



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΠΛΗΡΟΦΟΡΙΚΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΒΙΟΙΑΤΡΙΚΗ

Στατιστική ανάλυση δεδομένων που αφορούν τη διατροφική συμπεριφορά ενηλίκων

Ελένη Χατζηβασιλείου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Επιβλέπων
Αναπληρωτής Καθηγητής κ. Ιωάννης Τριανταφύλλου

Λαμία, Νοέμβριος 2020



UNIVERSITY OF THESSALY

SCHOOL OF SCIENCE

INFORMATICS AND COMPUTATIONAL BIOMEDICINE

Statistical Analysis of Data on Adult's Diet Behavior

Eleni Chatzivasiliou

Master thesis

Supervisor

Assistant Professor Mr. Ioannis Triantafyllou

Lamia, November 2020



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΔΙΑΤΜΗΜΑΤΙΚΟ ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΠΛΗΡΟΦΟΡΙΚΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΒΙΟΙΑΤΡΙΚΗ
ΚΑΤΕΥΘΥΝΣΗ:**

«ΥΠΟΛΟΓΙΣΤΙΚΗ ΙΑΤΡΙΚΗ ΚΑΙ ΒΙΟΛΟΓΙΑ»

Στατιστική ανάλυση δεδομένων που αφορούν τη διατροφική συμπεριφορά ενηλίκων

Ελένη Χατζηβασιλείου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Επιβλέπων
Αναπληρωτής Καθηγητής κ. Ιωάννης Τριανταφύλλου**

Λαμία, Νοέμβριος 2020

«Υπεύθυνη Δήλωση μη λογοκλοπής και ανάληψης προσωπικής ευθύνης»

Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, και γνωρίζοντας τις συνέπειες της λογοκλοπής, δηλώνω υπεύθυνα και ενυπογράφως ότι η παρούσα εργασία με τίτλο [«τίτλος εργασίας»] αποτελεί προϊόν αυστηρά προσωπικής εργασίας και όλες οι πηγές από τις οποίες χρησιμοποίησα δεδομένα, ιδέες, φράσεις, προτάσεις ή λέξεις, είτε επακριβώς (όπως υπάρχουν στο πρωτότυπο ή μεταφρασμένες) είτε με παράφραση, έχουν δηλωθεί κατάλληλα και ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή. Αναλαμβάνω πλήρως, ατομικά και προσωπικά, όλες τις νομικές και διοικητικές συνέπειες που δύναται να προκύψουν στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής.

Η ΔΗΛΟΥΣΑ

Ημερομηνία

Υπογραφή

Στατιστική ανάλυση δεδομένων που αφορούν τη διατροφική συμπεριφορά ενηλίκων

Ελένη Χατζηβασιλείου

Τριμελής Επιτροπή:

Όνοματεπώνυμο: Αναπληρωτής Καθηγητής κ. Ιωάννης Τριανταφύλλου

Όνοματεπώνυμο:

Όνοματεπώνυμο:

ΕΥΧΑΡΙΣΤΙΕΣ

Στα πλαίσια αυτής της Μεταπτυχιακής Διπλωματικής Εργασίας θα ήθελα να ευχαριστήσω όλους όσους με στήριξαν και με καθοδήγησαν μέχρι και σήμερα καθ' όλη τη διάρκειά της φοίτησης μου στο Διατμηματικό Μεταπτυχιακό Πρόγραμμα Πληροφορική και Υπολογιστική Βιοϊατρική.

Αρχικά θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα καθηγητή και καθοδηγητή της παρούσας εργασίας κ. Τριανταφύλλου Ιωάννη για την απρόσκοπτη επικοινωνία και συμβολή των γνώσεων του αλλά και της στήριξης του καθ' όλη την διάρκεια της διεκπεραίωσης της εργασίας.

Θερμότατες είναι και οι ευχαριστίες μου προς την κα. Τσολάκη Αικατερίνη, την κλινική Διατροφολόγο του Πανεπιστημιακού Γενικού Νοσοκομείου Λάρισας, για την παροχή δεδομένων των ασθενών, καθώς και την απαραίτητη καθοδήγηση για την διαδικασία της παροχής των δεδομένων.

Ακόμη, θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους μου, που μου στάθηκαν σε όλη την διάρκεια των μεταπτυχιακών μου σπουδών μου.

Λαμία, Νοέμβριος 2020

Ελένη Χατζηβασιλείου

ΠΕΡΙΛΗΨΗ

Ο σκοπός της διατριβής ήταν η ανάλυση δεδομένων που αφορούν τη διατροφική συμπεριφορά ενηλίκων. Η παραπάνω ανάλυση διακρίνεται σε 2 στάδια:

- Στο 1^ο στάδιο πραγματοποιήθηκε Περιγραφική Στατιστική Ανάλυση για κάθε δείκτη - μεταβλητή. Αναλυτικότερα, σε κάθε μεταβλητή, μελετήθηκε ο μέσος όρος, διακύμανση κα.
- Στο 2^ο Στάδιο πραγματοποιήθηκε συσχέτιση μεταξύ των μεταβλητών των δεδομένων που είχαμε στη διάθεσή μας και στη συνέχεια έγινε προσπάθεια ανάπτυξης σχετικών μοντέλων πρόβλεψης διαφόρων ασθενειών – παθήσεων των ασθενών. Όσο αφορά τα μοντέλα πρόβλεψης δύναται να εμπλουτιστούν και με άλλες μεταβλητές, προκειμένου να παρέχουν στους ειδικούς ακριβέστερα και πιο ολοκληρωμένα αποτελέσματα. Τα αποτελέσματα αυτά είναι σημαντικά διότι παρέχουν εξοικονόμηση χρόνου και χρήματος

Λέξεις Κλειδιά: Περιγραφική Στατιστική Ανάλυση, Επαγωγική Στατιστική Συμπερασματολογία, συσχέτιση, μοντέλα πρόβλεψης.

Abstract

The aim of the present MSc thesis was the statistical analysis of data on adult's diet behavior. In the present MSc, the statistical package IBM SPSS was used. Via this package Several Descriptive Statistics for each parameter were materialized. Furthermore, through Inferential Statistical Methods, for each illness (parameter) prediction models were constructed. These models are important for the results estimation thanks to the accuracy and speed that they offer.

Key Words: Descriptive Statistics, Inferential Statistics, Prediction models

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1.	ΕΙΣΑΓΩΓΗ – ΠΕΡΙΓΡΑΦΗ ΔΕΔΟΜΕΝΩΝ	1
1.1	ΕΡΕΥΝΗΤΙΚΑ ΕΡΩΤΗΜΑΤΑ.....	1
1.2	ΑΣΘΕΝΕΙΕΣ – ΠΑΘΗΣΕΙΣ	1
1.3	ΔΕΔΟΜΕΝΑ	5
1.4	ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ.....	6
1.5	ΣΤΟΧΟΙ ΤΗΣ ΔΙΑΤΡΙΒΗΣ.....	6
2.	ΕΡΓΑΛΕΙΑ ΣΤΑΤΙΣΤΙΚΗΣ.....	7
2.1	ΟΡΙΣΜΟΣ.....	7
2.2	ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ	7
2.3	ΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ.....	10
3.	ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ.....	16
3.1	ΠΟΣΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ.....	16
3.2	ΠΟΙΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ.....	26
4.	ΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ.....	32
4.1	ΕΛΕΓΧΟΣ ΚΑΝΟΝΙΚΟΤΗΤΑΣ.....	32
4.2	ΑΝΑΛΥΣΗ ΣΥΣΧΕΤΙΣΗΣ.....	32
4.3	ΜΟΝΤΕΛΑ ΠΡΟΒΛΕΨΗΣ	44
4.4	ΒΕΛΤΙΣΤΑ ΜΟΝΤΕΛΑ ΠΡΟΒΛΕΨΗΣ.....	60
5.	ΣΥΜΠΕΡΑΣΜΑΤΑ.....	62
5.1	ΕΛΕΓΧΟΣ ΚΑΝΟΝΙΚΟΤΗΤΑΣ.....	62
5.2	ΑΝΑΛΥΣΗ ΣΥΣΧΕΤΙΣΗΣ.....	62
5.3	ΒΕΛΤΙΣΤΟ ΚΑΙ ΜΗ ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ ΤΗΣ ΜΕΤΑΒΛΗΤΗΣ ΔΜΣ.....	63
5.4	ΜΟΝΤΕΛΑ ΠΡΟΒΛΕΨΗΣ ΜΕ ΠΟΙΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ ΣΤΟΧΟΥΣ...	65

5.5	ΒΑΣΙΚΕΣ ΕΠΙΣΗΜΑΝΣΕΙΣ – ΣΥΜΒΟΥΛΕΣ.....	66
	ΒΙΒΛΙΟΓΡΑΦΙΑ.....	68
	ΑΛΦΑΒΗΤΙΚΟ ΕΥΡΕΤΗΡΙΟ ΣΗΜΑΝΤΙΚΩΝ ΟΡΩΝ.....	70

1. ΕΙΣΑΓΩΓΗ – ΠΕΡΙΓΡΑΦΗ ΔΕΔΟΜΕΝΩΝ

Σε αυτό το Κεφάλαιο παρουσιάζονται τα ερευνητικά ερωτήματα, τα οποία αναλύονται στα ακόλουθα κεφάλαια και πραγματοποιείται βασική περιγραφή των παθήσεων – ασθενειών των ερωτημάτων αυτών.

1.1 ΕΡΕΥΝΗΤΙΚΑ ΕΡΩΤΗΜΑΤΑ

Τα ερευνητικά ερωτήματα, τα οποία μελετώνται, είναι τα παρακάτω:

- Ο Δείκτης Μάζας Σώματος (ΔΜΣ) και η ηλικία επιδρούν στην εμφάνιση οποιασδήποτε μορφής Καρκίνου;
- Ο ΔΜΣ και η ηλικία επιδρούν στην εμφάνιση Υψηλής Αρτηριακής Πίεσης;
- Ο ΔΜΣ και η ηλικία επιδρούν στην εμφάνιση Χοληστερίνης;
- Ο ΔΜΣ και η ηλικία επιδρούν στην εμφάνιση Σακχάρου;

Σε αυτό το σημείο είναι επιθυμητό να επισημανθεί ότι τα μοντέλα πρόβλεψης που αναπτύχθηκαν για τα παραπάνω ερωτήματα – ασθένειες, περιλαμβάνουν κι άλλες μεταβλητές όπως βάρος, ύψος, φύλο (ουσιαστικά ο ΔΜΣ περιλαμβάνει το βάρος και το ύψος του/της ασθενούς). Επιπλέον εξετάστηκε αν η εμφάνιση κάθε ασθένειας επηρεάζεται από την εμφάνιση ή μη, κάποιας άλλης.

1.2 ΑΣΘΕΝΕΙΕΣ - ΠΑΘΗΣΕΙΣ

Οι παθήσεις οι οποίες θα αναλυθούν είναι ο Σακχαρώδης Διαβήτης, η Υψηλή Αρτηριακή Πίεση, Χοληστερίνη και Καρκίνος. Επιπλέον θα γίνει αναφορά στο ΔΜΣ ως μέρος εξέτασης των δεδομένων μας.

ΚΑΡΚΙΝΟΣ

...«Ο καρκίνος, επίσης γνωστός ως κακοήθης όγκος ή κακοήθες νεόπλασμα, είναι μία ομάδα ασθενειών που έχουν σχέση με την αφύσικη κυτταρική ανάπτυξη με πιθανότητα εισβολής ή διάδοσης σε άλλα μέρη του σώματος. Δεν είναι όλοι οι όγκοι καρκινικοί. Ο καλοήθης όγκος δεν κάνει μεταστάσεις σε άλλα μέλη του σώματος. Υπάρχουν πάνω από 100 διαφορετικά γνωστά είδη καρκίνου που επηρεάζουν τους ανθρώπους.

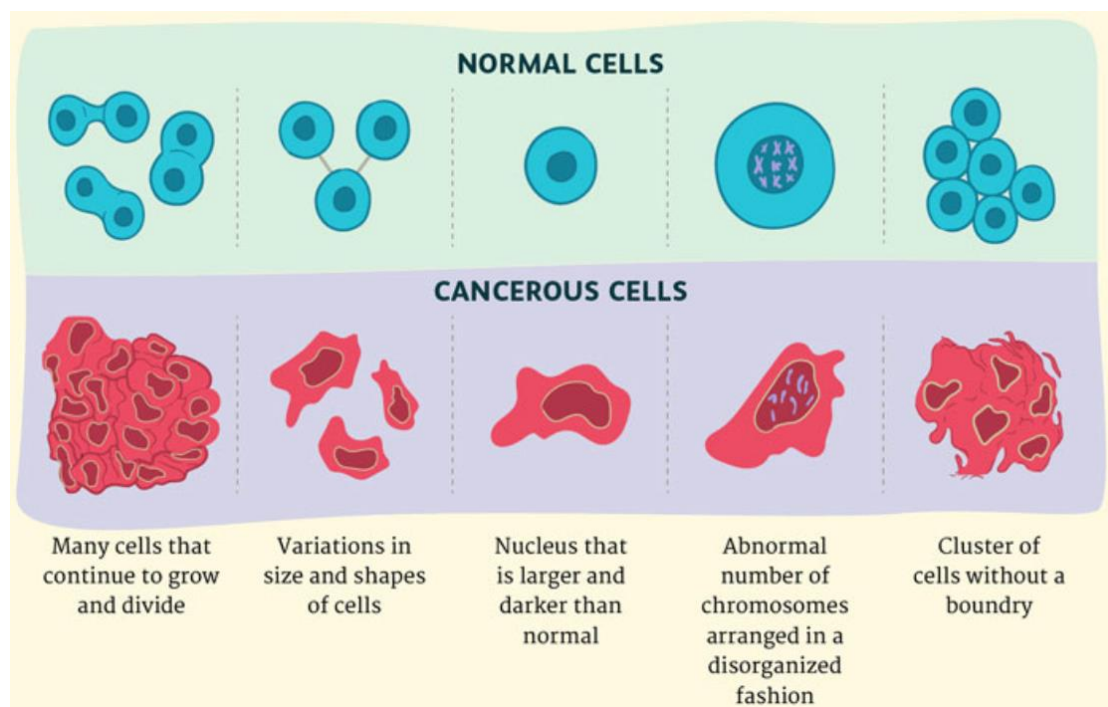
Ο καρκίνος είναι ένα από τα σοβαρότερα προβλήματα υγείας που παρατηρούνται σήμερα στις αναπτυγμένες χώρες. Οι στατιστικές δείχνουν ότι αποτελεί τη δεύτερη πιο συχνή αιτία θανάτου μετά τις καρδιοπάθειες. Συνήθως προσβάλλει ανθρώπους μεγάλης ηλικίας, υπάρχουν όμως και μορφές καρκίνου που εμφανίζονται σε νεαρής ηλικίας άτομα, ακόμη και σε παιδιά.

Ο όρος «καρκίνος» δεν αποδίδεται σε μία και μόνη ασθένεια, αλλά σε μια ομάδα ασθενειών που χαρακτηρίζονται από τον ανεξέλεγκτο πολλαπλασιασμό των κυττάρων. Σε αντίθεση με τα φυσιολογικά κύτταρα στο σώμα μας, τα οποία αυξάνονται, διαιρούνται και πεθαίνουν με έναν αυστηρά ελεγχόμενο τρόπο, τα καρκινικά κύτταρα διαφέρουν διότι συνεχίζουν να διαιρούνται ανεξέλεγκτα. Αυτό έχει ως αποτέλεσμα την ανάπτυξη μιας μάζας κυττάρων, που ονομάζεται όγκος. Οι όγκοι μπορεί να είναι καλοήθεις ή κακοήθεις.

Γενικά, τα χαρακτηριστικά των καρκινικών κυττάρων είναι ότι:

- Ανθίστανται στην απόπτωση.

- Πολλαπλασιάζονται παρουσία ή μη αυξητικών παραγόντων.
- Ανθίστανται σε σήματα που σταματούν τον κυτταρικό πολλαπλασιασμό.
- Ανθίστανται στο μηχανισμό γήρανσης.
- Κάνουν μεταστάσεις.
- Δημιουργούν αγγεία για την αιμάτωσή τους.»...
(www.el.wikipedia.org)



Εικόνα 1: Διαφορά υγιών και καρκινικών κυττάρων
(πηγή: www.iubas.gr/pathiseis/karkinos-neoplasia/)

ΑΡΤΗΡΙΑΚΗ ΠΙΕΣΗ

...«Η υπέρταση (HTN) ή υψηλή αρτηριακή πίεση, μερικές φορές αρτηριακή υπέρταση, είναι μία χρόνια πάθηση κατά την οποία η αρτηριακή πίεση στις αρτηρίες είναι αυξημένη. Αυτή η αύξηση της πίεσης κάνει την καρδιά να εργάζεται πιο εντατικά από το φυσιολογικό για να κυκλοφορεί το αίμα μέσω των αιμοφόρων αγγείων. Η αρτηριακή πίεση περιλαμβάνει δύο μετρήσεις, τη συστολική και τη διαστολική, που εξαρτώνται από το εάν ο καρδιακός μυς συστέλλεται (συστολή) ή χαλαρώνει μεταξύ των παλμών (διαστολή). Η φυσιολογική αρτηριακή πίεση σε κατάσταση ηρεμίας κυμαίνεται από 100 έως 140 mmHg συστολική (ανώτατη μέτρηση) και από 60 έως 90 mmHg διαστολική (κατώτατη μέτρηση). Υπάρχει υψηλή αρτηριακή πίεση εάν είναι μόνιμως σε επίπεδα 140/90 mmHg ή παραπάνω.

Η υπέρταση ταξινομείται είτε ως πρωτογενής (ουσιαστική) υπέρταση είτε ως δευτεροπαθής υπέρταση. Περίπου το 90–95% των περιπτώσεων κατηγοριοποιούνται ως "πρωτογενής υπέρταση", που σημαίνει υψηλή αρτηριακή πίεση χωρίς προφανές υποκείμενο ιατρικό αίτιο. Άλλες καταστάσεις που επηρεάζουν τα νεφρά, τις αρτηρίες, την καρδιά ή το ενδοκρινικό σύστημα προκαλούν το υπόλοιπο 5–10% των περιπτώσεων (δευτεροπαθής υπέρταση).

Η υπέρταση είναι ένας σημαντικός παράγοντας κινδύνου για το εγκεφαλικό, το έμφραγμα του μυοκαρδίου, την καρδιακή ανεπάρκεια, τα ανευρύσματα των αρτηριών, την περιφερική αρτηριακή νόσο και είναι η αιτία της χρόνιας νεφρικής νόσου. Ακόμα και η μέτρια αύξηση της αρτηριακής πίεσης σχετίζεται με το μειωμένο προσδόκιμο ζωής. Οι αλλαγές στον τρόπο διατροφής και ζωής μπορούν να βελτιώσουν τον έλεγχο της αρτηριακής πίεσης και να μειώσουν τον κίνδυνο των επιπλοκών που σχετίζονται με την υγεία.»...

(www.el.wikipedia.org)

BLOOD PRESSURE CATEGORY	SYSTOLIC mm Hg (upper number)		DIASTOLIC mm Hg (lower number)
NORMAL	LESS THAN 120	and	LESS THAN 80
ELEVATED	120 – 129	and	LESS THAN 80
HIGH BLOOD PRESSURE (HYPERTENSION) STAGE 1	130 – 139	or	80 – 89
HIGH BLOOD PRESSURE (HYPERTENSION) STAGE 2	140 OR HIGHER	or	90 OR HIGHER
HYPERTENSIVE CRISIS (consult your doctor immediately)	HIGHER THAN 180	and/or	HIGHER THAN 120

Εικόνα 2: Στάδια και τιμές Αρτηριακής Πίεσης

(πηγή: www.heart.org)

ΧΟΛΗΣΤΕΡΙΝΗ

...«Η χοληστερίνη είναι μια κηρώδης οργανική ουσία που ανήκει στην κατηγορία των λιπιδίων. Είναι απαραίτητη για τη σωστή λειτουργία του οργανισμού μας. Είναι χρήσιμη επειδή:

- Είναι βασικό συστατικό των κυτταρικών μεμβρανών και είναι απαραίτητη για να ρυθμιστεί σωστά η διαπερατότητα της μεμβράνης.
- Είναι δομικό συστατικό αρκετών ορμονών (ανδρογόνα, οιστρογόνα, κορτιζόλη, κορτικοστερόνη, αλδοστερόνη κλπ).
- Συμβάλλει στην παραγωγή της βιταμίνης D.
- Συμβάλλει στην παραγωγή χολικών οξέων, που βοηθούν στην πέψη των λιπών.
- Συμβάλλει στον μεταβολισμό των λιποδιαλυτών βιταμινών

Η χοληστερίνη υπάρχει στο αίμα σε σύμπλοκα, τις λιποπρωτεΐνες. Υπάρχουν διάφορα είδη χοληστερίνης, που δεν είναι όλα επιβλαβή. Τα βασικά είναι:

Η LDL χοληστερίνη, που είναι βλαβερή, μεταφέρει τη χοληστερίνη από το ήπαρ στα κύτταρα. Όταν είναι αυξημένη είναι πιθανό να προκαλέσει απόφραξη των αγγείων, αυξάνοντας τον κίνδυνο για καρδιαγγειακές παθήσεις.

Η HDL χοληστερίνη είναι ωφέλιμη, αφού απομακρύνει την LDL χοληστερίνη από τα τοιχώματα των αγγείων και τη μεταφέρει στο συκώτι από όπου αποβάλλεται με τη χολή. Έτσι, ελαττώνεται ο κίνδυνος για εμφάνιση καρδιαγγειακών παθήσεων. Άτομα με υψηλά επίπεδα HDL χοληστερίνης έχουν χαμηλό κίνδυνο να πάθουν έμφραγμα και εγκεφαλικό επεισόδιο.»... (www.mikroviologos.gr)

Φυσιολογικές τιμές χοληστερίνης

Ενήλικες		
Ολική χοληστερίνη	LDL-χοληστερίνη	HDL-χοληστερίνη
<170 Επιθυμητή	<100 Ιδανική	>=60 Επιθυμητή
	100-129 Σχεδόν φυσιολογική	<40 Χαμηλή
	130-159 Οριακά υψηλή	
	160-189 Υψηλή	
	>=190 Πολύ υψηλή	

Παιδιά		
Ολική χοληστερίνη	LDL-χοληστερίνη	HDL-χοληστερίνη
<170 Επιθυμητή	<110 Επιθυμητή	>40 Επιθυμητή
170-200 Οριακά υψηλή	110-130 Οριακά υψηλή	
>200 Υψηλή	>130 Υψηλή	<35 Χαμηλή

Εικόνα 3: Φυσιολογικές τιμές Αρτηριακής Πίεσης

(πηγή:www.mikroviologos.gr)

ΣΑΚΧΑΡΩΔΗΣ ΔΙΑΒΗΤΗΣ

...«Ο σακχαρώδης διαβήτης είναι μια νόσος που χαρακτηρίζεται από αύξηση του σακχάρου (γλυκόζης) στο αίμα, που κατά κύριο λόγο οφείλεται στην ελλιπή παραγωγή ινσουλίνης. Η ινσουλίνη είναι μια ορμόνη που εκκρίνεται από το πάγκρεας. Ένας από τους κύριους ρόλους της είναι να βοηθάει το σάκχαρο του αίματος να εισέρχεται στα κύτταρα ώστε να χρησιμοποιηθεί ως “καύσιμο” για την παραγωγή ή την αποθήκευση ενέργειας. Υπάρχουν δυο βασικοί τύποι διαβήτη:

- Ο διαβήτης Τύπου 1 εμφανίζεται συνήθως σε νεαρά άτομα (αν και μπορεί να εμφανιστεί σε οποιαδήποτε ηλικία). Χαρακτηρίζεται από καταστροφή των β-κυττάρων του παγκρέατος που παράγουν την ινσουλίνη, με αποτέλεσμα να είναι αναγκαία η χορήγηση ινσουλίνης με ενέσεις για να μπορεί το άτομο να επιβιώσει.
- Ο διαβήτης Τύπου 2 αφορά κυρίως άτομα μεγαλύτερης ηλικίας, συνήθως παχύσαρκα και συχνά υπάρχει κληρονομική προδιάθεση για την εμφάνισή του. Για τη θεραπεία του απαιτείται απώλεια βάρους (με σωστή διατροφή και άσκηση) και χορήγηση αρχικά δισκίων και σε προχωρημένα στάδια χορήγηση ινσουλίνης. Ο διαβήτης της κύησης είναι ουσιαστικά μια ειδική περίπτωση διαβήτη που παρουσιάζεται μόνο κατά τη διάρκεια της κύησης και κατά κανόνα υποχωρεί αμέσως μετά τον τοκετό. Αντιμετωπίζεται με δίαιτα και αν είναι απαραίτητο με

ινσουλίνη, ενώ οι γυναίκες που παρουσιάζουν διαβήτη κήσης έχουν μεγαλύτερη πιθανότητα να εμφανίσουν τύπου 2 διαβήτη σε μεγαλύτερη ηλικία.»...
(www.novartis.gr)

Blood Test Levels for Diagnosis of Diabetes and Prediabetes

	A1C (percent)	Fasting Plasma Glucose (mg/dL)	Oral Glucose Tolerance Test (mg/dL)
Diabetes	6.5 or above	126 or above	200 or above
Prediabetes	5.7 to 6.4	100 to 125	140 to 199
Normal	About 5	99 or below	139 or below

Definitions: mg = milligram, dL = deciliter
For all three tests, within the prediabetes range, the higher the test result, the greater the risk of diabetes.

