



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Συσχέτιση επιδόσεων οδικής ασφάλειας με δείκτες υγείας και οικονομίας στην Ευρωπαϊκή Ένωση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μυρτώ Εμ. Δαμιανού

Επιβλέποντες: Διονύσιος-Δημήτριος Κουτσούρης
Καθηγητής Ε.Μ.Π.

Γεώργιος Γιαννής
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2018



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Συσχέτιση επιδόσεων οδικής ασφάλειας με δείκτες υγείας και οικονομίας στην Ευρωπαϊκή Ένωση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μυρτώ Εμ. Δαμιανού

Επιβλέποντες: Διονύσιος-Δημήτριος Κουτσούρης
Καθηγητής Ε.Μ.Π.

Γεώργιος Γιαννής
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή της ... Ιουλίου 2018

.....
Κουτσούρης Δ.
Καθηγητής Ε.Μ.Π.

.....
Γιαννής Γ.
Καθηγητής Ε.Μ.Π.

.....
Ματσόπουλος Γ.
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2018

.....
Μυρτώ Εμ. Δαμιανού
Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών

Copyright © Μυρτώ Εμ. Δαμιανού, 2018.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Στην παρούσα διπλωματική εργασία μελετάται η συσχέτιση των οδικών ατυχημάτων και της οδικής συμπεριφοράς, εν γένει, με τον κλάδο της υγείας. Στόχος της είναι η συσχέτιση των επιδόσεων της οδικής ασφάλειας με ιατρικούς, οικονομικούς και κοινωνικούς δείκτες στην Ευρωπαϊκή Ένωση. Για το λόγο αυτό, μελετώνται οι 27 χώρες-μέλη της Ευρωπαϊκής Ένωσης για τα έτη 2008-2014. Μετά τον καθορισμό της ύπαρξης συσχέτισης μεταξύ, του υπό εξέταση φαινομένου και των δεικτών από τους τομείς της υγείας, της οικονομίας και της κοινωνίας, καθορίζονται κάποια στατιστικά μοντέλα, σύμφωνα με την εφαρμογή της θεωρίας της γραμμικής παλινδρόμησης και του γραμμικού μεικτού μοντέλου. Τέλος, παρατίθενται τα σχετικά συμπεράσματα, καθώς και οι περιορισμοί οι οποίοι προέκυψαν κατά την ερευνητική διαδικασία από όπου προκύπτουν ποικίλες προτάσεις για περαιτέρω έρευνα στον εν λόγω κλάδο.

Λέξεις κλειδιά: θνησιμότητα, οδικά ατυχήματα, υγεία, οικονομία, παλινδρόμηση, μικτό μοντέλο, ευρωπαϊκή ένωση

Abstract

The purpose of this diploma thesis is to identify potential relationships between road accidents and road behavior, as far as the health sector is concerned. Its aim is to relate the performance of road safety to medical, economic and social indicators related to countries in the European Union. Thus, the 27 European Union Member States are studied for an indicative time frame, from 2008 to 2014. After determining the correlation between road fatalities and the aforementioned indicators, the statistical models were defined, according to the application of the theory of linear regression and the linear mixed model. Finally, the conclusions drawn are presented, as well as the limitations of the current research process, leading to various proposals for further research in this field.

Key words: mortality, road accidents, health, economy, regression, mixed model, European union

Αναλυτική περίληψη

Αντικείμενο της παρούσας Διπλωματικής Εργασίας αποτέλεσε η **διερεύνηση της επιρροής κοινωνικών, οικονομικών και ιατρικών παραγόντων στα οδικά ατυχήματα** και συγκεκριμένα στην **ετήσια μεταβολή του αριθμού των νεκρών στα οδικά ατυχήματα** ως προς τον πληθυσμό.

Μετά τον καθορισμό του επιδιωκόμενου στόχου, ξεκίνησε η βιβλιογραφική αναζήτηση ερευνών συναφών με το αντικείμενο της παρούσας Διπλωματικής Εργασίας σε χώρες της Ευρωπαϊκής Ένωσης, αλλά και παγκοσμίως, καθώς και των διαθέσιμων στοιχείων που ήταν απαραίτητα για τη συγκεκριμένη διερεύνηση.

Στη συνέχεια, ακολούθησε η συλλογή των απαραίτητων για την εργασία στοιχείων από βάσεις δεδομένων διάφορων οργανισμών, όπως EUROSTAT, CARE και UNDP, προκειμένου να αναπτυχθεί η σχετική βάση δεδομένων. Τα στοιχεία που συλλέχθηκαν περιελάμβαναν τον αριθμό των νεκρών στα οδικά ατυχήματα, τον πληθυσμό, το κατά κεφαλήν Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.), τον ετήσιο αριθμό νεκρών σε οδικά ατυχήματα, το πλήθος των οδικών ατυχημάτων εντός αστικού δικτύου, την προσδοκώμενη διάρκεια ζωής των ανθρώπων, τα παραγωγικά χρόνια ζωής των ανθρώπων, το κατά κεφαλήν Α.Ε.Π., το δείκτη ανθρώπινης ανάπτυξης, τις ετήσιες δαπάνες για τον τομέα της υγείας, τις δημόσιες δαπάνες για τον τομέα της υγείας και τον αριθμό των κλινών στα νοσοκομεία των 27 κρατών-μελών της Ευρωπαϊκής Ένωσης για τη χρονική περίοδο 2008-2014.

Τη συλλογή των δεδομένων ακολούθησε η επεξεργασία τους προκειμένου να επιλεγεί η κατάλληλη μεθοδολογία, αλλά και να κωδικοποιηθούν, ώστε να εισαχθούν στο ειδικό στατιστικό λογισμικό (Rstudio©). Η στατιστική ανάλυση πραγματοποιήθηκε μέσω Γραμμικής Παλινδρόμησης και Γραμμικού Μικτού Μοντέλου. Αρχικά, η ανάλυση πραγματοποιήθηκε για το σύνολο των κρατών, σύμφωνα με τη **λογαριθμισμένη γραμμική παλινδρόμηση, με εξαρτημένη μεταβλητή την ετήσια μεταβολή του φυσικού λογαρίθμου του αριθμού των νεκρών στα οδικά ατυχήματα** και ως **ανεξάρτητες μεταβλητές:**

- το κατά κεφαλήν Α.Ε.Π.
- το πλήθος των οδικών ατυχημάτων εντός αστικού δικτύου συναρτήσει του συνόλου των θανατηφόρων οδικών ατυχημάτων
- την προσδοκώμενη διάρκεια ζωής των ανθρώπων
- τα παραγωγικά χρόνια ζωής των ανθρώπων
- το δείκτη ανθρώπινης ανάπτυξης
- τις ετήσιες δαπάνες για τον τομέα της υγείας
- τις δημόσιες δαπάνες για τον τομέα της υγείας
- τον αριθμό των κλινών στα νοσοκομεία ως προς τον πληθυσμό

Στη συνέχεια, δημιουργήθηκαν στατιστικά μοντέλα μέσω της εφαρμογής του **Γραμμικού Μικτού Μοντέλου**, τα οποία περιελάμβαναν ως **εξαρτημένη μεταβλητή την ετήσια μεταβολή του φυσικού λογαρίθμου του αριθμού των νεκρών στα οδικά ατυχήματα** και ως **ανεξάρτητες μεταβλητές:**

- το κατά κεφαλήν Α.Ε.Π.
- το πλήθος των οδικών ατυχημάτων εντός αστικού δικτύου συναρτήσει του συνόλου των θανατηφόρων οδικών ατυχημάτων
- την προσδοκώμενη διάρκεια ζωής των ανθρώπων
- τα παραγωγικά χρόνια ζωής των ανθρώπων
- το δείκτη ανθρώπινης ανάπτυξης
- τις ετήσιες δαπάνες για τον τομέα της υγείας
- τις δημόσιες δαπάνες για τον τομέα της υγείας
- τον αριθμό των κλινών στα νοσοκομεία ως προς τον πληθυσμό

Από αυτά τα μοντέλα προέκυψαν τα δύο τελικά μοντέλα, τα οποία έχουν ως εξαρτημένη μεταβλητή την ετήσια μεταβολή του φυσικού λογαρίθμου του αριθμού των νεκρών στα οδικά ατυχήματα και ως ανεξάρτητες μεταβλητές τα παραγωγικά χρόνια ζωής των ανθρώπων, το δείκτη ανθρώπινης ανάπτυξης, τις ετήσιες δαπάνες για τον τομέα της υγείας ανά κάτοικο και τον αριθμό των κλινών στα νοσοκομεία ως προς τον πληθυσμό.

Πρέπει να σημειωθεί, ότι η επιλογή των μοντέλων, τόσο για το σύνολο των κρατών όσο και για την κάθε ομάδα ξεχωριστά, έγινε έπειτα από αρκετές δοκιμές συνδυασμών ανεξάρτητων και εξαρτημένων μεταβλητών. Στο προηγούμενο κεφάλαιο έγινε η ανάπτυξη και επεξήγηση των τελικών μοντέλων που προέκυψαν από τη στατιστική επεξεργασία, αφού πρώτα ελέγχθηκε κατά πόσο αυτά πληρούν τους απαραίτητους στατιστικούς ελέγχους.

Από τα διάφορα στάδια εκπόνησης της Διπλωματικής Εργασίας προκύπτει ότι, σύμφωνα με τα τέσσερα μοντέλα τα οποία αναπτύχθηκαν, **η θνησιμότητα στα οδικά ατυχήματα έχει άμεση σχέση με τον κλάδο της υγείας και μάλιστα σε μεγάλο βαθμό.**

Όπως φαίνεται από τα αποτελέσματα της λογαριθμισμένης γραμμικής παλινδρόμησης, **ιδιαίτερη έμφαση** πρέπει να δοθεί και σε **παράγοντες κοινωνικού και οικονομικού χαρακτήρα**, όπως ο δείκτης ανθρώπινης ανάπτυξης, τα παραγωγικά χρόνια ζωής και η ανεργία. Παράλληλα, εφόσον το φαινόμενο το οποίο μελετάται εμπίπτει στον τομέα των συγκοινωνιών, είναι χρήσιμο στο μοντέλο να εντάσσονται και παράγοντες από αυτό τον τομέα, όπως η αναλογία ατυχημάτων εντός αστικού ιστού ως προς το σύνολο των θανατηφόρων οδικών ατυχημάτων.

Σύμφωνα με τα τελικά μοντέλα του Γραμμικού Μικτού Μοντέλου, η **επιρροή των ιατρικών παραγόντων διαφέρει ανάλογα** με την **ομάδα κρατών** που μελετάται. Όταν εμφανίζεται το πλήθος των νοσοκομειακών κλινών ως προς τον πληθυσμό ως ο μόνος ιατρικός παράγοντας, η πρώτη και η δεύτερη ομάδα χωρών εμφανίζουν μεγάλη ομοιομορφία στις τιμές των συντελεστών τους, ενώ η τρίτη ομάδα φαίνεται να επηρεάζεται περισσότερο από το συγκεκριμένο παράγοντα και ο αντίστοιχος συντελεστής της να είναι αισθητά μικρότερος από τις άλλες δύο (το οποίο είναι και το επιθυμητό). Αντίθετα, όταν προστίθεται ένας ακόμα παράγοντας από τον κλάδο της υγείας στο στατιστικό μοντέλο, η συνολική εικόνα του μοντέλου είναι σχετικά καλύτερη.

Επίσης, παρατηρείται, ότι, **κάθε ομάδα χωρών επηρεάζεται σε διαφορετικό βαθμό** από τους **ίδιους παράγοντες**, γεγονός αναμενόμενο και πλήρως συμβατό με την πρόταση διάσπασης και μελέτης του συνόλου των χωρών σε μικρότερες ομάδες, ώστε να είναι πιο ρεαλιστικά τα

αποτελέσματα και με όσο το δυνατόν μικρότερο σφάλμα. Αυτό επιβεβαιώνεται από την καλύτερη (μικρότερη) τιμή του κριτηρίου AIC μεταξύ των μοντέλων με ομάδες κρατών και των μοντέλων του συνόλου των χωρών άνευ ομάδων.

Τέλος, η μέθοδος του **Γραμμικού Μικτού Μοντέλου** είναι **κατάλληλη** για την **ανάλυση** παραμέτρων **οδικών ατυχημάτων** σε **σχέση με ιατρικούς, κοινωνικούς και οικονομικούς παράγοντες**. Τα **τελικά μαθηματικά μοντέλα**, τα οποία αναπτύχθηκαν με τη χρήση της μεθόδου αυτής, θεωρούνται γενικά **αξιόπιστα**, αφού είχαν **καλή προσαρμογή στα δεδομένα** και τα αποτελέσματα τους κρίθηκαν από το σύστημα ικανοποιητικά.

Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω τον κ. Δημήτριο Κουτσούρη, καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου (Ε.Μ.Π.) της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, για την ανάθεση της παρούσας διπλωματικής εργασίας.

Θα ήθελα επίσης να ευχαριστήσω τον κ. Γεώργιο Γιαννή, καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου (Ε.Μ.Π.) της σχολής Πολιτικών Μηχανικών, για την συμβολή του στη διαμόρφωση του θέματος της εργασίας καθώς και στην διαδικασία συλλογής των υπό μελέτη δεδομένων.

Στη συνέχεια, θα επιθυμούσα να ευχαριστήσω ιδιαίτερω τον υποψήφιο διδάκτορα του Εργαστηρίου Βιοϊατρικής Τεχνολογίας, Παναγιώτη Κατρακάζα, για την πολύτιμη βοήθειά του, την καθοδήγηση και τις συμβουλές του, καθώς και για την υπομονή του καθ' όλη τη διαδικασία συλλογής των στοιχείων, επεξεργασίας τους και εν τέλει συγγραφής της εν λόγω διπλωματικής εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου για την πολύτιμη στήριξη και βοήθεια που μου παρείχαν σε όλα τα χρόνια των σπουδών μου.

Πίνακας Περιεχομένων

Κατάλογος Πινάκων	13
Κατάλογος Διαγραμμάτων	14
1. Εισαγωγή	15
1.1 Γενική Ανασκόπηση.....	15
1.2 Στόχος της Διπλωματικής Εργασίας.....	20
1.3 Μεθοδολογία.....	21
1.4 Δομή Διπλωματικής Εργασίας.....	24
2. Βιβλιογραφική Ανασκόπηση.....	25
2.1 Εισαγωγή	25
2.2 Συναφείς Έρευνες και Μεθοδολογίες.....	25
2.3 Σύνοψη-Κριτική Αξιολόγηση	31
3. Θεωρητικό Υπόβαθρο	33
3.1 Εισαγωγή	33
3.2 Βασικές Έννοιες Στατιστικής.....	33
3.3 Συσχέτιση Μεταβλητών – Συντελεστής Συσχέτισης.....	34
3.4 Κανονική Κατανομή.....	35
3.5 Μαθηματικά Μοντέλα.....	35
3.5.1 Γραμμική Παλινδρόμηση.....	35
3.5.2 Λογαριθμισμένη Γραμμική Παλινδρόμηση.....	37
3.5.3 Γενικό Γραμμικό Μοντέλο	37
3.6 Αυτοσυσχέτιση	38
3.6.1 Βασικές Αρχές Παλινδρομήσεων.....	40
3.6.2 Κριτήρια Αποδοχής Σχέσεων.....	41
4. Συλλογή και Επεξεργασία Στοιχείων	44
4.1 Εισαγωγή	44
4.2 Συλλογή Στοιχείων.....	44
4.2.1 Βάσεις Δεδομένων	44
4.2.2 Προβλήματα που παρουσιάστηκαν κατά τη συλλογή των στοιχείων και επίλυσή τους....	45
4.3 Επεξεργασία Στοιχείων	46
5. Εφαρμογή Μεθοδολογίας και Αποτελέσματα	57

5.1 Εισαγωγή	57
5.2 Βάση Δεδομένων	57
5.3 Γραμμικό Μοντέλο.....	58
5.4 Πίνακας Αυτοσυσχέτισης	59
5.5 Λογαριθμισμένη Γραμμική Παλινδρόμηση.....	62
5.6 Γραμμικό Μικτό Μοντέλο.....	72
5.6.1 Ανάπτυξη Μαθηματικού Μοντέλου	72
5.6.2 Προετοιμασία Εφαρμογής Γραμμικού Μικτού Μοντέλου.....	73
5.6.3 Αποτελέσματα Γραμμικού Μικτού Μοντέλου	74
6. Συμπεράσματα.....	97
6.1 Σύνοψη Αποτελεσμάτων.....	97
6.2 Συμπεράσματα.....	98
6.3 Περιορισμοί Παρούσας Έρευνας.....	100
6.4 Προτάσεις	101
6.5 Προτάσεις για Περαιτέρω Έρευνα	102
Βιβλιογραφία	103

Κατάλογος Πινάκων

Πίνακας 1.1	Αριθμός νεκρών ανά εκατομμύριο πληθυσμού στην Ε.Ε την περίοδο 2007-2016...	16
Πίνακας 1.2	Τροχαία ατυχήματα στην Ε.Ε. έτη 2008-2016.....	18
Πίνακας 4.1	Ανθρώπινες απώλειες για τα έτη 2008-2014 στις 28 χώρες της Ε.Ε	46
Πίνακας 4.2	Βάση δεδομένων στο excel No.1	48
Πίνακας 4.3	Pivot table Πληθυσμού	49
Πίνακας 4.4	Pivot table Δαπανών υγείας ανά κάτοικο	50
Πίνακας 4.5	Pivot table Απωλειών λόγω ατυχημάτων προς πληθυσμό	51
Πίνακας 4.6	Pivot table Πλήθους νοσοκομειακών κλινών προς πληθυσμό	52
Πίνακας 4.7	Pivot table Ατυχημάτων εντός αστικού ιστού προς θανάσιμα ατυχήματα	53
Πίνακας 4.8	Βάση Δεδομένων excel No.2	54
Πίνακας 5.1	Πίνακας αυτοσυσχέτισης	60
Πίνακας 5.2	74

Κατάλογος Διαγραμμάτων

Διάγραμμα 1.1 Αριθμός νεκρών ανά εκατομμύριο πληθυσμού για το έτος 2016 για τις 28 χώρες της Ε.Ε.	17
Διάγραμμα 1.2 Τροχαία ατυχήματα έτη 2008-2016 στην Ε.Ε.	19
Διάγραμμα 1.3 Σχηματική απεικόνιση σταδίων Διπλωματικής Εργασίας	23
Διάγραμμα 4.1 Θνησιμότητα στα οδικά ατυχήματα και πλήθος κλινών στα νοσοκομεία ως προς τις προς μελέτη χώρες.....	55
Διάγραμμα 4.2 Θνησιμότητα στα οδικά ατυχήματα και πλήθος ατυχημάτων εντός αστικού ιστού ως προς τις προς μελέτη χώρες.....	55
Διάγραμμα 4.3 Θνησιμότητα στα οδικά ατυχήματα και ποσοστό ανεργίας ως προς τις προς μελέτη χώρες	56
Διάγραμμα 5.1 Διάγραμμα αυτοσυσχέτισης	61
Διάγραμμα 5.2	68
Διάγραμμα 5.3	69
Διάγραμμα 5.4	70
Διάγραμμα 5.5	71
Διάγραμμα 5.6 Ανάλυση ευαισθησίας ως προς τις νοσοκομειακές κλίνες για το σύνολο των χωρών αναφορικά με την ανεργία	92
Διάγραμμα 5.7 Ανάλυση ευαισθησίας ως προς τις δαπάνες για τον τομέα της υγείας ως ποσοστό του ΑΕΠ για το σύνολο των χωρών αναφορικά με το δείκτη ανθρώπινης ανάπτυξης	93
Διάγραμμα 5.8 Ανάλυση ευαισθησίας ως προς τα cluster για δαπάνες υγείας ως ποσοστό Α.Ε.Π.....	94
Διάγραμμα 5.9 Ανάλυση ευαισθησίας ως προς τα cluster για δείκτη ανθρώπινης ανάπτυξης	95
Διάγραμμα 5.10 Ανάλυση ευαισθησίας ως προς τα cluster για πλήθος κλινών ανά πληθυσμό ..	95

1. Εισαγωγή

1.1 Γενική Ανασκόπηση

Οι νεκροί, οι τραυματίες και οι υλικές ζημιές, εξαιτίας των οδικών ατυχημάτων, έχουν τεράστιο κόστος τόσο οικονομικό όσο και κοινωνικό, με αποτέλεσμα ο περιορισμός τους να αποτελεί προτεραιότητα για κάθε χώρα. Κάθε χρόνο περίπου 1.24 εκατομμύρια άνθρωποι χάνουν τη ζωή τους σε οδικά ατυχήματα παγκοσμίως, ενώ μέχρι το 2030 ο αριθμός αυτός αναμένεται να υπερδιπλασιαστεί φτάνοντας τα 3.6 εκατομμύρια ανθρώπινες απώλειες το χρόνο εάν δεν ληφθούν κατάλληλα μέτρα αντιμετώπισης. Το πρόβλημα αυτό παρατηρείται εντονότερα στον αναπτυσσόμενο κόσμο. Πιο συγκεκριμένα, οι αναπτυσσόμενες χώρες κατέχουν το 90% των οδικών ατυχημάτων, ενώ αντιπροσωπεύουν μόλις το 50% της οδικής κυκλοφορίας στον κόσμο. Τα οδικά ατυχήματα σύντομα θα καταστούν η πέμπτη κυριότερη αιτία θανάτου στις χώρες αυτές, η οποία θα υπερβαίνει το HIV / AIDS, την ελονοσία, τη φυματίωση και άλλες γνωστές αιτίες θανάτου, σύμφωνα με την πιο πρόσφατη μελέτη του Παγκόσμιου Οργανισμού Υγείας (World Health Organization, WHO) με τίτλο «Global Burden of Disease»[1].

Όσον αφορά την Ευρώπη, το 2016 ο ετήσιος αριθμός νεκρών από οδικά ατυχήματα στην Ευρώπη των 27 χωρών ήταν 22.070 και στην Ελλάδα 824. Παρατηρείται ότι, στην Ελλάδα αντιστοιχεί το 3.74% του συνολικού αριθμού ατυχημάτων της Ευρώπη των 27 χωρών¹. Μάλιστα, την τελευταία δεκαετία σημειώνεται μείωση του αριθμού των ανθρώπινων απωλειών εξαιτίας οδικών ατυχημάτων, τόσο στα πλαίσια της Ευρώπης όσο και στην Ελλάδα.

Στην μελέτη του παραπάνω φαινομένου, υψίστης σημασίας είναι η συνεισφορά του κλάδου της στατιστικής τόσο στην επεξεργασία των μετρήσεων όσο και στην αποδοτικότερη παρουσίαση των αποτελεσμάτων. Μέσω της στατιστικής έχουν διαμορφωθεί ποικίλα εργαλεία τα οποία συμβάλλουν στη μελέτη του φαινομένου αυτού τόσο με τη μορφή δεικτών όσο και σχέσεων μεταξύ μεγεθών. Χαρακτηριστικά μεγέθη για την παρούσα μελέτη αποτελούν ο αριθμός των νεκρών ανά εκατομμύριο πληθυσμού για το σύνολο των 27 χωρών της Ευρώπης καθώς και οι δείκτες οι οποίοι υποδεικνύουν τη μεταβολή των διαφόρων μεγεθών στο πέρασμα των χρόνων.

Παρακάτω παρατίθενται κάποιοι ενδεικτικοί πίνακες τέτοιων μεγεθών μαζί με τα αντίστοιχα διαγράμματά τους. Οι πίνακες αυτοί καθώς επίσης και τα αντίστοιχα διαγράμματα προέκυψαν ύστερα από συλλογή των δεδομένων από τις διάφορες βάσεις δεδομένων και την κατάλληλη επεξεργασία τους, προκειμένου να καταστεί κατανοητότερο το θέμα της παρούσας Διπλωματικής Εργασίας και τα σχετικά σε αυτό μεγέθη.

Σύμφωνα με τον παρακάτω πίνακα είναι εμφανής η μείωση, κατά μέσο όρο, των θυμάτων των τροχαίων ατυχημάτων με το πέρασμα των χρόνων.

¹Τα δεδομένα προήλθαν από τον παρακάτω σύνδεσμο « https://ec.europa.eu/transport/road_safety/road-safety-facts-figures-0_en»

Country	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2015-2016
AT	83	82	76	66	62	63	54	51	56	50	0.11
BE	101	88	88	77	78	69	65	65	65	56	0.14
BG	133	141	121	105	89	82	83	91	98	99	-0.01
CY	117	106	89	73	85	59	51	52	67	54	0.19
CZ	119	104	86	77	74	71	62	65	70	58	0.17
DE	60	54	51	45	49	45	41	42	43	39	0.09
DK	75	74	55	46	40	30	34	32	31	37	-0.19
EE	146	99	73	59	76	66	61	59	51	54	-0.06
EL	145	139	130	112	103	89	80	73	73	76	-0.04
ES	85	68	59	53	44	41	36	36	36	39	-0.08
FI	72	65	52	51	54	47	48	42	49	47	0.04
FR	73	67	66	64	63	58	51	53	54	54	0.00
HR	144	154	127	99	97	91	86	73	82	73	0.11
HU	122	99	82	74	64	61	60	63	65	62	0.05
IE	78	63	53	47	41	35	41	42	35	39	-0.11
IT	88	81	72	70	65	63	57	56	56	54	0.04
LT	228	155	116	95	97	101	86	91	83	66	0.20
LU	97	72	97	64	64	65	84	64	64	56	0.13
LV	190	144	117	103	86	87	88	106	95	80	0.16
MT	30	22	37	31	39	22	40	24	26	53	-1.04
NL	43	41	39	32	33	34	28	28	31	31	0.00
PL	146	143	120	102	110	93	88	84	77	80	-0.04
PT	84	80	80	84	68	61	61	61	57	54	0.05
RO	133	148	137	117	100	102	93	91	95	97	-0.02
SE	52	43	39	28	34	30	27	28	27	27	0.00
SI	146	106	84	67	69	63	61	52	58	63	-0.09
SK	123	113	71	65	60	65	46	54	57	51	0.11
UK	50	43	38	30	31	28	28	29	28	28	0.00
EU	87	79	70	63	61	56	52	51.4	51.6	50.5	0.02

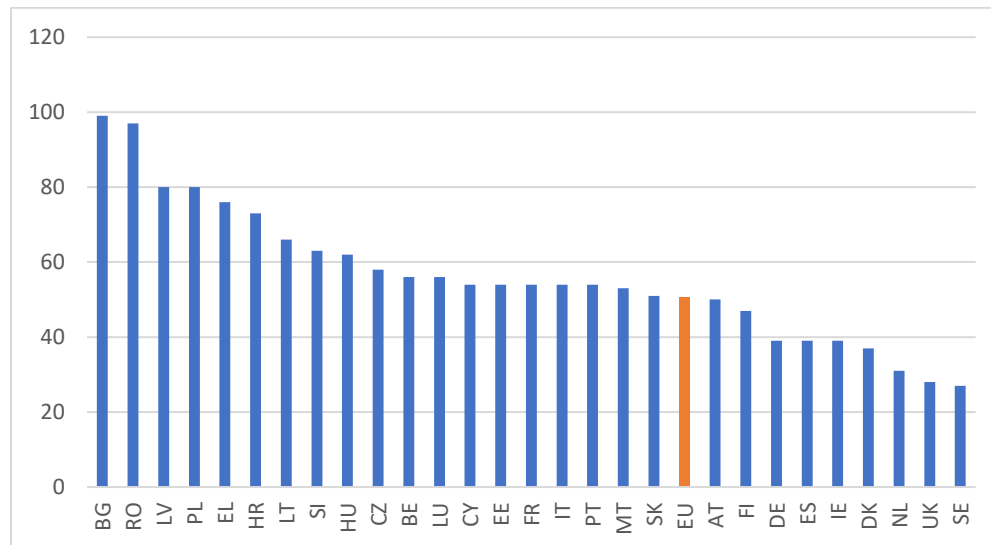
Πίνακας 1.1 Αριθμός νεκρών ανά εκατομμύριο πληθυσμού στην Ε.Ε την περίοδο 2007-2016²

Από τον παραπάνω πίνακα φαίνεται πως η Ελλάδα βρίσκεται στις πρώτες χώρες σε ότι αφορά τον αριθμό των νεκρών εξαιτίας οδικών ατυχημάτων γεγονός το οποίο καθίσταται ακόμη πιο εμφανές για το έτος 2016 με τη βοήθεια του παρακάτω διαγράμματος. Παράλληλα, βοριοευρωπαϊκές

² Τα δεδομένα προήλθαν από τον παρακάτω σύνδεσμο

«https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/statistics/historical_evol_popul.pdf»

χώρες όπως η Αυστρία, η Φινλανδία, η Ιρλανδία, η Νορβηγία και η Σουηδία φαίνεται πως έχουν μικρότερο αριθμό νεκρών ανά εκατομμύριο πληθυσμού από το μέσο όρο της Ευρώπης και αυτό εξαιτίας των ποικίλων και αυστηρών μέτρων αντιμετώπισης του φαινομένου που έχουν λάβει τα τελευταία χρόνια. Ιδιαίτερη εντύπωση προκαλεί και η θέση της Ισπανίας στις χώρες με το μικρότερο αριθμό θυμάτων από τροχαία ατυχήματα, γεγονός που οφείλεται στην πολιτική οδικής ασφάλειας την οποία είχε εφαρμόσει μεταξύ 2000-2010.



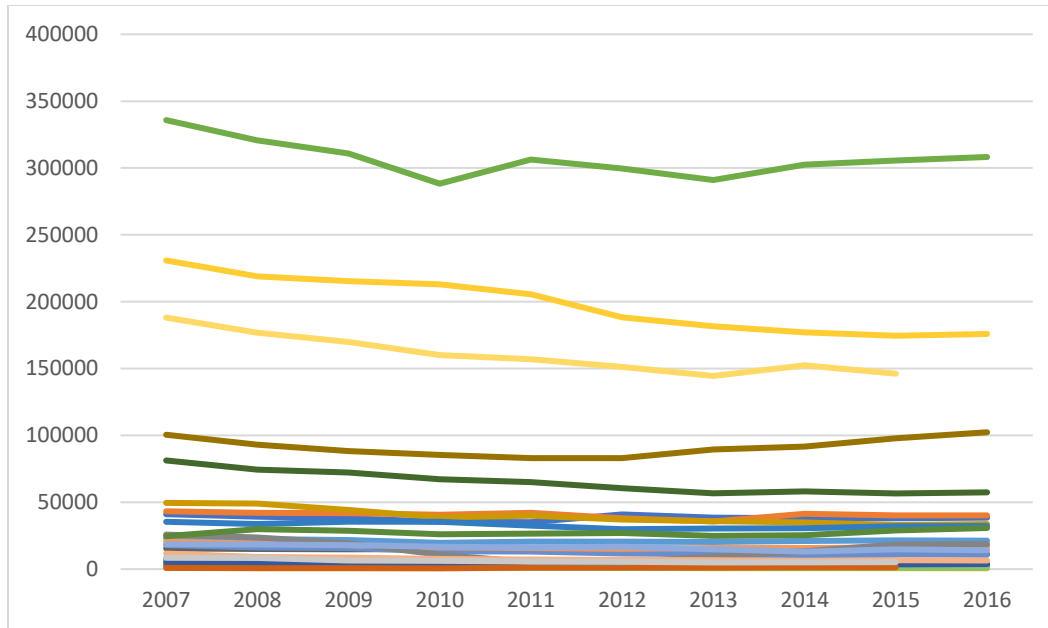
Διάγραμμα 1.1 Αριθμός νεκρών ανά εκατομμύριο πληθυσμού για το έτος 2016 για τις 28 χώρες της Ε.Ε.

Στη συνέχεια παρατίθεται και ο πίνακας του πλήθους των τροχαίων ατυχημάτων στην Ε.Ε. τα χρόνια 2008-2016.

Από τον παρακάτω πίνακα και το αντίστοιχο διάγραμμα φαίνεται ότι το πλήθος των τροχαίων ατυχημάτων διατηρείται γύρω από κάποιες σταθερές τιμές ανά χώρα χωρίς να υπάρχουν ιδιαίτερες αποκλίσεις μεταξύ των ετών.

Country	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
AT	41096	39173	37925	35348	35129	40831	38502	37957	37960	38466
BE	43239	42115	41944	40569	42119	38057	35632	41481	40303	40096
BG	8010	8045	7068	6610	6638	6717	7016	7015	7226	-
CY	1468	1392	1197	1198	1058	919	774	758	660	650
CZ	23060	22481	21706	19675	20487	20503	20342	21054	21561	21387
DE	335845	320614	310806	288297	306266	299637	291105	302435	305659	308145
DK	5549	5020	4174	3498	3525	3124	2984	2881	2853	2882
EE	2449	1868	1506	1348	1508	1383	1382	1436	1391	-
EL	15499	15083	14789	15032	13849	12398	12109	11690	11440	11318
ES	100508	93161	88251	85503	83027	83115	89519	91570	97756	102362
FI	6657	6881	6414	6072	6408	5725	5334	5299	5164	-
FR	81272	74487	72315	67288	65024	60437	56812	58191	56600	57515
HR	18033	16290	15731	13274	13229	11774	11228	10607	11038	10779
HU	20634	19174	17863	16308	15827	15174	15691	15847	16331	16627
IE	5467	6736	6615	5779	5230	5610	4976	5797	-	-
IT	230871	218963	215403	212997	205638	188228	181660	177031	174539	175791
LT	6448	4796	3827	3530	3266	3392	3391	3256	3031	-
LU	954	927	869	876	962	1019	949	908	983	941
LV	4780	4196	3160	3193	3386	3358	3489	3728	3692	3792
MT	942	764	636	577	1140	1270	1208	1449	1377	-
NL	25819	23708	19378	10778	5134	4966	9522	13358	18523	18749
PL	49536	49054	44195	38832	40069	37046	35847	34970	32967	33664
PT	35311	33613	35484	35426	32541	29867	30339	30604	31953	32299
RO	24661	29861	28612	25995	26647	26928	24827	25355	28944	30751
SE	18548	18462	18027	16627	16274	16636	14942	13091	14703	14086
SI	11640	9165	8717	7659	7257	6857	6568	6263	6578	6494
SK	8483	8343	6461	6131	5775	5370	5111	5064	5172	-
UK	188105	176814	169805	160080	157068	151346	144480	152407	146203	-
EU	1314884	1251186	1202878	1128500	1124481	1081687	1055739	1081502	1084607	926794

Πίνακας 1.2 Τροχαία ατυχήματα στην Ε.Ε. έτη 2008-2016



Διάγραμμα 1.2 Τροχαία ατυχήματα έτη 2008-2016 στην Ε.Ε.

Ιδιαίτερο ενδιαφέρον έγκειται στην ερμηνεία και αιτιολόγηση των προαναφερθέντων αποτελεσμάτων. Επιχειρώντας μία ιστορική αναδρομή διαπιστώνει κανείς ότι από τις αρχές της δεκαετίας του '70, οι περισσότερες βόρειες και δυτικές ευρωπαϊκές χώρες είχαν ξεκινήσει συντονισμένες προσπάθειες βελτίωσης της οδικής ασφάλειας τόσο σε εθνικό όσο και σε τοπικό επίπεδο. Αποτέλεσμα αυτών των προσπαθειών ήταν μία γενική τάση μείωσης του αριθμού των νεκρών στις χώρες της Ε.Ε., με εξαίρεση την Ελλάδα, την Ισπανία και την Πορτογαλία λόγω της μεγάλης αύξησης του αριθμού των οχημάτων και στις οποίες η πτωτική τάση ακολούθησε αργότερα.

Εκτός, όμως, από τις πρωτοβουλίες σε επίπεδο Ευρωπαϊκής Ένωσης αλλά και του κάθε κράτους μέλους ξεχωριστά η μείωση που παρατηρείται ίσως οφείλεται και στην οικονομική κρίση που ξεκίνησε το 2008 και επηρέασε θετικά την εξέλιξη της οδικής ασφάλειας μέσω μιας σειράς επιπτώσεων όπως μείωση της κινητικότητας, λιγότερη κίνηση βαρέων οχημάτων, λιγότεροι άπειροι και ηλικιωμένοι οδηγοί με σχετικά υψηλότερους κινδύνους, μείωση ελεύθερου χρόνου οδήγησης, ασφαλέστερη συμπεριφορά οδήγησης. Τα οικονομικά μεγέθη μιας χώρας έχουν σημαντική επιρροή στην οδική ασφάλεια, όμως θα ήταν εσφαλμένο να θεωρηθεί ότι είναι και ο μοναδικός παράγοντας επιρροής της. Δείκτες με κοινωνικό χαρακτήρα, όπως η ανεργία, το επίπεδο μόρφωσης των πολιτών, ο δείκτης ανθρώπινης ανάπτυξης (HDI), τα παραγωγικά χρόνια ζωής των ανθρώπων και ο βαθμός ικανοποίησης από τη ζωή διαδραματίζουν με τη σειρά τους σημαντικό ρόλο στη διακύμανση των επιδόσεων της οδικής ασφάλειας από κράτος σε κράτος και όλοι μαζί αναδεικνύουν την πολυπλοκότητα και ρευστότητα που διακρίνει το συγκεκριμένο ζήτημα καθώς επίσης και τη δυσκολία στη δημιουργία αξιόπιστων προβλέψεων για τη μελλοντική πορεία της οδικής ασφάλειας.

Παρόλα αυτά υπάρχουν και άλλοι παράγοντες οι οποίοι δύναται να επηρεάζουν τον αριθμό των τροχαίων ατυχημάτων και κατ' επέκταση το βαθμό θνησιμότητας, πέραν των κοινωνικών και οικονομικών παραγόντων και αυτοί δεν είναι άλλοι από παράγοντες σχετιζόμενους με την υγεία. Συνεπώς, στην παρούσα Διπλωματική Εργασία επιχειρείται μία προσπάθεια αναζήτησης των σχέσεων διαφόρων ιατρικών, οικονομικών και κοινωνικών δεικτών με τον αριθμό των νεκρών σε οδικά ατυχήματα και ανάδειξης των δεσμών που συνδέουν αυτά τα μεγέθη, με ιδιαίτερη έμφαση στον τρόπο επηρεασμού των υπολοίπων παραγόντων από δείκτες σχετικούς με την υγεία και την νοσηλεία.

1.2 Στόχος της Διπλωματικής Εργασίας

Στόχος της παρούσας Διπλωματικής Εργασίας είναι η συσχέτιση των επιδόσεων της οδικής ασφάλειας με ιατρικούς, οικονομικούς και κοινωνικούς δείκτες στην Ευρωπαϊκή Ένωση.

Επιχειρείται δηλαδή μία προσπάθεια **σύνδεσης του αριθμού των νεκρών σε οδικά ατυχήματα** στις χώρες της Ευρωπαϊκής Ένωσης με **ιατρικά μεγέθη**, όπως είναι οι δαπάνες στον τομέα της υγείας και ο αριθμός των κλινών, με **οικονομικά μεγέθη**, όπως είναι το κατά κεφαλήν ακαθάριστο εγχώριο προϊόν (Α.Ε.Π.), και με **κοινωνικούς δείκτες**, όπως ο δείκτης ανθρώπινης ανάπτυξης (Human Development Index HDI) ο οποίος καθορίζεται από το προσδόκιμο ζωής, το επίπεδο μόρφωσης και την ποιότητα ζωής σε μία χώρα.

Για την επίτευξη του στόχου αυτού, αναπτύσσονται και εφαρμόζονται μαθηματικά μοντέλα που τεκμηριώνουν τη σχέση της οδικής ασφάλειας με ιατρικούς, οικονομικούς και κοινωνικούς δείκτες τόσο για το σύνολο των χωρών της Ευρωπαϊκής Ένωσης όσο και για τρεις ομάδες κρατών που δημιουργήθηκαν από τη διαίρεση των είκοσι επτά κρατών μελών με βάση τον πληθυσμό, το κατά κεφαλήν Α.Ε.Π. και τον αριθμό νεκρών από οδικά ατυχήματα ανά εκατομμύριο πληθυσμού της κάθε χώρας. Την διαίρεση των χωρών την χρησιμοποιούμε έτοιμη σύμφωνα με την εργασία της Υπατίας Μίχου[2].

Πιο συγκεκριμένα, για τις ανάγκες της ανάπτυξης των προαναφερθέντων μαθηματικών μοντέλων αξιοποιήθηκαν οι εξής μεταβλητές για τα έτη 2008-2014 :

- ο αριθμός των νεκρών σε οδικά ατυχήματα σύμφωνα με τη EUROSTAT
- το πλήθος των θανάσιμων οδικών ατυχημάτων από την CARE database
- το πλήθος των οδικών ατυχημάτων εντός αστικού δικτύου από την CARE database
- το προσδόκιμο ζωής των ανθρώπων από τη EUROSTAT
- οι δαπάνες για τον τομέα της υγείας ως προς το Α.Ε.Π. από την World bank
- οι δαπάνες για τον τομέα της υγείας ως προς το πλήθος των πολιτών από την World bank
- οι δημόσιες δαπάνες για τον τομέα της υγείας από τους World Development Indicators
- ο αριθμός κλινών στα νοσοκομεία από τη EUROSTAT
- ο πληθυσμός των χωρών σύμφωνα με την EUROSTAT
- το κατά κεφαλήν Α.Ε.Π σύμφωνα με τη EUROSTAT
- ο δείκτης ανθρώπινης ανάπτυξης με στοιχεία από το πρόγραμμα ανάπτυξης των Ηνωμένων Εθνών

- τα παραγωγικά χρόνια ζωής των ανθρώπων από τη EUROSTAT
- ο δείκτης ανεργίας από τη EUROSTAT

Επισημαίνεται ωστόσο ότι οι παράγοντες που επηρεάζουν την οδική ασφάλεια σε κάθε χώρα είναι πολυάριθμοι και πολυδιάστατοι, συνεπώς δεν περιορίζονται μόνο σε δείκτες ιατρικής, οικονομικής και κοινωνικής φύσεως αλλά εκτείνονται από στοιχεία όπως ο αριθμός κυκλοφορούντων οχημάτων μέχρι και κλιματολογικούς παράγοντες.

1.3 Μεθοδολογία

Για την εκπόνηση της παρούσας Διπλωματικής Εργασίας ακολουθήθηκε συγκεκριμένη μεθοδολογία, τα στάδια της οποίας παρουσιάζονται στη συνέχεια

Αρχικά μετά την οριστικοποίηση του επιδιωκόμενου στόχου πραγματοποιήθηκε βιβλιογραφική ανασκόπηση τόσο σε ελληνικό όσο και σε διεθνές επίπεδο. Στη φάση αυτή πραγματοποιήθηκε αναζήτηση παρεμφερών ερευνών, επιστημονικών άρθρων καθώς και γενικών πληροφοριών σχετικά με το εξεταζόμενο αντικείμενο που θα μπορούσαν να φανούν χρήσιμες για τη συγκεκριμένη έρευνα. Μέσω των ερευνών αυτών καταβλήθηκε προσπάθεια να αποκτηθεί μια σχετική εμπειρία στην επεξεργασία τέτοιων θεμάτων, καθώς επίσης και να αποφασιστεί η μέθοδος με βάση, την οποία θα πραγματοποιηθεί η επεξεργασία των στοιχείων και να επιτευχθεί ο επιδιωκόμενος στόχος.

