



Εισαγωγή στο SPSS

Εισαγωγή στο SPSS

Ζιντζαράς Ηλίας, M.Sc., Ph.D.

*Καθηγητής Βιομαθηματικών-Βιομετρίας
Εργαστήριο Βιομαθηματικών
Τμήμα Ιατρικής
Πανεπιστήμιο Θεσσαλίας*

*Institute for Clinical Research and Health Policy Studies
Tufts University School of Medicine
Boston, MA, USA*

*Θεόδωρος Μπρότσης, MSc, PhD
Εντεταλμένος Διδάσκων
(<http://biomath.med.uth.gr>)
Πανεπιστήμιο Θεσσαλίας
Email: tmprotsis@uth.gr*



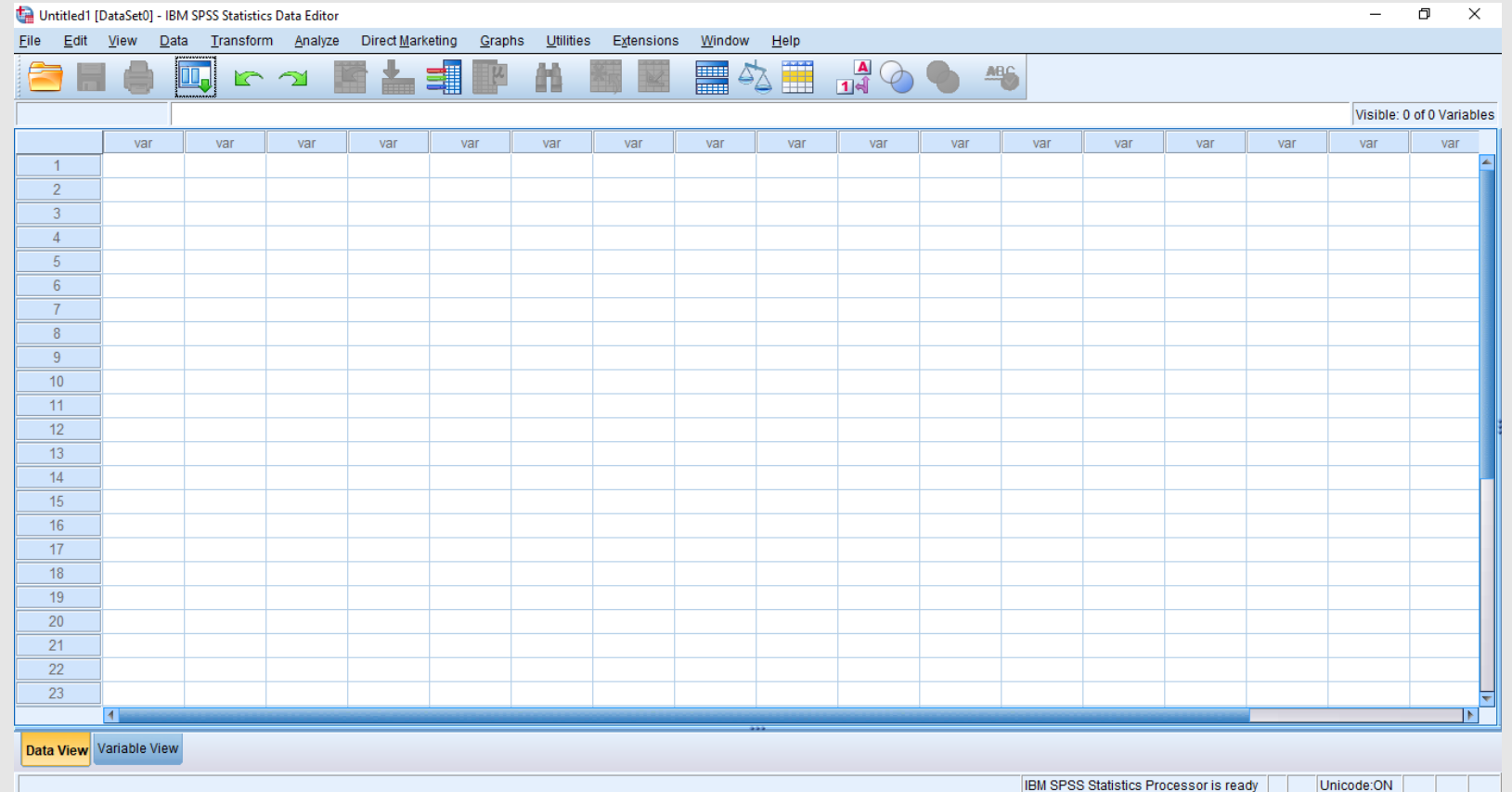
Περιεχόμενο

- Τι είναι το SPSS
- Εισαγωγή/Επεξεργασία δεδομένων στον Data Editor
- Ελλείπουσες τιμές / Missing values
- Περιγραφική Στατιστική
- Διαγράμματα
 - Ιστόγραμμα
 - Box plot
 - Διάγραμμα διασποράς / Scatter Plot
 - Pie Chart ...
- Επιλογή περιπτώσεων / Select cases
- Μετασχηματισμός και τροποποίηση μεταβλητών
- Στατιστικές διαδικασίες / One Sample T test
- Αποθήκευση δεδομένων



Data Editor

- Είναι ένα λογισμικό επεξεργασίας στατιστικών δεδομένων
- Τα δεδομένα αποθηκεύονται σε μεταβλητές
- Κάθε στήλη είναι μία μεταβλητή-μονοδιάστατος πίνακας δεδομένων με όνομα το περιεχόμενο του πάνω κελιού της στήλης και κάθε γραμμή μία περίπτωση (**case**)

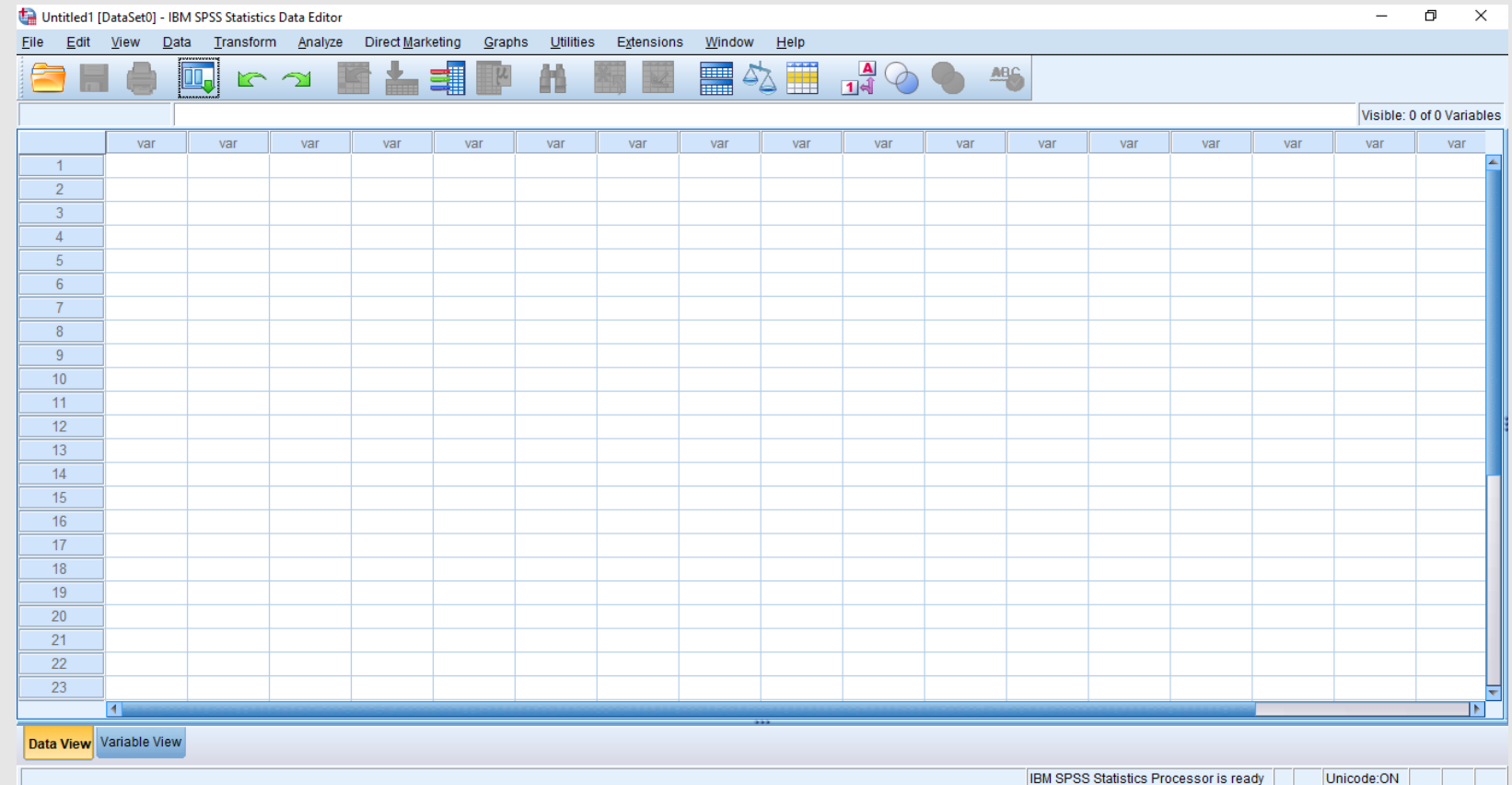


Data Editor



Εισαγωγή δεδομένων στο παράθυρο Data Editor

- Τα στατιστικά δεδομένα που επεξεργαζόμαστε είναι ένα σύνολο αριθμητικών ή αλφαριθμητικών τιμών
- Τα στατιστικά δεδομένα μπορεί να
 - είναι αποθηκευμένα σε ένα αρχείο δεδομένων .sav του SPSS
 - Είναι αποθηκευμένα σε ένα αρχείο δεδομένων .xls ή .xlsx του Excel
 - να μην είναι σε ηλεκτρονική μορφή



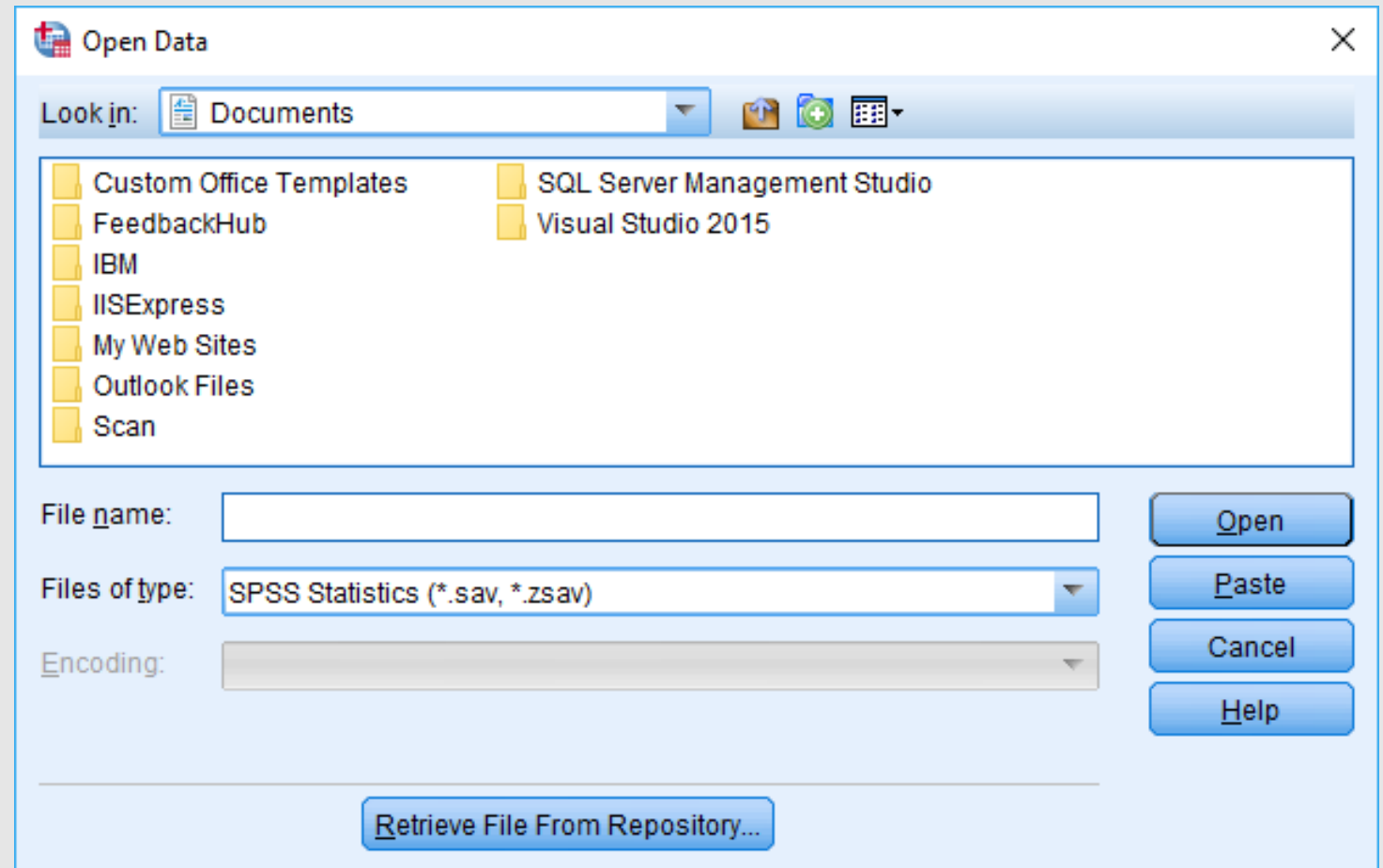
Data Editor



Άνοιγμα αρχείου

Αν τα δεδομένα είναι αποθηκευμένα σε ένα αρχείο δεδομένων του SPSS (*.sav) τότε επιλέγουμε από το μενού του Data Editor διαδοχικά

File -> Open -> Data



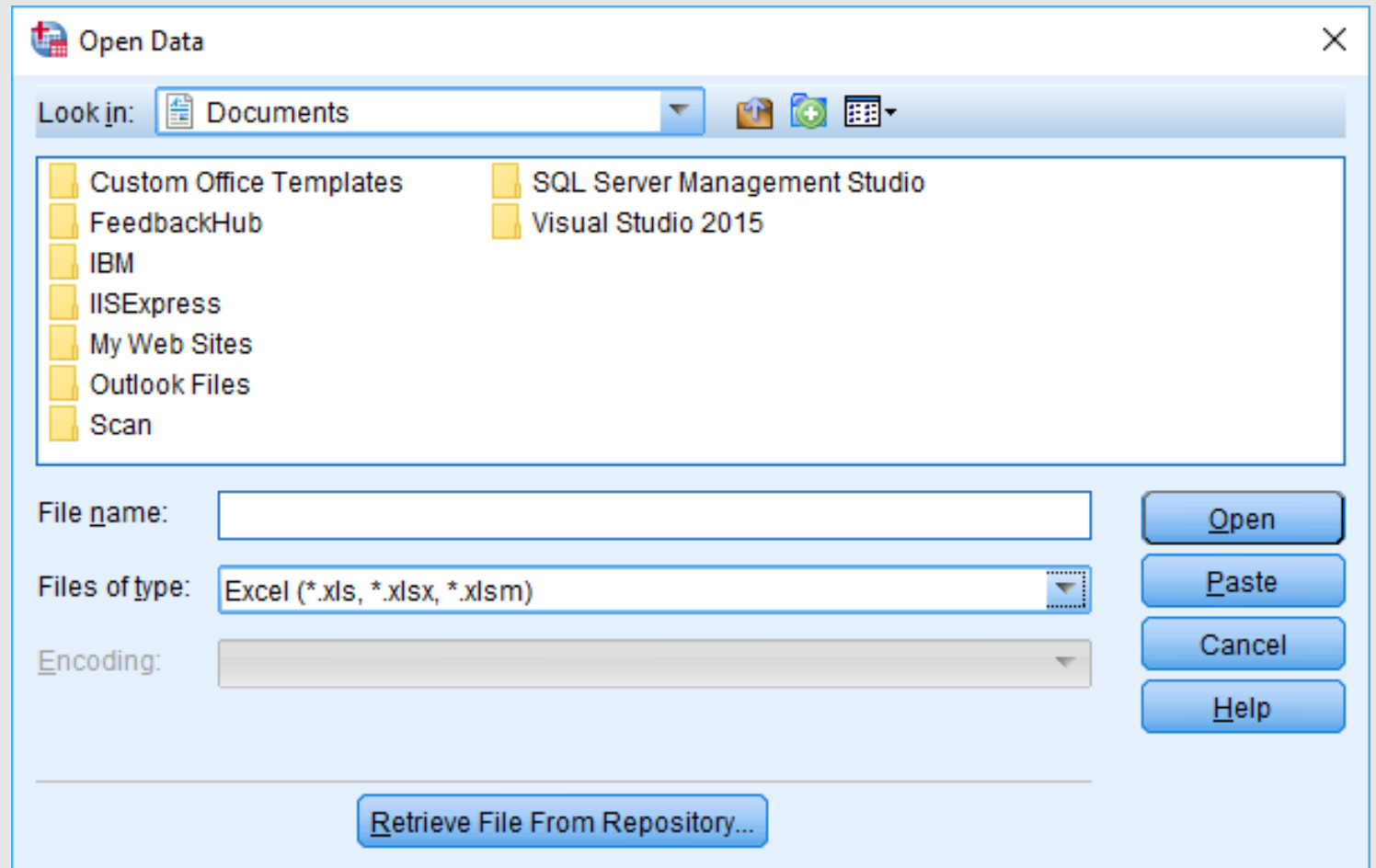


Άνοιγμα αρχείου Excel

Αν τα δεδομένα είναι αποθηκευμένα σε ένα αρχείο δεδομένων του Excel (*.xls ή *.xlsx) τότε επιλέγουμε από το μενού του Data Editor διαδοχικά

File -> Open -> Data

και στο παράθυρο που εμφανίζεται επιλέγουμε Excel (*.xls, *.xlsx, *.xlsm) στο πεδίο **Files of type**



