

# Introduction à la Data science

Semestre 3

Code Apogée : LE204MXC

3 ECTS

16h de CM et 12h de TD

**Master Santé Publique**

**Parcours M2 B3H**

Biostatistics, Bioinformatics, Biomathematics for Health

**Responsables de l'UE :**

**Alexandre SAÏDI**

**DOMAINE SCIENCES TECHNOLOGIE SANTE**

**Composante de gestion de l'UE :**

**Médecine Lyon Est**

Service des Etudes Interdisciplinaires en Santé

## PROGRAMME DE L'ENSEIGNEMENT

### Objectifs de la formation

Donner aux étudiants un aperçu et la pratique des bases de cette discipline. Un tour non exhaustif des méthodes les plus utilisées sur des données transactionnelles (ce sont en général des données non-image car les données images nécessitent d'autres traitements).

**Mots-clés :** Fouille de données, Extraction de connaissances à partir de données, Data Science, Machine learning.

### Programme

- La problématique scientifique de l'apprentissage artificiel (Machine Learning)
- Formalisation et fondements Algorithmiques / Statistiques
- Quelques applications typiques
- Apprentissage supervisé, Non-(et semi-) supervisé,
- Classification, régression, association (Introduction à la fouille de textes) - Arbres de décision / de régression
- Méthodes statistiques (Bayes, SVM, Méthodes à base de noyaux)
- Méthodes et indicateurs d'évaluation statistiques
- Méthodes de Classification (Clustering)

### Compétences

- Initiation & pratique de Data Science : en présence des données, l'élève devrait être capable de choisir les méthodes applicables. Il devrait ensuite pouvoir évaluer les résultats et choisir la meilleure méthode.
- Rappel des éléments statistiques (de base).
- Apporter un outil d'aide à la décision.

### Modalités d'évaluations de l'unité d'enseignement :

Session 1 : Ecrit 2h 67% examen terminal + Rapport 33% contrôle continu

Session 2 : pas de session2

(Modalités données à titre indicatif, les modalités précises sont votées tous les ans par le CA sur proposition du CFVU)

**Vos enseignants dans cette UE : A. SAÏDI**