DATA 5000 B (Fall, 2025)
Introduction to Data Science

Instructor: Ahmed El-Roby

Email: ahmed.elroby@carleton.ca

Office Location: HP 5433

Best Ways to be in Touch: in class, or via email

Class Location: Virtual via Zoom. https://carleton-ca.zoom.us/j/93188821639

Discord Server URL: https://discord.gg/QR2BWBZY

Lecture Times: Thursday, 11:35 AM - 2:25 PM

Course Website:
https://people.scs.carleton.ca/~ahmedelroby/DATA5000_F25/

Graduate courses only:
- Brightspace access for University of Ottawa Students; please see information here: https://gradstudents.carleton.ca/faculty-of-graduate-and-postdoctoral-affairs-access-to-brightspace/
- University of Ottawa Students who need access to SCS IT resources such, as OpenStack and Nextcloud, must submit a request to SCS Tech Support SCS.Tech.Support@cunet.carleton.ca. The request must be sent from their @cmail.carleton.ca email address and the email should say which resource is required and for which course (including section).

Important dates and deadlines can be found here: Dates, Deadlines, and Regulations—Registrar's Office, including class suspension for fall, winter breaks, and statutory holidays.

Course Calendar Description
The course covers topics relevant to data science: working with data, exploratory data analysis, data mining, machine learning. The concepts are illustrated using Python. Students also receive hands-on tutorials. Students will be evaluated by their course projects. The course will be lecture-based and will also o er some hands-on tutorials. The project component will be flexible and will involve data collection, manipulation, and analysis.

Learning Material(s) and Other Course/Lab-Related Resources
No textbooks are needed in the course.

Students are not required to purchase textbooks or other learning materials for this course.

Topics Covered and Learning Outcomes

- Foundations of Data Science and the Data Science Lifecycle
- Exploratory Data Analysis and Data Visualization
- Data Wrangling and Cleaning Techniques
- Introduction to Probability and Statistical Inference
- Supervised Learning: Regression and Classification
- Unsupervised Learning: Clustering and Dimensionality Reduction
- Working with Real-world Datasets (e.g., open data, health, transportation)
- Introduction to Natural Language Processing
- Reproducibility and Scientific Computing Practices
- Introduction to Scalable Data Science (e.g., MapReduce and Spark)

By the end of this course, students will be able to:

- Explain the key stages in the data science workflow and apply them to real-world data problems.
- Perform exploratory data analysis and generate meaningful visualizations using Python libraries such as Pandas, Matplotlib, and Seaborn.
- Clean and prepare raw datasets for analysis through appropriate data wrangling techniques.
- Apply fundamental statistical concepts to analyze data and interpret results.
- Implement basic machine learning models for classification, regression, and clustering tasks, and evaluate their performance.
- Use real-world datasets to draw insights and communicate findings effectively through visual, written, and oral presentations.
- Develop reproducible code and workflows using version control and scientific computing tools.
- Gain exposure to scalable data processing frameworks and appreciate their role in handling large datasets.

Assessment Scheme
Grade Breakdown

COMPONENT	GRADE VALUE	DATE
Project Report	50 %	December 11
Project Presentation	15 %	December 4
Paper Presentation	15 %	November 6
Project Proposal Presentation	15 %	October 16
Presentation Discussion	10%	October 16 and December 4

Late and Missed Work Policies

Late Work

The project report can be accepted up to 4 days later (until December 15). Each day passing will incur a 10% penalty. No submissions will be accepted after December 15.

Missed Work

The rest of the coursework comprises presentations, which if missed, cannot be presented later.

The final project must be submitted. Please refer to academic considerations form and longer-term accommodation for more information.

School of Computer Science Laptop Requirement (only applies to on-campus courses)
Every student that has been enrolled in a 1000-level (i.e., first year) course offered is required to have a laptop. This includes COMP1001, COMP1005, and COMP1006. For more information, please visit https://carleton.ca/scs/scs-laptop-requirement/ and then review the requirements at https://carleton.ca/scs/scs-laptop-requirement/laptop-specs/.

Graduate Academic Advisors
The Graduate Advisors for the School of Computer Science are available in Room 5302 HP; or by email at grad.scs@carleton.ca. The graduate advisors can assist with understanding your academic audit and the remaining courses required to meet graduation requirements.

SCS Computer Laboratory
Students taking a COMP course can access the SCS computer labs. The lab schedule and location can be found at: https://carleton.ca/scs/tech-support/computer-laboratories/. All SCS computer lab and technical support information can be found at: https://carleton.ca/scs/tech-support/. Technical support staff may be contacted in-person or virtually, see this page for details: https://carleton.ca/scs/tech-support/contact-it-support/.

Mental Health and Wellness
The Carleton Wellness Website is a wonderful resource link to include in the course outline for students.

Academic Accommodations and Regulations

Academic Accommodation

Carleton is committed to providing academic accessibility for all individuals. You may need special arrangements to meet your academic obligations during the term. The accommodation request processes are outlined on the Academic Accommodations website (https://students.carleton.ca/course-outline/).

Chat GPT/Generative AI Usage

- **AI use in this course:** Students may use AI tools for basic word processing and formatting functions, including:

- Grammar and spell checking (e.g., Grammarly, Microsoft Word Editor).

- Basic formatting and design suggestions (e.g., Microsoft Word's formatting tools, PowerPoint Design editor).

Documenting AI use: It is not necessary to document the use of AI for the permitted purposes listed above. If you have questions about a specific use of AI that isn't listed above, please consult your instructor.

Why have I adopted this policy? This policy ensures that student voices and ideas are prioritized and authentically represented, maintaining the integrity of the work produced by students while allowing basic support to enhance clarity, correctness, layout, and flow of ideas. The goal of adopting a limited use of AI is to help students develop foundational skills in writing and critical thinking by practicing substantive content creation without relying on AI support.

Academic Integrity

Misconduct in scholarly activity will not be tolerated and will result in consequences as outlined in Carleton University's Academic Integrity Policy. A list of standard sanctions in the Faculty of Science can be found here.

Additional details about this process can be found on the Faculty of Science Academic Integrity website.

Students are expected to familiarize themselves with and abide by Carleton University's Academic Integrity Policy.

> If you are unsure of the expectations regarding academic Integrity (how to use and cite references, if unauthorized collaboration with lab- or classmates is permitted (and, if so, to what degree), then you must ASK your instructor. Sharing assignment or quiz specifications or posting them online (to sites like Chegg, CourseHero, OneClass, etc.) is ALWAYS considered academic misconduct. You are NEVER permitted to post, share, or upload course materials without explicit permission from your instructor. Academic integrity offences are reported to the office of the Dean of Science. Information, process and penalties for such offences can be found on the ODS webpage: https://science.carleton.ca/students/academic-integrity/.

Student Rights & Responsibilities

Students are expected to act responsibly and engage respectfully with other students and members of the Carleton and the broader community. See the [7 Rights and Responsibilities Policy](#) for details regarding the expectations of non-academic behaviour of students. Those who participate with another student in the commission of an infraction of this Policy will also be held liable for their actions.

Student Concerns

If you have any concerns regarding this course, your first point of contact is me. Please email me or visit during my student hours, and I will do my best to address your concerns. If I cannot resolve the issue, the next point of contact is the School of Computer Science at studentconcerns@scs.carleton.ca. If the concern remains unresolved, the final point of contact is the Office of the Dean of Science at ODScience@carleton.ca. Please follow this order of contact.

Note: You can also bring your concerns to Ombuds services.