφορες ομάδες δεδομένων γίνεται προσθήκη ορισμένων ονομαστικών μεταβλητών (nominal) οι οποίες αφορούν χαρακτηριστικά του δείγματος για τα οποία δεν υπάρχει κάποια συγκεκριμένη διάταξη απλώς αποτελούν ομάδες με διαφορετικά μεταξύ τους χαρακτηριστικά . Η παραπάνω διαδικασία γίνεται αντίστοιχα και για τους μετρητές δίνοντας από μια αριθμητική τιμή στον καθένα. Όπως αναλύεται και σε επόμενη ενότητα αρκεί λοιπόν να ορίσουμε τις μεταβλητές τέτοιου τύπου στο προγραμματιστικό περιβάλλον ως παράγοντα προκειμένου να μην αντιμετωπίζεται ως αριθμητικό στοιχείο (numerical) αλλά ως παράγοντας (factor) ο οποίος με τη σειρά του δηλώνει συνθήκες.

Οδός	Μετρητές																
ΑΘΗΝΩΝ	MS633	MS634	MS636	MS637	MS638	MS639	MS640	MS641	MS644	MS645	MS646	MS648					
ΑΜΑΛΙΑΣ	MS356	MS358															
ΑΜΦΙΘΕΑΣ	MS899	MS901	MS904	MS905													
ΑΧΑΡΝΩΝ	MS556																
ΒΑΣ. ΚΩΝΣΤΑΝΤΙΝΟΥ	MS219	MS237	MS238	MS239	MS240	MS241	MS242	MS243	MS244								
ΒΑΣ. ΣΟΦΙΑΣ	MS248	MS249	MS251	MS252	MS253	MS254	MS255	MS256	MS390	MS391	MS392	MS393	MS394				
ΕΡΑΤΟΣΘΕΝΟΥΣ	MS318																
ΗΠΕΙΡΟΥ	MS582																
ΙΕΡΑ ΟΔΟΣ	MS656	MS659	MS957	MS958													
ΚΑΛΥΠΡΟΗΣ	MS223	MS224	MS226	MS227	MS228	MS229	MS230	MS767	MS768	MS769							
ΚΩΝΣΤΑΝΤΙΝΟΥΠΟΛΕΩΣ	MS657	MS658															
Λ. ΑΘΗΝΩΝ	MS653																
Λ. ΑΛΕΞΑΝΔΡΑΣ	MS257	MS407	MS413	MS414	MS415	MS416	MS417	MS418	MS419	MS421	MS423	MS426	MS427				
Λ. ΚΗΦΙΣΙΑΣ	MS258	MS259	MS260	MS261	MS272	MS277	MS278	MS280	MS281	MS290	MS291	MS295	MS300	MS302	MS308	MS315	MS428
Λ. ΚΗΦΙΣΟΥ	MS106	MS108	MS110	MS112	MS114	MS116	MS118	MS120	MS122	MS124	MS126	MS127	MS128	MS130			
Λ. ΚΩΝΣΤΑΝΤΙΝΟΥΠΟΛΕΩΣ	MS466	MS970															
Λ. ΜΕΣΣΟΓΕΙΩΝ	MS435	MS437	MS440	MS443	MS448	MS449	MS451	MS452	MS454	MS456	MS457	MS459	MS460	MS462	MS463	MS464	MS486
Λ. ΣΥΓΓΡΟΥ	MS175	MS179	MS225	MS349	MS352	MS353	MS354										
ΛΕΝΟΡΜΑΝ	MS977	MS986															
ΜΙΧΑΛΑΚΟΠΟΥΛΟΥ	MS438																
ΠΕΙΡΑΙΩΣ	MS339	MS377	MS700	MS701	MS703	MS704	MS707	MS709	MS712	MS725							
ΠΕΤΡΟΥ ΡΑΛΛΗ	MS670	MS671	MS672	MS673	MS710	MS948	MS953	MS954									
ΣΤΑΔΙΟΥ	MS367	MS371	MS395														
ΧΑΜΟΣΤΕΡΝΑΣ	MS772	MS773															

Πίνακας 5: Μετρητές ανά οδό.

Στη συνέχεια ακολουθεί η χαρτογράφηση των παραπάνω μετρητών προκειμένου να γίνει κατανοητή η γεωγραφική τους θέση και άρα να γίνει αντιληπτό ποιο κομμάτι του κυκλοφοριακού φόρτου της Αθήνας μελετά η συγκεκριμένη διπλωματική εργασία.



Εικόνα 1: Γεωγραφική τοποθεσία μετρητών.

#### 4.2.2 Δεδομένα για την τιμή της βενζίνης

Σε περιόδους μεταβλητότητας στις τιμές των καυσίμων είναι συχνό φαινόμενο το κόστος ανά λίτρο να είναι αρκετά διαφορετικό για τους οδηγούς ανάλογα με την περιοχή που θα γίνει ο ανεφοδιασμός. Έτσι λοιπόν για τις τιμές της βενζίνης τα δεδομένα συλλέχθηκαν από το **Παρατηρητήριο Τιμών Υγρών Καυσίμων** του Υπουργείου Ανάπτυξης και Ανταγωνιστικότητας. Στην ιστοσελίδα της υπηρεσίας για κάθε μέρα του μήνα είναι διαθέσιμη η μέση τιμή της βενζίνης ανά νομό η οποία περιλαμβάνει στη τιμή και την φορολογία. Άρα η τιμή της βενζίνης δίνεται σε **ευρώ ανά λίτρο**. Για κάθε μέρα η υπηρεσία παρέχει ένα έντυπο στο οποίο αναγράφεται η τιμή ανά νομό για όλα τα καύσιμα, μέσω του PowerBi έγινε εισαγωγή όλων των εντύπων, ορίστηκε το πεδίο του οποίου τη τιμή συλλέγουμε, το οποίο είναι κοινό για όλα τα έντυπα και ύστερα δημιουργήθηκε ένα λογιστικό φύλλο με στήλη την τιμή της βενζίνης στον νομό της Αττικής για κάθε ημέρα της περιόδου Νοέμβριος 2021 – Ιούλιος 2022.

#### 4.2.3 Δεδομένα θερμοκρασίας και βροχόπτωσης

Για το ύψος της βροχόπτωσης δεδομένα συλλέχθηκαν από την ιστοσελίδα της οποίας η διεύθυνση δίνεται στη βιβλιογραφία, ορίζοντας το γεωγραφικό πλάτος και ύψος ή επιλέγοντας την πόλη παρέχει καταγραφές για κλιματικά δεδομένα όπου επιλέγεται η θερμοκρασία και το ύψος βροχής.

#### 4.2.4 Δεδομένα ενεργών κρουσμάτων Covid-19

Τα δεδομένα για τα ενεργά κρούσματα δηλαδή τον συνολικό αριθμό των ασθενών με Covid 19 ανά ημέρα προκύπτουν από ιστοσελίδα. Ο αριθμός αυτός δεν αναφέρεται στα θετικά τέστ ανίχνευσης κορονοϊού της ημέρας αλλά στο πόσοι άνθρωποι νοσούν ανά ημέρα για αυτό και τα νούμερα είναι αρκετά υψηλά.

#### 4.3 Επεξεργασία δεδομένων

Εν συνεχεία της συλλογής των δεδομένων ακολουθεί η επεξεργασία αυτών με στόχο τη σύνθεση της ολοκληρωμένης βάσης δεδομένων επί της οποίας θα εφαρμοστούν τα μαθηματικά πρότυπα. Πέρα από την διαχείριση για τις ελλιπείς τιμές προστίθονται ανεξάρτητες μεταβλητές οι οποίες δεν απαιτούν τη συλλογή δεδομένων και υπολογίζονται νέες ανεξάρτητες μεταβλητές οι οποίες θα εκφράσουν καλύτερα τα ήδη υπάρχοντα δεδομένα. Όπως αναφέρθηκε και παραπάνω οι προβληματικές μετρήσεις έχουν αφαιρεθεί από το δείγμα και πλέον υπάρχουν οι μετρήσεις για 209 ημέρες οι οποίες δεν είναι συνεχείς. Για κάθε ημερομηνία μέτρησης του κυκλοφοριακού φόρτου σε διπλανή στήλη προστίθεται για την ίδια ημερομηνία η μέση τιμή της βενζίνης.



Παρακάτω αναλύεται η κάθε μεταβλητή της βάσης δεδομένων και περιγράφεται ο τρόπος που αυτή υπολογίστηκε. Ενδεικτικά τυπώνεται η αρχή της βάσης δεδομένων για ευκολότερη ανάγνωση.

id	deviceId	Date	Traffic	Log_Volume	Weekend	Gasoline	Diesel	Ukraine_War	Gasoline_1weekbf	Gasoline_1monthbf	Public_Holidays	FuelPass	Rain	Covid	Temperature	Schools_Vacation	Log_Covid	Daktylios	FuelDelta	FuelDiff_Today	Yest	GasDiff_Today_Weekbf	road_id
1	MS633	11/1/2021	38.900	4.590	0	1.744	1.491	1	1.739	1.717	0	0	0.000	747.595	21.300	0	5.873666388	0	0.0000	0.000	0.000	0.0046	1
1	MS633	11/2/2021	34.884	4.543	0	1.745	1.491	1	1.741	1.719	0	0	2.800	754.451	23.100	0	5.877631039	0	0.0006	0.001	0.001	0.0040	1
1	MS633	11/3/2021	39.477	4.596	0	1.747	1.492	1	1.742	1.721	0	0	0.000	760.592	25.100	0	5.881151753	0	0.0011	0.002	0.001	0.0047	1
1	MS633	11/4/2021	40.579	4.608	0	1.747	1.492	1	1.743	1.723	0	0	0.000	767.376	25.000	0	5.885028212	0	0.0000	0.000	0.000	0.0036	1
1	MS633	11/5/2021	41.176	4.615	0	1.747	1.492	1	1.744	1.725	0	0	0.000	774.265	25.000	0	5.88889628	0	0.0000	0.000	0.000	0.0026	1
1	MS633	11/6/2021	35.499	4.550	1	1.746	1.491	1	1.745	1.727	0	0	0.000	780.621	24.200	0	5.89244023	0	-0.0006	-0.001	-0.001	0.0009	1
1	MS633	11/7/2021	36.870	4.567	1	1.745	1.490	1	1.746	1.729	0	0	0.000	784.904	23.300	0	5.894816542	0	-0.0006	-0.001	-0.001	-0.0006	1
1	MS633	11/8/2021	38.938	4.590	0	1.745	1.490	1	1.746	1.730	0	0	0.000	792.239	21.700	0	5.898856218	0	0.0000	0.000	0.000	-0.0009	1
1	MS633	11/9/2021	38.694	4.588	0	1.744	1.488	1	1.746	1.732	0	0	0.000	801.208	20.800	0	5.903745277	0	-0.0006	-0.001	-0.001	-0.0020	1
1	MS633	11/10/2021	38.779	4.589	0	1.742	1.487	1	1.746	1.733	0	0	0.000	808.297	17.700	0	5.907570967	0	-0.0011	-0.002	-0.002	-0.0039	1
1	MS633	11/11/2021	38.041	4.580	0	1.740	1.488	1	1.745	1.734	0	0	0.000	815.068	16.800	0	5.911193843	0	-0.0011	-0.002	-0.002	-0.0051	1
1	MS633	11/12/2021	39.739	4.599	0	1.738	1.487	1	1.744	1.735	0	0	0.000	821.900	18.800	0	5.91481898	0	-0.0011	-0.002	-0.002	-0.0061	1
1	MS633	11/13/2021	32.794	4.516	1	1.736	1.487	1	1.743	1.736	0	0	0.000	828.032	18.900	0	5.918047121	0	-0.0012	-0.002	-0.002	-0.0069	1
1	MS633	11/14/2021	27.996	4.447	1	1.735	1.486	1	1.741	1.736	0	0	0.000	831.874	19.200	0	5.920057551	0	-0.0006	-0.001	-0.001	-0.0064	1
1	MS633	11/15/2021	38.398	4.584	0	1.735	1.486	1	1.740	1.737	0	0	0.000	838.824	18.800	0	5.923670848	0	0.0000	0.000	0.000	-0.0050	1
1	MS633	11/16/2021	37.737	4.577	0	1.734	1.486	1	1.739	1.737	0	0	2.200	847.188	17.900	0	5.927979796	0	-0.0006	-0.001	-0.001	-0.0046	1
1	MS633	11/17/2021	32.617	4.513	0	1.734	1.486	1	1.737	1.738	0	0	0.000	853.841	16.700	0	5.931377005	0	0.0000	0.000	0.000	-0.0031	1
1	MS633	11/18/2021	41.089	4.614	0	1.733	1.486	1	1.736	1.738	0	0	0.000	861.117	18.800	0	5.935062163	0	-0.0006	-0.001	-0.001	-0.0030	1
1	MS633	11/19/2021	38.519	4.586	0	1.733	1.485	1	1.735	1.739	0	0	1.000	868.868	15.700	0	5.938953803	0	0.0000	0.000	0.000	-0.0020	1
1	MS633	11/20/2021	34.928	4.543	1	1.733	1.486	1	1.734	1.739	0	0	0.000	874.812	17.200	0	5.941914732	0	0.0000	0.000	0.000	-0.0013	1
1	MS633	11/21/2021	29.937	4.476	1	1.732	1.485	1	1.734	1.739	0	0	0.000	878.920	20.000	0	5.943949347	0	-0.0006	-0.001	-0.001	-0.0019	1
1	MS633	11/22/2021	38.385	4.584	0	1.732	1.486	1	1.733	1.739	0	0	0.000	886.207	21.300	0	5.947535176	0	0.0000	0.000	0.000	-0.0014	1
1	MS633	11/23/2021	37.160	4.570	0	1.731	1.483	1	1.733	1.739	0	0	5.100	894.555	21.400	0	5.951607047	0	-0.0006	-0.001	-0.001	-0.0020	1
1	MS633	11/24/2021	34.800	4.542	0	1.730	1.481	1	1.733	1.739	0	0	6.700	901.661	15.800	0	5.955043285	0	-0.0006	-0.001	-0.001	-0.0026	1

Πίνακας 6: Βάση δεδομένων.

