Maxine Lanier
DATA 3320
4/24/23

Weather Results

Problem

An investigation on whether it rains more in Seattle, WA, or in St. Louis, MO. To determine this, we compare using daily precipitation data from St. Louis and Seattle acquired from the National Centers for Environmental Information climate data. This data has entries from 2017 to 2022.

Analysis

Before comparing the precipitation data between the two cities, the data was cleaned and prepared.

Some of the first steps of this cleaning process were to identify what data in each dataset would be important for our question, make sure this data was represented in the same way for each city, and make sure that each city was evenly represented when analyzing. We converted the data types of "date" to make this easier to use later on. When examining the initial data sets, it was found that there wasn't precipitation data for all the dates and that St. Louis had many more observations due to it being reported from multiple stations, as opposed to Seattle data coming from just one station. To fix these problems, we selected the relevant subsets of the data, limiting the data to the years 2018-2022 coming from just one station for St. Louis. To account for the missing values, we decided to resolve this after joining the two data sets so that the missing precipitation data could be replaced all in one.

We joined the two data sets and changed the form of the frame so that each day has two rows with an observation, one for each city. This form made it more simple to identify where a date is missing. After adding a column that includes the day of the year for each observation (out of 365), the missing values were filled in with the average precipitation across all years for that missing day of the year.

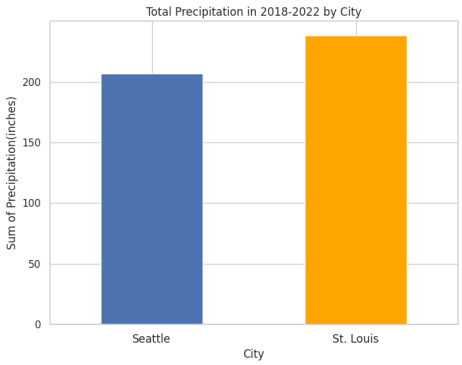
Now that the data sets are combined and the missing values are filled in, we make some finishing touches to the data frame to make it most useful for analysis later. This included renaming the columns, putting the city names into a more readable form, and creating the column "month" which identifies the number of the month (1-12) each observation is from. The file was then clean and ready for analysis to answer the question.

The resulting weather data was analyzed using bar charts that visually represent the amounts of rain each city is getting. The two main variables that were used to determine where it rained more were the inches of rain and the count of days including or not including rain. The distinction is that the first is a measure of the amount of rain, and the other is a measure of how much time a city spends in the rain.

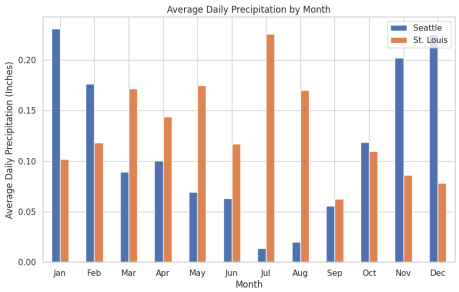
Three plots were created. The first shows a bar for each city that represents the total amount of precipitation in inches that said city got over the full 5 years of data. The second chart

shows a bar for each month of the year, for each city, showing the difference between the average daily precipitations during that month. The third chart also has a bar for each month of the year, visualizing for each city the average number of days in that month when it does not rain.

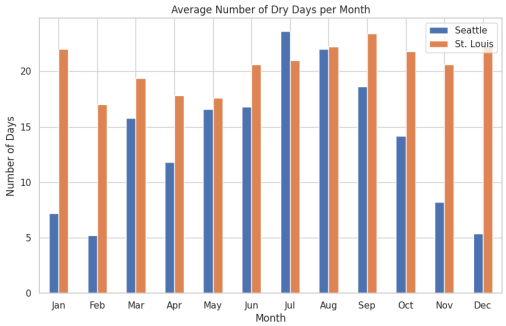
Conclusion



The chart above shows that over the entire 5 years of data, St. Louis received more rain in inches than Seattle.



The above chart shows us that St. Louis gets the majority of its rain in the summer months, while Seattle is the least rainy during that time. For 5 months, Seattle is rainier on a daily basis, and St. Louis is rainier for 7 months. Overall, St. Louis receives more rain in inches on an average daily basis.



This chart shows us that for 11 out of 12 months, Seattle spends more days per month in the rain than St. Louis.

While it rains more in volume in St. Louis, Seattle has many more rainy days. For someone who has to consider extreme conditions, for example, if they were looking to buy a

house and didn’t want to weatherproof it against heavy rain during select seasons, St. Louis would be rainier. For someone who is asking this question because they like to be outside often or would be affected by constant slightly rainy days, they would consider Seattle to be rainier.