

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Σχολή Χρηματοοικονομικής και Στατιστικής



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ
ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

Διερεύνηση παραγόντων που σχετίζονται με εμφάνιση
καρδιακών επεισοδίων

Βλάρα Βαλεντίνα

Διπλωματική Εργασία

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

Πειραιάς
Ιούλιος 2023

UNIVERSITY OF PIRAEUS
School of Finance and Statistics



DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE
POSTGRADUATE PROGRAM IN APPLIED STATISTICS

**Exploration of factors related to the coronary
heart incidents**

Vlara Valentina

M.Sc. Dissertation

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

PIRAEUS, GREECE

July 2023

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

Τα μέλη της επιτροπής ήταν:

1. Τζαβελάς Γεώργιος, Αναπληρωτής Καθηγητής (επιβλέπων)
2. Βερροπούλου Γεωργία, Καθηγήτρια
3. Λαβδανίτη Μαρία, Αναπληρώτρια Καθηγήτρια, Διεθνές Πανεπιστήμιο της Ελλάδος

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

Ευχαριστίες

Θα ήθελα να ευχαριστήσω πολύ τον επιβλέποντα καθηγητή μου κ. Γεώργιο Τζαβελά για την καθοδήγηση και τη βοήθεια του καθ' όλη τη διάρκεια προετοιμασίας της διπλωματικής μου εργασίας αλλά και την υπομονή του όλο αυτό το διάστημα της εκπόνησής της.

ΠΕΡΙΛΗΨΗ

Τα καρδιαγγειακή νοσήματα ή αλλιώς οι παθήσεις της καρδιάς αποτελούν μία από τις κυριότερες αιτίες θανάτου στον κόσμο, γεγονός που καθιστά συνεχή την ανάγκη για έρευνα των παραγόντων που οδηγούν σε καρδιολογικά νοσήματα. Τα τελευταία χρόνια ο αριθμός της θνησιμότητας λόγω καρδιαγγειακών νοσημάτων συνεχώς αυξάνεται γεγονός που αποδίδεται κυρίως στον τρόπο ζωής των αναπτυγμένων χωρών. Η παρούσα εργασία έχει ως σκοπό να μελετήσει δημογραφικούς και κλινικούς παράγοντες καθώς επίσης και τους παράγοντες κατανάλωσης καπνού και αλκοόλ οι οποίοι είναι δυνατό να σχετίζονται με την εμφάνιση καρδιακών επεισοδίων.

Για την πραγματοποίηση αυτής της μελέτης χρησιμοποιήθηκαν 2 βάσεις δεδομένων οι οποίες περιέχουν άτομα τα οποία εμφάνισαν ή όχι κάποιο καρδιολογικό πρόβλημα καθώς επίσης την ηλικία τους, τις συνήθειες τους σχετικά με το κάπνισμα και το αλκοόλ καθώς επίσης και κλινικά δεδομένα τα οποία αφορούν την αρτηριακή τους πίεση συστολική (ΣΑΠ) και διαστολική (ΔΑΠ), την τιμή της χοληστερόλης και της γλυκόζης του αίματος. Η 1^η βάση μελέτης αφορά γυναίκες και άνδρες ώστε να ελεγχθεί αν υπάρχουν σημαντικές διαφορές εμφάνισης καρδιαγγειακών νοσημάτων ανάμεσα στα δύο φύλα. Η 2^η βάση δεδομένων περιέχει μόνο άνδρες για να μελετηθούν οι παράγοντες κινδύνου που οδηγούν σε νοσήματα της καρδιάς καθώς σύμφωνα με μελέτες οι άνδρες τείνουν να εμφανίζουν κάποια στεφανιαία νόσο σε νεαρότερη ηλικία από τις γυναίκες.

Η βασική στατιστική μέθοδος που χρησιμοποιήθηκε για την πρόβλεψη της καρδιαγγειακής νόσου είναι πολλαπλή λογιστική παλινδρόμηση από την οποία και προέκυψε ότι παράγοντες που αφορούν κλινική κατάσταση του ατόμου, όπως η αρτηριακή πίεση και η τιμή της χοληστερόλης καθώς επίσης και η κατανάλωση αλκοόλη ή η καπνιστική συνήθεια αποτελούν παράγοντες κινδύνου εμφάνισης στεφανιαίας νόσου. Τέλος, συγκρίνονται και οι διαφορές ανάμεσα στα δύο φύλα σχετικά με την εμφάνιση παθήσεων της καρδιάς όπου οι γυναίκες φαίνεται να εμφανίζουν καρδιολογικά νοσήματα σε μεγαλύτερη ηλικία από τους άνδρες, ωστόσο οι ίδιοι παράγοντες κινδύνου είναι σημαντικοί και για τις γυναίκες.

ABSTRACT

Cardiovascular diseases or heart diseases are one of the main causes of death in the world, which makes the need for continuous research into the factors that lead to heart diseases. In recent years, the number of deaths due to cardiovascular diseases is constantly increasing, which is mainly attributed to the lifestyle of developed countries. The purpose of this research study is to investigate demographic and clinical factors as well as tobacco and alcohol consumption factors that may be associated with the occurrence of heart attacks.

To carry out this study, 2 databases were used containing subjects who did or did not develop a heart problem, their age, smoking and alcohol habits as well as clinical data on systolic (SBP) and diastolic (DBP) blood pressure, the value of cholesterol and blood glucose. The 1st database includes women and men in order to test whether there are significant gender differences in the incidence of cardiovascular disease. The 2nd database contains exclusively male gender for the study of risk factors leading to heart disease as according to studies men tend to develop coronary heart disease at a younger age than women.

The basic statistical method used to predict cardiovascular disease is multiple logistic regression from which it emerged that factors related to the individual's clinical condition, such as blood pressure and cholesterol value, as well as alcohol consumption or smoking habit are risk factors for coronary heart disease. Finally, gender differences in the occurrence of heart disease are compared, finding that women develop heart diseases at an older age than men, however the same risk factors are significant for women.

Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1

| | | |
|---------|--|----|
| 1.1 | ΕΙΣΑΓΩΓΗ..... | 10 |
| 1.1.1 | ΑΙΤΙΕΣ ΘΑΝΑΤΟΥ ΣΕ ΠΑΓΚΟΣΜΙΑ ΚΛΙΜΑΚΑ..... | 10 |
| 1.2 | ΚΑΡΔΙΑΚΟ ΕΠΕΙΣΟΔΙΟ..... | 11 |
| 1.2.1 | ΤΥΠΟΙ ΚΑΡΔΙΑΚΟΥ ΕΠΕΙΣΟΔΙΟΥ..... | 12 |
| 1.2.2 | ΣΥΝΗΘΗ ΣΥΜΠΤΩΜΑΤΑ ΚΑΡΔΙΑΚΟΥ ΕΠΕΙΣΟΔΙΟΥ..... | 13 |
| 1.3 | ΠΑΡΑΓΟΝΤΕΣ ΚΙΝΔΥΝΟΥ ΕΜΦΑΝΙΣΗΣ ΚΑΡΔΙΑΚΩΝ ΕΠΕΙΣΟΔΙΩΝ..... | 13 |
| 1.3.1 | ΤΡΟΠΟΠΟΙΗΣΙΜΟΙ ΠΑΡΑΓΟΝΤΕΣ ΚΙΝΔΥΝΟΥ..... | 14 |
| 1.3.1.1 | ΚΑΠΝΙΣΜΑ..... | 14 |
| 1.3.1.2 | ΚΑΤΑΝΑΛΩΣΗ ΑΛΚΟΟΛ..... | 14 |
| 1.3.1.3 | ΔΙΑΤΡΟΦΙΚΕΣ ΣΥΝΗΘΕΙΕΣ..... | 14 |
| 1.3.1.4 | ΦΥΣΙΚΗ ΔΡΑΣΤΗΡΙΟΤΗΤΑ..... | 15 |
| 1.3.1.5 | ΔΥΣΛΙΠΙΔΑΙΜΙΑ..... | 15 |
| 1.3.2 | ΜΗ ΤΡΟΠΟΠΟΙΗΣΙΜΟΙ ΠΑΡΑΓΟΝΤΕΣ ΚΙΝΔΥΝΟΥ..... | 16 |
| 1.3.2.1 | ΦΥΛΟ..... | 16 |
| 1.3.2.2 | ΗΛΙΚΙΑ..... | 16 |
| 1.3.2.3 | ΚΛΗΡΟΝΟΜΙΚΟΤΗΤΑ..... | 16 |

ΚΕΦΑΛΑΙΟ 2

| | | |
|-------|---|----|
| 2.1 | ΕΙΣΑΓΩΓΗ..... | 17 |
| 2.1.1 | ΠΕΡΙΓΡΑΦΗ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ..... | 17 |
| 2.2 | ΔΗΜΟΓΡΑΦΙΚΗ ΑΝΑΛΥΣΗ..... | 19 |
| 2.2.1 | ΔΗΜΟΓΡΑΦΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΓΙΑ ΤΗΝ 1 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΥ ΑΦΟΡΑ ΑΝΔΡΕΣ ΚΑΙ ΓΥΝΑΙΚΕΣ..... | 19 |
| 2.2.2 | ΔΗΜΟΓΡΑΦΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΓΙΑ ΤΗΝ 1 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΥ ΑΦΟΡΑ ΜΟΝΟ ΑΝΔΡΕΣ..... | 22 |
| 2.3 | ΣΩΜΑΤΙΚΑ ΣΤΟΙΧΕΙΑ ΚΑΙ ΑΘΛΗΤΙΚΗ ΔΡΑΣΤΗΡΙΟΤΗΤΑ..... | 25 |
| 2.3.1 | ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΑ ΣΩΜΑΤΙΚΑ ΣΤΟΙΧΕΙΑ ΑΝΔΡΩΝ ΚΑΙ ΓΥΝΑΙΚΩΝ ΤΗΣ 1 ^{ΗΣ} ΒΑΣΗΣ ΔΕΔΟΜΕΝΩΝ..... | 25 |

| | | |
|-------|--|----|
| 2.3.2 | ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΑ ΣΩΜΑΤΙΚΑ ΣΤΟΙΧΕΙΑ ΑΝΔΡΩΝ 2 ^{ΗΣ} ΒΑΣΗΣ ΔΕΔΟΜΕΝΩΝ..... | 29 |
| 2.4 | ΚΑΠΝΙΣΤΙΚΕΣ ΣΥΝΗΘΕΙΕΣ ΚΑΙ ΚΑΤΑΝΑΛΩΣΗ ΑΛΚΟΟΛ..... | 25 |
| 2.4.1 | ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΗΝ 1 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΥ ΑΦΟΡΑ ΑΝΔΡΕΣ ΚΑΙ ΓΥΝΑΙΚΕΣ..... | 25 |
| 2.4.2 | ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΗΝ 2 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΥ ΑΦΟΡΑ ΜΟΝΟ ΑΝΔΡΕΣ..... | 37 |
| 2.5 | ΚΛΙΝΙΚΑ ΣΤΟΙΧΕΙΑ..... | 40 |
| 2.5.1 | ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΗΝ 1 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΥ ΑΦΟΡΑ ΑΝΔΡΕΣ ΚΑΙ ΓΥΝΑΙΚΕΣ..... | 40 |
| 2.5.2 | ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΤΗΝ 2 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΥ ΑΦΟΡΑ ΜΟΝΟ ΑΝΔΡΕΣ..... | 45 |

ΚΕΦΑΛΑΙΟ 3^Ο

| | | |
|-------|---|----|
| 3.1 | ΕΛΕΓΧΟΣ ΑΝΕΞΑΡΤΗΣΙΑΣ ΜΕΤΑΞΥΠΟΙΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ..... | 52 |
| 3.1.1 | ΠΙΝΑΚΕΣ ΣΥΝΑΦΕΙΑΣ..... | 52 |
| 3.1.2 | ΕΛΕΓΧΟΣ ΑΝΕΞΑΡΤΗΣΙΑΣ Χ ² | 53 |
| 3.2 | ΕΛΕΓΧΟΣ ΣΥΣΧΕΤΙΣΗΣ ΠΟΣΟΤΙΚΗΣ ΜΕΤΑΒΛΗΤΗΣ ΜΕ ΜΙΑ ΠΟΙΟΤΙΚΗ..... | 54 |
| 3.3 | ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΙΣΗ..... | 55 |

ΚΕΦΑΛΑΙΟ 4^Ο

| | | |
|-------|--|----|
| 4.1 | ΕΙΣΑΓΩΓΗ..... | 57 |
| 4.2 | ΣΥΓΚΡΙΣΗ ΟΛΩΝ ΤΩΝ ΜΕΤΑΒΛΗΤΩΝ ΜΕ ΤΗΝ ΥΠΑΡΞΗ ΚΑΡΔΙΑΓΓΕΙΑΚΗΣ ΝΟΣΟΥ ΓΙΑ ΤΗΝ 1 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΥ ΑΦΟΡΑ ΑΝΔΡΕΣ ΚΑΙ ΓΥΝΑΙΚΕΣ..... | 57 |
| 4.2.1 | ΕΛΕΓΧΟΣ ΣΥΣΧΕΤΙΣΗΣ ΠΟΙΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ ΜΕ ΤΗΝ ΜΕΤΒΛΗΤΗ ‘ΚΑΡΔΙΑΓΓΕΙΑΚΗ ΝΟΣΟΣ’..... | 57 |
| 4.2.2 | ΕΛΕΓΧΟΣ ΣΥΣΧΕΤΙΣΗΣ ΠΟΣΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ ΜΕ ΤΗΝ ΜΕΤΒΛΗΤΗ ‘ΚΑΡΔΙΑΓΓΕΙΑΚΗ ΝΟΣΟΣ’..... | 67 |
| 4.3 | ΣΥΓΚΡΙΣΗ ΟΛΩΝ ΤΩΝ ΜΕΤΑΒΛΗΤΩΝ ΜΕ ΤΗΝ ΥΠΑΡΞΗ ΣΤΕΦΑΝΙΑΙΑΣ ΝΟΣΟΥ ΓΙΑ ΤΗΝ 2 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΥ ΑΦΟΡΑ ΜΟΝΟ ΑΝΔΡΕΣ | 73 |
| 4.3.1 | ΕΛΕΓΧΟΣ ΣΥΣΧΕΤΙΣΗΣ ΠΟΙΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ ΜΕ ΤΗΝ ΜΕΤΒΛΗΤΗ ‘ΣΤΕΦΑΝΙΑΙΑ ΝΟΣΟΣ’..... | 73 |

| | |
|---|----|
| 4.3.2 ΕΛΕΓΧΟΣ ΣΥΣΧΕΤΙΣΗΣ ΠΟΣΟΣΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ ΜΕ ΤΗΝ ΜΕΤΒΛΗΤΗ ‘ΣΤΕΦΑΝΙΑΙΑ ΝΟΣΟΣ’..... | 77 |
| 4.4 ΕΛΕΓΧΟΣ ΣΥΣΧΕΤΙΣΗΣ ΤΗΣ ΗΛΙΚΙΑΣ ΤΩΝ ΑΤΟΜΩΝ ΠΟΥ ΕΜΦΑΝΙΣΑΝ ΚΑΡΔΙΑΓΓΕΙΑΚΗ ΝΟΣΟ ΑΝΑΛΟΓΑ ΜΕ ΤΟ ΦΥΛΟ ΤΟΥΣ..... | 84 |

ΚΕΦΑΛΑΙΟ 5^ο

| | |
|---|----|
| 5.1 ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ ΜΕ ΕΞΑΡΤΗΜΕΝΗ ΜΕΤΑΒΛΗΤΗ ΤΗΝ ΚΑΡΔΙΑΓΓΕΙΑΚΗ ΝΟΣΟ ΓΙΑ ΤΗΝ 1 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΣΕ ΑΝΔΡΕΣ ΚΑΙ ΓΥΝΑΙΚΕΣ..... | 86 |
| 5.2 ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ ΜΕ ΕΞΑΡΤΗΜΕΝΗ ΜΕΤΑΒΛΗΤΗ ΤΗΝ ΣΤΕΦΑΝΙΑΙΑ ΝΟΣΟ ΓΙΑ ΤΗΝ 2 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΠΟΥ ΑΦΟΡΑ ΜΟΝΟ ΑΝΔΡΕΣ..... | 91 |
| 5.3 ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ ΜΕ ΕΞΑΡΤΗΜΕΝΗ ΜΕΤΑΒΛΗΤΗ ΤΗΝ ΚΑΡΔΙΑΓΓΕΙΑΚΗ ΝΟΣΟ ΓΙΑ ΤΗΝ 1 ^Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΜΟΝΟ ΣΤΟΥΣ ΑΝΔΡΕΣ..... | 95 |

ΚΕΦΑΛΑΙΟ 6^ο

| | |
|--|-----|
| 6.1 ΕΙΣΑΓΩΓΗ..... | 100 |
| 6.2 ΣΥΜΠΕΡΑΣΜΑΤΑ ΑΠΟ ΠΕΡΙΓΡΑΦΙΚΗ ΑΝΑΛΥΣΗ..... | 100 |
| 6.3 ΣΥΜΠΕΡΑΣΜΑΤΑ ΤΩΝ ΣΥΣΧΕΤΙΣΕΩΝ..... | 101 |
| 6.4 ΣΥΜΠΕΡΑΣΜΑΤΑ ΛΟΓΙΣΤΙΚΗΣ ΠΑΛΙΝΔΡΟΜΙΣΗΣ..... | 101 |

ΒΙΒΛΙΟΓΡΑΦΙΑ

1^ο ΚΕΦΑΛΑΙΟ

ΕΙΣΑΓΩΓΗ

1.1 Εισαγωγή

Στο κεφάλαιο αυτό θα παρουσιάσουμε τις κύριες αιτίες θανάτου σε παγκόσμιο επίπεδο δίνοντας έμφαση στα καρδιακά επεισόδια, τα είδη και τα συμπτώματά τους καθώς επίσης θα ασχοληθούμε και με τους παράγοντες που οδηγούν στην εμφάνιση τους.

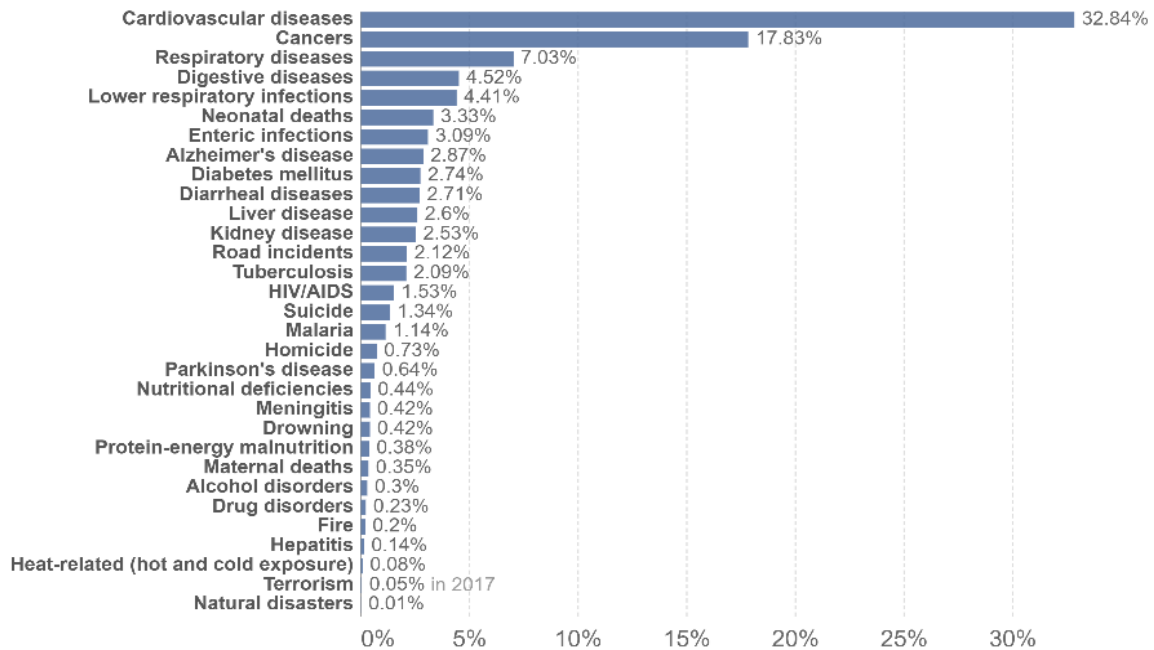
1.1.1 Αιτίες θανάτου σε παγκόσμια κλίμακα

Κύρια αιτία θανάτου αποτελούν τα τελευταία 20 χρόνια οι καρδιακές παθήσεις (Muacevic A. and Adler John R , 2022)σε παγκόσμιο επίπεδο, αντιπροσωπεύοντας περίπου το ένα τρίτο όλων των θανάτων (Ritchie H., Spooner F. and Roser M., 2018).

Οι καρδιακές παθήσεις αφορούν μια σειρά ασθενειών οι οποίες επηρεάζουν την καρδιά και τα αιμοφόρα αγγεία. Οι ασθένειες αυτές περιλαμβάνουν την υπέρταση, δηλαδή την εμφάνιση υψηλής αρτηριακής πίεσης, τη στεφανιαία καρδιακή νόσο η οποία οδηγεί στην πρόκληση καρδιακού επεισοδίου, την εγκεφαλοαγγειακή νόσο δηλαδή τη σοβαρή στένωση των εγκεφαλικών αρτηριών με αποτέλεσμα το εγκεφαλικό επεισόδιο, την καρδιακή ανεπάρκεια και άλλες καρδιακές παθήσεις [πηγή: Βικιπαίδεια: Καρδιαγγειακές παθήσεις [<https://el.wikipedia.org/wiki>]].

Share of deaths by cause, World, 2019

Data refers to the specific cause of death, which is distinguished from risk factors for death, such as air pollution, diet and other lifestyle factors. This is shown by cause of death as the percentage of total deaths.



Source: IHME, Global Burden of Disease (2019)

OurWorldInData.org/causes-of-death • CC BY

Εικόνα 1.1: Διαγραμματική απεικόνιση των συχνότερων αιτιών θανάτου στον κόσμο (Πηγή: IHME, Global Burden Disease(2019))

1.2 Καρδιακό επεισόδιο

Ένα καρδιακό επεισόδιο συμβαίνει όταν η ροή του αίματος προς την καρδιά ξαφνικά ελαττώνεται ή σταματά τελείως με αποτέλεσμα η καρδιά να μην μπορεί να πάρει το οξυγόνο που χρειάζεται (Βικιπαίδεια: Καρδιαγγειακές Παθήσεις). Αν δεν αντιμετωπιστεί γρήγορα, ο μυς της καρδιάς σταματά να συστέλλεται και με την πάροδο του χρόνου αρχίζει προοδευτικά να νεκρώνεται. Το καρδιακό επεισόδιο συμβαίνει ξαφνικά αλλά πρόκειται για μία διεργασία χρόνων η οποία λέγεται αθηροσκλήρωση. Η αθηροσκλήρωση είναι μία συχνή βλάβη και χαρακτηρίζεται από την συσσώρευση λιπαρών ουσιών (αθηρώματα) στα τοιχώματα των αρτηριών γύρω από την καρδιά. Καθώς αυξάνεται η συγκέντρωση αθηρώματος οδηγεί στην δημιουργία αθηρωματικής πλάκας η οποία διογκώνεται με το πέρασμα του χρόνου, μειώνοντας έτσι την εσωτερική διάμετρο των αρτηριών και δυσχεραίνοντας την ροή του αίματος προς τον καρδιακό μυ.

1.2.1 Τύποι καρδιακού επεισοδίου

Το έμφραγμα του μυοκαρδίου γνωστό και ως καρδιακό επεισόδιο είναι το στεφανιαίο σύνδρομο στο οποίο υπάρχει νέκρωση καρδιακού μυ και χωρίζεται σε τρεις τύπους. Το διατοιχωματικό (STEMI), το υπενδοκάρδιο (NSTEMI) και ο στεφανιαίος σπασμός ή ασταθής στηθάγχη.

- Διατοιχωματικό (STEMI)

Όταν οι περισσότεροι άνθρωποι αναφέρονται σέ ένα καρδιακό επεισόδιο συνήθως εννοούν το έμφραγμα τύπου STEMI, δηλαδή το οξύ έμφραγμα του μυοκαρδίου το οποίο μπορεί να διαγνωσθεί μέσω της ανύψωσης του τμήματος ST του καρδιογραφήματος. Αυτός ο τύπος εμφράγματος εμφανίζεται όταν μια στεφανιαία αρτηρία αποφραχθεί πλήρως και μεγάλο μέρος του καρδιακού μυ σταματά να λαμβάνει αίμα. Πρόκειται για ένα σοβαρό έμφραγμα το οποίο μπορεί να προκαλέσει σημαντική βλάβη στην καρδιά.

- Υπενδοκάρδιο (NSTEMI)

Σε αυτό τον τύπο εμφράγματος παρατηρείται μερική στένωση της προσβεβλημένης αρτηρίας και σε αυτή την περίπτωση δεν θα εμφανιστεί καμία αλλαγή στο τμήμα ST του καρδιογραφήματος. Οι εξετάσεις που ακολουθούν ενός εμφράγματος τύπου NSTEMI είναι η στεφανιογραφία η οποία θα μπορέσει να δείξει το βαθμό απόφραξης της αρτηρίας καθώς και η εξέταση αίματος ώστε να ελεγχθεί αν είναι αυξημένα τα επίπεδα της πρωτεΐνης τροπονίνης. Σε αυτό τον τύπο εμφράγματος παρόλο που μπορεί να προκαλέσει μικρότερη βλάβη στην καρδιά εξακολουθεί να αποτελεί μία σοβαρή κατάσταση.

- Στεφανιαίος σπασμός ή ασταθής στηθάγχη

Ο σπασμός της στεφανιαίας αρτηρίας γνωστός και ως ασταθής στηθάγχη ή σιωπηλή καρδιακή προσβολή πρόκειται για μία σύντομη και παροδική σύσπαση σε μία από τις αρτηρίες της καρδιάς η οποία μειώνει τη διάμετρο του αυλού της αρτηρίας με αποτέλεσμα να ελαττώνεται σημαντικά ή να σταματά η ροή του αίματος. Η διάγνωση της σιωπηλής καρδιακής προσβολής μπορεί να γίνει μόνο με τα αποτελέσματα εξετάσεων αίματος και απεικόνισης. Κατά τη διάρκεια του σπασμού της στεφανιαίας αρτηρίας δεν προκαλείται μόνιμη βλάβη στην καρδιά και παρόλο που δεν πρόκειται για μία σοβαρή κατάσταση αυξάνει τον κίνδυνο εμφάνισης καρδιακού επεισοδίου που μπορεί να είναι πιο σοβαρό.

1.2.2 Συνήθη συμπτώματα εμφάνισης ενός καρδιακού επεισοδίου

Τα συμπτώματα ενός καρδιακού επεισοδίου ποικίλουν από άτομο σε άτομο, καθώς κάποιος μπορεί να έχει έντονα συμπτώματα ενώ κάποιος άλλος να αντιμετωπίσει ηπιότερα συμπτώματα ή και καθόλου. Τα πιο κοινά συμπτώματα ενός καρδιακού επεισοδίου είναι τα παρακάτω:

- 1) Το πιο σύνηθες σύμπτωμα είναι η δυσφορία στο στήθος ο οποίος είναι μια αίσθηση σαν πόνος ή πίεση ή σφίξιμο στο στήθος. Ο πόνος αυτός διαρκεί αρκετά λεπτά ή μπορεί να φύγει και να επιστρέψει ξανά.
- 2) Ένα άλλο κλασικό σύμπτωμα του καρδιακού επεισοδίου είναι ο πόνος ή η ενόχληση η οποία εξαπλώνεται από τον ώμο στο χέρι και στην πλάτη συνήθως στην αριστερή πλευρά του σώματος.
- 3) Η δυσκολία στην αναπνοή έπειτα από μια δραστηριότητα η οποία στο παρελθόν δεν προκαλούσε κάποιο πρόβλημα ενδεχομένως είναι ένα σύμπτωμα που θα πρέπει να διερευνηθεί.
- 4) Ο πόνος στο λαιμό ή το σαγόι δεν σχετίζεται με την καρδιά συνήθως όμως αν αυτοί οι πόνοι ακολουθήσουν έπειτα από πόνο στο στήθος είναι πιθανό να πρόκειται για σύμπτωμα ενός καρδιακού εμφράγματος.
- 5) Άλλα συμπτώματα τα οποία μπορεί να συμβούν κατά τη διάρκεια ενός καρδιακού επεισοδίου είναι η ναυτία, ο πόνος στο στομάχι, η ξαφνική ζαλάδα ή η απότομη εφίδρωση τα οποία συνήθως εμφανίζονται συνδυαστικά με κάποιο από τα παραπάνω συμπτώματα.

Τόσο στους άνδρες αλλά και στις γυναίκες το πιο σύνηθες σύμπτωμα της καρδιακής προσβολής είναι ο πόνος στο στήθος ή η δυσφορία, ωστόσο οι γυναίκες ενδέχεται να εμφανίσουν μεμονωμένα ή συνδυαστικά κάποια από τα συμπτώματα που σχετίζονται λιγότερο με την καρδιακή προσβολή όπως δύσπνοια, ναυτία ή εμετό και πόνο στην πλάτη. Κάποια από τα καρδιακά επεισόδια συμβαίνουν ξαφνικά αλλά πολλοί άνθρωποι έχουν προειδοποιητικά σημάδια ή συμπτώματα ώρες, ημέρες ή και εβδομάδες νωρίτερα.

1.3 Παράγοντες κινδύνου εμφάνισης καρδιακών επεισοδίων

Ως παράγοντες κινδύνου ορίζουμε τους παράγοντες οι οποίοι αυξάνουν την πιθανότητα εμφάνισης ενός καρδιακού επεισοδίου και τους κατηγοριοποιούμε σε μη τροποποιήσιμους δηλαδή παράγοντες στους οποίους δεν μπορεί να παρέμβει το άτομο ή ο ιατρός και σε τροποποιήσιμους (Καλιακούδα Μ. ,2018).

1.3.1 Τροποποιήσιμοι παράγοντες κινδύνου

Οι παράγοντες κινδύνου στους οποίους μπορούμε να επέμβουμε είτε εμείς είτε οι θεράποντες ιατροί είναι οι τροποποιήσιμοι παράγοντες και αφορούν τις συνήθειες και τον τρόπο ζωής του ατόμου. Οι σημαντικότεροι τροποποιήσιμοι παράγοντες κινδύνου που οδηγούν στην εμφάνιση ενός καρδιακού επεισοδίου είναι το κάπνισμα, η υπερβολική κατανάλωση αλκοόλ, οι κακές διατροφικές συνήθειες, η έλλειψη σωματικής δραστηριότητας και η δυσλιπιδαιμία. Δες για παράδειγμα Mozaffarian D. et.al (2008).

1.3.1.1 Κάπνισμα

Το κάπνισμα προκαλεί αδιαμφισβήτητα βλαβερές επιπτώσεις στην υγεία του ατόμου και αποτελεί έναν από τους ισχυρότερους παράγοντες κινδύνου εμφάνισης ενός καρδιακού επεισοδίου. Σημαντικό κίνδυνο αντιμετωπίζουν και τα άτομα τα οποία εκτίθενται στο παθητικό κάπνισμα. Παρά τη λανθασμένη άποψη που έχουμε να συνδέουμε τις δυσμενείς επιπτώσεις του καπνίσματος με τον καρκίνο, ένας καπνιστής είναι πιθανότερο να προσβληθεί από ένα καρδιακό επεισόδιο παρά από τον καρκίνο του πνεύμονα. Στο σύνολο των καπνιστών οι γυναίκες φαίνεται να είναι πιο επιρρεπείς στην εμφάνιση καρδιακής προσβολής σε σχέση με τους άνδρες.

1.3.1.2 Κατανάλωση Αλκοόλ

Είναι σαφές ότι η υπερβολική κατανάλωση αλκοόλ με την πάροδο του χρόνου οδηγεί στην εμφάνιση υψηλής αρτηριακής πίεσης (υπέρταση) η οποία αυξάνει το κίνδυνο καρδιακής προσβολής.

1.3.1.3 Διατροφικές Συνήθειες

Οι διατροφικές μας συνήθειες συνδέονται άμεσα με καρδιακές παθήσεις όπως η αθηροσκλήρωση. Η κατανάλωση διατροφής (Micha R. & Mozaffarian D. , 2010) με υψηλή περιεκτικότητα σε κορεσμένα λιπαρά (π.χ. τυρί) και τρανς λιπαρά (π.χ. συσκευασμένα γλυκά και fast food) οδηγεί στην αύξηση της χοληστερόλης επιβαρύνοντας σημαντικά την καρδιά και αυξάνοντας τον κίνδυνο εμφάνισης ενός καρδιακού επεισοδίου. Επίσης, η υπέρμετρη κατανάλωση αλατιού (νάτριο) μπορεί να οδηγήσει στην αύξηση της αρτηριακής πίεσης γεγονός που επιβαρύνει ιδιαίτερα τη λειτουργία του καρδιακού μυ.

1.3.1.4 Φυσική Δραστηριότητα

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (ΠΟΥ) η σωματική αδράνεια αποτελεί τον τέταρτο κύριο παράγοντα κινδύνου για την παγκόσμια θνησιμότητα. Πολλές μελέτες υποστηρίζουν ότι η τακτική άσκηση έχει ευεργετικά αποτελέσματα στην υγεία του ανθρώπου και αποτελεί παράγοντα πρόληψης και θεραπείας των καρδιακών νοσημάτων. Ακόμη και η μέτρια σωματική δραστηριότητα όπως το περπάτημα 3 ώρες την εβδομάδα συμβάλει σημαντικά στη μείωση θνησιμότητας από καρδιακά επεισόδια.

1.3.1.5 Δυσλιπιδαιμία

Η δυσλιπιδαιμία είναι η μη φυσιολογική ποσότητα λιπιδίων στο αίμα, συγκεκριμένα αυξημένα επίπεδα «κακής» LDL χοληστερόλης και τριγλυκεριδίων ή μειωμένα επίπεδα «καλής» HDL χοληστερόλης αποτελούν έναν από τους βασικότερους παράγοντες κινδύνου εμφάνισης ενός καρδιακού επεισοδίου.

