Project Report on Cross Data Analytic on Air Quality in Victoria

Junkai Zhang (1166036)

Yufei Li (1338325)

Ming Hsiu Tsai (1397638)

YuCheng Chien (1375065)

Fan Pu (1372387)

May 21, 2024

Abstract

This research endeavours to uncover the intricate relationship between air quality and geo-social factors, specifically focusing on chronic obstructive pulmonary disease (COPD), the number of vehicles per dwelling, and real estate prices in Victoria, Australia. With the advent of cloud computing technologies, this study leverages innovative methodologies to analyse vast datasets efficiently. By deploying Kubernetes (k8s) for orchestration and utilising ElasticSearch and Fission for data management and analysis, the research aims to provide insights into how these geo-social factors interplay with air quality. The findings of this research are expected to contribute significantly to environmental policy-making, urban planning, and public health initiatives. By elucidating the impact of geo-social factors on air quality, stakeholders can develop targeted interventions and strategies to mitigate air pollution and improve overall public health and well-being in Victoria, Australia, and potentially in other regions facing similar challenges.

Keywords: Air Quality, Kubernetes, Elastic Search, Fission, Data, Victoria, Australia, COPD, Real estate.

Links:

Github Repo: https://github.com/Lisherex/comp90024_ass2/tree/master

Youtube: https://youtu.be/rRkdXxZe9ww

Contents

1 Introduction

In this project, our primary objective was to address the fundamental question: How do socioeconomic factors, specifically respiratory system disease admissions and regional property prices, fluctuate under different air quality conditions? To systematically explore this question, we delineated it into six minor scenarios, each aiming to dissect a specific aspect of the relationship between air quality and socioeconomic indicators.

These scenarios include:

1. Investigating the correlation between air quality and Chronic Obstructive Pulmonary Disease (COPD) admissions.
2. Examining the association between air quality and regional average property prices.
3. Analysing the interplay between air quality and Respiratory System Disease (RSD) admissions.
4. Exploring the relationship between air quality and the number of vehicles per dwelling.
5. Understanding the connection between regional average property prices and COPD admissions.
6. Investigating the relationship between regional average property prices and the number of vehicles per dwelling.

By dissecting the overarching question into these minor scenarios, we aimed to gain a holistic understanding of how air quality interacts with socioeconomic factors and influences social behaviours. To achieve this, we aggregated data from various sources, including the Environmental Protection Agency (EPA) and SUDO databases.

To facilitate our analysis, we leveraged advanced cloud computing technologies. Specifically, we developed a Kubernetes Cluster on the Melbourne Research Cloud and deployed data management indexes on Elasticsearch. Additionally, we utilised Fission functions to streamline data processing tasks. Through these technological innovations, we were able to unveil the socioeconomic impact of air quality, shedding light on its implications for public health and urban planning initiatives.

2 System Architecture & Infrastructure

This section aims to briefly introduce the infrastructure on which this project operates and outline the architecture of the application. A detailed discussion of the challenges encountered during development will be addressed in the Discussion section.

2.1 Melbourne Research Cloud (MRC)

This application has been developed and is currently operational on virtual machines provided by the MRC. As a 24/7 private cloud service, MRC facilitates continuous and secure deployment of scripts, applications, and research projects. The platform, through its user-friendly interface, offers a diverse array of operating system images, thereby addressing a wide range of user requirements.

For this specific project, six instances with a total of eleven virtual CPUs (VCPUs) were allocated. Additionally, the pre-configured images provided by MRC were leveraged, which significantly expedited the configuration process. This efficient allocation of resources, combined with the availability of customizable operating system images, has ensured that the application operates optimally, meeting the project's computational and operational demands.

2.2 Kubernetes

This system architecture diagram illustrates the core components and their interactions within this project's Kubernetes cluster. Our K8s system embeds 3 nodes (Node 0, Node 1, Node 2), each running distinct controllers and Pods:

1. **Services:** Kibanakibana and elasticsearchmaster are the main services for this project. These services facilitate routing traffic to the corresponding Pods.

2. **Label Selector:** The Label Selector mechanism is employed to route service traffic to Pods with matching labels, ensuring that the appropriate Pods receive the requests.

3. **Node Distribution:**

    - **Node 0:** Hosts a ReplicaSet that manages a Pod named kibanakibana5bdc8d7b97-pfz49.
    - **Node 1:** Hosts a StatefulSet that manages a Pod named elastic-master-0.

