Assignment 3 – Creating Interactive Visualization Software

- Anurag Banerjee and Vaibhav Walvekar

Introduction

The domain we choose was the literacy data for thirty four state and union territories of India for the years 1981, 1991, 2001 and 2011 and also the national literacy rate in these years. We choose this data since it shows the trends of how the literacy rate in India has increased for the various states over the years. The data is further split into the number of years and into males and females. Thus the data also shows the disparity among literacy rates of men and women in some states and also some drastic increases or decreases in literacy rates of some states in particular decades. Going forward we would like to explore the reasons for these trends in a different visualization study.

For now, the data we have consists of state and union territory names, year, and three columns to show the percentages of literacy, namely overall literacy, literacy percentage among males and literacy percentage among females. This data has been extracted from the census conducted by Indian government in years 1981, 1991, 2001 and 2011. Since the data is in percentages, the representations are also in percentage formats. We choose this data mainly because we are aware of these statistics and we can graphically see how the literacy rates increase over the years. Also we wanted to explore about if there are any trends among states, such as drastic increase in female literacy rate or drop in average literacy rate. The data is present in .csv format.

Interactivities Included and Reasons for their Inclusion

1. **Dropdown for gender:** We have included a drop down for gender values which gives us options for filtering data based on average values, males or females. Default value selected in Males.
2. **Dropdown for year:** Data can be filtered for the following years: 1981, 1991, 2001 & 2011. Default value selected is 1991.
   When combining both of these filters, we can get the data for the particular year for selected gender or the overall data.

3. **Tool-tip:** When hovering over the bar for any of the states, we are able to see the exact literacy value via a tool tip. This gives us a better idea as compared to viewing the Y-axis on a bar chart.
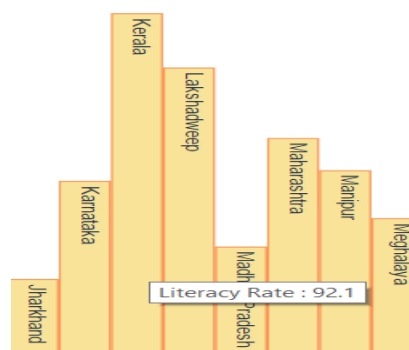
4. **Slider:** The slider has the percentage range of the minimum and maximum values among all the three columns i.e. Overall literacy, Literacy percentage among males and literacy percentage among females. When changing the range of the slider, graphs of only those states are visible whose values lie in the slider range. The graphs are also rearranged accordingly.

Literacy Percentage Range: **20% - 100%**

5. **On-click graph:** We have implemented a separate graph with its own X & Y axes to the right of the existing axes. When clicking on a particular state on the main axis, the year data along with overall, male and female literacy percentages are displayed here. The Y axis of this graph displays the name of the state which is selected. Initially on loading it depicts the Literacy % for India. The screenshot attached is after we have clicked on Rajasthan state bar present in the first axes. This is a grouped bar chart.
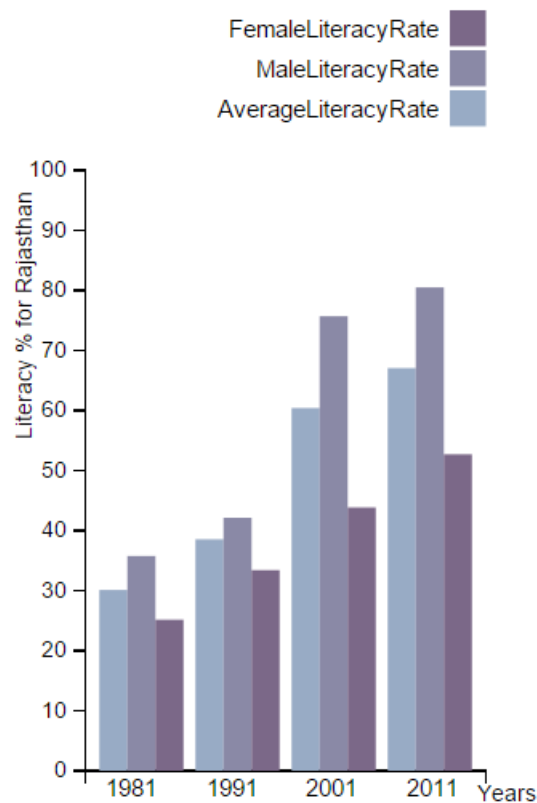

