**Capstone Project**

Exploring Takapuna PM10 data using a cloud database API

Hariprasanth Sinnavajoumouny

Date:4/8/2024

Supervisory team

Dr Sara Zandi (NZSE)

Dr Louis Boamponsem (Auckland City Council)

Contents

[Abstract 3](#_Toc127282232)

[Intro 3](#_Toc127282233)

[Literature review 3](#_Toc127282234)

[Problem definition 3](#_Toc127282235)

[Possible solutions/Research methods 4](#_Toc127282236)

[Schedule for experiments / implementation 4](#_Toc127282237)

[Experiments and data analysis 4](#_Toc127282238)

[Discussion 5](#_Toc127282239)

[Conclusions 5](#_Toc127282240)

[References 5](#_Toc127282241)

# Abstract

This project endeavors to advance the management and analysis of environmental data by focusing on PM10 (particulate matter less than 10 microns) in Takapuna, Auckland. Utilizing a cloud database API provided by Auckland Council, the project facilitates real-time data streaming and visualization. By implementing state-of-the-art data processing and visualization tools, the system aims to provide a comprehensive dashboard for monitoring PM10 levels. The project underscores the significance of API-driven data analytics in environmental science, offering a robust tool for policymakers and researchers to make informed decisions and enhance public health responses to air quality issues.

**Keywords; PM10, API, Cloud Database, Visualization**

# Introduction

Air pollution poses significant challenges to environmental sustainability and public health worldwide. Particulate matter with a diameter of 10 micrometers or less (PM10) is a major component of air pollution and can have adverse effects on respiratory health, visibility, and overall air quality. In urban areas like Takapuna, New Zealand, understanding PM10 concentrations is essential for mitigating environmental impacts and safeguarding public health.

Air quality monitoring is a critical aspect of public health management, as airborne pollutants can have significant impacts on human health. Particulate matter (PM) is one of the primary pollutants tracked globally. PM10, specifically, refers to particulate matter with a diameter of 10 micrometers or less. Because these particles are small enough to be inhaled, they can settle in the respiratory system and cause various health issues. Epidemiological studies have linked PM10 exposure to respiratory and cardiovascular diseases, leading to increased hospitalizations, mortality rates, and a higher prevalence of chronic illnesses. Moreover, PM10 serves as an indicator of pollution from various sources, such as vehicle emissions, industrial activities, and natural processes (e.g., dust storms, wildfires).

**Takapuna: Overview and Importance**

Takapuna is a coastal suburb located on Auckland's North Shore in New Zealand. Known for its beaches, residential neighborhoods, and vibrant commercial areas, Takapuna experiences significant traffic and commercial activity, making it susceptible to air pollution. The suburb's geographic location near the ocean and hills further complicates air quality, as prevailing winds, and topography influence pollutant dispersion.

Air quality monitoring in Takapuna is vital because of its proximity to dense traffic routes and urban centers. Vehicle emissions, particularly from diesel engines, significantly contribute to PM10 levels, potentially impacting residents, and visitors. The presence of schools, parks, and commercial zones amplifies the need for clean air.

The monitoring site at Takapuna, situated at Westlake Girls High School, has been in operation since 1995. It is one of the longest-running monitoring stations in Auckland. This continuous dataset provides insights into PM10 patterns and trends over time, making it valuable for studying long-term pollution changes. Understanding these trends can guide environmental policies, inform stakeholders, and shape public awareness campaigns.

In response, many countries, including New Zealand, have established air quality monitoring networks to measure and report PM10 concentrations. These networks help set safe exposure standards and provide valuable data for public awareness, policymaking, and pollution control measures. Regular monitoring can identify pollution hotspots, detect trends over time, and highlight areas requiring urgent intervention to protect public health.

# The primary goal of this project is to develop a dynamic and responsive tool that leverages a cloud database API to analyze and visualize PM10 data in Takapuna. This initiative is part of a broader effort to enhance environmental monitoring and management practices within the region. By the end of this project, we aim to provide the Auckland Council and local stakeholders with a powerful analytical tool that not only tracks real-time PM10 levels but also integrates with existing environmental management systems to support ongoing air quality initiatives.

