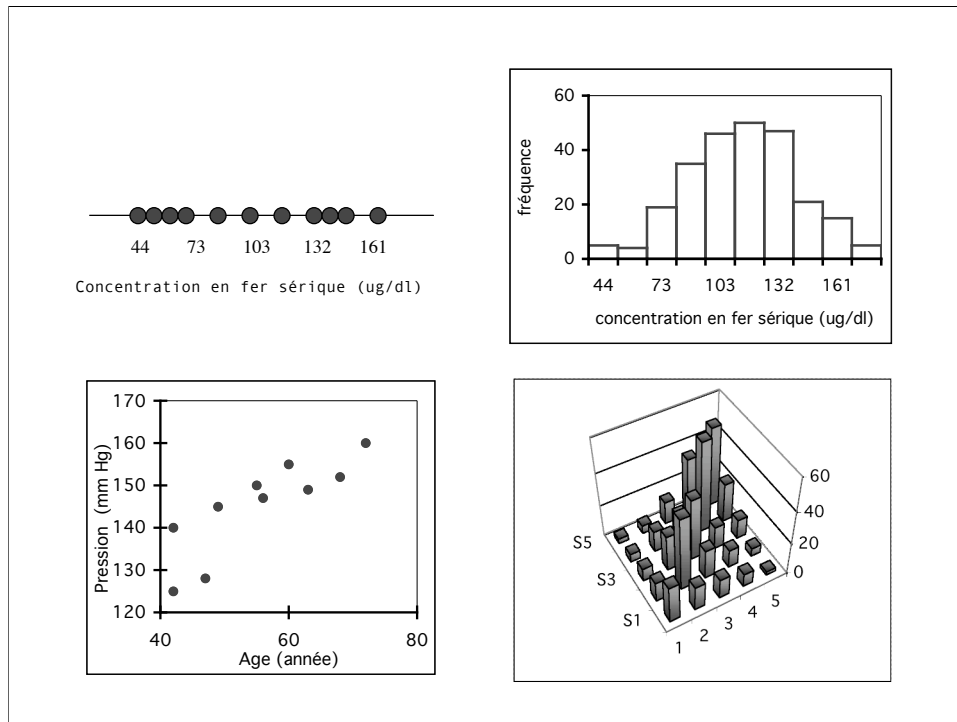
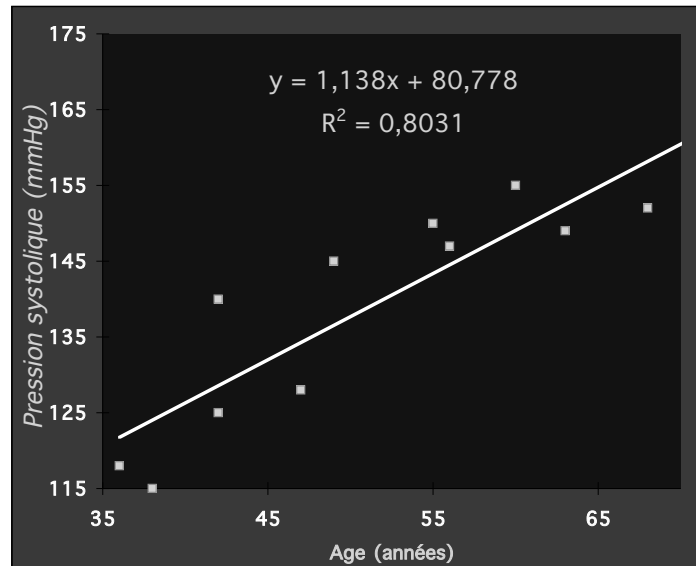




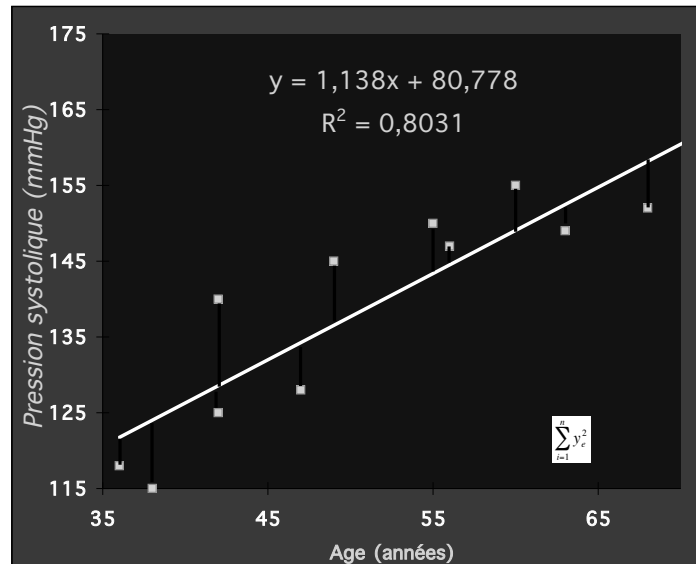
De ces 4 types de graphiques, l'usage retient essentiellement l'histogramme (en haut à droite) pour représenter une dimension et le diagramme de dispersion (en bas à gauche) pour représenter 2 dimensions

