**Background**

Data's true power is its ability to definitively answer questions. As earth’s weather cannot be controlled, we can use historical data on weather to predict weather patterns and make better data-driven decisions.

In this challenge, Python requests, APIs, and JSON traversals were used to answer a fundamental question: *"What is the weather like as we approach the equator ?"*

The following two websites were used to obtain API keys:

1. VacationPY - <https://www.geoapify.com/>
2. WeatherPY - <https://openweathermap.org/api>

A map of the world

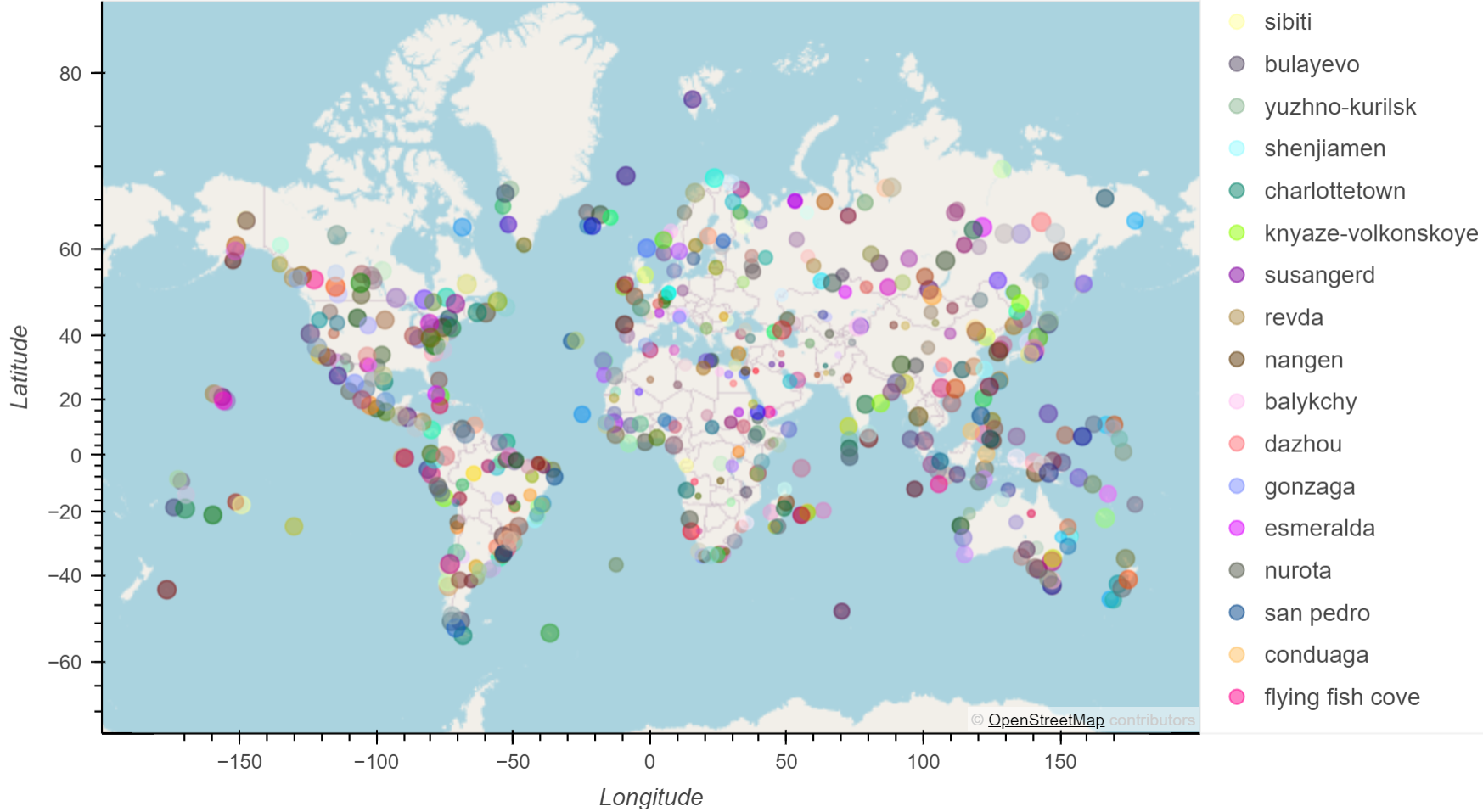
Description automatically generated

**VacationPy**

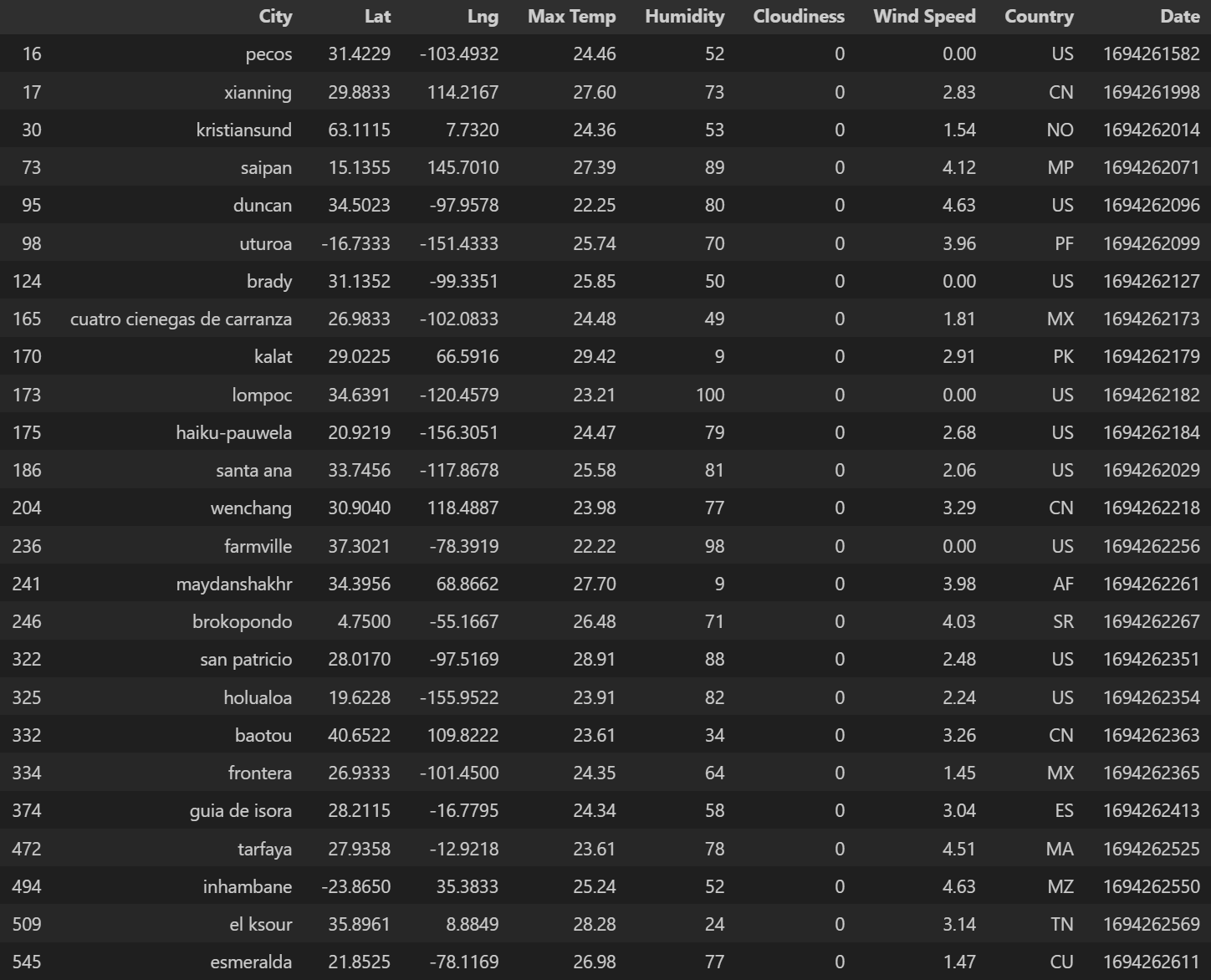
For this part of the challenge, I used Jupyter notebook, the geoViews Python library, and the Geoapify API to plan future vacations. The code provided for this challenge was used to import the required libraries and load the CSV file with the weather and coordinates data for each city.

I mainly used the Geoapify API and the geoViews Python library to create map visualisations.

