Visualizing Air Quality Data: Extending The Air Quality Stripes

Feasibility Study

Duyu Yang

8 March

1 Introduction

Air pollution, particularly fine particulate matter (PM2.5), has been a significant environmental and public health concern worldwide. PM2.5 consists of tiny airborne particles that can penetrate deep into the respiratory system, contributing to various health issues such as respiratory diseases, cardiovascular conditions, and premature mortality. Tracking and understanding historical PM2.5 trends is crucial for assessing the impact of industrialization, policy interventions, and environmental changes over time.

This project aims to visualize historical trends of PM2.5 levels across different countries and regions worldwide. By presenting these trends through effective data visualization, the goal is to raise awareness of air quality issues and encourage greater attention to environmental protection.

So far, we have conducted an initial exploration of global PM2.5 datasets. Additionally, we have attempted to replicate the Air Quality Stripes visualization [1] and explored dynamic and interactive visualization techniques to enhance data representation.

2 Background and Literature Review

Particulate matter less than 2.5 micrometers in diameter, known as PM2.5, is a significant air pollutant due to its ability to penetrate deep into the respiratory tract and enter the bloodstream. Sources of PM2.5 include vehicle emissions, industrial processes, wildfires, and other combustion activities. Exposure to PM2.5 has been linked to various health issues, such as respiratory infections, heart disease, lung cancer, and premature death [2].

Previous work, notably Air Quality Stripes [1], has provided a novel visual representation of PM2.5 trends across different regions from 1850 to 2022. The previous work uses a combination of satellite observations and computer model simulations to estimate historical air pollution levels, given the scarcity of direct PM2.5 measurements before the 21st century. The satellite dataset integrates ground-based and remote sensing data with modeled air quality trends to create an intuitive and accessible visualization.

While Air Quality Stripes [1] provides valuable static visualizations for exploring air quality trends by location, there is an opportunity to enhance user interaction by adding dynamic and interactive elements. This would not only make the exploration process more engaging and intuitive but also attract a wider audience, encouraging deeper insights and greater awareness of air quality trends. This project builds upon Air Quality Stripes [1] by introducing a web-based interactive visualization system. The goal is to allow users to dynamically explore air pollution trends across different regions, compare multiple locations, and gain deeper insights into the historical evolution of PM2.5 levels.

By improving interactivity, the project aims to make air pollution data more accessible and informative for researchers, policymakers, and the general public.

Air quality visualization has evolved from static charts to interactive web-based tools, enabling users to explore PM2.5 trends across regions and compare air quality over time. Many studies highlight the advantages of interactive, animated visualizations in making complex environmental data more accessible [3].

Several web-based tools have successfully visualized air pollution data interactively. Lu et al. developed a nationwide interactive air quality map of China, using Leaflet.js and D3.js to enable time-series exploration and pollutant comparisons [4].

Comparative visualization is another key approach. AtmoVis integrates multiple interactive views—heat calendars, line charts, and spatial overlays—to compare pollution trends across locations [5]. A clustering-based visualization tool for Taiwan grouped monitoring stations with similar pollution patterns, informing region-specific policies [6].

To enhance user engagement, many systems incorporate animated time-series visualizations. An air quality visualization system was developed based on feature extraction from multivariate time-series data, utilizing deep learning techniques to help analysts explore and analyze complex air quality datasets. This system integrates multiple interactive views, allowing users to dynamically examine pollutant trends, detect anomalies, and correlate different atmospheric variables [7]. The CSIRO Smoke and Fire visualization used WebGL to render dynamic pollution plumes in a global 3D view [8]. Such temporal animations improve understanding of pollution cycles and extreme events.

Interactive air quality tools not only enhance public awareness but also support data-driven policymaking [9]. By allowing users to explore "what-if" scenarios and visually compare pollution across regions, these systems empower stakeholders with actionable insights. The integration of D3.js, WebGL, and mapping libraries has proven crucial in building engaging, scalable, and informative air quality visualizations [10].

