

AirAware: Smart Air Quality Prediction System

AI-driven air quality prediction and environmental insights

Team-1:

Rajalakshmi

Rahul Varma

N Sai Sreya

Divija Nandana

Lokesh



The Air Quality Crisis

Air pollution continues to pose significant health risks globally, affecting millions of people daily. Traditional air quality monitoring systems provide only static, retrospective data, limiting their utility for proactive decision-making.

Communities and individuals need predictive capabilities to make informed choices about outdoor activities, health precautions, and environmental interventions.



Project Objectives

Predictive Analytics

Leverage machine learning algorithms to forecast future air quality values with high accuracy, enabling proactive responses.

Visual Intelligence

Create an interactive dashboard that transforms complex environmental data into accessible, actionable insights.

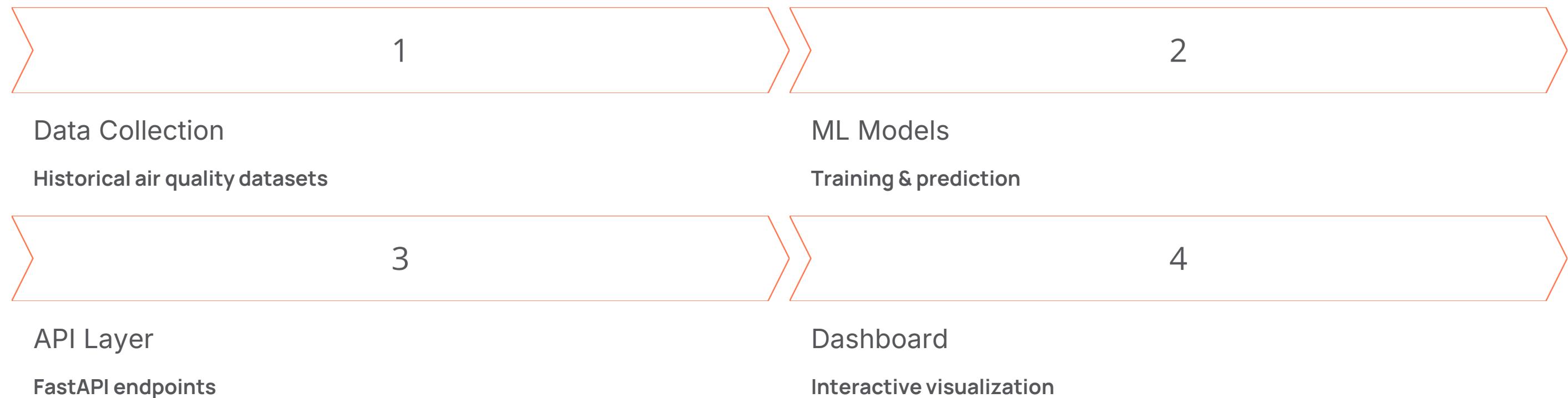
Explainable AI

Provide transparent, user-friendly explanations of model predictions and air quality metrics for diverse audiences.



System Architecture

AirAware integrates data collection, machine learning, backend services, and frontend visualization into a cohesive prediction platform.





DATA

Data Collection & Preparation

Dataset Composition

- PM2.5, PM10 particulate matter
- NO₂, SO₂, CO, O₃ gas pollutants
- Temporal and location metadata
- Historical AQI measurements

Preprocessing Pipeline

Raw environmental data undergoes rigorous cleaning to remove outliers and handle missing values. Features are normalized using standard scaling techniques to ensure optimal model performance.

Target variable: Future PM2.5 concentration, selected for its health significance and predictive value.

Machine Learning Pipeline

01

Feature Engineering

Created temporal features (hour, day, month), lag variables, and rolling averages to capture pollution patterns.

03

Model Training

Implemented multiple algorithms with hyperparameter tuning using cross-validation techniques.

