

**Name:** The Anh Nguyen  
**ASU ID:** 1215126926

## **Assignment #2: Interactive Visualization Exercise**

### **Introduction**

This assignment is to design an interactive visualization to convey a story from the provided data set. I enjoy analyzing match stat of every tennis match. Therefore, I visualize every match through some measurements.

### **Missing Value**

There are some missing values in the dataset, so I use R to fill them by mean value of the column because the mean is a reasonable estimate for a randomly selected observation from a normal distribution and it is also fast to calculate and implement.

### **Story and Visualization**

I visualize all of tennis match from the dataset with three dropdown selection in order, which are year, round and name of two players on the top of HTML page. After choosing a match, you just click the “Update” button next to the dropdown selection to see the bar chart of the match.

### **Chart and Analysis**

Double side bar chart will be used in this assignment because in my opinion, this type of chart is suitable to provides a general overview to the reader, compare information of two players and give reader a good view to see the difference between them. From double side bar chart, we can examine every measurement of two players and then draw a conclusion of which information is important for a winner in Australian Open Tournament.

### **Normalization**

Some features are not measured in percentage, so I need to normalize them. I convert them into percentage by dividing each of information of player by sum of both information of that feature and multiplying them by 100.

For example, if double fault of first player is 5 and that of second one is 15, so sum of them is 20, double fault of first player will be converted to 25 ( $5 \times 100 / 20$ ) and that of second one will be converted to 75. They will be equivalent to percentage measurements and they become percentage of bar's width.

### **Size and Color**

The chart is placed inside a box with black border in the middle of the page. Moreover, there are also two pictures of players on both sides of chart and their name below. The list of measurements is in the middle of the chart.

Above the chart, some information about match is showed, such as year, round and result of the match. Background of visualization is white to show information clearly. As black, blue and red are contrast colors, it makes chart best for visualization. Blue in the left is used to show information of winner, while red demonstrates measurements of loser. All text color is black to see information clearly with red and blue chart.

### **Interactive Visualization**

When you choose the match and click “Update” button, the bar chart of the match will be shown and the bars are expanded towards two sides gradually.

When you “mouseover” the bar, the dark shade of its bar's color appears and the number showed information becomes bigger to highlight the information. Moreover, when you “mouseover” the name of features in the middle of the chart, bars in both side of that features are also “mouseover” as above to give you a clear look of measurement that you want to compare two players.

### **Browser**

My HTML file works effectively on Firefox

### **Image**

All images of tennis players are collected from **www.atptour.com** and Google