# Statistiques — A
# Corrigé



**Exercice 1.** *À partir des recensements effectués tous les dix ans, on a établi le tableau suivant qui donne l'évolution de la population française en millions d'individus entre 1851 et 1911. Peu de données sont disponibles pour l'année 1871.*

| Année : | 1851 | 1861 | 1881 | 1891 | 1901 | 1911 |
|---|---|---|---|---|---|---|
| Rang de la décennie : $x_i$ | 0 | 1 | 3 | 4 | 5 | 6 |
| Population en millions : $y_i$ | 35 | 37,4 | 37,7 | 39,9 | 39 | 39,6 |

*Source : INSEE*

## Partie A : Approximation de la population en 1871.

1. *Placer sur le graphique donné au verso le nuage de points de coordonnées ($x_i$ ; $y_i$ ).* Voir le graphique.
2. *Donner une équation de la droite d'ajustement affine de y en fonction de x obtenue par la méthode des moindres carrés. Les coefficients seront arrondis au millième.* À la calculatrice, on trouve $y = 0,701x + 35,881$.
3. *On décide d'ajuster ce nuage de points par la droite (d) d'équation $y = 0,7x + 35,9$. Tracer cette droite sur ce même graphique.*
   - On prend une première valeur $x = 0$ (par exemple), et on trouve $y = 0,7\times 0+35,9 = 35,9$. Donc le point de coordonnées $(0; 35,9)$ est sur la droite.
   - On prend une seconde valeur $x = 10$ (par exemple), et on trouve $y = 0,7 \times 10 + 35,9 = 42,9$. Donc le point de coordonnées $(10; 42,9)$ est sur la droite.

   Il ne reste plus qu'à tracer la droite passant par ces deux points. Voir sur le graphique.
4. *A l'aide de ce modèle, estimer la population en 1871.* L'année 1871 correspond à l'année d'indice $x = 2$, donc $y = 0,7 \times 2 + 35.9 = 37,3$. Donc en 1871, il y avait environ $37,3$ millions d'habitants en France.

   Remarque : Il était aussi possible de lire cette information graphiquement (voir sur le graphique).

**Partie B : Évolution de la population après 1911.**

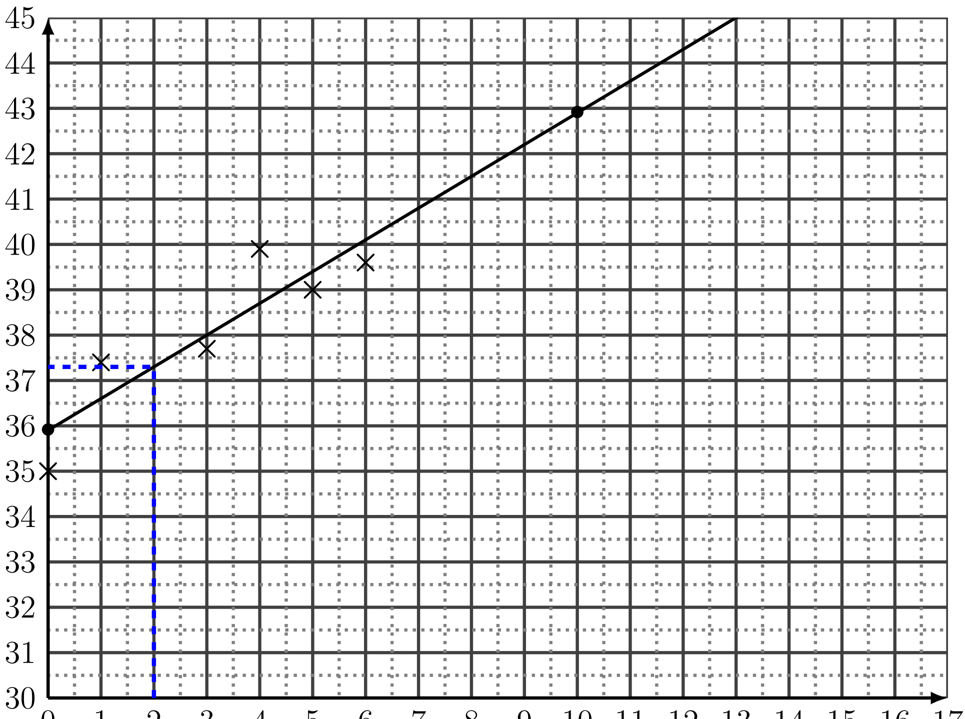
*On étudie un autre ajustement de cette série statistique, par la fonction $f$ définie par $f(x) = 0,1x^2 + 0,1x + 36,2$ (où $x$ est le rang de l'année, et $f(x)$ la population estimée cette année là). Par exemple, pour l'année 1911 de rang 6, ce modèle estime la population à $f(6) \approx 40,4$ millions d'habitants. D'autre part, les données de l'INSEE (Institut National de la Statistique et des Études Économiques) montrent qu'en 2011 il y avait 65,2 millions d'habitants en France.*

