

**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](mailto:etienne.st-onge@uqo.ca)

**5. Plan détaillé du cours sur 15 semaines :**

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

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

## 8. Principales références :

### Références :

- Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. Wes McKinney, 3<sup>rd</sup> edition, O'Reilly.
- W<sup>3</sup>School - Data Science

## 9. Page Web du cours :

<https://moodle.uqo.ca>