

Sigle : INF1823 Gr. 20
Titre : Introduction à la science des données
Session : Automne 2024 Horaire et local
Professeur : St-Onge, Etienne

1. Description du cours paraissant à l'annuaire :

Objectifs

Au terme de ce cours, l'étudiant.e comprendra l'utilité et la portée de l'utilisation des données dans la prise de décision dans différents domaines d'applications. Il/elle comprendra les principes méthodologiques du traitement et de l'analyse des données.

Contenu

Fondements de la science des données : objectifs, problématiques, enjeux et défis. Techniques de collecte et de prétraitement des données. Structure et visualisation des données. Analyse exploratoire des données: extraction des règles d'association, classification et prédiction, regroupement. Métadonnées. Interprétation, validation et communication des résultats d'une analyse de données. Étude d'applications de la science des données (ex. : cybersécurité, santé, recherche d'information, intelligence d'affaires, réseaux sociaux, etc.).

Descriptif – Annuaire

2. Objectifs spécifiques du cours :

Les objectifs spécifiques visés dans ce cours sont :

- Acquérir des bases théoriques et pratiques requises pour le traitement de données numérique.
- Se familiariser à quelques applications courantes, telles que *Pandas*, *Seaborn*, *Scikit-learn*.
- S'initier par la pratique à des techniques de prétraitement, de manipulation et d'analyse de données.

3. Stratégies pédagogiques :

- Le cours sera donné en présentiel, par le biais de cours magistraux.
- Travaux pratiques en Python pour appliquer des techniques de traitement de données numérique.
- Python sera aussi accessible à distance via le laboratoire virtuel de l'UQO : <https://uqo.ca/docs/39679>

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

Sur demande. L'étudiant(e) peut envoyer un courriel pour fixer un rendez-vous : etienne.st-onge@uqo.ca

5. Plan détaillé du cours sur 15 semaines :

Semaine	Thèmes	Dates
1	Fondements de la science des données <ul style="list-style-type: none"> - Généralités et objectifs - Problématiques, enjeux et défis <u>Introduction à Python et Numpy</u>	4 sept. 2024
2	Concepts mathématiques <ul style="list-style-type: none"> - Statistiques et distribution des données - Notions de base de l'algèbre linéaire <u>Statistiques et calcul matriciel</u>	11 sept. 2024

3	Représentation et manipulation des données <ul style="list-style-type: none"> - Données numériques, métadonnées - Opérations sur les tables de données (filtrage, agrégation, fusion) Création et manipulation de DataFrames Pandas.	18 sept. 2024
4	Préparation et visualisation des données <ul style="list-style-type: none"> - Techniques de nettoyage des données (prétraitement) - Gestion des valeurs manquantes et traitement des valeurs aberrantes Visualisation des données	25 sept. 2024
5	Introduction à l'apprentissage machine <ul style="list-style-type: none"> - Concepts de base, type d'apprentissage - Modèles de régression Applications et régression avec Scikit-learn	02 oct. 2024
6	Séries temporelles <ul style="list-style-type: none"> - Concepts de corrélation et d'autocorrélation - Modélisation de séries temporelles Analyse de séries temporelles	9 oct. 2024
7	Semaine d'études	16 oct. 2024
8	Examen mi-session	23 oct. 2024
9	Mégadonnées - Big Data <ul style="list-style-type: none"> - Principes et importance du traitement de données à grande échelle - Comparaison des approches traditionnelles et modernes Étude de cas (<i>Big Data</i>)	30 oct. 2024
10	Traitement de données massives <ul style="list-style-type: none"> - Stratégies et optimisation des tâches de traitement - Méthodes de transformation pour l'analyse de ces données Exercices de traitement de données massives	6 nov. 2024
11	Gestion de données <ul style="list-style-type: none"> - Structures de données pour la gestion d'informations - Techniques de recherche et de récupération d'informations Création et optimisation de structures de données	13 nov. 2024
12	Analyse et interprétation des modèles <ul style="list-style-type: none"> - Approches pour interpréter les résultats suivant l'analyse de données Étude d'applications	20 nov. 2024
13	Validation et communication de résultats <ul style="list-style-type: none"> - Validation des modèles des résultats - Meilleures pratiques pour la communication des résultats Création de rapports et de présentations	27 nov. 2024
14	Examen final	4 déc. 2024
15	Présentations orales et remise du rapport du projet de session	11 déc. 2024

6. Évaluation du cours :

- Devoirs : 15 %
- Projet de session : 20 %
- Examen mi-session : 30 %
- Examen final : 35 %

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 :

Références :

- Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. Wes McKinney, 3rd edition, O'Reilly.
- W³School - Data Science

9. Page Web du cours :

<https://moodle.uqo.ca>