

## **Data Visualization**

### **Abhinethri Tumkur Umesh**

#### **Story:**

The provided data set has statistics about Australian Open Tennis for 10 years i.e. from 2008 to 2017. After performing the analysis of the data, 3 charts are built to convey the results and the identified patterns. Tableau, the reporting tool has been used to produce the static visualization.

Chart 1 is targeted for common people to get the over view of Australian Open Tennis i.e. to know the participants, the winner and the corresponding year. Thus, this chart helps to get the quick summary within few seconds. From the chart1, one can easily conclude that the in the year 2010, players were Roger and Andy and the winner was Roger. If users would like to know the pattern of various attributes for winners and losers, then they can refer Chart 2 and Chart 3.

Chart 2 helps the viewer to compare the pattern of attributes for losers and winners (for 10 years). The comparison includes important attributes such as error, double fault and ace. Analyzing these patterns will help the viewers to find out what could be the reason for losing the match and can prioritize the attributes. For example, if we refer errors section then we could see that most of the times, player with less errors have won the match. Also, we could find that Ace didn't play an important role for winning the matches. Similarly, pattern could be found for other attributes. These patterns mainly help the tennis player to concentrate on specific areas to improve their performance. Labels in the graph represent the actual value.

Chart 3 is very much like Chart2 in terms of representing the pattern except for the attributes whose values are in terms of percentage. For meaningful information, it is always recommended to compare the entities which are represented in the same/similar way.

#### **Design:**

##### **Chart1:**

Dots are used to represent the common view as it is simple and easy to understand. Blue color is used by large companies and it reflects depth and stability.

##### **Chart2 and Chart3:**

Dual lines (Winner and loser) have been used because it makes the visualization better and comparison of patterns clear.

Green line has been used to represent the winner and red line has been used to represent the loser because this is the generic notation and can be easily understood by common audience.

#### **Justification:**

The represented static visualization targets the common users as well as tennis players. Chart 1 targets the users who are interested in overview. Chart2 and Chart3 could be used by common users if more information has to be fetched about winners and losers but these charts are mainly useful for tennis players as they can work on areas which requires more effort from their side to play better in their upcoming matches. This information would have been tough to understand by looking at raw data but the issue has been resolved through simple static visualization.

Chart1:

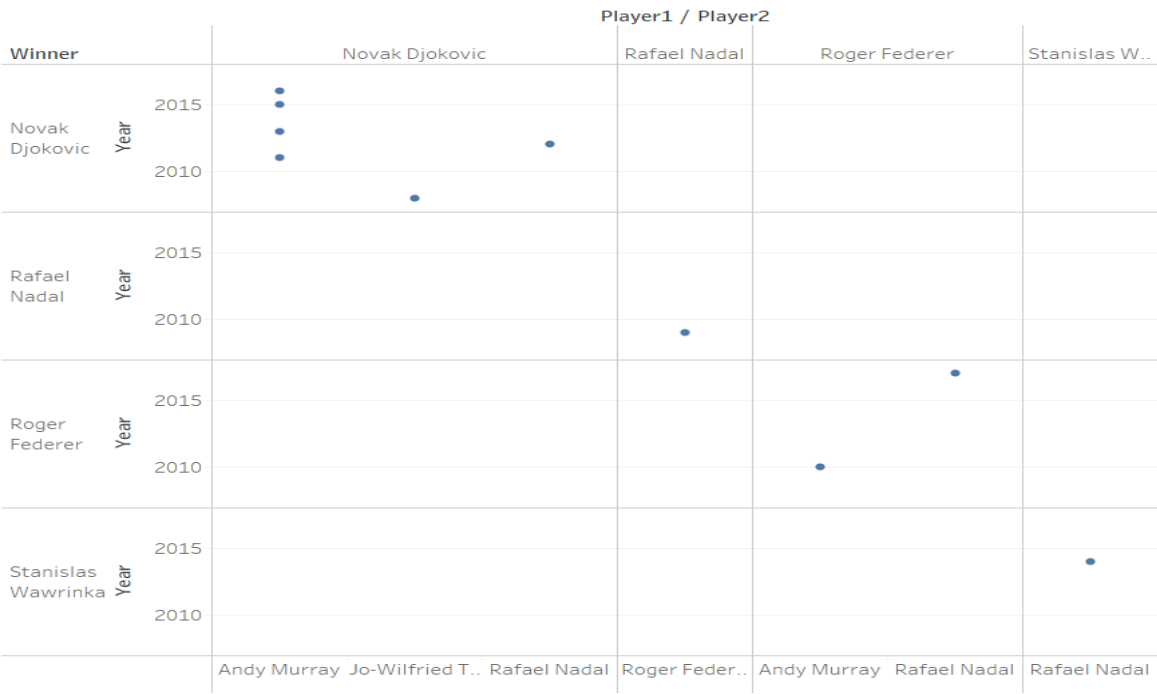


Chart2:

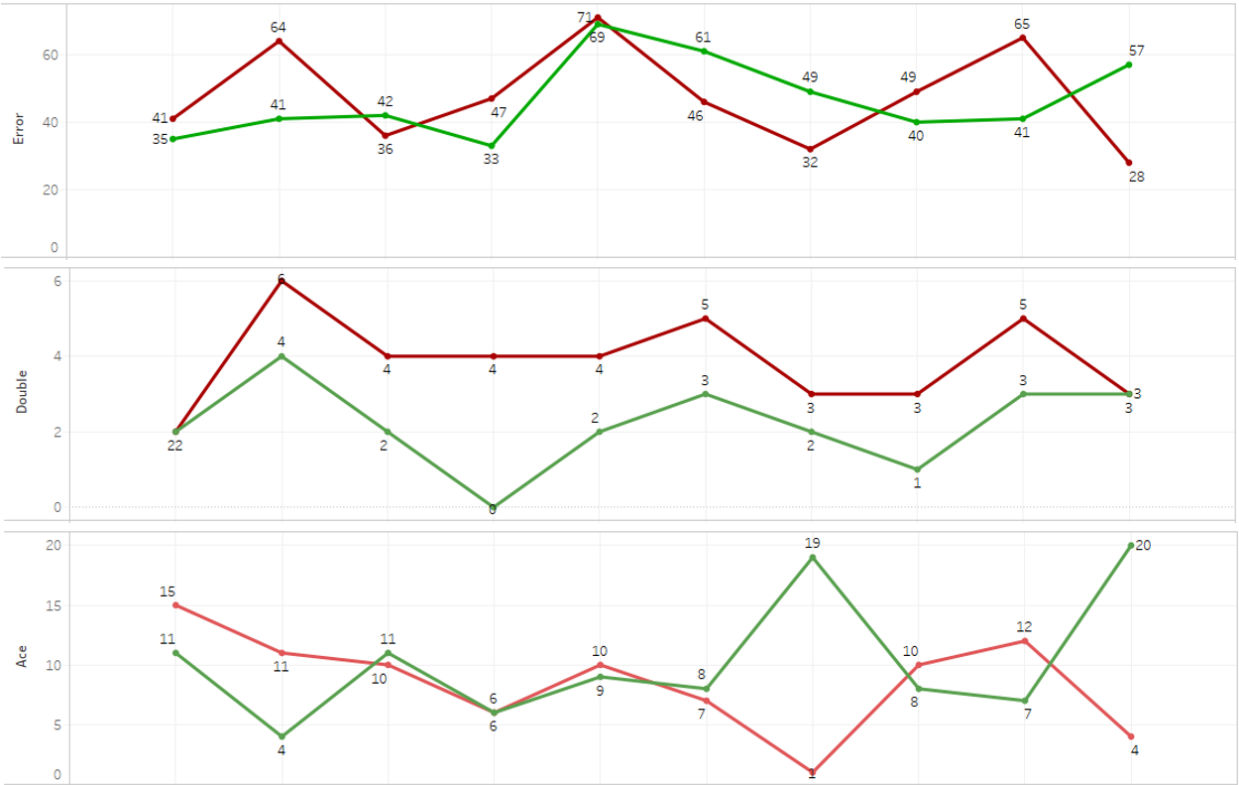


Chart3:

