Global Health Statistics Analysis and Visualization

1. Project Overview

This project presents an analysis of a global health dataset with the aim of understanding disease impact, mortality rates, recovery outcomes, and healthcare access across countries. The project applies standard data analysis techniques and visualization practices to transform raw health data into structured insights.The results are summarized through exploratory data analysis and an executive-style dashboard created using Power BI.

2. Objectives

The objectives of this project are:

- To analyze disease-wise population impact

- To compare mortality and recovery rates across countries

- To assess healthcare access levels

- To examine the relationship between healthcare access and mortality

- To present key metrics and patterns through a clear dashboard

3. Dataset Description

The dataset contains global health-related records covering multiple diseases and countries. The main attributes include:

- Country – Name of the country

- Disease Name – Name of the disease

- Disease Category – Classification of diseases

- Population Affected – Number of individuals affected

- Mortality Rate (%) – Percentage of deaths

- Recovery Rate (%) – Percentage of recoveries

- Healthcare Access (%) – Measure of access to healthcare services

The dataset was cleaned and standardized prior to analysis to ensure consistency and accuracy.

4. Tools and Technologies

- Python

  - Pandas for data manipulation

  - Matplotlib and Seaborn for exploratory visualization

- Power BI
  - KPI visualization
  - Dashboard creation
  - Interactive filtering

5. Methodology

5.1 Data Preparation

- Verified data completeness and structure
- Ensured appropriate data types for numerical and categorical variables
- Prepared a clean dataset for analysis

5.2 Exploratory Data Analysis

Exploratory analysis was conducted to:

- Identify diseases with the highest population impact
- Compare mortality rates across countries
- Analyze recovery trends
- Examine healthcare access in relation to mortality

The EDA results guided the selection of metrics and visuals used in the dashboard.

6. Dashboard Design

A Power BI dashboard was developed to present summarized health indicators in a concise and readable format.

Dashboard Components

- Key Performance Indicators
  - Average Population Affected per Record
  - Average Mortality Rate
  - Average Recovery Rate
  - Average Healthcare Access
- Visualizations
  - Top 5 Diseases by Population Affected
  - Top 10 Countries by Median Mortality Rate
  - Healthcare Access vs Mortality (Scatter Plot)

- Filters
  - Country
  - Disease Category

Figure 1: Executive Power BI dashboard summarizing global health indicators.

7. Key Observations

- A limited number of diseases account for a significant proportion of the affected population

- Mortality rates vary considerably across countries

- Countries with higher healthcare access generally show lower mortality rates

- Disease categories exhibit differences in mortality and recovery outcomes

8. Conclusion

The project provides a structured analysis of global health statistics using Python and Power BI. By combining exploratory analysis with dashboard visualization, the project highlights key patterns in disease burden, mortality, and healthcare access across countries. The dashboard offers a clear summary of health indicators suitable for high-level interpretation.

9. Future Scope

- Inclusion of time-based data for trend analysis

- Region-level aggregation and comparison

- Integration of additional health and socioeconomic indicators

- Extension to predictive or risk-based analysis