Analyse de l'information chiffrée

1 Rappels : proportion et pourcentage

1.1 Proportion d'une sous-population

Définition 1 : La proportion, notée p, d'une sous-population de taille n dans une population de taille N vaut :

$$p = \frac{n}{N} = \frac{\text{Sous-effectif}}{\text{Effectif total}}$$

Pour exprimer cette proportion en pourcentage, on multiplie le nombre obtenu par 100.

■ **Exemple 1 :** Sur les 400 élèves de 1ère, 280 ont choisi de continuer les mathématiques.
La proportion d'élèves ayant continué les mathématiques est de $\frac{280}{400} = 0{,}7$. Cela représente 70% des élèves.
■

Propriété 1 : Soit p la proportion d'une sous-population dans une population de taille N.
L'effectif n de cette sous-population vaut $n = p \times N$.

■ **Exemple 2 :** Les jeunes de moins de 16 ans représentaient $1{,}27\%$ de la population française en 2018. Cette population s'élève à environ 67 millions habitants.
Le nombre de jeunes de moins de 16 ans dans la population française vaut donc $\frac{1.27}{100} \times 67000000 \simeq 850900$.
En 2018, il y avait environ 850 000 jeunes de moins de 16 ans dans la population française. ■

Propriété 2 : A, B et C sont trois ensembles tels que $A \subset B \subset C$. On note

- p_{AB} la proportion de A dans B
- p_{BC} la proportion de B dans C.

Alors la proportion de A dans C vaut $p_{AC} = p_{AB} \times p_{BC}$.

Démonstration 1.1 : Appelons n_A, n_B et n_C les effectifs des populations A, B et C. On a alors

$$p_{AC} = \frac{n_A}{n_C} = \frac{n_A}{n_B} \times \frac{n_B}{n_C} = p_{AB} \times p_{BC}$$

□

■ **Exemple 3 :** Dans un lycée, 45 % des élèves font un sport et un tiers d'entre eux font du badminton.
La proportion d'élèves du lycée faisant du badminton est donc $\frac{45}{100} \times \frac{1}{3} = \frac{15}{100} = 0{,}15$.
Autrement dit, 15% des élèves du lycée font du badminton. ■

2 Tableaux à double entrée

Définition 2 : Un tableau à double entrée (ou tableau croisé d'effectifs) est un tableau qui donne conjointement en lignes et en colonnes les effectifs des différentes valeurs de deux caractères d'une population.

■ **Exemple 4 :** Le tableau croisé ci-dessous présente les résultats des candidats aux baccalauréats en 2021.

	Général	Technologique	Professionnel	Total
Admis	379930	147896	186395	714221
Recalé	6029	6407	19206	31642
Total	385959	154303	205601	745863

La population étudiée est l'ensemble des personnes ayant passé leur baccalauréat en 2021. Les caractères étudiés sont le type de baccalauréat et leur réussite ou non.

D'après ce tableau, le taux de réussite au baccalauréat, toutes filières confondues, est de $\frac{714221}{745863} \simeq 0.958$ soit environ $95{,}8\%$ de réussite. Le taux de réussite en filière générale est lui de $\frac{379930}{385959} \simeq 0.984$ soit environ 98.4% de réussite.

■

■ **Exemple 5 :** On réalise un sondage auprès de 2000 personnes à propos d'un film. Parmi ces personnes, 40% déclarent avoir aimé le film. 35% des personnes interrogées ont moins de 35 ans et, parmi celles-ci, quatre cinquièmes déclarent avoir aimé le film. Ces résultats peuvent être résumées dans un tableau à double-entrée.

	ont aimé	n'ont pas aimé	Total
- de 35 ans	560	140	700
35 ans ou +	240	1060	1300
Total	800	1200	2000

Pour s'y retrouver, il est préférable de rappeler à chaque fois l'effectif de la population considérée.