Εικόνα 4: Στάδια και τιμές Αρτηριακής Πίεσης
(πηγή:www.medivizor.com)

1.3 ΔΕΔΟΜΕΝΑ

Σύμφωνα με επιστήμονες υγείας, η διατροφή παίζει σημαντικό ρόλο στην εμφάνιση ή μη, διαφόρων ασθενειών. Έτσι κατόπιν αιτήσεως μου προκειμένου να εκπονήσω την παρούσα Διπλωματική, έλαβα από την κα. Αικατερίνη Τσολάκη, η οποία αποτελεί Κλινική Διαιτολόγος του Πανεπιστημιακού Γενικού Νοσοκομείου Λάρισας, τα δεδομένα που περιγράφονται παρακάτω.

Αναλυτικότερα, τα δεδομένα διακρίνονται σε ποσοτικές και ποιοτικές μεταβλητές.

ΠΟΣΟΤΙΚΕΣ

Οι μεταβλητές ΒΑΡΟΣ, ΗΛΙΚΙΑ, ΥΨΟΣ και ΔΜΣ είναι ποσοτικές.

ΠΟΙΟΤΙΚΕΣ

Οι 5 (ΦΥΛΟ, ΚΑΡΚΙΝΟΣ, ΠΙΕΣΗ, ΣΑΚΧΑΡΟ και ΧΟΛΗΣΤΕΡΙΝΗ) αποτελούν ποιοτικές μεταβλητές. Συγκεκριμένα, οι μεταβλητές ΦΥΛΟ, ΚΑΡΚΙΝΟΣ και ΧΟΛΗΣΤΕΡΙΝΗ αποτελούν Δίτιμες ή δυαδικές ή διχοτομικές (binaries) ή διμερείς διότι «παίρνουν» 2 τιμές. Οι υπόλοιπες μεταβλητές ΣΑΚΧΑΡΟ και ΠΙΕΣΗ αποτελούν τακτικές (ordinals) διότι αποτελούνται από τρεις κατηγορίες μεταξύ των οποίων παρατηρείται μορφή ανισότητας.

Κατά την μελέτη, οι μεταβλητές ΔΜΣ, ΚΑΡΚΙΝΟΣ, ΣΑΚΧΑΡΟ, ΠΙΕΣΗ και ΧΟΛΗΣΤΕΡΙΝΗ αποτέλεσαν τις μεταβλητές στόχους. Επιπλέον όταν δεν αποτελούσαν μεταβλητές στόχους,

χρησιμοποιήθηκαν - εισήχθησαν ως ανεξάρτητες, έτσι ώστε να αναπτυχθεί λεπτομερέστερα μοντέλα πρόβλεψης.

1.4 ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

Η παρούσα διπλωματική πραγματοποιήθηκε μέσω του προγράμματος IBM SPSS Statistics.

SPSS

...«Τα αρχικά του σημαίνουν Superior Performance Software System και για πρώτη φορά αναπτύχθηκε το 1965 στο Stanford της Καλιφόρνια. Πρόκειται για ένα στατιστικό πακέτο ανάλυσης δεδομένων που δίνει στο χρήστη τη δυνατότητα δημιουργίας αναφορών, μοντελοποίησης και ανάλυσης δεδομένων αλλά και τη γραφική αναπαράσταση αυτών. Το πακέτο αυτό προσφέρει ένα μεγάλο αριθμό στατιστικών συναρτήσεων μέσα από ένα φιλικό για το χρήστη, γραφικό περιβάλλον.»...

(www.anavasis.gr)

1.5 ΣΤΟΧΟΙ ΤΗΣ ΔΙΑΤΡΙΒΗΣ

Οι στόχοι της διατριβής είναι η πραγματοποίηση:

- Περιγραφική Στατιστική Ανάλυση.
- Επαγωγική Στατιστική Συμπερασματολογία.

Συγκεκριμένα, τελικός στόχος είναι η ανάπτυξη μοντέλων πρόβλεψης για κάθε ασθένεια – πάθηση με σκοπό την υποβοήθηση των ειδικών στην εξαγωγή γρήγορων και αξιόπιστων συμπερασμάτων ή διαγνώσεων.

2. ΕΡΓΑΛΕΙΑ ΣΤΑΤΙΣΤΙΚΗΣ

Στο Κεφάλαιο αυτό, αρχικά θα πραγματοποιηθεί σύντομη ανάλυση της Στατιστικής και στη συνέχεια θα παρουσιασθούν τα εργαλεία εκείνα που χρησιμοποιήθηκαν στη παρούσα Διπλωματική.

2.1 ΟΡΙΣΜΟΣ

Στατιστική ονομάζεται η επιστήμη η οποία βοηθά στην ανάλυση και κατανόηση δεδομένων προερχόμενα είτε από διάφορα πειράματα είτε από παρατήρηση (καταγραφή φαινομένων – γεγονότων).

Η στατιστική έρευνα διακρίνεται σε 3 στάδια τα οποία παρουσιάζονται παρακάτω:

- Σχεδιασμός πειραμάτων (Experimental Design).
- Περιγραφικής Στατιστικής Ανάλυσης (Descriptive Statistics).
- Επαγωγική Στατιστική Συμπερασματολογία (Inferential Statistics).

2.2 ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ

Η Περιγραφική Στατιστική Ανάλυση ασχολείται με τη συνοπτική και αποτελεσματική παρουσίαση των δεδομένων μιας στατιστικής έρευνας υπό μορφή πινάκων και διαγραμμάτων.

Σε αυτό το στάδιο μελετήθηκαν τα παρακάτω:

- Μέτρα Κεντρικής Τάσης ή Θέσης:

- Η μέση τιμή (Mean) εκφράζει την κεντρική τάση ενός δείγματος. Υπολογίζεται από τον τύπο:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Συνάρτηση 1

- Η διάμεσος (Median) περιγράφει κι εκείνη την κεντρική τάση ενός δείγματος. Ο υπολογισμός της εξαρτάται από τον αριθμό των παρατηρήσεων.

Έτσι αν το πλήθος n των παρατηρήσεων είναι περιττό, ισχύει: $\delta = x_{((n+1)/2)}$

Αν το n είναι άρτιο, ισχύει: $\delta = \frac{x_{(n/2)} + x_{(n/2+1)}}{2}$

- Η κορυφή (Mode) ισούται με την τιμή της παρατήρησης του δείγματος που εμφανίζει τη μεγαλύτερη συχνότητα.

- Μέτρα Διασποράς:

- Η διασπορά (Variance) εκφράζει το μέγεθος των αποκλίσεων μεταξύ των παρατηρήσεων ενός δείγματος.

Υπολογίζεται από τον τύπο:

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_v - \bar{x})^2}{v - 1}$$

Συνάρτηση 2

- Η τυπική απόκλιση (Standard Deviation) αποτελεί την τετραγωνική ρίζα της διασποράς και εκφράζει το μέγεθος της απόκλισης της καθεμιάς παρατήρησης του δείγματος από τη μέση τιμή.

Υπολογίζεται από τον τύπο:

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_v - \bar{x})^2}{v - 1}}$$

Συνάρτηση 3

- Το εύρος (Range) αποτελεί την μέγιστη δυνατή απόκλιση μεταξύ των παρατηρήσεων ενός δείγματος.

Υπολογίζεται από τον τύπο: $R = x_{\max} - x_{\min}$

- Το ενδοτεταρτημοριακό εύρος (Interquartile Range) εκφράζει το μέγεθος αποκλίσεων μεταξύ των παρατηρήσεων του δείγματος.

Υπολογίζεται από τον τύπο: $IQR = Q_3 - Q_1$

- Ο συντελεστής μεταβλητότητας (Coefficient of Variability) εκφράζει την ομοιογένεια των παρατηρήσεων ενός δείγματος, όπως παρακάτω:

Αν το αποτέλεσμα είναι $\leq 10\%$, τότε το δείγμα χαρακτηρίζεται ως ομοιογενές.

Αν το αποτέλεσμα είναι $\geq 10\%$, τότε το δείγμα χαρακτηρίζεται ως ανομοιογενές.

Στη περίπτωση σύγκρισης 2 δειγμάτων ως προς την ομοιογένεια, πιο ομοιογενές χαρακτηρίζεται το δείγμα με τη μικρότερη μεταβλητότητα δηλαδή με το μικρότερο CV. Υπολογίζεται από τον τύπο:

$$CV = \frac{s}{\bar{x}} 100\%$$

Συνάρτηση 4

- Μέτρα Ασυμμετρίας

- Ο συντελεστής ασυμμετρίας (Skewness) μας πληροφορεί σχετικά με το βαθμό συμμετρίας των παρατηρήσεων ενός δείγματος γύρω από τη μέση τιμή. Υπολογίζεται από τον τύπο:

$$a_3 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{s^3}$$

Συνάρτηση 5

Ωστόσο, το παραπάνω αποτέλεσμα δεν αρκεί για να καταλήξουμε σε κάποια ερμηνεία. Για να είναι δυνατό αυτό, θα πρέπει να τον διαιρέσουμε με το αντίστοιχο τυπικό σφάλμα ασυμμετρίας:

$$\tau_\alpha = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}$$

Συνάρτηση 6

Χαρακτηρισμός του δείγματος βάσει της ασυμμετρίας:

Αν $\pi_3 = \frac{a_3}{\tau_\alpha} > 2$, τότε το δείγμα παρουσιάζει θετική ασυμμετρία.

Αν $\pi_3 = \frac{a_3}{\tau_\alpha} < -2$, τότε το δείγμα παρουσιάζει αρνητική ασυμμετρία.

Αν $-2 < \pi_3 < 2$, τότε το δείγμα παρουσιάζει μηδενική ασυμμετρία.

Γενικότερα όταν υπάρχει ισοκατανομή των παρατηρήσεων γύρω από τη μέση τιμή, το δείγμα χαρακτηρίζεται ως συμμετρικό, διαφορετικά ασύμμετρο.

- Μέτρα Κύρτωσης

- Ο συντελεστής κύρτωσης (Kurtosis) εκφράζει τον βαθμό συγκέντρωσης των παρατηρήσεων ενός δείγματος γύρω από την κορυφή. Υπολογίζεται από τον τύπο:

$$a_4 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{s^4}$$

Συνάρτηση 7

Για να διαπιστωθεί αν το δείγμα είναι μεσόκυρτο, λεπτόκυρτο ή πλατύκυρτο, πρέπει να αξιολογηθεί, όπως και στη περίπτωση της ασυμμετρίας, το πηλίκο του συντελεστή κύρτωσης με το αντίστοιχο τυπικό σφάλμα του.

$$\pi_4 = \frac{\alpha_4}{\tau_k}, \text{ όπου } \tau_k = \tau_\alpha \sqrt{\frac{4(v^2 - 1)}{(v - 3)(v + 5)}}$$

Συνάρτηση 8

Έτσι:

Αν $\pi_4 > 2$, τότε το δείγμα χαρακτηρίζεται ως λεπτόκυρτο.

Αν $\pi_4 < -2$, τότε το δείγμα χαρακτηρίζεται ως πλατύκυρτο.

Αν $-2 < \pi_4 < 2$, τότε το δείγμα χαρακτηρίζεται ως μεσόκυρτο.

2.3 ΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ

Η Επαγωγική Στατιστική Συμπερασματολογία πραγματεύεται την ανάλυση και εξαγωγή συμπερασμάτων. Σε αυτό το σημείο πρέπει να υπογραμμιστεί ότι τα εξαγόμενα συμπεράσματα της Επαγωγικής Στατιστικής Συμπερασματολογίας είναι ισχυρότερα και ασφαλέστερα από εκείνα της Περιγραφικής Στατιστικής Ανάλυσης. Αυτό σημαίνει ότι σε περίπτωση διχογνωμίας κι όχι μόνο, τα συμπεράσματα της Επαγωγικής Στατιστικής Συμπερασματολογίας υπερισχύουν.

Τα εργαλεία που χρησιμοποιήθηκαν είναι τα παρακάτω:

1. Έλεγχος ικανοποιητικής προσαρμογής των δεδομένων στην Κανονική Κατανομή.
2. Έλεγχος Συσχέτισης μεταξύ των μεταβλητών.
3. Κατασκευή κατάλληλου προβλεπτικού μοντέλου για την τιμή της μεταβλητής στόχου.

2.3.1 Έλεγχος Κανονικής Κατανομής:

- Διαχωρισμό των μεταβλητών σε ποσοτικές, ποιοτικές ή ονομαστικές.
- Έλεγχος Κατανομής των μεταβλητών.
- Πραγματοποίηση παραμετρικού ή μη ελέγχου βάσει της κανονικότητας των ποσοτικών μεταβλητών.

2.3.2 Έλεγχος Συσχέτισης:

Συσχέτιση (Correlation) ονομάζεται η εξάρτηση 2 μεταβλητών μεταξύ τους και διαχωρίζεται σε:

- Θετική, όταν η αύξηση ή μείωση της τιμής της 1 μεταβλητής προκαλεί αύξηση ή μείωση της άλλης.

- Αρνητική, όταν η αύξηση ή μείωση της τιμής της 1 μεταβλητής προκαλεί μείωση ή αύξηση της άλλης.

Υπάρχουν διάφοροι συντελεστές συσχέτισης, η χρήση των οποίων βασίζεται στο είδος των μεταβλητών καθώς και στη παρουσία ή μη Κανονικής Κατανομής. Οι σημαντικότεροι και πιο εύχρηστοι είναι οι παρακάτω:

Ο Συντελεστής Συσχέτισης Spearman (rho) χρησιμοποιείται για τον μη-παραμετρικό έλεγχο της στατιστικής εξάρτησης μεταξύ δύο μεταβλητών. Χρησιμοποιείται σε ποιοτικές μεταβλητές καθώς και σε ποσοτικές μεταβλητές, οι οποίες δεν παρουσιάζουν Κανονική Κατανομή. Υπολογίζεται από τον τύπο:

$$rho = 1 - \frac{6 \sum_{i=1}^n \delta_i^2}{n(n^2 - 1)}$$

Συνάρτηση 9

όπου n ο αριθμός των παρατηρήσεων και $\delta_i = x_i - y_i$

Ο Συντελεστής Συσχέτισης Pearson (r) αφορά ποσοτικές μεταβλητές, οι οποίες παρουσιάζουν Κανονική Κατανομή. Υπολογίζεται από τον τύπο:

$$r = \frac{\sum_{i=1}^n x_i y_i - \frac{\left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n}}{\sqrt{\left(\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}\right)\left(\sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}\right)}}$$

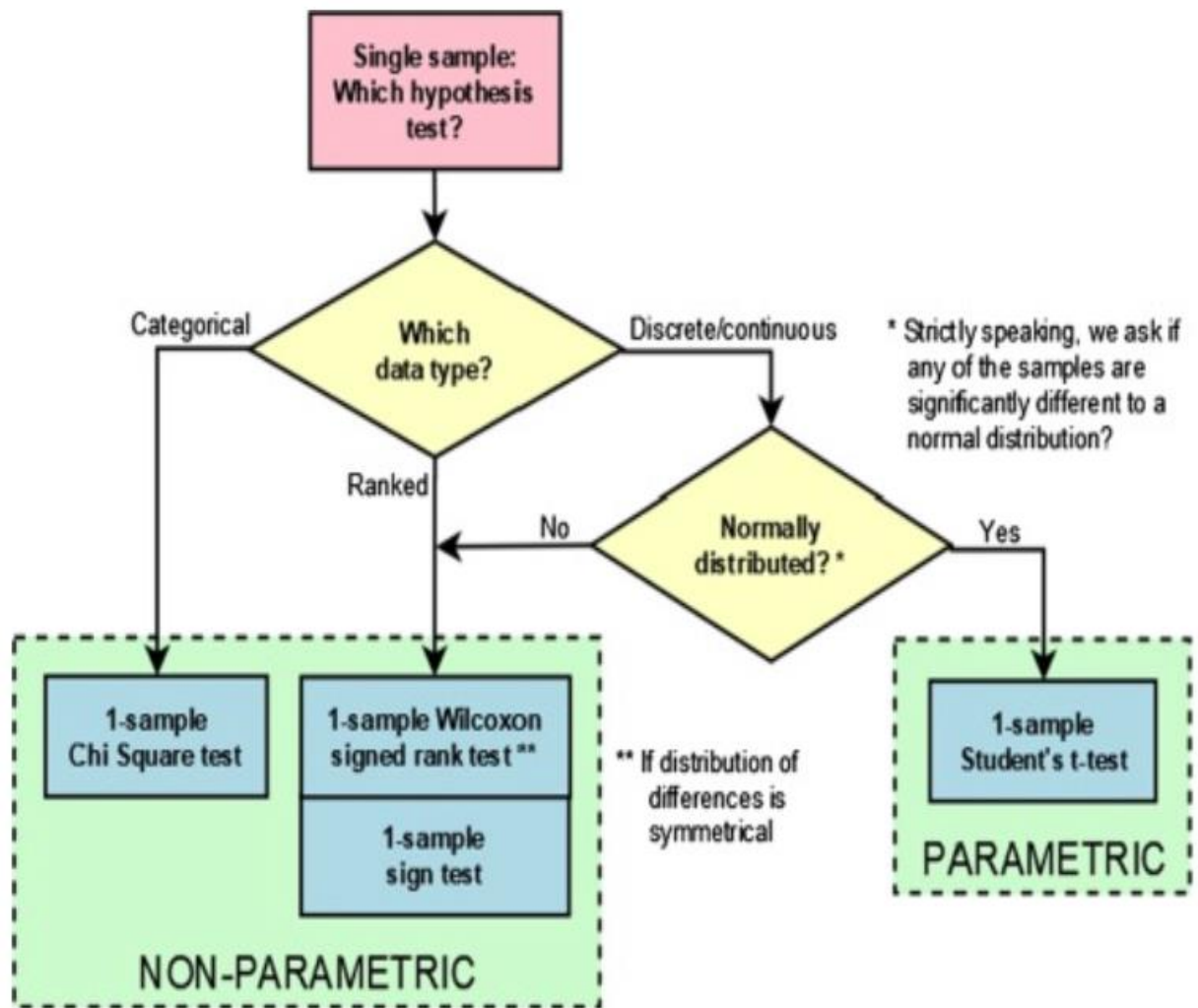
Συνάρτηση 10

όπου n ο αριθμός των παρατηρήσεων και x, y οι τιμές των 2 μεταβλητών.

Επιλογή Κατάλληλης Διαδικασίας Ελέγχου Συσχέτισης

Η επιλογή του ανωτέρω συντελεστή εξαρτάται από:

- Το είδος των μεταβλητών.
- Παρουσία ή μη Κανονικής Κατανομής.



Εικόνα 5: Διαδικασία επιλογής παραμετρικού ή μη ελέγχου (πηγή: www.doi.org/10.1016/B978-0-12-369492-8.50014-5)

Parametric v nonparametric tests

PARAMETRIC

- Evaluate hypotheses for a particular parameter, usually the population mean
- Quantitative data
- Require assumptions about the distributional characteristics of the population distribution
 - Normal shape
 - Equal variance
- More powerful than nonparametric tests when assumptions met

NONPARAMETRIC

- Evaluate hypotheses for *entire* population distributions
- Quantitative, ranked, qualitative data
- Require no assumptions ("distribution-free") so used with non-normal distributions and when variances of the groups are not equal
- Generally easy to compute

Εικόνα 6: Χαρακτηριστικά παραμετρικού και μη ελέγχου
(πηγή: www.nysora.com/)

2.3.3 Ανάλυση Παλινδρόμησης

Ανάλυση Παλινδρόμησης (Regression Analysis) ονομάζεται η διαδικασία η οποία πραγματοποιείται με σκοπό τη δημιουργία μοντέλου πρόβλεψης της μεταβλητής στόχου βάσει των υπολοίπων μεταβλητών.

Η μεταβλητή στόχος (y) ονομάζεται εξαρτημένη μεταβλητή και οι υπόλοιπες ανεξάρτητες (x). Γενικά το μοντέλο πρόβλεψης είναι της μορφής:

$$y = \beta_0 + \beta_1 x$$

Συνάρτηση 11

όπου β_1 : ο συντελεστής διεύθυνσης της ευθείας
και β_0 : σημείο τομής της y στον άξονα yy' .

2.3.4 Λογιστική Παλινδρόμηση

Η Λογιστική Παλινδρόμηση (Logistic Regression) αναφέρεται σε μη γραμμική σχέση μιας εξαρτημένης μεταβλητής βάσει κάποιων ανεξάρτητων. Αναλυτικότερα, η εξαρτημένη μεταβλητή είναι κατηγορική και συνήθως δυαδικής μορφής. Με άλλα λόγια, το είδος της εξαρτημένης μεταβλητή αποτελεί και την βασική διαφορά με την Γραμμική Παλινδρόμηση.

Στην παρούσα μελέτη παρατηρήθηκαν 2 είδη κατηγορικών εξαρτημένων μεταβλητών: Δίτιμες (2 τιμές) και Τακτικές (3 και περισσότερες τιμές με σχέση ανισότητας μεταξύ τους).

Τα μοντέλα πρόβλεψης ανά περίπτωση είναι τα παρακάτω:

$$f(z) = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}}$$

$$z = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_v x_v + \beta_0$$

x_1, \dots, x_v : μεταβλητές

$\alpha_1, \dots, \alpha_v$: συντελεστές

β_0 : σταθερά

Συνάρτηση 12: Μοντέλο Πρόβλεψης για δίτιμη εξαρτημένη μεταβλητή

$$\log it[P(y \leq k)] = \log \frac{P(y \leq k)}{1 - P(y \leq k)}$$

$$P(y \leq k) = p_1 + p_2 + p_3 + \dots + p_k, \quad k = 1, 2, \dots, K$$

$$P(y \leq k) = \frac{e^z}{1 + e^z}$$

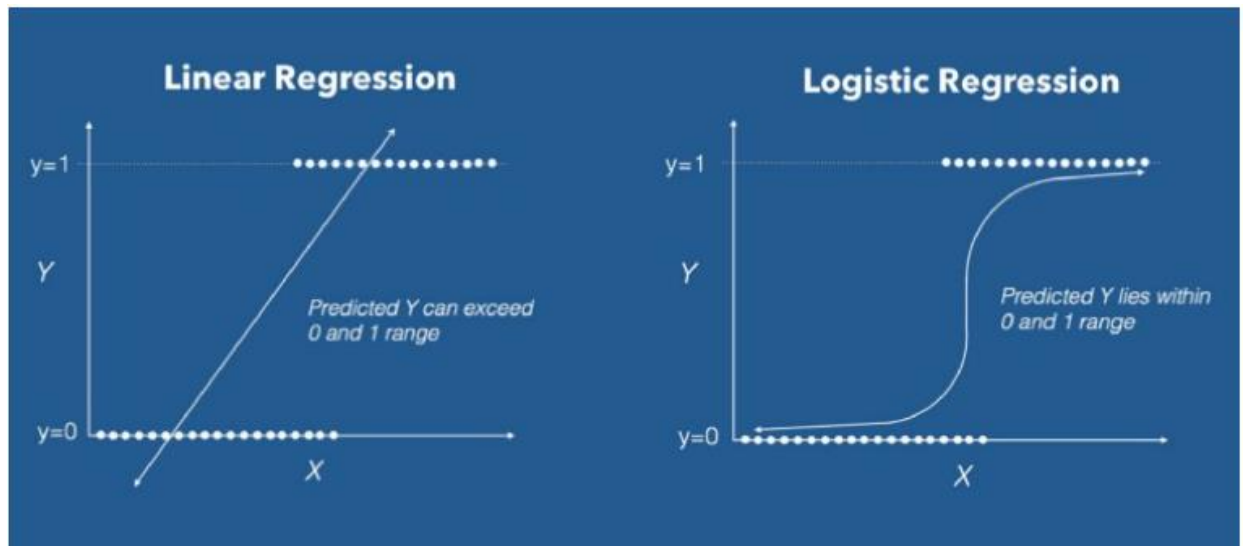
$$z = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_v x_v + \beta_0$$

x_1, \dots, x_v : μεταβλητές

$\alpha_1, \dots, \alpha_v$: συντελεστές

β_0 : σταθερά

Συνάρτηση 13: Μοντέλο Πρόβλεψης για τακτική εξαρτημένη μεταβλητή



Εικόνα 7: Γραφική διαφορά Γραμμικής και Λογιστικής Παλινδρόμησης
(πηγή: www.medium.com)

3. ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ

Στο παρόν Κεφάλαιο, αφού πρώτα κατηγοριοποιηθούν οι μεταβλητές που συνθέτουν τα ερευνητικά ερωτήματα (αναλύθηκαν στο 1^ο Κεφάλαιο), σε ποσοτικές, ποιοτικές κα, θα πραγματοποιηθεί η περιγραφική στατιστική ανάλυση τους, όπως αναλύθηκε στο 2^ο Κεφάλαιο.

3.1 ΠΟΣΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ

Ενδεικτικά στον παρακάτω πίνακα, παρουσιάζονται αποτελέσματα που αφορούν την Περιγραφική Στατιστική Ανάλυση, κατόπιν χρήσης του SPSS.

Statistics

ΗΛΙΚΙΑ		
N	Valid	1000
	Missing	0
Mean		45,55
Median		46,00
Mode		47
Std. Deviation		14,291
Variance		204,236
Skewness		,004
Std. Error of Skewness		,077
Kurtosis		-,484
Std. Error of Kurtosis		,155
Range		83
Percentiles	25	35,00
	50	46,00
	75	56,00

Πίνακας 1: Μέτρα Κεντρικής Τάσης της μεταβλητής ΗΛΙΚΙΑ

3.1.1 ΜΕΤΡΑ ΚΕΝΤΡΙΚΗΣ ΤΑΣΗΣ Ή ΘΕΣΗΣ

Μεταβλητή	Μέση τιμή (Mean)	Διάμεσος (Median)	Κορυφή (Mode)
Ηλικία	45,55	46	47
Δείκτης Μάζας Σώματος (ΔΜΣ)	32,19	31,18	32,5327
Ύψος	1,667	1,650	1,6
Βάρος	89,262	87	73

Πίνακας 2: Μέτρα Κεντρικής Τάσης ποσοτικών μεταβλητών

ΗΛΙΚΙΑ

Η μέση τιμή του δείγματος ισούται με 45,55 και αποτελεί την αντιπροσωπευτική τιμή των παρατηρήσεων.

