RStudio

Shweta Ajay Shinde
Masters in Data Analytics, San Jose State University
Data 230: Data Visualization
Instructor: Venkata Duvvuri
November 13, 2024

RStudio

Analysis of TB Deaths Across Regions and Years

Executive Summary

This report analyzes regional trends in tuberculosis (TB) in various areas over time. Visualizations such as charts illustrate the impact of these diseases on different populations. Key findings reveal significant regional differences and trends in TB and HIV-related mortality. These insights are essential for targeting public health interventions and resource allocation.

Introduction

This report looks at how tuberculosis (TB) and deaths from TB vary across different regions over time. We used charts to make the data easier to understand and to spot patterns and trends. The analysis was conducted using **RStudio**, which allowed us to process and visualize large datasets efficiently. By analyzing this information, we can see how TB affect different populations. These findings are important for improving public health strategies and making sure resources are used where they are needed most.

Methodology

Data was extracted from the TB_Burden_Country.csv file, which includes detailed records on TB and HIV-related deaths. The analysis focused on the following columns like:
Country or Territory, Region,
Estimated Total Population,
Method for Prevalence Estimates etc.
This data was used to identify trends and regional impacts.

Summary of Visualization

Line Plot: Time Series of TB Prevalence Over Time by Country

Description:

This line chart visualizes the trend of tuberculosis (TB) prevalence over time for selected countries. The chart allows users to choose a specific country and observe the TB prevalence per 100,000 population across multiple years. This visualization helps in understanding the historical patterns of TB prevalence in each country and evaluating the effectiveness of public health interventions.

Method:

- The dataset containing TB prevalence rates by country and year was loaded into a Shiny application.
- A selectInput dropdown was implemented to allow country selection.
- A line plot was generated using ggplot2, showing TB prevalence over time for the selected country. For code, please click here.

Insights:

- Figure 1 (Brazil): The line plot for Brazil shows a steady decline in TB prevalence from the early 1990s to around 2010, indicating progress in TB control measures.
- Figure 2 (India): For India, the TB prevalence remained constant from 1990 to 1995, after which it steadily decreased until around 2010, suggesting positive effects from health interventions during this period.

RStudio

- Figure 3 (Andorra): For Andorra, the trend does not follow a consistent pattern, with fluctuations in TB prevalence over time. This highlights potential gaps or challenges in public health strategies.
- Overall, the downward trend in several countries suggests significant improvements in TB control and healthcare efforts over the years.
- Monitoring TB trends by country helps identify regions with high prevalence and assess the impact of health policies and programs over time.

