PART A

In the last assignment you answered a few questions based on the plots. In this assignment we will make the plots answer our questions for us. We will be using scales and axes extensively in this assignment. We have provided you with a dataset along with the assignment files. You will create visualizations based on questions to answer the problems.

Note: For each of the questions below, refer to the textbook Visualization Analysis and Design (Tamara Munzner) (Ch. 10) and determine appropriate color schemes for each of the plots below. Type your reasoning for choosing a specific color scheme below the plot in the "div" tag provided in index.html. Along with the color schemes, mention the following for each plot: keys, values, marks, and channels.

Question 1:

Problem Statement:
[20 points] Create a bar plot to show variation of participation after the year 1980. If an athlete participates in 3 different events the number of participations for that athlete is 3. Provide appropriate scale and axes to the plot. Do NOT hard code the scale/axes. Change the color of bars according to the setting given below:
Participations>= 2000 : Color1 Participations in range 1800-2000 : Color2 Participations less than 1800 : Color 3 Color 1, Color 2, and Color 3 to be determined by you.

I chose this color scheme because of its naturally light aesthetic. A white background typically is easy on the eyes when other colors are lighter toned. In addition, orange and blue create a nice contrast and green contrasts with orange as well. The keys are the years and the values are the number of events. The marks are the lines of the barchart and the channels include the colors, width and lengths of the bars.
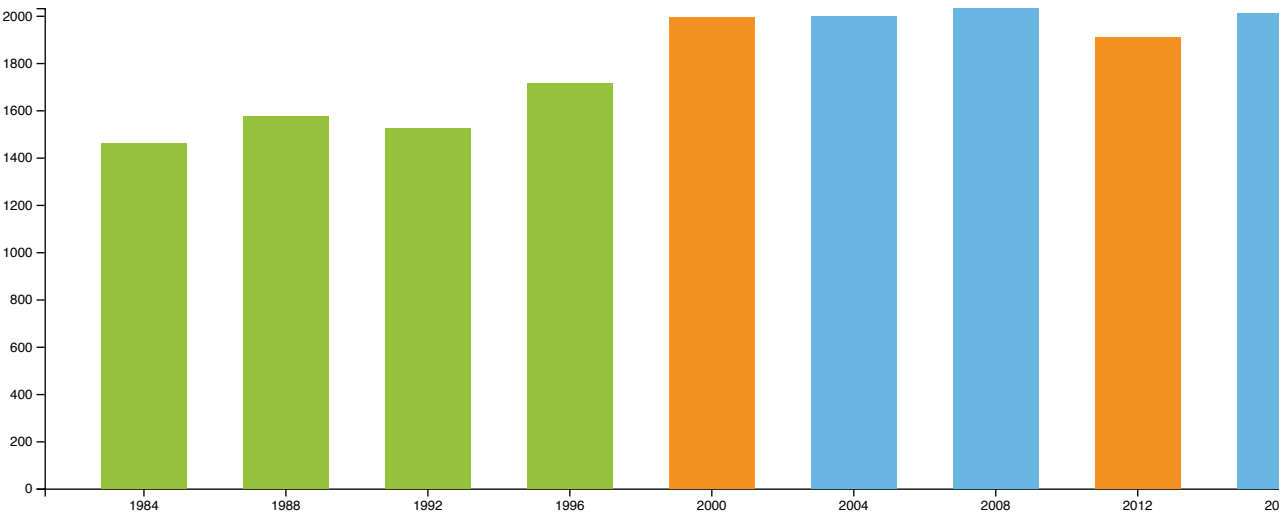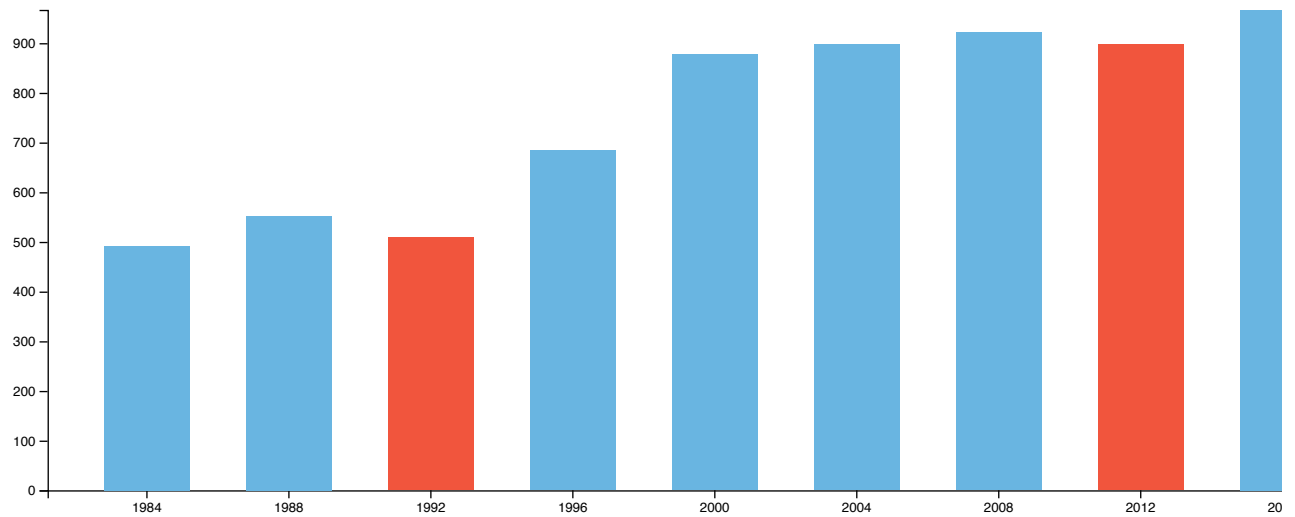
Plot:

