

Ch 7 : Séries statistiques à deux variables quantitatives

I. Définition et représentation

1) Définition :

Définition : Une **série statistique à deux variables** est une série statistique étudiant simultanément deux caractères sur un même échantillon de n individus extraits d'une population.

On peut représenter une série statistique à deux variables à l'aide d'un tableau de la forme :

Valeur du 1 ^{er} caractère	x_1	x_2	...	x_n
Valeur du 2 nd caractère	y_1	y_2	...	y_n

Exemple :

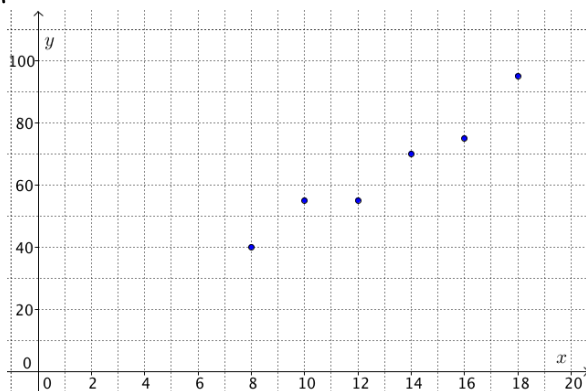
Le tableau suivant présente l'évolution du budget publicitaire et du chiffre d'affaire d'une société au cours des 6 dernières années :

Budget publicitaire en milliers d'euros x_i	8	10	12	14	16	18
Chiffre d'affaire en milliers d'euros y_i	40	55	55	70	75	95

2) Nuage de point :

Définition : Dans un repère du plan, l'ensemble des points M_i de coordonnées $(x_i; y_i)$, est appelé le **nuage de points** associé à la série statistiques $\{(x_1; y_1), (x_2; y_2), \dots, (x_n; y_n)\}$ à deux variables. Il représente la série statistique à deux variables.

Exemple :



On place les points de coordonnées $(x_i; y_i)$ pour obtenir le nuage de point de la série statistique précédente.

Définition : Soit \bar{x} et \bar{y} les moyennes des valeurs prises respectivement par chacun des deux caractères de la série.

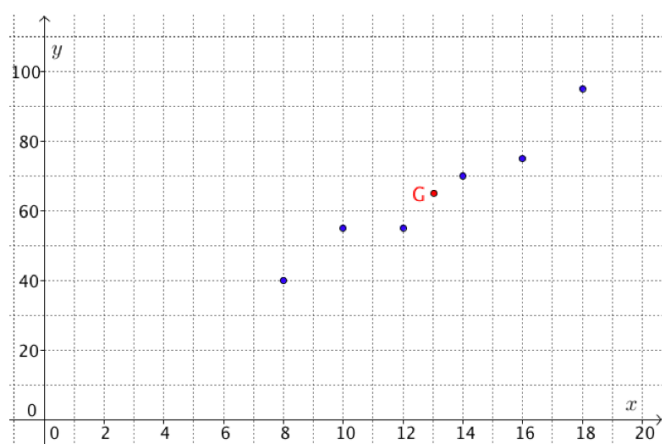
On appelle **point moyen du nuage de points** de la série statistique, le point G de coordonnées $(\bar{x} ; \bar{y})$.

Example :

$$\bar{x} = \frac{8+10+12+14+16+18}{6} = 13$$

$$\bar{y} = \frac{40+55+55+70+75+95}{6} = 65$$

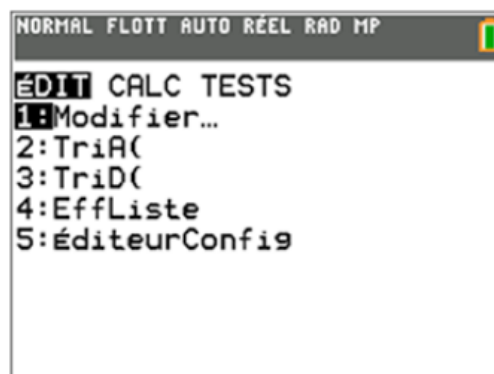
Le point moyen G du nuage de points a pour coordonnées $(13 ; 65)$. On peut placer ce point dans le repère.



Avec la calculatrice :

1- Entrer ces données dans la calculatrice.

