






4.00 crédits	15.0 h + 5.0 h	Q2
--------------	----------------	----

Enseignants	Symul Laura ;
Langue d'enseignement	Français
Lieu du cours	Louvain-la-Neuve
Préalables	Concepts et outils équivalents à ceux enseignés dans les UEs LSTAT2020 Logiciels et programmation statistique de base LSTAT2110 Analyse des données
Thèmes abordés	<ul style="list-style-type: none"> • La normalisation de données omics (que ce soit génomiques ou métabolomique) • Les méthodes mathématiques et statistiques pour le prétraitement de données spectrales (ex : modèles semi-paramétrique de lissage pour correction de ligne de base, alignement de pics) • L'organisation d'expériences pour analyser la qualité informatique de données omics et leur analyse par modèles à composantes de variance, méthodes de classification et méthodes multivariées telles • ASCA, ANOVA-PCA • La modélisation de données de grande dimension dans un but de recherche de biomarqueurs ou de prédiction par modèle PLS, O-PLS, ICA, arbres de décision • Les méthodes pour tests multiples (FDR) • Les méthodes d'intégration de données (analyse de données multitableaux) • Revue et utilisation des packages R les plus courants dans le domaine (ex : bioconductor) • Application sur des bases de données réelles.
Acquis d'apprentissage	<p>A la fin de cette unité d'enseignement, l'étudiant est capable de :</p> <p>Eu égard au référentiel AA du programme « Le master en statistique, orientation biostatistique », cette activité permet aux étudiants de maîtriser</p> <ul style="list-style-type: none"> • De manière prioritaire les AA suivants: 1.4, 2.2, 2.5, 2.6, 3.1, 3.2, 3.3, 3.4, 3.5, 4.5, 5.1, 5.2, 5.6 • De manière secondaire les AA suivants: 4.1, 5.7 <p>1</p> <p>Eu égard au référentiel AA du programme de « Le master en statistique, orientation générale », cette activité permet aux étudiants de maîtriser</p> <ul style="list-style-type: none"> • De manière prioritaire les AA suivants : 1.3, 1.4, 2.2, 2.4, 2.5, 2.6, 3.1, 3.2, 3.3, 3.4, 5.1, 5.5, 5.6 • De manière secondaire les AA suivants : 1.2, 1.5, 4.1
Modes d'évaluation des acquis des étudiants	<p>Dans le cadre de ce cours, les étudiant-es sont évalué-es de deux manières :</p> <ul style="list-style-type: none"> • l'évaluation continue certificative incluant: <ul style="list-style-type: none"> • des travaux obligatoires à remettre en cours de quadrimestre selon un calendrier fixé en début de quadrimestre (40% de la note finale) • et un projet final à présenter oralement lors du dernier cours (40% de la note finale) • un examen oral à livre ouvert (20% de la note finale) <p>Une note globale de 10/20 ou plus ne sera acquise que si l'étudiant a obtenu pour chaque travail ou partie de l'évaluation une note individuelle de minimum 8/20.</p>
Méthodes d'enseignement	<p>Le cours est composé d'une série d'activité qui amènent l'étudiant à se plonger activement dans le monde des données -omiques. Il propose:</p> <ul style="list-style-type: none"> • des exposées par des spécialistes actifs dans le domaine, • des mini-projets de traitement de données à réaliser chaque semaine, • un travail interactif sur ordinateur durant le cours, • une visite de laboratoire, • un projet final sur des données proposées par les différents intervenants du cours ou de data repositories. <p>Les modalités prévues évolueront en fonction de la situation sanitaire.</p>
Contenu	<p>Après avoir revu les bases de la biologie moléculaire, le cours présente une série de méthodes -omiques et surtout les méthodes de traitement de données liées:</p> <ul style="list-style-type: none"> • Bases de biologie moléculaire.

	<ul style="list-style-type: none"> • Révision de méthodes multivariées utiles en méthodes -omiques (PCA, Clustering...) et application en R + RMarkdown. • Méthode d'acquisition de données de transcriptomique (micro-damiers, q-PCR, RNA-Seq...). • Prétraitement et analyse de données transcriptomiques (correction de background, normalisation,... + tests d'hypothèses avec correction de multiplicité). • Utilisation de modèle de prédiction et classification émanant de la chimiontrie et du machine learning pour l'analyse de données omique (PLS, O-PLS, arbres...). • Acquisition et traitement de données protéomiques. • Acquisition et traitement de données métabolomiques (dont prétraitement détaillé de données 1H-NMR). • Traitement de données métagénomiques.
Ressources en ligne	Site Moodle: https://moodleucl.uclouvain.be/course/view.php?id=10846
Faculté ou entité en charge:	LSBA

Programmes / formations proposant cette unité d'enseignement (UE)				
Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Master [120] en science des données, orientation statistique	DATS2M	4		
Master [120] en statistique, orientation biostatistiques	BSTA2M	4		
Master [120] en statistique, orientation générale	STAT2M	4		
Master [120] : bioingénieur en chimie et bioindustries	BIRC2M	5		
Certificat d'université : Statistique et science des données (15/30 crédits)	STAT2FC	4		
Master [120] : bioingénieur en sciences agronomiques	BIRA2M	4		