

ΑΝΑΛΥΣΗ ΧΡΟΝΟΣΕΙΡΩΝ ΚΑΙ ΔΕΔΟΜΕΝΩΝ ΠΕΡΙΒΑΛΛΟΝΤΙΚΩΝ ΚΙΝΔΥΝΩΝ

Καθ. Θεόδωρος Καρακασίδης
Δρ Αθανάσιος Φράγκου

Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών
«Βιώσιμη Διαχείριση Περιβαλλοντικών Αλλαγών και
Κυκλική Οικονομία»

Συσχέτιση

- **Συσχέτιση**: Σχέση που συνδέει δυο ποσοτικές μεταβλητές.

Παράδειγμα: το βάρος των ανθρώπων μπορεί να σχετίζεται με το δείκτη μάζας σώματος ή η θερμοκρασία σε ένα τόπο μπορεί να σχετίζεται με την υγρασία

Βασικά χαρακτηριστικά συσχέτισης

- Το είδος της συσχέτισης
- Η κατεύθυνση τα συσχέτισης
- Η ένταση της συσχέτισης.

Συσχέτιση

Τα είδη της συσχέτισης (correlation):

- Γραμμική συσχέτιση (linear correlation)

Όταν αυξάνονται οι τιμές της μίας μεταβλητής, οι τιμές της άλλης μεταβλητής αυξάνονται ή μειώνονται.

- Μη γραμμική συσχέτιση (non linear correlation)

Σχέσεις μεταξύ μεταβλητών που δεν μπορούν να περιγράψουν γραμμικά και περιγράφονται με διαφορετικά πρότυπα

Συσχέτιση

Κατεύθυνση της συσχέτισης

Ο τρόπος που **μεταβάλλονται** οι τιμές της μίας μεταβλητής όταν **μεταβάλλονται** οι τιμές της άλλης.

Παράδειγμα:

Έστω ότι τίθεται το ερώτημα -> Όταν **αυξάνεται** η βροχόπτωση, **αυξάνεται** και ο δείκτης υγρασίας στην ατμόσφαιρα ή όχι;

Απάντηση: **Θετική συσχέτιση** -> αυξάνεται ο δείκτης υγρασίας

Αρνητική συσχέτιση -> μειώνεται ο δείκτης υγρασίας

Συσχέτιση

Ένταση της συσχέτισης :

- Πόσο ισχυρή είναι η σχέση μεταξύ των μεταβλητών

Ισχυρή συσχέτιση -> όταν οι τιμές της μίας μεταβλητής σχετίζονται σε μεγάλο βαθμό με τις τιμές της άλλης.

Μεγάλη ή μικρή μεταβολή της μίας μεταβλητής -> μεγάλη ή μικρή μεταβολή της άλλης.

Προσοχή: Ο συντελεστής συσχέτισης δεν αποκαλύπτει σχέση αιτίου - αιτιατού

Συντελεστές Συσχέτισης

Συντελεστής γραμμικής συσχέτισης του Pearson

$$r_{X,Y} = \frac{S_{X,Y}}{\sqrt{S_X^2 S_Y^2}} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

- Παραμετρικός συντελεστής (δεδομένα ακολουθούν κανονική κατανομή)

Ανεξάρτητος των μονάδων μέτρησης των μεταβλητών

- Η **τιμή** του βρίσκεται μεταξύ του 1 και του -1, δηλαδή $-1 \leq r \leq 1$
 1. Εάν $r = 0$, -> οι δυο μεταβλητές είναι γραμμικά **ασυσχέτιστες**.
 2. Εάν $r = 1$ -> πλήρης **θετική συσχέτιση** – ομόρροπη μεταβολή
 3. Εάν $r = -1$ -> πλήρης **αρνητική συσχέτιση** – αντίρροπη μεταβολή

Συντελεστές Συσχέτισης

Στατιστική σημαντικότητα του συντελεστή συσχέτισης του Pearson

- Υπολογίζουμε την τιμή του συντελεστή και αξιολογούμε τη στατιστική σημαντικότητά του με έλεγχο υποθέσεων (θεωρητικός συντελεστής συσχέτισης ρ)