Ύστερα από τη μελέτη των δεδομένων που συγκεντρώθηκαν από τη βιβλιογραφική ανασκόπηση, πραγματοποιήθηκε η συλλογή των στοιχείων που απαιτούνταν για την εκπόνηση της Διπλωματικής Εργασίας. Ενδεικτικά μερικά από τα στοιχεία που συλλέχθηκαν ήταν για τις 27 χώρες μέλη της Ε.Ε. και για τη χρονική περίοδο 2008 έως 2014 ο πληθυσμός, ο αριθμός των νεκρών στα οδικά ατυχήματα, το κατά κεφαλήν Α.Ε.Π. και οι δαπάνες για τον τομέα της υγείας. Για τη συλλογή τους αξιοποιήθηκαν βάσεις δεδομένων της ευρωπαϊκής στατιστικής υπηρεσίας (EUROSTAT), όπως επίσης και άλλων διεθνών οργανισμών, όπως ο οργανισμός οικονομικής συνεργασίας και ανάπτυξης (OECD), η παγκόσμια τράπεζα (World Bank) και το πρόγραμμα ανάπτυξης των Ηνωμένων Εθνών (UNDP). Αξίζει να σημειωθεί ότι στα στοιχεία που συλλέχθηκαν πραγματοποιήθηκε ο σχετικός έλεγχος της αξιοπιστίας τους, έτσι ώστε να είναι ομοιόμορφα και συγκρίσιμα.

Στο επόμενο στάδιο, ταξινομήθηκαν τα συλλεχθέντα στοιχεία σε μια ενιαία βάση δεδομένων, στην οποία για κάθε κράτος και για συγκεκριμένο έτος αντιστοιχούν τα κατάλληλα δεδομένα. Σε περιορισμένο αριθμό κρατών και ετών τα δεδομένα ήταν ελλιπή, με αποτέλεσμα να μην μπορούν να χρησιμοποιηθούν για την εξαγωγή του μαθηματικού μοντέλου, εξ 'ου και ο δειγματικός χώρος περιορίστηκε στις 27 χώρες μέλη και στο χρονικό πλαίσιο 2008-2014. Παράλληλα, τα στοιχεία υπέστησαν κατάλληλη επεξεργασία για τη χρησιμοποίησή τους στο επόμενο στάδιο.

Την επεξεργασία των στοιχείων ακολούθησε η στατιστική τους ανάλυση με τη χρήση ειδικού στατιστικού λογισμικού. Αρχικά, χρησιμοποιήθηκε το μοντέλο της λογαριθμισμένης γραμμικής παλινδρόμησης με εξαρτημένη μεταβλητή την ετήσια μεταβολή του αριθμού των νεκρών στα οδικά ατυχήματα ως προς τον πληθυσμό και ανεξάρτητες μεταβλητές το προσδόκιμο ζωής, το

HDI και την ετήσια μεταβολή του κατά κεφαλήν Α.Ε.Π.. Ύστερα από δοκιμές που πραγματοποιήθηκαν τόσο για το σύνολο των κρατών όσο και για τις τρεις ομάδες στις οποίες χωρίστηκαν, προέκυψαν ποικίλα μαθηματικά μοντέλα που περιέγραφαν άλλα σε μεγαλύτερο και άλλα σε μικρότερο επίπεδο εμπιστοσύνης την επιρροή των παραπάνω ανεξάρτητων μεταβλητών στη μεταβολή του αριθμού των νεκρών στα οδικά ατυχήματα.

Προκειμένου όμως να ληφθεί υπόψη η επίδραση του κάθε κράτους ή της κάθε ομάδας κρατών στα τελικά μοντέλα, επιλέχθηκε να χρησιμοποιηθεί ένα μοντέλο που θα συνυπολόγιζε την αλληλεπίδραση των ανεξάρτητων μεταβλητών. Η τελική επεξεργασία επομένως έγινε με το Γραμμικό Μικτό Μοντέλο. Στη φάση αυτή συσχετίστηκαν η μεταβολή του αριθμού των νεκρών στα οδικά ατυχήματα ως προς τον πληθυσμό (ως εξαρτημένη μεταβλητή), με το αριθμό κλινών στα νοσοκομεία ως προς τον πληθυσμό, τον δείκτη ανθρώπινης ανάπτυξης HDI, τα παραγωγικά χρόνια ζωής των ανθρώπων καθώς και με το κράτος ή την ομάδα των κρατών και τέλος με το έτος ως σταθερό παράγοντα για κάθε μοντέλο.

Μετά το πέρας της παραπάνω διαδικασίας, προέκυψαν τέσσερις εξισώσεις, μία για το σύνολο των κρατών και μία για κάθε ομάδα κρατών ανάλογα με τη γεωγραφική τους κατανομή. Μετά την αξιολόγηση και ερμηνεία των αποτελεσμάτων, εξήχθησαν τα αντίστοιχα συμπεράσματα για τον βαθμό και τον τύπο της επιρροής των εκάστοτε ανεξάρτητων μεταβλητών στην εξαρτημένη, καθώς και για τη σύγκριση μεταξύ των χωρών της Ε.Ε.. Έτσι, προέκυψαν σημαντικές πληροφορίες για το υπό εξέταση πρόβλημα καθώς και διατυπώθηκαν αξιόλογες προτάσεις για περαιτέρω έρευνα.

Στη συνέχεια παρουσιάζονται σχηματικά τα στάδια της μεθοδολογίας που ακολουθήθηκαν για τις ανάγκες της παρούσας Διπλωματικής Εργασίας.



Διάγραμμα 1.3 Σχηματική απεικόνιση σταδίων Διπλωματικής Εργασίας

1.4 Δομή Διπλωματικής Εργασίας

Το **πρώτο κεφάλαιο** αποτελεί την εισαγωγή της Διπλωματικής Εργασίας και έχει σκοπό να παρουσιάσει στον αναγνώστη το γενικότερο πλαίσιο του αντικειμένου με το οποίο ασχολείται. Ξεκινά με μία αναφορά στις έννοιες οδικές μεταφορές, οδικά ατυχήματα και οδική ασφάλεια, ενώ επισημαίνεται η σημαντική μείωση που παρατηρείται την τελευταία δεκαετία στην εξέλιξη των οδικών ατυχημάτων στην Ευρώπη. Στη συνέχεια παρουσιάζεται ο επιδιωκόμενος στόχος της Διπλωματικής Εργασίας και η μεθοδολογία που ακολουθήθηκε για την εκπόνησή της. Ολοκληρώνεται με την παρούσα αναφορά στη δομή της .

Στο **δεύτερο κεφάλαιο** παρουσιάζονται τα αποτελέσματα της βιβλιογραφικής ανασκόπησης τα οποία προέκυψαν από την αναζήτηση και καταγραφή ερευνών με αντικείμενο συναφές με αυτό της Διπλωματικής Εργασίας. Εξετάζονται έρευνες από την Ελλάδα και το εξωτερικό, οι οποίες έχουν δημοσιευθεί σε επιστημονικά περιοδικά, βιβλία, άρθρα και πρακτικά συνεδρίων, ενώ δίνεται ιδιαίτερη έμφαση στο πλαίσιο έρευνας, τη μεθοδολογία και τα βασικά αποτελέσματά τους. Στο τέλος παρατίθεται σύνθεση των βασικών σημείων όλων των εργασιών που εξετάστηκαν με στόχο αφενός μεν να προκύπτει η αναγκαιότητα εξέτασης του αντικειμένου της Διπλωματικής Εργασίας και αφετέρου να αιτιολογείται η επιλογή της υιοθετηθείσας μεθοδολογίας.

Στο **τρίτο κεφάλαιο**, το οποίο αποτελεί το θεωρητικό υπόβαθρο της Διπλωματικής Εργασίας, παρουσιάζεται και αναλύεται σε βάθος η επιλεγείσα μεθοδολογία, ενώ δίνεται ιδιαίτερη έμφαση στις μαθηματικές και στατιστικές θεωρίες στις οποίες αυτή βασίζεται.

Στο **τέταρτο κεφάλαιο** περιγράφεται η διαδικασία με την οποία συλλέχθηκαν και επεξεργάστηκαν τα στοιχεία στα οποία στηρίχθηκε η Διπλωματική Εργασία. Παρουσιάζονται οι πηγές των στοιχείων, το εύρος τους (περίοδος και περιοχή αναφοράς) καθώς και τα βασικά χαρακτηριστικά τους (υπο-κατηγορίες κλπ.). Επιπλέον περιγράφεται η κωδικοποίηση και ο τρόπος εισαγωγής τους στον ηλεκτρονικό υπολογιστή, ενώ γίνεται σύντομη αναφορά και στα ηλεκτρονικά προγράμματα που αξιοποιήθηκαν καθώς και στα κρίσιμα σημεία της λειτουργίας αυτών.

Το **πέμπτο κεφάλαιο** περιλαμβάνει την αναλυτική περιγραφή της εφαρμογής της μεθοδολογίας και την παρουσίαση του συνόλου των αποτελεσμάτων της. Περιγράφεται η διαδικασία ανάπτυξης των μαθηματικών μοντέλων και αναλύονται τα τελικά μοντέλα στα οποία κατέληξε η διαδικασία αυτή, ενώ επιχειρείται και η ερμηνεία των αποτελεσμάτων.

Στο **έκτο κεφάλαιο**, το οποίο αποτελεί το σημαντικότερο της Διπλωματικής Εργασίας, παρατίθενται τα συμπεράσματα που προέκυψαν από την ερμηνεία των εξαγόμενων μοντέλων. Επισημαίνεται η χρησιμότητα των βασικών αποτελεσμάτων και κατατίθενται προτάσεις οι οποίες αφορούν στην αξιοποίηση των αποτελεσμάτων της Διπλωματικής Εργασίας και στα περιθώρια για περαιτέρω έρευνα στο συγκεκριμένο τομέα.

2. Βιβλιογραφική Ανασκόπηση

2.1 Εισαγωγή

Το παρόν κεφάλαιο αφορά στη βιβλιογραφική ανασκόπηση που πραγματοποιήθηκε για τις ανάγκες της παρούσας Διπλωματικής Εργασίας και περιλαμβάνει τα αποτελέσματα που προέκυψαν από έρευνες συναφείς τόσο με το αντικείμενο της παρούσας εργασίας όσο και με τη μεθοδολογία που αξιοποιήθηκε. Συγκεκριμένα παρουσιάζονται έρευνες που αφορούν στη συσχέτιση των οδικών ατυχημάτων με ιατρικούς, οικονομικούς και κοινωνικούς δείκτες οι οποίες αναζητήθηκαν σε βάσεις δεδομένων όπως η «pubmed» με τη χρήση λέξεων κλειδιών όπως «road traffic accidents», «crashes», «healthcare», «epidemiology» και «regression analysis». Καθώς τα αρχικά αποτελέσματα δεν προσέφεραν τα αναμενόμενα εργαλεία και μεθοδολογίες, η έρευνα εξειδικεύτηκε στις έρευνες των τελευταίων δέκα ετών, διατηρώντας την κλίμακα σε παγκόσμιο επίπεδο. Έτσι εντοπίστηκαν οι πρώτες σχετικές έρευνες οι οποίες οδήγησαν την αναζήτηση προς έρευνες σχετικές με τα οδικά ατυχήματα και τις επιπτώσεις τους στον άνθρωπο όσον αφορά τον ιατρικό κλάδο και τις κοινωνικο-οικονομικές συνέπειες. Αναφορικά με τις μεθοδολογίες, η αναζήτηση βασίστηκε και στις προηγούμενες μελέτες αλλά κυρίως σε μελέτες σχετικές με τα οδικά ατυχήματα και κοινωνικούς και οικονομικούς παράγοντες. Τα τελικά στοιχεία που εξετάζονται προκύπτουν από έρευνες που έχουν πραγματοποιηθεί στην Ελλάδα αλλά και στο εξωτερικό. Για κάθε εργασία γίνεται συνοπτική αναφορά στη μεθοδολογία που ακολουθήθηκε και στα συμπεράσματα που προέκυψαν. Τέλος, με βάση τη βιβλιογραφική ανασκόπηση προσδιορίστηκε το ακριβές αντικείμενο της Διπλωματικής Εργασίας και επιχειρήθηκε να επιλεγεί η καταλληλότερη μεθοδολογία.

2.2 Συναφείς Έρευνες και Μεθοδολογίες

Στο υποκεφάλαιο αυτό παρουσιάζονται έρευνες συναφείς με το αντικείμενο της Διπλωματικής Εργασίας με στόχο τον προσδιορισμό ενός αντικειμένου το οποίο δεν έχει καλυφθεί πλήρως (τουλάχιστον στην Ελλάδα), συμπληρώνοντας με αυτόν τον τρόπο τις υπάρχουσες εργασίες. Επιπλέον επιτρέπει τον έλεγχο ώστε να διαπιστωθεί εάν τα αποτελέσματα της Διπλωματικής Εργασίας συμφωνούν με εκείνα της διεθνούς βιβλιογραφίας. Αντίστοιχα η ανασκόπηση συναφών μεθοδολογιών έχει στόχο την επιλογή της καταλληλότερης μεθοδολογίας για την αντιμετώπιση του αντικειμένου της παρούσας Διπλωματικής Εργασίας.

Το **1986** δημοσιεύθηκε η έρευνα των Jacobs και Cutting[3], οι οποίοι διεξήγαγαν μια διατμηματική-συγχρονική μελέτη για να εξετάσουν τη **σχέση ανάμεσα στα ποσοστά νεκρών σε οδικά ατυχήματα και κάποια οικονομικά και κοινωνικά χαρακτηριστικά** σε επιλεγμένες αναπτυσσόμενες χώρες. Αυτά τα χαρακτηριστικά περιλαμβάνουν εκτός από το κατά κεφαλήν Α.Ε.Π., τον αριθμό κυκλοφορούντων οχημάτων, την πυκνότητα του οδικού δικτύου, την πυκνότητα οχημάτων ανά χιλιόμετρο οδικού δικτύου, τον πληθυσμό ανά ιατρό και τον πληθυσμό ανά νοσοκομειακή κλίνη. Τα αποτελέσματά τους έδειξαν ότι τα ποσοστά νεκρών σε οδικά

ατυχήματα δε σχετίζονται μόνο με το **κατά κεφαλήν Α.Ε.Π.** αλλά και με την **πυκνότητα οχημάτων** και τον **πληθυσμό ανά νοσοκομειακή κλίνη.**

Το **2009** δημοσιεύθηκε η έρευνα των Sánchez-Mangas R, García-Ferrrer A, de Juan A και Arroyo AM[4] αναφορικά με την πιθανότητα θνησιμότητας στα οδικά ατυχήματα και τη σημασία της γρήγορης ιατρικής ανταπόκρισης. Σύμφωνα με την έρευνα, ο αριθμός των θανάτων σε οδικά ατυχήματα στην Ισπανία υπερβαίνει τους τρεις χιλιάδες ανθρώπους κάθε χρόνο. Οι δημόσιες αρχές έχουν εφαρμόσει ορισμένες πολιτικές με στόχο τη μείωση αυτού του αριθμού. Μεταξύ αυτών, βρίσκονται η βελτίωση των προτύπων ορθής οδικής κυκλοφορίας και ορισμένες νομικές αλλαγές που ενθαρρύνουν την προσεκτική οδική συμπεριφορά. Ωστόσο, λιγότερη προσοχή έχει δοθεί στην ταχεία ιατρική περίθαλψη έκτακτης ανάγκης, η οποία μπορεί να είναι κρίσιμη για τη μείωση του αριθμού των θανατηφόρων ατυχημάτων που προκλήθηκαν από τροχαία ατυχήματα. Στην εργασία αυτή, χρησιμοποιήθηκε δείγμα περισσότερων των 1400 ατυχημάτων που σημειώθηκαν στους ισπανικούς δρόμους τον Μάιο του 2004. Στόχος είναι η ανάλυση του βαθμού στον οποίο η μείωση του χρονικού διαστήματος μεταξύ της σύγκρουσης και της άφιξης των υπηρεσιών έκτακτης ανάγκης στη σκηνή σύγκρουσης σχετίζεται με χαμηλότερη πιθανότητα θανάτου. Τα αποτελέσματά υποδηλώνουν ότι η **μείωση του ιατρικού χρόνου απόκρισης των 10 λεπτών** μπορεί στατιστικά να συνδυαστεί με μια **μέση μείωση της πιθανότητας θανάτου κατά ένα τρίτο**, τόσο σε αυτοκινητόδρομους όσο και σε συμβατικούς δρόμους.

Το **2012** δημοσιεύθηκε η έρευνα των Zuhir Bodalal, Riyad Bendardaf και Mohammed Ambarek[5] αναφορικά με την **παρατήρηση και την μελέτη των τάσεων των οδικών ατυχημάτων** (RTA's) τα τελευταία δέκα χρόνια στη Βεγγάζη της Λιβύης. Στα πλαίσια της έρευνας αυτή πραγματοποιήθηκε μια αναδρομική ανάλυση χρησιμοποιώντας τα αρχεία ασθενών του νοσοκομείου Al-Jalaa (του κύριου κέντρου τραυματισμών στη Βεγγάζη) με περισσότερες από 21.753 περιπτώσεις RTA, όπου αξιοποιώντας **εργαλεία στατιστικής ανάλυσης** όπως ο μέσος όρος, η τυπική απόκλιση, οι ποσοστιαίες μεταβολές όπως επίσης κατάλληλους πίνακες και διαγράμματα, προέκυψαν τα τελικά αποτελέσματα. Στη συνέχεια, τα ετήσια στοιχεία συγκρίθηκαν μεταξύ τους και παρατηρήθηκαν αλλαγές στις τάσεις. Από την έρευνα προέκυψε ότι τα RTA αντιπροσωπεύουν ένα αυξανόμενο ποσοστό του φορτίου του Al-Jalaa κατά τη διάρκεια των ετών. Περίπου το 41% των περιπτώσεων αυτών χρειάστηκε να υποβληθεί σε χειρουργική επέμβαση. Η νεότερη ηλικιακή ομάδα (ηλικίας 20-29 ετών) αποτελούσε την πλειονότητα των περιπτώσεων, ενώ υπήρχε μια τάση αύξησης της μέσης ηλικίας των ασθενών που συμμετείχαν σε κάποιο ατύχημα. Οι άρρενες ασθενείς βρέθηκαν νεότεροι σε ηλικία από τους αντίστοιχους θήλεις ασθενείς. Στην παρούσα έρευνα, οι άνδρες αποτελούσαν το 81,5%, ενώ οι γυναίκες το 18,5% των ασθενών με RTA. Ενώ, όσον αφορά τη διάρκεια των ασθενών, οι περισσότεροι ασθενείς παρέμειναν στο νοσοκομείο για λιγότερο από μια εβδομάδα. Οι επιβαίνοντες (οδηγοί και επιβάτες) εισήχθησαν συχνότερα από τους πεζούς. Κατά τη διάρκεια των ετών υπήρξε μια τάση για αυξημένη συμμετοχή των επιβατών οχημάτων και μείωση του ποσοστού των πεζών που έπρεπε να νοσηλευτούν. Επιπλέον, σημειώθηκε μείωση των θανάτων των πεζών. Συνολικά, οι περισσότεροι ασθενείς με RTA απελευθερώθηκαν και προτάθηκαν για παρακολούθηση σε κλινικές εξωτερικών ασθενών, ωστόσο υπήρξε μια αυξανόμενη τάση προς την αύξηση των διαφυγόντων ασθενών και των ασθενών οι οποίοι έφευγαν παρά τις αντιρρήσεις των ιατρών. Παρόλα αυτά, υπάρχουν τόσο ενθαρρυντικά συμπεράσματα, όσο και σημεία που απαιτούν περαιτέρω έμφαση και δράση. Σύμφωνα με την προαναφερθείσα μελέτη, πρέπει να εξετασθεί η

δημόσια εκπαίδευση, η κατάρτιση στη στήριξη της ζωής και η διαφοροποίηση των μεταφορών (εκτός από τη χρήση των οδών), ως ένα μέσο βελτίωσης της σημερινής κατάστασης.

Το **2013** δημοσιεύθηκε η έρευνα των Anna García-Altés, a Josep M Suelvesb και Eneko Barberí[6] σε συνεργασία με το Ινστιτούτο Έρευνας Βιοϊατρικής (IIB Sant Pau) και την ερευνητική ομάδα CIBERESP αναφορικά με την **αξιολόγηση της εφαρμογής νέων κανόνων οδικής συμπεριφοράς** μεταξύ 2000 και 2010 στην Καταλονία της Ισπανίας **και τον συσχετισμό τους με πιθανά οικονομικά κέρδη για την κοινωνία**. Σύμφωνα με τα αποτελέσματα της έρευνας για το διάστημα 2000-2010, παρατηρήθηκε μείωση των συγκρούσεων με θύματα, μείωση των περιστατικών όπου απαιτείτο νοσηλεία καθώς και έντονη μείωση του πλήθους των θανάτων λόγω τροχαίων ατυχημάτων (της τάξης του 57%). Σύμφωνα με τα αποτελέσματα υπολογίστηκε ότι τα συνολικά κέρδη ανέρχονται στο ποσό των 18000 εκατομμυρίων Ευρώ, από τα οποία το 97% προήλθε από μείωση χαμένης παραγωγικότητας. Από το υπόλοιπο 3%, το 63% σχετίζεται με κόστη νοσηλείας σε μονάδες εντατικής θεραπείας, το 15% με διαδικασίες προσαρμογής σε περιπτώσεις αναπηρίας ενώ το 8.1% σχετίζεται με έξοδα νοσοκομειακής περίθαλψης. Για την επεξεργασία των δεδομένων χρησιμοποιήθηκε η **μέθοδος Ανάλυσης Κόστους** με χρονικό ορίζοντα τα 10 χρόνια μεταξύ 2000 και 2010. Στην επεξεργασία συμπεριελήφθησαν το κόστος **εισαγωγής στο νοσοκομείο, χρήσης ασθενοφόρου, εντατικής θεραπείας, αστυνομίας, πυροσβεστικής, τροχαίας, προσαρμογής σε περιπτώσεις αναπηρίας, απώλειας παραγωγικότητας λόγω μόνιμης εισαγωγής σε ινστιτούτο αποκατάστασης, θανάτου και άδειας ασθενείας** καθώς επίσης και κόστη υλικών και διοικητικών δαπανών.

Το **2013** δημοσιεύθηκε η έρευνα των Jose ´ I. Castillo-Manzano • Mercedes Castro-Nun ˆo και Xavier Fageda[7] σχετικά με την **εκτίμηση των καθοριστικών παραγόντων των ποσοστών θνησιμότητας των οδικών μεταφορών στην Ευρώπη των είκοσι επτά χωρών**, για την περίοδο 1999-2009. Μέσω του ελέγχου των χαρακτηριστικών των διαφόρων χωρών και των πολιτικών οδικής ασφάλειας που εφαρμόζουν, εξετάζουμε την επίδραση των μεταβλητών που σχετίζονται με τα εθνικά συστήματα υγείας, τον αριθμό των νοσοκομειακών κλινών ανά τετραγωνικό χιλιόμετρο και το ποσοστό των δαπανών για την υγεία σε σχέση με το ακαθάριστο εγχώριο προϊόν (ΑΕΠ). Τα δεδομένα αυτά μοντελοποιήθηκαν μέσω ενός **μοντέλου στατιστική ανάλυσης fixed effect** της μορφής $Y_{it} = \alpha + \beta_k X_{it} + \lambda_k Z_{it} + \mu_i + \nu_t + \varepsilon_{it}$, όπου το Y_{it} αντιπροσωπεύει το λογάριθμο της θνησιμότητας εξαιτίας των οδικών ατυχημάτων προς τον πληθυσμό, το X_{it} είναι ο πίνακας με τα οικονομικά στοιχεία κάθε χώρας και το Z_{it} εκφράζει μεταβλητές σχετικές με τις εφαρμοζόμενες πολιτικές οδικής ασφάλειας. Τα μ_i είναι σταθεροί όροι ανά χώρα σχετικοί με χρονικώς ανεξάρτητες μεταβλητές οι οποίες παραλείπονται από την μελέτη, τα ν_t εκφράζουν ετήσιες τάσεις οι οποίες εμφανίζονται σε όλες τις χώρες του χρησιμοποιηθέντος δείγματος και ε_{it} είναι το μέσο μηδενικό σφάλμα. Σε συνέχεια της ανάλυσης του παραπάνω μοντέλου, διαπιστώνεται ότι η **πυκνότητα των νοσοκομειακών κλινών συμβάλλει στην πτώση των θανάτων που σχετίζονται με την κυκλοφορία**. Επιπλέον, η ποιότητα των γενικών ιατρικών εγκαταστάσεων και της τεχνολογίας που συνδέονται με την αύξηση των δαπανών για την υγεία δύναται να αποτελέσει σημαντικό παράγοντα για τη μείωση των θανάτων εξαιτίας οδικών ατυχημάτων.

Το **2013** δημοσιεύθηκε η έρευνα των Serkan Emre Erođlu, M.D., Siddika Nihal Toprak, M.D., Ebru Akođlu, M.D., ˆOzge Ecmel Onur, M.D., Arzu Denizbaşı, M.D., ˆCıđdem ˆOzpolat, M.D. και Haldun Akođlu, M.D.[8], αναφορικά με **τα ατυχήματα με μοτοσυκλέτες και τον αντίκτυπό τους**

στο κόστος της υγειονομικής περίθαλψης. Τα δεδομένα συλλέχθηκαν από τα τροχαία ατυχήματα τα οποία σημειώθηκαν στην Τουρκία το έτος 2011, εκ των οποίων το 7.58% ήταν ατυχήματα μοτοσυκλετιστών. Τα ατυχήματα με μοτοσυκλέτα που σημειώθηκαν με ή χωρίς σύγκρουση μεταξύ της 1ης Ιουλίου 2010 και της 30ης Ιουνίου 2011 μελετήθηκαν μακροπρόθεσμα μέσω της επιθεώρησης των ασθενών που επισκέφθηκαν την υπηρεσία έκτακτης ανάγκης. Τα έξοδα υγειονομικής περίθαλψης που αφορούν κάθε άτομο που τραυματίστηκε σε ατύχημα μοτοσυκλέτας διερευνήθηκαν μέσω εντύπων. Τα δεδομένα αναλύθηκαν χρησιμοποιώντας τις **συχρότητες Kolmogorov-Smirnov, Mann-Whitney U και chi-square δοκιμές** σε διάστημα εμπιστοσύνης της τάξης του 95% στο πρόγραμμα SPSS v16.0. Μελετήθηκαν ενενήντα ένα άτομα εξ' αυτών που συμμετείχαν σε ατυχήματα, με μέση ηλικία τα 28,47 έτη. Η μέση δαπάνη υγειονομικής περίθαλψης για αυτούς τους 91 ασθενείς που μελετήθηκαν μεταξύ της λήψης και της απόρριψης ήταν 253,02 δολάρια ΗΠΑ (διάμεσος, US \$ 55,90, εύρος, US \$ 11,52 - 7137,19). Σύμφωνα με τη μελέτη αυτή, δεν υπήρξε οριστική συσχέτιση μεταξύ του κόστους υγειονομικής περίθαλψης και του χρόνου ατυχήματος, του τύπου της μοτοσυκλέτας, τη φύσης της επιφάνειας του οδοστρώματος, του προστατευτικού εξοπλισμού, του καιρού ή του φυσικού φωτός. Από την έρευνα προκύπτει ότι, ο κίνδυνος ατυχήματος αυξάνεται στους νέους ενήλικες. Συνεπώς, αυξάνεται και το κόστος της υγειονομικής περίθαλψης. Επομένως, είναι σημαντικό να επαναπροσδιοριστούν οι κανονισμοί σχετικά με την ηλικία και την εκπαίδευση που απαιτούνται για την παραλαβή άδειας εκμετάλλευσης μοτοσυκλέτας και, εάν είναι απαραίτητο, να ρυθμιστούν.

Το **2015** δημοσιεύθηκε η έρευνα των Palmera-Suárez R, López-Cuadrado T, Almazán-Isla J, Fernández-Cuenca R, Alcalde-Cabero E και Galán I[9] αναφορικά με **την εκτίμηση του μεγέθους και της κατανομής της πρόκλησης αναπηρίας λόγω οδικών ατυχημάτων**, σύμφωνα με τους βασικούς **κοινωνικούς, οικονομικούς και ιατρικούς παράγοντες**. Η έρευνα αυτή διεξήχθη στην Ισπανία με αντιπροσωπευτικό δείγμα 91.846 νοικοκυριών και 20.425 ατόμων με αναπηρία ηλικίας άνω των 15 ετών, εκ των οποίων οι 443 εμφάνιζαν αναπηρία λόγω οδικού ατυχήματος. Στην έρευνα αυτή χρησιμοποιήθηκαν εργαλεία στατιστική ανάλυσης όπως η **γραμμική παλινδρόμηση** (με $p < 0.05$) για την εξαγωγή των παρακάτω αποτελεσμάτων. Σύμφωνα με την έρευνα, ο κίνδυνος αναπηρίας ποσοτικοποιείται στο 2.1 ανά 1000 κατοίκους, είναι υψηλότερος μεταξύ των ατόμων ηλικίας 31 έως 64 ετών, ενώ και στα δύο φύλα εμφανίζεται έντονο σημείο καμψής στην ηλικία των 16 ετών. Παράλληλα, η αναλογία πιθανοτήτων είναι υψηλότερη στους συμμετέχοντες με δευτεροβάθμια εκπαίδευση από ότι σε όσους έχουν μόνο πρωτοβάθμια εκπαίδευση. Αντίστοιχα, οι πιθανότητες είναι αυξημένες σε όσους έχουν χαμηλότερα επίπεδα εισοδήματος και χαμηλότερες σε όσους έχουν αυξημένα επίπεδα εισοδήματος. Στο δείγμα αυτό, μόνο το 24% των συμμετεχόντων με αναπηρία εργάζονται, ενώ όπως παρατηρήθηκε σε σύγκριση με άλλες αιτίες πρόκλησης αναπηρίας, τα οδικά ατυχήματα προκαλούν μεγαλύτερο βαθμό αναπηρίας όσον αφορά την κινητικότητα, μεγαλύτερα προβλήματα σχετικά με την ανεξάρτητη μετακίνηση των πασχόντων και ιδιαίτερος εκτός σπιτιού, μεγαλύτερη ανάγκη για ιατρικές και κοινωνικές υπηρεσίες, καθώς επίσης και μεγαλύτερες αλλαγές στην οικονομική κατάσταση των παθόντων. Συμπερασματικά, η πιθανότητα αναπηρίας λόγω τροχαίων ατυχημάτων στην Ισπανία είναι χαμηλότερη από ό,τι σε άλλες ανεπτυγμένες χώρες, ενώ τα άτομα μέσης ηλικίας και κοινωνικο-οικονομικά μειονεκτούντα αποτελούν την πιο ευπαθή ομάδα. Η αναπηρία εξαιτίας οδικών ατυχημάτων σχετίζεται άμεσα με μεγαλύτερη απαίτηση για κοινωνική και υγειονομική περίθαλψη, προβλήματα πρόσβασης και μετακίνησης και σημαντικές αλλαγές στην οικονομική δραστηριότητα.

Όσον αφορά τις **μεθοδολογίες** υπάρχει πλήθος χρήσιμων ερευνών από τον τομέα της στατιστικής ανάλυσης των οδικών ατυχημάτων οι οποίες παρέχουν τα κατάλληλα εργαλεία και μεθόδους για την παρούσα Διπλωματική Εργασία.

Το **1995** δημοσιεύθηκε η έρευνα των Söderland και Zwi[10] στην οποία διεξήχθη **ανάλυση πολλαπλής παλινδρόμησης** κατά την οποία χρησιμοποιήθηκαν δεδομένα από **83 χώρες** για το έτος **1990**. Ως εξαρτημένη μεταβλητή θεωρήθηκαν εναλλάξ τα ετήσια ποσοστά θανάτων που σχετίζονται με οδική κυκλοφορία ανά εκατό χιλιάδες πληθυσμού, οι ετήσιοι θάνατοι που σχετίζονται με οδική κυκλοφορία ανά χίλια τετράτροχα οχήματα, ο λόγος της θνησιμότητας στη μέση ηλικία προς τη θνησιμότητα του συνολικού πληθυσμού, ο λόγος της ανδρικής θνησιμότητας προς τη γυναικεία και ο λόγος των θανάσιμων τραυματισμών προς το συνολικό αριθμό τραυματισμών. Εισήχθησαν επίσης διάφορες επεξηγηματικές μεταβλητές ανάλογα με το αποτέλεσμα που προέκυπτε, όπως ο κατά κεφαλήν αριθμός οχημάτων, η πυκνότητα οδικού δικτύου (χιλιόμετρα ανά τετραγωνικά χιλιόμετρα), το κατά κεφαλήν Α.Ε.Π., η δαπάνη για υγεία ως ποσοστό του Α.Ε.Π. και η πυκνότητα πληθυσμού. Συμπερασματικά, το κεφαλήν Α.Ε.Π. συσχετίζεται θετικά με τους θανάτους σε οδικά ατυχήματα ανά εκατό χιλιάδες πληθυσμού, αλλά αρνητικά με τους θανάτους ανά χίλια τετράτροχα οχήματα υποδεικνύοντας ότι σε όρους ανά όχημα το αυξημένο εισόδημα μειώνει τους νεκρούς σε οδικά ατυχήματα. Επιπροσθέτως, ο αριθμός νεκρών από οδικά ατυχήματα στις νεαρές και στις πολύ μεγάλες ηλικίες σχετίζεται άμεσα με την πυκνότητα του πληθυσμού. **Τέλος το κατά κεφαλήν Α.Ε.Π. και οι δαπάνες στον τομέα της υγείας ως ποσοστό του Α.Ε.Π. σχετίζονται με ένα μειούμενο ποσοστό θανάσιμων τραυματισμών ανάμεσα στα θύματα οδικών ατυχημάτων.**

Το **2010** δημοσιεύθηκε η έρευνα των Teik Hua Law, Robert B. Noland και Andrew W. Evans[11] αναφορικά με τη **σχέση του κατά κεφαλήν εισοδήματος και των τροχαίων ατυχημάτων** σε 60 χώρες κατά την περίοδο 1972-2004. Στη συγκεκριμένη έρευνα χρησιμοποιήθηκαν δεδομένα σε μορφή χρονοσειράς αλλά και διατμηματικά δεδομένα, με αποτέλεσμα να επιλεγεί η **αρνητική διωνυμική κατανομή** ως καταλληλότερη για την δημιουργία του μοντέλου. Με τη χρήση αυτής της κατανομής εξασφαλίστηκε η απόκλιση των παραμέτρων διασποράς και η αιτιολόγηση της ετερογένειας των δεδομένων. Για την επεξεργασία των δεδομένων χρησιμοποιήθηκε το **LR τεστ** της μορφής $LR = 2(LLP LLNB)$, μέσω του οποίου εξακριβώθηκε ότι η μηδενική υπόθεση (όταν η διακύμανση ισούται με τη μέση τιμή) δύναται να απορριφθεί όταν η σημαντικότητα των αποτελεσμάτων είναι στο 0.01, γεγονός που συμφωνεί με την αρνητική διωνυμική κατανομή. Τα αποτελέσματα που προέκυψαν είναι ότι **ο αριθμός των θανάτων από τροχαία ατυχήματα αυξάνεται με την αύξηση της κινητικότητας στα αρχικά στάδια της οικονομικής ανάπτυξης. Στη συνέχεια όμως, λόγω των προόδων των τεχνικών, πολιτικών και πολιτικών θεσμών, ο αριθμός των θανάτων μειώνεται με την αύξηση του κατά κεφαλήν εισοδήματος.** Η ποιότητα των πολιτικών θεσμών καθώς και οι βελτιώσεις στην ιατρική περίθαλψη και την τεχνολογία φαίνεται να έχουν αντίκτυπο στους θανάτους από τροχαία ατυχήματα. Τελικά, τα αποτελέσματα δείχνουν σημάδια μιας καμπύλης Kuznet ως τη σχέση μεταξύ του κατά κεφαλήν εισοδήματος και των θανάτων από τροχαία ατυχήματα τόσο για τις πολύ ανεπτυγμένες όσο και για τις λιγότερο ανεπτυγμένες χώρες και υποστηρίζει την υπόθεσή ότι οι αλλαγές στην θεσμική ποιότητα και τις ιατρικές βελτιώσεις αποτελούν τη βάση της σχέσης Kuznet. Τα στοιχεία που παρουσιάζονται σε αυτή τη μελέτη υποδηλώνουν ότι η μείωση των επιπέδων διαφθοράς καθώς και οι βελτιώσεις στην

ιατρική περίθαλψη και την τεχνολογία δύναται να βοηθήσουν στη μείωση των θανάτων από τροχαία ατυχήματα

Όσον αφορά τις μεθόδους επεξεργασίας, η συνάρτηση πυκνότητας πιθανότητας της αρνητικής διωνυμικής κατανομής για κάθε χώρα i έτος t εκφράστηκε ως εξής:

$$p(Y_{it}) = \frac{\Gamma(a_i^{-1}\lambda_{it} + Y_{it})}{\Gamma(a_i^{-1}\lambda_{it})Y_{it}!} \left(\frac{1}{1+a_i^{-1}}\right)^{Y_{it}} \left(\frac{a_i^{-1}}{1+a_i^{-1}}\right)^{a_i^{-1}\lambda_{it}} \quad \text{Εξίσωση 2.1}$$

όπου $\Gamma()$ η συνάρτηση Γάμμα και a_i ο ρυθμός υπερβολικής διασποράς.

Η μέση τιμή και η διακύμανση της αρνητικής διωνυμικής κατανομής είναι:

$$E(Y_{it}) = \lambda_{it}$$

$$\text{Var}(Y_{it}) = \lambda_{it}(1+a_i)$$

Σύμφωνα με τις παραπάνω σχέσεις προκύπτει ότι υπάρχει υψηλός συσχετισμός μεταξύ του εισοδήματος και των οχημάτων κατά κεφαλήν (0,91), των ηλικιακών μεταβλητών και των οχημάτων ανά κάτοικο (0,8), και τέλος των μεταβλητών που εκφράζουν ηλικία και των ιατρικών μεταβλητών (0,81-0,84).

Το **2016** δημοσιεύθηκε η έρευνα των Constantinos Antoniou, George Yannis, Eleonora Papadimitriou και Sylvain Lassarre[12] σε συνεργασία με το Εθνικό Μετσόβιο Πολυτεχνείο της Αθήνας και το GRETIA-IFSTAR της Γαλλίας, αναφορικά με τη διερεύνηση της **στατιστικής σχέσης του ΑΕΠ και της ανεργίας, με ορισμένους δείκτες θνησιμότητας στα τροχαία ατυχήματα**. Στην συγκεκριμένη έρευνα αξιοποιήθηκε η μέθοδος των χρονολογικών σειρών για περίοδο 38 ετών (1975-2012) για τις 30 χώρες της Ευρωπαϊκής Ένωσης, η οποία βασίστηκε στην εκτίμηση μακροπρόθεσμων μοντέλων, τα οποία διαμορφώθηκαν χωριστά για κάθε χώρα. Τέλος με τη χρήση μεθόδων όπως ο **εκτιμητής μέσου όρου κοινών αποτελεσμάτων (Pesaran)** εξήχθησαν τα ζητούμενα αποτελέσματα, σύμφωνα με τα οποία, η σχέση μεταξύ οδικής ασφάλειας, η οποία υπολογίζεται μέσω του αριθμού των θανάτων στα τροχαία ατυχήματα, και της οικονομικής ανάπτυξης, η οποία μετράται μέσω του ΑΕΠ, για το σύνολο των τριάντα χωρών είναι μια σχέση αρκετά σύνθετη. Σύμφωνα με την συγκεκριμένη αναφορά, πολλές φορές παραβλέπεται η διαφορετική φύση των δύο αυτών χρονοσειρών και υιοθετείται μια σιωπηρή συμφωνία ώστε η ελαστικότητα στο ΑΕΠ να θεωρείται ομοιογενής. Μια άλλη ευρέως γνωστή λύση είναι είτε η βραχυπρόθεσμη μελέτη έναντι της μακροπρόθεσμης, όπου το φαινόμενο αυτό δεν είναι τόσο έντονο, είτε ο περιορισμός της μελέτης του εκάστοτε φαινομένου σε μεμονωμένες χώρες.

Όσον αφορά τις μεθόδους επεξεργασίας, αναφέρεται στην έρευνα πως για δεδομένα μίας μόνο χώρας και μιας αρκετά μεγάλης χρονοσειράς, μπορούν να αξιοποιηθούν μονόπλευρα «univariate» μοντέλα χρονοσειρών, όπως **αυτορρυθμιζόμενη παλινδρόμηση με γραμμική τάση, ARIMAX** και διμερή δομικά μοντέλα. Αντίστοιχα, σε περιπτώσεις όπου τα δεδομένα περιλαμβάνουν μεγάλο αριθμό χωρών (όπως στην προς μελέτη περίπτωση) αλλά μόνο ένας μικρός αριθμός χρονικών παρατηρήσεων είναι διαθέσιμος, τότε η **ανάλυση διατομής «cross-sectional» και μικρού πλαισίου**, μπορεί να πραγματοποιηθεί με τη βοήθεια αρνητικής διωνυμικής κατανομής ή

κατανομής Poisson. Τέλος σε περιπτώσεις όπου τα στοιχεία επικαλύπτονται μεταξύ των χωρών και ο αριθμός των παρατηρήσεων είναι επαρκής, μπορεί να εφαρμοστεί μοντέλο μακροχρόνιας ανάλυσης με παλινδρόμηση, συνενώνοντας τις χρονοσειρές. Συνεπώς στην παρούσα έρευνα εφαρμόστηκε μια ειδική μορφή ενός γενικού δυναμικού μοντέλου, το λεγόμενο αυτορρυθμιζόμενο μοντέλο καθυστερημένης κατανομής, όπως φαίνεται παρακάτω:

$$\log\text{FAT}_i = a_i + \beta_0 \log\text{GDP}_{it} + v_{it} \quad \text{Εξίσωση 2.2}$$

Το 2017 δημοσιεύθηκε η έρευνα των Chinmoy Pala, Shigeru Hirayamaa, Sangolla Naraharib, Manoharan Jeyabharathb, Gopinath Prakashb και Vimalathithan Kulothunganb[13] αναφορικά με **τα οδικά ατυχήματα και το συσχετισμό τους με το Α.Ε.Π.** Για την έρευνα αυτή αξιοποιήθηκαν δεδομένα από την βάση δεδομένων του Παγκόσμιου Οργανισμού Υγείας και για την επεξεργασία τους χρησιμοποιήθηκε «cluster analysis». Σύμφωνα με την έρευνα, **τα οδικά ατυχήματα αυξάνονται κάθε χρόνο στις χώρες με χαμηλό και μεσαίο εισόδημα** σε σύγκριση με τις **χώρες υψηλού εισοδήματος, οι οποίες παρουσιάζουν μειούμενες τάσεις**. Για την επεξεργασία των δεδομένων χρησιμοποιήθηκε μια **τεχνική αυτόματης χαρτογράφησης (SOM)** για τη χαρτογράφηση των μη γραμμικών σχέσεων μεταξύ διαφορετικών χαρακτηριστικών. Συνολικά, 176 χώρες με 44 ιδιότητες εξετάστηκαν από τη βάση δεδομένων «Global Status Report on Road Traffic Crashes» της Παγκόσμιας Οργάνωσης Υγείας (WHO). Τελικά, παρατηρήθηκαν πολύ διαφορετικά και μοναδικά πρότυπα αιτίου-αποτελέσματος για τρεις ομάδες αποτελεσμάτων SOM. Οι χώρες υψηλού εισοδήματος βρέθηκαν να έχουν χαμηλότερο συνολικό αριθμό θανάτων ανά 100.000 κατοίκους. Ο αριθμός των θανάτων που αφορούν τα οχήματα τεσσάρων τροχών ήταν υψηλότερος σε ορισμένες χώρες του Περσικού Κόλπου, των Ηνωμένων Πολιτειών, της Αυστραλίας και της Νέας Ζηλανδίας. Παράλληλα, χώρες με έλλειψη νόμων για την προστασία των ποδηλάτων είχαν υψηλότερα ποσοστά θανάτου μεταξύ των ποδηλάτων. Το ποσοστό των θανάτων από ποδηλάτες ήταν επίσης υψηλότερο σε περιοχές χωρίς απαίτηση κράνους και χωρίς επενδύσεις για την βελτίωση της υποδομής, ενώ το ποσοστό των θανάτων των πεζών ήταν υψηλό όταν δεν υπήρχε πολιτική για τον διαχωρισμό των χρηστών των οδών, ιδίως στις αφρικανικές χώρες χαμηλού εισοδήματος.

