

ΤΕΙ ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ  
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΟΙΚΟΝΟΜΙΑΣ  
ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ



[ΘΕΜΑ  
ΠΤΥΧΙΑΧΗΣ  
ΕΡΓΑΣΙΑ]

Το SPSS ΩΣ ΕΡΓΑΛΕΙΟ ΓΙΑ ΤΗΝ ΕΠΙΛΥΣΗ  
ΠΡΟΒΛΗΜΑΤΩΝ ΣΤΑΤΙΣΤΙΚΗΣ

Ομάδα εργασίας

Αυγούστη Ελένη

Γαρουφαλής Σωκράτης

Παναγιωταράκος Νίκος

Υπεύθυνος Καθηγητής

κ. Βάσιου

ΠΑΤΡΑ 2015



## Περιεχόμενα

<b><u>Περιεχόμενα εικόνων .....</u></b>	<b><u>6</u></b>
<b><u>Περίληψη .....</u></b>	<b><u>10</u></b>
<b><u>Abstract.....</u></b>	<b><u>11</u></b>
<b><u>Εισαγωγή .....</u></b>	<b><u>12</u></b>
<b><u>Κεφάλαιο 1: Εισαγωγή στο στατιστικό πακέτο SPSS.....</u></b>	<b><u>14</u></b>
1.1 Τι είναι το SPSS.....	14
1.2 Βασικές προδιαγραφές υλικού και μνήμης.....	17
<b><u>Κεφάλαιο 2ο: Βασικές Έννοιες Στατιστικής.....</u></b>	<b><u>18</u></b>
2.1 Η έννοια της Στατιστικής Επιστήμης .....	18
2.2 Η έννοια της Μεταβλητής .....	18
2.3 Είδη Μεταβλητών.....	19
2.4 Πληθυσμός και δείγμα.....	21
2.5 Αριθμητικά Περιγραφικά Μέτρα .....	22
2.6 Η έννοια της γραμμικής παλινδρόμησης και της συσχέτισης.....	224
2.6.1 Η έννοια της γραμμικής παλινδρόμησης .....	225
2.6.2 Συντελεστής συσχέτισης .....	226
<b><u>Κεφάλαιο 3<sup>ο</sup>: Βασικές Λειτουργίες του SPSS .....</u></b>	<b><u>27</u></b>
3.1 Εκκίνηση του SPSS για Windows.....	27
3.2 Βασικά εργαλεία του SPSS για Windows .....	31
3.2.1 Παράθυρο Συντάκτη Επεξεργασίας Δεδομένων (Data Editor) .....	31
3.2.2 Παράθυρο Συντάκτη Επεξεργασίας Αποτελεσμάτων (Output Navigator) .....	31
3.2.3 Παράθυρο Συντάκτη Επεξεργασίας Πινάκων (Pivot Table Editor) .....	32

3.2.4 Παράθυρο Συντάκτη Επεξεργασίας Κειμένου(Text Output Editor) .....	33
3.2.5 Παράθυρο Συντάκτη Επεξεργασίας Γραφημάτων (Chart Editor) .....	33
3.2.6 Συντάκτης Εντολών (Syntax Editor) .....	34
3.3 Το κύριο Μενού του Spss για Windows και οι λειτουργίες του ....	35
3.4 Βασικές εργασίες με το SPSS για Windows .....	41
3.4.1. Ανάγνωση, Εισαγωγή Δεδομένων.....	41
3.4.2 Μετασχηματισμός Δεδομένων .....	42
3.4.3 Γραφική Παρουσίαση Δεδομένων .....	42
3.4.4 Ανάλυση Δεδομένων .....	42
3.5 Αποθήκευση και Έξοδος από το SPSS.....	42
3.6 Data Editor και SPSS.....	43
3.6.1 Βασικές λειτουργίες του Data Editor .....	44
3.7 Στατιστική Ανάλυση.....	53
3.7.1 Περιγραφή εντολής Frequencies .....	53
3.7.2 Περιγραφή εντολής Descriptives.....	57
3.7.3 Περιγραφή εντολής Crosstabs .....	59
3.7.4 Παρουσίαση εντολής Correlation.....	61
3.7.5 Παρουσίαση εντολής Regression .....	64
3.7.6 Το t-test.....	66
3.7.7 Παρουσίαση των αποτελεσμάτων στατιστικής ανάλυσης .....	69
3.8 Παραγωγή γραφημάτων και διαγραμμάτων.....	71
3.8.1 Μονοδιάστατα διαγράμματα .....	71
<b><u>Κεφάλαιο 4ο: Πεδία εφαρμογής Spss.....</u></b>	<b>75</b>

<b><u>4.1 Βασικά στοιχεία ανάλυσης με το Spss.....</u></b>	<b><u>75</u></b>
<b><u>Κεφάλαιο 5ο: Εφαρμογή στο Spss.....</u></b>	<b><u>77</u></b>
<u>5.1 Προετοιμασία.....</u>	<u>77</u>
<u>5.2 Στατιστική Ανάλυση παραδείγματος.....</u>	<u>80</u>
<b><u>Βιβλιογραφία .....</u></b>	<b><u>103</u></b>
<b><u>Ηλεκτρονική Βιβλιογραφία.....</u></b>	<b><u>105</u></b>

## Περιεχόμενα εικόνων

Εικόνα 3.1.1: Εκκίνηση Spss.....	27
Εικόνα 3.1.2: Παράθυρο επεξεργασίας δεδομένων.....	28
Εικόνα 3.1.3: Παράθυρο προβολής μεταβλητών .....	29
Εικόνα 3.1.4: Παράθυρο του Data Editor.....	30
Εικόνα 3.2.2.1: Παράθυρο του Output Navigator .....	32
Εικόνα 3.2.3.1: Παράθυρο του Pivot Table Editor .....	33
Εικόνα 3.2.5.1: Παράθυρο του Chart Editor.....	34
Εικόνα 3.3.1: Μενού File.....	35
Εικόνα 3.3.2: Μενού Edit .....	36
Εικόνα 3.3.3: Μενού Data .....	36
Εικόνα 3.3.4: Μενού Transform .....	37
Εικόνα 3.3.5: Μενού Analyze .....	38
Εικόνα 3.3.6: Μενού Graphs .....	38
Εικόνα 3.3.7: Μενού Utilities .....	39
Εικόνα 3.3.8: Μενού Add-ons .....	39
Εικόνα 3.3.9: Μενού Window .....	40
Εικόνα 3.3.10: Μενού Help .....	41
Εικόνα 3.5.1: Αποθήκευση αρχείου δεδομένων .....	43
Εικόνα 3.6.1.1: Επιλογή τύπου μεταβλητής .....	46
Εικόνα 3.6.1.2: Καθορισμός μεταβλητής Date (Ημερομηνία) στο πλαίσιο διαλόγου Variable Type (Είδος Μεταβλητής) .....	49
Εικόνα 3.7.1.1: Παράθυρο εντολής Frequencies .....	54
Εικόνα 3.7.1.2: Παράθυρο επιλογών εντολής Frequencies .....	55

Εικόνα 3.7.1.3 : Παράθυρο γραφημάτων εντολής Frequencies .....	56
Εικόνα 3.7.2.1: Παράθυρο εντολής Descriptives .....	58
Εικόνα 3.7.3.1: Παράθυρο εντολής Crosstabs .....	60
Εικόνα 3.7.4.1: Παράθυρο εντολής Correlate.....	61
Εικόνα 3.7.4.2: Παράθυρο εντολής Crosstabs.....	63
Εικόνα 3.7.5.1: Παράθυρο εντολής Regression .....	65
Εικόνα 2.7.6.1: Το t-τεστ για εξαρτημένα δείγματα.....	67
Εικόνα 3.7.6.2: Το t-τεστ για ανεξάρτητα δείγματα.....	68
Εικόνα 3.7.7.1: Παράθυρο Viewer Outline .....	70
Εικόνα 3.8.1.1.1: Παράθυρο Chart Bar .....	72
Εικόνα 3.8.1.4.1: Παράθυρο επιλογής γραφικής παράστασης.....	74
Εικόνα 3.8.1.4.2: Παράθυρο επιλογής γραφικής παράστασης.....	74
Εικόνα 5.1.1: Μενού μεταβλητών Variable View.....	78
Εικόνα 5.2.1: Επιλογή μεταβλητών στην εντολή Descriptives .....	80
Εικόνα 5.2.2: Επιλογή περιγραφικών στοιχείων στην εντολή Descriptives.....	81
Εικόνα 5.2.3: Αποτελέσματα εντολής Descriptives .....	81
Εικόνα 5.2.4: Στατιστικά στοιχεία κατηγοριών οχήματος .....	82
Εικόνα 5.2.5: Προετοιμασία εντολής Frequencies.....	82
Εικόνα 5.2.6: Αποτελέσματα εντολής Frequencies.....	83
Εικόνα 5.2.7: Κωδικοποίηση μεταβλητών ομαδοποίησης.....	85
Εικόνα 5.2.8: Πίνακας Συχνοτήτων ομαδοποιημένης μεταβλητής SM1.....	86
Εικόνα 5.2.9: Ιστόγραμμα ομαδοποιημένης μεταβλητής SM1 .....	87
Εικόνα 5.2.10: Πίνακας Συχνοτήτων ομαδοποιημένης μεταβλητής M1 .....	88

Εικόνα 5.2.11: Ιστόγραμμα ομαδοποιημένης μεταβλητής M1 .....	88
Εικόνα 5.2.12: Πίνακας Συχνοτήτων ομαδοποιημένης μεταβλητής L1 .	89
Εικόνα 5.2.13: Ιστόγραμμα ομαδοποιημένης μεταβλητής L1 .....	90
Εικόνα 5.2.14: Επιλογή μεταβλητών συσχέτισης.....	91
Εικόνα 5.2.15: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μικρών αυτοκινήτων και τιμής πώλησης .....	91
Εικόνα 5.2.16: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μικρών αυτοκινήτων και δείκτη ανεργίας.....	92
Εικόνα 5.2.17: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μικρών αυτοκινήτων και τιμής καυσίμου .....	92
Εικόνα 5.2.18: Αποτελέσματα εντολής Συσχέτισης των πωλήσεων μεσαίων αυτοκινήτων και τιμής πώλησης.....	93
Εικόνα 5.2.19: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μεσαίων αυτοκινήτων και δείκτη ανεργίας.....	93
Εικόνα 5.2.20: Αποτελέσματα εντολής Συσχέτισης των πωλήσεων μεσαίων αυτοκινήτων και τιμής καυσίμου.....	94
Εικόνα 5.2.21: Αποτελέσματα εντολής Συσχέτισης των πωλήσεων μεγάλων αυτοκινήτων και τιμής πώλησης .....	94
Εικόνα 5.2.22: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μεγάλων αυτοκινήτων και δείκτη ανεργίας.....	95
Εικόνα 5.2.23: Αποτελέσματα εντολής Συσχέτισης των πωλήσεων μεσαίων αυτοκινήτων και τιμής καυσίμου.....	95
Εικόνα 5.2.24: Δημιουργία μεταβλητών στο Variable View .....	96
Εικόνα 5.2.25: Κωδικοποίηση μεταβλητών στο Variable View .....	96
Εικόνα 5.2.26: Εισαγωγή τιμών μεταβλητών στο Data View.....	97
Εικόνα 5.2.27: Εισαγωγή τιμών μεταβλητής count στο Data View.....	97
Εικόνα 5.2.28: Επεξεργασία εντολής Weight Cases .....	98



Εικόνα 5.2.29: Πίνακας συχνοτήτων μεταβλητής year .....	98
Εικόνα 5.2.30: Πίνακας συχνοτήτων μεταβλητής megethos.....	99
Εικόνα 5.2.31: Επεξεργασία εντολής Crosstabs .....	100
Εικόνα 5.2.32 Πίνακας αποτελεσμάτων Crosstabs .....	100
Εικόνα 5.2.33: Γράφημα αποτελεσμάτων Crosstabs.....	101

## Περίληψη

Η παρούσα εργασία πραγματεύεται την επεξήγηση των βασικών εντολών στο στατιστικό πακέτο SPSS, την εισαγωγή δεδομένων στο πρόγραμμα, την στατιστική επεξεργασία τους και την ερμηνεία των παραγόμενων αποτελεσμάτων. Στα πλαίσια της πτυχιακής αυτής εργασίας έγινε εκτενής βιβλιογραφική ανασκόπηση σχετικά με την χρήση των βασικών εργαλείων και εντολών που είναι χρήσιμα στη στατιστική επεξεργασία των εισαγόμενων δεδομένων και επισήμανση των βασικών εννοιών στατιστικής που κρίνονται απαραίτητες για την ορθή χρήση του προγράμματος και τη σωστή ερμηνεία των αποτελεσμάτων. Στη συνέχεια, μελετήθηκαν τα κυριότερα πεδία εφαρμογής του στατιστικού πακέτου σε διάφορους τομείς και τέλος, πραγματοποιήθηκε μια εφαρμογή των βασικών εντολών σε δεδομένα που αντλήθηκαν από το διαδίκτυο. Για την ποσοτική ανάλυση των δεδομένων χρησιμοποιήθηκε το SPSS 20.

## **Abstract**

This paper deals with the explanation of the basic commands of the statistical package called SPSS, the state of entering data into the program, their statistical process and the interpretation of the output. As far as this thesis is concerned, an extensive literature review on the use of the basic tools and commands that are useful in statistical processing of the input of data were taken into consideration, as also the identifying key statistical concepts that are needed to the proper use of the program and the correct interpretation of the results. Then, the main fields of the application of the statistical package across different sectors were recognized and finally an application of basic the commands to data that were sourced from the internet, took place. For the quantitative data analysis the SPSS 20 was used.

## Εισαγωγή

Η ευρεία χρήση και εφαρμογή των στατιστικών μεθόδων, καθώς και η εκτεταμένη ανάπτυξη των πληροφοριακών συστημάτων και των Ηλεκτρονικών Υπολογιστών (Η/Υ), αποτέλεσαν την αρχή για την ανάπτυξη ενός αρκετά μεγάλου αριθμού εξειδικευμένων προϊόντων στατιστικού λογισμικού. Έτσι, πακέτα όπως το SPSS (Statistical Package for Social Sciences), το Statgraphics, το Statistica, το Splus, το SAS κ.ά εμφανίστηκαν στη διεθνή αλλά και στην ελληνική αγορά λογισμικού, προσφέροντας νέες δυνατότητες στην προσέγγιση της στατιστικής επιστήμης και στη διαδικασία υποστήριξης αποφάσεων με Η/Υ, ακόμη και σε αυτούς που δεν έχουν μνηθεί σε βάθος στη στατιστική θεωρία.

Το στατιστικό πακέτο SPSS αποτελεί προϊόν της εταιρείας SPSS Inc. και είναι ένα από τα πλέον αξιόπιστα συστήματα λογισμικού που υποστηρίζουν στατιστικές μεθόδους. Περιλαμβάνει ένα πλήρες σύνολο εργαλείων στατιστικής ανάλυσης και παρουσίασης δεδομένων, που μπορεί να αξιοποιηθεί με απόλυτη επιτυχία σε ένα ευρύ φάσμα εφαρμογών όπως στην επιστημονική έρευνα, στην έρευνα αγοράς, στη διοίκηση επιχειρήσεων, στο σχεδιασμό και τις προβλέψεις, στη βελτίωση της ποιότητας των προϊόντων, στην επεξεργασία και παραγωγή αναφορών από βάσεις δεδομένων, κ.ά.

Η πρώτη εμφάνιση του SPSS έγινε πριν από 20 χρόνια περίπου και είχε σχέση με την υποστήριξη ερευνητικών προβλημάτων από το πεδίο των κοινωνικών επιστημών. Η ανάπτυξη βέβαια έγινε αρχικά σε μεγάλα υπολογιστικά συστήματα (mainframes). Με την τεράστια εξέλιξη των προσωπικών υπολογιστών και με την παράλληλη επίλυση προβλημάτων ταχύτητας και χώρου αποθήκευσης, αναπτύχθηκαν σταδιακά νέες κατάλληλες εκδόσεις για κάθε τύπο λειτουργικού συστήματος (όπως το SPSS/PC+ για το DOS και το SPSS for Windows) που στη συνέχεια συμπληρώθηκαν και επεκτάθηκαν.

Σήμερα το SPSS λειτουργεί στο σύγχρονο περιβάλλον των Windows και αποτελεί ένα πανίσχυρο και εύχρηστο εργαλείο που μπορεί να καλύψει κάθε ανάγκη στατιστικής επεξεργασίας.

Τα ιδιαίτερα χαρακτηριστικά του πακέτου, που το έχουν καταστήσει ένα από τα πλέον δημοφιλή στο χώρο του, είναι το ευρύ φάσμα στατιστικών μεθόδων που διαθέτει από τις πιο απλές μέχρι τις πιο εξειδικευμένες, η ευκολία διαχείρισης δεδομένων και αρχείων, η δημιουργία πινάκων και γραφικών και τέλος, η δυνατότητα συνεργασίας με όλα τα γνωστά και δημοφιλή προγράμματα της αγοράς.

# Κεφάλαιο 1: Εισαγωγή στο στατιστικό πακέτο SPSS

## 1.1 Τι είναι το SPSS

Το SPSS (Statistical Package for Social Sciences) είναι ένα ολοκληρωμένο στατιστικό πακέτο και μπορεί κανείς να το κατατάξει στην ίδια κατηγορία με τα SAS, GENSTAT, STATISTICA, NAG και STATGRAPHICS. Οι σύγχρονες τεχνολογικές εξελίξεις στους υπολογιστές έχουν επιτρέψει τη χρήση του λογισμικού σε κάθε είδους υπολογιστή. Αρχικά, αναπτύχθηκε στο Πανεπιστήμιο Stanford της Καλιφόρνιας το 1965 για την επίλυση στατιστικών προβλημάτων, που αφορούν τις κοινωνικές επιστήμες. Στη συνέχεια, συμπληρώθηκε με στατιστικά εργαλεία που χρησιμοποιούνται σε διάφορους επιστημονικούς και επαγγελματικούς τομείς με αποτέλεσμα σήμερα να χρησιμοποιείται σχεδόν σε κάθε είδους στατιστική ανάλυση και να έχει καταξιωθεί ως ένα από τα φιλικότερα και αποτελεσματικότερα στατιστικά προγράμματα.

Για να υποδηλώσει τη χρήση του προγράμματος σε όλους τις τομείς που έχουν σχέση με τη στατιστική ανάλυση δεδομένων, το πρόγραμμα ονομάστηκε Superior Performance Software System (σύστημα λογισμικού μέγιστης παραγωγικότητας). Με το SPSS για Windows έγινε ένα μεγαλύτερο βήμα προς τα εμπρός. Αυτή η έκδοση του SPSS έχει όλες τις δυνατότητες της έκδοσης για τους μεγαλύτερους υπολογιστές.

Η πρώτη έκδοση του πακέτου κυκλοφόρησε το 1965 και έτρεχε σε μεγάλα υπολογιστικά συστήματα. Στη συνέχεια, στις αρχές τις δεκαετίας του 80 έγινε προσαρμογή του πακέτου για μικροϋπολογιστές. Η πρώτη έκδοση του πακέτου για μικροϋπολογιστές ήταν η SPSS/PC+ ver. 1.0. και έτρεχε σε περιβάλλον λειτουργικού συστήματος MS-DOS. Στη συνέχεια, ακολούθησαν τέσσερις εκδόσεις σε περιβάλλον λειτουργικού συστήματος MS-DOS, οι οποίες περιελάμβαναν νέες στατιστικές επεξεργασίες, νέες δυνατότητες ελέγχου του περιβάλλοντος εργασίας και των βημάτων, που ακολουθούνται σε μία στατιστική ανάλυση και νέες δυνατότητες βοήθειας στην εξήγηση και στην επιλογή των εντολών του SPSS/PC+.

Στις αρχές της δεκαετίας του 90 με την έλευση του γραφικού περιβάλλοντος των παραθύρων (Windows 3.0, Windows 3.1) και στη συνέχεια με την εμφάνιση του νέου λειτουργικού συστήματος Windows 9, το SPSS εμφανίζει εκδόσεις, οι οποίες

μπορούν και τρέχουν θαυμάσια κάτω από το παραθυρικό περιβάλλον. Το SPSS για Windows γίνεται πακέτο πολύ φιλικό προς τον χρήστη, επειδή όπως όλες οι παραθυρικές εφαρμογές λειτουργεί με ένα σύστημα μενού και πλαισίων διαλόγου (dialogs boxes). Επίσης, είναι πολύ πιο συμπληρωμένο σε σχέση με τις εκδόσεις του πακέτου σε περιβάλλον MS-DOS. Έτσι δημιουργώντας την έκδοση για τα Windows το SPSS κάνει την ανάλυση δεδομένων διαθέσιμη στις μάζες δηλαδή ότι έκανε ο Bill Gates για τους υπολογιστές.

Στο παραθυρικό περιβάλλον, η πρώτη έκδοση που κυκλοφόρησε ήταν το SPSS για τα Windows ver. 5.0 και στην συνέχεια ακολούθησαν μέχρι σήμερα 8 εκδόσεις. Η τελευταία έκδοση είναι η 22.0. Οι σημειώσεις καλύπτουν τη προτελευταία έκδοση 20.0 του SPSS για τα Windows, αλλά παραμένουν κατάλληλες και για τις παλαιότερες εκδόσεις. Αυτό οφείλεται στο γεγονός ότι η εταιρεία κατασκευής του πακέτου φροντίζει, ώστε σε κάθε νέα έκδοση ο βασικός κορμός του πακέτου να παραμένει σταθερός και απλά να συμπληρώνονται καινούργιες στατιστικές διαδικασίες.

Στην Ελλάδα το στατιστικό πακέτο SPSS για Windows υποστηρίζεται από την SPSS Hellas, η οποία είναι αποκλειστική αντιπρόσωπος των προϊόντων SPSS εκτός από την Ελλάδα στην Κύπρο και στη Μάλτα. Όπως ήδη έχει αναφερθεί, το SPSS για Windows καλύπτει ένα ευρύ πεδίο στατιστικών εφαρμογών και επεξεργασίας δεδομένων όλων των κατηγοριών και επιπέδων, τα οποία χωρίζονται σε επιμέρους τμήματα (modules).

Τα τρία κύρια τμήματα του SPSS, που αναφέρονται στη συνέχεια και είναι τα εξής :

**1. Βασικό τμήμα (Base Module):** Αυτό περιλαμβάνει τεχνικές διαχείρισης και επεξεργασίας δεδομένων. Μία σειρά απλών στατιστικών αναλύσεων, όπως κατανομές συχνοτήτων, πίνακες διπλής, τριπλής εισόδου και παρουσίασης αποτελεσμάτων.

**2. Επαγγελματικό Τμήμα (Professional Module):** Αυτό περιλαμβάνει στατιστικές τεχνικές παραμετρικής (π.χ. t-test, ανάλυση της διακύμανσης (Analyssis of Variance), συσχέτιση (Correlation) και παλινδρόμηση (Regression)) και μη παραμετρικής ανάλυσης (π.χ. Wilcoxon, XP , Friedman, Mann-Whitney και Kruskal –Wallis) και

**3. Προχωρημένο Τμήμα (Advanced Statistics):** Αυτό περιλαμβάνει πολύπλοκες στατιστικές αναλύσεις (εφαρμογές), όπως, παραγοντική ανάλυση (Factor analysis), ανάλυση κατά συστάδες (Cluster analysis), μη γραμμική παλινδρόμηση, διακριτική ανάλυση (Discriminant analysis) κλπ.

Επίσης εκτός από τα τρία κύρια τμήματα το SPSS για Windows περιλαμβάνει και άλλα, όπως:

**1. Τμήμα Χρονολογικών Σειρών (Trend Module):** Αυτό περιλαμβάνει τεχνικές επεξεργασίας και ανάλυσης δεδομένων σε χρονολογικές σειρές.

**2. Τμήμα Παραγωγής Πινάκων (Table Module):** Με αυτό μπορεί κανείς να δημιουργήσει πίνακες απλούς ή περίπλοκους, μίας, δύο ή τριών διαστάσεων έτοιμους προς παρουσίαση ή δημοσίευση.

**3. Τμήμα των Γραφικών (Graphic Module):** Αυτό περιλαμβάνει περισσότερα από 50 είδη διαγραμμάτων δισδιάστατα ή τρισδιάστατα σε μορφή **ραβδογράμματος (bar chart)**, **ιστογράμματος**, **γραμμικών διαγραμμάτων (line chart)**, **διαγραμμάτων διασποράς (scatterplots)** κλπ. Για την δημιουργία και επεξεργασία αυτών των διαγραμμάτων το SPSS για Windows διαθέτει ειδικό ενσωματωμένο συντάκτη διαγραμμάτων και υποστηρίζει όλα τα γνωστά πρότυπα της Microsoft για το περιβάλλον των Windows, όπως Dynamic Data Exchange (DDE), Object Linking and Embedding (OLE) , Object Database Connectivity (ODBC), τα οποία επιτρέπουν να μοιράζεται κανείς τα δεδομένα και τα διαγράμματα με άλλα πακέτα, που τρέχουν σε παραθυρικό περιβάλλον.

**4. Τμήμα Κατηγορικών Δεδομένων (Categories Module):** Αυτό περιλαμβάνει στατιστικές τεχνικές για την ανάλυση δεδομένων έρευνας αγοράς.

**5. Τμήμα για Εισαγωγή δεδομένων (Data Entry):** Με αυτό μπορεί κάποιος να δημιουργήσει φόρμες για την εισαγωγή των δεδομένων σε αρχεία μορφής δεδομένων SPSS για Windows, δηλαδή σε αρχεία SAV.

**6. Τμήμα Επεξεργασίας Δεδομένων Ποιοτικού Ελέγχου (QI Analyst Module):** Αυτό το τμήμα περιλαμβάνει στατιστικές τεχνικές για δεδομένα προερχόμενα από το χώρο του ποιοτικού ελέγχου.

**7. Τμήμα δημιουργίας εξελιγμένων Αναφορών (Report Writer):** Με αυτό μπορεί κανείς να δημιουργήσει εύκολα και γρήγορα επαγγελματικές αναφορές.



## 1.2 Βασικές προδιαγραφές υλικού και μνήμης

Για να λειτουργήσει το SPSS για Windows χρειάζεται ένας υπολογιστής συμβατός με IBM και μνήμη RAM τουλάχιστον 64MB. Θα πρέπει στον υπολογιστή να έχουν εγκατασταθεί τα Microsoft Windows ως λειτουργικό σύστημα. Ο υπολογιστής θα πρέπει να διαθέτει μια μονάδα δισκετών 3.5 ιντσών, ένα σκληρό δίσκο και μία μονάδα CD-ROM.

Επίσης πρέπει εκτός από το πληκτρολόγιο να υπάρχει ένα ποντίκι, ώστε η αποδοτικότητά να είναι αυξημένη. Προαιρετικά πρέπει να έχει εγκατασταθεί ένας εκτυπωτής (printer) ή ένας σχεδιογράφος (plotter). Το SPSS για Windows υποστηρίζει σχεδόν όλους τους εκτυπωτές και τους σχεδιογράφους, οι οποίοι είναι συμβατοί με τα Microsoft Windows.<sup>1</sup>

---

<sup>1</sup> Νέλλας Ε., «Ανάλυση Δεδομένων με χρήση του πακέτου SPSS/PC για WINDOWS», Τμήμα Αγροτικής Οικονομίας και Ανάπτυξης, Γεωπονικό Πανεπιστήμιο Αθηνών, Αθήνα, 2005

## Κεφάλαιο 2ο: Βασικές Έννοιες Στατιστικής

### 2.1 Η έννοια της Στατιστικής Επιστήμης

Ας ξεκινήσουμε, δίνοντας έναν ακόμη ορισμό της **Στατιστικής Επιστήμης**, που πρόκειται για μία επιστήμη με αντικείμενο τη συλλογή, οργάνωση, παρουσίαση, ανάλυση και ερμηνεία δεδομένων, με στόχο την εξαγωγή συμπερασμάτων που ερμηνεύουν διάφορα προβλήματα.

Η στατιστική ακολουθεί την **επαγωγική λογική**. Δηλαδή, επιχειρεί την εξαγωγή συμπερασμάτων για ένα σύνολο ύστερα από παρατήρηση ενός μέρους του συνόλου. Με άλλα λόγια, ο ερευνητής που αντιμετωπίζει ένα πρόβλημα ή θέτει ένα ερώτημα προς απάντηση, πρέπει να πάρει κάποιες μετρήσεις των ιδιοτήτων (χαρακτηριστικών) από το σύνολο των ατόμων στο οποίο αναφέρεται η έρευνά του. Τις περισσότερες όμως φορές η καταγραφή των ιδιοτήτων ολόκληρου του συνόλου είναι αδύνατη. Έτσι υπάρχει η ανάγκη ενός επαγωγικού εργαλείου, μιας διαδικασίας δηλαδή, η οποία να καταγράφει μετρήσεις σε ένα αντιπροσωπευτικό υποσύνολο του αρχικού συνόλου και στη συνέχεια χρησιμοποιώντας τις μετρήσεις αυτές να εξάγει συμπεράσματα για το αρχικό σύνολο.

### 2.2 Η έννοια της Μεταβλητής

Με τον όρο «μεταβλητή» ονομάζουμε κάθε τι το οποίο μεταβάλλεται, ποικίλλει ή παραλλάσσει. Και αυτό το «κάθε τι» μπορεί να είναι ένα χαρακτηριστικό, μια ιδιότητα, μια ικανότητα, ή τέλος ένας παράγοντας που μας ενδιαφέρει στο πλαίσιο της κοινωνικής έρευνας που διεξάγουμε. Θέλουμε για παράδειγμα, να μετρήσουμε το δείκτη ευφυΐας κάποιων ατόμων, να παρακολουθήσουμε την εξέλιξη ασθενών, να καταγράψουμε τις στάσεις και γενικά την πολιτική συμπεριφορά κάποιων κοινωνικών ομάδων, να αξιολογήσουμε τις ικανότητες κάποιων μαθητευόμενων ατόμων, να μετρήσουμε το χρόνο αντίδρασης σε κάποιο ερέθισμα, να καταγράψουμε τη θέληση του εκλογικού σώματος κτλ. Και όλα αυτά, φυσικά επιβάλλουν να εισαχθούμε σε μια διαδικασία μέτρησης με τη χρήση μεταβλητών. Πρακτικά, όταν κάνουμε λόγο για μια μεταβλητή, ξεχωρίζουμε σε αυτήν δύο πράγματα: το όνομα και την τιμή της. Για παράδειγμα όταν δηλώνουμε :  $V1 = 3$  είναι

σαφές ότι πρόκειται για μια μεταβλητή της οποίας το όνομα είναι V1 και η τιμή που τώρα της αποδίδεται, είναι 3.

## 2.3 Είδη Μεταβλητών

Ανάλογα με το πώς μεταβάλλεται μια μεταβλητή διακρίνεται και το είδος στο οποίο ανήκει. Έτσι, αν μεταβάλλεται από την άποψη της ποιότητας τη διακρίνουμε σε ποιοτική (qualitative), ενώ αν μεταβάλλεται από την άποψη της ποσότητας τη διακρίνουμε σε ποσοτική (quantitative).

- Στις **ποιοτικές** μεταβλητές κατατάσσονται το φύλο, (άνδρας ή γυναίκα), η θρησκεία (Χριστιανός, Μουσουλμάνος, κτλ.), ο τόπος διαμονής (αστικός, ημιαστικός, αγροτικός), η οικογενειακή κατάσταση (έγγαμος, άγαμος, διαζευγμένος, χήρος), το μορφωτικό επίπεδο (απόφοιτος Δημοτικού, Γυμνασίου, Λυκείου, Πανεπιστημίου κτλ.), η πολιτική τοποθέτηση (δεξιά, κέντρο, αριστερά) καθώς και πλήθος άλλων μεταβλητών οι οποίες συγκροτούν κατηγορίες κατά το δυνατόν ευδιάκριτες και οπωσδήποτε αμοιβαία αποκλειόμενες, αν η μεταβλητή διαθέτει δύο επίπεδα (διχοτομική).
- Στις **ποσοτικές** μεταβλητές κατατάσσονται η ηλικία, το βάρος, το ύψος, το εισόδημα, ο βαθμός ευφυΐας κτλ. Στην περίπτωση των ποσοτικών μεταβλητών αυτό που αναζητούμε επειγόντως είναι η μονάδα και μάλιστα η κατάλληλη μονάδα μέτρησης. Για παράδειγμα, ένας φοιτητής πήρε 80 μονάδες στα Μαθηματικά και 65 μονάδες στη Βοτανική. Διόλου απίθανο ο συγκεκριμένος φοιτητής να είναι καλύτερος στη Βοτανική. Εξαρτάται από το είδος των χρησιμοποιούμενων μονάδων. Και για να είμαστε πιο ακριβείς το ερώτημα πού είναι καλύτερος ο εν λόγω φοιτητής, θα το απαντήσουμε σωστά μόνο αν ανάγουμε τις μονάδες των Μαθηματικών και της Βοτανικής σε ένα ενιαίο (και άρα αντικειμενικό ως ένα βαθμό) σύστημα μέτρησης. Οι ποσοτικές μεταβλητές διακρίνονται παραπέρα σε συνεχείς (continuous) και σε διακριτές (discrete).
- Μια **συνεχής ποσοτική** μεταβλητή συνήθως κινείται ανάμεσα σε μια ελάχιστη και σε μια μέγιστη τιμή από το σύνολο των πραγματικών αριθμών (real numbers). Αυτό που τη χαρακτηρίζει «συνεχή» είναι η ικανότητά της να μπορεί να λάβει κάθε τιμή ανάμεσα στην ελάχιστη και στη μέγιστη τιμή της. Για παράδειγμα, το βάρος ενός πειραματόζωου είναι δυνατόν να είναι 2 κι-

λά, 45 γραμμάρια και 9 χιλιοστά του γραμμαρίου και κτλ.... Άλλο παράδειγμα, η ηλικία ενός μικροβίου στον αέρα μπορεί να είναι 25 δευτερόλεπτα, 0,3 εκατοστά του δευτερολέπτου κτλ.. Με άλλα λόγια, μπορεί να γίνει λόγος για ατελείωτα δεκαδικά ψηφία και επομένως υπάρχει «συνέχεια» στις τιμές μιας μεταβλητής. Τελειώνοντας με αυτές τις μεταβλητές θα πρέπει να επισημάνουμε κάτι το οποίο, λίγο- πολύ είναι φανερό, από τα προηγούμενα. Δεν υπάρχει μέγιστη ακρίβεια στις μετρήσεις με συνεχείς ποσοτικές μεταβλητές.

- Αντίθετα με τις συνεχείς, οι **διακριτές ποσοτικές** μεταβλητές δεν «κτενίζουν» όλες τις δυνατές τιμές ανάμεσα σε ένα ελάχιστο και σε μέγιστο, αλλά λαμβάνουν ορισμένες, διακριτές ή απαριθμητές τιμές, οι οποίες προκύπτουν από τις μετρήσεις και είναι πάντα ακέραιοι αριθμοί. Στην περίπτωση των διακριτών ποσοτικών μεταβλητών είναι σαφής η ύπαρξη ακρίβειας, σε αντίθεση με της συνεχείς ποσοτικές μεταβλητές.

Ανάλογα με τη θέση της σε ένα στατιστικό μοντέλο μια μεταβλητή είναι δυνατόν να διακριθεί σε ανεξάρτητη ή εξαρτημένη.

- Θα χαρακτηρίζουμε μια μεταβλητή ως **ανεξάρτητη** (independent), όταν αυτή έχει εισαχθεί από τον ερευνητή για να εκτιμηθεί η επίδρασή της πάνω σε μια άλλη και σε άλλες μεταβλητές οι οποίες χαρακτηρίζονται ως εξαρτημένες. Με άλλα λόγια, η ανεξάρτητη μεταβλητή αποτελεί την αιτία και η αυτόνομη δράση της επηρεάζει άλλη ή άλλες μεταβλητές. Σε ορισμένα στατιστικά μοντέλα την ανεξάρτητη μεταβλητή τη συναντάμε και με το όνομα predictor variable, για το λόγο ότι προϋποθέτει ή προβλέπει την τιμή της εξαρτημένης μεταβλητής.
- Ανάλογα, θα χαρακτηρίζουμε μια μεταβλητή ως εξαρτημένη (dependent), όταν η τιμή της επηρεάζεται ή προσδιορίζεται από τις μεταβολές που συμβαίνουν στην ανεξάρτητη ή στις ανεξάρτητες μεταβλητές. Με άλλα λόγια, η εξαρτημένη μεταβλητή εκφράζει το αποτέλεσμα της δράσης άλλης ή άλλων μεταβλητών οι οποίες την επηρεάζουν. Σε ορισμένα στατιστικά μοντέλα η εξαρτημένη μεταβλητή απαντάται και με τον όρο response variable.

