Fun with Flags
(Inspired by Dr. Sheldon Cooper)

1. Description of the Data

The data for this project was assisted from the following source:

Lichman, M. (2013). UCI Machine Learning Repository [https://archive.ics.uci.edu/ml/datasets/Flags]. Irvine, CA: University of California, School of Information and Computer Science.

This data file contains details of various nations and their flags. The fields in this file are space separated instead of comma separated. There are 30 attributes for each country documented. At the time of the collection of this dataset, there were only 194 countries. Many flags have been redesigned ever since. However, from basic research on Google, we noticed that there were no significant differences in what we wish to present in the visualization of this project. Hence we decided to use the dataset to the exact values.

We converted the comma separated data into JSON files except for one visualization in which we have used a Tab Separated (tsv) format. The data was scraped by writing python scripts. In the original dataset, we had a column for every color with a value being a binary 0 or 1 to indicate if the color appears on the flag or not. We have modified this to a JSON object format where a country name is associated with a list of only the colors that are present on the flag. Also, since the number of countries is a very large number to visualize, we have divided them as per the continents, which is indicated by the attribute 'landmass' in the original data.
For example: An item in the 'continent to color' JSON looks like the following:

```
{
        "name": "Asia",
        "color": ["red", "green", "gold", "black", ....."red", "gold"]
}
```

To represent the number of shapes used by countries of a particular religion, we have calculated the percentage of the countries in a religious category to the total countries in that category to know how a religion can influence the elements on a flag. This data is in the TSV format. A row in this file indicates the following.

religion	element	percentage
1	1	30.77
.	.	.
.	.	.
6	8	37.5

Religion 1 corresponds to Catholic while 6 is for all others religions. The numbers in between are assigned to various other religions which is documented in the file README.txt. Element 1 is for the star shaped structures while 8 is for stripes on a flag. A row in this file can be read as, 30.77% of the Catholic countries have stars on their country flags.

The dataset has an attribute called 'mainhue' which tells us which is the dominant or the main hue in a country's flag. In addition to this there is information about all the colors present in a flag as described above. When we designed the visualizations for country vs color, to show which part of the world used which color the most, it was difficult to decide between the 'main hue' or the 'all colors' information. We performed our experiments on both types of data and noticed that when we consider all colors, the results were more promising. Hence, our final visualizations represent all the colors. To support this decision, we notice that many countries have stripes or bars distributed equally. In this case, picking just one of the colors and drawing conclusions on it did not justify our purpose. Hence we chose all the colors in the flag as they speak about a country as a whole.

2. Description of the Visualizations

Based on the inspiration from the examples shown in the class and also those listed on the D3. Gallery, we have constructed the following visualizations. The colors used follow hex codes listed on the website color brewer.

2.1 Stripe Flag

We create this stripe flag according to the percentage of the color used in flags all over the world. We have one scale here which is the ordinal scale to make each stripe with their color accordingly. Each color has an index, the domain of this scale is their index, as for the range, we list the hex of color according to the index, therefore, we can use this scale color to fill the stripe.
For the height of each stripe, we use the data about color and number (colorset). And we use a for loop and the index of each color to convert each number to a percentage and then to the height they should be. What's more, the label of each percentage of the color is also generated from this dataset.

2.2 Donut Chart

We implemented this donut chart to show two information: one is for which is the most common number of colors used in the flags; the other one is for which is the most common number of shapes used in the flags. We choose number from 1 to 7 and over 8 which including all the flags over the world.
In this donut chart, we use arc function to create donuts to show each data by transform attribute, like [attr("transform","translate("+ ((width/10) + t1*120) +","+ (200) +")")]. Therefore, we created 8 donuts in a line. We get all the data which is number of colors and number of shapes from the dataset we cleaned, and read the JSON file, calculate the percentage of each number and generate the arc based on the data. We add the description text such as the number of color and the number of shape for each donut using transform attribute too.
From the donut charts, we can learn that in the number of color part, there are most of the countries use three colors in their flags, and in the number of shape part, we can learn that most of the countries in the world use only one shape in their flags.