Afin d'étudier la relation qui pourrait exister entre l'âge et la pression sanguine, un médecin mesure sur 12 femmes d'âges différents la pression sanguine systolique (mm Hg).



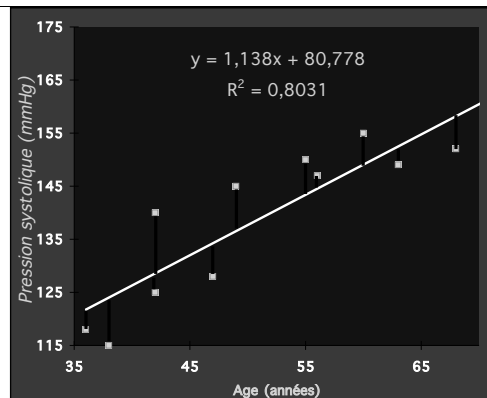
La droite des moindres carrés modélise la relation entre X et Y par une équation simple $Y_m = B_1X + B_0$

Afin d'étudier la relation qui pourrait exister entre l'âge et la pression sanguine, un médecin mesure sur 12 femmes d'âges différents la pression sanguine systolique (mm Hg).



Dans la droite des moindres carrés la valeur des paramètres B_1 et B_0 est obtenue en minimisant globalement les écarts en Y (ordonnée). Ces écarts s'écrivent $Y_o - Y_m = Y_e$, Y_o étant la valeur mesurée dans l'expérience, $Y_m = B_1X + B_0$ et Y_e l'écart entre les deux valeurs. La valeur optimale du couple (B_1, B_0) correspond à la valeur minimale de la somme des carrés des écarts au modèle. Ce modèle implique que les valeurs de X ne soient pas aléatoires.

X	Y _o	Y _m	Y _e
56	147	144,5	2,5
42	125	128,6	-3,6
72	160	162,7	-2,7
36	118	121,7	-3,7
63	149	152,5	-3,5
47	128	134,3	-6,3
55	150	143,4	6,6
49	145	136,5	8,5
38	115	124,0	-9,0
42	140	128,6	11,4
68	152	158,2	-6,2
60	155	149,1	5,9
SCE	2500,7	2008,2	492,5
R ²	0,8031		



X non aléatoire

Variabilité = information + bruit

SCE Y_o = SCE Y_m + SCE Y_e

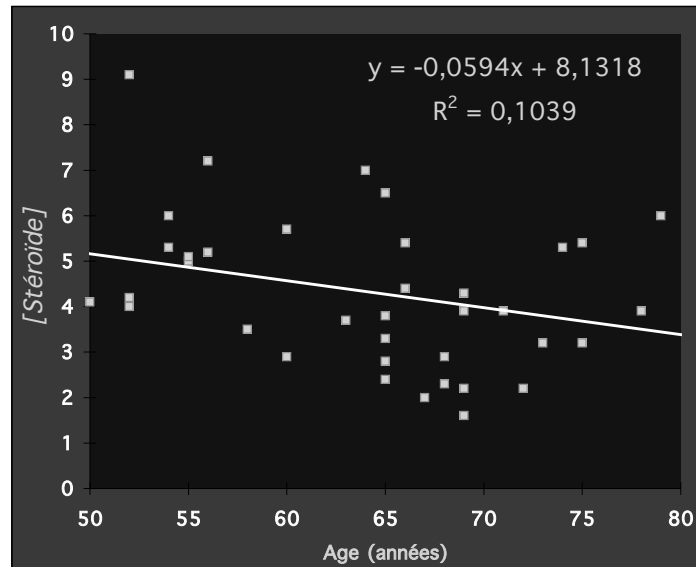
2500,7 = 2008,2 + 492,5

$$R^2 = \frac{\text{information}}{\text{Variabilité}} = \frac{\text{SCE Y}_m}{\text{SCE Y}_o} = \frac{2008,2}{2500,7} = 0,8031$$

La variabilité de Y_o est décomposée en deux SCE additives: la variabilité expliquée par le modèle et la variabilité inexpliquée par le modèle (bruit).