## 2.4 Πληθυσμός και δείγμα

Ο όρος **πληθυσμός** (*population*) αναφέρεται σε ολόκληρο το σύνολο των ατόμων για το οποίο ενδιαφερόμαστε. Ένας πληθυσμός μπορεί να είναι *πραγματικός* όπως το σύνολο των ατόμων που γεννήθηκαν το 1990, ή *θεωρητικός*, όπως το σύνολο των βρεφών που γεννήθηκαν ή θα γεννηθούν από μητέρες που καπνίζουν ένα πακέτο τσιγάρα ημερησίως.

Ο όρος **δείγμα** (*sample*) αναφέρεται σε ένα υποσύνολο του πληθυσμού. Οι περισσότερες στατιστικές έρευνες στηρίζονται σε δείγματα αφού, όπως αναφέραμε και προηγουμένως, οι ιδιότητες του πληθυσμού είναι συνήθως αδύνατο να καταγραφούν. Θέλουμε λοιπόν συχνά να εξάγουμε συμπεράσματα για τον πληθυσμό βασιζόμενοι στο δείγμα που έχουμε επιλέξει. Για παράδειγμα, θέλουμε να εκτιμήσουμε το πραγματικό ποσοστό των θεατών ενός αγώνα ποδοσφαίρου που μεταδόθηκε τηλεοπτικά, βασιζόμενοι σε ένα δείγμα ατόμων που ρωτήθηκε σχετικά.

Στο σημείο αυτό είναι ανάγκη να επισημάνουμε τα εξής:

- Το δείγμα θα πρέπει να είναι τυχαίο, πράγμα που σημαίνει ότι κάθε στοιχείο (παρατηρησιακή μονάδα) του πληθυσμού, θα πρέπει να έχει ίσες δυνατότητες (πιθανότητα) να συμπεριληφθεί στο δείγμα. Με άλλα λόγια, η επιλογή ενός στοιχείου του πληθυσμού, είναι ανεξάρτητη από την επιλογή ενός άλλου στοιχείου του πληθυσμού.
- Το δείγμα θα πρέπει να είναι αντιπροσωπευτικό, που σημαίνει ότι το δείγμα θα πρέπει να έχει τα χαρακτηριστικά και τις ιδιότητες του πληθυσμού από τον οποίο προέρχεται. Κι αυτό εξασφαλίζεται στην περίπτωση που έχει επιλεγεί στη βάση των επιστημονικών μεθόδων δειγματοληψίας.
- Ο πληθυσμός θα πρέπει να καθορίζεται στην αρχή της κάθε έρευνας με τη μεγαλύτερη δυνατή ακρίβεια. Μόνο τότε, είναι δυνατή η επιλογή αντιπροσωπευτικού δείγματος.<sup>2</sup>

---

<sup>2</sup> Σωσσίδου Ε., Ψευτογιάννη Δ., «Μεθοδολογία Έρευνας & Στατιστική με τη χρήση του SPSS 13.00 for Windows», Θεσσαλονίκη, 2007

## 2.5 Αριθμητικά Περιγραφικά Μέτρα<sup>3</sup>

Τα **στατιστικά μέτρα** (*statistics - measures*) είναι αριθμοί που υπολογίζονται από τα δεδομένα και η τιμή τους αντιπροσωπεύει κάποια τάση ή συμπεριφορά του δείγματος. Θα μπορούσαμε να πούμε ότι είναι αριθμοί - μέτρα που περιγράφουν το σύνολο του δείγματος με τον ίδιο τρόπο που ένα στερεό σώμα όσο πολύπλοκο και αν είναι στην δομή του μπορεί να περιγραφεί συνοπτικά μόνο από το βάρος του και από τον όγκο του. Τα στατιστικά μέτρα διακρίνονται στα **μέτρα κεντρικής τάσης** (*measures of central tendency*), στα **μέτρα μεταβλητότητας** (*measures of dispersion*) και στα **μέτρα σχήματος** (*measures of shape*). Τις τρεις αυτές κατηγορίες περιγράφουμε αναλυτικά στη συνέχεια.

**(α) Μέτρα κεντρικής τάσης:** Χρησιμοποιούνται για να περιγράψουν τη *θέση* του συνόλου των δεδομένων. Δηλαδή, η τιμή τους είναι η πιο αντιπροσωπευτική για να σχηματίσει κάποιος μία πρώτη εικόνα για το ύψος των τιμών των παρατηρήσεων μίας μεταβλητής. Τα πλέον γνωστά μέτρα κεντρικής τάσης που υπολογίζονται και από το SPSS είναι τρία:

**1. Η μέση τιμή (*mean*).** Γνωστή και σαν **αριθμητικός μέσος όρος**, είναι το άθροισμα των τιμών όλων των παρατηρήσεων διαιρεμένο με το πλήθος των παρατηρήσεων, δηλαδή:

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$$

όπου  $N$  ο αριθμός των περιπτώσεων και  $X_i$  η τιμή της μεταβλητής για την  $i$ -οστή περίπτωση. Η μέση τιμή υπολογίζεται και ερμηνεύεται στατιστικά στις ποσοτικές μεταβλητές. Απαιτείται ιδιαίτερη προσοχή στην περίπτωση ποιοτικών μεταβλητών που οι τιμές τους έχουν κωδικοποιηθεί με αριθμούς, όπου η μέση τιμή δεν έχει νόημα. Εξαίρεση αποτελεί η περίπτωση όπου η ποιοτική μεταβλητή παίρνει δύο μόνο τιμές (π.χ. "ΝΑΙ" και "ΟΧΙ") οι οποίες έχουν κωδικοποιηθεί με τους αριθμούς 0 και 1. Η μέση τιμή σε αυτή την περίπτωση παριστά το ποσοστό των περιπτώσεων που έχουν δηλωθεί σαν 1 στα δεδομένα.

---

<sup>3</sup> Δαφέρμος Β., «Κοινωνική Στατιστική & Μεθοδολογία Έρευνας με το SPSS», Εκδόσεις ΖΗΤΗ, Θεσσαλονίκη, 2011

**2. Η διάμεσος (*median*).** Είναι η τιμή της μεταβλητής για την οποία το 50% των τιμών είναι μεγαλύτερο από αυτή και το υπόλοιπο 50% μικρότερο. Αν το πλήθος των παρατηρήσεων είναι περιττό, η διάμεσος είναι η μεσαία παρατήρηση, όταν αυτές διαταχθούν σε αύξουσα σειρά. Στην περίπτωση που το πλήθος είναι άρτιο, η διάμεσος υπολογίζεται από τη μέση τιμή των δύο μεσαίων παρατηρήσεων. Χρησιμοποιείται αποκλειστικά για ποσοτικές μεταβλητές.

**3. Η επικρατούσα τιμή (*mode*).** Είναι η πιο συχνά εμφανιζόμενη τιμή (ή και τιμές) στο δείγμα. Είναι δυνατόν να υπάρχουν περισσότερες από μία επικρατούσες τιμές, το SPSS όμως εμφανίζει μόνο τη μία από αυτές. Χρησιμοποιείται και για τις ποσοτικές αλλά και για τις ποιοτικές μεταβλητές.

**(β) Μέτρα μεταβλητότητας:** Χρησιμοποιούνται για να περιγράψουν τις διαφορές που υπάρχουν ανάμεσα στις τιμές μίας μεταβλητής. Δηλαδή, προσδιορίζουν αν οι παρατηρήσεις είναι συγκεντρωμένες γύρω από μία αντιπροσωπευτική τιμή ή παρουσιάζουν μεγάλη διασπορά. Τα μέτρα μεταβλητότητας που υπολογίζονται στο SPSS και χρησιμοποιούνται αποκλειστικά για ποσοτικές μεταβλητές είναι:

**1. Το εύρος (*range*).** Είναι η διαφορά ανάμεσα στη **μέγιστη** (*maximum*) και στην **ελάχιστη** (*minimum*) από τις παρατηρούμενες τιμές.

**2. Η διακύμανση ή διασπορά (*variance*).** Είναι το πλέον συνηθισμένο μέτρο μεταβλητότητας. Υπολογίζεται αθροίζοντας τα τετράγωνα των διαφορών από τη μέση τιμή όλων των παρατηρήσεων και στη συνέχεια διαιρώντας το άθροισμα με το πλήθος των παρατηρήσεων ελαττωμένο κατά ένα. Ο τύπος υπολογισμού δηλαδή είναι:

$$s^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}$$

**3. Η τυπική απόκλιση (*standard deviation*).** Είναι η τετραγωνική ρίζα της διακύμανσης. Έχει πιο φανερή σημασία από τη διακύμανση γιατί εκφράζεται σε μονάδες μέτρησης ίδιες με αυτές των παρατηρήσεων, ενώ η διακύμανση στα τετράγωνα των μονάδων αυτών.

(γ) **Μέτρα σχήματος:** Είναι μέτρα που περιγράφουν το σχήμα της μεταβλητής όταν αυτή παρασταθεί με ιστόγραμμα ή ραβδόγραμμα. Στο SPSS υπολογίζονται τα δύο παρακάτω μέτρα σχήματος:

**1. Η λοξότητα (*skewness*).** Για πολλές μεταβλητές το μεγαλύτερο πλήθος των παρατηρήσεων συγκεντρώνεται σε μία κεντρική τιμή. Όσο αυξάνει η απόσταση από την μεσαία αυτή τιμή, μειώνεται η συχνότητα των παρατηρήσεων. Αν αυτή η συμπεριφορά είναι ίδια για τις μεγάλες και για τις μικρές τιμές, η μεταβλητή παρουσιάζει στο σχήμα της μία συμμετρία ως προς την κεντρική τιμή. Τέτοιο παράδειγμα είναι η κανονική κατανομή. Υπάρχουν όμως μεταβλητές που δεν είναι συμμετρικές και εμφανίζουν μία "ουρά" προς τη μία μόνο κατεύθυνση. Αυτές ονομάζονται **λοξές (*skewed*)**. Ανάλογα προς τα πού εμφανίζεται η "ουρά" (μεγάλες ή μικρές τιμές), η μεταβλητή ονομάζεται **θετικά λοξή** (λοξή προς τα δεξιά) ή **αρνητικά λοξή** (λοξή προς τα αριστερά). Αν λοιπόν η λοξότητα που υπολογίζει το SPSS είναι 0, αυτό σημαίνει ότι έχουμε συμμετρική κατανομή συχνοτήτων. Τιμή θετική ή αρνητική δείχνει αντίστοιχα θετικά ή αρνητικά λοξή κατανομή συχνοτήτων.

**2. Η κύρτωση (*kurtosis*).** Είναι ο βαθμός στον οποίο οι παρατηρήσεις συνωστίζονται γύρω από μία κεντρική τιμή. Η κύρτωση της κανονικής κατανομής είναι ίση με 0. Μία τιμή θετική δείχνει κατανομή συχνοτήτων **λεπτόκυρτη** που σημαίνει ότι η γραφική παράσταση της μεταβλητής παρουσιάζει "αιχμηρή" κορυφή και πολλές ακραίες τιμές. Αντίθετα, μία αρνητική τιμή δείχνει κατανομή **πλατύκυρτη** δηλαδή "αμβλεία" κορυφή και σχετικά κοντές ουρές.

## 2.6 Η έννοια της γραμμικής παλινδρόμησης και της συσχέτισης<sup>45</sup>

Με τον όρο "ανάλυση παλινδρόμησης" περιγράφονται όλες οι στατιστικές διαδικασίες που έχουν ως στόχο την ανίχνευση της σχέσης δύο ή περισσότερων μεταβλητών αλλά και την καταγραφή μοντέλου πρόβλεψης. Πιο συγκεκριμένα η ανάλυση παλινδρόμησης προσπαθεί να καταλάβει τον τρόπο που μεταβάλλεται η αναμενόμενη (μέση) τιμή μίας εξαρτημένης μεταβλητής όταν αλλάζει η τιμή μίας ή περισσότερων ανεξάρτητων μεταβλητών. Το αποτέλεσμα μίας τέτοιας διαδικασίας είναι ένα μοντέλο πρόβλεψης της μορφής  $Y = f(X_1, X_2, \dots, X_k)$ .

---

<sup>4</sup> <http://www.de.teipat.gr>

<sup>5</sup> <http://www.lib.teicrete.gr>



Η συνάρτηση  $f$  μπορεί να έχει οποιαδήποτε μορφή ωστόσο ο ερευνητής πρέπει να προσπαθεί να εντοπίζει το απλούστερο δυνατό μοντέλο, δηλαδή ένα της μορφής  $Y = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k$  και να το εγκαταλείπει μόνο όταν αυτό δεν είναι ικανοποιητικό στην πρόβλεψή του.

### 2.6.1 Γραμμική παλινδρόμηση

Το πρώτο βήμα για τη μελέτη της σχέσης δύο συνεχών μεταβλητών είναι ο υπολογισμός του συντελεστή γραμμικής συσχέτισης Pearson. Το επόμενο βήμα είναι η εύρεση της ευθείας γραμμικής παλινδρόμησης της μίας μεταβλητής πάνω στην άλλη, ιδιαίτερα αν μεταξύ των μεταβλητών υπάρχει ισχυρή γραμμική συσχέτιση (συντελεστής μεγαλύτερος κατά απόλυτη τιμή από το 0,7). Η ευθεία γραμμικής παλινδρόμησης είναι η ευθεία η οποία βρίσκεται όσο το δυνατόν “κοντύτερα” στα σημεία του διαγράμματος διασποράς. Το “κοντύτερα” είναι υποκειμενικό και μπορεί να οριστεί με διάφορους τρόπους. Η πλέον συνηθισμένη επιλογή είναι ο εντοπισμός της ευθείας για την οποία το άθροισμα των κάθετων αποστάσεων όλων των σημείων από αυτήν είναι ελάχιστο. Η ευθεία αυτή ονομάζεται και ευθεία ελαχίστων τετραγώνων. Η εξίσωσή της ευθείας ελαχίστων τετραγώνων, όπως και η εξίσωση κάθε ευθείας έχει τη μορφή  $Y = \alpha \cdot X + \beta$ . Στην παραπάνω εξίσωση το  $X$  ονομάζεται ανεξάρτητη μεταβλητή (independent variable) ενώ το  $Y$  εξαρτημένη (dependent variable) διότι οι τιμές του λαμβάνονται από τις τιμές του  $X$ . Οι συντελεστές του γραμμικού μοντέλου υπολογίζονται από τα ζεύγη τιμών των μεταβλητών  $X$  και  $Y$  ως εξής :  $\alpha = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$  και  $\beta = \bar{y} - \alpha \cdot \bar{x}$

Η ευθεία γραμμικής παλινδρόμησης δίνει πρόβλεψη της άγνωστης τιμής της εξαρτημένης μεταβλητής. Η ποιότητα της πρόβλεψης αυτής ελέγχεται από την τιμή του συντελεστή προσδιοριστίας (Coefficient of determination)  $R^2$  ( $R$  – Square) το οποίο είναι καλό να υπολογίζεται μαζί με του συντελεστές της εξίσωσης και το οποίο ερμηνεύεται ως το ποσοστό της μεταβλητότητας των τιμών της εξαρτημένης μεταβλητής που προσδιορίζεται από τις τιμές της ανεξάρτητης μεταβλητής, και είναι καλό να είναι μεγάλο δηλαδή κοντά στη μονάδα.

## 2.6.2 Συντελεστές συσχέτισης

Όπως ήδη έχει αναφερθεί, το πρώτο βήμα στη μελέτη της παλινδρόμησης είναι η κατασκευή του διαγράμματος διασποράς. Στην συνέχεια η μελέτη συνεχίζεται με τον υπολογισμό του συντελεστή συσχέτισης.

Συντελεστής Συσχέτισης καλείται η ποσοτική μέτρηση της έντασης της (γραμμικής) σχέσης μεταξύ δυο μεταβλητών. Ο συντελεστής αυτός καλείται γραμμικός συντελεστής συσχέτισης και συμβολίζεται με  $\rho$  ενώ η αντίστοιχη εκτίμησή του με  $r$ .

Υπάρχει ένα κάποιο πλήθος συντελεστών συσχέτισης. Ο συντριπτικά πιο κοινός από αυτούς είναι ο ) Συντελεστής Συσχέτισης Pearson. Είναι ένα αριθμητικό μέτρο ή δείκτης του μεγέθους της συσχέτισης μεταξύ δύο συνόλων τιμών. Κυμαίνεται σε μέγεθος από +1.00 μέχρι -1.00 περνώντας και από το 0.00.

Το πρόσημο “+“ σημαίνει θετική συσχέτιση — δηλαδή, Οι Τιμές μιας μεταβλητής αυξάνονται όταν αυξάνονται και της άλλης.

Ένα πρόσημο “—“ σημαίνει αρνητική συσχέτιση — δηλαδή, οι τιμές μιας μεταβλητής αυξάνονται καθώς μειώνονται της άλλης.

Συντελεστής συσχέτισης 0.00 σημαίνει μια τέλεια συσχέτιση μεταξύ των δύο μεταβλητών. Με άλλα λόγια, ένα γράφημα διασποράς των δύο μεταβλητών θα δείξει ότι ΟΛΑ τα σημεία προσαρμόζονται απόλυτα σε μια ευθεία γραμμή.

Τιμή 1.00 σημαίνει ότι τα σημεία του γραφήματος διασποράς είναι κατανεμημένα τυχαία γύρω από οποιοδήποτε ευθεία σχεδιαστεί ή εί ναι διατεταγμένα έτσι ώστε να πλησιάζουν κάποια καμπύλη.

Ένας συντελεστής συσχέτισης -0.5 σημαίνει ότι υπάρχει μια μέτρια αρνητική γραμμική σχέση μεταξύ των δύο μεταβλητών.

## Κεφάλαιο 3<sup>ο</sup>: Βασικές Λειτουργίες του SPSS

### 3.1 Εκκίνηση του SPSS για Windows

Για να ξεκινήσει το SPSS για Windows ακολουθείται η εξής διαδικασία:

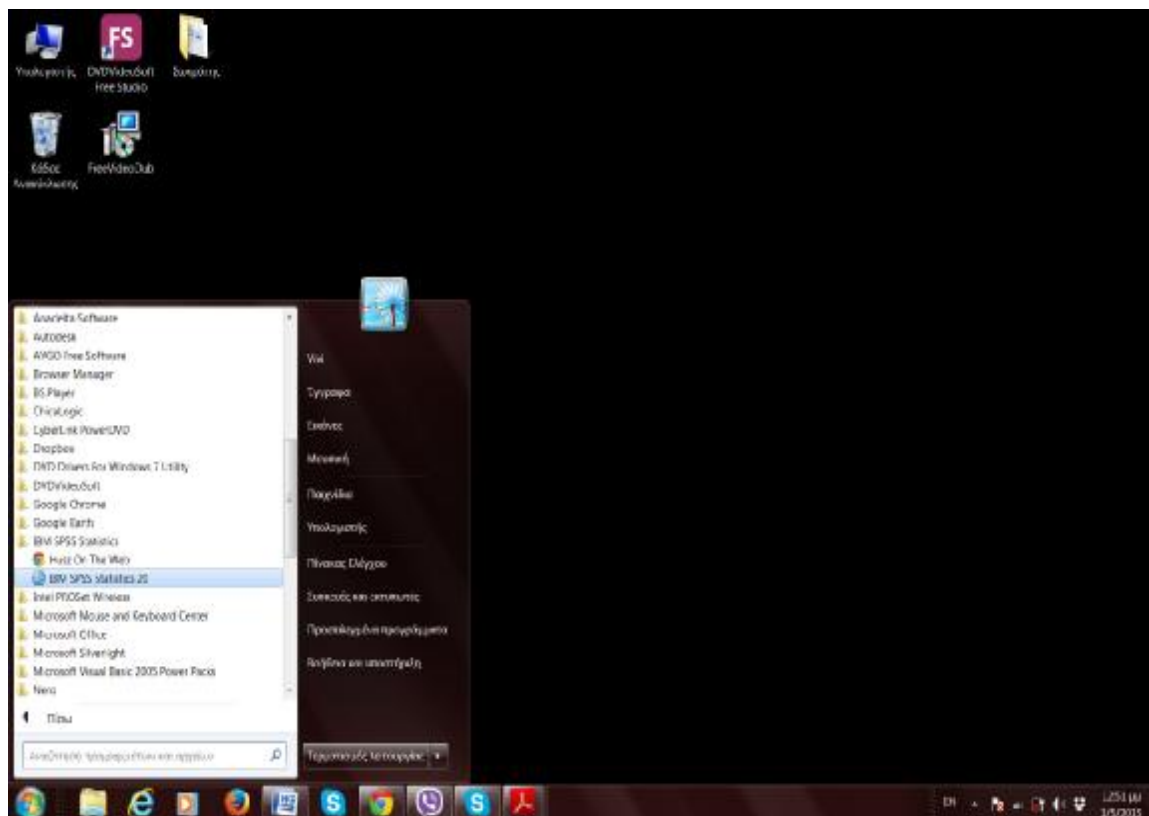
**Βήμα 1<sup>ο</sup>:** Τοποθετείται ο δείκτης του ποντικιού στο κουμπί Έναρξη (Start) της γραμμής εργασιών (TaskBar) των Windows. Κάνοντας κλικ στο αριστερό κουμπί του ποντικιού ανοίγει το μενού Έναρξη (Start).

**Βήμα 2<sup>ο</sup>:** Στην συνέχεια επιλέγονται τα Προγράμματα (Programs)

**Βήμα 3<sup>ο</sup>:** Κάνοντας κλικ στην επιλογή IBM SPSS Statistics ( Εικόνα 3.1.1) και μετά στην επιλογή του δεύτερου μενού IBM SPSS Statistics 20 γίνεται η εκκίνηση του προγράμματος.

Στην περίπτωση που δεν υπάρχει η επιλογή IBM SPSS Statistics πρέπει να αναζητηθεί το αρχείο Spsswin.exe μέσα στην εξερεύνηση των Windows και κάνοντας διπλό κλικ επάνω του, γίνεται η εγκατάσταση και η ενεργοποίηση του.

Εικόνα 3.1.1: Εκκίνηση Spss

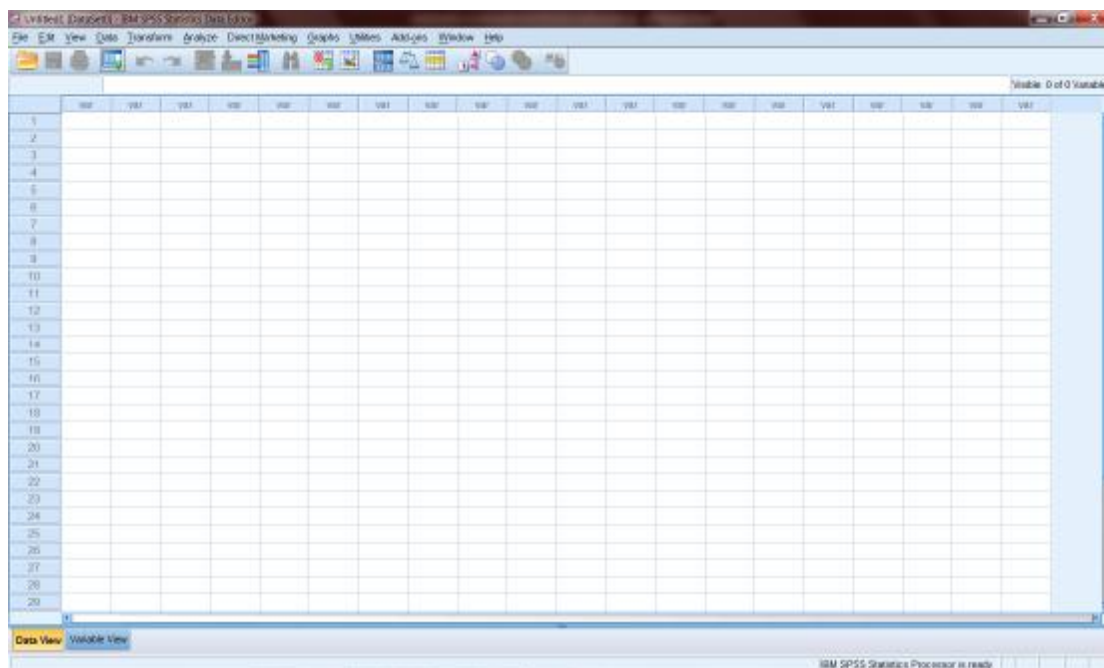


Στη συνέχεια, και ύστερα από λίγα δευτερόλεπτα εμφανίζεται το παράθυρο της Επεξεργασίας Δεδομένων ( **Data Editor** ) που μοιάζει με Spreadsheet (Λογιστικό Φύλλο) και ένα κενό αρχείο το οποίο δεν έχει όνομα και ονομάζεται ενεργό αρχείο. Το παράθυρο της Επεξεργασίας Δεδομένων αποτελείται από δύο τμήματα ή από δύο επιμέρους παράθυρα, τα οποία μπορεί να επιλέξει κάποιος πατώντας στην αντίστοιχη καρτέλα, που βρίσκεται στην κάτω αριστερή γωνία. Τα τμήματα αυτά είναι τα εξής:

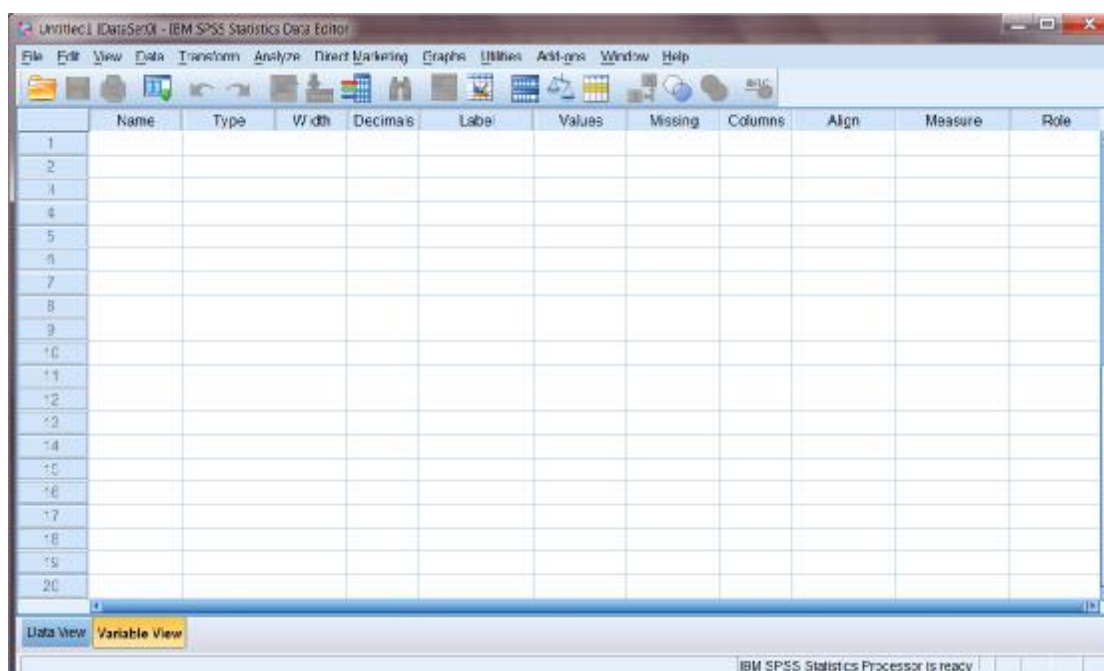
- Το τμήμα το οποίο ονομάζεται **Data View (Προβολή Δεδομένων)**, που μοιάζει με τα λογιστικά φύλλα, και στο οποίο καταχωρούνται τα στατιστικά δεδομένα (εικόνα 3.1.2) και
- Το τμήμα το οποίο περιέχει πληροφορίες για τη δομή των στατιστικών δεδομένων και ονομάζεται **Variable View (Προβολή Μεταβλητών)** (Εικόνα 3.1.3).

Είναι σημαντικό να αναφέρουμε ότι το παράθυρο του Συντάκτη Δεδομένων (Data Editor) διαθέτει τα δύο τμήματα Προβολής Δεδομένων View) από την έκδοση 10 και μετά. Στις προηγούμενες εκδόσεις ο Συντάκτης Δεδομένων (Data Editor) διέθετε μόνο το Data View. Αυτή ίσως είναι η πιο σημαντική αλλαγή που εμφάνισε το Spss for Windows στις τελευταίες του εκδόσεις. (Data View) και Προβολή Μεταβλητών (Variable View).

Εικόνα 3.1.2: Παράθυρο επεξεργασίας δεδομένων



Εικόνα 3.1.3: Παράθυρο προβολής μεταβλητών



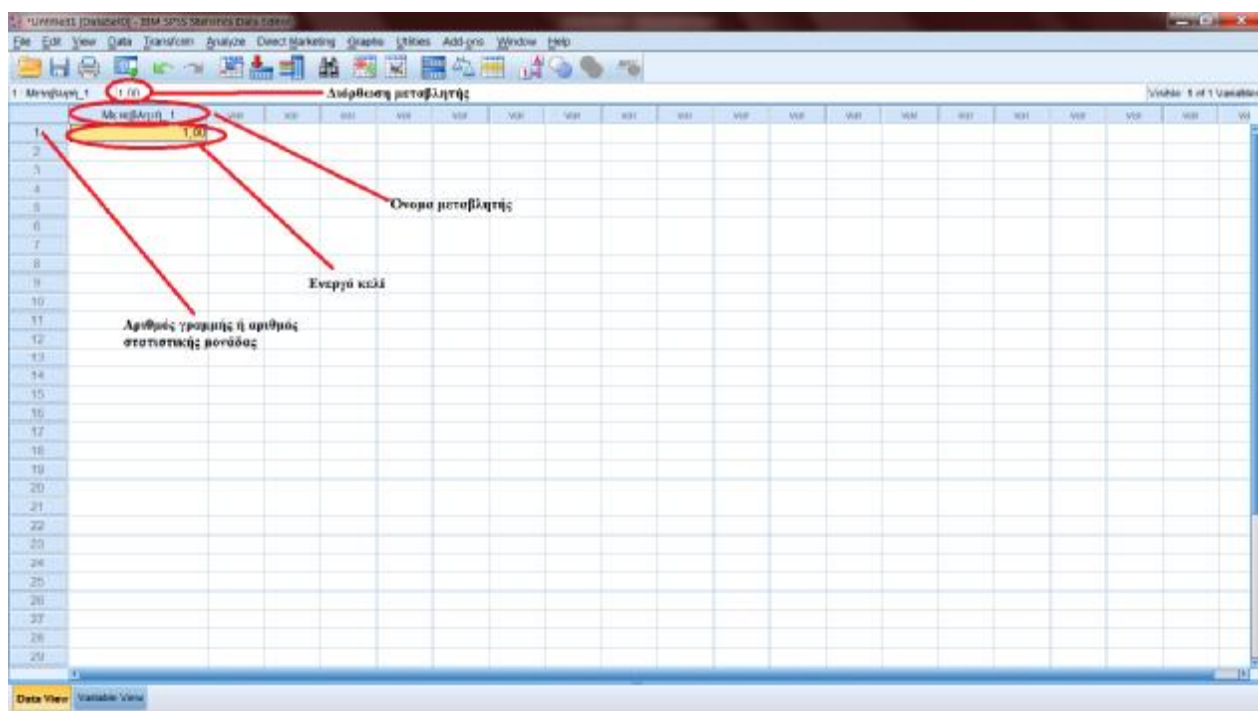
Τα κύρια μέρη του παραθύρου του Data Editor του SPSS είναι τα εξής:

- **Η γραμμή τίτλου**, όπου αναγράφεται ο τίτλος του παραθύρου
- **Η γραμμή μενού**, η οποία αποτελείται από στοιχεία τα οποία υποκρύπτουν μενού με επιλογές
- **Τα κουμπιά αποκατάστασης μεγέθους παραθύρου** με τα οποία γίνεται η μεγιστοποίηση, η ελαχιστοποίηση και το κλείσιμο του παραθύρου
- **Το εικονίδιο μενού ελέγχου**, το οποίο βρίσκεται στο πάνω αριστερό άκρο του παραθύρου και κάνοντας κλικ πάνω του εμφανίζεται ένα μενού που περιλαμβάνει εντολές χειρισμού του παραθύρου
- **Η γραμμή εργαλείων**, η οποία περιέχει μια συλλογή από εργαλεία με τα οποία εκτελούνται με εύκολο τρόπο συχνά χρησιμοποιούμενες εντολές και διαδικασίες
- **Η γραμμή διόρθωσης**, η οποία βρίσκεται κάτω από τις γραμμές των εργαλείων. Σε αυτή την γραμμή εμφανίζονται τα δεδομένα, που εισάγει ο χρήστης καθώς και άλλες χρήσιμες πληροφορίες σχετικά με τα κελιά.
- **Η γραμμή κατάστασης**, στην οποία εμφανίζονται χρήσιμα μηνύματα

- **Οι γραμμές κύλισης(scroll bars)**, οι οποίες περιβάλλουν δεξιά και κάτω δεξιά το παράθυρο. Με κλικ στα βέλη των γραμμών επιτυγχάνεται η μετακίνηση μιας γραμμής ή μιας στήλης κάθε φορά
- **Τα κελιά(cells)**, τα οποία αποτελούν το μεγαλύτερο μέρος του παραθύρου. Η θέση τους προσδιορίζεται από τις τομές των αντίστοιχων γραμμών και στηλών.

Όλα τα ανωτέρω περιγράφονται συνοπτικά στην παρακάτω εικόνα 3.1.4.

Εικόνα 3.1.4: Παράθυρο του Data Editor



## 3.2 Βασικά εργαλεία του SPSS για Windows

Όπως ήδη έχει αναφερθεί το SPSS είναι ένα Στατιστικό Πακέτο, δηλαδή ένα δυναμικό λογισμικό για στατιστική ανάλυση με αυξημένες δυνατότητες στη διαχείριση αρχείων σε γραφικό περιβάλλον.

Η επικοινωνία ενός χρήστη με το πρόγραμμα μπορεί να γίνει, είτε μέσω των μενού και απλών παραθύρων διαλόγου(κυρίως αφορά τον απλό χρήστη), είτε μέσω συγκεκριμένων εντολών (command), οι οποίες μοιάζουν με απλές αγγλικές προτάσεις και είναι εύκολες στην κατανόηση (κυρίως αφορά τον έμπειρο χρήστη).

Το SPSS είναι πλαισιωμένο με μία ομάδα **editors** (**συντακτών**), οι οποίοι διευκολύνουν τον χρήστη σε όλες τις εργασίες (καταχώρηση δεδομένων, διόρθωση και διαγραφή δεδομένων, δημιουργία και επεξεργασία διαγραμμάτων, δημιουργία και επεξεργασία πινάκων). Στη συνέχεια, θα γίνει μια σύντομη παρουσίαση αυτών.

### 3.2.1 Παράθυρο Συντάκτη Επεξεργασίας Δεδομένων (Data Editor)

Όπως, ήδη έχει αναφερθεί χρησιμοποιείται για τον **ορισμό, εισαγωγή και διόρθωση δεδομένων** και ανοίγει με το ξεκίνημα του SPSS για Windows. Στο data editor μπορεί να έχει κανείς ανοικτό μόνο ένα ενεργό αρχείο δεδομένων (**active file**). Το αρχείο αυτό είναι προσωρινό εκτός και εάν ζητήσει κάποιος τη διατήρησή του. Αλλάζει όμως εάν εκτελεσθεί κάποια άλλη εντολή ορισμού-ανάγνωσης αρχείου. Το ενεργό αρχείο περιλαμβάνει τα εξής:

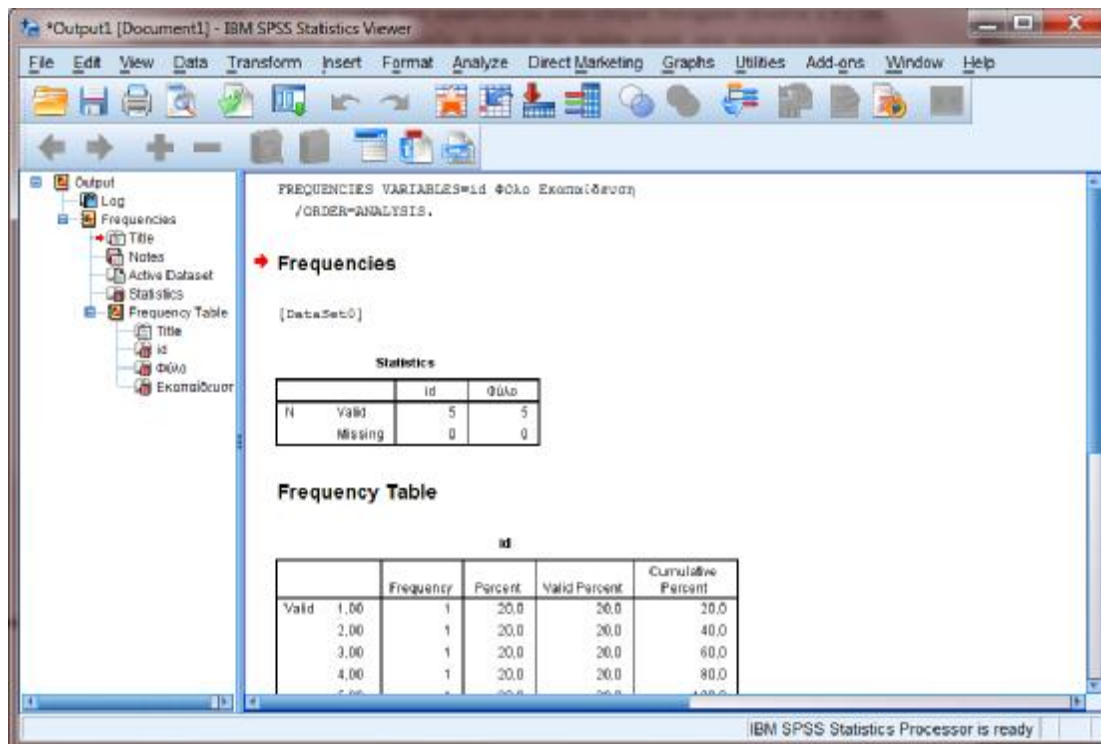
- Την περιγραφή των μεταβλητών (name, label, missing values, κλπ.)
- Τα δεδομένα που το SPSS έχει διαβάσει
- Τους μετασχηματισμούς που ίσως έχει ζητήσει κάποιος. Στο ίδιο ενεργό αρχείο μπορεί κανείς να ορίσει πολλές διαφορετικές διαδικασίες (procedures).

