Air Quality(AQI) & Health Impact Analysis Using Power BI

Power BI Analytics Report

Project Report

Submitted By:

Pankaj Kumar A

Course:

Data Pre-Processing and Business Analytics

Table of Contents

1. Introduction

Air pollution is one of the most significant environmental and public health challenges in India. Rapid urbanization, industrial development, increased vehicular emissions, and seasonal factors have resulted in frequent deterioration of air quality across major cities. Poor air quality not only affects the environment but also has serious short-term and long-term health consequences.

This project focuses on analyzing **Air Quality Index (AQI) data** across Indian cities and transforming large volumes of raw data into meaningful insights using **Power BI**. The objective is to understand pollution patterns, compare cities, identify dominant pollutants, assess health risks, and present the findings through an interactive and user-friendly dashboard.

---

2. Project Objectives

The key objectives of this project are:

- To analyze overall air quality trends across multiple years
- To identify the most polluted and cleanest cities
- To understand the distribution of AQI categories
- To analyze dominant pollutants affecting air quality
- To assess health risk exposure using derived indicators
- To study daily AQI variations using time-based analysis
- To design an interactive Power BI dashboard with slicers and navigation buttons

---

3. Dataset Description

The dataset used in this project contains approximately **4,00,000+ records** of air quality observations collected across multiple Indian cities.

Key attributes include:

- City
- State
- AQI (Air Quality Index)
- AQI Status (Good, Satisfactory, Moderate, Poor, Severe)
- Prominent Pollutant
- Meteorological attributes (temperature, humidity, wind speed, etc.)
- Time attributes (Year, Month, Day – derived where required)

The dataset spans several years, allowing both **long-term trend analysis** and **short-term daily analysis**.

---

4. Data Preparation and Modeling Methodology

4.1 Data Cleaning

- Removed null and inconsistent records
- Standardized city and state names
- Ensured AQI and numerical fields were properly formatted

4.2 Feature Engineering

- Created date hierarchy (Year, Month, Day) for time-based analysis
- Categorized AQI values into standard AQI status levels
- Created calculated measures for averages, counts, and rankings

4.3 Health Impact Derivation

Since the dataset did not contain direct health impact data, a **Health Impact Score** was derived based on AQI ranges. Higher AQI values were mapped to higher health risk scores, representing increased potential health impact.

This approach reflects real-world analytical practice where health risk is inferred from pollution severity.

---

5. Dashboard Structure and Design Approach

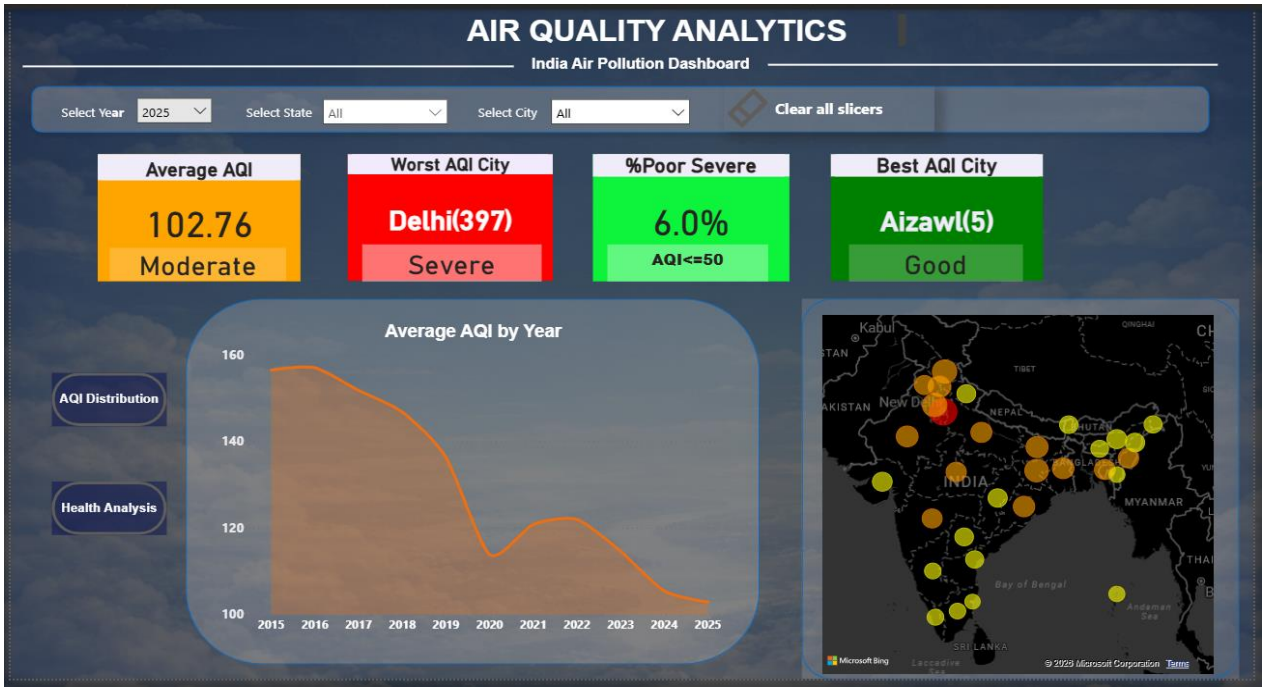
The dashboard is divided into **three structured pages**, each focusing on a specific analytical layer:

