**Mujahid Abbas 22i-1969**

**Danish Shah 22i-1860**

**Usama Bin Sajid 19i-1864**

**Comprehensive Documentation for TB Data Dashboard**

**Overview**

The TB Data Dashboard provides a rich, interactive visual exploration of tuberculosis (TB) data, leveraging modern web technologies such as HTML, CSS, and D3.js. The dashboard integrates multiple visualizations, including a force-directed graph, to help users understand and analyze data trends effectively.

**1. Dashboard Design Choices**

**1.1. Layout**

* **Responsive Grid Layout**: A grid system organizes visualizations into a two-column layout for wider screens. On smaller devices, visualizations stack vertically to enhance readability.
* **Containerization**: Each visualization is placed inside a dedicated container for separation, styled with rounded corners, shadows, and a white background for contrast.

**1.2. Aesthetics**

* **Dynamic Animated Background**:
  + Aesthetic appeal is enhanced with animated circles created using SVG and D3.js.
  + Circles randomly transition positions, sizes, and opacity, providing a subtle motion background.
* **Color Scheme**:
  + A dark blue header and footer frame the content, paired with light and neutral tones for the background and text.

**1.3. Interactive Features**

* **Visualization Integration**:
  + Multiple visualizations are embedded using <iframe> elements, each representing distinct aspects of TB data.
  + Includes a force-directed graph, hierarchical tree map, sunburst chart, timeline visualization, and a global map.
* **Interactive Controls**:
  + For the force-directed graph, users can filter nodes by group, reset focus, and explore node details via tooltips.

**2. Force-Directed Graph Design**

**2.1. Data Visualization**

* **Nodes and Links**:
  + Nodes represent entities (e.g., countries or regions) and are sized uniformly.
  + Links connect nodes to the central "TB Cases" node, with link width proportional to TB case values.
* **Color Coding**:
  + A D3 ordinal scale (d3.scaleOrdinal(d3.schemeCategory10)) assigns colors to groups for quick differentiation.

**2.2. Interaction**

* **Zoom and Pan**:
  + Users can zoom in and out or pan across the graph for a better view.
* **Tooltip**:
  + On mouse hover, a tooltip displays details about a node, such as:
    - **Node Identifier**: Country or region name.
    - **Value**: Numeric TB case data.
    - **Group**: Node classification (e.g., region or metric).
    - **Timestamp**: Relevant date information.
* **Click to Focus**:
  + Clicking a node dims unconnected nodes and links, allowing users to focus on specific relationships.
* **Legend**:
  + The legend provides an overview of group colors and toggles node visibility by group.

**3. Data Preprocessing**

**3.1. Input Data**

* **Source File**: tb\_preprocessfile\_2.csv
* **Columns**:
  + country: Name of the country.
  + new.sp.cur: Numeric TB case data for the country.
  + region: Group classification, e.g., region or custom grouping.
  + date: Date of the data record.

**3.2. Transformation**

* **Node Data**:
  + Each country is mapped as a node with attributes:
    - id: Country name.
    - value: TB case data from new.sp.cur.
    - group: Region or classification.
    - timestamp: Derived from date.
* **Links**:
  + Each country node connects to the central "TB Cases" node with a link attribute:
    - source: Country name.
    - target: "TB Cases".
    - value: TB case data.

**4. Key Features of Force-Directed Graph Code**

**4.1. Core D3.js Concepts**

* **Force Simulation**:
  + The graph uses d3.forceSimulation to position nodes dynamically.
  + Forces include:
    - d3.forceLink: Connects nodes with links, setting link distances.
    - d3.forceManyBody: Creates repulsion between nodes.
    - d3.forceCenter: Centers the graph within the SVG container.

**4.2. Dynamic Styling**

* **Node Styling**:
  + Nodes are styled with color from the colorScale based on their group.
  + Hover interactions highlight nodes and display tooltips.
* **Link Styling**:
  + Link widths scale proportionally to their values using Math.sqrt(d.value).
  + Links are semi-transparent for better readability.

**4.3. User Interactions**

* **Drag Behavior**:
  + Nodes can be dragged, updating their positions dynamically within the force simulation.
* **Double-Click Reset**:
  + Double-clicking resets the graph view, restoring all nodes and links to full opacity.

**4.4. Filtering and Legends**

* The legend allows users to filter nodes by group. Clicking a group in the legend toggles its visibility.