### 3.2.2 Παράθυρο Συντάκτη Επεξεργασίας Αποτελεσμάτων (Output Navigator)

Χρησιμοποιείται για τη **διαχείριση των παραχθέντων αποτελεσμάτων**. Όλα τα αποτελέσματα, πίνακες γραφικά κλπ εμφανίζονται στον Output Navigator (Εικόνα 3.2.2.1) για περαιτέρω επεξεργασία και διαφύλαξη. Ανοίγει την πρώτη φορά ,που

παράγεται κάποιο αποτέλεσμα. Μπορεί όμως να έχει κανείς ανοιχτά περισσότερα του ενός αρχεία αποτελεσμάτων.

Εικόνα 3.2.2.1: Παράθυρο του Output Navigator

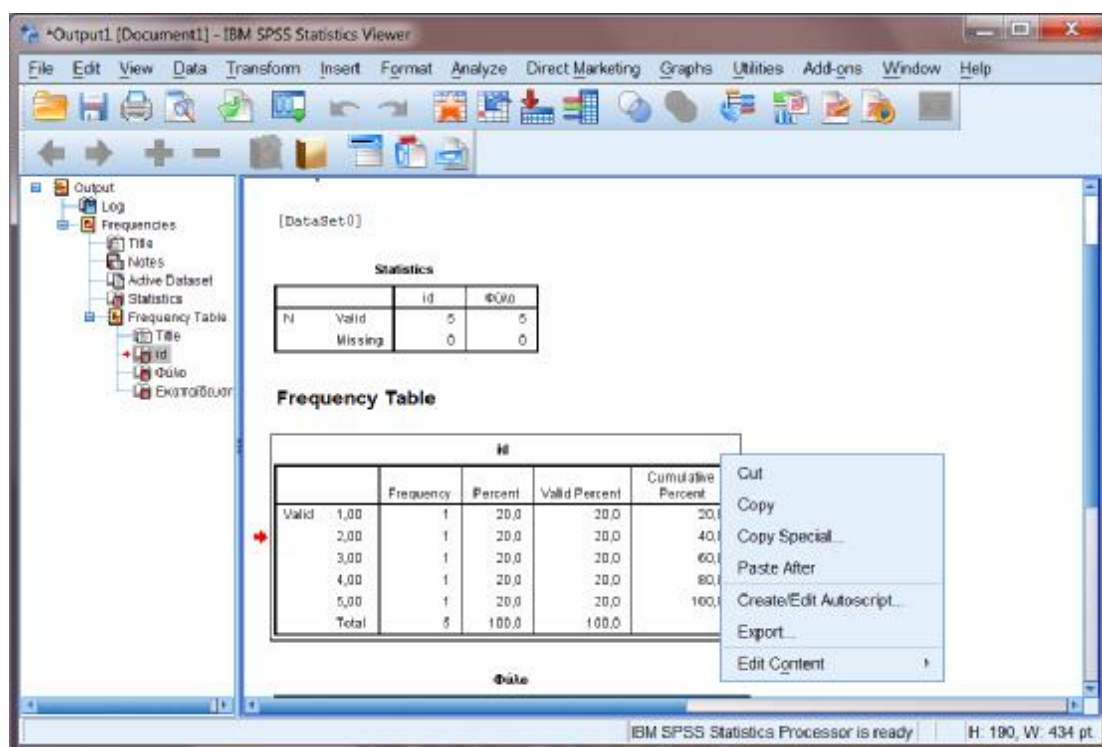


### 3.2.3 Παράθυρο Συντάκτη Επεξεργασίας Πινάκων (Pivot Table Editor)

Χρησιμοποιείται για τη **διαχείριση πινάκων**, που εμφανίζονται στον Output Navigator. Τα αποτελέσματα που εμφανίζονται σε πίνακες μπορούν να τροποποιηθούν πολύ εύκολα με τη χρήση του **Pivot Table Editor**. Με αυτό μπορεί κάποιος να κάνει διάφορες εργασίες, όπως απλές διορθώσεις κειμένου, να αντιμετωπίσει τις γραμμές με τις στήλες του πίνακα με μία κίνηση, να χρωματίσει, να αποκρύψει και να εμφανίσει επιλεκτικά τμήματα του πίνακα κλπ.



Εικόνα 3.2.3.1: Παράθυρο του Pivot Table Editor



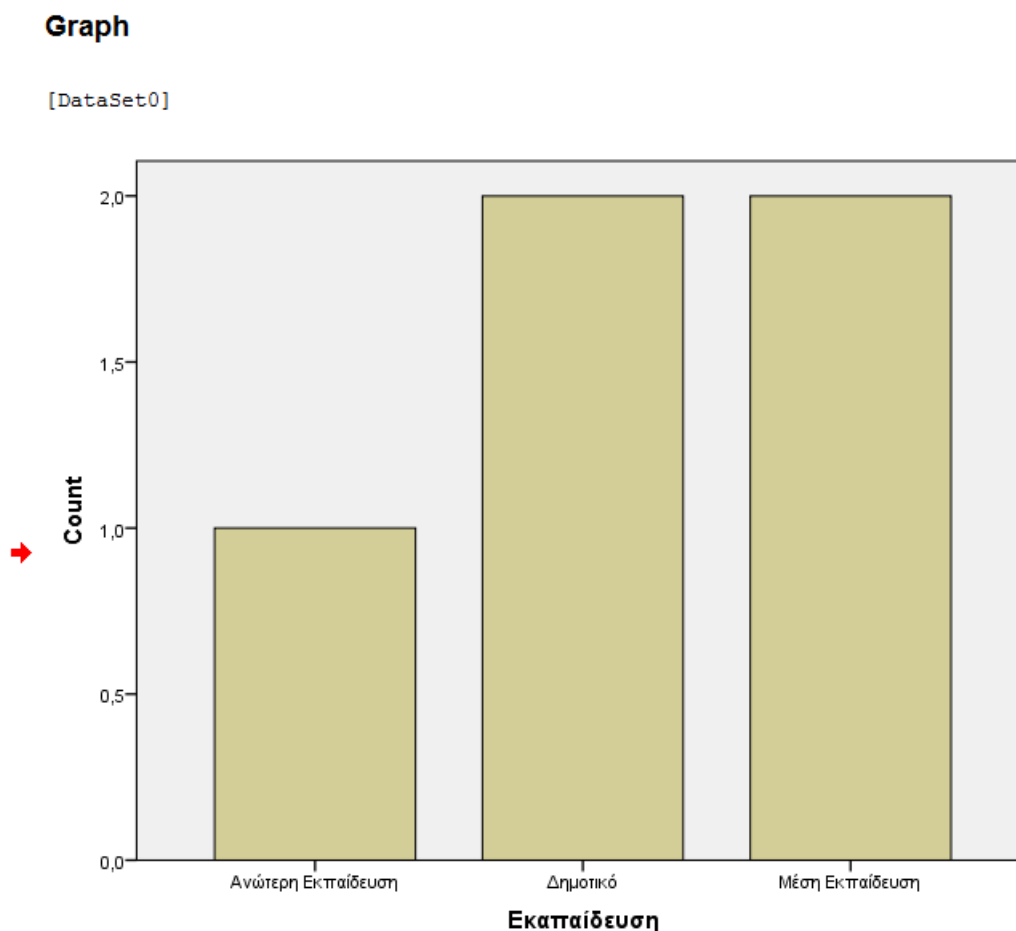
### 3.2.4 Παράθυρο Συντάκτη Επεξεργασίας Κειμένου(Text Output Editor)

Ο **Text Output Editor** χρησιμοποιείται για τη **διαχείριση απλού κειμένου**. Τα απλά κείμενα, που παράγονται έξω από πίνακες μπορούν να τροποποιηθούν με τη βοήθεια του. Με αυτό μπορεί κάποιος να κάνει διάφορες επεμβάσεις στο κείμενο, όπως διορθώσεις, αλλαγές γραμματοσειρών (fonts) κλπ

### 3.2.5 Παράθυρο Συντάκτη Επεξεργασίας Γραφημάτων (Chart Editor)

Ο **Chart Editor** χρησιμοποιείται στη **διαχείριση και καθορισμό διαγραμμάτων και γραφημάτων**. Όπως ήδη έχει αναφερθεί, το Spss για Windows μπορεί να παράγει μεγάλη ποικιλία διαγραμμάτων και γραφημάτων, τα οποία μπορούν να τροποποιηθούν με την χρήση του Chart Editor. Μπορεί κανείς να αλλάξει χρώματα, γράμματα, άξονες αλλά και τύπο γραφήματος.

Εικόνα 3.2.5.1: Παράθυρο του Chart Editor



### 3.2.6 Συντάκτης Εντολών (Syntax Editor)

Ο **Syntax Editor** χρησιμοποιείται για τη **διαχείριση των εντολών (commands)** του **SPSS**. Οι διαδικασίες που επιλέγονται μέσα από τα μενού μπορεί να παρουσιάζονται στο Syntax Editor με μορφή σειράς εντολών του Spss πατώντας απλά στο κουμπί Paste (Επικόλληση). Στη συνέχεια, μπορεί κανείς να τροποποιήσει τις εντολές ή ακόμα να γράψει νέες, που δεν έχει τη δυνατότητα να τις επιλέξει από τα μενού. Οι βασικοί κανόνες σύνταξης των εντολών του SPSS είναι οι εξής:

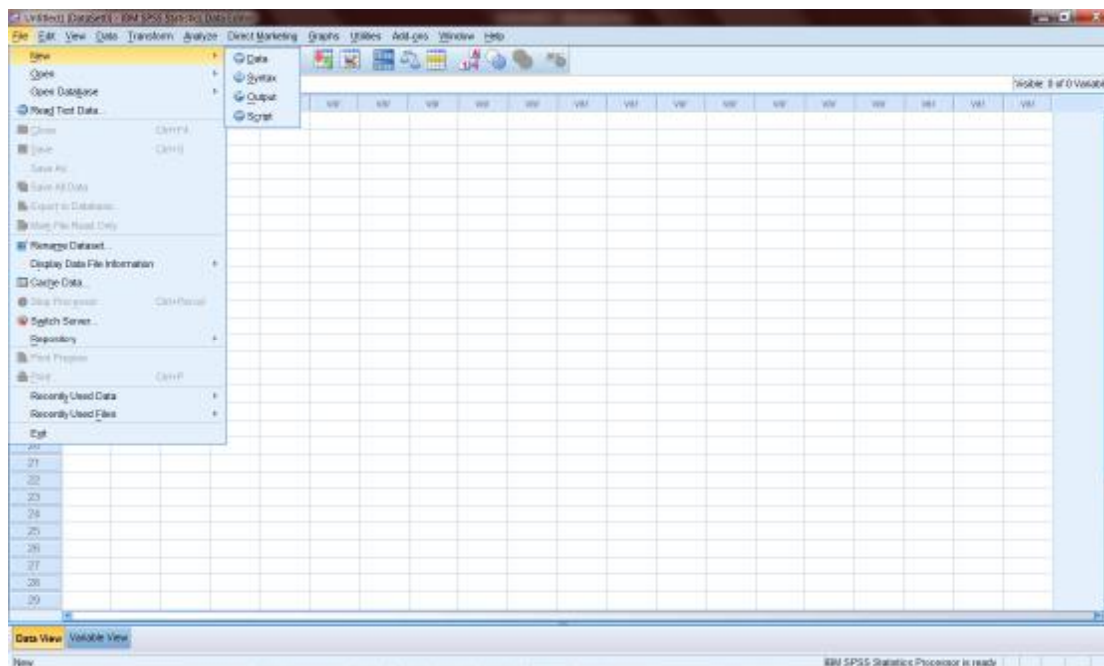
- Κάθε εντολή ξεκινά από νέα γραμμή και τελειώνει με **τελεία (.)**
- Οι υπό-εντολές ξεχωρίζουν με **slash (/)**
- Οι μεταβλητές πρέπει να είναι επακριβώς ορισμένες
- Τα κείμενα πρέπει να περιέχονται εντός αποστρόφων
- Κάθε γραμμή δεν πρέπει να περιλαμβάνει άνω των 80 χαρακτήρων

- Κάθε γραμμή που αντιπροσωπεύει σχόλια θα πρέπει να αρχίζει με το σύμβολο \* , στη συνέχεια να μεσολαβεί ένα κενό και να συνεχίζει το κείμενο.

### 3.3 Το κύριο Μενού του Spss για Windows και οι λειτουργίες του

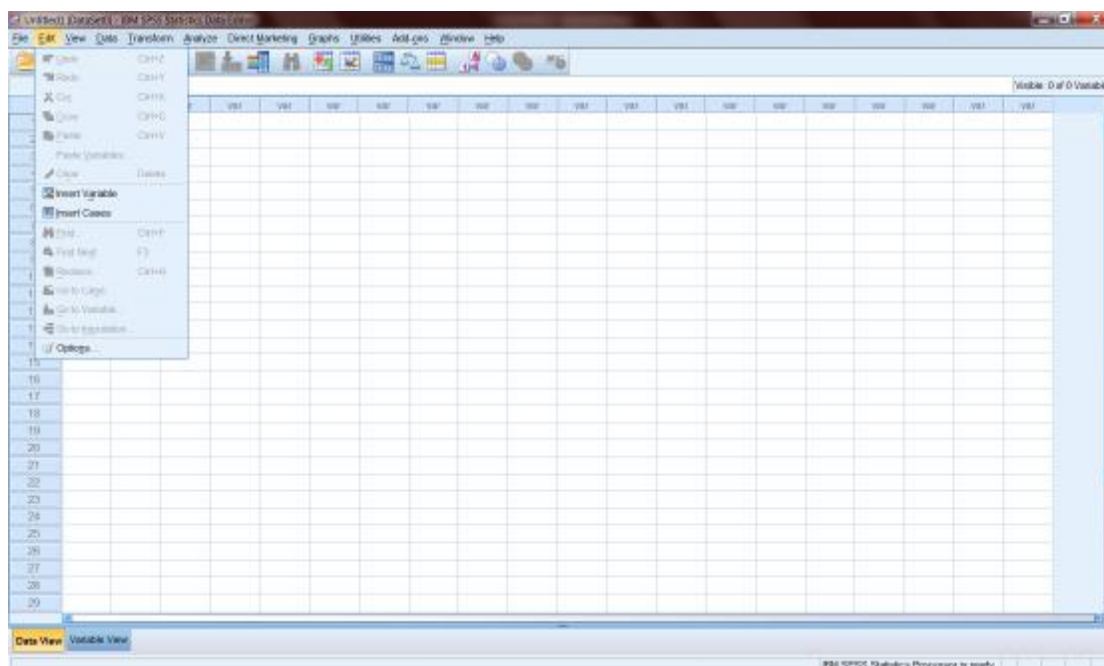
Στο κεφάλαιο αυτό θα γίνει μία σύντομη παρουσίαση του κυρίως μενού, που εμφανίζεται με το παράθυρο του Data Editor του Spss για Windows. Με το συγκεκριμένο υπομενού μπορεί κανείς να δημιουργήσει νέο αρχείο του Spss, να διαβάσει, να γράφει και να τυπώνει αρχεία δεδομένων, αρχεία κειμένων και διαγράμματα. Επίσης, με την εντολή Exit μπορεί κανείς να εγκαταλείψει το περιβάλλον του Spss.

Εικόνα 3.3.1: Μενού File



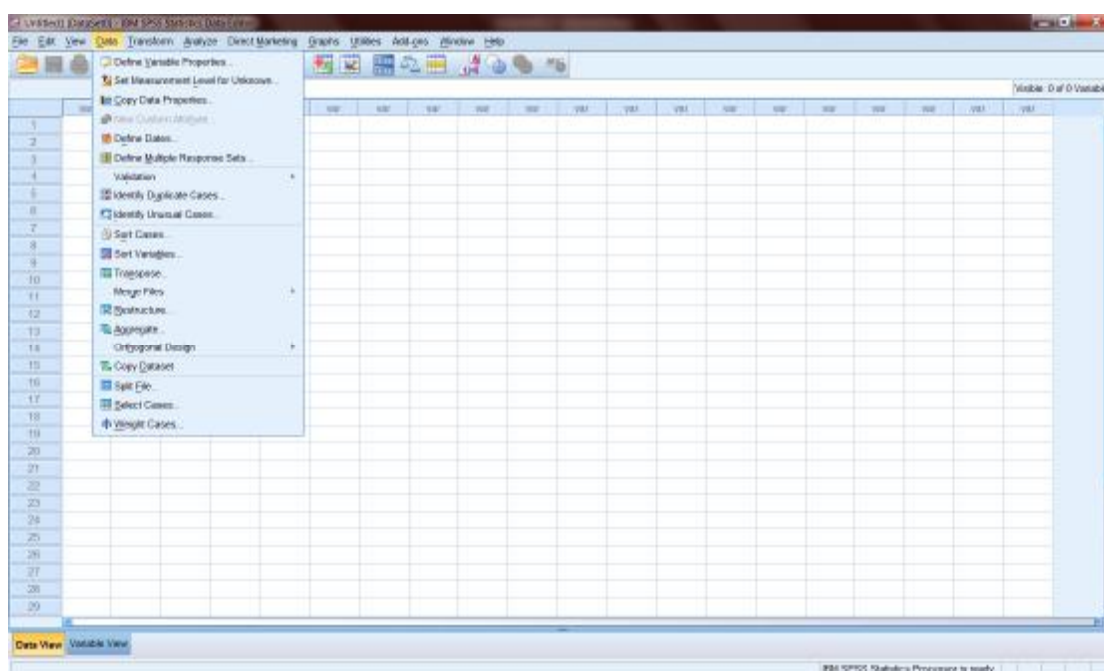
Με τις εντολές του μενού Edit, μπορεί κανείς να διορθώσει και να ψάξει κείμενο στα παράθυρα **syntax** και **output** Επίσης, περιλαμβάνει την εντολή **Options**: για βασικούς καθορισμούς του τρόπου λειτουργίας του Spss.

Εικόνα 3.3.2: Μενού Edit



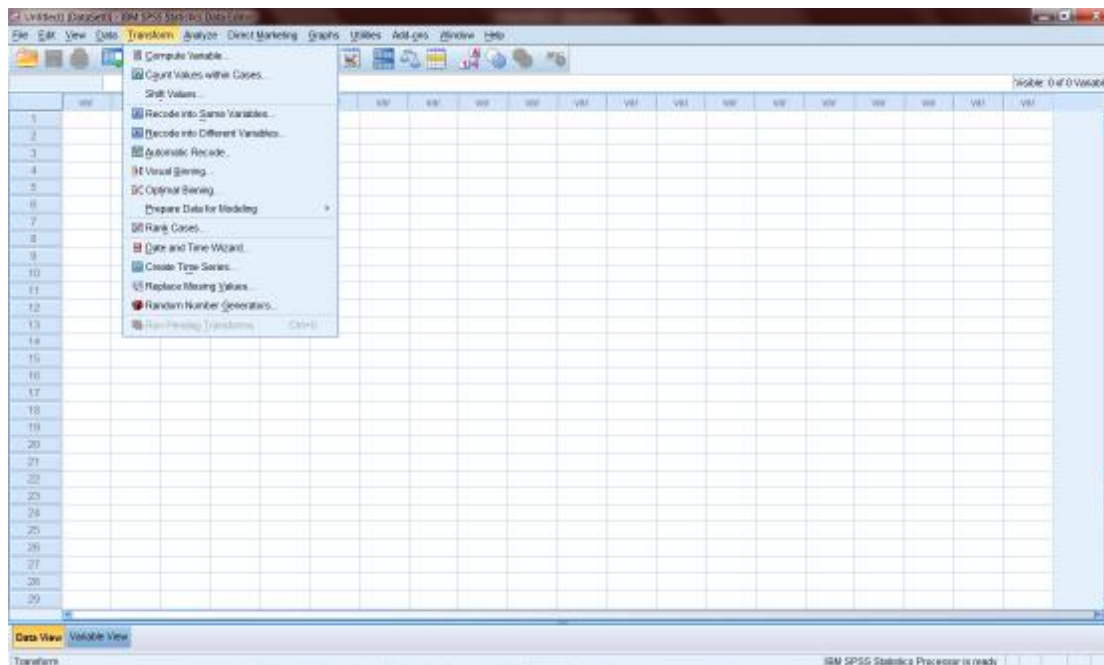
Το μενού Data υποστηρίζει τεχνικές για τον ορισμό μεταβλητών, για την εισαγωγή μεταβλητών και εγγραφών, για τη ταξινόμηση των εγγραφών, για την ομαδοποίηση εγγραφών του αρχείου δεδομένων, για το διαχωρισμό, την ένωση και την αναστροφή αρχείων.

Εικόνα 3.3.3: Μενού Data



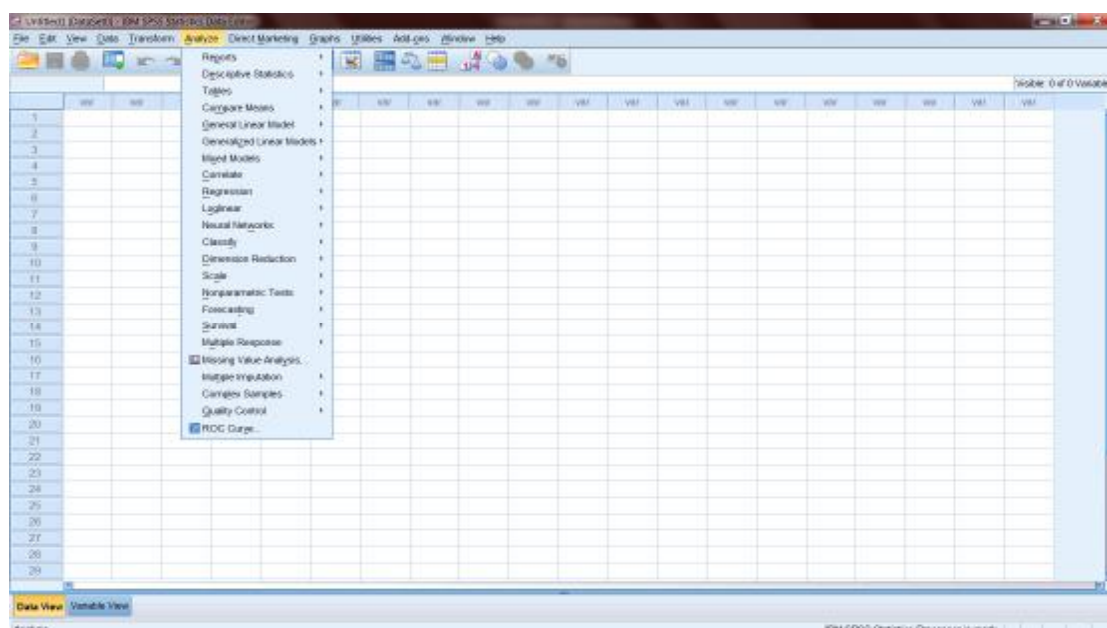
Το μενού Transform παρέχει τεχνικές για μετασχηματισμό δεδομένων χωρίς αλλαγή της δομής του αρχείου.

Εικόνα 3.3.4: Μενού Transform



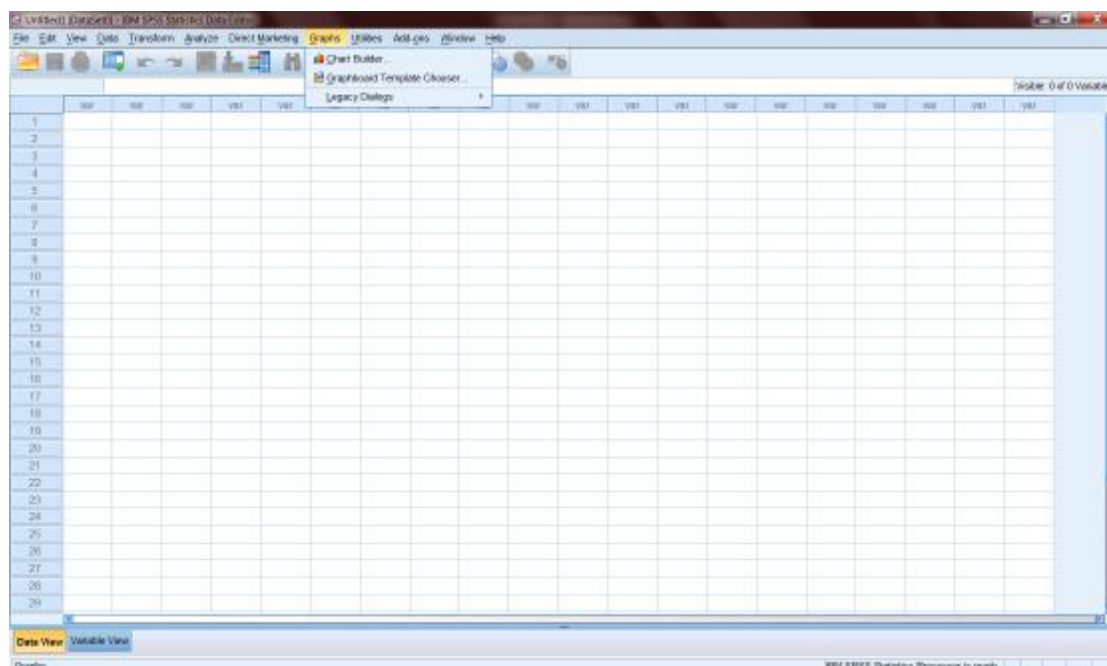
Το μενού Analyze παρέχει όλες τις τεχνικές στατιστικών αναλύσεων, μέσα από τα υπομενού, όπως περιγραφικές στατιστικές διαδικασίες frequencies, explore κλπ. καθώς και επαγωγικές τεχνικές όπως έλεγχο μέσων τιμών, ανάλυση διασποράς κλπ. Οι εκδόσεις πριν από την έκδοση 10 στο μενού αντί για την λέξη Analyze χρησιμοποιούν την λέξη Statistics. Επίσης το υπομενού Summarise της εντολής Statistics στην εντολή Analyze έχει αλλάξει σε Descriptive Statistics.

Εικόνα 3.3.5: Μενού Analyze



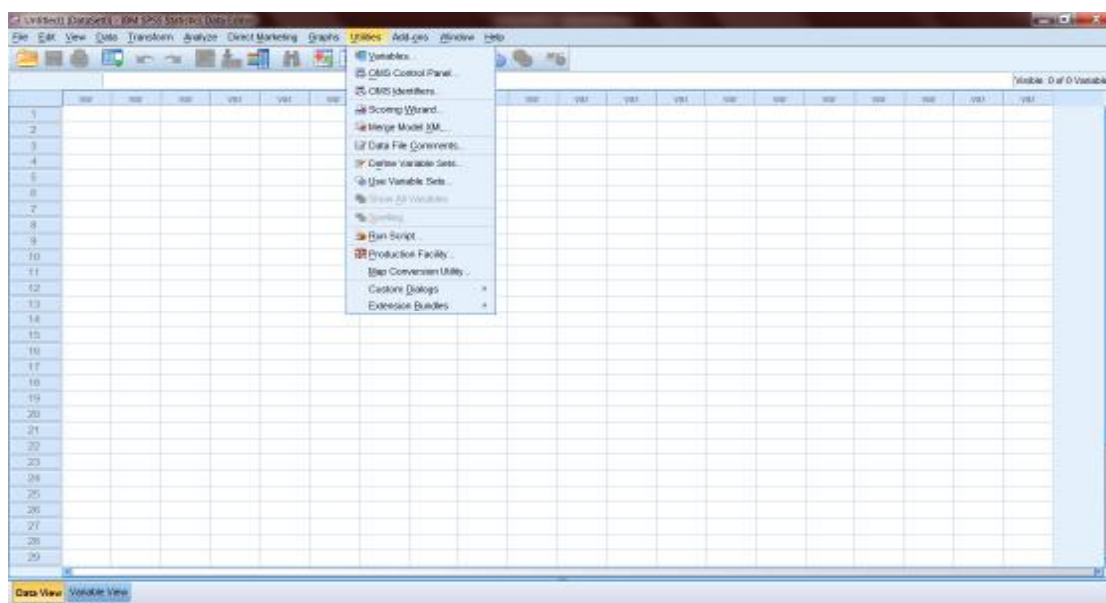
Το μενού Graphs επιτρέπει την δημιουργία διαγραμμάτων από τα δεδομένα για καλύτερη παρουσίαση πχ. Ιστογράμματα, Ραβδογράμματα, Κυκλικά διαγράμματα κλπ.

Εικόνα 3.3.6: Μενού Graphs



Το μενού Utilities υποστηρίζει διάφορες συναρτήσεις, όπως λίστα εντολών, επιλογή γραμματοσειρών, πληροφορίες για το αρχείο δεδομένων, την εμφάνιση του **Data Editor** και την επιλογή του παράθυρου προορισμού για τα αποτελέσματα και τη σύνταξη των εντολών.

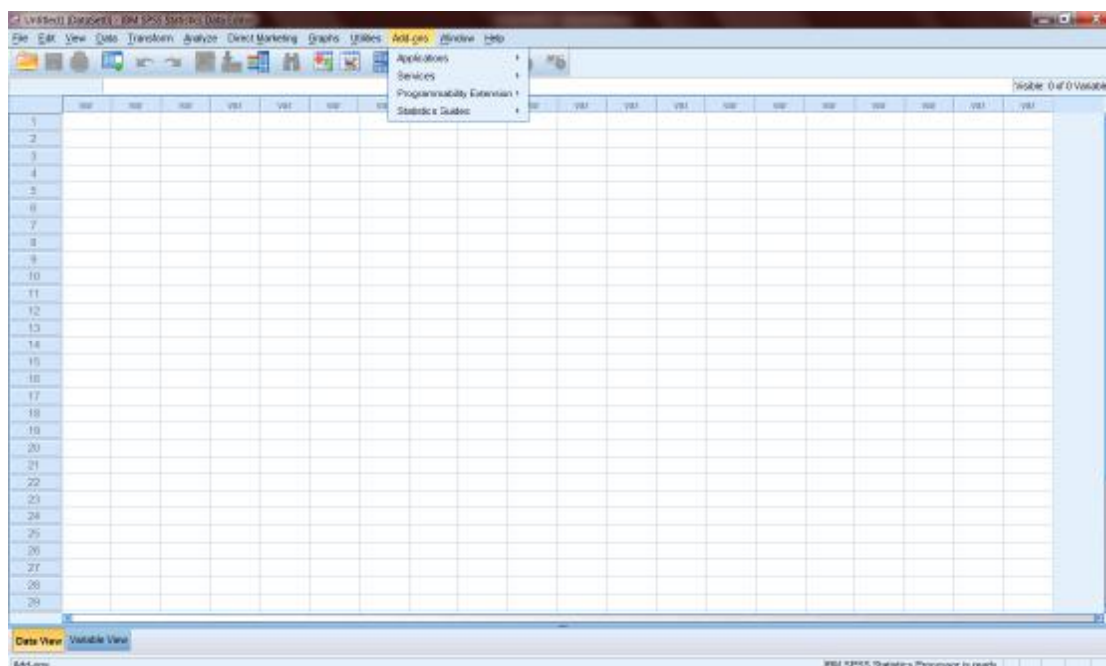
Εικόνα 3.3.7: Μενού Utilities



Το μενού **Add-ons** διαθέτει κάποιες πρόσθετες λειτουργίες όπως οδηγίες στατιστικής επεξεργασίας και επιλογές εφαρμογής του στατιστικού πακέτου SPSS.

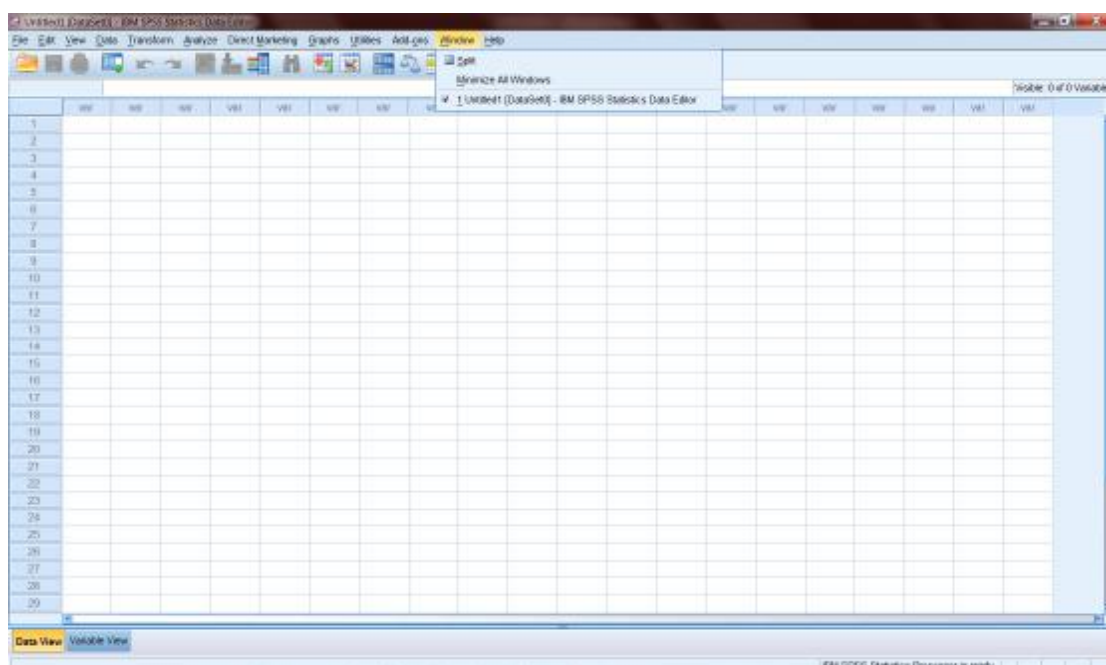
Εικόνα 3.3.8: Μενού Add-ons





Το μενού **Window** επιτρέπει τη ενεργοποίηση οποιουδήποτε παράθυρου του SPSS για Windows και τροποποιεί τον τρόπο εμφάνισής τους

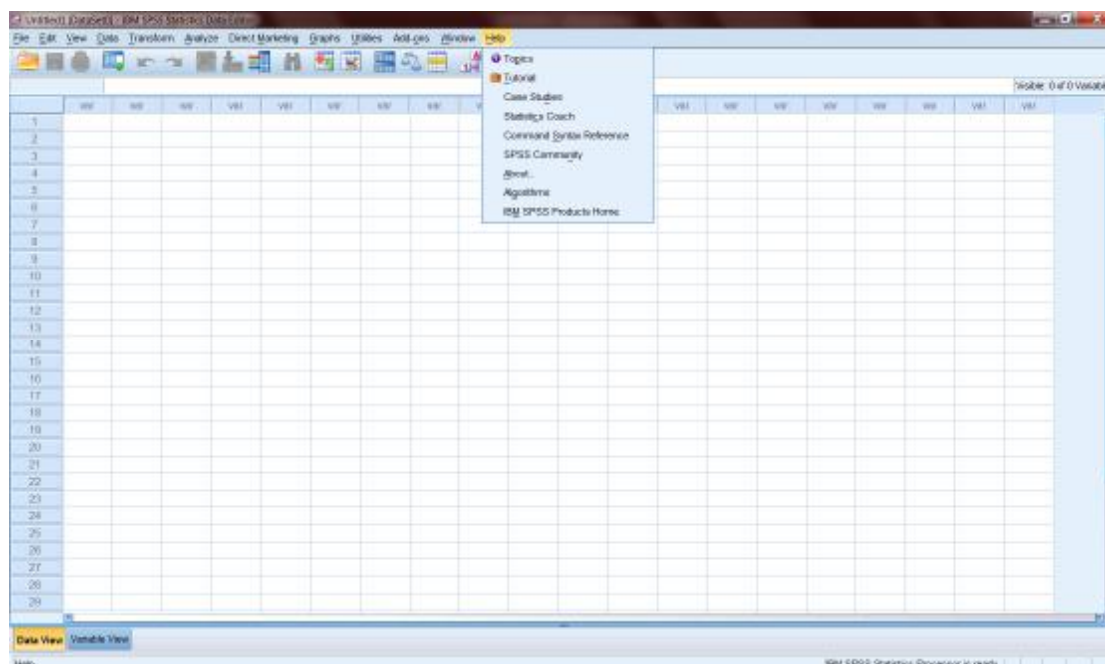
Εικόνα 3.3.9: Μενού Window



Το μενού **Help** προσφέρει βοήθεια για τη χρήση γενικά του Spss για Windows, αλλά προσφέρει και άμεση βοήθεια σε κάθε φάση.



