

# Chapitre 1 GÉNÉRALITÉS

## 1.1 La définition de l'enquête

En termes très généraux, une enquête est une méthode de recueil d'information sur un grand nombre de personnes, en interrogeant seulement quelques-unes d'entre elles.

## 1.2 L'intérêt de l'enquête

Lorsqu'on a besoin des informations sur le terrain, on fait un recensement. Cependant, les recensements prennent beaucoup de temps et coûtent cher. Avant de planifier un recensement, il faut s'assurer **i)** qu'il est le meilleur et le seul moyen d'obtenir les informations dont vous avez besoin, et **ii)** que vous avez le temps, l'argent, le personnel et les ressources voulues pour effectuer correctement cette étude dans les délais dont vous disposez.

Au cas où l'on ne peut pas faire un recensement, on peut procéder à une enquête. Les conclusions des enquêtes peuvent être utilisées par les planificateurs et les décideurs politiques des gouvernements, dans les organisations privées (lucratives ou à but non lucratif), ainsi que par les sociologues et les autres spécialistes des sciences sociales.

Une enquête comprend plusieurs étapes liées entre elles, notamment, la définition des objectifs, la sélection d'une base de sondage, le choix du plan d'échantillonnage, la conception du questionnaire, la collecte et le traitement des données, l'analyse et la diffusion des données, et la documentation de l'enquête.

La durée d'une enquête peut être répartie en plusieurs phases.

La première est la **planification** : la conduite d'une enquête est une opération complexe qui comporte de nombreuses étapes et procédure différentes. Il est essentiel pour le succès de l'enquête de procéder à une planification précoce et minutieuse du travail qu'elle implique.

Planifier une enquête n'est évidemment pas le travail d'une seule personne; il faut constituer une équipe de planification de l'enquête. Cette équipe réfléchit à la façon dont on va mener l'enquête. Elle doit avoir une idée claire de la raison pour laquelle on l'effectue.

Elle doit également savoir comment influencer sur les décisions concernant le projet et savoir de quelles informations spécifiques dont on a besoin pour le faire.

En Dehors de la phase de la planification, il existe aussi celle de la **conception** et de l'**élaboration** puis, celle de la **mise en œuvre**.

L'objectif de ce cours est de donner un aperçu des activités comprises dans le déroulement d'une enquête statistique.

## **Chapitre 2** DÉMARCHE DE L'ENQUÊTE

À première vue les gens pensent que le déroulement d'une enquête consiste simplement à poser des questions et à compiler les réponses pour obtenir des statistiques. Il faut cependant faire une enquête étape par étape, appliquer des procédures et des formules précises pour que les résultats donnent de l'information exacte et significative. Il faut bien connaître les tâches particulières, leurs liens et leur pertinence pour comprendre le processus complet.

Voici les étapes d'une enquête :

- ➔ définition du problème et formulation des objectifs ;
- ➔ sélection d'une base de sondage ;
- ➔ choix d'un plan d'échantillonnage ;
- ➔ conception du questionnaire ;
- ➔ collecte des données ;
- ➔ saisie et codage des données ;
- ➔ vérification et imputation (traitement des données) ;
- ➔ estimation (pondération) ;
- ➔ analyse des données ;
- ➔ diffusion des données ;
- ➔ documentation (concepts et définition, sur la qualité des données, etc.).

Voici maintenant une brève description de chaque étape.

### **2.1 La définition du problème et la formulation des objectifs**

La première tâche de la planification d'une enquête est d'éclaircir pourquoi cette étude, quel est le problème et quels sont les objectifs le mieux et le plus clairement possible. Un énoncé clair des objectifs oriente toutes les étapes ultérieures de l'enquête. Ces étapes devraient être planifiées de façon à garantir que les résultats en bout de ligne correspondent aux objectifs originaux.

### **2.2 L'échantillonnage**

Pour recueillir des informations sur une population statistique, on dispose de deux méthodes :

- ➔ La méthode exhaustive ou recensement où chaque individu de la population est étudié selon le ou les caractères étudiés;
- ➔ La méthode des sondages ou échantillonnage qui conduit à n'examiner qu'une fraction de la population, un échantillon.

L'échantillonnage est la technique par laquelle on choisit une partie **représentative** de la population pour déterminer les caractéristiques de l'ensemble de la population étudiée.

Ce n'est que si un échantillon est statistiquement représentatif que les données des ménages de l'échantillon reflètent la situation véritable de l'ensemble des ménages de la zone étudiée.

Dans la plupart des études, il n'est pas possible de recueillir des données auprès de chaque ménage ou personne relevant du domaine d'étude. Le recensement complet d'un groupe donné est une tâche considérable qui dépasse généralement les possibilités de la plupart des **budgets** et du **temps** de l'enquête.

C'est pourquoi, dans la plupart des cas, on sélectionne une partie seulement du groupe à étudier, pour donner des informations fiables sur l'ensemble du groupe. Les procédures scientifiques d'échantillonnage permettent de procéder à des généralisations sur un grand nombre de personnes à partir d'informations recueillies auprès d'un petit nombre seulement.

Avant d'échantillonner, il faut :

- ➔ Définir la population d'intérêt ou population cible et l'unité d'enquête ;
- ➔ Fixer la taille de l'échantillon ;
- ➔ Sélectionner une base de sondage ;
- ➔ Fixer une méthode d'échantillonnage.

### 2.2.1 définir la population d'intérêt et l'unité d'enquête.

La population d'intérêt, c'est l'ensemble des individus (unités statistiques) que l'on veut étudier.

