Tableau Project Report - Group 1

Scale and Risk: A Data-Driven Analysis of Global Tuberculosis Incidence and Resource

Team Members (Group 1):

Jennifer Sun

Linyu Zhao

Sangeeta Sheshachala

Sevgi Cosan Senbaslar

1. Introduction and Background

Tuberculosis (TB) is one of the most common infectious diseases in the world. We chose to examine the burden of tuberculosis by nation because it demonstrates the connections between infectious illnesses and variations in surveillance, health systems, and socioeconomic circumstances. These elements contribute to the uneven progress in the fight to eradicate tuberculosis. Our data covers the period from 1990 to 2013, when international TB control initiatives increased HIV/TB service integration, diagnosis, and treatment. Our goal is to determine the locations of the biggest burdens, how regional trends have evolved, and what the near future might hold if present trends persist.

Our guiding questions are:

1. Which countries and regions carry the highest **TB incidence and mortality**—measured both as rates per 100,000 and absolute counts?
2. How have TB rates changed over time (1990–2013), and are declines uniform across WHO regions (AFR, AMR, EMR, EUR, SEA, WPR)?
3. How strongly are HIV/TB co-infection and case detection rate associated with national TB rates?
4. What would happen in a what-if scenario where countries improve case detection or reduce incidence by a policy-relevant margin?

Answering these questions can help public health stakeholders prioritize resources, evaluate whether declines are on track, and communicate progress and gaps to non-technical audiences.

2. Datasets

We use a single, country-year panel dataset titled **TB_Burden_Country.csv** (5,120 rows; 219 countries/territories; 1990–2013; WHO regions). This data, compiled and harmonized by the **World Health Organization (WHO)**, offers standardized TB indicators with uncertainty bounds. Given that the WHO spearheads global TB control efforts, the data provides a credible, authoritative source for our analysis.

Core identifiers include:

- Country or territory name
- ISO 3-character country/territory code
- Region (WHO geographical region)
- Year

Population and burden metrics include:

- **Estimated incidence (all forms) per 100 000 population:** This is our primary measure of risk and burden, defined as the number of new and relapse TB cases per 100,000 people in a given year.
- **Estimated number of incident cases (all forms):** The absolute count of new and relapse cases, which measures the scale of the problem irrespective of population size.
- **Estimated mortality of TB cases (all forms, excluding HIV) per 100 000 population** and corresponding absolute counts.
- **Estimated HIV in incident TB (percent):** The proportion of incident TB cases who are also HIV-positive.
- **Case detection rate (all forms), percent:** The percentage of estimated new and relapse cases that were officially notified to national authorities.

Size & Coverage: The dataset covers 219 unique countries/territories across six WHO regions over 24 years (1990–2013).

Limitations: The data consists of estimates, which depend heavily on the quality of country reporting and surveillance systems; thus, under-diagnosis and under-reporting are likely variables. The provided **uncertainty intervals** must be respected in interpretation to communicate confidence transparently. Furthermore, because the data ends in 2013, trend statements are historical, and any forecasts we generate (in Section 3) will be illustrative projections.

Why this dataset fits: This dataset directly addresses our core questions about the location, scale, and time trends of TB, making it highly suitable for our project goals.

3. Data Story

To build a foundational understanding of the TB epidemic, the initial three charts are designed to answer our first two guiding questions: **Where is the burden highest?** and **How has incidence changed over time**?

