

Statistique multidimensionnelle

Intitulé du cours	Statistique multidimensionnelle
Niveau /semestre	A1/P4
Composante	École d'Économie de Toulouse
Enseignant responsable	Sandrine CASANOVA
Volume Horaire CM	15
Volume Horaire TD	7,5
Langue du cours	Français

Coordonnées de l'équipe pédagogique :

Sandrine Casanova : sandrine.casanova@tse-fr.eu, bureau T 215

Les étudiants peuvent poser des questions à la fin du cours, par mail ou prendre rendez-vous par mail pour rencontrer l'enseignante.

Objectifs du cours :

A l'issue de ce cours, l'étudiant connaîtra cinq méthodes d'analyse des données qui permettent de résumer l'information contenue dans de grands tableaux de données et saura appliquer ces méthodes sur des données réelles à l'aide du logiciel R. Plus largement, l'étudiant aura acquis un savoir-faire dans l'exploration de données à l'aide de méthodes statistiques multidimensionnelles descriptives. Un certain nombre de résultats seront démontrés.

Prérequis :

Statistique descriptive.

Modalités pratiques de gestion du cours :

Les ordinateurs et tablettes sont acceptés en cours.

Une participation active en cours est attendue de la part des étudiants et le logiciel R devra être installé sur vos ordinateurs personnels dès les premiers cours.

Des exercices seront donnés sur Moodle et seront à faire en autonomie en dehors des cours. Des corrigés seront ensuite mis à disposition sur Moodle.

L'acceptation des arrivées tardives est laissée à l'appréciation de l'enseignante.

Modalités d'évaluation :

L'évaluation de mi-parcours compte pour 40% et l'évaluation terminale compte pour 60%.

L'évaluation de mi-parcours est sous la forme d'un projet avec le logiciel R.

L'évaluation terminale est un examen écrit sans documents portant sur l'ensemble du cours.

Bibliographie/références :

Pour les prérequis :

« Statistique pour économistes et gestionnaires ». Brigitte Tribout, Pearson Education.

Pour l'apprentissage du logiciel R :

R peut être téléchargé **gratuitement** sur le site du CRAN : <https://cran.r-project.org/>

Plusieurs manuels en anglais y sont accessibles.

Un manuel en ligne en français :

https://cran.r-project.org/doc/contrib/Paradis-rdebuts_fr.pdf

« Analyse de données avec R », François Husson, Sébastien Lê et Jérôme Pagès, PUR.

« Statistiques avec R », Pierre-André Cornillon et autres, PUR.

Pour le cours :

- « Statistique : méthodes pour décrire, expliquer et prévoir », Michel Tenenhaus. Dunod.
« Probabilités, analyse des données et statistique », Gilbert Saporta. Technip.

Planification des séances :**Plan du cours :**

- Introduction à la statistique multidimensionnelle,
- Analyse en composantes principales (ACP),
- Analyse factorielle des correspondances simple et multiple (AFC et AFCM),
- Classification non supervisée : agrégation autour des moyennes mobiles (méthode k-means) et classification ascendante hiérarchique (CAH).

Enseignement à distance :

Si la situation sanitaire l'exige, les enseignements pourront avoir lieu à distance sous forme de classe en ligne interactive, QCM et exercices en ligne, TP à distance, forum sous Moodle.