Data Analytics and Visualization Bootcamp

Openweathermap.org API Data Analysis

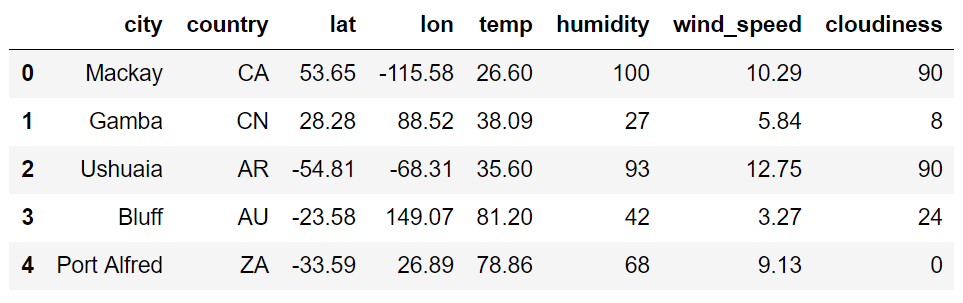
Jeff Olson

10/9/2018

Summary:

The following api data was collected and analyzed for the purposes of this body of work. Openweathermap raw data was collected using api methodologies and results evaluated by temperature, humidity, cloudiness, and windspeed charted against latitudes. The data was collected using api access to gather relevant data from 606 cities worldwide on 10/9/2018. Citipy was used to randomly generate the sample of cities based on latitude and longitude for this data set. A list compression loop was used to collect the data with a 1 second per loop throttle to prevent oversampling issues with the vendor. The data was parsed using json methods and cast into lists and then assembled into a dataframe (see figure 1). Plots were constructed from the dataframe sample to analyze results. Observations based on this series of scatter plots to showcase the following relationships. Units are imperial in this dataset (Fahrenheit, humidity %, windspeed in mph ,% cloudiness).

**Figure 1) Data Frame Header:**

****

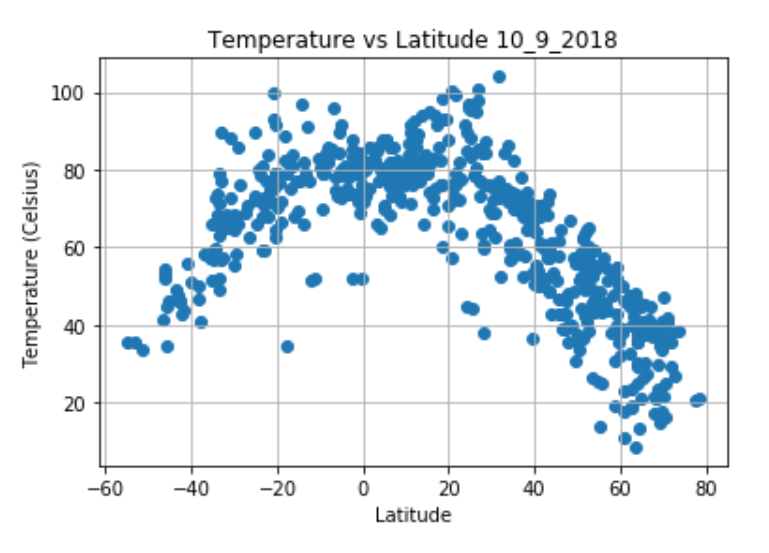
**Trends Discussion:**

1 - There is a definite trend to the data for temperature vs latitude shown in **Chart 1) Temperature vs Latitude**. This data was collected on 10\_9\_2018 just past Autumnal equinox. There is a clear curvature to the data with a maximum temperature near the equator where the sun is at its zenith for the longest period.

