

GUIDE N°1: BIENVENUE DANS LE MONDE DES DONNÉES

L'ère des données

Les données sont partout, qu'on les remarque ou non. Elles influencent notre vie, souvent à notre insu. Que nous soyons actifs sur les réseaux sociaux ou simples consommateurs, notre comportement est recensé, scruté.

Les données et les statistiques inspirent souvent un sentiment de véracité scientifique. Pourtant, elles servent aussi à manipuler et à nous influencer. D'où l'importance de bien comprendre les données et leur mode de production.

Ce petit guide trace les premiers pas dans le monde des données. Il s'adresse principalement aux étudiant(e)s amené(e)s à réaliser un travail de recherche et permet de réfléchir aux questions suivantes:

Qu'est-ce que des données? Quelles approches choisir? Produire des données ou réutiliser des données existantes?

Qu'est-ce que des données ?

Les données consistent en des observations ou des mesures collectées à des fins de recherches scientifiques. Ce sont des sources d'information au sujet d'une population spécifique.

Les données sont des informations. Toutefois, c'est le traitement et l'analyse des données qui leur donnent un sens, une signification et qui permettent d'obtenir des résultats de recherche. Elles sont des observations et des chiffres recueillis et analysés pour mieux comprendre le monde. Elles permettent ainsi d'établir des faits, d'appuyer des décisions, des priorités, des jugements.

Les données se rapportent à une question ou problématique particulière. Elles peuvent donc se référer à différentes unités (un individu, un groupe d'individus ou même des objets) et il existe différentes manières de les récolter.

Dans ce qui suit, nous allons distinguer différents types de données et manières de les obtenir:

- les approches quantitatives, qualitatives et les approches mixtes
- les données primaires et secondaires



©Jumpeestudio - Fotolia

LE RECUEIL DES DONNÉES: QUELLE APPROCHE CHOISIR POUR MA RECHERCHE?

L'approche qualitative génère des données non numériques, tandis que l'approche quantitative procure des données numériques ou des informations qui peuvent être converties en chiffres. Le choix de l'une, de l'autre, ou des deux approches combinées dépend de la problématique et des informations recherchées. Chacune présente des avantages et inconvénients.

L'approche qualitative

L'approche qualitative vise à découvrir des comportements, des événements ainsi que les raisons de ces différents types de comportements et événements. Cette approche convient à des recherches dont le but est de comprendre en profondeur comment et pourquoi un événement ou phénomène particulier s'est déroulé.

Les données qualitatives peuvent être recueillies au travers de différents outils de collecte tels que des grilles d'entretiens (canevas de questions) utilisées durant des interviews individuels ou collectifs ou encore des observations.



- L'approche qualitative permet d'aborder des phénomènes en profondeur et de façon détaillée.
- L'approche qualitative permet de saisir le verbal et le non verbal (comportements, gestes...).
- L'approche qualitative permet de produire des résultats avec un échantillon de petite taille.



- L'approche qualitative ne permet souvent qu'une généralisation limitée.
- Les résultats de l'analyse de données qualitatives ne permettent pas de quantifier un phénomène à grande échelle.
- Le recueil de données qualitatives représente souvent un grand investissement de la part du chercheur (temps pour mener et transcrire les entretiens).



Données qualitatives: qu'est-ce que c'est?

Les données qualitatives consistent en des données descriptives relatives à des paroles (dites ou écrites) ou à des comportements observés.

En voici quelques exemples:

- Des documents écrits comme des transcriptions d'entretiens, des articles de presse, des livres ou encore des rapports;
- Des notes d'observation;
- Des enregistrements audio ou vidéo.

L'approche quantitative

L'approche quantitative vise à mesurer des comportements ou des caractéristiques et se focalise sur le comptage, les classifications de caractéristiques.

Elle est privilégiée pour des recherches dont le but est de répondre à une question par une évidence numérique.

Dans cette approche, des méthodes mathématiques et statistiques sont utilisées pour aboutir à des résultats.

Cette approche permet de produire des données quantitatives également appelées numériques qui permettent de dresser un portrait global d'une situation, d'une population.



- L'approche quantitative permet de quantifier des phénomènes.
- L'approche quantitative considère des populations de grandes tailles et permet ainsi des généralisations.
- L'approche quantitative permet l'observation de tendances sur la durée.



- L'approche quantitative se limite à mesurer de façon générale certains types d'informations (ex: sentiment, comportement, émotion...).
- L'approche quantitative ne permet pas de saisir un phénomène autant en profondeur que l'approche qualitative.
- L'approche quantitative exige des connaissances en statistiques. Les analyses sont souvent effectuées à l'aide de logiciels spécialisés (ex: SPSS).