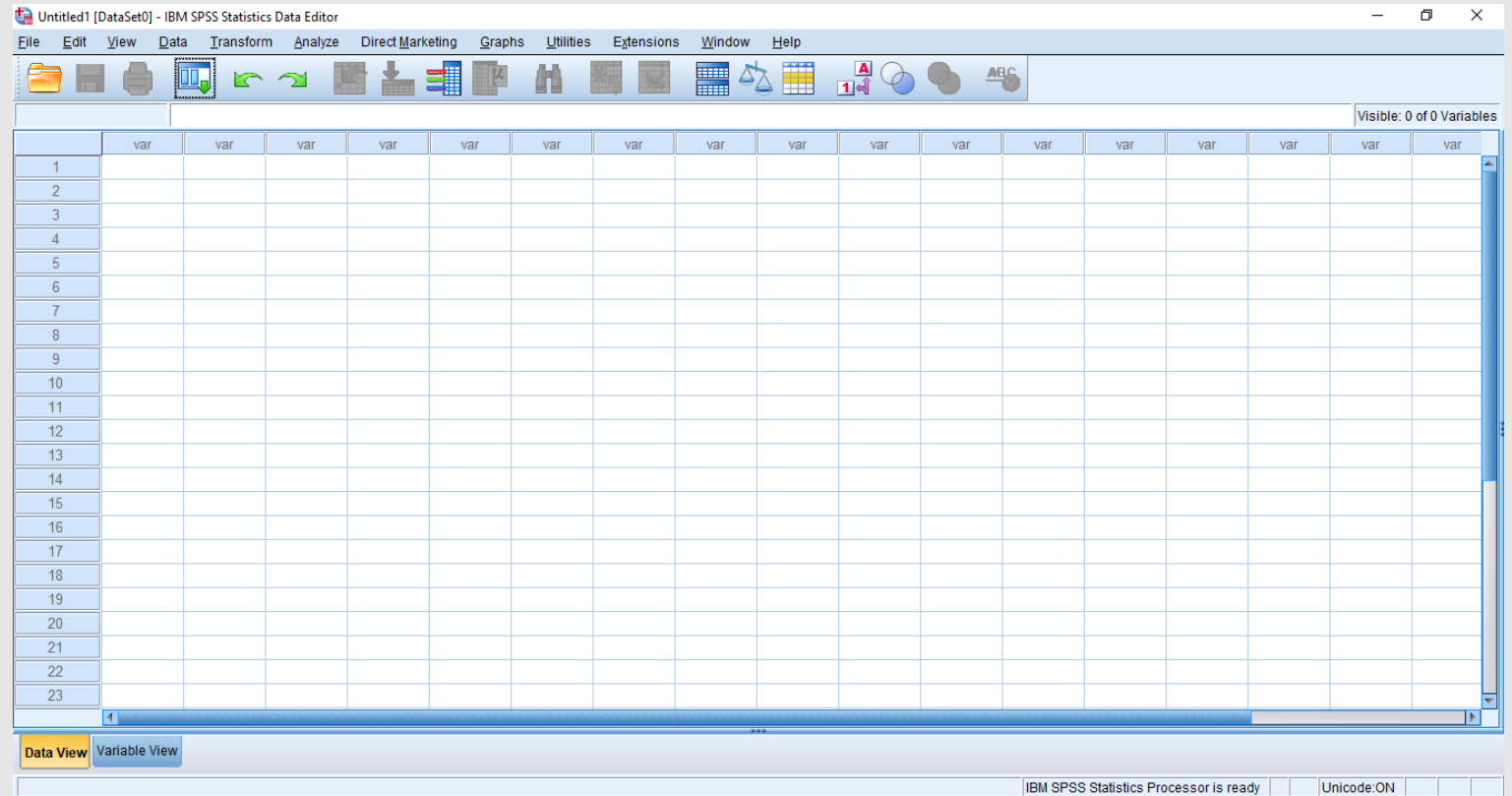


Εισαγωγή δεδομένων στο παράθυρο Data Editor

Αν τα δεδομένα **δεν** είναι σε ψηφιακή μορφή, τότε τα πληκτρολογούμε στις στήλες του **Data Editor**

Στον **Data Editor** τα δεδομένα εμφανίζονται με έναν από τους παρακάτω δύο τρόπους:

- **Data View**
- **Variable View**

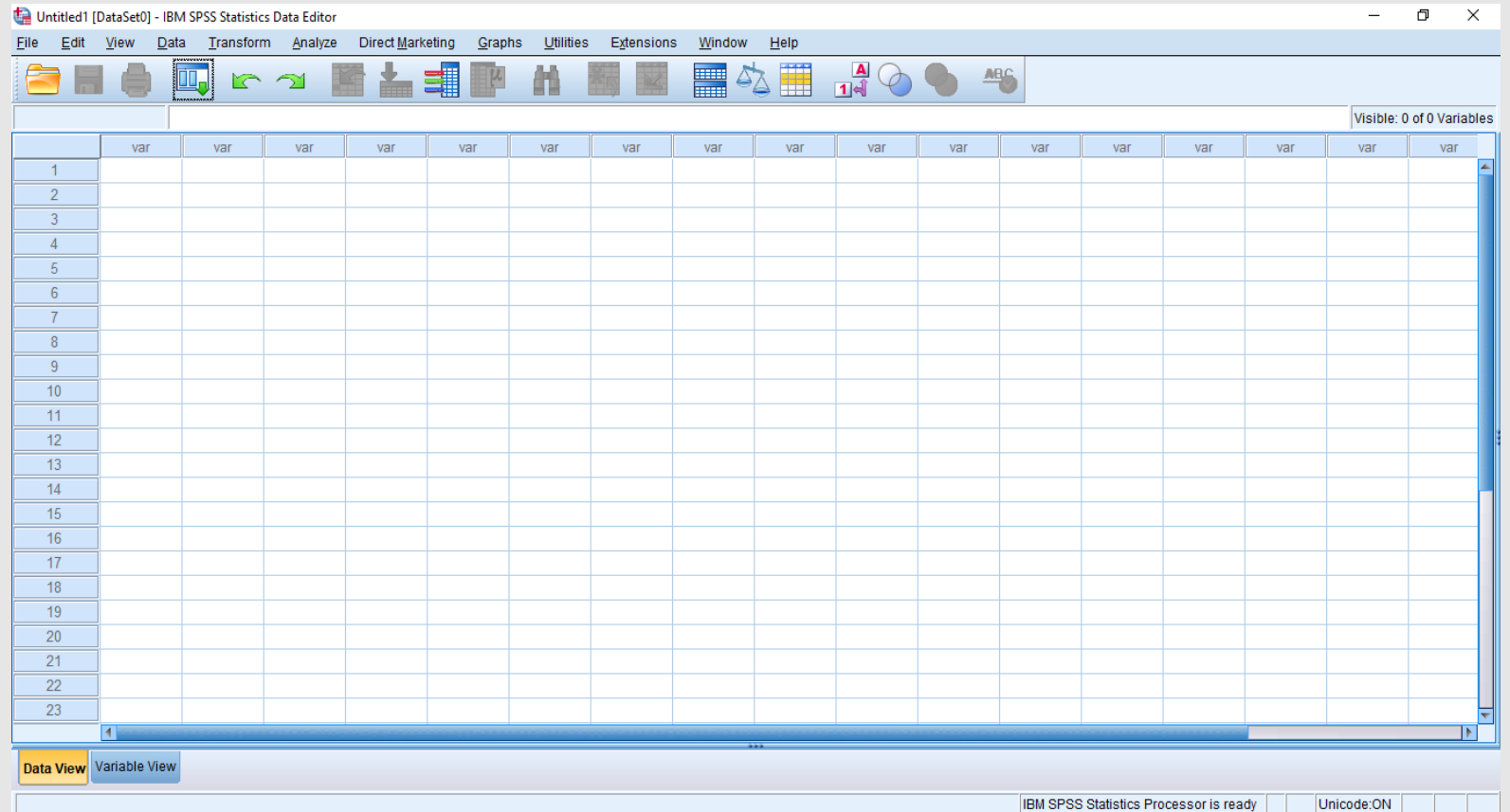


Data Editor



Data View / Variable View

- **Data View:** Εδώ εμφανίζονται οι τιμές των δεδομένων
- **Variable View:** Εδώ ρυθμίζουμε τις μεταβλητές



Data Editor

Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1											
2											
3	■	Name: Όνομα μεταβλητής									
4											
5	■	Type: Ο τύπος της μεταβλητής									
6											
7	■	Width: Ο αριθμός των ψηφίων									
8											
9	■	Decimals: ο αριθμός των δεκαδικών									
10											
11	■	Label: Η ετικέτα/περιγραφή της μεταβλητής									
12											
13	■	Values: Οι ετικέτες των τιμών των μεταβλητών									
14											
15	■	Missing: Κωδικοποίηση των ελλειπουσών τιμών									
16											
17	■	Columns: Το πλάτος της στήλης									
18											
19	■	Align: Η στοίχιση των στοιχείων των κελιών (Left, Right, Center)									
20											
21	■	Measure: Το επίπεδο μέτρησης της μεταβλητής (Ordinal, Nominal, Scale)									
22											
23	■	Role: Ο ρόλος της μεταβλητής σε μία ανάλυση									
24											
25											

Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON

Data Editor



Παράδειγμα (Data View)

Στην διπλανή εικόνα φαίνονται το φύλο και τα επίπεδα αιμοσφαιρίνης 20 ατόμων (data1.sav).

	sex	hgb	var
1	0	10.80	
2	0	10.30	
3	0	11.50	
4	1	11.60	
5	1	11.10	
6	1	11.40	
7	0	12.00	
8	0	12.00	
9	0	12.10	
10	1	12.10	
11	1	12.90	
12	1	12.90	
13	1	12.30	
14	0	12.40	
15	0	13.50	
16	0	13.60	
17	0	14.10	
18	1	15.10	
19	1	12.90	
20	1	15.10	



Παράδειγμα (Variable View)

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	sex	Numeric	1	0	Sex	{0, Male}...	None	8	Right	Nominal	Input
2	hgb	Numeric	8	2		None	None	8	Right	Scale	Input

Value Labels

Value:

Label:

Spelling...

Add

Change

Remove

0 = "Male"
1 = "Female"

OK Cancel Help

- **Scale:** Ποσοτικές μεταβλητές όπως π. χ. το βάρος, το ύψος κλπ.
- **Nominal:** Κατηγορικές μεταβλητές (ποιοτικές) π. χ. το φύλο κλπ.
- **Ordinal:** Διατεταγμένες μεταβλητές π. χ.
 - 0 = «πολύ δυσανεσθημένος»
 - 1 = «δυσανεσθημένος»
 - 2 = «αδιάφορος»
 - 3 = «ικανοποιημένος»
 - 4 = «πολύ ικανοποιημένος»

Measure

Nominal

Scale

Ordinal

Nominal



Ελλείπουσες τιμές (missing values)

- Ελλείπουσες τιμές (missing values) λέμε τις μη καταχωρημένες τιμές μία μεταβλητής.
- Το SPSS καταχωρεί αυτόματα το σύμβολο του δεκαδικού σημείου (κόμμα ή τελεία ανάλογα με τις ρυθμίσεις των windows).
- Αν για παράδειγμα το εύρος των δυνατών απαντήσεων είναι 0 – 250 (π. χ. το ύψος ατόμων σε cm) τότε σαν ελλείπουσες τιμές θα μπορούσαμε να χρησιμοποιήσουμε τις: -1 = «δεν θέλω να απαντήσω», 999 = «δεν γνωρίζω».
- Στις αλφαριθμητικές μεταβλητές το κενό θεωρείται έγκυρη καταχώρηση. Αν θέλουμε να θεωρηθεί ελλείπουσα τιμή, τότε πρέπει να καταχωρηθεί ως Missing Values.

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Column
1	sex	Numeric	1	0	Sex	{0, Male}...	None	8
2	hgb	Numeric	8	2		None	None	8

Missing Values

☐ No missing values

☒ Discrete missing values

-1 999

☐ Range plus one optional discrete missing value

Low: High:

Discrete value:

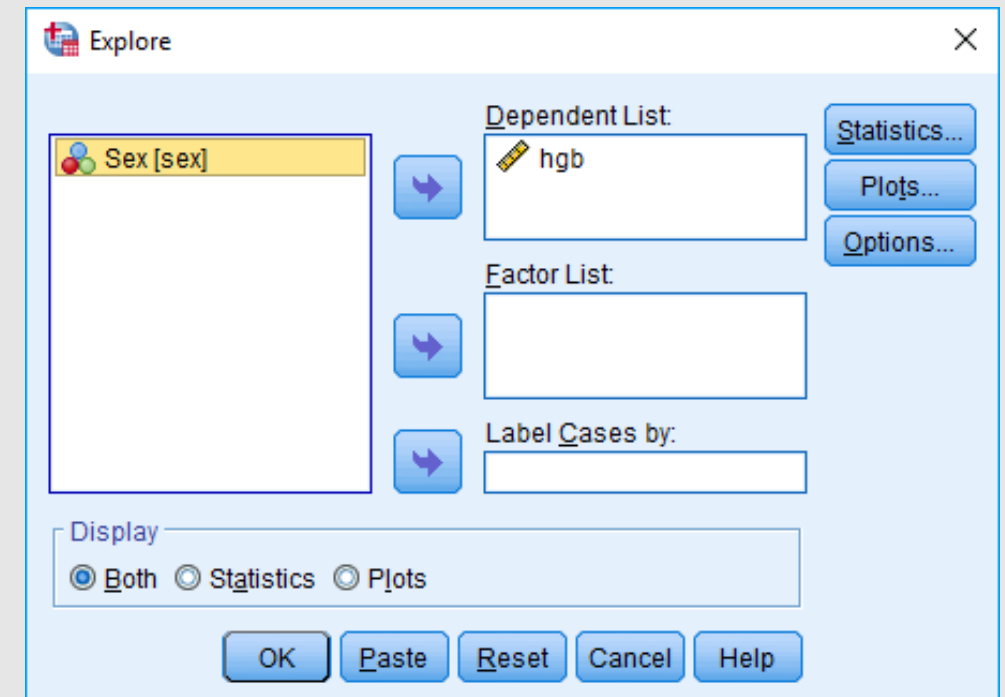
OK Cancel Help

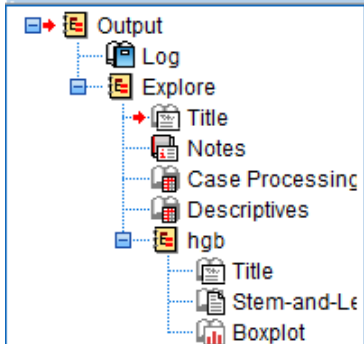


Explore

Με την διαδικασία αυτή λαμβάνουμε την τιμή κάποιων στατιστικών συναρτήσεων για μία μεταβλητή όπως sample size, mean, median, minimum, maximum, standard deviation, variance, range, standard error κ.α.

- Επιλέγουμε **Analyze -> Descriptive Statistics -> Explore**
- Σέρνουμε την μεταβλητή hgb στο πλαίσιο **Dependent List**
- Πατάμε **OK**





Explore

Case Processing Summary

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
hgb	20	100.0%	0	0.0%	20	100.0%

Descriptives

		Statistic	Std. Error
hgb	Mean	12.4850	.29175
	95% Confidence Interval for Mean	Lower Bound	11.8744
		Upper Bound	13.0956
	5% Trimmed Mean	12.4611	
	Median	12.2000	
	Variance	1.702	
	Std. Deviation	1.30476	
	Minimum	10.30	
	Maximum	15.10	
	Range	4.80	
	Interquartile Range	1.83	
	Skewness	.565	.512
	Kurtosis	-.033	.992

12.49±1.30

hgb

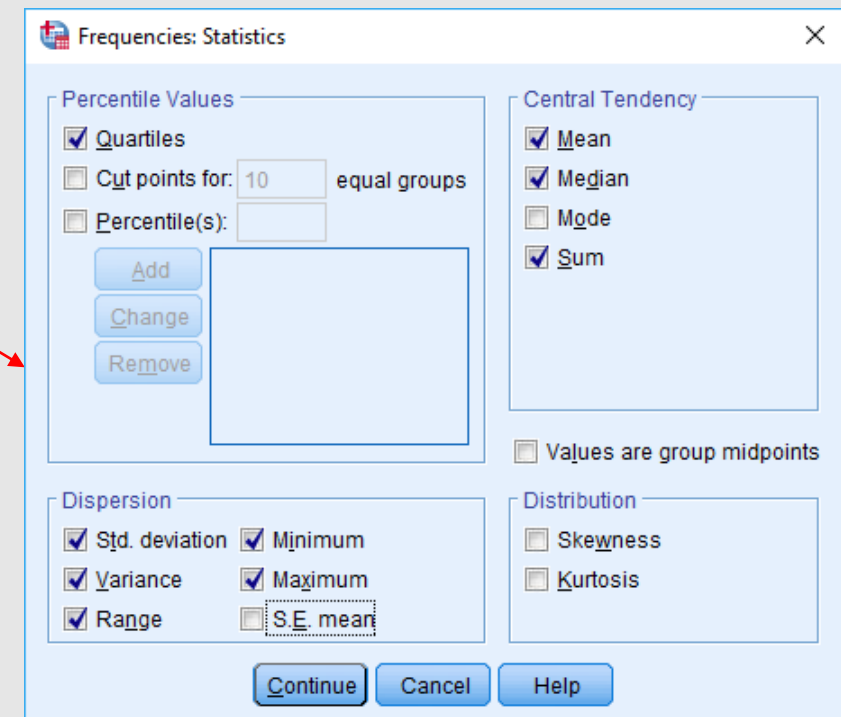
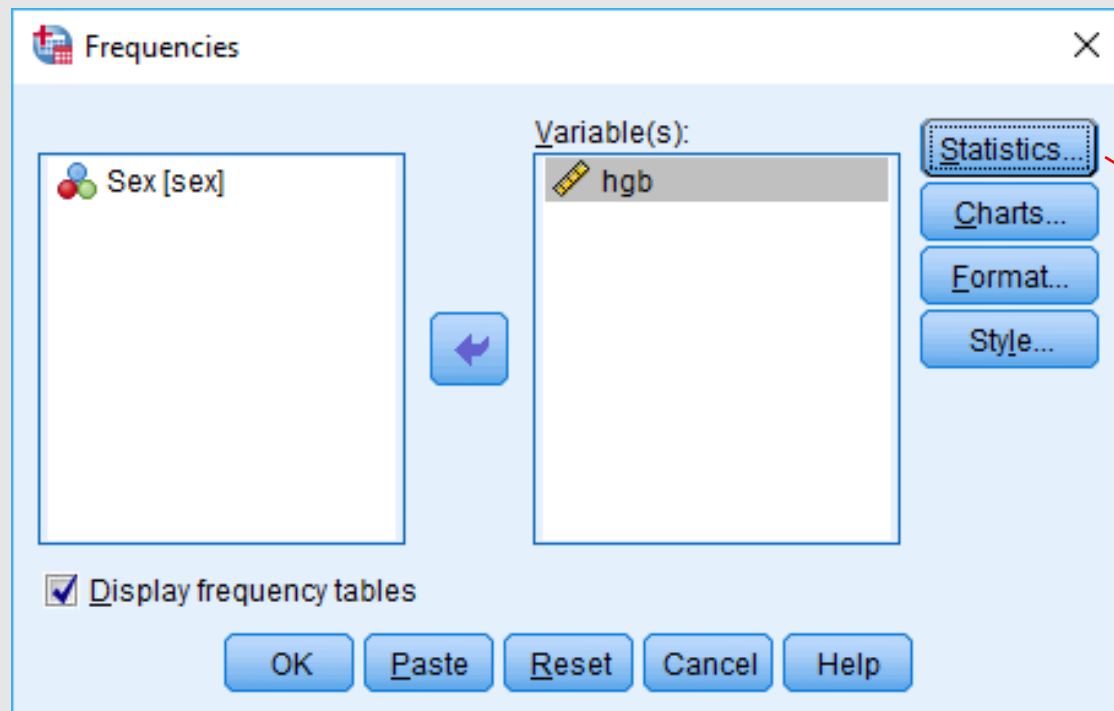
- Η **αριστερή πλευρά** του Viewer περιέχει ένα γενικό περίγραμμα των αποτελεσμάτων.
- Με κλικ πάνω σε κάποιο αποτέλεσμα μεταφερόμαστε σε αυτό
- Η **δεξιά πλευρά** περιέχει στατιστικούς πίνακες, διαγράμματα και κείμενο
- Μέσω **copy – paste** είναι εφικτή η μεταφορά ενός αποτελέσματος σε άλλο πρόγραμμα π. χ. MS Word

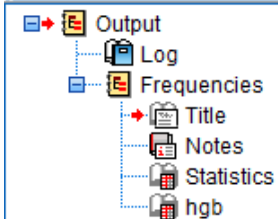


Frequencies

Με την διαδικασία αυτή λαμβάνουμε επίσης την τιμή κάποιων στατιστικών συναρτήσεων για μία μεταβλητή όπως sample size, mean, median, minimum, maximum, standard deviation, variance, range, standard error κ.α.