Chart 1: World Map – TB Incidence Rate (per 100,000) *(Sangeeta)*

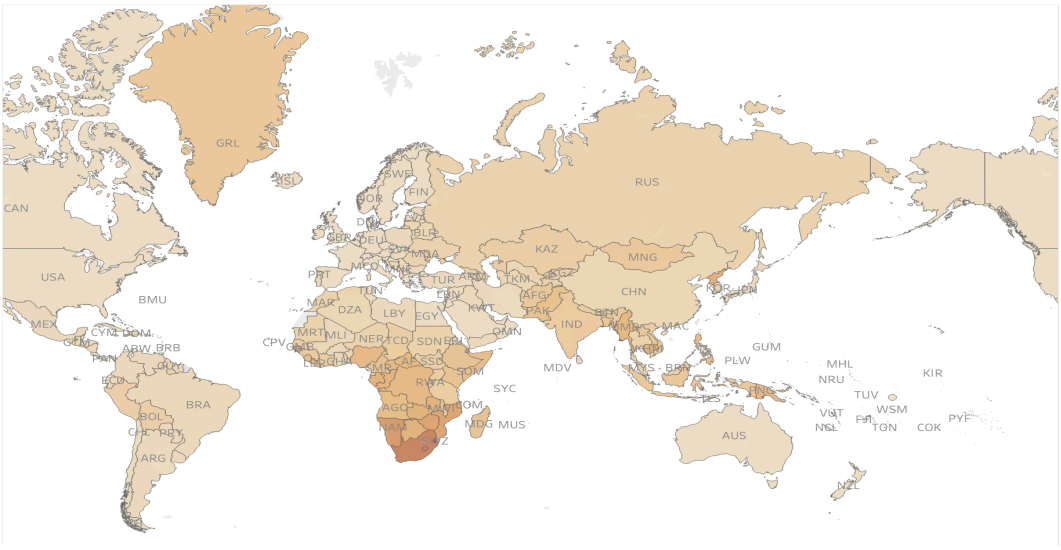


Chart 1 - World Map

This chloropleth map visualizes the **Estimated incidence (all forms) per 100 000 population** for **2013**, directly addressing our question of *where* the per-capita risk is highest. The map reveals stark regional inequalities, with the highest-incidence countries clustered predominantly in the **African Region (AFR)** and the **South-East Asian Region (SEA)**. This spatial context is vital for policymakers, as it immediately highlights the primary geographic hotspots requiring prioritized resource allocation.

Chart 2: Top 10 Countries by Absolute TB Cases *(Sangeeta)*

While the incidence rate communicates **risk**, this horizontal bar chart complements the map by focusing on **scale**, showing the **Estimated number of incident cases (all forms)** for the top 10 countries in **2013**. This view is crucial, as some highly populated countries may have a moderate rate but contribute the largest number of total cases globally. It clearly identifies that **India**, **China**, and **Indonesia** drive the majority of the global TB burden. For instance, India alone accounted for approximately **2,100,000** incident cases. The concentration of cases in these

few nations is critical for national and international planning, as success here will significantly impact worldwide disease control outcomes.

