STAT 301 – Statistical Modelling for Data Science
2024 Winter Term 1 (September – December, 2024)

Explanatory and predictive data analysis with multiple explanatory variables. Choosing the right methods to apply based on the statistical question and data at hand. Trade-offs between model-based and non-model based approaches. Emphasis placed on case studies and real data sets, as well as reproducible and transparent workflows when writing computer scripts for analysis and reports.

Prerequisites: STAT 201 and one of MATH 100, MATH 102, MATH 104, MATH 110, MATH 120, MATH 180, MATH 184, SCIE 001.

Teaching & Learning Activities: The course is structured in weekly lectures and tutorials. The lectures will be expositive with the use of in-class activities. Students are expected to attend lectures and tutorials. This course will have plenty of synchronous activities that students must work on during the lectures and tutorials. Students will work on activities in Jupyter Notebooks and iClickers.. Students need to access to a computer. If a student does not have their own laptop or chromebook, students may be able to loan a laptop from the UBC library

Instructor: Dr. W. A. Lasantha Premarathna (Email: wpremara@stat.ubc.ca).

- Please use the email only for personal matters that you would want to discuss with the instructor.
- If you send me an email, make sure to include **STAT 301** in the subject line.
- Please use **office hours** and **Piazza Discussion Board** for questions regarding worksheet & tutorial assignment problems, projects and class note examples etc.

Class Room: Please refer the Canvas Course page for the class room information

Lectures & Tutorials: in-person

Note: If the university is closed due to extreme weather conditions (e.g., due to heavy snow), lectures will run via Zoom or recorded lectures will be posted in the canvas page. Detailed information on such days will be posted through canvas announcements. Please check the canvas course page regularly.

Instructor Office Hours:

- **Online office hour:** 9:30am – 10:30am on Friday (Zoom link can be found through the zoom tab)
- **in-person:** Just after the class on Tuesday & Thursday. (around 30 minutes each day). I will leave if there are no students to ask questions.

Teaching assistants & TA office hours: Please check the Canvas course page.

Course Website: canvas.ubc.ca

Please check the **Canvas** website regularly to keep up-to-date with the course. Everything you need will be available through Canvas and you should get familiar with all the tabs as soon as possible.
If you have any problems related to technical issues, please use **?Help** (see the left side menu in the Canvas course page) to report the problem or to contact IT service.

Learning Outcomes

By the end of the course, students are expected to be able to:

- Describe real-world examples of explanatory modelling (e.g. A/B testing optimization & regression with variable selection) and predictive modelling problems.
- Explain the trade-offs between model-based and non-model based approaches, and describe situations where each might be the preferred approach.
- Explain the difference between creating models for explanation vs prediction, in the context of both how you choose and evaluate models as well as how you interpret the results.
- Choose & apply a suitable method (e.g., regression, GLM's, sample size estimation, controlling for multiple testing, peeking, bandit algorithms, variable selection, model diagnostics) based on the statistical question and data at hand. Discuss the advantages and disadvantages of different methods that may be suitable for a given problem.
- Correctly interpret computer output when performing the statistical analyses presented in this course, in the context of the statistical question being asked and the audience being reported to.
- Identify the assumptions / conditions required for each method to produce reliable results. Choose techniques to check (or at least be able to falsify) those assumptions. Discuss the consequence(s) of mapping the wrong methods to the question and/or data type.

Assessments

Most weeks there will be two assignments: (1) a worksheet; and (2) a tutorial. These assignments will be due every Monday at 11:59pm. The worksheets are fully autograded with visible tests to help you identify points that need more clarification. Therefore, reach out to the teaching team if you don't understand why you are getting an answer wrong in the worksheet. On the other hand, the tutorials are not fully autograded. You can access the assignments through Canvas.

