

Εφαρμογή του SPSS

Γιαννούλα Φλώρου

Καθηγήτρια

Τμήμα Λογιστικής και Χρηματοοικονομικής

Μεταπτυχιακό Πρόγραμμα

Λογιστική και Ελεγκτική

Βασικά σημεία

- Εισαγωγή δεδομένων- δήλωση μεταβλητών
- Παρουσίαση με πίνακα – διάγραμμα
- Επεξεργασία μιας-μιας μεταβλητής
- Επεξεργασία δύο μεταβλητών μαζί
- Μετασχηματισμοί μεταβλητών
- Έλεγχος υποθέσεων

Εισαγωγή δεδομένων- δήλωση μεταβλητών

- Τα στατιστικά δεδομένα αντιστοιχούν σε αντικείμενα και μεταβλητές που χαρακτηρίζουν ένα αντικείμενο.
- Η αρχική οθόνη του SPSS αποτελείται από δύο «καρτέλες»,
 - Data view (εισαγωγή δεδομένων σε κάθε γραμμή αντιστοιχεί ένα αντικείμενο σε κάθε στήλη αντιστοιχεί μια μεταβλητή)
 - Variable view (δήλωση μεταβλητών και κωδικών τους)

παράδειγμα

- Θέλουμε να καταγράψουμε τις απαντήσεις 15 ατόμων στο παρακάτω ερωτηματολόγιο:
 - Φύλο
 - Οικογενειακή κατάσταση
 - Αριθμός παιδιών
 - Μηνιαίο οικογενειακό εισόδημα
 - Μηνιαίες υποχρεώσεις σε τραπεζικά δάνεια
 - Αισιοδοξία για το μέλλον της οικονομίας

παράδειγμα

μεταβλητή	Οι δυνατές απαντήσεις
Φύλο	Άνδρας γυναίκα
Οικογενειακή κατάσταση	Άγαμος έγγαμος διαζευγμένος χήρος
Αριθμός παιδιών	0,1,2,3,4,.....
Μηνιαίο οικογενειακό εισόδημα
Μηνιαίες υποχρεώσεις σε τραπεζικά δάνεια
Αισιοδοξία για το μέλλον της οικονομίας	Καθόλου, λίγο, μέτρια ,πολύ, πάρα πολύ

Παράδειγμα variable view

Μεταβλητή (name)	Τύπος (type)	Περιγραφή (label)	κωδικοί απαντήσεων (values)	Είδος (Measure)
filo	numeric	Φύλο	1=Άνδρας 2=γυναίκα	nominal
oik_kat	numeric	Οικογενειακή κατάσταση	1=Άγαμος 2= έγγαμος 3= διαζευγμένος 4=χήρος	nominal
paidia	numeric	Αριθμός παιδιών	0,1,2,3,4,.....	scale
eisodima	numeric	Μηνιαίο οικογενειακό εισόδημα	scale
daneia	numeric	Μηνιαίες υποχρεώσεις σε τραπεζικά δάνεια	
ais_melon	numeric	Αισιοδοξία για το μέλλον της οικονομίας	1= Καθόλου, 2= λίγο, 3= μέτρια, 4=πολύ, 5=πάρα πολύ	nominal

Παράδειγμα data view

filo	oik_kat	paidia	eisodima	daneia	ais_melon
1	1	0	1500	1010,53	3
1	2	2	1300	906,78	4
2	2	1	900	315,45	2
1	1	0	800	120,44	2
2	4	2	950	159,06	1
2	1	1	780	457,90	1
2	3	3	1650	825.47	3
1	1	0	1450	543.02	5
1	2	3	1100	785,21	4

*px_eisagogi.sav [DataSet2] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

10 : daneia

	filo	oik_kat	paidia	eisodima	daneia	ais_meon	var	var	var	var
1	Ανδρας	άγαμος	0	1500,00	Μηνιαίες υποχρεώσεις σε τραπεζικά δάνεια					
2	Ανδρας	έγγαμος	2	1300,00	906,78	πολύ				
3	Γυναίκα	έγγαμος	1	900,00	315,45	λιγο				
4	Ανδρας	άγαμος	0	80,00	120,44	λιγο				
5	Γυναίκα	χήρος	2	950,00	159,07	καθόλου				
6	Γυναίκα	άγαμος	1	780,00	457,90	καθόλου				
7	Γυναίκα	διαζευγμένος	3	150,00	825,47	μέτρια				
8	Ανδρας	άγαμος	0	1450,00	543,02	πέρα πολύ				
9	Ανδρας	έγγαμος	3	1100,00	785,21	πολύ				
10										
11										
12										
13										
14										
15										
16										

Data View Variable View

*untitled2 [DataSet2] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Col	Align	Measure	Role
1	filo	Numeric	8	0	Φύλο	{1, Ανδρας}...	None	8	Right	Nominal	Input
2	oik_kat	Numeric	8	0	Οικογενειακή κ...	{1, άγαμος}...	None	8	Right	Nominal	Input
3	paidia	Numeric	8	0	Αριθμός παιδιών	None	None	8	Right	Unknown	Input
4	eisodima	Numeric	8	2	Μηνιαίο εισογμ...	None	None	8	Right	Unknown	Input
5	daneia	Numeric	8	2	Μηνιαίες υποχρ...	None	None	8	Right	Unknown	Input
6	ais_meon	Numeric	8	0	Αισιοδοξία για ...	{1, καθόλου}...	None	8	Right	Ordinal	Input
7											
8				0							
9											
10											
11											
12											
13											
14											
15											
16											

Data View Variable View

Παρουσίαση με πίνακα – διάγραμμα

Προκειμένου να έχουμε μια πρώτη εικόνα των απαντήσεων, παρουσιάζουμε κάθε μεταβλητή με πίνακα ή και διάγραμμα.

Αυτό γίνεται από το menu

Analyze

Descriptive Statistics

Frequency

(επιλογή μεταβλητής και επιλογή διαγράμματος από το πλήκτρο chart)

Τα αποτελέσματα της παρουσίασης εμφανίζονται σε νέο παράθυρο με τίτλο Output

πίνακας συχνοτήτων

The screenshot shows the IBM SPSS Statistics Data Editor interface. The main window displays a data table with columns 'filo' and 'oik_kat'. The 'Frequencies' dialog box is open, showing the variable 'Φύλο [filo]' selected in the 'Variable(s):' list. The 'Display frequency tables' checkbox is checked. The dialog box includes buttons for 'Statistics...', 'Charts...', 'Format...', and 'Bootstrap...'. The 'OK' button is highlighted.

	filo	oik_kat
1	Ανδρας	άγαμο
2	Ανδρας	έγγαμο
3	Γυναίκα	έγγαμο
4	Ανδρας	άγαμο
5	Γυναίκα	χήρο
6	Γυναίκα	άγαμο
7	Γυναίκα	διαζευγμέν
8	Ανδρας	άγαμο
9	Ανδρας	έγγαμο
10		
11		
12		
13		
14		

Παράθυρο output

The screenshot shows the IBM SPSS Statistics Viewer interface. The main window displays the 'Output' pane with a tree view on the left and a 'Frequencies' report on the right. The report includes a summary table and a detailed table for the variable 'Οικογενειακή κατάσταση'.