**Research Aim and Problem Statement**

The aim of this industry project is to delve into PM10 data in Takapuna, utilizing a cloud database API for exploration and analysis. The project seeks to address specific data analytic challenges related to PM10 monitoring, visualization, and decision-making processes in environmental management. By using the capabilities of a cloud database API, a data visualization tool will be developed to directly access PM10 data, enabling real-time in-depth exploration and analysis. This project will involve identifying or developing a customizable data visualization tool capable of leveraging the API data, allowing users to uncover meaningful insights into PM10 pollution patterns.

Key questions driving this project include:

What are the temporal patterns of PM10 concentrations in Takapuna?

How can a cloud database API be utilized to integrate and access PM10 data for analysis?

What insights can be gained from the analysis of PM10 data, and how can they inform environmental management decisions in Takapuna?

The specific goals include:

**Access and Integration:** Seamlessly connecting visualization tools with the cloud database API requires stable access, proper authentication, and data extraction protocols to ensure accurate PM10 data retrieval.

**Data Cleaning and Preparation:** PM10 data obtained through the API often requires extensive cleaning, aggregation, and normalization to ensure consistency and comparability across different timeframes and data units.

**Visualization Compatibility:** Visualization tools need to handle the data structures provided via the API, supporting flexible data manipulation. Compatibility issues can limit visual analysis, making it challenging to reveal PM10 trends, seasonal changes, and unexpected pollution spikes.

**Scalable Cloud Platform:** Leveraging a scalable cloud platform is critical for processing large datasets, ensuring secure storage, and enabling real-time analysis, particularly when data from multiple sites or pollutants are included.

**Customizable Visualization:** Finding tools with customizable visualization features can enhance data exploration by allowing stakeholders to adjust visual elements (e.g., time periods, pollutants) dynamically, thus adding value.

Therefore, identifying an API-integrated visualization tool that provides customizable features for cloud-based PM10 data analysis will empower policymakers, researchers, and the public to explore trends, pinpoint sources, and support decision-making on air quality management.

# 

Project Questions (Complex Business Needs) (300 words)

Project Questions and Complex Business Needs:

What are the spatial and temporal patterns of PM10 concentrations in Takapuna?

This question aims to understand the distribution and variation of PM10 concentrations across different locations and time periods within Takapuna. The complexity lies in the dynamic nature of air pollution, influenced by factors such as traffic patterns, industrial activities, weather conditions, and topography. Addressing this question is crucial for identifying high-risk areas and prioritizing intervention strategies to mitigate air pollution's adverse effects on public health and environmental quality.

How can a cloud database API be utilized to integrate and access PM10 data for analysis?

This question focuses on leveraging technological solutions to streamline data management and analysis processes. The complexity arises from the need to integrate data from multiple sources, ensure data consistency and quality, and facilitate real-time access to PM10 data for analysis. By exploring the capabilities of a cloud database API, this question addresses the challenge of efficiently managing large volumes of environmental data and harnessing it for informed decision-making.

What insights can be gained from the analysis of PM10 data, and how can they inform environmental management decisions in Takapuna?

This question emphasizes the practical application of data analytics in addressing environmental management challenges. The complexity lies in translating raw data into actionable insights that can guide policy formulation and intervention strategies. By uncovering patterns, trends, and correlations in PM10 data, this question aims to inform evidence-based decision-making processes, enhance environmental monitoring and regulation, and ultimately improve air quality and public health outcomes in Takapuna.

Scope:

The scope of this industry project encompasses the collection, integration, analysis, and visualization of PM10 data in Takapuna, focusing on understanding spatial and temporal patterns, leveraging cloud database API for data access, and deriving actionable insights to inform environmental management decisions. The project will utilize advanced data analytics techniques and visualization tools to address complex business needs related to air pollution monitoring and management in urban areas.

**Research Questions in a nutshell.**

To guide the exploration and ensure comprehensive analysis of Takapuna's PM10 data using the API and cloud platform, the following research questions are proposed:

**Temporal Patterns:** What patterns can be identified in Takapuna's PM10 concentrations over time (daily, monthly, or yearly), and how do seasonal variations impact PM10 levels?

**Correlation with Other Pollutants:** How do PM10 concentrations correlate with other monitored pollutants (e.g., NO2, SO2) and environmental factors (temperature, wind)?

**API Data Accessibility:** What challenges or advantages arise from using the cloud API for accessing and analyzing Takapuna's historical PM10 data?

**Customizable Visualizations:** How can customizable visualizations aid in identifying significant PM10 pollution periods, seasonal anomalies, and patterns for stakeholders?

**Decision-making Support:** How effective is a cloud API-based visualization platform with real-time data visualization in supporting policymakers, researchers, and the public in understanding PM10 pollution and making informed air quality management decisions?

Literature Review (700 words)

Takapuna PM10 Analysis

Air Quality Index (AQI) and PM10 Monitoring

The Air Quality Index (AQI) is a standardized tool to communicate air pollution levels to the public. PM10 is one of the key components of AQI due to its health risks, as emphasized by Zandi. In her thesis, Zandi mentions that PM10 concentrations contribute significantly to health risks like respiratory conditions, especially in urban areas like Takapuna​​. Thus, monitoring and analysing PM10 levels remains a priority.

Exploratory Data Analysis (EDA) of PM10 in Takapuna and Other Sites

Zandi’s thesis explores the PM10 data through a comprehensive exploratory data analysis. This includes summarizing data distribution across different monitoring sites and performing statistical comparisons. In Takapuna, for instance, PM10 levels were tracked using daily average values to reveal temporal trends, seasonal variations, and sources of pollution​​. Zandi identifies how wind conditions, traffic, and nearby industrial activities influence PM10 concentrations.

PM10 Data Visualization Recommendations

Zandi’s research demonstrates the utility of different visualizations in understanding PM10 patterns across multiple sites:

**Time Series Plots:** Daily and monthly average PM10 levels are plotted to identify temporal trends and seasonal fluctuations.

**Matrix Plots:** Correlation matrix plots visually compare different PM10 metrics and meteorological factors.

**Rose Diagrams:** Wind roses and pollution roses illustrate the direction and concentration of PM10 influenced by wind speed and direction.

**Geospatial Mapping:** Spatial distribution maps show PM10 concentration variations across sites over time.

These visualization techniques effectively highlight PM10 pollution patterns, sources, and correlations, providing stakeholders with actionable insights to improve air quality in Takapuna.

**Second Literature Review**

The research by Cao et al. (2021) wonderfully demonstrates the capabilities of Python in the realm of data visualization, notably via Matplotlib and Pyecharts. These libraries convert complex datasets into visually appealing and understandable formats, which perfectly complements the objectives of our PM10 data project in Takapuna. Intriguingly, the article also introduces the benefits of using **Dask** for processing large datasets efficiently, suggesting its potential to streamline our exploratory data analysis (EDA). The incorporation of Dask promises to enhance our handling of extensive environmental data, ensuring swift processing and real-time visualization. This advancement is particularly exciting as it promises to refine how we present and interpret air quality data, thereby aiding stakeholders in making quick, well-informed decisions.

**Third Literature Review**

In my review of "Seaborn: statistical data visualization" by Waskom (2021), I was particularly impressed by how Seaborn advances Python's data visualization capabilities by integrating tightly with Matplotlib and pandas. This synergy is crucial for effectively managing the complex data structures typical in environmental analysis. Seaborn's ability to automatically link data features to visual elements like colour and size significantly eases the creation of detailed, multidimensional visual analyses. Its facility for generating comparative multi-panel figures is especially valuable for assessing temporal and variable-specific trends in PM10 pollution within Takapuna. Utilizing Seaborn can dramatically enhance our exploratory and analytic phases, allowing us to produce more intuitive and statistically nuanced visualizations that support thorough and effective environmental policy formulation.

