

Indicateurs de position

$$\bar{x} = \frac{1}{N} \sum n_i x_i = \sum f_i x_i \cdot \text{moyenne}$$

- Médiane Me · 50 % au-dessus / au-dessous · robuste aux extrêmes
- Mode · valeur la plus fréquente
- Quartiles Q_1, Q_2, Q_3 · partagent en 4 parts égales
- Déciles $D_1 \dots D_9$ · en 10 parts

Si $\bar{x} > Me$ · distribution **asymétrique à droite** (quelques grandes valeurs tirent la moyenne).

Salaires · préférer la **médiane** · les très hauts revenus tirent la moyenne.

Indicateurs de dispersion

Étendue = $x_{\max} - x_{\min}$ · sensible aux extrêmes
IQR = $Q_3 - Q_1$ · écart interquartile · 50 % centraux

$$\sigma^2 = \frac{1}{N} \sum n_i (x_i - \bar{x})^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma = \sqrt{\sigma^2} \cdot \text{écart-type} \cdot \text{même unité que } x$$

σ faible · données concentrées · σ élevé · forte dispersion.

Évolutions · taux & coefficients

$$t = \frac{V_A - V_D}{V_D} \times 100 \cdot \text{taux d'évolution (\%)}$$

$$CM = \frac{V_A}{V_D} = 1 + \frac{t}{100} \cdot \text{coefficient multiplicateur}$$

$$TVAM = \left(\frac{V_A}{V_D} \right)^{1/n} - 1 \cdot \text{annuel moyen sur } n \text{ années}$$

Élasticité · $e = \frac{\Delta Y/Y}{\Delta X/X}$ · sensibilité de Y à X (en %).

Indice base 100 · $I_t = \frac{V_t}{V_0} \times 100$ · facilite la comparaison.

Piège · ne pas additionner des taux successifs · les multiplier (via coefficients).

Mémo · % variation

t (%)	CM
+10 %	× 1,10
+25 %	× 1,25
+50 %	× 1,50
+100 %	× 2
-10 %	× 0,90
-50 %	× 0,50

Niveau ≠ croissance · une variable à -3% reste positive si son niveau est élevé.

Représentations graphiques

variable	Graphique
qualitative · discrète	diagramme en barres
continue (classes)	histogramme
évolution temporelle	courbe
résumé 5 nombres	boîte à moustaches

Histogramme · c'est l'aire qui compte, pas la hauteur · si classes inégales, calculer la **densité** $d_i = f_i/a_i$.

Boxplot · min · Q_1 · Me · Q_3 · max · outliers au-delà de $1,5 \times IQR$.

Corrélation & régression

$$r = \frac{Cov(x,y)}{\sigma_x \sigma_y} \cdot r \in [-1; 1]$$

- $|r|$ proche de 1 · forte corrélation linéaire
- $|r|$ proche de 0 · pas de corrélation linéaire
- signe de r · sens (positif ou négatif)

Droite des moindres carrés · $y = ax + b$

$$a = \frac{Cov(x,y)}{\sigma_x^2} \cdot b = \bar{y} - a\bar{x}$$

Corrélation ≠ causalité · variable cachée possible (glaces ↔ noyades · chaleur).

Lecture critique · erreurs fréquentes

Erreur	Conséquence	Remède
Confondre corrélation et causalité	fausse explication	chercher la variable cachée
Extrapoler hors plage	prédiction absurde	rester dans la zone observée
Comparer effectifs bruts	biais sur populations	passer aux fréquences
Ignorer la dispersion	moyennes trompeuses	donner σ ou IQR
Biais de sélection	échantillon non représentatif	vérifier la méthode
Paradoxe de Simpson	tendance qui s'inverse	stratifier par sous-groupes