- 2000 personnes ont été interrogées : on place donc le nombre 2000 dans la cellule" Total ; Total".
- 40 % des 2000 interrogées ont déclaré avoir aimé le film, ce qui représente $\frac{40}{100} \times 2000 = 800$ personnes. On place donc le nombre 800 dans la case "ont aimé ; Total".
- Il y a donc $2000 - 800 = 1200$ personnes qui n'ont pas aimé le film. On place le nombre 1200 dans la case "n'ont pas aimé ; Total".
- 35% des 2000 personnes interrogés ont moins de 35 ans, soit $\frac{35}{100} \times 2000 = 700$ personnes.
- Par conséquent, $2000 - 700 = 1300$ des personnes interrogées ont 35 ans ou +.
- Quatre cinquièmes des 700 personnes ayant moins de 35 ans ont aimé le film, ce qui représente $\frac{4}{5} \times 700 = 560$ personnes. On place ce nombre dans la cellule "ont aimé ; - de 35 ans"
- Il y a donc $700 - 560 = 140$ personnes qui ont - de 35 ans et qui n'ont pas aimé le film. Par ailleurs, $800 - 560 = 240$ personnes ont 35 ans ou + et ont aimé le film.
- Enfin, $1300 - 240 = 1060$ personnes ont 35 ans ou + et n'ont pas aimé le film. On aurait pu également faire le calcul en utilisant la colonne pour obtenir le même résultat (ce qui nous permet de vérifier que notre tableau est cohérent).

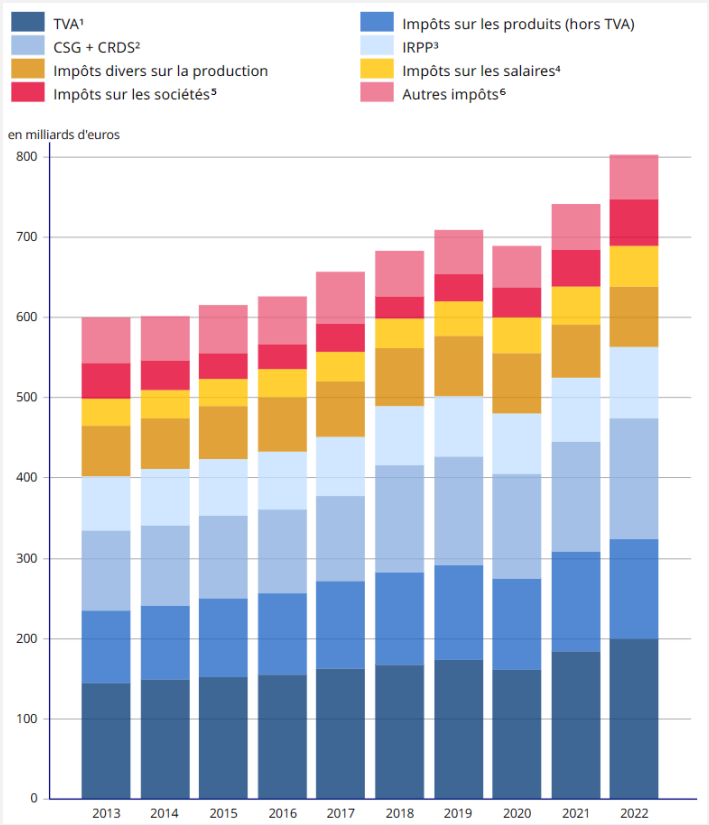
■

3 Représentation graphique des données

3.1 Diagramme en barres

Définition 3 : Un diagramme en bâtons ou diagramme en barres permet de représenter les effectifs de plusieurs variables catégorielles (qui ne se rapportent pas forcément à des quantités).

■ **Exemple 6 :** Le diagramme ci-dessous est un diagramme empilé qui montre l'ensemble des impôts en France entre les années 2013 et 2022.



Sur ce graphique, la hauteur totale des barres représente la somme totale des impôts perçues par l'État.

Par exemple, en 2022, l'État a perçu près de 800 milliards d'euros d'impôts au total. En 2013, ce montant était d'environ 600 milliards d'euros. La tendance de ces montants est à la hausse, la seule exception étant en 2020 (saurez-vous expliquer pourquoi ?)

Pour connaître le montant représentée par une catégorie de revenu, il faut calculer la différence entre les limites hautes et basses d'une zone du graphique.

Par exemple, l'impôt sur les sociétés en 2022 est représenté par la zone qui s'étend d'environ 690 à 730. Or, $730 - 690 = 40$, on peut donc estimer que l'impôt sur les sociétés en 2022 s'élève environ à 40 millions d'euros.

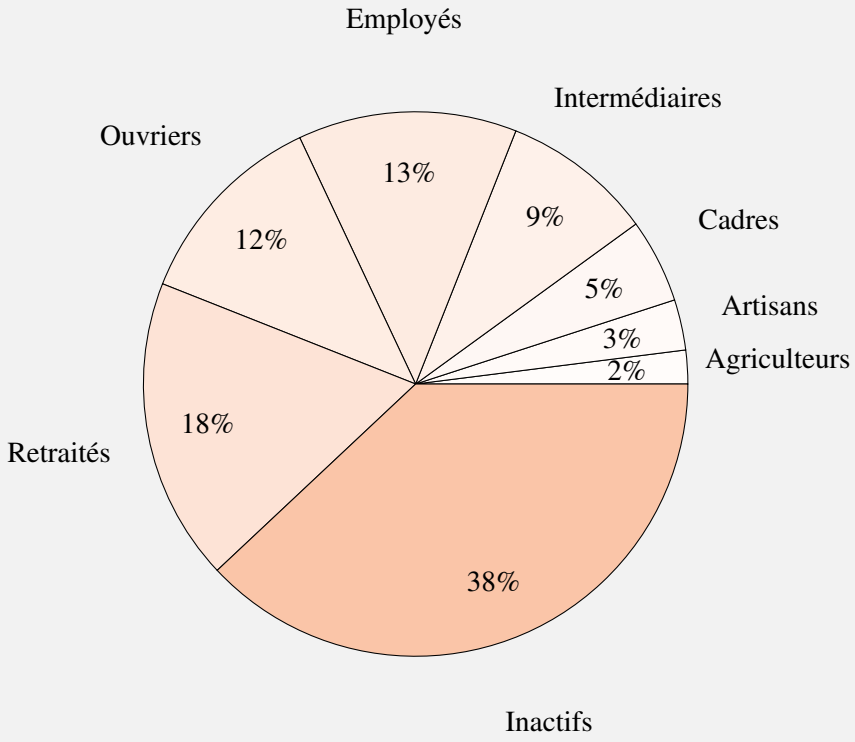
■

Les détails complets de ce graphique se trouvent à l'adresse : https://www.insee.fr/fr/statistiques/2381408

3.2 Diagramme circulaire

Définition 4 : Un diagramme circulaire ou diagramme en secteurs est un graphique permettant de représenter un petit nombre de valeurs ou de classes par des angles proportionnels à la fréquence de ces valeurs dans la population étudiée.

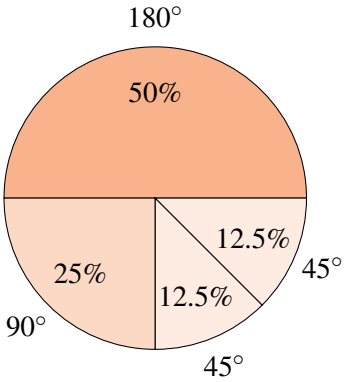
■ **Exemple 7 :** Le diagramme ci-dessous montre la répartition en pourcentage de la population française par catégorie socioprofessionnelles pour l'année 1999.

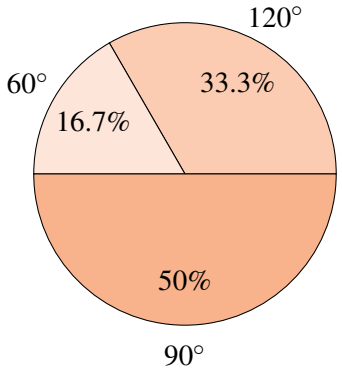


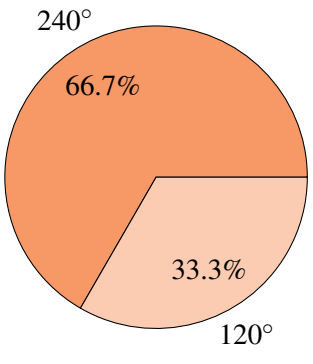
Rappelons qu'un tour complet correspond à 360°. Les ouvriers représentent 12% de la population.

Le secteur qui représente les ouvriers a donc un angle de $\frac{12}{100} \times 360 = 43.2$ degrés. ■

Pour lire efficacement ce genre de diagramme, il est important de se représenter mentalement différents angles remarquables et les proportions auxquelles ces angles correspondent.







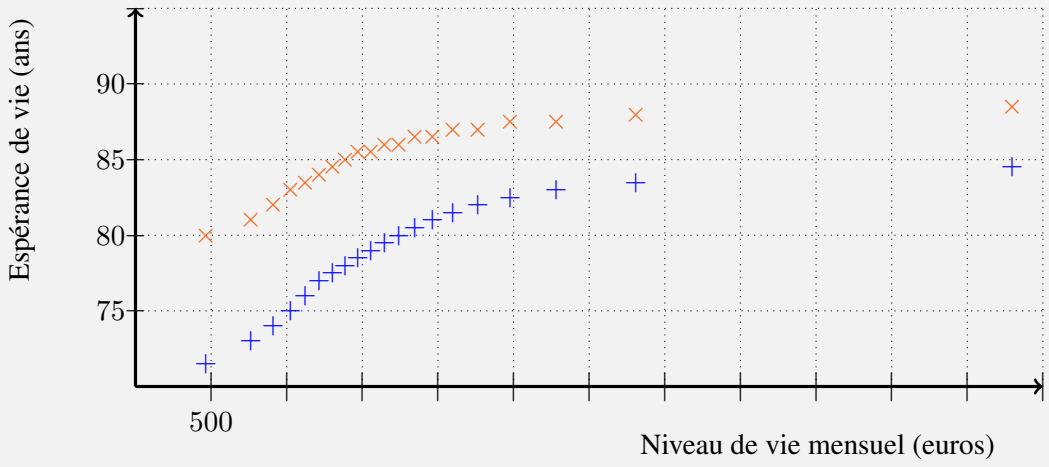
3.3 Nuage de points

Définition 5 : On considère une série statistique composée de deux caractères. Le nuage de points associé à cette série est l'ensemble des points ayant pour abscisses les valeurs du premier caractère et pour ordonnées les valeurs du second caractère.

■ **Exemple 8 :** Le site de l'INSEE indique les données suivantes quant à l'espérance de vie à la naissance selon le sexe et le niveau de vie mensuel en euros.

Niveau de vie	466	762	910	...	5796
Femmes	80.0	81.1	82.0	...	88.3
Hommes	71.7	73.0	74.2	...	84.4

Les données complètes peuvent être trouvées à l'adresse suivante : https://www.insee.fr/fr/statistiques/6046588.
Pour visualiser ces données, il est possible d'utiliser un nuage de points, en représentant en abscisses le niveau de vie mensuel et en ordonnées l'espérance de vie.



En regardant ce nuage de points, on s'aperçoit que l'espérance de vie augmente avec le niveau de vie mensuel. Par ailleurs, l'espérance de vie chez les femmes est nettement supérieure à celle des hommes.

Bien que nous n'ayons pas la précision que peut nous donner le tableau complet, un nuage de points nous permet d'analyser un comportement, une tendance, un lien entre différentes variables.

■

Exercices : Information chiffrée

1 Proportions

► Exercice 1
A la dernière question d'une interrogation, 12 élèves ont répondu juste et 13 ont répondu faux. Quel est le pourcentage de bonnes réponses ?

► Exercice 2
M. Machin télécharge (légalement bien sûr) un fichier d'une taille de 360 Mo. Son téléchargement est terminé à 62%. Combien de Mo lui reste-t-il à télécharger ?

► Exercice 3
Sur un site , on peut lire l'affirmation suivante : « en mai 2021, le nombre d'utilisateurs actifs des réseaux sociaux s'élève à 4,33 milliards, soit 55,1% de la population mondiale ».

1. A combien d'habitants s'élève la population mondiale ?
2. Selon ce même site, 99% des socionautes se connectent à leurs réseaux via leurs appareils mobiles. Combien cela représente-il d'utilisateurs ?

► Exercice 4
Dans le journal télévisé de France 2 daté du 2 avril 2020, un journaliste déclare « 1,5 millions de Français ont été contaminés par le virus, soit 9% de la population »

1. Supposons que cette affirmation soit vraie, quelle serait alors la population française ?
2. La population française s'élève en réalité à 67,39 millions d'habitants.
   (a) Quel était alors le véritable pourcentage de personnes contaminées par le virus ?
   (b) Si 9% de la population avait été contaminée, combien cela représenterait-il de personnes ?

► Exercice 5
Dans un lycée, 60% des élèves sont des filles et 20% d'entre elles sont internes. Quelle est la proportion de filles internes dans le lycée ?

► Exercice 6
Dans une boîte de bonbons, 20% sont des Dragibus et 5% d'entre eux sont des Dragibus noirs.

1. Quelle est la proportion de Dragibus noirs dans une boîte de bonbon ?
2. Dans une boîte de 200 bonbons, combien aurait-on de Dragibus noirs ?

► Exercice 7
Dans une société, les cadres représentent 40% des employés, et les cadres supérieurs représentent un cinquième des cadres. Quelle est la proportion de cadres supérieurs parmi les employés ?

► Exercice 8
Dans une classe comportant 40% de filles, 20% de ces filles font du sport. 30% des garçons en font également. Quel pourcentage de la classe fait du sport ?

