**AIR QUALITY ANALYSIS**

**MINI PROJECT REPORT**

**Submitted by**

**VARSHINI S H (22ADR114)**

**MOHAMMED AADHIL A (22ADR065)**

**SASVATH K R (22ADR098)**

***for***

**22ADF01 DATA ANALYSIS**

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE**



**KONGU ENGINEERING COLLEGE**

**(Autonomous)**

**PERUNDURAI ERODE – 638 060**

**NOVEMBER 2024**

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE KONGU ENGINEERING COLLEGE**

**(Autonomous)**

**PERUNDURAI ERODE – 638 060**

**NOVEMBER 2024**

**Department of Artificial Intelligence**

**22ADF01 – Data Analysis Project Report**

Signature of course in-charge Signature of the HOD

Submitted for the continuous Assessment viva voice examination held on \_\_\_\_\_\_\_\_\_\_\_\_

**EXAMINER I EXAMINER II**

**ABSTRACT**

In the modern era, monitoring and analyzing air quality is essential for public health and environmental sustainability. With increasing urbanization and industrialization, cities worldwide are facing critical air pollution challenges that directly impact the health and well-being of their populations. This project focuses on analyzing air quality data across various cities, by examining critical pollutants along with their corresponding Air Quality Index (AQI). The dataset includes records from multiple cities of India, spanning over years from 2015 to 2020, providing valuable insights into trends and patterns of air pollution levels. The aim of this analysis is to assess air quality trends and to aid in policy-making and environmental interventions. By analyzing pollutant concentration levels, seasonal variations, and AQI values, this project strives to provide actionable insights into the severity of air pollution in different regions. Data preprocessing techniques such as handling missing values, segregating data by city, and performing fill methods (mean, mode, forward fill) have been applied to ensure the completeness and reliability of the dataset. The project leverages data visualizations and dashboards to communicate key findings and trends effectively. These visual tools provide a comprehensive overview of air quality fluctuations and pollutant behavior over time, allowing stakeholders to make data-driven decisions for improving air quality management. Ultimately, this analysis contributes to a better understanding of air pollution dynamics, supporting initiatives to enhance public health and reduce pollution-related risks. The findings will also assist in empowering cities to take preventive measures to mitigate environmental harm and improve the quality of life for their residents.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **CHAPTER**  **No.** | **TITLE** | **PAGE NO.** |
|  | **ABSTRACT** | 1 |
| 1. | **INTRODUCTION** | 3 |
|  | 1.1 INTRODUCTION | 3 |
|  | 1.1. DATA COLLECTION | 4 |
|  | 1.2 PROBLEM STATEMENT | 5 |
|  | 1.3 BUSINESS OBJECTIVE | 6 |
| **2.** | **DATA PREPARATION AND MODELING** | 7 |
|  | 2.1 DATA CLEANING | 7 |
|  | 2.2 DATA TRANSFORMATION | 8 |
|  | 2.3 DATA DISTRIBUTION USING CHARTS | 9 |
|  | 2.4 DATA MODELLING | 12 |
| **3.** | **DATA ANALYSIS AND INTERPRETATION** | 13 |
|  | 3.1 DATA ANALYSIS | 13 |
|  | 3.2 PUBLISHING DASBOARDS | 22 |
|  | 3.3 INFERENCE | 24 |
| **4.** | **CONCLUSION** | 29 |
|  | 4.1 RECOMMENDATIONS | 29 |
| **5.** | **REFERENCES** | 33 |

**CHAPTER 1**

**INTRODUCTION**

* 1. **INTRODUCTION**

The importance of air quality monitoring has grown significantly in recent years due to the increasing levels of air pollution caused by rapid urbanization, industrialization, and population growth. Air pollution poses serious health risks and environmental challenges, making it crucial to track and analyze pollutant levels in urban areas. Cities worldwide are facing the challenge of maintaining healthy air quality levels, as rising pollution impacts the well-being of residents and the sustainability of the environment. This report focuses on the analysis of air quality data across several cities, including Ahmedabad, with a particular emphasis on pollutants such as PM2.5, PM10, NO, NO2, CO, and SO2. The dataset spans multiple years, providing valuable insights into trends and variations in pollutant concentrations. By examining this data, the report aims to identify pollution hotspots, understand the correlation between various pollutants and their effect on the overall Air Quality Index (AQI), and offer a clear picture of air quality across different regions. The primary goal of this analysis is to provide actionable insights that can help inform air quality management strategies. By leveraging data visualizations and dashboards, the analysis seeks to make air quality trends easier to interpret for decision-makers, allowing them to take informed actions to mitigate pollution. This report will explore the relationships between pollutants, seasonal changes, and their contribution to AQI values, providing a comprehensive overview of air quality dynamics. Ultimately, this work aims to support initiatives aimed at reducing air pollution and promoting healthier, more sustainable environments in urban areas.

**1.1.1 DATA COLLECTION**

For this project, the air quality dataset is sourced from Kaggle, specifically from the “Clean Air - India’s Air Quality” dataset, which provides extensive air quality data across multiple cities in India. The dataset comprises 25,555 rows and 16 columns, tracking pollutant measurements in 21 cities from the year 2015 to 2020. This dataset offers a valuable foundation for analysing air quality patterns and identifying pollution trends across different regions.

Key attributes in the dataset include:

* **City**: The name of the city where air quality data was recorded, enabling geographic comparisons and regional analysis.
* **Date**: The specific date of each record, allowing for temporal analysis, including the identification of seasonal pollution trends.
* **PM2.5** and **PM10**: Concentrations of fine (PM2.5) and coarse (PM10) particulate matter, which are essential for evaluating air pollution and its impact on respiratory health.
* **NO**, **NO2**, and **NOx**: Measurements of nitrogen oxides, which are critical indicators of pollution from transportation and industrial sources.
* **CO**: Carbon monoxide levels, predominantly from combustion processes, which have significant health implications.
* **SO2**: Sulfur dioxide concentration, commonly from industrial emissions, which contributes to acid rain and affects air quality.
* **O3**: Ozone levels, relevant to smog formation and public health impacts.
* **AQI**: The Air Quality Index value, which provides an overall measure of air quality based on pollutant levels, with higher values indicating worse air quality.
* **AQI\_Bucket**: The AQI category (Good, Moderate, Poor, Satisfactory, Severe, Very Poor), simplifying air quality assessment for public understanding.

The dataset includes some missing values across attributes, necessitating preprocessing to ensure data accuracy and reliability. Techniques such as mean, mode, forward-fill, and backward-fill are applied to handle these missing entries effectively. The variety of pollutant measurements and AQI categories provides a comprehensive basis for analyzing air quality trends, identifying seasonal patterns, and understanding pollution sources across cities. This robust dataset supports a detailed analysis aimed at generating insights that inform air quality management strategies and policy-making efforts.

**1.2 PROBLEM STATEMENT**

With increasing urbanization and industrial activity, air quality has become a critical concern for public health and environmental sustainability. As cities across India experience rising levels of air pollution, there is a growing need for comprehensive air quality analysis to support informed decision-making and policy development. This project aims to provide a structured analysis of air quality data, offering insights into pollution trends and health risks associated with various pollutants.