1. *Le modèle étudié dans la partie A reste-t-il valable jusqu'en 2011 ? Justifier.*

   Calculons la population en 2011 : cette année correspond à l'indice 16, donc $y = 0,7 \times 16 + 35,9 = 47,1$. Selon ce modèle, il y aurait 47,1 millions d'habitants en 2011, ce qui est bien inférieur à la réalit (65,2 millions). Le modèle n'est donc plus valide en 2011.

2. *Le nouveau modèle donné par la fonction $f$ est-il meilleur ? Justifier.*

   Avec le nouveau modèle, la population en 2011 est estimée à $f(16) = 0,1 \times 16^2 + 0,1 \times 16 + 36,2 = 63,4$ millions, ce qui est très proche de la valeur réelle : ce nouveau modèle est meilleur que le précédent.



**Exercice 1.** *En France, le temps moyen quotidien, en heures, passé par une personne devant un écran d'ordinateur, de tablette ou de smartphone est donné dans le tableau suivant :*

| *Année* | *2013* | *2014* | *2015* | *2016* | *2017* |
|---|---|---|---|---|---|
| *Rang de l'année $x_i$* | *0* | *1* | *2* | *3* | *4* |
| *Temps en heures passé devant un écran $y_i$* | *2,78* | *3,27* | *3,52* | *3,77* | *3,97* |

**Partie A : Approximation du temps passé devant un écran en 2018.**

1. *Placer sur le graphique donné au verso le nuage de points de coordonnées ($x_i$ ; $y_i$ ).* Voir le graphique.
2. *À l'aide de la calculatrice, déterminer une équation de la droite d'ajustement de $y$ en $x$ par la méthode des moindres carrés. On arrondira les coefficients au millième.* À la calculatrice, on trouve $y = 0,288x + 2,886$.

*Dans la suite de l'exercice, on prend la droite d'équation $y = 0,3x + 2,9$ comme ajustement du nuage de points.*

3. *Tracer cette droite dans le repère donné en annexe à rendre avec la copie.*
   — On prend une première valeur $x = 0$ (par exemple), et on trouve $y = 0,3 \times 0 + 2,9 = 2,9$. Donc le point de coordonnées $(0; 2,9)$ est sur la droite.
   — On prend une seconde valeur $x = 10$ (par exemple), et on trouve $y = 0,3 \times 10 + 2,9 = 5,9$. Donc le point de coordonnées $(10; 5,9)$ est sur la droite.

   Il ne reste plus qu'à tracer la droite passant par ces deux points. Voir sur le graphique.
4. *En utilisant cet ajustement, déterminer une estimation du temps quotidien passé devant un écran en 2018.* L'année 2018 correspond à l'indice 5. Or $y = 0,3 \times 5 + 2,9 = 4,4$. Donc le temps moyen passé est estimé à 4,4h.

Remarque : On pouvait aussi lire cette information graphiquement (voir le graphique).

**Partie B : Évolution du temps passé devant un écran après 2017.**

*On étudie un autre ajustement de cette série statistique, par la fonction $f$ définie par $f(x) = -0,04x^2 + 0,45x + 2,8$ (où $x$ est le rang de l'année, et $f(x)$ le temps passé devant un écran estimé cette année là). Par exemple, pour l'année 2017 de rang 4, ce modèle estime le temps passé à $f(4) \approx 3,96$ heures.*

1. *Calculer les estimations du temps passé en 2017 pour chacun des deux modèles (celui de la partie A, et celui donné par la fonction $f$). Lequel des deux semble le meilleur ?* L'année 2017 correspond à l'indice 4.
   — Selon le modèle de la partie précédente, le temps passé est estimé à $y = 0,3 \times 4 + 2,9 = 4,2$.
   — Selon le nouveau modèle, le temps passé est $f(4) \approx 3,96$ (d'après l'énoncé).

   La vraie valeur étant 3,97, pour cette année-là, le nouveau modèle est plus précis que l'ancien.
2. *Pourquoi peut-on affirmer que le modèle donné par la fonction $f$ n'est pas valide en 2030 ?* L'année 2030 correspond à l'indice 17. Or $f(17) = -0,04 \times 17^2 + 0,45 \times 17 + 2,8 = -1,11$. Cela correspond à un temps négatif, ce qui est impossible : le nouveau modèle n'est pas valide en 2030.