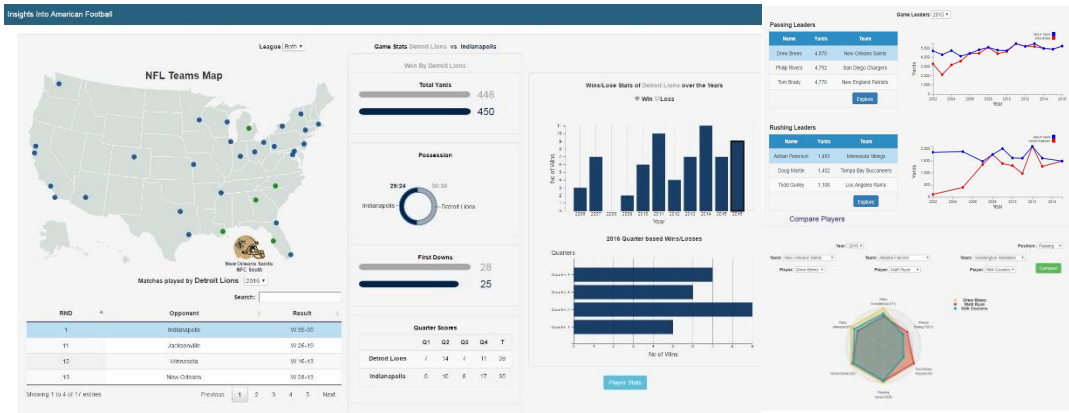


Insights into American Football

Maheedhar Gunnam and Mahesh Reddy Kukunooru, Computer Science, Old Dominion University



Abstract— this article presents a visualization interface to explore the football data and find some interesting aspects of the game over the past 10 years. Most of the sport websites which are followed merely show the game statistics in a tabular data format, which get more difficult for the user to perceive at a first glance. Different idioms like multi-line chart, bar chart, radar and donut charts are used to effectively visualize the football dataset, which aims at providing a platform for the users to help them explore and find some interesting insights that may go unnoticed.

1 INTRODUCTION

American football, aka football is one of the most popular and competitive sports prominent in the United States. It is played at various levels, out of which professional and college football are the most followed forms of the game. The most popular league out of all is the National Football League (NFL) consisting of 32 teams divided equally into Conferences. Many people don't get to follow every game, thus there is a need of a source, which describes the important events that occur in an entire game in an interesting way possible to the user in a short span of time.

In the recent times, most of the websites that we follow for day-to-day sports news only show statistics of team, players in each game, which is restricted in some cases. For example, if we want to look at a player's performance trends over the seasons or if the user wants to compare two or more players based on their skill set, there isn't an elegant platform available. We provide a platform which addresses the above needs, which also becomes an efficient tool to explore more insights that may go unnoticed.

In this article we are providing the user with an interface which visualizes past 10 years of NFL data. The following can be explored from this visualization implementation:

- Super bowl winner, when a season is selected

- Compare win/ loss of a team over the years
- Compare quarter based win/ losses of a team in a particular year
- View the statistics of each game
- Compare different players statistics in a season
- Compare the stats of a player with the highest player stats of that season

We have used tools like D3, jQuery, CSS for implementing the visualizations. This visualization is provided with animated transitions that pleases the user and allow him to perceive the data effectively. Several user interactions are provided like selection based on season, position of player, team etc., which makes the user drill down to the depths of data and make some interesting findings.

2 RELATED WORK




Tools or websites, which does quantitative analysis of a game, can be very useful to those who cannot follow the entire game or to those who are new to the game. Currently most analysis on sports data shows the statistics of team, players in each game, but they do not provide much of visualizations to represent it. If we provide the user with suitable charts which can effectively visualize the underlying dataset, it becomes easy for the user to identify the interesting aspects of the game, rather than merely going through the data which is arranged in the form of rows and columns of tables. We are taking advantage of information visualization to showcase the data presented by the sports based websites into something attractive and effective. The figure below describes the ESPN website which shows various tables which represents the statistics of a team, players



- Maheedhar Gunnam is with the Department of Computer Science, Old Dominion University, Norfolk. E-mail: mgunn001@odu.edu.
- Mahesh Reddy Kukunooru is with the Department of Computer Science, Old Dominion University, Norfolk. E-mail: mkuku002@odu.edu.

in a particular season.