**Fourth Literature Review**

Exploring "Research on Python Data Visualization Technology" by Cao et al. (2021) has provided profound insights into the evolution of data visualization tools within the Python ecosystem, focusing on Matplotlib and Pyecharts. This paper details the essential role these tools play in converting complex data into visually intuitive formats, which is crucial for diverse fields including environmental science. For our PM10 pollution analysis project in Takapuna, the application of these tools is invaluable. They offer not only robustness and precision but also dynamic, interactive capabilities that can transform how we engage with and communicate our environmental data findings. By adopting these technologies, we can achieve a deeper, more accessible understanding of data trends and distributions, enhancing our ability to inform and influence environmental strategies and actions effectively.

**Fifth Literature**

In my research on PM10 pollution analysis in Takapuna, I've identified the need for advanced visualization techniques to communicate complex environmental data more effectively. This requirement is echoed in the study by Encalada-Malca et al. (2021), where they demonstrate the importance of employing sophisticated visualization tools to analyse spatio-temporal data on PM10 concentrations. Inspired by this approach, my project will incorporate several advanced visualization strategies to enhance the interpretability and accessibility of air quality data in Takapuna.

**Interactive Dashboards**

Interactive dashboards are invaluable for dynamically exploring data. They allow users to manipulate variables to see real-time changes in PM10 levels across different areas of Takapuna. This capability not only aids in immediate data interpretation but also in long-term monitoring, making it an essential tool for both public health officials and environmental scientists.

**3D Air Quality Mapping**

The integration of 3D GIS technology, as referenced by Encalada-Malca et al. (2021), provides a vivid illustration of how environmental factors influence air pollution dispersion. In Takapuna, using 3D mapping to represent PM10 data will help visualize the impact of geographical features and urban structures on pollution levels, thereby aiding in more strategic urban planning and community health initiatives.

**Real-time Visualization Feeds**

The adoption of real-time visualization feeds will enable continuous monitoring of air quality, an approach crucial for responding to acute pollution events and for assessing the effectiveness of pollution control measures. Implementing such technology ensures that data is not only up-to-date but also actionable, supporting swift decision-making in critical situations.

These techniques, informed by the methodological framework presented by Encalada-Malca et al. (2021), will significantly improve our understanding of PM10 pollution dynamics in Takapuna, facilitating more effective communication and decision-making processes related to air quality management.

Sixth Literature

Seventh Literature

**For final report**

The upcoming sections of this document provide a structured overview of the project, starting with a Literature Review in **Section 2** that explores existing research on PM10 data analysis, cloud database APIs, and real-time data visualization's role in public health. **Section2.5 project problems, gap** …

**Section 2.6:** proposing the possible solutions.

**Section 3** details the technical architecture of the data streaming pipeline from the cloud database API to the Business Intelligence (BI) tool, explaining the integration process and technologies employed. In **Section 4**, we discuss the design and functionality of the data visualization dashboard, illustrating how it aids in interpreting PM10 data effectively. **Section 4.5** – discuss about the possible solutions and it suitability in the environment Finally, **Section 5** concludes the document by summarizing the project's achievements and discussing future enhancements that could further support environmental monitoring efforts in Takapuna Health

Project planning (300 words)

* Project phases and work breakdown structure.

• Project management

* + including time-scheduling
  + Gantt chart for 16 weeks and chosen project type (e.g., AI, Big Data, Deep Learning, etc.)

• Project milestones/ project deliverables, dependencies, etc.

# Discussion with possible solutions (600 words)

Conclusion and future directions (400 words)

# Possible solutions/Research methods

## Schedule for experiments / implementation (Time and Project Management)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| W1 | Capstone project | Features | Challenges/Actions | Time |
|  |  |  |  | 1 hour |
|  |  |  |  |  |
|  |  |  |  | 2 days |
|  |  |  |  |  |
| W2 | **Project selection** | **Task Performed** | **Challenges/Actions** | **Time** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| W3 | **Selecting Aucklnd Council project 2** | **Task performed** | **Challenges/Actions** | **Time** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| W4 | **Research performed** | **Task performed** | **Challenges/Actions** | **Time** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

# Reference

May, R. M., Goebbert, K. H., Thielen, J. E., Leeman, J. R., Camron, M. D., Bruick, Z., Bruning, E. C., Manser, R. P., Arms, S. C., & Marsh, P. T. (2022). MetPy: A Meteorological Python Library for Data Analysis and Visualization: Bulletin of the American Meteorological Society. *Bulletin of the American Meteorological Society*, *103*(10), E2273–E2284. https://doi.org/10.1175/BAMS-D-21-0125.1

