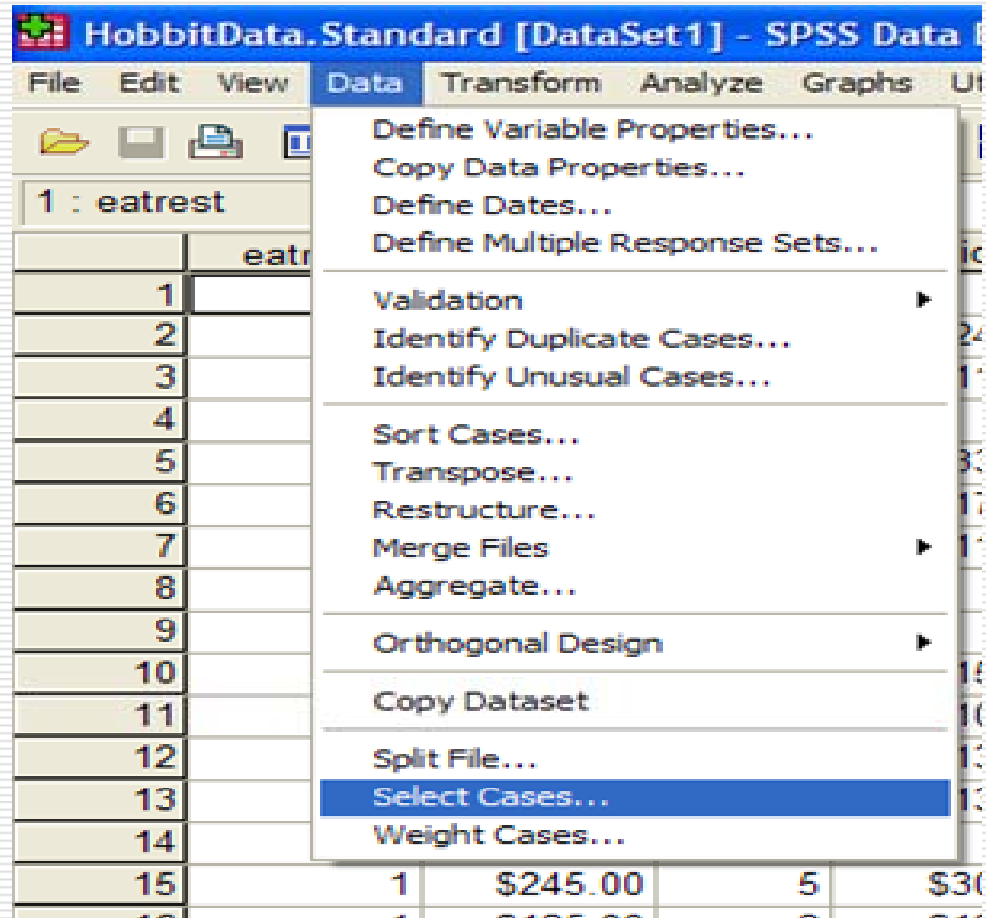


ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ

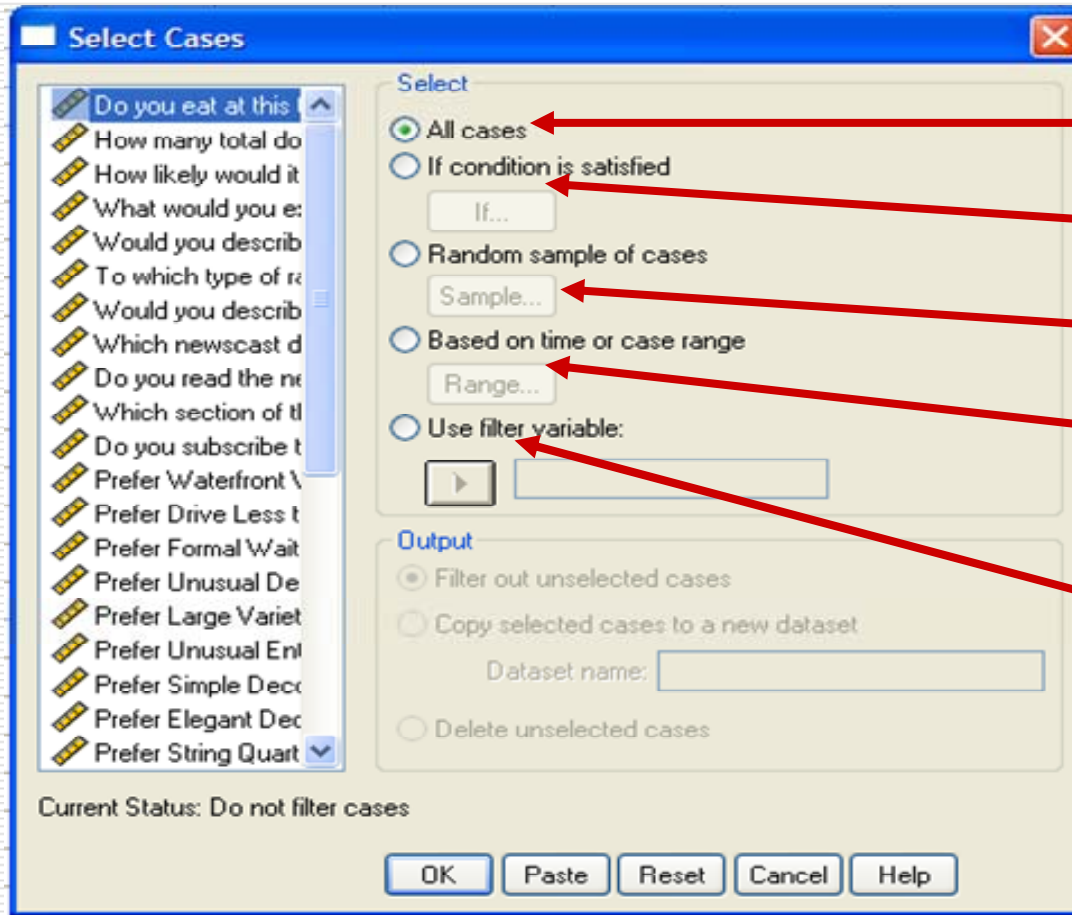
4. Διαχείριση και μετασχηματισμός δεδομένων με το SPSS

Επιλογή δεδομένων

- Για στατιστική ανάλυση ενός υποσυνόλου των δεδομένων (τα υπόλοιπα είτε αγνοούνται είτε διαγράφονται)
- Ορισμός φίλτρου



Κριτήριο επιλογής δεδομένων



Εργαζόμαστε με όλες
τα δεδομένα

Ορισμός φίλτρου με
συνθήκη

Εξαγωγή τυχαίου
δείγματος

Επιλογή από
Χρονοσειρές

Ορισμός μεταβλητής
ως φίλτρο

Ορισμός συνθήκης

The image shows the SPSS 'Select Cases' dialog box with the 'If...' condition selected. A secondary 'Select Cases: If' dialog box is open, showing a list of variables on the left and a central area for building the condition. The central area contains a keypad with arithmetic and logical operators, and a 'Functions' list on the right. Red arrows point from yellow text boxes to these elements.

Συγκρίσεις (Comparisons): Points to the comparison operators (<, >, <=, >=, =, ~=) in the keypad.

Σύνταξη συνθήκης (Condition Syntax): Points to the main text input area in the 'Select Cases: If' dialog.

Συναρτήσεις (Functions): Points to the 'Functions' list containing ABS, ANY, ARSIN, ARTAN, CDFNORM, and CDF.BERNOULLI.