New England Patriots Schedule - 2016

Season: 2016 [Add to Calendar](#) [More NFL Teams](#)

2016 POSTSEASON SCHEDULE									
RND	DATE	OPPONENT	RESULT	HI PASSING	HI RUSHING	HI RECEV			
DIV	Sat, Jan 14	VS  Houston	W 34-16	Brady 287	Lewis 41	Edelman 1			
CONF	Sun, Jan 22	VS  Pittsburgh	W 36-17	Brady 384	Blount 47	Hogan 180			
SB	Sun, Feb 5	@  Atlanta	W 34-28 OT	Brady 466	Blount 31	White 110			

2016 REGULAR SEASON SCHEDULE									
WK	DATE	OPPONENT	RESULT	HI PASSING	HI RUSHING	HI RECEV			
1	Sun, Sep 11	@  Arizona	W 23-21	Garoppolo 264	Blount 70	Edelman 6			
2	Sun, Sep 18	VS  Miami	W 31-24	Garoppolo 232	Blount 123	Bennett 11			

TEAM LEADERS

PASSING YARDS	RUSHING YARDS	RECEIVING YARDS	TACKLES	INTERCEPTIONS
 3554	 1161	 1106	 92	 4

PASSING														RUSHING										
NAME	ATT	COMP	PCT	YDS	AVG	YDS/G	LONG	TD	INT	INT%	SACK	YDS	RATE	NAME	ATT	YDS	AVG	LONG	20+	TD	YDS/G	FUM	FUM%	LDN
Tom Brady	432	291	67.4	3554	8.2	288.9	79	29	6.5	2	0.5	15	112.2	LeGarrette Blount	299	1161	3.9	44	7	18	72.6	2	1	67
Jimmy Garoppolo	63	43	68.3	502	8.0	81.2	37	4	6.3	0	0.0	3	113.3	Dion Lewis	64	282	4.4	15	0	0	40.6	1	0	14
Jacoby Brissett	55	34	61.8	400	7.3	116.0	58	0	0.0	0	0.0	6	83.9	James White	35	166	4.3	16	0	0	10.4	0	0	8
Totals	550	368	66.9	4456	8.1	269.3	79	33	3.8	2	0.4	24	109.3	Jacoby Brissett	16	83	5.2	27	1	1	27.7	2	1	5
Opponents	596	368	61.7	4033	6.8	237.9	66	21	3.5	13	2.2	34	84.4											

Figure 2: Team schedule and comparison of player statistics (<http://www.espn.com/>)

3 DATASET

A lot of football data is available on the sports websites like ESPNs, which provide extensive data covering all the matches from the beginning of football that also include several changes that took place over the years like a team moving from one city to the other and players being drafted by other teams.

Most of the data required for our visualization is obtained from ESPN website through which the data is crawled to get the required attributes in a format required that makes it easier to draw all the charts. Out of many libraries that we have in several languages, we chose Beautiful Soup library in python as it suits effectively to the dataset that we are dealing with. Later with the data that is gathered there were a lot of entries where the data doesn't match with the attribute, where there is a need of cleaning the data, which we did using a data cleaning and refining tool named Open Refine.

For this visualization, we confined ourselves to the NFL data from 2006 to 2016 as the data sources doesn't provide us with all the statistics of each game happened before. For all the years in between the years 2006 and 2016, the data being collected includes the list of super bowl winners of every season; statistics of games played by each team with all the other teams in a season like total yards, first downs, possession, quarter scores of a team and also the player statistics like passing yards of a quarterback, rushing yards of a rush back or the number of sacks of a defender. The following attributes can be derived from the data that is being crawled: number of wins of a team in a season, number of losses of a team in a season, No of wins/ losses in the each of the Quarters.

4 VISUALIZATION

The interface which we are presenting can be mainly divided into team based and player based visualizations. The

home page of the interface has a US map showing all the teams that participate in NFL based on their cities represented by a dot. The teams can be further filtered down based on the league, NFC, AFC and both. On hover of a dot, the helmet of that particular team pops up with a smooth transition which shows the team name and the conference it belongs to. The other teams which belongs to the same conference also gets highlighted so that it gives the user a glimpse of the all teams present in that conference. Below the US map, user can find a dropdown to select a season to find out the Super Bowl winner (NFL champion) of that season.

