

DSCI_V 200 Syllabus

Term 2025 Winter 2: January - April 2026

Course info

Instructors:

Katie Burak

Website: <https://katieburak.github.io/>

Email: kburak@stat.ubc.ca

Gabriela V. Cohen Freue

Website: <https://gcohenfr.github.io/>

Email: gcohen@stat.ubc.ca

Office hours:

See [Canvas](#) for times and locations.

Course webpage:

WWW: <https://ubc-stat.github.io/dsci-200/>

Canvas: <https://canvas.ubc.ca/courses/>

Lectures/Labs:

See [Canvas](#) for times and locations.

Prerequisite:

DSCI 100

Course objectives

The course objective is to train students in navigating the many aspects of data needed to successfully work through the data stage of the data analysis lifecycle. Students will broaden their applied statistical knowledge base and skillset required to dive into statistical and computational modelling in subsequent courses. Students will learn how to explore different layers of the data structure, consider the limitations introduced by the study design, acknowledge security and ownership features related to data, and address problems encountered in the data at hand. The course also aims to demonstrate how simulation studies can be carried out to examine properties of estimators and algorithms (i.e., beyond the data stage). The course reinforces and refines the computational skills (i.e., writing computer scripts), tools and resources learned in DSCI_V 100 to analyze data, as well as to draw and communicate appropriate conclusions.

Learning outcomes

- Constructively reflect on how the data was collected for the statistical question being asked. Identify improvements and, where improvements cannot be made, clearly discuss the limitations of the study design with regards to the conclusions that can be drawn for the question being asked.
- Determine the data acquisition strategy needed for a given data source and use common data science tools to write reproducible computer scripts to read the data into the chosen programming language from that data source.
- Identify when data simulation is a useful technique for assessing an analysis method, as well as plan and carryout appropriate simulations based on the task at hand.
- Identify outliers and data anomalies, justify and apply strategies for managing the these, and reflect on the consequences with regards to the conclusions of the chosen method.
- Identify when and why data are missing, justify and apply strategies for managing the missing data, and reflect on the consequences with regards to the conclusions of the chosen method.
- Plan and carry out exploratory data analysis using statistical and visualization techniques for the purpose of generating hypotheses and planning a further valid statistical analysis.
- Appraise the impact of measurement scales, units, significant digits, and measurement error on data analysis, as well as using simple methods to mitigate these effects on results of the analysis.
- Evaluate and justify the data privacy needs for a given data analysis, and where needed, choose and apply an appropriate data privacy technique. Reflect on the consequences with regards to the conclusions of the chosen method.
- Evaluate and justify who owns the data for a given case.

Assessments

The course will have exams, worksheets, iClicker questions and case studies for assessments. All official due dates are available on Canvas.

Exams

This course has one midterm and one final exam.

Note: Instructors reserve the rights to scale grades in order to maintain equity among sections according the UBC campus wide policies and regulations.

Worksheets

- All official due dates are available on Canvas.
- To open an assignment, click the link (e.g. [worksheet_EDA](#)) from Canvas.
- To submit your assignment, just make sure your work is saved on our server (File -> Save Notebook to be sure).
- At the deadline, our server will automatically snapshot your work.

- You must access the lecture and tutorial worksheets through our Canvas course page (as opposed to the worksheets publicly available via Github), otherwise your worksheets may not be marked!

iClicker

During each lecture there will be iClicker questions to help check your understanding of the course material. iClicker grades will be based on participation. You must attend the section you are registered in. It is your responsibility to make sure that the student ID and name associated with your iClicker account matches the Canvas gradebook. If you need help connecting to iClicker please see iClicker Cloud Student Guide.

Students who participate in at least 75% of iClicker assessments will receive full iClicker credit. Students who participate in at least 50% but less than 75% will receive half credit, and students who participate in under 50% will receive no credit. This participation-based grading scheme is designed to accommodate unforeseen absences or personal matters that may arise during the semester.

Case studies

The case studies act as mini-projects and will provide additional practice with the data science skills we teach in the class. It is an extension to the worksheets by providing little to no prompts and scaffolding code. You will only be able to use the dataset(s) we have provided you to use. Further details regarding the case studies will be announced during the term.

Course breakdown

Deliverable	Percent Grade
Worksheets	7%
iClicker	2%
Case studies	15%
Midterm	25%
Final	50%
Bonus regrade percent	1%

If a student's score on the final exam is higher than the student's score on the midterm exam, 10% of the midterm's weight will be transferred to the final exam when calculating the course grade. In this case, the midterm will count for 15% of the course grade and the final exam will count for 60%. If the student's final exam score is not higher than the midterm exam score, the original exam weights remain in effect, with the midterm counting for 25% and the final exam counting for 50% of the final grade. All other components of the course grade remain unchanged.

Policies

Code of Conduct

All participants in our course and communications are expected to show respect and courtesy to others. To creating a friendly and respectful place for learning, teaching and contributing, you are expected to read and follow the DSCI 200 Code of Conduct.

Late Registration

Students who register for the class late have 1 week from their registration date on Canvas to complete all prior assignments.

Late Assignments / Absences

For examinations, students must be present at the invigilation venue (in class, examination centre, etc) to take exams; otherwise they will be considered to have missed the exam and will be assigned a grade of zero.

