

Name : Melita Saldanha
SBU ID : 112551230

Course : CSE 564 – Visualization and Visual Analytics
Mini Project : 1 – World Statistics Visualization

Report

Resources:

- YouTube video link of Demonstration:
<https://youtu.be/xQcC6IWYpSw>
- Link to Individual Dataset Files:
https://drive.google.com/drive/folders/1xjwy416dVxxsNV5zF_SgF8Xh2lwcpm9G?usp=sharing

Data Collection:

The dataset consists of various features about each country in the world. Each feature had an attribute: Country and its value for each year in a certain range.

The following features are being used:

- **Country:** List of all countries.
- **Region:** Region the country belongs to. [1]
- **Year:** The year for which the row contains data.
- **Gender:** Some attributes have gender specific values.
- **Population Density:** Average number of people on each square km of land in the country. [2]
- **Health Expenditure:** Percentage of health expenditure paid by: [3]
 - Government
 - Private Firms
 - Individual
- **Life Expectancy:** Average number of years a newborn child would live. [4]
- **Education Expenditure:** In terms of percentage of Gross National Income (GNI). [5]
- **Mean years in school:** Number of years of school attended by people in age group 15-24. [6]
- **Average Income:** Adjusted National Annual Income. [7]
- **Employment Rate:** Percentage of population of the country employed during the given year. [8]
- **Income Inequality:** Gini Coefficient showing Income Inequality in a society. [9]
- **Child Mortality Rate:** Death of children under 5 years as per 1000 live births. [10]
- **Adult Mortality Rate:** Death of adults 1000 live adults. [11]
- **Suicides:** Mortality due to self-afflicted injury as per 100,000 standard population. [12]
- **Murders:** Mortality due to inter-personal violence as per 100,000 standard population. [13]
- **Military Expenditures:** In terms of percentage of GDP. [14]
- **Corruption Perception:** Transparency International's Score of perceptions of corruption. [15]

Data Processing:

1. **Cleaning Individual Datasets:** Missing values of each dataset was handled. Since the data is a time-dependent data, forward fill and back fill methods were used to propagate non-null values forward or backward.
2. **Merging All Datasets:** Each row representing a country was replicated to handle the data for different years. Similarly, each row representing a year was replicated to handle gender specific data. For example:

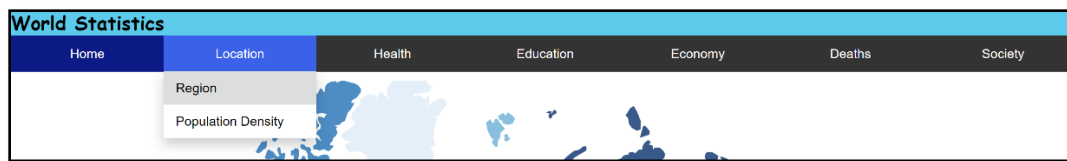
Country	Year	Gender	Population_Density	Life_Expectancy
United States	1950	Male	17.4	65.5
United States	1950	Female	17.4	71.2
United States	1951	Male	17.6	65.6
United States	1951	Female	17.6	71.4

Code Implementation:

The entire page has been design using HTML, CSS, JavaScript, JQuery and D3.js. Following are its features:

1. Menu to choose attribute and update chart

The top navigation bar is used to choose the attribute that is to be displayed. The attributes have been divided into categories to be easily accessible.



Menu Bar

2. Bar chart for Categorical Attributes

If categorical variables are selected, a bar graph is displayed.

```
xAxis = d3.svg.axis()
    .scale(x)
    .orient("bottom");

yAxis = d3.svg.axis()
    .scale(y)
    .orient("left")
    .ticks(10);

svg = d3.select("body").append("svg")
    .attr("width", width + margin.left + margin.right)
    .attr("height", height + margin.top + margin.bottom)
    .append("g")
    .attr("transform",
        "translate(" + margin.left + "," + margin.top + ")");

x.domain(filtered_data.map(function(d) { return d.key; }));
y.domain([0, d3.max(filtered_data, function(d) { return d.values; })]);

svg.selectAll("bar")
    .data(filtered_data)
    .enter().append("rect")
    .style("fill", "steelblue")
    .on("mouseover", onMouseOver)
    .on("mouseout", onMouseOut)
    .attr("x", function(d) { return x(d.key); })
    .attr("width", x.rangeBand())
    .attr("y", function(d) { return y(d.values); })
    .attr("height", function(d) { return height - y(d.values); });
```

