CHAPITRE 5

RELATION ENTRE DEUX VARIABLES CONTINUES:

LE COEFFICIENT R DE PEARSON

INTRODUCTION

- Étudier UNE variable à la fois = **statistique univariée**.
- Étudier le comportement simultané de DEUX variables = **statistique bivariée**.

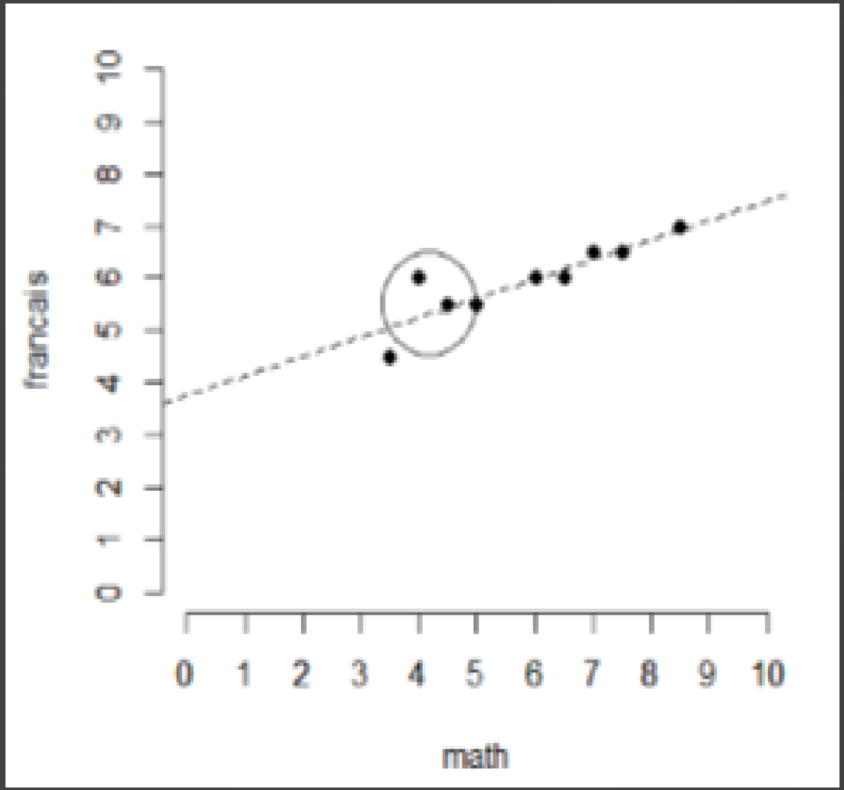
EXEMPLE: Y A-T-IL UN LIEN ENTRE LES NOTES AUX COURS DE MATH ET DE FRANÇAIS?

i	Math (X_i)	Français (Z_i)
1	8.5	7.0
2	4.0	6.0
3	4.5	5.5
4	6.0	6.0
5	3.5	4.5
6	6.5	6.0
7	7.0	6.5
8	5.0	5.5
9	7.5	6.5
10	7.5	6.5

REPRÉSENTATION GRAPHIQUE

REPRÉSENTER GRAPHIQUEMENT LA RELATION ENTRE LES COURS DE MATH ET DE FRANÇAIS

- Nuage de points



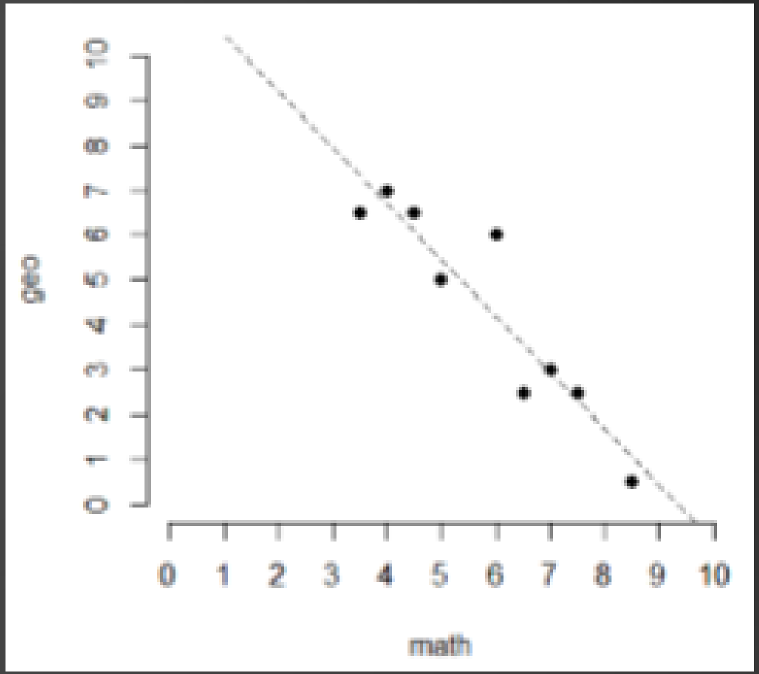
Le nuage peut-il être simplifié par une droite?