► Exercice 9

Au football, lors de transfert de joueurs, il arrive que les clubs négocient un pourcentage à verser en cas de revente du joueur en question. En 2019, Nabil Fékir est ainsi transféré de l'Olympique Lyonnais au Bétis Séville contre une indemnité de 20 millions d'euros et un pourcentage à la revente de 20%

1. Si le Bétis Séville revend Nabil Fékir pour 30 millions d'euros, combien le club devra-t-il verser à l'Olympique Lyonnais pour respecter ce pourcentage à la revente ?
2. Si le Bétis Séville souhaite empocher 35 millions d'euros pour la vente de ce joueur après avoir reversé le pourcentage accordé à l'Olympique Lyonnais, à quel prix doit-il vendre le joueur ?

2 Tableaux à double entrée

► Exercice 10

Dans un collège de 500 élèves, 60% sont des filles. 80 % d'entre elles sont droitières. On sait également que 15% des garçons sont gauchers. Compléter le tableau suivant avec les **effectifs**.

	Gauchers	Droitiers	Total
Garçons			
Filles			
Total			

► Exercice 11

Lors d'une opération de promotions exceptionnelles d'un grand magasin de bricolage, on s'intéresse aux ventes de deux articles particuliers du rayon « Outillage motorisé » : une meuleuse et une scie sauteuse.

Pendant cette période de promotions, une enquête réalisé sur 300 clients de ce magasin montre que :

- 63 clients ont acheté une meuleuse;
- 80 clients ont acheté une scie sauteuse;
- 5% des clients ayant acheté une scie sauteuse ont aussi acheté une meuleuse.

Chaque client a acheté au plus une scie sauteuse et au plus une meuleuse.

1. Compléter le tableau croisé d'effectifs ci-dessous.

	Achète une meuleuse	N'achète pas de meuleuse	Total
Achète une scie sauteuse			
N'achète pas de scie sauteuse			
Total			

2. Quel est le pourcentage de clients ayant acheté une meuleuse ?

► Exercice 12

On a interrogé 1200 personnes pour savoir si elles possédaient ou non un ordinateur ou une télévision. 73% déclarent avoir une télévision. Parmi eux 75% ont également un ordinateur. Parmi ceux n'ayant pas de télévision, un quart n'a pas non plus d'ordinateur.

Compléter le tableau suivant avec **les effectifs**.

	Ordinateur	Pas d'ordinateur	Total
Télévision			
Pas de télévision			
Total			

► Exercice 13

Un fabriquant d'ampoules possède deux machines notées A et B. La machine A fournit 65% de la production et la chine B fournit le reste. Certaines ampoules présentent un défaut de fabrication :

- À la sortie de la machine A, 8% des ampoules présentent un défaut;
- À la sortie de la machine B, 4% des ampoules présentent un défaut.

La production quotidienne du fabriquant est de 15000 ampoules par jour.

1. Combien d'ampoules proviennent de chacune des machines?
2. Compléter le tableau croisé des effectifs :

	Machine A	Machine B	Total
Avec défaut			
Sans défaut			
Total			

3. Calculer la fréquence en pourcentage des ampoules ayant un défaut.

► Exercice 14

On fait une prise de sang à 300 personnes pour connaître leur groupe sanguin (A, B, AB ou O) et leur rhésus (+ ou -). 10% de ces personnes sont du groupe B et 4% du groupe AB. Le reste est équitablement réparti entre les groupes A et O. Parmi les individus du groupe B, 10% des personnes ont le rhésus -. Cette proportion s'élève à 25% pour le groupe AB. On sait par ailleurs que les individus O+ représentent 36% de la population testée et que les individus A+ en représentent 38%.

Résumer toutes ces informations dans un tableau croisé d'effectifs.