L'unité d'enquête n'est pas forcément les personnes. Elle peut être les hôpitaux, les entreprises, les pays, etc.

Attention : La population observée est différente de la population cible<sup>1</sup> et les conclusions ne s'appliqueront qu'à la population réellement observée. L'utilisateur des résultats doit en être informé surtout dans le cas d'une étude de panel (**vrai panel et pseudo panel**).

### 2.2.2 Fixer la taille de l'échantillon

Voici l'une des questions les plus souvent posées à un statisticien : Combien de personnes doit-on questionner dans notre échantillon? Les gestionnaires sont anxieux d'obtenir une réponse à cette question fondamentale pendant la phase de la planification de l'enquête parce qu'elle a des répercussions directes sur les considérations opérationnelles, notamment sur le budget et le temps de réalisation de l'enquête.

---

<sup>1</sup> La population cible est la population que nous **voulons observer**, tandis que la population observée est la population **qu'on a observé**.

---

Il n'y a pas de solution magique ou de recette parfaite pour déterminer la taille de l'échantillon. Il s'agit plutôt d'un processus de compromis au cours duquel les besoins de précision des estimations sont pondérés en tenant compte de diverses contraintes opérationnelles, par exemple, le temps et les ressources disponibles (financière, humaines, etc.).

### 2.2.3 Sélectionner une base de sondage

La base de sondage est l'outil qu'on utilise pour avoir accès à la population. C'est une liste d'unités (individus ou groupe d'individus) qui couvre toute la population avec une identification de chaque unité. Cette liste doit être **exhaustive** (sinon défaut de couverture) et **sans doublon**. Une base de sondage doit **couvrir** correctement la population cible, être **mis à jour** à chaque instant, être de bonne **qualité**, être aussi **stable** que possible dans le temps, **accessible** et **facile** à utiliser.

Il existe trois types de base de sondage :

- Les listes : c'est une **liste physique**, notamment, un fichier de données, un imprimé d'ordinateur ou un annuaire téléphonique ; ou une **liste conceptuelle**, par exemple une liste de tous les véhicules qui entrent au stationnement d'un centre commercial entre 9 h et 20 h pendant une journée en particulier,

- Les bases aréolaires : c'est une liste géographique dont les unités correspondent à des secteurs géographiques et dont les unités composantes sont des ménages, des fermes, des entreprises, etc.

- Les bases de sondage multiples : c'est une combinaison de deux ou plusieurs bases (des listes et des bases aréolaires ou deux listes ou plus). Les bases de sondage multiples sont habituellement utilisées lorsqu'aucune base unique ne peut fournir la couverture nécessaire de la population cible. Lors d'une enquête, on peut utiliser la base des entreprises téléphoniques et celle des mairies pour avoir les informations - presque - complètes d'une population.

Un organisme statistique peut habituellement utiliser, approfondir ou créer une base de sondage. La base choisie détermine la définition de la population de l'enquête et peut avoir des répercussions sur les méthodes de collecte des données, de sélection et d'estimation de l'échantillon, ainsi que sur le coût de l'enquête et la qualité des résultats. Par exemple, si on utilise une liste de numéros de téléphone pour sélectionner un échantillon de ménages, tous les ménages n'ayant pas le téléphone seront alors exclus de la population observée.

### 2.2.4 Fixer une méthode d'échantillonnage

Il existe deux types de méthodes d'échantillonnage : L'échantillonnage probabiliste ou aléatoire et l'échantillonnage non probabiliste ou empirique. Les échantillons aléatoires sont constitués par tirage au sort dans la population mère pour laquelle on dispose d'une base de sondage. Alors

que pour l'échantillonnage non probabiliste, on n'a pas besoin de la base de sondage. Une autre différence qui existe entre les deux tient au fait que dans le cas de l'échantillonnage probabiliste chaque unité a une « chance » d'être sélectionnée et que cette chance peut être quantifiée, ce qui n'est pas vrai pour l'échantillonnage non probabiliste; dans ce cas, chaque unité incluse à l'intérieur d'une population n'a pas une chance égale d'être sélectionnée.

#### a) L'échantillonnage probabiliste

Il est plus complexe, prend plus de temps et est habituellement plus coûteux que l'échantillonnage non probabiliste. On distingue 4 méthodes :

- L'échantillonnage aléatoire simple ;
- L'échantillonnage systématique ;
- L'échantillonnage stratifié ;
- L'échantillonnage par grappe.

##### ➔ Echantillonnage aléatoire simple

L'échantillonnage aléatoire simple est une méthode qui consiste à prélever au hasard et de façon indépendante,  $n$  individus ou unités d'échantillonnage d'une population à  $N$  individus.

Dans ce cas, chaque membre d'une population a une chance égale d'être inclus à l'intérieur de l'échantillon. Cette méthode est appliquée quand la population possède des caractéristiques homogènes et que nous détenons une base de sondage.

Le choix du tirage peut se faire avec remise (un individu peut être choisi plusieurs fois) ou sans remise (un individu déjà choisi ne peut l'être de nouveau).

Si l'on note  $n$  la taille de l'échantillon et  $N$  la taille de la population, on peut tirer :

$N^n$  échantillon avec remise ou  $C_N^n$  échantillon quand on fait un tirage sans remise.

Avantage de cette méthode : On peut espérer un échantillon «représentatif» puisque la méthode donne à chaque individu de la population une chance égale.

