

## SS 4850G/9850: Advanced Data Analysis

### Course Outline

#### 1. Course Information

**Course Information**

Term: Winter 2025

Location: VAC-100

Lectures: Monday, Wednesday, Friday 12:30 – 1:20

**List of Prerequisites:**

-- For SS4850G: A minimum mark of 60% in both Statistical Sciences 3843A/B and Statistical Sciences 3859A/B.

Unless you have either the requisites for this course or written special permission from your Dean to enroll in it, you may be removed from this course, and it will be deleted from your record. This decision may not be appealed. You will receive no adjustment to your fees in the event that you are dropped from a course for failing to have the necessary prerequisites.

#### 2. Instructor Information

Instructors	Email	Office	Phone	Office Hours
Dr. Guowen Huang	<a href="mailto:ghuan23@uwo.ca">ghuan23@uwo.ca</a>	WSC 270	NA	MF 10:00 – 11:00 F 14:00 – 15:00

Students must use their Western (@uwo.ca) email addresses when contacting their instructors and have SS4850G or SS9850 in the subject line.

**Communication:**

Please use the Discussion board (you can find it by clicking on Communications tab) on the OWL site (<https://westernu.brightspace.com/>) to post questions about course material (e.g., lecture, reading, etc). This helps other students who may have the same question, and I encourage you to help each other if you can. The Discussion board will be monitored on a regular basis and the instructor or TAs will interject with corrections or responses as necessary. As this is an open Forum, please be respectful of your peers, instructor(s), and TAs. Derogatory, discriminatory, or otherwise inappropriate language or topics will be removed and dealt with at the instructor's discretion.

## Instructor Policies

I expect respect and inclusivity in our classroom. To promote this, I have several policies that you should follow:

- Please do not be late. If you must be, please enter quietly and choose a seat closest to the door to limit the disruption.
- Put your phones on silently and refrain from being on them. If you must answer your phone, please do so quietly and immediately leave the room. I understand there are sometimes emergencies, but this should not disrupt other students.
- Treat others as you would like to be treated. When speaking for the first time in a new group, introduce yourself with your name, and pronouns if you feel comfortable. Do not judge others for their work in this class, their questions, or anything they choose to share.
- When others are speaking, whether it be instructor or student, listen and do not speak over them.
- You can expect a response to a message to the instructor within approximately 48 hours during the work week (during busy times, it may take a little longer). Note that messages will not be answered within the 24-hour period before exams or project deadlines, nor can I guarantee responses over weekends/holiday.

## 3. Course Syllabus, Schedule, Delivery Mode

### Course Description

Modern methods of data analysis including ridge regression, lasso, linear discriminant analysis, nonparametric regression, bootstrap, EM algorithm, Bayesian inference, classification trees, neural networks, dimensionality reduction techniques, and clustering.

### Key Sessional Dates

Classes begin: January 6, 2025

Reading Week: February 15– 23, 2025

Classes end: April 4, 2025

Exam period: April 7 – 30, 2025

### Tentative Course Schedule

Week	Date	Topics	Reminders
1	Jan 6-10	- Statistical learning - Supervised vs unsupervised learning	
2	Jan 13-17	Supervised Learning - Ridge regression and LASSO	
3	Jan 20-24	- Classification: Discriminant Analysis, Logistic regression, Naïve Bayes, kNN - ROC	
4	Jan 27-31	- Mitigating class imbalance	Jan 31 – Assignment 1
5	Feb 3-7	- Nonparametric regression	
6	Feb 10-14	- Model assessment	
7	Feb 17-21	Reading week	
8	Feb 24-28	- Overview of tree-based methods and neural networks	Midterm test – Feb 27, 19:00-21:00
9	Mar 3-7	Inference - Bootstrap versus maximum likelihood - EM algorithm	March 7 – Assignment 2
10	Mar 10-14	- Introduction to Bayesian Inference	

11	Mar 17-21	- More on Bayesian Inference	First part of final project is due Friday Mar 21
12	Mar 24-28	Unsupervised Learning - Clustering: K-means, hierarchical clustering	
13	Mar 31- Apr 4	- Dimensionality reduction techniques: PCA, NMF, MDS	Assignment 3 – April 7 Final project due April 11

## 4. Course Materials

Chapters and sections from the following books (all available online at UWO library):

- The Elements of Statistical Learning – also available at author’s webpage: <https://web.stanford.edu/~hastie/ElemStatLearn/>
- An Introduction to Statistical Learning with Applications in R – also available at author’s webpage: <https://trevorhastie.github.io/ISLR/>
- Kuhn, Max, and Kjell Johnson. Applied predictive modeling. New York: Springer, 2013
- Hoff, Peter D. A first course in Bayesian statistical methods. New York: Springer, 2009

This course requires the use of **R**, **RStudio**, and **RMarkdown** for data analysis and report generation. All assignments will involve coding in R and producing technical reports using RMarkdown. Students are expected to have prior knowledge of these tools or be willing to learn them as part of the course. Please ensure that the latest version of R is installed on your computer along with RStudio, which can be downloaded here: <https://posit.co/download/rstudio-desktop/>.

