

## Programming in R Syllabus

**Instructors:** Jaime Ramirez-Cuellar - Daniel Cullen - Travis Cyronek  
**Office:** North Hall 2119  
**Contact:** jramirezcuellar@ucsb.edu - dcullen@ucsb.edu - travis\_cyronek@ucsb.edu  
**Classroom:** North Hall 1122 - EY Lab  
**Dates:** September 24, 25, 28 and October 5  
**Time:** 9/24, 9/25: 10 am - 1 pm. 9/28: 10 am - 4 pm, 10/5: 8 am - 6 pm.  
**Website:** [jaimeramirezcuellar.wordpress.com/teaching](http://jaimeramirezcuellar.wordpress.com/teaching)  
[dancullen.me/teaching](http://dancullen.me/teaching)

## Course Description

The UCSB Department of Economics recognizes the growing need for proficiency in data analytics in the workforce and for undergraduates to be able to participate in research. To supplement students coursework, the Department has decided to host the first ever Data Hack 2018: Data Analytics Boot Camp & Hackathon.

In the competition, students will learn to work with data in the programming language R. They will build their resumes with a desirable and applicable skillset. These skills will prepare them for research opportunities with professors. Finally, they will apply the skills learned in a competition with prizes.

Participants must be in either the Economics or Economics & Accounting majors and must have completed ECON 140A. They must complete the Boot Camp in order to participate in the competition.

## Recommended Materials

These materials are recommended. **Students are not required to purchase any of these materials.**

- [www.datacamp.com](http://www.datacamp.com)
- Kleiber, C., & Zeileis, A. (2008). Applied econometrics with R. Springer Science & Business Media. (online UCSB Library)
- Varian, H. R. (2014). Big data: New tricks for econometrics. Journal of Economic Perspectives, 28(2), 3-28.
- Wickham, H. (2016). ggplot2: elegant graphics for data analysis. Springer. (online UCSB library)
- Wilkinson, L., & Wills, G. (2005). The grammar of graphics (2nd ed., Statistics and computing). New York: Springer. (online UCSB library)

## Topic List and Schedule

### Introduction to R

- Using R Studio. Save Projects, set working directory, organize files and directories. R Objects and variables types. Loops and functions

### Data Management in R

- Reading data into R. Merge and append data. Create new variables. Subset data. Summary Statistics.

### Data visualization and regression

- Graphics with the **base** package. Introduction to the **ggplot2** package. The grammar of graphics. Data, aesthetics, geometries, facets, statistics, coordinates and themes.
- Common plots: Scatter plots (**geom\_point**), bar plots (**geom\_histogram**), line plots (**geom\_line**), boxplots (**geom\_boxplot**), density plots (**geom\_density**).
- Regression. Formula notation. Marginal effects. R squared. Hypothesis testing.

## Competition

Students will receive two handouts with instructions for the competition. The first handout contains general instructions and will be distributed on the last day of class. The second handout will be distributed at 8 am on the day of the competition.

## General Schedule

Table 1: General Schedule (Tentative)

Day	Time	Activity	Instructor
Sep 24	10 AM - 1 PM	Introduction to R	DC
Sep 25	10 AM - 1 PM	Data Management in R	JRC
Sep 28	10 AM - 4 PM	Data Visualization and Analysis in R	DC, JRC, TC
Sep 28	3:30 PM	Students receive general topic	DC, JRC, TC

Notes: DC = Daniel Cullen, TC = Travis Cyronnek, JRC = Jaime Ramirez-Cuellar