



Carleton
University

Department
of Economics

ECON 3880 - An Introduction to Economic Data Science

Summer 2024

Instructor: Firouz Fallahi
e-mail: FirouzFallahi@cunet.carleton.ca
Lecture hours: 6:05–9:00 p.m. on Tuesdays and Thursdays (online)
Office hours: After the class or by appointment - Online
TA: TBA
Office hours: TBA

Welcome to Data Science!

Course Description

This course introduces you to fundamental skills and practical training necessary to analyze complex data and carry out research using data science. Data science contributes to knowledge generation, better decision making, and turning data into insight. In this course, you will learn how to use R programming language to collect, process, and analyze data, as well as how to estimate statistical models and present the findings. This course provides a hands-on learning opportunity to learn and use R for data manipulation, analysis, and visualization. R is an open source language of programming widely used in academia and industry by economists and other professionals and can be used for classical techniques, artificial intelligence (e.g. machine learning), and also big data analysis. The course is designed for economics students with no or minimal programming experience; however, knowledge of introductory statistics and elementary mathematics is required. This course enables you to learn and practice skills that will help you with employability after graduation.

Learning Outcomes

By the end of this course, you will learn and demonstrate competency in programming tools and concepts in R and should be able to create R codes to do the following:

1. access, import, and clean data
2. organize, manipulate, and transform data into an appropriate format
3. calculate descriptive statistics of data
4. perform statistical analysis and estimate models
5. plot and visualize data and findings
6. learn about selected libraries in R
7. explore and gain ability to use various sources of economic data
8. produce reports in R Notebook
9. work with different data types including text, time series, and geospatial data

Prerequisites

The prerequisites for this course are ECON 1002 or ECON 1000 (or equivalent), ECON/MATH 1401 (or equivalent), and ECON 2210 (or equivalent) with a grade of C- or higher. Students who believe they have taken a similar background course or courses from another university must provide appropriate documentation to the Department of Economics Undergraduate Administrator.

Organization

This is an online course. It will be delivered **SYNCHRONOUSLY** through Zoom and will not be recorded. Students will be required to have access to a computer/laptop and Internet in order to successfully engage in and complete the course.

Please note that Brightspace and the Carleton email system will be used extensively as a means of communication. Therefore, students are strongly advised to access Brightspace and to check their Carleton email at regular intervals in order to check for new information. To access Brightspace and the Carleton email system, students require a MyCarletonOne account. For questions about MyCarletonOne accounts, students should access carleton.ca/its/get-started or contact the ITS Service Desk <https://carleton.ca/its>

Textbooks

There is no required textbook, as the course material is drawn from a variety of sources, including the books listed below. Lecture slides/notes and the code developed during the lecture will be posted after each lecture.

1. *R for Data Science* by Wickham, H., Cetinkaya-Rundel, M., & Golemund, G. (2023). Free e-book is available at <https://r4ds.hadley.nz/> . If you prefer, you can purchase a PDF copy online or a hardcopy through Amazon
2. *Using R for Introductory Econometrics* by Florian Heiss (2020). Available for free at <http://www.urfie.net>. If you prefer, you can purchase a hardcopy through Amazon
3. *Introduction to Econometrics with R* by Hanck, C., Arnold, M., Gerber, A., and Schmelzer, M. (2024), available at <https://www.econometrics-with-r.org>

Very useful websites and packages

- <https://swirlstats.com> Swirl is an R package that teaches R programming and data science interactively right in the R console
- <https://www.r-bloggers.com/>
- <https://r-coder.com>
- <https://moderndive.com>
- <https://stackoverflow.com/> This is an online community for programming questions and answers to learn and share knowledge
- <https://www.kaggle.com/> Machine learning and data science community

Software

You will be required to use R and RStudio in this course. R is an open source language of programming and can be downloaded from <https://cran.r-project.org/>. RStudio IDE (integrated development environment) is a user friendly environment for developing and running R code. It also is an open source software and is available at <https://www.rstudio.com> or <https://posit.co/>. Setup information will be covered in the first session of the course.

Course Outline

The following outline is tentative and may change depending on the pace of the lectures.

| Topics | Readings | Assignment |
|--|--|--------------|
| <ul style="list-style-type: none"> • Introduction to R, RStudio, and R Markdown/Notebook | R for Data Science (R4DS): Introduction https://rmarkdown.rstudio.com/lesson-1.html https://bookdown.org/yihui/rmarkdown/notebook.html | |
| <ul style="list-style-type: none"> • Data visualization with ggplot2: <ul style="list-style-type: none"> ○ Plotting data ○ Adding layers ○ Faceting and customizing plots | R4DS Ch. 1, 9, 10 | Assignment 1 |
| <ul style="list-style-type: none"> • Data wrangling in with dplyr/tidyverse: <ul style="list-style-type: none"> ○ Importing and inspecting the data ○ R data structures (vector, matrix, data frame, and list) ○ Subsetting and creating new data frame ○ Creating new variables ○ Data type conversion ○ Merging, ○ Reshaping and transforming data ○ Pivoting data | R4DS Ch. 3, 5, 7, 19 | Assignment 2 |
| <ul style="list-style-type: none"> • Data summarization and cross tabulation: <ul style="list-style-type: none"> ○ Descriptive statistics (mean, median, variance, covariance and correlation) ○ Filter and select data ○ Mutate new variable ○ Grouped summaries | R4DS Ch. 3 | Assignment 3 |
| <ul style="list-style-type: none"> • Intro to Monte Carlo simulation | Heiss section 1.10 | |
| <ul style="list-style-type: none"> • Regression analysis: <ul style="list-style-type: none"> ○ Linear regression (univariate and multivariate) ○ Basics of logistic regression | Heiss section 2.1, 2.2, 3.1, 17.1 Hanck <i>et al</i> Ch. 4, 6, 11 | Assignment 4 |
| <ul style="list-style-type: none"> • Selected data sources in economics: <ul style="list-style-type: none"> ○ World Development Indicators ○ Statistics Canada's Socioeconomic Data (<i>formerly CANSIM</i>) ○ Statistics Canada's survey data | | Assignment 5 |
| <ul style="list-style-type: none"> • Brief introduction to selected advanced topics including GitHub, Text mining, and Shiny App (if time permit) | | |