The primary focus of this analysis is to examine air quality across major cities, identifying pollutant levels, seasonal patterns, and potential sources of pollution. By exploring trends in particulate matter (PM2.5 and PM10), nitrogen oxides, carbon monoxide, sulfur dioxide, and other pollutants, this analysis seeks to reveal critical factors impacting air quality in different regions. Additionally, the project aims to identify periods or areas with particularly high pollution levels, providing valuable information for targeted interventions. This structured approach will help government agencies and environmental organizations prioritize their efforts to improve air quality and safeguard public health

**1.3 BUSINESS OBJECTIVE**

The business objective of this air quality analysis is to provide actionable insights that support data-driven strategies for improving air quality management and public health. By understanding pollutant trends, seasonal fluctuations, and regional air quality variations, this analysis aims to empower policymakers, environmental agencies, and urban planners in making informed decisions that address air quality concerns effectively.

Leveraging insights from this analysis, stakeholders can identify high-risk areas and periods, enabling targeted interventions such as stricter emissions controls, pollution reduction policies, and public awareness campaigns. Additionally, understanding pollution patterns across cities will assist in allocating resources more effectively and prioritizing efforts in areas with severe air quality challenges. In summary, this project aims to support efforts to enhance environmental quality and protect public health by providing a comprehensive and data-informed view of air quality across major urban centers in India.

**CHAPTER 2**

**DATA PREPARATION AND MODELING**

**2.1 DATA CLEANING**

**Data Cleaning** is essential for identifying and rectifying errors, inconsistencies, and missing values to enhance the reliability of the dataset. For the air quality dataset, containing detailed pollutant information across 21 cities, we implemented the following steps to ensure data accuracy and consistency :

* **Handling Missing Values**: Missing values in this dataset, particularly for columns like PM2.5, PM10, NO2, and AQI, could significantly affect analysis if not appropriately addressed. To manage missing values, we applied several methods based on the nature of the data:
  + **Imputation**: For key pollutant columns with moderate levels of missing data, such as PM2.5 and PM10, we used mean or median imputation to fill gaps, considering the typical distribution of these pollutants.
  + **Forward Fill (ffill) and Backward Fill (bfill)**: For time-series data such as daily pollutant measurements, we applied forward fill and backward fill techniques to carry values forward or backward when gaps were minimal, preserving the temporal flow of data.
  + **Mode Imputation**: For categorical columns like AQI\_Bucket, where the mode provided a reasonable estimate, missing values were replaced with the most frequent entry.
* **Ensuring Data Consistency**: Standardizing data is crucial for maintaining uniformity across the dataset.
  + **City Names**: To ensure consistency across cities, all city names were standardized, preventing duplicate or misspelled entries.
  + **Pollutant Measurement Units**: We verified that each pollutant was recorded in consistent units across the dataset, crucial for valid comparison and aggregation.
  + **Handling Outliers and Inconsistent Values**: Outliers and improbable values in pollutant data, such as extreme CO, SO2, and O3 values, can skew results. We implemented the following steps to address these:
  + **Outlier Detection**: For pollutants like PM2.5 and PM10, which can exhibit spikes due to seasonal or geographic factors, extreme values were reviewed within the context of typical city and seasonal levels. Outliers that appeared implausible, given known environmental standards, were flagged and reviewed.

These data-cleaning procedures ensured a well-prepared and reliable dataset, forming a robust foundation for further analysis of air quality trends across India’s urban areas.

**2.2 DATA TRANSFORMATION**

Data Transformation is a crucial step in preparing the air quality dataset for analysis. This process involved structuring and modifying the data to enhance interpretability and enable more granular insights:

* **Regional Segmentation**:
  + **City Segregation**: The dataset was segmented based on cities to allow focused analysis on specific locations. This step provided clearer insights into city-level air quality patterns and comparisons.
  + **Regional Categorization**: Cities were grouped into broader regions—**Northern**, **Southern**, and **Coastal**—enabling comparative analysis across different geographic areas. This categorization allowed for the examination of regional air quality trends and the impact of geography on pollutant levels.
* **Date Transformation**:
  + **Year, Month, and Day Extraction**: The Date column was split into Year, Month, and Day components. This made it easier to analyze seasonal trends, yearly changes, and monthly averages, giving a better understanding of how air quality fluctuates over time.
* **Creating Measures for Averages**:
  + **Pollutant Averages**: New measures were created to calculate average pollutant levels (e.g., Average\_PM2.5, Average\_PM10, etc.), allowing for quick comparisons and trend analysis. These averages provided a summary view, enabling a high-level assessment of air quality across different time periods and regions.
  + **Regional Averages**: Regional average measures for each pollutant were also calculated for Northern, Southern, and Coastal cities, offering insights into how air quality compares across different areas of the country.

These transformations prepared the dataset for a thorough analysis of air quality across Indian cities, allowing for focused regional and seasonal insights. The creation of averages and other derived measures provided a foundational view, supporting a detailed and meaningful examination of trends in air quality across the nation.

**2.3 DATA DISTRIBUTION USING CHARTS**

The various visualizations, including bar charts, line graphs, pie charts, and maps, reveal both patterns and critical insights into the air quality levels across various Indian cities. This interactive setup allows users to filter and explore pollution trends, which can be useful for decision-making, policy planning, and identifying areas needing intervention.

**Types of Charts Used**

* **Bar Charts:** These were used to represent AQI severity and concentration levels of specific pollutants in cities like Bengaluru, Hyderabad, and Delhi. Bar charts also display AQI levels comparison between north and south Indian cities, helping to highlight regional pollution disparities.
* **Line Charts:** The line charts capture temporal pollution trends, such as monthly or annual variations in pollutant concentrations. For instance, the chart tracking Hyderabad's nitrogen compound levels and the concentration of particulate matter over time in Delhi allow for seasonality and year-over-year trend analysis.
* **Pie Charts:** Pie charts illustrate proportional data, like the pollutants' distribution in coastal regions and average AQI distribution by city. This helps in quickly understanding the share of each pollutant or AQI category.
* **Map Visualization:** A world map showing "City and AQI Bucket" visualizes the global context, placing the Indian air quality data within a broader environmental landscape, which can be used for cross-regional comparisons.

**Examples of Specific Distributions**

* **AQI Severity Levels:** The dashboard’s AQI severity chart highlights cities with severe air quality concerns. This section allows stakeholders to identify regions where immediate intervention is necessary to address hazardous air quality conditions.
* **Regional Pollution Trends:** A comparison between north and south AQI levels shows disparities, possibly due to varying industrial activity, population density, and natural factors affecting pollutant dispersion. This can guide localized strategies for air quality improvement.
* **Pollutant Type Distribution in Specific Cities:** Concentration data for pollutants like PM2.5, PM10, nitrogen oxides, and ozone across cities such as Chennai and Hyderabad offers insights into the primary pollutants affecting each location. These insights can inform targeted actions, like vehicle emissions control or industrial regulations, based on the dominant pollutants in each region.
* **Average AQI by City:** By visualizing AQI levels for different cities, the dashboard identifies areas where air quality consistently falls below safe levels, signaling long-term health risks for residents. This can help prioritize cities for air quality monitoring enhancements.
* **Seasonal and Altitudinal Variations:** Trends in pollutants by year or season (e.g., pollutant concentration in Chennai during summer) highlight how air quality worsens during certain times. The altitude-based distribution of PM2.5 and PM10 concentrations further shows how natural factors impact pollutant concentration levels.

**Key Insights and Strategic Implications**

This dashboard provides valuable insights into air pollution patterns across India, revealing areas with urgent air quality issues and opportunities for targeted interventions. Key insights include the north-south AQI disparity, seasonal pollution spikes, and city-specific pollutant types. Policymakers and environmentalists can use these insights to prioritize resources, design pollution control measures, and improve public awareness campaigns, ultimately aiming for healthier and more sustainable urban environments across India.