### OWL

Students are responsible for checking the course OWL site (<http://owl.uwo.ca>) on a regular basis for news and updates. This is the primary method by which information will be disseminated to all students in the class.

All course material will be posted to OWL: <http://owl.uwo.ca>.

If students need assistance with the course OWL site, they can seek support on the OWL Help page. Alternatively, they can contact the Western Technology Services Helpdesk. They can be contacted by phone at 519-661-3800 or ext. 83800.

## 5. Methods of Evaluation

### Course Evaluation Components

Component	Weight	Deadlines/Due dates
Assignment 1	5%	Jan 31, due at 11:55 pm
Assignment 2	7.5%	Mar 7, due at 11:55 pm
Assignment 3	7.5%	Apr 7, due at 11:55 pm
Midterm Test	35%	Feb 27, 19:00-21:00. <b>There will be one extra question in the midterm test for Graduate students.</b>
Final Exam: data analysis project report (individual or in pairs)	45% (10% for first part and 35% for complete final report)	- First project part (dataset description and methodology plan) is due Friday March 21 - Complete final project report is due Friday April 11

iClicker might be used for in class feedback and non-graded assessments

## Assignments

- **Assignments will be available on the course OWL site. However, you will not submit your solutions to OWL. Instead, assignments must be submitted through [Gradescope](#),** an online collaborative grading system. It is your responsibility to make sure that your assignment is successfully uploaded and legible. Submissions that cannot be read by the grader will receive a grade of zero. Assignment questions must be assigned properly to each page or the submission will not be graded.
- After receiving the grades from an assignment, **students will have seven days to submit any regrade requests on that assignment.** After this seven-day period, regrade requests will NOT be accepted. Regrade requests must be made using the Gradescope tool “Regrade Request”.
- **Assignment submissions are due 11:55 pm (Eastern Time) on the due date.** Assignments that are up to 24 hours late will receive a deduction of 15% on their mark. Late assignments up to 48 hours will receive a deduction of 30% on their mark. No credit will be given for submissions beyond 48 hours of the deadline time unless valid academic accommodation is obtained.
- Each student will have **2 late coupons** worth 24 hours each that they can use at their own discretion for whatever reason (Note: I do not need to know the reason) towards their assignments. Students can use them together on one assignment for a 48-hour extension, or on separate assignments for 24-hour extensions. **No extensions will be given for any reason beyond the use of late coupons; therefore, I suggest saving them until you absolutely need them.** You do not need to tell me when you will use the coupons; they will automatically come off the first two late days.
- Solutions to assignments **will not** be posted; however, TAs will provide comments on incorrect answers using Gradescope, which will allow students to find out the correct solutions. In addition, students can ask the instructor and TAs for more details on solutions via the Regrade Request tool on Gradescope and during office hours.

## Midterm test

- There will be a **2-hour in-person closed-book midterm test scheduled for February 27, 19:00-21:00, though the date is tentative and may be adjusted if significant scheduling conflicts arise for students.** Students will be allowed one single-sided, A4 letter-sized, handwritten cheat sheet. Handwritten means pen to paper, no printing from a computer of any kind. Photocopying or scanning is also not permitted. Any violations of this requirement will result in a student not being permitted to use such a cheat sheet.

## Final exam

- The final exam consists of a project report, which can be conducted individually or in pairs (two students). The project involves a comprehensive, advanced statistical analysis of a dataset. Each student (or pair of students) will work on a different dataset. More details will be given in the first weeks of classes. Submissions of the first part and complete report are **due 11:55 pm (Eastern Time) on the due dates and must be conducted via Gradescope.** No credit will be given for submissions beyond due dates unless valid academic accommodation is obtained (see Section 6 for details on accommodation). Those students obtaining appropriate accommodation must contact the instructor immediately to set a new deadline.

## **Rounding of marks**

Across the Sciences Undergraduate Education programs, we strive to maintain high standards that reflect the effort that both students and faculty put into the teaching and learning experience during this course. All students will be treated equally and evaluated based only on their actual achievement. Final grades on this course, irrespective of the number of decimal places used in marking individual assignments and tests, will be calculated to one decimal place and rounded to the nearest integer, e.g., 74.4 becomes 74, and 74.5 becomes 75. Marks WILL NOT be bumped to the next grade or GPA, e.g. a 79 will NOT be bumped up to an 80, an 84 WILL NOT be bumped up to an 85, etc. The mark attained is the mark you achieved, and the mark assigned; requests for mark “bumping” will be denied.

## **Student Absences**

If you are unable to meet a course requirement due to illness or other serious circumstances, please follow the procedures below.

**Assignments:** The assignments have late coupons to account for any extensions needed, for any reason. No extensions will be given beyond this.

