

Statistiques à deux variables

1 Nuage de points

Définition :

Soit une série statistique à deux variables x et y chacune de n valeurs x_1, x_2, \dots, x_n et y_1, y_2, \dots, y_n .

Exemple :

x_i	0	1	2	3	4	5	6	7	8	9	10
y_i	12	12	14	15	15,5	16	17	17	19	20	22

Soit un repère orthogonal (O, I, J) , on appelle nuage de points l'ensemble des n points (non reliés) M_i de coordonnées $(x_i; y_i)$ pour i entier variant de 1 à n .

Exercice :

Faire le graphique qui présente le nuage de points de l'exemple.

Définition :

Soit une série statistique à deux variables x et y chacune de n valeurs x_1, x_2, \dots, x_n et y_1, y_2, \dots, y_n et \bar{x} et \bar{y} les moyennes arithmétiques respectives des n valeurs x_i et des n valeurs y_i .

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

Exercice :

Calculer les coordonnées du point moyen G du nuage de points donné dans l'exemple.

2 Ajustement affine

Soit la série statistique à deux variables x et y qui donne l'évolution de la population française exprimée en millier d'habitants suivant le temps exprimé en années.

Évolution de la population jusqu'en 2013, France hors Mayotte

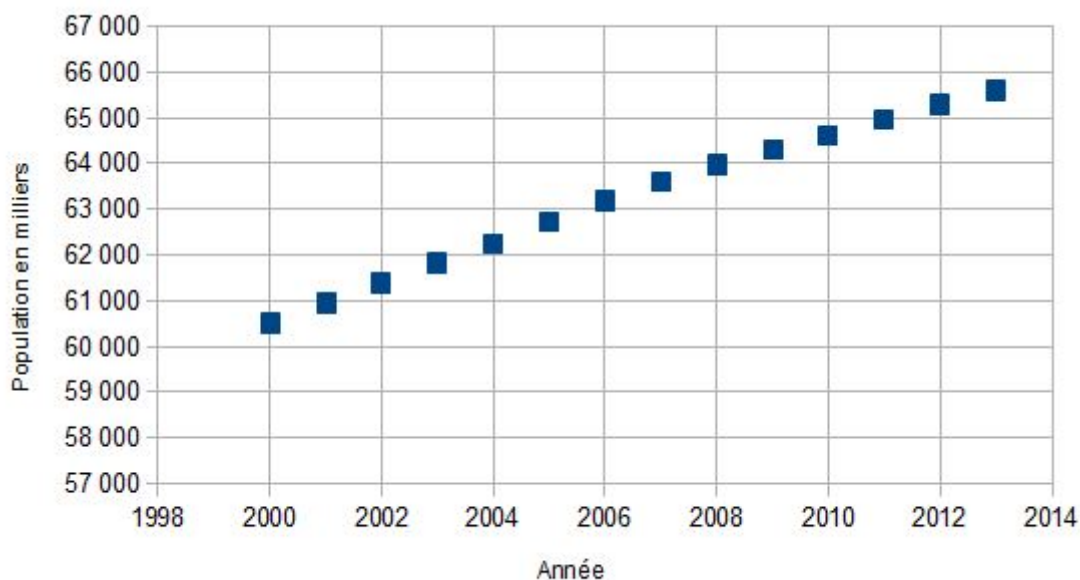
en milliers

Année	Ensemble
2000	60 508
2001	60 941
2002	61 385
2003	61 824
2004	62 251
2005	62 731
2006	63 186
2007	63 601
2008	63 962
2009	64 305
2010	64 613
2011	64 949
2012	65 281
2013	65 586

Champ : France hors Mayotte

Source : Insee, estimations de population

On a représenté le nuage de points de cette série sur le graphique suivant :



Un ajustement affine est une droite qui doit être "la plus proche" possible du nuage de points. Cette droite représentera le nuage avec une erreur (la somme des distances de la droite aux points), cette erreur étant la plus faible possible. On pourra ainsi se servir des données

statistiques pour prévoir l'avenir, dans notre exemple, il s'agit de prévoir la population française dans quelques années à venir.

Attention, on établit un ajustement affine si le nuage de points est assez "allongé" et on établit un lien statistique avec des variables qui semblent avoir une interaction commune.

La calculatrice permet d'obtenir un ajustement affine par la méthode dite des moindres carrés.

Exercice :

1. Donner l'équation de la droite obtenue par la calculatrice (c'est à dire par la méthode des moindres carrés) (voir protocole d'utilisation de la calculatrice) ($y = mx + p$, arrondir m à 10^{-4} et p à 10^{-3}).
2. Tracer cette droite sur le graphique précédent.
3. Vérifier que cette droite passe par le point G , point moyen du nuage (arrondir les coordonnées de G à 10^{-1}).
4. prévoir la population en 2015.