



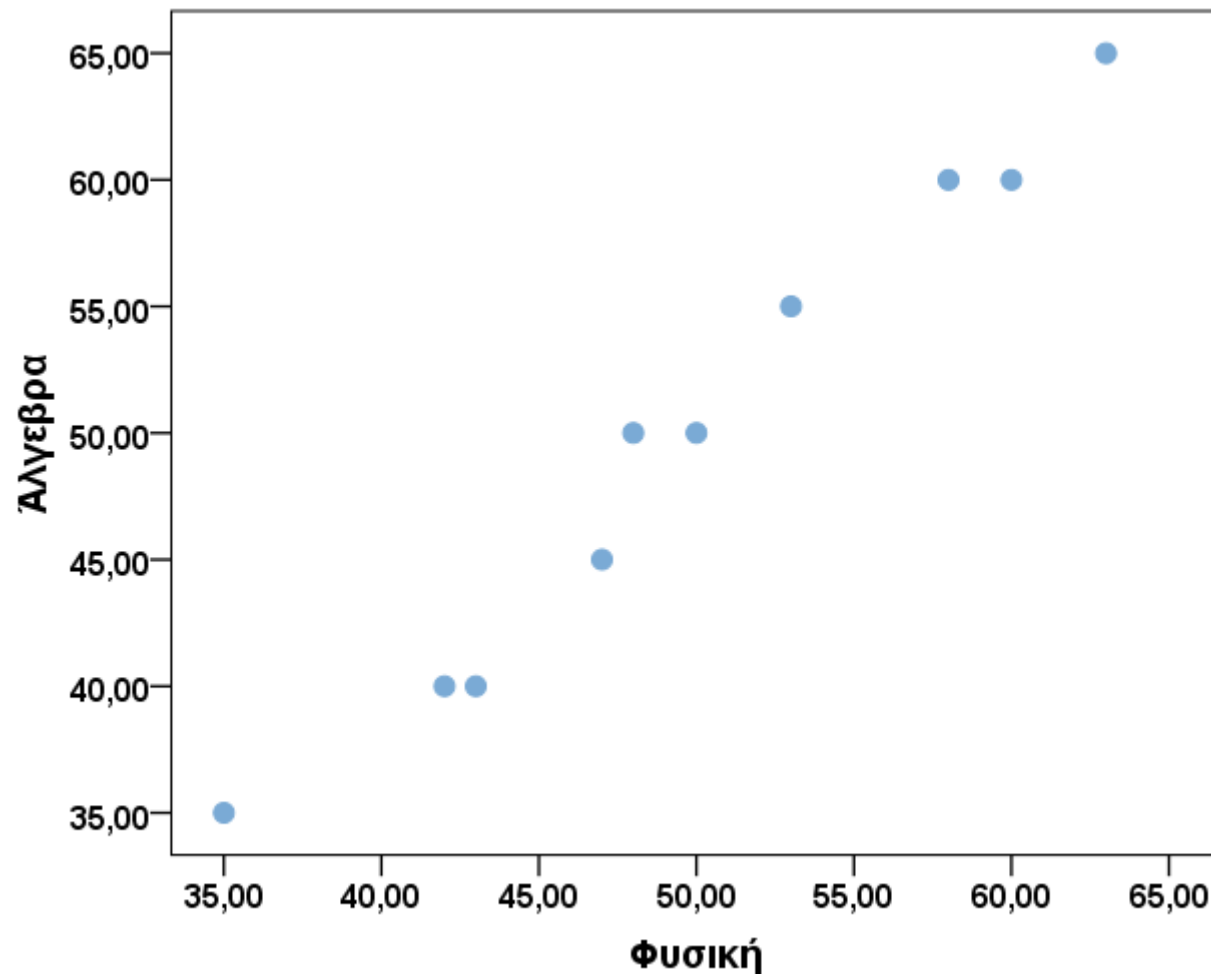
**Παιδαγωγικό Τμήμα Δημοτικής Εκπαίδευσης
Δημοκρίτειο Πανεπιστήμιο Θράκης
Αλεξανδρούπολη**

Οι επιδόσεις δέκα μαθητών σε τέσσερα μαθήματα

Μαθητής	Άλγεβρα	Φυσική	Νέα Ελληνικά	Μουσική
A	65	63	35	61
B	60	58	38	35
Γ	60	60	40	46
Δ	55	53	42	40
E	50	48	53	50
ΣΤ	50	50	49	60
Z	45	47	54	60
H	40	42	62	40
Θ	40	43	60	55
I	35	35	63	41

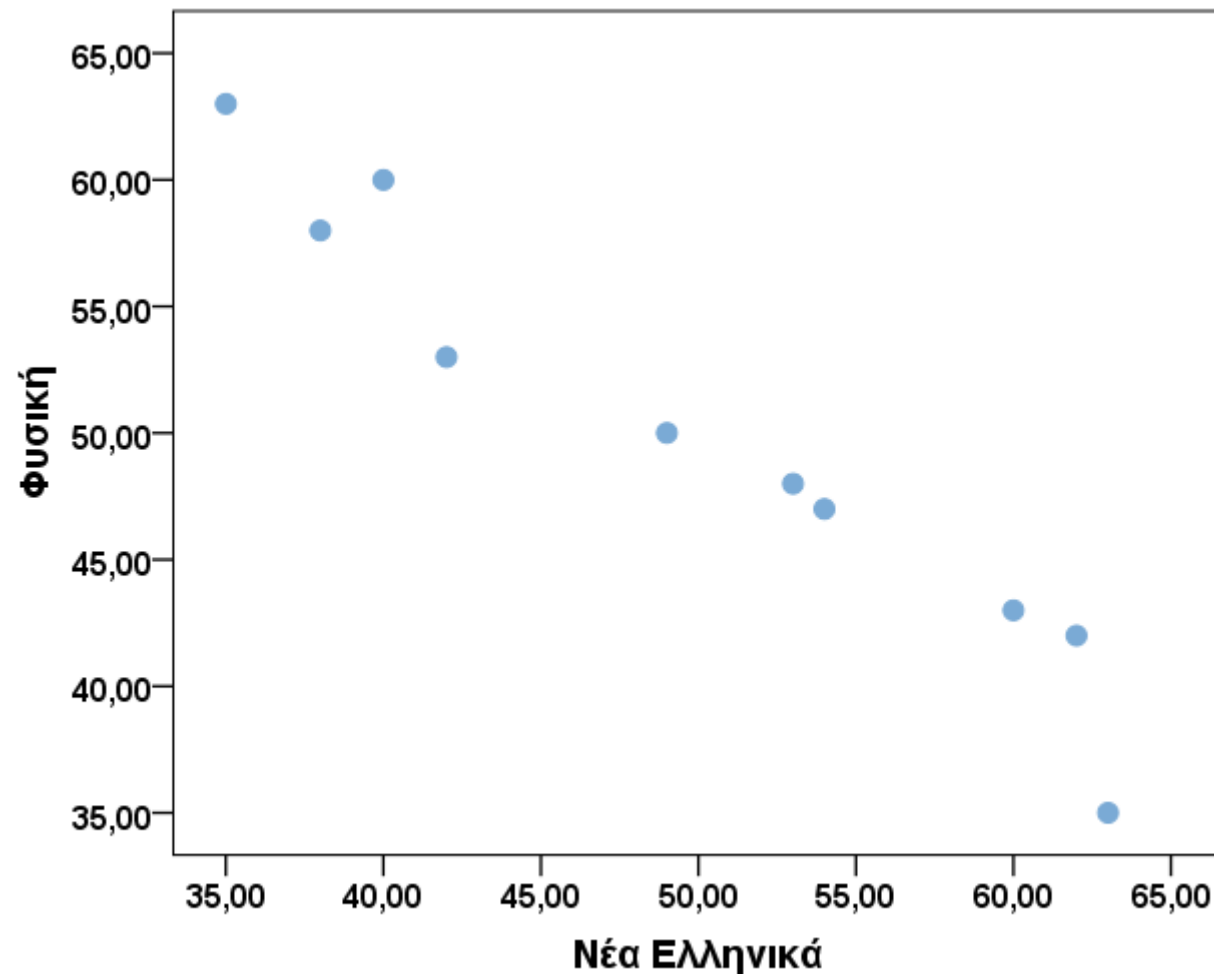
Διάγραμμα Διασποράς

- Ποιά είναι η σχέση ανάμεσα στις επιδόσεις στην Άλγεβρα και την Φυσική;