2.3 Σύνοψη-Κριτική Αξιολόγηση

Στο κεφάλαιο αυτό παρουσιάστηκαν τα αποτελέσματα της βιβλιογραφικής ανασκόπησης που έγινε για τις ανάγκες της παρούσας Διπλωματικής Εργασίας. Πραγματοποιήθηκε ανάλυση των ερευνών που έχουν διεξαχθεί σε χώρες της Ε.Ε. και παγκοσμίως με σκοπό τη διερεύνηση της επιρροής των ιατρικών, οικονομικών και κοινωνικών παραγόντων στην οδική ασφάλεια μιας χώρας ή περιοχής.

Παρατηρείται ότι υπάρχει μεγάλη σχέση μεταξύ του κλάδου υγείας και των οδικών ατυχημάτων, τόσο ως προς την πιθανότητα θνησιμότητας μετά από ένα ατύχημα, όσο και ως προς τα έξοδα, κοινωνικά και οικονομικά, εξαιτίας της πρόκλησης ενός ατυχήματος.

Όσον αφορά τους υπόλοιπους παράγοντες οι οποίοι σχετίζονται με τα οδικά ατυχήματα, παρατηρείται μεγάλη σχέση μεταξύ του ΑΕΠ μιας χώρας και του βαθμού θνησιμότητας ως προς

τον πληθυσμό. Από τις παραπάνω έρευνες συμπεραίνεται ότι όσο πιο ανεπτυγμένη είναι μια χώρα συγκριτικά με τις άλλες, τόσο μικρότερος είναι ο αριθμός των νεκρών εξαιτίας οδικών ατυχημάτων. Σε αυτό το σημείο οφείλεται να τονιστεί και η ύπαρξη της καμπύλης Kuznet έτσι όπως προκύπτει από τις παραπάνω έρευνες, σύμφωνα με την οποία ο αριθμός των θανάτων από τροχαία ατυχήματα αυξάνεται με την αύξηση της κινητικότητας στα αρχικά στάδια της οικονομικής ανάπτυξης, ενώ στη συνέχεια, λόγω των προόδων των τεχνικών, πολιτικών και πολιτικών θεσμών, ο αριθμός των θανάτων μειώνεται με την αύξηση του κατά κεφαλήν εισοδήματος.

Παράλληλα, πολύ έντονη είναι και η προσπάθεια που γίνεται στην Ισπανία, σύμφωνα με τις παραπάνω έρευνες σχετικά με την μείωση των οδικών ατυχημάτων και των επιπτώσεών τους. Όπως φαίνεται, η επένδυση τόσο στο οδικό δίκτυο όσο και στην επιβολή νέας πολιτικής οδικής ασφάλειας μπορεί να οδηγήσει σε μεγάλη οφέλη την κοινωνία τόσο από άποψη χαμμένης παραγωγικότητας και γενικότερα οικονομικών δεικτών όσο και ανθρώπινων απωλειών.

Τέλος, ο συσχετισμός των οδικών ατυχημάτων με τους νεότερους σε ηλικία ενήλικες καθιστά την μελέτη του φαινομένου αυτού υψίστης σημασίας, τόσο για την κοινωνία και την ανάπτυξη της όσο και για το κάθε άτομο ξεχωριστά.

Όσον αφορά στα μαθηματικά μοντέλα, χρησιμοποιήθηκαν τόσο απλά μοντέλα παλινδρόμησης, γραμμικά και μη γραμμικά, όσο και στατιστικές μέθοδοι ανάλυσης χρονοσειρών, όπως η ARIMA.

Ως συνέπεια των παραπάνω προέκυψε η ανάγκη για τη διερεύνηση της επιρροής και των ιατρικών παραγόντων, πέραν των οικονομικών και κοινωνικών δεικτών, όπως το κατά κεφαλήν Α.Ε.Π., η ανεργία και ο δείκτης ανθρώπινης ανάπτυξης, στην οδική ασφάλεια στις χώρες της Ε.Ε. Επιλέχθηκε να γίνει χρήση λογαριθμισμένων γραμμικών μοντέλων παλινδρόμησης με εξαρτημένη μεταβλητή τον αριθμό των νεκρών σε οδικά ατυχήματα ως προς τον πληθυσμού για κάθε μία χώρα της Ε.Ε. και ως ανεξάρτητες μεταβλητές κατάλληλους ιατρικούς, οικονομικούς και κοινωνικούς δείκτες. Θεωρήθηκε σκόπιμο να μελετηθούν και να αναπτυχθούν μοντέλα για τις χώρες της Ε.Ε. τόσο στο σύνολο τους όσο και σε ομάδες χωρών οι οποίες θα μοιράζονται κατά το δυνατόν κοινά πληθυσμιακά, οικονομικά, κοινωνικά και συγκοινωνιακά στοιχεία.

3. Θεωρητικό Υπόβαθρο

3.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζεται το θεωρητικό υπόβαθρο πάνω στο οποίο βασίστηκε η στατιστική ανάλυση των στοιχείων της παρούσας Διπλωματικής Εργασίας. Αρχικά παρατίθεται μια σύντομη περιγραφή βασικών εννοιών της στατιστικής και στη συνέχεια πραγματοποιείται μια λεπτομερής ανάλυση της μεθοδολογίας που χρησιμοποιήθηκε για να περιγράψει τη συσχέτιση των επιδόσεων της οδικής ασφάλειας με οικονομικούς και κοινωνικούς δείκτες στις χώρες της Ευρωπαϊκής Ένωσης. Το κεφάλαιο κλείνει παρουσιάζοντας τους απαραίτητους στατιστικούς ελέγχους και τα κριτήρια αποδοχής του μοντέλου που προέκυψε από την εφαρμογή της επιλεγείσας μεθοδολογίας.

3.2 Βασικές Έννοιες Στατιστικής

Ο όρος **πληθυσμός** (population) αναφέρεται στο σύνολο των παρατηρήσεων του χαρακτηριστικού που ενδιαφέρει τη στατιστική έρευνα. Πρόκειται για ένα σύνολο στοιχείων που είναι τελείως καθορισμένα. Ένας πληθυσμός μπορεί να είναι πραγματικός ή θεωρητικός.

Ο όρος **δείγμα** (sample) αναφέρεται σε ένα υποσύνολο του πληθυσμού. Οι περισσότερες στατιστικές έρευνες στηρίζονται σε δείγματα, αφού οι ιδιότητες του συνόλου του πληθυσμού είναι συνήθως αδύνατο να καταγραφούν. Όλα τα στοιχεία που ανήκουν στο δείγμα ανήκουν και στον πληθυσμό χωρίς ωστόσο να ισχύει το αντίστροφο. Τα συμπεράσματα που θα προκύψουν από τη μελέτη του δείγματος θα ισχύσουν με ικανοποιητική ακρίβεια για ολόκληρο τον πληθυσμό μόνο εάν το δείγμα είναι αντιπροσωπευτικό του πληθυσμού.

Με τον όρο **μεταβλητές** (variables) εννοούνται τα χαρακτηριστικά που ενδιαφέρουν να μετρηθούν και να καταγραφούν σε ένα σύνολο στοιχείων. Οι μεταβλητές διακρίνονται στις παρακάτω κατηγορίες :

- **Ποιοτικές μεταβλητές** (qualitative variables) των οποίων οι τιμές δε δίδονται με αριθμούς αλλά με διακριτικό είδος, παραδείγματος χάριν το επάγγελμα, η υπηκοότητα κλπ.
- **Ποσοτικές μεταβλητές** (quantitative variables) των οποίων οι τιμές είναι αριθμοί. Οι ποσοτικές μεταβλητές διακρίνονται με τη σειρά τους σε δύο μεγάλες κατηγορίες, τις **διακριτές** και τις **συνεχείς**. Σε μια διακριτή μεταβλητή η μικρότερη μη μηδενική διαφορά που μπορούν να έχουν δύο τιμές είναι σταθερή ποσότητα, ενώ αντιθέτως σε μία συνεχή μεταβλητή δύο τιμές μπορούν να διαφέρουν κατά οποιαδήποτε μικρή ποσότητα. Στην πράξη συνεχής θεωρείται μία μεταβλητή όταν μπορεί να λάβει όλες τις τιμές σε ένα διάστημα, διαφορετικά θεωρείται διακριτή.

Μέτρα κεντρικής τάσης (measures of central tendency) είναι αριθμοί που μας δείχνουν γύρω από ποια τιμή βρίσκονται τα δεδομένα μας. Τα κυριότερα μέτρα κεντρικής τάσης είναι :

- Ο **μέσος όρος** (mean) δηλαδή το άθροισμα των τιμών των δεδομένων μας διαιρεμένο με το πλήθος τους.
- Η **διάμεσος** (median) δηλαδή η μεσαία τιμή όταν τα δεδομένα μας έχουν διαταχθεί σε αύξουσα σειρά
- Η **επικρατούσα τιμή** (mode) δηλαδή η τιμή που εμφανίζεται στα δεδομένα μας με τη μεγαλύτερη συχνότητα

Μέτρα διασποράς (measures of variability) είναι αριθμοί που μας δείχνουν την απόσταση των δεδομένων μας από το κέντρο τους. Τα κυριότερα μέτρα διασποράς είναι :

- Το **εύρος** (range) δηλαδή η διαφορά της μικρότερης από τη μεγαλύτερη τιμή των δεδομένων μας
- Η **διακύμανση** (variance) δηλαδή μία εκτίμηση του μέσου όρου των τετραγώνων των αποστάσεων των δεδομένων μας από τον μέσο όρο τους
- Η β (standard deviation) δηλαδή η τετραγωνική ρίζα της διακύμανσης

Μέτρα αξιοπιστίας :

- **Επίπεδο εμπιστοσύνης** (confidence level) δηλαδή η αναλογία των περιπτώσεων που μια εκτίμηση θα είναι σωστή
- **Επίπεδο σημαντικότητας** (significance level) δηλαδή η αναλογία των περιπτώσεων που ένα συμπέρασμα είναι εσφαλμένο.

3.3 Συσχέτιση Μεταβλητών – Συντελεστής Συσχέτισης

Στη συνέχεια θεωρούνται δύο τυχαίες και συνεχείς μεταβλητές X , Y . Ο βαθμός της γραμμικής συσχέτισης των δύο αυτών μεταβλητών X και Y με διασπορά σ_x και σ_y αντίστοιχα και συνδιασπορά $\sigma_{xy} = \text{Cov} [X,Y]$ καθορίζεται με τον συντελεστή συσχέτισης (correlation coefficient) ρ , που ορίζεται ως εξής:

$$\rho = \left(\frac{\sigma_{xy}}{\sigma_x} \right) \left(\frac{1}{\sigma_y} \right) \quad \text{Εξίσωση 3.1}$$

Ο συντελεστής συσχέτισης ρ εκφράζει τον βαθμό και τον τρόπο που οι δύο μεταβλητές συσχετίζονται. Δεν εξαρτάται από την μονάδα μέτρησης των X και Y και παίρνει τιμές στο διάστημα $[-1,1]$. Τιμές κοντά στο 1 δηλώνουν ισχυρή θετική συσχέτιση, τιμές κοντά στο -1 δηλώνουν ισχυρή αρνητική συσχέτιση και τιμές κοντά στο 0 δηλώνουν γραμμική ανεξαρτησία των X και Y .

Η εκτίμηση του συντελεστή συσχέτισης ρ γίνεται με την αντικατάσταση στην ανωτέρω εξίσωση της συνδιασποράς σ_{xy} και των διασπορών σ_x , σ_y , από όπου προκύπτει τελικά η έκφραση της εκτιμήτριας r :

$$r(X, Y) = \frac{(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}))}{\left[(\sum_{i=1}^n (x_i - \bar{x})^2)^{\frac{1}{2}} (\sum_{i=1}^n (y_i - \bar{y})^2)^{\frac{1}{2}} \right]} \quad \text{Εξίσωση 3.2}$$

3.4 Κανονική Κατανομή

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

$$F(x) = \left(\frac{1}{(2\pi\sigma)^{1/2}} \right) e^{[-(x-\mu)^2/2\sigma^2]} \quad \text{Εξίσωση 3.3}$$

όπου μ και σ είναι σταθερές ίσες με τη μέση τιμή και την τυπική απόκλιση αντίστοιχα.

3.5 Μαθηματικά Μοντέλα

3.5.1 Γραμμική Παλινδρόμηση

Ο κλάδος της στατιστικής ο οποίος εξετάζει τη σχέση μεταξύ δύο ή περισσότερων μεταβλητών ώστε να είναι δυνατή η πρόβλεψη της μιας από τις υπόλοιπες ονομάζεται **ανάλυση παλινδρόμησης** (regression analysis). Με τον όρο **εξαρτημένη μεταβλητή** εννοείται η μεταβλητή της οποίας η τιμή πρόκειται να προβλεφθεί, ενώ με τον όρο **ανεξάρτητη** γίνεται αναφορά σε εκείνη τη μεταβλητή η οποία χρησιμοποιείται για την πρόβλεψη της εξαρτημένης μεταβλητής. Η ανεξάρτητη μεταβλητή δεν θεωρείται τυχαία, αλλά παίρνει καθορισμένες τιμές. Η εξαρτημένη μεταβλητή θεωρείται τυχαία και επηρεάζεται από την ανεξάρτητη. Προκειμένου να προσδιοριστεί αν μια ανεξάρτητη μεταβλητή ή συνδυασμός ανεξάρτητων μεταβλητών προκάλεσε τη μεταβολή της εξαρτημένης μεταβλητής, κρίνεται απαραίτητη η ανάπτυξη μαθηματικών μοντέλων.

Η ανάπτυξη ενός μαθηματικού μοντέλου αποτελεί μια στατιστική διαδικασία που συμβάλλει στην παραγωγή εξισώσεων που περιγράφουν τη σχέση μεταξύ των ανεξάρτητων μεταβλητών και της εξαρτημένης. Επισημαίνεται ότι η επιλογή της μεθόδου ανάπτυξης ενός μοντέλου βασίζεται στο αν η εξαρτημένη μεταβλητή είναι συνεχής ή διακριτό μέγεθος.

Στην περίπτωση που η εξαρτημένη μεταβλητή είναι συνεχές μέγεθος και ακολουθεί την κανονική κατανομή, μία από τις πλέον διαδεδομένες στατιστικές τεχνικές είναι η γραμμική παλινδρόμηση. Η απλούστερη περίπτωση γραμμικής παλινδρόμησης είναι η **απλή γραμμική παλινδρόμηση** (simple linear regression).

Στην απλή γραμμική παλινδρόμηση υπάρχει μόνο μία ανεξάρτητη μεταβλητή x και μία εξαρτημένη μεταβλητή y , που προσεγγίζεται ως μια γραμμική συνάρτηση του x . Η τιμή y_i της y , για κάθε τιμή x_i της x , δίνεται από τη σχέση :

$$y_i = \alpha + \beta * x_i + \epsilon_i \quad \text{Εξίσωση 3.4}$$

Το πρόβλημα της γραμμικής παλινδρόμησης είναι η εύρεση των παραμέτρων α και β που εκφράζουν καλύτερα τη γραμμική εξάρτηση της y από τη x . Κάθε ζεύγος τιμών (α, β) καθορίζει μία διαφορετική γραμμική σχέση που εκφράζεται γεωμετρικά από ευθεία γραμμή και οι δύο παράμετροι ορίζονται ως εξής :

- Ο σταθερός όρος α είναι η τιμή του y για $x=0$
- Ο συντελεστής β του x είναι η κλίση (slope) της ευθείας ή αλλιώς ο **συντελεστής παλινδρόμησης** (regression coefficient). Εκφράζει τη μεταβολή της μεταβλητής y όταν η μεταβλητή x μεταβληθεί κατά μία μονάδα.

Ο όρος ϵ_i ονομάζεται 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 . Συνεπώς είναι απαραίτητο να συμπεριληφθεί στο μοντέλο ο όρος του σφάλματος ϵ_i . Αυτό γίνεται τόσο για να αντιπροσωπευθούν στο μοντέλο τυχόν μεταβλητές που έχουν παραλειφθεί, όσο και για να ληφθεί υπόψη κάθε σφάλμα προσέγγισης που σχετίζεται με τη γραμμική συναρτησιακή μορφή. Το ϵ_i μπορεί συχνά να αναφέρεται και ως σφάλμα, απόκλιση, υπόλοιπο κλπ.

Στην περίπτωση που η τυχαία μεταβλητή y εξαρτάται γραμμικά από περισσότερες από μία μεταβλητές x γίνεται αναφορά στην **πολλαπλή γραμμική παλινδρόμηση** (multiple linear regression). Η εξίσωση που περιγράφει τη σχέση μεταξύ εξαρτημένης και ανεξαρτήτων μεταβλητών είναι η εξής :

$$y_i = \beta_0 + \beta_1 * X_{1i} + \beta_2 * X_{2i} + \beta_3 * X_{3i} + \dots + \beta_k * X_{ki} + \epsilon_i \quad \text{Εξίσωση 3.5}$$

Η εκτίμηση της πολλαπλής γραμμικής παλινδρόμησης δεν διαφέρει ουσιαστικά από το μοντέλο της απλής γραμμικής παλινδρόμησης. Ωστόσο ένα καινούριο στοιχείο στην πολλαπλή γραμμική παλινδρόμηση είναι ότι πριν προχωρήσει κανείς στην εκτίμηση των παραμέτρων πρέπει να ελέγξει εάν πράγματι πρέπει να συμπεριληφθούν όλες οι ανεξάρτητες μεταβλητές στο μοντέλο. Εκείνο που απαιτείται να εξασφαλιστεί είναι η μηδενική συσχέτιση των ανεξάρτητων μεταβλητών ($\rho(x_i, x_j) \rightarrow 0$, για κάθε $i \neq j$).

Στη γραμμική παλινδρόμηση οι παράμετροι εκτιμώνται με τη μέθοδο των ελαχίστων τετραγώνων, δηλαδή οι συντελεστές υπολογίζονται έτσι ώστε το άθροισμα των τετραγώνων των διαφορών των παρατηρούμενων και των υπολογιζόμενων να είναι το ελάχιστο.

Προκειμένου το μοντέλο να μπορεί να προσεγγίσει την επιρροή των ανεξάρτητων μεταβλητών στην εξαρτημένη με όσο το δυνατόν πιο ορθό και αξιόπιστο τρόπο, θα πρέπει να ελέγχεται ότι πληρούνται κάθε φορά οι παρακάτω τέσσερις υποθέσεις:

1. Η υπόθεση της **γραμμικότητας**, που δηλώνει ότι η σχέση μεταξύ δύο μεταβλητών X και Y είναι κατά προσέγγιση γραμμική.

2. Η υπόθεση της **ανεξαρτησίας**, που δηλώνει ότι τα υπόλοιπα (σφάλματα, αποκλίσεις) για διαφορετικές παρατηρήσεις πρέπει να είναι ανεξάρτητα μεταξύ τους.
3. Η υπόθεση της **κανονικότητας**, που δηλώνει ότι η απόκλιση πρέπει να είναι (προσεγγιστικά) κανονικά κατανομημένη.
4. 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 Λογαριθμισμένη Γραμμική Παλινδρόμηση

Μέσω της λογαριθμισμένης κανονικής παλινδρόμησης (lognormal regression) δίνεται η δυνατότητα ανάπτυξης ενός μοντέλου που συσχετίζει δύο ή περισσότερες μεταβλητές. Η σχέση που συνδέει την εξαρτημένη με τις ανεξάρτητες μεταβλητές είναι γραμμική. Στη λογαριθμισμένη κανονική παλινδρόμηση οι συντελεστές των μεταβλητών του μοντέλου είναι οι συντελεστές της γραμμικής παλινδρόμησης. Υπολογίζονται από την ανάλυση της παλινδρόμησης με βάση την αρχή των ελαχίστων τετραγώνων.

Η λογαριθμισμένη κανονική 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 η εξής:

$$\log(y_i) = \beta_0 + \beta_1 * X_{1i} + \beta_2 * X_{2i} + \beta_3 * X_{3i} + \dots + \beta_k * X_{ki} + \varepsilon_i \quad \text{Εξίσωση 3.6}$$

όπου y είναι η εξαρτημένη μεταβλητή, $\beta_0, \beta_1, \dots, \beta_k$ οι συντελεστές μερικής παλινδρόμησης, $X_{1i}, X_{2i}, \dots, X_{ki}$ οι εξαρτημένες μεταβλητές και ε_i το σφάλμα παλινδρόμησης.

3.5.3 Γενικό Γραμμικό Μοντέλο

Το γενικό γραμμικό μοντέλο (General Linear Model) μπορεί να θεωρηθεί ως επέκταση της γραμμικής πολλαπλής παλινδρόμησης για μία 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:

$$x_{ij} = g_{i1} * \beta_{1j} + g_{i2} * \beta_{2j} + \dots + g_{ik} * \beta_{kj} + \varepsilon_{ij} \quad \text{Εξίσωση 3.7}$$

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

Το γενικό γραμμικό μοντέλο βασίζεται στην υπόθεση ότι τα σφάλματα (ϵ_{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:

- μία **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 μετρείται η εξαρτημένη μεταβλητή.

Από μαθηματική σκοπιά δεν υπάρχει διάκριση μεταξύ των δύο τύπων μεταβλητών. Η παραπάνω εξίσωση μπορεί να γραφτεί σε μορφή πίνακα ως ένα πολυμεταβλητό γενικό γραμμικό μοντέλο:

$$X = G * \beta + \epsilon \quad \text{Εξίσωση 3.8}$$

όπου X είναι ένας πίνακας δεδομένων, ο οποίος έχει στοιχεία x_{ij} σε κάθε στήλη για κάθε παράγοντα j και σε κάθε σειρά για κάθε παρατήρηση i . Ο πίνακας G αποτελείται από τους συντελεστές g_{ik} και ονομάζεται στη διεθνή ορολογία design matrix, ενώ $\beta = [x_1, x_2, \dots, x_j]$ είναι πίνακας παραμέτρων, όπου x_j είναι ένα διάνυσμα στήλη με παραμέτρους για τους παράγοντες j . Επιπλέον, ϵ είναι ένας πίνακας με κανονικά κατανεμημένους όρους σφαλμάτων.

Η παραπάνω εξίσωση δεν περιλαμβάνει σταθερό όρο, καθώς μπορεί να απομακρυνθεί με δύο τρόπους:

- με μέση διόρθωση του πίνακα δεδομένων
- προσθέτοντας μία στήλη με άσους στον πίνακα B

Σε αυτήν την περίπτωση και εφόσον τα σφάλματα είναι κανονικά κατανεμημένα, οι υπολογισμοί των ελαχίστων τετραγώνων αποτελούν υπολογισμούς μέγιστης πιθανότητας και χαρακτηρίζονται και αυτοί από κανονική κατανομή. Ειδικά, χρησιμοποιείται η μέθοδος **ανάλυσης διασποράς** (analysis of variance - ANOVA).

3.6 Αυτοσυσχέτιση

Μία από τις υποθέσεις των παραπάνω μαθηματικών μοντέλων είναι η υπόθεση της ανεξαρτησίας των υπολοίπων. Με άλλα λόγια, τα υπόλοιπα (σφάλματα) για διαφορετικές παρατηρήσεις πρέπει να είναι ανεξάρτητα μεταξύ τους. Η υπόθεση αυτή ουσιαστικά σημαίνει ότι οι διάφορες τιμές του σφάλματος ϵ δεν συσχετίζονται. Δηλαδή το σφάλμα της περιόδου t δε συσχετίζεται με το σφάλμα μιας οποιασδήποτε άλλης περιόδου s . Εάν αυτή η υπόθεση δεν ικανοποιείται, τότε έχουμε το φαινόμενο της αυτοσυσχέτισης (autocorrelation) ή αυτοπαλινδρόμησης (autoregression). Η αυτοσυσχέτιση είναι συνηθισμένο φαινόμενο όταν χρησιμοποιούνται στοιχεία χρονοσειρών.

Το σφάλμα ε της εξίσωσης, ουσιαστικά παριστάνει την επίδραση όλων των παραγόντων που δεν μπορούν να περιληφθούν στην εξεταζόμενη σχέση. Συχνά όμως, η επίδραση πολλών από αυτούς τους παράγοντες μπορεί να μην εξαντλείται στην τρέχουσα περίοδο, αλλά να διαχέεται και σε μελλοντικές περιόδους. Στην περίπτωση αυτή, οι διαδοχικές τιμές του σφάλματος θα συσχετίζονται. Η αυτοσυσχέτιση μπορεί επίσης να οφείλεται στην παράλειψη ερμηνευτικών (ανεξάρτητων) μεταβλητών, στην εσφαλμένη εξειδίκευση της μαθηματικής μορφής του υποδείγματος καθώς και σε πολλούς άλλους λόγους.

Η σχέση εξάρτησης, εάν υπάρχει, ανάμεσα στις διαδοχικές τιμές του σφάλματος μπορεί να πάρει διάφορες μορφές. Εάν η τιμή του σφάλματος στην περίοδο t εξαρτάται από την τιμή του στην περίοδο $t - 1$, δηλαδή:

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t \quad \text{Εξίσωση 3.9}$$

όπου u_t μια τυχαία ερμηνευτική μεταβλητή και ρ μια παράμετρος, τότε έχουμε αυτοσυσχέτιση πρώτης τάξεως ή πρώτου βαθμού (first-order autocorrelation) ή αυτοπαλίνδρομο σχήμα πρώτου βαθμού (first-order autoregressive scheme), που συμβολίζεται ως AR(1). Ο συντελεστής ρ ονομάζεται συντελεστής αυτοσυσχέτισης πρώτης τάξεως.

Εάν η τιμή του σφάλματος στην περίοδο t εξαρτάται όχι μόνο από την τιμή του στην περίοδο $t - 1$ αλλά και από την τιμή του στην περίοδο $t - 2$, δηλαδή,

$$\varepsilon_t = \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + u_t \quad \text{Εξίσωση 3.10}$$

τότε έχουμε αυτοσυσχέτιση δεύτερης τάξης ή AR(2) κ.ο.κ.

Οι εκτιμητές (συντελεστές) που προκύπτουν από τη μέθοδο των ελαχίστων τετραγώνων όταν το υπόδειγμα χαρακτηρίζεται από αυτοσυσχέτιση εξακολουθούν να είναι γραμμικοί, αμερόληπτοι και συνεπείς. Το πρόβλημα που δημιουργείται αναφέρεται κυρίως στις εκτιμήσεις των διακυμάνσεων τους και την αποτελεσματικότητά τους. Οι διακυμάνσεις είναι μεροληπτικές και οι εκτιμητές δεν είναι αποτελεσματικοί. **Συνεπώς δεν μπορούν να χρησιμοποιηθούν οι συντελεστές που προκύπτουν από τη μέθοδο των ελαχίστων τετραγώνων.**

Ένα **μικτό μοντέλο** είναι ένα στατιστικό μοντέλο το οποίο περιέχει σταθερούς και τυχαίους παράγοντες. Αυτά τα μοντέλα είναι χρήσιμα στην επαναλαμβανόμενη μέτρηση όπου λαμβάνονται τιμές κάτω από τις ίδιες συνθήκες με μία μόνο μεταβλητή y το χρόνο να μεταβάλλεται. Λόγω των πλεονεκτημάτων τους όσον αφορά την αντιμετώπιση των ελλειπουσών τιμών, τα μοντέλα μικτών αποτελεσμάτων προτιμώνται συχνά έναντι των πιο παραδοσιακών προσεγγίσεων, όπως τα επαναλαμβανόμενα μέτρα ANOVA. Η εξίσωση που περιγράφει τη σχέση μεταξύ εξαρτημένης και ανεξαρτητών μεταβλητών είναι η εξής:

$$y = \beta * X + Z*u + \varepsilon \quad \text{Εξίσωση 3.11}$$

το y είναι ένας γνωστός πίνακας με τις τιμές της εξαρτημένης μεταβλητής και μέση τιμή $E(y)=X\beta$ το β είναι ένα άγνωστο διάνυσμα με σταθερούς συντελεστές το u είναι ένα άγνωστο διάνυσμα με τυχαίους συντελεστές με μέση τιμή $E(u)=0$ και διακύμανση πίνακα $\text{var}(u)=G$

ο ε είναι ένα άγνωστο διάνυσμα με τυχαίους συντελεστές με μέση τιμή $E(\varepsilon)=0$ και διακύμανση πίνακα $\text{var}(u)=R$ και X και Z είναι γνωστοί πίνακες συσχέτισης της μεταβλητής y με τις μεταβλητές β και u , αντίστοιχα.

3.6.1 Βασικές Αρχές Παλινδρομήσεων

Οι γραμμικές παλινδρομήσεις στηρίζονται στη **μέθοδο των ελαχίστων τετραγώνων** και έχουν ως βασικό κριτήριο αξιοπιστίας την τιμή του **συντελεστή συσχέτισης R^2** . Ο συντελεστής αυτός εκφράζει το πόσο αξιόπιστο είναι το πρότυπο που δημιουργήθηκε με την παλινδρόμηση και το ποσοστό της μεταβολής της εξαρτημένης μεταβλητής που μπορεί να εξηγηθεί από τη σχέση που προκύπτει.

Στην περίπτωση της απλής παλινδρόμησης ο συντελεστής αυτός εκφράζει το βαθμό της προσαρμογής της ευθείας παλινδρόμησης στις τιμές του δείγματος. Εάν πρόκειται για πολλαπλή παλινδρόμηση τότε ο ίδιος συντελεστής εκφράζει το βαθμό προσαρμογής της επιφάνειας παλινδρόμησης στα δειγματικά δεδομένα. Ο συντελεστής R^2 υπολογίζεται από το πηλίκο της παλινδρομικής μεταβολής προς τη συνολική μεταβολή. Παλινδρομική μεταβολή είναι το άθροισμα των αποκλίσεων των τιμών που προέκυψαν από την παλινδρόμηση από το μέσο όρο του δείγματος. Ουσιαστικά πρόκειται για την επεξηγούμενη απόκλιση από το μοντέλο αφού μπορούν να εξηγηθούν με βάση την ευθεία παλινδρόμησης και τείνουν να ακολουθήσουν μία ορισμένη συμπεριφορά. Η συνολική μεταβολή είναι το άθροισμα των τετραγώνων των αποκλίσεων των τιμών του δείγματος από το μέσο όρο του δείγματος. Η διαφορά της παλινδρομικής μεταβολής από τη συνολική μεταβολή καλείται υπόλοιπη μεταβολή και ισούται με το άθροισμα των τετραγώνων των αποκλίσεων των τιμών του δείγματος από τις αντίστοιχες τιμές που δίνει η ευθεία παλινδρόμησης.

Το πεδίο τιμών της ποσότητας R^2 είναι από 0 μέχρι και 1. Η τιμή 0 αντιστοιχεί στην περίπτωση που δεν υπάρχει καμία γραμμική συσχέτιση, ενώ όταν η τιμή φτάσει τη μονάδα, τότε υπάρχει απόλυτη γραμμική συσχέτιση μεταξύ των μεταβλητών. Ο συντελεστής R^2 διαφοροποιείται σε περιπτώσεις όπου η απόκλιση μιας τιμής που προέκυψε από την αντίστοιχη τιμή του δείγματος είναι σημαντική. Σφάλμα απόκλισης καλείται η διαφορά της πραγματικής από την εκτιμώμενη τιμή. Πρόκειται δηλαδή για σημεία τα οποία είναι απομακρυσμένα από την ευθεία παλινδρόμησης σε σχέση με τα υπόλοιπα που είναι ομοιόμορφα διασπαρμένα γύρω από την ευθεία. Τέτοια σημεία, τα οποία καλούνται outliers, χρήζουν προσεκτικής μελέτης ούτως ώστε να δικαιολογηθεί η παρουσία τους.

Βασική προϋπόθεση της μεθόδου είναι πως οι ανεξάρτητες μεταβλητές σε μία εξίσωση δεν πρέπει να σχετίζονται μεταξύ τους, δηλαδή να είναι μεταξύ τους **γραμμικώς ανεξάρτητες**. Η μη πλήρωση αυτής της συνθήκης είναι μία από τις συνηθέστερες πηγές σφαλμάτων, αφού τότε είναι δύσκολο να εξακριβωθεί η επίδραση αποκλειστικά κάθε μεταβλητής στο τελικό αποτέλεσμα.

Στην περίπτωση που είναι συνέπεια σφαλμάτων κατά τις μετρήσεις των στοιχείων, την κωδικοποίηση ή τον υπολογισμό των μεταβλητών, τότε απλά γίνεται η κατάλληλη διόρθωση και υπολογίζεται η νέα εξίσωση παλινδρόμησης. Εάν μία τέτοια επεξήγηση δεν είναι εφικτή, θα

πρέπει να εξεταστεί το ενδεχόμενο να μην έχει ληφθεί υπόψη η επιρροή παραγόντων η οποία είναι καθοριστική στην εξαρτημένη μεταβλητή στα σημεία εκείνα.

Είναι σημαντικό να δίνεται η κατάλληλη προσοχή τόσο στον εντοπισμό όσο και στην επεξεργασία των σημείων αυτών, καθώς τείνουν να έλκουν την ευθεία παλινδρόμησης και συνεπώς να οδηγούν σε διαφορετικά αποτελέσματα[14].

3.6.2 Κριτήρια Αποδοχής Σχέσεων

Η κάθε σχέση παλινδρόμησης που προκύπτει πρέπει να ελέγχεται κατά πόσο παρέχει αξιόπιστα αποτελέσματα και λογικά συμπεράσματα. Ο έλεγχος αποδοχής ή όχι μιας εξίσωσης παλινδρόμησης πραγματοποιείται με τη βοήθεια των ακόλουθων κριτηρίων:

- **Της τιμής του συντελεστή συσχέτισης R^2 .** Όπως έχει ήδη αναφερθεί, η εξαρτημένη και οι ανεξάρτητες μεταβλητές βρίσκονται σε ισχυρή συσχέτιση μόνο για υψηλές τιμές του R^2 . Συνεπώς εξίσωση με **χαμηλή τιμή** του συντελεστή συσχέτισης οδηγεί σε **απόρριψη της σχέσης** ως προσεγγιστικής και αναξιόπιστης.
- **Των προσήμων στους μερικούς συντελεστές παλινδρόμησης.** Η ορθή εισαγωγή της κάθε μεταβλητής στην εξίσωση ελέγχεται από το πρόσημο με το οποίο εμφανίζεται σε αυτή, ώστε να παρέχει συμπεράσματα που να μην αντικρούονται με τις επικρατούσες αντιλήψεις και λογική.

Επιπροσθέτως σημαντικό ρόλο όσον αφορά την καταλληλότητα του μαθηματικού προτύπου στο οποίο καταλήγει η μέθοδος της γραμμικής παλινδρόμησης διαδραματίζει ο στατιστικός δείκτης t-test, ο οποίος παρέχει πληροφορίες σχετικά με το αν πρέπει μια μεταβλητή να συμπεριληφθεί στο πρότυπο ή όχι. Ο δείκτης t-test αναφέρεται σε καθεμία από τις μεταβλητές ξεχωριστά και είναι το αποτέλεσμα της διαίρεσης της εκτιμώμενης για το συντελεστή τιμής με την τυπική απόκλιση της. Η τυπική απόκλιση είναι ένα μέγεθος που παρουσιάζει τη συνέπεια με την οποία έχει υπολογιστεί η τιμή του συγκεκριμένου συντελεστή. Αυτό σημαίνει ότι η πραγματική τιμή του συντελεστή για επίπεδο εμπιστοσύνης 95% βρίσκεται στο διάστημα με κέντρο την υπολογιζόμενη τιμή του συντελεστή και άκρα την τιμή αυτή συν ή πλην την τυπική απόκλιση.

Ο δείκτης t-test στην ουσία δείχνει αν η πραγματική τιμή του συγκεκριμένου συντελεστή διαφέρει σημαντικά από το μηδέν ή όχι. Αν η επιρροή αυτή είναι σημαντική τότε η συγκεκριμένη μεταβλητή πρέπει να συμπεριληφθεί στην ανάπτυξη του μαθηματικού προτύπου. Σε αντίθετη περίπτωση πρέπει να αποκλειστεί. Οι τιμές που μπορεί να λάβει κυμαίνονται από μείον άπειρο έως συν άπειρο.

Όσο μεγαλύτερη είναι η απόλυτη τιμή του τόσο μεγαλύτερη είναι η επιρροή της συγκεκριμένης μεταβλητής στο τελικό αποτέλεσμα. Ανάλογα με το επίπεδο σημαντικότητας στο οποίο μας ενδιαφέρει να βρίσκονται τα αποτελέσματα της έρευνας υπάρχουν πίνακες που δίνουν την τιμή του t-test πάνω από την οποία η συγκεκριμένη μεταβλητή πρέπει να συμπεριληφθεί στο μαθηματικό πρότυπο. Προκύπτει ότι για διάστημα εμπιστοσύνης 95% μία

μεταβλητή μπορεί να παραμείνει στο πρότυπο αν η απόλυτη τιμή του δείκτη t του συντελεστή της είναι **μεγαλύτερη από 1,645**.

Αντίστοιχα, για το γραμμικό μικτό μοντέλο ισχύουν τα παραπάνω κριτήρια μαζί με κάποια νέα εργαλεία-κριτήρια, τα οποία προσφέρει η μέθοδος του μικρού μοντέλου.

Για τη στατιστική εμπιστοσύνη του μικτού μοντέλου χρησιμοποιείται η μέθοδος της μεγιστοποίησης της πιθανοφάνειας. Για να επιτευχθεί υψηλή πιθανοφάνεια πρέπει ο λογάριθμος των συναρτήσεων πιθανοφάνειας $L = -2 \text{ Restricted Log Likelihood}$ να είναι όσο το δυνατόν μικρότερος και γενικά προτιμώνται τα μοντέλα με το μικρότερο λογάριθμο συνάρτησης πιθανοφάνειας L . Σύμφωνα με το κριτήριο του λόγου πιθανοφάνειας, υπολογίζεται η διαφορά $LRT = - (L(b) - L(0))$, όπου $L(b) = L$ (μοντέλου με p μεταβλητές) και $L(0) = L$ (μοντέλου χωρίς τις p μεταβλητές) και εάν είναι μεγαλύτερη από την τιμή του κριτηρίου χ^2 για p βαθμούς ελευθερίας σε επίπεδο σημαντικότητας 5%, τότε το μοντέλο είναι στατιστικά προτιμότερο από το μοντέλο χωρίς τις μεταβλητές και γίνεται αποδεκτό. Πρέπει να σημειωθεί εδώ, ότι στα αποτελέσματα χρησιμοποιούνται και άλλα κριτήρια, τα οποία προκύπτουν από το παραπάνω (AIC, AICC, CAIC και BIC).

Το κριτήριο πληροφοριών Akaike (AIC) είναι ένας εκτιμητής της σχετικής ποιότητας των στατιστικών μοντέλων για ένα προσδιορισμένο σύνολο δεδομένων. Δεδομένης της συλλογής μοντέλων για τα δεδομένα, το AIC υπολογίζει την ποιότητα κάθε μοντέλου σε σχέση με κάθε άλλο μοντέλο. Έτσι, το AIC παρέχει ένα μέσο επιλογής μοντέλου. Το AIC βασίζεται στη θεωρία των πληροφοριών, προσφέρει μια εκτίμηση των σχετικών πληροφοριών που χάνονται, όταν ένα δεδομένο μοντέλο χρησιμοποιείται για να αντιπροσωπεύει τη διαδικασία που δημιούργησε τα δεδομένα. Έτσι, ασχολείται με το αντιστάθμισμα μεταξύ της καλής προσαρμογής του μοντέλου και της απλότητας του μοντέλου. Το AIC δεν παρέχει δοκιμή ενός μοντέλου υπό την έννοια της μηδενικής υπόθεσης. Το κριτήριο αυτό πληροφορεί για την ποιότητα ενός μοντέλου σε σχέση με άλλα μοντέλα, αλλά δεν πληροφορεί για την απόλυτη ποιότητά του. Έτσι, εάν όλα τα υποψήφια μοντέλα δεν ταιριάζουν καλά, η AIC δε θα δώσει καμία προειδοποίηση γι' αυτό.

Δοθέντος ενός στατιστικού μοντέλου μερικών δεδομένων, έστω k ο αριθμός των εκτιμώμενων παραμέτρων στο μοντέλο και \hat{L} η μέγιστη τιμή της συνάρτησης πιθανότητας για το μοντέλο. Στη συνέχεια, η τιμή AIC του μοντέλου είναι η ακόλουθη. [1] [2]

$$AIC = 2k - 2\ln(\hat{L}) \quad \text{Εξίσωση 5.1}$$

Δεδομένης μιας σειράς υποψηφίων μοντέλων για τα δεδομένα, το προτιμώμενο μοντέλο είναι εκείνο με την ελάχιστη τιμή AIC. Έτσι, το AIC ανταμείβει την τάση προσαρμογής (όπως αξιολογείται από τη συνάρτηση πιθανότητας), αλλά περιλαμβάνει, επίσης, μια ποινή που είναι μια αυξανόμενη συνάρτηση του αριθμού των εκτιμώμενων παραμέτρων. Η αύξηση του αριθμού παραμέτρων στο μοντέλο σχεδόν πάντα βελτιώνει την καλοσύνη της τοποθέτησης. Το φαινόμενο αυτό ονομάζεται υπερφόρτωση και η ύπαρξη ποινής εξασφαλίζει την αποφυγή αυτού του φαινομένου.

Στην παρούσα Διπλωματική Εργασία επιλέγεται να εξετάζεται όχι μόνο το **log Likelihood**, αλλά και το **AIC**.

Οι **έλεγχοι των σταθερών επιδράσεων** (tests of fixed effects) γίνονται με τα F-tests, όπου η τιμή σημαντικότητας των συντελεστών αυτών θα πρέπει να είναι $\text{sig.} < 0,5$, ενώ ο έλεγχος των συντελεστών των μεταβλητών των σταθερών επιδράσεων γίνεται με το t-test. Προκειμένου η μεταβλητή να γίνει αποδεκτή θα πρέπει κι εδώ να ισχύει $\text{sig.} < 0,5$.

