

DOSSIER 1 - Notions de statistiques

1 Exploration

1.1 Se rendre à l'école

On a réalisé une enquête sur la mobilité des élèves du secondaire en Wallonie. Voici, sous forme de graphes, les résultats obtenus avec les réponses des élèves d'une école qui ont répondu au questionnaire.

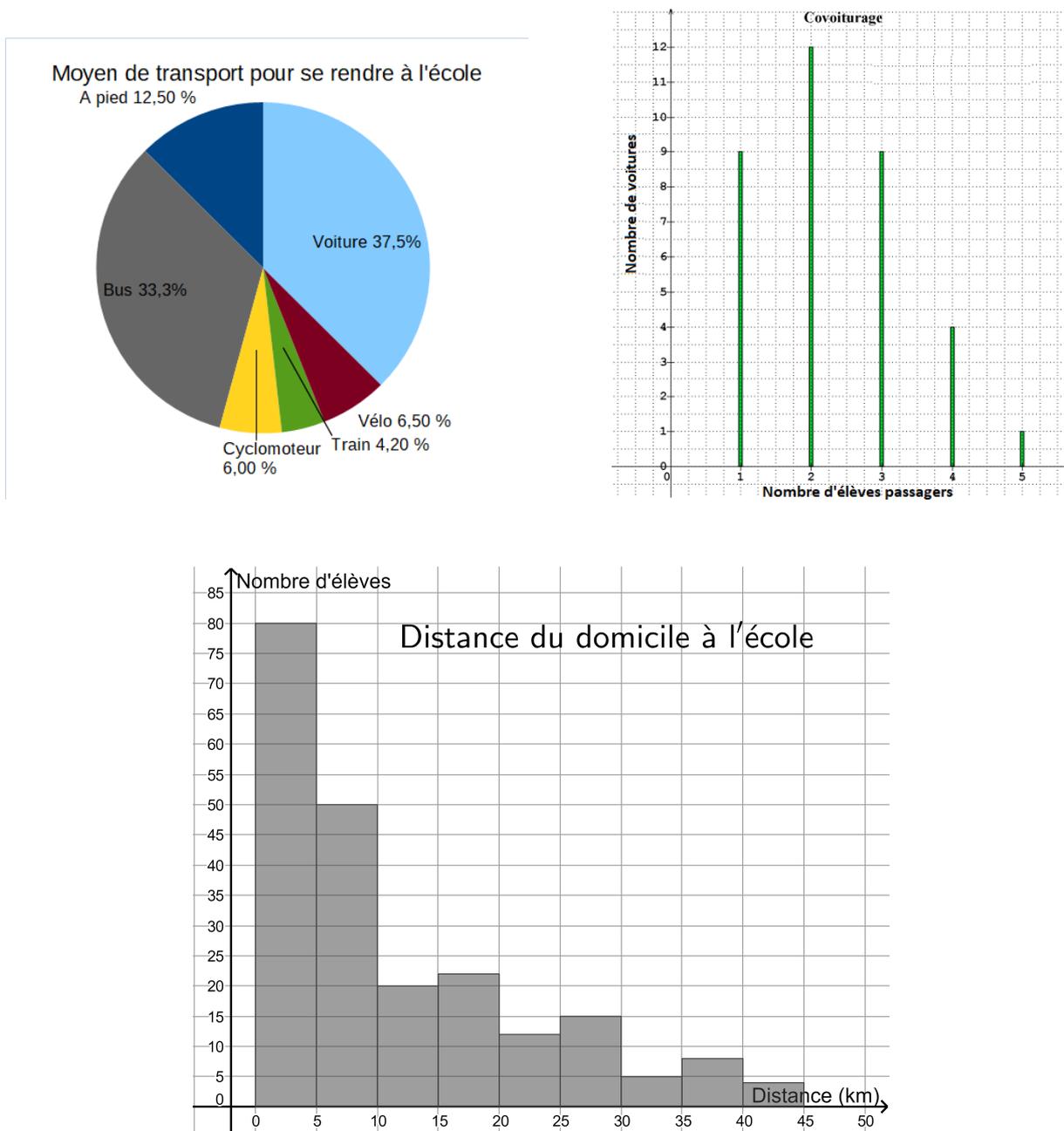


FIGURE 1 – Graphiques sur la mobilité des élèves d'une école

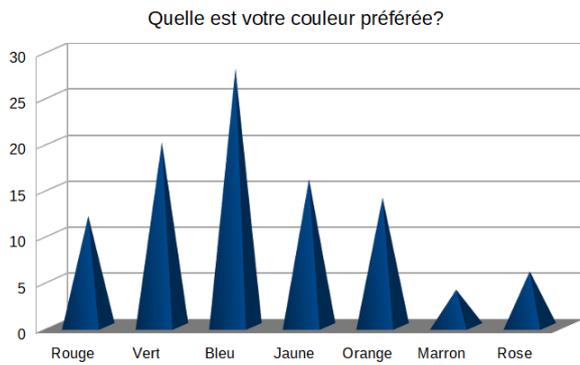
Sur base des graphique de la figure 1, réponds aux questions suivantes :

1. Sur quelle population porte l'étude réalisée ?
2. Vous n'avez, ici, que les résultats d'un échantillon de cette population. Quel est l'effectif de cet échantillon ? Autrement dit, combien d'élèves ont-été interrogés ?
3. Quelles sont les différentes variables statistiques étudiées ? Pour chacune d'elle, donnez son type (qualitative, quantitative discrète, quantitative continue) ainsi que les différentes modalités (les valeurs possibles) rencontrées.
4. Quel est le pourcentage des élèves qui vont à l'école à pied ou à vélo ?
5. Parmi les élèves qui vont à l'école en voiture, combien effectuent du covoiturage ?
6. Quel est le nombre total des élèves qui vont à l'école en voiture ?
7. Quel est la fréquence (le pourcentage) d'élèves qui habitent entre 20 et 25 km de l'école ?
8. Combien d'élèves habitent à moins de 25 km de l'école ?
9. Peut-on dire que la moitié des élèves vont à l'école en transports en commun ?
10. Récupérez les informations relatives au moyen de transport et représentez les sur un diagramme similaire à celui de la distance domicile/école. Vous aurez ainsi réalisé un histogramme.

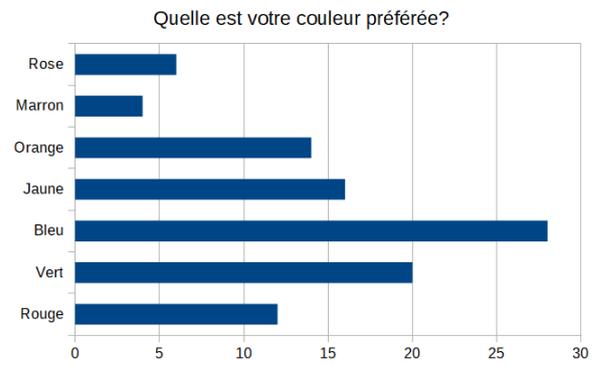
1.2 Des représentations

Sur base des graphes donnés à la figure 2, réponds aux questions.

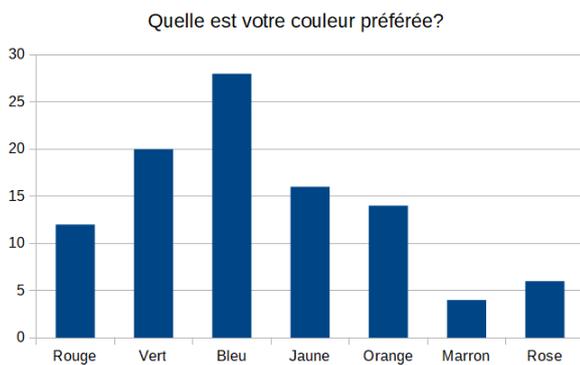
1. Quels sont les deux graphes que vous trouvez les plus beaux ? Pour quelle raison ?
2. Quel est l'avantage du graphe f par rapport au graphe e ?
3. Quel est l'avantage du graphe e par rapport au graphe f ?
4. Quel est le graphe qui n'a pas de sens pour un mathématicien ?
5. Quelle différence y a-t-il entre les graphes b et c ? Lequel préférer ? Justifiez.
6. Quels sont les deux graphes qui correspondent le plus aux exigences d'un mathématicien ? Justifiez le plus complètement possible.



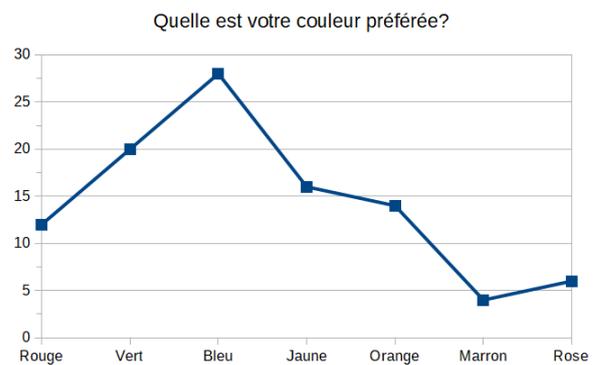
a



b



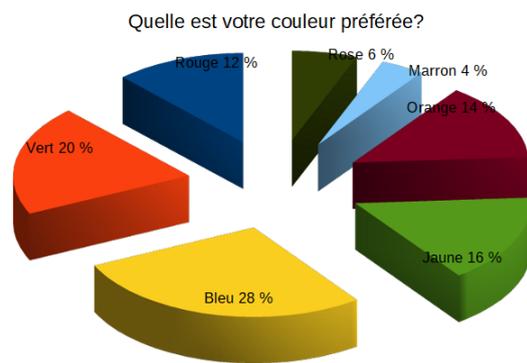
c



d



e



f

FIGURE 2 – Différents graphiques présentant les mêmes données

2 Savoirs

2.1 Définitions

Depuis des siècles, l'homme a souvent tenu une comptabilité des personnes et des choses avec l'espoir d'utiliser les données du passé pour la résolution des problèmes du présent et la prévision des événements à venir. Cependant, ce n'est qu'à la fin du 19^e siècle que l'on a trouvé la méthode correcte à employer pour la résolution des problèmes de l'interprétation des données recueillies.

- La **statistique** est une science qui a pour objet la collecte, le traitement, l'interprétation et la présentation d'ensemble d'observations relatives à un même phénomène et pouvant être quantifiées.

Exemple d'étude : une enquête sur la mobilité des élèves.

- La **population** est l'ensemble des éléments auxquels se rapporte une étude statistique.

Exemple : les élèves des écoles secondaires wallonnes.

- Un **échantillon** est un ensemble d'individus représentatifs d'une population. On choisit un/des échantillon(s) lorsqu'il est impossible d'effectuer l'étude sur toute la population.

Exemple : Les élèves de l'IST Manage, de sainte-Ursule à Namur et de l'athénée d'Arlon

- L'**effectif** (total) n de la population est le nombre d'individus de cette population.

Exemple : le nombre d'élèves des écoles wallonnes. L'effectif de l'échantillon est le nombre d'élèves des 3 écoles sélectionnées pour former l'échantillon.

- La **Variable statistique** ou le caractère est la caractéristique étudiée sur une population.

Exemple : le moyen de transport pour se rendre à l'école.

- Les **modalités** x_k d'un caractère sont les différentes formes (ou valeurs) qu'il peut prendre.

Exemple : la voiture, à pied, le bus, le train, le vélo, la trottinette.

- L'**effectif** n_k d'une modalité x_k est le nombre d'individus rencontrés pour cette modalité .

Exemple : Quand 8 élèves vont en trottinette à l'école, alors l'effectif de la modalité trottinette est de 8.

- La **fréquence** f_k d'une modalité x_k est définie par $f_k = \frac{n_k}{n}$.

Exemple : Lorsque 8 élèves sur 200 vont en trottinette à l'école, alors la fréquence associée à la trottinette est de $\frac{8}{200}$ et donc de 4%.

2.2 Types de variables statistiques

Une variable statistique est **qualitative** est une variable dont les différentes valeurs ne sont pas mesurables ; le plus souvent, ces valeurs sont exprimées par des mots.

Exemple : le moyen de transport pour se rendre à l'école.

Une variable statistique **quantitative** est une variable dont les différentes valeurs sont mesurables ; ces valeurs sont exprimées par des nombres.

- Une variable quantitative **discrète** prend des valeurs isolées.

Exemple : le nombre d'élèves par voiture se rendant à l'école.

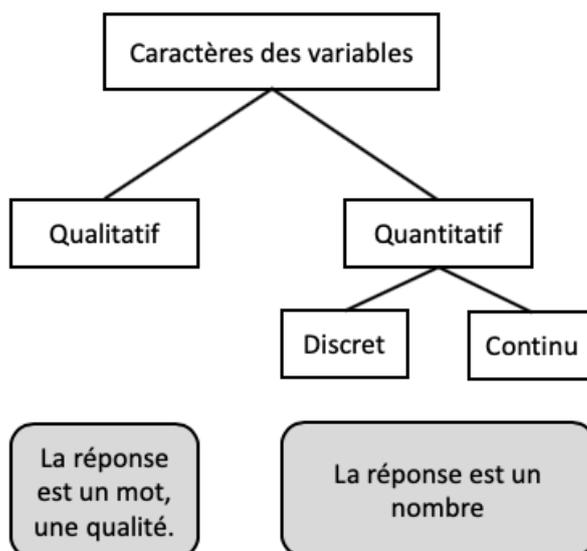
- Une variable quantitative **continue** peut prendre toute valeur d'un intervalle.

Exemple : la distance du domicile à l'école.

- Quand une variable discrète prend un grand nombre de valeurs différentes, on effectue l'étude statistiques comme si la variable était continue.

Exemple : le nombre d'amis sur FaceBook. Les valeurs sont trop nombreuses pour qu'on s'y retrouve. Nous pouvons définir des intervalles de 10, 50 ou 100 amis, par exemple.

Ces notions sont illustrées par le schéma suivant :



2.3 Graphiques

1. L'histogramme

L'histogramme permet de décrire les effectifs (ou les fréquences) observé(e)s. Il est utilisé pour présenter des données qualitatives et il a les caractéristiques suivantes :

- Chaque rectangle est associée à une valeur ou modalité de la variable.
- La hauteur d'un rectangle est proportionnelle à son effectif (ou à sa fréquence).
- La largeur des rectangles n'a pas de signification mais doit être uniforme.
- Le diagramme doit avoir un titre et les axes doivent être identifiés selon ce qu'ils représentent.
- L'axe des abscisses ne comporte pas de flèche car les modalités ne sont pas des nombres.

Exemple : Une étude a été faite sur les animaux de compagnie préférés parmi les élèves de cinquième TQ.

Animal de compagnie	oiseau	Chat	Chien	Poisson
Points accumulés	10	20	25	30

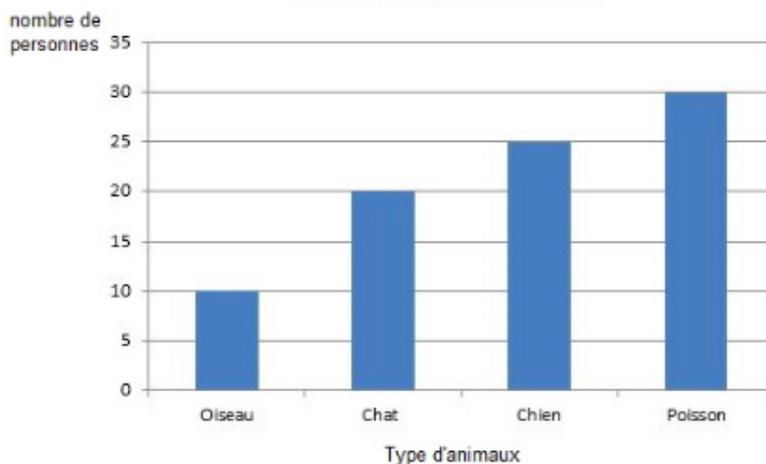


FIGURE 3 – Animal de compagnie préféré

2. Le diagramme circulaire

Le diagramme circulaire permet d'illustrer qu'un tout est partagé en parties. On l'utilise pour représenter des données qualitatives. Il a les caractéristiques suivantes :

- Chaque secteur du cercle est en lien avec une valeur ou une modalité généralement présentée avec sa fréquence (en pour-cent).
- L'angle d'un secteur circulaire représente la proportion d'une catégorie par rapport au tout (360°).
- Il doit y avoir un titre et une légende qui associe le contenu des secteurs à une valeur ou une modalité.

Exemple : On a interrogé 160 élèves d'une école seconde au sujet de leur saison préférée et voici le tableau de distribution des résultats obtenus :

Modalité	Effectif	Fréquence (%)	Angle au centre (en degrés)
Hiver	48	$\frac{48}{160} * 100 = 30$	$\frac{30}{100} * 360 = 108$
Automne	24	15	54
Printemps	16	10	36
Eté	72	45	162
Total	160	100	360

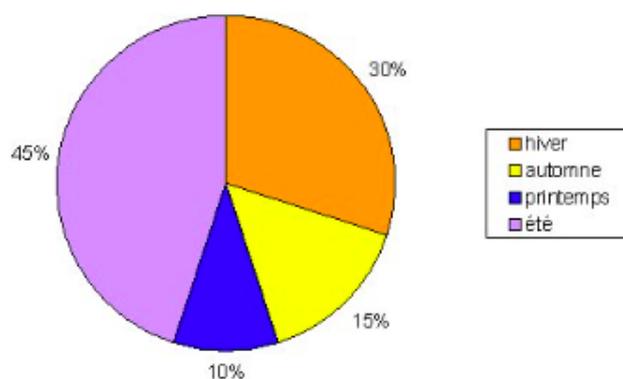


FIGURE 4 – Saison préférée

3 Applications

1. Quelles marques de téléphone ?

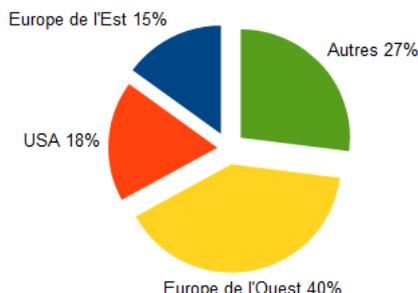
Nous avons examiné la répartition des différentes marques de téléphones dans la classe dans le tableau suivant.

Marque	Nombre d'élèves	Fréquence
Samsung	7	
Huawei	2	
Apple	9	
Xiaomi	1	
Motorola	3	

- Complète le tableau avec le pourcentage relatif à chaque marque.
- On souhaite représenter graphiquement ces résultats. Montre les choix possibles en les réalisant proprement.
- Quels sont les avantages et les inconvénients de chacun de ces graphes ?

2. **Provenance de voitures**

De l'ensemble des voitures particulières immatriculées en Belgique, on a pu créer le diagramme circulaire donnant le pays de construction du véhicule :



- (a) Le caractère étudié est-il qualitatif ou quantitatif? Justifie.
- (b) Si on t'informe qu'en Belgique, 5 700 000 voitures particulières sont immatriculées, détermine le nombre de voitures en fonction de leur provenance.
- (c) Trace un histogramme correspondant au diagramme circulaire.

3. **À propos de chaussures**

Les chaussures portées par les jeunes dépendent de l'activité pratiquée, de la tenue vestimentaire ou encore de la saison, de la tendance du moment, etc. Une grande marque a soumis un petit questionnaire aux passants d'une rue commerciale un samedi.

Voici le nombre de paires de chaussures que possèdent les 80 premières filles et les 80 premiers garçons âgés entre 15 et 18 ans qui ont répondu au questionnaire.

Nombre de paires de chaussures	2	3	4	5	6	8	10	15
Nombre de filles	5	0	11	34	8	5	15	2
Nombre de garçons	6	26	16	22	2	8	0	0

- (a) Quelle est la population de cette enquête?
- (b) Quel est l'échantillon dont les informations sont données ci-dessus?
- (c) Le caractère étudié est-il qualitatif ou quantitatif? Justifie.

4. **Nationalités à l'université.**

Dans une université belge, on a relevé les nationalités d'un certain nombre d'étudiants. Les résultats sont donnés à la table 1 dont il faut compléter les fréquences.

k	Modalités x_k	Effectifs n_k	Fréquences (%) f_k
1	Belge	6051	
2	Française	556	
3	Luxembourgeoise	267	
4	Espagnole	168	
5	Allemande	245	
6	Autres	213	
	Total	7500	

TABLE 1 – Nationalités d'étudiants

5. Quels type de variable et quel traitement statistique ?

Pour les variables suivantes, indique si elle sont qualitative, quantitatives discrètes ou continues. Certaines variables quantitatives discrètes sont traitées statistiquement comme des variables continues. Repère-les.

- a) Le temps qu'il fait
- b) La température extérieure
- c) La taille des élèves de l'école
- d) Le nombre de voitures qui passent à un carrefour sur une journée dans un petit village des Ardennes
- e) Le nombre de pépins contenus dans une mandarine
- f) La vitesse des escargots
- g) Le salaire des travailleurs belges
- h) Le nombre de voitures qui passent devant la pompe à essence de Manage sur une journée