Εικόνα 3.3.10: Μενού Help



### 3.4 Βασικές εργασίες με το SPSS για Windows

Χρησιμοποιώντας τα μενού ή τις εντολές του SPSS ζητάει κανείς να εκτελεστούν κυρίως τέσσερις κατηγορίες εργασιών, συνήθως με τη παρακάτω σειρά.

#### 3.4.1. Ανάγνωση, Εισαγωγή Δεδομένων

Τα δεδομένα μας μπορεί να υπάρχουν γραμμένα σε κάποιο χαρτί ή να είναι ήδη καταχωρημένα σε ηλεκτρονική μορφή με χρήση κάποιου λογισμικού πχ. Spreadsheet (Lotus, Excel κλπ), πακέτα διαχείρισης βάσεων δεδομένων (Dbase, Access, Oracle κλπ.), του ίδιου του SPSS ή σε ASCII μορφή (σειριακά αρχεία σταθερού ή ελεύθερου Format) κλπ. Η εισαγωγή δεδομένων γίνεται κατευθείαν στο **Data Editor**. Η ανάγνωση άλλων αρχείων γίνεται επιλέγοντας από το μενού **File - Open** ή **File - Open - Database** ή **File - Read Text Data** .οπότε εμφανίζονται τα δεδομένα πάλι στο Data editor. Με την εισαγωγή των δεδομένων μας στο **SPSS για Windows** ορίζεται ένα ενεργό αρχείο( **active file**), όπως αναφέρθηκε πιο πάνω.

### 3.4.2 Μετασχηματισμός Δεδομένων

Στο στάδιο αυτό αποφασίζονται όλες οι αναγκαίες τροποποιήσεις, ώστε να μετασχηματισθούν τα στοιχεία, που έχουν ήδη εισαχθεί στο περιβάλλον του SPSS. Ο μετασχηματισμός γίνεται κυρίως από τα μενού **Data**, **Edit**, **Transform** και **Utilities**. Στη φάση αυτή χρησιμοποιούνται οι εντολές Διαχείρισης Δεδομένων (Data Manipulation) του SPSS, όπως πχ. η **Sort** για ταξινόμηση, η **Compute** για δημιουργία νέων μεταβλητών, η **Recode** για επανακωδικοποίηση κλπ.

### 3.4.3 Γραφική Παρουσίαση Δεδομένων

Στο στάδιο αυτό γίνεται μια πρώτη γραφική παρουσίαση των δεδομένων για έλεγχο των μεταβλητών. Σε αυτή τη φάση θα χρησιμοποιηθεί το μενού **Graphs**.

### 3.4.4 Ανάλυση Δεδομένων

Το τελευταίο στάδιο είναι το καθαρά στατιστικό μέρος του SPSS για Windows. Τα δεδομένα είναι στην κατάλληλη μορφή ώστε να μπορεί κανείς να χρησιμοποιήσει τις κατάλληλες στατιστικές διαδικασίες από το μενού **Analyse**. Επίσης, μπορεί κανείς να χρησιμοποιήσει κατευθείαν τις εντολές Ανάλυσης Δεδομένων (**Procedures Commands**) του SPSS για Windows όπως πχ. Descriptives, Means, Regression κλπ ζητώντας την εκτέλεση κάποιων συγκεκριμένων αναλύσεων και υπολογισμών.

Εκτός των εργασιών που έχουν ήδη αναφερθεί με τη σειρά που συνήθως ακολουθείται, μπορεί κανείς να καθορίζει το γενικό περιβάλλον λειτουργίας του SPSS από το μενού **View**, **Edit**, κλπ ή με αντίστοιχες εντολές λειτουργιών (**Operation commands**) όπως πχ. Display, Set κλπ.

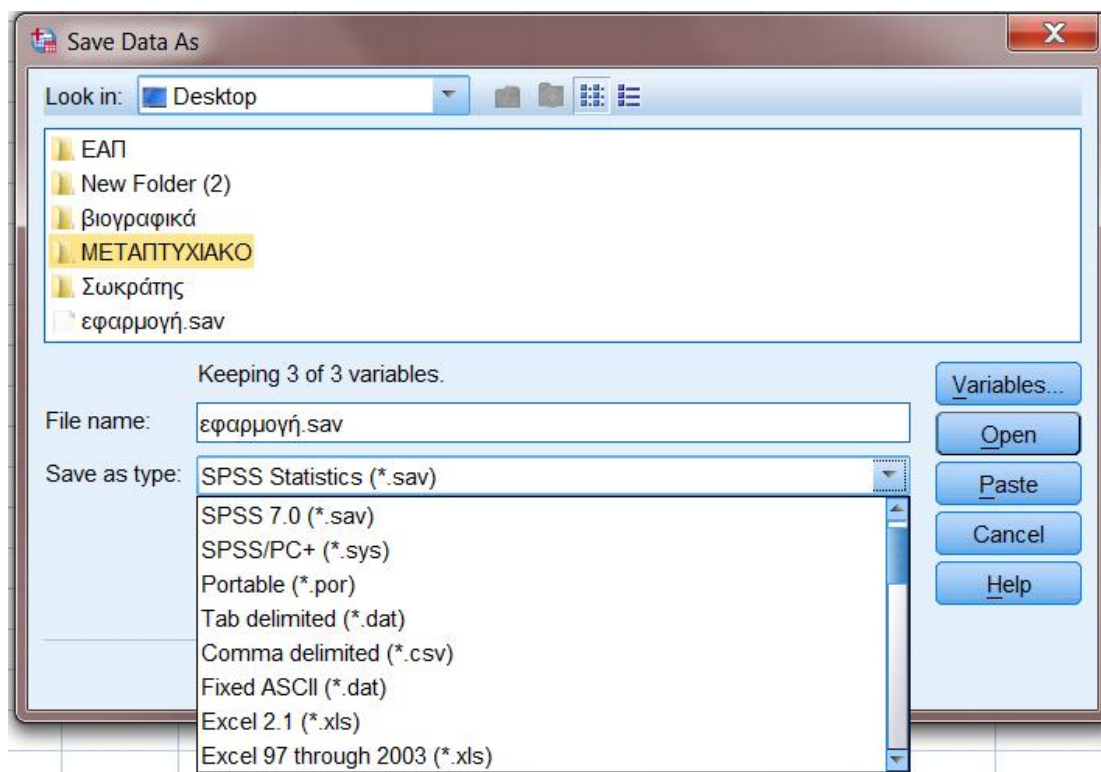
## 3.5 Αποθήκευση και Έξοδος από το SPSS

Για να αποθηκεύσουμε σε ένα νέο αρχείο δεδομένων ή για να αποθηκεύσουμε σε ένα διαφορετικό format, εκτελούμε διαδοχικά τα παρακάτω βήματα:

- Έχουμε ενεργό το Data Editor παράθυρο.
- Επιλέγουμε File - Save As

και έχουμε ως αποτέλεσμα την εμφάνιση της θυρίδας διαλόγου Save Data As, που δίνεται στην Εικόνα 3.5.1.

Εικόνα 3.5.1: Αποθήκευση αρχείου δεδομένων



Στο παράθυρο αυτό καθορίζουμε πλήρως το όνομα και τη διεύθυνση του αρχείου στο οποίο θέλουμε να αποθηκευτούν τα δεδομένα μας. Ο τρόπος καθορισμού είναι παρόμοιος με αυτόν που χρησιμοποιήσαμε και για το άνοιγμα του αρχείου. Η σημαντική δυνατότητα που έχουμε εδώ είναι η αποθήκευση των δεδομένων σε τύπο αρχείου που μπορεί να χρησιμοποιηθεί από άλλη εφαρμογή (Excel, dBASE κλπ). Ο τύπος του αρχείου δηλώνεται στη θυρίδα «Αποθήκευση ως».

Η έξοδος από το πρόγραμμα μπορεί να γίνει με τρεις τρόπους. Από το μενού **File** επιλέγοντας την εντολή **Exit** ή Κάνοντας διπλό κλικ στο εικονίδιο μενού ελέγχου ή επιλέγοντας το κουμπί 'x' άνω δεξιά.

### 3.6 Data Editor και SPSS

Όπως ήδη αναφέρθηκε με το ξεκίνημα του Spss για Windows ανοίγει το παράθυρο του **Data Editor**. Περιέχει ένα πλέγμα γραμμών και στηλών μέσα στο οποίο εμφανίζονται οι εγγραφές του αρχείου δεδομένων. Στο **Data Editor** μπορεί κανείς να δημιουργήσει νέο αρχείο γράφοντας κατευθείαν σ' αυτό τα δεδομένα, να διορθώσει

τα δεδομένα υπάρχοντος αρχείου, να προσθέσει ή να διαγράψει μεταβλητές, να αλλάξει τις ιδιότητες τους κλπ.

Κάθε **στήλη** του Data Editor αποτελεί μία **μεταβλητή**, δηλαδή ένα μετρούμενο χαρακτηριστικό ενός πληθυσμού. Για παράδειγμα κάθε ερώτηση ενός ερωτηματολογίου είναι μία μεταβλητή. Κάθε **γραμμή** του Data Editor είναι μία **εγγραφή** ή όπως ονομάζεται στο Spss **case**. Για παράδειγμα, όλες οι απαντήσεις ενός ατόμου ενός ερωτηματολογίου είναι ένα **case**. Το κάθε κελί είναι η διασταύρωση μιας στήλης και μιας γραμμής δηλαδή μιας **μεταβλητής** και μιας **case**. Τα κελιά στο SPSS για Windows περιέχουν μόνο τιμές σε αντίθεση με τα κελιά στα λογιστικά φύλλα στα οποία μπορεί κάποιος να εισαγάγει και συναρτήσεις.

Θα μπορούσε κάποιος να παρομοιάσει το αρχείο δεδομένων με ένα παραλληλόγραμμο, του οποίου οι διαστάσεις καθορίζονται κάθε φορά σε σχέση με τον αριθμό των εγγραφών και τον αριθμό των μεταβλητών. Τα κελιά τα οποία δεν έχουν τιμές μέσα στα όρια του αρχείου δεν θεωρούνται κενά. Στην περίπτωση που οι μεταβλητές περιέχουν αριθμητικά δεδομένα τα κενά κελιά μετατρέπονται σε τιμές system missing και παρουσιάζονται με τελεία, ενώ στις αλφαριθμητικές μεταβλητές τα κενά θεωρούνται έγκυρες τιμές.

### 3.6.1 Βασικές λειτουργίες του Data Editor

Όπως προαναφέρθηκε είναι σκόπιμο να δημιουργηθεί ένα αρχείο δεδομένων για το SPSS, και για το σκοπό αυτό υπάρχουν δύο τρόποι. Ο πρώτος τρόπος είναι να εισαχθούν τα δεδομένα απευθείας στο άδαιο αρχείο (active file), που δημιουργείται αυτόματα με το ξεκίνημα του SPSS για Windows. Σε πρώτη φάση καθορίζονται οι μεταβλητές, στη συνέχεια εισάγονται τα δεδομένα και τέλος αποθηκεύονται. Επίσης, οποιαδήποτε στιγμή, μπορεί ο χρήστης να προβεί στη δημιουργία ενός νέου αρχείου επιλέγοντας από το μενού **File - New Data**.

Ο δεύτερος τρόπος είναι να χρησιμοποιήσει κανείς ένα άλλο λογισμικό (βάσεις δεδομένων ή λογιστικά φύλλα ή Editor ή κειμενογράφο) για την εισαγωγή των στατιστικών δεδομένων και ύστερα να τα μεταφέρει στο πρόγραμμα του SPSS επιλέγοντας από το μενού την εξής ακολουθία εντολών:

**File-Open-Data** ή **File-Open Database** ή **File-Read Text Data**

Για παράδειγμα, στη περίπτωση που τα δεδομένα έχουν καταχωρηθεί σε ένα λογιστικό φύλλο Excel για μπορέσει κάποιος να τα εισάγει στο SPSS για Windows θα ακολουθήσει τα ακόλουθα βήματα:

**Βήμα 1:** Καταρχάς επιλέγοντας το μενού File-Open - Data εμφανίζει το πλαίσιο διαλόγου Open File. Σε αυτό το πλαίσιο διαλόγου ορίζει στη περιοχή Τύπος Αρχείου (File Type) την επιλογή Excel(\*.xls), που είναι το είδος αρχείου του MsExcel.

**Βήμα 2:** Στη συνέχεια, εντοπίζει τη θέση του αρχείου στον υπολογιστή και πατώντας στο κουμπί Άνοιγμα μεταφέρονται τα δεδομένα στον Data Editor στο τμήμα Data View και η τεκμηρίωση των μεταβλητών στο τμήμα Variable View.

Στη συνέχεια, απαραίτητο βήμα είναι ο ορισμός ή η αλλαγή των μεταβλητών και των χαρακτηριστικών τους στο SPSS για Windows η οποία γίνεται στο τμήμα Variable View του Data Editor. Η ενεργοποίηση του Variable View γίνεται με τους εξής δύο τρόπους:

- Ο πρώτος τρόπος είναι να κάνει κάποιος διπλό κλικ στην αρχή της στήλης του Data Editor, όπου εισάγονται τα δεδομένα της αντίστοιχης Variable
- Ο δεύτερος τρόπος είναι να επιλέξει κάποιος το αντίστοιχο κουμπί Variable View που βρίσκεται κάτω αριστερά του Data Editor. Τα 10 χαρακτηριστικά που καλείται κάποιος να ορίσει για κάθε μεταβλητή είναι τα εξής: Όνομα, Τύπος, Εύρος, Δεκαδικά, Επιγραφή, Τιμές, Κενά, Στήλες, Ευθυγράμμιση, και Μέτρηση (Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, και Measure).

Στη συνέχεια, γίνεται επεξήγηση των 10 αυτών χαρακτηριστικών που είναι απαραίτητο να συμπληρώσει ο χρήστης:

### **1. Name (Όνομα Μεταβλητής)**

Στο αντίστοιχο πλαίσιο πληκτρολογεί κάποιος το όνομα, που έχει προαποφασίσει για την κάθε μεταβλητή και πατάει το πλήκτρο Enter ή κάποιο από τα βελάκια στο πληκτρολόγιο προκειμένου να καταχωρηθεί το όνομα και να μεταφερθεί σε κάποια επόμενη θέση. Για την ονομασία των μεταβλητών θα πρέπει κανείς να ακολουθεί τους παρακάτω κανόνες:

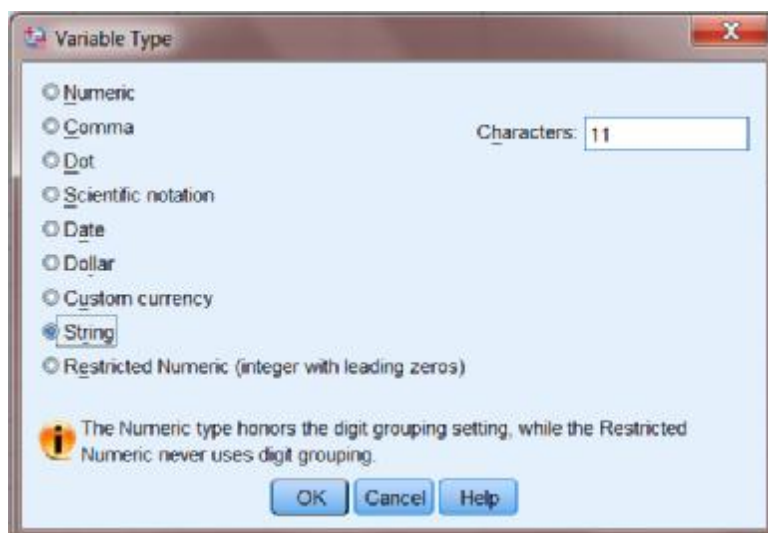
- Χρήση ονομάτων μέχρι οκτώ χαρακτήρες
- Το όνομα θα πρέπει να αρχίζει με γράμμα. Οι υπόλοιποι χαρακτήρες μπορεί να είναι γράμματα, αριθμοί, σύμβολα εκτός από @, \_, \$, !, ?, \* και το κενό.

- Τα ονόματα των μεταβλητών δεν πρέπει να τελειώνουν σε τελεία
- Όλα τα ονόματα των μεταβλητών πρέπει να είναι μοναδικά· δεν μπορεί δύο μεταβλητές να έχουν το ίδιο όνομα
- Τα ονόματα των μεταβλητών μπορεί να γραφτούν είτε με μικρά είτε με κεφαλαία γράμματα. Για παράδειγμα, οι λέξεις **COLOUR**, **Colour**, και **colour** είναι ακριβώς ίδιες για το SPSS για Windows.
- Επειδή έχουν ξεχωριστή σημασία για το SPSS, κάποιες λέξεις δεν μπορούν να χρησιμοποιηθούν για ονόματα μεταβλητών, συμπεριλαμβανομένων των: **all**, **ne**, **eq**, **to**, **le**, **It**, **gt**, **and**, **not**, **ge** και **width**.
- Θα προσέξετε ότι καθώς πατάτε enter ή κάποιο απ' τα βελάκια, στα δεξιά της μεταβλητής, της οποίας μόλις πληκτρολογήσατε το όνομα, οχτώ από τις υπόλοιπες εννιά στήλες γεμίζουν με default πληροφορίες. Οι αρχικές ρυθμίσεις συνήθως θα πρέπει να μεταβληθούν, αλλά αυτό μπορεί να γίνει εύκολα αφού έχετε πληκτρολογήσει τα ονόματα κάθε μεταβλητής.

## 2. Type( Τύπος της μεταβλητής )

Αφού δώσει κάποιος το όνομα σε μία μεταβλητή στη συνέχεια, είτε πατώντας το πλήκτρο **Enter**, είτε το πλήκτρο **Tab** ή κάνοντας κλικ στο κελί της διπλανής στήλης μεταφέρεται στην επόμενη στήλη η οποία έχει τον τίτλο **Type**. Σε αυτή τη στήλη θα πρέπει κανείς να καταχωρήσει το είδος της μεταβλητής. Εξορισμού το SPSS για Windows σ' όλες τις νέες μεταβλητές επιλέγει το τύπο **Numeric (Αριθμητική)**, επειδή θεωρεί ότι οι περισσότερες μεταβλητές, που δημιουργεί κανείς για να συμμετέχουν στις αναλύσεις είναι αριθμητικές. Αν κάποιος προσέξει στα δεξιά της λέξης **Numeric** θα δει ένα μικρό γκρι κουτάκι από όπου κάποιος μπορεί να διαλέξει μεταξύ οκτώ επιλογών ενεργοποιώντας κάποια από αυτές.

Εικόνα 3.6.1.1: Επιλογή τύπου μεταβλητής



Η κάθε μία από τις 8 επιλογές αναλύεται στη συνέχεια.

- **Numeric (Αριθμητική)**

Χρησιμοποιείται εάν η απάντηση, που θα δοθεί σε μια μεταβλητή είναι κάποιος αριθμός. Αυτός ο αριθμός δεν εμφανίζει κάποιο κόμμα ή τελεία για τον προσδιορισμό των χιλιάδων, αλλά μόνο μια τελεία για τον προσδιορισμό των δεκαδικών ψηφίων. (Τα δεκαδικά εμφανίζονται με τελεία γιατί το πρόγραμμα είναι Αμερικανικό και στις Ηνωμένες Πολιτείες της Αμερικής συνηθίζεται τα δεκαδικά να απεικονίζονται με τελεία, αντί με το κόμμα που χρησιμοποιείται στην Ελλάδα). Το μέγιστο μήκος είναι 60 ψηφία και το μέγιστο μήκος των δεκαδικών είναι 16 ψηφία.

- **Comma (Κόμμα)**

Χρησιμοποιείται για τους μεγάλους αριθμούς και ο διαχωρισμός των χιλιάδων γίνεται με το κόμμα. Επομένως με την επιλογή **Comma** ορίζει κάποιος, ότι μια **Numeric** μεταβλητή θα εμφανίζει με κόμμα τις χιλιάδες και με τελεία το δεκαδικό κομμάτι. Δηλαδή ο αριθμός χίλια θα εμφανιστεί με τη μορφή 1,000 και αν προσθέσει κάποιος δύο δεκαδικά γίνεται 1,000.00 .

- **Dot (Τελεία)**

Είναι και αυτή μια **Numeric** μεταβλητή, που όμως διαφέρει από τις προηγούμενες αφού εμφανίζεται με τελεία για τον προσδιορισμό των χιλιάδων, ενώ τα δεκαδικά προσδιορίζονται με κόμμα. Δηλαδή στις μεταβλητές Dot ο αριθμός χίλια θα εμφανιστεί με τη μορφή 1.000 και αν προσθέσει κάποιος δεκαδικά γίνεται 1.000,00, ακριβώς όπως συνηθίζεται στην Ελλάδα.

- **Scientific notation (Επιστημονική Σημείωση)**

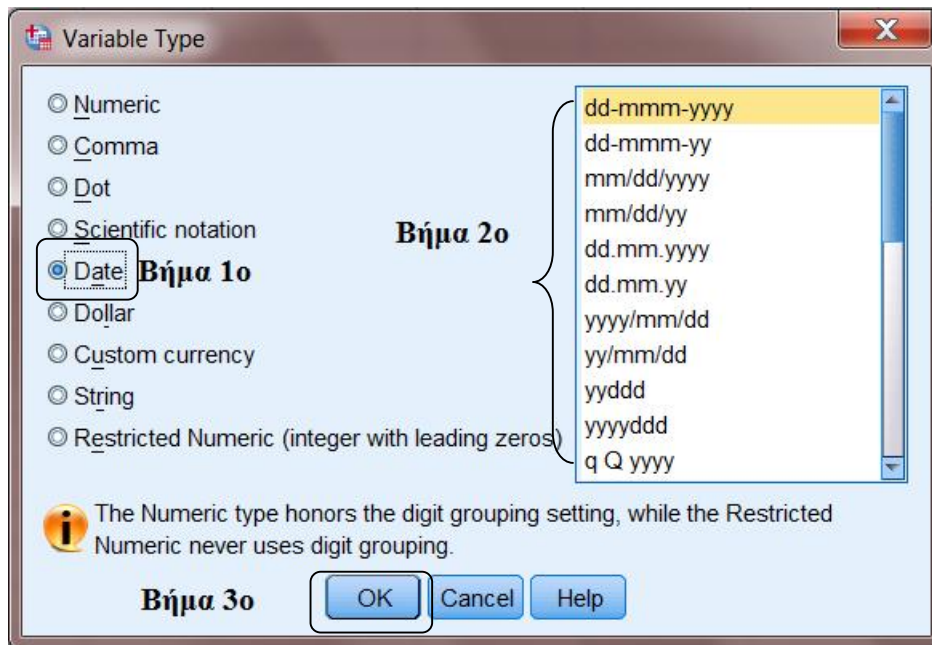
Σ' αυτές τις μεταβλητές οι τιμές εμφανίζονται με τη μορφή επιστημονικής σημείωσης, π.χ. 8.35E2, 8.35D2, 8.35E+2 και 8.35+2.

- ο **Date (Ημερομηνία)**

Σε περίπτωση, που η μεταβλητή που καταχωρεί κάποιος αναφέρεται σε κάποια ημερομηνία ενεργοποιεί το είδος μεταβλητής Date. Τότε εμφανίζεται το πλαίσιο, που επιδεικνύει το Βήμα 2, απ' όπου επιλέγει κάποιος τη μορφή με την οποία θέλει να εμφανίζει την ημερομηνία. Τέλος, πατώντας το κουμπί OK (Βήμα 3) κατοχυρώνεται ο τύπος της ημερομηνίας.



Εικόνα 3.6.1.2: Καθορισμός μεταβλητής Date (Ημερομηνία) στο πλαίσιο διαλόγου Variable Type (Είδος Μεταβλητής)



- **Dollar (Δολάριο)**

Σε περίπτωση, που κάποιος αναφέρεται σε νομίσματα δολαρίου ενεργοποιεί το αντίστοιχο είδος μεταβλητή και μπορεί κανείς να επιλέξει τη μορφή και το δεκαδικό ψηφίο που θέλει να έχει η μεταβλητή. Τέλος πατώντας το κουμπί OK, καταχωρείται το είδος της μεταβλητής.

- **Custom currency (Συνηθισμένο νόμισμα)**

Σε περίπτωση που αναφέρεται κάποιος σ' άλλο νόμισμα εκτός δολαρίου τότε επιλέγει **Custom currency** και ακολουθεί την ίδια διαδικασία μ' αυτή της μεταβλητής **Dollar**.

- **String (Αλφαριθμητική σειρά χαρακτήρων)**

Οι μεταβλητές String αποτελούν μία σειρά από χαρακτήρες που συμπεριλαμβάνουν γράμματα, σύμβολα, αριθμούς και κενά. Για παράδειγμα, εάν θέλει κάποιος να καταγράψει τα ονόματα όλων των ατόμων ή επιχειρήσεων που συμμετείχαν σε μια έρευνα, τότε αυτή η μεταβλητή πρέπει να είναι της μορφής String.

- **Restricted Numeric (Αριθμητική με περιορισμούς)**

Η διαφορά της με την αριθμητική είναι ότι δεν χρησιμοποιεί ομαδοποίηση ψηφίων.

#### 4. Width (Εύρος) και Decimals (Δεκαδικά)

Στη στήλη **Width** καθορίζουμε το μέγιστο αριθμό ή το μέγιστο string, που μπορεί να έχει κάποιος για τις τιμές κάθε μεταβλητής. Όταν κάποιος κάνει κλικ στο κελί υπό

τον τίτλο **width** στα δεξιά του ονόματος κάθε μεταβλητής το κελί τονίζεται κι εμφανίζονται δύο βελάκια στο δεξί περιθώριο του κελιού. Αυτά χρησιμοποιούνται για να αυξηθεί ή να μειωθεί το εξορισμού αρχικό εύρος των 8 χαρακτήρων μιας μεταβλητής. Στη στήλη **Decimals** ορίζει κανείς τον αριθμό των δεκαδικών ψηφίων, που θέλει να υπάρχει στα δεδομένα του. Εξορισμού κάθε μεταβλητή έχει δύο δεκαδικά ψηφία. Για τις μεταβλητές string ο αριθμός των δεκαδικών είναι εξ' αρχής 0.

### 5. Label ( Ετικέτα της μεταβλητής)

Η στήλη **Label** επιτρέπει σε κάποιον να τοποθετήσει ετικέτες (περιγραφή του τι αντιπροσωπεύει η μεταβλητή) στις μεταβλητές των οποίων η σημασία δεν είναι ξεκάθαρη από το όνομα τους. Πολλές φορές η σημασία είναι ξεκάθαρη από το ίδιο το όνομα (π.χ. **colour, sex, age**) και δεν απαιτείται ετικέτα. Άλλες φορές η σημασία **δεν** είναι ξεκάθαρη και μια ετικέτα είναι πολύ χρήσιμη. Τα κελιά στη στήλη **Label** είναι απλώς πλαίσια κειμένου και πληκτρολογεί κάποιος την περιγραφή που θέλει. Το μέγιστο μήκος είναι **256 χαρακτήρες**, αλλά κάτι τόσο μεγάλο είναι δυσκίνητο.

### 6. Value (Ετικέτα των τιμών μιας μεταβλητή)

Η επιλογή **Value** επιτρέπει να εντοπίζει κάποιος το **επίπεδο** μιας μεταβλητής (π.χ. **sex:1=θηλυκό, 2=αρσενικό· answer: 1=ναι, 2=όχι, 3=ίσως**). Η εισαγωγή ετικετών **Value** για μεταβλητές, που έχουν πολλές διακριτές ομάδες είναι εξαιρετικά κρίσιμη για τη σαφήνεια της ερμηνείας του αποτελέσματος. Το SPSS για Windows μπορεί να κάνει τις πράξεις, είτε βάλει κάποιος ετικέτες είτε όχι. Αν δεν βάλει ετικέτες κάποιος δεν πρόκειται να θυμάται μετά από κάποιο διάστημα τι σημαίνει μια οποιαδήποτε τιμή σε μια μεταβλητή.

### 7. Missing (Κενά -Ορισμός των τιμών που λαμβάνονται ως μη έγκυρες)

Ο σκοπός της στήλης **Missing** είναι να προσδιορίζει διαφορετικούς τύπους κενών στα δεδομένα σας. Μπορεί κανείς να διακρίνει δύο κατηγορίες **Missing**.

- **System Missing Values.** Τιμές που θεωρούνται από το σύστημα σαν μη έγκυρες. Για παράδειγμα, κάθε κενό σε αριθμητική μεταβλητή ή μετά από αδύνατους υπολογισμούς δηλώνεται με τελεία.
- **User Missing Values.** Κανονικές τιμές τις οποίες μπορεί να δηλώσει ο χρήστης σαν μη έγκυρες π.χ. περιπτώσεις στις οποίες δεν θα έπρεπε να απαντήσει κάποιος ή ερωτώμενοι, που αρνήθηκαν να απαντήσουν σε μια

ερώτηση μπορεί να κωδικοποιηθούν με το 8 κι εκείνοι, που ήθελαν να απαντήσουν κάτι διαφορετικό από αυτό που είχε στη σκέψη του ο ερευνητής να κωδικοποιηθούν με το 9. Αν εισάγει κάποιος αυτές τις τιμές στην στήλη **Missing**, τότε μπορεί να ορίσει ότι τα 8 και τα 9 δεν θα συμμετέχουν στις αναλύσεις που ακολουθούν.

### **8. Column (Καθορισμός της μορφοποίησης μιας μεταβλητής)**

Για να καθορίσει ή να τροποποιήσει κάποιος το πλάτος εμφάνισης μιας μεταβλητής παράθυρο **Data View του Data Editor** θα πρέπει, είτε να πατήσει και να σύρει τα όρια μιας στήλης (ο κέρσορας του ποντικιού παίρνει την μορφή σταυρονήματος), είτε να κάνει κλικ στην επιλογή **Column** του παραθύρου **Variable View**. Στη δεύτερη επιλογή κάνοντας κλικ στο κελί της επιλογής column για την αντίστοιχη μεταβλητή τονίζεται το κελί και εμφανίζονται δύο βελάκια τα οποία επιτρέπουν στον χρήστη να αυξήσει ή να μειώσει το αριθμό που καθορίζει το πλάτος εμφάνισης της συγκεκριμένης μεταβλητής.

### **9. Align (Καθορισμός της ευθυγράμμισης μιας μεταβλητή)**

Η στήλη **Align** παρέχει ένα drop-down μενού που επιτρέπει σε κάποιον να ευθυγραμμίσει τα δεδομένα σε κάθε κελί, είτε δεξιά είτε αριστερά ή στο κέντρο. Εξορισμού οι αριθμητικές μεταβλητές είναι ευθυγραμμισμένες δεξιά, οι μεταβλητές string στα αριστερά.

### **10. Measure (Καθορισμός της κλίμακας μέτρησης μιας μεταβλητή)**

Η στήλη **Measure** παρέχει ένα drop-down μενού που επιτρέπει σε κάποιον να διαλέξει τρεις επιλογές βασίζομενος στη φύση των δεδομένων του:

- **Nominal** (ονομαστική κλίμακα).
- **Ordinal** (ιεραρχική κλίμακα), και
- **Scale** (περιλαμβάνονται οι γνωστές από τη στατιστική κλίμακες Interval (διαστήματος) και Ratio (Αναλογίας)

Οι **Ονομαστικές Κλίμακες** αντιπροσωπεύουν το χαμηλότερο επίπεδο μέτρησης και απλά ταξινομούν τα δεδομένα σε κατηγορίες, δηλαδή κάθε τιμή προσδιορίζει απλά μία ξεχωριστή κατηγορία, αλλά δεν έχουν καμία εσωτερική διάταξη (λιγότερο σε περισσότερο).

Οι **Τακτικές Κλίμακες** ή **Ιεραρχικές Κλίμακες** ταξινομούν τα δεδομένα, όπως και οι Ονομαστικές Κλίμακες αλλά επιπλέον καθορίζουν και μία σειρά (εσωτερική διάταξη) μεταξύ των κατηγοριών. Για παράδειγμα, όταν κάποιος κατηγοριοποιήσει ένα πληθυσμό με βάση το επίπεδο εκπαίδευσης ξέρει, ότι οι απόφοιτοι λυκείου είναι μεγαλύτερη κατηγορία από του απόφοιτους δημοτικού, αλλά μικρότερη κατηγορία σε σχέση με τους απόφοιτους πανεπιστημίων.

Η επιλογή **Scale** στο SPSS είναι η αρχικά καθορισμένη για τις αριθμητικές μεταβλητές. Οι μεταβλητές που έχουν δύο μόνο κατηγορίες ή παίρνουν δύο τιμές, 0 και 1 ονομάζονται **Δίτιμες** ή **Διχοτομικές** ή μεταβλητές **Binary** και αποτελούν μία ιδιαίτερη κατηγορία όσον αφορά τη χρήση τους στη στατιστική. Έχουν την έννοια της κατάταξης και της απόστασης (πάντα σταθερή διαφορά), αλλά δεν είναι αναλογικές. Όλες οι ονομαστικές και τακτικές μεταβλητές μπορούν να μετατραπούν σε μία ή περισσότερες μεταβλητές δύο κατηγοριών (0,1).

Στο SPSS για Windows, ανεξάρτητα από το τι λέει η Στατιστική, ισχύουν οι εξής κανόνες:

- Για τις Διχοτομικές Μεταβλητές ορίζει κάποιος **ΠΑΝΤΑ** Κλίμακα Μέτρησης **Ordinal**
- Οι **Μεταβλητές String** μετρώνται σε Κλίμακα Nominal (χρησιμοποιούνται για κατηγοριοποίηση, αλλά δεν συμμετέχουν στις περισσότερες αναλύσεις).
- Οι μεταβλητές με προκαθορισμένες κατηγορίες τιμών μετρώνται σε Κλίμακα Ordinal. Αν υπάρχουν περισσότερες από 24 κατηγορίες (default), τότε η κλίμακα αλλάζει σε Scale.
- Οι **Μεταβλητές Numeric** οι τιμές των οποίων είναι καθαροί αριθμοί μετρώνται σε Κλίμακα Scale.

Μερικές φορές μπορεί να είναι δύσκολο να διαλέξει κάποιος μεταξύ κλίμακας και ιεράρχησης. Αν συμβεί αυτό δεν πρέπει να ανησυχεί. Σε όλες τις αναλύσεις, το SPSS για Windows αντιμετωπίζει και τις δύο κατηγορίες μεταβλητών με τον ίδιο τρόπο. Εκείνο που διαφέρει είναι τα στατιστικά και οι αναλύσεις, που έχουν έννοια για κάθε περίπτωση.

Αξίζει να σημειωθεί, ότι τα στατιστικά μέτρα που εφαρμόζονται σε μεταβλητές ενός επιπέδου μετρησιμότητας μπορούν πάντα να χρησιμοποιηθούν για μεταβλητές υψηλότερου επιπέδου μετρησιμότητας, αλλά όχι για μεταβλητές χαμηλότερου επιπέδου μετρησιμότητας. Ένας άλλος τρόπος υποδιαίρεσης των κλιμάκων μέτρησης των μεταβλητών είναι η κατηγοριοποίηση τους σε δύο επίπεδα: στις ασυνεχείς και στις συνεχείς. Στις ασυνεχείς ανήκουν οι ονομαστικές και διατεταγμένες κλίμακες, ενώ στις συνεχείς ανήκουν οι κλίμακες διαστήματος και οι αναλογικές.<sup>6</sup>

### **3.7 Στατιστική Ανάλυση**

Στο στάδιο αυτό θα ασχοληθούμε με την βασική στατιστική ανάλυση των στοιχείων και ερμηνεία των βασικών εντολών για την απόκτηση των αποτελεσμάτων αυτών. Για να προβεί κανείς στη στατιστική ανάλυση των στοιχείων του, επιλέγει από την γραμμή εντολών την εντολή Analyze και στη συνέχεια Descriptive Statistics δηλαδή περιγραφή στατιστική. Με την εντολή αυτή μπορεί κανείς να υπολογίσει στο SPSS για Windows τις κατανομές συχνοτήτων των τιμών διαφόρων μεταβλητών, υπό μορφή πινάκων συχνοτήτων και διαγραμμάτων. Επίσης, μπορεί κάποιος να εμφανίσει διάφορα βασικά μέτρα περιγραφικής στατιστικής (μέση τιμή, διάμεσο, τυπική απόκλιση, διάμεσο, επικρατούσα τιμή κλπ).