Με βάση λοιπόν τα παραπάνω ο ερευνητής είναι πλέον σε θέση να αξιολογήσει τα αποτελέσματα της έρευνάς του και να ελέγξει την αξιοπιστία τους. Αν συμβαίνει αυτό, τότε η στατιστική ανάλυση έχει ολοκληρωθεί. Σε αντίθετη περίπτωση πρέπει να συνεχιστεί μέχρις ότου οδηγήσει σε αποτελέσματα που να περνούν με επιτυχία όλους τους στατιστικούς ελέγχους που προαναφέρθηκαν

4. Συλλογή και Επεξεργασία Στοιχείων

4.1 Εισαγωγή

Στο κεφάλαιο αυτό παρουσιάζεται η διαδικασία συλλογής και επεξεργασίας των στοιχείων της Διπλωματικής Εργασίας, έτσι ώστε να δοθεί μία πλήρης εικόνα για την ποιότητα και την αξιοπιστία τους. Επιπλέον γίνεται αναφορά στα προβλήματα που προέκυψαν κατά τη συλλογή τους καθώς και στους τρόπους με τους οποίους αντιμετωπίστηκαν. Στο υποκεφάλαιο της επεξεργασίας των στοιχείων παρουσιάζεται η μέθοδος κωδικοποίησης και εισαγωγής τους στον υπολογιστή, όπως επίσης και η αρχική επεξεργασία που υπέστησαν στο πρόγραμμα Microsoft® Excel 2016 προκειμένου να μπορούν να αξιοποιηθούν από το εργαλείο στατιστικής ανάλυσης Rstudio©

4.2 Συλλογή Στοιχείων

Για την επίτευξη του στόχου της Διπλωματικής Εργασίας απαιτήθηκε η συλλογή στοιχείων που αφορούσαν τα 27 κράτη-μέλη της Ευρωπαϊκής Ένωσης. Στα στοιχεία αυτά περιλαμβάνονταν οι πιο πρόσφατες διαθέσιμες πληροφορίες σχετικά με τον πληθυσμό, ο ετήσιος αριθμός νεκρών σε οδικά ατυχήματα, το πλήθος των οδικών ατυχημάτων εντός αστικού δικτύου, η προσδοκώμενη διάρκεια ζωής των ανθρώπων, τα παραγωγικά χρόνια ζωής των ανθρώπων, το κατά κεφαλήν Α.Ε.Π., ο δείκτης ανθρώπινης ανάπτυξης, οι ετήσιες δαπάνες για τον τομέα της υγείας, οι δημόσιες δαπάνες για τον τομέα της υγείας και ο αριθμός των κλινών στα νοσοκομεία. Η διαδικασία της συλλογής δεν ήταν εφικτή με τη χρήση μόνο μίας βάσης δεδομένων, συνεπώς αναζητήθηκαν περισσότερες από μία βάσεις διεθνών οργανισμών στις οποίες γίνεται λεπτομερής αναφορά στη συνέχεια.

4.2.1 Βάσεις Δεδομένων

Το μεγαλύτερο μέρος των στοιχείων προήλθε έπειτα από αναζήτηση σε βάσεις δεδομένων της EUROSTAT, της στατιστικής υπηρεσίας της Ευρωπαϊκής Επιτροπής. Στις βάσεις αυτής της υπηρεσίας είναι δυνατή η εύρεση στοιχείων σχετικά με διάφορους τομείς δραστηριοτήτων, όπως επιστημονικούς, κοινωνικούς, ιατρικούς, οικονομικούς, μεταφορών και αναπτυξιακούς ενώ τα στοιχεία αυτά καλύπτουν ένα μεγάλο χρονικό εύρος.

Συγκεκριμένα για τις ανάγκες της παρούσας Διπλωματικής Εργασίας αντλήθηκαν από τη βάση της EUROSTAT για τα έτη 2008-2014 στοιχεία σχετικά με τον πληθυσμό των κρατών της Ε.Ε., τα παραγωγικά χρόνια ζωής των ανθρώπων, τη διάρκεια ζωής, τα ποσοστά ανεργίας, τον αριθμό κλινών των νοσοκομείων, τις ετήσιες δαπάνες για τον τομέα της υγείας και τις δημόσιες δαπάνες για τον τομέα της υγείας.

Στην παρούσα Διπλωματική Εργασία αξιοποιήθηκε άλλη μια βάση δεδομένων η CARE, η οποία είναι μια κοινοτική βάση δεδομένων η οποία δημιουργήθηκε από την Ευρωπαϊκή Επιτροπή και

περιλαμβάνει δεδομένα σχετικά με τα οδικά ατυχήματα που έχουν ως αποτέλεσμα θάνατο ή τραυματισμό (δεν υπάρχουν στατιστικά στοιχεία για ατυχήματα που προκαλούν ζημιές μόνο). Η μεγάλη διαφορά μεταξύ της CARE και των περισσότερων άλλων υφιστάμενων διεθνών βάσεων δεδομένων είναι το υψηλό επίπεδο διαχωρισμού, δηλαδή η CARE περιλαμβάνει λεπτομερή στοιχεία για ατομικά ατυχήματα που συλλέγονται από τα κράτη μέλη. Αυτή η δομή επιτρέπει τη μέγιστη ευελιξία και το δυναμικό όσον αφορά την ανάλυση των πληροφοριών που περιέχονται στο σύστημα και ανοίγει ένα σύνολο νέων δυνατοτήτων στον τομέα της ανάλυσης ατυχημάτων. Σκοπός του συστήματος CARE είναι να παράσχει ένα ισχυρό εργαλείο που θα επιτρέψει τον εντοπισμό και τον ποσοτικό προσδιορισμό των προβλημάτων οδικής ασφάλειας σε όλους τους ευρωπαϊκούς δρόμους, την αξιολόγηση της αποτελεσματικότητας των μέτρων οδικής ασφάλειας, τον προσδιορισμό της συνάφειας των κοινοτικών δράσεων και τη διευκόλυνση της ανταλλαγής εμπειριών στο συγκεκριμένο πεδίο. Από την συγκεκριμένη βάση αντλήθηκαν δεδομένα σχετικά με τον ετήσιο αριθμό νεκρών σε οδικά ατυχήματα και το πλήθος των οδικών ατυχημάτων εντός αστικού δικτύου.

Μία άλλη βάση δεδομένων η οποία αξιοποιήθηκε είναι αυτή του Προγράμματος Ανάπτυξης των Ηνωμένων Εθνών (United Nations Development Programme, UNDP). Ο οργανισμός αυτός αποτελεί ένα παγκόσμιο αναπτυξιακό δίκτυο του Ο.Η.Ε. που στοχεύει στην ανάπτυξη μέσω της σύνδεσης των χωρών με τη γνώση, την εμπειρία και τις κατάλληλες πηγές που θα βοηθήσουν τους ανθρώπους των χωρών αυτών να εξασφαλίσουν μια καλύτερη ζωή. Το κύριο αντικείμενο του δικτύου είναι να συντελέσει στην εξεύρεση λύσεων σε προκλήσεις, όπως η δημοκρατική διακυβέρνηση, η μείωση της φτώχειας, η καταπολέμηση του AIDS, η ενέργεια και το περιβάλλον, η προστασία των ανθρωπίνων δικαιωμάτων και η ενίσχυση της θέσης των γυναικών. Στη βάση δεδομένων αυτού του οργανισμού αναζητήθηκαν τα πιο πρόσφατα διαθέσιμα στοιχεία σχετικά με το δείκτη ανθρώπινης ανάπτυξης (Human Development Index, HDI) για τις χώρες τις Ε.Ε. τα οποία αφορούσαν τα έτη 2008-2014. Ο συγκεκριμένος στατιστικός δείκτης χρησιμοποιείται για την κατάταξη των χωρών με βάση την "ανθρώπινη ανάπτυξη". Αποτελεί ένα σύνθετο μέτρο που κατασκευάζεται με βάση τρεις επί μέρους δείκτες, οι οποίοι σχετίζονται με το προσδόκιμο ζωής, τον βαθμό εκπαίδευσης και την ποιότητα ζωής. Βάσει του HDI γίνεται χαρακτηρισμός μιας χώρας σε υπανάπτυκτη, αναπτυσσόμενη ή αναπτυγμένη. Χρησιμοποιείται επίσης για την μέτρηση της επίδρασης των οικονομικών πολιτικών στην ποιότητα ζωής.

4.2.2 Προβλήματα που παρουσιάστηκαν κατά τη συλλογή των στοιχείων και επίλυσή τους

Λόγω της ύπαρξης διαφορετικών συστημάτων συλλογής και επεξεργασίας των στοιχείων οδικών ατυχημάτων ανά χώρα, είναι δυνατό να παρουσιαστούν προβλήματα, που είναι πολύ πιθανό να οδηγήσουν στη διατύπωση εσφαλμένων συμπερασμάτων έπειτα από τη χρήση τους. Ένα τέτοιο πρόβλημα οφείλεται στους διαφορετικούς ορισμούς που έχουν τα κράτη για τα διάφορα επιμέρους στοιχεία. Στην παρούσα Διπλωματική Εργασία αυτό το πρόβλημα αντιμετωπίστηκε με τη χρήση στοιχείων από διεθνείς βάσεις δεδομένων, όπου τα στοιχεία τα οποία προέρχονται από την κάθε χώρα έχουν υποστεί την κατάλληλη επεξεργασία, ώστε να γίνουν ομοιόμορφα και συγκρίσιμα. Τα μόνα προβλήματα τα οποία παρουσιάστηκαν κατά τη συλλογή των δεδομένων ήταν η μη αναφορά όλων των απαιτούμενων στοιχείων για τα ίδια έτη (το ίδιο χρονικό πλαίσιο) και η

έλλειψη πληροφορίας αναφορικά με τον τομέα της υγείας. Πιο συγκεκριμένα, παρατηρήθηκε έλλειψη δεδομένων στους περισσότερους από τους παράγοντες οι οποίοι επιλέχθηκαν να μελετηθούν και να αναλυθούν για τις 27 χώρες-μέλη της Ε.Ε.. Εξαιτίας αυτού του φαινομένου το πλαίσιο παρατήρησης και έρευνας περιορίστηκε στα έτη 2008-2014 για τα οποία υπάρχουν περισσότερα δεδομένα, ενώ σε μερικές περιπτώσεις στις οποίες απουσίαζε μια μόνο τιμή εντός του προαναφερθέντος χρονικού πλαισίου χρησιμοποιήθηκε γραμμική παλινδρόμηση για την συμπλήρωση της τιμής σύμφωνα με τις προηγούμενες τιμές και στο τέλος η τιμή αυτή ελέγχθηκε ώστε να βρίσκεται εντός των αναμενόμενων τιμών για τον εκάστοτε δείκτη . Παράλληλα, σημειώνεται μεγάλη έλλειψη δεδομένων σχετικά με τα οικονομικά στοιχεία στον κλάδο της υγείας, όπως οι δαπάνες και οι αμοιβές του προσωπικού. Αυτό το ζήτημα επιλύθηκε με την αξιοποίηση του συνολικού ποσού δαπανών για τον κλάδο της υγείας τόσο ως μέρος των δημόσιων δαπανών όσο και αναλογικά με τον πληθυσμό των διαφόρων χωρών.

4.3 Επεξεργασία Στοιχείων

Για την επεξεργασία των δεδομένων απαιτήθηκε σε πρώτη φάση η συγκέντρωση των απαραίτητων στοιχείων μέσω των προαναφερθέντων βάσεων δεδομένων. Τα στοιχεία τα οποία συλλέχθηκαν από αυτές τις βάσεις βρίσκονταν τις περισσότερες φορές ήδη σε μορφή πινάκων, ενώ τις λίγες φορές όπου τα δεδομένα χρειάστηκε να συμπληρωθούν χειροκίνητα σε πίνακες εντός του εργαλείου excel αυτό και έγινε. Ένα παράδειγμα τέτοιου πίνακα παρατίθεται παρακάτω για την καλύτερη και ορθότερη κατανόηση.

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
EU	43,15 0	39,57 7	35,35 9	31,50 6	30,68 7	28,24 4	25,95 5	25,97 7	25,96 6	22,07 0
AT	691	679	633	552	523	531	455	430	479	432
BE	1,071	944	944	840	862	770	723	727	732	637
BG	1,006	1,061	901	776	657	601	601	661	708	-
CY	89	82	71	60	71	51	44	45	57	46
CZ	1,221	1,076	901	802	773	742	654	688	734	611
DE	4,949	4,477	4,152	3,648	4,009	3,600	3,339	3,377	3,459	3,206
DK	406	406	303	255	220	167	191	182	178	211
EE	196	132	98	79	101	87	81	78	67	-
EL	1,612	1,553	1,456	1,258	1,141	988	879	795	793	824
ES	3,822	3,098	2,714	2,479	2,060	1,902	1,680	1,688	1,689	1,810
FI	380	344	279	272	292	255	258	229	266	-
FR	4,620	4,275	4,273	3,992	3,963	3,653	3,268	3,384	3,459	3,471
HR	619	664	548	426	418	393	368	308	348	307
HU	1,232	996	822	740	638	605	591	626	644	607
IE	338	280	238	212	186	162	188	193	-	-
IT	5,131	4,725	4,237	4,114	3,860	3,753	3,401	3,381	3,428	3,283
LT	740	499	370	299	296	302	256	267	242	-
LU	46	35	48	32	33	34	45	35	36	32

LV	419	316	254	218	179	177	179	212	188	158
MT	12	9	15	13	16	9	17	10	11	-
NL	709	677	644	537	546	562	476	476	531	533
PL	5,583	5,437	4,572	3,908	4,189	3,571	3,357	3,202	2,938	3,026
PT	974	885	840	937	891	718	637	638	593	563
RO	2,800	3,065	2,796	2,377	2,018	2,042	1,861	1,818	1,893	1,913
SE	471	397	358	266	319	285	260	270	259	270
SI	293	214	171	138	141	130	125	108	120	130
SK	661	606	384	371	325	352	251	295	310	-
UK	3,059	2,645	2,337	1,905	1,960	1,802	1,770	1,854	1,804	-

Πίνακας 4.1 Ανθρώπινες απώλειες για τα έτη 2008-2014 στις 28 χώρες της Ε.Ε.[15]

Στη συνέχεια και αφού συγκεντρώθηκαν όλα τα απαραίτητα δεδομένα στην παραπάνω μορφή, ξεκίνησε η δημιουργία μιας κοινής βάσης δεδομένων με τη χρήση του προγράμματος Microsoft®Excel 2016. Η βάση αυτή περιείχε για κάθε κράτος τα αντίστοιχα στοιχεία, δηλαδή αριθμό νεκρών σε οδικά ατυχήματα, πληθυσμό, κατά κεφαλήν Α.Ε.Π, καθώς επίσης και άλλους ιατρικούς, οικονομικούς, κοινωνικούς και συγκοινωνιακούς δείκτες. Στη συνέχεια παρατίθεται ο σχετικός πίνακας, ο οποίος παρουσιάζει την τελική μορφή των διαφόρων στοιχείων εντός της βάσης δεδομένων.

Στον παρακάτω πίνακα γίνεται αντιληπτός ο τρόπος με τον οποίο δημιουργήθηκε η βάση δεδομένων. Στην αρχική στήλη τοποθετήθηκαν τα 27 κράτη της Ε.Ε., στην δεύτερη στήλη οι χρονολογίες και στις επόμενες στήλες βρίσκονται οι διάφορες κατηγορίες δεικτών. Πιο συγκεκριμένα, ο κωδικός κάθε χώρας αναγράφεται στην πρώτη στήλη τόσες φορές όσες είναι και οι διαφορετικές χρονολογίες για τις οποίες υπάρχουν δεδομένα προς μελέτη. Έτσι, για κάθε χώρα και κάθε έτος υπάρχουν τα απαιτούμενα δεδομένα όπως ο πληθυσμός, το ΑΕΠ και τα ποσοστά ανεργίας.

NAME	YEAR	POPULATION	GDP	FATALITIES	URBAN ACCIDENTS	HDI	WORKING LIFE	PRODUCTIVITY	LIFE EXPECTANCY	UNEMPLOYMENT	HOSPITAL BEDS	PUBLIC HEALTH EXPENDITURE	HEALTH EXPENDITURE PER CAPITA	HEALTH EXPENDITURE %GDP
AT	2008	8307989	51629.646	679	189	0.87	35.7	117.5	80.6	172	64018	75.4046557	5286.915229	10.60393
AT	2009	8335003	47785.563	633	173	0.87	35.9	117.8	80.5	223	64069	75.628005	5153.788544	11.1944373
AT	2010	8351643	46757.13	552	141	0.88	36	115.4	80.7	203	64008	75.2334998	5049.56234	11.1652482
AT	2011	8375164	51192.565	523	139	0.88	36.2	115.7	81.1	194	64417	75.1096872	5431.67401	10.9355731
AT	2012	8408121	48381.444	531	151	0.89	36.5	117.7	81.1	209	64691	75.3673343	5239.517579	11.1682198
AT	2013	8451860	50533.288	455	115	0.89	36.7	117	81.3	231	64825	75.15506	5478.24495	11.1377909
AT	2014	8507786	51390.001	430	123	0.89	36.6	116.5	81.6	245	64815	77.8639644	5580.493943	11.2054733
BE	2008	10666866	48850.72	944	274	0.88	32.1	128.2	79.8	333	70410	77.0621569	4565.948425	9.59842645
BE	2009	10753080	45176.012	944	257	0.88	32.1	129.6	80.2	380	70265	77.3865609	4574.69481	10.3943401
BE	2010	10839905	44691.258	840	246	0.88	32.5	131.3	80.3	406	70170	77.6803252	4419.444255	10.1656407
BE	2011	11000638	47950.947	862	281	0.89	32.1	129.7	80.7	347	70138	76.8143208	4864.154446	10.4234956
BE	2012	11075889	44900.088	770	213	0.89	32.2	130.7	80.5	369	70032	77.6707129	4587.91449	10.53975
BE	2013	11137974	46613.195	723	179	0.89	32.5	130.2	80.7	417	69940	77.8094635	4813.421123	10.572903
BE	2014	11180840	47634.682	727	188	0.9	32.6	130.6	81.4	423	69924	77.8686497	4884.066261	10.5947509
BG	2008	7518002	7152.785	1,061	447	0.77	32.5	41.3	73.3	202	49507	58.5386082	483.3947231	6.61156197
BG	2009	7467119	6859.679	901	312	0.77	32.1	41.3	73.7	240	50041	55.3517635	471.574667	6.77819389
BG	2010	7421766	6743.737	776	312	0.78	31.6	41.3	73.8	352	48934	55.686333	488.2120682	7.24456053
BG	2011	7369431	7836.578	657	235	0.78	31.1	42.2	74.2	376	47391	54.6225225	532.5046849	6.88297734
BG	2012	7327224	7401.944	601	233	0.78	31.6	43.5	74.4	410	48308	56.2902987	519.4822426	7.10569778
BG	2013	7284552	7703.481	601	227	0.79	32	42.9	74.9	436	49522	51.9892531	606.0252134	7.93021922
BG	2014	7245677	7877.094	661	251	0.79	32	43.8	74.5	385	51505	54.5739524	661.8465273	8.44291326
HR	2008	4311967	15889.111	664	414	0.8	31.7	70.5	76	166	24282	84.8686542	1265.142379	7.7045206
HR	2009	4309796	14142.275	548	316	0.8	31.8	68.5	76.3	180	23806	85.6897671	1193.487039	8.17701771
HR	2010	4302847	13505.029	426	265	0.81	31.6	67.2	76.7	224	24831	85.6582204	1143.851235	8.24785313
HR	2011	4289857	14537.955	418	252	0.82	31.4	70.2	77.2	256	25030	80.2069061	1131.667419	7.79909419
HR	2012	4275984	13234.667	393	230	0.82	31.2	72.4	77.3	292	25129	81.6086049	1029.521053	7.79722563
HR	2013	4262140	13573.74	368	213	0.82	31.1	73.4	77.8	320	24933	81.7936167	1060.833882	7.82827445
HR	2014	4246809	13468.729	308	191	0.82	32.3	70.7	77.9	325	25036	81.8670982	1050.334391	7.80324176
CY	2008	776333	19006.1745	82	57	0.85	36.4	94.9	80.6	15	2977	41.3632339	2241.265392	6.89217292
CY	2009	796930	18673.5297	71	42	0.85	36.8	95.2	81	22	3040	44.736319	2185.842041	7.40220132
CY	2010	819140	19299.5149	60	42	0.85	36.9	91.5	81.5	26	2958	47.3503428	2038.202786	7.2339653
CY	2011	839751	19730.975	71	40	0.85	36.6	89.9	81.2	34	2988	46.5369026	2235.669386	7.53731961
CY	2012	862011	19489.7159	51	31	0.85	36.3	88.7	81.1	52	2990	45.8618069	1964.815338	7.43887767
CY	2013	865878	18140.4881	44	30	0.85	36.3	86.8	82.5	69	2938	46.5336698	1889.867889	7.45866684
CY	2014	858000	17605.5676	45	34	0.85	36.8	85	82.3	70	2912	45.2252289	1819.112996	7.36761162
CZ	2008	10343422	22819.082	1,076	444	0.86	33.7	77.7	77.3	230	74924	83.7389819	1486.182785	6.81889577
CZ	2009	10425783	19787.766	901	329	0.86	33.9	79.2	77.4	352	74607	83.759706	1483.943674	7.84539515

Πίνακας 4.2 Βάση δεδομένων στο excel No.1

Από τον παραπάνω πίνακα καθίσταται ευκολότερη, καθότι γίνεται με αυτόματο τρόπο εντός του εργαλείου excel, η διαμόρφωση των pivot tables οι οποίοι παρουσιάζουν πιο αποδοτικά και συνοπτικά τις τιμές ενός μόνο παράγοντα για όλες τις χώρες και για όλα τα έτη. Αυτοί οι πίνακες έχουν στην πρώτη στήλη τον κωδικό της χώρας και στην πρώτη γραμμή τα έτη, ενώ όλοι οι εσωτερικοί συνδυασμοί περιέχουν την ζητούμενη πληροφορία ανά χώρα και ανά έτος. Ένας τέτοιος πίνακας παρατίθεται παρακάτω.

Product of POPULATION	Column Labels						
Row Labels	2008	2009	2010	2011	2012	2013	2014
AT	8307989	8335003	8351643	8375164	8408121	8451860	8507786
BE	10666866	10753080	10839905	11000638	11075889	11137974	11180840
BG	7518002	7467119	7421766	7369431	7327224	7284552	7245677
CY	776333	796930	819140	839751	862011	865878	858000
CZ	10343422	10425783	10462088	10486731	10505445	10516125	10512419
DE	82217837	82002356	81802257	80222065	80327900	80523746	80767463
DK	5475791	5511451	5534738	5560628	5580516	5602628	5627235
EE	1338440	1335740	1333290	1329660	1325217	1320174	1315819
EL	11060937	11094745	11119289	11123392	11086406	11003615	10926807
ES	45668939	46239273	46486619	46667174	46818219	46727890	46512199
FI	5300484	5326314	5351427	5375276	5401267	5426674	5451270
FR	64007193	64350226	64658856	64978721	65276983	65600350	65942267
HR	4311967	4309796	4302847	4289857	4275984	4262140	4246809
HU	10045401	10030975	10014324	9985722	9931925	9908798	9877365
IE	4457765	4521322	4549428	4570881	4589287	4609779	4637852
IT	58652875	59000586	59190143	59364690	59394207	59685227	60782668
LU	483799	493500	502066	511840	524853	537039	549680
LV	2191810	2162834	2120504	2074605	2044813	2023825	2001468
NL	16405399	16485787	16574989	16655799	16730348	16779575	16829289
PL	38115641	38135876	38022869	38062718	38063792	38062535	38017856
PT	10553339	10563014	10573479	10572721	10542398	10487289	10427301
RO	20635460	20440290	20294683	20199059	20095996	20020074	19947311
SE	9182927	9256347	9340682	9415570	9482855	9555893	9644864
SI	2010269	2032362	2046976	2050189	2055496	2058821	2061085
SK	5376064	5382401	5390410	5392446	5404322	5410836	5415949
UK	61571647	62042343	62510197	63022532	63495303	63905297	64351155

Πίνακας 4.3 Pivot table Πληθυσμού

Product of HEALTH EXPENDITURE PER CAPITA	Column Labels						
Row Labels	2008	2009	2010	2011	2012	2013	2014
AT	5287	5154	5050	5432	5240	5478	5580
BE	4566	4575	4419	4864	4588	4813	4884
BG	483	472	488	533	519	606	662
CY	2241	2186	2038	2236	1965	1890	1819
CZ	1486	1484	1410	1546	1411	1413	1379
DE	4743	4753	4700	5025	4754	5098	5411
DK	6396	6465	6267	6522	6204	6638	6463
EE	1075	1009	903	988	1078	1214	1248
EL	3094	2879	2459	2528	2049	2017	1743
ES	3072	2995	2847	2954	2651	2644	2658
FI	4287	4137	4001	4399	4255	4519	4612
FR	4828	4722	4584	4994	4699	4955	4959
HR	1265	1193	1144	1132	1030	1061	1050
HU	1146	977	1020	1099	991	1021	1037
IE	5326	4962	4238	4310	4079	4157	4239
IT	3624	3488	3384	3559	3242	3295	3258
LU	8304	8126	7964	8334	7551	7981	8138
LV	1019	817	739	827	827	869	921
NL	5462	5355	5286	5649	5456	5688	5694
PL	957	815	868	931	870	882	910
PT	2458	2404	2352	2332	2000	2059	2097
RO	541	457	483	508	468	535	557
SE	4886	4357	4694	6662	6522	7000	6808
SI	2298	2271	2083	2225	2069	2115	2161
SK	1406	1483	1378	1415	1377	1417	1455
UK	3864	3491	3491	3649	3649	3685	3935

Πίνακας 4.4 Pivot table Δαπανών υγείας ανά κάτοικο

Για την καλύτερη και ορθότερη ανάλυση και επεξεργασία των δεδομένων, τα στοιχεία αυτά διαμορφώθηκαν κατάλληλα ώστε να είναι δυνατή η δημιουργία του ζητούμενου στατιστικού μοντέλου. Αρχικά διαμορφώθηκαν κάποιες νέες κατηγορίες δεδομένων αξιοποιώντας τα αρχικά δεδομένα και είτε τον πληθυσμό κάθε χώρας ανά έτος, είτε τα οδικά ατυχήματα στα οποία σημειώθηκαν νεκροί. Έτσι προέκυψαν μεταβλητές όπως η θνησιμότητα ανά πληθυσμό (Fat.Per.Pop) και οι νοσοκομειακές κλίνες ανά πληθυσμό (Hospital.Beds.Per.Population) αλλά και το πλήθος των ατυχημάτων εντός αστικού δικτύου προς το σύνολο των θανατηφόρων ατυχημάτων (Urban.Per.Fatal). Παρακάτω παρατίθενται οι σχετικοί ατομικοί πίνακες καθώς επίσης και η ανανεωμένη βάση δεδομένων.

Product of FAT.PER.POP	Column Labels	2008	2009	2010	2011	2012	2013	2014
AT	Row Labels	8.17E-05	7.59E-05	6.61E-05	6.24E-05	6.32E-05	5.38E-05	5.05E-05
BE		8.85E-05	8.78E-05	7.75E-05	7.84E-05	6.95E-05	6.49E-05	6.50E-05
BG		1.41E-04	1.21E-04	1.05E-04	8.92E-05	8.20E-05	8.25E-05	9.12E-05
CY		1.06E-04	8.91E-05	7.32E-05	8.45E-05	5.92E-05	5.08E-05	5.24E-05
CZ		1.04E-04	8.64E-05	7.67E-05	7.37E-05	7.06E-05	6.22E-05	6.54E-05
DE		5.45E-05	5.06E-05	4.46E-05	5.00E-05	4.48E-05	4.15E-05	4.18E-05
DK		7.41E-05	5.50E-05	4.61E-05	3.96E-05	2.99E-05	3.41E-05	3.23E-05
EE		9.86E-05	7.34E-05	5.93E-05	7.60E-05	6.56E-05	6.14E-05	5.93E-05
EL		1.40E-04	1.31E-04	1.13E-04	1.03E-04	8.91E-05	7.99E-05	7.28E-05
ES		6.78E-05	5.87E-05	5.33E-05	4.41E-05	4.06E-05	3.60E-05	3.63E-05
FI		6.49E-05	5.24E-05	5.08E-05	5.43E-05	4.72E-05	4.75E-05	4.20E-05
FR		6.68E-05	6.64E-05	6.17E-05	6.10E-05	5.60E-05	4.98E-05	5.13E-05
HR		1.54E-04	1.27E-04	9.90E-05	9.74E-05	9.19E-05	8.63E-05	7.25E-05
HU		9.91E-05	8.19E-05	7.39E-05	6.39E-05	6.09E-05	5.96E-05	6.34E-05
IE		6.28E-05	5.26E-05	4.66E-05	4.07E-05	3.53E-05	4.08E-05	4.16E-05
IT		8.06E-05	7.18E-05	6.95E-05	6.50E-05	6.32E-05	5.70E-05	5.56E-05
LU		7.23E-05	9.73E-05	6.37E-05	6.45E-05	6.48E-05	8.38E-05	6.37E-05
LV		1.44E-04	1.17E-04	1.03E-04	8.63E-05	8.66E-05	8.84E-05	1.06E-04
NL		4.13E-05	3.91E-05	3.24E-05	3.28E-05	3.36E-05	2.84E-05	2.83E-05
PL		1.43E-04	1.20E-04	1.03E-04	1.10E-04	9.38E-05	8.82E-05	8.42E-05
PT		8.38E-05	7.95E-05	8.86E-05	8.43E-05	6.81E-05	6.07E-05	6.12E-05
RO		1.49E-04	1.37E-04	1.17E-04	9.99E-05	1.02E-04	9.30E-05	9.11E-05
SE		4.32E-05	3.87E-05	2.85E-05	3.39E-05	3.01E-05	2.72E-05	2.80E-05
SI		1.06E-04	8.41E-05	6.74E-05	6.88E-05	6.32E-05	6.07E-05	5.24E-05
SK		1.13E-04	7.13E-05	6.88E-05	6.03E-05	6.51E-05	4.64E-05	5.45E-05
UK		4.30E-05	3.77E-05	3.05E-05	3.11E-05	2.84E-05	2.77E-05	2.88E-05

Πίνακας 4.5 Pivot table Απωλειών λόγω ατυχημάτων προς πληθυσμό

Product of HOSPITAL BEDS PER POPULATION							
Row Labels	Column Labels						
	2008	2009	2010	2011	2012	2013	2014
AT	0.007705595	0.00768674	0.007664121	0.007691431	0.007693871	0.007669909	0.007618316
BE	0.006600814	0.006534407	0.006473304	0.006375812	0.006322924	0.006279419	0.006253913
BG	0.006585127	0.006701514	0.006593309	0.006430754	0.006592947	0.006798222	0.007108376
CY	0.003834695	0.003814639	0.003611104	0.003558198	0.003468633	0.003393088	0.003393939
CZ	0.007243638	0.007156009	0.00704888	0.00684198	0.006664544	0.00645561	0.006462547
DE	0.008202843	0.008229398	0.008245164	0.00838389	0.008346328	0.008290225	0.008250067
DK	0.003582131	0.003501074	0.003506038	0.003135078	0.003109748	0.003077306	0.002696529
EE	0.005625953	0.005364817	0.005266671	0.00535024	0.005515323	0.004997069	0.005003728
EL	0.004850584	0.004930623	0.004481222	0.004467342	0.004428847	0.004226793	0.004224473
ES	0.003217373	0.003164193	0.00312363	0.003056538	0.002990161	0.002956543	0.002964319
FI	0.006584304	0.006266998	0.005866846	0.005534786	0.005310791	0.004883102	0.004538575
FR	0.00014539	0.000142703	0.000145904	0.000142801	0.000139038	0.000139893	0.000137818
HR	0.005631305	0.005523695	0.00577083	0.005834693	0.005876776	0.005849878	0.00589525
HU	0.007099965	0.00713789	0.007171528	0.007177148	0.006992501	0.00702618	0.006976557
IE	0.004887875	0.002833906	0.002733091	0.002627065	0.002547672	0.002567802	0.002585033
IT	0.003802286	0.003699353	0.003648918	0.003518152	0.003430015	0.0033421	0.003211261
LU	0.005624237	0.005513678	0.005419606	0.00534542	0.005212888	0.005113223	0.004995634
LV	0.007713716	0.006673651	0.005621305	0.005837738	0.005854814	0.005767791	0.005635364
NL	0.004710035	0.004669477	0.004479171	0.004367002	0.004251436	0.004190213	0.004129827
PL	0.006621298	0.006655544	0.006613283	0.006628034	0.006629713	0.006604473	0.006625939
PT	0.003392576	0.003373564	0.003371265	0.00336725	0.003397235	0.003382952	0.003310732
RO	0.006848745	0.006956995	0.00663898	0.006483965	0.00658355	0.006661014	0.006698597
SE	0.002816313	0.002771396	0.002737059	0.00271529	0.002628639	0.002606245	0.002551824
SI	0.004768516	0.004619748	0.004576018	0.004630305	0.004548294	0.004554548	0.004539357
SK	0.006594416	0.00654875	0.006465185	0.006057177	0.00591434	0.005806127	0.00578809
UK	0.003345306	0.003277213	0.002940816	0.002887412	0.002816602	0.002766453	0.002740029

Πίνακας 4.6 Pivot table Πλήθους νοσοκομειακών κλινών προς πληθυσμό

Product of URBAN PER FATAL	Column Labels						
Row Labels	2008	2009	2010	2011	2012	2013	2014
AT	0.300476948	0.300347222	0.273786408	0.282520325	0.301397206	0.264367816	0.299270073
BE	0.315304948	0.290723982	0.312977099	0.351689612	0.295833333	0.264792899	0.274452555
BG	0.477054302	0.41711216	0.519702037	0.511119291	0.591933571	0.576690647	0.540365694
CY	0.721518987	0.65625	0.75	0.597014925	0.607843137	0.731707317	0.772727273
CZ	0.451678535	0.388888889	0.392183288	0.392706872	0.38573508	0.397689769	0.375601926
DE	0.306290989	0.316783036	0.298494243	0.299409237	0.314666667	0.312040882	0.30844054
DK	0.341269841	0.339483395	0.337662338	0.331730769	0.378205128	0.341040462	0.273809524
EE	0.359649123	0.228915663	0.2	0.268817204	0.203679793	0.225024568	0.305555556
EL	0.527285613	0.49845679	0.519264448	0.531874405	0.549559471	0.57002457	0.542625169
ES	0.265845596	0.276051346	0.281730978	0.251772687	0.269615272	0.302350806	0.291868386
FI	0.336448598	0.298039216	0.261410788	0.271062271	0.238297872	0.233606557	0.298076923
FR	0.314009662	0.316481294	0.305720453	0.300520976	0.303307738	0.308609272	0.315638907
HR	0.707692308	0.640973631	0.65920398	0.654545455	0.647887324	0.649390244	0.672535211
HU	0.470258137	0.400265957	0.419106317	0.415630551	0.388170055	0.42962963	0.413612565
IE	0.244094488	0.254545455	0.243243243	0.220930233	0.236842105	0.195530726	0.340782123
IT	0.474988527	0.476214448	0.460346164	0.482300885	0.455761024	0.451755773	0.474015748
LU	0.264705882	0.227272727	0.103448276	0.233333333	0.205882353	0.365853659	0.290322581
LV	0.340350877	0.299559471	0.393939394	0.323170732	0.329192547	0.313609467	0.379120879
NL	0.385714286	0.380234506	0.398	0.448076923	0.390977444	0.440789474	0.347252747
PL	0.513986014	0.523889961	0.517408676	0.518940397	0.508934073	0.516329197	0.496276236
PT	0.578696255	0.57423477	0.564760793	0.589588378	0.595202399	0.601709402	0.575456053
RO	0.723917137	0.712662338	0.709938184	0.699119912	0.679759956	0.70006035	0.706971006
SE	0.278873239	0.264880952	0.269076305	0.273972603	0.337209302	0.222672065	0.263779528
SI	0.365	0.415584416	0.472440945	0.364341085	0.344262295	0.456896552	0.412371134
SK	0.504504505	0.523809524	0.455072464	0.47608507	0.464949344	0.455640271	0.453706506
UK	0.445674457	0.462748727	0.335016835	0.34789644	0.375743163	0.332531569	0.364318707

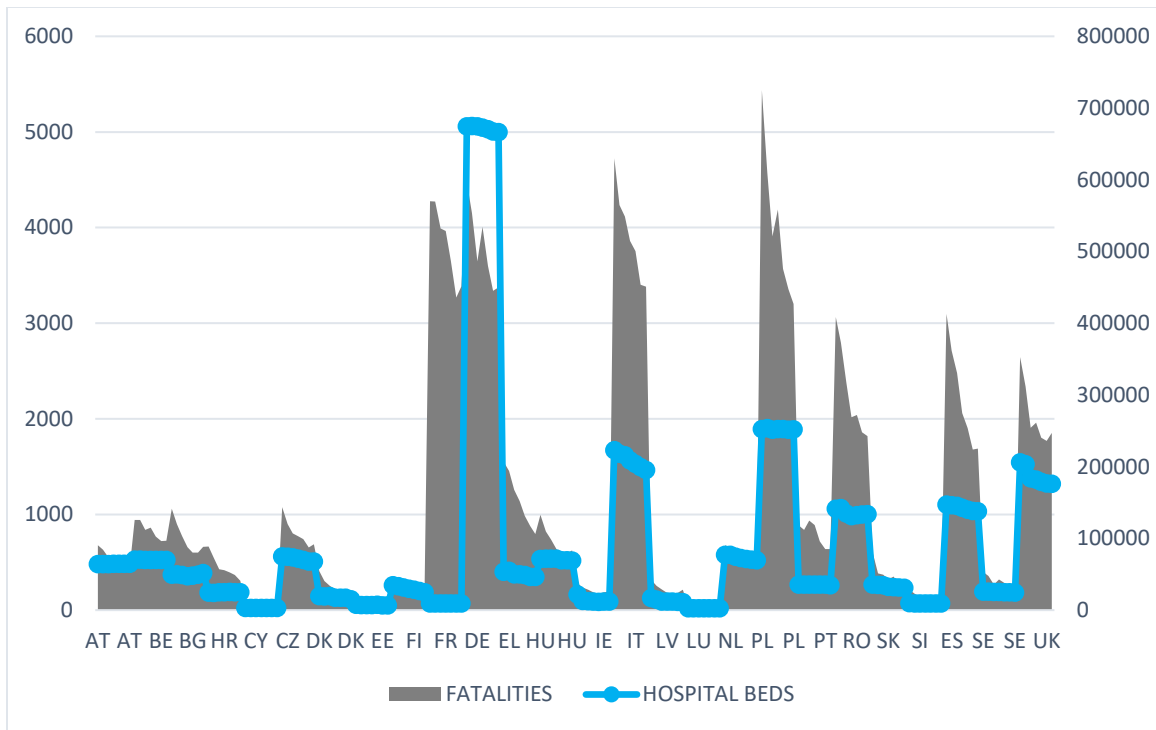
Πίνακας 4.7 Pivot table Ατυχημάτων εντός αστικού ιστού προς θανάσιμα ατυχήματα

NAME	COUNTRY	YEAR	POPULATION	GDP	LOG.FAT.PER POP	URBAN PER FATAL	HDI	WORKIN G LIFE	PRODUCT IVITY	LIFE EXPECTANCY	UNEMPLO YMENT	HOSPITAL BEDS PER POPULATION	PUBLIC HEALTH EXPENDITURE	HEALTH EXPENDITURE PER CAPITA	HEALTH EXPENDITURE %GDP
AT	1.00	2008.00	8307989.00	51629.65	-4.09	0.30	0.87	35.70	117.50	80.60	172.00	0.01	75.40	5286.92	10.60
AT	1.00	2009.00	8335003.00	47785.56	-4.12	0.30	0.87	35.90	117.80	80.50	223.00	0.01	75.63	5153.79	11.19
AT	1.00	2010.00	8351643.00	46757.13	-4.18	0.27	0.88	36.00	115.40	80.70	203.00	0.01	75.23	5049.56	11.17
AT	1.00	2011.00	8375164.00	51192.57	-4.20	0.28	0.88	36.20	115.70	81.10	194.00	0.01	75.11	5431.67	10.94
AT	1.00	2012.00	8408121.00	48381.44	-4.20	0.30	0.89	36.50	117.70	81.10	209.00	0.01	75.37	5239.52	11.17
AT	1.00	2013.00	8451860.00	50533.29	-4.27	0.26	0.89	36.70	117.00	81.30	231.00	0.01	75.16	5478.24	11.14
AT	1.00	2014.00	8507786.00	51390.00	-4.30	0.30	0.89	36.60	116.50	81.60	245.00	0.01	77.86	5580.49	11.21
BE	2.00	2008.00	10666866.00	48850.72	-4.05	0.32	0.88	32.10	128.20	79.80	333.00	0.01	77.06	4565.95	9.60
BE	2.00	2009.00	10753080.00	45176.01	-4.06	0.29	0.88	32.10	129.60	80.20	380.00	0.01	77.39	4574.69	10.39
BE	2.00	2010.00	10839905.00	44691.26	-4.11	0.31	0.88	32.50	131.30	80.30	406.00	0.01	77.68	4419.44	10.17
BE	2.00	2011.00	11000638.00	47950.95	-4.11	0.35	0.89	32.10	129.70	80.70	347.00	0.01	76.81	4864.15	10.42
BE	2.00	2012.00	11075889.00	44900.09	-4.16	0.30	0.89	32.20	130.70	80.50	369.00	0.01	77.67	4587.91	10.54
BE	2.00	2013.00	11137974.00	46613.20	-4.19	0.26	0.89	32.50	130.20	80.70	417.00	0.01	77.81	4813.42	10.57
BE	2.00	2014.00	11180840.00	47634.68	-4.19	0.27	0.90	32.60	130.60	81.40	423.00	0.01	77.87	4884.07	10.59
BG	3.00	2008.00	7518002.00	7152.79	-3.85	0.48	0.77	32.50	41.30	73.30	202.00	0.01	58.54	483.39	6.61
BG	3.00	2009.00	7467119.00	6859.68	-3.92	0.42	0.77	32.10	41.30	73.70	240.00	0.01	55.35	471.57	6.78
BG	3.00	2010.00	7421766.00	6743.74	-3.98	0.52	0.78	31.60	41.30	73.80	352.00	0.01	55.69	488.21	7.24
BG	3.00	2011.00	7369431.00	7836.58	-4.05	0.51	0.78	31.10	42.20	74.20	376.00	0.01	54.62	532.50	6.88
BG	3.00	2012.00	7327224.00	7401.94	-4.09	0.59	0.78	31.60	43.50	74.40	410.00	0.01	56.29	519.48	7.11
BG	3.00	2013.00	7284552.00	7703.48	-4.08	0.58	0.79	32.00	42.90	74.90	436.00	0.01	51.99	606.03	7.93
BG	3.00	2014.00	7245677.00	7877.09	-4.04	0.54	0.79	32.00	43.80	74.50	385.00	0.01	54.57	661.85	8.44
HR	4.00	2008.00	4311967.00	15889.11	-3.81	0.71	0.80	31.70	70.50	76.00	166.00	0.01	84.87	1265.14	7.70
HR	4.00	2009.00	4309796.00	14142.28	-3.90	0.64	0.80	31.80	68.50	76.30	180.00	0.01	85.69	1193.49	8.18
HR	4.00	2010.00	4302847.00	13505.03	-4.00	0.66	0.81	31.60	67.20	76.70	224.00	0.01	85.66	1143.85	8.25
HR	4.00	2011.00	4289857.00	14537.96	-4.01	0.65	0.82	31.40	70.20	77.20	256.00	0.01	80.21	1131.67	7.80
HR	4.00	2012.00	4275984.00	13234.67	-4.04	0.65	0.82	31.20	72.40	77.30	292.00	0.01	81.61	1029.52	7.80
HR	4.00	2013.00	4262140.00	13573.74	-4.06	0.65	0.82	31.10	73.40	77.80	320.00	0.01	81.79	1060.83	7.83
HR	4.00	2014.00	4246809.00	13468.73	-4.14	0.67	0.82	32.30	70.70	77.90	325.00	0.01	81.87	1050.33	7.80
CY	5.00	2008.00	776333.00	19006.17	-3.98	0.72	0.85	36.40	94.90	80.60	15.00	0.00	41.36	2241.27	6.89
CY	5.00	2009.00	796930.00	18673.53	-4.05	0.66	0.85	36.80	95.20	81.00	22.00	0.00	44.74	2185.84	7.40
CY	5.00	2010.00	819140.00	19299.51	-4.14	0.75	0.85	36.90	91.50	81.50	26.00	0.00	47.35	2038.20	7.23
CY	5.00	2011.00	839751.00	19730.97	-4.07	0.60	0.85	36.60	89.90	81.20	34.00	0.00	46.54	2235.67	7.54
CY	5.00	2012.00	862011.00	19489.72	-4.23	0.61	0.85	36.30	88.70	81.10	52.00	0.00	45.86	1964.82	7.44
CY	5.00	2013.00	865878.00	18140.49	-4.29	0.73	0.85	36.30	86.80	82.50	69.00	0.00	46.53	1889.87	7.46
CY	5.00	2014.00	858000.00	17605.57	-4.28	0.77	0.85	36.80	85.00	82.30	70.00	0.00	45.23	1819.11	7.37
CZ	6.00	2008.00	10343422.00	22819.08	-3.98	0.45	0.86	33.70	77.70	77.30	230.00	0.01	83.74	1486.18	6.82
CZ	6.00	2009.00	10425783.00	19787.77	-4.06	0.39	0.86	33.90	79.20	77.40	352.00	0.01	83.76	1483.94	7.85
CZ	6.00	2010.00	10462088.00	19831.40	-4.12	0.39	0.86	33.90	77.10	77.70	384.00	0.01	84.18	1410.04	7.43
CZ	6.00	2011.00	10486731.00	21736.84	-4.13	0.39	0.86	33.90	77.40	78.00	351.00	0.01	84.02	1545.57	7.50
CZ	6.00	2012.00	10505445.00	19739.90	-4.15	0.39	0.87	34.30	76.20	78.10	367.00	0.01	84.33	1411.48	7.55

