

Introduction

- INTRODUCTION
- POURQUOI ANALYSER DES DONNÉES ?
- LE PROCÉDÉ STATISTIQUE



À LA FIN DE CE CHAPITRE VOUS DEVEZ ÊTRE CAPABLE DE :

- EXPLIQUER CE QU'ON ENTEND PAR STATISTIQUE
- EXPLIQUER L'IMPORTANCE DES STATISTIQUES
- DÉFINIR LES TERMES POPULATION, RECENSEMENT ET ENQUÊTE
- EXPLIQUER LE PROCÉDÉ STATISTIQUE

CHAPITRE 1

INTRODUCTION

Introduction

comprendre les statistiques

Ce cours est destiné à faire comprendre ce que sont l'analyse de données et la manière de présenter les résultats d'une analyse dans un rapport. Les ouvrages techniques contiennent régulièrement des résultats basés sur des analyses statistiques. Quel que soit votre domaine professionnel vous avez une certaine connaissance en analyse statistique de base. Les méthodes qui constituent l'analyse statistique ne sont pas simplement une liste de procédures statistiques et de formules, C'est plutôt le raisonnement sur lequel s'appuie souvent la recherche scientifique.

À mesure qu'on étudie la statistique, on doit établir des définitions pour son vocabulaire et ses concepts. Dans l'usage courant, les mots peuvent avoir des significations assez larges et être utilisés dans des cas très différents. En statistique, les mots doivent être définis très soigneusement de manière à ce que chacun sache précisément de quoi on parle.

Pourquoi analyser des données ?

La **population observée** est l'ensemble des unités pour lesquelles l'information doit être rassemblée. On l'appelle également **champ** de la collecte. Il est bien entendu souhaitable que la population observée coïncide avec la population-cible, mais ça n'est pas toujours possible. Par exemple, la population-cible dans un recensement de la population peut être la population entière d'un pays, mais ceci risque de ne pas être pratique car se déplacer sur des îles isolées peut se révéler trop onéreux. Dans ce cas, la population peut être réduite aux personnes se trouvant sur les îles regroupant le plus fort pourcentage de la population totale ou de la population-cible. Les résultats de la collecte peuvent alors seulement prétendre représenter la population observée, et il est alors important que les utilisateurs des résultats de l'enquête soient informés de la couverture de la population observée.

La **population-cible** est l'étendue de l'enquête ...

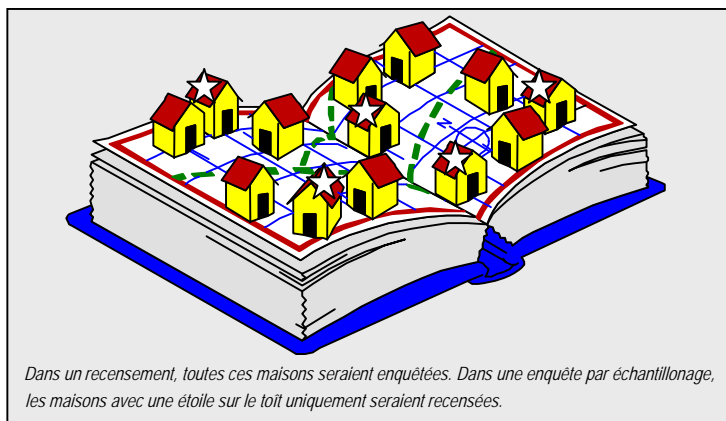
l'ensemble d'unités pour lesquelles on veut des informations.

La population **observée** est la couverture de la collecte ...

la série d'unités dont on tire les informations.

deux types d'enquêtes

Les termes **recensement** et **échantillon** se réfèrent à l'étendue de la collecte entreprise. Dans un recensement, les données sont collectées à partir de toutes les unités de la population observée. De ce fait, tant que la population est bien définie et que les données sont bien collectées à partir de chaque unité, un résultat exact sera obtenu. Dans un échantillon, les données requises sont collectées à partir d'un sous-ensemble d'unités dans la population, et on les utilise pour estimer les résultats de l'ensemble de la population. Ce procédé introduit une erreur appelée **erreur d'échantillonnage**, mais ces erreurs peuvent être contrôlées et l'échantillonnage est souvent préférable pour son faible coût et sa rapidité.



AVANTAGES DES SONDAGES PAR RAPPORT AUX RECENSEMENTS

- 0 Coûts réduits – en terme d'argent et de personnel
- 0 Délais réduits pour produire les résultats
- 0 Permet de collecter les données de manière plus précise (moins d'absence de réponse, moins d'erreurs de collecte et de traitement)
- 0 Permet de poser des questions plus précises

INCONVENIENTS DES SONDAGES PAR RAPPORT AUX RECENSEMENTS

- 0 Les résultats (ou estimations) sont l'objet d'erreurs d'échantillonnage
- 0 Les tableaux croisés détaillés et les résultats de petites zones géographiques et sous-populations peuvent se révéler trop inexacts pour être utiles
- 0 Le sondage ne fournit pas de base de sondage à utiliser par la suite
- 0 Il est difficile d'en communiquer les résultats

