

OBJECTIFS

- Identifier deux variables quantitatives liées aux mêmes individus.
- Représenter une série par un nuage de points.
- Lire une tendance et repérer une valeur atypique.
- Calculer et placer le point moyen.
- Exploiter un tableur ou un fichier de données.
- Interpréter sans confondre liaison et causalité.

ACTIVITE DE DEPART

Une entreprise relève, pour six campagnes, la dépense publicitaire x , en milliers d'euros, et le chiffre d'affaires y , en milliers d'euros :

x	4	6	7	9	10	12
y	52	58	61	70	73	80

1. Chaque colonne décrit-elle le même individu ?
2. Comment représenter simultanément x et y ?
3. Quelle tendance semble apparaître ?

1. Série statistique à deux variables quantitatives

Définition (Série à deux variables) On observe sur les mêmes individus deux caractères quantitatifs :

$$x_1, \dots, x_n \quad \text{et} \quad y_1, \dots, y_n.$$

À chaque individu correspond un couple

$$(x_i; y_i).$$

L'ensemble de ces couples forme une **série statistique à deux variables quantitatives**.

Exemple Dans l'activité, chaque individu est une campagne publicitaire. Pour la quatrième campagne, le couple est

$$(9; 70).$$

Il signifie : 9 milliers d'euros dépensés et 70 milliers d'euros de chiffre d'affaires.

Méthode Avant toute représentation :

1. préciser les individus ;
2. nommer les deux variables ;
3. indiquer leurs unités ;
4. vérifier que les valeurs sont appariées dans le même ordre.

Campagne	x	y
1	4	52
2	6	58
3	7	61
4	9	70
5	10	73
6	12	80

Vigilance. Il est interdit de réordonner une seule ligne du tableau : cela détruirait l'association entre les deux valeurs d'un même individu.

Propriété (Rôle des variables)

Dans un graphique, on place généralement :

- la variable explicative ou chronologique en abscisse ;
- la variable étudiée en ordonnée.

Ce choix dépend du contexte et doit être annoncé.

⇒ TD S22 – Partie A : individus, variables, unités et couples de données.

2. Représenter un nuage de points

Définition (Nuage de points) Dans un repère, on place les points

$$M_i(x_i; y_i)$$

associés à tous les individus. L'ensemble de ces points est appelé **nuage de points**.

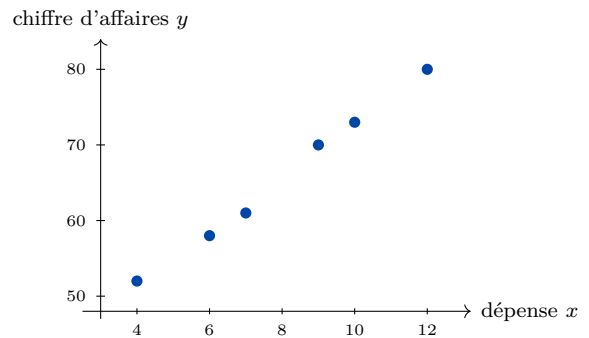
Méthode Pour construire un nuage :

1. choisir les axes et écrire les unités ;
2. choisir des échelles adaptées aux données ;
3. placer chaque couple $(x_i ; y_i)$;
4. ne pas relier les points ;
5. ajouter un titre ou une légende.

Propriété (Lecture qualitative) Un

nuage peut suggérer :

- une tendance croissante ;
- une tendance décroissante ;
- l'absence de tendance nette ;
- un ou plusieurs points atypiques.



Exemple Le nuage monte globalement de la gauche vers la droite. On observe une tendance croissante : les campagnes aux dépenses les plus élevées sont ici associées aux chiffres d'affaires les plus élevés.

Vigilance. Une tendance ne prouve pas que l'une des variables cause l'autre. D'autres facteurs peuvent intervenir.

⇒ TD S22 – Partie B : construction, lecture et critique de nuages de points.

