Report following Python API challenge – WeatherPy (Part 1)

Look across all previously generated figures and tables and write at least three observations or inferences that can be made from the data.

Of the potential 500+ random cities selected for this assignment; the purpose is to determine what the weather is like as we approach the equator. There were three days of observation in the study done on September 12th ,15th, and 16th. On the first day there were 688 records, followed by 696 the next day, and on the last day 687. During each day, we created a statistical summary table showing all averages and quantiles after first inspecting the data for any humidity conditions greater than 100%. If so, we would omit these from consideration. In the three days, there were no omissions. The average humidity remained near 70%, maximum temperature of 68 degrees Fahrenheit, 18mph winds, and partly cloudy. Although the average values held steady, it was significant that the greatest maximum temperature for a city rose by 10 degrees while the lowest dropped by 7 degrees. Therefore, it seems that a location might be a factor, and so we take a closer look at the correlation between latitude and all the fields observed above.

In the first part of the analysis, plots were created per day to visually show each correlation. The result showed little dependency with respect to cloudiness due to the evenly wide distributed scatter points across the ranges of latitude. Wind conditions having a more dependent factor than cloudiness showed a more moderate speed distribution across the ranges of latitudes. There were higher winds speeds observed within 60-80 degrees positive latitude. Humidity tended to balloon upward with a significant majority of cities having humidity conditions above 60% and even a larger cluster seen within 40-60 degrees positive latitude. In contrast, lesser humid conditions seen in the southern hemisphere such as within latitudes of minus 20-40 degrees. Maximum temperature is the most prevalent factor with respect to changes in latitude. As expected, the temperature increases moving away from the equator and towards the north but decreases south of the equator. This further explains the change in seasons being opposite to on another based on which side of the hemisphere a city is in. It is summer for most cities in the northern hemisphere but winter in the south.

In the second part of the analysis, the correlation was further divided between cities in the northern and southern hemispheres. The calculations to support this consisted of the correlation coefficient ‘r’ and its p-value. The coefficient tests if two datapoints have any kind of dependent relationship and measured between -1 (strong negative) to 1 (strong positive) whereas p-value proves if its statistically significant. The lower the p-value, the more significant it is and especially if it is less than 5%.

In the northern hemisphere, the results show that as latitude increases meaning that as cities are further away from the equator, there is a stronger correlation with maximum temperature compared to the other factors. However, depending on the day; it will determine how significant that is because its p-value also changes. This happens because on the first day it is much lower than on the third day observed (9.03e-119 to 5.78e-101) which means that it is more significant. The plot further supports this claim showing a sharper decline in trend as latitude increases, and a further reduction in its r-squared value (0.67 to 0.62). The trend for cloudiness and wind speed are relatively flat showing a smaller r-squared value (<0.01). Finally, the trend for humidity increased by slightly better margins.