Η υψηλή χοληστερίνη στο αίμα ή διαφορετικά υπερλιπιδαιμία αποτελεί ένα σημαντικό παράγοντα κινδύνου ο οποίος αυξάνει τις πιθανότητες εμφάνισης καρδιακών επεισοδίων ή άλλων σοβαρών προβλημάτων υγείας.

Στον παρακάτω πίνακα δίνονται οι τιμές αναφοράς των λιπιδίων σύμφωνα με τις οδηγίες της Ευρωπαϊκής Καρδιολογικής Εταιρείας.

| Παράμετρος | Τιμές Αναφοράς |
|--------------------------|--|
| Ολική Χοληστερόλη | < 170 mg/dL |
| LDL Χοληστερόλη | < 116 mg/dL για άτομα χαμηλού κινδύνου < 100 mg/dL για άτομα μέτριου κινδύνου < 70 mg/dL για άτομα υψηλού κινδύνου < 55 mg/dL για άτομα υψηλού κινδύνου |
| HDL Χοληστερόλη | > 40 mg/dL για άνδρες > 50 mg/dL για γυναίκες |
| Τριγλυκερίδια | < 150 mg/dL |

Πίνακας 1. 1: Τιμές αναφορές των λιπιδίων

1.3.2 Μη τροποποιήσιμοι παράγοντες κινδύνου

Στους μη τροποποιήσιμους παράγοντες κινδύνου ανήκουν οι γενετικοί κίνδυνοι οι οποίοι σχετίζονται με τις γενετικές προδιαθέσεις του ατόμου (κληρονομικό ιστορικό) και οι δημογραφικοί κίνδυνοι είναι εκείνοι που αφορούν το φύλο, την ηλικία και την φυλή του κάθε ατόμου.

1.3.2.1 Φύλο

Παρόλο που η καρδιακή νόσος λανθασμένα θεωρούμε ότι είναι η ασθένεια του άνδρα, τόσες γυναίκες όσοι και άνδρες πεθαίνουν κάθε χρόνο από καρδιακή ανακοπή (Albert, C.M., Chae, C.U. et al., 2003). Ωστόσο, μελέτες υποστηρίζουν ότι είναι λιγότερες οι γυναίκες με τον τύπο της εκτεταμένης αθηροσκλήρωσης που απαιτεί ιατρική παρέμβαση σε σχέση με τους άνδρες.

1.3.2.2 Ηλικία

Έχει παρατηρηθεί ότι οι άνδρες άνω των ετών και οι γυναίκες άνω των 55 ετών είναι πιο πιθανό να υποστούν καρδιακή προσβολή από ότι οι νεότεροι άνδρες και γυναίκες.

1.3.2.3 Κληρονομικότητα

Σε μελέτες που έχουν γίνει για το αν σχετίζεται η εμφάνιση ενός καρδιακού επεισοδίου με το οικογενειακό ιστορικό έχει διαπιστωθεί ότι εάν συγγενής πρώτου βαθμού ενός ατόμου έχει υποστεί καρδιακό έμφραγμα σε ηλικία νεότερη των 55-60 ετών τότε αυξάνεται ο κίνδυνος εμφάνισης σε σύγκριση με αυτόν στον γενικό πληθυσμό.

2^ο ΚΕΦΑΛΑΙΟ

ΠΕΡΙΓΡΑΦΙΚΗ ΑΝΑΛΥΣΗ ΔΕΙΓΜΑΤΩΝ

2.1 Εισαγωγή

Στο κεφάλαιο αυτό θα γίνει παρουσίαση των μεταβλητών που περιέχονται στις βάσεις δεδομένων που θα χρησιμοποιηθούν στην παρούσα εργασία, επίσης θα ασχοληθούμε με την περιγραφική ανάλυση των μεταβλητών και την σύγκριση τους ώστε να καταφέρουμε να δημιουργήσουμε μια σφαιρική εικόνα για τα δεδομένα.

2.1.1 Περιγραφή των βάσεων δεδομένων

- Βάση δεδομένων σε άνδρες και γυναίκες που παρουσίασαν καρδιαγγειακή νόσο.

Τα δεδομένα που θα χρησιμοποιήσουμε στην παρούσα έρευνα από την 1^η βάση δεδομένων αποτελούνται από 65.662 παρατηρήσεις και 12 μεταβλητές οι οποίες είναι οι εξής:

- 1) age: Ηλικία ατόμου σε έτη
- 2) Height: Ύψος ατόμου σε cm
- 3) Weight: Βάρος ατόμου σε kg
- 4) Gender: Φύλο ατόμου
- 5) Systolic blood pressure: Συστολική αρτηριακή πίεση
- 6) Diastolic blood pressure: Διαστολική αρτηριακή πίεση
- 7) Cholesterol: Χοληστερόλη (1: normal, 2: above normal, 3: well above normal)
- 8) Glucose: Γλυκόζη (1: normal, 2: above normal, 3: well above normal)
- 9) Smoking: Κάπνισμα (1: ναι , 0: όχι)
- 10) Alcohol intake: Κατανάλωση αλκοόλ (1: ναι , 0: όχι)
- 11) Physical activity: Φυσική δραστηριότητα (1: ναι , 0: όχι)
- 12) Presence or absence of cardiovascular disease: Εμφάνιση καρδιαγγειακής νόσου (1: ναι , 0: όχι)

- Βάση δεδομένων της μελέτης “Western Collaborative Group Study”

Οι μεταβλητές που θα χρησιμοποιήσουμε για την παρούσα μελέτη από την πρώτη βάση δεδομένων αποτελούν ένα μέρος των μεταβλητών που χρησιμοποιήθηκαν στην προοπτική μελέτη “Western Collaborative Group Study” η οποία ξεκίνησε το 1960 με 3.154 άνδρες ηλικίας 39-59 ετών οι οποίοι εργάζονταν σε 11 εταιρίες της Καλιφόρνια. Αυτή η μελέτη εξέτασε τους παράγοντες κινδύνου εμφάνισης της

στεφανιαίας νόσου σε σχέση με τα άτομα που έχουν προσωπικότητα τύπου A. Η προσωπικότητα τύπου A και B για την ανθρώπινη συμπεριφορά προέκυψε ως παράγωγο της έρευνας που περιγράφει τους δύο αντίθετους τύπους προσωπικότητας και επιχείρησαν οι καρδιολόγοι Meyer Friedman και RH Rosenman να συνδέσουν επιστημονικά τους τύπους συμπεριφοράς με την πιθανότητα εμφάνισης καρδιακών παθήσεων. Σύμφωνα με τη μελέτη ένα άτομο με προσωπικότητα τύπου A εστιάζει στην κοινωνική αποδοχή, στα επαγγελματικά επιτεύγματα ενώ συχνά είναι αυταρχικό, ανυπόμονο και επιρρεπής σε συναισθήματα οργής. Αντίθετα ένα άτομο με προσωπικότητα τύπου B είναι πιο ανεκτικό, ήρεμο, δεν έχει άγχος και επίσης δηλώνει υψηλότερα επίπεδα ικανοποίησης στη ζωή.

Τα δεδομένα που θα χρησιμοποιήσουμε στην παρούσα έρευνα από την 2^η βάση δεδομένων αποτελούνται από 3.155 παρατηρήσεις και 12 μεταβλητές οι οποίες είναι οι εξής:

- 1) age : Ηλικία ατόμου σε έτη
- 2) bmi : Δείκτης μάζας σώματος
- 3) chd69 : Στεφανιαία νόσος (1: ναι , 2: όχι)
- 4) chol : Χοληστερίνη ορού νηστείας
- 5) dbp : Διαστολική αρτηριακή πίεση (ΔΑΠ)
- 6) sbp : Συστολική αρτηριακή πίεση (ΣΑΠ)
- 7) dibpat : Συμπεριφορά (1: Τύπος A , 2: Τύπος B)
- 8) height : Ύψος ατόμου
- 9) ncigs : Αριθμός τσιγάρων την ημέρα
- 10) smoke : Κάπνισμα (1: ναι , 2: όχι)
- 11) typchd69 : Τύπος στεφανιαίας νόσου (1: έμφραγμα ή θάνατος, 2: σιωπηλό έμφραγμα, 3: στηθάγχη)
- 12) weight : Βάρος ατόμου

Ο Δείκτης Μάζας Σώματος γνωστός και ως BMI, αποτελεί έναν ιατρικό δείκτη ο οποίος υπολογίζεται διαιρώντας το βάρος (σε κιλά) του ατόμου με το τετράγωνο του ύψους (σε μέτρα)του. Επιπλέον, εάν θέλαμε να δημιουργήσουμε μια μεταβλητή η οποία να συνδυάζει τις διατροφικές συνήθειες του ατόμου και την φυσική του δραστηριότητα ο ΔΜΣ αποτελεί την πλέον καταλληλότερη.

Οι κατηγορίες στις οποίες κατατάσσουμε τα άτομα ανάλογα με το ΔΜΣ που έχουν είναι λιποβαρείς, φυσιολογικό βάρος, υπέρβαρη και παχύσαρκοι.

Οι κατηγορίες του δείκτη μάζας σώματος (BMI) φαίνονται στον παρακάτω πίνακα.

| BMI (Kg / m ²) | Κλινική εκτίμηση θρέψης |
|----------------------------|-------------------------|
| < 18,5 | Λιποβαρής |
| 18,5 – 24,9 | Υγιές βάρος |
| 25 – 29,9 | Υπέρβαρος |
| ≥ 30 | Παχύσαρκος |

Πίνακας 2. 1: Κατηγοριοποίηση της μεταβλητής του δείκτη μάζας σώματος (ΔΜΣ)

2.2 Δημογραφική Ανάλυση

Θα παρουσιάσουμε την στατιστική ανάλυση για την ηλικία των ατόμων και στις δύο βάσεις δεδομένων όπου στην 1^η αφορά τόσο των ανδρών όσο και των γυναικών ενώ στη δεύτερη μόνο των ανδρών. Επιπλέον, θα δούμε τα στατιστικά στοιχεία για τον τύπο προσωπικότητας των ανδρών της 2^{ης} βάσης δεδομένων.

2.2.1 Δημογραφικά χαρακτηριστικά για την 1^η βάση δεδομένων που αφορά άνδρες και γυναίκες

- Ηλικία

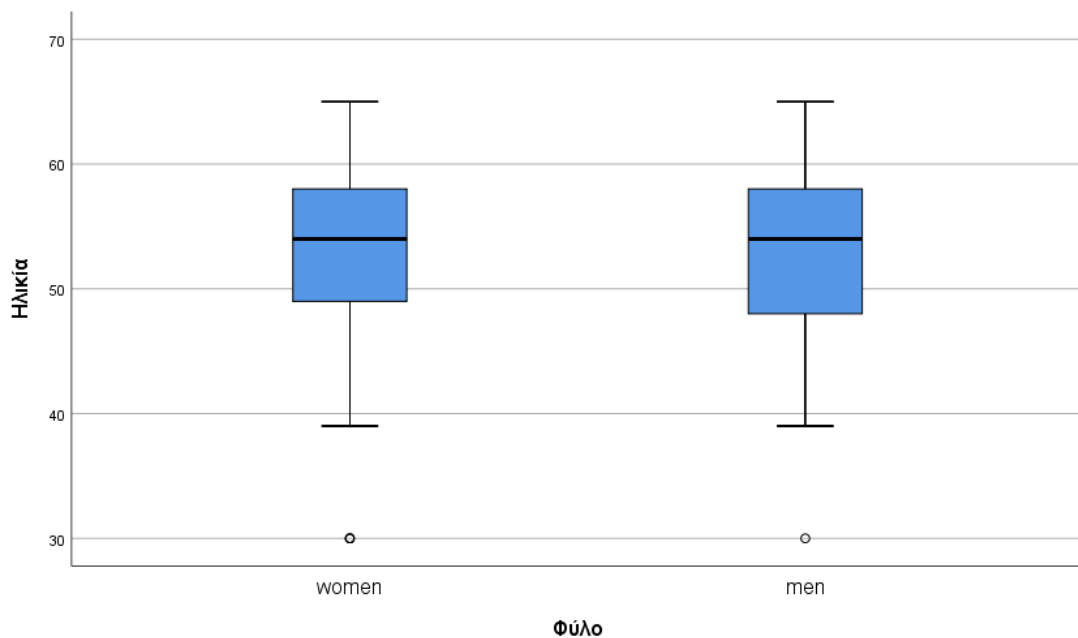
Από το δείγμα των 65.662 ατόμων της βάσης δεδομένων, παρατηρούμε στον πίνακα που ακολουθεί ότι το ηλικιακό εύρος κυμαίνεται μεταξύ 30 έως 65 ετών. Επίσης, ο μέσος όρος ηλικίας είναι $53,33 \pm 6,76$ έτη και η διάμεσος του δείγματος είναι στα 54 έτη.

Παρατηρούμε ότι ο συντελεστής ασυμμετρίας (skewness = -0.304) έχει αρνητικό πρόσημο άρα έχουμε ελαφρώς αρνητική ασυμμετρία που σημαίνει ότι τα περισσότερα άτομα του δείγματος έχουν ηλικία μεγαλύτερη από εκείνη του μέσου όρου. Το παραπάνω συμπέρασμα επιβεβαιώνεται και από το διάγραμμα όπου παρατηρούμε ότι η αριστερή ουρά είναι μεγαλύτερη από την δεξιά

Statistics

| | | |
|------------------------|---------|-------|
| Ηλικία | | |
| N | Valid | 65662 |
| | Missing | 0 |
| Mean | | 53,33 |
| Median | | 54,00 |
| Std. Deviation | | 6,764 |
| Skewness | | -,304 |
| Std. Error of Skewness | | ,010 |
| Kurtosis | | -,822 |
| Std. Error of Kurtosis | | ,019 |
| Minimum | | 30 |
| Maximum | | 65 |

Πίνακας 2. 2: Περιγραφικά μέτρα για την μεταβλητή “ηλικία”



Διάγραμμα 2. 1: Θηκογράμματα της μεταβλητής “ηλικία” ανά φύλο

Στα παραπάνω θηκογράμματα παρατηρούμε ότι τόσο στις γυναίκες όσο και στους άνδρες του συνόλου δεδομένων μας η κατανομή της μεταβλητής της ηλικίας δεν είναι συμμετρική. Παρατηρούμε ότι η διακύμανση είναι μεγαλύτερη στους άνδρες σε σχέση

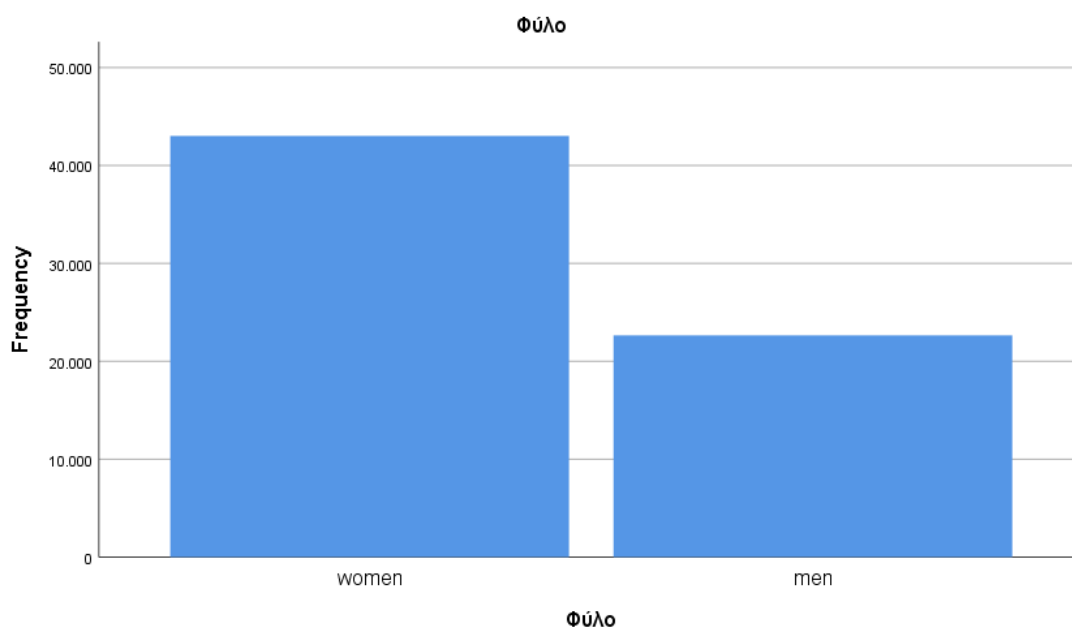
με τις γυναίκες καθώς, το σύνολο των ανδρών περιέχει μικρότερες τιμές από εκείνο των γυναικών, δηλαδή στο σύνολο μας περιέχονται άνδρες με νεαρότερη ηλικία από εκείνη των γυναικών. Η μέγιστη ηλικία φαίνεται να είναι παρόμοια τόσο σε άνδρες όσο και σε γυναίκες

- Φύλο

Παρατηρούμε παρακάτω στον πίνακα συχνοτήτων που αφορά το φύλο του συνόλου των ατόμων καθώς και στο ραβδόγραμμα συχνοτήτων που ακολουθεί το μεγαλύτερο ποσοστό το οποίο αποτελεί το 65,5% του συνόλου δηλαδή, 43.015 από τα 65.662 άτομα είναι γυναίκες. Αντίστοιχα το 34,5% του δείγματος το οποίο αποτελείται από 22.647 άτομα είναι άνδρες.

| Φύλο | | | | | |
|-------|-------|-----------|---------|---------------|--------------------|
| | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | women | 43015 | 65,5 | 65,5 | 65,5 |
| | men | 22647 | 34,5 | 34,5 | 100,0 |
| | Total | 65662 | 100,0 | 100,0 | |

Πίνακας 2. 3: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “Φύλο”



Διάγραμμα 2. 2: Ραβδόγραμμα συχνοτήτων για την μεταβλητή “Φύλο”

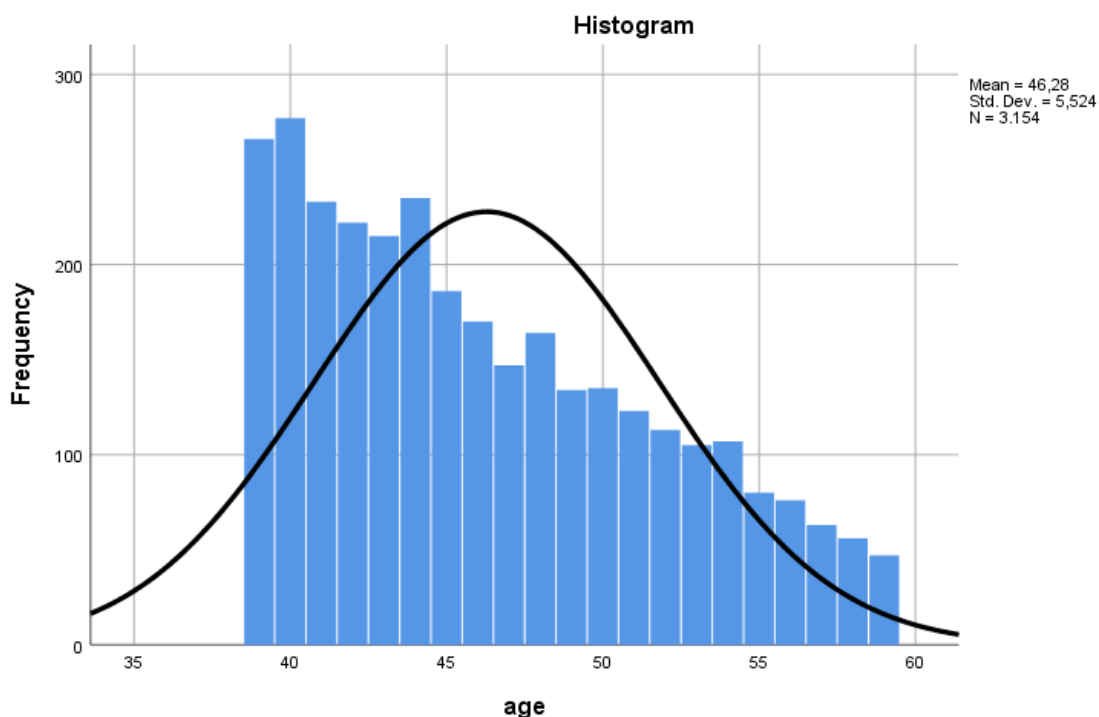
2.2.2 Δημογραφικά χαρακτηριστικά για την 2^η βάση δεδομένων που αφορά μόνο άνδρες

- Ηλικία

Με βάση το δείγμα των 3.154 ανδρών της 1^{ης} βάσης δεδομένων στον πίνακα 2.1 που ακολουθεί παρατηρούμε ότι το εύρος της ηλικίας κυμαίνεται στα 39 – 59 έτη. Επιπλέον, ο μέσος όρος ηλικίας είναι $46,28 \pm 5,52$ και η διάμεσος του δείγματος βρέθηκε στα 45 έτη. Επιπλέον, παρατηρούμε ότι ο συντελεστής ασυμμετρίας (skewness = 0.527) έχει θετικό πρόσημο δηλαδή έχουμε ελαφρώς θετική ασυμμετρία που σημαίνει ότι οι περισσότεροι άνδρες του δείγματός μας έχουν ηλικία μικρότερη από εκείνη του μέσου όρου. Το παραπάνω συμπέρασμα επιβεβαιώνεται και από το ιστόγραμμα που ακολουθεί όπου παρατηρούμε ότι η δεξιά ουρά είναι μεγαλύτερη από την αριστερή.

| | | Statistics |
|------------------------|---------|------------|
| Ηλικία | | |
| N | Valid | 3154 |
| | Missing | 0 |
| Mean | | 46,28 |
| Std. Error of Mean | | ,098 |
| Median | | 45,00 |
| Std. Deviation | | 5,524 |
| Variance | | 30,515 |
| Skewness | | ,527 |
| Std. Error of Skewness | | ,044 |
| Kurtosis | | -,772 |
| Std. Error of Kurtosis | | ,087 |
| Range | | 20 |
| Minimum | | 39 |
| Maximum | | 59 |

Πίνακας 2. 4: Περιγραφικά μέτρα για την μεταβλητή “ηλικία” των ανδρών



Διάγραμμα 2. 3: Διάγραμμα συχνοτήτων της μεταβλητής “ηλικία” στους άνδρες

- Τύπος Προσωπικότητας

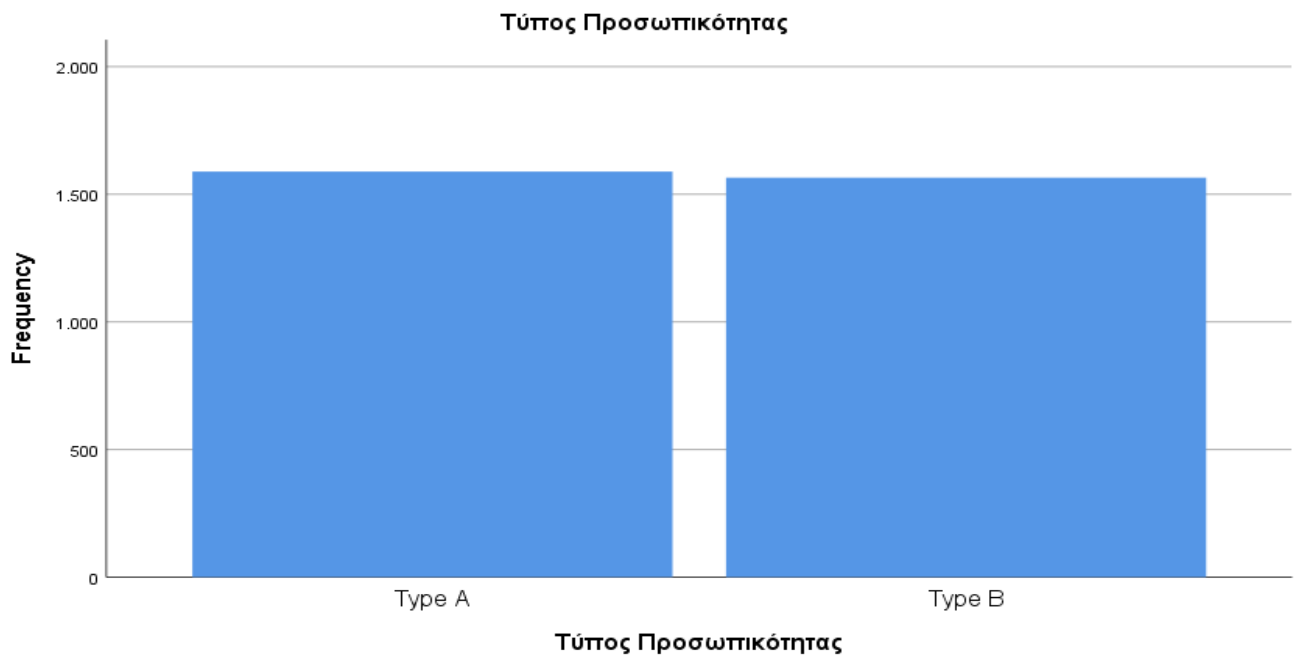
Η έννοια της προσωπικότητας τύπου A και τύπου B περιγράφει δύο διαφορετικούς τύπους προσωπικότητας. Τα άτομα με προσωπικότητες που μπορεί να είναι πιο ανταγωνιστικές, πολύ οργανωμένες, φιλόδοξες και επιθετικές χαρακτηρίζονται ως Τύπος A, ενώ εκείνα με προσωπικότητες πιο χαλαρές, υπομονετικές και ευέλικτες χαρακτηρίζονται Τύπου B.

Όπως μπορούμε να παρατηρήσουμε τόσο στον πίνακα συχνοτήτων που αφορά τις δύο πιθανές προσωπικότητες των ανδρών όσο και στο ραβδόγραμμα συχνοτήτων που ακολουθεί τα ποσοστά των ανδρών του δείγματος μας είναι σχεδόν ίσα, Πιο συγκεκριμένα το 50,4% των ανδρών, δηλαδή 1.589 από τους 3.154 ανήκουν στην κατηγορία προσωπικότητας τύπου A και το 49,6%, δηλαδή 1.565 άνδρες στην προσωπικότητα τύπου B.

Τύπος Προσωπικότητας

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|--------|-----------|---------|---------------|--------------------|
| Valid | Type A | 1589 | 50,4 | 50,4 | 50,4 |
| | Type B | 1565 | 49,6 | 49,6 | 100,0 |
| | Total | 3154 | 100,0 | 100,0 | |

Πίνακας 2. 5: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “Προσωπικότητας” στους άνδρες



Διάγραμμα 2. 4: Ραβδόγραμμα συχνοτήτων για την μεταβλητή “Προσωπικότητα” των ανδρών

2.3 Σωματικά στοιχεία και αθλητική δραστηριότητα

Σε αυτή την ενότητα θα παρουσιάσουμε τα στατιστικά στοιχεία που αφορούν το βάρος και το δείκτη μάζας σώματος (ΔΜΣ) καθώς επίσης και για τη φυσική δραστηριότητα των ατόμων.

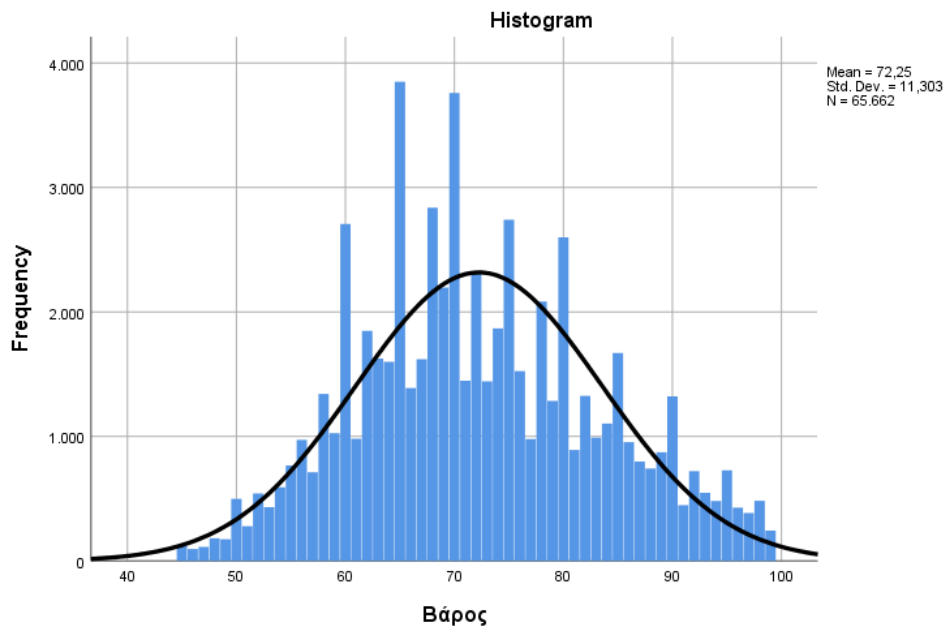
2.3.1 Αποτελέσματα για τα σωματικά στοιχεία ανδρών και γυναικών της 1^{ης} βάσης δεδομένων.

- Βάρος

Στον πίνακα 2.6 παραπάνω παρατηρούμε ότι το μέσο βάρος των ατόμων που συμμετείχαν στην έρευνα είναι $72,25 \pm 11,303$ κιλά και το μέγιστο βάρος είναι στα 99 κιλά ενώ το ελάχιστο στα 45 κιλά. Στο δείγμα υπάρχει θετική ασυμμετρία ($skewness = 0,227$), το οποίο σημαίνει ότι οι περισσότερες παρατηρήσεις του δείγματος βρίσκονται δεξιά της μέσης τιμής, δηλαδή το μεγαλύτερο πλήθος των ατόμων του δείγματος έχει βάρος άνω των 72,25 κιλών.

| Statistics | | |
|------------------------|---------|--------|
| Βάρος | | |
| N | Valid | 65662 |
| | Missing | 0 |
| Mean | | 72,25 |
| Median | | 71,00 |
| Std. Deviation | | 11,303 |
| Skewness | | ,227 |
| Std. Error of Skewness | | ,010 |
| Kurtosis | | -,504 |
| Std. Error of Kurtosis | | ,019 |
| Minimum | | 45 |
| Maximum | | 99 |

Πίνακας 2. 6: Περιγραφικά μέτρα για την μεταβλητή “βάρος”



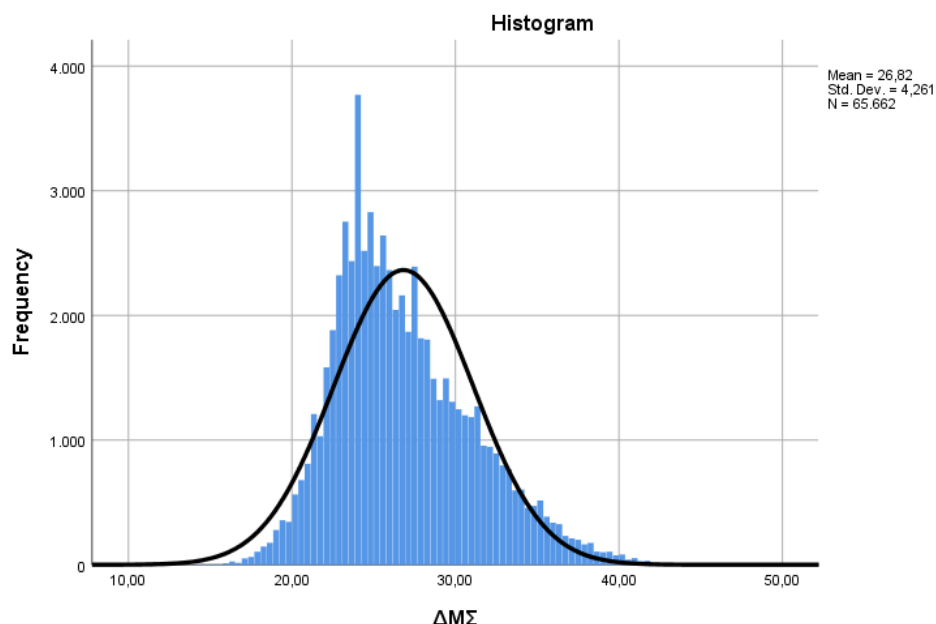
Διάγραμμα 2. 5: Ιστόγραμμα συχνοτήτων της μεταβλητής “βάρος”

- Δείκτης Μάζας Σώματος (ΔΜΣ)

| Statistics | | |
|------------------------|---------|---------|
| ΔΜΣ | | |
| N | Valid | 65662 |
| | Missing | 0 |
| Mean | | 26,8236 |
| Median | | 26,0400 |
| Std. Deviation | | 4,26082 |
| Skewness | | ,658 |
| Std. Error of Skewness | | ,010 |
| Kurtosis | | ,222 |
| Std. Error of Kurtosis | | ,019 |
| Minimum | | 13,76 |
| Maximum | | 42,92 |

Πίνακας 2. 7: Περιγραφικά μέτρα για την μεταβλητή “ΔΜΣ”

Παρατηρούμε ότι η μέση τιμή του ΔΜΣ 26,82 kg/m² με τυπική απόκλιση 4,26 kg/m², ενώ η μέγιστη τιμή για τον ΔΜΣ είναι 42,92 kg/m² και η ελάχιστη 13,76 kg/m² δηλαδή στο δείγμα μας υπάρχουν λιποβαρή αλλά και υπέρβαρα άτομα. Επίσης, στο ιστόγραμμα παρατηρούμε ότι στο δείγμα η κατανομή της μεταβλητής του ΔΜΣ είναι θετικά ασύμμετρη, (skewness = 0,658), δηλαδή για τα περισσότερα άτομα του δείγματος η τιμή του ΔΜΣ είναι πάνω από το μέσο όρο.



