

Data Intensive Computing Term Project

By

Pruthvi Mulagala – 50208595

Vamsi Nagabhiru - 50204813

Dashboard Link - <http://dic-term-proj.s3-website-us-west-2.amazonaws.com/>

Data used: ATP World Tournament details with the list of players and their details –
ATP_Data.csv

Win Set Dashboard

Win Sets vs Loss Sets:

Linear regression of the number of sets won versus and the number of sets lost by the players starting from the year 2000 to 2016. Filters are provided to limit both the set counts. We have also divided the data into 5 clusters

Sets won per year and prediction for the upcoming years:

Details of the sets won per year of the top players with filters provided for the surface the player on and the series. Then based on the current data the values for the upcoming years are predicted.

Sets won in Grand slams:

Details of the number of sets won by the top players in the four major grand slams – US Open, Australian Open, French Open and Wimbledon. Filters are provided for the year of the grand slam and the round which they played in.

Prediction Dashboard

Wins predicted for the upcoming years:

Based on the current trends of the number of wins of each player in a single player, we predicted the wins the players could win for the upcoming years. For the display purpose we have showed the expected number of wins for the top 5 players.

World rank prediction for the upcoming years:

Similar to the way we predicted the number of wins for the top players above, here too we predicted the minimum rank a player could achieve in the upcoming year based on the previous year ranks. For the display purpose we showed only the top 7 players in the dashboard.

ATP Story Book

Based on the results we obtained in the above dashboards we prepared a story book by imbibing all of them together. In addition to including all the sheets in the above dashboards, we have also included the details of the number of matches won by different players on different turfs. Below are the list of sheets included in the story book:

- Win Sets vs Loss Sets
- Sets won per year with predictions for upcoming years
- Sets won in major Grand Slams
- Matches won on different turfs
- Number of wins of the top players with prediction for upcoming years
- World rank trends of the top players with prediction for the upcoming years

From the story book we can make the following observations:

- Players with the most wins are usually the players who won the most number of sets in any given year
- Surprising players who win a lot of sets are the ones who lost the most number of sets.
- Roger Federer leads the pack with most number of win and loss sets. Players like Nadal has won the second most number of sets while there are quite a few players who lost more number of sets over him. On the other hand David Ferrer has lost the second most number of sets while winning the fourth most number of sets.
- Number of sets won has also a clear reflection on the number of games won by the players.
- The trends show that Wimbledon has been predominantly dominated by Federer in the early years whereas Djokovic and Andy Murray lead the pack in the recent years
- French Open charts for the most number of sets won is lead by Rafael Nadal followed by Federer and Djokovic over the years
- Trends for Australian Open and US Open is also similar to that of Wimbledon
- While to the different kind of courts, Federer leads all the players with the most number of matches won on Grass, Hard and Carpet turfs whereas Nadal leads the pack for clay courts.
- Number of wins have a direct correlation to the rank of the player with Djokovic leading and Federer leading the charts with the top honours.