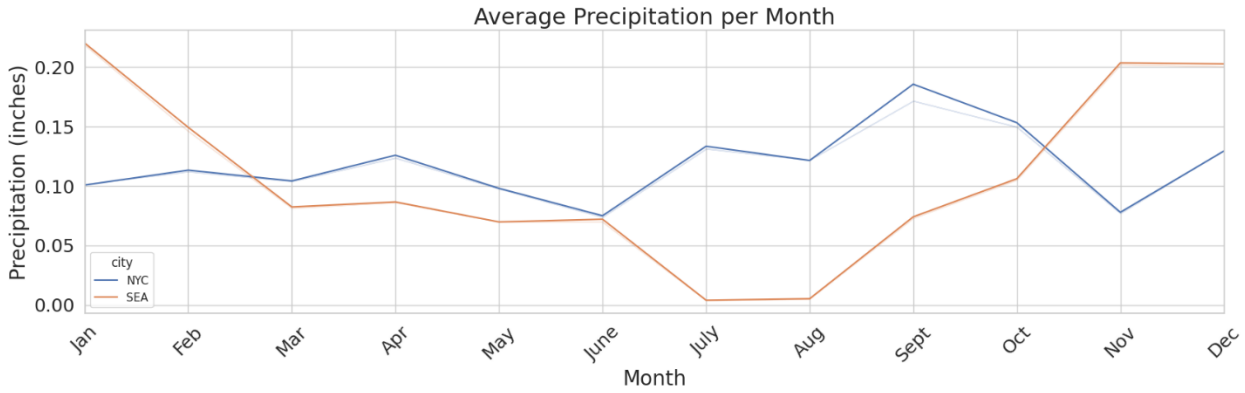
Seattle Weather Project Report
David Stanko
DATA 3320

The Problem

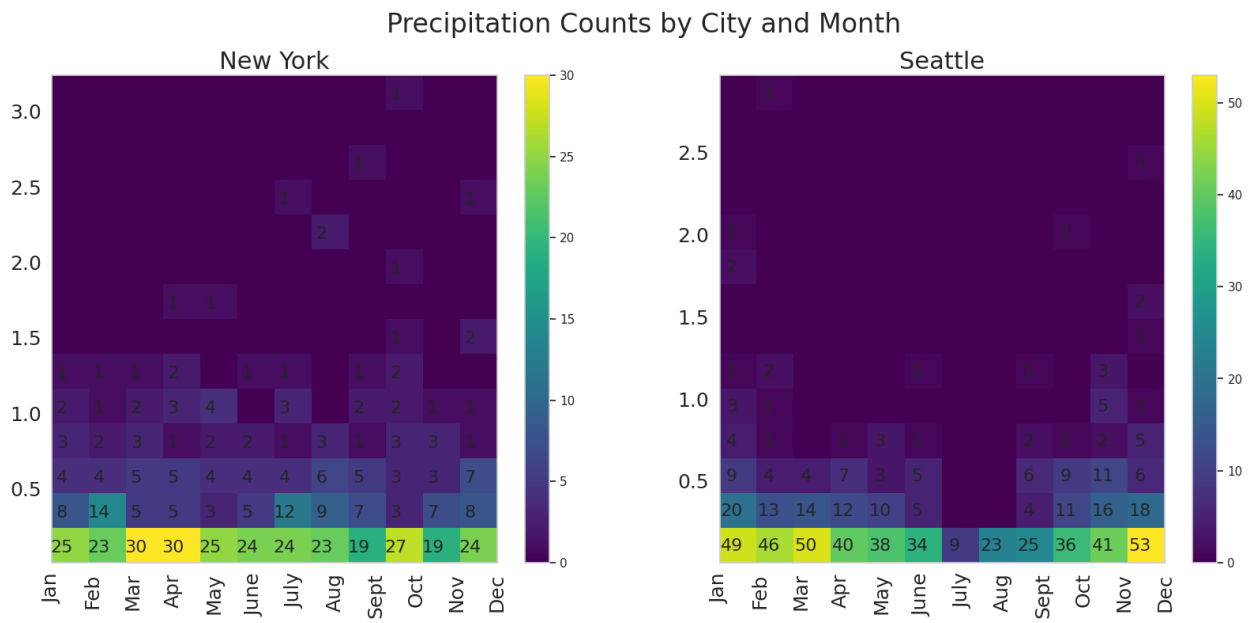
The problem is to determine whether it rains more in New York or Seattle. We did this by analyzing weather datasets from both cities. Both datasets are daily precipitation datasets that were obtained from NOAA's Climate Data Online Search tool. The data was collected over a four-year period, from January 1, 2020 to January 1, 2024. After I cleaned the data, the dataset included the following columns: `date`, `city`, `precipitation`, `month`, and `year`. `date` was the date that each data point was collected. `precipitation` was measured in inches.