**2.4 DATA MODELING**

Data Modelling for the air quality analysis project primarily used statistical techniques to understand pollution trends, identify regional variations, and assess seasonal patterns across various cities. Rather than complex machine learning models, we focused on applying straightforward statistical approaches to extract actionable insights from the dataset.

Our analysis employed key statistical methods to provide a comprehensive understanding of air quality patterns. Descriptive statistics, including averages, standard deviations, and ranges, were calculated for pollutants such as PM2.5, PM10, CO, SO₂, and NO₂ across various cities, highlighting areas with consistently high pollution levels. We also performed trend analysis by breaking down the date field into year, month, and day components, enabling us to uncover seasonal and annual fluctuations in pollutant levels, which often correlated with factors like increased winter pollution or significant reductions in other months. Finally, correlation analysis was conducted to examine relationships between pollutants, such as PM2.5 and NO₂, helping to identify pollutant pairs that tend to rise together and offering insights for targeted pollution control strategies. The statistical methods provided an informative view of air quality levels across various cities and timeframes without the need for machine learning. This approach allowed us to pinpoint high-risk periods and areas, support targeted interventions, and suggest potential areas for future pollution mitigation efforts.

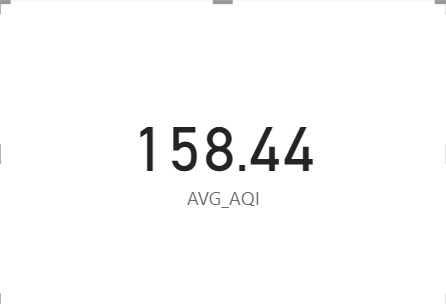
**CHAPTER 3**

**DATA ANALYSIS AND INTERPRETATION**

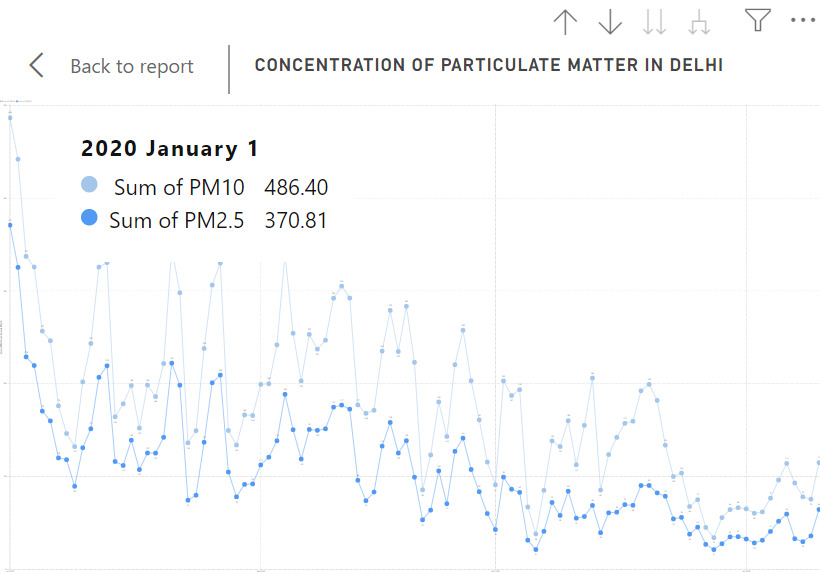
**3.1 DATA ANALYSIS**

Data analysis is central to this project, providing a quantitative foundation for understanding air quality patterns across various cities. Key analyses included descriptive statistics, trend analysis, and the exploration of specific questions related to pollutant levels, seasonal trends, and regional differences in air quality. By breaking down pollutant data and examining correlations, we aimed to uncover meaningful insights that could inform pollution control and public health strategies.

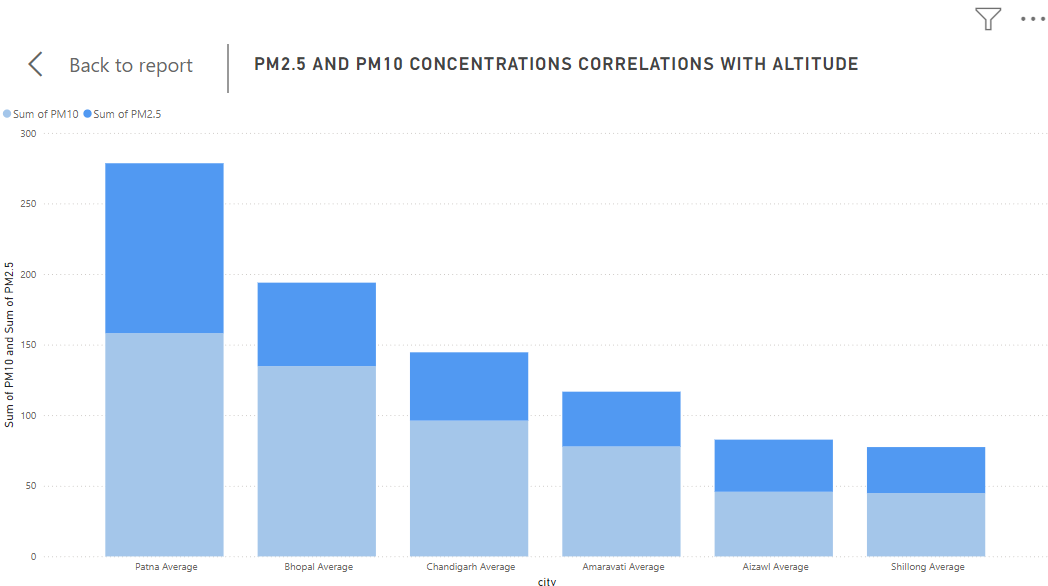
**1. What is the Average AQI of India ?**



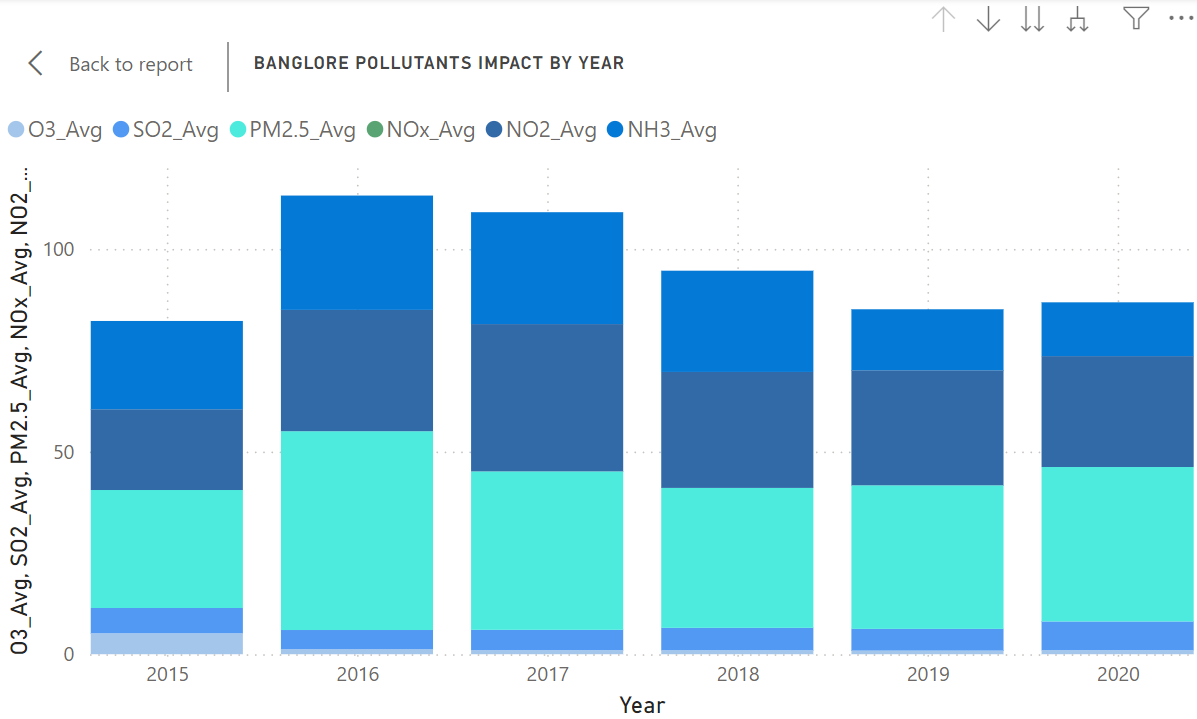
**2. Analyse the concentration of Particulate matter in Delhi in the years in 2020.**