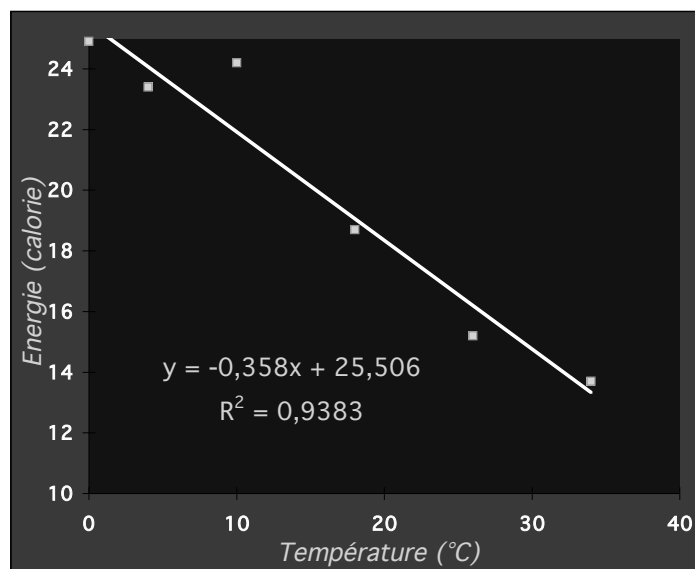
Le coefficient de détermination R² exprime la proportion de la variabilité de Y_o qui est expliquée par le modèle. Il mesure la qualité du modèle.

Pour 39 individus adultes, un endocrinologue a noté l'âge et la concentration urinaire d'un stéroïde particulier.



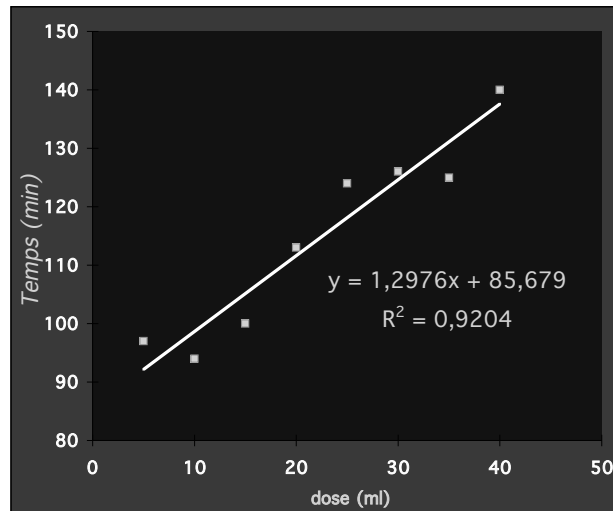
La qualité du modèle linéaire est médiocre : il explique seulement 10% de la variabilité de Y_0

Un physiologiste a mesuré la quantité d'énergie métabolisée en 10 heures (calories) par un moineau soumis à différentes températures (° Celsius).



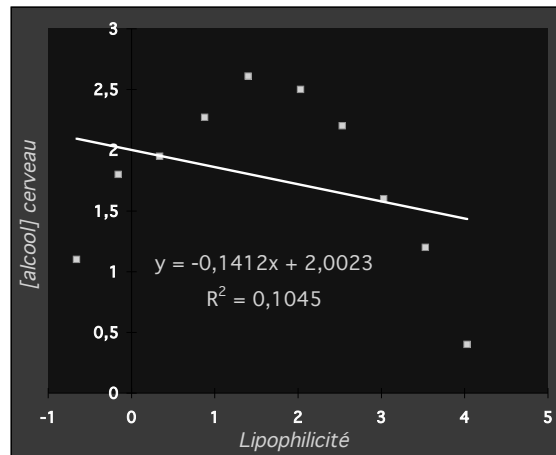
La qualité du modèle linéaire est bonne (93% de la variabilité expliquée)

Une étude est menée en vue de déterminer l'influence de l'alcool sur le comportement humain. Huit individus prélevés au hasard reçoivent chacun une dose (ml) de whisky. Après une demi-heure, le temps (secondes) mis par chaque sujet pour réaliser un test est mesuré.



