David Reynolds

Comp574

Project Milestone 1

**How was the data collected?**

I used the OpenWeather 5-day weather forecast API. To do this I used my OpenWeather API key along with the zip codes for the two cities that I wanted data from there I was able to get the URL for each to make the API call. Once the call was made, I was able to get the json result and turn that into dataframes and then the CVSs needed.

NOTE: There are two CSV files since I want to use two different locations and do not want to merge the dataframes

**What is the data and what information does EACH column represent?**

The data is forecasted weather data from the Concord, NH and Venezia, Italy (part of Venice) over the next five days as of 11/04/2021 with data from every 3 hour increment between start and stop times.

* dt Time that forecast data was requested, formatted in unix, UTC (seconds since January 01, 1970)
* main.temp Temperature. Currently in Kelvin
* main.feels\_like This temperature parameter accounts for the human perception of weather. Currently in Kelvin
* main.temp\_min Minimum temperature forecasted for the time at the moment of calculation. Currently in Kelvin
* main.temp\_max Maximum temperature forecasted for the time at the moment of calculation. Currently in Kelvin
* main.pressure Atmospheric pressure on the sea level calculated in hectopascals (hPa)
* main.sea\_level Atmospheric pressure on the sea level, calculated in hectopascals (hPa)
* main.grnd\_level Atmospheric pressure on the ground level calculated in hectopascals (hPa)
* main.humidity Humidity as a percentage (%)
* weather - values are dictionary stored in list all as a string -- str(list(dict))
  + weather.id Weather condition id
  + weather.main Group of weather parameters (Rain, Snow, Extreme)
  + weather.description Weather condition within the group.
* clouds.all Cloudiness cover in percent (%)
* wind.speed Wind speed which is meters/sec
* wind.deg Wind direction, degrees (meteorological)
* wind.gust Wind gust which is meter/sec
* visibility Average visibility, formatted in meters
* pop Probability of precipitation which includes rain and snow
* rain.3h Rain volume that the city had in last 3 hours, formatted as mm (millimeters)
* snow.3h Snow volume that the city had in last 3 hours, formatted as mm (millimeters)
* sys.pod Time of day d = day n = night
* dt\_txt Time that forecast data was requested, formatted as string version of datetime which is UTC

**Questions to be answered with the data/ things to do with data:**

* Convert Temperature temperature, wind speed, pressure columns to Imperial system (Fahrenheit, mph, psi)
* Are there any blank, NaN, Null values that could skew results?
* Warmest/coolest day?
* Trend of Weather over the 5 days for each location:
  + Is it getting colder or is precipitation expected more often?
* Which city is supposed to have better conditions over the next 5 days overall?
* Conditions to compare for the same time:
* Chance of precipitation
  + If so, volume?
  + Rain or Snow?
* Temperature
  + too hot, too cold, or just, right?
* Wind
  + Calm or fast
  + Gust speed
* Cloudy or Clear?
* Can wind direction be converted to cardinal directions? (N,S,E,W,NE,NW,SE,SW)