

# ΒΙΟΣΤΑΤΙΣΤΙΚΗ (ΕΡΓΑΣΤΗΡΙΟ)

με χρήση του Στατιστικού Πακέτου SPSS v19

ΓΕΝΙΚΟ ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

Τουβιάς Θωράς ΕΠΙΣΤ. ΣΥΝΕΡΓ.  
Παπαγεωργίου Έφη ΕΠΙΚ. ΚΑΘ.

ΚΕΦΑΛΑΙΟ 2 Περιγραφική Στατιστική







**Εφαρμογή 2.1.** Να υπολογιστούν οι πίνακες και τα ιστογράμματα συχνοτήτων των παρατηρήσεων της **Εφαρμογής 1.2** καθώς και τα δειγματικά μέτρα θέσης και διασποράς αυτών.

- Από το μενού **Analyze** επιλέγουμε **Descriptive Statistics** > **Frequencies**<sup>[A]</sup> και στο παράθυρο διαλόγου, επιλέγοντας όλες τις μεταβλητές μας (αριστερή λίστα)<sup>[B]</sup>, τις τοποθετούμε (με το βέλος απόθεσης) στην λίστα των επιλεγμένων μεταβλητών **Variable(s)**<sup>[Γ]</sup>. Ελέγχουμε επίσης ότι είναι τσεκαρισμένη η ενέργεια **Display frequency tables**. Επιλέγοντας στη συνέχεια **Statistics ...** >

The image shows a sequence of three screenshots from the SPSS software interface, illustrating the steps to perform a frequency analysis:

- [A]** The **Analyze** menu is open, and **Descriptive Statistics** > **Frequencies...** is selected.
- [B]** The **Frequencies** dialog box is shown. The variables **Επάγγελμα [Job]**, **Ημερομίσθιο [Wage]**, and **Οικ. Κατάσταση [...]** are selected in the left list. The **Display frequency tables** checkbox is checked.
- [Γ]** The **Statistics** sub-dialog box is open, showing the same variables in the **Variable(s)** list. The **Display frequency tables** checkbox is also checked.





► ... εμφανίζεται νέο παράθυρο διαλόγου<sup>[Δ]</sup>, στο οποίο για τα δειγματικά μέτρα θέσης τσεκάρουμε τις επιλογές: **Quartiles** (τεταρτημόρια), **Mean** (μέση τιμή), **Median** (διάμεσος) και **Mode** (επικρατούσα τιμή), ενώ για τα δειγματικά μέτρα διασποράς τσεκάρουμε αντίστοιχα: **Std. deviation** (τυπική απόκλιση), **Variance** (διασπορά), **Range** (εύρος τιμών), **Minimum** και **Maximum** (ελάχιστη και μέγιστη τιμή). Με **Continue** επιστρέφουμε στο προηγούμενο παράθυρο διαλόγου και επιλέγοντας **Charts**<sup>[Ε]</sup>, τσεκάρουμε **Bar charts**<sup>[Ζ]</sup>. Επικυρώνουμε με **Continue** και τέλος με **OK**.

The image shows three overlapping dialog boxes from the SPSS software interface, illustrating the steps to generate bar charts for frequency statistics.

- Frequencies: Statistics** (labeled [Δ]): This dialog box is used to select the statistical measures to be calculated. Under "Central Tendency", the options Mean, Median, and Mode are checked. Under "Dispersion", Std. deviation, Variance, Range, Minimum, and Maximum are checked. The "Continue" button is highlighted with a yellow arrow.
- Frequencies** (labeled [Ε]): This is the main dialog box for the Frequencies procedure. The "Charts..." button is highlighted with a yellow arrow.
- Frequencies: Charts** (labeled [Ζ]): This sub-dialog box allows the user to choose the chart type. "Bar charts" is selected under the "Chart Type" section. The "Continue" button is highlighted with a yellow arrow.

Yellow arrows indicate the flow of the process: from the "Continue" button in the Statistics dialog to the "Charts" button in the main Frequencies dialog, and then to the "Continue" button in the Charts sub-dialog.



Εμφανίζεται λοιπόν στον **Viewer** ο πίνακας των (επιλεγμένων) περιγραφικών στατιστικών <sup>[A]</sup>, οι 3 πίνακες συχνοτήτων για κάθε μια μεταβλητή <sup>[B]</sup> καθώς και τα 3 αντίστοιχα ιστογράμματα.

		Επάγγελμα	Ημερομίσθιο	Οικ. Κατάσταση
N	Valid	20	20	20
	Missing	0	0	0
Mean			75,75	1,90
Median			75,00	2,00
Mode			70 <sup>a</sup>	2
Std. Deviation			13,599	1,071
Variance			184,934	1,147
Range			50	4
Minimum			50	0
Maximum			100	4
Percentiles	25		66,25	1,00
	50		75,00	2,00
	75		87,50	2,00

a. Multiple modes exist. The smallest value is shown

[A]

**Επάγγελμα**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Δάσκαλος	4	20,0	20,0	20,0
Δημ. Υπάλληλος	6	30,0	30,0	50,0
Εργάτης	6	30,0	30,0	80,0
Ιερέας	2	10,0	10,0	90,0
Οδηγός	2	10,0	10,0	100,0
Total	20	100,0	100,0	

**Ημερομίσθιο**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 50	1	5,0	5,0	5,0
60	3	15,0	15,0	20,0
65	1	5,0	5,0	25,0
70	4	20,0	20,0	45,0
75	2	10,0	10,0	55,0
80	4	20,0	20,0	75,0
90	3	15,0	15,0	90,0
100	2	10,0	10,0	100,0
Total	20	100,0	100,0	

