

MATH 145: Biostatistics

Assignment to complete for Tues., Feb. 4, 2020

Things to do before our next class meeting:

- **Read** Section 1.1 from the Lock5 text. Focus on concepts, particularly on being able to identify when an object does/does not fit a certain definition. From the “Exercises for Section 1.1”, do numbers 2 (marked as 1.2, in the text), 4, 6, 8, 10, 15 and 17. (Be ready to discuss your responses on Tues.)
- Import into RStudio the file found at

<http://www.calvin.edu/~scofield/data/csv/ssurv.csv>

which contains data collected (mostly) from Calvin students back in 2004.

- Identify which are the categorical variables.
- Pick some pair of (two) categorical variables, for which you think it would be interesting to know if an association exists between them. Be able to name the variables, and designate which one you are thinking of as the explanatory variable.
- Find some data of interest to you, data which includes cases and variables in table format. (You are welcome to find yours via the internet.) Choose a set that is not overly large (neither an abundance of cases nor of variables, as you are asked to enter this data into a spreadsheet yourself), but which has at least two variables. Enter this data into a spreadsheet program, and save it to your computer as a .csv file.

Upload this file to your workspace on the RStudio server. You will need to be logged into the RStudio server. Then look for the “Upload” option on the “Files” tab. Finally, use the `read.csv()` command to bring the data into RStudio. Your command will likely need to look something like this:

```
read.csv("/home/tls22/myDat.csv")
```