α) $H_0: \rho \leq 0$, $H_1: \rho > 0$ με κρίσιμη περιοχή $t \geq t_{n-2, \alpha}$

β) $H_0: \rho \geq 0$, $H_1: \rho < 0$ με κρίσιμη περιοχή $t \leq -t_{n-2, \alpha}$

γ) $H_0: \rho = 0$, $H_1: \rho \neq 0$ με κρίσιμη περιοχή $|t| \geq t_{n-2, \alpha/2}$

Με στατιστική συνάρτηση $t = r \sqrt{\frac{n-2}{1-r^2}}$

Συντελεστές Συσχέτισης

Προσοχή: Για την εφαρμογή του συντελεστή συσχέτισης του Pearson τα δεδομένα μας πρέπει να ακολουθούν την τυπική κανονική κατανομή.

Για να κανονικοποιήσουμε τα δεδομένα κατά την τυπική κανονική κατανομή $N(0,1)$ αφαιρούμε κάθε μέτρηση ξεχωριστά από τη μέση τιμή και διαιρούμε με την τυπική απόκλιση

$$Y = \frac{x_i - \bar{X}}{\sigma}$$

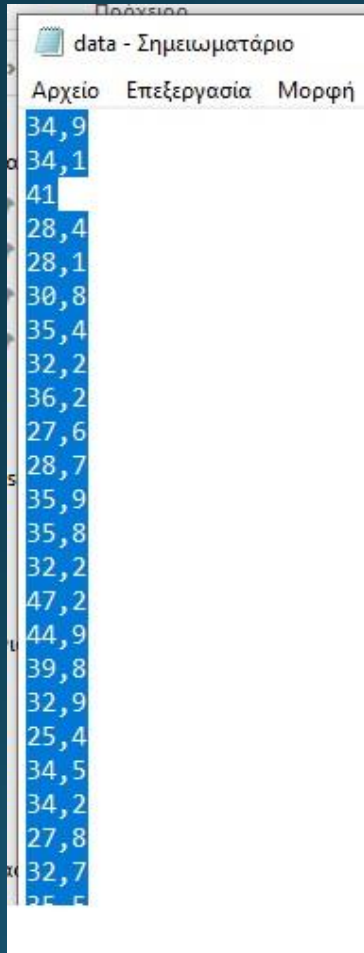
Σημείωση μεθοδολογίας: Κανονικοποιούμε κατά την τυπική κανονική κατανομή τα δεδομένα, προκειμένου να μπορούν να «συγκριθούν». Με την κανονικοποίηση οι τιμές των δεδομένων κυμαίνονται σε παρόμοια διαστήματα

Συντελεστές Συσχέτισης Εφαρμογή

- Διαιτολόγος θέλει να δημιουργήσει μια νέα δίαιτα και μελετά τη σχέση που έχουν μεταξύ τους το ύψος (HG), ο λιπώδης ιστός (FMSS) και ο δείκτης μάζας σώματος (BMI). Από δείγμα 50 ατόμων κατέγραψε αυτές τις ποσότητες. Να μελετηθεί η συσχέτιση με τη χρήση του SPSS.

Συντελεστές Συσχέτισης (SPSS)

- Κανονικοποίηση των δεδομένων με αντιγραφή από το αρχείο που τα περιέχει και εισαγωγή τους στο Excel



data - Σημειωματάριο

Αρχείο	Επεξεργασία	Μορφή
34,9		
34,1		
41		
28,4		
28,1		
30,8		
35,4		
32,2		
36,2		
27,6		
28,7		
35,9		
35,8		
32,2		
47,2		
44,9		
39,8		
32,9		
25,4		
34,5		
34,2		
27,8		
32,7		
25,5		



	A	B	C
1	ΑΡΧΙΚΑ ΔΕΔΟΜΕΝΑ		
2	ΔΕΔΟΜΕΝΑ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ
3	34,9	34,442	5,706070101
4	34,1		
5	41		
6	28,4		
7	28,1		
8	30,8		
9	35,4		
10	32,2		
11	36,2		
12	27,6		
13	28,7		
14	35,9		
15	35,8		
16	32,2		
17	47,2		
18	44,9		
19	39,8		
20	32,9		
21	25,4		
22	34,5		
23	34,2		
24	27,8		
25	32,7		