***Step 1: Created a map that displayed a point for every city in the `city\_data\_df` DataFrame. The size of the point was the humidity in each city.***

******

***Step 2: Narrowed down the `city\_data\_df` DataFrame to find the ideal weather condition***



***Step 3: Created a new DataFrame called `hotel\_df`.***

***A screenshot of a black screen

Description automatically generated***

***Step 4: For each city, used the Geoapify API to find the first hotel located within 10,000 metres of my coordinates.***

A screenshot of a black screen

Description automatically generated

***Step 5: Added the hotel name and the country as additional information in the hover message for each city in the map.***

A map of the world

Description automatically generated

**WeatherPY**

A Starter Code was provided so as to be able to generate random geographic coordinates and a list of cities by using the `citipy` library.

***Created Plots to Showcase the Relationship Between Weather Variables and Latitude***

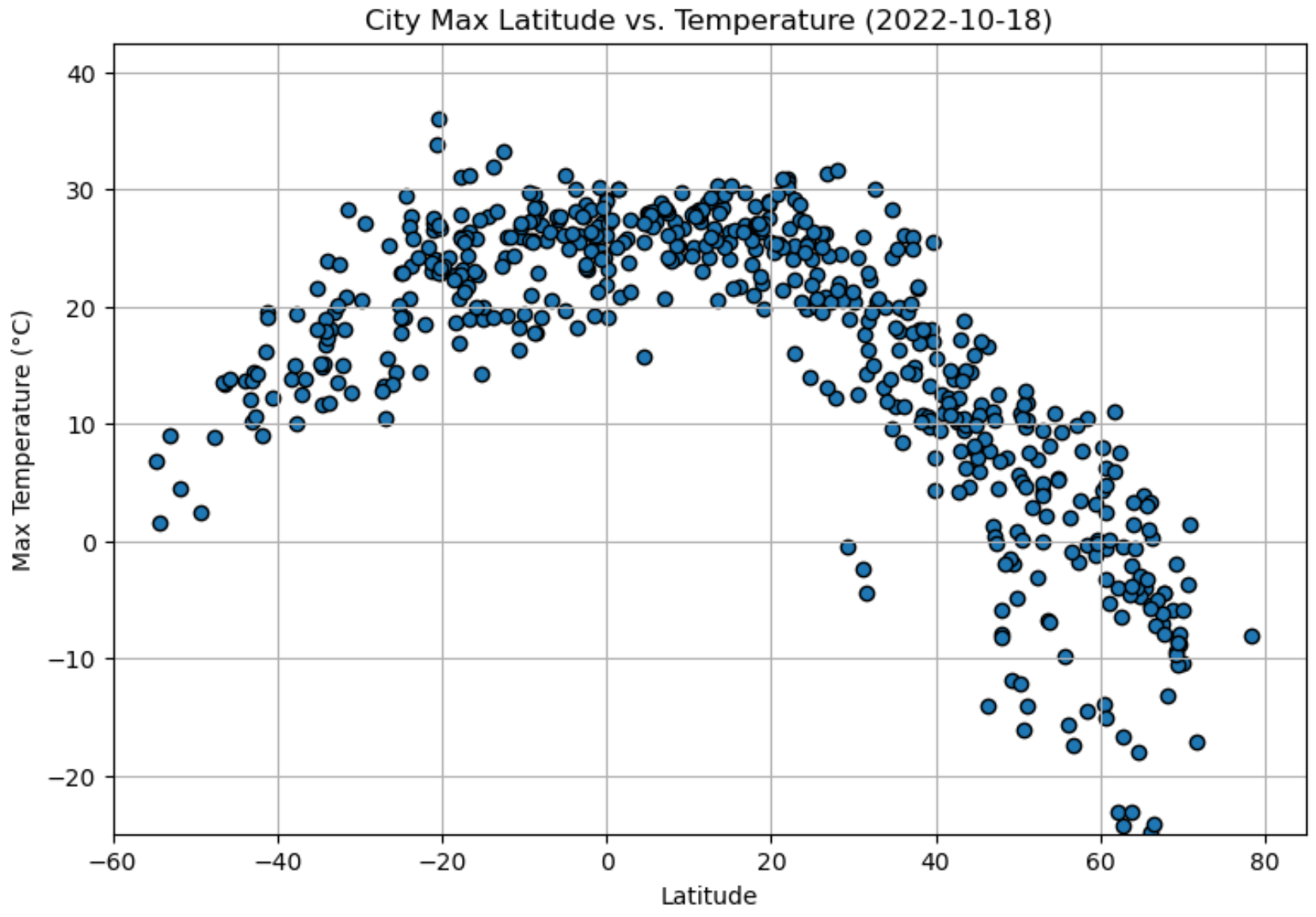
To fulfill the first part of my analysis, I used the OpenWeatherMap API to retrieve the weather data from the cities list generated in the starter code. The cities weather data was then converted into a Pandas DataFrame and a sample of the data was displayed:

