

STAT4604A [35334] Statistical Computing (LEC) Fall 2020

From the calendar:

Statistical Computing (Honours)

Statistical computing techniques, pseudo-random number generation, tests for randomness, numerical algorithms in statistics; optimization techniques; environments for data analysis, efficient programming techniques; statistics with mainstream software.

Includes: Experiential Learning Activity

Prerequisite(s): STAT 3553 or STAT 3503 or permission of the School.

Lectures three hours a week, laboratory one hour a week.

Lectures will be at least somewhat synchronous, students are expected to be online during class time.

Instructor

Dr. Dave Campbell,

Professor

School of Mathematics and Statistics

Tentative Marking Scheme:

Assignments 30%

Weekly in-class quizzes: 20%

In class tests: 20%

Final Project: 30%

There will not be a final exam

Textbook:

Grolemund and Wickham, (2016) "R for Data Science"

<https://r4ds.had.co.nz>

Fall term is unusual at Carleton, See the University plan for the term:

<https://students.carleton.ca/2020/05/plans-for-the-fall-2020-term-timetable-and-scheduling-tool-will-open-on-june-8/>

Schedule:

Wednesdays and Fridays 4:05-5:25pm. See the course page on culearn.carleton.ca for class login details and resources.

Office Hours:

Immediately before class Wednesdays and Fridays. Other times are possible to accommodate student schedules and different time zones.

For best results, use culearn to contact the instructor.

Software:

R: <https://muug.ca/mirror/cran/> (install R first)

Rstudio: <https://rstudio.com/products/rstudio/> (install this after you have installed R)

Approximate course outline;

Week 1-3:

R for Data Science Chapter 1-5

ggplot, dplyr, intro to tidyverse, exploratory data analysis

Week 4:

Chapter 21

RMarkdown and reproducible workflows.

Week 5-7:

Chapter 7-9

tidyverse, wrangling diverse data types

Week 8-9:

Data acquisition and cleaning, scraping data tables from html,
Regular expressions, stringr

Week 10:

Permutation tests

Bootstrap

Week 11:

For loops, parallel for loops, and avoiding for loops

Week 12:

Maximum likelihood estimation

Week 13:

Optimization techniques