**Department of Mathematics & Statistics**
**Concordia University**

# MAST 387
Data Science Lab
***Fall 2024***

**Instructor**: Dr. J. Zhang, Office: LB 921.13 (SGW), Phone: (514) 848-2424, Ext. 3328
Email: junxi.zhang@concordia.ca

**Class Schedule:** Tuesdays & Thursdays, 11:45-13:00.
Mid-term break: no class between October 15, 2024, and October 20, 2024.

**Office Hours:** Tuesdays & Thursdays, 13:10-15:00.
Remote office hours are available through zoom by appointment.

**Objectives:** The objective of data science is to extract meaningful insights from data using mathematics, statistics, and computer science. This course is a hands-on introduction to the essential tools needed for this purpose, with particular emphasis on data preparation, visualization, and exploration. A tentative list of topics to be covered includes:

- Python fundamentals, NumPy, Pandas
- Data cleaning, preparation, and wrangling
- Data exploration: summary statistics, visualization, dimension reduction
- Database creation and management
- Structured data: time series, text, images, videos, networks
- Introduction to machine learning
- Extensive illustrations and examples with real-world datasets

The course will include extensive programming in Python, presented via Jupyter notebooks.

**Pre-requisites:** STAT 380; MAST 333 or STAT 360. Previous experience with programming is assumed, but not necessarily in Python.

**Textbooks:** There will *not* be one main textbook for this course. Lectures will be based on different sources, including but not limited to the following textbooks:

*Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter*, by W. McKinney, O'Reilly Media, 3rd Edition, 2022.
https://concordiauniversity.on.worldcat.org/oclc/1341278534

*Data Science from Scratch: First Principles with Python*, by J. Grus, O'Reilly Media, 2nd Edition, 2019.
https://concordiauniversity.on.worldcat.org/oclc/1097183567

*Data Science Using Python and R*, by C.D. Larose and D.T. Larose, John Wiley and Sons, 2019.
https://concordiauniversity.on.worldcat.org/oclc/1089273491

**Assessment:** Students will be evaluated based on homework assignments, a mid-term project, and a final project, according to the following grading scheme:

20% Homework + 30% Mid-term project + 50% Final project.

If the grading scheme for this course includes graded assignments, a reasonable and representative subset of each assignment may be graded. Students will not be told in advance which subset of the assigned problems will be marked and should therefore attempt all assigned problems.

**Plagiarism:** Cases of plagiarism will be treated according to the University policy.

**Academic Integrity and the Academic Code of Conduct**
This course is governed by Concordia University's policies on Academic Integrity and the Academic Code of Conduct as set forth in the Undergraduate Calendar and the Graduate Calendar. Students are expected to familiarize themselves with these policies and conduct themselves accordingly. "Concordia University has several resources available to students to better understand and uphold academic integrity. Concordia's website on academic integrity can be found at the following address, which also includes links to each Faculty and the School of Graduate Studies: https://www.concordia.ca/conduct/academic-integrity.html" *[Undergraduate Calendar, Sec 17.10.2]*

**Behaviour**
All individuals participating in courses are expected to be professional and constructive throughout the course, including in their communications.

Concordia students are subject to the Code of Rights and Responsibilities which applies both when students are physically and virtually engaged in any University activity, including classes, seminars, meetings, etc. Students engaged in University activities must respect this Code when engaging with any members of the Concordia community, including faculty, staff, and students, whether such interactions are verbal or in writing, face to face or online/virtual. Failing to comply with the Code may result in charges and sanctions, as outlined in the Code.

**Intellectual Property**

Content belonging to instructors shared in online courses, including, but not limited to, online lectures, course notes, and video recordings of classes remain the intellectual property of the faculty member. It may not be distributed, published or broadcast, in whole or in part, without the express permission of the faculty member. Students are also forbidden to use their own means of recording any elements of an online class or lecture without express permission of the instructor. Any unauthorized sharing of course content may constitute a breach of the Academic Code of Conduct and/or the Code of Rights and Responsibilities. As specified in the Policy on Intellectual Property, the University does not claim any ownership of or interest in any student IP. All university members retain copyright over their work.

**Extraordinary circumstances**

In the event of extraordinary circumstances and pursuant to the Academic Regulations the University may modify the delivery, content, structure, forum, location and/or evaluation scheme. In the event of such extraordinary circumstances, students will be informed of the change.