STAT 3280 Fall 2021 HW4

Due by the end of Oct 13, eastern time. Submit your homework by sending it to our GTA, Ruizhong Miao (rm9dd@virginia.edu), with the subject "STAT 3280-HW4: 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: 7 points.

1. (7 pts) The data set `NASA` contains 24×24 grid of central America atmosphere measurements from NASA. In the data folder, you can find the `readme.txt` which contains the basic information about the data set. There are several measurements and for this problem, you are supposed to work on the High Cloud Coverage. Design a nice one-page figure to visualize the High Cloud Coverage average over years for each of the 12 months (e.g. average January HCC over years, average February HCC over years etc.) over the whole spatial area, located by the latitude and longitude over a map background. That means, for each of the 24×24 locations, you have 12 numbers to show. And you need to show the information for all the 596 locations according to geographical locations. The data structure might be seemingly complicated, so we will go through the basic data processing together. But once you fully understand the DataStructure.R, things will be easier. We will also illustrate some basic ideas about how to visualize it in the lecture. **No words are allowed for this question except the necessary captions/labels/legend etc.** Due the large volume of data you need to show, it is acceptable if one has to zoom in to see all details clearly. But you need to ensure: 1) you do have all details visible after zooming in (hint: use pdf as your output figure); 2) without zooming in, the high-level trend should still be visible. (2.5 pts for all details available, 2.5 pts for visible high-level pattern, 2 pts for correct map layout)