

Computational Methods Syllabus

Computer Camp

Nick Eubank & Claire Evans

July 26, 2018

Hello and welcome to Vanderbilt!

1 Learning Goals

By the end of Computer Camp, our goal is for you to be able to:

- Load data in R
- Manipulate the data in basic ways (like tabulating results)
- Find help online when you get stuck in R
- Compile basic LaTeX documents
- Choose good tools for organizing papers, checking algebra, etc. during your first year.

2 Required Programming Background

None.

We have *absolutely no expectations that students will have any experience with programming!* Computer Camp is required for all students, and we fully recognize that many students have never worked with statistical software. That's FINE! If you know how to use google and email, you have plenty of experience – everything else we'll take care of.

At times during the course, we may make statements about how the tools we're learning (like R) compare to different tools (like Stata or Python) if we find there are students with experience in these other tools. However, if we make those comparisons, it is only to help those students the traps one can fall in if one's background is in another language. It is in no way because we *expect* all students to have experience with other tools.

3 R

We'll spend most of Computer Camp teaching you to use a program for statistical analysis called *R* (yup, just the letter).

Why R? Because it's currently the most-used statistical software in political science, and using the same tool as those around you will make it easier to collaborate and get help.

It is worth emphasizing that we're not teaching you R because we think it is the best. The reality is that there are lots of tools for statistical programming, and each has its own strengths and weaknesses (e.g. R, Stata, SPSS, Python, Julia, Matlab, etc.). People develop really strong opinions about what language is *best*, and sometimes pass judgement on people who use other languages. We would like to discourage this type of thinking. Personally, I (Nick) regularly work in at least four different programming languages depending on which is best suited to the task at hand, so I think I have reasonable authority to say: there is no single *best* language for all purposes.

As a result, over the course of your career you may find yourself gravitating to one tool or another as required by your research. But in providing you with a firm foundation in a very popular language like R, we feel confident that we will not only be providing you with tools that will allow you to do most everything you'll want to do in graduate school, but we will also be providing you with *generalizable* skills around data manipulation that you will find useful if you later change platforms.

4 Class Organization

In this class we will be “flipping the classroom” – you will be required to review tutorials on R between classes, and each afternoon we will focus on doing exercises that allow you to get hands on experience with the skills you've read about in an environment where help will be available. These tutorials aren't very long, and we **strongly** recommend that while you read through them you do so with an open R session so you can just play around a little, trying out the things you learn. The research on learning to program is exceedingly clear on this point: **the only way to learn to program is to actually program**, so the more time you spend playing with R, making mistakes, and troubleshooting, the more you will learn. We'll do lots of exercises in class, but the more you play on your own too, the more you will learn.

In particular, we will primarily relying on a set of excellent R tutorials written by Simon Edjmejr which you can find at:

- <https://sejdemyr.github.io/r-tutorials/basics/>

as well as a few supplementary tutorials from Kelly Black which you can find at:

- <https://www.cyclismo.org/tutorial/R/>

5 Schedule

Class 1: Getting Started

Thursday, August 9th

- Discuss course goals
- Do baseline knowledge assessment
- Install R on personal laptops
- Look at tutorial pages!

Homework to be done before Class 2:

- Edjemyr Tutorial: Intro to R
We'll go over most of this material on the first day, but please skim it at home for completeness.
- Edjemyr Tutorial: Vectors
- Edjemyr Tutorial: DataFrames

Don't worry! This will be BY FAR the biggest set of homeworks. Other nights will be better, but it's hard to do much until we've covered all this.

Class 2: Basic R Manipulations

Friday, August 10th

- read in CSV
- look at data, play with couple data types, get summary stats.

Homework to be done before Class 3:

- Black Tutorial Indexing
- Edjemyr Tutorial: Merging
- Black Tutorial: OLS

Class 3: Working with Multiple Data Sources

Monday, August 13th

- Merge datasets
-

Homework to be done before Class 3:

- Edjmeyr Tutorial: ggplot

Class 4: Ad Hoc Day

Tuesday, August 14th

- ?

Wednesday, August 15th: NO CLASS

Class 5: LaTeX

Thursday, August 16th

- Teach latex!

Class 5: Tools You'll Want in Profession

Friday, August 17th

- Wolfram Alpha
- Mendeley / Papers