



4BM019 (3 ECTS)

Biostatistique et Bioanalyse sous R

Initiez-vous à l'analyse de données biologiques sous R à travers des cours et des projets

1^{er} semestre 2024

Lundi 9 septembre de 9h à 12h (CM)
Mardi 10 septembre de 9h à 12h (CM)
Mercredi 11 septembre de 9h à 12h (CM)
Jeudi 12 septembre de 9h à 12h (CM)

Vendredi 13 septembre de 13h45 à 18h (TD)

Lundi 28 octobre de 9h à 18h (TP, projets)
Mardi 29 octobre de 9h à 18h (TP, projets)

Responsable pédagogique
Nicolas Tchitchek
Maître de Conférences en Immunologie des Systèmes
nicolas.tchitchek@sorbonne-universite.fr



Objectifs pédagogiques

Poussée par les révolutions biotechnologiques successives, l'analyse de données est aujourd'hui une compétence essentielle dans l'ensemble des métiers de la biologie moderne.

Cette unité d'enseignement a pour objectif de former les étudiants de première année du master BMC à l'analyse des données biologiques en utilisant la programmation en R, afin qu'ils soient capables d'interpréter des données de faible ou de grande dimension.

Les étudiants acquerront des compétences en biostatistiques (statistique descriptive et inférentielle) et en bioanalyse (analyses univariées, multivariées et intelligence artificielle), essentielles dans différents contextes des sciences du vivant. Ces compétences leur permettront de concevoir des stratégies d'analyse pour leurs projets académiques et leurs futurs métiers, d'être autonomes dans leur démarche scientifique, et de communiquer et discuter les résultats obtenus.

Programme pédagogique

Cette unité d'enseignement s'organise en quatre demi-journées de cours magistraux, une demi-journée de travaux dirigés, et deux jours de travaux pratiques sous forme de projets à réaliser en groupes d'étudiants.

Le programme pédagogique détaillé inclut :

- L'environnement de travail RStudio
- Les fondamentaux de la programmation sous R
- L'utilisation de R pour l'analyse de données
- Les notions de base en statistique
- La statistique inférentielle et les tests d'hypothèses
- Les comparaisons d'échantillons biologiques
- Les associations entre variables biologiques
- Les analyses multivariées et la réduction de dimensionnalité
- Les représentations graphiques avancées
- Une introduction à l'intelligence artificielle
- Les stratégies d'analyse et les méthodologies de travail
- La communication des résultats scientifiques

Organisation

Cette unité d'enseignement de 3 ECTS se déroulera au premier semestre de l'année 2024/2025.

Les modalités de contrôle des connaissances seront les suivantes :

- Un contrôle des connaissances sous forme de QCM (CC1, 33 % de la note finale), qui aura lieu le lundi 14 octobre 2024, de 18h30 à 19h30.
- Un projet d'analyse de données biologiques (CC2, 33 % de la note finale), à rendre pour le lundi 11 novembre 2024, avant 23h59.
- Un contrôle des connaissances sous forme de partiel (CC3, 34 % de la note finale) qui aura lieu le mercredi 18 décembre 2024, de 10h45 à 12h45.

Le projet consistera en l'analyse de trois jeux de données parmi un large choix et sera réalisé en équipes de trois étudiants. L'évaluation comprendra la restitution d'un rapport d'analyse ainsi que les scripts d'analyse. Les étudiants pourront également proposer leurs propres jeux de données à analyser et interpréter, sous réserve de validation par l'enseignant.