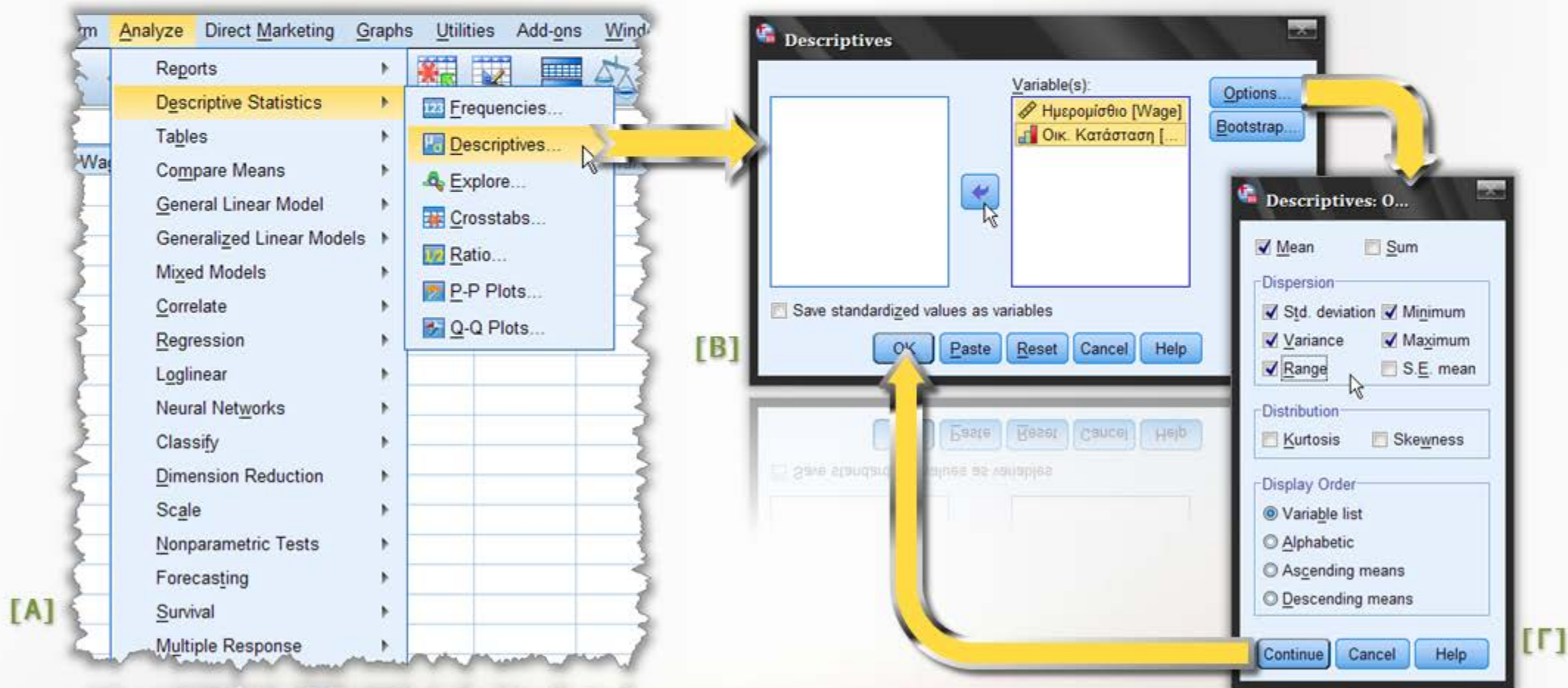
**Οικ. Κατάσταση**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Άτεκνος	2	10,0	10,0	10,0
1 τέκνο	4	20,0	20,0	30,0
2 τέκνα	10	50,0	50,0	80,0
Τρίτεκνος	2	10,0	10,0	90,0
Πολύτεκνος	2	10,0	10,0	100,0
Total	20	100,0	100,0	

[B]

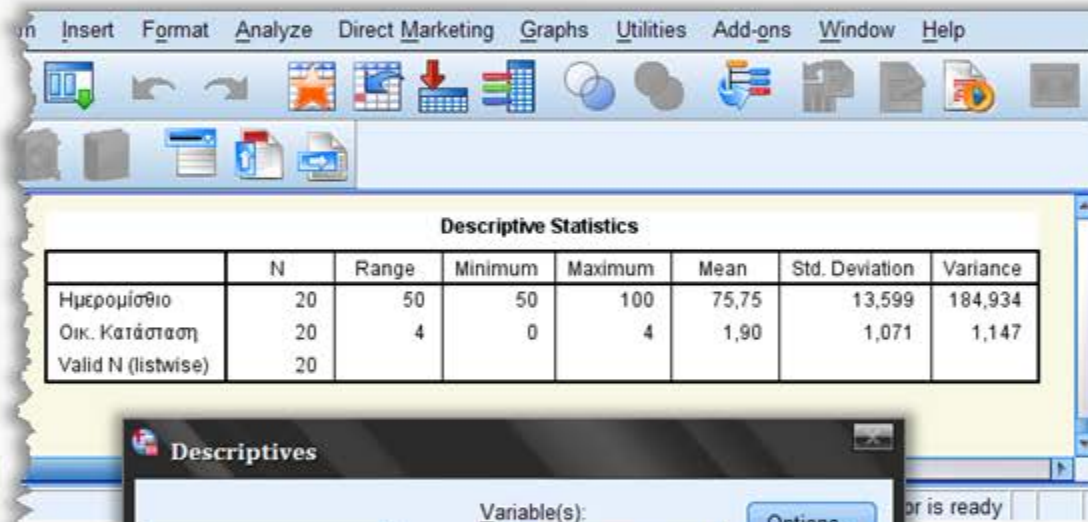


Εναλλακτικά, εάν δεν επιθυμούσαμε διαγράμματα ή εκατοστημόρια (όπως τα τεταρτημόρια ή η διάμεσος), μπορούμε, πιο άμεσα, από το menu **Analyze** να επιλέξουμε **Descriptive Statistics** > **Descriptives**<sup>[A]</sup>. Στο εμφανιζόμενο παράθυρο διαλόγου μεταφέρουμε από τη λίστα όλων των (μη-ποιοτικών) μεταβλητών μας αυτές που επιθυμούμε (με το βέλος απόθεσης) στη λίστα **Variable(s)** των επιλεγμένων<sup>[B]</sup>. Με χρήση του **Options** τσεκάρουμε στο νέο παράθυρο διαλόγου τις ενέργειες **Mean**, **Std. deviation**, **Variance**, **Range**, **Minimum** και **Maximum**<sup>[Γ]</sup>. Επικυρώνουμε τις επιλογές μας με **Continue** και τέλος με **OK**.