3 Preliminary Investigations and Findings

As a starting point, we have conducted an initial exploration of global PM2.5 data, examining historical trends across different regions and visualizing its trends. As shown in Figure 1, a slider allows users to adjust the year of the PM2.5 concentration distribution map. The next task was to replicate previous static visualizations, specifically the PM2.5 air quality stripes visualization developed in prior research [1]. This process involved processing and formatting air quality datasets and ensuring that the generated visualizations accurately reflect historical trends. Figure 2 illustrates the results of this replication, showing the recreated PM2.5 air quality stripes visualization. By successfully reproducing these static visualizations, we have established a baseline for further development.

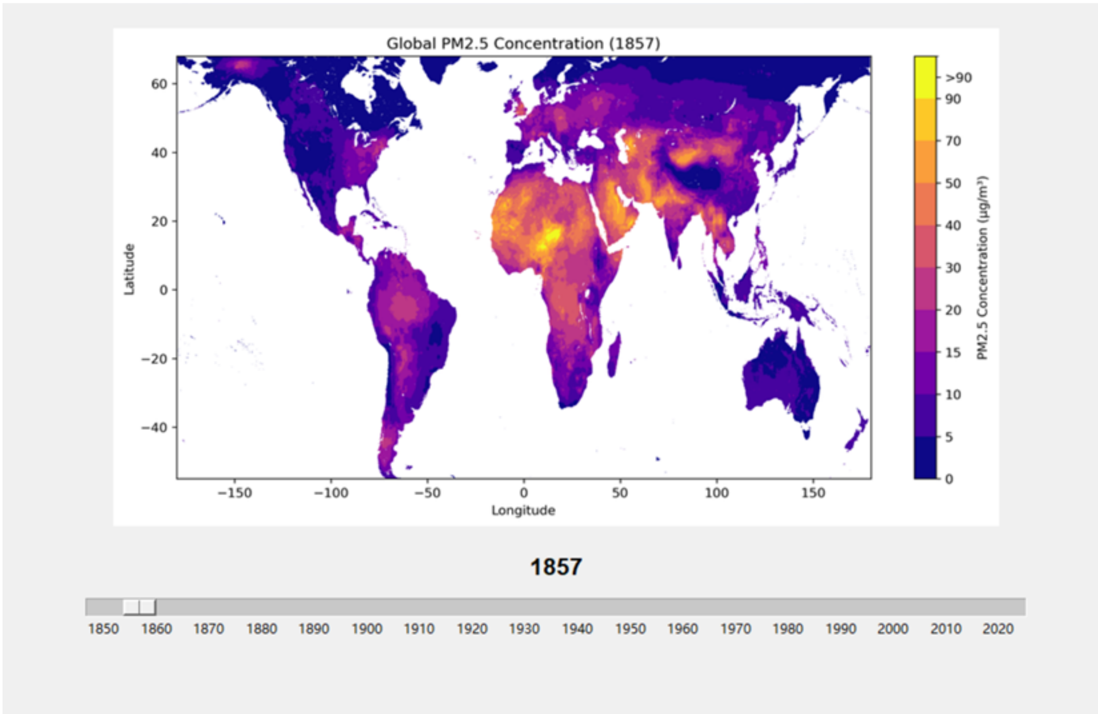

Figure 1: Interactive slider to allow the user to explore the data in different years.