Statistics
Οικογενειακή κατάσταση

N	Valid	Missing
	9	0

Οικογενειακή κατάσταση

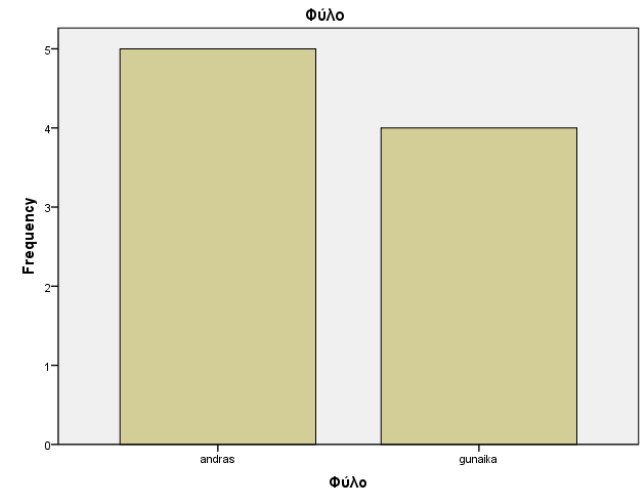
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid άγαμος	4	44,4	44,4	44,4
έγγαμος	3	33,3	33,3	77,8
διαζευγμένος	1	11,1	11,1	88,9
χήρος	1	11,1	11,1	100,0

IBM SPSS Statistics Processor is ready

παράδειγμα

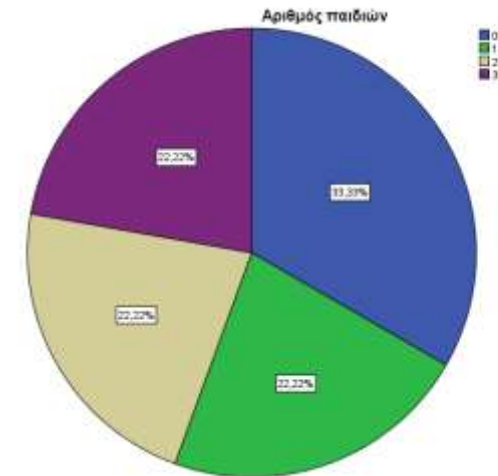
Φύλο

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	andras	5	55,6	55,6	55,6
	gunaika	4	44,4	44,4	100,0
	Total	9	100,0	100,0	



Αριθμός παιδιών

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3	33,3	33,3	33,3
	1	2	22,2	22,2	55,6
	2	2	22,2	22,2	77,8
	3	2	22,2	22,2	100,0
	Total	9	100,0	100,0	



Επιλογή τύπου γραφήματος

The screenshot shows the PASW Statistics Data Editor interface. The 'Graphs' menu is open, displaying various chart options. The 'Line...' option is highlighted. The data table below shows the following information:

	filo	oik_kat	paidia	eisodima	gariera	ais_meron
1	1	1	0	1500	1010,53	3
2	1	2	2	1300	906,78	4
3	2	2	1	900	315,45	2
4	1	1	0	800	120,44	2
5	2	4	2	950	159,06	1
6	2	1	1	780	457,90	1
7	2	3	3	1650	825,47	3
8	1	2	3	1100	785,21	4
9	1	1	0	1450	543,02	5
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						

Επεξεργασία μιας-μιας μεταβλητής

- Για κάθε μια ποσοτική μεταβλητή, υπολογίζουμε μέση τιμή, διάμεσο, τυπική απόκλιση, κλπ.
- Για κάθε μια ποιοτική μεταβλητή, **δεν** υπολογίζουμε μέση τιμή, διάμεσο, τυπική απόκλιση, κλπ. Υπολογίζουμε μόνο πίνακα συχνοτήτων με ποσοστά.

Εντολή descriptives

The screenshot shows the IBM SPSS Statistics Data Editor interface. The 'Reports' menu is open, and the 'Descriptive Statistics' option is selected. The 'Descriptives...' option is highlighted in the submenu. The 'Descriptives' dialog box is open, showing the 'Variable(s):' list with the following variables: Φύλο [filo], Οικογενειακή κατάστ..., Αριθμός παιδιών [ρα..., Μηνιαίο οικογενειακ..., Μηνιαίες υποχρεώσε..., and Αισιοδοξία για το μέλ... The 'Save standardized values as variables' checkbox is unchecked. The 'Options...' and 'Bootstrap...' buttons are visible on the right side of the dialog box. The 'OK', 'Paste', 'Reset', 'Cancel', and 'Help' buttons are at the bottom.

File Edit View Data Transform Reports Descriptive Statistics Tables Compare Means General Linear Model Generalized Linear Models Mixed Models Correlate Regression Loglinear Neural Networks Classify Dimension Reduction Scale Nonparametric Tests Forecasting Survival Multiple Response Missing Value Analysis... Multiple Imputation Complex Samples Quality Control Utilities Add-ons Window Help

123 Frequencies...
Pd Descriptives...
Explore...
Crosstabs...
Ratio...
P-P Plots...
Q-Q Plots...

159,07 καθόλου

Descriptives

Variable(s):

Φύλο [filo]
Οικογενειακή κατάστ...
Αριθμός παιδιών [ρα...
Μηνιαίο οικογενειακ...
Μηνιαίες υποχρεώσε...
Αισιοδοξία για το μέλ...

Options...
Bootstrap...

Save standardized values as variables

OK Paste Reset Cancel Help

	filo	oik_kat
1	Ανδρας	άγαμα
2	Ανδρας	έγγαμο
3	Γυναίκα	έγγαμο
4	Ανδρας	άγαμα
5	Γυναίκα	χήρα
6	Γυναίκα	άγαμα
7	Γυναίκα	διαζευγμέν
8	Ανδρας	άγαμα
9	Ανδρας	έγγαμο
10		
11		
12		
13		
14		

Data View Variable View

9:50 μμ

παράδειγμα

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Αριθμός παιδιών	9	0	3	1,33	1,225
Μηνιαίο οικογενειακό εισόδημα	9	780	1650	1158,89	325,709
Μηνιαίες υποχρεώσεις σε τραπεζικά δάνεια	9	120,44	1010,53	569,3178	329,41794
Valid N (listwise)	9				

Αισιοδοξία για το μέλλον της οικονομίας

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	katholou	2	22,2	22,2	22,2
	ligo	2	22,2	22,2	44,4
	metria	2	22,2	22,2	66,7
	polu	2	22,2	22,2	88,9
	para polu	1	11,1	11,1	100,0
	Total	9	100,0	100,0	

Επεξεργασία δύο μεταβλητών μαζί

- Για δύο ποιοτικές μεταβλητές που υποθέτουμε ότι σχετίζονται υπολογίζουμε πίνακα διασταυρώσεων με ποσοστά. (crosstab)
- Για δύο ποσοτικές μεταβλητές που υποθέτουμε ότι σχετίζονται υπολογίζουμε συντελεστή συσχέτισης. (correlation)
- Για μία ποιοτική και μία ποσοτική μεταβλητή που υποθέτουμε ότι σχετίζονται υπολογίζουμε μέσες τιμές ξεχωριστά για κάθε κατηγορία της ποιοτικής μεταβλητής. (explore)

crosstab

The screenshot shows the IBM SPSS Statistics Data Editor interface. The menu path is: **Reports** > **Descriptive Statistics** > **Crosstabs...**. The Crosstabs dialog box is open, showing the following settings:

- Row(s):** Φύλο [filo]
- Column(s):** Οικογενειακή κατάσταση ...
- Layer 1 of 1:** Previous, Next

The data table in the background is as follows:

	filo	oik_kat
1	Ανδρας	άγαμο
2	Ανδρας	έγγαμο
3	Γυναίκα	έγγαμο
4	Ανδρας	άγαμο
5	Γυναίκα	χήρο
6	Γυναίκα	άγαμο
7	Γυναίκα	διαζευγμένο
8	Ανδρας	άγαμο
9	Ανδρας	έγγαμο
10		
11		
12		
13		
14		
15		

παράδειγμα

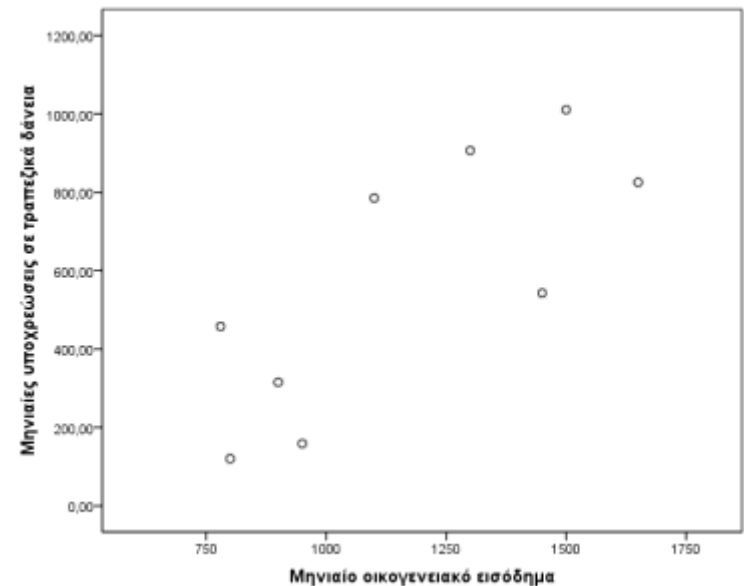
Φύλο * Αισιοδοξία για το μέλλον της οικονομίας Crosstabulation

