**Weather & Climate Trends Analysis**

**1. Introduction**

This project analyses weather and climate trends across multiple cities using Python’s core libraries: **Matplotlib, NumPy, Seaborn, and Pandas**. The dataset includes key weather parameters such as temperature, humidity, rainfall, wind speed, and snowfall, enabling insightful visualizations and trend analysis.

**2. Objectives**

* Generate a **clean dataset** with realistic weather trends.
* Perform **data cleaning** to remove duplicates and null values.
* Visualize temperature trends, humidity correlations, and weather patterns.
* Store **results in Excel** for structured analysis.
* Export final insights as a **PDF report**.

**3. Dataset Description**

The dataset contains daily weather records for five cities over a **one-year period (365 days)**. The fields included are:

A screenshot of a weather report

Description automatically generated

**4. Implementation**

### **1. Data Generation**

* **Objective:** The code simulates weather data for five cities (New York, Los Angeles, Chicago, Houston, Miami) over a year (365 days). The dataset includes attributes like temperature, humidity, rainfall, snowfall, wind speed, and weather conditions.
* **Generated Data:**
  + **Temperature (°C):** Average, max, and min temperatures.
  + **Humidity (%):** Percentage of humidity.
  + **Rainfall (mm):** Precipitation in millimeters.
  + **Snowfall (cm):** Snowfall in centimeters (only for cold temperatures).
  + **Wind Speed (km/h):** Wind speed in kilometers per hour.
  + **Weather Condition:** A random selection of weather conditions (e.g., Sunny, Cloudy, Rainy, etc.).

### **2. Data Cleaning**

* **Operations:**
  + Remove duplicates.
  + Drop missing values (if any).
  + Reset the index to ensure clean, sequential indexing.

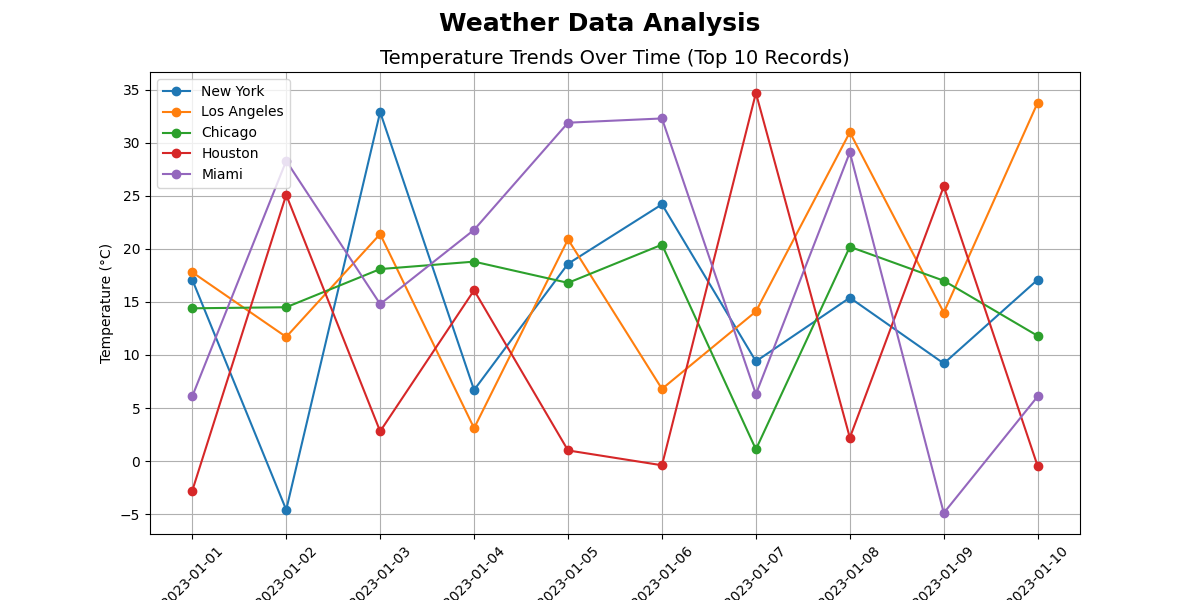
After cleaning, the dataset is saved into an Excel file (weather\_data.xlsx).

### **3. Visualizations**

The visualizations focus on various aspects of the weather data, providing both detailed and comparative insights across different cities. Below are the 13 specific visualizations generated from the data:

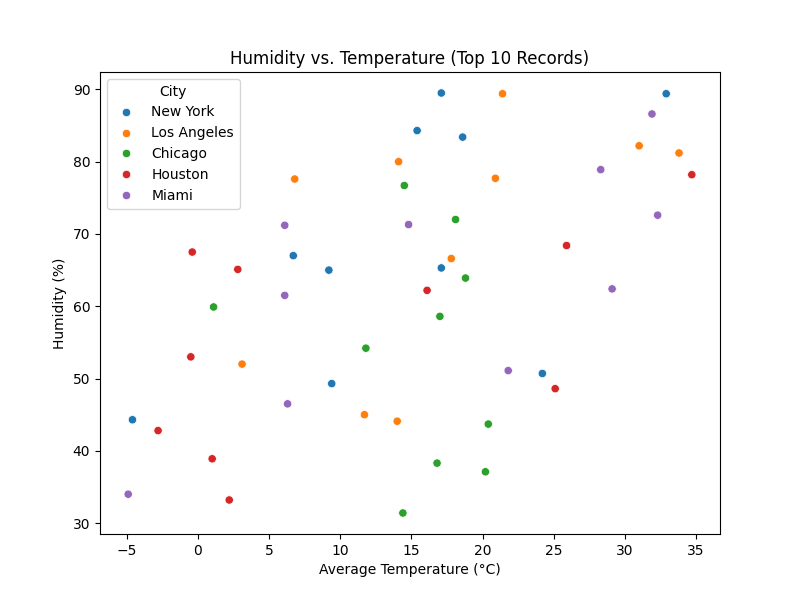
#### **Visualization 1: Temperature Trends Over Time (Top 10 Records)**

* **Description:** This line plot shows the average temperature over time for each city, focusing on the first 10 records for each city.
* **Purpose:** To track temperature variations over time for different cities.
* **Details:**
  + **X-axis:** Date.
  + **Y-axis:** Average Temperature (°C).
  + **arker:** Circles ('o') to highlight data points.
  + **Legend:** Differentiates cities.

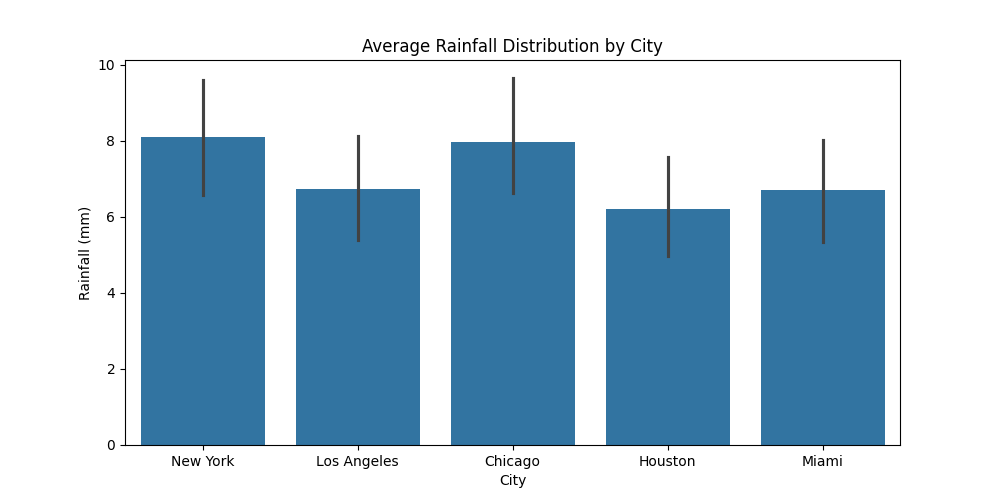


#### **Visualization 2: Humidity vs. Temperature Scatter Plot (Top 10 Records)**

* **Description:** This scatter plot visualizes the relationship between temperature and humidity for the top 10 records in each city.
* **Purpose:** To examine how humidity levels relate to temperature.
* **Details:**
  + **X-axis:** Average Temperature (°C).
  + **Y-axis:** Humidity (%).
  + **Colour:** Points are coloured by city to differentiate them.

