

# USA Weather Data Analysis

by Joshua Nguyen

**Abstract** The United States of America possesses diverse geographical climates. Therefore, it is beneficial to track day-to-day weather - we don't want to be caught in the rain! However, forecasting potential dangerous weather like snow storms and thunderstorms can be incredibly important because it warns us of danger.

## Project Overview

1. This project explores different statistical metrics to analyze and assay *USA Weather Data*. Additionally, this project will display the results that is easily read and interpreted by the non-technical population who may not easily understand statistical vernacular.
2. This project aims to provide analytical and visual **real-time** data for current weather in the USA.
3. This project aims to provide measuring tools for meteorologists to better develop warning systems that help save lives.
4. This project utilizes **Python** for data extraction, transformation and loads it into a **Google Cloud Service** as the data warehouse. **Tableau** was utilized for the data analytics and visualization.

## Objectives

The overall goal of this project is to build a weather report that showcases real time weather metrics along with visual analytical representation about current weather data in USA state capitals.

## Questions of Interest

1. Which USA city has the highest temperature currently?
2. Which USA city are currently raining?
3. The main ingredients for *tornadoes* are high moisture, warm temperature and low pressures. Currently, which USA cities meets these criteria?

## Extraction, Transformation and Preparation

### Extracting The Data

This project uses **OpenWeatherMap** for the weather API.

### Import Libraries

```
1 # Import Libraries
2
3 import requests
4 import datetime
5 import schedule
6 import time
7 import pandas as pd
```

```
1 # Setting API Key
2 api_key = '3492844594fa3402b5c5d03276ce4f55'
3
```

```
4 # Creating List of USA Cities
5 cities = ['Montgomery', 'Juneau', 'Phoenix', 'Little Rock', 'Sacramento', 'Denver',
6 'Hartford', 'Dover', 'Tallahassee', 'Atlanta', 'Honolulu', 'Boise', 'Springfield',
7 'Indianapolis', 'Des Moines', 'Topeka', 'Frankfort', 'Baton Rouge', 'Augusta', 'Annapolis',
8 'Boston', 'Lansing', 'Saint Paul', 'Jackson', 'Jefferson City', 'Helena', 'Lincoln',
9 'Carson City', 'Concord', 'Trenton', 'Santa Fe', 'Albany', 'Raleigh', 'Bismarck',
10 'Columbus', 'Oklahoma City', 'Salem', 'Harrisburg', 'Providence', 'Columbia', 'Pierre',
11 'Nashville', 'Austin', 'Houston', 'Salt Lake City', 'Montpelier', 'Richmond', 'Olympia',
12 'Charleston', 'Madison', 'Cheyenne', 'Los Angeles', 'New York City', 'Pittsburgh',
13 'Buffalo', 'Dallas', 'San Diego', 'Seattle']
14
15 # Creating Empty List For Data
16 cities_weather = []
```

## Make API Request

```
1 # API Call
2 def get_response(city):
3     response = requests.get(
4         f"https://api.openweathermap.org/data/2.5/forecast?q={city}
5         &units=imperial&APPID={api_key}"
6     )
7     return response
```

## Data Extraction

```
1 # Data Extraction
2 def get_data(newlist, cities):
3     for city in cities:
4         city_data = get_response(city).json()
5
6         # City Details
7         city_details = city_data['city']
8         lat = city_details['coord']['lat']
9         lon = city_details['coord']['lon']
10        population = city_details['population']
11        sunrise = city_details['sunrise']
12        sunset = city_details['sunset']
13
14        # Converting timestamp
15        sunset_time = convert_datetime(sunset)[1]
16        sunrise_time = convert_datetime(sunrise)[1]
17
18        city_weather = city_data['list']
19        for dt in range(0, len(city_weather)):
20            dt_data = city_weather[dt]
21
22            # Getting Date and Time
23            date_time = dt_data['dt']
24            _date = convert_datetime(date_time)[0]
25            _time = convert_datetime(date_time)[1]
26
27            # Getting Data
28            weather_data = {
```

```
29         'date' : _date,
30         'time' : _time,
31         'sunrise_time' : sunrise_time,
32         'sunset_time' : sunset_time,
33         'city' : city,
34         'lat' : lat,
35         'lon' : lon,
36         'population' : population,
37         'weather' : dt_data['weather'][0]['main'],
38         'temp' : dt_data['main']['temp'],
39         'temp_feels_like' : dt_data['main']['feels_like'],
40         'pressure' : dt_data['main']['pressure'],
41         'humidity' : dt_data['main']['humidity'],
42         'wind' : dt_data['wind']['speed']
43     }
44     newlist.append(weather_data)
45
46     return newlist
```

