

Final Report

Data Visualization 2020

| | |
|-----------------|----------|
| Berk Atıl | s2458691 |
| Joris Jager | s2169770 |
| Melike Oğuz | s2128039 |
| Sjoerd van Veen | s2136090 |
| Lukas Vugts | s1842218 |

Soccer

Soccer is one of the, if not the biggest sports in the world and comes with a lot of data. Although many of the statistics, like player- or club statistics, can be found all over the internet, there are still some subjects interesting to investigate further.

For the assignment the behavior and loyalty of fans and their effect in home or away games was investigated. The effect of the referee(s) and amount of broadcasting revenue is also looked at. Finally a comparison between men and women is made with regards to transfer fees and amount of viewers. All of the visualizations are brought together via a websiteⁱ with an interactive stadium. There are five dots placed on a stadiumⁱⁱ, each connected to a different topic relevant to their placement on the image.

Fans

These visualizations were all made with Tableau. Pre-processing was done with Excel.

World Cup attendance vs. Total Prize Money

The visualization compares the average attendance per game during the FIFA Men's World Cup with the total prize money. While the attendance seems to grow steadily, the prize pool has greatly increased over the last few tournaments.

The graph showing the prize pool is colored, from yellow to green. This was chosen because green can be associated with both money and football. The gradient is added as an extra way of visualizing the increase in the prize pool. A trend line is also added to make it easier to see the increase in average per game attendance.

The average attendance was found on [statista.com](https://www.statista.com), which used data from [sport.de](https://www.sport.de), a German sports news outletⁱⁱⁱ. The prize money was also found on [statista.com](https://www.statista.com), which got the information from a FIFA document^{iv}.

'Big Five' average game attendance vs European football market size

As already suggested in the World Cup visualization, more and more money is going around in the football market. This visualization shows that the size of the European football market does not have a great correlation with the average game attendance in the five greatest national competitions in Europe (and possibly the world).

Green is again used as color for the graph displaying money. To make the increase in market size more apparent, an area graph is used. This also makes the visualization more clear. Blue is again used to visualize the average game attendance. This is done as it is a neutral color.

Both datasets were acquired from statista.com, a report by Deloitte was used to make both datasets^v ^{vi}.

Football-related arrests across English professional football competitions

This visualization allows us to compare the total attendance and football-related arrests across multiple English football leagues. The amount of arrests per 100,000 attendees is also visualized.

The visualization shows multiple different leagues. These are all represented by different symbols. This allows the viewer to compare between different leagues or different seasons within the same league. The color of the symbols is based on the amount of arrests per 100,000 attendees. The color is a gradient from green to red, where red means more arrests. The dataset used was published by the UK Home Office^{vii}.

Referees

In soccer, referees are a very important component and may have effects on the results. In order to understand this, different years and different competitions were analyzed through perspectives. All data is retrieved from WhoScored^{viii} (This link is the example for Premier League)

Fouls

It can be interesting to compare the average number of fouls per referee. Therefore, in order to show this bar charts were used because it is very easy to see the difference.

Premier League → the average is usually around 21 fouls per game. The most interesting referee is Simon Hooper in 2017-2018 because in his games there are approximately 29 fouls per game although the average is 2.

Bundesliga → Usually average is around 28, but there is Patrick Ittrich which is similar to Simon Hooper and Manuel Gräfe's average is 20 which is far away from the average. Also, in 2018-109 average fouls decreased. Maybe the foul standards in the Bundesliga gets close to the Premier League.

La Liga → Average is around 28 and there is no interesting referee.

Serie A → Average is around 27 and in 2017-2018 at Luca Banti's matches the average is around 19. To sum up, it can be inferred that the referees in Premier League have higher standards on fouls

Cards

The average yellow card per game is higher in Serie A and La Liga

Usually, every year and every league, there are a few referees whose average number of yellow cards 1 above the average. This is a significant difference because usually the average number of cards is around 3-4. Hence, when the average is 4, having a value 5 makes an important difference.