Inconvénients: la méthode n'est applicable que lorsqu'il existe une liste exhaustive de toute la population.

##### ➔ Echantillonnage aléatoire systématique

L'échantillonnage systématique est une méthode qui exige aussi l'existence d'une liste de la base de sondage où chaque individu est numéroté de 1 jusqu'à  $N$ .

Notons  $n$ , le nombre d'individus que doit comporter l'échantillon (la taille de l'échantillon).

L'entier voisin de  $\frac{N}{n}$  sera noté «  $r$  » et appelé « raison de sondage » ou « pas de sondage » qui correspond aux nombres d'échantillons possibles que l'on peut tirer. On choisit au hasard un entier naturel  $d$  entre 1 et  $r$  (cet entier sera le point de départ). L'individu dont le numéro

correspond à  $d$  est le premier individu. Pour sélectionner les autres, il suffit d'ajouter à  $d$  la raison de sondage.

### Exemples

On a une population de 400 individus, on veut un échantillon de 100 individus.

De ce fait,  $r = 400/100=4$

On n'a donc que 4 échantillons possibles.

1, 5, 9, 13, 17, ..., 397

2, 6, 10, 14, 18, ..., 398

3, 7, 11, 15, 19, ..., 399

4, 8, 12, 16, 20, ..., 400

Avantages : facile à sélectionner parce qu'un seul individu est choisi au hasard.

On peut obtenir une bonne précision parce que la méthode permet de répartir l'échantillon dans l'ensemble de la liste.

Inconvénients : Les données peuvent être biaisées à cause de la périodicité s'il n'y a pas homogénéité des caractéristiques de la population.

#### ↪ Echantillonnage stratifié

On divise la population en groupes homogènes (appelés strates), qui sont mutuellement exclusifs (selon l'âge, le sexe, la province de résidence, le revenu, etc.).

**Les variables de stratification** doivent être simples à utiliser, facile à observer et étroitement liées au thème de l'enquête<sup>2</sup>.

Proportionnellement à son importance dans la population, on calcule combien il faut d'individus au sein de l'échantillon pour représenter chaque strate et dans chacune des strates, on sélectionne des échantillons indépendants. On peut utiliser n'importe quelle des méthodes d'échantillonnage. La méthode d'échantillonnage peut varier d'une strate à une autre.

### Exemples :

Choisir par échantillonnage stratifié, 10 élèves dans un groupe de 60, en tenant compte du fait que 30 d'entre eux sont en CP2, 18 en CE1 et 12 en CE2.

Solution : On aura dans les 10 élèves, 5 du CP2, 3 de CE1 et 2 de CE2.

#### ↪ Echantillonnage par grappe et à plusieurs degrés

On sélectionne **au hasard** un certain nombre de grappes (des groupes qui ne sont pas forcément homogènes à l'intérieur) pour représenter la population.

- Si on enquête tous les individus des grappes choisies, on parle de sondage par grappe

<sup>2</sup> Lutter contre le paludisme (âge), analyse de la pauvreté (quartier), analyse du chômage (sexe).

- Si, dans chaque grappe, on tire encore un autre échantillon, on parle de sondage à deux degrés. On peut avoir un sondage à plusieurs degrés (degré supérieur à 2).

Avantages : La méthode ne nécessite pas forcément une base de sondage puisque seuls les individus inclus dans les grappes comptent. Elle permet de limiter l'échantillon à des groupes compacts ce qui permet de réduire les coûts de déplacement, de suivi et de supervision.

Inconvénients : La méthode peut entraîner des résultats imprécis (moins précis que les méthodes précédentes) puisque les unités voisines ont tendance à se ressembler. Elle ne permet pas de contrôler la taille finale de l'échantillon.

### **b) L'échantillonnage non probabiliste**

On oppose aux méthodes aléatoires les méthodes non aléatoires ou empiriques.

Dans le cas de l'échantillonnage probabiliste, chaque unité a une chance d'être sélectionnée. Dans celui de l'échantillonnage non probabiliste, on suppose que la distribution des caractéristiques à l'intérieur de la population est égale. C'est ce qui fait que le chercheur croit que n'importe quel échantillon serait représentatif et que les résultats, par conséquent, seront exacts.

On ne peut mesurer la fiabilité d'un échantillonnage non probabiliste; la seule façon de mesurer la qualité des données en résultant consiste à comparer certains des résultats de l'enquête à l'information dont on dispose au sujet de la population. Encore une fois, rien ne fournit l'assurance que les estimations ne dépasseront pas un niveau acceptable d'erreur. Les statisticiens hésitent à utiliser les méthodes d'échantillonnage non probabiliste parce qu'il n'existe aucun moyen de mesurer la précision des échantillons en découlant.

Cette méthode est utilisée :

- pour des études exploratoires;
- pour réduire les coûts;
- pour l'analyse des petits échantillons
- quand il est impossible ou non envisageable d'utiliser la méthode aléatoire.

La méthode d'échantillonnage non-probabiliste est utilisée lorsqu'il n'est pas possible de constituer une liste exhaustive (base de sondage) de toutes les unités du sondage.

Il existe plusieurs méthodes d'échantillonnage non probabiliste, les plus connues sont les suivantes : la méthode de quotas, la méthode de volontaires, la méthode de boule neige, la méthode d'échantillonnage ciblé et la méthode de panel ouvert.

