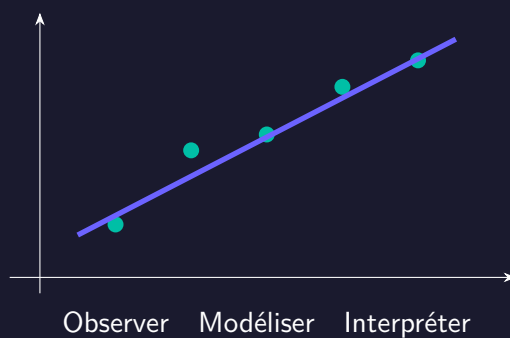


Statistiques à deux variables

Statistiques ■ Comprendre et s'entraîner



Première STMG et STI2D — commun
Année scolaire 2026–2027

 Cours

 15 exercices

 Problème guidé

 Corrigés

Table des matières

1	Pourquoi ce chapitre ?	3
2	L'idée avant la formule	3
3	Le cours formel	3
3.1	Représenter deux variables par un nuage	3
3.2	Le point moyen du nuage	4
3.3	Construire un ajustement affine	6
3.4	Interpolation et extrapolation	7
3.5	Association, causalité et qualité d'ajustement	8
4	Boîte à outils	9
5	Exercices progressifs	11
6	Problème — Pour aller plus loin	13
7	Indices des exercices	14
8	Corrigés détaillés	16
9	Corrigé du problème	22
10	Variante autonome et bilan	24
11	Fiche-mémoire	24
12	Indices des pauses et des preuves	25
13	Explications des pauses et des preuves	27

1 Pourquoi ce chapitre ?

Relier deux grandeurs aide à décrire une tendance et à préparer une décision. Tu vas construire un modèle, calculer ses prévisions et en discuter les limites. **Ton parcours.** Enseignement commun de mathématiques pour les séries STMG et STI2D. Les situations de gestion et les situations techniques mobilisent ici les mêmes notions du programme.

- Représenter deux variables par un nuage.
- Le point moyen du nuage.
- Construire un ajustement affine.
- Interpolation et extrapolation.
- Association, causalité et qualité d'ajustement.

2 L'idée avant la formule

Un nuage conserve les couples observés. La droite résume une tendance ; elle ne prouve ni une loi universelle ni une causalité. Nomme les grandeurs et leurs unités avant tout calcul.

3 Le cours formel

3.1 Représenter deux variables par un nuage

Intuition | Construire la notion pas à pas

Conserver l'association entre deux mesures. Une série à deux variables relève deux informations sur chaque individu : par exemple sa durée de travail et sa production. Le couple $(x_i; y_i)$ doit rester associé. Trier les deux colonnes séparément créerait des couples qui n'ont jamais été observés.

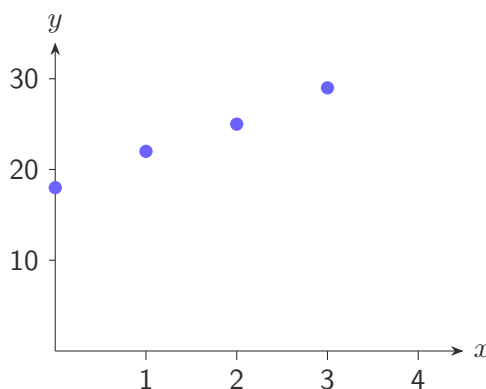
Lire un point dans son contexte. La première coordonnée se lit sur l'axe horizontal et la seconde sur l'axe vertical. On indique le nom et l'unité de chaque variable. Le point $(3; 42)$ peut ainsi signifier « après 3 heures, 42 pièces », ce qui est différent de 42 heures et 3 pièces.

Observer avant de calculer. Le nuage montre une tendance générale, une éventuelle courbure, des regroupements ou des observations atypiques. Une tendance croissante n'exige pas que tous les points montent un à un : les mesures peuvent fluctuer autour de cette tendance.

Évaluer une droite avec les bons écarts. Pour des pas égaux en x , un alignement affine exact exige des différences égales en y . Si les pas en x changent, on compare les quotients $\Delta y / \Delta x$. Une apparence graphique dépend aussi des échelles ; on complète donc l'observation par quelques calculs.

Définition | Comprendre la notion

Une série à deux variables quantitatives contient des couples $(x_i; y_i)$ observés sur les mêmes individus. Un nuage de points place chaque couple dans un repère. Annonce les variables, leurs unités et l'échelle des axes. Invertir les colonnes modifie la question étudiée : prédire y à partir de x n'est pas la même opération que prédire x à partir de y . Le nuage permet d'observer une tendance, une courbure, des groupes ou des valeurs atypiques. Une tendance croissante ne signifie pas que chaque point est au-dessus du précédent ni qu'une relation de cause à effet est démontrée.



💡 Exemple | Exemple commenté

On relève les couples durée/température $(0; 18)$, $(1; 22)$, $(2; 25)$, $(3; 29)$. Le temps en minutes est placé horizontalement et la température en degrés Celsius verticalement. Le point $(2; 25)$ signifie qu'au bout de deux minutes, la température mesurée est 25 degrés.

Pour les trois pas d'une minute, les températures augmentent de 4, puis 3, puis 4 degrés. La tendance est croissante et proche d'une augmentation régulière, mais les quatre points ne sont pas exactement alignés puisque ces accroissements diffèrent.

Une droite pourrait fournir une approximation à examiner. Elle ne doit pas effacer la distinction entre valeurs mesurées et valeurs prévues. Le graphique seul ne prouve pas non plus ce qui cause l'évolution : cette interprétation dépend du protocole et du contexte.

⚠️ Attention | Comprendre une erreur fréquente

Une observation atypique peut être vraie. Avant d'utiliser le mot « erreur », distingue une erreur de saisie vérifiée d'une mesure inhabituelle mais réelle.

À toi d'essayer : Pause 1 — Mets la notion à l'épreuve

Un nuage à deux variables conserve...

1. Les valeurs de x triées indépendamment de y
2. Seulement les deux moyennes
3. Les couples d'observations associés

Tu peux chercher, prendre un indice ou lire l'explication, à ton rythme.