			Αισιοδοξία για το μέλλον της οικονομίας					Total
			katholou	ligo	metria	polu	para polu	
Φύλο	andras	Count	0	1	1	2	1	5
		% within Φύλο	,0%	20,0%	20,0%	40,0%	20,0%	100,0%
	gunaika	Count	2	1	1	0	0	4
		% within Φύλο	50,0%	25,0%	25,0%	,0%	,0%	100,0%
Total		Count	2	2	2	2	1	9
		% within Φύλο	22,2%	22,2%	22,2%	22,2%	11,1%	100,0%

Correlations

		Μηνιαίο οικογενειακό εισόδημα	Μηνιαίες υποχρεώσεις σε τραπεζικά δάνεια
Μηνιαίο οικογενειακό εισόδημα	Pearson Correlation	1	,767*
	Sig. (2-tailed)		,016
	N	9	9
Μηνιαίες υποχρεώσεις σε τραπεζικά δάνεια	Pearson Correlation	,767*	1
	Sig. (2-tailed)	,016	
	N	9	9

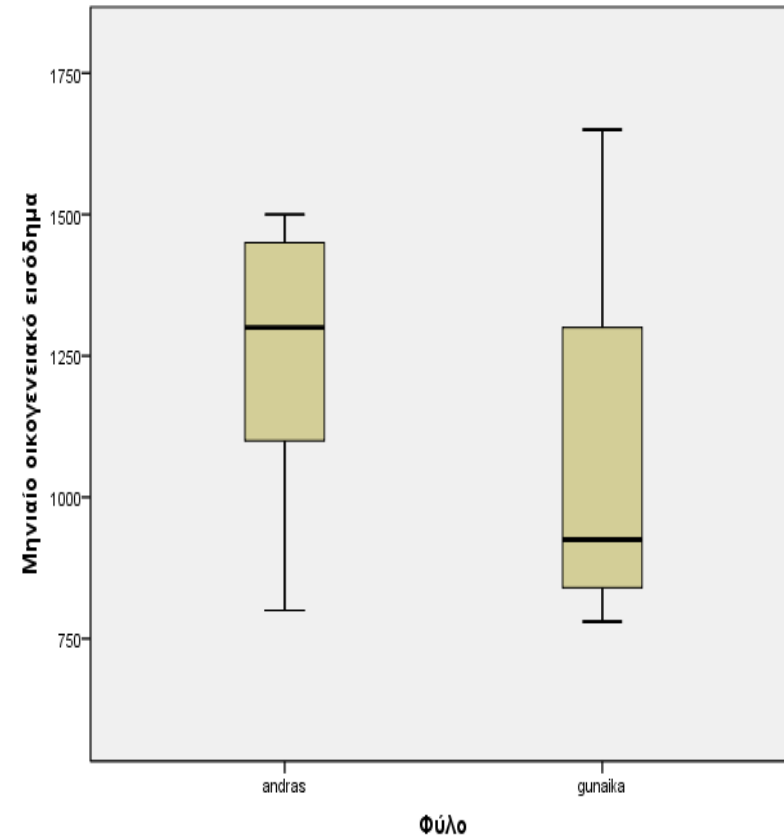
*. Correlation is significant at the 0.05 level (2-tailed).



Σχέση ποσοτικής-ποιοτικής μεταβλητής

Descriptives

Φύλο		Statistic	Std. Error
Μηνιαίο οικογενειακό εισόδημα	andras	Mean	1230,00
		95% Confidence Interval for Mean	
		Lower Bound	874,44
		Upper Bound	1585,56
		5% Trimmed Mean	1238,89
		Median	1300,00
		Variance	82000,000
		Std. Deviation	286,356
		Minimum	800
		Maximum	1500
		Range	700
		Interquartile Range	525
		Skewness	-,906
	Kurtosis	-,163	2,000
gunaika	Mean	1070,00	196,596
		95% Confidence Interval for Mean	
		Lower Bound	444,34
		Upper Bound	1695,66
		5% Trimmed Mean	1053,89
		Median	925,00
		Variance	154600,000
		Std. Deviation	393,192
		Minimum	780
		Maximum	1650
		Range	870
		Interquartile Range	665
		Skewness	1,800
	Kurtosis	3,414	2,619



Μετασχηματισμοί μεταβλητών

- Κάποιες φορές μπορεί να θέλουμε να αλλάξουμε μια ή περισσότερες μεταβλητές, να την χωρίσουμε σε ομάδες, να μετασχηματίσουμε τη μονάδα μέτρησης κλπ
- Κάποιες φορές υπολογίζουμε νέα μεταβλητή από ήδη υπάρχουσες (άθροισμα ή διαφορά δύο, υπολογισμός φόρου κλπ)
- Κάποιες φορές θέλουμε να επιλέξουμε μόνο κάποια αντικείμενα με χαρακτηριστικές ιδιότητες ως προς κάποια μεταβλητή

Επιλογή με συνθήκη

The screenshot displays the PASW Statistics Data Editor interface. The main window shows a data table with columns: `filo`, `oik_kat`, `paidia`, `eisodima`, `daneia`, `ais_melon`, and several empty `var` columns. The first row contains the values: 1, 1, 1, 0, 1500, 1010,53, 3.

Two dialog boxes are open over the data table:

- Select Cases:** The **Select** section has **If condition is satisfied** selected. The **Output** section has **Filter out unselected cases** selected.
- Select Cases: If:** The condition `ais_melon > 3` is entered in the text field.

Buttons for **OK**, **Paste**, **Reset**, **Cancel**, and **Help** are visible at the bottom of the **Select Cases** dialog. Buttons for **Continue**, **Cancel**, and **Help** are visible at the bottom of the **Select Cases: If** dialog.

Data View Variable View

PASW Statistics Processor is ready

Υπολογισμός νέας μεταβλητής

The screenshot displays the PASW Statistics Data Editor interface. A 'Compute Variable' dialog box is open, showing the following configuration:

- Target Variable: eisod_net
- Numeric Expression: eisodima-daneaia
- Function group: All
- Functions and Special Variables: (empty)

The background data table is as follows:

	filo	oik_kat	paidia
1	1	1	1
2	1	2	2
3	2	2	2
4	1	1	1
5	2	4	4
6	2	1	1
7	2	3	3
8	1	2	2
9	1	1	1
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			

Μεταφορά αποτελεσμάτων σε αρχείο word ή excel

- Απλώς χρησιμοποιούμε την αντιγραφή και επικόλληση ή ειδική επικόλληση μέσω του «ποντικιού».
- Η εισαγωγή των δεδομένων μπορεί να γίνει και ανοίγοντας ένα αρχείο excel από το menu file open.

Έλεγχος υποθέσεων

- Όταν θέλουμε να κάνουμε έλεγχο υποθέσεων για δύο μέσες τιμές επιλέγουμε :
- Analyze Compare means
 - Independent samples T-test για ανεξάρτητα δείγματα
 - Paired samples T-test για μεταβλητές μετρούμενες κατά ζεύγη.

*px_eisagogi.sav [DataSet2] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Utilities Add-ons Window Help

3 :

	filo	oik_kat
1	Ανδρας	άγαμο
2	Ανδρας	έγγαμο
3	Γυναίκα	έγγαμο
4	Ανδρας	άγαμο
5	Γυναίκα	χήρο
6	Γυναίκα	άγαμο
7	Γυναίκα	διαζευγμέν
8	Ανδρας	άγαμο
9	Ανδρας	έγγαμο
10		
11		
12		
13		
14		

159,07 καθόλου
457,90 καθόλου
825,47 μέτρια
543,02 πάρα πολύ
785,21 πολύ

var var var

Visible:

Compare Means

- Means...
- One-Sample T Test...
- Independent-Samples T Test...
- Paired-Samples T Test...
- One-Way ANOVA...

