

Données

Inventaire et qualité

Véronique Tremblay

Retour sur les étapes d'un projet d'analyse de données

1. Définition des objectifs
2. Données
 - Inventaire et qualité
 - Constitution de la base de données
 - Exploration et traitement préliminaire
3. Élaboration et validation des modèles
4. Mise en oeuvre
5. Suivi de la performance et amélioration

Inventaire et qualité des données

Qu'est-ce qu'on veut dire par qualité des données?

- Est-ce que les données sont **représentatives** de la population cible?
- Est-ce que les données permettent de tirer des conclusions de **causalité**?
- Est-ce que les données sont **fiables**?

Principales sources de données

- Laboratoire
- Médecine
- Marketing

Données internes

- Base de données sur la production
- Base de données des clients
- Listes de transactions
- Information sur les visites web
- Listes de clients potentiels
- Base de données des employés
- Données de sondage

- Statistique Canada
 - Recensement
 - Librairie CANSIM2R
- Institut de la statistique du Québec
- Ressources naturelles Canada (Geogratis)

Autres sources officielles

- Ministères et organismes
- Gouvernement ouvert Canada Québec

Autre données disponibles

- Données sur les réseaux sociaux
- Données diverses en ligne
- Données payantes (Poste Canada, Environix)

- Représentatives?
- Permettent de tirer des conclusions de causalité?
- Fiables?