

Données : modèle, qualité et nettoyage (PROJ702_MATH)



En bref

- > **Méthodes d'enseignement:** En présence
- > **Ouvert aux étudiants en échange:** Oui

Présentation

Description

Avant de commencer une analyse de données, il faut bien comprendre les données que l'on va explorer, savoir les nettoyer et en évaluer la qualité. Ce travail fastidieux mais indispensable peut prendre jusqu'à 80% du temps dédié à un projet en science des données.

Objectifs

Le but de ce cours est de préciser les différents modèles de données qui permettent de faire les bons choix pour nettoyer les données et pour remplacer (ou pas) les données manquantes en se référant à la distribution théorique des données. Nous allons également évaluer la qualité des données pour pouvoir avoir en sortie des analyses ciblées et des méthodes prédictives pertinentes pour "faire parler les données".

Heures d'enseignement

CM	Cours Magistral	1,5h
TD	Travaux Dirigés	3h
TP	Travaux Pratiques	8h

Plan du cours

Le cours sera orienté projet avec une partie théorique et une grande partie de manipulation des données en TD et TP sur des données réelles.

Infos pratiques

Contacts

Responsable du cours

Laurent Vuillon

☎ +33 4 79 75 87 33

✉ Laurent.Vuillon@univ-savoie.fr

Lieux

› Le Bourget-du-Lac (73)

Campus

› Le Bourget-du-Lac / campus Savoie Technolac