[Indices, p. 25](#) [Explication, p. 27](#)

3.2 Le point moyen du nuage

🧠 Intuition | Construire la notion pas à pas

Résumer séparément les deux coordonnées. Le point moyen G a pour abscisse la moyenne des x et pour ordonnée la moyenne des y . On effectue donc deux moyennes de n valeurs, pas une seule moyenne de tous les nombres mélangés.

Respecter les unités. Si x est un temps et y une température, $\bar{x}$ reste un temps et $\bar{y}$ une température. Additionner directement un temps et une température ne produit pas une coordonnée interprétable.

Utiliser G comme n'importe quel point. Pour qu'une droite $y = ax + b$ passe par G, les coordonnées doivent satisfaire $\bar{y} = a\bar{x} + b$. Si la pente a est connue, on isole $b = \bar{y} - a\bar{x}$.

Connaître les limites de ce résumé. Le point moyen situe le centre du nuage, mais ne donne ni sa dispersion ni sa tendance. Deux nuages opposés peuvent avoir le même centre. La droite de régression affine des moindres carrés avec constante passe par G ; cette propriété ne vaut pas automatiquement pour une droite choisie à l'œil.

Définition | Comprendre la notion

Pour n couples, le point moyen est $G(\bar{x}; \bar{y})$, avec $\bar{x} = (x_1 + \dots + x_n)/n$ et $\bar{y} = (y_1 + \dots + y_n)/n$. Calcule séparément les moyennes de chaque coordonnée, en gardant leurs unités. Le point moyen résume le centre du nuage, mais ne suffit pas à décrire sa forme. La droite de régression affine des moindres carrés avec constante passe par G. Une droite choisie au jugé ne possède pas automatiquement cette propriété. Si une droite de pente a doit passer par G, son ordonnée à l'origine est $b = \bar{y} - a\bar{x}$.

Exemple | Exemple commenté

Pour $(1; 4), (2; 7), (3; 7)$, on calcule $\bar{x} = (1 + 2 + 3)/3 = 2$ et $\bar{y} = (4 + 7 + 7)/3 = 6$. Le point moyen est $G(2; 6)$. Il n'est pas nécessaire qu'un point observé possède exactement ces coordonnées.

Une droite de pente 1,5 qui passe par G s'écrit $y = 1,5x + b$. En remplaçant les coordonnées, on obtient $6 = 1,5 \times 2 + b = 3 + b$, donc $b = 3$. La droite est $y = 1,5x + 3$.

Le contrôle en x égal à 2 donne $1,5 \times 2 + 3 = 6$. Ce contrôle établit le passage par G, mais ne suffit pas à prouver que la pente 1,5 est celle d'une régression : elle était imposée ici par l'énoncé.

Attention | Comprendre une erreur fréquente

Le point moyen n'est pas obtenu en faisant une moyenne des six nombres d'un tableau de trois couples. Les abscisses et les ordonnées doivent conserver leurs rôles et leurs unités.

À toi d'essayer : Pause 2 — Mets la notion à l'épreuve

Le point moyen de (0; 2) et (4; 8) est. . .

1. (4; 10)
2. (2; 5)
3. (5; 2)

Tu peux chercher, prendre un indice ou lire l'explication, à ton rythme.

[Indices](#), p. 25 [Explication](#), p. 27

3.3 Construire un ajustement affine

Intuition | Construire la notion pas à pas

Remplacer un nuage par une tendance. Un ajustement affine utilise une droite $\hat{y} = ax + b$ pour prévoir y à partir de x . Le symbole $\hat{y}$ désigne une valeur prévue, tandis que y désigne une observation. Les deux ne coïncident pas nécessairement.

Mesurer ce que la droite manque. Le résidu est $e_i = y_i - \hat{y}_i$. Il est positif lorsque l'observation se situe au-dessus de la droite et négatif lorsqu'elle se situe en dessous. Additionner les résidus signés peut masquer des erreurs par compensation ; les carrés évitent cette compensation.

Distinguer les méthodes. La méthode de Mayer relie les points moyens de deux groupes de points ordonnés. La méthode des moindres carrés cherche la droite qui minimise la somme des carrés des résidus verticaux. Ces critères ne donnent pas nécessairement la même droite.

Évaluer le résultat au-delà de la formule. On observe le nuage avant l'ajustement et les résidus après. Une courbure régulière des résidus suggère qu'une droite est mal adaptée. Une petite erreur sur les observations ne garantit pas une bonne prévision loin de la zone étudiée.

Définition | Comprendre la notion

Un ajustement remplace un nuage imparfaitement aligné par une relation approchée $y = ax + b$. Une droite au jugé, une droite de Mayer ou une régression peuvent donner des ajustements différents. La méthode de Mayer partage les points ordonnés en deux groupes et relie leurs points moyens. La méthode des moindres carrés choisit a, b pour minimiser la somme des carrés des écarts verticaux $\sum (y_i - (ax_i + b))^2$; on utilise un outil numérique pour la déterminer. L'écart $y_i - \hat{y}_i$ s'appelle un résidu. Un nuage courbé ou des résidus organisés signalent qu'une droite peut être inadéquate. Arrondis les coefficients seulement à la précision utile.

Exemple | Exemple commenté

Pour les points (0; 1), (1; 3), (2; 4), une régression affine donne $\hat{y} = 1,5x + 7/6$. Les prévisions sont $7/6$, $8/3$ et $25/6$. Les résidus observation moins prévision sont donc $-1/6$, $1/3$ et $-1/6$.

Leur somme est nulle, mais les erreurs ne sont pas toutes nulles. La somme des carrés vaut $1/36 + 1/9 + 1/36 = 1/6$. Ce calcul rend visible une dispersion que la somme signée annulerait.

Le point moyen est $G(1; 8/3)$. La droite passe bien par lui puisque $1,5 + 7/6 = 8/3$. Pour un affichage, on peut arrondir la constante à 1,1667 ; pour contrôler les résidus, garder $7/6$ évite les petites erreurs d'arrondi.

⚠ Attention | Comprendre une erreur fréquente

Un résidu négatif n'est pas une erreur « plus petite que zéro » à annuler. Il indique le côté de la droite où se trouve l'observation ; son carré contribue positivement au critère d'ajustement.

À toi d'essayer : Pause 3 — Mets la notion à l'épreuve

Les moindres carrés minimisent. . .

