

# BLOC 1

## 6.1 - Traitement des données et probabilités

### 6 - RÉSULTAT D'APPRENTISSAGE GÉNÉRAL

Recueillir et traiter des données statistiques pour faire des prédictions et prendre des décisions éclairées.

#### 6.1 Démontrer une compréhension des concepts de population statistique et de variable statistique.

- *Population et individu statistique appropriés à une étude*
- *Types de variables statistiques et modalités (continue, discrète, ordinale, nominale)*
- *Définition des variables statistiques appropriées pour répondre aux questions d'une étude (compréhension des objectifs d'une étude)*

[http://www.atelier.on.ca/edu/resources/guides/GEE\\_math\\_M\\_3\\_TDP.pdf](http://www.atelier.on.ca/edu/resources/guides/GEE_math_M_3_TDP.pdf)

Dans de nombreux domaines comme les sports, la médecine, les affaires et les spectacles, il est important de savoir interpréter les données qu'on recueille. En effet, on se sert de ces données pour prendre des décisions à tous les jours. Comme exemple :

- Quelle émission devrait-on programmer aux heures d'écoute maximum?
- À quelles intersections faut-il installer des feux de circulation?
- Quels types de lancers doit-on utiliser contre tel(le) joueur(euse) de base-ball?

La **statistique** est la branche des mathématiques qui est consacrée à l'étude des données, la classification et l'analyse de donnée afin de faire une conclusion qui apportera à des prédictions.

La **population** est l'ensemble des personnes, des objets ou des faits sur lesquels porte une étude statistique.

Le **caractère** est sur quoi l'étude porte, de quoi on parle, ce qu'on veut savoir.

L'**échantillon** est un sous-ensemble d'une population.

La **taille** d'une population ou d'un échantillon est le nombre d'éléments qu'il comporte.

Exemple : Une étude a été faite auprès de 1000 élèves de 12<sup>e</sup> année du Nouveau-Brunswick pour savoir s'ils savent ce qu'ils vont faire une fois fini l'école.

#### a) *Quelle est la population de cette étude statistique ?*

Pour répondre à cette question, il faut se demander : « Qui est-ce que l'étude vise? » ou « Nous allons recueillir nos réponses auprès de qui? »

La population de cette étude statistique est donc l'ensemble de tous les élèves de 12<sup>e</sup> année du Nouveau-Brunswick.

b) *Quel est le caractère de cette étude statistique?*

Pour répondre à cette question, il faut se demander : « Sur quoi porte l'étude? » ou « Que veut-on savoir dans cette étude statistique? »

Le caractère de cette étude statistique est de savoir si les élèves de 12<sup>e</sup> année du Nouveau-Brunswick savent ce qu'ils vont faire une fois fini l'école.

c) *Quelle est l'échantillon de cette étude statistique ?*

Il serait trop long de vérifier auprès de tous les élèves de 12<sup>e</sup> de la province alors on se limite à 1000 élèves de 12<sup>e</sup> année de la province. Alors, l'échantillon de cette étude statistique regroupe 1000 élèves de 12<sup>e</sup> année du Nouveau-Brunswick.

d) *Quelle est la taille de cet échantillon ?*

Donc, la taille de cet échantillon est de 1000.

### Exercice

1. *Quelle est la population, le caractère, l'échantillon et la taille de l'échantillon de cette étude statistique ?*

a) Lundi dernier, un contrôleur de la qualité dans une usine d'ampoules a vérifié une ampoule sur 10 pour savoir si elles éclairaient comme il faut. Cette journée-là, l'usine a produit 600 ampoules.

Population : Ampoules dans une usine.

Caractère : Savoir si elles éclairent comme il faut.

Échantillon : 1 ampoule sur 10 produites cette journée là.

$$\frac{1}{10} = \frac{x}{600}$$

Taille :  $10x = 600$

$$x = 60 \text{ ampoules}$$

b) Certains habitants d'une ville aimeraient bien avoir une piscine municipale. Les conseillers municipaux demandent à une firme d'effectuer un sondage auprès de 1200 personnes de la population pour déterminer si le pourcentage de la population désirant une piscine est assez élevé pour en justifier la construction.

Population : Habitants d'une ville.

Caractère : Savoir si elles veulent avoir une piscine municipale.

