

## Objectifs des cours du Master 2 Mention économétrie, statistiques en FOAD

### ○ Logiciels statistiques :

- **R** : Apprendre à utiliser les outils de gestion de données ainsi que les outils statistiques « de base » avec le logiciel R.
- **Python** : Apprendre les techniques de base du logiciel Python.
- **SAS** : Donner les principes de base de programmation SAS et de savoir rechercher seul les différentes syntaxes afin d'être autonome en entreprise.

### ○ Base de données :

Les principes des bases de données sont abordés selon deux aspects : la modélisation et l'exploitation. Deux modèles sont étudiés : le modèle Entité-Association et le modèle relationnel. Pour l'exploitation, on étudie les principes du langage SQL et on les applique à la version proposée par Oracle™.

### ○ Econométrie :

Expliquer une variable qualitative ou une variable tronquée.

### ○ Data-mining :

Familiariser les étudiants aux enjeux et aux principales techniques de modélisation utilisée en data-mining. Il a pour but de décrire les différentes démarches d'un processus de traitement de données ainsi que de présenter les principes de base de l'apprentissage automatique essentiellement dans le cas de l'apprentissage supervisé et non supervisé.

### ○ Techniques de sondage :

Apporter les bases de l'inférence sous un plan de sondage. Différents plans de sondage et méthodes d'estimation sont étudiés en détail et mis en œuvre avec les logiciels R et SAS (automatique essentiellement dans le cas de l'apprentissage supervisé et non supervisé).

- **Modèles linéaires gaussiens :**

Etudier différentes formes du modèle linéaire gaussien qui a pour but de prédire ou d'expliquer une variable réponse mesurée sur une échelle continue en fonction de variables explicatives qui peuvent être quantitatives ou qualitatives. savoir-faire par l'application de ces techniques avec le logiciel R.

- **Modèles non paramétriques :**

Donner une culture générale dans le domaine de la statistique non paramétrique et à acquérir un savoir-faire par l'application de ces techniques avec le logiciel R.

- **Séries temporelles :**

Modéliser des séries temporelles, macroéconomiques, ou financières et faire des prévisions.

- **Scoring :**

Le scoring est une méthode statistique permettant de prédire une variable catégorielle (le plus souvent binaire) en utilisant des variables quantitatives ou qualitatives. Le scoring est très utilisé en marketing et dans le domaine bancaire.

Il existe différentes techniques pour construire un score : régression logistique, modèles probit, analyse discriminante, modèles additifs généralisés, la régression logistique restant la technique principale utilisée dans les entreprises. Le scoring y est utilisé comme un outil pour l'évaluation du risque et pour la réduction des coûts.

- **Big Data :**

Comprendre les défis du traitement massif de données et avoir des bases sur la conception et le déploiement d'applications BigData notamment avec les applications du framework Hadoop.

- **Econométrie spatiale :**

Etre capable de réaliser (via le logiciel R) et d'interpréter les résultats d'une analyse mobilisant les outils de l'économétrie spatiale.

- **Web mining :**

L'objectif de ce module "web mining" est double : il s'agira dans un premier temps de dresser un panorama assez large de ce qu'est le "web mining" et des techniques mises en œuvre pour fouiller ces données. Dans un second temps, les étudiants se confronteront eux-mêmes à cette problématique d'extraction de connaissances à travers la mise en place d'un TP/projet.

- **Géostatistique :**

Appréhender la spatialisation d'une variable quantitative dans un modèle probabiliste au second ordre, mener des calculs dans ce modèle.

Analyser la corrélation d'une variable spatialisée, estimer un modèle géostatistique à partir d'un échantillon d'une réalisation.

Prédire avec un modèle géostatistique.

Appréhender la corrélation spatiale d'objets ponctuels spatialisés par un modèle probabiliste.

Tester l'hypothèse de répartition aléatoire vis à vis de la régularité ou de l'agrégation

Estimer les caractéristiques (intensité, fonctions d'ordre 2, ...) d'un processus ponctuel à partir de son observation dans une fenêtre.

- **Analyse des durées de vie :**

Savoir analyser des données issues d'un essai thérapeutique avec les tests statistiques classiques adéquats et calculer le nombre de sujets nécessaires.

Savoir reconnaître une situation où apparaissent des données censurées, pouvoir analyser de tels jeux de données en mettant en œuvre une modélisation appropriée.

- **Mémoire, Rapport de stage :**

Le stage a pour finalité de favoriser l'insertion professionnelle en permettant d'acquérir et de valoriser des compétences.

Le mémoire et le rapport de stage ont pour objectif de mettre en pratique les techniques acquises pendant la formation.