



Βιοστατιστική (Ε)

Ενότητα 2: Περιγραφική στατιστική

Δρ.Ευσταθία Παπαγεωργίου, Αναπληρώτρια Καθηγήτρια
Τμήμα Ιατρικών Εργαστηρίων



Το περιεχόμενο του μαθήματος διατίθεται με άδεια Creative Commons εκτός και αν αναφέρεται διαφορετικά



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



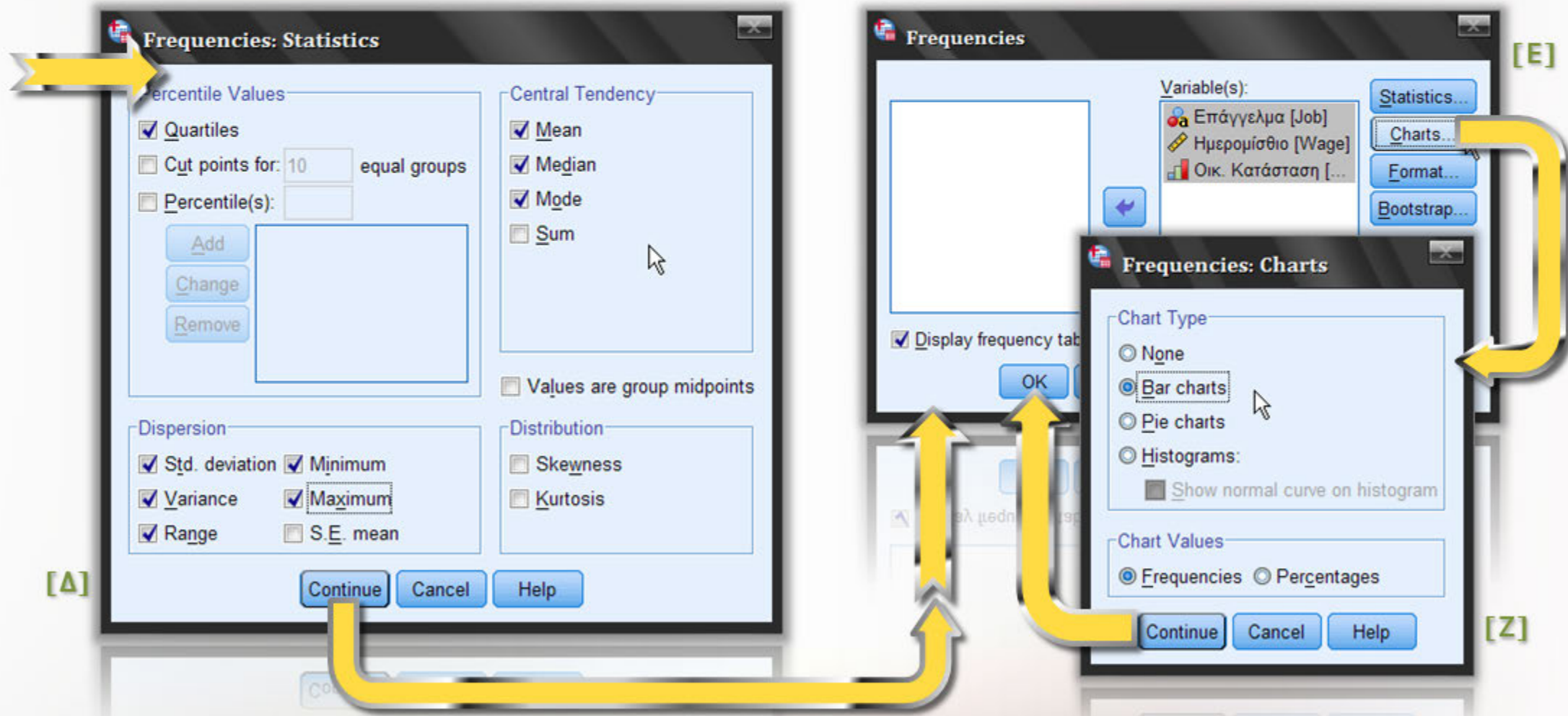
Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

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

The image illustrates the steps to access the Frequencies dialog box in SPSS. It shows the menu path from **Analyze** to **Descriptive Statistics** to **Frequencies**. The dialog box is shown with the variables **Επάγγελμα [Job]**, **Ημερομίσθιο [Wage]**, and **Οικ. Κατάσταση [...]** selected in the left list. The **Display frequency tables** checkbox is checked. The **Statistics...** button is highlighted, indicating the next step in the process.

► ... εμφανίζεται νέο παράθυρο διαλόγου^[Δ], 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**^[Ε], τσεκάρουμε **Bar charts**^[Ζ]. Επικυρώνουμε με **Continue** και τέλος με **OK**.