Relation **linéaire**

REPRÉSENTER GRAPHIQUEMENT LA RELATION ENTRE LES COURS DE MATH ET DE GÉO

i	math(X_i)	geographie(W_i)
1	8.5	0.5
2	4.0	7.0
3	4.5	6.5
4	6.0	6.0
5	3.5	6.5
6	6.5	2.5
7	7.0	3.0
8	5.0	5.0
9	7.5	2.5
10	7.5	2.5

Nuage de points



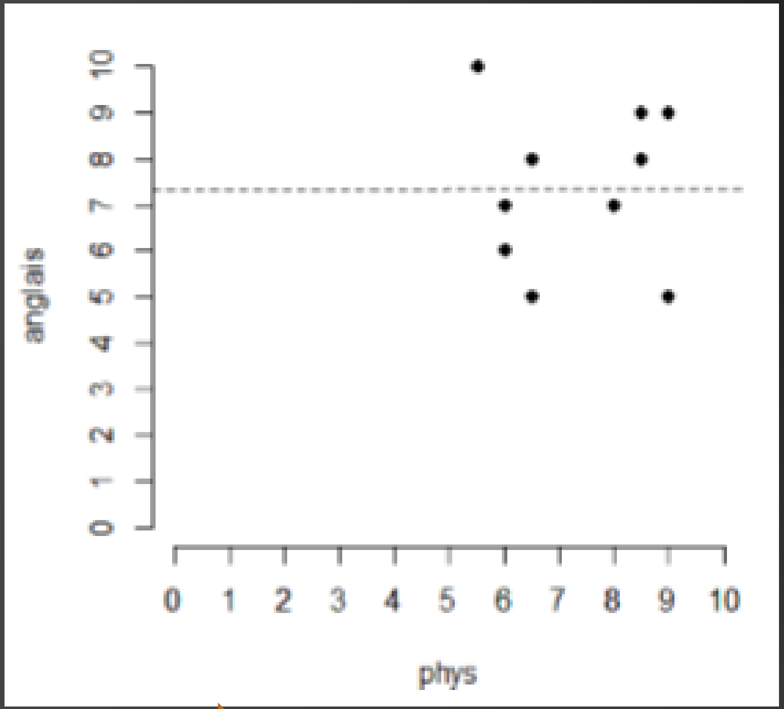
Relation **linéaire**

REPRÉSENTER GRAPHIQUEMENT LA RELATION ENTRE LES COURS D'ÉDUCATION PHYSIQUE ET D'ANGLAIS

i	**education physique (V_i)**	**Anglais (V_i)**
1	5.5	10
2	9.0	5
3	6.0	6
4	6.0	7
5	6.5	5
6	6.5	8
7	8.0	7
8	8.5	9
9	8.5	8
10	9.0	9

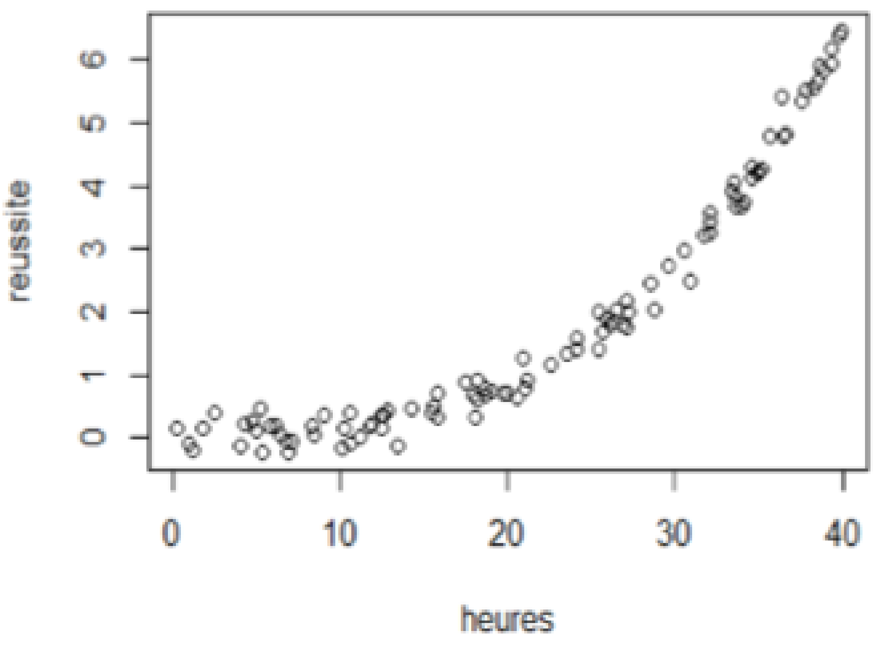
ERRATUM: même lettre pour anglais et éducation physique

Nuage de points



Pas de relation

RELATION ENTRE LE NOMBRE D'HEURES D'ÉTUDES ET LA RÉUSSITE



Le nuage peut-il être simplifié par une droite?