Επιλέγουμε **Analyze -> Descriptive Statistics -> Frequencies** και στην συνέχεια **Statistics**





FREQUENCIES VARIABLES=hgb

Run Script

/NTILES=4

/STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM MEAN MEDIAN SUM

/ORDER=ANALYSIS.

→ Frequencies

Statistics

hgb

N	Valid	20
	Missing	0
Mean		12.4850
Median		12.2000
Std. Deviation		1.30476
Variance		1.702
Range		4.80
Minimum		10.30
Maximum		15.10
Sum		249.70
Percentiles	25	11.5250
	50	12.2000
	75	13.3500

12.49±1.30

hgb

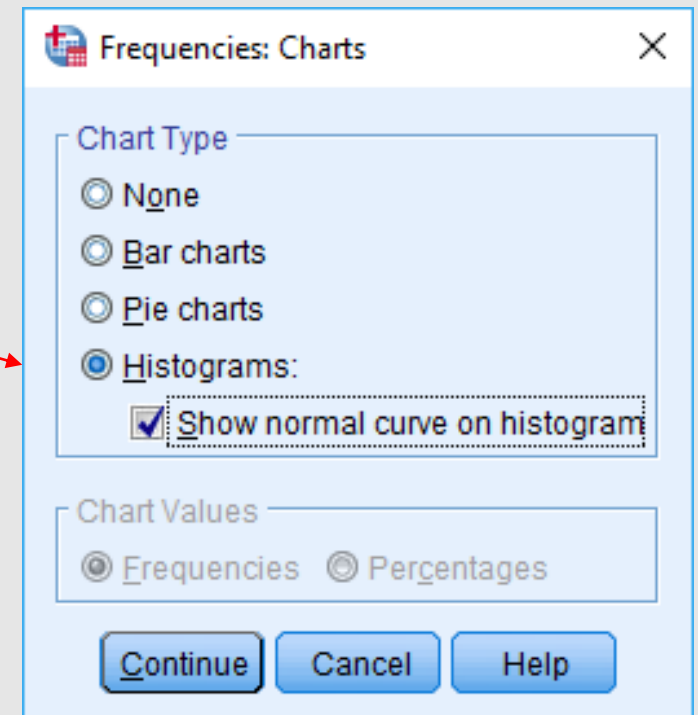
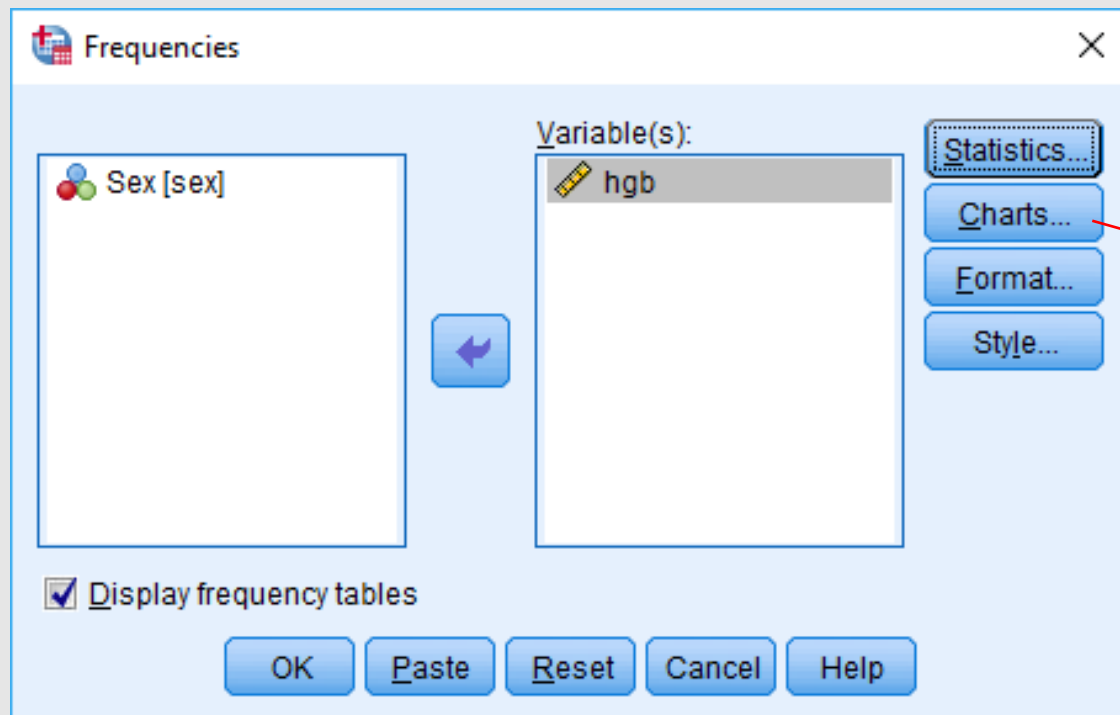
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	10.30	1	5.0	5.0	5.0

Αρχείο εξόδου



Histogram

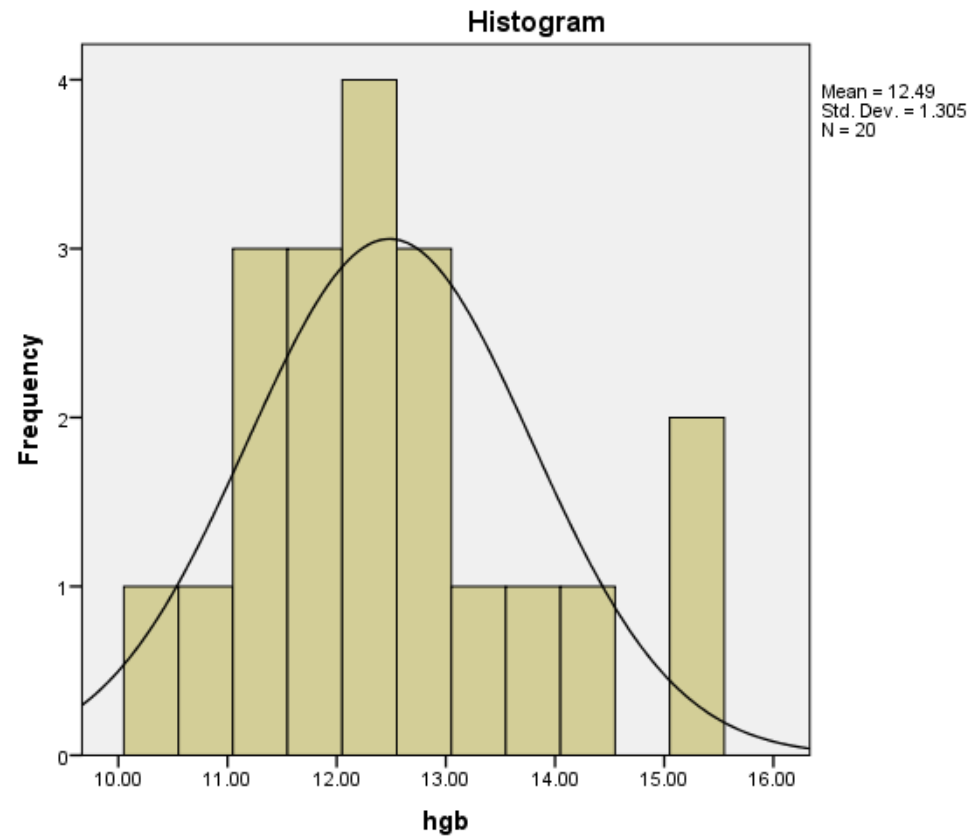
Επιλέγουμε **Analyze -> Descriptive Statistics -> Frequencies** και στην συνέχεια **Charts**





- Output
 - Log
 - Frequencies
 - Title
 - Notes
 - Statistics
 - hgb
 - Histogram

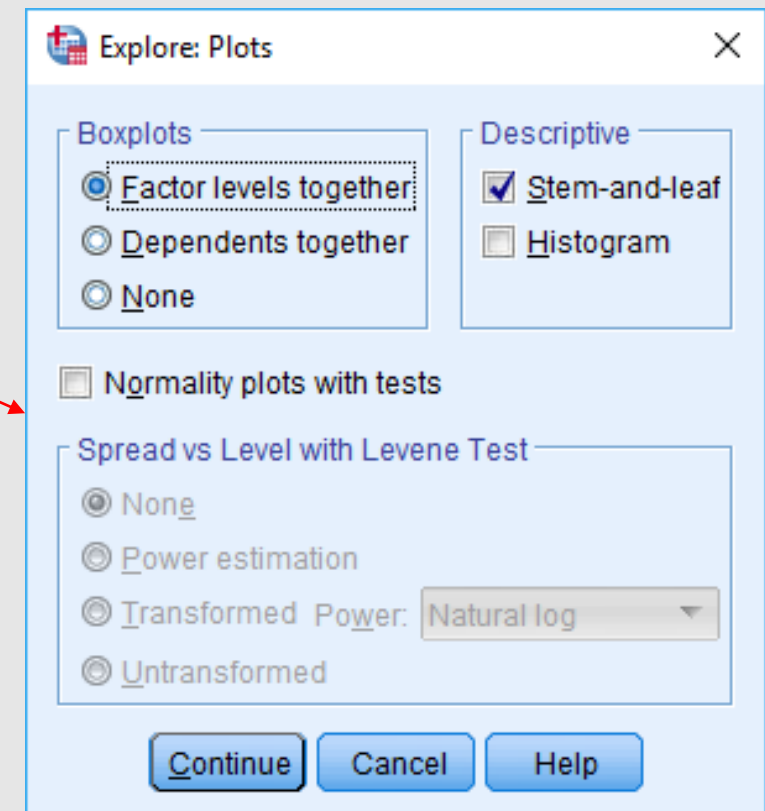
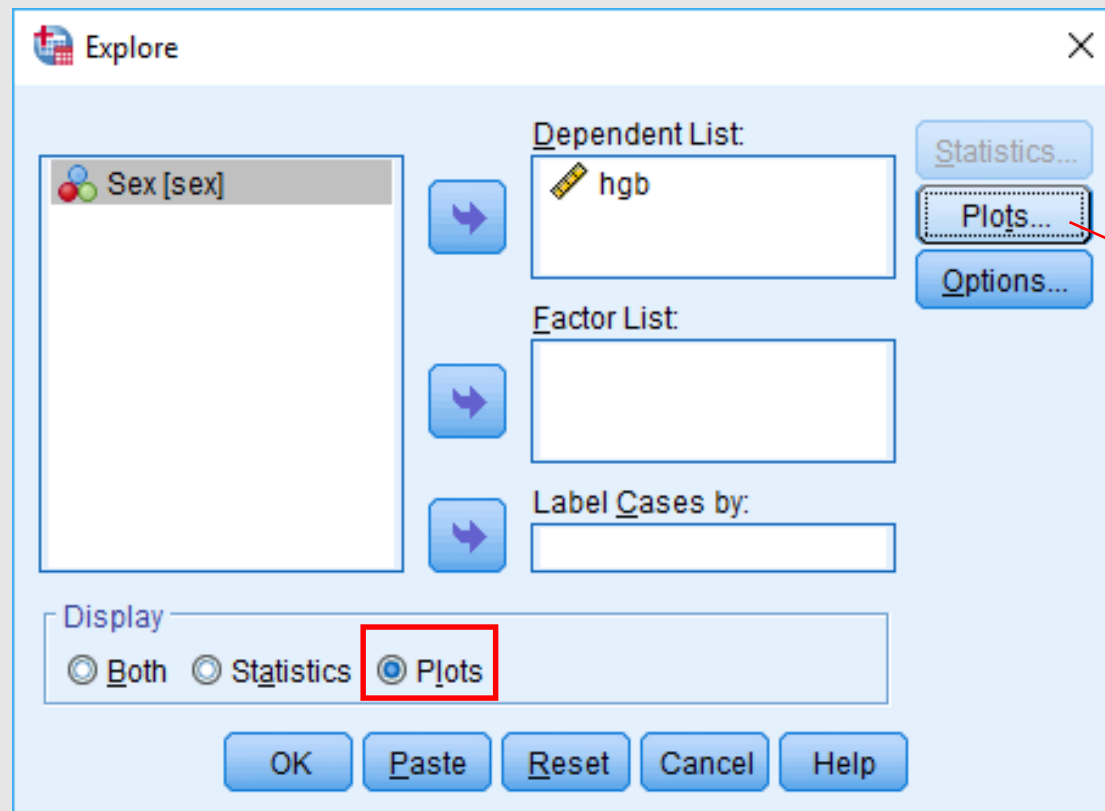
15.10	2	10.0	10.0	100.0
Total	20	100.0	100.0	

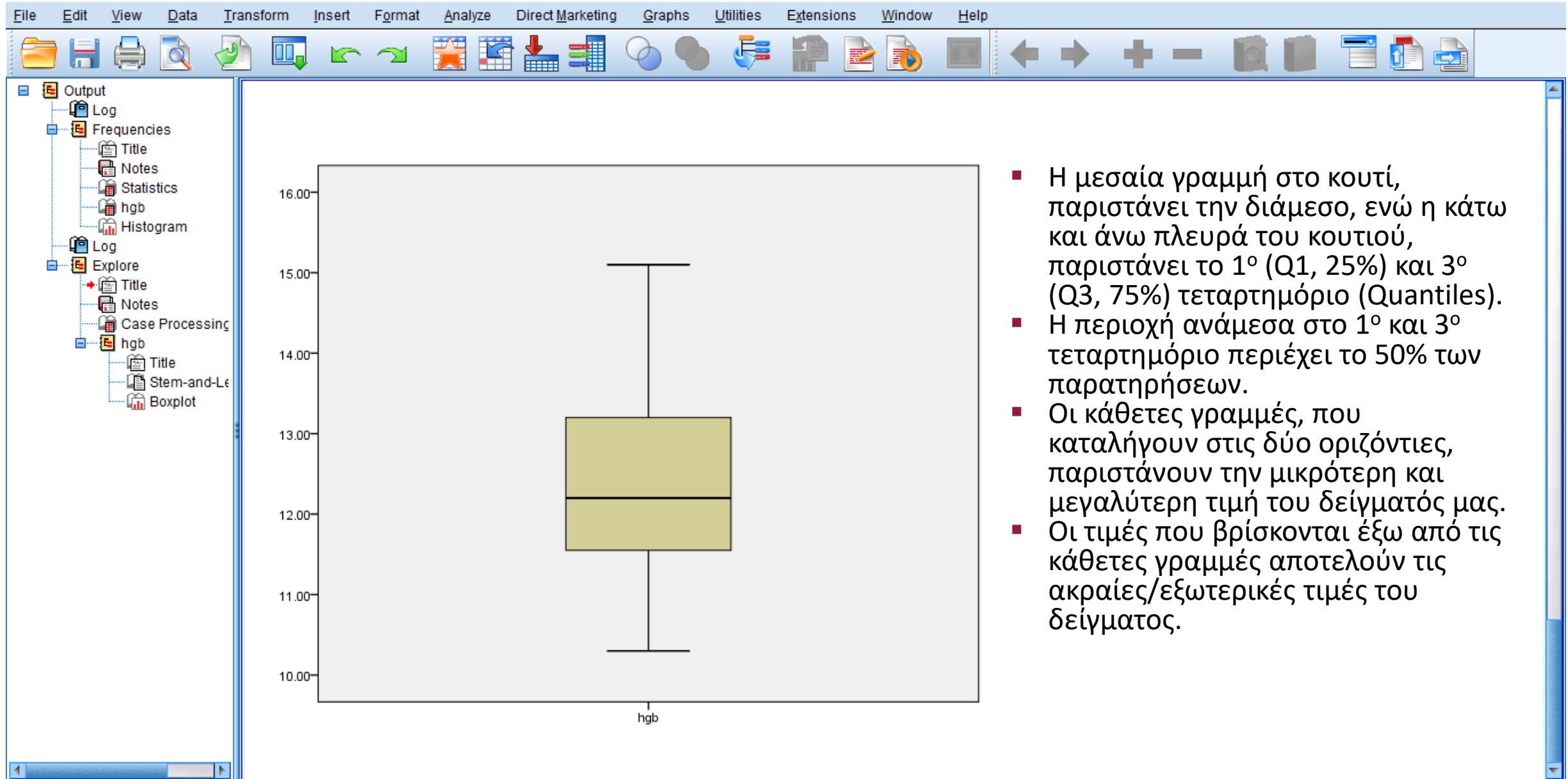




Box Plot

Επιλέγουμε **Analyze -> Descriptive Statistics -> Explore** και στην συνέχεια **Plots**



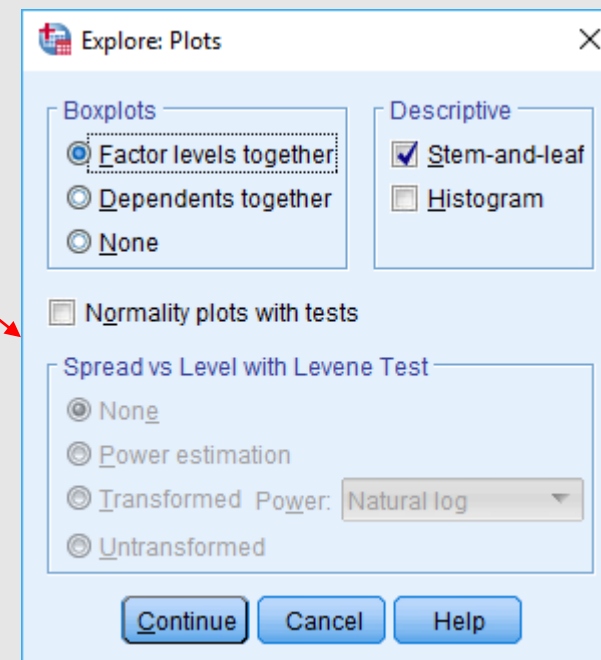
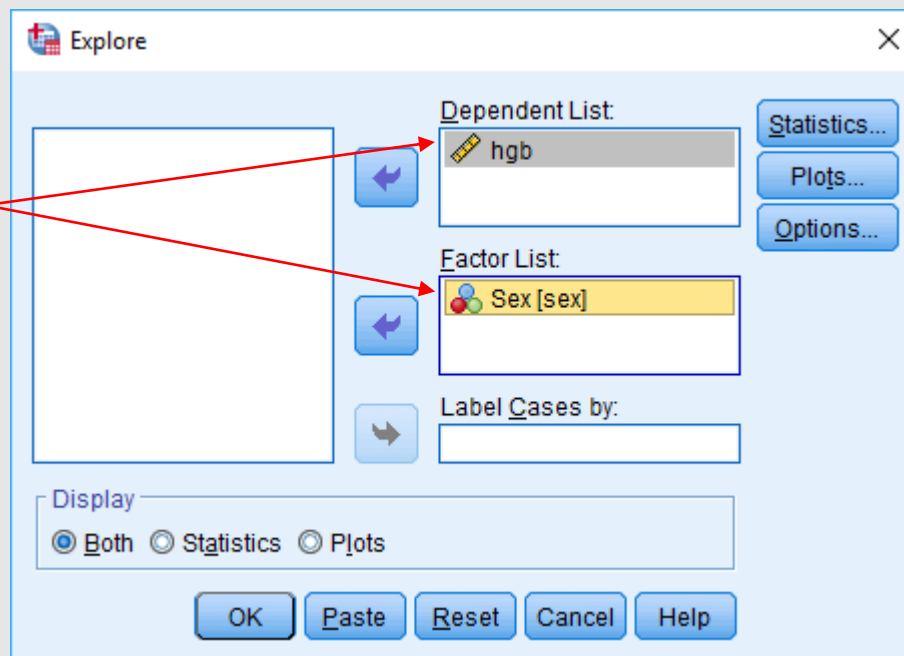
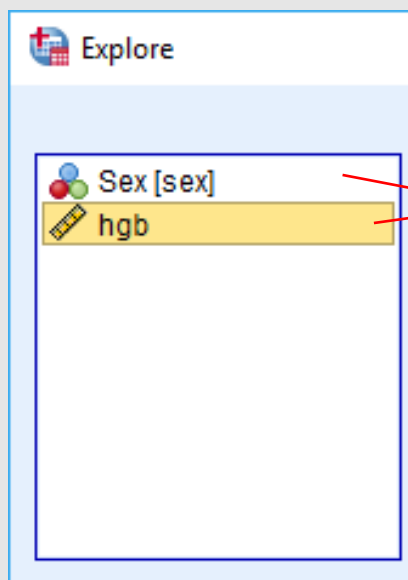


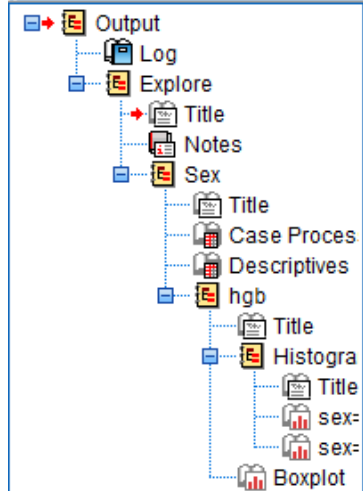
- Η μεσαία γραμμή στο κουτί, παριστάνει την διάμεσο, ενώ η κάτω και άνω πλευρά του κουτιού, παριστάνει το 1^ο (Q1, 25%) και 3^ο (Q3, 75%) τεταρτημόριο (Quantiles).
- Η περιοχή ανάμεσα στο 1^ο και 3^ο τεταρτημόριο περιέχει το 50% των παρατηρήσεων.
- Οι κάθετες γραμμές, που καταλήγουν στις δύο οριζόντιες, παριστάνουν την μικρότερη και μεγαλύτερη τιμή του δείγματός μας.
- Οι τιμές που βρίσκονται έξω από τις κάθετες γραμμές αποτελούν τις ακραίες/εξωτερικές τιμές του δείγματος.



