



4BM018 (3 ECTS)

Python pour le Traitement de Données Biologiques

1^{er} semestre 2025

10h de vidéos interactives en autonomie
4 séances de TP de 4h en présentiel

Initiez-vous à la programmation en Python à travers des enseignements interactifs et des travaux pratiques.



Responsable pédagogique

Nicolas Tchitchek

Maître de Conférences

nicolas.tchitchek@sorbonne-universite.fr

Objectifs pédagogiques

L'exploitation des données biologiques complexes en sciences du vivant nécessite des compétences en programmation informatique appliquée.

Cette unité d'enseignement vise à former les étudiants de 1^{ère} année de la mention de master BMC à la programmation en langage Python pour le traitement automatique de données biologiques. L'objectif est de leur permettre de travailler sur des projets variés et de s'adapter aux besoins des laboratoires de recherche et de l'industrie, tout en accompagnant les avancées en analyse des données biologiques.

Les étudiants développeront leurs connaissances et compétences en programmation, allant de la manipulation de données biologiques à grande échelle à visualisation des résultats d'analyses. Les applications aborderont des domaines clés tels que l'analyse de données génomiques et transcriptomiques, la chimioinformatique, le repliement des protéines, l'imagerie biologique et l'interprétation de données complexes générées par des technologies de pointe.

Programme pédagogique

Les enseignements se dérouleront sous forme de séances vidéo interactives à travailler en autonomie, complétées par des travaux pratiques permettant aux étudiants d'appliquer les concepts abordés.

Le programme pédagogique inclut les éléments suivants :

- Présentation de l'environnement de développement (Spyder et VSCode).
- Introduction aux bases de Python (syntaxe, variables, types de données).
- Structures de contrôle (conditions, boucles, fonctions).
- Manipulation de fichiers (lecture/écriture de fichiers CSV ou texte).
- Manipulation de données tabulaires (listes, dictionnaires, DataFrames avec Pandas).
- Nettoyage et filtrage de données (Pandas).
- Calculs scientifiques et statistiques de base (NumPy).
- Visualisation de données (graphiques simples avec Matplotlib et complexes avec Seaborn).
- Organisation et documentation du code (scripts clairs, commentaires, bonnes pratiques).
- Principes de reproductibilité (gestion des environnements, sauvegarde du code).
- Mise en pratique sur des jeux de données biologiques variés.
- Introduction aux ressources et communautés Python (documentation, notebooks partagés).

Organisation

Cette unité d'enseignement de 3 ECTS se déroulera les mercredi après-midi du 1^{er} semestre de l'année.

Les modalités de contrôle des connaissances sont les suivantes :

- Un contrôle des connaissances sous forme de QCM (CC1, 30% de la note finale).
- L'évaluation des scripts rédigés au cours des travaux pratiques (CC2, 10 % de la note finale).
- Un contrôle des compétences sous forme de problèmes à résoudre (CC3, 60% de la note finale).

Les séances de travaux pratiques sont organisées en 4 sessions de 4 heures, au cours desquelles les étudiants pourront appliquer les notions et compétences abordées dans les vidéos et démonstrations suivies en autonomie. Des forums et des séances de questions/réponses seront également organisés régulièrement avec les enseignants pour accompagner les étudiants tout au long de l'UE.

Le CC1 se tiendra à l'issue de la deuxième séance de TP (de 18h à 19h) et le CC3 sera réalisé pendant la dernière séance de TP (de 16h à 18h).