

Preliminary
Overview &
Document

This document will be finalized by the weekend of the first week of classes. The information here is tentative and *not* final.

Welcome to **MIE 1626 – Data Science Methods!** The course introduces core techniques in data science, in lectures, tutorials, and mini-projects. The lectures, tutorials, and mini-projects will support students' work on their course project and on future research and professional work involving data. Students will select a dataset of interest to them and produce an analysis and a project report. Students will use domain knowledge and technical expertise to produce their analyses.

Website &
Forum

Website: <http://www.cs.toronto.edu/~guerzhoy/1626s26>

Forum: <https://piazza.com/utoronto.ca/summer2026/mie1626>

All course handouts will be posted on the course website. *Students are responsible for reading all announcements on the course forum on Piazza.*

References

There is no required textbook. Lecture slides will be posted on the course website. The following are recommended as references.

- Andrew Gelman and Jennifer Hill. **Data Analysis using Regression and Multi-level/Hierarchical Models**. Cambridge University Press, 2006
- Cosma Rohilla Shalizi. **Advanced Data Analysis from an Elementary Point of View**. Cambridge University Press (forthcoming). Preprint available at <https://www.stat.cmu.edu/~cshalizi/ADAfaEPoV/>
- Fred Ramsey and Daniel Schafer. **The Statistical Sleuth: A Course in Methods of Data Analysis**, 3rd ed. Brooks/Cole, 2013
- Kieran Healy. **Data Visualization: A Practical Introduction**. Princeton University Press, 2026

Instructor
Contact
Grading

Instructor	Email	Office	Office Hours
Michael Guerzhoy	guerzhoy@cs.toronto.edu	BA 2028	TBA

Lectures and tutorials take place on Mondays and Fridays, 2:00pm–5:00pm. Students are expected to attend both lecture and tutorial blocks.

The grading scheme and tentative schedule for the course is as follows.

	Worth	Due
Mini-Project 1	5%	May 16
Preliminary Project Proposal	2%	May 19
Mini-Project 2	5%	May 23
Test 1	10%	May 27
Mini-Project 3	5%	May 30
Mini-Project 4	5%	Jun 6
Preliminary Project Presentation + Document	8%	Jun 15
Test 2	30%	To be scheduled after classes end
Course Project	30%	Jul 2

Late work

Each student starts the term with 7 “grace days.” You can use grace days to submit late work with no penalty. For example, a student who submits their work 25 hours after the deadline will use up two grace days. You cannot use more than 3 grace days at a time. Grace days do not apply to the preliminary project presentation, tests, or the final project.

**Course
project**

Each student (or pair of students) in the course will work on a large-scale data science project. Students are encouraged to pick a dataset of interest to them. Students should aim to achieve one or several of the following.

- Obtain interesting insights about the dataset
- Apply interesting data analysis techniques to the dataset
- Demonstrate a new data analysis technique and apply it to the dataset
- Collect a novel dataset

Course projects will contain

- A description of the dataset being analyzed, and a summary of the relevant domain knowledge. For example, if working with audio data, a summary of how sound is generated, perceived, and processed would be appropriate. If working with hospital data involving length-of-stay times, it would be appropriate to summarize the relevant literature on hospital lengths of stay.
- A summary of previous work that is related to the data analysis project.
- An exploratory analysis of the dataset being analyzed.
- A description of the method being applied, and a summary of related work.
- A description of the outcomes of the analysis, or an analysis of a predictive system that was built.

Students will submit a preliminary project proposal by May 19 and a preliminary project document, which should contain a full description of the project and initial exploratory data analysis results, by Jun 15.

Each student will give a preliminary project presentation to the class on Jun 15.