We have chosen above mentioned interactivities because we thought dropdowns on years and gender helped us get to comparisons about similar entities among states and union territories. Slider helped us to focus on particular range of states in a given year and may be categorize them as high or medium or low literacy states and see whether they continue or rise to higher brackets in following decades. Tooltip adds the convenience of knowing the literacy rates on hover over the bar graphs. This helps as it reduces efforts of viewer to check value of literacy rate on the Y axis. Finally, the on click feature adds to interactive visualization by detailing out all literacy rates over the years for a particular state. Thus

helping us study any trends for particular state in one go. We think all these interactivities blend well together and showcase information present in our data succinctly.

Challenges faced & Final Output

We started the work between us by initially learning the basics of JavaScript. Moving on to D3, we started by building basic rectangles and circles. Then we learnt importing the data and creating simple visualizations on them. Following that we split the tasks for creating the different visualizations. Vaibhav focused on generating the sliders and dropdowns while Anurag focused on graphs and tooltips. After completing those, we focused on combining the two codes to generate a single working visualization for the data. Once our basic interactive visualization was in place we explored the opportunity of having on click feature which would bring in more detailing to the visualization, thus we together developed on click feature for our visualization. The resultant view helps in knowing granular details of particular states though it is not connected to any other interactive feature other than on click.

Literacy Rate comparison among Indian State and Union Territories over past four decades
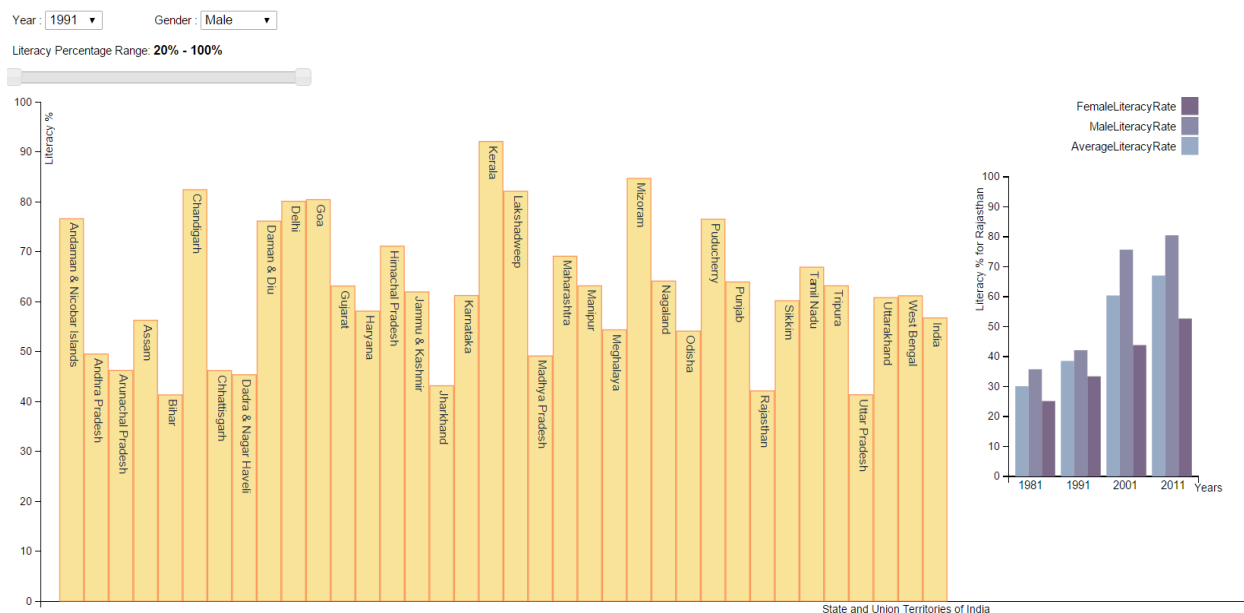
We spent almost 60 hours in total on developing the visualization. The most time consuming aspects were implementing the logic for slider and dropdowns. Generating the visualizations were not so difficult, however, customizing them for our data was what we spent most time on. Also, a lot of time was spent on debugging random errors, identifying the source of which was an issue.