UNIVERSITY of GUELPH

DATA*6200 Data Manipulation and Visualization
Fall 2022

0   COURSE PREAMBLE

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 the Wellness & Education Promotion Centre (inside the J.T. Powell Building) 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-campuses/how-u-of-g-is-preparing-for-your-safe-return/
https://news.uoguelph.ca/return-to-campuses/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).

1 INSTRUCTOR

Instructor: Ayesha Ali

Office: MACN 509

Phone extension: 53896

E-mail: aali@uoguelph.ca (best way to reach me)

2 AIMS & OBJECTIVES

2.1 Calendar Description

This course provides a hands-on introduction to the manipulation and visualization of complex data sets using a programming language. Efficient techniques for importing and exporting data in various formats, data acquisition, data integrity, and good analysis practices are discussed. Several programming tools and libraries are introduced to restructure, transform and fuse disparate data types for visualization and data summaries in table format. Basics of manipulating space-time data is covered.

2.2 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 analyst's 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 data storytelling: visual thinking and visualization design including focus, data, aesthetics, scale, statistics, and facets
- Tools and techniques for visualization 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

2.3   Learning Outcomes

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. Understand implications of data manipulation and visualization.

6. Create insightful data graphics using contemporary software tools and packages and appropriate design choices.

7. Understand what a dashboard is and its role in knowledge translation.

8. Demonstrate basics of importing, geoprocessing and mapping of spatial data.

9. Generate basic data graphics for spatial data over time.

3 TEACHING AND LEARNING ACTIVITIES

3.1 Timetable

Lectures: Tu/Th 2:30—3 :50pm EST in MACS 301
Office Hours: See Courselink; office hours may change.

Due to the Fall Study Break, lecture for Tuesday, October 11 is rescheduled for Thurs. December 1

3.2 Course Topics and Schedule

Week	Topic
Week 1	Review of data types and data structures; Importing data; Ethics of data extraction and manipulation.
Week 2-4	Data manipulation, workflow, inspection and verification; Selection of data subsets; Diversity, inclusion.
Weeks 3-5	Data summaries and Grammar of Graphics; Ethical concerns – privacy, optics and implications.
Weeks 6-7	Data visualization, generation and interpretation; Introduction to dashboards.
Weeks 8-10	Data reduction techniques for visualization of big data.
Week 11	Spatially structured data handling; basic visualization of spatial-temporal data (e.g. spatial data over time slices).
Weeks 12	Student presentations of final project progress reports.

4 LEARNING RESOURCES

4.1 Course Website

Course material, news, announcements, and grades will be regularly posted to the DATA*6100 Courselink site. You are responsible for checking the site regularly.

4.2 Readings and Resources

Required Text: None

Recommended Texts:
  1. Hadley Wickham and Garrett Grolemund. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, 1st Ed., O'reilly Media Inc., 2017.

2. Claus O. Wilke, *Fundamentals of Data Visualization*, O'reilly Media Inc., 2019.

3. Other relevant readings may/will be provided throughout the semester.

5 ASSESSMENT

5.1 Dates and Distribution

Assignment	Due Date	Weighting	Learning Outcome(s) Assessed
Assignment 1	Week 3	15%	1,2,8
Assignment 2	Week 6	15%	2,3,4,8
Case Study 1	Weeks 6-8	10%	2,5,8
Case Study 2	Weeks 9-11	10%	2,6,8
Project − Proposal	Week 9	5%	1-7
Project − Overview, preliminaries	Week 12	7%	1-7
Project − Final project submission	Week 14	38%	1-8

5.2 Assessment Descriptions

Assignments provide hands-on experience with analyzing data using relevant software and critically evaluating model performance while **case studies** provide opportunities for group work, literature review, analysis, and presentation. The **final project** will allow students to apply the techniques learned in the course to real world problems, but with clear stop points for feedback, through the **proposal** and **project overview** assessments. **Final projects** will be in the form of formal statistical reports that include a section on data ethics. **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.

Assignments and project proposal and overview/preliminaries will be due on Thursdays at 11:59pm EST in weeks 3, 6, 9 and 12, respectively. Case Study presentations will be scheduled for during class time in weeks 8 and 11, with specific dates and times depending on the number of groups. Final presentation slides must be submitted by 11:59pm EST on Wednesdays in weeks 8 and 11. The final project will be due Tuesday, December 13, 2022 at 11:59pm.

5.3 Course Grading Policies

Passing grade: In order to pass the course, students must obtain a grade of 65% or higher on the total mark of all assessments.

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.

Accommodation of Religious Obligations: If you are unable to meet an in-course requirement due to religious obligations, please email the course instructor within two weeks of the start of the semester to make alternate arrangements. See the graduate calendar for information on regulations and procedures for Academic Accommodation of Religious Obligations: [https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/sec_d0e2228.shtml]

---

6   UNIVERSITY STATEMENTS

6.1   Email Communication

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

6.2   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#, and e-mail contact. See the Undergraduate Calendar for information on regulations and procedures for Academic Consideration.

Academic Consideration, Appeals and Petitions

[https://calendar.uoguelph.ca/undergraduate-calendar/undergraduate-degree-regulations-procedures/academic-consideration-appeals-petitions/]

6.3   Drop Date

Courses that are one semester long must be dropped by the end of the last class day; two-semester courses must be dropped by the last day of class in the second semester. The regulations and procedures for changing graduate course registration are available in the Undergraduate and Graduate Calendars.

Undergraduate Calendar – Dropping Courses
[https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-drop.shtml]
Graduate Calendar – Registration Changes
[https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/genreg-reg-regchg.shtml]

6.4   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.

6.5   Accessibility

The University promotes the full participation of students who experience disabilities in their academic programs. To that end, the provision of academic accommodation is a shared responsibility between the University and the student.

When accommodations are needed, the student is required to first register with Student Accessibility Services (SAS). Documentation to substantiate the existence of a disability is required, however, interim accommodations may be possible while that process is underway.

Accommodations are available for both permanent and temporary disabilities. It should be noted that common illnesses such as a cold or the flu do not constitute a disability.

Use of the SAS Exam Centre requires students to make a booking at least 14 days in advance, and no later than November 1 (fall), March 1 (winter) or July 1 (summer). Similarly, new or changed accommodations for online quizzes, tests and exams must be approved at least a week ahead of time.

More information: www.uoguelph.ca/sas

6.6   Academic Integrity

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.

Undergraduate Calendar – Academic Misconduct

[https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-amisconduct.shtml]

Graduate Calendar – Academic Misconduct
[https://www.uoguelph.ca/registrar/calendars/graduate/2018-2019/genreg/sec_d0e2632.shtml]

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 username and password
- Recording lectures without the permission of the instructor

6.7   Recording of Materials

Presentations that are made in relation to course work—including lectures—cannot be recorded or copied without the permission of the presenter, whether the instructor, a student, or guest lecturer. Material recorded with permission is restricted to use for that course unless further permission is granted.

6.8   Other University Resources

The Academic Calendars [https://www.uoguelph.ca/academics/calendars] are the source of information about the University of Guelph's procedures, policies, and regulations that apply to undergraduate, graduate, and diploma programs.