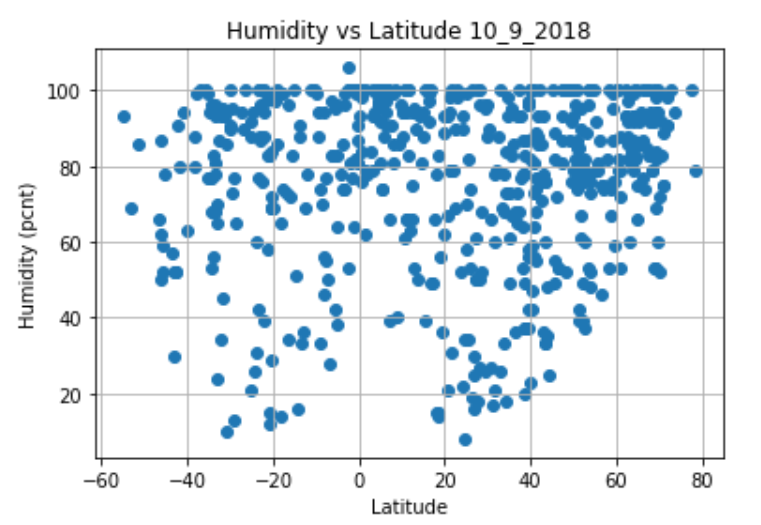
2 - The wind speed and humidity in **Chart 2) Humidity vs Latitude** and **Chart 3) Windspeed vs Latitude** seem to be different at the equator than the poles. The distribution of speeds near the equator shows lower dispersion but the distribution increases as you move away from the equator. Similarly, for humidity, the closer you are to the equator the lower the humidity. Certainly, there are more arid regions near the equator worldwide.

3 - There is no correlation of cloudiness vs latitude as can be seen in **Chart 4) Cloudiness vs Latitude**.

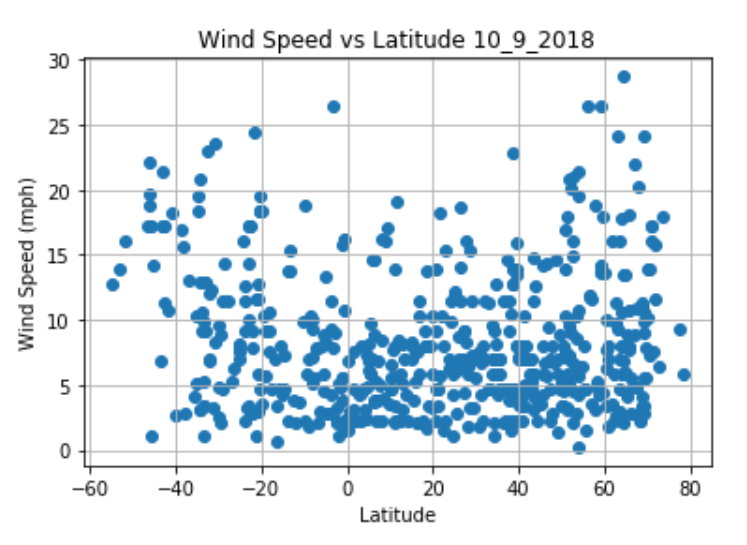
**Chart 1) Temperature vs Latitude**



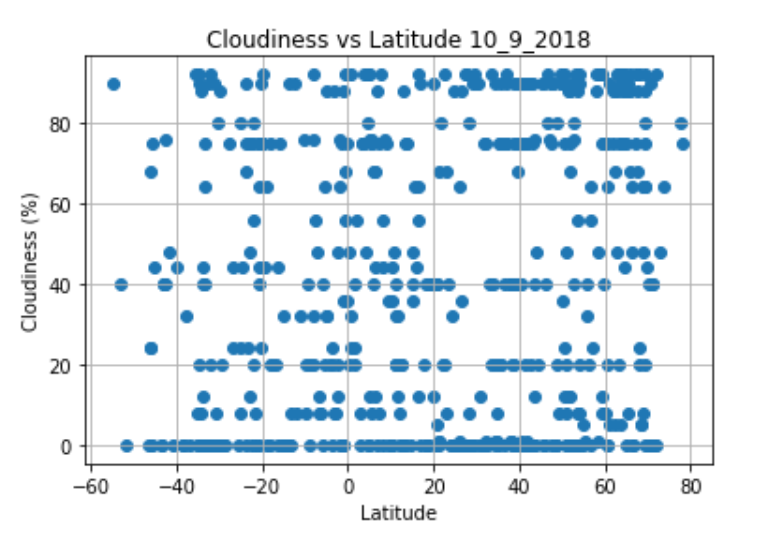
**Chart 2) Humidity vs Latitude**



**Chart 3) Windspeed vs Latitude**



**Chart 4) Cloudiness vs Latitude**

****

**What are some of the limitations of this dataset?**

First, it is a small sample and more historical data could be reviewed to add validity. Secondly, the data could be sliced along on longitude as well particularly for windspeed correlation to jet stream. And Lastly, the distributions could be compared with t tests to prove a statistical significance.