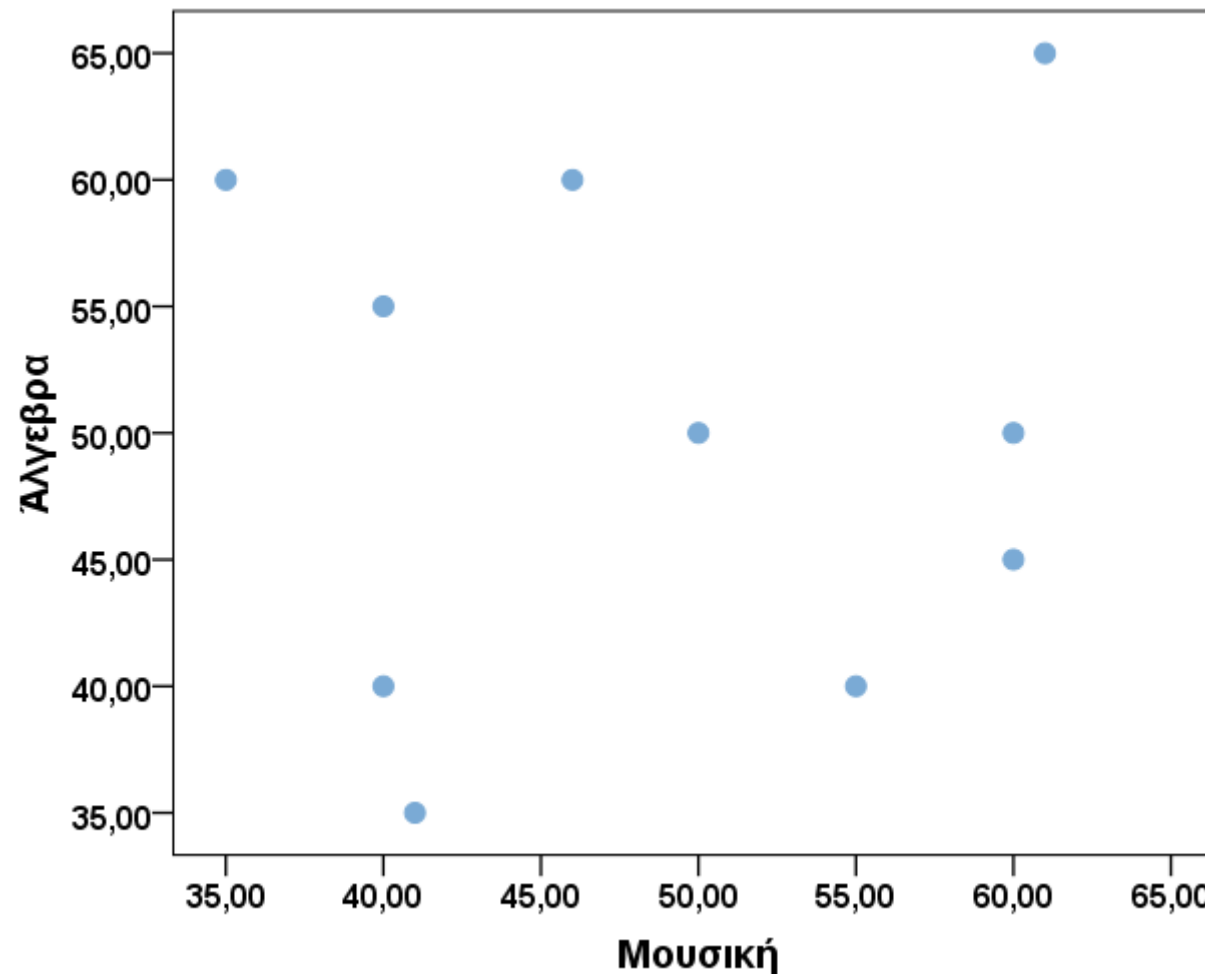
Διάγραμμα Διασποράς

- Ποιά είναι η σχέση ανάμεσα στις επιδόσεις στην Φυσική και τα Νέα Ελληνικά;



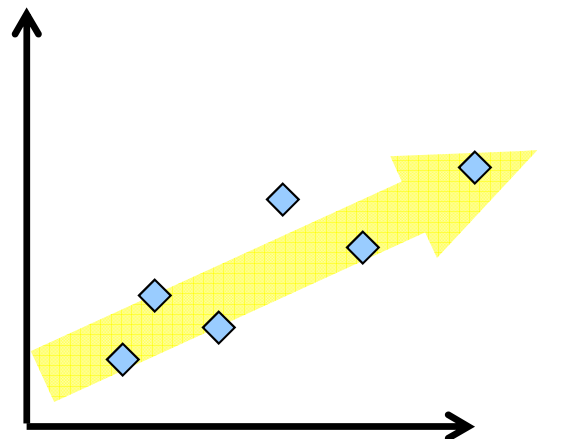
Διάγραμμα Διασποράς

- Ποιά είναι η σχέση ανάμεσα στις επιδόσεις στην Άλγεβρα και την Μουσική;

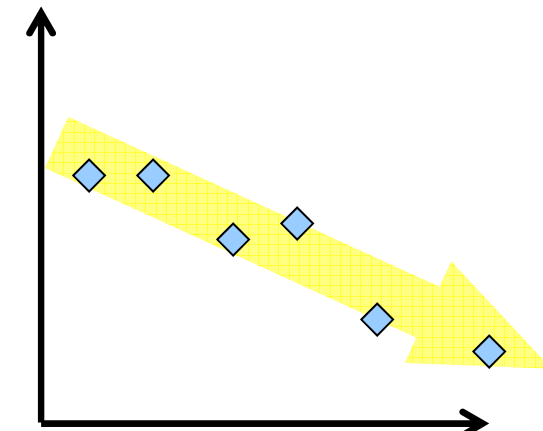


Διάγραμμα Διασποράς

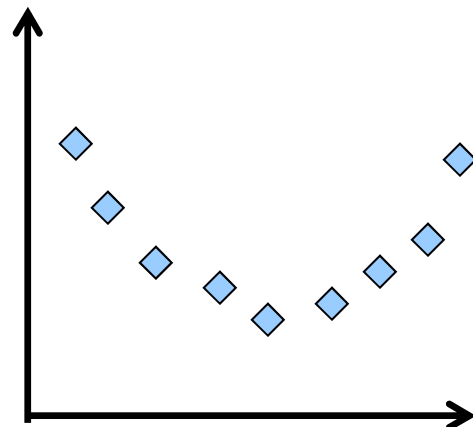
- Η Γραμμικότητα, η Κατεύθυνση και η Ένταση της σχέσης είναι τρεις έννοιες που μας ενδιαφέρουν



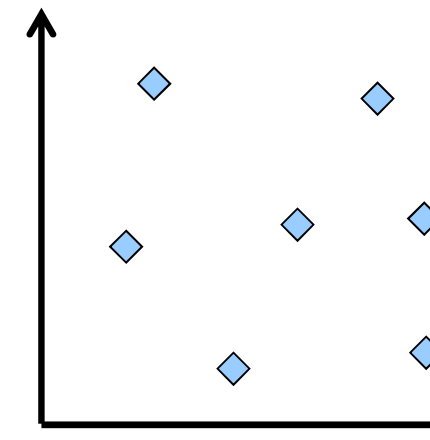
Θετική Γραμμική Σχέση



Αρνητική Γραμμική Σχέση



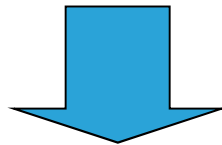
Μη Γραμμική Σχέση



Απουσία Γραμμικής Σχέσης

Πώς θα μετρήσουμε τη αλληλεξάρτηση δύο μεταβλητών

- Το στατιστικό κριτήριο που χρησιμοποιούμε για να διαπιστώσουμε αν υπάρχει **αλληλεξάρτηση** μεταξύ δύο μεταβλητών, ονομάζεται



ΔΕΙΚΤΗΣ ΣΥΣΧΕΤΙΣΗΣ
(CORRELATION COEFFICIENT)

Δείκτες Συσχέτισης

Οι σημαντικότεροι δείκτες συσχέτισης είναι:

➤ **Δείκτης γραμμικής συσχέτισης r του Pearson**

Μπορεί να χρησιμοποιηθεί όταν και οι δύο μεταβλητές είναι συνεχείς ποσοτικές.

➤ **Δείκτης ρ του Spearman**

Μπορεί να χρησιμοποιηθεί όταν η μία από τις δύο ή και οι δύο μεταβλητές είναι σε κλίμακα ιεράρχησης.

➤ **Ο δείκτης Συμφωνίας W του Kendall**

Μπορεί να χρησιμοποιηθεί για την αξιολόγηση του βαθμού συμφωνίας μεταξύ βαθμολογητών στην κατάταξη μιας σειράς θεμάτων.

Δείκτες Συσχέτισης

Οι σημαντικότεροι δείκτες συσχέτισης είναι:

➤ **Δείκτης Point-biserial r_{pb}**

Μπορεί να χρησιμοποιηθεί όταν η μία μεταβλητή είναι διχοτομική (π.χ. ναι/όχι) και η άλλη συνεχής.

➤ **Δείκτης biserial r_{pb}**

Μπορεί να χρησιμοποιηθεί όταν η μία μεταβλητή είναι διχοτομική, αλλά έχει προκύψει από τη συγχώνευση των τιμών μιας συνεχούς (π.χ. υψηλό/χαμηλό) και η άλλη συνεχής.

Η Ερμηνεία του Δείκτη Συσχέτισης r

Ο δείκτης συσχέτισης r αποτελείται από δύο στοιχεία:

➤ **ένα πρόσημο**

(πληροφορίες για την κατεύθυνση της σχέσης)

➤ **μία αριθμητική τιμή από 0 ως 1**

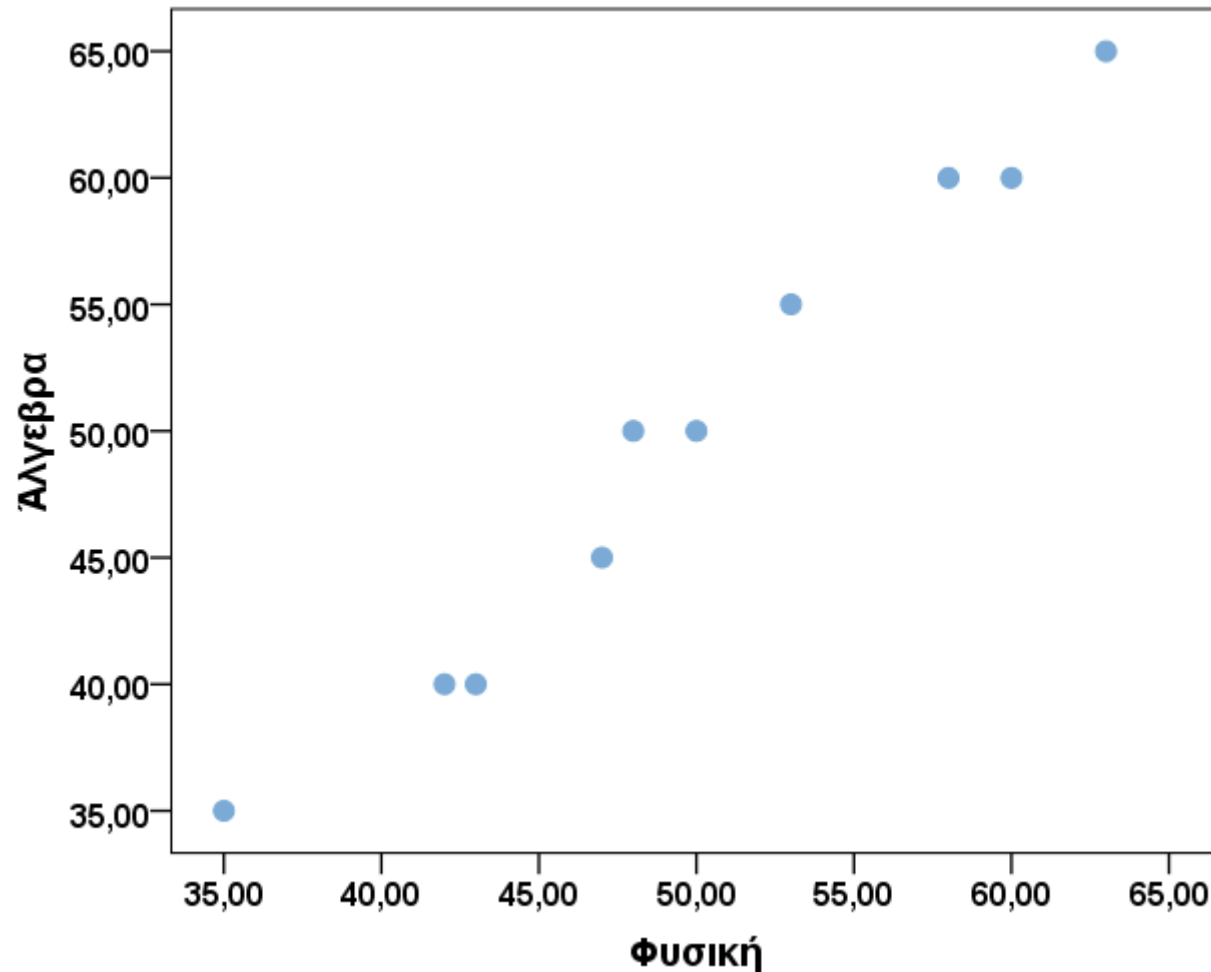
(πληροφορίες για το βαθμό/ένταση της σχέσης)

Θετική συσχέτιση (+) Όταν αυξάνεται η μια μεταβλητή, αυξάνεται και η άλλη

Αρνητική συσχέτιση (-) Όταν αυξάνεται η μια μεταβλητή, μειώνεται η άλλη

Διάγραμμα Διασποράς

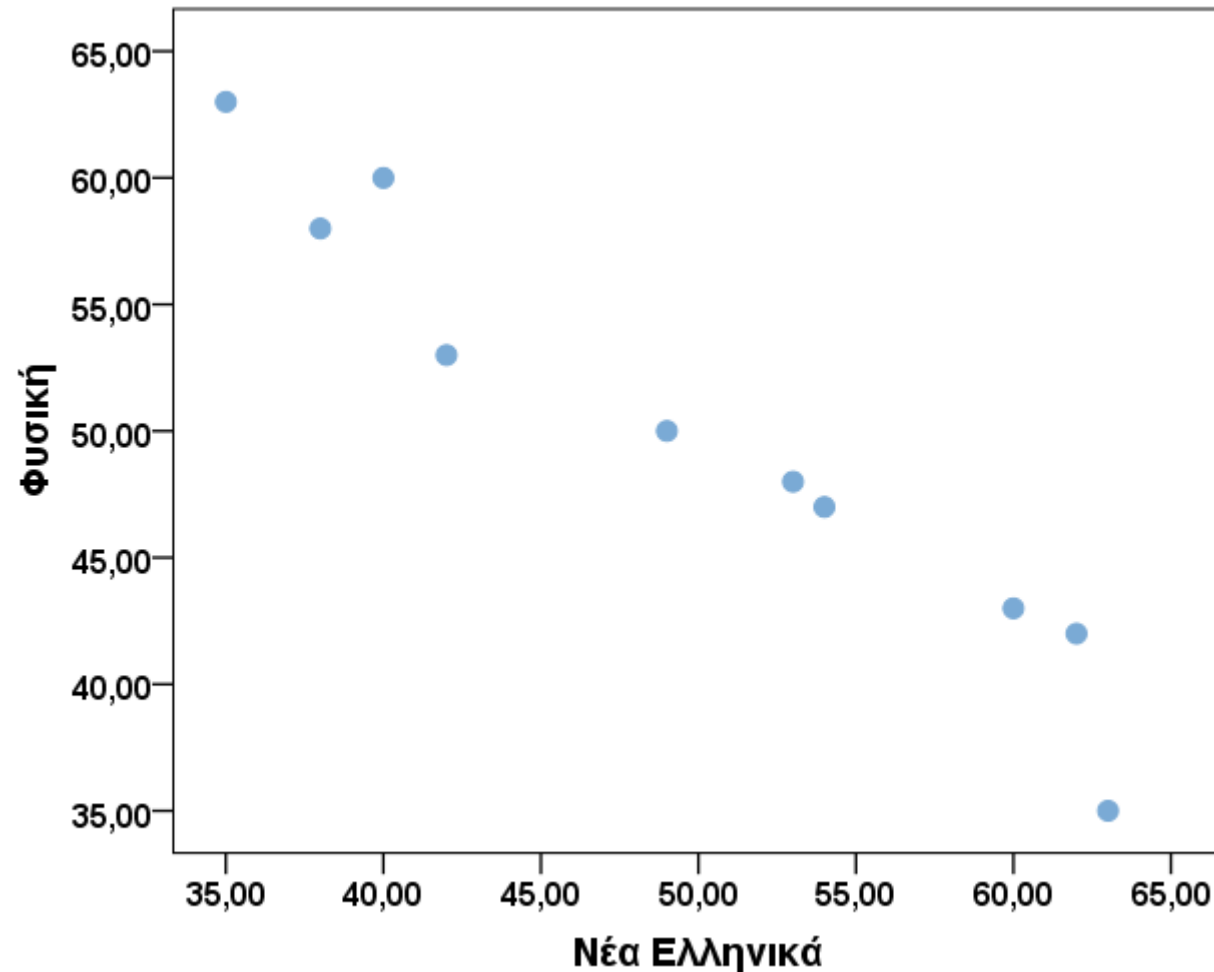
- Ποιά είναι η σχέση ανάμεσα στις επιδόσεις στην Άλγεβρα και την Φυσική;



Πολύ υψηλή Θετική Συσχέτιση $r = + 0,98$

Διάγραμμα Διασποράς

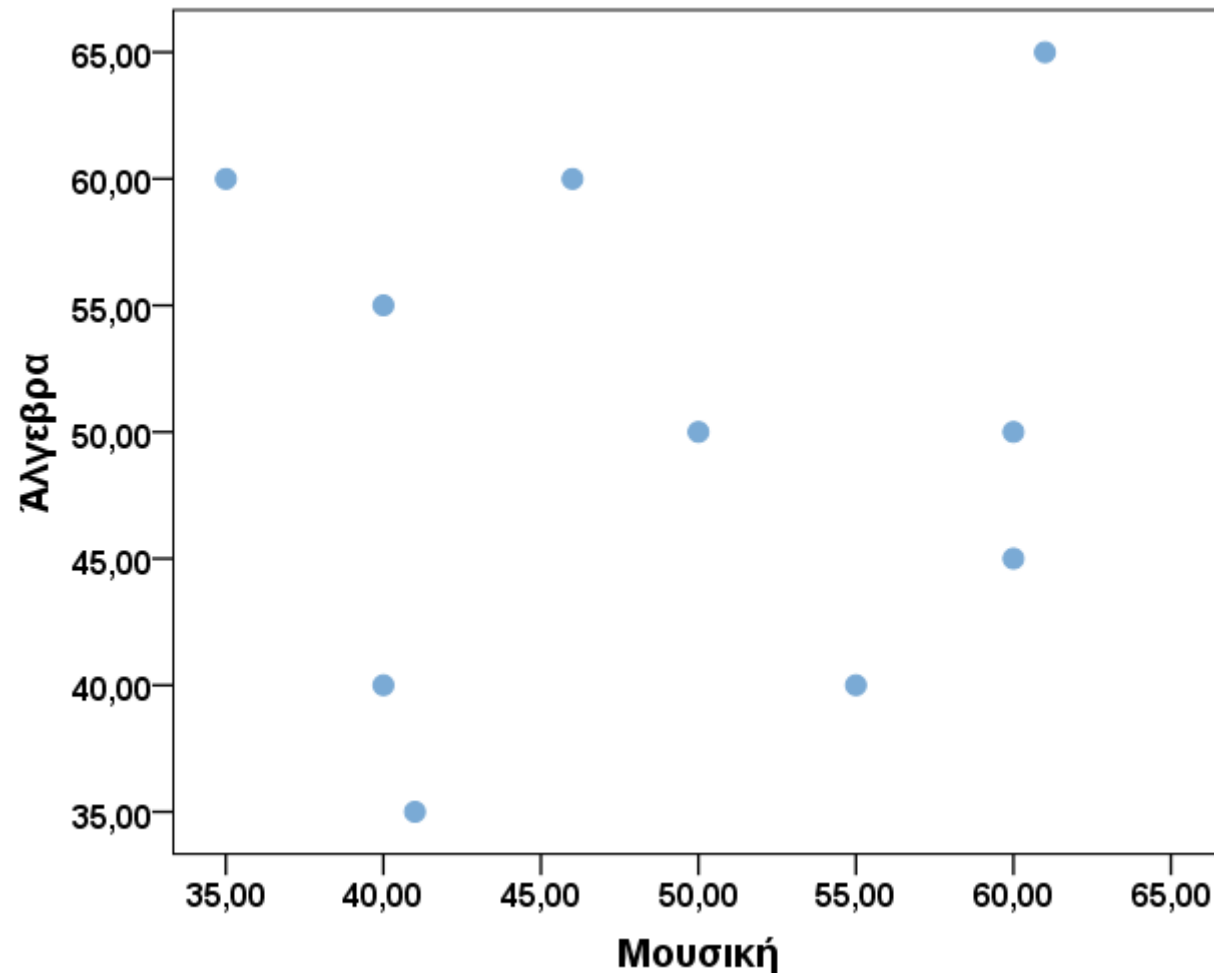
- Ποιά είναι η σχέση ανάμεσα στις επιδόσεις στην Φυσική και τα Νέα Ελληνικά;