Échantillon : 1200 personnes de la population.

Taille : 1200 personnes

2. Dans chacun des cas suivants, détermine la population et le caractère de l'étude statistique.

a) Une commission de transport en commun effectue un sondage auprès de ses usagers pour savoir s'ils sont satisfaits de l'entretien des autobus.

Population : les usagers du transport en commun.

Caractère : savoir s'ils sont satisfaits de l'entretien des autobus.

b) Martin vérifie dans le journal la liste des meilleurs compteurs parmi les joueurs de la Ligue nationale de hockey.

Population : Compteurs de la ligue nationale de hockey.

Caractère : savoir qui sont les meilleurs compteurs.

c) Le capitaine d'une équipe de soccer demande aux joueurs si l'heure de la pratique est convenable.

Population : l'équipe de soccer.

Caractère : savoir si l'heure de la pratique est convenable.

3. A la fin des Jeux Olympiques, la ville hôte désire savoir ce que les athlètes ont le plus apprécié de leur séjour. Parmi les quatre échantillons suivants, lequel est-il préférable de choisir ?

a) Demander à 300 athlètes choisis au hasard parmi ceux qui sont à la cafétéria le dernier midi.

b) Demander à 200 athlètes choisis au hasard parmi tous les athlètes participant aux Jeux.

c) Demander à 275 athlètes choisis au hasard à la sortie de la cérémonie de fermeture.

d) Demander à 250 athlètes choisis au hasard parmi tous les athlètes participant aux Jeux.

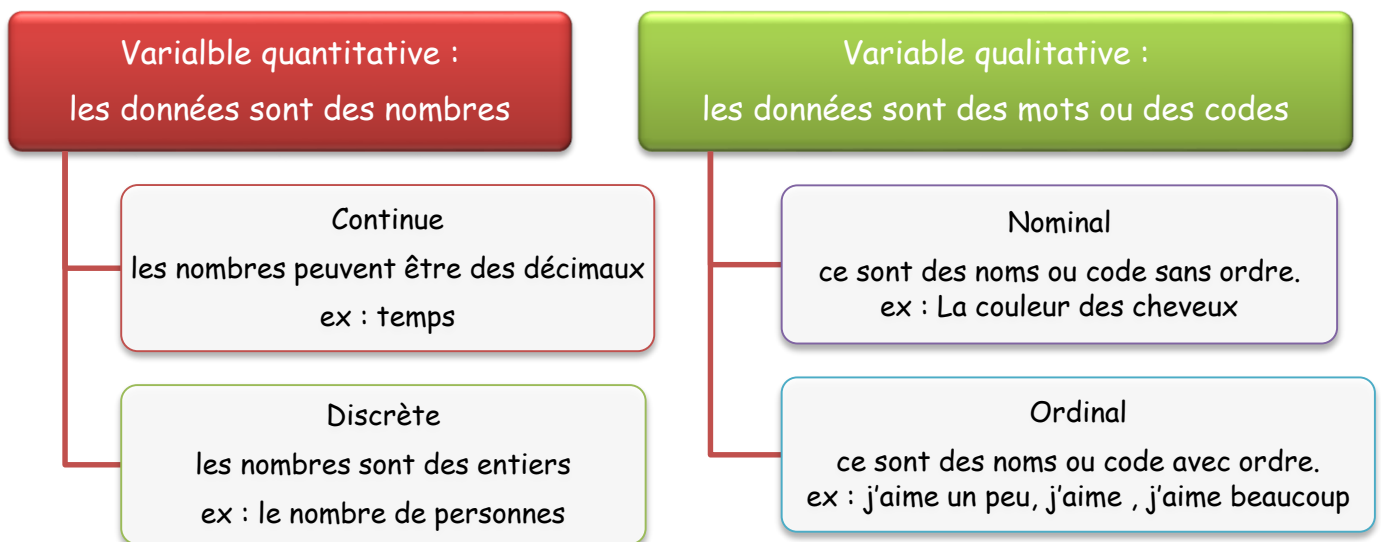
Le caractère est ce sur quoi porte la recherche de données.

Exemple : On désire connaître la saveur préférée de gomme à mâcher chez les adolescents.  
Alors, le caractère de l'étude est la saveur de gomme à mâcher.

Les modalités sont les différentes valeurs que la variable étudiée peut prendre.

