DALHOUSIE UNIVERSITY

School of Information Management
INFO 6270 Introduction to Data Science
Winter 2022

Course Type: Face to face
Room: McCain 1170 (lectures 8:35-9:25 am) & Rowe 3080 (labs 9:35-11:25 am)
Day and time: Thursday, 8:35 am - 11:25 am
Instructor: Philippe Mongeon (he/him)
Office: Rowe 4032
Contact info: PMongeon@dal.ca
Preferred method of contact: Walk-in or scheduled meeting, email
Office hours: If my office door is open, I'm available. You can also book a virtual or in-person meeting anytime you want: https://calendly.com/pmongeon/

Students are encouraged to take advantage of my office hours even if their questions are not strictly related to the content of the course. I try to respond to emails promptly but may take up to 72 hours.

COURSE DESCRIPTION

This course serves as an introduction to data science, an increasingly important set of skills and techniques for business intelligence, effective governance, and the research process. The amount of data we generate increases year on year. As computers have begun to play roles in many aspects of our daily life, our actions and interactions leave digital traces. This has led both to an explosion in the amount of data that we generate and an increased interest in analyzing and understanding that data. This class will give you an introduction to the skills you need to effectively collect, manipulate, and analyze data yourself. Rather than being constrained to using any specific data analysis software, we will focus on using the flexible programming language R. You will receive a thorough introduction to R, learning how to use a variety of its built-in capabilities as well as a number of available data analysis packages. By the course's end you should be capable enough that you will be able to begin teaching yourself and expanding your data science skills.

IMPORTANT NOTES

1. This course is intended for an audience with little-to-no programming experience.
2. This course has few readings and emphasizes on hands-on learning.

COURSE PRE-REQUISITES

There are no pre-requisites for this course.

COURSE OBJECTIVES

This course will introduce students to computer programming, computational thinking, data analysis and basic statistics programming. It will give students a foundation from which to build upon in order to learn more advanced and specialized computational research and data analytic skills in the future.

LEARNING OUTCOMES

Upon completion of the course, students will:

1. Understand the principles of data analysis

2. Know how to use R
3. Know how to manipulate (wrangle) data
4. Know how to use descriptive statistics
5. Know how to visualize data
6. Know how to plan a data science project
7. Know how to communicate data science projects
8. Know how to perform text-mining tasks in R
9. Know how to perform network analysis in R
10. Understand the basics of inferential statistics
11. Understand the basics of machine learning

TECHNOLOGY USED

- R

INSTRUCTIONAL METHODS

- Labs
- Self-guided tutorials
- Lectures

LEARNING MATERIALS

There is no required textbook for this course. The course website will contain everything you need. The However, we will be drawing from several open resources, which can be valuable to further develop your understanding and skills, these are all available online for free (see link in the table below):

Title	Note
Grolemund, G. (2014). *Hands-on programming with R*. O'Reilly. https://rstudio-education.github.io/hopr/	Covers the very basics of R programming
Hadley, W. (2016). *ggplot2*. Springer Science+Business Media, LLC. https://ggplot2-book.org/	Good reference for data visualization with the ggplot2 package
Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach*. O'Reilly. https://www.tidytextmining.com/	The text-mining module draws heavily from this resource.
Wickham, H., & Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data*. O'Reilly. https://r4ds.had.co.nz/	Comprehensive resource that covers a lot of what we will do in the course.
Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R Markdown: The definitive guide*. Taylor & Francis, CRC Press. https://bookdown.org/yihui/rmarkdown/	Comprehensive resource for making beautiful and effective R Markdown documents and websites.

METHODS OF EVALUATION

Detailed instructions regarding each assignment will be provided. Assessment of all assignments is directly related to attention to the instructions, clarity of expression and presentation, and evidence of significant analysis and reflection.

See also the SIM Grading Policy.

COMPONENT	DETAILS	DUE DATE	VALUE
Weekly assignments	Every week, we will work on small in-class exercises together following the presentation and scrum meeting. 10 of the exercises must be submitted on Brightspace for 5 points each. The exercises are due 11 days following their release, always on a Friday at 11:55 pm.	Weekly	50%
Individual project	**Propose** a small independent data science project.	2021-02-12	10%
	Complete a small independent data science project.	2021-04-15	40%

INTEGRATION OF MI Competencies

PROGRAM COMPETENCY	COURSE LEARNING OUTCOME	COURSE ASSESSMENT
Research and Evaluation	1-11	All assignments
Collaborate & communicate	7	All assignments
Organize, Plan & Manage	6	Individual project

CLASS POLICIES

Attendance
This is an asynchronous course and presence to the supervised labs is not required. However, attendance and punctuality to your team meetings will be evaluated through the peer-assement process.