Figure 1:
Brazil

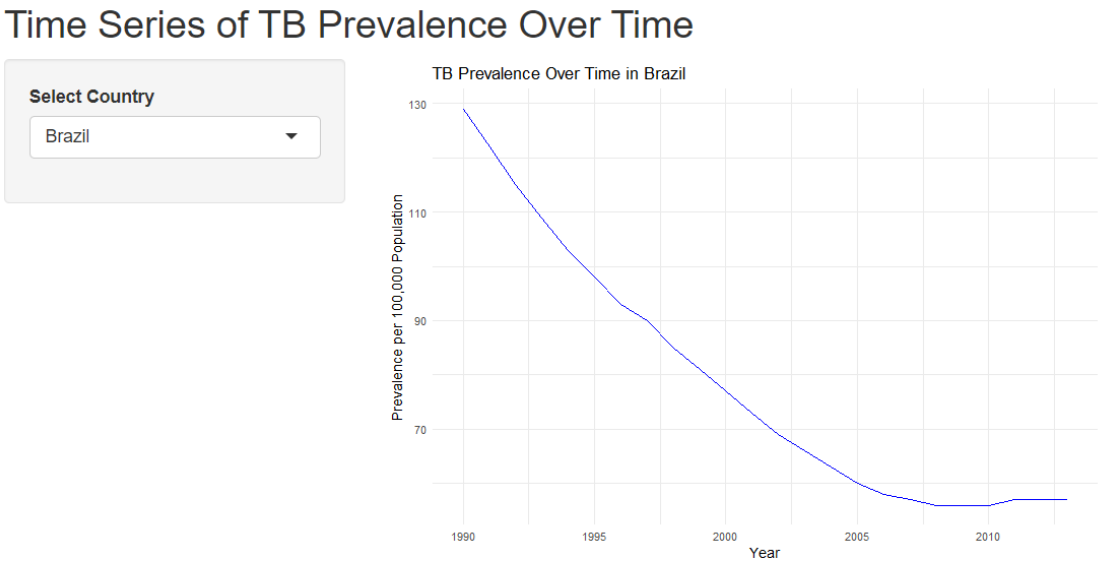