A screen shot of a black screen

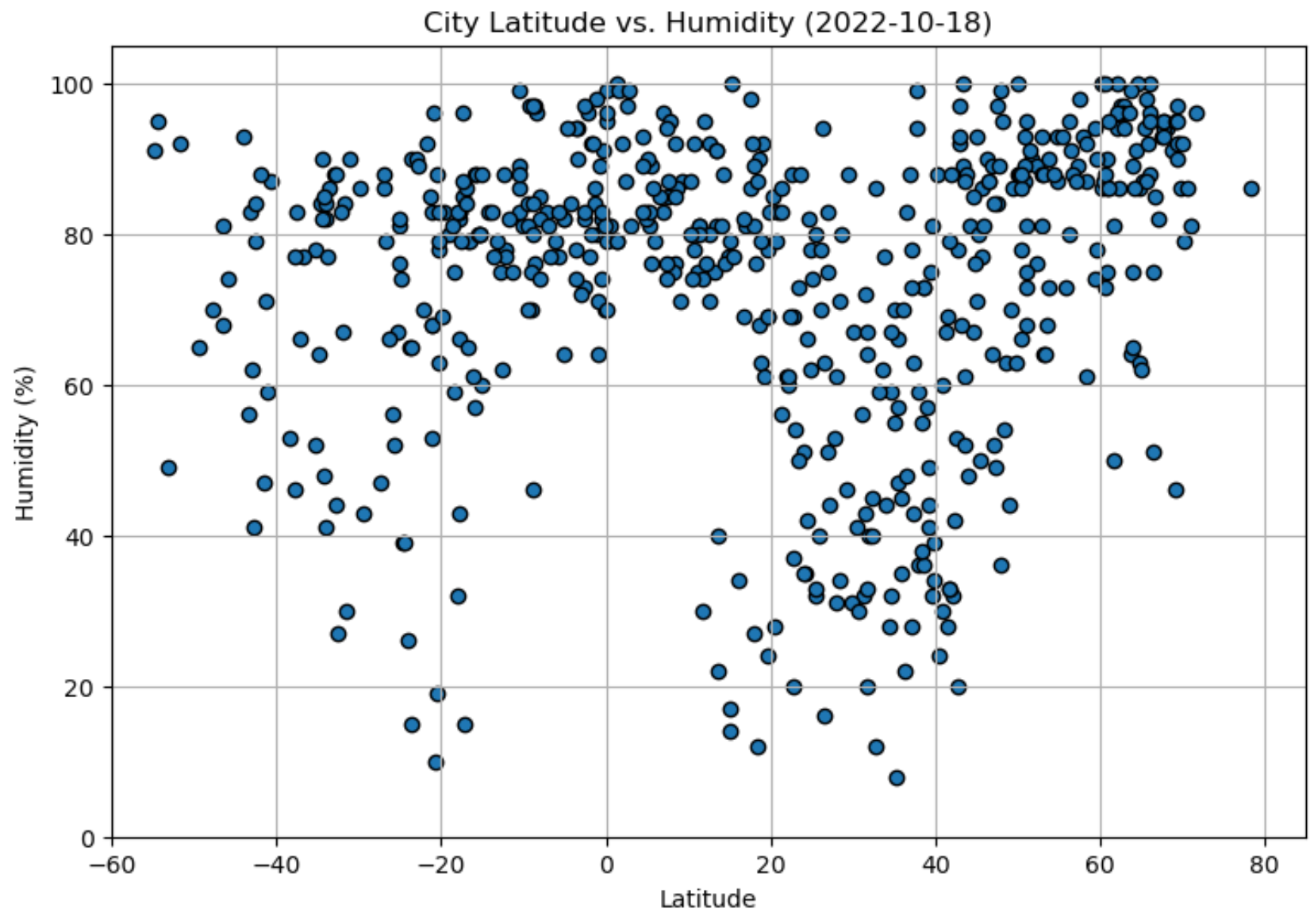
Description automatically generated

I then created a series of scatter plots to showcase the following relationships:

* Latitude vs. Temperature



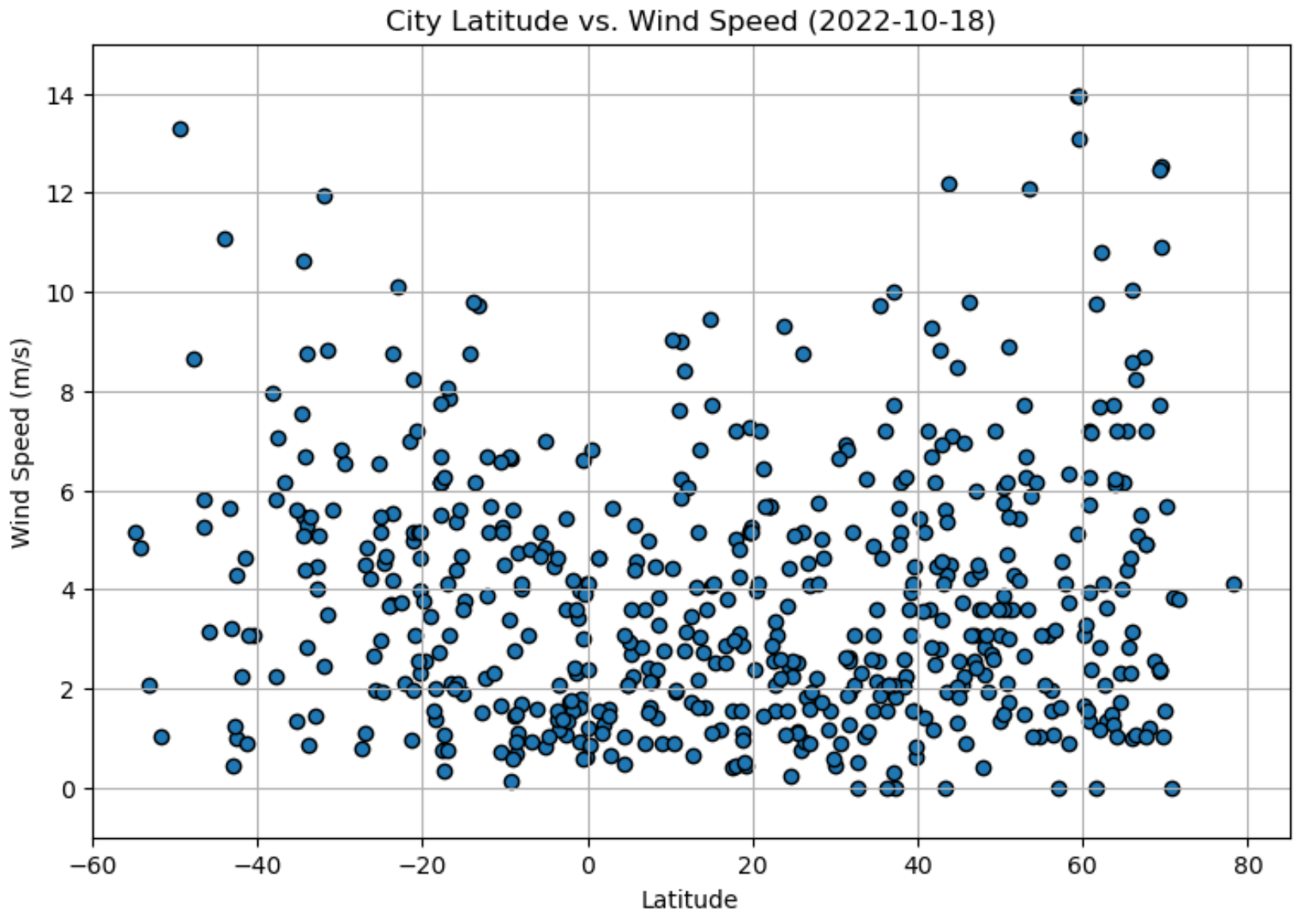
* Latitude vs. Humidity



* Latitude vs. Cloudiness



* Latitude vs. Wind Speed



***Computed Linear Regression for Each Relationship***

For the second part of my analysis, I ran a linear regression on each relationship listed above after separating the plots by Northern (greater than or equal to 0 degrees latitude) and Southern Hemisphere (less than 0 degrees latitude).

