

# Statistiques pour Big Data

## Infos pratiques

---

- > ECTS : 3.0
- > Nombre d'heures : 24.0
- > Langue(s) d'enseignement : Français
- > Niveau d'étude : BAC +4
- > Période de l'année : Enseignement septième semestre
- > Méthodes d'enseignement : En présence
- > Forme d'enseignement : Cours magistral
- > Ouvert aux étudiants en échange : Oui
- > Composante : Sciences économiques, gestion, mathématiques et informatique
- > Code ELP : 4E7GASBI

## Présentation

---

Ce cours porte sur l'apprentissage supervisé: construire des modèles à partir de des données qui prédisent un résultat à l'aide d'une collection d'input. Le cours aura pour plan :

- 1) Introduction à l'apprentissage statistique et fondements théoriques
- 2) Présentation des régressions pénalisées pour Big Data
- 3) Introduction aux réseaux de neurones
- 3) Lien entre l'économétrie non paramétrique et l'apprentissage machine
- 4) Présentation d'algorithmes d'apprentissage

## Objectifs

---

Savoir choisir le modèle approprié pour réaliser des prévisions, savoir exploiter les données de façon intelligente. Savoir utiliser les langages statistiques (Python, R) dans le but d'une tâche précise

## Évaluation

---

Évaluation écrite + Dossier

## Pré-requis nécessaires

---

Économétrie Avancée, Algèbre linéaire, base de programmation

## Compétences visées

---

Choisir le modèle approprié pour réaliser des prévisions, être capable de collecter et d'exploiter les données de façon intelligente, Utilisation des langages statistiques pour faire de l'apprentissage machine

## Bibliographie

---

The Elements of Statistical Learning, T. Hastie, R. Tibshirani, J. Friedman (Springer 2009).

Économétrie & Machine Learning, A.Charpentier, E. Flachaire et A. Ly (2018),

## Ressources pédagogiques

---

Avoir un ordinateur est un plus pour appliquer les exemples pendant le cours, utilisation de Moodle pour présenter les slides et mise à disposition d'un notebook jupyter