Ανάλυση δεδομένων κατά ομάδες

Επιλέγουμε **Analyze -> Descriptive Statistics -> Explore**



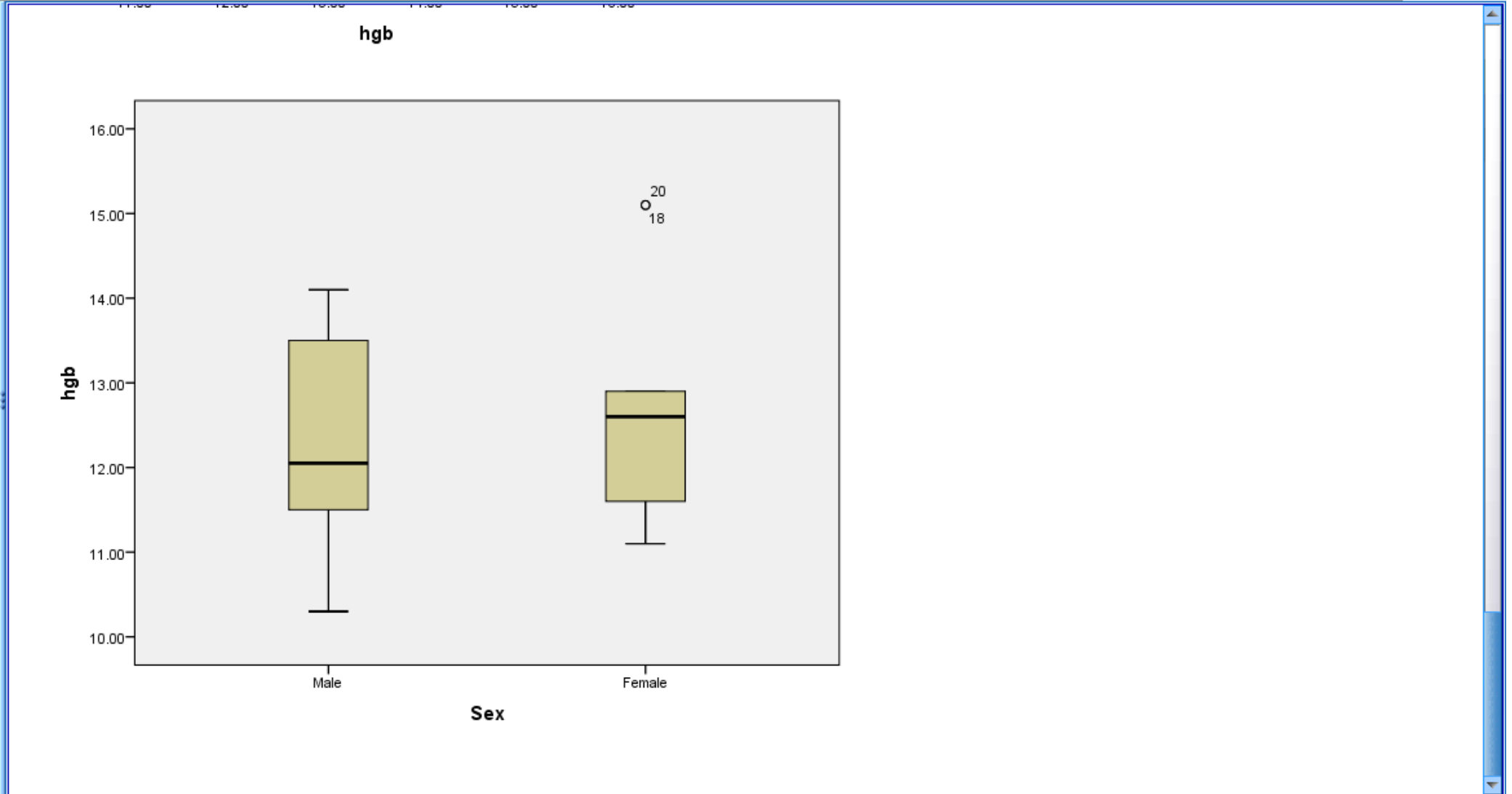
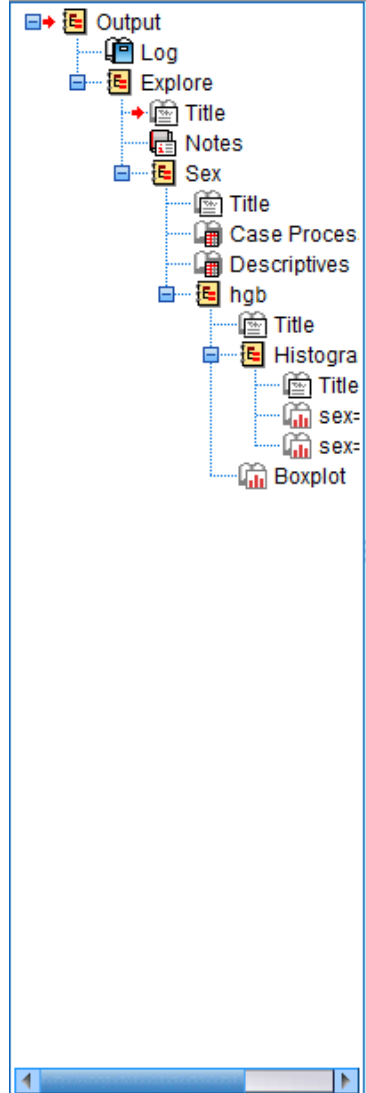


hgb	Male	10	100.0%	0	0.0%	10	100.0%
	Female	10	100.0%	0	0.0%	10	100.0%

Descriptives

Sex		Statistic		Std. Error
hgb	Male	Mean	12.2300	.38645
		95% Confidence Interval for Mean	Lower Bound	11.3558
			Upper Bound	13.1042
		5% Trimmed Mean	12.2333	
		Median	12.0500	
		Variance	1.493	
		Std. Deviation	1.22207	
		Minimum	10.30	
		Maximum	14.10	
		Range	3.80	
		Interquartile Range	2.20	
		Skewness	.048	.687
		Kurtosis	-.748	1.334
	Female	Mean	12.7400	.44227
		95% Confidence Interval for Mean	Lower Bound	11.7395
			Upper Bound	13.7405
		5% Trimmed Mean	12.7000	
		Median	12.6000	
		Variance	1.956	
		Std. Deviation	1.39857	
		Minimum	11.10	
		Maximum	15.10	

 12.23 ± 1.22 12.74 ± 1.40





Ανάλυση δύο κατηγορικών μεταβλητών

- Με την εντολή **crosstabs** παρουσιάζεται σε μορφή πίνακα η σχέση δύο ποιοτικών μεταβλητών
 - Τέτοιες είναι οι μεταβλητές που έχουν διακριτές κατηγορίες, δηλ. αυτές χωρίς καμιά σειρά κατάταξης/ιεράρχησης, όπως είναι το φύλο (γυναίκα, άνδρας), απαντήσεις (ναι, όχι) κ.α.
- Μπορεί να χρησιμοποιηθεί και για συνεχή δεδομένα μόνο όταν αυτά είναι ομαδοποιημένα σε επιμέρους κατηγορίες, όπως η ηλικία (0-19, 20-39, 40-59, 60-79, 80-99)
- Μπορεί να παρουσιαστεί η σχέση δύο μεταβλητών με βάση κάποια τρίτη μεταβλητή (διαστρωμάτωση)
- Με την εντολή **crosstabs** εκτελείται και η γνωστή **χ^2 δοκιμασία (chi-square test)**



Visible: 3 of 3 Variables

	section	gender	ethnicit	var
1	1	1	2	
2	1	1	2	
3	1	1	2	
4	1	1	2	
5	1	1	2	
6	1	1	2	
7	1	1	3	
8	1	1	3	
9	1	1	4	
10	1	1	4	
11	1	1	4	
12	1	1	4	
13	1	1	4	
14	1	1	4	
15	1	1	4	
16	1	1	4	
17	1	1	4	
18	1	1	5	
19	1	1	5	
20	1	1	5	
21	1	2	2	
22	1	2	3	
23	1	2	3	
24	1	2	3	
25	1	2	3	
26	1	2	3	
27	1	2	4	

Value Labels

Value:

Label:

Spelling...

Add
Change
Remove

1 = "Female"
2 = "Male"

OK Cancel Help

Value Labels

Value:

Label:

Spelling...

Add
Change
Remove

1 = "American/Indian"
2 = "Asian"
3 = "Afro-American"
4 = "Caucasian"
5 = "Hispanic"

OK Cancel Help

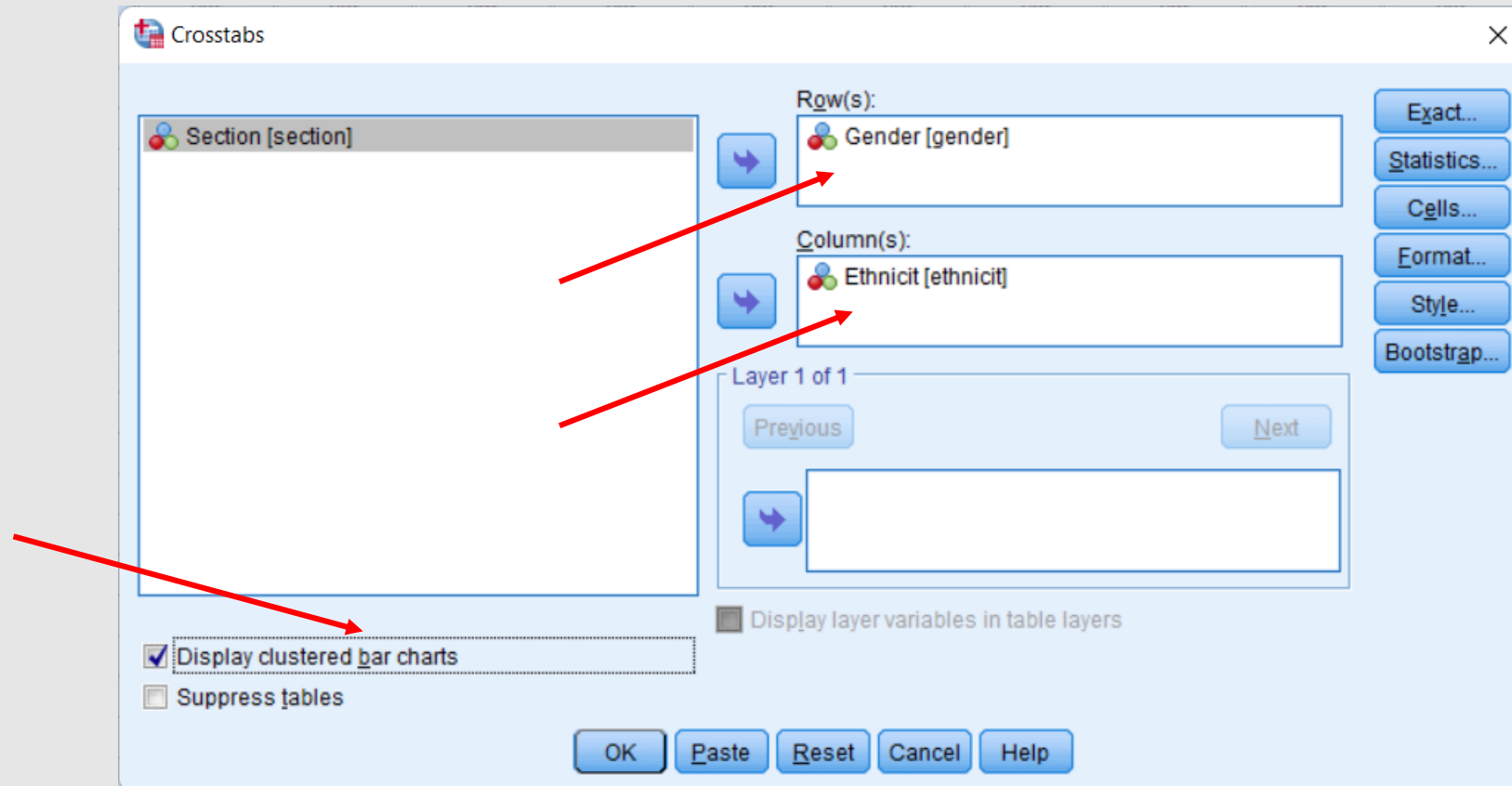


	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	section	Numeric	8	0	Section	None	None	8	Right	Nominal	Input
2	gender	Numeric	8	0	Gender	{1, Female}...	None	8	Right	Nominal	Input
3	ethnicit	Numeric	8	0	Ethnicit	{1, America...	None	8	Right	Nominal	Input



Ανάλυση δύο κατηγορικών μεταβλητών, χωρίς διαστρωμάτωση

Επιλέγουμε **Analyze -> Descriptive Statistics -> Crosstabs** (crosstabs.sav)



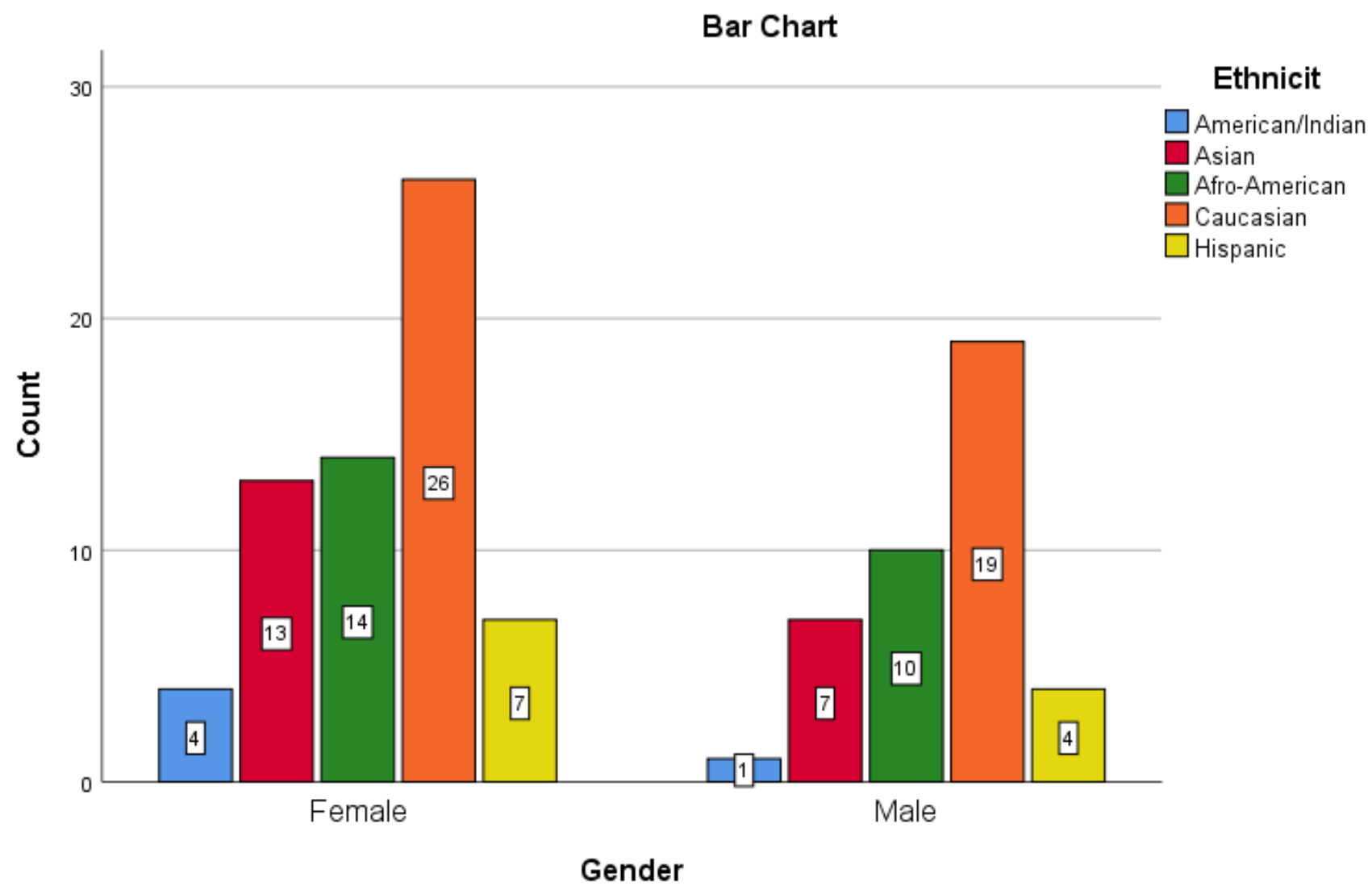
Case Processing Summary

	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender * Ethnicit	105	100.0%	0	0.0%	105	100.0%

Gender * Ethnicit Crosstabulation

Count

		Ethnicit					Total
		American/Indian	Asian	Afro-American	Caucasian	Hispanic	
Gender	Female	4	13	14	26	7	64
	Male	1	7	10	19	4	41
Total		5	20	24	45	11	105





Ανάλυση δύο κατηγορικών μεταβλητών, με διαστρωμάτωση

IBM SPSS Statistics Data Editor window titled "crosstabs.sav [DataSet1]". The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Extensions, Window, and Help. The toolbar contains icons for file operations, data manipulation, and analysis. The status bar at the bottom indicates "IBM SPSS Statistics Processor is ready" and "Unicode: ON".