Η διάμεσος του ισούται με 46. Αποτελεί έναν διαφορετικό τρόπο περιγραφής της κεντρικής τάσης και πρακτικά εκφράζεται ότι το 50% των παρατηρήσεων έχει τιμή μεγαλύτερη ή ίση από τη διάμεσο (το υπόλοιπο 50% έχει τιμή μικρότερη ή ίση από τη διάμεσο).

Η κορυφή του ισούται με 47 και είναι μοναδική.

ΔΜΣ

Η μέση τιμή του ισούται με 32,19 και αποτελεί την αντιπροσωπευτική τιμή των παρατηρήσεων.

Η διάμεσος του ισούται με 31,18. Αποτελεί έναν διαφορετικό τρόπο περιγραφής της κεντρικής τάσης και πρακτικά εκφράζεται ότι το 50% των ανωτέρω παρατηρήσεων έχει τιμή μεγαλύτερη ή ίση από τη διάμεσο (το υπόλοιπο 50% έχει τιμή μικρότερη ή ίση από τη διάμεσο).

Η κορυφή του ισούται με 32,5327 και είναι μοναδική.

ΥΨΟΣ

Η μέση τιμή του ισούται με 1,667 και αποτελεί την αντιπροσωπευτική τιμή των ανωτέρω παρατηρήσεων.

Η διάμεσος του ισούται με 1,650. Αποτελεί έναν διαφορετικό τρόπο περιγραφής της κεντρικής τάσης και πρακτικά εκφράζεται ότι το 50% των ανωτέρω παρατηρήσεων έχει τιμή μεγαλύτερη ή ίση από τη διάμεσο (το υπόλοιπο 50% έχει τιμή μικρότερη ή ίση από τη διάμεσο).

Η κορυφή (Mode) του ισούται με 1,6 και είναι μοναδική.

ΒΑΡΟΣ

Η μέση τιμή του ισούται με 89,262 και αποτελεί την αντιπροσωπευτική τιμή των ανωτέρω παρατηρήσεων.

Η διάμεσος του ισούται με 87. Αποτελεί έναν διαφορετικό τρόπο περιγραφής της κεντρικής τάσης και πρακτικά εκφράζεται ότι το 50% των ανωτέρω παρατηρήσεων έχει τιμή μεγαλύτερη ή ίση από τη διάμεσο (το υπόλοιπο 50% έχει τιμή μικρότερη ή ίση από τη διάμεσο).

Η κορυφή του ισούται με 73 και είναι μοναδική.

3.1.2 ΜΕΤΡΑ ΔΙΑΣΠΟΡΑΣ

Μεταβλητή	Διασπορά (Variance)	Τυπική Απόκλιση (Standard Deviation)	Εύρος (Range)	Ενδοτεταρτημοριακό Εύρος (Interquartile Range)	Συντελεστής μεταβλητότητας (Coefficient of Variability)
Ηλικία	204,236	14,291	83	21	31,37%
Δείκτης Μάζας Σώματος (ΔΜΣ)	75,156	8,669	128,47	8,42	26,93%
Ύψος	0,008	0,0912	1,3	0,06	5,47%
Βάρος	464,474	21,55	150,5	27,8	24,14%

Πίνακας 3: Μέτρα Διασποράς ποσοτικών μεταβλητών

ΗΛΙΚΙΑ

Η διασπορά του ισούται με 204,236.

Η τυπική απόκλιση αποτελεί την τετραγωνική ρίζα της διασποράς και ισούται με 14,291. Το εύρος του ισούται με 83 και παρουσιάζει τη μέγιστη δυνατή απόκλιση μεταξύ 2 οποιονδήποτε παρατηρήσεων του δείγματος.

Το ενδοτεταρτημοριακό εύρος ισούται με τη διαφορά Q_3-Q_1 , $56-35=21$. Αυτό σημαίνει ότι σε μια περιοχή τιμών με πλάτος 21 εκ., συσσωρεύεται το 50% ασθενών.

Ο συντελεστής μεταβλητικότητας ισούται με 31,37%. Το CV είναι μεγαλύτερο από 10% οπότε το δείγμα χαρακτηρίζεται ως αναμοιογενές. Αυτό σημαίνει ότι οι παρατηρήσεις του δείγματος παρουσιάζουν μεταξύ τους μικρές αποκλίσεις.

ΔΜΣ

Η διασπορά του ισούται με 75,156.

Η τυπική απόκλιση αποτελεί την τετραγωνική ρίζα της διασποράς και ισούται με 8,669.

Το εύρος του ισούται με 128,47 και παρουσιάζει τη μέγιστη δυνατή απόκλιση μεταξύ 2 οποιονδήποτε παρατηρήσεων του ανωτέρου δείγματος.

Το ενδοτεταρτημοριακό εύρος ισούται με τη διαφορά Q_3-Q_1 , $35,64-27,22=8,42$. Αυτό σημαίνει ότι σε μια περιοχή τιμών με πλάτος 8,42 εκ., συσσωρεύεται το 50% ασθενών.

Ο συντελεστής μεταβλητικότητας ισούται με 26,93%. Το CV είναι μεγαλύτερο από 10% οπότε το δείγμα χαρακτηρίζεται ως αναμοιογενές. Αυτό σημαίνει ότι οι παρατηρήσεις του δείγματος παρουσιάζουν μεταξύ τους μικρές αποκλίσεις.

ΥΨΟΣ

Η διασπορά του ισούται με 0,008.

Η τυπική απόκλιση αποτελεί την τετραγωνική ρίζα της διασποράς και ισούται με 0,0912.

Το εύρος του ισούται με 1,3 και παρουσιάζει τη μέγιστη δυνατή απόκλιση μεταξύ 2 οποιονδήποτε παρατηρήσεων του ανωτέρου δείγματος.

Το ενδοτεταρτημοριακό εύρος (Interquartile Range) ισούται με τη διαφορά Q_3-Q_1 , $1,680-1,620=0,06$. Αυτό σημαίνει ότι σε μια περιοχή τιμών με πλάτος 0,06 εκ., συσσωρεύεται το 50% ασθενών.

Ο συντελεστής μεταβλητικότητας ισούται με 5,47%. Το CV είναι μικρότερος από 10% οπότε το δείγμα χαρακτηρίζεται ως ομοιογενές. Αυτό σημαίνει ότι οι παρατηρήσεις του παρουσιάζουν μεταξύ τους μικρές αποκλίσεις.

ΒΑΡΟΣ

Η διασπορά του ισούται με 464,474.

Η τυπική απόκλιση αποτελεί την τετραγωνική ρίζα της διασποράς και ισούται με 21,55.

Το εύρος του ισούται με 150,5 και παρουσιάζει τη μέγιστη δυνατή απόκλιση μεταξύ 2 οποιονδήποτε παρατηρήσεων του ανωτέρου δείγματος.

Το ενδοτεταρτημοριακό εύρος ισούται με τη διαφορά Q_3-Q_1 , 27,8. Αυτό σημαίνει ότι σε μια περιοχή τιμών με πλάτος 27,8 εκ., συσσωρεύεται το 50% ασθενών.

Ο συντελεστής μεταβλητικότητας ισούται με 24,14%. Το CV είναι μεγαλύτερος από 10% οπότε το δείγμα χαρακτηρίζεται ως ανομοιογενές. Αυτό σημαίνει ότι οι παρατηρήσεις του παρουσιάζουν μεταξύ τους μεγάλες αποκλίσεις.

3.1.3 ΜΕΤΡΑ ΑΣΥΜΜΕΤΡΙΑΣ ΚΑΙ ΚΥΡΤΩΣΗΣ

Μεταβλητή	Συντελεστής Ασυμμετρίας (Skewness)	Τυπικό Σφάλμα Ασυμμετρίας	Συντελεστής Κύρτωσης (Kurtosis)	Τυπικό Σφάλμα Κύρτωσης
Ηλικία	0,004	0,077	-0,484	0,155
Δείκτης Μάζας Σώματος (ΔΜΣ)	4,317	0,077	43,211	0,155
Ύψος	-2,094	0,077	26,717	0,155
Βάρος	0,776	0,077	1,181	0,155

Πίνακας 4: Μέτρα Ασυμμετρίας και Κύρτωσης ποσοτικών μεταβλητών

ΗΛΙΚΙΑ

Μέτρα Ασυμμετρίας

Ο συντελεστής ασυμμετρίας ισούται με 0,004. Για να πραγματοποιηθεί ο έλεγχος της ασυμμετρίας προχωρούμε στη διαίρεση του συντελεστή προς το αντίστοιχο τυπικό σφάλμα του και ισούται με 0,077. Παρατηρείται ότι το πηλίκο είναι μικρότερο του 2 και μεγαλύτερο του -2, οπότε έχει μηδενική ασυμμετρία. Αυτό σημαίνει ότι οι περισσότερες παρατηρήσεις του δείγματος βρίσκονται κοντά στη μέση τιμή του.

Μέτρα Κύρτωσης

Ο συντελεστής κύρτωσης ισούται με -0,484. Στη συνέχεια θα πραγματοποιηθεί έλεγχος της κύρτωσης, υπολογίζοντας το πηλίκο του συντελεστή προς το αντίστοιχο τυπικό σφάλμα του (0,155). Παρατηρείται ότι το πηλίκο είναι μικρότερο του -2, οπότε πλατύκυρτο. Αυτό σημαίνει ότι οι παρατηρήσεις του δείγματος που είναι συσσωρευμένες γύρω από την κορυφή του δείγματος είναι ιδιαίτερα μικρές.

ΔΜΣ

Μέτρα Ασυμμετρίας

Ο συντελεστής ασυμμετρίας (Skewness) ισούται με 4,317. Για να πραγματοποιηθεί ο έλεγχος της ασυμμετρίας προχωρούμε στη διαίρεση του συντελεστή προς το αντίστοιχο τυπικό σφάλμα του και ισούται με 0,077. Παρατηρείται ότι το πηλίκο είναι μεγαλύτερο του 2, οπότε έχει θετική ασυμμετρία. Αυτό σημαίνει ότι οι περισσότερες παρατηρήσεις του δείγματος είναι μικρότερες ή ίσες της μέσης τιμής του.

Μέτρα Κύρτωσης

Ο συντελεστής κύρτωσης (Kurtosis) ισούται με 43,211. Στη συνέχεια θα πραγματοποιηθεί έλεγχος της κύρτωσης, υπολογίζοντας το πηλίκο του συντελεστή προς το αντίστοιχο τυπικό

σφάλμα του (0,155). Παρατηρείται ότι το πηλίκο είναι μικρότερο του -2, οπότε πλατύκυρτο. Αυτό σημαίνει ότι οι παρατηρήσεις του δείγματος που είναι συσσωρευμένες γύρω από την κορυφή του δείγματος είναι ιδιαίτερα μικρές.

ΥΨΟΣ

Μέτρα Ασυμμετρίας

Ο συντελεστής ασυμμετρίας (Skewness) ισούται με -2,094. Για να πραγματοποιηθεί ο έλεγχος της ασυμμετρίας προχωρούμε στη διαίρεση του συντελεστή προς το αντίστοιχο τυπικό σφάλμα του και ισούται με 0,077. Παρατηρείται ότι το πηλίκο είναι μικρότερο του -2, οπότε έχει αρνητική ασυμμετρία. Αυτό σημαίνει ότι οι περισσότερες παρατηρήσεις του δείγματος είναι μεγαλύτερες της μέσης τιμής του.

Μέτρα Κύρτωσης

Ο συντελεστής κύρτωσης (Kurtosis) ισούται με 26,717. Στη συνέχεια θα πραγματοποιηθεί έλεγχος της κύρτωσης, υπολογίζοντας το πηλίκο του συντελεστή προς το αντίστοιχο τυπικό σφάλμα του (0,155). Παρατηρείται ότι το πηλίκο είναι μεγαλύτερο του 2, οπότε λεπτόκυρτο. Αυτό σημαίνει ότι οι περισσότερες παρατηρήσεις του δείγματος συσσωρευμένες γύρω από την κορυφή του δείγματος.

ΒΑΡΟΣ

Μέτρα Ασυμμετρίας

Ο συντελεστής ασυμμετρίας (Skewness) ισούται με 0,776. Για να πραγματοποιηθεί ο έλεγχος της ασυμμετρίας προχωρούμε στη διαίρεση του συντελεστή προς το αντίστοιχο τυπικό σφάλμα του και ισούται με 0,077. Παρατηρείται ότι το πηλίκο είναι μεγαλύτερο του 2, οπότε έχει θετική ασυμμετρία. Αυτό σημαίνει ότι οι περισσότερες παρατηρήσεις του δείγματος είναι μικρότερες ή ίσες της μέσης τιμής του.

Μέτρα Κύρτωσης

Ο συντελεστής κύρτωσης (Kurtosis) ισούται με 1,181. Στη συνέχεια θα πραγματοποιηθεί έλεγχος της κύρτωσης, υπολογίζοντας το πηλίκο του συντελεστή προς το αντίστοιχο τυπικό σφάλμα του (0,155). Παρατηρείται ότι το πηλίκο είναι μεγαλύτερο του 2, οπότε λεπτόκυρτο. Αυτό σημαίνει ότι οι περισσότερες παρατηρήσεις του δείγματος συσσωρευμένες γύρω από την κορυφή του δείγματος.

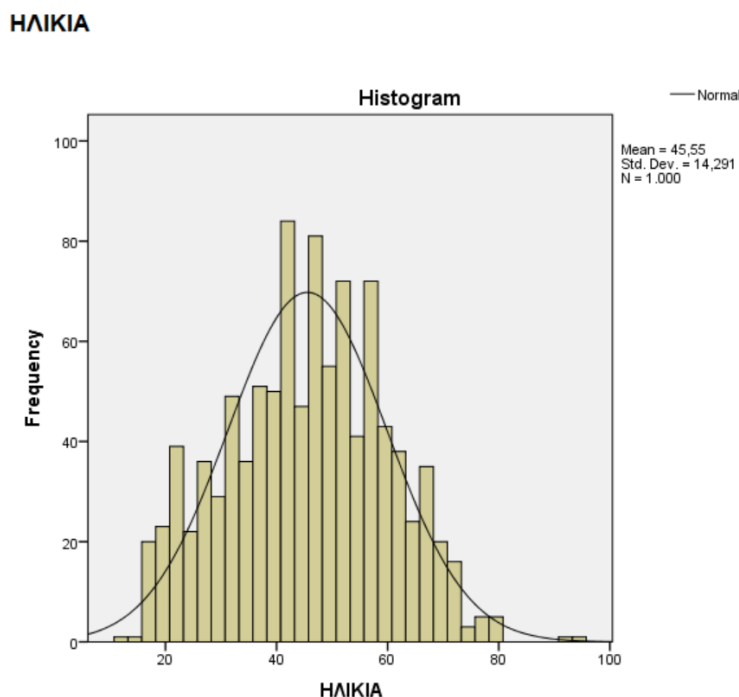
3.1.4 ΚΑΤΑΝΟΜΗ

Στα παρακάτω διαγράμματα παρουσιάζεται η κατανομή των ανωτέρω μεταβλητών με Ραβδόγραμμα και P-P Plot.

ΗΛΙΚΙΑ

Ραβδόγραμμα

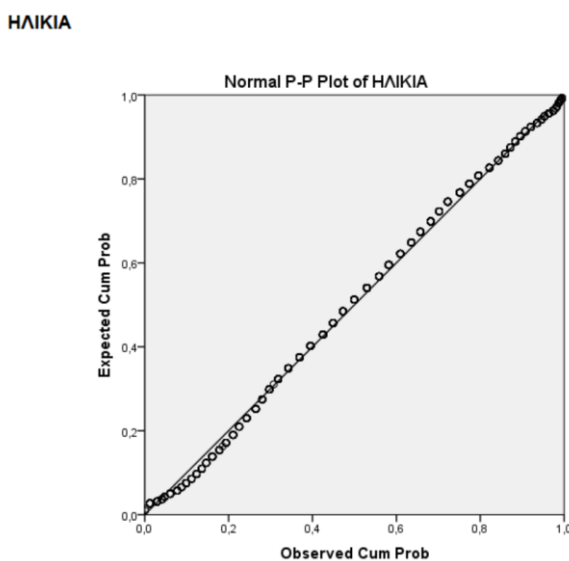
Στο παρακάτω διάγραμμα παρατηρείται ότι οι τιμές προσαρμόζονται ικανοποιητικά στην Κανονική Κατανομή.



Διάγραμμα 1: Έλεγχος Κατανομής μεταβλητής ΗΛΙΚΙΑΣ με Ραβδόγραμμα

P-P Plot

Εναλλακτικός τρόπος απεικόνισης και εξέτασης της προσαρμοστικότητας των δεδομένων στην Κανονική Κατανομή είναι το P-P Plot. Παρατηρούμε ότι τα σημεία είναι κοντά στη διχοτόμο της γωνίας 1^{ου} τεταρτημορίου, οπότε ισχύει η ανωτέρω διαπίστωση.

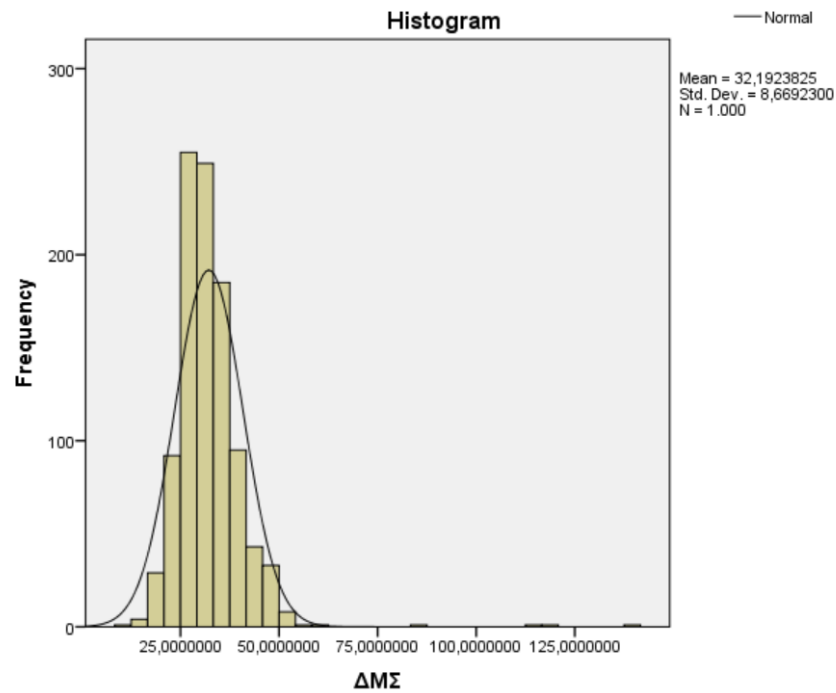


Διάγραμμα 2: Έλεγχος Κατανομής μεταβλητής ΗΛΙΚΙΑΣ με P-P Plot

ΔΜΣ

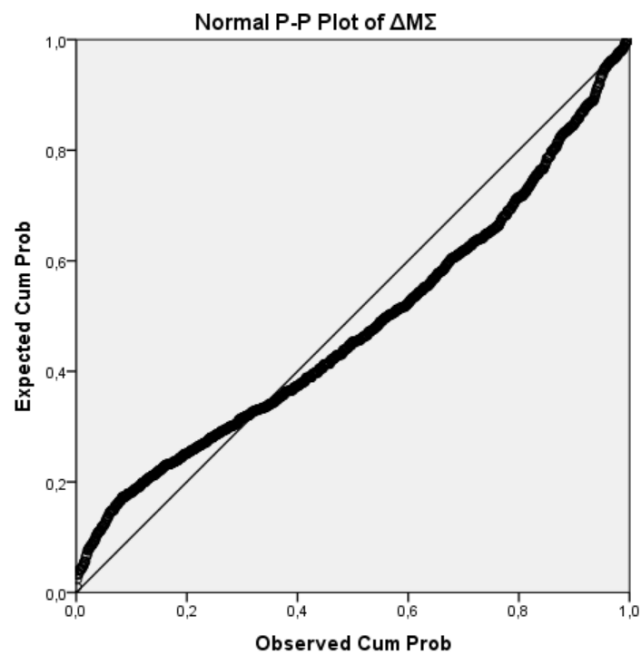
Στις παρακάτω εικόνες παρατηρείται ότι οι τιμές της μεταβλητής προσαρμόζονται ικανοποιητικά στην Κανονική Κατανομή.

ΔΜΣ



Διάγραμμα 3: Έλεγχος Κατανομής μεταβλητής ΔΜΣ με Ραβδόγραμμα

ΔΜΣ

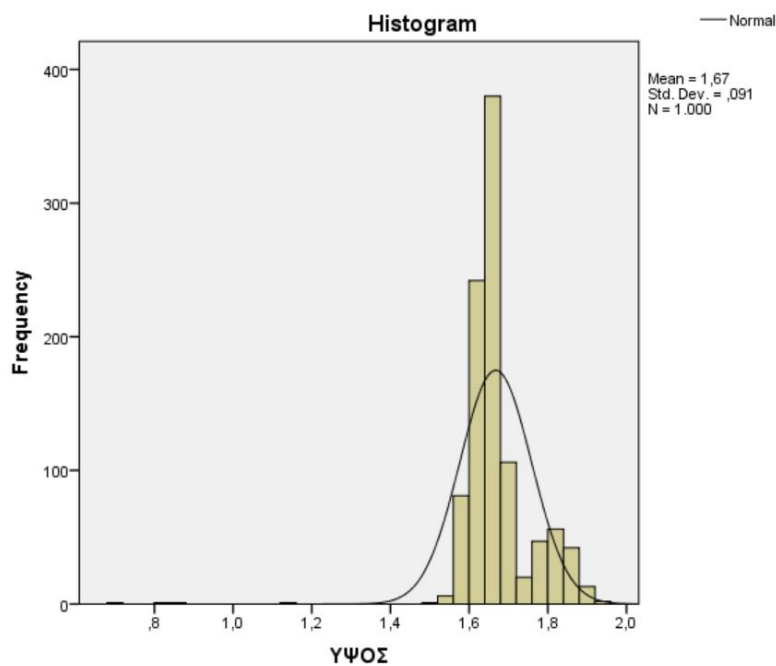


Διάγραμμα 4: Έλεγχος Κατανομής μεταβλητής ΔΜΣ με P-P Plot

ΥΨΟΣ

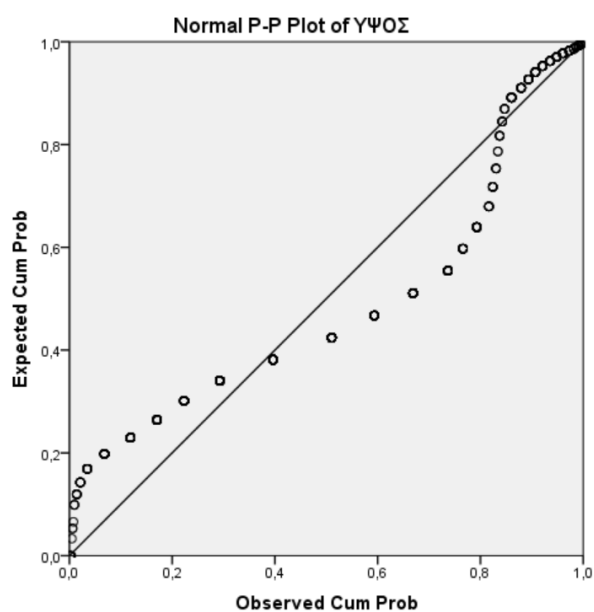
Στις παρακάτω εικόνες παρατηρείται ότι οι τιμές της μεταβλητής δεν προσαρμόζονται ικανοποιητικά στην Κανονική Κατανομή.