- Η μεταβλητή **id** είναι μια ποιοτική μεταβλητή (dummy variable) στην οποία δίνεται μοναδικός αριθμός για κάθε μετρητή προκειμένου να μπορεί να ληφθεί ως παράγοντας στο προγραμματιστικό περιβάλλον της R.
- Η μεταβλητή **Volume** αντιστοιχεί στον **ημερήσιο κυκλοφοριακό φόρτο** για τον κάθε μετρητή και την κάθε μέρα σε μονάδες **διελεύσεις οχημάτων ανά ημέρα**. Υπολογίζεται αθροίζοντας όλες τις ωριαίες μετρήσεις του κυκλοφοριακού φόρτου.
- Η μεταβλητή **Log\_Volume** αναφέρεται στον λογάριθμο της μεταβλητής Volume, είναι δηλαδή ο λογάριθμος του κυκλοφοριακού φόρτου για κάθε μετρητή και ημέρα. Υπολογίστηκε καθώς υπάρχει πολύ μεγάλη διαφορά τάξης μεγέθους ανάμεσα στην εξαρτημένη μεταβλητή και την ανεξάρτητη μεταβλητή της βενζίνης.
- Προκειμένου να είναι αντιληπτό ποιες μετρήσεις αφορούν σαββατοκύριακα προστίθεται η ποιοτική μεταβλητή **Weekend** η οποία μέσω της φόρμουλας weekday του excel τυπώνει τη τιμή 0 για εργάσιμες και 1 για Σάββατο ή Κυριακή.
- Για τις τιμές της βενζίνης και του πετρελαίου κίνησης σε κάθε ημέρα αντιστοιχίζεται η εκάστοτε τιμή και προκύπτουν οι δυο μεταβλητές **Gasoline** και **Diesel**.
- Ακολουθεί άλλη μια ποιοτική μεταβλητή **Ukraine\_War** η οποία παίρνει τη τιμή 1 για τις ημερομηνίες που ο πόλεμος στην Ουκρανία υφίσταται.
- Υπολογίζεται ο μέσος όρος της τιμής της βενζίνης για ένα μήνα πριν από την εκάστοτε ημερομηνία αλλά και για επτά μέρες πριν, ύστερα αντιστοιχείται σε κάθε ημερομηνία ο μέσος όρος. Εκφράζεται έτσι η χρονική διαφορά ανάμεσα στην αντίληψη των οδηγών για το αυξημένο κόστος και την αλλαγή στις μετακινήσεις τους. Οι οδηγοί δεν

αντιλαμβάνονται αμέσως τις αυξήσεις στις τιμές της βενζίνης καθώς δεν ανεφοδιάζουν τα οχήματα τους κάθε μέρα αλλά και επειδή όσοι πρόκειται να μειώσουν τις μετακινήσεις τους χρειάζεται χρόνος για να αλλάξουν τις συνήθειες τους. Οι μεταβλητές αυτές είναι η **Gasoline\_1monthbf** και η **Gasoline\_1weekbf**

- Οι δημόσιες αργίες εκφράζονται μέσω της μεταβλητής **Public\_Holidays** και όταν πρόκειται για ήμερες αργιών παίρνει τιμή ίση με 1.
- Η κρατική επιδότηση για την αντιμετώπιση της ακρίβειας εκφράζεται μέσω της ποιοτικής μεταβλητής **FuelPass** όπου πάλι για τις περιόδους όπου χορηγήθηκε η μεταβλητή παίρνει τιμή ίση με 1.
- Το ύψος της βροχής για την εκάστοτε ημέρα δίνεται με την ποσοτική συνεχή μεταβλητή **Rain**.
- Τα ενεργά κρούσματα κορονοϊού δίνονται με την διακριτή ποσοτική μεταβλητή **Covid**.
- Οι σχολικές διακοπές προκύπτουν με τη ποιοτική μεταβλητή **Schools\_Vacation** η οποία παίρνει τη τιμή 1 όποτε ισχύουν σχολικές διακοπές ή αργίες.
- **Log covid** είναι ο λογάριθμος των ενεργών κρουσμάτων.
- Εάν ο εκάστοτε μετρητής βρίσκεται εκτός ή εντός του δακτυλίου εκφράζεται μέσω της ποιοτικής μεταβλητής **Daktlios** η οποία παίρνει τιμή 1 για μετρητές εντός.

Η ίδια βάση δημιουργείται επιπρόσθετα ξεχωριστά για μετρητές εκτός και εντός του δακτυλίου καθώς έχει ενδιαφέρον η διαφορά στα χαρακτηριστικά του κυκλοφοριακού φόρτου για αυτές τις δύο περιπτώσεις και μέσω της ανάλυσης προκύπτουν εξίσου ενδιαφέροντα συμπεράσματα.

#### 4.4 Περιγραφική στατιστική

Πριν από οποιαδήποτε στατιστική ανάλυση είναι αρκετά σημαντικό να υπολογιστούν κάποια περιγραφικά στατιστικά καθώς και να σχεδιαστούν ορισμένα διαγράμματα. Έτσι δίνεται στον ερευνητή μια καλύτερη ιδέα για τα δεδομένα που θα χρησιμοποιήσει στην ανάλυση του ενώ εμφανίζονται τάσεις ανάμεσα στην εξαρτημένη μεταβλητή και τις ανεξάρτητες αρκετά πριν γίνει οποιαδήποτε προσαρμογή μοντέλου. Προκύπτουν λοιπόν πληροφορίες για το δείγμα οι οποίες βοηθάνε αρκετά όχι μόνο στην εμπέδωση του φαινομένου που μελετάται αλλά και στην επιλογή του μαθηματικού προτύπου. Ακολουθεί σειρά από πίνακες και διαγράμματα μαζί με αναλυτική περιγραφή των υπολογισμών τους καθώς και συμπεράσματα που προκύπτουν από αυτά.

Road_id	Min	Median	Mean	Max	Μέσος όρος οδού/Μέσος όρος δείγματος
ΑΘΗΝΩΝ	10,143	37,966	39,100	82,161	1.30
ΑΜΑΛΙΑΣ	14,385	36,466	35,940	53,200	1.20
ΑΜΦΙΘΕΑΣ	3,432	21,851	19,881	33,565	0.66
ΑΧΑΡΝΩΝ	2,036	6,479	6,266	8,629	0.21
ΒΑΣ. ΚΩΝΣΤΑΝΤΙΝΟΥ	3,368	26,631	26,159	40,196	0.87
ΒΑΣ. ΣΟΦΙΑΣ	7,483	28,342	29,791	57,219	0.99
ΕΡΑΤΟΣΘΕΝΟΥΣ	769	2,957	2,928	4,164	0.10
ΗΠΕΙΡΟΥ	2,513	8,734	8,031	11,557	0.27
ΙΕΡΑ ΟΔΟΣ	3,977	12,364	12,271	19,273	0.41
ΚΑΛΛΙΡΡΟΗΣ	4,464	24,766	23,839	40,288	0.79
ΚΩΝΣΤΑΝΤΙΝΟΥΠΟΛΕΩΣ	6,184	16,636	16,341	22,569	0.54
Λ. ΑΘΗΝΩΝ	19,102	55,652	52,081	71,724	1.74
Λ. ΑΛΕΞΑΝΔΡΑΣ	5,968	23,319	23,155	37,916	0.77
Λ. ΚΗΦΙΣΙΑΣ	10,929	37,457	40,274	90,930	1.34
Λ. ΚΗΦΙΣΟΥ	8,201	54,675	54,436	97,729	1.81
Λ. ΚΩΝΣΤΑΝΤΙΝΟΥΠΟΛΕΩΣ	9,490	21,346	20,220	28,609	0.67
Λ. ΜΕΣΟΓΕΙΩΝ	7,404	29,938	29,419	52,879	0.98
Λ. ΣΥΓΓΡΟΥ	6,551	26,032	29,750	61,215	0.99
ΛΕΝΟΡΜΑΝ	5,592	14,752	14,069	18,890	0.47
ΜΙΧΑΛΑΚΟΠΟΥΛΟΥ	12,842	39,693	35,802	49,733	1.19
ΠΕΙΡΑΙΩΣ	2,494	17,073	17,339	31,864	0.58
ΠΕΤΡΟΥ ΡΑΛΛΗ	1,963	22,190	21,796	53,355	0.73
ΣΤΑΔΙΟΥ	8,060	21,511	22,470	38,540	0.75
ΧΑΜΟΣΤΕΡΝΑΣ	8,485	27,272	26,311	37,162	0.88

Πίνακας 7: Περιγραφικά στατιστικά ανά οδό.

Στον παραπάνω πίνακα δίνονται ορισμένα περιγραφικά στατιστικά για κάθε οδό όπως αυτά υπολογίστηκαν στο περιβάλλον της γλώσσας R . Ο υπολογισμός έγινε για κάθε οδό ξεχωριστά μέσω της βιβλιοθήκης **tapply** στην οποία ορίζεται ο παράγοντας που υποδεικνύει τις διαφορετικές ομάδες μετρήσεων δηλαδή στην συγκεκριμένη περίπτωση η ποιοτική μεταβλητή **road\_id**. Από τον πίνακα λοιπόν αντιλαμβανόμαστε πώς οι κυκλοφοριακοί φόρτοι ανάμεσα στις οδούς ποικίλουν αρκετά ενώ παράλληλα παρατηρώντας τα ελάχιστα και μέγιστα κάθε οδού γίνεται αντιληπτό το εύρος τιμών του φόρτου μιας οδού. Ενδεικτικά η οδός λεωφόρος Κηφισίας έχει μέσο ημερήσιο κυκλοφοριακό φόρτο 40.247 διελεύσεις οχημάτων ενώ η λεωφόρος

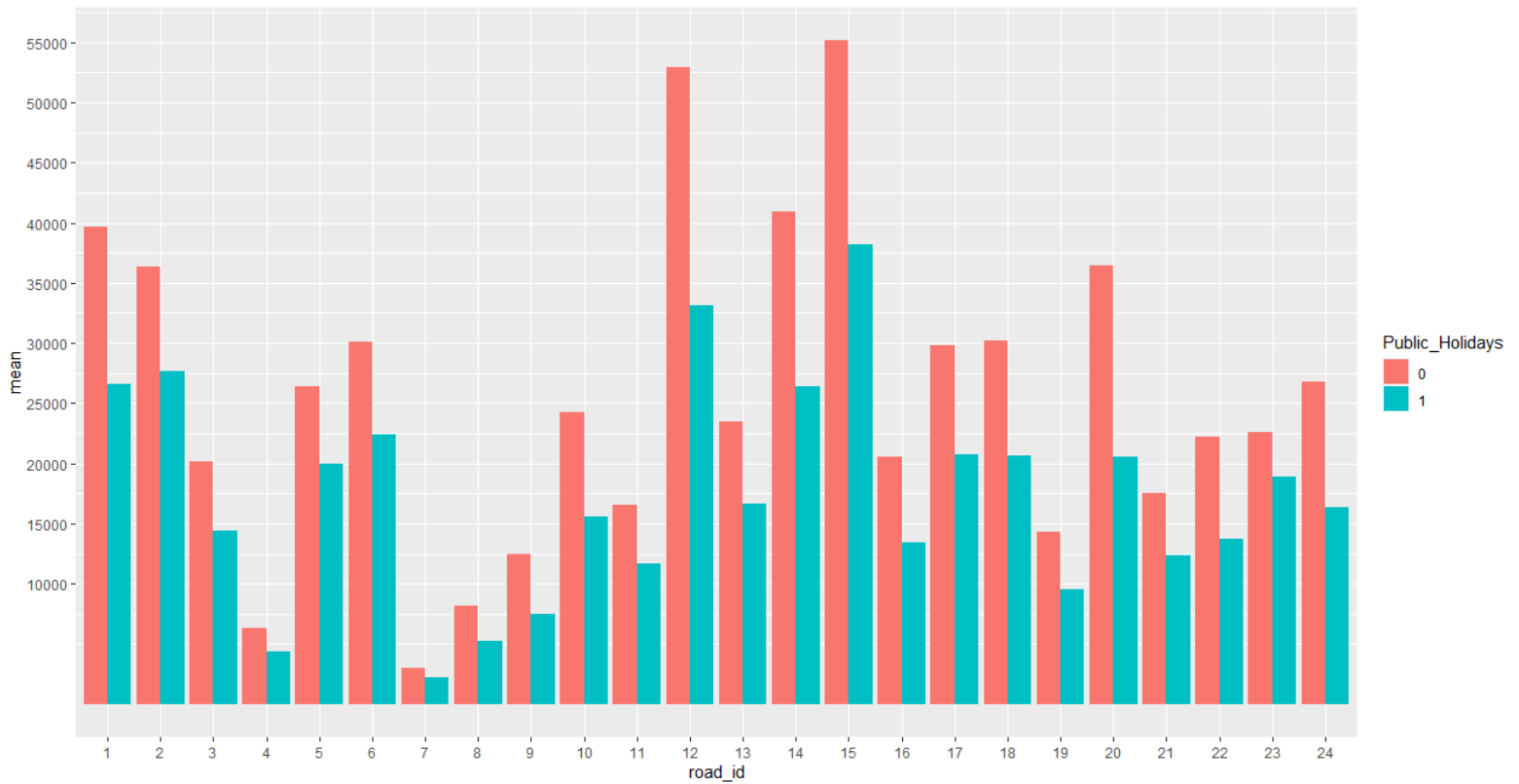
Μεσογείων έχει αντίστοιχα 29.750 διελεύσεις. Η τελευταία στήλη ισούται με πηλίκο του μέσου ημερήσιου κυκλοφοριακού φόρτου κάθε οδού προς τον ολικό μέσο ημερήσιο φόρτο.



Γράφημα 3: Μέσος όρος κυκλοφοριακών φόρτων ανά οδό για εργάσιμες και σαββατοκύριακα.

Το παραπάνω διάγραμμα προκύπτει στο προγραμματιστικό περιβάλλον R. Στον οριζόντιο άξονα τα νούμερα αντιπροσωπεύουν την κάθε οδό ανάλογα με το νούμερο που της αντιστοιχεί όπως αυτά ορίστηκαν σε προηγούμενη παράγραφο. Στον κάθετο άξονα βρίσκονται οι τιμές για τους μέσους όρους των κυκλοφοριακών φόρτων κάθε οδού. Τα χρώματα αντιπροσωπεύουν με κόκκινο χρώμα το μέσο ημερήσιο κυκλοφοριακό φόρτο για τις εργάσιμες ενώ με μπλε ανοιχτό σχεδιάζονται οι μέσοι όροι ημερήσιου κυκλοφοριακού φόρτου για τα σαββατοκύριακα. Σε όλες τις οδούς ο κυκλοφοριακός φόρτος είναι ανώτερος για τις εργάσιμες, σημαντική λεπτομέρεια είναι η απόκλιση που έχουν ορισμένες οδοί όπως για παράδειγμα η οδός Μιχαλακοπούλου για την οποία υπάρχει διαφορά άνω των 10.000 διελεύσεων ανά ημέρα ανάλογα με το εάν πρόκειται για εργάσιμη η σαββατοκύριακο.





Γράφημα 4: Μέσος όρος κυκλοφοριακών φόρτων ανά οδό για εργάσιμες και αργίες

Όπως και στο προηγούμενο γράφημα έτσι και εδώ παρατηρείται μείωση των μέσων όρων των κυκλοφοριακών φόρτων για τις ημέρες που είναι αργίες. Σημαντική είναι η απόκλιση ανάμεσα στις εργάσιμες και τις αργίες.

## 5. Ανάπτυξη μαθηματικών προτύπων

### 5.1 Εισαγωγή

Στο παρόν κεφάλαιο δίνεται αναλυτική περιγραφή της εφαρμογής της μεθοδολογίας για την ανάλυση του ερευνητικού αντικειμένου καθώς και παρουσιάζονται τα **αποτελέσματα** της διπλωματικής εργασίας. Ύστερα από την ανασκόπηση στη διεθνή βιβλιογραφία, την περιγραφή του θεωρητικού υποβάθρου για την ανάλυση των στοιχείων και την περιγραφή συλλογής και επεξεργασίας των στοιχείων, επιλέχθηκε η κατάλληλη τεχνική ανάλυσης.