Διάγραμμα 2. 6: Ιστόγραμμα συχνοτήτων της μεταβλητής "ΔΜΣ"

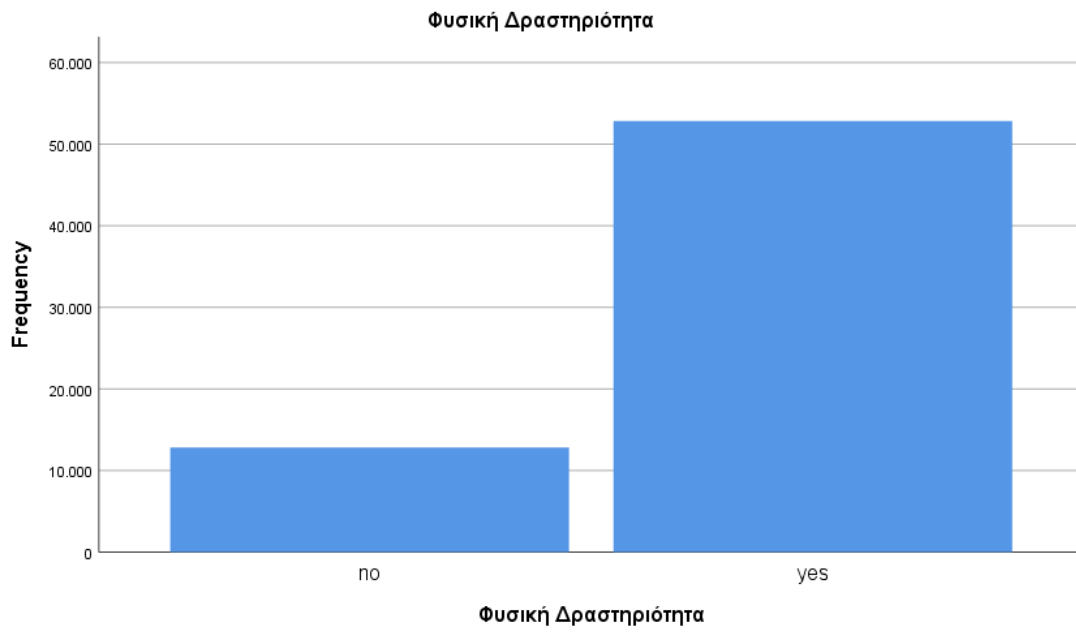
- Φυσική Δραστηριότητα

Παρατηρούμε παρακάτω στον πίνακα που αφορά τη μεταβλητή για την σωματική άσκηση του συνόλου των ατόμων καθώς και στο ραβδόγραμμα συχνοτήτων που ακολουθεί το μεγαλύτερο ποσοστό το οποίο αποτελεί το 80,5% δηλαδή, 52.834 από τα 65.662 άτομα γυμνάζονται. Ενώ μόλις το 19,5% του δείγματος το οποίο αποτελείται από 12.828 άτομα απάντησαν ότι δεν έχουν καμία σωματική δραστηριότητα.

Φυσική Δραστηριότητα

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|----------|-----------|---------|---------------|--------------------|
| Valid no | 12828 | 19,5 | 19,5 | 19,5 |
| yes | 52834 | 80,5 | 80,5 | 100,0 |
| Total | 65662 | 100,0 | 100,0 | |

Πίνακας 2. 8: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής "Φυσική Δραστηριότητα"



Διάγραμμα 2. 7 :Ραβδόγραμμα συχνοτήτων για την μεταβλητή “Φυσική Δραστηριότητα”

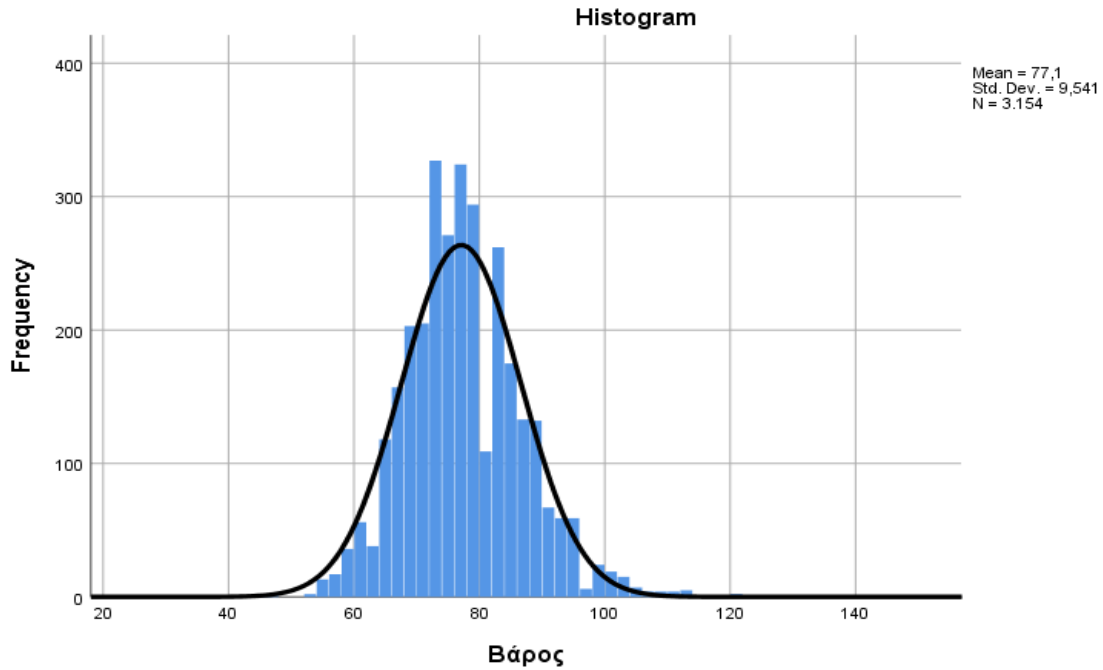
2.3.2 Αποτελέσματα για τα σωματικά στοιχεία της 2^{ης} βάσης που αφορά άνδρες

- Βάρος

Στον παρακάτω πίνακα παρατηρούμε ότι ο μέσος όρος βάρους των ανδρών είναι 77,1 ± 9,54 κιλά με μέγιστο βάρος στα 145 κιλά και ελάχιστο στα 35. Επιπλέον, παρατηρούμε ότι στη μεταβλητή του βάρους υπάρχει θετική ασυμμετρία (skewness = 0,607) το οποίο σημαίνει ότι οι περισσότερες παρατηρήσεις βρίσκονται αριστερά του μέσου όρου, δηλαδή οι περισσότεροι άντρες του δείγματός μας έχουν βάρος λιγότερο των 77 κιλών.

| Statistics | | |
|------------------------|---------|--------|
| Βάρος | | |
| N | Valid | 3154 |
| | Missing | 0 |
| Mean | | 77,10 |
| Median | | 77,00 |
| Std. Deviation | | 9,541 |
| Variance | | 91,029 |
| Skewness | | ,607 |
| Std. Error of Skewness | | ,044 |
| Kurtosis | | 1,965 |
| Std. Error of Kurtosis | | ,087 |
| Minimum | | 35 |
| Maximum | | 145 |

Πίνακας 2. 9: Περιγραφικά μέτρα για την μεταβλητή “βάρος” των ανδρών



Διάγραμμα 2. 8: Ιστόγραμμα συχνοτήτων της μεταβλητής "βάρους" στους άνδρες

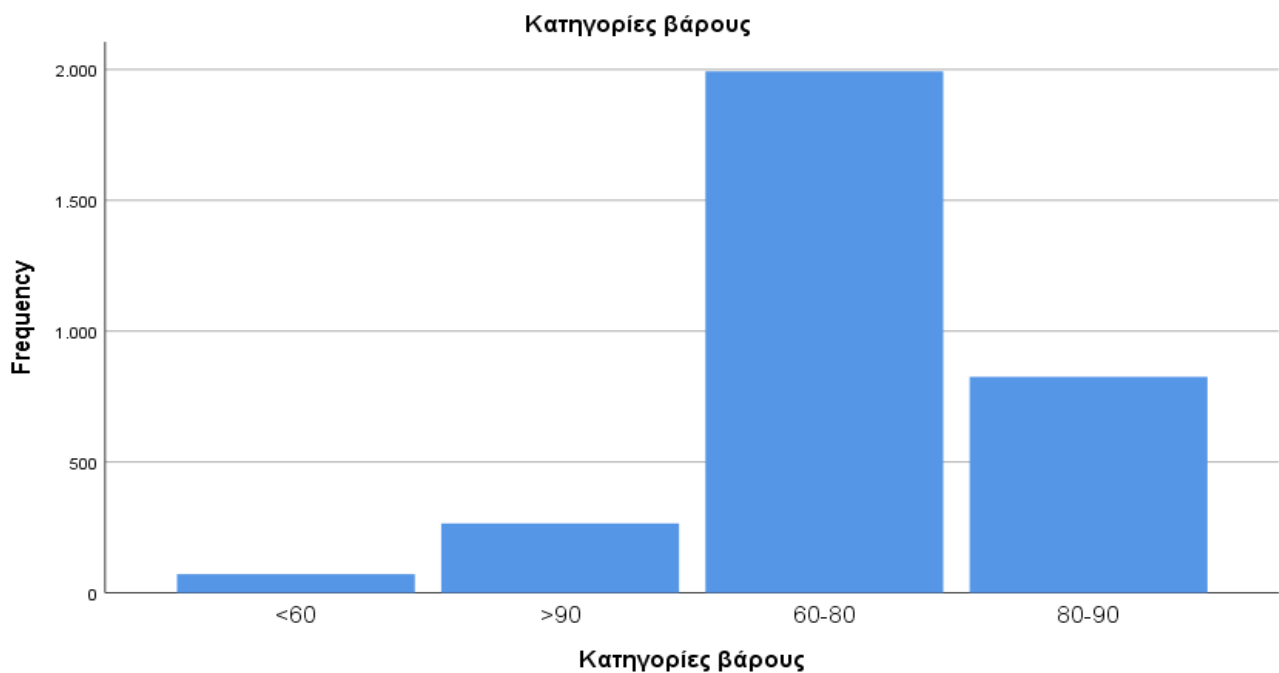
Για την καλύτερη ανάλυση του δείγματος των αντρών δημιουργήσαμε μία μεταβλητή κατηγοριοποίησης του βάρους με τις εξής τέσσερις κατηγορίες:

- < 60: οι άνδρες με βάρος μικρότερο των 60 κιλών
- > 90: οι άνδρες με βάρος μεγαλύτερο των 90 κιλών
- 60-80: οι άνδρες με βάρος μεταξύ 60 και 80 κιλών
- 80-90: οι άνδρες με βάρος μεταξύ 80 και 90 κιλών

Κατηγορίες βάρους

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|-------|-----------|---------|---------------|--------------------|
| Valid | <60 | 71 | 2,3 | 2,3 | 2,3 |
| | >90 | 265 | 8,4 | 8,4 | 10,7 |
| | 60-80 | 1993 | 63,2 | 63,2 | 73,8 |
| | 80-90 | 825 | 26,2 | 26,2 | 100,0 |
| Total | | 3154 | 100,0 | 100,0 | |

Πίνακας 2. 10: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής "Κατηγορίες Βάρους" στους άνδρες



Διάγραμμα 2. 9: Ραβδόγραμμα συχνοτήτων για την μεταβλητή “ Κατηγορίες Βάρους ” των ανδρών

Από τον πίνακα συχνοτήτων καθώς και από το ραβδόγραμμα παραπάνω παρατηρούμε ότι το μεγαλύτερο ποσοστό το οποίο αποτελεί το 63,2% του δείγματος δηλαδή 1.993 άνδρες από τους 3.154 έχουν βάρος μεταξύ 60 και 80 κιλών, το 26,2%, δηλαδή 825 άνδρες ζυγίζουν μεταξύ 80 και 90 κιλών ενώ 8,4% δηλαδή 265 από τους 3.154 άνδρες ζυγίζουν περισσότερο από 90 κιλά. Τέλος το μικρότερο ποσοστό του δείγματος μας το οποίο αντιστοιχεί στο 2,3%, δηλαδή 71 άνδρες έχουν βάρος κάτω από 60 κιλά

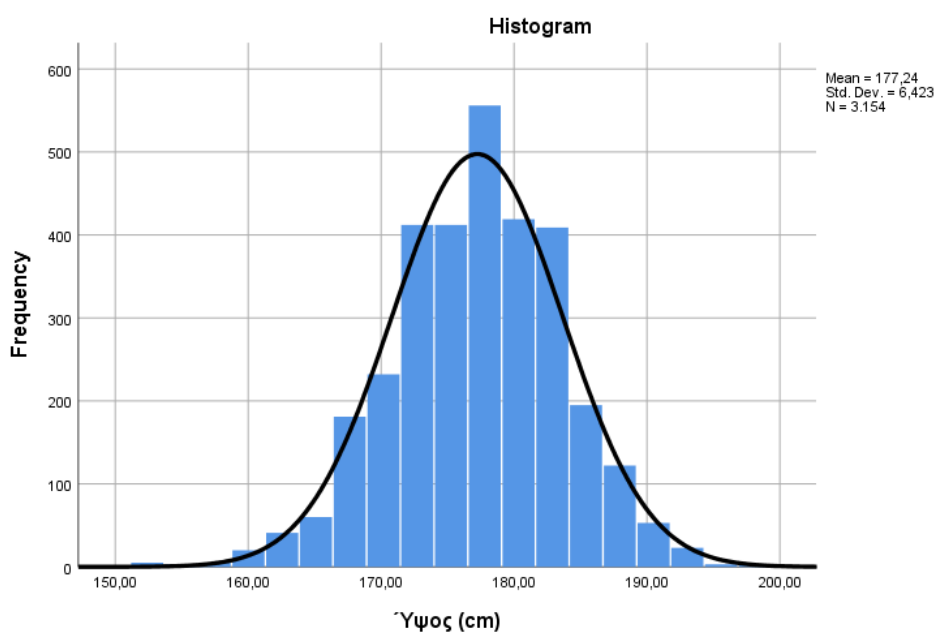
- Ύψος

Στον παρακάτω πίνακα παρατηρούμε ότι η μέση τιμή του ύψους του δείγματος των ανδρών σε εκατοστά είναι $177,24 \pm 6,4$ με το μέγιστο ύψος τους να φτάνει στα 198,12 εκατοστά και το ελάχιστο στα 152,40 εκατοστά. Όπως φαίνεται στον πίνακα αλλά και το ιστόγραμμα παρακάτω για την μεταβλητή του ύψους των ανδρών ο συντελεστής ασυμμετρίας είναι αρνητικός ($skewness = -0.199$) το οποίο σημαίνει ότι το μεγαλύτερο ποσοστό του δείγματος των ανδρών βρίσκεται δεξιά του μέσου όρου του ύψους, δηλαδή οι περισσότεροι άνδρες του δείγματός μας έχουν ύψος μεγαλύτερο των 177,23 εκατοστών.

Statistics

| Ύψος (cm) | | |
|------------------------|---------|----------|
| N | Valid | 3154 |
| | Missing | 0 |
| Mean | | 177,2355 |
| Median | | 177,8000 |
| Std. Deviation | | 6,42288 |
| Variance | | 41,253 |
| Skewness | | -,199 |
| Std. Error of Skewness | | ,044 |
| Kurtosis | | ,280 |
| Std. Error of Kurtosis | | ,087 |
| Minimum | | 152,40 |
| Maximum | | 198,12 |

Πίνακας 2. 11: Περιγραφικά μέτρα για την μεταβλητή “βάρος” των ανδρών



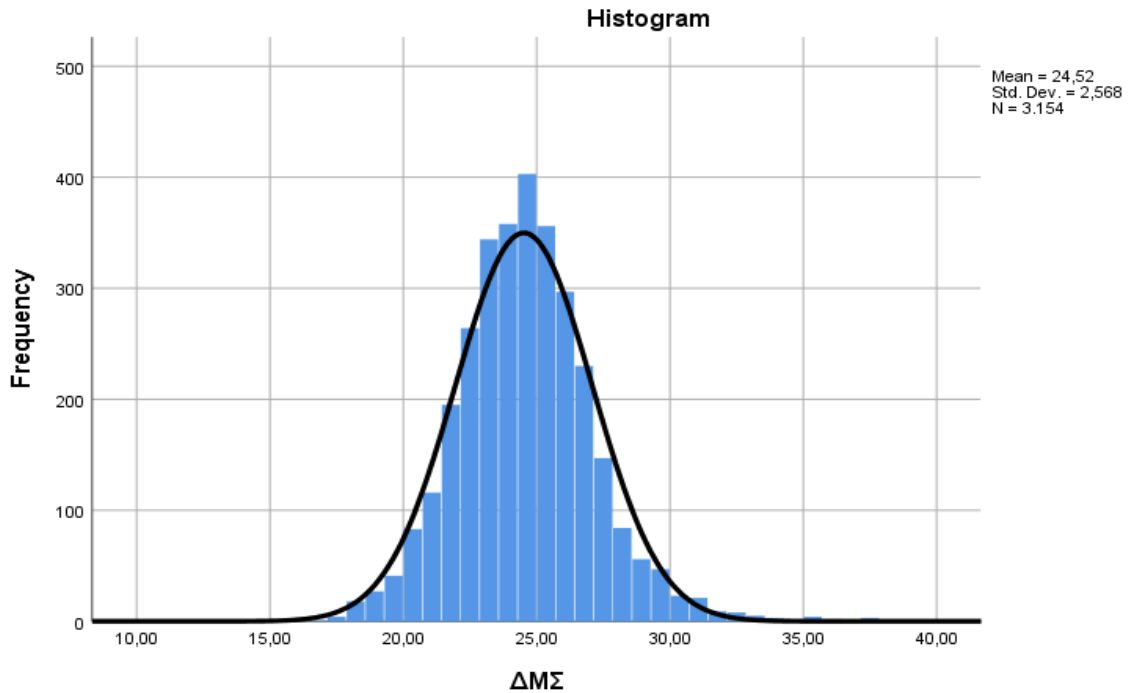
Διάγραμμα 2. 10: Ιστόγραμμα συχνοτήτων της μεταβλητής “ύψους” στους άνδρες

- Δείκτης Μάζας Σώματος (ΔΜΣ)

Παρατηρείται στον παρακάτω πίνακα ότι η μέση τιμή του ΔΜΣ στους άνδρες 24,51kg/m² με τυπική απόκλιση 2,57 kg/m². Η μέγιστη τιμή για τον ΔΜΣ είναι 38,95 kg/m² και η ελάχιστη 11,19 kg/m² δηλαδή στο δείγμα μας υπάρχουν λιποβαρείς αλλά και υπέρβαροι άνδρες. Επίσης, στο ιστόγραμμα παρατηρούμε ότι στο δείγμα η κατανομή της μεταβλητής του ΔΜΣ είναι θετικά ασύμμετρη, (skewness=0,533), δηλαδή για τους περισσότερους άνδρες του δείγματος η τιμή του ΔΜΣ είναι πάνω από το μέσο όρο.

| Statistics | | |
|------------------------|---------|---------|
| ΔΜΣ | | |
| N | Valid | 3154 |
| | Missing | 0 |
| Mean | | 24,5186 |
| Median | | 24,3900 |
| Std. Deviation | | 2,56755 |
| Variance | | 6,592 |
| Skewness | | ,533 |
| Std. Error of Skewness | | ,044 |
| Kurtosis | | 1,980 |
| Std. Error of Kurtosis | | ,087 |
| Minimum | | 11,19 |
| Maximum | | 38,95 |

Πίνακας 2. 12: Περιγραφικά μέτρα για την μεταβλητή “ΔΜΣ” των ανδρών



Διάγραμμα 2. 11: Ιστόγραμμα συχνοτήτων της μεταβλητής “ΔΜΣ” στους άνδρες

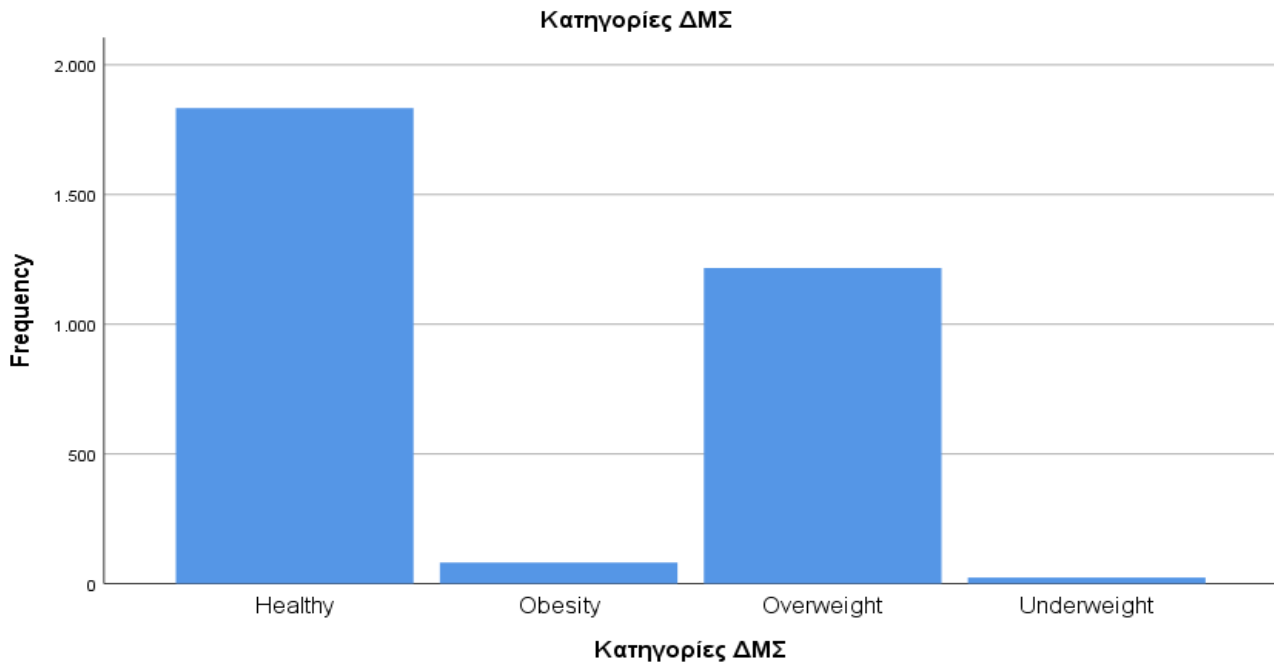
- Κατηγοριοποίηση του Δείκτη Μάζας Σώματος

Παρατηρούμε ότι το 58,1% των ανδρών, δηλαδή οι 1.833 άνδρες από τους 3.154 του δείγματός μας έχουν φυσιολογικό βάρος με βάση τον δείκτη μάζας σώματος. Το 38,6% του δείγματος ανήκει στην κατηγορία των υπέρβαρων, δηλαδή 1.217 άνδρες και μόλις το 0,7%, δηλαδή 23 άνδρες είναι λιποβαρείς. Επίσης, το 2,6% των ανδρών δηλαδή 81 από τους 3.154 άνδρες βρίσκονται στην κατηγορία των παχύσαρκων.

Κατηγορίες ΔΜΣ

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|-------------|-----------|---------|---------------|--------------------|
| Valid | Healthy | 1833 | 58,1 | 58,1 | 58,1 |
| | Obesity | 81 | 2,6 | 2,6 | 60,7 |
| | Overweight | 1217 | 38,6 | 38,6 | 99,3 |
| | Underweight | 23 | ,7 | ,7 | 100,0 |
| | Total | 3154 | 100,0 | 100,0 | |

Πίνακας 2. 13: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “Κατηγορίες ΔΜΣ” στους άνδρες



Διάγραμμα 2. 12: Ραβδόγραμμα συχνοτήτων για την μεταβλητή “Κατηγορίες ΔΜΣ” των ανδρών

2.4 Καπνιστικές συνήθειες και κατανάλωση αλκοόλ

2.4.1 Αποτελέσματα για την 1^η βάση δεδομένων που αφορά άνδρες και γυναίκες

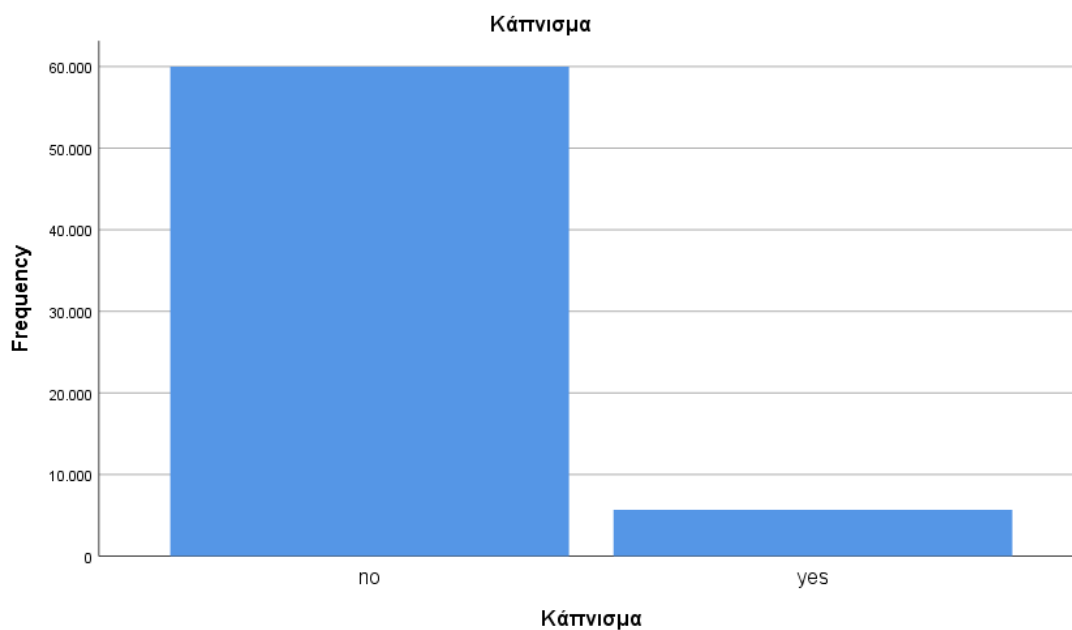
- Κάπνισμα

Παρατηρούμε ότι το δείγμα μας που αποτελείται από 65.662 άτομα εκ των οποίων οι 59.992 δεν είναι καπνιστές και αποτελούν το 91,4% του συνόλου ενώ οι 5.670 καπνίζουν, δηλαδή το 8,6% του συνολικού δείγματος.

Κάπνισμα

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|----------|-----------|---------|---------------|--------------------|
| Valid no | 59992 | 91,4 | 91,4 | 91,4 |
| yes | 5670 | 8,6 | 8,6 | 100,0 |
| Total | 65662 | 100,0 | 100,0 | |

Πίνακας 2. 14: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “ Κάπνισμα”



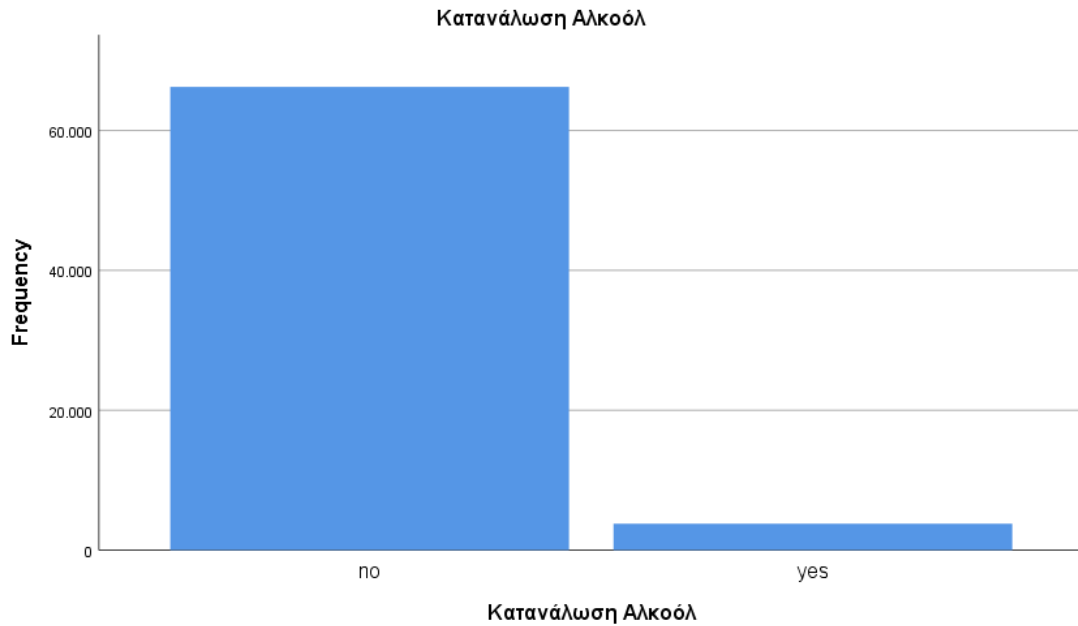
Διάγραμμα 2. 13: Ραβδόγραμμα συχνοτήτων για την μεταβλητή “Κάπνισμα”

- Κατανάλωση αλκοόλ

Παρατηρούμε ότι το 94,8% του δείγματος μας, δηλαδή 62.238 από τα 65.662 άτομα δεν καταναλώνουν αλκοόλ ενώ το 5,2%, δηλαδή 3.424 άτομα καταναλώνουν.

| Αλκοόλ | | | | | |
|--------|-------|-----------|---------|---------------|--------------------|
| | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | no | 62238 | 94,8 | 94,8 | 94,8 |
| | yes | 3424 | 5,2 | 5,2 | 100,0 |
| | Total | 65662 | 100,0 | 100,0 | |

Πίνακας 2. 15: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “Κατανάλωση Αλκοόλ”



Διάγραμμα 2. 14: Ραβδόγραμμα συχνοτήτων για την μεταβλητή “Κατανάλωσης Αλκοόλ”

2.4.1 Αποτελέσματα για την 2^η βάση δεδομένων που αφορά μόνο άνδρες

- Κάπνισμα

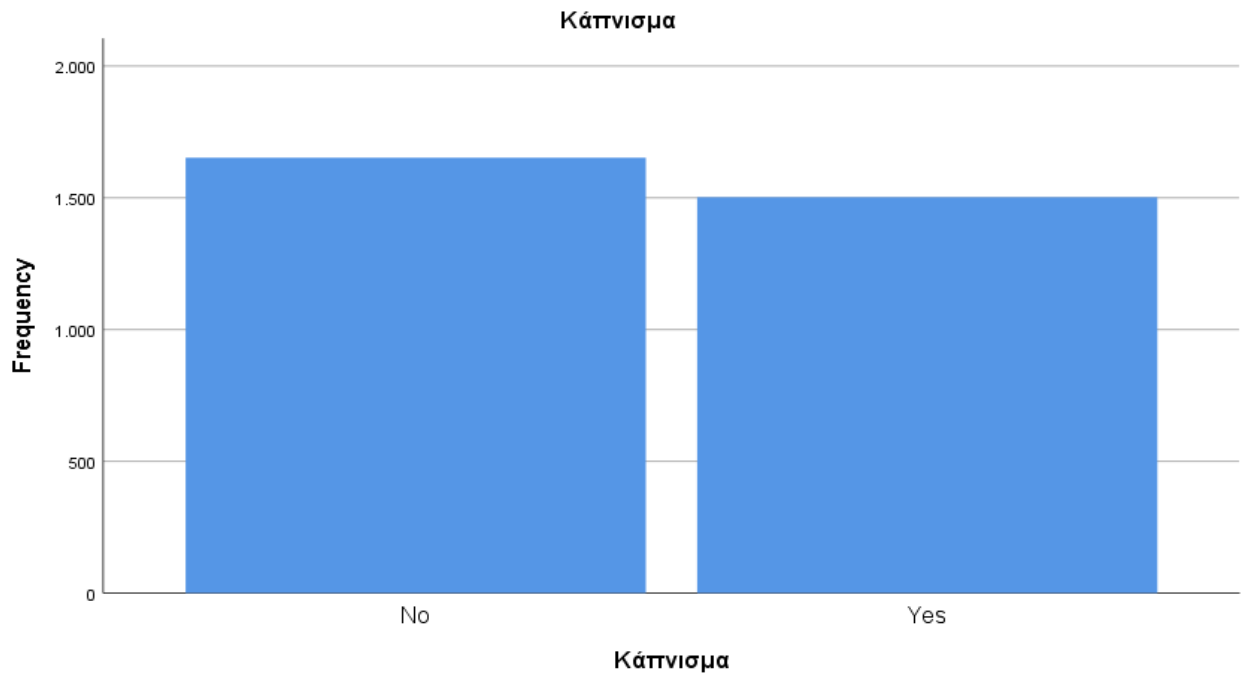
Το κάπνισμα όπως μας είναι γνωστό αποτελεί έναν από τους κύριους παράγοντες που οδηγούν σε καρδιαγγειακά νοσήματα και στην εμφάνιση ενός ξαφνικού καρδιακού επεισοδίου

Στο δείγμα μας που αποτελείται από 3.154 άνδρες εκ των οποίων οι 1.652 δεν είναι καπνιστές και αποτελούν το 52,4% του συνόλου ενώ οι 1.502 καπνίζουν, δηλαδή το 47,6%.

Κάπνισμα

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|-------|-----------|---------|---------------|--------------------|
| Valid | No | 1652 | 52,4 | 52,4 | 52,4 |
| | Yes | 1502 | 47,6 | 47,6 | 100,0 |
| | Total | 3154 | 100,0 | 100,0 | |

Πίνακας 2. 16: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “ Κάπνισμα” στους άνδρες



Διάγραμμα 2. 15: Ραβδόγραμμα συχνοτήτων για την μεταβλητή “Κάπνισμα” των ανδρών

- Αριθμός τσιγάρων

Στον παρακάτω πίνακα παρατηρούμε για τους άντρες καπνιστές ότι ο μέσος όρος τσιγάρων που καπνίζουν ημερησίως είναι $24,36 \pm 11,48$ τσιγάρα. Επιπλέον ο ελάχιστος αριθμός των τσιγάρων στους καπνιστές είναι 1 ημερησίως ενώ ο μέγιστος είναι 99 ημερησίως.

Όπως φαίνεται στο διάγραμμα αλλά και στον πίνακα για τον αριθμό των τσιγάρων ημερησίως υπάρχει θετική ασυμμετρία ($skewness = 0,671$), δηλαδή στο δείγμα μας είναι περισσότεροι οι άνδρες καπνιστές οι οποίοι καπνίζουν λιγότερα τσιγάρα ημερησίως σε σχέση με τον μέσο όρο που είδαμε παραπάνω.