1. La somme des carrés des résidus verticaux
2. La somme des abscisses
3. Le nombre de points du nuage

Tu peux chercher, prendre un indice ou lire l'explication, à ton rythme.

[Indices, p. 25](#) [Explication, p. 27](#)

3.4 Interpolation et extrapolation**🧠 Intuition | Construire la notion pas à pas**

Situer la demande par rapport aux observations. Une interpolation utilise le modèle à l'intérieur de l'intervalle des abscisses observées. Une extrapolation le prolonge à l'extérieur. Dans les deux cas, le résultat reste une estimation ; l'extrapolation ajoute l'hypothèse que la tendance continue hors de la zone connue.

Décoder les dates et les unités. Le rang x peut être le nombre d'années après une date de référence, et non l'année du calendrier. Si 2022 porte le rang zéro, 2026 porte le rang 4. Remplacer x par 2026 produirait une erreur de sens, même si le calcul était arithmétiquement exact.

Savoir quelle variable on cherche. Pour prévoir y à partir d'une valeur de x , on remplace x dans la formule. Pour trouver quand un objectif y sera atteint, on résout une équation en x , puis on reconvertit le rang en date si nécessaire.

Appliquer les limites du phénomène. Un effectif ne peut pas être négatif, une proportion reste entre zéro et un, une capacité peut plafonner. Une prévision qui viole ces limites signale que le modèle ne peut pas être prolongé jusque-là. Une valeur admissible ne suffit toutefois pas à garantir sa fiabilité.

📖 Définition | Comprendre la notion

Une interpolation estime une valeur pour une abscisse à l'intérieur de l'intervalle observé. Une extrapolation utilise le modèle en dehors de cet intervalle ; elle demande davantage de prudence. Pour prévoir y , remplace x dans le modèle. Pour chercher la valeur de x correspondant à un objectif, résous l'équation du modèle. Si x est une date codée, reviens ensuite à la date réelle. Même à l'intérieur de l'intervalle, une prévision n'est pas une certitude individuelle. Vérifie son ordre de grandeur, son unité et la pertinence du modèle ; évite les extrapolations incompatibles avec une limite physique ou un marché saturé.

 Exemple | Exemple commenté

Les années 2020 à 2024 sont codées de x égal à 0 à x égal à 4, avec le modèle $y = 12x + 80$. Pour 2023, le rang est $2023 - 2020 = 3$; la prévision vaut $12 \times 3 + 80 = 116$. Comme 3 appartient à $[0; 4]$, il s'agit d'une interpolation.

Pour 2027, le rang est 7 et la prévision vaut $12 \times 7 + 80 = 164$. Le rang 7 est hors de l'intervalle observé : c'est une extrapolation. Le calcul utilise la même formule, mais la justification est plus fragile puisque l'on suppose la poursuite de la tendance trois ans après la dernière observation.

Pour un objectif de 140, l'équation $12x + 80 = 140$ donne $x = 5$, donc l'année 2025 selon le modèle. Ce résultat constitue lui aussi une extrapolation.

 Attention | Comprendre une erreur fréquente

Une formule peut fournir un nombre sans que ce nombre ait un sens dans le contexte. Vérifier le domaine observé et les limites physiques fait partie de la réponse, pas seulement du commentaire final.

À toi d'essayer : Pause 4 — Mets la notion à l'épreuve

Prédire hors de l'intervalle observé est...

1. Une interpolation
2. Une mesure exacte
3. Une extrapolation

Tu peux chercher, prendre un indice ou lire l'explication, à ton rythme.

[Indices, p. 25](#) [Explication, p. 27](#)

3.5 Association, causalité et qualité d'ajustement

 Intuition | Construire la notion pas à pas

Distinguer trois questions. Un ajustement peut bien reproduire les données observées, servir à une prévision et proposer une explication du phénomène. Ces trois qualités ne se déduisent pas automatiquement l'une de l'autre. Une droite proche des points ne prouve pas qu'une variable agit sur l'autre.

Comprendre ce que mesure R^2 . Pour une régression affine avec constante et des ordonnées non toutes identiques, R^2 compare l'erreur restante à la variation initiale des ordonnées. Une valeur proche de 1 indique que le modèle reproduit une grande part de cette variation sur les données utilisées. Ce n'est pas la probabilité que la prévision soit correcte.

Chercher d'autres explications. Une association peut provenir d'une cause commune, d'une sélection des individus ou d'une influence dans l'autre sens. Pour soutenir une conclusion causale, il faut un protocole et des informations adaptés, au-delà d'un nuage de points.

Examiner la quantité et la qualité des données. Deux points peuvent être parfaitement alignés sans apporter beaucoup d'information sur la suite du phénomène. Un point extrême peut aussi influencer fortement une droite. Le nombre d'observations, leurs conditions de recueil et les résidus complètent donc l'indicateur numérique.

Définition | Comprendre la notion

Deux variables peuvent évoluer ensemble sans que l'une cause directement l'autre. Une troisième variable, un effet de sélection ou une coïncidence peuvent expliquer l'association. Pour une régression affine avec constante, le coefficient de détermination R^2 mesure la part de variation des ordonnées expliquée par le modèle sur les données. Une valeur proche de 1 traduit un bon ajustement au sens de ce critère, mais ne prouve ni causalité ni validité future. Regarde toujours le nuage, le nombre d'observations, les résidus et les conditions de recueil. Un point extrême peut influencer fortement une régression. Des populations différentes peuvent masquer des relations différentes.

Exemple | Exemple commenté

Pendant les périodes chaudes, les ventes de glaces et le nombre de noyades peuvent tous deux augmenter. Un nuage croissant montrerait une association entre ces deux variables. Il ne démontrerait pas que manger une glace provoque une noyade.

Une explication possible est la saison : la chaleur favorise l'achat de glaces et la fréquentation des lieux de baignade. Cette variable commune peut faire évoluer les deux quantités ensemble. Pour analyser un effet propre, il faudrait examiner notamment l'exposition à la baignade et les autres différences entre périodes.

Même un ajustement très précis de ce nuage ne supprimerait pas ce problème d'interprétation. La qualité de reproduction des observations et la preuve d'une cause sont deux questions distinctes.