02

Train-Test Split

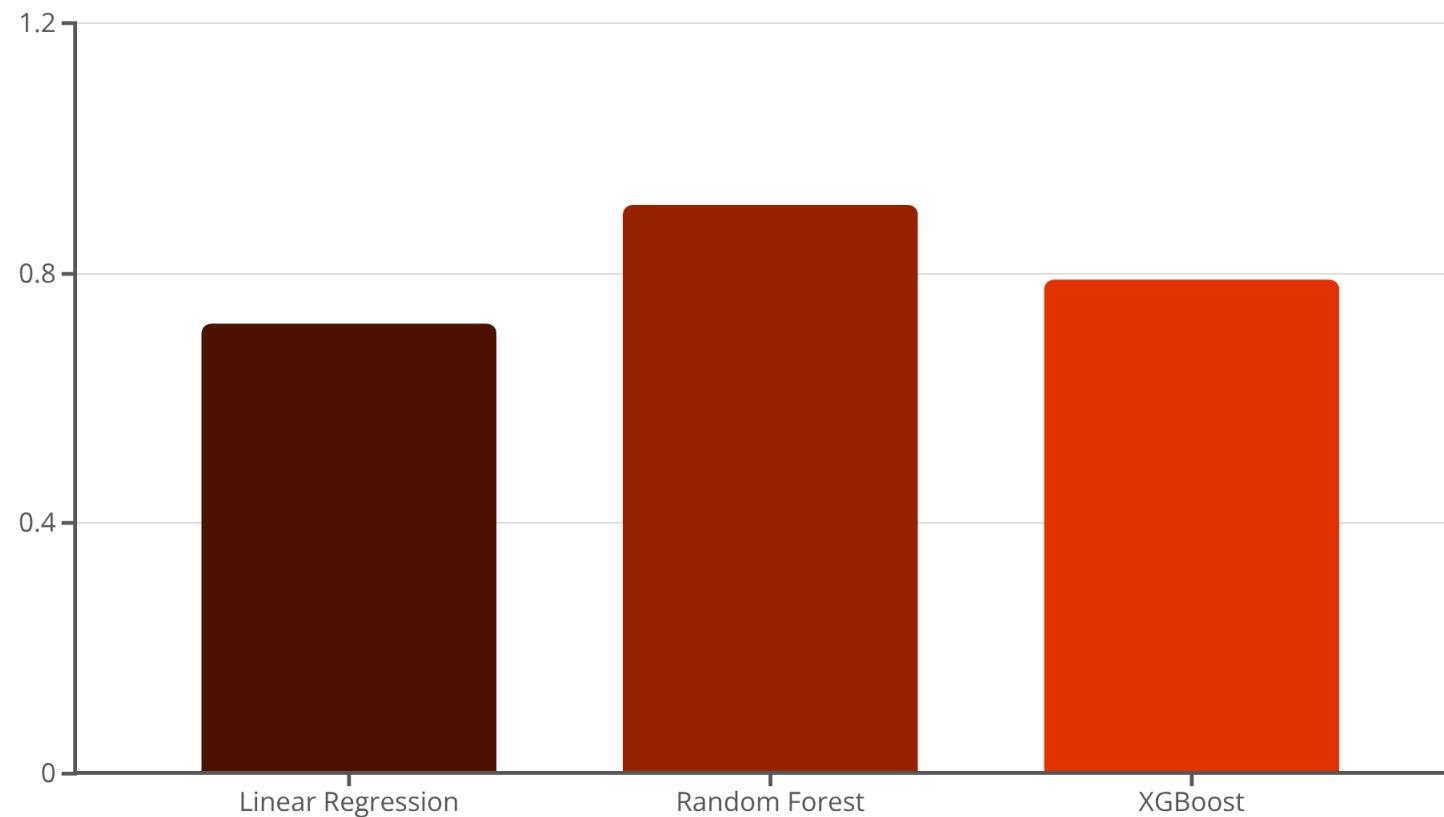
Divided dataset chronologically (80/20) to preserve temporal relationships and prevent data leakage.

04

Evaluation & Selection

Compared models using R^2 , MSE, and RMSE metrics to identify optimal predictor.

Model Performance Comparison



Model Selection

Random Forest emerged as the best-performing model, achieving 91% accuracy in predicting future PM2.5 levels.

This ensemble method effectively captures non-linear relationships and complex interactions between pollutants, weather patterns, and temporal factors.

Technology Stack

Each technology component was strategically selected to optimize performance, scalability, and user experience.

React Frontend

Powers the interactive dashboard with responsive components, real-time chart updates, and seamless user interactions.

FastAPI Backend

Provides high-performance REST APIs for predictions, metrics retrieval, and data processing with automatic documentation.

Supabase

Stores user login credentials.

Scikit-learn ML

Enables model training, evaluation, and persistence with robust implementations of Random Forest and regression algorithms.

LLM Integration

Powers the intelligent chatbot using OpenAI/Gemini APIs for natural language understanding and contextual responses.