2.3 Bar Chart

We created this bar chart to show the relationship between continents and colors, religions and color, continents and religions. At last, it shows the relationship between religion and color in the flag. There are two bar charts in each side of the svg element: one is six continents (Yes, because the Antarctica doesn't have any country and flag!) and their top 4 colors in the flags, the other one is six main religions (If you are finding Buddhism and Hindu, they are in Other religions) and their top 4 colors in the flags. The middle chart connects the continents and religions which represents the relationship between them, for example, if the biggest religion in Asia is Muslim, the line between Asia and Muslim will be the darkest line out of the Asia.
For the continents vs colors and religions vs colors bar charts, they have the same kinds of scale function. yScale linear scale function helps we show the continent names or religion names, xScale linear scale function helps we generate width of each bars. We also use ybars ordinal scale and ybars ordinal scale to set the length of first node and last node, and the interval between each node by different ranges attribute which are rangePoints and rangeRoundBands.
Besides the scales, we create a dictionary named Dic_cont_color and Dic_reli_color to store the keys and counts for each continent and religion, we also use sorting functions to locate the top 4 color and their index to implement the bars combining with for loop and different kind of scales.
For the continents vs religions line chart, we use the same method like before to create a dataset including the continents and their top 3 religions. In the implementation process, we use two for loop to calculate the percentage of each religion in the continent and use these numbers in the opacity of lines to show different relationships.
From these combined chart, we can show that the religions do influence the color of the flags. For example, let's see the top three religions of Asia, they are Muslim, Others and Marxist. The top color among these three religions is red, and the second largest number of color for Muslim is green which is also in the top 4 colors for Marxist. At the same time, the second largest color of Asia is green. This phenomenon does work on other religions which can learn from this chart.

2.4 Heat Map

We implemented this heat map chart to show the relationship between each religion and all the common shapes in the flags. There are stars, circles, triangles, texts, crosses and so on. And this is also a chart we can learn the percentage of shape used in all the countries if we look the chart column by column.
We create an ordinal scale function to show the 5 color in the heat map. We also wrote functions to calculate position of each little segments in the map and add a rectangle to show the numbers and their corresponding transition color to make the map more understandable.
From the map, it is obvious to know that star is the most popular shape in all the religions which also means star is the most popular shape in all over the world. In addition, stripe is a common pattern found in the country flags and that's why we can find at least 18 flags in the final color flag.

2.5 Color Flag

We create this flag using the data we get from the previous results. The most popular colors in flags are red, white, blue, green and gold, and the most of the countries only use one shape in their flags which star is the most common shape.

Therefore, we pick 5 colors above and allocate the area for them according to their percentage and make the stripe flag. We can find out 14 flags just using the different combinations of the stripes. What's more, we can add one star according to our data and find out another 4 flags in this color flag.
We create an object contains the coordinates, width, height and color and a for loop to implement the stripe flags and these 18 countries' flags.

3. The Story

A flag consists of colors and/or elements, in this project we have shown you religion does influence the flag of a country and there is an easy method to find that out in many flags all over the world when you see the visualizations we created.

How is religion related to the flag?

Let's look the third chart, what's the top 3 religions for Europe? You can make a best guess and find out the answer from the three lines out of the Europe, they are Catholic, Marxist and Other Christian. For Europe, the top 4 color in flags are red, gold, black and blue, then what's the top 4 color about Catholic, Marxist and Other Christian? They are red, gold, blue, black and white which are same with the Europe top 4 color! Is Europe the only continent which influenced by religion? NO! NO! NO! You'll find all the continents top 4 colors are the same with their religions' top color. Religion also influences shapes on the flag! Let's see the heat map. Stars and stripes are commonly used in every religion, crescent is preferred by Muslim for in their culture, crescent and stars are the symbol of Muslim country. The star-and-crescent in the Flag of Pakistan is stated as symbolizing "progress and light" (while the green color is stated as representing Islam)
Crescent has long been a religious cult in Arabic area, for the day is too hot so people usually do all the work at night under the light of moon; Also in the history of Muslim country, crescent and star was one the symbolic of the power, so people still worship them.
Similarly crosses are preferred by the Catholic. It is raised from Christian-civilization. For the death of Christ was on the cross for the happiness for his people, people use cross to remember him.

How can we find out 18 flags in a simple stripe flag?

If you want to design a 'world flag' like us, which contains as many as flags, how do we do this flag? Let's make it step by step.

1. First, which color should we choose?

Looking at the first stripe chart, you will know red, white, blue, green and gold are really popular among all the countries. Therefore, if we chose only these colors.

2. How many colors to choose?

From the donut chart, we learn that 75 out of 194 countries use three color in their flag. Hence, looking for three color in a flag in a particular order was easier!

3. Should we have any elements on the flag, if yes, how many?

First of all, how many elements can you choose? Looking at the donut chart again, we have already found that 59 out of 194 countries use only one shape which means most of the countries use one element! It was better for us to choose just one shape.

4. What elements can we use?

Let's look the heat map, stars and stripes are the most popular shapes over the flags. The stripes pattern and the combination of colors was taken care of in the previous steps. Therefore, you just added the most common element, a star, to our flag.

Congratulation! You have known the secrets with flags! Dr. Sheldon Cooper will be proud of you! It's not easy to create a flag which contains 18 flags! Try it for yourself and see how many flags you can put in a simple flag!

Team :

Devi Snigdha Muppala (dm754)

Di Jiang (dj327)

Ziwei Zhao (zz384)

TA:

Linna Li (ll577)