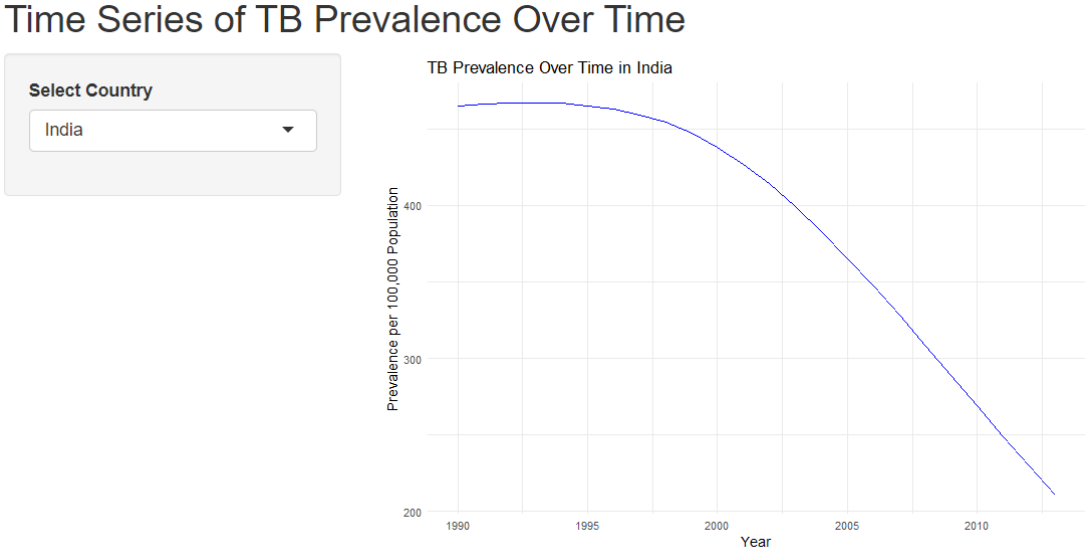
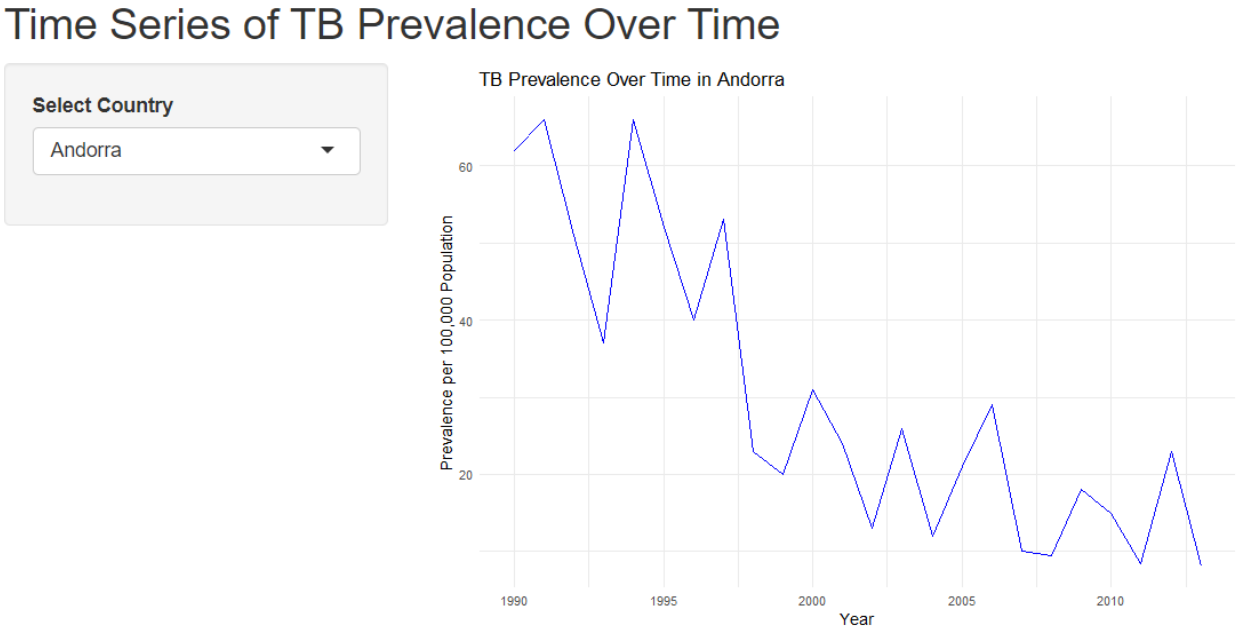