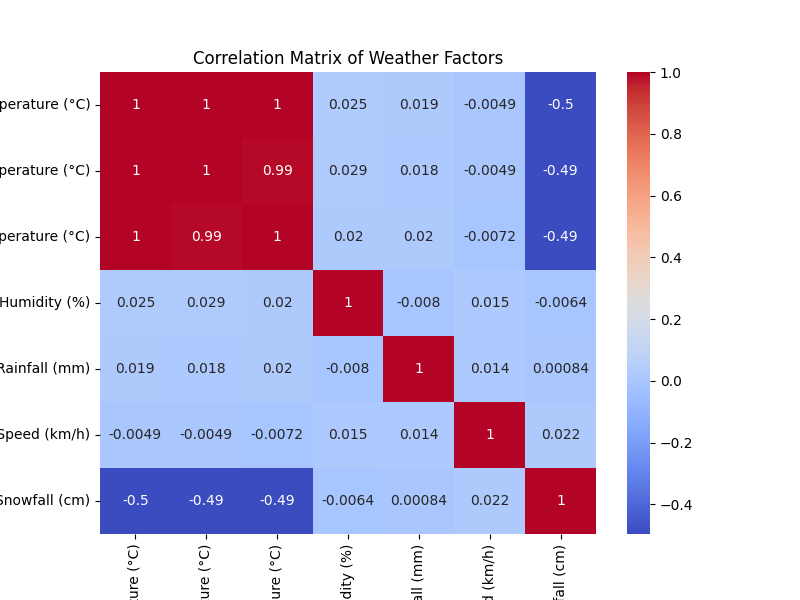


#### **Visualization 3: Rainfall Distribution by City**

* **Description:** A bar plot showing the average rainfall for each city.
* **Purpose:** To compare the average rainfall across cities.
* **Details:**
  + **X-axis:** Cities.
  + **Y-axis:** Average Rainfall (mm).
  + **Estimation Method:** The mean of rainfall for each city.

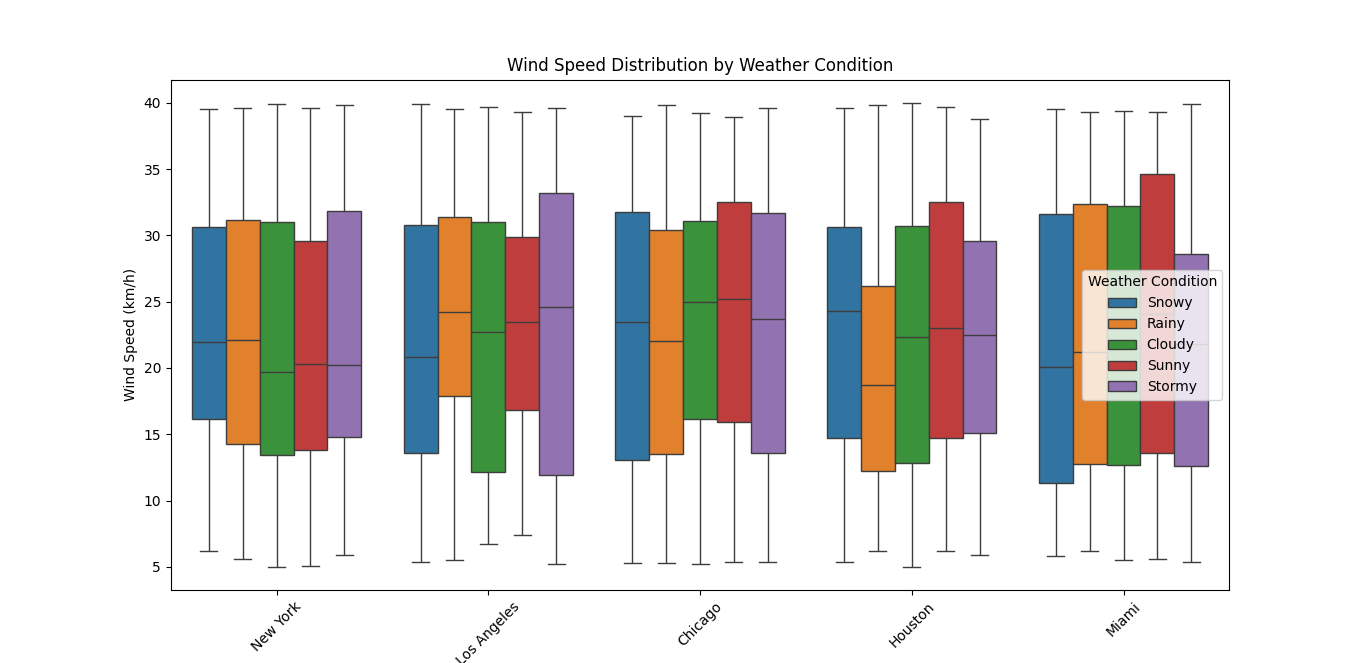
#### **Visualization 4: Correlation Heatmap**

* **Description:** A heatmap showing correlations between numerical weather attributes (temperature, humidity, rainfall, wind speed, snowfall).
* **Purpose:** To identify how different weather factors are interrelated.
* **Details:**
  + **Colour Gradient:** Cool-to-warm colours to show the strength of correlations.
  + **Annotations:** Correlation coefficients displayed on the heatmap.