Data View Variable View

10:04 μμ

Αποτελέσματα ελέγχου υποθέσεων

Independent Samples Test

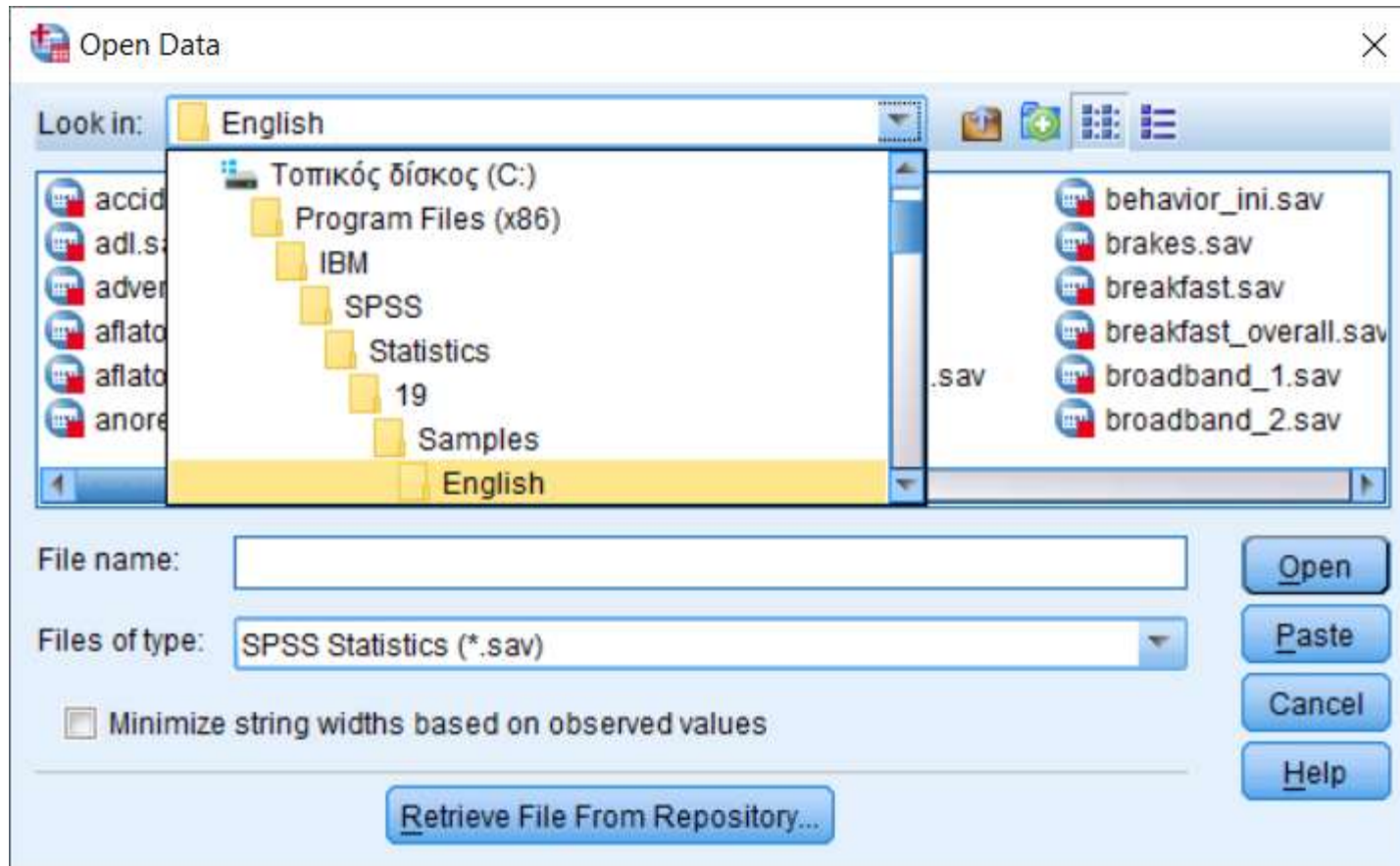
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
antikeimeniki axia oikopedou	Equal variances assumed	7,040	,010	1,517	77	,133	4834,79	3187,700	-1512,734	11182,311
	Equal variances not assumed			1,508	62,140	,137	4834,79	3206,813	-1575,252	11244,829

- Προσέχουμε την τιμή t και την πιθανότητα να κάνουμε λάθος απορρίπτοντας την H0, που αντιστοιχεί στην τιμή sig.
- Αν $sig > 0.05$ δεν μπορούμε να απορρίψουμε την H0.

Εφαρμογή με έτοιμο αρχείο δεδομένων

- Data credit card

Επιλογή αρχείου από ενσωματωμένα δεδομένα



Credit card αρχείο δεδομένων

- Δεδομένα για κατόχους χρεωστικών καρτών, 26280 καταγραφές, ισομοιρασμένες σε δύο ετη, σε 12 μήνες κάθε έτους και ισομοιρασμένες ανά κατηγορία δαπάνης
- Ηλικία, φύλο, είδος κάρτας ημερομηνία έκδοσης κάρτας, τύπος αγορών, πλήθος αντικειμένων, ποσό που δαπανήθηκε

Ενέργειες 1

- Εντοπίστε το είδος των μεταβλητών
- Παρουσιάστε μια ποιοτική μεταβλητή με πίνακα συχνοτήτων και γράφημα
- Παρουσιάστε μια ποσοτική μεταβλητή με γράφημα, και θηκόγραμμα. «Μοιάζει» στην κανονική κατανομή; (γραφικά και με έλεγχο υποθέσεων)
- Βρείτε μέτρα θέσεως και διασποράς για την ποσοτική μεταβλητή.
- Γράψτε τα συμπεράσματα σας σε word μεταφέροντας κι τα αντίστοιχα από το spss

Παρουσιάστε μια ποιοτική μεταβλητή με πίνακα συχνοτήτων και γράφημα

The screenshot shows the IBM SPSS Statistics Data Editor interface. The 'Frequencies' dialog box is open, showing the variable 'Type of transaction' selected in the 'Variable(s):' list. The 'Display frequency tables' checkbox is checked. The background shows a data table with columns 'card', 'card_data', 'Items', and 'spent'.

card	card_data	Items	spent
169	Mastercard	09/11/20	295.33
170	Mastercard	09/11/20	.00
171	Mastercard	09/11/20	234.67
172	Mastercard	09/11/20	540.95
173	Mastercard	09/11/20	305.72
174	Mastercard	09/11/20	622.73
175	Mastercard	09/11/20	410.07
176	Mastercard	09/11/20	121.29
177	Mastercard	09/11/20	150.48
178	Mastercard	09/11/20	59.12
179	Mastercard	09/11/20	298.38
180	Mastercard	09/11/20	469.90
181	Mastercard	05/11/20	188.32
182	Mastercard	09/11/20	662.79
183	Mastercard	09/11/20	143.12
184	Mastercard	09/11/20	175.83

The screenshot shows the 'Frequencies: Charts' dialog box. The 'Chart Type' section has 'Bar charts' selected. The 'Chart Values' section has 'Frequencies' selected. The 'Display frequency tables' checkbox is checked. The 'Variable(s):' list contains 'Type of transaction'.

Variable(s):
Type of transaction [...]

Chart Type:
 None
 Bar charts
 Pie charts
 Histograms:
 Show normal curve on histogram

Chart Values:
 Frequencies Percentages

Display frequency tables

OK Paste Reset Cancel Help

The screenshot shows the 'Frequencies: Charts' dialog box. The 'Chart Type' section has 'Bar charts' selected. The 'Chart Values' section has 'Frequencies' selected. The 'Display frequency tables' checkbox is checked. The 'Variable(s):' list contains 'Type of transaction'.

Variable(s):
Type of transaction [...]

