

Data Science 2000B

Introduction to Data Science

Course Outline

1. Course Information

Course Information

Data Science 2000B

Winter Semester 2024/25

Lecture: Thursday 10:30am – 12:30pm, BGSB-0165

Lab: Friday 12:30pm-2:30pm, UCC-41 (TBD) or

Lab: Friday 2:30pm-4:30pm, SSC-3028, (TBD)

List of Prerequisites:

1.0 courses from Mathematics, Calculus, or Applied Mathematics (numbered 1000 and higher) with a minimum mark of 60%. [Data Science 1000A/B](#) (with a minimum grade of 60%) can be used to meet 0.5 of the 1.0 mathematics course requirements.

Unless you have either the requisites for this course or written special permission from the Undergraduate Chair of Data science to enroll in it, you may be removed and withdrawn from this course in accordance with university policy. This may be done after the add/drop deadline of the academic term, and the course will be marked as withdrawn (WDN) on your academic record. This decision may not be appealed.

2. Instructor Information

Instructors	Email	Office	Office Hours
Prof. Jörn Diedrichsen (Course Coordinator)	jdiedric@uwo.ca	WIRB 4138	Upon request only
Amanda Lily	arivet3@uwo.ca		TBD
Mariam Abdelmagied	mabde8@uwo.ca		TBD
Yin Quingrong	gyin33@uwo.ca		-

All subject-specific questions must be asked in the OWL forum, we will not answer questions sent to us by email. Emails are reserved **strictly** for private and confidential communications. We will not answer any course-related questions via email. Office hours are to be determined and will be posted on OWL.

3. Course Syllabus, Schedule, Delivery Mode

The course covers three basic concepts of data science together with the corresponding techniques:

- Sampling to estimate properties of a population (Bootstrap),
- Random assignment and experiments to make causal inferences (Randomization test),
- Model selection to enable good predictions (Cross-validation).

Emphasizes practical data handling and programming skills in Python.

Week starting with	Lecture (Thursday)	Lab section (Friday)	Assignment (due Wednesday, 12pm on the next week)
Week 1 Jan 05	Data Basics (O1.2): Examining numerical data (O1.6), Mean, Median, Mode, Standard deviation	Python and Pandas basics Jupyter notebooks Variables	HW 1 Pandas Basics Histograms Descriptive Statistics
Week 2 Jan 12	Examining categorical data (O1.7) Probability (O2.1-O2.2, O2.4)	Cross-tabulation Numpy arrays and indexing	HW2 Analyzing categorical data Probabilities Descriptive by Category
Week 3 Jan 19	Estimation, Confidence Intervals, Bootstrap Types of studies (O1.4, 1.5)	Functions 1, Randomization 1 Lab-quiz 1 (covering HW1+2)	HW3 Sampling and Bootstrap
Week 4 Jan 26	Foundations of Inference (R2) Hypothesis testing	Dictionaries, Functions Randomization	HW4 Randomization Test
Week 5 Feb 02	Test for independence Monte-carlo Decision errors and Power	Randomization Lab-quiz 2 (covering HW3 +4)	HW5 More Randomization test and Monte Carlo simulation
Week 6 Feb 9	Bayesian Inference	Discrete and continuous probability distributions. Calculating Bayes Factors.	HW6 Bayesian inference + Midterm practice
Feb 16	Reading week		
Feb 23	Midterm		
Week 7 Mar 02	Simple linear regression: Cost function and optimization parameter estimation	Loss function Optimization Lab-quiz 3 (covering HW5+6)	HW7 Optimization of squared error cost function
Week 8 Mar 9	Outliers and robust techniques,	Functions Optimization	HW8

	Polynomial regression, Cross-validation	Cross-validation	Optimization of Mean-absolute deviation cost function (Median regression) Polynomial regression Crossvalidation
Week 9 Mar 16	Multiple regression: Model comparison and correlation of regressors	Review of HW7 and 8 Bootstrap for Multiple regression Lab-quiz 4 (covering HW 7 + 8)	HW9 Multiple regression Bootstrap for multiple regression
Week 10 Mar 23	Discrete variables and logistic regression	Dummy coding Logistic regression	HW10 Regression with discrete variables: dummy coding Logistic regression
Week 11 Mar 30	Model selection, Stepwise regression Regularization	Stepwise regression Regularization Lab-quiz 5 (covering HW 9+10)	HW11 Final practice
Apr 12 – 30	Final (TBD)		

4. Course Materials

All course material will be posted to OWL: <https://westernu.brightspace.com/>

Students are responsible for checking the course OWL site (<https://westernu.brightspace.com/>) regularly for news and updates. This is the primary method by which information will be disseminated to all students in the class.

Helpful resources are:

Python for Data Analysis: (P)

<https://www.amazon.ca/Python-Data-Analysis-Wrangling-IPython/dp/1491957662>

Open Intro Statistics (O):

https://www.openintro.org/stat/textbook.php?stat_book=os

Open Intro Stat with Randomization and Simulation (S):

https://www.openintro.org/stat/textbook.php?stat_book=isrs

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.