Visible: 3 of 3 Variables

	section	gender	ethnicit	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1	1	Female	Asian															
2	1	Female	Asian															
3	1	Female	Asian															
4	1	Female	Asian															
5	1	Female	Asian															
6	1	Female	Asian															
7	1	Female	Afro-Ameri...															
8	1	Female	Afro-Ameri...															
9	1	Female	Caucasian															
10	1	Female	Caucasian															
11	1	Female	Caucasian															
12	1	Female	Caucasian															
13	1	Female	Caucasian															
14	1	Female	Caucasian															
15	1	Female	Caucasian															
16	1	Female	Caucasian															
17	1	Female	Caucasian															
18	1	Female	Hispanic															
19	1	Female	Hispanic															
20	1	Female	Hispanic															
21	1	Male	Asian															
22	1	Male	Afro-Ameri...															
23	1	Male	Afro-Ameri...															
24	1	Male	Afro-Ameri...															
25	1	Male	Afro-Ameri...															
26	1	Male	Afro-Ameri...															
27	1	Male	Caucasian															

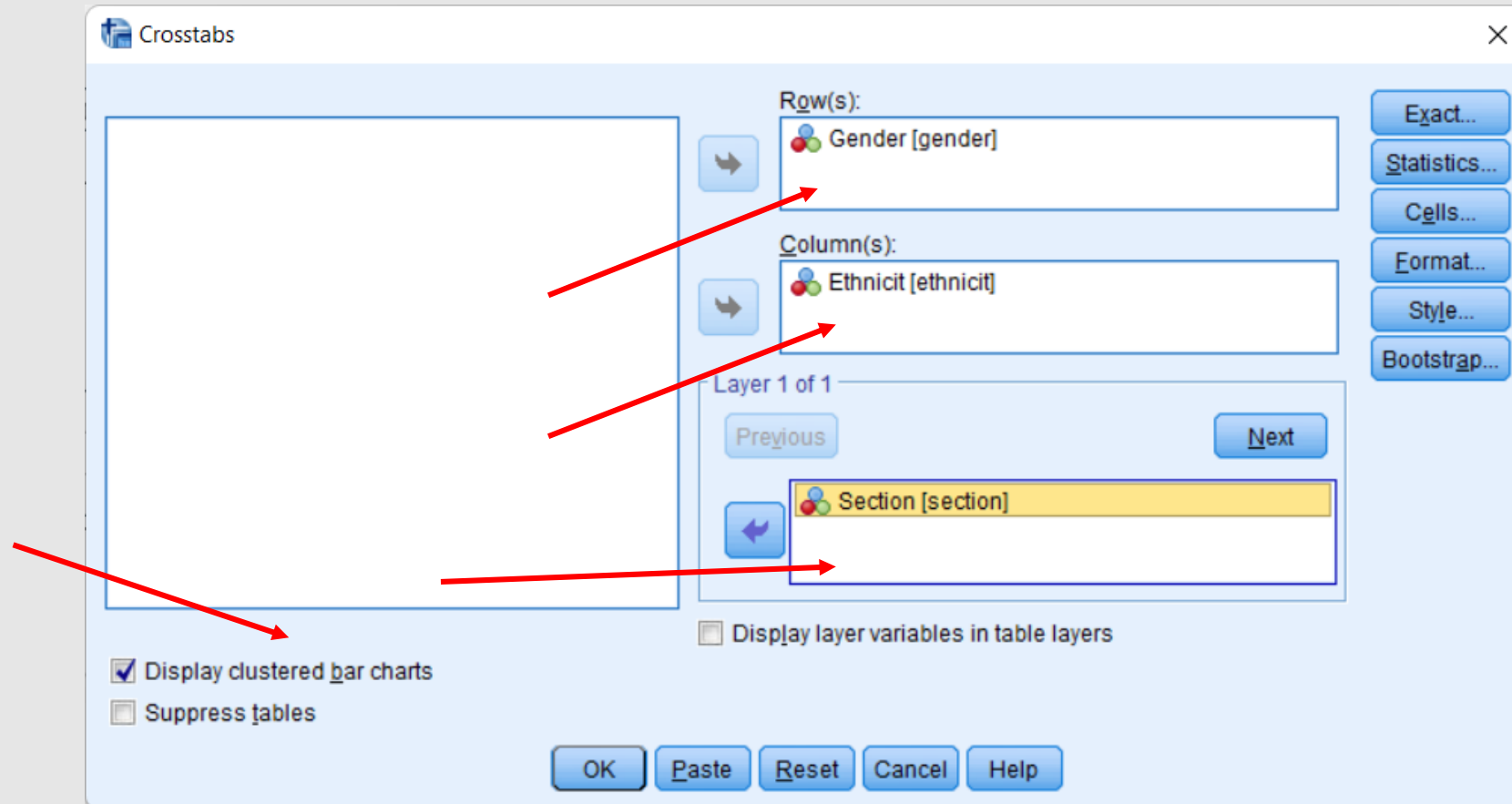
Data View Variable View

IBM SPSS Statistics Processor is ready Unicode: ON



Ανάλυση δύο κατηγορικών μεταβλητών, με διαστρωμάτωση

Επιλέγουμε **Analyze -> Descriptive Statistics -> Crosstabs** (crosstabs.sav)



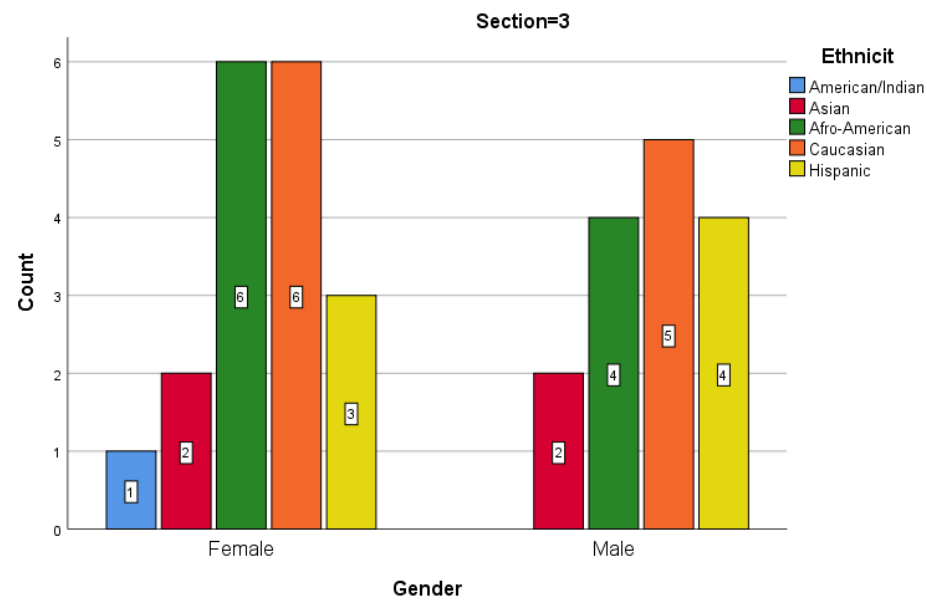
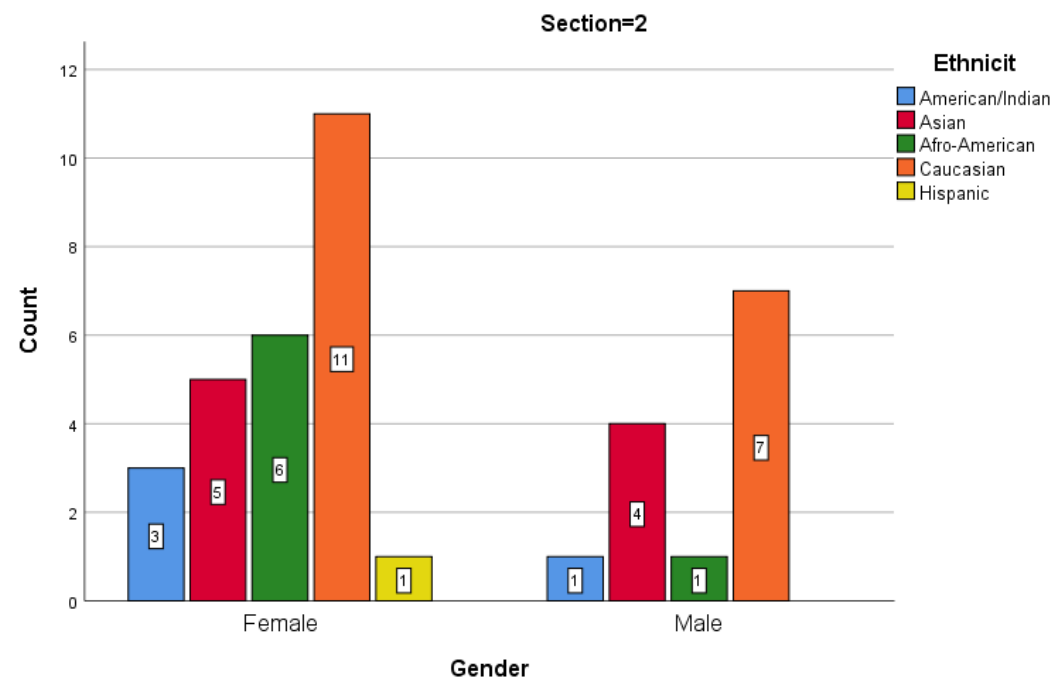
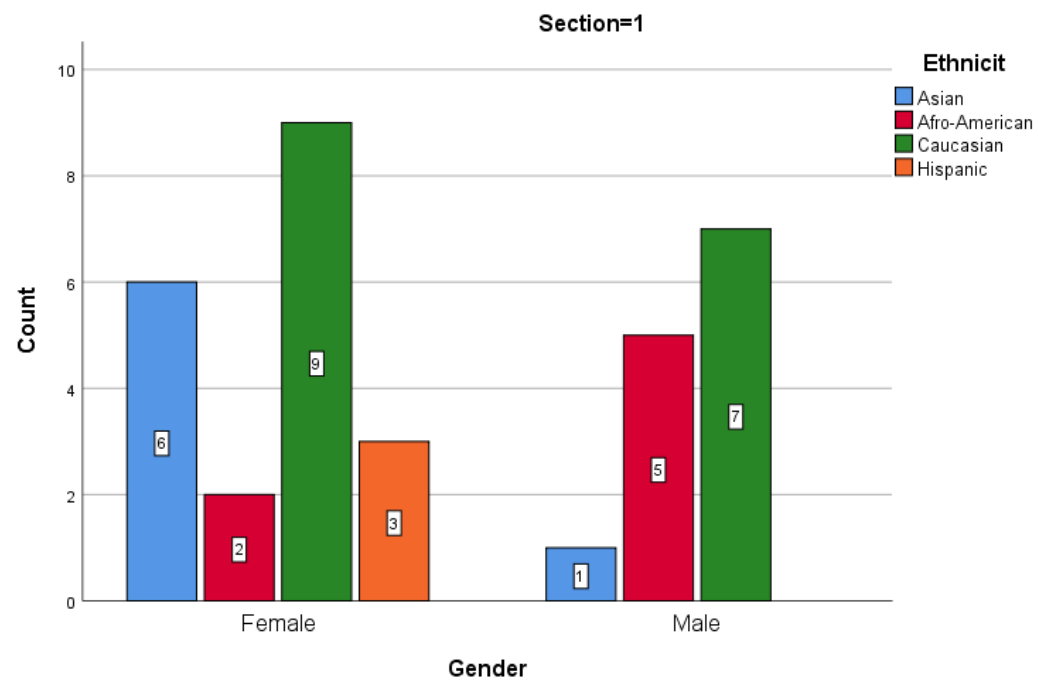
Case Processing Summary

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender * Ethnicit * Section	105	100.0%	0	0.0%	105	100.0%

Gender * Ethnicit * Section Crosstabulation

Count

Section			Ethnicit					Total
			American/Indi an	Asian	Afro-American	Caucasian	Hispanic	
1	Gender	Female		6	2	9	3	20
		Male		1	5	7	0	13
	Total			7	7	16	3	33
2	Gender	Female	3	5	6	11	1	26
		Male	1	4	1	7	0	13
	Total		4	9	7	18	1	39
3	Gender	Female	1	2	6	6	3	18
		Male	0	2	4	5	4	15
	Total		1	4	10	11	7	33
Total	Gender	Female	4	13	14	26	7	64
		Male	1	7	10	19	4	41
	Total		5	20	24	45	11	105





Scatter Plot

Το βάρος του σώματος και ο όγκος πλάσματος 8 ανδρών είναι:

A/A	βάρος Kg (x)	όγκος lt (y)
1	58.0	2.75
2	70.0	2.86
3	74.0	3.37
4	63.5	2.76
5	62.0	2.62
6	70.5	3.49
7	71.0	3.05
8	66.0	3.12

Θέλουμε να κατασκευάσουμε ένα **Scatter plot** για να δούμε αν υπάρχει σχέση μεταξύ όγκου πλάσματος και βάρος σώματος.

Καταχωρήστε τους όγκους πλάσματος και βάρος σώματος από τα 8 άτομα, και δώστε στη μία μεταβλητή το όνομα **volume** και στην άλλη το όνομα **weight**



Scatter Plot

Data View

*spss_data.sav [DataSet1] - IBM SPSS Statistics

File Edit View Data Transform

1 : volume 58.00

	weight	volume
1	58.00	2.75
2	70.00	2.86
3	74.00	3.37
4	63.50	2.76
5	62.00	2.62
6	70.50	3.49
7	71.00	3.05
8	66.00	3.12

Variable View

*spss_data.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	weight	Numeric	8	2	Body weight	None	None	8	Right	Scale	Input
2	volume	Numeric	8	2	Plasma volume	None	None	8	Right	Scale	Input

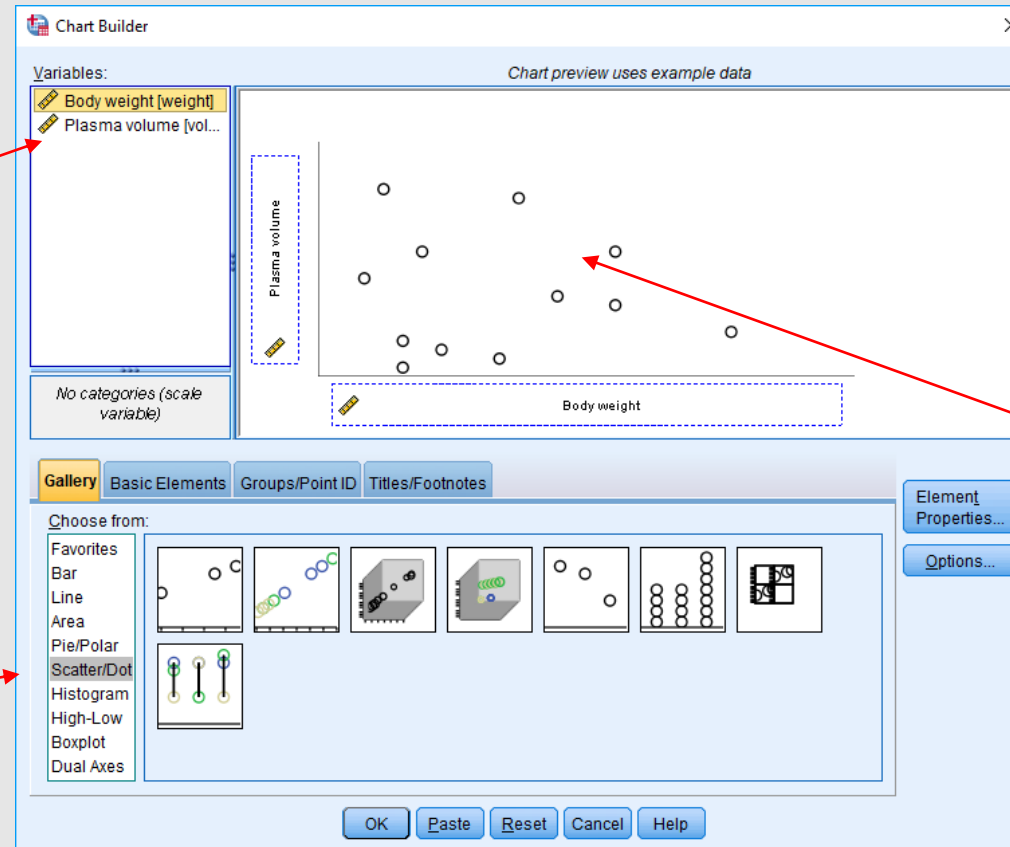


Scatter Plot

Επιλέγουμε **Graphs -> Chart Builder** και στην συνέχεια **Charts**

3. Με Drag & Drop
σέρνουμε τις μεταβλητές
στους άξονες

1. Επιλέγουμε Scatter/Dot

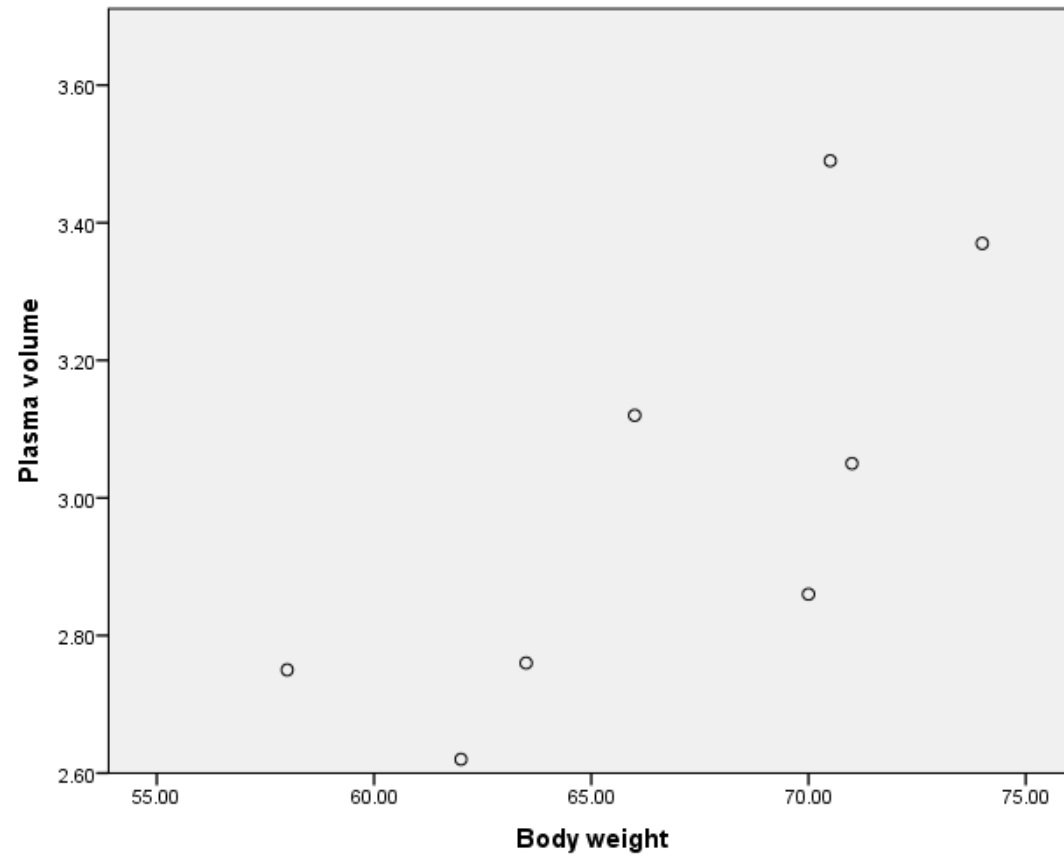


2. Με Drag & Drop
σέρνουμε το διάγραμμα
στο πάνω μέρος



- Output
 - Log
 - GGraph
 - Title
 - Notes
 - Graph

→ **GGraph**





Περιγραφική Στατιστική

Παράδειγμα (data2.sav)