Η μέθοδος που επιλέχθηκε για την στατιστική ανάλυση όπως ήδη παρουσιάστηκε αναλυτικά στο Κεφάλαιο 3 είναι τα **γραμμικά μικτά μοντέλα (linear mixed effects)**. Μέσω κατάλληλων στατιστικών ελέγχων τα διάφορα μοντέλα χαρακτηρίζονται στατιστικά σημαντικά ή μη στατιστικά σημαντικά δίνοντας έτσι μια καλύτερη εικόνα για την επιρροή των παραμέτρων στις κυκλοφοριακές συνθήκες της Αθήνας. Για την επίτευξη των παραπάνω έγινε χρήση ειδικών πακέτων της γλώσσας προγραμματισμού R όπως αυτά παρουσιάζονται αναλυτικά παρακάτω.

Επιπρόσθετα στο συγκεκριμένο κεφάλαιο παρουσιάζονται και περιγράφονται τα αποτελέσματα που προέκυψαν από την εφαρμογή των στατιστικών μοντέλων ενώ γίνεται και επεξήγηση αυτών προκειμένου να δοθούν στο επόμενο κεφάλαιο τα συμπεράσματα.

Σημαντικό τμήμα του κεφαλαίου καταλαμβάνει το εδάφιο που αφορά στην εμφάνιση των αποτελεσμάτων και διακρίνεται στις εξής τρεις φάσεις :

- Παρουσίαση των εξαγόμενων στοιχείων
- Περιγραφή των αποτελεσμάτων
- Εξήγηση των αποτελεσμάτων

Τέλος γίνεται ανάλυση της ευαισθησίας και της ελαστικότητας επί των μοντέλων ώστε να ερμηνευθεί πλήρως η σημασία της μεταβολής κάθε ανεξάρτητης μεταβλητής στην τιμή της εξαρτημένης μεταβλητής, δηλαδή στον **ημερήσιο κυκλοφοριακό φόρτο**.

#### 5.1.2 Λογισμικό R και βιβλιοθήκες

Η γλώσσα προγραμματισμού της R αποτελεί ένα από τα πιο συνηθισμένα εργαλεία στατιστικής ανάλυσης για ερευνητές. Αυτό οφείλεται αρχικώς στις δυνατότητες που παρέχει το περιβάλλον της γλώσσας ειδικά στη διαχείριση δεδομένων καθώς με απλές εντολές επιτρέπει την επεξεργασία μεγάλων βάσεων δεδομένων, την ποικιλία εντολών στατιστικής φύσεως στη βασική έκδοση (base R) (R Core Team, 2022). Σημαντικό ρόλο στη χρήση της γλώσσας προγραμματισμού R έχουν οι διάφορες βιβλιοθήκες που μπορεί να προσθέσει στον κώδικα ο κάθε ερευνητής οι οποίες παρέχουν αμέτρητες δυνατότητες δίχως να απαιτείται η σύνταξη κώδικα εκ νέου. Παρακάτω αναφέρονται συνοπτικά τα βασικά «πακέτα» που χρησιμοποιούνται κατά την εκπόνηση της συγκεκριμένης διπλωματικής εργασίας.

- **Library(readxl):** Βιβλιοθήκη με στόχο την εισαγωγή και αποθήκευση αρχείων τύπου Excel στο προγραμματιστικό περιβάλλον της R.(Hadley Wickham and Jennifer Bryan, 2022)
- **Library(lme4):** Βιβλιοθήκη μέσω της οποίας γίνεται η εφαρμογή των γραμμικών μικτών μοντέλων, επιτρέπει τον καθορισμό της φόρμουλας του μοντέλου, επιστρέφει το μοντέλο με τους συντελεστές ενώ καθορίζονται random και fixed effects.(Bates et al., 2015)
- **Library(lmerTest):** Βιβλιοθήκη για την διεξαγωγή στατιστικών τεστ και υπολογισμό της στατιστικής σημαντικότητας κάθε μεταβλητής για τα γραμμικά μοντέλα που δίνει το πακέτο lme4.(Kuznetsova et al., 2017)
- **Library(jtools):** Υπολογισμός του adjusted R<sup>2</sup> για κάθε μοντέλο που προκύπτει από την εφαρμογή της βιβλιοθήκης lme4.(Jacob A. Long, 2022)
- **Library(dplyr):** Διαχείριση βάσης δεδομένων.(Wickham H et al., n.d.)
- **Library(ggplot2):** Σχεδίαση ποικίλων γραφημάτων.(Hadley Wickham, 2016)

### 5.1.3 Εισαγωγή βάσης δεδομένων στο προγραμματιστικό περιβάλλον R

Η εισαγωγή της εκάστοτε βάσης δεδομένων όπως προέκυψε από το κεφάλαιο 4 γίνεται με τον παρακάτω κώδικα. Δηλώνουμε ποιο αρχείο τύπου excel θέλουμε να εισάγουμε προκειμένου να αποθηκευτεί στην μνήμη του περιβάλλοντος και να γίνει επί αυτού η εφαρμογή των στατιστικών μοντέλων. Πριν γράψουμε την εντολή για την εισαγωγή της βάσης στο περιβάλλον της R θα πρέπει να φορτώσουμε σε αυτό τις διάφορες βιβλιοθήκες που θα χρησιμοποιήσουμε.

```

1 #Πακέτα απαραίτητα για την στατιστική ανάλυση
2
3 library(lmerTest)
4 library(readxl)
5 library(jtools)
6 library(dplyr)
7 library(lme4)

```

Εικόνα 2: Βιβλιοθήκες προγραμματισμού

Για την εισαγωγή της βάσης δεδομένων αρκεί η αυτή να βρίσκεται στον ίδιο φάκελο με αυτόν που εργαζόμαστε και με την παρακάτω γραμμή κώδικα αποθηκεύεται στο περιβάλλον με όνομα df, ακρωνύμιο για τον όρο data frame.

```

26 #Importing database
27 df <- read_excel("Deviceid_Database.xlsx")

```

Εικόνα 3: Εισαγωγή βάσης δεδομένων στο προγραμματιστικό περιβάλλον R

Όπως αναφέρθηκε εκτενώς και στο κεφάλαιο 4 η βάση δεδομένων περιέχει αρκετές κατηγορικές μεταβλητές οι οποίες προκειμένου το εκάστοτε στατιστικό μοντέλο να τις αντιληφθεί και να μην τις λάβει ως συνεχείς θα πρέπει να ορίσουμε την κάθε στήλη τέτοιας μεταβλητής ως παράγοντα. Η παραπάνω διαδικασία γίνεται μέσω της βασικής R με την εντολή **as.factor**. Οι στήλες που θα οριστούν ως παράγοντες είναι :

- Weekend
- Vacations
- Ukraine War
- FuelPass
- Schools\_Vacation
- Id

```

27 #Factors
28
29 df$weekend=as.factor(df$weekend)
30 df$vacations=as.factor(df$vacations)
31 df$id=as.factor(df$id)
32 df$FuelPass=as.factor(df$FuelPass)
33 df$Schools_vacation=as.factor(df$Schools_vacation)
34 df$ukraine_war=as.factor(df$ukraine_war)

```

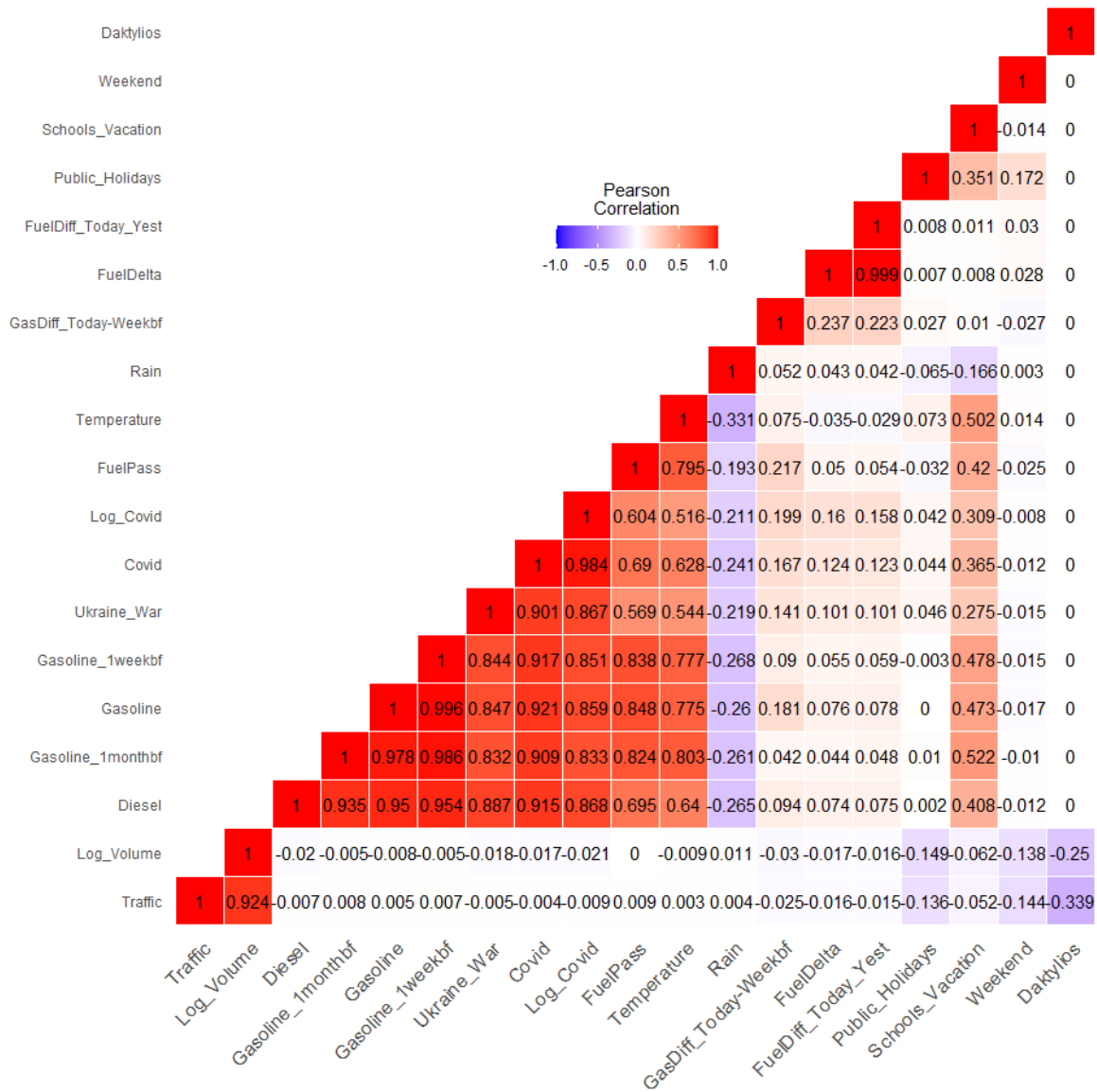
Εικόνα 4: Ορισμός ποιοτικών μεταβλητών ως παράγοντες

#### 5.1.4 Συσχέτιση μεταβλητών (Pearson Correlation)

Για την ανάπτυξη των στατιστικών μοντέλων παλινδρόμησης είναι απαραίτητο να διερευνηθεί η συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών. Όπως αναφέρεται και στο Κεφάλαιο 3 οι τιμές του συντελεστή συσχέτισης Pearson κυμαίνονται στο διάστημα [-1,1] και η συσχέτιση ανάλογα της τιμής του r χαρακτηρίζεται ως εξής :

- Υψηλή συσχέτιση :  $0.00 \leq |r| \leq 0.30$
- Μέτρια συσχέτιση :  $0.31 \leq |r| \leq 0.70$
- Υψηλή συσχέτιση :  $0.71 \leq |r| \leq 1.00$

Παρακάτω παρουσιάζονται τριγωνικοί θερμικοί χάρτες, με θερμό χρώμα απεικονίζεται η θετική συσχέτιση ενώ με ψυχρό χρώμα η αρνητική συσχέτιση. Ανάλογα την ένταση του χρώματος απεικονίζεται το μέγεθος της συσχέτισης ενώ επί του πίνακα αναγράφονται και οι τιμές του συντελεστή συσχέτισης ανάμεσα στις ανεξάρτητες μεταβλητές. Οι χάρτες υπολογίστηκαν για κάθε επιμέρους βάση ξεχωριστά.



Εικόνα 5: Θερμικός χάρτης συσχέτισης μεταβλητών

Οι πληροφορίες από τον παραπάνω χάρτη είναι ιδιαίτερα σημαντικές για την επιλογή των ανεξάρτητων μεταβλητών του μοντέλου. Επιλέγονται οι ανεξάρτητες μεταβλητές **GasDiff\_Today-weekbf, Weekend, Public\_Holidays, Temperature, Schools\_Vacation** οι οποίες καθώς δεν έχουν υψηλή συσχέτιση μπορούν να συνυπάρξουν σε ένα μοντέλο παλινδρόμησης δίχως να παρουσιάζεται **πολυσυγγραμμικότητα** η οποία θα μείωνε την ακρίβεια υπολογισμού των εκτιμώμενων συντελεστών και άρα την στατιστική ισχύ του μοντέλου.

## 5.2 Εφαρμογή στατιστικών μοντέλων

Όπως αναφέρθηκε και παραπάνω ύστερα από δοκιμές και ανασκόπηση στη βιβλιογραφία η συγκεκριμένη διπλωματική εργασία επιχειρεί να διερευνήσει την επιρροή της τιμής της βενζίνης στον κυκλοφοριακό φόρτο της Αθήνας εφαρμόζοντας **γραμμικά μικτά μοντέλα**. Η βιβλιοθήκη που χρησιμοποιείται για να εφαρμοστούν τα παραπάνω μοντέλα είναι η **lme4** στην οποία ορίζεται η φόρμουλα του μοντέλου, δηλαδή οι ανεξάρτητες μεταβλητές και η εξαρτημένη ενώ ορίζονται οι παράγοντες σταθερών (fixed effects factors) και οι παράγοντες τυχαίων επιδράσεων (random effects factors). Η βάση δεδομένων περιέχει μετρήσεις κυκλοφοριακού φόρτου από μετρητές σε διαφορετικές οδούς άρα όπως αναφέρεται στο κεφάλαιο 4 υπάρχει μεγάλη διακύμανση ανάμεσα στους κυκλοφοριακούς φόρτους που καταγράφει ο κάθε μετρητής. Επομένως, εφαρμόζεται ένα πολύ-επίπεδο μοντέλο έτσι λαμβάνονται υπόψιν οι συσχετίσεις ανάμεσα στις μετρήσεις του ίδιου μετρητή και διαχωρίζονται από αυτές ανάμεσα στους μετρητές. Λαμβάνοντας υπόψιν τα παραπάνω εφαρμόζεται το μοντέλο των τυχαίων σταθερών ( random intercept model ) όπου ενδεικτικά η εξίσωση για j ομάδες του δείγματος είναι της μορφής:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + \varepsilon_{ij}$$

- Ολική μέση τιμή για όλους τους μετρητές «ομάδες» (  $\beta_0$  )
- Συντελεστής παλινδρόμησης ( $\beta_1$ ) κοινός για όλες τις ομάδες
- Σφάλμα παλινδρόμησης (υπόλοιπο)
- Υπόλοιπο επιπέδου ομάδας  $u_j$

Αν τώρα θεωρήσουμε πως ο σταθερός όρος για μια ομάδα ισούται με  $\beta_{0j} = \beta_0 + u_j$  τότε προκύπτει το μοντέλο τυχαίων σταθερών ως :

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + \varepsilon_{ij}$$

Η παραπάνω επιλογή επιβεβαιώνεται με έναν **έλεγχο λόγου πιθανοφάνειας** όπου συγκρίνουμε το **μηδενικό μοντέλο (null model)** αυτό δηλαδή που δεν εμπεριέχει τυχαίες επιδράσεις ανά μετρητή με το μικτό μοντέλο το οποίο μέσω του τυχαίου σταθερού όρου **ομαδοποιεί τα δεδομένα ανά μετρητή (data clustering)**.