Page	Focus
Page 1	Overall air quality summary and trends
Page 2	City-level AQI comparison and distribution
Page 3	Health impact, pollutant dominance and daily AQI trends

Navigation buttons were implemented to allow seamless movement between pages, creating a clean and app-like user experience.

---

6. Page 1: Air Quality Overview



Purpose

This page provides a **high-level summary** of air quality conditions across India.

Key Visuals

- KPI Cards:

1)Average AQI

Represents the **mean Air Quality Index** across the selected filters (Year, State, City).

Purpose:

- Provides a quick understanding of **overall pollution severity**
- Serves as a benchmark indicator

Interpretation:

- Lower value → Better air quality
- Higher value → Poorer air quality

2)Best AQI City

The city with the **lowest average AQI** within the selected filters.

Purpose:

- Highlights cities with the **cleanest air**
- Encourages positive benchmarking

3)Worst AQI City

The city with the **highest average AQI** within the selected filters.

Purpose:

- Identifies pollution hotspots
- Helps prioritize regulatory or health interventions

Interpretation:

- Higher AQI values indicate severe air quality issues

4)Percentage of Poor/Severe AQI

Shows the **percentage of observations** where AQI falls under **Poor or Severe categories**.

Purpose:

- Quantifies the **extent of critical air quality**
- Useful for risk communication

Yearly AQI Trend (Line Chart)

The area chart shows the **year-wise variation in average AQI levels**. It helps identify long-term trends, such as improvement or deterioration in air quality over time. The shaded area highlights the magnitude of pollution levels, while slicers allow analysis by specific states or cities.

City-level India AQI Map (Bubble Map)

The bubble map displays **city-wise air quality across India**. Each bubble represents a city, where **bubble size and color indicate AQI severity**. This visualization helps quickly identify pollution hotspots and compare air quality across regions using interactive filters.

Methodology

- Yearly AQI trends are calculated using average AQI values
- The map visual highlights pollution hotspots using bubble size and color
- Global slicers (Year, State, City) enable interactive filtering

This page answers the question:
“What is the overall air quality situation?”

7. Page 2: City-Level AQI Analysis and Distribution

Purpose

This page focuses on comparing cities and understanding how air quality is distributed across different AQI categories.

7.1 Top 10 Worst Polluted Cities

- Visual: Horizontal Bar Chart
- Metric: Average AQI by City
- Sorted in descending order

This visualization highlights cities experiencing **consistently high pollution levels**.

7.2 Top 10 Cleanest Cities

- Visual: Horizontal Bar Chart
- Metric: Average AQI by City
- Sorted in ascending order

This provides a positive contrast by highlighting cities with **better air quality**.

