Interactive Data Visualization Dashboard for TB Using D3

1. Project Objective

To create an interactive dashboard visualizing Tuberculosis (TB) relationships across multiple dimensions (geography, time, and type) using D3.js. The goal is to uncover patterns, provide insightful analysis, and facilitate exploration through dynamic visualizations.

---

2. Dataset Details

The dataset updated_tb_dataset.csv contains the following key attributes:

- ***latitude*** and ***longitude***: Geographic coordinates for plotting.
- ***year***: Indicates the year of the recorded event.
- ***country***: Location of the recorded event.
- ***Relationship Type***: Type of relationship between entities.
- ***Path***: Descriptive information about relationships.
- ***TB Cause***: Primary causes of TB spread.
- ***People Affected*** and ***Deaths***: Metrics for impact.
- ***Deaths Due to Cause Treatment Available***: Availability of treatments for deaths due to specific causes.

---

3. Visualizations

3.1. Force-Directed Graph

The Force-Directed Graph, implemented in the file *Forced.html*, provides a clear and dynamic visualization of the relationships between Tuberculosis (TB) causes and affected countries. This graph is a crucial tool for mapping connections and understanding the complexities of TB spread and its impact across regions. The dataset utilized for this visualization includes key attributes such as *country*, *TB Cause*, *People Affected*, and *year*.

3.1.1. Data Preparation and Grouping

To prepare for visualization, the data was grouped by country and TB causes. This grouping involved aggregating critical metrics like the number of people affected, ensuring the dataset was structured and ready for meaningful representation. By organizing the data in this way, relationships between entities could be effectively depicted in the graph layout.

3.1.2. Node Representation

In the graph, nodes are used to represent entities—either countries or TB causes. A distinct color-coding scheme differentiates these entities: blue nodes signify countries, while green nodes represent causes. The size of each node dynamically corresponds to the "People Affected" metric, with larger nodes indicating higher values. This feature allows users to easily identify significant countries or causes at a glance.

3.1.3. Edge Representation

Edges between the nodes symbolize relationships between countries and causes. The thickness of each edge varies proportionally to the number of people affected, visually emphasizing the strength and intensity of the connection. This approach ensures that users can quickly identify the most impactful relationships in the dataset.

3.1.4. Interactivity Features

The graph is designed to be interactive, enhancing user engagement and exploration. Hovering over nodes or edges displays detailed information through tooltips, such as the country, cause, and number of people affected. Clicking on a node zooms in for a closer view and reveals detailed insights in a side panel. Filters are also integrated, enabling users to adjust the graph dynamically based on criteria such as year, country, or cause. This functionality makes the graph highly adaptable to specific analysis needs.