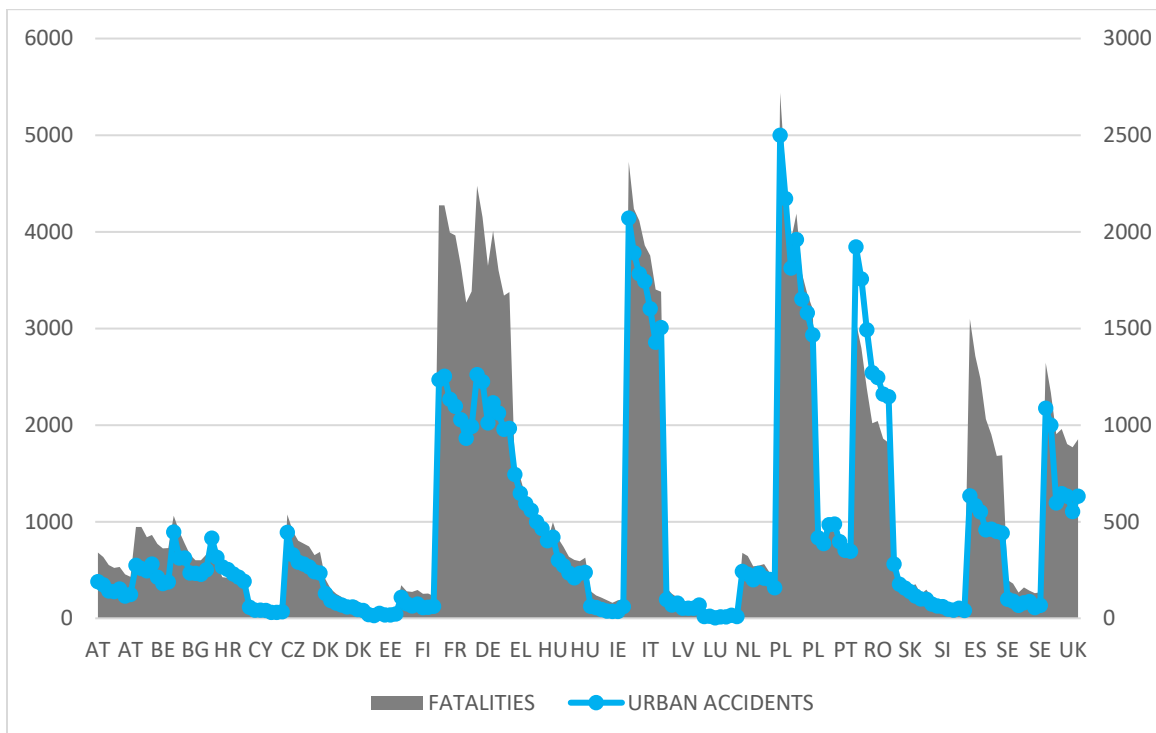
Πίνακας 4.8 Βάση Δεδομένων excel No.2

Αυτή είναι η δεύτερη βάση δεδομένων η οποία θα αξιοποιηθεί σαν αρχείο εισόδου για την εφαρμογή της λογαριθμισμένης γραμμικής παλινδρόμησης, όπου τα ατυχήματα εντός αστικού ιστού έχουν αντικατασταθεί από τα ατυχήματα εντός αστικού ιστού ως προς το σύνολο των θανατηφόρων ατυχημάτων και οι νοσοκομειακές κλίνες έχουν αντικατασταθεί με τις νοσοκομειακές κλίνες ως προς τον πληθυσμό της κάθε χώρας.

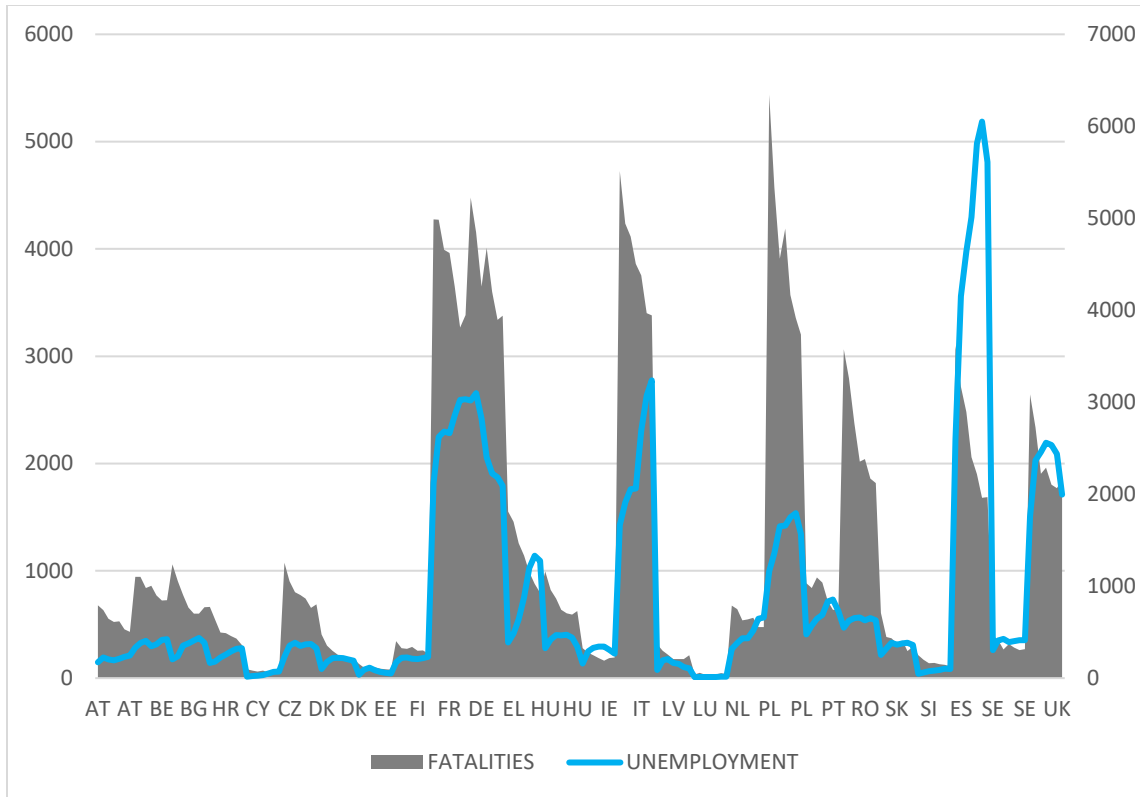
Στη συνέχεια παρατίθενται κάποια ενδεικτικά διαγράμματα των υπό μελέτη παραγόντων ως προς την θνησιμότητα εξαιτίας των οδικών ατυχημάτων.



Διάγραμμα 4.1 Θνησιμότητα στα οδικά ατυχήματα και πλήθος κλινών στα νοσοκομεία ως προς τις προς μελέτη χώρες



Διάγραμμα 4.2 Θνησιμότητα στα οδικά ατυχήματα και πλήθος ατυχημάτων εντός αστικού ιστού ως προς τις προς μελέτη χώρες



Διάγραμμα 4.3 Θνησιμότητα στα οδικά ατυχήματα και ποσοστό ανεργίας ως προς τις προς μελέτη χώρες

Από τα παραπάνω διαγράμματα φαίνεται ότι υπάρχει σχέση μεταξύ της θνησιμότητας στα οδικά ατυχήματα και των παραγόντων των νοσοκομειακών κλινών, των οδικών ατυχημάτων εντός αστικού ιστού και της ανεργίας. Μέσα από τα διαγράμματα καθίσταται κατανοητό πως τα σημεία μεγιστοποίησης του ενός μεγέθους τείνουν να είναι σημεία μεγίστου και στο άλλο μέγεθος, γεγονός που οδηγεί στο συμπέρασμα ότι η αύξηση του ενός παράγοντα είτε της ανεργίας, είτε των ατυχημάτων εντός αστικού ιστού, δύναται να αυξήσει το βαθμό θνησιμότητας στα οδικά ατυχήματα και αντίστοιχα για τις νοσοκομειακές κλίνες, η αύξηση της θνησιμότητας λόγω οδικών ατυχημάτων φαίνεται να μπορεί να προκαλέσει αύξηση του αριθμού των νοσοκομειακών κλινών. Μέσα από την παραπάνω διαδικασία διαμόρφωσης των προαναφερθέντων διαγραμμάτων διαπιστώνονται κάποιες αρχικές σχέσεις μεταξύ των προς μελέτη παραγόντων και της θνησιμότητας στα οδικά ατυχήματα. Τα συμπεράσματα αυτά φάνηκαν ιδιαίτερος χρήσιμα και κατέστησαν σκόπιμο να διερευνηθεί στην παρούσα Διπλωματική Εργασία η συσχέτιση όλων αυτών των δεικτών με τις επιδόσεις οδικής ασφάλειας μέσω της δημιουργίας ενός μαθηματικού μοντέλου που θα περιελάμβανε όσο το δυνατόν περισσότερους δείκτες από κάθε κατηγορία (υγεία, οικονομία, κοινωνία). Επιπλέον επιβεβαίωσαν την δυνατότητα ομαδοποίησης των κρατών της Ε.Ε σε σύνολα που παρουσίαζαν παρόμοια ιατρικά, οικονομικά, κοινωνικά και συγκοινωνιακά χαρακτηριστικά, ούτως ώστε να δημιουργηθεί και για κάθε ένα από αυτά τα σύνολα κρατών ένα αντίστοιχο μαθηματικό μοντέλο.

5. Εφαρμογή Μεθοδολογίας και Αποτελέσματα

5.1 Εισαγωγή

Το κεφάλαιο αυτό περιλαμβάνει την αναλυτική περιγραφή της μεθοδολογίας και την παρουσίαση των αποτελεσμάτων της παρούσας Διπλωματικής Εργασίας. Μετά τη συλλογή και την επεξεργασία των στοιχείων στο πρόγραμμα Microsoft Excel, όπως αναφέρθηκε στο προηγούμενο κεφάλαιο, ακολούθησε η στατιστική ανάλυση των δεδομένων. Η στατιστική ανάλυση πραγματοποιήθηκε αρχικά με το γραμμικό μοντέλο παλινδρόμησης, το οποίο πολύ σύντομα αντικαταστάθηκε από το λογαριθμισμένο μοντέλο παλινδρόμησης, διότι αυτό θεωρήθηκε ιδανικότερο. Στη συνέχεια και αφού εξήχθησαν τα τελικά αποτελέσματα για το σύνολο των κρατών, επιχειρήθηκε μια στατιστική ανάλυση χρησιμοποιώντας το Γραμμικό Μικτό Μοντέλο, σε τρία διαφορετικά σετ χωρών, σύμφωνα με την εργασία της Υπατίας Μίχου[2] και σύμφωνα με το διαχωρισμό που προτείνεται εκεί.

Στη συνέχεια περιγράφονται αναλυτικά τα βήματα που ακολουθήθηκαν κατά την εφαρμογή της μεθοδολογίας και παρουσιάζεται η διαδικασία ανάπτυξης των κατάλληλων μοντέλων. Η επεξεργασία πραγματοποιήθηκε σε τέσσερις φάσεις, από τις οποίες προέκυψαν τα αντίστοιχα μοντέλα. Αναπόσπαστο μέρος των αποτελεσμάτων αποτελούν οι στατιστικοί έλεγχοι που απαιτούνται για την αποδοχή ή μη των μοντέλων.

Τέλος, στο παρόν κεφάλαιο παρουσιάζονται και περιγράφονται τα αποτελέσματα που προέκυψαν από τη στατιστική επεξεργασία και επιχειρείται προσπάθεια εξήγησής τους με βάση τη λογική, την εμπειρία και τα στοιχεία από τη σχετική βιβλιογραφία. Η παρουσίαση των αποτελεσμάτων διακρίνεται σε τρεις φάσεις:

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

Αξίζει να σημειωθεί, ότι πριν την επιλογή του τελικού μοντέλου κατά τη στατιστική ανάλυση, αναπτύχθηκαν ποικίλα μοντέλα λογαριθμισμένης κανονικής παλινδρόμησης, ορισμένα από τα οποία παρατίθενται στη συνέχεια, που έπαιξαν σημαντικό ρόλο στην εξέλιξη της μεθοδολογίας και στον καθορισμό του τελικού μοντέλου.

5.2 Βάση Δεδομένων

Όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο, τα απαιτούμενα στοιχεία δεν ήταν πλήρη για ορισμένα κράτη και έτη με συνέπεια να μη συμπεριληφθούν στην ανάλυση. Συγκεκριμένα, τα κράτη αυτά είναι η Μάλτα και η Λιθουανία.

Για την πραγματοποίηση της παλινδρόμησης αξιοποιήθηκαν από την ήδη υπάρχουσα Βάση Δεδομένων Νο. 1 (βλ. Κεφ. 4 σελ. 42), για κάθε κράτος (country) και για τα αντίστοιχα έτη (year) ο αριθμός των νεκρών στα οδικά ατυχήματα (fatalities) και ο πληθυσμός (population), ώστε να προκύψει ο αριθμός των νεκρών στα οδικά ατυχήματα ανά πληθυσμό (FatPerPop) και τελικά ο φυσικός λογάριθμος του λόγου αυτού (Log(FatPerPop)), ώστε να χρησιμοποιηθεί σαν εξαρτημένη μεταβλητή. Σαν ανεξάρτητες μεταβλητές χρησιμοποιούνται το πλήθος των ατυχημάτων εντός αστικού δικτύου προς το σύνολο των θανατηφόρων ατυχημάτων (Urban.Per.Fatal), το προσδόκιμο ζωής των ανθρώπων (Life Expectancy), οι δαπάνες για τον τομέα της υγείας ως προς το Α.Ε.Π.(Health Expenditure per GDP), οι δαπάνες για τον τομέα της υγείας ως προς το πλήθος των πολιτών (health Expenditure per Capita), οι δημόσιες δαπάνες για τον τομέα της υγείας (Public Health Expenditure), οι νοσοκομειακές κλίνες ανά πληθυσμό (Hospital.Beds.Per.Population), το κατά κεφαλήν Α.Ε.Π (GDP), ο δείκτης ανθρώπινης ανάπτυξης (HDI), τα παραγωγικά χρόνια ζωής των ανθρώπων (Working Life) και ο δείκτης ανεργίας (Unemployment).

5.3 Γραμμικό Μοντέλο

Μετά τη συλλογή και την επεξεργασία των στοιχείων στο excel, με τη διαδικασία που αναπτύχθηκε στο προηγούμενο κεφάλαιο, ακολούθησε η στατιστική ανάλυση των στοιχείων. Σε πρώτο επίπεδο διερευνήθηκε η συσχέτιση του αριθμού των νεκρών στα οδικά ατυχήματα ως προς τον πληθυσμό, το Α.Ε.Π., τις δαπάνες στην υγεία, ως προς το ΑΕΠ και τον αριθμό των κλινών στα νοσοκομεία, με τη μέθοδο του γραμμικού μοντέλου.

Το παρακάτω μοντέλο έδωσε κάποια αποτελέσματα, τα οποία, όμως δεν ήταν πολύ ικανοποιητικά.

lm(formula = FATALITIES ~ POPULATION + GDP + `HEALTH EXPENDITURE %GDP` + `HOSPITAL BEDS`)					
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1.286e+03	2.355e+02	5.460	1.59e-07	***
POPULATION	5.367e-05	2.999e-06	17.893	< 2e-16	***
GDP	-5.258e-03	2.155e-03	-2.440	0.015661	*
`HEALTH EXPENDITURE %GDP`	-1.098e+02	3.056e+01	-3.593	0.000423	***
`HOSPITAL BEDS`	8.232e-05	4.916e-04	0.167	0.867198	
Multiple R-squared: 0.8152, Adjusted R-squared: 0.8111					
F-statistic: 195.3 on 4 and 177 DF, p-value: < 2.2e-16					

Μοντέλο 5.1

Έτσι, δημιουργήθηκαν μερικά πειραματικά μοντέλα, όπως φαίνεται παρακάτω, όπου οι ανεξάρτητες μεταβλητές άλλαζαν διαρκώς προκειμένου να βρεθεί ο καταλληλότερος συνδυασμός τους.

```
lm(formula = FATALITIES ~ POPULATION + `HEALTH EXPENDITURE %GDP` +
`HOSPITAL BEDS`)
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1.394e+03	2.345e+02	5.943	1.44e-08	***
POPULATION	5.417e-05	3.034e-06	17.854	< 2e-16	***
`HEALTH EXPENDITURE %GDP`	-1.442e+02	2.748e+01	-5.248	4.34e-07	***
`HOSPITAL BEDS`	1.127e-04	4.982e-04	0.226	0.821	

Multiple R-squared: 0.809, Adjusted R-squared: 0.8058
F-statistic: 251.4 on 3 and 178 DF, p-value: < 2.2e-16

Μοντέλο 5.2

```
lm(formula = FATALITIES ~ POPULATION + `WORKING LIFE` + `HEALTH EXPENDI
TURE%GDP` + `HOSPITAL BEDS`)
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	4.340e+03	4.630e+02	9.373	< 2e-16	***
POPULATION	5.044e-05	2.733e-06	18.456	< 2e-16	***
`WORKING LIFE`	-1.101e+02	1.547e+01	-7.117	2.65e-11	***
`HEALTH EXPENDITURE %GDP`	-4.022e+01	2.835e+01	-1.418	0.158	
`HOSPITAL BEDS`	4.139e-04	4.426e-04	0.935	0.351	

Multiple R-squared: 0.8515, Adjusted R-squared: 0.8482
F-statistic: 253.8 on 4 and 177 DF, p-value: < 2.2e-16

Μοντέλο 5.3

Καθώς τα μοντέλα αυτά προέκυψαν με διερευνητικό τρόπο και χωρίς την εφαρμογή κάποιας συγκεκριμένης μεθοδολογίας, αποτέλεσαν απλώς μια πρώτη ενδεικτική επαφή με το αντικείμενο και παρατίθενται μόνο για λόγους πληρότητας της εργασίας, καθώς και για την καλύτερη κατανόηση των βημάτων που ακολουθήθηκαν στην παρούσα Διπλωματική Εργασία.

5.4 Πίνακας Αυτοσυσχέτισης

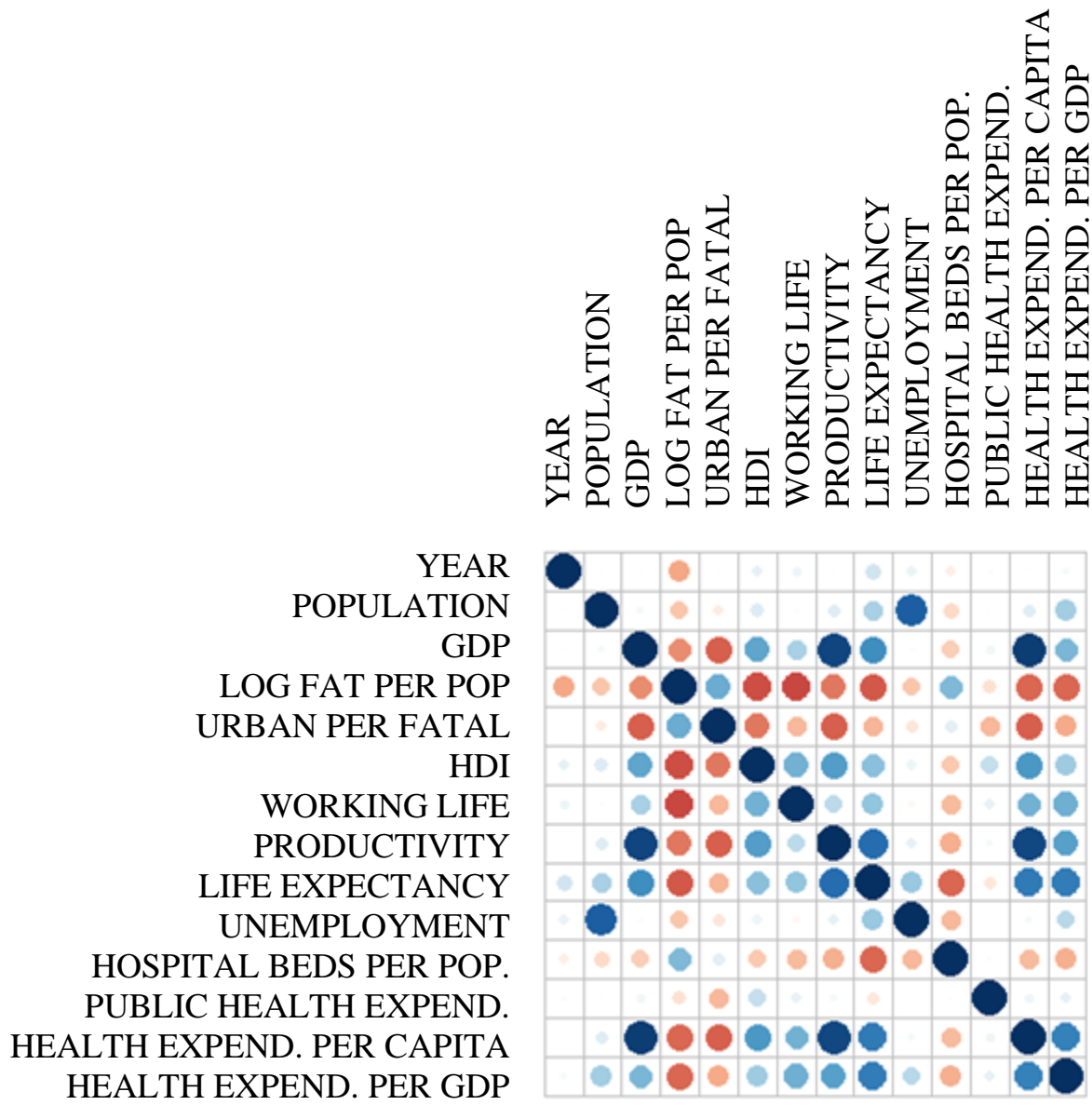
Προκειμένου να αντιμετωπιστεί το προαναφερθέν πρόβλημα, διαμορφώθηκε ένας πίνακας αυτοσυσχέτισης, ώστε να καταστεί κατανοητό ποιο δείκτες έχουν άμεση σχέση με τη θνησιμότητα στα οδικά ατυχήματα και σε ποιο βαθμό.

Παρατηρήθηκε, επίσης, ότι, προκειμένου να προκύψουν ορθότερα αποτελέσματα θα πρέπει τα μεγέθη να είναι συγκρίσιμα και συνεπώς απαιτείται η αξιοποίηση της Βάσης Δεδομένων Νο. 2 (βλ. Κεφ. 4 σελ. 48), όπου τα ατυχήματα εντός αστικού ιστού έχουν αντικατασταθεί από τα ατυχήματα εντός αστικού ιστού ως προς το σύνολο των θανατηφόρων ατυχημάτων και οι νοσοκομειακές κλίνες έχουν αντικατασταθεί με τις νοσοκομειακές κλίνες ως προς τον πληθυσμό της κάθε χώρας. Ως εξαρτημένη μεταβλητή παραμένει ο φυσικός λογάριθμος θνησιμότητας προς τον πληθυσμό, ενώ αναπτύσσονται νέα, βελτιωμένα μοντέλα.

Σύμφωνα με τις παραπάνω παρατηρήσεις παρατίθενται παρακάτω ο πίνακας αυτοσυσχέτισης για τη Βάση Δεδομένων Νο. 2, καθώς επίσης και το αντίστοιχο διάγραμμα αυτοσυσχέτισης, το οποίο καθιστά ευκολότερη την κατανόηση των σχέσεων των μεταβλητών μεταξύ τους.

	YEAR	POPULATION	GDP	LOG.FAT. PER.POP	URBAN PER FATAL	HDI	WORKING LIFE	PRODUCTIVI TY	LIFE EXPECTANCY	UNEMPLOY MENT	HOSPITAL BEDS PER POPULATION	PUBLIC HEALTH EXPENDITURE	HEALTH EXPENDITURE PER CAPITA	HEALTH EXPENDITURE %GDP
YEAR	1.00	0.00	-0.01	-0.39	-0.01	0.10	0.07	0.01	0.19	0.08	-0.09	-0.02	0.01	0.03
POPULATION	0.00	1.00	0.05	-0.27	-0.10	0.15	0.02	0.13	0.33	0.82	-0.19	0.01	0.14	0.35
GDP	-0.01	0.05	1.00	-0.47	-0.60	0.53	0.33	0.91	0.61	-0.02	-0.25	0.05	0.95	0.45
LOG.FAT.PER.P OP	-0.39	-0.27	-0.47	1.00	0.49	-0.65	-0.67	-0.52	-0.61	-0.28	0.45	-0.16	-0.57	-0.57
URBAN PER FATAL	-0.01	-0.10	-0.60	0.49	1.00	-0.53	-0.33	-0.60	-0.33	-0.13	0.12	-0.32	-0.59	-0.38
HDI	0.10	0.15	0.53	-0.65	-0.53	1.00	0.48	0.56	0.41	0.08	-0.27	0.24	0.57	0.35
WORKING LIFE	0.07	0.02	0.33	-0.67	-0.33	0.48	1.00	0.26	0.40	-0.03	-0.32	0.09	0.47	0.48
PRODUCTIVITY	0.01	0.13	0.91	-0.52	-0.60	0.56	0.26	1.00	0.76	0.10	-0.35	0.04	0.90	0.54
LIFE EXPECTANCY	0.19	0.33	0.61	-0.61	-0.33	0.41	0.40	0.76	1.00	0.38	-0.58	-0.12	0.70	0.70
UNEMPLOYME NT	0.08	0.82	-0.02	-0.28	-0.13	0.08	-0.03	0.10	0.38	1.00	-0.32	0.00	0.04	0.28
HOSPITAL BEDS PER POPULATION	-0.09	-0.19	-0.25	0.45	0.12	-0.27	-0.32	-0.35	-0.58	-0.32	1.00	-0.01	-0.31	-0.35
PUBLIC HEALTH EXPENDITURE	-0.02	0.01	0.05	-0.16	-0.32	0.24	0.09	0.04	-0.12	0.00	-0.01	1.00	0.07	0.10
HEALTH EXPENDITURE PER CAPITA	0.01	0.14	0.95	-0.57	-0.59	0.57	0.47	0.90	0.70	0.04	-0.31	0.07	1.00	0.68
HEALTH EXPENDITURE %GDP	0.03	0.35	0.45	-0.57	-0.38	0.35	0.48	0.54	0.70	0.28	-0.35	0.10	0.68	1.00

Πίνακας 5.1 Πίνακας αυτοσυσχέτισης



Διάγραμμα 5.1 Διάγραμμα αυτοσυσχέτισης

Μέσω του διαγράμματος γίνονται πιο κατανοητές οι σχέσεις μεταξύ των δεικτών, καθώς τα κελιά με τα πιο έντονα χρώματα υποδηλώνουν μεγαλύτερη συσχέτιση μεταξύ των μεγεθών τους, είτε αρνητική (κόκκινο χρώμα) είτε θετική (μπλε χρώμα). Αρνητική σχέση σημαίνει, πως όσο το ένα μέγεθος αυξάνεται το άλλο μειώνεται, π.χ. όσο αυξάνεται ο δείκτης του HDI αναμένεται να μειώνονται τα τροχαία ατυχήματα ανά πληθυσμό, ενώ θετική σημαίνει, πως όσο το ένα μέγεθος αυξάνεται, τότε θα αυξάνεται και το άλλο μέγεθος, π.χ. όσο αυξάνεται η παραγωγικότητα αναμένεται να αυξάνονται οι δαπάνες για τον τομέα της υγείας ανά κάτοικο.

5.5 Λογαριθμισμένη Γραμμική Παλινδρόμηση

Στη συνέχεια, σύμφωνα με τη Βάση Δεδομένων Νο.2, τον παραπάνω πίνακα και το αντίστοιχο διάγραμμα αυτοσυσχέτισης, δημιουργήθηκαν κάποια νέα στατιστικά μοντέλα με τη μέθοδο της λογαριθμισμένης γραμμικής παλινδρόμησης. Αρχικά, εξετάστηκαν οι δείκτες, οι οποίοι φαίνεται να έχουν μεγαλύτερη συσχέτιση με τις ανθρώπινες απώλειες ανά πληθυσμό, αυτή τη φορά με τη μορφή της λογαριθμισμένης σχέσης $\text{Log}(\text{FatPerPop})$.

Καθώς στο βήμα αυτό παρήχθησαν πάνω από τριάντα διαφορετικά μοντέλα στατιστικής ανάλυσης, επιλέγεται να παρουσιαστούν μόνο ορισμένα από αυτά, τα οποία και παρατίθενται παρακάτω.

```
lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + LIFE.EXPECTANCY)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.930239	0.216464	-4.297	2.84e-05 ***
WORKING.LIFE	-0.023097	0.003093	-7.468	3.51e-12 ***
HDI	-0.997816	0.155914	-6.400	1.34e-09 ***
LIFE.EXPECTANCY	-0.020337	0.003075	-6.613	4.25e-10 ***

Multiple R-squared: 0.6656, Adjusted R-squared: 0.66
F-statistic: 118.1 on 3 and 178 DF, p-value: < 2.2e-16

Μοντέλο 5.4

```
lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + LIFE.EXPECTANCY + HEALTH.EXPENDITURE.PER.CAPITA)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-7.333e-01	3.173e-01	-2.311	0.022 *
WORKING.LIFE	-2.352e-02	3.134e-03	-7.503	2.92e-12 ***
HDI	-1.049e+00	1.674e-01	-6.268	2.71e-09 ***
LIFE.EXPECTANCY	-2.228e-02	3.836e-03	-5.809	2.87e-08 ***
HEALTH.EXPENDITURE.PER.CAPITA	4.943e-06	5.818e-06	0.850	0.397

Multiple R-squared: 0.667, Adjusted R-squared: 0.6594
F-statistic: 88.62 on 4 and 177 DF, p-value: < 2.2e-16

Μοντέλο 5.5

```
lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + LIFE.EXPECTANCY + HEALTH.EXPENDITURE..GDP)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.186862	0.269939	-4.397	1.89e-05 ***
WORKING.LIFE	-0.021558	0.003230	-6.674	3.09e-10 ***
HDI	-1.003138	0.155300	-6.459	9.82e-10 ***
LIFE.EXPECTANCY	-0.016566	0.003883	-4.266	3.23e-05 ***
HEALTH.EXPENDITURE..GDP	-0.010342	0.006548	-1.579	0.116

Multiple R-squared: 0.6703, Adjusted R-squared: 0.6628
F-statistic: 89.95 on 4 and 177 DF, p-value: < 2.2e-16

Μοντέλο 5.6

Παρατηρείται, ότι ενώ τα μοντέλα εμφανίζουν αρκετά καλά αποτελέσματα, το πλήθος των δεικτών είναι αρκετά περιορισμένο. Έτσι, ξεκινά η εισαγωγή μεταβλητών στα μοντέλα, οι οποίες δεν έχουν τόσο μεγάλη σχέση αλληλεπίδρασης με τη θνησιμότητα ανά πληθυσμό, αλλά έχουν σχέση με τους υπόλοιπους δείκτες, οι οποίοι εμφανίζονται ως ανεξάρτητες μεταβλητές.

Μία από τις πρώτες επιλογές λόγω και του διαγράμματος αυτοσυσχέτισης είναι ο δείκτης της ανεργίας, **UNEMPLOYMENT**.

```
lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + LIFE.EXPECTANCY +
  UNEMPLOYMENT)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.227e+00	2.235e-01	-5.491	1.37e-07 ***
WORKING.LIFE	-2.533e-02	3.044e-03	-8.319	2.30e-14 ***
HDI	-1.004e+00	1.505e-01	-6.672	3.12e-10 ***
LIFE.EXPECTANCY	-1.526e-02	3.264e-03	-4.674	5.83e-06 ***
UNEMPLOYMENT	-2.611e-05	6.972e-06	-3.746	0.000243 ***

Multiple R-squared: 0.6902, Adjusted R-squared: 0.6832
F-statistic: 98.57 on 4 and 177 DF, p-value: < 2.2e-16

Μοντέλο 5.7

```
lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + LIFE.EXPECTANCY +
  UNEMPLOYMENT + PUBLIC.HEALTH.EXPENDITURE)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.062e+00	2.374e-01	-4.476	1.36e-05 ***
WORKING.LIFE	-2.498e-02	3.026e-03	-8.256	3.46e-14 ***
HDI	-9.171e-01	1.559e-01	-5.882	2.00e-08 ***
LIFE.EXPECTANCY	-1.701e-02	3.362e-03	-5.059	1.05e-06 ***
UNEMPLOYMENT	-2.487e-05	6.947e-06	-3.581	0.000443 ***
PUBLIC.HEALTH.EXPENDITURE	-1.563e-03	8.043e-04	-1.944	0.053493 .

Multiple R-squared: 0.6967, Adjusted R-squared: 0.6881
F-statistic: 80.85 on 5 and 176 DF, p-value: < 2.2e-16

Μοντέλο 5.8

Στο μοντέλο 5.8 διακρίνεται ξεκάθαρα η σχέση αλληλεπίδρασης μεταξύ του δείκτη των δημοσίων εξόδων για τον τομέα της υγείας (PUBLIC.HEALTH.EXPENDITURE) και της λογαριθμισμένης τιμής της θνησιμότητας ανά πληθυσμό. Αναλυτικότερα, παρατηρείται, ότι ο δείκτης PUBLIC.HEALTH.EXPENDITURE έχει μικρή σχέση με το προς μελέτη μέγεθος LOG.FAT.PER.POP εξαιτίας της μικρής τιμής που έχει στις στήλες t-value και Pr(>|t|), ενώ καμία από τις τιμές των άλλων δεικτών δεν εμφανίζει σημαντικές διαφοροποιήσεις σε σχέση με το προηγούμενο μοντέλο. Η παρατήρηση αυτή οδηγεί στο συμπέρασμα, ότι ο δείκτης PUBLIC.HEALTH.EXPENDITURE έχει σχέση με το δείκτη **LOG.FAT.PER.POP**, η οποία, όμως, είναι **συγκριτικά μικρότερη** από τη σχέση των υπόλοιπων εξεταζόμενων δεκτών του συγκεκριμένου δείγματος. Ταυτόχρονα, παρατηρείται, ότι ο δείκτης PUBLIC.HEALTH.EXPENDITURE δεν έχει μεγάλη σχέση, έως καθόλου σχέση, με τα υπόλοιπα εξεταζόμενα μεγέθη εντός του μοντέλου εφόσον δεν παρατηρούνται διαφοροποιήσεις στις τιμές

τους. Οι δύο αυτές παρατηρήσεις μπορούν πολύ εύκολα να επιβεβαιωθούν μέσω του διαγράμματος αυτοσυσχέτισης.

Στη συνέχεια, επιλέγεται να εισαχθεί άλλο ένα νέο μέγεθος στην ανάλυση, αυτό των κλινών ανά πληθυσμό **HOSPITAL.BEDS.PER.POPULATION**.

lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + LIFE.EXPECTANCY + UNEMPLOYMENT + HOSPITAL.BEDS.PER.POPULATION)					
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-1.365e+00	2.755e-01	-4.956	1.68e-06	***
WORKING.LIFE	-2.494e-02	3.080e-03	-8.098	8.97e-14	***
HDI	-1.005e+00	1.506e-01	-6.672	3.16e-10	***
LIFE.EXPECTANCY	-1.395e-02	3.604e-03	-3.870	0.000153	***
UNEMPLOYMENT	-2.511e-05	7.075e-06	-3.549	0.000496	***
HOSPITAL.BEDS.PER.POPULATION	4.184e+00	4.873e+00	0.859	0.391764	
Multiple R-squared: 0.6915, Adjusted R-squared: 0.6827					
F-statistic: 78.89 on 5 and 176 DF, p-value: < 2.2e-16					

Μοντέλο 5.9

lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + LIFE.EXPECTANCY + UNEMPLOYMENT + HOSPITAL.BEDS.PER.POPULATION + HEALTH.EXPENDITURE..GDP)					
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-1.615e+00	3.259e-01	-4.954	1.71e-06	***
WORKING.LIFE	-2.340e-02	3.255e-03	-7.190	1.81e-11	***
HDI	-1.010e+00	1.502e-01	-6.721	2.45e-10	***
LIFE.EXPECTANCY	-1.048e-02	4.342e-03	-2.415	0.01678	*
UNEMPLOYMENT	-2.391e-05	7.105e-06	-3.365	0.00094	***
HOSPITAL.BEDS.PER.POPULATION	5.196e+00	4.911e+00	1.058	0.29153	
HEALTH.EXPENDITURE..GDP	-9.140e-03	6.430e-03	-1.421	0.15700	
Multiple R-squared: 0.695, Adjusted R-squared: 0.6845					
F-statistic: 66.46 on 6 and 175 DF, p-value: < 2.2e-16					

Μοντέλο 5.10

Το αποτέλεσμα της εισαγωγής του νέου μεγέθους, **HOSPITAL.BEDS.PER.POPULATION**, φαίνεται να είναι μια μικρή μείωση στην τιμή του R^2 από το μοντέλο 5.7 στο μοντέλο 5.9, γεγονός το οποίο υποδηλώνει τον αρνητικό τρόπο επίδρασης του δείκτη αυτού στο μοντέλο είτε εξαιτίας της περιορισμένης συσχέτισής του με την εξαρτημένη μεταβλητή **LOG.FAT.PER.POP**, είτε λόγω κακού συνδυασμού με κάποιον από τους υπόλοιπους δείκτες.

Παρατηρείται, ότι με την εισαγωγή στο μοντέλο 5.10 ενός ακόμα δείκτη, του **HEALTH.EXPENDITURE..GDP**, αυξάνεται το R^2 και συνολικά το μοντέλο έχει καλύτερη συμπεριφορά.

Συνεχίζοντας στην ίδια κατεύθυνση, επιδιώκεται η μελέτη ενός ακόμα μοντέλου μετά την προσθήκη άλλης μίας μεταβλητής, η οποία έχει μεγάλη σχέση αλληλεπίδρασης με το **LOG.FAT.PER.POP**, της παραγωγικότητας **PRODUCTIVITY**.


```
lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + UNEMPLOYMENT +
  HOSPITAL.BEDS.PER.POPULATION + PRODUCTIVITY + HEALTH.EXPENDITURE..GDP)
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2.390e+00	1.351e-01	-17.685	< 2e-16	***
WORKING.LIFE	-2.537e-02	3.380e-03	-7.507	2.96e-12	***
HDI	-8.839e-01	1.697e-01	-5.210	5.26e-07	***
UNEMPLOYMENT	-2.977e-05	7.087e-06	-4.201	4.22e-05	***
HOSPITAL.BEDS.PER.POPULATION	7.831e+00	4.559e+00	1.718	0.0876	.
PRODUCTIVITY	-8.897e-04	3.749e-04	-2.373	0.0187	*
HEALTH.EXPENDITURE..GDP	-1.109e-02	6.040e-03	-1.836	0.0680	.

Multiple R-squared: 0.6946, Adjusted R-squared: 0.6842
F-statistic: 66.35 on 6 and 175 DF, p-value: < 2.2e-16

Μοντέλο 5.11

Όπως φαίνεται, η συγκεκριμένη μεταβλητή δε συμβάλλει σημαντικά στο μοντέλο 5.11, όπου αντικαθιστά το δείκτη της διάρκειας ζωής (LIFE.EXPECTANCY) αναφορικά με το μοντέλο 5.10.

Αυτό είναι αναμενόμενο, καθώς ο νεοεισαχθείς δείκτης PRODUCTIVITY έχει μεγάλη σχέση με το δείκτη LIFE.EXPECTANCY καθώς επίσης και με τους άλλους δείκτες, οι οποίοι χρησιμοποιούνται στο μοντέλο 5.10, όπως οι HDI και HEALTH.EXPENDITURE..GDP. Εξαιτίας της μεγάλης σχέσης αλληλεπίδρασης μεταξύ LIFE.EXPECTANCY και PRODUCTIVITY, αυτοί οι δύο δείκτες δε μελετώνται μαζί στο ίδιο μοντέλο, καθώς το αναμενόμενο αποτέλεσμα είναι η δημιουργία ενός μοντέλου χειρότερου από τα προηγούμενα, εξαιτίας των ισχυρών σχέσεων εξάρτησης μεταξύ των ανεξάρτητων μεταβλητών του.