Έστω ο διπλανός πίνακας που περιέχει τις μεταβλητές

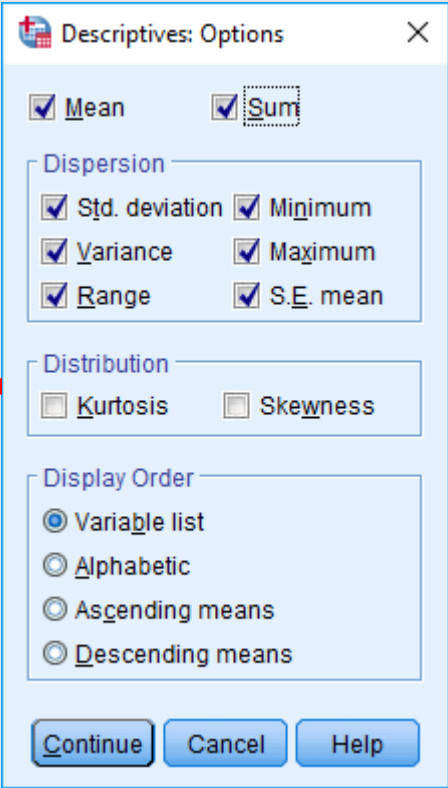
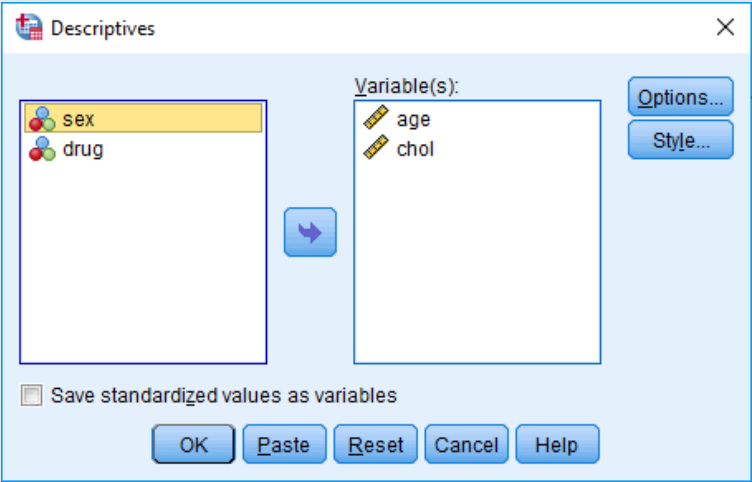
- chol (δείκτη χοληστερίνης)
- sex (φύλο)
- age (ηλικία)
- drug (ποσότητα ενός ληφθέντος φαρμάκου)
 - 0: καθόλου
 - 1: ελάχιστη
 - 2: μικρή
 - 3: μέτρια

	chol	sex	age	drug
1	161	0	31	0
2	163	0	.	0
3	169	0	39	0
4	169	1	41	1
5	170	1	35	0
6	.	0	31	2
7	174	1	33	1
8	176	1	28	0
9	.	0	.	3
10	195	0	41	1
11	233	1	54	0
12	234	0	53	1
13	239	0	44	3
14	244	1	50	1
15	.	0	52	2
16	249	0	47	2
17	.	1	49	2
18	256	0	46	1
19	.	1	63	3
20	258	1	64	3



Descriptive

Επιλέγουμε από το μενού **Analyze -> Descriptive Statistics -> Descriptive** και κάνουμε τις ρυθμίσεις όπως στις διπλανές εικόνες:



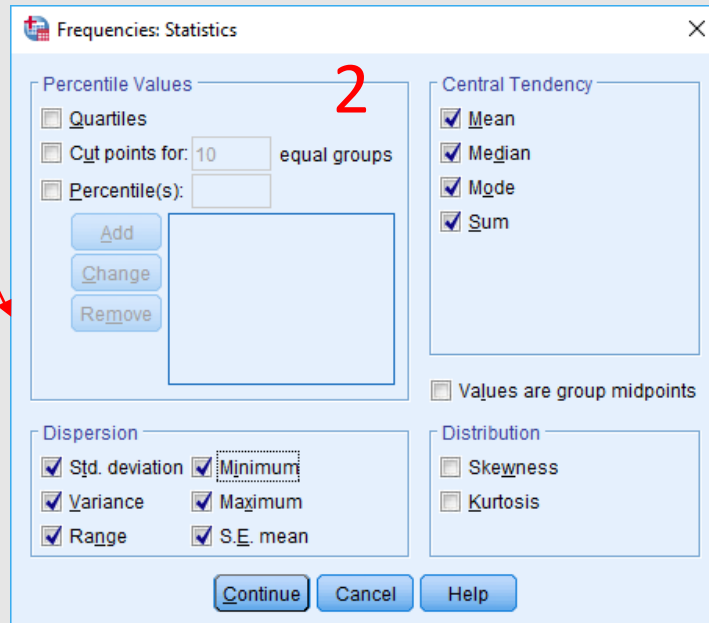
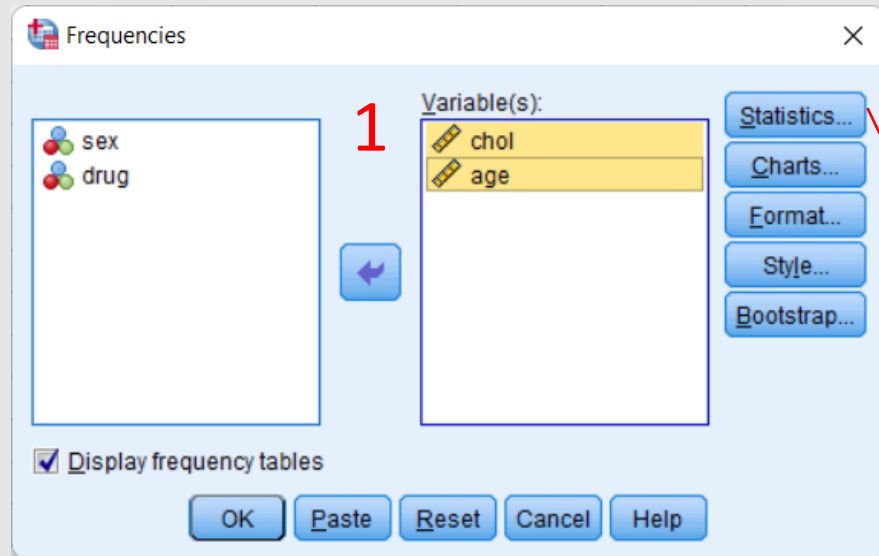
Πατάμε **Continue** και στην συνέχεια **OK** οπότε εμφανίζεται ο ακόλουθος στατιστικός πίνακας

Descriptive Statistics								
	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
chol	15	97	161	258	206.00	10.012	38.778	1503.714
age	18	36	28	64	44.50	2.501	10.612	112.618
Valid N (listwise)	14							



Frequencies

Επιλέγουμε από το μενού **Analyze -> Descriptive Statistics -> Frequencies** και κάνουμε τις ρυθμίσεις όπως στις εικόνες (1) και (2):



Statistics		
	chol	age
N	Valid 15	18
	Missing 5	2
Mean	206.00	44.50
Std. Error of Mean	10.012	2.501
Median	195.00	45.00
Mode	169	31 ^a
Std. Deviation	38.778	10.612
Variance	1503.714	112.618
Range	97	36
Minimum	161	28
Maximum	258	64
Sum	3090	801

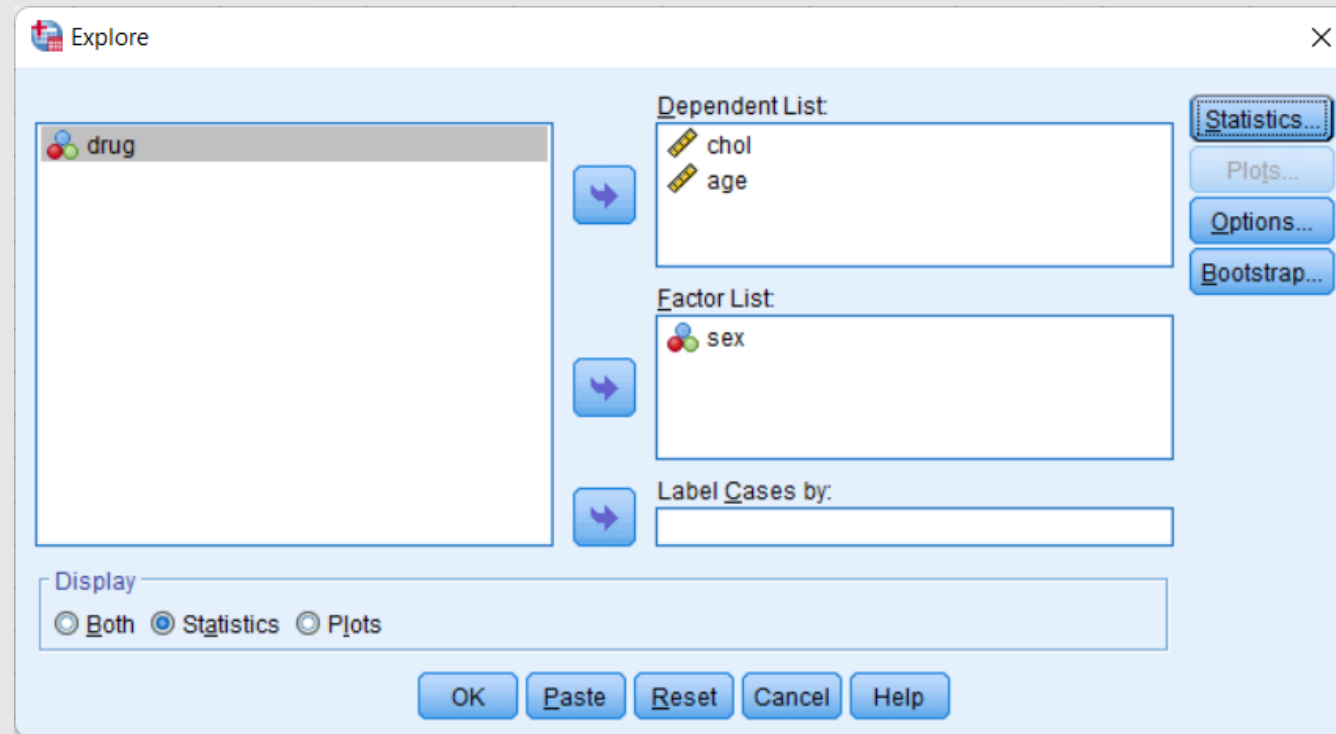
a. Multiple modes exist. The smallest value is shown

Πατάμε **Continue** και στην συνέχεια **OK** οπότε εμφανίζεται ο στατιστικός πίνακας (3)



Explore

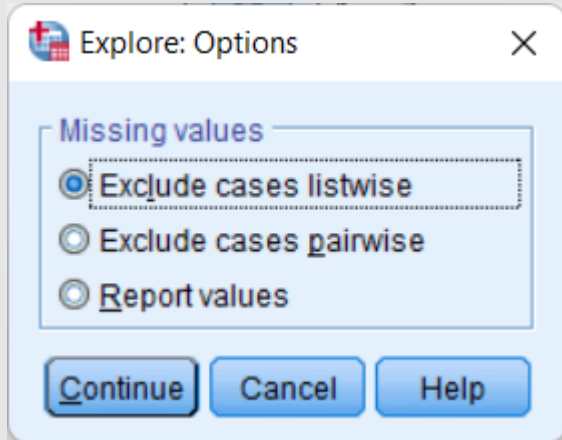
Επιλέγουμε από το μενού **Analyze -> Descriptive Statistics -> Explore**



Από το πλήκτρο **Options** γίνεται η διαχείριση των ελλειπουσών τιμών.
Πατώντας το κουμπί εμφανίζεται το παράθυρο της επόμενης διαφάνειας



Explore



Exclude cases listwise

- Στα περιγραφικά μέτρα που θα προκύψουν χρησιμοποιούνται μόνο οι περιπτώσεις που είναι **ταυτόχρονα έγκυρες** σε όλες τις μεταβλητές που έχουν εισαχθεί στις Dependent List και Factor List

Exclude cases pairwise

- Στα περιγραφικά μέτρα που θα προκύψουν χρησιμοποιούνται όλες οι περιπτώσεις που είναι **έγκυρες** για **κάθε** μεταβλητή που έχει εισαχθεί στην Dependent List

Report values

- Οι παρατηρήσεις με ελλείπουσες τιμές αναλύονται ως ξεχωριστή κατηγορία



Explore

Exclude cases listwise

Case Processing Summary						
	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
chol	14	70.0%	6	30.0%	20	100.0%
age	14	70.0%	6	30.0%	20	100.0%

Descriptives				
		Statistic		Std. Error
chol	Mean		209.07	10.236
	95% Confidence Interval for Mean	Lower Bound	186.96	
		Upper Bound	231.19	
	5% Trimmed Mean		209.02	
	Median		214.00	
	Variance		1466.995	
	Std. Deviation		38.301	
	Minimum		161	
	Maximum		258	
	Range		97	
	Interquartile Range		76	
age	Skewness		.002	.597
	Kurtosis		-2.038	1.154
	Mean		43.29	2.665
	95% Confidence Interval for Mean	Lower Bound	37.53	
		Upper Bound	49.04	
	5% Trimmed Mean		42.98	
	Median		42.50	
	Variance		99.451	
	Std. Deviation		9.972	
	Minimum		28	
	Maximum		64	
	Range		36	
	Interquartile Range		16	
	Skewness		.382	.597
	Kurtosis		-.102	1.154



Explore

Exclude cases pairwise

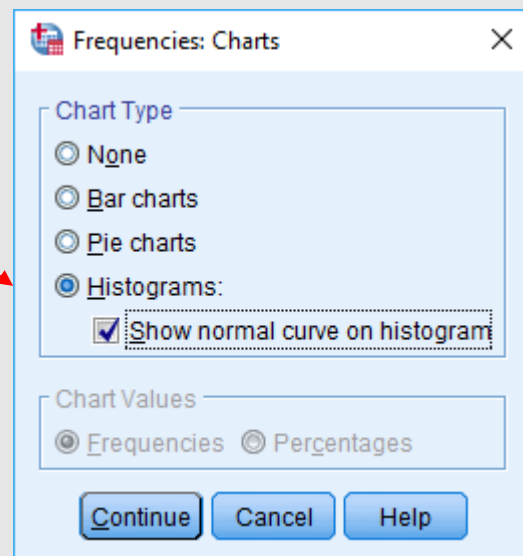
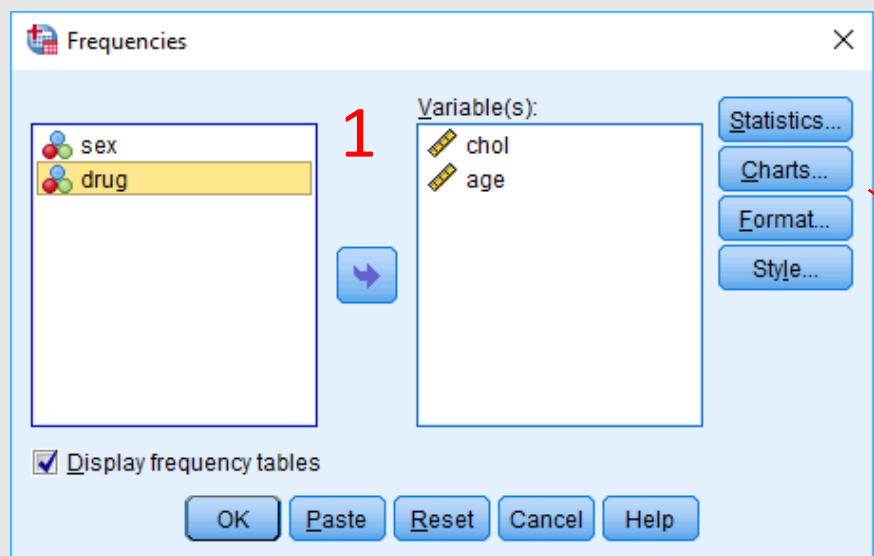
Case Processing Summary						
	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
chol	15	75.0%	5	25.0%	20	100.0%
age	18	90.0%	2	10.0%	20	100.0%

Descriptives				
		Statistic		Std. Error
chol	Mean		206.00	10.012
	95% Confidence Interval for Mean	Lower Bound	184.53	
		Upper Bound	227.47	
	5% Trimmed Mean		205.61	
	Median		195.00	
	Variance		1503.714	
	Std. Deviation		38.778	
	Minimum		161	
	Maximum		258	
	Range		97	
	Interquartile Range		75	
	Skewness		.138	.580
age	Kurtosis		-2.009	1.121
	Mean		44.50	2.501
	95% Confidence Interval for Mean	Lower Bound	39.22	
		Upper Bound	49.78	
	5% Trimmed Mean		44.33	
	Median		45.00	
	Variance		112.618	
	Std. Deviation		10.612	
	Minimum		28	
	Maximum		64	
	Range		36	
	Interquartile Range		18	
	Skewness		.214	.536
	Kurtosis		-.658	1.038

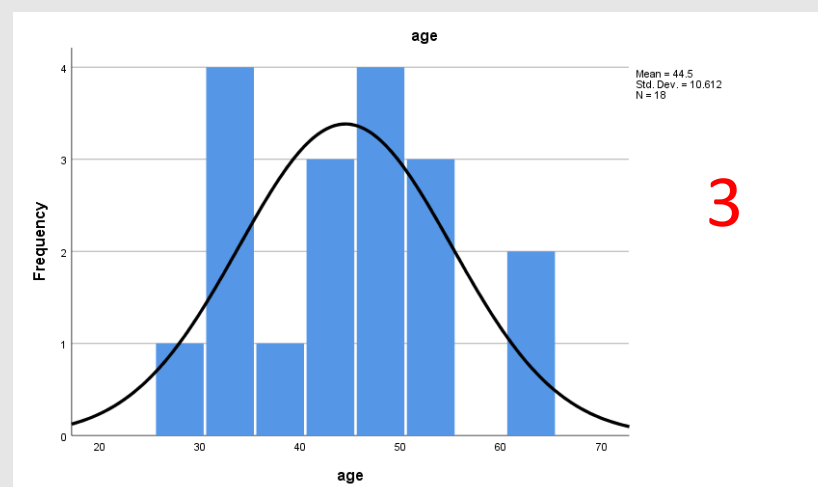
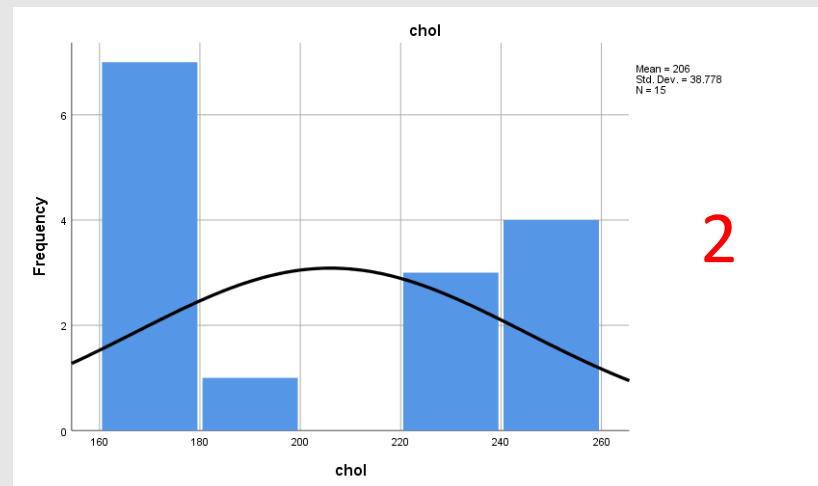