άμεσα
υπολογισμός
Μέσης τιμής
και τυπ.
Απόκλισης
των
πραγματικών
τιμών των
δεδομένων

Συντελεστές Συσχέτισης (SPSS)

- Αυτόματος υπολογισμός κανονικοποιημένων δεδομένων

	A	B	C
1	ΑΡΧΙΚΑ ΔΕΔΟΜΕΝΑ		
2	ΔΕΔΟΜΕΝΑ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ
3	34,9	34,442	5,706070101
4	34,1		
5	41		
6	28,4		
7	28,1		
8	30,8		
9	35,4		
10	32,2		
11	36,2		
12	27,6		
13	28,7		
14	35,9		
15	35,8		
16	32,2		
17	47,2		
18	44,9		
19	39,8		
20	32,9		
21	25,4		
22	34,5		
23	34,2		
24	27,8		
25	32,7		



D	E	F	G
ΚΑΝΟΝΙΚΟΠΟΙΗΜΕΝΑ ΔΕΔΟΜΕΝΑ			
ΚΑΝ_ΔΕΔΟΜ		ΜΕΣΗ ΤΙΜΗ	ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ
0,0802654		1,23346E-15	1
-0,059936172			
1,149302389			
-1,058872375			
-1,111447965			
-0,638267658			
0,167891383			
-0,392914906			
0,308092955			
-1,199073947			
-1,006296785			
0,255517366			
0,237992169			
-0,392914906			
2,235864575			
1,832785054			
0,939000031			
-0,270238531			
-1,584628271			
0,010164614			
-0,042410976			

Υπολογισμός
Μέσης τιμής
(ίσης με
μηδέν) και
τυπ.
Απόκλισης
(ίσης με
μονάδα)

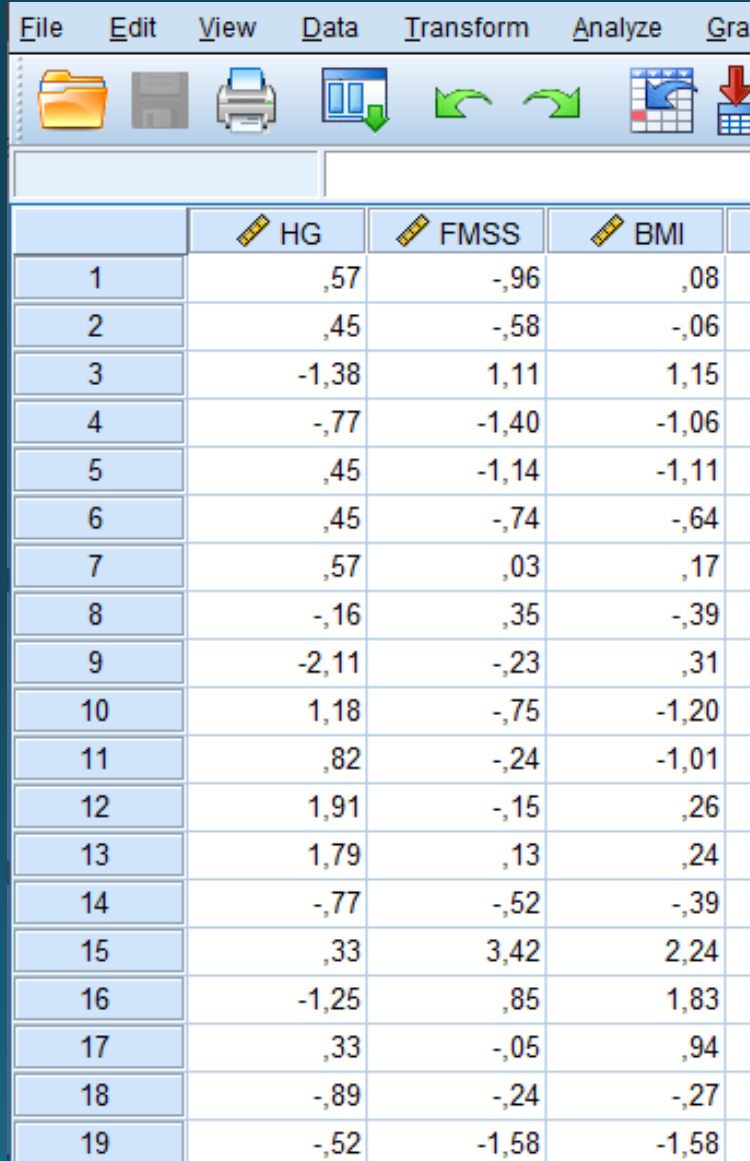
Συντελεστές Συσχέτισης (SPSS)