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

Statistics

		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 ^a	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]

Frequency Table

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]

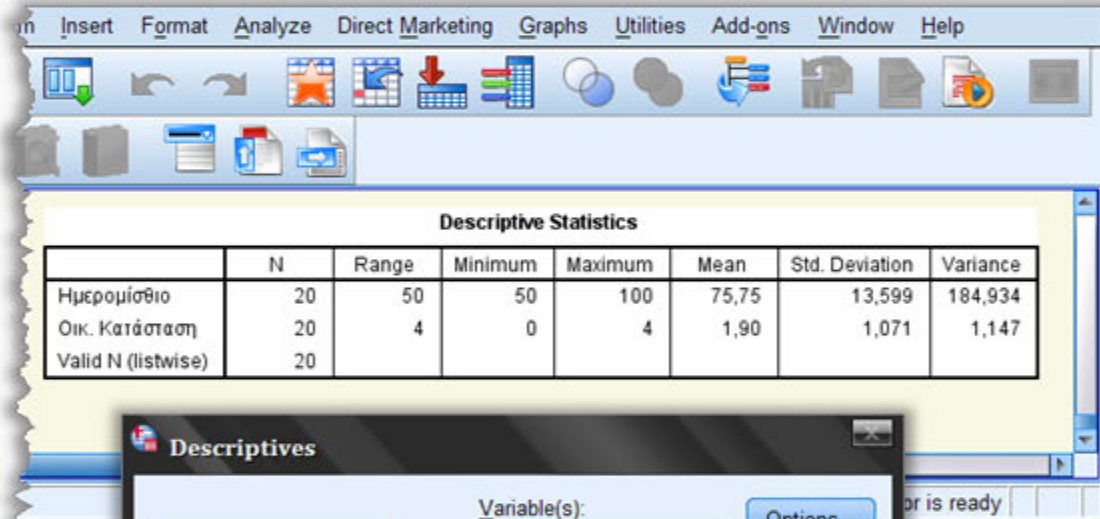
Εναλλακτικά, εάν δεν επιθυμούσαμε διαγράμματα ή εκατοστημόρια (όπως τα τεταρτημόρια ή η διάμεσος), μπορούμε, πιο άμεσα, από το menu **Analyze** να επιλέξουμε **Descriptive Statistics** > **Descriptives**^[A]. Στο εμφανιζόμενο παράθυρο διαλόγου μεταφέρουμε από τη λίστα όλων των (μη-ποιοτικών) μεταβλητών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

The image illustrates the steps to generate descriptive statistics in SPSS:

- [A]** The **Analyze** menu is open, and **Descriptive Statistics** > **Descriptives...** is selected.
- [B]** The **Descriptives** dialog box is shown. The variable **Ημερομίσθιο [Wage]** has been moved to the **Variable(s)** list.
- [Γ]** The **Descriptives: Options** dialog box is shown. The following options are checked: **Mean**, **Std. deviation**, **Variance**, **Range**, **Minimum**, and **Maximum**.

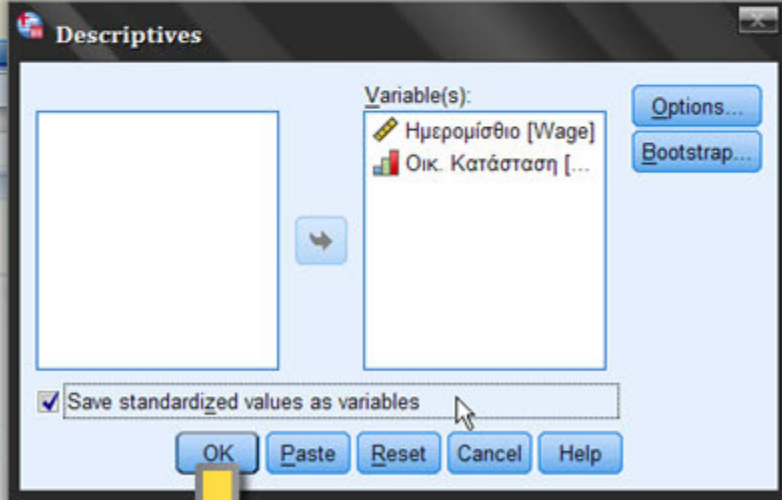
Εμφανίζεται λοιπόν στον **Viewer** ο πίνακας των (επιλεγμένων) περιγραφικών στατιστικών^[A]. Εάν στο προηγούμενο παράθυρο διαλόγου **Descriptives** είχαμε επιλέξει επιπλέον την ενέργεια **Save standardized values as variables**^[B] τότε δημιουργούνται νέες μεταβλητές (όσες και οι επιλεγμένες) οι οποίες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

[A]



	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						

[B]