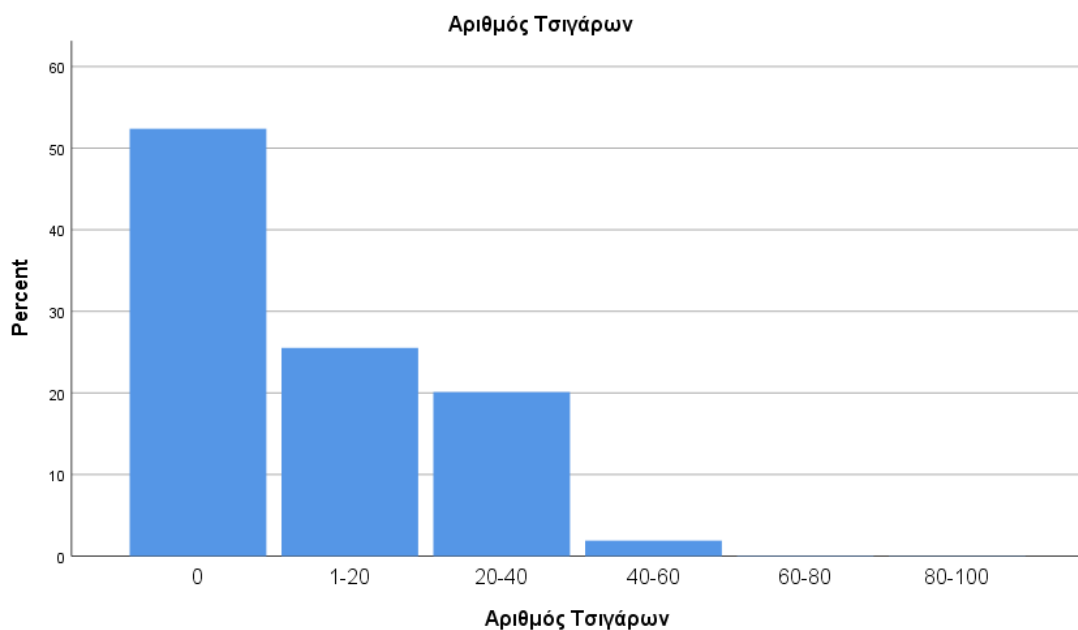
Για την καλύτερη μελέτη του δείγματος θα κατηγοριοποιήσουμε το σύνολο δεδομένων μας ως εξής:

- Άνδρες που δεν καπνίζουν καθόλου
- Άνδρες που καπνίζουν 1-20 τσιγάρα ημερησίως
- Άνδρες που καπνίζουν 20-40 τσιγάρα ημερησίως
- Άνδρες που καπνίζουν 40-60 τσιγάρα ημερησίως
- Άνδρες που καπνίζουν 60-80 τσιγάρα ημερησίως
- Άνδρες που καπνίζουν 80-100 τσιγάρα ημερησίως

Statistics

| Αριθμός τσιγάρων | | |
|------------------------|---------|---------|
| N | Valid | 1502 |
| | Missing | 1652 |
| Mean | | 24,36 |
| Median | | 20,00 |
| Std. Deviation | | 11,477 |
| Variance | | 131,712 |
| Skewness | | ,671 |
| Std. Error of Skewness | | ,063 |
| Kurtosis | | 1,585 |
| Std. Error of Kurtosis | | ,126 |
| Minimum | | 1 |
| Maximum | | 99 |

Πίνακας 2. 17: Περιγραφικά μέτρα για την μεταβλητή “αριθμός τσιγάρων” στους άντρες



Διάγραμμα 2. 16: Ραβδόγραμμα συχνοτήτων για την μεταβλητή “Αριθμός τσιγάρων” των ανδρών

2.5 Κλινικά στοιχεία

2.5.1 Αποτελέσματα για την 1^η βάση δεδομένων που αφορά άνδρες και γυναίκες

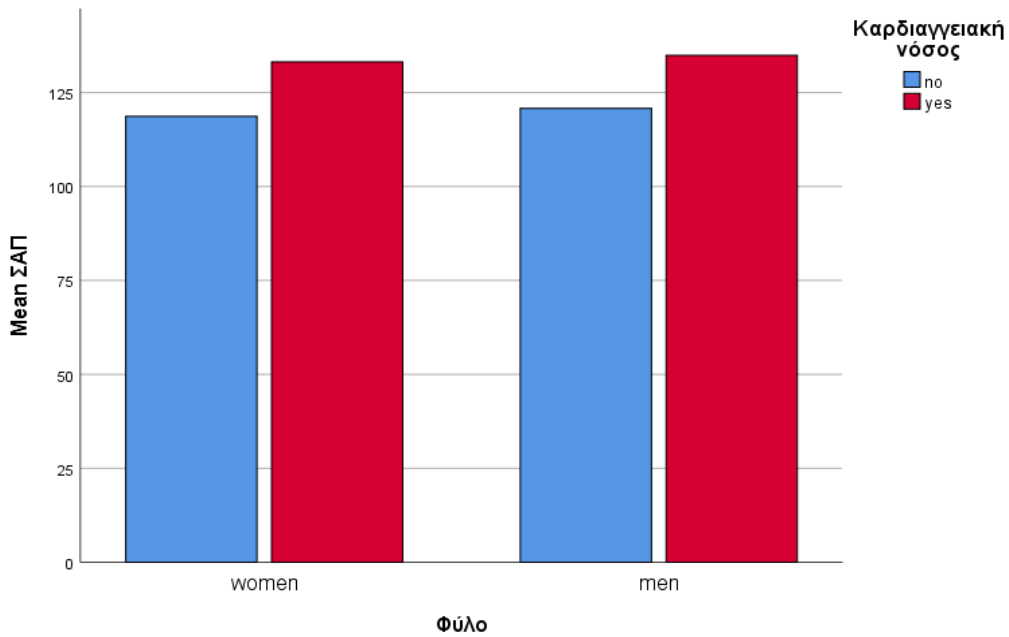
- Συστολική Αρτηριακή Πίεση

Στον πίνακα 2.18 παρατηρούμε ότι η μέση Συστολική Αρτηριακή Πίεση (ΣΑΠ) του δείγματός μας είναι $126,43 \pm 16,714$ με μέγιστη τιμή 240 και ελάχιστη 60. Επίσης, στο ιστόγραμμα παρατηρούμε ότι στο δείγμα μας η κατανομή είναι θετικά ασύμμετρη. ($skewness=0,913$), δηλαδή τα περισσότερα άτομα του δείγματος εμφάνισαν τιμές ΣΑΠ πάνω από το μέσο όρο.

Όπως φαίνεται και στο ραβδόγραμμα παρακάτω τόσο οι άνδρες όσο και οι γυναίκες που εμφάνισαν καρδιαγγειακή νόσο παρουσίαζαν υψηλότερες τιμές ΣΑΠ. Τέλος, οι άνδρες παρουσίασαν υψηλότερες τιμές ΣΑΠ από τις γυναίκες είτε εμφάνισαν καρδιαγγειακή νόσο είτε όχι.

| Statistics | | |
|------------------------|---------|--------|
| ΣΑΠ | | |
| N | Valid | 65662 |
| | Missing | 0 |
| Mean | | 126,43 |
| Median | | 120,00 |
| Std. Deviation | | 16,714 |
| Skewness | | ,913 |
| Std. Error of Skewness | | ,010 |
| Kurtosis | | 1,809 |
| Std. Error of Kurtosis | | ,019 |
| Minimum | | 60 |
| Maximum | | 240 |

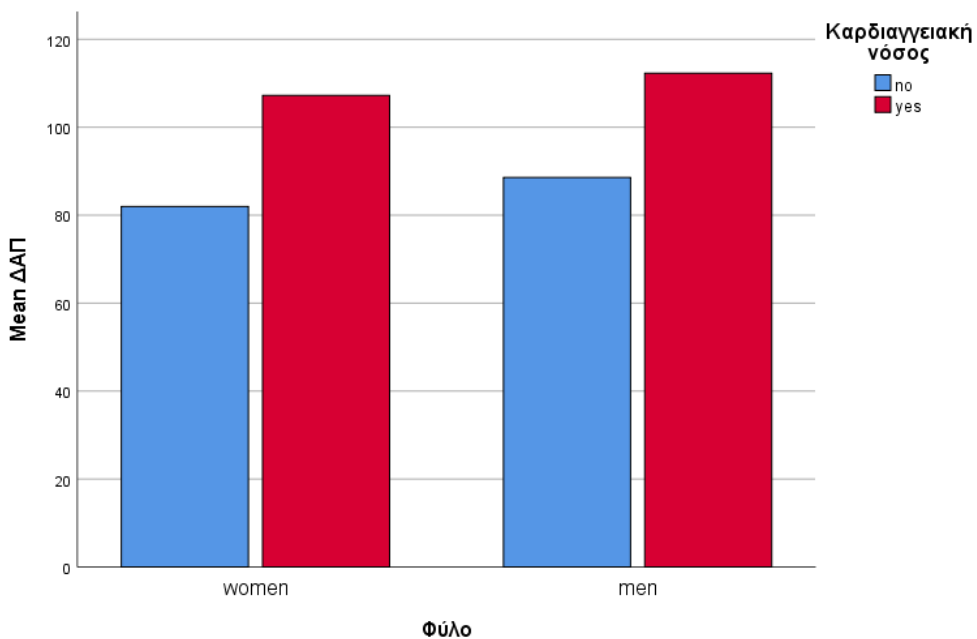
Πίνακας 2. 18: Περιγραφικά μέτρα για την μεταβλητή “Συστολική Αρτηριακή Πίεση”



Διάγραμμα 2. 17: Ραβδόγραμμα μέσης τιμής της μεταβλητής “Συστολική Αρτηριακή Πίεση”

- Διαστολική Αρτηριακή Πίεση

Από τα παρακάτω παρατηρούμε ότι η μέση Διαστολική Αρτηριακή Πίεση (ΔΑΠ) του δείματός μας είναι $81,34 \pm 9,88$ με μέγιστη τιμή 190 και ελάχιστη 10. Επίσης, στο ιστόγραμμα παρατηρούμε ότι στο δείγμα μας η κατανομή είναι θετικά ασύμμετρη. ($skewness=0,588$), δηλαδή τα περισσότερα άτομα του δείγματος εμφάνισαν τιμές ΔΑΠ πάνω από το μέσο όρο.



Διάγραμμα 2. 18: Ραβδόγραμμα μέσης τιμής της μεταβλητής “Διαστολική Αρτηριακή Πίεση”

| Statistics | | |
|------------------------|---------|-------|
| ΔΑΠ | | |
| N | Valid | 65662 |
| | Missing | 0 |
| Mean | | 81,34 |
| Median | | 80,00 |
| Std. Deviation | | 9,880 |
| Skewness | | ,588 |
| Std. Error of Skewness | | ,010 |
| Kurtosis | | 4,924 |
| Std. Error of Kurtosis | | ,019 |
| Minimum | | 10 |
| Maximum | | 190 |

Πίνακας 2. 19: Περιγραφικά μέτρα για την μεταβλητή “Διαστολική Αρτηριακή Πίεση”

Όπως φαίνεται και στο ραβδόγραμμα 2.18 τόσο οι άνδρες όσο και οι γυναίκες που εμφάνισαν καρδιαγγειακή νόσο παρουσίαζαν υψηλότερες τιμές ΔΑΠ. Τέλος, οι άνδρες παρουσίασαν υψηλότερες τιμές ΔΑΠ από τις γυναίκες είτε εμφάνισαν καρδιαγγειακή νόσο είτε όχι.

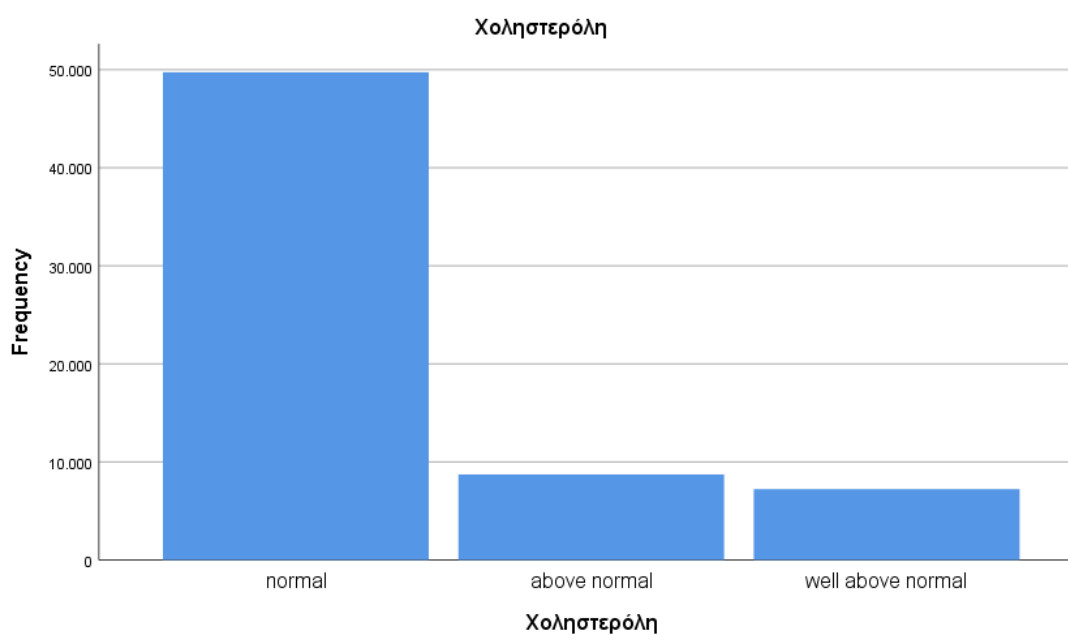
- Χοληστερόλη

Στον παρακάτω πίνακα παρατηρούμε ότι από το σύνολο των ατόμων το 75,7% δηλαδή 49.729 άτομα εμφάνισαν κανονική τιμή χοληστερόλης, το 13,3%, δηλαδή οι 8.717 άτομα εμφάνισαν τιμή χοληστερόλης πάνω από το κανονικό ενώ το 11,0%, δηλαδή 7.216 άτομα εμφάνισαν τιμή χοληστερόλης πολύ πάνω από το κανονικό.

Χοληστερόλη

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|-------------------|-----------|---------|---------------|--------------------|
| Valid | normal | 49729 | 75,7 | 75,7 | 75,7 |
| | above normal | 8717 | 13,3 | 13,3 | 89,0 |
| | well above normal | 7216 | 11,0 | 11,0 | 100,0 |
| | Total | 65662 | 100,0 | 100,0 | |

Πίνακας 2. 20: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “Χοληστερόλη”



Διάγραμμα 2. 19: Ραβδόγραμμα για την μεταβλητή “Χοληστερόλη”

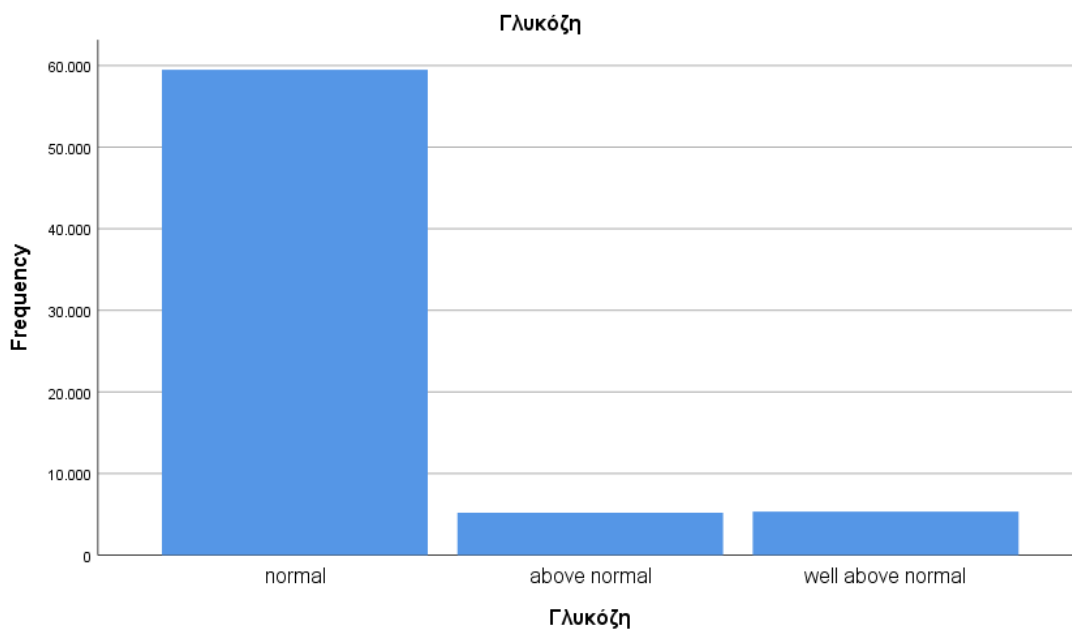
- Γλυκόζη

Στον παρακάτω πίνακα παρατηρούμε ότι από το σύνολο των ατόμων το 85,7% δηλαδή 56.259 άτομα εμφάνισαν κανονική τιμή γλυκόζης, το 7,0%, δηλαδή οι 4.580 άτομα εμφάνισαν τιμή γλυκόζης πάνω από το κανονικό ενώ το 7,3%, δηλαδή 4.823 άτομα εμφάνισαν τιμή χοληστερόλης πολύ πάνω από το κανονικό.

Γλυκόζη

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|-------------------|-----------|---------|---------------|--------------------|
| Valid | normal | 56259 | 85,7 | 85,7 | 85,7 |
| | above normal | 4580 | 7,0 | 7,0 | 92,7 |
| | well above normal | 4823 | 7,3 | 7,3 | 100,0 |
| | Total | 65662 | 100,0 | 100,0 | |

Πίνακας 2. 21: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “Γλυκόζη”



Διάγραμμα 2. 20: Ραβδόγραμμα για την μεταβλητή “Γλυκόζη”

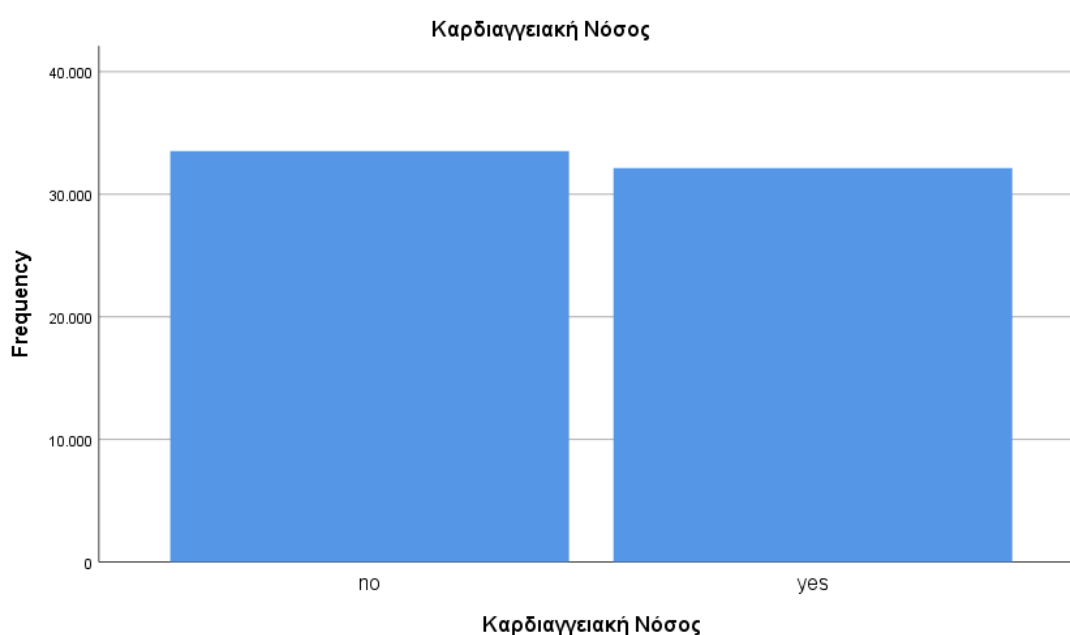
- Υπαρξη καρδιαγγειακού νοσήματος

Παρατηρούμε ότι το 51,1%, δηλαδή 33.524 άτομα από το σύνολο του δείγματος δεν έχουν εμφανίσει κάποια καρδιαγγειακή νόσο και αντίστοιχα 48,9% του συνόλου, δηλαδή 32.138 άτομα έχουν προσβληθεί από καρδιαγγειακή νόσο.

Καρδιαγγειακή Νόσος

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|----------|-----------|---------|---------------|--------------------|
| Valid no | 33524 | 51,1 | 51,1 | 51,1 |
| yes | 32138 | 48,9 | 48,9 | 100,0 |
| Total | 65662 | 100,0 | 100,0 | |

Πίνακας 2. 22: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “Καρδιαγγειακή νόσος”



Διάγραμμα 2. 21: Ραβδόγραμμα για την μεταβλητή “Καρδιαγγειακή νόσος”

2.5.2 Αποτελέσματα για τη 2^η βάση δεδομένων που αφορά άνδρες

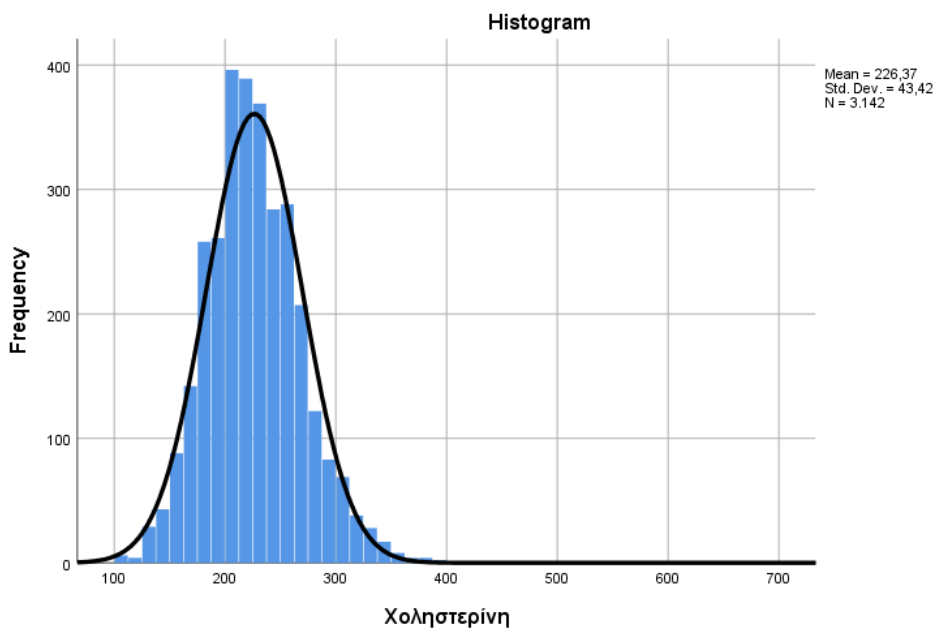
- Ολική χοληστερόλη

Παρατηρούμε παρακάτω ότι η μέση ολική χοληστερόλη στους άνδρες είναι $226,37 \pm 43,12$ mg/dL καθώς επίσης ότι η μέγιστη τιμή ολικής χοληστερόλης στο δείγμα μας είναι 645 mg/dL ενώ η ελάχιστη 103 mg/dL. Από το ιστόγραμμα μπορούμε να δούμε πως η “ουρά” προς τα δεξιά μας δίνει θετικό πρόσημο ($skewness = 0,677$) συνεπώς η κατανομή είναι θετικά ασύμμετρη, δηλαδή οι περισσότεροι άνδρες του δείγματος έχουν τιμές χοληστερίνης πάνω από τη μέση τιμή.

Statistics

| Χοληστερίνη | | |
|------------------------|---------|----------|
| N | Valid | 3142 |
| | Missing | 12 |
| Mean | | 226,37 |
| Median | | 223,00 |
| Std. Deviation | | 43,420 |
| Variance | | 1885,333 |
| Skewness | | ,677 |
| Std. Error of Skewness | | ,044 |
| Kurtosis | | 3,039 |
| Std. Error of Kurtosis | | ,087 |
| Minimum | | 103 |
| Maximum | | 645 |

Πίνακας 2. 23: Περιγραφικά μέτρα για την μεταβλητή “χοληστερίνη” των ανδρών



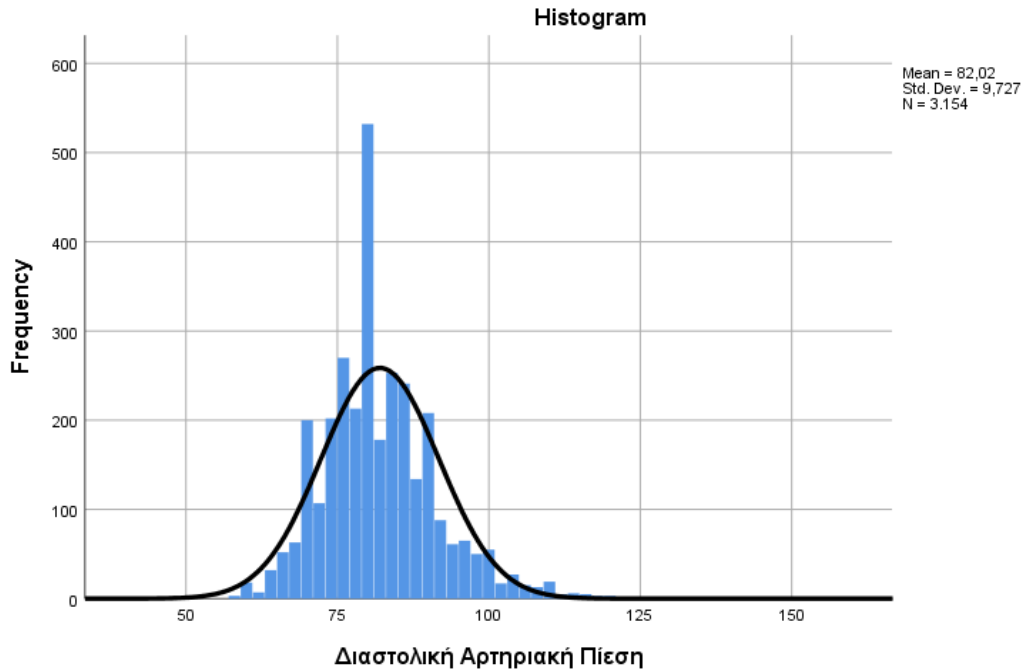
Διάγραμμα 2. 22: Ιστόγραμμα συχνοτήτων της μεταβλητής “χοληστερίνη” στους άνδρες

- Διαστολική Αρτηριακή Πίεση (ΔΑΠ)

Από τα παρακάτω παρατηρούμε ότι η μέση Διαστολική Αρτηριακή Πίεση (ΔΑΠ) του δείγματός μας είναι $82,02 \pm 9,73$ με μέγιστη τιμή 150 και ελάχιστη 58. Επίσης, στο ιστόγραμμα παρατηρούμε ότι στο δείγμα η κατανομή είναι θετικά ασύμμετρη. ($skewness=0,928$), δηλαδή οι περισσότεροι άνδρες εμφάνισαν τιμές ΔΑΠ κάτω από το μέσο όρο.

| Statistics | | |
|----------------------------|---------|--------|
| Διαστολική Αρτηριακή Πίεση | | |
| N | Valid | 3154 |
| | Missing | 0 |
| Mean | | 82,02 |
| Median | | 80,00 |
| Std. Deviation | | 9,727 |
| Variance | | 94,612 |
| Skewness | | ,928 |
| Std. Error of Skewness | | ,044 |
| Kurtosis | | 2,369 |
| Std. Error of Kurtosis | | ,087 |
| Minimum | | 58 |
| Maximum | | 150 |

Πίνακας 2. 24: Περιγραφικά μέτρα για την μεταβλητή “Διαστολική Αρτηριακή Πίεση” των ανδρών



Διάγραμμα 2. 23: Ιστόγραμμα συχνοτήτων της μεταβλητής “Διαστολική Αρτηριακή Πίεση” στους

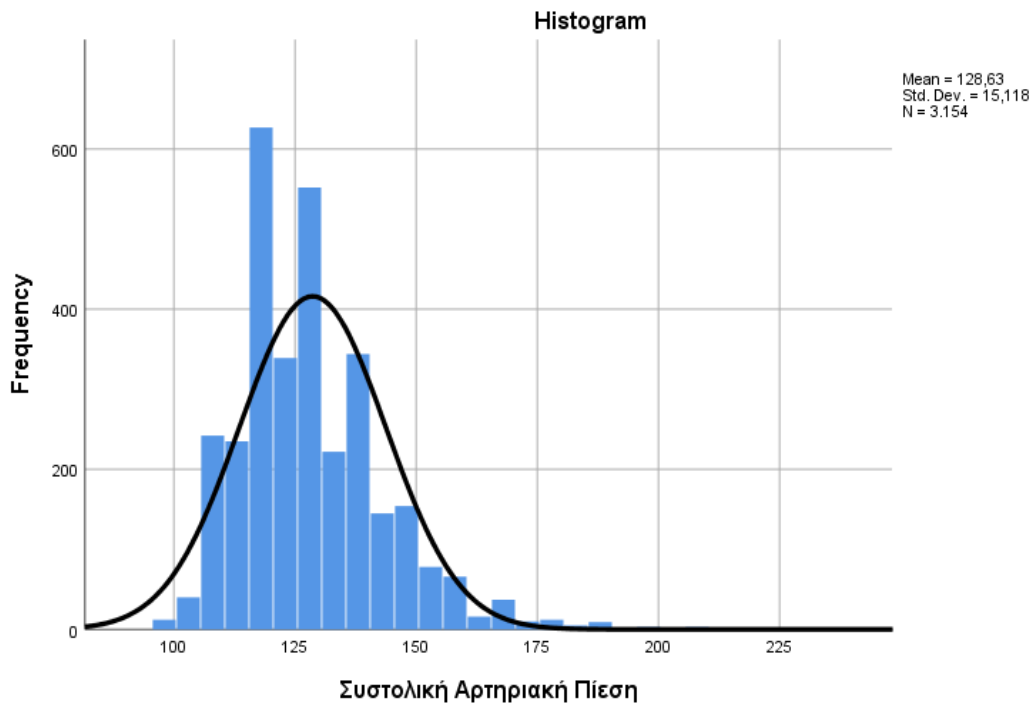
- Συστολική Αρτηριακή Πίεση (ΣΑΠ)

Από τα παρακάτω παρατηρούμε ότι η μέση Συστολική Αρτηριακή Πίεση (ΣΑΠ) του δείγματός μας είναι $128,63 \pm 15,12$ με μέγιστη τιμή 230 και ελάχιστη 98. Επίσης, στο ιστόγραμμα παρατηρούμε ότι στο δείγμα μας η κατανομή είναι θετικά ασύμμετρη. (skewness=1,205), δηλαδή οι περισσότεροι άνδρες εμφάνισαν τιμές ΣΑΠ κάτω από το μέσο όρο.

Statistics

| Συστολική Αρτηριακή Πίεση | | |
|---------------------------|---------|---------|
| N | Valid | 3154 |
| | Missing | 0 |
| Mean | | 128,63 |
| Median | | 126,00 |
| Std. Deviation | | 15,118 |
| Variance | | 228,546 |
| Skewness | | 1,205 |
| Std. Error of Skewness | | ,044 |
| Kurtosis | | 2,799 |
| Std. Error of Kurtosis | | ,087 |
| Minimum | | 98 |
| Maximum | | 230 |

Πίνακας 2. 25: Περιγραφικά μέτρα για την μεταβλητή “Συστολική Αρτηριακή Πίεση” των ανδρών



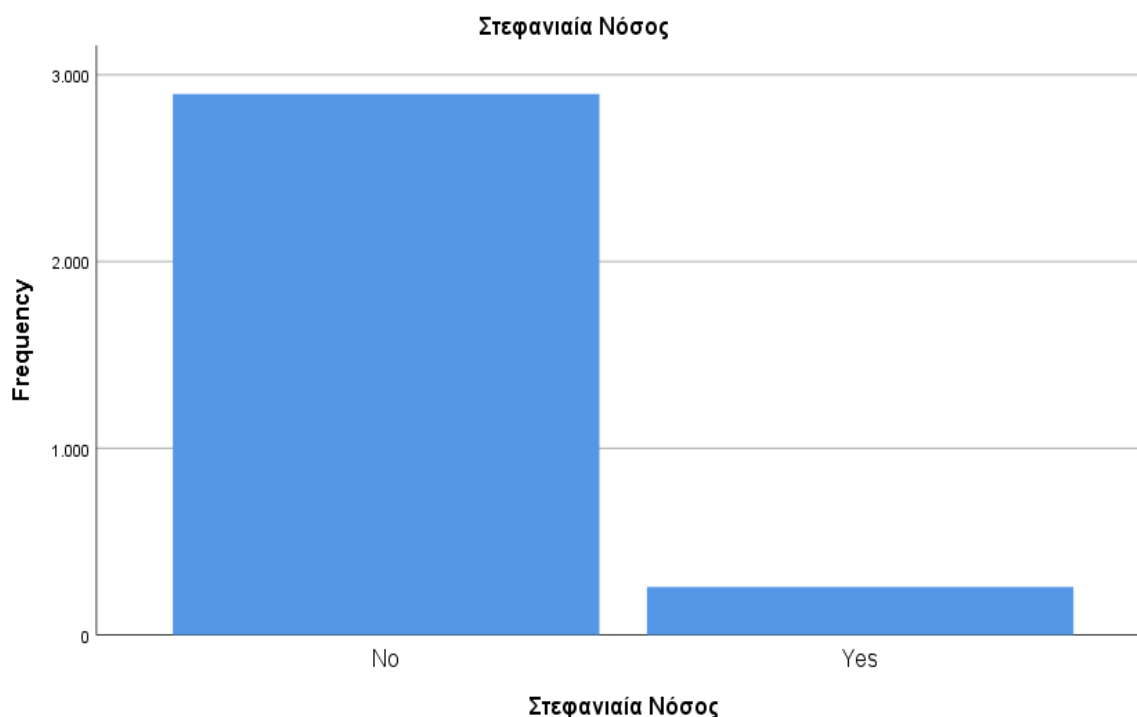
Διάγραμμα 2. 24: Ιστόγραμμα συχνοτήτων της μεταβλητής “Συστολική Αρτηριακή Πίεση” στους

- Υπαρξη στεφανιαίας νόσου

Παρατηρούμε ότι το 91,9%, δηλαδή οι 2.897 από τους 3.154 άνδρες του δείγματος δεν έχουν εμφανίσει κάποια στεφανιαία νόσο ενώ το 8,1%, δηλαδή οι 257 από τους 3.154 έχουν προσβληθεί από κάποια στεφανιαία νόσο.

| Στεφανιαία Νόσος | | | | | |
|------------------|-------|-----------|---------|---------------|--------------------|
| | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | No | 2897 | 91,9 | 91,9 | 91,9 |
| | Yes | 257 | 8,1 | 8,1 | 100,0 |
| | Total | 3154 | 100,0 | 100,0 | |

Πίνακας 2. 26: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “Στεφανιαία νόσος” στους άνδρες



Διάγραμμα 2. 25: Ραβδόγραμμα για την μεταβλητή “Στεφανιαία νόσος” στους άνδρες

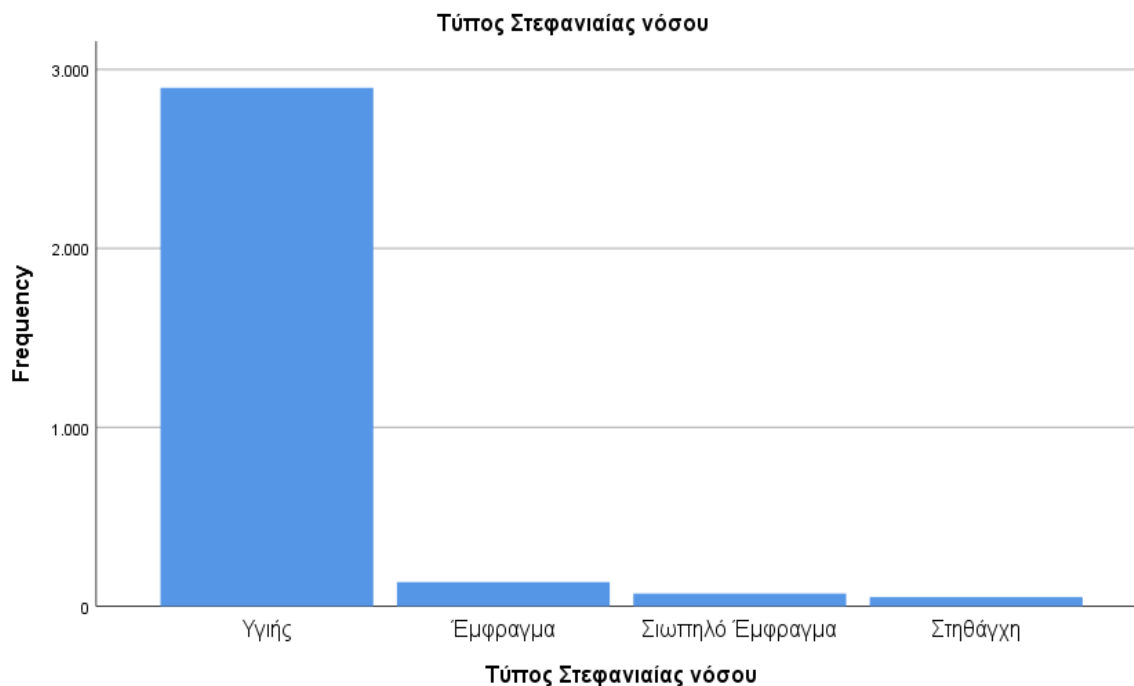
- Τύπος στεφανιαίας νόσου

Στον παρακάτω πίνακα παρατηρούμε ότι το 91,9%, δηλαδή οι 2.897 από τους 3.154 άνδρες του δείγματος δεν έχουν εμφανίσει κάποια στεφανιαία νόσο, το 4,3%, δηλαδή 135 από τους 3.154 έπαθαν έμφραγμα και το 2,3%, δηλαδή 71 άνδρες εμφάνισαν σιωπηλό έμφραγμα.

