STAT 2150

Statistics and Computing (A01)

Fall Term 2018

Class Time:	Tuesday / Thursday 1:00 p.m. - 2:15 p.m.
Location:	208 Armes
CRN:	17120

Instructor:	Saman Muthukumarana
Office:	371 Machray Hall
	Telephone: 474-6274
	Email: `Saman_Muthukumarana@UManitoba.CA`

Office Hours:	Monday 10:00 - 11:30 a.m.
	Thursday 11:00 - 12:00 p.m.
	(Or by appointment.)

Calendar Description: (Lab required) This course is recommended for students in mathematically rich disciplines, including Statistics, Mathematics, Actuarial Science, Computer Science, and related interdisciplinary programs. Topics to be covered include: exploratory data analysis and visualization, graphical methods, random number generation, random variables, simple statistical models and computing, Monte Carlo methods, large sample and simulation-based inference, statistical software packages.

Prerequisite: [STAT 1150 (C) or STAT 2000 (B) or STAT 2001 (B)] and [a C or better in one of: MATH 1230, MATH 1500, MATH 1501, MATH 1510, the former MATH 1530, or MATH 1690 (C)].

Course Website: The course website is accessible via UM Learn management system.

Computing: This course will expose you to R computing language throughout the course. R is available in lab computers at 311 Machray Hall. R is a free software environment for statistical computing and runs on Windows, Linux, UNIX and Mac. You can download your own copy from R Project (CRAN) homepage at `http://www.r-project.org/`. The introductory tutorial for R can be found here. The official R Short Reference Card contains some basic useful functions. RStudio can be downloaded from `https://www.rstudio.com/`.

Grading Scheme: The minimum percentage final mark required to receive each of the various letter grades are A+ (90%), A (80%), B+ (75%), B (70%), C+ (65%), C (60%), D (50%). The final mark will be determined as follows. You will have access to your assignments, quizzes and mid term marks via UM Learn Grade-book.

UM Learn Assignments (4)	10%
In Class Mid-term Test	30%
Computer Lab Quizzes (2)	10%
Final Examination	50%

Assignments: All assignments must be submitted by their due date using **UM Learn Dropbox**. Assignments will be marked using Brightspace **Assignment Grader** and specific submission instructions will be posted on UM Learn. You are expected to work independently on the assignments. Assignments will include problems which require data analysis using R and your submission must accompany R code written by you.

Mid-term Test: The in-class mid-term test is tentatively scheduled on **October 30, 2018**. There will be no makeup tests for any reason. If you miss the exam due to a legitimate reason, your exam weight will transfer to the final exam. Your midterm exam will be marked and returned to you using Crowdmark.

Computer Labs and Quizzes: You will attend a mandatory computer lab session once per week starting from 2nd week. The lab sessions will be held in 311 Machray Hall computer lab. You must attend the lab session in which you are registered. The lab quizzes are tentatively scheduled during **Oct 15, 16, 17 and Nov 19, 20, 21** lab sessions and **you must write the quizzes in the lab section in which you are registered**. There will be no makeup quizzes for any reason. If you miss the 1st quiz due to a legitimate reason, your quiz weight will transfer to the mid term. If you miss the 2nd quiz due to a legitimate reason, your quiz weight will transfer to the final.

Final Exam: The final exam covers all course materials and will be 3 hours in length. It will be scheduled by Registrar's Office. Non-programmable calculators are allowed in the midterm and final. No other electronic devices can be in your possession during the midterm and final exam.

Class and Computer Lab Attendance: I will introduce, discuss and demonstrate R codes in classes to understand the course material. Computer Lab sessions will provide you with various practical hands-on computing experience by analyzing real data and conducting simulation studies. Therefore, I encourage you to attend classes and labs regularly to avoid falling behind. The exams will also resemble in part on problems discussed during classes.

Textbook: You will have access to my lecture notes via course website and there is no required textbook for this course due to its broad nature in blend of theory and computing. However, listed below are textbooks that are excellent resources for this course. Note that e-copies of these books are freely available from SpringerLink via UoM Library server. For some lectures, specific sections from these books will be assigned as reading assignments.

- *Understanding Statistics Using R*,
  Randall Schumacker and Sara Tomek, Springer (2013).

- *Statistical Analysis and Data Display: An Intermediate Course with Examples in R* (Second Edition),
  Richard M. Heiberger and Burt Holland, Springer (2015).

