

Fiche

I. Qu'est-ce qu'une série statistique à deux variables ?

- **Définition** : Une série statistique à deux variables est une série statistique étudiant simultanément deux caractères sur un même échantillon de n individus extraits d'une population.
- On présente cette série statistique à deux variables à l'aide d'un tableau statistique de la forme :

Valeurs du premier caractère	x_1	x_2	x_3	...	x_n
Valeurs du second caractère	y_1	y_2	y_3	...	y_n

- Exemple : On a relevé la température à différentes altitudes sur une route de montagne. On obtient les données suivantes :

Altitude (en km)	0,4	0,8	1,2	1,5	1,9	2
Température (en °C)	8,6	6,5	2,8	1,2	-1	-1,9

- Les deux caractères étudiés sont l'altitude en km et la température en °C.
- Remarques : Quand sur une population, on étudie un seul caractère dont les valeurs sont relevées à des instants différents, on forme une série statistique à deux variables où le temps est une des deux variables. Une telle série est dite chronologique.

II. Comment compléter un tableau croisé d'effectifs ?

- **Définition** : Un tableau croisé d'effectifs est un tableau à double entrée qui permet d'étudier la répartition d'une population selon deux caractères. L'un est représenté en ligne et l'autre en colonne.
- Exemple : catégorisation des Lillois.
 - Population : la ville de Lille en 2019.
 - Individus : habitants.
 - Premier caractère étudié : sexe.
 - Second caractère étudié : âge.

Source : <https://www.insee.fr/fr/statistiques/2011101?geo=COM-59350>

	Hommes	Femmes	Total
Strictement moins de 30 ans	58 382	61 775	120 157
Plus de 30 ans	54 835	59 483	114 318
Total	113 217	121 258	234 475

- Il y avait $\frac{59\,483}{234\,475}$ Lilloises âgées de plus de 30 ans parmi les 234 475 habitants de Lille en 2019. Cela représente une proportion de $\frac{59\,483}{234\,475}$ soit environ 25,37 % (ou presque un quart des habitants).

 [Exercice n°1](#)

 [Exercice n°2](#)

III. Comment représenter graphiquement des données de manière pertinente ?

Diagramme en barres

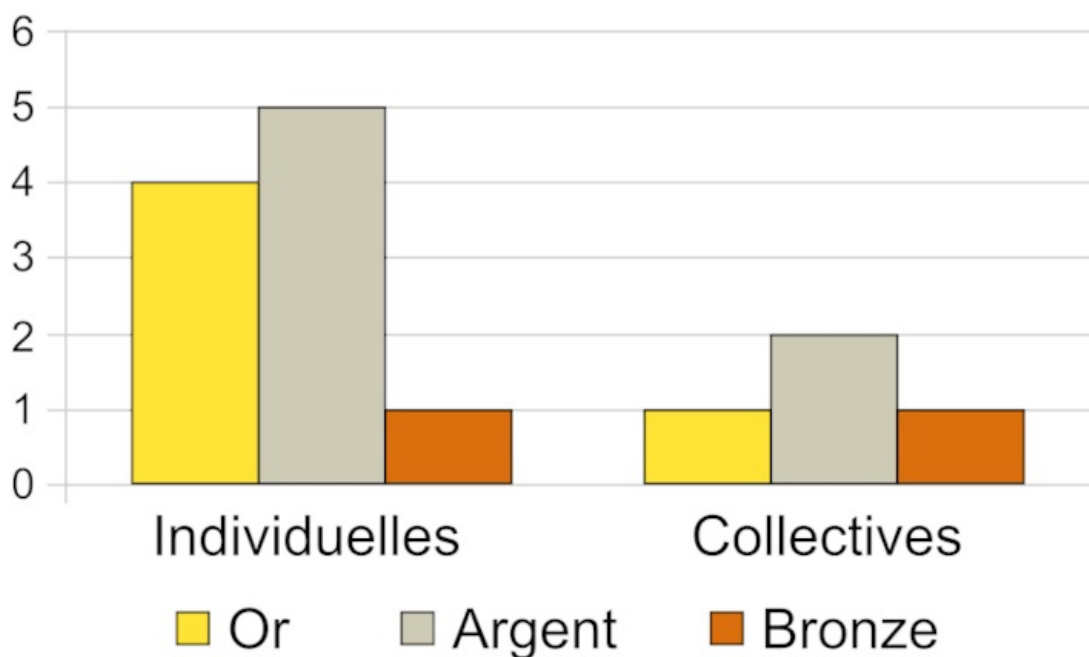
- Le premier caractère sera positionné sur l'axe des abscisses.
- Le second caractère sera positionné sur l'axe des ordonnées.
- On pourra utiliser un tableur pour construire le diagramme (créer le tableau sur une feuille de calcul, sélectionner le tableau, puis chercher le menu INSERTION et le bouton DIAGRAMME ; suivre ensuite les indications concernant les options de présentation du diagramme).

