

Analyse des données (spécifique Magistère)

Intitulé du cours	Analyse des données
Niveau /semestre	L3 / S1
Composante	École d'Économie de Toulouse
Enseignant responsable	Sandrine CASANOVA
Volume Horaire CM	12 h
Volume horaire TD	
Volume horaire TP	12 h
Langue du cours	Français
Langue des TD et/ou TP	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 cette partie, l'étudiant connaîtra trois 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 RStudio. Plus largement, l'étudiant aura acquis un savoir-faire dans l'exploration de données à l'aide de méthodes statistiques multidimensionnelles.

Prérequis :

Statistique descriptive, calcul matriciel.

Modalités pratiques de gestion du cours :

Les ordinateurs et tablettes sont acceptés en cours et en TP.

Il est attendu des étudiants une participation active en cours et en TP.

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

Modalités d'évaluation :

L'évaluation de ce cours est en contrôle continu.

Le contrôle continu se compose de trois projets à réaliser en binôme avec le logiciel R et d'un examen individuel de vingt minutes sous la forme d'un QCM en ligne.

Bibliographie/références :

Pour les prérequis :

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

Pour le téléchargement (gratuit) du logiciel RStudio :

<http://www.rstudio.com/products/rstudio/download/>

Il est recommandé d'installer préalablement le logiciel R qui 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

Pour le cours

« Statistique : méthodes pour décrire, expliquer et prévoir », Michel Tenenhaus. Dunod

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

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

Plan du cours :

- Introduction à l'analyse des données
- Analyse en composantes principales (ACP)
- Analyse factorielle des correspondances (AFC)
- Classification non supervisée

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.