**5. Supporting Files**

**5.1. HTML and CSS**

* force.html: Provides the structural layout for the force-directed graph visualization.
* force.css: Styles the visualization and its components, such as the tooltip and legend.

**5.2. JavaScript**

* force.js: Implements the force-directed graph using D3.js, including data processing, simulation setup, and interactivity.

**5.3. Data**

* tb\_preprocessfile\_2.csv: Contains the preprocessed TB case data, used as input for the visualization.

**6. Future Enhancements**

* **Enhanced Filters**:
  + Add multi-criteria filters (e.g., filter by both group and value ranges).
* **Additional Visualizations**:
  + Include trend analyses or correlation heatmaps to provide deeper insights.
* **Dynamic Data Loading**:
  + Fetch updated data from an API endpoint to ensure real-time data visualization.

**3. Map Graph**

**Purpose**

The **Interactive Map of TB Treatment Success Rates** visualizes the geographic distribution of TB treatment success rates across different countries. Users can filter by year or country to focus on specific subsets of data.

**Files**

* **HTML**: map.html
* **JavaScript**: map.js
* **CSS**: map2.css

**Key Features**

1. **Dynamic Filtering**:
   * Dropdown menus for selecting countries and years.
   * Updates the map to reflect filtered data in real-time.
2. **Color Encoding**:
   * Treatment success rates (TSR) are color-coded using a blue gradient.
   * Missing data is represented with a gray fill.
3. **Tooltip**:
   * Displays detailed information about the country, TSR, region, and year on hover.
4. **Zoom and Pan**:
   * Allows users to zoom in and pan around the map for better exploration.
5. **Legend**:
   * A dynamic legend explains the color scale for TSR values.

**Data Preprocessing**

* GeoJSON data for the world map.
* CSV data (tx\_2022-08-29.csv) containing:
  + iso3: ISO country codes for matching with GeoJSON features.
  + c.new.tsr: Treatment success rate values.
  + g\_whoregion: WHO region classifications.
  + year, country: Additional attributes for filtering.

**Implementation Highlights**

* **Projection**: The geoNaturalEarth1 projection ensures a visually appealing and accurate world map.
* **Event Handling**:
  + Mouse events for tooltips.
  + Zoom events using d3.zoom() for smooth interactions.
* **Dynamic Data Binding**:
  + Filters update the GeoJSON properties, dynamically changing map colors.

**4. Sunburst Chart**

**Purpose**

The **Sunburst Chart** provides a hierarchical view of TB metrics (e.g., cases per 100k population or total population) by country and year. It supports metric-based filtering and country selection to enable detailed analysis.

**Files**

* **HTML**: sun.html
* **JavaScript**: sun.js
* **CSS**: sun2.css

**Key Features**

1. **Metric Selection**:
   * Dropdown to select the metric to visualize (e.g., tb\_cases\_per\_100k, pop).
2. **Country Filtering**:
   * Dropdown to filter by a specific country or display all countries.
3. **Zoom and Pan**:
   * Allows users to zoom into specific sections of the sunburst for focused exploration.
4. **Tooltip**:
   * Provides details on hover, including the name, metric value, and population.
5. **Interactive Updates**:
   * Dynamically updates the chart based on selected metric and country.

**Data Preprocessing**

* CSV data (tb\_preprocessfile\_1.csv) containing:
  + tb\_cases\_per\_100k: TB cases per 100k population.
  + pop: Total population.
  + country, year: Hierarchical attributes.

**Implementation Highlights**

* **Hierarchy Creation**:
  + Data is grouped by country and year to form a nested structure suitable for the sunburst layout.
  + The createHierarchy function transforms flat CSV data into a hierarchical format.
* **Partition Layout**:
  + d3.partition() is used to calculate positions for hierarchical data.
* **Arc Generator**:
  + d3.arc() generates arcs for the sunburst slices, handling radial dimensions and angles.
* **Event Handling**:
  + Click events for zooming into specific nodes.
  + Mouseover events for tooltips.

**Common Design Choices**

**Accessibility**

* **Tooltips**:
  + Designed for clarity and ease of use.
  + Ensures that key data points are visible on hover.
* **Color Scales**:
  + High-contrast color palettes for better readability.

**Responsiveness**

* SVG elements scale dynamically with the window size to ensure consistent visualization.

**Data Integration**

* Ensures that real-world data is accurately mapped to visual elements through robust preprocessing and validation.