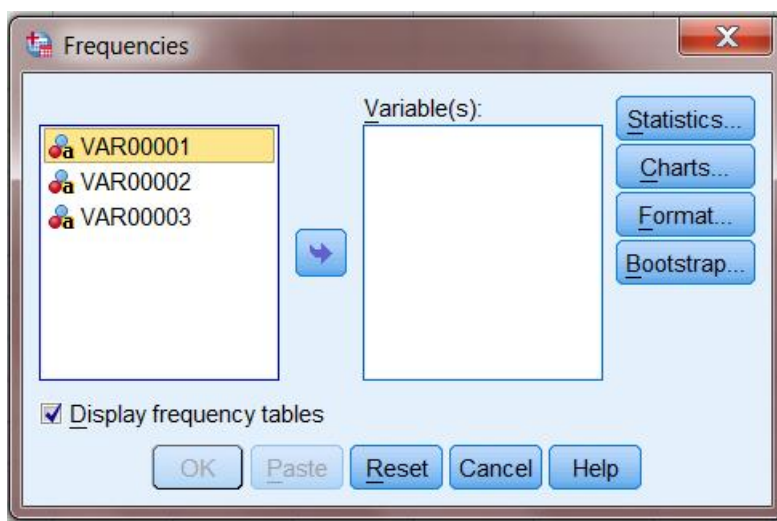
#### **3.7.1 Περιγραφή εντολής Frequencies**

Η επιλογή της συγκεκριμένης εντολής γίνεται ύστερα από την εξής ακολουθία εντολών Analyze - Descriptive Statistics - Frequencies. Έτσι, ανοίγει το παράθυρο Frequencies το οποίο περιέχει διάφορες επιλογές.

---

<sup>6</sup> Νέλλας Ε., «Ανάλυση Δεδομένων με χρήση του πακέτου SPSS/PC για WINDOWS», Τμήμα Αγροτικής Οικονομίας και Ανάπτυξης, Γεωπονικό Πανεπιστήμιο Αθηνών, Αθήνα, 2005

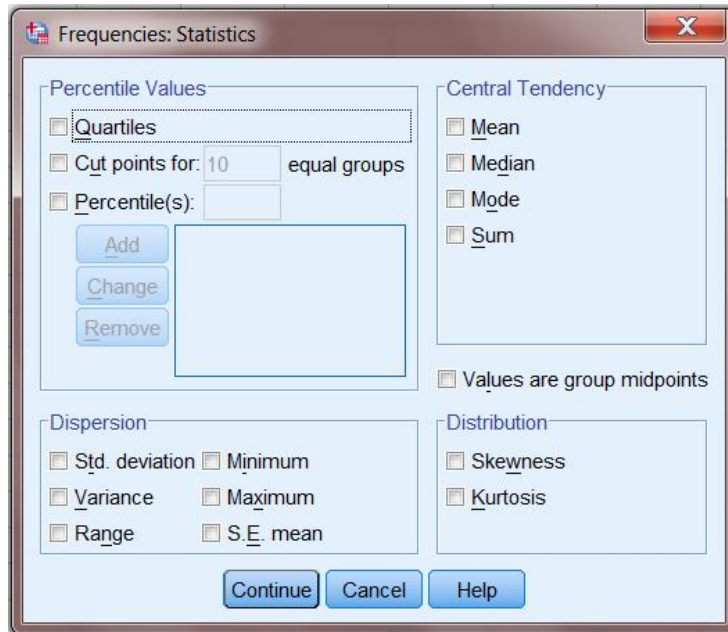
Εικόνα 3.7.1.1: Παράθυρο εντολής Frequencies



Στο αριστερό πλαίσιο επισυνάπτονται τα ονόματα όλων των μεταβλητών ενώ στο δεξί πλαίσιο (Variables), επικολλούνται οι μεταβλητές για τις οποίες επιθυμεί κάποιος να αποκτήσει πίνακες συχνοτήτων. Ένας τρόπος για να γίνει αυτό είναι με διπλό κλικ στη μεταβλητή στο αριστερό πλαίσιο ή διαφορετικά μαρκάρουμε τη μεταβλητή που επιθυμούμε να αναλύσουμε στο αριστερό πλαίσιο, πατάμε το βελάκι και τη μεταφέρουμε στο δεξί πλαίσιο των μεταβλητών προς ανάλυση. Αξίζει να σημειωθεί ότι στην επιλογή Statistics, μπορεί κάποιος να επιλέξει Στατιστικά Μέτρα, που θέλει να υπολογίσει για τις μεταβλητές, όπως φαίνεται στην εικόνα 3.7.1.2. Δηλαδή, μπορεί κανείς να κάνει υπολογισμό συχνοτήτων και να συμπεριλάβει τεταρτημόρια, Σημεία τομής για ισοδύναμες ομάδες, ποσοστά, μέση τιμή (Mean), διάμεσο (Median), επικρατούσα τιμή, άθροισμα, τυπική απόκλιση (St. Deviation), διακύμανση, εύρος (Range), ελάχιστο και μέγιστο, τυπικό σφάλμα μέσου, ασυμμετρία ή κύρτωση. Συνοπτικά, η διάμεσος αποτελεί τη μεσαία τιμή μιας κατανομής, η τυπική απόκλιση δείχνει το ακριβές ποσοστό των περιπτώσεων που βρίσκονται κάτω ή πάνω από τη διάμεσο και καθορίζεται από το ποσοστό των τιμών που ισούνται με τη διάμεσο. Δηλαδή, η τυπική απόκλιση δείχνει πόσο «απλώνονται» οι τιμές της κατανομής γύρω από τον μέσο όρο. Η τυπική απόκλιση εκφράζεται σε μονάδες μέτρησης ίδιες με αυτές του μέσου όρου, και όσο μικρότερη είναι τόσο πιο κοντά στο μέσο όρο είναι οι περισσότερες τιμές της κατανομής. Αν πολλές από τις τιμές της κατανομής βρίσκονται μακριά από τον μέσο όρο, τότε η τυπική απόκλιση είναι μεγαλύτερη. Το εύρος είναι η έκταση του μικρότερου

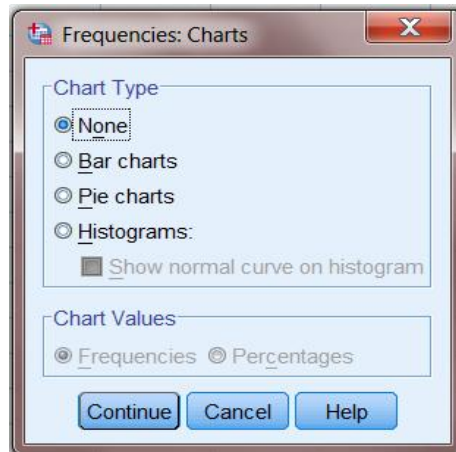
διαστήματος που περιέχει όλα τα δεδομένα και προκύπτει από τη διαφορά της μικρότερης παρατηρούμενης τιμής από τη μεγαλύτερη παρατηρούμενη τιμή.

Εικόνα 3.7.1.2: Παράθυρο επιλογών εντολής Frequencies



Στη συνέχεια, αφού πατήσουμε το πλήκτρο Continue, μπορούμε να μεταβούμε στην επιλογή Charts, όπως διακρίνουμε στην εικόνα 3.7.1.3. Η εντολή αυτή μας δίνει τη δυνατότητα επιλογής της γραφικής απεικόνισης δεδομένων των επιλεγόμενων μεταβλητών. Μια κλασική τεχνική απεικόνισης είναι η γραφική τους παράσταση σε στήλες (Bar), σύμφωνα με τη συχνότητα εμφάνισης τους. Το διάγραμμα αυτό ονομάζεται Bar Chart και επιτρέπει σε κάποιον να έχει μια οπτική αναπαράσταση των συχνοτήτων κάθε κατηγορίας μιας μεταβλητής. Χρησιμοποιείται και σε περιπτώσεις, όπου έχουμε μια μεταβλητή με σχετικά μεγάλο αριθμό κατηγοριών. Μια εναλλακτική γραφική απεικόνιση είναι το διάγραμμα της πίτας (Pie Chart).

Εικόνα 3.7.1.3 : Παράθυρο γραφημάτων εντολής Frequencies



Η πίτα (Pie Chart) απεικονίζει συχνότητες ή ποσοστά σε κάθε κατηγορία. Η πίτα είναι πιο κατάλληλη για τη παρουσίαση μεταβλητών, που έχουν σχετικά λίγες κατηγορίες, διαφορετικά γίνεται δυσδιάκριτη.

Αφού κάποιος ορίσει τις επιλογές του, επιλέγοντας Continue επανέρχεται στο αρχικό παράθυρο των frequencies. Εκεί μπορεί να διατηρήσει την ήδη επιλεγμένη εντολή Display Frequency Tables ή να την ακυρώσει (αποεπιλέγοντάς την) όταν οι μεταβλητές του έχουν πάρα πολλές πιθανές τιμές, πράγμα, που σημαίνει ότι θα προκύψει ένας εξαιρετικά μεγάλος και δύσχρηστος πίνακας, πολλά από τα κελιά του οποίου δεν αποκλείεται να είναι κενά (ανάλογα τον αριθμό των παρατηρήσεων που έχει συγκεντρώσει). Στη συνέχεια, με την επιλογή Format επιλέγεται ο τρόπος εμφάνισης των αποτελεσμάτων(εάν οι τιμές των μεταβλητών στο πίνακα συχνοτήτων θα εμφανίζονται κατά αύξουσα ή κατά φθίνουσα σειρά όσον αφορά την απόλυτη τιμή τους ή την τιμή της συχνότητά τους κλπ.). Τέλος, κάνοντας κλικ στο OK, ΤΟ SPSS για Windows θα τρέξει την ανάλυση, που του έχει ζητηθεί και μόλις αυτή ολοκληρωθεί θα ανοίξει το παράθυρο των αποτελεσμάτων με τα αριθμητικά και γραφικά αποτελέσματα της ανάλυσης.

Μέσω αυτής της επιλογής το SPSS μας δίνει για κάθε τιμή της μεταβλητής, τα ονόματα των οποίων βρίσκονται στην 1η στήλη, τη συχνότητα (Frequency), τη σχετική συχνότητα ως ποσοστό (Percent) επί του συνόλου των περιπτώσεων, το ποσοστό των περιπτώσεων για τα οποία έχουμε τιμή στο αρχείο δεδομένων, δηλαδή δεν είναι ελλειπείς (Valid Percent) και την αθροιστική σχετική συχνότητα



(Cumulative Percent) οποιουδήποτε τύπου μεταβλητών. Η στήλη Cumulative Percent αφορά στην αθροιστική σχετική συχνότητα των τιμών της μεταβλητής στο δείγμα και δεν έχει νόημα για όλες τις μεταβλητές (παρότι το SPSS υπολογίζει την αθροιστική σχετική συχνότητα κάθε φορά που θα του ζητήσετε να κατασκευάσει μία κατανομή συχνότητας χρησιμοποιώντας την εντολή Frequencies). Η αθροιστική σχετική συχνότητα προφανώς δεν έχει εφαρμογή σε μεταβλητές που έχουν μόνο δύο κατηγορίες αλλά και γενικώς σε όσες μεταβλητές δεν έχει νόημα η άθροιση των σχετικών συχνοτήτων.<sup>7</sup>

Σε περίπτωση που δεν υπάρχουν ελλειπείς τιμές, τότε η στήλη των σχετικών συχνοτήτων ταυτίζεται με τη στήλη Percent και Valid Percent. Στο σημείο αυτό είναι απαραίτητο να διευκρινίσουμε ότι η χρήση αυτής της επιλογής θα πρέπει να γίνεται για διακριτές μεταβλητές ή μεταβλητές με περιορισμένο αριθμό κατηγοριών.

### 3.7.2 Περιγραφή εντολής Descriptives

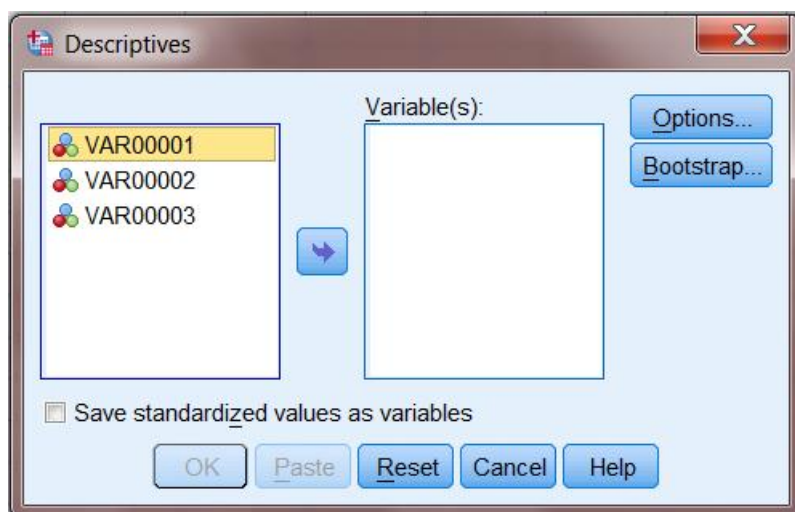
Αυτή η εντολή εφαρμόζεται μόνο σε αριθμητικές μεταβλητές. Με αυτή μπορεί κάποιος να υπολογίσει τα στατιστικά μέτρα, που υπολογίζει η εντολή Frequencies εκτός από τη διάμεσο και την επικρατούσα τιμή. Επίσης, υπολογίζει επιπλέον τις τυποποιημένες τιμές (τα z-scores).

Η επιλογή της συγκεκριμένης εντολής γίνεται ύστερα από την εξής ακολουθία εντολών Analyze - Descriptive Statistics – Descriptives. Στο παράθυρο που ανοίγει και φαίνεται στην εικόνα 3.7.2.1, η επιλογή των μεταβλητών που θέλουμε να αναλύσουμε γίνεται με την ίδια διαδικασία με την εντολή Frequencies. Κατά τον ίδιο τρόπο στην επιλογή Options επιλέγονται τα στατιστικά μέτρα που πρέπει να εμφανισθούν και η σειρά ταξινόμησης στην παρουσίαση των αποτελεσμάτων. Τέλος, με την επιλογή «Save standardized values as variables» ζητάει κανείς να υπολογισθούν και να διατηρηθούν οι τυποποιημένες τιμές Z.

---

<sup>7</sup> Ρούσσοις Π., Ευσταθίου Γ., «Σύντομο Εγχειρίδιο SPSS 16.00» Πρόγραμμα Ψυχολογίας, Τμήμα ΦΠΨ, Εθνικό Καποδιστριακό Πανεπιστήμιο Αθηνών, Αθήνα, 2008

Εικόνα 3.7.2.1: Παράθυρο εντολής Descriptives



### 3.7.3 Περιγραφή εντολής Crosstabs

Με την εντολή Crosstabs μπορεί κάποιος να δημιουργήσει πίνακες συνάφειας και να παράγει διάφορα στατιστικά ελέγχου των πιθανών σχέσεων των υπό εξέταση μεταβλητών. Είναι ένας καλός τρόπος παρουσίασης δύο μεταβλητών, ώστε να μπορεί κάποιος να διερευνήσει την ύπαρξη σχέσης ανάμεσά τους. Γενικά, στις περιπτώσεις που κάποιος θέλει να δει τα στοιχεία, που προκύπτουν μετά την διασταύρωση δύο ή περισσότερων μεταβλητών δημιουργεί πίνακες διπλής εισόδου ή και μεγαλύτερους με την βοήθεια της εντολής Crosstabs. Μέσα στα κελιά του πίνακα μπορεί κανείς να ζητήσει την εμφάνιση, εκτός του αριθμού-count των εγγραφών, τα ποσοστά γραμμών και στηλών, τις αναμενόμενες – expected τιμές σε περίπτωση ανεξαρτησίας, καθώς επίσης και τα υπόλοιπα – residuals (δηλ αναμενόμενη τιμή κελιού – πραγματική τιμή). Τέλος, μπορεί να ζητήσει κάποιος τον υπολογισμό στατιστικών μέτρων για να κάνει ελέγχους συσχέτισης.

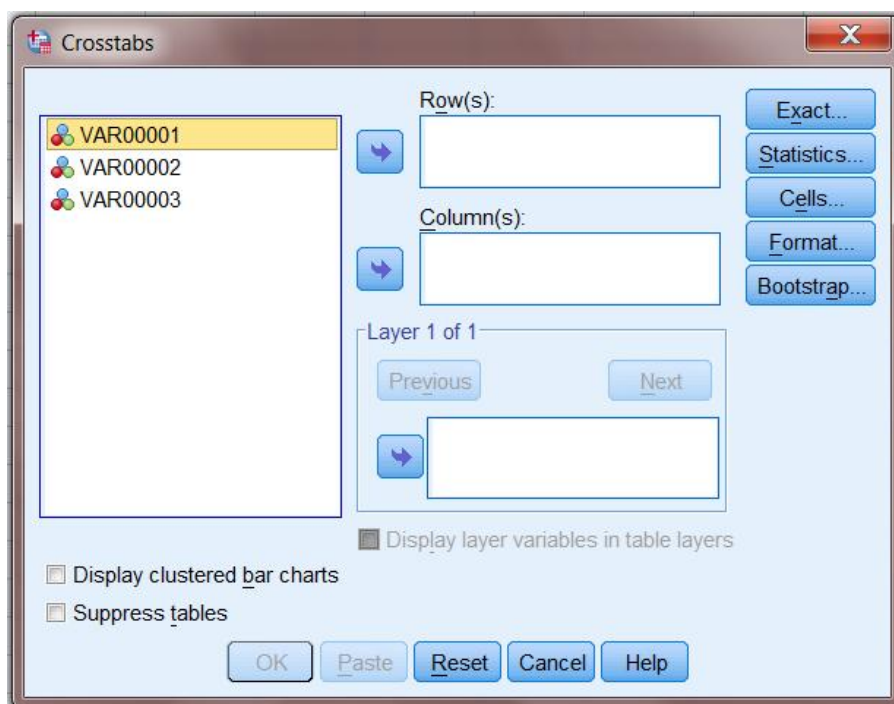
Η επιλογή της συγκεκριμένης εντολής γίνεται ύστερα από την εξής ακολουθία εντολών **Analyse - Descriptive Statistics – Crosstabs**. Όπως φαίνεται στην εικόνα 3.7.3.1, επιλέγουμε με τον τρόπο που προαναφέρθηκε τις μεταβλητές που θέλουμε να είναι στη γραμμή και τις μεταβλητές που θέλουμε να είναι σε στήλη. Με την εντολή Cells μπορεί κάποιος να ορίσει, ποια ποσοστά θέλει να εμφανίζονται στον πίνακα διπλής εισόδου. Με την επιλογή **Statistics** ορίζει κάποιος ποια μέτρα συνάφειας(συσχέτισης) θέλει να χρησιμοποιήσει.

Τα μέτρα ή συντελεστές συνάφειας είναι τα στατιστικά μέτρα με τα οποία μπορεί κανείς να διακρίνει, εάν υπάρχει σχέση μεταξύ δύο ή περισσότερων μεταβλητών και σε πιο βαθμό μπορεί κάποιος γνωρίζοντας τις τιμές της μιας μεταβλητής να προβλέψει τις τιμές της άλλης μεταβλητής. Αυτά τα μέτρα ή συντελεστές συνάφειας παίρνουν τιμές από  $-1$  μέχρι  $+1$ . Το 0 δείχνει παντελή έλλειψη σχέσης, το  $+1$  δείχνει τέλεια θετική σχέση και το  $-1$  τέλεια αρνητική σχέση μεταξύ δυο μεταβλητών. Η επιλογή του κατάλληλου μέτρου εξαρτάται από διάφορους παράγοντες εκ των οποίων δύο είναι πολύ σημαντικοί:

- **το επίπεδο μέτρησης** της μεταβλητής και
- αν η σχέση είναι **συμμετρική ή ασύμμετρη**.

Στην πράξη το αν μια σχέση είναι συμμετρική ή ασύμμετρη εξαρτάται από την ερώτηση, που θέλει κανείς να απαντήσει. Αν η ερώτηση είναι “**υπάρχει σχέση μεταξύ δύο μεταβλητών**”, τότε η σχέση είναι συμμετρική. Αν η ερώτηση είναι **αν μία μεταβλητή επηρεάζει μία άλλη** τότε η σχέση είναι ασύμμετρη.

Εικόνα 3.7.3.1: Παράθυρο εντολής Crosstabs



Στα πλαίσια Rows και Columns βάζουμε τις μεταβλητές που θα συμμετάσχουν στην ανάλυση. Συνήθως είναι προτιμότερο να βάλουμε στο κουτί Rows τη μεταβλητή με τις περισσότερες κατηγορίες, ώστε ο πίνακας διπλής εισόδου που θα προκύψει να εκτείνεται καθ' ύψος και όχι κατά πλάτος.

Στο δεξιό τμήμα του κουτιού βρίσκονται τέσσερα κουμπιά, από τα οποία το σημαντικότερο είναι το κουμπί Statistics. Πρέπει να κάνουμε κλικ σε αυτό και να επιλέξουμε Chi-square, αλλιώς το στατιστικό κριτήριο δε θα υπολογιστεί.<sup>8</sup>

Στον πίνακα των αποτελεσμάτων παρατίθεται και ένας πίνακας που ονομάζεται **chi-square** δηλαδή  $\chi^2$ . Αυτός μας δίνει τη στατιστική σημαντικότητα δηλαδή μας

<sup>8</sup> Ρούσσος Π., Ευσταθίου Γ., «Σύντομο Εγχειρίδιο SPSS 16.00» Πρόγραμμα Ψυχολογίας, Τμήμα ΦΠΨ, Εθνικό Καποδιστριακό Πανεπιστήμιο Αθηνών, Αθήνα, 2008

βοηθάει στην ερμηνεία, εάν υπάρχει συστηματική σχέση (όχι τυχαία), άρα εξάρτηση μεταξύ των δυο εξεταζόμενων μεταβλητών. Στο SPSS για Windows δίνεται κατευθείαν, η σημαντικότητα Significance, η οποία για να στοιχειοθετηθεί θα πρέπει να είναι μικρότερη από το 5% (sign  $\leq 0.05$ ). Σημαντικότητα της τάξης του 5% σημαίνει, ότι στις 100 περιπτώσεις συσχέτισης των δύο μεταβλητών μόνο σε πέντε περιπτώσεις θα κάνει κάποιος λάθος (έχοντας την κατηγορία της μίας μεταβλητής μπορεί να προβλέψει κανείς με ακρίβεια 95% σε ποια κατηγορία της δεύτερης μεταβλητής ανήκει η εξεταζόμενη περίπτωση).

Στον πίνακα του **x<sup>2</sup> (chi-square)** υπάρχουν εκτός από το **Pearson chi-square** και άλλοι δύο δείκτες του **x<sup>2</sup>**. Ο δείκτης **Likelihood Ratio** χρησιμοποιείται στις περιπτώσεις, που οι μεταβλητές μετριοούνται σε λογαριθμική κλίμακα. Για μεγάλα δείγματα ο δείκτης **Likelihood Ratio** δίνει το ίδιο αποτέλεσμα με το **Pearson chi-square**. Ο τρίτος δείκτης **Linear by Linear Association** χρησιμοποιείται όταν και οι δύο μεταβλητές είναι ποσοτικές.<sup>9</sup>

### 3.7.4 Παρουσίαση εντολής Correlation

Στην περίπτωση δύο ποσοτικών (συνεχών) μεταβλητών ένας τρόπος μελέτης της σχέσης των δύο αυτών μεγεθών μπορεί να υποστηριχθεί από την εντολή Correlation.

Ο συντελεστής συσχέτισης (Correlation Coefficient) μετράει ποσοτικά αφενός την ύπαρξη κάποιας γραμμικής σχέσης και αφετέρου την ένταση αυτής. Είναι γνωστός στην στατιστική ανάλυση με το σύμβολο  $r$  και η τιμή του κυμαίνεται από -1 (τέλεια αρνητική συσχέτιση) έως +1 (τέλεια θετική συσχέτιση). Τιμές κοντά στο 0 δηλώνουν

τη μη ύπαρξη γραμμικής συσχέτισης. Ο συντελεστής συσχέτισης πρέπει να χρησιμοποιείται αποκλειστικά για γραμμικές σχέσεις, διαφορετικά μπορεί να δώσει τελείως λανθασμένη εικόνα.<sup>10</sup>

Ο συντελεστής συσχέτισης χρησιμοποιείται είτε για να εκφράσει την ένταση της γραμμικής σχέσης είτε για να ελεγχθεί από το δείγμα αν υπάρχει η ίδια σχέση και

---

<sup>9</sup> Νέλλας Ε., «Ανάλυση Δεδομένων με χρήση του πακέτου SPSS/PC για WINDOWS», Τμήμα Αγροτικής Οικονομίας και Ανάπτυξης, Γεωπονικό Πανεπιστήμιο Αθηνών, Αθήνα, 2005

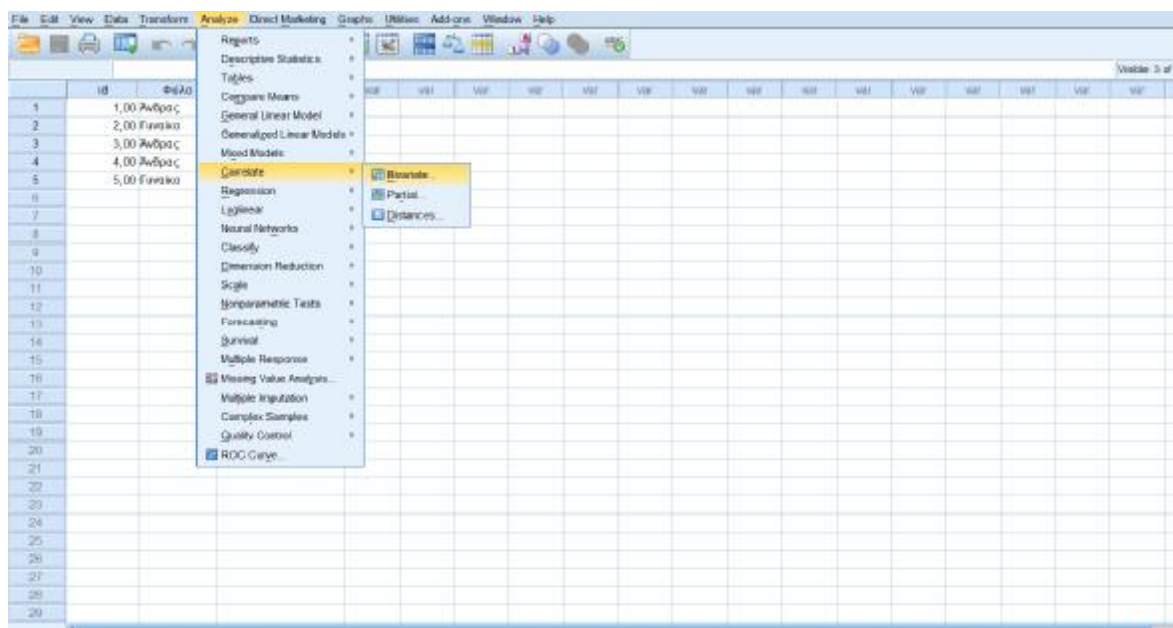
<sup>10</sup> Τσομπανάκη Ε., «Εισαγωγή στη χρήση του στατιστικού πακέτου SPSS 14.00»

στον πληθυσμό. Η ύπαρξη θετικής συσχέτισης δεν προϋποθέτει ότι η μια μεταβλητή προκαλεί την άλλη.<sup>11</sup>

Για να υπολογίσουμε το συντελεστή συσχέτισης, επιλέγουμε Analyze - Correlate – Bivariate (Εικόνα 3.7.4.1).

Εικόνα 3.7.4.1: Παράθυρο εντολής

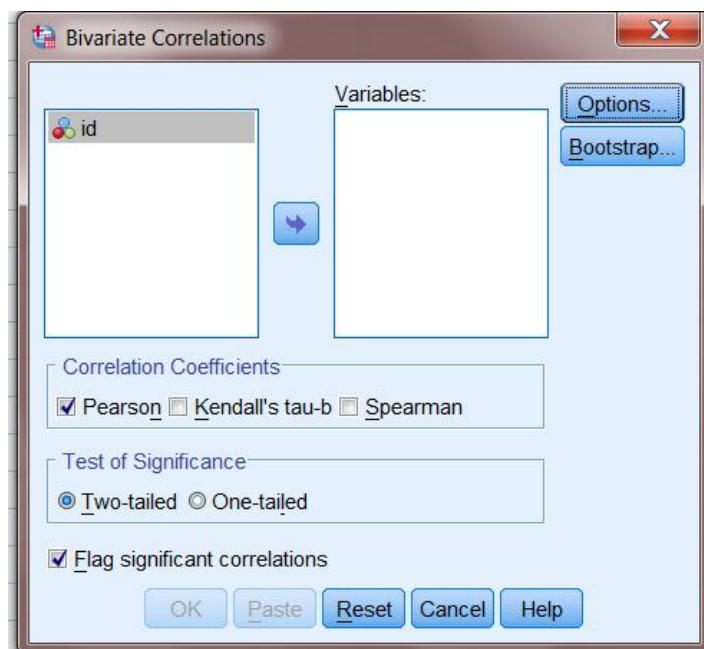
#### Correlate



Στη συνέχεια, επιλέγουμε τις μεταβλητές που θέλουμε να εντοπίσουμε αν υπάρχει συσχέτιση καθώς και τους συντελεστές Pearson, Kendall's tau-b και Spearman, όπως φαίνεται στην Εικόνα 3.7.4.2.

<sup>11</sup>Φαχρίδης Γ., «Εισαγωγή στο πρόγραμμα SPSS for Windows», Παιδαγωγικό Τμήμα Δημοτικής Εκπαίδευσης, Δημοκρίτειο Πανεπιστήμιο Θράκης, Αλεξανδρούπολη

Εικόνα 3.7.4.2: Παράθυρο εντολής Crosstabs



Ο συντελεστής συσχέτισης **Pearson** δείχνει τη γραμμική σχέση μεταξύ 2 μεταβλητών. Απαραίτητες προϋποθέσεις χρήσης του είναι:

- Ύπαρξη μεταβλητών scale
- Οι μεταβλητές να κατανέμονται κανονικά στον πληθυσμό (ειδικά αν το δείγμα είναι μικρό)
- Οι μεταβλητές να έχουν γραμμική σχέση (αυτό διαπιστώνεται από το διάγραμμα διασποράς – scatter diagram )

Η χρήση του Pearson πρέπει να αποφεύγεται όταν η μία ή και οι δύο μεταβλητές είναι nominal.

Το μέτρο που καθορίζει αυτή τη σχέση είναι ο συντελεστής συσχέτισης του Pearson, ο οποίος παίρνει τιμές από -1 έως +1. Όσο μεγαλύτερη είναι η απόλυτη τιμή του συντελεστή αυτού, τόσο πιο ισχυρή είναι η σχέση των δύο μεταβλητών καθώς και η πρόβλεψη της μίας μεταβλητής με βάση την άλλη. Όταν ο συντελεστής παίρνει αρνητικές τιμές, τότε έχουμε αρνητική συσχέτιση, δηλαδή όταν οι τιμές της μίας μεταβλητής αυξάνονται, οι τιμές της άλλης μειώνονται. Όταν ο συντελεστής παίρνει

θετικές τιμές, τότε έχουμε θετική συσχέτιση, δηλαδή όταν οι τιμές της μίας μεταβλητής αυξάνονται, αυξάνονται και της άλλης. Εδώ, πρέπει να επισημάνουμε ότι τέλειες συσχετίσεις στις οποίες η τιμή του συντελεστή είναι  $-1$  ή  $+1$  είναι αδύνατον να βρεθούν. Συσχετίσεις όπου η απόλυτη τιμή του συντελεστή συσχέτισης Pearson βρίσκεται στο διάστημα  $[0,0.2]$  χαρακτηρίζονται ως ασήμαντες, στο διάστημα  $(0.2,0.4]$  ως μέτριες, στο  $(0.4,0.7]$  ως σημαντικές και στο  $(0.7,1)$  ως ισχυρές.

**ΠΡΟΣΟΧΗ:** Στην περίπτωση που ο συντελεστής συσχέτισης είναι 0 δεν σημαίνει ότι δεν υπάρχει συσχέτιση, απλώς λέμε ότι δεν υπάρχει γραμμική συσχέτιση μεταξύ των μεταβλητών.

Ο συντελεστής συσχέτισης Kendall's Tau b χρησιμοποιείται στην περίπτωση που οι 2 μεταβλητές είναι ordinal (π.χ. κατάταξη 5 διαφορετικών μορφών διδασκαλίας πάνω στο ίδιο αντικείμενο από φοιτητές ανάλογα με τη στάση δύο επιβλεπόντων καθηγητών απέναντι σ' αυτές). Ο συντελεστής αυτός στηρίζεται στην απόσταση μεταξύ των ζευγών των παρατηρήσεων που έχουν την ίδια κατεύθυνση και των ζευγών των παρατηρήσεων που έχουν αντίθετη κατεύθυνση. Αν τα ζεύγη προς την ίδια κατεύθυνση υπερέχουν αυτών της αντίθετης κατεύθυνσης τότε ο Kendall tau έχει θετικό πρόσημο. Οι τιμές που παίρνει είναι από  $-1$  έως  $+1$ .

### 3.7.5 Παρουσίαση εντολής Regression<sup>12</sup>

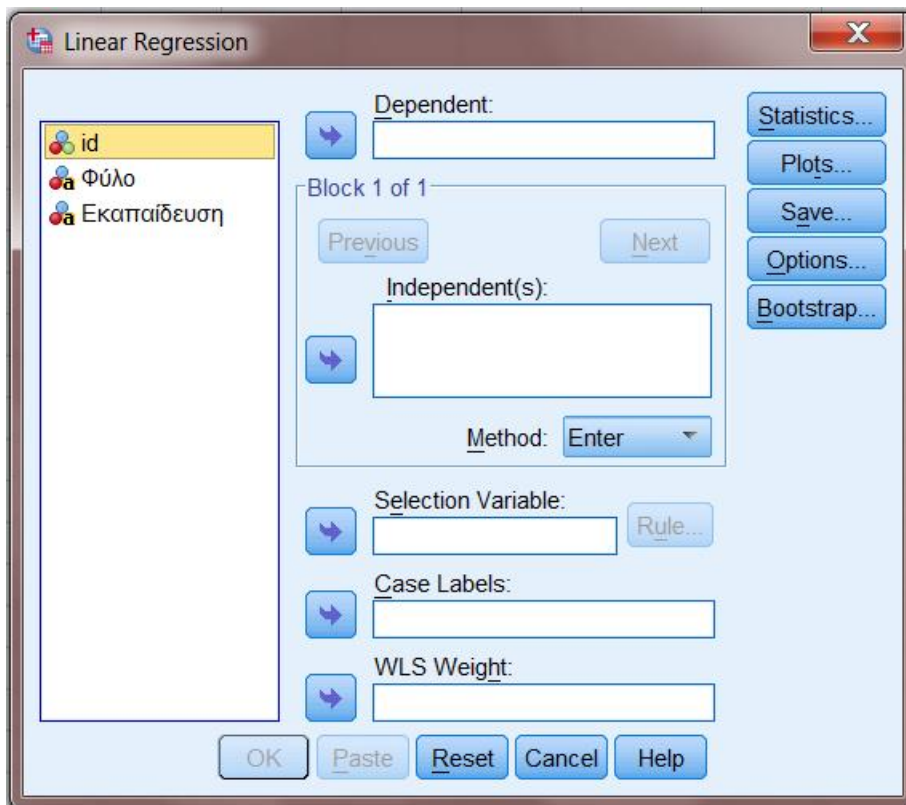
Για την παρουσίαση της απλής γραμμικής παλινδρόμησης επιλέγουμε την εξής ακολουθία εντολών: Analyze – Regression – Linear. Αφού ανοίξει το πλαίσιο διαλόγου που φαίνεται στην εικόνα 2.7.5.1, και τοποθετούμε στο πλαίσιο κειμένου Dependent την εξαρτημένη μεταβλητή και στο πλαίσιο Independents την ανεξάρτητη μεταβλητή. Στην περίπτωση που θέλουμε να εκτελέσουμε πολλαπλή γραμμική παλινδρόμηση, μετακινούμε όλες τις μεταβλητές που επιθυμούμε.

---

<sup>12</sup> Ρούσσος Π., Ευσταθίου Γ., «Σύντομο Εγχειρίδιο SPSS 16.00» Πρόγραμμα Ψυχολογίας, Τμήμα ΦΠΨ, Εθνικό Καποδιστριακό Πανεπιστήμιο Αθηνών, Αθήνα, 2008



Εικόνα 3.7.5.1: Παράθυρο εντολής Regression



Στην απλή γραμμική παλινδρόμηση υπάρχει μια εξαρτημένη μεταβλητή (dependent) που συχνά αποκαλείται και μεταβλητή response και μια ανεξάρτητη μεταβλητή (independent) που συχνά αποκαλείται και μεταβλητή explanatory. Η εξαρτημένη μεταβλητή μπορεί να είναι ordinal ή scale. Η ανεξάρτητη μπορεί να είναι επίσης ordinal ή scale. Η συνάρτηση της απλής γραμμικής παλινδρόμησης έχει τη μορφή:

$$Y = \alpha + \beta \cdot X$$

Όπου

Y = εξαρτημένη μεταβλητή

X= ανεξάρτητη μεταβλητή

$\alpha$  = σταθερά του μοντέλου

$\beta$  = συντελεστής παλινδρόμησης

Ο συντελεστής παλινδρόμησης  $\beta$  αντιπροσωπεύει το μέγεθος της μεταβολής της εξαρτημένης μεταβλητής όταν η ανεξάρτητη μεταβλητή μεταβληθεί κατά μία μονάδα (π.χ. αν η βαθμολογία ενός φοιτητή αυξηθεί κατά 1 μονάδα στα μαθηματικά τότε η βαθμολογία του στη στατιστική μεταβάλλεται κατά  $\beta$ ). Η σταθερά  $a$  δείχνει την τιμή της εξαρτημένης μεταβλητής όταν η ανεξάρτητη μεταβλητή γίνει 0 (δηλ. αν κάποιος πάρει 0 στα μαθηματικά τότε ο βαθμός του στη στατιστική θα είναι  $a$ ).