Τύπος Στεφανιαίας νόσου

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|------------------|-----------|---------|---------------|--------------------|
| Valid | Υγιής | 2897 | 91,9 | 91,9 | 91,9 |
| | Έμφραγμα | 135 | 4,3 | 4,3 | 96,1 |
| | Σιωπηλό Έμφραγμα | 71 | 2,3 | 2,3 | 98,4 |
| | Στηθάγχη | 51 | 1,6 | 1,6 | 100,0 |
| | Total | 3154 | 100,0 | 100,0 | |

Πίνακας 2. 27: Παρατηρούμενες συχνότητες και ποσοστά της μεταβλητής “Τύπος Στεφανιαίας νόσου” στους άνδρες



Διάγραμμα 2. 26: Ραβδόγραμμα για την μεταβλητή “Τύπος Στεφανιαίας νόσου” στους άνδρες

3^ο ΚΕΦΑΛΑΙΟ

Μεθοδολογία

3.1 Έλεγχος ανεξαρτησίας μεταξύ ποιοτικών μεταβλητών.

Η σχέση μεταξύ δύο ποιοτικών μεταβλητών ονομάζεται συνάφεια (association ή contiguity). Για να ελέγξουμε αν υπάρχει εξάρτηση ή γενικότερα συσχέτιση μεταξύ δύο ή περισσότερων ποιοτικών μεταβλητών χρησιμοποιούμε τον έλεγχο ανεξαρτησίας χ^2 . Η σχέση εξάρτησης μεταξύ δύο ή περισσότερων μεταβλητών μπορεί να περιγραφεί μέσω των πινάκων συνάφειας (Σαχλάς Α. και Μπερσίμης Σ., 2017).

3.1.1 Πίνακες συνάφειας

Πίνακα συνάφειας ονομάζουμε τον πίνακα συχνοτήτων που προκύπτει αν ταξινομήσουμε ταυτόχρονα άτομα ή περιπτώσεις ή στοιχεία, βάσει των τιμών δύο ή περισσότερων ποιοτικών μεταβλητών.

Ο πιο απλός πίνακας συνάφειας είναι ο 2x2 που αποτελείται από δύο ποιοτικές μεταβλητές με δύο κατηγορίες η κάθε μία και συνολικά έχει τέσσερα κελιά. Οι στήλες του πίνακα αποτελούν τις κατηγορίες της μίας μεταβλητής και οι γραμμές της έταιρης. Με O_{ij} συμβολίζονται οι συχνότητες εμφάνισης των συνδυασμών των i γραμμών με τις j στήλες και ονομάζονται παρατηρημένες τιμές (observed values). Επίσης, εμφανίζονται τα περιθώρια αθροίσματα των στηλών και των γραμμών.

| Μεταβλητή 1 | Μεταβλητή 2 | | |
|-------------|---------------|---------------|---------------|
| | E | A | Σύνολο |
| E | O_{11} | O_{12} | Σ_{1j} |
| A | O_{21} | O_{22} | Σ_{2j} |
| Σύνολο | Σ_{i1} | Σ_{i2} | n |

Πίνακας 2. 28: Παράδειγμα 2x2 πίνακα συνάφειας

Οι αναμενόμενες τιμές (expected values) συμβολίζονται με E_{ij} και υπολογίζονται πολλαπλασιάζοντας το σύνολο μιας στήλης με το σύνολο μιας γραμμής και διαιρώντας το με το γενικό σύνολο του δείγματος.

Οι σχετικές συχνότητες προκύπτουν όταν διαιρούνται τις παρατηρημένες συχνότητες με το σύνολο του δείγματος.

Εάν πάλι διαιρέσουμε τα περιθώρια αθροίσματα με το μέγεθος του δείγματος θα προκύψουν οι συνολικές σχετικές συχνότητες.

Τέλος, οι σχετικές συχνότητες ανά γραμμή ή στήλη υπολογίζονται όταν διαιρεθεί η παρατηρημένη συχνότητα ενός κελιού με το σύνολο της αντίστοιχης γραμμής ή της στήλης. Πρόκειται ουσιαστικά και δεσμευμένες πιθανότητες, δεσμεύοντας ως προς κάποιο συγκεκριμένο χαρακτηριστικό.

3.1.2 Έλεγχος ανεξαρτησίας χ^2

Αρχικά, για να εφαρμόσουμε τον έλεγχο ανεξαρτησίας χ^2 (chi-square test) θα πρέπει να έχουμε τα δεδομένα σε μορφή πίνακα συνάφειας. Σε ιατρικές μελέτες όπως αυτή της παρούσας εργασίας, ενδιαφερόμαστε να ερευνήσουμε εάν υπάρχει σχέση ανάμεσα σε δύο κατηγορικές μεταβλητές (Τζαβελάς Γ., 2018). Μέσω του στατιστικού ελέγχου ανεξαρτησίας χ^2 θα είμαστε σε θέση να απαντήσουμε εάν δύο ποιοτικές μεταβλητές είναι ανεξάρτητες ή όχι. Εάν ορίσουμε ως περιθώριες πιθανότητες τις $p_{ij} = O_{ij}/n$, $p_{i1} = O_{i1}/n$ και $p_{1j} = O_{1j}/n$,

τότε μπορούμε να ορίσουμε και μαθηματικά τις υποθέσεις του ελέγχου χ^2 ως εξής:

$$H_0: p_{ij} = p_{i.} \cdot p_{.j}$$

δηλαδή η πιθανότητα του κελιού να είναι ίση με το γινόμενο των δύο περιθωρίων πιθανοτήτων

$$H_1: p_{i|j} \neq p_{i.}$$

δηλαδή ότι δοθέντος ενός επιπέδου μιας μεταβλητής η πιθανότητα των επιπέδων της άλλης είναι σταθερή

Προκειμένου να εφαρμόζεται σωστά ο έλεγχος χ^2 θα πρέπει να ικανοποιούνται οι παρακάτω συνθήκες:

- Το μέγεθος του δείγματος δεν πρέπει να είναι μικρότερο του τετραπλάσιου του αριθμού των κελιών του πίνακα συνάφειας
- Καμία από τις αναμενόμενες συχνότητες δεν πρέπει να είναι μικρότερη του 1.
- Το ποσοστό των αναμενόμενων συχνοτήτων οι οποίες είναι μικρότερες του 5 δεν πρέπει να είναι μεγαλύτερο από το 20% με 25%.

Στην περίπτωση που το μέγεθος του δείγματος είναι μικρός και δεν μπορούν να καλυφθούν οι παραπάνω υποθέσεις αντί του ελέγχου ανεξαρτησίας χ^2 χρησιμοποιείται ο ακριβής έλεγχος του Fischer (Fischer's exact test)

Τέλος, στην περίπτωση που η μία μεταβλητή είναι κατηγορική και η άλλη διατάξιμη ο έλεγχος χ^2 δεν μπορεί να χρησιμοποιηθεί γιατί σε αυτή την περίπτωση δεν εξετάζουμε

αν υπάρχει εξάρτηση μεταξύ των μεταβλητών αλλά αν υπάρχει τάση στα ποσοστά ως προς τις κατηγορίες της διατάξιμης μεταβλητής. Για τον παραπάνω λόγο σε αυτή την περίπτωση εφαρμόζουμε έλεγχο χ^2 για τάση (linear trend test). Η μηδενική υπόθεση του ελέγχου αυτού είναι ότι δεν υπάρχει γραμμική τάση στα ποσοστά καθώς αυξάνονται οι κατηγορίες της διατάξιμης μεταβλητής η οποία μαθηματικά γράφεται ως εξής:

$$H_0: p_{1|1} = p_{1|2} = \dots = p_{1|k}$$

έναντι της εναλλακτικής υπόθεσης

$$H_1: p_{1|1} \leq p_{1|2} \leq \dots \leq p_{1|k}$$

όπου $p_{1|i} = a + bx_i$, x_i είναι τα σκόρ των κατηγοριών της διατάξιμης μεταβλητής.

3.2 Έλεγχος συσχέτισης ποσοτικής μεταβλητής με μια ποιοτική

Στην περίπτωση που η ανεξάρτητη μεταβλητή είναι ποσοτική τότε δεν μπορεί να χρησιμοποιηθούν οι παραπάνω έλεγχοι. Δεδομένου ότι η εξαρτημένη μεταβλητή είναι δίτιμη και ότι οι τιμές της ανεξάρτητης μεταβλητής δεν κατανέμονται κανονικά θα γίνει χρήση του ελέγχου U των Mann-Whitney.

Προκειμένου να ελεγχθεί η κανονικότητα των τιμών της ανεξάρτητης μεταβλητής ως προς την εξαρτημένη μεταβλητή θα γίνει χρήση του μη παραμετρικού ελέγχου Kolmogorov-Smirnov.

Ο έλεγχος U των Mann – Whitney χρησιμοποιείται για να εξετασθεί η μηδενική υπόθεση ότι δύο ανεξάρτητα δείγματα προέρχονται από τον ίδιο πληθυσμό (ή έχουν την ίδια κατανομή ή έχουν την ίδια διάμεσο).

Έστω ότι από ένα πληθυσμό έχει ληφθεί δείγμα x_1, x_2, \dots, x_n μεγέθους n και από έναν δεύτερο πληθυσμό δείγμα y_1, y_2, \dots, y_m μεγέθους m . Ο έλεγχος U των Mann – Whitney βασίζεται στη σύγκριση κάθε παρατήρησης x_i από το πρώτο δείγμα με κάθε παρατήρηση y_i από το δεύτερο δείγμα, έτσι ο αριθμός του συνόλου των συγκρίσεων που πραγματοποιούνται είναι nm .

Αν τα δύο δείγματα έχουν την ίδια διάμεσο, τότε κάθε x_i έχει την ίδια πιθανότητα να είναι μεγαλύτερο ή μικρότερο από κάθε y_i . Πιο συγκεκριμένα,

$$H_0: P(x_i > y_i) = 0.5$$

$$H_1: P(x_i > y_i) \neq 0.5$$

Προκειμένου να υπολογιστεί η τιμή της συνάρτησης U των Mann – Whitney, μετρίεται ο αριθμός των φορών που μια παρατήρηση x_i από το πρώτο δείγμα είναι μεγαλύτερη από μια παρατήρηση y_i από το δεύτερο δείγμα, ο οποίος στη συνέχεια συμβολίζεται με U_x . Με όμοιο τρόπο, γίνεται η μέτρηση του αριθμού των φορών που μια παρατήρηση x_i από το πρώτο δείγμα είναι μικρότερη από μια παρατήρηση y_i από το δεύτερο δείγμα. Ο αριθμός αυτός συμβολίζεται με U_y . Στη μηδενική υπόθεση αναμένεται τα U_x και U_y να είναι περίπου ίσα.

Θα τεθεί $U = \min\{U_x, U_y\}$ ώστε να ολοκληρωθεί ο έλεγχος των Mann – Whitney. Για να παρθεί η απόφαση εάν πρέπει να απορριφθεί ή όχι η μηδενική υπόθεση, θα βρεθεί η κρίσιμη τιμή U_{cr} του ελέγχου. Εάν ισχύει ότι $U < U_{cr}$, τότε η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας 5%. Στην περίπτωση όπου ο αριθμός των παρατηρήσεων στο δείγμα μπορεί να δώσει αριθμό συνολικών συγκρίσεων (nm) μεγαλύτερο του 20 μπορεί να ακολουθηθεί η κανονική προσέγγιση. Για τους υπολογισμούς χρησιμοποιούνται οι

$$\mu_U = \frac{nm}{2}$$

$$\sigma_U = \sqrt{\frac{nm(n+m+1)}{12}}$$

Εάν βρεθούν τουλάχιστον δύο ίσες παρατηρήσεις στα δείγματα θα προστεθεί μισή μονάδα στα U_x και U_y για κάθε ζευγάρι παρατηρήσεων που θα βρεθεί. Στην περίπτωση όπου θα ακολουθηθεί η κανονική προσέγγιση είναι απαραίτητο να γίνει κατάλληλος επαναπροσδιορισμός της τυπικής απόκλισης ώστε να ολοκληρωθεί ο έλεγχος.

3.3 Λογιστική παλινδρόμηση

Η μέθοδος της λογιστικής παλινδρόμησης χρησιμοποιείται για να περιγράψει τη σχέση της πιθανότητας ενός χαρακτηριστικού (μεταβλητή απόκρισης), με έναν ή περισσότερους παράγοντες (ερμηνευτικές μεταβλητές). Η εξαρτημένη μεταβλητή είναι δίτιμη και φανερώνει την παρουσία ή απουσία ενός χαρακτηριστικού ή ασθένειας ενώ οι ανεξάρτητες μεταβλητές μπορεί να είναι συνεχείς ή ποιοτικές. Όταν είναι συνεχείς ονομάζονται συμμεταβλητές, ενώ όταν είναι ποιοτικές, παράγοντες.

Πολλαπλό λογιστικό μοντέλο

Η μέθοδος που θα χρησιμοποιήσουμε επιτρέπει την πρόβλεψη των τιμών της δίτιμης εξαρτημένης μεταβλητής από ένα πλήθος ανεξάρτητων μεταβλητών οι οποίες μπορεί να είναι ποσοτικές ή και ποιοτικές. Η μεταβλητή πρόβλεψης λαμβάνει την τιμή 1 με πιθανότητα επιτυχίας p και την τιμή 0 με πιθανότητα αποτυχίας $1 - p$.

Το μοντέλο πολλαπλής λογιστικής παλινδρόμησης είναι:

$$p = \frac{e^z}{1+e^z}$$

όπου $z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$

όπου X_1, X_2, \dots, X_k οι ερμηνευτικές μεταβλητές του μοντέλου

Η ισότητα (1) ισοδυναμεί με:

$$\frac{p}{1-p} = e^z$$

και

$$\text{Log}\left(\frac{p}{1-p}\right) = z$$

Το $\frac{p}{1-p}$ είναι ο λόγος των συμπληρωματικών πιθανοτήτων (odds ratio). Η εκτίμηση των παραμέτρων του μοντέλου της λογιστικής παλινδρόμησης γίνεται με τη μέθοδο της μέγιστης πιθανοφάνειας.

Εάν ο λόγος συμπληρωματικών πιθανοτήτων παίρνει τιμές κοντά στη μονάδα ή ισοδύναμα όταν το $\log(\text{odds ratio})$ παίρνει τιμές κοντά στο μηδέν, τότε η ερμηνευτική μεταβλητή χ δεν έχει ιδιαίτερη προβλεπτική ικανότητα.

Η ερμηνεία των συντελεστών του μοντέλου είναι ίδια με αυτών της γραμμικής παλινδρόμησης. Ο σταθερός όρος b_0 είναι πρακτικά η τιμή του λογαρίθμου του λόγου των συμπληρωματικών πιθανοτήτων όταν η ανεξάρτητη μεταβλητή b_1 είναι μηδέν. Ο συντελεστής b_1 αποτελεί τη κλίση της ευθείας και αντιπροσωπεύει πρακτικά τη μεταβολή που θα επέλθει στο λογάριθμο του λόγου συμπληρωματικών πιθανοτήτων όταν η ερμηνευτική μεταβλητή αυξηθεί κατά μία μονάδα.

4^ο ΚΕΦΑΛΑΙΟ

ΔΙΕΡΕΥΝΗΣΗ ΣΥΣΧΕΤΙΣΕΩΝ ΤΩΝ ΜΕΤΑΒΛΗΤΩΝ

4.1 Εισαγωγή

Στο κεφάλαιο αυτό θα μελετήσουμε τις συσχετίσεις μεταξύ των μεταβλητών στην κάθε βάση δεδομένων για να εξάγουμε συμπεράσματα για τις μεταβλητές με τις οποίες σχετίζονται. Δηλαδή, να βρούμε τις ποσοτικές και ποιοτικές μεταβλητές οι οποίες επηρεάζουν την εμφάνιση της καρδιαγγειακής νόσου για την 1^η βάση δεδομένων που αφορά άνδρες και γυναίκες και την εμφάνιση στεφανιαίας νόσου για την 2^η βάση δεδομένων που αφορά τους άνδρες.

4.2 Συσχετίσεις όλων μεταβλητών με την ύπαρξη Καρδιαγγειακής νόσου για την 1^η βάση που αφορά άνδρες και γυναίκες

4.2.1 Έλεγχος συσχέτισης ποιοτικών μεταβλητών με την μεταβλητή “Καρδιαγγειακή νόσος”

- Έλεγχος συσχέτισης του φύλου με την εμφάνιση καρδιαγγειακής νόσου

| Φύλο * Καρδιαγγειακή Νόσος Crosstabulation | | | | | |
|--|-------|----------------|---------------------|---------|---------|
| | | | Καρδιαγγειακή Νόσος | | |
| | | | no | yes | Total |
| Φύλο | women | Count | 22005 | 21010 | 43015 |
| | | Expected Count | 21961,5 | 21053,5 | 43015,0 |
| | men | Count | 11519 | 11128 | 22647 |
| | | Expected Count | 11562,5 | 11084,5 | 22647,0 |
| Total | | Count | 33524 | 32138 | 65662 |
| | | Expected Count | 33524,0 | 32138,0 | 65662,0 |

Πίνακας 4. 1 Πίνακας συνάφειας των μεταβλητών Φύλο - Καρδιαγγειακή νόσος

| Chi-Square Tests | | | | | |
|------------------------------------|-------------------|----|--|--------------------------|--------------------------|
| | Value | df | Asymptotic Significance (2- sided) | Exact Sig. (2- sided) | Exact Sig. (1- sided) |
| Pearson Chi-Square | ,511 ^a | 1 | ,475 | | |
| Continuity Correction ^b | ,499 | 1 | ,480 | | |
| Likelihood Ratio | ,511 | 1 | ,475 | | |
| Fisher's Exact Test | | | | ,475 | ,240 |
| Linear-by-Linear Association | ,511 | 1 | ,475 | | |
| N of Valid Cases | 65662 | | | | |

Πίνακας 4. 2: Πίνακας ελέγχου – χ^2 για τις μεταβλητές Φύλο - Καρδιαγγειακή νόσος

Από τον παραπάνω πίνακα παρατηρούμε ότι αποδεχόμαστε την μηδενική υπόθεση του ελέγχου χ^2 του Pearson καθώς είναι στατιστικά σημαντικό ($p\text{-value} = 0,47 > 0,05$) σε επίπεδο σημαντικότητας 5%. Επομένως, οι μεταβλητές “Φύλο” και “Καρδιαγγειακή Νόσος” δεν συσχετίζονται μεταξύ τους.

- **Έλεγχος συσχέτισης του επιπέδων χοληστερόλης με την εμφάνιση καρδιαγγειακής νόσου**

Στον παρακάτω πίνακα παρατηρούμε ότι απορρίπτουμε την μηδενική υπόθεση του ελέγχου χ^2 για γραμμική τάση καθώς το $p\text{-value}$ είναι μικρότερο από το 5% ($p\text{-value} \approx 0 < 0,05$). Επομένως, τα δεδομένα μας παρουσιάζουν ισχυρή ένδειξη γραμμικής τάσης των ατόμων που εμφάνισαν καρδιαγγειακή νόσο και το επίπεδο που είναι η τιμή της χοληστερόλης του.

Χοληστερόλη * Καρδιαγγειακή Νόσος Crosstabulation

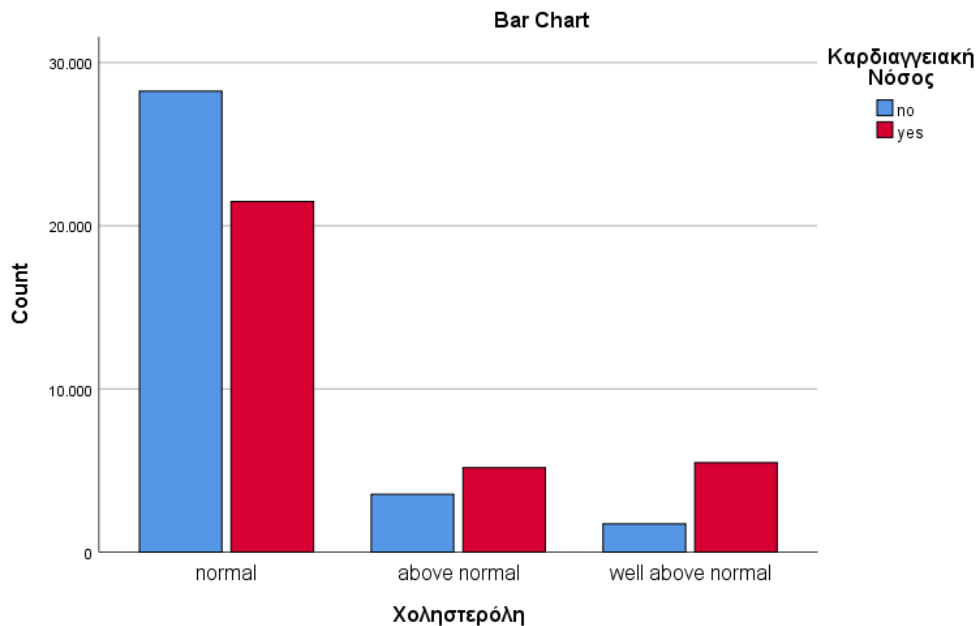
| | | Καρδιαγγειακή Νόσος | | | |
|-------------|-------------------|---------------------|---------|---------|---------|
| | | no | yes | Total | |
| Χοληστερόλη | normal | Count | 28248 | 21481 | 49729 |
| | | Expected Count | 25389,3 | 24339,7 | 49729,0 |
| | above normal | Count | 3544 | 5173 | 8717 |
| | | Expected Count | 4450,5 | 4266,5 | 8717,0 |
| | well above normal | Count | 1732 | 5484 | 7216 |
| | | Expected Count | 3684,2 | 3531,8 | 7216,0 |
| Total | Count | 33524 | 32138 | 65662 | |
| | Expected Count | 33524,0 | 32138,0 | 65662,0 | |

Πίνακας 4. 3: Πίνακας συνάφειας των μεταβλητών Χοληστερόλη - Καρδιαγγειακή νόσος

Chi-Square Tests

| | Value | df | Asymptotic Significance (2-sided) |
|------------------------------|-----------------------|----|-----------------------------------|
| Pearson Chi-Square | 3148,279 ^a | 2 | ,000 |
| Likelihood Ratio | 3250,606 | 2 | ,000 |
| Linear-by-Linear Association | 3148,065 | 1 | ,000 |
| N of Valid Cases | 65662 | | |

Πίνακας 4. 4: Πίνακας ελέγχου χ^2 για τις μεταβλητές Χοληστερόλη - Καρδιαγγειακή νόσος



Διάγραμμα 4. 1: Ραβδόγραμμα συσχέτισης των μεταβλητών Καρδιαγγειακή νόσος – Χοληστερόλη

Στο παραπάνω διάγραμμα παρατηρούμε ότι το μεγαλύτερο ποσοστό των ατόμων τα οποία εμφανίζουν φυσιολογικές τιμές χοληστερόλης δεν εμφάνισαν κάποια καρδιαγγειακή νόσο.

- Έλεγχος συσχέτισης του επιπέδων γλυκόζης με την εμφάνιση καρδιαγγειακής νόσου

Στον παρακάτω πίνακα παρατηρούμε ότι απορρίπτουμε την μηδενική υπόθεση του ελέγχου χ^2 για γραμμική τάση καθώς το p-value είναι μικρότερο από το 5% (p-value $\approx 0 < 0,05$). Επομένως, τα δεδομένα μας παρουσιάζουν ισχυρή ένδειξη γραμμικής τάσης των ατόμων που εμφάνισαν καρδιαγγειακή νόσο και το επίπεδο που είναι η τιμή της γλυκόζης τους.

Γλυκόζη * Καρδιαγγειακή Νόσος Crosstabulation

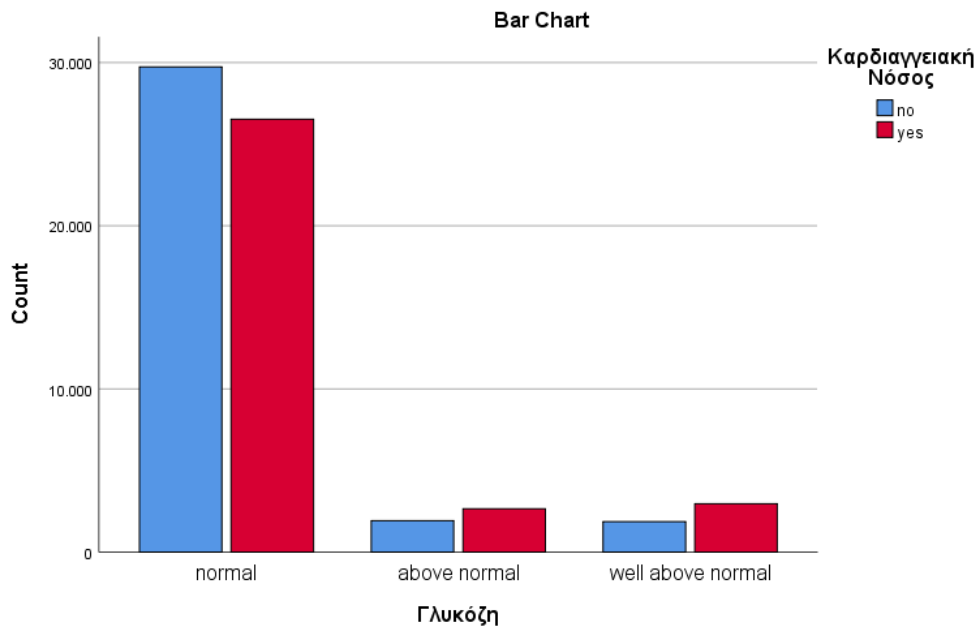
| | | Καρδιαγγειακή Νόσος | | Total | |
|---------|-------------------|---------------------|---------|---------|---------|
| | | no | yes | | |
| Γλυκόζη | normal | Count | 29734 | 26525 | 56259 |
| | | Expected Count | 28723,3 | 27535,7 | 56259,0 |
| | above normal | Count | 1928 | 2652 | 4580 |
| | | Expected Count | 2338,3 | 2241,7 | 4580,0 |
| | well above normal | Count | 1862 | 2961 | 4823 |
| | | Expected Count | 2462,4 | 2360,6 | 4823,0 |
| Total | Count | 33524 | 32138 | 65662 | |
| | Expected Count | 33524,0 | 32138,0 | 65662,0 | |

Πίνακας 4. 5: Πίνακας συνάφειας των μεταβλητών Γλυκόζη - Καρδιαγγειακή νόσος

Chi-Square Tests

| | Value | df | Asymptotic Significance (2-sided) |
|------------------------------|----------------------|----|-----------------------------------|
| Pearson Chi-Square | 518,890 ^a | 2 | ,000 |
| Likelihood Ratio | 521,451 | 2 | ,000 |
| Linear-by-Linear Association | 499,646 | 1 | ,000 |
| N of Valid Cases | 65662 | | |

Πίνακας 4. 6: Πίνακας ελέγχου χ^2 για τις μεταβλητές Γλυκόζη - Καρδιαγγειακή νόσος



Διάγραμμα 4. 2: Ραβδόγραμμα συσχέτισης των μεταβλητών Καρδιαγγειακή νόσος – Γλυκόζη

Στο παραπάνω διάγραμμα παρατηρούμε ότι το μεγαλύτερο ποσοστό των ατόμων τα οποία εμφανίζουν φυσιολογικές τιμές γλυκόζης δεν εμφάνισαν κάποια καρδιαγγειακή νόσο.

- **Έλεγχος συσχέτισης του καπνίσματος με την εμφάνιση καρδιαγγειακής νόσου**

Στον παρακάτω πίνακα παρατηρούμε ότι απορρίπτουμε την μηδενική υπόθεση του ελέγχου χ^2 του Pearson καθώς δεν είναι στατιστικά σημαντικό ($p\text{-value} \approx 0 < 0,05$) σε επίπεδο σημαντικότητας 5%. Επομένως, οι μεταβλητές “Κάπνισμα” και “Καρδιαγγειακή Νόσος” συσχετίζονται μεταξύ τους.

