Assignment 7

gzahn

December 30, 2018

Assignment 7

In this assignment, you will use R (within R-Studio) to:

- Use the tidyr package to convert "wide" data to "long" data using a simple example
- Do the same for a more complex data set
- Use the tidyr package to convert the instructor's base R code into tidy format, making it more readible
- Begin to look at correlations and causations using a real data set of religious identity from Utah in 2010

All file paths should be relative, starting from the Assignment_7 directory!! (where you found this file)

This means that you need to create a new R-Project named "Assignment_7.Rproj" in your Assignment_7 directory, and work from scripts within that.

For credit...

1. Push a completed version of your Rproj and R-script (details at end of this assignment) to GitHub
2. Your score will also depend on whether any files generated in this workflow are found in your repository
3. Submit your answers to the questions in part 2 of Assignment_6_messy_code.R as plain text in Canvas

---

Your tasks:

- Using the dplyr verbs, change the code found in PART 1 of the file "Assignment_7_messy_code.R" into a version with improved readibility. Some of the code will stay the same as the original "messy" version, but most of the code chunks can be done better using dplyr in a "tidy framework."
- use things like: arrange() desc() group_by() summarize() filter() select() %>%
- Run the code in PART 2 of the same R script to generate correlation plots and answer the 5 questions (on Canvas)

Here's a reminder of how gather() can be used...

```
library(tidyr)
library(dplyr)
utah <- read.csv("Data/Utah_Religions_by_County.csv")
names(utah)
utah_long = gather(utah,key = Religion, value = Proportion, -c(1:3) )
```

And don't forget about the amazing combination of group_by() %>% summarize()