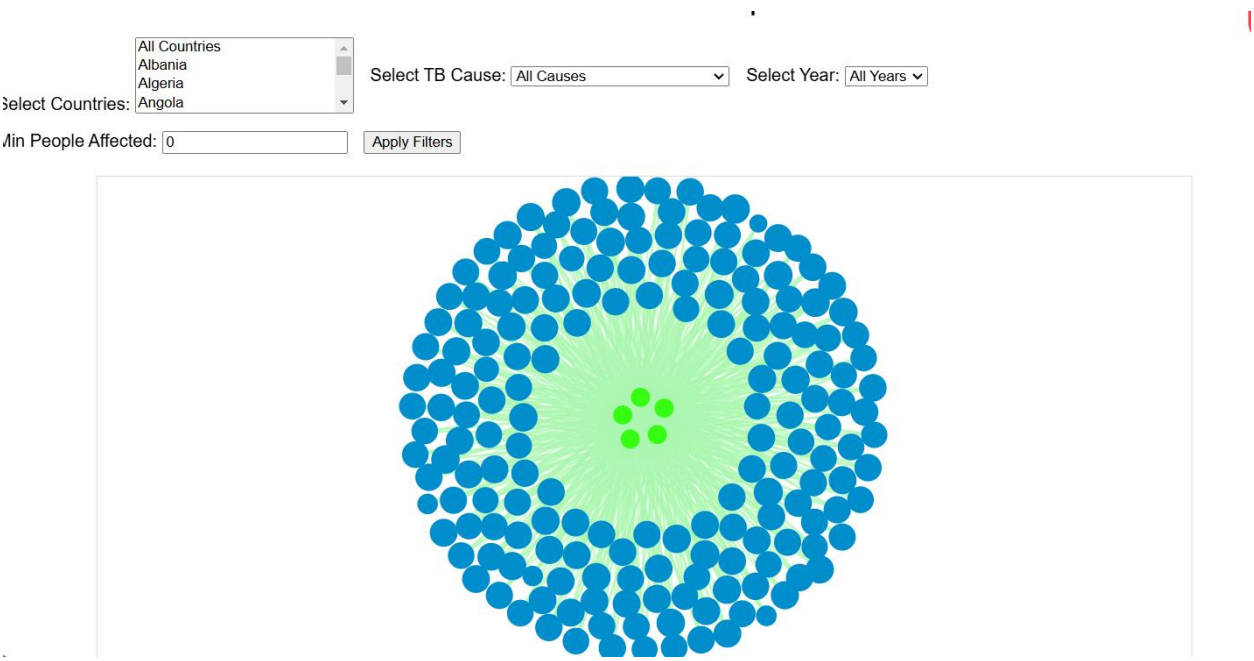
3.1.5. D3.js Implementation

The Force-Directed Graph leverages several powerful D3.js features. The *forceSimulation* API enables a responsive and dynamic layout, allowing nodes and edges to adjust naturally as data changes. The *d3.zoom* module facilitates smooth zooming and panning across the graph. Additionally, dynamic tooltips and a detailed side panel provide immediate, context-sensitive insights, making the graph both informative and user-friendly.

3.1.6. Conclusion

This Force-Directed Graph serves as an effective tool for visualizing TB relationships globally. By combining insightful data representation with robust interactivity, it offers a visually engaging way to explore and understand the complex dynamics of TB spread.

File Reference: Forced.html Details.



3.2. Geographic Map Chart Overview

The Geographic Map Chart, implemented in the file *Map.html*, provides an intuitive visualization of the geographical distribution of Tuberculosis (TB) cases. This visualization aims to help users analyze the spread of TB across various regions, emphasizing key metrics such as deaths, people affected, and the underlying causes of the disease. The dataset used for this chart includes critical attributes like *latitude*, *longitude*, *TB Cause*, *Deaths*, *People Affected*, *country*, and *year*.

3.2.1. Map Integration

This visualization is built using an integration of Leaflet.js and D3.js, combining the mapping capabilities of Leaflet with the flexibility of D3 for data-driven visual elements. Leaflet serves as the foundation for rendering an interactive world map, while D3 handles data processing and dynamic overlays, creating a seamless user experience.

3.2.2. Marker Placement

Markers on the map represent individual data points, such as countries or regions affected by TB. These markers are displayed either as circles or pins, depending on user preference. Each marker is color-coded to correspond with specific TB causes, enhancing visual differentiation. To improve clarity, marker opacity is adjustable, allowing users to focus on prominent causes or densely populated areas without visual clutter.

3.2.3. Interactive Filters

The Geographic Map Chart includes several interactive filters to enhance exploration and analysis:

- Dropdown Menus: Users can filter data by region or specific TB causes using dropdown menus.
- Year Range Slider: A slider allows users to focus on data from a specific time frame, enabling trend analysis across different periods.
- Death Count Threshold Filter: This feature helps users isolate regions or causes based on the minimum number of deaths recorded, highlighting the most severe cases.

These filters ensure that users can tailor the visualization to their specific analytical needs, making it highly adaptable and insightful.

3.2.4. Legend

A dynamic legend accompanies the map, providing a clear reference for the color-coded TB causes. Users can interact with the legend to toggle the visibility of specific causes directly on the map, simplifying the exploration of complex datasets.

3.2.5. D3.js Implementation

D3.js plays a crucial role in the Geographic Map Chart by enabling efficient data processing and interactivity. The *d3.csv* function is used to load and preprocess the dataset. Event listeners handle user interactions such as hovering and clicking on markers, displaying detailed information in tooltips. Additionally, D3's SVG capabilities are employed to create customized markers that align seamlessly with the map interface.

3.2.6. Conclusion

The Geographic Map Chart combines the strengths of Leaflet.js and D3.js to deliver a powerful and interactive visualization of TB cases worldwide. Its dynamic filters, customizable markers, and informative legend make it an essential tool for exploring the geographical dimensions of TB impact.

File Reference: Map.html Details.

3.3. Hierarchical Tree Map Overview

The Hierarchical Tree Map, implemented in the file *Treemap.html*, visualizes hierarchical relationships within Tuberculosis (TB) treatment data. This chart is designed to help users explore the availability and impact of treatments across different years and regions. By structuring data hierarchically, it provides a comprehensive view of trends and disparities in treatment-related outcomes. The dataset used for this visualization includes attributes such as *Treatment Available*, *year*, and *Deaths*.

3.3.1. Hierarchy Construction

The tree map represents data hierarchically by grouping it according to treatment availability (e.g., "Yes," "No," or "Limited") and year. Each rectangle in the tree map corresponds to a data group, with its size proportionate to the number of deaths in that category. This layout makes it easy to compare the magnitude of TB-related deaths across different treatments and time periods. Users can instantly identify which treatment categories have the most significant impact based on the size of the rectangles.

3.3.2. Breadcrumb Navigation

To facilitate deeper exploration, the tree map includes breadcrumb navigation. Users can drill down into specific hierarchical levels, such as narrowing the view to a single treatment type or year. Breadcrumbs provide a clear path for users to navigate back to the root level, ensuring they do not lose context while exploring complex data hierarchies. This feature enhances usability and promotes seamless exploration.

3.3.3. Color Coding

Distinct colors are assigned to each treatment availability category, such as bright green for "Yes," bright red for "No," and bright blue for "Limited." This color-coding enables quick visual differentiation, helping users identify trends and disparities at a glance. The consistent use of colors across the tree map ensures clarity and aids in interpreting the data efficiently.

3.3.4. Filters

The tree map is equipped with multiple filters to allow users to customize the visualization:

- A dropdown menu enables filtering data by specific countries.
- Checkboxes let users select or deselect treatment availability types.
- A slider allows users to adjust the year range, focusing on data from specific time periods.

These filters make the visualization highly interactive and adaptable to various analytical needs.

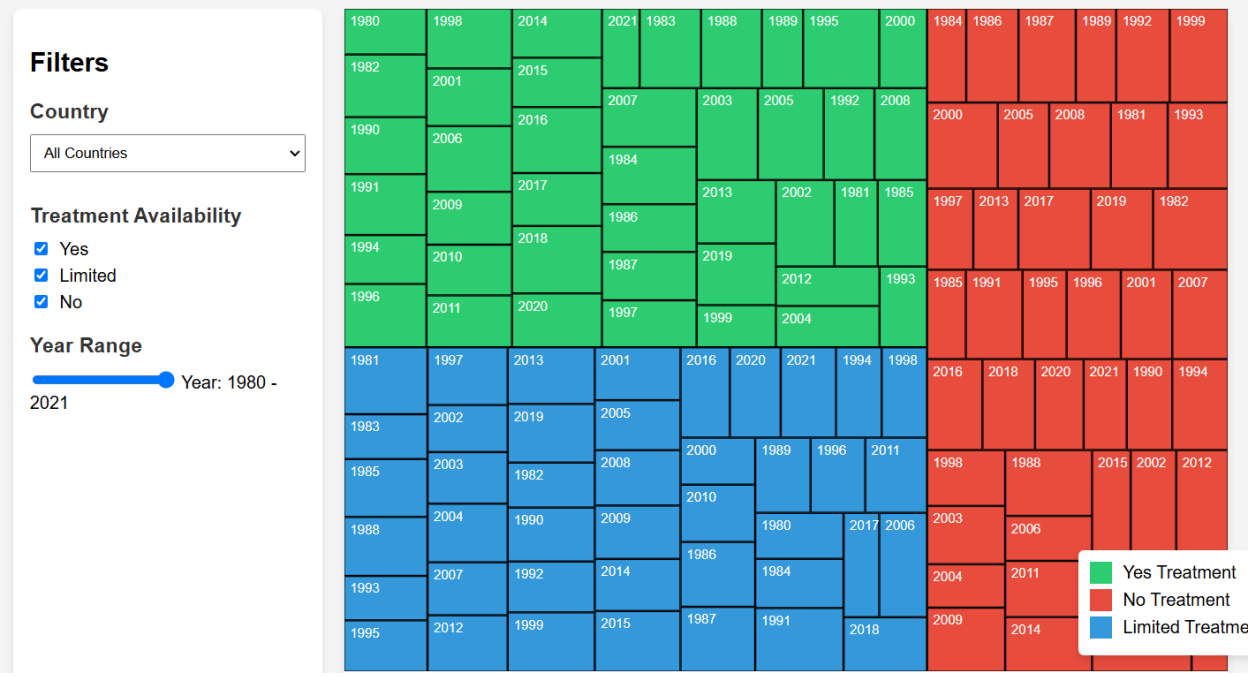
3.3.5. D3.js Implementation

The tree map leverages D3.js for its structure and interactivity. The *d3.hierarchy* function constructs the hierarchical data structure, while *d3.treemap* renders the tree map layout dynamically based on the data. Tooltips provide additional details on hover, displaying metrics like treatment availability, year, and deaths. Breadcrumb navigation is powered by dynamic interactivity, ensuring users can drill down or return to previous levels smoothly.