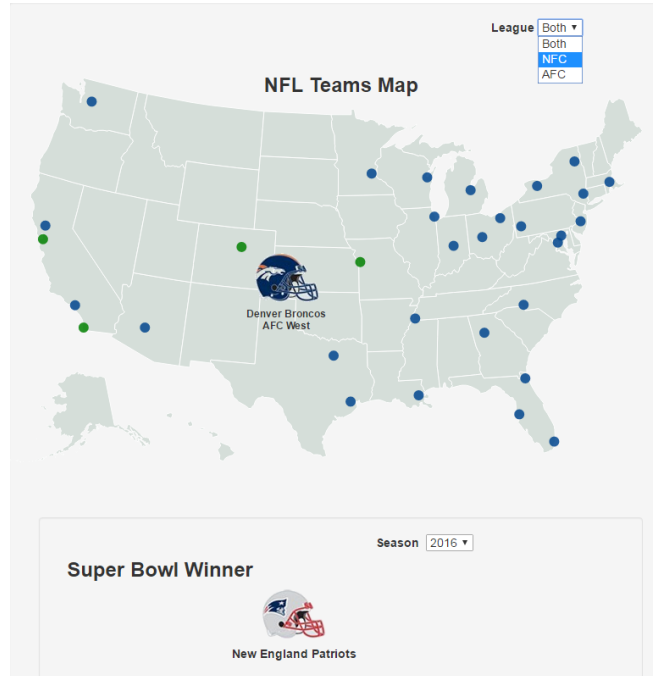


Figure 3: US map showing NFL teams

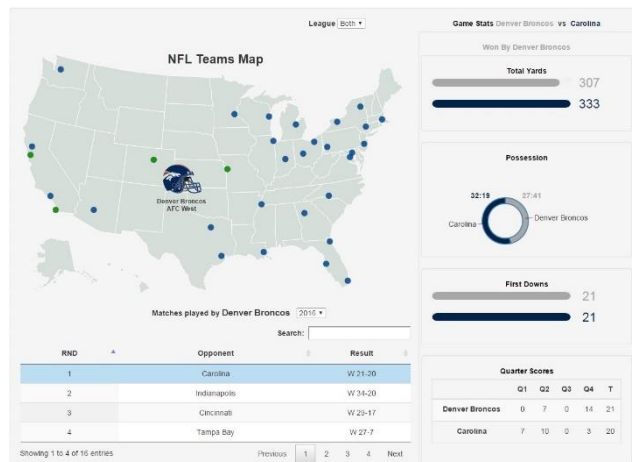


Figure 4: Game stats (Denver Vs Carolina 2016)

4.1 Team based statistics

When the user clicks a dot on US map which represent a team, it navigates to a screen wherein it shows a table that has a list of all the games played by the selected team in a particular season that can be selected from a dropdown. By

default the most recent season, in this case 2016 is selected automatically. On click of any of the game from the table shows the statistics of the game to the right of the map. The statistics of a game being shown include comparison of total yards, possession, first downs and quarter scores of both the teams as shown in the Figure 4 above.

The rightmost section of the screen shows the win/loss trends of the selected team over the seasons using a vertical bar chart, where the category (win/loss) is selected using a radio button at the top. A selection of any of the bar from the above chart would populate a horizontal bar chart which shows the number of quarter based wins/losses as shown in Figure 5s. For example, if a team won 11 games in a season, and among them, it dominated quarter four 7 times. So selection of that season generates a horizontal bar chart which shows the number of games a team dominated a particular quarter given the team went on to win the game.

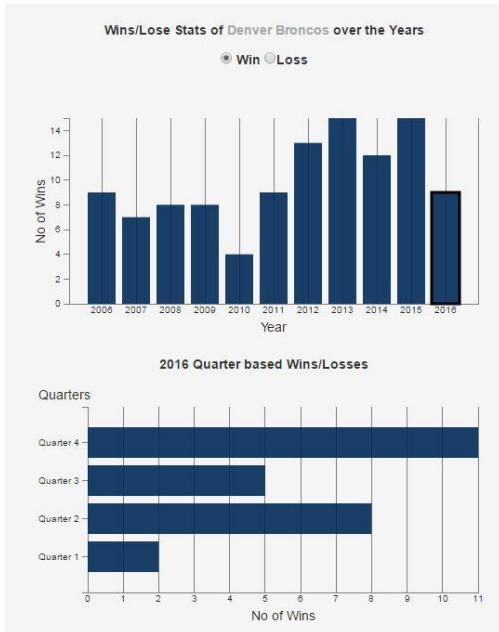


Figure 5: Charts with linked highlighting

4.2 Players based statistics

To the right of US map showing NFL teams on the landing page, there is a section which shows the game leaders based on player positions (Quarter Back, Rush Back, Wide Receiver etc.). A dropdown is provided to select the season to update the game leaders according to that season. If a player is selected, a multiline chart is generated which is juxtaposed next to the player category table that compares the number of yards of the selected player with the yards of the player who has the highest in that season as shown in Figure 6.

When the user clicks the explore button present in the category leaders table generates a radar plot which compares the three players in the list by default. Here all the dropdowns can be changed to select any three players based on the season and the player position. We thought that it is irrelevant to

compare the statistics of a Quarter back with that of a Wide Receiver or a defense player as the player position matters a lot in sports which are spatio-temporal. Radar chart can be used to compare multiple attributes with different scales at a time with the attributes are separated by equal angle and the radius is used for the scale.



Figure 6: Player performance trend

5 DESIGN DECISIONS

The whole article talks about comparing the player and the team statistics which can be thought as an abstraction task. In the multiline chart in the above Figure 6, we are comparing the selected player statistics with that of the one with the highest in that season and in radar chart we are comparing the player statistics with each other which is an abstraction task. For the comparison, we have the following visual encoding:

- Multi-line chart to compare the player statistics with the highest of the season.
- Radar chart to compare different attributes of different players.

