

STAT 3280 Fall 2021 Homework 1

Due by the end of Sep 11, eastern time. Submit your homework by sending it to our GTA, Ruizhong Miao (rm9dd@virginia.edu), with the subject “STAT 3280-HW1: names”, where the “names” should be replaced by your **last name(s)** of the group. Each group only has to submit it once. Make sure you include everyone’s **name AND computing id on the first page**. **Missing any part of these will result in missing grades.** Please use **a separate page for each problem**. And the answer to each problem cannot be longer than one page (with reasonable font size, line space, margins etc.). You can explain how you did it in R by submitting your code with detailed explanations, but only include this part in an appendix. The GTA will not be guaranteed to look at your appendix, so make sure you explains things clearly in your main text. Notice that you are working on a visualization task. So, for each problem, make sure that your **plain language explanations should not exceed 1/2 of the paper in total** for each problem. The main results would be your figures. **You are required to use R basic graphics tools for this homework** – that means, do not use ggplot2 package or other extensions based on the it. Superb submissions for Q2 may earn invitations for presentations.

Total points: 10.

1. (5 pts) Use figures to give **exploratory analysis** of the two AddHealth datasets introduced in the lecture:
 1. (2 pts) Generate **one figure** exploring the relation pattern about the depression and anxiety and grade in `dt1`. Use one or two sentences to summarize your findings from the figure.
 2. (3 pts) Based on `dt2`, generate figures to explore the marijuana usage pattern cross gender and education level. Does the similar pattern hold if you check the Cocaine usage instead of marijuana?
2. (5 pts) Do you know that the North Carolina–Virginia football rivalry (UNC vs. UVa) is named the “South’s Oldest Rivalry”? The relevant background and historical data can be found on Wikipedia.

https://en.wikipedia.org/wiki/South%27s_Oldest_Rivalry#Other_sports

You can download the data in whatever way you want (e.g. simply copy-paste). Now, suppose you are invited to create a one-page graph summary for the entire head-to-head history as a UVa Daily poster. For this problem, you need to collect your data (from whatever resources, e.g., the Wiki page). With clearness and the space limit in mind, you want to include as much information as possible. Note that as a poster, you want to demonstrate some interesting patterns or observations. You can use sentences to explain things, but you should only use them parsimoniously (no more than 4 sentences) because

you need to tell the story and catch readers' eyes with your figures. Feel free to design your layout. You do not have to generate everything as one figure. A better way is to generate a few figures and then combine them with more straightforward tools such as Microsoft Word or Powerpoint, with concise explanations. Keep your readers in mind as well. (2 pts for basically correct information, 1 pt for exciting patterns or observations, 1 pt for clearness and conciseness, 1 pt for reasonably good designs.)