Evaluation

There will be **five** assignments and a final exam for this course. The overall grade will be constructed as follows:

| Evaluation Format | Weight | Date |
|---------------------|--------------|--------------------------------------|
| Class Participation | 10% | |
| Assignments | 4 @ 15% each | Due one week after they are assigned |
| Final exam | 30% | Scheduled by the Registrar's Office |

- ✓ The **lowest** grade of the five assignments **will be dropped** from calculation of the course grade and there will be no deferred assignments. In addition, late submissions will not be graded.
- ✓ For active learning, you may work in groups of up to 3 members to complete the assignments and submit one solution for the group. All assignments must be prepared using R Markdown/Notebook and submitted as a knitted HTML files electronically.
- ✓ Assignments will mainly consist of exercises designed to get you more comfortable in using R for data analysis.
- ✓ You are not allowed to post the assignment questions on the Stack Exchange, R help forums, or the like.
- ✓ The final exam will be online during the examination period and will be scheduled by Examination Services. In addition, it will be a cumulative, open note, open book, and open internet exam; however, you are not allowed to communicate with anyone except the instructor and TAs.
- ✓ The deferred final exam, should one be required as per the university regulations, will be written on a date determined by the University. An application to write a deferred final exam must be made at the Registrar's office.
- ✓ Please note that "Standing in a course is determined by the course instructor subject to the approval of the Faculty Dean. This means that grades submitted by an instructor may be subject to revision. No grades are final until they have been approved by the Dean."

Notes

- ✓ You are expected to attend and participate in all classes - this helps you learn better and also achieve better grades.
- ✓ For verification and security purposes, I will not reply to emails originating from non-Carleton e-mail accounts, and please include the course name “ECON 3880” in the subject line of your email.

Academic Integrity

Please ensure that you are familiar with and comply with the Carleton University Academic Integrity Policy:

<https://carleton.ca/secretariat/wp-content/uploads/Academic-Integrity-Policy-2021.pdf>

- **Plagiarism**

Please be aware that plagiarism is serious offense at Carleton and should be recognized and avoided. For more details, see the link above.

Requests for Academic Accommodation

You may need special arrangements to meet your academic obligations during the term. For an accommodation request, the processes are as follows:

- i. **Pregnancy obligation**

Please contact me with any requests for academic accommodation during the first two weeks of class, or as soon as possible after the need for accommodation is known to exist. For more details, visit the Equity Services website:

carleton.ca/equity/wp-content/uploads/Student-Guide-to-Academic-Accommodation.pdf

- ii. **Religious obligation**

Please write to me with any requests for academic accommodation during the first two weeks of class, or as soon as possible after the need for accommodation is known to exist. For more details, see the link above.

- iii. **Academic Accommodations for Students with Disabilities**

If you have a documented disability requiring academic accommodations in this course, please contact the Paul Menton Centre for Students with Disabilities (PMC) at 613-520-6608 or pmc@carleton.ca for a formal evaluation or contact

your PMC coordinator to send me your Letter of Accommodation at the beginning of the term. You must also contact the PMC no later than two weeks before the first scheduled test or exam requiring accommodation (if applicable). After requesting accommodation from PMC, contact me as soon as possible to ensure accommodation arrangements are made. For more details, visit <https://carleton.ca/pmc/>

iv. Survivors of Sexual Violence

As a community, Carleton University is committed to maintaining a positive learning, working and living environment where sexual violence will not be tolerated, and its survivors are supported through academic accommodations as per Carleton's Sexual Violence Policy. For more information about the services available at the university and to obtain information about sexual violence and/or support, visit: <https://carleton.ca/equity/focus/sexual-violence-prevention-survivor-support/>

v. Accommodation for Student Activities

Carleton University recognizes the substantial benefits, both to the individual student and for the university, that result from a student participating in activities beyond the classroom experience. Reasonable accommodation must be provided to students who compete or perform at the national or international level. Please contact me with any requests for academic accommodation during the first two weeks of class, or as soon as possible after the need for accommodation is known to exist. See also <https://carleton.ca/senate/wp-content/uploads/Accommodation-for-Student-Activities-1.pdf>

If you have problems with this course and need help, please let me know!