- Exemple : On considère le nombre de médailles françaises récoltées au JO de Pékin de 2022. On entre les données dans une feuille de calcul.

	A	B	C	D
1		Or	Argent	Bronze
2	Individuelles	4	5	1
3	Collectives	1	2	1

- Puis, en sélectionnant les cellules de A1 à D3 et en choisissant le type de diagramme (Diagramme en barres), et en reparamétrant les couleurs, on peut obtenir ceci :

JO PÉKIN 2022 Médailles françaises



- On observe que la France été plus performante dans les épreuves individuelles que dans les épreuves collectives.

Nuage de points

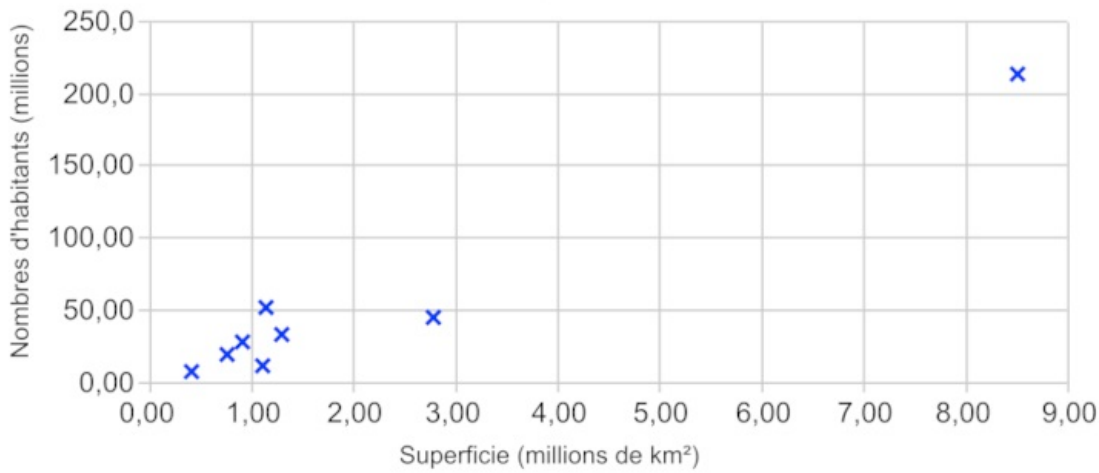
- Définition** : Dans un repère orthogonal du plan, l'ensemble des points $M(x_i; y_i)$ où l'entier i varie de 1 à n est appelé le nuage de points de la série statistique.

- Exemple : D'après le site web <https://donnees.banquemondiale.org/>, on a les données suivantes concernant quelques pays d'Amérique du Sud.

Pays	Brésil	Argentine	Pérou	Colombie	Bolivie	Venezuela	Chili	Paraguay
Superficie (millions de km ²)	8,51	2,78	1,29	1,14	1,1	0,912	0,756	0,407
Nombre d'habitants en 2021 (millions)	214	45,8	33,7	51,5	12	28,2	19,5	6,7

- On peut obtenir le nuage de points suivant à l'aide d'un tableur :

Pays d'Amérique du Sud 2021 - Banque Mondiale



- On peut supputer une corrélation positive entre la superficie et le nombre d'habitants. En effet les points semblent placés dans une bande qui « monte ». Les deux caractères semblent liés, mais attention : on ne prouve pas de causalité entre eux.

Exercice n°3

Diagramme circulaire

- **Définition** : Un diagramme circulaire ne permet de visualiser la répartition que d'un seul caractère dans une population. Les différentes valeurs du caractère seront représentées par des secteurs angulaires (souvent de couleurs différentes) dont la taille est proportionnelle à l'effectif.

- Exemple : Observons la production d'électricité en France le jeudi 15 juin 2023 à 8 h.

Source : <https://www.rte-france.com/>



Type	Fioul	Charbon	Gaz	Hydraulique	Nucléaire	Solaire	Éolien	Bioénergies	Total
Puissance (MW)	136	17	3 257	1 014,5	30 703	2 182	1 964	668	49 072
Mesure d'angle (degré)	1	0,12	23,9	74,42	225,24	16,01	14,41	4,9	360

- L'angle plein mesure 360°. On calcule la somme des puissances produites : on trouve une production de 49 072 MW. Enfin, pour remplir la dernière ligne du tableau, on va utiliser la proportionnalité.