Ιστόγραμμα

Επιλέγουμε από το μενού **Analyze -> Descriptive Statistics -> Frequencies** και κάνουμε τις ρυθμίσεις όπως στην διπλανή εικόνα (1):



Πατάμε **Continue** και στην συνέχεια **OK** οπότε εμφανίζονται τα γραφήματα (2) και (3)



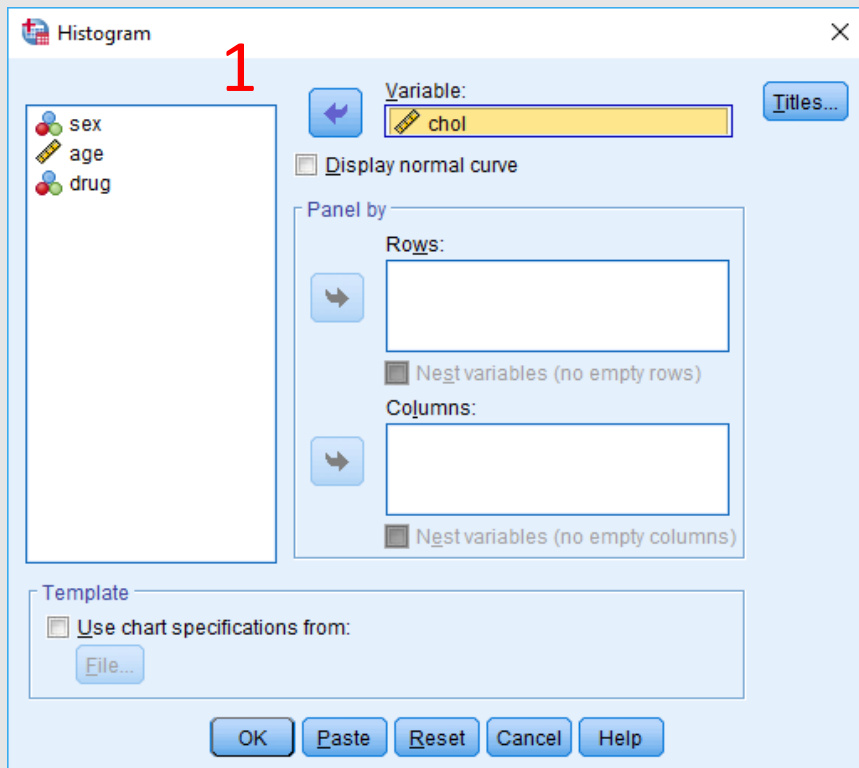
2

3

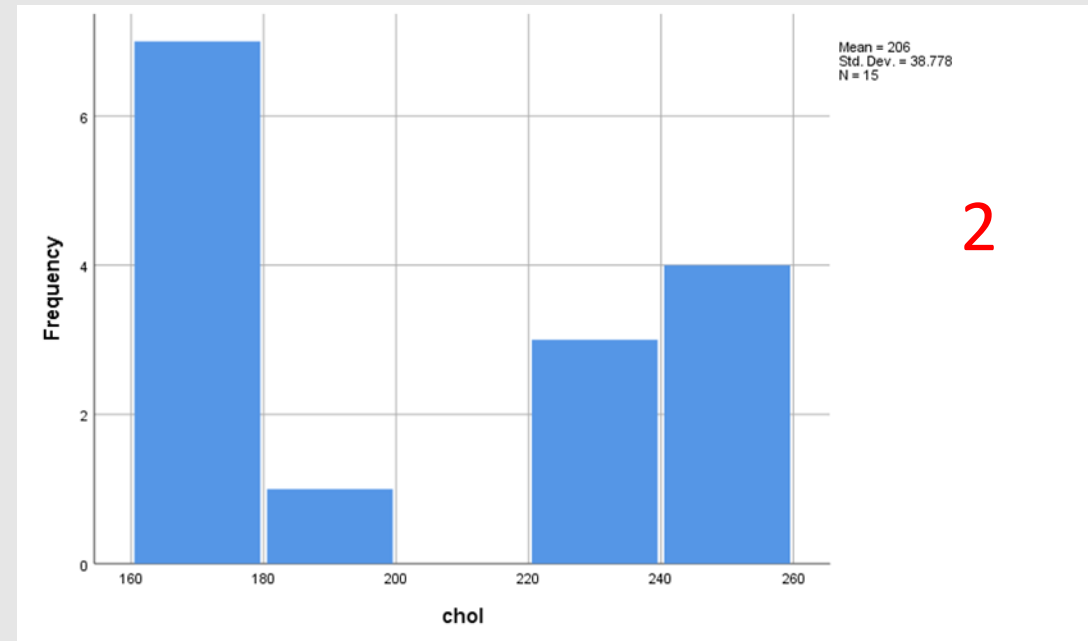


Ιστόγραμμα

Επιλέγουμε από το μενού **Graphs -> Legacy Dialogs -> Histogram** και κάνουμε τις ρυθμίσεις όπως στην παρακάτω εικόνα (1):



Πατάμε **OK** οπότε εμφανίζεται το ιστόγραμμα (2)

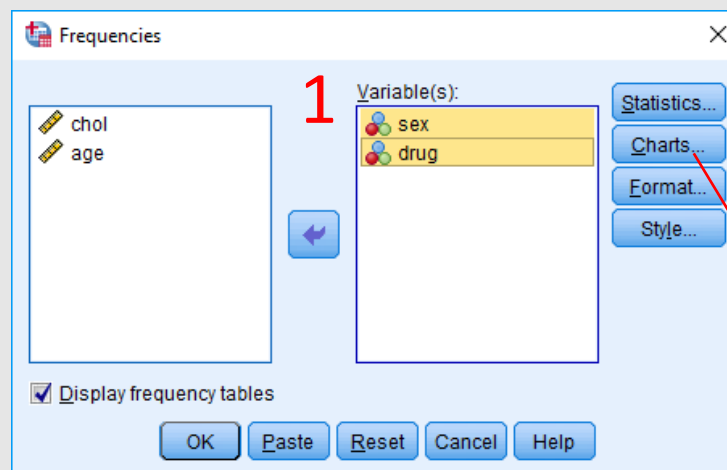
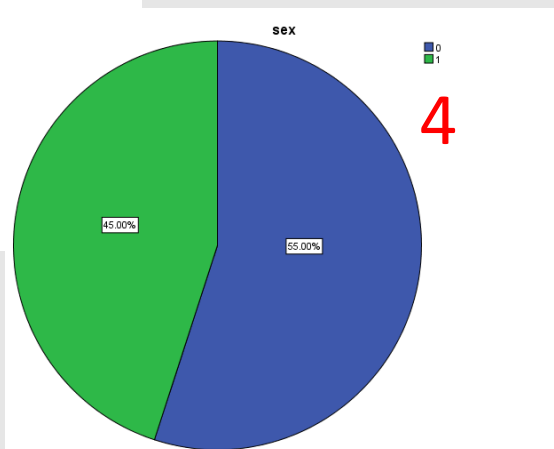
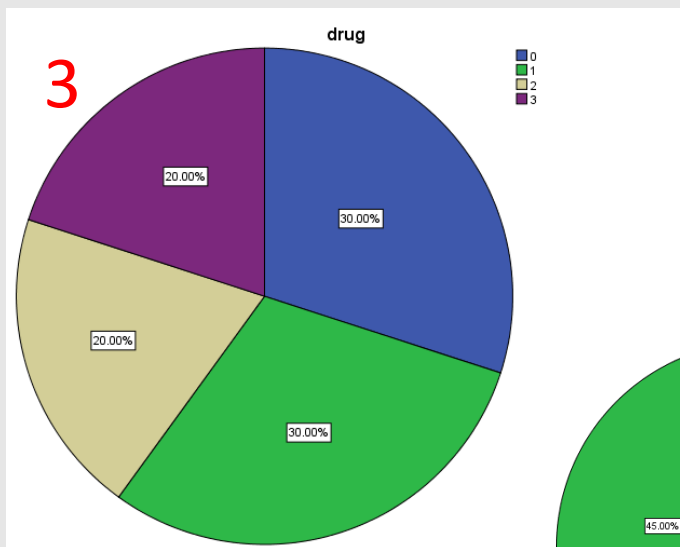


Σημείωση: Μπορούμε να αλλάξουμε διάφορα χαρακτηριστικά του ιστογράμματος, αλλά και κάθε γραφήματος, που παράγεται κάνοντας διπλό κλικ στο ιστόγραμμα και αλλάζοντας κατάλληλα διάφορες παραμέτρους (π. χ. το πλήθος των κλάσεων)

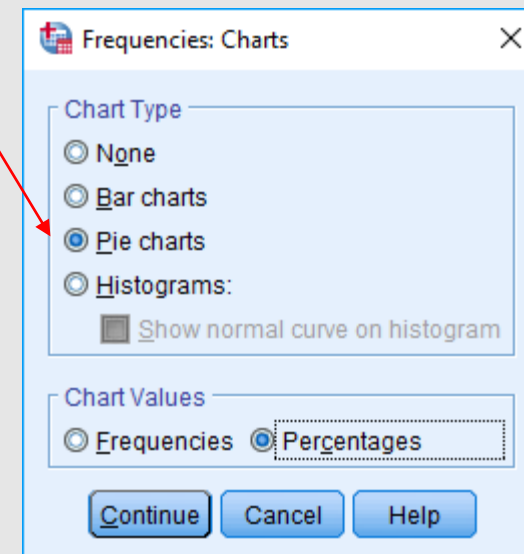


Pie chart

Επιλέγουμε από το μενού **Analyze -> Descriptive Statistics -> Frequencies** και κάνουμε τις ρυθμίσεις όπως στην διπλανή εικόνα (1):



Πατάμε **Continue** και στην συνέχεια **OK** οπότε εμφανίζονται τα γραφήματα (3) και (4)



Επιλογή περιπτώσεων (select cases)



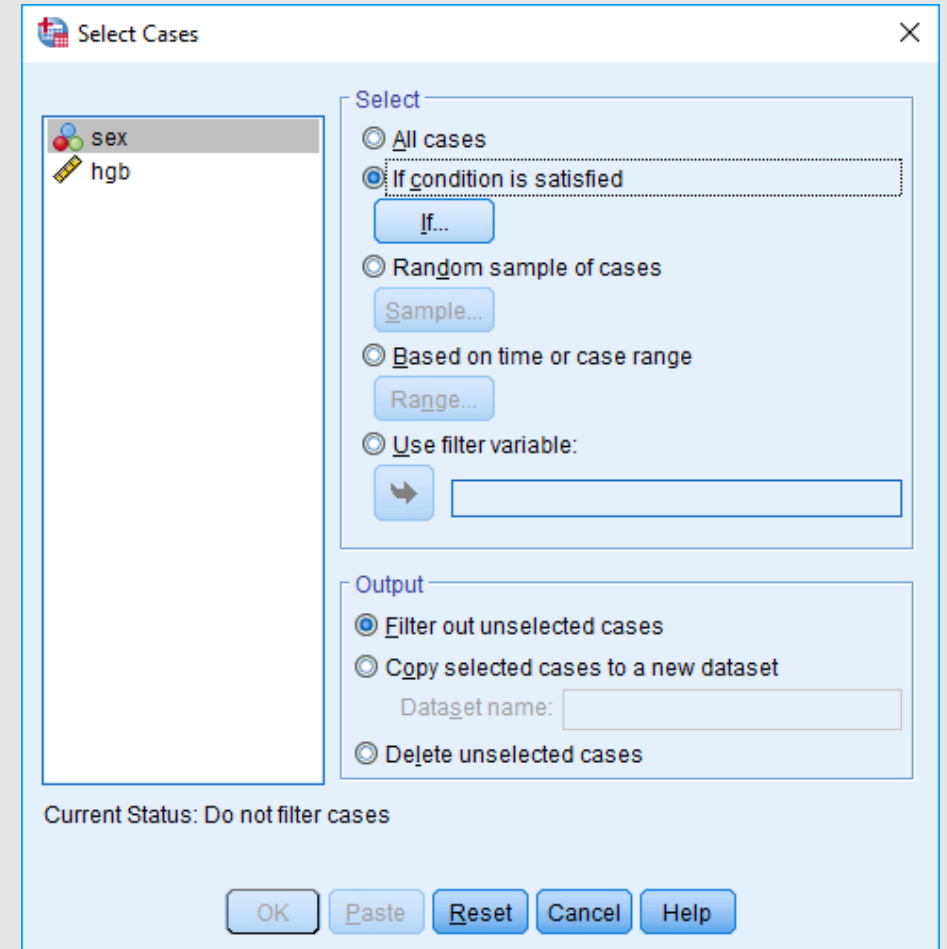
Επιλογή περιπτώσεων (select cases)

Έστω ότι στο παράδειγμα της διαφάνειας 10 θέλουμε να υπολογίσουμε τη μέση τιμή και την τυπική απόκλιση των επιπέδων αιμοσφαιρίνης μόνο για τις γυναίκες

Επιλέγουμε διαδοχικά **Data -> Select cases**

Στο παράθυρο που εμφανίζεται επιλέγουμε **If condition is satisfied**

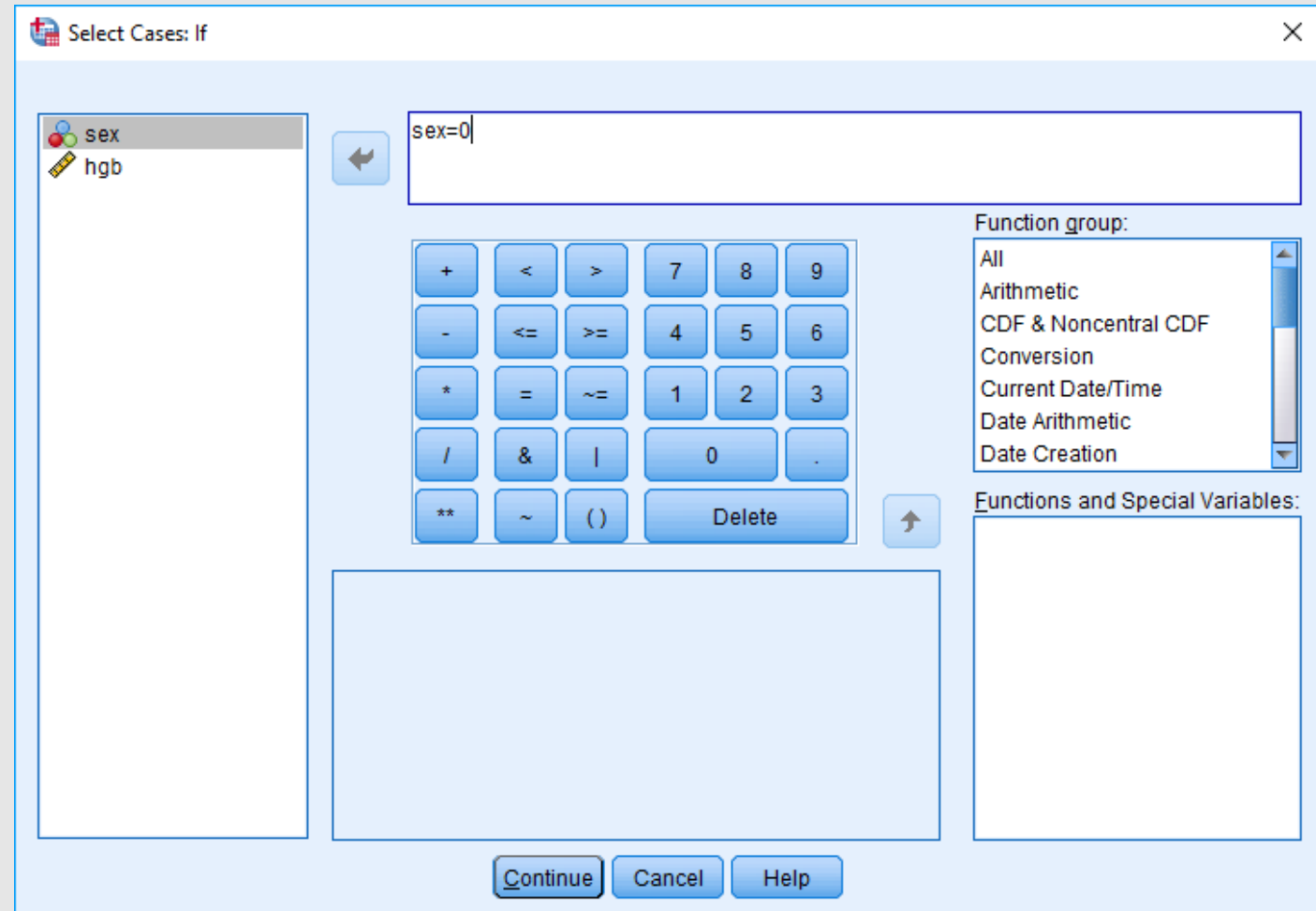
Κάνουμε κλικ στο κουμπί **if** οπότε ανοίγει το παράθυρο **Select Cases: if ...**





Επιλογή περιπτώσεων (select cases)

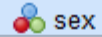


... εισάγουμε την συνθήκη **sex = 0**, πατάμε **Continue** και μετά **OK**





Επιλογή περιπτώσεων (select cases)

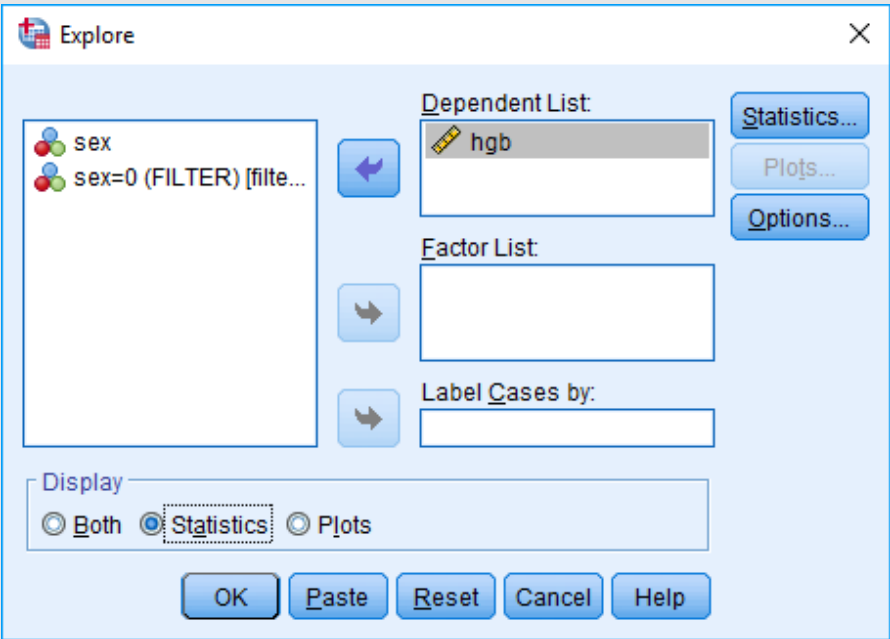
Με τον τρόπο αυτό οι ενέργειες που θα γίνουν στη συνέχεια θα εφαρμοστούν μόνο στις τιμές των οποίων οι αντίστοιχες τιμές της μεταβλητής **sex** είναι **0**

	 sex	 hgb	 filter_\$
1	0	10.00	1
2	0	10.30	1
3	0	11.50	1
4	1	11.60	0
5	1	11.10	0
6	1	11.40	0
7	0	12.00	1
8	0	12.00	1
9	0	12.10	1
10	1	12.10	0
11	1	12.90	0
12	1	12.90	0
13	1	12.30	0
14	0	12.40	1
15	0	13.50	1
16	0	13.60	1
17	0	14.10	1
18	1	15.10	0
19	1	12.90	0
20	1	15.10	0



