

AIR QUALITY HEALTH DATA ANALYSIS REPORT

STUDENT NAME: ALBIN REGI JOHN

ROLL NUMBER : 2462604

COLLEGE EMAIL ID:

albin.regi@btech.christuniversity.in

COURSE : FUNDAMENTALS OF DATA ANALYTICS

INSTRUCTOR NAME: DHIRAJ ALATE

INSTITUTION : CHRIST UNIVERSITY BANGLORE

DATE OF SUBMISSION : 18/08/2025

Abstract

This project analyzes an air quality and health dataset to uncover patterns and insights related to environmental conditions and their impact on human well-being. The study explores regional variations in air quality, identifies key contributing factors, and examines associated health outcomes. Microsoft Excel is employed to design an interactive dashboard using PivotTables, charts, and slicers, enabling dynamic data exploration. The analysis is structured around five key research questions, guiding the discovery of trends and correlations within the dataset. The final deliverable includes a comprehensive dashboard and report, offering stakeholders a clear, data-driven tool to monitor air quality and assess its health implications, thereby supporting informed decision-making and policy development.

Objectives

- Clean and prepare the raw dataset for analysis.
- Formulate and answer five key business questions based on the dataset.
- Create a user-friendly dashboard that visualizes key sales metrics.
- Use charts and graphs to communicate insights effectively.
- Summarize the findings and their implications for business decision-making.

Scope of the Project

This project is focused on cleaning, analyzing, and visualizing the provided air quality and health dataset. The scope is limited to the given dataset and does not involve advanced programming techniques such as Python or R. All tasks—including data preparation, manipulation, visualization, and dashboard creation—were carried out exclusively in Microsoft Excel.

Tools & Technologies Used

- Microsoft Excel: Data manipulation, analysis, and dashboard creation
- PivotTables: Summarizing sales data for analysis
- Charts & Graphs: Data visualization of key sales trends

Data Cleaning & Preparation

The initial dataset contained records related to air quality indicators and associated health factors. The following steps were carried out to ensure data accuracy and consistency:

- Verified and corrected data types (e.g., dates, numeric pollution measures, and health statistics).
- Removed duplicate entries to maintain reliability of results.
- Checked for missing values and ensured completeness of key attributes.
- Standardized region names and health-related categories where inconsistencies were identified.
- Created calculated fields such as Air Quality Index (AQI) categories and health impact ratios for deeper analysis.

Dashboard Design Strategy

The dashboard was developed to highlight the most critical environmental and health metrics. Key visualizations include:

- **Line chart** to show air quality trends over time.
- **Bar chart** to compare pollution levels across different regions.
- **Pie chart** to illustrate the proportion of health impacts (e.g., respiratory cases) by category.
- **Slicers** for filtering by date, region, or pollutant type.

Challenges Faced & Solutions

- Challenge: Handling missing or inconsistent entries

Solution: Ensured data quality by cleaning and standardizing records.

- Challenge: Choosing effective chart types

Solution: Tested multiple visualizations and selected the most intuitive charts.

- Challenge: Ensuring dashboard usability

Solution: Added slicers and filters for interactivity and better insights.

Questions & Solutions

QUESTION NO 1:

WHICH REGION HAS THE HIGHEST PERCENTAGE OF NO2 ,CO AND O3 AS PER THE DASHBOARD ?

SOLUTION:

AS PER THE DASHBOARD EAST REGION HAS THE HIGHEST PERCENTAGE OF NO2 CO AND O3

QUESTION NO 2 :

WHICH REGION HSAS THE LOWEST PERCENTAGE OF NO2 ,CO AND O3 AS PER THE DASHBOARD ?

SOLUTION:

AS PER THE DASHBOARD CENTRAL REGION HAS THE LOWEST PERCENTAGE OF NO2 AND O3 AND WEST REGION HAS THE LOWEST PERCENTAGE OF CO

QUESTION 3:

WHICH FACTOR HAS HIGHEST VALUE OF SUM IN ALL REGIONS AS PER THE DASHBOARD ?

SOLUTION:

AS PER THE DASHBOARD "SUM OF VALUES OF GREEN COVER PERCENTAGE " HAS THE HIGHEST VALUE IN ALL REGIONS

QUESTION 4:

WHICH FACTOR HAS LOWEST VALUE OF SUM IN ALL REGIONS AS PER THE DASHBOARD ?

SOLUTION:

AS PER THE DASHBOARD "SUM OF VALUES OF TEMPERATURE " HAS THE LOWEST VALUE IN ALL REGIONS

QUESTION 5:

WHICH REGION HAS THE VALUE OF SOME OF AQI BY REGION AS HIGHEST ?