JavaScript Code Bar

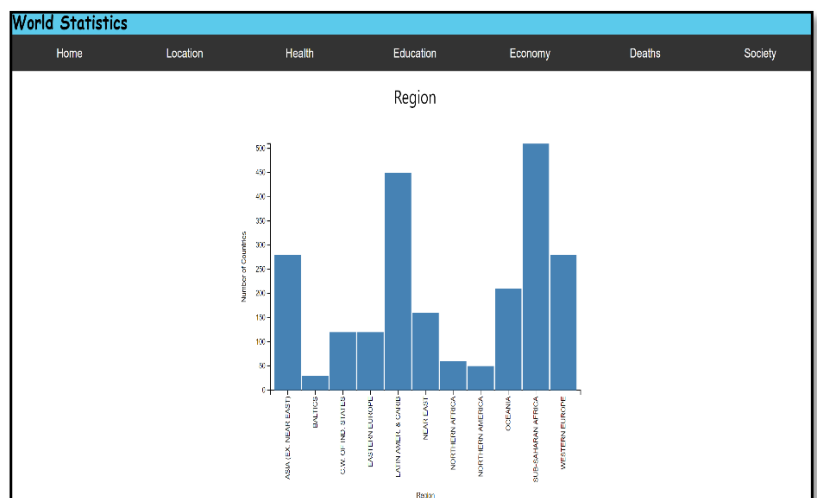


Chart to display Categorical Values

3. Drop down to choose various date ranges

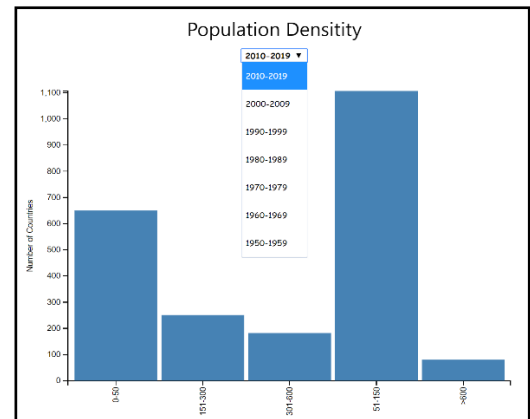
Since the date range is quite large, comparing data in the same date range is more valid. The dropdown menu consists of various date ranges to choose from and the graph changes accordingly.

```
d3.csv("data/Final_Dataset.csv", function(error, data) {

    var selection = elements[0]
    filtered_data = getFilteredData(elements[0], data)
    createElements();

    var selector = d3.select(".date-dropdown")
    .append("select")
    .attr("id", "date-dropdown-selector")
    .on("change", function(d){
        selection = document.getElementById("date-dropdown-selector");
        filtered_data = getFilteredData(selection.value, data)
        d3.selectAll("rect").remove();
        d3.selectAll("text").remove();
        d3.selectAll("svg").remove();
        createElements();
        y.domain([0, d3.max(filtered_data, function(d){
            return +d.values;})]);
    });
});
```

JavaScript Code Snippet for Dropdown Selector



Drop down to choose date range

4. Radio buttons to switch between charts

Mouse-over displays focuses on the bar and displays its value

Radio buttons are provided to choose one of the sub-attributes of Health Expenditure upon choosing which the graph changes.