Citation Style
SIM courses use APA as the default standard citation style. Unless the instructor provides alternative written instructions, please use the APA citation style in your assignments to briefly identify (cite) other people's ideas and information and to indicate the sources of these citations in the References list at the end of the assignment. For more information on APA style, consult Dalhousie Library website at https://libraries.dal.ca/help/style-guides.html or the APA's Frequently Asked Questions about APA

Late penalties for assignments
A penalty for late assignments will be assessed, unless prior permission has been given by the instructor to submit an assignment late, which normally will be for extended illness, medical, or family emergencies only (see below). Late submissions will be assessed a penalty of five percent per day, including weekends. Assignments will not normally be accepted seven days or more after the due date; in such cases the student will receive a grade of zero.

Missed or Late Academic Requirements due to Student Absence:
Dalhousie University recognizes that students may experience short-term physical or mental health conditions, or other extenuating circumstances that may affect their ability to attend required classes, tests, exams or submit other coursework.

Dalhousie students are asked to take responsibility for their own short-term absences (3 days or less) by contacting their instructor by phone or email prior to the academic requirement deadline or scheduled time **AND** by submitting a completed Student Declaration of Absence form to their instructor in case of missed or late academic requirements. Only 2 separate Student Declaration of Absence forms may be submitted per course during a term.

SIM GRADING POLICY

A+	90-100	Demonstrates original work of distinction.
A	85-89	Demonstrates high-level command of the subject matter and an ability for critical analysis.
A-	80-84	Demonstrates above-average command of the subject matter.
B+	77-79	Demonstrates average command of the subject matter.
B	73-76	Demonstrates acceptable command of the subject matter.
B-	70-72	Demonstrates minimally acceptable command of the subject matter.
F	<70	Unacceptable for credit towards a Master's degree.

ACCOMMODATION POLICY FOR STUDENTS

The Student Accessibility Centre is Dalhousie's centre of expertise for student accessibility and accommodation. The advising team works with students on the Halifax campus who request an accommodation as a result of a disability, religious obligation, or any barrier related to any other characteristic protected under Human Rights legislation (NS, NB, PEI, NFLD).
If there are aspects of the design, instruction, and/or experiences within this course that result in barriers to your inclusion please contact the Student Accessibility Centre. Please visit www.dal.ca/access for more information and to obtain the Request for Accommodation form.

A note-taker may be required as part of a student's accommodation. Visit https://www.dal.ca/campus_life/academic-support/accessibility/accommodations-/classroom-accommodation.html for more details.

Please note that your classroom may contain accessible furniture and equipment. It is important that these items remain in the classroom, undisturbed so that students who require their use will be able to fully participate.

ACADEMIC INTEGRITY

At Dalhousie University, we are guided in all of our work by the values of academic integrity: honesty, trust, fairness, responsibility and respect. As a student, you are required to demonstrate these values in all of the work you do. The University provides policies and procedures that every member of the university community is required to follow to ensure academic integrity.

The commitment of the Faculty of Management is to graduate future leaders of business, government and civil society who manage with integrity and get things done. This is non-negotiable in our community and it starts with your first class at Dalhousie University. So when you submit any work for evaluation in this course or any other, please ensure that you are familiar with your obligations under the Faculty of Management's Academic Integrity Policies and that you understand where to go for help and advice in living up to our standards. You should be familiar with the Faculty of Management Professor and Student Contract on Academic Integrity, and it is your responsibility to ask questions if there is anything you do not understand.

Dalhousie offers many ways to learn about academic writing and presentations so that all members of the University community may acknowledge the intellectual property of others. Knowing how to find, evaluate, select, synthesize and cite information for use in assignments is called being "information literate." Information literacy is taught by Dalhousie University Librarians in classes and through Dalhousie Libraries' online Citing & Writing tutorials.

Do not plagiarize any materials for this course. For further guidance on what constitutes plagiarism, how to avoid it, and proper methods for attributing sources, please consult the University Secretariat's Academic Integrity page.

Please note that Dalhousie subscribes to plagiarism detection software that checks for originality in submitted papers. Any paper submitted by a student at Dalhousie University may be checked for originality to confirm that the student has not plagiarized from other sources. Plagiarism is considered a very serious academic offence that may lead to loss of credit, suspension or expulsion from the University, or even the revocation of a degree. It is essential that there be correct attribution of authorities from which facts and opinions have been derived. At Dalhousie, there are University Regulations that deal with plagiarism and, prior to submitting any paper in a course; students should read the Policy on Academic Dishonesty contained in the Calendar.

