

Zhi Wen Huang
Metis: Data Science Bootcamp
Data Engineering Project: Weather Dashboard

Abstract:

In this project, I created a dashboard to get current weathers for the capitals in the United States for all 50 states. Each state located using their latitude and longitude. Using this information, we can get the current weather status using the OpenWeatherMap API. This request will allow us to get the most current weather and it will constantly update. This application are built through Streamlit, a very useful tools when building dashboard or related websites that able to create maps and have neat styles when creating the dashboard.

Design:

Created dashboard using Streamlit with Python and building visuals for the application. Data are collected through API and stored in database using MongoDB and SQLite. MongoDB are storing JSON files which are non-relational data. SQLite stores relational data and will called using Ssqlalchemy in Python.

Data:

Data are collected using API from OpenWeatherMap and NOAA. OpenWeatherMap is to get current weather data and actively updating the status of the weather condition. NOAA allows me to get historical data for each stations in the US. The file are saved as CVS file which then I put them in the SQLite database.

Tools:

Streamlit is the main tool I use to deploy the application. Besides that, I've use Seaborn to create visuals and other Python library to extract information. Data manipulation are done mostly in pandas and numpy to get a better analysis on the data. I've also use Spark on historical data because it was too large. I was manage to found an error in the data and only taking the relevent data. This leads to a better algorithm time complexity when modifying the data.

Communication:

The dashboard is the way to communicate my collection of data and presenting them with maps, visuals, and other functionality. There is map showing the location which you can choose using the search box. This will constantly updates the status base on the location that has been chosen. There are also visuals of time series showing the average temperature and average wind speed throughout the month. Another big time series plot to show the entire year of data showing the temperature.