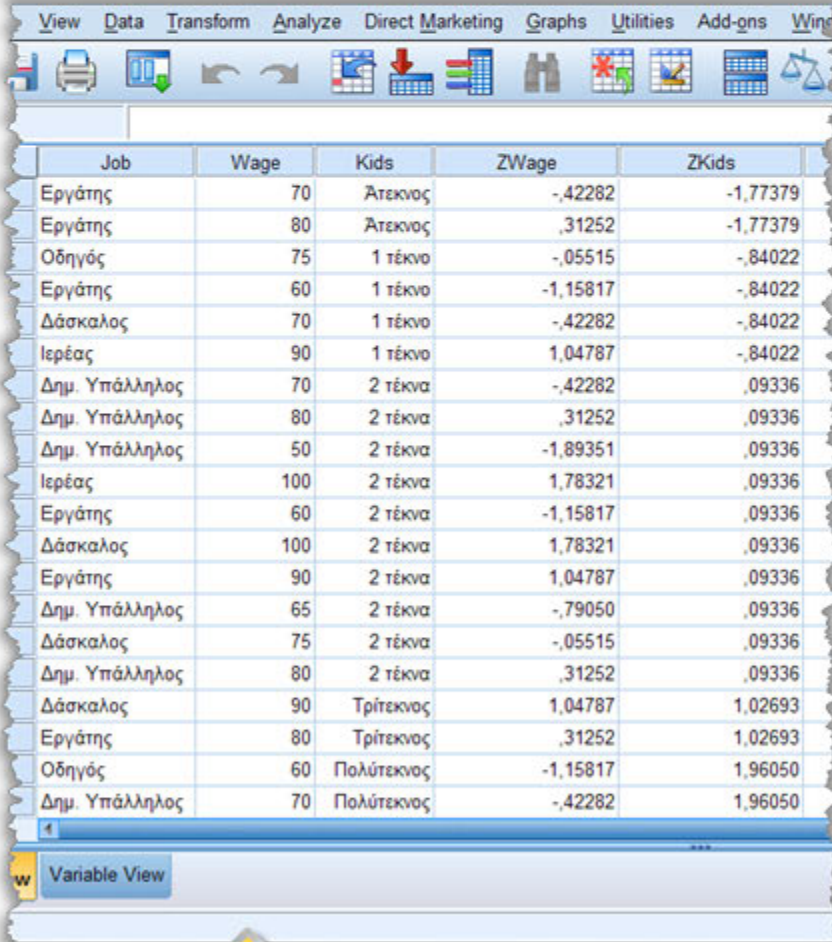


Variable(s):
 Ημερομίσθιο [Wage]
 Οικ. Κατάσταση [...]

Save standardized values as variables

OK Paste Reset Cancel Help

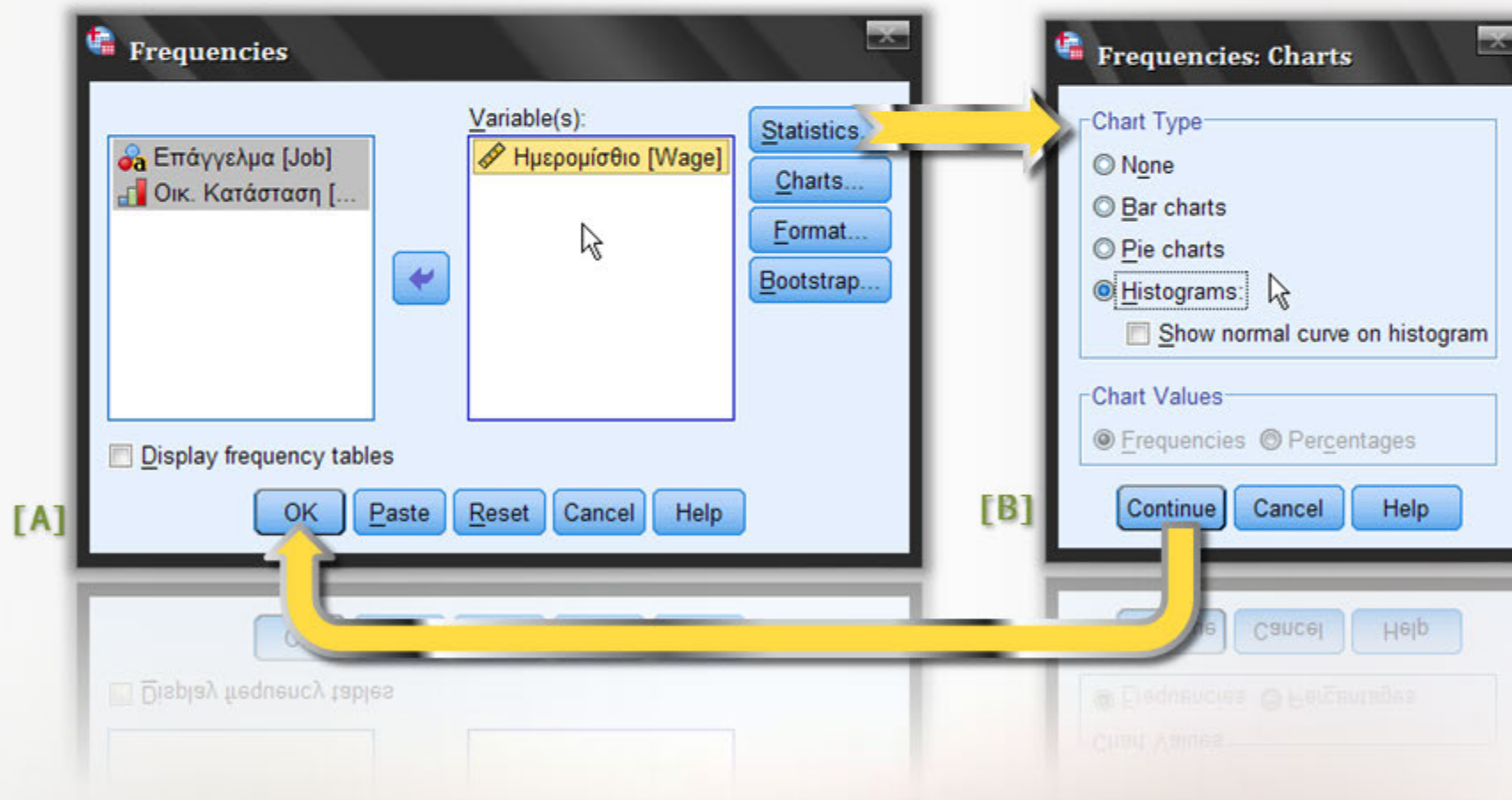
[Γ]