ΥΨΟΣ



Διάγραμμα 5: Έλεγχος Κατανομής μεταβλητής ΥΨΟΣ με Ραβδόγραμμα

ΥΨΟΣ

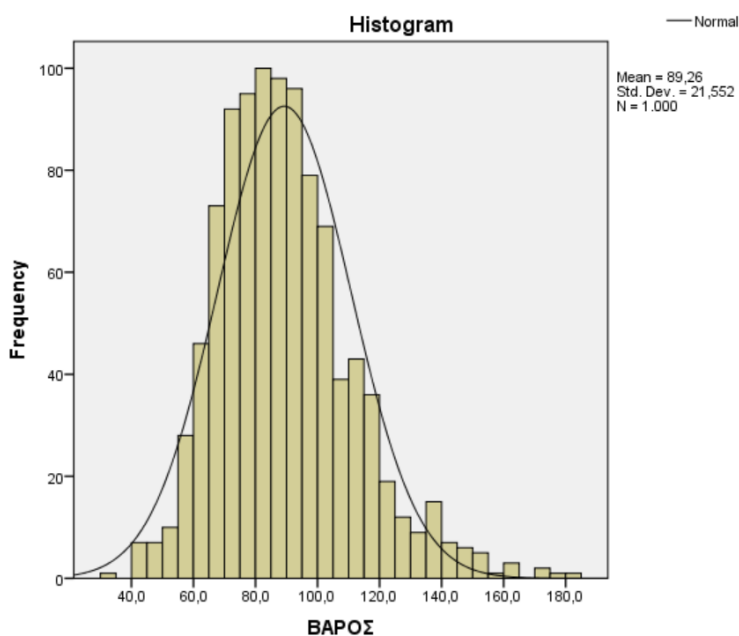


Διάγραμμα 6: Έλεγχος Κατανομής μεταβλητής ΥΨΟΣ με P-P Plot

ΒΑΡΟΣ

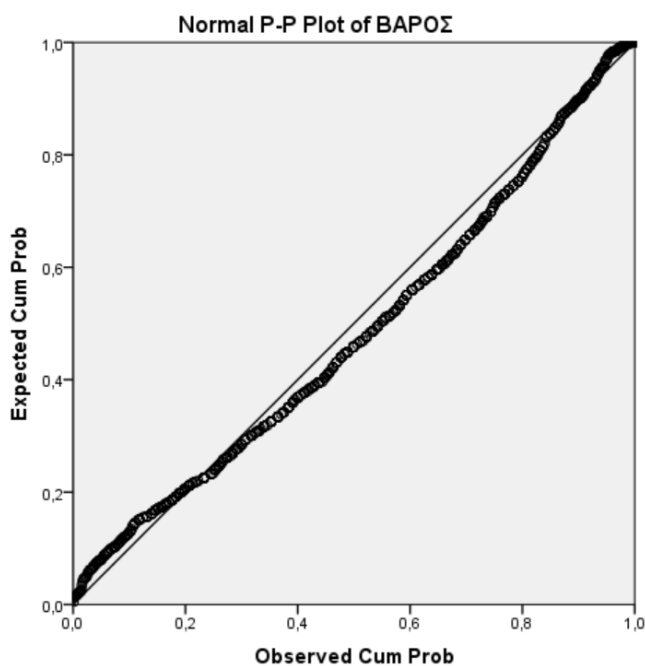
Στις παρακάτω εικόνες παρατηρείται ότι οι τιμές της μεταβλητής προσαρμόζονται ικανοποιητικά στην Κανονική Κατανομή.

ΒΑΡΟΣ



Διάγραμμα 7: Έλεγχος Κατανομής μεταβλητής ΒΑΡΟΣ με Ραβδόγραμμα

ΒΑΡΟΣ



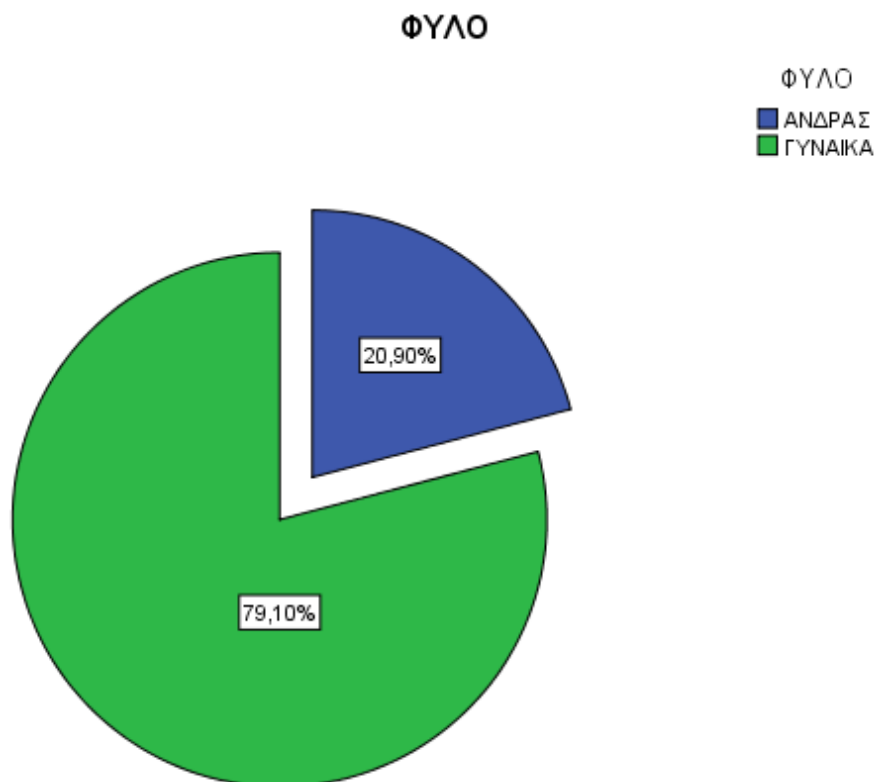
Διάγραμμα 8: Έλεγχος Κατανομής μεταβλητής ΒΑΡΟΣ με P-P Plot

3.2 ΠΟΙΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ

3.2.1 ΓΕΝΙΚΕΥΜΕΝΗ ΑΝΑΛΥΣΗ

3.2.1.1 ΦΥΛΟ

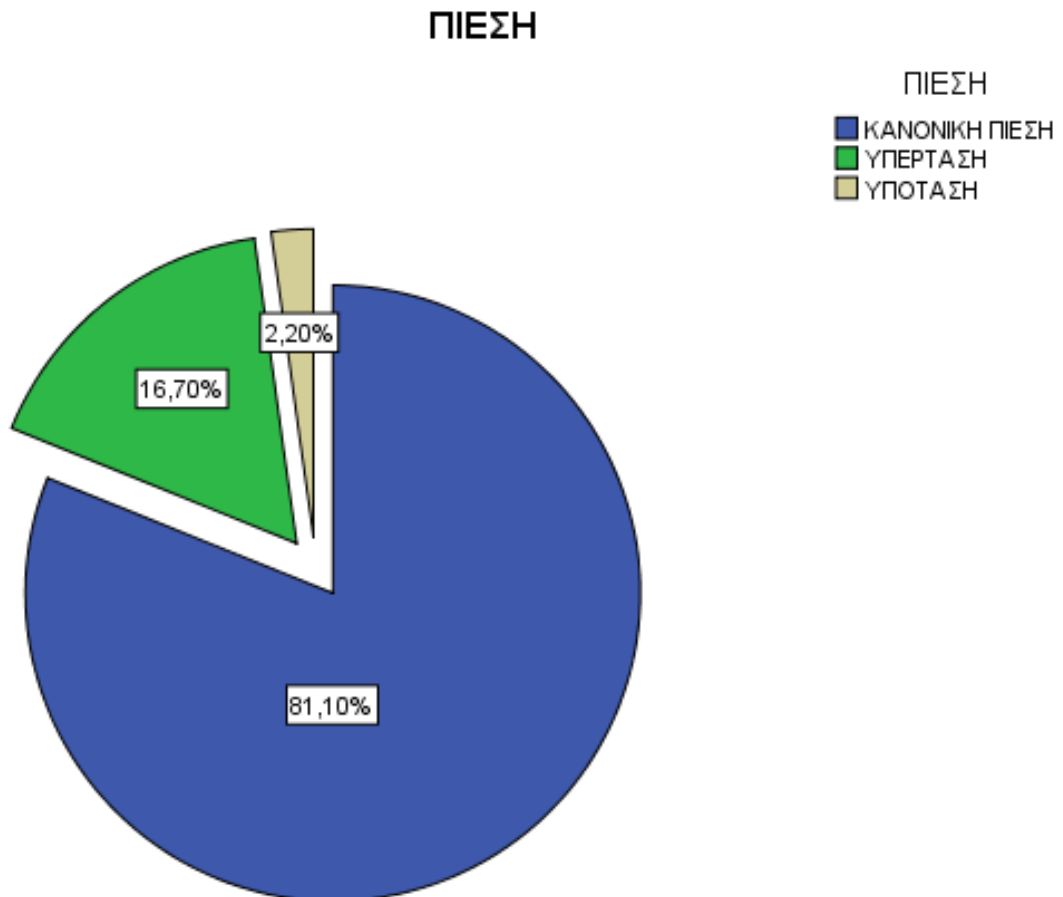
Σύμφωνα με το παρακάτω διάγραμμα σε μορφή Πίτας, το 20,90% των ασθενών του δείγματός μας είναι άνδρες, ενώ το 79,10% αντίστοιχα, είναι γυναίκες.



Διάγραμμα 9: Φύλο ασθενών

3.2.1.2 ΠΙΕΣΗ

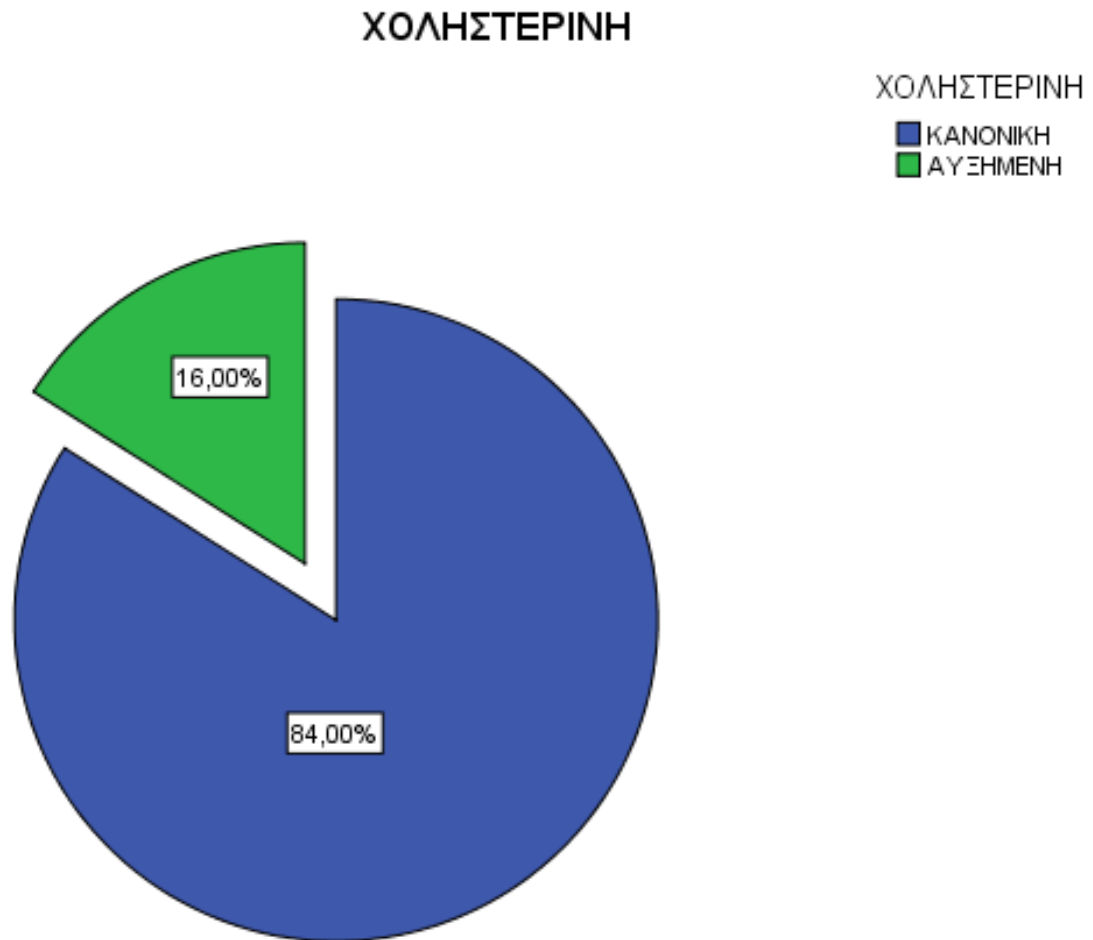
Σύμφωνα με το παρακάτω διάγραμμα σε μορφή Πίτας, στο 81,10% των ασθενών του δείγματός μας παρατηρούνται κανονικές τιμές στην Πίεσή τους. Στο 16,70% και 2,20% των ασθενών εμφανίζεται Υπέρταση και Υπόταση αντίστοιχα.



Διάγραμμα 10: Ποσοστά Πίεσης ασθενών

3.2.1.3 ΧΟΛΗΣΤΕΡΙΝΗ

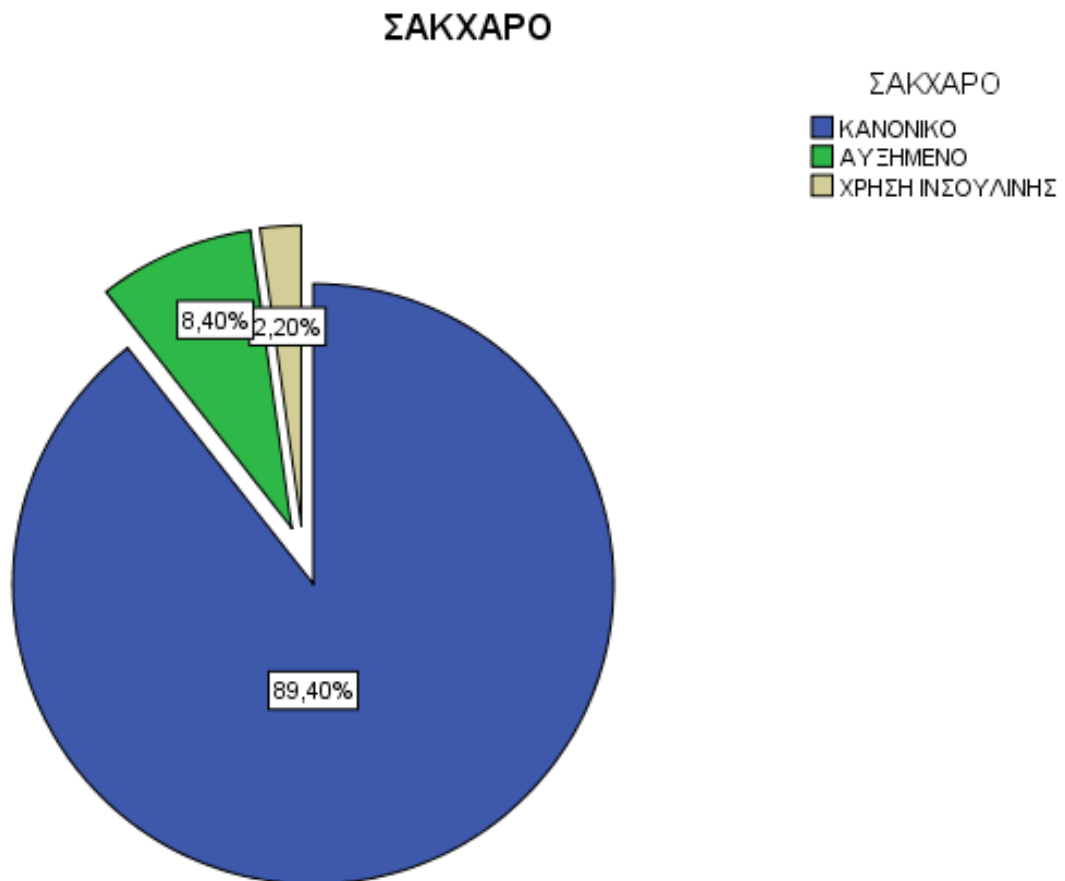
Σύμφωνα με το παρακάτω διάγραμμα σε μορφή Πίτας, στο 84% και 16% των ασθενών του δείγματός μας παρατηρούνται κανονικές τιμές και υψηλές τιμές χοληστερίνης



Διάγραμμα 11: Ποσοστά Χοληστερίνης ασθενών

3.2.1.4 ΣΑΚΧΑΡΟ

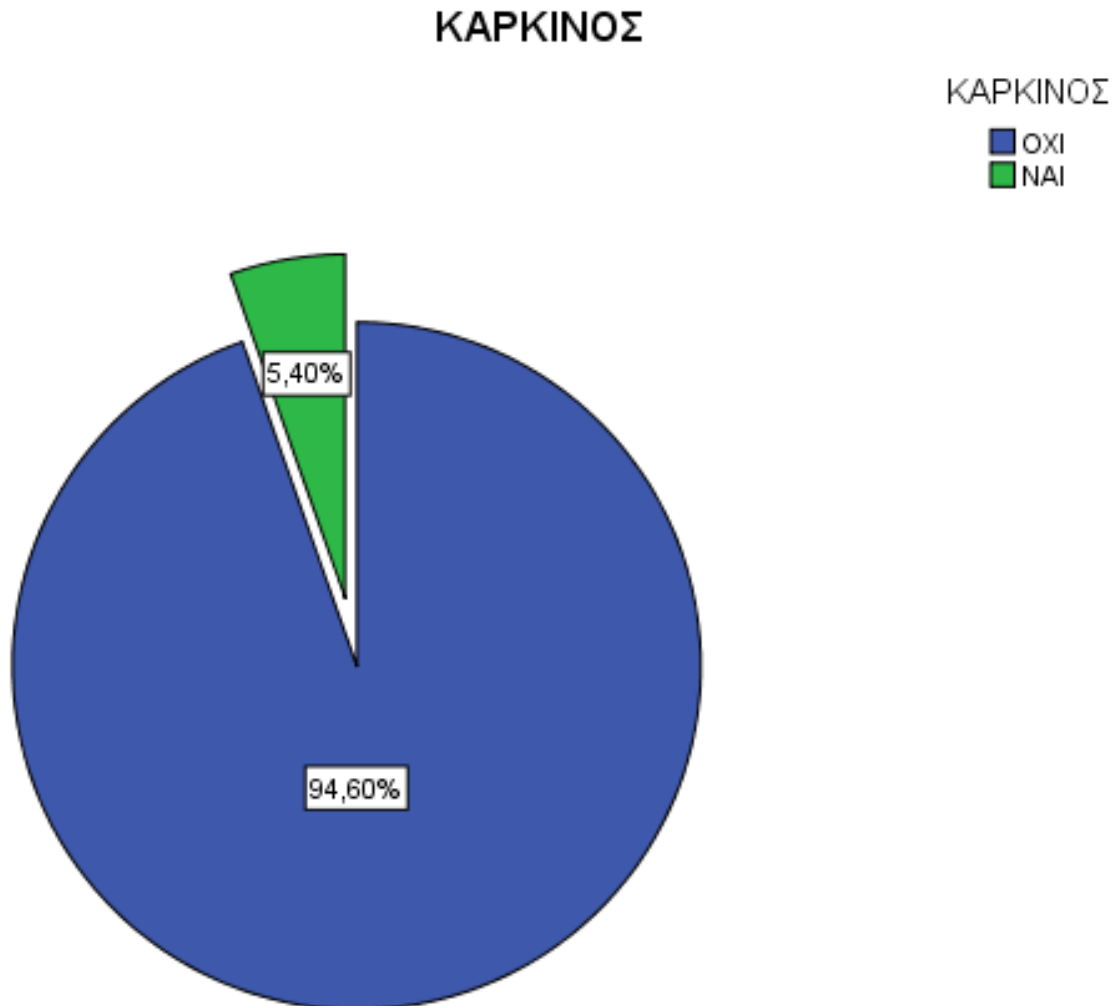
Σύμφωνα με το παρακάτω διάγραμμα σε μορφή Πίτας, στο 89,40% των ασθενών του δείγματός μας παρατηρούνται κανονικές τιμές στο Σάκχαρο. Στο 8,40% εμφανίζεται αυξημένη τιμή Σακχάρου και το 2,20% των ασθενών κάνει χρήση ποσότητας Ινσουλίνης.



Διάγραμμα 12: Ποσοστά Σακχάρου ασθενών

3.2.1.5 ΚΑΡΚΙΝΟΣ

Σύμφωνα με το παρακάτω διάγραμμα σε μορφή Πίτας, το 94,60 των ασθενών του δείγματός μας δεν έχει νοσήσει από κάποια μορφή Καρκίνου. Αντίστοιχα το 5,40% έχει νοσήσει.



Διάγραμμα 13: Εμφάνιση Καρκίνου στους ασθενείς

3.2.2 ΑΝΑΛΥΣΗ ΠΑΘΗΣΕΩΝ ΒΑΣΕΙ ΦΥΛΟΥ

3.2.2.1 ΑΡΙΘΜΗΤΙΚΑ

Στον παρακάτω Πίνακα, παρουσιάζονται αναλυτικά αριθμητικά δεδομένα των ασθενών ανά Φύλο και Πάθηση. Επιπλέον σε κάθε πάθηση – ασθένεια δίνονται αναλυτικότερες πληροφορίες.

Φύλο	Πίεση			Χοληστερίνη		Σάκχαρο			Καρκίνος	
	LOW	MID.	HIGH	MID.	HIGH	MID.	HIGH	EX. HIGH	NO	YES
Άνδρες	0	171	38	178	31	164	32	13	203	6
Γυναίκες	22	640	129	662	129	730	52	9	743	48

Πίνακας 5: Αναλυτικός Πίνακας Ασθενών ανά Πάθηση

Στους παρακάτω Πίνακες, παρουσιάζονται αναλυτικά ποσοστά ασθενών βάσει του Φύλου ανά Πάθηση. Επιπλέον σε κάθε πάθηση δίνονται αναλυτικότερες πληροφορίες.

Φύλο	Πίεση			Χοληστερίνη	
	LOW	MID.	HIGH	MID.	HIGH
Άνδρες	0	0,81818182	0,181818	0,851675	0,148325
Γυναίκες	0,027813	0,8091024	0,163085	0,836915	0,163085

Πίνακας 6: Ποσοστιαία αναλογία ασθενών σχετικά με Πίεση και Χοληστερίνη

Φύλο	Σάκχαρο			Καρκίνος	
	MID.	HIGH	MID.	HIGH	MID.
Άνδρες	0,784689	0,15311	0,784689	0,15311	0,784689
Γυναίκες	0,922882	0,06574	0,922882	0,06574	0,922882

Πίνακας 7: Ποσοστιαία αναλογία ασθενών σχετικά με Σάκχαρο και Καρκίνο

4. ΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ

Στο παρόν Κεφάλαιο, πραγματοποιείται επαγωγική στατιστική ανάλυση των δεδομένων. Αρχικά, όπως πραγματοποιήθηκε στο προηγούμενο Κεφάλαιο (3^ο), θα κατηγοριοποιηθούν οι μεταβλητές των ερευνητικών ερωτημάτων.

4.1 ΕΛΕΓΧΟΣ ΚΑΝΟΝΙΚΟΤΗΤΑΣ

4.1.1 ΠΟΣΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ΗΛΙΚΙΑ	,038	1000	,002	,990	1000	,000
ΔΜΣ	,057	1000	,000	,968	1000	,000
ΥΨΟΣ	,204	1000	,000	,743	1000	,000

a. Lilliefors Significance Correction

Πίνακας 8: Έλεγχος Κατανομής

Όπως φαίνεται στον παραπάνω πίνακα, η κανονικότητα των δεδομένων σε κάθε μεταβλητή απορρίπτεται σε επίπεδο σημαντικότητας 1% επειδή τα αντίστοιχα p-values και των 2 στατιστικών ελέγχων (Kolmogorov-Smirnov και Shapiro-Wilk) είναι ίσα με 0,000 δηλαδή μικρότερα από 0,01.

Με άλλα λόγια τα δεδομένα δεν κατανέμονται σύμφωνα με την Κανονική Κατανομή, κατά συνέπεια η ανάλυση τους δύναται να πραγματοποιηθεί με μη παραμετρικές μεθόδους. Εν ολίγοις, δεν δύναται να εφαρμόσουμε ανάλυση συσχέτισης κατά Pearson, αλλά κατά Spearman.

4.2 ΑΝΑΛΥΣΗ ΣΥΣΧΕΤΙΣΗΣ

4.2.1 ΠΟΣΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ

Correlations

			ΒΑΡΟΣ	ΗΛΙΚΙΑ	ΔΜΣ	ΥΨΟΣ
Spearman's rho	ΒΑΡΟΣ	Correlation Coefficient	1,000	,124**	,913**	,465**
		Sig. (2-tailed)	.	,000	,000	,000
		N	1000	1000	1000	1000
	ΗΛΙΚΙΑ	Correlation Coefficient	,124**	1,000	,168**	-,086**
		Sig. (2-tailed)	,000	.	,000	,007
		N	1000	1000	1000	1000
	ΔΜΣ	Correlation Coefficient	,913**	,168**	1,000	,143**
		Sig. (2-tailed)	,000	,000	.	,000
		N	1000	1000	1000	1000
	ΥΨΟΣ	Correlation Coefficient	,465**	-,086**	,143**	1,000
		Sig. (2-tailed)	,000	,007	,000	.
		N	1000	1000	1000	1000

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

Πίνακας 9: Συσχέτιση Ποσοτικών μεταβλητών

Από τον παραπάνω πίνακα, στον οποίο παρουσιάζεται η συσχέτιση όλων των δυνατών συνδυασμών μεταξύ των ποσοτικών μεταβλητών μας, καταλήγουμε στα παρακάτω συμπεράσματα:

Υπάρχει θετική συσχέτιση μεταξύ των μεταβλητών ΒΑΡΟΣ και ΗΛΙΚΙΑ. Η συσχέτιση μεταξύ των 2 μεταβλητών ισούται με 0,124 και κρίνεται στατιστικά σημαντική διότι $p\text{-value} < 0,001 < 0,01$ (επίπεδο σημαντικότητας 1%). Αυτό σημαίνει ότι η αύξηση της μεταβλητής ΒΑΡΟΣ προκαλεί αύξηση της μεταβλητής ΗΛΙΚΙΑ κατά 0,124.

Υπάρχει θετική συσχέτιση μεταξύ των μεταβλητών ΒΑΡΟΣ και ΔΜΣ. Η συσχέτιση μεταξύ των 2 μεταβλητών ισούται με 0,913 και κρίνεται στατιστικά σημαντική διότι $p\text{-value} < 0,001 < 0,01$ (επίπεδο σημαντικότητας 1%). Αυτό σημαίνει ότι η αύξηση της μεταβλητής ΒΑΡΟΣ προκαλεί αύξηση της μεταβλητής ΔΜΣ κατά 0,913.