**Future Enhancements**

1. **Additional Metrics**:
   * Include more granular metrics, such as age-specific or region-specific data.
2. **Animations**:
   * Smooth transitions for map updates and sunburst changes.
3. **Export Options**:
   * Allow users to export visualizations as PNG or PDF.
4. **Enhanced Filtering**:
   * Multi-select filters for year and country.

**5. Timeline Graph**

**Purpose**

The timeline graph provides an interactive visualization of TB treatment metrics over time. Users can explore the data by filtering countries and metrics to analyze trends.

**Design Choices**

**Visualization Type**

* A **line chart** was chosen for its effectiveness in showing temporal trends.
* Data points are marked with **circles**, allowing users to identify specific values.

**Interactivity**

1. **Zoom and Pan:**
   * Users can zoom in and out to focus on specific periods.
   * Panning allows exploration of the entire timeline without resetting filters.
2. **Filters:**
   * **Country Filter:** Dropdown allows users to select a specific country or view data for all countries.
   * **Metric Filter:** Dropdown lets users switch between different metrics, such as:
     + Treatment Success Rate (TSR)
     + Current Smear-Positive Cases
     + Deaths
3. **Tooltip:**
   * Displays year and metric value when hovering over a data point.
   * Dynamically updates its position based on cursor movement.
4. **Click Interaction:**
   * Clicking on a data point zooms into the region around that point for a detailed view.

**Aesthetic Choices**

* Line color: **Steelblue** for clear visibility against the white background.
* Tooltip: Styled for readability with a semi-transparent background.
* Axes: X-axis uses integer year values; Y-axis scales dynamically based on the data.

**Data Preprocessing**

**Input File:**

* tb\_preprocessfile\_2.csv

**Steps:**

1. **Parsing:**
   * Convert year and metric values to numerical formats.
2. **Filtering:**
   * Filter data based on user-selected country and metric.
3. **Grouping:**
   * Average the selected metric for each year when viewing data for all countries.
4. **Domain Calculation:**
   * X-axis domain: Range of years in the data.
   * Y-axis domain: Minimum to maximum value of the selected metric.

**Key Features**

* **Dynamic Scales:** Automatically adjust based on the filtered dataset.
* **Responsive Axes:** Update with transitions for a smooth user experience.
* **Click-to-Zoom:** Centers the visualization on the selected data point.

**6. Tree Map Visualization**

**Purpose**

The tree map provides a hierarchical representation of TB-related data, showing the distribution of cases or population across different WHO regions.

**Design Choices**

**Visualization Type**

* A **tree map** was chosen for its ability to represent hierarchical data effectively.
* Each rectangle represents a country, sized based on the selected metric.

**Interactivity**

1. **Zoom and Pan:**
   * Users can zoom into specific regions or countries to explore data at finer granularity.
   * Panning is restricted within the tree map boundaries.
2. **Filters:**
   * **Metric Filter:** Dropdown to select between:
     + TB Cases per 100k
     + Population
3. **Tooltip:**
   * Displays the country name and the selected metric value when hovering over a rectangle.
4. **Click Interaction:**
   * Clicking on a rectangle zooms into the specific node (region or country).

**Aesthetic Choices**

* Colors: **Discrete color scheme** using D3’s schemeCategory10, grouped by WHO regions.
* Text labels: Displayed only if the rectangle is large enough for readability.
* Borders: White stroke to separate rectangles for clarity.

**Data Preprocessing**

**Input File:**

* tb\_preprocessfile\_1.csv

**Steps:**

1. **Parsing:**
   * Convert numeric columns (tb\_cases\_per\_100k, pop) to integers.
2. **Grouping:**
   * Group countries by g\_whoregion for hierarchical structuring.
3. **Hierarchy Creation:**
   * Use d3.hierarchy to represent the data structure.
   * Summarize the selected metric (sum) at each node.
4. **Layout Calculation:**
   * Use d3.treemap to calculate positions and sizes for rectangles.**Key Features**

* **Dynamic Updates:** Tree map layout and coloring adjust automatically when the metric is changed.
* **Hierarchical Zoom:** Enables users to focus on specific regions or countries.
* **Responsive Labels:** Dynamically sized to fit within rectangles.

**General Considerations**

**Performance**

* Transitions and animations are implemented for a smooth user experience.
* Filtering and zooming are optimized to handle large datasets.

**Scalability**

* The designs can accommodate additional metrics or data without significant code changes.

**Accessibility**

* Tooltips provide textual feedback for interactive elements.
* Intuitive controls make the visualizations user-friendly.