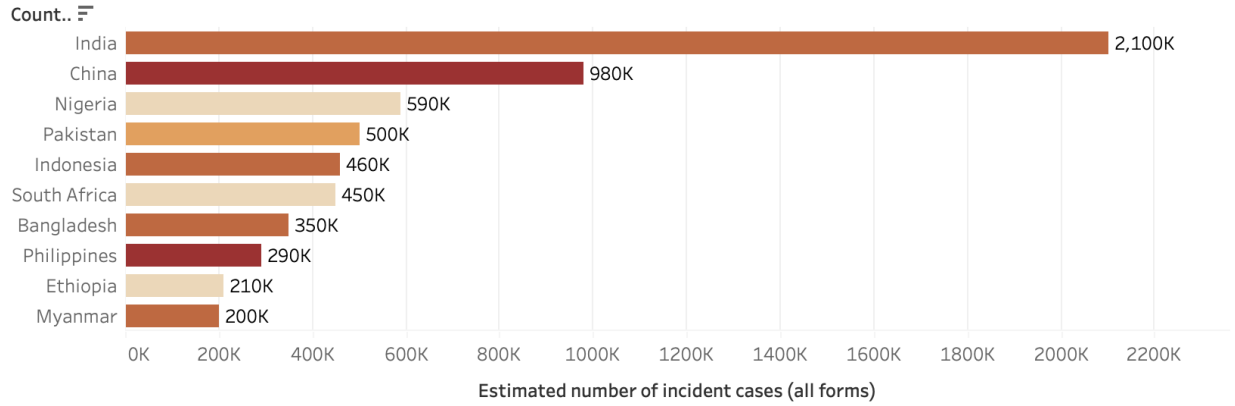


Chart2- Top 10 Countries

Chart 3: Global TB Incidence Trends (1990–2013) *(Sangeeta)*

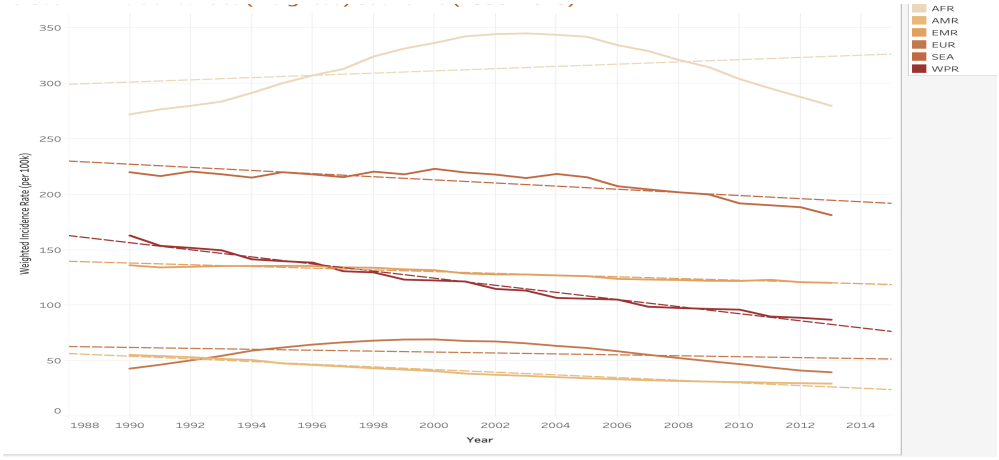


Chart 3 - Trend Line - Global TB incidence rate (weighted) over time (1990–2013)

This line chart tracks the change in **Global TB Incidence Rate (per 100,000)** from **1990 to 2013**, broken down by the six **WHO Regions**. The variable used on the Y-axis is the **Average Estimated incidence (all forms) per 100 000 population** for the countries within each respective region. The primary narrative is that while the **global incidence has steadily declined** over this period, progress is **highly uneven**. The **South-East Asian Region (SEA)** and **African Region (AFR)** maintain the highest incidence levels. Notably, the AFR line shows an initial **rise** in the 1990s before declining, which is largely attributed to the devastating **HIV epidemic** that increased TB susceptibility. In contrast, the European Region (EUR), Region of the Americas (AMR), and Western Pacific Region (WPR) maintained consistently low and steadily declining rates. This visualization confirms progress but underscores the need for region-specific public health strategies to address structural challenges.

Chart 4: TB Deaths (Excluding HIV) — Rate vs. Total Count for Top 10 Countries *(Linyu)*

The dual-axis chart highlights the importance of considering population size when interpreting mortality trends — a lower rate may still correspond to a high total death count. On the left y-axis, the measure represents the estimated mortality rate of TB cases per 100,000 population. The right y-axis displays the estimated total number of deaths from TB. As the goal is to demonstrate the greatest effect, we have filtered the countries list by displaying the data for the top 10 countries in terms of the death counts(y-axis). We conclude that in countries

with large populations, the absolute number of deaths remains high even though both the mortality rate(per 100K) and the total deaths have been declining.

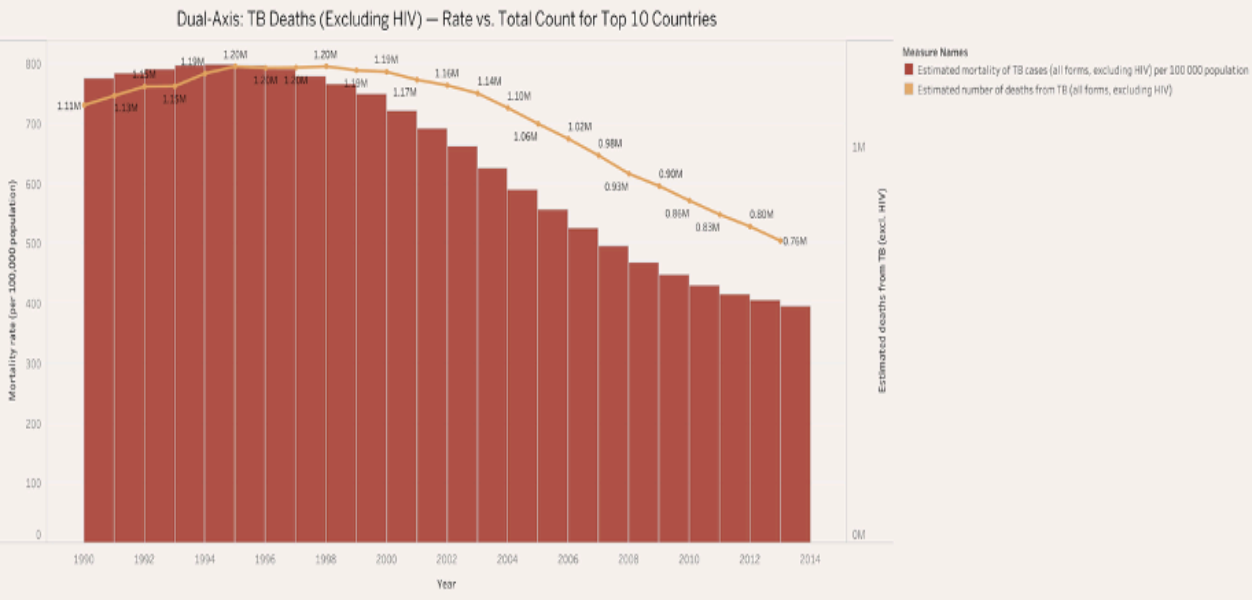


Chart 5: Forecast - Project incidence rate 3-5 years beyond 2013 *(Linyu)*

The forecast chart displays the trend of TB incidence rates for the top 10 countries, along with the projected rates from 2013 to 2018. The incidence rate refers to the *estimated incidence (all forms) per 100,000 population*. As shown in the middle of the chart, the incidence rate began to decline around 2003. Based on the historical trend, the TB incidence rate is projected to decrease by approximately 15–20%.

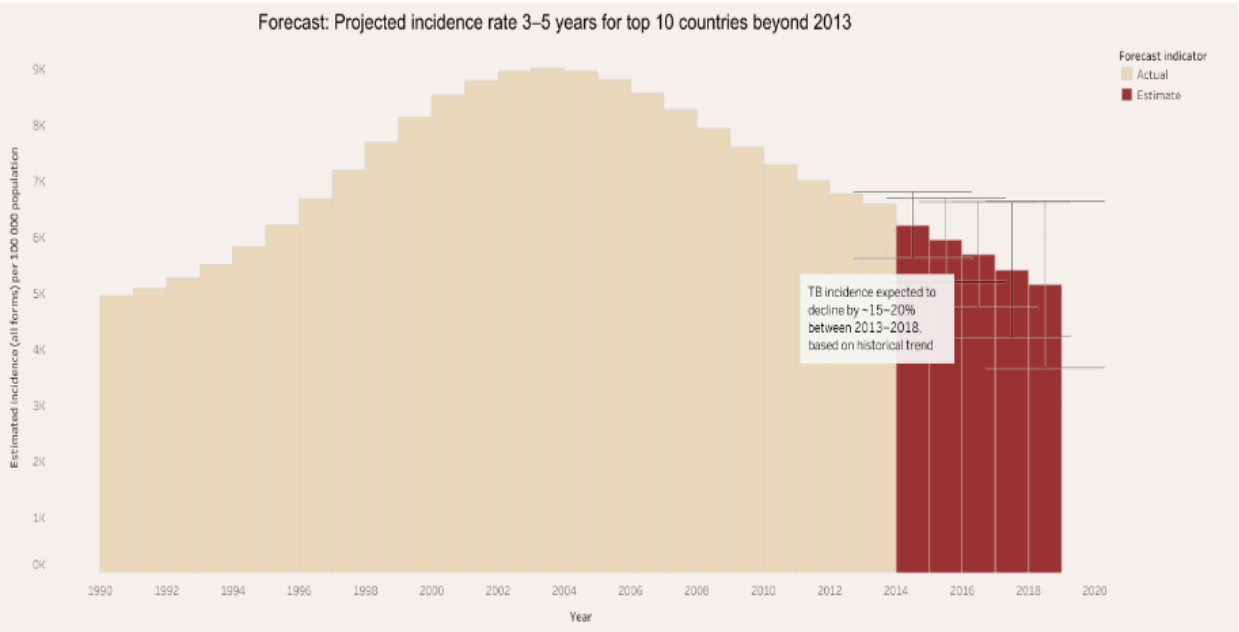
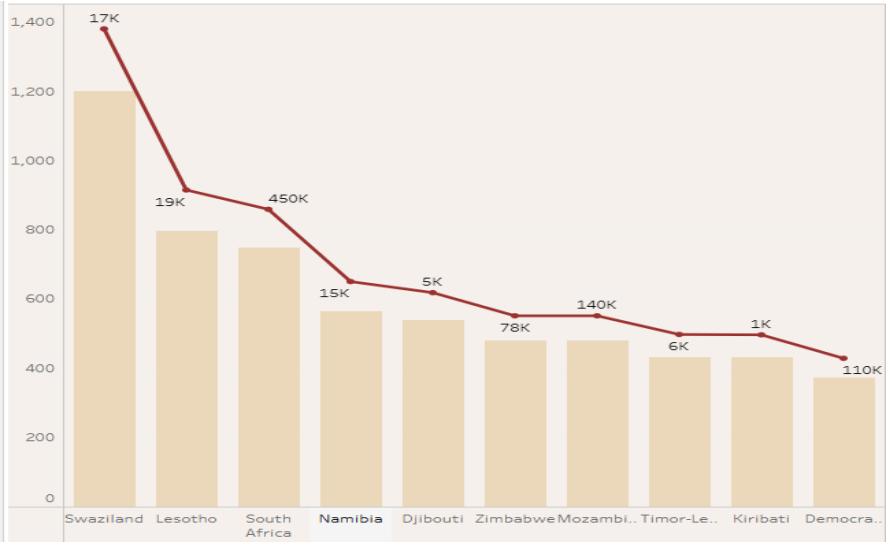
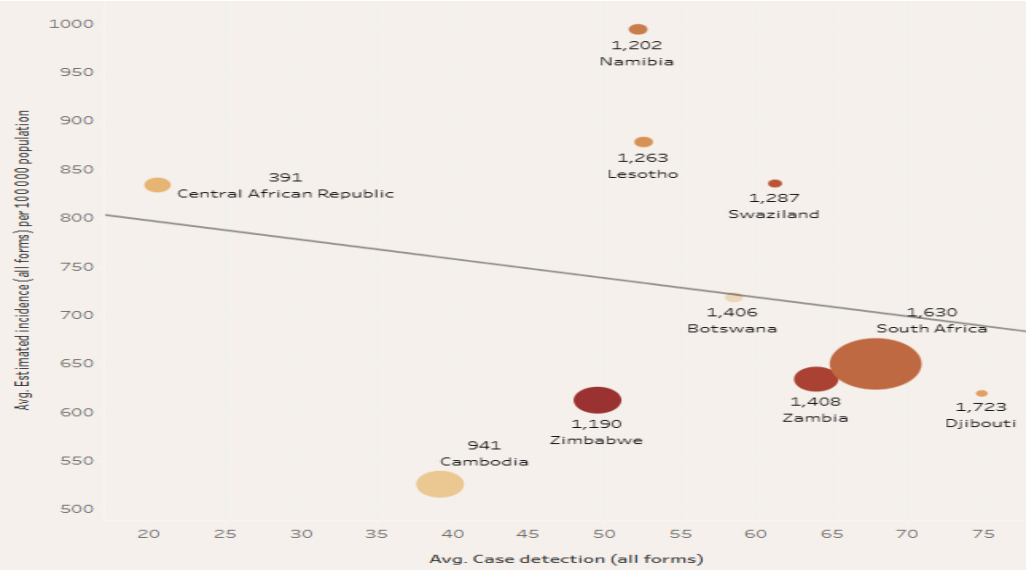


Chart 6: Impact of Policy - Driven Incidence Reduction Top 10 Countries *(Sevgi)*



The visual I have created here for illustrating the potential impact of public health interventions on disease incidence across countries. The graph shows a comparative view between two measures: the base incidence rate (representing the current estimated burden) and the adjusted incidence rate, which dynamically updates according to the selected value of the “Incidence Reduction (%)” parameter (0-30%). By sliding the parameter, users can simulate different intervention scenarios such as enhanced detection or preventive measures and observe the related decline in incidence rates. As the reduction percentage increases, the graph shortening of the adjusted bars reflects the potential decrease in disease burden. This interactivity enables user to have an understanding of how different levels of intervention could influence outcomes, thereby supporting evidence-based policy discussions and decision-making.

Chart 7: AVG Case Detection Top 10 Countries *(Sevgi)*



I created this visualization to highlight the effectiveness of tuberculosis management across countries by checking the relationship between case detection performance and overall disease incidence. Each data point shows a different country, Size proportional to population and color denoting regional classification, allowing users rapid identification of regional trends and outliers. A linear regression line is added to help the users to

highlight the general direction of the relationship between detection rates and incidence levels. A negative trend would suggest that stronger surveillance and case detection systems are contributing to reduced disease transmission, while a flat or positive trend may imply that, despite improved detection, underlying issues such as reinfection, healthcare accessibility, or drug resistance continue to hinder progress. With this analysis I got a clearer understanding of how national tuberculosis control strategies're different in effectiveness across regions.

4. Dashboard Overview *(Jennifer)*

The dashboard provides a unified platform for exploring the global burden of TB, integrating geographic, temporal, and analytical views to support public health analysis. The dashboard displays the map and top-country chart at the top to present the "what" and "where," while the trend and mortality graphs in the center address "how rates have changed." The bottom row focuses on "why"—through detection rates, policy simulations, and forecasts. Filters for *Region* and *Year* enable comparative analysis across countries and time periods. Users can select regions or adjust the policy slider to update all visualizations dynamically. Hovering reveals tooltips with key statistics, including TB rate, population, and case counts. Highlighting links in the map and scatter plot for direct country-level analysis.

The dashboard combines various viewpoints into a single analytical picture, exposing trends, causes, and possible future developments pertaining to tuberculosis. A Fibonacci-based pattern is used in the design to gradually shift focus from a broad overview to in-depth projections. Both professional analysis and successful public communication are made easier by this planned interaction.

5. Summary and Conclusions

From 1990 to 2013, global TB incidence rates per 100,000 generally trend downward, but declines are uneven across regions. Countries with the highest rates are not always those with the highest absolute case counts, which is why we present both views. Mortality (excluding HIV) may fall in rate terms while staying high in absolute terms in populous settings. HIV-associated TB remains a critical dimension where available, underscoring integrated programs. Our forecast suggests continued improvement if historical patterns persist, while the what-if analysis shows how policy-relevant reductions (e.g., 10–20%) could substantially lower incidence in selected countries.

Limitations include under-reporting, between-country differences in surveillance, and the dataset's endpoint at 2013. We address this by (a) visualizing uncertainty bounds, (b) labeling forecasts as illustrative, and (c) focusing conclusions on the historical window while using scenarios to discuss potential futures. Overall, the data and visuals support a clear story: progress is real yet uneven, and targeted investments in high-burden geographies can accelerate declines.

6. Contributions

Our group collaborated to divide the analytical and visualization work across team members. **Sangeeta** created **Graphs 1, 2, and 3**, which included the global TB incidence map, the Top 10 countries comparison, and the regional trend analysis over time. **Linyu** developed Chart 4 and Chart 5, focusing on the mortality vs. death count dual-axis chart and the projected TB incidence forecast. **Sevgi** created **Graphs 6 and 7**, including the parameter-driven What-If analysis and the scatter plot examining case detection rate vs. incidence. **Jennifer** developed the overall project idea and built the integrated dashboard, to ensure the visualizations told a clear and cohesive story. All group members participated in discussion, interpretation of results, and review of the final report and dashboard.