Data Processing
Homework 7: Design Process

Vincent Erich
10384081

December 11, 2015

1 Introduction

In this document I will briefly describe the design process while making the dashboard. I will describe my design choices per step.

2 Step 1

I have used four datasets from The World Bank for the dashboard:

- Population ages 0-14 (% of total)[1]
- Population ages 15-64 (% of total)[2]
- Population ages 65 and above (% of total)[3]
- Population, total[4]

I have written a Python script to process these datasets and create a single JSON file. Since DataMaps[5] expects a specific data format, the data in this JSON file has a format that is compatible with DataMaps. The format of the data is as follows:

[1] Source `http://data.worldbank.org/indicator/SP.POP.0014.TO.ZS`
[2] Source `http://data.worldbank.org/indicator/SP.POP.1564.TO.ZS`
[3] Source `http://data.worldbank.org/indicator/SP.POP.65UP.TO.ZS`
[4] Source `http://data.worldbank.org/indicator/SP.POP.TOTL`
[5] `http://datamaps.github.io/`

```
{
    'fills': {
        'popCategory1': 'color 1',
        'popCategory2': 'color 2',
        ...
        'popCategory9': 'color 9',
        'defaultFill': 'default color'
    },
    'data': {
        '1960': {
            'ANG': {
                'total': 'total population',
                'fillKey': 'popCategory2',
                'precentages': [
                    ['0-14', 'percentage 0-14'],
                    ['15-64', 'percentage 15-64'],
                    ['65+', 'percentage 65+']
                ]
            },
            'AUS': {
                ...
            },
            ... // Other countries (country codes).
        },
        ... // Other years.
        '2014': {
            ...
        }
    }
}
```

This format makes it easy to select all the data from a specific year. If you load the data with D3 (using `d3.json`) and bind the data to a variable (e.g., `loaded_data`), then you can select all the data from a specific year with

```
var data_specific_year = loaded_data['data']['specific_year'];
```

The data format for the percentages (i.e., a list of lists) is convenient to work with in D3.

3 Step 2

The dashboard consists of an interactive choropleth map and an interactive bar chart. The choropleth map encodes the total population, and the bar chart encodes the distribution of age groups in a country.

The interactive choropleth map is the same as the one that I have made for the previous assignment (week 6); see that map for a discussion of the design

choices that I have made. See the choropleth map from week 5 (SVG3) for a discussion of the color scheme that I have used.

I have made a number of design choices while implementing the interactive bar chart. First, the fill color and stroke color of the bars are the same as the fill color and stroke color of the country that is currently selected. I have done this to clarify that the data shown in the bar chart relates to the country that is currently selected (hovered on). Second, the bars in the bar chart are not ordered from low to high or vice versa (in terms of percentages). I have done this to preserve consistency (i.e., the bar for the age group 0-14 is always on the left, the bar for the age group 15-64 is always in the middle, and the bar for the age group 65+ is always on the right). Third, the maximum value on the y-axis is equal to the percentage of the age group with the highest percentage. This allows for better comparison between the age groups. Fourth, I have added an interactive tooltip to the bar chart that shows the exact percentage for an age group (bar). The tooltip appears when the mouse enters a bar and disappears when the mouse exits a bar. Furthermore, when the mouse enters a bar, a black border is drawn around it. This black border disappears when the mouse exits the bar. I have done this to (further) clarify to which bar the tooltip relates.

4 Step 3

The interactive choropleth map and the interactive bar chart are connected with each other. If you hover over a country on the choropleth map, the distribution of age groups in that country is shown on the right, below the legend (as a bar chart). When the distribution (bar chart) is drawn, the bars move from 0% to the corresponding percentage (i.e., the bars are animated). I have done this to draw the user's attention to the bar chart.

When you click on the interactive choropleth map, the map interaction will be disabled (i.e., no popups and highlights on hover). This allows you to explore the interactive bar chart. In order to clarify this, when the map interaction is disabled, the fill color for all the countries changes to light-grey, except for the country that was last hovered on (and for which the distribution of age groups is shown); this country remains highlighted.

5 Step 4

The dashboard also has a third interactive element: you can select the year for which you want to see the data. This was a fairly obvious choice, since the data from The World Bank goes back to 1960. This also clarifies the structure of the data as explained in Section 2.