Κάπνισμα * Καρδιαγγειακή Νόσος Crosstabulation

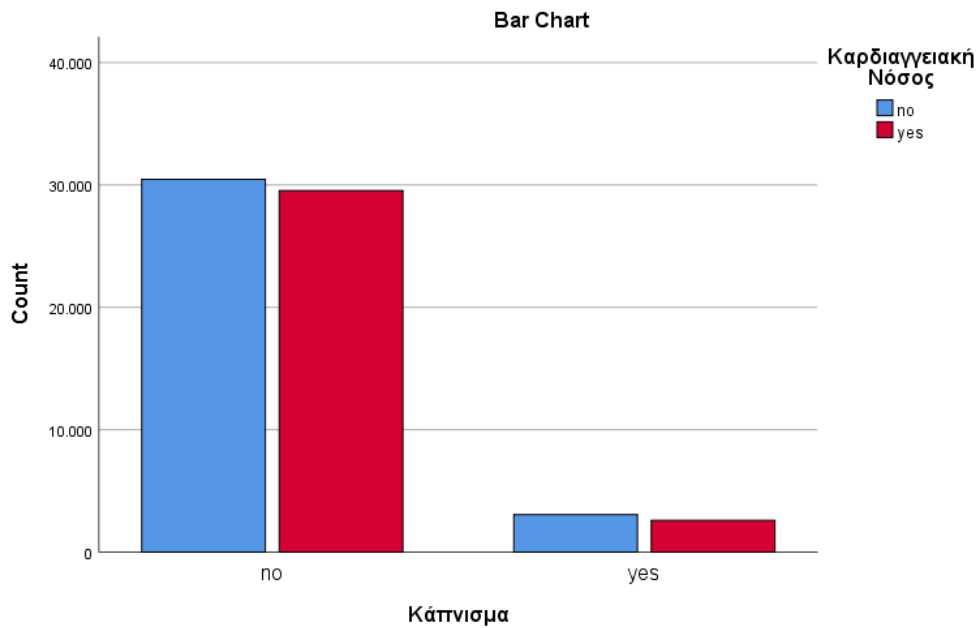
| | | Καρδιαγγειακή Νόσος | | | |
|----------|----------------|---------------------|---------|---------|---------|
| | | no | yes | Total | |
| Κάπνισμα | no | Count | 30456 | 29536 | 59992 |
| | | Expected Count | 30629,2 | 29362,8 | 59992,0 |
| | yes | Count | 3068 | 2602 | 5670 |
| | | Expected Count | 2894,8 | 2775,2 | 5670,0 |
| Total | Count | 33524 | 32138 | 65662 | |
| | Expected Count | 33524,0 | 32138,0 | 65662,0 | |

Πίνακας 4. 7: Πίνακας συνάφειας των μεταβλητών Καπνιστής - Καρδιαγγειακή νόσος

Chi-Square Tests

| | Value | df | Asymptotic Significance (2- sided) | Exact Sig. (2- sided) | Exact Sig. (1- sided) |
|------------------------------------|---------------------|----|--|--------------------------|--------------------------|
| Pearson Chi-Square | 23,162 ^a | 1 | ,000 | | |
| Continuity Correction ^b | 23,029 | 1 | ,000 | | |
| Likelihood Ratio | 23,193 | 1 | ,000 | | |
| Fisher's Exact Test | | | | ,000 | ,000 |
| Linear-by-Linear Association | 23,162 | 1 | ,000 | | |
| N of Valid Cases | 65662 | | | | |

Πίνακας 4. 8: Πίνακας ελέγχου χ^2 για τις μεταβλητές Καπνιστής - Καρδιαγγειακή νόσος



Διάγραμμα 4. 3: Ραβδόγραμμα συσχέτισης των μεταβλητών Καρδιαγγειακή νόσος – Κάπνισμα

- Έλεγχος συσχέτισης της κατανάλωσης αλκοόλ με την εμφάνιση καρδιαγγειακής νόσου

Αλκοόλ * Καρδιαγγειακή Νόσος Crosstabulation

| | | Καρδιαγγειακή Νόσος | | Total | |
|--------|-----|---------------------|---------|---------|---------|
| | | no | yes | | |
| Αλκοόλ | no | Count | 31697 | 30541 | 62238 |
| | | Expected Count | 31775,9 | 30462,1 | 62238,0 |
| | yes | Count | 1827 | 1597 | 3424 |
| | | Expected Count | 1748,1 | 1675,9 | 3424,0 |
| Total | | Count | 33524 | 32138 | 65662 |
| | | Expected Count | 33524,0 | 32138,0 | 65662,0 |

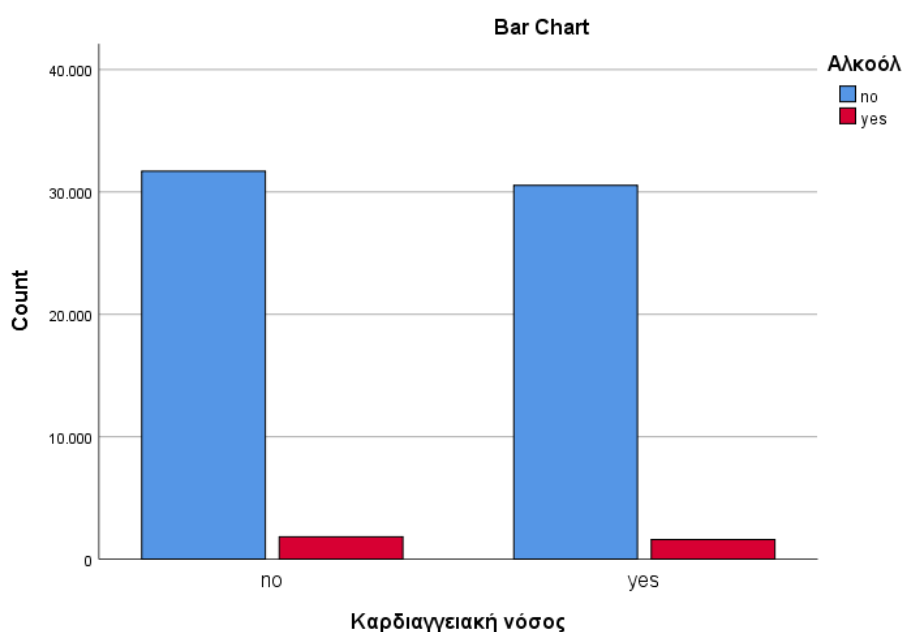
Πίνακας 4. 9: Πίνακας συνάφειας των μεταβλητών Κατανάλωση Αλκοόλ - Καρδιαγγειακή νόσος

Στον παρακάτω πίνακα παρατηρούμε ότι απορρίπτουμε την μηδενική υπόθεση του ελέγχου χ^2 του Pearson ($p\text{-value} = 0,006 < 0,05$) σε επίπεδο σημαντικότητας 5%. Επομένως, οι μεταβλητές “Κατανάλωση Αλκοόλ” και “Καρδιαγγειακή Νόσος” συσχετίζονται μεταξύ τους

Chi-Square Tests

| | Value | df | Asymptotic Significance (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) |
|------------------------------------|--------------------|----|-----------------------------------|----------------------|----------------------|
| Pearson Chi-Square | 7,669 ^a | 1 | ,006 | | |
| Continuity Correction ^b | 7,572 | 1 | ,006 | | |
| Likelihood Ratio | 7,676 | 1 | ,006 | | |
| Fisher's Exact Test | | | | ,006 | ,003 |
| Linear-by-Linear Association | 7,669 | 1 | ,006 | | |
| N of Valid Cases | 65662 | | | | |

Πίνακας 4. 10: Πίνακας ελέγχου – χ^2 για τις μεταβλητές Κατανάλωση αλκοόλ - Καρδιαγγειακή νόσος



Διάγραμμα 4.4: Ραβδόγραμμα συσχέτισης των μεταβλητών Καρδιαγγειακή νόσος – Αλκοόλ

- Έλεγχος συσχέτισης της Φυσική Δραστηριότητα με την εμφάνιση καρδιαγγειακής νόσου

Φυσική Δραστηριότητα * Καρδιαγγειακή νόσος Crosstabulation

| | | Καρδιαγγειακή νόσος | | Total | |
|----------------------|-----|---------------------|---------|---------|---------|
| | | no | yes | | |
| Φυσική Δραστηριότητα | no | Count | 6093 | 6735 | 12828 |
| | | Expected Count | 6549,4 | 6278,6 | 12828,0 |
| | yes | Count | 27431 | 25403 | 52834 |
| | | Expected Count | 26974,6 | 25859,4 | 52834,0 |
| Total | | Count | 33524 | 32138 | 65662 |
| | | Expected Count | 33524,0 | 32138,0 | 65662,0 |

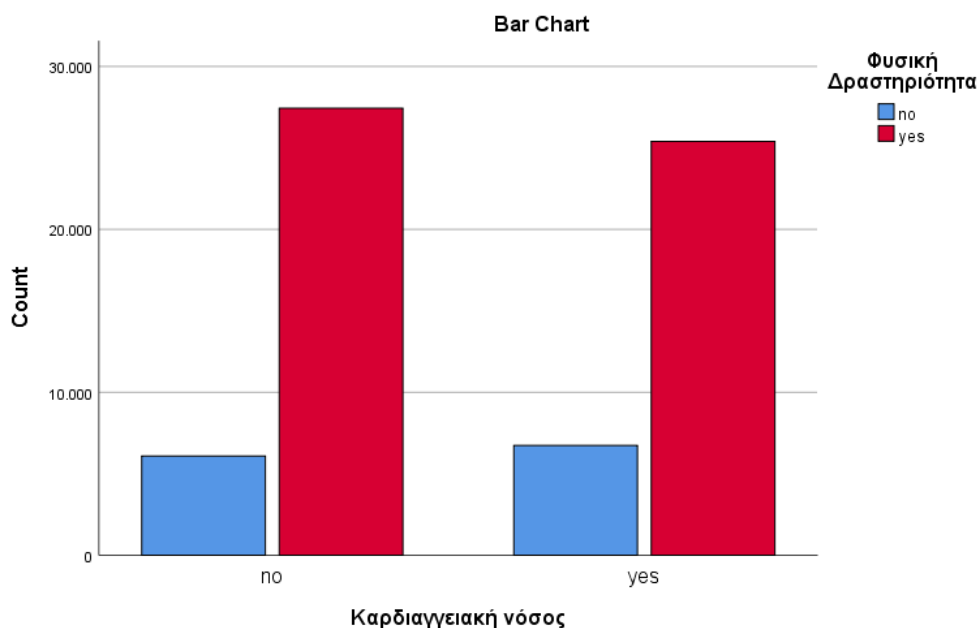
Πίνακας 4.11: Πίνακας συνάφειας των μεταβλητών Φυσική Δραστηριότητα - Καρδιαγγειακή νόσος

Στον παρακάτω πίνακα παρατηρούμε ότι απορρίπτουμε την μηδενική υπόθεση του ελέγχου χ^2 καθώς το p-value είναι μικρότερο από το 5% (p-value $\approx 0 < 0,05$). Επομένως, οι μεταβλητές “Φυσική Δραστηριότητα” και “Καρδιαγγειακή Νόσος” συσχετίζονται μεταξύ τους.

Chi-Square Tests

| | Value | df | Asymptotic Significance (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) |
|------------------------------------|---------------------|----|-----------------------------------|----------------------|----------------------|
| Pearson Chi-Square | 80,754 ^a | 1 | ,000 | | |
| Continuity Correction ^b | 80,577 | 1 | ,000 | | |
| Likelihood Ratio | 80,748 | 1 | ,000 | | |
| Fisher's Exact Test | | | | ,000 | ,000 |
| Linear-by-Linear Association | 80,752 | 1 | ,000 | | |
| N of Valid Cases | 65662 | | | | |

Πίνακας 4.12: Πίνακας ελέγχου – χ^2 για τις μεταβλητές Φυσική Δραστηριότητα - Καρδιαγγειακή νόσ



Διάγραμμα 4.5: Ραβδόγραμμα συσχέτισης των μεταβλητών Καρδιαγγειακή νόσος – Φυσική Δραστηριότητα

4.2.2 Έλεγχος συσχέτισης ποσοτικών μεταβλητών με την μεταβλητή “Καρδιαγγειακή νόσος”

| One-Sample Kolmogorov-Smirnov Test | | | | | | |
|------------------------------------|----------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | | Ηλικία | Βάρος | ΔΜΣ | ΣΑΠ | ΔΑΠ |
| N | | 65662 | 65662 | 65662 | 65662 | 65662 |
| Normal Parameters ^{a,b} | Mean | 53,33 | 72,25 | 26,8236 | 126,43 | 81,34 |
| | Std. Deviation | 6,764 | 11,303 | 4,26082 | 16,714 | 9,880 |
| Most Extreme Differences | Absolute | ,088 | ,070 | ,073 | ,247 | ,267 |
| | Positive | ,060 | ,070 | ,073 | ,247 | ,267 |
| | Negative | -,088 | -,035 | -,038 | -,161 | -,239 |
| Test Statistic | | ,088 | ,070 | ,073 | ,247 | ,267 |
| Asymp. Sig. (2-tailed) | | ,000 ^c | ,000 ^c | ,000 ^c | ,000 ^c | ,000 ^c |

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

Πίνακας 4. 13: Πίνακας ελέγχου κανονικότητας για ποσοτικές μεταβλητές

Παρατηρούμε από τον παραπάνω πίνακα ότι σε επίπεδο σημαντικότητας 5% θα απορρίψουμε την μηδενική υπόθεση ($p\text{-value} < 0.05$) για όλες τις μεταβλητές, καθώς

το p-value είναι περίπου μηδέν. Επομένως, καμία από τις ποσοτικές μας μεταβλητές δεν ακολουθεί την κανονική κατανομή δηλαδή για την συσχέτιση των μεταβλητών θα χρησιμοποιήσουμε μη παραμετρικούς ελέγχους υποθέσεων.

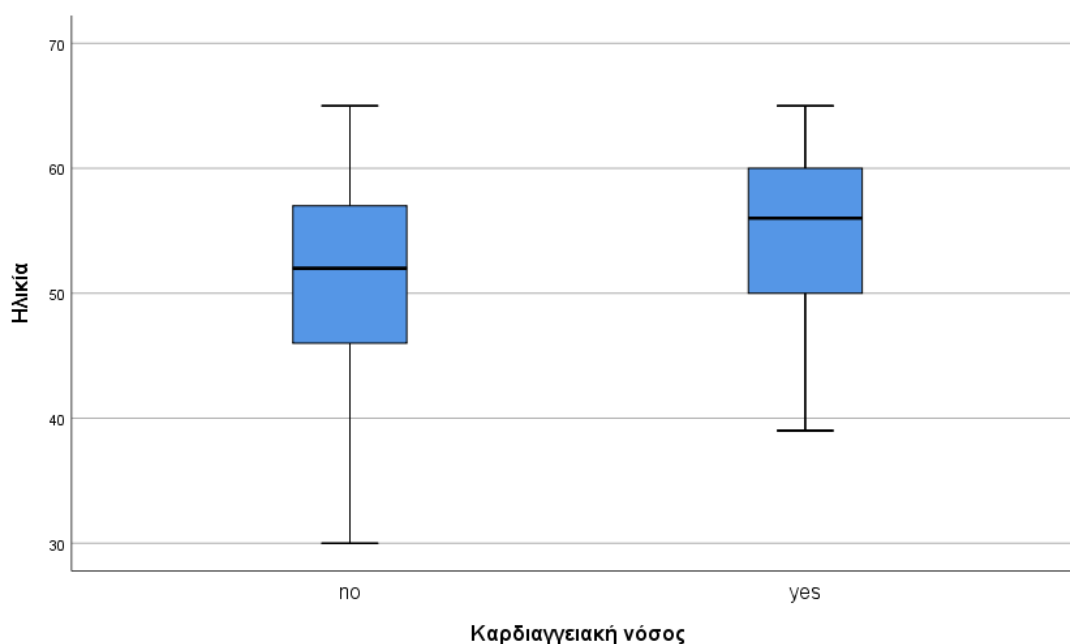
- Έλεγχος συσχέτισης της ηλικίας με την εμφάνιση καρδιαγγειακής νόσου

| Test Statistics ^a | |
|------------------------------|---------------|
| Ηλικία | |
| Mann-Whitney U | 388890763,000 |
| Wilcoxon W | 950836813,000 |
| Z | -61,819 |
| Asymp. Sig. (2-tailed) | ,000 |

a. Grouping Variable: Καρδιαγγειακή νόσος

Πίνακας 4. 14: Πίνακας συσχέτισης των μεταβλητών Ηλικία - Στεφανιαίας νόσος

Παρατηρούμε ότι για την μεταβλητή της ηλικίας το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές της ηλικίας και της εμφάνισης της καρδιαγγειακής νόσου συσχετίζονται.



Διάγραμμα 4. 6: Θηκογράμματα συσχέτισης των μεταβλητών Ηλικία – καρδιαγγειακή νόσο

- Έλεγχος συσχέτισης του βάρους με την εμφάνιση στεφανιαίας νόσου

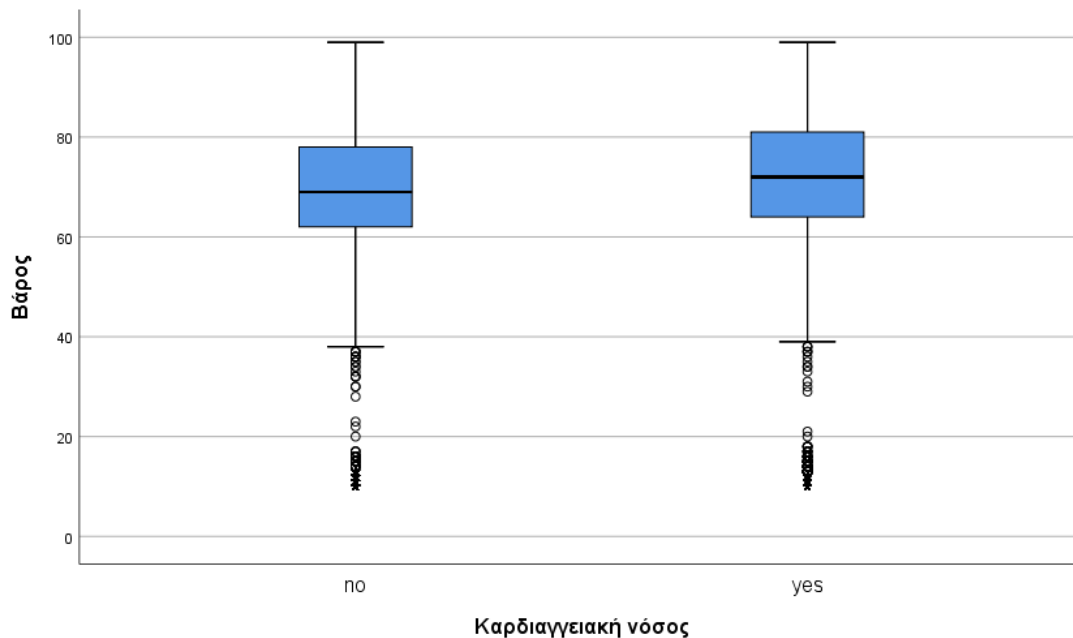
Test Statistics^a

| Βάρος | |
|------------------------|----------------|
| Mann-Whitney U | 438118947,000 |
| Wilcoxon W | 1000064997,000 |
| Z | -41,444 |
| Asymp. Sig. (2-tailed) | ,000 |

a. Grouping Variable: Καρδιαγγειακή νόσος

Πίνακας 4. 15: Πίνακας συσχέτισης των μεταβλητών Βάρος - Στεφανιαίας νόσος

Στον πίνακα παρατηρούμε ότι το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές του βάρους και της εμφάνισης της καρδιαγγειακής νόσου συσχετίζονται.



Διάγραμμα 4. 7:Θηκογράμματα συσχέτισης των μεταβλητών Βάρος– καρδιαγγειακή νόσος

- Έλεγχος συσχέτισης του Δείκτη Μάζας Σώματος (ΔΜΣ) με την εμφάνιση στεφανιαίας νόσου

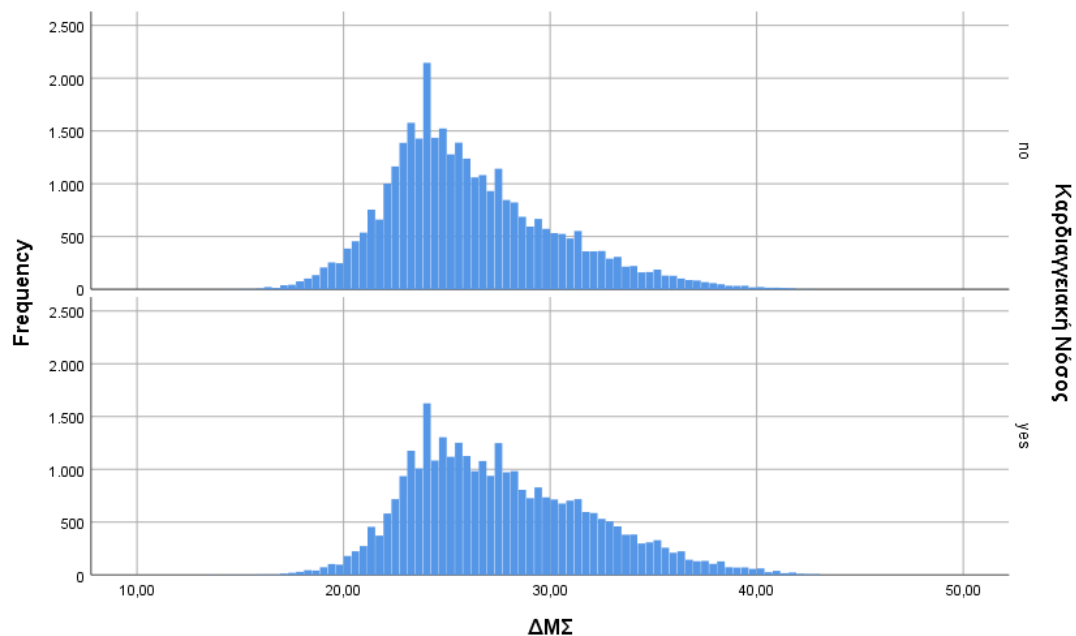
Test Statistics^a

| | ΔΜΣ |
|------------------------|---------------|
| Mann-Whitney U | 426125358,000 |
| Wilcoxon W | 988071408,000 |
| Z | -46,363 |
| Asymp. Sig. (2-tailed) | ,000 |

a. Grouping Variable: Καρδιαγγειακή νόσος

Πίνακας 4. 16: Πίνακας συσχέτισης των μεταβλητών ΔΜΣ - Στεφανιαίας νόσος

Στον πίνακα παρατηρούμε ότι το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές του ΔΜΣ και της εμφάνισης της καρδιαγγειακής νόσου συσχετίζονται.



Διάγραμμα 4. 8: Ραβδόγραμμα συσχέτισης των μεταβλητών ΔΜΣ – καρδιαγγειακή νόσος

- Έλεγχος συσχέτισης της ΣΑΠ με την εμφάνιση στεφανιαίας νόσου

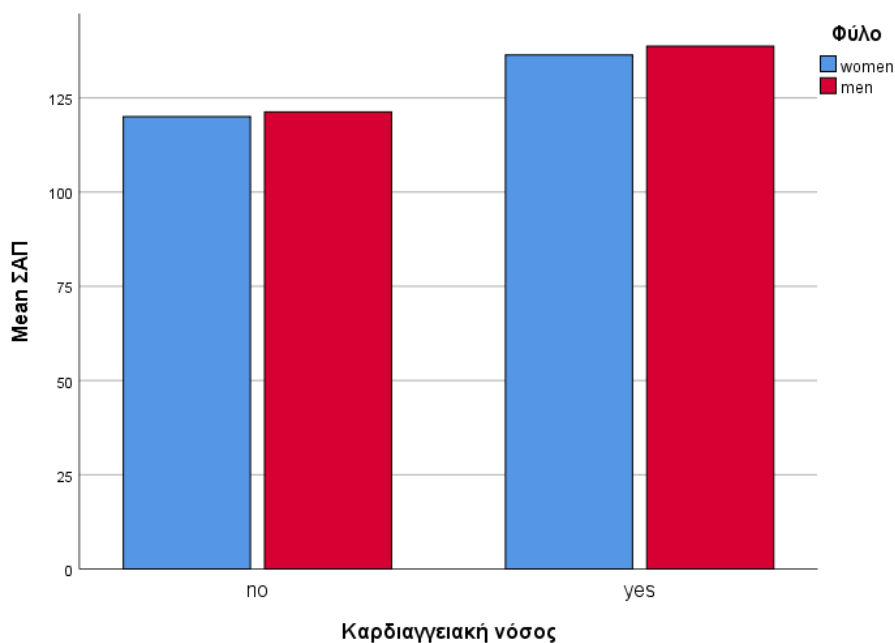
Test Statistics^a

| | ΣΑΠ |
|------------------------|---------------|
| Mann-Whitney U | 268076953,000 |
| Wilcoxon W | 830023003,000 |
| Z | -115,862 |
| Asymp. Sig. (2-tailed) | ,000 |

a. Grouping Variable: Καρδιαγγειακή νόσος

Πίνακας 4. 17: Πίνακας συσχέτισης των μεταβλητών ΣΑΠ - Στεφανιαίας νόσος

Στον πίνακα παρατηρούμε ότι το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές της ΣΑΠ και της εμφάνισης της καρδιαγγειακής νόσου συσχετίζονται.



Διάγραμμα 4. 9: Ραβδόγραμμα συσχέτισης των μεταβλητών ΣΑΠ – καρδιαγγειακή νόσος

- Έλεγχος συσχέτισης της ΔΑΠ με την εμφάνιση στεφανιαίας νόσου

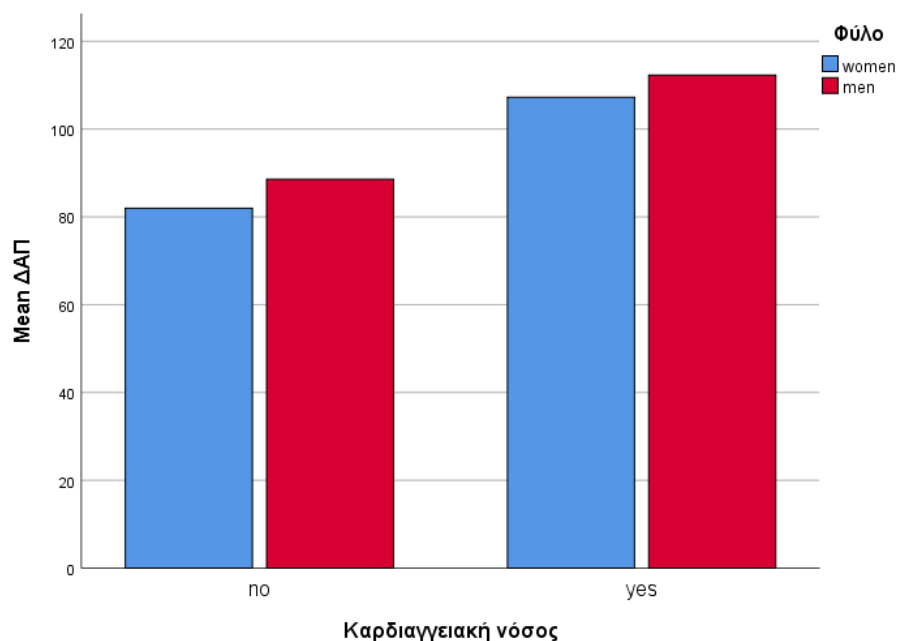
Test Statistics^a

| | ΔΑΠ |
|------------------------|---------------|
| Mann-Whitney U | 331685689,000 |
| Wilcoxon W | 893631739,000 |
| Z | -91,976 |
| Asymp. Sig. (2-tailed) | ,000 |

a. Grouping Variable: Καρδιαγγειακή νόσος

Πίνακας 4. 18: Πίνακας συσχέτισης των μεταβλητών ΔΑΠ - Στεφανιαίας νόσος

Στον πίνακα παρατηρούμε ότι το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές της ΔΑΠ και της εμφάνισης της καρδιαγγειακής νόσου συσχετίζονται.



Διάγραμμα 4. 10: Ραβδόγραμμα συσχέτισης των μεταβλητών ΔΑΠ – καρδιαγγειακή νόσος

4.3 Συσχετίσεις όλων μεταβλητών με την ύπαρξη Στεφανιαίας Νόσου για την 2^η βάση που αφορά μόνο άνδρες

4.3.1 Έλεγχος συσχέτισης ποιοτικών μεταβλητών με την μεταβλητή “Στεφανιαία Νόσος

- Έλεγχος συσχέτισης του καπνίσματος με την εμφάνιση στεφανιαίας νόσου

Κάπνισμα * Στεφανιαία Νόσος Crosstabulation

| | | Στεφανιαία Νόσος | | Total | |
|----------|-----|------------------|--------|-------|--------|
| | | no | yes | | |
| Κάπνισμα | no | Count | 1554 | 98 | 1652 |
| | | Expected Count | 1517,4 | 134,6 | 1652,0 |
| | yes | Count | 1343 | 159 | 1502 |
| | | Expected Count | 1379,6 | 122,4 | 1502,0 |
| Total | | Count | 2897 | 257 | 3154 |
| | | Expected Count | 2897,0 | 257,0 | 3154,0 |

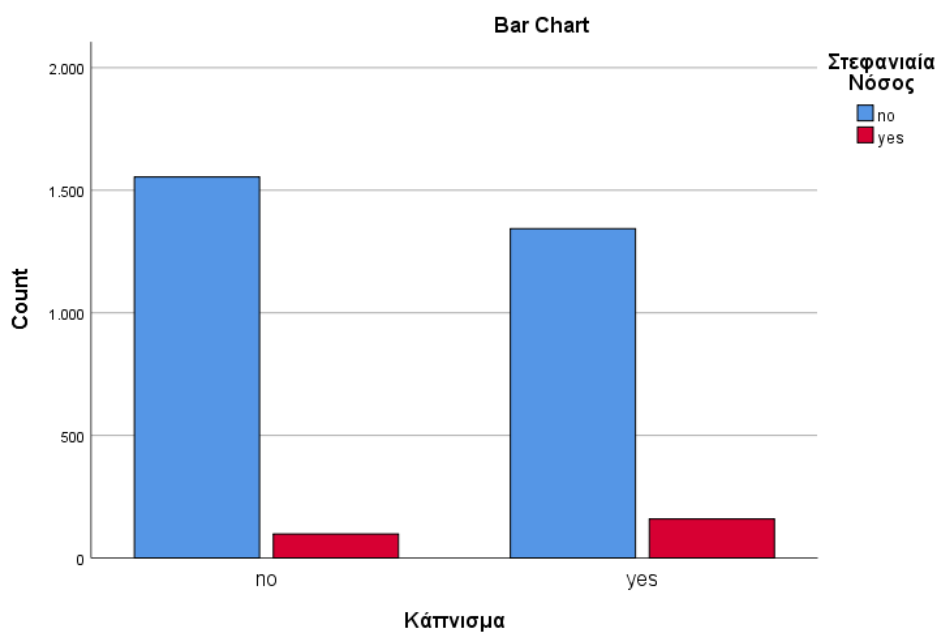
Πίνακας 4.19: Πίνακας συνάφειας των μεταβλητών Κάπνισμα - Στεφανιαία νόσος

Από τον πίνακα 4.20 παρατηρούμε ότι απορρίπτουμε την μηδενική υπόθεση του ελέγχου χ^2 του Pearson καθώς δεν είναι στατιστικά σημαντικό ($p\text{-value} \approx 0 < 0,05$) σε επίπεδο σημαντικότητας 5%. Επομένως, οι μεταβλητές “Κάπνισμα” και “Στεφανιαία Νόσος” συσχετίζονται μεταξύ τους.

Chi-Square Tests

| | Value | df | Asymptotic Significance (2- sided) | Exact Sig. (2- sided) | Exact Sig. (1- sided) |
|------------------------------------|---------------------|----|--|--------------------------|--------------------------|
| Pearson Chi-Square | 22,764 ^a | 1 | ,000 | | |
| Continuity Correction ^b | 22,147 | 1 | ,000 | | |
| Likelihood Ratio | 22,863 | 1 | ,000 | | |
| Fisher's Exact Test | | | | ,000 | ,000 |
| Linear-by-Linear Association | 22,757 | 1 | ,000 | | |
| N of Valid Cases | 3154 | | | | |

Πίνακας 4.20: Πίνακας ελέγχου – χ^2 για τις μεταβλητές Κάπνισμα - Στεφανιαία νόσος



Διάγραμμα 4.11: Ραβδόγραμμα συσχέτισης των μεταβλητών Κάπνισμα – Στεφανιαία Νόσος

- Έλεγχος συσχέτισης του τύπου Προσωπικότητας με την εμφάνιση στεφανιαίας νόσου

Τύπος Προσωπικότητας * Στεφανιαία Νόσος Crosstabulation

| | | Στεφανιαία Νόσος | | Total | |
|----------------------|--------|------------------|--------|-------|--------|
| | | no | yes | | |
| Τύπος Προσωπικότητας | Type A | Count | 1411 | 178 | 1589 |
| | | Expected Count | 1459,5 | 129,5 | 1589,0 |
| | Type B | Count | 1486 | 79 | 1565 |
| | | Expected Count | 1437,5 | 127,5 | 1565,0 |
| Total | | Count | 2897 | 257 | 3154 |
| | | Expected Count | 2897,0 | 257,0 | 3154,0 |

Πίνακας 4.21: Πίνακας συνάφειας των μεταβλητών Τύπος Προσωπικότητας - Στεφανιαία νόσος

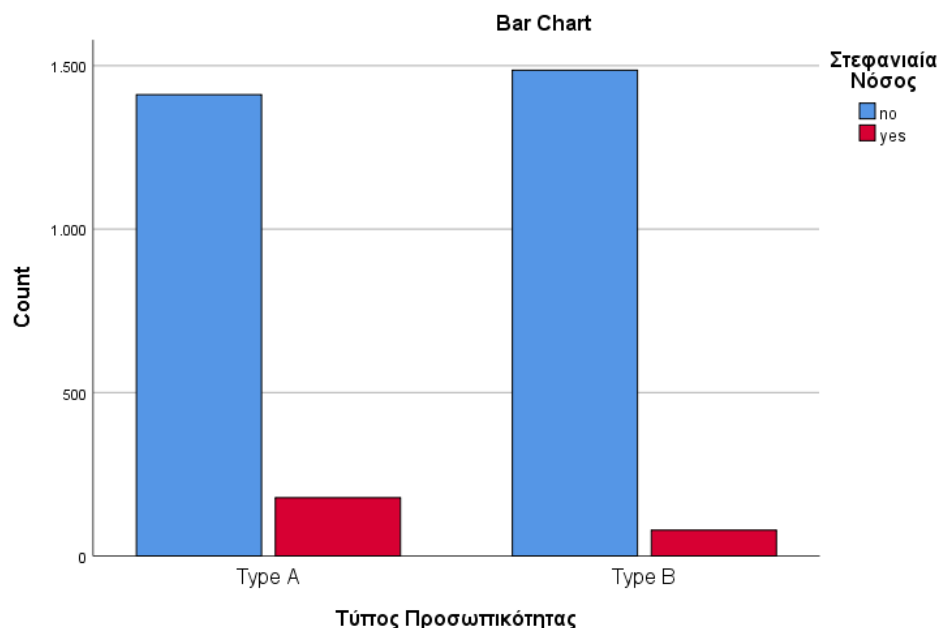
Chi-Square Tests

| | Value | df | Asymptotic Significance (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) |
|------------------------------------|---------------------|----|-----------------------------------|----------------------|----------------------|
| Pearson Chi-Square | 39,898 ^a | 1 | ,000 | | |
| Continuity Correction ^b | 39,080 | 1 | ,000 | | |
| Likelihood Ratio | 40,900 | 1 | ,000 | | |
| Fisher's Exact Test | | | | ,000 | ,000 |
| Linear-by-Linear Association | 39,885 | 1 | ,000 | | |
| N of Valid Cases | 3154 | | | | |

Πίνακας 4.22: Πίνακας ελέγχου χ^2 για τις μεταβλητές Τύπος Προσωπικότητας - Στεφανιαία νόσος

Από τον παραπάνω πίνακα παρατηρούμε ότι απορρίπτουμε την μηδενική υπόθεση του ελέγχου χ^2 του Pearson καθώς δεν είναι στατιστικά σημαντικό ($p\text{-value} \approx 0 < 0,05$) σε

επίπεδο σημαντικότητας 5%. Επομένως, οι μεταβλητές “Τύπος Προσωπικότητας” και “Στεφανιαία Νόσος” συσχετίζονται μεταξύ τους.