L'approche mixte

L'approche mixte combine des approches quantitatives et qualitatives soit simultanément soit successivement. Elle permet ainsi de rassembler des données qualitatives et quantitatives sur une même thématique et d'enrichir les résultats de recherche:

- lorsque des données qualitatives sont nécessaires pour interpréter des résultats quantitatifs;
- lorsque des données quantitatives sont nécessaires pour pouvoir généraliser des résultats qualitatifs;
- quand une thématique jusqu'ici peu développée doit être à la fois explorée au travers de données qualitatives et mesurée par des données quantitatives.



Données quantitatives: qu'est-ce que c'est?

Les données quantitatives sont des informations numériques ou statistiques.

Elles proviennent le plus souvent:

- d'enquêtes par questionnaires
- de registres ou dossiers tenus par des institutions publiques ou privées (ex: tribunaux, hôpitaux, entreprises)
- automatiquement, sans qu'on s'en rende compte (ex: Facebook, carte bancaire, carte de fidélité...).

PRODUCTION OU RÉUTILISATION DE DONNÉES: QUELLES DONNÉES POUR MA RECHERCHE?

Les données primaires sont des données recueillies par les chercheurs eux-mêmes contrairement aux données secondaires, qui consistent en des données disponibles et recueillies par d'autres. Le choix de produire des données primaires ou d'utiliser des données secondaires dépend de la question de recherche et de la disponibilité des données.

Les données primaires: la production de données

Les données primaires sont collectées "sur le terrain". Elles sont spécialement produites en vue de réaliser une enquête ou une étude, c'est-à-dire pour répondre à une problématique spécifique. Elles sont ainsi obtenues directement au travers d'enquêtes, d'observations ou d'expérimentations, que ce soit au travers d'une approche qualitative ou quantitative.



Comment produire des données primaires?

- Identifier les questions de recherche et les méthodes pour y répondre.
- S'assurer qu'il n'existe pas encore de données sur la thématique.
- Déterminer les personnes à solliciter.
- Déterminer une approche et un outil pour recueillir des données : questionnaires papier ou en ligne, interviews en face à face, interviews téléphoniques, focus group,...
- Recueillir les données puis les préparer pour l'analyse: transcription, anonymisation, sauvegarde dans une base de données, ...
- Pour aller plus loin sur ce sujet, vous pouvez consulter le Guide n°3 au sujet de la gestion des données.



- Les données primaires permettent de récolter les informations qui correspondent exactement aux problématiques de recherche.
- La collecte de données primaires permet une meilleure connaissance des données (qualité, processus de récolte,...).
- La production de données primaires permet d'avoir des données récentes et exclusives (pour l'analyse et les publications).



- La collecte de données primaires peut avoir un coût important.
- La collecte et la préparation de données primaires peuvent être longues.
- La collecte de données primaires implique des compétences tout au long du processus d'enquête et des différentes étapes (développement de l'outil de collecte, traitement des données, nettoyage des données...).

Les données secondaires: la réutilisation de données

Les données secondaires sont des données de “seconde main”. Elles ont déjà été recueillies, traitées, interprétées par d'autres. Elles constituent une information déjà existante et accessible.

Les données secondaires peuvent provenir de différentes sources: travaux de recherche, articles ou autres publications scientifiques, statistiques officielles (nationales ou internationales), instituts de sondage,...

Selon les données et informations nécessaires à une recherche, il s'agit de déterminer la ou les sources disponibles et s'assurer de la pertinence des données et de leur réutilisation pour sa propre recherche.



- Les données secondaires sont peu coûteuses.
- Les données secondaires sont directement accessibles. Le recours à des données secondaires permet d'éviter les processus de mise en oeuvre d'une enquête.
- Il existe énormément de données secondaires offrant encore de nombreuses possibilités d'analyses.



- Les données secondaires peuvent ne pas correspondre exactement à la problématique de recherche.
- Les données secondaires peuvent être anciennes ou inexistantes.
- La qualité des données peut être difficile à estimer (documentation lacunaire, erreurs de traitement, données incohérentes...).



www.forscenter.ch, dataservice@fors.unil.ch



Où trouver des données secondaires?

- [FORS](#) Le centre de compétences suisse en sciences sociales met à disposition des données secondaires provenant d'enquêtes nationales et internationales. Pour savoir comment obtenir des données chez FORS, consultez le Guide n°2.
- [OFS](#) L'office fédérale de la statistique permet d'accéder à des données secondaires officielles concernant la Suisse.
- [Eurostat](#) L'office de la statistique de l'Union européenne permet d'accéder à des données de pays européens y compris la Suisse.
- [OCDE](#) L'Organisation de Coopération et de Développement Economique donne accès à des données sur les pays membres et non membres de l'OCDE.
- [UNdata](#) Ce sont les données internationales des Nations Unies.
- [La Banque mondiale](#) Cette banque de données contient des données internationales sur le développement.