- Το τελικό φύλλο Excel με την κανονικοποίηση των δεδομένων

	A	B	C	D	E	F	G
1	ΑΡΧΙΚΑ ΔΕΔΟΜΕΝΑ			ΚΑΝΟΝΙΚΟΠΟΙΗΜΕΝΑ ΔΕΔΟΜΕΝΑ			
2	ΔΕΔΟΜΕΝΑ	ΜΕΣΗ ΤΙΜΉ	ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ	ΚΑΝ_ΔΕΔΟΜ		ΜΕΣΗ ΤΙΜΉ	ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ
3	34,9	34,442	5,706070101	0,0802654		1,23346E-15	1
4	34,1			-0,059936172			
5	41			1,149302389			
6	28,4			-1,058872375			
7	28,1			-1,111447965			
8	30,8			-0,638267658			
9	35,4			0,167891383			
10	32,2			-0,392914906			
11	36,2			0,308092955			
12	27,6			-1,199073947			
13	28,7			-1,006296785			
14	35,9			0,255517366			
15	35,8			0,237992169			
16	32,2			-0,392914906			
17	47,2			2,235864575			
18	44,9			1,832785054			
19	39,8			0,939000031			
20	32,9			-0,270238531			
21	25,4			-1,584628271			
22	34,5			0,010164614			
23	34,2			-0,042410976			

Συντελεστές Συσχέτισης (SPSS)

- Κατασκευή νέου αρχείου με τις μεταβλητές medical_norm.sav



The image shows a screenshot of the SPSS software interface. The menu bar includes File, Edit, View, Data, Transform, Analyze, and Graph. Below the menu bar is a toolbar with icons for file operations (folder, save, print, save as), navigation (undo, redo), and data management (refresh, download). The main window displays a data table with 19 rows and 4 columns. The first column contains row numbers from 1 to 19. The second column is labeled 'HG', the third 'FMSS', and the fourth 'BMI'. Each cell in the data table contains a numerical value.

	HG	FMSS	BMI
1	,57	-,96	,08
2	,45	-,58	-,06
3	-1,38	1,11	1,15
4	-,77	-1,40	-1,06
5	,45	-1,14	-1,11
6	,45	-,74	-,64
7	,57	,03	,17
8	-,16	,35	-,39
9	-2,11	-,23	,31
10	1,18	-,75	-1,20
11	,82	-,24	-1,01
12	1,91	-,15	,26
13	1,79	,13	,24
14	-,77	-,52	-,39
15	,33	3,42	2,24
16	-1,25	,85	1,83
17	,33	-,05	,94
18	-,89	-,24	-,27
19	-,52	-1,58	-1,58

Συντελεστές Συσχέτισης (SPSS)