Διάγραμμα 4. 12: Ραβδόγραμμα συσχέτισης των μεταβλητών Τύπος Προσωπικότητας – Στεφανιαία Νόσο

4.3.2 Έλεγχος συσχέτισης ποσοτικών μεταβλητών με την μεταβλητή “Στεφανιαία νόσος”

One-Sample Kolmogorov-Smirnov Test

| | Ηλικία | ΔΜΣ | Βάρος | Χοληστερόλη | ΔΑΠ | ΣΑΠ | Αριθμός Τσιγάρων | |
|----------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------|
| N | 3154 | 3154 | 3154 | 3142 | 3154 | 3154 | 3154 | |
| Normal Parameters ^{a,b} | Mean | 46,28 | 24,5186 | 77,10 | 226,37 | 82,02 | 128,63 | 11,60 |
| | Std. Deviation | 5,524 | 2,56755 | 9,541 | 43,420 | 9,727 | 15,118 | 14,518 |
| Most Extreme Differences | Absolute | ,119 | ,065 | ,076 | ,037 | ,121 | ,113 | ,312 |
| | Positive | ,119 | ,065 | ,076 | ,037 | ,121 | ,113 | ,312 |
| | Negative | -,094 | -,026 | -,046 | -,020 | -,053 | -,069 | -,212 |
| Test Statistic | ,119 | ,065 | ,076 | ,037 | ,121 | ,113 | ,312 | |
| Asymp. Sig. (2-tailed) | ,000 ^c | ,000 ^c | ,000 ^c | ,000 ^c | ,000 ^c | ,000 ^c | ,000 ^c | |

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

Πίνακας 4. 23: Έλεγχος Κανονικότητας Συνεχών μεταβλητών

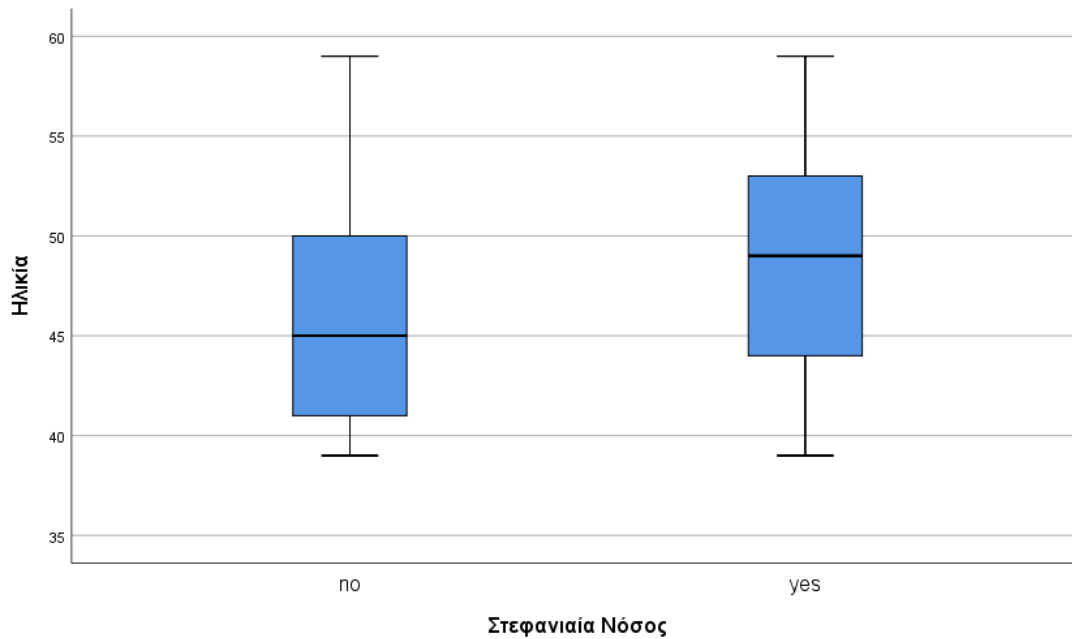
Παρατηρούμε από τον παραπάνω πίνακα ότι σε επίπεδο σημαντικότητας 5% θα απορρίψουμε την μηδενική υπόθεση ($p\text{-value} < 0.05$) για όλες τις μεταβλητές, καθώς το $p\text{-value}$ είναι περίπου μηδέν. Επομένως, καμία από τις ποσοτικές μας μεταβλητές δεν ακολουθεί την κανονική κατανομή δηλαδή για την συσχέτιση των μεταβλητών θα χρησιμοποιήσουμε μη παραμετρικούς ελέγχους υποθέσεων.

- Έλεγχος συσχέτισης της ηλικίας με την εμφάνιση στεφανιαίας νόσου

| Test Statistics ^a | |
|------------------------------|-------------|
| | Ηλικία |
| Mann-Whitney U | 282978,500 |
| Wilcoxon W | 4480731,500 |
| Z | -6,394 |
| Asymp. Sig. (2-tailed) | ,000 |

Πίνακας 4. 24: Πίνακας συσχέτισης της ηλικίας με την εμφάνιση στεφανιαίας νόσου

Παρατηρούμε ότι για την μεταβλητή της ηλικίας το $p\text{-value}$ είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές της ηλικίας και της εμφάνισης της στεφανιαίας νόσου συσχετίζονται.



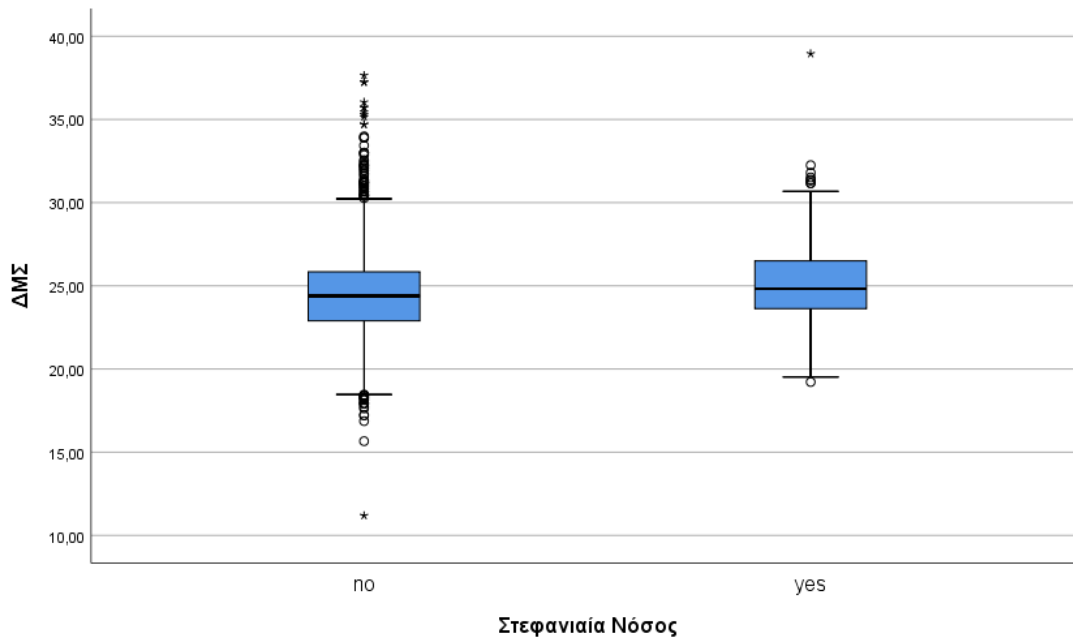
Διάγραμμα 4. 43: Θηκογράμματα συσχέτισης των μεταβλητών Ηλικία – Στεφανιαία Νόσος

- Έλεγχος συσχέτισης του ΔΜΣ με την εμφάνιση στεφανιαίας νόσου

| Test Statistics ^a | |
|------------------------------|-------------|
| | ΔΜΣ |
| Mann-Whitney U | 322985,500 |
| Wilcoxon W | 4520738,500 |
| Z | -3,522 |
| Asymp. Sig. (2-tailed) | ,000 |

Πίνακας 4. 25: Πίνακας συσχέτισης του ΔΜΣ με την εμφάνιση στεφανιαίας νόσου

Παρατηρούμε ότι για την μεταβλητή του ΔΜΣ το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές του ΔΜΣ και της εμφάνισης της στεφανιαίας νόσου συσχετίζονται.



Διάγραμμα 4.14: Θηκογράμματα συσχέτισης των μεταβλητών ΔΜΣ – Στεφανιαία Νόσος

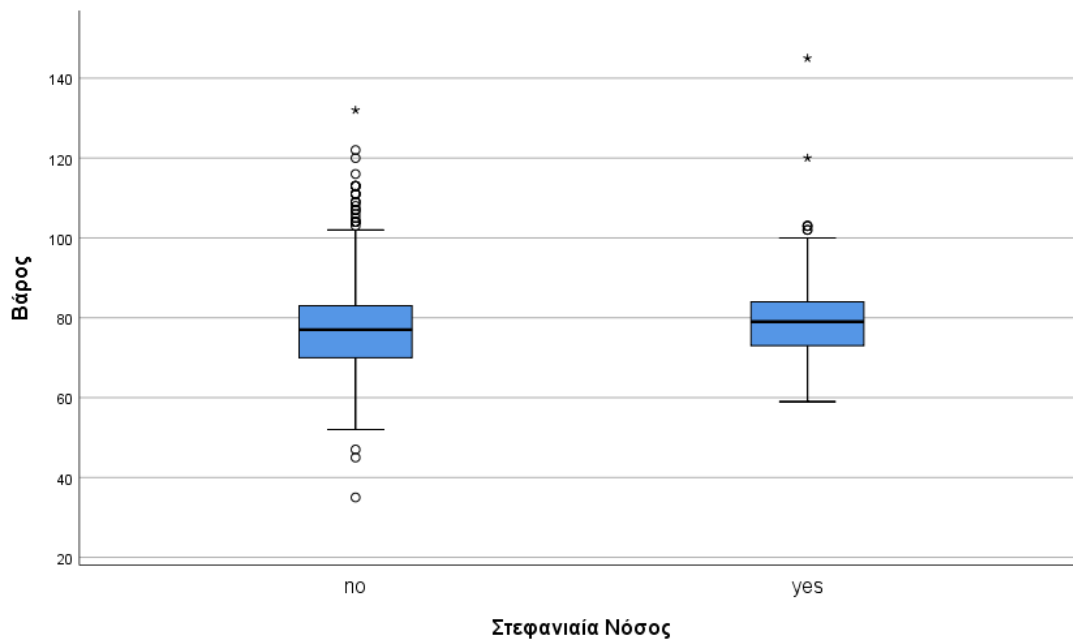
- Έλεγχος συσχέτισης του βάρους ενός ατόμου με την εμφάνιση στεφανιαίας νόσου

Test Statistics^a

| | Βάρος |
|------------------------|-------------|
| Mann-Whitney U | 320292,500 |
| Wilcoxon W | 4518045,500 |
| Z | -3,719 |
| Asymp. Sig. (2-tailed) | ,000 |

Πίνακας 4. 26: Πίνακας συσχέτισης του βάρους με την εμφάνιση στεφανιαίας νόσου

Παρατηρούμε ότι για την μεταβλητή του βάρους το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές του βάρους και της εμφάνισης της στεφανιαίας νόσου συσχετίζονται.



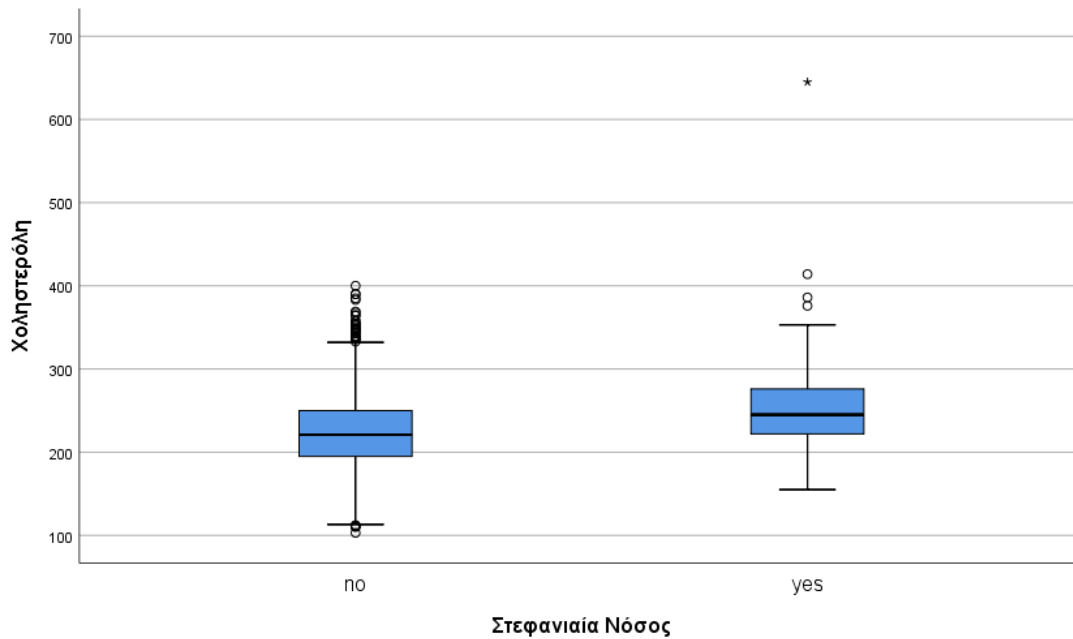
Διάγραμμα 4.15: Θηκογράμματα συσχέτισης των μεταβλητών Βάρος – Στεφανιαία Νόσος

- Έλεγχος συσχέτισης της τιμής της χοληστερίνης με την εμφάνιση στεφανιαίας νόσου

| Test Statistics ^a | |
|------------------------------|-------------|
| | Χοληστερίνη |
| Mann-Whitney U | 250258,500 |
| Wilcoxon W | 4413313,500 |
| Z | -8,645 |
| Asymp. Sig. (2-tailed) | ,000 |

Πίνακας 4. 27: Πίνακας συσχέτισης της χοληστερίνης με την εμφάνιση στεφανιαίας νόσου

Παρατηρούμε ότι για την μεταβλητή της χοληστερίνης το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές της χοληστερίνης και της εμφάνισης της στεφανιαίας νόσου συσχετίζονται.



Διάγραμμα 4.16: Θηκογράμματα συσχέτισης των μεταβλητών Χοληστερόλης – Στεφανιαία Νόσος

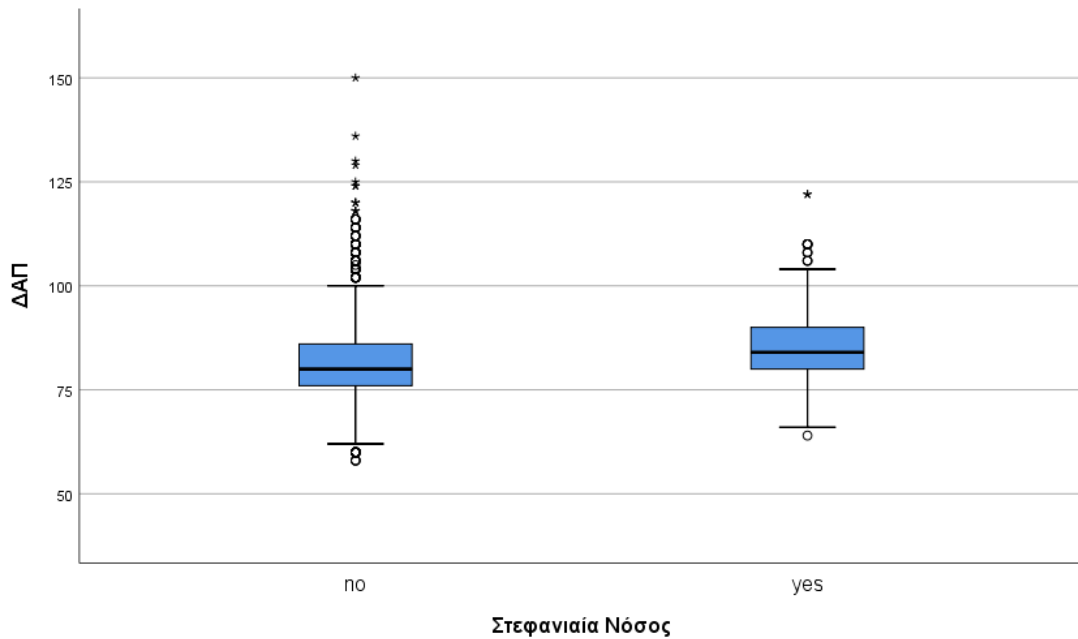
- Έλεγχος συσχέτισης της Διαστολικής αρτηριακής πίεσης με την εμφάνιση στεφανιαίας νόσου

Test Statistics^a

| | Διαστολική Αρτηριακή Πίεση |
|------------------------|----------------------------------|
| Mann-Whitney U | 296588,500 |
| Wilcoxon W | 4494341,500 |
| Z | -5,430 |
| Asymp. Sig. (2-tailed) | ,000 |

Πίνακας 4. 28: Πίνακας συσχέτισης της ΔΑΠ με την εμφάνιση στεφανιαίας νόσου

Παρατηρούμε ότι για την μεταβλητή της ΔΑΠ το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές της ΔΑΠ και της εμφάνισης της στεφανιαίας νόσου συσχετίζονται.



Διάγραμμα 4.17: Θηκογράμματα συσχέτισης των μεταβλητών ΔΑΠ – Στεφανιαία Νόσος

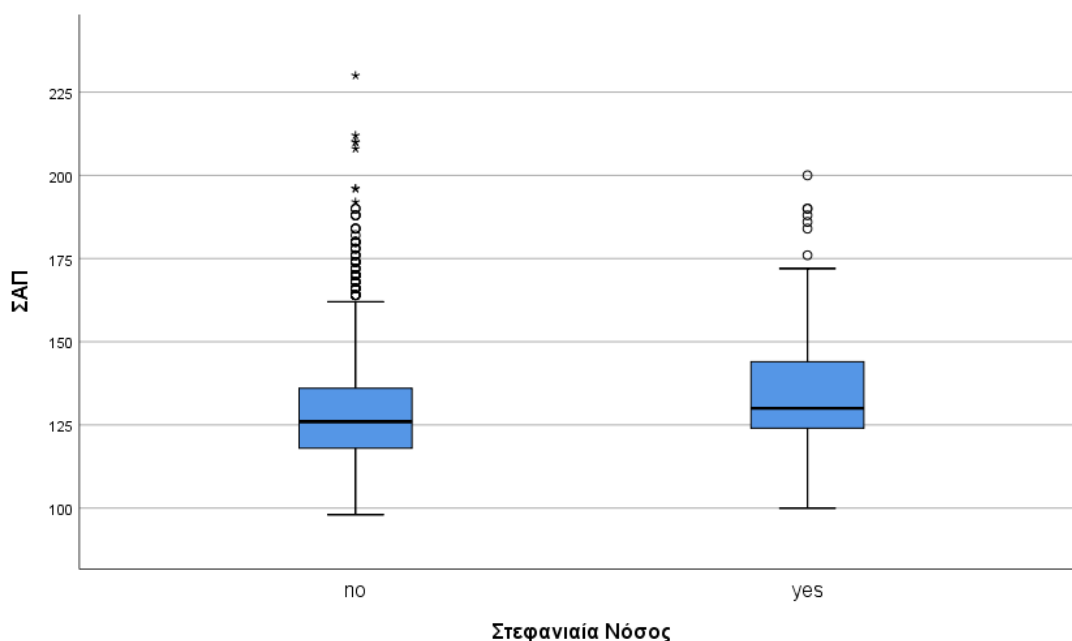
- Έλεγχος συσχέτισης της Συστολικής αρτηριακής πίεσης με την εμφάνιση στεφανιαίας νόσου

Test Statistics^a

| | Συστολική Αρτηριακή Πίεση |
|------------------------|---------------------------------|
| Mann-Whitney U | 274441,500 |
| Wilcoxon W | 4472194,500 |
| Z | -7,005 |
| Asymp. Sig. (2-tailed) | ,000 |

Πίνακας 4. 29: Πίνακας συσχέτισης της ΣΑΠ με την εμφάνιση στεφανιαίας νόσου

Παρατηρούμε ότι για την μεταβλητή της ΣΑΠ το p-value είναι περίπου μηδέν, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές της ΣΑΠ και της εμφάνισης της στεφανιαίας νόσου συσχετίζονται.



Διάγραμμα 4.18: Θηκογράμματα συσχέτισης των μεταβλητών ΣΑΠ – Στεφανιαία Νόσος

- Έλεγχος συσχέτισης του ημερήσιου αριθμού των τσιγάρων με την εμφάνιση στεφανιαίας νόσου

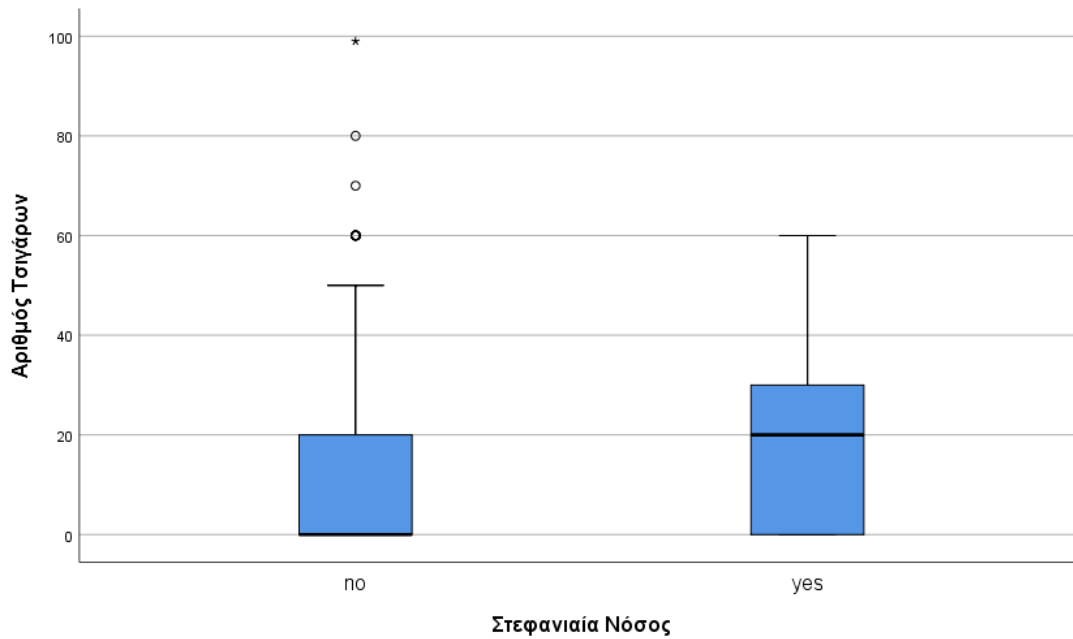
Test Statistics^a

| | Αριθμός τσιγάρων |
|------------------------|---------------------|
| Mann-Whitney U | 90570,000 |
| Wilcoxon W | 993066,000 |
| Z | -3,191 |
| Asymp. Sig. (2-tailed) | ,001 |

a. Grouping Variable: Στεφανιαία Νόσος

Πίνακας 4. 30: Πίνακας συσχέτισης του ημερήσιου αριθμού τσιγάρων με την εμφάνιση στεφανιαίας νόσου

Παρατηρούμε ότι για την μεταβλητή του αριθμού των ημερήσιων τσιγάρων το $p\text{-value} = 0.001$, άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως οι μεταβλητές του αριθμού των ημερήσιων τσιγάρων και της εμφάνισης της στεφανιαίας νόσου συσχετίζονται.



Διάγραμμα 4.19: Θηκογράμματα συσχέτισης των μεταβλητών Αριθμός Τσιγάρων – Στεφανιαία Νόσος

4.4 Έλεγχος συσχέτισης της ηλικίας των ατόμων που εμφάνισαν Καρδιαγγειακή νόσο ανάλογα με το φύλο τους

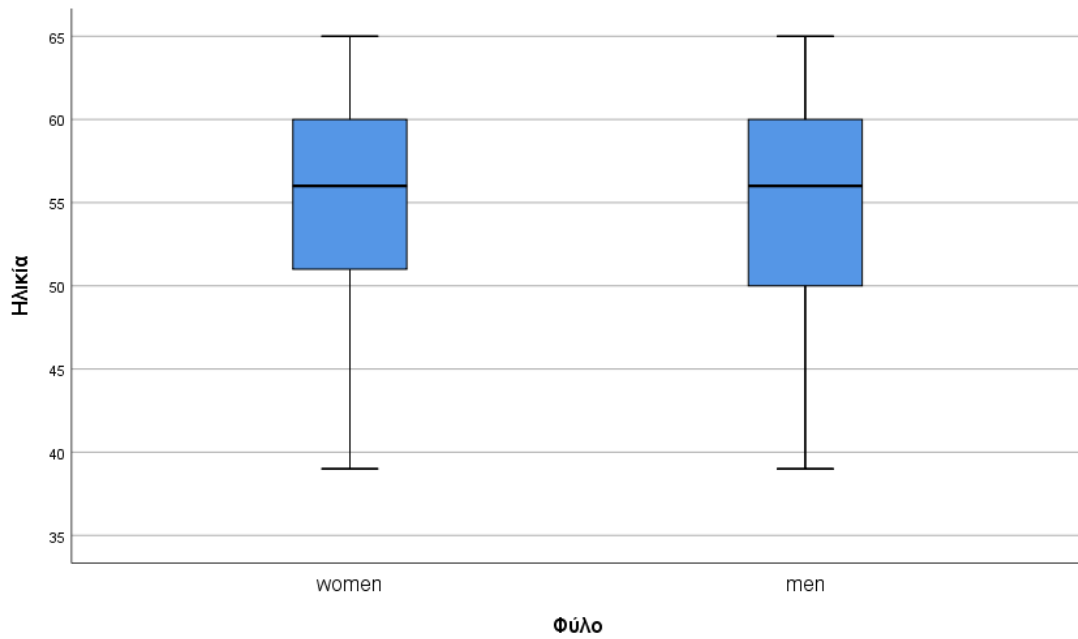
| Test Statistics ^a | |
|------------------------------|---------------|
| | Ηλικία |
| Mann-Whitney U | 111840935,500 |
| Wilcoxon W | 173762691,500 |
| Z | -6,408 |
| Asymp. Sig. (2-tailed) | ,000 |

a. Grouping Variable: Φύλο

Πίνακας 4. 31: Πίνακας συσχέτισης της ηλικίας με το φύλο τους για τα άτομα που εμφάνισαν καρδιαγγειακή νόσο

Παρατηρούμε ότι για την μεταβλητή της ηλικία ανάμεσα στα δύο φύλα έχουμε p-value περίπου 0 , άρα σε επίπεδο σημαντικότητας 5% απορρίπτουμε την μηδενική υπόθεση

ότι τα δύο δείγματα έχουν την ίδια κατανομή. Επομένως η μεταβλητή της ηλικίας ανάμεσα σε άνδρες και γυναίκες δεν ακολουθεί την ίδια κατανομή, δηλαδή η ηλικία εμφάνισης καρδιαγγειακής νόσου σχετίζεται με το φύλο του ατόμου και όπως φαίνεται και στο θηκόγραμμα παρακάτω οι άνδρες εμφάνισαν τη νόσο σε νεαρότερη ηλικία από τις γυναίκες.



Διάγραμμα 4. 20: Θηκογράμματα συσχέτισης των μεταβλητών Ηλικίας – Φύλο για τα άτομα που εμφάνισαν καρδιαγγειακή νόσο

5^ο ΚΕΦΑΛΑΙΟ

ΠΑΛΙΝΔΡΟΜΗΣΗ

5.1 Λογιστική παλινδρόμηση με εξαρτημένη μεταβλητή την εμφάνιση καρδιαγγειακής νόσου για την 1^η βάση δεδομένων

➤ Πίνακας Case processing Summary

Στον παρακάτω πίνακα παρατηρούμε ότι από τις συνολικά 65.662 περιπτώσεις έχουν χρησιμοποιηθεί όλες οι περιπτώσεις καθώς δεν υπάρχουν ελλείπουσες τιμές (Missing Cases).

| Unweighted Cases ^a | | N | Percent |
|-------------------------------|----------------------|-------|---------|
| Selected Cases | Included in Analysis | 65662 | 100,0 |
| | Missing Cases | 0 | ,0 |
| | Total | 65662 | 100,0 |
| Unselected Cases | | 0 | ,0 |
| Total | | 65662 | 100,0 |

a. If weight is in effect, see classification table for the total number of cases.

Πίνακας 5. 1: Πίνακας περιγραφής του δείγματος

➤ Έλεγχος Omnibus

Ο έλεγχος Omnibus εξετάζει την σημαντικότητα του μοντέλου σε σχέση με το null μοντέλο (είναι το μοντέλο το οποίο περιέχει μόνο τον σταθερό όρο). Οι υποθέσεις του ελέγχου είναι οι εξής:

H_0 : το null μοντέλο ξεπερνά το υπό εξέταση μοντέλο

H_1 : το υπό εξέταση μοντέλο ξεπερνά το null μοντέλο

Omnibus Tests of Model Coefficients

| | | Chi-square | df | Sig. |
|--------|-------|------------|----|------|
| Step 1 | Step | 17611,428 | 13 | ,000 |
| | Block | 17611,428 | 13 | ,000 |
| | Model | 17611,428 | 13 | ,000 |

Πίνακας 5. 2: Πίνακας ελέγχου Omnibus

Από τον παραπάνω πίνακα παρατηρούμε ότι το p-value είναι σχεδόν μηδέν, επομένως απορρίπτουμε την μηδενική υπόθεση. Δηλαδή το μοντέλο που προσαρμόσαμε είναι καλύτερο από το μοντέλο που περιέχει μόνο την σταθερά (null).

➤ Πίνακας προσαρμογής του μοντέλου

Model Summary

| Step | -2 Log likelihood | Cox & Snell R Square | Nagelkerke R Square |
|------|------------------------|----------------------|---------------------|
| 1 | 73386,175 ^a | ,235 | ,314 |

Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

Πίνακας 5. 3: Πίνακας ποσοστού προσαρμογής του μοντέλου

Όπως παρατηρούμε στον παραπάνω πίνακα το μοντέλο πολλαπλής παλινδρόμησης που προσαρμόσαμε το 31,4% της μεταβλητότητας της εξαρτημένης μεταβλητής ερμηνεύεται από τις επεξηγηματικές μεταβλητές.

➤ **Ακρίβεια πρόβλεψης του μοντέλου**

Classification Table^a

| Observed | | Predicted | | Percentage Correct |
|--------------------|------------------------|---------------------------|-------|--------------------|
| | | Καρδιαγγειακή νόσος no | yes | |
| Step 1 | Καρδιαγγειακή νόσος no | 26664 | 6860 | 79,5 |
| | yes | 10933 | 21205 | 66,0 |
| Overall Percentage | | | | 72,9 |

a. The cut value is ,500

Πίνακας 5. 4: Πίνακας ακρίβειας πρόβλεψης του μοντέλου

Στον παραπάνω πίνακα ταξινόμησης μπορούμε να συμπεράνουμε ότι το ποσοστό ορθής ταξινόμησης του μοντέλου είναι 72,9%.

Παρατηρούμε ότι το μοντέλο προβλέπει πως 37.597 (26.664+10.933) άτομα δεν εμφάνισαν την καρδιαγγειακή νόσο και 28.065(6.860+21.205) εμφάνισαν την καρδιαγγειακή νόσο.

Επιπλέον, 26.664 άτομα όντως δεν εμφάνισαν καρδιαγγειακή νόσο και 21.205 εμφάνισαν καρδιαγγειακή νόσο.

➤ **Πίνακας συντελεστών παλινδρόμησης του μοντέλου**

Το βέλτιστο μοντέλο είναι :

$$\log\left(\frac{p}{1-p}\right) = -11,413 + 0,052*\text{Ηλικία} + 0,007*\text{Βάρος} + 0,014*\text{ΔΜΣ} + 0,055*\text{ΣΑΠ} + 0,016*\text{ΔΑΠ} - 1,091*\text{Χοληστερόλη}(1) - 0,704*\text{Χοληστερόλη}(2) + 0,311*\text{Γλυκόζη}(1) + 0,339*\text{Γλυκόζη}(2) + 0,144*\text{Κάπνισμα}(1) + 0,213*\text{Αλκοόλ} + 0,228*\text{Καθιστική Ζωή}$$

| | | Variables in the Equation | | | | | |
|---------------------|-------------------------|---------------------------|------|----------|----|------|--------|
| | | B | S.E. | Wald | df | Sig. | Exp(B) |
| Step 1 ^a | Ηλικία | ,052 | ,001 | 1380,770 | 1 | ,000 | 1,053 |
| | Φύλο(1) | ,022 | ,023 | ,906 | 1 | ,341 | 1,022 |
| | Βάρος | ,007 | ,002 | 21,199 | 1 | ,000 | 1,007 |
| | ΔΜΣ | ,014 | ,004 | 11,686 | 1 | ,001 | 1,015 |
| | ΣΑΠ | ,055 | ,001 | 3740,271 | 1 | ,000 | 1,056 |
| | ΔΑΠ | ,016 | ,001 | 133,030 | 1 | ,000 | 1,016 |
| | Χοληστερόλη | | | 958,713 | 2 | ,000 | |
| | Χοληστερόλη(1) | -1,091 | ,037 | 856,718 | 1 | ,000 | ,336 |
| | Χοληστερόλη(2) | -,704 | ,044 | 257,930 | 1 | ,000 | ,495 |
| | Γλυκόζη | | | 58,930 | 2 | ,000 | |
| | Γλυκόζη(1) | ,311 | ,041 | 57,160 | 1 | ,000 | 1,365 |
| | Γλυκόζη(2) | ,339 | ,054 | 39,799 | 1 | ,000 | 1,403 |
| | Κάπνισμα(1) | ,144 | ,036 | 15,886 | 1 | ,000 | 1,155 |
| | Αλκοόλ(1) | ,213 | ,044 | 23,392 | 1 | ,000 | 1,237 |
| | Φυσική Δραστηριότητα(1) | ,228 | ,022 | 103,195 | 1 | ,000 | 1,256 |
| | Constant | -11,635 | ,146 | 6394,198 | 1 | ,000 | ,000 |

a. Variable(s) entered on step 1: Ηλικία, Φύλο, Βάρος, ΔΜΣ, ΣΑΠ, ΔΑΠ, Χοληστερόλη, Γλυκόζη, Κάπνισμα, Αλκοόλ, Φυσική Δραστηριότητα.

Πίνακας 5. 5: Πίνακας των συντελεστών του μοντέλου

Από τον παραπάνω πίνακα βλέπουμε ότι το μοντέλο λογιστικής παλινδρόμησης μέσω του ελέγχου Wald έκρινε ότι μόνο η μεταβλητή του φύλου είναι στατιστικά μη σημαντική στην εμφάνιση καρδιαγγειακής νόσου με $p\text{-value} = 0,341 > 0,05$

Ερμηνεία των συντελεστών του μοντέλου (Πίνακας 5.6)

- Αν αυξηθεί κατά μία μονάδα η ηλικία ενός ατόμου τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.052} = 1,053 > 1$. Δηλαδή, η πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 5,3%.
- Αν αυξηθεί κατά μία μονάδα το βάρος ενός ατόμου τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.007} = 1,007 > 1$. Δηλαδή, η πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 0,7%.