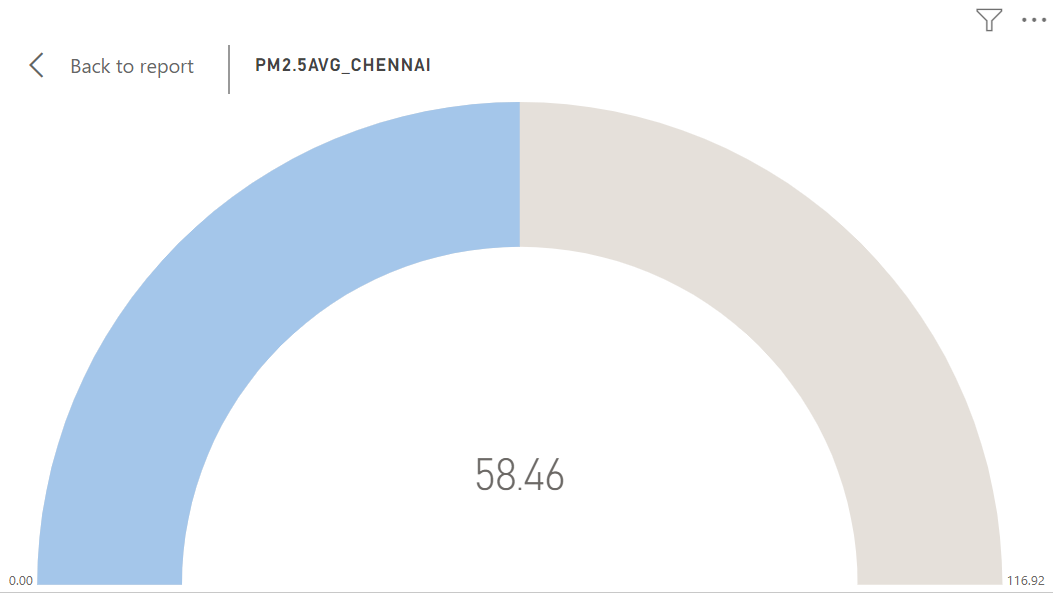
**3.What trends, if any, can be observed in seasonal variations of PM2.5 and PM10 at higher altitudes?**



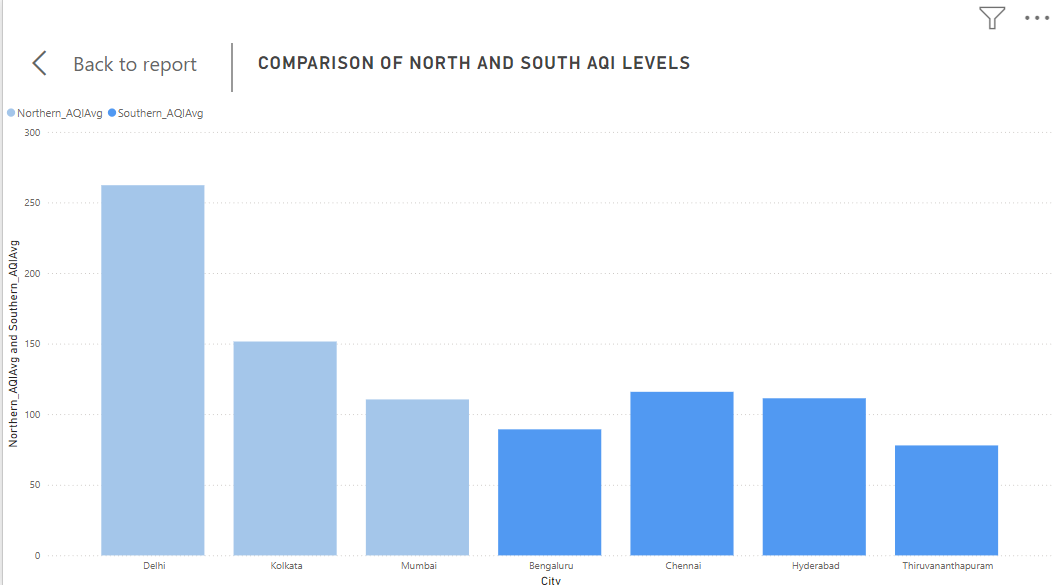
**4. Anayze the impact of pollutants in Banglore in the years between 2015 and 2020.**



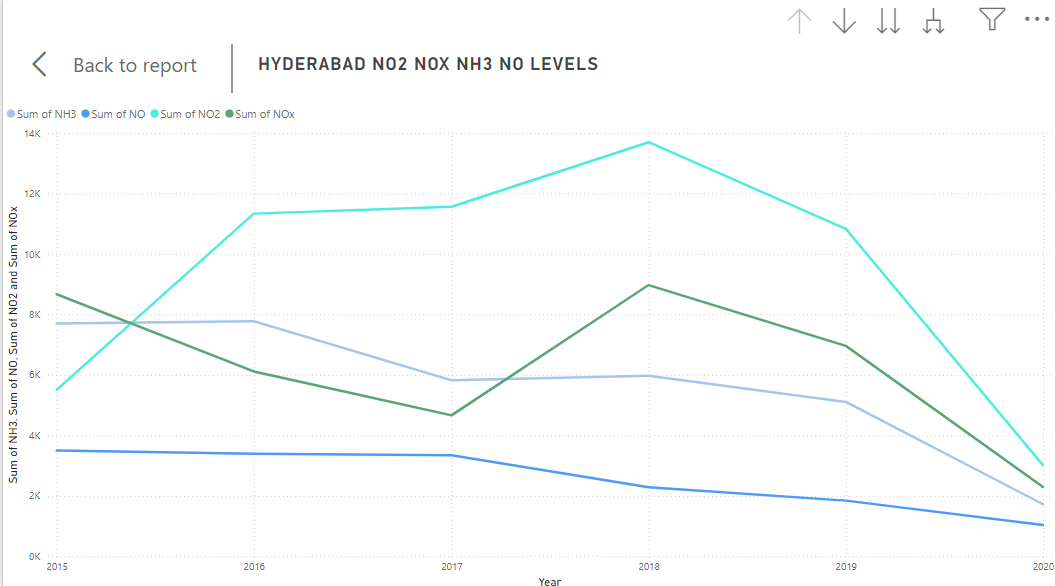
**5. What is the Average PM2.5 concentration in Chennai?**