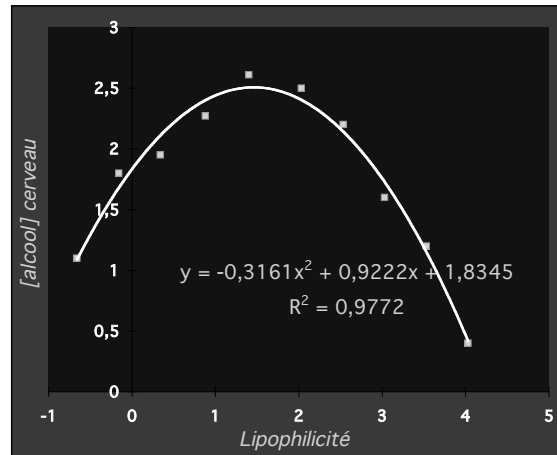
La qualité du modèle linéaire est bonne (92% de la variabilité expliquée)

Un neurophysiologiste étudie le passage de différents alcools du sang vers le cerveau en fonction de leur lipophilicité (mesurée par le coefficient de partage). Il injecte à 10 patients des alcools différents et note, pour chaque patient, la lipophilicité de l'alcool et sa concentration dans le cerveau.



La qualité du modèle linéaire est médiocre : il explique seulement 10% de la variabilité de Y_o

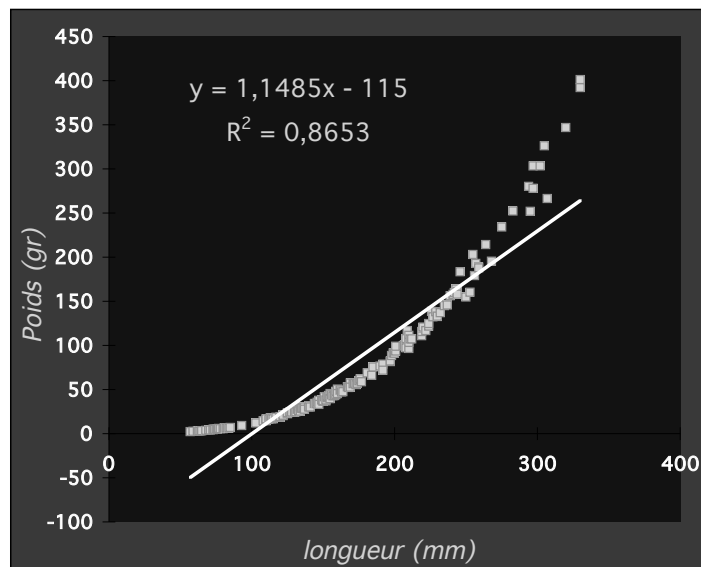
Un neurophysiologiste étudie le passage de différents alcools du sang vers le cerveau en fonction de leur lipophilicité (mesurée par le coefficient de partage). Il injecte à 10 patients des alcools différents et note, pour chaque patient, la lipophilicité de l'alcool et sa concentration dans le cerveau.



La qualité du modèle quadratique est bonne : il explique 97% de la variabilité de Y_o

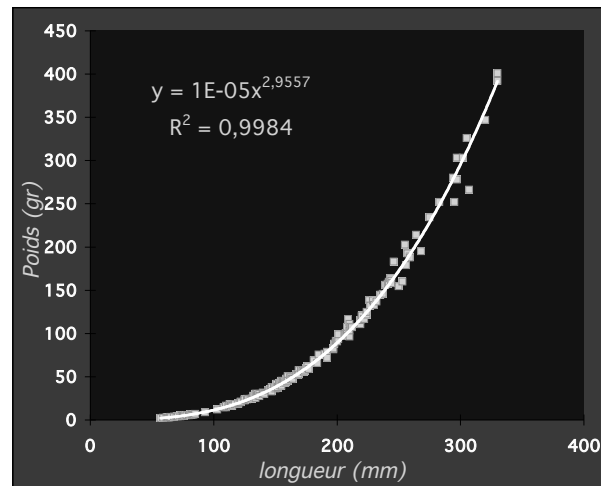
La référence de R^2 est le modèle, qui n'est pas nécessairement linéaire

Des écologistes ont capturé des truites (*Salmo Trutta Fario*) dans le Samson et ont mesuré la longueur à la fourche (mm) et le poids (gr) de chaque animal.



La qualité du modèle linéaire est bonne (86% de la variabilité expliquée). CE R2 élevé n'est pas la preuve que le modèle linéaire est le plus adéquat.