Next, I created a DataFrame with the Northern Hemisphere data (Latitude >= 0)

A screen shot of a black screen

Description automatically generated

and a DataFrame with the Southern Hemisphere data (Latitude < 0):

A screenshot of a computer

Description automatically generated

Next, I plotted a series of scatter plots with linear regression lines, the models’ formulas, and the R values on the Northern and Southern Hemispheres. I also provided analyses of what relationships the plots indicated.

**NOTE:** An R-value of 1 indicates a perfect positive linear relationship, while an R-value of -1 indicates a perfect negative linear relationship. An R-value of 0 suggests no linear relationship.

* Northern Hemisphere: Temperature (C) vs. Latitude

The R-value is: 0.7418385760864274

A diagram of a graph

Description automatically generated

* Southern Hemisphere: Temperature (C) vs. Latitude

The R-value is: 0.5181619771464621

A red line with blue dots

Description automatically generated

**OBSERVATION:**

There is an inverse correlation between the latitudes and the temperatures.

In Temperature vs. Latitude Linear Regression Plots on the Northern Hemisphere and the Southern Hemisphere, the two R-values of 0.7418385760864274 & 0.5181619771464621 respectively indicate a moderately strong positive linear relationship between the "x" and "y" variables i.e. between the Latitude and the Temperature.

The two linear regression equations describe how "y" (Temperature) changes with changes in "x" (Latitude). This means that the lines are reasonably good fits for the data, and can be used to make predictions about "y" (Temperature) based on the values of "x" (Latitude).

* Northern Hemisphere: Humidity (%) vs. Latitude Linear Regression Plot

The R-value is: 0.02825154105884831

A diagram of a graph

Description automatically generated with medium confidence

* Southern Hemisphere: Humidity (%) vs. Latitude Linear Regression Plot

The R-value is: 0.057362528661052216

A diagram of a red line and blue dots

Description automatically generated

**OBSERVATION:**

There is in Humidity vs. Latitude Linear Regression Plots on the Northern Hemisphere and the Southern Hemisphere, the two R-values of 0.02825154105884831 & 0.057362528661052216 respectively indicate very weak linear relationships between the "x" and "y" variables i.e. between the Latitude and the Humidity.

The two low R-values suggest that the data points are not well explained by the linear models. This suggests that these two linear models are not good fits for the data and may not be very useful for making predictions about Humidity based on the values of the Temperatures.

* Northern Hemisphere: Cloudiness (%) vs. Latitude Linear Regression Plot

The R-value is: 0.02596306246371301

A diagram of a graph

Description automatically generated

* Southern Hemisphere: Cloudiness (%) vs. Latitude Linear Regression Plot

The R-value is: 0.014027567648827993

A diagram of a graph

Description automatically generated

**OBSERVATION:**

From the plots on the two graphs above, we can determine that there is no correlation between cloudiness levels and the latitudes.

In the two Cloudiness vs. Latitude Linear Regression Plots, the two linear regression equations of y = 0.33x + 51.59 & y = 0.32x + 66.11 describe how "y" (Cloudiness) changes with changes in "x" (Latitude) but the r-values of 0.02596306246371301 and 0.014027567648827993 indicate very weak linear relationships between the two variables.

This suggests that the linear model is not a good fit for the data, and it may not be very useful for making predictions about "y"(Cloudiness) based on the values of "x" (Latitude).

* Northern Hemisphere: Wind Speed (m/s) vs. Latitude Linear Regression Plot

The R-value is: 0.03947205138213818

A graph of a number of blue dots

Description automatically generated

* Southern Hemisphere: Wind Speed (m/s) vs. Latitude Linear Regression Plot

The R-value is: 0.07706443649353777

A graph of a number of points

Description automatically generated with medium confidence

**OBSERVATION:**

The two scatter plots above for Wind Speed vs. Latitude show that there is no correlation between Wind Speeds and Latitudes.

R-values of 0.03947205138213818 and 0.07706443649353777 indicate very weak linear relationships between "x" (Latitude) and "y" (Wind Speed).

The coefficients of "0.03" and "-0.05" (the slopes) are very small rates of changes in "y" (Wind Speed) with respect to changes in "x" (Latitude).