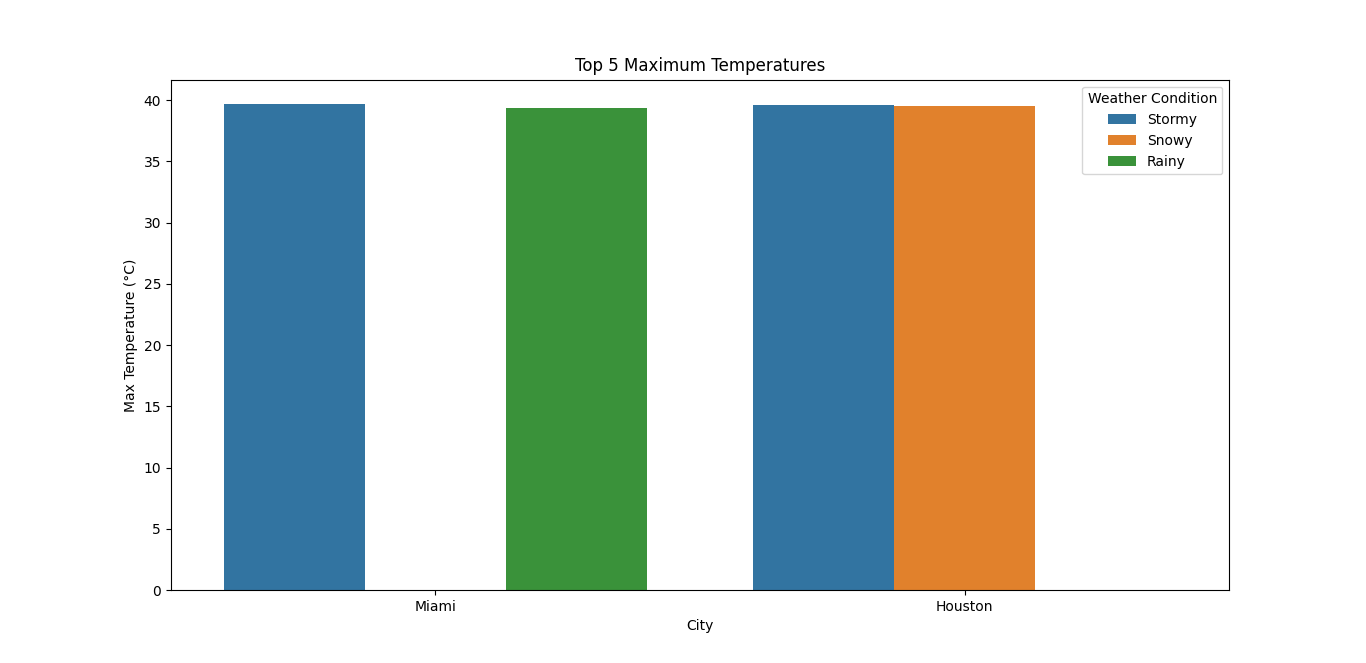
#### **Visualization 5: Wind Speed & Rainfall Impact Box Plot**

* **Description:** A box plot showing the distribution of wind speeds across cities, categorized by weather condition.
* **Purpose:** To analyse how weather conditions impact wind speeds.
* **Details:**
  + **X-axis:** City.
  + **Y-axis:** Wind Speed (km/h).
  + **Hue:** Weather Condition.