FEATURES

Dashboard & Visualization Capabilities

Prediction Interface

- Real-time PM2.5 forecasting
- AQI category classification
- Interactive time-series charts
- Actual vs. predicted comparisons

AI Chatbot Assistant

- Pollutant explanations and health impacts
- AQI interpretation and recommendations
- Project methodology Q&A
- Contextual environmental insights



Team Contributions

Sreya – Frontend Development

Designed and implemented the complete React dashboard, creating interactive charts, prediction cards, and heatmap visualizations with emphasis on responsive design and user experience.

Rajalakshmi – Backend Development

Developed FastAPI endpoints for predictions, metrics, and data retrieval. Integrated ML models with frontend and ensured efficient API design and data flow.

Rahul – Machine Learning

Built the complete ML pipeline from feature engineering to model deployment. Trained and evaluated multiple algorithms, selecting Random Forest as the optimal predictor.

Lokesh – Backend Development

Assisted in API development and testing. Managed data handling logic and ensured seamless integration between machine learning models and backend services.

Nandana – Database Management

Designed Supabase schema for air quality and user credentials storage. Optimized data retrieval queries for analytics and ensured database scalability for future expansion.

Outputs

Air Aware

Welcome back

Sign in to continue to AirAware

 you@example.com

 Password

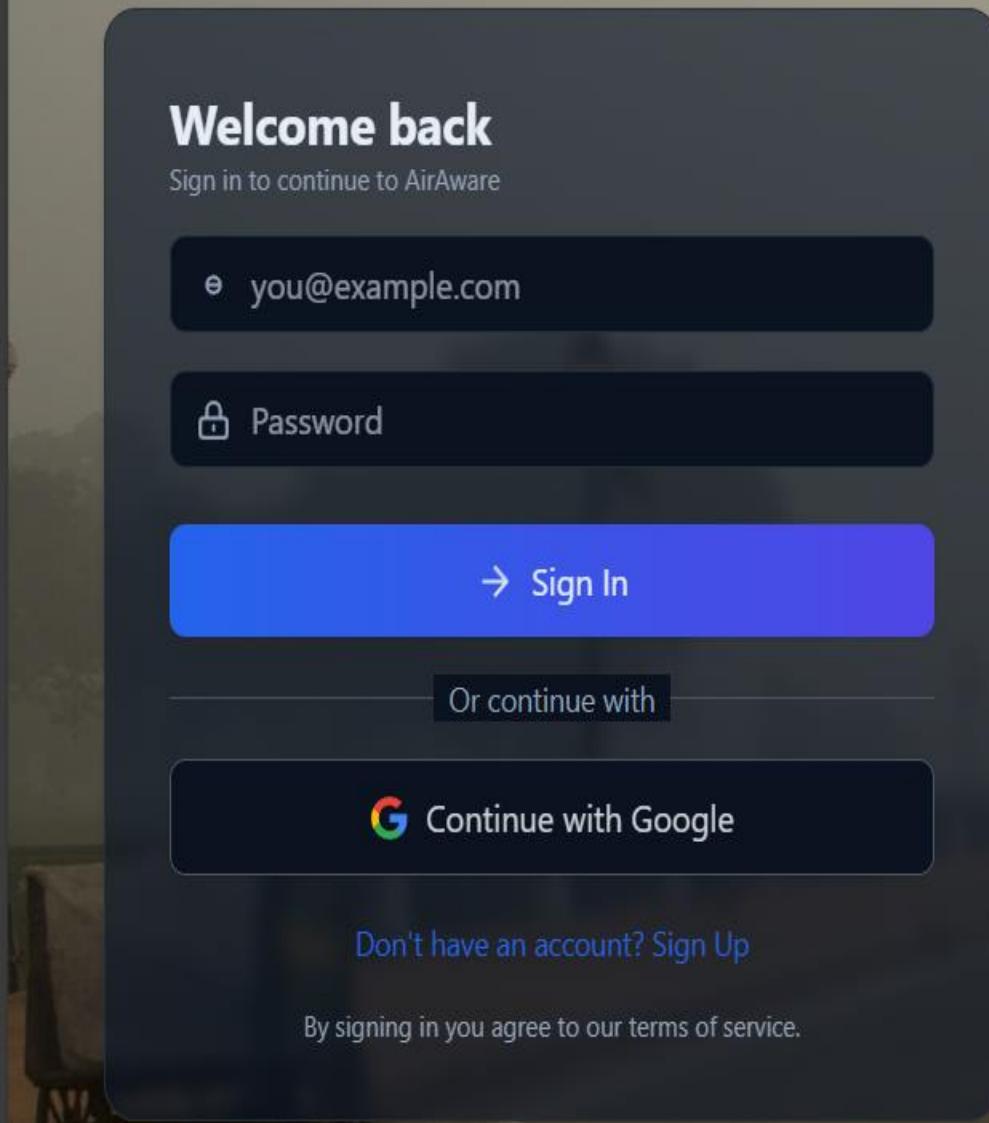
→ Sign In

Or continue with

 Continue with Google

[Don't have an account? Sign Up](#)

By signing in you agree to our terms of service.