Σε ένα τέτοιο μοντέλο μπορεί το R^2 να μην επηρεαστεί ιδιαίτερα, παρόλα αυτά θα παρατηρηθούν διαφορές στην ατομική αξιολόγηση των διαφόρων μεταβλητών και στον τρόπο με τον οποίο επηρεάζουν το συνολικό μοντέλο.

Ένα τέτοιο μοντέλο παρουσιάζεται παρακάτω.

```
lm(formula = LOG.FAT.PER.POP ~ WORKING LIFE + HDI + UNEMPLOYMENT +
  HOSPITAL BEDS PER POPULATION + PRODUCTIVITY + HEALTH EXPENDITURE %GDP
  + `LIFE EXPECTANCY`)
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-1.921e+00	4.449e-01	-4.319	2.63e-05	***
WORKING LIFE`	-2.457e-02	3.454e-03	-7.115	2.80e-11	***
HDI	-9.225e-01	1.731e-01	-5.329	3.03e-07	***
UNEMPLOYMENT	-2.669e-05	7.614e-06	-3.505	0.000581	***
HOSPITAL BEDS PER POPULATION	5.727e+00	4.938e+00	1.160	0.247796	
PRODUCTIVITY	-5.133e-04	5.065e-04	-1.013	0.312285	
HEALTH EXPENDITURE %GDP	-8.565e-03	6.455e-03	-1.327	0.186283	
LIFE EXPECTANCY	-6.481e-03	5.870e-03	-1.104	0.271035	

Multiple R-squared: 0.6968, Adjusted R-squared: 0.6846
F-statistic: 57.12 on 7 and 174 DF, p-value: < 2.2e-16

Μοντέλο 5.12

Μέσω του μοντέλου 5.12 φαίνεται το αποτέλεσμα χρήσης εντός του ίδιου μοντέλου δύο ισχυρά εξαρτημένων μεταβλητών και καθίσταται κατανοητός, εξαιτίας των κακών στην πλειοψηφία αποτελεσμάτων, ο λόγος για τον οποίο κάτι τέτοιο πρέπει να αποφεύγεται. Παράλληλα, μέσω αυτού του μοντέλου προκύπτει η χρησιμότητα του πίνακα αυτοσυσχέτισης, εφόσον αυτός βοηθά το μελετητή στην αποφυγή χρήσης εντός του ίδιου μοντέλου μεταβλητών με μεγάλη σχέση εξάρτησης.

Τέλος, προστίθεται άλλος ένας δείκτης στη διαδικασία εύρεσης του κατάλληλου μοντέλου για την περιγραφή του προβλήματος της παρούσας εργασίας, ο αριθμός των οδικών ατυχημάτων εντός αστικού ιστού προς το σύνολο των θανάσιμων οδικών ατυχημάτων, **URBAN.PER.FATAL**.

```
lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + UNEMPLOYMENT +
  URBAN.PER.FATAL + HOSPITAL.BEDS.PER.POPULATION +
  PUBLIC.HEALTH.EXPENDITURE)
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2.602e+00	1.816e-01	-14.326	< 2e-16	***
WORKING.LIFE	-2.626e-02	3.107e-03	-8.451	1.09e-14	***
HDI	-9.305e-01	1.678e-01	-5.546	1.06e-07	***
UNEMPLOYMENT	-3.050e-05	6.969e-06	-4.377	2.07e-05	***
URBAN.PER.FATAL	1.809e-01	6.556e-02	2.759	0.00641	**
HOSPITAL.BEDS.PER.POPULATION	1.341e+01	4.532e+00	2.959	0.00352	**
PUBLIC.HEALTH.EXPENDITURE	-9.102e-06	8.238e-04	-0.011	0.99120	

Multiple R-squared: 0.6801, Adjusted R-squared: 0.6691
F-statistic: 61.99 on 6 and 175 DF, p-value: < 2.2e-16

Μοντέλο 5.13

Παρατηρείται, ότι το άνωθεν μοντέλο υπ' αριθμόν 5.13, αποτελεί μια από τις καλύτερες μοντελοποιήσεις του προβλήματος μέχρι στιγμής τόσο από πλευράς συνολικού αποτελέσματος, όσο και από πλευράς επιλογής μεταβλητών, ώστε να μην υπάρχουν μεγάλες σχέσεις εξάρτησης μεταξύ των ανεξάρτητων μεταβλητών.

Παρακάτω παρατίθεται άλλη μια προσπάθεια ένταξης της μεταβλητής **URBAN.PER.FATAL** στο μοντέλο.

```
lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI +
  HOSPITAL.BEDS.PER.POPULATION + URBAN.PER.FATAL + UNEMPLOYMENT +
  HEALTH.EXPENDITURE.PER.CAPITA)
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2.646e+00	1.727e-01	-15.315	< 2e-16	***
WORKING.LIFE	-2.517e-02	3.135e-03	-8.029	1.38e-13	***
HDI	-8.533e-01	1.710e-01	-4.989	1.46e-06	***
HOSPITAL.BEDS.PER.POPULATION	1.170e+01	4.593e+00	2.546	0.0117	*
URBAN.PER.FATAL	1.275e-01	6.983e-02	1.825	0.0697	.
UNEMPLOYMENT	-3.180e-05	6.938e-06	-4.583	8.70e-06	***
HEALTH.EXPENDITURE.PER.CAPITA	-9.099e-06	5.114e-06	-1.779	0.0769	.

Multiple R-squared: 0.6857, Adjusted R-squared: 0.675
F-statistic: 63.64 on 6 and 175 DF, p-value: < 2.2e-16

Μοντέλο 5.14

Και αυτό το μοντέλο είναι αρκετά καλό, χωρίς όμως να προσφέρει τα επιδιωκόμενα αποτελέσματα.

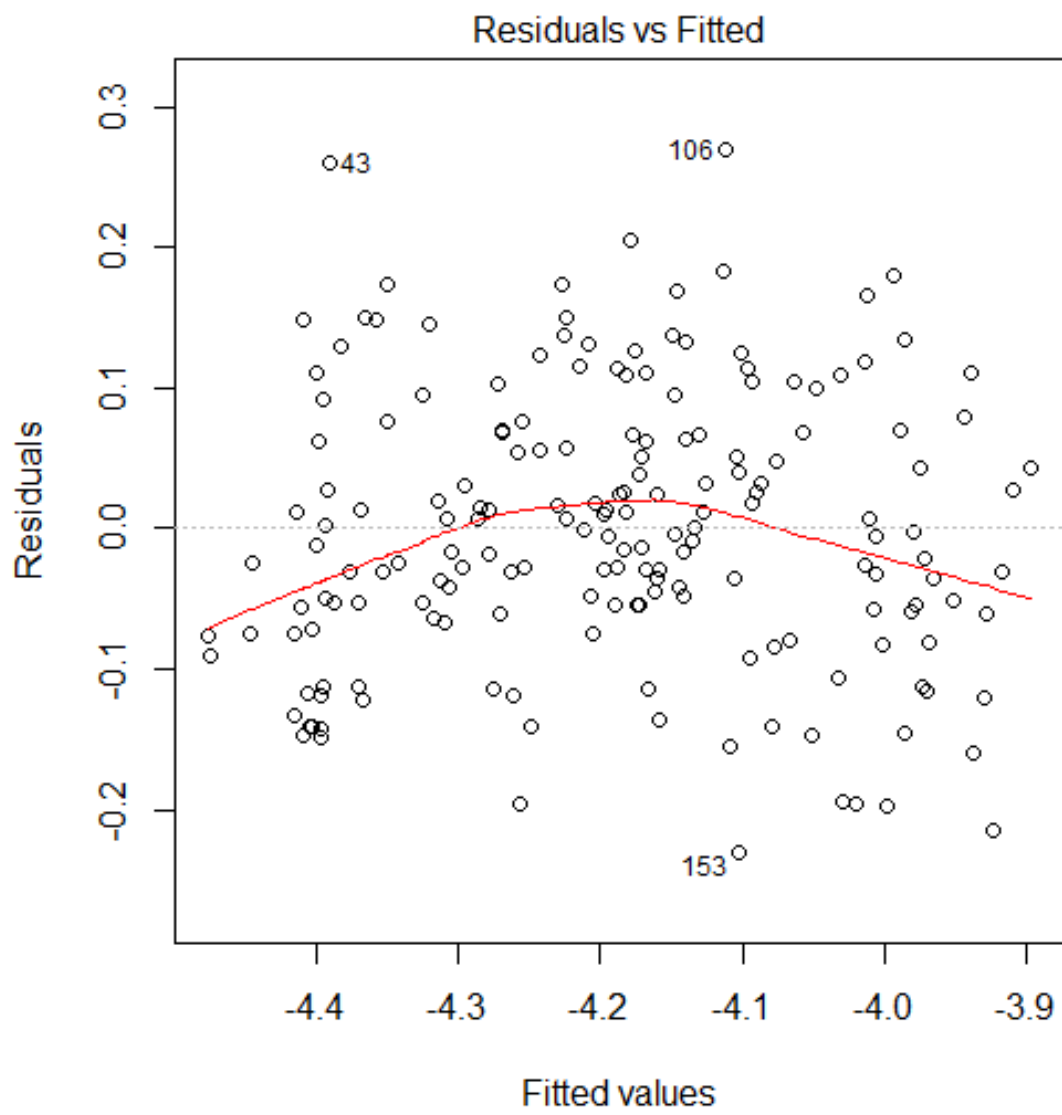
Μετά από αρκετή μελέτη και επεξεργασία των παραπάνω δεκατεσσάρων μοντέλων, έγινε περισσότερο κατανοητό ποιες μεταβλητές επηρεάζουν περισσότερο τα υπάρχοντα μοντέλα και ποιες αλληλεπιδράσεις είναι ορθότερο να αποφεύγονται. Έτσι προέκυψε το παρακάτω τελικό μοντέλο, το οποίο θεωρείται, ότι περιγράφει με τον καλύτερο δυνατό τρόπο το πρόβλημα του συσχετισμού των οδικών ατυχημάτων με παράγοντες (δείκτες) από τους κλάδους της υγείας, της οικονομίας και της κοινωνίας.

lm(formula = LOG.FAT.PER.POP ~ WORKING.LIFE + HDI + UNEMPLOYMENT + URBAN.PER.FATAL + HOSPITAL.BEDS.PER.POPULATION + HEALTH.EXPENDITURE..GDP)					
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2.569e+00	1.693e-01	-15.175	< 2e-16	***
WORKING.LIFE	-2.291e-02	3.264e-03	-7.017	4.78e-11	***
HDI	-9.255e-01	1.633e-01	-5.666	5.90e-08	***
UNEMPLOYMENT	-2.567e-05	7.027e-06	-3.653	0.000342	***
URBAN.PER.FATAL	1.435e-01	6.362e-02	2.256	0.025328	*
HOSPITAL.BEDS.PER.POPULATION	1.153e+01	4.484e+00	2.572	0.010930	*
HEALTH.EXPENDITURE..GDP	-1.525e-02	5.457e-03	-2.795	0.005765	**
Multiple R-squared: 0.6937, Adjusted R-squared: 0.6832					
F-statistic: 66.06 on 6 and 175 DF, p-value: < 2.2e-16					

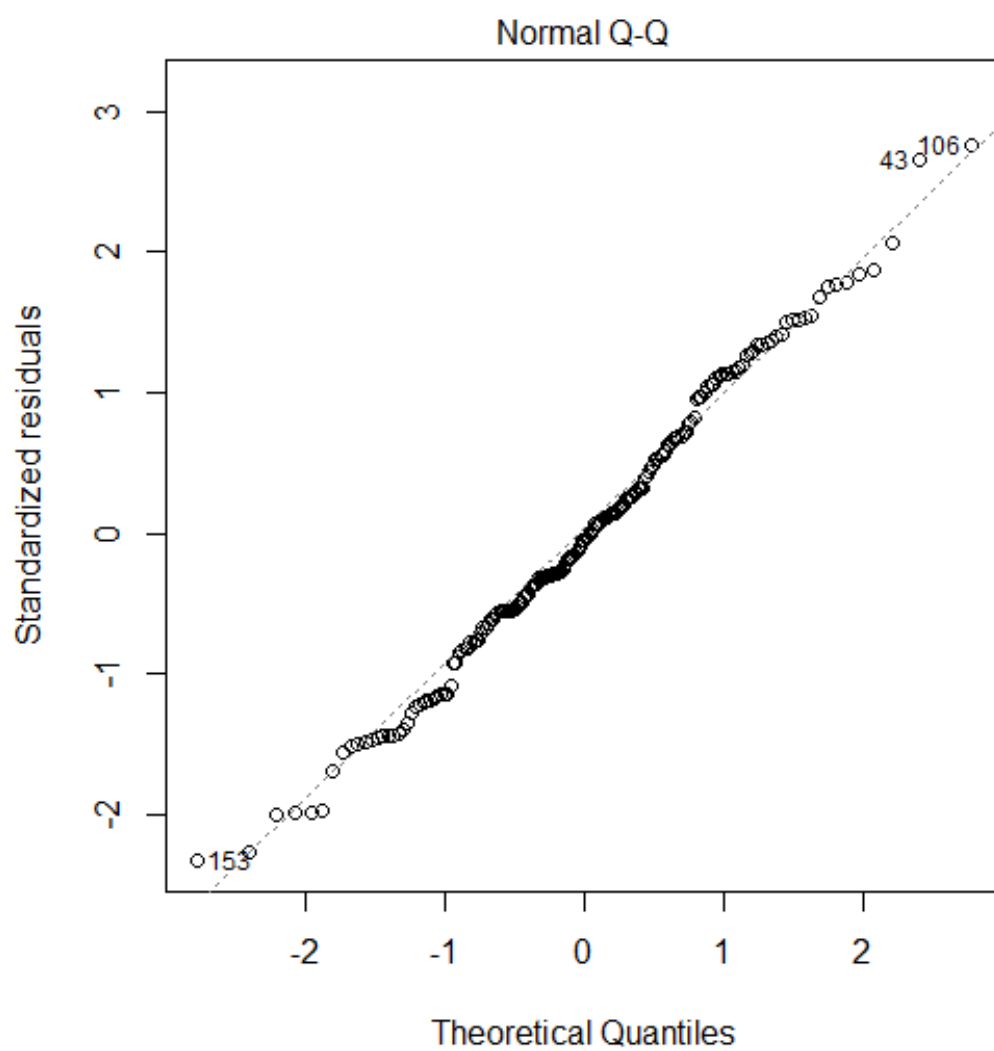
Μοντέλο 5.15

Το επιλεγθέν μοντέλο χρησιμοποιεί δείκτες και από τις τρεις κατηγορίες παραγόντων, οι οποίες επιδιώκεται να εξεταστούν στην παρούσα Διπλωματική Εργασία με ιδιαίτερη έμφαση στον τομέα της υγείας (HOSPITAL.BEDS.PER.POPULATION, HEALTH.EXPENDITURE..GDP), χωρίς, όμως, να παραλείπεται ο τομέας της οικονομίας (WORKING.LIFE, HDI), της κοινωνίας (UNEMPLOYMENT, HDI) ή των συγκοινωνιών (URBAN.PER.FATAL).

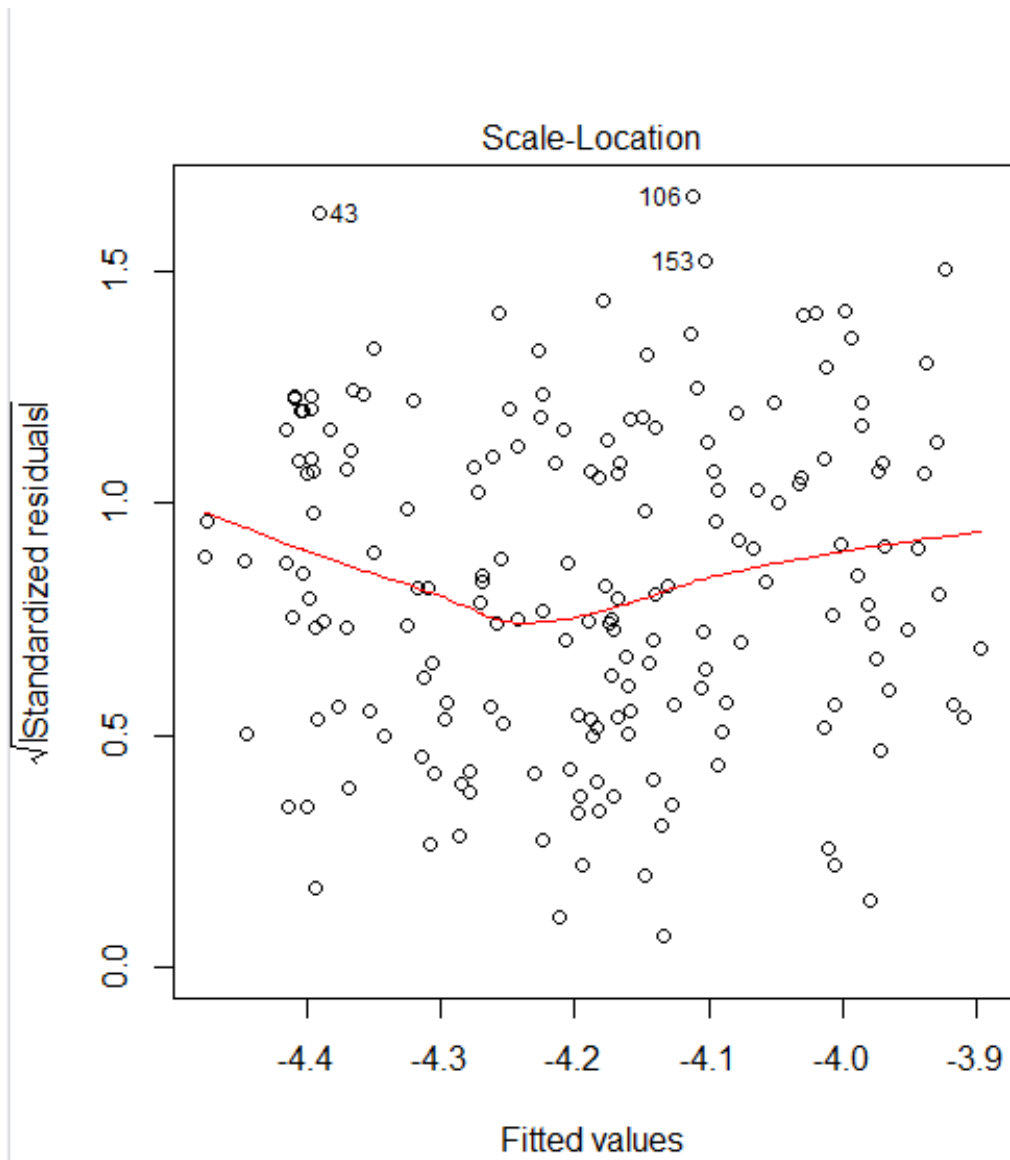
Για το τελικό αυτό μοντέλο παρατίθενται και τα σχετικά διαγράμματα της λογαριθμισμένης παλινδρόμησης.



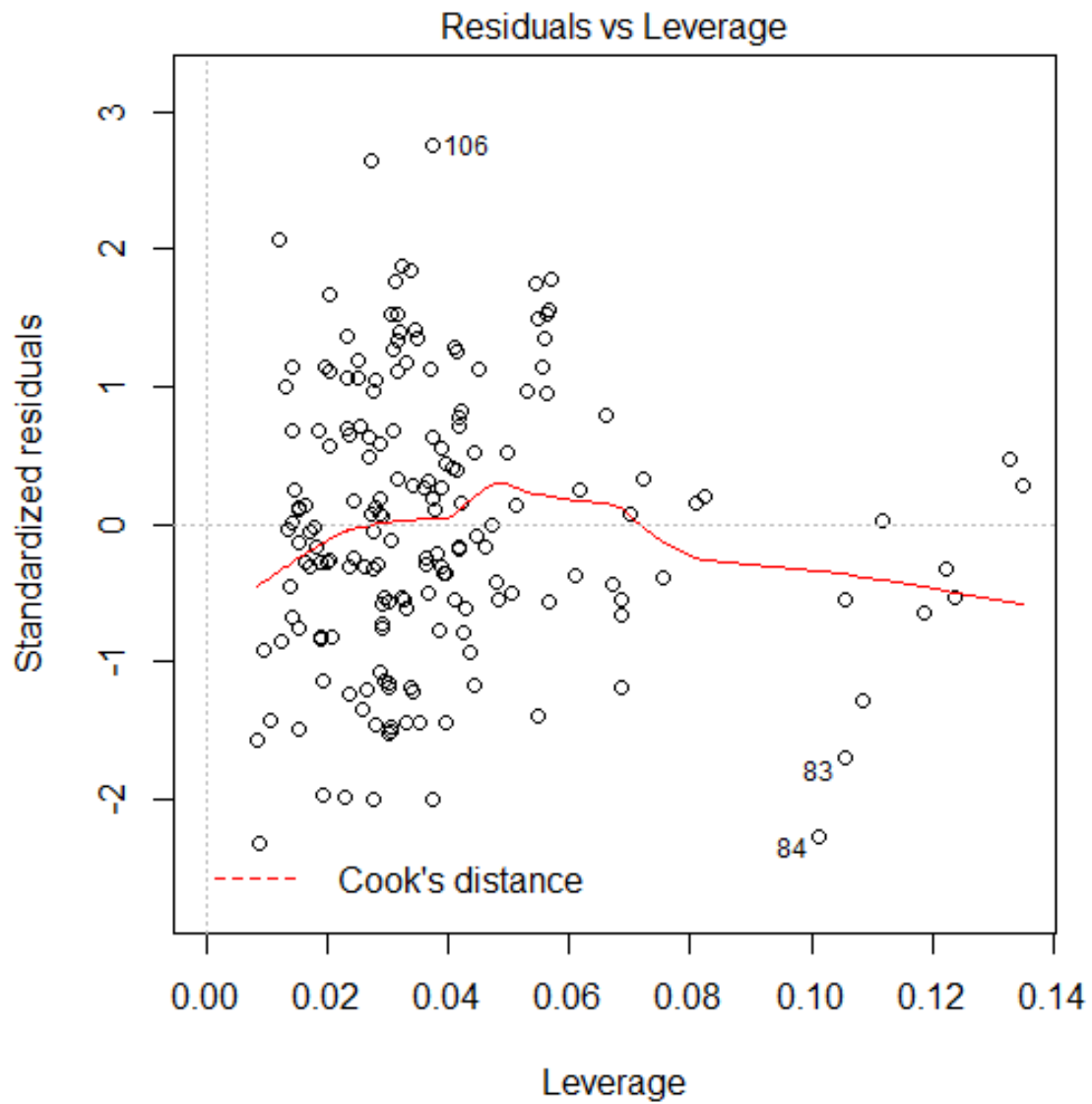
Διάγραμμα 5.2



Διάγραμμα 5.3



Διάγραμμα 5.4



Διάγραμμα 5.5

5.6 Γραμμικό Μικτό Μοντέλο

Σε συνέχεια της προηγούμενης προσπάθειας, επιδιώκεται η διαμόρφωση ενός νέου μοντέλου για κάθε μια από τις τρεις κατηγορίες κρατών, στις οποίες έχουν διαχωριστεί τα 27 κράτη-μέλη της Ε.Ε..

Ένα πλεονέκτημα της μεθόδου αυτής έναντι της λογαριθμισμένης γραμμικής παλινδρόμησης είναι ότι επιτρέπει οι όροι σφάλματος της μεταβλητής να εμφανίζουν συσχέτιση. Επιπλέον, δίνεται η δυνατότητα προσδιορισμού μιας μεταβλητής επαναλαμβανόμενων επιδράσεων (π.χ. έτος) και της δομής της συνδιακύμανσης αυτής της μεταβλητής. Τέλος, μπορεί να συμπεριληφθεί στο μοντέλο και η επίδραση του παράγοντα της ομάδας του κράτους, καθώς και να προσδιοριστεί η αλληλεπίδραση της μεταβλητής αυτής με τις άλλες ανεξάρτητες μεταβλητές.

Μετά τη στατιστική ανάλυση με τη μέθοδο της λογαριθμισμένης γραμμικής παλινδρόμησης πραγματοποιήθηκαν ορισμένες αρχικές δοκιμές με το Γραμμικό Μικτό Μοντέλο. Παρατηρήθηκε, ότι το συγκεκριμένο μαθηματικό μοντέλο χρησιμοποιεί μια ομάδα χωρών ως σημείο αναφοράς, με το οποίο συγκρίνει τις υπόλοιπες και προκύπτει εάν υπάρχει στατιστικά σημαντική διαφορά στην επιρροή τους στο μοντέλο. Παρότι έχει ήδη προκύψει ένα αποδεκτό στατιστικά μοντέλο για λόγους πληρότητας και περαιτέρω μελέτης του φαινομένου, επιλέχθηκε να γίνει η σύγκριση αυτή μεταξύ των ομάδων των κρατών. Για το λόγο αυτό, δημιουργήθηκε η μεταβλητή **cluster** των κρατών, η οποία και χρησιμοποιήθηκε στο μοντέλο. Σε όλα τα κράτη της πρώτης ομάδας και για κάθε έτος δόθηκε στη μεταβλητή η τιμή 1, στα κράτη της δεύτερης ομάδας δόθηκε η τιμή 2 και στα κράτη της τρίτης ομάδας δόθηκε η τιμή 3. Η πρώτη ομάδα κρατών επιλέχθηκε να είναι η ομάδα αναφοράς.

5.6.1 Ανάπτυξη Μαθηματικού Μοντέλου

Στην παρούσα διπλωματική εργασία ως μεταβλητή-υποκείμενο επιλέγεται η εκάστοτε ομάδα κρατών και ως επαναλαμβανόμενη μεταβλητή το έτος. Ως εξαρτημένη μεταβλητή επιλέγεται ο φυσικός λογάριθμος της θνησιμότητας ως προς τον πληθυσμό, ενώ ως παράγοντες ορίζονται, όπως και πριν, το πλήθος των ατυχημάτων εντός αστικού δικτύου προς το σύνολο των θανατηφόρων ατυχημάτων (Urban.Per.Fatal), το προσδόκιμο ζωής των ανθρώπων (Life Expectancy), οι δαπάνες για τον τομέα της υγείας ως προς το Α.Ε.Π. (Health Expenditure per GDP), οι δαπάνες για τον τομέα της υγείας ως προς το πλήθος των πολιτών (health Expenditure per Capita), οι δημόσιες δαπάνες για τον τομέα της υγείας (Public Health Expenditure), οι νοσοκομειακές κλίνες ανά πληθυσμό (Hospital.Beds.Per.Population), το κατά κεφαλήν Α.Ε.Π (GDP), ο δείκτης ανθρώπινης ανάπτυξης (HDI), τα παραγωγικά χρόνια ζωής των ανθρώπων (Working Life) και ο δείκτης ανεργίας (Unemployment). Τα πιθανά μοντέλα προκύπτουν από τους παραπάνω συνδυασμούς μεταξύ εξαρτημένης μεταβλητής και ανεξάρτητων μεταβλητών, δηλαδή συμμεταβλητών και παραγόντων.

Επόμενο βήμα είναι ο στατιστικός έλεγχος των μοντέλων με σκοπό την επιλογή εκείνου που προσαρμόζεται καλύτερα στα δεδομένα.

Στην παρούσα Διπλωματική Εργασία επιλέγεται να εξετάζεται όχι μόνο το **log Likelihood**, αλλά και το **AIC**.

Οι **έλεγχοι των σταθερών επιδράσεων** (tests of fixed effects) γίνονται με τα F-tests, όπου η τιμή σημαντικότητας των συντελεστών αυτών θα πρέπει να είναι $\text{sig.} < 0,5$.

Μετά τον έλεγχο των μοντέλων βάσει των ανωτέρω κριτηρίων, ο έλεγχος συνεχίζεται, εξετάζοντας το μέγεθος επιρροής του σταθερού όρου. Φυσικά, πρέπει να εξεταστεί, εάν τα αποτελέσματα του μοντέλου οδηγούν σε λογικά συμπεράσματα και εάν μπορούν να ερμηνευτούν με βάση τις επικρατούσες συνθήκες και αντιλήψεις.

5.6.2 Προετοιμασία Εφαρμογής Γραμμικού Μικτού Μοντέλου

Όπως και προηγουμένως, τα μοντέλα, τα οποία διαμορφώνονται, βασίζονται στον πίνακα αυτοσυσχέτισης. Έχοντας, ήδη, διαμορφώσει ένα ορθό μοντέλο περιγραφής του φαινομένου μέσω παλινδρόμησης, είναι γνωστές οι μεταβλητές, οι οποίες έχουν καθοριστικό ρόλο. Συνεπώς, οι μεταβλητές αυτές αναμένεται να αξιοποιηθούν και στα νέα μοντέλα μικτού τύπου.

Μετά τις πρώτες προσπάθειες σχηματισμού στατιστικών μοντέλων μικτού τύπου, παρατηρείται μεγάλη απόκλιση μεταξύ αναμενόμενων και τελικών αποτελεσμάτων, κάτι το οποίο φαίνεται να οφείλεται στη διαφορά στην τάξη μεγέθους, την οποία έχουν οι χρησιμοποιούμενες μεταβλητές.

Για την αντιμετώπιση του προβλήματος αυτού διαμορφώθηκε μια νέα βάση δεδομένων, βασισμένη στη βάση δεδομένων No. 2, στην οποία εφαρμόστηκε μια συνάρτηση κανονικοποίησης των τιμών των μεταβλητών. Το αποτέλεσμα φαίνεται παρακάτω.

NAME	CLUSTER	COUNTRY	YEAR	POPULATION	GDP	LOG.FAT. PER.POP	URBAN PER FATAL	HDI	WORKING LIFE	PRODUCT IVITY	LIFE EXPECTANCY	UNEMPLOYMENT	HOSPITAL BEDS PER POPULATION	PUBLIC HEALTH EXPENDITURE	HEALTH EXPENDITURE PER CAPITA	HEALTH EXPENDITURE %GDP
AT	1	1	2008	8307989	51630	0.58	-0.75	0.30	0.37	0.78	0.50	-0.60	1.43	0.32	1.02	1.07
AT	1	1	2009	8335003	47786	0.40	-0.75	0.33	0.44	0.80	0.47	-0.55	1.42	0.34	0.96	1.41
AT	1	1	2010	8351643	46757	0.06	-0.94	0.47	0.47	0.71	0.54	-0.57	1.41	0.30	0.91	1.40
AT	1	1	2011	8375164	51193	-0.08	-0.88	0.54	0.54	0.72	0.68	-0.58	1.42	0.29	1.09	1.26
AT	1	1	2012	8408121	48381	-0.05	-0.75	0.59	0.64	0.79	0.68	-0.57	1.42	0.31	1.00	1.40
AT	1	1	2013	8451860	50533	-0.44	-1.00	0.68	0.71	0.77	0.75	-0.55	1.41	0.29	1.11	1.38
AT	1	1	2014	8507786	51390	-0.59	-0.76	0.68	0.68	0.75	0.85	-0.54	1.38	0.57	1.16	1.42
BE	1	2	2008	10666866	48851	0.78	-0.65	0.40	-0.86	1.17	0.22	-0.46	0.85	0.48	0.68	0.48
BE	1	2	2009	10753080	45176	0.76	-0.82	0.44	-0.86	1.22	0.36	-0.42	0.81	0.52	0.68	0.95
BE	1	2	2010	10839905	44691	0.45	-0.66	0.54	-0.73	1.28	0.40	-0.40	0.78	0.55	0.61	0.82
BE	1	2	2011	11000638	47951	0.48	-0.39	0.57	-0.86	1.22	0.54	-0.45	0.73	0.46	0.82	0.97
BE	1	2	2012	11075889	44900	0.19	-0.78	0.62	-0.83	1.26	0.47	-0.43	0.70	0.55	0.69	1.03
BE	1	2	2013	11137974	46613	0.02	-1.00	0.64	-0.73	1.24	0.54	-0.39	0.68	0.56	0.79	1.05
BE	1	2	2014	11180840	47635	0.02	-0.93	0.73	-0.69	1.25	0.78	-0.38	0.67	0.57	0.83	1.07
BG	2	3	2008	7518002	7153	1.92	0.48	-1.44	-0.73	-1.94	-2.07	-0.57	0.84	-1.41	-1.28	-1.26
BG	2	3	2009	7467119	6860	1.54	0.06	-1.41	-0.86	-1.94	-1.93	-0.54	0.90	-1.73	-1.29	-1.16
BG	2	3	2010	7421766	6744	1.19	0.78	-1.33	-1.03	-1.94	-1.90	-0.44	0.84	-1.70	-1.28	-0.89
BG	2	3	2011	7369431	7837	0.79	0.72	-1.27	-1.20	-1.91	-1.76	-0.42	0.76	-1.81	-1.26	-1.10
BG	2	3	2012	7327224	7402	0.59	1.29	-1.22	-1.03	-1.86	-1.69	-0.40	0.84	-1.64	-1.26	-0.97
BG	2	3	2013	7284552	7703	0.61	1.18	-1.12	-0.90	-1.88	-1.51	-0.37	0.95	-2.08	-1.22	-0.49
BG	2	3	2014	7245677	7877	0.85	0.93	-1.03	-0.90	-1.85	-1.65	-0.42	1.11	-1.81	-1.20	-0.19
HR	2	4	2008	4311967	15889	2.13	2.10	-0.85	-1.00	-0.90	-1.12	-0.60	0.34	1.28	-0.91	-0.62
HR	2	4	2009	4309796	14142	1.66	1.63	-0.85	-0.97	-0.97	-1.02	-0.59	0.28	1.37	-0.94	-0.34
HR	2	4	2010	4302847	13505	1.05	1.76	-0.76	-1.03	-1.01	-0.88	-0.55	0.41	1.36	-0.96	-0.30
HR	2	4	2011	4289857	14538	1.01	1.73	-0.64	-1.10	-0.91	-0.70	-0.53	0.45	0.81	-0.97	-0.56
HR	2	4	2012	4275984	13235	0.87	1.68	-0.61	-1.17	-0.83	-0.66	-0.50	0.47	0.95	-1.02	-0.57
HR	2	4	2013	4262140	13574	0.72	1.69	-0.56	-1.20	-0.79	-0.49	-0.47	0.45	0.97	-1.00	-0.55
HR	2	4	2014	4246809	13469	0.29	1.85	-0.50	-0.79	-0.89	-0.45	-0.47	0.48	0.98	-1.01	-0.56
CY	2	5	2008	776333	19006	1.21	2.19	-0.06	0.61	-0.02	0.50	-0.73	-0.60	-3.16	-0.44	-1.09
CY	2	5	2009	796930	18674	0.79	1.74	0.01	0.74	-0.01	0.64	-0.72	-0.62	-2.82	-0.46	-0.80
CY	2	5	2010	819140	19300	0.31	2.39	-0.09	0.78	-0.15	0.82	-0.72	-0.72	-2.55	-0.54	-0.89
CY	2	5	2011	839751	19731	0.67	1.32	-0.04	0.68	-0.20	0.71	-0.71	-0.75	-2.63	-0.44	-0.72
CY	2	5	2012	862011	19490	-0.21	1.40	-0.04	0.57	-0.25	0.68	-0.70	-0.80	-2.70	-0.57	-0.77
CY	2	5	2013	865878	18140	-0.58	2.27	-0.04	0.57	-0.31	1.17	-0.68	-0.84	-2.63	-0.61	-0.76
CY	2	5	2014	858000	17606	-0.50	2.55	0.03	0.74	-0.38	1.10	-0.68	-0.84	-2.77	-0.64	-0.82
CZ	2	6	2008	10343422	22819	1.17	0.31	0.09	-0.32	-0.64	-0.66	-0.55	1.19	1.17	-0.80	-1.14
CZ	2	6	2009	10425783	19788	0.72	-0.13	0.11	-0.25	-0.59	-0.63	-0.44	1.14	1.17	-0.80	-0.54
CZ	2	6	2010	10462088	19831	0.43	-0.11	0.15	-0.25	-0.66	-0.52	-0.42	1.08	1.21	-0.84	-0.78
CZ	2	6	2011	10486731	21737	0.33	-0.11	0.20	-0.25	-0.65	-0.42	-0.45	0.97	1.20	-0.77	-0.74
CZ	2	6	2012	10505445	19740	0.22	-0.16	0.21	-0.11	-0.69	-0.38	-0.43	0.88	1.23	-0.84	-0.71

Πίνακας 5.2

Παρά την κανονικοποίηση των μεταβλητών, ο πίνακας αυτοσυσχέτισης εξακολουθεί να είναι εφαρμόσιμος και σε αυτόν θα βασιστούν τα μοντέλα στη συνέχεια.

Στο σημείο αυτό κρίνεται σκόπιμο να τονιστεί, ότι οι μεταβλητές CLUSTER, COUNTRY και YEAR ορίζονται στο σύστημα ως μεταβλητές διακριτής τιμής (factors) προκειμένου το σύστημα να γνωρίζει, ότι οι μεταβλητές αυτές παίρνουν διακριτές τιμές εντός ενός συνόλου, το οποίο αναγνωρίζει αυτόματα το σύστημα και δεν είναι συνεχείς όπως οι υπόλοιπες.

5.6.3 Αποτελέσματα Γραμμικού Μικτού Μοντέλου

Στο παρόν κομμάτι της Διπλωματικής Εργασίας πραγματοποιείται έρευνα και επεξεργασία των δεδομένων με σκοπό την εύρεση του σωστότερου συνδυασμού μεταβλητών (fixed effect και random effect), ώστε να προκύψει το καταλληλότερο μοντέλο περιγραφής του φαινομένου.

Και αυτή τη φορά διαμορφώθηκαν πολλά μοντέλα και μελετήθηκαν αρκετοί διαφορετικοί συνδυασμοί, εκ των οποίων οι σημαντικότεροι θα παρουσιαστούν και θα αναλυθούν παρακάτω.

Αρχικά, παρατίθενται μοντέλα με μόνο μια μεταβλητή για να φανεί ο τρόπος λειτουργίας του χρησιμοποιηθέντος μοντέλου, καθώς και η σημασία της κάθε μεταβλητής.

Formula: LOG.FAT.PER.POP ~ (PUBLIC.HEALTH.EXPENDITURE + 1 CLUSTER)-1					
REML criterion at convergence: 446.3					
Scaled residuals:					
Min	1Q	Median	3Q	Max	
-2.27703	-0.76748	0.02843	0.66151	2.99407	
Random effects:					
Groups	Name	Variance	Std.Dev.	Corr	
CLUSTER	(Intercept)	0.63544	0.7971		
	PUBLIC.HEALTH.EXPENDITURE	0.02805	0.1675	1.00	
	Residual	0.63591	0.7974		
Number of obs: 182, groups: CLUSTER, 3					
Fixed effects:					
	Estimate	Std. Error	t value		
(Intercept)	0.3181	0.2497	1.274		

Μοντέλο 5.16

Όπως φαίνεται και από το παραπάνω μοντέλο, πραγματοποιείται μια προσπάθεια διαμόρφωσης **τριών διαφορετικών συναρτήσεων, μια για κάθε ομάδα χωρών (cluster)**, όπου οι μεταβλητές με **σταθερή επίδραση (fixed effects)**, στη συγκεκριμένη περίπτωση μόνο το intercept, θα έχουν **σταθερό τρόπο επηρεασμού** του μοντέλου, ενώ οι μεταβλητές με **μεταβλητή επίδραση (random effects)**, εδώ οι δημόσιες δαπάνες για την υγεία, θα επηρεάζουν **με διαφορετικό τρόπο το κάθε μοντέλο**. Όπως φαίνεται, η πρώτη και η τρίτη ομάδα χωρών επηρεάζονται με τον ίδιο σχεδόν τρόπο, εφόσον ο συντελεστής της μεταβλητής των δημοσίων δαπανών είναι στη μία περίπτωση 0.63544 και στην άλλη 0.63591. Αντιθέτως, στη δεύτερη κατηγορία ο συντελεστής έχει τιμή 0.02805, γεγονός που υποδηλώνει, ότι σε αυτές τις χώρες οι δημόσιες δαπάνες για την υγεία ενδεχομένως και να μην επηρεάζουν τόσο έντονα τη θνησιμότητα στα οδικά ατυχήματα ανά πληθυσμό.

Μέσα από το άνωθεν μοντέλο καθίσταται κατανοητή η ανάγκη διαμόρφωσης γραμμικών μοντέλων μικτού τύπου, η οποία δεν είναι άλλη από την καλύτερη περιγραφή του φαινομένου. Ενώ στην λογαριθμισμένη παλινδρόμηση εξάγεται μία μόνο συνάρτηση περιγραφής του φαινομένου για όλες τις χώρες, συνεπώς για το μέσο όρο των τιμών όλων των χωρών, με αποτέλεσμα να προκύπτει ένα πιο γενικό μοντέλο προκειμένου να μπορεί να ικανοποιήσει εντός ενός διαστήματος τις περισσότερες χώρες, εδώ δημιουργούνται τρία μοντέλα, τα οποία πάλι βασίζονται σε μέσους όρους, αλλά εφόσον το δείγμα έχει ήδη κατηγοριοποιηθεί και από ένα σύνολο πλέον υπάρχουν τρία διαφορετικά σύνολα με κοινά χαρακτηριστικά εντός του κάθε συνόλου, οι μέσοι όροι και τα αποτελέσματα αναμένεται να είναι πιο εύστοχα και να ικανοποιούν καλύτερα την κάθε κατηγορία.

Στη συνέχεια, επιχειρείται μια προσπάθεια ανάδειξης της σημασίας του διαχωρισμού του συνόλου των χωρών σε τρεις ομάδες και της ορθής επιλογής των ομάδων αυτών μέσω μοντέλων μικτού τύπου μίας μόνο μεταβλητής.

```

Formula: LOG.FAT.PER.POP ~ (HEALTH.EXPENDITURE..GDP + 1 | CLUSTER) - 1
REML criterion at convergence: 425.6

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.11507 -0.80223 -0.03142  0.67158  2.09087

Random effects:
  Groups   Name                Variance Std.Dev. Corr
  CLUSTER (Intercept)          0.4061   0.6373
          HEALTH.EXPENDITURE..GDP 0.1441   0.3796  -0.43
  Residual                    0.5541   0.7444
Number of obs: 182, groups: CLUSTER, 3

Fixed effects:
              Estimate Std. Error t value
(Intercept)  -0.3176     0.3406  -0.932

```

Μοντέλο 5.17

Και πάλι παρατηρείται έντονη ομοιότητα μεταξύ της πρώτης ομάδας και της τρίτης ομάδας.

