**Data Visualization**

FOOTBALL DATA ANALYSIS

Zakeer Ahmed Shaik

U62607023

**INTRODUCTION**

Football clubs spend millions on player transfers in the hope of winning their country’s league or Champions League. We know that not every transfer is successful. In fact, most of the transfers were failures. They pay huge amounts to acquire the player, and yet they do not win the local league. In a few cases, they even put their highly valued players on the bench.

**Ambitiousness of the Project**

 Teams pay players weekly wages in the thousands of dollars, and if they feel that the player is unfit for the formation, they simply put him on the bench. They are still supposed to pay him the full amount as per the contract, even though he is not being played.

 Neymar was sold to PSG for $263 million. PSG has been winning League 1 for a long time. The main goal behind acquiring Neymar is to win the Champions League. Keeping aside the fact that they paid $260 million for Neymar, they never won the Champions League.

 While Atletico Madrid built an entire, huge, and beautiful stadium for $300 million, which cost only $40 million more than Neymar’s transfer value. Stadium filling will be the main source of income for most of the clubs.

So, this project mainly focuses on teams’ performance compared to the money spent on players.

* Comparing teams’ total wages vs their performance.
* Comparing teams’ total value vs their performance.

Did huge money help big teams win their leagues?

Is money really a big factor in winning cups?

**METHODOLOGY**

1. Player details and attributes

* The data source is Kaggle, which addresses almost every aspect of each player.
* This table has different attributes like name, age, club, value, wage, and how good each player is at different aspects like shooting, dribbling, etc.
* A lot of data cleaning has been done for the columns like club name, value, wage, etc.
* [Data Source](https://www.kaggle.com/datasets/stefanoleone992/fifa-22-complete-player-dataset).

1. Teams

* This table consists of club names from the top 5 leagues and their league IDs.
* This table was self-created with the help of reliable sources on the Internet.

1. Leagues

* This table consists of the names of the top 5 leagues and the countries they belong to.
* This table was self-created with the help of reliable sources on the Internet.

1. Managers

* This table consists of the names of the managers of the top 5 leagues.
* This table was self-created with the help of reliable sources on the Internet.

1. Stadiums

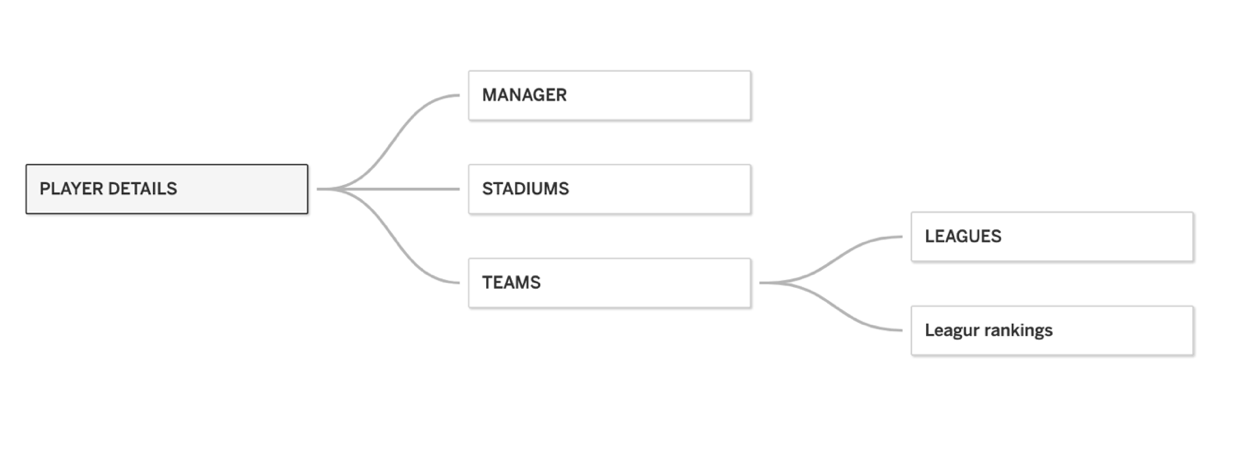
* This table has different stadiums the club owns, their country, and their capacities.
* This table was self-created with the help of reliable sources on the Internet.