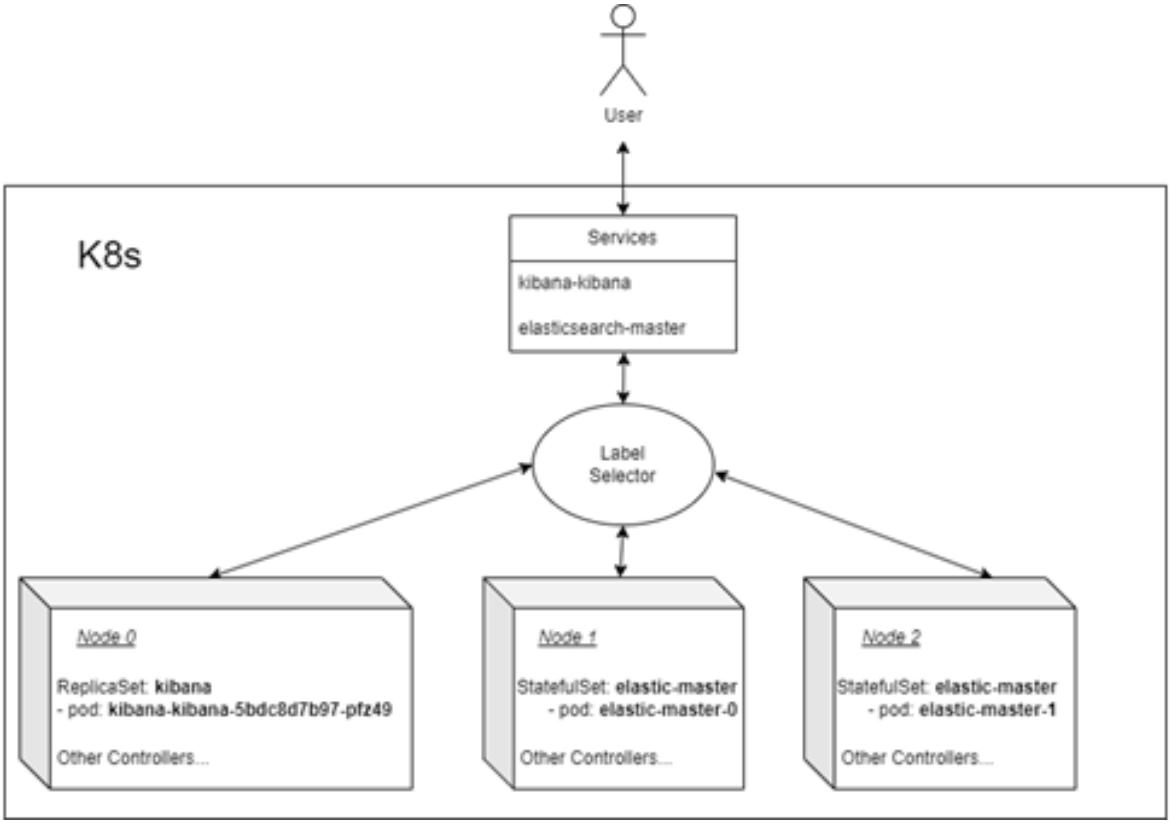

Figure 1: System Architecture of K8s. (Communication through headless service is omitted.)

- **Node 2:** Hosts a StatefulSet that manages a Pod named elastic-master-1.

This architecture ensures that different types of applications (stateless and stateful) can be efficiently operated and managed. Services utilise the Label Selector to route traffic to the correct Pods, facilitating both intra-cluster and inter-cluster communication. The Pods on each node are managed by their respective controllers (ReplicaSet and StatefulSet), ensuring high availability and stability of the applications.

2.3 Elastic Search

To facilitate the analysis and storage of big data, ElasticSearch was employed. With the aid of Elasticsearch, this project flexibly processes and integrates data from various sources. Its RESTful searching feature, combined with Fission, enables us to ensure data consistency and integrity through a backend-driven architecture, while also reducing the complexity of front-end development.

2.4 Fission: Serverless Framework

This project used Fission to develop relevant RESTful API and fucntionalities. As an open-source, Kubernetes-native serverless framework, it designed to enable rapid development and deployment of functions without the need for managing underlying server infrastructure, facilitating an event-driven architecture and supports multiple programming languages, thereby providing developers with a streamlined approach to building, updating, and scaling applications.

3 Server-side RESTful API design

In our system, the server-side utilises a RESTful architecture to interact effectively with the front-end, specifically our Jupyter Notebook interface. REST (Representational State Transfer) is a software architectural style that emphasises the separation of concerns. This setup allows our servers to expose data and functionalities consistently and predictably, making them accessible to a wide variety of clients.

3.1 Design Principles

Our choice of RESTful APIs is driven by their simplicity, flexibility, and widespread acceptance. These APIs use standard HTTP methods such as GET, PUT, POST. This compatibility simplifies development and debugging while enhancing support for various network infrastructures, including caches and proxies. The adaptability of RESTful APIs is essential for catering to our project's diverse data interaction needs, particularly for dynamic data queries and presentations.

3.2 Implementation Details

The RESTful API on our server is structured to support a variety of data interactions:

- **Data Retrieval:** Through HTTP GET requests, front-end applications like our Jupyter Notebooks can request specific datasets from the server. This is evident in functions like `get_airquality_vehicles` and `get_airquality_copd`, which fetch relevant environmental data.

- **Data Update:** Using HTTP PUT requests, the server receives data from the front-end, which might include user-defined parameters for data analysis. This is managed by functions such as `putHousePrice` and `putVehicle`, which update the data stored in our backend.

3.3 Security Considerations

Although our project does not mandate authentication or authorization mechanisms for front-end interactions, we secure our API by enforcing HTTPS on all data transfers to prevent eavesdropping or tampering during transmission.

3.4 Integration with Jupyter Notebook

Our Jupyter Notebook interfaces with the server by issuing HTTP requests, and the server responds with data in JSON format. This method ensures real-time data updates and enhances the flexibility and dynamism of data visualisation in the notebooks.

3.5 Fission and RESTful APIs

We use Fission to encapsulate code into functions that can be invoked through HTTP, simplifying the deployment of serverless functions on our Kubernetes cluster. Each function, such as get_airquality_vehicles or putHousePrice, is exposed as a RESTful endpoint. This modular approach aligns with the client-server communication model using standard HTTP methods, facilitating straightforward interactions between our Jupyter Notebooks and server-side functions.

