DATA 1517 - Foundation of Data Science

Instructor: Mohamad Elmasri **MohamadElmasri@cunet.carleton.ca**
Office: 5249 Herzberg Laboratories

Lectures: Tuesdays and Thursdays 10:05 – 11:25 am. In person location at Southam Hall 517. Attendance will be mandatory for both lectures.

Tutorials: Thursday 2:35 pm - 3:25 pm, in person at Azrieli Pavilion 132 (or online). Days and times differ. Please ensure you attend your own tutorial. Announcement to be made in brightspace.

Labs: Will be posted online every week on Wednesday, for 10 weeks. Submissions are online in PDF format, the Friday 9 days after the submission. Refer to Brightspace for in-person lab hours. Install RStudio to complete the labs.

Office Hours: Tuesdays 11:30 - 12:30 - at 5249 Herzberg Laboratories (on the weeks with in-person lecture, otherwise online with – see brightspace announcement).

Evaluation: Your final grade will be calculated as:

- Term Mark (60%)
  - Tests (20%, 10% first test, 10% second test)
  - Assignments (30 %)
  - Labs (10%)
- Final Examination Mark (40%)

Assignments

There will be 4 assignments, each counting equally toward the term mark. No late assignments will be accepted. Assignments are to be uploaded to brightspace, **no in-hand assignments. See assignment submission regulation.**

Tentative Assignments Schedule

Assignment	Available	Due
Assignment 1	Monday, September 15th	Friday, October 3rd
Assignment 2	Monday, September 29th	Friday, October 17th

Assignment 3	Monday, October 13th	Friday, November 7th
Assignment 4	Monday, November 10th	Friday, December 5th

Tests

There will be 2 midterm tests in class on Thursday October 16th, 2025 and Thursday November 20th, 2025. You must have a valid reason for missing a test. Any missed test would

NOTE

1. A mark of at least 45% is required on the final exam in order to receive a passing grade.
2. A mark of at least 45% is required in your term work in order to receive a passing grade.
3. At least 3 assignments must be handed in, or an automatic grade of F will be given for the course.

Solutions: Solutions to all assignments, tests, and review problems will be posted on Brightspace.
The outline below is approximate only. Some topics may require more time than specified, while others may require less. More information on exact coverage will be provided in class.

DATA 1517 **Approximate** Weekly Outline

WEEK	SECTIONS	TOPICS
1	Introduction	Introduction to the course, Cause and Effect.
2	Foundations	Early concepts of data and variables.Tables, Sequences, Ranges
3	Programming Basics	Basic programming constructs in R.
4	Data Visualization	Working with data in a computational environment.
5	Empirical probability	Ratio, proportions, coin toss, and iterations, and loops
6	Sampling	Continuous and deterministic, randomization.
7	Distributions	Introduction to data distributions and initial statistical concepts.
8	Inference I	Hypothesis testing, decision rules, A/B testing. Test statistics
9	Inference II	Confidence intervals, properties of the mean and median.
10	Normal Distribution	The Normal Distribution and its properties and Central limit theorem
11	Modeling I	Correlation and Regression.
12	Modeling II	Classification I - K-NN

WEEK	SECTIONS	TOPICS
1	Introduction	Introduction to the course, Cause and Effect.
2	Foundations	Early concepts of data and variables.Tables, Sequences, Ranges
3	Programming Basics	Basic programming constructs in R.
4	Data Visualization	Working with data in a computational environment.
5	Empirical probability	Ratio, proportions, coin toss, and iterations, and loops
13	Advanced Modeling	Training, evaluation, prediction and overfitting.

WARNING: The above weekly schedule is subject to change. Make sure you keep up to date with any changes in material covered, order of presentation, etc.

Lecture format:

- In person lectures:
  - September 4th to September 25th, 2025.
  - November 4thth to November 29th, 2025.
- Online Lectures:
  - September 30th to October 30th, 2025
  - December 2nd to December 5th, 2025.

References:

Coding and Mathematics Resources

R references

- R for Data Science
- Advanced R – For enthusiast
- Cursor – The future of IDE for data science and programming.

LaTeX

- Quick Guide to Overleaf and LaTeX

Books

Because data science is a relatively new and rapidly evolving discipline there is no single ideal textbook for this subject. Instead we plan to use reading from a collection of books all of which are free. However, we have listed a few optional books that will provide additional context for those who are interested.

- Data Science a First introduction. Free online resource.
- Principles and Techniques of Data Science, the Data 100 textbook.
- Introduction to Statistical Learning (Free online PDF) This book is a great reference for the machine learning and some of the statistics material in the class
- Data Science from Scratch This more applied book covers many of the topics in this class using Python but doesn't go into sufficient depth for some of the more mathematical material.
- Doing Data Science This book provides a unique case-study view of data science using R.
- Computational and Inferential Thinking: The Foundations of Data Science

Academic affairs

Fall 2025 Important Dates

- September 3, 2025: Fall term begins.
- September 16, 2025: Last day for registration and course changes for fall term and fall/winter courses.
- September 30, 2025: Last day to withdraw from fall term and fall/winter courses with a full fee adjustment.
- October 13, 2025: Statutory holiday. University closed.
- October 20-24, 2025: Fall break. No classes.
- December 5, 2025: Fall term ends. Last day of classes.
- December 8-20, 2025: Final examination period

Academic Accommodations for Students with Disabilities
Students with disabilities requiring academic accommodation must register with the Paul Menton Centre for Students with Disabilities (PMC) for a formal evaluation of disability- related needs. Registered PMC students are required to contact the PMC every term to have a Letter of Accommodation sent to the instructor by their coordinator. In addition, students are expected to confirm their need for accommodation with the instructor no later than two weeks before the first assignment is due, or two weeks before the first test. If you require accommodations only for the final examination in this course, you must request accommodations by March 15. For more details see the following webpage: https://students.carleton.ca/course-outline/

Academic Integrity
Students are required to be familiar with the **Academic Integrity Policy at Carleton University.** Students who violate the standards of academic integrity relating to any coursework will be required to meet with the Associate Dean of Science.

Course Contents

The outline given is approximate only. Some topics may require more time than specified while others may require less. More information on exact coverage will be given in class.

You are responsible for all material covered in class whether it is in the textbook or not unless you are specifically told otherwise. Course notes will be posted in Brightspace - BUT not any extra explanations given in class. You are also responsible for any announcements made in class. Most of them will also be posted in Brightspace.