Furthermore, the University's Senate has affirmed the right of any instructor to require that student assignments be submitted in both written and computer-readable format, e.g.: a text file or as an email attachment, and to submit any paper to a check such as that performed by the plagiarism detection software. As a student in this class, you are to keep an electronic copy of any paper you submit, and the course instructor may require you to submit that electronic copy on demand. The use of third-party originality checking software does not preclude instructor use of alternate means to identify lapses in originality and attribution. The result of such assessment may be used as evidence in any disciplinary action taken by the Senate.

Finally:

If you suspect cheating by colleagues or lapses in standards by a professor, you may use the confidential email: ManagementIntegrity@dal.ca which is read only by the Assistant Academic Integrity Officer.

> **Faculty of Management clarification on plagiarism versus collaboration:**
>
> There are many forms of plagiarism, for instance, copying on exams and assignments. There is a clear line between group work on assignments when explicitly authorized by the professor and copying solutions from others. It is permissible to work on assignments with your friends but only when the professor gives you permission in the specific context of the assignment. University rules clearly stipulate that all assignments should be undertaken individually unless specifically authorized.
>
> Specific examples of plagiarism include, but are not limited to, the following:
> - Copying a computer file from another student, and using it as a template for your own solution
> - Copying text written by another student
> - Submitting the work of someone else, including that of a tutor as your own
>
> An example of acceptable collaboration includes the following:
> - When authorized by the professor, discussing the issues and underlying factors of a case with fellow students, and then each of the students writing up their submissions individually, from start to finish.

UNIVERSITY STATEMENTS
This course is governed by the academic rules and regulations set forth in the University Calendar and the Senate.

ACCESSIBILITY
The Advising and Access Centre serves as Dalhousie's Centre for expertise on student accessibility and accommodation. Our work is governed by Dalhousie's Student Accommodation Policy, to best support the needs of Dalhousie students. Our teams work with students who request accommodation as a result of disability, religious obligation, an experienced barrier related to any other characteristic protected under Canadian Human Rights legislation.

STUDENT CODE OF CONDUCT
Everyone at Dalhousie is expected to treat others with dignity and respect. The Code of Student Conduct allows Dalhousie to take disciplinary action if students don't follow this community expectation. When appropriate, violations of the code can be resolved in a reasonable and informal manner—perhaps through a restorative justice process. If an informal resolution can't be reached, or would be inappropriate, procedures exist for formal dispute resolution.

DIVERSITY AND INCLUSION
Every person at Dalhousie has a right to be respected and safe. We believe inclusiveness is fundamental to education. Dalhousie is strengthened in our diversity and dedicated to achieving equity. We are committed to being a respectful and inclusive community where everyone feels welcome and supported, which is why our university prioritizes fostering a culture of diversity and inclusiveness.

RECOGNITION OF MI'KMAQ TERRITORY
Dalhousie University is located in Mi'kma'ki, the ancestral and unceded territory of the Mi'kmaq. We are all Treaty people. For more information about the purpose of territorial acknowledgements, or information about alternative territorial acknowledgements if your class is offered outside of Nova Scotia, please visit https://native-land.ca/.

The Elders in Residence program provides students with access to First Nations elders for guidance, counsel and support. Visit the office in the McCain Building (room 3037) or contact the programs at elders@dal.ca or 902-494-6803 (leave a message).

[FAIR DEALING POLICY](#)
The Dalhousie University Fair Dealing Policy provides guidance for the limited use of copyright protected material without the risk of infringement and without having to seek the permission of copyright owners. It is intended to provide a balance between the rights of creators and the rights of users at Dalhousie.

COURSE SCHEDULE

Date of Class	Topics
Week 1 (Jan 6)	Course introduction Getting started with R
Week 2 (Jan 13)	Getting started with R pt.2
Week 3 (Jan 20)	Wrangling data
Week 4 (Jan 27)	Wrangling data pt.2
Week 5 (Feb 3)	Analyze data (descriptive statistics)
Week 6 (Feb 10)	Visualize data
Week 7 (Feb 17)	Text-mining
Reading week (Feb 21-25)	No class
Week 8 (Mar 3)	Text-mining pt.2
Week 9 (Mar 10)	Network analysis
Week 10 (Mar 17)	Basic inferential statistics
Week 11 (Mar 24)	Basic machine learning
Week 12 (Mar 31)	Basic machine learning pt.2 Course wrap-up