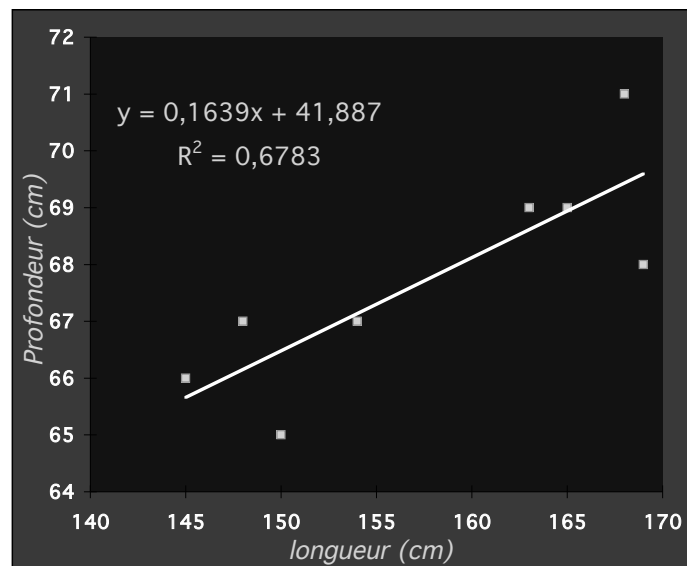
Des écologistes ont capturé des truites (*Salmo Trutta Fario*) dans le Samson et ont mesuré la longueur à la fourche (mm) et le poids (gr) de chaque animal. *La droite des moindres rectangles serait plus adaptée, bien que l'on puisse considérer que la taille est constante durant Δt tandis que le poids est variable dans ce même Δt .*



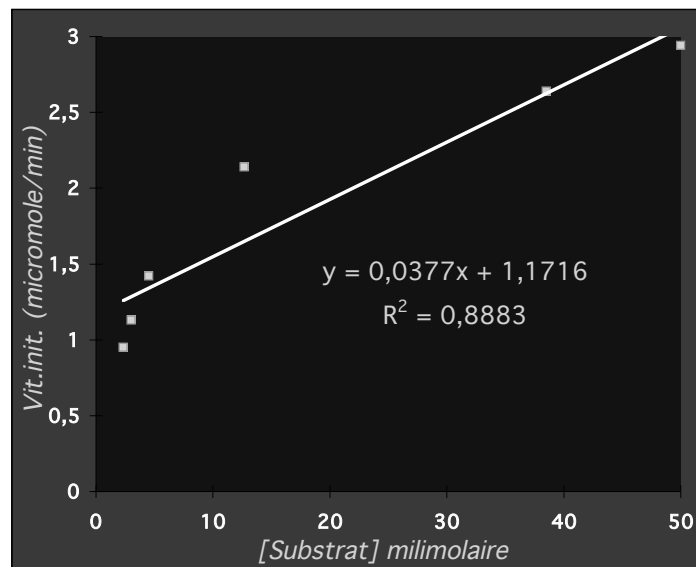
La qualité du modèle puissance est meilleure que celle du modèle linéaire : il explique 97% de la variabilité de Y_o

La référence de R^2 est le modèle, qui n'est pas nécessairement linéaire

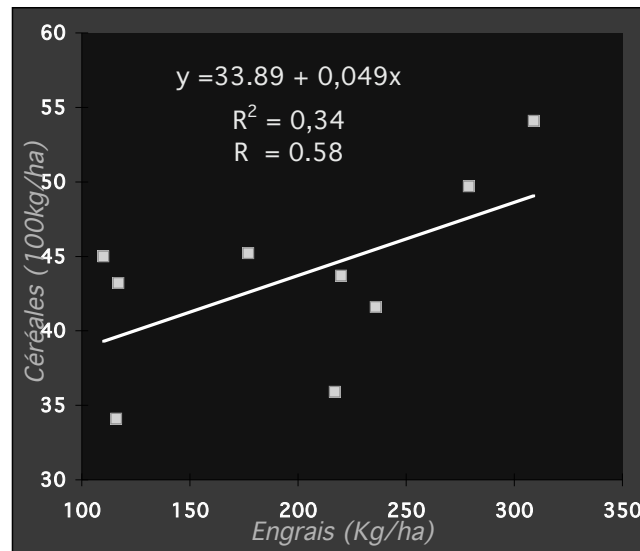
Longueur du corps (cm) et profondeur de poitrine (cm)
de 22 vaches laitières.



