STAT 3280 Fall 2021 HW5

Due by the end of Oct 20, eastern time. Submit your homework by sending it to our GTA, Ruizhong Miao (rm9dd@virginia.edu), with the subject "STAT 3280-HW5: 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 can use any software or packages for this homework.**

Total points: 5 points.

1. (5 pts) We will introduce the way to generate the US airline plot in class and talk about the airport data. Now, based on the larger data set (`airport.zip`) on Collab, select the top 100 airports globally with the largest number of flights (in or out). Generate the global airline network data. You can also find more information about the data sets at `https://openflights.org/data.html`. Watch out for the arcs you draw. A flight always selects the shortest path between two airports, so that is what you want to plot. For example, you need to make sure that the flights crossing the Atlantic Ocean are corrected plotted. Plotting such flights to be crossing the Pacific Ocean will not be valid. More details in the lecture. (Correct setup of the world map 1 pt, correct setup of the airports 2 pts, correct setup of the airlines 2 pts.)