- επιλέγουμε τις μεταβλητές

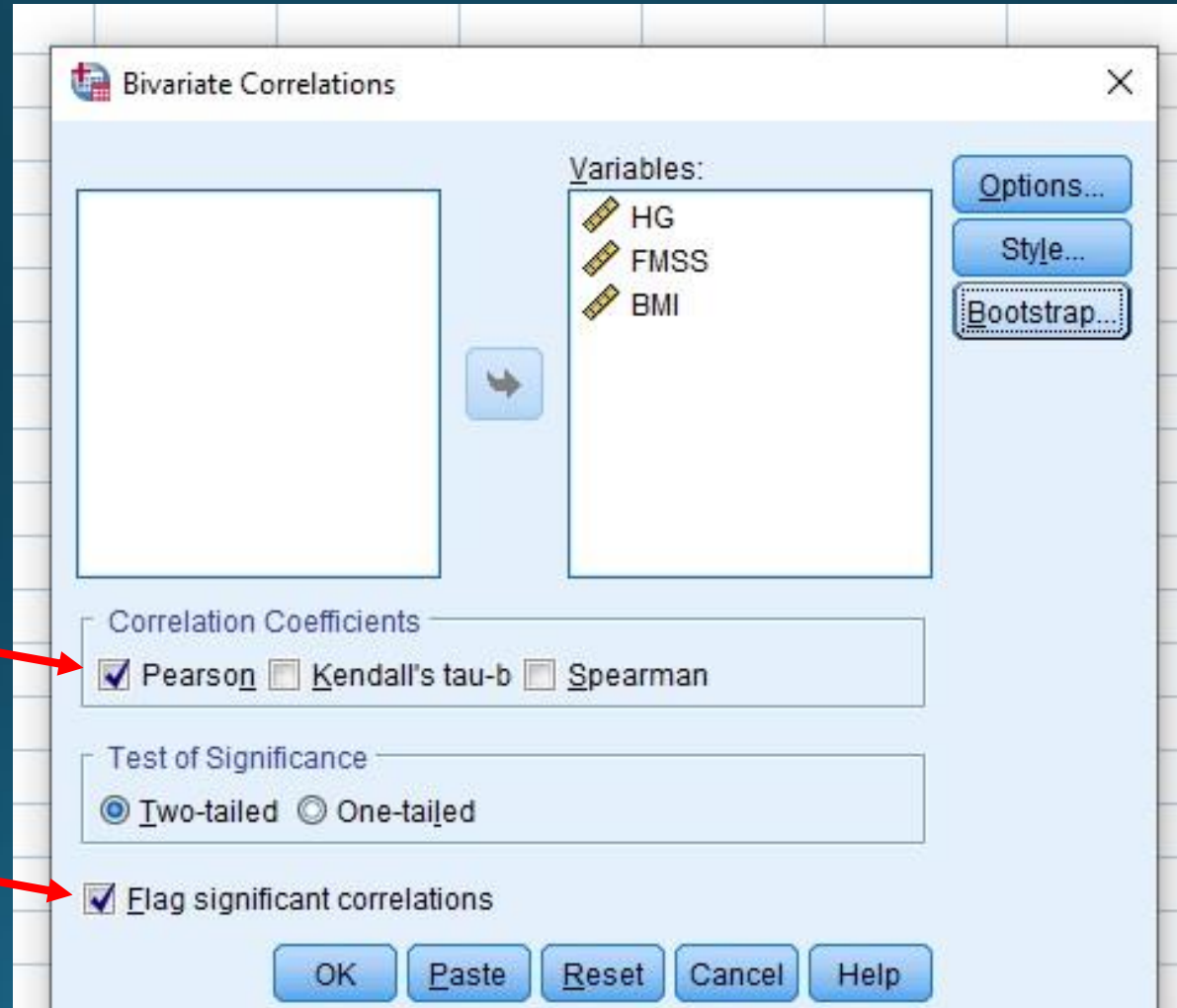
The screenshot shows the SPSS Statistics Data Editor interface. The 'Analyze' menu is open, and the 'Correlate' option is selected, which has opened a sub-menu with 'Bivariate...' highlighted. The data table below shows the following values:

	HG	FMSS
1	,57	-,96
2	,45	-,58
3	-1,38	1,11
4	-,77	-1,40
5	,45	-1,14
6	,45	-,74
7	,57	,03
8	-,16	,35
9	-2,11	-,23
10	1,18	-,75
11	,82	-,24
12	1,91	-,15
13	1,79	,13
14	-,77	-,52
15	,33	3,42
16	-1,25	,85
17	,33	-,05
18	-,89	-,24

- Επιλογή: **Analyze - > Correlate -> Bivariate**

Συντελεστές Συσχέτισης (SPSS)

- επιλέγουμε τις μεταβλητές



συντελεστής



σημαντικότητα



Συντελεστές Συσχέτισης (SPSS)