	Groupe A	Groupe B	Groupe AB	Groupe O	Total
Rhésus +					
Rhésus -					
Total					

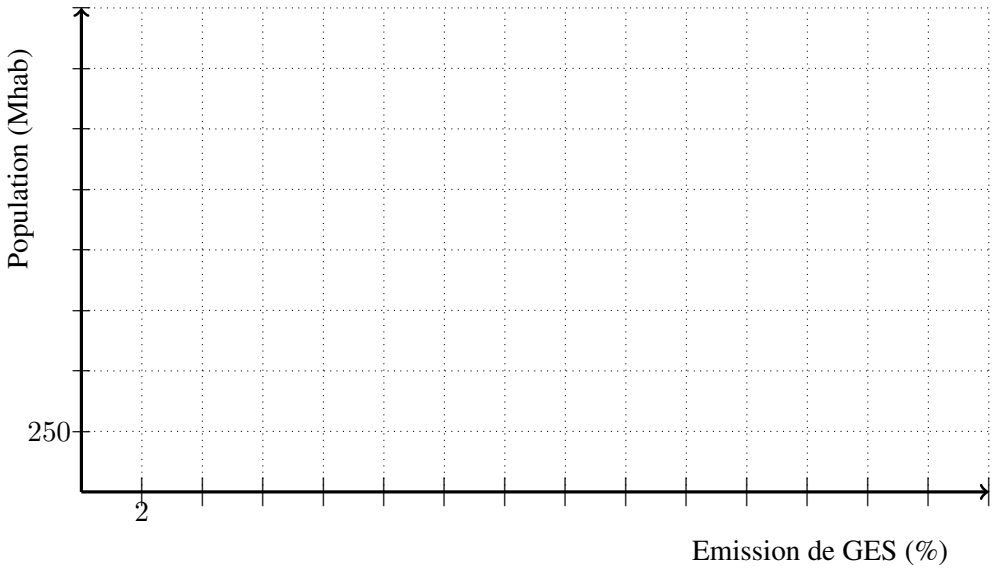
3 Représentation graphique des données

► Exercice 15

Dans son rapport *Mitigation of Climate Change* d'avril 2022, le GIEC (Groupe intergouvernemental sur l'évolution du climat) donne le tableau suivant, présentant la répartition des émissions de gaz à effet de serre (GES) par zone géographique en 2019.

	Population (Millions)	Emission de GES
Afrique	1292	9%
Amérique latine	646	10%
Amérique du Nord	366	12%
Asie de l'Est	1471	27%
Asie du Sud	1836	8%
Asie du Sud-Est	674	9%
Australie, Japon, Nouvelle-Zélande	157	3%
Europe	620	8%
Europe de l'Est	291	6%
Moyen-Orient	252	5%

1. Quelle proportion de la population mondiale se trouve en Amérique du Nord ? Comparer avec la part de gaz à effet de ce serre émis depuis ce continent.
2. Représenter le nuage de points (Population, Émission) sur le graphique suivant



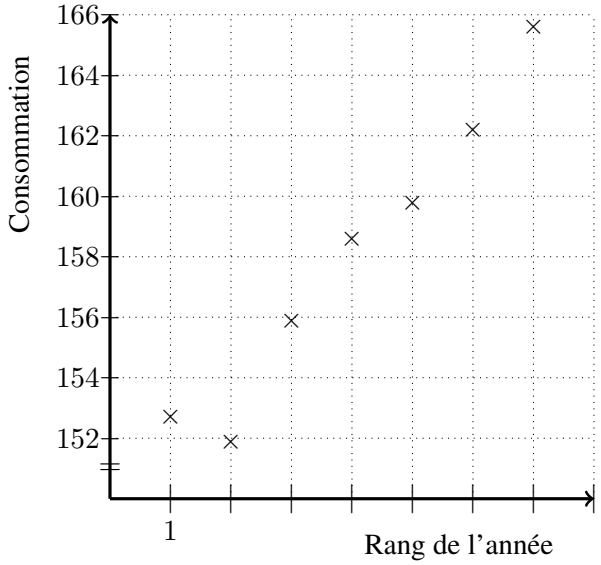
3. Que pensez-vous de l'affirmation suivante : « plus une zone est peuplée, plus il y a de gaz à effets de serre émis depuis cette zone » ?

► Exercice 16

L'interpolation et l'extrapolation linéaire sont des méthodes statistiques permettant d'estimer la valeur prise par une fonction continue en approchant celle-ci par une fonction affine. On considère par exemple le tableau suivant, qui donne l'évolution du pôle "Alimentation et boissons non alcoolisées" dans la consommation des ménages en France entre 2013 et 2019. Les données de consommation sont indiquées en milliard d'euros.

Rang de l'année	1	2	3	4	5	6	7
Consommation	152.7	151.9	154.9	158.6	159.8	162.2	165.6

On a représenté le nuage de points correspondant ci-dessous



Les points semblent presque alignés. Nous allons essaye d'approcher ce nuage de points par une droite.