Relation **non linéaire**

REPRÉSENTATION ALGÉBRIQUE DE LA RELATION **LINÉAIRE**

RAPPEL CALCULER LA MOYENNE ET L'ÉCART-TYPE DES DEUX SÉRIES SUIVANTES:

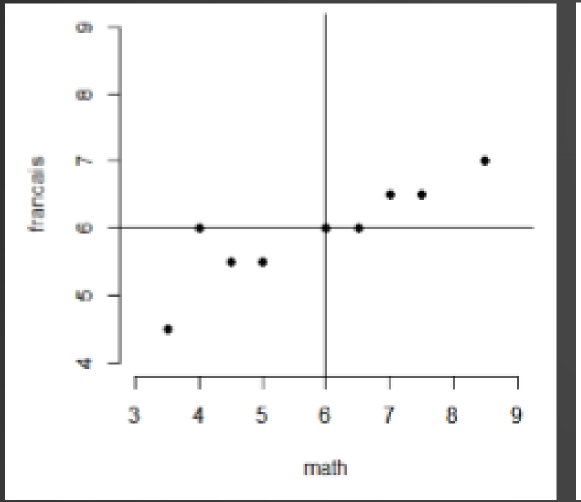
i	Math (X_i)	Français (Z_i)
1	8.5	7.0
2	4.0	6.0
3	4.5	5.5
4	6.0	6.0
5	3.5	4.5
6	6.5	6.0
7	7.0	6.5
8	5.0	5.5
9	7.5	6.5
10	7.5	6.5

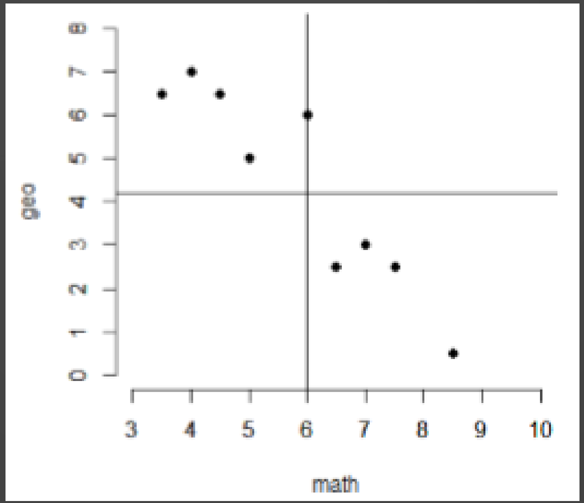
- $\bar{X}$=6
- $\bar{Y}$=6

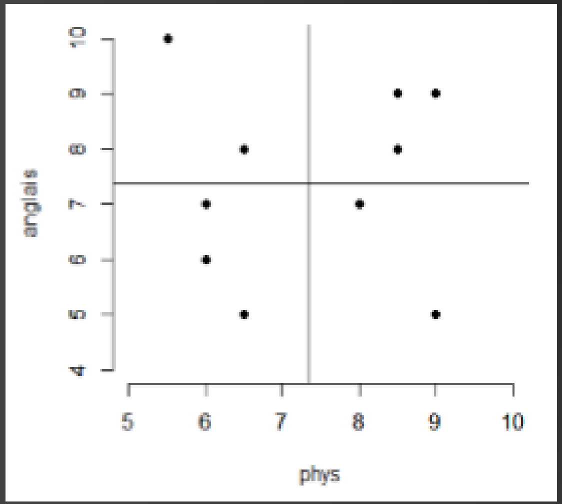
- S_X^2 =2.55
- S_Y^2 =2

CALCUL DE LA COVARIANCE ENTRE DEUX COURS

$$S_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{n}$$







LIMITE DE LA COVARIANCE

$$S_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{n}$$

- Mesurez la relation entre la longueur et la largeur des 5 objets suivants:

Objet	Longueur	Largeur
1	2	1
2	4	3
3	6	4
4	8	7
5	10	8
Moyenne:	**6**	**4,6**

S_{XY} = 72000

LIMITE DE LA COVARIANCE

$$S_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{n}$$

- Mesurez la relation entre la longueur et la largeur des 5 objets suivants:

Objet	Longueur	Largeur
1	200	100
2	400	300
3	600	400
4	800	700
5	1000	800
Moyenne:	**600**	**460**