- Ainsi pour déterminer, par exemple, la mesure d'angle du secteur angulaire représentant la puissance produite par les bioénergies on calcule :

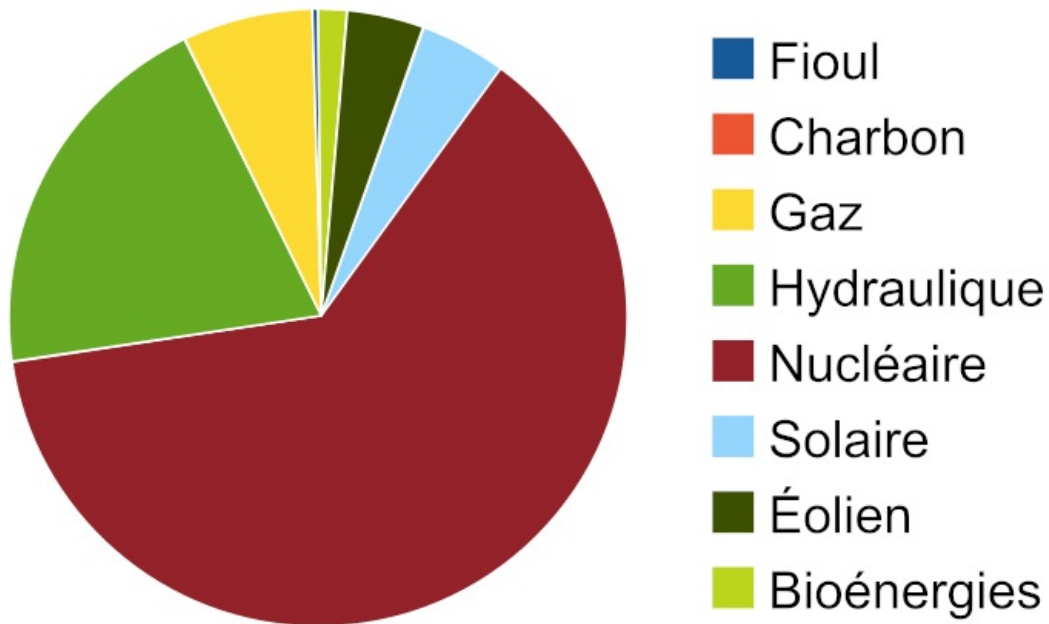
$$\frac{668 \times 360}{49\,072}$$

On trouve environ 4,9 degrés.

- Une fois le tableau complété on construit un cercle. On trace un premier rayon (souvent horizontal) de ce cercle. On construit le premier secteur angulaire en utilisant un rapporteur. Puis on construit les suivants à la suite. On obtient alors le diagramme circulaire suivant :

Production électrique française

15 juin 2023 à 8h

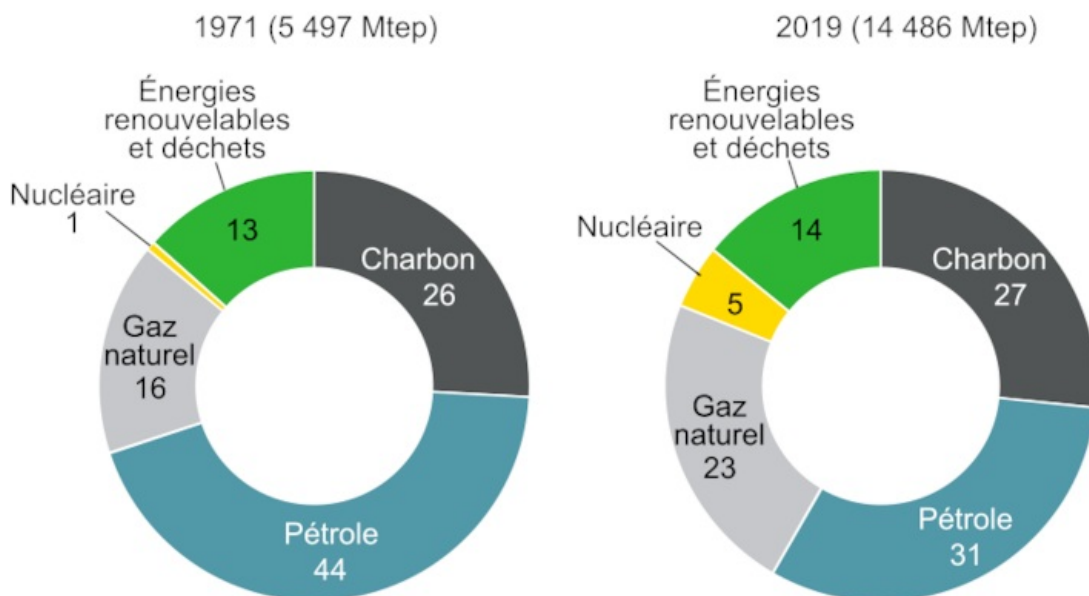


- La domination de la filière nucléaire dans la production d'électricité française apparaît clairement.