Οι έλεγχοι των σταθερών επιδράσεων (test of fixed effects) γίνονται μέσω των t-test όπου με τον δείκτη  $t$  προσδιορίζεται η στατιστική σημαντικότητα των ανεξάρτητων μεταβλητών. Ο συντελεστής  $t$  τυπώνεται αυτόματα από τον αλγόριθμο του πακέτου lme4. Όπως αναφέρθηκε και στο Κεφάλαιο 3, όσο μεγαλύτερη είναι η τιμή του  $t$  ( κατά απόλυτη τιμή ), τόσο μεγαλύτερη είναι η επιρροή της συγκεκριμένης μεταβλητής στο τελικό αποτέλεσμα και η ανεξάρτητη μεταβλητή θεωρείται τότε στατιστικά σημαντική. Για τη συγκεκριμένη ανάλυση κριτήριο για την αποδοχή μιας ανεξάρτητης μεταβλητής ως στατιστικά σημαντική είναι η τιμή του  $t$  να είναι κατά απόλυτη τιμή μεγαλύτερη από  $t = 1,7$  . Επιπρόσθετα μέσω του πακέτου lmerTest δίνεται η δυνατότητα για τις ανεξάρτητες μεταβλητές του εκάστοτε μοντέλου να τυπωθεί το p-value μέσω της μεθόδου Satterthwaite.

Μετά τον έλεγχο στατιστική εμπιστοσύνης αξιολογείται η ποιότητα προσαρμογής του μοντέλου που καθορίζεται βάσει του συντελεστή προσαρμογής  $R^2$  ο οποίος τυπώνεται με τη βοήθεια του πακέτου `jtools` για κάθε μοντέλο.

Σημειώνεται ότι στα αποτελέσματα κάθε μοντέλου και άρα οι εκτιμήσεις των συντελεστών προκύπτουν από τη μέθοδο της **μεγιστοποίησης της πιθανοφάνειας (maximum likelihood estimation «ML»)** σε αντίθεση με την προκαθορισμένη επιλογή της συνάρτησης `lmer`. Σύμφωνα με το **κριτήριο του λόγου πιθανοφάνειας** υπολογίζεται η διαφορά  $LRT=2(\text{LogLik}(1)-\text{LogLik}(0))$  και εάν είναι μεγαλύτερη από την τιμή του κριτηρίου  $\chi^2$  για  $n$  βαθμούς ελευθερίας σε επίπεδο σημαντικότητας 5% τότε το μοντέλο είναι στατιστικά προτιμότερο από το μοντέλο χωρίς τις μεταβλητές (0) και γίνεται αποδεκτό. Παράλληλα συνάρτηση `lme4` τυπώνει και άλλα μεγέθη που θα μπορούν να ληφθούν στην επιλογή μοντέλου αλλά αυτό εξετάζει κατά περίπτωση, αυτά είναι το AIC, BIC όπου προτιμάται το μοντέλο με το μικρότερο αυτών.

Τέλος πριν παρουσιαστούν τα μοντέλα που επιλέχθηκαν σημειώνεται πως κατά την εκπόνηση της συγκεκριμένης διπλωματικής εργασίας έγιναν πολλές δοκιμές ανάμεσα σε ανεξάρτητες μεταβλητές οι οποίες μεταβλητές υπάρχουν όλες στη βάση δεδομένων όπως αυτή παρουσιάζεται στο Κεφάλαιο 4. Τα μοντέλα που πέτυχαν στατιστική σημαντικότητα και έδωσαν ορθά αποτελέσματα ήταν αρκετά, όμως βασικός στόχος της συγκεκριμένης ανάλυσης είναι να υπολογιστεί ένα μοντέλο που να μπορεί να προσαρμοστεί στα δεδομένα ανεξάρτητα εάν αυτά προκύπτουν από μετρητές εκτός η εντός του δακτυλίου. Μάλιστα το παραπάνω στοιχείο είναι που κάνει τη συγκεκριμένη έρευνα πιο περίπλοκη αφού υπάρχει αρκετά μεγάλη διακύμανση ανάμεσα στις μετρήσεις της κάθε συσκευής μέτρησης εξαιτίας της διαφορετικής της τοποθεσίας σε ολόκληρη την Αθήνα.

#### Μοντέλο A: Επιρροή της τιμής της βενζίνης στην κυκλοφορία της ευρύτερης περιοχής της Αθήνας

Παρουσιάζεται το μοντέλο που εφαρμόζεται σε ολόκληρη τη βάση δεδομένων, αυτή δηλαδή που περιέχει όλους τους μετρητές υπό εξέταση στην ευρύτερη περιοχή της Αθήνας, δίχως να γίνεται διαχωρισμός ανάμεσα σε αυτούς που βρίσκονται εντός η εκτός Δακτυλίου. Η πρώτη εφαρμογή γίνεται για τις ανεξάρτητες μεταβλητές: **GasDiff\_Today-Weekbf, Weekend, Public\_Holidays, Schools\_Vacation**. Στο μοντέλο δίνεται τυχαία επίδραση μέσω του σταθερού όρου για κάθε διαφορετικό μετρητή προσθέτοντας στην φόρμουλα τον όρο  $(1|id)$  έτσι ο αλγόριθμος της `lme4` θα προσαρμόσει ένα μοντέλο τυχαίων σταθερών όρων στα δεδομένα όπως περιγράφεται παραπάνω. Έτσι το μοντέλο υπολογίζει στην ουσία δύο διακυμάνσεις πρώτα την  $\sigma_u^2$  η οποία αναφέρεται στην ανεξήγητη διακύμανση του επιπέδου 2 και την  $\sigma_e^2$  η οποία αναφέρεται στην ανεξήγητη διακύμανση του επιπέδου 1. Με άλλα λόγια η πρώτη είναι η διακύμανση μεταξύ των μετρητών και η δεύτερη η διακύμανση μεταξύ της επαναλαμβανόμενης μέτρησης.

Log_Volume ~ Weekend + Public_Holiday + Schools_Vacation + Temperature +GassDiff_Today-Weekbf+ (1   id)				
Fixed Effects				
	Estimate	Std.Error	t-value	p-value
Intercept	4.442	0.01692	262.62	0.000
Weekend1	-0.059	0.00084	-71.00	0.000
Public_Holidays1	-0.132	0.00202	-65.31	0.000
Schools_Vacations1	-0.017	0.00105	-15.88	0.000
Temperature	0.001	0.00006	10.59	0.000
GasDiff_Today-Weekbf	-0.333	0.01777	-18.73	0.000
Random Effects				
Groups	Variance	Std.Dev	R <sup>2</sup>	0.91
id	0.04499	0.21211		
Residual	0.004613	0.06792		

Πίνακας 1: Αποτελέσματα μοντέλου για την επιρροή της τιμής της βενζίνης στην κυκλοφορία της ευρύτερης περιοχής της Αθήνας

Μαθηματική σχέση Μοντέλου A

$$\begin{aligned} \text{Log\_Volume} = & 4.442 - 0.059 * (\text{Weekend}) - 0.132 * (\text{Public}_{\text{Holidays}1}) - 0.017 \\ & * (\text{Schools}_{\text{Vacations}1}) + 0.001 * (\text{Temperature}) - 0.333 \\ & * (\text{GassDiff}_{\text{Today}} - \text{Weekbf}) \end{aligned}$$

Ποιότητα Μοντέλου A

Στο στατιστικό μοντέλο που προέκυψε τηρούνται οι παρακάτω έλεγχοι :

- ✓ Ο **συντελεστής προσδιορισμού R<sup>2</sup>** προκύπτει αρκετά υψηλός, ίσος με 0,91. Άρα το μοντέλο ερμηνεύει τη διακύμανση της εξαρτημένης μεταβλητής , δηλαδή του κυκλοφοριακού φόρτου μέσω των ανεξάρτητων αυτών μεταβλητών κατά 91%.
- ✓ Οι ανεξάρτητες μεταβλητές είναι **στατιστικά σημαντικές** σύμφωνα με τον έλεγχο στατιστικής σημαντικότητας του κριτηρίου *t* της κατανομής student. Κάθε ανεξάρτητη μεταβλητή έχει δείκτη *t* αρκετά **μεγαλύτερο (κατά απόλυτη τιμή)** της κρίσιμης τιμής 1,7 για επίπεδο εμπιστοσύνης 0,95.
- ✓ Οι μεταβλητές που εισάγονται στο πρότυπο και τα πρόσημα τους εξηγούνται λογικά.

Σχολιασμός αποτελεσμάτων Μοντέλου A

Παρατηρώντας τους συντελεστές του μοντέλου προκύπτει μια σειρά από σημαντικές παρατηρήσεις για τον **κυκλοφοριακό φόρτο** της Αθήνας και τις επιδράσεις σε αυτόν.



- Η ανεξάρτητη μεταβλητή **Weekend1** αναφέρεται στην επίδραση που έχουν στον κυκλοφοριακό φόρτο τα **Σαββατοκύριακα**, όπως αναφέρθηκε και στο Κεφάλαιο 4 πρόκειται για κατηγορική μεταβλητή ( **dummy variable** ) και παίρνει τιμή 1 όταν η μέτρηση του κυκλοφοριακού φόρτου γίνεται εντός Σαββατοκύριακου. Η τιμή ίση με - 0,059 δηλώνει πώς τα σαββατοκύριακα ο λογάριθμος του φόρτου μειώνεται κατά 0,059.
- Παρόμοια επίδραση αλλά μεγαλύτερη κατά μέγεθος έχουν οι **αργίες**, ο συντελεστής της ανεξάρτητης μεταβλητής **Public\_Holidays** όπως είναι λογικό υποδεικνύει ότι οι κάτοικοι **μειώνουν** τις μετακινήσεις στην πόλη της Αθήνας κατά τις περιόδους αργιών, πιο συγκεκριμένα τις ημέρες αυτές λογάριθμος του κυκλοφοριακού φόρτου μειώνεται κατά 0,132.
- Για τις **σχολικές διακοπές** όπου να σημειωθεί στη μεταβλητή δεν συμμετέχουν μόνο οι κλασσικές καθορισμένες σχολικές διακοπές αλλά και **σχολικές αργίες**, οι καθημερινές μετακινήσεις που αφορούν μαθητές σχολείων δεν είναι λίγες για αυτό και κατά τις σχολικές διακοπές, αργίες ο λογάριθμος του φόρτου μειώνεται κατά 0,017
- Η **θερμοκρασία** αντιθέτως με όλους τους προηγούμενους παράγοντες έχει αυξητική επίδραση στον κυκλοφοριακό φόρτο, για αύξηση της θερμοκρασίας κατά ένα βαθμό κελσίου σημειώνεται αύξηση στον λογάριθμο του κυκλοφοριακού φόρτου κατά 0,001.
- Σύμφωνα με το μοντέλο η **αύξηση** της τιμής της **βενζίνης** επιδρά στατιστικά σημαντικά στον **κυκλοφοριακό φόρτο** της Αθήνας και μάλιστα τον **μειώνει**. Σύμφωνα με τον συντελεστή και το πρόσημο αυτού όταν η διαφορά μεταξύ της τιμής της εκάστοτε ημέρας και του μέσου όρου της τιμής της προηγούμενης εβδομάδας είναι θετική, άρα η τιμή της βενζίνης έχει ανέβει τότε παρατηρείται μείωση στον κυκλοφοριακό φόρτο. Αντίθετα όταν η διαφορά είναι αρνητική άρα και η τιμή της βενζίνης έχει μειωθεί τότε παρατηρείται αύξηση στον κυκλοφοριακό φόρτο. Αν δηλαδή η τιμή της βενζίνης ήταν 1 ευρώ ακριβότερη από τον μέσο όρο της τιμής της βενζίνης την προηγούμενη εβδομάδα θα είχαμε πτώση του λογαρίθμου του κυκλοφοριακού φόρτου κατά 0,333.

#### Σχετική επιρροή Μοντέλου Α

Σε συνέχεια του παραπάνω μοντέλου υπολογίζεται η σχετική επιρροή η αλλιώς ελαστικότητα όταν αναφερόμαστε σε ποσοστιαία επιρροή, για όλες τις ανεξάρτητες μεταβλητές του μοντέλου. Οι σχετικές επιρροές της κάθε ανεξάρτητης μεταβλητής για το συγκεκριμένο στατιστικό μοντέλο καθώς και η ανηγμένη σχετική επιρροή αυτών δίνονται στον ακόλουθο πίνακα.

