



# MAT4376E Topics in Statistics

# MAT5314E Topics in Probability and Statistics

Techniques of Data Analysis | Fall 2024 | Instructor: P. Boily

In October 2012, the *Harvard Business Review* published an article calling data science the **sexiest job of the 21st century**, a long cry from the business-as-usual practice of data geeks playing a supporting role in organizations. Today's data scientists are not just number-crunchers – as a combination of **data hacker, analyst, communicator, and trusted adviser**, they discover meaningful relationships in ever-growing masses of information, and play a leading role in the decision-making processes.

In **groups**, you will apply specific data analysis approaches and investigate real-world datasets through the “Multiple I’s” framework.

---

## PRE-REQUISITES

Programming proficiency (R/Python/Matlab/etc.); MAT2122, MAT2141, MAT2371, MAT2375 or MAT2377, and MAT3375 (or permission).

## COURSE SCHEDULE

LEC: MON 10:00-11:20 (STE G0103)  
LEC: WED 08:30-09:50 (STE G0103)  
OFF: by appointment (ZOOM, SLACK)

---

## DELIVERABLES

### Projects:

- 1) Data Visualization, 04-Oct
- 2) Bayesian Data Analysis, 25-Oct
- 3) Queueing Systems, 13-Nov
- 4) Anomaly Detection & Outlier Analysis, 06-Dec
- 5) Graduate Project, 13-Dec

**4376E:** only projects 1-4 must be completed.

**5314E:** all projects have to be completed. Graduate project topics must be approved by 20-Sep.

All projects are weighed equally.

See **Course Logistics** for more details.

## NOTES

**Initiative** and **independence** (and the other “multiple I’s”) are rewarded: projects which stay solely within the suggested guidelines can score at most a B+.

**Engagement** is mandatory in this course. Students who do not participate in at least 70% of the sessions will be docked 10 marks.

---

## EXPECTATIONS

You are expected to spend 8-10 hours [4376] or 12-15 hours [5314] per week on this course.

**Teamwork** is crucial to insightful data analysis; in this course, all projects are done in teams of 3 or 4 (which can change from one project to another). Be advised that a grade is given to the whole group (independently of the quantity and quality of the work performed by each person).

You may need to use methods or concepts that have not been discussed in the lectures. More details will be provided in class.

One objective is to learn to navigate tight deadlines, and to plan your analysis/reporting accordingly (**10 page limit/report, no exception**). Do not wait before starting work on your projects.

There may be times when you are unable to deliver the projects by the deadline due to reasons outside your control. You are requested to inform me (and to submit the work you have already completed) as soon as you become aware of such a situation (within reason) so that we can discuss alternatives.