

# Master Mathématiques appliquées, statistique

## Projet intégré

### Informations

Composante : Faculté des Sciences

### Langue(s) d'enseignement

Français

### Contenu

Renfort « statistique exploratoire pour les données catégorielles » : (resp. Nicolas PECH)

- prétraitement des données de type catégoriel
- données manquantes et méthodes d'imputations usuelles,
- analyses uni et bivariées,
- analyse des correspondances simple et multiple

Renfort « Ingénierie des données pour l'apprentissage » : (resp. Pierre PUDLO)

- discrétisation de variables, transformation de variables (logarithmique, quadratique)
- transformation min-max, Box-Cox, Yeo-Johnson,...
- introduction aux données textuelles
- introduction aux splines
- introduction à l'imputation de données manquantes pour l'apprentissage

### Compétences à acquérir

- Mobiliser les savoir et savoir-faire de différentes UE pour mener une petite étude
- Fournir des livrables (rapport, code, données)
- Restituer des résultats obtenus et dégager des conclusions d'une étude
- Mobiliser les méthodes vues dans le renfort

### Modalités d'organisation

Projet et compléments de formation en renfort par parcours

### Bibliographie, lectures recommandées

Saporta (2011) Probabilités, Analyse des données et Statistique. Technip

J-J Dreesbeke, M. Lejeune, G. Saporta (2005) Modèles statistiques pour données qualitatives. Technip

Bressoux (2008) Modélisation statistique appliquée aux sciences sociales. DeBroeck

Waal, Ton de et al. (2012) Handbook of Statistical Data Editing and Imputation. Wiley publication

Zheng, A., & Casari, A. (2018). Feature engineering for machine learning: principles and techniques for data scientists. O'Reilly Media.

### VOLUME HORAIRE

- Volume total: 20 heures
- Travaux dirigés: 20 heures

### Codes Apogée

- SMSAU25C [ELP]

### Pour plus d'informations

[Aller sur le site de l'offre de formation...](#)



Dernière modification le 07/10/2024