```

Formula: LOG.FAT.PER.POP ~ (HEALTH.EXPENDITURE.PER.CAPITA + 1 | CLUSTER) - 1
REML criterion at convergence: 404.7

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.10363 -0.77637 -0.05225  0.72258  2.26795

Random effects:
  Groups   Name                Variance Std.Dev. Corr
  CLUSTER (Intercept)          0.4084   0.6391
          HEALTH.EXPENDITURE.PER.CAPITA 0.3019   0.5495  -0.50
  Residual                    0.4883   0.6988
Number of obs: 182, groups: CLUSTER, 3

Fixed effects:
              Estimate Std. Error t value
(Intercept)  -0.2085     0.3271  -0.637

```

Μοντέλο 5.18

Στο άνωθεν μοντέλο φαίνεται, πως ο παράγοντας των δαπανών για την υγεία ανά κάτοικο τείνει να επηρεάσει με παρόμοιο τρόπο και τις τρεις ομάδες χωρών, εφόσον οι συντελεστές εμφανίζουν μεγάλη ομοιότητα, όπου και πάλι, όμως, η πρώτη και η τρίτη ομάδα εμφανίζουν αρκετά κοινή συμπεριφορά.

Formula: LOG.FAT.PER.POP ~ (PRODUCTIVITY + 1 | CLUSTER) - 1

REML criterion at convergence: 419.4

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.21012	-0.65465	0.03143	0.64760	2.28652

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.9128	0.9554	
	PRODUCTIVITY	1.6662	1.2908	-0.89
	Residual	0.5226	0.7229	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-0.1458	0.2655	-0.549

Μοντέλο 5.19

Στο άνωθεν μοντέλο παρατηρείται, ότι και οι τρεις κατηγορίες επηρεάζονται σε διαφορετικό βαθμό από τον παράγοντα της παραγωγικότητας, μια παρατήρηση η οποία δικαιολογεί και τον διαχωρισμό στις τρεις αυτές διαφορετικές κατηγορίες.

Formula: LOG.FAT.PER.POP ~ (LIFE.EXPECTANCY + 1 | CLUSTER) - 1

REML criterion at convergence: 381.8

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.27594	-0.69216	-0.04075	0.67451	2.42454

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.5925	0.7698	
	LIFE.EXPECTANCY	0.7609	0.8723	-0.58
	Residual	0.4233	0.6507	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-0.4178	0.3672	-1.138

Μοντέλο 5.20

Και πάλι παρατηρούνται ομοιότητες μεταξύ πρώτης και τρίτης κατηγορίας.

Formula: LOG.FAT.PER.POP ~ (WORKING.LIFE + 1 | CLUSTER) - 1

REML criterion at convergence: 356

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.51129	-0.79633	0.03386	0.63236	2.97170

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.7281	0.8533	
	WORKING.LIFE	0.3225	0.5679	0.86
	Residual	0.3680	0.6066	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	0.5997	0.2621	2.288

Μοντέλο 5.21

Στο συγκεκριμένο μοντέλο παρατηρείται ομοιότητα στο βαθμό επίδρασης του παράγοντα των παραγωγικών χρόνων εργασίας μεταξύ της δεύτερης και της τρίτης ομάδας.

Formula: LOG.FAT.PER.POP ~ (HDI + 1 | CLUSTER) - 1

REML criterion at convergence: 343

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.20775	-0.76258	-0.07331	0.68467	2.51389

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.1385	0.3721	
	HDI	2.0988	1.4487	-0.22
	Residual	0.3393	0.5825	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	0.3108	0.2224	1.398

Μοντέλο 5.22

Το συγκεκριμένο μοντέλο παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς υπάρχει πολύ μεγάλη ομοιότητα μεταξύ πρώτης και τρίτης κατηγορίας, οι οποίες όμως ταυτόχρονα εμφανίζουν πολύ μεγάλη απόκλιση με τη δεύτερη κατηγορία. Είναι γνωστό, πως όσο πιο κοντά είναι η τιμή της απόκλισης στο μηδέν, τόσο περισσότερο επηρεάζει ο παράγοντας αυτός το φαινόμενο και συνεπώς το μοντέλο καθίσταται ορθότερο. Στη συγκεκριμένη περίπτωση, τείνει να προκύψει το συμπέρασμα, ότι ενώ για την πρώτη και για τη δεύτερη κατηγορία το HDI είναι ένας πολύ καθοριστικός παράγοντας, κάτι τέτοιο δεν ισχύει για τη δεύτερη κατηγορία στην οποία το HDI εμφανίζεται να μην μπορεί να προβλέψει την εξέλιξη το φαινομένου και συνεπώς δεν επιδρά τόσο καθοριστικά σε αυτό.

```

Formula: LOG.FAT.PER.POP ~ (HOSPITAL.BEDS.PER.POPULATION + 1 | CLUSTER) - 1
REML criterion at convergence: 394.9

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.10951 -0.77204  0.01298  0.64667  2.64306

Random effects:
Groups   Name                    Variance Std.Dev. Corr
CLUSTER (Intercept)                0.4403   0.6636
        HOSPITAL.BEDS.PER.POPULATION 0.2357   0.4855   0.06
Residual                                0.4569   0.6760
Number of obs: 182, groups: CLUSTER, 3

Fixed effects:
              Estimate Std. Error t value
(Intercept)  -0.2702     0.3868  -0.698

```

Μοντέλο 5.23

Στο συγκεκριμένο μοντέλο παρατηρείται και πάλι μεγάλη ομοιότητα στο βαθμό επηρεασμού των ομάδων ένα και τρία από τον παράγοντα των νοσοκομειακών κλινών ανά πληθυσμό, καθώς επίσης και γενικά μια έντονη συσχέτιση και των τριών ομάδων με το συγκεκριμένο δείκτη, κάτι το οποίο θα αξιοποιηθεί στα επόμενα πιο πολύπλοκα μοντέλα.

```

Formula: LOG.FAT.PER.POP ~ (URBAN.PER.FATAL + 1 | CLUSTER) - 1
REML criterion at convergence: 432.9

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.48110 -0.67639  0.07778  0.76797  2.34626

Random effects:
Groups   Name                    Variance Std.Dev. Corr
CLUSTER (Intercept)                0.9001   0.9487
        URBAN.PER.FATAL 0.2394   0.4893  -1.00
Residual                                0.5796   0.7613
Number of obs: 182, groups: CLUSTER, 3

Fixed effects:
              Estimate Std. Error t value
(Intercept)   0.8117     0.1226   6.621

```

Μοντέλο 5.24

Σο παρόν μοντέλο παρατηρείται μεγάλη απόκλιση στο βαθμό επηρεασμού των τριών ομάδων από τον παράγοντα των οδικών ατυχημάτων εντός αστικού ιστού ως προς το σύνολο των θανατηφόρων ατυχημάτων.

Formula: LOG.FAT.PER.POP ~ (UNEMPLOYMENT + 1 | CLUSTER) - 1

REML criterion at convergence: 433.1

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.23292	-0.70619	0.05571	0.71897	2.30311

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.3037	0.5510	
	UNEMPLOYMENT	0.5702	0.7551	-0.02
Residual		0.5749	0.7582	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	0.07488	0.33702	0.222

Μοντέλο 5.25

Σο συγκεκριμένο μοντέλο παρατηρείται έντονη ομοιότητα στο βαθμό επηρεασμού της δεύτερης και τρίτης ομάδας χωρών από τον παράγοντα της ανεργίας.

Συνολικά, παρατηρείται, ότι οι μεταβλητές, οι οποίες ατομικά μπορούν να εκφράσουν καλύτερα το φαινόμενο το οποίο μελετάται, είναι οι HEALTH.EXPENDITURE.PER.CAPITA, WORKING.LIFE, UNEMPLOYMENT και HOSPITAL.BEDS.PER.POPULATION. Παρότι, οι μεταβλητές αυτές δύναται ατομικά να διαμορφώσουν ένα αρκετά καλό μοντέλο, παρακάτω γίνεται προσπάθεια διαμόρφωσης ενός σύνθετου περιγραφικού μοντέλου, στο οποίο θα υπάρξει πλήθος μεταβλητών σταθερής και μεταβλητής επίδρασης. Για την ορθή και τεκμηριωμένη διαμόρφωση τέτοιων μοντέλων οφείλουν να γίνουν κάποια ενδεικτικά μοντέλα στην αρχή, καθώς και να μελετηθούν οι αλληλεπιδράσεις των παραγόντων, όπως αυτές προκύπτουν από τον πίνακα αλληλεπίδρασης στην αρχή του παρόντος κεφαλαίου.

Στη συνέχεια, λοιπόν, μελετώνται μοντέλα με περισσότερες μεταβλητές, όπου όλες οι μεταβλητές πλην μιας ασκούν σταθερή επίδραση, ενώ μόνο μια μεταβλητή έχει μεταβλητή επίδραση, διαχωρίζοντας έτσι τα τρία μοντέλα.

Σύμφωνα με τις παραπάνω παρατηρήσεις, καθώς και το διάγραμμα αλληλεπίδρασης, επιλέγεται στην παρακάτω ανάλυση να μελετηθούν οι παράμετροι WORKING.LIFE, HDI, HOSPITAL.BEDS.PER.POPULATION, URBAN.PER.FATAL και UNEMPLOYMENT με σκοπό να βρεθεί ο κατάλληλος συνδυασμός τους.

Παράλληλα, στα νέα αυτά μοντέλα εισάγεται και η μεταβλητή του χρόνου μέσω της μεταβλητής YEAR. Η προσθήκη αυτή γίνεται, καθώς υπάρχει επαναληψιμότητα των μετρήσεων των δεδομένων για τα έτη 2008-2014 και κρίνεται σκόπιμο να εισαχθεί και αυτή η παράμετρος ως παράγοντας (factor).

Formula: $y_{01} = \text{LOG.FAT.PER.POP} \sim (\text{WORKING.LIFE} + 1 \mid \text{CLUSTER}) + \text{YEAR} - 1 + \text{HDI} + \text{HOSPITAL.BEDS.PER.POPULATION} + \text{URBAN.PER.FATAL} + \text{UNEMPLOYMENT}$

REML criterion at convergence: 231.2

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.39314	-0.71350	-0.02258	0.64967	2.82778

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.1104	0.3322	
	WORKING.LIFE	0.2227	0.4719	0.92
Residual		0.1615	0.4018	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
YEAR2008	0.90021	0.11273	7.986
YEAR2009	0.60171	0.11327	5.312
YEAR2010	0.27907	0.11345	2.460
YEAR2011	0.24752	0.11356	2.180
YEAR2012	0.06543	0.11474	0.570
YEAR2013	-0.04730	0.11550	-0.410
YEAR2014	-0.05760	0.11563	-0.498
HDI	-0.27232	0.04157	-6.551
HOSPITAL.BEDS.PER.POPULATION	0.17507	0.03679	4.759
URBAN.PER.FATAL	0.10004	0.03713	2.694
UNEMPLOYMENT	-0.08188	0.05337	-1.534

Correlation of Fixed Effects:

	YEAR2008	YEAR2009	YEAR2010	YEAR2011	YEAR2012	YEAR2013	YEAR2014	HDI	HOSPIT
URBAN.									
YEAR2009	0.510								
YEAR2010	0.508	0.516							
YEAR2011	0.508	0.515	0.518						
YEAR2012	0.509	0.518	0.521	0.522					
YEAR2013	0.508	0.517	0.522	0.523	0.530				
YEAR2014	0.509	0.516	0.521	0.522	0.529	0.534			
HDI	0.029	0.036	0.009	-0.010	-0.019	-0.052	-0.062		
HOSPITA.BE	-0.074	-0.061	-0.045	-0.033	-0.042	-0.042	-0.042	0.086	
URBAN.P.FA	-0.023	0.025	0.019	0.012	0.010	-0.022	-0.038	0.350	
UNEMPLOYMEN	-0.070	-0.131	-0.152	-0.154	-0.188	-0.207	-0.190	0.160	
0.251	0.007								

Μοντέλο 5.26

Formula: $y_{02} = \text{LOG.FAT.PER.POP} \sim (\text{HDI} + 1 \mid \text{CLUSTER}) + \text{YEAR} - 1 + \text{WORKING.LIFE} + \text{HOSPITAL.BEDS.PER.POPULATION} + \text{URBAN.PER.FATAL} + \text{UNEMPLOYMENT}$

REML criterion at convergence: 252.7

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.4527	-0.6510	0.0454	0.6573	2.6303

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.0521	0.2283	
	HDI	0.8928	0.9449	-0.81
	Residual	0.1832	0.4280	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
YEAR2008	0.68658	0.11807	5.815
YEAR2009	0.40059	0.11871	3.375
YEAR2010	0.10762	0.12025	0.895
YEAR2011	0.07712	0.12088	0.638
YEAR2012	-0.07906	0.12215	-0.647
YEAR2013	-0.16672	0.12370	-1.348
YEAR2014	-0.15204	0.12439	-1.222
WORKING.LIFE	-0.31960	0.04156	-7.690
HOSPITAL.BEDS.PER.POPULATION	0.01305	0.04232	0.308
URBAN.PER.FATAL	0.10951	0.03959	2.766
UNEMPLOYMENT	-0.19873	0.05624	-3.534

Correlation of Fixed Effects:

	YEAR2008	YEAR2009	YEAR2010	YEAR2011	YEAR2012	YEAR2013	YEAR2014	WORKIN
HOSPIT URBAN.								
YEAR2009	0.494							
YEAR2010	0.497	0.505						
YEAR2011	0.497	0.505	0.515					
YEAR2012	0.497	0.507	0.518	0.522				
YEAR2013	0.497	0.507	0.520	0.525	0.532			
YEAR2014	0.498	0.506	0.520	0.525	0.533	0.541		
WORKING.LIF	0.039	0.041	0.060	0.072	0.053	0.053	0.039	
HOSPITAL.BE	-0.133	-0.115	-0.122	-0.123	-0.136	-0.150	-0.162	0.110
URBAN.PER.F	0.007	0.053	0.043	0.036	0.029	-0.006	-0.023	0.031
UNEMPLOYMEN	-0.066	-0.125	-0.156	-0.165	-0.202	-0.227	-0.211	0.017
	0.218	0.078						

Μοντέλο 5.27

Formula: $y_{03} = \text{LOG.FAT.PER.POP} \sim (\text{HOSPITAL.BEDS.PER.POPULATION} + 1 \mid \text{CLUSTER})$
 $+ \text{YEAR} - 1 + \text{WORKING.LIFE} + \text{HDI} + \text{URBAN.PER.FATAL} + \text{UNEMPLOYMENT}$

REML criterion at convergence: 256.3

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.3111	-0.5606	0.1255	0.6436	2.7646

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.1180	0.3435	
	HOSPITAL.BEDS.PER.POPULATION	0.0655	0.2559	0.03
Residual		0.1885	0.4342	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
YEAR2008	0.50495	0.21722	2.325
YEAR2009	0.21924	0.21755	1.008
YEAR2010	-0.11170	0.21773	-0.513
YEAR2011	-0.15485	0.21781	-0.711
YEAR2012	-0.33292	0.21853	-1.523
YEAR2013	-0.44815	0.21902	-2.046
YEAR2014	-0.44083	0.21892	-2.014
WORKING.LIFE	-0.34718	0.04526	-7.671
HDI	-0.21832	0.04659	-4.686
URBAN.PER.FATAL	0.11361	0.04052	2.803
UNEMPLOYMENT	-0.06113	0.06700	-0.912

Correlation of Fixed Effects:

	YEAR2008	YEAR2009	YEAR2010	YEAR2011	YEAR2012	YEAR2013	YEAR2014	WORKIN
HDI								
URBAN.								
YEAR2009	0.845							
YEAR2010	0.843	0.846						
YEAR2011	0.843	0.846	0.847					
YEAR2012	0.841	0.846	0.847	0.847				
YEAR2013	0.839	0.844	0.846	0.847	0.848			
YEAR2014	0.840	0.844	0.846	0.846	0.848	0.849		
WORKING.LIF	-0.050	-0.042	-0.031	-0.025	-0.040	-0.043	-0.054	
HDI	0.000	-0.005	-0.025	-0.037	-0.045	-0.065	-0.069	
-0.254								
URBAN.PER.F	-0.006	0.023	0.024	0.022	0.021	0.004	-0.007	0.107
0.270								
UNEMPLOYMEN	-0.029	-0.075	-0.094	-0.099	-0.124	-0.139	-0.128	0.064
0.268	-0.019							

Μοντέλο 5.28

Formula: $y_{04} = \text{LOG.FAT.PER.POP} \sim (\text{URBAN.PER.FATAL} + 1 \mid \text{CLUSTER}) + \text{YEAR} - 1 + \text{WORKING.LIFE} + \text{HDI} + \text{HOSPITAL.BEDS.PER.POPULATION} + \text{UNEMPLOYMENT}$

REML criterion at convergence: 275.7

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.36255	-0.65521	0.04391	0.68740	2.60966

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.03420	0.1849	
	URBAN.PER.FATAL	0.07802	0.2793	0.74
Residual		0.21675	0.4656	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
YEAR2008	0.52159	0.11905	4.381
YEAR2009	0.24989	0.11887	2.102
YEAR2010	-0.08755	0.11936	-0.733
YEAR2011	-0.13919	0.11939	-1.166
YEAR2012	-0.29892	0.12025	-2.486
YEAR2013	-0.39512	0.12251	-3.225
YEAR2014	-0.39457	0.12234	-3.225
WORKING.LIFE	-0.43451	0.04434	-9.800
HDI	-0.23658	0.04833	-4.895
HOSPITAL.BEDS.PER.POPULATION	0.10102	0.04061	2.487
UNEMPLOYMENT	-0.18173	0.05968	-3.045

Correlation of Fixed Effects:

	YEAR2008	YEAR2009	YEAR2010	YEAR2011	YEAR2012	YEAR2013	YEAR2014
WORKIN HDI							
HOSPIT							
YEAR2009	0.406						
YEAR2010	0.407	0.410					
YEAR2011	0.406	0.409	0.415				
YEAR2012	0.406	0.412	0.417	0.418			
YEAR2013	0.410	0.416	0.424	0.424	0.431		
YEAR2014	0.411	0.415	0.422	0.423	0.430	0.443	
WORKING.LIF	-0.020	-0.032	-0.006	0.000	-0.035	-0.036	-0.058
HDI	-0.073	-0.057	-0.096	-0.113	-0.117	-0.158	-0.165
0.279							
HOSPITAL.BE	-0.052	-0.040	-0.020	-0.008	-0.020	-0.032	-0.027
0.280	0.121						
UNEMPLOYMEN	-0.053	-0.124	-0.134	-0.132	-0.175	-0.207	-0.187
0.185	0.129	0.250					

Μοντέλο 5.29

```

Formula: y05 = LOG.FAT.PER.POP ~ (UNEMPLOYMENT + 1 | CLUSTER) + YEAR - 1 +
WORKING.LIFE + HDI + HOSPITAL.BEDS.PER.POPULATION + URBAN.PER.FATAL

REML criterion at convergence: 267.8

Scaled residuals:
  Min       1Q   Median       3Q      Max
-2.1284 -0.6815  0.0396  0.7087  2.3677

Random effects:
 Groups Name      Variance Std.Dev. Corr
 CLUSTER (Intercept) 0.03227 0.1796
          UNEMPLOYMENT 0.30622 0.5534  0.82
 Residual          0.20461 0.4523
Number of obs: 182, groups: CLUSTER, 3

Fixed effects:
              Estimate Std. Error t value
YEAR2008      0.53716    0.11639   4.615
YEAR2009      0.25618    0.11550   2.218
YEAR2010     -0.05709    0.11429  -0.500
YEAR2011     -0.09194    0.11389  -0.807
YEAR2012     -0.25415    0.11413  -2.227
YEAR2013     -0.36004    0.11451  -3.144
YEAR2014     -0.36101    0.11469  -3.148
WORKING.LIFE -0.36582    0.04239 -8.629
HDI           -0.32457    0.05168 -6.281
HOSPITAL.BEDS.PER.POPULATION 0.11098    0.03879  2.861
URBAN.PER.FATAL 0.13141    0.04364  3.011

Correlation of Fixed Effects:
              YEAR2008 YEAR2009 YEAR2010 YEAR2011 YEAR2012 YEAR2013 YEAR2014
WORKIN HDI    HOSPIT
YEAR2009      0.410
YEAR2010      0.397    0.402
YEAR2011      0.391    0.397    0.395
YEAR2012      0.384    0.393    0.394    0.393
YEAR2013      0.379    0.388    0.391    0.392    0.396
YEAR2014      0.382    0.389    0.391    0.392    0.396    0.400
WORKING.LIF -0.023    -0.025    -0.015    -0.006    -0.033    -0.041    -0.061
HDI           0.127    0.119    0.073    0.045    0.023    -0.017    -0.021
0.258
HOSPITAL.BE -0.060    -0.037    -0.020    -0.012    -0.020    -0.020    -0.020
0.232 0.125
URBAN.PER.F -0.106    -0.042    -0.034    -0.035    -0.032    -0.062    -0.086
0.043 0.221 0.118

```

Μοντέλο 5.30

Όπως φαίνεται από τα παραπάνω μοντέλα, με την εισαγωγή περισσότερων μεταβλητών η ανάλυση καθίσταται πιο περίπλοκη και συνεπώς δυσκολότερη. Προκειμένου, λοιπόν, να αποφευχθεί η σύγκριση των στοιχείων ένα προς ένα, από εδώ και στο εξής θα αξιοποιηθούν τα κριτήρια AIC και log Likelihood, μέσω των οποίων θα συγκρίνονται τα μοντέλα, ώστε τελικά να προκύψει το ορθότερο και αυτό με τα λιγότερα σφάλματα μοντέλο.

AIC(y01) = **261.1841**
AIC(y02) = 282.6933
AIC(y03) = 286.3494

$$AIC(y04) = 305.6524$$

$$AIC(y05) = 297.7588$$

Για το κριτήριο αυτό αναζητείται το μοντέλο με τη **μικρότερη θετική τιμή AIC**. Συνεπώς, στη συγκεκριμένη περίπτωση φαίνεται, πως το **μοντέλο 5.26** εξασφαλίζει τα **καλύτερα αποτελέσματα**. Παρατηρείται, ότι ο δείκτης WORKING.LIFE επιδρά με ομοιόμορφο τρόπο και στις τρεις κατηγορίες χωρών, γεγονός που συμβάλει σημαντικά στην τελική αξιολόγηση του δείκτη μέσω του κριτηρίου του AIC. Αντίθετα, όσον αφορά το δείκτη HDI, ο οποίος εμφανίζει τα αμέσως καλύτερα συνολικά αποτελέσματα, παρατηρείται, ότι όπως είχε προκύψει και προηγουμένως στο μοντέλο 5.22, ο δείκτης αυτός προκαλεί πολύ διαφορετικές αντιδράσεις στις τρεις κατηγορίες χωρών, ενώ για την δεύτερη κατηγορία δεν είναι σε θέση να περιγράψει ορθά το φαινόμενο. Παρόλα αυτά, στην περίπτωση του μοντέλου 5.27, η συνολική εικόνα του μοντέλου είναι πιο κοντά στο φαινόμενο το που επιχειρείται να μοντελοποιηθεί παρά την απόκλιση, η οποία φαίνεται να υπάρχει αν μελετηθεί μόνο η επίδραση του HDI στο μοντέλο μέσω των συντελεστών στην ανάλυση μεταβλητών επιδράσεων, στους οποίους η έντονη διαφορά διατηρείται. Ακόμη, παρατηρείται, ότι οι συντελεστές στην ανάλυση μεταβλητών επιδράσεων του μοντέλου 5.27 είναι πιο κοντά στις ιδανικές τιμές εξαιτίας της προσθήκης περαιτέρω παραμέτρων από ότι ήταν στο μοντέλο 5.22.

$$\log Lik(y01) = -115.5921 \text{ (df=15)}$$

$$\log Lik(y02) = -126.3466 \text{ (df=15)}$$

$$\log Lik(y03) = -128.1747 \text{ (df=15)}$$

$$\log Lik(y04) = -137.8262 \text{ (df=15)}$$

$$\log Lik(y05) = -133.8794 \text{ (df=15)}$$

Όπως φαίνεται από το κριτήριο μεγιστοποίησης της πιθανοφάνειας, ορθότερο μοντέλο είναι το μοντέλο 5.29, καθώς έχει **κατ' απόλυτη τιμή την μεγαλύτερη τιμή**. Παρουσιάζεται, επομένως, μια αντίθεση στην καταλληλότητα των μοντέλων, έτσι όπως αυτή ελέγχεται με τα δοθέντα κριτήρια.

Και τα δύο κριτήρια είναι ορθά και ο λόγος για τον οποίο οδηγούν σε διαφορετικά συμπεράσματα είναι η αξιοποίηση διαφορετικών μεθόδων υπολογισμού της ζητούμενης τιμής, καθώς επίσης και το γεγονός, ότι το αποτέλεσμά τους σε καμία περίπτωση δεν είναι τα ίδια, καθότι υπολογίζουν διαφορετικά μεγέθη. Το **AIC** κριτήριο χρησιμοποιείται για τη **σύγκριση των μοντέλων**, τα οποία **έχει μελετήσει ο ερευνητής**, ενώ η **συνάρτηση μέγιστης πιθανοφάνειας** πραγματοποιεί έναν **υπολογισμό** μεταξύ του **εκάστοτε μοντέλου** και του **μοντέλου** που θα προέκυπτε, εάν **δεν υπήρχαν οι μεταβλητές**, οι οποίες προσδίδουν **βαθμούς ελευθερίας**.

Στην **παρούσα έρευνα επιλέγεται** τα αποτελέσματα και η αξιολόγηση των μοντέλων να βασιστούν περισσότερο στο **AIC κριτήριο**, διότι το ενδιαφέρον έγκειται στην επιλογή μέσω ενός δοθέντος συνόλου μοντέλων, του καταλληλότερου, το οποίο συμπίπτει με τον ορισμό του κριτηρίου αυτού.

Συνεχίζοντας την ανάλυση, διαμορφώνονται τρία νέα μοντέλα, τα οποία βασίζονται στην μεταβλητή των δαπανών για τον τομέα της υγείας ανά κάτοικο (HEALTH.EXPENDITURE.PER.CAPITA) και έχουν ελαφρώς διαφορετική σύνθεση από τα προηγούμενα. Αυτό οφείλεται στο γεγονός, ότι η μεταβλητή αυτή έχει έντονη σχέση

αλληλεπίδρασης με μερικές από τις μεταβλητές των προηγούμενων μοντέλων (5.25-5.30) και για το λόγο αυτό διαμορφώνονται τα παρακάτω μοντέλα σύμφωνα με το διάγραμμα αλληλεπίδρασης, ενώ παράλληλα επιχειρείται μια προσπάθεια ανάδειξης κάποιας πιθανής μεγαλύτερης σχέσης των ιατρικών παραγόντων με το προς μελέτη φαινόμενο.

```

Formula: LOG.FAT.PER.POP ~ (HEALTH.EXPENDITURE.PER.CAPITA + 1 |
CLUSTER) + YEAR - 1 + HOSPITAL.BEDS.PER.POPULATION

REML criterion at convergence: 344.4

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.22049 -0.69202  0.00263  0.79918  1.68174

Random effects:

   Groups   Name                Variance Std.Dev. Corr
   CLUSTER (Intercept)          0.2895   0.5380
           HEALTH.EXPENDITURE.PER.CAPITA 0.2242   0.4735  -0.35
   Residual                    0.3292   0.5737

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:
              Estimate Std. Error t value
YEAR2008      0.56073   0.31374   1.787
YEAR2009      0.21687   0.31379   0.691
YEAR2010     -0.14773   0.31395  -0.471
YEAR2011     -0.15023   0.31405  -0.478
YEAR2012     -0.39831   0.31423  -1.268
YEAR2013     -0.51006   0.31430  -1.623
YEAR2014     -0.53763   0.31434  -1.710
HOSPITAL.BEDS.PER.POPULATION 0.15931   0.04877   3.267

Correlation of Fixed Effects:
              YEAR2008 YEAR2009 YEAR2010 YEAR2011 YEAR2012 YEAR2013 YEAR2014
YEAR2009      0.871
YEAR2010      0.871    0.871
YEAR2011      0.871    0.871    0.871
YEAR2012      0.871    0.871    0.872    0.871
YEAR2013      0.871    0.871    0.871    0.872    0.872
YEAR2014      0.871    0.871    0.871    0.872    0.872    0.872
HOSPITAL.BE -0.001    0.013    0.026    0.027    0.032    0.036    0.040

AIC = 368.4254
logLik = -172.2127 (df=12)

```

Μοντέλο 5.31

Formula: LOG.FAT.PER.POP ~ (HEALTH.EXPENDITURE.PER.CAPITA + HDI + WORKING.LIFE + 1 | CLUSTER) + YEAR - 1 + HOSPITAL.BEDS.PER.POPULATION

REML criterion at convergence: 208.3

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.6723	-0.5250	-0.1071	0.7348	2.8062

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.3067	0.5538	
	HEALTH.EXPENDITURE.PER.CAPITA	0.1312	0.3623	-0.86
	HDI	0.4063	0.6375	-0.15 -0.37
	WORKING.LIFE	0.2321	0.4818	0.66 -0.95 0.65
Residual		0.1418	0.3766	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
YEAR2008	0.79211	0.09190	8.619
YEAR2009	0.47425	0.09273	5.115
YEAR2010	0.16922	0.09242	1.831
YEAR2011	0.11944	0.08970	1.332
YEAR2012	-0.04058	0.09210	-0.441
YEAR2013	-0.15048	0.09074	-1.658
YEAR2014	-0.15083	0.09085	-1.660
HOSPITAL.BEDS.PER.POPULATION	0.11787	0.03084	3.822

Correlation of Fixed Effects:

	YEAR2008	YEAR2009	YEAR2010	YEAR2011	YEAR2012	YEAR2013	YEAR2014
YEAR2009	0.359						
YEAR2010	0.356	0.363					
YEAR2011	0.334	0.341	0.340				
YEAR2012	0.350	0.357	0.357	0.338			
YEAR2013	0.336	0.343	0.345	0.327	0.346		
YEAR2014	0.331	0.338	0.342	0.325	0.345	0.337	
HOSPITAL.BE	-0.027	0.003	0.016	0.018	0.018	0.019	0.020

AIC = 246.3007

logLik = -104.1504 (df=19)

Μοντέλο 5.32


```

Formula: LOG.FAT.PER.POP ~ (HEALTH.EXPENDITURE.PER.CAPITA + 1 | CLUSTER) +
YEAR - 1 + HOSPITAL.BEDS.PER.POPULATION + HDI + WORKING.LIFE

REML criterion at convergence: 269.3

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.46730 -0.65327  0.02667  0.70659  2.39673

Random effects:
   Groups Name                    Variance Std.Dev. Corr
   CLUSTER (Intercept)            0.17335  0.4163
   HEALTH.EXPENDITURE.PER.CAPITA  0.08821  0.2970  -0.81
   Residual                        0.20930  0.4575
Number of obs: 182, groups: CLUSTER, 3

Fixed effects:
              Estimate Std. Error t value
YEAR2008      0.51502   0.17513   2.941
YEAR2009      0.18166   0.17600   1.032
YEAR2010     -0.16203   0.17700  -0.915
YEAR2011     -0.19272   0.17559  -1.098
YEAR2012     -0.38463   0.17743  -2.168
YEAR2013     -0.48648   0.17699  -2.749
YEAR2014     -0.47969   0.17717  -2.708
HOSPITAL.BEDS.PER.POPULATION  0.07395   0.04000   1.849
HDI           -0.21590   0.04624  -4.670
WORKING.LIFE  -0.36903   0.04617  -7.993

Correlation of Fixed Effects:
              YEAR2008 YEAR2009 YEAR2010 YEAR2011 YEAR2012 YEAR2013 YEAR2014
HOSPIT HDI
YEAR2009      0.738
YEAR2010      0.739      0.741
YEAR2011      0.737      0.738      0.740
YEAR2012      0.739      0.741      0.743      0.741
YEAR2013      0.737      0.739      0.742      0.740      0.743
YEAR2014      0.736      0.738      0.741      0.740      0.743      0.743
HOSPITAL.BE -0.010      0.010      0.027      0.033      0.034      0.039      0.041
HDI          -0.053     -0.056     -0.077     -0.087     -0.093     -0.108     -0.116
-0.008
WORKING.LIF -0.057     -0.060     -0.056     -0.035     -0.061     -0.057     -0.071
0.229 -0.254

AIC = 297.3452
logLik = -134.6726 (df=14)

```

Μοντέλο 5.33

Παρατηρείται, ότι το μοντέλο 5.32 είναι το καλύτερο αναφορικά με τα μοντέλα 5.31-5.33 σύμφωνα, με το AIC κριτήριο. Συγκρίσει με το μοντέλο 5.26, φαίνεται να έχει μικρότερο AIC, άρα να θεωρείται ορθότερο. Όμως, εδώ εντοπίζεται η χρησιμότητα του κριτηρίου της πιθανοφάνειας, από όπου διαπιστώνεται, ότι τα δύο αυτά μοντέλα έχουν διαφορετικούς βαθμούς ελευθερίας (15 το 5.27 και 14 το 5.32), γεγονός το οποίο έχει άμεση επίδραση στα αποτελέσματά τους και καθιστά τη σύγκρισή τους κατά τι πιο δύσκολη. Επομένως, για τα επόμενα βήματα σημειώνονται τα δύο αυτά μοντέλα και επιδιώκεται είτε ο συνδυασμός τους, είτε η εύρεση κάποιου καλύτερου.

Σύμφωνα με τα προηγούμενα, προκύπτουν τα δύο παρακάτω μοντέλα, τα οποία θεωρούνται τα καλύτερα που μπορούν να προκύψουν λαμβάνοντας υπόψιν, ότι στην παρούσα Διπλωματική Εργασία σκοπός είναι η εύρεση του ορθότερου στατιστικού μοντέλου περιγραφής του φαινομένου της θνησιμότητας στα οδικά ατυχήματα αναφορικά με παράγοντες από τους τομείς της υγείας, της οικονομίας και της κοινωνίας. Στο πρώτο μοντέλο αξιοποιείται ένας παράγοντας από κάθε κλάδο, ενώ στο δεύτερο μοντέλο δίνεται μεγαλύτερη έμφαση στον τομέα της υγείας. Εξ' ου και εισάγεται ένας δεύτερος παράγοντας του κλάδου της υγείας, ο οποίος φαίνεται να συμβάλει στο προτεινόμενο μοντέλο.

```

Formula: LOG.FAT.PER.POP ~ (HOSPITAL.BEDS.PER.POPULATION + HDI +
WORKING.LIFE + 1 | CLUSTER) + YEAR - 1

REML criterion at convergence: 203.8

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.65557 -0.67301 -0.05993  0.71688  3.01882

Random effects:

Groups   Name                                Variance Std.Dev. Corr
CLUSTER  (Intercept)                            0.28318  0.5321
          HOSPITAL.BEDS.PER.POPULATION  0.05136  0.2266    1.00
          HDI                             0.31647  0.5626   -0.65 -0.70
          WORKING.LIFE                    0.13972  0.3738   -0.18 -0.25  0.86
Residual 0.14153  0.3762

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

             Estimate Std. Error t value
YEAR2008    0.27260    0.10017   2.721
YEAR2009   -0.04425    0.09880  -0.448
YEAR2010   -0.34892    0.09731  -3.586
YEAR2011   -0.37002    0.09651  -3.834
YEAR2012   -0.55670    0.09728  -5.723
YEAR2013   -0.65328    0.09708  -6.729
YEAR2014   -0.65130    0.09792  -6.652

Correlation of Fixed Effects:

             YEAR2008 YEAR2009 YEAR2010 YEAR2011 YEAR2012 YEAR2013
YEAR2009    0.450
YEAR2010    0.441    0.433
YEAR2011    0.435    0.428    0.420
YEAR2012    0.438    0.431    0.424    0.420
YEAR2013    0.435    0.428    0.421    0.418    0.423
YEAR2014    0.437    0.430    0.424    0.420    0.427    0.427

Coefficients:
$CLUSTER
(Intercept) HOSPITAL.BEDS.PER.POPULATION HDI WORKING.LIFE
1  0.55734240 0.25158355 -0.7662175 -0.4002601
2  0.71344859 0.29498309 -0.2837264  0.0878406
3 -0.09878058 -0.01905677 -0.5217944 -0.4952544

