



ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Στάθης Κλωνάρης

Στην Καθημερινότητα.....

- ▶ ...πολύ συχνά συναντάμε στατιστική και στατιστική πληροφορία. Για παράδειγμα:
- ▶ Ερωτηματολόγια Πελατών - Ιατρικά Νέα - Πολιτικές Σφυγμομετρήσεις - Οικονομικές Προβλέψεις Έρευνα Αγοράς-Αθλητικά Δεδομένα
- ▶ Τι μπορούμε να συμπεράνουμε από αυτά τα δεδομένα;
- ▶ Πως διαχωρίζουμε έγκυρα από πλασματικά αποτελέσματα;
- ▶ Τι είναι Στατιστική;



Τι είναι Στατιστική;

- ▶ Στατιστική είναι ένας τρόπος με τον οποίο αντλούμε πληροφορίες από δεδομένα. (*Gerald Keller*)



- ▶ Στατιστική: «η συλλογή και η ανάλυση ποσοτικών δεδομένων για την συναγωγή συμπερασμάτων σχετικά με ένα θέμα (οικονομικό, πολιτικό, κοινωνικό κ.τ.λ.)» (Γ. Μπαμπινιώτης)



Παράδειγμα : Για το Μάθημα Στατιστικής...

Ο διευθυντής ενός γραφείου πιστεύει ότι ο χρόνος που χάνουν καθημερινά οι εργαζόμενοι διαβάζοντας και σβήνοντας ανεπιθύμητα μηνύματα ηλεκτρονικής αλληλογραφίας (spam) ξεπερνά τα 25 λεπτά. Για να ελέγξει την υπόθεση αυτή κατέγραψε τον χρόνο που έχασαν λόγω της ανεπιθύμητης αλληλογραφίας.



Εφαρμογές Στατιστικής στις Επιχειρήσεις

- ▶ Αποστολή του συγκεκριμένου μαθήματος είναι να αναδείξει το ρόλο της στατιστικής ανάλυσης (ανάλυσης δεδομένων) σε όλες τις πλευρές της λειτουργίας μίας επιχείρησης και ιδιαίτερα στον αγροδιατροφικού τομέα.
- ▶ Η στατιστική ανάλυση παίζει ένα σημαντικό ρόλο ουσιαστικά σε όλες τις πτυχές των οικονομικών και των επιχειρήσεων
- ▶ Σημαντικές εφαρμογές σε επίπεδο της επιχείρησης έχει στους τομείς της οικονομίας, μάρκετινγκ, ανθρωπίνων πόρων, παραγωγής και λογιστικής.



Βασικές έννοιες της Στατιστικής

Πληθυσμός

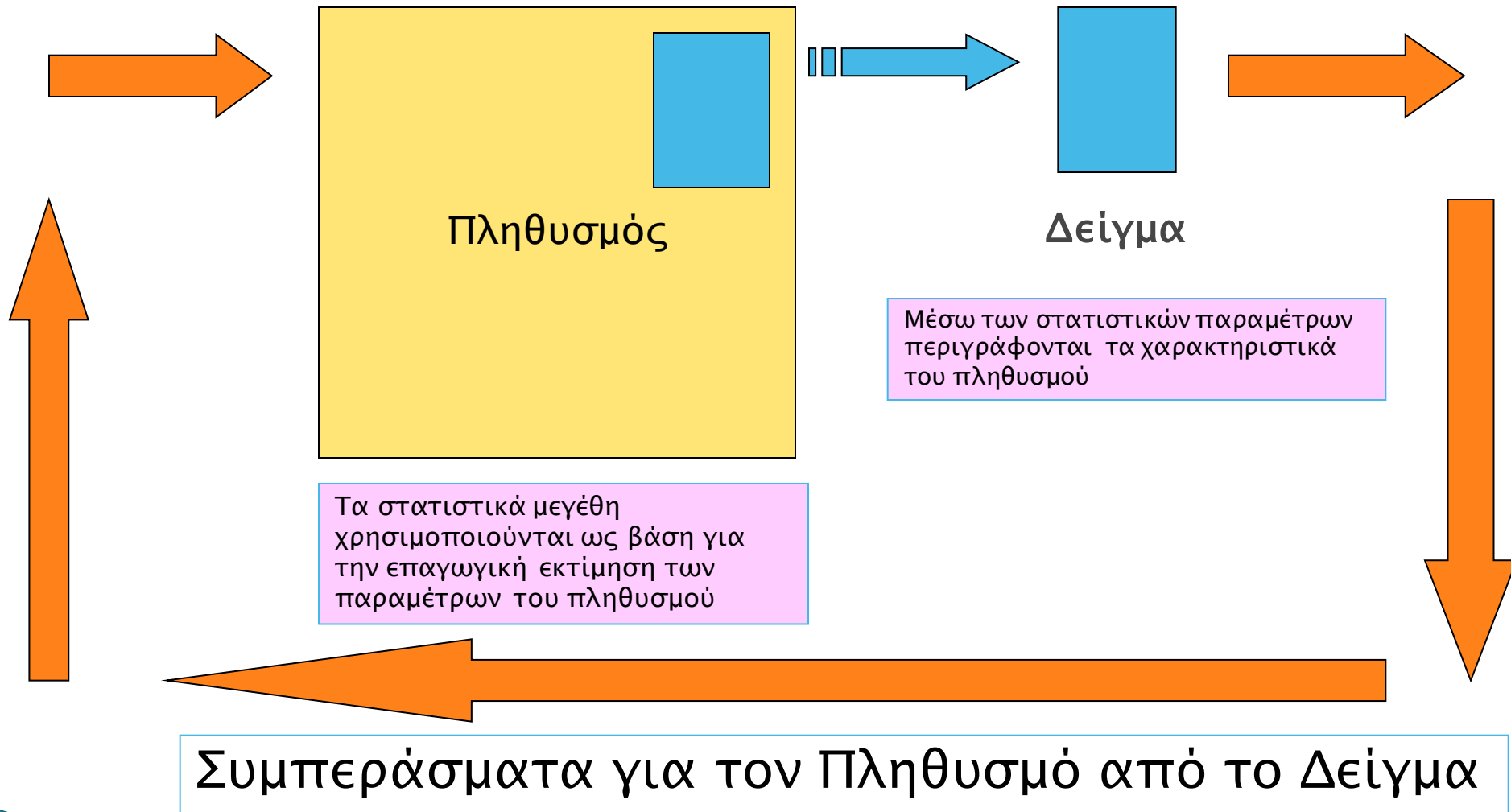
- ❖ ένας *πληθυσμός* (population) είναι το σύνολο όλων των στοιχείων που μας ενδιαφέρουν.
- ❖ Συχνά είναι πολύ μεγάλος; Ορισμένες φορές είναι μη πεπερασμένος. π.χ. 30.000–40.000 γιαούρτια ημερησίως
- ❖ Κάθε περιγραφικό μέγεθος ενός πληθυσμού ονομάζεται **παράμετρος** (*parameter*)

Δείγμα

- ❖ Ένα *δείγμα* είναι ένα υποσύνολο δεδομένων συλλεγόμενο από τον πληθυσμό.
- ❖ Ενδεχομένως πολύ μεγάλο, αλλά αρκετά μικρότερο από τον πληθυσμό. π.χ. ένα δείγμα 850 γιαούρτια.
- ❖ Κάθε περιγραφικό μέγεθος του δείγματος ονομάζεται **στατιστικό μέγεθος** (*statistic*)



Πληθυσμός και Δείγμα



Περιγραφική Στατιστική

- ▶ ...είναι μέθοδοι που οργανώνουν, συνοψίζουν, και παρουσιάζουν δεδομένα με εύκολο και πληροφοριακό τρόπο. Αυτοί οι μέθοδοι περιλαμβάνουν:
 - Γραφικές και
 - Αριθμητικές Τεχνικές .
- ▶ Η μέθοδος η οποία εφαρμόζεται εξαρτάται από το τι πληροφορία θέλουμε να εξάγουμε. Ενδιαφέρεστε για ...
 - μέτρα κεντρικής τάσης; και/ή
 - μέτρα μεταβλητότητας (διασποράς);
- ▶ Η περιγραφική στατιστική δίνει λύση στα παραπάνω θέματα.



Επαγωγική Στατιστική ή Στατιστική Συμπερασματολογία

- ▶ Η περιγραφική στατιστική περιγράφει το σύνολο των δεδομένων το οποίο αναλύεται, αλλά δεν μας επιτρέπει να βγάλουμε συμπεράσματα ή να κάνουμε παρεμβολές σχετικές με τα δεδομένα. Για αυτό τον λόγο χρειαζόμαστε έναν άλλον κλάδο της στατιστικής: **επαγωγική στατιστική** (*inferential statistics*),
- ▶ Η επαγωγική στατιστική είναι επίσης μία σειρά μεθόδων, αλλά χρησιμοποιείται για την εξαγωγή συμπερασμάτων σχετικά με τα χαρακτηριστικά του **πληθυσμού** βασισμένα σε δεδομένα από το **δείγμα**.



Επαγωγική Στατιστική

- ▶ Επαγωγική Στατιστική είναι η **διαδικασία** η οποία εκτιμά, προβλέπει, ή αποφασίζει σχετικά με τον πληθυσμό βασισμένη στο δείγμα.



Τι μπορούμε να *συμπεράνουμε* σχετικά με τις παραμέτρους του πληθυσμού βασισμένοι σε στατιστικά στοιχεία του δείγματος;



Επαγωγική Στατιστική

Χρησιμοποιούμε στατιστικά στοιχεία για να εξάγουμε συμπεράσματα σχετικά με τους παραμέτρους.

Συνεπώς, μπορούμε να εκτιμήσουμε, να προβλέψουμε, ή να αποφασίσουμε σχετικά με τον πληθυσμό βασισμένοι σε δεδομένα του δείγματος.

Επομένως, ότι γνωρίζουμε σχετικά με το δείγμα μπορούμε να το εφαρμόσουμε στον πληθυσμό από τον οποίο επιλέχθηκε.



Επαγωγική Στατιστική

▶ Αιτιολογία:

- • Είναι ανέφικτο και δαπανηρό να ερευνήσουμε όλα τα άτομα από έναν μεγάλο πληθυσμό.
- • Είναι απλό και οικονομικό να πάρουμε ένα δείγμα και να κάνουμε εκτιμήσεις σχετικά με τον πληθυσμό από το δείγμα.

▶ Όμως:

- Τέτοια συμπεράσματα και εκτιμητές δεν είναι πάντα αληθή.
- Για αυτό τον λόγο, κατασκευάζουμε στην επαγωγική στατιστική «μέτρα αξιοπιστίας», ονομαστικά **επίπεδο εμπιστοσύνης** και **επίπεδο σημαντικότητας**.



Περιγραφικές Τεχνικές

- ▶ Η περιγραφική στατιστική ασχολείται με την αναδιάταξη, την σύνοψη και την παρουσίαση των δεδομένων με τρόπο που διευκολύνει την αναγνώριση χρήσιμων πληροφοριών.
- ▶ Οι μέθοδοι της κάνουν χρήση γραφικών τεχνικών και αριθμητικών μέτρων περιγραφής (όπως μέσες τιμές) για να συνοψίσουμε και να παρουσιάσουμε τα δεδομένα.
- ▶ Οι μέθοδοι της περιγραφικής στατιστικής εφαρμόζονται σε κάθε σύνολο δεδομένων είτε προκύπτει για το σύνολο του πληθυσμού είτε για ένα δείγμα.

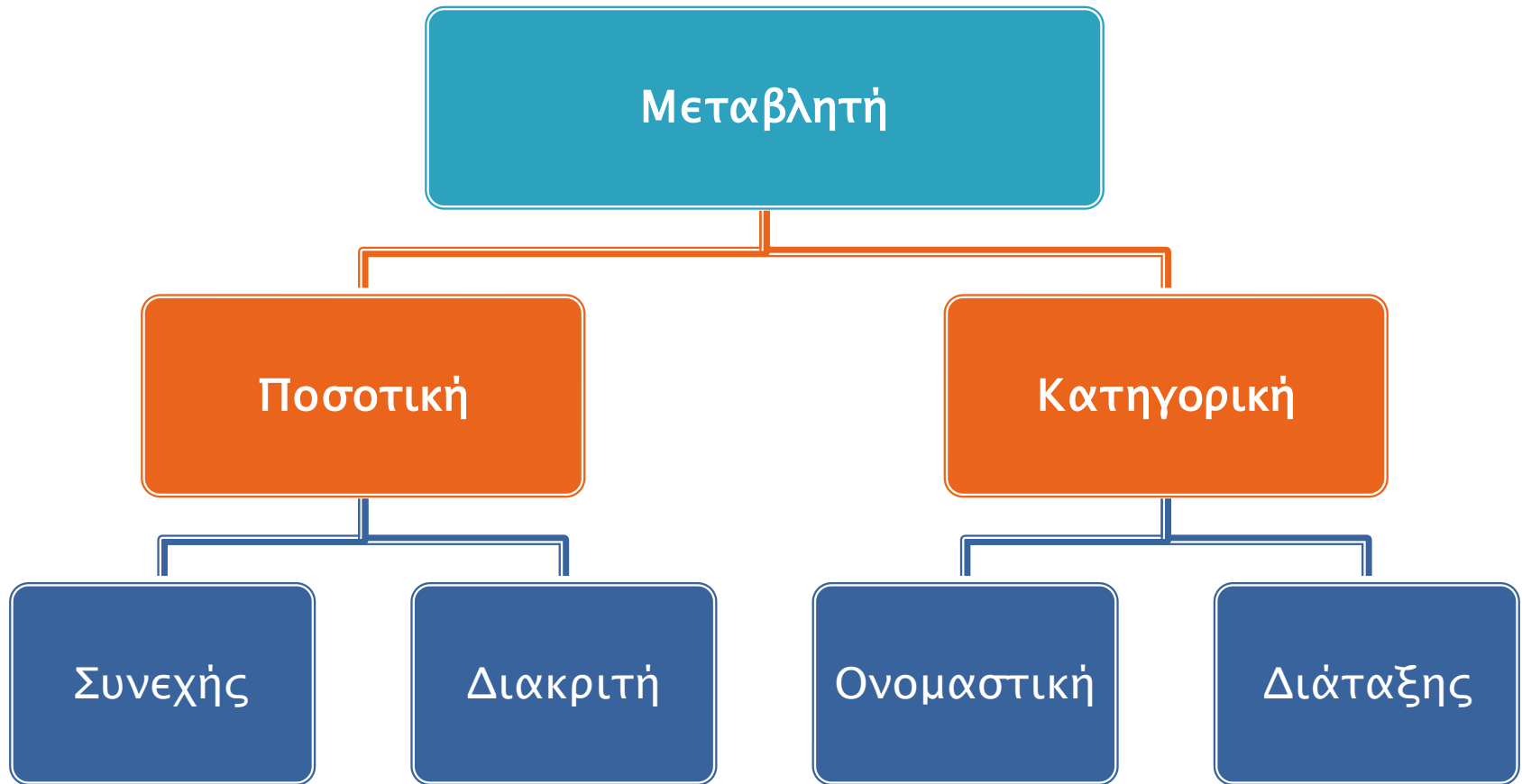


Ορισμοί

- ▶ Η **μεταβλητή** είναι κάποιο χαρακτηριστικό του πληθυσμού ή του δείγματος.
Π.χ Χαμένος χρόνος
Τυπικά σημαίνονται με κεφαλαία γράμματα: X, Y, Z...
- ▶ Οι **τιμές** της μεταβλητής είναι το εύρος των πιθανών τιμών για μία μεταβλητή.
Π.χ. Τα λεπτά για σβήσουν τα spam μηνύματα (0..90)
- ▶ **Δεδομένα** είναι οι τιμές που παρατηρούνται από μία μεταβλητή.
Π.χ. Λεπτά της ώρας: {67, 74, 71, 83, 93, 55, 48}



Είδη Μεταβλητών



Τύποι Δεδομένων και πληροφοριών¹⁶

Ποσοτικά ή Συνεχή

- ✓Οι τιμές είναι πραγματικοί αριθμοί
- ✓Όλοι οι αριθμητικοί υπολογισμοί επιτρέπονται
- ✓Μπορούν να χρησιμοποιηθούν και ως διατακτικά ή ονομαστικά (π.χ. ηλικία, βάρος, τιμές, ποσότητες κτλ)

Διακριτά

- Οι Τιμές είναι πραγματικοί αριθμού χωρίς να περιλαμβάνουν την τιμή μηδέν
- Όλοι οι αριθμητικοί υπολογισμοί επιτρέπονται εκτός της διαίρεση
- Μπορούν να χρησιμοποιηθούν και ως διατακτικά ή ονομαστικά (π.χ. μεταβολές θερμοκρασίας, σκορ GMAT, GRE κτλ)

Διατακτικά

- ✓Οι τιμές αντιπροσωπεύουν μόνο διάταξη και όχι μέγεθος
- ✓Επιτρέπονται μόνο υπολογισμοί σε σχέση με τη διάταξη
- ✓Μπορούν να χρησιμοποιηθούν και ως ονομαστικά αλλά όχι ποσοτικά (π.χ. ένταση γνώμης, ένταση γνώσης, ηλικιακές κατηγορίες κτλ.)

Ονομαστικά

- ✓Οι τιμές αντιπροσωπεύουν μη συγκρίσιμες κατηγορίες
- ✓Επιτρέπεται μόνο η μέτρηση των συχνοτήτων
- ✓Δεν μπορούν να χρησιμοποιηθούν ως διατακτικά ή ως ποσοτικά δεδομένα (π.χ. φύλο, επάγγελμα, θρήσκευμα κτλ)



Γραφικά & Πίνακες για Ονομαστικά Δεδομένα ...

- ▶ Ο μόνος επιτρεπόμενος υπολογισμός για ονομαστικά δεδομένα είναι να καταμετρήσουμε τη συχνότητα για κάθε τιμή της μεταβλητής.
- ▶ Μπορούμε να συνοψίσουμε τα δεδομένα σε έναν πίνακα ο οποίος παρουσιάζει τις κατηγορίες με τις συχνότητές τους, καλώντας τον *πίνακας ή κατανομή συχνοτήτων*.
- ▶ Ο *πίνακας ή η κατανομή σχετικών συχνοτήτων* καταγράφει τις κατηγορίες και την αναλογία με την οποία η καθεμία εμφανίζεται.



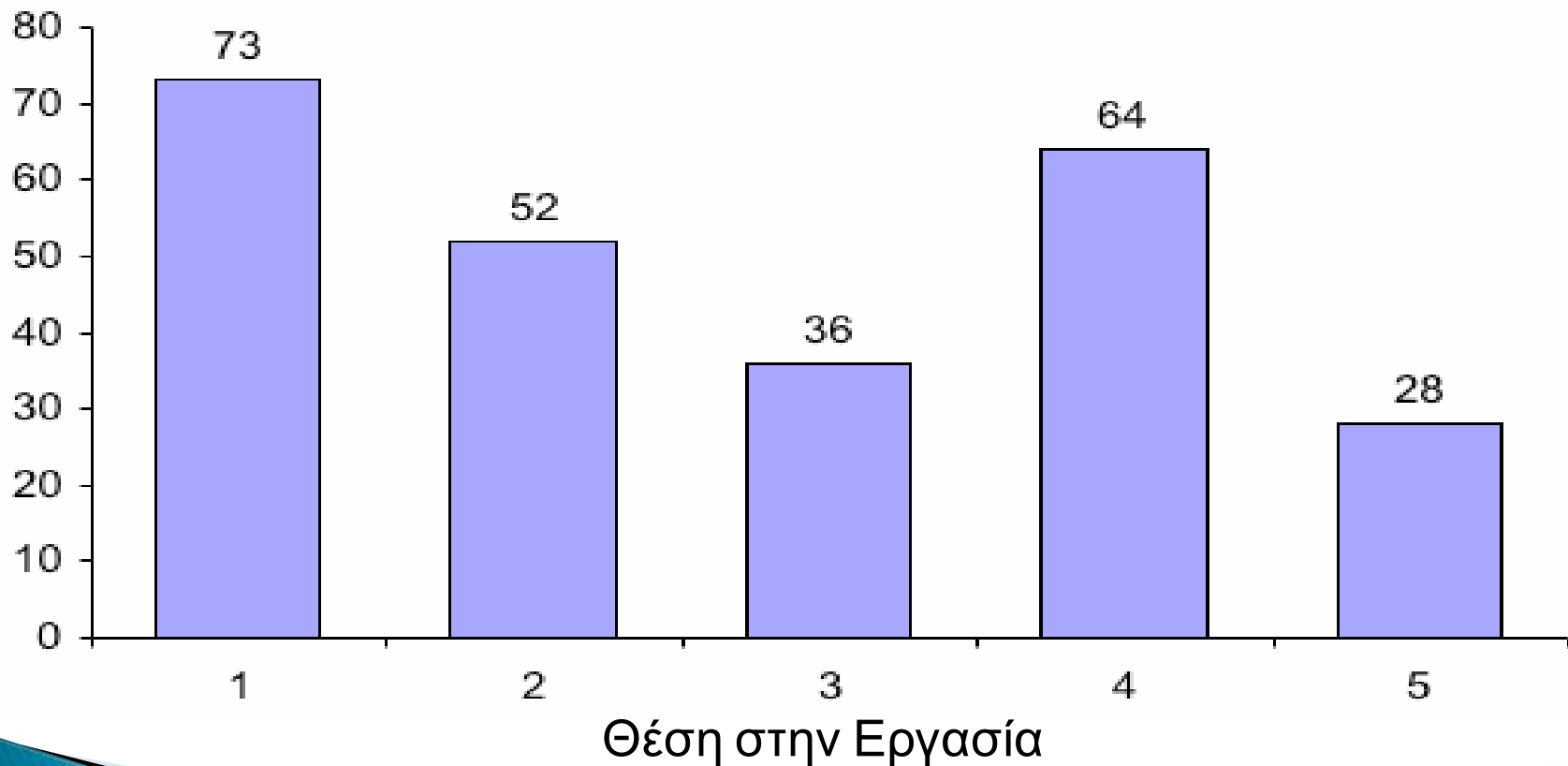
Πίνακας Συχνοτήτων ή Σχετικών Συχνοτήτων ¹⁸

Θέση στην Εργασία	Συχνότητα	Σχετική Συχνότητα
Accounting	73	28.9%
Finance	52	20.6
General management	36	14.2
Marketing/Sales	64	25.3
Other	28	11.1
Total	253	100



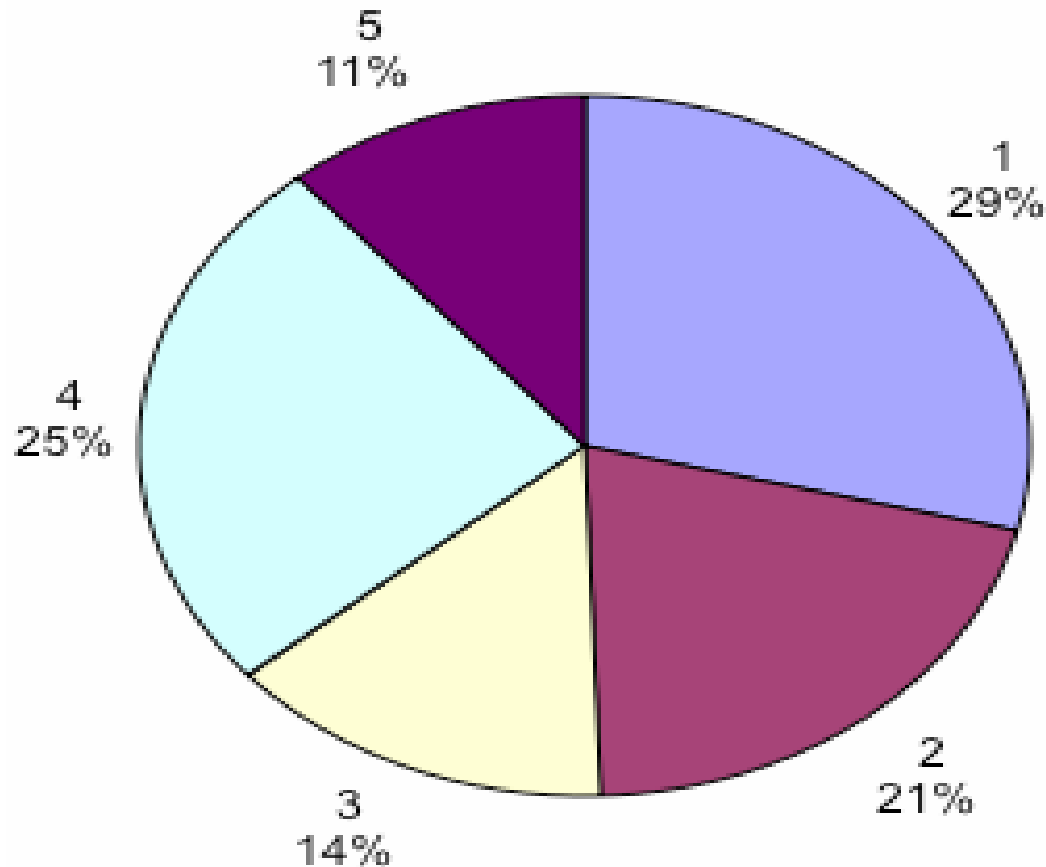
Ραβδογράμματα (Bar charts)

Τα ραβδογράμματα συχνά χρησιμοποιούνται για να παρουσιάσουν συχνότητες



Κυκλικό Διάγραμμα (Pie Chart)

Τα Κυκλικά Διαγράμματα δείχνουν
σχετικές συχνότητες ...

