

Data Science 6200

Please note that the ongoing COVID-19 pandemic may necessitate a revision of the format of course offerings and academic schedules. Any such changes will be announced via Courselink and/or class email. This includes on-campus scheduling during the semester, midterms and final examination schedules. All University-wide decisions will be posted on the COVID-19 website and circulated by email.

Mental Health One out of every five students in Canada experiences some sort of mental health issue at some point in their academic career. If you find yourself facing a mental health crisis, or just need to talk to someone, please consider taking advantage of one of the following resources available to University of Guelph students:

Counseling Services: (x53244) is located at Health Services (J.T. Powell Building) and offers individual and group counselling sessions by appointment or walk-in.

Student Support Network: is located in Raithby House (across from the cannon) and offers confidential, peer-based, drop-in support.

Good2Talk: (1-866-925-5454) is a free, 24/7 student hotline that provides professional counselling and referrals for mental health, addictions and well-being.

Here 24/7: (1-844-437-3247) specializes in assessment, referral and appointment booking and is available 24/7 for crisis support.

You are not alone and you will not be judged for asking for help.

For information on current safety protocols, follow these links:

<https://news.uoguelph.ca/return-to-campus/how-u-of-g-is-preparing-for-your-safe-return/>

<https://news.uoguelph.ca/return-to-campus/spaces/#ClassroomSpaces>

Please note, these guidelines may be updated as required in response to evolving University, Public Health or government directives. Medical notes will not normally be required for singular instances of academic consideration, although students may be required to provide supporting documentation for multiple missed assessments or when involving a large part of a course (e.g. final exam or major assignment).

Data Science 6200: Fall 2021, Outline

Data Science Program

General Information

Course Title: Data Manipulation and Visualization

Course Description: An analysis is only as good as the data on which it is based and the tables and graphics used to summarize the data and analysis results. Data manipulation and visualization are key components to the data life cycle, and often where most of the analysts time is spent. This course will cover the basics of accessing, cleaning, re-shaping, re-formatting, transforming, selecting, filtering and tidying data in preparation for effective analysis. Generating informative and powerful data tables/graphics will be introduced for checking and maintaining data integrity as well as for insightful exploratory data analysis. Critical assessment of data tables and visualizations as well as good analysis practices will be developed, including review of ethical considerations arising from manipulation and visualization of data. The course ends with the acquisition and manipulation of spatially structured data. Mastering these skills can facilitate and drastically simplify downstream analyses of data.

Course topics include:

- Review of data basics: data types, objects, data structures, and reading/writing data
- Fundamentals of data manipulation: select, filter, transform, arrange, mutate and summarize
- Tools and techniques for data manipulation: group-wise calculations, joining tables, piping, chain operators, working with distributed and backend databases, dealing with dates and character strings
- Fundamentals of visualization: focus, data, aesthetics, scale, statistics, and facets
- Tools and techniques for based on grammar of graphics for multi-dimensional data
- Introduction to spatially structured data: geoprocessing and mapping for raster data, including aggregation, reclassification, cropping, masking and extraction
- Ethical issues including intentional and inadvertent deception, P-hacking, cherry-picking data, and inadvertent and intentional violation of privacy.

Prerequisite(s): Enrollment in the data science program or permission of instructor.

Credit Weight: 0.5

Academic Department: Mathematics & Statistics

Campus: University of Guelph

Semester Offering: Fall 2021

Class Schedule and Location:

Synchronous Lectures through CourseLink 11:30-12:20 MWF

Instructor Information

Instructor Name: Daniel Ashlock

Instructor Email: dashlock@uoguelph.ca

Office location and office hours: Online via Google Meet

Office hours can be found on the class page on CourseLink, they may change.

Course Content**Specific Learning Outcomes:**

This course will introduce data manipulation and visualization. Upon successful completion of this course, students will have demonstrated the ability to:

1. Perform data import and pre-processing of raw data: inspection, cleaning, verification, and reporting.
2. Process disparate multi-dimensional data using selection, filtering, rearranging, grouping, transformation.
3. Understand the fundamentals of effective tables and graphs for diverse data types.
4. Demonstrate understanding of good and ethical data practices.
5. Create insightful data graphics using contemporary software tools and packages and appropriate design choices.
6. Demonstrate basic knowledge of importing, geoprocessing and mapping spatial data.
7. Understand ethical implications of data manipulation and visualization.

Lecture Content:

Week	Content
1	Discussion of data structures and algorithms for data.
2	Data manipulation via modelling
3	Data validation: do not just debug your code, debug your data.
4	Selection of subsets, representativity as a quality.
5	Data summaries, how to tell people what's in there.
6	Graphics, techniques and visualizations.
7	Ethics of data – gathering and presenting. Diversity and inclusion.
8	Data visualization, more techniques and graphics.
9	Synthetic data, generating it, recognizing your assumptions.
10	Data reduction techniques for big data and its visualization.
11	Spatially structured data – another dimension.
12	Student presentations of final project progress reports.

