MSDS DATA 5100

Communicate the Results | Weather

Submitted by Gurpreet Kaur

Problem Statement

We will use the data science methodology to investigate whether it rains more in Seattle, WA than in Vancouver, BC (where I used to live). The problem is simple; As I have lived in Vancouver for more than seven years, I think that it rains too much in Vancouver. However, my in-laws have lived in Seattle for more than a decade, and they debate it rains more in Seattle. I want to use data to determine whether it rains more in Seattle, WA than in Vancouver, BC.

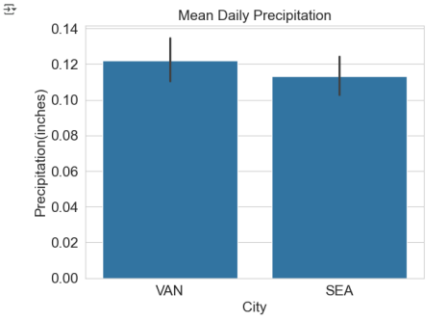
Data Used

We will use daily precipitation measured in Seattle and Vancouver from January 1, 2018, to December 31, 2022. The data sets were downloaded from the National Centers for Environmental Information NOAA tool : Search | Climate Data Online (CDO) | National Climatic Data Center (NCDC).

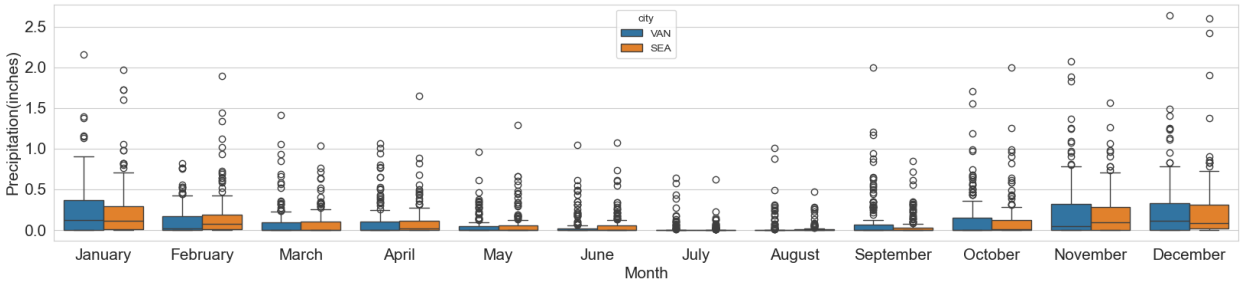
Analysis

The data downloaded was cleaned for the five years. There was some missing data for Seattle at the start of the year 2018 from January to March which was used by calculating the the mean across years of values on that day. Both datasets have one weather station and daily DATE values. The format for the date values was changed to standardize, and the precipitation was visualized to see how precipitation varies over time. Both the datasets were merged to compare the precipitation. I have used various graphs to compare the data sets but the below have been found to be more effective.

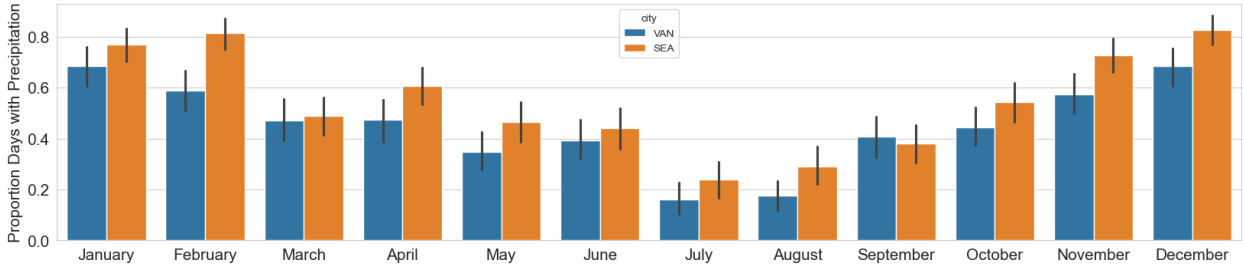
1. **Bar Plot of Mean Precipitation for each City**



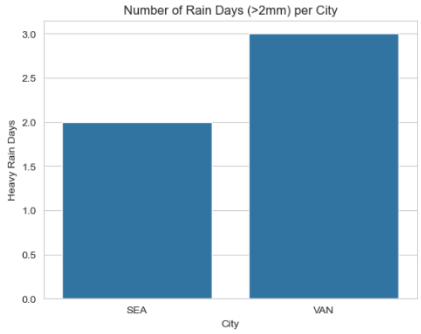
2. **Box Plot of Monthly Precipitation by City**



3. **Proportion of days with precipitation by each month**



4. **Number of Rainy days with more than 2 mm of precipitation**



Results

The bar plot comparing mean daily precipitation shows that Vancouver experiences slightly higher average precipitation than Seattle, although the difference is small. This suggests both cities receive a comparable amount of daily rainfall on average.

The box plot of monthly precipitation reveals greater variability in rainfall for both cities during the winter months, especially from October to January. Vancouver tends to have more frequent extreme precipitation events, as seen by the higher number of outliers, particularly in the colder months. In contrast, summer months show less variation and lower overall precipitation.

The graph showing the proportion of days with precipitation each month, highlights a clear seasonal pattern: both cities have more rainy days in the fall and winter, peaking in December, while July and August are the driest months. Seattle generally shows a higher proportion of rainy days in winter months compared to Vancouver.

The final graph illustrates Vancouver consistently shows a higher count compared to Seattle across most months. This indicates that while both cities experience frequent rain, Vancouver tends to have more days with substantial rainfall. The difference is especially noticeable during the fall and winter months, where Vancouver often has significantly more heavy rain days.

Conclusion

In conclusion, although both Vancouver and Seattle have similar average precipitation levels, Vancouver tends to experience more intense rainfall and a greater number of days with significant precipitation, especially in the wetter seasons. Seattle, on the other hand, has more frequent but lighter rainy days. Overall, the data highlights the strong seasonal pattern of precipitation in the Pacific Northwest, with much wetter winters and notably drier summers.