Ανεξάρτητες μεταβλητές	$\beta_i$	t	Σχετική επιρροή	
			$e_i$	$e_i^*$
Weekend1	-0.059	-71.000	-0.1275913	3.38
Public_Holidays1	-0.132	-65.310	-0.2622657	6.94
Schools_Vacation1	-0.017	-15.880	-0.0377897	1.00
Temperature	0.001	10.590	0.0002959	-4.85
GasDiff_Today-Weekbf	-0.333	-18.730	-0.0000610	1.00

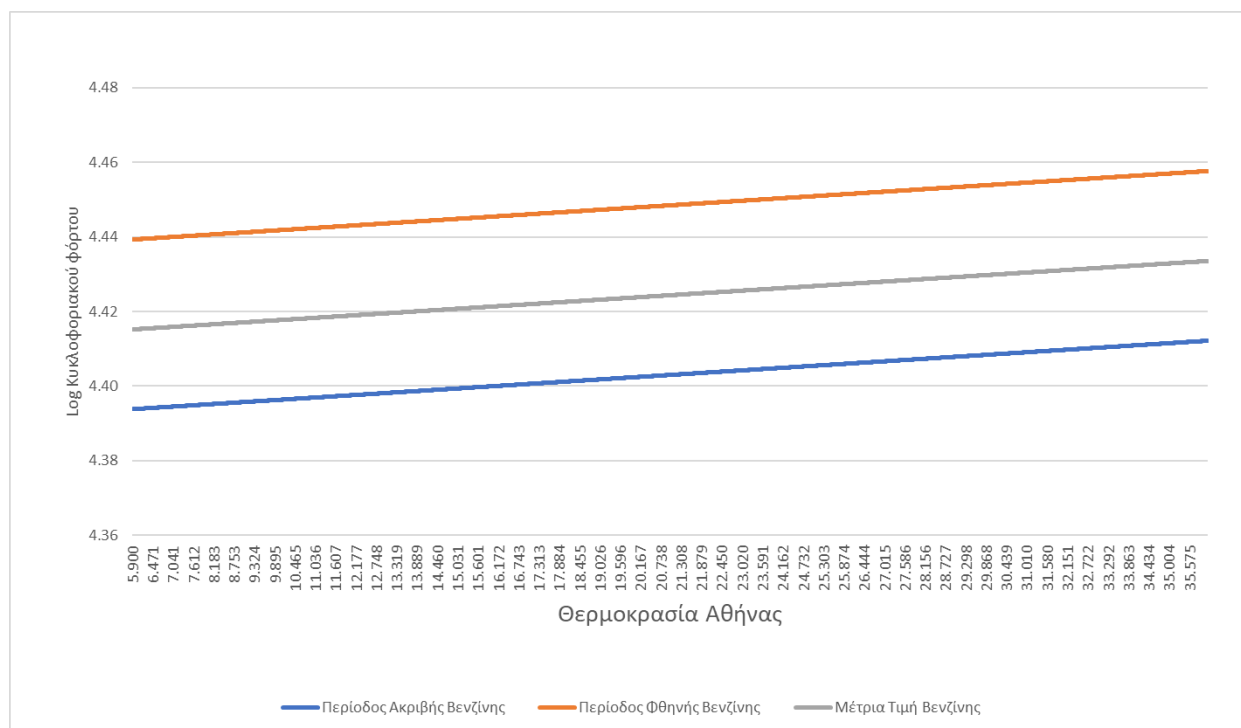
Πίνακας 2: Σχετική επιρροή ανεξάρτητων μεταβλητών για εφαρμογή στατιστικού μοντέλου στην ευρύτερη περιοχή της Αθήνας

Στον πίνακα βρίσκεται το είδος και το μέγεθος της επιρροής της κάθε ανεξάρτητης μεταβλητής για το συγκεκριμένο μαθηματικό πρότυπο. Ανηγγεμένη σχετική επιρροή είναι η  $e_i^*$  η οποία εκφράζει το βαθμό της σχετικής επιρροής της κάθε ανεξάρτητης μεταβλητής ως προς την ανεξάρτητη μεταβλητή που επηρεάζει κατά ελάχιστο την εξαρτημένη μεταβλητή. Τα παραπάνω αποτελέσματα παρουσιάζουν αρκετό ενδιαφέρον και δίνουν μια πιο σφαιρική ερμηνεία στους παράγοντες της ανάλυσης. Η σύγκριση στις σχετικές επιρροές γίνεται μόνο για ανεξάρτητες μεταβλητές του ίδιου είδους.

- Για τις ποσοτικές συνεχείς ανεξάρτητες μεταβλητές η επιρροή της θερμοκρασίας είναι μεγαλύτερη συγκριτικά με την επιρροή της τιμής της βενζίνης στον κυκλοφοριακό φόρτο. Πιο συγκεκριμένα για μια αύξηση της τάξης του 1% στην τιμή της θερμοκρασίας ο κυκλοφοριακός φόρτος της Αθήνας αυξάνει κατά 0,0296%
- Μπορεί η **θερμοκρασία** να έχει μεγαλύτερη επιρροή στον κυκλοφοριακό φόρτο συγκριτικά με την **τιμή της βενζίνης**, όμως η επιρροή της βενζίνης αν αναλογιστούμε πώς μια αύξηση 10% στη τιμή της βενζίνης εντός μια εβδομάδας προκαλεί συνολική μείωση κατά 0,06% του κυκλοφοριακού φόρτου δεν είναι αμελητέα. Βέβαια κατά τη σύγκριση με την επιρροή της θερμοκρασίας, η τιμή της βενζίνης επηρεάζει τον κυκλοφοριακό φόρτο κατά 4.85 φορές λιγότερο.
- Η μεταβλητή **Public\_Holidays** παρουσιάζει τη μεγαλύτερη επιρροή σχετικά με τις υπόλοιπες ανεξάρτητες μεταβλητές ίδιου είδους. Με βάση τη τιμή της σχετικής επιρροής ο κυκλοφοριακός φόρτος μειώνεται περίπου κατά το 1/3 τις ημέρες που είναι αργίες.
- Η μεταβλητή των **σχολικών διακοπών** εμφανίζει τη μικρότερη επιρροή της κατηγορίας, σε σχέση με την προηγούμενη μεταβλητή εμφανίζει 6,94 φορές μικρότερη επιρροή.
- Αρκετά μεγαλύτερη επιρροή από αυτή των σχολικών διακοπών παρουσιάζουν τα Σαββατοκύριακα και πιο συγκεκριμένα επηρεάζουν τον κυκλοφοριακό φόρτο κατά 3.38 φορές παραπάνω.

## Ανάλυση Ευαισθησίας

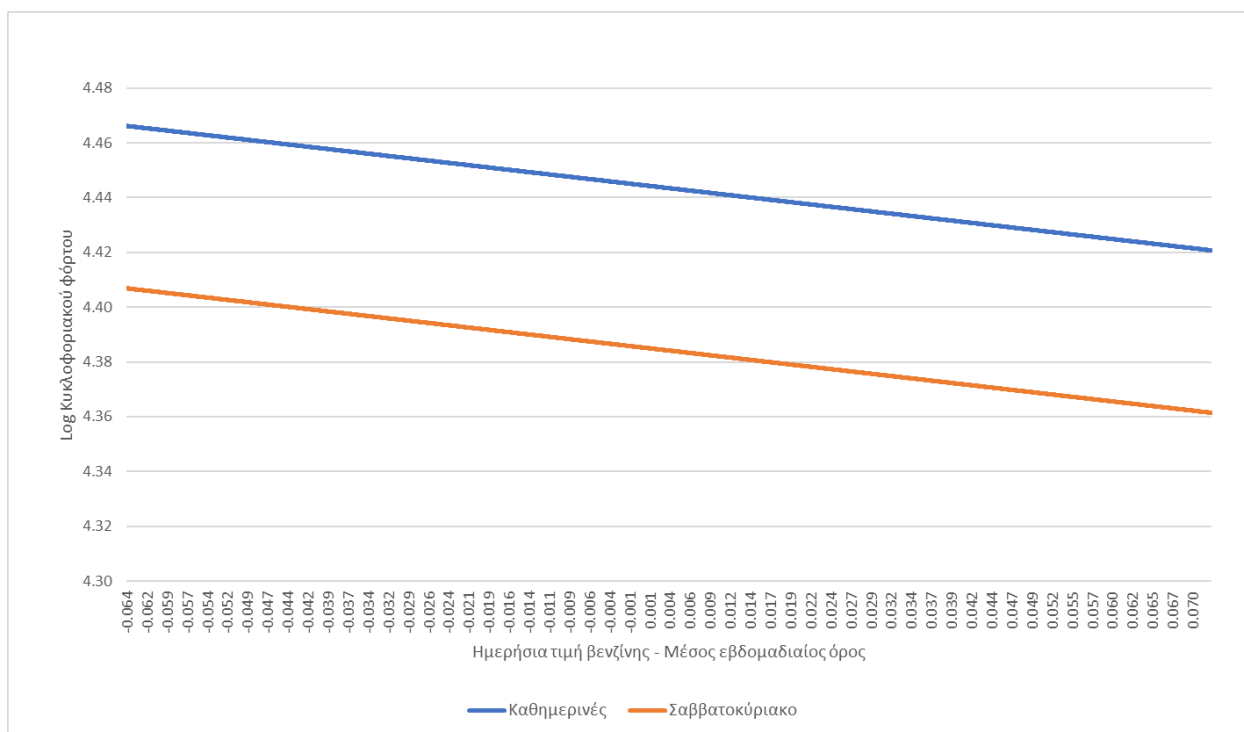
Με στόχο να γίνει πιο αντιληπτός ο τρόπος με τον οποίο οι ανεξάρτητες μεταβλητές επηρεάζουν την εξαρτημένη αναπτύχθηκαν τα διαγράμματα ευαισθησίας. Μέσω των διαγραμμάτων περιγράφεται η ευαισθησία της εξεταζόμενης εξαρτημένης μεταβλητής, στην συγκεκριμένη περίπτωση του λογαρίθμου του κυκλοφοριακού φόρτου όταν μεταβάλλεται κάποια από τις συνεχείς ανεξάρτητες μεταβλητές. Πιο συγκεκριμένα παρουσιάζονται γραφήματα που στον οριζόντιο άξονα υπάρχουν είτε τιμές της διαφοράς της βενζίνης όπως αυτή έχει οριστεί από το Κεφάλαιο 4, είτε τιμές της θερμοκρασίας. Στα διαγράμματα πέρα από τις συνεχείς ανεξάρτητες μεταβλητές εμφανίζονται και οι κατηγορικές καθώς για κάθε διαφορετική τιμή αυτών σχηματίζεται διαφορετική καμπύλη οι οποίες μιας και πρόκειται για καμπύλες του ίδιου μοντέλου έχουν κοινή κλίση σε κάθε διάγραμμα. Οι συνδυασμοί που μπορούν να γίνουν στα παραπάνω διαγράμματα για τις ανεξάρτητες μεταβλητές είναι αρκετοί οπότε επιλέχθηκαν ορισμένα τα οποία παρουσιάζουν ενδιαφέρον και έχουν εξηγηματικό χαρακτήρα.



Γράφημα 5: Ευαισθησία λογαρίθμου κυκλοφοριακού φόρτου για επίπεδα βενζίνης και άνοδο της θερμοκρασίας

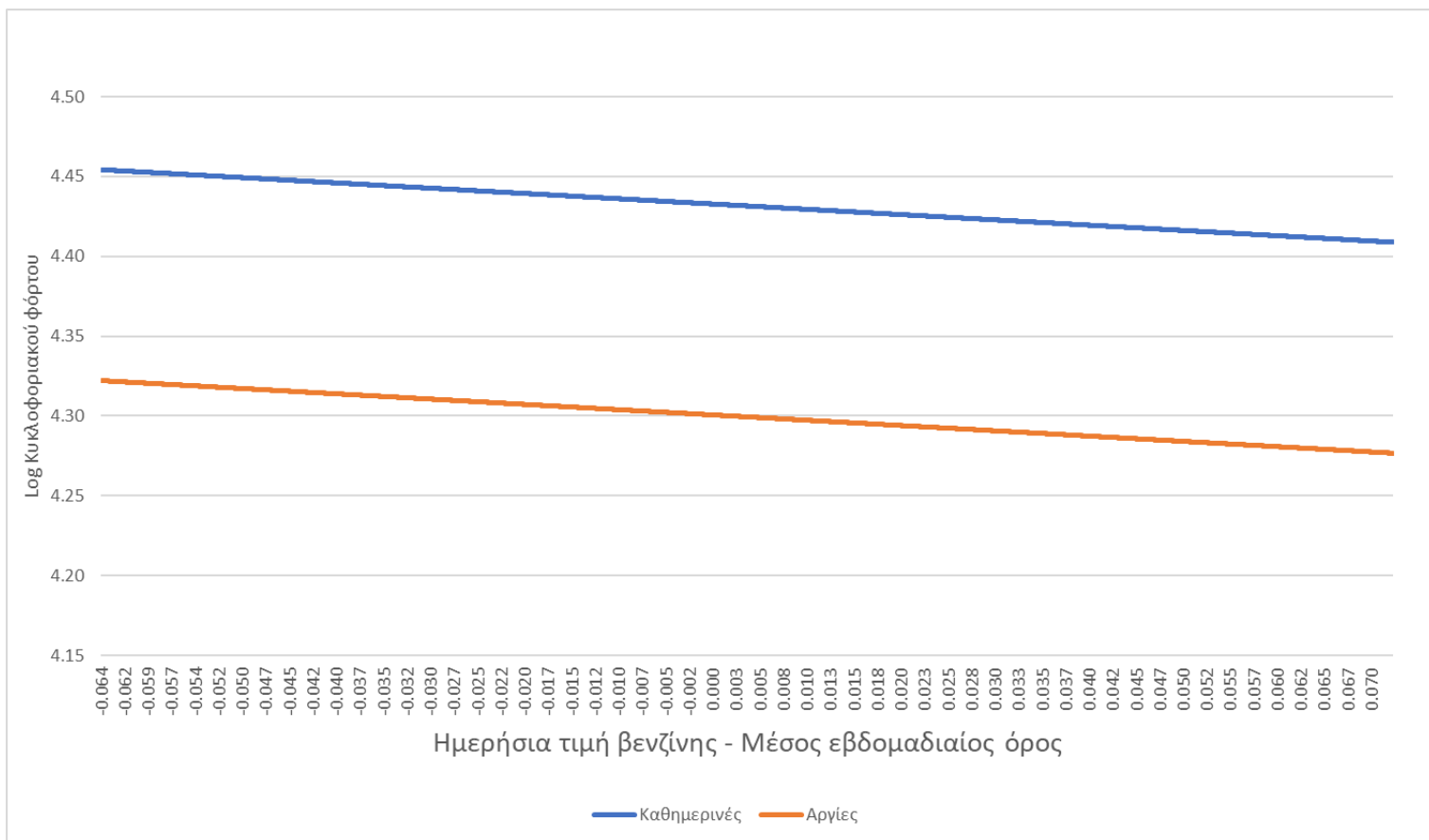
Για το παραπάνω διάγραμμα δημιουργήθηκαν 3 επίπεδα τιμών βενζίνης τα οποία εκφράζουν το καθένα μια διαφορετική περίοδο τιμολόγησης. Περίοδος ακριβής βενζίνης θεωρείται η

περίοδος όπου παρατηρείται μεγαλύτερη διαφορά ανάμεσα στην ημερήσια τιμή του καυσίμου και του μέσου όρου των προηγούμενων επτά ημερών. Ως περίοδος μέτριας τιμής της βενζίνης θεωρείται τέτοια ώστε η διαφορά να ισοδυναμεί με τον μέσο όρο των συνολικών διαφορών στη τιμή της βενζίνης ενώ περίοδος φθηνής βενζίνης θεωρείται περίοδος όπου η διαφορά μεταξύ της ημερήσιας και του εβδομαδιαίου μέσου όρου ελαχιστοποιείται. Έτσι προκύπτει πώς όσο αυξάνεται η θερμοκρασία ο κυκλοφοριακός φόρτος αυξάνει και φτάνει σε διαφορετικές τιμές ανάλογα την περίοδο τιμολόγησης της βενζίνης.



Γράφημα 6: Ευαισθησία λογαρίθμου κυκλοφοριακού φόρτου για εργάσιμες και σαββατοκύριακα με αύξηση τιμής της βενζίνης

Όσο αυξάνει προς τα θετικά η τιμή της βενζίνης τότε σημαίνει πώς η ημερήσια τιμή είναι πιο υψηλή από τον μέσο όρο της προηγούμενης εβδομάδας. Η αύξηση στη τιμή οδηγεί με τη σειρά της σε μείωση του κυκλοφοριακού φόρτου. Η διαφορά ανάμεσα στις δύο καμπύλες προκύπτει από τη διαφορά στους κυκλοφοριακούς φόρτους για τις εργάσιμες ημέρες και τα Σαββατοκύριακα. Οι εργάσιμες έχουν μεγαλύτερο κυκλοφοριακό φόρτο άρα το γράφημα ξεκινάει πιο ψηλά για την μπλε καμπύλη και σταδιακά μειώνεται όπως ακριβώς συμβαίνει για την πορτοκαλί καμπύλη η οποία αναφέρεται στον κυκλοφοριακό φόρτο για Σαββατοκύριακα.