Πολύ υψηλή Αρνητική Συσχέτιση $r = - 0,97$

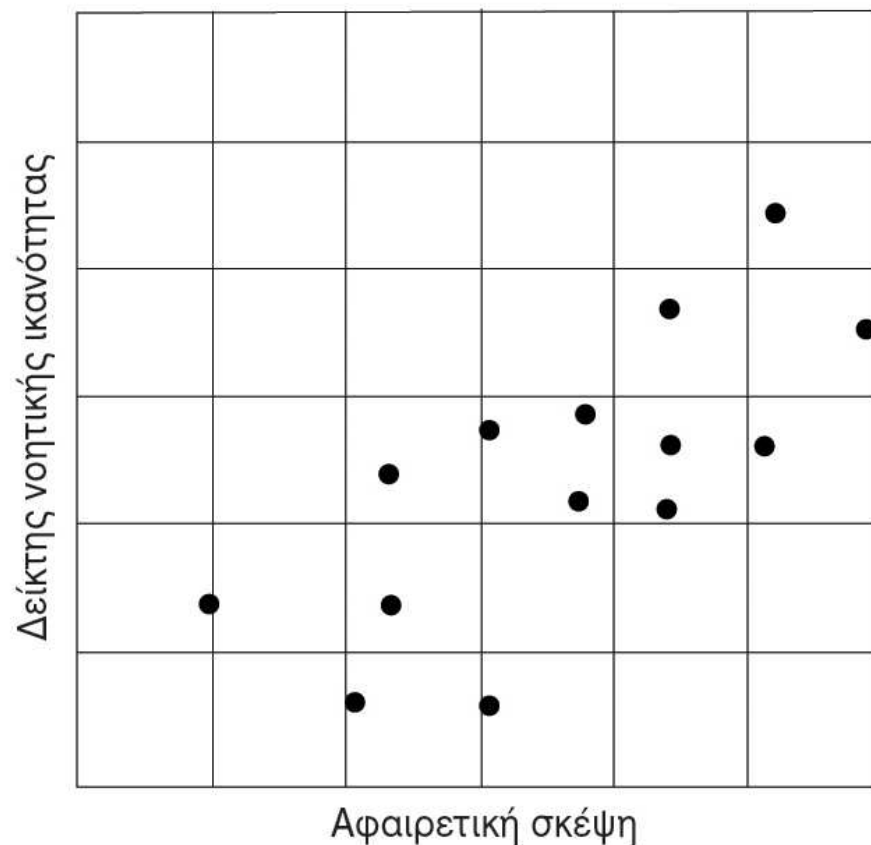
Διάγραμμα Διασποράς

- Ποιά είναι η σχέση ανάμεσα στις επιδόσεις στην Άλγεβρα και την Μουσική;



Μηδενική Συσχέτιση $r = 0,06$ (απουσία γραμμικής σχέσης)

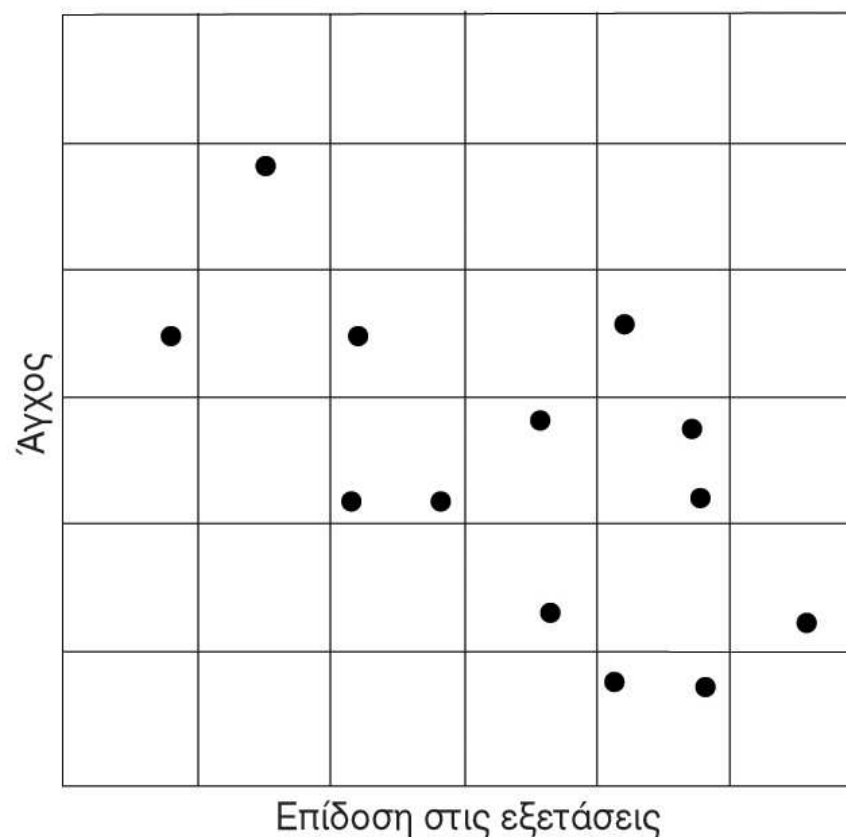
Το Διάγραμμα Διασποράς για τη Νοημοσύνη και την Αφαιρετική Σκέψη



© ΕΛΛΗΝΙΚΑ ΓΡΑΜΜΑΤΑ - ΡΟΥΣΣΟΣ & ΤΣΑΟΥΣΗΣ

Μέτρια Θετική Συσχέτιση $r = + 0,60$

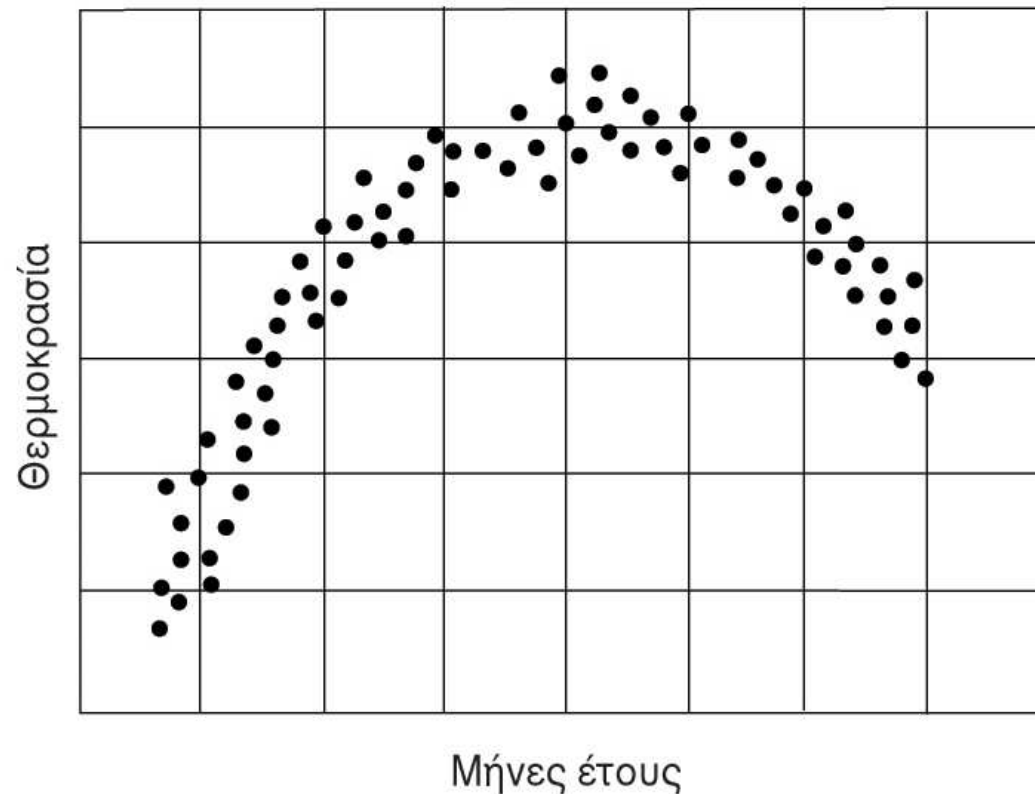
Διάγραμμα Διασποράς ανάμεσα στο άγχος και την επίδοση στις εξετάσεις



© ΕΛΛΗΝΙΚΑ ΓΡΑΜΜΑΤΑ - ΡΟΥΣΣΟΣ & ΤΣΑΟΥΣΗΣ

Μέτρια Αρνητική Συσχέτιση $r = - 0,52$

Διάγραμμα Διασποράς ανάμεσα στη θερμοκρασία και τους μήνες του έτους



© ΕΛΛΗΝΙΚΑ ΓΡΑΜΜΑΤΑ - ΡΟΥΣΣΟΣ & ΤΣΑΟΥΣΗΣ

Θετική Καμπυλόγραμμα Συσχέτιση (ακατάλληλος ο r)

Πληροφορίες που παίρνουμε από το δείκτη συσχέτισης r του Pearson

- Αν υπάρχει γραμμική συσχέτιση μεταξύ δύο μεταβλητών
- Το είδος (κατεύθυνση) της συσχέτισης
- Το βαθμό/ένταση της συσχέτισης

