DALHOUSIE UNIVERSITY

CSCI 1109 — Practical Data Science
Course Syllabus

Instructor Information

Instructor:	Frank Rudzicz	**Office:**	1577 Barrington
E-mail:	`frank@dal.ca`	**Office Hours:**	TBD
Class Meeting Time:	TR 8h35-9h55	**Room No:**	Studley LSC 208.
Course Homepage:	Here		
Teaching Assistants:	N/A		

Important Dates

See Academic dates & deadlines and course website for any changes.

8 January	First lecture
22 January	Last day to add CSCI 1109
16–20 February	Winter study break – no lectures or office hours
9 April	Last lecture
9 April	Final exam

Course Overview

Data is transforming every domain of society, from health care and science to business and the arts. This course provides a hands-on introduction to the methods and mindsets of practical data science. Students will learn how to acquire, clean, analyze, and visualize data, while developing a critical awareness of the opportunities and risks of data-driven decision making. This course introduces students to the practice of data science through a combination of programming, mathematics, and applied case studies. Using Python and its data science libraries, students will learn how to acquire, clean, analyze, and visualize data, while developing an awareness of the opportunities and risks of data-driven decision making. Emphasis is placed on hands-on problem solving, clear communication of results, and critical reflection on the ethical and societal dimensions of data science.

Learning Outcomes

Upon completion of this course, students will be able to:

- Write and debug Python programs for data analysis using packages such as *pandas*, *NumPy*, and *Matplotlib*.
- Formulate questions and hypotheses that can be addressed with data, and design analyses to investigate them.
- Apply key concepts in descriptive and inferential statistics, including probability distributions, correlation, and hypothesis testing.
- Implement basic machine learning methods such as regression, classification, and clustering using *scikit-learn*.
- Construct, interpret, and critique data visualizations as tools for exploration and storytelling.
- Analyze the structure and properties of networks using *NetworkX*.
- Evaluate the limitations of data science methods and articulate ethical risks of data misuse and algorithmic bias.
- Present clear, reproducible analyses of real-world datasets, both in written form and through visual or oral presentation.

Course Rationale

Data is ubiquitous, and the ability to work with it is becoming an essential skill across disciplines. This course equips students with the practical and conceptual foundations of data science, enabling them to manipulate and analyze diverse datasets and to interpret results critically. In doing so, it prepares students both for more advanced study in computer science and for thoughtful engagement with data in their academic, professional, and civic lives.

Course Topics

1. Foundations of Data Science and Python
   (a) What is data science?
   (b) Python fundamentals: variables, data types, functions, and control flow
   (c) Working with tabular data using `pandas`
   (d) Data cleaning and preparation
2. Mathematical and Statistical Principles
   (a) Descriptive statistics and visualization
   (b) Probability distributions and hypothesis testing
   (c) Correlation, causation, and inference
3. Machine Learning Concepts
   (a) Supervised vs. unsupervised learning
   (b) Regression and classification models
   (c) Clustering techniques (e.g., k-means)
4. Networks and Structure in Data
   (a) Graph theory fundamentals: nodes, edges, degree, clustering coefficient
   (b) Network analysis with `NetworkX`
   (c) Applications in social and information networks
5. Ethics and Responsible Data Science
   (a) Algorithmic bias and fairness
   (b) Data privacy and responsible use
   (c) Communicating results and ethical reflections

Course Materials

- Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong (2020) Mathematics for Machine Learning, Cambridge University Press.
- Kevin Murphy (2022) Probabilistic machine learning.

Assessments

- Participation and discussion: 10%
  - Engagement in in-class activities, including submission of in-class module notebooks. These low-stakes activities provide opportunities for reflection and retrieval practice. Notebooks can be completed and submitted after class, but with a maximum score of 33%.

  - This is measured in teams and is maintained on the class Leaderboard.

- Weekly quizzes: 25%
  - Short quizzes covering core concepts, readings, and technical skills. The focus is on spaced retrieval across topics to strengthen long-term understanding.
  - Your lowest quiz mark will be discarded

- Practical Assignments (x5): 35%
  - Assignments are completed individually and weighted in increasing order of performance: the lowest-to-highest performing assignments contribute 5%, 6%, 7%, 8%, and 9% of the course grade. This structure reduces the penalty for struggles while rewarding consistent improvement. Assignments are worked on during the labs.

- Midterm exam: 15%
  - A mid-course assessment emphasizing mastery of core programming, statistics, and conceptual material.

- Final exam: 15%
  - A written and practical exam testing cumulative understanding.