We have many interactions between the charts where if the user selects specific sections of a chart.



Figure 7: Radar chart comparing players

6 ANALYSIS

Idiom	Donut Chart (Figure 4)
What: Data	Table: one quantitative attribute(possesion time), One categorical attribute(Team name)
Why: Task	Part-whole relationship
How: Encode	Area marks with angle channel; radial layout
How: Manipulate	Hover – Show()
Scale	2 teams

Donut chart is used to show the division of time between the two teams. As the total time is 60minutes, which is a constant value, it can be easily shown using a donut chart. On hover of the donut chart the possession stats of the teams are presented to the user.

Idiom	Multi-line Chart (Figure 6)
What: Data	Table: one quantitative attribute(Yards), One ordered key attribute(Year)
Why: Task	Show trends of player yards over years
How: Encode	Dot chart with connection between dots
How: Manipulate	Hover – Show respective year and yards of player
Scale	Key attributes – depends on how many years of data is available for that player, maximum of 15(2002-2016)

To know how a player evolved over the years, we can just compare the performance of the player with the year leader's stats. In this case we considered only the Yards as the attribute being compared. For example in Figure 6, when Drew Bees is selected, his yards over the seasons are compared with the leader's yards in that particular season. We have observed that Adrian Peterson, who is the important player of Minnesota Vikings passing yards over the years fluctuated, which is not a usual trend that we observed; that is the players who performs well in a season normally does well in the consecutive years, which failed in case of Peterson, so we looked into the reason to find that he was injured in those seasons.

Idiom	Radar Chart (Figure 7)
What: Data	Table: one quantitative attribute(attribute value), Two categorical attribute(Player name, attribute)
Why: Task	Compare different players
How: Encode	Area marks with angle channel;
How: Manipulate	Hover – (Show statistics of a particular player)
Scale	3 players can be compared with a maximum number of 6 attributes

This is the plot that we used to compare the performances of more than 2 players at a time. A user is allowed to choose a season, position of the players, different teams, and different players to be compared. This comparison is made considering around 6 attributes of each player, each of which has different scales. The main reason behind using the Radar plot but not the multiline chart is because the attributes taken into account has different scales.

Idiom	Vertical Bar Chart (Figure 5)
What: Data	Table: one quantitative attribute(No.of wins/ losses), One categorical attribute(Year)
Why: Task	Compare wins/ losses of teams
How: Encode	Line marks, express categorical attribute with horizontal position (wins/ losses),quantitative attribute with vertical alignment
How: Manipulate	Hover-Shows wins/ losses ,Select – Navigation to Horizontal bar chart
Scale	11 key attributes (11 years- from 2006-2016)

This is the plot drawn mainly to give the user an ability to see how a team is performing over the past several seasons. On selection of any of the bar in this chart (selection) would update (Hightlight) the Underlying horizontal bar chart which shows the quarter based trends.

Idiom	Horizontal Bar Chart (Figure 5)
What: Data	Table: one quantitative attribute(No.of wins/ losses), One categorical attribute(Quarter)
Why: Task	Compare quarter based wins/ losses
How: Encode	Line marks,express quantitative attribute with horizontal position (wins/ losses),categorical attribute with vertical alignment
How: Manipulate	Hover-shows wins/ losses
Scale	4 key attributes (4 quarters)

7 CONCLUSION

While going through the visualizations, we found some interesting aspects like Detroit Lions went winless in 2008, which was considered to be one of the worst performances in the NFL history. Later when we researched over the web to find the reason of such performance, we found that the main player John Kitna, the quarter back of the team was injured in the 5th week itself. Also a naive user can now easily explore many of the interesting insights like how a player's performance changed over the years using this platform.

8 FINAL THOUGHTS

Initially when we were crawling through the web for the desired dataset, we thought to keep to entire data in a single JSON object structure that we designed. But later when we started implementing the charts, it became hectic to iterate over such a large JSON object for every single plot as each of them requires new set of attributes. So, as a solution we divided the JSON object into small objects which are easy to access and process.

9 ACKNOWLEDGEMENTS

We wish to thank Dr. Michele Weigle, Associate Professor, Computer Science at Old Dominion University for her continual efforts in providing some interesting ideas to get to the final system throughout the project.

10 REFERENCES

- [1] [Multiline Chart Animated](#)
- [2] [Radar Chart](#)
- [3] [Radar Chart with different scales](#)
- [4] [Horizontal Bar Chart](#)
- [5] [Progress bar](#)
- [6] [Donut chart with animation](#)
- [7] [Basic US Map](#)
- [8] [http:// www.espn.com/](http://www.espn.com/)