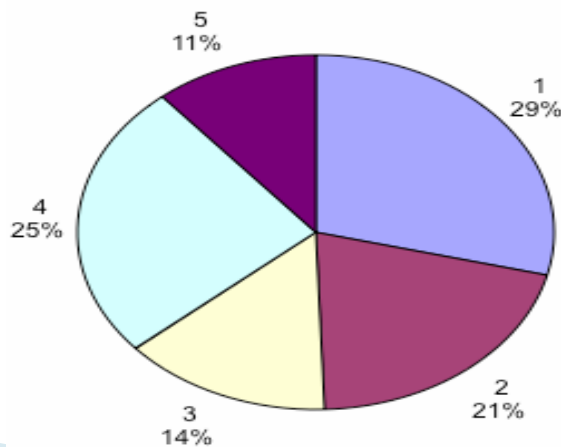


Παρουσίαση Ονομαστικών δεδομένων

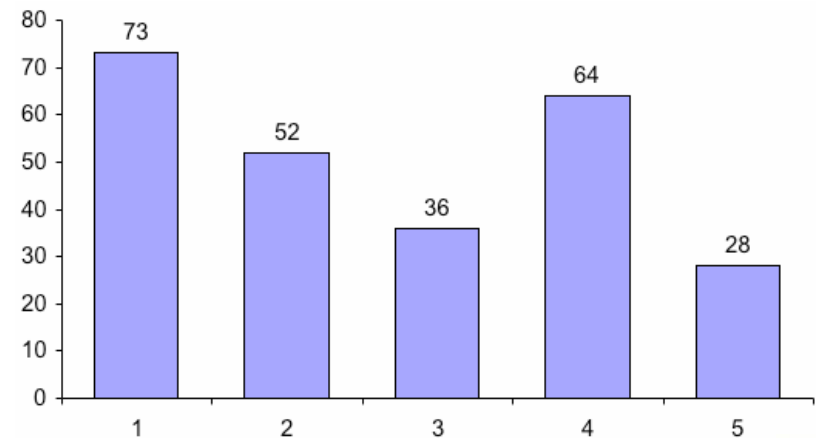
Περιοχή	Συχνότητα	Σχετική Συχνότητα
Accounting	73	28.9%
Finance	52	20.6
General management	36	14.2
Marketing/Sales	64	25.3
Other	28	11.1
Total	253	100

Και τα τρία δείχνουν την ίδια **πληροφόρηση**, (βασισμένα στα ίδια **δεδομένα**).
Απλά διαφορετική **παρουσίαση**.

Κυκλικό Διάγραμμα



Ραβδόγραμμα

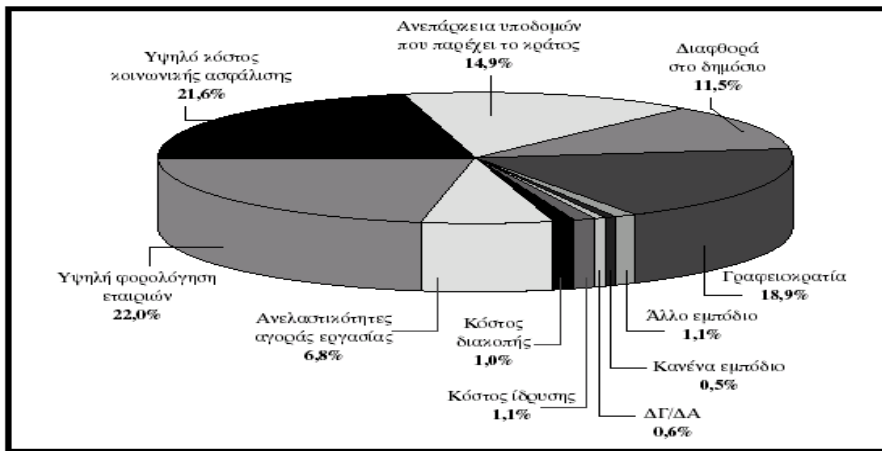


.....παράδειγμα Κυκλικών διαγραμμάτων

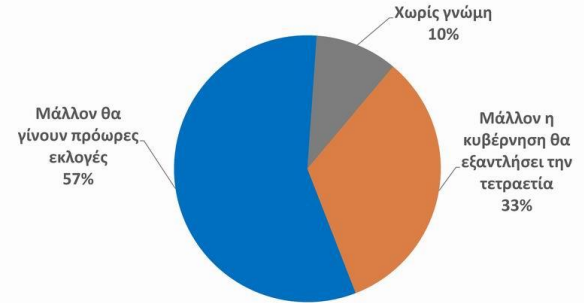
Δ.23 ΓΝΩΜΗ ΓΙΑ ΤΗ ΔΙΕΞΑΓΩΓΗ ΠΡΩΩΡΩΝ ΕΚΛΟΓΩΝ

Εσείς προσωπικά τι πιστεύετε; Η σημερινή Κυβέρνηση μάλλον θα εξαντλήσει την τετραετία, ή μάλλον θα γίνουν πρόωρες εκλογές;

Τα Εμπόδια στις Σχέσεις των Επιχειρήσεων με το Δημόσιο



Πηγή: ΣΕΒ-ICAP

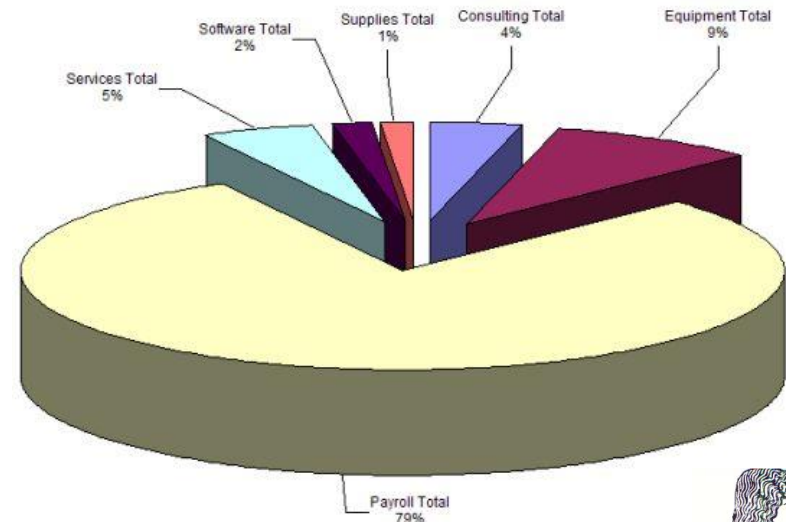


public issue www.publicissue.gr

2011065

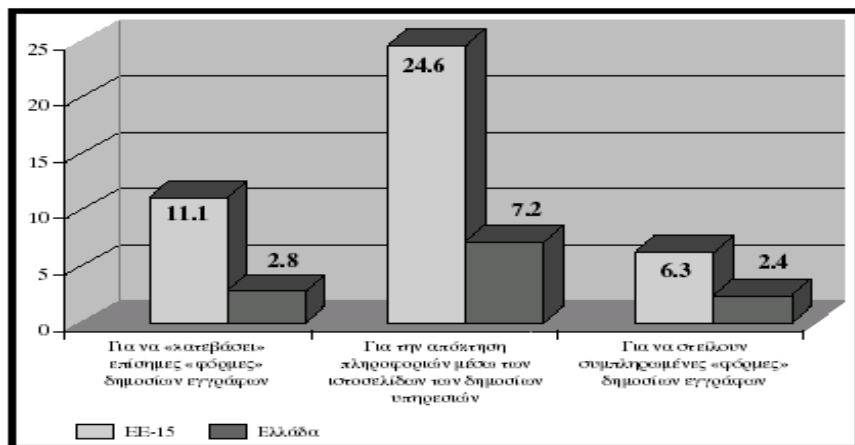
ΣΚΑΪ Η ΚΑΘΗΜΕΡΙΝΗ

3-month Expenditure Breakdown

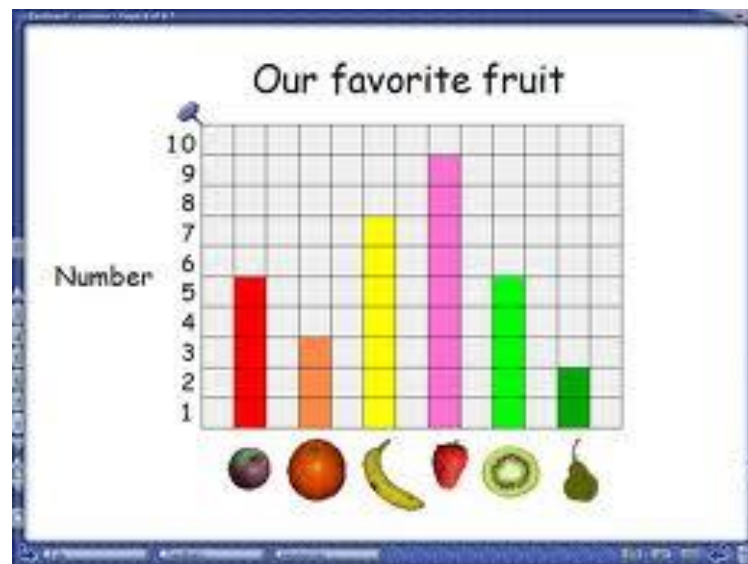


...παράδειγμα Ραβδογραμμάτων

Ποσοστό των πολιτών που χρησιμοποιεί το διαδίκτυο διαδραστικά με τις δημόσιες υπηρεσίες, ανά τύπο δραστηριότητα, 2004



Πηγή: Eurostat



Κατασκευή Πίνακα συχνοτήτων στο SPSS

Παράδειγμα:

Τα παρακάτω δεδομένα αφορούν την οικογενειακή κατάσταση 20 ανδρών (1: Έγγαμοι, 2: Άγαμοι 3: Διαζευγμένος ή χήρος)

1	1	2	1	3
1	2	2	2	1
3	3	1	2	3
1	1	2	2	1

Πίνακες Συχνοτήτων

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

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

Save this document

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

Data View Variable View

*Untitled1 [DataSet0] - 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	Columns	Align	Measure	Role
1	MS	Numeric	8	0	Marital Status	{1, Έγγαμος... None		8	Right	Unknown	Input
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											

Data View Variable View

Πίνακας Συχνοτήτων

Analyze > Descriptive Statistics > Frequencies

The screenshot shows the IBM SPSS Statistics Data Editor interface. The 'Analyze' menu is open, and the 'Frequencies...' option is selected. The 'Frequencies' dialog box is displayed, showing the variable 'Marital Status [MS]' selected in the 'Variable(s):' list. The 'Display frequency tables' checkbox is checked and circled in red. The 'OK' button is highlighted.

	MS	var
1	1	
2	1	
3	2	
4	1	
5	3	
6	1	
7	2	
8	2	
9	2	
10	1	
11	3	
12	3	
13	1	
14	2	
15	3	
16	1	
17	1	
18	2	
19	2	
20	1	

Πίνακας Συχνοτήτων

Marital Status

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Έγγαμος	9	45,0	45,0	45,0
Άγαμος	7	35,0	35,0	80,0
Διαζευγμένος ή Χήρος	4	20,0	20,0	100,0
Total	20	100,0	100,0	

Συχνότητες

Σχετικές
Συχνότητες

Αθροιστικές
Συχνότητες

Σχέση μεταξύ Δύο Ονομαστικών Μεταβλητών ...

- ▶ Μέχρι τώρα έχουμε κοιτάξει σε τεχνικές πινάκων και γραφικές για μία μεταβλητή (για ονομαστικά ή για διαστημικά δεδομένα).
- ▶ Ένας *πίνακας διπλής εισόδου* (*cross tabs*) χρησιμοποιείται για να περιγράψει την σχέση μεταξύ **δύο** ονομαστικών μεταβλητών.
- ▶ Ένας πίνακας συνάφειας καταγράφει την *συχνότητα* για **κάθε παρατήρηση** των τιμών των δύο μεταβλητών ...



Πίνακας Διπλής εισόδου...

- ▶ Παράδειγμα:, ένα δείγμα αναγνωστών εφημερίδων ρωτήθηκε να αναφέρει ποια εφημερίδα διαβάζουν: Globe και Mail (1), Post (2), Star (3), ή Sun (4), και να υποδείξουν αν ήταν εργάτες (blue-collar worker) (1), υπάλληλοι (white-collar worker) (2), ή επαγγελματίες (professional) (3).

Reader	Newspaper	Occupation
1	2	2
2	4	1
3	1	2
.	.	.
.	.	.
352	2	3
353	3	1
354	3	2

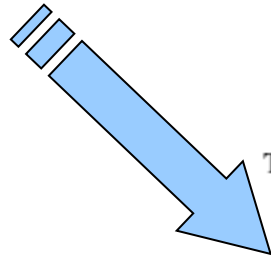


Table 2.9 Contingency Table of Frequencies for Example 2.8

	Occupation			
Newspaper	Blue Collar	White Collar	Professional	Total
G&M	27	29	33	89
Post	18	43	51	112
Star	38	21	22	81
Sun	37	15	20	72
Total	120	108	126	

Αυτή η απάντηση των αναγνωστών βρίσκεται ως μέρος του συνολικού αριθμού στον πίνακα συνάφειας ...



Πίνακας Διπλής Εισόδου...

- ▶ Ερμηνεία: Η σχετική συχνότητα στις στήλες 2 & 3 είναι παραπλήσιες, αλλά υπάρχουν μεγάλες διαφορές ανάμεσα στις στήλες 1 και 2 και ανάμεσα στις στήλες 1 και 3.

Table 2.10 Column Relative Frequencies for Example 2.8

Newspaper	Occupation		
	Blue Collar	White Collar	Professional
G&M	$27/120 = .23$	$29/108 = .27$	$33/126 = .26$
Post	$18/120 = .15$	$43/108 = .40$	$51/126 = .40$
Star	$38/120 = .32$	$21/108 = .19$	$22/126 = .17$
Sun	$37/120 = .31$	$15/108 = .14$	$20/126 = .16$

παραπλήσια

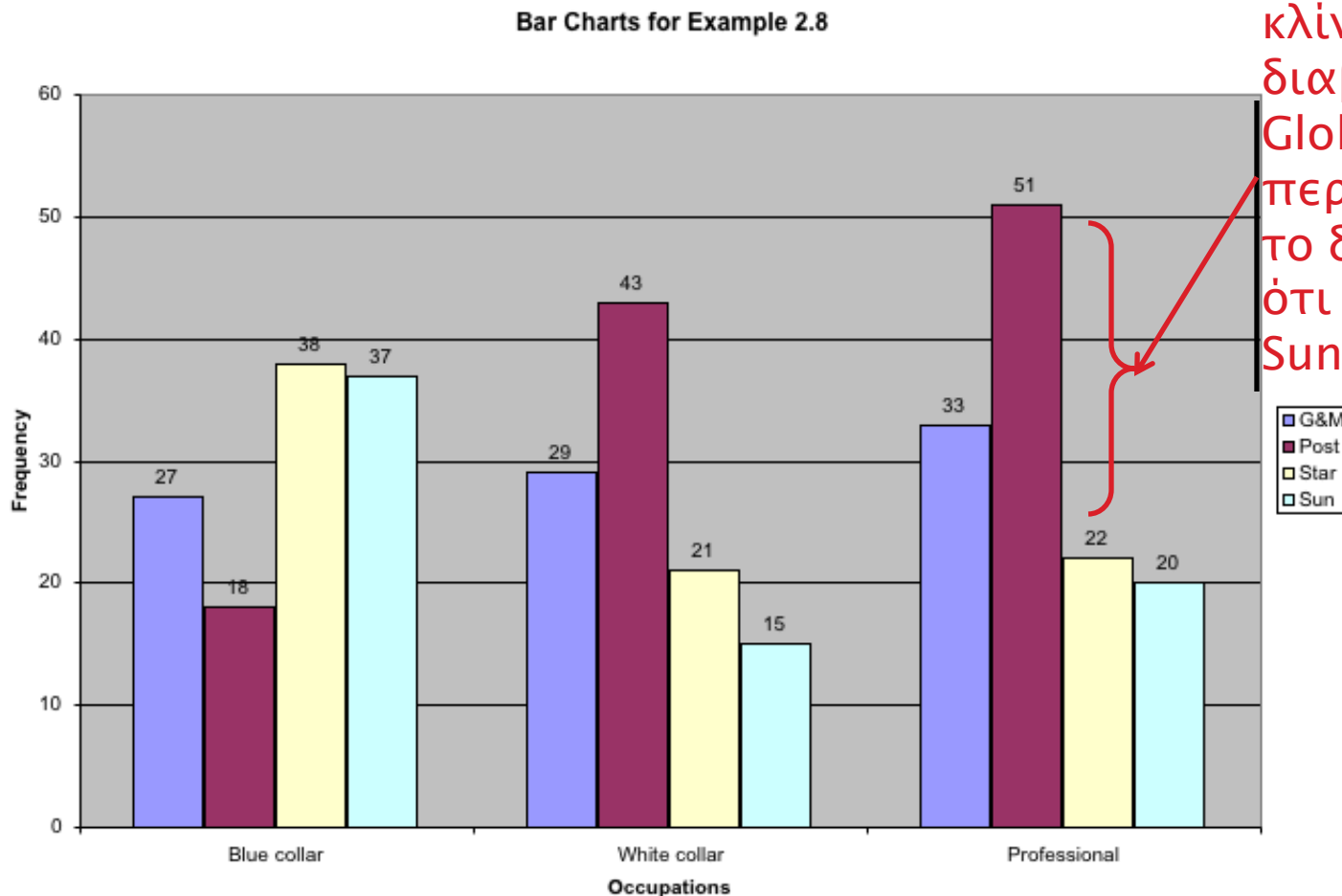
ανόμοια

- ▶ Αυτό μας λέει ότι οι εργάτες έχουν την τάση να διαβάζουν διαφορετικές εφημερίδες από τους υπαλλήλους και τους επαγγελματίες, και οι υπάλληλοι και οι επαγγελματίες έχουν κοινές προτιμήσεις για εφημερίδες.



Γράφημα για τη Σχέση Μεταξύ Δύο Ονομαστικών μεταβλητών ...

- ▶ Χρησιμοποιούμε τα δεδομένα από τον πίνακα συνάφειας για να δημιουργήσουμε τις ράβδους (μπάρες) ...



Οι επαγγελματίες κλίνουν να διαβάσουν την Globe & Mail περισσότερο από το διπλάσιο από ότι την Star ή την Sun...

Πίνακες Συνάφειας

- ◆ Έστω ότι στο προηγούμενο παράδειγμα οι 10 πρώτοι ανήκουν σε μια εισοδηματική τάξη ($1: < 15.000\text{€}/\text{ετησίως}$) ενώ οι υπόλοιποι 10 σε μεγαλύτερη εισοδηματική τάξη ($2: > 15.000\text{€}/\text{ετησίως}$)
- ◆ Μπορούμε να κατασκευάσουμε το **πίνακα συνάφειας** (contingency table), όπου απεικονίζει τη διμεταλητή κατανομή συχνοτήτων για δύο κατηγορικές μεταβλητές οικογενειακή κατάσταση και εισοδηματική κλάση

Πίνακες Συνάφειας

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

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

1 : IN 1

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

Data View Variable View

*Untitled1 [DataSet0] - 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	Columns	Align	Measure	Role
1	MS	Numeric	8	0	Marital Status	{1, Έγγαμος...	None	8	Right	Nominal	Input
2	IN	Numeric	8	0	Income	None	None	8	Right	Ordinal	Input
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											

Data View Variable View

Πίνακες Συνάφειας

The screenshot shows the IBM SPSS Statistics Data Editor interface. The 'Analyze' menu is open, and 'Crosstabs...' is selected. The data editor shows a table with two columns: 'MS' and 'IN'. The data is as follows:

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

The 'Crosstabs' dialog box is shown. The 'Row(s):' field contains 'Marital Status [MS]' and the 'Column(s):' field contains 'Income [IN]'. The 'Cells...' button is circled in red. Other buttons include 'Exact...', 'Statistics...', 'Format...', and 'Bootstrap...'. The 'Display clustered bar charts' and 'Suppress tables' options are unchecked. The 'OK', 'Paste', 'Reset', 'Cancel', and 'Help' buttons are at the bottom.

The 'Crosstabs: Cell Display' dialog box is shown. It has several sections:

- Counts:** Observed, Expected, Hide small counts (Less than 5).
- z-test:** Compare column proportions, Adjust p-values (Bonferroni method).
- Percentages:** Row, Column, Total.
- Residuals:** Unstandardized, Standardized, Adjusted standardized.
- Noninteger Weights:** Round cell counts, Round case weights, Truncate cell counts, Truncate case weights, No adjustments.

The 'Continue', 'Cancel', and 'Help' buttons are at the bottom.

Πίνακες Συνάφειας

Martial Status * Income Crosstabulation

			Income		Total
			1	2	
Martial Status	Έγγαμος	Count	5	4	9
		% within Martial Status	55,6%	44,4%	100,0%
		% within Income	50,0%	40,0%	45,0%
		% of Total	25,0%	20,0%	45,0%
	Άγαμος	Count	4	3	7
		% within Martial Status	57,1%	42,9%	100,0%
		% within Income	40,0%	30,0%	35,0%
		% of Total	20,0%	15,0%	35,0%
	Διαζευγμένος ή Χήρος	Count	1	3	4
		% within Martial Status	25,0%	75,0%	100,0%
		% within Income	10,0%	30,0%	20,0%
		% of Total	5,0%	15,0%	20,0%
Total		Count	10	10	20
		% within Martial Status	50,0%	50,0%	100,0%
		% within Income	100,0%	100,0%	100,0%
		% of Total	50,0%	50,0%	100,0%

5 Άνδρες είναι Παντρεμένοι και έχουν ετήσιο εισόδημα <15.000€

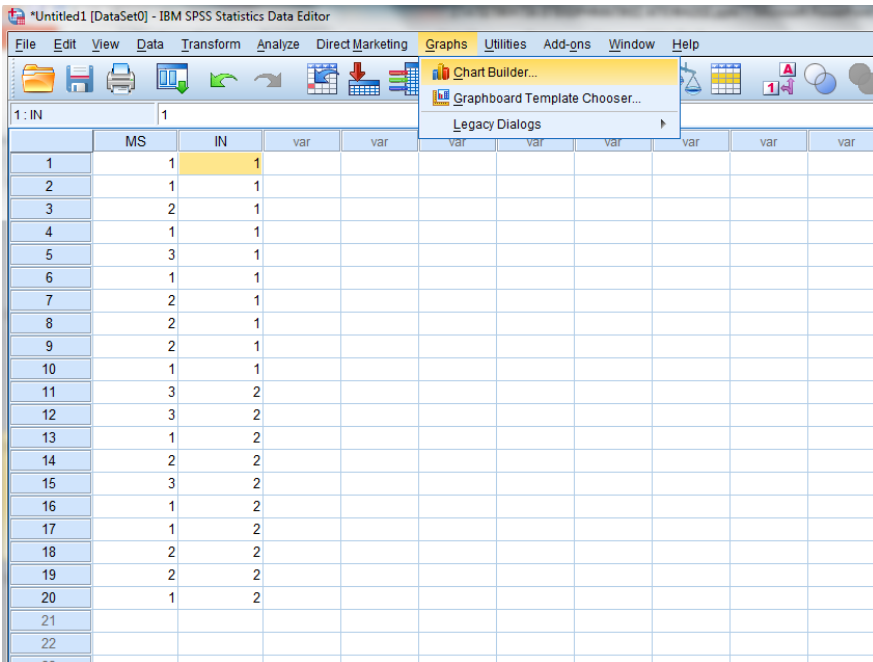
Το 55,6% των παντρεμένων ανδρών έχουν ετήσιο εισόδημα <15.000€

Το 50% των ανδρών με ετήσιο εισόδημα <15.000€ είναι παντρεμένοι

Το 25% των ανδρών είναι παντρεμένοι και έχουν ετήσιο εισόδημα <15.000€

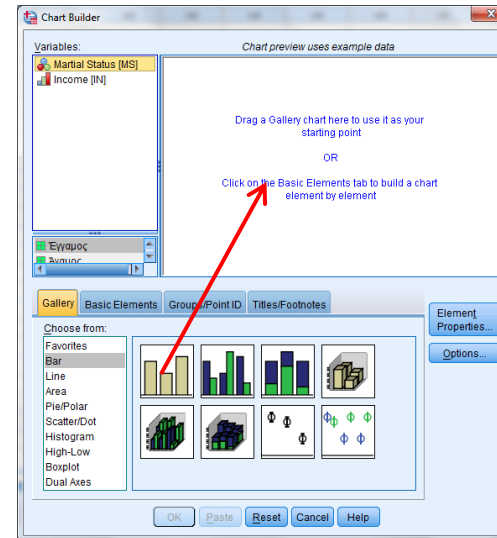
Γραφικές Μέθοδοι

► Ραβδόγραμμα 1 μεταβλητή



The screenshot shows the IBM SPSS Statistics Data Editor interface. The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Add-ons, Window, and Help. The 'Graphs' menu is open, showing options like 'Chart Builder...', 'Graphboard Template Chooser...', and 'Legacy Dialogs'. The data grid below shows two variables: MS (Marital Status) and IN (Income). The data is as follows:

	MS	IN	var	var	var	var	var	var	var	var	var
1	1	1									
2	1	1									
3	2	1									
4	1	1									
5	3	1									
6	1	1									
7	2	1									
8	2	1									
9	2	1									
10	1	1									
11	3	2									
12	3	2									
13	1	2									
14	2	2									
15	3	2									
16	1	2									
17	1	2									
18	2	2									
19	2	2									
20	1	2									
21											
22											
23											



Γραφικές Μέθοδοι

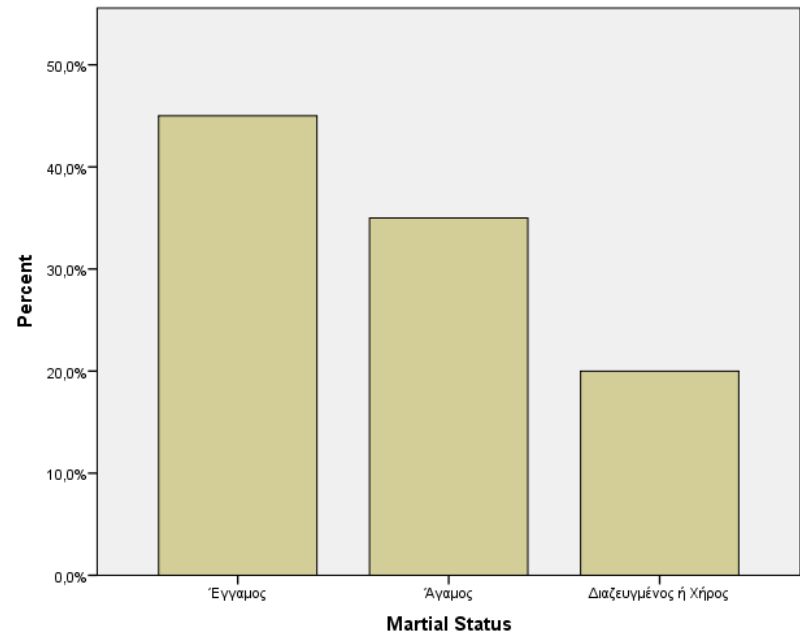
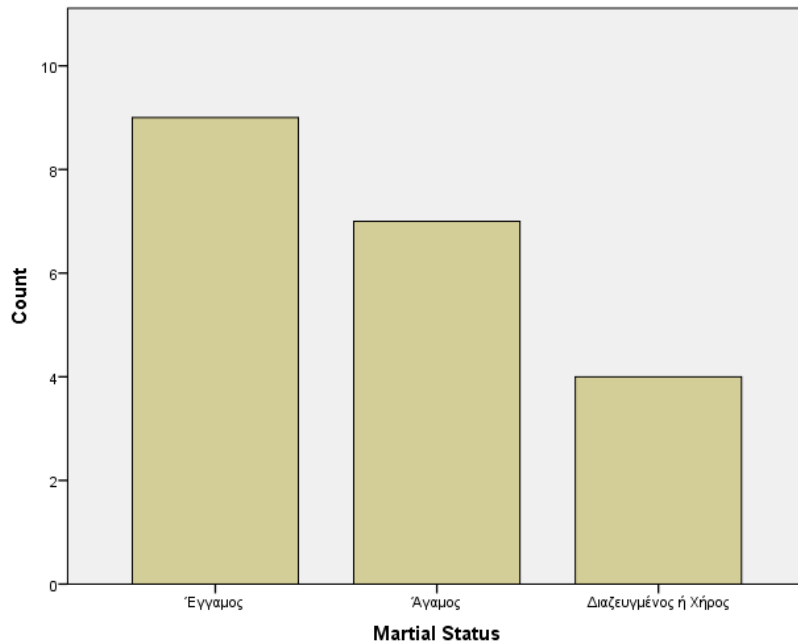
► Ραβδόγραμμα 1 μεταβλητή

The screenshot displays the IBM SPSS Statistics Data Editor interface. The main window shows a data table with columns 'MS' and 'IN'. The 'Chart Builder' dialog box is open, showing a preview of a bar chart with three bars. The 'Variables' list includes 'Marital Status (MS)' and 'Income (IN)'. The 'Choose from:' list is highlighted with a red box, and the 'Count' statistic is selected in the 'Element Properties' dialog box. The 'Element Properties' dialog box also shows the 'Statistics' section with 'Count' selected, and the 'Bar Style' set to 'Bar'.

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

Γραφικές Μέθοδοι

► Ραβδόγραμμα 1 μεταβλητή



Γραφικές Μέθοδοι

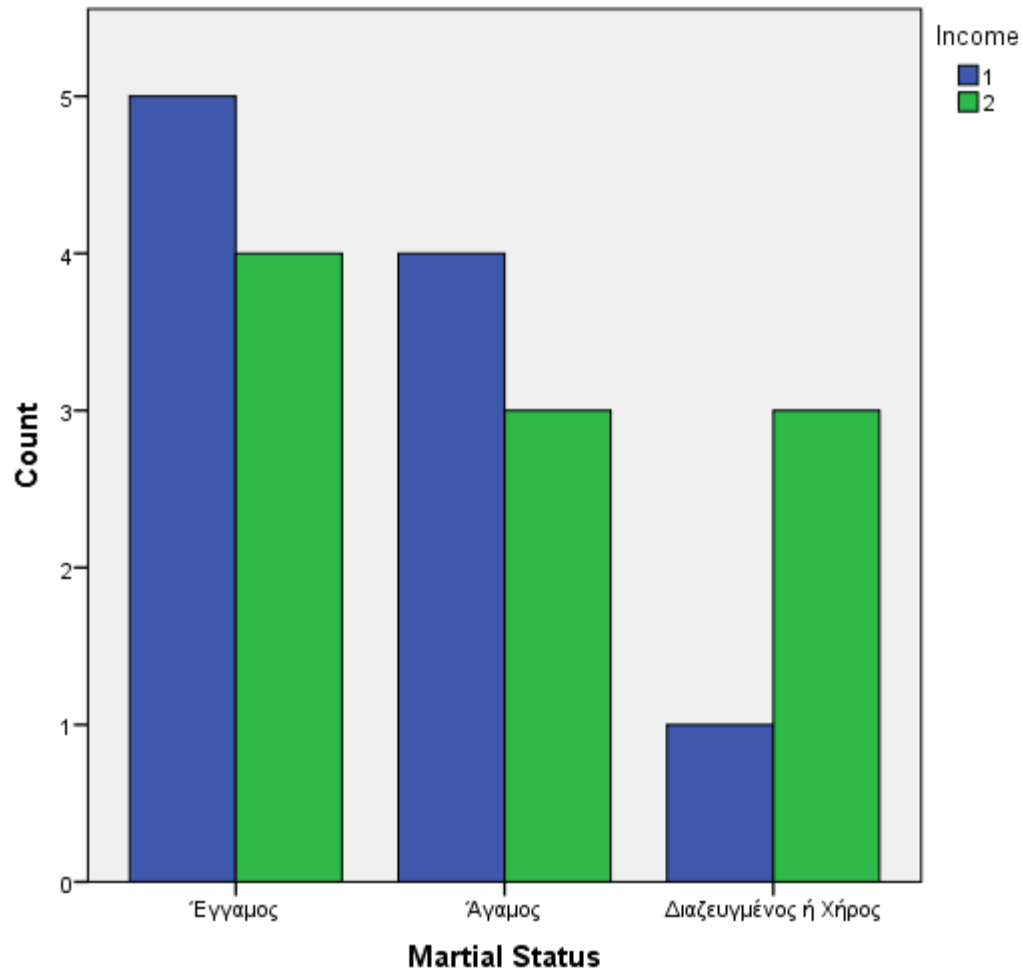
► Ραβδόγραμμα 2 μεταβλητές

The screenshot shows the IBM SPSS Statistics Data Editor interface. The main window displays a data table with columns 'MS' and 'IN'. The 'Chart Builder' dialog box is open, showing a 2D bar chart configuration. The chart displays 'Income [IN]' on the Y-axis and 'Marital Status [MS]' on the X-axis. The X-axis categories are 'Έγγαμος', 'Άγαμος', and 'Διαζευγμένος ή Χηρός'. The Y-axis is labeled 'Count'. The chart is clustered by 'Marital Status'. The 'Element Properties' dialog box is open, showing 'Bar1' with 'X-Axis1 (Bar1)', 'Y-Axis1 (Bar1)', and 'Group Color (Bar1)'. The 'Statistics' section shows 'Variable: Count' and 'Statistic: Count'. The 'Display error bars' section is checked, and 'Confidence intervals' is selected. The 'Bar Style' is set to 'Bar'.

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

Γραφικές Μέθοδοι

► Ραβδόγραμμα 2 μεταβλητές



Γραφικές Μέθοδοι

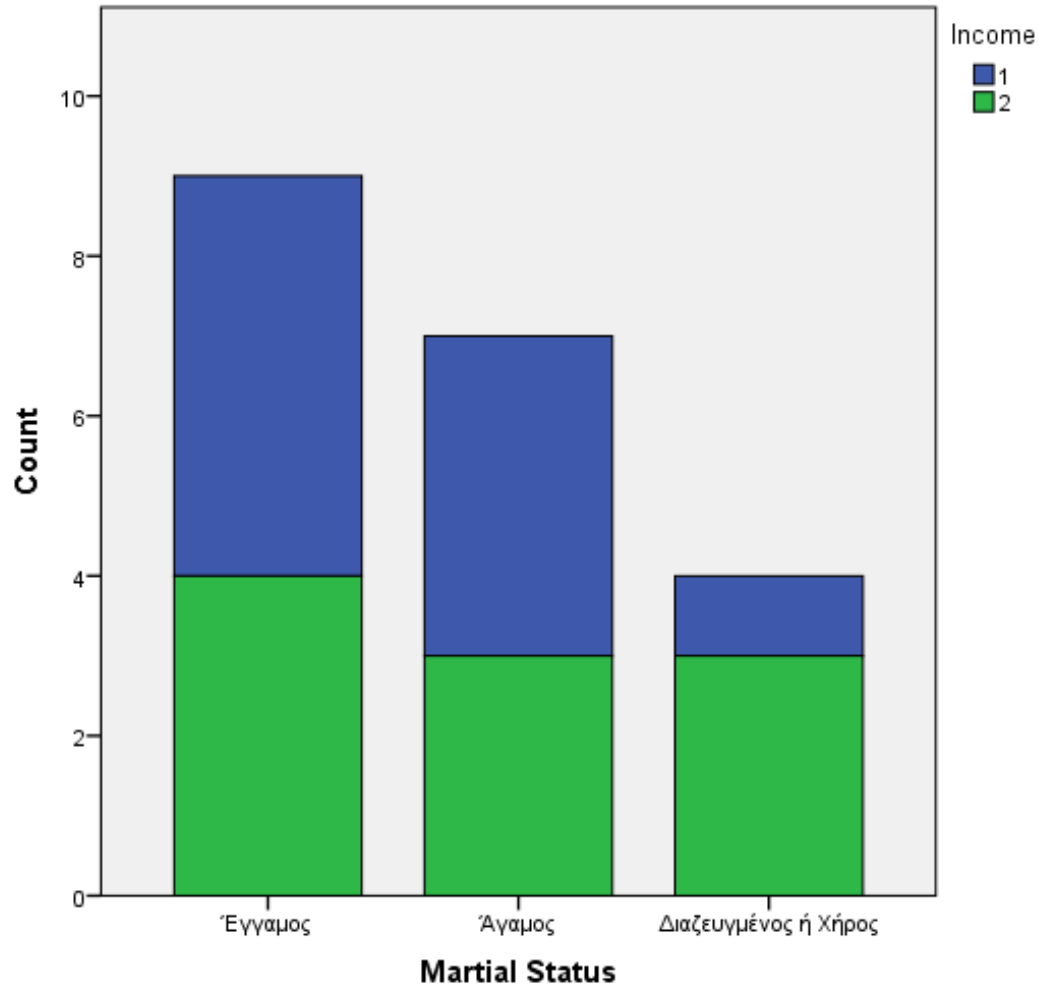
- ▶ Ραβδόγραμμα 2 μεταβλητές κατά στοιβές

The screenshot shows the IBM SPSS Statistics Data Editor interface. The main window displays a data table with columns 'MS' (Marital Status) and 'IN' (Income). The 'Chart Builder' dialog box is open, showing a 2D stacked bar chart. The variables 'Marital Status [MS]' and 'Income [IN]' are selected. The chart preview shows three bars representing marital statuses: 'Εγγάμος' (Married), 'Αγαμος' (Single), and 'Διαζευγμένος ή Χήρος' (Divorced or Widowed). The 'Income' variable is stacked within each bar. The 'Element Properties' dialog box is also open, showing settings for the 'Bar' element, including 'Confidence intervals' and 'Standard deviation' options.

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

Γραφικές Μέθοδοι

- ▶ Ραβδόγραμμα 2 μεταβλητές κατά στοιβες



Γραφήματα για ποσοτικά συνεχή δεδομένα

43

- ▶ Ιστογράμματα

- ▶ Θηκογράμματα

- ▶ Φυλλογραφήματα

- ▶ Αθροιστικό Πολυγωνικό Διάγραμμα
(*ogive*=αψίδα)



Κατασκευάζοντας Ένα Ιστόγραμμα ...

- ▶ Συλλέξτε τα Δεδομένα ✓
- ▶ Δημιουργήστε μία κατανομή συχνοτήτων για τα δεδομένα ...
- ▶ Πως;
- ▶ α) Καθορίστε τον αριθμό των **τάξεων** ...
- ▶ Πως; -> Κοιτάξτε τον παρακάτω πίνακα

Με 200 παρατηρήσεις, θα πρέπει να έχουμε μεταξύ 7 & 10 τάξεις...

Αριθμός των Παρατηρήσεων	Αριθμός Τάξεων
< 50	5-7
50-200	7-9
200-500	9-10
500-1000	10-11
1000-5000	11-13
5000-50000	13-17
> 50000	17-20

Εναλλακτικά, μπορούμε να χρησιμοποιήσουμε τον τύπο του Sturges: Αριθμός των τάξεων = $1 + 3.3 \log_{10}(n)$



Κατασκευάζοντας Ένα Ιστόγραμμα ... (συνεχεία)

- ▶ Συλλέξτε τα Δεδομένα ✓
- ▶ Δημιουργήστε μία κατανομή συχνοτήτων για τα δεδομένα ...

Πως;

α) Καθορίστε τον αριθμό των *τάξεων* [8] ✓

β) Καθορίστε το μήκος της κάθε τάξης ...

Πως;

Κοιτάξτε το εύρος των δεδομένων, δηλαδή,

Εύρος = Μέγιστη Παρατήρηση - Ελάχιστη Παρατήρ.

Εύρος = \$119.63 - \$0 = \$119.63

Τότε κάθε τάξη θα έχει μήκος:

Εύρος ÷ (# τάξεις) = 119.63 ÷ 8 ≈ 15



Κατασκευάζοντας Ένα Ιστόγραμμα ...(συνεχεία)

- ▶ Συλλέξτε τα Δεδομένα ✓
- ▶ Δημιουργήστε μία κατανομή συχνοτήτων για τα δεδομένα ...

Πως;

α) Καθορίστε τον αριθμό των *τάξεων* [8] ✓

β) Καθορίστε το μήκος της κάθε τάξης [15] ✓

γ) Τοποθέτησε τα δεδομένα στις τάξεις ...

Κάθε παρατήρηση μπορεί να ανήκει σε μόνο μία τάξη,

Οι τάξεις περιέχουν παρατηρήσεις κάτω φραγμάτων $<$ και \leq άνω φραγμάτων.

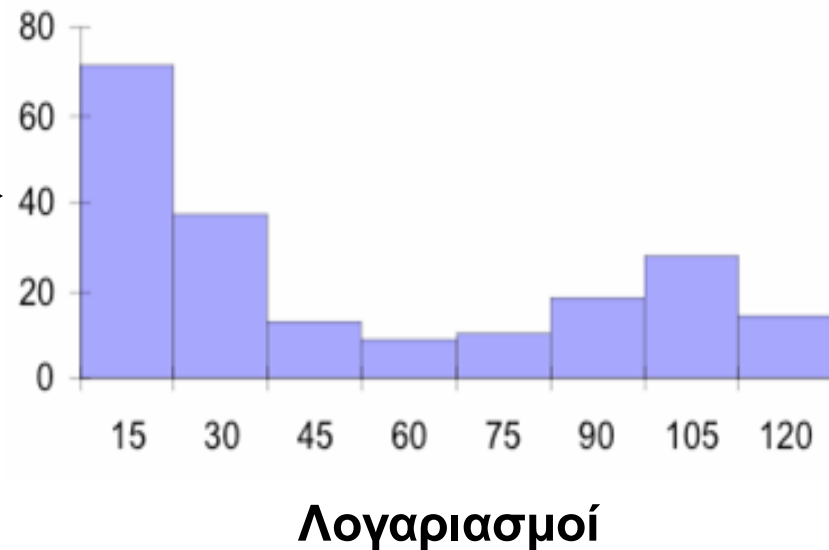


Κατασκευάζοντας Ένα Ιστόγραμμα ...(συνεχεία)

<u>Φράγματα τάξεων</u>	<u>Συχνότητα</u>
0 to 15	71
15 to 30	37
30 to 45	13
45 to 60	9
60 to 75	10
75 to 90	18
90 to 105	28
105 to 120	14
<u>Total</u>	<u>200</u>



Ιστόγραμμα Μακροπρόθεσμων Λογαριασμών

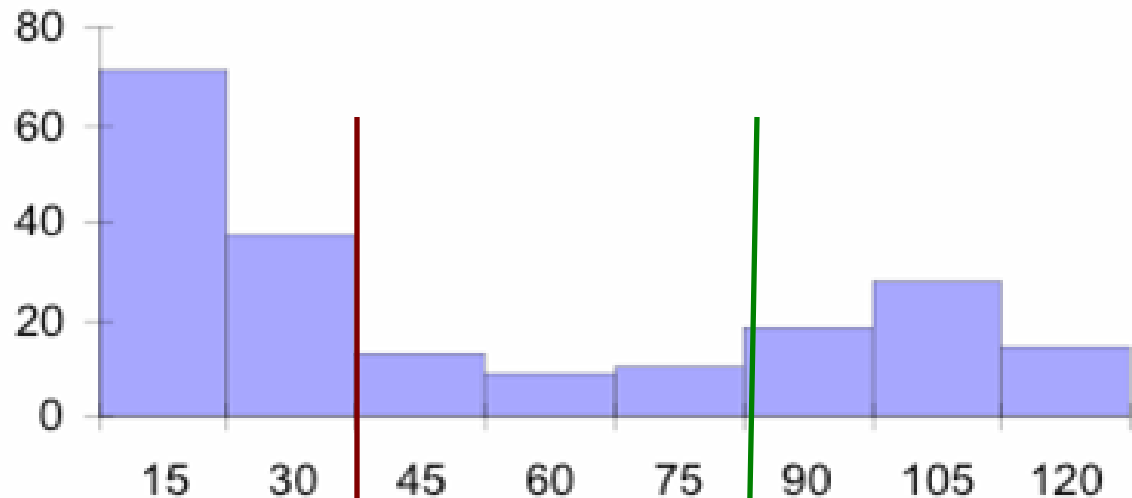


▶ 3) Σχεδιάστε ένα ιστόγραμμα ...



Ερμηνεία...

Ιστόγραμμα Μακροπρόθεσμων Λογαριασμών



Σχεδόν μισοί ($71+37=108$) από τους λογαριασμούς είναι «μικροί», π.χ. μικρότεροι από \$30

Λογαριασμοί

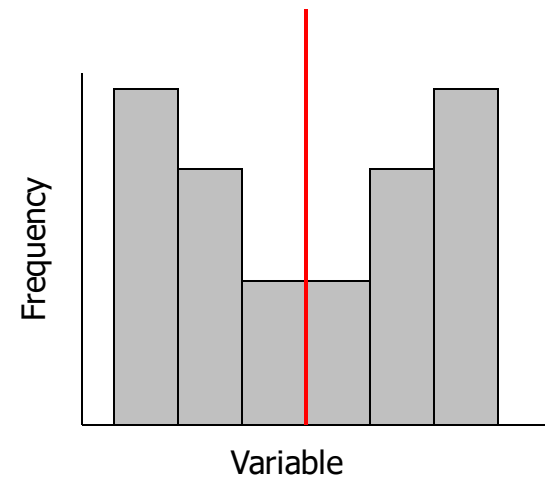
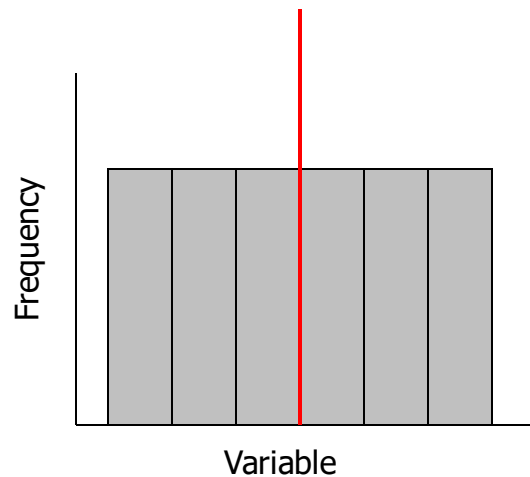
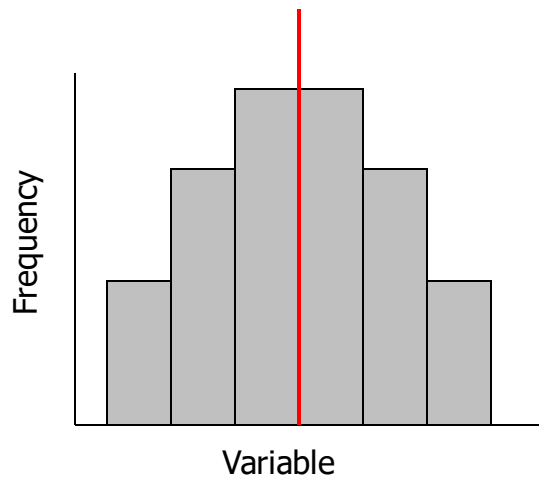
$(18+28+14=60) \div 200 = 30\%$
π.χ. σχεδόν ένα τρίτο των τηλεφωνικών λογαριασμών είναι \$90 ή περισσότερο.

Υπάρχουν μόνο μερικοί τηλεφωνικοί λογαριασμοί στη μέση.



▶ Συμμετρία

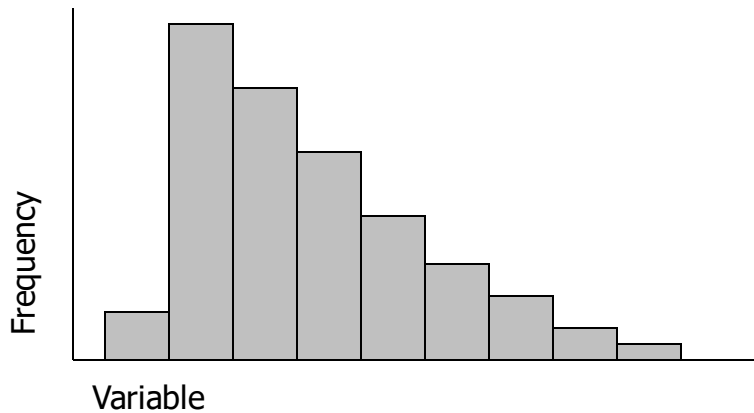
- ▶ Ένα ιστογράμμα λέγεται να είναι *συμμετρικό* εάν, όταν τραβήξουμε μία **κάθετη γραμμή** στο κέντρο του ιστογράμματος, τα δύο χωρισμένα κομμάτια είναι ακριβώς τα ίδια σε μορφή και σε μέγεθος:



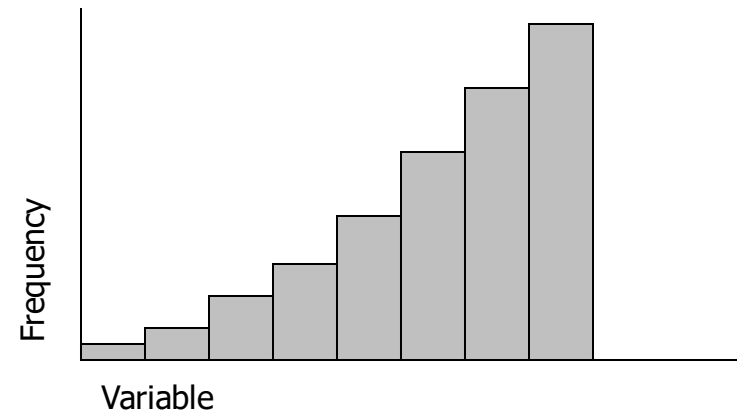
Μορφές Ιστογραμμάτων ...