Attention | Comprendre une erreur fréquente

$R^2 = 0,94$ ne signifie pas « 94 pour cent de chances que la prochaine valeur soit exacte ». L'indicateur décrit la variation reproduite sur une série, pas une probabilité individuelle.

À toi d'essayer : Pause 5 — Mets la notion à l'épreuve

Un R^2 proche de 1 prouve une causalité. . .

1. Oui si les variables sont positives
2. Non
3. Oui

Tu peux chercher, prendre un indice ou lire l'explication, à ton rythme.

[Indices, p. 26](#) [Explication, p. 27](#)

4 Boîte à outils

Méthode | Représenter deux variables par un nuage

1. Identifie les deux variables et garde chaque paire de mesures associée.
2. Place x horizontalement et y verticalement, avec des axes nommés et gradués.
3. Décris la tendance, la courbure, les groupes et les points atypiques.
4. Compare des taux de variation si tu recherches un alignement exact.
5. Vérifie les données atypiques sans les supprimer uniquement pour améliorer l'apparence.

 Méthode | Le point moyen du nuage

1. Compte le nombre de couples et calcule la moyenne des abscisses.
2. Calcule séparément la moyenne des ordonnées avec le même nombre d'observations.
3. Écris G avec les coordonnées dans le bon ordre et leurs unités.
4. Pour une droite passant par G, remplace x et y par ces deux moyennes.
5. Contrôle le passage par G et ne déduis pas la forme du nuage de son seul centre.

 Méthode | Construire un ajustement affine

1. Observe le nuage et précise la méthode d'ajustement utilisée.
2. Calcule ou relève les coefficients a et b avec une précision suffisante.
3. Calcule les prévisions $\hat{y}_i$, puis les résidus $y_i - \hat{y}_i$.
4. Compare les carrés des résidus et leur organisation, pas seulement leur somme signée.
5. Présente les valeurs obtenues comme des prévisions dépendant du modèle.

 Méthode | Interpolation et extrapolation

1. Relève l'intervalle observé, le codage de x et les unités.
2. Convertis la demande en abscisse et distingue interpolation et extrapolation.
3. Calcule une image ou résous l'équation correspondant à l'objectif.
4. Reconvertis le résultat en date ou en quantité et arrondis avec une précision justifiée.
5. Contrôle les limites physiques et indique la dépendance au modèle.

 Méthode | Association, causalité et qualité d'ajustement

1. Précise si la question concerne l'ajustement, la prévision ou une explication causale.
2. Compare les indicateurs uniquement sur des données et une variable comparables.
3. Examine le nuage, les résidus, les valeurs atypiques et la taille de la série.
4. Cherche une cause commune, une sélection ou une causalité inverse avant de conclure.
5. Limite la conclusion à ce que le protocole et les données permettent réellement d'affirmer.

5 Exercices progressifs

Les exercices 1 à 5 consolident les bases ; 6 à 10 demandent plusieurs étapes ; 11 à 15 t'invitent à choisir, justifier et contrôler. Cherche avant de consulter les indices.

Exercice 1 — Placer un couple ★☆☆

Un relevé indique 3 heures et 42 pièces produites. Quel point représente-t-il si le temps est en abscisse ?

Exercice 2 — Calculer G ★☆☆

Trouve le point moyen de $(1; 3)$, $(3; 5)$, $(5; 7)$.

Exercice 3 — Utiliser une droite fournie ★☆☆

Un ajustement donne $y = 2,5x + 8$. Quelle valeur prévoit-il pour $x = 6$?

Exercice 4 — Identifier la zone ★☆☆

On observe des durées entre 2 et 12 min. Prédire à 8 min est-il une interpolation ? Et à 20 min ?

Exercice 5 — Lire R carré ★☆☆

Deux ajustements affines sur les mêmes données donnent $R^2 = 0,94$ et $R^2 = 0,62$. Lequel explique davantage la variation selon ce critère ?

Exercice 6 — Identifier une tendance ★★★

Les couples sont $(1; 9)$, $(2; 7)$, $(3; 6)$, $(4; 3)$. Décris la tendance et explique pourquoi une droite exacte n'est pas évidente.

Exercice 7 — Une droite passant par G ★★★

Une droite de pente 4 passe par $G(2; 11)$. Trouve son équation.

Exercice 8 — Un ajustement de Mayer ★★★

Les deux sous-groupes ont pour points moyens $G_1(2; 7)$ et $G_2(6; 15)$. Détermine la droite de Mayer.

Exercice 9 — Une date codée ★★★

Le rang 0 correspond à 2022 et le modèle est $y = 6x + 50$. Quelle prévision donne-t-il pour 2026 ?

Exercice 10 — Une conclusion abusive ★★★

Une enquête associe possession d'un vélo et bonne santé. Peut-on en déduire que donner un vélo améliore forcément la santé ?

Exercice 11 — Un point atypique ★★★

Une série donne $(1; 12)$, $(2; 14)$, $(3; 16)$, $(4; 50)$, $(5; 20)$. Peut-on supprimer le quatrième point parce qu'il gêne l'alignement ?

Exercice 12 — Ce que le centre ne dit pas ★★★

Les nuages A : $(0; 0)$, $(2; 2)$ et B : $(0; 2)$, $(2; 0)$ ont-ils le même point moyen ? La même tendance ?

Exercice 13 — Comparer des résidus ★★ ★

Pour $(0; 1), (1; 2), (2; 5)$, compare les sommes des carrés des résidus des modèles A : $y = 2x$ et B : $y = 2x + 1$.

Exercice 14 — Un modèle qui devient impossible ★★ ★

Un ajustement $p(x) = 110 - 5x$ donne un pourcentage. Peut-on l'utiliser pour toute valeur positive de x ?

Exercice 15 — Trop peu de données ★★ ★

Deux points distincts ont un ajustement affine parfait. Est-ce une preuve solide de prévision ?

6 Problème — Pour aller plus loin

Fréquentation d'un atelier

x compte les séances et y les participants. Les couples $(x;y)$ mesurés sont : $(1;20)$, $(2;28)$, $(3;36)$, $(4;44)$.