About AirAware

AirAware is a lightweight air-quality dashboard focused on making local AQI and pollutant information understandable and actionable. It combines public data, weather context and simple ML predictions to help users monitor conditions, download reports, and receive alerts when air quality is poor.

How it works

AirAware aggregates real-time readings from public APIs and community sensors, normalizes pollutant concentrations, computes an AQI value, and surfaces trend-based predictions. The chatbot provides heuristic guidance (not medical advice) and the report export creates a printable snapshot of the current reading.

Data sources

- World Air Quality Index (WAQI) feeds
- OpenWeatherMap (for weather and contextual signals)
- Optional community sensor networks where available

Privacy & limitations

This demo stores only minimal, local data in your browser (a lightweight profile and theme preference). No user data is transmitted to any third-party by this app unless you explicitly use OAuth sign-in. Readings are fetched from public APIs and may have gaps or delays.

Interpretation & safety

AQI categories indicate population-level risk. Sensitive groups (children, elderly, people with respiratory conditions) should take extra precautions when AQI is elevated. The recommendations shown in the app are heuristic and should not replace professional medical advice.

Contributing & source

This project is open-source and intended as a demo. Contributions are welcome — see the repository README for how to run locally, add data sources, or improve models and UI.

For development notes, see the project README. This is a demonstration tool — use it as a guide, not a certified monitoring system.



Current Air Quality

Last updated: 2 minutes ago • North Delhi

430

AQI Index

0 Good 50 100 150 200 300+ Hazardous

Hazardous

What this means

Health warnings of emergency conditions. Sensitive groups should consider reducing prolonged outdoor activities.

Quick Stats

Current AQI **430**

Category **Hazardous**

Last Updated 12:08:56 am

7-Day Avg **430 AQI**

Real-time Updates

Manual refresh only

Weather

Temp: 14.8°C

Wind: 3.1 km/h

[Download AQI Report](#)