To submit your assignment, make sure your work is saved **on our server** (i.e., accessed using the link from Canvas) before the deadline. Our server will automatically snapshot at the due date/time. Also, **please DO NOT rename the assignments files.**

Course Assessment:

Assessment	Date	Percentage
Worksheets	refer the Canvas Course page	4%
Tutorials	refer the Canvas Course page	8%
Project	Final Report Due: December 6	18%
Midterm	Thursday, October 24 (12:30pm – 1:50pm) at the regular class room	25%
Final Exam (you must pass the final to pass the course)	To be scheduled by Classroom Services. Exam schedule is released at least 3 weeks before the final day of classes. **Exam period:** December 10 - 21	45%
Bonus points		1%

Note: Please refer the Canvas Course page for deadlines

Lectures' Worksheets: fully auto-graded with visible tests to help you identify points that need more clarification.

Tutorials' Worksheets : only a few exercises will have visible tests.

Group Project: A project that you will work with your group throughout the term. Details about the Group Project will be made available to you on Canvas. Although this is a group project, some related assignments will be submitted individually. You can (and are encouraged to) discuss it with your group members. However, *every* student will submit their own assignment and will receive an individual grade.

Exams: All the exams will be on Canvas with lockdown browser. You will be able to have 1 letter-size page (one-sided) cheatsheet (you can write or print). You **are not allowed** to access any webpage or files in your computer or other electronic devices. The types of questions can vary: reasoning, multiple-choice, multiple-answer, dropdown, true or false. Although most questions will be about the content, you can expect some coding questions. The coding question will not be overly complicated, and we will only check your familiarity with the main functions and packages we use in the course. We are not trying to test your software development skills here!!! Please don't spend energy trying to memorize everything. If you have done the worksheets and tutorial, this should not be a problem for you.

Bonus points: Students that have answered the most statistics-related questions in Piazza in a way that explains concepts well but does not reveal the answer to an assignment, lab, or webwork question will get a bonus 1% added to their grade. When you answer question, teaching team endorse your answers as "good answer". I add this 1% if you have more than 10 Endorsed Answers".

Course Policies

Policy regarding missing the midterm:

- There will be no make-up exam.
- Students who miss an exam should notify the instructor prior to (if possible) or immediately after the exam to request an Academic Concession. Students must supply a supporting document (for example, a doctor's note will be sufficient in case of a medical emergency) within 3 days of the day of exam. Failing to contact the instructor may result in a grade of zero on the Midterm.
- If your request is approved by the instructor, your midterm weight will be moved to the final exam.

Deferred Exam Policy if you miss the Final Exam:

- The policy (UBC policy) is that students who miss the final exam MUST report to their faculty advising office within 48 hours to apply for deferred standing. They must also notify the instructor to receive instructions as to when they will write their deferred final. But they will not be granted a deferred final unless they are granted deferred standing by their faculty advising office.

Late/Absence

- Regular attendance to lectures and tutorials is expected of students. Students who are unavoidably absent from numerous classes because of illness or other reasons should inform the instructor(s) of the course as soon as possible, preferably prior to the start of the lecture/tutorial.
- Late submissions will receive a grade of 0. But note that:
    - At the end of the semester, one lecture worksheet and one tutorial worksheet will be dropped.
    - If you have extenuating circumstances and need concessions **beyond** dropping one worksheet and one tutorial, please contact the instructor. Note that concessions will only be considered after the "free drop" has occurred.

Autograder Policy

Many of the questions in assignments are graded automatically by software. The grading computer has the same hardware setup as the server 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. Furthermore, students are responsible for making sure their assignments are reproducible and run from beginning to end on the auto-grading computer. 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.

Tip: when you're done with the assignment, click "Restart and Run All" and check that the visible tests are all working and that your notebook runs in less than 5 minutes.

