**Link to the notebook**

<https://github.com/mulombiannuar/MCSC_2108_Data_Visualization/blob/main/TB_Burden_Comprehensive_Notebook_Selected_Countries.ipynb>

**Data Visualization Practice Summary**

This is a complete data visualization assignment that was undertaken using the 'TB\_Burden\_Country.csv' dataset provided. The assignment was intended to implement important concepts of data visualization such as data discovery and preparation, visual components, principles of chart design, as well as advanced methods of visualization through Python packages Pandas, Matplotlib, and Seaborn which are my preferred tools for visualization for now

**Data Exploration and Preparation**

Firstly, the dataset was read in and it was explored to know its contents as well as its structure. The preliminary examination involved **reviewing for missing values, data types, as well as duplicate records**. **Columns were also renamed for clarity** whenever necessary while missing data points were appropriately dealt with. **Descriptive statistics was generated** to summarize distribution of main numerical variables such as TB incidence, prevalence, as well as mortality rates.

**Data Visualization Techniques**

**Bar charts** were used to compare TB incidence rates across countries, providing a clear visual comparison that helps identify nations with the highest or lowest burden. **Line graphs** illustrated trends in TB cases over time, which are essential for observing progress, stability, or regression in TB control efforts.

**Histograms** were applied to display the distribution of incidence and mortality rates, helping to assess data spread and detect skewness or concentration patterns. **Box plots** further explored variability and outliers, highlighting countries with unusual TB burden levels that may require policy attention. **Scatter plots** visualized the relationship between TB incidence and mortality, supporting the identification of possible correlations or risk clusters. Lastly, a **heatmap** was employed to show correlations among multiple TB indicators, offering a holistic view of how key variables interact.

These visualization techniques were chosen to **simplify complex datasets, make hidden patterns visible, and support evidence-based decision-making in global TB management**. Each visualization provided unique insights that collectively enhanced data interpretation and storytelling.

**Insights and Interpretation**

The visualizations revealed high TB-burden countries, TB case trends over time, as well as TB prevalence and socio-economic correlation measures. As an example, elevated TB occurrence was observed among countries with low GDP per capita, which brought forth the socio-economic dimension of health. Choice of coloring, appropriate scaling, as well as proper legends, enhanced interpretability as well as visual excellence.

**Conclusion**

This exercise in data visualization imparted hands-on experience of exploring, preparing, and visualizing data appropriately with Python. With TB\_Burden\_Country dataset, different methods of visualizations like bar charts, line plots, heatmaps, and correlation plots were used to discover patterns, relationships, as well as differences in tuberculosis burden with respect to country and year. The preparation of the data delivered credibility with cleaning, filling missing values, and devising common formats so that appropriate insights can be extracted from visualizations.

The research unveiled essential understanding of TB's global spread and temporal patterns of cases with a focus on highlighting how visualization makes massive datasets both meaningful and actionable. Through conversion of data into visual narratives, decision-makers and epidemiological researchers can more easily spot high-risk locations, track elimination target achievements, as well as make evidence-based decisions. In general, this exercise highlights that good data visualization is more than pretty pictures it is a critical analytical tool that connects data investigation with decisive action.