3.6 Example Use Case

For instance, to retrieve vehicle-related air quality data, a Jupyter Notebook might send an HTTP GET request to the get_airquality_vehicles function deployed in Fission. Fission processes this request and routes it to the designated function, returning the execution results directly to the notebook. This interaction demonstrates the practical application of our RESTful API in real-world scenarios.

Our project employs Fission to implement RESTful APIs, significantly decoupling backend functionalities and enabling dynamic data interactions. This architecture not only improves

the maintainability and scalability of the system but also ensures flexible and efficient data processing.

4 Data Collection

4.1 Environmental Protection Authority : Air Quality

The Environmental Protection Authority (EPA) Air Quality Data API offers a vital resource for monitoring and analysing air quality conditions in Victoria, Australia. With growing concerns over the impacts of air pollution on public health and the environment, access to accurate and real-time air quality data is essential for informed decision-making and effective environmental management.

In this project, EPA Air Quality Data API was used for collecting real time particulate matter data and provided a universal station_id index for cross dataset merge. The air pollution data retrieved from EPA contains the geographical information of EPA monitoring stations, 32 digit unique station id, the time interval of latest air quality data around Victoria and the value on targeted pollution. This analytic utilises the geographical location of EPA monitoring stations as the place of interest to link the relationship between air quality and other key factors.

4.2 Spatial Urban Data Observatory

The Spatial Urban Data Observatory (SUDO) represents a pioneering platform at the forefront of urban data analysis and decision-making. Developed to address the complexities of modern urban environments, SUDO serves as a comprehensive repository of spatial data, offering valuable insights into various aspects of urban life, infrastructure, and sustainability.

SUDO harnesses the power of spatial data science, integrating diverse datasets from sources such as satellite imagery, sensor networks, administrative records, and social media streams. By leveraging advanced analytical techniques and visualisation tools, SUDO enables stakeholders to explore, analyse, and interpret urban data in unprecedented ways.

4.3 Chronic Obstructive Pulmonary Disease

The dataset used for gathering data about chronic obstructive pulmonary disease is the comprehensive collection provided by the Population Health Information Development Unit (PHIDU).

This dataset encapsulates a wide array of indicators spanning demographic and social aspects, health status, disability, carers, deaths, as well as the utilisation and provision of health and welfare services.

Central to understanding the dataset are the statistical methodologies employed. Age-standardised rates and ratios, predominantly based on the Australian standard, serve as the foundation for analysis. Quintiles of socioeconomic disadvantage are calculated based on the 2021 Index of Relative Socio-economic Disadvantage (IRSD), while Remoteness Areas utilise the ABS Remoteness Structure of either 2016 or 2021.

Additionally, modelled estimates, provided by the Australian Bureau of Statistics (ABS) and PHIDU, offer insights into health risk factors in areas where administrative data sets are lacking. These estimates, though indicative, are invaluable for informing policy, program development, and decision-making processes, particularly in areas where localised data is scarce.

4.4 Number of Vehicle by dwelling

This part of data is collected from SUDO, ABS Census - E39 Structure Of Dwelling By Number Of Motor Vehicles By Number Of Persons (Usually Resident) (SLA) 1991.

The 1991 Census Expanded Community Profiles offer a comprehensive view of Australian communities, providing detailed insights into various demographic and housing characteristics. These profiles serve as invaluable resources for researchers, policymakers, and analysts seeking to understand the intricacies of Australian society at a local level. One notable aspect of the 1991 Census Expanded Community Profiles is the inclusion of 44 tables, each offering more detailed information compared to the basic community profiles. These tables delve deeper into demographic, social, and housing data for Statistical Local Areas (SLAs) across Australia.

The granularity of this dataset allows for a nuanced understanding of housing dynamics, particularly in relation to transportation infrastructure and household composition. By examining the distribution of motor vehicles across different household sizes, researchers can discern patterns of car ownership, household size, and potentially, socioeconomic status.

It's important to note that the data is delineated by SLA 1991 boundaries, providing a snapshot of communities within this geographic framework. The periodicity of the data is 5-yearly, offering insights into longitudinal trends and changes over time.

4.5 Property Price

The housing market serves as a fundamental aspect of any urban landscape, reflecting economic dynamics, societal trends, and individual preferences. As a crucial component of urban studies and economics, analysing housing data offers valuable insights into various aspects of a city's development and livability. In this context, the dataset under examination provides a comprehensive glimpse into Melbourne's vibrant housing market.

Sourced from a publicly available GitHub repository, the dataset encapsulates a wealth of information regarding property prices across Melbourne's diverse suburbs. Composing variables such as suburb, address, number of rooms, property type, price, selling method, and more, this dataset offers a rich repository for exploration and analysis.

4.6 Data Processing

To collect EPA air quality data, we utilised the EPA environment monitoring API. Through this API, we obtained comprehensive air quality data comprising station information including ID, name, and geographical location of the EPA monitoring stations, along with monitoring records on selected pollutants. Subsequently, we employed an "update-airquality" Fission function to upload the acquired data to a designated index on Elasticsearch. This ensures that the air quality data from the EPA API is regularly updated and replaces outdated records. To facilitate meaningful comparisons regarding air quality, our focus was specifically on the pollutant "PM 2.5", resulting in 90 EPA monitoring stations with valid air quality data.

In contrast, the process for gathering SUDO data differed as there is no API provided by SUDO for data retrieval. We obtained COPD and vehicle per dwelling data separately, downloading them and subsequently uploading them to a specific index on Elasticsearch using the "upload-csv-file" function in Fission. To optimise the utilisation of SUDO data, various data processing actions were performed, including linking the geographical location in SUDO data entries to the geographical location in EPA air monitoring stations. This standardisation of geographical representation across different datasets enabled the merging of separate datasets to unveil underlying relationships.