1. Place les points et calcule leur point moyen G .
2. Détermine la droite passant par les points extrêmes et contrôle les points intermédiaires.
3. Prévois y en $x = 2,5$ et $x = 6$; distingue interpolation et extrapolation.
4. En $x = 5$, une nouvelle observation donne $y = 55$. Calcule l'écart observé moins prévu.
5. Cet écart rend-il le modèle forcément inutile ? La relation prouve-t-elle une causalité ?

7 Indices des exercices

Exercice 1

Indice 1. La première coordonnée correspond à l'axe horizontal.

Indice 2. Le temps est placé en abscisse.

Exercice 2

Indice 1. Il y a trois couples.

Indice 2. Calcule deux moyennes, pas une moyenne des six nombres.

Exercice 3

Indice 1. Remplace l'abscisse dans l'équation.

Indice 2. Distingue valeur prédite et valeur mesurée.

Exercice 4

Indice 1. Compare l'abscisse à l'intervalle observé.

Indice 2. À l'intérieur : interpolation ; à l'extérieur : extrapolation.

Exercice 5

Indice 1. Compare les valeurs sur une base de données identique.

Indice 2. Un bon ajustement observé n'est pas une garantie future.

Exercice 6

Indice 1. Compare les variations de y pour les mêmes pas de x .

Indice 2. Des points exactement alignés ont un taux de variation constant.

Exercice 7

Indice 1. Utilise la forme $y = ax + b$.

Indice 2. Remplace x et y par les coordonnées de G.

Exercice 8

Indice 1. La droite passe par les deux points moyens.

Indice 2. Calcule pente puis ordonnée à l'origine.

Exercice 9

Indice 1. Calcule l'écart à l'année de référence.

Indice 2. 2022 porte le rang 0.

Exercice 10

Indice 1. La causalité peut aller dans plusieurs sens.

Indice 2. Cherche une variable tierce ou un effet de sélection.

Exercice 11

Indice 1. Un point atypique n'est pas automatiquement faux.

Indice 2. Cherche une justification dans les données ou la mesure.

Exercice 12

Indice 1. Calcule les deux centres.

Indice 2. Observe ensuite la relation entre les coordonnées.

Exercice 13

Indice 1. Calcule observation moins prévision pour chaque point.

Indice 2. Élève les résidus au carré avant de les additionner.

Exercice 14

Indice 1. Un pourcentage a des bornes naturelles.

Indice 2. Résous séparément $p(x) \geq 0$ et $p(x) \leq 100$.

Exercice 15

Indice 1. Deux points imposent toujours une droite.

Indice 2. Un ajustement parfait peut être obtenu avec très peu d'information.

8 Corrigés détaillés

Corrigé 1

 *Démonstration / Solution expliquée*

Comprendre la question. Le temps doit être la première coordonnée puisque l'énoncé le place en abscisse.

Construire la réponse.

1. La durée est 3 heures : on place 3 sur l'axe horizontal.
2. La production associée est 42 pièces : on place 42 sur l'axe vertical.
3. Le point qui réunit ces deux informations est (3; 42).

Vérifier. Lire le point dans l'autre sens donnerait 42 heures et 3 pièces, ce qui ne correspond pas au relevé. Les unités permettent de repérer cette inversion.

Conclure. Le relevé est représenté par le point (3; 42).

Corrigé 2

 *Démonstration / Solution expliquée*

Comprendre la question. Il y a trois couples : on calcule deux moyennes de trois nombres.

Construire la réponse.

1. La moyenne des abscisses vaut $(1 + 3 + 5)/3 = 9/3 = 3$.
2. La moyenne des ordonnées vaut $(3 + 5 + 7)/3 = 15/3 = 5$.
3. On associe ces résultats dans l'ordre : $G(3; 5)$.

Vérifier. Le point central de cette liste symétrique est lui aussi (3; 5), ce qui confirme le calcul. Cette coïncidence avec une observation n'est pas une règle générale.

Conclure. Le point moyen est $G(3; 5)$.

Corrigé 3

 *Démonstration / Solution expliquée*

Comprendre la question. La droite est fournie ; on cherche la valeur qu'elle prévoit pour une abscisse précise.

Construire la réponse.

1. On remplace x par 6 : $\hat{y} = 2,5 \times 6 + 8$.
2. La multiplication donne 15 ; l'addition donne $15 + 8 = 23$.

Vérifier. La droite part de 8 à x égal à zéro et gagne 2,5 par unité d'abscisse. Six unités ajoutent 15, ce qui confirme 23.

Conclure. Le modèle prévoit 23 pour x égal à 6 ; ce n'est pas nécessairement une valeur mesurée.

Corrigé 4

 *Démonstration / Solution expliquée*

Comprendre la question. Le critère est la position de l'abscisse demandée par rapport à l'intervalle observé.

Construire la réponse.

1. Les observations couvrent $[2; 12]$ minutes.
2. Comme $2 \leq 8 \leq 12$, 8 minutes est à l'intérieur : on interpole.
3. Comme $20 > 12$, 20 minutes est à l'extérieur : on extrapole.

Vérifier. Le type de prévision dépend de la durée demandée, pas de la valeur de y que donnera le modèle.

Conclure. À 8 minutes, c'est une interpolation ; à 20 minutes, une extrapolation.

Corrigé 5

 *Démonstration / Solution expliquée*

Comprendre la question. On compare deux indicateurs fournis pour la même variable sur les mêmes données.

Construire la réponse.

1. Le premier indicateur vaut 0,94 et le second 0,62 ; on a $0,94 > 0,62$.
2. Selon ce critère, le premier ajustement explique une plus grande part de la variation observée et laisse une part résiduelle plus faible.
3. Les compléments sont $1 - 0,94 = 0,06$ et $1 - 0,62 = 0,38$, soit respectivement 6 et 38 pour cent de variation non expliquée au sens de cet indicateur.

Vérifier. Ces pourcentages ne sont ni des taux de prédictions exactes ni des probabilités de succès. La comparaison concerne l'ajustement sur la série utilisée.

Conclure. Le premier explique davantage la variation selon R^2 , sans garantie automatique de meilleure prévision hors de cette série.

Corrigé 6

 *Démonstration / Solution expliquée*

Comprendre la question. On décrit le sens général de l'évolution et l'on teste si les points sont exactement alignés.

Construire la réponse.