Chart Type:
 None
 Bar charts
 Pie charts
 Histograms:
 Show normal curve on histogram

Chart Values:
 Frequencies Percentages

Continue Cancel Help

Παρουσιάστε μια ποσοτική μεταβλητή με γράφημα «Μοιάζει» στην κανονική κατανομή; (γραφικά και με έλεγχο υποθέσεων)

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

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

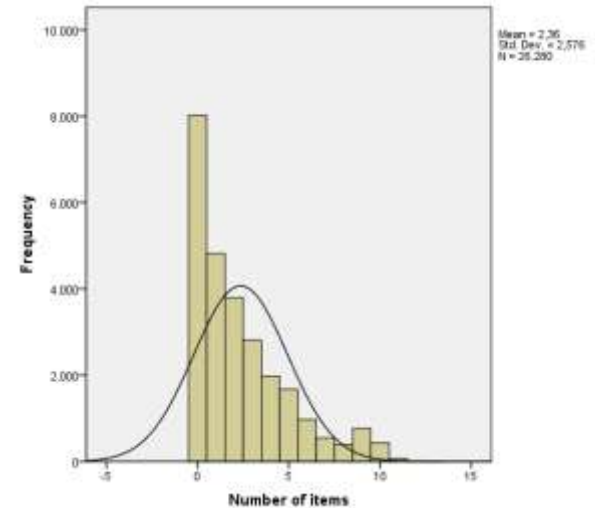
Chart Builder...
Graphboard Template Chooser...

Legacy Dialogs

Variable: 11 of 11 Variables

	card	card_date	month	quarter	year	type	item
169	Mastercard	09/11/2004	May	Q2	2008	Travel	
170	Mastercard	09/11/2004	May	Q2	2008	Other	
171	Mastercard	09/11/2004	June	Q2	2007	Grocery	
172	Mastercard	09/11/2004	June	Q2	2007	Retail	
173	Mastercard	09/11/2004	June	Q2	2007	Entertainment	
174	Mastercard	09/11/2004	June	Q2	2007	Travel	
175	Mastercard	09/11/2004	June	Q2	2007	Other	
176	Mastercard	09/11/2004	June	Q2	2008	Grocery	
177	Mastercard	09/11/2004	June	Q2	2008	Retail	
178	Mastercard	09/11/2004	June	Q2	2008	Entertainment	
179	Mastercard	09/11/2004	June	Q2	2008	Travel	
180	Mastercard	09/11/2004	June	Q2	2008	Other	5 469.90
181	Mastercard	09/11/2004	July	Q3	2007	Grocery	3 188.32
182	Mastercard	09/11/2004	July	Q3	2007	Retail	10 662.79
183	Mastercard	09/11/2004	July	Q3	2007	Entertainment	1 143.12
184	Mastercard	09/11/2004	July	Q3	2007	Travel	1 129.83

Bar...
3-D Bar...
Line...
Area...
Pie...
High-Low...
Boxplot...
Error Bar...
Population Pyramid...
ScatterDot...
Histogram...



Data View Variable View

Histogram... IBM SPSS Statistics Data Editor

Histogram

Variable: Number of Items [item]

Display normal curve

Panel by: Rows

Use chart specifications from: [None]

OK Paste Reset Cancel Help

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

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Regions
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis
Multiple Imputation
Complex Samples
Quality Control

1-Sample K-S...

year	type	item	spent	var	var	var
2008	Travel	2	295.33			
2008	Other	0	.00			
2007	Grocery	4	234.67			
2007	Retail	6	540.95			
2007	Entertainment	2	305.72			
2007	Travel	5	622.73			
2007	Other					
2008	Grocery					
2008	Retail					
2008	Entertainment					
2008	Travel					
2007	Travel	1	129.83			

Chi-square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...

IBM SPSS Statistics Processor is ready

Παρουσιάστε μια ποσοτική μεταβλητή με θηκόγραμμα.

IBM SPSS Statistics Data Editor

	card	card_date	month	quarter	year	type	items	var	var
169	Mastercard	09/11/2004	May	Q2	2008	Travel			
170	Mastercard	09/11/2004	May	Q2	2008	Other			
171	Mastercard	09/11/2004	June	Q2	2007	Grocery			
172	Mastercard	09/11/2004	June	Q2	2007	Retail			
173	Mastercard	09/11/2004	June	Q2	2007	Entertainment			
174	Mastercard	09/11/2004	June	Q2	2007	Travel			
175	Mastercard	09/11/2004	June	Q2	2007	Other			
176	Mastercard	09/11/2004	June	Q2	2008	Grocery			
177	Mastercard	09/11/2004	June	Q2	2008	Retail			
178	Mastercard	09/11/2004	June	Q2	2008	Entertainment			
179	Mastercard	09/11/2004	June	Q2	2008	Travel			
180	Mastercard	09/11/2004	June	Q2	2008	Other	5	469.90	
181	Mastercard	09/11/2004	July	Q3	2007	Grocery	3	188.32	
182	Mastercard	09/11/2004	July	Q3	2007	Retail	10	662.79	
183	Mastercard	09/11/2004	July	Q3	2007	Entertainment	1	143.12	
184	Mastercard	09/11/2004	July	Q3	2007	Travel	1	129.63	

Legacy Dialogs: Bar, Line, Area, Pie, High-Low, **Boxplot**, Engr Bar, Population Pyramid, Scatter/Dot, Histogram

Boxplot

Simple

Clustered

Data in Chart Are

Summaries for groups of cases

Summaries of separate variables

Define Cancel Help

Define Simple Boxplot: Summaries of Separate Variables

Boxes Represent: Number of Items [items]

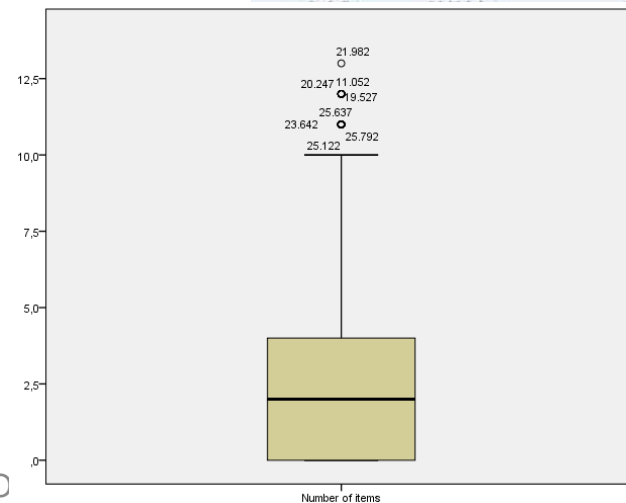
Label Cases by:

Panel by:

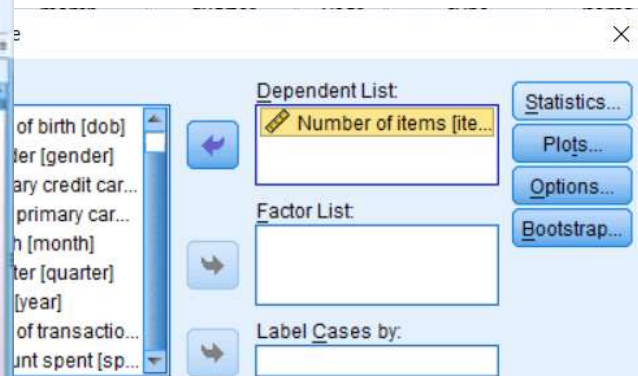
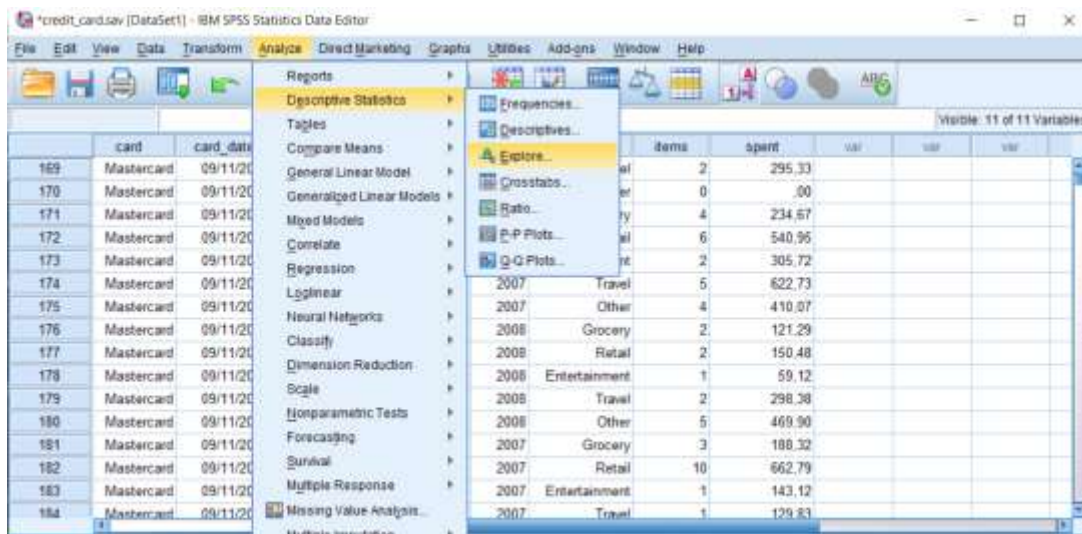
Rows:

Columns:

OK Paste Reset Cancel Help



μέτρα θέσεως και διασποράς για την ποσοτική μεταβλητή

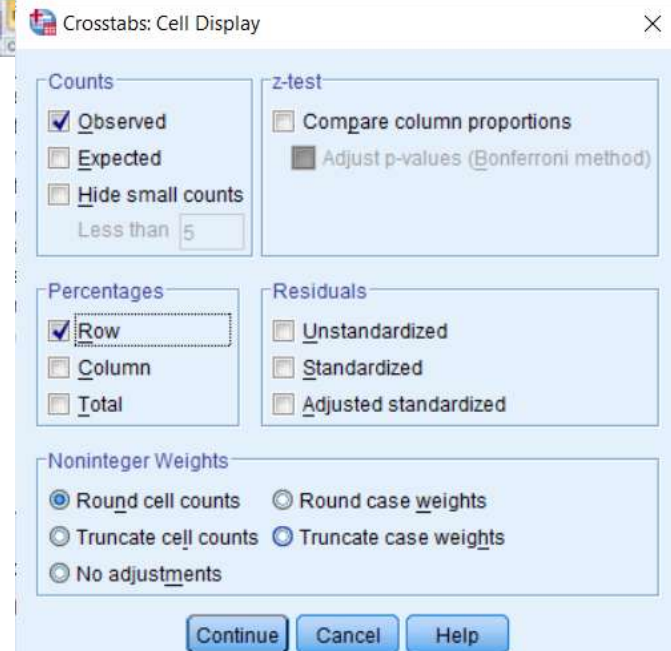
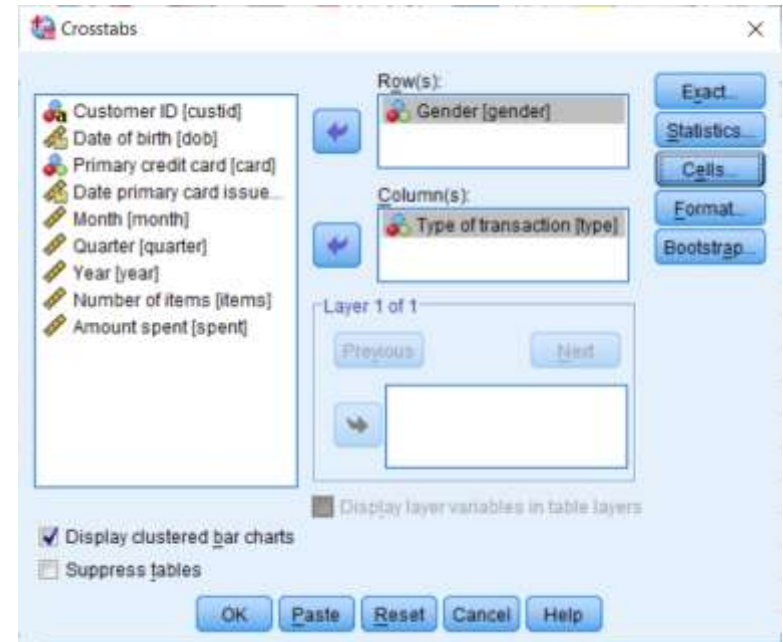
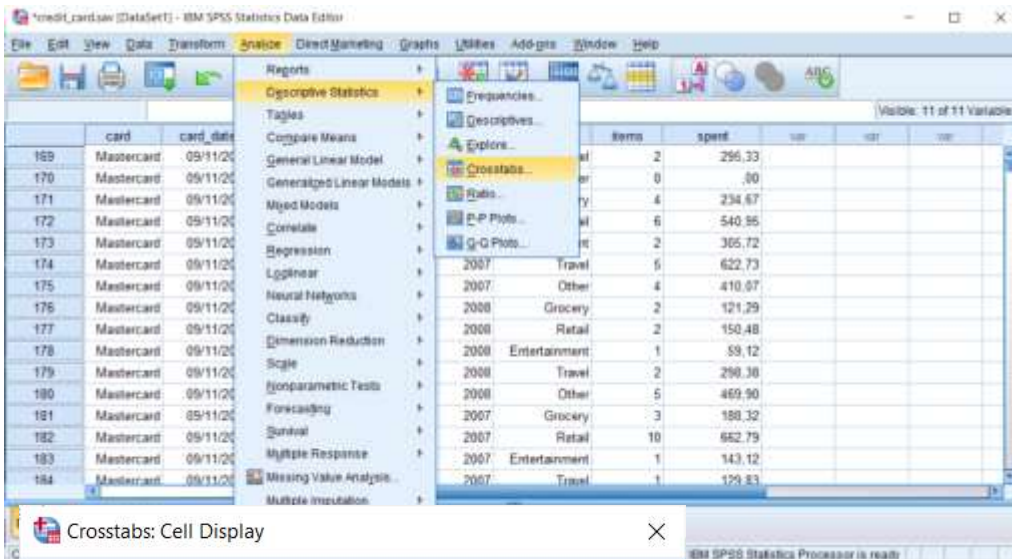


Descriptives			Statistic	Std. Error
Number of items	Mean		2,36	,016
	95% Confidence Interval for Mean	Lower Bound	2,33	
		Upper Bound	2,39	
5% Trimmed Mean		2,10		
Median		2,00		
Variance		6,638		
Std. Deviation		2,576		
Minimum		0		
Maximum		13		
Range		13		
Interquartile Range		4		
Skewness		1,260	,015	
Kurtosis		,999	,030	

Ενέργειες 2

- Περιγράψτε δύο ποιοτικές μεταβλητές αλληλοσχετιζόμενες. (πίνακας και γράφημα). Ελέγξτε αν μπορούν να θεωρηθούν ανεξάρτητες.
- Περιγράψτε μία ποσοτική μεταβλητή σε σχέση με μια ποιοτική μεταβλητή. (πίνακας και γράφημα)
- Δώστε ένα διάστημα εμπιστοσύνης για τη μέση τιμή μιας ποσοτικής μεταβλητής στον συνολικό πληθυσμό
- Συγκρίνετε τις μέσες τιμές της ποσοτικής μεταβλητής σε δύο κατηγορίες ποιοτικής μεταβλητής

Περιγράψτε δύο ποιοτικές μεταβλητές αλληλοσχετιζόμενες.