Property price data was sourced from a GitHub repository, reformatted from .csv to .json format, and then uploaded via the "uploadcsvfile" function within Fission for further processing.

5 API Introduction

Item	Method	URL	Ingress	Namespace
1	[GET]	/airquality/copd	true	default
2	[GET]	/airquality/houseprice	true	default
3	[GET]	/airquality/rsd	true	default
4	[GET]	/airquality/vehicle	true	default
5	[GET]	/houseprice/copd	true	default
6	[GET]	/houseprice/vehicle	true	default
7	[POST]	/post-json-data-from-local/index	false	default
8	[PUT]	/update-airquality	false	default

Table 1: API Endpoints and Methods

1. **[GET] /airquality/copd**

   This endpoint retrieves air quality station data specifically related to Chronic Obstructive Pulmonary Disease (COPD). It is designed to help researchers and healthcare professionals understand the correlation between air quality levels and the prevalence or severity of COPD in different regions. The data can be used for epidemiological studies, public health assessments, and the development of intervention strategies to mitigate the impact of poor air quality on COPD patients.

2. **[GET] /airquality/houseprice**

   This endpoint provides access to air quality station data in relation to house prices. It aims to assist urban planners, real estate professionals, and researchers in analysing how air quality influences real estate market values. The data can be used to identify trends and correlations between environmental quality and property prices, aiding in market analysis and policy-making to improve urban living conditions.

3. **[GET] /airquality/rsd**

   This endpoint retrieves data from air quality stations that is relevant to Respiratory System Disease (RSD). The information is valuable for public health officials and researchers studying the impact of air pollution on respiratory diseases. By examining this data, it is possible to identify areas with higher risks of RSD outbreaks and to formulate strategies for improving air quality to protect public health.

4. **[GET] /airquality/vehicle**

   This endpoint provides data on air quality in relation to the number of vehicles per

dwelling. It is useful for urban planners, environmental scientists, and policy-makers interested in understanding the impact of vehicular emissions on air quality. The data can help in developing sustainable urban transportation policies and initiatives to reduce vehicle emissions and improve air quality.

5. **[GET] /houseprice/copd**

   This endpoint offers house price data linked with COPD statistics. It helps researchers, healthcare economists, and policy analysts explore the relationship between real estate values and the health status of populations, particularly those affected by COPD. The data can inform decisions on healthcare resource allocation and community health initiatives aimed at improving living conditions for COPD patients.

6. **[GET] /houseprice/vehicle**

   This endpoint retrieves house price data in relation to the number of vehicles per dwelling. It is designed to assist urban planners, real estate analysts, and environmental scientists in studying how vehicle density in residential areas affects property values. The data can be used to develop strategies for sustainable urban development and to promote environmental policies that enhance property values by reducing traffic congestion and pollution.

7. **[POST] /post-json-data-from-local/{index}**

   This endpoint allows users to upload JSON data from local sources to ElasticSearch. It is a vital tool for data engineers and system administrators responsible for maintaining and updating the database. By facilitating the ingestion of local data, this endpoint ensures that the system's datasets are comprehensive and up-to-date, supporting robust data analysis and application functionality.

8. **[PUT] /update-airquality**

   This endpoint updates the air quality station data from the Environmental Protection Agency (EPA). It is crucial for maintaining accurate and current data in the system. Environmental scientists, data analysts, and policy-makers rely on this endpoint to ensure they are working with the latest air quality information, which is essential for environmental monitoring, public health assessments, and policy development aimed at improving air quality standards.

5.1 Testing & Error Handling

Time Efficiency and Team Dynamics:

- **Testing is essential because it dramatically reduces the time spent debugging and troubleshooting in later stages of the development cycle.** By catching bugs early, teams can avoid costly delays close to deployment.
- **Proper testing mitigates internal friction by clearly identifying issues without ambiguity about their origin, whether in the user interface, the database, or the integration points.**

Preventing and Managing Malfunctions

- **FaaS applications, like any software, are susceptible to malfunctions due to various factors including poor coding practices or unforeseen usage scenarios.**
- Regular testing ensures that the system can handle both expected and unexpected user behaviours, and helps in maintaining robust error handling and system stability.

Testing Statements from Iterative Development

- **First Iteration:** Basic tests to ensure that each endpoint responds with a status code of 200, confirming that the endpoint is reachable and functioning as expected.
- **Second Iteration:** Introduction of tests for JSON payload integrity and the correctness of responses, validating not just connectivity but also functional correctness.
- **Third Iteration:** Enhanced tests for data accuracy, including checks against expected data values and structures returned by API calls, ensuring data retrieved from Elasticsearch is correctly processed and delivered.

6 Results & System Functionality Illustration

In this section, we will discuss and demonstrate the functionality of our data analytic system and present the result and image outcome along with our findings and recommendations on each scenario.

6.1 Air Quality vs COPD Admissions

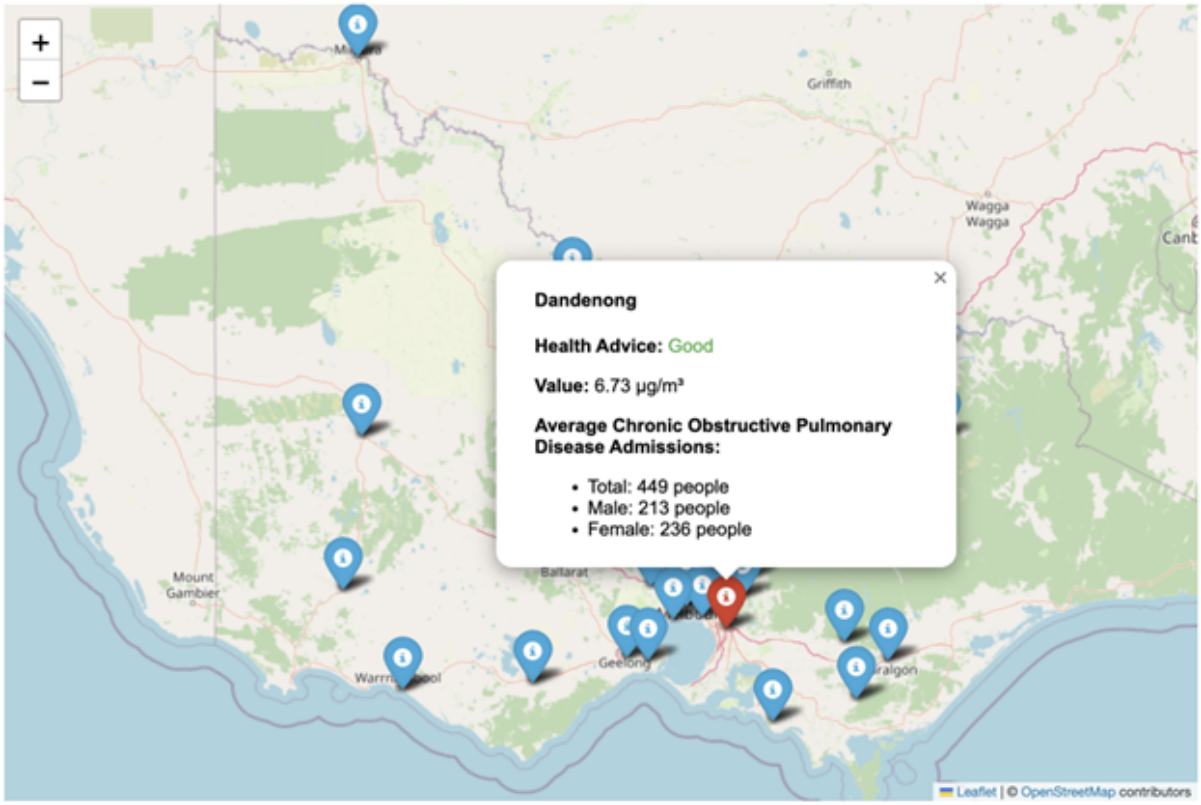

Figure 2: Map for Air Quality vs COPD Admissions.

The popup indicates a notably high number of COPD admissions (449 total) in the area, suggesting a significant prevalence of the disease in Dandenong. Interestingly, the air quality levels recorded suggest that air pollution was not a significant issue on the day of the data snapshot. This discrepancy implies that the high incidence of COPD in the region might be influenced by factors other than current outdoor air pollution levels, such as occupational exposures, socioeconomic and lifestyle factors, genetic predispositions, and overall health conditions.

6.2 Air Quality vs Average House Price

Altona North features exceptional air quality, with a PM2.5 concentration of just 2.45 µg/m³, a characteristic shared throughout Victoria. This uniformity in air quality across the region implies that while all suburbs, including Altona North, benefit from excellent air conditions, this factor does not directly influence the observed high average house prices. Instead, property values are more likely shaped by factors such as location, amenities, and socioeconomic status, rather than by environmental quality. This consistency highlights that in Victoria, good air

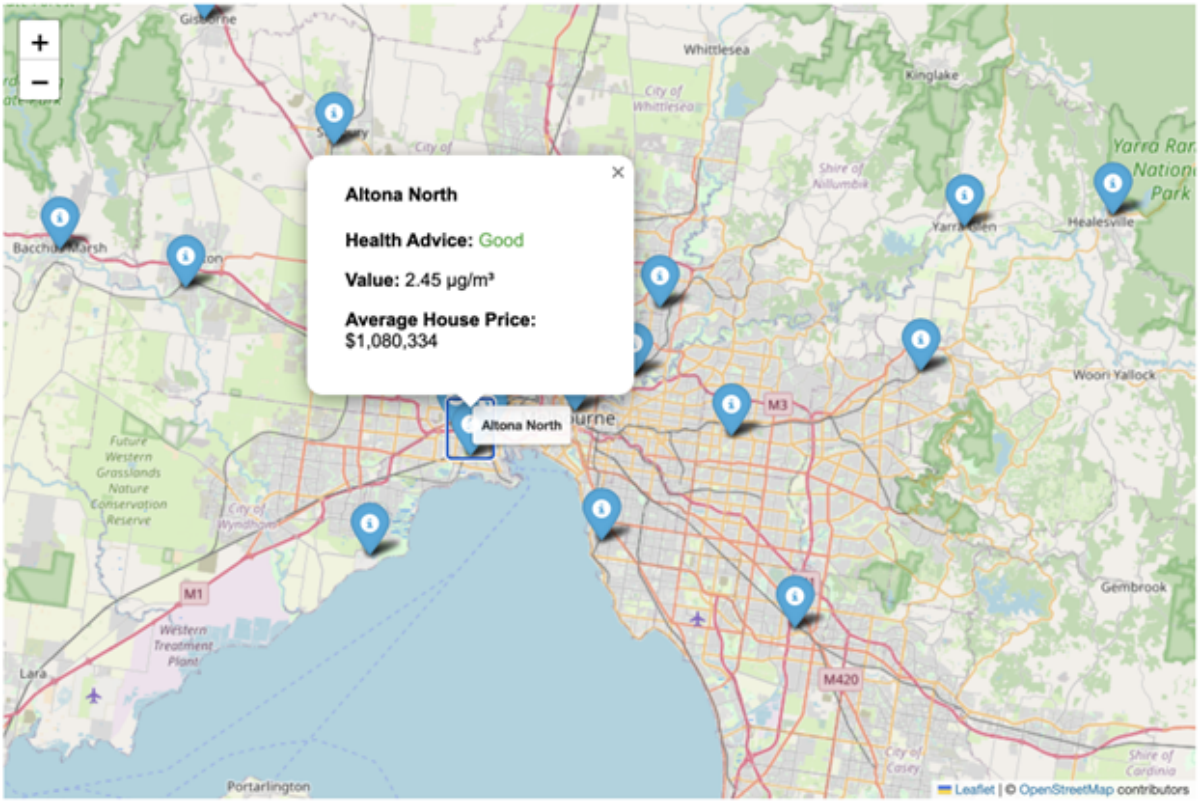

Figure 3: Map for Air Quality vs Average House Price.

quality is a standard condition and does not serve as a key differentiator in real estate pricing dynamics.

6.3 Air Quality vs Respiratory System Disease Admissions

Wallan is noted for having the highest number of respiratory system disease admissions in Victoria, with a total of 3,361 cases, where males slightly outnumber females. Despite this high incidence of respiratory diseases, the air quality in Wallan remains classified as good, with a PM2.5 value of 7.65 µg/m³. This suggests that the high rate of respiratory disease admissions may not be directly linked to current levels of air pollution.

This scenario indicates that other factors besides air quality might be influencing the high rates of respiratory illnesses in Wallan. Potential contributing factors could include genetic predispositions, indoor air quality issues, occupational exposures, or lifestyle choices such as smoking. The consistent pattern of good air quality across Victoria, including in regions with high disease rates, reinforces the need for further investigation into these alternative influences.

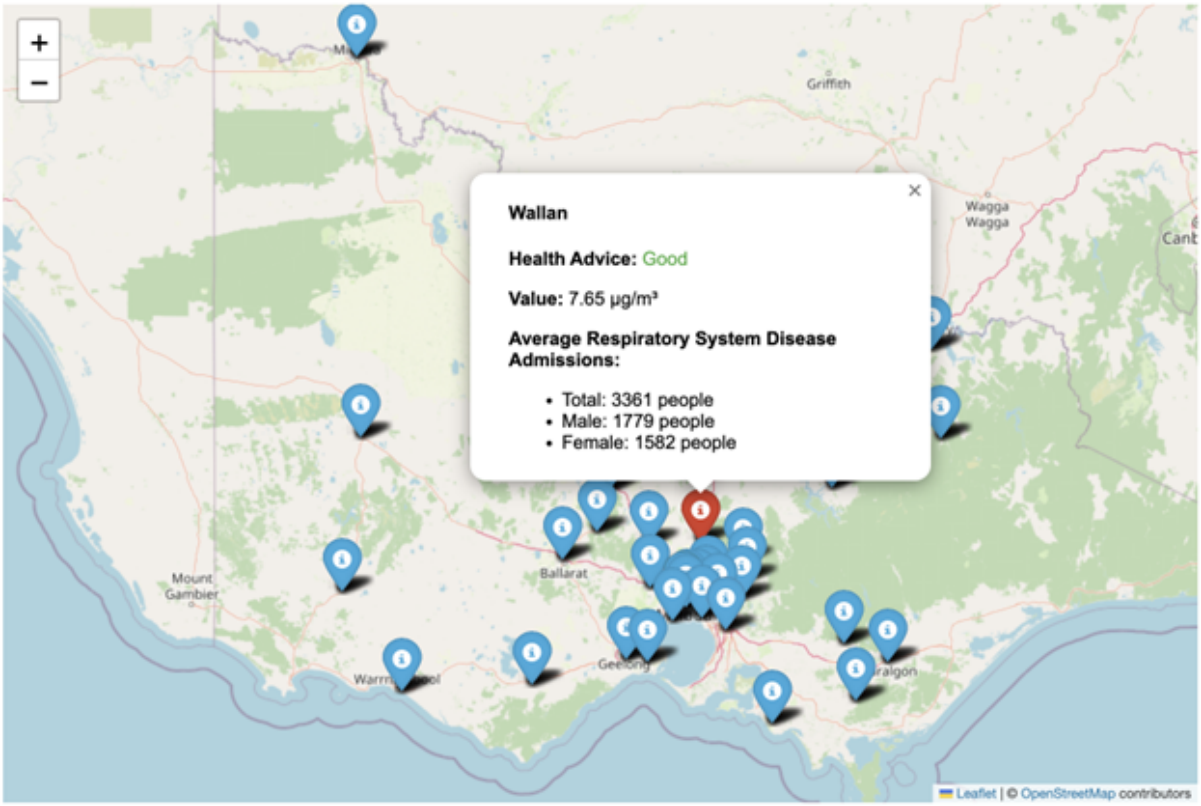

Figure 4: Map for Air Quality vs Respiratory System Disease Admissions.