The average red card per game is also lower in Premier League (0.1, unlike 0.2)

In Serie A, there is no referee who has not used Red Cards in any season, there are in other leagues though.

Wins

The rates vary a lot from one referee to another. Hence, it is possible that the referees are also important about the result of the matches

It is obvious that usually, the rate of home team wins is the largest. However, there are some interesting results.

For example in the Premier League, there are some referees whose game in, one of the results. does

not exist at all. For example in 2016-2017 there is no draw result in Graham Scott's matches. Also there is one similar example in Serie A in 2016-2017.

Home vs Away

In soccer there exists a phenomenon called home advantage. People claim that playing at your own stadium has an effect on the results. Is this phenomenon real or is it a hoax? To investigate this three parameters were visualized.

The most important parameter is the amount of points per game. The amount of points is an important factor for a good result. To visualize this in a clear way, we decided to make a radar graph. In the radar graph you can clearly visualize the difference of point per game away and points per game home by overlapping the two and making the colors transparent.

The second factor was goals scored. Can you see a difference in goals scored home and away. This shows how different teams tend to play at home. Do they play attractive football and score more goals at home? A pie chart can clearly show the difference between two variables by making them two different colours. By only having two different types(home and away) the difference between the two becomes clear. To give more insight information about goals scored, subgraphs of the separate teams were added.

The last parameter was the amount of goals conceded. Are teams playing more defensive and because of this concede less goals. To make a clear visualisation of these parameters, a pie chart was made. This is because of the easy way to visualize the difference between two variables. Here we also added subgraphs to give more insight information.

To gather all the information about this topic the website soccerstats^{ix} was used. On this website was a clear overview of the performance of teams of different competitions of the past years. They also had a special page of each competition called home/away tables. These tables showed the different results of the teams and they separated it based on home and away. I used the information of five competitions to be able to make a valid conclusion about if home advantage exists or not.

Men vs Women

In the past few years women football started to become more popular. There are still a lot of differences within men vs women football. Different types of data are investigated.

First, the world ranking of the countries around the world based on the amount of points. The first thing that pops out is that the women's world ranking scores are generally higher than the men's ranking scores. Secondly, we can see that the women world ranking top 20 has more variety around the world, meaning the countries are from all around the world. For men's top 20 it is morley distributed within just Europe and Mid- and South Amerika^x. The reason to choose a tree map is that it is very clear within one look which country has the most amount of points. The bigger the rectangles the more points.

Secondly, viewers of the world cup. The world cup is within men's football typically a big thing. For the past few years it has had a lot more attention to women's football as well. In general the men's world cup has a lot more viewers than women's football, when looking at viewers for at least 1

minute, total audience reached, average viewers live match. However, when looking at viewers from digital platforms and the finals the women's football outreaches the men's^{xi}. Bar graphs are chosen to display the differences between men's football vs women's football per category very easy and quickly.

Lastly, the transfer fees. It is generally known that there is not that much money involved in women's football than in men's. To see the real difference, two types of data is looked at. The first one is the amount of transfers made in men's football vs women's over the past two years. There is only data available from 2018 on for the women's football, therefore it was not possible to look at a bigger time period^{xii}.

We can clearly see that the men's football has about 30 times more transfers than the women's. The second one is the actual transfer fees. The differences here are about 15000 times more money in men's football than in women's football. Since the difference is so big, it was hard to display any difference between the two data sets. Because the women's transfer fee seemed like non-existing. Therefore, the circles on such a graph are used.

Media

The value of the European soccer leagues has grown over the past years. The value of the TV rights differs massively per competition, which is why it's investigated.

Five largest Leagues^{xiii}

By comparing the five largest competitions in Europe, namely the Premier League, La Liga, Serie A, Ligue 1 and the Bundesliga, great differences in value can be seen. The line graphs indicate the amount of money in millions each of the clubs receive. In the first graph the amount of money received is compared to the result in the competition. It's clear to see that a better result doesn't mean more money. This is due to the way the distribution is built up since it's based on results of previous years.