Υπάρχει θετική συσχέτιση μεταξύ των μεταβλητών ΒΑΡΟΣ και ΥΨΟΣ. Η συσχέτιση μεταξύ των 2 μεταβλητών ισούται με 0,465 και κρίνεται στατιστικά σημαντική διότι $p\text{-value} < 0,001 < 0,01$ (επίπεδο σημαντικότητας 1%). Αυτό σημαίνει ότι η αύξηση της μεταβλητής ΒΑΡΟΣ προκαλεί αύξηση της μεταβλητής ΥΨΟΣ κατά 0,465.

Υπάρχει θετική συσχέτιση μεταξύ των μεταβλητών ΗΛΙΚΙΑ και ΔΜΣ. Η συσχέτιση μεταξύ των 2 μεταβλητών ισούται με 0,168 και κρίνεται στατιστικά σημαντική διότι $p\text{-value} < 0,001 < 0,01$ (επίπεδο σημαντικότητας 1%). Αυτό σημαίνει ότι η αύξηση της μεταβλητής ΒΑΡΟΣ προκαλεί αύξηση της μεταβλητής ΥΨΟΣ κατά 0,168.

Υπάρχει αρνητική συσχέτιση μεταξύ των μεταβλητών ΗΛΙΚΙΑ και ΥΨΟΣ. Η συσχέτιση μεταξύ των 2 μεταβλητών ισούται με -0,086 και κρίνεται στατιστικά σημαντική διότι $p\text{-value} < 0,001 < 0,01$ (επίπεδο σημαντικότητας 1%). Αυτό σημαίνει ότι η αύξηση της μεταβλητής ΒΑΡΟΣ προκαλεί μείωση της μεταβλητής ΥΨΟΣ κατά 0,086.

Υπάρχει θετική συσχέτιση μεταξύ των μεταβλητών ΔΜΣ και ΥΨΟΣ. Η συσχέτιση μεταξύ των 2 μεταβλητών ισούται με 0,143 και κρίνεται στατιστικά σημαντική διότι $p\text{-value} < 0,001 < 0,01$ (επίπεδο σημαντικότητας 1%). Αυτό σημαίνει ότι η αύξηση της μεταβλητής ΔΜΣ προκαλεί αύξηση της μεταβλητής ΥΨΟΣ κατά 0,143.

4.2.2 ΠΟΙΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ

Στην παρούσα υποπαράγραφο πραγματοποιήθηκε έλεγχος ανεξαρτησίας χ^2 . Μετά από κάθε αποτέλεσμα, το οποίο παρουσιάζεται σε πίνακα, ακολουθεί η επεξήγηση του. Επιπλέον παρουσιάζονται διαγράμματα σε μορφή Ραβδόγραμμα ανά περίπτωση.

ΦΥΛΟ * ΚΑΡΚΙΝΟΣ

Crosstab

Count		ΚΑΡΚΙΝΟΣ		Total
		NO	YES	
ΦΥΛΟ	ΑΝΔΡΕΣ	203	6	209
	ΓΥΝΑΙΚΕΣ	743	48	791
Total		946	54	1000

Chi-Square Tests

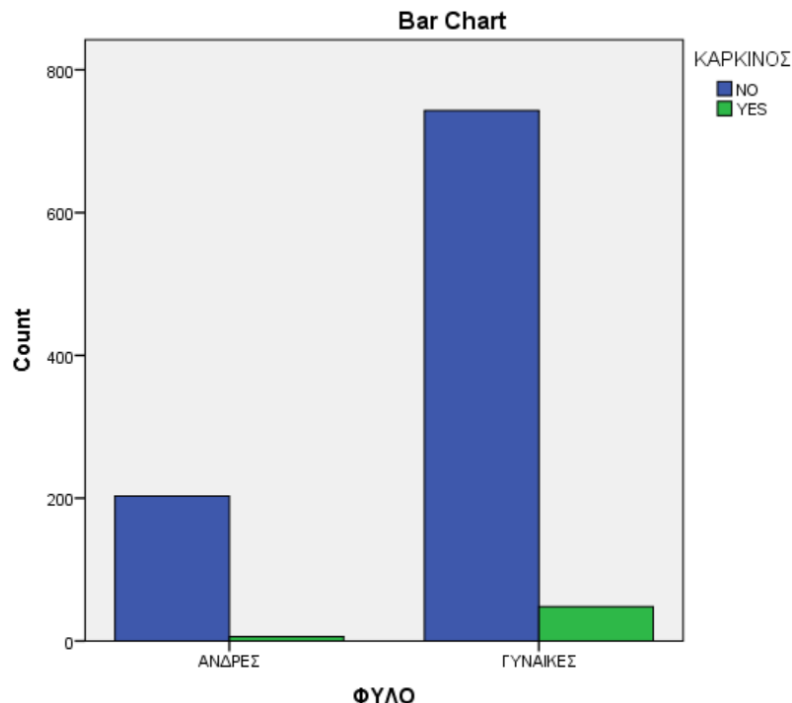
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3,309 ^a	1	,069		
Continuity Correction ^b	2,712	1	,100		
Likelihood Ratio	3,797	1	,051		
Fisher's Exact Test				,084	,043
Linear-by-Linear Association	3,305	1	,069		
N of Valid Cases	1000				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 11,29.

b. Computed only for a 2x2 table

Πίνακας 10: Συσχέτιση μεταβλητών ΦΥΛΟ και ΚΑΡΚΙΝΟΣ

Δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών ΦΥΛΟ και ΚΑΡΚΙΝΟΣ. Οι 2 μεταβλητές δεν σχετίζονται - εξαρτώνται διότι $p\text{-value}=0,069 > 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 14: Συσχέτιση μεταβλητών ΦΥΛΟ και ΚΑΡΚΙΝΟΣ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΓΥΝΑΙΚΕΣ – ΟΧΙ ΚΑΡΚΙΝΟΣ υπερτερεί των υπολοίπων. Ακολουθεί ο αντίστοιχος συνδυασμός για τους ΑΝΔΡΕΣ

ΦΥΛΟ * ΣΑΚΧΑΡΟ

Crosstab

Count		ΣΑΚΧΑΡΟ			Total
		ΚΑΝΟΝΙΚΟ	ΑΥΞΗΜΕΝΗ	ΧΡΗΣΗ ΙΝΣΟΥΛΙΝΗΣ	
ΦΥΛΟ	ΑΝΔΡΕΣ	164	32	13	209
	ΓΥΝΑΙΚΕΣ	730	52	9	791
Total		894	84	22	1000

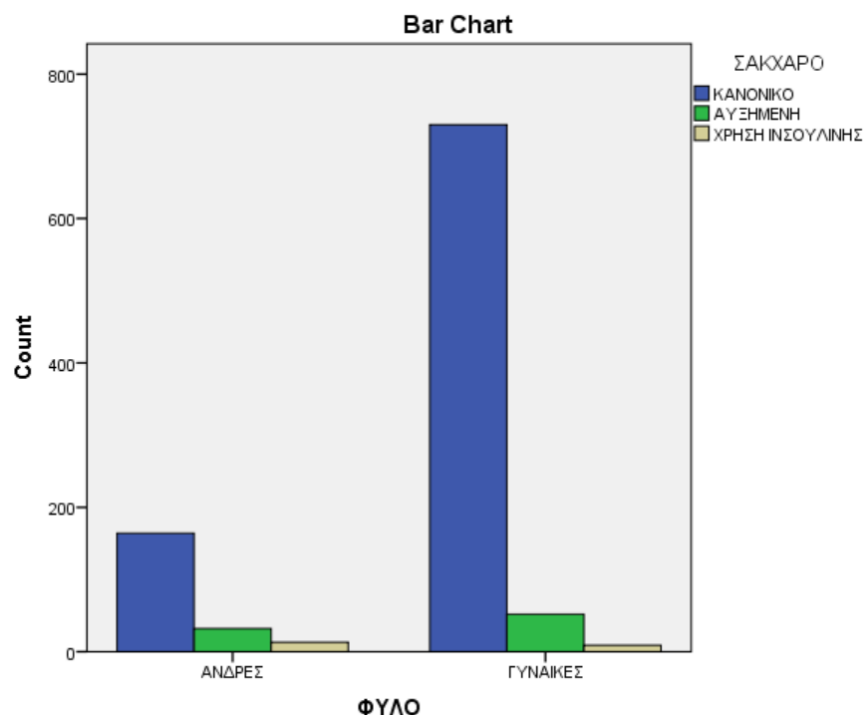
Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	37,965 ^a	2	,000
Likelihood Ratio	31,729	2	,000
Linear-by-Linear Association	37,917	1	,000
N of Valid Cases	1000		

a. 1 cells (16,7%) have expected count less than 5. The minimum expected count is 4,60.

Πίνακας 11: Συσχέτιση μεταβλητών ΦΥΛΟ και ΣΑΚΧΑΡΟ

Υπάρχει συσχέτιση μεταξύ των μεταβλητών ΦΥΛΟ και ΣΑΚΧΑΡΟ. Οι 2 μεταβλητές σχετίζονται – εξαρτώνται ισχυρά διότι $p\text{-value} < 0,001 < 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 15: Συσχέτιση μεταβλητών ΦΥΛΟ και ΣΑΚΧΑΡΟ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΓΥΝΑΙΚΕΣ – ΚΑΝΟΝΙΚΟ ΣΑΚΧΑΡΟ υπερτερεί των υπολοίπων. Ακολουθεί ο αντίστοιχος συνδυασμός για τους ΑΝΔΡΕΣ.

ΦΥΛΟ * ΧΟΛΗΣΤΕΡΙΝΗ

Crosstab

Count		ΧΟΛΗΣΤΕΡΙΝΗ		Total
		ΚΑΝΟΝΙΚΟ	ΑΥΞΗΜΕΝΟ	
ΦΥΛΟ	ΑΝΔΡΕΣ	178	31	209
	ΓΥΝΑΙΚΕΣ	662	129	791
Total		840	160	1000

Chi-Square Tests

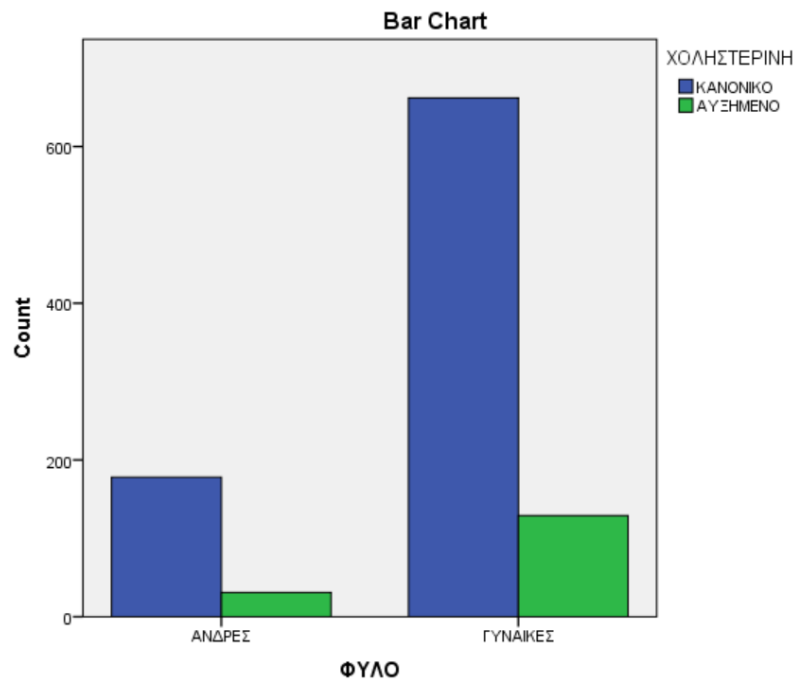
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,268 ^a	1	,605		
Continuity Correction ^b	,169	1	,681		
Likelihood Ratio	,272	1	,602		
Fisher's Exact Test				,672	,345
Linear-by-Linear Association	,268	1	,605		
N of Valid Cases	1000				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 33,44.

b. Computed only for a 2x2 table

Πίνακας 12: Συσχέτιση μεταβλητών ΦΥΛΟ και ΧΟΛΗΣΤΕΡΙΝΗ

Δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών ΦΥΛΟ και ΧΟΛΗΣΤΕΡΙΝΗ. Οι 2 μεταβλητές δεν σχετίζονται - εξαρτώνται διότι $p\text{-value} = 0,605 > 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 16: Συσχέτιση μεταβλητών ΦΥΛΟ και ΧΟΛΗΣΤΕΡΙΝΗ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΓΥΝΑΙΚΕΣ – ΚΑΝΟΝΙΚΗ ΧΟΛΗΣΤΕΡΙΝΗ υπερτερεί των υπολοίπων. Ακολουθεί ο αντίστοιχος συνδυασμός για τους ΑΝΔΡΕΣ.

ΦΥΛΟ * ΠΙΕΣΗ

Crosstab

Count		ΠΙΕΣΗ			Total
		ΚΑΝΟΝΙΚΗ	ΥΠΕΡΤΑΣΗ	ΥΠΟΤΑΣΗ	
ΦΥΛΟ	ΑΝΔΡΕΣ	171	38	0	209
	ΓΥΝΑΙΚΕΣ	640	129	22	791
Total		811	167	22	1000

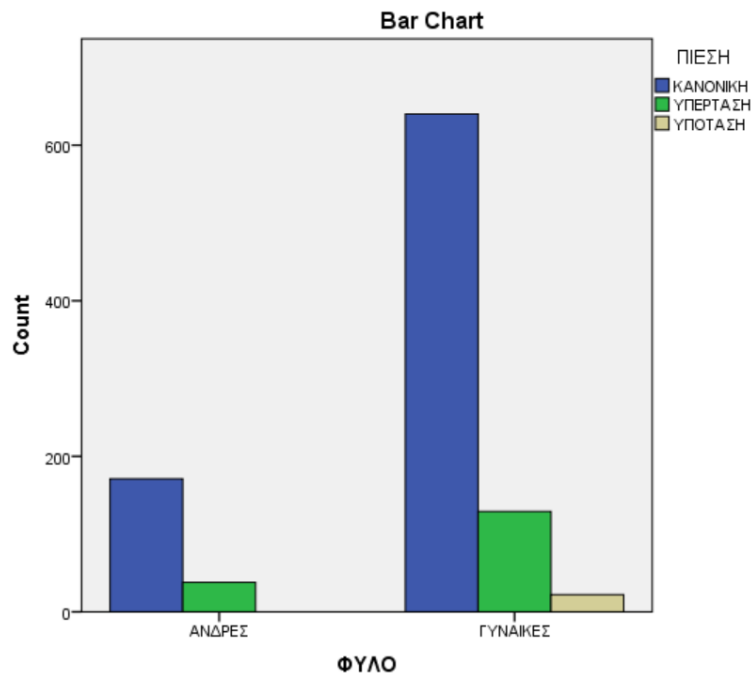
Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6,177 ^a	2	,046
Likelihood Ratio	10,673	2	,005
Linear-by-Linear Association	1,068	1	,301
N of Valid Cases	1000		

a. 1 cells (16,7%) have expected count less than 5. The minimum expected count is 4,60.

Πίνακας 13: Συσχέτιση μεταβλητών ΦΥΛΟ και ΠΙΕΣΗ

Υπάρχει συσχέτιση μεταξύ των μεταβλητών ΦΥΛΟ και ΠΙΕΣΗ. Οι 2 μεταβλητές σχετίζονται – εξαρτώνται ισχυρά διότι $p\text{-value} = 0,046 < 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 17: Συσχέτιση μεταβλητών ΦΥΛΟ και ΠΙΕΣΗ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΓΥΝΑΙΚΕΣ – ΚΑΝΟΝΙΚΗ ΠΙΕΣΗ υπερτερεί των υπολοίπων. Ακολουθεί ο αντίστοιχος συνδυασμός για τους ΑΝΔΡΕΣ.

ΚΑΡΚΙΝΟΣ * ΣΑΚΧΑΡΟ

Crosstab

Count		ΣΑΚΧΑΡΟ			Total
		ΚΑΝΟΝΙΚΟ	ΑΥΞΗΜΕΝΗ	ΧΡΗΣΗ ΙΝΣΟΥΛΙΝΗΣ	
ΚΑΡΚΙΝΟΣ	NO	844	80	22	946
	YES	50	4	0	54
Total		894	84	22	1000

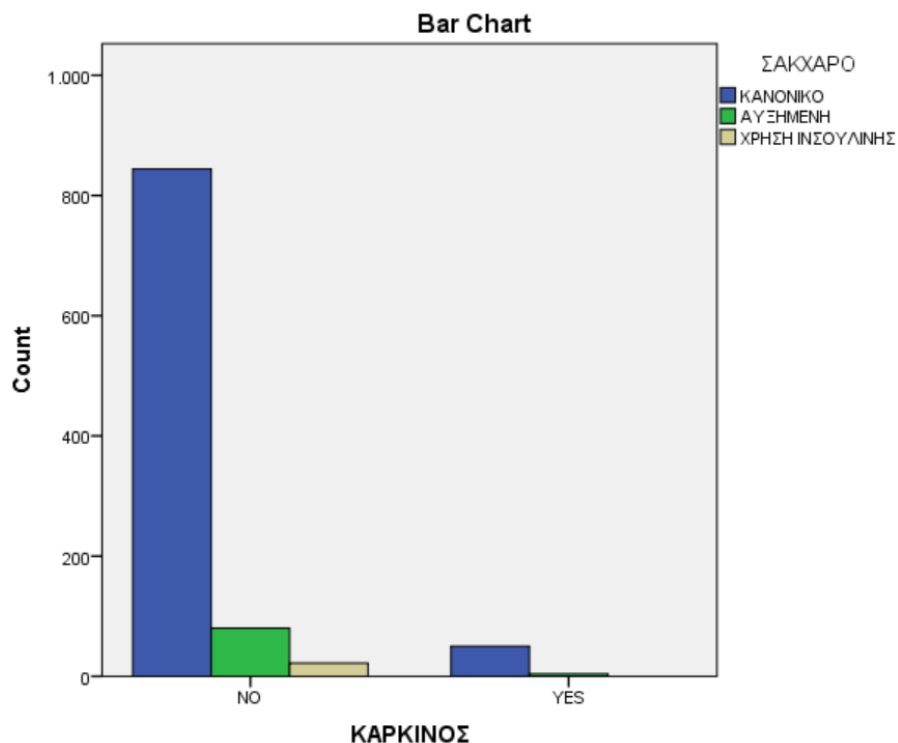
Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1,388 ^a	2	,500
Likelihood Ratio	2,577	2	,276
Linear-by-Linear Association	1,066	1	,302
N of Valid Cases	1000		

a. 2 cells (33,3%) have expected count less than 5. The minimum expected count is 1,19.

Πίνακας 14: Συσχέτιση μεταβλητών ΚΑΡΚΙΝΟΣ και ΣΑΚΧΑΡΟ

Δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών ΚΑΡΚΙΝΟΣ και ΣΑΚΧΑΡΟ. Οι 2 μεταβλητές δεν σχετίζονται - εξαρτώνται διότι $p\text{-value} = 0,500 > 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 18: Συσχέτιση μεταβλητών ΚΑΡΚΙΝΟΣ και ΣΑΚΧΑΡΟ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΚΑΝΟΝΙΚΟ ΣΑΚΧΑΡΟ – ΟΧΙ ΚΑΡΚΙΝΟΣ υπερτερεί των υπολοίπων. Ακολουθεί ο συνδυασμός ΑΥΞΗΜΕΝΟ ΣΑΚΧΑΡΟ – ΟΧΙ ΚΑΡΚΙΝΟΣ.

ΚΑΡΚΙΝΟΣ * ΧΟΛΗΣΤΕΡΙΝΗ

Crosstab

Count		ΧΟΛΗΣΤΕΡΙΝΗ		Total
		ΚΑΝΟΝΙΚΟ	ΑΥΞΗΜΕΝΟ	
ΚΑΡΚΙΝΟΣ	NO	798	148	946
	YES	42	12	54
Total		840	160	1000

Chi-Square Tests

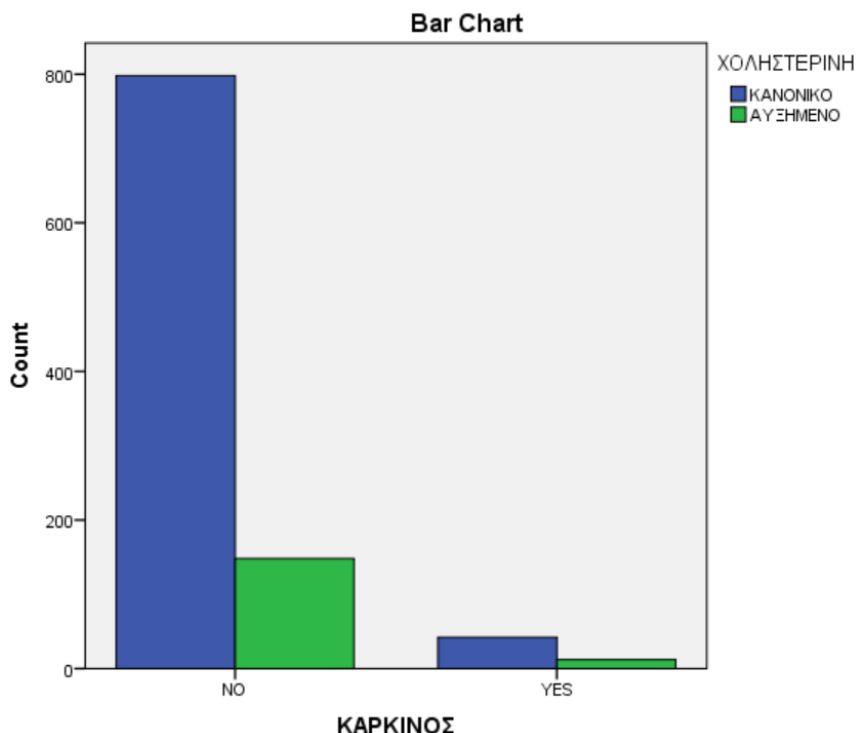
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,644 ^a	1	,200		
Continuity Correction ^b	1,191	1	,275		
Likelihood Ratio	1,509	1	,219		
Fisher's Exact Test				,249	,138
Linear-by-Linear Association	1,643	1	,200		
N of Valid Cases	1000				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 8,64.

b. Computed only for a 2x2 table

Πίνακας 15: Συσχέτιση μεταβλητών ΚΑΡΚΙΝΟΣ και ΧΟΛΗΣΤΕΡΙΝΗ

Δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών ΚΑΡΚΙΝΟΣ και ΧΟΛΗΣΤΕΡΙΝΗ. Οι 2 μεταβλητές δεν σχετίζονται - εξαρτώνται διότι $p\text{-value} = 0,200 > 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 19: Συσχέτιση μεταβλητών ΚΑΡΚΙΝΟΣ και ΧΟΛΗΣΤΕΡΙΝΗ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΚΑΝΟΝΙΚΗ ΧΟΛΗΣΤΕΡΙΝΗ – ΟΧΙ ΚΑΡΚΙΝΟΣ υπερτερεί των υπολοίπων. Ακολουθεί ο συνδυασμός ΑΥΞΗΜΕΝΗ ΧΟΛΗΣΤΕΡΙΝΗ – ΟΧΙ ΚΑΡΚΙΝΟΣ.

ΚΑΡΚΙΝΟΣ * ΠΙΕΣΗ

Crosstab

Count		ΠΙΕΣΗ			Total
		ΚΑΝΟΝΙΚΗ	ΥΠΕΡΤΑΣΗ	ΥΠΟΤΑΣΗ	
ΚΑΡΚΙΝΟΣ	NO	771	154	21	946
	YES	40	13	1	54
Total		811	167	22	1000

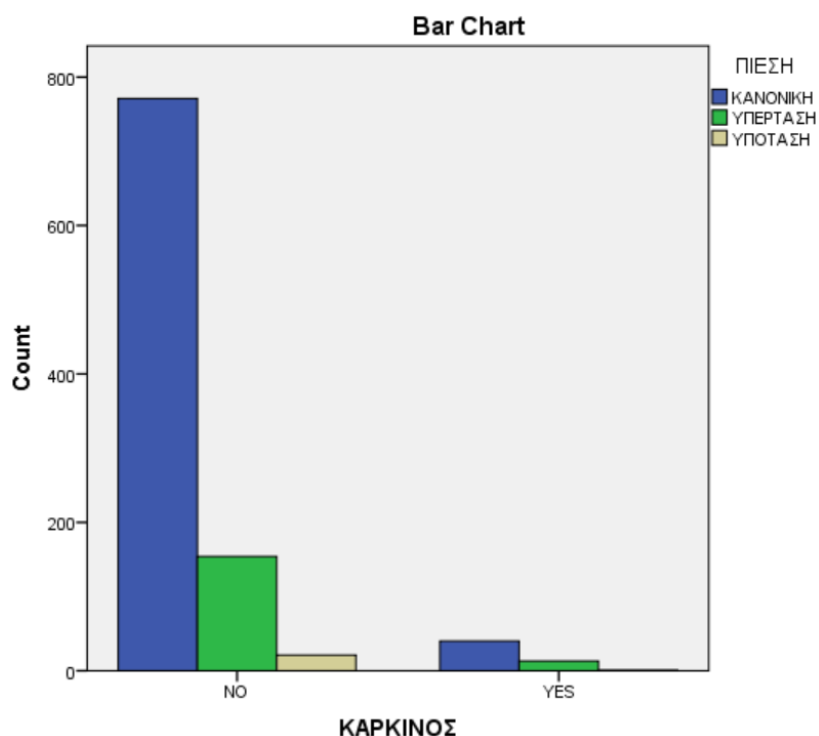
Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2,238 ^a	2	,327
Likelihood Ratio	2,037	2	,361
Linear-by-Linear Association	1,208	1	,272
N of Valid Cases	1000		

a. 1 cells (16,7%) have expected count less than 5. The minimum expected count is 1,19.