Γράφημα 7: Ευαισθησία λογαρίθμου κυκλοφοριακού φόρτου για εργάσιμες και αργίες με αύξηση τιμής της βενζίνης

Αντίστοιχα στο παραπάνω γράφημα φαίνεται η μείωση του κυκλοφοριακού φόρτου για αυξητική τάση της τιμής της βενζίνης για δύο καμπύλες. Η πρώτη με χρώμα μπλε αναπαριστά τις μετακινήσεις για κανονικές ημέρες του χρόνου ενώ η πορτοκαλί αναφέρεται σε μετακινήσεις για τις ημέρες που είναι αργίες. Όπως είναι λογικό υπάρχει σημαντική διαφορά ανάμεσα στους κυκλοφοριακούς φόρτους για τις δύο αυτές περιπτώσεις ενώ και στις δύο παρατηρείται μείωση του κυκλοφοριακού φόρτου όσο η ημερήσια τιμή της βενζίνης είναι μεγαλύτερη από το μέσο όρο των προηγούμενων επτά ημερών.

## Εφαρμογή μοντέλου για τη βάση δεδομένων με μετρητές εκτός Δακτυλίου

Οι μετρητές έχουν ποικίλες τοποθεσίες με αποτέλεσμα κάθε μετρητής να καταγράφει διαφορετικούς κυκλοφοριακούς φόρτους. Η διαφορά αυτή δεν αφορά μόνο τις τιμές των μετρήσεων κατά απόλυτες τιμές αλλά η διαφορετική τοποθεσία δίνει και διαφορετικά κυκλοφοριακά χαρακτηριστικά. Η κυκλοφορία εκτός του δακτυλίου μπορεί να προϋποθέτει μετακινήσεις παραπάνω χιλιομέτρων και άρα οι οδηγοί να είναι πιο σκεπτικοί στο κόστος μετακίνησης. Στη συγκεκριμένη διπλωματική το μοντέλο εφαρμόζεται σε κάθε βάση δεδομένων ξεχωριστά προκειμένου να γίνει μια συγκριτική αξιολόγηση της επιρροής των ανεξάρτητων μεταβλητών στον κυκλοφοριακό φόρτο. Έτσι εφαρμόζεται το ίδιο μοντέλο σε διαφορετική βάση με τα παρακάτω αποτελέσματα.

Log_Volume ~ Weekend + Public_Holiday + Schools_Vacation + Temperature +GassDiff_Today-Weekbf+ (1   id)				
Fixed Effects				
	Estimate	Std.Error	t-value	p-value
Intercept	4.496	0.0247	181.72	0.000
Weekend1	-0.077	0.0011	-69.00	0.000
Public_Holidays1	-0.137	0.0027	-50.43	0.000
Schools_Vacations1	-0.020	0.0014	-14.04	0.000
Temperature	0.001	0.0001	9.69	0.000
GasDiff_Today-Weekbf	-0.354	0.0238	-14.89	0.000
Random Effects				
Groups	Variance	Std.Dev	R <sup>2</sup>	0.92
id	0.054251	0.23292		
Residual	0.004654	0.06822		

Πίνακας 3: Αποτελέσματα στατιστικού μοντέλου για εφαρμογή σε μετρητές εκτός Δακτυλίου.

Μαθηματική σχέση μοντέλου για μετρητές εκτός Δακτυλίου

$$\begin{aligned} \text{Log}_{\text{Volume}} = & 4.496 - 0.077 * (\text{Weekend}) - 0.137 * (\text{PublicHolidays}_1) - 0.02 \\ & * (\text{SchoolsVacations}_1) + 0.001 * (\text{Temperature}) - 0.354 \\ & * (\text{GassDiff}_{\text{Today}} - \text{Weekbf}) \end{aligned}$$

Ποιότητα μοντέλου για μετρητές εκτός Δακτυλίου

Στο στατιστικό μοντέλο που προέκυψε τηρούνται οι παρακάτω έλεγχοι :

- **Συντελεστής προσδιορισμού**  $R^2$  αρκετά υψηλός, άρα το μοντέλο ερμηνεύει τη διασπορά της εξαρτημένης μεταβλητής κατά 92%
- Στατιστικά σημαντικές ανεξάρτητες μεταβλητές, όλες οι ανεξάρτητες μεταβλητές έχουν δείκτη t μεγαλύτερο από 1.7.
- Οι μεταβλητές και τα πρόσημα τους έχουν λογική ερμηνεία.

Σχολιασμός αποτελεσμάτων μοντέλου για μετρητές εκτός Δακτυλίου

- Αρχικά καθώς αυξάνει η διαφορά ανάμεσα στην ημερήσια τιμή της **βενζίνης** και του μέσου όρου των προηγούμενων επτά ημερών μειώνεται ο κυκλοφοριακός φόρτος. Επίσης το αρνητικό πρόσημο του συντελεστή δίνει και την αντίθετη κατεύθυνση, αν δηλαδή η διαφορά είναι αρνητική άρα υπάρχει μείωση της ημερήσιας τιμής συγκριτικά με τις προηγούμενες επτά τότε ο κυκλοφοριακός φόρτος αυξάνει.
- **Τα σαββατοκύριακα** παρατηρείται μείωση του κυκλοφοριακού φόρτου, πιο αναλυτικά για τις ημέρες Σάββατο και Κυριακή ο λογάριθμος του κυκλοφοριακού φόρτου μειώνεται κατά 0,077
- **Οι αργίες** παρουσιάζουν μείωση του κυκλοφοριακού φόρτου, ενδεικτικά σε περιόδους αργιών ο λογάριθμος κυκλοφοριακού φόρτου μειώνεται κατά 0,137.
- Μείωση στον κυκλοφοριακό φόρτο προκύπτει και κατά τις περιόδους των **σχολικών διακοπών**.
- Η **θερμοκρασία** με τη σειρά της όσο αυξάνει, προκαλεί παραπάνω μετακινήσεις όπως προκύπτει από το θετικό συντελεστή της ανεξάρτητης μεταβλητής **Temperature**.

Σχετική επιρροή των ανεξάρτητων μεταβλητών για μετρητές εκτός Δακτυλίου

Ανεξάρτητες μεταβλητές	β <sub>i</sub>	t	Σχετική επιρροή	
			e <sub>i</sub>	e <sub>i</sub> <sup>*</sup>
Weekend1	-0.077	-69.000	-0.1626443	3.65
Public_Holidays1	-0.137	-50.430	-0.2697022	6.05
Schools_Vacation1	-0.020	-14.040	-0.0445455	1.00
Temperature	0.001	9.690	0.0003627	-5.95
GasDiff_Today-Weekbf	-0.354	-14.890	-0.0000610	1.00

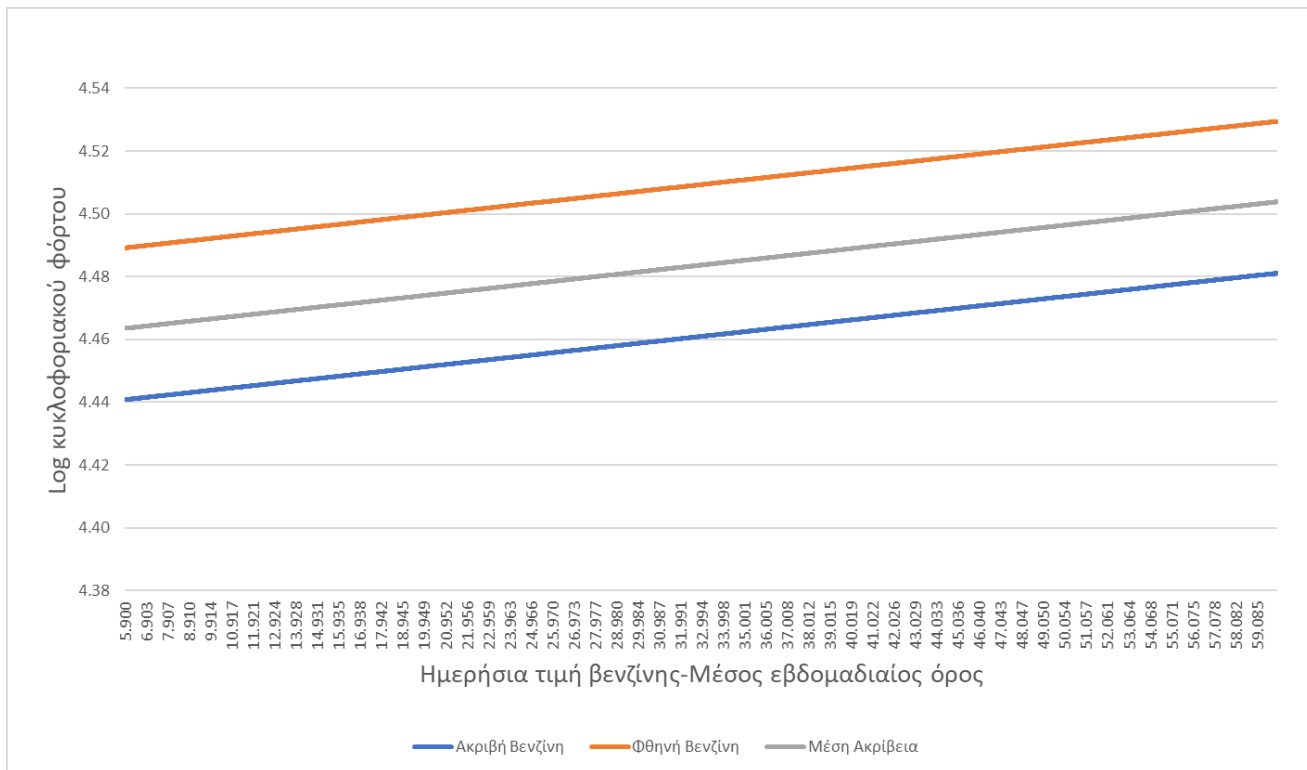
Πίνακας 4: Σχετική επιρροή ανεξάρτητων μεταβλητών για εφαρμογή στατιστικού μοντέλου σε μετρητές εκτός Δακτυλίου.

- Η **θερμοκρασία** παρουσιάζει μεγαλύτερη επιρροή από τη τιμή της βενζίνης κατά 5,95 φορές ενώ η σχετική της επιρροή δείχνει πώς για μια αύξηση της θερμοκρασίας της τάξης του 1% ο κυκλοφοριακός φόρτος αυξάνει κατά 0,03267% .
- Η ανεξάρτητη μεταβλητή **GassDiff** παρουσιάζει τη χαμηλότερη επιρροή για τις ποσοτικές συνεχείς μεταβλητές αν και το πλήθος τους είναι μικρό, ενδεικτικά για μια αύξηση της τάξης του 10% της ανεξάρτητης μεταβλητής GassDiff τότε ο κυκλοφοριακός φόρτος μειώνεται κατά 0,06%.
- Οι **αργίες** έχουν τη μεγαλύτερη σχετική επιρροή για τις κατηγορικές ανεξάρτητες μεταβλητές όπου για τη συγκεκριμένη βάση δηλαδή τις οδούς ( μετρητές ) εκτός δακτυλίου για τις αργίες παρατηρείται μια μείωση του κυκλοφοριακού φόρτου της τάξης του 27%.
- Η ανεξάρτητη μεταβλητή των **σχολικών διακοπών** παρουσιάζει την μικρότερη σχετική επιρροή για τις ποιοτικές μεταβλητές.
- Τα **σαββατοκύριακα** έχουν την αμέσως μεγαλύτερη επιρροή για της ανεξάρτητες μεταβλητές της ίδιας κατηγορίας, η οποία είναι 3.65 φορές μεγαλύτερη από αυτή των σχολικών διακοπών.

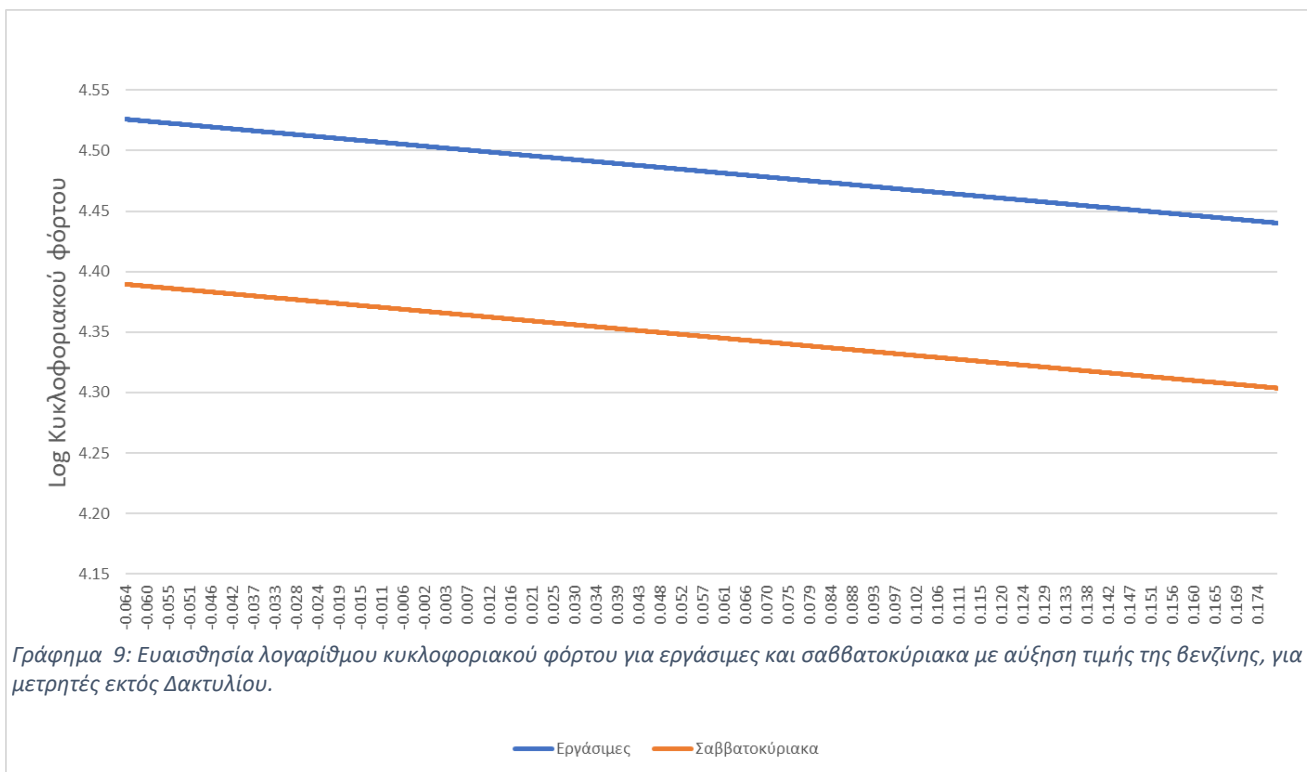
Ανάλυση ευαισθησίας των ανεξάρτητων μεταβλητών για μετρητές εκτός Δακτυλίου

Όπως και για την εφαρμογή του μαθηματικού προτύπου στη βάση δεδομένων για όλους τους μετρητές έτσι και εδώ γίνεται μια ανάλυση της ευαισθησίας της εξαρτημένης μεταβλητής, δηλαδή του λογαρίθμου του ημερήσιου κυκλοφοριακού φόρτου όταν μεταβάλλεται μια εκ των ανεξάρτητων συνεχών μεταβλητών και οι υπόλοιπες παραμένουν σε σταθερή τιμή. Τα διαγράμματα που παρουσιάζονται είναι για τους ίδιους συνδυασμούς που παρουσιάστηκαν στην εφαρμογή του μοντέλου για τη βάση δεδομένων καθολικά.