Στα αποτελέσματα της ανωτέρω διαδικασίας παρουσιάζονται σε μορφή πινάκων τα ακόλουθα στοιχεία:

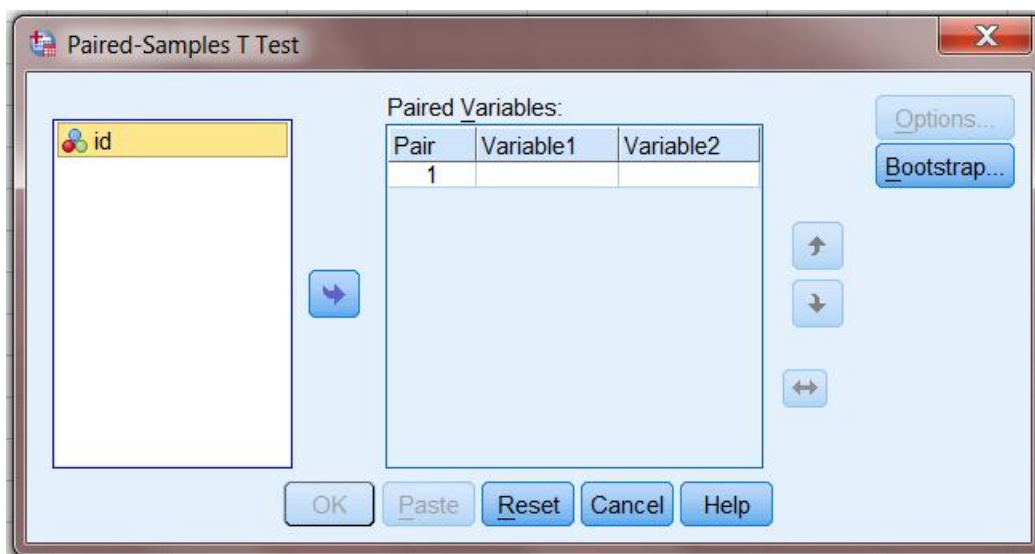
- Πίνακας 1: Παρουσίαση των μεταβλητών της ανάλυσης
- Πίνακας 2: Παρουσιάζεται η ποσότητα  $R$ , που αποτελεί τον δείκτη συσχέτισης μεταξύ των μεταβλητών, τον συντελεστή προσδιορισμού  $R^2$  και τον συντελεστή Adjusted  $R^2$ .
- Πίνακας 3: Ο πίνακας ANOVA δείχνει τη στατιστική σημαντικότητα του λόγου  $F$ . Αυτό το τεστ αξιολογεί την υπόθεση αν υπόθεση είναι στατιστικά σημαντική ή τυχαία.
- Πίνακας 4: Ο πίνακας Coefficients παρουσιάζει με το σύμβολο  $B$  την κλίση της γραμμής παλινδρόμησης. Ακόμη, στον πίνακα υπάρχουν τα αποτελέσματα από 2  $t$ -τεστ: το πρώτο από αυτά εξετάζει την υπόθεση ότι η τιμή της σταθεράς είναι ίση με το μηδέν, ενώ το δεύτερο ελέγχει την υπόθεση ότι η κλίση είναι ίση με το μηδέν. Στατιστικά σημαντική κλίση σημαίνει στατιστικά προβλέψιμη.

### 3.7.6 Το $t$ -test

Υπάρχουν δύο ειδών  $t$ -τεστ, ένα για ανεξάρτητα δείγματα και ένα για εξαρτημένα δείγματα. Προκειμένου να ενημερωθεί το SPSS για αυτήν την πληροφορία, χρειάζεται να ακολουθηθούν δύο διαφορετικοί τρόποι εισαγωγής δεδομένων. Για εξαρτημένα δείγματα πρέπει να εισαχθούν σε δύο στήλες του επεξεργαστή δεδομένων του SPSS. Για ανεξάρτητα δείγματα χρησιμοποιείται μια στήλη για τις μετρήσεις και μια δεύτερη στήλη στην οποία καθορίζεται η υπο-ομάδα στην οποία ανήκουν οι συμμετέχοντες.

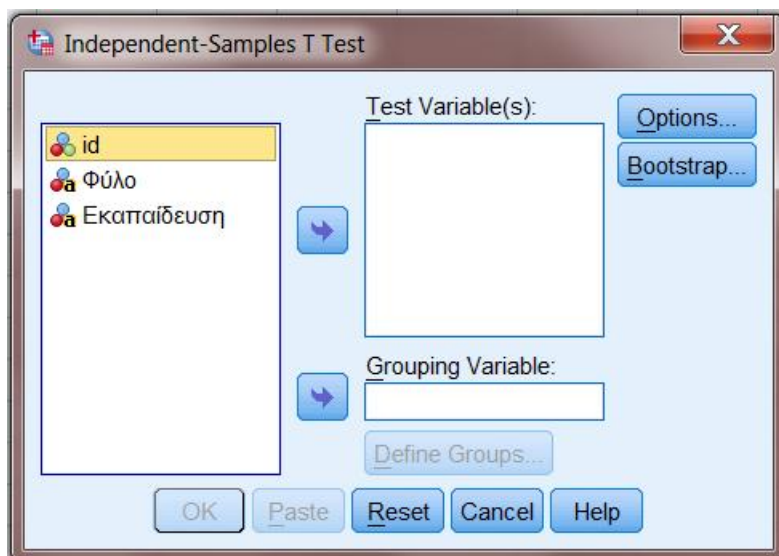
Το t-τεστ για **εξαρτημένα δείγματα** εφαρμόζεται στις περιπτώσεις που έχει μετρηθεί η ίδια μεταβλητή στους ίδιους συμμετέχοντες σε δύο διαφορετικές χρονικές στιγμές ή όταν έχουν ληφθεί μετρήσεις από τους ίδιους συμμετέχοντες για δύο μεταβλητές, χρησιμοποιώντας την ίδια κλίμακα μέτρησης. Αφού εισάγουμε τα δεδομένα στον πίνακα, επιλέγουμε την ακολουθία εντολών: Analyze – Compare Means – Paired Samples T Test, όπως φαίνεται στην ακόλουθη εικόνα.

Εικόνα 3.7.6.1: Το t-τεστ για εξαρτημένα δείγματα



Το **t-τεστ για ανεξάρτητα δείγματα**, εφαρμόζεται στις περιπτώσεις που θέλουμε να συγκρίνουμε δύο ομάδες μετρήσεων από διαφορετικά άτομα (δείγματα) ως προς μία εξαρτημένη μεταβλητή. Η διαδικασία που ακολουθείται στο Spss είναι η ακόλουθη: Σε μία στήλη περνάμε την ανεξάρτητη μεταβλητή για να δηλώσουμε την ομάδα στην οποία ανήκουν οι συμμετέχοντες και σε μια δεύτερη στήλη (Εξαρτημένη μεταβλητή) καταχωρείται η μέτρηση που έχουμε για κάθε συμμετέχοντα. Για να υπολογιστεί το στατιστικό κριτήριο t, ακολουθούμε την εξής διαδικασία: Analyze – Compare Means – Independent –Samples T Test, οπότε και παίρνουμε την ακόλουθη εικόνα.

Εικόνα 3.7.6.2: Το t-τεστ για ανεξάρτητα δείγματα



Αξίζει να αναφέρουμε ότι το **τεστ Wilcoxon** είναι το μη παραμετρικό ισοδύναμο t-τεστ για εξαρτημένα δείγματα το οποίο εφαρμόζεται στις ίδιες περιπτώσεις και είναι αναγκαίο να έχουν καταχωρηθεί τα δεδομένα με τον ίδιο τρόπο. Η ακολουθία εντολών είναι η ακόλουθη: Analyze – Nonparametric Tests – 2 Related Samples.

Επίσης το **τεστ Mann-Whitney U** είναι το μη παραμετρικό ισοδύναμο του t-τεστ ανεξάρτητων ομάδων το οποίο επίσης εφαρμόζεται στις ίδιες περιπτώσεις και χρειάζεται να έχουν καταχωρηθεί τα δεδομένα με τον ίδιο ακριβώς τρόπο. Η ακολουθία εντολών είναι η εξής: Analyze – Compare Means – Means.

Το τεστ Kruskal – Wallis H είναι το μη παραμετρικό ισοδύναμο της ανάλυσης διακύμανσης μονής κατεύθυνσης το οποίο εφαρμόζεται στις ίδιες περιπτώσεις και απαιτείται να έχουν καταχωρηθεί τα δεδομένα με τον ίδιο ακριβώς τρόπο. Η ακολουθία εντολών είναι η εξής: Analyze – Nonparametric Tests – K Independent Samples.

Το κριτήριο Friedman είναι το μη παραμετρικό ισοδύναμο για την ανάλυση διακύμανσης επαναληπτικών μετρήσεων το οποίο εφαρμόζεται στις ίδιες περιπτώσεις και απαιτείται να έχουν καταχωρηθεί τα δεδομένα με τον ίδιο ακριβώς τρόπο. Η ακολουθία εντολών είναι η εξής: Analyze – Nonparametric Tests – K Related Samples.

### 3.7.7 Παρουσίαση των αποτελεσμάτων στατιστικής ανάλυσης

Τα αποτελέσματα της στατιστικής ανάλυσης δεδομένων στο στατιστικό πρόγραμμα SPSS, παρουσιάζονται στο μενού **Output**, όπως προαναφέρθηκε. Το μενού αυτό αποτελεί ταυτόχρονα και τον αντικειμενικός στόχος όλων των αναλύσεων δεδομένων. Το SPSS έχει μια μακρά και πλούσια ιστορία προσπαθειών προκειμένου να δημιουργήσει μια μορφή των αποτελεσμάτων που θα είναι ξεκάθαρη και ταυτόχρονα περιεκτική. Από την έκδοση 7.5 και μετά χρησιμοποιεί editor που εμφανίζει τα αποτελέσματα σε μορφή πινάκων. Ο Output Navigator ή Παράθυρο Αποτελεσμάτων αποτελείται από δύο μέρη. Το αριστερό μέρος, το οποίο μας δείχνει τη διάρθρωση των αποτελεσμάτων και το δεξιό μέρος στο οποίο περιέχονται τα ίδια τα αποτελέσματα μετά την εκτέλεση των στατιστικών διαδικασιών και των διαδικασιών παραγωγής γραφικών. Το αρχείο αποτελεσμάτων ανοίγει αυτόματα μόλις ξεκινήσει το SPSS ή μόλις εκτελεσθεί η πρώτη εντολή παραγωγής αποτελεσμάτων. Μπορεί κανείς να διορθώσει τα περιεχόμενα ενός παραθύρου αποτελεσμάτων, όταν βέβαια δεν χρησιμοποιείται από τον επεξεργαστή. Όπως ήδη αναφέρθηκε μπορεί ακόμα κάποιος να αποθηκεύσει τα αποτελέσματα σαν αρχείο αποτελεσμάτων με προέκταση \*.spo ή να τα εξάγει με την μορφή **Hypertext (υπερκειμένου)** και να τα χρησιμοποιήσει απευθείας στο Word ή σε ένα άλλο λογισμικό. Επίσης, μπορεί να έχει κάποιος ανοιχτά περισσότερα του ενός αρχεία αποτελεσμάτων, αλλά μόνο ένα από αυτά μπορεί να καθορίσει σαν αρχείο προορισμού των αποτελεσμάτων που παράγονται.

Ο **Output Navigator** χρησιμοποιείται για τα εξής:

- Παρουσίαση των αποτελεσμάτων
- Εμφάνιση ή απόκρυψη των πινάκων και γραφικών
- Αλλαγή της σειράς εμφάνισης των αποτελεσμάτων με μετακίνηση των επιλεγμένων στοιχείων.
- Μετάβαση στο Pivot Editor ή Text output editor ή Chart editor.
- Μεταφορά στοιχείων από το Spss σε άλλες εφαρμογές

Στο παράθυρο αποτελεσμάτων υπάρχει μια στάνταρ μπάρα εργαλείων για την εύκολη εκτέλεση εργασιών. Όπως φαίνεται στην Εικόνα 3.7.7.1, με τη βοήθεια αυτής της μπάρας μπορούμε να κάνουμε προεπισκόπηση των αποτελεσμάτων εκτύπωσης, να

εξάγουμε τα αποτελέσματα σε Hypertext, να επιλέξουμε το τελευταίο παραχθέν αποτέλεσμα κ.ά. Το τελευταίο κουμπί με το θαυμαστικό καθορίζει το παράθυρο προορισμού. Εάν είναι έντονο και κάποιος το πατήσει, το συγκεκριμένο παράθυρο αποτελεσμάτων γίνεται παράθυρο προορισμού. Μπορεί να υπάρχουν ανοιχτά πολλά παράθυρα αποτελεσμάτων, αλλά μόνο **ένα** μπορεί να καθορισθεί σαν παράθυρο προορισμού των αποτελεσμάτων. Κάνοντας κλικ στο θαυμαστικό, που υπάρχει στη μπάρα εικονιδίων το θαυμαστικό μπαίνει πριν τον τίτλο. Το παράθυρο προορισμού δεν πρέπει να συγχέεται με το **ενεργό** παράθυρο, το οποίο είναι απλά το τελευταίο παράθυρο, που έχει επιλέξει κάποιος και είναι φωτεινή η μπάρα του τίτλου του. Με χρήση των εντολών **Close**, **Save** ή **Save As** του μενού **File** μπορεί κάποιος να κλείσει ή να αποθηκεύσει τα αποτελέσματά του, αφού φυσικά είναι ενεργό το παράθυρο Output.

Εικόνα 3.7.7.1: Παράθυρο Viewer Outline



Μετακινεί ένα επιλεγμένο τμήμα κατά ένα επίπεδο πριν ή μετά στο σχήμα διάρθρωσης

Ανοίγει ή κλείνει ένα επιλεγμένο τμήμα

Εμφανίζει ή κρύβει συγκεκριμένα τμήματα

Εισάγει επικεφαλίδα, τίτλο ή κείμενο

Όπως έχει αναφερθεί το αριστερό μέρος του παραθύρου των αποτελεσμάτων μας δείχνει τη διάρθρωσή τους. Είναι στην ουσία ένας βοηθητικός πίνακας περιεχομένων.

Οι ενέργειες, που εφαρμόζονται πάνω στον πίνακα αλλάζουν αντίστοιχα και την εμφάνιση των αποτελεσμάτων. Για την καλύτερη διαχείριση των αποτελεσμάτων εκτός από τη Standard μπάρα υπάρχει και η μπάρα διαχείρισης των αποτελεσμάτων, που αφορά το μέρος της διάρθρωσης του Navigator. Η μπάρα αυτή δεν εμφανίζεται παρά μόνον αν ακυρώσει κάποιος την επιλογή **Outline Toolbar** από το μενού **View**. Με τη βοήθεια της μπάρας αυτής μπορούμε να εισάγουμε έναν τίτλο κειμένου, να αποκρύψουμε τα περιεχόμενα επιλεγμένου τμήματος, να ανοίγουμε ή να κλείνουμε επιλεγμένα τμήματα στη λίστα διάρθρωσης κ.ά.

Είναι ακόμη δυνατή η επεξεργασία των πινάκων με τα παραγόμενα αποτελέσματα.

Συγκεκριμένα μπορούμε να προβούμε στις ακόλουθες ενέργειες:

- Αλλαγή των γραμμών σε στήλες και αντίστροφα
- Αλλαγή της μεταβλητής ελέγχου σε μεταβλητή ορισμού γραμμής ή στήλης
- Αλλαγή της γραμμής σε μεταβλητή ελέγχου ή στήλης
- Αλλαγή της στήλης σε μεταβλητή ελέγχου ή γραμμής
- Τροποποίηση των επιπέδων διάταξης
- Απόκρυψη ή εμφάνιση των κελιών του πίνακα ή και των περιεχομένων τους
- Εμφάνιση ή όχι πλέγματος, επικεφαλίδων, τίτλων, υποσημειώσεων
- Τροποποίηση του μεγέθους του πίνακα
- Επαναφορά πίνακα στην αρχική του μορφή
- Απόκρυψη ή εμφάνιση ετικετών των Διαστάσεων ή των κατηγοριών
- Αλλαγές στην εμφάνιση των πινάκων

### **3.8 Παραγωγή γραφημάτων και διαγραμμάτων**

Τα διαγράμματα είναι πολύ σημαντικά αλλά συχνά παραμελημένα. Αν και είναι πολύ δύσκολο να αναπαραστήσει κάποιος γραφικά όλες μαζί τις μεταβλητές του, μπορεί εναλλακτικά να κάνει διαγράμματα για την κάθε μεταβλητή ξεχωριστά και για ζευγάρια μεταβλητών.

Η χρησιμότητα των διαγραμμάτων είναι μεγάλη και μπορούν να κατανοηθούν σχεδόν από όλους. Σημαντική είναι και η πληροφορία που περικλείουν, καθώς μπορεί να εντοπίσει κάποιος σχέσεις μεταβλητών και να ερμηνεύσει τη συμπεριφορά των υπό μελέτη αντικειμένων.

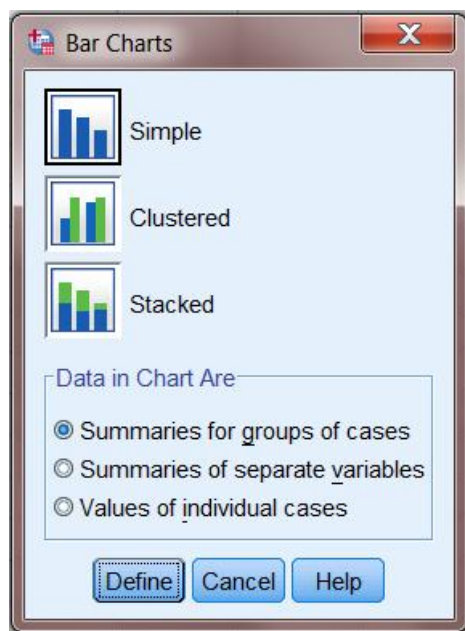
Επιπλέον, με τη χρήση των σύγχρονων Η/Υ μπορεί κανείς να κάνει πολύπλοκα τρισδιάστατα διαγράμματα, ενώ υπάρχουν και μερικά γραφήματα για την ταυτόχρονη απεικόνιση πολλών μεταβλητών (scatter plot matrix, star plot, chernoff faces).

#### **3.8.1 Μονοδιάστατα διαγράμματα**

##### **3.8.1.1 Ραβδογράμματα (Chart Bar)**

Το συγκεκριμένο διάγραμμα περιγράφει τις κατηγορίες μιας ποιοτικής μεταβλητής με ράβδους (Bars). Το ύψος κάθε ράβδου συνήθως είναι ανάλογο του πραγματικού αριθμού ή του ποσοστού που αντιστοιχεί στην κάθε κατηγορία.

Εικόνα 3.8.1.1.1: Παράθυρο Chart Bar



Εδώ πρέπει να επιλέξουμε ποιο είδος ραβδογράμματος επιθυμούμε. Τα διαθέσιμα είναι το απλό (**Simple**), το συγκεντρωτικό (**Clustered**) και το συσσωρευμένο (**Stacked**). Στη συγκεκριμένη περίπτωση, μας ενδιαφέρει το απλό. Πρέπει ακόμη να επιλέξουμε και τον τρόπο παράστασης των δεδομένων από τις ράβδους του γραφήματος (**Data in Chart Are**). Υπάρχουν οι εξής επιλογές:

- **Summaries for groups of cases:** Οι ράβδοι παριστούν περιληπτικά στοιχεία για τις κατηγορίες μίας μεταβλητής
- **Summaries of separate variables:** Οι ράβδοι παριστούν περιληπτικά στοιχεία για περισσότερες μεταβλητές. Κάθε ράβδος αντιστοιχεί σε μία μεταβλητή.
- **Values of individual cases:** Κάθε ράβδος αντιστοιχεί σε μία περίπτωση (case) μίας μεταβλητής.

### 3.8.1.2 Κυκλικό Διάγραμμα (Pie Chart)

Αυτός ο τύπος Διαγράμματος περιγράφει τις κατηγορίες μιας ποιοτικής μεταβλητής με κομμάτια μιας πίτας (κύκλου). Το κομμάτι κάθε κατηγορίας είναι ανάλογο του αριθμού των αντικειμένων, που ανήκουν σε κάθε κατηγορία. Εάν θέλει κάποιος να τονίσει μια κατηγορία μπορεί να αποκόψει κάποιο κομμάτι.

### 3.8.1.3 Ιστόγραμμα (Histogram)

Το ιστόγραμμα απεικονίζει την κατανομή μιας ποσοτικής μεταβλητής με τη βοήθεια ράβδων. Κάθε ράβδος αντιστοιχεί σε ένα διάστημα τιμών και το ύψος της είναι

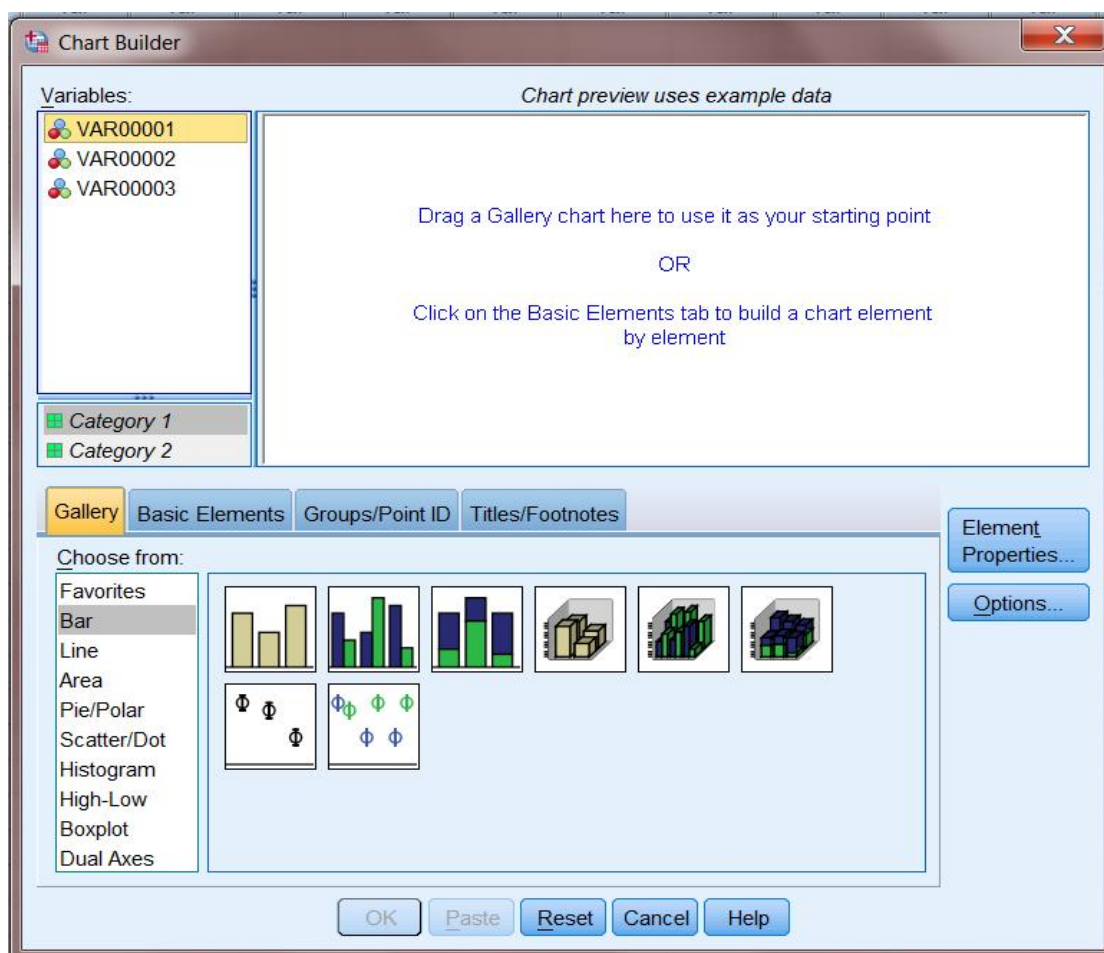


ανάλογο των αντικειμένων, που ανήκουν σε αυτό το διάστημα. Συχνά απεικονίζει κάποιος τη γραμμή της κανονικής κατανομής για σύγκριση.

### 3.8.1.4 Θηκόγραμμα ή Διάγραμμα Πλαισίου και Απολήξεων (Box Plot)

Αυτή η κατηγορία διαγραμμάτων περιγράφουν περιληπτικά την κατανομή των ποσοτικών μεταβλητών. Κάθε πλαίσιο κουτί απεικονίζει το 1ο Τεταρτημόριο, την διάμεσο και το τρίτο τεταρτημόριο. Οι απολήξεις υποδεικνύουν τα όρια των ακραίων τιμών. Οι τιμές των φραγμάτων των απολήξεων θεωρούνται ακραίες και υποδεικνύονται στο γράφημα με ξεχωριστά σημεία. Συμμετρικά Box Plots πλησιάζουν την κανονική κατανομή.

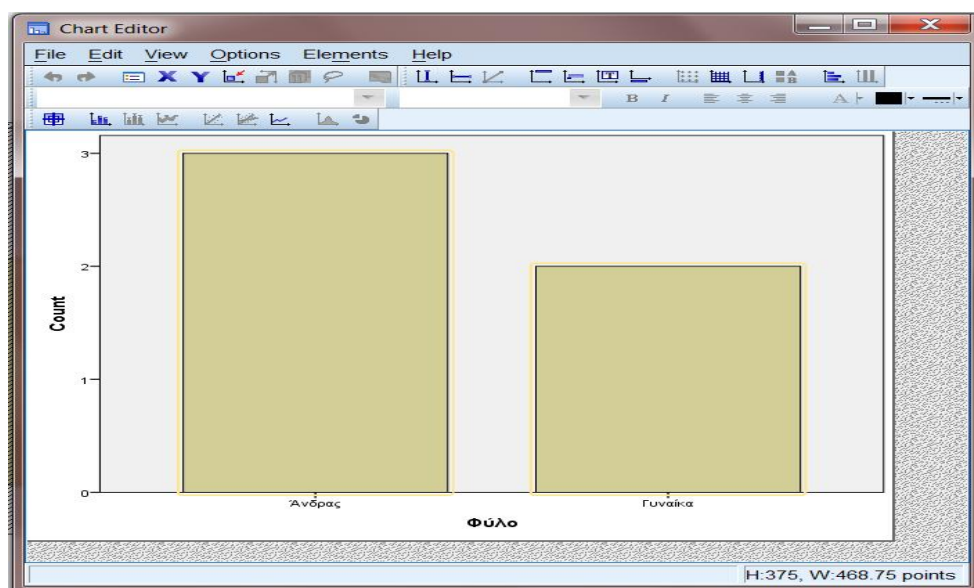
Εικόνα 3.8.1.4.1: Παράθυρο επιλογής γραφικής παράστασης



Είναι σημαντικό να αναφέρουμε ότι επιλέγοντας το παραγόμενο γράφημα, μπορούμε να το επεξεργαστούμε και να το μορφοποιήσουμε.

Το SPSS ως εργαλείο για την επίλυση προβλημάτων της Στατιστικής

Εικόνα 3.8.1.4.2: Παράθυρο επιλογής γραφικής παράστασης



## Κεφάλαιο 4ο: Πεδία εφαρμογής Spss

### 4.1 Βασικά στοιχεία ανάλυσης με το Spss<sup>13</sup>

Το Spss είναι, όπως προαναφέρθηκε, ένα στατιστικό πακέτο το οποίο μας βοηθά στην εκτέλεση πολύπλοκων στατιστικών αναλύσεων. Έτσι, τα βασικά στοιχεία τα οποία αναλύονται με τη βοήθεια του λογισμικού αυτού είναι:

- Υπολογισμός συχνοτήτων (Frequencies): Η ανάλυση αυτή παράγει πίνακες συχνοτήτων.
- Υπολογισμός μέτρων περιγραφικής στατιστικής (Descriptives): Αυτή η ανάλυση παρουσιάζει μέτρα περιγραφικής στατιστικής όπως (maximum, minimum, mean, standard deviation κ.ά.)
- Συγκρίσεις μέσων όρων (t-tests): Αυτές οι αναλύσεις μας επιτρέπουν να συγκρίνουμε δύο μέσους όρους και να γενικεύσουμε προς τον πληθυσμό από τον οποίο πήραμε τα δείγματα.
- Υπολογισμός Συσχετίσεων (correlation): Αυτή η ανάλυση μας επιτρέπει να διαγνώσουμε τον βαθμό της σχέσης που υπάρχει ανάμεσα σε δύο ή περισσότερες μεταβλητές.
- Υπολογισμός Παλινδρόμησης (Linear regression analysis): Η γραμμική παλινδρόμηση κάνει εκτίμηση των συντελεστών της γραμμικής εξίσωσης.

Έτσι, λοιπόν, το SPSS είναι ένα στατιστικό πακέτο ανάλυσης δεδομένων, το οποίο προσφέρει στο χρήστη δυνατότητες για δημιουργία αναφορών, ανάλυση και μοντελοποίηση δεδομένων καθώς και για τη γραφική αναπαράστασή τους. Διαθέτει πολλές στατιστικές συναρτήσεις για ανάλυση δεδομένων μέσα από ένα εύχρηστο περιβάλλον.

Με τη βοήθεια του SPSS όλα τα στάδια της αναλυτικής διαδικασίας ολοκληρώνονται κάτω από ένα ενοποιημένο περιβάλλον εργασίας, καλύπτοντας την ανάλυση κομμάτι-κομμάτι και από το ένα άκρο έως το άλλο.

---

<sup>13</sup> Σπανούδης Γ., «Εισαγωγή στο SPSS», Τμήμα Ψυχολογίας, ΚΕΔΙΜΑ, 2013

Διάφοροι τομείς, όπως ο χρηματοοικονομικός τομέας, ο τομέας των τηλεπικοινωνιών, ο τομέας της υγείας, της εκπαίδευσης και της έρευνας, ο ευρύτερος δημόσιος τομέας και ο τομέας της βιομηχανίας αποτελούν τα βασικότερα πεδία εφαρμογών της τεχνολογίας του SPSS.<sup>14</sup>

Το SPSS (Statistical Package for Social Sciences) είναι ένα στατιστικό πρόγραμμα με ευρύτατη χρήση σε όλους τους ερευνητικούς χώρους και ιδιαίτερα στο χώρο των κοινωνικών επιστημών. Δίνει λύσεις και απαντήσεις σε θέματα που απαιτούν χρήση της Στατιστικής Επιστήμης, αλλά δεν διδάσκει στο χρήστη Στατιστική. Αντίθετα, αυτός που γνωρίζει Στατιστική μπορεί ευκολότερα να κατανοήσει το πρόγραμμα, τη λειτουργία του και να ερμηνεύσει τα αποτελέσματα που θα του δώσει αυτό.

Η στατιστική ανάλυση στο σύνολό της περιλαμβάνει:

- Στατιστική Ανάλυση Δεδομένων Ερωτηματολογίων. Το κομμάτι αυτό περιλαμβάνει την παρουσίαση των αποτελεσμάτων (π.χ. μέσες τιμές δείγματος), με πίνακες ή / και με γραφήματα, τα οποία προέκυψαν από την κατάλληλη επεξεργασία των δεδομένων.
- Έλεγχοι Υποθέσεων. Περιλαμβάνονται παραμετρικοί και μη έλεγχοι υποθέσεων στόχων (όπως t-test, Mann-Whitney), οι οποίοι αποσκοπούν στην εύρεση στατιστικά σημαντικών συσχετίσεων και διαφορών.
- Μοντελοποίηση. Κατά τη μοντελοποίηση εφαρμόζονται κατάλληλα μοντέλα, τα οποία προκύπτουν από τα αποτελέσματα κατά το στάδιο των ελέγχων υποθέσεων, με στόχο να εξεταστεί ταυτόχρονα η συμπεριφορά πολλών παραγόντων. Παραδείγματα αποτελούν η εκθετική και η μη γραμμική παλινδρόμηση.<sup>15</sup>

---

<sup>14</sup> <http://pacific.jour.auth.gr/spss/>

<sup>15</sup> <http://www.pioneeringedu.com/statistics/>

## Κεφάλαιο 5ο: Εφαρμογή στο Spss

### 5.1 Προετοιμασία

Η ακόλουθη εφαρμογή στηρίζεται σε επεξεργασία στατιστικών στοιχείων που αντλήθηκαν από το Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, σε μάθημα σχετικό με τις μεταφορές. Το συγκεκριμένο παράδειγμα αφορά τις πωλήσεις αυτοκινήτων τριών μεγεθών (small, medium, large) στις χώρες της Ευρωπαϊκής Ένωσης για κάθε μήνα πέντε τελευταίων ετών από το 2004 έως το 2008 και ερευνά και εξετάζει τους παράγοντες που επέδρασαν στην αγορά συγκεκριμένου τύπου αυτοκινήτων. Η έρευνα διεξήχθη καθ' όλη τη διάρκεια των 5 αυτών ετών, ενώ οι πιθανοί παράγοντες από τους οποίους επηρεάζονται οι πωλήσεις των αυτοκινήτων είναι η ανεργία («UN»), και η τιμή του καυσίμου («FUEL»).

Αφού συμπληρώθηκαν τα ερωτηματολόγια, έγινε η κατάλληλη προετοιμασία για την εισαγωγή των στοιχείων τους σε έναν υπολογιστή και την επεξεργασία τους με τη βοήθεια του προγράμματος SPSS.

Πιο συγκεκριμένα, για τη στατιστική επεξεργασία ερωτηματολογίων δημιουργήθηκε ένας σαφής πίνακας κωδικοποίησης. Ο πίνακας αυτός αντιστοιχίζει κάθε ερώτηση του ερωτηματολογίου σε μια μεταβλητή. Για παράδειγμα, η ερώτηση «Τιμή Καυσίμου» αντιστοιχίζεται στη μεταβλητή «Fuel».

Οι μεταβλητές λαμβάνουν διάφορες τιμές. Για παράδειγμα, η μεταβλητή «Μήνας» έχει δώδεκα πιθανές τιμές: μία για κάθε μήνα. Στον πίνακα κωδικοποίησης αντιστοιχίζουμε σε κάθε τιμή μιας μεταβλητής έναν αριθμό, για παράδειγμα, στην τιμή «Ιανουάριος» της μεταβλητής «Μήνας» μπορούμε να αντιστοιχήσουμε τον αριθμό «1» και στην τιμή «Φεβρουάριος» τον αριθμό «2», κ.λ.π.

Σε περίπτωση που είχαμε μεταβλητές που έπαιρναν διάφορες τιμές, όπως π.χ. επαγγελματική κατάσταση, πιθανότατα θα είχαμε μεγάλο εύρος τιμών, οπότε και η ομαδοποίηση θα ήταν αναγκαία.

Με βάση τα παραπάνω, λοιπόν, όσα προαναφέρθηκαν, ύστερα από την εισαγωγή των δεδομένων μέσω αρχείου .xls (Microsoft Excel), πρωταρχικό μας βήμα είναι η δήλωση των μεταβλητών μας και ο καθορισμός του τύπου τους από το μενού

Variable View. Δηλώνουμε τα ονόματα των μεταβλητών, τον τύπο τους, το πλήθος των δεκαδικών ψηφίων. Η συμπλήρωση των στηλών Label και Value δεν είναι αναγκαίες, καθώς έχουν καθαρά επεξηγηματικό χαρακτήρα και βοηθούν τον χρήστη στην κατανόηση των μεταβλητών.

Έτσι, βλέπουμε ότι όλες οι μεταβλητές έχουν αριθμητικό χαρακτήρα (Εικόνα 5.1.1).

