Lab Session Econ 220  
Probability and Statistics for Economists - Lab  

Fall 2018

Time and Location:  
LAB 1 F: 9:00 am – 9:50 am, White Hall 110  
LAB 2 F: 10:00 am – 10:50 am, White Hall 110  
LAB 3 F: 3:00 pm – 3:50 pm, White Hall 102

Note: You may only attend the section in which you are registered; all sections are full and we cannot have an overcrowded classroom for any of the lab sessions.

Office Hours: My regular office hours for the Fall Semester are Friday, 11:30am – 2:30pm.

Questions: You can come to my office hours or ask questions at the end of the class if we finish early. However, good programming skills come with a lot of self-teaching and fixing your code on your own – so the rules for asking questions are the following:

1- First see if R itself has the answer to your question (In the first class you will learn how to use R help)  
2- If not solved by (1), look for online resources (these will also be provided in the introductory classes)  
3- You can simultaneously ask your colleagues using the Canvas discussion board.  
4- If after trying these you couldn’t find an answer to your question, we can look at it together. But you must prove that you tried 1-3 first. I will reply to emails within a 24 hours window during weekdays.

I will be continuously monitoring Canvas and I will discuss some of the questions at the beginning of each class and give participatory points for solving questions on your own.

Textbook: There is no required textbook for this class. We will create the notes for the class in each lecture. I use “Discovering Statistics with R” by Andy Field as guidance book for some topics but I also rely on online sources and I will reference those in each class.

IMPORTANT: You need to bring your laptop for every class. This will allow you to run your code in class, make sure that you are able to replicate what we do in class on your own, and take notes on your code.

Discussion Board: This term we will be using the Canvas discussion board for class discussion. The system is highly catered to getting you help quickly and efficiently from classmates and/or myself. Rather than emailing questions to the teaching staff, I encourage you to post your questions on the
The discussion and help etiquette will be on the discussion board as a pinned discussion.

**Honor Code:** The honor code is in effect throughout the semester. By taking this course, you affirm that it is a violation of the code to cheat on exams, to plagiarize, to deviate from the teacher’s instructions about collaboration on work that is submitted for grades, to give false information to a faculty member of the TA, and to undertake any other form of academic misconduct. You agree that the teacher is entitled to move you to another seat during examinations, without explanation. You also affirm that if you witness others violating the code you have a duty to report them to the honor council.

**Grading:** This lab represents 25% of your grade for Eco 220. I will provide a total grade to your Statistics instructor. I will do random quizzes in class, and participation in Canvas and in class will be part of your grade. You will have to submit a final project at the end of the semester in which you apply the skills learned in the lab using a real dataset.

<table>
<thead>
<tr>
<th>Grade distribution</th>
<th>% of your total grade</th>
<th>% of your lab session grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quizzes/Homework and participation</td>
<td>10%</td>
<td>40%</td>
</tr>
<tr>
<td>Final Project</td>
<td>15%</td>
<td>60%</td>
</tr>
</tbody>
</table>

Attendance is not required but it is highly recommended that you don’t miss any class. Quizzes and participation points will be randomly given during class.

**Important dates:**

**October 12th:** Choose the dataset for your final project, submit the source of your dataset and the code that shows that you have access to the data and are able to load it into R. Details on Canvas.

**December 7th:** Final project submission.

**Canvas:** I will be using Canvas for all the course information. Please make sure you check it every day and get the notifications. All the relevant information for the course will be there.

**Course Description.** Data analysis is a highly valued skill in the job market. In this class, you will learn how to organize and analyze data using computer-programming software. We will focus primarily on the R language, which has powerful tools to manage and analyze data. We will develop the skills to apply the concepts you learn in class and present data and analysis in an organized manner. We will also learn how to think algorithmically within the R environment, with the aim to prepare you for more advanced data analysis tasks in your future classes (Econ 320) or in other endeavors. Your programming skills will only improve when you struggle to figure things out on your own. This class will provide the guidance to develop the basics of R for statistical analysis and, hopefully, the desire to keep practicing and challenging yourself with new and different quantitative problems and questions.
Topics

I will list the topics that we will cover during the semester. Each topic will take one to three classes.

- Introduction to R programming
  - The basics of R:
  - Write an R script, create a working directory
  - Objects in R: Variables, vectors, matrices, data frames and lists

- Upload and manage data
  - Importing data into R
  - *Tidyverse package*
  - *dplyr* package for data visualization
  - Simple tables in R

- More on data visualization: Graphs in R
  - *ggplot2* package
  - Histograms

- Distributions in R
  - Sampling from a population of different distributions
  - Histograms from random samples
  - Descriptive statistics: Mean, Variance, Covariance, Correlation

- Central limit theorem and Monte Carlo Simulation

- Hypothesis testing