AIC = 239.8169
logLik = -101.9085 (df=18)

```

Μοντέλο 5.34

Formula: LOG.FAT.PER.POP ~ (HEALTH.EXPENDITURE.GDP + HOSPITAL.BEDS.PER.POPULATION + HDI + WORKING.LIFE + 1 | CLUSTER) + YEAR - 1

REML criterion at convergence: 174

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.6933	-0.5630	0.0007	0.5693	3.3073

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
CLUSTER	(Intercept)	0.34732	0.5893	
	HEALTH.EXPENDITURE.GDP	0.09546	0.3090	-0.57
	HOSPITAL.BEDS.PER.POPULATION	0.02136	0.1461	0.87 -0.09
	HDI	0.28441	0.5333	-0.57 -0.36 -0.90
	WORKING.LIFE	0.18467	0.4297	0.01 -0.83 -0.49
0.82	Residual	0.11834	0.3440	

Number of obs: 182, groups: CLUSTER, 3

Fixed effects:

	Estimate	Std. Error	t value
YEAR2008	0.26153	0.09305	2.811
YEAR2009	-0.08483	0.08562	-0.991
YEAR2010	-0.39967	0.08640	-4.626
YEAR2011	-0.43004	0.08618	-4.990
YEAR2012	-0.61851	0.08590	-7.200
YEAR2013	-0.71458	0.08561	-8.347
YEAR2014	-0.71245	0.08652	-8.235

Correlation of Fixed Effects:

	YEAR2008	YEAR2009	YEAR2010	YEAR2011	YEAR2012	YEAR2013
YEAR2009	0.423					
YEAR2010	0.430	0.384				
YEAR2011	0.428	0.382	0.388			
YEAR2012	0.424	0.378	0.385	0.385		
YEAR2013	0.420	0.373	0.381	0.382	0.381	
YEAR2014	0.424	0.375	0.385	0.386	0.386	0.385

Coefficients:

\$CLUSTER	(Intercept)	HEALTH.EXPENDITURE.GDP	HOSP.BEDS.PER.POP	HDI	WORKING.LIFE
1	0.6157860	0.09791985	0.21203353	-0.7996972	-0.47026905
2	0.6448171	-0.20927343	0.13416923	-0.3006895	0.03336522
3	-0.4711452	0.47728719	-0.00663056	-0.3386403	-0.57137856

AIC(a1) = 220.0372

logLik(a1) = -87.0186 (df=23)

Μοντέλο 5.35

Παρατηρείται, ότι τα δύο αυτά μοντέλα εμφανίζουν τα καλύτερα χαρακτηριστικά μέχρι στιγμής από κάθε άλλο μοντέλο που έχει μελετηθεί, ενώ η μεταξύ τους διαφορά είναι πολύ μικρή.

Παράλληλα, παρατηρούνται και οι αλληλεπιδράσεις που έχουν οι διάφοροι παράγοντες μεταξύ τους, εφόσον με την είσοδο του δείκτη HEALTH.EXPENDITURE.GDP εμφανίζονται διαφορές

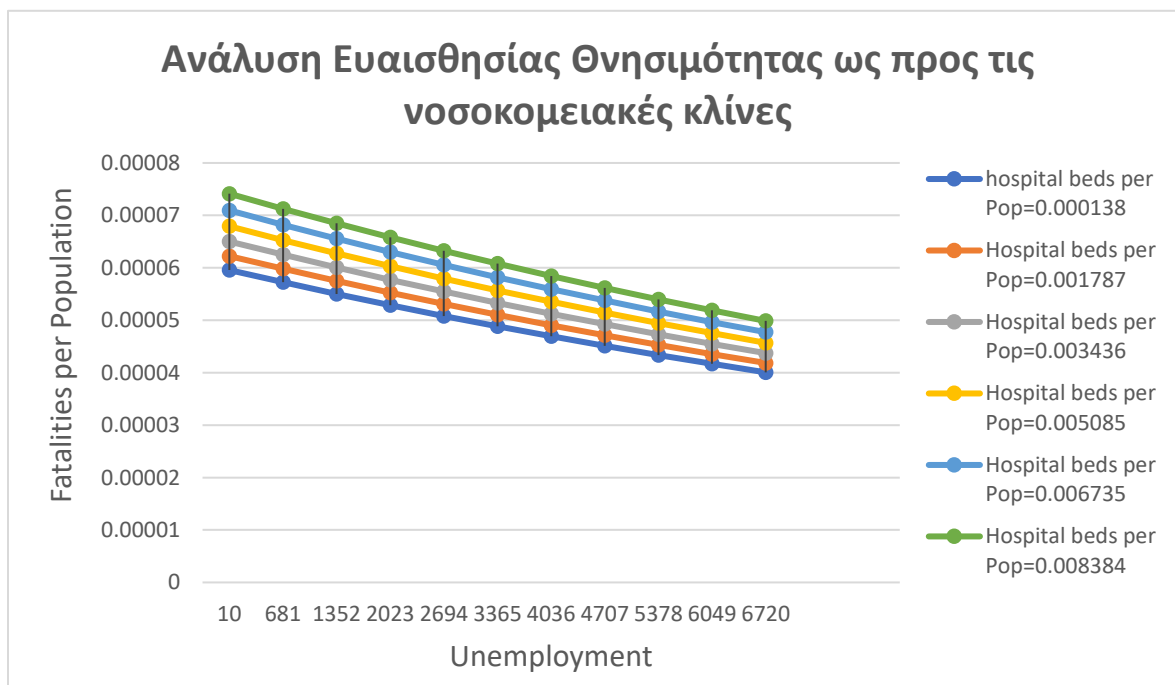
στις τιμές των υπόλοιπων δεικτών αναφορικά με την κάθε ομάδα χωρών. Εμφανής είναι η επίδραση του δείκτη αυτού, ιδιαιτέρως στον δείκτη HOSPITAL.BEDS.PER.POPULATION, ο οποίος, ενώ φαίνεται να μην έχει πλέον τόσο ομοιόμορφη συσχέτιση με το μοντέλο, έχει συνολικά μικρότερες τιμές, γεγονός που καθιστά το νεότερο μοντέλο καλύτερο του προηγούμενου, τουλάχιστον σε ό,τι αφορά την επίδραση του δείκτη HOSPITAL.BEDS.PER.POPULATION στο μοντέλο.

Στα πλαίσια της παρούσας Διπλωματικής Εργασίας επιλέγεται ως τελικό στατιστικό μοντέλο το **Μοντέλο 5.35** προκειμένου να αναδειχθεί η **σημασία των παραγόντων υγείας στη θνησιμότητα στα τροχαία ατυχήματα**, καθώς επίσης και ο **βαθμός** στον οποίο ο κλάδος της υγείας **επιηρεάζει** και δύναται να **καθορίσει την εξέλιξη του φαινομένου** αυτού.

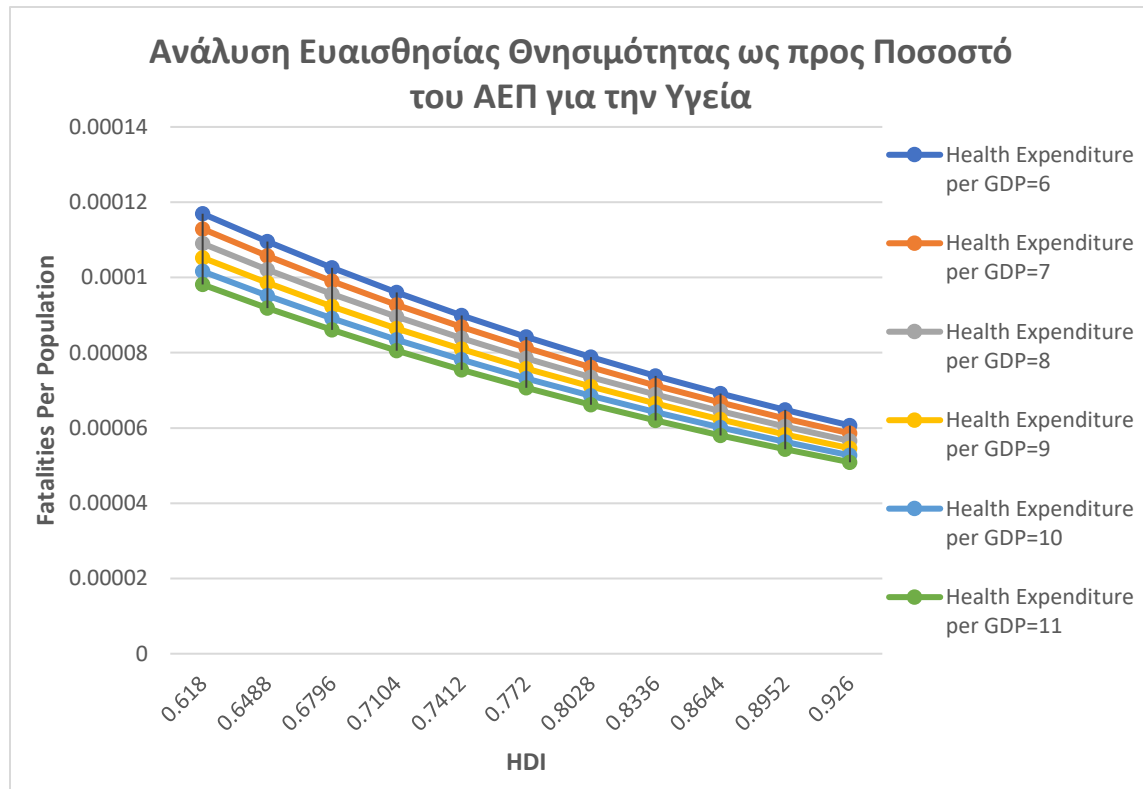
5.7 Ανάλυση Ευαισθησίας

Η Ανάλυση Ευαισθησίας μελετά τις συνέπειες που υφίσταται η βέλτιστη λύση ενός γραμμικού μοντέλου, ως συνέπεια αλλαγών στις τιμές των παραμέτρων του.

Στο παρόν υποκεφάλαιο παρουσιάζονται ορισμένα διαγράμματα ευαισθησίας, που σχεδιάστηκαν με σκοπό την καλύτερη κατανόηση της επιρροής των ανεξάρτητων μεταβλητών στην εξαρτημένη μεταβλητή, τόσο ως προς το σύνολο των κρατών-μελών, όσο και ως προς κάθε ομάδα κρατών (cluster) ξεχωριστά, για το Μοντέλο 5.15 της γραμμικής παλινδρόμησης.



Διάγραμμα 5.6 Ανάλυση ευαισθησίας ως προς τις νοσοκομειακές κλίνες για το σύνολο των χωρών αναφορικά με την ανεργία



Διάγραμμα 5.7 Ανάλυση ευαισθησίας ως προς τις δαπάνες για τον τομέα της υγείας ως ποσοστό του ΑΕΠ για το σύνολο των χωρών αναφορικά με το δείκτη ανθρώπινης ανάπτυξης

Στα παραπάνω διαγράμματα παρουσιάζονται η **ετήσια μεταβολή στον αριθμό των νεκρών** στα οδικά ατυχήματα ως προς τον πληθυσμό, σε σχέση με την **ανεργία** για τις διάφορες τιμές **των νοσοκομειακών κλινών ως προς τον πληθυσμό** (Διάγραμμα 5.6) και η **ετήσια μεταβολή στον αριθμό των νεκρών** στα οδικά ατυχήματα ως προς τον πληθυσμό, σε σχέση με τον **δείκτη ανθρώπινης ανάπτυξης (HDI)** για τις διάφορες τιμές **των δαπανών για τον τομέα της υγείας ως ποσοστό του Α.Ε.Π.** (Διάγραμμα 5.7).

Το διάγραμμα για παράδειγμα των δαπανών για τον τομέα της υγείας ως ποσοστό του Α.Ε.Π. προκύπτει, εάν χρησιμοποιώντας τους συντελεστές που προέκυψαν από τις αναλύσεις των λογαριθμισμένων γραμμικών μοντέλων διαμορφωθεί η τελική εξίσωση. Στην εξίσωση αυτή επιλέγεται να διατηρηθεί σταθερή η τιμή των δαπανών για τον τομέα της υγείας ως ποσοστό του Α.Ε.Π. και ο σταθερός όρος, ενώ δίνονται τιμές στη μεταβλητή του δείκτη ανθρώπινης ανάπτυξης και στις υπόλοιπες μεταβλητές εφόσον υπάρχουν, τιμή ίση με τη μέση τιμή τους. Αφού προκύψουν όλα τα διαφορετικά αποτελέσματα, επιλέγεται να παρουσιαστούν όλα σε ένα κοινό διάγραμμα με τα διαφορετικά χρώματα να αντιπροσωπεύουν τις διαφορετικές τιμές των δαπανών για τον τομέα της υγείας ως ποσοστό του Α.Ε.Π.. Πρέπει να σημειωθεί ότι η εξαρτημένη μεταβλητή επιλέγεται να απολογριθμοποιηθεί, προκειμένου να είναι πιο εμφανής η επιρροή της εξαρτημένης μεταβλητής στη μεταβολή του αριθμού των νεκρών στα οδικά ατυχήματα

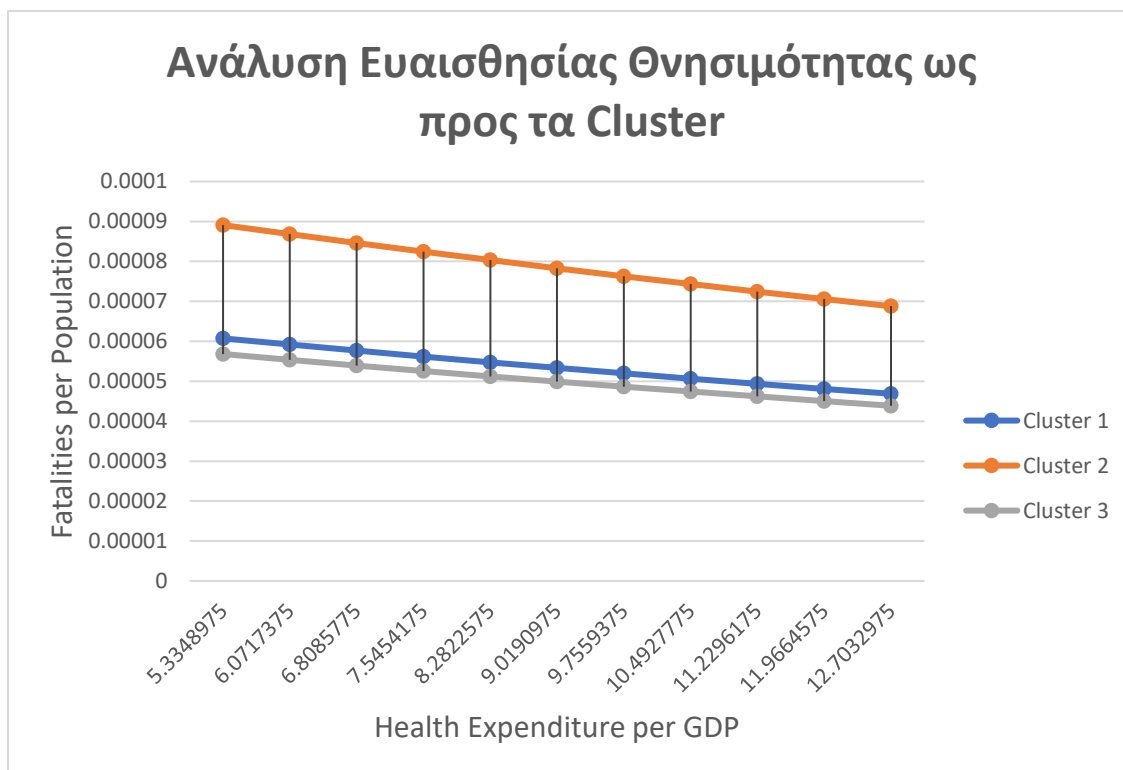
Στα παραπάνω διαγράμματα γίνεται εμφανής η σχέση αλληλεπίδρασης των ανεξάρτητων μεταβλητών τόσο με την εξαρτημένη μεταβλητή, όσο και μεταξύ τους.

Αρχικά, παρατηρείται ότι υπάρχει **αρνητική σχέση** μεταξύ **θνησιμότητας** και **ανεργίας** όπως επίσης και **θνησιμότητας** με **δείκτη ανθρώπινης ανάπτυξης**, γεγονός το οποίο είναι πλήρως συμβατό με το διάγραμμα αυτοσυσχέτισης που έχει παρατεθεί παραπάνω. Αυτό σημαίνει ότι όσο αυξάνεται το ένα μέγεθος το άλλο μειώνεται και αντίστροφα.

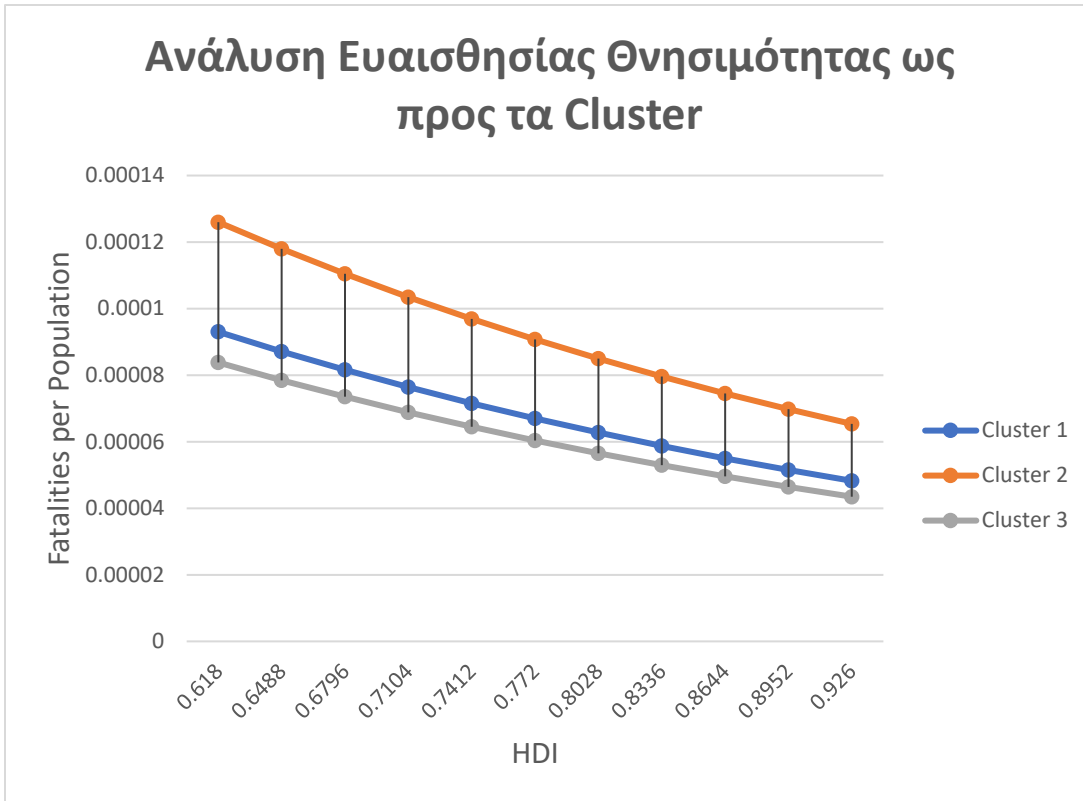
Επίσης, καθίσταται πιο κατανοητή η σχέση του πλήθους των νοσοκομειακών κλινών ως προς τον πληθυσμό με τη θνησιμότητα και των δαπανών για τον τομέα της υγείας ως ποσοστό του Α.Ε.Π. με τη θνησιμότητα. Ενώ λοιπόν αναφορικά με τις **νοσοκομειακές κλίνες** είναι **θετική** η **σχέση** τους με τη **θνησιμότητα**, δηλαδή όσο το ένα μέγεθος αυξάνεται, αυξάνεται και το άλλο, η σχέση των **δαπανών για τον τομέα της υγείας** ως ποσοστό του Α.Ε.Π. εμφανίζει **αρνητική σχέση** με την **θνησιμότητα**, δηλαδή όσο το ένα μέγεθος αυξάνεται το άλλο μειώνεται και αντίθετα. Οι παρατηρήσεις αυτές είναι και πάλι σε πλήρη συμφωνία με το διάγραμμα αυτοσυσχέτισης που έχει παρατεθεί παραπάνω.

Τέλος και στα δύο διαγράμματα παρατηρείται μια **τάση συγκέντρωσης** των **τιμών** των συναρτήσεων **γύρω από μία τιμή**, καθώς αυξάνεται είτε η ανεργία, είτε ο δείκτης ανθρώπινης ανάπτυξης για κάθε τιμή των μεταβλητών είτε των νοσοκομειακών κλινών, είτε των δαπανών για τον τομέα της υγείας ως ποσοστό του Α.Ε.Π. αντίστοιχα.

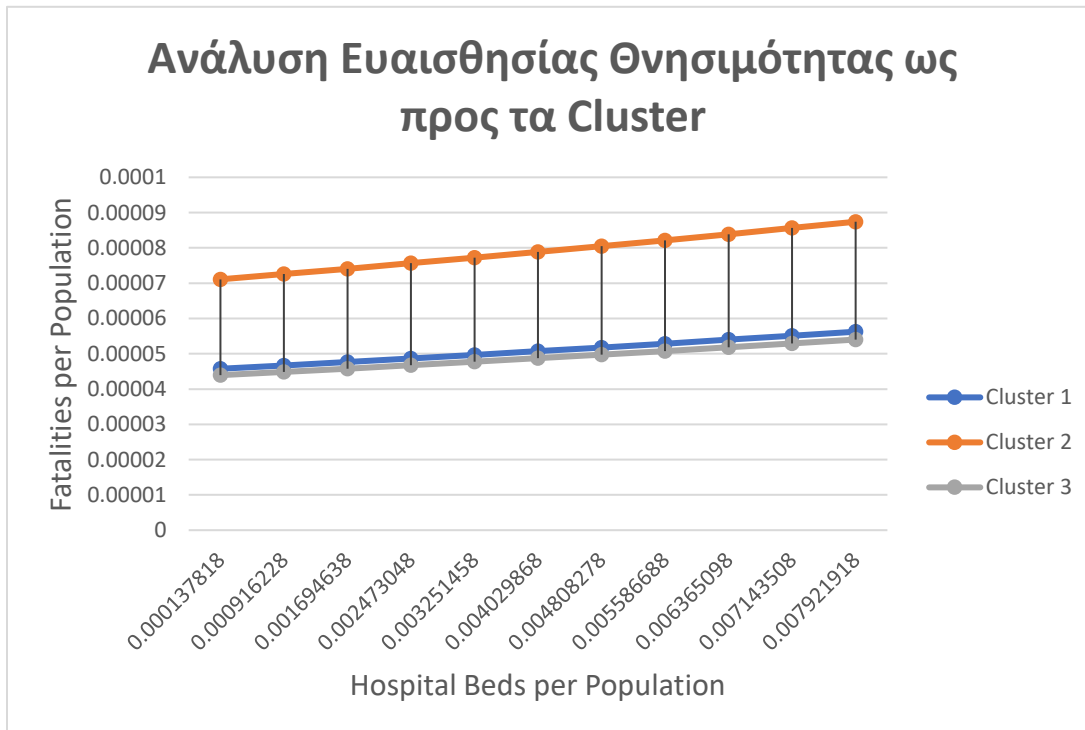
Στη συνέχεια παρατίθενται διαγράμματα με κατηγοριοποίηση των χωρών ως προς τα **cluster**.



Διάγραμμα 5.8 Ανάλυση ευαισθησίας ως προς τα cluster για δαπάνες υγείας ως ποσοστό Α.Ε.Π.



Διάγραμμα 5.9 Ανάλυση ευαισθησίας ως προς τα cluster για δείκτη ανθρώπινης ανάπτυξης



Διάγραμμα 5.10 Ανάλυση ευαισθησίας ως προς τα cluster για πλήθος κλινών ανά πληθυσμό

Στο σύνολό τους οι τρεις καμπύλες, για κάθε ένα από τα παραπάνω διαγράμματα, έχουν την ίδια μορφή όπως ακριβώς είναι και το αναμενόμενο. Η μόνη τους διαφορά είναι η μετατόπιση που έχουν υποστεί η κάθε μία ανάλογα τις τιμές της. Η μετατόπιση αυτή δικαιολογεί και την προσπάθεια διαχωρισμού των χωρών σε υποκατηγορίες προκειμένου να μελετηθούν καλύτερα, καθώς όπως φαίνεται και από τα διαγράμματα κάποιες ομάδες χωρών παρουσιάζουν ομοιότητες ενώ άλλες όχι.

Στα παραπάνω διαγράμματα παρατηρείται μια συμφωνία στις τιμές των cluster 1 και cluster 2, η οποία έχει σαν αποτέλεσμα οι καμπύλες τους να βρίσκονται πολύ κοντά και τελικά να συγκλίνουν. Αντίστοιχα, φαίνεται και η καμπύλη του cluster 2 να τείνει να συγκλίνει με τις άλλες δύο καμπύλες, απλώς η ταχύτητα σύγκλισής της είναι πολύ μικρότερη. Το σημείο σύγκλισης για τις τρεις καμπύλες φαίνεται να βρίσκεται σε μεγάλες τιμές των δαπανών υγείας ως ποσοστό Α.Ε.Π. και σε μεγάλες τιμές του δείκτη ανθρώπινης ανάπτυξης αντίστοιχα. Αντίθετα, για την μεταβλητή των νοσοκομειακών κλινών, το σημείο σύγκλισης βρίσκεται σε μικρές τιμές της μεταβλητής αυτής.

Τέλος όσον αφορά την μορφή της καμπύλης κάθε φορά, αυτή εξαρτάται από τη συσχέτιση της εξαρτημένης μεταβλητής επί του άξονα y με την ανεξάρτητη μεταβλητή επί του άξονα x .

6. Συμπεράσματα

6.1 Σύνοψη Αποτελεσμάτων

Αντικείμενο της παρούσας Διπλωματικής Εργασίας αποτέλεσε η διερεύνηση της επιρροής κοινωνικών, οικονομικών και ιατρικών παραγόντων στα οδικά ατυχήματα και συγκεκριμένα στην ετήσια μεταβολή του αριθμού των νεκρών στα οδικά ατυχήματα ως προς τον πληθυσμό.

Μετά τον καθορισμό του επιδιωκόμενου στόχου, ξεκίνησε η βιβλιογραφική αναζήτηση ερευνών συναφών με το αντικείμενο της παρούσας Διπλωματικής Εργασίας σε χώρες της Ευρωπαϊκής Ένωσης, αλλά και παγκοσμίως, καθώς και των διαθέσιμων στοιχείων που ήταν απαραίτητα για τη συγκεκριμένη διερεύνηση.

Στη συνέχεια, ακολούθησε η συλλογή των απαραίτητων για την εργασία στοιχείων από βάσεις δεδομένων διάφορων οργανισμών, όπως EUROSTAT, CARE και UNDP, προκειμένου να αναπτυχθεί η σχετική βάση δεδομένων. Τα στοιχεία που συλλέχθηκαν περιελάμβαναν τον αριθμό των νεκρών στα οδικά ατυχήματα, τον πληθυσμό, το κατά κεφαλήν Ακαθάριστο Εγχώριο Προϊόν (Α.Ε.Π.), τον ετήσιο αριθμό νεκρών σε οδικά ατυχήματα, το πλήθος των οδικών ατυχημάτων εντός αστικού δικτύου, την προσδοκώμενη διάρκεια ζωής των ανθρώπων, τα παραγωγικά χρόνια ζωής των ανθρώπων, το κατά κεφαλήν Α.Ε.Π., το δείκτη ανθρώπινης ανάπτυξης, τις ετήσιες δαπάνες για τον τομέα της υγείας, τις δημόσιες δαπάνες για τον τομέα της υγείας και τον αριθμό των κλινών στα νοσοκομεία των 27 κρατών-μελών της Ευρωπαϊκής Ένωσης για τη χρονική περίοδο 2008-2014.

Τη συλλογή των δεδομένων ακολούθησε η επεξεργασία τους προκειμένου να επιλεγεί η κατάλληλη μεθοδολογία, αλλά και να κωδικοποιηθούν, ώστε να εισαχθούν στο ειδικό στατιστικό λογισμικό (Rstudio©). Η στατιστική ανάλυση πραγματοποιήθηκε μέσω Γραμμικής Παλινδρόμησης και Γραμμικού Μικτού Μοντέλου. Αρχικά, η ανάλυση πραγματοποιήθηκε για το σύνολο των κρατών, σύμφωνα με **τη λογαριθμισμένη γραμμική παλινδρόμηση, με εξαρτημένη μεταβλητή την ετήσια μεταβολή του φυσικού λογαρίθμου του αριθμού των νεκρών στα οδικά ατυχήματα και ως ανεξάρτητες μεταβλητές:**

- το κατά κεφαλήν Α.Ε.Π.
- το πλήθος των οδικών ατυχημάτων εντός αστικού δικτύου συναρτήσει του συνόλου των θανατηφόρων οδικών ατυχημάτων
- την προσδοκώμενη διάρκεια ζωής των ανθρώπων
- τα παραγωγικά χρόνια ζωής των ανθρώπων
- το δείκτη ανθρώπινης ανάπτυξης
- τις ετήσιες δαπάνες για τον τομέα της υγείας
- τις δημόσιες δαπάνες για τον τομέα της υγείας
- τον αριθμό των κλινών στα νοσοκομεία ως προς τον πληθυσμό

Οι ομάδες των κρατών προέκυψαν από την εργασία της Υπατίας Μίχου[2].

Στη συνέχεια, δημιουργήθηκαν στατιστικά μοντέλα μέσω της εφαρμογής του **Γραμμικού Μικτού Μοντέλου**, τα οποία περιελάμβαναν ως **εξαρτημένη μεταβλητή την ετήσια μεταβολή του φυσικού λογαρίθμου του αριθμού των νεκρών στα οδικά ατυχήματα** και ως **ανεξάρτητες μεταβλητές:**

- το κατά κεφαλήν Α.Ε.Π.
- το πλήθος των οδικών ατυχημάτων εντός αστικού δικτύου συναρτήσει του συνόλου των θανατηφόρων οδικών ατυχημάτων
- την προσδοκώμενη διάρκεια ζωής των ανθρώπων
- τα παραγωγικά χρόνια ζωής των ανθρώπων
- το δείκτη ανθρώπινης ανάπτυξης
- τις ετήσιες δαπάνες για τον τομέα της υγείας
- τις δημόσιες δαπάνες για τον τομέα της υγείας
- τον αριθμό των κλινών στα νοσοκομεία ως προς τον πληθυσμό

Από αυτά τα μοντέλα προέκυψαν τα δύο τελικά μοντέλα, τα οποία έχουν ως εξαρτημένη μεταβλητή την ετήσια μεταβολή του φυσικού λογαρίθμου του αριθμού των νεκρών στα οδικά ατυχήματα και ως ανεξάρτητες μεταβλητές τα παραγωγικά χρόνια ζωής των ανθρώπων, το δείκτη ανθρώπινης ανάπτυξης, τις ετήσιες δαπάνες για τον τομέα της υγείας ανά κάτοικο και τον αριθμό των κλινών στα νοσοκομεία ως προς τον πληθυσμό.

Πρέπει να σημειωθεί, ότι η επιλογή των μοντέλων, τόσο για το σύνολο των κρατών όσο και για την κάθε ομάδα ξεχωριστά, έγινε έπειτα από αρκετές δοκιμές συνδυασμών ανεξάρτητων και εξαρτημένων μεταβλητών. Στο προηγούμενο κεφάλαιο έγινε η ανάπτυξη και επεξήγηση των τελικών μοντέλων που προέκυψαν από τη στατιστική επεξεργασία, αφού πρώτα ελέγχθηκε κατά πόσο αυτά πληρούν τους απαραίτητους στατιστικούς ελέγχους.

6.2 Συμπεράσματα

Από τα διάφορα στάδια εκπόνησης της Διπλωματικής Εργασίας προέκυψαν αποτελέσματα άμεσα συνδεδεμένα με τον κύριο στόχο που τέθηκε αρχικά. Τα σημαντικότερα συμπεράσματα, τα οποία προκύπτουν από την ανάλυση των αποτελεσμάτων των διαμορφωθέντων μαθηματικών μοντέλων είναι τα εξής:

1. Από τα τέσσερα μοντέλα τα οποία αναπτύχθηκαν διαπιστώθηκε ότι η θνησιμότητα στα οδικά ατυχήματα έχει άμεση σχέση με τον κλάδο της υγείας και μάλιστα σε μεγάλο βαθμό.
2. Η συσχέτιση αυτή εξηγείται, εάν σκεφτεί κανείς ότι εξαιτίας των οδικών ατυχημάτων και των συνεπειών τους, όπως ο τραυματισμός των ανθρώπων, σε μικρότερο ή μεγαλύτερο βαθμό, καθίσταται επιτακτική η ανάγκη παροχής ιατρικής βοήθειας. Συνεπώς, όσο αυξάνεται η ανάγκη παροχής ιατρικής βοήθειας είναι λογικό να αυξάνονται και οι

υπηρεσίες υγείας προκειμένου να καλύψουν τη ζήτηση. Ακολούθως, αυξάνονται οι επενδύσεις σε αυτό τον κλάδο και κατά συνέπεια το πλήθος του έμφυχου και άψυχου δυναμικού. Αντίστοιχα, χάρη στην καλύτερη ενημέρωση και εφαρμογή μέτρων πρόληψης, μειώνονται οι ιατρικοί λόγοι, οι οποίοι θα μπορούσαν να οδηγήσουν σε οδικά ατυχήματα, όπως προβλήματα όρασης, νευρικές βλάβες και σπάνιες παθήσεις του μυοσκελετικού συστήματος.

3. Όπως φαίνεται από τα αποτελέσματα της λογαριθμισμένης γραμμικής παλινδρόμησης, ιδιαίτερη έμφαση πρέπει να δοθεί και σε παράγοντες κοινωνικού και οικονομικού χαρακτήρα, όπως ο δείκτης ανθρώπινης ανάπτυξης, τα παραγωγικά χρόνια ζωής και η ανεργία. Παράλληλα, εφόσον το φαινόμενο το οποίο μελετάται εμπίπτει στον τομέα των συγκοινωνιών, είναι χρήσιμο στο μοντέλο να εντάσσονται και παράγοντες από αυτό τον τομέα, όπως η αναλογία ατυχημάτων εντός αστικού ιστού ως προς το σύνολο των θανατηφόρων οδικών ατυχημάτων.
4. Όπως προκύπτει από τα τελικά μοντέλα του Γραμμικού Μικτού Μοντέλου, η επιρροή των ιατρικών παραγόντων διαφέρει ανάλογα με την ομάδα κρατών που μελετάται. Όταν εμφανίζεται το πλήθος των νοσοκομειακών κλινών ως προς τον πληθυσμό ως ο μόνος ιατρικός παράγοντας, η πρώτη και η δεύτερη ομάδα χωρών εμφανίζουν μεγάλη ομοιομορφία στις τιμές των συντελεστών τους, ενώ η τρίτη ομάδα φαίνεται να επηρεάζεται περισσότερο από το συγκεκριμένο παράγοντα και ο αντίστοιχος συντελεστής της να είναι αισθητά μικρότερος από τις άλλες δύο (το οποίο είναι και το επιθυμητό). Αντίθετα, όταν προστίθεται ένας ακόμα παράγοντας από τον κλάδο της υγείας στο στατιστικό μοντέλο, η συνολική εικόνα του μοντέλου είναι σχετικά καλύτερη.
5. Αναφορικά με τους χρησιμοποιηθέντες δείκτες από τους τομείς της κοινωνίας και της οικονομίας παρατηρείται, ότι ενώ και οι δύο επηρεάζονται από κάθε μεταβολή (προσθαφαίρεση) μεταβλητών στο σύστημα, τα παραγωγικά χρόνια ζωής των ανθρώπων τείνουν να παραμένουν σταθερά στις τιμές τους. Αντίθετα, ο δείκτης ανθρώπινης ανάπτυξης επηρεάζεται πιο έντονα και ειδικά όταν υπάρχουν δύο παράγοντες από τον κλάδο της υγείας. Τότε, οι τιμές των συντελεστών για το δείκτη αυτό, για τις ομάδες χωρών δύο και τρία, τείνουν να συγκλίνουν σε μια τιμή κοντά στο -0.3.
6. Από τα παραχθέντα μοντέλα διακρίνεται τόσο η σχέση μεταξύ των δεικτών (ανεξάρτητων μεταβλητών), κάθε κλάδου ξεχωριστά, με την εξαρτημένη μεταβλητή, όσο και οι σχέσεις αλληλεπίδρασης των δεικτών από τους διαφορετικούς κλάδους μεταξύ τους.
7. Παρατηρείται, ότι, κάθε ομάδα χωρών επηρεάζεται σε διαφορετικό βαθμό από τους ίδιους παράγοντες, γεγονός αναμενόμενο και πλήρως συμβατό με την πρόταση διάσπασης και μελέτης του συνόλου των χωρών σε μικρότερες ομάδες, ώστε να είναι πιο ρεαλιστικά τα αποτελέσματα και με όσο το δυνατόν μικρότερο σφάλμα. Αυτό επιβεβαιώνεται από την καλύτερη (μικρότερη) τιμή του κριτηρίου AIC μεταξύ των μοντέλων με ομάδες κρατών και των μοντέλων του συνόλου των χωρών άνευ ομάδων.

8. Η υποδιαίρεση του συνόλου των κρατών-μελών της Ευρωπαϊκής Ένωσης σε τρεις ομάδες με βάση τον πληθυσμό, το κατά κεφαλήν Α.Ε.Π. και τον αριθμό νεκρών από οδικά ατυχήματα ανά εκατομμύριο πληθυσμού της κάθε χώρας, αποτελεί μια κατάλληλη μέθοδο προσέγγισης της επιρροής των ιατρικών, οικονομικών και κοινωνικών παραγόντων κάθε ομάδας κρατών στην οδική ασφάλεια.
9. Η μέθοδος του Γραμμικού Μικτού Μοντέλου είναι κατάλληλη για την ανάλυση παραμέτρων οδικών ατυχημάτων σε σχέση με ιατρικούς, κοινωνικούς και οικονομικούς παράγοντες. Τα τελικά μαθηματικά μοντέλα, τα οποία αναπτύχθηκαν με τη χρήση της μεθόδου αυτής, θεωρούνται γενικά αξιόπιστα, αφού είχαν καλή προσαρμογή στα δεδομένα και τα αποτελέσματά τους κρίθηκαν από το σύστημα ικανοποιητικά.

6.3 Περιορισμοί Παρούσας Έρευνας

Στο σημείο αυτό πρέπει να τονιστεί, ότι τα παραπάνω αποτελέσματα και συμπεράσματα ισχύουν για τις 27 χώρες της Ε.Ε., οι οποίες μελετήθηκαν, ενώ η αναγωγή των συμπερασμάτων αυτών για ευρύτερα σύνολα χωρών ή ακόμα και σε παγκόσμια κλίμακα θα πρέπει να γίνει ύστερα από εκτεταμένη μελέτη και επισταμένη έρευνα. Οι συγκεκριμένες χώρες επελέγησαν λόγω του πλήθους των δεδομένων τα οποία έχουν δημοσιευθεί αναφορικά με αυτές, στους τομείς της υγείας, της οικονομίας, της κοινωνίας και των συγκοινωνιών. Χώρες όπως η Μάλτα και η Λιθουανία εμφανίζουν έλλειψη στοιχείων και ως εκ τούτου αποκλείστηκαν από το δειγματικό χώρο.

Αναφορικά με το χρονικό πλαίσιο της μελέτης, καθότι αυτό είναι περιορισμένο στα έτη 2008-2014 συνίσταται ιδιαίτερη προσοχή σε πιθανά συμπεράσματα και ισχυρισμούς, οι οποίοι θα μπορούσαν να διατυπωθούν για μετέπειτα χρονικές περιόδους, όπως αυτή της συγγραφής της παρούσας έρευνας. Κάτι τέτοιο, θα μπορούσε να οδηγήσει σε εσφαλμένα συμπεράσματα, εάν δεν πραγματοποιηθεί κάτω από τις κατάλληλες προϋποθέσεις και ύστερα από σχετική έρευνα για την προς μελέτη χρονική περίοδο. Η επιλογή της προαναφερθείσας χρονικής περιόδου έγινε σύμφωνα με το πλήθος των δεδομένων, τα οποία ήταν διαθέσιμα και είτε παρουσίαζαν λιγότερες ελλείψεις στο χρονικό αυτό πλαίσιο, είτε κάποιες πιθανές ελλείψεις μπόρεσαν να συμπληρωθούν μέσω της εφαρμογής γραμμικής παλινδρόμησης, για την πρόβλεψη της τιμής τους, σύμφωνα με προηγούμενες και επόμενες χρονιές.

Συνεπώς, οι περιορισμοί αυτοί θα πρέπει να λαμβάνονται υπόψιν, κάθε φορά που επιχειρείται αξιολόγηση των αποτελεσμάτων της παρούσας έρευνας ή κάποια αναφορά στην παρούσα έρευνα και τα αποτελέσματά της.

6.4 Προτάσεις

Με βάση τα συμπεράσματα που προέκυψαν από τη σύνθεση των αποτελεσμάτων, αλλά και σε συνδυασμό με άλλα στοιχεία, είναι δυνατό να διατυπωθούν οι συνολικές προτάσεις της Διπλωματικής Εργασίας, όπως αυτές συνοψίζονται παρακάτω.

Οι φορείς της πολιτείας που είναι υπεύθυνοι για την εφαρμογή και την παρακολούθηση της προόδου του εθνικού προγράμματος οδικής ασφάλειας οφείλουν να λαμβάνουν υπόψη την επιρροή τόσο των οικονομικών και κοινωνικών παραγόντων κατά τη φάση της αξιολόγησης της εφαρμογής του προγράμματος αυτού, όσο και των ιατρικών παραγόντων, ώστε να μπορούν να μελετούν καλύτερα και πιο σφαιρικά την επιρροή των συγκεκριμένων μέτρων που εφαρμόστηκαν συγκριτικά με τους τομείς, οι οποίοι επηρεάζουν το φαινόμενο της θνησιμότητας στα οδικά ατυχήματα.

Επιπλέον, οι αρμόδιοι φορείς και τα όργανα λήψης αποφάσεων στην προσπάθειά τους να μειωθεί ο αριθμός των οδικών ατυχημάτων και κατ' επέκταση οι ανθρώπινες απώλειες εξαιτίας αυτών, θα πρέπει να λαμβάνουν υπόψη και την επιρροή μακροσκοπικών χαρακτηριστικών (π.χ. Α.Ε.Π., κοινωνία). Με αυτόν τον τρόπο θα είναι σε θέση να θέτουν τους σωστούς στόχους, οι οποίοι θα οδηγήσουν σε επιδόσεις οδικής ασφάλειας καλύτερες από τις αναμενόμενες και θα μπορούν να συγκριθούν σε ευρωπαϊκό επίπεδο.

Θα πρέπει, επομένως, ο κλάδος της υγείας να μην μελετάται ερευνητικά μόνο σχετικά με την αποκατάσταση των ασθενών, εξαιτίας οδικών ατυχημάτων, αλλά και ως ένας εκ των παραγόντων επίδρασης στα οδικά ατυχήματα. Εξάλλου, ο τομέας της υγείας διαδραματίζει καθοριστικό ρόλο στην πρόληψη τόσο μέσω της ενημέρωσης, όσο και μέσω των ενδεδειγμένων και επαναλαμβανόμενων εξετάσεων.

Παράλληλα, η μείωση του αριθμού των ατυχημάτων που παρατηρείται σε περιόδους οικονομικής ύφεσης δε θα πρέπει να αποτελεί καθησυχαστικό παράγοντα για τους αρμόδιους φορείς, αλλά να λαμβάνεται υπόψη η επιρροή της οικονομικής κατάστασης της χώρας και να εφαρμόζονται μέτρα που θα στοχεύουν στην περαιτέρω μείωση τόσο του αριθμού, όσο και της σοβαρότητας των ατυχημάτων.

Τέλος, σύμφωνα και με την παρατεθείσα βιβλιογραφία, ο αριθμός των θανάτων από τροχαία ατυχήματα αυξάνεται με την αύξηση της κινητικότητας στα αρχικά στάδια της οικονομικής ανάπτυξης, στη συνέχεια, όμως, λόγω των προόδων των τεχνικών, πολιτικών και πολιτικών θεσμών, ο αριθμός των θανάτων μειώνεται με την αύξηση του κατά κεφαλήν εισοδήματος. Σε καμία περίπτωση, όμως, δε θα πρέπει αυτή η παρατήρηση να λειτουργήσει ανασταλτικά στην περαιτέρω μελέτη αναφορικά με την οδική ασφάλεια και τη λήψη νέων και καλύτερα προσαρμοσμένων, στα πλαίσια της κάθε χώρας, μέτρων οδικής ασφάλειας.

6.5 Προτάσεις για Περαιτέρω Έρευνα

Στο πλαίσιο της παρούσας Διπλωματικής Εργασίας μελετήθηκε η επιρροή ποικίλων παραγόντων από τους τομείς της υγείας, της οικονομίας και της κοινωνίας στα οδικά ατυχήματα στις χώρες της Ευρωπαϊκής Ένωσης για τα έτη 2008-2014 τόσο στο σύνολο τους, όσο και σε επιλεγμένες ομάδες. Αρκετά ενδιαφέροντα θα ήταν η επέκταση της συγκεκριμένης έρευνας τόσο στο χρονικό πλαίσιο, εικοσαετία, πενηκονταετία, όσο και γεωγραφικά, με μια αντίστοιχη μελέτη να πραγματοποιείται είτε συγκριτικά μεταξύ των ηπείρων, είτε για το σύνολο των χωρών με κριτήριο διάκρισης τον παράγοντα της ανάπτυξης, είτε ακόμα και μεταξύ των χωρών με κριτήριο διάκρισης τη μορφή και τη λειτουργικότητα του συστήματος υγείας.

Με αυτόν τον τρόπο θα μπορούσε να ελεγχθεί, εάν οι παράμετροι, οι οποίες χρησιμοποιήθηκαν στην παρούσα εργασία έχουν την ίδια βαρύτητα παγκοσμίως, καθώς επίσης και να συγκριθούν τα μέτρα οδικής ασφάλειας μεταξύ των χωρών με παρόμοια χαρακτηριστικά, με σκοπό την αναγνώριση πιθανών παραλείψεων και τάσεων, οι οποίες ενδέχεται να συνδέονται με πολιτιστικά, γεωγραφικά και οικονομικά χαρακτηριστικά.

Επιπλέον, σε επόμενη φάση θα μπορούσε να γίνει διερεύνηση της συσχέτισης της μεταβολής των ίδιων παραγόντων και του αριθμού των νεκρών στα οδικά ατυχήματα με χρήση άλλων στατιστικών μεθόδων, όπως με κάποια μέθοδο ανάλυσης χρονοσειρών.

Τέλος, η ανάλυση που πραγματοποιήθηκε στο πλαίσιο της παρούσας Διπλωματικής Εργασίας θα μπορούσε να επαναληφθεί μετά από λίγα χρόνια, όπου περισσότερα στοιχεία θα είναι διαθέσιμα, ώστε να επαληθευτεί η ισχύς των αποτελεσμάτων, αλλά και να αξιολογηθούν καλύτερα οι επιπτώσεις των ισχυόντων νόμων οδικής ασφάλειας και της τρέχουσας οικονομικής ύφεσης στην οδική ασφάλεια.

Βιβλιογραφία

- [1] T. Hundley, “puitzercenter.org.” [Online]. Available: <http://puitzercenter.org/projects/roads-kill-traffic-safety-world-health-organization-united-nations-fatalities-puitzer-center-reporting-interactive-map-data-visualization>.
- [2] Υ. Μίχου, “Συγκριτική διερεύνηση του κόστους των οδικών ατυχημάτων στην Ευρωπαϊκή Ένωση.”
- [3] G. D. Jacobs and C.A. Cutting, “Further research on accident rates in developing countries.”
- [4] R. Sánchez-Mangas, A. García-Ferrer, A. De Juan, and A. M. Arroyo, “The probability of death in road traffic accidents. How important is a quick medical response?,” *Accid. Anal. Prev.*, vol. 42, no. 4, pp. 1048–1056, 2010.
- [5] Z. Bodalal, R. Bendardaf, and M. Ambarek, “A study of a decade of road traffic accidents in benghazi - libya: 2001 to 2010,” *PLoS One*, vol. 7, no. 7, pp. 1–7, 2012.
- [6] A. García-Altés, a J. M. Suelvesb, and E. Barberí, “Ahorro asociado a 10 años de políticas de seguridad vial en Cataluña, España,” *Bull. World Health Organ.*, vol. 91, no. 1, pp. 28–35, 2013.
- [7] J. I. Castillo-Manzano, M. Castro-Nuño, and X. Fageda, “Can health public expenditure reduce the tragic consequences of road traffic accidents? The EU-27 experience,” *Eur. J. Heal. Econ.*, vol. 15, no. 6, pp. 645–652, 2014.
- [8] S. E. Eroğlu *et al.*, “Details of motorcycle accidents and their impact on healthcare costs,” *Ulus. Travma ve Acil Cerrahi Derg.*, vol. 19, no. 5, pp. 423–428, 2013.
- [9] R. Palmera-Suárez, T. López-Cuadrado, J. Almazán-Isla, R. Fernández-Cuenca, E. Alcalde-Cabero, and I. Galán, “Disability related to road traffic crashes among adults in Spain,” *Gac. Sanit.*, vol. 29, pp. 43–48, 2015.
- [10] N. Soderlund and A. B. Zwi, “Traffic-related mortality in industrialized and less developed countries,” *Bull. World Health Organ.*, vol. 73, no. 2, pp. 175–182, 1995.
- [11] T. H. Law, R. B. Noland, and A. W. Evans, “The sources of the Kuznets relationship between road fatalities and economic growth,” *J. Transp. Geogr.*, vol. 19, no. 2, pp. 355–365, 2011.
- [12] C. Antoniou, G. Yannis, E. Papadimitriou, and S. Lassarre, “Relating traffic fatalities to GDP in Europe on the long term,” *Accid. Anal. Prev.*, vol. 92, pp. 89–96, 2016.

- [13] C. Pal, S. Hirayama, S. Narahari, M. Jeyabharath, G. Prakash, and V. Kulothungan, “An insight of World Health Organization (WHO) accident database by cluster analysis with self-organizing map (SOM),” *Traffic Inj. Prev.*, vol. 19, pp. S15–S20, 2018.
- [14] Ι. Ζαντίρης, “Συσχέτιση επιδόσεων οδικής ασφάλειας με οικονομικούς και κοινωνικούς δείκτες στην ευρωπαϊκή ένωση,” 2016.
- [15] “Eurostat.” [Online]. Available:
http://ec.europa.eu/eurostat/data/database?p_p_id=NavTreeprotletprod_WAR_NavTreeprotletprod_INSTANCE_nPqeVbPXRmWQ&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_pos=1&p_p_col_count=2.