THE UNIVERSITY OF BRITISH COLUMBIA

DATA 301/COSC 301/DATA 501:
Introduction to Data Analytics
Fall 2019 Term 1

Instructor:	Dr. Irene Vrbik
Office:	SCI 393
email:	irene.vrbik@ubc.ca
Lectures:	Mon Wed Fri (11:30 – 12:30 in ASC 140)
Office Hours:	Friday 2:00 – 3:00 PM

Course Description

Official Calendar: Techniques for computation, analysis, and visualization of data using software. Manipulation of small and large data sets. Automation using scripting. Real-world applications from life sciences, physical sciences, economics, engineering, or psychology. No prior computing background is required. Credit will be granted for only one of COSC 301, DATA 301 or DATA 501. [3-2-0] *Prerequisite*: Either (a) third-year standing, or (b) one of COSC 111 or COSC 122.

Specific description: This course provides an introduction to data analytics to train students with practical industrial techniques for data manipulation, analysis, reporting, and visualization. This is **not** an introduction to programming using Python. Programming techniques will be taught to automate data analysis. Introduction to programming courses are COSC 111 or COSC 123. Prior computing experience is not required, but is helpful including COSC 122 or COSC 111. See course descriptions here.

Marking and Evaluation

Graduate students are responsible for a substantial data analytics project. Optional or bonus assignment questions for undergraduates will often be required for grad students.

Undergraduate (Data 301/Cosc 301) Evaluation

Clickers	5 %	
Assignments	20 %	
Midterm 1	15 %	
Midterm 2	15 %	(not cumulative)
Final Exam	45 %	(cumulative 3 hours)

Graduate Student (Data 501) Evaluation

Clickers	5 %
Assignments	15 %
Project	10 %
Midterm 1	15 %
Midterm 2	15 %
Final Exam	40 %

- A graduate student that receives less than 60% on the project will fail the course and be awarded a maximum overall mark of 45.
- A student must either a) pass the exam with a grade of at least 50% or b) receive an average grade of at least 50% on the exams (midterms and final) to pass the course. Otherwise, the student will be assigned a maximum overall grade of 45.

Required Materials

- A clicker is required.
- Course notes available on Canvas.
- All notes are available online. A textbook is not required.

Software:

- Microsoft Excel (see Office 365 for UBC Students)
- Microsoft Access (part of Office 365 for Windows) or LibreOffice Base
- R
- Tableau (which is installed on the lab computers are available for student download.

Other recommended Software: I strongly recommend installing Python and Jupyter using the Anaconda and that you use Rstudio for running R.

Labs and Assignments

Please check your registration to determine your lab/tutorial section and time. You must register for a lab and attend the one you are registered in. While labs are not mandatory (i.e. attendance will not be taken) you are highly encouraged to attend. Labs will provide guidance on carrying out assignments which involve techniques discussed in lecture. It is likely that assignments will take longer than the scheduled lab time to complete.

Labs assignments are to be submitted electronically through Canvas. Late assignments will have 10% deducted for each day (which includes weekends) beyond the due date. Assignments that are more than 5 days overdue will **not** be accepted.

Tentative Course Schedule

Below is an outline of the topics. The professor is not bound to these topics and timelines as they only serve as a general reference.

Week	Description of Topics
1	Introduction to course. What is data analytics?
	Data Representation: Data and metadata; file formats and encoding; text/binary files
2	Excel I: Intro to Excel – formulas, formatting, aggregate functions
	Excel II: Data Analysis in Excel – sorting, filtering, charts, what-if scenarios, pivot tables
3	Excel III: Excel scripting – macros, VBA
	Databases I: Introduction to Relational Databases; Creating a Database
4	Databases II: Querying using SQL
	Databases III: Advanced Querying using SQL
5	Command Line: Linux/Unix Command Line Introduction
	Python I: Introduction to Python
6	**Midterm 1: Wednesday October 9**
	Python II: Decisions and Loops
7	Python III: Reading and Writing Files
	Python IV: Data Analysis with Python
8	Python V: Data Analysis with Map-Reduce
	R I: Data Analysis with R – Brief statistics intro/review
9	R II: Data Analysis with R (cont.)
	Midterm 2: Monday November 4
10	R III: Data Analysis with R (cont.)
	GIS*
11	Data Visualization I: Reporting
	Data Visualization II: Tableau
12+	Open science (guest lecture by Dr. Jason Pither Nov 20th)
	Review for final exam
	Data 501 final project presentations

this topic is subject to be removed/changed

Grievances and Complaints Procedures

A student who has a complaint related to this course should attempt to resolve the matter with the instructor first. Students may talk first to someone other than the instructor if they do not feel, for whatever reason, that they can directly approach the instructor. If the complaint is not resolved to the student's satisfaction, the student should go to the departmental chair John Braun.