Εικόνα 5.1.1: Μενού μεταβλητών Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	YEAR	Numeric	4	0		None	None	11	Right	Scale	Input
2	MONTH	Numeric	11	0		None	None	11	Right	Scale	Input
3	SM	Numeric	11	0		None	None	11	Right	Scale	Input
4	SM_PR	Numeric	11	0		None	None	11	Right	Scale	Input
5	M	Numeric	11	0		None	None	11	Right	Scale	Input
6	M_PR	Numeric	11	0		None	None	11	Right	Scale	Input
7	L	Numeric	11	0		None	None	11	Right	Scale	Input
8	L_PR	Numeric	11	0		None	None	11	Right	Scale	Input
9	UN	Numeric	11	1		None	None	11	Right	Scale	Input
10	IMM	Numeric	11	4		None	None	11	Right	Scale	Input
11	PIFI	Numeric	11	3		None	None	11	Right	Scale	Input
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											
28											
29											
30											
31											

Πιο συγκεκριμένα η μεταβλητή «YEAR» αντιπροσωπεύει την εξεταζόμενη χρονιά, παίρνει τιμές από 1 έως 5 καθώς τα εξεταζόμενα έτη είναι πέντε και είναι αριθμητικού χαρακτήρα. Για παράδειγμα, ο αριθμός ένα αντιστοιχεί στο έτος 2004, ο αριθμός δύο στο έτος 2005 κ.ο.κ.

Στη συνέχεια, η μεταβλητή «MONTH» αντιπροσωπεύει κάθε μήνα για καθεμιά χρονιά και παίρνει τιμές από 1 έως 12, όσοι δηλαδή και οι ημερολογιακοί μήνες και είναι αριθμητικού χαρακτήρα.

Η μεταβλητή «SM» αντιπροσωπεύει τις πωλήσεις των μικρών αυτοκινήτων, είναι αριθμητικού χαρακτήρα και στο παράδειγμά μας έχουμε τιμές για κάθε μήνα του

κάθε εξεταζόμενο έτος. Οι τιμές αυτές δε χρειάστηκαν κωδικοποίηση, καθώς αποτελούν τις πραγματικές τιμές πωλήσεων.

Η μεταβλητή «SM\_PR» αντιπροσωπεύει την μέση τιμή πώλησης μικρών αυτοκινήτων για κάθε μήνα του κάθε εξεταζόμενου έτους και είναι αριθμητικού χαρακτήρα. Οι τιμές αυτές δε χρειάστηκαν κωδικοποίηση, καθώς αποτελούν τις πραγματικές τιμές.

Η μεταβλητή «M» αντιπροσωπεύει τις πωλήσεις των μεσαίων αυτοκινήτων, είναι αριθμητικού χαρακτήρα και στο παράδειγμά μας έχουμε τιμές για κάθε μήνα του κάθε εξεταζόμενου έτους. Οι τιμές αυτές δε χρειάστηκαν κωδικοποίηση, καθώς αποτελούν τις πραγματικές τιμές πωλήσεων.

Η μεταβλητή «M\_PR» αντιπροσωπεύει την μέση τιμή πώλησης μεσαίου μεγέθους αυτοκινήτων για κάθε μήνα του κάθε εξεταζόμενου έτους και είναι αριθμητικού χαρακτήρα. Οι τιμές αυτές δε χρειάστηκαν κωδικοποίηση, καθώς αποτελούν τις πραγματικές τιμές.

Η μεταβλητή «L» αντιπροσωπεύει τις πωλήσεις των μεγάλων αυτοκινήτων, είναι αριθμητικού χαρακτήρα και στο παράδειγμά μας έχουμε τιμές για κάθε μήνα του κάθε εξεταζόμενου έτους. Οι τιμές αυτές δε χρειάστηκαν κωδικοποίηση, καθώς αποτελούν τις πραγματικές τιμές πωλήσεων.

Η μεταβλητή «L\_PR» αντιπροσωπεύει την μέση τιμή πώλησης μεγάλου μεγέθους αυτοκινήτων για κάθε μήνα του κάθε εξεταζόμενου έτους και είναι αριθμητικού χαρακτήρα. Οι τιμές αυτές δε χρειάστηκαν κωδικοποίηση, καθώς αποτελούν τις πραγματικές τιμές.

Η μεταβλητή «UN» αντιπροσωπεύει το μέσο ποσοστό ανεργίας για κάθε μήνα του κάθε εξεταζόμενου έτους και είναι αριθμητικού χαρακτήρα.

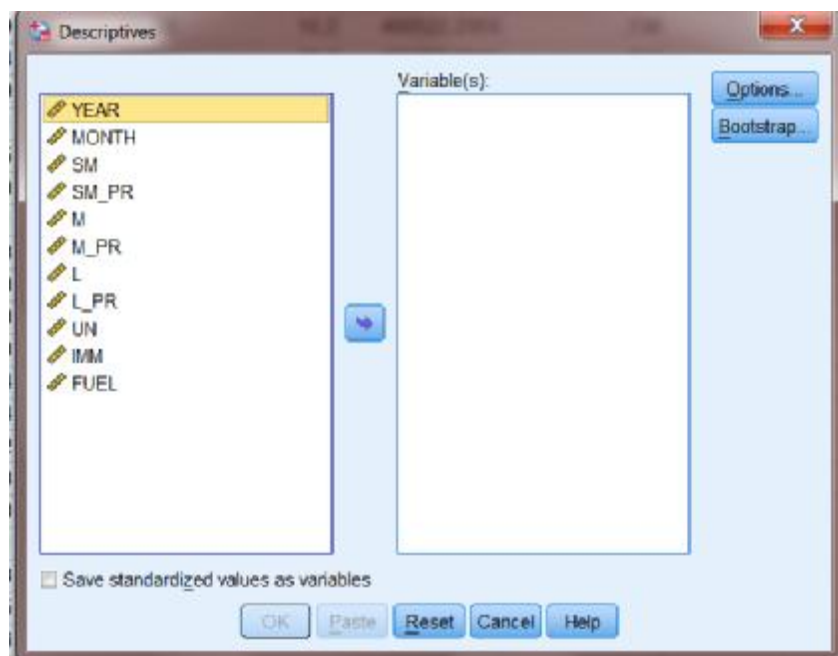
Η μεταβλητή «FUEL» αντιπροσωπεύει την μέση τιμή της τιμής της βενζίνης ανά λίτρο για κάθε μήνα του κάθε εξεταζόμενου έτους και είναι αριθμητικού χαρακτήρα.

## 5.2 Στατιστική Ανάλυση παραδείγματος<sup>16</sup>

Αρχικά, θα ασχοληθούμε με βασική στατιστική ανάλυση των στοιχείων του παραδείγματός μας. Η έμφαση δίνεται στη χρήση του στατιστικού πακέτου SPSS για την απόκτηση των αποτελεσμάτων και λιγότερο στην ερμηνεία των αποτελεσμάτων και των στατιστικών εννοιών.

Για να προχωρήσουμε σε βασική στατιστική ανάλυση, επιλέγουμε την εντολή Analyze – Descriptive Statistics -Descriptives, από το μενού εντολών, οπότε και ανοίγει το ακόλουθο παράθυρο.

Εικόνα 5.2.1: Επιλογή μεταβλητών στην εντολή Descriptives



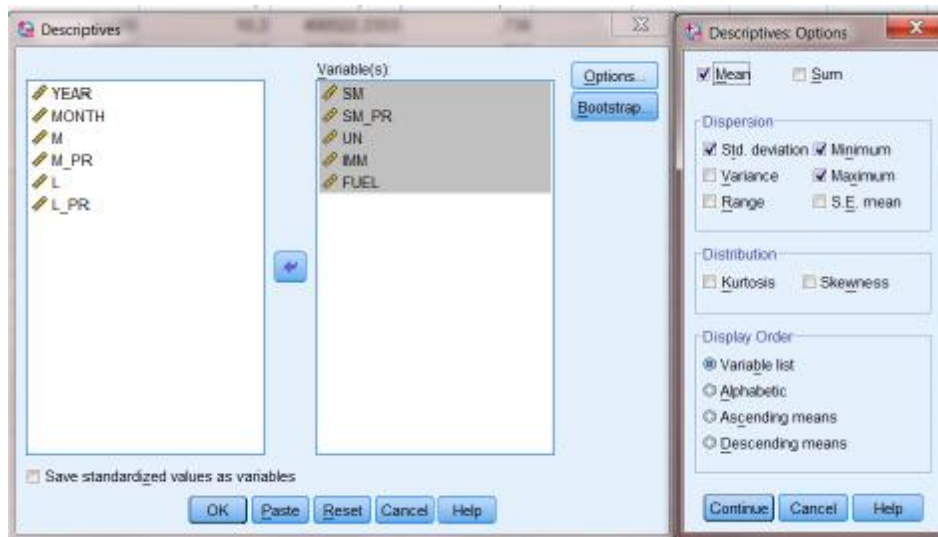
Στη συνέχεια μεταφέρουμε στο δεξί ορθογώνιο πλαίσιο (Variables), τις μεταβλητές τις οποίες θέλουμε να αναλύσουμε, επιλέγοντας καθεμιά μεταβλητή και πατώντας το βέλος. Στο παράθυρο που ανοίγει αν πατήσουμε την επιλογή Options μπορούμε να ζητήσουμε τα περιγραφικά μέτρα που θέλουμε να αναλύσουμε. Έτσι, στο παράδειγμά μας, επιλέγουμε να αναλύσουμε πως επηρεάζονται οι πωλήσεις των μικρών

<sup>16</sup> Μητρόπουλος Ι., Σημειώσεις για το εργαστήριο του μαθήματος «Εισαγωγή στη Στατιστική των επιχειρήσεων», Τεχνολογικό Εκπαιδευτικό Ίδρυμα Δυτικής Ελλάδας, Τμήμα Διοίκησης Επιχειρήσεων, Πάτρα, 2015



αυτοκινήτων από την τιμή τους, την ανεργία, τον αριθμό των μεταναστών και την τιμή του καυσίμου.

Εικόνα 5.2.2: Επιλογή περιγραφικών στοιχείων στην εντολή Descriptives



Στον πίνακα των αποτελεσμάτων που παίρνουμε με το που πατήσουμε OK, εμφανίζονται οι μεταβλητές των που χρησιμοποιήθηκαν στην πρώτη στήλη, το πλήθος των στοιχείων από τα οποία απαρτίζεται η κάθε μεταβλητή, οι ελάχιστες και μέγιστες τιμές για κάθε μεταβλητή, ο μέσος όρος και η τυπική απόκλιση. Παρατηρούμε μεγάλη διακύμανση στον αριθμό των πωλήσεων ανά μήνα (από 2.138 πωλήσεις τον μήνα έως 12.499 τον μήνα), στον αριθμό των μεταναστών καθώς και στην τιμή της βενζίνης (από 0.681 έως 1.01). Η τιμή των οχημάτων (SM\_PR) δεν διαφοροποιείται σημαντικά.

Εικόνα 5.2.3: Αποτελέσματα εντολής Descriptives

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
SM	60	2138	12499	7257,42	2271,247
SM_PR	60	11195	12512	11991,38	330,316
UN	60	9,3	11,4	10,240	,6214
FUEL	60		1,010	,78107	,074455
Valid N (listwise)	60				

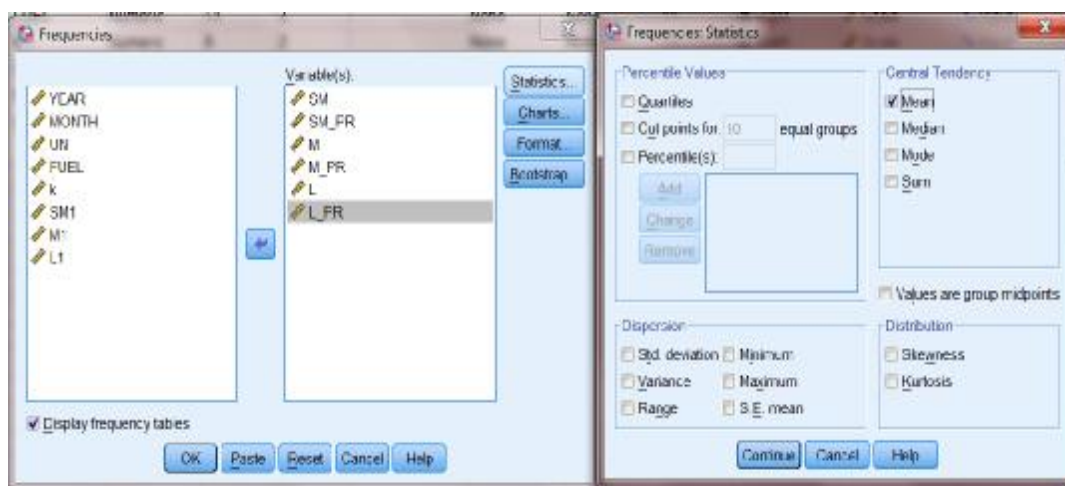
Αν επιλέξουμε ως μεταβλητές τις πωλήσεις των μικρών αυτοκινήτων (SM), των μεσαίων (M) και των μεγάλων αυτοκινήτων (L) παίρνουμε τον ακόλουθο συγκεντρωτικό πίνακα, όπου βλέπουμε το μέγιστο και το ελάχιστο των πωλήσεων κάθε κατηγορίας καθώς και την διαφορά του. Έτσι, βλέπουμε ότι τη μεγαλύτερη διακύμανση στις πωλήσεις ανά μήνα έχουν τα μικρά αυτοκίνητα, ενώ στα μεγάλα αυτοκίνητα η διακύμανση είναι πολύ μικρότερη, όπως βέβαια και σοι συνολικές πωλήσεις.

Εικόνα 5.2.4: Στατιστικά στοιχεία κατηγοριών οχήματος

		SM	M	L
N	Valid	60	60	60
	Missing	0	0	0
Range		10361	8948	3536
Minimum		2138	2132	1658
Maximum		12499	11080	5194

Ακόμη, επαναλαμβάνουμε την ίδια διαδικασία και τώρα ως μεταβλητές επιλέγω τις ακόλουθες όπως φαίνεται στην εικόνα.

Εικόνα 5.2.5: Προετοιμασία εντολής Frequencies



Επομένως, προκύπτει ο ακόλουθος πίνακας όπου φαίνεται για κάθε παράμετρο η μέση τιμή. Διαπιστώνουμε ότι οι περισσότερες πωλήσεις είναι αυτές των μικρών αυτοκινήτων ενώ με μικρή διαφορά ακολουθούν οι πωλήσεις των μεσαίων

αυτοκινήτων. Τα μεγάλα αυτοκίνητα φαίνεται ότι προσελκύουν μικρότερο αγοραστικό κοινό. Επίσης, είναι ευδιάκριτη η διακύμανση των τιμών, η οποία διαφορά κυμαίνεται κοντά στις 5000€ από τη μια κατηγορία στην άλλη.

Εικόνα 5.2.6: Αποτελέσματα εντολής Frequencies

		Statistics					
		SM	M	L	SM_PR	M_PR	L_PR
N	Valid	60	60	60	60	60	60
	Missing	0	0	0	0	0	0
Mean		7257,42	6539,17	3205,37	11991,38	17567,38	24973,80

Στη συνέχεια, ομαδοποιούμε για κάθε μια κατηγορία ξεχωριστά τις πωλήσεις κατά παρόμοιες κατηγορίες. Ομαδοποίηση είναι μια διαδικασία με την οποία ομαδοποιούμε (χωρίζουμε σε ομάδες) τα δεδομένα μας. Τη διαδικασία αυτή την χρησιμοποιούμε για να έχουμε καλύτερη παρουσίαση των αποτελεσμάτων μιας στατιστικής ανάλυσης και των διαγραμμάτων, στη περίπτωση που έχουμε μεγάλο μέγεθος δείγματος (π.χ. 100 παρατηρήσεις) ή στη περίπτωση που έχουμε μεγάλη διαφοροποίηση (ποικιλία) στα δεδομένα μας (αυτό συμβαίνει όταν έχουμε ποσοτικές συνεχείς μεταβλητές).

Η διαδικασία που ακολουθήσαμε για να προβούμε στην ομαδοποίηση των δεδομένων μας είναι η ακόλουθη:

- 1) Προσδιορίζουμε τον αριθμό των τάξεων ( $k$ ) με τον κανόνα του Sturges:  $k = 1 + 3,322 \log n$  όπου  $k$  = αριθμός τάξεων και  $n$  = ο αριθμός παρατηρήσεων όπου στην περίπτωσή μας  $n=60$  για όλες τις κατηγορίες αυτοκινήτων

Ο υπολογισμός του  $k$  μπορεί να γίνει και μέσω του SPSS με συγκεκριμένη ακολουθία εντολών, ωστόσο για συντομία χρόνου επιλέξαμε να γίνει χειρωνακτικά.

- 2) Στη συνέχεια υπολογίζουμε το πλάτος των τάξεων από τον τύπο  $\delta = R / k$ , όπου  $\delta$  = πλάτος των τάξεων,  $R$  = εύρος μεταβολής (δηλαδή  $R = \text{Max} - \text{Min}$ ), και  $k$  = αριθμός τάξεων

Για τα μικρά αυτοκίνητα  $\max=12499$  ,  $\min= 2138$  , και  $k=7$ .

Για τα μεσαία αυτοκίνητα είναι  $\max=11080$  ,  $\min= 2132$  , και  $k=7$ .

Για τα μεγάλα αυτοκίνητα είναι  $\max=12499$  ,  $\min= 5194$ , και  $k=7$ .

Αντίστοιχα υπολογίζεται το πλάτος για καθεμία από τις κατηγορίες αυτοκινήτων με βάση τα δεδομένα που προαναφέρθηκαν.

Διαφορετικά υπολογίζεται αν ακολουθήσουμε την εξής διαδικασία: ANALYSE - DESCRIPTIVE STATISTICS – FREQUENCIES

Επιλέγουμε τις μεταβλητές SM, M, L και προσέχουμε να μην έχει τσεκαριστεί η ένδειξη DISPLAY FREQUENCIES TABLES για να μην δημιουργήσουμε πίνακα συχνοτήτων. Επιλέγουμε την ένδειξη και εμφανίζεται το παράθυρο Frequencies: Statistics από το οποίο επιλέγουμε τις ενδείξεις Range (εύρος μεταβολής), Minimum (η ελάχιστη τιμή των δεδομένων) και Maximum (η μέγιστη τιμή των δεδομένων). Στη συνέχεια πατάμε CONTINUE και O.K.. Έτσι προκύπτει ο πίνακας 5.2.6 που προηγήθηκε όπου έχουμε έτοιμο το εύρο μεταβολής R.

Οι τάξεις που δημιουργεί το SPSS είναι της μορφής ( ], δηλαδή κλειστές στο πάνω όριο. Προσέχουμε ελάχιστη και η μέγιστη τιμή των δεδομένων να περιέχονται στις τάξεις που δημιουργήσαμε.

Πιο συγκεκριμένα, για τα μικρά αυτοκίνητα έχουμε:

Κατηγορία 1: από 2138 έως 3619 οχήματα

Κατηγορία 2: από 3619 έως 5100 οχήματα

Κατηγορία 3: από 5100 έως 6581 οχήματα

Κατηγορία 4: από 6581 έως 8062 οχήματα

Κατηγορία 5: από 8062 έως 9543 οχήματα

Κατηγορία 6: από 9543 έως 11024 οχήματα

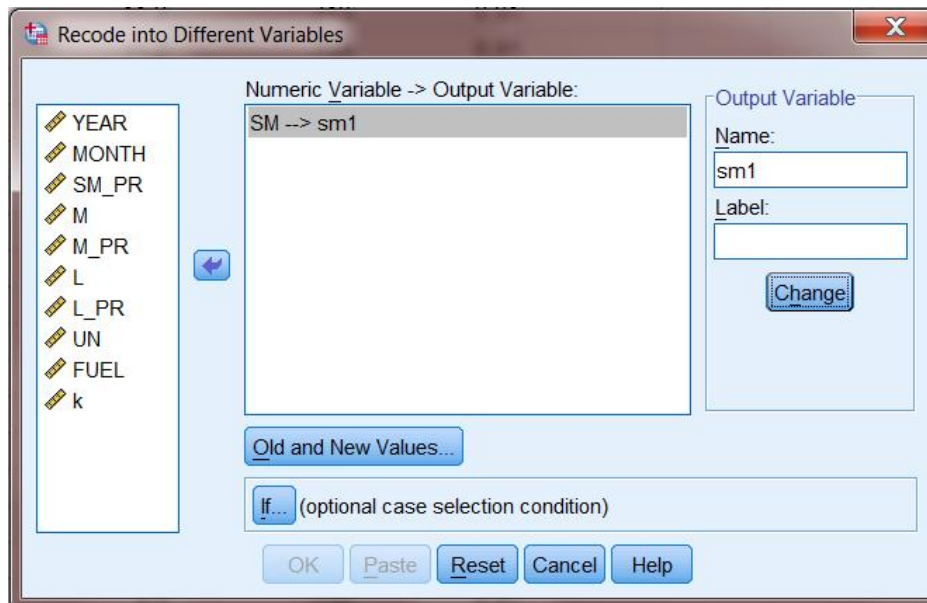
Κατηγορία 7: από 11024 έως 12505 οχήματα,

- 3) Για να πραγματοποιηθεί η ομαδοποίηση στο αρχείο μας ακολουθούμε την εξής διαδικασία: Πατάμε στο μενού TRANSFORM - RECODE - INTO DIFFERENT VARIABLES. Αν επιλέξουμε INTO SAME VARIABLES η

ομαδοποίηση θα γίνει στην ίδια στήλη με αποτέλεσμα να χαθούν τα αρχικά δεδομένα (και σε περίπτωση που έχει γίνει κάποιο λάθος στην ομαδοποίηση δεν θα έχουμε διαθέσιμα τα αρχικά δεδομένα για να το διορθώσουμε). Επιλέγουμε INTO DIFFERENT VARIABLES επειδή θέλουμε η ομαδοποίηση να γίνει σε διαφορετική στήλη.

Επιλέγουμε από το παράθυρο αριστερά τη μεταβλητή sm και με πάτημα στο μαύρο βέλος αυτή μεταφέρεται στο παράθυρο δεξιά. Στη θέση NAME γράφουμε το όνομα που θα δώσουμε στη νέα στήλη δηλαδή sm1 και πατώντας στην ένδειξη το όνομα που δώσαμε, πηγαίνει δίπλα στο όνομα της προηγούμενης μεταβλητής στη θέση του ερωτηματικού.

Εικόνα 5.2.7: Κωδικοποίηση μεταβλητών ομαδοποίησης



Στη συνέχεια πατάμε στην επιλογή **Old and New Values**. Πατάμε στην επιλογή (OLD VALUE) RANGE και στα δύο παράθυρα γράφουμε το κατώτερο και το ανώτερο όριο της πρώτης τάξης, δηλ. 2138 through 3619. Πατάμε δεξιά στην ένδειξη (NEW VALUE) VALUE και δίνουμε τον κωδικό ο οποίος αντικαθιστά την πρώτη τάξη. Δηλαδή, για την ομάδα 2138– 3619 ο κωδικός είναι 1 και συνεχίζουμε μέχρι να τελειώσουν όλες οι τάξεις. Οπότε πλέον στο Data Editor δημιουργήθηκε μια καινούρια στήλη η sm1.

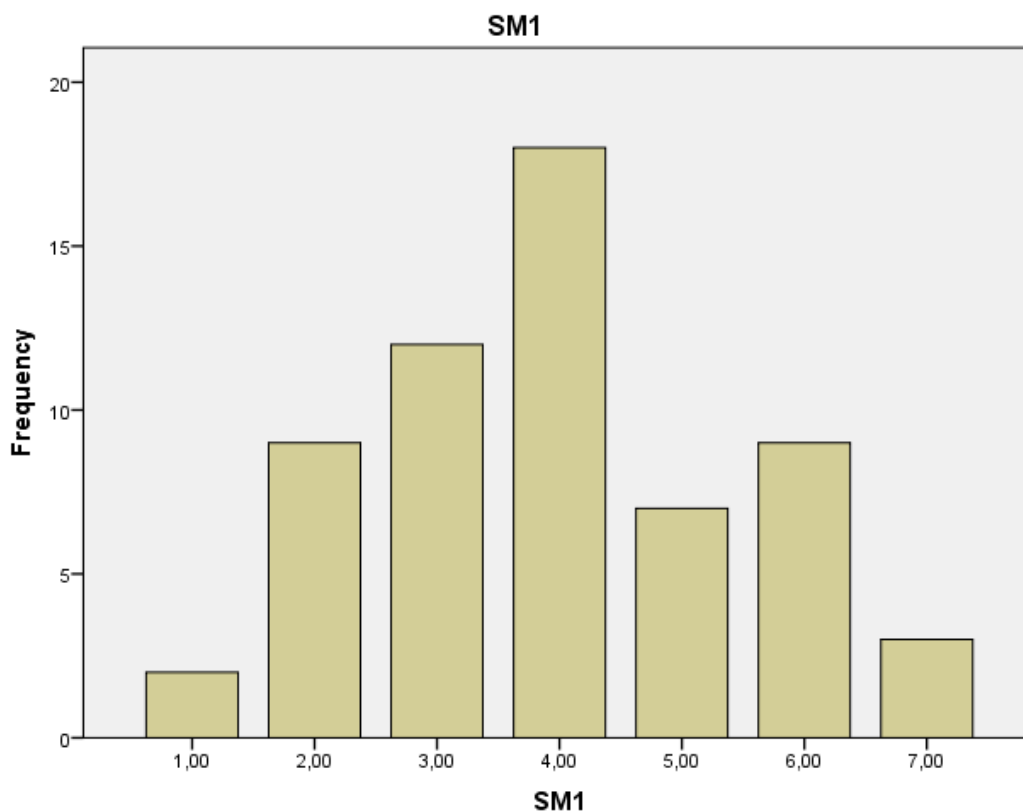
Στη συνέχεια, μπορούμε κανονικά να προχωρήσουμε σε πίνακες κατανομής συχνοτήτων και ιστογράμματα με την εξής ακολουθία εντολών: ANALYSE - DESCRIPTIVE STATISTICS – FREQUENCIES. Επιλέγουμε από το παράθυρο αριστερά την μεταβλητή SM1 και με πάτημα στο μαύρο βέλος αυτή μεταφέρεται στο παράθυρο δεξιά. Τσεκάρουμε την ένδειξη DISPLAY FREQUENCY TABLES και στη συνέχεια Ο.Κ. Οπότε προκύπτει ο ακόλουθος πίνακας συχνοτήτων:

Εικόνα 5.2.8: Πίνακας Συχνοτήτων ομαδοποιημένης μεταβλητής SM1

SM1					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	2	3,3	3,3	3,3
	2,00	9	15,0	15,0	18,3
	3,00	12	20,0	20,0	38,3
	4,00	18	30,0	30,0	68,3
	5,00	7	11,7	11,7	80,0
	6,00	9	15,0	15,0	95,0
	7,00	3	5,0	5,0	100,0
	Total	60	100,0	100,0	

Όπως βλέπουμε από τον πίνακα το μεγαλύτερο ποσοστό συγκεντρώνεται στην κατηγορία 4 με ποσοστό 30%, ενώ τα μικρότερα ποσοστά στις «ακραίες» κατηγορίες, κάτι το οποίο είναι περισσότερο ευδιάκριτο στο ακόλουθο ιστόγραμμα.

Εικόνα 5.2.9: Ιστόγραμμα ομαδοποιημένης μεταβλητής SM1



Όσον αφορά στις πωλήσεις των μεσαίων αυτοκινήτων έχουμε ακολουθείται ακριβώς η ίδια διαδικασία με τα μικρά οχήματα, οπότε έχουμε:

Κατηγορία 1: από 2132 έως 3411 οχήματα

Κατηγορία 2: από 3411 έως 4690 οχήματα

Κατηγορία 3: από 4690 έως 5969 οχήματα

Κατηγορία 4: από 5969 έως 7248 οχήματα

Κατηγορία 5: από 7248 έως 8527 οχήματα

Κατηγορία 6: από 8527 έως 9806 οχήματα

Κατηγορία 7: από 9806 έως 11085 οχήματα.

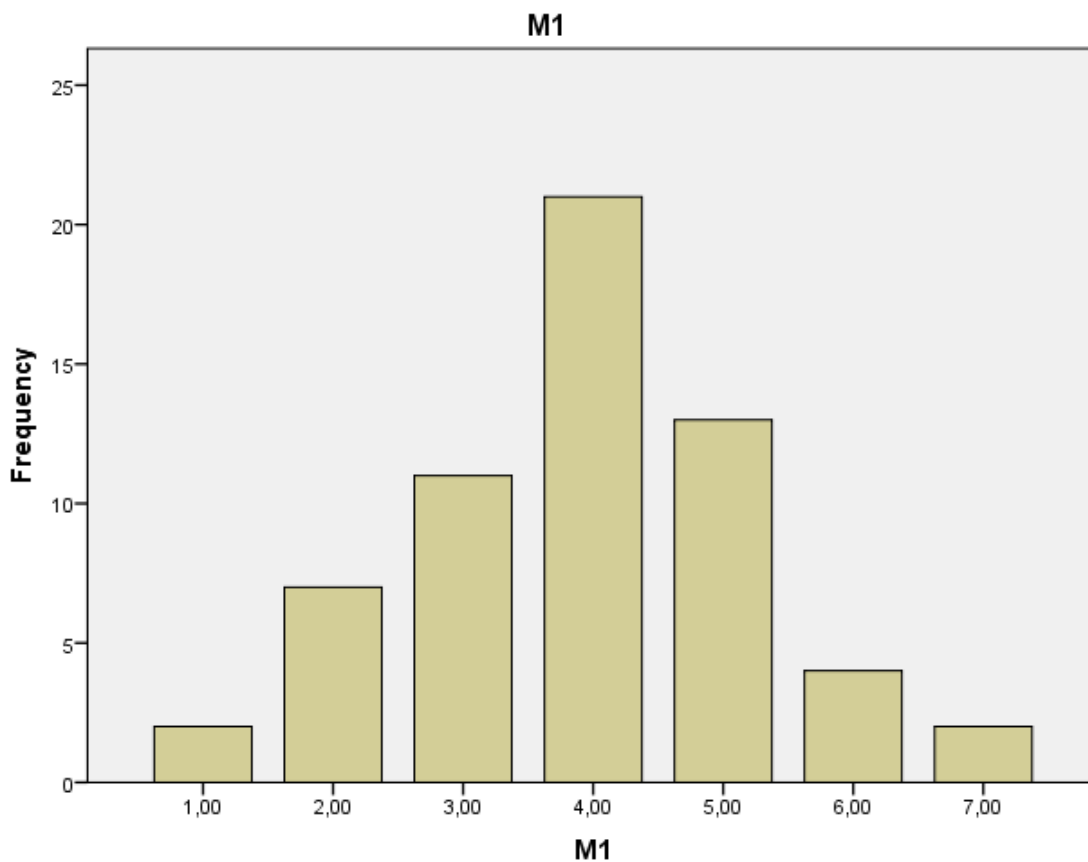
Οπότε προκύπτει ο ακόλουθος πίνακας συχνοτήτων:

Εικόνα 5.2.10: Πίνακας Συχνοτήτων ομαδοποιημένης μεταβλητής M1

		M1			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	2	3,3	3,3	3,3
	2,00	7	11,7	11,7	15,0
	3,00	11	18,3	18,3	33,3
	4,00	21	35,0	35,0	68,3
	5,00	13	21,7	21,7	90,0
	6,00	4	6,7	6,7	96,7
	7,00	2	3,3	3,3	100,0
Total		60	100,0	100,0	

Όπως βλέπουμε από τον πίνακα το μεγαλύτερο ποσοστό συγκεντρώνεται πάλι στην κατηγορία 4 με ποσοστό 35%, ενώ τα μικρότερα ποσοστά στις «ακραίες» κατηγορίες με ποσοστό 3.3% , κάτι το οποίο είναι περισσότερο ευδιάκριτο στο ακόλουθο ιστόγραμμα όπου παρατηρείται σχετική συμμετρία.

Εικόνα 5.2.11: Ιστόγραμμα ομαδοποιημένης μεταβλητής M1





Όσον αφορά στις πωλήσεις των μεγάλων αυτοκινήτων έχουμε:

Κατηγορία 1: από 1658 έως 2164 οχήματα

Κατηγορία 2: από 2164 έως 2670 οχήματα

Κατηγορία 3: από 2670 έως 3176 οχήματα

Κατηγορία 4: από 3176 έως 3682 οχήματα

Κατηγορία 5: από 3682 έως 4188 οχήματα

Κατηγορία 6: από 4188 έως 4694 οχήματα

Κατηγορία 7: από 4694 έως 5200 οχήματα,

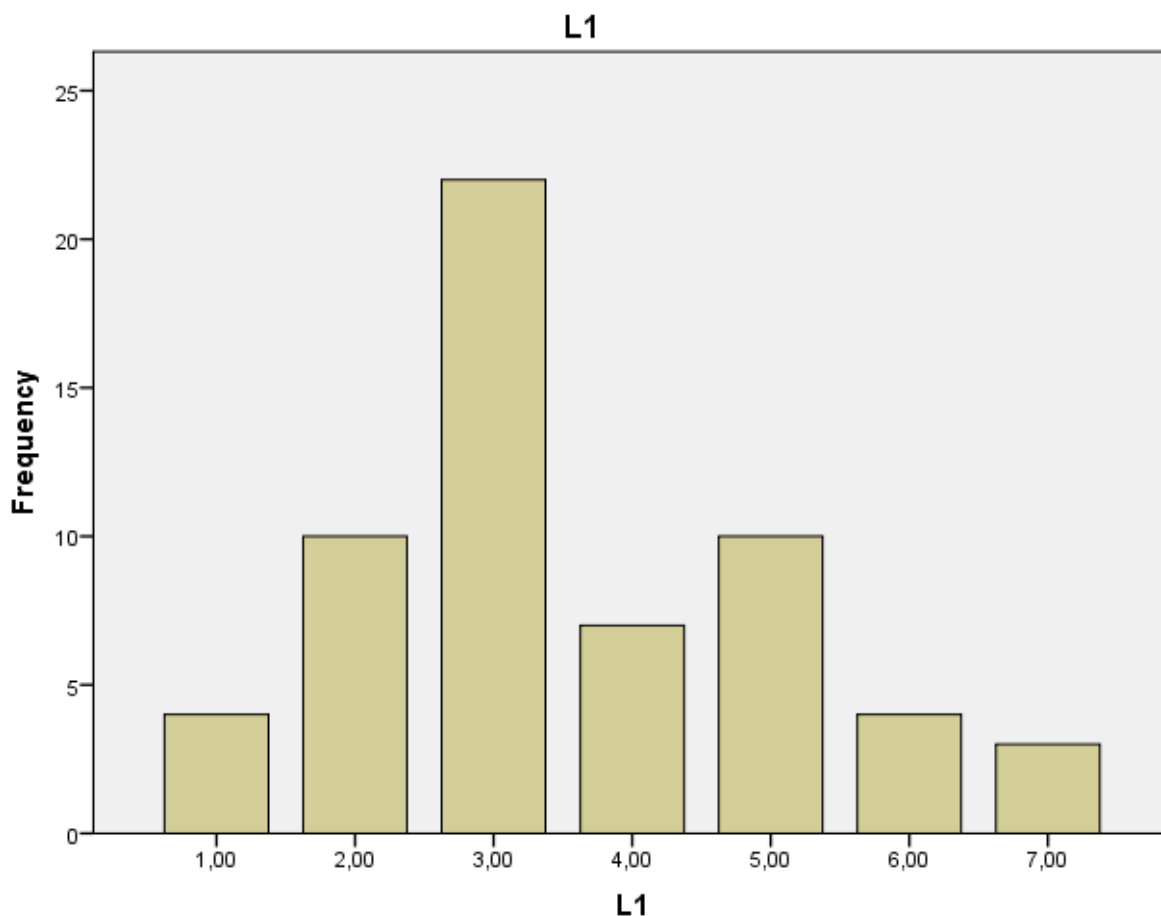
Οπότε προκύπτει ο ακόλουθος πίνακας συχνοτήτων:

Εικόνα 5.2.12: Πίνακας Συχνοτήτων ομαδοποιημένης μεταβλητής L1

		L1			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	4	6,7	6,7	6,7
	2,00	10	16,7	16,7	23,3
	3,00	22	36,7	36,7	60,0
	4,00	7	11,7	11,7	71,7
	5,00	10	16,7	16,7	88,3
	6,00	4	6,7	6,7	95,0
	7,00	3	5,0	5,0	100,0
Total		60	100,0	100,0	

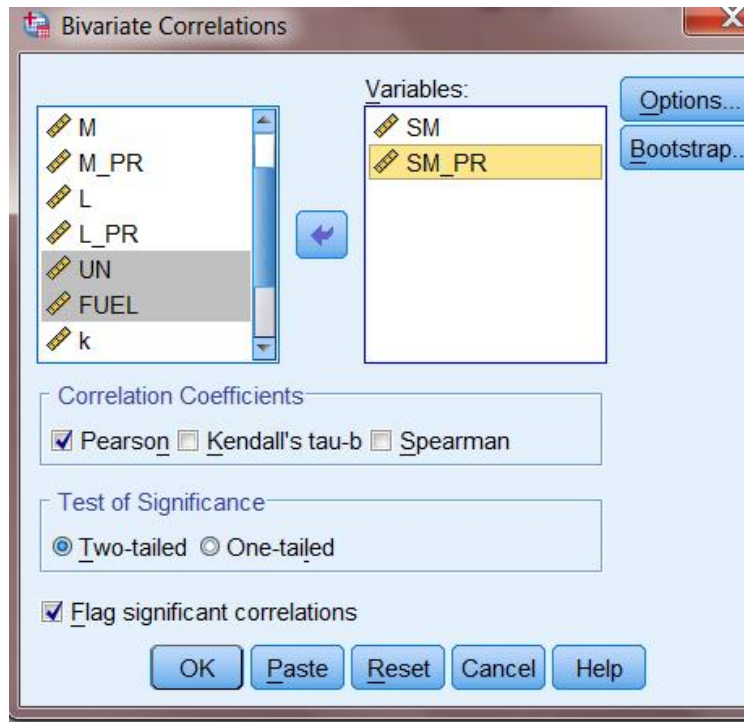
Όπως βλέπουμε από τον πίνακα το μεγαλύτερο ποσοστό συγκεντρώνεται στην κατηγορία 3 με ποσοστό 36,7%, ενώ τα μικρότερα ποσοστά στις «ακραίες» κατηγορίες με ποσοστό 5-6,7% , κάτι το οποίο είναι περισσότερο ευδιάκριτο στο ακόλουθο ιστόγραμμα. Έτσι, συμπεραίνουμε ότι οι πωλήσεις μεγάλων αυτοκινήτων είναι συγκεντρωμένες σε πιο χαμηλές τιμές συγκριτικά με τις άλλες δυο κατηγορίες.

