Sigle : INF1833 Gr. 01
Titre : Analyse statistique et visualisation de données
Session : Automne 2025 Horaire et local
Professeur : Yapi, N'Dah Daniel

1. Description du cours paraissant à l'annuaire :

Objectifs

Au terme de ce cours, l'étudiant.e comprendra les principales méthodes statistiques utilisées en analyse des données. Il/elle sera en mesure d'appliquer ces méthodes pour résoudre des problèmes pratiques à l'aide d'un langage de programmation (ex. Python ou R).

Contenu

Types de données et prétraitement : échantillonnage, nettoyage. Statistique descriptive pour la visualisation et l'interprétation des données. Statistique inférentielle: sélection d'un modèle statistique, estimation de paramètres d'un modèle à partir de données, prédiction. Maximum de vraisemblance. Modèles linéaires pour la régression et la classification. Évaluation d'un modèle statistique (ex. biais, consistance, etc.). Analyse de variance. Processus stochastiques. Introduction à la statistique Bayésienne. Ce cours comporte des séances obligatoires de travaux dirigés (TD).

[Descriptif – Annuaire](#)

2. Objectifs spécifiques du cours :

Au terme de cette activité, l'étudiant(e) aura acquis des connaissances sur :

- **Comprendre** les principales méthodes statistiques appliquées à l'analyse des données.
- **Appliquer** ces méthodes pour résoudre des problèmes concrets, en utilisant un langage de programmation tel que Python ou R.
- **Maîtriser** la visualisation de données pour l'exploration, la description et l'interprétation des ensembles de donnée.
- **Utiliser** la statistique inférentielle pour la sélection de modèles, l'estimation de paramètres et la prédiction.
- **Appliquer** les principes du maximum de vraisemblance.
- **Construire et évaluer** des modèles linéaires pour la régression et la classification (incluant l'analyse de variance).
- **S'initier** aux processus stochastiques et à la statistique Bayésienne.

3. Stratégies pédagogiques :

Les formules pédagogiques suivantes seront utilisées :

- Cours magistraux (3 heures par semaine) en mode présentiel.
- Lectures et discussions
- Séances de travaux dirigés avec Python ou R (8 séances de 2 heures chacune)

4. Heures de disponibilité ou modalités pour rendez-vous :

Sur rendez-vous par courriel : yapida01@uqo.ca.

5. Plan détaillé du cours sur 15 semaines :

Semaine	Thèmes	Dates
1	Cours en ligne Zoom Introduction à l'analyse statistique, Python et manipulation de données	02 sept 2025

	<ul style="list-style-type: none"> • Introduction au cours et aux objectifs. • Cycle de vie de l'analyse de données. • Types de données (qualitatives, quantitatives). • Présentation de python 	
2	<p>Statistiques descriptives et visualisation.</p> <ul style="list-style-type: none"> • Collecte des données • Mesures de tendance centrale • Mesures de dispersion • Introduction à la visualisation de données <p>Travaux dirigés 1</p>	09 sept 2025
3	<p>Cours en ligne Zoom</p> <p>Statistiques descriptives et visualisation suite</p> <ul style="list-style-type: none"> • Mesures de forme (asymétrie, aplatissement). • Corrélation et covariance. • Visualisation des relations entre variables : nuages de points, matrices de corrélation. <p>Travaux dirigés 2</p>	16 sept 2025
4	<ul style="list-style-type: none"> • Lectures • Travail pratique sur les meilleurs pratiques de la visualisation de données 	23 sept 2025
5	<p>Cours en ligne Zoom</p> <p>Modélisation inférentielle</p> <ul style="list-style-type: none"> • Ajustement de modèles • Estimation des paramètres - maximum de vraisemblance • Sélection d'un modèle statistique • Qualité des estimateurs : biais, variance, consistance • Prédiction : Distinction entre l'inférence sur les paramètres et la prédiction d'observations futures, incluant la construction d'intervalles de prédiction. • <p>Travaux dirigés 3</p>	30 sept 2025
6	<p>Modèles Linéaires I : Régression Linéaire Simple et Multiple</p> <ul style="list-style-type: none"> • Définition du modèle de régression linéaire simple. • Estimation des coefficients par MCO. <p>Extension à la régression linéaire multiple (interprétation des coefficients, sélection de variables)</p> <p>Travaux dirigés 4</p>	07 oct 2025
7	Semaine d'études	14 oct 2025
8	Examen de mi-session	21 oct 2025
9	<p>Analyse de Variance (ANOVA)</p> <ul style="list-style-type: none"> • Principe de l'ANOVA à un facteur et à deux facteurs. <p>Travaux dirigés 5</p>	28 oct 2025
10	Processus Stochastiques et Séries Temporelles	04 nov 2025