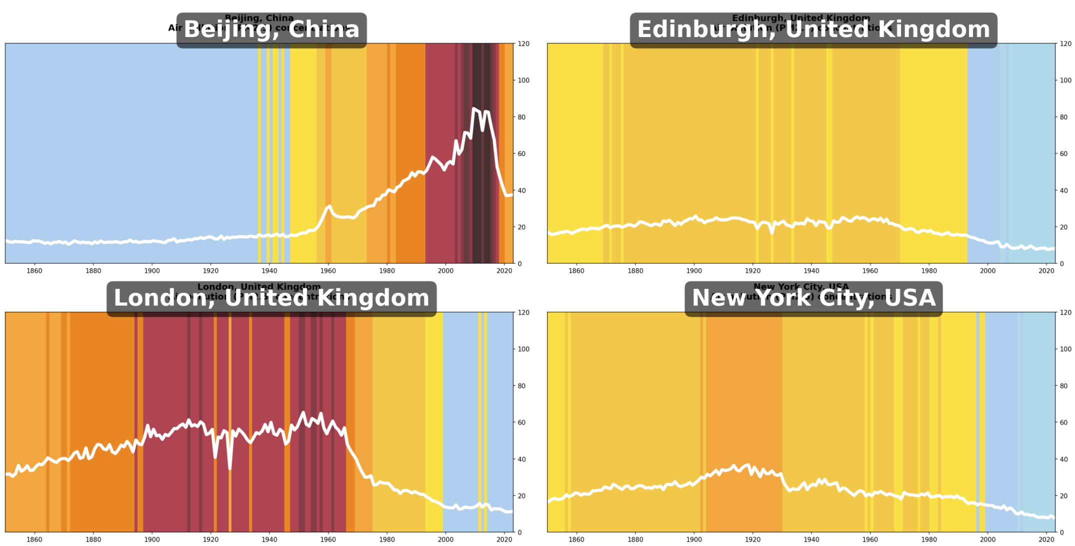

Figure 2: Reproduction of Air Quality Stripes

Building on this static representation, we have developed a dynamic prototype for 10 cities based on previous work [1], including Edinburgh, London, Beijing, and other major areas. As shown in Figure 3, this proof-of-concept animation allows for a more engaging way to present time-series air pollution data, offering users an intuitive understanding of historical changes. The dynamic prototype and code can be found on this project's GitLab [11].

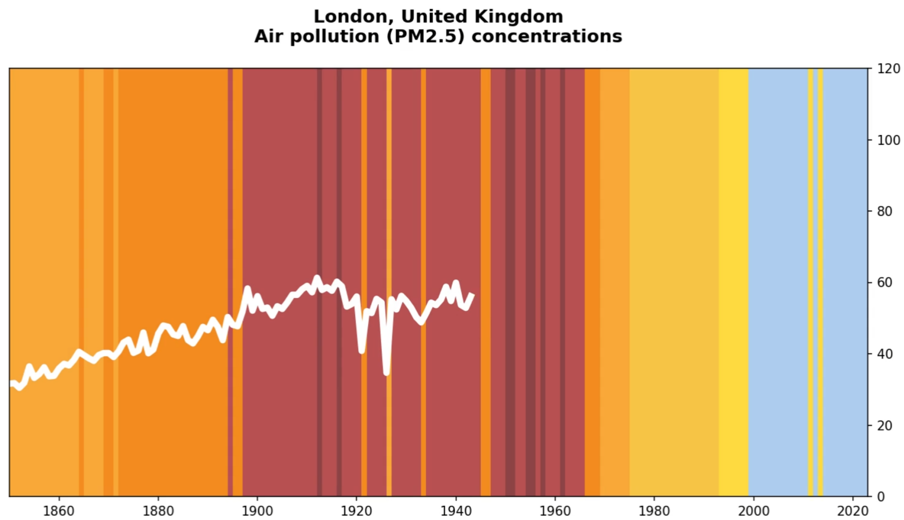

Figure 3: Still from an animated Air Quality Stripes: PM2.5 Concentrations in London Over Time [11]

The insights gained from this early-stage research have helped refine the scope of this project. The successful replication of static visualizations confirms that the dataset is correctly formatted, while the initial animation experiments suggest promising directions for interactive enhancements. Moving forward, we will focus on implementing user-driven interactions, optimizing performance for real-time web rendering, and developing comparative visual analytics tools to allow meaningful cross-region air quality analysis.

4 Project Proposal

This project aims to explore different visualization techniques and animation methods for representing historical PM2.5 trends, building upon the foundation of Air Quality Stripes [1], which serves as both the starting point and a key reference for this research. In addition to investigating novel ways to visualize air pollution data through dynamic and interactive animations, this project also aims to experiment with the development of a full-fledged system or dashboard that integrates these techniques into a functional and interactive platform.