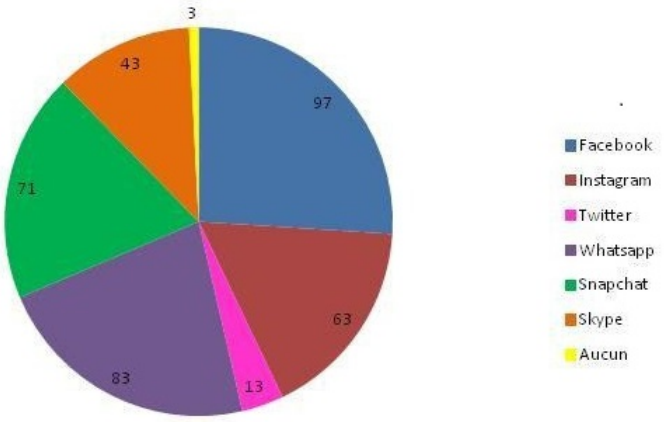
1. Calculer les coordonnées du point dont l'abscisse est égal à la moyenne des abscisses des trois premiers points et dont l'ordonnée est égal à la moyenne des ordonnées des trois premiers points. Placer ce point sur le graphique.
2. Faire de même avec les quatre derniers points.
3. Tracer la droite qui relie les deux points ainsi placés. Cette droite vous semble-t-elle approcher le nuage de points de manière satisfaisante ?

On admet que cette droite admet pour équation réduite $y = 2.96x + 148.42$, les coefficients ayant été arrondis au centième. y représente la consommation pour l'année de rang x.

4. Selon ce modèle, quelle consommation en milliards d'euros peut-on prévoir pour l'alimentation et les boissons non alcoolisées en 2023 en France ?
5. Selon ce modèle, en quelle année cette consommation dépassera-t-elle 200 milliards d'euros ?

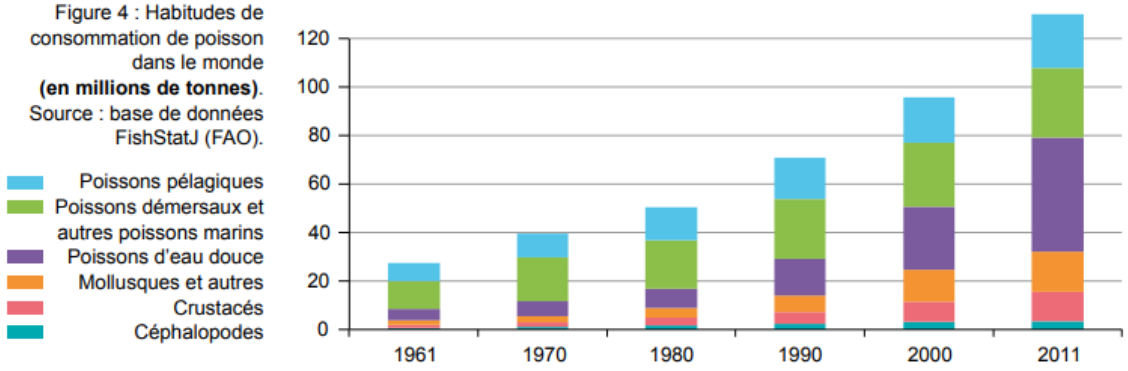
► Exercice 17

Une élève a réalisé un sondage en 2015 auprès de 100 élèves pour connaître les réseaux sociaux utilisés par ses camarades et a ensuite réalisé le diagramme suivant. Celui-ci vous semble-t-il pertinent ? Pourquoi ?



► Exercice 18

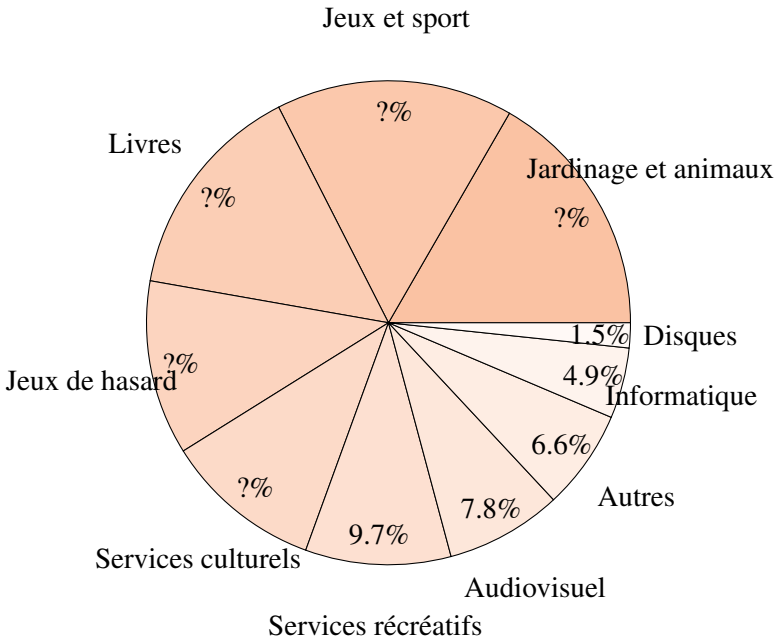
Le diagramme ci-dessous est tiré du rapport *Nourrir l'humanité à l'horizon 2050*, rédigé par la WWF (World Wildlife Fund).