Looking at the second graph, here the distribution is sorted by amount of money. For all the competitions a great difference between the top teams and the lower teams can be seen. Although the total value of the competition differs greatly, the decline of distribution is roughly the same for each of the competition. Note: The Bundesliga only has 18 teams competing, thus it stops at 18.

Premier League^{xiv}

The first thing you notice when looking at the previous graph is the enormous amount of money the Premier League has in comparison to the other competition. This difference is the result of the British broadcast channels bidding for the rights to broadcast the premier league. The total value of the TV rights has tripled since 2007-2010. The amount of games has also increased over the past years and is graphed with a line. This shows that the amount of money per game has increased a lot.^{xv}

In the second graph the distribution of the TV money for the 2018/19 season is visualized. This shows all teams receive around 92 million from equally sharing some of the budget. The other differences come from the distribution of the Facility fees, which are based on how often a club's matches are broadcasted. The rest comes from the Merit Payments, based on the position where the club finished. The differences between clubs can be seen by filtering on fee.

In the third graph the total value is visualized by things that can be bought with the immense amount of money. By comparing it to the total value of the Eredivisie^{xvi}, which is approximately €80 million, the difference with the Netherlands becomes clear. The a380^{xvii}, the largest passenger plane in the world, and FC Barcelona^{xviii} are recognizable subjects and therefore chosen to compare the amount of money with.

References

-
- ⁱ <http://dataviz.lukasvugts.nl/>
- ⁱⁱ https://www.psv.nl/upload/357645f1-d381-4487-99fc-e5b6abcf07ba_philipsstadion-lucht-ronaldotter2.jpg
- ⁱⁱⁱ <https://www.statista.com/statistics/264441/number-of-spectators-at-football-world-cups-since-1930/>
- ^{iv} <https://www.statista.com/statistics/328497/fifa-world-cup-prize-money/>
- ^v <https://www.statista.com/statistics/261213/european-soccer-leagues-average-attendance/>
- ^{vi} <https://www.statista.com/statistics/261223/european-soccer-market-total-revenue/>
- ^{vii} <https://data.gov.uk/dataset/e74d7ef0-ac2a-46c6-802a-935882284bab/statistics-on-football-related-arrests-and-football-banning-orders>
- ^{viii} <https://www.whoscored.com/Regions/252/Tournaments/2/Seasons/7811/Stages/17590/RefereeStatistics/England-Premier-League-2019-2020>
- ^{ix} <https://www.soccerstats.com/homeaway.asp?league=netherlands>
- ^x <https://www.fifa.com/fifa-world-ranking/ranking-table/men/rank/id11909/>
- ^{xi} <https://www.fifa.com/worldcup/news/more-than-half-the-world-watched-record-breaking-2018-world-cup>
<https://img.fifa.com/image/upload/rvgxekduqpeo1ptbgcng.pdf>
<https://www.cnn.com/2019/07/10/us-viewership-of-the-womens-world-cup-final-was-higher-than-the-mens.html#:~:text=According%20to%20a%20statement%20from,a%2022%25%20U.S.%20viewership%20boost>
- ^{xii} <https://resources.fifa.com/image/upload/women-transfers-in-itms-september-2018.pdf?cloudid=huljca9blppbjl4wpbh%E2%80%8B> <https://resources.fifa.com/image/upload/global-transfer-market-report-2019-women.pdf?cloudid=dm3mxtl95tjrufvuohwn>
<https://resources.fifa.com/image/upload/global-transfer-market-report-2019-men.pdf?cloudid=x2wrqjstwjailnncnod>
- ^{xiii} [Premier League Bundesliga Ligue 1 La Liga Serie A](#)
- ^{xiv} <https://www.premierleague.com/news/1225126>
- ^{xv} <https://www.sbibarcelona.com/newsdetails/index/403>
- ^{xvi} Eredivisie [Image Amount](#)
- ^{xvii} A380 [Image Amount](#)
- ^{xviii} FC Barcelona [Image Amount](#)