Πίνακας 16: Συσχέτιση μεταβλητών ΚΑΡΚΙΝΟΣ και ΠΙΕΣΗ

Δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών ΚΑΡΚΙΝΟΣ και ΠΙΕΣΗ. Οι 2 μεταβλητές δεν σχετίζονται - εξαρτώνται διότι $p\text{-value} = 0,327 > 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 20: Συσχέτιση μεταβλητών ΚΑΡΚΙΝΟΣ και ΠΙΕΣΗ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΚΑΝΟΝΙΚΗ ΠΙΕΣΗ – ΟΧΙ ΚΑΡΚΙΝΟΣ υπερτερεί των υπολοίπων. Ακολουθεί ο συνδυασμός ΑΥΞΗΜΕΝΗ ΠΙΕΣΗ – ΟΧΙ ΚΑΡΚΙΝΟΣ.

ΣΑΚΧΑΡΟ * ΧΟΛΗΣΤΕΡΙΝΗ

Crosstab

Count		ΧΟΛΗΣΤΕΡΙΝΗ		Total
		ΚΑΝΟΝΙΚΟ	ΑΥΞΗΜΕΝΟ	
ΣΑΚΧΑΡΟ	ΚΑΝΟΝΙΚΟ	769	125	894
	ΑΥΞΗΜΕΝΗ	50	34	84
	ΧΡΗΣΗ ΙΝΣΟΥΛΙΝΗΣ	21	1	22
Total		840	160	1000

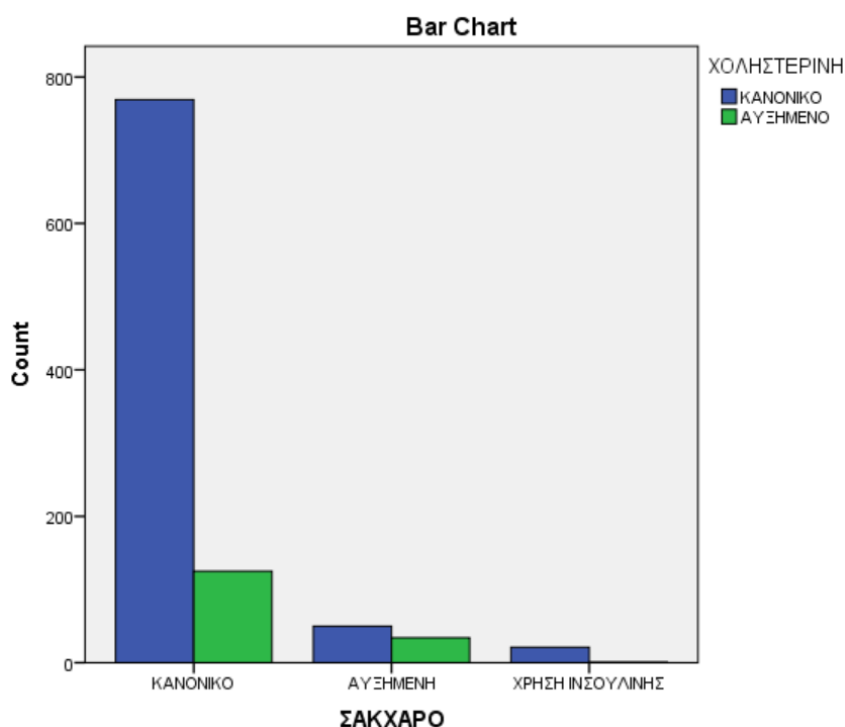
Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	42,299 ^a	2	,000
Likelihood Ratio	34,328	2	,000
Linear-by-Linear Association	11,505	1	,001
N of Valid Cases	1000		

a. 1 cells (16,7%) have expected count less than 5. The minimum expected count is 3,52.

Πίνακας 17: Συσχέτιση μεταβλητών ΣΑΚΧΑΡΟ και ΧΟΛΗΣΤΕΡΙΝΗ

Υπάρχει συσχέτιση μεταξύ των μεταβλητών ΣΑΚΧΑΡΟ και ΧΟΛΗΣΤΕΡΙΝΗ. Οι 2 μεταβλητές σχετίζονται – εξαρτώνται ισχυρά διότι $p\text{-value} < 0,001 < 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 21: Συσχέτιση μεταβλητών ΣΑΚΧΑΡΟ και ΧΟΛΗΣΤΕΡΙΝΗ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΚΑΝΟΝΙΚΟ ΣΑΚΧΑΡΟ - ΚΑΝΟΝΙΚΗ ΧΟΛΗΣΤΕΡΙΝΗ υπερτερεί των υπολοίπων. Ακολουθεί ο συνδυασμός ΚΑΝΟΝΙΚΟ ΣΑΚΧΑΡΟ - ΑΥΞΗΜΕΝΗ ΧΟΛΗΣΤΕΡΙΝΗ.

ΣΑΚΧΑΡΟ * ΠΙΕΣΗ

Crosstab

Count		ΠΙΕΣΗ			Total
		ΚΑΝΟΝΙΚΗ	ΥΠΕΡΤΑΣΗ	ΥΠΟΤΑΣΗ	
ΣΑΚΧΑΡΟ	ΚΑΝΟΝΙΚΟ	728	144	22	894
	ΑΥΞΗΜΕΝΗ	64	20	0	84
	ΧΡΗΣΗ ΙΝΣΟΥΛΙΝΗΣ	19	3	0	22
Total		811	167	22	1000

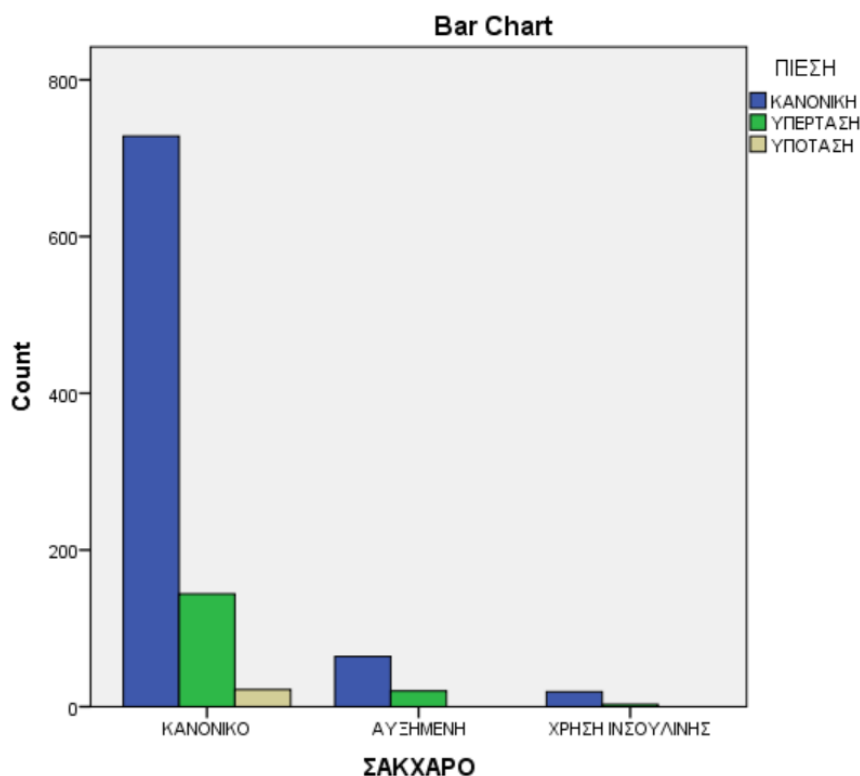
Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5,799 ^a	4	,215
Likelihood Ratio	7,836	4	,098
Linear-by-Linear Association	,031	1	,860
N of Valid Cases	1000		

a. 3 cells (33,3%) have expected count less than 5. The minimum expected count is ,48.

Πίνακας 18: Συσχέτιση μεταβλητών ΣΑΚΧΑΡΟ και ΠΙΕΣΗ

Δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών ΣΑΚΧΑΡΟ και ΠΙΕΣΗ. Οι 2 μεταβλητές δεν σχετίζονται - εξαρτώνται διότι $p\text{-value}=0,215 > 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 22: Συσχέτιση μεταβλητών ΣΑΚΧΑΡΟ και ΠΙΕΣΗ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΚΑΝΟΝΙΚΟ ΣΑΚΧΑΡΟ - ΚΑΝΟΝΙΚΗ ΠΙΕΣΗ υπερτερεί των υπολοίπων. Ακολουθεί ο συνδυασμός ΚΑΝΟΝΙΚΟ ΣΑΚΧΑΡΟ - ΥΠΕΡΤΑΣΗ.

ΧΟΛΗΣΤΕΡΙΝΗ * ΠΙΕΣΗ Crosstabulation

Count		ΠΙΕΣΗ			Total
		ΚΑΝΟΝΙΚΗ	ΥΠΕΡΤΑΣΗ	ΥΠΟΤΑΣΗ	
ΧΟΛΗΣΤΕΡΙΝΗ	ΚΑΝΟΝΙΚΟ	698	122	20	840
	ΑΥΞΗΜΕΝΟ	113	45	2	160
Total		811	167	22	1000

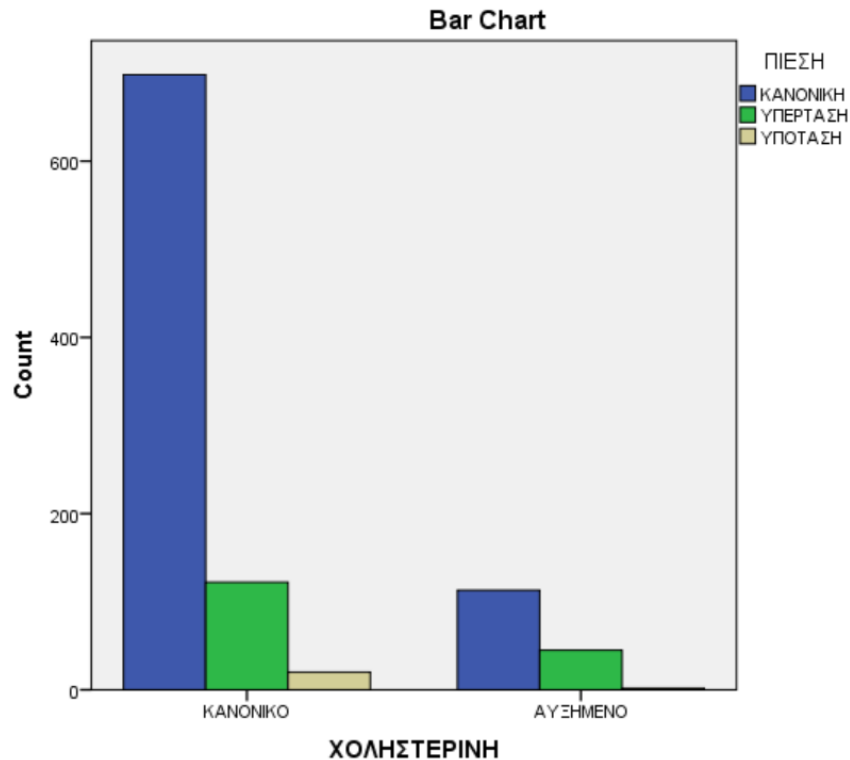
Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	18,246 ^a	2	,000
Likelihood Ratio	16,419	2	,000
Linear-by-Linear Association	8,202	1	,004
N of Valid Cases	1000		

a. 1 cells (16,7%) have expected count less than 5. The minimum expected count is 3,52.

Πίνακας 19: Συσχέτιση μεταβλητών ΧΟΛΗΣΤΕΡΙΝΗ και ΠΙΕΣΗ

Υπάρχει συσχέτιση μεταξύ των μεταβλητών ΧΟΛΗΣΤΕΡΙΝΗ και ΠΙΕΣΗ. Οι 2 μεταβλητές σχετίζονται – εξαρτώνται ισχυρά διότι $p\text{-value} < 0,001 < 0,05$ (επίπεδο σημαντικότητας 5%).



Διάγραμμα 23: Συσχέτιση μεταβλητών ΧΟΛΗΣΤΕΡΙΝΗ και ΠΙΕΣΗ

Στο παραπάνω διάγραμμα παρουσιάζονται το μέγεθος ανά συνδυασμού. Συγκεκριμένα ο αριθμός των παρατηρήσεων του συνδυασμού ΚΑΝΟΝΙΚΗ ΧΟΛΗΣΤΕΡΙΝΗ - ΚΑΝΟΝΙΚΗ ΠΙΕΣΗ υπερτερεί των υπολοίπων. Ακολουθεί ο συνδυασμός ΚΑΝΟΝΙΚΗ ΧΟΛΗΣΤΕΡΙΝΗ - ΥΠΕΡΤΑΣΗ.

4.3 ΜΟΝΤΕΛΑ ΠΡΟΒΛΕΨΗΣ

Στην παρούσα υποπαράγραφο παρουσιάζονται τα μοντέλα πρόβλεψης ανά πάθηση – ασθένεια. Συγκεκριμένα, λόγω του είδους της μεταβλητής στόχου (ΔΜΣ) πραγματοποιείται Ανάλυση Γραμμικής Παλινδρόμηση. Επιπλέον θεωρείται σημαντική η ανάπτυξη μοντέλων πρόβλεψης της μεταβλητής ΔΜΣ βάσει της μεταβλητής ΗΛΙΚΙΑ.

4.3.1 ΠΟΣΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	ΒΑΡΟΣ, ΗΛΙΚΙΑ, ΥΨΟΣ ^b	.	Enter

a. Dependent Variable: ΔΜΣ

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,939 ^a	,882	,881	2,984596160

a. Predictors: (Constant), ΒΑΡΟΣ, ΗΛΙΚΙΑ, ΥΨΟΣ

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	66208,210	3	22069,403	2477,533	,000 ^b
	Residual	8872,183	996	8,908		
	Total	75080,393	999			

a. Dependent Variable: ΔΜΣ

b. Predictors: (Constant), ΒΑΡΟΣ, ΗΛΙΚΙΑ, ΥΨΟΣ

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	104,284	1,815		57,446	,000
	ΗΛΙΚΙΑ	-,001	,007	-,002	-,176	,860
	ΥΨΟΣ	-64,255	1,136	-,676	-56,542	,000
	ΒΑΡΟΣ	,393	,005	,978	81,520	,000

a. Dependent Variable: ΔΜΣ

Πίνακας 20: Μοντέλο Πρόβλεψης μεταβλητής ΔΜΣ

Από τον παραπάνω πίνακα παρατηρούμε ότι η μεταβλητή ΗΛΙΚΙΑ δεν είναι στατιστικώς σημαντική. Θα πραγματοποιηθεί εκ νέου Ανάλυση Γραμμικής Παλινδρόμησης με τις ανεξάρτητες μεταβλητές ΒΑΡΟΣ και ΥΨΟΣ.

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	ΒΑΡΟΣ, ΥΨΟΣ ^b	.	Enter

a. Dependent Variable: ΔΜΣ

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,939 ^a	,882	,882	2,983145312

a. Predictors: (Constant), ΒΑΡΟΣ, ΥΨΟΣ

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	66207,934	2	33103,967	3719,900	,000 ^b
	Residual	8872,458	997	8,899		
	Total	75080,393	999			

a. Dependent Variable: ΔΜΣ

b. Predictors: (Constant), ΒΑΡΟΣ, ΥΨΟΣ

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	104,205	1,759		59,242	,000
	ΥΨΟΣ	,393	,005	,977	82,229	,000

a. Dependent Variable: ΔΜΣ

Πίνακας 21: Μοντέλο Πρόβλεψης μεταβλητής ΔΜΣ

Σύμφωνα με τον παραπάνω πίνακα, μετά την εισαγωγή των μεταβλητών ΥΨΟΣ και ΒΑΡΟΣ, η ζητούμενη γραμμική εξίσωση της εξαρτημένης μεταβλητής ΔΜΣ είναι:

$$y = 0,393x_1 + 104,205$$

$$x_1 : \text{ΥΨΟΣ}$$

Μοντέλο Πρόβλεψης 1: Μεταβλητή στόχος ΔΜΣ

4.3.2 ΠΟΙΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ

Στην παρούσα υποπαράγραφο παρουσιάζονται τα μοντέλα πρόβλεψης ανά πάθηση – ασθένεια. Συγκεκριμένα, λόγω του είδους της μεταβλητής στόχου πραγματοποιείται Δίτιμη Παλινδρόμηση (Binary Regression).

ΚΑΡΚΙΝΟΣ

Πραγματοποιείται Δίτιμη Παλινδρόμηση διότι η μεταβλητή ΚΑΡΚΙΝΟΣ «παίρνει» 2 τιμές.

Block 0: Beginning Block

Classification Table^{a,b}

Observed		Predicted		
		ΚΑΡΚΙΝΟΣ		Percentage Correct
		NO	YES	
Step 0	ΚΑΡΚΙΝΟΣ NO	946	0	100,0
	YES	54	0	,0
Overall Percentage				94,6

a. Constant is included in the model.

b. The cut value is ,500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-2,863	,140	418,799	1	,000	,057

Variables not in the Equation

	Score	df	Sig.
Step 0 Variables ΦΥΛΟ(1)	12,858	1	,000
ΔΜΣ	,765	1	,382
ΥΨΟΣ	2,231	1	,135
ΧΟΛΗΣΤΕΡΙΝΗ(1)	1,644	1	,200
Overall Statistics	24,730	10	,006

Πίνακας 22: Μοντέλο Πρόβλεψης μεταβλητής ΚΑΡΚΙΝΟΣ

Το Block 0 λαμβάνει υπόψη μόνο τη σταθερά του μοντέλου μας, οπότε τα αποτελέσματα που μας δίνει, είναι το ποσοστό κάποιος ασθενής να έχει οποιαδήποτε μορφή Καρκίνου ($946/1000 = 94,6\%$ δεν έχει τη νόσο και $5,4\%$ έχει τη νόσο). Χρησιμοποιώντας μόνο τη σταθερά, η πρόβλεψή μας θα ήταν σωστή $94,6\%$.

Επιπλέον μας δίνεται πληροφορία σχετικά (πίνακα Variables in the Equation) με το αναμενόμενο odd. Συγκεκριμένα, $\ln(odds) = -2,863$ και το αναμενόμενο odd ισούται με : $Exp(B_0) = e^{-2,863} = 0,057$.

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	28,056	10	,002
	Block	28,056	10	,002
	Model	28,056	10	,002

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	392,202 ^a	,028	,081

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	11,436	8	,178

Πίνακας 23: Μοντέλο Πρόβλεψης μεταβλητής ΚΑΡΚΙΝΟΣ

Ο έλεγχος των Hosmer-Lemeshow είναι ένα στατιστικό τεστ καλής προσαρμογής για μοντέλα λογιστικής παλινδρόμησης. Συγκεκριμένα, το μοντέλο είναι καλά προσαρμοσμένο διότι $p\text{-value}=0,178 > 0,05$.

Classification Table^a

	Observed	Predicted			
		ΚΑΡΚΙΝΟΣ		Percentage Correct	
		NO	YES		
Step 1	ΚΑΡΚΙΝΟΣ	NO	946	0	100,0
		YES	54	0	,0
	Overall Percentage				94,6

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	ΦΥΛΟ(1)	,049	,012	16,627	1	,000	1,050	1,026	1,075
	ΔΜΣ	-35,405	8302,127	,000	1	,997	,000	,819	1,196
	ΥΨΟΣ	,328	1,081	,092	1	,761	1,388	,167	11,552
	ΧΟΛΗΣΤΕΡΙΝΗ(1)	-,057	,358	,025	1	,874	,945	,468	1,907

a. Variable(s) entered on step 1: ΦΥΛΟ, ΗΛΙΚΙΑ, ΔΜΣ, ΥΨΟΣ, ΒΑΡΟΣ, ΣΑΚΧΑΡΟ, ΧΟΛΗΣΤΕΡΙΝΗ, ΠΙΕΣΗ.

Πίνακας 24: Μοντέλο Πρόβλεψης μεταβλητής ΚΑΡΚΙΝΟΣ

Στο Block 1 και συγκεκριμένα στον παραπάνω πίνακα, παρουσιάζονται οι μεταβλητές με τα αντίστοιχα p-value (στήλη Sig.) Παρατηρούμε ότι τα p-value των μεταβλητών ΔΜΣ, ΥΨΟΣ και ΧΟΛΗΣΤΕΡΙΝΗ είναι μεγαλύτερα του 0,05 (επίπεδο σημαντικότητας 5%). Αυτό σημαίνει ότι οι ανεξάρτητες μεταβλητές δεν παρουσιάζουν στατιστική σημαντικότητα και δεν

διαδραματίζουν στατιστικά σημαντικό ρόλο στο μοντέλο. Αντιθέτως το p-value της μεταβλητής ΦΥΛΟ είναι μικρότερο του 0,05 (επίπεδο σημαντικότητας 5%), δηλαδή η συγκεκριμένη μεταβλητή διαδραματίζει στατιστικά σημαντικό ρόλο στο μοντέλο. Σύμφωνα με τα παραπάνω, η παραπάνω διαδικασία πρέπει να πραγματοποιηθεί εκ νέου εισάγοντας μόνο τη μεταβλητή ΦΥΛΟ, επειδή είναι η μόνη μεταβλητή που διαδραματίζει στατιστικά σημαντικό ρόλο στο μοντέλο.

Block 0: Beginning Block

Classification Table^{a,b}

Observed		Predicted		
		ΚΑΡΚΙΝΟΣ		Percentage Correct
		NO	YES	
Step 0	ΚΑΡΚΙΝΟΣ NO	946	0	100,0
	YES	54	0	,0
Overall Percentage				94,6

a. Constant is included in the model.

b. The cut value is ,500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-2,863	,140	418,799	1	,000	,057

Variables not in the Equation

	Score	df	Sig.
Step 0 Variables ΦΥΛΟ(1)	3,309	1	,069
Overall Statistics	3,309	1	,069

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

	Chi-square	df	Sig.
Step 1 Step	3,797	1	,051
Block	3,797	1	,051
Model	3,797	1	,051

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	416,461 ^a	,004	,011

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Πίνακας 25: Μοντέλο Πρόβλεψης μεταβλητής ΚΑΡΚΙΝΟΣ

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	,000	0	.

Contingency Table for Hosmer and Lemeshow Test

		ΚΑΡΚΙΝΟΣ= NO		ΚΑΡΚΙΝΟΣ= YES		Total
		Observed	Expected	Observed	Expected	
Step 1	1	203	203,000	6	6,000	209
	2	743	743,000	48	48,000	791

Classification Table^a

Observed	ΚΑΡΚΙΝΟΣ	Predicted		Percentage Correct
		ΚΑΡΚΙΝΟΣ		
		NO	YES	
Step 1	ΚΑΡΚΙΝΟΣ NO	946	0	100,0
	YES	54	0	,0
Overall Percentage				94,6

a. The cut value is ,500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a ΦΥΛΟ(1)	-2,739	,149	338,372	1	,000	,065

a. Variable(s) entered on step 1: ΦΥΛΟ.

Πίνακας 26: Μοντέλο Πρόβλεψης μεταβλητής ΚΑΡΚΙΝΟΣ

Από τους παραπάνω πίνακες προκύπτουν τα εξής:

- Στο Block 0 δεν υπάρχει κάποια ουσιαστική διαφορά.
- Στο Block 1 (έλεγχος των Hosmer-Lemeshow) παρατηρείται ότι το μοντέλο δεν είναι καλά προσαρμοσμένο διότι $p\text{-value} < 0,05$. Επιπλέον το $p\text{-value}$ της μεταβλητής ΦΥΛΟ είναι μικρότερο του 0,05 (επίπεδο σημαντικότητας 5%), δηλαδή η συγκεκριμένη μεταβλητή διαδραματίζει στατιστικά σημαντικό ρόλο στο μοντέλο.

