April 4, 2024

City of Worcester Air Quality Monitor Dashboard

Danielle Hall

Last Updated for Assignment 1 – 04/11/2024

OBJECTIVES

- Develop a web-based dashboard that displays live air quality (Particulate Matter 2.5) readings from all public air quality sensors in the town of Worcester.
- I will utilize all current air quality monitors and their live data readings.
- By creating a readable and easily accessible web-based dashboard, it is hoped that residents will have a better understanding of their local air quality.
- APIs will be used to query live data from PurpleAir and AirNOW
- A python script will be developed to read the data into tables
- The mapping platform chosen for this will be mapbox.

ASSIGNMENTS

Complete the following assignments. Deliverables will include pushing to your public GitHub and updating the README at every step.

Project Proposal

1. **Fill in the highlighted portions of this prompt. Utilize the Group Prompts for inspiration.**
    - Propose at least 5 vector datasets and 2 raster images that will address your topic.

To present current air quality, I will index live air quality readings at each of the six air quality sensors in Worcester. The sensors are from different providers; MassDEP, AirNOW and PurpleAir
Data will be queried via API key

Sensor Name	Provider	Address	Capability (What does it Measure)	Frequency of Observation	Date of Installation
Worcester – Summer Street	MassDEP	Summer Street	PM2.5, PM10, NOX, SO2, CO	PM2.5 continuous is continuous	01/01/2004
Worcester - Summer Street	AirNOW (same sensor as MassDEP)	Summer Street	PM2.5, PM10, NOX, SO2, CO	PM2.5 continuous is continuous	01/01/2004
Forest Grove	PurpleAir	Tattan Farm Road (Northwest Worcester)	PM2.5	10 Minute Average	
Batters Eye 2	PurpleAir	Polar Park (Within stadium)	PM2.5	10 Minute Average	
Batters Eye Polar Park	PurpleAir	Polar Park (Outside of stadium area)	PM2.5	10 Minute Average	
DEP9A2C	PurpleAir	Summer Street	PM2.5	10 Minute Average	

To visualize the data on an aesthetically pleasing dashboard, it is anticipated that neighborhood block group and city boundary vector files will be needed. If time allows, demographic information would also be desired to include within the dashboard (via joining Census data to neighborhood block group data for Worcester).

- What relationships will you analyze? Propose at least 3 spatial queries.

I will develop a web-based database for air quality in Worcester. This will include aspects of spatial database management. Querying the publicly available data involves using API keys, which will be read into a python or HTTP file (from my understanding). This file would format the data and import it into a spatial database, possibly using pgAdmin. From there, data can be cleaned and exported.
Visualization of data will be on the web-mapping platform chosen.

2. **Create a new Final Project repository and invite Jon & Kunal to collaborate on GitHub.**

Due Friday, April 5 @ 5 pm (10 Points)

Assignment 1 – Data Acquisition, Processing, & Database Setup

2. **Find and Process Geospatial Data**
   - I will query data from two air quality providers, PurpleAir and AirNOW. Live data feeds will be accessed via API for PurpleAir and RSS for AirNOW data. Data queried consists of air quality readings (Particulate Matter 2.5).
   - Images of locations of sensors from which data is being queried from is in github README.

3. **Set Up Database Schema**
   - Create schema for your chosen topic.
   - What attributes should you be mindful of?
   - Potential schema for resulting data frames may look like the following, containing fields of sensor index and current air quality (PM2.5) reading

Sensor_index	Pm2.5_atm
1435	6.8
0679	12.7
9324	8.9

4. **Pre-process the Data**
   - Process the data to align different datasets temporally and spatially.
   - Be sure to capture the details in your README.
   - Data frames may need to be normalized once they are created and their schemas compared, PurpleAir versus AirNOW.

Due Friday, April 12 @ 5 pm (10 Points)

Assignment 2 – Import Spatial Data & Normalize Tables

1. **Import your data into PostgreSQL tables/schema created in Assignment 1.**
2. **Normalize your tables (1NF up to possibly 4NF, depending on your data) and explain the logic in your README.**
   - Even if normalization is not required, explain why in your README.

Due Friday, April 19 @ 5 pm (20 Points)

Assignment 3 - Spatial Queries & Presentation

Perform spatial analyses to determine:

- _____

- _____
- _____

Spatial Analysis & Presentation are Due Thursday, April 25 @ 10:15 am (40 Points)

Final GitHub Repo & README are Due Friday, May 3 @ 5 pm (30 Points)

Total: 110 Points

NO LATE SUBMISSIONS ACCEPTED AFTER MAY 3 -- Plan accordingly.