

DATA 301 / IGS 501 - Introduction to Data Analytics

Winter 2015 Term 2

Instructor: Dr. Ramon Lawrence
Class Schedule: 2:00 p.m. – 3:30 p.m. Tuesday/Thursday
Location: EME 2141
Lab time/location: L01 - 9:30 a.m. to 11:30 a.m. Friday in SCI 234
L02 - 8:30 a.m. to 10:30 a.m. Wednesday in SCI 234
L03 - 5:00 p.m. to 7:00 p.m. Tuesday in SCI 234
Office Hours: 9:00-10:00 a.m. Tuesdays/Thursdays or by appointment
Office Location: ASC 349
Phone: 807-9390
E-mail: ramon.lawrence@ubc.ca (preferred contact method)
Course URL: <https://people.ok.ubc.ca/rlawrenc/teaching/301/>

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. [3-2-0]

Prerequisite: Third-year standing.

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.

Marking and Evaluation

Clickers	5 %	
Assignments	35 %	(weekly assignments)
Two Midterm Exams	30 %	(in class, 15% each)
Final Exam	30 %	(cumulative, three hours)

IGS Graduate Student Evaluation:

Assignments	15 %	(weekly assignments)
Clickers	5 %	
Project	20 %	
Two Midterm Exams	30 %	(in class, 15% each)
Final Exam	30 %	(cumulative, three hours)

- Graduate students are responsible for a substantial data analytics project. Optional or bonus assignment questions for undergraduates will often be required for grad students.
- A student must receive a combined 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.

Textbook and Reference Material:

- A *clicker* is required. All notes are available online.
- A text book is *not required*.

Expectations

- Attend **all** classes and prepare before attending class.
- Read the notes **before** the lecture.
- Learn the material in the course by completing all assignments.
- Enjoy attending class and feel free to participate according to your own personalities. Feel free to ask questions by raising your hand or speaking out at appropriate times.
- Please actively participate in class discussions, questions, and problem solving exercises.
- **I want all students to pass the course, receive a good grade, and feel the course was beneficial.**

Homework Expectation

For this course, it is expected that you will spend *at least six hours per week in out-of-class preparation*.

Grievances and Complaints Procedures

A student who has a complaint related to this course should follow the procedures summarized below.

- The student 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 at SCI 388, 807-8032.

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 behavior 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 President's 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, the student's final exam will be worth more of the final grade. Further information on Academic Concession can be found under Policies and Regulation in the Okanagan Academic Calendar <http://www.calendar.ubc.ca/okanagan/index.cfm?tree=3,48,0,0>.

Course Objectives

Course Format: Interactive classes consisting of topic introduction, understanding evaluation using clickers, and concept mastery with in-class exercises. Practical skills and applications of topics are covered in computer labs as well as practice 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 computational thinking approach to problem solving and use programs to solve data tasks.

Course Outline

The course has a substantial amount of material to be covered in a short time. This requires the student make a strong effort to prepare before class so that the material can be practiced in class. 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.

Date	Topics Covered and Description
January 5 (T)	First day of classes. Introduction to course. What is data analytics?
January 7 (TH)	Data Representation I: Data and metadata ; file formats and encoding ; text/binary files
January 12 (T)	Excel I: Introduction to Excel – formulas, formatting, aggregate functions
January 14 (TH)	Excel II: Data Analysis in Excel – sorting, filtering, charts, what-if scenarios, pivot tables
January 19 (T)	Excel III: Excel scripting – macros, VBA
January 21 (TH)	Databases I: Introduction to Relational Databases ; Creating a Database
January 26 (T)	Databases II: Querying using SQL
January 28 (TH)	Databases III: Advanced Querying using SQL
February 2 (T)	Linux/Unix Command Line Introduction
February 4 (TH)	Python I: Introduction to Python
February 9 (T)	No classes during Midterm Break.
February 11 (TH)	No classes during Midterm Break.
February 16 (T)	Python II: Decisions and Loops
February 18 (TH)	Midterm Exam #1
February 23 (T)	Python III: Reading and Writing Files
February 25 (TH)	Python IV: Data Analysis with Python
March 1 (T)	Python V: Data Analysis with Map-Reduce
March 3 (TH)	Data Analysis with R – Brief statistics intro/review
March 8 (T)	Data Analysis with R (cont.)
March 10 (TH)	Data Analysis with R (cont.)
March 15 (T)	Midterm Exam #2
March 17 (TH)	GIS
March 22 (T)	GIS (cont.)
March 24 (TH)	Data Visualization I: Reporting
March 29 (T)	Data Visualization II: Tableau
March 31 (TH)	Web Data Services ; Biological and Scientific Data Sets
April 5 (T)	Data Integration - Putting it all together
April 7 (TH)	Demonstrations of data analytics projects. Review for final exam.

Laboratory times: The laboratory time will be spent performing assignments and practice questions using data analytics software.

Week	Dates	Topics Covered and Description
1	January 4 – 8	No Lab First Week of Class
2	January 11 – 15	Lab 1: Excel – Analyzing and Reporting Data
3	January 18 – 22	Lab 2: Excel – Pivot tables and What-If Scenarios
4	January 25 – 29	Lab 3: Excel – Macros and VBA Scripting
5	February 1 – 5	Lab 4: Creating a Relational Database in Microsoft Access and MySQL
6	February 8 - 12	No Lab During Midterm Break
7	February 15 - 19	Lab 5: Introduction to Python
8	February 22 - 26	Lab 6: Python Data Analysis
9	Feb. 29 - March 4	Lab 7: Python Map-Reduce
10	March 7 - 11	Lab 8: Introduction to R
11	March 14 – 17	Lab 9: GIS
12	March 21 – 25	Lab 10: Visualization with Tableau Note: No lab for L01 on Good Friday, March 25th. Make up on April 1st.
13	Mar. 28 – April 1	Lab 10: Visualization with Tableau (cont.)
14	April 4 – 8	Bonus Lab: Putting it all together on your own data set