#### ➔ La méthode de quotas

La méthode de quotas est une technique d'échantillonnage qui s'impose à chaque fois qu'il faut utiliser un échantillon représentatif alors qu'il n'est pas facile de disposer d'une liste de sondage

et que la structure de la population à étudier est connue. Elle repose non plus sur le hasard, au sens statistique du terme, mais sur le raisonnement de l'analyste qui choisit des éléments composants l'échantillon en fonction de critères de sélection jugés pertinents en regard du problème à résoudre. Les critères de sélection sont appelés « variable de quotas ou de contrôle ». Cette méthode consiste à construire un échantillon représentatif à partir de certains critères clés appelés « variable de quota ou de contrôle ».

Cette méthode est largement utilisée dans les enquêtes d'opinion et les études de marché notamment parce qu'il ne suppose pas de liste des individus de la population. On parle aussi d'échantillonnage dirigé ou par choix raisonné. On demande aux enquêteurs de faire un nombre d'entrevues dans divers groupes établis en tenant compte de la répartition de l'échantillon à partir des variables de quotas. L'enquêteur doit respecter son quota.

Il faut noter qu'on ne doit retenir qu'un nombre restreint de quotas (au maximum trois variables de quotas).

Avantages : L'échantillonnage par quotas est généralement moins coûteux que l'échantillonnage aléatoire. Il est également facile à administrer.

Inconvénients : Certaines unités peuvent n'avoir aucune chance d'être sélectionnées (voir la méthode de la **théorie des valeurs extrêmes**).

➔ La méthode de volontaires ou de convenance

On prélève l'échantillon à partir d'un groupe de volontaires (Exemples : expériences médicales ou psychologiques). Dans ce cas, l'échantillon est composé de toutes **les personnes** que l'analyse considère **concernées** (population cible) **et volontaires** (acceptent d'y participer) par le sujet de l'étude.

Avantages : Cette méthode est très facile à appliquer car l'enquêteur ne se préoccupe pas de convaincre les enquêtés. Il ne s'intéresse qu'à ceux qui ont accepté de répondre.

Inconvénients : échantillon biaisé car aucune stratégie de sélection n'est suivie pour s'assurer que l'échantillon ainsi constitué est bien représentatif de la population cible.

➔ L'échantillonnage ciblé ou échantillonnage à la place

Cette méthode consiste à obtenir un échantillon d'éléments recrutés en **des lieux où la probabilité de rencontrer les personnes** concernées par la thématique de l'étude **est très élevée**<sup>3</sup>.

Avantages : Ce type d'échantillon est largement utilisé pour travailler sur des cibles structurellement bien identifiées mais dispersées ou de faible effectif.

---

<sup>3</sup> Exemple d'une étude auprès des prostitués ou auprès des enfants de la rue.

---

Inconvénients : échantillon biaisé.

➔ La méthode de boule de neige

C'est une méthode dont l'échantillon est composé d'**individus recommandés par des personnes déjà enquêtées**.

Avantages : Ce type d'échantillon est largement utilisé en milieu industriel. Il est également utile pour atteindre des populations très particulières ou pour pénétrer des milieux fermés.

Inconvénients : difficulté d'obtention d'un échantillon représentatif si vous n'avez pas pu avoir des recommandations.

➔ La méthode d'access panel ou de panel ouvert

L'access panel est un échantillon composé d'**individus volontaires qui ont été recrutés au cours des études précédemment réalisées**. L'usage de cette méthode est fréquent pour les études par Internet.

Avantages : Son principale avantage est de ne pas avoir à rechercher au hasard des personnes à interrogées, mais de sélectionner sur une liste celles qui répondent aux caractéristiques recherchées et de les contacter.

Inconvénients : Si la base n'est pas actualisée, on risque de ne pas retrouver les enquêtés au cas où ils avaient changé d'adresse ou de numéro de téléphone.

### **2.3 L'élaboration du questionnaire**

Si toutes les phases de l'enquête sont importantes et doivent être exécutées avec beaucoup de soin, la réussite finale de l'enquête est fonction de **la nature et de la qualité des données recueillies**. Des données de qualité exigent essentiellement un questionnaire bien conçu et des **procédures uniformes** pour le remplir. Le questionnaire est au cœur du processus de collecte des données. Il a des répercussions importantes sur la qualité des données parce qu'il constitue le moyen de collecte des données. Il a aussi des répercussions sur l'image de marque que l'organisme statistique projette dans le public.

Les questions posées doivent être **conformes à l'énoncé des objectifs** de l'enquête et permettre la collecte d'information utile pour l'analyse des données. Elles doivent **répondre à tous les besoins d'information**, mais chaque question devrait avoir une justification explicite pour être inscrite dans le questionnaire. Il faut savoir **pourquoi chaque question** est posée et **à quoi servira cette information**. La formulation de **la question doit être claire**. Les questions doivent être **réparties en séquences logiques** pour le répondant. Les questions doivent être formulées pour être **faciles à comprendre** et permettre au répondant d'y répondre précisément. Enfin, le questionnaire devrait être **mis à l'essai** avant son application, à l'aide de groupes de discussion, d'un prétest et d'autres méthodes décrites dans ce chapitre.

---

Le type d'informations que nous obtiendrons lors des entrevues avec les enquêtés dépendra de **la nature des questions** que nous leur poserons et de la manière dont nous les poserons.

La rédaction d'un questionnaire doit répondre à de nombreux critères : les questions doivent **être formulées de façon claire et simple** ; elles doivent être **ordonnées d'une façon logique** ; elles doivent enfin donner lieu à des réponses précises et objectives.

**La longueur du questionnaire** et sa facilité d'utilisation sur le terrain sont aussi des critères importants.

Pour concevoir un questionnaire, il faut respecter les étapes suivantes :

*2.3.1 Consultation avec les experts du domaine, les utilisateurs des données et les répondants* : Le statisticien ne connaît pas tout, mais il est recherché partout. Dans son travail, il est dans l'obligation de travailler avec des gens qui maîtrisent le domaine dont il effectue l'étude. Le processus de consultation avec les **utilisateurs des données** commence lors de la formulation des objectifs de l'enquête au cours de la phase de planification et continue pendant la conception et l'élaboration du questionnaire. Cette consultation approfondie est particulièrement importante pour les grandes enquêtes, sinon toutes, d'un organisme statistique. Une compréhension approfondie de l'utilisation des données devrait permettre à l'organisme statistique d'élaborer un questionnaire bien conçu qui répond aux besoins des utilisateurs.

Il faudrait consulter non seulement les utilisateurs des données, mais aussi les **répondants**, les **experts de la matière** de l'étude et **ceux qui ont procédé à des enquêtes semblables auparavant**, avant de formuler la version provisoire du questionnaire. Ils devraient pouvoir donner une rétroaction sur le genre d'information que les répondants peuvent fournir et aider à préciser les concepts à étudier. Rencontrer les répondants **peut aider à identifier les questions et les préoccupations importantes** pour eux et à obtenir des répercussions sur les décisions pertinentes à la matière du questionnaire. Cette intervention peut aussi **aider à identifier les expressions et le langage qu'utilisent les répondants** pour décrire les concepts de l'enquête, et **donner une bonne idée de la façon dont les catégories de questions et réponses devraient être formulées**. Ces discussions peuvent se dérouler pendant des consultations approfondies ou en groupe de discussion ;

*2.3.2 Examen des questionnaires précédents* : D'autres enquêtes sont une bonne source d'information pour l'élaboration d'une enquête. L'examen des questions posées dans d'autres enquêtes sur le même sujet ou un sujet semblable peut être un bon point de départ lorsqu'il faut formuler une question (c.-à-d. rédiger une question). Lorsque l'on souhaite comparer les résultats de différentes enquêtes, il est préférable d'utiliser les mêmes questions. Il faudrait aussi examiner la documentation sur la qualité des données de ces enquêtes pour évaluer l'efficacité

---

du questionnaire (par exemple, les problèmes de rédaction des questions, le fardeau de réponse, les taux de refus, etc.).

**2.3.4 Rédaction de la version provisoire du questionnaire** : Il est important de procéder à la mise à l'essai de toutes les versions (c.-à-d. les versions dans toutes les langues) du questionnaire auprès de répondants bien avant le début de la « vraie » collecte des données. Répondre à une question est un processus complexe. Les **répondants doivent d'abord comprendre la question**. Ils doivent ensuite faire un **effort de mémoire** ou fouiller des dossiers pour extraire l'information demandée. Ils doivent aussi réfléchir à la réponse exacte à la question et déterminer s'ils sont disposés à révéler l'information, en tout ou en partie. Ils répondent alors à la question. Chacun de ces processus peut être une source d'erreur.

**2.3.5 Examen du questionnaire auprès des utilisateurs des données** ;

**2.3.6 Mise à l'essai et révision du questionnaire** : Les méthodes appliquées aux mises à l'essai des questionnaires (matière, présentation, etc.) sont habituellement axées sur de petits échantillons subjectifs non probabilistes de répondants tirés de la population cible. Voici les méthodes décrites dans les sections suivantes :

➔ **Prétest** : Le prétest (parfois intitulé essai préliminaire) est facile, le coût est raisonnable, et c'est une étape fondamentale de l'élaboration d'un questionnaire. S'il n'y a pas d'autres mises à l'essai du questionnaire, il faudrait au moins faire un prétest. La taille de l'échantillon du prétest peut varier de 20 à 100 répondants ou plus. Si le principal objectif est de repérer des problèmes de rédaction ou de séquence, très peu d'interviews sont nécessaires. Il faut en faire davantage (de 50 à 100) pour déterminer les catégories de réponse aux questions fermées, à partir des réponses aux questions ouvertes du prétest. Lors du prétest, le répondant n'est pas informé, il remplit simplement le questionnaire ou répond à l'interview pour refléter la situation lors de la collecte réelle des données. Le prétest indique seulement là où il y a un problème. Sans aller plus loin, il ne détermine pas pourquoi il y a un problème ou comment le corriger. La mise à l'essai non officiel n'identifiera peut-être pas non plus tous les problèmes du questionnaire.

**Voici à quoi sert le prétest d'un questionnaire** : découvrir l'ordre ou la rédaction médiocres des questions, corriger des sauts, fermer certaines questions ouvertes, retrouver les modalités manquantes (NSP), reconnaître les problèmes des non réponse (ne veut pas répondre) et des hors-champs (ne peut pas répondre), estimer la longueur d'un questionnaire, repérer les erreurs de présentation ou d'instructions du questionnaire.

➔ **groupes de discussion** : Un groupe de discussion considère un sujet sélectionné par les participants choisis dans la population d'intérêt. Au cours des premières étapes de

---

l'élaboration du questionnaire, les groupes de discussion peuvent aider à identifier les questions saillantes de la recherche. On fait appel à eux pour évaluer la compréhension du langage et de la rédaction des questions et des instructions de la part du répondant, ainsi que d'autres formulations et mises en forme des questions. Un animateur qui connaît bien les techniques d'interview des groupes et l'objectif de la discussion oriente le groupe de discussion. Chaque groupe comprend habituellement de six à douze personnes et la taille optimale est de sept à neuf personnes. Une séance en groupe de discussion demande habituellement deux heures environ. Le groupe de discussion est enregistré sur bande sonore (et parfois sur bande vidéo) que les observateurs peuvent entendre dans une salle derrière un miroir d'observation. Il est recommandé que ceux qui élaborent le questionnaire observent le groupe de discussion. **Les observateurs n'interviennent pas dans la discussion du groupe, mais leurs observations peuvent servir à l'animateur à la fin de la séance du groupe de discussion.**

➔ **essai pilote** : Une enquête pilote se déroule pour observer toutes les étapes du processus de l'enquête, y compris l'administration du questionnaire. Une enquête pilote est une « simulation » qui applique la version finale du plan d'enquête à petite échelle du début à la fin, y compris le traitement et l'analyse des données. Elle permet à l'organisme statistique de considérer les résultats du questionnaire pendant toutes les étapes de l'enquête (collecte, vérification, imputation, traitement, analyse des données, etc.). Le questionnaire est habituellement soumis à des essais approfondis à l'aide des méthodes susmentionnées avant l'enquête pilote.

*2.3.7 Touche finale apportée au questionnaire* : La conception du questionnaire est un processus itératif : des modifications sont continuellement apportées pendant l'élaboration et la mise à l'essai du questionnaire. Les objectifs et les besoins d'information sont formulés et réévalués, les répondants et les utilisateurs des données sont consultés, la version préliminaire des questions proposées est formulée et mise à l'essai, les questions sont examinées et révisées jusqu'à la formulation de la version finale du questionnaire. Lorsqu'il est décidé qu'il n'y aura pas d'autres modifications apportées au questionnaire, l'étape finale du processus est franchie. La touche finale est alors apportée au questionnaire et il est imprimé ou programmé, selon la méthode de saisie des données appliquée.

Une autre décision dans la conception d'un questionnaire concerne le type de réponse qu'il faut associer à chaque question : une réponse libre pour une question ouverte et une réponse imposée pour une question fermée.

**La formulation d'une question ouverte** laisse au répondant toute la latitude pour construire librement sa réponse, avec ses propres mots et spontanément : aucune réponse lui est imposée. L'usage d'une question ouverte s'impose dans différents cas de figure :

- Quand le contenu issu de la spontanéité constitue même l'objet de l'information recherchée. **C'est le cas pour une étude qui veut savoir si les répondants ont une bonne connaissance sur un problème, par exemple ;**

- Lorsque la diversité des réponses est beaucoup plus étendue pour établir une liste exhaustive des modalités de la question (la profession d'une personne);

- Quand on n'est pas sûr de connaître toutes les réponses possibles à la question.

Les différentes formes des questions ouvertes sont :

- Question à réponse textuelles : quel est le plus important problème au Congo ?

- Question à réponse numérique : Au cours des 12 derniers mois, combien de véhicules avez-vous lavés ?

**La formulation d'une question fermée** contient les modalités de réponses attendues entre lesquelles le répondant doit impérativement choisir. sa rédaction implique donc que l'analyste connaisse les réponses qui peuvent être données et qu'il sélectionne celles qui l'intéressent en regard des objectifs de l'étude.

Il existe plusieurs types de questions fermées :

- Question dichotomique : La question dichotomique est la version la plus simple d'une question fermée. Il s'agit souvent d'une question qui n'a que deux réponses possibles (oui/non, pauvre/non pauvre, masculin/féminin, etc.) ;

- Question à choix unique : c'est une question qui demande au répondant de **ne choisir qu'une seule réponse** dans la liste. Exemple : Quel est le réseau téléphonique que vous utilisez le plus quand vous appelez à l'étranger ? (MTN, Airtel, Azur, Congo Telecom, NSP) ;

- Questions à choix multiples: c'est une question qui demande au répondant de choisir **au moins une réponse** dans la liste. Exemple : quels sont les réseaux téléphoniques que vous utilisez ? (MTN, Airtel, Azur, Congo Telecom) ;

- Question avec classement (ou à réponse ordonnée) : La question avec classement est un autre genre de question fermée et elle demande au répondant d'établir l'ordre des catégories de réponse, par exemple : "Quelles sont, selon vous, les priorités au Congo ? Veuillez les classer par ordre de priorité en inscrivant « 1 » à la réponse qui serait la plus utile, selon vous, « 2 » à la méthode qui serait la plus utile en second lieu, et ainsi de suite. (lutte contre le chômage, lutte contre la pauvreté, lutte contre l'inflation, lutte contre les inégalités, la santé, l'éducation, l'autosuffisance alimentaire).

---

Un questionnaire doit être bien ordonné. Il est souvent séparé par des sections. Chaque section est constituée des questions.

Les sections doivent être ordonnées. Il faut commencer par des sections ayant des questions introductives.

Exemple :

- 1- Nous voulons d'abord obtenir des renseignements généraux sur vous ;
- 2- Nous voulons maintenant savoir.....
- 3- Les questions suivantes portent sur .....
- 4- Pour ne pas trop abuser de votre temps, nous terminerons par deux ou trois petites questions.

