DS 1000A (Data Science Concepts)
Section 650
Summer 2022 Course Outline

1. Course information:

Instructor	Day/Time	Location	Contact	Office hours
Joseph Raaymakers, M.Sc.		Online, Asynchronous	Use Messages in OWL. (to Instructor's Role) or email jraaymak@uwo.ca	Online via Zoom by appointment (email to arrange a time)

Calendar description: Students will learn how to visualize and analyze continuous and categorical data from various domains, using modern data science tools. Concepts of distributions, sampling, estimation, confidence intervals, experimental design, inference, correlation will be introduced in a practical, data-driven way.

Requisites: One or more of Ontario Secondary School MCV4U, MHF4U, MDM4U, Mathematics 0109A/B, Mathematics 0110A/B, Mathematics 1229A/B, or equivalent.

Anti-requisites: Statistical Sciences 1023A/B, the former Statistical Sciences 1024A/B.

Extra Information: Lectures and laboratory tutorials posted each week on OWL.

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.

Course delivery

This course is distance studies, with online asynchronous learning. To be successful in the course, students should expect to complete independent study tasks. Lectures on course material and tutorials in Python will be posted each week. Assessments will be conducted online as determined by the course instructor.

When deemed necessary, tests and examinations in this course will be conducted using a remote proctoring service. By taking this course, you are consenting to the use of this software and acknowledging that you will be required to provide personal information (including some biometric data) and that the session will be recorded. Completion of this course will require you to have a reliable internet connection and a device that meets the technical requirements for this service. More

information about this remote proctoring service, including technical requirements, is available on Western's Remote Proctoring website at:
https://remoteproctoring.uwo.ca.

2. Course Objectives and Schedule

By the end of this course, a successful student will be able to:

- Understand and correctly use foundational vocabulary associated with Statistics and Data Science.
- Interpret, create and critically evaluate graphical and numerical data summaries.
- Understand and appreciate probability, chance, randomness, and 'average'.
- Understand, assess, and critique the conclusions of data analyses.
- Apply concepts learned in this course to future courses, careers, and everyday life.

→ **Course schedule:** see Pages 6 to 8.

3. Communication

- To communicate with the instructor you can **use email or OWL Messages (to Instructor's Role)**.

You should 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/holidays.

Students should check OWL (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. Students are responsible for checking OWL on a regular basis.

The Forum tool is enabled on the OWL website. Please use this Forum to post and respond to questions about course content (e.g., lecture, readings, etc.). The Forum 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 TA's. Derogatory, discriminatory, or otherwise inappropriate language or topics will be removed and dealt with at the instructor's discretion.

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. Course Materials

Required text: The Basic Practice of Statistics, 9th Ed, 2021, by D. S. Moore; W. I. Notz

UWO bookstore listing:
https://bookstore.uwo.ca/textbook-search?campus=UWO&term=W2021A&courses%5B0%5D=002_UW/DATASCI1000A

Laboratory tutorials: Python and Jupyter Notebook are the main tools for laboratory tutorials. Instructions on how to install them are available on OWL under Resources.

6. Methods of Evaluation

Component	Weight	Deadlines/Due dates
Assignment 1	6%	**Assign. 1 due June 1**
Assignment 2	6%	**Assign. 2 due June 29**
Assignment 3	8%	**Assign. 3 due July 27**
Midterm 1	20%	**June 8, 7:00 – 8:30 pm**
Midterm 2	20%	**July 6, 7:00 – 8:30 pm**
Final exam (and closed book)	40%	During exam period (Aug 2 to 5, to be scheduled by the Registrar's Office). Cumulative.

i) 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** (https://www.gradescope.com/) an on-line 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.

- 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".

- **Students must submit at least 2 out of the 3 assignments to write the final exam.**

- **Assignment submissions are due 11:55pm (Eastern Time) on the due date**. No credit will be given for submissions beyond this time unless a valid academic accommodation is obtained (see Section 7 for details on accommodation).

- Solutions to assignments **will not** be posted; however, TAs will provide comments on wrong 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 during office hours.

ii) Midterms and final exam

- There will be two **1.5-hour online midterm exams to be scheduled for June 8 and July 6 from 7:00 to 8:30 pm**.

- If you do poorly on one of the tests, the weight of that test will be reduced to 15% of your final grade and the weight of the other test will be increased to 25%.

- **There are no make-ups for the tests**. If you miss one of the tests, the weight of the marks for that test will be shifted to the final exam. **If you miss both tests, you will not be permitted to write the final exam.**

- There will be a **3-hour online final exam during the exam period (August 2 to 5) and its time will be scheduled by the Registrar's Office**. The exam will be cumulative and cover all material.

iii) 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.

7. Accommodation and accessibility

- If you are unable to submit an assignment by the due date/time, you must seek approval for the absence as soon as possible and contact your instructor immediately. Approval can be granted either through a self-reporting of absence (SRA) or via the Dean's Office/Academic Counselling unit of your Home Faculty. Those **students obtaining such appropriate accommodation will have to submit their assignment via OWL messages within 48 hours of the end of the accommodation period**. Failure to submit a missed assignment within 48 hours of a valid academic accommodation period will result in a grade of 0%.

- If you miss the Midterm you must request accommodation from Academic Counseling in your Faculty's Dean's Office. Students who obtain appropriate accommodation will have the weight of the midterm redistributed to the final exam.

- If you miss the Final Exam you must request accommodation from Academic Counseling in your Faculty's Dean's Office and contact the instructor immediately.