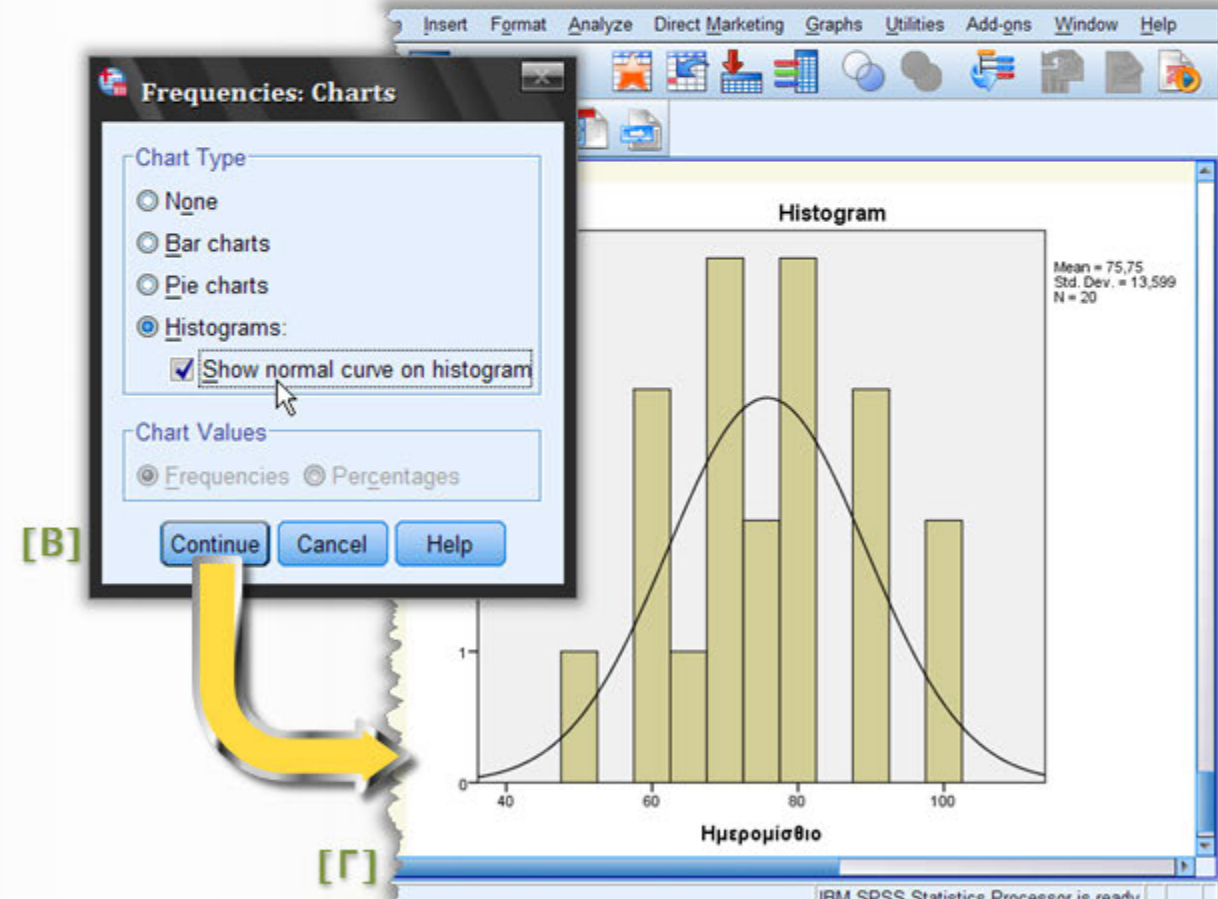
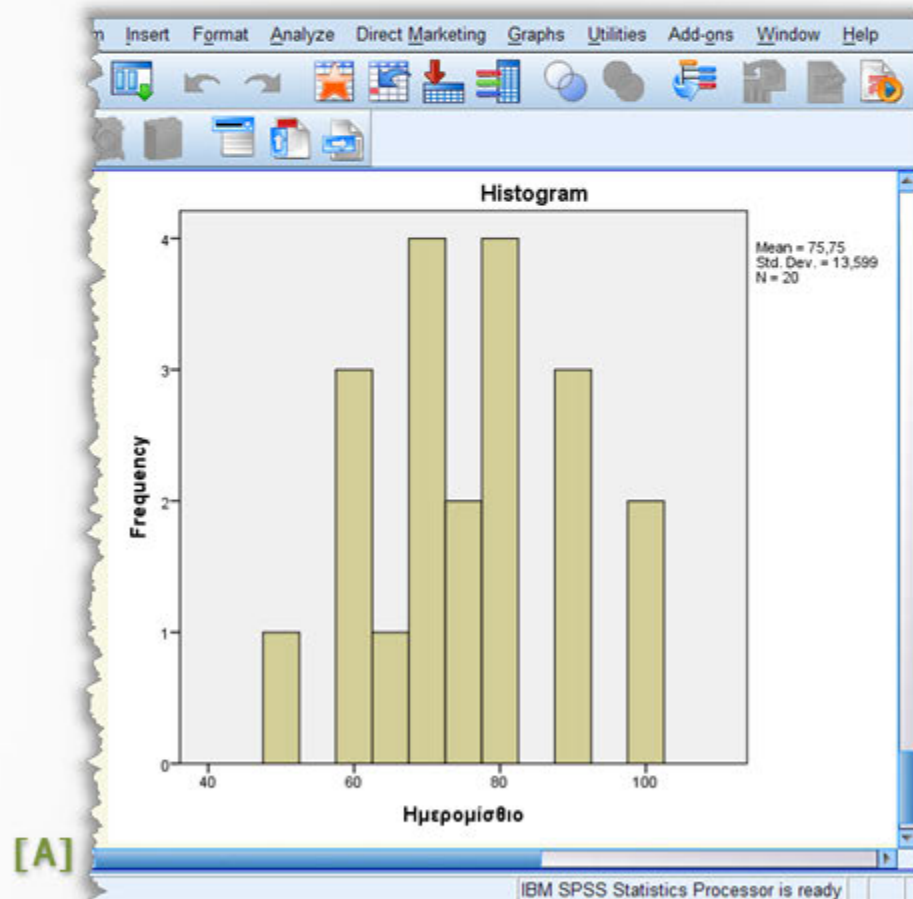
Job	Wage	Kids	ZWage	ZKids
Εργάτης	70	Άτεκνος	-.42282	-.177379
Εργάτης	80	Άτεκνος	.31252	-.177379
Οδηγός	75	1 τέκνο	-.05515	-.84022
Εργάτης	60	1 τέκνο	-1,15817	-.84022
Δάσκαλος	70	1 τέκνο	-.42282	-.84022
Ιερέας	90	1 τέκνο	1,04787	-.84022
Δημ. Υπάλληλος	70	2 τέκνα	-.42282	.09336
Δημ. Υπάλληλος	80	2 τέκνα	.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**^[B] (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^[A]. Εάν επιπλέον, στο παράθυρο διαλόγου **Frequencies: Charts**, είχαμε επιλέξει και την ενέργεια **Show normal curve on histogram**^[B], τότε μαζί με το ιστόγραμμα θα εμφανίζονταν και η συνάρτηση πυκνότητας πιθανότητας^[Γ] 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 *leaf-and-stem*.

