TITLE:

AIR QUALITY ANALYSIS OF INDIAN CITIES (2010–2023)
USING SQL, TABLEAU AND EXCEL

PRESENTED BY:

PUNEET KUMAR (IOT-B) : 2202901550056
GIRISH KUMAR (CSDS-A) : 2202901540064

DATE: 26/07/2025

OBJECTIVE:

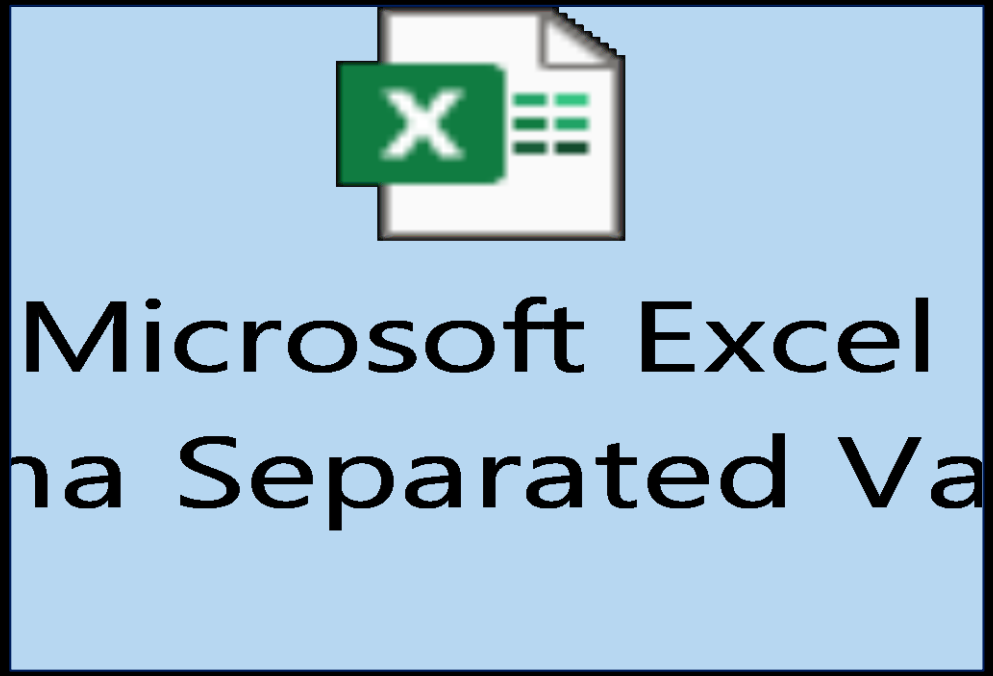
- Analyze air quality data across major Indian cities.
- Calculate overall pollution using AQI (Air Quality Index).
- Identify key pollutants and trends.
- Visualize and communicate insights using Tableau dashboards.

Source:

Kaggle - *Time-Series Air Quality Data of India (2010–2023)*

- **Total Files:** 100+ CSVs from CPCB sensors
- **Features:**
- Pollutants: PM2.5, PM10, NO, NO_2, SO_2, CO, O_3, Benzene, etc.
- Weather data: Temperature, RH, WS, BP
- Columns: 25 per file
- Time Range: 2010 to 2023
- Cities Covered: Delhi, Kolkata, Noida, Mumbai, Channai etc.

Click Excel icon to See Dataset



TOOLS & TECHNOLOGIES USED:

- **SQL :** For structured queries on cleaned data
- **Tableau:** Final data visualization & dashboards
- **Excel :** for final Cleaning and understanding data
- **Python:** Data wrangling and AQI computation

Data Cleaning & Preprocessing:

- Renamed inconsistent column headers
- Removed rows with all pollutants missing
- Handled missing values via imputation/dropping
- Converted date columns to datetime format
- Added station_id to every dataset for identification
- Data Merging In a File

Approach:

- Used pollutant-specific AQI breakpoints (CPCB formula)
- Calculated AQI for PM2.5, PM10, NO_2, SO_2, CO, O_3
- Final AQI = maximum of all sub-AQIs per row