**Midterm Test & Final Exam (Undergraduate):** For medical absences, get a Student Medical Certificate (SMC) signed by a licensed medical or mental health. Get appropriate documentation for non-medical absences. Note that in all cases, students are required to contact their instructors within 24 hours of the end of the period covered to get approved unless otherwise instructed in the course outline. If the midterm test is missed and proper documentation is obtained, the student will be eligible to write the makeup test at another time selected by the instructor. If the deadline for the final project (partial or final report) is missed and proper documentation is obtained, the student must then contact the instructor immediately to set a new deadline.

For procedures on how to submit Academic Consideration requests, please see the information posted on the Office of the Registrar’s webpage:

[https://registrar.uwo.ca/academics/academic\\_considerations/](https://registrar.uwo.ca/academics/academic_considerations/)

All requests for Academic Consideration must be made within 48 hours after the assessment date or submission deadline.

**Midterm Test & Final Exam (Graduate):** You must provide valid documentation for medical-related absences or adequate supporting documentation for non-medical absences to the instructor as soon as possible and no later than 48 hours after the missed assessment.

Documentation for medical illness must include the completion of a Western Student Medical Certificate (SMC) or, where that is not possible, equivalent documentation, by a health care practitioner.

## 6. Additional Statements

### Religious Accommodation

When conflicts with a religious holiday that requires an absence from the University or prohibits certain activities, students should request an accommodation for their absence in writing to the course instructor and/or the Academic Advising office of their Faculty of Registration. This notice should be made as early as possible but not later than two weeks prior to the writing or the examination (or one week prior to the writing of the test).

Please visit the Diversity Calendars posted on our university's EDID website for the recognized religious holidays:

<https://www.edi.uwo.ca>.

### Accommodation Policies

Students with disabilities are encouraged to contact Accessible Education, which provides recommendations for accommodation based on medical documentation or psychological and cognitive testing. The policy on Academic Accommodation for Students with Disabilities can be found at:

[https://www.uwo.ca/univsec/pdf/academic\\_policies/appeals/Academic\\_Accommodation\\_disabilities.pdf](https://www.uwo.ca/univsec/pdf/academic_policies/appeals/Academic_Accommodation_disabilities.pdf).

### Academic Policies

The website for Registrar Services is <https://www.registrar.uwo.ca/>.

In accordance with policy,

[https://www.uwo.ca/univsec/pdf/policies\\_procedures/section1/mappl13.pdf](https://www.uwo.ca/univsec/pdf/policies_procedures/section1/mappl13.pdf),

the centrally administered e-mail account provided to students will be considered the individual's official university e-mail address. It is the responsibility of the account holder to ensure that e-mail received from the University at their official university address is attended to in a timely manner.

A scientific calculator will be allowed on the final exam. A graphics calculator will not be allowed.

**Scholastic offences** are taken seriously, and students are directed to read the appropriate policy, specifically, the definition of what constitutes a Scholastic Offence, at the following Web site:

[http://www.uwo.ca/univsec/pdf/academic\\_policies/appeals/scholastic\\_discipline\\_undergrad.pdf](http://www.uwo.ca/univsec/pdf/academic_policies/appeals/scholastic_discipline_undergrad.pdf).

### Support Services

Please visit the Science & Basic Medical Sciences Academic Counselling webpage for information on adding/dropping courses, academic considerations for absences, appeals, exam conflicts, and many other academic related matters: <https://www.uwo.ca/sci/counselling/>.

Students who are in emotional/mental distress should refer to Mental Health@Western (<https://uwo.ca/health/>) for a complete list of options about how to obtain help.

Western is committed to reducing incidents of gender-based and sexual violence and providing compassionate support to anyone who has gone through these traumatic events. If you have experienced sexual or gender-based violence (either recently or in the past), you will find information about support services for survivors, including emergency contacts at

[https://www.uwo.ca/health/student\\_support/survivor\\_support/get-help.html](https://www.uwo.ca/health/student_support/survivor_support/get-help.html).

To connect with a case manager or set up an appointment, please contact [support@uwo.ca](mailto:support@uwo.ca).

Please contact the course instructor if you require lecture or printed material in an alternate format or if any other arrangements can make this course more accessible to you. You may also wish to contact Accessible Education at

[http://academicsupport.uwo.ca/accessible\\_education/index.html](http://academicsupport.uwo.ca/accessible_education/index.html)

if you have any questions regarding accommodations.

Learning-skills counsellors at the Student Development Centre (<https://learning.uwo.ca>) are ready to help you improve your learning skills. They offer presentations on strategies for improving time management, multiple-choice exam preparation/writing, textbook reading, and more. Individual support is offered throughout the Fall/Winter terms in the drop-in Learning Help Centre, and year-round through individual counselling.

Western University is committed to a thriving campus as we deliver our courses in the mixed model of both virtual and face-to-face formats. We encourage you to check out the Digital Student Experience website to manage your academics and well-being: <https://www.uwo.ca/se/digital/>.

Additional student-run support services are offered by the USC, <https://westernusc.ca/services/>.