3.3.6. Conclusion

The Hierarchical Tree Map is a powerful tool for analyzing TB treatment data. Its ability to represent data hierarchically, combined with interactivity and intuitive navigation, makes it an excellent resource for identifying trends, comparing treatment outcomes, and exploring relationships in TB-related metrics.

File Reference: Treemap.html Details.



3.4. Timeline Visualization Overview

The Timeline Visualization, implemented in the file *tl.html*, provides an animated view of the evolution of Tuberculosis (TB) cases over time. This interactive chart allows users to analyze the trends of TB cases, deaths, and other relevant metrics across different countries and years. The dataset used for this visualization includes *year*, *country*, *People Affected*, and *Deaths*, which are the primary metrics displayed in the timeline.

3.4.1. Line Graph Creation

At the core of the timeline visualization is a line graph that connects data points for each country over time. The X-axis represents the years, and the Y-axis represents the selected metric (either "People Affected" or "Deaths"). Each data point corresponds to a specific year and country, and the line connects these points, showing how TB cases or deaths have evolved over time. The line graph provides an easy way to observe trends and fluctuations in the data.

3.4.2. Animation Features

One of the most engaging aspects of the Timeline Visualization is its animation. Users can interact with the chart using a slider or play/pause button to animate the data over the years. The animation progresses year by year, allowing users to observe changes in TB metrics in a dynamic and fluid manner. Additionally, a toggle switch is included, which lets users switch between the "People Affected" and "Deaths" metrics. This feature ensures that users can explore the data from multiple perspectives, depending on the metric they are most interested in.

3.4.3. Tooltip Functionality

To enhance the interactivity, the timeline visualization includes a tooltip feature. When users hover over a data point (i.e., a circle representing a country for a given year), a tooltip appears, displaying detailed information about the selected country, the year, and the value of the chosen metric. This feature helps users quickly access more granular data without cluttering the visualization. It provides important contextual information that is essential for in-depth analysis.

3.4.4. Filters

The Timeline Visualization also incorporates several filtering options to enhance user experience and customization. A dropdown menu allows users to select specific countries, enabling them to focus on particular regions or compare multiple countries. The year slider allows users to filter the data for a specific time range, and the metric toggle enables users to switch between "People Affected" and "Deaths" for deeper insights.

3.4.5. D3.js Features

The Timeline Visualization utilizes several features of D3.js to create a dynamic and interactive experience. The *d3.line* function is used to generate the line paths connecting data points. The *d3.scaleLinear* function creates the linear scales for both the X and Y axes, ensuring that the graph is properly scaled and aligned. The *d3.axis* module is used to generate the X and Y axes dynamically, and the transition functions allow for smooth animation between different years and metrics. The tooltip is implemented using D3.js's event listeners to display context-sensitive data when users interact with the graph.

3.4.6. Conclusion

The Timeline Visualization provides an interactive and animated way to explore the evolution of Tuberculosis cases and deaths over time. By offering dynamic filtering, smooth transitions, and real-time updates, the visualization allows users to engage with the data and gain valuable insights. The combination of interactivity and detailed data presentation makes this tool a powerful resource for analyzing TB trends across various countries and years.

File Reference: tl.html Details.

3.5. Sunburst Chart for TB Causes and Treatment Availability

The Sunburst Chart, implemented in the file *sun.html*, provides an interactive and hierarchical visualization of Tuberculosis (TB) causes and treatment availability. This chart allows users to explore the relationships between different TB causes, the availability of treatments, and the number of people affected by these causes across various countries. The dataset used for this chart includes attributes such as *TB Cause*, *Treatment Available*, *People Affected*, and *Country*.