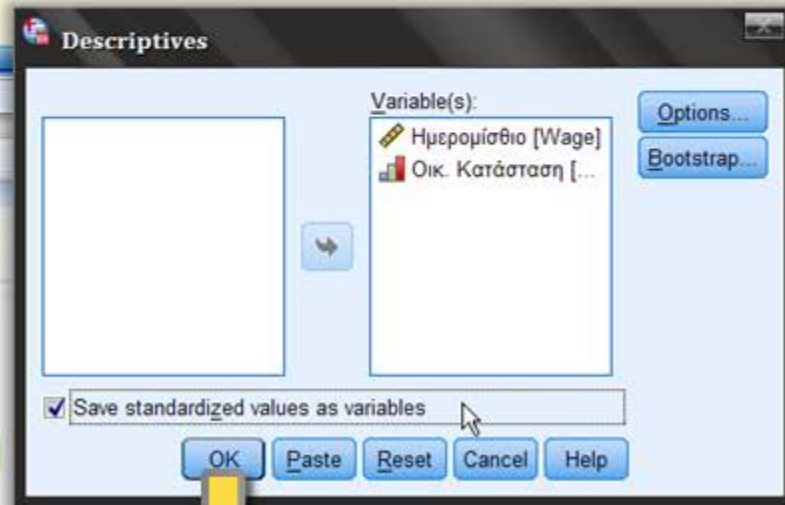


Εμφανίζεται λοιπόν στον **Viewer** ο πίνακας των (επιλεγμένων) περιγραφικών στατιστικών<sup>[A]</sup>. Εάν στο προηγούμενο παράθυρο διαλόγου **Descriptives** είχαμε επιλέξει επιπλέον την ενέργεια **Save standardized values as variables**<sup>[B]</sup> τότε δημιουργούνται νέες μεταβλητές (όσες και οι επιλεγμένες) οι οποίες περιέχουν τις αντίστοιχες κανονικοποιημένες τιμές της απόστασης (διαφοράς) των παρατηρήσεων από τους μέσους όρους των αντίστοιχων μεταβλητών.



Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
Ημερομίσθιο	20	50	50	100	75,75	13,599	184,934
Οικ. Κατάσταση	20	4	0	4	1,90	1,071	1,147
Valid N (listwise)	20						



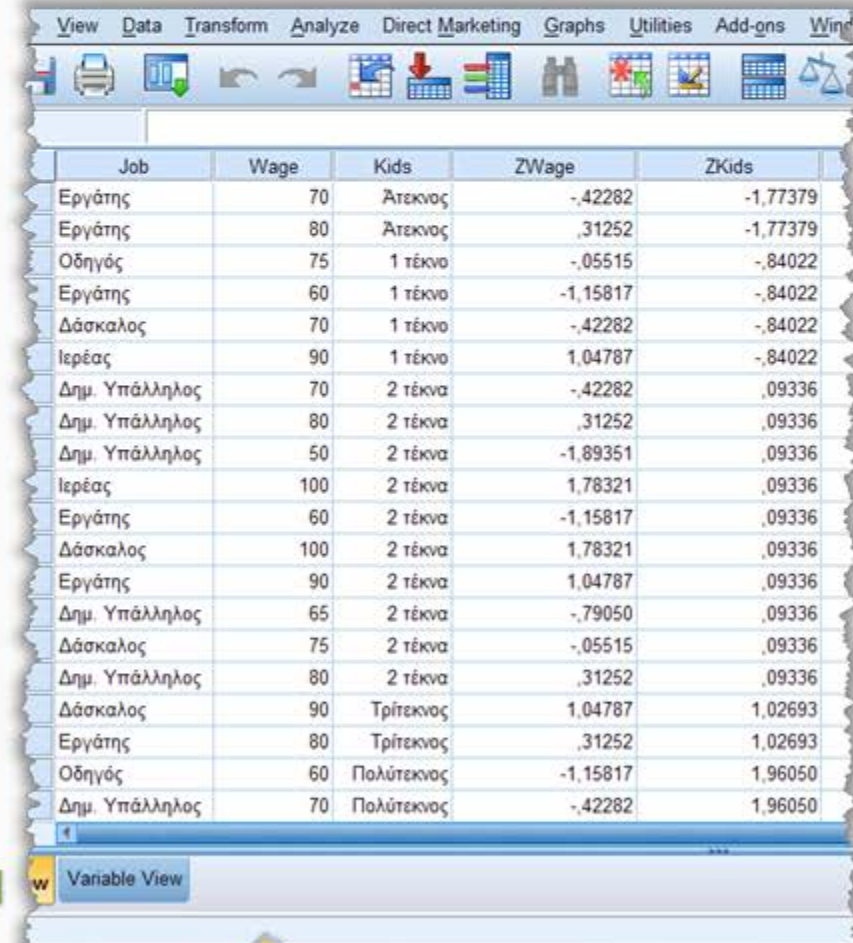
Descriptives

Variable(s):

- Ημερομίσθιο [Wage]
- Οικ. Κατάσταση [...]

