Revised Proposal for Group B8

Group name: Blue Cookie

Title: The Climate of Boston over Time

Purpose

We would like to explore how the climate and temperature change in a given location over the years. Perhaps we could add more locations for comparisons. For example, we could analyze at Boston with questions such as: Is there a trend in temperature increasing? If so, how visible is the change? Or how much does the temperature vary over the year in an East Coast location vs a location on the West Coast? We can also look at other variables such as precipitation and weather conditions and see how they change depending on season/month.

We also want to explore if today (3/24) is the hottest/coldest day among all 3/24's (from 2010 to 2022). Currently, we are interested in addressing this question for Boston. However, if possible, we would expand our search to different cities as well.

Data

There are many government sites that contain information about the temperature and past temperatures. For example, the National Oceanic and Atmospheric Administration in accordance with the National Centers for Environmental Information contains a record of information that can be used. We will most likely be using downloading ".csv" files because the files with a large amount of information tend to be in this format. Any difficulties with wrangling are most likely going to be due to the large number of observations in the dataset and/or combining multiple datasets. We will be joining multiple files from multiple sources (because other sites have different climate information for cities). We will have to be joining multiple datasets from different cities into a large dataset that we can use for creating visuals for comparisons.

A challenge that we saw was dealing with the url. Unlike the Emily Dickinson's poems examples, there is no list of urls for us to utilize. Therefore, if we want to compare different cities max/min temperature of a chosen day, we will have to craft the urls ourselves. Although the links share the same pattern, the final element for each city is different. We suspect this will be very time consuming and prone to error. This is why our current proposal only focuses on Boston.

Shiny app

Most likely, we would do plots over time to show how our variable of interest changes over time. For example, if we do something with the temperature, we could have an interactive scatterplot where we can change the response variables. Additionally, we could look into comparisons between different cities where a user can input multiple cities from a pull-down menu and visualize the differences in a variable such as precipitation.

The scatterplot will have time as the explanatory variable and average temperature as the response variable. The pull-down menu of different cites would allow a user to see the relationship between time of the year and

temperature for multiple cities. This will be interactive because the user can select and deselect different cities to compare.

We could also do another plot that compares the different variables to each other. For example, we could use another scatterplot to plot the relationship between variables such as precipitation, temperature, weather conditions, etc where the user can select which variables to use as the explanatory variable and which to use as the response variable.

Another interactive widget we would make will be an interactive table using the shiny package, specifically the renderDataTable function. Picking Boston as our city of interest, we will built an interactive table. It will have a slider to determine the number of rows to display. It will also have filtering function (powered by the interactive table) so that users can find information of interest. This table will show entries for different variables of the chosen date (say 3/24), such as max/min temp, precipitation, etc. throughout chosen years (say from 2002 to 2022).