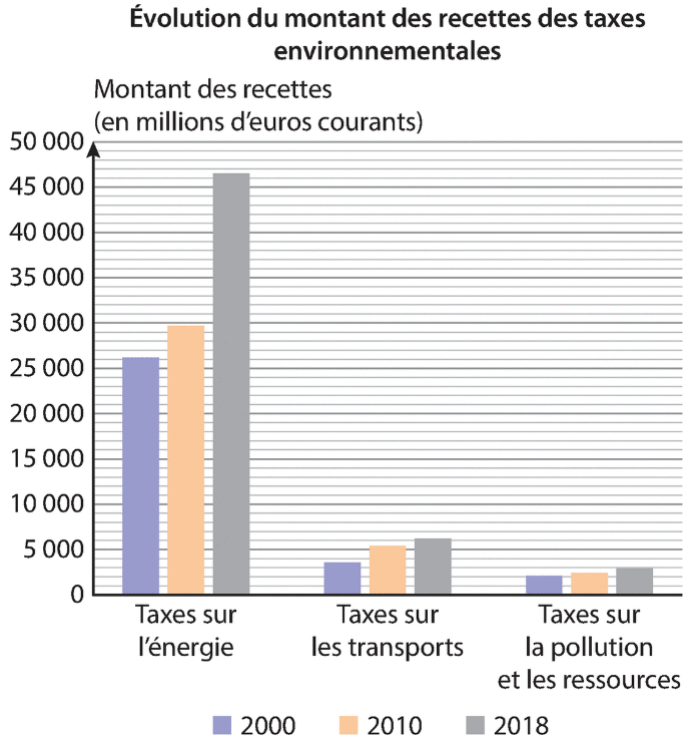
1. Vrai ou Faux ? La consommation générale de poissons a plus que triplé entre 1961 et 2011. Justifier.
2. Donner une estimation de la part de poissons d'eau douce consommée parmi l'ensemble de la consommation de poissons en 2011.

► Exercice 19

Le diagramme circulaire ci-dessous donne la part des dépenses culturelles de loisirs en France en millions d'euros.



1. A l'aide du graphique, donner une estimation du pourcentage des dépenses culturelles réservé au jardinage et aux animaux.
2. A l'aide du graphique, donner une estimation du pourcentage des dépenses culturelles réservé à l'ensemble Jardinage, jeux et sports, et livres.
3. D'après l'INSEE, les dépenses totales culturelles en France s'élevaient à 97351 millions d'euros en 2021.
   (a) Donner le montant des dépenses réservées à l'informatique.
   (b) Donner une estimation des dépenses dédiées au jardinage et aux animaux.



► Exercice 20

On considère le graphique ci-contre, publié dans Eurostat le 31 janvier 2020.

1. En 2018, donner les montant des taxes sur l'énergie, les transports et la pollution des ressource.
2. Quel type de taxes semble avoir le plus augmenté en pourcentage entre 2000 et 2018 ?
3. Vérifier votre intuition pour le calcul.
   On rappelle que pour calculer le pourcentage d'augmentation d'une valeur de départ D à une valeur d'arrivée A, on effectue le calcul
   $$\frac{A-D}{D} \times 100$$

► Exercice 21

Le Baromètre de la santé commandé par l'INVS (Institut de Veille sanitaire) fournit les données suivantes sur le prix moyen du paquet de 20 cigarettes en France, en uros, et le nombre total de paquets de 20 cigarettes vendus en France, en milliards.

Année	2004	2007	2010	2013	2016
Prix du paquet	5	5.13	5.65	6.7	7
Paquets vendus	2.75	2.75	2.74	2.38	2.25

1. Construire un nuage de points associé à ce tableau.
2. La hausse du prix a-t-elle fait diminuer le tabagisme ?