Save standardized values as variables

OK Paste Reset Cancel Help

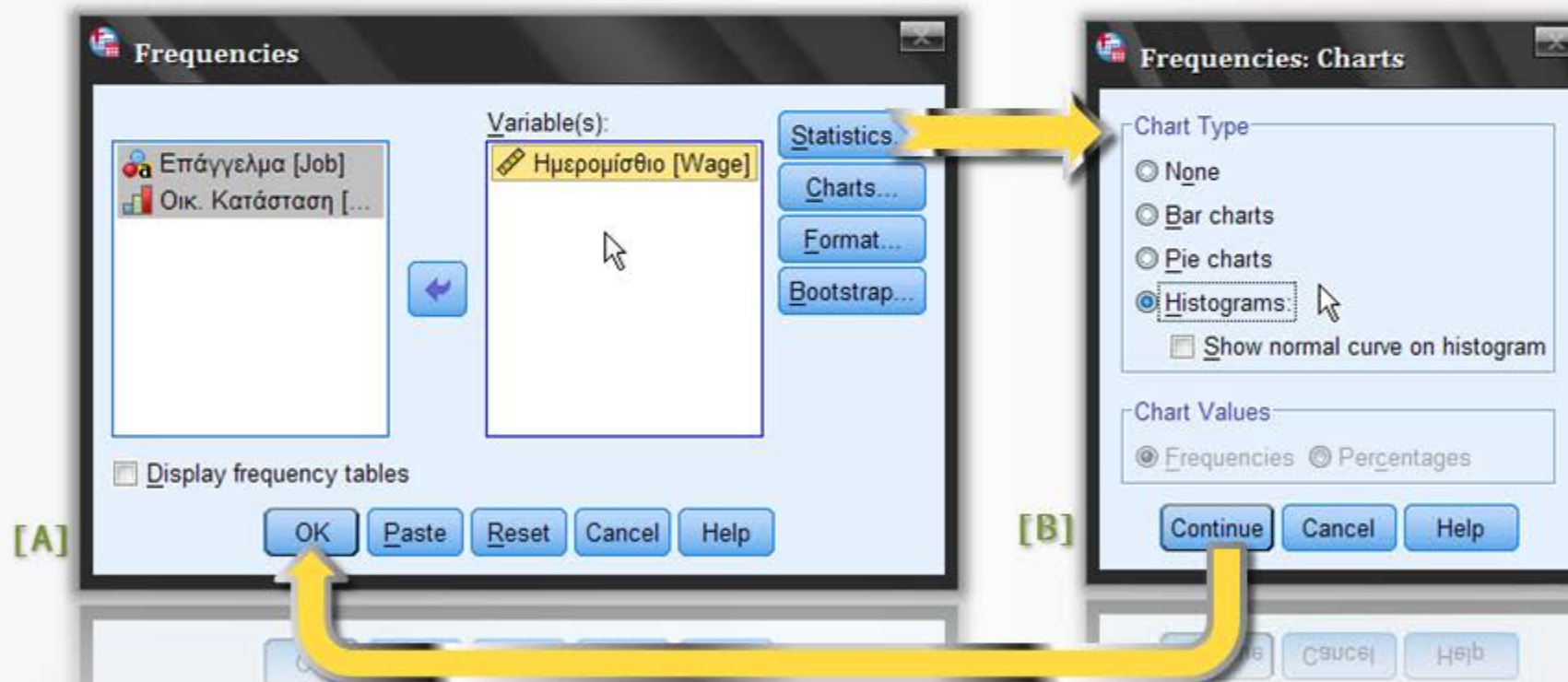


Job	Wage	Kids	ZWage	ZKids
Εργάτης	70	Άτεκνος	-.42282	-.177379
Εργάτης	80	Άτεκνος	.31252	-.177379
Οδηγός	75	1 τέκνο	-.05515	-.84022
Εργάτης	60	1 τέκνο	-1,15817	-.84022
Δάσκαλος	70	1 τέκνο	-.42282	-.84022
Ιερέας	90	1 τέκνο	1,04787	-.84022
Δημ. Υπάλληλος	70	2 τέκνα	-.42282	.09336
Δημ. Υπάλληλος	80	2 τέκνα	.31252	.09336
Δημ. Υπάλληλος	50	2 τέκνα	-1,89351	.09336
Ιερέας	100	2 τέκνα	1,78321	.09336
Εργάτης	60	2 τέκνα	-1,15817	.09336
Δάσκαλος	100	2 τέκνα	1,78321	.09336
Εργάτης	90	2 τέκνα	1,04787	.09336
Δημ. Υπάλληλος	65	2 τέκνα	-.79050	.09336
Δάσκαλος	75	2 τέκνα	-.05515	.09336
Δημ. Υπάλληλος	80	2 τέκνα	.31252	.09336
Δάσκαλος	90	Τρίτεκνος	1,04787	1,02693
Εργάτης	80	Τρίτεκνος	.31252	1,02693
Οδηγός	60	Πολύτεκνος	-1,15817	1,96050
Δημ. Υπάλληλος	70	Πολύτεκνος	-.42282	1,96050