AQI Categorization:

AQI Ranges (CPCB Standard):

AQI	Category
0–50	Good
51–100	Satisfactory
101–200	Moderate
201–300	Poor
301–400	Very Poor
401–500	Severe

Questions:

Trend Analysis

- **How has the PM2.5 level changed over time (2010–2023) in major cities like Delhi, Kolkata, Mumbai, etc.?**
- **Which year recorded the worst average AQI in India, and which pollutant contributed most?**
- **What is the seasonal trend of PM10, NO2, and SO2 across all stations?** (use line chart or area plot)

Location-Based Analysis

- **Which city shows the highest air pollution levels?**

Pollutant-wise Analysis

- **Which pollutant (PM2.5, NO2, CO, etc.) contributes the most to overall pollution in various cities?**
- **What is the correlation between PM2.5 and other pollutants like CO, NOx, or Benzene?**
  (scatter plot with trendline)

Categorical AQI Distribution

- **What is the percentage of days falling under each AQI category (Good, Moderate, Poor, etc.) per city?**
- **How many "Hazardous" AQI days occurred in Delhi in the last 5 years?**

Impact and External Factors

- **Does rainfall (RF) significantly impact PM2.5 levels in monsoon vs. non-monsoon seasons?**
- **Is there any visible impact of temperature or humidity on AQI levels?**

TABLEAU DASHBOARD:

- AQI Trends over Time (City-wise)
- Monthly & Yearly AQI comparison
- City-wise AQI Category heatmap
- Top Polluted Cities & Time Periods
- Correlation between AQI & Weather
- Interactive Filters:
- City | Year | Month | Pollutant Type

Tableau Public - Book3

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Tableau Desktop Public Edition Buy Tableau

Show Me

Dashboard Layout

Default
Phone
Device Preview

Size
Fit to height: width: 1600

Range
Minimum size
Width 1600 px Height 460 px
Maximum size
Width 1600 px Height 750 px

Sheet 11

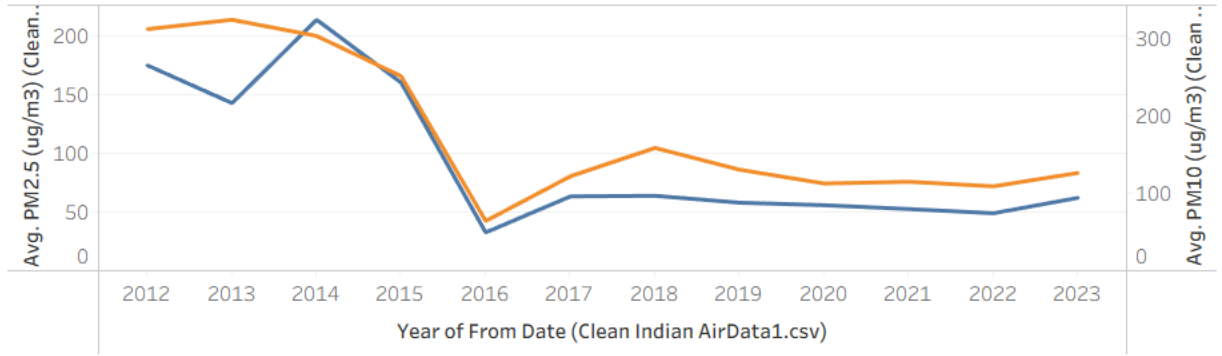
Objects
Extension
Pulse Metric
Image
Blank
Workflow
Web Page