7.3 AQI Status Distribution

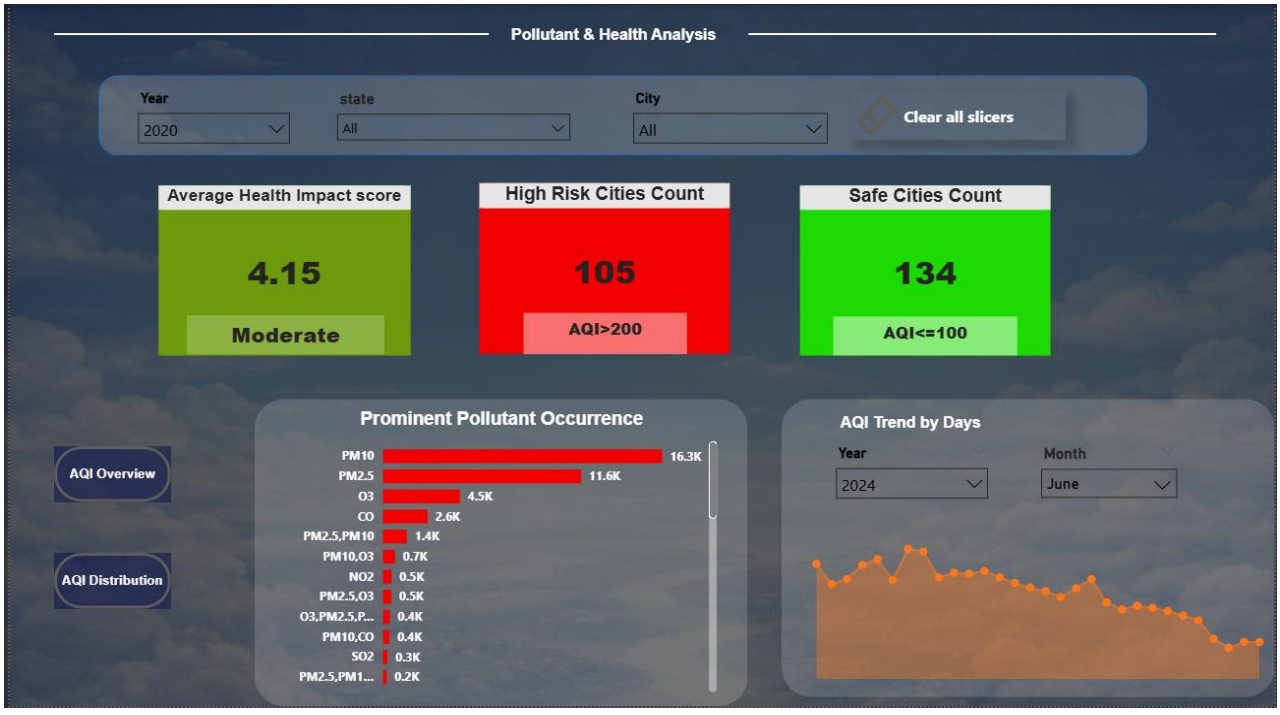
- Visual: Donut / Pie Chart
- Categories:
  - Good
  - Satisfactory
  - Moderate
  - Poor
  - Very Poor
  - Severe

This visualization shows how frequently different AQI categories occur, helping users understand overall air quality distribution.

This page answers:
“Which cities are most polluted, which are cleanest, and how is AQI distributed?”

---

8. Page 3: Health Impact, Pollutants and Daily AQI Trends



This page focuses on **health exposure, pollutant dominance, and short-term AQI behavior**.

8.1 Health-Focused KPI Row

- **Average Health Impact Score**

  A **derived metric** representing potential health risk based on AQI ranges.

Purpose:

- Translates pollution data into **health relevance**
- Bridges environmental and public health analysis

Interpretation:

- Lower score → Lower health risk
- Higher score → Higher potential health impact

- **High-Risk Cities Count (AQI > 200)**

Counts the number of cities where **AQI > 200** (Poor or Severe).

Purpose:

- Identifies how many cities face **serious health threats**
- Useful for regional risk assessmen

- **Safe Cities Count (AQI ≤ 100)**

Counts cities where **AQI ≤ 100** (Good or Satisfactory).

Purpose:

- Highlights relatively safe urban environments
- Acts as a positive performance indicator

These KPIs summarize pollution-related health risk at a glance.

8.2 Prominent Pollutants Analysis

- Visual: Bar Chart
- Dimension: Prominent Pollutant
- Measure: Count of occurrences

This analysis identifies pollutants such as **PM2.5 and PM10** that most frequently dominate air quality.

8.3 Daily AQI Trend Analysis

- Visual: Area Chart
- X-axis: Day
- Y-axis: Average AQI
- Separate slicers for:
    - Year
    - Month

This allows detailed analysis of **daily AQI fluctuations**, revealing short-term spikes caused by environmental or human factors.

This page answers:
"How does AQI vary daily and what are the health implications?".

---

10. Key Insights and Findings

- PM2.5 emerges as the most dominant pollutant across cities
- A small group of cities contributes disproportionately to high AQI levels
- Moderate to Poor AQI conditions occur frequently
- Higher AQI values correlate with increased health risk
- Daily AQI trends show noticeable short-term fluctuations

---

11. Tools and Technologies Used

- Power BI Desktop
- Power Query
- DAX (Data Analysis Expressions)
- CSV / Excel data sources

---

13. Conclusion

This project demonstrates how large-scale air quality data can be effectively analyzed and visualized using Power BI. By combining AQI trends, city comparisons, pollutant analysis, and health-focused indicators, the dashboard provides a comprehensive understanding of air pollution patterns and risks.

The interactive design enables both high-level monitoring and detailed exploration, making the dashboard suitable for analytical and decision-support purposes.