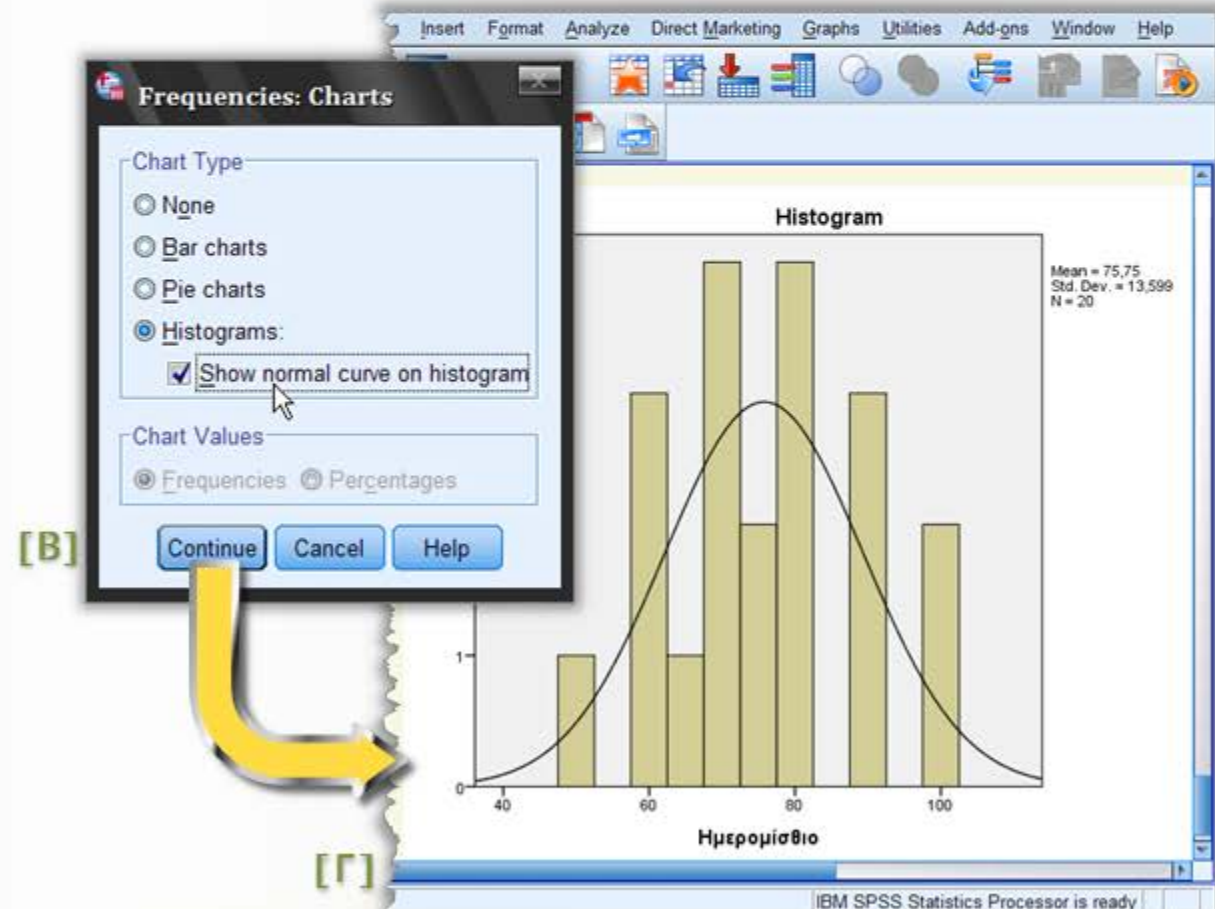
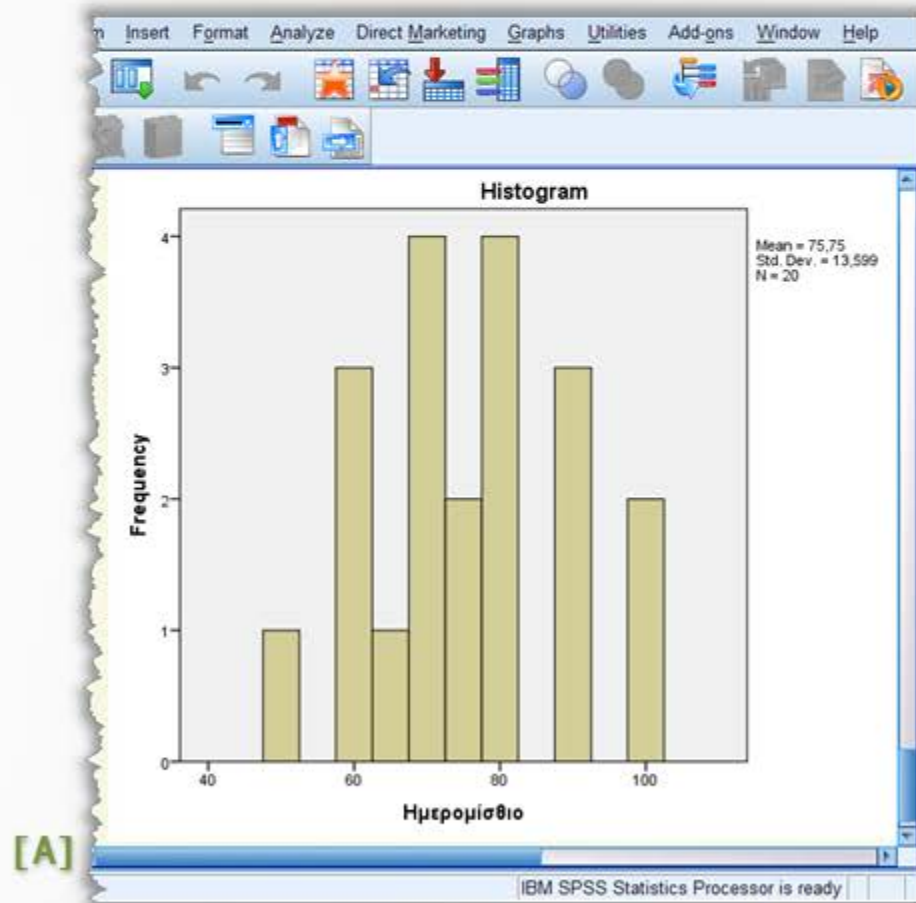


**ΣΗΜΕΙΩΣΗ:** Στην περίπτωση που μια μεταβλητή μας εμπεριέχει πολλές και διαφορετικές αριθμητικές τιμές, όπως στην περίπτωσή μας η μεταβλητή **Wage**, για το ιστόγραμμα συχνοτήτων της, όταν την επιλέγουμε στο παράθυρο διαλόγου **Frequencies<sup>[B]</sup>** (και έχοντας από-επιλέξει την ενέργεια **Display frequency tables**), είναι προτιμότερο στο παράθυρο διαλόγου **Frequencies: Charts** να τσεκάρουμε την ενέργεια **Histograms**. Επικυρώνουμε πάλι τις επιλογές μας με **Continue** και τέλος με **OK**.





Κατά αυτόν τον τρόπο γίνεται αυτόματα η κατηγοριοποίηση των τιμών της επιλεγμένης μεταβλητής (στο παράδειγμά μας η **Wage**) ενώ επάνω δεξιά (στο ιστόγραμμα του **Viewer**) εμφανίζονται η μέση τιμή, η τυπική απόκλιση και το πλήθος των παρατηρήσεων<sup>[Α]</sup>. Εάν επιπλέον, στο παράθυρο διαλόγου **Frequencies: Charts**, είχαμε επιλέξει και την ενέργεια **Show normal curve on histogram**<sup>[Β]</sup>, τότε μαζί με το ιστόγραμμα θα εμφανίζονταν και η συνάρτηση πυκνότητας πιθανότητας<sup>[Γ]</sup> με μέση τιμή και τυπική απόκλιση τις εμφανιζόμενες δειγματικές.