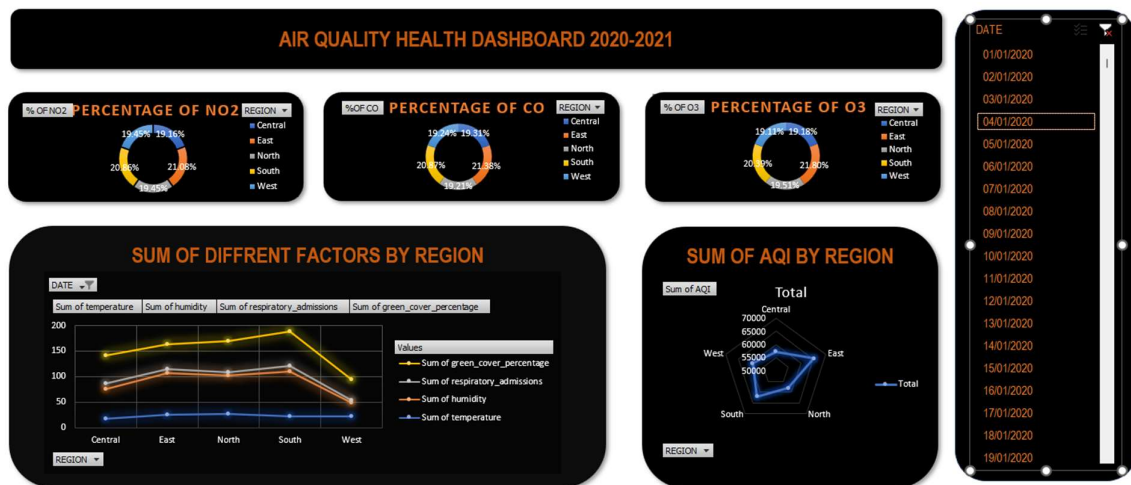
SOLUTION:

AS PER THE DASHBOARD "EAST REGION " HAS THE HIGHEST AQI ON COMPARING WITH OTHER REGIONS

Outcome

The project successfully delivered a clean and structured dataset, five key research questions answered, and a comprehensive dashboard in Excel. The dashboard provides insights into regional air quality patterns, pollutant levels, health impacts, and overall environmental trends. These insights can support policymakers, healthcare professionals, and researchers in identifying high-risk regions, understanding pollution–health correlations, and developing data-driven strategies for improving public health and environmental management.

Screenshots of Final Output



QUESTION NO 1:

WHICH REGION HAS THE HIGHEST PERCENTAGE OF NO₂, CO AND O₃ AS PER THE DASHBOARD ?

SOLUTION:

AS PER THE DASHBOARD EAST REGION HAS THE HIGHEST PERCENTAGE OF NO₂ CO AND O₃

QUESTION NO 2 :

WHICH REGION HSAS THE LOWEST PERCENTAGE OF NO₂, CO AND O₃ AS PER THE DASHBOARD ?

SOLUTION:

AS PER THE DASHBOARD CENTRAL REGION HAS THE LOWEST PERCENTAGE OF NO₂ AND O₃ AND WEST REGION HAS THE LOWEST PERCENTAGE OF CO

QUESTION 3:

WHICH FACTOR HAS HIGHEST VALUE OF SUM IN ALL REGIONS AS PER THE DASHBOARD ?

SOLUTION:

AS PER THE DASHBOARD "SUM OF VALUES OF GREEN COVER PERCENTAGE " HAS THE HIGHEST VALUE IN ALL REGIONS

QUESTION 4:

WHICH FACTOR HAS LOWEST VALUE OF SUM IN ALL REGIONS AS PER THE DASHBOARD ?

SOLUTION:

AS PER THE DASHBOARD "SUM OF VALUES OF TEMPERATURE " HAS THE LOWEST VALUE IN ALL REGIONS

QUESTION 5:

WHICH REGION HAS THE VALUE OF SOME OF AQI BY REGION AS HIGHEST ?

SOLUTION:

AS PER THE DASHBOARD "EAST REGION " HAS THE HIGHEST AQI ON COMPARING WITH OTHER REGIONS

Conclusion

This project provided hands-on experience in cleaning and analyzing air quality and health data using Microsoft Excel. It enhanced skills in data preparation, visualization, and dashboard creation, while demonstrating how Excel-based tools can be effectively used to derive meaningful insights. The analysis of the dataset highlighted the relationship between environmental factors and health outcomes, showcasing how accessible tools like Excel can support evidence-based decision-making in environmental and public health contexts.

References

- Microsoft Excel Official Documentation
- Online tutorials for PivotTables and Dashboard
- Project Submission Guidelines provided by the institution