# Appendix 1: Capstone Project Checklist

##### Purpose

This checklist ensures that a selected workplace is adequate and appropriate for the purposes of the Capstone Project. The Checklist will be completed by the NZSE Capstone Project Leader in conjunction with the student and the Capstone Project Capstone Project Leader.

|  |  |  |
| --- | --- | --- |
| **Criteria** | **Yes/No** | **Initials** |
| Is the selected workplace suitable for the learning objectives of the Capstone Project? | Yes |  |
| Have the student and Capstone Project Supervisors received a copy of the Capstone Project Handbook? | Yes |  |
| Have the student and the Capstone Project Supervisors been fully briefed on the information in the Capstone Project Handbook and the learning and assessment requirements of the Capstone Project?   * roles and responsibilities of all parties. * requirements for liaison and contact with the NZSE Capstone Project Leader. * problem-solving process to be used in the event of any difficulty; and * learning and assessment requirements of the Capstone Project. | Yes |  |
| Have the students and the Capstone Project Supervisors discussed and agreed on the project description and deliverables? | Yes |  |
| Have the students and the Capstone Project Supervisors read and signed the Internship contract? | Yes |  |
| Have the roles and responsibilities of all parties in the Capstone Project been explained to the student? | Yes |  |
| Have the students read and signed the confidentiality agreement (**Appendix 3**)? | Yes |  |
| Have the students been inducted on the Health and Safety policies and procedure of the organisation (if applicable)? | Yes |  |
| Have the student and the Capstone Project Supervisors been advised of the problem-solving process to be used in the event of any difficulty? | Yes |  |
| Have the requirements for liaison and contact with the New Zealand School of Education Capstone Project Leader been discussed and acknowledged by the  student? | Yes |  |

##### Comments

NZSE Capstone Project Leader Name & Signature Date

# Appendix 2: Capstone Project MOU

##### Student Details

Name Hariprasanth Sinnavajoumouny

Student ID 764707774 Cohort GDDA7123C

Mobile Number 0225429096

Email 764707774@nzse.ac.nz

##### Industry Partner Details

Organisation Name Auckland Council

Physical Address

Website

##### Capstone Project Supervisor/s

Name Louis Boamponsem (External)

Mobile Number

Email louis.boamponsem@aucklandcouncil.govt.nz

Name Sara Zandi (NZSEG)

Mobile Number

Email sara@nzse.ac.nz

##### NZSE Capstone Project Leader

Name Mohammad Nourozifard

Mobile Number

Email

##### Capstone Project Description

Capstone Project Dates from 16/04/2024 to 02/08/2024

Weekly hours of work 15

List below the description of the tasks, roles and responsibilities of the student undertaking the for the Capstone Project. Else, attach the description provided by the company. (These may be added to or amended during the period of the Capstone Project. If this is the case a new contract should be signed by all parties – see section 4 of this document under **Inadequate, inappropriate or insufficient work)**

##### Capstone Project tasks (milestones), roles and responsibilities:

**Exploring Takapuna PM10 data using a cloud database API [NB---- This topic can be varied from 10 different sites across particulate matter (PM2.5), toxic gases (NO2, CO, O3, SO2), black carbon]**

This capstone project aims to delve into PM10 (particulate matter with a diameter of 10 micrometres or less) data in Takapuna, utilising a cloud database API for exploration and analysis. Application Programming Interfaces (API) play a crucial role in data analytics and technology integration. By leveraging the capabilities of a cloud database API, a data visualisation tool will be built to directly access PM10 data, enabling real-time in-depth exploration and analysis. The outcomes of this study will not only enhance our understanding of PM10 concentrations in Takapuna but also demonstrate the utility of API in environmental data analysis and decision-making processes.

This is a joint agreement between the student, the organisation and the NZSE.