- Αν αυξηθεί κατά μία μονάδα ο ΔΜΣ ενός ατόμου τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.016} = 1,016 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 1,5%.
- Αν αυξηθεί κατά μία μονάδα η ΣΑΠ ενός ατόμου τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.055} = 1,056 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 5,6%.
- Αν αυξηθεί κατά μία μονάδα η ΔΑΠ ενός ατόμου τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.015} = 1,016 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 1,6%.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων με τιμή χοληστερόλης φυσιολογική είναι 0,339 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου είναι μειωμένη κατά $1 - 33,6\% = 66,4\%$ σε σχέση με ένα άτομο που εμφανίζει τιμή χοληστερόλης πολύ πάνω από το φυσιολογικό.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων με τιμή χοληστερόλης πάνω από το φυσιολογικό είναι 0,495 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου είναι μειωμένη κατά $1 - 49,5\% = 50,5\%$ σε σχέση με ένα άτομο που εμφανίζει τιμή χοληστερόλης πολύ πάνω από το φυσιολογικό.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων με φυσιολογική τιμή γλυκόζης είναι 1,355 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου είναι 35,5% πιο πιθανή σε σχέση με ένα άτομο που εμφανίζει τιμή γλυκόζης πολύ πάνω από το φυσιολογικό.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων με τιμή γλυκόζης πάνω από το φυσιολογικό είναι 1,396 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου είναι 39,6% πιο πιθανή σε σχέση με ένα άτομο που εμφανίζει τιμή γλυκόζης πολύ πάνω από το φυσιολογικό.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων που καπνίζουν είναι 1,155 δηλαδή έχουν 15,5% μεγαλύτερη σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου σε σχέση με τα άτομα που δεν καπνίζουν.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων που καταναλώνουν αλκοόλ είναι 1,237 δηλαδή έχουν 23,7% μεγαλύτερη σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου σε σχέση με τα άτομα που δεν καταναλώνουν αλκοόλ.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων που έχουν Φυσική Δραστηριότητα είναι 1,256 δηλαδή έχουν 25,6% μεγαλύτερη σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου σε σχέση με τα άτομα που δεν ασκούνται.

5.2 Λογιστική παλινδρόμηση με εξαρτημένη μεταβλητή την εμφάνιση στεφανιαίας νόσου για την 2^η βάση δεδομένων

Όπως είδαμε στο κεφάλαιο 4 όλες οι μεταβλητές της βάσης συσχετίζονται με την μεταβλητή “Στεφανιαία Νόσος”, επομένως μπορούμε να τις συμπεριλάβουμε όλες στο λογιστικό μοντέλο.

➤ Πίνακας Case processing Summary

Στον παρακάτω πίνακα παρατηρούμε ότι από τις συνολικά 3.154 περιπτώσεις έχουν χρησιμοποιηθεί οι 3.142 καθώς υπάρχουν 12 ελλείπουσες τιμές (Missing Cases).

| Case Processing Summary | | | |
|-------------------------------|----------------------|------|---------|
| Unweighted Cases ^a | | N | Percent |
| Selected Cases | Included in Analysis | 3142 | 99,6 |
| | Missing Cases | 12 | ,4 |
| Total | | 3154 | 100,0 |
| Unselected Cases | | 0 | ,0 |
| Total | | 3154 | 100,0 |

Πίνακας 5. 6: Πίνακας περιγραφής του δείγματος

➤ Έλεγχος Omnibus

Ο έλεγχος Omnibus εξετάζει την σημαντικότητα του μοντέλου σε σχέση με το null μοντέλο (είναι το μοντέλο το οποίο περιέχει μόνο τον σταθερό όρο). Οι υποθέσεις του ελέγχου είναι οι εξής:

H_0 : το null μοντέλο ξεπερνά το υπό εξέταση μοντέλο

H_1 : το υπό εξέταση μοντέλο ξεπερνά το null μοντέλο

Omnibus Tests of Model Coefficients

| | | Chi-square | df | Sig. |
|--------|-------|------------|----|------|
| Step 1 | Step | 193,942 | 8 | ,000 |
| | Block | 193,942 | 8 | ,000 |
| | Model | 193,942 | 8 | ,000 |

Πίνακας 5. 7: Πίνακας ελέγχου Omnibus

Από τον παραπάνω πίνακα παρατηρούμε ότι το p-value είναι σχεδόν μηδέν, επομένως απορρίπτουμε την μηδενική υπόθεση. Δηλαδή το μοντέλο που προσαρμόσαμε είναι καλύτερο από το μοντέλο που περιέχει μόνο την σταθερά (null).

➤ Έλεγχος Hosmer – Lemeshow

Ο έλεγχος των Hosmer – Lemeshow εξετάζει εάν το μοντέλο που κατασκευάσαμε προσαρμόζεται καλά στα δεδομένα ή όχι δηλαδή μας δείχνει πόσο κοντά βρίσκονται οι παρατηρούμενες και οι προβλεπόμενες πιθανότητες. Οι υποθέσεις του ελέγχου είναι οι εξής:

H_0 : το μοντέλο έχει καλή προσαρμογή
 H_1 : το μοντέλο δεν έχει καλή προσαρμογή

Hosmer and Lemeshow Test

| Step | Chi-square | df | Sig. |
|------|------------|----|------|
| 1 | 9,036 | 8 | ,339 |

Πίνακας 5. 8: Πίνακας ελέγχου καλής προσαρμογής Hosmer - Lemeshow

Παρατηρούμε ότι $p\text{-value} = 0.339 > 0.05$, επομένως σε επίπεδο σημαντικότητας 5% αποδεχόμαστε την μηδενική υπόθεση του ελέγχου. Άρα το μοντέλο μας προσαρμόζεται καλά στα δεδομένα.

➤ **Ακρίβεια πρόβλεψης του μοντέλου**

Classification Table^a

| | | Observed | Predicted | | Percentage Correct |
|--------------------|------------------|----------|------------------|----|--------------------|
| | | | Στεφανιαία Νόσος | No | |
| Step 1 | Στεφανιαία Νόσος | No | 2880 | 5 | 99,8 |
| | | Yes | 254 | 3 | 1,2 |
| Overall Percentage | | | | | 91,8 |

a. The cut value is ,500

Πίνακας 5. 9: Πίνακας ακρίβειας πρόβλεψης του μοντέλου

Από τον παραπάνω πίνακα μπορούμε να συμπεράνουμε ότι το ποσοστό ορθής ταξινόμησης του μοντέλου είναι 91,8%.

Παρατηρούμε ότι το μοντέλο προβλέπει πως 3.134 (2880+254) άτομα δεν εμφάνισαν την στεφανιαία νόσο και 8 (5+ 3) εμφάνισαν την στεφανιαία νόσο.

Επίσης, 2.880 άτομα όντως δεν εμφάνισαν κάποια στεφανιαία νόσο και 3 εμφάνισαν κάποια στεφανιαία νόσο.

➤ **Πίνακας συντελεστών παλινδρόμησης του μοντέλου**

Το βέλτιστο μοντέλο είναι :

$$\log\left(\frac{p}{1-p}\right) = -11,920 + 0,64*\text{Ηλικία} + 0,026*\text{Βάρος} + 0,011*\text{Χοληστερίνη} + 0,018*\text{ΣΑΠ} - 0,598*\text{Κάπνισμα} + 0,686*\text{Τύπος προσωπικότητας}$$

| | | Variables in the Equation | | | | | 95% C.I. for EXP(B) | | |
|----------------|-------------------------------|---------------------------|------|--------|----|------|---------------------|-------|-------|
| Step | | B | S.E. | Wald | df | Sig. | Exp(B) | Lower | Upper |
| 1 ^a | Ηλικία | ,064 | ,012 | 27,676 | 1 | ,000 | 1,066 | 1,041 | 1,091 |
| | ΔΜΣ | -,024 | ,046 | ,281 | 1 | ,596 | ,976 | ,892 | 1,068 |
| | Βάρος | ,026 | ,012 | 4,742 | 1 | ,029 | 1,027 | 1,003 | 1,051 |
| | Χοληστερίνη | ,011 | ,002 | 52,878 | 1 | ,000 | 1,011 | 1,008 | 1,014 |
| | Διαστολική Αρτηριακή Πίεση | -,001 | ,011 | ,016 | 1 | ,898 | ,999 | ,978 | 1,020 |
| | Συστολική Αρτηριακή Πίεση | ,018 | ,006 | 8,285 | 1 | ,004 | 1,019 | 1,006 | 1,031 |
| | Κάπνισμα(1) | -,598 | ,142 | 17,772 | 1 | ,000 | ,550 | ,416 | ,726 |
| | Τύπος Προσωπικότητας(1) | ,686 | ,145 | 22,507 | 1 | ,000 | 1,986 | 1,496 | 2,637 |
| | Constant | -11,920 | ,977 | 148,88 | 1 | ,000 | ,000 | | |

9

a. Variable(s) entered on step 1: Ηλικία, ΔΜΣ, Βάρος, Χοληστερίνη, Διαστολική Αρτηριακή Πίεση, Συστολική Αρτηριακή Πίεση, Κάπνισμα, Τύπος Προσωπικότητας.

Πίνακας 5. 10: Πίνακας των συντελεστών του μοντέλου

Από τον παραπάνω πίνακα βλέπουμε ότι από τον έλεγχο Wald η μεταβλητή “ΔΜΣ” (p-value = 0,596) και η μεταβλητή “ΔΑΠ” (p-value = 0,55) είναι στατιστικά μη σημαντικές για την πρόβλεψη την πιθανότητας εμφάνισης στεφανιαίας νόσου καθώς το p-value > 0,05.

Ερμηνεία των συντελεστών του μοντέλου (Πίνακας 5.11)

- Αν αυξηθεί κατά μία μονάδα η ηλικία του ατόμου τότε ο σχετικός λόγος θετικού αποτελέσματος (Στεφανιαία Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.064} = 1,066 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης στεφανιαίας νόσου αυξάνεται κατά 6,6%.
- Αν αυξηθεί κατά μία μονάδα το βάρος ενός ατόμου τότε ο σχετικός λόγος θετικού αποτελέσματος (Στεφανιαία Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.026} = 1,027 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης στεφανιαίας νόσου αυξάνεται κατά 5,1%.
- Αν αυξηθεί κατά μία μονάδα η τιμή της χοληστερίνης ενός ατόμου τότε ο σχετικός λόγος θετικού αποτελέσματος (Στεφανιαία Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.011} = 1,011 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης στεφανιαίας νόσου αυξάνεται κατά 1,1%

- Αν αυξηθεί κατά μία μονάδα η τιμή της $\Sigma\text{ΑΠ}$ ενός ατόμου τότε ο σχετικός λόγος θετικού αποτελέσματος (Στεφανιαία Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.018} = 1,019 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης στεφανιαίας νόσου αυξάνεται κατά 1,9%.
- Ο σχετικός λόγος πιθανοτήτων θετικού αποτελέσματος (Στεφανιαία Νόσος = 1) ενός ατόμου που δεν καπνίζει (Κάπνισμα = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{-0.598} = 0,55 < 1$. Δηλαδή η σχετική πιθανότητα εμφάνισης στεφανιαίας νόσου ενός ατόμου που δεν καπνίζει μειώνεται κατά 45% σε σχέση με ένα άτομο που καπνίζει.
- Ο σχετικός λόγος πιθανοτήτων θετικού αποτελέσματος (Στεφανιαία Νόσος = 1) ενός ατόμου με προσωπικότητα Τύπου Α (Τύπος Προσωπικότητας = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.686} = 1,986 > 1$. Δηλαδή η σχετική πιθανότητα εμφάνισης στεφανιαίας νόσου είναι διπλάσια σε ένα άτομο με προσωπικότητα τύπου Α αυξάνεται κατά 98,6% σε σχέση με ένα άτομο με προσωπικότητα τύπου Β.

5.3 Λογιστική παλινδρόμηση με εξαρτημένη μεταβλητή την εμφάνιση καρδιαγγειακής νόσου μόνο για τους άνδρες της 2^{ης} βάσης δεδομένων

➤ Πίνακας Case processing Summary

Στον παρακάτω πίνακα παρατηρούμε ότι από τους 22.647 άνδρες που υπάρχουν στο σύνολο δεδομένων έχουν χρησιμοποιηθεί και οι 22.647 καθώς δεν υπάρχουν ελλείπουσες τιμές (Missing Cases).

| Case Processing Summary | | | |
|-------------------------------|----------------------|-------|---------|
| Unweighted Cases ^a | | N | Percent |
| Selected Cases | Included in Analysis | 22647 | 100,0 |
| | Missing Cases | 0 | ,0 |
| | Total | 22647 | 100,0 |
| Unselected Cases | | 0 | ,0 |
| Total | | 22647 | 100,0 |

Πίνακας 5. 12: Πίνακας περιγραφής του δείγματος

➤ Έλεγχος Omnibus

Ο έλεγχος Omnibus εξετάζει την σημαντικότητα του μοντέλου σε σχέση με το null μοντέλο (είναι το μοντέλο το οποίο περιέχει μόνο τον σταθερό όρο). Οι υποθέσεις του ελέγχου είναι οι εξής:

H_0 : το null μοντέλο ξεπερνά το υπό εξέταση μοντέλο

H_1 : το υπό εξέταση μοντέλο ξεπερνά το null μοντέλο

Omnibus Tests of Model Coefficients

| | | Chi-square | df | Sig. |
|--------|-------|------------|----|------|
| Step 1 | Step | 5954,031 | 12 | ,000 |
| | Block | 5954,031 | 12 | ,000 |
| | Model | 5954,031 | 12 | ,000 |

Πίνακας 5. 13: Πίνακας ελέγχου Omnibus

Από τον παραπάνω πίνακα παρατηρούμε ότι το p-value είναι σχεδόν μηδέν, επομένως απορρίπτουμε την μηδενική υπόθεση. Δηλαδή το μοντέλο που προσαρμόσαμε είναι καλύτερο από το μοντέλο που περιέχει μόνο την σταθερά (null).

➤ Πίνακας προσαρμογής του μοντέλου

Model Summary

| Step | -2 Log likelihood | Cox & Snell R Square | Nagelkerke R Square |
|------|------------------------|----------------------|---------------------|
| 1 | 25434,626 ^a | ,231 | ,308 |

Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

Πίνακας 5. 14: Πίνακας ποσοστού προσαρμογής του μοντέλου

Όπως παρατηρούμε στον παραπάνω πίνακα το μοντέλο πολλαπλής παλινδρόμησης που προσαρμόσαμε το 30,8% της μεταβλητότητας της εξαρτημένης μεταβλητής ερμηνεύεται από τις επεξηγηματικές μεταβλητές

➤ **Ακρίβεια πρόβλεψης του μοντέλου**

Classification Table^a

| Observed | | Predicted | | Percentage Correct |
|--------------------|------------------------|---------------------------|------|--------------------|
| | | Καρδιαγγειακή νόσος no | yes | |
| Step 1 | Καρδιαγγειακή νόσος no | 9213 | 2306 | 80,0 |
| | yes | 3803 | 7325 | 65,8 |
| Overall Percentage | | | | 73,0 |

Πίνακας 5.15: Πίνακας ακρίβειας πρόβλεψης του μοντέλου

Από τον παραπάνω πίνακα μπορούμε να συμπεράνουμε ότι το ποσοστό ορθής ταξινόμησης του μοντέλου είναι 73%.

Παρατηρούμε ότι το μοντέλο προβλέπει πως 13.016(9.213+3.803) άνδρες δεν εμφάνισαν την καρδιαγγειακή νόσο και 9.631(2.306+7.325) εμφάνισαν την καρδιαγγειακή νόσο.

Επίσης, 9.213 άνδρες όντως δεν εμφάνισαν κάποια καρδιαγγειακή νόσο και 7.325 εμφάνισαν κάποια καρδιαγγειακή νόσο.

➤ **Πίνακας συντελεστών παλινδρόμησης του μοντέλου**

Το βέλτιστο μοντέλο είναι :

$$\log\left(\frac{p}{1-p}\right) = -12,031 + 0,045 * \text{Ηλικία} + 0,006 * \text{Βάρος} + 0,031 * \Delta\text{Μ}\Sigma + 0,056 * \Sigma\text{ΑΠ} +$$

$$0,017 * \Delta\text{ΑΠ} - 1,043 * \text{Χοληστερόλη}(1) - 0,61 * \text{Χοληστερόλη}(2) + 0,368 * \text{Γλυκόζη}(1) + 0,43 * \text{Γλυκόζη}(2) + 0,152 * \text{Κάπνισμα} + 0,23 * \text{Αλκοόλ} + 0,297 * \text{Καθιστική Ζωή}$$

| | | Variables in the Equation | | | | | |
|---------------------|--------------------------|---------------------------|------|----------|----|------|--------|
| | | B | S.E. | Wald | df | Sig. | Exp(B) |
| Step 1 ^a | Ηλικία | ,045 | ,002 | 399,870 | 1 | ,000 | 1,046 |
| | Βάρος | ,006 | ,003 | 5,281 | 1 | ,022 | 1,006 |
| | ΔΜΣ | ,031 | ,008 | 13,963 | 1 | ,000 | 1,032 |
| | ΣΑΠ | ,056 | ,002 | 1341,210 | 1 | ,000 | 1,057 |
| | ΔΑΠ | ,017 | ,002 | 55,231 | 1 | ,000 | 1,017 |
| | Χοληστερόλη | | | 307,168 | 2 | ,000 | |
| | Χοληστερόλη(1) | -1,043 | ,065 | 259,465 | 1 | ,000 | ,352 |
| | Χοληστερόλη(2) | -,610 | ,077 | 62,812 | 1 | ,000 | ,543 |
| | Γλυκόζη | | | 28,442 | 2 | ,000 | |
| | Γλυκόζη(1) | ,368 | ,071 | 26,652 | 1 | ,000 | 1,444 |
| | Γλυκόζη(2) | ,430 | ,093 | 21,246 | 1 | ,000 | 1,537 |
| | Κάπνισμα(1) | ,152 | ,040 | 14,245 | 1 | ,000 | 1,164 |
| | Αλκοόλ(1) | ,230 | ,055 | 17,610 | 1 | ,000 | 1,259 |
| | Φυσική Δραστηριότητα (1) | ,297 | ,038 | 59,436 | 1 | ,000 | 1,345 |
| | Constant | -12,031 | ,248 | 2344,901 | 1 | ,000 | ,000 |

a. Variable(s) entered on step 1: Ηλικία, Βάρος, ΔΜΣ, ΣΑΠ, ΔΑΠ, Χοληστερόλη, Γλυκόζη, Κάπνισμα, Αλκοόλ, Φυσική Δραστηριότητα.

Πίνακας 5.16: Πίνακας των συντελεστών του μοντέλου

Από τον παραπάνω πίνακα βλέπουμε ότι από τον έλεγχο Wald όλες οι μεταβλητές είναι στατιστικά σημαντικές για την πρόβλεψη την πιθανότητας εμφάνισης κάποια καρδιαγγειακής νόσου στους άνδρες του συνόλου δεδομένων καθώς τα p-value < 0,05.

Ερμηνεία των συντελεστών του μοντέλου (Πίνακας 5.16)

- Αν αυξηθεί κατά μία μονάδα η ηλικία ενός άνδρα τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.046} = 1,046 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 4,6%.
- Αν αυξηθεί κατά μία μονάδα το βάρος ενός άνδρα τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.006} = 1,006 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 0,6%.

- Αν αυξηθεί κατά μία μονάδα ο ΔΜΣ ενός άνδρα τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.032} = 1,032 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται 3,2%.
- Αν αυξηθεί κατά μία μονάδα η ΣΑΠ ενός άνδρα τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.056} = 1,057 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 5,7%.
- Αν αυξηθεί κατά μία μονάδα η ΔΑΠ ενός άνδρα τότε ο σχετικός λόγος θετικού αποτελέσματος (Καρδιαγγειακή Νόσος = 1) πολλαπλασιάζεται με τον εκθετικό συντελεστή $e^{0.017} = 1,017 > 1$. Δηλαδή, η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 1,7%.
- Ο σχετικός λόγος πιθανοτήτων των ανδρών με τιμή χοληστερόλης φυσιολογική είναι 0,352 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου είναι μειωμένη κατά $(0,352-1)$ 64,8% σε σχέση με έναν άνδρα που εμφανίζει τιμή χοληστερόλης πολύ πάνω από το φυσιολογικό.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων με τιμή χοληστερόλης πάνω από το φυσιολογικό είναι 0,543 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου είναι μειωμένη κατά 45,7% σε σχέση με ένα άτομο που εμφανίζει τιμή χοληστερόλης πολύ πάνω από το φυσιολογικό.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων με φυσιολογική τιμή γλυκόζης είναι 1,444 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου είναι 44.4% φορές πιο πιθανή σε σχέση με ένα άτομο που εμφανίζει τιμή γλυκόζης πολύ πάνω από το φυσιολογικό.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων με τιμή γλυκόζης πάνω από το φυσιολογικό είναι 1,537 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου είναι 53,7% πιο πιθανή σε σχέση με ένα άτομο που εμφανίζει τιμή γλυκόζης πολύ πάνω από το φυσιολογικό.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων που καπνίζουν είναι 1,164 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 16,4% σε σχέση με κάποιον που δεν καπνίζει.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων που καταναλώνουν αλκοόλ είναι 1,259 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 25,9% σε σχέση με κάποιον που δεν καταναλώνει.
- Ο σχετικός λόγος πιθανοτήτων των ατόμων που έχουν Φυσική Δραστηριότητα είναι 1,345 δηλαδή η σχετική πιθανότητα εμφάνισης καρδιαγγειακής νόσου αυξάνεται κατά 34,5% σε σχέση με κάποιον που δεν ασκείται.

6^ο ΚΕΦΑΛΑΙΟ

ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 Εισαγωγή

Η παρούσα εργασία έχει ως στόχο να μελετήσει τους παράγοντες που σχετίζονται με την εμφάνιση των διάφορων καρδιαγγειακών νοσημάτων, τα οποία αποτελούν μία από τις κυριότερες αίτιες θανάτου στον κόσμο και για τα οποία έχουν γίνει παρόμοιες μελέτες, Δες για παράδειγμα Ruan, Y. , Yanfei G. , et al., (2018), Panagiotakos D., Fitzgerald A., et al. , (2007), Panagiotakos, D., Pitsavos, C., et al.,(2002), Panagiotakos D. and Toutouzas P., (2008), Pitsavos C., Panagiotakos D., et al. ,(2003). Η μελέτη πραγματοποιήθηκε με βάση τα σετ δεδομένων δύο ερευνών, όπου στην πρώτη περιέχονται τόσο άνδρες αλλά και γυναίκες ενώ στη δεύτερη παρουσιάζονται στοιχεία που αφορούν αποκλειστικά άνδρες.

6.2 Συμπεράσματα από Περιγραφική Ανάλυση

Σύμφωνα με το κεφάλαιο 2 που προηγήθηκε έχουν προκύψει συνοπτικά τα παρακάτω αποτελέσματα για τις 2 βάσεις δεδομένων. Στην πρώτη βάση που αφορά άνδρες και γυναίκες σε ποσοστό 34,5% και 65,5% αντίστοιχα βρέθηκε ότι ο μέσος όρος ηλικίας του δείγματος είναι τα 53 έτη ενώ αντίστοιχα στη βάση που αφορά αποκλειστικά άνδρες η μέση ηλικία είναι τα 46 έτη. Το μέσο βάρος για τη βάση που αφορά άνδρες και γυναίκες είναι 72,25kg ενώ για τη βάση που αφορά μόνο άνδρες είναι 77,1kg. Αντίστοιχα η μέση τιμή για τον Δείκτη Μάζας Σώματος (ΔΜΣ) είναι 26,81 στην πρώτη βάση ενώ στη δεύτερη 24,51. Παρατηρούμε ότι η μέση τιμή του ΔΜΣ για την βάση που αφορά άνδρες και γυναίκες υποδεικνύει ότι κατά μέσο όρο τα άτομα ανήκουν στην κατηγορία των Υπέρβαρων και αντίστοιχα στη βάση που αφορά μόνο άνδρες περίπου το 39% είναι υπέρβαροι ή παχύσαρκοι. Επίσης, είδαμε ότι στον πληθυσμό που αφορά και τα δύο φύλα το 91,4% δεν καπνίζει και το 94 % δεν καταναλώνει αλκοόλ ενώ στο δείγμα που περιέχει μόνο άνδρες το 47,6% καπνίζουν με μέσο όρο 24 τσιγάρα ημερησίως. Σχετικά με την αρτηριακή πίεση για την οποία τα φυσιολογικά όρια είναι όταν η τιμή της συστολικής είναι μικρότερη από 120 και της διαστολικής μικρότερη από 80, παρατηρήσαμε στη βάση με τα δύο φύλα η μέση τιμή της συστολικής είναι 126,43 και η μέση τιμή της διαστολικής 81,34 και αντίστοιχα στη βάση με τους άνδρες 128,63 και 82,02. Επιπλέον, για την βάση των δύο φύλων παρατηρήθηκε ότι το 75,7% των ατόμων έχει φυσιολογική την τιμή της χοληστερόλης και το 85,7% διατηρεί φυσιολογική την τιμή της γλυκόζης στο αίμα, ενώ για τη βάση των ανδρών

διαπιστώθηκε η μέση τιμή χοληστερόλης 226,37 mg/dL με τη φυσιολογική τιμή να ορίζεται μικρότερη του 190. Τέλος, από το σύνολο των ατόμων στην 1^η βάση δεδομένων το 49,9% εμφάνισε καρδιαγγειακή νόσο ενώ στο δείγμα των ανδρών το 91,9% δεν εμφάνισε κάποια στεφανιαία νόσο.

6.3 Συμπεράσματα των Συσχετίσεων

Σύμφωνα με τη μελέτη που πραγματοποιήσαμε η εμφάνιση καρδιαγγειακής νόσου για την 1^η βάση δεδομένων που αφορά και τα δύο φύλα φαίνεται να μην επηρεάζεται από το φύλο του ατόμου αλλά ούτε και από την κατανάλωση αλκοόλ. Αντίθετα οι παράγοντες που επηρεάζουν και τις δύο βάσεις είναι η ηλικία των ατόμων, το κάπνισμα, το βάρος καθώς και ο ΔΜΣ αλλά και η κλινική εικόνα των ατόμων που αφορά την αρτηριακή τους πίεση και τις τιμές χοληστερόλης και γλυκόζης. Τέλος σημαντικό είναι και ο παράγοντας του τύπου προσωπικότητας που έχει ένας άνδρας στη 2^η βάση δεδομένων.

6.4 Συμπεράσματα Λογιστικής Παλινδρόμησης

- Προσαρμογή μοντέλου πολλαπλής παλινδρόμησης για την 1^η βάση των δύο φύλων

Σύμφωνα με τα αποτελέσματα του μοντέλου μας μέσω του ελέγχου Wald όλες οι μεταβλητές είναι στατιστικά σημαντικές για την πρόβλεψη της πιθανότητας εμφάνισης καρδιαγγειακής νόσου. Δηλαδή, στον πληθυσμό που περιέχει τόσο άνδρες όσο και γυναίκες προκειμένου να προβλέψουμε την ύπαρξη καρδιαγγειακής νόσου απαιτούνται οι εξής επεξηγηματικές μεταβλητές:

- ❖ Ηλικία
- ❖ Βάρος
- ❖ Δείκτης Μάζας Σώματος
- ❖ Συστολική Αρτηριακή Πίεση
- ❖ Διαστολική Αρτηριακή Πίεση
- ❖ Χοληστερόλη
- ❖ Γλυκόζη
- ❖ Κάπνισμα
- ❖ Κατανάλωση Αλκοόλ
- ❖ Καθιστική Ζωή

- Προσαρμογή μοντέλου πολλαπλής παλινδρόμησης για την 1^η βάση των δύο φύλων απομονώνοντας μόνο τον πληθυσμό των ανδρών

Διαπιστώσαμε ότι στο σύνολο των ανδρών για την 1^η βάση δεδομένων που περιείχε και τα δύο φύλα η μεταβλητή που δεν είναι στατιστικά σημαντική για την πρόβλεψη της καρδιαγγειακής νόσου είναι το κάπνισμα. Σύμφωνα με τον έλεγχο του Wald στους άνδρες αυτού του συνόλου είναι στατιστικά σημαντικοί οι παράγοντες της ηλικίας, του βάρους, του ΔΜΣ, της αρτηριακής πίεσης, της τιμής γλυκόζης και χοληστερόλης καθώς και η κατανάλωση αλκοόλ.

- ❖ Ηλικία
- ❖ Βάρος
- ❖ Δείκτης Μάζας Σώματος
- ❖ Συστολική Αρτηριακή Πίεση
- ❖ Διαστολική Αρτηριακή Πίεση
- ❖ Χοληστερόλη
- ❖ Γλυκόζη
- ❖ Κάπνισμα
- ❖ Κατανάλωση Αλκοόλ
- ❖ Καθιστική Ζωή

- Προσαρμογή μοντέλου πολλαπλής παλινδρόμησης για την 2^η βάση που αφορά αποκλειστικά άνδρες

Για την πολλαπλή γραμμική παλινδρόμηση στη βάση που αφορά μόνο άνδρες προέκυψε από τον έλεγχο του Wald ότι ο ΔΜΣ και η τιμή της Διαστολικής Αρτηριακής Πίεσης είναι στατιστικά μη σημαντικοί παράγοντες και την πρόβλεψη της στεφανιαίας νόσου. Ωστόσο, και σε αυτό το δείγμα συμπεράναμε ότι είναι στατιστικά σημαντικοί παράγοντες η ηλικία, το βάρος, τιμή της χοληστερίνης και το κάπνισμα. Επίσης, προέκυψε στατιστικά σημαντικός ο παράγοντας της προσωπικότητας που έχει ένας άνδρας, συμπεραίνοντας ότι τα άτομα με αυταρχικό χαρακτήρα που είναι ανυπόμονα και συχνά εμφανίζουν συναισθήματα οργής έχουν περισσότερες πιθανότητες εμφάνισης στεφανιαίας νόσου.

- ❖ Ηλικία
- ❖ Βάρος
- ❖ Χοληστερόλη
- ❖ Συστολική Αρτηριακή Πίεση
- ❖ Διαστολική Αρτηριακή Πίεση
- ❖ Κάπνισμα
- ❖ Τύπος Προσωπικότητας

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική Βιβλιογραφία

Βικιπαίδεια: Καρδιαγγειακές Παθήσεις

<https://el.wikipedia.org/wiki/%CE%9A%CE%B1%CF%81%CE%B4%CE%B9%CE%B1%CE%B3%CE%B3%CE%B5%CE%B9%CE%B1%CE%BA%CE%AD%CF%82%CF%80%CE%B1%CE%B8%CE%AE%CF%83%CE%B5%CE%B9%CF%82>

Καλιακούδα Μαριλένα (2018). Παράγοντες κινδύνου που σχετίζονται με τα καρδιαγγειακά νοσήματα, Διπλωματική Εργασία Πανεπιστημίου Πειραιώς, Τμήμα Οικονομικής Επιστήμης.

Σαχλάς Α. και Μπερσίμης Σ. (2017). *Ανάλυση Ιατρικών Δεδομένων*, Πανεπιστημιακές Σημειώσεις

Τζαβελάς Γ. (2018). *Βιοστατιστική και Στατιστικές Μέθοδοι στην Επιδημιολογία*, Πανεπιστημιακές σημειώσεις

Ξενόγλωσση Βιβλιογραφία

Albert, C.M., Chae, C.U., Grodstein, F., Rose, L.M., Rexrode, K.M., Ruskin, J.N., et al., (2003). Prospective study of sudden cardiac death among women in the United States. *Circulation* , **107**(16):2096-101.

[Micha R.](#) & [Mozaffarian D.](#) (2010). Saturated Fat and Cardiometabolic Risk Factors, Coronary Heart Disease, Stroke, and Diabetes: a Fresh Look at the Evidence, *Lipids* **45**: 893-905

Mozaffarian D., Wilson Peter W.F. and Kannel William B. (2008). Lifestyle Risk factors for Cardiovascular Disease, *Circulation*, **117**(23) :3031-3038.

Muacevic A. and Adler John R , (2022). A Systematic Review of Major Cardiovascular Risk Factors: A Growing Global Health Concern, *Cureus*, **14**(100):e30119

Ritchie H., Spooner F. and Roser M. (2018). Causes of Death, *Our World in Data*. Available from: <https://ourworldindata.org/causes-of-death>.

[Ruan Y.](#), [Yanfei G.](#), [ZhengvY](#), [Huang Z.](#), [Shuangyuan S](#), [Kowal P](#), [Yan S](#), and [Wu F](#). (2018). Cardiovascular disease (CVD) and associated risk factors among older adults in six low-and middle-income countries . *BMC Public Health*. **18** : 778

Panagiotakos D., Fitzgerald A., Pitsavos C., Pipilis A., Graham I., Stefanadis Ch. (2007). Statistical Modelling of 10-Year Fatal Cardiovascular Disease Risk in Greece: The Hellenic SCORE (a Calibration of the ESC SCORE Project). *Hellenic J Cardiol* **48**: 55-63, 2007

Panagiotakos, D., Pitsavos, C., Chrisohoou, C., Stefanadis, C., Preve Toutouzas, P. (2002) . Risk Stratification of Coronary Heart Disease in Greece: Final Results from the CARDIO2000 Epidemiological Study, *Preventive Medicine* **35**(6) : 548-556

Panagiotakos D. , Toutouzas P. (2008). Importance of LDL/HDL cholesterol ratio as a predictor for coronary heart disease events in patients with heterozygous familial hypercholesterolaemia: a 15-year follow-up (1987-2002), *Current Medical Research and Opinion* **19**(2): 89-94.

Pitsavos C., Panagiotakos D., Chrysohoou C., Stefanadis C. (2003). Epidemiology of cardiovascular risk factors in Greece: aims, design and baseline characteristics of the ATTICA study, *BMC Public Health* **3**:32