6.4 Air Quality vs Vehicle Data per Dwelling

Based on the data provided and the figure, there is no clear relationship between air quality (measured by PM2.5 concentrations) and the number of vehicles per dwelling across various locations in Victoria. Air quality remains good across most sites, regardless of variations in vehicle ownership. This suggests that factors other than vehicle density, such as environmental management practices and vehicle emission standards, may play a significant role in maintaining low pollution levels. Further analysis would be needed to explore the specific influences on air quality in these areas.

6.5 Average House Price vs COPD Admissions

Correlation Between House Prices and COPD Admissions:

- **The data suggests that there is a relationship between house prices and the number of COPD admissions.**
- Mid-range house prices ($600,000 - $800,000) show higher COPD admissions for both males and females, which contributes significantly to the total admissions.

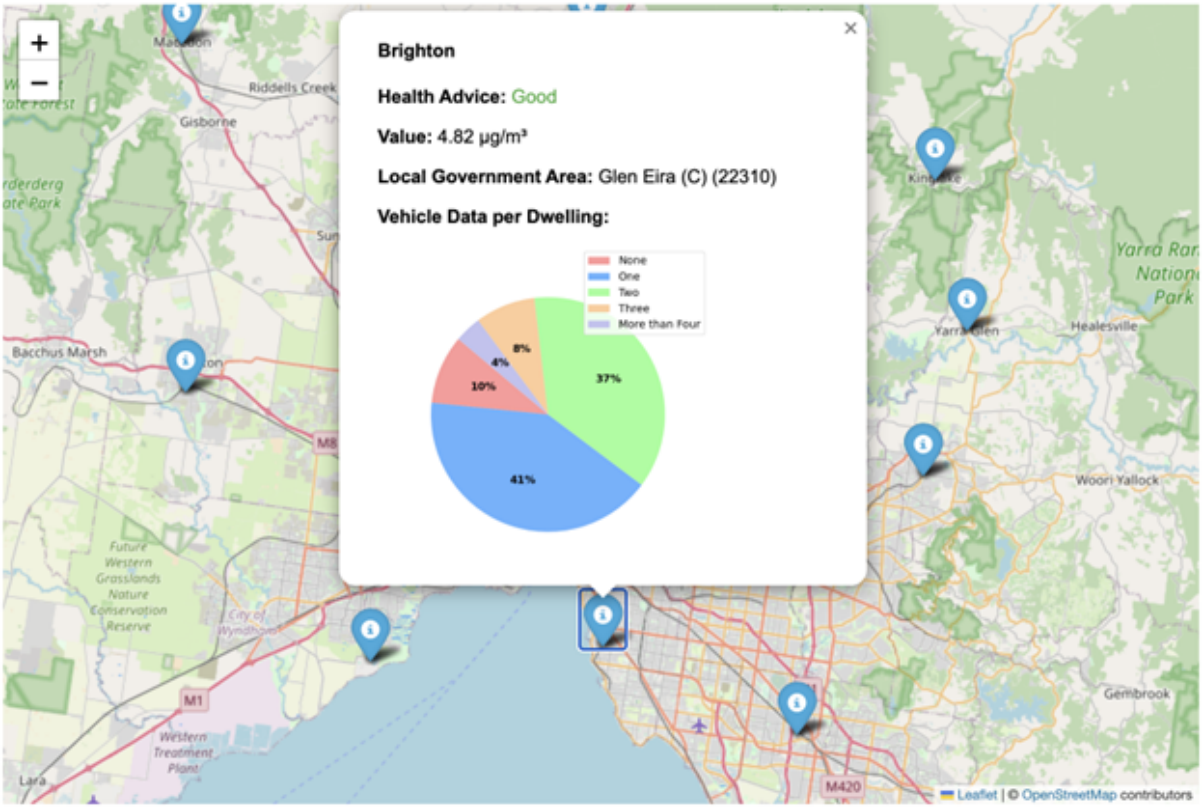

Figure 5: Map for Air Quality vs Vehicle Data per Dwelling.

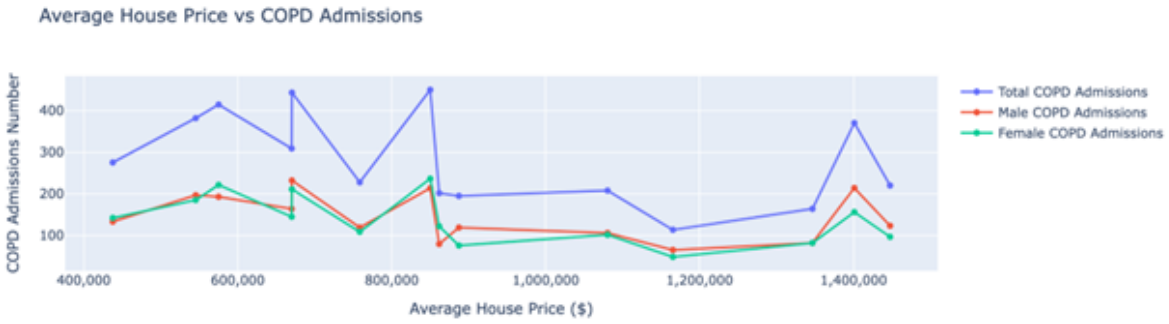

Figure 6: Average Houre Price vs. COPD Admission.

- There are peaks in COPD admissions at specific house price points ($600,000 and $1,400,000), indicating potential socioeconomic factors affecting health outcomes.

Gender Differences in COPD Admissions

- Both male and female COPD admissions show similar patterns, peaking at the same house price points.
- Female admissions slightly exceed male admissions at some points, suggesting that fac-

tors affecting COPD admissions might impact genders differently but generally follow a similar trend.