Η ζητούμενη εξίσωση που έχει το μοντέλο μας θα ήταν η παρακάτω:

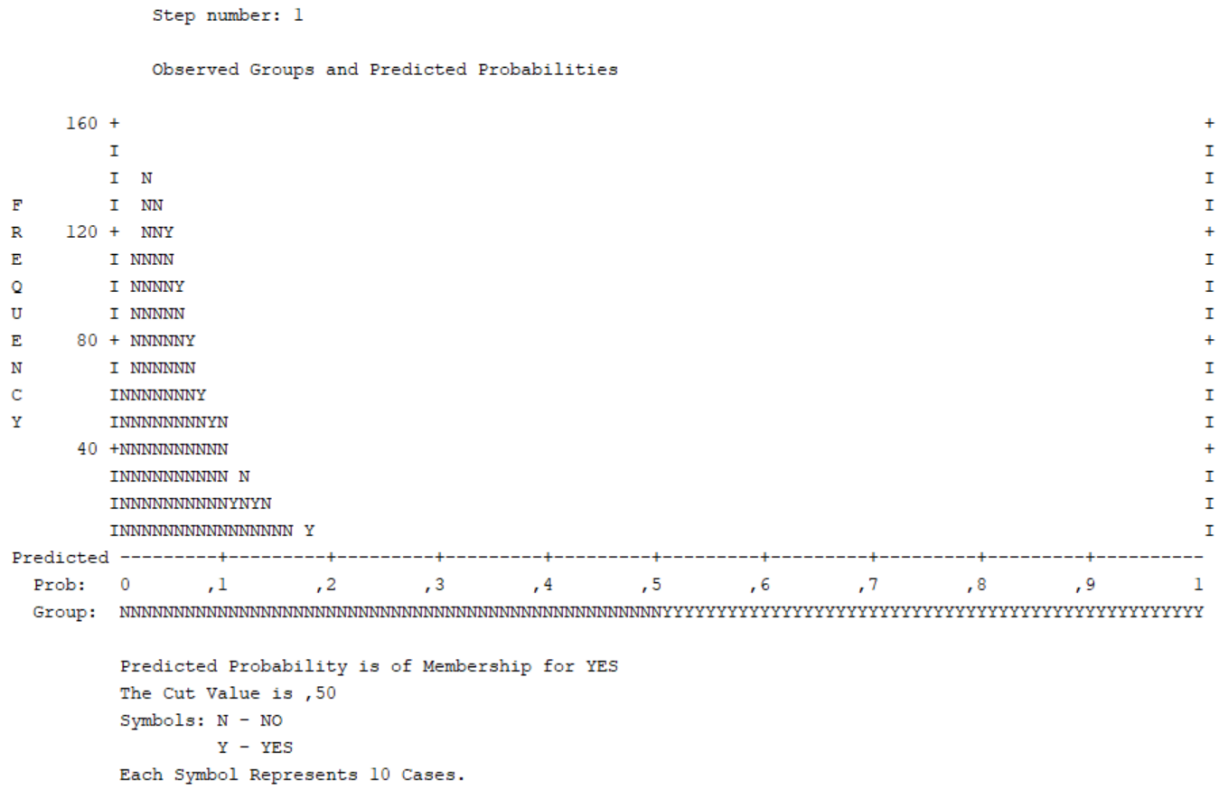
$$f(z) = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}}$$

$$z = -2,739x_1$$

$$x_1 : \text{ΦΥΛΟ}$$

Μοντέλο Πρόβλεψης 2: Μεταβλητή στόχος ΚΑΡΚΙΝΟΣ

Ιστόγραμμα Εκτιμώμενων Πιθανοτήτων



Διάγραμμα 24: Ιστόγραμμα Εκτιμώμενων Πιθανοτήτων μεταβλητής ΚΑΡΚΙΝΟΣ

Στο παραπάνω διάγραμμα ο οριζόντιος άξονας του διαγράμματος αντιστοιχεί στις εκτιμώμενες από το μοντέλο πιθανότητες. Όλες οι παρατηρήσεις είναι συγκεντρωμένες μεταξύ 0-0,2. Αυτό σημαίνει ότι όλες οι παρατηρήσεις έχουν μικρή πιθανότητα υλοποίησης.

ΧΟΛΗΣΤΕΡΙΝΗ

Πραγματοποιείται Δίτιμη Παλινδρόμηση διότι η μεταβλητή ΧΟΛΗΣΤΕΡΙΝΗ παίρνει 2 τιμές.

Block 0: Beginning Block

Classification Table^{a,b}

Observed		Predicted			
		ΧΟΛΗΣΤΕΡΙΝΗ		Percentage Correct	
		ΚΑΝΟΝΙΚΟ	ΑΥΞΗΜΕΝΟ		
Step 0	ΧΟΛΗΣΤΕΡΙΝΗ	ΚΑΝΟΝΙΚΟ	840	0	100,0
		ΑΥΞΗΜΕΝΟ	160	0	,0
Overall Percentage					84,0

a. Constant is included in the model.

b. The cut value is ,500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	
Step 0	Constant	-1,658	,086	369,562	1	,000	,190

Variables not in the Equation

	Score	df	Sig.	
Step 0	Variables			
	ΦΥΛΟ(1)	63,961	1	,000
	ΔΜΣ	,833	1	,362
	ΥΨΟΣ	17,873	1	,000
	Overall Statistics	101,660	10	,000

Πίνακας 27: Μοντέλο Πρόβλεψης μεταβλητής ΧΟΛΗΣΤΕΡΙΝΗ

Το Block 0 λαμβάνει υπόψη μόνο τη σταθερά του μοντέλου μας, οπότε τα αποτελέσματα που μας δίνει, είναι το ποσοστό κάποιος ασθενής να έχει αυξημένα ποσοστά Χοληστερίνης ($840/1000 = 84\%$ έχει κανονικά και $5,4\%$ έχει αυξημένα). Χρησιμοποιώντας μόνο τη σταθερά, η πρόβλεψή μας θα ήταν σωστή 84% .

Επιπλέον μας δίνεται πληροφορία σχετικά (πίνακα Variables in the Equation) με το αναμενόμενο odd. Συγκεκριμένα, $\ln(odds) = -1,658$ και το αναμενόμενο odd ισούται με : $Exp(B_0) = e^{-1,658} = 0,190$.

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	99,810	10	,000
	Block	99,810	10	,000
	Model	99,810	10	,000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	779,530 ^a	,095	,162

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	19,834	8	,011

Πίνακας 28: Μοντέλο Πρόβλεψης μεταβλητής ΧΟΛΗΣΤΕΡΙΝΗ

Σύμφωνα με τον έλεγχο των Hosmer-Lemeshow, το μοντέλο δεν είναι καλά προσαρμοσμένο διότι $p\text{-value}=0,011 < 0,05$.

Classification Table^a

Observed		Predicted			
		ΧΟΛΗΣΤΕΡΙΝΗ		Percentage Correct	
		ΚΑΝΟΝΙΚΟ	ΑΥΣΗΜΕΝΟ		
Step 1	ΧΟΛΗΣΤΕΡΙΝΗ	ΚΑΝΟΝΙΚΟ	827	13	98,5
		ΑΥΣΗΜΕΝΟ	147	13	8,1
Overall Percentage					84,0

a. The cut value is ,500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)		
							Lower	Upper	
Step 1 ^a	ΦΥΛΟ(1)	,051	,008	42,437	1	,000	1,052	1,036	1,068
	ΔΜΣ	-12,180	6,169	3,899	1	,048	,000	,954	1,016
	ΥΨΟΣ	,764	,783	,953	1	,329	2,147	,463	9,952

a. Variable(s) entered on step 1: ΦΥΛΟ, ΗΛΙΚΙΑ, ΔΜΣ, ΥΨΟΣ, ΒΑΡΟΣ, ΚΑΡΚΙΝΟΣ, ΣΑΚΧΑΡΟ, ΠΙΕΣΗ.

Πίνακας 29: Μοντέλο Πρόβλεψης μεταβλητής ΧΟΛΗΣΤΕΡΙΝΗ

Παρατηρούμε ότι τα $p\text{-value}$ των μεταβλητών ΔΜΣ και ΦΥΛΟ είναι μικρότερα του 0,05 (επίπεδο σημαντικότητας 5%). Αυτό σημαίνει ότι οι ανεξάρτητες μεταβλητές παρουσιάζουν στατιστική

σημαντικότητα και διαδραματίζουν στατιστικά σημαντικό ρόλο στο μοντέλο. Αντιθέτως το p-value της μεταβλητής ΥΨΟΣ είναι μεγαλύτερο του 0,05 (επίπεδο σημαντικότητας 5%), δηλαδή η συγκεκριμένη μεταβλητή δεν διαδραματίζει στατιστικά σημαντικό ρόλο στο μοντέλο. Σύμφωνα με τα παραπάνω, η παραπάνω διαδικασία πρέπει να πραγματοποιηθεί εκ νέου εισάγοντας μόνο τις μεταβλητές ΦΥΛΟ και ΔΜΣ, επειδή διαδραματίζουν στατιστικά σημαντικό ρόλο στο μοντέλο.

Block 0: Beginning Block

Classification Table^{a,b}

Observed		Predicted		
		ΧΟΛΗΣΤΕΡΙΝΗ		Percentage Correct
		ΚΑΝΟΝΙΚΟ	ΑΥΣΗΜΕΝΟ	
Step 0	ΧΟΛΗΣΤΕΡΙΝΗ	840	0	100,0
	ΚΑΝΟΝΙΚΟ			
	ΑΥΣΗΜΕΝΟ	160	0	,0
Overall Percentage				84,0

a. Constant is included in the model.

b. The cut value is ,500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-1,658	,086	369,562	1	,000	,190

Variables not in the Equation

	Score	df	Sig.
Step 0 Variables ΦΥΛΟ(1)	,268	1	,605
ΔΜΣ	3,588	1	,058
Overall Statistics	3,775	2	,151

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

	Chi-square	df	Sig.
Step 1 Step	3,353	2	,187
Block	3,353	2	,187
Model	3,353	2	,187

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	875,986 ^a	,003	,006

a. Estimation terminated at iteration number 4 because parameter estimates changed by less

Πίνακας 30: Μοντέλο Πρόβλεψης μεταβλητής ΧΟΛΗΣΤΕΡΙΝΗ

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	15,157	8	,056

Contingency Table for Hosmer and Lemeshow Test

		ΧΟΛΗΣΤΕΡΙΝΗ = ΚΑΝΟΝΙΚΟ		ΧΟΛΗΣΤΕΡΙΝΗ = ΑΥΞΗΜΕΝΟ		Total
		Observed	Expected	Observed	Expected	
Step 1	1	90	86,518	10	13,482	100
	2	90	85,544	10	14,456	100
	3	87	85,960	14	15,040	101
	4	77	84,797	23	15,203	100
	5	86	84,465	14	15,535	100
	6	84	84,098	16	15,902	100
	7	78	83,681	22	16,319	100
	8	79	83,201	21	16,799	100
	9	90	82,495	10	17,505	100
	10	79	79,242	20	19,758	99

Classification Table^a

	Observed	Predicted		
		ΧΟΛΗΣΤΕΡΙΝΗ		Percentage Correct
		ΚΑΝΟΝΙΚΟ	ΑΥΞΗΜΕΝΟ	
Step 1	ΧΟΛΗΣΤΕΡΙΝΗ ΚΑΝΟΝΙΚΟ	840	0	100,0
	ΑΥΞΗΜΕΝΟ	159	1	,6
Overall Percentage				84,1

a. The cut value is ,500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a ΦΥΛΟ(1)	-,095	,218	,190	1	,663	,909
ΔΜΣ	-2,149	,303	50,279	1	,000	,117

a. Variable(s) entered on step 1: ΦΥΛΟ, ΔΜΣ.

Πίνακας 31: Μοντέλο Πρόβλεψης μεταβλητής ΧΟΛΗΣΤΕΡΙΝΗ

Από τους παραπάνω πίνακες προκύπτουν τα εξής:

- Στο Block 0 δεν υπάρχει κάποια ουσιαστική διαφορά.
- Στο Block 1 (έλεγχος των Hosmer-Lemeshow) παρατηρείται ότι το μοντέλο είναι καλά προσαρμοσμένο διότι $p\text{-value} > 0,05$. Επιπλέον το $p\text{-value}$ της μεταβλητής ΔΜΣ είναι μικρότερο του 0,05 (επίπεδο σημαντικότητας 5%), δηλαδή η συγκεκριμένη μεταβλητή διαδραματίζει στατιστικά σημαντικό ρόλο στο μοντέλο. Αντίθετα το $p\text{-value}$ της μεταβλητής ΦΥΛΟ είναι μεγαλύτερο του 0,05 (επίπεδο σημαντικότητας 5%), δηλαδή η συγκεκριμένη μεταβλητή δεν διαδραματίζει στατιστικά σημαντικό ρόλο στο μοντέλο. Αυτό σημαίνει ότι η παραπάνω διαδικασία πρέπει να πραγματοποιηθεί εκ νέου εισάγοντας μόνο τη μεταβλητή ΔΜΣ, επειδή μόνο εκείνη διαδραματίζει στατιστικά σημαντικό ρόλο στο μοντέλο.

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	3,161	1	,075
	Block	3,161	1	,075
	Model	3,161	1	,075

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	876,179 ^a	,003	,005

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	15,424	8	,051

Contingency Table for Hosmer and Lemeshow Test

		ΧΟΛΗΣΤΕΡΙΝΗ = ΚΑΝΟΝΙΚΟ		ΧΟΛΗΣΤΕΡΙΝΗ = ΑΥΣΗΜΕΝΟ		Total
		Observed	Expected	Observed	Expected	
Step 1	1	94	86,218	6	13,782	100
	2	85	85,464	15	14,536	100
	3	80	85,115	20	14,885	100
	4	87	84,815	13	15,185	100
	5	85	84,493	15	15,507	100
	6	76	84,983	25	16,017	101
	7	81	83,742	19	16,258	100
	8	84	83,307	16	16,693	100
	9	86	82,579	14	17,421	100
	10	82	79,284	17	19,716	99

Πίνακας 32: Μοντέλο Πρόβλεψης μεταβλητής ΧΟΛΗΣΤΕΡΙΝΗ

Classification Table^a

Observed	Observed	Predicted		Percentage Correct
		ΧΟΛΗΣΤΕΡΙΝΗ		
		ΚΑΝΟΝΙΚΟ	ΑΥΣΗΜΕΝΟ	
Step 1	ΧΟΛΗΣΤΕΡΙΝΗ ΚΑΝΟΝΙΚΟ	840	0	100,0
	ΑΥΣΗΜΕΝΟ	159	1	,6
	Overall Percentage			84,1

a. The cut value is ,500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a ΔΜΣ	-2,175	,298	53,362	1	,000	,114

a. Variable(s) entered on step 1: ΔΜΣ.

Πίνακας 33: Μοντέλο Πρόβλεψης μεταβλητής ΧΟΛΗΣΤΕΡΙΝΗ

Από τους παραπάνω πίνακες προκύπτουν τα εξής:

- Στο Block 0 δεν υπάρχει κάποια ουσιαστική διαφορά.
- Στο Block 1 (έλεγχος των Hosmer-Lemeshow) παρατηρείται ότι το μοντέλο είναι καλά προσαρμοσμένο διότι p-value>0,05. Επιπλέον το p-value της μεταβλητής ΔΜΣ είναι μικρότερο

ΣΑΚΧΑΡΟ

Πραγματοποιείται Τακτική Παλινδρόμηση διότι η μεταβλητή ΣΑΚΧΑΡΟ παίρνει 3 τιμές.

Warnings

There are 1991 (66,6%) cells (i.e., dependent variable levels by combinations of predictor variable values) with zero frequencies.

Πίνακας 35: Μοντέλο Πρόβλεψης μεταβλητής ΣΑΚΧΑΡΟ

Στον παραπάνω πίνακα αναφέρεται ο αριθμός των κελιών με μηδενική (0) συχνότητα, ο οποίος ισούται με 1991.

Case Processing Summary

		N	Marginal Percentage
ΣΑΚΧΑΡΟ	ΚΑΝΟΝΙΚΟ	894	89,4%
	ΑΥΞΗΜΕΝΗ	84	8,4%
	ΧΡΗΣΗ ΙΝΣΟΥΛΙΝΗΣ	22	2,2%
Valid		1000	100,0%
Missing		0	
Total		1000	

Πίνακας 36: Μοντέλο Πρόβλεψης μεταβλητής ΣΑΚΧΑΡΟ

Παραπάνω παρουσιάζεται η κατανομή των 3 κατηγοριών της μεταβλητής ΣΑΚΧΑΡΟ.

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	783,019			
Final	679,866	103,154	8	,000

Link function: Logit.

Πίνακας 37: Μοντέλο Πρόβλεψης μεταβλητής ΣΑΚΧΑΡΟ

Στον παραπάνω πίνακα το αποτέλεσμα της Μηδενικής Υπόθεσης, η οποία απορρίπτεται σε επίπεδο σημαντικότητας 5%, διότι $p\text{-value} < 0,001 < 0,05$. Συγκεκριμένα, υπάρχει χαρακτηριστική διαφορά μεταξύ του Βασικού και Τελικού μοντέλου. Σε αυτό το σημείο πρέπει να αναφερθεί ότι το βασικό μοντέλο δεν αποτελείται από καμία ανεξάρτητη μεταβλητή, σε αντίθεση με το τελικό το οποίο αποτελείται από όλες τις δυνατές ανεξάρτητες μεταβλητές.

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	2400,300	1982	,000
Deviance	678,479	1982	1,000

Link function: Logit.

Πίνακας 38: Μοντέλο Πρόβλεψης μεταβλητής ΣΑΚΧΑΡΟ

Στον παραπάνω πίνακα παρατηρούμε ότι $p\text{-value}(\text{Pearson}) < 0,001 < 0,05$, το οποίο σημαίνει ότι δεν υπάρχει καλή προσαρμογή του μοντέλου πρόβλεψής (θα παρουσιαστεί παρακάτω)

Parameter Estimates

	Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval		
						Lower Bound	Upper Bound	
Threshold	[ΣΑΚΧΑΡΟ = 0]	19,307	6,728	8,234	1	,004	6,119	32,495
	[ΣΑΚΧΑΡΟ = 1]	21,101	6,739	9,804	1	,002	7,892	34,309
Location	ΗΛΙΚΙΑ	,062	,009	45,271	1	,000	,044	,080
	ΔΜΣ	,080	,037	4,589	1	,032	,007	,153
	ΥΨΟΣ	8,532	3,829	4,965	1	,026	1,027	16,037
	ΒΑΡΟΣ	-,030	,015	3,735	1	,053	-,060	,000
	ΚΑΡΚΙΝΟΣ	-,667	,557	1,434	1	,231	-1,759	,425
	ΧΟΛΗΣΤΕΡΙΝΗ	,758	,245	9,559	1	,002	,277	1,238
	ΠΙΕΣΗ	-,178	,256	,485	1	,486	-,679	,323
	ΦΥΛΟ	-,345	,437	,624	1	,430	-1,202	,511

Link function: Logit.

Πίνακας 39: Μοντέλο Πρόβλεψης μεταβλητής ΣΑΚΧΑΡΟ

Το μοντέλο πρόβλεψης που προκύπτει από τον παραπάνω πίνακα είναι:

$$\log it[P(y \leq k)] = \log \frac{P(y \leq k)}{1 - P(y \leq k)}$$

$$P(y \leq k) = p_1 + p_2 + p_3 + \dots + p_k, \quad k = 1, 2, \dots, K$$

$$P(y \leq k) = \frac{e^z}{1 + e^z}$$

$$z = 0,062x_1 + 0,080x_2 + 8,532x_3 - 0,030x_4 - 0,667x_5 + 0,758x_6 - 0,178x_7 - 0,345x_8$$

$$x_1 : \text{ΗΛΙΚΙΑ}, \quad x_2 : \text{ΔΜΣ}, \quad x_3 : \text{ΥΨΟΣ}, \quad x_4 : \text{ΒΑΡΟΣ}, \quad x_5 : \text{ΚΑΡΚΙΝΟΣ},$$

$$x_6 : \text{ΧΟΛΗΣΤΕΡΙΝΗ}, \quad x_7 : \text{ΠΙΕΣΗ}, \quad x_8 : \text{ΦΥΛΟ}$$

Μοντέλο Πρόβλεψης 4: Μεταβλητή στόχος ΣΑΚΧΑΡΟ

Σε αυτό το σημείο θα πρέπει να σημειωθεί ότι τα p-value των μεταβλητών ΗΛΙΚΙΑ, ΔΜΣ, ΥΨΟΣ, ΒΑΡΟΣ και ΧΟΛΗΣΤΕΡΙΝΗ είναι μικρότερα του 0,05 (επίπεδο σημαντικότητας 5%). Αυτό σημαίνει ότι οι ανεξάρτητες μεταβλητές παρουσιάζουν στατιστική σημαντικότητα και διαδραματίζουν στατιστικά σημαντικό ρόλο στο μοντέλο. Αντιθέτως τα p-value των μεταβλητών ΚΑΡΚΙΝΟΣ, ΠΙΕΣΗ και ΦΥΛΟ είναι μεγαλύτερα του 0,05 (επίπεδο σημαντικότητας 5%), δηλαδή οι συγκεκριμένες μεταβλητές δεν διαδραματίζουν στατιστικά σημαντικό ρόλο στο μοντέλο.

ΠΙΕΣΗ

Πραγματοποιείται Τακτική Παλινδρόμηση διότι η μεταβλητή ΠΙΕΣΗ παίρνει 3 τιμές.

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	1105,504			
Final	1065,413	40,091	8	,000

Link function: Logit.

	Chi-Square	df	Sig.
Pearson	2386,358	1984	,000
Deviance	1065,413	1984	1,000

Link function: Logit.

Πίνακας 40: Μοντέλο Πρόβλεψης μεταβλητής ΠΙΕΣΗ

Σύμφωνα με τα αποτελέσματα του παραπάνω πίνακα, υπάρχει χαρακτηριστική διαφορά μεταξύ του Βασικού και Τελικού μοντέλου και επιπλέον δεν υπάρχει καλή προσαρμογή του μοντέλου πρόβλεψης που θα παρουσιαστεί παρακάτω.

	Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Threshold [ΠΙΕΣΗ = 0]	-4,491	8,664	,269	1	,604	-21,472	12,490
[ΠΙΕΣΗ = 1]	-2,111	8,664	,059	1	,808	-19,092	14,870
Location ΗΛΙΚΙΑ	,026	,007	15,335	1	,000	,013	,040
ΔΜΣ	-,159	,117	1,838	1	,175	-,388	,071
ΥΨΟΣ	-5,104	5,138	,987	1	,320	-15,174	4,966
ΒΑΡΟΣ	,066	,041	2,545	1	,111	-,015	,147
ΚΑΡΚΙΝΟΣ	,236	,329	,517	1	,472	-,408	,881
ΧΟΛΗΣΤΕΡΙΝΗ	,476	,205	5,383	1	,020	,074	,879
ΦΥΛΟ	,484	,387	1,564	1	,211	-,275	1,242
ΣΑΚΧΑΡΟ	-,147	,221	,442	1	,506	-,579	,286

Link function: Logit.

Πίνακας 41: Μοντέλο Πρόβλεψης μεταβλητής ΠΙΕΣΗ

Το μοντέλο πρόβλεψης που προκύπτει από τον παραπάνω πίνακα είναι:

$$\log it[P(y \leq k)] = \log \frac{P(y \leq k)}{1 - P(y \leq k)}$$

$$P(y \leq k) = p_1 + p_2 + p_3 + \dots + p_k, \quad k = 1, 2, \dots, K$$

$$P(y \leq k) = \frac{e^z}{1 + e^z}$$

$$z = 0,026x_1 - 0,159x_2 - 5,104x_3 + 0,066x_4 + 0,236x_5 + 0,476x_6 + 0,484x_7 - 0,147x_8$$

$$x_1 : \text{ΗΛΙΚΙΑ}, \quad x_2 : \text{ΔΜΣ}, \quad x_3 : \text{ΥΨΟΣ}, \quad x_4 : \text{ΒΑΡΟΣ}, \quad x_5 : \text{ΚΑΡΚΙΝΟΣ},$$

$$x_6 : \text{ΧΟΛΗΣΤΕΡΙΝΗ}, \quad x_7 : \text{ΦΥΛΟ}, \quad x_8 : \text{ΣΑΚΧΑΡΟ}$$

Μοντέλο Πρόβλεψης 5: Μεταβλητή στόχος ΠΙΕΣΗ

Σε αυτό το σημείο θα πρέπει να σημειωθεί ότι τα p-value των μεταβλητών ΗΛΙΚΙΑ, και ΧΟΛΗΣΤΕΡΙΝΗ είναι μικρότερα του 0,05 (επίπεδο σημαντικότητας 5%). Αυτό σημαίνει ότι οι ανεξάρτητες μεταβλητές παρουσιάζουν στατιστική σημαντικότητα και διαδραματίζουν στατιστικά σημαντικό ρόλο στο μοντέλο. Αντιθέτως τα p-value των μεταβλητών ΔΜΣ, ΥΨΟΣ, ΒΑΡΟΣ, ΚΑΡΚΙΝΟΣ, ΠΙΕΣΗ και ΦΥΛΟ είναι μεγαλύτερα του 0,05 (επίπεδο σημαντικότητας 5%), δηλαδή οι συγκεκριμένες μεταβλητές δεν διαδραματίζουν στατιστικά σημαντικό ρόλο στο μοντέλο.

4.4 ΒΕΛΤΙΣΤΟ ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ

Σε αυτή την υποπαράγραφο, γίνεται προσπάθεια ανάπτυξης του Βέλτιστου μοντέλου πρόβλεψης της μεταβλητής ΔΜΣ.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,706 ^a	,499	,498	6,141593030
2	,939 ^b	,882	,882	2,983145312

a. Predictors: (Constant), ΒΑΡΟΣ

b. Predictors: (Constant), ΒΑΡΟΣ, ΥΨΟΣ

Πίνακας 42: Βέλτιστο Μοντέλο Πρόβλεψης μεταβλητής ΔΜΣ

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6,838	,828		8,260	,000
	ΒΑΡΟΣ	,284	,009	,706	31,504	,000
2	(Constant)	104,205	1,759		59,242	,000
	ΒΑΡΟΣ	,393	,005	,977	82,229	,000
	ΥΨΟΣ	-64,234	1,130	-,676	-56,860	,000

a. Dependent Variable: ΔΜΣ

Excluded Variables^a

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	ΗΛΙΚΙΑ	,063 ^b	2,798	,005	,088	,991
	ΥΨΟΣ	-,676 ^b	-56,860	,000	-,874	,839
2	ΗΛΙΚΙΑ	-,002 ^c	-,176	,860	-,006	,980

a. Dependent Variable: ΔΜΣ

b. Predictors in the Model: (Constant), ΒΑΡΟΣ

c. Predictors in the Model: (Constant), ΒΑΡΟΣ, ΥΨΟΣ

Πίνακας 43: Βέλτιστο Μοντέλο Πρόβλεψης μεταβλητής ΔΜΣ

Βάσει των παραπάνω πινάκων, το βέλτιστο γραμμικό μοντέλο είναι εκείνο που περιέχει τις μεταβλητές ΒΑΡΟΣ και ΥΨΟΣ, επειδή $R_2(Square) > R_1(Square)$.

$$y = 0,393x_1 - 64,234x_2 + 104,205$$

$$x_1 : \text{ΒΑΡΟΣ}, \quad x_2 : \text{ΥΨΟΣ}$$

Μοντέλο Πρόβλεψης 6: Βέλτιστο Μοντέλο Πρόβλεψης μεταβλητής ΔΜΣ

5. ΣΥΜΠΕΡΑΣΜΑΤΑ

5.1 ΕΛΕΓΧΟΣ ΚΑΝΟΝΙΚΟΤΗΤΑΣ

Στο 3^ο Κεφάλαιο, πραγματοποιήθηκε έλεγχος της κανονικότητας των δεδομένων βάσει του διαγράμματος P-P Plots και του Ραβδογράμματος. Κατά τον έλεγχο εξήλθαν λάθος συμπεράσματα. Συγκεκριμένα, διαπιστώθηκε η κανονική κατανομή των μεταβλητών ΔΜΣ, ΒΑΡΟΣ, ΥΨΟΣ και ΗΛΙΚΙΑ.