- *Introduction to Statistics and Data Analysis*,
  Christian Heumann, Michael Schomaker and Shalabh, Springer (2016).

- *Data Wrangling with R* (Second Edition),
  Bradley C. Boehmke, Springer (2016).

Course Outline: The course aims to cover the following topics.

- Exploratory data analysis

  - Summarizing categorical, numerical, bi-variate and time series data

  - Quantiles and Normal Quantile Plots, Skewness and Kurtosis

  - Visualization techniques for complex and high dimensional data: Contour plots, 3D plots, matrix plots, Parallel coordinates, Multidimensional scaling, Pivot tables, Slice and dice

- Discovering Statistics Using R

  - Data exploration and representation in R

  - Reading data from various sources and writing data

  - Data wrangling in R

  - Loops/if/while and other control-flow constructs

  - Basics of writing R functions

- Random Variables and Probability Distributions

  - Conditional Probability

  - Discrete and continuous random variables and their distributions

  - The expectation and variance of a discrete and continuous random variables

- Estimation

  - Estimation methods

  - Properties of estimators

  - Confidence intervals

  - Pivotal quantity

- Hypothesis Testing

  - Terminology and notation

  - Types of hypotheses and types of errors

  - Size and power of tests

  - Cross-tabulations and tests of association

  - Goodness-of-Fit tests

- Data Simulation and Re-sampling Methods

  - Generating data from distributions

- Evaluating statistical procedures and understanding large sample results

- Introduction to permutation and bootstrapping

- cross-validation and their uses

- Bootstrap confidence intervals and hypothesis tests

- Introduction to Monte Carlo methods

- Regression and the Analysis of Variance

  - Inference for simple linear regression

  - Building the ANOVA table

  - Assumptions behind ANOVA

  - Assessing the assumptions and associated diagnostics

- Use of R Markdown for Analyses and Reports

Important Dates: These dates are tentative and subject to change at the discretion of the instructor and/or based on the learning needs of the students but such changes are subject to Section 2.8 of the ROASS Procedure.

	Preliminary Schedule
Date	Task
Sep 6	First class in 208 Armes
Sep 10,11,12	First Lab for sections B01, B02 and B03 in 311 Machray
Sep 27	Assignment 1 Due
Oct 11	Assignment 2 Due
Oct 15,16,17	Lab Quiz 1
Oct 30	In class Mid-Term Test
Nov 8	Assignment 3 Due
Nov 13 – 16	Fall Term Break
Nov 19	Voluntary Withdrawal (VW) Deadline
Nov 19, 20, 21	Lab Quiz 2
Dec 4	Assignment 4 Due
Dec 6	Last Class – Review

Other Important Information:

- **Academic Dishonesty:** It is important that you understand what constitutes academic dishonesty and that you are familiar with the very serious consequences. Links to resources that describe academic dishonesty (including plagiarism, cheating, inappropriate collaboration and examination impersonation, as well as typical penalties) can be found at: www.umanitoba.ca/science/undergrad/resources/webdisciplinedocuments.html.

- **Student Accessibility Services:** If you are a student with a disability, please contact SAS for academic accommodation supports and services such as note-taking, interpreting, assistive technology and exam accommodations. Students who have, or think they may have, a disability (e.g. mental illness, learning, medical, hearing, injury-related, visual) are invited to contact SAS to arrange a confidential consultation. The details can be found at http://umanitoba.ca/student/saa/accessibility/.

- **Copyrighted Material:** All course notes, assignments, tests, exams, computer codes and solutions are the intellectual property of your instructor or the Department of Statistics. Reproduction or distribution of these materials is strictly forbidden without their consent.

- **Recording of Class Lectures:** Your instructor and the University of Manitoba hold copyright over the course materials, presentations and lectures which form part of this course. No audio or video recording of lectures or presentations is allowed in any format, openly or surreptitiously, in whole or in part without permission from your instructor.

- **Use of Electronics in the Classroom:** It is the general University of Manitoba policy that all technology resources are to be used in a responsible, efficient, ethical and legal manner. A student may use technology in the classroom setting only for educational purposes approved by the instructor and/or the University of Manitoba Student Accessibility Services.

- **ROASS Schedule:** Schedule A of the Responsibilities of Academic Staff with regards to Students (ROASS) policies of the University of Manitoba lists resources and policies for students. It is important that you familiarize yourself with these resources and policies. Schedule A is available at http://www.stats.umanitoba.ca/files/pages/2016/09/Schedule-A-ROASS-Statistics.pdf.