Notes

- A minimum C grade is required in this course if it is core to your FCS degree, or if it will be used as a prerequisite for a subsequent CSCI course.
- As of 2019, students who receive a grade lower than C in the same required CS course twice, will be dismissed.
- The grade conversion scale in Section 17.1 of the Academic Regulations, Undergraduate Calendar will be used.
- A 24-hour 'silence policy' will be in effect – we do not guarantee that the instructors or TAs will respond to your request within 24 hours of an assignment's due time.
- **It is up to the discretion of the instructor to use remote proctoring in online testing. Students may be required to download proctoring software onto their devices. Students who cannot meet system requirements for remote proctoring should contact the instructor for an alternate assessment. (Typical system requirements are: (i) Mac OS or Windows, (ii) a web-cam, and (iii) an internet connection.)**

Recommended Software and Tools

- Python (scikit-learn, PyTorch, SHAP, MLFlow)
- Jupyter Notebooks, Git, podman (optional)

Student Declaration of Absence

The Student Declaration of Absence policy shall apply.

Academic Standards

Failure to properly attribute sources in your work will be treated as an academic standards issue and points may be deducted for not following citation requirements. For example, forgetting to quote text taken from other sources, **including large language models**, failure to include in-text citations, or a failure to include required information in the citations or references. Please see the resources on proper citation provided by the Dalhousie Writing Center (`https://dal.ca.libguides.com/c.php?g=257176&p=5001261`).

Please note that if it appears that the error was made with intent to claim other people's work as your own such as a lack of both citations and references, an allegation of plagiarism will be submitted to the Faculty Academic Integrity Officer, which could result in consequences such as a course failure.

Tentative Schedule of Topics

1. Introduction to Data Science & Python [Lec 1]
2. Variables, Loops, Functions in Python [Lec 2]
3. Working with Data in `pandas` [Lec 3]
4. Filtering, Grouping, Aggregating Data [Lec 4]
5. Data Cleaning: Missing Values, Duplicates [Lec 5]
6. Data Transformation & Feature Scaling [Lec 6]
7. Mathematical Foundations: Vectors, Matrices [Lec 7]
8. Descriptive Statistics: Mean, Variance, Distributions [Lec 8]
9. Data Visualization with `Matplotlib` [Lec 9]
10. Storytelling with `Seaborn` Visualizations [Lec 10]
11. Probability Distributions & Hypothesis Testing [Lec 11]
12. Correlation vs. Causation [Lec 12]
13. Introduction to Machine Learning [Lec 13]
14. k-Nearest Neighbors and Classification [Lec 14]
15. Linear Regression Models [Lec 15]
16. Logistic Regression & Error Metrics [Lec 16]
17. Clustering with k-Means [Lec 17]
18. Graph Theory: Nodes, Edges, Clustering Coefficient [Lec 18]
19. Network Analysis with `NetworkX` [Lec 19]
20. Ethics in Data Science: Bias & Fairness [Lec 20]
21. Data Privacy & Responsible Use [Lec 21]
22. Course Review & Future Directions [Lec 22]

Responsible Computing Policy

Usage of all computing resources in the Faculty of Computer Science must be within the Dalhousie Acceptable Use Policies and the Faculty of Computer Science Responsible Computing Policy.

Use of Plagiarism Detection Software

All submitted assignment may be passed through a plagiarism detection software, such as the Moss Software Similarity Detection System (`https://theory.stanford.edu/~aiken/moss/`), or similar systems. If a student does not wish to have their assignments passed through plagiarism detection software, they should contact the instructor for an alternative. Please note, that code not passed through plagiarism detection software will necessarily receive closer scrutiny. `https://cdn.dal.ca/content/dam/dalhousie/pdf/dept/university_secretariat/policy-repository/OriginalitySoftwarePolicy.pdf`

Use of Artificial Intelligence Tools

You may use AI-driven tools to assist you in learning but remember that your objective is to understand, achieve, and apply the course competencies and outcomes. While you may use tools for learning, specific assessments in this course will disallow the use of AI-driven tools to assert that you have attained course learning outcomes. This is because a graduate must be able to analyze, assess and produce work unassisted by AI technology. Where tools are allowed: you must acknowledge all tools used to assist you. If applicable, you must provide links to chat logs. Using AI-driven tools where prohibited constitutes an academic offence.

Culture of Respect

Every person has a right to respect and safety. We believe inclusiveness is fundamental to education and learning. Misogyny and other disrespectful behaviour in our classrooms, on our campus, on social media, and in our community is unacceptable. As a community, we must stand for equality and hold ourselves to a higher standard.