##### The student agrees to:

* take responsibility for their own learning
* act ethically and responsibly at all times
* abide by organisational policy
* students read and signed the confidentiality agreement **Appendix 3**
* attend and participate in meetings with the Capstone Project Supervisors and/or NZSE Capstone Project Leader
* speak with their Capstone Project Supervisors should they have any concerns or if the Capstone Project is not meeting the learning needs of the student. If this is not an option, the student is to speak with their NZSE Capstone Project Leader; and
* take 30 minutes per day to spend time on reflective practice and completion of practice requirements.

##### The organisation agrees to:

* provide students with opportunities to meet the learning outcomes as specified in the course descriptor
* familiarise students with organisational policy, structure, accountability systems, codes of conduct and reporting systems at the commencement of the placement
* ensure the student has regular weekly professional supervision
* ensure the student is not left unsupervised or alone in a dangerous situation
* capstone Project Supervisors conduct mid-placement and closure interviews at the end of placement.
* communicate to the NZSE Capstone Project Leader, any sign of difficulty regarding the student’s performance, attendance or other aspects relevant to the success of their Capstone Project/ Capstone Project provide learning opportunities which acknowledge the student’s status as a beginning practitioner; and provide a healthy and safe workplace for the student.

##### NZSE agrees to:

* ensure the organisation is fully informed of everyone’s roles and responsibilities
* keep the NZSE Capstone Project Leader up to date with factors affecting Capstone Project arrangements meet (physically if applicable) with the Capstone Project Supervisors at least three times during the Capstone Project and use the checkpoint checklist in **Appendix 4**; and
* complete the final assessment of the student’s work and award a grade.

##### Student

Signature A close-up of a black text

Description automatically generated Date

##### Capstone Project Supervisors (Industrial supervisor)

Signature Date

##### Capstone Project Supervisors

Signature Date

##### NZSE Capstone Project Leader

Signature Date

This is a joint agreement between the student, the business organisation, and the New Zealand School of Education.

# Appendix 3: Confidentiality

As a student you must always be aware of the confidentiality of information gained during the course of your duties. It is expected that you understand the importance of treating information in a discreet and confidential manner and your attention is drawn to the following:

1. Written records and correspondence must be kept securely at all times, including when not in use
2. Business organisational documentation must not be submitted as appendices to the Final Assessment Report unless prior agreement from the organisation is received
3. Information regarding the business must not be disclosed either orally or in writing to unauthorised persons. It is particularly important that the authenticity of phone, e-mail and text enquirers should be checked
4. Conversation relating to confidential matters should not take place in situations where they may be heard by passers-by i.e., in corridors, reception areas etc.
5. The same confidentiality must also be preserved in dealing with matters relating to departmental personnel.

##### I have read and accept the terms of the above on confidentiality:

Signed: A close-up of a black text

Description automatically generated

Printed Name: HARIPRASANTH SINNVAJOUMOUNY

A logo with text on it

Description automatically generated

API Usage for research - Agreement

This API Usage Agreement ("Agreement") is made and entered into on 19th April 2024, by and between:

Dr Louis Boamponsem /Auckland Council], ("API Provider"), and

Hariprasanth Sinnavajoumouny., ("Student").

Agreement:

**License and Usage:**

API Provider grants Student permission to use the API for research purposes only.

Student agrees to use the API responsibly and in compliance with any provided guidelines.

**Confidentiality:**

Student agrees to keep the API URL, key, and access credentials confidential.

Student will not share the API URL, key, or access credentials with anyone.

**Termination:**

API Provider may terminate this Agreement at any time with notice to Student.

Upon termination, Student will stop using the API immediately.

**Liability:**

Student is responsible for any misuse of the API and any resulting consequences.

API Provider is not liable for any damages from Student's use of the API.

**Signatures:**

Both parties agree to the terms of this Agreement.

API Provider:

Dr Louis Boamponsem

A line drawing of a person shaking hands

Description automatically generated

19/4/2024

Student:

Hariprasanth Sinnavajoumouny

A close-up of a black text

Description automatically generated

22/4/2024

135 Albert Street | Private Bag 92300, Auckland 1142 | **aucklandcouncil.govt.nz** | **Ph** +64 272025907