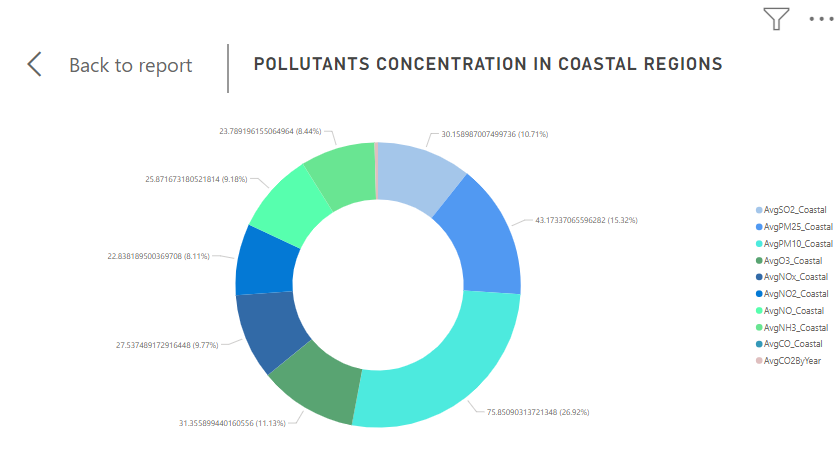
**6.What are the average AQI values for the northern and southern regions, and how do they reflect regional air quality differences?**



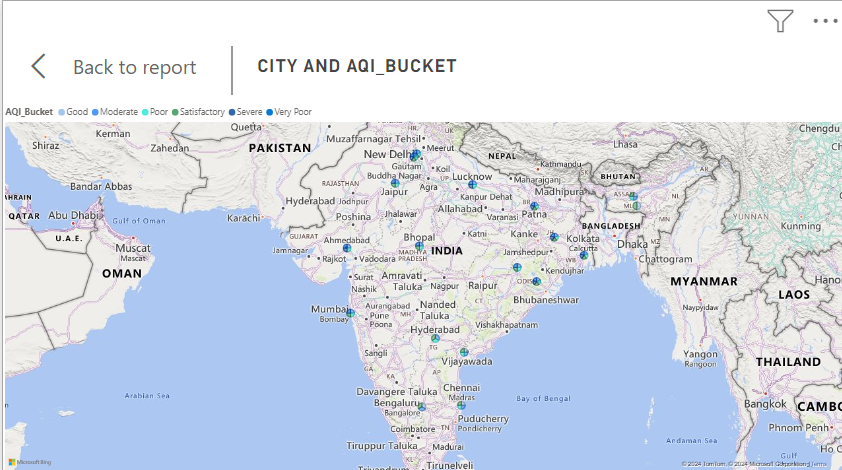
**7.Analyze the concentration of Nitrogen compounds in Hyderabad.**



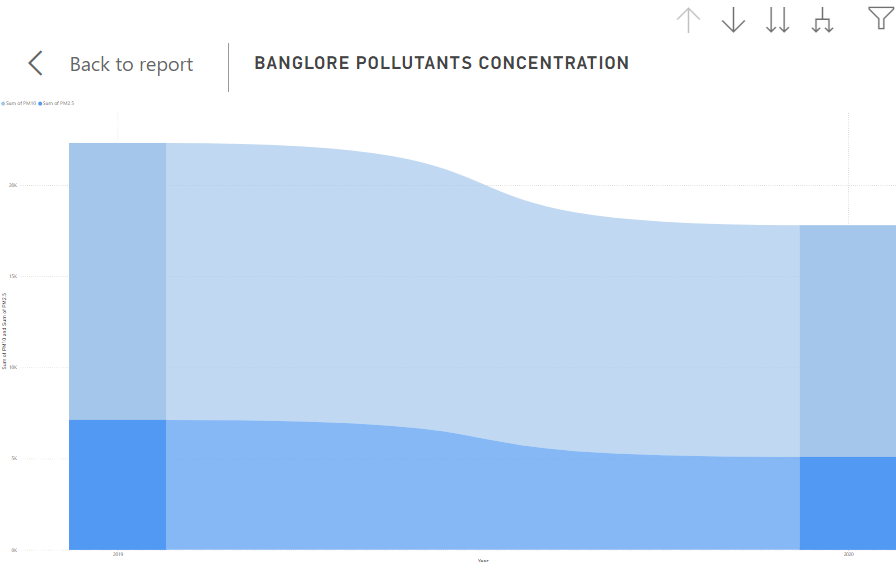
**8.Analyze the concentration of pollutants Coastal region.**



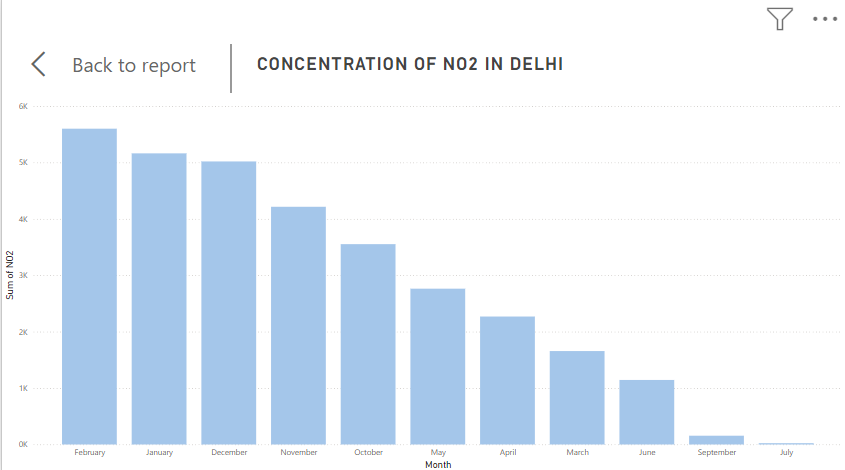
**9.Analyze the AQI Bucket for cities of India**



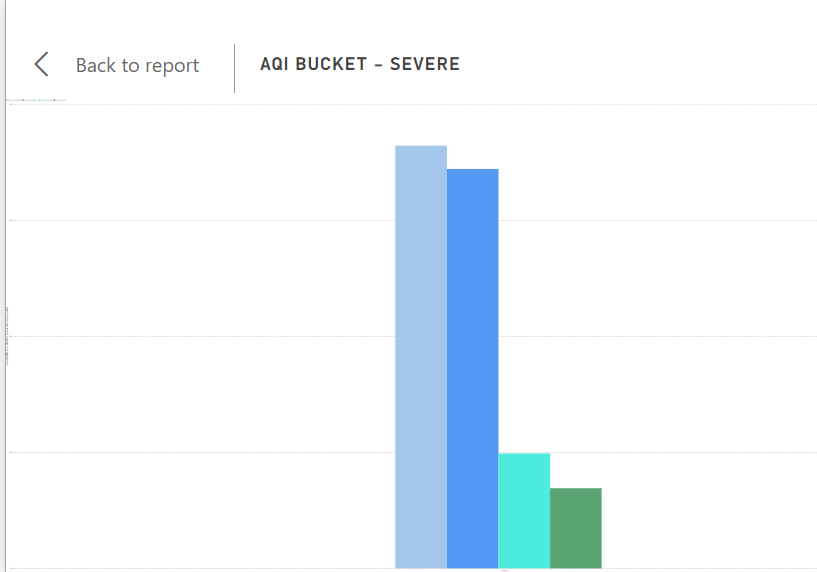
**10.Did the concentrations of the pollutants in Banglore significantly change during the initial phase of the COVID-19 lockdown ?**



