Week 4 Instructions and Checklist

Stat 331

Data Analysis Project

Collaboration

(40 points)

As this is a *group* project, you are required to meet **in person** with your group *at least twice*.

Upload evidence of your presence at these **face-to-face** meetings to the Canvas assignments.

Your Writeup

(200 points)

Introduction (40 points): Introduce your chosen dataset. Who created it, and how? What information does it contain? Provide some relevant summary information about some of the variables. This summary should include at least 2 plots.

Analysis (100 points): Come up with at least three interesting insights into the dataset. Insights can be something like "The Northeast has more private colleges than the West" or "Cereals on the first shelf tend to be more sugary". Support each of your insights with at least one plot, with appropriate summary calculations, and with a clear discussion.

Style (60 points) Your writeup should read like a fun article. Format your headers and text and chunks nicely, and don't display any code or output that does not flow with your narrative. Make sure plots have clear labels and titles, make sure your discussion is in complete sentences. You should not say things like "I chose to do my project on this dataset" - pretend you are writing this article for Buzzfeed, not for class. Feel free to search around for fun ways to make R Markdown output even prettier!

Peer Reviews

(60 points) Due: Sunday, Feb 2 at midnight

Review 3 of your peers projects on Canvas. Leave each project two compliments, and two suggestions for improvement.

Conference Videos

2019: Hilary Parker

Watch this talk from the 2019 RStudio Conference: https://resources.rstudio.com/rstudio-conf-2019/cultivating-creativity-in-data-work

Enter your discussion into the "quiz" on Canvas.

2020: Welcome and Keynotes

Best option: Attend the San Luis Obispo livestream party!

8am-11am on **Weds, January 29th**, on campus (room TBD)

This event includes:

- Breakfast
- Free RStudio swag
- A special Q&A session with RStudio employees before the talks.

This event is limited to 35 people - sign up at the meetup link to ensure your spot!

If you can not make this watch party, you can livestream the talks, or watch them later (I will post a link when it exists.)

Two more talks

Choose two talks from 2019 that interest you: https://resources.rstudio.com/rstudio-conf-2019

Or, you can livestream talks from this year's conference at https://www.twitch.tv/thomas_mock

2020 Conf Info

Watch them, and put your discussion into the quiz question on Canvas.

Other Assignments

Lab 2

due: Friday, Jan 24 at midnight *10% reduction if turned in by Sunday, Jan 26 at midnight 50% reduction after that*

PA 5

due: Sunday, Feb 2 at midnight

Housekeeping

Lab 0

If you missed the first set of questions, you may re-submit Lab 0 for **full credit**

Due: Friday, Jan 24 at midnight.

Lab 1

If you uploaded Lab 1 in the wrong format (recieving a score of 0 and a comment on Canvas), you may **email your Lab1.html file** to kelly.bodwin+stat331+lab1@gmail.com. If the time stamp on this file is before the original due date, you will not be deducted in points.

Due: Sunday, Feb 2 at midnight