Students who will miss an exam must provide a self-declaration of academic concession prior to the exam (see Canvas homepage for the academic concession form) and make arrangements with the instructor. Failing to present a declaration within a reasonable timeframe before the exam will result in a grade of zero.

There will be no extensions for the worksheets; late assignments will receive a grade of zero. Instead, the lowest worksheet grade from the semester will be dropped. This policy is meant to cover illness/unexpectancies. The worksheet that you were not able to complete before the deadline will be covered by this policy. However, if you have extenuating circumstances and require further accommodations for subsequent requests, please contact the course instructor with supporting documents, and we will deal with them case by case.

Students who miss a lecture for their registered section will receive an iClicker grade of 0 for that lecture. There will be no make ups or accommodations for missing i-clickers. Instead, we will drop the 2 lowest grades on iClicker lectures to accommodate late registration and/or unforeseeable events.

For all other assignments including the case studies, a late submission will receive a 50% penalty.

Autograder Policy

Many of the questions in assignments are graded automatically by software. The grading computer has exactly the same hardware setup as the server that students work on. No assignment, when completed, should take longer than 5 minutes to run on the server. The

autograder will automatically stop (time out) for each student assignment after a maximum of 5 minutes; any ungraded questions at that point will receive a score of 0.

Students are responsible for making sure their assignments are reproducible, and run from beginning to end on the autograding computer. In particular, please ensure that any data that needs to be downloaded is done so by the assignment notebook with the correct filename to the correct folder. A common mistake is to manually download data when working on the assignment, making the autograder unable to find the data and often resulting in an assignment grade of 0.

In short: whatever grade the autograder returns after 5 minutes (assuming the teaching team did not make an error) is the grade that will be assigned.

Re-grading

Students may review their midterm and final exams during the designated CBTF exam review windows. During this period, you may submit an issue if you believe a grading error occurred. All regrade requests should be reasonable and supported by clear justification.

To account for minor grading errors throughout the course, every student will get a **bonus of one percentage point** at the end of the semester. If you have a question about your score on a case study assignment or believe a grading error occurred, you may submit a regrade request through Gradescope. We only accept worksheet regrade requests for major errors in grading. If you think there is a grading error of more than 10% on a single worksheet, please submit a copy of your worksheet on Gradescope along with a regrade request. Regrade requests must be submitted within one week of the grade release date. After that deadline, the grade is final.

Device/Browser

Students are responsible for using a device and browser compatible with all functionality of Canvas. Chrome or Firefox browsers are recommended.

Missed Final Exam

Students who miss the final exam must report to their faculty advising office within 72 hours of the missed exam, and must supply supporting documentation. Only your faculty advising office can grant deferred standing in a course. You must also notify your instructor prior to (if possible) or immediately after the exam. Your instructor will let you know when you are expected to write your deferred exam. Deferred exams will **ONLY** be provided to students who have applied for and received deferred standing from their faculty.

Academic Concession Policy

Please see UBC's concession policy for detailed information on dealing with missed coursework and exams under circumstances of an acute and unanticipated nature.

See our Canvas homepage for the academic concession form.

Academic Integrity

The academic enterprise is founded on honesty, civility, and integrity. As members of this enterprise, all students are expected to know, understand, and follow the codes of conduct regarding academic integrity. At the most basic level, this means submitting only original work done by you and acknowledging all sources of information or ideas and attributing them to others as required. This also means you should not cheat, copy, or mislead others about what is your work. Violations of academic integrity (i.e., misconduct) lead to the breakdown of the academic enterprise, and therefore serious consequences arise and harsh sanctions are imposed. For example, incidences of plagiarism or cheating may result in a mark of zero on the assignment or exam and more serious consequences may apply if the matter is referred to the President's Advisory Committee on Student Discipline. Careful records are kept in order to monitor and prevent recurrences.

A more detailed description of academic integrity, including the University's policies and procedures, may be found in the Academic Calendar at <http://calendar.ubc.ca/vancouver/index.cfm?tree=3,54,111,0>.

Use of GenAI

Generative AI tools, such as large language models like ChatGPT, Claude, or Gemini, can be useful for exploring ideas, reviewing concepts, and troubleshooting problems. In this course, you may use these tools for those purposes, provided that any use of AI assistance in your coursework is properly cited. However, it is not permitted to complete assignments by copying and pasting AI-generated responses. All submitted work must reflect your own understanding and original effort.

Plagiarism

Students must correctly cite any code or text that has been authored by someone else or by the student themselves for other assignments. Cases of plagiarism may include, but are not limited to:

- the reproduction (copying and pasting) of code or text with none or minimal reformatting (e.g., changing the name of the variables)
- the translation of an algorithm or a script from a language to another
- the generation of code and/or text by automatic code-generation software or large language model

An “adequate acknowledgement” requires a detailed identification of the (parts of the) code or text reused and a full citation of the original source code that has been reused.

The above attribution policy applies only to assignments. No code or text may be copied (with or without attribution) from any source during an exam. Answers must always be in your own words. At a minimum, copying will result in a grade of 0 for the related assignment.

Repeated plagiarism of any form could result in larger penalties, including failure of the course.

Resources

For additional information, please check out these useful student resources, the survival tips from your TAs, and the Frequently Asked Questions. If you want to use any of this material elsewhere, please read the license.

Attribution

Parts of this syllabus have been copied and derived from the [DSCI 100 syllabus](#) and the [UBC MDS Policies](#).