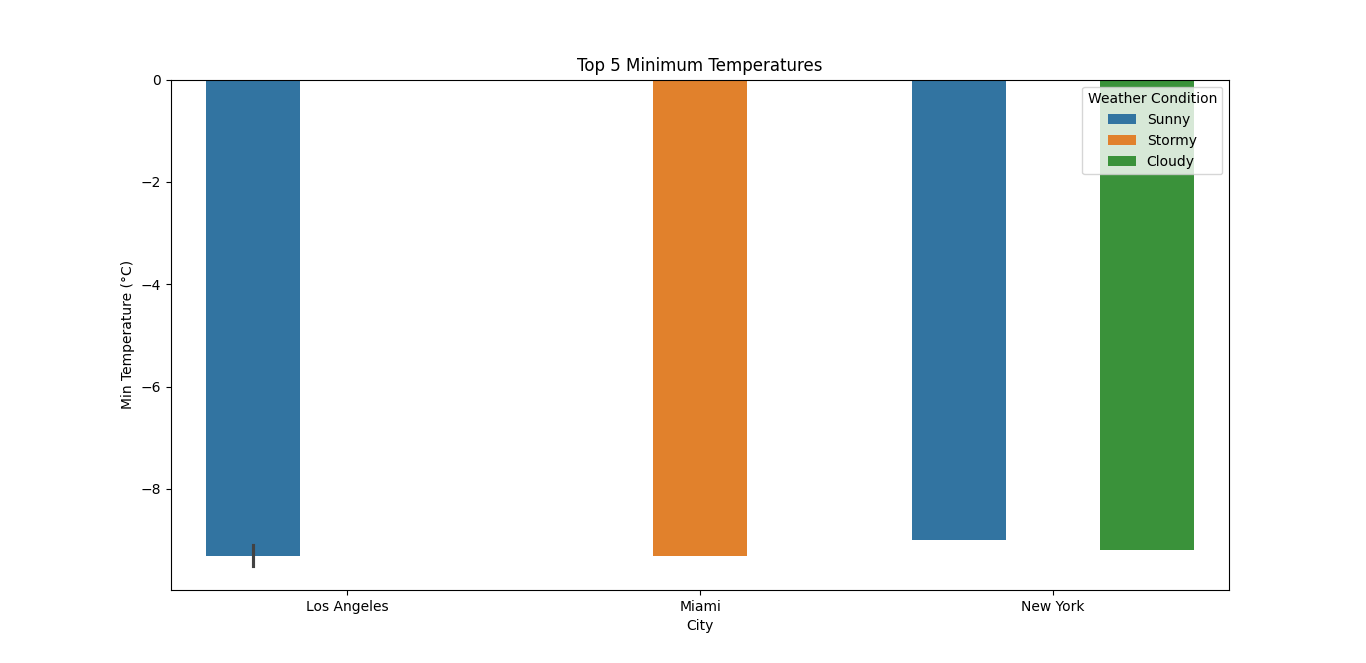
#### **Visualization 6: Top 5 Max Temperatures**

* **Description:** A bar plot displaying the top 5 highest maximum temperatures across cities.
* **Purpose:** To highlight the cities with the highest temperatures.
* **Details:**
  + **X-axis:** Cities.
  + **Y-axis:** Max Temperature (°C).
  + **Hue:** Weather Condition to show the type of weather for each reading.



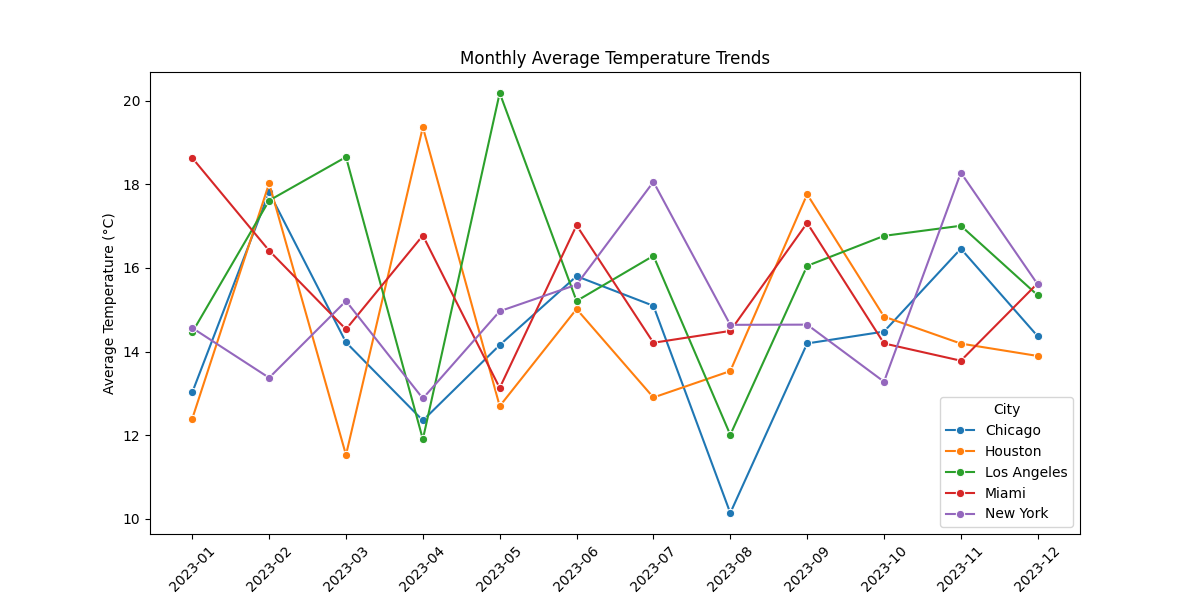
#### **Visualization 7: Top 5 Min Temperatures**

* **Description:** A bar plot displaying the top 5 lowest minimum temperatures across cities.
* **Purpose:** To highlight the cities with the coldest temperatures.
* **Details:**
  + **X-axis:** Cities.
  + **Y-axis:** Min Temperature (°C).
  + **Hue:** Weather Condition.