▶ Λοξότητα

- ▶ Ένα λοξό ιστογράμμα είναι αυτό που έχει μια εκτεταμένη μακριά ουρά προς τα αριστερά ή προς τα δεξιά:



Λοξή προς τα Δεξιά



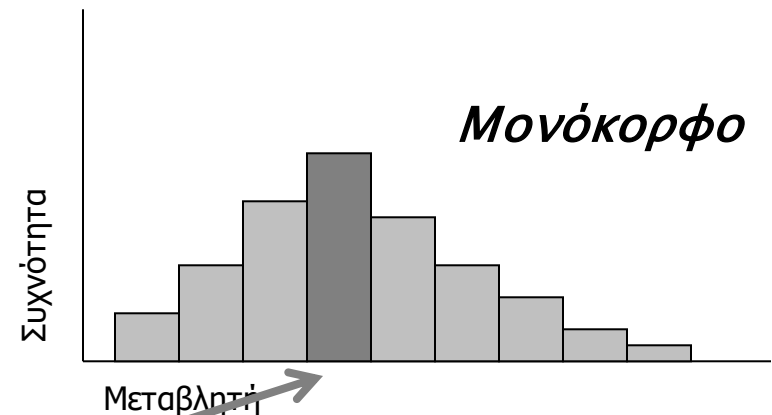
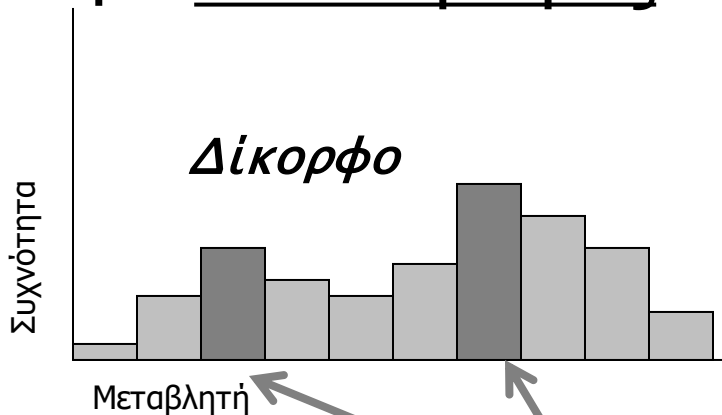
Λοξή προς τα Αριστερά



Μορφές Ιστογραμμάτων ...

► Κορυφές

- Ένα *μονόκορφο ιστογράμμα* είναι ένα ιστογράμμα με μία κορυφή, ενώ ένα *δίκορφο ιστογράμμα* είναι ένα ιστογράμμα με δύο κορυφές:



Η *κλάση κορυφής* είναι η κλάση με τον μεγαλύτερο αριθμό παρατηρήσεων

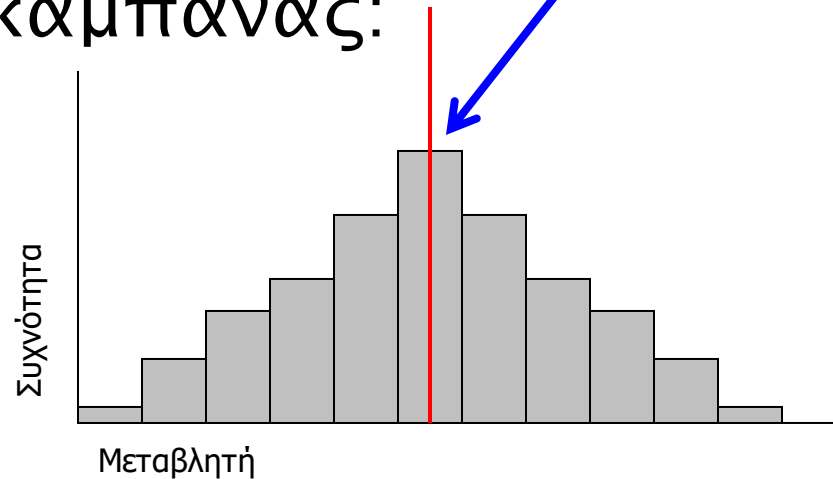


Μορφές Ιστογραμμάτων ...

- ▶ **Σχήμα Καμπάνας**
- ▶ Μία ειδική περίπτωση ενός **συμμετρικού** και **μονόκορφου** ιστογράμματος είναι αυτό με το σχήμα της καμπάνας:

Πολλές στατιστικές τεχνικές απαιτούν ότι ο πληθυσμός έχει το σχήμα της καμπάνας.

Σχεδιάζοντας το ιστογράμμα βοηθάει να επαληθεύσουμε την μορφή του πληθυσμού υπό εξέταση.



Σχήμα Καμπάνας

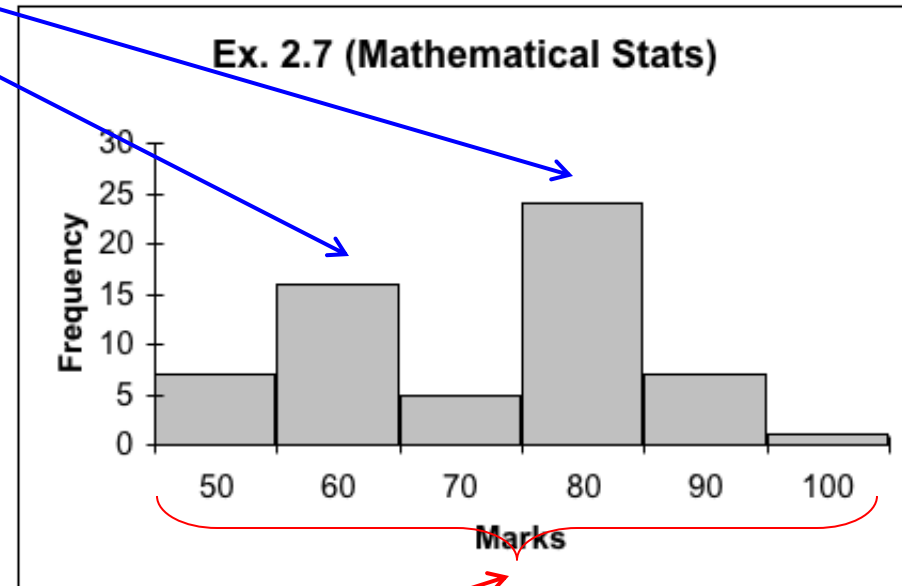
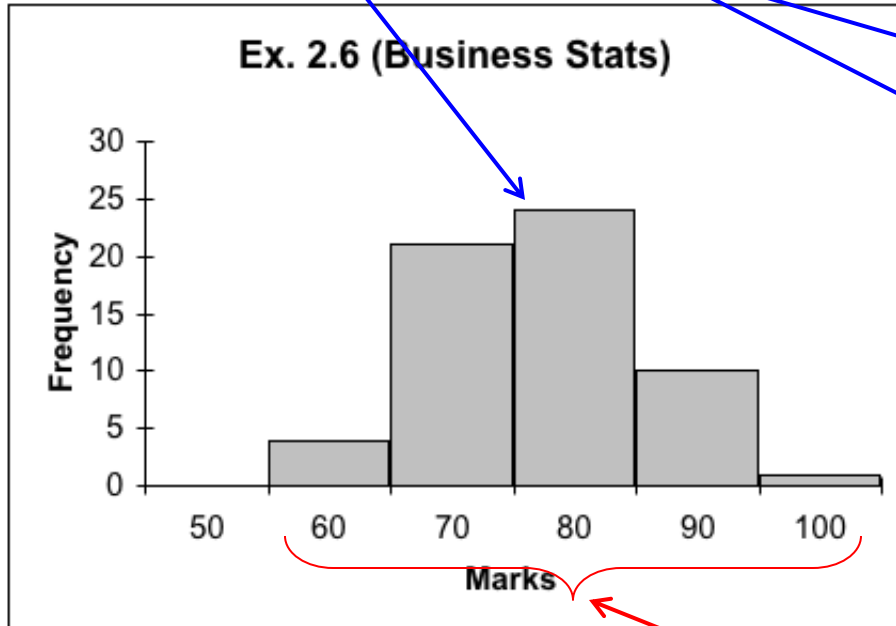


Σύγκριση Ιστογραμμάτων ...

- ▶ Συγκρίνουμε και αντιπαραθέτουμε τα ακόλουθα ιστογράμματα βασισμένοι σε δεδομένα

μονόκορφο & δίκορφο

Τα δύο μαθήματα έχουν διαφορετικές μορφές ιστογραμμάτων ...



Το εύρος των βαθμών (στενότερο | πλατύτερο)



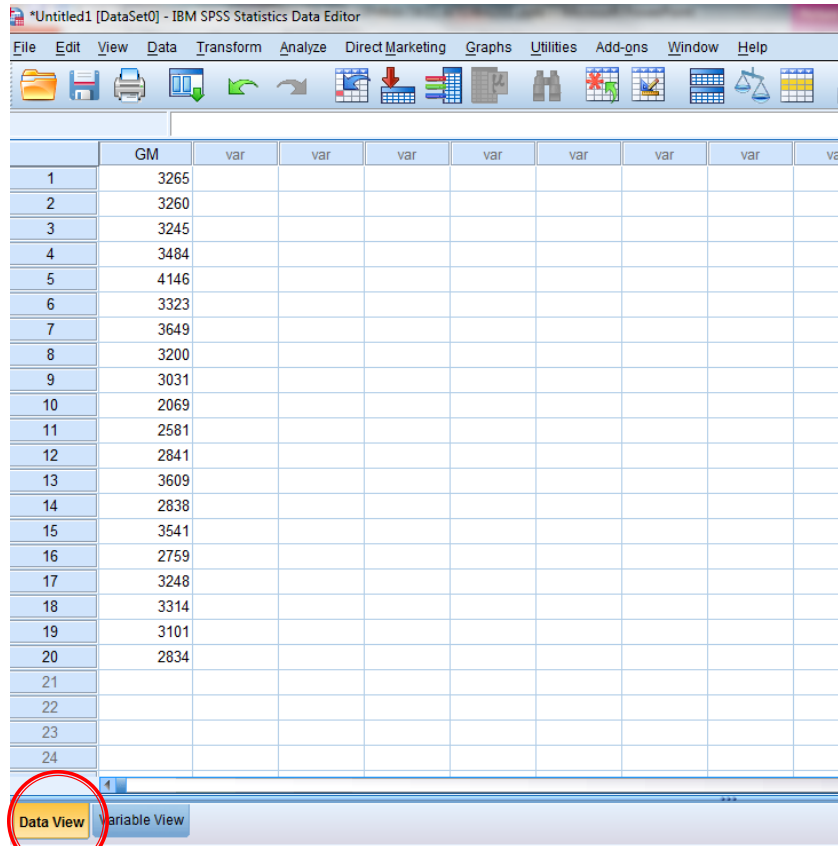
Ποσοτικές Μεταβλητές: Κατασκευή Ιστογραμμάτων στο SPSS

Παράδειγμα:

Τα δεδομένα αφορούν το ακαθάριστο κέρδος (τζίρος σε χιλ €) 20 γαλακτοκομικών εταιρειών το 2014 στην Ελλάδα

3265	3260	3245	3484	4146
3323	3649	3200	3031	2069
2581	2841	3606	2838	3541
2759	3248	3314	3101	2834

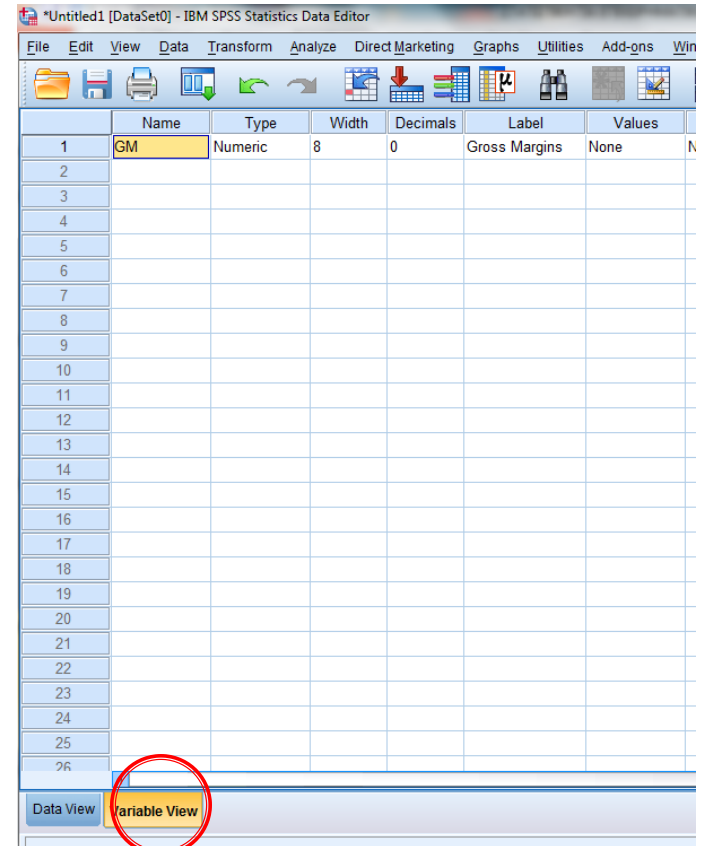
Ποσοτικές Μεταβλητές: Κατασκευή Ιστογραμμάτων στο SPSS



The screenshot shows the SPSS Data Editor window in Data View. The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains icons for file operations, editing, and analysis. The data grid has 24 rows and 10 columns. The first column is labeled 'GM' and contains numerical values. The other columns are labeled 'var'.

	GM	var	var	var	var	var	var	var	var
1	3265								
2	3260								
3	3245								
4	3484								
5	4146								
6	3323								
7	3649								
8	3200								
9	3031								
10	2069								
11	2581								
12	2841								
13	3609								
14	2838								
15	3541								
16	2759								
17	3248								
18	3314								
19	3101								
20	2834								
21									
22									
23									
24									

The 'Data View' tab is selected and circled in red.



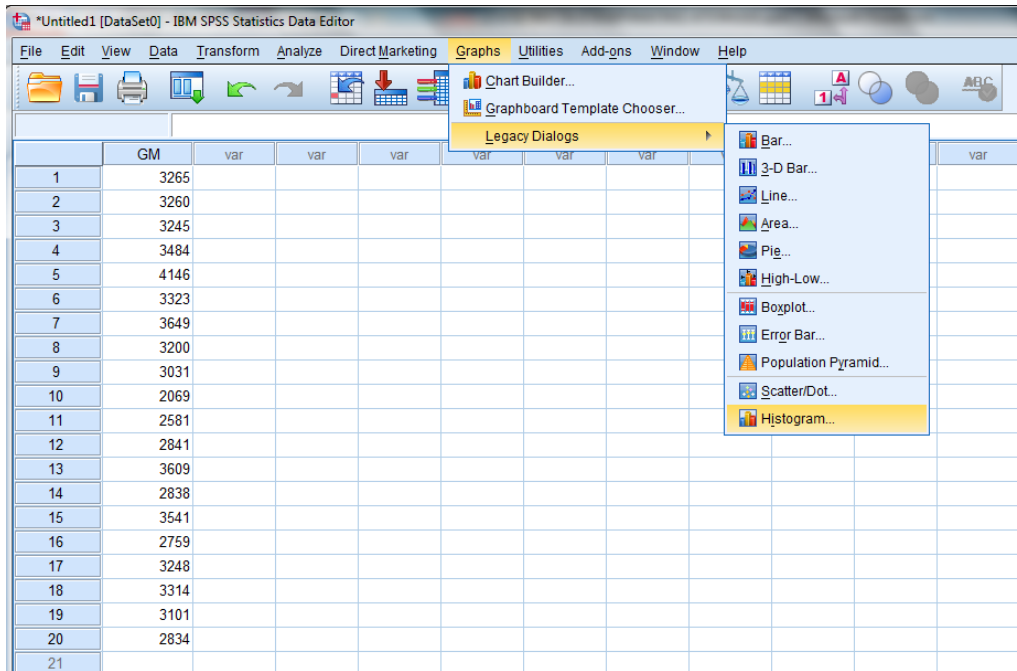
The screenshot shows the SPSS Data Editor window in Variable View. The menu bar and toolbar are the same as in the Data View screenshot. The variable grid has 26 rows and 7 columns. The first row is for the variable 'GM', which is highlighted in yellow. The other rows are empty.

	Name	Type	Width	Decimals	Label	Values
1	GM	Numeric	8	0	Gross Margins	None
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						

The 'Variable View' tab is selected and circled in red.

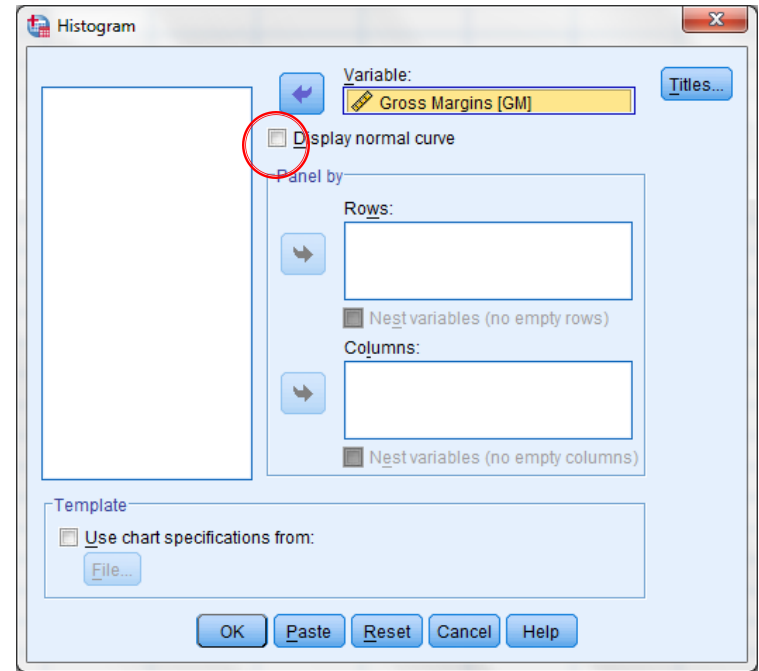
Ποσοτικές Μεταβλητές: Κατασκευή Ιστογραμμάτων στο SPSS

Graphs>Legacy Dialogs>Histogram..



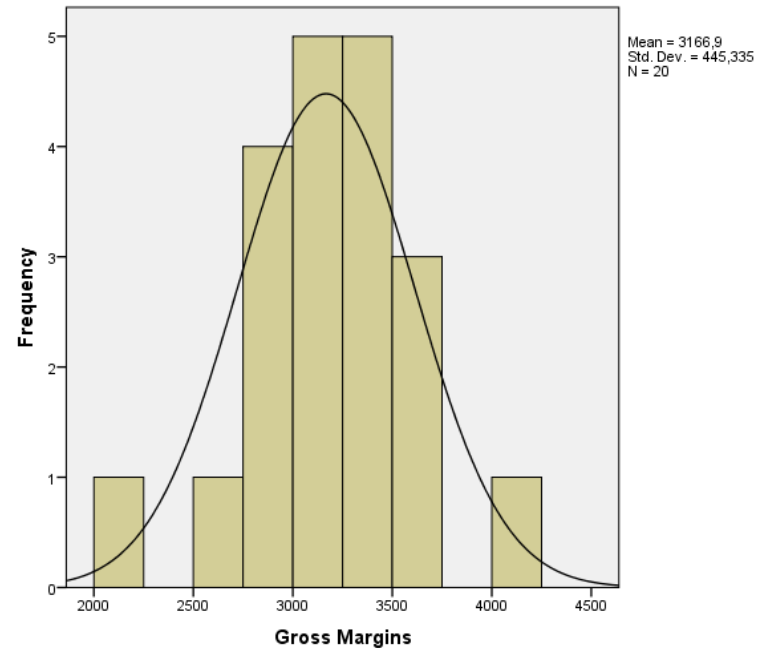
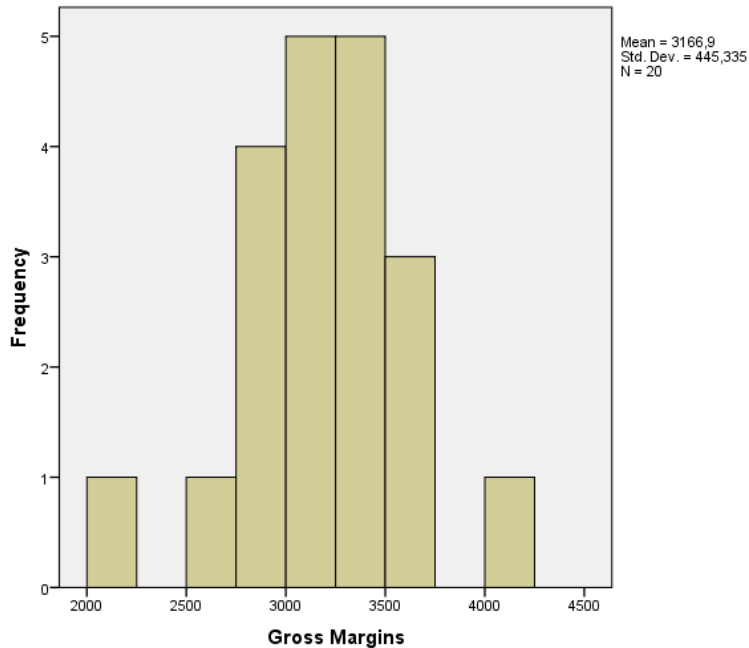
The screenshot shows the SPSS Statistics Data Editor interface. The 'Legacy Dialogs' menu is open, and the 'Histogram...' option is highlighted. The data editor displays a table with 21 rows and 5 columns. The first column is labeled 'GM' and contains numerical values ranging from 2069 to 3265. The other four columns are labeled 'var'.

	GM	var	var	var	var
1	3265				
2	3260				
3	3245				
4	3484				
5	4146				
6	3323				
7	3649				
8	3200				
9	3031				
10	2069				
11	2581				
12	2841				
13	3609				
14	2838				
15	3541				
16	2759				
17	3248				
18	3314				
19	3101				
20	2834				
21					



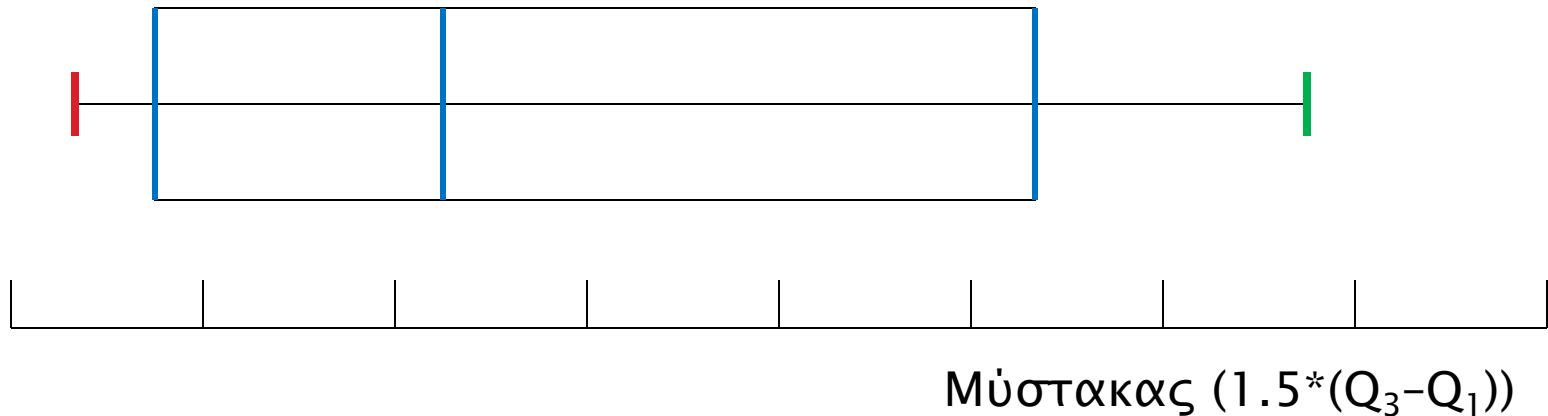
The screenshot shows the 'Histogram' dialog box in SPSS. The 'Variable:' field is set to 'Gross Margins [GM]'. The 'Display normal curve' checkbox is checked and circled in red. The 'Panel by' section is empty. The 'Rows' and 'Columns' sections are also empty. The 'Template' section has the 'Use chart specifications from:' checkbox unchecked. The 'OK', 'Paste', 'Reset', 'Cancel', and 'Help' buttons are visible at the bottom.

Ποσοτικές Μεταβλητές: Κατασκευή Ιστογραμμάτων στο SPSS



Θηκόγραμμα (Box Plot)...

- ▶ Το **θηκόγραμμα** (*box plot*) είναι μια γραφική τεχνική η οποία σχεδιάζει **πέντε** στατιστικές:
- ▶ την **μικρότερη** την **μεγαλύτερη** παρατήρηση, το **πρώτο**, **δεύτερο**, και τρίτο **τεταρτημόριο**.

