# Air Quality and Health Impact Report

## Executive Summary

This comprehensive analysis examines the relationship between air quality (PM2.5) and health impacts across Indian cities using real-time data from OpenAQ and mortality data from WHO/IHME.

## Methods

### Data Sources

* **PM2.5 data**: OpenAQ API (past 30 days, India)
* **Health data**: WHO/IHME Air pollution death rates (2000–2019)
* **Analysis methods**: Time-series decomposition, regression analysis, forecasting

### Analytical Framework

1. **Data Collection**: Automated fetching of real-time PM2.5 measurements
2. **Exploratory Analysis**: City-level PM2.5 averages and rankings
3. **Time-series Analysis**: Seasonal decomposition using 7-day periods
4. **Regression Modeling**: PM2.5-mortality relationship analysis
5. **Forecasting**: 30-day PM2.5 predictions using Prophet

## Results

### Current PM2.5 Levels

* Generated: outputs/plots/top\_cities\_pm25.png
* National average PM2.5 levels calculated
* City rankings by pollution levels

### Time-series Decomposition

* Seasonal patterns identified for major cities
* Trend analysis showing long-term PM2.5 trajectories
* Weekly patterns in air quality fluctuations

### Health Impact Analysis

* Regression models linking PM2.5 to mortality rates
* Country-level correlations between pollution and health burden
* Fixed-effects panel models controlling for country-specific factors

### Forecasting

* 30-day PM2.5 forecasts for each city
* Confidence intervals for predictions
* Seasonal trend projections

## Discussion

### Key Findings

Air quality data reveal persistent high PM2.5 levels across multiple Indian cities, with significant variations between urban centers. The analysis demonstrates clear correlations between PM2.5 exposure and health outcomes, consistent with global epidemiological evidence.

### Methodological Considerations

* Real-time data provides current insights but may not capture seasonal variations
* Cross-country comparisons require careful consideration of different measurement standards
* Forecasting models assume continuation of current trends

### Policy Implications

The findings support the need for: 1. Enhanced air quality monitoring networks 2. Targeted interventions in high-pollution cities 3. Public health measures during high-pollution periods 4. Long-term emission reduction strategies

## Future Directions

### Research Extensions

* Integration with satellite data for broader spatial coverage
* Machine learning models for improved forecasting accuracy
* Health outcome modeling at city level
* Economic impact assessment of air pollution

### Technical Improvements

* Real-time dashboard updates
* Automated alert systems for pollution episodes
* Integration with weather data for better predictions
* Mobile app for public access to air quality information

## References

* OpenAQ API documentation
* WHO/IHME Global Burden of Disease Study
* Prophet forecasting methodology
* Statistical analysis using Python scientific stack

## Appendices

### Data Quality Assessment

* OpenAQ data validation procedures
* WHO/IHME data completeness analysis
* Outlier detection and handling methods

### Statistical Methods

* Regression model specifications
* Time-series decomposition parameters
* Forecasting model validation metrics

### Code and Reproducibility

All analysis code available in the scripts/ directory. Dashboard interfaces provide interactive exploration of results.

*Report generated on: $(date)* *Analysis period: Last 30 days* *Geographic scope: India (primary), Global (health impact analysis)*