**Seattle-New York City Rain Data Analysis**

**Introduction**

This project set out to answer a simple question: Does it rain more in Seattle or New York City? Since this is a broad question that could be answered in any number of ways, the first step in tackling this issue was to synthesize smaller questions that could help break it into more digestible data points for readers and for the analysis itself. Those smaller questions were namely: Where does it rain more in terms of average rainfall, what is the annual precipitation trend in both cities, what are the minimum and maximum precipitation days, and where does it rain more month-to-month in each city?

**Data Description**

The data used for this project is from the National Centers for Environmental Information’s [NOAA Climate Data](https://www.ncei.noaa.gov/cdo-web/search?datasetid=GHCND). The initial data sets included thousands of entries from a myriad of stations across New York City and Seattle, with data regarding the position of each station, snowfall, temperature, and more. Since the research question focuses solely on precipitation however, all extra data columns were dropped from the set. Also, to create a more cohesive and representative precipitation column, all weather recording stations’ data was averaged out for each day, so that there is one representative precipitation number for each day in each city. There were also more recording stations in New York City, so averaging out the precipitation was an appropriate method to ensure the data from either city could be compared fairly and accurately.

**Analysis Methods**

To answer the question of where it rains more on average, multiple methods were used. First, a new data frame was created which included the yearly precipitation trends for each city. Second, a smaller data frame was created that aggregated the total precipitation for each city in the entire data set. To answer the question of what the maximum and minimum precipitation days were, simple minimum and maximum functions were used on the data set. Since the minimum was 0 inches in either city, an additional variable was created to group each city by the size of how many rainless days they had, and those figures were compared. In terms of determining the annual trends, a new data frame was pulled from the original that contained the total precipitation in each year for each city. Finally, a line graph made using Matplotlib displayed the monthly trends for each city in the set.

**Data Analysis**

Every year on record featured more total precipitation in New York City than Seattle, and 2023 featured over 18 inches more of rain in NYC than Seattle. In the entire set, there were over 25 more inches of rain in NYC (195.16 inches in NYC vs. 169.33 inches in Seattle). In terms of monthly trends, every month of the year is rainier in NYC except for the winter months – November, December, January, and February were rainier in Seattle. In the summer months of July, August, and September in particular, there is far more rain in NYC.

A graph showing the number of months and months

Description automatically generated

In all, there were 94 more rainless days in Seattle. The rainiest day in the whole data set (9/2/2021) belonged to NYC with 4.24 inches recorded. No other day saw even 3 inches in the entire four-year span, with all but 1 of the top 10 rainiest days in the dataset being in New York City.

A graph with numbers and dots

Description automatically generated

**Conclusion**

From every metric that was analyzed, New York City is clearly rainier than Seattle. With more total precipitation, far less rainless days, and more consistent rain throughout the course of a year, New York City is simply a wetter place to live between the two cities.