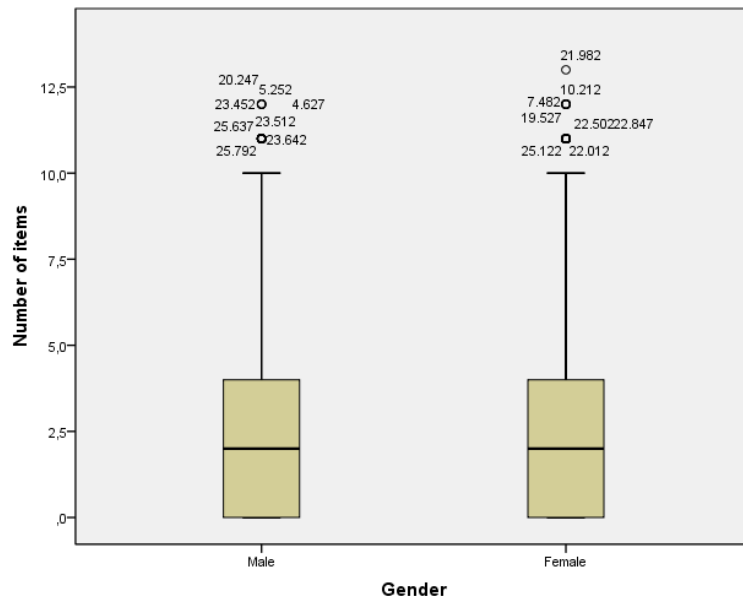
Gender * Type of transaction Crosstabulation

		Type of transaction					Total	
		Grocery	Retail	Entertainment	Travel	Other		
Gender	Male	Count	2688	2688	2688	2688	2688	13440
		% within Gender	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%
	Female	Count	2568	2568	2568	2568	2568	12840
		% within Gender	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%
Total		Count	5256	5256	5256	5256	5256	26280
		% within Gender	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%

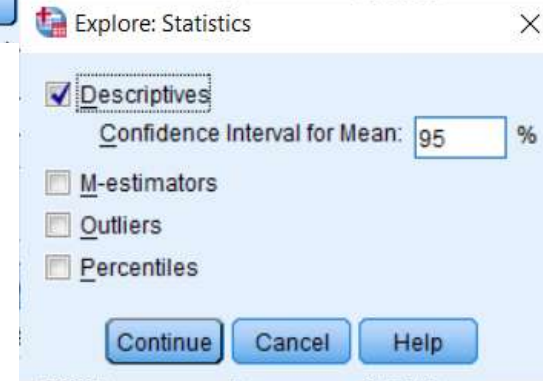
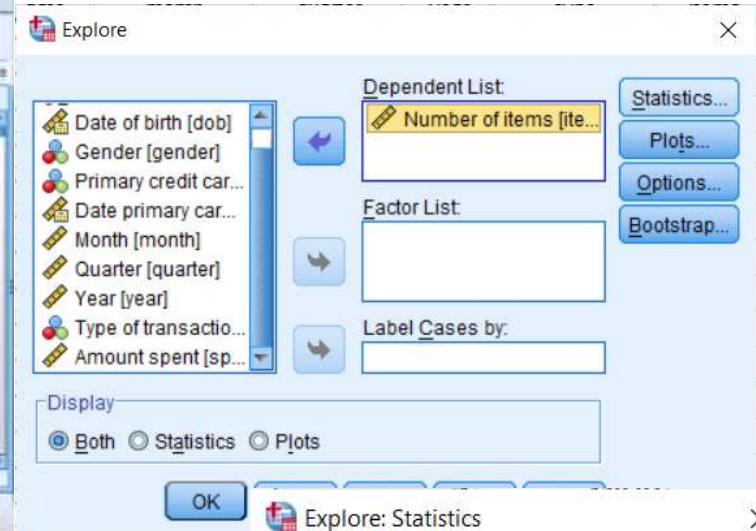
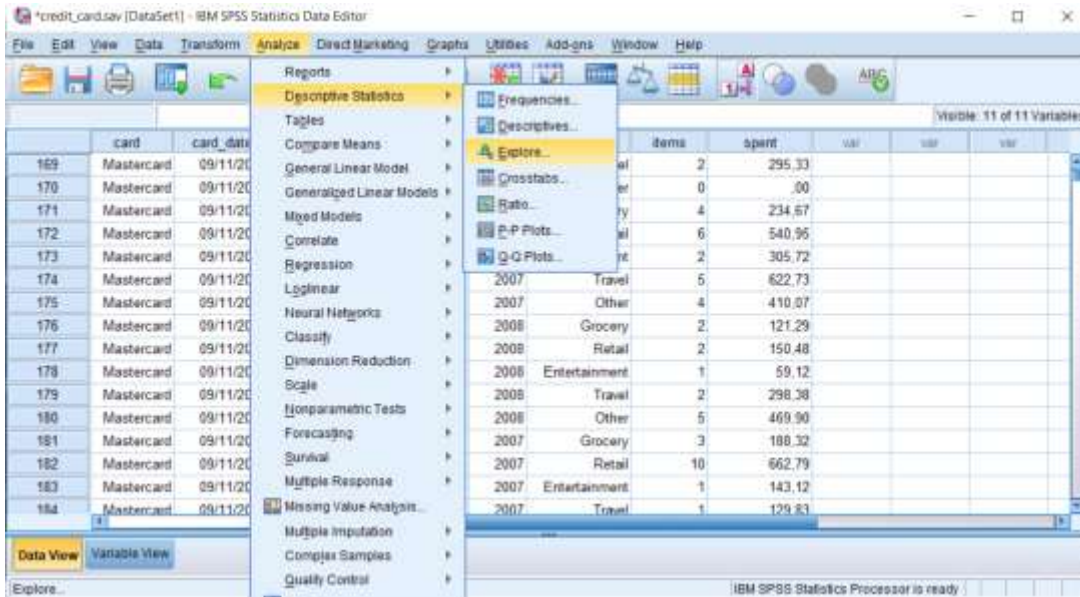
Περιγράψτε μία ποσοτική μεταβλητή σε σχέση με μια ποιοτική μεταβλητή.

The screenshot shows the IBM SPSS Statistics interface. The 'Explore' dialog box is open, with 'Number of items [items]' in the 'Dependent List' and 'Gender [gender]' in the 'Factor List'. The background data table is as follows:

card	card_date	items	spent
169	Mastercard	09/11/20	
170	Mastercard	09/11/20	
171	Mastercard	09/11/20	
172	Mastercard	09/11/20	
173	Mastercard	09/11/20	
174	Mastercard	09/11/20	
175	Mastercard	09/11/20	
176	Mastercard	09/11/20	
177	Mastercard	09/11/20	
178	Mastercard	09/11/20	
179	Mastercard	09/11/20	
180	Mastercard	09/11/20	
181	Mastercard	09/11/20	
182	Mastercard	09/11/20	
183	Mastercard	09/11/20	
184	Mastercard	09/11/20	



Δώστε ένα διάστημα εμπιστοσύνης για τη μέση τιμή μιας ποσοτικής μεταβλητής στον συνολικό πληθυσμό



Descriptives			Statistic	Std. Error
Number of items	Mean		2,36	,016
	95% Confidence Interval for Mean	Lower Bound	2,33	
		Upper Bound	2,39	
	5% Trimmed Mean		2,10	
	Median		2,00	
	Variance		6,638	
	Std. Deviation		2,576	
	Minimum		0	
	Maximum		13	
	Range		13	
Interquartile Range		4		
Skewness		1,260	,015	
Kurtosis		,999	,030	

Συγκρίνετε τις μέσες τιμές της ποσοτικής μεταβλητής σε δύο κατηγορίες ποιοτικής μεταβλητής

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Number of items	Male	13440	2,34	2,565	,022
	Female	12840	2,38	2,589	,023

