

Assignment 1

Issued: March 1, 2025

Due: March 15, 2025

In this assignment, you will apply the learned techniques to visualize data. Specifically, you will use the dataset provided below to plot **3** different types of charts.

About The Dataset

S&M-HSTPM2d5 is an open-source dataset for anyone interested in researching air pollution in Chinese cities Foshan, Cangzhou, and Tianjin. This open-source dataset offers high spatial-temporal resolution PM2.5 measurements, complete with timestamps and GPS locations of both mobile and static devices. Researchers can use this data to create fine-grained city-scale air pollution maps, providing critical environmental insights for city managers and residents alike. By leveraging this information, city managers can make informed decisions when it comes to planning, managing, and developing their cities. Meanwhile, residents can use this data to steer clear of polluted areas and reduce their exposure to harmful pollutants. We encourage you to take advantage of this dataset and explore the many interesting phenomena and conclusions that it can reveal. Here is the URL of the provided dataset: https://zenodo.org/records/4028130#.Y_tUUGRBwuU

In this assignment, we use the data collected in Cangzhou, the smallest(1.7mb) portion of this dataset. Please go to the above link and download the file "Cangzhou.zip".

Name	Size	Download all
Cangzhou.zip md5:800f5de7ce64c78ee198716ca079c1c	1.7 MB	Preview Download

Requirements

- 1.1. You need to select a programming tool that helps you create the visualization application.
- 1.2. You need to create 3 different types of graphs based on the given dataset:
 - (a) A Line Chart that shows how the PM2.5 level changes over time.
 - i. Use the '.csv' files in the static folder and filter the data by time stamps between **2019-01-01 00:00:00** and **2019-01-01 12:00:00**.
 - ii. The x-axis and the y-axis represent time and the value of PM2.5, respectively.
 - iii. Plot one line for each static sensor and distinguish the sensors with colors.
 - (b) A Scatter Plot that shows how vehicles carrying mobile sensors move in the city.
 - i. Use the '.csv' files in the mobile folder and filter the data by time stamps between **2019-01-02 10:00:00** and **2019-01-02 10:20:00**.

- ii. The x-axis and the y-axis represent longitude and the latitude, respectively.
 - iii. Use scatters to represent the mobile sensors' location and distinguish the sensors with colors.
 - iv. Time evolution should be represented by the change of the scatters' transparency (earlier time stamp represented by higher transparency).
 - (c) A 3D Histogram that shows the distribution of PM2.5 levels in the whole city at a specific time range.
 - i. Use the '.csv' files in both the *mobile* and *static* folders and filter the data by time stamps between **2019-01-01 09:00:00** and **2019-01-01 09:10:00**.
 - ii. Convert the GPS coordinates of the data to grid coordinates with a suitable resolution and aggregate the data in the same grid.
 - iii. The x-axis, y-axis, and z-axis represent longitude, latitude, and the value of PM2.5, respectively.
- 1.3. You are required to write a report that includes a description and interpretation of each graph. The report should contain, at minimum, the following information:
- (a) The graphs created for each visualization task.
 - (b) A simple description of each graph. When describing each graph, provide a brief overview of what the graph represents and what data it displays. You may also want to highlight any important features or trends that are visible in the graph.
 - (c) An interpretation of your findings for each graph. When interpreting your findings, consider what insights the data provides and what implications these findings may have.
- 1.4. It is important to be transparent about any external resources used in completing an assignment. If you have discussed the assignment with anyone or used any large language models, such as ChatGPT, Github Copilot, or BLOOM, please specify how you used the external resource(s) - for example, did you use them to brainstorm ideas, write code, or proofread your work?

If you have completed the homework unaided, please state that explicitly in your report by saying, "I have completed this homework unaided."

Please note that the above requirements represent the minimum criteria for the assignment. We strongly encourage you to think creatively about how to better achieve the visualization goal and to add additional content to your graphs. You will be rewarded with points for "creativity", so don't be afraid to think outside the box!

Submission

To ensure that your submission is complete and easy to manage, please make sure to follow these guidelines:

- 1.1. Your submission should consist of two parts: **the programming part** and **the report part** (Both Chinese and English are acceptable). Please compress both parts into a single .zip format file, which should be named using the following format: **'[Your Name]_[Your Student Number].zip'**.

- 1.2. The programming part of your submission should be placed in a folder named 'VisualizationApp'. This folder should contain all of your code, and the entrance of your visualization application should be named 'main.htm', 'main.html', 'main.ipynb', or 'main.py'.
- 1.3. The report part should be integrated into a single PDF file, named 'report.pdf'. Please make sure to put this file outside of the programming part folder, so it's easy to locate.
- 1.4. To ensure that your file is easy to upload and download, please limit the size of your uploaded file to no more than 50MB.

Grading

Important! Your homework will be graded based on the following criteria:

- Completeness (20% * 3): Have you completed all three visualization tasks using your chosen visualization tools? Have you provided a description and interpretation for each graph in the report?
- Clarity(30%): Have you written clear and concise descriptions for each graph? Have you made your graphs easy to read and understand? Have you used appropriate colors, sizes, shapes, and legends to distinguish different elements on your graphs? Have you added informative titles and labels to your graphs?
- Creativity(10%): Have you gone beyond the minimum requirements of the assignment to create a particularly engaging design?