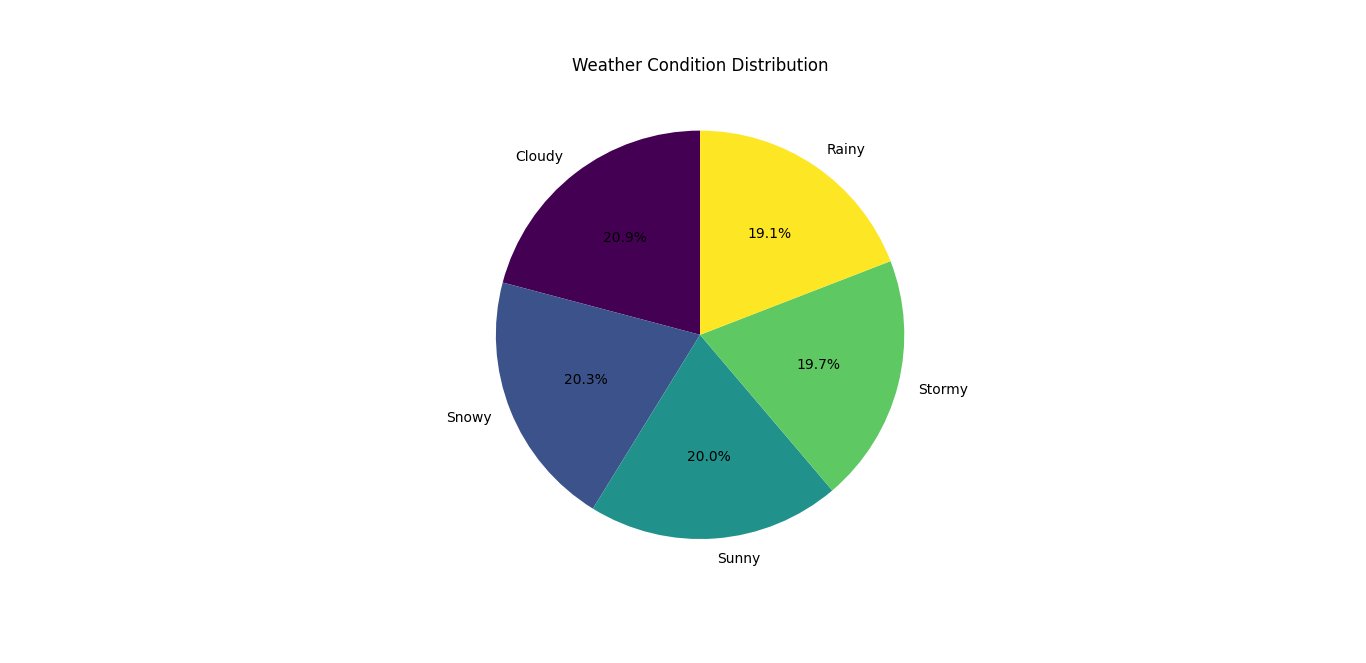
#### **Visualization 8: Monthly Average Temperature Trends**

* **Description:** A line plot showing the average temperature for each city by month.
* **Purpose:** To track temperature trends over the course of the year.
* **Details:**
  + **X-axis:** Month (formatted as YYYY-MM).
  + **Y-axis:** Average Temperature (°C).
  + **Hue:** City.