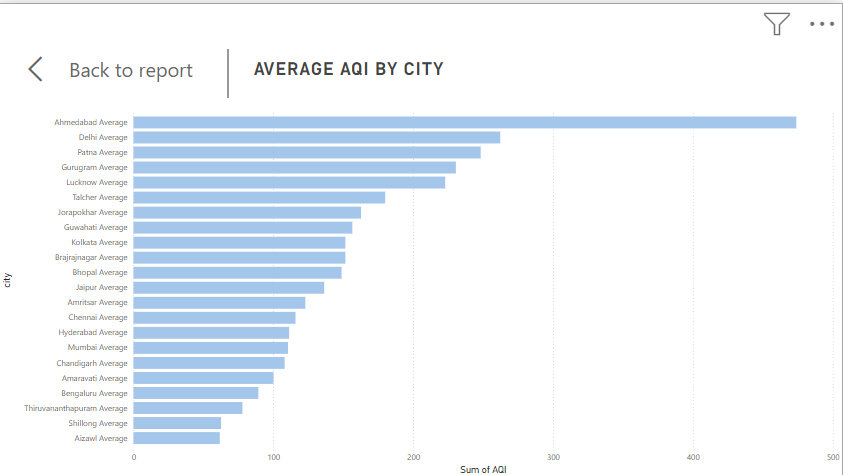
**11.Analyze the concentration of NO2 in Delhi with AQI bucket as Very Poor.**



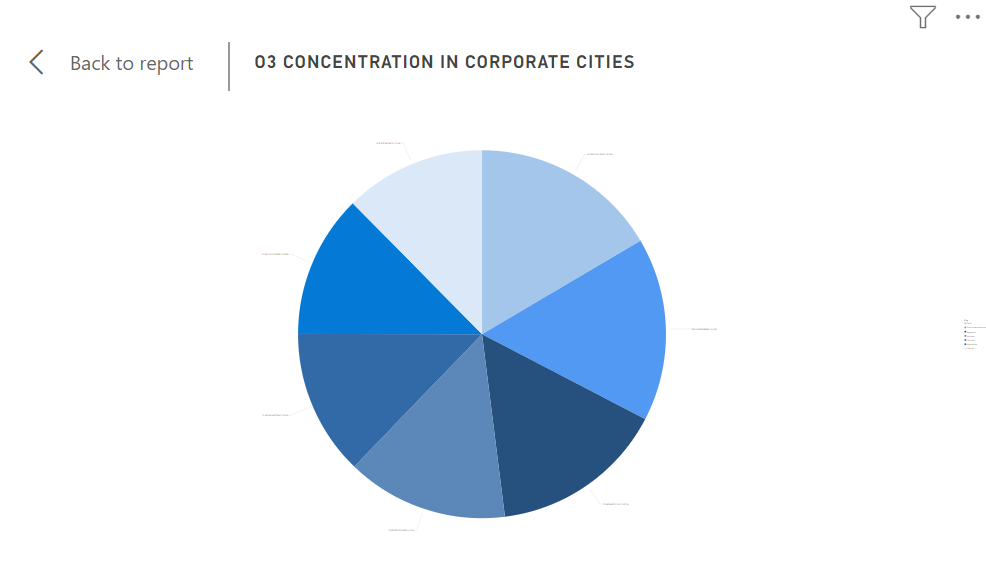
**12. Analyze the AQI Bucket of severe based on pollutants Nox, O3, PM2.5 , PM10.**



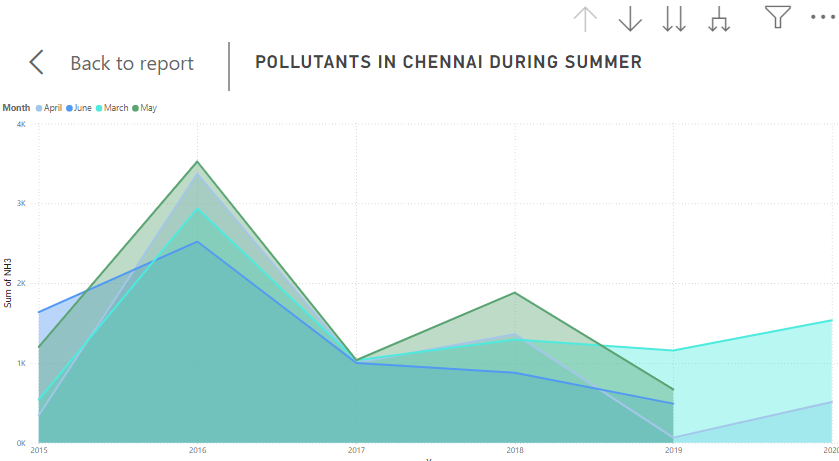
**13.Which city has the highest average Air Quality Index (AQI) across all the listed cities in the dataset?**



**14.Analyze the O3 concentration Corporate cities**



**15.Analyze the pollutants concentration in Chennai during summer.**



* **Descriptive Statistics**:
  + **Average AQI for India**: The average AQI for India was calculated as 158.44, providing a benchmark for assessing overall air quality.
  + **City-specific Pollutant Concentrations**: For instance, in Delhi during 2020, the highest recorded concentrations were 486.40 for PM10 and 370.81 for PM2.5 on January 1, indicating severe pollution levels in early winter.
  + **Regional AQI Differences**: Northern cities like Delhi, Kolkata, and Mumbai had average AQIs of 262.26, 151.54, and 110.51, respectively, while southern cities like Bengaluru, Chennai, and Hyderabad showed lower AQI levels, with Thiruvananthapuram at 77.92, highlighting regional disparities in pollution.
* **Trend Analysis**:
  + **Seasonal Variations at Higher Altitudes**: There is a marked seasonal variation in PM2.5 and PM10 concentrations at higher altitudes, with PM2.5 levels significantly exceeding PM10 levels, indicating altitude-related pollutant accumulation trends.
  + **Pollutant Trends in Bangalore**: Between 2015 and 2020, Bangalore saw relatively constant pollutant concentrations, with the highest overall levels in 2016 and the lowest in 2015. Notably, O3 was only present in 2015, showing variations in pollutant behaviour over time.
  + **Impact of COVID-19 Lockdown on Bangalore's Air Quality**: Analysis indicated minimal change in pollutant levels during the initial COVID-19 lockdown phase, though a slight downward trend was observed from 2018 to 2020.
* **Specific Questions Analysed**:
  + **How Does AQI Differ Across Coastal vs. Inland Cities?** Coastal cities like Chennai and Mumbai had higher PM10 concentrations, while inland cities like Delhi and Jaipur displayed greater levels of pollutants like NOx, reflecting the impact of geographic factors on air quality.
  + **How Did AQI Buckets Change in Delhi During 2015?** The concentration of NO2 was highest in Delhi in February 2015 when the AQI bucket was categorized as "Very Poor," underscoring the link between seasonal factors and pollutant spikes.
  + **Which Pollutants Are Predominant in the AQI Bucket ‘Severe’?** Across cities, the AQI Bucket labeled "Severe" had the highest concentrations of PM10, followed closely by PM2.5, suggesting that particulate matter significantly impacts poor air quality ratings.

These analyses provided actionable insights into pollution patterns, supporting regional comparisons, and temporal trends in pollutant levels. By examining specific pollutants and their variations across seasons, cities, and regions, we aimed to better understand the factors that contribute to air pollution in urban Indian settings.

**3.2 PUBLISHING DASHBOARDS**

To make the air quality analysis results accessible and interactive, we developed dashboards to visualize key metrics and trends across cities and regions. These dashboards enable users to filter, explore, and derive insights from the data interactively, providing an intuitive understanding of complex air quality metrics. The dashboards were created using Power BI, offering a user-friendly interface for data visualization and exploration.