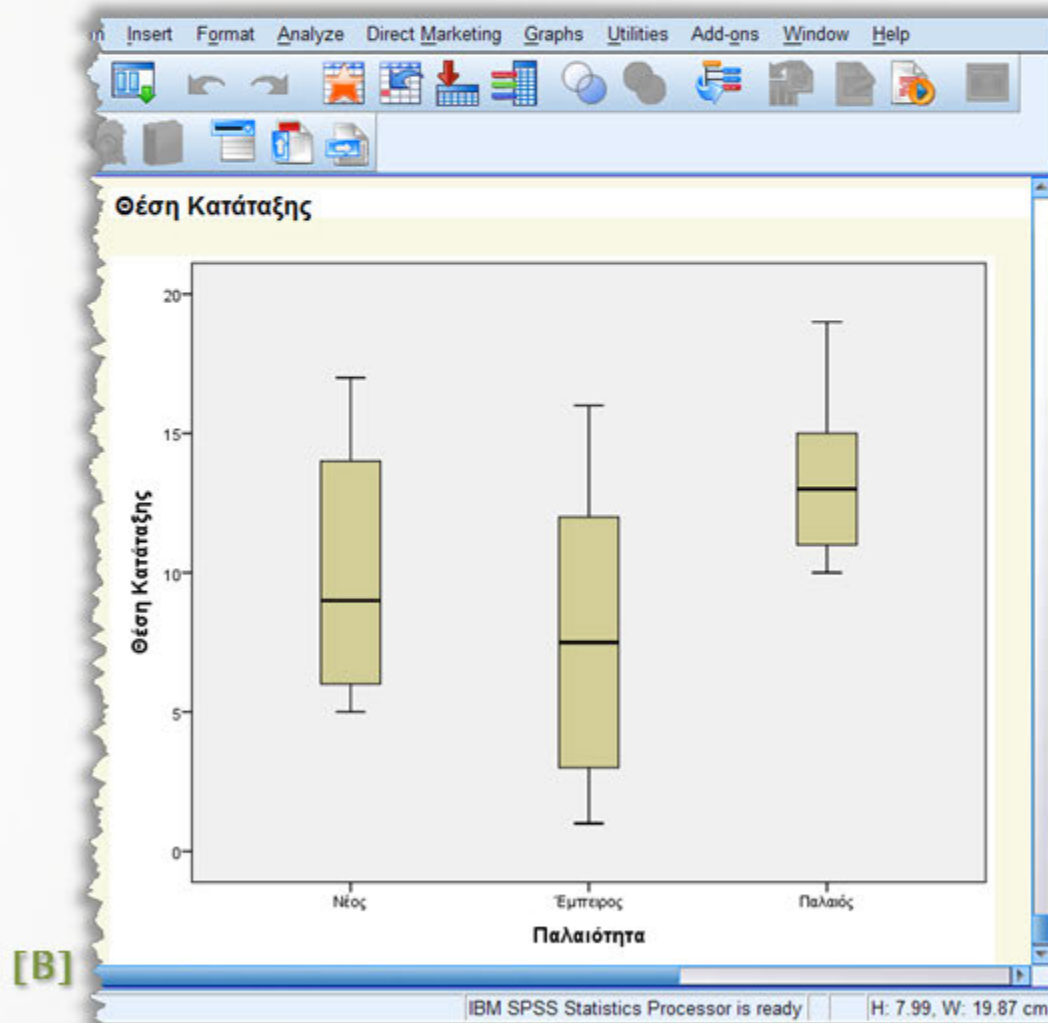
- Από το μενού **Analyze** επιλέγουμε **Descriptive Statistics** > **Explore**^[A] και στο εμφανιζόμενο παράθυρο διαλόγου τοποθετούμε (με τα αντίστοιχα βέλη απόθεσης) τις μεταβλητές **Θέση Κατάταξης [Rank]** και **Παλαιότητα [Age_Class]** στις λίστες **Dependent List** και **Factor List** αντίστοιχα^[B].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