Avant de commencer son interview, il faut d'abord se présenter en commençant par :

- 1- Donner le titre ou le sujet de l'enquête (ne pas dire enquête, mais plutôt étude) ;
- 2- Identifier le commanditaire ;
- 3- Exprimer l'objectif de l'enquête ;
- 4- Expliquer pourquoi il est important de remplir ce questionnaire (ne pas mentir) ;
- 5- Souligner comment seront utilisées les données ;
- 6- Préciser comment le répondant peut avoir accès aux résultats de l'étude ;
- 7- Indiquer que les réponses seront confidentielles ;

Vous devrez maîtriser votre questionnaire et éviter d'être trop lent.

La première question que l'on pose doit être facile à répondre, fermée et doit rendre le répondant à l'aise (pas stressant). Ne pas commencer par exemple à demander l'âge de l'enquêté.

Poser les questions les plus délicats quand le répondant se sent à l'aise et confiant (gentil, rassuré).

## **2.4 Collecte des données**

Il existe plusieurs éléments de collecte de données : **Papier, tablette, téléphone, ordinateur (internet)**. La méthode sur support papier est intitulée Interview Papier et Crayon (IPC) et la méthode sur support électronique est intitulée Auto-Interview Assistée par Ordinateur (AIAO) ou Interview sur Place Assistée par Ordinateur (IPAO) ou encore Interview Téléphonique Assistée par Ordinateur (ITAO). La différence entre la AIAO, la ITAO et la IPAO se reconnaît au niveau des différentes méthodes de collecte des données.

Voici les méthodes élémentaires de collecte des données :

### **2.4.1 Autodénombrement**

Le répondant remplit le questionnaire d'enquête par autodénombrement sans l'aide d'un intervieweur.

---

Divers moyens peuvent servir à envoyer le questionnaire au répondant et à le retourner à l'expéditeur : le service postal, le télécopieur, un moyen électronique (y compris Internet) ou un enquêteur. (Si le questionnaire est retourné par télécopieur ou sur support électronique, une ligne sécuritaire ou le chiffage est alors nécessaire pour garantir la confidentialité des données du répondant). La méthode sur support papier est intitulée Interview Papier et Crayon (IPC) et la méthode sur support électronique est intitulée Auto-Interview Assistée par Ordinateur (AIAO).

#### 2.4.2 *Interviews sur place ou le face à face*

Un intervieweur aide le répondant à remplir le questionnaire. L'interview se déroule sur place, habituellement à la résidence du répondant ou en milieu de travail, même si elle peut avoir lieu dans un endroit public (p. ex., aéroport, centre commercial). La méthode sur support papier est intitulée Interview Papier et Crayon (IPC) et la méthode assistée par ordinateur est intitulée Interview sur Place Assistée par Ordinateur (IPAO).

#### 2.4.3 *Interviews téléphoniques*

Un intervieweur aide le répondant à remplir le questionnaire au téléphone. La méthode sur support papier est intitulée interview papier et crayon (IPC) et la méthode assistée par ordinateur est intitulée Interview Téléphonique Assistée par Ordinateur (ITAO).

### 2.5 L'équipe de l'enquête

Au moment où notre questionnaire sera prêt à être prétesté sur le terrain, nous devons avoir rassemblé tout notre personnel pour l'enquête. Outre notre équipe de planification de l'enquête, il nous faudra un personnel de terrain (enquêteurs, contrôleurs et superviseurs), un personnel de traitement des données (codificateurs, tabulateurs et le personnel de saisie de données) et un personnel d'analyse de données.

Les effectifs dont nous aurons besoin dépendront évidemment de la dimension de l'enquête et du délai prévu pour la collecte et le traitement des données. Pour choisir l'équipe d'enquête, nous devons prendre en compte les facteurs suivants :

➔ La personnalité : les membres de l'équipe doivent **être** capable de bien travailler ensemble, de s'adapter, de s'aider les uns les autres et d'avoir une attitude amicale envers les personnes interrogées ;

➔ Le niveau d'instruction des enquêteurs ;

➔ L'honnêteté des enquêteurs et des contrôleurs ;

➔ L'esprit managérial des superviseurs ;

➔ L'expérience antérieure des **enquêteurs** ;

➔ Les langues et dialectes parlés et écrits ;

→ L'état de santé des enquêteurs ;

→ **Le sexe** : si ce facteur est important pour des raisons culturelles dans la composition des équipes d'enquêteurs.

Choisir des personnes qui connaissent bien la zone de projet est un atout.

## 2.6 La formation des agents

Nous devons prévoir un temps et des moyens suffisants pour apprendre aux **enquêteurs**, aux **contrôleurs** et aux **superviseurs** à accomplir leurs tâches correctement, efficacement et de manière standardisée.

L'enquêteur est un agent de terrain. Il a pour objectif de :

- 1- Prendre l'information à la base ;
- 2- Maitriser toutes les questions du questionnaire ;
- 3- Poser clairement les questions aux enquêtés ;
- 4- Transmettre dans le questionnaire toutes les réponses avec le maximum de fidélité ;
- 5- Retourner tous les questionnaires et matériels<sup>4</sup> remis par le contrôleur en parfait état;
- 6- Respecter toutes les instructions transmises par le contrôleur.

Le contrôleur est agent qui suit et encadre les enquêteurs sur trois aspects :

1- Organisationnel : il organise le travail de son équipe, sensibilise les unités de son échantillon, s'assure que chaque membre de son équipe dispose du matériel nécessaire au bon fonctionnement de l'enquête et transmet les informations importantes à son superviseur ;

2- Technique : il s'assure que l'enquêteur a bien rempli ses questionnaires et a couvert toutes les unités à enquêter de son échantillon ;

3- Pratique : il instaure au sein de son équipe un climat de confiance susceptible de favoriser le bon déroulement du travail. En cas de problème ou de conflit, il est le responsable.