What we all need to do [1]:

1. **Be Ready to Act:** This starts with promising yourself to speak up to help prevent it from happening again. Whatever it takes, summon your courage to address the issue. Try to approach the issue with open-ended questions like "Why did you say that?" or "How did you develop that belief?"
2. **Identify the Behaviour:** Use reflective listening and avoid labeling, name-calling, or assigning blame to the person. Focus the conversation on the behaviour, not on the person. For example, "The comment you just made sounded racist, is that what you intended?" is a better approach than "You're a racist if you make comments like that."
3. **Appeal to Principles:** This can work well if the person is known to you, like a friend, sibling, or co-worker. For example, "I have always thought of you as a fair-minded person, so it shocks me when I hear you say something like that."
4. **Set Limits:** You cannot control another person's actions, but you can control what happens in your space. Do not be afraid to ask someone "Please do not tell racist jokes in my presence anymore" or state "This classroom is not a place where I allow homophobia to occur." After you have set that expectation, make sure you consistently maintain it.
5. **Find or be an Ally:** Seek out like-minded people that support your views, and help support others in their challenges. Leading by example can be a powerful way to inspire others to do the same.
6. **Be Vigilant:** Change can happen slowly, but do not let this deter you. Stay prepared, keep speaking up, and do not let yourself be silenced.

University Statements

This course is governed by the academic rules and regulations set forth in the University Calendar and the Senate.
`https://academiccalendar.dal.ca/Catalog/ViewCatalog.aspx?pageid=viewcatalog&catalogid=117&loaduseredits=False`

Territorial Acknowledgement

Dalhousie University is located in Mi'kma'ki, the ancestral and unceded territory of the Mi'kmaq. We are all Treaty people.

Internationalization

At Dalhousie, 'thinking and acting globally' enhances the quality and impact of education, supporting learning that is "interdisciplinary, cross-cultural, global in reach, and orientated toward solving problems that extend across national borders."

`https://www.dal.ca/about-dal/internationalization.html`

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.

(read more: `http://www.dal.ca/dept/university_secretariat/academic-integrity.html`)

[1] Source: Speak Up! ©2005 Southern Poverty Law Center. First Printing. This publication was produced by Teaching Tolerance, a project of the Southern Poverty Law Center. Full "Speak Up" document found at: `http://www.dal.ca/dept/dalrespect.html` Revised by Susan Holmes from a document provided April 2015 by Lyndsay Anderson, Manager, Student Dispute Resolution, Dalhousie University 902.494.4140 `lyndsay.anderson@dal.ca www.dal.ca/think`.

Accessibility

The Student Accessibility Centre is Dalhousie's centre of expertise for matters related to student accessibility and accommodation. If there are aspects of the design, instruction, and/or experiences within this course (online or in-person) that result in barriers to your inclusion please contact:
https://www.dal.ca/campus_life/academic-support/accessibility.html
for all courses offered by Dalhousie with the exceptionof Truro.

Conduct in the Classroom — Culture of Respect

Substantial and constructive dialogue on challenging issues is an important part of academic inquiry and exchange. It requires willingness to listen and tolerance of opposing points of view. Consideration of individual differences and alternative viewpoints is required of all class members, towards each other, towards instructors, and towards guest speakers. While expressions of differing perspectives are welcome and encouraged, the words and language used should remain within acceptable bounds of civility and respect.

Diversity and Inclusion — Culture of Respect

Every person at Dalhousie has a right to be respected and safe. We believe inclusiveness is fundamental to education. We stand for equality. Dalhousie is strengthened in our diversity. We are a respectful and inclusive community. We are committed to being a place where everyone feels welcome and supported, which is why our Strategic Direction prioritizes fostering a culture of diversity and inclusiveness (Strategic Priority 5.2).
(read more:http://www.dal.ca/cultureofrespect.html)

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.
(read more: https://cdn.dal.ca/content/dam/dalhousie/pdf/dept/university_secretariat/policy-repositc
Code%20of%20Student%20Conduct%20rev%20Sept%202021.pdf)

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. (read more: https://www.dal.ca/dept/university_secretariat/policies/academic/fair-dealing-policy-.html)

Originality Checking Software

The course instructor may use Dalhousie's approved originality checking software and Google to check the originality of any work submitted for credit, in accordance with the Student Submission of Assignments and Use of Originality Checking Software Policy. Students are free, without penalty of grade, to choose an alternative method of attesting to the authenticity of their work, and must inform the instructor no later than the last day to add/drop classes of their intent to choose an alternate method. (read more: https://cdn.dal.ca/content/dam/dalhousie/pdf/dept/university_secretariat/policy-repository/OriginalitySoftwarePolicy.pdf)

Student Use of Course Materials

These course materials are designed for use as part of the CSCI courses at Dalhousie University and are the property of the instructor unless otherwise stated. Third party copyrighted materials (such as books, journal articles, music, videos, etc.) have either been licensed for use in this course or fall under an exception or limitation in Canadian Copyright law. Copying this course material for distribution (e.g. uploading material to a commercial third party website) may lead to a violation of Copyright law.

Learning and Support Resources

Please see https://www.dal.ca/campus_life/academic-support.html