Επιλογή περιπτώσεων (select cases)

Οπότε επιλέγοντας και πάλι διαδοχικά **Analyze -> Descriptive Statistics -> Explore** και στην συνέχεια **OK** εμφανίζεται το αποτέλεσμα της διπλανής εικόνας οι υπολογισμοί των οποίων έγιναν μόνο για τις περιπτώσεις όπου η τιμή της μεταβλητής **sex** είναι 0



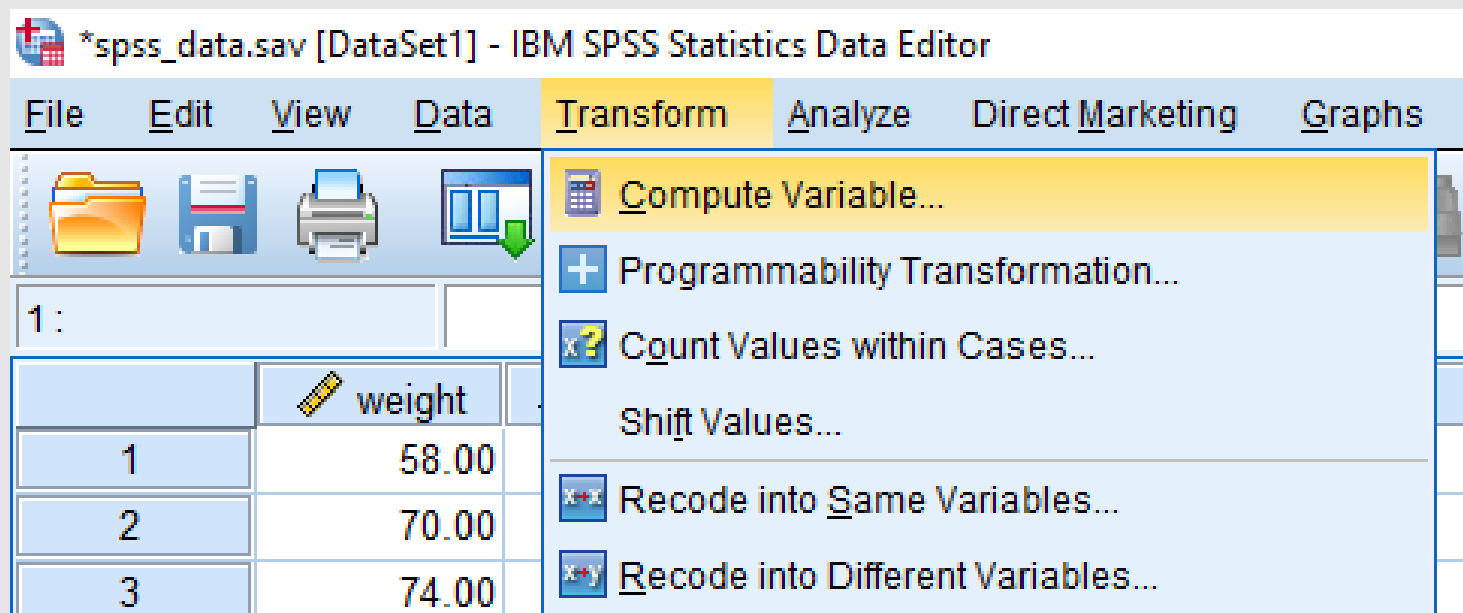
Case Processing Summary						
	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
hgb	10	100.0%	0	0.0%	10	100.0%

Descriptives			
hgb	Statistic	Std. Error	
		Mean	Std. Error
Mean	12.1500	.42564	
95% Confidence Interval for Mean	Lower Bound	11.1871	
	Upper Bound	13.1129	
5% Trimmed Mean	12.1611		
Median	12.0500		
Variance	1.812		
Std. Deviation	1.34598		
Minimum	10.00		
Maximum	14.10		
Range	4.10		
Interquartile Range	2.33		
Skewness	-.206	.687	
Kurtosis	-.651	1.334	



Μετασχηματισμός και τροποποίηση μεταβλητών

- Αν θέλουμε να τροποποιήσουμε/μετασχηματίσουμε μια μεταβλητή, είτε υπολογίζοντας τον λογάριθμο της είτε διαιρώντας την με ένα αριθμό, αυτό μπορούμε να το πετύχουμε, επιλέγοντας από το μενού επιλογών **Transform > Compute Variable**.
- Το SPSS έρχεται με μια μεγάλη λίστα συναρτήσεων που μπορούμε να εφαρμόσουμε σε μια μεταβλητή. Μας δίνει επίσης την δυνατότητα να δημιουργήσουμε και νέες κάνοντας διάφορους συνδυασμούς.

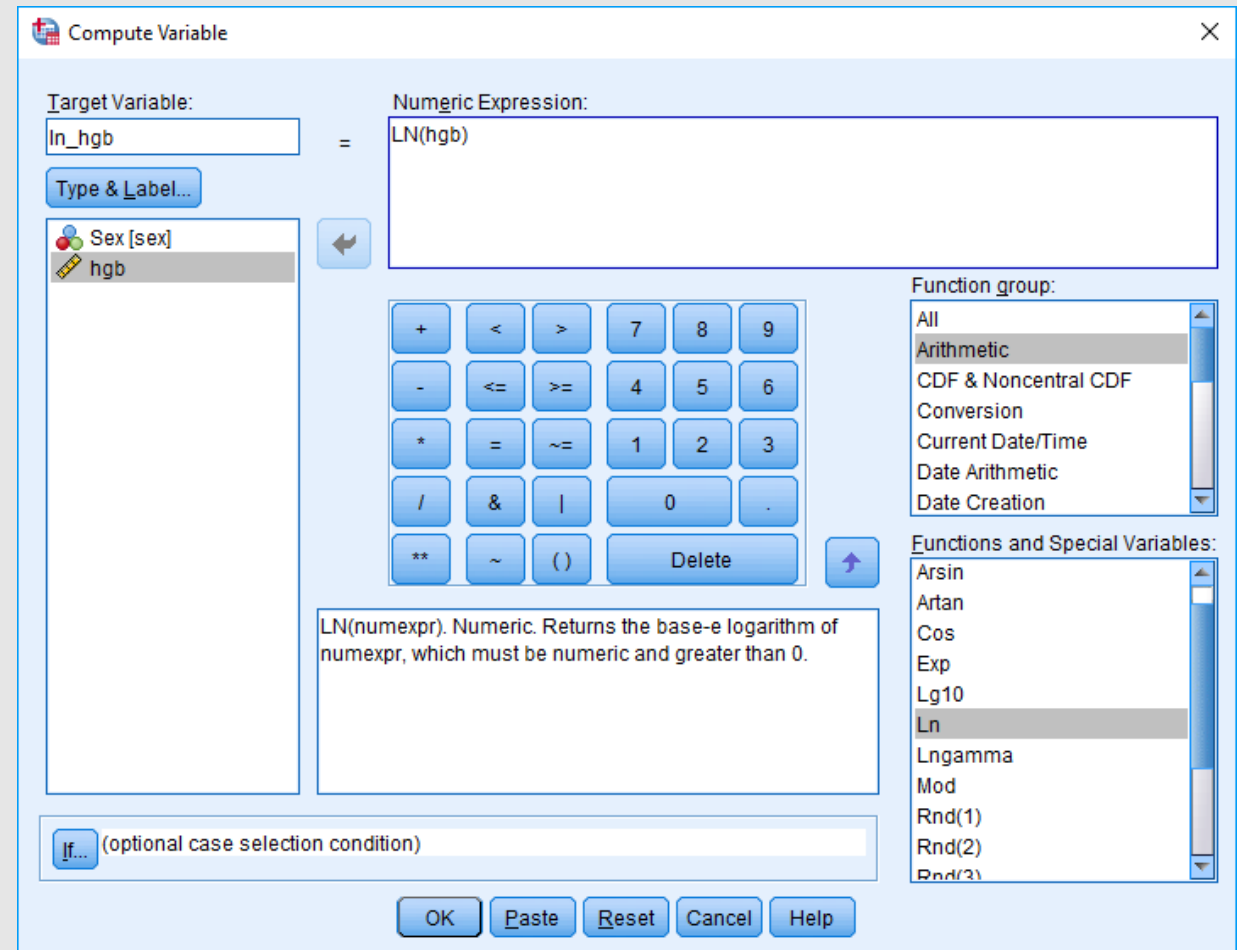




Μετασχηματισμός και τροποποίηση μεταβλητών

Για να μετασχηματίσουμε τα δεδομένα της μεταβλητής **hgb** σε φυσικούς λογαρίθμους, ακολουθούμε τα παρακάτω βήματα:

1. Στο πλαίσιο **Target Variable** ορίζουμε τη νέα μεταβλητή **ln_hgb**
2. Επιλέγουμε τη συνάρτηση **LN(numexpr)** και την σέρνουμε στο πλαίσιο **Numeric Expression**
3. Σέρνουμε την μεταβλητή **hgb** στο πλαίσιο **Numeric Expression**
4. Κάνουμε κλικ στο **OK**





Μετασχηματισμός και τροποποίηση μεταβλητών

Αυτόματα δημιουργείται στον **Data Editor** η μεταβλητή **ln_hbg** και τιμές τις ζητούμενες





*spss_data.sav [DataSet1] - IBM SPSS Statistics Data Editor

	sex	hgb	ln_hgb
1	0	10.80	2.38
2	0	10.30	2.33
3	0	11.50	2.44
4	1	11.60	2.45
5	1	11.10	2.41
6	1	11.40	2.43
7	0	12.00	2.48
8	0	12.00	2.48
9	0	12.10	2.49
10	1	12.10	2.49
11	1	12.90	2.56
12	1	12.90	2.56
13	1	12.30	2.51
14	0	12.40	2.52
15	0	13.50	2.60
16	0	13.60	2.61
17	0	14.10	2.65
18	1	15.10	2.71
19	1	12.90	2.56
20	1	15.10	2.71



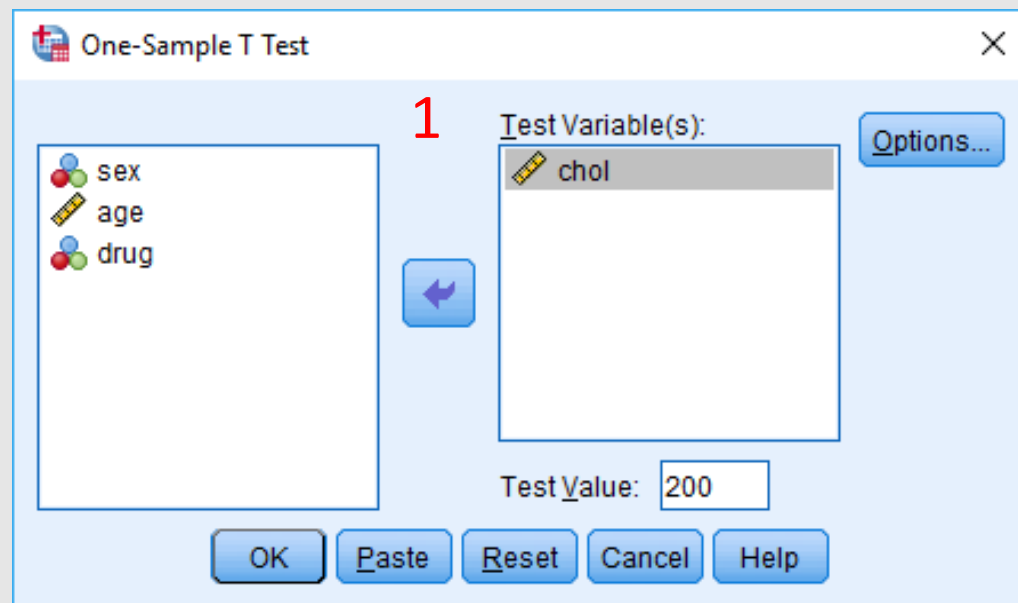
Στατιστικές διαδικασίες / One Sample T test

- Για να εφαρμόσουμε μία στατιστική διαδικασία σε μία ή περισσότερες μεταβλητές χρησιμοποιούμε το μενού Analyze και στην συνέχεια το αντίστοιχο υπομενού
- Για παράδειγμα για την διεξαγωγή T test επιλέγουμε διαδοχικά Analyze -> Compare means -> One Sample T test οπότε εμφανίζεται το παράθυρο της επόμενης διαφάνειας

	 chol	 sex	 age	 drug
1	161	0	31	0
2	163	0	19	0
3	169	0	39	0
4	169	1	41	1
5	170	1	35	0
6	173	0	31	2
7	174	1	33	1
8	176	1	28	0
9	195	0	49	3
10	195	0	41	1
11	233	1	54	0
12	234	0	53	1
13	239	0	44	3
14	244	1	50	1
15	248	0	52	2
16	249	0	47	2
17	249	1	49	2
18	256	0	46	1
19	256	1	63	3
20	258	1	64	3



One Sample T test



Κάνοντας τις ρυθμίσεις όπως φαίνονται στην εικόνα (1) εμφανίζεται το αποτέλεσμα της εικόνας (2)

2

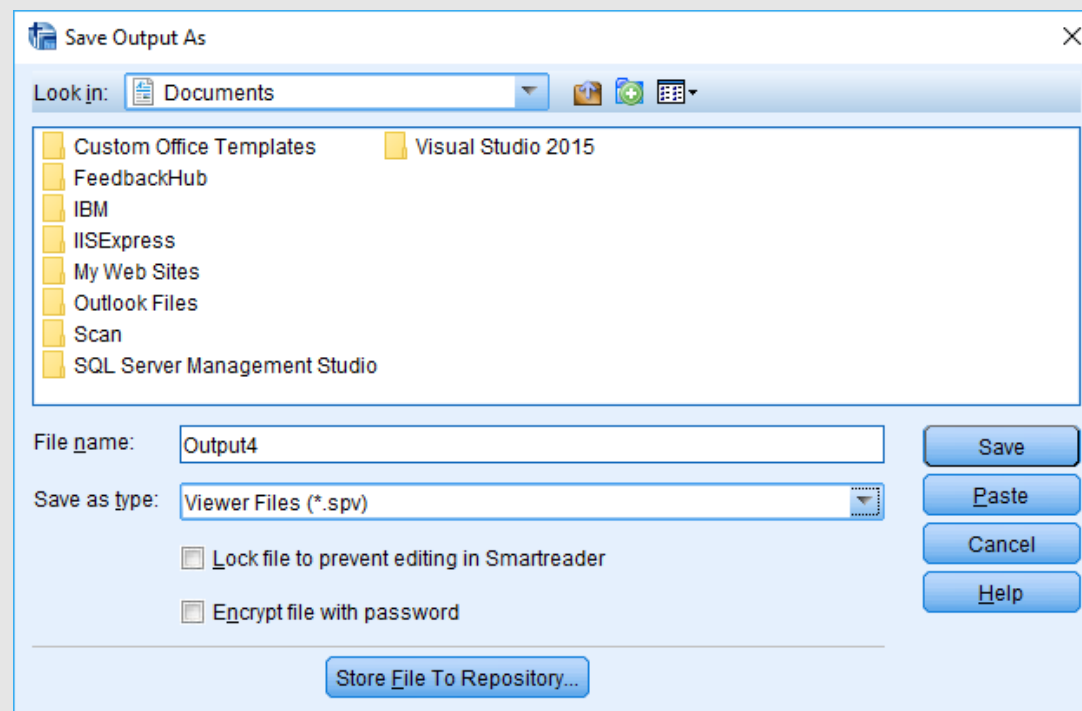
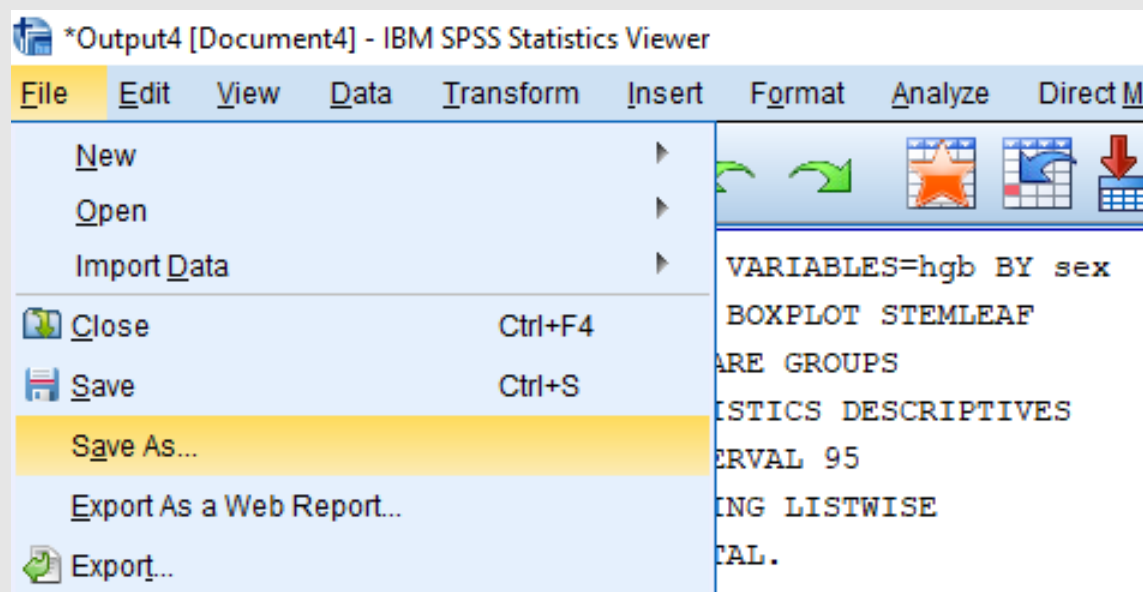
One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
chol	20	210.55	38.362	8.578

One-Sample Test					
Test Value = 200					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower Upper
chol	1.230	19	.234	10.550	-7.40 28.50



Αποθήκευση δεδομένων

- Μπορούμε να αποθηκεύσουμε στο δίσκο το output.
- Από το μενού επιλογών διαλέγουμε **File -> Save As** και στο πεδίο **File name**, δίνουμε το όνομα του αρχείου π.χ. output.srv. Έτσι θα μπορέσουμε αργότερα να ανακτήσουμε τα αποτελέσματα της ανάλυσης που έχουμε κάνει.





Αποθήκευση δεδομένων

- Μπορούμε επίσης να αποθηκεύσουμε και τα δεδομένα του Data Viewer σε κάποιο αρχείο, για μελλοντική χρήση.
- Έτσι από το μενού επιλογών, διαλέγουμε **File -> Save As**. Στο πεδίο **File name**, δίνουμε το όνομα του αρχείου που θέλουμε να αποθηκεύσουμε τα δεδομένα π.χ. data.sav.