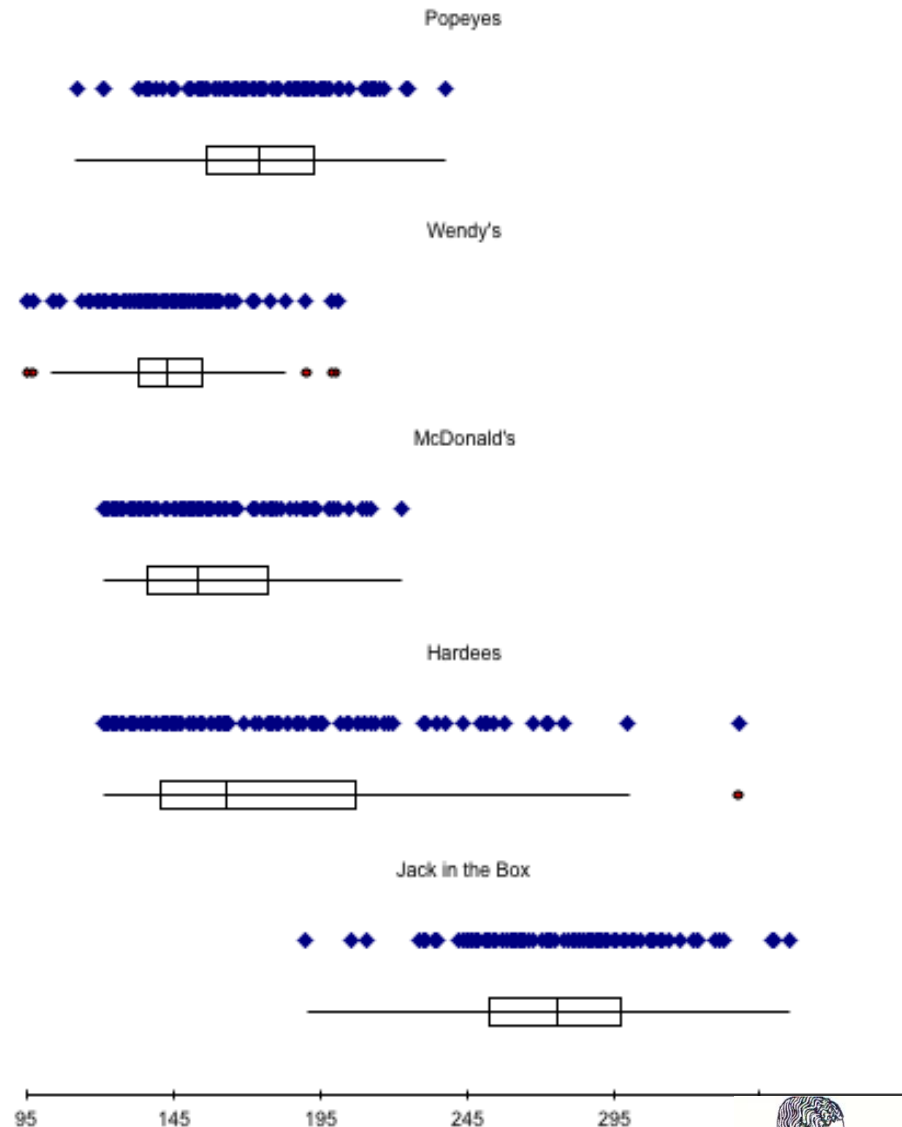


Οι προεκτεινόμενες γραμμές στα αριστερά και στα δεξιά καλούνται **μύστακες** (**whiskers**). Κάθε σημείο που πέφτει έξω από τους μύστακες καλείται ακραία τιμή (outlier). Οι μύστακες επεκτείνονται προς τα έξω στο μικρότερο από το (1.5 φορά το ενδοτεταρτημοριακό εύρος) ή στην ποιο τελευταία τιμή η οποία δεν είναι ακραία.



Θηκόγραμμα (Box Plot)...

- ▶ Αυτά τα θηκογράμματα είναι από δεδομένα με χρόνος εξυπηρέτησης πελατών.
- ▶ Ο χρόνος εξυπηρέτησης στα Wendy's είναι ο μικρότερος και λιγότερο μεταβλητός.
- ▶ Στα Hardee's ο χρόνος εξυπηρέτησης έχει την μεγαλύτερη μεταβλητότητα, ενώ στα Jack-in-the-Box έχει τον μεγαλύτερο χρόνο εξυπηρέτησης.



Ποσοτικές Μεταβλητές: Κατασκευή Θηκογραμμάτων στο SPSS

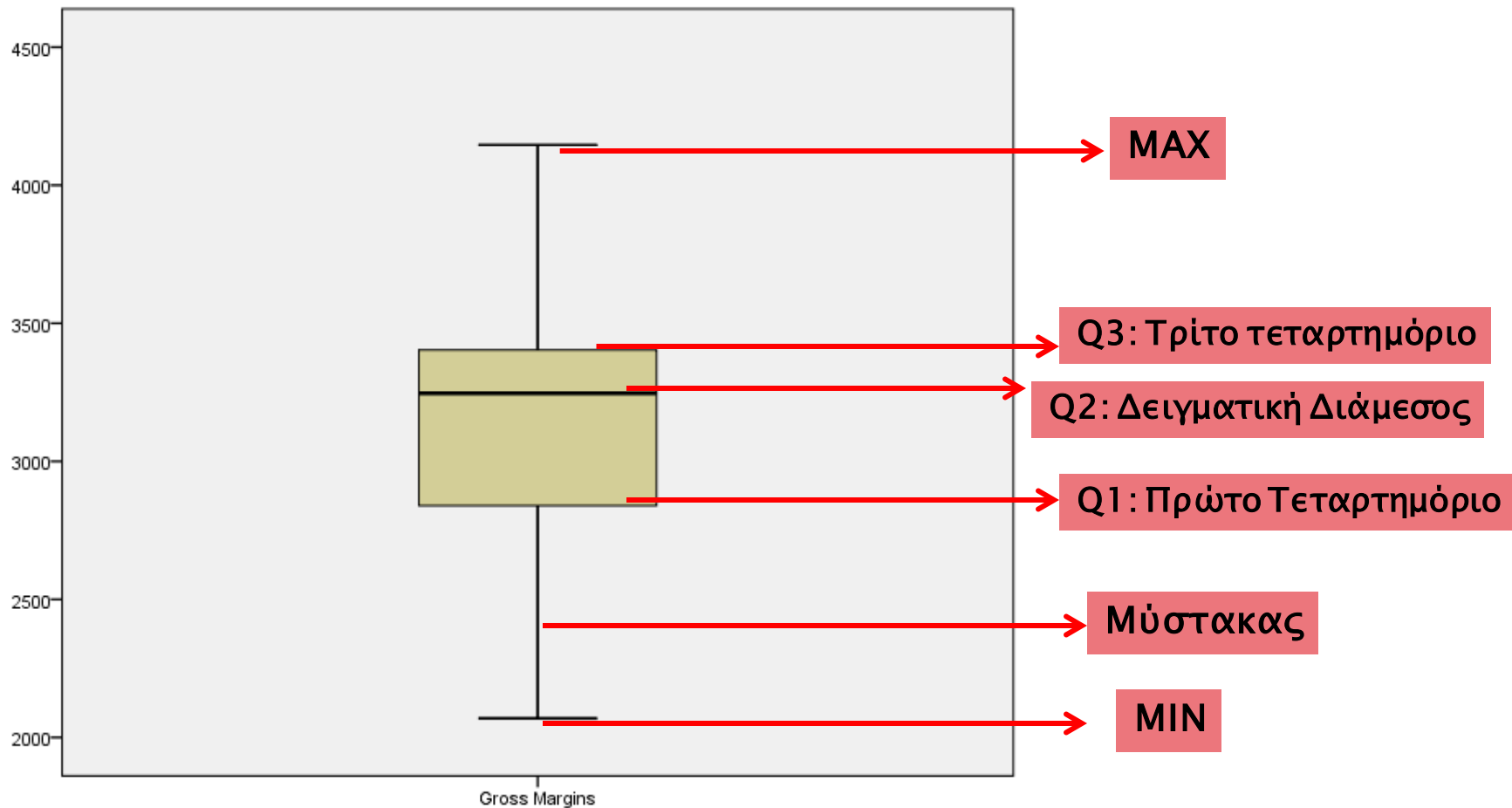
The screenshot shows the IBM SPSS Statistics Data Editor interface. The main window displays a data table with 21 rows and 8 columns. The first column is labeled 'GM' and contains numerical values. The 'Graphs' menu is open, and 'Boxplot...' is highlighted. The 'Legacy Dialogs' submenu is also visible.

	GM	var	var	var	var	var	var
1	3265						
2	3260						
3	3245						
4	3484						
5	4146						
6	3323						
7	3649						
8	3200						
9	3031						
10	2069						
11	2581						
12	2841						
13	3609						
14	2838						
15	3541						
16	2759						
17	3248						
18	3314						
19	3101						
20	2834						
21							

The 'Boxplot' dialog box is shown. The 'Simple' option is selected and highlighted with a red box. The 'Summaries of separate variables' radio button is also selected and highlighted with a red box. The 'Define' button is visible at the bottom.

The 'Define Simple Boxplot: Summaries of Separate Variables' dialog box is shown. The 'Boxes Represent' field contains 'Gross Margins [GM]'. The 'Label Cases by' field is empty. The 'Panel by' section has 'Rows' and 'Columns' fields, both empty. The 'Nest variables (no empty rows)' and 'Nest variables (no empty columns)' checkboxes are unchecked. The 'Options...' button is visible in the top right corner. The 'OK' button is highlighted at the bottom.

Ποσοτικές Μεταβλητές: Κατασκευή Θηκογραμμάτων στο SPSS

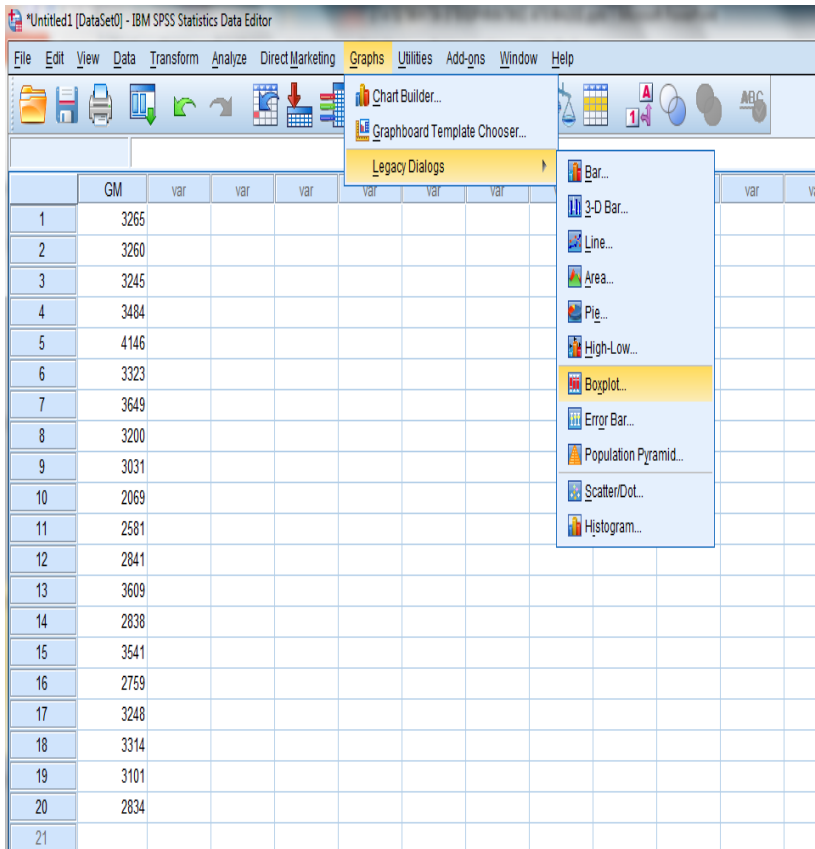


Ποσοτικές Μεταβλητές: Κατασκευή Θηκογραμμάτων στο SPSS-2 Μεταβλητές

Τα θηκογράμματα είναι χρήσιμα για να συγκρίνουμε επίσης δύο δείγματα. Έστω ότι επιπλέον του 1^{ου} παραδείγματος, έχουμε και τα ακαθάριστα κέρδη 20 οινοποιείων στην Ελλάδα το 2014

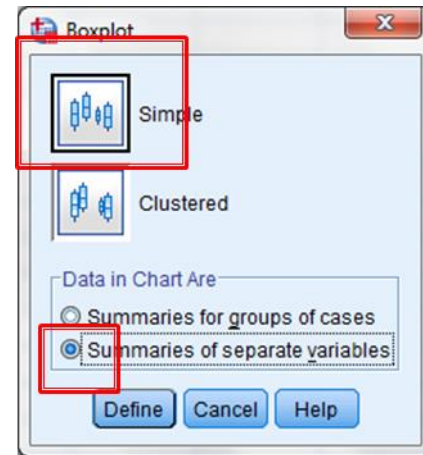
3257	3328	2576	2753	3262
3658	2836	3245	3236	3205
3601	3305	3478	3023	2842
3105	4153	2066	3541	2838

Ποσοτικές Μεταβλητές: Κατασκευή Θηκογραμμάτων στο SPSS-2 Μεταβλητές

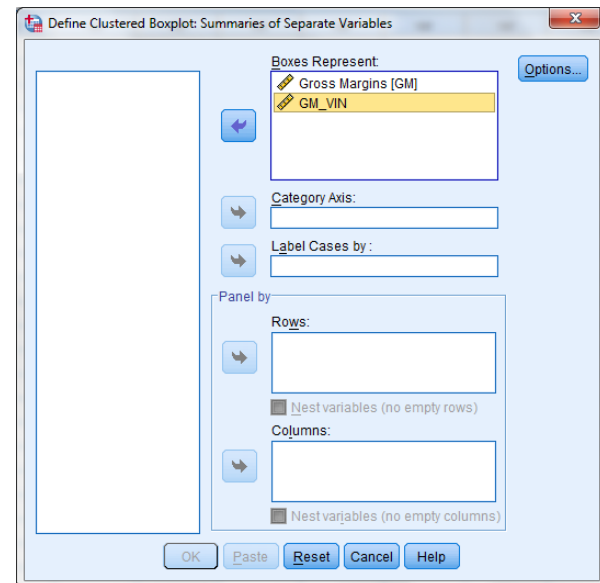


The screenshot shows the IBM SPSS Statistics Data Editor interface. The 'Graphs' menu is open, and 'Boxplot...' is selected. The data table below shows the following values for the 'GM' variable:

	GM	var	var	var	var	var	var
1	3265						
2	3260						
3	3245						
4	3484						
5	4146						
6	3323						
7	3649						
8	3200						
9	3031						
10	2069						
11	2581						
12	2841						
13	3609						
14	2838						
15	3541						
16	2759						
17	3248						
18	3314						
19	3101						
20	2834						
21							

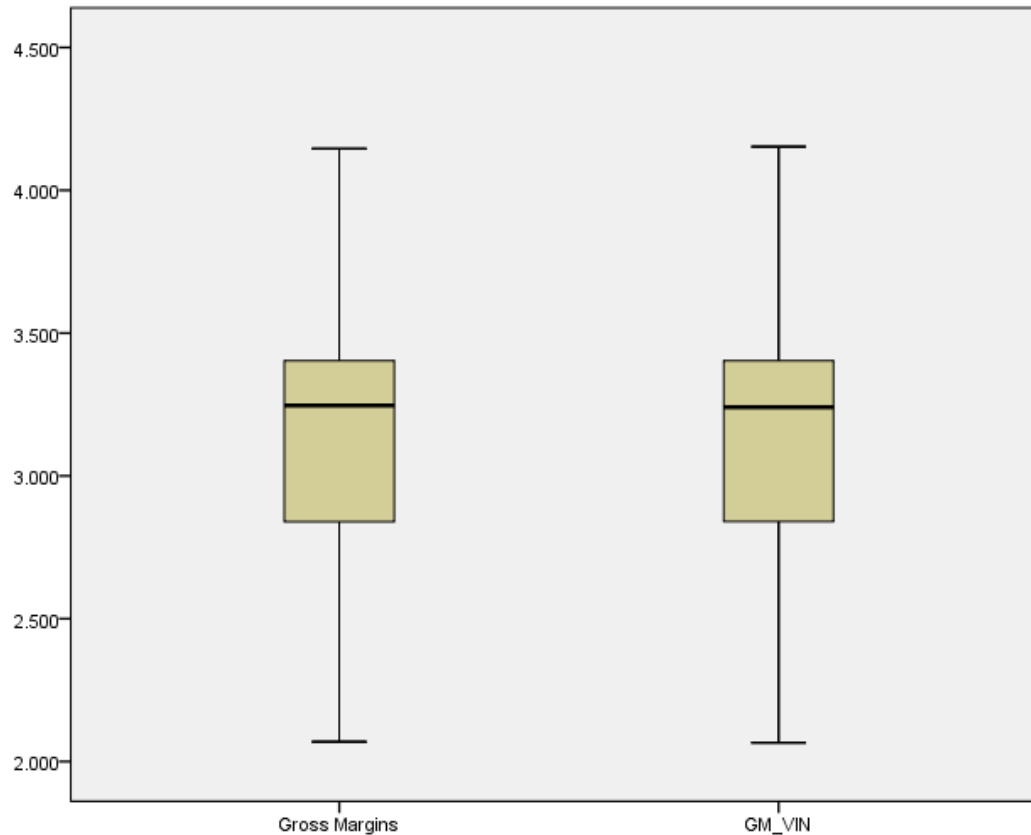


The 'Boxplot' dialog box is shown. The 'Simple' option is selected, and 'Summaries of separate variables' is chosen under 'Data in Chart Are'. The 'Define' button is highlighted.



The 'Define Clustered Boxplot: Summaries of Separate Variables' dialog box is shown. The 'Boxes Represent' list contains 'Gross Margins (GM)' and 'GM_VIN'. The 'Category Axis' and 'Label Cases by' fields are empty. The 'Panel by' section has 'Rows' and 'Columns' fields, both empty. The 'Nest variables' checkboxes are unchecked. The 'OK' button is highlighted.

Ποσοτικές Μεταβλητές: Κατασκευή Θηκογραμμάτων στο SPSS-2 Μεταβλητές



Ποσοτικές Μεταβλητές: Κατασκευή Θηκογραμμάτων στο SPSS-2 Μεταβλητές Μία ποσοτική και μία κατηγορική

Ισοδύναμα θα μπορούσαμε να είχαμε θεωρήσει ότι έχουμε μία ποσοτική μεταβλητή (GM) και μία κατηγορική μεταβλητή (1: Γαλακτοκομική επιχείρηση και 2: Οινοποιεία) και να συγκρίνουμε τις τιμές της ποσοτικής μεταβλητής ανά κατηγορία της κατηγορικής μεταβλητής

Ποσοτικές Μεταβλητές: Κατασκευή Θηκογραμμάτων στο SPSS-2 Μεταβλητές Μία ποσοτική και μία κατηγορική

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

	GM	SEC	var	var	var	var	var	v
1	3265	1						
2	3260	1						
3	3245	1						
4	3484	1						
5	4146	1						
6	3323	1						
7	3649	1						
8	3200	1						
9	3031	1						
10	2069	1						
11	2581	1						
12	2841	1						
13	3609	1						
14	2838	1						
15	3541	1						
16	2759	1						
17	3248	1						
18	3314	1						
19	3101	1						
20	2834	1						
21	3257	2						
22	3328	2						
23	2576	2						
24	2753	2						

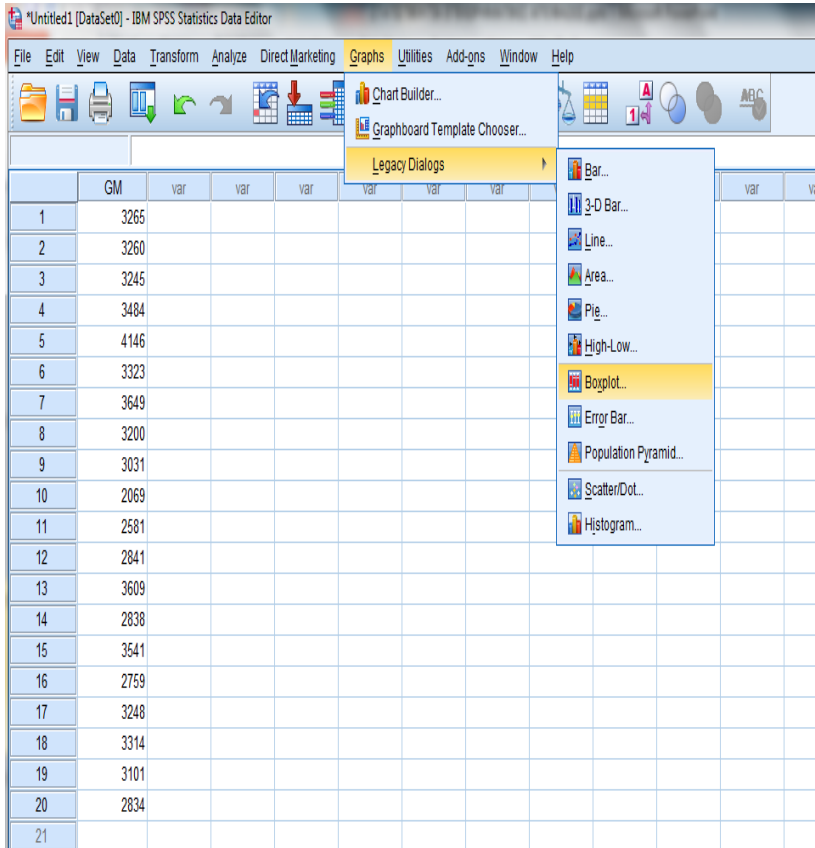
Data View Variable View

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	GM	Numeric	8	0	Gross Margins	None	None	8	Right	Scale	Input
2	SEC	Numeric	8	0	Sector	None	None	8	Right	Nominal	Input
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											

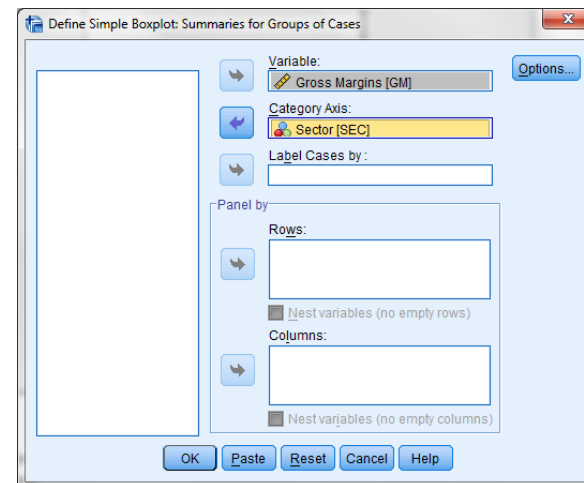
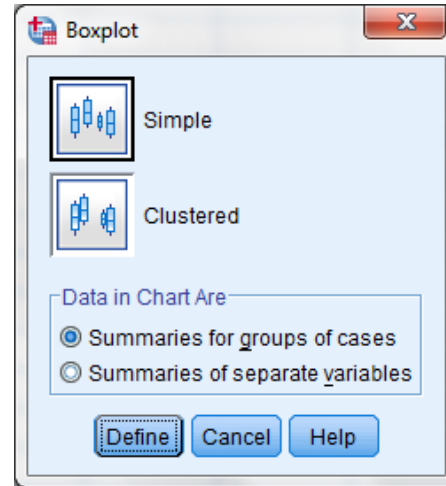
Data View Variable View

Ποσοτικές Μεταβλητές: Κατασκευή Θηκογραμμάτων στο SPSS-2 Μεταβλητές Μία ποσοτική και μία κατηγορική

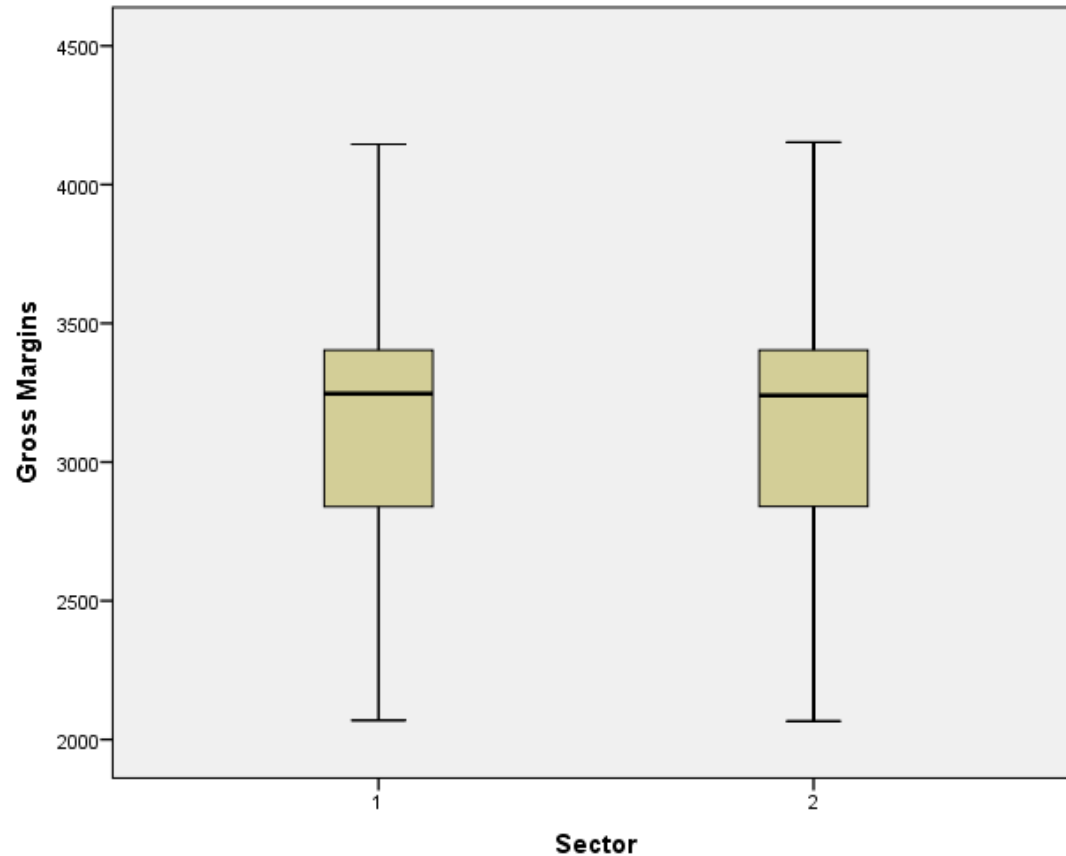


The screenshot shows the SPSS Statistics Data Editor interface. The 'Graphs' menu is open, and 'Boxplot...' is highlighted. The data table below shows a single variable 'GM' with 21 rows of values.

