



Nom du Syllabus :	Data mining (analyse des données)
Niveau :	DU
Enseignante :	Sandrine Casanova

OBJECTIFS :

A l'issue du cours, l'étudiant connaîtra les méthodes de base d'analyse des données et saura appliquer ces méthodes sur de vraies données à l'aide des logiciels R et SAS. Plus largement, l'étudiant aura acquis un savoir-faire dans l'exploration de données à l'aide de méthodes statistiques bidimensionnelles et multidimensionnelles.

PLAN DU COURS :

Ce cours est divisé en 5 cours et travaux pratiques (logiciels R et SAS)

1. Statistique descriptive
2. Analyse en composantes principales (ACP)
3. Analyse factorielle des correspondances (AFC et AFCM)
4. Classification : agrégation par les moyennes mobiles, classification ascendante hiérarchique
5. Analyse factorielle discriminante (AFD)

PRE-REQUIS :

Statistique descriptive : moyenne, écart-type, corrélations, tableaux croisés

BIBLIOGRAPHIE :

« Statistique : méthodes pour décrire, expliquer et prévoir »

Auteur : Michel Tenenhaus.

Editeur : Dunod



« L'analyse de données avec R" »

Auteurs : François Husson, Sébastien Lê et Jérôme Pagès.

Editeur : Presses Universitaires de Rennes

« Probabilités, Analyse des données et Statistique »

Auteur : G. Saporta

Editeur : Technip

« Approche pragmatique de la classification »

Auteur : J.P. Nakache, J. Confais

Editeur : Technip