1. Quand x passe de 1 à 4, les ordonnées passent de 9 à 3 : elles diminuent à chaque étape.
2. Les pas d'abscisse valent tous 1. Les variations d'ordonnée sont $7 - 9 = -2$, $6 - 7 = -1$ et $3 - 6 = -3$.
3. Une droite affine imposerait le même quotient $\Delta y / \Delta x$ à chaque pas ; ici ces quotients valent -2, -1 et -3.

Vérifier. Les différences négatives confirment la tendance décroissante, mais leur inégalité exclut un alignement exact des quatre points.

Conclure. Le nuage est décroissant ; une droite peut éventuellement l'ajuster, mais ne passe pas exactement par tous les points.

Corrigé 7

 *Démonstration / Solution expliquée*

Comprendre la question. La pente est donnée ; il ne reste qu'à déterminer la constante b en utilisant le point G .

Construire la réponse.

1. Une droite de pente 4 s'écrit $y = 4x + b$.
2. Comme elle passe par $(2; 11)$, on a $11 = 4 \times 2 + b = 8 + b$.
3. En soustrayant 8, on trouve $b = 3$.

Vérifier. Pour x égal à 2, $4 \times 2 + 3 = 11$; la pente reste bien 4. L'ordonnée à l'origine 3 est différente de l'ordonnée 11 du point G .

Conclure. L'équation est $y = 4x + 3$.

Corrigé 8

 *Démonstration / Solution expliquée*

Comprendre la question. La droite de Mayer doit passer par les deux points moyens fournis.

Construire la réponse.

1. La pente est la variation verticale divisée par la variation horizontale : $a = (15 - 7)/(6 - 2) = 8/4 = 2$.
2. On écrit alors $y = 2x + b$. Au point $G_1(2; 7)$, $7 = 4 + b$, donc $b = 3$.
3. L'équation obtenue est $y = 2x + 3$.

Vérifier. Pour le second point, $2 \times 6 + 3 = 15$. Les deux points moyens satisfont donc l'équation, comme l'exige la méthode.

Conclure. La droite de Mayer est $y = 2x + 3$.

Corrigé 9

 *Démonstration / Solution expliquée*

Comprendre la question. x compte les années écoulées depuis 2022 ; il ne représente pas directement l'année.

Construire la réponse.

1. Le rang de 2026 est $x = 2026 - 2022 = 4$.
2. On remplace ce rang : $\hat{y} = 6 \times 4 + 50 = 24 + 50 = 74$.

Vérifier. Au rang zéro, la formule donne 50 pour 2022. Quatre hausses annuelles de 6 ajoutent 24, ce qui retrouve 74. L'intervalle observé n'étant pas précisé ici, on ne peut pas classer cette prévision en interpolation ou extrapolation.

Conclure. Le modèle prévoit 74 pour 2026.

Corrigé 10

 *Démonstration / Solution expliquée*

Comprendre la question. Une association observée ne suffit pas à garantir l'effet d'une action sur chaque personne.

Construire la réponse.

1. Les personnes déjà en bonne santé peuvent être plus enclines à choisir le vélo : le sens de l'influence peut être inverse de celui supposé.
2. L'âge, le lieu de vie, les revenus ou l'activité générale peuvent aussi influencer à la fois possession d'un vélo et santé.
3. Donner un vélo n'implique pas son utilisation ; le mot « forcément » ajoute une garantie individuelle absente des données.
4. Il faudrait un protocole permettant de distinguer ces explications pour estimer un effet causal.

Vérifier. Toutes ces explications sont compatibles avec une association positive. Le constat statistique seul ne permet donc pas de choisir entre elles.

Conclusion. Non : on peut décrire l'association, mais pas affirmer que donner un vélo améliore forcément la santé.

Corrigé 11

 *Démonstration / Solution expliquée*

Comprendre la question. Le quatrième point s'écarte d'une tendance visible ; cela justifie une enquête sur la donnée, pas sa suppression automatique.

Construire la réponse.

1. Les points de rangs 1, 2, 3 et 5 suivent $y = 2x + 10$: les images sont 12, 14, 16 et 20.
2. À x égal à 4, cette tendance donnerait 18, alors que l'observation vaut 50. L'écart est $50 - 18 = 32$.
3. On vérifie la saisie, l'unité, l'appareil et les circonstances de mesure. Une erreur documentée pourrait justifier une correction.
4. Si la mesure est valide, elle doit être conservée et expliquée : elle peut révéler un événement particulier ou une limite du modèle.

Vérifier. Supprimer le point seulement parce qu'il gêne l'alignement sélectionnerait les données en fonction du résultat souhaité et pourrait masquer un phénomène réel.

Conclusion. Le point (4; 50) est atypique, mais on ne le supprime pas sans justification liée aux données.

Corrigé 12

 *Démonstration / Solution expliquée*

Comprendre la question. On compare à la fois les centres et les sens d'évolution de deux nuages.

Construire la réponse.

1. Pour A, $\bar{x} = (0 + 2)/2 = 1$ et $\bar{y} = (0 + 2)/2 = 1$.
2. Pour B, $\bar{x} = (0 + 2)/2 = 1$ et $\bar{y} = (2 + 0)/2 = 1$. Les deux centres sont donc (1; 1).
3. Dans A, y augmente de 0 à 2 lorsque x augmente : la tendance est croissante.
4. Dans B, y diminue de 2 à 0 : la tendance est décroissante.

Vérifier. La droite de A est $y = x$ et celle de B $y = 2 - x$. Elles passent toutes deux par (1;1) malgré leurs pentes opposées.

Conclure. Les nuages ont le même point moyen, mais des tendances opposées. Le centre ne détermine donc pas la relation.

Corrigé 13

 *Démonstration / Solution expliquée*

Comprendre la question. On compare les deux modèles avec le même critère sur les mêmes trois observations.

Construire la réponse.

1. A prévoit 0, 2, 4. Les résidus sont $1 - 0 = 1$, $2 - 2 = 0$, $5 - 4 = 1$. Leur somme des carrés vaut $1^2 + 0^2 + 1^2 = 2$.
2. B prévoit 1, 3, 5. Les résidus sont $1 - 1 = 0$, $2 - 3 = -1$, $5 - 5 = 0$. Leur somme des carrés vaut $0^2 + (-1)^2 + 0^2 = 1$.
3. Comme $1 < 2$, B a la plus faible erreur quadratique parmi ces deux modèles sur cette série.