	GM	var	var	var	var	var	var
1	3265						
2	3260						
3	3245						
4	3484						
5	4146						
6	3323						
7	3649						
8	3200						
9	3031						
10	2069						
11	2581						
12	2841						
13	3609						
14	2838						
15	3541						
16	2759						
17	3248						
18	3314						
19	3101						
20	2834						
21							



Ποσοτικές Μεταβλητές: Κατασκευή Θηκογραμμάτων στο SPSS-2 Μεταβλητές Μία ποσοτική και μία κατηγορική



Απεικόνιση Φυλλογραφημάτων... 69

- ▶ Διατηρεί πληροφορίες σχετικά με ατομικές παρατηρήσεις που θα χανόταν με την δημιουργία ενός ιστογράμματος.
- ▶ Διασπάστε κάθε παρατήρηση σε δύο μέρη, σε ένα **στέλεχος** και σε ένα **φύλλο**.

Π.χ. Τιμή Παρατήρησης: 42.19

Υπάρχουν αρκετοί τρόποι να την διασπάσουμε ...

Μπορούμε να την διασπάσουμε στο κόμμα:

Ή να την διασπάσουμε στην θέση των «δεκάδων»

(ενώ στρογγυλοποιούμε στον κοντινότερο ακέραιο, στην θέση «μονάδων»)

Στέλεχος	Φύλλο
42	19
4	2



Απεικόνιση Φυλλογραφημάτων

- ▶ Συνεχίστε την διαδικασία για όλες τις παρατηρήσεις. Ύστερα χρησιμοποιήστε τα «φύλλα» για τις ομάδες και κάθε φύλλο γίνεται μέρος του ιστογράμματος, π.χ. στο παρακάτω παράδειγμα:

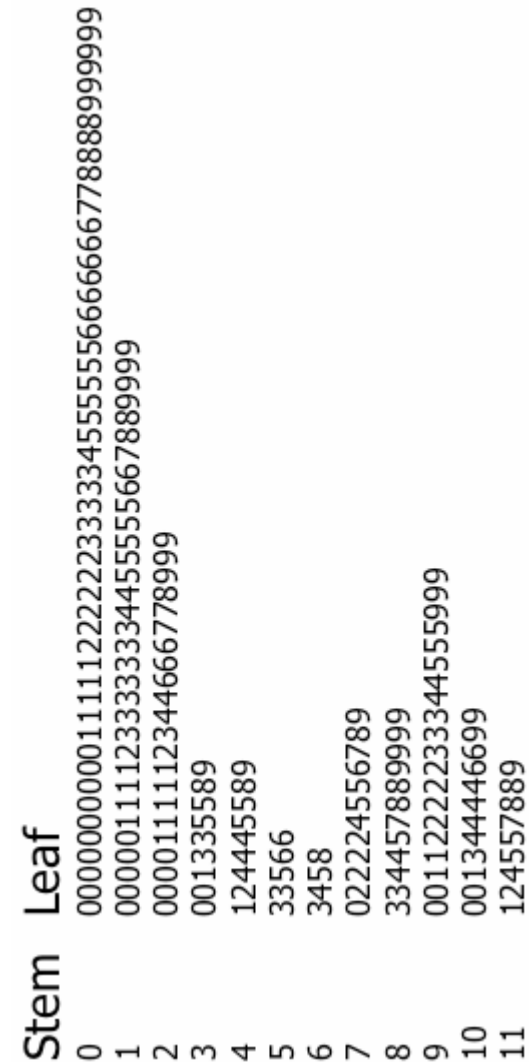
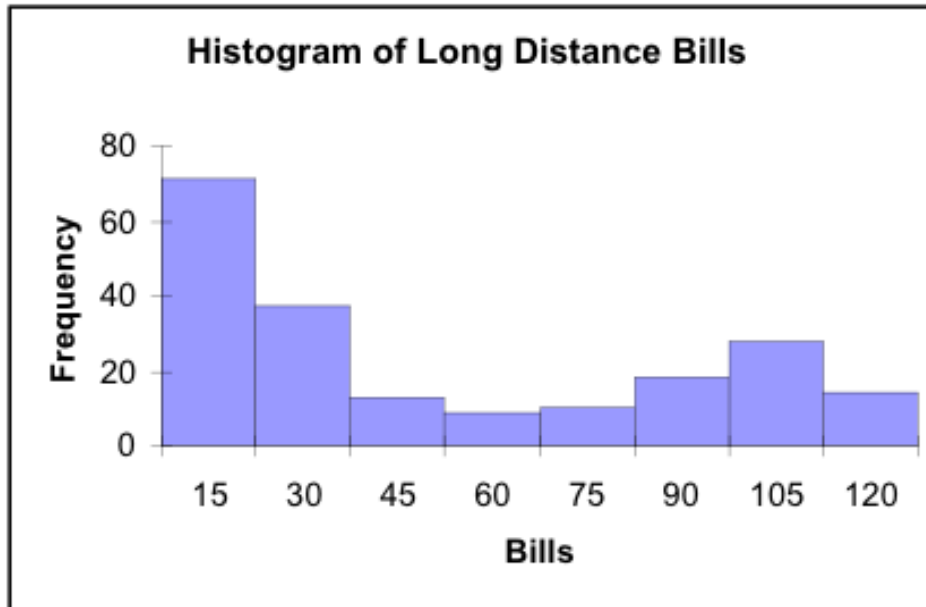
Στελέχη Φύλλα

0	0000000000111112222223333345555556666666778888999999
1	000001111233333334455555667889999
2	0000111112344666778999
3	001335589
4	12 4445589
5	33566
6	3458
7	022224556789
8	334457889999
9	00112222233344555999
10	001344446699
11	124557889

Έτσι, έχουμε ακόμα πρόσβαση στις τιμές των αρχικών δεδομένων!



Ιστόγραμμα και Φυλλογραφήματα...⁷¹



Κατασκευή Φυλλογράμματος στο SPSS

The screenshot shows the SPSS Statistics Data Editor interface. The 'Analyze' menu is open, and the 'Explore...' option is selected. The data table below shows two variables: GM and SEC.

	GM	SEC
1	3265	
2	3260	
3	3245	
4	3484	
5	4146	
6	3323	
7	3649	
8	3200	
9	3031	
10	2069	
11	2581	
12	2841	
13	3609	
14	2838	
15	3541	
16	2759	
17	3248	
18	3314	
19	3101	
20	2834	
21	3257	
22	3328	
23	2576	2
24	2722	2

The 'Explore' dialog box is shown with the following settings:

- Dependent List: Gross Margins [GM]
- Factor List: Sector [SEC]
- Label Cases by: (empty)
- Display: Both Statistics Plots

The 'Plots...' button is circled in red.

The 'Explore: Plots' sub-dialog box is shown with the following settings:

- Boxplots: Factor levels together Dependents together None
- Descriptive: Stem-and-leaf Histogram
- Normality plots with tests
- Spread vs Level with Levene Test: None Power estimation Transformed Power: Natural log Untransformed

Κατασκευή Φυλλογράμματος στο SPSS

Gross Margins Stem-and-Leaf Plot for
SEC= 1

Frequency	Stem & Leaf
1.00	2 . 0
5.00	2 . 57888
10.00	3 . 0122222334
3.00	3 . 566
1.00	4 . 1

Stem width: 1000
Each leaf: 1 case(s)

Gross Margins Stem-and-Leaf Plot for
SEC= 2

Frequency	Stem & Leaf
1.00	2 . 0
5.00	2 . 57888
10.00	3 . 0122222334
3.00	3 . 566
1.00	4 . 1

Stem width: 1000
Each leaf: 1 case(s)

Η Καμπύλη S ή καμπύλη αθροιστικών συχνοτήτων...

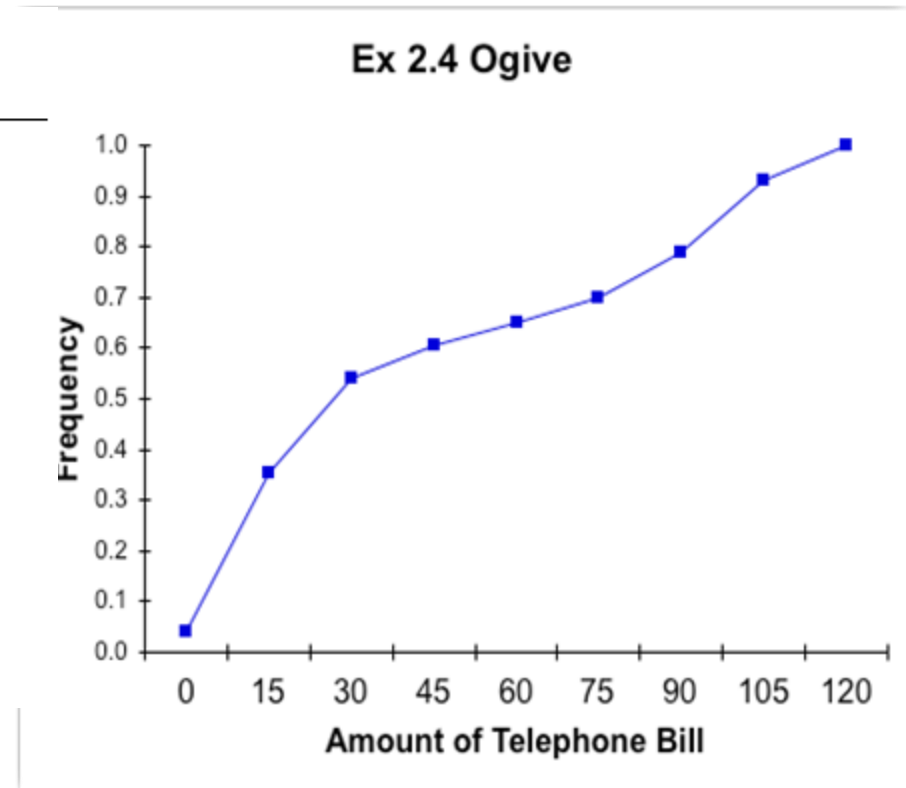
- ▶ Είναι ένα γράφημα μιας **κατανομής αθροιστικών συχνοτήτων**.
- ▶ 1) Υπολογίστε σχετικές συχνότητες.
- ▶ 2) Υπολογίστε **αθροιστικές σχετικές συχνότητες**.
- ▶ 3) Σχεδιάστε σε γράφημα τις αθροιστικές σχετικές συχνότητες ...



Η Καμπύλη S...

Table 2.8 Cumulative Relative Frequency Distribution for Example 2.4

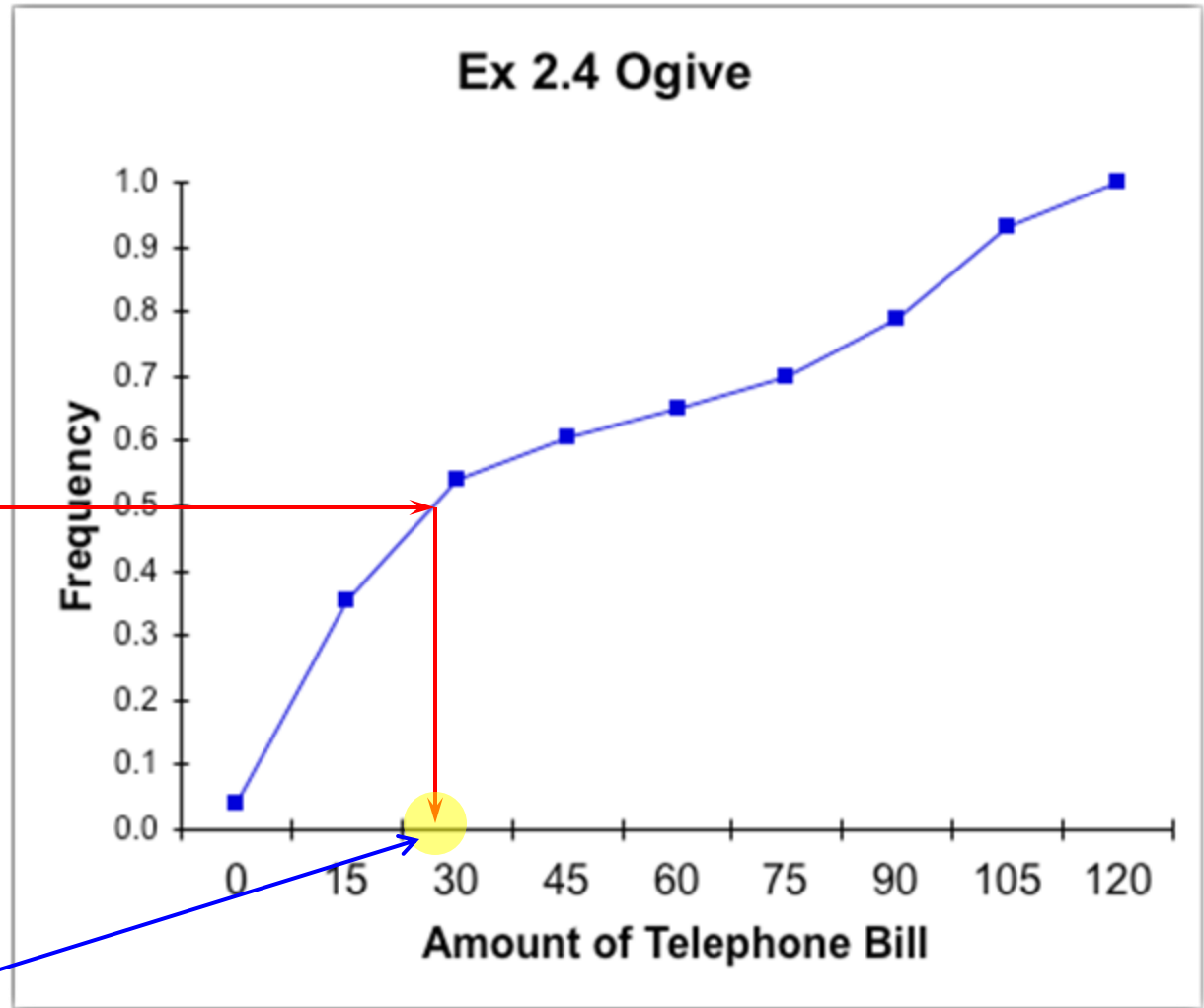
Class Limits	Relative Frequency	Cumulative Relative Frequency
0 to 15	$71/200 = .355$	$71/200 = .355$
15 to 30	$37/200 = .185$	$108/200 = .540$
30 to 45	$13/200 = .065$	$121/200 = .605$
45 to 60	$9/200 = .045$	$130/200 = .650$
60 to 75	$10/200 = .05$	$140/200 = .700$
75 to 90	$18/200 = .09$	$158/200 = .790$
90 to 105	$28/200 = .14$	$186/200 = .930$
105 to 120	$14/200 = .07$	$200/200 = 1.00$



Η Καμπύλη S ...

Η καμπύλη S μπορεί να χρησιμοποιηθεί για να απαντήσει ερωτήσεις όπως:

Τι ποσό τηλεφώνου είναι το 50^ο ποσοστημόριο;



“περίπου \$35”



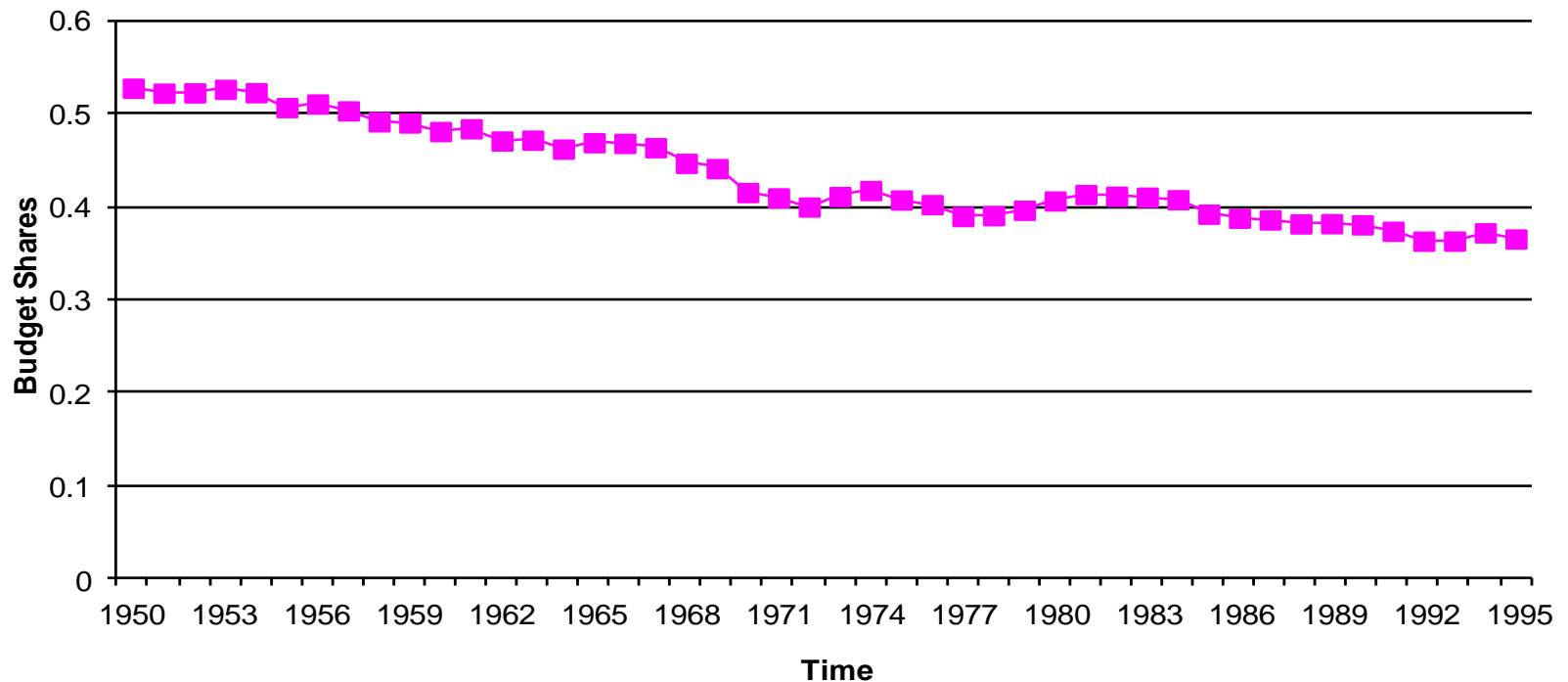
Δεδομένα Χρονολογικών Σειρών ...

- ▶ Παρατηρήσεις οι οποίες παίρνονται την ίδια χρονική στιγμή ονομάζονται **διαστρωματικά δεδομένα**.
- ▶ Παρατηρήσεις οι οποίες παίρνονται σε διαδοχικές χρονικές στιγμές καλούνται δεδομένα **χρονολογικών σειρών**.
- ▶ Δεδομένα χρονολογικών σειρών παρουσιάζονται με **χρονοδιαγράμματα (*line chart*)**, τα οποία σχεδιάζουν την τιμή της μεταβλητής στον κάθετο άξονα σε σχέση ως προς τις χρονικές περιόδους στον οριζόντιο άξονα.



Γραμμικό Διάγραμμα

Εξέλιξη του μεριδίου δαπάνης για αγορά τροφίμων την περίοδο 1950-1995



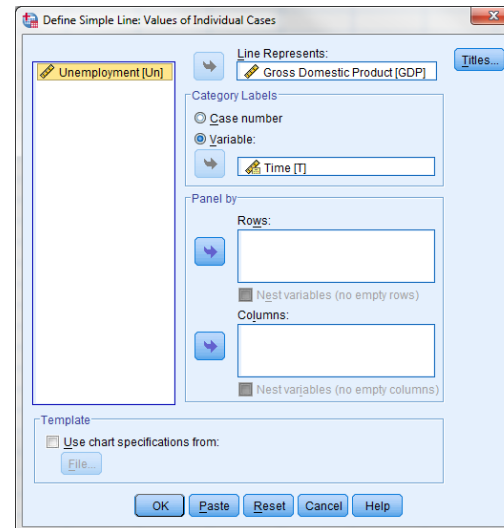
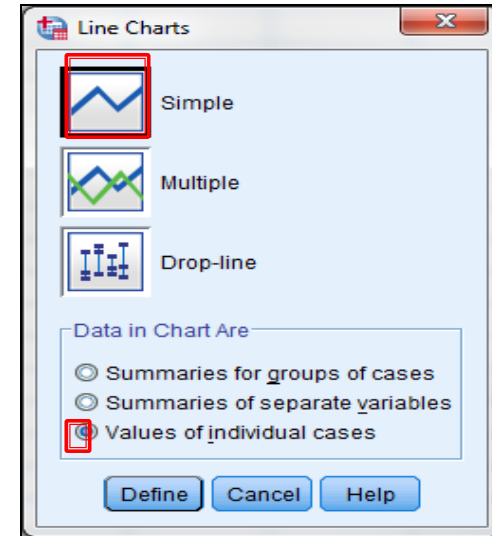
—■— Food at Home



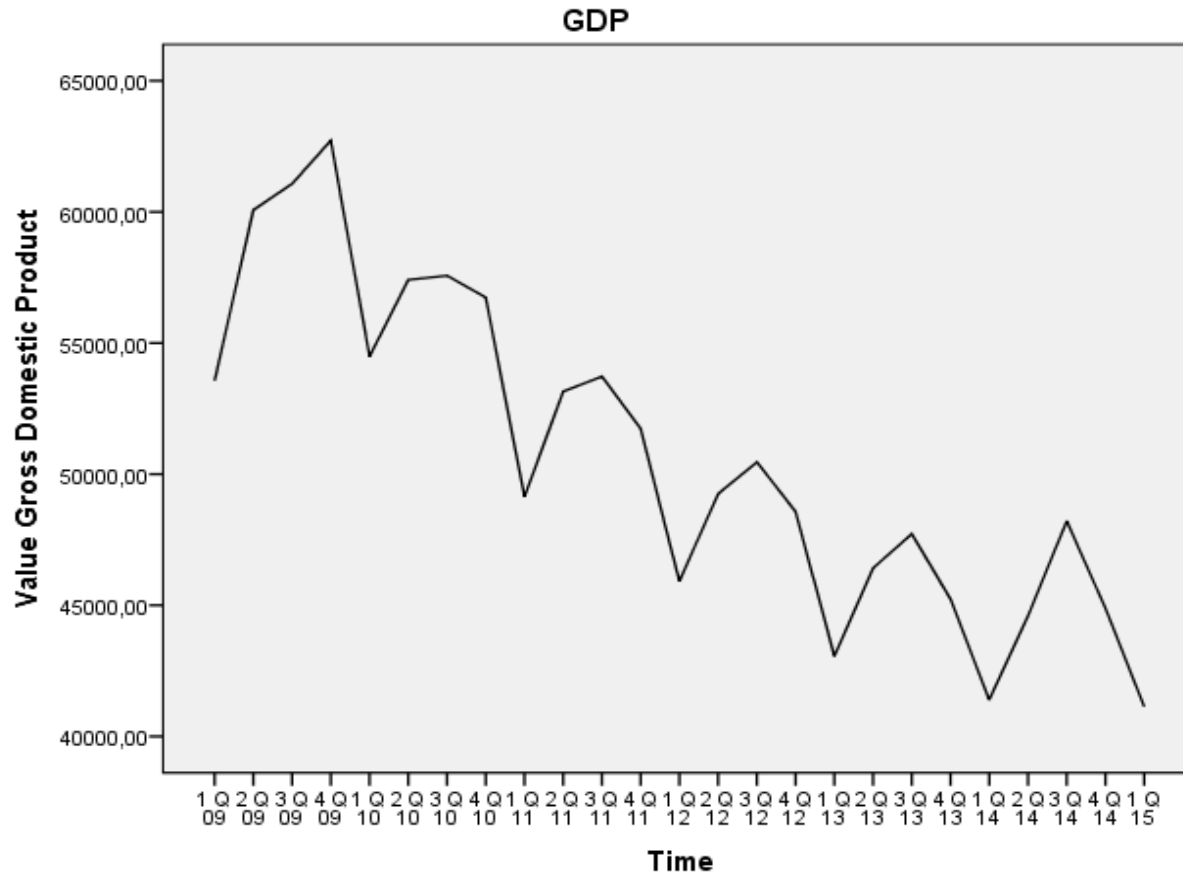
Γραμμικό Διάγραμμα στο SPSS

Graphs > Legacy Dialogs > Line

	GDP	Un	T	var
1	53563.00	476.70	1 Q 09	
2	60074.00	456.60	2 Q 09	
3	61072.00	477.90	3 Q 09	
4	62722.00	528.50	4 Q 09	
5	54490.00	600.20	1 Q 10	
6	57411.00	604.50	2 Q 10	
7	57570.00	631.80	3 Q 10	
8	56739.00	720.70	4 Q 10	
9	49139.00	799.60	1 Q 11	
10	53152.00	815.60	2 Q 11	
11	53727.00	883.40	3 Q 11	
12	51734.00	1028.60	4 Q 11	
13	45922.00	1119.00	1 Q 12	
14	49251.00	1163.00	2 Q 12	
15	50464.00	1218.30	3 Q 12	
16	48567.00	1279.90	4 Q 12	
17	43059.00	1335.90	1 Q 13	
18	46420.00	1327.80	2 Q 13	
19	47722.00	1320.20	3 Q 13	
20	45236.00	1337.10	4 Q 13	
21	41394.00	1342.30	1 Q 14	
22	44591.00	1280.10	2 Q 14	
23	48213.00	1229.37	3 Q 14	
24	44884.00	1245.85	4 Q 14	
25	41135.00	1272.54	1 Q 15	
26				
27				