* **Overview of Dashboards**:
  + **Main Dashboard**: The primary dashboard includes several interactive components that allow users to filter data by city, pollutant type, year, and AQI category. This flexibility enables users to view specific subsets of data, making it easier to explore air quality trends across different cities and time periods.
  + **Key Metrics and Insights**: The dashboard displays essential metrics such as average AQI by city, pollutant concentration distributions, seasonal patterns, and regional comparisons. Visualizations include bar charts for pollutant contributions, line graphs for seasonal trends, and heat maps to illustrate AQI variations across regions.
* **Explanation of Tools Used**: Power BI was chosen for its robust data visualization capabilities, which support the creation of interactive and visually appealing dashboards. Power BI’s filtering and drill-down features allow users to explore detailed air quality metrics, making it easier to interpret and analyze data at both regional and city levels.
* **Insights Presented on the Dashboard**:
  + **Regional AQI Distribution**: Pie charts and bar graphs display the AQI distribution across regions (Northern, Southern, Coastal), providing a quick overview of which areas experience higher pollution levels.
  + **Pollutant Contribution by City**: Bar charts and line graphs show the impact of different pollutants (PM2.5, PM10, NO2, CO) on overall AQI by city, offering insights into which pollutants are the primary contributors to poor air quality in specific locations.
  + **Seasonal Trends in Pollutant Levels**: Line charts visualize seasonal patterns, showing how pollutant levels fluctuate across months, helping stakeholders understand seasonal effects on air quality.
  + **City Comparisons by Pollutant**: Comparative charts display pollutant concentrations across cities, enabling users to assess how factors like geography (e.g., coastal vs. inland) influence pollution levels.

These dashboards provide a comprehensive view of air quality metrics, supporting interactive data exploration that informs regional air quality management and intervention strategies.

The Dashboard (Figure 1) presents a comprehensive analysis of air quality data across various cities, broken down by multiple factors. It includes visualizations that show the distribution of AQI levels, pollutant concentrations (such as PM2.5, PM10, CO, and NO2), and seasonal patterns across different regions, with minimum and maximum pollutant values highlighted. Key metrics, such as average AQI by region, pollutant contribution by city, and AQI bucket categories, provide deeper insights into air quality trends. Additionally, the dashboard displays data on seasonal pollutant variations, regional comparisons, and pollutant concentration distributions for coastal, northern, and southern cities. These visualizations collectively provide a detailed overview of air quality across India, allowing for comparisons and pattern recognition across cities, regions, and time periods, supporting data-driven insights into pollution management and public health strategies.



Figure-1 Dashboard

**3.3 INFERENCE**

1. The average AQI for India is calculated as 158.44, placing it in the “Unhealthy” category. This indicates that air quality in many parts of the country may pose health risks, particularly to sensitive groups like children, the elderly, and individuals with respiratory conditions. This average reflects high levels of pollution in urban areas and industrial zones, where pollutants like PM2.5 and PM10 are prevalent. A high national AQI average underscores the need for targeted pollution control measures, especially in areas contributing most to the elevated AQI. Maintaining and monitoring this average can help guide policy adjustments to improve air quality over time.
2. In Delhi, the concentration of particulate matter peaked on January 1, 2020, with PM10 levels at 486.40 and PM2.5 levels at 370.81. These values far exceed safe limits, indicating severe pollution during this period, likely exacerbated by winter weather conditions that trap pollutants close to the ground. High PM levels contribute significantly to respiratory issues and visibility reduction in Delhi. This data highlights the need for urgent measures in Delhi to reduce particulate emissions, especially from vehicles and construction, which are major contributors to these pollutants.
3. Seasonal variations reveal that at higher altitudes, the concentration of PM2.5 tends to be significantly higher than PM10. This trend may be due to limited industrial activity but an accumulation of fine particulate matter, which can linger longer in the air at elevated regions. PM2.5 poses severe health risks as it penetrates deeper into the lungs. Seasonal analysis shows elevated PM levels during colder months, which could be linked to increased heating activities and atmospheric conditions that inhibit pollutant dispersion.
4. The data indicates that Bangalore experienced consistent pollutant concentrations from 2015 to 2020, with the highest levels in 2016. Ozone (O3) was observed only in 2015, suggesting potential inconsistencies in O3 measurement or a notable shift in emissions. The constancy in pollutant levels implies that urban sources of pollution, such as traffic emissions, remain significant contributors. Observing slight declines in 2019 and 2020 could indicate the early impact of pollution control measures or a reduction in traffic during the COVID-19 pandemic.
5. The average PM2.5 concentration in Chennai is calculated as 58.46, which is above the recommended safe limit, indicating moderate air pollution. While this level is relatively lower than that in heavily industrialized cities, it still presents health risks, especially for vulnerable populations. PM2.5 is primarily from traffic emissions and industrial activities, which are consistent sources in Chennai. This moderate yet persistent pollution level highlights the need for long-term strategies to reduce emissions from urban sources to maintain healthier air quality.
6. Northern regions, particularly Delhi and Kolkata, have higher average AQI values (262.26 and 151.54, respectively) than southern cities like Bengaluru and Thiruvananthapuram, with AQIs around 89.31 and 77.92. This difference reflects the influence of industrial activity, population density, and climatic conditions in the northern region. Northern cities suffer from higher pollution levels due to factors like winter inversion layers and proximity to industrial hubs. The cleaner air in southern cities emphasizes the benefits of geographic and meteorological conditions in dispersing pollutants more effectively.
7. In Hyderabad, nitrogen compounds, including NO, NO2, NOx, and NH3, showed varying concentration trends from 2015 to 2020. In 2015, NOx was the most concentrated nitrogen compound, while NO2 reached peak concentrations by 2018. This suggests a steady increase in traffic and industrial emissions, as NO2 is primarily a byproduct of combustion processes. Elevated NO2 levels can aggravate respiratory problems and contribute to ground-level ozone formation. Monitoring nitrogen compounds is essential in Hyderabad to control their harmful impacts and reduce urban air pollution.
8. Coastal regions exhibit a high concentration of PM10, with CO2 being the least concentrated pollutant. The elevated PM10 levels may be attributed to sea salt, dust, and coastal industrial activities, while the lower CO2 concentration could be due to effective dispersion over the sea. Coastal cities face unique air quality challenges, balancing natural marine aerosols and urban emissions. These findings suggest the importance of specific pollution control measures that address both natural and anthropogenic pollutants in coastal areas.
9. The AQI Bucket categorization reveals that cities like Ahmedabad, Delhi, and Lucknow in northern India frequently fall into higher AQI categories, indicating poor air quality. Southern cities, such as Bengaluru and Pondicherry, are generally in lower AQI categories, reflecting comparatively cleaner air. This regional variation suggests that cities in the north face more significant pollution challenges, likely due to industrial density, climatic conditions, and vehicular emissions. The AQI bucket data highlights areas where stringent pollution control measures are urgently needed.
10. The analysis indicates minimal changes in pollutant concentrations in Bangalore during the initial COVID-19 lockdown, with only slight variations from 2018 to 2020. The data suggests that while emissions from vehicles may have decreased, other sources such as industrial activities continued to contribute to overall pollution levels. The limited impact of the lockdown underscores the importance of comprehensive pollution control measures that go beyond temporary reductions in traffic emissions.
11. When Delhi’s AQI Bucket was categorized as “Very Poor” in February 2015, NO2 concentrations were at their peak, indicating a strong correlation between NO2 emissions and overall poor air quality. High NO2 levels are a marker of vehicular and industrial emissions, which are prevalent sources in Delhi. This data highlights the need for stringent NO2 emission controls in urban areas, particularly during seasons when atmospheric conditions worsen air quality.
12. In cases where the AQI Bucket is labeled “Severe,” PM10 is the dominant pollutant, followed closely by PM2.5. This finding emphasizes that particulate matter is a significant driver of extremely poor air quality. High levels of PM10 and PM2.5 in the “Severe” AQI category suggest that emissions from dust, vehicles, and industrial activities need strict regulation to prevent reaching dangerous air quality levels.
13. Ahmedabad has the highest average AQI across all cities in the dataset, reflecting a persistent pollution problem. This may be due to its dense industrial zones and high vehicular emissions. A consistently high AQI indicates long-term exposure to unhealthy air, necessitating focused mitigation efforts in Ahmedabad to reduce emissions and improve public health.
14. Among corporate cities, Delhi exhibits the highest O3 concentration, while Kolkata has the lowest. Elevated ozone levels are likely due to high NOx and VOC emissions in densely populated and industrially active areas. Ozone at ground level contributes to smog and respiratory issues, indicating that cities with high O3 concentrations should focus on controlling precursor pollutants to reduce ozone formation.
15. Pollutant levels in Chennai peak during May, reflecting an accumulation of pollutants due to dry, stagnant summer conditions that hinder dispersion. Higher temperatures in summer can also increase photochemical reactions, raising levels of secondary pollutants like ozone. Monitoring and addressing pollutant levels during this season is essential to ensure public health and manage pollution-related risks effectively.