The image illustrates the configuration of the SPSS Explore dialog box and its Plots sub-dialog box. It is divided into three main sections:

- [A] Analyze Menu:** Shows the 'Analyze' menu with 'Descriptive Statistics' > 'Explore...' selected.
- [B] Explore Dialog Box:** Shows the 'Explore' dialog box with 'Θέση Κατάταξης [Rank]' in the 'Dependent List' and 'Παλαιότητα [Age_Class]' in the 'Factor List'. The 'Display' section has 'Both' selected.
- [Γ] Explore: Plots Dialog Box:** Shows the 'Explore: Plots' sub-dialog box with 'Factor levels together' selected under 'Boxplots' and 'Stem-and-leaf' and 'Histogram' deselected under 'Descriptive'.

Yellow arrows indicate the flow of the process: from the menu to the main dialog, then to the sub-dialog, and finally to the 'OK' button.

Εμφανίζεται λοιπόν στον **Viewer** ο πίνακας των ζητούμενων περιγραφικών στατιστικών **[A]** καθώς και τα boxplots **[B]** της μεταβλητής **Rank** για κάθε κατηγορία της μεταβλητής **Age_Class**.



[B]

Descriptives

Παλαιότητα		Statistic	Std. Error	
Θέση Κατάταξης	Νέος	Mean	10,20	2,311
		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	

IBM SPSS Statistics Processor is ready | H: 7.99, W: 19.87 cm

[A]

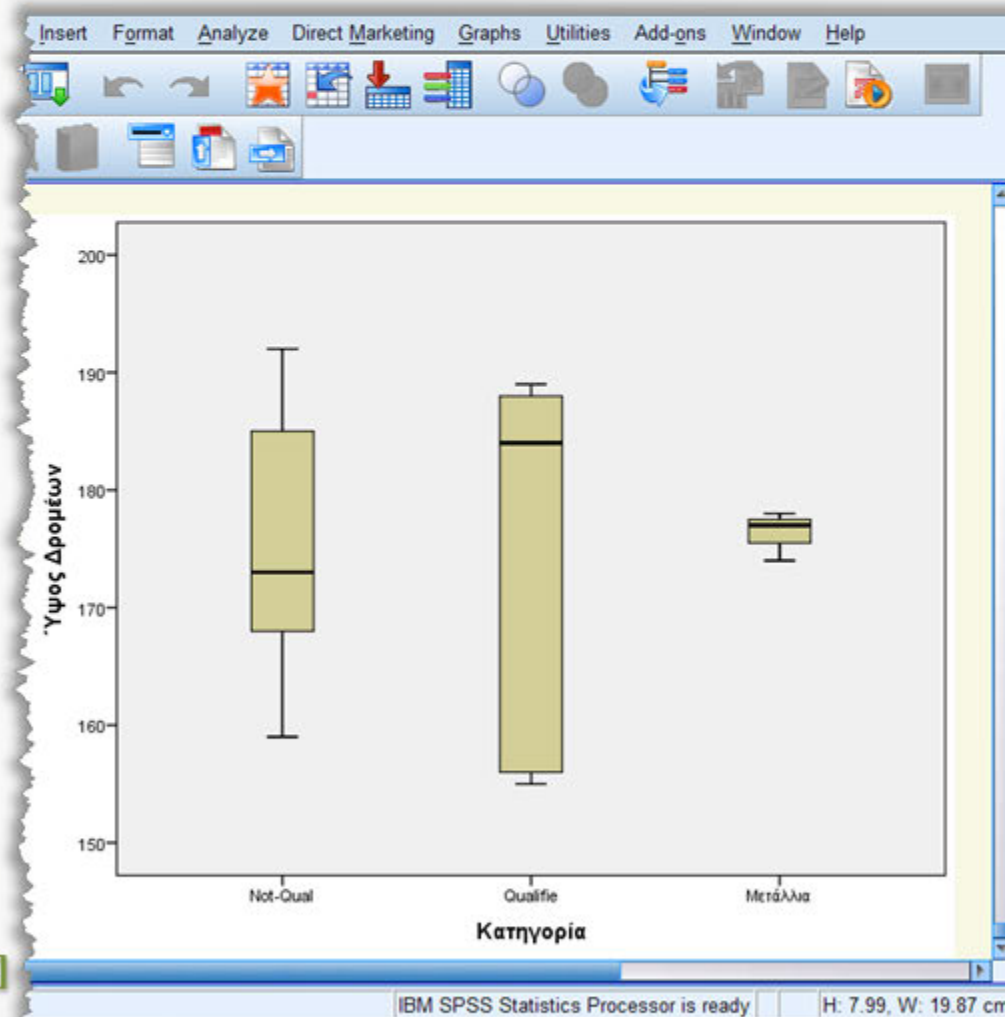
- Για το δεύτερο ζητούμενο, αλλάζουμε πρώτα τις τιμές **Χρυσό Μετάλλιο**, **Αργυρό Μετάλλιο** και **Χάλκινο Μετάλλιο** σε **Μετάλλια**.

