

Tina Chau
DATA 3320

Introduction

The data science question we are concerned with is whether it rains more in Seattle or St. Louis. To accomplish this, we will be using a refined data set which can be referenced from my github repository [here](#).

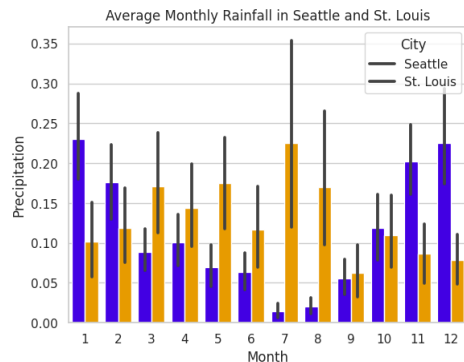
This environmental data is sourced from the NOAA National Centers for Environmental Information, which includes records of daily precipitation. We can use their [website](#) to request records of daily precipitation from Seattle and St. Louis (or other locations of interest) for the last 5 years (2018 - 2022).

Some questions about the dataset in the analysis are:

- How does the average precipitation per month for each city compare to one another?
- Are there any days with 0 precipitation data for either city?
- What is the maximum amount of precipitation recorded in each city, and are there any significant outliers in the data?

Analysis

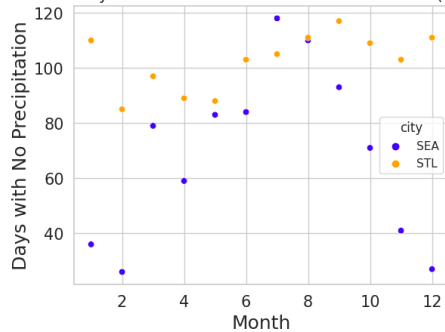
→ How does the average precipitation per month for each city compare to one another?



To create this barplot, I used Seaborn and stored the code in `monthly_avg`. I created a dataframe containing the precipitation data for both cities. Then I used the `palette` argument to set the colors according to the city. Overall, it looks like St. Louis does have a higher average than the Seattle dataset does. However, it's worth noting again that the two cities exhibit distinct seasonal patterns in their precipitation levels. Seattle tends to experience higher rainfall averages during the winter months, particularly from November to February, whereas St. Louis sees higher rainfall averages during the spring and summer months, particularly from March to August. This implies that the two cities experience different weather patterns, and maybe there are other factors to examine these patterns more closely and investigate any potential factors driving these trends, such as geographical location, local weathering systems, or climate change.

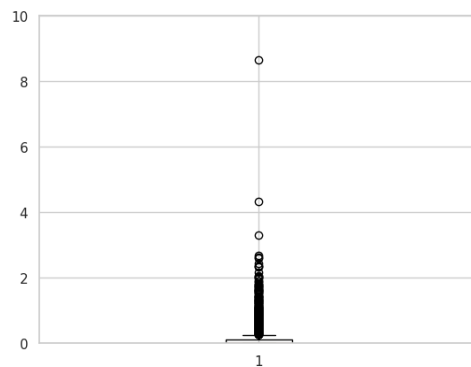
→ Are there any days with 0 precipitation data for either city?

Number of Days with No Rainfall in Seattle and St. Louis (2018-2022)



To create this scatterplot, I created a dataframe that groups the data by city and month and the count function was set to 0 to count the number of times that precipitation value was equal to 0. The plot shows that St. Louis has a higher recorded number of no precipitation days throughout the years 2018 and 2022 compared to Seattle, suggesting that rainfall in St. Louis may be less consistent and more sporadic. St. Louis may have periods of heavy rainfall followed by longer dry spells, while Seattle may have more consistent rainfall throughout the year. The higher average in St. Louis could also be influenced by a few heavy rainfall events, whereas Seattle's average is based on more consistent rainfall throughout the year.

→ What is the maximum amount of precipitation recorded in each city, and are there any significant outliers in the data?



To create this boxplot, I create a dataframe to only include data from the years 2018-2022. I then grouped it by city and set the precipitation to compute the maximum value for each of the cities. The box plot shows that the maximum precipitation for both cities falls within the range of 0 to 2, with an outlier in St. Louis at 8.64. Seattle's maximum value is much lower at 2.60. This suggests that St. Louis is more prone to extreme rainfall events, while both cities generally experience moderate levels of precipitation. However, the outlier may skew the mean and standard deviation of the St. Louis dataset, so it should be considered when analyzing the data.

Conclusion

We originally wanted to find out which cities rain more, Seattle or St. Louis. Here's some key findings and takeaways from the analyses:

- St. Louis generally experiences higher daily precipitation averages than Seattle throughout the year, with significant increases during the spring and summer months and for Seattle during fall and winter months.

- The higher average precipitation in St. Louis may be influenced by more consistent rainfall events, while Seattle's average is based on more variable rainfall with lower amounts throughout the year.
- St. Louis does have an outlier that does influence the mean precipitation value and the standard deviation. This outlier may skew these values and should be taken into consideration.

Both cities exhibit distinct seasonal patterns in precipitation. However, we also noted that there's various factors that influence these findings, such as the geographical location, the local weather systems, as well as climate change. Even though both cities do have occurring rainfall throughout the years, St. Louis ultimately is the city that rains the most out of the two.