Εικόνα 5.2.13: Ιστόγραμμα ομαδοποιημένης μεταβλητής L1



Ακολουθεί ο προσδιορισμός των συσχετίσεων των πωλήσεων καθεμιάς κατηγορίας αυτοκινήτου συγκριτικά με διάφορες παράγοντες. Αυτό γίνεται για να δούμε από ποιους παράγοντες επηρεάζονται οι πωλήσεις τους κατά την τελευταία πενταετία. Οι εξεταζόμενοι παράγοντες είναι η τιμή πώλησης του οχήματος, η ανεργία και η τιμή καυσίμου. Έτσι, ύστερα από την ακολουθία των εντολών Analyze – Correlate – Bivariate, έχουμε τον ακόλουθο πίνακα, όπου αρχικά επιλέγουμε τις μεταβλητές προς συσχέτιση. Παρουσιάζεται ενδεικτικά ο πίνακας μιας συσχέτισης μεταβλητών και με όμοιο τρόπο γίνονται και οι υπόλοιπες συσχετίσεις.

Εικόνα 5.2.14: Επιλογή μεταβλητών συσχέτισης



Έτσι, παίρνουμε ως αποτέλεσμα τον ακόλουθο πίνακα.

Εικόνα 5.2.15: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μικρών αυτοκινήτων και τιμής πώλησης

		Correlations	
		SM_PR	SM
SM_PR	Pearson Correlation	1	-,240
	Sig. (2-tailed)		,065
	N	60	60
SM	Pearson Correlation	-,240	1
	Sig. (2-tailed)	,065	
	N	60	60

Σύμφωνα με τη θεωρία που παρατίθεται σε προηγούμενο κεφάλαιο, βλέπουμε ότι η τιμή των μικρών αυτοκινήτων χαρακτηρίζεται από μέτρια συσχέτιση με την τιμή των μεσαίων αυτοκινήτων αφού είναι  $\text{Pearson} = -0.240$  και ανήκει στο διάστημα  $(0.2-0.4]$ . Επίσης, ο μειωτικός συντελεστής στο συντελεστή Pearson υποδηλώνει ότι όσο αυξάνεται η τιμή των οχημάτων, τόσο μειώνονται οι πωλήσεις, δηλαδή μιλάμε και δυο μεταβλητές αντιστρόφως ανάλογες.

Εικόνα 5.2.16: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μικρών αυτοκινήτων και δείκτη ανεργίας

		Correlations	
		SM	UN
SM	Pearson Correlation	1	,337**
	Sig. (2-tailed)		,008
	N	60	60
UN	Pearson Correlation	,337**	1
	Sig. (2-tailed)	,008	
	N	60	60

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

Παρατηρώ ότι η ανεργία χαρακτηρίζεται επίσης από μέτρια συσχέτιση με τις πωλήσεις των μικρών αυτοκινήτων αφού είναι  $\text{Pearson} = 0.337$  και ανήκει στο διάστημα  $(0.2-0.4]$ , αλλά περισσότερη συγκριτικά με την τιμή πώλησης. Ακόμη η συσχέτιση κρίνεται στατιστικά σημαντική για επίπεδο σημαντικότητας  $\alpha=0.01$ .

Εικόνα 5.2.17: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μικρών αυτοκινήτων και τιμής καυσίμου

		Correlations	
		SM	FUEL
SM	Pearson Correlation	1	,226
	Sig. (2-tailed)		,082
	N	60	60
FUEL	Pearson Correlation	,226	1
	Sig. (2-tailed)	,082	
	N	60	60

Παρατηρώ ότι η τιμή του καυσίμου χαρακτηρίζεται επίσης από μέτρια συσχέτιση με τις πωλήσεις των αυτοκινήτων αφού είναι  $\text{Pearson} = 0.226$  και ανήκει στο διάστημα  $(0.2-0.4]$ . Το θετικό πρόσημο υποδηλώνει ότι η αύξηση της τιμής του καυσίμου, αυξάνει τις πωλήσεις των μικρών αυτοκινήτων που είναι και περισσότερο οικονομικά στην κατανάλωση ενέργειας.

Στη συνέχεια, εξετάζουμε πως εξαρτώνται οι πωλήσεις των μεσαίων αυτοκινήτων από τους παράγοντες της τιμής, της ανεργίας και της τιμής της βενζίνης.

Εικόνα 5.2.18: Αποτελέσματα εντολής Συσχέτισης των πωλήσεων μεσαίων αυτοκινήτων και τιμής πώλησης

		M	M_PR
M	Pearson Correlation	1	,281*
	Sig. (2-tailed)		,029
	N	60	60
M_PR	Pearson Correlation	,281*	1
	Sig. (2-tailed)	,029	
	N	60	60

\*. Correlation is significant at the 0.05 level (2-tailed).

Βλέπουμε ότι οι πωλήσεις των μεσαίων αυτοκινήτων χαρακτηρίζονται από μέτρια συσχέτιση με την τιμή πώλησης αφού ο συντελεστής Pearson ισούται με 0.281 και ανήκει στο διάστημα (0.2-0.4]. Ακόμη η συσχέτιση είναι σημαντική για επίπεδο σημαντικότητας 0.05.

Εικόνα 5.2.19: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μεσαίων αυτοκινήτων και δείκτη ανεργίας

		M	UN
M	Pearson Correlation	1	,231
	Sig. (2-tailed)		,076
	N	60	60
UN	Pearson Correlation	,231	1
	Sig. (2-tailed)	,076	
	N	60	60

Παρατηρώ και εδώ ότι ο δείκτης ανεργίας χαρακτηρίζεται από μέτρια συσχέτιση με τις πωλήσεις των αυτοκινήτων αφού είναι  $\text{Pearson} = 0.231$  και ανήκει στο διάστημα (0.2-0.4].

Εικόνα 5.2.20: Αποτελέσματα εντολής Συσχέτισης των πωλήσεων μεσαίων αυτοκινήτων και τιμής καυσίμου

		M	FUEL
M	Pearson Correlation	1	-,404**
	Sig. (2-tailed)		,001
	N	60	60
FUEL	Pearson Correlation	-,404**	1
	Sig. (2-tailed)	,001	
	N	60	60

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

Παρατηρώ στο σημείο αυτό ότι η τιμή του καυσίμου επηρεάζει σημαντικά τις πωλήσεις των μεσαίων αυτοκινήτων αφού είναι Pearson= -0.404, και ανήκει στο διάστημα (0.4-0.6]. επομένως η συσχέτιση κρίνεται σημαντική. Το αρνητικό πρόσημο υποδηλώνει ότι η αύξηση της τιμής του καυσίμου, μειώνει τις πωλήσεις των μεσαίων αυτοκινήτων.

Εικόνα 5.2.21: Αποτελέσματα εντολής Συσχέτισης των πωλήσεων μεγάλων αυτοκινήτων και τιμής πώλησης

		L	L_PR
L	Pearson Correlation	1	,199
	Sig. (2-tailed)		,127
	N	60	60
L_PR	Pearson Correlation	,199	1
	Sig. (2-tailed)	,127	
	N	60	60

Όπως βλέπουμε στον παραπάνω πίνακα η τιμή των μεγάλων αυτοκινήτων δεν επηρεάζει σημαντικά τις πωλήσεις τους αφού ο συντελεστής Pearson ισούται με 0.199 και ανήκει στο διάστημα [0.0-0.2]. Η συσχέτιση πώλησης μεγάλων οχημάτων και τιμής τους χαρακτηρίζεται ως ασήμαντη. Αυτό δικαιολογείται από το γεγονός ότι τα μεγάλα αυτοκίνητα απευθύνονται σε μεμονωμένο κοινό και σε ιδιαίτερες ανάγκες.

Εικόνα 5.2.22: Αποτελέσματα εντολής Συσχέτισης πωλήσεων μεγάλων αυτοκινήτων και δείκτη ανεργίας

Correlations			L	UN
L	Pearson Correlation		1	,123
	Sig. (2-tailed)			,349
	N		60	60
UN	Pearson Correlation		,123	1
	Sig. (2-tailed)		,349	
	N		60	60

Παρατηρούμε ότι και εδώ ότι ο δείκτης ανεργίας δεν επηρεάζει σημαντικά τις πωλήσεις των μεγάλων οχημάτων αφού ο συντελεστής Pearson ισούται με 0.123 και ανήκει στο διάστημα [0.0-0.2]. Η συσχέτιση κρίνεται ασήμαντη.

Εικόνα 5.2.23: Αποτελέσματα εντολής Συσχέτισης των πωλήσεων μεσαίων αυτοκινήτων και τιμής καυσίμου

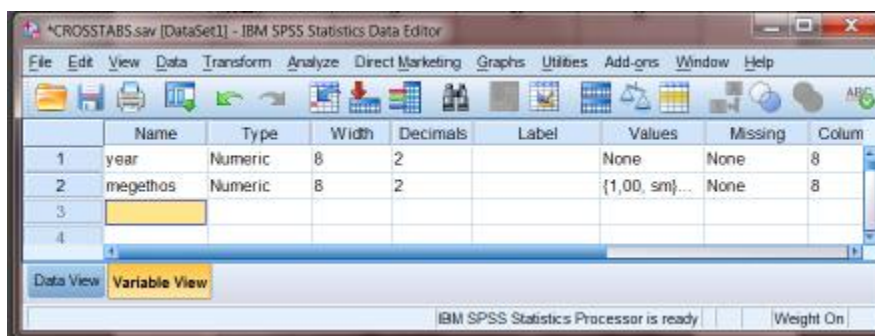
Correlations			L	FUEL
L	Pearson Correlation		1	-,253
	Sig. (2-tailed)			,051
	N		60	60
FUEL	Pearson Correlation		-,253	1
	Sig. (2-tailed)		,051	
	N		60	60

Παρατηρώ ότι η τιμή του καυσίμου χαρακτηρίζεται από μέτρια συσχέτιση με τις πωλήσεις των μεσαίων αυτοκινήτων αφού είναι  $\text{Pearson} = -0.253$  και ανήκει στο διάστημα [0.2-0.4]. Το αρνητικό πρόσημο υποδηλώνει ότι η αύξηση της τιμής του καυσίμου, μειώνει τις πωλήσεις των μεγάλων αυτοκινήτων.

Τέλος, θα εφαρμόσουμε την εντολή Crosstabs όπου πέρα της εμφάνισης των ποσοστιαίων κατανομών στα τρία είδη αυτοκινήτου, θα φανεί αν υπάρχει σχέση αλληλεπίδρασης μεταξύ των μεταβλητών. Πρωταρχικό μας βήμα είναι η σωστή εισαγωγή των δεδομένων σε νέο φύλο επεξεργασίας Excel. Έτσι, λοιπόν, ακολουθούμε την εξής μεθοδολογία:

- 1) Δημιουργούμε στο Variable View τις μεταβλητές «year» και «megethos», οι οποίες είναι αριθμητικού χαρακτήρα. Η μεταβλητή year περιέχει τα έτη από 2004 έως 2008 και η μεταβλητή megethos τα είδη αυτοκινήτου δηλαδή s, m, l.

Εικόνα 5.2.24: Δημιουργία μεταβλητών στο Variable View



Επίσης, για ευκολία στην επεξεργασία των δεδομένων, στην καρτέλα Values της μεταβλητής megethos θα συμβολίσουμε με αριθμούς (εφόσον ορίσαμε τη μεταβλητή μας αριθμητική) τα μεγέθη, δηλαδή αντιστοιχίζουμε τον αριθμό ένα με τη κατηγορία s, τον αριθμό 2 με την κατηγορία m και τον αριθμό 3 με την κατηγορία l.

Εικόνα 5.2.25: Κωδικοποίηση μεταβλητών στο Variable View

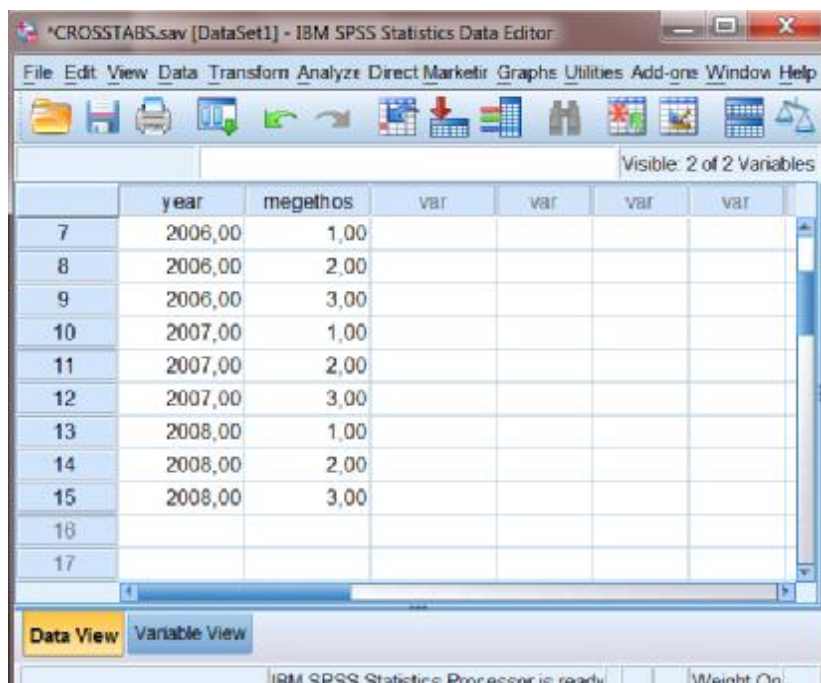


Και στο φύλο **Data View** εισάγουμε ως τιμές για τις μεταβλητές αυτές, όλους τους δυνατούς συνδυασμούς των τιμών τους, όπως φαίνεται στην εικόνα που



ακολουθεί. Έτσι, για κάθε έτος έχουμε 3 κατηγορίες μεγεθών, δηλαδή, s, m, l οι οποίες συμβολίζονται με 1,2,3 αντίστοιχα όπως προείπαμε.

Εικόνα 5.2.26: Εισαγωγή τιμών μεταβλητών στο Data View

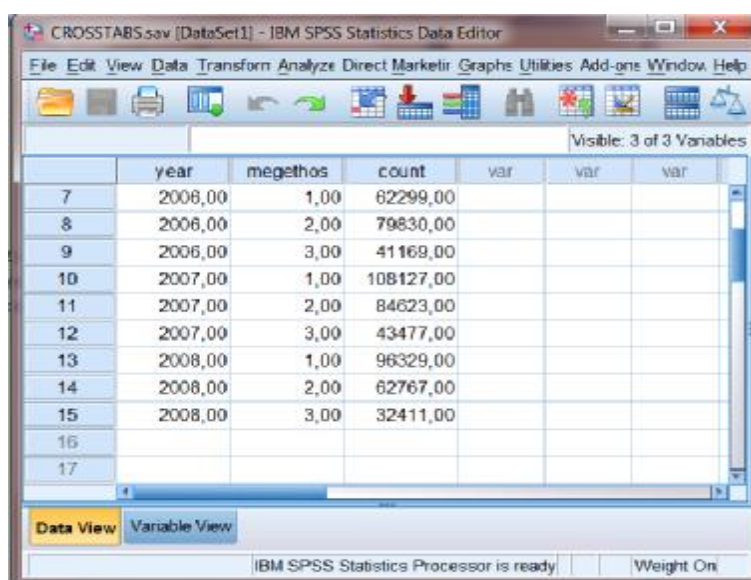


The screenshot shows the IBM SPSS Statistics Data Editor window for a file named 'CROSSTABS.sav'. The 'Data View' tab is active, displaying a table with 17 rows and 7 columns. The first two columns are labeled 'year' and 'megethos'. The 'year' column contains values 2006,00, 2007,00, and 2008,00. The 'megethos' column contains values 1,00, 2,00, and 3,00. The remaining columns are labeled 'var'. The status bar at the bottom indicates 'IBM SPSS Statistics Processor is ready' and 'Weight On'.

	year	megethos	var	var	var	var
7	2006,00	1,00				
8	2006,00	2,00				
9	2006,00	3,00				
10	2007,00	1,00				
11	2007,00	2,00				
12	2007,00	3,00				
13	2008,00	1,00				
14	2008,00	2,00				
15	2008,00	3,00				
16						
17						

- 2) Στη συνέχεια, δημιουργούμε την μεταβλητή **COUNT (Numeric)** στην οποία για τιμές εισάγουμε την συχνότητα εμφάνισης των συνδυασμών των τιμών, των μεταβλητών s, m, l ανά έτος:

Εικόνα 5.2.27: Εισαγωγή τιμών μεταβλητής count στο Data View

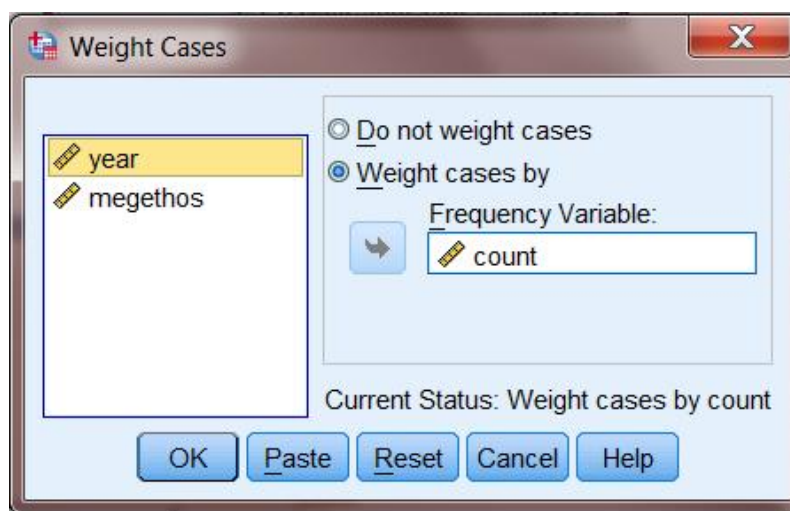


The screenshot shows the IBM SPSS Statistics Data Editor window for the same file 'CROSSTABS.sav'. The 'Data View' tab is active, and a third column labeled 'count' has been added. This column contains numerical values representing the frequency of each combination of year and megethos. The status bar at the bottom indicates 'IBM SPSS Statistics Processor is ready' and 'Weight On'.

	year	megethos	count	var	var	var
7	2006,00	1,00	62299,00			
8	2006,00	2,00	79830,00			
9	2006,00	3,00	41169,00			
10	2007,00	1,00	108127,00			
11	2007,00	2,00	84623,00			
12	2007,00	3,00	43477,00			
13	2008,00	1,00	96329,00			
14	2008,00	2,00	62767,00			
15	2008,00	3,00	32411,00			
16						
17						

Χρησιμοποιούμε την εντολή **Weight Cases**, η οποία σταθμίζει τις παρατηρήσεις ενός αρχείου με βάση τις τιμές μίας μεταβλητής συχνοτήτων (count variable). Η στάθμιση γίνεται με προσομοιωμένες επαναλήψεις των παρατηρήσεων του αρχείου, σύμφωνα με τις τιμές της μεταβλητής στάθμισης. Δηλαδή οι τιμές της μεταβλητής στάθμισης υποδεικνύουν τον αριθμό των επαναλήψεων που θα προκύψουν για κάθε παρατήρηση του αρχικού αρχείου μετά τη στάθμισή του. **Η στάθμιση γίνεται μέσω της εντολής Data - Weight Cases:**

Εικόνα 5.2.28: Επεξεργασία εντολής Weight Cases



Με τη εκτέλεση της εντολής Weight Cases δεν έχουν προστεθεί cases στο φύλλο των δεδομένων. Οι επαναλήψεις έχουν υλοποιηθεί εσωτερικά, γεγονός που είναι περισσότερο ευδιάκριτο με τους πίνακες των συχνοτήτων που παρατίθενται στη συνέχεια.

Εικόνα 5.2.29: Πίνακας συχνοτήτων μεταβλητής year

year				
	Frequency	Percent	Valid Percent	Cumulative Percent
2004,00	204610	20,1	20,1	20,1
2005,00	201389	19,8	19,8	39,9
2006,00	183298	18,0	18,0	57,9
2007,00	236227	23,2	23,2	81,2
2008,00	191507	18,8	18,8	100,0
Total	1017031	100,0	100,0	

Σύμφωνα με τον πίνακα παρατηρούμε ότι οι συνολικές πωλήσεις αυτοκινήτων ήταν περισσότερες κατά το έτος 2007 με ποσοστό που ανέρχεται στο 23,2%, ενώ οι λιγότερες πωλήσεις έγιναν το έτος 2006 με ποσοστό 18%. Γενικά, οι διαφορές από χρονιά σε χρονιά θεωρούνται αμελητέες.

Εικόνα 5.2.30: Πίνακας συχνοτήτων μεταβλητής megethos

megethos					
	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	sm	435445	42,8	42,8	42,8
	m	389264	38,3	38,3	81,1
	l	192322	18,9	18,9	100,0
	Total	1017031	100,0	100,0	

Όσον αφορά στη συχνότητα αγοράς των διαφόρων κατηγοριών συνολικά και για τα πέντε έτη διαπιστώνουμε ότι κυριαρχεί η αγορά μικρών οχημάτων σε ποσοστό που ανέρχεται στο 42,8%, ακολουθεί η αγορά μεσαίου τύπου οχημάτων σε ποσοστό 38,3% και τέλος με μεγάλη διαφορά η αγορά μεγάλων οχημάτων σε ποσοστό 18,9%.

Στην συνέχεια για την διερεύνηση της σχέσης μεταξύ των μεταβλητών δημιουργούμε τον πίνακα συνάφειας (crosstabulation) μέσω της εντολής Analyze – Descriptive Statistics – Crosstabs.

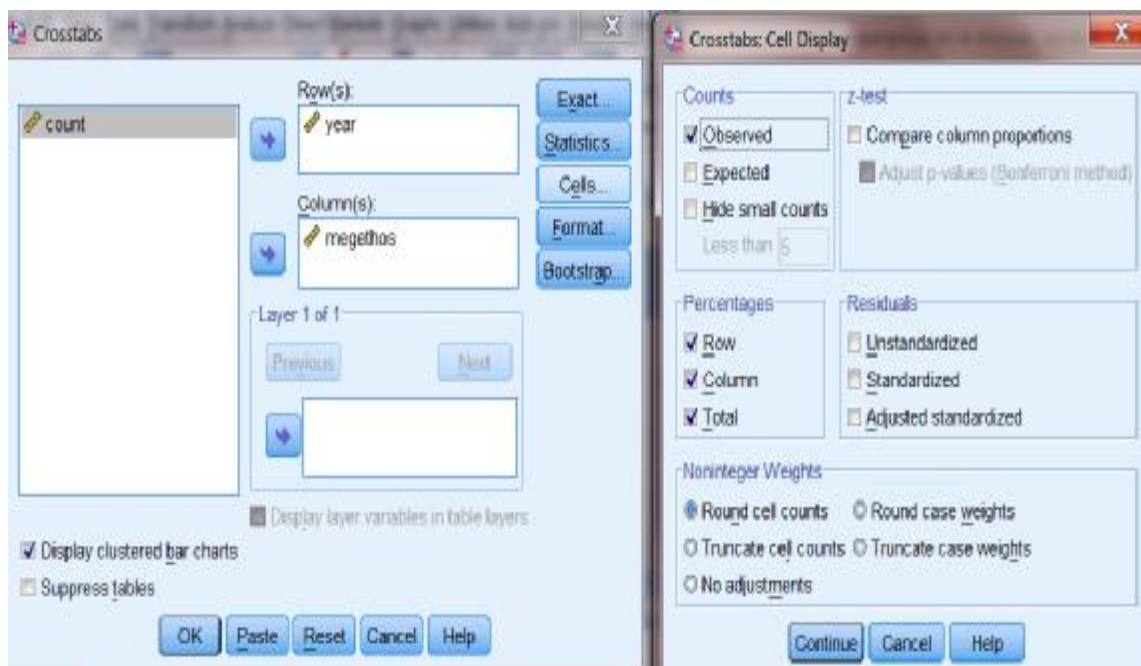
Στο παράθυρο Row(s) τοποθετούμε την μεταβλητή η οποία και θέλουμε να αποτελεί τις γραμμές του πίνακα μας, και αντίστοιχα στο παράθυρο Column(s) τοποθετούμε την μεταβλητή που θέλουμε να αποτελεί τις στήλες. Επίσης για την δημιουργία του γραφήματος των δυο μεταβλητών επιλέγουμε «Display clustered bar charts» για την παραγωγή γραφήματος.

Από την ένδειξη Percentages μπορούμε να επιλέξουμε:

- ROW, αν θέλουμε να εμφανίζονται τα ποσοστά των γραμμών,
- Column, αν θέλουμε να εμφανίζονται τα ποσοστά των στηλών, και
- Total, αν θέλουμε να εμφανίζονται τα συνολικά ποσοστά.

Επιλέγουμε την εμφάνιση και των τριών ποσοστών:

Εικόνα 5.2.31: Επεξεργασία εντολής Crosstabs



Κ έτσι παίρνουμε τον πίνακα των αποτελεσμάτων στο παράθυρο Output:

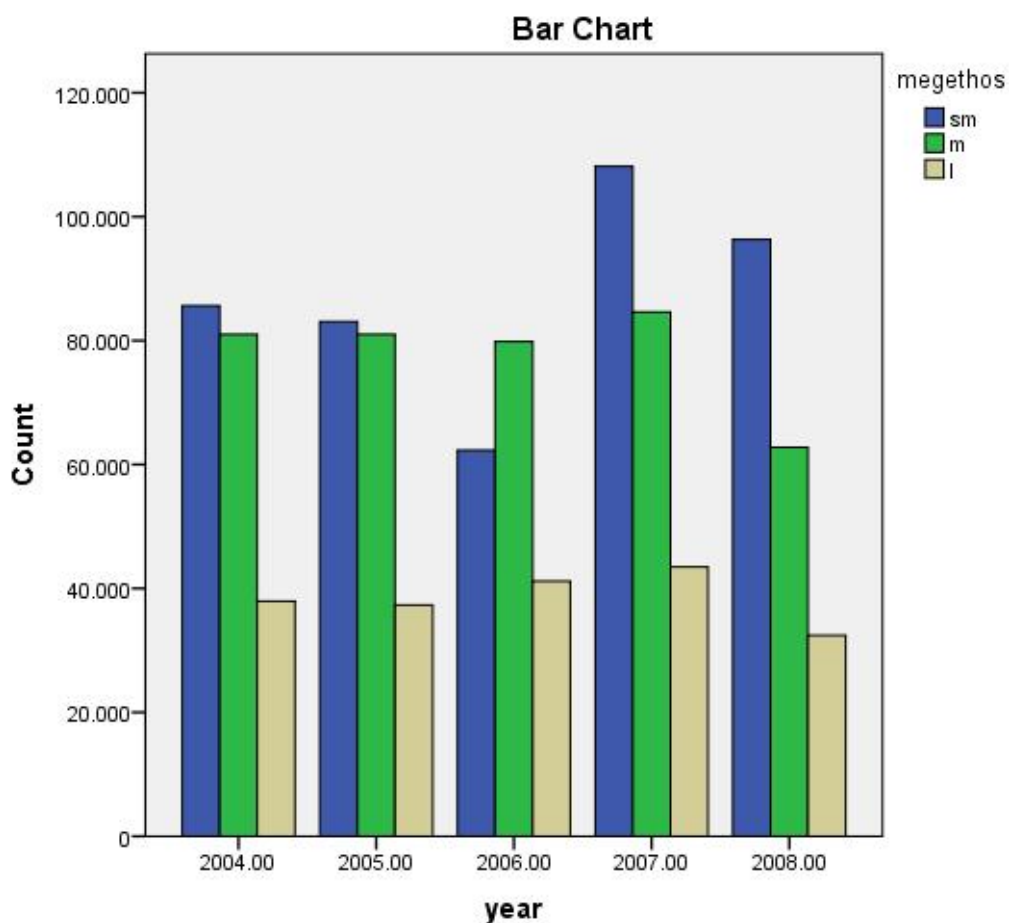
Εικόνα 5.2.32: Πίνακας αποτελεσμάτων Crosstabs

		year * megethos Crosstabulation				
		megethos			Total	
		sm	m	l		
year	2004,00	Count	85639	81022	37949	204610
		% within year	41,9%	39,6%	18,5%	100,0%
		% within megethos	19,7%	20,8%	19,7%	20,1%
		% of Total	8,4%	8,0%	3,7%	20,1%
	2005,00	Count	83051	81022	37316	201389
		% within year	41,2%	40,2%	18,5%	100,0%
		% within megethos	19,1%	20,8%	19,4%	19,8%
		% of Total	8,2%	8,0%	3,7%	19,8%
	2006,00	Count	62299	79830	41169	183298
		% within year	34,0%	43,6%	22,5%	100,0%
		% within megethos	14,3%	20,5%	21,4%	18,0%
		% of Total	6,1%	7,8%	4,0%	18,0%
2007,00	Count	108127	84623	43477	236227	
	% within year	45,8%	35,8%	18,4%	100,0%	
	% within megethos	24,8%	21,7%	22,6%	23,2%	

	% of Total	10,6%	8,3%	4,3%	23,2%
	Count	96329	62767	32411	191507
2008,00	% within year	50,3%	32,8%	16,9%	100,0%
	% within megethos	22,1%	16,1%	16,9%	18,8%
	% of Total	9,5%	6,2%	3,2%	18,8%
	Count	435445	389264	192322	1017031
Total	% within year	42,8%	38,3%	18,9%	100,0%
	% within megethos	100,0%	100,0%	100,0%	100,0%
	% of Total	42,8%	38,3%	18,9%	100,0%

Καθώς και το αντίστοιχο γράφημα:

Εικόνα 5.2.33: Γράφημα αποτελεσμάτων Crosstabs



Παρατηρούμε ότι η σειρά % within year δείχνει την κατανομή των τριών κατηγοριών οχημάτων για κάθε έτος. Έτσι, για παράδειγμα για το έτος 2005 οι πωλήσεις των

μικρών οχημάτων αποτελούσαν το 41,2% των πωλήσεων, των μεσαίων οχημάτων το 40,2% και των μεγάλων οχημάτων το 18,5%.

Η σειρά %within megethos δείχνει τη κατανομή κάθε κατηγορίας οχήματος ξεχωριστά μέσα στο πέρασμα των ετών, οπότε βλέπουμε και την μεταβολή τους στο χρόνο. Για παράδειγμα, τα μεσαία αυτοκίνητα το 2004 αποτελούσαν το 20,8%, το 2005 το 20,8%, το 2006 το 20,5%, το 2007 ανέβηκαν στο 21,7% και τέλος έπεσαν σημαντικά στο 16,1 %.

Η στήλη %of Total δείχνει το συνολικό αποτέλεσμα που είναι ότι από το σύνολο του δείγματός μας, για το έτος 2004 τα μικρά οχήματα αποτελούσαν το 8.4% του συνόλου των παρατηρήσεων, τα μεσαία το 8,0% και τα μεγάλα το 3.7%. Ακόμη, στο σύνολο του δείγματος οι πωλήσεις τόσο των μικρών όσο και των μεσαίων και μεγάλων οχημάτων ήταν περισσότερες κατά το 2007, καθώς παρατηρήθηκε και συνολικότερη αύξηση στις πωλήσεις οχημάτων.

Τα έτη 2004 και 2005 δεν υπάρχουν αξιόλογες μεταβολές στην κατανομή των πωλήσεων ανά κατηγορία οχήματος, σε αντίθεση με το έτος 2006 όπου παρατηρείται μείωση των πωλήσεων των μικρών αυτοκινήτων και παράλληλη αύξηση των πωλήσεων των μεσαίων και μεγάλων οχημάτων. Πιο συγκεκριμένα, οι πωλήσεις των μικρών οχημάτων από 41,9% το 2004 έπεσαν στο 34% το 2006, των μεσαίων από 39,6% το 2004 ανέβηκαν στο 43,6% το 2006 και των μεγάλων από 18,5% το 2004 ανήλθαν σε 22,5% το 2006. Επίσης, τα έτη 2007 και 2008 παρατηρείται μεγάλη αύξηση των πωλήσεων των μικρών οχημάτων αφού από 34% το 2006 ανήλθαν σε 50,3% το 2008, και παράλληλη πτώση των μεσαίων οχημάτων όπου από 43,6% το 2006 ανήλθαν σε 32.8% το 2008. Τα μεγάλα οχήματα επίσης σημειώνουν καθοδική πορεία, ωστόσο η διαφορά κρίνεται ασήμαντη συγκριτικά με των άλλων δυο κατηγοριών.

## Βιβλιογραφία

1. Μποζίκη Ε., Νομικού Α., « Η Αξιολόγηση της Εφαρμογής των συστημάτων Εσωτερικού Ελέγχου (ΣΕΕ) του Κύκλου Αγορών – Αποθεμάτων στις Εισηγμένες επιχειρήσεις του ΧΑΑ», Ανώτατο Τεχνολογικό Εκπαιδευτικό Ίδρυμα Σερρών, Σέρρες, 2011
2. Σωσσίδου Ε., Ψευτογιάννη Δ., «Μεθοδολογία Έρευνας & Στατιστική με τη χρήση του SPSS 13.00 for Windows», Θεσσαλονίκη, 2007
3. Σπανούδης Γ., «Εισαγωγή στο SPSS», Τμήμα Ψυχολογίας, ΚΕΔΙΜΑ, 2013
4. Τσομπανάκη Ε., «Εισαγωγή στη χρήση του στατιστικού πακέτου SPSS 14.00»
5. Νέλλας Ε., «Ανάλυση Δεδομένων με χρήση του πακέτου SPSS/PC για WINDOWS», Τμήμα Αγροτικής Οικονομίας και Ανάπτυξης, Γεωπονικό Πανεπιστήμιο Αθηνών, Αθήνα, 2005
6. Φαχρίδης Γ., «Εισαγωγή στο πρόγραμμα SPSS for Windows», Παιδαγωγικό Τμήμα Δημοτικής Εκπαίδευσης, Δημοκρίτειο Πανεπιστήμιο Θράκης, Αλεξανδρούπολη
7. Ρούσσοις Π., Ευσταθίου Γ., «Σύντομο Εγχειρίδιο SPSS 16.00» Πρόγραμμα Ψυχολογίας, Τμήμα ΦΠΨ, Εθνικό Καποδιστριακό Πανεπιστήμιο Αθηνών, Αθήνα, 2008
8. Εμβαλωτής Α., Κατσής Α., Σιδερίδης Γ., «Στατιστική Μεθοδολογία Εκπαιδευτικής Έρευνας», Α' Έκδοση, Ιωάννινα, 2006
9. Νικήτα Ε., « Εφαρμογές Στατιστικής και Εφαρμογές με το SPSS», Θεσσαλονίκη, 2012
10. Δαφέρμος Β., «Κοινωνική Στατιστική & Μεθοδολογία Έρευνας με το SPSS», Εκδόσεις ΖΗΤΗ, Θεσσαλονίκη, 2011

11. Μητρόπουλος Ι., Σημειώσεις για το εργαστήριο του μαθήματος «Εισαγωγή στη Στατιστική των επιχειρήσεων», Τεχνολογικό Εκπαιδευτικό Ίδρυμα Δυτικής Ελλάδας, Τμήμα Διοίκησης Επιχειρήσεων, Πάτρα, 2015



## Ηλεκτρονική Βιβλιογραφία

1. <http://www.pioneeringedu.com/statistics/>
2. <http://pacific.jour.auth.gr/spss/>
3. [https://www.youtube.com/watch?v=\\_peQHO9lIVg](https://www.youtube.com/watch?v=_peQHO9lIVg)
4. [https://www.youtube.com/watch?v=j\\_\\_6yrhyaYU](https://www.youtube.com/watch?v=j__6yrhyaYU)
5. <https://www.youtube.com/watch?v=tqfsDyWvFQ0>
6. <http://www.de.teipat.gr>
7. <http://www.lib.teicrete.gr>