

Partie 1 : Statistique et probabilités

Chapitre 1 : Organiser les données

- ❖ Qu'est-ce qui vous impressionne le plus : un saut à l'élastique ou un tour en voiture de rallye ?
- ❖ Les statistiques montrent-elles un lien entre le nombre de naissances d'enfants et la population de cigognes ?
- ❖ La fonte des glaces menacerait-elle beaucoup de villes de françaises ?

1 Activités

1.1 Extrêmes sensations

Une enquête est menée auprès d'adolescents. On leur pose la question suivante : « A ton avis, quelle activité te permettrait de trouver ta propre extrême sensation ou plutôt de te provoquer une poussée d'adrénaline ? »
Voici les résultats obtenus.

 Speed riding	135	 Via ferrata	90
 Parapente	315	 Kitewing	360
 Saut à l'élastique	900	 Télési nautique	45
 Canyoning	720	 Acrobranche	270
 Sport automobile	1 530	 Rafting	135

Questions :

1. Auprès de quelle population est menée cette enquête ? **Auprès d'adolescents**
2. Quel est le sujet de cette enquête **l'activité jugée la plus sensationnelle par les adolescents**
3. Combien de personnes ont été interrogées ? **4 500**
4. Comment pourrait-on regrouper les réponses pour réaliser une étude statistique de cette enquête ?
On pourrait regrouper les activités en catégorie, par exemple : sports de vol, d'eau, escalade, sport automobile
5. En utilisant les résultats de l'enquête, compléter la 2^e et la 3^e colonne du tableau.

Catégories	Effectif	Fréquence (en %)	Angle (en °)
Sports de vol	1 710	38	137
Sports d'eau	900	20	72
Escalade	360	8	29
Sports automobile	1 530	34	122
TOTAL	4 500	100	360

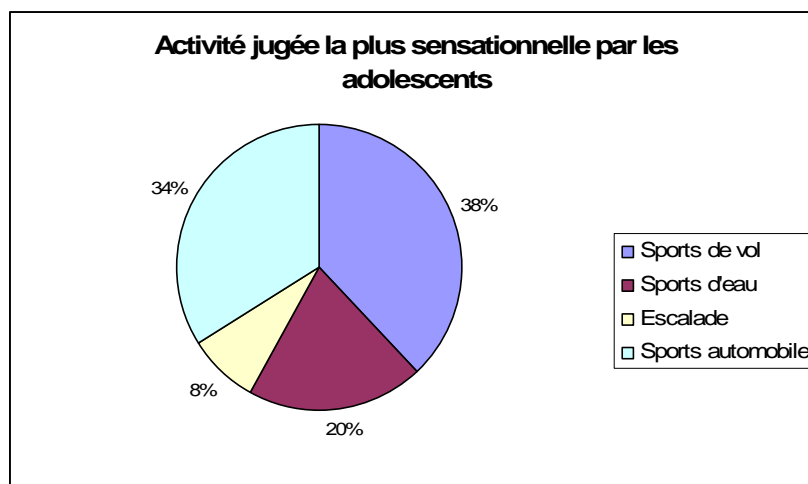
$$\frac{1710}{4500} \times 100 = 38$$

6. Comment doit-on compléter la 4^e colonne du tableau pour présenter les résultats de l'enquête sous forme d'un diagramme à secteur ou d'un « camembert » ?

Pour chaque catégorie, on calcule l'angle en degré correspondant. Un cercle = 360°;

2 méthodes : $\frac{1710}{4500} \times 360 \approx 137^\circ$ ou $\frac{38}{100} \times 360 \approx 137^\circ$