Les recensements – un peu d'histoire

Avant JC	3800	BABYLONE	<ul style="list-style-type: none"> ■ Entrepris tous les six ou sept ans; les objets recensés étaient les ânes, les boeufs, le beurre, le lait, le miel et la laine.
	2500	ÉGYPTE	<ul style="list-style-type: none"> ■ Pour estimer la main d'oeuvre disponible pour bâtir les pyramides.
	1491	ISRAËL	<ul style="list-style-type: none"> ■ Pour recenser les personnes aptes au service militaire, et à des fins fiscales.
	550	CHINE	<ul style="list-style-type: none"> ■ Entrepris par Confucius pour obtenir des informations sur l'agriculture, l'industrie et le commerce.
Après JC	1719	PRUSSE	<ul style="list-style-type: none"> ■ C'est le pays qui entreprend les premiers recensements systématiques d'Europe.
	1790	ETATS-UNIS	<ul style="list-style-type: none"> ■ Le premier recensement de l'Amérique.
	1801	FRANCE ANGLETERRE	<ul style="list-style-type: none"> ■ Les premiers recensements en France et en Angleterre.

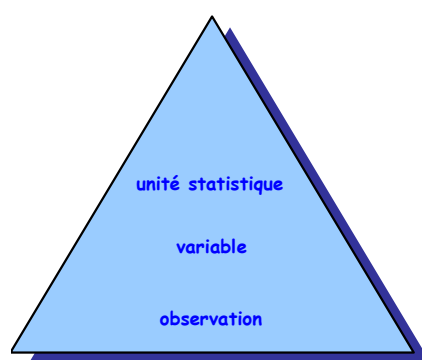
Choisir entre un sondage ou une enquête est une décision difficile à prendre

analyser les données d'enquête

Une fois que vous avez conduit votre enquête les données doivent être analysées. L'analyse statistique simplifie et résume des données. Il existe deux méthodes pour illustrer les séries de mesures : les graphiques et les chiffres. Dans ce cours nous apprendrons les deux techniques.

Le procédé statistique**voici un monde ordonné**

Les données statistiques sont l'observation de variables tirées d'unités statistiques. Les trois termes **observation**, **variable** et **unité statistique** ont une signification précise en statistique. Nous allons définir chacun de ces termes et en étudier quelques exemples.

**observer et enregistrer l'information**

Une **observation** est le résultat de l'observation et de l'enregistrement d'une information à partir d'une source donnée. En général une observation sera menée par une personne, mais elle pourrait l'être aussi par une machine (par exemple une machine pourrait enregistrer automatiquement la température à un endroit donné, à un horaire donné, chaque jour). Voici les observations les plus communes :

mesurer – mesurer les tailles d'enfants scolarisés avec un mètre;

compter – compter les visiteurs à une attraction touristique;

enregistrer les réponses à des entretiens – poser des questions à des personnes et enregistrer leurs réponses;

publiposter des questionnaires – envoyer un questionnaire à des entreprises pour demander leur nombre d'employés;

copier à partir de données existantes – se référer à des documents des douanes et copier des données sur les importations;

utiliser des instruments de mesure automatique – mesurer des températures sur une île isolée avec une machine automatique.

ensuite on a les variables

Observons la valeur d'une variable pour une unité de la population. Le terme **variable** veut dire une caractéristique ou une propriété des unités de la population et, bien sûr, une unité peut avoir de nombreuses variables différentes, dont quelques unes, ou toutes, seront observées. Par exemple, si une force de police entame un recrutement, cette police doit observer les valeurs de plusieurs variables pour chaque participant, de manière à sélectionner les personnes adéquates :

<i>Unités statistiques</i>	<i>Variables</i>	<i>Observations</i>
Postulants	Âge	années de vie
	Taille	hauteur en centimètres
	Poids	poids en kilos
	Vue	acuité visuelle
	Maladies existantes	signes vitaux, par ex. tests sanguins
	Situation de famille	statut marital légal
	Qualifications	qualifications professionnelles
	Casier judiciaire	crimes commis
	Expérience professionnelle	emplois occupés précédemment

les observations combinées de chaque unité statistique forment les données statistiques

Les valeurs des variables sont observées à partir d'**unités statistiques**. Les unités statistiques ne sont pas que des personnes, elles sont aussi n'importe quel détail de la population à observer. Par exemple, des objets, des groupes d'objets, des personnes, des familles, des zones géographiques, des bâtiments, des périodes et des pays.

Les valeurs d'une ou de plusieurs variables d'un certain nombre d'unités statistiques de la même population forment une série de données statistiques. Par exemple, les valeurs totales des exportations d'un pays pour un certain nombre d'années forment une série, avec les années comme unité statistique, alors que les valeurs des exportations de différents produits sur une année forment une autre série, avec les produits comme unité statistique.

ASTUCE

Dans chaque enquête statistique il est très important de définir clairement la population d'unités statistiques dans laquelle les données vont être collectées, les valeurs des variables qui seront observées, et comment l'observation va être faite.