- Students with disabilities work with Accessible Education (formerly SSD) which provides recommendations for accommodation based on medical documentation or psychological and cognitive testing. The Academic Accommodation for Students with Disabilities policy can be found at: https://www.uwo.ca/univsec/pdf/academic_policies/appeals/Academic Accommodation_disabilities.pdf

- Religious Accommodation: students should consult the University's list of recognized religious holidays, and should give reasonable notice in writing, prior to the holiday, to the Instructor and

an Academic Counsellor if their course requirements will be affected by a religious observance. Additional information is given in the Western Multicultural Calendar: https://multiculturalcalendar.com/ecal/index.php?s=c-univwo

8. Academic Policies

The website for Registrarial Services is http://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 his/her official university address is attended to in a timely manner.

If you are unable to meet a course requirement due to illness or other serious circumstances, you must seek approval for the absence as soon as possible. Approval can be granted either through a self-reporting of absence or via the Dean's Office/Academic Counselling unit of your Home Faculty. For further information, please consult the university's policy on academic consideration for student absences: https://www.westerncalendar.uwo.ca/PolicyPages.cfm?PolicyCategoryID=1&command=showCategory&SelectedCalendar=Live&ArchiveID=

Participants in this course are not permitted to record lectures and/or labs, except where recording is an approved accommodation, and the participant has the prior written permission of the instructor.

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.

Professionalism & Privacy

☒ Western students are expected to follow the Student Code of Conduct (https://www.uwo.ca/univsec/pdf/board/code.pdf). Additionally, the following expectations and professional conduct apply to this course:

☒ All course materials created by the instructor(s) are copyrighted and cannot be sold/shared

☒ Recordings are not permitted (audio or video) without explicit permission

☒ Permitted recordings are not to be distributed

Copyright Statement

Please be aware that all course materials created by the instructor(s) are copyrighted and **cannot be sold/shared**. Those include materials used in assignments, tests/quizzes, midterms, and finals. Any posting/sharing of such materials in part or whole without owner's consent is considered as violation of the Copyright Act and will be considered as a scholastic offence.

In addition, online services such as Chegg are actively monitored. Any questions that are coming out from assignments and are posted to an online service will be searched. Such an activity will be considered a scholastic offence and will result in an academic penalty.

9. Support Services

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

- 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 Student Accessibility Services (SAS) at 519-661-2147 if you have any questions regarding accommodations.

- Learning-skills counsellors at the Student Development Centre (http://www.sdc.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.

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

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

10. Tentative Course Schedule - Textbook: The Basic Practice of Statistics 9th Ed.

Week		Chapter	Topics	Lab tutorial
1	May 9-13	1 Sections: 1.1 to 1.6	- Categorical variables (pie charts and bar plots) - Quantitative variables (histograms, stem plots, time plots)	- Introduction to Python and Jupyter notebook - Data frames, arrays, types of variables
2	May 16-20	2 Sections: 2.1 to 2.5	- Mean, Median - Quartiles - Five-number summary	- Basic Python commands and functions
3	May 23-27	2 Sections: 2.5 to 2.8	- Inter-quartile range - Boxplots and spotting outliers - Standard deviation - Choosing measures of center and variability	- Import data - Histogram, pie chart and bar plot - Five-number summary - Boxplot - Calculating standard deviation
4	May 30 – Jun 3	3 Sections:	- The normal distribution - The 68-95-99.7 rule	- Generating normal data and plotting the corresponding histogram

		3.3 to 3.6	- Finding Normal Proportions	- Calculating normal proportions **Assign. 1 due June 1**
5	Jun 6 - 10	4 Sections: 4.1 to 4.6	- Explanatory and response variables - Displaying relationships: scatterplots - Measuring linear correlation (Pearson's correlation coefficient)	- Scatterplot - Linear correlation (Python function for Pearson's correlation coefficient) **Midterm 1 June 8, 7:00 – 8:30 pm**
6	Jun 13-17	5 Sections: 5.1, 5.7, 5.2, 5.3, 5.4, 5.8	- Regression lines - Least-squares regression line - Examples of software regression output - Caution about correlation and regression - Association does not imply causation	- Least-squares regression fit - Interpretation of result table
7	Jun 20-24		- Two-way contingency tables - Relative risk, odds ratio - Simpson's Paradox	- From raw data to a two-way table - Computing conditional and marginal proportions - Relative risk, odds ratio - Mosaic plot
8	Jun 27 – Jul 1	8 Sections: 8.1 to 8.7	- Population versus sample - Bad samples - Simple random sample - Stratified random sample - Why can we trust random samples? - Cautions about sample surveys	- Simple random sampling (SRS), systematic sampling, and stratified random sampling **Assign. 2 due June 29**
9	Jul 4-8	9 Sections: 9.1 to 9.7	- Observational studies versus random experiments	- Generating samples **Midterm 2 July 6, 7:00 – 8:30 pm**
10	Jul 11-15	12, 13 Sections: 12.1 to 12.7, 13.1 to 13.5	- Rules of probability - Addition - Independence and the multiplication rule - Conditional probability - Venn diagrams	- Venn diagram
11	Jul 18-22	15 Sections: 15.3 to 15.6	- Sampling distributions - Mean sampling distribution - Central limit theorem - Statistical significance	- Mean sampling distribution

12	Jul 25-29	16 Sections: 16.1 to 16.4 32 Sections: 32.3 and 32.4	- Quantifying estimation uncertainty - Confidence intervals (CIs) for a population mean - Bootstrap confidence intervals	- Building normal based CIs - Bootstrap samples - Bootstrap CIs **Assign. 3 due July 27**