Γραμμικό Διάγραμμα στο SPSS



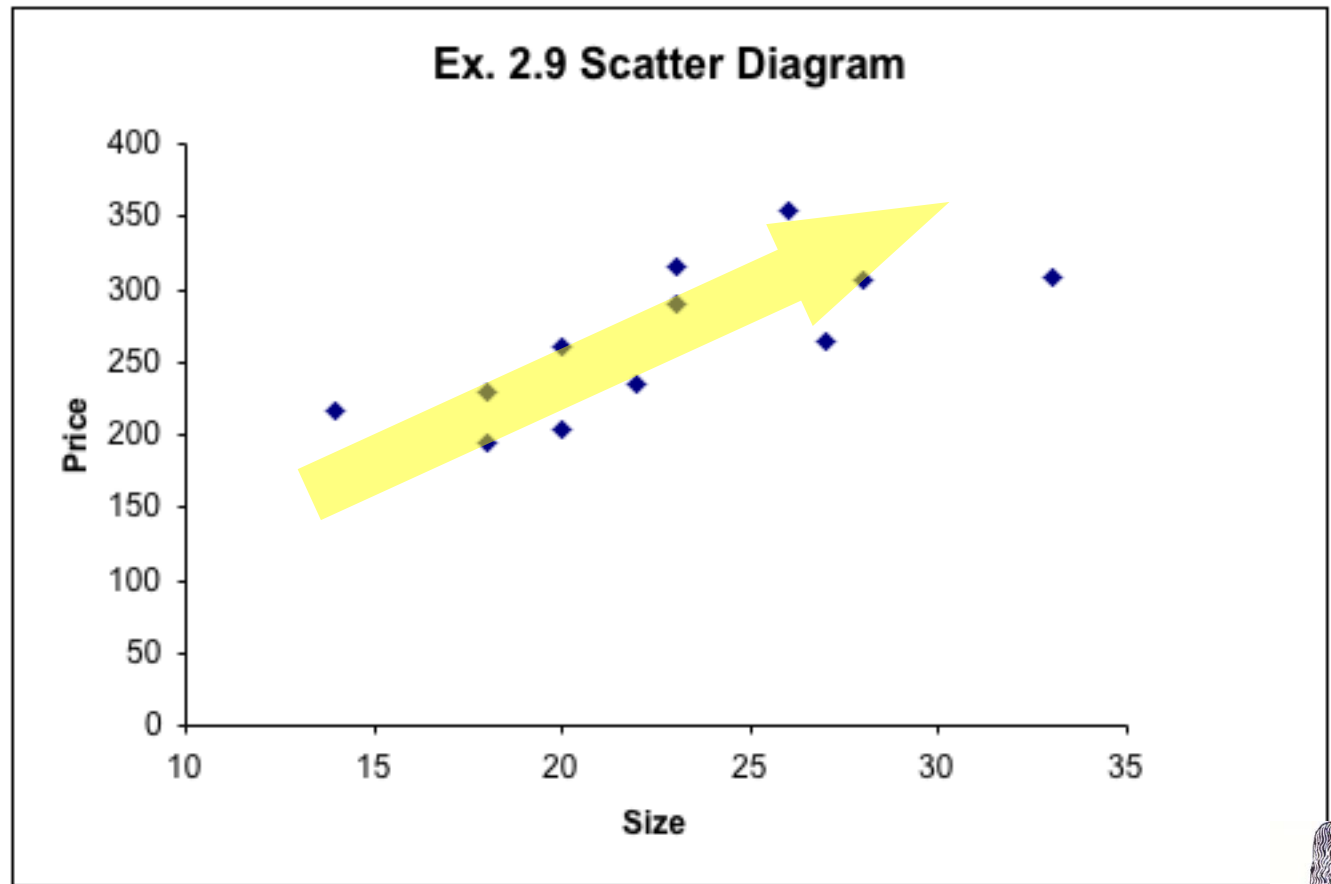
Γράφημα για τη Σχέση Μεταξύ Δύο Ποσοτικών μεταβλητών ...

- ▶ Πηγαίνοντας από ονομαστικά σε διαστημικά δεδομένα, μας ενδιαφέρει συχνά πως δύο διαστημικές μεταβλητές συσχετίζονται.
- ▶ Για να εξετάσουμε αυτή τη σχέση, κατασκευάζουμε ένα *διάγραμμα διασποράς* (scatter diagram ή scatterplot), το οποίο σχεδιάζει την μία μεταβλητή σε σχέση με την άλλη.
- ▶ Η *ανεξάρτητη* μεταβλητή συμβολίζεται με X και συνήθως τοποθετείται στον οριζόντιο άξονα, ενώ η άλλη μεταβλητή καλείται *εξαρτημένη* και παριστάνεται με Y στον κάθετο άξονα.



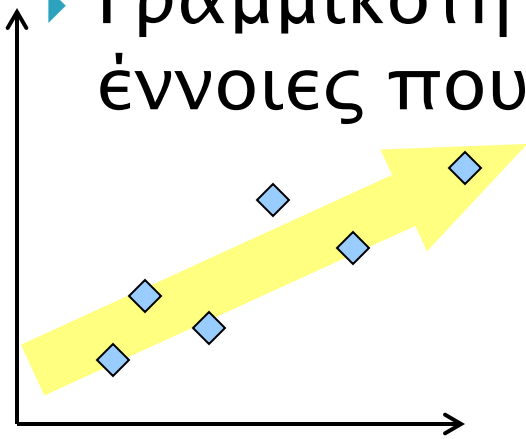
Διάγραμμα Διασποράς

- ▶ Παράδειγμα: Ένας μεσίτης σπιτιών επιθυμεί να ξέρει σε ποιο βαθμό η τιμή αγοράς συσχετίζεται με το μέγεθος του σπιτιού ...

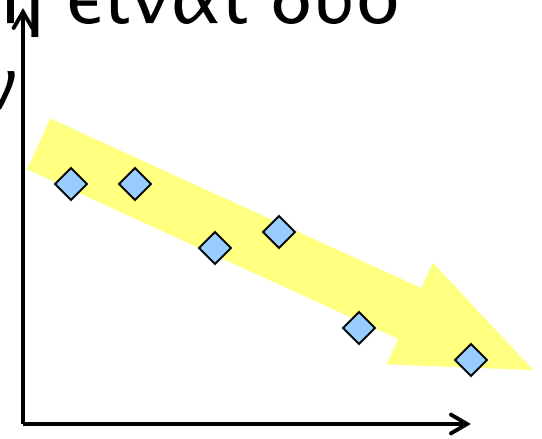


Διάγραμμα Διασποράς ...

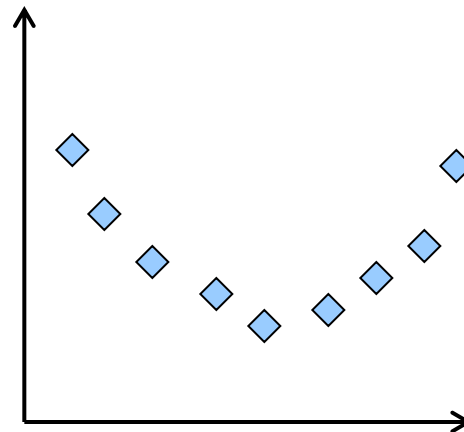
- ▶ Γραμμικότητα και Κατεύθυνση είναι δύο έννοιες που μας ενδιαφέρουν



Θετική Γραμμική Σχέση



Αρνητική Γραμμική Σχέση



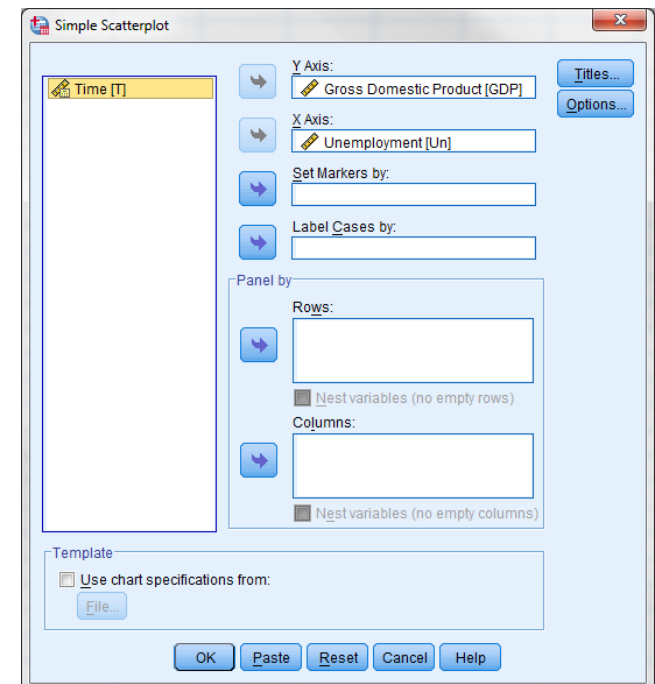
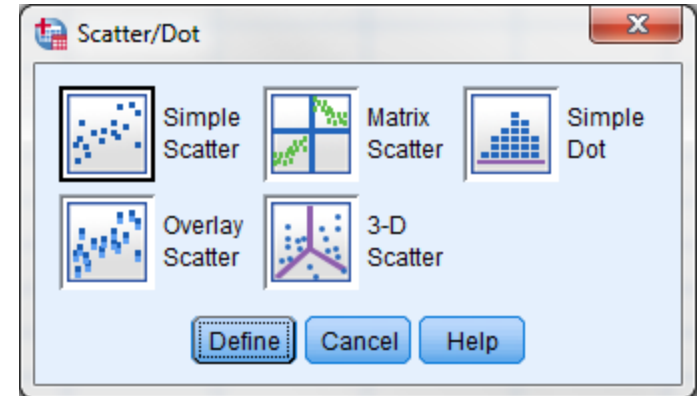
Ασθενή ή Μη-γραμμική Σχέση

Διαγράμματα διασποράς στο SPSS

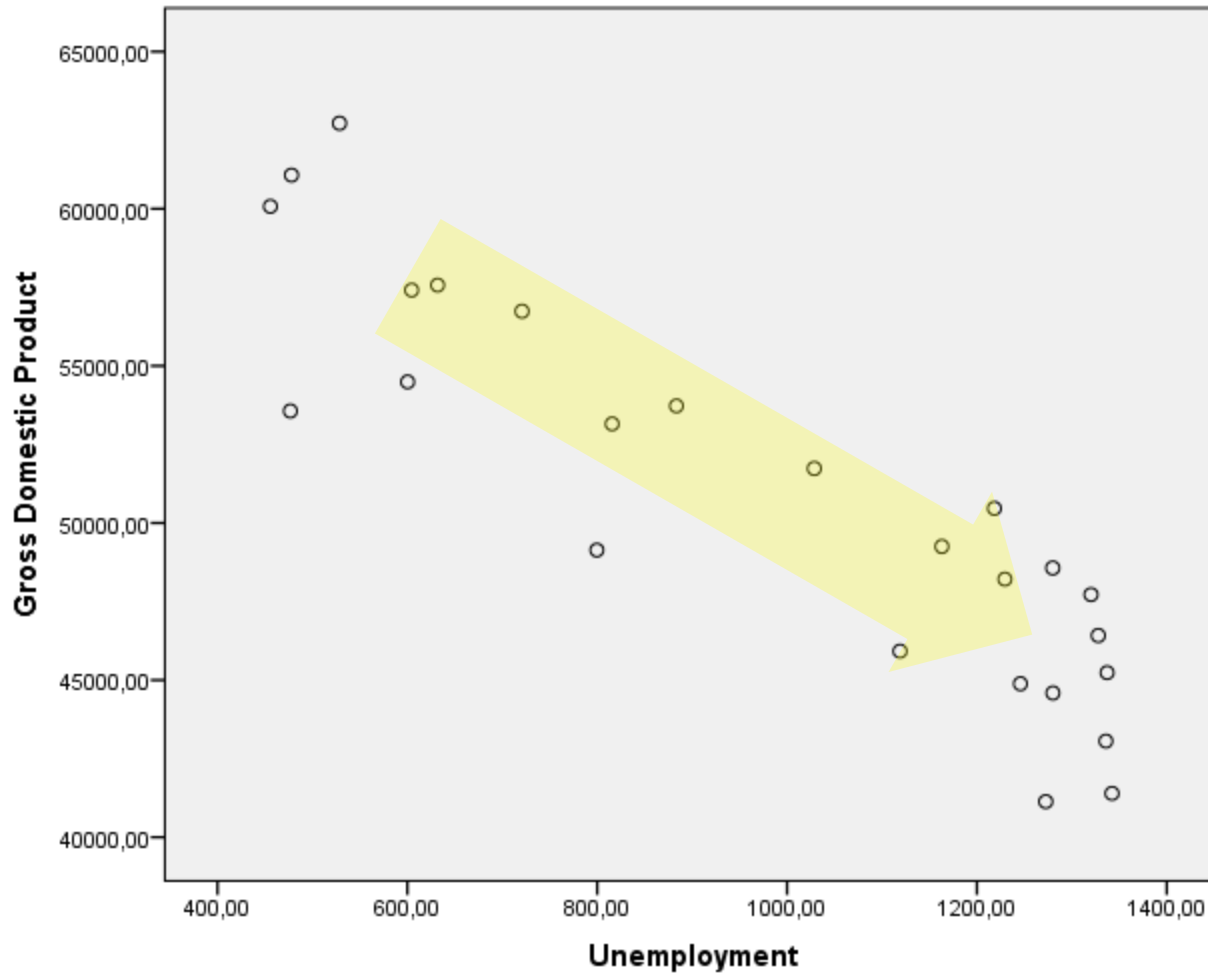
Graphs > Legacy Dialogs > Scatter/Dot

The screenshot shows the SPSS Data Editor interface. The 'Graphs' menu is open, and the 'Legacy Dialogs' sub-menu is selected. The 'Scatter/Dot...' option is highlighted. The background shows a data table with columns 'GDP', 'Un', and 'T'.

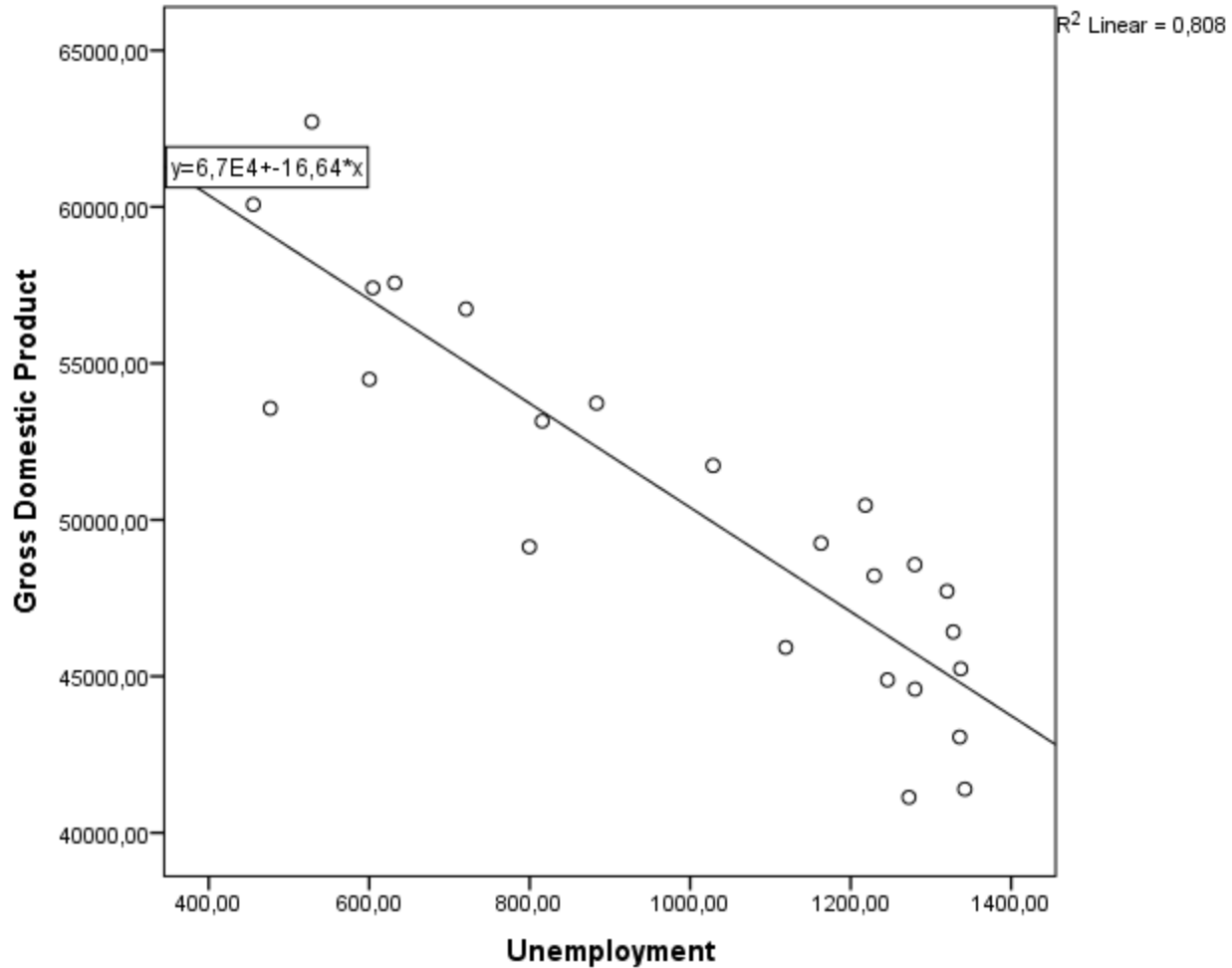
	GDP	Un	T	var
1	53563.00	476.70	1 Q 09	
2	60074.00	455.60	2 Q 09	
3	61072.00	477.90	3 Q 09	
4	62722.00	528.50	4 Q 09	
5	54490.00	600.20	1 Q 10	
6	57411.00	604.50	2 Q 10	
7	57570.00	631.80	3 Q 10	
8	56739.00	720.70	4 Q 10	
9	49139.00	799.60	1 Q 11	
10	53152.00	815.60	2 Q 11	
11	53727.00	883.40	3 Q 11	
12	51734.00	1028.60	4 Q 11	
13	45922.00	1119.00	1 Q 12	
14	49251.00	1163.00	2 Q 12	
15	50464.00	1218.30	3 Q 12	
16	48567.00	1279.90	4 Q 12	
17	43059.00	1335.90	1 Q 13	
18	46420.00	1327.80	2 Q 13	
19	47722.00	1320.20	3 Q 13	
20	45236.00	1337.10	4 Q 13	
21	41394.00	1342.30	1 Q 14	
22	44591.00	1280.10	2 Q 14	
23	48213.00	1229.37	3 Q 14	
24	44884.00	1245.85	4 Q 14	
25	41135.00	1272.54	1 Q 15	
26				
27				
28				
...				



Διαγράμματα διασποράς στο SPSS

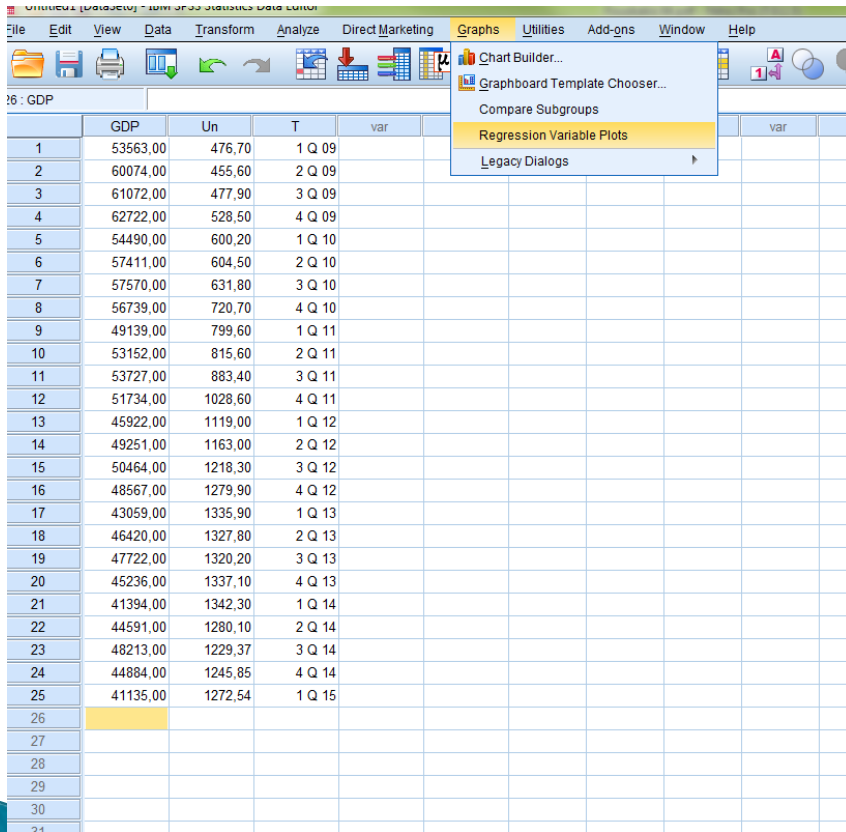


Διαγράμματα διασποράς στο SPSS

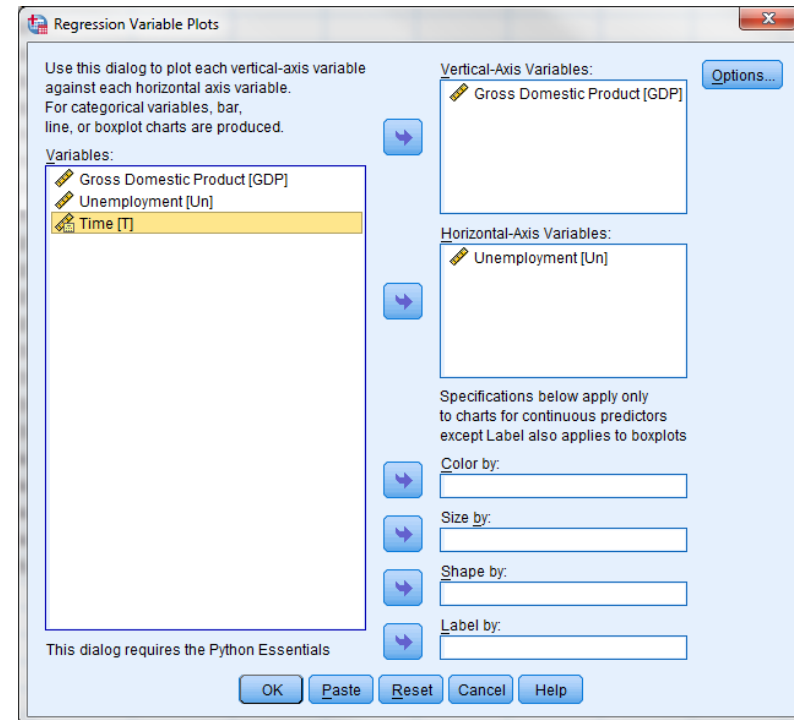


Συνδυασμός Διαγραμμάτων στο SPSS v22

Graphs > Regression Variable Plots



	GDP	Un	T	var
1	53563,00	476,70	1 Q 09	
2	60074,00	455,60	2 Q 09	
3	61072,00	477,90	3 Q 09	
4	62722,00	528,50	4 Q 09	
5	54490,00	600,20	1 Q 10	
6	57411,00	604,50	2 Q 10	
7	57570,00	631,80	3 Q 10	
8	56739,00	720,70	4 Q 10	
9	49139,00	799,60	1 Q 11	
10	53152,00	815,60	2 Q 11	
11	53727,00	883,40	3 Q 11	
12	51734,00	1028,60	4 Q 11	
13	45922,00	1119,00	1 Q 12	
14	49251,00	1163,00	2 Q 12	
15	50464,00	1218,30	3 Q 12	
16	48567,00	1279,90	4 Q 12	
17	43059,00	1335,90	1 Q 13	
18	46420,00	1327,80	2 Q 13	
19	47722,00	1320,20	3 Q 13	
20	45236,00	1337,10	4 Q 13	
21	41394,00	1342,30	1 Q 14	
22	44591,00	1280,10	2 Q 14	
23	48213,00	1229,37	3 Q 14	
24	44884,00	1245,85	4 Q 14	
25	41135,00	1272,54	1 Q 15	
26				
27				
28				
29				
30				
31				



Regression Variable Plots

Use this dialog to plot each vertical-axis variable against each horizontal axis variable. For categorical variables, bar, line, or boxplot charts are produced.