Η Ερμηνεία του Δείκτη Συσχέτισης r

Αν ο δείκτης είναι μικρότερος του ± 0.10

Δεν υπάρχει συσχέτιση

Αν ο δείκτης κυμαίνεται μεταξύ $\pm 0.10 - 0.40$

Χαμηλή συσχέτιση

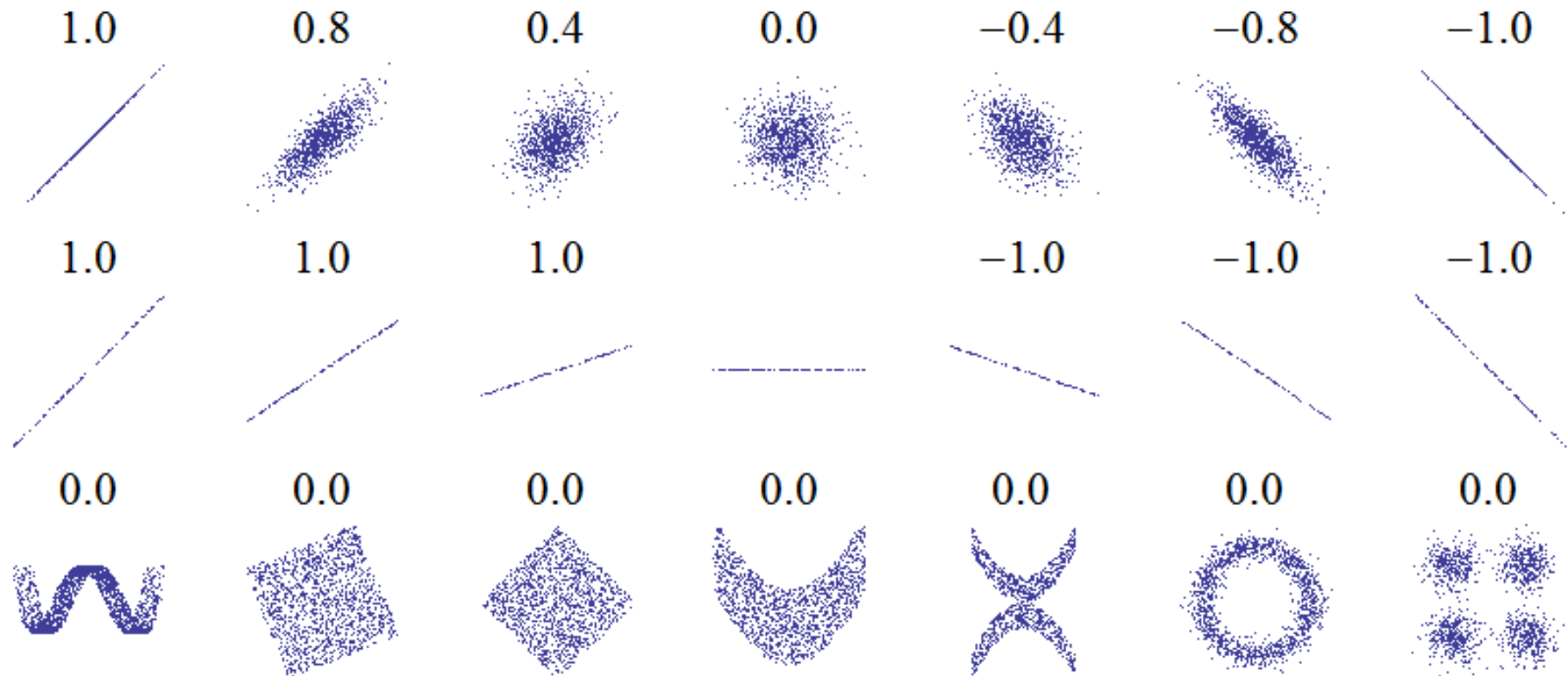
Αν ο δείκτης κυμαίνεται μεταξύ $\pm 0.40 - 0.70$

Μέτρια συσχέτιση

Αν ο δείκτης κυμαίνεται μεταξύ $\pm 0.70 - 0.99$

Υψηλή συσχέτιση

Καταλληλότητα Δείκτη Συσχέτισης r



Πηγή: Wikipedia

Ο δείκτης συσχέτισης αντανακλά την κατεύθυνση και την ένταση μιας γραμμικής σχέσης (επάνω σειρά), αλλά όχι την κλίση της (μεσαία σειρά), ούτε μη γραμμικές σχέσεις (κάτω σειρά).

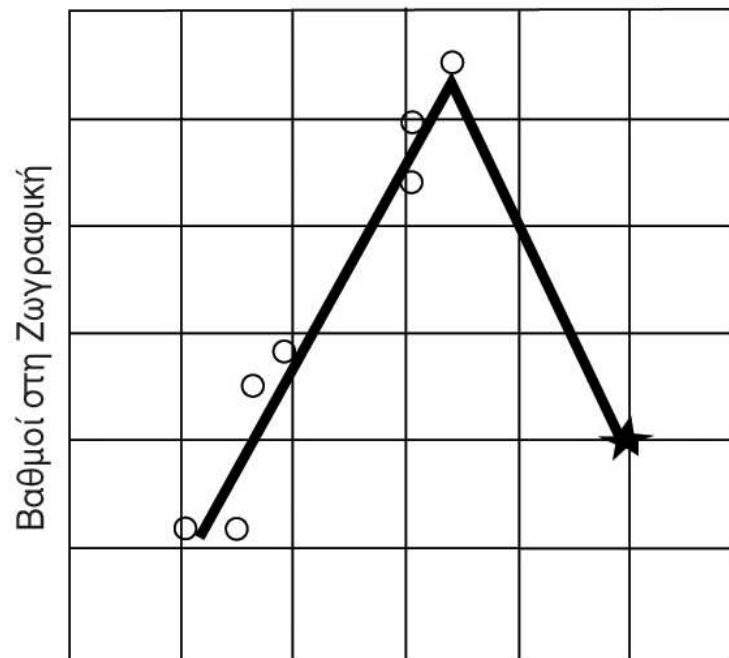
Παράγοντες που επηρεάζουν το δείκτη συσχέτισης r

- Η μη-γραμμική σχέση μεταξύ των μεταβλητών

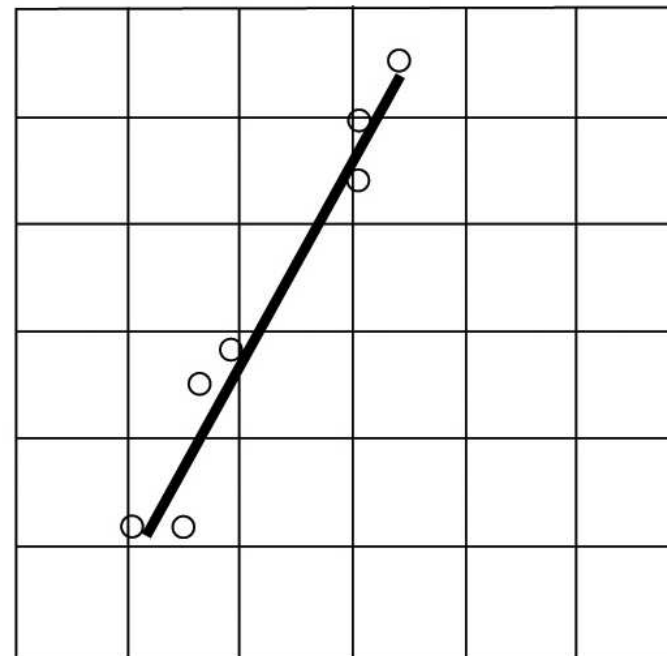
Εάν υπάρχει καμπυλόγραμμη (curvilinear) αντί για γραμμική σχέση ανάμεσα στις δύο μεταβλητές που μελετάμε είναι πιθανόν να βρεθεί χαμηλή ή και καθόλου συσχέτιση αλλά αυτό να μην ισχύει για **όλο το μήκος** της σχέσης

Παράγοντες που επηρεάζουν το δείκτη συσχέτισης r

- **Η ύπαρξη ακραίων τιμών**
Μπορεί να επηρεάσουν τη **μορφή** της συσχέτισης



(α)