**CHAPTER 4**

**CONCLUSION**

The analysis of air quality data across Indian cities reveals both critical challenges and areas for targeted improvement. While cities in southern and coastal regions, such as Chennai and Bengaluru, generally exhibit moderate AQI levels due to natural dispersion advantages, northern and inland cities, like Delhi and Ahmedabad, face severe air quality issues linked to industrial density, vehicular emissions, and seasonal weather patterns that trap pollutants. Our findings indicate that particulate matter (PM2.5 and PM10) and nitrogen oxides (NOx) are major contributors to poor air quality, especially in densely populated urban centers. Additionally, seasonal variations and higher pollution during colder months suggest the need for proactive measures during these high-risk periods. By implementing region-specific pollution control strategies and strengthening regulations on primary pollutants, cities can improve public health outcomes and reduce environmental risks. Addressing these air quality challenges will not only enhance quality of life but also support sustainable urban growth, positioning cities to meet environmental standards and improve overall well-being

**4.1 RECOMMENDATIONS**

The AQI dashboard provides a comprehensive view of air quality data and trends. By leveraging these insights, several targeted recommendations can help improve air quality monitoring, raise awareness, and promote better engagement with users. Here’s how you can optimize the AQI dashboard:

1. **Monitor and Improve Air Quality in Critical Areas**:
   * **Highlight High-Pollution Areas**: Use the dashboard to visually highlight regions with consistently high AQI scores. This will draw attention to critical areas that need intervention and allow local authorities to prioritize efforts in high-risk zones.
2. **Focus on Cities with Poor Air Quality**:
   * **Targeted Content for High-Pollution Areas**: The dashboard should provide insights into regions with the poorest air quality. By focusing efforts on these areas, both local governments and users can be informed about pollution sources, trends, and steps to take to mitigate the impact.
3. **Improve Air Quality Management with Actionable Insights**:
   * **Highlight Pollution Sources**: The dashboard can display real-time pollution data based on different sources (e.g., industrial emissions, traffic, construction activities) to help users understand the major contributors to high AQI levels in specific areas.
   * **Showcase Air Quality Improvements**: As efforts to improve air quality are made (e.g., increased green spaces, emission controls), the dashboard can track and display the impact of these actions over time. This can motivate communities to engage in air quality management practices and see measurable improvements.
4. **Personalized Recommendations for Individuals**:
   * **Custom Air Quality Alerts**: Allow users to set personalized alerts based on their specific health needs (e.g., asthma or heart conditions). This feature can send notifications when air quality is expected to dip below certain levels, helping users make informed decisions about outdoor activities.
   * **Health Advisory Notifications**: The dashboard can provide health advisories based on AQI levels. For example, it can recommend reducing outdoor physical activity or wearing protective masks when the AQI reaches hazardous levels.
5. **Identify and Optimize Solutions for Air Quality Improvement**:
   * **Support for Policy Decisions**: The dashboard can provide historical data and trends to help policymakers understand the effectiveness of current air quality control measures. This could guide decisions such as adjusting traffic patterns, controlling industrial emissions, or implementing stricter regulations in high-pollution zones.
   * **Encourage Sustainable Practices**: By providing users with insights into air quality trends, the dashboard can also promote sustainable behaviors like using public transportation, reducing energy consumption, and adopting eco-friendly technologies to minimize pollution.
6. **Optimize User Experience and Interaction with the Dashboard**:
   * **Advanced Filtering and Data Visualization**: The dashboard should offer filtering options for users to explore air quality data based on factors like location, time period, and pollutant type. This will make it easier for users to understand specific trends and identify the sources of pollution that most affect their region.
   * **Interactive Maps for Regional Insights**: Use interactive maps to allow users to visually explore air quality data across different regions, helping them compare pollution levels and access detailed information about their specific locations.
7. **Regular Updates and Adaptive Recommendations**:
   * **Seasonal and Event-Based Recommendations**: The dashboard can automatically update to reflect seasonal changes in air quality (e.g., poor air quality during winter due to heating) and provide users with relevant, timely advice. It can also suggest actions based on special events that may lead to a spike in pollution (e.g., festivals or public holidays).
   * **Regular Content Refreshes**: The AQI data should be updated regularly on the dashboard, ensuring that users always have the most accurate and relevant information to make decisions regarding their health and activities.

**Conclusion: Enhancing the AQI Dashboard for Better Engagement and Impact**

These recommendations focus on enhancing the AQI dashboard's usability and making it more actionable for users, authorities, and communities. By optimizing the real-time monitoring features, offering personalized health recommendations, and providing insights for improving air quality, the dashboard can play a key role in managing pollution levels and supporting better public health outcomes.Incorporating these recommendations will not only help users stay informed but also encourage more proactive measures for maintaining a healthier environment. The dashboard will continue to be a vital tool in understanding, managing, and improving air quality in real-time.

**REFERENCES**

1. Alade IO, Rahman MAA, Saleh TA (2019a) Predicting the specific heat capacity of alumina/ethylene glycol nanofluids using support vector regression model optimized with Bayesian algorithm. Sol Energy 183:74–82.Kumar Pavan, Rashid Ahmad, B.D. Chaudary and K. Sinha Mukul, "Enriched dashboard:-An Integration and Visualization Tool for Distributed NLP System on Heterogeneous Platform", *International Conference on Computer Science and its Applications (ICCSA)*, 2013.
2. Al-Jamimi HA, Saleh TA (2019) Transparent predictive modelling of catalytic hydrodesulfurization using an interval type-2 fuzzy logic. J Clean Prod 231:1079–1088.
3. Ayturan YA, Ayturan ZC, Altun HO, Kongoli C, Tuncez FD, Dursun S, Ozturk A (2020) Short-term prediction of PM2.5 pollution with deep learning methods. Global NEST J 22(1):126–131
4. Bellinger C, Jabbar MSM, Zaïane O, Osornio-Vargas A (2017) A systematic review of data mining and machine learning for air pollution epidemiology. BMC Public Health..