Exemple : Dans l'exemple des saveurs de gomme à mâcher, les modalités pourraient être : cerise, menthe, raisin, cannelle...

### Types de modalités



### Exercice

1. Entre 7h et 8h, on a relevé le nombre de passagers dans chacun des autobus de la ville passant à l'intersection des rues Rachel et Florence. Quelle est la population sous études?
  - a) L'ensemble des autobus de la ville.
  - b) L'ensemble des passagers qui prennent un autobus qui passe à l'intersection des rues Rachel et Florence entre 7 h et 8 h.
  - c) Le nombre de passagers qui prennent un autobus qui passe à l'intersection des rues Rachel et Florence entre 7h et 8h.
  - d) Les autobus de la ville qui passent à l'intersection des rues Rachel et Florence entre 7h et 8h.

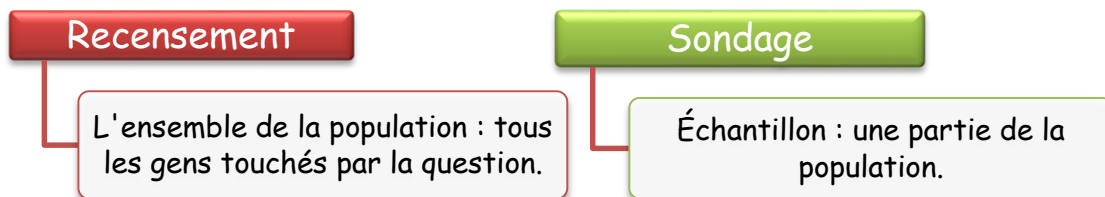
2. Indique le type de la variable de chaque étude ci-dessous.

Variable		Quantitative		Qualitative	
		Continue	Discrète	Nominale	Ordinale
a)	La vitesse du vent (km/h)	✓			
b)	Le prix d'une maison en dollars (\$)	✓			
c)	Le nombre d'ordinateurs vendus en une journée dans un commerce d'électronique		✓		
d)	La saison préférée d'une personne			✓	
e)	La quantité de précipitation quotidienne (mm)		✓		
f)	La lettre décrivant le résultat obtenu dans un cours de physique à l'université				✓

## 6.2 Choisir une méthode de collecte de données appropriée à une étude

- Méthodes d'enquête
  - recensement – population
  - sondage – échantillon
- Méthodes d'échantillonnage
  - probabiliste – aléatoire simple, en grappes, stratifié, systématique
  - non probabiliste – par quotas, auto sélectionné, de volontaires, par commodité

### Méthodes d'enquêtes



## Exercice

Pour chacune de ces études, détermine s'il est préférable de faire un recensement ou un sondage. Justifie ton choix.

1. Une compagnie veut connaître la durée de vie des ampoules qu'elle fabrique.

*Sondage*

2. Une école veut connaître les choix de cours des élèves de 11<sup>e</sup> année pour l'année suivante. *Recensement*

3. Une maison d'édition désire connaître l'opinion des jeunes Canadiens Français de 8 à 12 ans sur une collection de livres éducatifs qu'elle songe à publier. *Sondage*

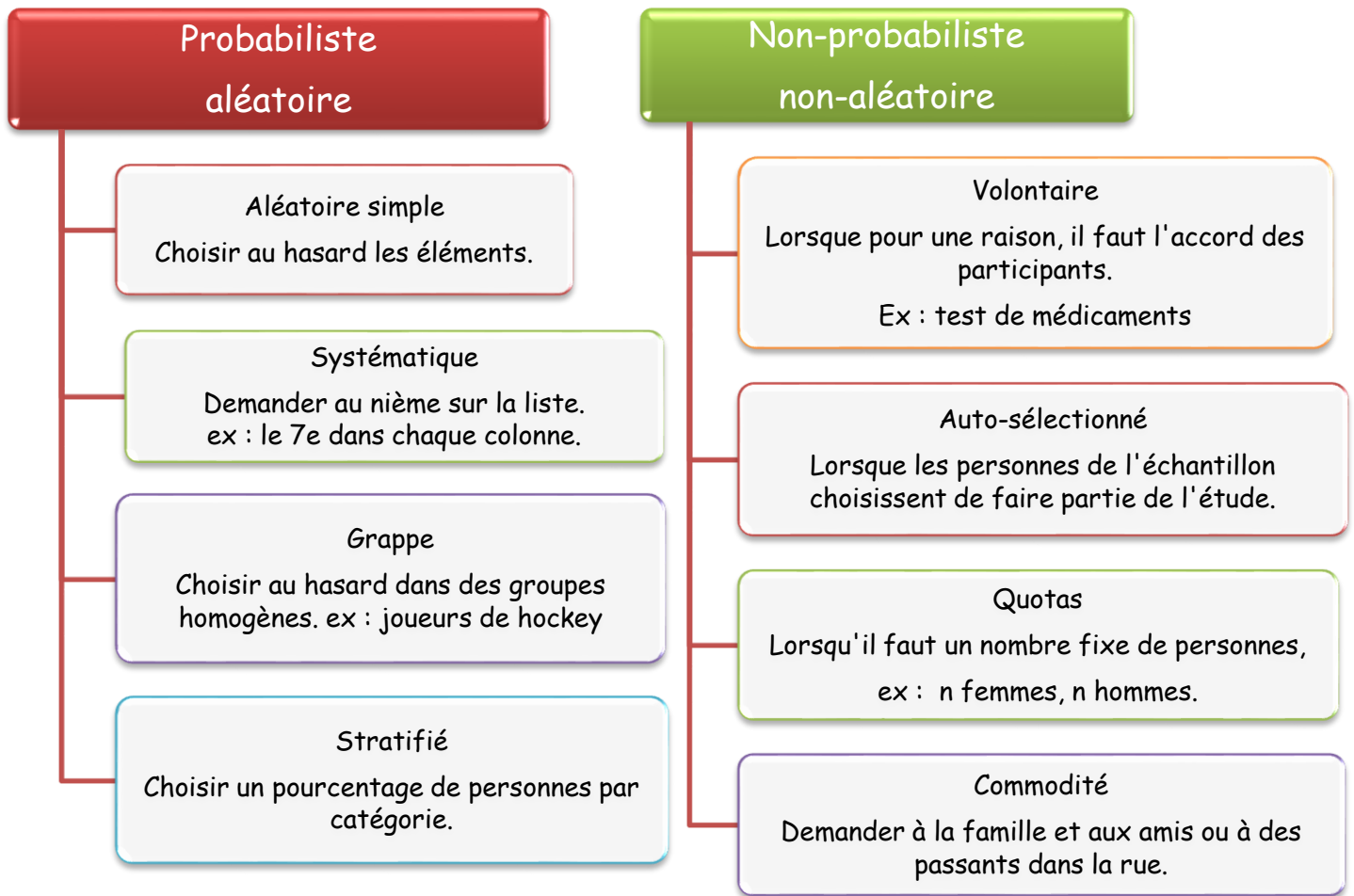
4. Un importateur de vin veut vérifier la qualité de ses réserves. *Sondage*

5. Dans l'émission de télévision Star Académie, le public détermine qui sera sauvé parmi les interprètes mis en danger. *Recensement*

6. L'organisation d'un festival d'été désire connaître l'origine des participants au festival. *Sondage*

7. Une université désire connaître le pourcentage de ses étudiants qui reçoit une bourse d'étude. *Recensement*

## Méthodes d'échantillonnage : Façons de choisir l'échantillon.



### Exemple : méthode de l'échantillon aléatoire

Lors d'une importante rencontre en athlétisme, les participants à l'épreuve du 100 mètres avaient les âges suivants : 19, 24, 25, 19, 23, 22, 25, 31, 27, 24, 28, 22, 20, 26, 31, 23, 27, 19, 20, 26

Les responsables désirent constituer un échantillon de 5 personnes pour passer des tests « anti-dopage ». Comme on compare les coureurs à l'aide de leur âge et que ces âges sont tous à peu près semblables, on procède de la façon suivante :

- On numérote les coureurs de 1 à 20.
- On place des nombres variant de 1 jusqu'à 20 dans un sac.
- On tire au hasard 5 numéros.

L'échantillon est constitué des 5 coureurs associés à ces numéros. (on aurait aussi pu placer les noms des coureurs dans un sac et tirer 5 noms au hasard).

Exemple : échantillon systématique

Sébastien, qui travaille pour une firme de sondage, doit effectuer un sondage par téléphone. Pour constituer son échantillon, il procède de la façon suivante :

- il prend l'annuaire téléphonique et il l'ouvre à la première page
- il choisit au hasard une colonne et un rang dans la colonne
- il téléphone à la personne qui occupe cette position ainsi qu'à tous les abonnés qui occupent cette position dans chacune des autres pages.

Exemple : échantillon par grappes

Une commission scolaire désire connaître les habiletés en français des élèves de 12<sup>e</sup> année. Comme les élèves sont déjà en groupes-classes dans les écoles, ils choisissent au hasard six groupes parmi les dix-neuf appartenant à ce district scolaire, et tous les élèves des six groupes choisis devront passer un examen de français.

Exemple : échantillon stratifié

La direction d'une école désire savoir qui est l'enseignant le plus populaire auprès des élèves. Dans cette école de 2000 élèves, ceux-ci sont répartis comme suit :

22% sont en 9<sup>e</sup>

18% sont en 10<sup>e</sup>

27% sont en 11<sup>e</sup>

33% sont en 12<sup>e</sup>

L'école décide de choisir un échantillon de 400 élèves dans lequel chaque niveau sera représenté dans le même rapport que dans l'école. Alors, il faudra choisir au hasard  $22\% \times 2000 = 44$  élèves de 9<sup>e</sup>,  $18\% \times 2000 = 36$  élèves de 10<sup>e</sup>,  $27\% \times 2000 = 54$  élèves de 11<sup>e</sup> et  $33\% \times 2000 = 66$  élèves de 12<sup>e</sup>.



Exercice : Nomme la méthode d'échantillonnage.

1. Justin demande à sa famille leur film préféré afin de savoir les préférences des canadiens. Non probabiliste - commodité
2. Kent donne à chaque client un site web où tu peux aller remplir un sondage à propos de la qualité du service. Non probabiliste - auto sélectionné
3. La cafétéria a demandé à au moins 50 personnes ce qu'il pensait du menu. Probabiliste - aléatoire
4. La psychologue de l'école cherche des élèves qui seraient intéressé à remplir un sondage sur la dépression. Non-probabiliste - auto sélectionné ou volontaire
5. Tilenoul cherche des gens près à essayé leur nouvelle aspirine. Non - probabiliste - volontaire
6. Google demande à sa liste d'abonnés pour ceux intéressé à recevoir des sondages à propos de leurs produits. Non- probabiliste - auto sélectionné
7. MurMart a demandé à 100 personnes par magasins s'ils préféreraient les caisses automatiques aux caisses normales. Probabiliste - stratifié
8. Mme Manon cherche des élèves intéressés à participer à une expérience au laboratoire. Non-probabiliste - auto sélectionné
9. Mme Elaine demande à ses élèves présents à la période d'étude ce qu'ils pensent de son enseignement. Non-probabiliste - commodité
10. Marc demande à 5 personnes par classe du secondaire s'ils aiment la planche à roulette longue. Probabiliste - par grappe

## Études statistique

Pour chaque étude, indique le caractère étudié, le type de caractère, la méthode d'enquête et la méthode d'échantillonnage à utiliser.

1. On fait une étude sur le poids des bébés à la naissance. On recueille les données de trois des hôpitaux de la province.

Caractère : poids des bébés à la naissance

Type : quantitatif continu

Méthode d'enquête : sondage

Méthode d'échantillonnage : probabiliste par grappe

2. On interroge un certain nombre de personne au hasard au sujet de leur intention de vote aux prochaines élections.

Caractère : Intention de vote aux prochaines élections

Type : qualitatif nominal

Méthode d'enquête : sondage

Méthode d'échantillonnage : probabiliste – aléatoire simple

3. On demande à tous les résidents d'un quartier combien de personnes compte leur famille.

Caractère : Combien de personnes compte leur famille

Type : quantitatif discrète

Méthode d'enquête : Recensement

Méthode d'échantillonnage : --

4. Une nutritionniste questionne deux classes par grade à l'école MFR afin de savoir la qualité de la nourriture servie à la cafétéria.

Caractère : la qualité de la nourriture à la cafétéria

Type : qualitatif – ordinale

Méthode d'enquête : sondage

Méthode d'échantillonnage : probabiliste – statifié

5. La direction questionne chaque élève afin de savoir quels types d'allergie ils ont.
- Caractère : quels types d'allergie*  
*Type : qualitatif nominal*  
*Méthode d'enquête : recensement*  
*Méthode d'échantillonnage : --*
6. Afin de connaître le nombre moyen de graines d'un tournesol géant, on compte le nombre de graines à chaque vingtième fleur dans en champ.
- Caractère : nombre de graines d'un tournesol géant*  
*Type : quantitatif discret*  
*Méthode d'enquête : sondage*  
*Méthode d'échantillonnage : probabiliste - systématique*
7. On fait une étude sur quelques poissons pour vérifier leur adaptation aux eaux polluées (peu adapté, plus ou moins adapté, très adapté).
- Caractère : l'adaptation des poissons aux eaux polluées*  
*Type : qualitatif ordinal*  
*Méthode d'enquête : sondage*  
*Méthode d'échantillonnage : probabiliste aléatoire*
8. Afin de connaître la quantité de tasse de café bue par personne par jour au Costa Rica, on questionne 10% des gens dans chaque ville.
- Caractère : la quantité de tasse de café bue par personne par jour au Costa Rica*  
*Type : quantitatif discret*  
*Méthode d'enquête : sondage*  
*Méthode d'échantillonnage : probabiliste stratifié*
9. On cherchant le pourcentage des élèves qui préfère la couleur bleu à l'école, on questionne toutes les classes de français à madame Mélanie.
- Caractère : préfère la couleur bleu à l'école*  
*Type : quantitatif continu*  
*Méthode d'enquête : sondage*  
*Méthode d'échantillonnage : non-probabiliste - commodité*

10. Afin de faire une liste d'appel, le directeur demande le numéro de téléphone à tous les employés de l'école.

*Caractère : le numéro de téléphone des employés de l'école*

*Type : qualitatif - nominal*

*Méthode d'enquête : recensement*

*Méthode d'échantillonnage : --*

11. Afin de connaître la taille moyenne des joueurs des équipes de hockey de la NHL, on questionne le 20<sup>e</sup> joueur de la liste, ensuite le 120<sup>e</sup>, le 220<sup>e</sup>, et ainsi de suite.

*Caractère : la taille moyenne des joueurs des équipes de hockey de la NHL*

*Type : quantitatif continu*

*Méthode d'enquête : sondage*

*Méthode d'échantillonnage : probabiliste - systématique*

12. On note la malléabilité de tous les types des roches qu'on extrait d'une carrière comme étant très malléable, malléable, peu malléable ou non malléable.

*Caractère : la malléabilité de tous les types de roches extrait d'une carrière*

*Type : qualitatif ordinal*

*Méthode d'enquête : recensement*

*Méthode d'échantillonnage : --*

13. Pour chaque 100 calculatrices fabriquées dans une usine, on en vérifie 10 afin de savoir s'ils sont défectueux ou non.

