

UM4EU001 ANALYSES DE DONNEES		
3 ECTS	Mots clefs	statistiques appliquées, statistiques descriptives, statistiques paramétriques, $\chi^2$ , modèle linéaire général, modèle linéaire généralisé, analyse factorielle
M1	Responsables	Matthieu DACHER (matthieu.dacher@sorbonne-universite.fr) et Dominique LAMY (dominique.lamy@sorbonne-universite.fr)
Paris	Intervenants	Céline ELLIEN, Éric GOBERVILLE (BOREA, Paris) + 1 ATER ou 1 moniteur·rice

## Descriptif

### Format de l'UE

#### Modalités d'enseignement

Les enseignements sont dispensés en présentiel et comprennent 10 heures de cours magistraux en 4 séances, 15 heures de travaux pratiques en 5 séances réalisés en salle informatique (2 demi-groupes) avec les logiciels R et RStudio. Un projet tutoré est mené tout au long du semestre, avec un suivi collectif (6 heures sur 2 séances). Tous les supports de cours et de TP sont mis à disposition sur la plateforme Moodle de l'UE.

#### Modalités d'évaluation

L'évaluation repose sur un projet en groupe (30 %), consistant en la réalisation et la présentation d'une analyse statistique complète, ainsi que sur deux examens écrits individuels : l'un portant sur les tests statistiques de base (30 %), l'autre sur les modèles linéaires et les méthodes multivariées (40 %). En cas de session de rattrapage, seule la note de projet est conservée.

## Résumé

Cette unité d'enseignement vise à initier les étudiant·es à l'analyse statistique appliquée et à leur donner les compétences nécessaires pour mobiliser les outils statistiques dans leurs travaux académiques et professionnels. Les étudiant·es apprennent à identifier et appliquer les principaux tests statistiques (tests de Student,  $\chi^2$ ) à partir d'exemples concrets, puis à mettre en œuvre et interpréter le modèle linéaire général (ANOVA, régression linéaire, ANCOVA) ainsi que le modèle linéaire généralisé (régression logistique, modèles log-linéaires). Une attention particulière est portée à l'analyse factorielle (ACP) comme méthode exploratoire des données.

L'enseignement repose sur une approche progressive, combinant cours, exercices guidés et projet tutoré, et met l'accent sur l'utilisation pratique du logiciel R comme outil d'analyse et de visualisation. Cette UE permet ainsi aux étudiant·es de développer une autonomie dans le traitement de données réelles, de communiquer efficacement les résultats obtenus et de les replacer dans le cadre de problématiques scientifiques ou appliquées.

## Objectifs d'apprentissage

Au terme de cette UE, l'étudiant·e sera capable de :

- Identifier et sélectionner les tests statistiques adaptés (tests de Student,  $\chi^2$ ) en fonction des données et des hypothèses.
- Mettre en œuvre et interpréter les modèles linéaires (ANOVA, régression linéaire, ANCOVA) et généralisés (régression logistique, modèles log-linéaires).
- Appliquer et analyser une méthode multivariée exploratoire (ACP) pour caractériser des jeux de données complexes.
- Utiliser efficacement le logiciel R pour réaliser, documenter et présenter une analyse statistique complète.

## Prérequis

L'unité d'enseignement est ouverte à l'ensemble des étudiant.es du portail SDUEE. Les étudiant.es doivent être capables de :

- Mobiliser des notions élémentaires de statistiques descriptives (moyennes, variances, représentations graphiques).
- Lire et interpréter des résultats simples issus de tests statistiques de base.

Les étudiants n'ayant pas acquis ces compétences sont invités à réviser les documents de rappels disponibles sur Moodle. Les premières séances incluent un rappel introductif de ces notions.

NB : Ce document est indicatif. Les détails du contenu et de la forme des enseignements et des évaluations peuvent évoluer d'une année à l'autre.