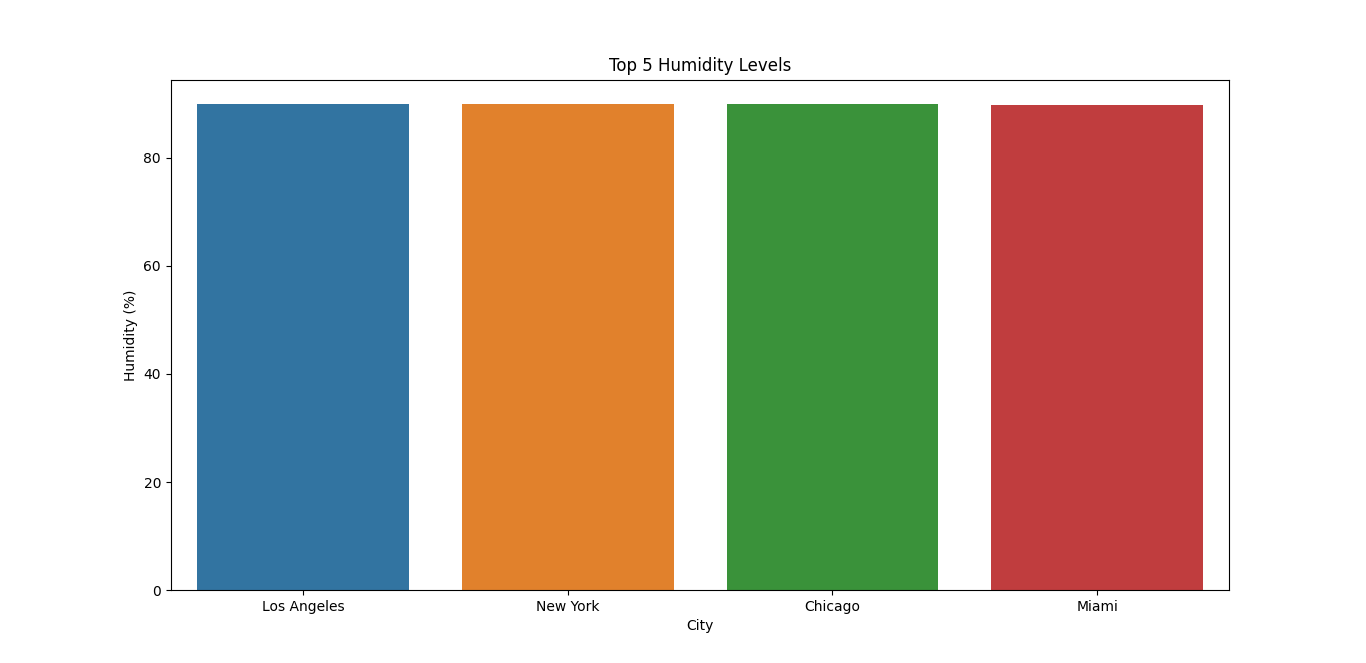
#### **Visualization 9: Weather Condition Distribution (Pie Chart)**

* **Description:** A pie chart showing the distribution of weather conditions across all data.
* **Purpose:** To visualize the proportion of each weather condition in the dataset.
* **Details:**
  + **Segments:** Represent the weather condition categories.
  + **Autopct:** Percentage displayed on each segment.



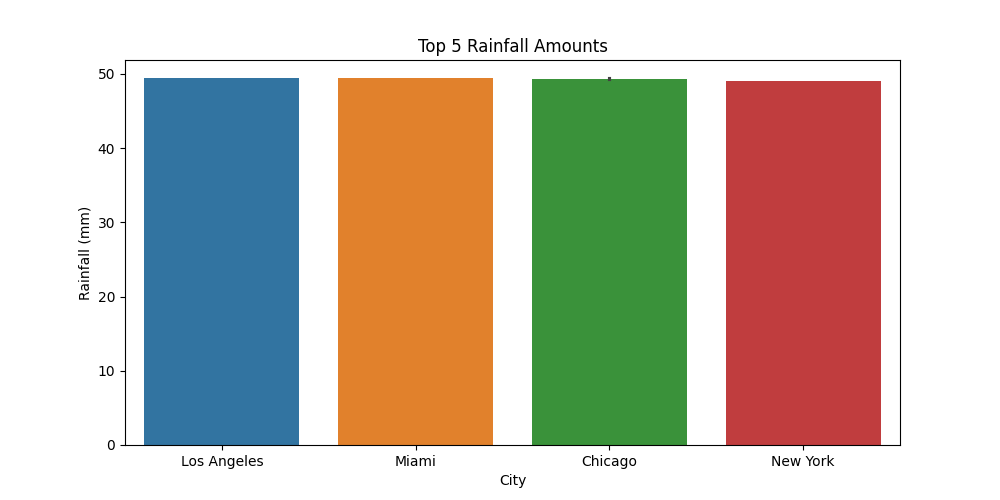
#### **Visualization 10: Top 5 Humidity Levels**

* **Description:** A bar plot showing the cities with the top 5 highest humidity levels.
* **Purpose:** To highlight the cities with the highest recorded humidity.
* **Details:**
  + **X-axis:** Cities.
  + **Y-axis:** Humidity (%).
  + **Hue:** City.