Figure 2:
India



RStudio

Figure 3:
Andorra



2. Scatter Plot: TB Incidence vs. Mortality Rate
Description:

This scatter plot visualizes the relationship between tuberculosis (TB) incidence and mortality rate for a specific year. Users can select a year from the dropdown menu to explore how TB incidence (per 100,000 people) correlates with the TB mortality rate (per 100,000 people) across different countries. This chart provides insights into the severity of TB's impact.

Method:

- The dataset containing TB incidence and mortality data by country and year was loaded into a Shiny application.
- A selectInput dropdown allowed users to choose a specific year for analysis.
- A scatter plot was generated using ggplot2, plotting TB incidence (x-axis) against TB mortality rate (y-axis) for the selected year. For code, please click here.

Insights:

- Figure 4 (1995):The plot indicates that the TB incidence and mortality was higher around 1250 per 1000,000.
- Figure 5,6 (2010-2013): The plot indicates that the TB incidence and mortality rate decreased over time to 75 per 1000,000.
- Overall, the downward trend throughout the years suggests significant improvements in TB control and healthcare efforts over the years.

RStudio

Figure 4:
1995

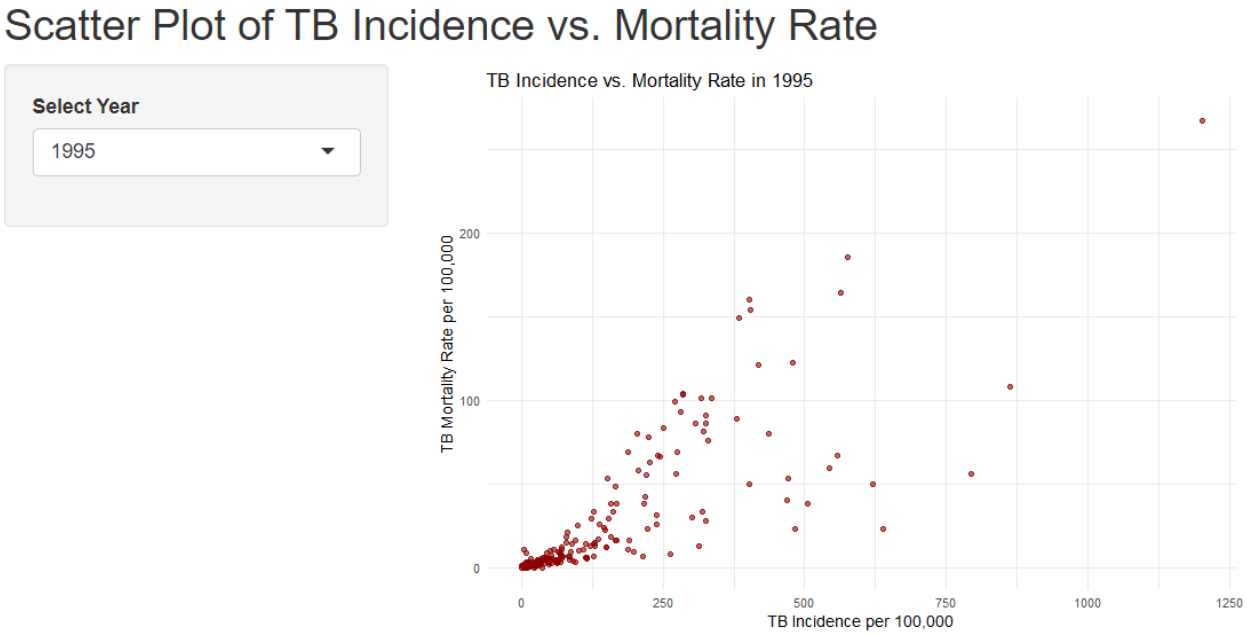
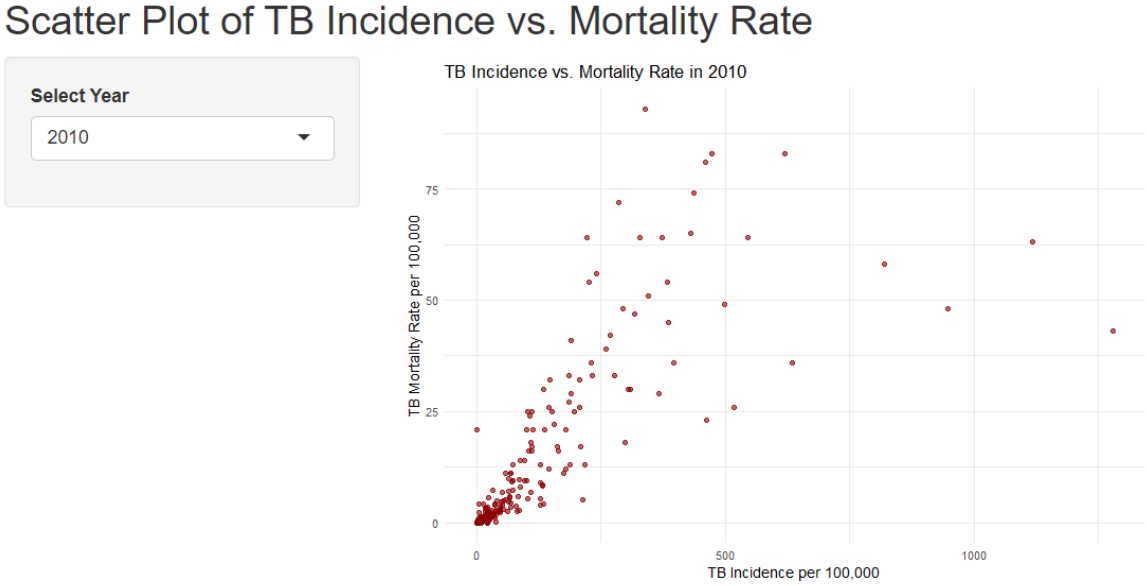
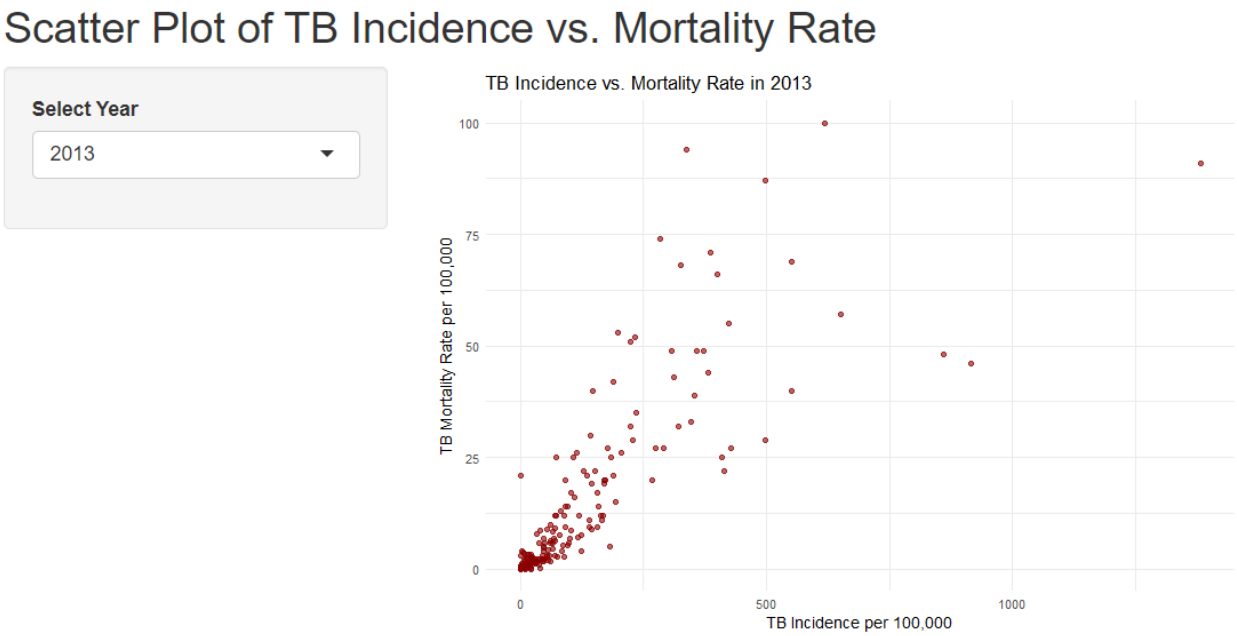