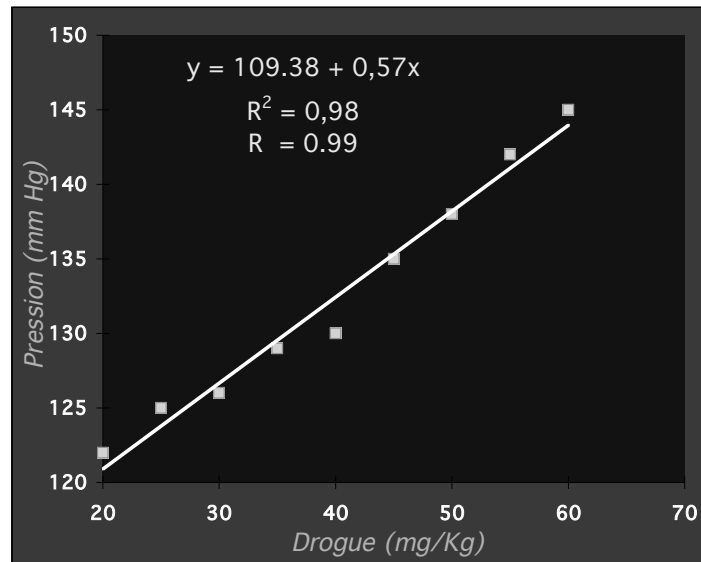
Un biochimiste a déterminé les vitesses initiales (micromoles/min) d'une enzyme pour différentes concentrations en substrat (millimolaires).



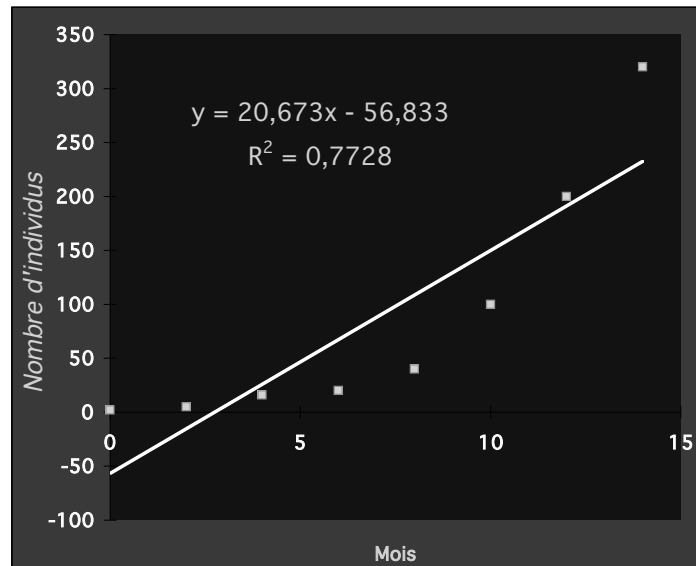
Les statistiques suivantes reprennent, pour 9 pays , en moyenne par hectare, les quantités d'engrais déversées et les quantités de céréales récoltées.



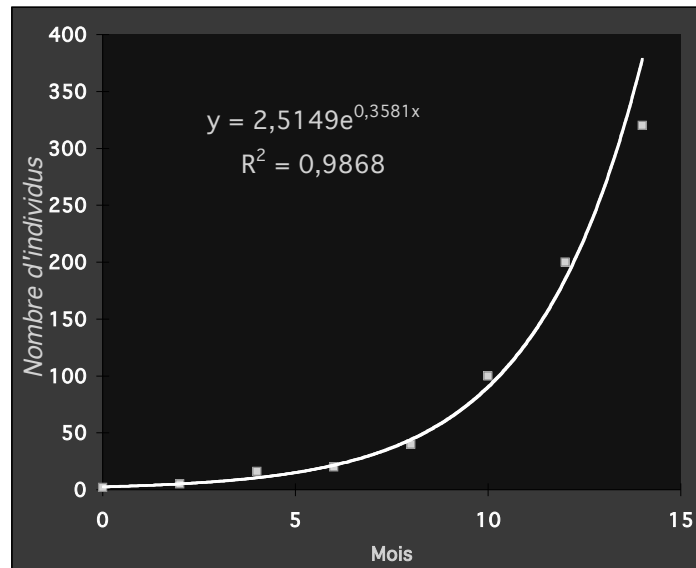
Une physiologiste s'intéresse à l'influence d'une drogue sur la pression sanguine. Elle décide de mesurer la pression sanguine de 9 individus ayant reçu chacun une quantité différente de drogue.