## DataFrame

```
1 get_data(cities_weather, cities)
2 weather_data = pd.DataFrame(cities_weather, columns=['date', 'time',
3 'sunrise_time', 'sunset_time', 'city', 'lat', 'lon', 'population', 'weather', 'temp',
4 'temp_feels_like', 'pressure', 'humidity', 'wind'])
```

## Transforming Data

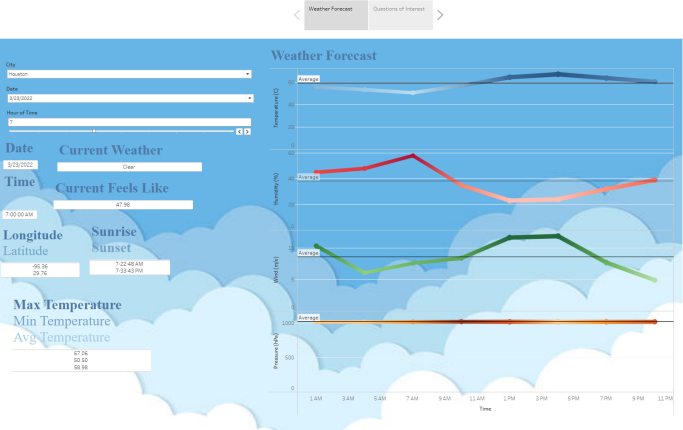
Here, we adjusted the date and columns into the proper format. That is, convert *unix timestamp* to a *datetime*. Additionally, the function will split datetime into two separate variables, date and time.

```
1 def convert_datetime(timestamp):
2     date_time = datetime.datetime.fromtimestamp(timestamp)
3     date_time = f"{date_time:%Y-%m-%d %H:%M:%S}"
4     _date = str(date_time).split()[0]
5     _time = str(date_time).split()[1]
6
7     return _date, _time
```

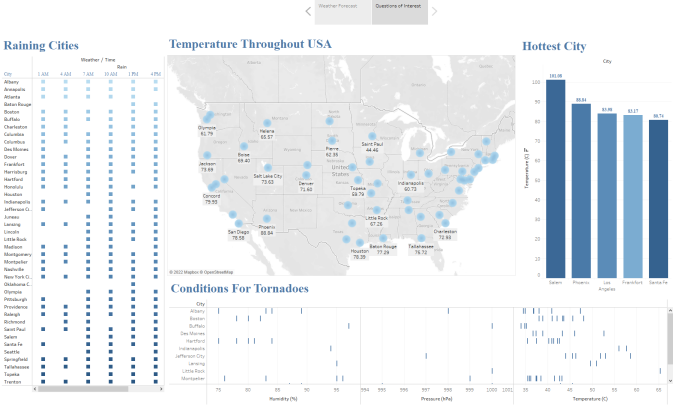
## Data Analysis and Visualizations

Importing the dataset into Tableau gives us a two-page dashboard. Click [HERE](#) to see the final dashboard.

Weather Data



Weather Data



Conclusion

The hottest city currently is **Salem, Massachusetts** with a temperature of **101.8 C**. In the *Raining Cities* section, we can see all currently raining cities.

Raining Cities		Weather / Time						
		Rain						
City		1 AM	4 AM	7 AM	10 AM	1 PM	4 PM	
Albany								
Annapolis								
Atlanta								
Baton Rouge								
Boston								
Buffalo								
Charleston								
Columbia								
Columbus								
Des Moines								
Dover								
Frankfort								
Harrisburg								
Hartford								
Honolulu								
Houston								
Indianapolis								
Jefferson C.								
Juniata								
Lansing								
Lincoln								
Little Rock								
Madison								
Montgomery								
Montpelier								
Nashville								
New York C.								
Oklahoma C.								
Olympia								
Pittsburgh								
Providence								
Raleigh								
Richmond								
Saint Paul								
Salem								
Santa Fe								
Seattle								
Springfield								
Tallahassee								
Topeka								
Trenton								

Similarly, cities that meet or almost meet the conditions for tornadoes are **Albany, Boston, Buffalo, Des Moines, Hartford, Indianapolis, Jefferson City, Lansing, Little Rock, Montpelier.**