Vérifier. Le carré de -1 vaut 1, pas -1. B n'est pas pour autant déclaré meilleur que toute droite possible, ni garanti pour des prévisions hors des observations.

Conclure. Selon la somme des carrés des résidus, B ajuste mieux ces données que A.

Corrigé 14

 *Démonstration / Solution expliquée*

Comprendre la question. Le résultat représente une proportion exprimée en pourcentage ; il doit être compris entre 0 et 100.

Construire la réponse.

1. La condition $110 - 5x \geq 0$ donne $-5x \geq -110$, puis $x \leq 22$ après division par -5 et inversion du sens.
2. La condition $110 - 5x \leq 100$ donne $-5x \leq -10$, puis $x \geq 2$.
3. Il faut satisfaire les deux conditions : $2 \leq x \leq 22$.

Vérifier. À x égal à 2, le modèle donne 100 ; à x égal à 22, il donne 0. Avant 2, il dépasse 100, et après 22 il devient négatif. Même dans l'intervalle admissible, la validité statistique dépend des observations ayant servi à l'ajustement.

Conclure. On ne peut pas utiliser ce modèle pour toute valeur positive de x ; son interprétation comme proportion impose au moins $x \in [2; 22]$.

Corrigé 15

 *Démonstration / Solution expliquée*

Comprendre la question. On évalue la quantité d'information derrière un ajustement parfait.

Construire la réponse.

1. Deux points d'abscisses distinctes déterminent une droite. Celle-ci peut passer exactement par chacun des deux points.

2. Les deux résidus sont alors nuls : un ajustement parfait est donc possible même avec très peu de données.
3. D'autres courbes peuvent passer par les mêmes points et donner des prévisions très différentes ailleurs.
4. Il faut davantage d'observations et une justification du modèle pour juger la stabilité de la tendance.

Vérifier. Par exemple, $(0; 0)$ et $(1; 1)$ appartiennent à la fois aux courbes $y = x$ et $y = x^2$; à x égal à 2, ces modèles prévoient respectivement 2 et 4.

Conclure. Un ajustement parfait sur deux points n'est pas une preuve solide de qualité prédictive.

9 Corrigé du problème

1. **Comprendre la question.** On conserve chaque couple : une abscisse doit rester associée à son observation.

Construire la réponse.

- (a) Les abscisses sont 1, 2, 3 et 4. Leur moyenne vaut $(1 + 2 + 3 + 4)/4 = 2,5$.
- (b) Les ordonnées sont 20, 28, 36, 44. Leur somme vaut 128 ; leur moyenne est donc $128/4 = 32$.
- (c) On place les quatre points dans un repère dont les axes sont nommés, puis le point moyen $G(2,5; 32)$.

Vérifier. Chaque coordonnée moyenne se trouve entre le minimum et le maximum des coordonnées correspondantes.

Conclure. Le point moyen est $G(2,5; 32)$; les ordonnées s'expriment en participants.

2. **Comprendre la question.** La droite doit passer par les deux points extrêmes ; on calcule sa pente avant son ordonnée à l'origine.

Construire la réponse.

- (a) La variation d'ordonnée vaut $44 - 20 = 24$, pour une variation d'abscisse $4 - 1 = 3$. La pente est donc $24/3 = 8$.
- (b) On écrit $y = 8x + \beta$. Le premier point impose $20 = 8 \times 1 + \beta$, donc $\beta = 12$.
- (c) Pour x égal à 2 et 3, la formule donne 28 et 36, exactement les deux observations intermédiaires.

Vérifier. Au point moyen, $8 \times 2,5 + 12 = 32$: G est aussi sur la droite.

Conclure. Le modèle affine est $y = 8x + 12$. Il reproduit ici les quatre observations.

3. **Comprendre la question.** Une prévision consiste à remplacer x dans le modèle ; son statut dépend de la plage observée.

Construire la réponse.

- (a) À $x = 2,5$, $y = 8 \times 2,5 + 12 = 32$. Cette abscisse est dans $[1 ; 4]$.
- (b) À $x = 6$, $y = 8 \times 6 + 12 = 60$. Cette abscisse est au-delà des observations.

Vérifier. La première prévision est encadrée par des observations proches. La seconde suppose que la tendance se prolonge au-delà des données disponibles.

Conclure. Les prévisions sont 32 participants par interpolation et 60 participants par extrapolation.

4. **Comprendre la question.** Le résidu demandé est observation moins prévision, dans cet ordre.

Construire la réponse.

- (a) À x égal à 5, la prévision est $8 \times 5 + 12 = 52$.
- (b) La nouvelle observation vaut 55. Le résidu est donc $55 - 52 = 3$.

Vérifier. En ajoutant le résidu à la prévision, $52 + 3 = 55$, on retrouve la mesure.

Conclure. La nouvelle observation dépasse la prévision de 3 participants.

5. **Comprendre la question.** Il faut distinguer une approximation utile, une loi exacte et une relation de cause à effet.

Construire la réponse.

- (a) Un écart non nul montre que le modèle ne reproduit pas exactement cette nouvelle observation. Il ne suffit pas à rendre toute utilisation impossible.
- (b) On compare sa taille à la précision attendue, à la dispersion des autres mesures et au coût d'une mauvaise décision.
- (c) Une tendance commune peut provenir d'une troisième variable ou de conditions de mesure. La droite seule ne démontre pas une causalité.

Vérifier. Il faut recueillir d'autres observations, notamment dans la plage d'utilisation prévue, pour voir si les écarts restent acceptables ou deviennent systématiques.

Conclure. Le modèle peut rester utile dans une plage définie, avec des limites explicites ; il ne prouve aucune causalité.

10 Variante autonome et bilan

Un autre jeu suit $y = 9x + 14$ pour x de 1 à 4. Donne son point moyen et sa prévision en $x = 3,5$.

À toi d'essayer : Vérifie ta démarche

Résous la variante, puis explique pourquoi ta méthode s'applique.

Tu peux chercher, prendre un indice ou lire l'explication, à ton rythme.