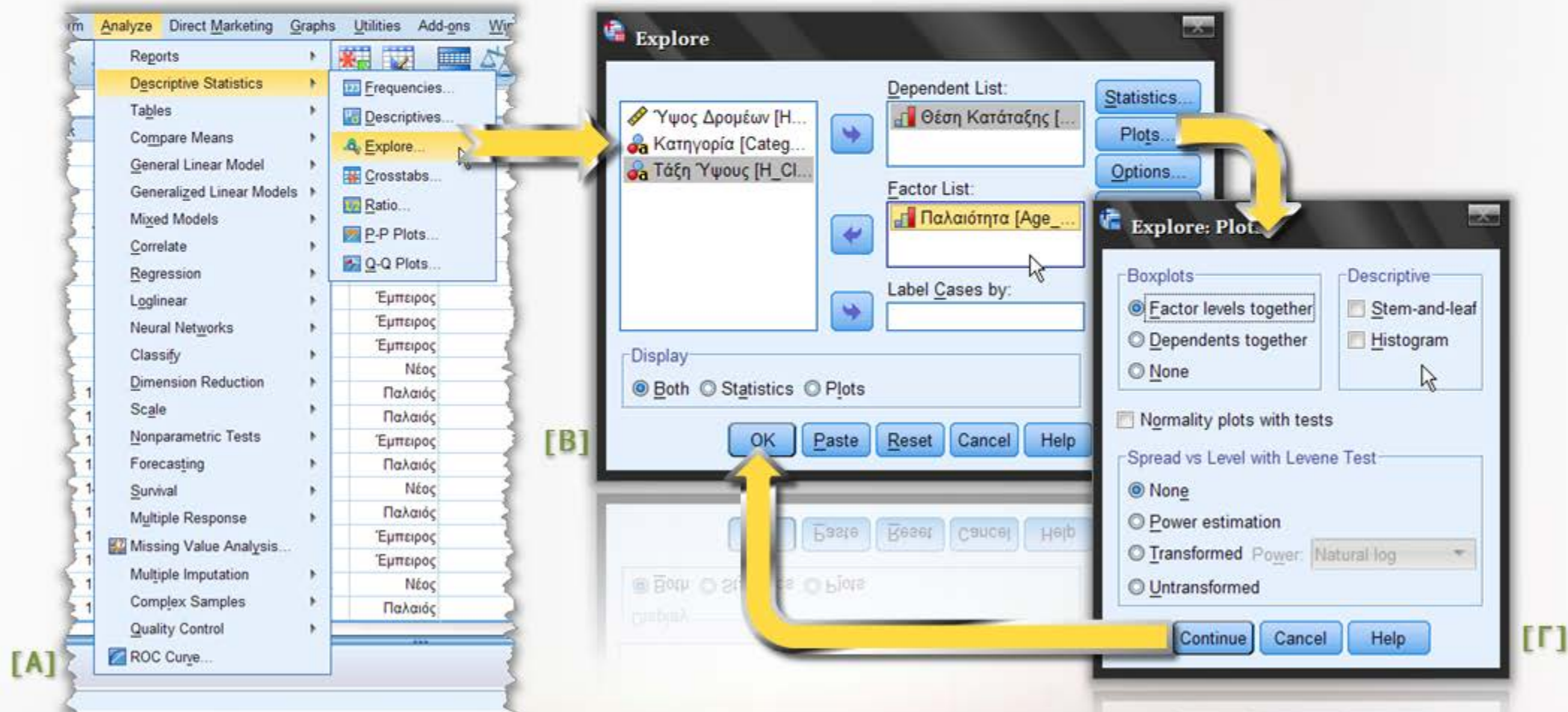




**Εφαρμογή 2.2.** Στην **Εφαρμογή 1.4** να περιγραφούν στατιστικά οι θέσεις κατάταξης των αθλητών ανά κατηγορία εμπειρίας και να δημιουργηθούν τα αντίστοιχα διαγράμματα *boxplots*. Επίσης, να γίνει η στατιστική περιγραφή των υψών τους ανά κατηγορία κατάταξης (μετάλλια, προκρινόμενοι, αποκλειόμενοι) και να παρουσιαστεί το αντίστοιχο διάγραμμα *boxplot*.