Technical Requirements

A laptop with internet connection and working web browser is required for homework assignments, labs, midterms and final. The computer must have an appropriate installation of Python (>3.6), including the required software packages (details available in the first week of classes).

To participate in the in-lecture and in-lab quizzes, all students must register on iClicker and have a laptop or mobile device with the iClickerApp installed: <https://wts.uwo.ca/iclicker/index.html>

If students need assistance with the course OWL site, they can seek support on the [OWL Brightspace 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

Grading Scheme and Assessment Dates

The overall course grade will be calculated as listed below:

Weekly Assignments:	10%
Lecture Questions:	10%
Lab Quizzes:	15%
Midterm:	30%.
Final:	35%

There will be a total of **11 weekly homework** assignments. Assignments are due on Wednesday at noon (see OWL site for exact dates). Assignments will be marked with Pass / Fail. Assignments will be marked with a pass (full points) if the student demonstrated a good-faith attempt to solve each of the questions. Incomplete assignments or assignment submitted after the due-date will be marked with a fail (0 points). *Only 10 of 11 assignments will be counted, so every student can drop one assignment, without submitting an Academic Consideration.* Each assignment therefore accounts for 1% of the course grade.

There will be **5 lab-quizzes**, held every other week (see Calendar). These quizzes cover the materials for the labs and probe if you have understood the materials covered in the previously submitted homework. Lab quizzes also require you to understand the materials covered in lecture and be able to apply them to real problems. *Only the best 4 of 5 lab quizzes will be counted, so each student can drop one lab-quiz, without submitting an Academic Consideration.* Each lab quiz therefore accounts for 3.75% of the course grade.

Each Lecture will include some short iClicker **Lecture questions**, distributed over the course of the lecture. These are not graded but designed to make the lectures more interactive – to get full points you need to make a response to at least 80% of the questions. *Only 10 of the 11 lecture quizzes will be counted, so every student can miss one lecture, without submitting an Academic Consideration.* The iClicker questions for each lecture therefore account for 1% of the course grade.

Note that, while we drop the lowest Assignment, lab-quiz, and lecture-quiz, all these elements are essential for achieving the learning outcomes of the course. It is the responsibility of the student to ensure that they have mastered the content covered in these class elements, even if the evaluation was missed.

For details of time, duration, and place of **Midterm and Final**, please check updates on OWL. Midterm and Final will contain both a closed-book and an open-book section.

Use of Generative AI Tools

The use of generative AI tools for Lab quizzes, Midterm, and Final is forbidden. Co-pilot, ChatGPT, Gemini, and other generative AI tools need to be disabled (Google Collab) or uninstalled (VSCode) during the entire duration of the tests. For weekly **homework assignments, it is the student's responsibility to use generative AI tools to promote, rather than hinder, their learning.** It is recommended that you first attempt to solve each homework assignment by yourself. Only if you

cannot find the answer after repeated tries, you can Co-Pilot. Ensure that you understand the answer, experiment with the code, and revise the answer yourself, if you are doing the homework assignments with AI help.

Evaluation Scheme for Missed Assessments

Given that we have built in an assignment flexibility (dropping the lowest one) for lab quizzes, lecture questions and homework assignments, **Academic Consideration requests without documentation will not be relevant in this course.** If a student has an approved Academic Consideration **with documentation**, the percentage of the missed elements will be reweighted to the remaining elements within the category.

If a student has Academic Considerations with documentation for the midterm (this needs to be submitted through faculty student counselling), the Final will be counted with 65%. **There will be no other makeup for the midterm.** If a student has an approved Academic Consideration for the Final, I will offer 2 possible dates for an **30-min oral examination** on the course content, no later than the end of the final period. Students will not be allowed to take an incomplete and retake the final in the next year.

Students must familiarize themselves with the University Policy on Academic Consideration – Undergraduate Students in First Entry Programs posted on the Academic Calendar:

https://www.uwo.ca/univsec/pdf/academic_policies/appeals/academic_consideration_Sep24.pdf,

This policy does not apply to requests for Academic Consideration submitted for **attempted or completed work**, whether online or in person.

The policy also does not apply to students experiencing longer-term impacts on their academic responsibilities. These students should consult [Accessible Education](#).

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/

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

All Academic Consideration requests relevant to this course must include supporting documentation.

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.

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/mapp113.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.

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:

https://www.uwo.ca/univsec/pdf/academic_policies/appeals/scholastic_discipline_undergrad.pdf.
https://www.uwo.ca/univsec/pdf/academic_policies/appeals/undergrad_scholastic_offence_procedure.pdf

Computer-marked multiple-choice tests and exams may be subject to submission for similarity review by software that will check for unusual coincidences in answer patterns that may indicate cheating.

iClicker will be used for synchronous lab- and lecture quizzes. Questions will be posted on the slides, discussed, and answers have to be submitted by each student with a personalized iClicker device or app. Answers will be used for grading. As will all other assignments, the outcome will only be seen by the course instructors.

Support Services

Please visit the Science & Basic Medical Sciences Academic Advising 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.

To connect with a case manager or set up an appointment, please contact support@uwo.ca.