7. Représenter cette enquête sur le camembert. Attention, ne pas oublier la légende.

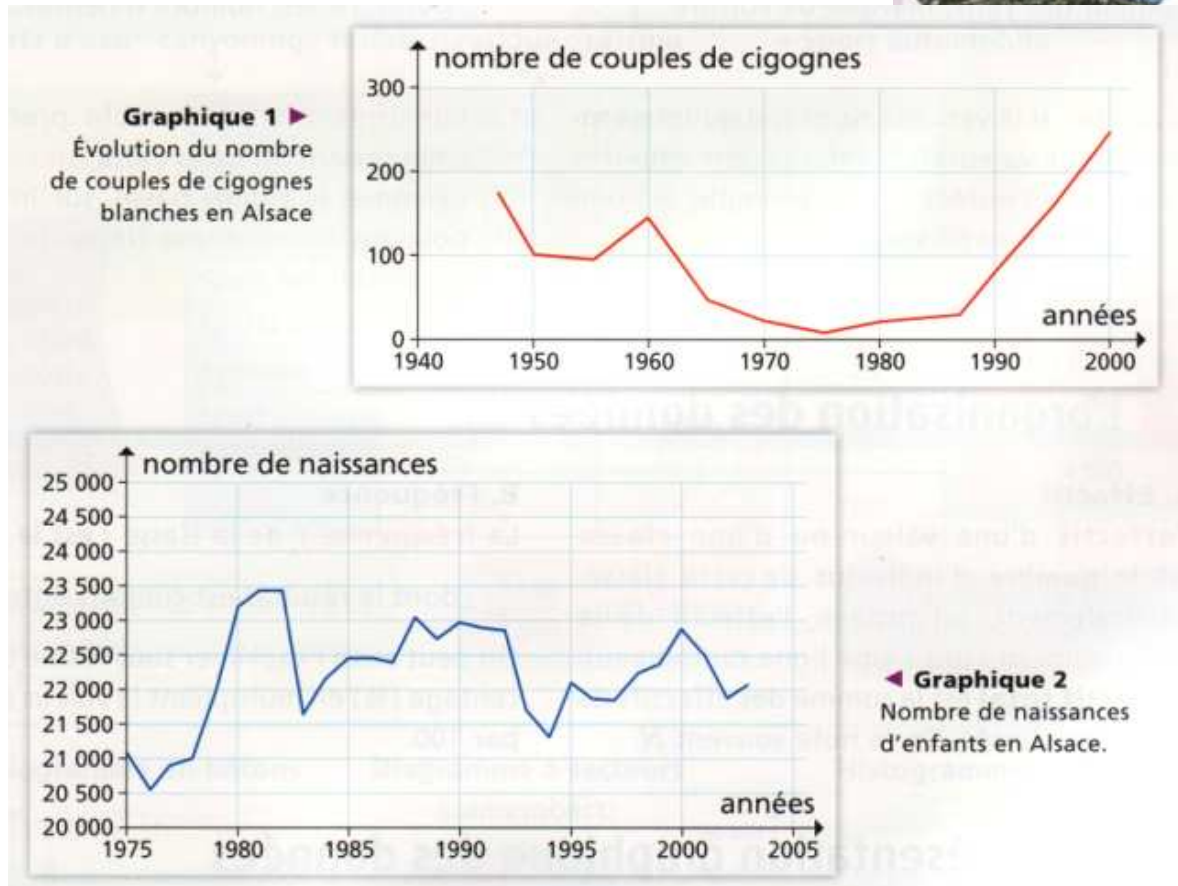


8. Quelle activité est jugée la plus « sensationnelle » par 1 adolescent sur 5 ?

$\frac{1}{5} = 0.2 = 20\%$ **donc cela correspond aux sports d'eau**

1.2 Les cigognes en Alsace

Selon la légende, la cigogne est connue pour apporter les bébés. Les futurs mamans doivent passer commande en mettant quelques morceaux de sucre sur le rebord de leur fenêtre. L'oiseau va alors chercher l'enfant auprès d'une source ou d'une mare, là où les lutins ramènent des profondeurs de la terre les âmes tombées du ciel avec la pluie, et réincarnées en nouveaux nés.



Questions :

1. Quel est le thème du graphique 1 ? Du graphique 2 ?

Le graphique 1 représente l'évolution du nombre de couples de cigognes en Alsace. Le graphique 2 représente l'évolution du nombre de naissances d'enfants en Alsace

2. D'après le graphique 1, en quelle année l'Alsace a-t-elle accueilli le plus grand nombre de cigognes ?

En 2 000

3. D'après le graphique 2, cette même année correspond-elle à un pic des naissances d'enfants en Alsace ?

Oui

4. Expliquer pourquoi la légende de la cigogne amenant les bébés peut être crédible cette année là.

Parce qu'en 2 000, on observe à la fois un pic du nombre de couples cigognes et un pic du nombre de naissances d'enfants

5. Peut-on faire la même remarque pour l'année 1980 ? **Non, il y a très peu de cigognes alors qu'il y a un pic du nombre de naissances d'enfants.**

6. Quel est le caractère étudié dans le premier graphique ?

Le caractère étudié est le nombre de couples de cigognes en Alsace.

7. Quel est le caractère étudié dans le second graphique ?

Le caractère étudié est le nombre de naissances d'enfants en Alsace.

Le caractère étudié est :

Graphique 1 qualitatif **quantitatif**

Graphique 2 qualitatif **quantitatif**

8. Comment pourrait-on modifier ces graphiques pour comparer plus efficacement les données ?

On pourrait d'abord comparer les deux graphiques sur une même période : de 1975 à 2000 par exemple.

9. On s'intéresse plus particulièrement aux périodes 1980 et 2000. Voici quelques chiffres issus des données des deux graphiques.