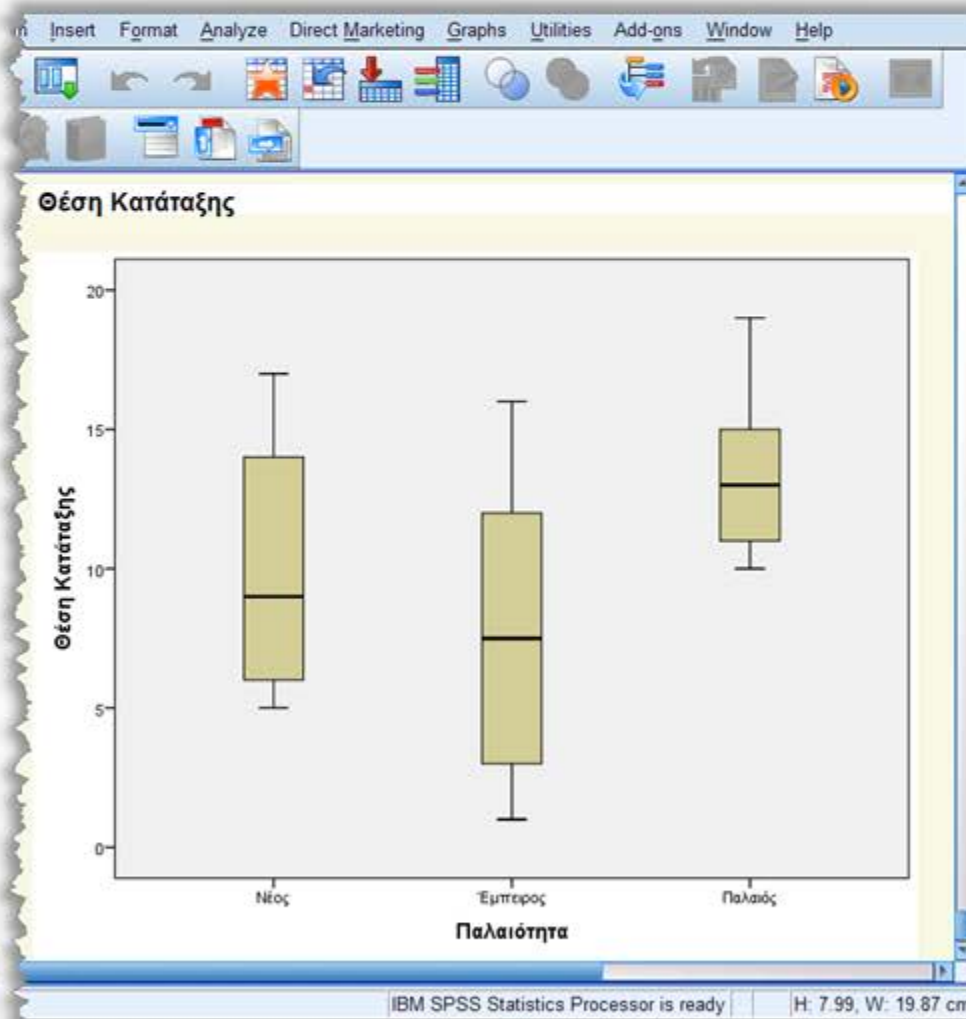


Από το menu **Analyze** επιλέγουμε **Descriptive Statistics** > **Explore<sup>[A]</sup>** και στο εμφανιζόμενο παράθυρο διαλόγου τοποθετούμε (με τα αντίστοιχα βέλη απόθεσης) τις μεταβλητές **Θέση Κατάταξης [Rank]** και **Παλαιότητα [Age\_Class]** στις λίστες **Dependent List** και **Factor List** αντίστοιχα<sup>[B]</sup>. Επιλέγοντας στη συνέχεια **Plots** στο νέο παράθυρο διαλόγου από-επιλέγουμε τις ενέργειες **Stem-and-leaf** και **Histogram<sup>[Γ]</sup>**. Επικυρώνουμε με **Continue** και τέλος με **OK**.





Εμφανίζεται λοιπόν στον **Viewer** ο πίνακας των ζητούμενων περιγραφικών στατιστικών<sup>[A]</sup> καθώς και τα boxplots<sup>[B]</sup> της μεταβλητής **Rank** για κάθε κατηγορία της μεταβλητής **Age\_Class**.



[B]

Descriptives

Παλαιότητα		Statistic	Std. Error
Θέση Κατάταξης	Νέος	Mean	10,20
		95% Confidence Interval for Mean	
		Lower Bound	3,78
		Upper Bound	16,62
		5% Trimmed Mean	10,11
		Median	9,00
		Variance	26,700
		Std. Deviation	5,167
		Minimum	5
		Maximum	17
	Range	12	
	Interquartile Range	10	
	Skewness	,462	,913
	Kurtosis	-2,054	2,000
Έμπειρος	Mean	7,70	1,732
		95% Confidence Interval for Mean	
		Lower Bound	3,78
		Upper Bound	11,62
		5% Trimmed Mean	7,61
		Median	7,50
		Variance	30,011
		Std. Deviation	5,478
		Minimum	1
		Maximum	16
	Range	15	
	Interquartile Range	10	
	Skewness	,492	,687
	Kurtosis	-1,049	1,334
Παλιός	Mean	13,60	1,600
		95% Confidence Interval for Mean	
		Lower Bound	9,16
		Upper Bound	18,04
		5% Trimmed Mean	13,50
		Median	13,00
		Variance	12,800
		Std. Deviation	3,578
		Minimum	10
		Maximum	19
	Range	9	
	Interquartile Range	7	
	Skewness	,871	,913
	Kurtosis	,148	2,000

[A]



- Για το δεύτερο ζητούμενο, αλλάζουμε πρώτα τις τιμές **Χρυσό Μετάλλιο**, **Αργυρό Μετάλλιο** και **Χάλκινο Μετάλλιο** σε **Μετάλλια**.

Rank	Category	H_Class	Age_Class
1	Μετάλλια	170-179	Έμπειρος
2	Μετάλλια	170-179	Έμπειρος
3	Μετάλλια	170-179	Έμπειρος
4	Qualified	Κάτω από 1,6μ.	Έμπειρος
5	Qualified	180-190	Νέος
6	Qualified	Κάτω από 1,6μ.	Νέος
7	Qualified	180-190	Έμπειρος
8	Qualified	180-190	Έμπειρος
8	Qualified	180-190	Έμπειρος
9	Not-Qualified	170-179	Νέος
10	Not-Qualified	190-200	Παλαιός
11	Not-Qualified	170-179	Παλαιός
12	Not-Qualified	160-170	Έμπειρος
13	Not-Qualified	160-170	Παλαιός
14	Not-Qualified	170-179	Νέος
15	Not-Qualified	190-200	Παλαιός
16	Not-Qualified	170-179	Έμπειρος
16	Not-Qualified	160-170	Έμπειρος
17	Not-Qualified	Κάτω από 1,6μ.	Νέος
19	Not-Qualified	190-200	Παλαιός