3.5.1. Chart Structure and Data Grouping

The Sunburst Chart is structured hierarchically, with the root representing the overall TB data, and child nodes representing different TB causes. Each TB cause is further subdivided based on the availability of treatment (e.g., "Treatment Available," "No Treatment," or "Treatment Limited"). The size of each segment in the chart corresponds to the number of people affected by TB within that category. To construct this hierarchy, the data is grouped by TB cause and treatment availability, summing up the number of affected individuals for each category. This grouping enables the chart to provide a clear overview of the global TB landscape in terms of both causes and the availability of treatment.

3.5.2. Interactivity and Tooltip Features

The Sunburst Chart is highly interactive, offering several ways to engage with the data:

- **Zooming and Panning**: Users can zoom in and out to explore specific subcategories in more detail. This zooming feature allows users to drill down into particular TB causes and treatment types.
- **Mouseover Effects**: When users hover over a segment, a tooltip displays additional information, including the category name and the number of people affected. This tooltip enhances the user experience by providing detailed, context-sensitive data without cluttering the visualization.
- **Segment Highlighting**: As users hover over or click on a segment, the opacity of that segment changes, making it easier to highlight specific parts of the chart and examine them more closely.

3.5.3. Color Coding and Aesthetic Features

Each level of the Sunburst Chart is color-coded for clarity. The root layer (representing the overall TB data) is white, while the first level of child nodes (representing TB causes) is assigned distinct colors. These colors are chosen to be easily distinguishable from one another, aiding users in differentiating between various TB causes. The second level, which represents treatment availability, inherits colors from its parent node but is made slightly lighter to indicate its relationship to the TB cause. This color hierarchy helps users visually follow the relationships between TB causes and treatment availability.

3.5.4. Dynamic Filters and Country Selection

To further enhance user interaction, the Sunburst Chart includes a filter for selecting specific countries. The country dropdown menu dynamically populates based on the available data, and users can select a country to view the TB causes and treatment availability data specific to that region. When a country is selected, the chart updates in real-time to reflect the filtered data, providing a customized view of the TB situation in the chosen country. The chart also updates if no country is selected, displaying global data instead.

3.5.5. D3.js Implementation

D3.js is used extensively throughout the Sunburst Chart to handle data processing, visualization, and interactivity:

- **d3.hierarchy**: This function is used to create the hierarchical structure needed for the Sunburst Chart, enabling easy grouping of data by TB cause and treatment availability.
- **d3.partition**: This layout is used to calculate the positioning of each segment within the Sunburst Chart, ensuring that each node is correctly placed based on its size and hierarchy.
- **d3.arc**: The arc generator creates the actual shapes for each segment, ensuring that each slice in the Sunburst Chart has the correct size and position.
- **d3.zoom**: This feature enables zooming and panning for exploring the chart in detail.

- **d3.select** and **d3.on**: These methods handle user interactions, including filtering, tooltips, and segment highlighting.

3.5.6. Conclusion

The Sunburst Chart for TB Causes and Treatment Availability is a powerful tool for exploring the relationships between TB causes, treatment availability, and the impact of TB across different regions. Its interactive features, such as zooming, panning, tooltips, and filtering, provide users with an engaging and insightful way to analyze the global and regional TB landscape. By combining hierarchical data representation with dynamic interactivity, this visualization offers a comprehensive overview of the state of TB treatment and the causes contributing to its spread.

File Reference: sun.html Details.

3.6. Conclusion

The Tuberculosis (TB) Dashboard, featuring a variety of visualizations such as the Force-Directed Graph, Geographic Map, Timeline Visualization, and Sunburst Chart, offers an interactive and comprehensive tool for exploring TB data across multiple dimensions. By leveraging D3.js, the dashboard allows users to analyze the relationships between TB causes, treatment availability, geographic distribution, and the evolution of TB cases over time. The dynamic filtering, zooming, panning, and tooltip functionalities

enhance user engagement, enabling tailored data exploration based on specific countries, years, and metrics like "People Affected" or "Deaths."

Together, these visualizations provide a clear and insightful view of global and regional TB patterns, highlighting key trends, disparities, and the effectiveness of treatment measures. Whether used for detailed trend analysis, geographic exploration, or understanding the impact of different TB causes, this dashboard serves as a powerful resource for researchers, policymakers, and healthcare professionals working to combat TB worldwide. The combination of interactivity, intuitive design, and robust data representation ensures that users can make informed decisions and gain a deeper understanding of the challenges posed by Tuberculosis.