STAT405: Statistical Computing with R

Location: 351 SH-DH

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.

Instructor: Professor Alexander Rakhlin   (rakhlin@wharton.upenn.edu)
TA: Matthew Olson   (maolson@wharton.upenn.edu)

There are three sections: 12-1:30 pm, 1:30-3:00 pm, and 3:00-4:30pm on Mondays and Wednesdays in 351 SHDH. Please attend the appropriate section.

Materials:

There are no textbooks for this course.
Lecture slides and other notes are in `Files/Notes`; R code introduced in lectures is in `Files/Rscripts`

Assignments and Grading:

There will be 4 homeworks and an in-class exam. 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*Exam.

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

	Assigned	Due
Homework 1	January 17	January 24
Homework 2	January 24	January 31
Homework 3	January 31	February 7
Homework 4	February 7	February 14

Tentative lecture plan:

1	Jan 10	What is R? Rstudio, help functionality, first look: data types, simple operations
2	Jan 17	Data structures: vectors, matrices, arrays, lists; subsetting, slicing and dicing
3	Jan 22	Dataframes, reading data into R, tables and sorting
4	Jan 24	More on data frames: combining data from different sources. Linear regression in R.
5	Jan 29	Logistic regression in R
6	Jan 31	Writing functions I
7	Feb 05	Writing functions II
8	Feb 07	Logic and flow control structures I
9	Feb 12	Logic and flow control structures II
10	Feb 14	Simulations in R: distributions and sampling
11	Feb 19	The R Ecosystem: working with packages. The ggplot graphics package.
12	Feb 21	Reproducible research and generating live reports in R
13	Feb 26	Review session
14	Feb 28	In-class exam