Furthermore, accessibility is an important consideration in visualization design. While this project primarily focuses on animation techniques, it is essential to ensure that the visualizations remain accessible to a broad audience, including individuals with visual impairments or other accessibility needs. Design choices such as color contrast, alternative text descriptions, and user-friendly interaction mechanisms will be considered where applicable to enhance inclusivity.

Similarly, performance optimization is crucial to ensure the visualizations run smoothly across different devices. Performance testing will be considered to evaluate factors such as frame rate, rendering efficiency, and responsiveness on both high-performance and lower-end hardware. This will help identify potential bottlenecks and ensure that the visualizations remain usable without requiring excessive computational resources.

4.1 Success Criteria

The success of this project will be determined by its ability to effectively explore and present different visualization and animation techniques for PM2.5 trends while maintaining a conceptual connection to Air Quality Stripes [1]. Key indicators of success include:

- The development of multiple visualization prototypes demonstrating varied approaches to visualize PM2.5 trends, including both static, dynamic, and interactive representations within a dashboard environment.
- The effectiveness of these visualizations will be evaluated based on their ability to clearly convey PM2.5 trends, computational performance across devices, and adherence to data visualization best practices.
- Smooth performance and optimized rendering of large-scale air quality datasets using appropriate technologies such as WebGL and D3.js.
- Ensuring accessibility considerations, such as colorblind-friendly design and alternative text for key data insights.
- Code modularity and maintainability to allow future expansion and improvements, particularly for integrating the visualizations into a more comprehensive dashboard system.

4.2 Minimum Viable Product

If certain visualization techniques prove infeasible due to performance constraints or implementation challenges, alternative approaches will be explored. The minimum viable product (MVP) for this project will include:

- At least one successfully implemented visualization technique that extends or enhances Air Quality Stripes while improving interpretability.

- A prototype of dashboard that allows users to explore PM2.5 data dynamically.
- An assessment of the visualization's effectiveness, covering aspects such as clarity, engagement, and computational performance.

5 Workplan

The project is structured into multiple phases, ensuring steady progress toward the exploration of different visualization and animation techniques for PM2.5 trends. The workplan includes key milestones and deliverables, with sufficient flexibility for iterative improvements and contingency planning. The project is structured into multiple phases, ensuring steady progress toward the exploration of different visualization and animation techniques for PM2.5 trends. The workplan includes key milestones and deliverables, with sufficient flexibility for iterative improvements and contingency planning. The Gantt chart for this plan can be found in the project's GitLab repository [12].

1. **Project Preparation (Jan 27 – Mar 14, 2025)**
    - Conduct a literature review on air pollution visualization.
    - Explore and analyze the PM2.5 dataset used in Air Quality Stripes.
    - Identify suitable visualization and animation techniques.
    - **Milestone:** Completion of project preparation.
2. **Data Preparation & Initial Prototyping (Mar 1 – Apr 20, 2025)**
    - Develop static baseline visualizations.
    - Create initial prototypes with different animation styles.
3. **Implementation & Optimization (Apr 10 – May 16, 2025)**
    - Develop multiple animated visualization prototypes.
    - Optimize performance for smooth rendering.
    - **Milestone:** Project Preparation Presentation Slides completed.
4. **Refinement & Comparative Analysis (May 1 – Jul 5, 2025)**
    - Conduct comparative analysis of visualization techniques.
    - Collect feedback and refine visualizations.
    - **Milestone:** Comparative analysis completed.
5. **Finalization & Dissertation Writing (Jun 15 – Aug 10, 2025)**
    - Finalize the best visualization approach.

- Complete dissertation writing and final visualization refinements.
- **Milestone:** Final dissertation submission.

6 Resources Estimation

The resource requirements for this project are minimal. The MVP only involves writing simple Python and JavaScript code to explore different visualization techniques, which can be run locally without significant computational demands.

If the project is expanded to include a web-based visualization, then hosting options will need to be considered. At that point, decisions will be made on where to host the website, such as GitHub Pages, Netlify, or other platforms.