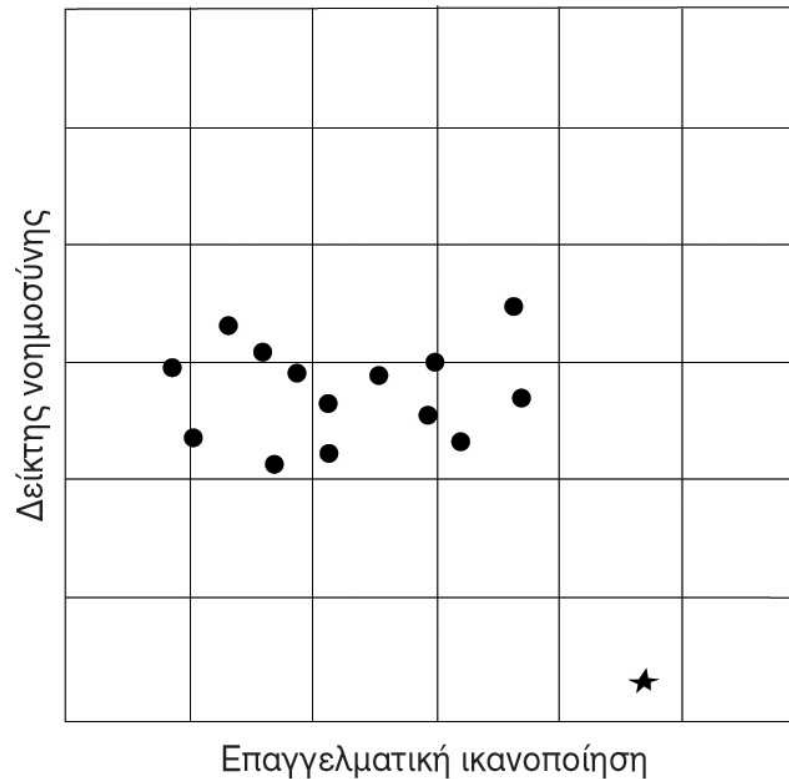
(β)

Παράγοντες που επηρεάζουν το δείκτη συσχέτισης r

➤ Η ύπαρξη ακραίων τιμών

Μπορεί να επηρεάσουν το **βαθμό** της συσχέτισης

Με την
ακραία τιμή
 $r = -0,48$



Χωρίς την
ακραία τιμή
 $r = 0,05$

Παράγοντες που επηρεάζουν το δείκτη συσχέτισης r

- **Η επίδραση ετερογενών υποομάδων ή άλλων μεταβλητών**

Η τιμή του δείκτη συσχέτισης δύο μεταβλητών μπορεί να μη αποδίδει την πραγματικότητα αν υπάρχουν διαφορές μεταξύ υποομάδων στο δείγμα ή επιδρούν και άλλες μεταβλητές στη σχέση των δύο μεταβλητών.

Για παράδειγμα, βρέθηκε ότι οι επιδόσεις των μαθητών μιας τάξης σε ένα μάθημα σχετίζονταν μέτρια θετικά με τις ώρες εργασίας στο σπίτι (r υψηλό). Όμως στην πραγματικότητα η σχέση αυτή οφειλόταν σε υψηλή συσχέτιση μόνο των επιδόσεων αγοριών με τις ώρες εργασίας και όχι και των κοριτσιών.

Προϋποθέσεις για τη χρήση του δείκτη

Pearson r

Κλίμακα Μέτρησης:	Ίσων Διαστημάτων ή αναλογική
Σχεδιασμός:	Εξαρτημένα Δείγματα
Γραμμικότητα:	Η συσχέτιση των δύο μεταβλητών να είναι γραμμική

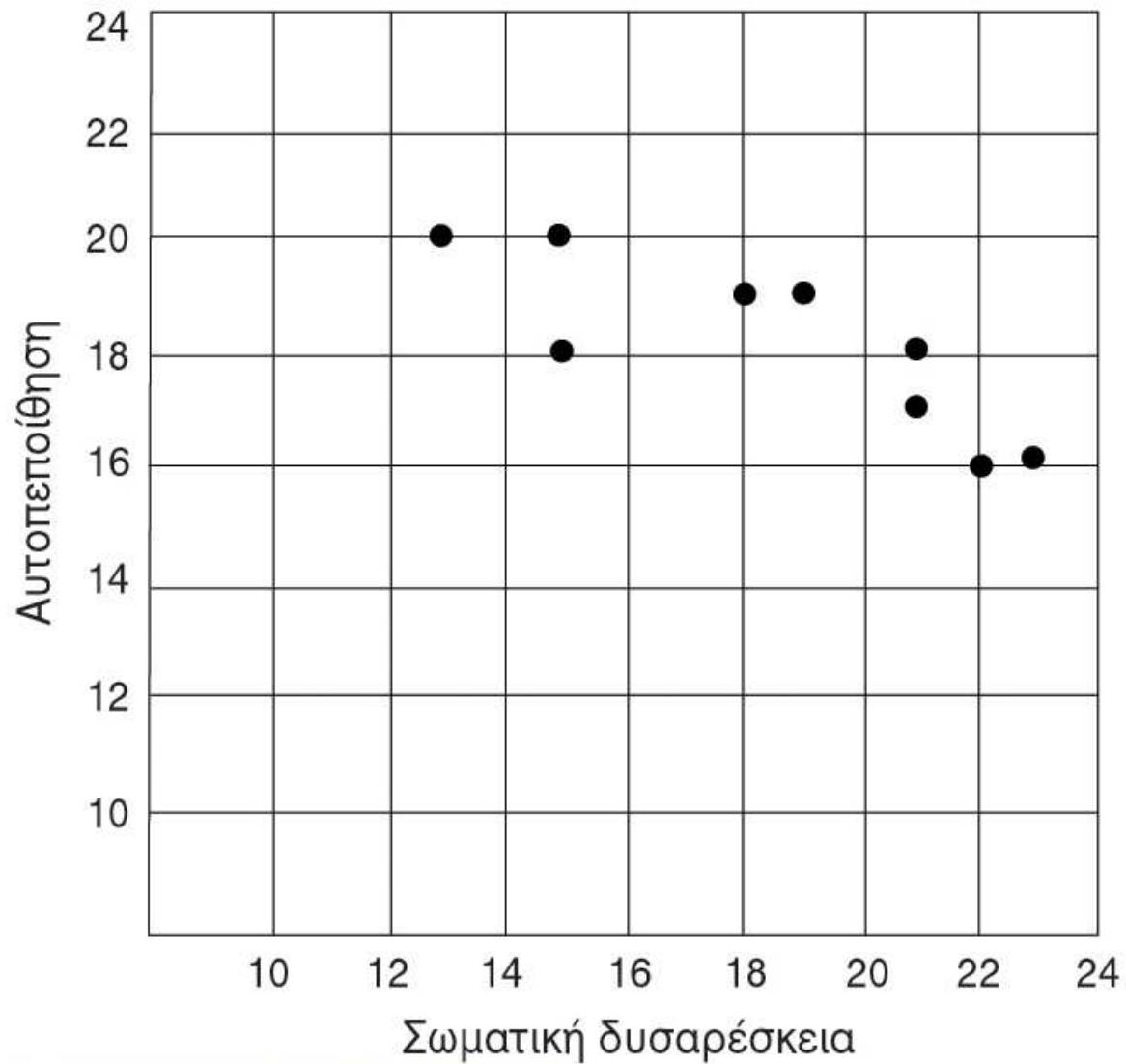
Παράδειγμα

- Ένας ερευνητής ενδιαφέρεται να μελετήσει εάν η δυσαρέσκεια του ατόμου για το σώμα του (X) **σχετίζεται** με τη γενικότερη αυτοπεποίθησή του (Y)
- 10 άτομα συμπλήρωσαν **δύο ερωτηματολόγια** που μετρούν αυτές τις μεταβλητές και συγκέντρωσε τα παρακάτω δεδομένα