➡ $S_{XY} = 7{,}2$

LIMITES DE LA COVARIANCE

- On étudie deux fois les 5 mêmes objets → la relation entre l et L est donc la même

- Pourtant, la mesure de covariance diffère

→ La mesure de covariance dépend de l'**unité de mesure**

SOLUTION = CORRÉLATION

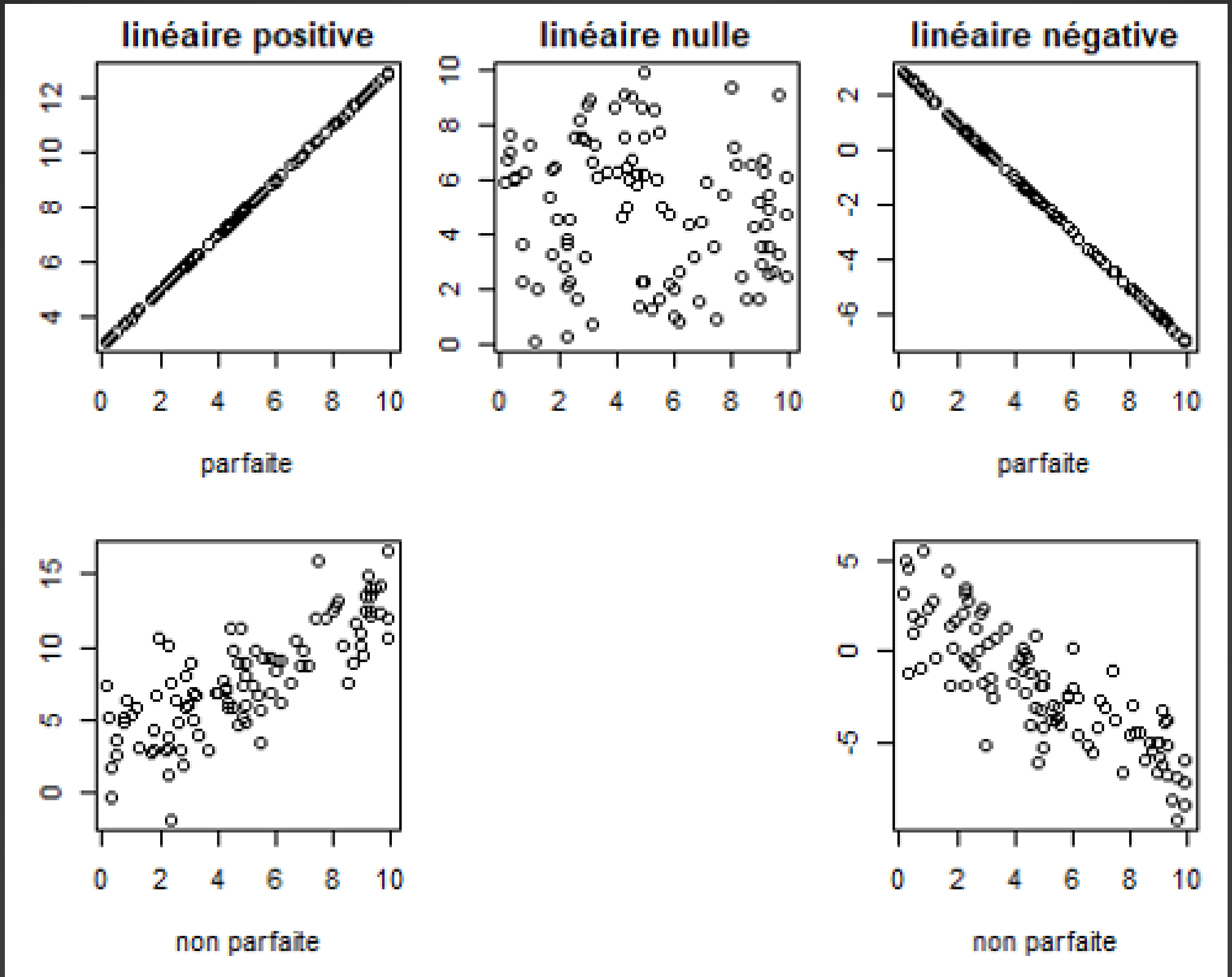
- Mesure standardisée de la covariance
- Mesure comprise entre -1 et 1

$$r_{XY} = \frac{s_{XY}}{s_X s_Y}$$

Objet	Longueur	Largeur
1	2	1
2	4	3
3	6	4
4	8	7
5	10	8
Moyenne:	**6**	**4,6**

Objet	Longueur	Largeur
1	200	100
2	400	300
3	600	400
4	800	700
5	1000	800
Moyenne:	**600**	**460**

Corrélation identique dans les deux cas = 0,988



linéaire positive
linéaire nulle
linéaire négative
parfaite
parfaite
non parfaite
non parfaite