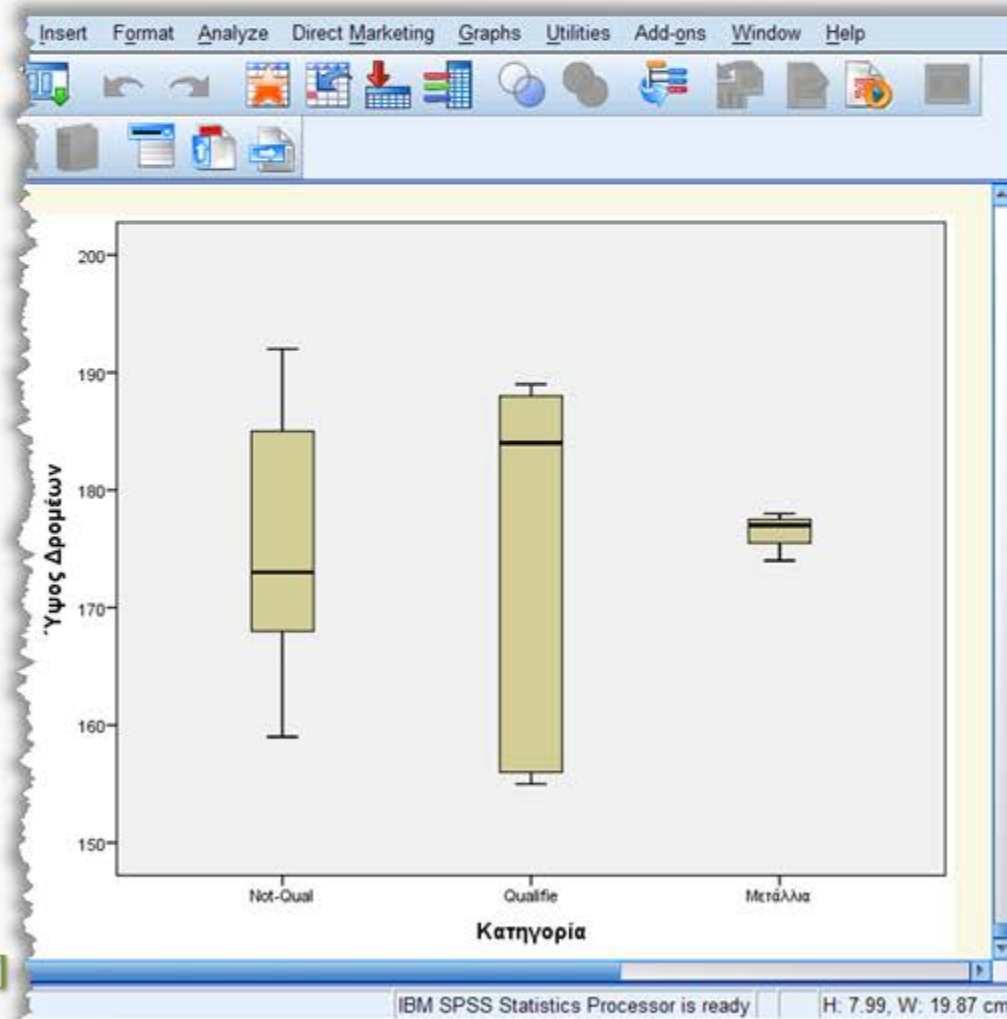
Στη συνέχεια, από το menu **Analyze** επιλέγουμε άλλη μια φορά **Descriptive Statistics** ➤ **Explore**<sup>[A]</sup> και στο εμφανιζόμενο παράθυρο διαλόγου τοποθετούμε (με τα αντίστοιχα βέλη απόθεσης) τις μεταβλητές **Ύψος Δρομέων [Height]** και **Κατηγορία [Category]** στις λίστες **Dependent List** και **Factor List** αντίστοιχα<sup>[B]</sup>. Επιλέγοντας στη συνέχεια **Plots** στο νέο παράθυρο διαλόγου τσεκάρουμε την ενέργεια **Histogram**<sup>[Γ]</sup>. Επικυρώνουμε με **Continue** και τέλος με **OK**.

The image shows a sequence of three screenshots from the SPSS software interface, illustrating the steps to generate a histogram for descriptive statistics.

- [A]** The first screenshot shows the **Analyze** menu with **Descriptive Statistics** expanded, and **Explore...** selected. A yellow arrow points from this menu item to the next dialog box.
- [B]** The second screenshot shows the **Explore** dialog box. The **Dependent List** contains 'Θέση Κατάταξης [...]' and 'Τάξη Ύψους [H\_Cl...]', and the **Factor List** contains 'Κατηγορία [Categ...]'.
- [Γ]** The third screenshot shows the **Explore: Plots** dialog box. The **Factor levels together** radio button is selected, and the **Histogram** checkbox is checked.



Εμφανίζεται λοιπόν στον **Viewer** ο πίνακας των ζητούμενων περιγραφικών στατιστικών<sup>[A]</sup> καθώς και τα ιστογράμματα και τα boxplots<sup>[B]</sup> της μεταβλητής **Height** για κάθε κατηγορία της μεταβλητής **Category**.



[B]

Κατηγορία		Statistic	Std. Error
Υψος Δρομέων	Not-Qual	Mean	176,09
	95% Confidence Interval for Mean	Lower Bound	168,35
		Upper Bound	183,84
	5% Trimmed Mean	176,16	
	Median	173,00	
	Variance	132,891	
	Std. Deviation	11,528	
	Minimum	159	
	Maximum	192	
	Range	33	
	Interquartile Range	24	
Skewness	,302	,661	
Kurtosis	-1,174	1,279	
Qualife	Mean	176,00	6,583
	95% Confidence Interval for Mean	Lower Bound	159,08
		Upper Bound	192,92
	5% Trimmed Mean	176,44	
	Median	184,00	
	Variance	260,000	
	Std. Deviation	16,125	
	Minimum	155	
	Maximum	189	
	Range	34	
	Interquartile Range	33	
Skewness	-,850	,845	
Kurtosis	-1,911	1,741	
Μετάλλια	Mean	176,33	1,202
	95% Confidence Interval for Mean	Lower Bound	171,16
		Upper Bound	181,50
	5% Trimmed Mean	-	
	Median	177,00	
	Variance	4,333	
	Std. Deviation	2,082	
	Minimum	174	
	Maximum	178	
	Range	4	
	Interquartile Range	-	
Skewness	-1,293	1,225	
Kurtosis	-		

[A]