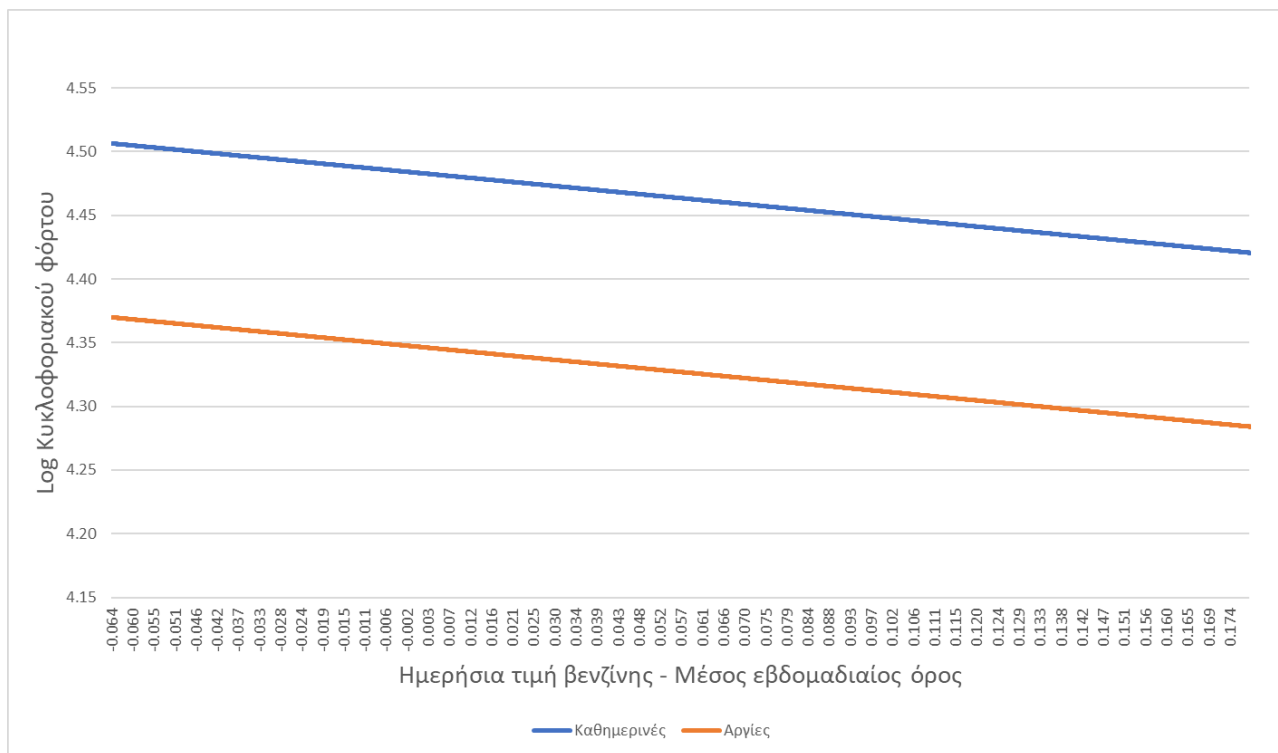




Γράφημα 8: Ευαισθησία λογαρίθμου κυκλοφοριακού φόρτου για επίπεδα βενζίνης και άνοδο της θερμοκρασίας για μετρητές εκτός Δακτυλίου.



Γράφημα 9: Ευαισθησία λογαρίθμου κυκλοφοριακού φόρτου για εργάσιμες και σαββατοκύριακα με αύξηση τιμής της βενζίνης, για μετρητές εκτός Δακτυλίου.



Γράφημα 10: Ευαισθησία λογαρίθμου κυκλοφοριακού φόρτου για εργάσιμες και αργίες με αύξηση τιμής της βενζίνης για μετρητές εκτός Δακτυλίου.

### Συνολικός σχολιασμός των διαγραμμάτων

Τα παραπάνω διαγράμματα δεν διαφέρουν σημαντικά από αυτά που προκύπτουν για την αρχική εφαρμογή του μοντέλου. Αύξηση στη διαφορά της τιμής της βενζίνης προκαλεί μείωση του κυκλοφοριακού φόρτου είτε πρόκειται για εργάσιμες είτε πρόκειται για αργίες ή σαββατοκύριακα. Επίσης στο διάγραμμα με τα τρία επίπεδα για τη τιμή της βενζίνης προκύπτει πώς ο κυκλοφοριακός φόρτος και οι μέγιστες τιμές αυτού κατά την άνοδο της θερμοκρασίας εξαρτάται από το κόστος ανεφοδιασμού. Για διαφορετικά επίπεδα ακρίβειας ο κυκλοφοριακός φόρτος αυξάνει με την ίδια κλίση αλλά καταλήγει σε διαφορετικές τιμές με υψηλότερη τιμή όταν η τιμή η ημερήσια τιμή της βενζίνης είναι χαμηλότερη από το μέσο όρος των προηγούμενων επτά ημερών και άρα η διαφορά που ορίζει την ανεξάρτητη μεταβλητή είναι αρνητική.

## Εφαρμογή μοντέλου για βάση δεδομένων με μετρητές εντός του Δακτυλίου

Το ίδιο μοντέλο εφαρμόζεται και στη βάση δεδομένων που αφορά μόνο τους μετρητές των οποίων η γεωγραφική θέση είναι εντός του δακτυλίου. Οι μετρητές αυτοί βρίσκονται δηλαδή στο κέντρο της πόλης όπως έχει αναλυθεί και στο Κεφάλαιο 4.

Log_Volume ~ Weekend + Public_Holiday + Schools_Vacation + Temperature + GassDiff_Today-Weekbf+ (1   id)				
Fixed Effects				
	Estimate	Std.Error	t-value	p-value
Intercept	4.372	0.01992	219.49	0.000
Weekend1	-0.037	0.00125	-29.32	0.000
Public_Holidays1	-0.127	0.00302	-42.05	0.000
Schools_Vacations1	-0.013	0.00157	-8.07	0.000
Termperature	0.0004	0.00009	4.97	0.000
GasDiff_Today-Weekbf	-0.306	0.02653	-11.53	0.000
Random Effects				
Groups	Variance	Std.Dev	R <sup>2</sup>	0.86
id	0.02637	0.1624		
Residual	0.00436	0.06603		

Πίνακας 5: Αποτελέσματα στατιστικού μοντέλου για εφαρμογή σε μετρητές εντός Δακτυλίου.

Μαθηματική σχέση μοντέλου για μετρητές εντός Δακτυλίου

$$\begin{aligned}
 \text{Log}_{\text{Volume}} = & 4.372 - 0.037 * (\text{Weekend}) - 0.127 * (\text{Public}_{\text{Holidays}1}) - 0.013 \\
 & * (\text{Schools}_{\text{Vacations}1}) + 0.0004 * (\text{Temperature}) - 0.306 \\
 & * (\text{GassDiff}_{\text{Today}} - \text{Weekbf})
 \end{aligned}$$

### Ποιότητα μοντέλου για μετρητές εντός Δακτυλίου

- Συντελεστής συσχέτισης  $R^2$  υψηλός και ίσος με 0.86, άρα το μοντέλο ερμηνεύει τη διακύμανση της εξαρτημένης μεταβλητής, του λογαρίθμου του κυκλοφοριακού φόρτου κατά 86%.
- Στατιστικά σημαντικές όλες οι ανεξάρτητες μεταβλητές με δείκτη  $t$  ανώτερο του 1.7.
- Λογικά πρόσχημα τα οποία συμφωνούν και με τα πρόσχημα του μοντέλου για τις προηγούμενες βάσεις δεδομένων.

### Σχολιασμός αποτελεσμάτων μοντέλου για μετρητές εντός του Δακτυλίου

- Κατά την αύξηση της ανεξάρτητης μεταβλητής **GasDiff\_Today-Weekbf** ο λογάριθμος του κυκλοφοριακού φόρτου μειώνεται κατά 0,306.
- Η ανεξάρτητη μεταβλητή **Temperature** έχει συντελεστή ίσο με 0,0004 άρα κατά την άνοδο της θερμοκρασίας οι μετακινήσεις αυξάνονται και έτσι ο λογάριθμος του κυκλοφοριακού φόρτου αυξάνει.
- Τα σαββατοκύριακα έχουν αρνητική επίδραση στον λογάριθμο του κυκλοφοριακού φόρτου με συντελεστή ίσο με -0,037 υπογραμμίζοντας έτσι πώς τις εργάσιμες ημέρες όπου η ανεξάρτητη μεταβλητή παίρνει τιμή ίση με μηδέν ο λογάριθμος του κυκλοφοριακού φόρτου είναι υψηλότερος.
- Η ανεξάρτητη μεταβλητή που εκφράζει τις **αργίες** έχει αρνητική τιμή συντελεστή και πιο συγκεκριμένα κατά τις ημέρες των αργιών ο λογάριθμος του κυκλοφοριακού φόρτου μειώνεται κατά 0,127.
- Οι σχολικές διακοπές επίσης έχουν αρνητικό συντελεστή άρα για τιμή της ανεξάρτητης μεταβλητής 1 τότε ο λογάριθμος του κυκλοφοριακού φόρτου μειώνεται κατά 0,013.

### Σχετική επιρροή των ανεξάρτητων μεταβλητών για μετρητές εντός Δακτυλίου

Ανεξάρτητες μεταβλητές	$\beta_i$	t	Σχετική επιρροή	
			$e_i$	$e_i^*$
Weekend1	-0.03655	-29.317	-0.081	2.80
Public_Holidays1	-0.12690	-42.045	-0.253	8.80
Schools_Vacation1	-0.01269	-8.068	-0.029	1.00
Temperature	0.00043	4.967	0.0002072	-3.70
GasDiff_Today-Weekbf	-0.30580	-11.527	-0.0000560	1.00

Πίνακας 6: Σχετική επιρροή ανεξάρτητων στατιστικού μοντέλου για μετρητές εντός του Δακτυλίου.

- Για τις συνεχείς ανεξάρτητες μεταβλητές η **θερμοκρασία** παρουσιάζει τη μεγαλύτερη σχετική επιρροή ίση με 0,02 άρα για κάθε μια αύξηση της θερμοκρασίας της τάξης του 10% παρατηρείται άνοδος στον λογάριθμο του ημερήσιου κυκλοφοριακού φόρτου κατά 0,2.
- Η ανεξάρτητη μεταβλητή της διαφοράς της τιμής της βενζίνης έχει αμέσως μικρότερη σχετική επιρροή στην εξαρτημένη μεταβλητή όσον αφορά τις συνεχείς ανεξάρτητες μεταβλητές, πιο συγκεκριμένα έχει μικρότερη σχετική επιρροή από τη θερμοκρασία κατά 3,7 φορές.
- Στις ποιοτικές ανεξάρτητες μεταβλητές τη μεγαλύτερη σχετική επιρροή έχει η ανεξάρτητη μεταβλητή των αργιών. Οι αργίες μειώνουν το λογάριθμο του κυκλοφοριακού φόρτου κατά 0.253.
- Η ανεξάρτητη μεταβλητή του **σαββατοκύριακου** έχει μικρότερη σχετική επιρροή από τις αργίες, η εξαρτημένη μεταβλητή του λογαρίθμου του κυκλοφοριακού φόρτου μειώνεται κατά 8,1% εντός του διημέρου.
- Την πιο μικρή σχετική επιρροή από τις ανεξάρτητες ποιοτικές μεταβλητές στην εξαρτημένη έχουν οι **σχολικές διακοπές** οι οποίες μειώνουν τον λογάριθμο κατά 0.029.

## 6. Συμπεράσματα

### 6.1 Σύνοψη αποτελεσμάτων

Αντικείμενο της διπλωματικής εργασίας αποτέλεσε η **ανάλυση της επιρροής της τιμής της βενζίνης στις κυκλοφοριακές συνθήκες της Αθήνας** και συγκεκριμένα στον κυκλοφοριακό φόρτο της πόλης.

Ύστερα από τον καθορισμό του στόχου της διπλωματικής εργασίας, ακολούθησε η **βιβλιογραφική ανασκόπηση** ώστε να μελετηθούν υφιστάμενες έρευνες και μεθοδολογίες με συναφή αντικείμενο. Στη συνέχεια, γίνεται η **συλλογή και επεξεργασία** όλων των απαραίτητων δεδομένων από διάφορες πηγές συνθέτοντας τις βάσεις με τις οποίες θα εκπονηθεί η στατιστική ανάλυση. Εν συνεχεία γίνεται η επιλογή της κατάλληλης **μεθοδολογίας**, η εφαρμογή των μαθηματικών προτύπων που αυτή περιλαμβάνει, ελέγχεται η στατιστική σημαντικότητα των μοντέλων και ερμηνεύονται τα αποτελέσματα. Η στατιστική ανάλυση πραγματοποιήθηκε με χρήση του **Γραμμικού Μικτού Μοντέλου** (Linear Mixed Effects Model). Ως εξαρτημένη μεταβλητή ορίστηκε ο **λογάριθμος του ημερήσιου κυκλοφοριακού φόρτου** (διελεύσεις οχημάτων ανά ημέρα) και ως ανεξάρτητες μεταβλητές ορίστηκαν η διαφορά της ημερήσιας τιμής της **βενζίνης** από τον μέσο όρο των επτά προηγούμενων ημερών (ευρώ ανά λίτρο), η θερμοκρασία ημέρας (βαθμούς κελσίου) και μια σειρά από μεταβλητές όπως η ημέρα της εβδομάδας, αν δηλαδή πρόκειται για εργάσιμες ή σαββατοκύριακα, οι αργίες και οι σχολικές διακοπές. Η εφαρμογή ενιαίου μοντέλου γίνεται σε τρεις διαφορετικές **περιοχές μελέτης**, η πρώτη αφορά την ευρύτερη περιοχή της Αθήνας, η δεύτερη την περιοχή εκτός Δακτυλίου Αθηνών ενώ η τελευταία αφορά την περιοχή εντός Δακτυλίου Αθηνών .

Σημειώνεται πώς η επιλογή του στατιστικού μοντέλου προκύπτει ύστερα από αρκετές δοκιμές συνδυασμών ανεξάρτητων μεταβλητών ενώ αρκετές ήταν και οι δοκιμές σχετικά με τον παράγοντα που θα καθορίσει την τυχαία επίδραση του σταθερού όρου (random intercept). Η ομαδοποίηση έγινε δηλαδή με βέλτιστο τρόπο προκειμένου οι ομάδες των παρατηρήσεων να περιέχουν μετρητές με κοινά χαρακτηριστικά κυκλοφοριακών συνθηκών μειώνοντας έτσι τις διαφορές κυκλοφοριακών συνθηκών που μετράει ο κάθε μετρητής.