συντελεστές
συσχέτισης
Pearson

		HG	FMSS	BMI
HG	Pearson Correlation	1	,172	,191
	Sig. (2-tailed)		,231	,185
	N	50	50	50
FMSS	Pearson Correlation	,172	1	,803**
	Sig. (2-tailed)	,231		,000
	N	50	50	50
BMI	Pearson Correlation	,191	,803**	1
	Sig. (2-tailed)	,185	,000	
	N	50	50	50

** . Correlation is significant at the 0.01 level (2-tailed).

στάθμη
σημαντικότητας

- Σημαντικός συντελεστής πρέπει $p\text{-value} < \alpha$

Εδώ $\alpha = 0.01$, άρα FMSS – BMI με $r = 0.803$

- **Ισχυρή θετική συσχέτιση** -> θετική (αρνητική) μεταβολή του λιπώδους ιστού επιφέρει ισχυρά θετική (αρνητική) μεταβολή του δείκτη μάζας σώματος

Συντελεστές Συσχέτισης (SPSS)

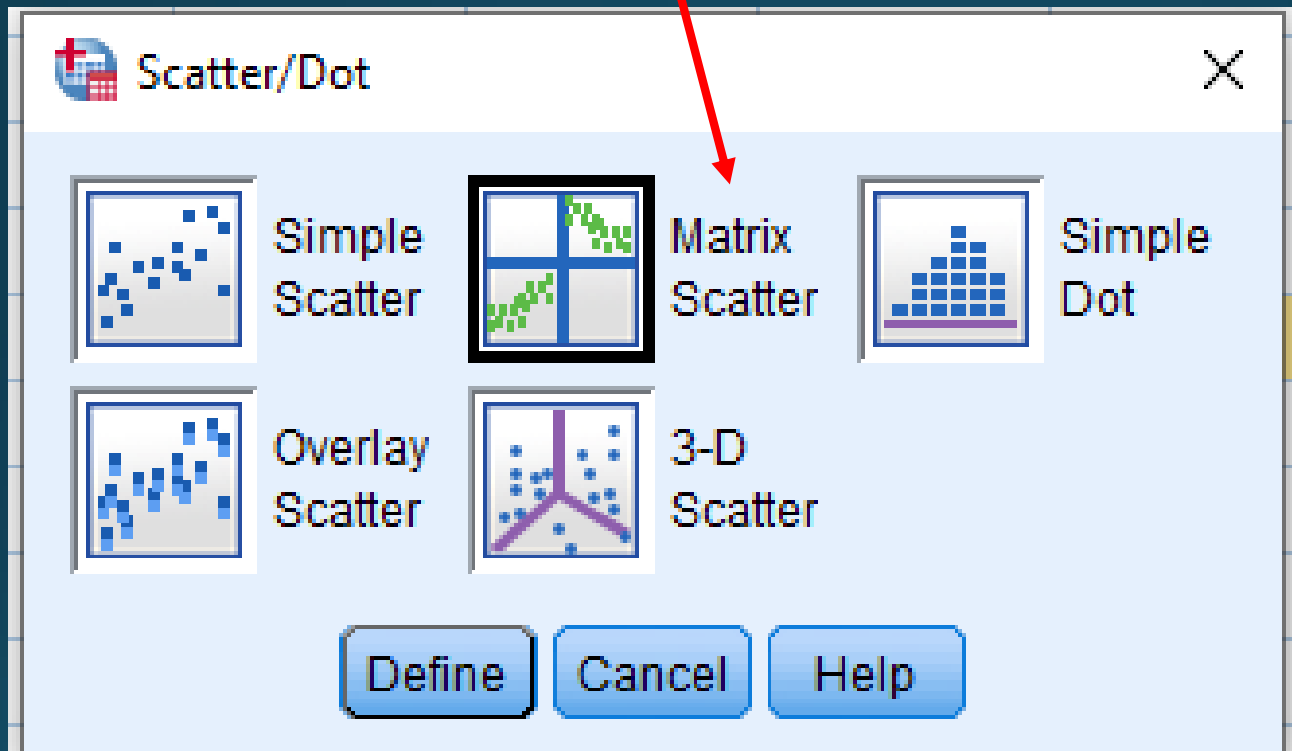
- Οπτική εικόνα γραμμικής συσχέτισης με **γραφήματα διασποράς**
- Όσο πιο **διάσπαρτες** είναι οι τιμές μεταξύ τους, τόσο πιο ασυσχέτιστες είναι οι μεταβλητές

The screenshot shows the IBM SPSS Statistics Data Editor interface. The main window displays a data table with columns labeled HG, FMSS, and BMI. The 'Graphs' menu is open, and the 'Scatter/Dot...' option is highlighted with a red arrow. The 'Legacy Dialogs' submenu is also visible, showing various chart options.