Il veille également au bon entretien et à la bonne conservation du matériel de l'enquête.

Le superviseur est le principal responsable de l'enquête dans sa zone d'étude. Il organise le travail concernant toute sa zone, assure le tirage des échantillons de sa zone, sensibilise les chefs de sa zone, transmet les instructions importantes de la coordination à ses contrôleurs, valide les travaux effectués par tous les enquêteurs après vérification des contrôleurs et récupère tout le matériel auprès des contrôleurs.

La formation des intervieweurs doit être soigneusement planifiée pour qu'ils aient tous un rendement uniforme et la même compréhension des concepts de l'enquête.

---

<sup>4</sup> Ça peut être des badges, tablettes, sacs, etc.

---

Un bon programme de formation et un manuel<sup>5</sup> clairement rédigé sont essentiels si l'on veut avoir un personnel de terrain efficace.

L'étendu de la formation nécessaire dépend de la complexité de l'échantillon et du questionnaire : les enquêteurs, les contrôleurs, les superviseurs et les agents de saisie doivent comprendre parfaitement toutes les procédures qu'ils auront à appliquer.

## **2.7 Le traitement et analyse des données**

Après la collecte des données, il est question de saisir les informations obtenues dans l'ordinateur. Avant cela, il est demandé de concevoir un masque de saisie. Ce masque de saisie peut se faire au même moment que l'enquête se déroule ou même avant l'enquête.

Une fois que les données sont saisies, il est indispensable de les traiter.

### *2.7.1 Codification de données*

La codification des données consiste essentiellement à traduire en chiffre les mots et réponses qui figurent sur les questionnaires.

Les résultats de notre enquête doivent être compilés par tabulations manuelles ou par ordinateur. La plupart ou la totalité des réponses figurant sur chaque questionnaire devrait être converties en code numérique avant de pouvoir être analysées de façon significative.

### *2.7.2 Apurement des données*

On apure des données pour résoudre :

- ➔ Des incohérences entre les données d'un même questionnaire ;
- ➔ Des entrées peu vraisemblables (les valeurs atypiques) ;
- ➔ Des entrées impossibles (des erreurs d'entrée des valeurs)

Ces vérifications peuvent se faire manuellement ou par ordinateur.

La méthode d'apurement est aussi appelée « traitement des données ». Elle consiste à revoir les données provenant du terrain afin de s'assurer qu'elles ne sont pas de mauvaise qualité. Il est question ici de traiter les hors-champs, les non-réponses et les valeurs atypiques.

### *2.7.3 Pondération*

Cette méthode consiste à ajuster les variables clés utilisées pour la section de l'échantillon à leurs proportions réelles dans la population.

Lorsque nous rendons compte des résultats, nous devons préciser si nous avons utilisé des pourcentages pondérés ou non.

---

<sup>5</sup> Ce manuel s'appelle : le manuel de l'enquêteur ;

#### 2.7.4 Analyse des données

Lorsqu'on a des données, on peut faire deux sortes d'analyses à savoir : l'analyse descriptive et l'analyse inférentielle.

### **2.8 Les erreurs**

Les méthodes d'échantillonnage peuvent être sources d'erreurs. Un certain nombre d'erreurs pourront être éliminées, certaines pourront être réduites, mais d'autres persisteront.

#### 2.8.1 Les erreurs dues aux instruments de mesure

Un instrument est fidèle s'il répond exactement de la même façon quand il est placé dans deux situations identiques.

Exemple le thermomètre, la balance.

Une question claire est dite fidèle quand tout le monde la comprend de la même façon.

Un instrument est valide lorsqu'il mesure vraiment ce qu'il est censé mesurer.

Cependant, les instruments fidèles sont difficiles à obtenir.

2.8.2 Les erreurs dues à l'organisation : Ce sont les erreurs qui se glissent lors de la collecte des données. Est-ce que les consignes ont été respectées? Les enquêteurs ont-ils agi de la même façon? Pour éviter ces erreurs il faut utiliser les mêmes instruments, et que tous les enquêteurs comprennent le questionnaire de la même façon.

2.8.3 Les erreurs dues à la méthode d'échantillonnage : Il faut toujours vérifier, à la lumière des objectifs de l'étude statistique, que la méthode d'échantillonnage est adaptée.

2.8.4 Les erreurs dues au phénomène de non-réponse : Même avec la meilleure méthode d'échantillonnage, il se présente toujours un certain nombre de non-répondants, ce qui peut entacher la représentativité de l'échantillon et amener des conclusions erronées.

2.8.5 L'erreur d'échantillonnage : Le fait d'étudier un échantillon plutôt qu'un autre engendre forcément une erreur. Cette erreur appelée erreur d'échantillonnage est inévitable.

2.8.6 L'erreur de mémorisation : elles en existent deux :

1- L'erreur de mémoire : le répondant ne se souvient plus de tous les événements.  
Exemple : Aviez-vous un téléphone au collège ? (Oui, Non, NSP).

2- L'erreur de télescopage : le répondant ne se rappelle plus exactement de la période de l'événement. soit il déclare l'événement avant (télescopage en aval), soit il déclare après (télescopage en amont). Exemple : En quel âge aviez-vous acquis le premier téléphone ?