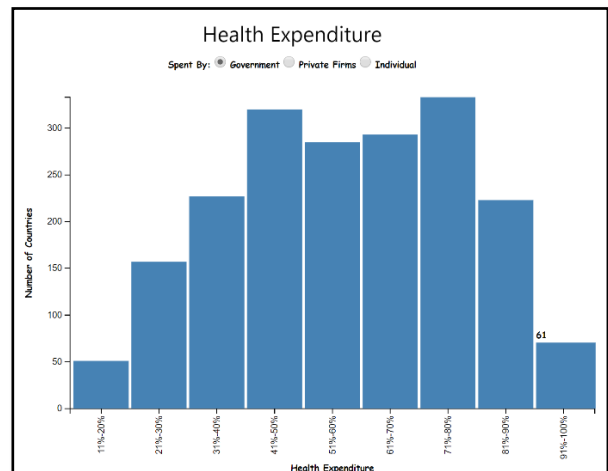
```
div id="Health_Expenditure" class="tabcontent">
<h3 align="center">Health Expenditure</h3>
<div class="chart-container" align="center">
    <form class="controls">
        Spent By:
        <label>
            <input type="radio" name="market" value="Health_Expenditure_Govt_Categories" checked> Government </label>
            <input type="radio" name="market" value="Health_Expenditure_Pvt_Categories"> Private Firms </label>
            <input type="radio" name="market" value="Health_Expenditure_Indv_Categories"> Individual </label>
        </label>
    </form>
    <svg class="chart js-chart"></svg>
</div>
</div>
```

HTML Code for Radio Buttons

```
//mouseover event handler function
function onMouseOver(d, i) {
    // d3.select(this).attr("fill", "orange");
    d3.select(this)
        .transition() // adds animation
        .duration(400)
        .attr('width', x.rangeBand() + 5)
        .attr("y", function(d) { return y(d.values) - 10; })
        .attr("height", function(d) { return height - y(d.values) + 10; });

    svg.append("text")
        .attr('class', 'val')
        .attr('x', function() {
            return x(d.key);
        })
        .attr('y', function() {
            return y(d.values) - 15;
        })
        .text(function() {
            return [d.values]; // Value of the text
        });
}
```

onMouseOver() Javascript Function



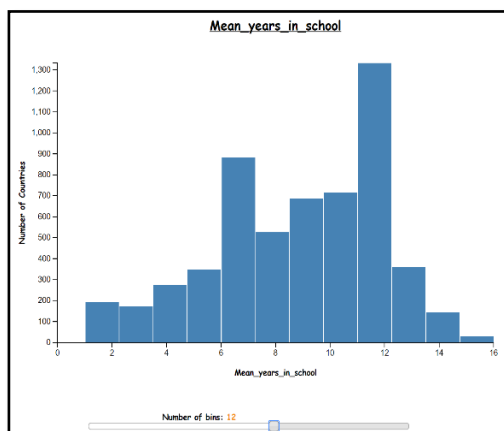
Radio Buttons used to switch charts,
Mouse-over used to focus on bar

5. Histogram for Numerical Attributes, Slider to increase / decrease number of bins in a Histogram

For numerical data, a histogram is created with 10 bins by default. A slider is provided, which on dragging allows the user to increase / decrease the number of bins in the range of 1 to 20.

```
<div id="Common" class="tabcontent">
  <h3 id="common-h3" align="center"> </h3>
  <div id="histogram_container" align="center">
    <div id="histogram"></div>
    <div>
      <p>
        <label for="bins">Number of bins:</label>
        <input type="text" id="bins" readonly style="border:0;
      </p>
      <div id="slider-range-max"></div>
    </div>
  </div>
</div>
```

HTML Code for Histogram and Slider



Histogram with Slider

```
// Get Values
var histogram_values = []

for (var i = 0; i < data.length; i++)
  histogram_values.push(parseInt(data[i][tabName]));

// var min = histogram_values.reduce(function(a, b) {return Math.min(a, b);})
var max = histogram_values.reduce(function(a, b) {return Math.max(a, b);});

x_gh_scale = d3.scale.linear().range([0, width])
  .domain([0, max]),
x_gh_axis = d3.svg.axis().scale(x_gh_scale).orient("bottom");

var new_data = d3.layout.histogram()
  .bins(x_gh_scale.ticks(10))
  (histogram_values);

var y_gh_scale = d3.scale.linear().range([height, 0])
  .domain([0, d3.max(new_data, function (d) {
    return d.length;
  })]);

var y_gh_axis = d3.svg.axis()
  .scale(y_gh_scale)
  .orient("left")
  .ticks(10);
```

JavaScript Code Snippet for Histogram

6. Country-wise Population Statistics on a Map

world-population.json: Consists of country name and co-ordinates, used to plot the map.