[Indices, p. 26](#) [Explication, p. 27](#)

- Je peux expliquer : représenter deux variables par un nuage.
- Je peux expliquer : le point moyen du nuage.
- Je peux expliquer : construire un ajustement affine.
- Je peux expliquer : interpolation et extrapolation.
- Je peux expliquer : association, causalité et qualité d'ajustement.

11 Fiche-mémoire

Représenter deux variables par un nuage.

1. Identifie les deux variables et garde chaque paire de mesures associée.
2. Place x horizontalement et y verticalement, avec des axes nommés et gradués.
3. Décris la tendance, la courbure, les groupes et les points atypiques.
4. Compare des taux de variation si tu recherches un alignement exact.
5. Vérifie les données atypiques sans les supprimer uniquement pour améliorer l'apparence.

Le point moyen du nuage.

1. Compte le nombre de couples et calcule la moyenne des abscisses.
2. Calcule séparément la moyenne des ordonnées avec le même nombre d'observations.
3. Écris G avec les coordonnées dans le bon ordre et leurs unités.
4. Pour une droite passant par G , remplace x et y par ces deux moyennes.
5. Contrôle le passage par G et ne déduis pas la forme du nuage de son seul centre.

Construire un ajustement affine.

1. Observe le nuage et précise la méthode d'ajustement utilisée.
2. Calcule ou relève les coefficients a et b avec une précision suffisante.
3. Calcule les prévisions $\hat{y}_i$, puis les résidus $y_i - \hat{y}_i$.
4. Compare les carrés des résidus et leur organisation, pas seulement leur somme signée.
5. Présente les valeurs obtenues comme des prévisions dépendant du modèle.

Interpolation et extrapolation.

1. Relève l'intervalle observé, le codage de x et les unités.
2. Convertis la demande en abscisse et distingue interpolation et extrapolation.
3. Calcule une image ou résous l'équation correspondant à l'objectif.
4. Reconvertis le résultat en date ou en quantité et arrondis avec une précision justifiée.
5. Contrôle les limites physiques et indique la dépendance au modèle.

Association, causalité et qualité d'ajustement.

1. Précise si la question concerne l'ajustement, la prévision ou une explication causale.
2. Compare les indicateurs uniquement sur des données et une variable comparables.
3. Examine le nuage, les résidus, les valeurs atypiques et la taille de la série.
4. Cherche une cause commune, une sélection ou une causalité inverse avant de conclure.
5. Limite la conclusion à ce que le protocole et les données permettent réellement d'affirmer.

12 Indices des pauses et des preuves

Chercher avec un indice fait partie du travail. Tu peux revenir à l'énoncé avec le lien sous chaque aide.

Pause 1 — Mets la notion à l'épreuve

Indice 1 — Une piste.

Chaque ligne correspond à un même individu ou à une même situation.

Indice 2 — Un pas de plus.

Séparer ou trier indépendamment les colonnes détruirait cette association.

[Revenir à la recherche, p. 4](#)

Pause 2 — Mets la notion à l'épreuve

Indice 1 — Une piste.

Moyenne d'abord 0 et 4 pour l'abscisse.

Indice 2 — Un pas de plus.

Moyenne ensuite 2 et 8 pour l'ordonnée.

[Revenir à la recherche, p. 6](#)

Pause 3 — Mets la notion à l'épreuve

Indice 1 — Une piste.

Les écarts sont mesurés verticalement entre observation et prévision.

Indice 2 — Un pas de plus.

On les élève au carré pour éviter les compensations entre signes.

[Revenir à la recherche, p. 7](#)

Pause 4 — Mets la notion à l'épreuve

Indice 1 — Une piste.

Compare l'abscisse demandée avec les bornes des données disponibles.

Indice 2 — Un pas de plus.

À l'extérieur, on prolonge la tendance au-delà des observations.

[Revenir à la recherche, p. 8](#)

Pause 5 — Mets la notion à l'épreuve**Indice 1 — Une piste.**

R^2 décrit l'écart entre un modèle et des observations.

Indice 2 — Un pas de plus.

Une cause commune peut produire un très bon alignement sans effet direct entre les variables.

[Revenir à la recherche, p. 9](#)

Vérifie ta démarche**Indice 1 — Une piste.**

Repère les données utiles et l'inconnue. Reviens à la méthode correspondante dans la boîte à outils.

Indice 2 — Un pas de plus.

Contrôle les unités, les bornes et la plausibilité du résultat avant de lire la correction.

[Revenir à la recherche, p. 24](#)

13 Explications des pauses et des preuves

Lis une étape, puis essaie de poursuivre avant de lire la suivante. Les calculs ci-dessous complètent le cours.

Pause 1 — Mets la notion à l'épreuve

Un nuage conserve les couples d'observations associés ; il ne se réduit pas aux moyennes des deux colonnes.

[Revenir à la recherche, p. 4](#)

Pause 2 — Mets la notion à l'épreuve

$$G((0+4)/2; (2+8)/2) = G(2; 5).$$

[Revenir à la recherche, p. 6](#)

Pause 3 — Mets la notion à l'épreuve

Les moindres carrés minimisent la somme des carrés des résidus verticaux.

[Revenir à la recherche, p. 7](#)

Pause 4 — Mets la notion à l'épreuve

Il s'agit d'une extrapolation, dont la validité dépend de la poursuite du modèle hors de la zone observée.

[Revenir à la recherche, p. 8](#)

Pause 5 — Mets la notion à l'épreuve

Non : un R^2 proche de 1 ne prouve pas une causalité.

[Revenir à la recherche, p. 9](#)

Vérifie ta démarche

Comprendre la question. On applique le nouveau modèle aux quatre abscisses de la variante.

Construire la réponse.

1. La moyenne des abscisses 1, 2, 3, 4 reste 2,5. Les ordonnées sont 23, 32, 41 et 50.
2. Leur moyenne vaut $9 \times 2,5 + 14 = 36,5$. Ainsi $G(2,5; 36,5)$.
3. À $x = 3,5$, on obtient $9 \times 3,5 + 14 = 45,5$.

Vérifier. L'abscisse 3,5 est comprise entre 1 et 4 : il s'agit d'une interpolation. Pour un modèle affine, l'image de la moyenne des abscisses égale la moyenne des images.

Conclure. Le point moyen est $G(2,5; 36,5)$ et la prévision vaut 45,5 participants.

[Revenir à la recherche, p. 24](#)