Your Responsibilities

Your responsibilities to this class and to your education as a whole, include attendance and participation. You have a responsibility to help create a classroom environment where all may learn. At the most basic level, this means you will respect the other members of the class and the instructor and treat them with the courtesy you hope to receive in return. Inappropriate classroom behaviour may include: disruption of the classroom atmosphere, engaging in non-class activities, talking on a cell-phone, inappropriate use of profanity in classroom discussion, use of abusive or disrespectful language toward the instructor, a student in the class, or about other individuals or groups.

Academic Integrity

The academic enterprise is founded on honesty, civility, and integrity. As members of this enterprise, all students are expected to know, understand, and follow the codes of conduct regarding academic integrity. At the most basic level, this means submitting only original work done by you and acknowledging all sources of information or ideas and attributing them to others as required. This also means you should not cheat, copy, or mislead others about what is your work. Violations of academic integrity (i.e., misconduct) lead to the breakdown of the academic enterprise, and therefore serious consequences arise and harsh sanctions are imposed. For example, incidences of plagiarism or cheating may result in a mark of zero on the assignment or exam and more serious consequences may apply if the matter is referred to the Presidents Advisory Committee on Student Discipline. Careful records are kept in order to monitor and prevent recurrences. **If you have any questions about how academic integrity applies to this course, please consult with your professor.**

Disability Services

If you require disability-related accommodations to meet the course objectives, please contact the Disability Resource Centre in UNC 227. More information is at: http://students.ok.ubc.ca/drc.

Equity, Human Rights, Discrimination and Harassment

UBC does not condone discrimination or harassment in classrooms, living or work environments on campus. For information about UBC's policies related to equity go to: http://equity.ok.ubc.ca/.

Missing an Exam

Only students who miss the final exam for a reason that corresponds to the University of British Columbia Okanagan's policy on excused absences from examinations will be permitted to take the final exam at a later time. A make-up exam may have a question format different from the regular exam. **There will be no make-up midterm exams**. If the reason for absence is satisfactory and appropriate documentation is provided, missed midterms may have weight shifted to the final exam. Further information on Academic Concession can be found here, under Policies and Regulation in the Okanagan Academic Calendar.

Course Objectives

Course Format: Interactive classes consisting of topic introduction and concept mastery with in-class exercises. Practical skills and applications of topics are covered in computer labs as well as practiced using data analytics systems and software.

Learning Outcomes:

- Ability to manipulate, extract, convert, and integrate data from different sources.

- Ability to perform advanced Excel analysis including what-if scenarios, pivot tables, and VBA scripting.
- Ability to use relational databases including creating tables and querying using SQL.
- Ability to use scripting programs to automate repetitive and large tasks and improve efficiency.

Course Objectives:

- Understand data representation formats and techniques and how to use them.
- Experience using a wide-range of data analytics tools including Excel, SQL databases, GIS, and visualization and reporting software.
- Develop a computational thinking approach to problem solving and use programs to solve data tasks.

Graduate Student Objectives:

- In addition to the skills learned in the cross-listed DATA 301 course, DATA 501 will extend the students skills in conducting independent research studies with data analytical techniques. These skills may also be applied to research in the students thesis work.
- The project will involve background literature review, data analysis of a real-world data set, and public presentation of the results (more details will be posted on Canvas).

Other Relevant Resources:

UBC Okanagan Disability Resource Centre

The Disability Resource Centre ensures educational equity for students with disabilities, injuries or illness. If you are disabled, have an injury or illness and require academic accommodations to meet the course objectives, please contact Earllene Roberts, the Diversity Advisor for the Disability Resource Centre located in Commons Corner in the University Centre building:

Where:	UNC 227A 250.807.9263
Email:	earllene.roberts@ubc.ca
Web:	http://students.ok.ubc.ca/drc/welcome.html

UBC Okanagan Equity and Inclusion Office

UBC Okanagan is a place where every student, staff and faculty member should be able to study and work in an environment that is free from discrimination and harassment. UBC prohibits discrimination and harassment on the basis of the following grounds: age, ancestry, colour, family status, marital status, physical or mental disability, place of origin, political belief, race, religion, sex, sexual orientation or unrelated criminal conviction. If you require assistance related to an issue of equity, discrimination or harassment, please contact the Equity and Inclusion Office or your administrative head of unit.

Where:	UNC 227C 250.807.9291
Email:	equity.ubco@ubc.ca
Web:	www.equity.ok.ubc.ca

Health & Wellness

At UBC Okanagan health services to students are provided by Health and Wellness. Nurses, physicians and counsellors provide health care and counselling related to physical health, emotional/mental health and sexual/reproductive health concerns. As well, health promotion, education and research activities are provided to the campus community. If you require assistance with your health, please contact Health and Wellness for more information or to book an appointment.

Where: UNC 337
Email: healthwellness.okanagan@ubc.ca
Web: www.students.ok.ubc.ca/health-wellness

Don't want to walk alone at night? Working alone in an office or a lab late at night? Not too sure how to get somewhere on campus? Call Safewalk at 250.807.8076. For more information, see: www.security.ok.ubc.ca.