Global Health Data Analytics & Dashboard Report

# Python Analytics Tasks

## 1. Data Cleaning

- Missing values were handled using techniques such as forward fill and imputation where applicable.  
- Inconsistent formats (e.g., numeric fields stored as strings) were corrected.  
- Outliers were detected using visualizations and statistical methods and either capped or removed.  
- Encoding (LabelEncoding) and feature scaling were applied.

## 2. Exploratory Data Analysis (EDA)

- Descriptive statistics (mean, median, std, etc.) were computed.  
- Distributions of numeric features were visualized using histograms.  
- Relationships among variables were explored with scatter plots and heatmaps.

## 3. Machine Learning / Clustering Model

- A clustering model (e.g., KMeans) was applied to group countries based on death rates.  
- Feature selection was based on domain knowledge and EDA insights.

## 4. Model Evaluation

- Evaluation used metrics like silhouette score to measure cluster cohesion and separation.  
- Visualizations of clusters were created to assess group characteristics.

## 5. Code Structuring & Documentation

- Code was modularized into functions such as `clean\_data`, `perform\_eda`, and `train\_model`.  
- Each section was clearly documented using markdown and in-code comments.

## 6. Innovation

- Custom visual analytics (e.g., seaborn pairplots, correlation matrix).  
- Model innovation included data transformation pipeline and feature engineering.

# Power BI Dashboard Tasks

## 1. Communication of Problem & Insights

- The problem of global mortality burden was addressed using WHO data.  
- Key insights were communicated through cards, tables, and summaries.

## 2. Interactivity Features

- Slicers for country and disease cause were added.  
- Drill-down features allowed user exploration at cause and gender levels.

## 3. Visuals Used

- Card visuals for KPIs such as total death rate, average rate, etc.  
- Clustered bar charts to show top countries by mortality rate.  
- Filled maps to visualize geographical patterns.  
- Tables to display detailed data records.

## 4. Design Clarity

- A consistent theme and layout was used.  
- Titles and axes were labeled appropriately.  
- Redundant or cluttered visuals were avoided.

## 5. Innovation

- Bookmarks for saved views like 'Cardio Focus' and 'Reset View'.  
- Buttons with action links to bookmarks.  
- Q&A visual added to allow natural language querying.  
- DAX measures used for rate calculations.