Pour entrer les données dans sa TI-83 Premium CE on appuie sur 

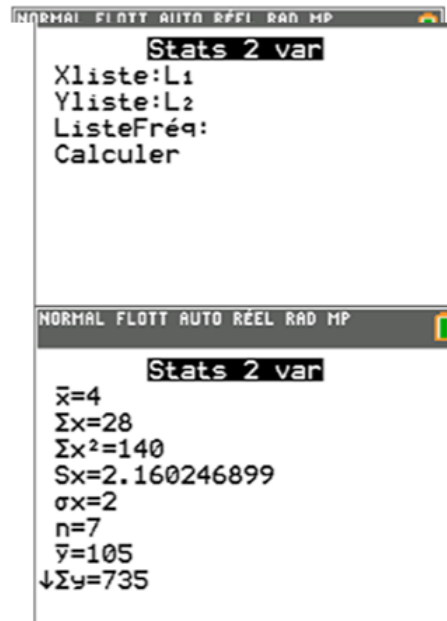


Puis on choisit **Modifier**, et on entre les données :

L1	L2	L3	L4	L5
1	87	----	-----	-----
2	92			
3	98			
4	103			
5	114			
6	119			
7	122			
-----	-----			

2°) Calculer les coordonnées de $G(\bar{x}, \bar{y})$ avec \bar{x} la moyenne des rangs et \bar{y} la moyenne des chiffres d'affaire.



On complète la boîte de dialogue et on valide :

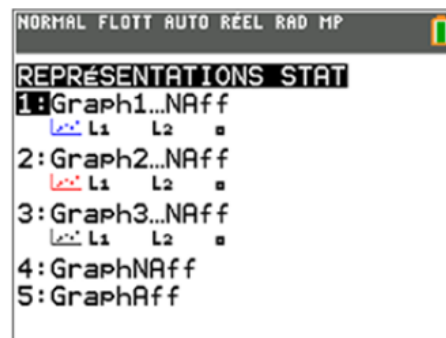


On obtient ainsi

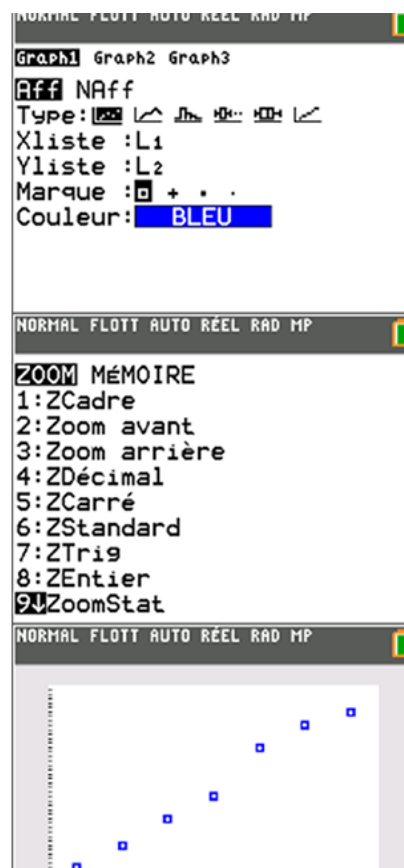
Donc $\bar{x}=4$ et $\bar{y}=105$, soit $G(4;105)$.



3°) Représenter graphiquement le nuage de points de cette série statistique.

Pour représenter graphiquement ce nuage de point sur sa TI-83 Premium CE, on appuie sur   et on sélectionne le premier graphe :



Puis on paramètre la boîte de dialogue de la façon suivante :



Afin d'ajuster la fenêtre correctement, on appuie sur   et on choisit ZoomStat :

On trouve le nuage de points ci-contre :

II. Ajustement affine

On cherche si il existe une relation, approximative, entre les deux variables x et y d'une série statistique. L'allure du nuage de points permet d'avoir une idée de la réponse. La disposition des points peut parfois suggérer la courbe d'une fonction f . Le lien approximatif entre x et y est donc de la forme $y=f(x)$.