Exercice n°4

- Remarque : Pour comparer deux séries de données, il faut donc deux diagrammes circulaires.
Exemple : Source : <https://www.statistiques.developpement-durable.gouv.f>

Mix énergétique primaire dans le monde, en %



- On observe que la proportion d'énergie fabriquée à partir du pétrole a diminué dans le monde entre 1971 et 2019 (diminution de 13 points de pourcentage).

IV. Comment manipuler des données avec un tableur ?

Tri

- On peut **trier les données** : on va modifier l'ordre des objets.
Il suffit de sélectionner « Données » puis « Trier... » et de choisir un descripteur. On peut choisir un tri dans l'ordre croissant ou

décroissant. On peut même choisir d'autres descripteurs pour effectuer un tri dans un tri ! Pour un tri simple, on peut aussi directement sélectionner la colonne voulue et choisir le bouton « A->Z » pour effectuer un tri dans l'ordre croissant.

Filter

- On peut aussi **filtrer** les données, c'est-à-dire afficher les objets qui contiennent une valeur particulière. Il suffit de sélectionner « Données » puis « Filtre ».
- On peut alors choisir AutoFiltre qui va mettre en place automatiquement autant de filtres que de valeurs différentes par descripteurs.
- On peut choisir Filtre standard afin de sélectionner soi-même un descripteur particulier et de n'afficher que certaines valeurs (conditions à choisir).
- Enfin, on peut mettre en place des filtres sur toutes les colonnes ou uniquement sur certaines.
- Les outils Tri et Filtre sont particulièrement utiles lorsqu'une collection de données contient des centaines ou des milliers de lignes.

ET

- La fonction logique ET renvoie VRAI si **tous les tests** sont vrais, sinon elle renvoie FAUX.

• On écrit :
=ET(test1 ;test2 ;...)

OU

- La fonction logique OU renvoie VRAI si **au moins un** des tests est vrai, sinon elle renvoie FAUX (cela signifie que tous les tests sont faux).

• On écrit :
=OU(test1 ;test2 ;...)

• Exemple : On considère la feuille de calcul suivante qui indique le montant des ventes réalisées par les différents salariés d'une entreprise durant les quatre trimestres de l'année 2022. Le directeur souhaite attribuer une prime pour encourager ses employés.

- Idée 1 : une prime est accordée si les montants trimestriels sont tous supérieurs à 700 €.
- Idée 2 : une prime est accordée si au moins un montant trimestriel est supérieur à 1 000 €.

	A	B	C	D	E	F	G
1	Salarié	T1	T2	T3	T4	Année	Prime
2	Jean	551	964	1293	1355	4163	
3	Kate	1456	1297	743	1476	4972	
4	Léonard	1097	742	1270	820	3929	
5	Corentin	1272	1024	1202	1402	4900	
6	Amandine	1201	973	1476	1371	5021	
7	Lison	1154	699	527	1055	3435	

• Pour déterminer les salariés ayant droit à la prime si l'idée 1 est retenue, alors il faudrait entrer la formule suivante dans la cellule G2 :

=ET(B2>700 ; C2>700 ; D2>700 ; E2>700)

• On utilisera ensuite la poignée de recopie en l'étirant vers le bas jusqu'à la cellule G7.

• Pour déterminer les salariés ayant droit à la prime si l'idée 2 est retenue, alors il faudrait entrer la formule suivante dans la cellule G2 :

=OU(B2>1000 ; C2>1000 ; D2>1000 ; E2>1000)

• On utilisera ensuite la poignée de recopie en l'étirant vers le bas jusqu'à la cellule G7.

Exercice n°5

NON

- La fonction logique NON renvoie FAUX si le test est vrai, sinon elle renvoie VRAI.

• On écrit :
=NON(test)

• Remarque : Cette fonction n'est pas très utile lorsque le test est simple. En effet, écrire =NON(A1>50) est plus long que d'écrire simplement =A1<=50.