Στο 4ο Κεφάλαιο πραγματοποιήθηκε στατιστικός έλεγχος, όπου διαπιστώθηκε η μη κανονική κατανομή των παραπάνω μεταβλητών. Επειδή ο στατιστικός έλεγχος Κανονικότητας είναι ισχυρότερος από το γραφικό έλεγχο, υπερισχύουν τα συμπεράσματα του 1ου.

Σε αυτό το σημείο πρέπει να αναφερθεί ότι αυτή η «εξέλιξη» στην μελέτη μας είναι σημαντική διότι η λάθος εκτίμηση της Κανονικότητας μας αναγκάζει να ακολουθήσουμε λανθασμένη διαδικασία. Στην παρούσα, διαπιστώθηκε έγκαιρα κι ακολουθήθηκε μη παραμετρική στατιστική μεθοδολογία.

5.2 ΑΝΑΛΥΣΗ ΣΥΣΧΕΤΙΣΗΣ

5.2.1 ΠΟΣΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ

Σύμφωνα με την παραπάνω διαπίστωση, εφαρμόστηκε ανάλυση συσχέτισης με μη παραμετρική μεθοδολογία (Spearman).

Οι συσχετίσεις μεταξύ των μεταβλητών συνοψίζονται στο παρακάτω πίνακα:

	ΒΑΡΟΣ	ΗΛΙΚΙΑ	ΔΜΣ	ΥΨΟΣ
ΒΑΡΟΣ	1	0,124	0,913	0,465
ΗΛΙΚΙΑ	0,124	1	0,168	-0,086
ΔΜΣ	0,913	0,168	1	0,143
ΥΨΟΣ	0,465	-0,086	0,143	1

Πίνακας 44: Συσχετίσεις μεταξύ ποσοτικών μεταβλητών

Από τον πίνακα παρατηρούνται τα εξής:

- Η συσχέτιση όλων των δυνατών συνδυασμών κρίνεται στατιστικά σημαντική σε επίπεδο σημαντικότητας 1%.
- Όλες οι συσχετίσεις μεταξύ των μεταβλητών, εκτός από 1 συνδυασμό, είναι θετικές. Αυτό σημαίνει ότι η αύξηση της τιμής της μίας μεταβλητής προκαλεί αύξησης στην τιμή της άλλης.
- Η συσχέτιση των μεταβλητών ΥΨΟΣ και ΗΛΙΚΙΑ είναι αρνητική. Αυτό σημαίνει ότι η αύξηση της τιμής της μίας μεταβλητής προκαλεί μείωση στην τιμή της άλλης.

5.2.2 ΠΟΙΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ

Οι συσχετίσεις μεταξύ των μεταβλητών συνοψίζονται παρακάτω:

	ΦΥΛΟ	ΚΑΡΚΙΝΟΣ	ΣΑΚΧΑΡΟ	ΧΟΛΗΣΤΕΡΙΝΗ	ΠΙΕΣΗ
ΦΥΛΟ		ΟΧΙ	ΝΑΙ	ΟΧΙ	ΝΑΙ
ΚΑΡΚΙΝΟΣ	ΟΧΙ		ΟΧΙ	ΟΧΙ	ΟΧΙ
ΣΑΚΧΑΡΟ	ΝΑΙ	ΟΧΙ		ΝΑΙ	ΟΧΙ
ΧΟΛΗΣΤΕΡΙΝΗ	ΟΧΙ	ΟΧΙ	ΝΑΙ		ΝΑΙ
ΠΙΕΣΗ	ΝΑΙ	ΟΧΙ	ΟΧΙ	ΝΑΙ	

Πίνακας 45: Συσχετίσεις μεταξύ ποιοτικών μεταβλητών

Από τον πίνακα παρατηρούνται τα εξής:

- Κάθε μεταβλητή συσχετίζεται με 2 άλλες μεταβλητές, οι οποίες δεν είναι σταθερές.
- Η μεταβλητή ΚΑΡΚΙΝΟΣ δεν συσχετίζεται με καμία άλλη μεταβλητή.

5.3 ΒΕΛΤΙΣΤΟ ΚΑΙ ΜΗ ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ ΤΗΣ ΜΕΤΑΒΛΗΤΗΣ ΔΜΣ

Στο 4^ο Κεφάλαιο πραγματοποιήθηκε η ανάπτυξη βέλτιστου και μη, μοντέλου πρόβλεψης για τη μεταβλητή ΔΜΣ (συναρτήσει ποσοτικών μεταβλητών).

	ΒΑΡΟΣ	ΗΛΙΚΙΑ	ΥΨΟΣ	ΣΤΑΘΕΡΑ
ΑΠΛΟ ΜΟΝΤΕΛΟ	-	-	0,393	104,205
ΒΕΛΤΙΣΤΟ ΜΟΝΤΕΛΟ	0,393	-0,001	-64,234	104,205

Πίνακας 46: Σύγκριση μεταξύ Απλού και Βέλτιστου Μοντέλου

Από την παραπάνω ανάπτυξη, παρατηρούνται τα εξής:

- Το βέλτιστο μοντέλο πρόβλεψης (Μέθοδος Stepwise) δεν ταυτίζεται με το μοντέλο πρόβλεψης (Μέθοδος Enter).
- Στο βέλτιστο μοντέλο πρόβλεψης, όσο αυξάνεται η μεταβλητή ΒΑΡΟΣ, αυξάνεται η μεταβλητή ΔΜΣ. Όσο αυξάνεται η μεταβλητή ΥΨΟΣ, μειώνεται η μεταβλητή ΔΜΣ. Τέλος, όσο αυξάνεται η μεταβλητή ΗΛΙΚΙΑ, παρατηρείται μια απειροελάχιστη μείωση της μεταβλητής ΔΜΣ.
- Στο απλό μοντέλο πρόβλεψης, όσο αυξάνεται η μεταβλητή ΥΨΟΣ, αυξάνεται η μεταβλητή ΔΜΣ.

Παρακάτω, πραγματοποιείται έλεγχος Κανονικότητας, Ανεξαρτησίας και Ομοσκεδαστικότητας των σφαλμάτων του βέλτιστου μοντέλου.

ΕΛΕΓΧΟΣ ΚΑΝΟΝΙΚΟΤΗΤΑΣ ΤΩΝ ΣΦΑΛΜΑΤΩΝ

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	,214	1000	,000	,486	1000	,000

a. Lilliefors Significance Correction

Πίνακας 47: Έλεγχος Κανονικότητας των Σφαλμάτων

Παρατηρώντας τον παραπάνω πίνακα, διαπιστώνουμε ότι σε επίπεδο σημαντικότητας 5% $p\text{-value} < 0,05$. Αυτό σημαίνει ότι τα σφάλματα του βέλτιστου μοντέλου πρόβλεψης της μεταβλητής δεν προσαρμόζονται ικανοποιητικά στην Κανονική Κατανομή.

ΕΛΕΓΧΟΣ ΑΝΕΞΑΡΤΗΣΙΑΣ ΤΩΝ ΣΦΑΛΜΑΤΩΝ

	Unstandardized Residual
Test Value ^a	-,62714
Cases < Test Value	500
Cases >= Test Value	500
Total Cases	1000
Number of Runs	471
Z	-1,898
Asymp. Sig. (2-tailed)	,058

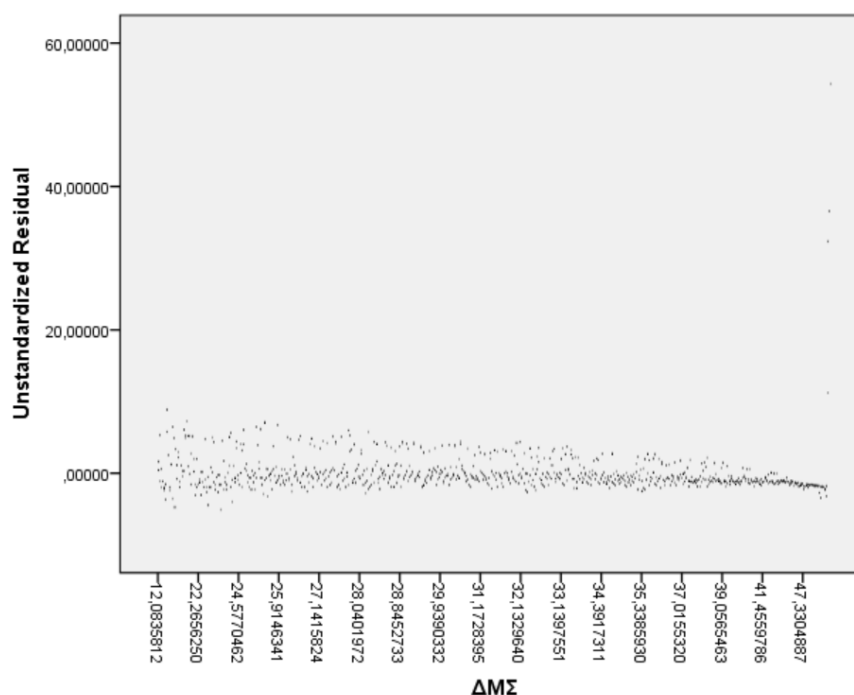
a. Median

Πίνακας 48: Έλεγχος Ανεξαρτησίας των Σφαλμάτων

Στον παραπάνω πίνακα, παρατηρούμε ότι $p\text{-value} = 0,058 > 0,05$. Αυτό σημαίνει ότι τα σφάλματα του βέλτιστου μοντέλου της μεταβλητής ΔΜΣ είναι ανεξάρτητα.

ΕΛΕΓΧΟΣ ΟΜΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑΣ ΤΩΝ ΣΦΑΛΜΑΤΩΝ

Ο έλεγχος ομοσκεδαστικότητας των σφαλμάτων πραγματοποιήθηκε με τη βοήθεια διαγράμματος (scatterplot), όπως παρακάτω:



Διάγραμμα 25: Έλεγχος Ομοσκεδαστικότητας των Σφαλμάτων

Στο παραπάνω διάγραμμα, παρατηρείται ότι υπάρχει διασπορά των παρατηρήσεων. Δηλαδή, δεν φαίνεται να παραβιάζεται η ομοσκεδαστικότητα των σφαλμάτων του βέλτιστου μοντέλου πρόβλεψης.

5.4 ΜΟΝΤΕΛΑ ΠΡΟΒΛΕΨΗΣ ΜΕ ΠΟΙΟΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ ΣΤΟΧΟΥΣ

Τα μοντέλα πρόβλεψης με μεταβλητές στόχους τις μεταβλητές ΚΑΡΚΙΝΟΣ, ΣΑΚΧΑΡΟ, ΧΟΛΗΣΤΕΡΙΝΗ και ΠΙΕΣΗ συνοψίζονται στο παρακάτω πίνακα:

	ΚΑΡΚΙΝΟΣ	ΣΑΚΧΑΡΟ	ΧΟΛΗΣΤΕΡΙΝΗ	ΠΙΕΣΗ
ΚΑΡΚΙΝΟΣ		0,667	-	0,236
ΣΑΚΧΑΡΟ	-			0,147
ΧΟΛΗΣΤΕΡΙΝΗ	-0,057	0,758		0,476
ΠΙΕΣΗ		-0,178		
ΦΥΛΟ	0,049	-0,345	0,051	0,484
ΗΛΙΚΙΑ	-	0,062	-	0,026
ΔΜΣ	-35,405	0,080	-12,180	-0,159
ΒΑΡΟΣ		-0,030		0,066
ΥΨΟΣ	0,328	8,532	0,764	-5,104

Πίνακας 49: Μοντέλα πρόβλεψης με ποιοτικές μεταβλητές στόχους

Ο παραπάνω πίνακας είναι συνδυαστικός. Παρουσιάζει τα μοντέλα πρόβλεψης με τους αντίστοιχους συντελεστές τους και ταυτόχρονα με τα χρώματα παρουσιάζεται η ύπαρξη ή μη,

της συσχέτισης μεταξύ των ποιοτικών μεταβλητών. Συγκεκριμένα το μπλε χρώμα δηλώνει την ύπαρξη συσχέτισης, ενώ το κόκκινο τη μη ύπαρξη.

Παρατηρούμε τα παρακάτω:

- Το μοντέλο πρόβλεψης της μεταβλητής ΚΑΡΚΙΝΟΣ εμπεριέχονται οι μεταβλητές ΧΟΛΗΣΤΕΡΙΝΗ και ΦΥΛΟ με τις οποίες δεν υπάρχει συσχέτιση.
- Το μοντέλο πρόβλεψης της μεταβλητής ΣΑΚΧΑΡΟ εμπεριέχονται οι μεταβλητές ΚΑΡΚΙΝΟΣ και ΠΙΕΣΗ με τις οποίες δεν υπάρχει συσχέτιση. Επιπλέον εμπεριέχονται οι μεταβλητές ΧΟΛΗΣΤΕΡΙΝΗ και ΦΥΛΟ με τις οποίες παρουσιάζεται συσχέτιση.
- Το μοντέλο πρόβλεψης της μεταβλητής ΧΟΛΗΣΤΕΡΙΝΗ εμπεριέχονται οι μεταβλητή ΦΥΛΟ με την οποία δεν υπάρχει συσχέτιση. Επιπλέον δεν εμπεριέχονται οι μεταβλητές ΣΑΚΧΑΡΟ και ΠΙΕΣΗ με τις οποίες παρουσιάζεται συσχέτιση.
- Το μοντέλο πρόβλεψης της μεταβλητής ΠΙΕΣΗ εμπεριέχονται οι μεταβλητές ΚΑΡΚΙΝΟΣ και ΣΑΚΧΑΡΟ με τις οποίες δεν υπάρχει συσχέτιση. Επιπλέον εμπεριέχονται οι μεταβλητές ΧΟΛΗΣΤΕΡΙΝΗ και ΦΥΛΟ με τις οποίες παρουσιάζεται συσχέτιση.
- Στα μοντέλα πρόβλεψης των μεταβλητών - στόχων ΣΑΚΧΑΡΟ και ΠΙΕΣΗ εμπεριέχονται όλες οι άλλες μεταβλητές.

5.5 ΒΑΣΙΚΕΣ ΕΠΙΣΗΜΑΝΣΕΙΣ – ΣΥΜΒΟΥΛΕΣ

Από την ανάλυση των συγκεκριμένων δεδομένων μπορούμε να εξάγουμε κάποια πρακτικά συμπεράσματα. Συγκεκριμένα παρατηρούμε ότι σε όλα τα μοντέλα πρόβλεψης των ασθενειών εμπεριέχονται οπωσδήποτε οι μεταβλητές ΔΜΣ , ΥΨΟΣ και ΦΥΛΟ, ενώ σε 2 από τις 4 ασθένειες εμπεριέχεται και η μεταβλητή ΒΑΡΟΣ. Βάσει αυτής της παρατήρησης μπορούμε να καταλήξουμε στο συμπέρασμα ότι η διατροφή παίζει σημαντικό ρόλο στην εμφάνιση ή μη, κάποιας από τις παθήσεις που μελετάμε. Οι παρακάτω συμβουλές αφορούν την ισορροπημένη διατροφή και πώς θα διατηρηθούμε υγιείς.

....«Δώστε ενέργεια στο σώμα σας συχνά

Η κατανάλωση των κατάλληλων τροφών σε μικρές ποσότητες ανά τακτά χρονικά διαστήματα παρέχει στον οργανισμό σας ενέργεια με σταθερό ρυθμό. Έτσι αποφεύγετε αυξομειώσεις στην έκκριση ινσουλίνης και των επιπέδων γλυκόζης στο αίμα, με αποτέλεσμα να μην έχετε έντονο αίσθημα πείνας, κούρασης και οι κρίσεις λιπαιμαργίας να αποτελούν παρελθόν.

Συνδυάστε πρωτεΐνη με σύνθετους υδατάνθρακες

Όταν συνδυάζετε στα γεύματά σας αυτές τις δύο ομάδες τροφών, στη σωστή αναλογία, εξασφαλίζετε σταθερή ροή ενέργειας στο σώμα σας, καθώς το σώμα μετατρέπει την ενέργεια από το φαγητό σχετικά αργά σε γλυκόζη. Με αυτό τον τρόπο εκκρίνεται λιγότερη ινσουλίνη και κατά συνέπεια το σώμα αποθηκεύει λιγότερο λίπος.

Τρώτε ποικιλία τροφών

Είναι συνηθισμένο να αγοράζουμε συνέχεια τα ίδια τρόφιμα. Για την ακρίβεια, εννέα στις δέκα φορές που πηγαίνουμε για ψώνια οι περισσότεροι από εμάς αγοράζουμε μόνο το 10% της ποικιλίας των τροφών που μας είναι διαθέσιμες.

Προσπαθήστε να συμπεριλαμβάνετε δύο νέα τρόφιμα στο καλάθι των αγορών σας την εβδομάδα.

Πίνετε άφθονο νερό

Πίνετε τόσο νερό που να μην αισθάνεστε διψασμένοι. Όταν διψάσετε έχετε ήδη αφυδατωθεί! Ελέγχετε το χρώμα στα ούρα σας για να βεβαιωθείτε ότι πίνετε αρκετό νερό. Όταν παράγετε ούρα με κίτρινο χρώμα χρειάζεστε περισσότερο. Μην ξεχνάτε ότι όταν ιδρώνετε πολύ, όπως το καλοκαίρι ή με έντονη άσκηση και όταν καταναλώνετε αλκοόλ χρειάζεστε περισσότερο νερό.

Ξεκινήστε την ημέρα σας με το αγαπημένο σας πρωινό

Τρώγοντας ένα ισορροπημένο πρωινό δίνετε στο σώμα σας τα καύσιμα που χρειάζεται για να έχετε ενέργεια την υπόλοιπη ημέρα και ενεργοποιείτε το μεταβολισμό σας. Διαλέξτε ένα πρωινό που να σας αρέσει ώστε να ξεκινήσετε την ημέρα σας ευχάριστα.

Αν αισθάνεστε ότι είναι δύσκολο να χωρέσετε το πρωινό στον περιορισμένο χρόνο σας διαλέξτε ένα πρωινό που δεν απαιτεί μεγάλη προετοιμασία. Σκεφτείτε ότι θα κερδίσετε τον χαμένο χρόνο όντας πιο αποδοτικοί την υπόλοιπη ημέρα.

Κινηθείτε!

Ξεκουραστείτε κάνοντας άσκηση! Όσο κουρασμένοι και αν αισθάνεστε ένας περίπατος στο αγαπημένο σας μέρος, δίπλα στην θάλασσα ή στο πάρκο είναι ένας ωραίος τρόπος να χαλαρώσετε. Όταν συνδυάσετε άσκηση με υγιεινή διατροφή έχετε μεγάλες πιθανότητες να χάσετε και να διατηρήσετε τα περιττά κιλά, και θα αισθάνεστε υγιής.

Προσπαθήστε να συμπεριλάβετε στο πρόγραμμά σας τουλάχιστον τριάντα λεπτά δραστηριότητας τρεις φορές την εβδομάδα.

Τρώτε λίπος για να χάσετε λίπος!

Είναι συνηθισμένο στην προσπάθειά μας να χάσουμε βάρος να αποφεύγουμε να τρώμε λιπαρές τροφές. Υπάρχουν όμως κάποια λιπαρά (ω-3 και ω-6) τα οποία το σώμα μας δεν μπορεί να παράγει και είναι απαραίτητα για την ομαλή λειτουργία του οργανισμού μας.

Ας αποφεύγουμε λοιπόν τα λιπαρά ζωικής προέλευσης και τα τηγανητά και ας καταναλώνουμε αρκετά λιπαρά ψάρια και λίπη φυτικής προέλευσης.

Και μην ξεχνάτε ότι είναι απόλυτα φυσιολογικό να 'παρεκτρέπεστε' που και που... Αν ακολουθείτε τις παραπάνω συμβουλές τουλάχιστον 80% του χρόνου σας μπορείτε να απολαμβάνετε την κοινωνική σας ζωή και παράλληλα να είστε υγιής.»....

(www.iatronet.gr)

ΒΙΒΛΙΟΓΡΑΦΙΑ

ΕΛΛΗΝΙΚΗ

1. Καρανδρέας Α., Τυρνεόπουλος Α. (2014), Πτυχιακή Εργασία με τίτλο: «Αναλυτική Λογιστική Παλινδρόμηση σε δεδομένα από έρευνα αγοράς», Σχολή Διοίκησης και Οικονομίας, Τ.Ε.Ι ΑΝ. ΜΑΚΕΔΟΝΙΑΣ & ΘΡΑΚΗΣ.
2. Κουτσομπλιάς Σ. (2019), Μεταπτυχιακή Διατριβή με τίτλο: «Στατιστική Ανάλυση βιοχημικών και αιματολογικών παραμέτρων σε συνάρτηση με την κατανάλωση αλκοόλ σε ασθενείς που εμφανίζουν ηπατική δυσλειτουργία», ΔΠΜΣ «Πληροφορική και Υπολογιστική Βιοϊατρική», Παν. ΘΕΣΣΑΛΙΑΣ.
3. Μπατσίδης Α. (2014), Πανεπιστημιακές Σημειώσεις: «Στατιστική Ανάλυση Δεδομένων με το SPSS», Τμήμα Μαθηματικών, Παν. ΙΩΑΝΝΙΝΩΝ.
4. Ξένη Μ., Μεταπτυχιακή Διατριβή με τίτλο: «Λογιστική Παλινδρόμηση & Διαχωριστική Ανάλυση», ΔΠΜΣ «Μαθηματικά των Υπολογιστών και των Αποφάσεων», Παν. ΠΑΤΡΩΝ.
5. Πετρίδης Δ. (2015), Ανάλυση πολυμεταβλητών τεχνικών, Τμήματος Τεχνολογίας Τροφίμων, ΑΛΕΞΑΝΔΡΕΙΟ Τ.Ε.Ι ΘΕΣΣΑΛΟΝΙΚΗΣ.
6. Σκουφά Α. (2008), Μεταπτυχιακή Διατριβή με τίτλο: «Λογιστική Παλινδρόμηση», ΠΜΣ «Στατιστική και Επιχειρησιακή Έρευνα», Παν. ΘΕΣΣΑΛΟΝΙΚΗΣ.
7. Τριανταφύλλου Ι. (2017), Πανεπιστημιακές Σημειώσεις: «Στατιστική Ανάλυση Δεδομένων με χρήση του πακέτου IBM SPSS Statistics ΔΠΜΣ «Πληροφορική και Υπολογιστική Βιοϊατρική», Παν. ΘΕΣΣΑΛΙΑΣ.

ΔΙΕΘΝΗΣ

8. Almquist Y., Ashir S. and Brännström L. (2016), A guide to quantitative methods.
9. Field A. (2018), Discovering Statistics Using IBM SPSS Statistics 5th Edition.
10. Reddy Ch., Alemayehu E. (2015), Ordinal logistic regression analysis to assess the factors that affect health status of students in Ambo University: a case of natural and computational sciences college

ΔΙΑΔΙΚΤΥΑΚΕΣ ΠΗΓΕΣ

11. https://repository.kallipos.gr/bitstream/11419/5082/1/08_chapter7.pdf
12. www.doi.org/10.1016/B978-0-12-369492-8.50014-5
13. www.el.wikipedia.org

14. www.heart.org
15. www.iatronet.gr
16. www.iubas.gr/pathiseis/karkinos-neoplasia
17. www.medium.com
18. www.medivizor.com
19. www.mikroviologos.gr
20. www.novartis.gr
21. www.nysora.com

ΑΛΦΑΒΗΤΙΚΟ ΕΥΡΕΤΗΡΙΟ ΣΗΜΑΝΤΙΚΩΝ ΟΡΩΝ

	A	
Ανάλυση Παλινδρόμησης.....		13
Ανεξαρτησία.....		64
Ασθένειες.....		1
	B	
Βέλτιστο Μοντέλο Πρόβλεψης.....		60 ,63
	Γ	
	Δ	
Δεδομένα.....		5
	E	
Έλεγχος.....		10, 32, 64, 65
Επαγωγική Στατιστική Συμπερασματολογία.....		10, 32
	Z	
	H	
	Θ	
	I	
	K	
Κανονικότητα		32, 64
Κατανομή.....		10, 32, 64
	Λ	
Λογιστική (Παλινδρόμηση).....		13
	M	
Μοντέλο Πρόβλεψης.....		43, 60 ,63
	N	
	Ξ	
	O	
Ομοσκεδαστικότητα.....		65
	Π	
Παλινδρόμηση.....		13
Περιγραφική Στατιστική Ανάλυση.....		7, 16
	P	
	Σ	
Συσχέτιση.....		10, 32, 62
Σφάλματα.....		64, 65
	T	
	Υ	
	X	
	Ω	