DS 1000B (Data Science Concepts)
Section 001
Winter 2022 Course Outline

1. Course information:

Instructor	Day/Time	Location	Contact	Office hours
Joseph Raaymakers, M.Sc.	Tu 8:30 am to 9:30 am, Th 8:30 am to 10:30 am	Online. The Zoom link will be announced in OWL. [1]	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: 1 hour of laboratory tutorial per week conducted by TAs.

Unless you have either the requisites forhe re4(is-3(you)5( )-3(h)-4(ave78-3(r)4(isi)11(t)-4(es)5(f)-4(or)65(h)t 0 0 1

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 <u>always use OWL Messages (to Instructor's Role)</u>.

You can expect a response to a message to the instructor within approximately 48 hours

8. Academic Policies

9. Support Services

- Please

5	Feb 7 - 11	4 Sections: 4.1 to 4.6	- Explanatory and response variables - Displaying relationships: scatterplots - Measuring linear correlation (Person's correlation coefficient)	- Scatterplot - Linear correlation (Python function for Pearson's correlation coefficient)
6	Feb 14-18	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	Feb 21- 25	Spring Reading Week		
████	████	████	- 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
9	March 7 - 11	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	**Midterm March 9, 7 9 pm**
10	March 14-18	9		

| 13 | April 04- April 08 | 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 April 06 |