*Caractère : calculatrices défectueuses*

*Type : qualitatif ordinal*

*Méthode d'enquête : sondage*

*Méthode d'échantillonnage : probabiliste aléatoire*

14. On mesure la taille de chaque 10<sup>e</sup> arbre de Noël vendu.

*Caractère : la taille des arbres de Noël vendus*

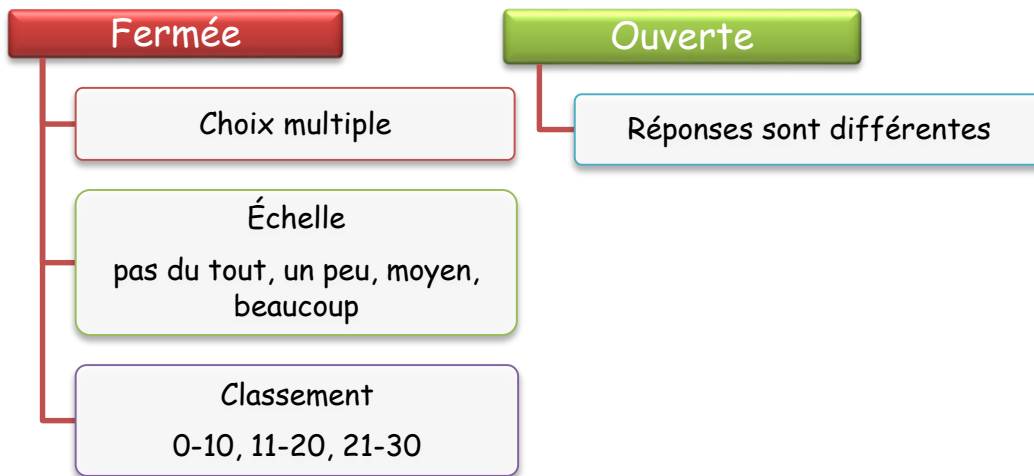
*Type : quantitatif continu*

*Méthode d'enquête : sondage*

*Méthode d'échantillonnage : probabiliste - systématique*

### 6.3 Choisir le type de questions approprié selon les données à cueillir.

- Fermée (choix multiple, échelle, classement)
- Ouverte



Exemple :

Question ouverte

*Quelle est la plus importante question à laquelle font face les jeunes aujourd'hui?*

Question fermée

*Parmi les questions auxquelles font face les jeunes aujourd'hui, laquelle est la plus importante?*

- Chômage
- Environnement
- Violence chez les jeunes
- Hausse des frais de scolarité
- Drogue dans les écoles
- Orientation professionnelle

Exercice : De quel type de question s'agit-il? (ouverte ou fermée)

### Boissons énergisantes

Une boisson énergisante, est destinée à donner un regain d'énergie à son consommateur, en utilisant un mélange de différents d'ingrédients stimulants tels que de la caféine, la taurine, la créatine, du sucre etc.

⇒ **Rappel** : Ce sondage concerne uniquement les boissons énergisantes telles que définies ci-dessus.

Comme le Red Bull , Monster , RockStar  etc.

⇒ Dans ce sondage un consommateur se définit comme étant quelqu'un qui en consomme sur une base régulière ou qui à déjà bu plus d'une canette au cours des 2 dernières années.

#### Questionnaire

1. Quel est votre sexe ? **Fermée, choix multiples**

Féminin       Masculin

2. Dans quel niveau scolaire êtes-vous? **Fermée - classement**

9<sup>e</sup>       10<sup>e</sup>       11<sup>e</sup>       12<sup>e</sup>

3. Êtes-vous un consommateur de boisson énergisante ? **Fermée, choix multiples**

oui       non

4. Quelle marque consommez-vous le plus ? \_\_\_\_\_

5. Quel est votre de degrés de satisfaction envers les boissons énergisantes ? **Fermée échelle**

très satisfait  
 satisfait  
 moyennement satisfait  
 peu satisfait  
 pas du tout satisfait

5. En moyenne combien d'argent dépensez-vous par semaine pour l'achat de ce produit ?  
**ouverte**

\_\_\_\_\_

6. Combien de canettes consommez-vous par semaine ? **Fermée - classement**

0 à 3       4 à 6       7 et plus

7. Êtes-vous dépendant de ce produit ? **Fermée - échelle**

(Très dépendant signifie que vous en buvez au moins une par jour)

très dépendant    dépendant    un peu dépendant    pas du tout dépendant

#### 6.4 Créer un outil de collection de données.

- [Sources de biais](#) (échantillonnage, questionnaire, représentation)

On désigne par biais toute erreur produite lors de la recherche d'informations, alors les résultats de l'étude statistique ne représentent pas la réalité.



#### Exercice

1. Explique pourquoi les maisons de sondage évitent de téléphoner à l'heure du souper.  
*Les gens sont occupés et vont répondre n'importe quoi pour avoir fini.*
2. Explique pourquoi les gouvernements n'organisent pas d'élections durant l'été ni en plein hiver ? Quel serait le meilleur mois pour tenir des élections ?  
*Trop de gens sont en vacances l'été, et l'hiver, il risquerait d'avoir une tempête.  
N'importe quand au printemps serait meilleur.*
3. Explique pourquoi les questions suivantes sont mal formulées. Corrige leur formulation afin que celles-ci n'influencent pas l'opinion des répondants.
  - a) Certains experts affirment que le produit A est plus efficace que le produit B. Êtes-vous d'accord avec l'opinion des experts ?  
*De cette façon, ça fait pencher les gens vers le produit A.*
  - b) Quel âge avez-vous ?  
*Il faudrait donner un choix de réponses car on pourrait avoir trop de réponses différentes.*