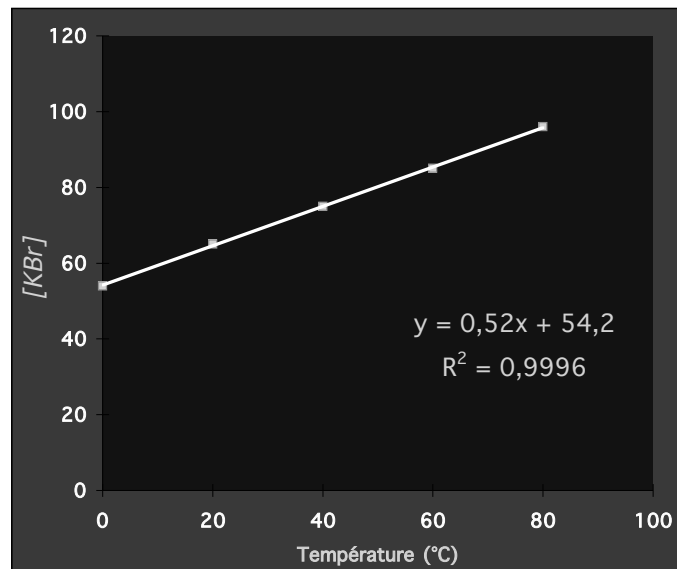
L'étude de la croissance d'une population de campagnols des champs (*Microtus Arvalis*) a fourni pour le nombre d'individus en fonction du temps les résultats suivants.



L'étude de la croissance d'une population de campagnols des champs (*Microtus Arvalis*) a fourni pour le nombre d'individus en fonction du temps les résultats suivants.

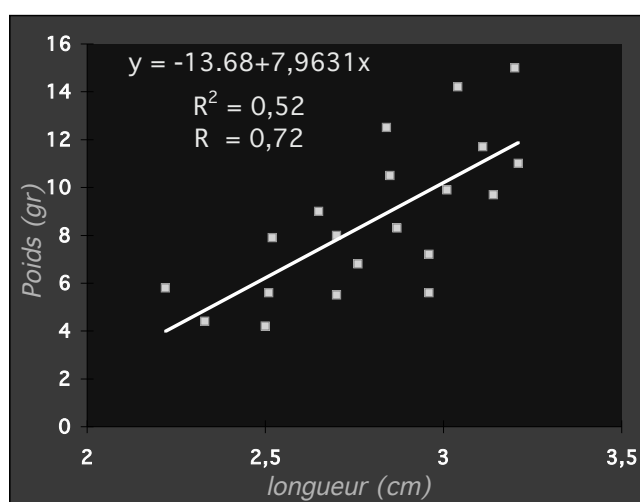


Un chimiste a mesuré, à différentes températures, la concentration de bromure de potassium (KBr) dissout dans 100 gr d'eau.

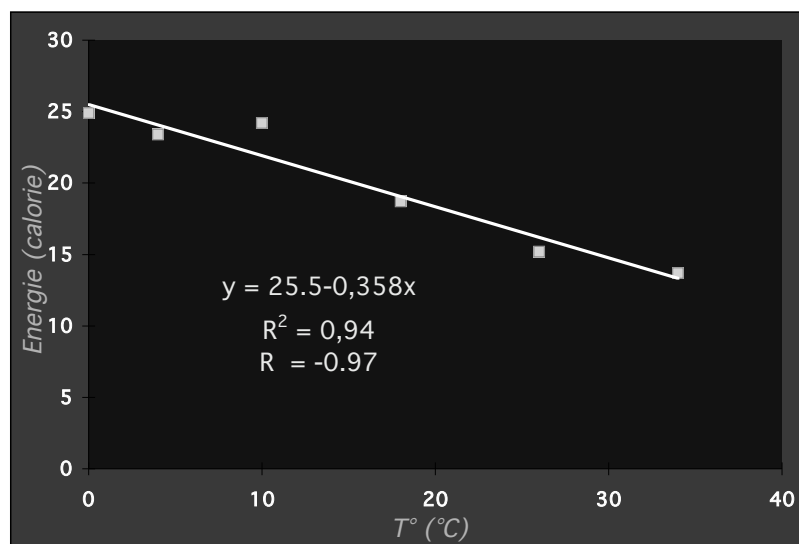


Un écologiste a mesuré la longueur et le poids de 20 coquilles vides pour une même espèce de mollusque marin.