Weight	Rank	Category	H_Class	Age_Class
174	1	Μετάλλια	170-179	Έμπειρος
177	2	Μετάλλια	170-179	Έμπειρος
178	3	Μετάλλια	170-179	Έμπειρος
156	4	Qualified	Κάτω από 1,6μ.	Έμπειρος
188	5	Qualified	180-190	Νέος
155	6	Qualified	Κάτω από 1,6μ.	Νέος
181	7	Qualified	180-190	Έμπειρος
187	8	Qualified	180-190	Έμπειρος
189	8	Qualified	180-190	Έμπειρος
171	9	Not-Qualified	170-179	Νέος
191	10	Not-Qualified	190-200	Παλαιός
179	11	Not-Qualified	170-179	Παλαιός
169	12	Not-Qualified	160-170	Έμπειρος
165	13	Not-Qualified	160-170	Παλαιός
173	14	Not-Qualified	170-179	Νέος
192	15	Not-Qualified	190-200	Παλαιός
179	16	Not-Qualified	170-179	Έμπειρος
167	16	Not-Qualified	160-170	Έμπειρος
159	17	Not-Qualified	Κάτω από 1,6μ.	Νέος
192	19	Not-Qualified	190-200	Παλαιός

- D Στη συνέχεια, από το menu **Analyze** επιλέγουμε άλλη μια φορά **Descriptive Statistics** ➤ **Explore**^[A] 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^[B]. Επιλέγοντας στη συνέχεια **Plots** στο νέο παράθυρο διαλόγου τσεκάρουμε την ενέργεια **Histogram**^[Γ]. Επικυρώνουμε με **Continue** και τέλος με **OK**.

[A]
[B]
[Γ]

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

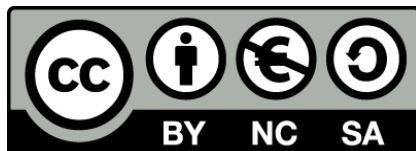


[B]

Descriptives						
Κατηγορία		Statistic	Std. Error			
Υψος Δρομέων	Not-Qual	Mean	176,09	3,476		
		95% Confidence Interval for Mean	Lower Bound	168,35		Upper Bound
	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
	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
	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]

Τέλος Ενότητας



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σημειώματα

Σημείωμα Αναφοράς

Copyright Τεχνολογικό Εκπαιδευτικό Ίδρυμα Αθήνας, Ευσταθία Παπαγεωργίου 2014. Ευσταθία Παπαγεωργίου. «Βιοστατιστική (Ε). Ενότητα 2: Περιγραφική στατιστική». Έκδοση: 1.0. Αθήνα 2014. Διαθέσιμο από τη δικτυακή διεύθυνση: ocp.teiath.gr.

Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση.

Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό. Οι όροι χρήσης των έργων τρίτων επεξηγούνται στη διαφάνεια «Επεξήγηση όρων χρήσης έργων τρίτων».

Τα έργα για τα οποία έχει ζητηθεί και δοθεί άδεια αναφέρονται στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

Επεξήγηση όρων χρήσης έργων τρίτων

© Δεν επιτρέπεται η επαναχρησιμοποίηση του έργου, παρά μόνο εάν ζητηθεί εκ νέου άδεια από το δημιουργό.

διαθέσιμο με άδεια CC-BY
Επιτρέπεται η επαναχρησιμοποίηση του έργου και η δημιουργία παραγώγων αυτού με απλή αναφορά του δημιουργού.

διαθέσιμο με άδεια CC-BY-SA
Επιτρέπεται η επαναχρησιμοποίηση του έργου με αναφορά του δημιουργού, και διάθεση του έργου ή του παράγωγου αυτού με την ίδια άδεια.

διαθέσιμο με άδεια CC-BY-ND
Επιτρέπεται η επαναχρησιμοποίηση του έργου με αναφορά του δημιουργού.
Δεν επιτρέπεται η δημιουργία παραγώγων του έργου.

διαθέσιμο με άδεια CC-BY-NC
Επιτρέπεται η επαναχρησιμοποίηση του έργου με αναφορά του δημιουργού.
Δεν επιτρέπεται η εμπορική χρήση του έργου.

διαθέσιμο με άδεια CC-BY-NC-SA
Επιτρέπεται η επαναχρησιμοποίηση του έργου με αναφορά του δημιουργού, και διάθεση του έργου ή του παράγωγου αυτού με την ίδια άδεια.
Δεν επιτρέπεται η εμπορική χρήση του έργου.

διαθέσιμο με άδεια CC-BY-NC-ND
Επιτρέπεται η επαναχρησιμοποίηση του έργου με αναφορά του δημιουργού.
Δεν επιτρέπεται η εμπορική χρήση του έργου και η δημιουργία παραγώγων του.

διαθέσιμο με άδεια CCO Public Domain
Επιτρέπεται η επαναχρησιμοποίηση του έργου, η δημιουργία παραγώγων αυτού και η εμπορική του χρήση, χωρίς αναφορά του δημιουργού.

διαθέσιμο ως κοινό κτήμα
Επιτρέπεται η επαναχρησιμοποίηση του έργου, η δημιουργία παραγώγων αυτού και η εμπορική του χρήση, χωρίς αναφορά του δημιουργού.

χωρίς σήμανση
Συνήθως δεν επιτρέπεται η επαναχρησιμοποίηση του έργου.

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο ΤΕΙ Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