Συνολική βάση δεδομένων					Βάση δεδομένων μετρητές εκτός δακτυλίου						Βάση δεδομένων μετρητές εντός δακτυλίου							
	Fixed Effects				Σχετική Επιρροή		Fixed Effects				Σχετική Επιρροή		Fixed Effects				Σχετική επιρροή	
	Estimate	Std.Error	t-value	p-value	$e_i$	$e_i^*$	Estimate	Std.Error	t-value	p-value	$e_i$	$e_i^*$	Estimate	Std.Error	t-value	p-value	$e_i$	$e_i^*$
Intercept	4.442	0.0169	262.62	0.000			4.496	0.0247	181.72	0.000			4.372	0.0199	219.49	0.000		
Weekend1	-0.059	0.0008	-71.00	0.000	-0.1276	3.376	-0.077	0.0011	-69.00	0.000	-0.1626	3.651	-0.037	0.0012	-29.32	0.000	-0.0807	2.803
Public_Holidays1	-0.132	0.0020	-65.31	0.000	-0.2623	6.940	-0.137	0.0027	-50.43	0.000	-0.2697	6.055	-0.127	0.0030	-42.05	0.000	-0.2534	8.799
Schools_Vacations1	-0.017	0.0011	-15.88	0.000	-0.0378	1.000	-0.020	0.0014	-14.04	0.000	-0.0445	1.000	-0.013	0.0016	-8.07	0.000	-0.0288	1.000
Termperature	0.001	0.0001	10.59	0.000	0.0003	-4.852	0.001	0.0001	9.69	0.000	0.0004	-5.947	0.0004	0.0001	4.97	0.000	0.0002	-3.697
GasDiff_Today-Weekbf	-0.333	0.0178	-18.73	0.000	-0.0001	1.000	-0.354	0.0238	-14.89	0.000	-0.0001	1.000	-0.306	0.0265	-11.53	0.000	-0.0001	1.000
Random Effects					Random Effects						Random Effects							
Groups	Variance	Std.Dev	R <sup>2</sup>		0.91		Variance	Std.Dev	R <sup>2</sup>		0.92		Variance	Std.Dev	R <sup>2</sup>		0.86	
id	0.04499	0.21211					0.054251	0.23292					0.02637	0.1624				
Residual	0.004613	0.06792					0.004654	0.06822					0.00436	0.06603				

Πίνακας 7: Αποτελέσματα στατιστικών μοντέλων

## 6.2 Συμπεράσματα

Κατά τα διάφορα στάδια εκπόνησης της παρούσας Διπλωματικής Εργασίας προέκυψε μία σειρά συμπερασμάτων που συνδέεται άμεσα με τους αρχικούς στόχους και τα ερωτήματά της. Τα **γενικά συμπεράσματα** συνοψίζονται ως εξής:

1. Παρατηρείται **αρνητική επιρροή της αύξησης στην τιμή της βενζίνης στον ημερήσιο κυκλοφοριακό φόρτο** στην ευρύτερη περιοχή της Αθήνας καθώς και εντός, εκτός δακτυλίου Αθηνών. Συγκεκριμένα, σε κάθε περιοχή μελέτης η διαφορά της μέσης ημερήσιας τιμής της βενζίνης από την μέση τιμή της προηγούμενης εβδομάδας. Δηλαδή, η ακρίβεια στο κόστος ανεφοδιασμού μειώνει τον ημερήσιο κυκλοφοριακό φόρτο, άρα το κόστος μετακίνησης όσον αφορά το κόστος καυσίμου όσο αυξάνει ο λογάριθμος του κυκλοφοριακού φόρτου μειώνει. Το παραπάνω ισχύει και αντίστροφα, όσο δηλαδή η τιμή της βενζίνης μειώνεται άρα η τιμή της βενζίνης είναι μειωμένη συγκριτικά με το μέσο όρο των επτά προηγούμενων ημερών τότε η ημερήσια κυκλοφορία αυξάνει.
2. **Εκτός του Δακτυλίου Αθηνών**, η τιμή της βενζίνης έχει ελαφρώς μεγαλύτερη επιρροή στον κυκλοφοριακό φόρτο σε σχέση με εντός. Η διαφορά αυτή ενδεχομένως να οφείλεται στις κυκλοφοριακές συνθήκες που επικρατούν στις οδούς εντός του δακτυλίου. Δεδομένου ότι η κυκλοφορία σε οδούς στο κέντρο της πόλης (εντός δακτυλίου) χαρακτηρίζεται συχνά από κυκλοφοριακή συμφόρηση, τότε ακόμα και αν η ακρίβεια της βενζίνης μειώνει τις μετακινήσεις η μεταβολή δεν μετακυλά στον κυκλοφοριακό φόρτο καθώς στις οδούς αυτούς στην πράξη ο κυκλοφοριακός φόρτος ορισμένες φορές αυξάνεται επειδή από την περιοχή συμφορημένης ροής η οδός περνά στην περιοχή μη συμφορημένης ροής.
3. Ύστερα από αρκετές δοκιμές σχετικά με την ανεξάρτητη μεταβλητή που θα εκφράσει την επιρροή του κόστους της βενζίνης στον κυκλοφοριακό φόρτο, είναι φανερό πως οι οδηγοί αντιλαμβάνονται και μειώνουν τις μετακινήσεις με ιδιωτικά οχήματα **ύστερα από**

**μία χρονική περίοδο και όχι άμεσα.** Αυτό συμβαίνει καθώς ο ανεφοδιασμός των οχημάτων τους δεν προκύπτει καθημερινά ενώ παράλληλα από τη στιγμή που παρατηρούν την αύξηση, χρειάζονται και κάποιο χρόνο για να τροποποιήσουν το πρόγραμμα μετακινήσεων τους.

4. Η **άνοδος της θερμοκρασίας αυξάνει τον ημερήσιο κυκλοφοριακό φόρτο** σε όλα τα στατιστικά μοντέλα. Συγκεκριμένα, όταν ο καιρός είναι αίθριος οι άνθρωποι τείνουν να κάνουν παραπάνω μετακινήσεις μιας και το κλίμα τους επιτρέπει πέρα από τις μετακινήσεις που αφορούν βιοποριστικές ανάγκες όπως η εργασία, να πραγματοποιούν και μετακινήσεις αναψυχής. Επίσης, εκτός Δακτυλίου η θερμοκρασία φαίνεται να έχει μεγαλύτερη επιρροή συγκριτικά με εντός Δακτυλίου, αυτό συμβαίνει ενδεχομένως εφόσον οι μετακινήσεις εκτός Δακτυλίου με προορισμό από και προς τα προάστια είναι παραπάνω για τις ημέρες με θερμό κλίμα αλλά και επειδή οι μετακινήσεις εντός Δακτυλίου είναι συνήθως με σκοπό την εξυπηρέτηση αναγκών και επαγγελματικών υποχρεώσεων άρα η θερμοκρασία δεν επηρεάζει τόσο την πραγματοποίησή τους.
5. Οι **αργίες** για όλα τα στατιστικά μοντέλα έχουν τη μεγαλύτερη σχετική επιρροή από τις κατηγορικές ανεξάρτητες μεταβλητές. Κάτι τέτοιο παρατηρείται ενδεχομένως εφόσον οι μετακινήσεις από και προς την εργασία δεν υφίστανται άρα ο κυκλοφοριακός φόρτος μειώνεται αρκετά.
6. Τα **σαββατοκύριακα** επηρεάζουν αρνητικά τον κυκλοφοριακό φόρτο για όλα τα στατιστικά μοντέλα. Σημαντική είναι η διαφορά που προκύπτει ανάμεσα στις μετακινήσεις εκτός και εντός δακτυλίου. Η διαφορά αυτή προκύπτει από τον εμπορικό χαρακτήρα που έχει το κέντρο της πόλης, πολλές φορές οι οδηγοί πραγματοποιούν μετακινήσεις για αγορές και αναψυχή στο κέντρο.
7. Οι **σχολικές διακοπές** για κάθε μοντέλο αποτελούν παράγοντα που μειώνει τον κυκλοφοριακό φόρτο σε κάθε εντός και εκτός Δακτυλίου αλλά και στην ευρύτερη περιοχή της Αθήνας. Οι σχολικές μετακινήσεις αποτελούν μέρος της κυκλοφορίας ασχέτως της τοποθεσίας με αποτέλεσμα να μην υπάρχει μεγάλη διαφορά στην επιρροή τους στον κυκλοφοριακό φόρτο για εκτός και εντός Δακτυλίου.
8. Τα **γραμμικά μικτά μοντέλα** είναι κατάλληλα για την ανάλυση της επιρροής της τιμής της βενζίνης στον ημερήσιο κυκλοφοριακό φόρτο ενώ ιδιαίτερα σημαντική είναι η δυνατότητα προσθήκης τυχαίων επιδράσεων ανά ομάδα παρατηρήσεων μειώνοντας έτσι τη συνολική διακύμανση.
9. Ο ημερήσιος **κυκλοφοριακός φόρτος** είναι πιο ευαίσθητος σε εξωγενείς παράγοντες εκτός του Δακτυλίου των Αθηνών σε σύγκριση με εντός. Αυτό συμβαίνει καθώς όπως έχει αναλυθεί το κέντρο της πόλης είναι πάντοτε πόλος έλξης μετακινήσεων για τους κατοίκους όλου του Λεκανοπεδίου με αποτέλεσμα να εξυπηρετεί παραπάνω ανάγκες για μεγαλύτερο πληθυσμό. Σημαντική λεπτομέρεια στο παραπάνω φαινόμενο αποτελεί πως οι περιοχές εντός Δακτυλίου διαθέτουν καλύτερη υποδομή Μέσων Μαζικής Μετακίνησης από τις περιοχές εκτός αλλά παρόλα αυτά σε περιόδους ακρίβειας της βενζίνης οι μετακινήσεις μειώνονται λιγότερο από ότι εκτός Δακτυλίου, αυτό



υποδεικνύει ενδεχομένως ότι οι οδηγοί προτιμούν να μειώσουν άσκοπες μετακινήσεις παρά να αλλάξουν μέσο μετακίνησης για την καθημερινότητα τους.

### 6.3 Προτάσεις

Σε συνέχεια των συμπερασμάτων και κατόπιν της εκπόνηση ολόκληρης της έρευνας είναι πλέον πλήρως αντιληπτή η επιρροή του κόστους της βενζίνης στον κυκλοφοριακό φόρτο. Η σφαιρικότητα με την οποία προσεγγίστηκε το ερευνητικό αντικείμενο επιτρέπει τη διατύπωση ορισμένων προτάσεων οι οποίες συνοψίζονται παρακάτω.

Ίσως η πιο καθολική λύση στην αντιμετώπιση της ακρίβειας των καυσίμων είναι η αναβάθμιση των υποδομών και τον δρομολογίων της συγκοινωνίας. Αν βελτιωθεί αρκετά το επίπεδο εξυπηρέτησης της δημόσιας συγκοινωνίας τότε οι οδηγοί σε περιόδους ακρίβειας θα είναι πιο ελαστικοί και θα προτιμήσουν να εξοικονομήσουν χρόνο και χρήμα χρησιμοποιώντας τις συγκοινωνίες.

Ο στόλος της Αθήνας αποτελείται από αυτοκίνητα διαφορετικής χρονολογίας και τεχνολογίες, συνεπώς η πολιτεία ύστερα από την εφαρμογή οικονομικών κριτηρίων θα μπορούσε να θεσπίσει ισχυρές επιδοτήσεις προκειμένου να ανανεωθεί ο στόλος και οι οδηγοί έχοντας πλέον οχήματα νέας τεχνολογίας με χαμηλή κατανάλωση καυσίμου θα μπορέσουν να ισοσκελίσουν τις ανοδικές τιμές.

Ένα μέτρο το οποίο εφαρμόστηκε με επιτυχία είναι και η επιδότηση για ανεφοδιασμό. Συνήθως οι περίοδοι ακρίβειας δεν διαρκούν μεγάλα χρονικά διαστήματα οπότε είναι πιο οικονομικό και εύκολο για την πολιτεία να χορηγήσει ατομικά επιδόματα ανεφοδιασμού παρά να επενδύσει σε αναβάθμιση υποδομών. Το δεύτερο βέβαια έχει επιπλέον αναρίθμητα πλεονεκτήματα ενώ παράλληλα τα έργα προσφέρουν θέσεις εργασίας δίνοντας έτσι μια κινητικότητα στην οικονομία η οποία συνήθως χωλαίνει τις περιόδους ακρίβειας.

### 6.4 Προτάσεις για περαιτέρω έρευνα

Η επέκταση της συγκεκριμένης ανάλυσης με παραπάνω μεταβλητές θα ήταν αρκετά σημαντική και ενδέχεται να προκύψουν αρκετά ενδιαφέροντα αποτελέσματα. Ενδεικτικά μια σημαντική προσθήκη θα ήταν η μέτρηση της κατανάλωσης και της μέσης ταχύτητας των οδηγών μέσω εφαρμογών τηλεματικής μελετώντας έτσι την αλλαγή στην οδική συμπεριφορά των οδηγών ανάλογα με την τιμή της βενζίνης. Ένα ακόμα βήμα περαιτέρω έρευνας θα ήταν η διεξαγωγή της ίδιας μελέτης σε αυτοκινητοδρόμους ώστε να διερευνηθεί η επιρροή της τιμής της βενζίνης σε μετακινήσεις μεγάλων αποστάσεων. Η προσθήκη ενός οικονομικού δείκτη ως ανεξάρτητη μεταβλητή όπως ο πληθωρισμός. Ακόμη θα μπορούσε να δοκιμαστεί μια διαφορετική στατιστική μέθοδος όπως αυτή της μεθόδου ανάλυσης χρονοσειρών.



## Βιβλιογραφία

- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using **lme4**. *J Stat Softw* 67.
- Bento, A.M., Hughes, J.E., Kaffine, D., 2013. Carpooling and driver responses to fuel price changes: Evidence from traffic flows in Los Angeles. *J Urban Econ* 77, 41–56.
- Delsaut, M., 2014. The Effect of Fuel Price on Demands for Road and Rail Travel: An Application to the French Case. In: *Transportation Research Procedia*. Elsevier, pp. 177–187.
- Hadley Wickham, 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Hadley Wickham and Jennifer Bryan, 2022. *readxl: Read Excel Files*.
- Jacob A. Long, 2022. *jtools: Analysis and Presentation of Social Scientific Data*.
- Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. *lmerTest Package: Tests in Linear Mixed Effects Models*. *J Stat Softw* 82.
- R Core Team, 2022. *R: A Language and Environment for Statistical Computing*.
- Wickham H, Francois R, Henry L, Müller K, n.d. Wickham H, François R, Henry L, Müller K (2022). *dplyr: A Grammar of Data Manipulation*. <https://dplyr.tidyverse.org>, <https://github.com/tidyverse/dplyr>.
- Zhang, T., Burke, P.J., 2020. The effect of fuel prices on traffic flows: Evidence from New South Wales. *Transp Res Part A Policy Pract* 141, 502–522.
- A. Σταθόπουλος, Μ. Καρλαύτης, 2016. *Σχεδιασμός Μεταφορικών Συστημάτων*, 2nd ed.