Tiled Floating

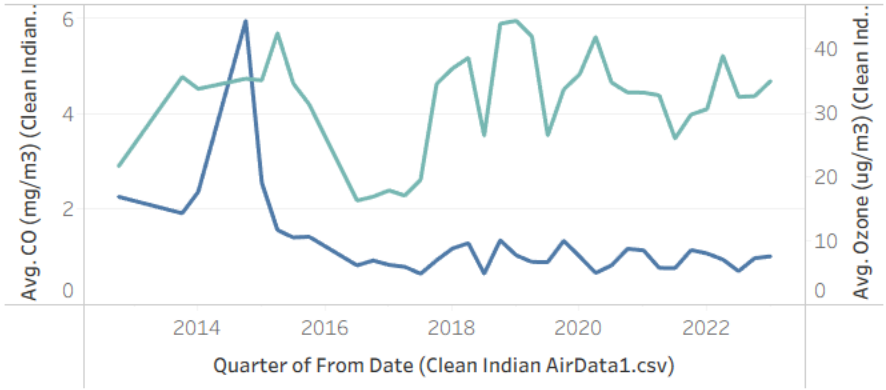
Show dashboard title

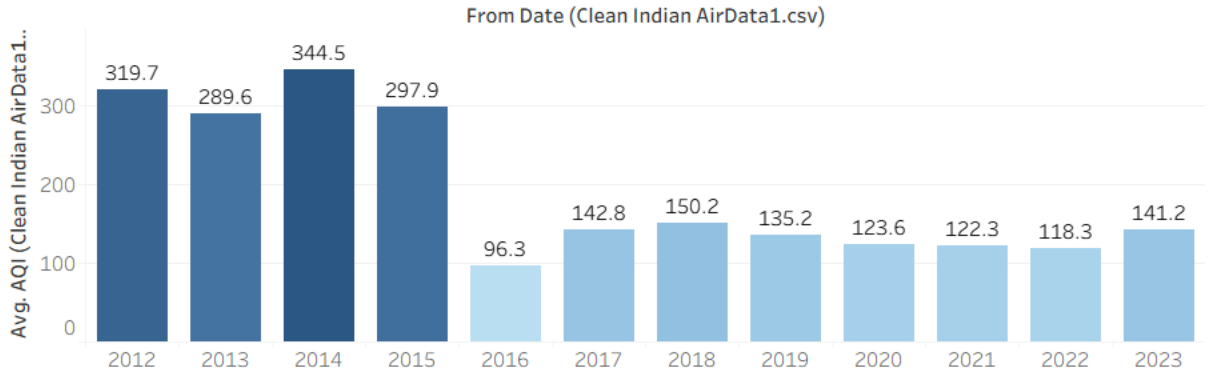
Air Quality Analysis of Indian States (2010–2023)

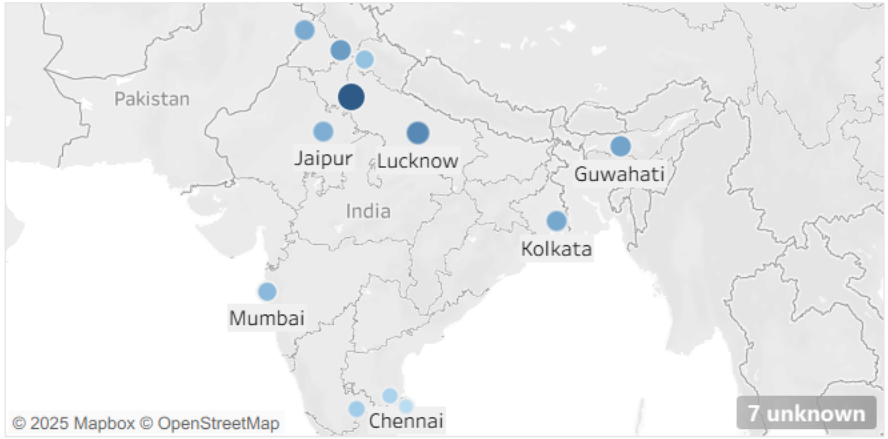
Average PM2.5 and PM10 Distribution By Year



Avg CO & Ozone distribution by year







The trend of average of CO (mg/m3) for From Month. Color shows details about cities. The marks are labeled by Station id.

Data Source | Sheet 1 | Sheet 2 | Sheet 3 | Sheet 4 | Sheet 5 | Sheet 6 | Sheet 7 | Sheet 8 | Sheet 9 | Dashboard 2 | Dashboard 3 | Sheet 10 | Dashboard 4 | Story 1 | Sheet 11

34°C Haze

Search

ENG IN

11:57 25-07-2025

Tableau Public - Book3

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Tableau Desktop Public Edition Buy Tableau

Fit Width Show Me

Dashboard Layout

Default
Phone
Device Preview

Size
Fit to height: width: 1620

Sheets
- Sheet 5
- Sheet 6
- Sheet 7
- Sheet 8
- Sheet 9
- Sheet 10
- Sheet 11

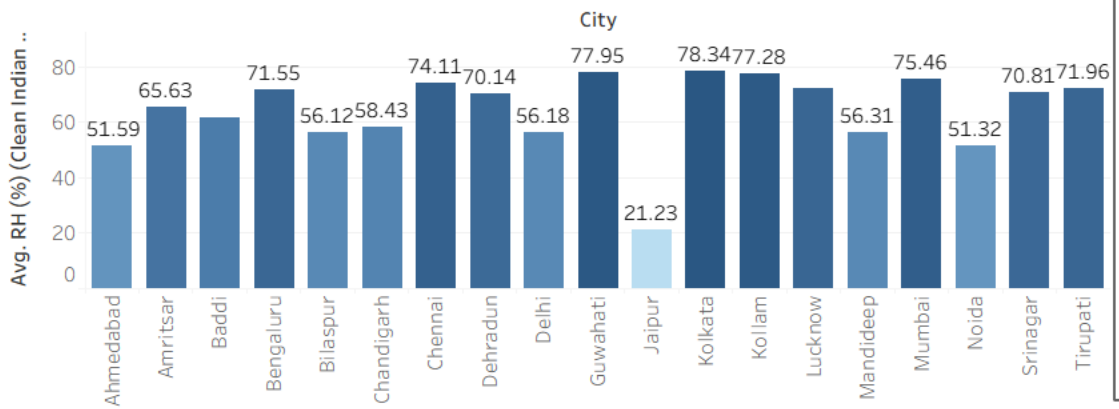
Objects
- Extension
- Pulse Metric
- Image
- Blank
- Workflow
- Web Page

Tiled Floating

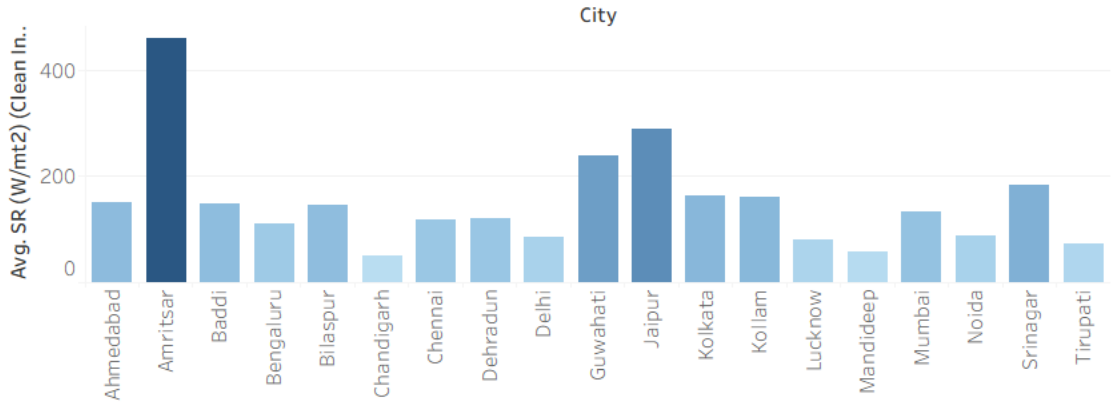
Show dashboard title

Air Quality Analysis of Indian States (2010–2023)

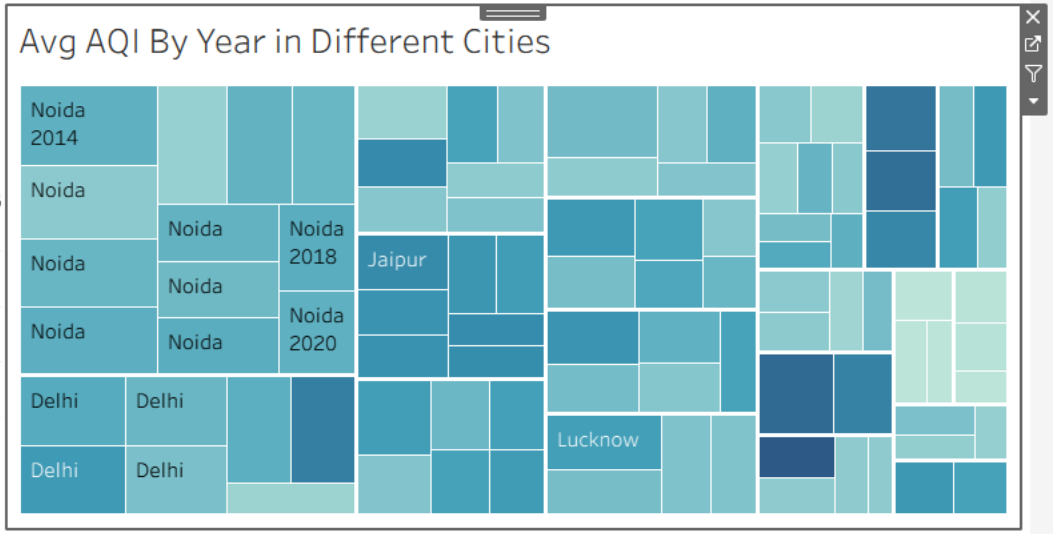
Avg Of Relative Humidity by cities



Solar Radiation (W/mt2) in Cities



Avg AQI By Year in Different Cities



Avg Of PM2.5 Distribution by year in Cities

From Date (Clean Indian AirData1.csv)

City	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Ahmeda..							52.4	46.3	52.8	44.4	49.8
Amritsar					72.4	55.7	50.5	46.4	51.0	51.0	69.1
Baddi										61.9	88.6
Bengaluru								21.9	30.9	38.3	34.7
Bilaspur									26.4	25.3	28.0
Chandig..									57.4	56.9	64.2
Chennai									15.8	26.5	28.4
Dehradun										36.9	31.9
Delhi			67.5			109.7	119.8	109.3	123.9	97.4	94.3
Guwahati							53.1	62.0	64.7	57.5	135.5
Jaipur					66.5	50.2	36.9	38.1	55.9	51.9	52.5
Kolkata						66.4	60.6	49.9	47.8	51.9	82.2
Kollam								40.3	37.9	32.8	36.0

Data Source | Sheet 1 | Sheet 2 | Sheet 3 | Sheet 4 | Sheet 5 | Sheet 6 | Sheet 7 | Sheet 8 | Sheet 9 | Dashboard 2 | Dashboard 3 | Sheet 10 | Dashboard 4 | Story 1 | Sheet 11

95 marks 1 row by 1 column SUM of AVG(Ozone (ug/m3) (Clean Indian AirData1.csv)): 3,105.21

34°C Haze Search ENG IN 11:58 25-07-2025

Tableau Public - Book3

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Tableau Desktop Public Edition Buy Tableau

Show Me

Dashboard Layout

Default
Phone
Device Preview

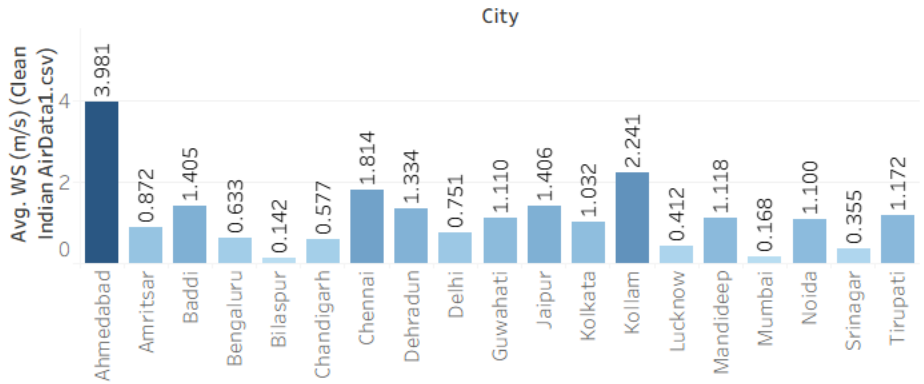
Size
Fit to height: width: 1420

Sheets
- Sheet 5
- Sheet 6
- Sheet 7
- Sheet 8
- Sheet 9
- Sheet 10
- Sheet 11

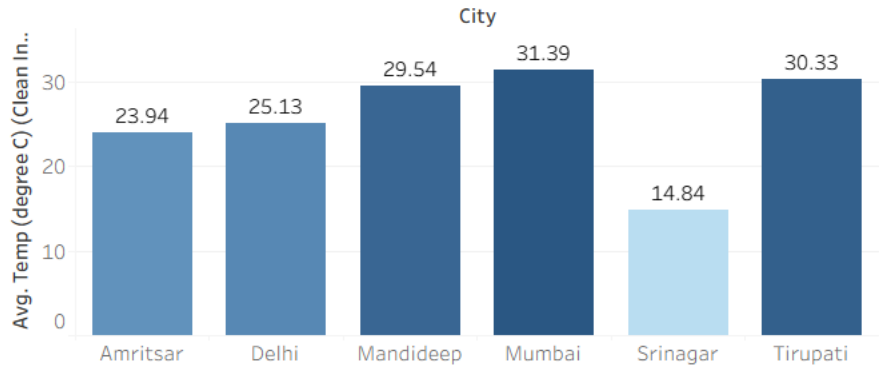
Objects
- Extension
- Pulse Metric
- Image
- Blank
- Workflow
- Web Page

Tiled Floating

Show dashboard title

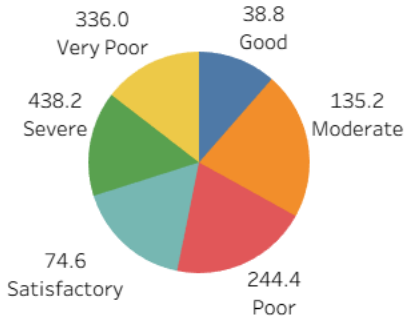


Average of WS (m/s) (Clean Indian AirData1.csv) for each City. Color shows average of WS (m/s) (Clean Indian AirData1.csv). The marks are labeled by average of WS (m/s) (Clean Indian AirData1.csv).



Average of Temp (degree C) (Clean Indian AirData1.csv) for each City. Color shows average of Temp (degree C) (Clean Indian AirData1.csv). The marks are labeled by average of Temp (degree C) (Clean Indian AirData1.csv). The view is filtered on City, which keeps 6 of 19 members.

Sheet 11



Data Source | Sheet 1 | Sheet 2 | Sheet 3 | Sheet 4 | Sheet 5 | Sheet 6 | Sheet 7 | Sheet 8 | Sheet 9 | Dashboard 2 | Dashboard 3 | Sheet 10 | Dashboard 4 | Story 1 | Sheet 11

34°C Haze

Search

ENG IN

12:09 25-07-2025

KEY INSIGHTS:

- PM2.5 and PM10 are the major contributors to pollution
- Delhi consistently has the highest AQI across years
- Winter months show the worst air quality (Oct–Jan)
- Monsoon tends to improve AQI significantly
- Several cities exceeded “Very Poor” AQI levels during lockdown recovery

CONCLUSION & IMPACT:

- Successfully calculated and classified AQI from raw sensor data
- Built scalable Tableau dashboards for monitoring pollution trends
- Project demonstrates skills in:
  - ✅ Data cleaning
  - ✅ AQI domain knowledge
  - ✅ Data visualization
  - ✅ Python & Tableau integration

FUTURE WORK:

- Real-time AQI prediction using ML models
- Integrate with Google Maps for live visualization
- Automate dashboard updates using APIs
- Add government policy impact analysis