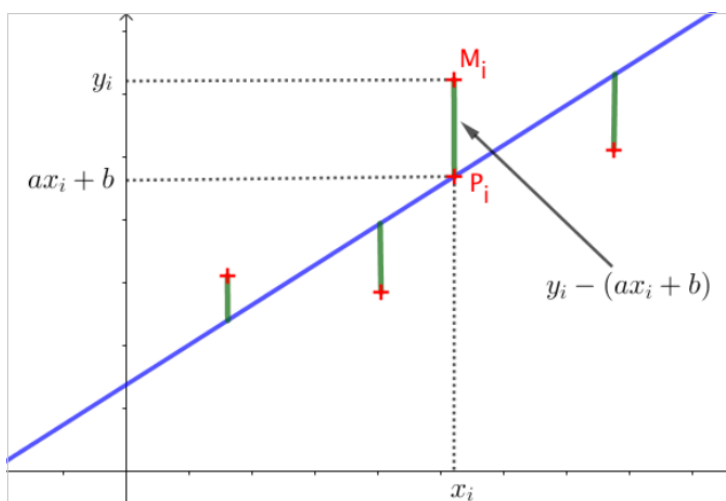
On cherche donc une courbe « passant au plus près » des points du nuage : on parle d'**ajustement** du nuage.

Dans le cas où les points du nuage sont proches de l'alignement, la courbe cherchée est une droite, représentant une fonction affine, qui réalise un **ajustement affine** du nuage.

Si la forme d'un nuage de points $M_i(x_i; y_i)$ suggère un ajustement affine, on peut trouver un grand nombre de droites ajustant ce nuage.


La **méthode des moindres carrés** consiste à déterminer la droite rendant la somme $M_1P_1^2 + M_2P_2^2 + \dots + M_nP_n^2$ la plus petite possible.

On admet que cette droite existe et qu'elle est unique.



Définition : La droite obtenue par la méthode des moindres carrés est appelée la **droite d'ajustement de y en x par la méthode des moindres carrés**.

Avec la calculatrice : Calculer l'équation de la droite d'ajustement de x en y :

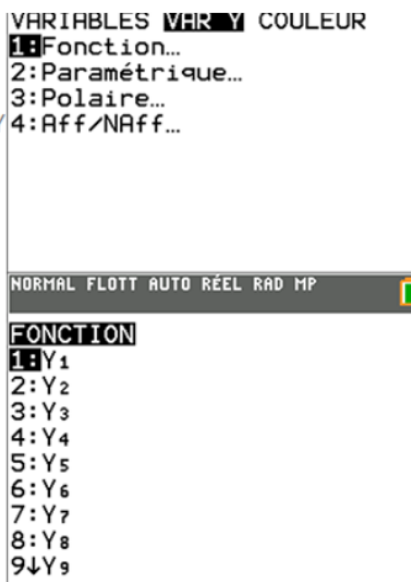
Pour trouver l'équation de la droite d'ajustement affine on appuie sur  puis on choisit l'onglet CALC.

Puis on sélectionne RegLin(ax+b) et on complète la boîte de dialogue :



Sur la ligne Enr regeQ , afin d'enregistrer le résultat dans une fonction (qu'on va tracer par la suite) on entre Y_1 accessible en appuyant sur **var** et choisir l'onglet VAR Y puis Fonction.

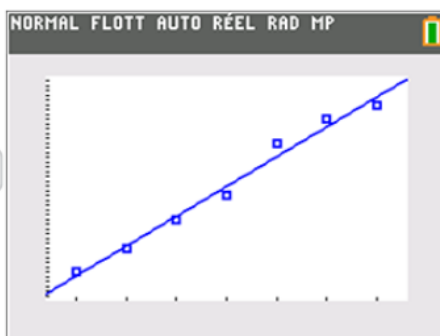
On choisit ici la première fonction Y_1 .



On obtient l'écran suivant :



Il suffit d'appuyer sur la touche **graphe**



Propriété : La droite d'ajustement de y en x par la méthode des moindres carrés passe par le point moyen du nuage de points de la série statistique.

Remarques :

- on peut obtenir une droite d'ajustement en utilisant d'autres méthodes
- Un ajustement affine permet d'obtenir des approximations pour des valeurs inconnues de cette série : par interpolation (le calcul est réalisé dans le domaine d'étude fourni par les valeurs de la série) ou par extrapolation (le calcul est réalisé en dehors du domaine d'étude). Des exemples seront traités dans les exercices.