Regrading

If you have concerns about the way your work was graded, please open a request within one week of having the grade returned to you. After this one-week window, we may deny your request for re-evaluation. Also, please keep in mind that your grade may go up or down due to re-grading. To open a regrade request, please follow the steps below:

1. Go to Piazza and click on New post.
2. In Post Type, select Question.
3. Make the post private to instructors and TAs only. In Post to select Individual Students(s)/Instructor(s). A text box will appear, where you must type Instructors.
4. In Select Folder(s) select the folder regrading
5. In Summary say the Assignment you want to be regraded, followed by the question and your name and student number. For example, lab 3 -> Q3 – Your Name (Your student Number)
6. Provide a brief reason for why the regrade is needed.
7. The TAs will see the request and will take a look at the assignment. If necessary, they will involve the instructors. Finally, once the TA is finished reassessing the assignment:
    - If the grade deserves more marks: the TA will update the mark on Canvas and comment on the question so everyone can see that the question has been addressed.
    - If your grade goes down or stays the same: the TA will answer the post on Piazza, giving the student a reason for their final decision

iClicker cloud:

- We will be using iClicker Cloud in lectures. iClicker Cloud is a response system that allows you to use your own computer or mobile device to respond to questions posed by instructors during class. You need to set up an iClicker Cloud account and add STAT 301 as a course to this account. To do so, please follow https://lthub.ubc.ca/guides/iclicker-cloud-student-guide for details.

Piazza Discussion Board:

- **Access Piazza:** Please go to **"Piazza"** in the left menu in the Canvas course page and it will open in a new window. Then you can sign up for the class page. You can use "Piazza Discussion Board" to post your questions and also to provide answers/hints to the questions posted there. This is where you can discuss ideas, strategies, and resources for solving the problems with your classmates. We, the teaching team will monitor the questions in the piazza page and post answers. But do not expect we will answer all your questions posted in Piazza page with detailed solutions. If you need more clarification, it's always better to contact TAs or me during our office hours. If you have any problems or feedback for the developers, email team@piazza.com.

Note: Students that have answered the most statistics-related questions in a way that explains concepts well but does not reveal the answer to an assignment, lab, or webwork question will get a bonus 1% added to their grade. When you answer question, teaching team endorse your answers as "good answer". I add this 1% if you have more than 10 "Endorsed Answers"

Academic Integrity:

- Discussion of ideas leaned in class is encourage (with other students, TAs or the instructor).
  This helps the leaning process. But individual work turned in by each student should be your own work. Do not copy or paraphrase solutions from other students or from other sources. Do Not provide your solutions to another student. Failure to comply with these rules will result in an automatic 0 for your work, and additional academic penalties.

For more information, please see

- Academic Honesty and Standards: https://vancouver.calendar.ubc.ca/campus-wide-policies-and-regulations/academic-honesty-and-standards
- Academic Integrity: https://academicintegrity.ubc.ca/about-academic-integrity/
- Academic Misconduct: https://academicintegrity.ubc.ca/regulation-process/academic-misconduct/
- Resources and Support: https://academicintegrity.ubc.ca/resources/

Reach Out for Success

- University students often encounter setbacks from time to time that can impact academic performance. Discuss your situation with your instructor or an academic advisor. Learn about how you can plan for success at: www.students.ubc.ca
- For help addressing mental or physical health concerns, including seeing a UBC counsellor or doctor, visit: https://students.ubc.ca/health-wellness

UBC policies and resources to support student success:

- UBC provides resources to support student learning and to maintain healthy lifestyles but recognizes that sometimes crises arise and so there are additional resources to access including those for survivors of sexual violence. UBC values respect for the person and ideas of all members of the academic community. Harassment and discrimination are not tolerated nor is suppression of academic freedom. UBC provides appropriate accommodation for students with disabilities and for religious and cultural observances. UBC values academic honesty and students are expected to acknowledge the ideas generated by others and to uphold the highest academic standards in all of their actions. Details of the policies and how to access support are available at https://senate.ubc.ca/policies-resources-support-student-success.

Land acknowledgement: *We acknowledge that the UBC Vancouver campus is situated within the traditional, ancestral and unceded territory of the xʷməθkʷəy̓əm (Musqueam).*