Hovering on the country displays its name and population.

```
svg.call(tip);

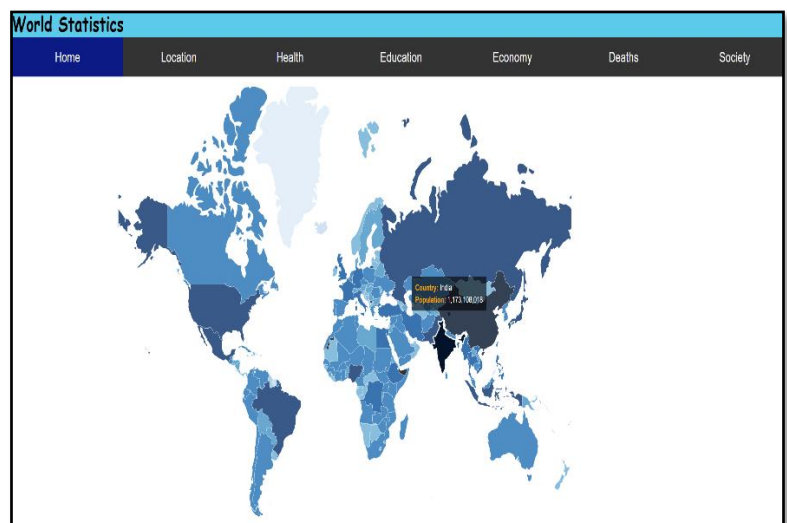
queue()
  .defer(d3.json, "data/world_countries.json")
  .defer(d3.tsv, "data/world_population.tsv")
  .await(ready);

function ready(error, data, population) {
  var populationById = {};

  population.forEach(function(d) { populationById[d.id] = +d.population; });
  data.features.forEach(function(d) { d.population = populationById[d.id] });

  svg.append("g")
    .attr("class", "countries")
    .selectAll("path")
    .data(data.features)
    .enter().append("path")
    .attr("d", path)
    .style("fill", function(d) { return color(populationById[d.id]); })
    .style("stroke", "white")
    .style("stroke-width", 1.5)
    .style("opacity", 0.8)
    // Tooltips
    .style("stroke", "white")
    .style("stroke-width", 0.3)
    .on('mouseover', function(d) {
      tip.show(d);
    });
}
```

JavaScript Code Snippet for Map



Country-wise Population Statistics on Hover

References for Datasets:

- [1] Region: Kaggle - <https://www.kaggle.com/fernandol/countries-of-the-world/version/1>
- [2] Population Density: <https://population.un.org/wpp/>
- [3] Health Expenditure: <https://www.who.int/gho/en/>
- [4] Life Expectancy: <https://population.un.org/wpp/>
- [5] Education Expenditure: UNESCO
- [6] Mean years in school: <http://ghdx.healthdata.org/record/global-educational-attainment-1970-2015;>
<http://www.healthmetricsandevaluation.org/>
- [7] Average Income: <http://gapm.io/dgdppc>
- [8] Employment Rate: <https://www.ilo.org/ilostat/>
- [9] Income Inequality: <http://gapm.io/ddgini>
- [10] Child Mortality Rate: <https://www.gapminder.org/data/documentation/gd005/>
- [11] Adult Mortality Rate: (1) United Nations Population Division. World Population Prospects 2017 Revision, (2) University of California, Berkeley and Max Plank Institute for Demographic Research. Human Mortality Database.
- [12] Suicides: <https://www.healthdata.org/>
- [13] Murders: <https://www.healthdata.org/>
- [14] Military Expenditure: <https://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS>
- [15] Corruption Perception: <https://www.transparency.org/research/cpi>