Labs:

No labs

Course Assignments and Final Project:

Assignments consist of small projects due at the end of weeks three, six, and nine, a final project proposal due in week 5, a paper analysis due in weeks 7 and week 11, and a final project due the Wednesday of the second week of finals. Assessments due in a week are due by midnight the Friday of that week.

Final project due date: Electronic submission as a PDF (no Word documents!) or paper submission by mail are acceptable. The final project is due on Wednesday, December 15th.

Course Resources

Required Texts: None

Recommended Texts: Given on Courselink

Other Resources: Paper listing on courselink

Course Policies

Grading Policies

Performance on the small projects is worth 40% of the grade, the final project proposal is worth 5%, the paper analysis are worth 35%, and the score on the final project is worth 20% of the grade.

Late assignments not accompanied by a reasonable medical or personal excuse are docked 5% per day late and are not accepted more than a week late.

Course Policy on Group Work:

Students are encouraged to work together but must each write up individually the material they turn in. Work on the final project may be done in consultation with other students or with faculty but the final project must be completely written by the student.

Course Policy regarding use of electronic devices and recording of lectures

Electronic recording of classes is expressly forbidden without consent of the instructor. When recordings are permitted they are solely for the use of the authorized student and may not be reproduced, or transmitted to others, without the express written consent of the instructor.

Policy Appendix

University Policies

Academic Accommodation of Religious Obligations

If you are unable to complete a course requirement due to religious obligations, please let the instructor know within the first two weeks of class. See the academic calendar for more information:

<https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-accomrelig.shtml>

Academic Consideration

When you find yourself unable to meet an in-course requirement because of illness or compassionate reasons, please advise the course instructor in writing, with your name, id, and e-mail contact. See the academic calendar for information on regulations and procedures for Academic Consideration:

<http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-ac.shtml>

Academic Misconduct

The University of Guelph is committed to upholding the highest standards of academic integrity and it is the responsibility of all members of the University community, faculty, staff, and students to be aware of what constitutes academic misconduct and to do as much as possible to prevent academic offences from occurring. University of Guelph students have the responsibility of abiding by the University's policy on academic misconduct regardless of their location of study; faculty, staff and students have the responsibility of supporting an environment that discourages misconduct. Students need to remain aware that instructors have access to and the right to use electronic and other means of detection. Please note: Whether or not a student intended to commit academic misconduct is not relevant for a finding of guilt. Hurried or careless submission of assignments does not excuse students from responsibility for verifying the academic integrity of their work before submitting it. Students who are in any doubt as to whether an action on their part could be construed as an academic offence should consult with a faculty member or faculty advisor.

The Academic Misconduct Policy is detailed in the Undergraduate Calendar:

<https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c03/index.shtml>

Accessibility

The University of Guelph is committed to creating a barrier-free environment. Providing services for students is a shared responsibility among students, faculty and administrators. This relationship is based on respect of individual rights, the dignity of the individual and the University community's shared commitment to an open and supportive learning environment. Students requiring service or accommodation, whether due to an identified, ongoing disability or a short-term disability should contact the Student Accessibility Services (SAS) as soon as possible. For more information, contact SAS at 519-824-4120 ext. 56208 or email sas@uoguelph.ca or see the website: <http://www.uoguelph.ca/sas>

Drop date

The last day to drop the class is the last day of classes.

E-mail Communication: As per university regulations, all students are required to check their uoguelph.ca e-mail account regularly: e-mail is the official route of communication between the University and its students.

When You Cannot Meet a Course Requirement

When you find yourself unable to meet an in-course requirement because of illness or compassionate reasons, please advise the course instructor (or designated person, such as a teaching assistant) in writing, with your name, id number, and e-mail contact. See the undergraduate calendar for information on regulations and procedures for Academic Consideration.

Copies of out-of-class assignments

Keep paper and/or other reliable back-up copies of all out-of-class assignments: you may be asked to resubmit

work at any time.

Inappropriate online behaviour will not be tolerated.

Examples of inappropriate online behaviour include:

- Posting inflammatory messages about your instructor or fellow students
- Using obscene or offensive language online
- Copying or presenting someone else's work as your own
- Adapting information from the Internet without using proper citations or references
- Buying or selling term papers or assignments
- Posting or selling course materials to course notes websites
- Having someone else complete your quiz or completing a quiz for/with another student
- Making false claims about lost quiz answers or other assignment submissions
- Threatening or harassing a student or instructor online
- Discriminating against fellow students, instructors or TAs
- Using the course website to promote profit-driven products or services
- Attempting to compromise the security or functionality of the learning management system
- Sharing your user name and password
- Recording lectures without the permission of the instructor