	HG	FMSS	BMI
1	,57	-,96	,0
2	,45	-,58	-,0
3	-1,38	1,11	1,
4	-,77	-1,40	-1,06
5	,45	-1,14	-1,11
6	,45	-,74	-,64
7	,57	,03	,17
8	-,16	,35	-,39
9	-2,11	-,23	,31
10	1,18	-,75	-1,20
11	,82	-,24	-1,01
12	1,91	-,15	,26
13	1,79	,13	,24
14	-,77	-,52	-,39
15	,33	3,42	2,24
16	-1,25	,85	1,83
17	,33	-,05	,94
18	-,89	-,24	-,27

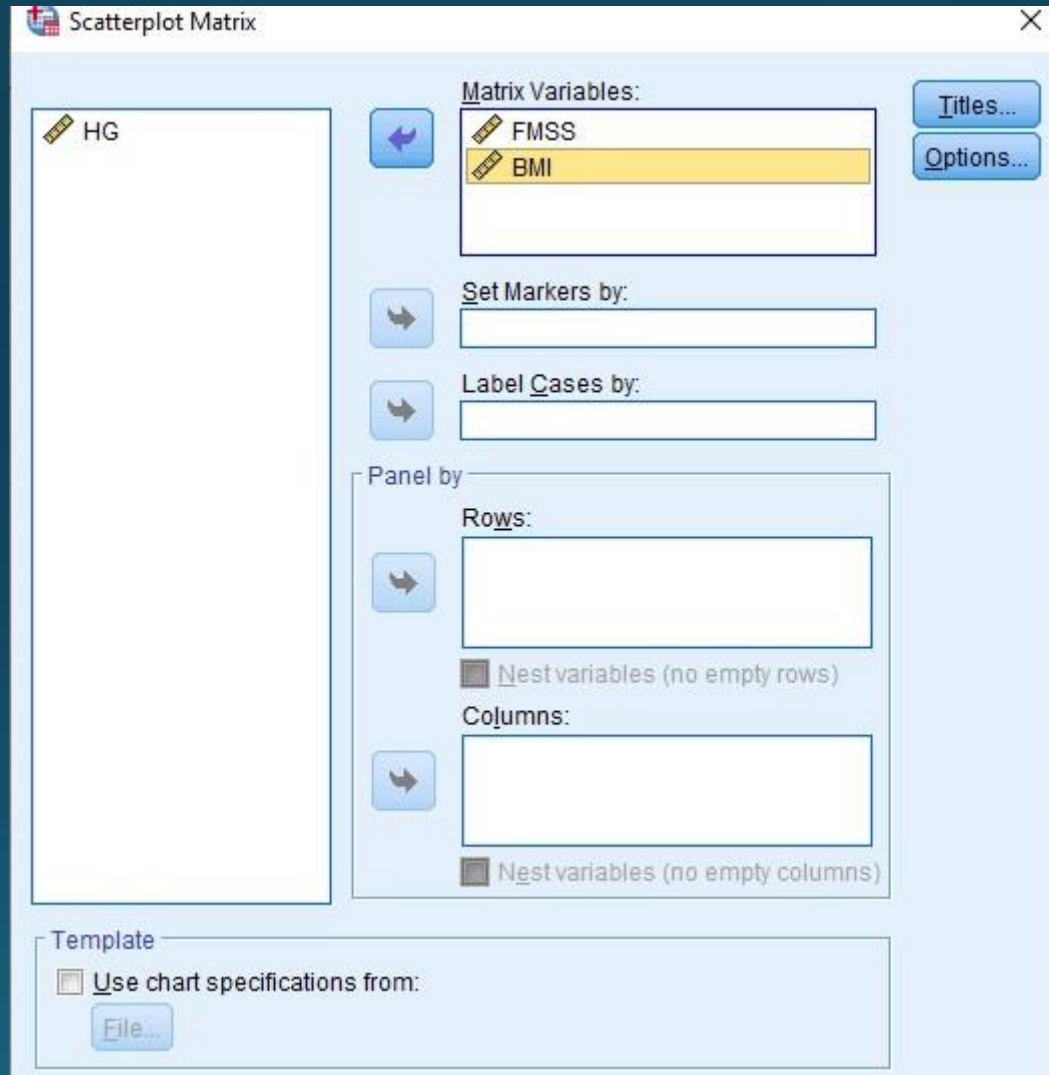
Συντελεστές Συσχέτισης (SPSS)