Années	Nombre de couples de cigognes	Nombre de naissances en Alsace
1980	20	23250
1981	20	23400
1982	20	23400
1983	20	21700
1984	20	22150
1985	20	22400
1986	20	22400
1987	30	22350
1988	50	23100
1989	60	22800

Années	Nombre de couples de cigognes	Nombre de naissances en Alsace
1990	80	23000
1991	95	22900
1992	100	22900
1993	110	21650
1994	140	21400
1995	170	22100
1996	190	22400
1997	200	22400
1998	220	22300
1999	230	22400
2000	250	22900

À partir des chiffres des tableaux ci-dessus, compléter les tableaux ci-dessous.

Années	Effectifs : nombre de couples de cigognes	Fréquences en% Arrondi à 0.1
[1980 ; 1985[100	4.8
[1985 ; 1990[180	8.7
[1990 ; 1995[525	25.4
[1995 ; 2000]	1260	61
TOTAL	2065	100
Années	Effectifs : nombre de naissances en Alsace	Fréquences en % Arrondi à 0.1
[1980 ; 1985[113 900	24,1
[1985 ; 1990[113 050	23,8
[1990 ; 1995[111 850	23,6
[1995 ; 2000]	112 100	23,7
TOTAL	473 300	100

10. À partir de ces données :

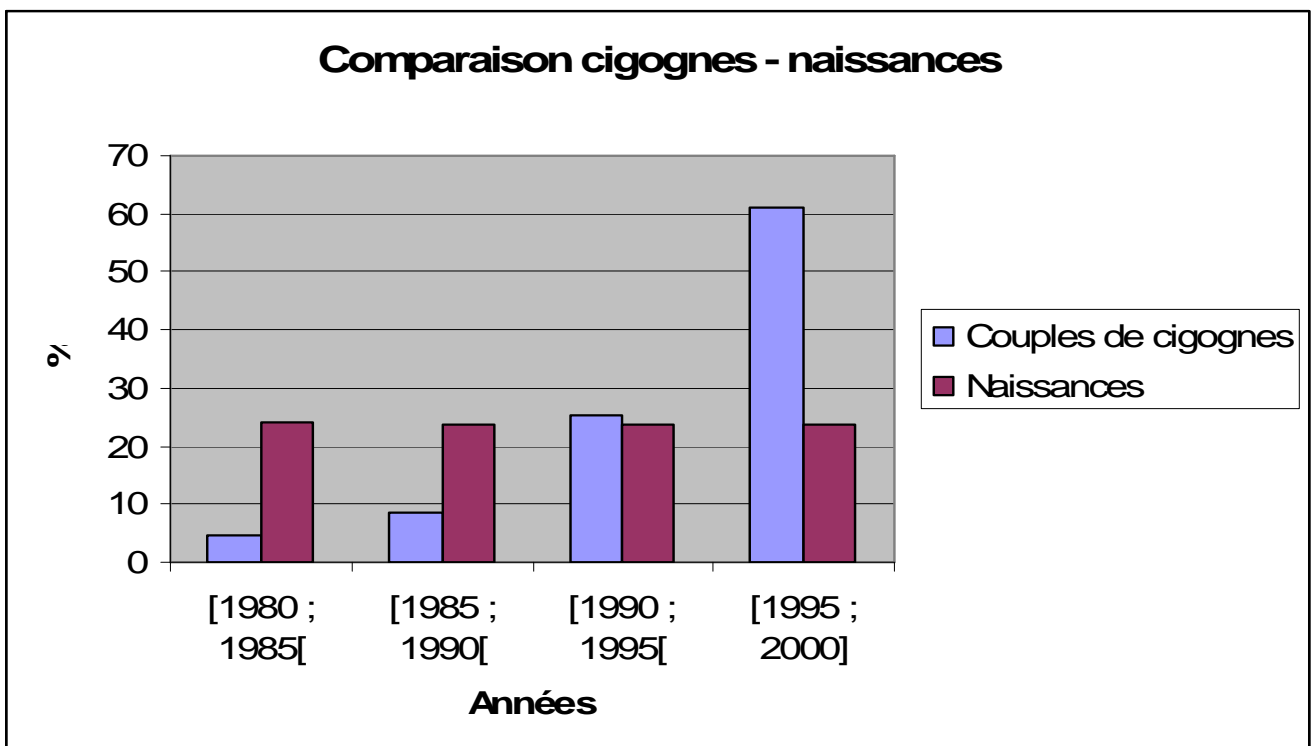
a) Quel est le pourcentage de couples de cigognes ayant niché en Alsace entre 1980 et 1990 ?

13,5 %

b) Quel est le pourcentage d'enfants nés en Alsace entre 1980 et 1990 ?