My Analysis

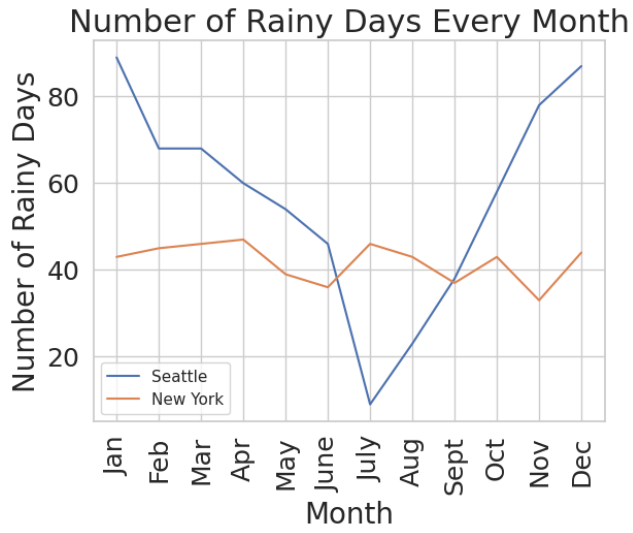
First, I looked at the mean precipitation across all the years. The mean precipitation was 0.118 inches in New York, and it was 0.106 inches in Seattle. So, just by looking at the means, it rains more in New York. Then, I wanted to know in which months it rains more in Seattle and in which months it rains more in New York. I averaged the precipitation over all the years by month, and I plotted that in the figure below. Seattle has a higher total amount of rainfall in January, February, November, and December. New York has a higher total amount of rainfall from March through October. So, when we only look at the total amount of rainfall for each month, it rains more in New York for most of the year.



Next, I wanted to know whether there are more light rain events or heavy rain events in each city during each month. First, I dropped the data where the amount of precipitation was zero, because I wanted to see a pattern on the days where it *did* rain. Then, for each city, I made a 2D histogram (shown below) with the month on the x-axis and the precipitation bins on the y-axis. Each grid square represents a pair of a month bin and a precipitation bin. The number of data points in each grid square is displayed, except for grid squares with no data points in them. Based on the 2D histograms, Seattle has more light rain events compared to New York. Out of the Seattle rain events, most of those events are light rain. However, when it rains in New York, some rain events are light and others are heavy. We can see this in the 2D histograms because, in Seattle, the precipitation counts in the first two precipitation bins (the bottom two rows) are high. In New York, the counts in those two bins are not as high, and New York has more data than Seattle in the higher precipitation bins.



Finally, I wanted to know how many rainy days there were per month in each city. In the plot below, I plotted the sum of the number of rainy days for each month over all the years. Based on this plot, Seattle had significantly more rainy days than New York almost every month. However, New York had more rain in July and August, and both cities had roughly the same amount of rain in September.



Conclusion

Based on the 2D histograms, when it rains in Seattle, it's more likely to be just light rain. Meanwhile, when it rains in New York, both light and heavy rain are likely. Further evidence for New York having more heavy rain events is that the mean rainfall is greater in New York, and for most months, New York has a higher mean rainfall by month. However, based on the last line plot, Seattle has a greater *number* of rainy days than New York. So, even though heavy rain is more likely in New York, it rains more often in Seattle.