- Επιλογή: Matrix Scatter - Define



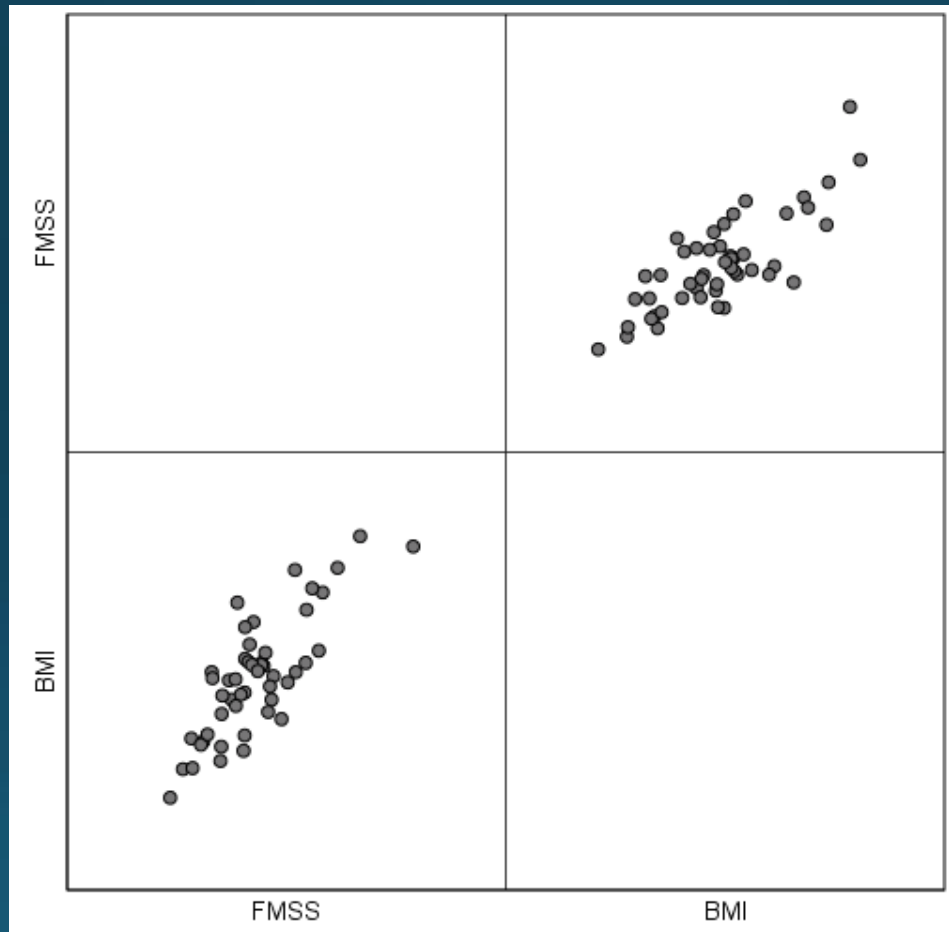
Συντελεστές Συσχέτισης (SPSS)

- Επιλογή των FMSS - BMI



Συντελεστές Συσχέτισης (SPSS)

- Οι τιμές τους **δεν** είναι διασκορπισμένες - > συσχετισμένες



- Να γίνει το διάγραμμα διασποράς για τις HG- BMI και HG - FMSS