47,9 %

c) Construire, sur le même graphique, les histogrammes des fréquences correspondants aux deux séries statistiques.



d) Quelle observation peut-on faire en comparant ces deux histogrammes ?

Le nombre de couples de cigognes varie mais le nombre de naissances d'enfants est stable.

e) En conclusion, qu'en est-il de la légende « La cigogne est connue pour apporter les bébés » ?

A partir des graphiques, on voit qu'il n'y a pas de lien entre le nombre de couples de cigognes et le nombre de naissances d'enfants. Donc la légende est infondée.

2 L'essentiel

2.1 Le vocabulaire

La statistique s'attache à étudier un certain nombre de paramètres : **caractères** ou **variables** d'un ensemble fini appelé **population**.

Les éléments de cette population étudiée sont appelés **individus**.

Exemple : dans l'activité « Extrêmes sensations » :

- La population est : les adolescents
- Le caractère ou variable étudié est : **l'activité sportive**.

Une variable peut être :

- **qualitative** : fait référence à une qualité (couleur des yeux, marque de voiture ...)
- **quantitative** : numérique (âge, taille, poids, notes, nombre d'heures....)

Si la variable est quantitative,

elle est quantitative **discrète** ou quantitative **continue**

La variable ne prend qu'un nombre fini de valeurs, ces valeurs sont appelées modalités et notées x , par exemple le nombre de frères et sœurs...	La variable prend ses valeurs dans un intervalle appelé classe , par exemple, le temps passé, sur Internet pour les élèves de 2 nd e.
---	---

- Exemple : [1980 ; 1985[est un **intervalle** ou une **classe** en statistique.
- Le caractère qui a le plus grand effectif s'appelle le **mode** La classe qui a le plus grand effectif s'appelle la **classe modale**.

2.2 L'organisation des données

2.2.1 L'effectif

L'effectif d'une valeur ou d'une classe est le nombre d'individus de cette classe.

Généralement, on note n_i l'effectif de la classe numéro i (ou i -ème ligne du tableau). L'effectif total est

la somme des effectifs de toutes les classes. On le note souvent **N**.

2.2.2 La fréquence

La fréquence f_i de la classe i est le rapport $\frac{n_i}{N}$ dont le résultat est compris entre 0 et 1. On peut aussi

l'exprimer sous forme de pourcentage (%) en multipliant la valeur obtenue par 100.

Exemple : dans l'activité « Extrêmes sensations », pour simplifier l'étude, on choisit de regrouper les différents sports en sous-catégories.

Tableau 1

Catégorie x_i	Effectif n_i (nombre d'adolescents)	Fréquence f_i (en %)	Angle (en°)
Sports de vol	1 710	38	137
Sports d'eau	900	20	72
Escalade	360	8	29
Sport automobile	1 530	34	122
Total	4 500	100	360

Calcul : $\frac{1710}{4500} \times 360$

1. En reprenant les données de l'activité 1, compléter la colonne effectif.

Compléter N = **4 500**

2. Calculer et compléter la colonne fréquence. Détailler ici le calcul de la fréquence de la catégorie « Sport automobile ».

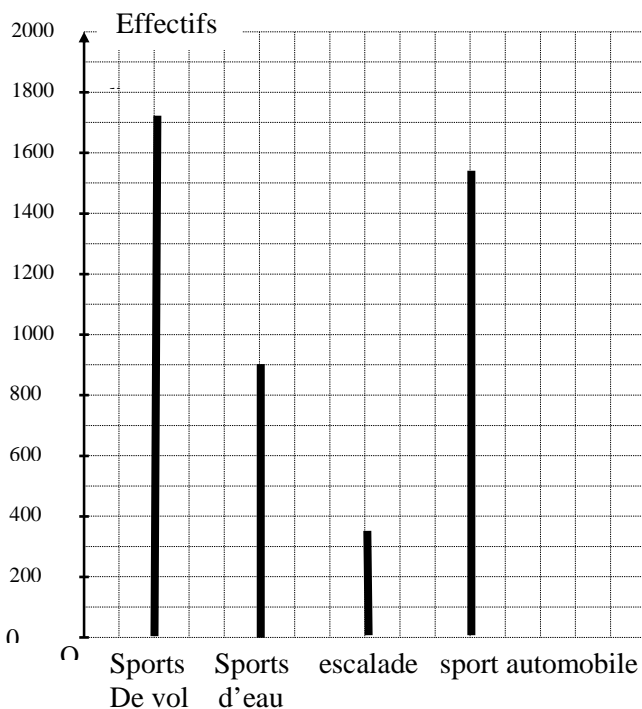
$$\frac{1530}{4500} \times 100 = 34$$

2.3 La représentation graphique des données

2.3.1 Le diagramme en bâtons

Le diagramme en bâtons est adapté aux cas **qualitatifs** et **quantitatif discret**. La hauteur de chaque bâton (un trait suffit) est **proportionnel** à l'effectif (ou à la fréquence) de la modalité correspondante.

Exemple : d'après le tableau de l'exemple précédent, construire ci-dessous le diagramme en bâtons des effectifs.



1. D'après le diagramme, déterminer le mode.

Le mode est : « sports de vol ».

2. Quel est le caractère étudié ?

Le sport qui procure le plus de sensations.

Quelle est la nature de ce caractère ?

Le caractère est qualitatif.

2.3.2 Le diagramme à secteurs ou camembert

Le « camembert » est adapté à l'étude de il est surtout efficace lorsque les variables sont

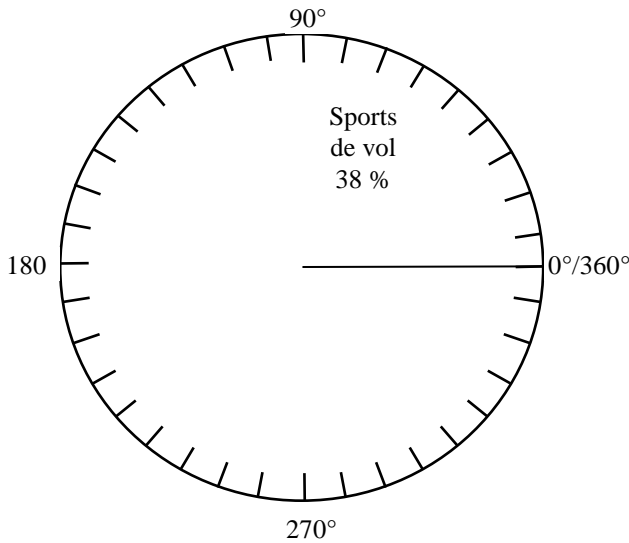
Les mesures des angles de chaque secteur sont à l'effectif (ou à la fréquence).

Un diagramme **circulaire** est composé d'un cercle qui fait au total °

Un diagramme **semi-circulaire** est d'un demi-cercle qui fait au total °

Exemple :

1. Compléter la colonne angle du tableau 1.
2. Construire ci-dessous le diagramme à secteurs correspondant.



3. Déterminer le mode de cette série statistique.

.....

4. Au départ de l'activité « Extrêmes sensations », combien y'a-t-il de sports différents ?

Etait-il judicieux de représenter ces sports directement (sans les regrouper) dans un diagramme circulaire ? Pourquoi ?

.....

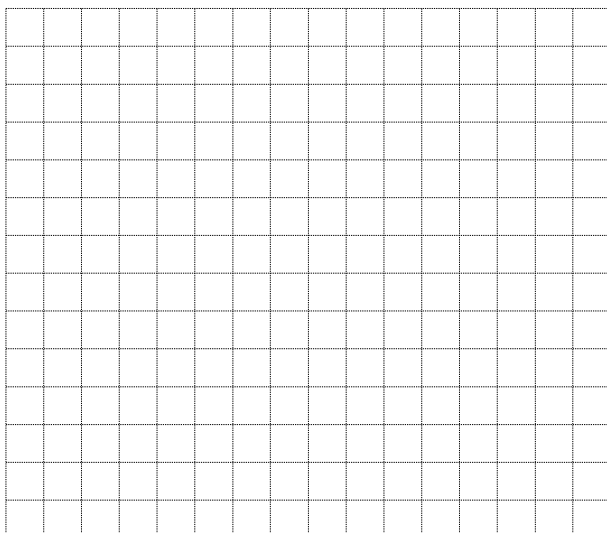
2.3.3 L'histogramme

L'histogramme est adapté à l'étude d'un caractère

Les aires des rectangles sont aux effectifs ou aux fréquences.

Exemple :

1. A partir des données du tableau de la page 5 (activité « Les cigognes en alsace ») construire ci-dessous l'histogramme.



2. Déterminer la classe modale.

.....

3. Compléter la phrase :

C'est entre 1995 et 2000 qu'il y a eu

.....

.....

3 Exercices

1 On réalise une étude sur l'équipe de France de Football en créant des fiches d'identité de chaque joueur, comme sur le modèle ci-contre.

1. Dans cette étude statistique, quelle est la population ?
2. Combien de caractères sont étudiés ?
3. Le caractère « nombre de clubs fréquentés » est-il qualitatif ou quantitatif ?
4. Citer trois caractères qualitatifs dans cette fiche.
5. Parmi les 6 caractères quantitatifs, lequel pourrait être considéré comme continu ?

Nom : BENZEMA
 Prénom : Karim
 Date de naissance : 19 décembre 1987
 Département de naissance : Rhône (69)
 Taille : 1,83 m
 Poids : 73 kg
 Poste : Attaquant
 Numéro : 10
 Nombre de sélections en équipe de France : 16
 Nombre de buts marqués en équipe de France : 4
 Club actuel : Olympique Lyonnais
 Nombre total de clubs fréquentés : 1



Vocabulaire et représentation

2 Une enquête est menée auprès de 2 500 jeunes de 13 à 16 ans sur leurs moyens de communication :

- 54 % utilisent prioritairement les SMS sur leur téléphone portable ;
- 27 % utilisent MSN ;
- 7 % utilisent la webcam ;
- 2 % utilisent les mails ;
- 10 % utilisent le téléphone.

1. Quelle est la population de cette série statistique ?
2. Quel est le caractère étudié ? Donner sa nature.
3. Construire un diagramme en bâtons permettant de représenter la répartition des effectifs de cette série statistique.

3 Indiquer, pour chaque variable, si elle est quantitative ou qualitative.

1. Les marques des voitures garées sur un

parking de supermarché.

2. La nationalité des touristes visitant le Mont-Saint-Michel.
3. L'âge des auditeurs de FUN Radio.
4. Les températures matinales relevées chaque jour sous abri à Avignon.

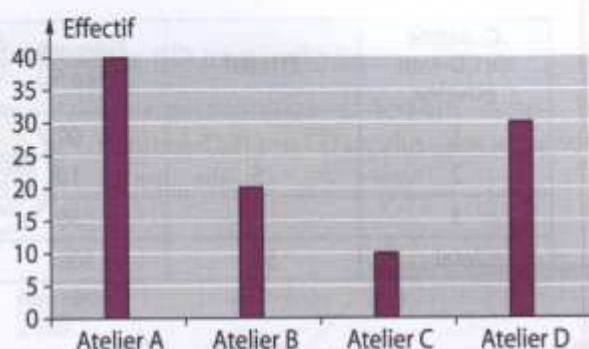
4 Un établissement de transfusion sanguine dresse le bilan de sa collecte de sang pour l'année écoulée.

Âge du donneur	Pourcentage
De 10 à 20 ans	4 %
Entre 20 et 30 ans	14 %
Entre 30 et 40 ans	24 %
Entre 40 et 50 ans	32 %
Entre 50 et 60 ans	26 %

1. Déterminer la population, le caractère étudié et sa nature.
2. À partir du tableau, représenter cette série statistique par un histogramme.

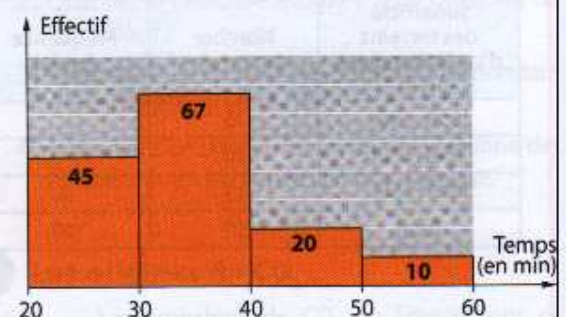
5 Fiche 2 p.141

Construisez le tableau statistique donnant la répartition des 100 employés dans les quatre ateliers (A, B, C et D), représentée par le diagramme en bâtons ci-dessous.



6 Fiche 2 p.141

L'histogramme de la série ci-dessous représente le nombre de pièces fabriquées en fonction du temps de fabrication.



- a. Calculez l'amplitude des classes.
- b. Calculez le nombre total de pièces fabriquées.

8 Fiches 2, 3 et 15 p. 141, 142 et 153



Complétez le tableau et construisez le diagramme en bâtons représentant cette série statistique.

Remarque : le diagramme en bâtons peut également être obtenu à l'aide d'une calculatrice graphique

Quantité de CD-rom achetée	Effectif	Fréquence (en %)
1		60
2	5	10
3		30
Total	50	100

9 Fiches 2, 3 et 14 p. 141, 142 et 153



Complétez le tableau et construisez l'histogramme représentant cette série statistique.

Remarque : l'histogramme peut également être obtenu à l'aide d'une calculatrice graphique.

Superficie des terrains d'un lotissement (en m ²)	Nombre de terrains	Fréquence (en %)
[500 ; 550[25	
[550 ; 600[10	
[600 ; 650[15	30
Total	50	100

10 Fiche 2 p. 141

Associez à chaque tableau une ou plusieurs représentations graphiques. Justifiez.

Tableau 1

Caractère	Effectif
A	10
B	20
C	30
D	40

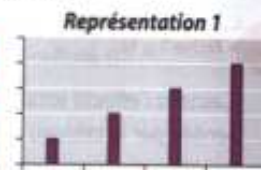


Tableau 2

Caractère	Effectif
1	10
2	20
3	30
4	40

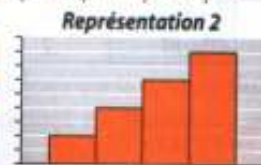
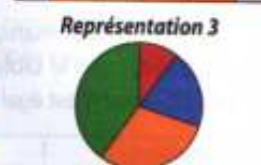


Tableau 3

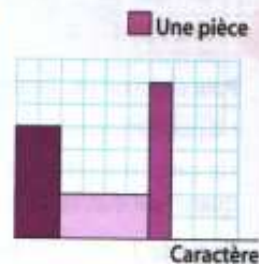
Caractère	Effectif
[0 ; 50[10
[50 ; 100[20
[100 ; 150[30
[150 ; 200[40



11 Fiche 2 p. 141

Complétez le tableau à partir des aires des rectangles de l'histogramme.

Longueur d'une pièce (en cm)	Effectif
[0 ; 10[
[10 ; 30[
[30 ; 40[
Total	



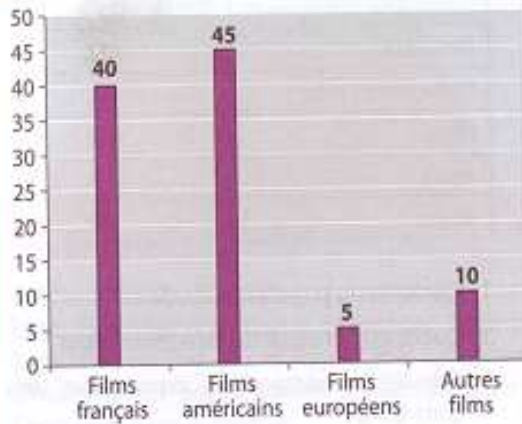
Testez-vous !

Parmi les trois réponses proposées, vous devez en choisir une seule.

- | | | | |
|---|---|---|--|
| 1. Le nombre de pièces dans une maison est | A. qualitatif | B. quantitatif discret | C. quantitatif continu |
| 2. Un diagramme en barres est choisi lorsque le caractère est | A. qualitatif | B. quantitatif discret | C. quantitatif continu |
| 3. Un histogramme est choisi lorsque le caractère est | A. qualitatif | B. quantitatif discret | C. quantitatif continu |
| 4. Le caractère étudié lors d'une élection municipale est | A. qualitatif | B. quantitatif discret | C. quantitatif continu |
| 5. Le centre de classe est | A. la différence entre les bornes de l'intervalle | B. la demi-somme des bornes de l'intervalle | C. la somme des valeurs appartenant à l'intervalle |
| 6. La fréquence est | A. la proportion en pourcentage | B. le nombre d'individus | C. le nombre de valeurs du caractère |

12 Le cinéma en France

Le diagramme en bâtons ci-dessous représente les parts de marché en pourcentage des 200 millions de tickets d'entrée au cinéma en France en 2009.



1. Quelle est la population ? Quel est le caractère étudié ? Précisez, en le justifiant, la nature du caractère.
2. Complétez le tableau statistique suivant.

Origine	Effectif n_i	Fréquence (en %)
Française		
Américaine		
Européenne		
Autres		
Total		

3. Tracez le diagramme en bâtons des fréquences. Présente-t-il des ressemblances avec le diagramme des effectifs ? Pourquoi ?

13 Le football et l'argent

Une étude statistique portant sur l'origine des recettes de l'industrie du football mondial, a donné les résultats ci-dessous.

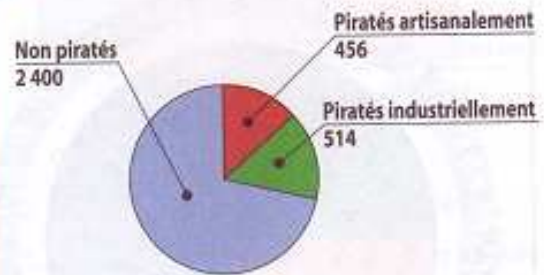
On estime les recettes annuelles de ce secteur d'activité à 165 milliards d'euros.

Origine des recettes	Proportion (en pourcentage)
Spectateurs	16
Subventions	4
Sponsors et publicité	24
Droits Télévision	56

1. Quelle est la nature du caractère étudié ?
2. Calculez le montant, en milliards d'euros, pour chaque origine des recettes.
3. Représentez graphiquement cette série statistique.

