MAT4376E Topics in Statistics
MAT5314E Topics in Probability and Statistics
Techniques of Data Analysis | Fall 2021 | 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 or individually, you will apply specific data analysis tasks 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 08:30-10:00 (MNT 103)
LEC: WED 13:00-14:30 (MNT 204)
OFF: by appointment (ZOOM, SLACK)

DELIVERABLES

Projects:
1) Data Visualization, 08-Oct
2) Bayesian Data Analysis, 29-Oct
3) Queueing Systems, 19-Nov
4) Anomaly Detection & Outlier Analysis, 10-Dec
5) Graduate Project, 17-Dec

4376E: only projects 1-4 must be completed.
5314E: all projects have to be completed. Graduate project topics must be approved by 24-Sep.

NOTES

Initiative and **independence** are rewarded: projects which stay solely within the suggested guidelines can score at most a A− [4376E] or a B+ [5314E].

Multiple I's Framework: intuition, initiative, innovation, interpretability, insights, integrity, independence, interaction, inquisitiveness.

EXPECTATIONS

You are expected to spend 8-10 hours [43756E] or 12-15 hours [5314E] per week on this course.

Team work is crucial to insightful data analysis. You are encouraged to work in teams of 2/3, but it is not mandatory for you to do so. Be advised that the grade will be given to the whole group (independently of the quantity and quality of the work performed by each person).

You may have 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 (12 page limit). 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.