The chart reveals significant peaks in COPD admissions at specific average house price points, suggesting that socioeconomic factors related to housing may influence health outcomes. Both male and female COPD admissions follow similar patterns, with slight variations in magnitude. Further research is necessary to understand the underlying causes of these trends and to explore potential interventions to mitigate the health impacts of housing-related socioeconomic factors.

6.6 Average House Price vs Vehicle Data per Dwelling

Figure 7: Average House Price vs Vehicle Data per Dwelling.

- **The data suggests that there is a relationship between house prices and the number of COPD admissions.**
- Mid-range house prices ($600,000 - $800,000) show higher COPD admissions for both males and females, which contributes significantly to the total admissions.
- There are peaks in COPD admissions at specific house price points ($600,000 and $1,400,000), indicating potential socioeconomic factors affecting health outcomes.

Gender Differences in COPD Admissions

- Both male and female COPD admissions show similar patterns, peaking at the same house price points.

- Female admissions slightly exceed male admissions at some points, suggesting that factors affecting COPD admissions might impact genders differently but generally follow a similar trend.

7 Discussion Future Improvements

7.1 Fault Tolerance

This application has strong resilience against single-point failures. Specifically, thanks to Kubernetes' node pressure eviction, workloads are automatically transferred or redistributed to other available nodes or instances to maintain service availability and continuity.

In our project, using Elasticsearch as an example, which is deployed as a StatefulSet, its health status is continuously monitored by the Controller Manager. If a node remains unresponsive for an extended period, it will be marked as NotReady. Subsequently, the Kubernetes scheduler will reschedule the related Elasticsearch pods to other nodes or evict them and then recreate them on other nodes.

Overall, this robust mechanism ensures that our application remains highly available and continues to function smoothly, even in the face of individual node failures.

7.2 Insufficient Data

Although data accuracy is not a primary requirement for this project, it is worth discussing the potential limitations.

The air quality data were collected through the EPA's database; however, the historical data is not accessible. The EPA updates its data every seven days, making it particularly inaccurate to correlate the latest air quality data with historical data, such as hospital admission, from SUDO.

Furthermore, upon examining the geographical distribution of EPA's monitor stations, stations are widely scattered across the Melbourne regions. The significant geographical span makes the air quality data between meteorological stations inaccurate.

For future improvements, it is necessary to collaborate with the EPA to obtain more detailed historical meteorological data. Additionally, to address the issue of incomplete coverage due to the uneven distribution of monitoring stations, more complex and refined data processing

methods, such as Spatial Interpolation Methods or deep learning, can be explored and utilized to enhance data quality.

7.3 Limitations of Using Research Cloud for Public-Facing Applications

Although MRC demonstrates strong capability in processing complex and large datasets, effectively balancing and reducing the computational load on the user side, as a research-oriented cloud service provider, MRC may face challenges in providing a comprehensive commercial solution comparable to platforms such as Azure.

Additionally, as a private cloud service at the University of Melbourne, MRC faces significant access challenges for external traffic from outside institutions, which limits the audience for the developed applications.

8 Team Contribution

Name	**Role**	**General Contribution**	**Specific Contribution**
Junkai Zhan	Backend Developer	Data Collection, Fission, ElasticSearch, K8s Cluster Construction, Report	Data Process, Index Management, EV
Yufei Li	DevOps Engineer		Automating deployment processes, PoC, EV
Ming Hsiu Tsai	Full-Stack Developer		MRC, Data Visualisation, End-to-end Test, EV
YuCheng Chien	Full-Stack Developer		Project management, Code review
Fan Pu	Backend Developer		End-to-end Test

Figure 8: Team work contribution

9 Discussion on teamwork

During the developmental stage, each member of Group 7 has demonstrated commendable commitment and enthusiasm towards system creation and group discussions. The active participation of every member, their prompt response to group calls, and punctuality in attending weekly development sessions have significantly contributed to the overall productivity and cohesion within the team. Working with such dedicated individuals has been a rewarding experience, as everyone consistently fulfills their individual responsibilities with diligence.

Despite the positive dynamics within the group, there remains an area for improvement in the distribution of workload throughout the development phase. It has been observed that a significant portion of the work tends to accumulate towards the final week, leading to potential challenges and stress in meeting deadlines. To enhance efficiency and mitigate such challenges, it is imperative to distribute the workload more evenly across the entire development period. By adopting a more balanced approach to task allocation, we can ensure smoother progress and alleviate undue pressure on team members, thereby optimizing our collective productivity and achieving better outcomes.

References

[1] PHIDU. (2024). *Admissions - Principal Diagnosis: Males (LGA) 2017-2018*. Retrieved May 10, 2024, from https://sudo.eresearch.unimelb.edu.au

[2] PHIDU. (2024). *Admissions - Principal Diagnosis: Females (LGA) 2016-2017*. Retrieved May 10, 2024, from https://sudo.eresearch.unimelb.edu.au

[3] PHIDU. (2024). *Admissions - Principal Diagnosis: Persons (LGA) 2016-2017*. Retrieved May 10, 2024, from https://sudo.eresearch.unimelb.edu.au

[4] Environment Protection Authority Victoria. (2024). *EPA*. Retrieved May 10, 2024, from https://epa.vic.gov.au

[5] Steven Obadja. (2024). *house_prices_melbourne*. Retrieved May 12, 2024, from GitHub Repository: https://github.com/stevenobadja/house_prices_melbourne