	<ul style="list-style-type: none"> • Introduction aux concepts de processus stochastiques. • Introduction aux séries temporelles (composantes, notions de stationnarité, modèles simples ARIMA/SARIMA). 	
	Travaux dirigés 6	
11	Introduction à la Statistique Bayésienne <ul style="list-style-type: none"> • Comparaison statistique fréquentiste vs. bayésienne. • Théorème de Bayes. Priors, vraisemblance, posteriors. • Réseau bayésien 	11 nov 2025
	Travaux dirigés 7	
12	Visualisation et Intelligence d'affaire <ul style="list-style-type: none"> • Utilisation de la visualisation pour la prise de décision • Création de graphique interactive • Tableau de bord 	18 nov. 2025
	Travail dirigés 8	
13	Introduction à la visualisation spatiale des données <ul style="list-style-type: none"> • Qu'est-ce qu'une donnée spatiale ? Systèmes de coordonnées (latitude/longitude). • Formats de données géospatiales (Shapefile, GeoJSON). • Manipulation et visualisation de données spatiales (filtrage, jointures spatiales). 	25 nov. 2025
14	Présentation des projets de session	02 dec. 2025
15	Examen final	09 dec. 2025

6. Évaluation du cours :

Outils d'évaluation	Pondération
Devoirs et/ou travaux pratiques	15 %
Projet de session	20 %
Examen de mi-session	30 %
Examen final	35 %

- Un projet à réaliser en duo ou individuellement sera assigné aux étudiant(e)s au cours de la session. Il devra être remis à la fin de celle-ci. Le travail sera effectué progressivement, en suivant les concepts abordés tout au long de la session.
- Tout appareil électronique personnel (cellulaires et autres outils de communication, lecteurs MP3, MP4...) est interdit pendant les examens et l'utilisation non autorisée d'un tel dispositif électronique sera considérée comme une fraude selon le terme de la procédure concernant les infractions relatives aux études et sanctions.
- Aucun délai pour les devoirs ne sera négociable (sauf force majeure) et une note de zéro (0) sera attribuée. Le rapport doit être remis via le site Moodle du cours.

7. Politiques départementales et institutionnelles :

- Politique du département d'informatique et d'ingénierie relative à la tenue des examens
- Note sur le plagiat et sur la fraude

- Politique relative à la qualité de l'expression française écrite chez les étudiants et les étudiantes de premier cycle à l'UQO
- Absence aux examens : cadre de gestion, demande de reprise d'examen (formulaire)

Tolérance **ZÉRO** en matière de violence à caractère sexuel.

Le Bureau d'intervention et de prévention en matière de harcèlement (BIPH) a pour mission d'accueillir, soutenir et guider toute personne vivant une situation de harcèlement, de discrimination ou de violence à caractère sexuel. Le BIPH oriente ses actions afin de prévenir les violences à caractère sexuel pour que nous puissions étudier, travailler et s'épanouir dans un milieu sain et sécuritaire.

Vous vivez ou êtes une personne témoin d'une situation de violence à caractère sexuel ? Vous êtes une personne membre de la communauté étudiante ou une personne membre du personnel, autant à Gatineau qu'à Ripon et St-Jérôme, l'équipe du BIPH est là pour vous, sans jugement et en toute confidentialité.

Ensemble, participons à une culture de respect.

Pour de plus amples renseignements consultez [UQO.ca/biph](https://uqo.ca/biph) ou écrivez-nous au Biph@uqo.ca

8. Principales références :

1. **Les documents du cours magistral**
2. Gandhi, P., Pruthi, J. (2020). Data Visualization Techniques: Traditional Data to Big Data. Dans : Anuncia, S., Gohel, H., Vairamuthu, S. (eds) Data Visualization (pp. 52–74). Springer
3. Petchey, O. L., Beckerman, A. P., Cooper, N., & Childs, D. Z. (2021). Insights from Data with R: An
4. Introduction for the Life and Environmental Sciences. Oxford University Press.
5. Nordmann, E., McAleer, P., Toivo, W., Paterson, H., & DeBruine, L. M. (2022). Data visualization
6. using R for researchers who do not use R. Advances in Methods and Practices in Psychological
7. Science, 5(2).
8. Silge, K., & Robinson, D. (2017). Text Mining with R: A Tidy Approach (2nd ed.). O'Reilly.
9. Welbers, K., Van Atteveldt, W., & Benoit, K. (2017). Text analysis in R. Communication Methods and
10. Measures, 11(4), 245–265
11. Wickham, H. (2019). Advanced R (2nd ed). Chapman and Hall/CRC.

9. Page Web du cours :

<https://moodle.uqo.ca>