3. Le point moyen

Définition (Point moyen) On note

$$\bar{x} = \frac{x_1 + \dots + x_n}{n} \quad \text{et} \quad \bar{y} = \frac{y_1 + \dots + y_n}{n}.$$

Le point

$$G(\bar{x}; \bar{y})$$

est appelé **point moyen** du nuage.

Méthode Pour calculer G :

1. calculer la moyenne des abscisses ;
2. calculer la moyenne des ordonnées ;
3. écrire le couple $G(\bar{x}; \bar{y})$;
4. conserver les unités de chaque coordonnée ;
5. placer G dans le même repère que le nuage.

Exemple Pour l'activité :

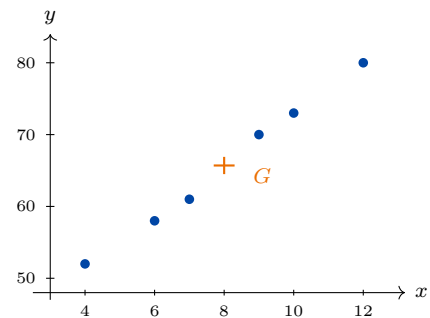
$$\bar{x} = \frac{4 + 6 + 7 + 9 + 10 + 12}{6} = 8$$

et

$$\bar{y} = \frac{52 + 58 + 61 + 70 + 73 + 80}{6} \approx 65,7.$$

Ainsi

$$G(8; 65,7).$$



Propriété (Interprétation) Les coordonnées de G sont les valeurs moyennes des deux variables. Le point moyen résume la position centrale du nuage, mais il ne décrit ni sa dispersion ni sa forme.

Vigilance. Le point moyen n'est pas nécessairement l'un des points observés.

⇒ TD S22 – Partie B : calcul et placement du point moyen.

4. Tableur, Python et données réelles

Méthode Avec un tableur. Si les valeurs de x sont dans A2 :A7 et celles de y dans B2 :B7 :

=MOYENNE(A2:A7) =MOYENNE(B2:B7)

donnent les coordonnées de G . On peut ensuite insérer un graphique de type **nuage de points**.

Propriété (Contrôles utiles)

- vérifier que les deux colonnes ont le même nombre de valeurs ;
- repérer les cellules vides ou les unités différentes ;
- contrôler quelques points du graphique avec le tableau ;
- distinguer les données observées d’une éventuelle valeur calculée.

Méthode Avec Python.

```
x = [4, 6, 7, 9, 10, 12]
y = [52, 58, 61, 70, 73, 80]
xbar = sum(x)/len(x)
ybar = sum(y)/len(y)
print(xbar, ybar)
```

Méthode Pour un fichier de données :

1. identifier les colonnes utiles ;
2. conserver uniquement les lignes complètes ;
3. extraire les deux listes ;
4. calculer, représenter puis interpréter.

Propriété (Esprit critique) Un nuage peut suggérer une liaison entre deux variables. La qualité de cette lecture dépend du nombre d’individus, de la dispersion, des valeurs atypiques et du domaine étudié. L’ajustement affine et les prévisions seront étudiés en N16.

A RETENIR

- Une série à deux variables quantitatives associe un couple $(x_i; y_i)$ à chaque individu.
- Le nuage de points représente les couples $M_i(x_i; y_i)$.
- On ne relie pas les points d’un nuage.
- Le point moyen est $G(\bar{x}; \bar{y})$.
- Une tendance peut être croissante, décroissante ou peu marquée.
- Une liaison observée ne prouve pas une causalité.

APPLICATIONS IMMEDIATES

1. Donner les individus d’un tableau « ancienneté / salaire ».
2. Placer le point (6 ; 58).
3. Calculer la moyenne de 2, 4, 6, 8.
4. Le point moyen est-il forcément observé ?
5. Décrire une tendance croissante.
6. Citer deux contrôles avant de tracer un nuage.
7. Expliquer pourquoi corrélation ne signifie pas causalité.

⇒ TD S22 : données, nuages, point moyen et interprétation critique.