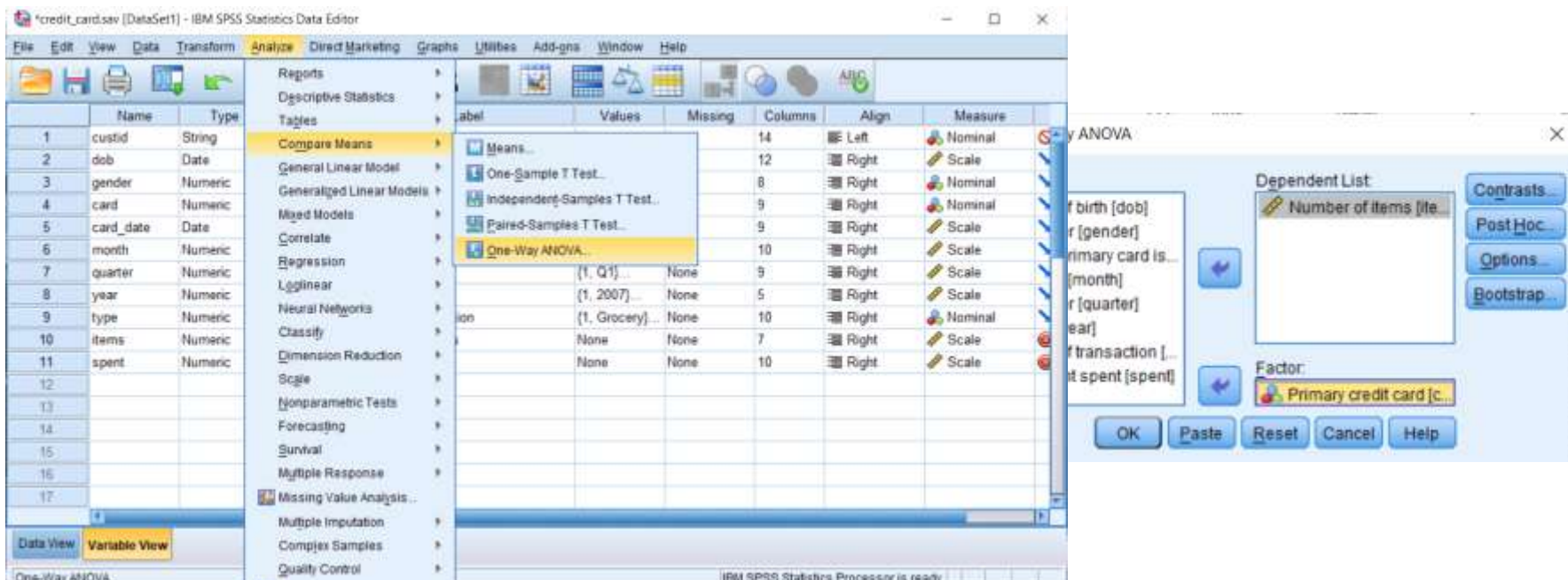
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval
									Lower Bound
Number of items	Equal variances assumed	,638	,424	-1,167	26278	,243	-,037	,032	
	Equal variances not assumed			-1,167	26198,957	,243	-,037	,032	

Ενέργειες 3

- Συγκρίνετε τις μέσες τιμές της ποσοτικής μεταβλητής σε περισσότερες από δύο κατηγορίες ποιοτικής μεταβλητής, ελέγχοντας αν αυτές μπορούν να θεωρηθούν ίσες μεταξύ τους στον πληθυσμό.
- Σχετίζεται το ποσό του λογαριασμού με το πλήθος των αντικειμένων που αγοράσθηκαν; Παρουσιάστε με διάγραμμα και υπολογίστε τον συντελεστή συσχέτισης

Συγκρίνετε τις μέσες τιμές της ποσοτικής μεταβλητής σε περισσότερες από δύο κατηγορίες ποιοτικής μεταβλητής, ελέγχοντας αν αυτές μπορούν να θεωρηθούν ίσες μεταξύ τους στον πληθυσμό



ANOVA

Number of items

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3512,746	4	878,187	134,997	,000
Within Groups	170925,505	26275	6,505		
Total	174438,252	26279			

Πολλαπλές συγκρίσεις για εντοπισμό διαφορών

One-Way ANOVA: Post Hoc Multiple Comparisons

Equal Variances Assumed

LSD S-N-K Waller-Duncan
 Bonferroni Tukey Type I/Type II Error Ratio: 100
 Sidak Tukey's-b Dunnett
 Scheffe Duncan Control Category: Last
 R-E-G-W F Hochberg's GT2 Test
 R-E-G-W Q Gabriel 2-sided < Control > Control

Equal Variances Not Assumed

Tamhane's T2 Dunnett's T3 Games-Howell Dunnett's C

Significance level: 0,05

Continue Cancel Help

geneous Subsets

		Number of items			
		N	Subset for alpha = 0.05		
			1	2	3
Scheffe ^{a,b}	Primary credit card				
	Discover	4080	1,69		
	Other	1800	1,71		
	Mastercard	8520		2,45	
	Visa	6240		2,58	2,58
	American Express	5640			2,67
	Sig.		,995	,252	,622

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 3982,205.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Επιλέξτε μόνο τις συναλλαγές που γίνονται με visa και κάντε έλεγχο υποθέσεων αν το μέσο πλήθος αντικειμένων επηρεάζεται από το φύλο

The screenshot shows the IBM SPSS Statistics Data Editor interface. The 'Data' menu is open, and 'Select Cases...' is highlighted. The main data grid shows columns for quarter, year, type, items, and spent. The status bar indicates 'Visible: 11 of 11 Variables'.

quarter	year	type	items	spent
Q2	2008	Travel	2	295.33
Q2	2008	Other	0	.00
Q2	2007	Grocery	4	234.67
Q2	2007	Retail	6	540.95
Q2	2007	Entertainment	2	305.72
Q2	2007	Travel	5	622.73
Q2	2007	Other	4	410.07
Q2	2008	Grocery	2	121.29
Q2	2008	Retail	2	150.48
Q2	2008	Entertainment	1	59.12
Q2	2008	Travel	2	298.38
Q2	2008	Other	5	469.90
Q3	2007	Grocery	3	188.32
Q3	2007	Retail	10	662.79
Q3	2007	Entertainment	1	143.12
Q3	2007	Travel	1	129.83

The 'Select Cases' dialog box is shown. The 'Date of birth (dob)' variable is selected. The 'If condition is satisfied' radio button is chosen. The 'Filter out unselected cases' option is selected under the 'Output' section.

The 'Select Cases: If' dialog box is shown. The condition 'card=2' is entered in the text box. The 'Primary credit card is...' variable is selected in the list on the left. A numeric keypad is visible for entering the condition.

The screenshot shows the filtered data in the IBM SPSS Statistics Data Editor. The data grid displays columns for customer ID, date of birth, gender, card type, and card data. The 'Primary credit card is...' variable is highlighted in yellow, showing only Visa transactions.

custid	dob	gender	card	card_data	
349	4397-SBL-YUO-317	12/02/1961	Female	Discover	07/25/2000
346	4397-SBL-YUO-317	12/02/1961	Female	Discover	07/25/2000
347	4397-SBL-YUO-317	12/02/1961	Female	Discover	07/25/2000
348	4397-SBL-YUO-317	12/02/1961	Female	Discover	07/25/2000
345	4397-SBL-YUO-317	12/02/1961	Female	Discover	07/25/2000
349	4397-SBL-YUO-317	12/02/1961	Female	Discover	07/25/2000
351	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004
352	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004
353	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004
354	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004
355	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004
356	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004
357	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004
358	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004
359	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004
370	8984-VGPBKS-02G	05/23/1981	Female	Visa	12/16/2004

Analyze, Compare means, independent samples T test

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Number of items	Male	3360	2,49	2,694	,046
	Female	2880	2,68	2,733	,051

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Number of items	Equal variances assumed	1,002	,317	-2,730	6238	,006	-,188	,069
	Equal variances not assumed			-2,727	6066,089	,006	-,188	,069

- Τα αποτελέσματα αφορούν μόνο 6240 άτομα με κάρτα visa από το σύνολο των 26280 ατόμων
- Για να επιλέξουμε ξανά όλα τα δεδομένα
- Data, Select Cases, τσεκάρουμε All cases

Για εξάσκηση

- Βρείτε τα μέτρα θέσεως και διασποράς της μεταβλητής ποσό πιστωτικής κάρτας;
- Υπάρχει διαφορά στα μέτρα θέσεως –διασποράς ανάλογα με τον τύπο της κάρτας; (πίνακας και διάγραμμα)
- Εμφανίστε το ιστόγραμμα της μεταβλητής ποσό πιστωτικής κάρτας;
- Ελέγξτε αν η μέση τιμή της μεταβλητής ποσό πιστωτικής κάρτας, είναι ίδια για άνδρες και γυναίκες
- Στη συνέχεια ελέγξτε αν οι μέσες τιμές ποσού πιστωτικής κάρτας διαφέρουν ανάλογα με το είδος των αγορών.
- Υπάρχει σχέση ανάμεσα στο είδος αγορών και στον τύπο κάρτας;
- Δημιουργείστε μια μεταβλητή που ομαδοποιεί το ποσό σε κατηγορίες, μικρότερο από A, ανάμεσα σε A και B, μεγαλύτερο από B. Τα A, B είναι το πρώτο και τρίτο τεταρτημόριο της μεταβλητής.
- Επιλέξτε μόνο όσες πιστωτικές κάρτες εκδόθηκαν μετά το 2000 και εμφανίστε το ιστόγραμμα του ποσού.