Statistical Computing with R

Professor Nancy Zhang

The goal of this course is to introduce students to the R programming language and related ecosystem. This course will provide a skill-set that is in demand in both the research and business environments. R is a platform that is used and required in other advanced classes taught at Wharton, so this class will prepare students for these higher level classes and electives.

Materials:

There are no textbooks for this course.

Lecture slides and other notes are in Files/Notes

R code introduced in lectures are in Files/Rscripts

Assignments and Grading:

4 homeworks and 1 project. All homeworks are graded and have equal weight. Homeworks will be posted online each Wednesday, and due the following Wednesday at midnight. Homeworks are submitted and graded online. Final grade is based on HW1+HW2+HW3+HW4+2*(Final Project).

You are encouraged to work together on the homeworks, but each student must submit independently written solutions. The tentative assignment and due dates are:

Assignment	Assign Date	Due Date
Homework 1	9/6	9/13
Homework 2	9/13	9/20
Homework 3	9/20	9/27
Homework 4	9/27	10/4
Final Project	10/4	10/18

Tentative lecture plan:

Lec. #	Date	Topics
1	30-Aug	What is R? Rstudio, help functionality, first look: data types, simple operations
2	6-Sep	Data structures: vectors, matrices, arrays, lists; subsetting, slicing and dicing
3	11-Sep	Dataframes, reading data into R, tables and sorting
4	13-Sep	More on data frames: combining data from different sources. Linear regression in R.
5	18-Sep	Logistic regression in R
6	20-Sep	Writing functions I
7	25-Sep	Writing functions II
8	27-Sep	Logic and flow control structures I
9	2-Oct	Logic and flow control structures II
10	4-Oct	Simulations in R: distributions and sampling
11	9-Oct	The R Ecosystem: working with packages. The ggplot graphics package.
12	11-Oct	Reproducible research and generating live reports in R
13	16-Oct	Case study and in-class activity
14	18-Oct	TBA