Vertical-Axis Variables:
Gross Domestic Product [GDP]

Horizontal-Axis Variables:
Unemployment [Un]

Variables:
Gross Domestic Product [GDP]
Unemployment [Un]
Time [T]

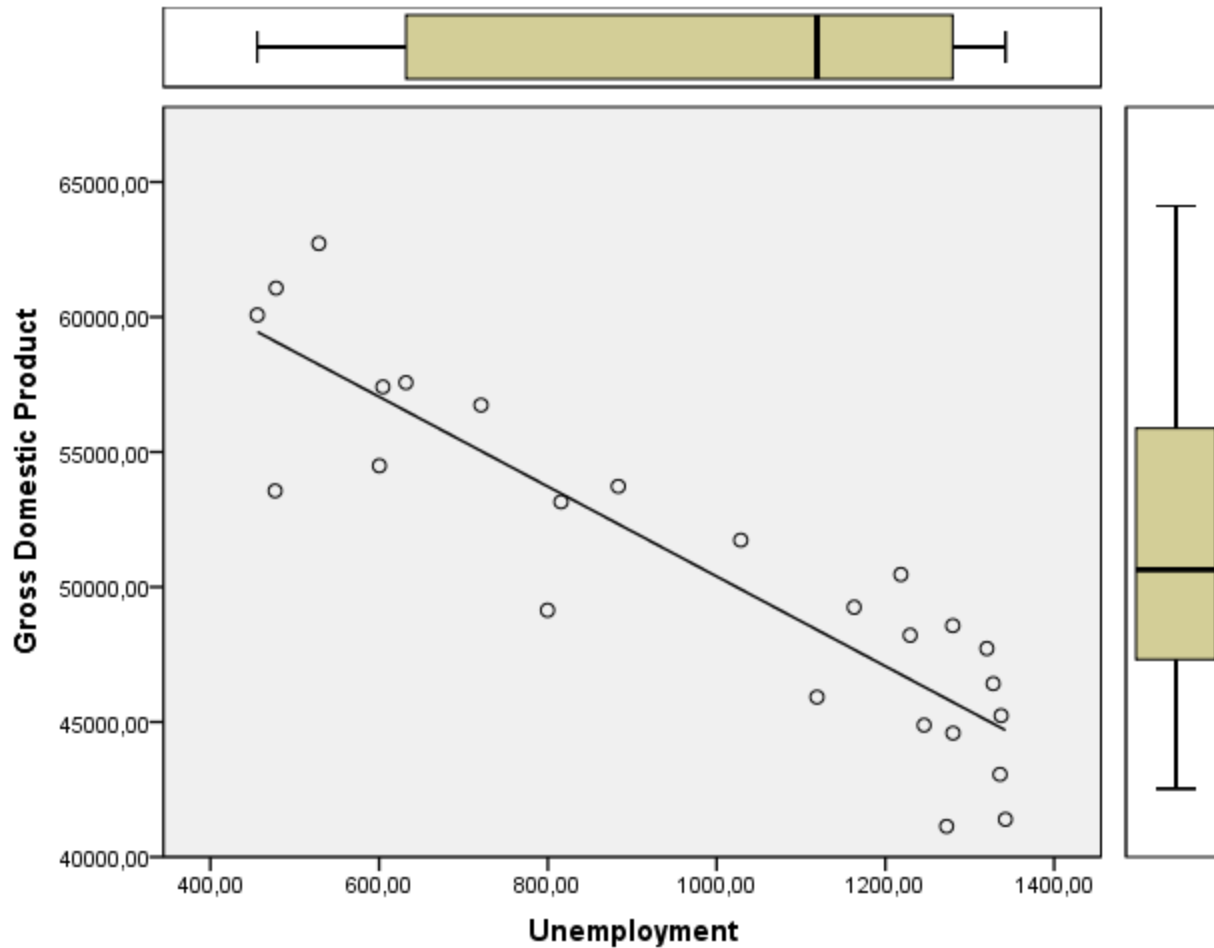
Specifications below apply only to charts for continuous predictors except Label also applies to boxplots

Color by:
Size by:
Shape by:
Label by:

This dialog requires the Python Essentials

OK Paste Reset Cancel Help

Συνδυασμός Διαγραμμάτων στο SPSS v22



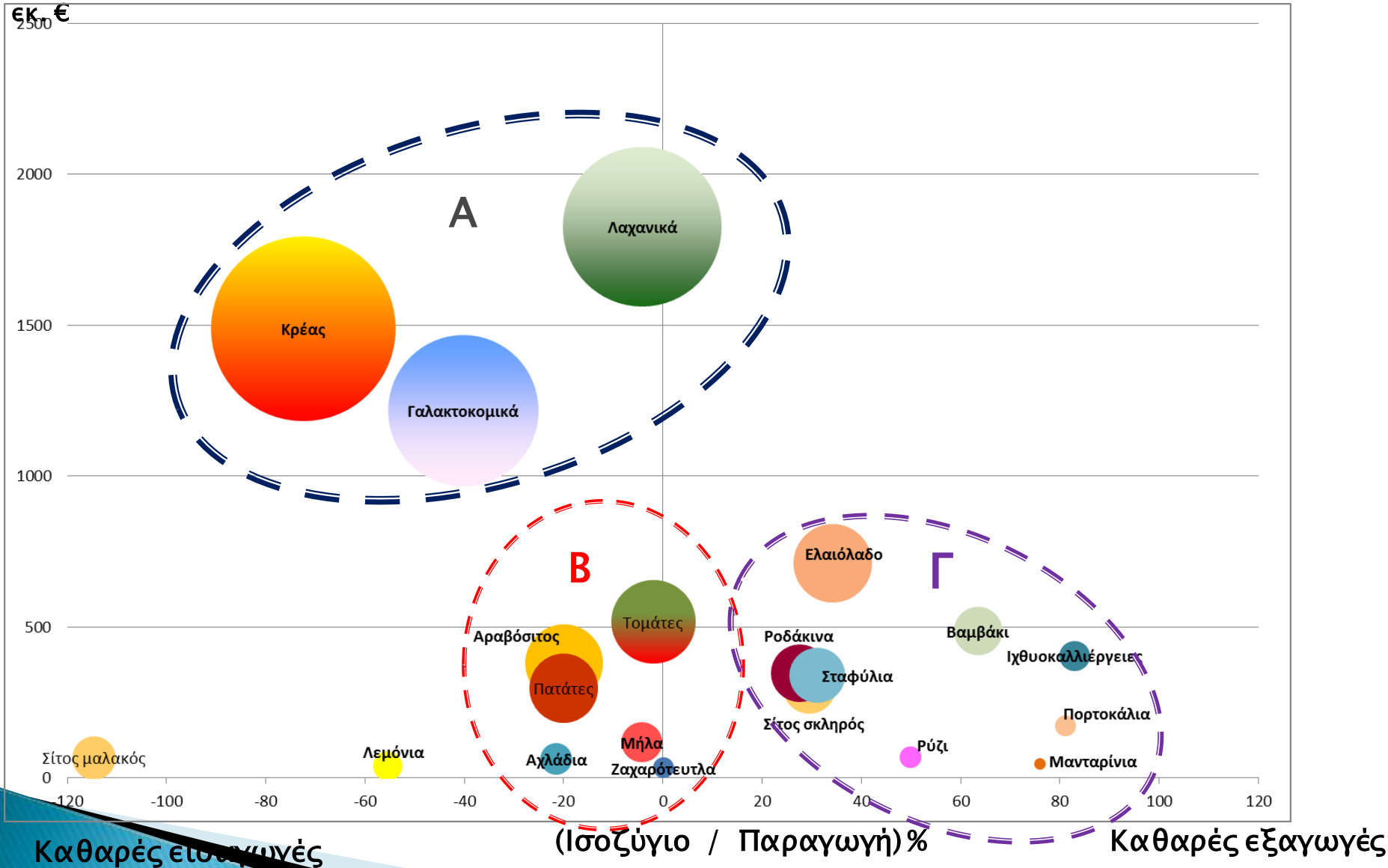
Διαγραμματική Τελειότητα

- ▶ Μεγάλα σύνολα δεδομένων παρουσιάζονται με συνέπεια και περιεκτικότητα
- ▶ Οι έννοιες και οι ιδέες που ο στατιστικός θέλει να εκφράσει γίνονται κατανοητές με σαφήνεια από τον θεατή
- ▶ Το γράφημα ενθαρρύνει το θεατή να συγκρίνει δύο ή περισσότερες μεταβλητές
- ▶ Η παρουσίαση ωθεί τον θεατή να εστιάσει την προσοχή του στην ουσία των δεδομένων και όχι στη μορφή του γραφήματος
- ▶ Δεν υπάρχει παραμόρφωση των πληροφοριών που περιέχονται στα δεδομένα

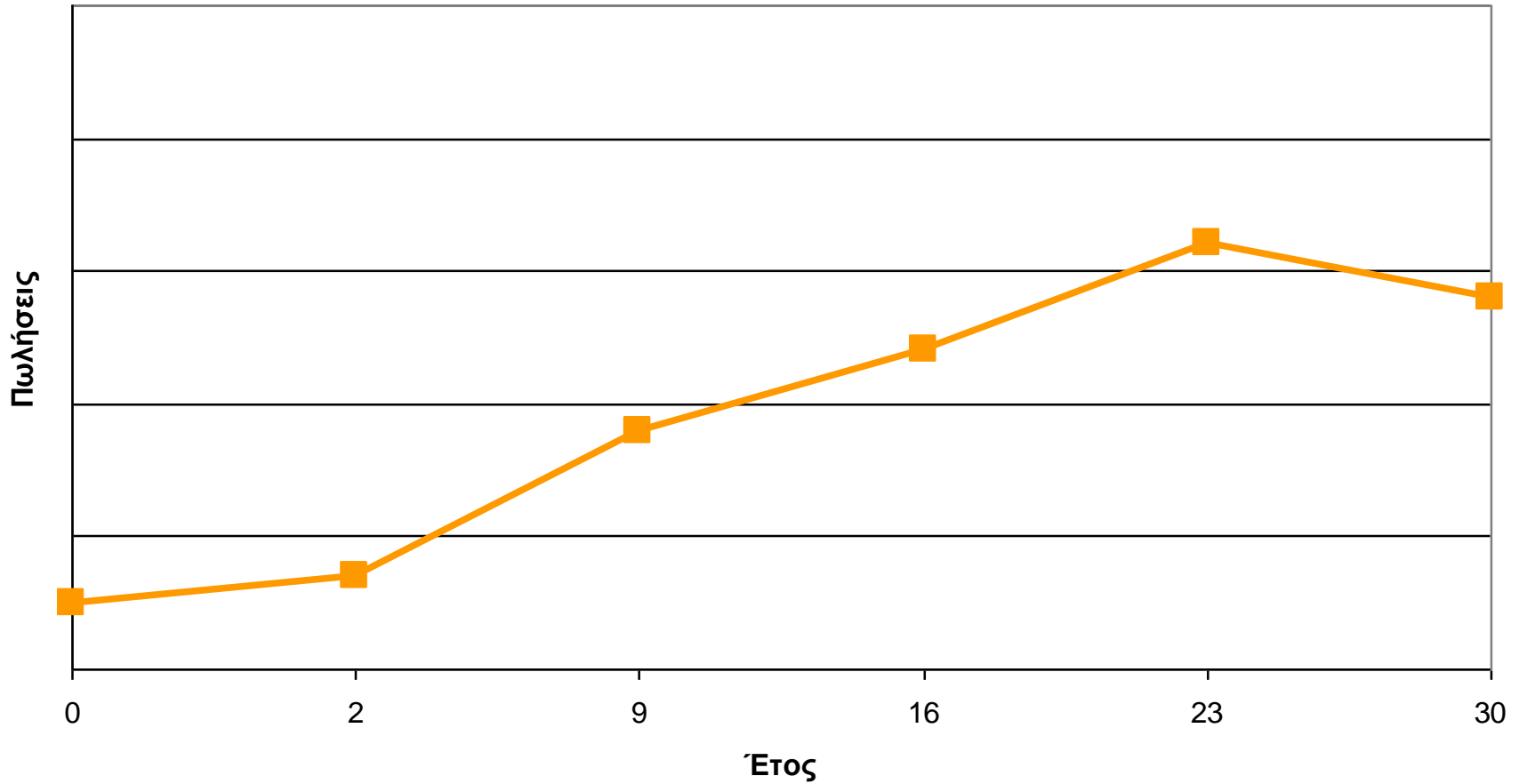
Τι παράγω

Αξία
παραγωγής σε
εκ. €

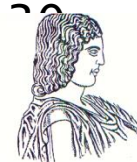
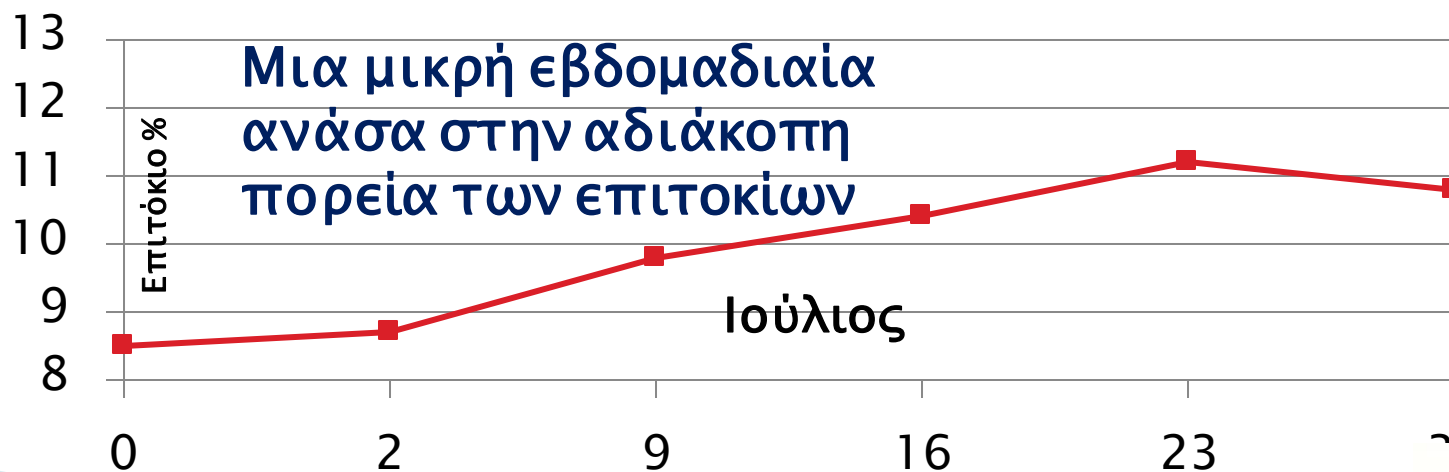
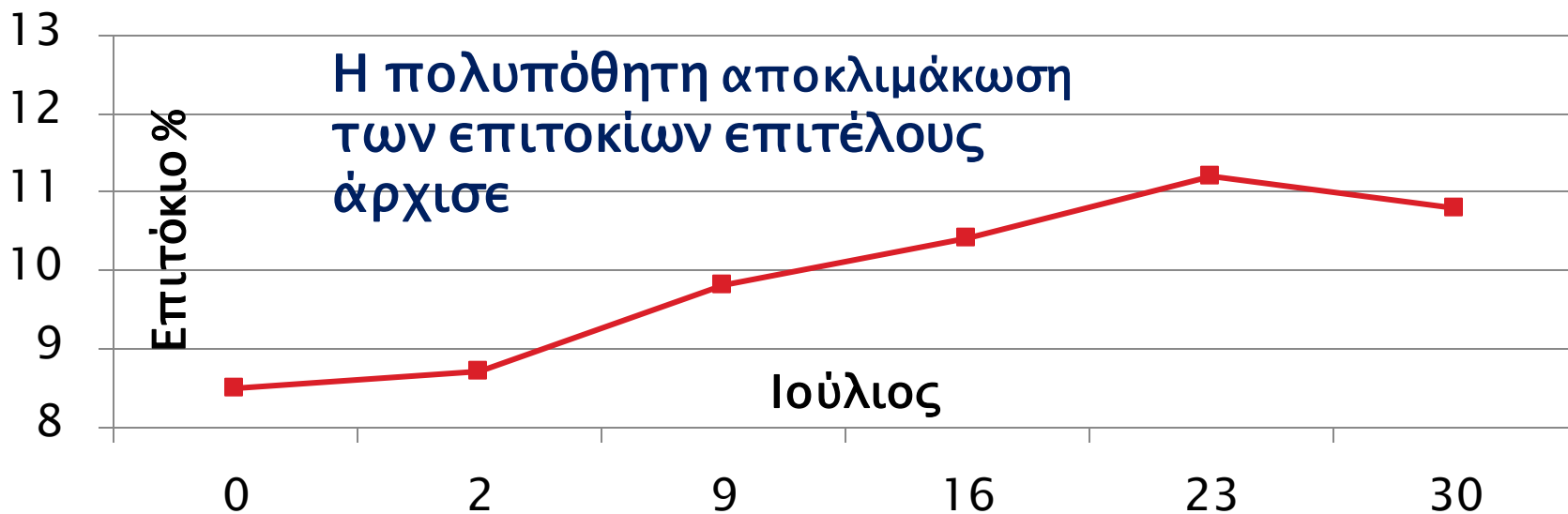
Πηγή: Eurostat, International trade



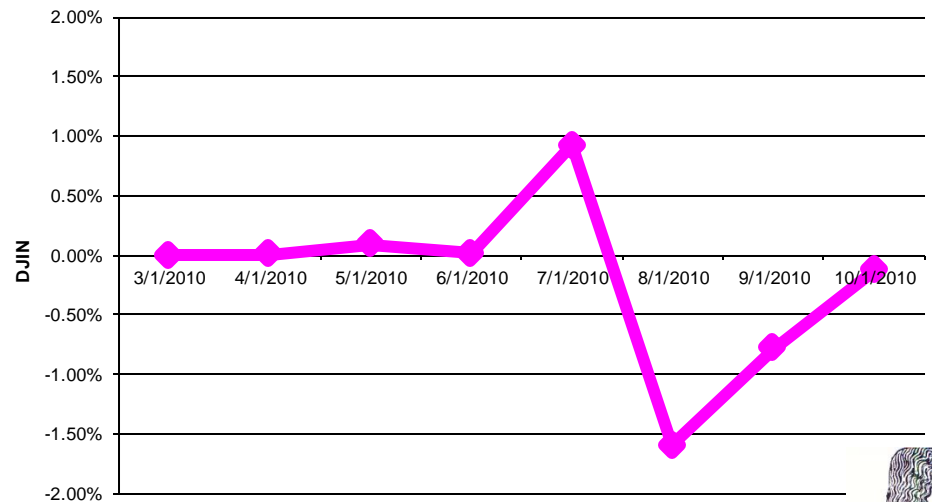
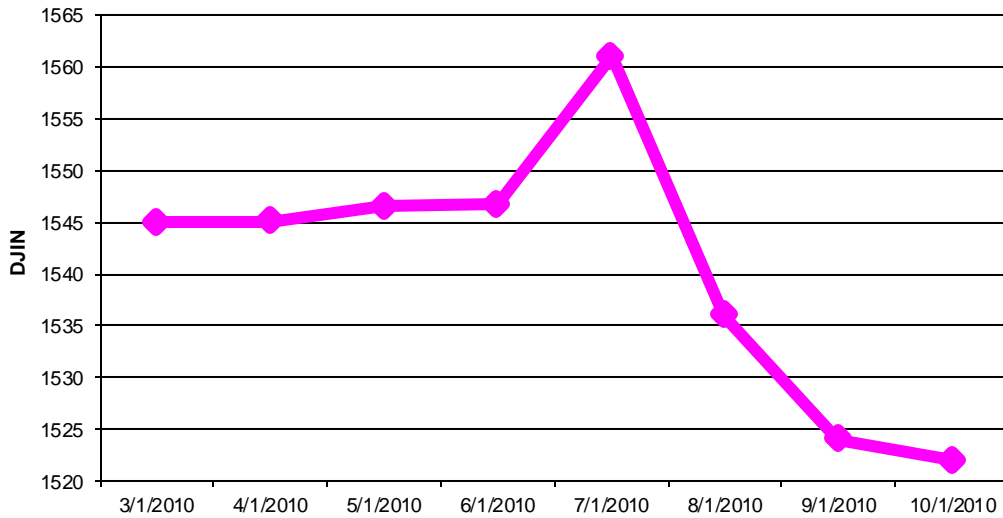
Διαγραμματική Παραπλάνηση



Διαγραμματική Παραπλάνηση



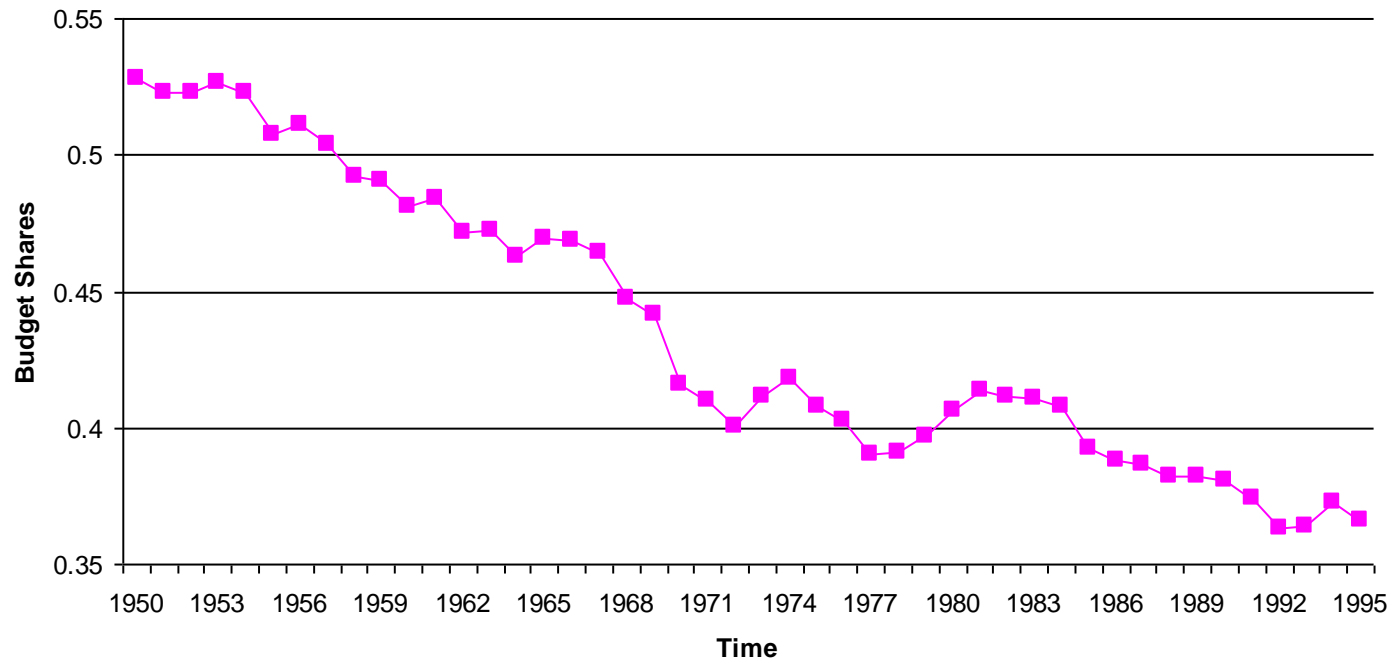
Διαγραμματική παραπλάνηση



Γραμμοδιάγραμμα ή Χρονοδιάγραμμα

Διαφορετική κλίμακα

Evolution of expenditure Shares over the period 1950-1995



—■— Food at Home



Περίληψη I...

- ▶ Παράγοντες που Προσδιορίζουν Πότε να Χρησιμοποιούμε Πίνακες Συχνοτήτων και Σχετικών Συχνοτήτων, Ραβδογράμματα και Κυκλικά Διαγράμματα
 - ▶ 1. Στόχος: Να περιγράψουμε δεδομένα μιας μεταβλητής.
 - ▶ 2. Τύπος Δεδομένων: Ονομαστικά

- ▶ Παράγοντες που Προσδιορίζουν Πότε να Χρησιμοποιούμε Ιστόγραμμα, Καμπύλη S , ή Απεικόνιση Στελεχών-Φύλλων
 - ▶ 1. Στόχος: Να περιγράψουμε δεδομένα μιας μεταβλητής.
 - ▶ 2. Τύπος Δεδομένων: Διαστημικά

- ▶ Παράγοντες που Προσδιορίζουν Πότε να Χρησιμοποιούμε Πίνακες Συνάφειας, ή Ραβδογράμματα
 - ▶ 1. Στόχος: Να περιγράψουμε την σχέση δύο μεταβλητών.
 - ▶ 2. Τύπος Δεδομένων: Ονομαστικά

- ▶ Παράγοντες που Προσδιορίζουν Πότε να Χρησιμοποιούμε ένα Διάγραμμα Διασποράς
 - ▶ 1. Στόχος: Να περιγράψουμε την σχέση δύο μεταβλητών.
 - ▶ 2. Τύπος Δεδομένων: Διαστημικά

Περίληψη II...

	Διαστημικά Δεδομένα	Ονομαστικά Δεδομένα
Δεδομένα μιας Μεταβλητής.	Ιστόγραμμα, Καμπύλη S, ή Απεικόνιση Στελεχών-Φύλλων	Πίνακες Συχνοτήτων και Σχετικών Συχνοτήτων, Ραβδογράμματα και Κυκλικά Διαγράμματα
Σχέση Μεταξύ Δύο Μεταβλητών	Διάγραμμα Διασποράς	Πίνακες Συνάφειας, ή Ραβδογράμματα