Τα δεδομένα της Έρευνας

Άτομα	Σωματική Δυσαρέσκεια (X)	Αυτοπεποίθηση (Y)
1	15	18
2	23	17
3	19	19
4	22	16
5	21	18
6	21	17
7	23	16
8	15	20
9	18	19
10	13	20

Το διάγραμμα Διασποράς



Υπολογισμός του δείκτη r του Pearson

- Ο δείκτης αυτός προκύπτει διαιρώντας την **συν-διακύμανση (co-variance)** των δύο μεταβλητών με το **γινόμενο των τυπικών αποκλίσεών** τους.

$$r_{XY} = \frac{COV_{XY}}{\sigma_X \sigma_Y}$$

Υπολογισμός του δείκτη r του Pearson

$$r_{XY} = \frac{COV_{XY}}{\sigma_X \sigma_Y}$$

Όπου:

$$COV_{XY} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Υπολογισμός του δείκτη r του Pearson

$$r_{XY} = \frac{COV_{XY}}{\sigma_X \sigma_Y}$$

Όπου (τύπος χωρίς τους μέσους όρους):

$$COV_{XY} = \frac{1}{n-1} \left(\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n} \right)$$

Υπολογισμός του δείκτη r του Pearson

Ένας ακόμη τύπος:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Υπολογισμοί

Άτομα	Τιμές Χ	Τιμές Υ	(Χ x Υ)
1	15	18	270
2	23	17	391
3	19	19	361
4	22	16	352
5	21	18	378
6	21	17	357
7	23	16	368
8	15	20	300
9	18	19	342
10	13	20	260
	ΣΧ=190	ΣΥ=180	ΣΧΥ=3379

τυπικές αποκλίσεις:

$$\sigma_X = 3,62$$

$$\sigma_Y = 1,49$$

$$COV_{XY} = \frac{3379 - \frac{180 \times 190}{10}}{9} = \frac{3379 - 3420}{9} = \frac{41}{9} = 4,55$$

$$r_{XY} = -\frac{4,55}{3,62 \times 1,49} = \frac{4,55}{5,39} = -0,84$$

Διατύπωση Αποτελεσμάτων

Αριθμός ατόμων

Τιμή r

$$r(10) = -0,84, p=0,002$$

Στατιστικά
σημαντικό
αποτέλεσμα

Διατύπωση των Υποθέσεων

➤ *Μηδενική Υπόθεση:*

Ο βαθμός δυσαρέσκειας που έχει το άτομο για το σώμα του **δε συσχετίζεται** με το βαθμό αυτοπεποίθησής του

$$H_0: r = 0$$

➤ *Εναλλακτική Υπόθεση:*

Ο βαθμός δυσαρέσκειας που έχει το άτομο για το σώμα του **συσχετίζεται** με το βαθμό αυτοπεποίθησής του **(διπλής κατεύθυνσης)**

$$H_A: r \neq 0$$

ρ : πληθυσμιακός δείκτης συσχέτισης
 r : δειγματικός δείκτης συσχέτισης

Τα αποτελέσματα από το SPSS

Correlations

		X	Y
X	Pearson Correlation	1,000	-,844**
	Sig. (2-tailed)	,	,002
	N	10	10
Y	Pearson Correlation	-,844**	1,000
	Sig. (2-tailed)	,002	,
	N	10	10

** . Correlation is significant at the 0.01 level

Τα αποτελέσματα στο χαρτί

Πίνακας 1: Συσχέτιση Μεταξύ Δυσαρέσκειας με το Σώμα και Αυτοεκτίμησης

	X	Y
X		-.84**
Y	-.84**	

X: Δυσαρέσκεια με το Σώμα, Y: Αυτοεκτίμηση. N= 10.

** : Συσχέτιση σημαντική στο επίπεδο $\alpha=0.01$ (διπλής κατεύθυνσης).

Παράδειγμα αναφοράς

Η ανάλυση έδειξε ότι μεγαλύτερα επίπεδα δυσαρέσκειας με το σώμα σχετίζονταν με χαμηλότερα επίπεδα αυτοεκτίμησης

[$r(10) = -0.84$, $p < 0.01$]

Δείκτης Συσχέτισης και Αιτιότητα

- ❑ Ο δείκτης συσχέτισης είναι ένα στατιστικό κριτήριο που μας πληροφορεί μόνο για τη **συμμεταβολή** των δύο μεταβλητών που μελετώνται και όχι για το εάν υπάρχει **αιτιώδης σχέση** μεταξύ τους

Δείκτης Συσχέτισης και Αιτιότητα

- Η υψηλή συσχέτιση δεν δηλώνει σχέσεις **αιτίου** και **αποτελέσματος**. Μπορεί να οφείλεται σε μια **τρίτη μεταβλητή**, την οποία δεν έχουμε συμπεριλάβει στην έρευνά μας, και η οποία να λειτουργεί ως **αίτιο**

ΤΟ IQ ΣΑΣ ΕΠΗΡΕΑΖΕΙ ΤΟ ΠΟΣΟ ΠΙΝΕΤΕ

**Γιατί οι έξυπνοι άνθρωποι
καταναλώνουν περισσότερο αλκοόλ**



Πηγή: <http://goo.gl/o9eUX4>

...και ναρκωτικά.

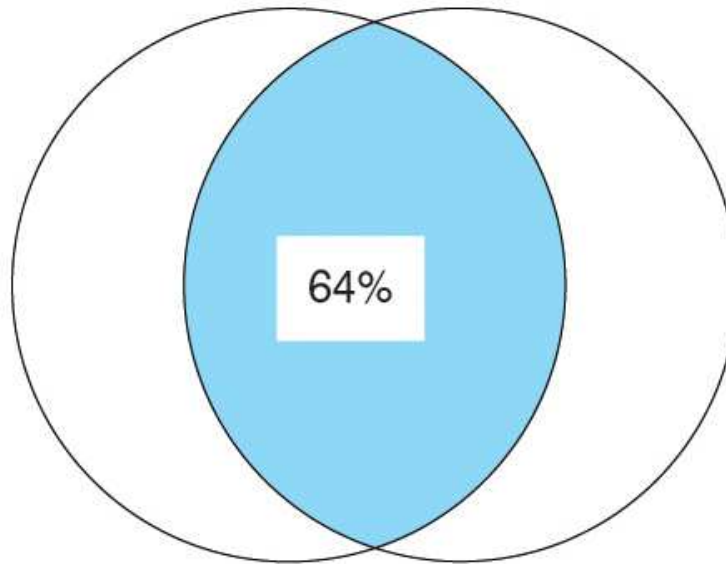
Ο δείκτης προσδιορισμού r^2

- Είναι ο δείκτης που μας δείχνει τι ποσοστό της **συνολικής διακύμανσης** της μεταβλητής X οφείλεται στη μεταβλητή Y , και συμβολίζεται με το γράμμα r^2 .
- Τον υπολογίζουμε εάν υψώσουμε στο τετράγωνο το **δείκτη συσχέτισης** των δύο μεταβλητών που μελετάμε και στη συνέχεια πολλαπλασιάζοντάς τον με το 100.

Πώς τον υπολογίζουμε;

Παράδειγμα:

Εάν $r = 0,80$, τότε $r^2 = 0,80^2 = 0,64 \times 100 = \mathbf{64\%}$



ΠΑΛΙΟΤΕΡΑ ΠΙΣΤΕΥΑ
ΟΤΙ Η ΣΥΣΧΕΤΙΣΗ
ΣΥΝΕΠΑΓΕΤΑΙ ΣΧΕΣΗ



ΥΣΤΕΡΑ ΔΙΑΒΑΣΑ
ΣΤΑΤΙΣΤΙΚΗ. ΤΩΡΑ
ΔΕΝ ΤΟ ΠΙΣΤΕΥΩ
ΠΙΑ.



ΦΑΙΝΕΤΑΙ ΟΤΙ ΤΟ
ΔΙΑΒΑΣΜΑ ΣΕ
ΒΟΗΘΗΣΕ.



ΧΜ, ΙΣΩΣ.

#END