Question 2:

Problem Statement:
[20 points] Next we look at the trend of female participation in Summer Olympics. Create a bar plot to show variation of female participation after the year 1980. Change the color of bar to Color 1 if the number of female participants drops from the previous year. Change the color of the bar to Color 2 if the number of participants increases from the previous year. Provide appropriate scale and axes to the plot. Do NOT hard code the scale/axes. Which year has the least and highest number of female participants? Color1 and Color 2 to be determined by you.

The year 1984 had the least number of female participants while the year 2016 had the most. I chose this light red and light blue color scheme because it matches the tone of the white background. In addition, the harsh red that denotes a year with a decline in particiapnts is used to emphasize that point. The keys are the years and the values are the number of participants. The marks are the lines of the barchart and the channels include the colors, width and lengths of the bars.

Plot:

Question 3:

Problem Statement:
[20 points] Create a scatter plot to show female and male participation in Summer Olympics. Provide appropriate scale and axes to the plot. Do NOT hard code the scale/axes. You are required to create ONE plot and derive the years in which the difference between female and male participation is minimum and maximum. [Extra Credit (+10)] : You can attempt to modify this plot to be a scatter line plot by using d3.line() to connect the data points.

The year 1992 resulted in the largest difference between female and male participation. I chose this light red and light blue color scheme because it matches the tone of the white background. The keys are the years and the values are the number of participants. The marks are the circles of the barchart and the channels include the colors(sexes) and heights of the circles.

**Plot:**
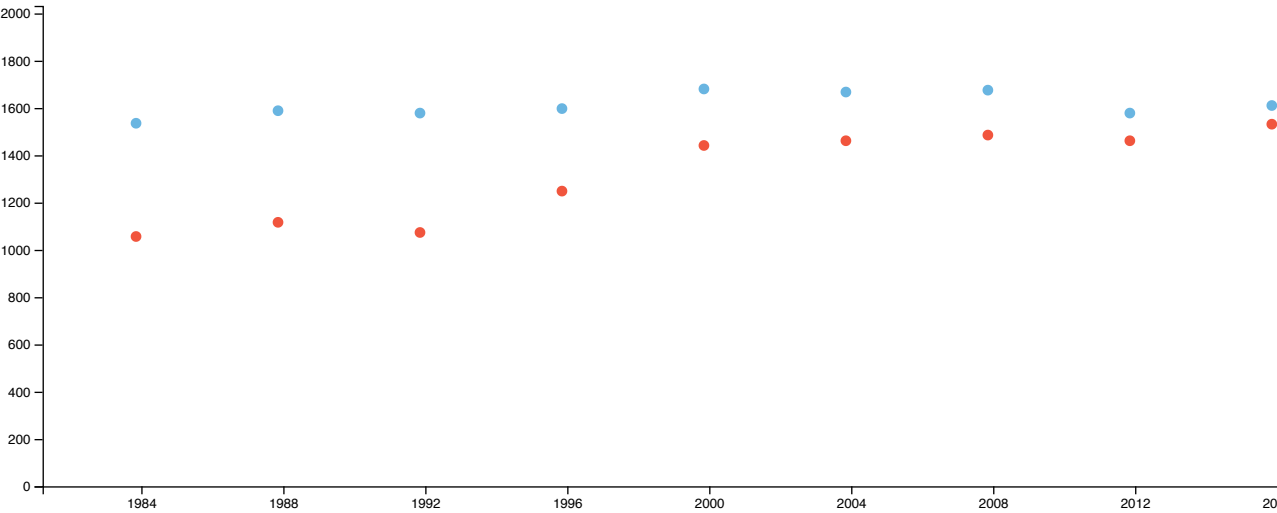
Question 4:

Problem Statement:

[20 points] One of the most important properties of data visualization is to provide information in a compact format that is easily readable. Let's say you wanted to analyse the performance of US Olympics team based on their medal tally. Generally you would plot three bar plots for Gold, Silver and Bronze and analyze those plots. But is there a better way to do this ? Create a stacked bar chart to show the medal tally after the year 1984 for United States of America. Provide appropriate scale and axes to the plot. Do NOT hard code the scale/axes. The structure of stack should be : Gold at the bottom, Bronze at the top and Silver between Gold and Bronze. Use the appropriate colors to represent the stacks:

Plot:
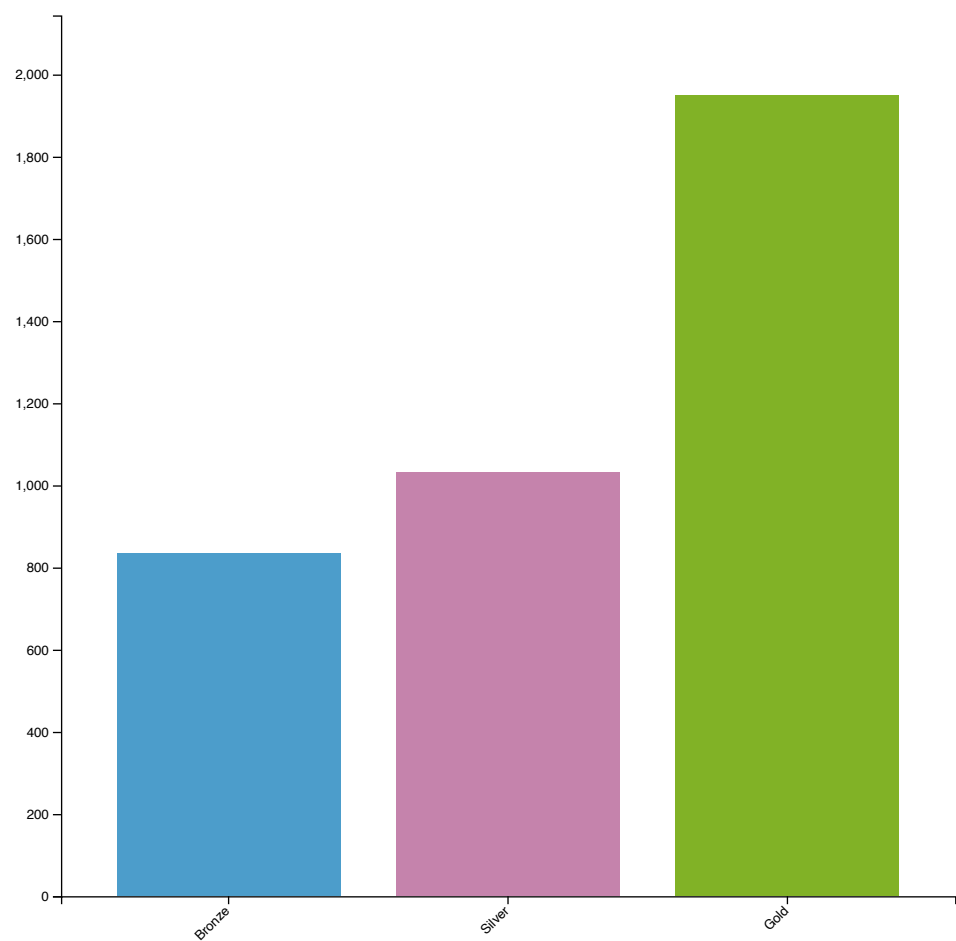
Question 5:

Problem Statement:
[20 points] There is a nice correlation between the height and weight of athletes for specific sports. In this problem we will try to visualize this correlation. We will see this relationship between height and weight for three sports - Basketball, Gymnastics and Judo. Create a scatter plot to show the relationship between height and weight for the three sports mentioned above. Use different colors to represent different sports. Now let us see if this correlation is dependent on gender. Make use of two shades of a color to represent male and female athletes. Example : If you decided to represent basketball with red, plot the data points corresponding to female athletes with pink and male athletes with red. Provide appropriate scale and axes to the plot. Do NOT hard code the scale/axes.

It seems that each sport concentrates in a certain area of the plot. Gymnists seem to weigh the least and are the shortest, those who do judo are the second, and basketball players seem to weight the most and are the tallest. The keys are the weights of participants and the values are the heights of participants. The marks are the circles of the barchart and the channels include the colors(sports), shades,(sexes) and heights of the circles.

Plot: