STAT3503A [35311] Regression Analysis, Fall 2020

From the calendar:
Review of simple and multiple regression with matrices, Gauss-Markov theorem, polynomial regression, indicator variables, residual analysis, weighted least squares, variable selection techniques, nonlinear regression, correlation analysis and autocorrelation. Computer packages are used for statistical analyses.

Includes: Experiential Learning Activity
Precludes additional credit for STAT 3553.

Prerequisite(s): i) STAT 2509 or STAT 2607 or ECON 2202 or ECON 2220 or equivalent; and ii) MATH 1102 or MATH 1107 or MATH 1119 or equivalent; or permission of the School.
Lectures three hours a week and one hour laboratory.

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

Textbooks:
http://users.stat.ufl.edu/~rohitpatra/4210/KNNL.pdf

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:
Tuesdays and Thursdays 8:35-9:55am. See the course page on culearn.carleton.ca for class login details and resources.

Office Hours:
Immediately after class Tuesdays and Thursdays. Other times are possible to accommodate student schedules and different time zones.
For best results, use culearn to contact the instructor and TA

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 1-2:**
**R for Data Science Chapter 1-3**
ggplot, dplyr, and intro to tidyverse:

**Week 3-4:**
**Applied Linear Statistical Models Chapter 1, 2**
Linear regression, correlation

**Week 3-4:**
**Applied Linear Statistical Models Chapter 5, 6, 7**
Gauss-Markov Theorem, Regression review with matrices

**Week 5-7:**
**Applied Linear Statistical Models Chapter 8-10**
polynomial regression, indicator variables,
Model selection
Why stepwise regression methods should never be used despite appearing in section 9.4 of this old textbook.
Better model selection alternatives

**Week 8-9**
**Applied Linear Statistical Models Chapter 11**
weighted least squares

**Week 10-13**
**Applied Linear Statistical Models Chapter 12, 13**
Nonlinear regression, autocorrelation, getting fancy