Pollutant Breakdown

PM2.5 425 µg/m³ ↗ 100%

PM10 430 µg/m³ ↗ 100%

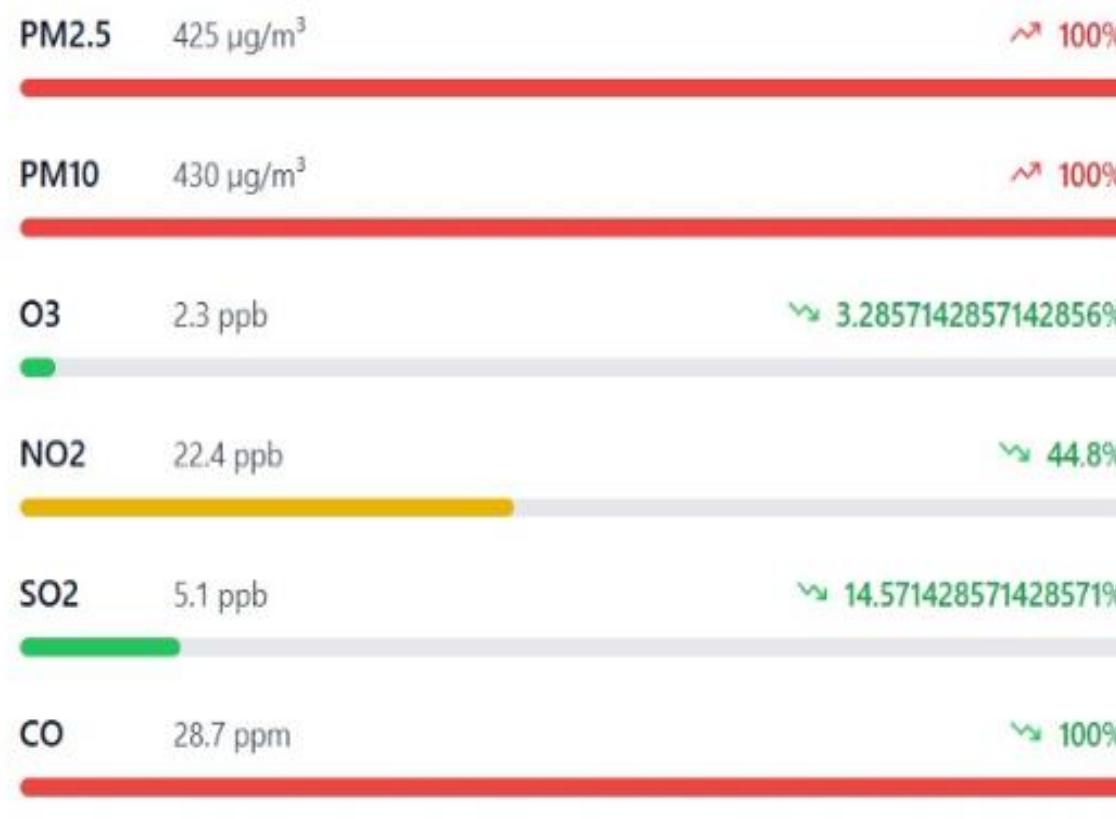
Health Recommendations

⚠ Limit outdoor activities

⚠ Everyone should avoid prolonged outdoor exertion

⚠ Keep windows closed and use air purifiers

Pollutant Breakdown



Legend: PM = Particulate Matter, O₃ = Ozone, NO₂ = Nitrogen Dioxide, SO₂ = Sulfur Dioxide, CO = Carbon Monoxide

Health Recommendations

- ⚠ Limit outdoor activities
- ⚠ Everyone should avoid prolonged outdoor exertion
- ⚠ Keep windows closed and use air purifiers

Sensitive Groups Include:

Children Elderly Asthma patients Heart disease Active outdoors

Activity Guidance:

- | | |
|--------------------|-------------------|
| 🚶 Walking: | ✗ Not recommended |
| 🏃 Running: | ✗ Not recommended |
| 🚴 Cycling: | ✗ Avoid |
| 🏠 Indoor workouts: | ✓ Safe |

North Delhi

Industrial and densely populated area with major commercial hubs.

Key Characteristics:

High vehicle density

Industrial emissions

Residential congestion

Construction activity





ML Model Integration Demo

Test the Gradient Boosting AQI Predictor

● API Status

Input Pollutants

PM2.5

45.2

$\mu\text{g}/\text{m}^3$

PM10

78.5

$\mu\text{g}/\text{m}^3$

NO2

38.7

$\mu\text{g}/\text{m}^3$

SO2

12.3

$\mu\text{g}/\text{m}^3$

CO

1.2

ppm

O3

42.1

$\mu\text{g}/\text{m}^3$

NH3

180

ppb

🎲 Randomize

↗ Predict AQI with ML Model

ML Predicted AQI

79

Moderate

ⓘ Air quality is acceptable. However, there may be a risk for some people

MODEL DETAILS

Algorithm: Gradient Boosting

Test R²: 0.9978

Features: 7 pollutants

Training samples: 151



Live ML Integration: This component connects to the Flask API (localhost:5000) running your trained Gradient Boosting model. Predictions are computed in real-time based on pollutant inputs.

 **Live ML Integration:** This component connects to the Flask API (localhost:5000) running your trained Gradient Boosting model. Predictions are computed in real-time based on pollutant inputs.

7-Day Forecast

Predicted air quality trends



█ Good (0-50) █ Moderate (51-100) █ Unhealthy (101-150)

 Updated hourly





Air Assistant

readings; check the dashboard for details.

tips

Avoid prolonged outdoor exertion when AQI is high; use masks if needed.

|Ask about AQI, tips, or fore

Send

AirAware AQI Report

Location: North Delhi

AQI: 430 (Hazardous)

Time: 30/12/2025, 12:08:56 am

Pollutants

- **PM2.5:** 425 $\mu\text{g}/\text{m}^3$
- **PM10:** 430 $\mu\text{g}/\text{m}^3$
- **O₃:** 2.3 ppb
- **NO₂:** 22.4 ppb
- **SO₂:** 5.1 ppb
- **CO:** 28.7 ppm

**SAI SREYA**

saisreya9999@gmail.com

Last seen: just now

[Logout](#)

Account

SAI SREYA

saisreya9999@gmail.com

Notification Threshold

600 AQI

AQI Snapshot

--

PM2.5: -- • PM10: --

Saved Locations

- North Delhi

Preferences

Real-time updates



Alert threshold (AQI)

600

[Save](#)

Actions

[Download AQI Report](#)[Clear Saved](#)

Recent activity

Viewed AQI for North Delhi

Downloaded AQI report