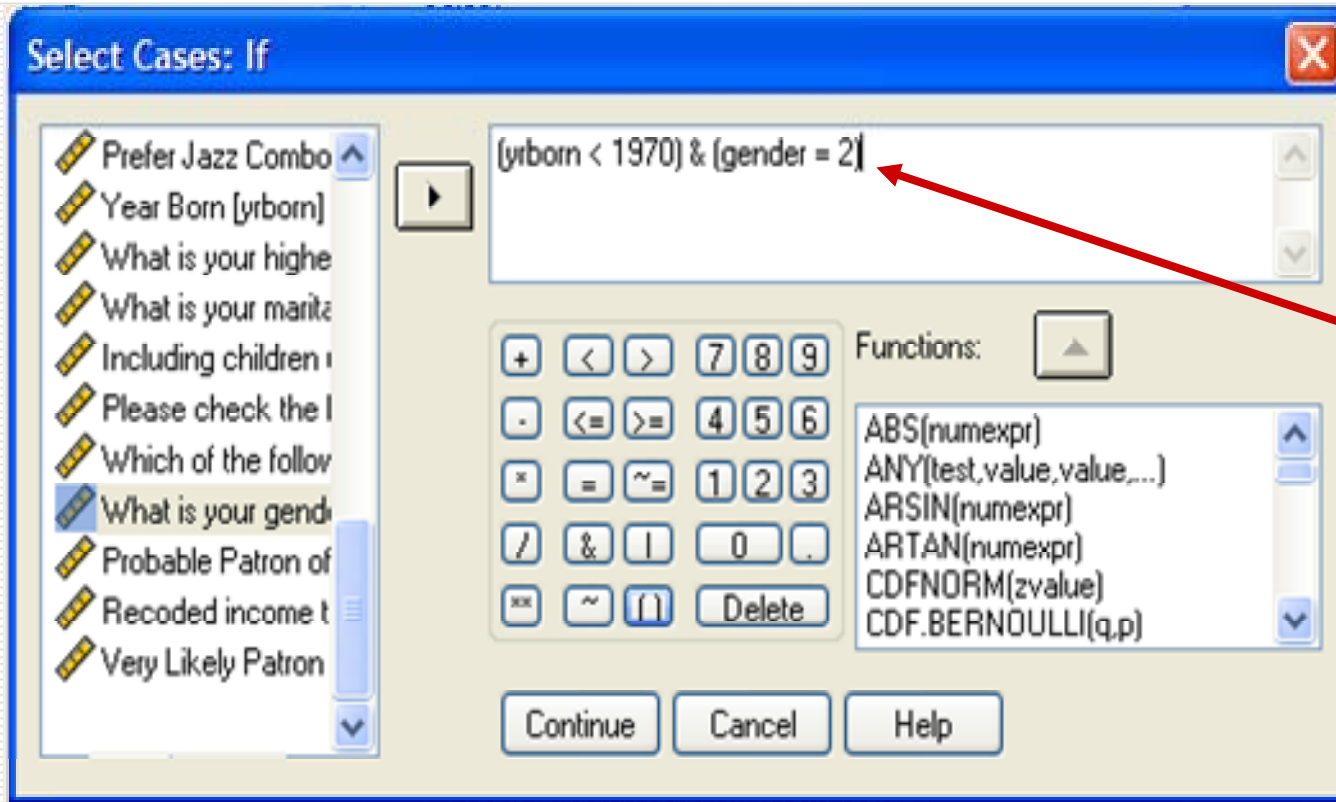
Αριθμητικοί τελεστές (Arithmetic Operators): Points to the arithmetic operators (+, -, *, /, **) in the keypad.

Λογικοί τελεστές (Logical Operators): Points to the logical operators (&, |, ~, ()) in the keypad.

Current Status: Do not filter cases

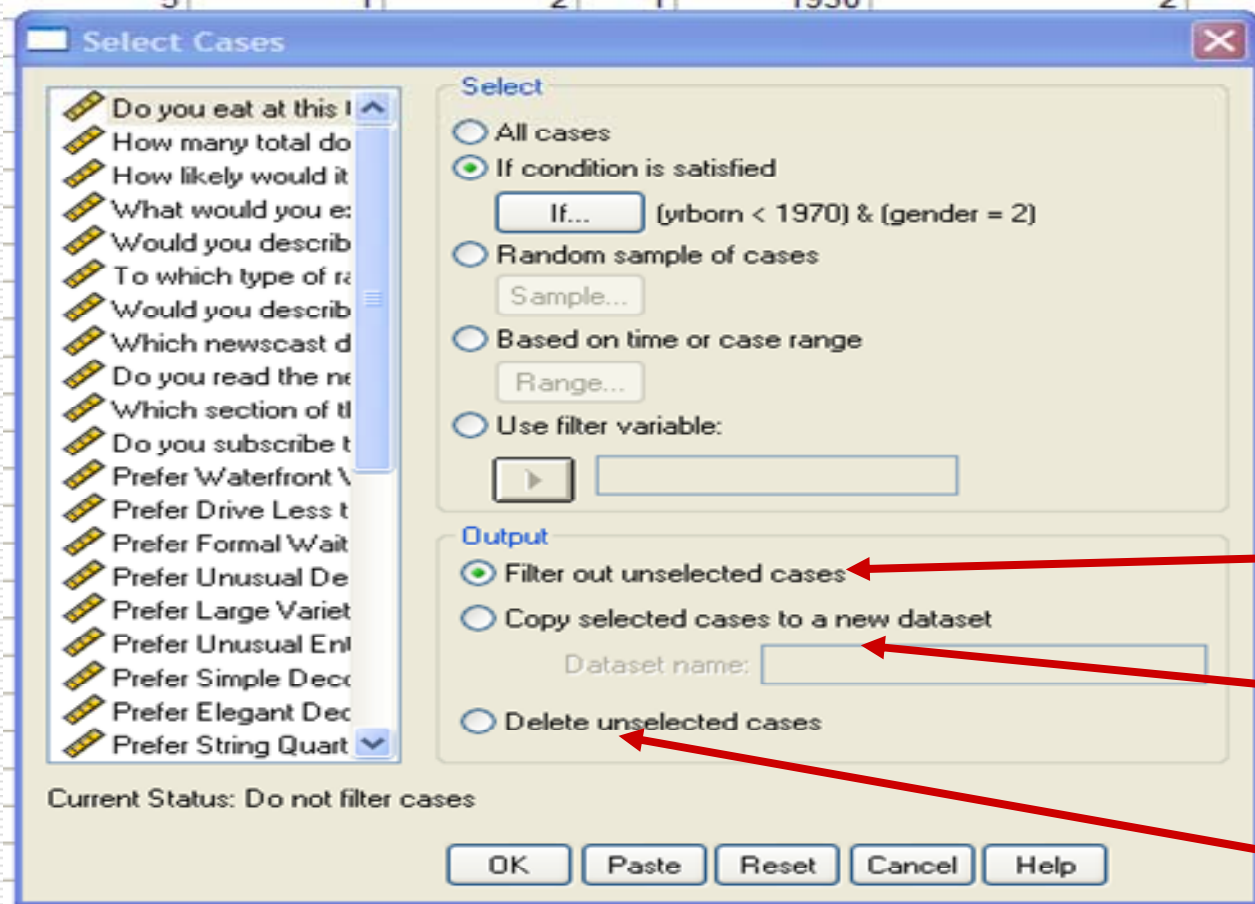
Buttons: OK, Paste, Reset, Cancel, Help

Σύνταξη συνθήκης



**Επιλογή
γυναικών
που έχουν
γεννηθεί
πριν από το
1970**

Αποτελέσματα επιλογής (φίλτρου)



Δημιουργία μεταβλητής - φίλτρου

Δημιουργία νέου αρχείου

Διαγραφή μη επιλεγμένων (!)

**Μεταβλητές
συνθήκης**

Νέα μεταβλητή

	yrborn	educa	marital	famsize	zipcode	income	gender	hcpatron	reincome	VLPatron	filter_\$	
1	1975		2	2	4	4	2	2	20,0	0	0	
2	1947		6	3	2	2	6	2	1	125,0	0	1
3	1968		6	2	4	3	4	1	2	62,5	0	0
4	1930		2	2	2	1	2	2	2	20,0	0	1
5	1951		6	2	2	2	7	1	1	175,0	1	0
6	1965		7	2	2	3	4	2	2	62,5	0	1
7	1968		6	2	2	3	4	1	2	62,5	0	0
8	1973		3	2	4	4	2	1	2	20,0	0	0
9	1975		5	2	2	4	2	2	2	20,0	0	0
10	1959		7	1	7	3	3	1	2	37,5	0	0
11	1959		6	2	2	3	3	2	2	37,5	0	1
12	1962		6	1	3	3	3	2	2	37,5	0	1
13	1957		6	2	4	3	4	1	2	62,5	0	0
14	1973				1	4	2	2	2	20,0	0	0
15	1949		6	1	1	2	6	1	1	125,0	1	0
16	1956		6	2	4	3	4	2	2	62,5	0	1

Επιλεγμένη

Μη επιλεγμένες

Στατιστική σε επιλεγμένα δεδομένα

What is your highest level of education?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Less than High School	11	2,8	2,8	2,8
Some High School	14	3,5	3,5	6,3
High School Graduate	14	3,5	3,5	9,8
Some College (No Degree)	14	3,5	3,5	13,3
Associate Degree	14	3,5	3,5	16,8
Bachelor's Degree	238	59,5	59,5	76,3
Master's Degree	86	21,5	21,5	97,8
Doctorate Degree	9	2,3	2,3	100,0
Total	400	100,0	100,0	

Αφορά όλα τα δεδομένα

What is your highest level of education?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Less than High School	2	1,1	1,1	1,1
Some High School	5	2,7	2,7	3,8
High School Graduate	4	2,2	2,2	6,0
Some College (No Degree)	4	2,2	2,2	8,2
Associate Degree	7	3,8	3,8	12,0
Bachelor's Degree	113	61,7	61,7	73,8
Master's Degree	46	25,1	25,1	98,9
Doctorate Degree	2	1,1	1,1	100,0
Total	183	100,0	100,0	

Αφορά μόνο τα επιλεγμένα δεδομένα (μετά το Select cases)

Αναίρεση διαγραφής

- ❑ Διαγραφή μεταβλητής – φίλτρου
- ❑ Data → Select → All cases
- ❑ Προσοχή στην διαγραφή των μη επιλεγμένων!

Μετασχηματισμός δεδομένων

- ❑ Δημιουργία νέων μεταβλητών από μετασχηματισμούς των υπαρχόντων
- ❑ Επανακωδικοποίηση υπαρχόντων μεταβλητών
- ❑ Κωδικοποίηση αλφαριθμητικών μεταβλητών (string variables)

Η διαδικασία Compute

The image shows the SPSS 'Compute Variable' dialog box. The 'Target Variable' is 'age' and the 'Numeric Expression' is '2006-yrborn'. The 'Function group' is 'Arithmetic'. The 'Functions and Special Variables' list includes 'Abs', 'Arsin', 'Artan', 'Cos', 'Exp', 'Lg10', 'Ln', 'Lgamma', 'Mod', 'Rnd', 'Sin', and 'Sort'. The 'If...' button is also visible.

Ορισμός νέας μεταβλητής (Definition of new variable) - points to the Target Variable field.

Υπολογισμός νέας μεταβλητής (Calculation of new variable) - points to the Numeric Expression field.

Αριθμητικοί & λογικοί τελεστές & συναρτήσεις (Arithmetic & logical operators & functions) - points to the calculator keypad.

Ορισμός συνθήκης επιλογής δεδομένων (Definition of data selection condition) - points to the 'If...' button.

Δημιουργία της μεταβλητής

*HobbitData.Standard [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

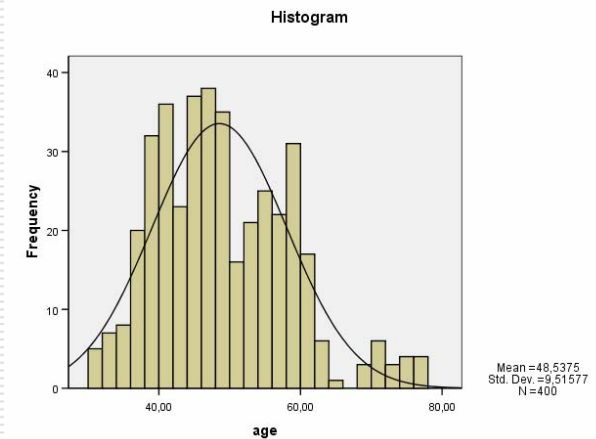
1 : eatrest 1

	gender	hcpatron	reincome	VLPatron	age	var
1	2	2	20,0	0	31,00	
2	2	1	125,0	0	59,00	
3	1	2	62,5	0	38,00	
4	2	2	20,0	0	76,00	
5	1	1	175,0	1	55,00	
6	2	2	62,5	0	41,00	
7	1	2	62,5	0	38,00	
8	1	2	20,0	0	33,00	
9	2	2	20,0	0	31,00	
10	1	2	37,5	0	47,00	
11	2	2	37,5	0	47,00	
12	2	2	37,5	0	44,00	
13	1	2	62,5	0	49,00	
14	2	2	20,0	0	33,00	
15	1	1	125,0	1	57,00	
16	2	2	62,5	0	50,00	
17	1	2	37,5	0	39,00	
18	2	2	20,0	0	35,00	
19	1	2	37,5	0	45,00	

Η νέα μεταβλητή που υπολογίστηκε από την Compute

Statistics

age		
N	Valid	Missing
	400	0
Mean	48,5375	
Minimum	31,00	
Maximum	78,00	



Διαδικασία Recode

The screenshot shows the SPSS Data Editor interface. The 'Transform' menu is open, and the 'Recode' option is selected. A sub-menu is displayed with two options: 'Into Same Variables...' and 'Into Different Variables...'. The 'Into Different Variables...' option is highlighted. In the background, a data table is visible with columns 'lisradio' and 'radprogr'.

	lisradio	radprogr	
	1	1	
	1	4	
	1	3	
	2	.	
	1	2	
	1	3	
	1	3	
	1	1	
	4	4	

Μετασχηματισμός της ίδιας μεταβλητής (!)

Δημιουργία νέας μεταβλητής

(Επανα)κωδικοποίηση μεταβλητής

Standard [DataSet1] - SPSS Data Editor

Transform Analyze Graphs Utilities Window Help

Compute...
Recode
Visual Bander...
Count...
Rank Cases...
Automatic Recode...
Date/Time...
Create Time Series...
Replace Missing Values...
Random Number Generators...
Run Pending Transforms

Recode into Different Variables

Numeric Variable -> Output Variable:
age -> rec_age

Output Variable
Name:
rec_age
Label:
Age Categories
Change

Old and New Values...
If... (optional case selection condition)

OK Paste Reset Cancel Help

Υπάρχουσα και νέα μεταβλητή

Ορισμός νέας μεταβλητής

Ορισμός κωδικοποιήσεων

Ορισμός νέων τιμών

Recode into Different Variables: Old and New Values

Old Value

Value:

System-missing

System- or user-missing

Range:

through

Range, LOWEST through value:

Range, value through HIGHEST:

71

All other values

New Value

Value: 5

System-missing

Copy old value(s)

Old -> New:

Add

Change

Remove

Lowest thru 40 -> 1
41 thru 50 -> 2
51 thru 60 -> 3
61 thru 70 -> 4
71 thru Highest -> 5

Output variables are strings Width: 8

Convert numeric strings to numbers ('5'>5)

Continue Cancel Help

Δημιουργία μεταβλητής

*HobbitData.Standard [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities V

1 : eatrest 1

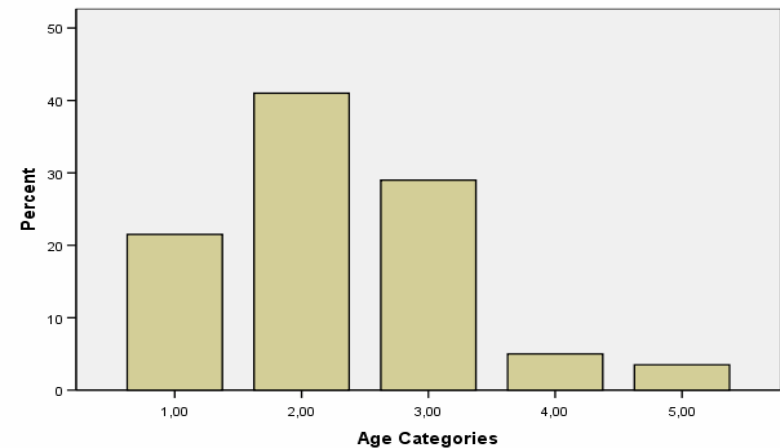
	reincome	VLPatron	age	rec age	year
1	20,0	0	31,00	1,00	
2	125,0	0	59,00	3,00	
3	62,5	0	38,00	1,00	
4	20,0	0	76,00	5,00	
5	175,0	1	55,00	3,00	
6	62,5	0	41,00	2,00	
7	62,5	0	38,00	1,00	
8	20,0	0	33,00	1,00	
9	20,0	0	31,00	1,00	
10	37,5	0	47,00	2,00	
11	37,5	0	47,00	2,00	
12	37,5	0	44,00	2,00	
13	62,5	0	49,00	2,00	
14	20,0	0	33,00	1,00	
15	125,0	1	57,00	3,00	
16	62,5	0	50,00	2,00	
17	37,5	0	39,00	1,00	

Η νέα μεταβλητή που υπολογίστηκε από τη Recode

Age Categories

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1,00	86	21,5	21,5	21,5
2,00	164	41,0	41,0	62,5
3,00	116	29,0	29,0	91,5
4,00	20	5,0	5,0	96,5
5,00	14	3,5	3,5	100,0
Total	400	100,0	100,0	

Age Categories



Εφαρμογή: Γραφική σύγκριση τιμών συνεχούς μεταβλητής ως προς κατηγορική

The image shows the SPSS Data Editor interface. On the left, the 'Graphs' menu is open, and 'Boxplot...' is selected. A red arrow points from this menu item to the 'Boxplot' dialog box. In the 'Boxplot' dialog box, the 'Simple' chart type is selected, and the 'Summaries for groups of cases' radio button is chosen. A second red arrow points from this radio button to the 'Define Simple Boxplot: Summaries for Groups of Cases' dialog box. In this second dialog box, the 'Variable' field contains 'How many total dollars' and the 'Category Axis' field contains 'Age Categories [rec_age]'. Red text labels 'Συνεχής μεταβλητή' (Continuous variable) and 'Κατηγορική μεταβλητή' (Categorical variable) point to these two fields respectively. The 'Define Simple Boxplot' dialog box also shows a list of variables on the left and options for 'Panel by' (Rows and Columns) on the right.