14 Albums CD et piratage

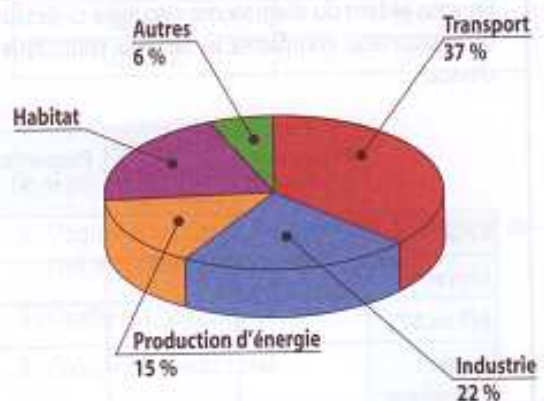
La représentation ci-dessous donne le nombre d'albums CD vendus dans le monde en milliards.



1. Quelle est la population étudiée ?
2. Quel est le caractère étudié ? Est-il qualitatif ou quantitatif ?
3. Construisez le tableau statistique correspondant à ce diagramme en secteurs.
4. Complétez ce tableau en ajoutant la colonne des fréquences en exprimées en pourcentage.

15 Les origines du CO₂

Les émissions de CO₂ en France sont de 350 millions de tonnes. Le diagramme en secteurs ci-dessous en donne la distribution.



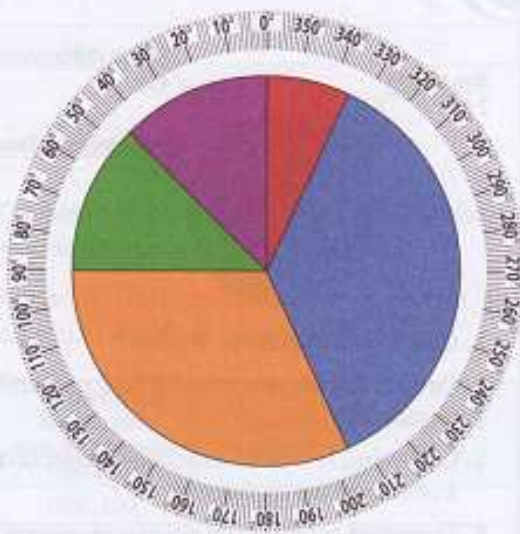
1. Quelle est la population étudiée ?
2. Quelle est la nature du caractère étudié ?
3. Calculez le pourcentage du poste **habitat**.

- Calculez la quantité de CO_2 en millions de tonnes due au transport.
- Construisez le tableau statistique donnant les quantités de CO_2 en millions de tonnes pour toutes les autres origines.

16 Après le Bac



Les statistiques nationales montrent que sur 1 000 bacheliers, 877 continuent leurs études dans l'enseignement supérieur.



- Classes préparatoires aux grandes écoles
- Université
- IUT et STS
- Autres formations
- Arrêt des études

En vous aidant du diagramme circulaire ci-dessus et du rapporteur, complétez le tableau statistique ci-dessous.

	Proportion (en degrés)	Nombre de bacheliers pour 1 000	Proportion (en %)
CPGE			
Université			
IUT ou STS			
Autres formations			
Arrêt des études			
Total			

17 Les exportations européennes



Le montant des exportations de l'Union européenne s'élève à 1 105 milliards de dollars.

La répartition des clients de L'UE, en pourcentage des exportations, est donnée dans le tableau ci-dessous.

Client	Proportion des exportations (en %)
États-Unis	22
Union européenne	34
Chine	4
Japon	4
Autres	36

- Quelle est la population étudiée ?
- Quelle est la nature du caractère étudié ?
- Calculez le montant des exportations vers les États-Unis.
- Représentez cette série par un diagramme semi-circulaire.

18 Le lancer de poids

Le tableau ci-dessous donne les résultats obtenus par un groupe d'élèves au lancer de poids en EPS.

Distance (en m)	Nombre d'élèves	Fréquence (en %)
[2 ; 3[15	
[3 ; 4[23	
[4 ; 5[54	
[5 ; 6[60	
[6 ; 7[32	
[7 ; 8[10	
Total		

- Quelle est la population étudiée ?
- Quelle est la nature du caractère étudié ?
- Quelle est l'amplitude de classe ?
- Complétez le tableau ci-dessus.
- Représentez cette série par un histogramme. Échelle : 1 cm pour 1 m ; 1 cm pour 10 élèves.

19 Les fournisseurs d'accès à Internet



L'ADSL (*Asymmetric Digital Subscriber Line*) permet, à partir d'une ligne téléphonique, de transmettre et recevoir des signaux numériques à des débits élevés. Cette technologie est mise en œuvre par des fournisseurs d'accès à Internet dits « haut-débit ».