



**PRÉFET
DE LA RÉGION
BOURGOGNE-
FRANCHE-COMTÉ**

*Liberté
Égalité
Fraternité*

Direction régionale
de l'économie, de l'emploi,
du travail et des solidarités

LE STATISTICIEN FACE AUX ENJEUX POSÉS PAR L'UTILISATION DE SOURCES ADMINISTRATIVES

**1. Origine et usages
spécifiques**

**2. Des opportunités à saisir
pour le statisticien**

**3. Des enjeux techniques et
statistiques**

1. Origine et usages spécifiques

Que sont les données administratives?

- Le terme « administrative » qualifie d'abord les données par leur origine, et plus que cela par leur statut. Elles viennent de la sphère administrative au sens large.
- Les données administratives sont souvent issues de processus de gestion interne à l'administration concernée. Mais elles peuvent provenir de déclarations administratives.
- Contrairement aux données statistiques dont le seul but est d'informer, les données administratives doivent leur existence aux actions qu'elles permettent de déclencher.
- Leur usage peut aller au-delà de l'administration concernée.

2. Des opportunités à saisir pour le statisticien

Les données administratives de plus en plus présentes dans l'environnement du statisticien.

Un important flux d'informations disponibles facilement et rapidement .

Leurs atouts :

- Exhaustivité du champ couvert.
- Fraîcheur des données.

Les opportunités pour le statisticien:

- Compléter les données d'enquête quand les appariements sont possibles.
- Voire les substituer aux enquêtes
- Explorer de nouveaux domaines non couverts jusque là par les enquêtes statistiques.

⇒ gain de temps et de budget

3. Des enjeux techniques et statistiques

La plupart des points de vigilance issus de leur caractéristiques :

Un **important flux** d'informations disponibles facilement et rapidement .

les données administratives doivent leur existence aux actions qu'elles permettent de déclencher => **Pas de finalité statistique**

Leurs atouts :

- Exhaustivité du **champ couvert**.
- **Fraîcheur** des données utiles pour la conjoncture mais pas seulement.

⇒ Des enjeux techniques et statistiques qui engendrent une nouvelle façon de travailler pour le statisticien

- Gérer un volume important d'information
- Un processus de production qui diffère des données statistiques et dont il n'existe pas de cadre général

⇒ Déficit de méthode

- Des données « vivantes » :

⇒ Gestion de la temporalité : nécessité de figer les données à un moment T

⇒ Des modifications qui peuvent intervenir n'importe quand et qui ne sont pas prévisibles

⇒ Perte de maîtrise sur les données (définition des concepts, champ couvert, etc.)

Nécessité de travailler avec les organismes concentrant ces flux de données, bien au-delà de la simple fourniture de données.

Une profusion de données de toutes sortes

- Un contexte de profusion de données qui s'accroît (des données venant aussi du secteur privé, d'internet, etc.)

BIG DATA

- Des opportunités et enjeux semblables à ceux qu'on vient de voir
- Des méthodes nouvelles ou remises au goût du jour d'analyse de ces données (méthodes de classification, d'analyses de réseau, analyse textuelle et plus largement les méthodes de machine learning).



**PRÉFET
DE LA RÉGION
BOURGOGNE-
FRANCHE-COMTÉ**

Liberté

Égalité

Fraternité

**Direction régionale
de l'économie, de l'emploi,
du travail et des solidarités**