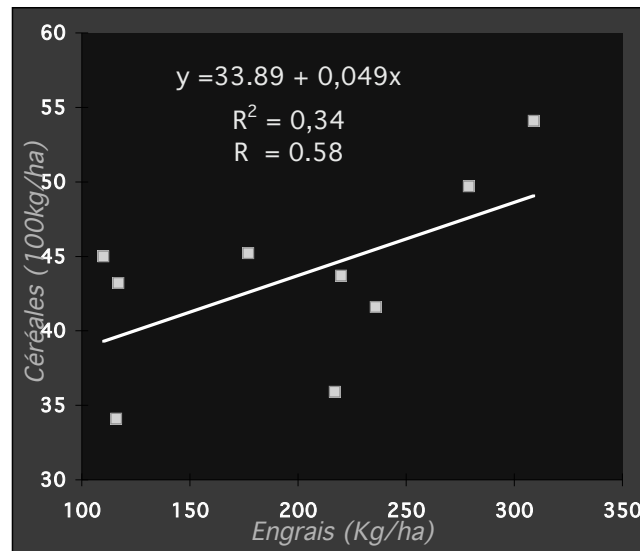
La droite des moindres rectangles serait plus adaptée.



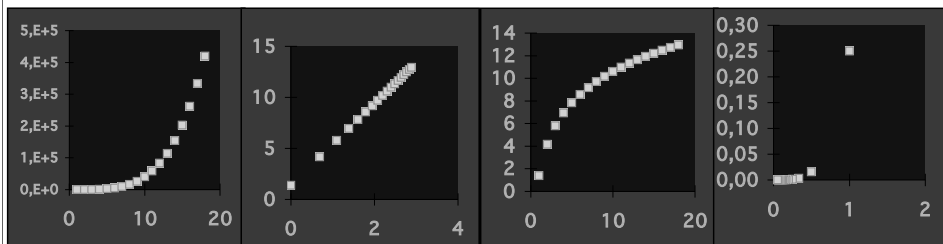
Un physiologiste a mesuré la quantité d'énergie métabolisée en 10 heures (calories) par un moineau soumis à différentes températures (° Celsius).



Les statistiques suivantes reprennent, pour 9 pays , en moyenne par hectare, les quantités d'engrais déversées et les quantités de céréales récoltées.



$$y = ax^b$$



y x

Log(y) Log(x)

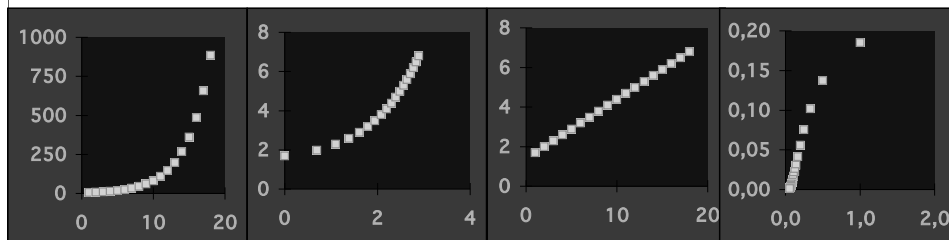
Log(y) x

1/y 1/x

$$\log(y) = \log(a) + b \log(x)$$

Fonction puissance (ex: taille, poids)

$$y = ae^{bx}$$



y x

Log(y) Log(x)

Log(y) x

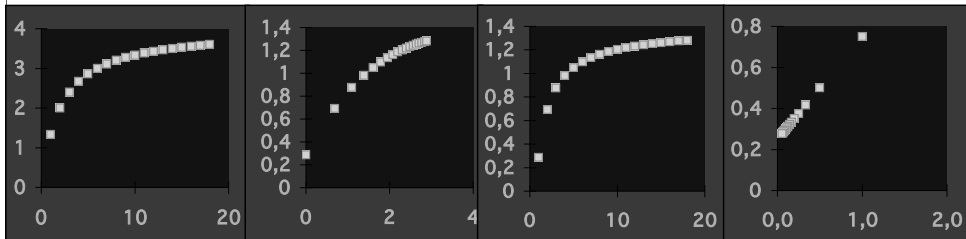
1/y

1/x

$$\log(y) = \log(a) + bx$$

Fonction exponentielle (ex: croissance de population)

$$y = \frac{ax}{b + x}$$



y x

Log(y) Log(x)

Log(y) x

1/y 1/x

$$\frac{1}{y} = \frac{1}{a} + \frac{b}{a} \frac{1}{x}$$

Fonction hyperbolique (ex: cinétique enzymatique)