

Executive Summary

This project presents the design and implementation of a Real-Time Air Quality Index (AQI) Monitoring and Prediction System developed through effective teamwork and clear role distribution. The objective of the project was to build a reliable, automated system capable of collecting live air quality data, performing short-term predictions, and presenting insights through an interactive dashboard.

The team collaboratively designed an end-to-end data pipeline where live air pollution and weather data are fetched from the OpenWeatherMap API using Python-based automation. The backend system is deployed on a cloud environment, ensuring continuous execution without manual intervention. Data is structured and stored in Google Sheets, enabling seamless live connectivity with the visualization layer.

A machine learning approach was incorporated using Linear Regression to predict short-term PM2.5 levels based on historical pollution data and time-based trends. This model was selected to balance simplicity, interpretability, and effectiveness within the project timeline. Additionally, a rule-based logic was implemented to generate meaningful health advisories from raw AQI values.

The visualization component was developed using Power BI, where interactive dashboards display key pollution metrics, trend analysis, predictions, and geographical insights. The system is automated through scheduled execution, ensuring timely and consistent data updates.

This project demonstrates strong teamwork, practical application of data analytics, machine learning, cloud deployment, and business intelligence tools, and successfully meets academic and technical objectives within defined constraints.