7 Risk Analysis

Risk management is a continuous process throughout the dissertation work. The key risks in this project relate to technical challenges, time management, and resource availability. The table below summarizes the identified risks, their probability and impact levels, and the planned corrective actions.

Risk	**Probability**	**Impact**	**Corrective Action**
Learning Curve for Visualization Libraries (D3.js, WebGL, Python Animation)	Probable	Moderate	Allocate additional time for learning in early phases; use simpler alternatives if necessary.
Technical Issues (Bugs, Unexpected Performance Problems)	Possible	Moderate	Follow best practices in debugging; consult documentation and online resources.
Delays in Prototype Development	Possible	Moderate	Prioritize core functionalities; create an initial simple prototype before refining further.
Overestimation of Scope	Possible	Severe	Focus on MVP first; cut non-essential features if needed.
Resource Constraints (Computing Power, Storage)	Unlikely	Low	Use local computing resources, as heavy processing is not required for MVP.
Unexpected Personal Commitments or Health Issues	Possible	Moderate	Buffer time is included in the schedule; adjust milestones if needed.

References

[1] Air Quality Stripes. *Air Quality Stripes.* Available at: https://airqualitystripes.info/ (Accessed: 4 March 2025).

[2] U.S. Environmental Protection Agency. *Health and Environmental Effects of Particulate Matter (PM).* Available at: https://www.epa.gov/pm-pollution/health-and-environmental-effects-particulate-matter-pm (Accessed:5 March 2025).

[3] Li, H., Chen, J., Zhang, X., and Xu, H. (2016). *A Visualization Approach to Air Pollution Data Exploration – A Case Study of Air Quality Index (PM2.5) in Beijing, China.* Atmosphere, 7(3), 35. doi:10.3390/atmos7030035.

[4] Lu, W., He, J., and Sun, H. (2017). *An Interactive Web Mapping Visualization of Urban Air Quality Monitoring Data of China.* Atmosphere, 8(8), 148. doi:10.3390/atmos8080148.

[5] Powley, B., Anslow, C., and Pearce, D. J. (2022). *AtmoVis: Web-Based Visualization of Air Quality Data with Interconnected Windows.* Workshop on Visualization in Environmental Sciences (EnvirVis). doi:10.2312/envirvis.20221055.

[6] Ashouri, M., Phoa, F. K. H., Chen, C.-H., and Shmueli, G. (2023). *An interactive clustering-based visualization tool for air quality data analysis.* Aerosol and Air Quality Research, 23(12), 230124. doi:10.4209/aaqr.230124.

[7] Luo, X., Jiang, R., Yang, B., Qin, H., and Hu, H. (2024). *Air quality visualization analysis based on multivariate time series data feature extraction.* Journal of Visualization, 27, 567–584. doi:10.1007/s12650-024-00981-3.

[8] Commonwealth Scientific and Industrial Research Organisation (CSIRO). (2015). *Smoke and Fire: WebGL-Based Global Air Pollution Visualization.* Retrieved from https://research.csiro.au/static/airquality/smoke-and-fire/.

[9] U.S. Environmental Protection Agency. (n.d.). *Environmental Benefits Mapping and Analysis Program (BenMAP)*. Retrieved from https://en.wikipedia.org/wiki/BenMAP.

[10] Appsilon. (2023). *Air Quality vs Respiratory Diseases Dashboard: An R Shiny and D3.js Visualization Case Study*. Retrieved from https://www.appsilon.com/post/air-quality-vs-respiratory-diseases.

[11] s2504850, *Week4_0220_0225 - Learning.* GitLab repository, 2024. [Online]. Available: https://git.ecdf.ed.ac.uk/msc-24-25/s2504850/-/tree/main/Learning/Week4_0220_0225?ref_type=heads.

[12] s2504850, *Gantt Chart Updated.* GitLab repository, 2025. [Online]. Available: https://git.ecdf.ed.ac.uk/msc-24-25/s2504850/-/blob/main/Work%20plan/gantt_chart_updated.png