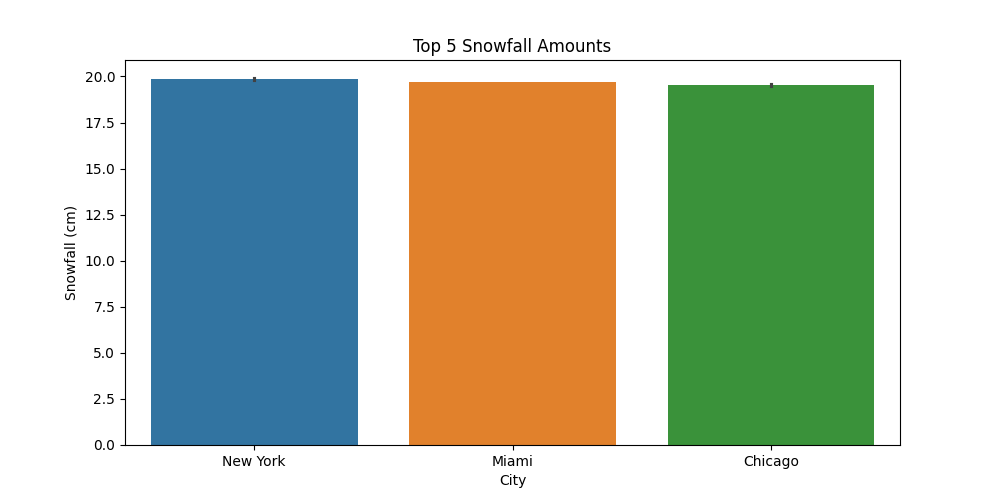
#### **Visualization 11: Top 5 Rainfall Amounts**

* **Description:** A bar plot displaying the cities with the top 5 highest rainfall amounts.
* **Purpose:** To show cities with the highest rainfall.
* **Details:**
  + **X-axis:** Cities.
  + **Y-axis:** Rainfall (mm).
  + **Hue:** City.



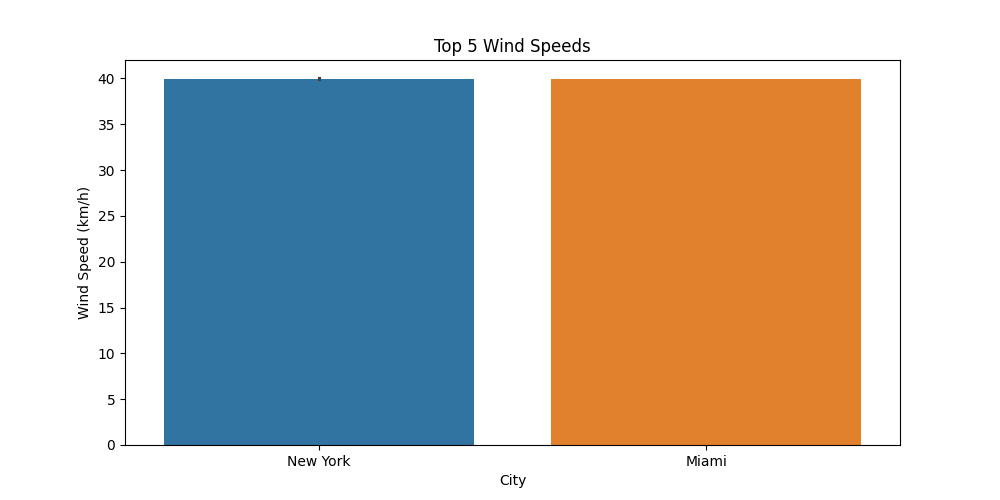
#### **Visualization 12: Top 5 Snowfall Amounts**

* **Description:** A bar plot displaying the cities with the top 5 highest snowfall amounts.
* **Purpose:** To highlight cities with significant snowfall, particularly for colder regions.
* **Details:**
  + **X-axis:** Cities.
  + **Y-axis:** Snowfall (cm).
  + **Hue:** City.



#### **Visualization 13: Top 5 Wind Speeds**

* **Description:** A bar plot showing the cities with the top 5 highest wind speeds.
* **Purpose:** To emphasize cities with the highest wind speeds.
* **Details:**
  + **X-axis:** Cities.
  + **Y-axis:** Wind Speed (km/h).
  + **Hue:** City.



### **Conclusion:**

The visualizations provide a comprehensive view of the weather data across cities over a year. They allow for:

* Identifying trends in temperature, humidity, and weather conditions.
* Comparing cities based on weather attributes like rainfall, snowfall, and wind speed.
* Gaining insights into potential correlations between weather factors.