Figure 5:
2010



RStudio

Figure 6:
2013



3.Geospatial Map: TB Prevalence by Country
Description:

The geospatial map visualizes the TB prevalence across different countries for the most recent year available. The map uses color coding to represent the prevalence of TB (all forms) per 100,000 population, making it easy to identify countries with high and low TB burden. Users can interact with the map to explore how TB prevalence is distributed globally, providing valuable insights into the geographical areas where TB is most prevalent.

TB Prevalence: The number of TB cases per 100,000 people, which helps to assess the extent of TB infection within each country. The map provides a clear and interactive way to explore the global burden of TB, highlighting the regions most affected and those that have managed to control TB effectively.

Method:

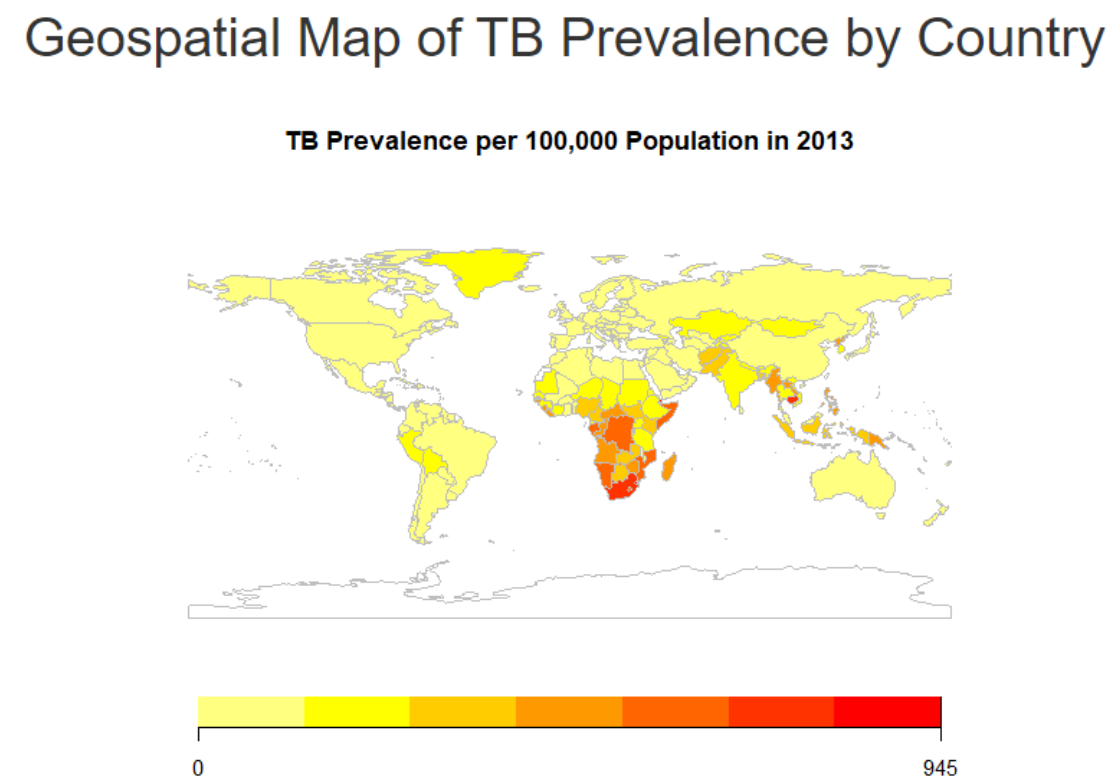
- The dataset containing TB prevalence data by country and year is loaded into the Shiny app.
- joinCountryData2Map() function from the rworldmap package is used to join the TB data with the world map using the ISO3 country codes.
- The map is plotted with the TB prevalence column mapped to the color scale.
- The color palette used is heat (or another color palette of choice), which visually indicates higher and lower prevalence regions.
- The Shiny app uses plotOutput to render the map in the UI. Users can explore the map interactively by hovering over different countries to view their respective TB prevalence. For code, please click here.

RStudio

Insights:

- Africa: Countries like South Africa, Nigeria, and Mozambique show high TB prevalence rates, which are compounded by the ongoing HIV epidemic in these regions. These countries require more intensive TB control measures.(Refer Figure 7)
- Asia: Countries like India and China also show TB prevalence, although their healthcare systems are rapidly improving, leading to some decrease in the overall burden.
- Countries like Germany, Canada, and Australia show significantly lower TB prevalence, indicating successful TB control programs, effective healthcare infrastructure, and early detection mechanisms. These countries often have stronger health systems, better access to diagnostic tools, and public health initiatives that reduce TB transmission.

Figure 7:
Geo Graph



Conclusion:

The combination of all the graphs provides a comprehensive understanding of the global TB burden. The scatter plot showing TB incidence vs. mortality rate reveals that both TB cases and related deaths have decreased over time, especially from 1995 to 2013. This downward trend suggests that global TB control efforts, such as improved diagnosis, treatment, and healthcare access, have been effective. The geospatial map of TB prevalence shows how TB is distributed across the world. High-prevalence regions like Sub-Saharan Africa and Southeast Asia still face significant challenges due to factors like poverty and HIV co-infection, while countries with strong healthcare systems, such as Germany and Canada, show much lower TB rates. Overall, these visualizations highlight the progress made in TB control and the areas where more work is needed to further reduce the global TB burden.

References

Prof. Venkata Duvvuri (2024). RStudio. *San Jose State University*. https://sjsu.instructure.com/courses/1595119/pages/week-14-ppt?module_item_id=15676005

R Coder. (n.d.). R graphs: How to create a graph in R. *R Coder*. https://r-coder.com/r-graphs/