1. League Rankings

* This table contains the league table positions of the clubs in their respective leagues.
* The data is collected from sources like
  1. <https://www.premierleague.com/>
  2. <https://www.laliga.com/en-GB/laliga-santander/standing>
  3. <https://www.espn.com/soccer/standings/_/league/ita.1>
  4. <https://www.ligue1.com/>

 All these tables are joined for analyzing different aspects of the football world. These tables are joined based on foreign keys in the tables.

Five links have been used to make joins between these tables.



**ANALYSIS**

**Total team wage vs Team Rankings**

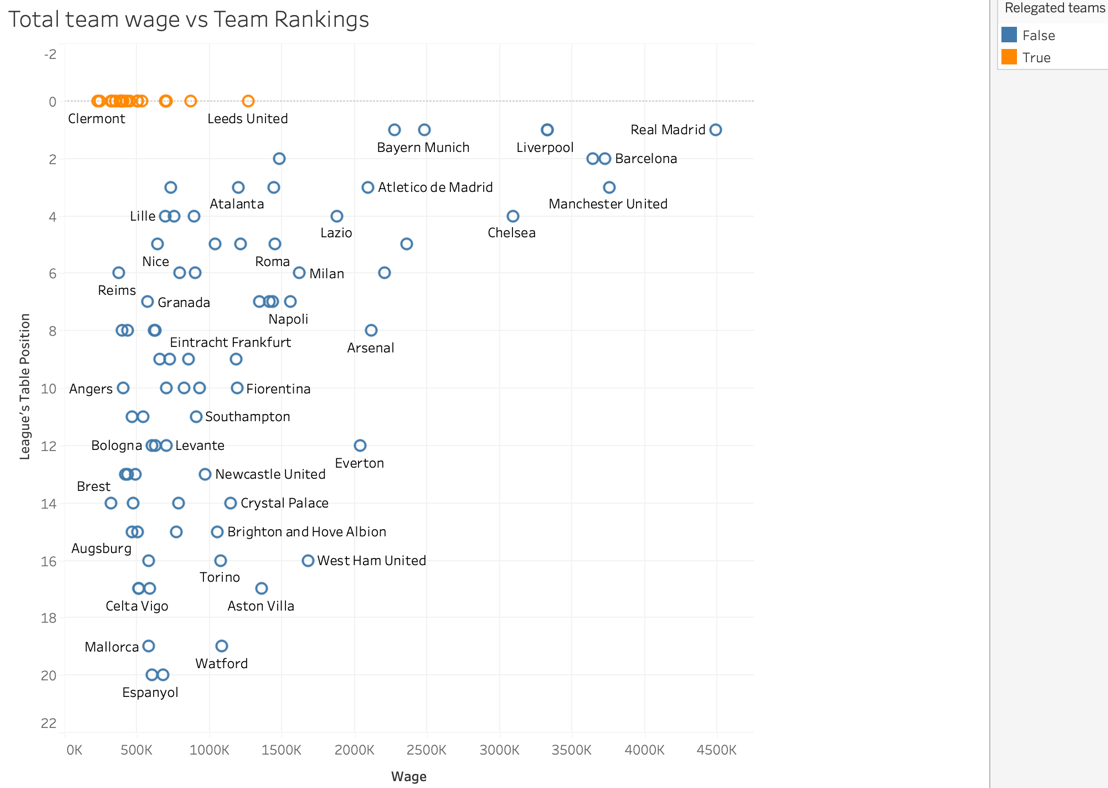
This visualization is one of the key ones in the project, as it plots the total amount a team is spending on salaries, including all the benched players, versus the positions in their respective leagues.

So, according to this, we can see that the team that is to the most right and bottom has the worst management of money and performance. For example, we can see that Everton is paying 2 million in wages and yet remained at the 12th position in the table, obviously not qualifying to play in the Champions League.

The team that is top and to the most left is very good at managing money and performance. For example, we can see that Rennes paid a total wage of 750k and ended up in third place in the table. They are even qualified to play in the Champions League.

Real Madrid paid 450,000 in wages, and they won the league, qualified for the Champions League, and even won the Champions League. Even though they spent huge amounts, they had excellent results.

The teams on X-axis are relegated teams.



**Total team Value vs Team Rankings**

This visualization is also one of the key ones in the project, as it plots the total value of players on a team, including all the benched players, versus the positions in their respective leagues.

 So, according to this, we can see that the team that is to the most right and to the bottom has the worst management performance, even with highly valued players. For example, we can see that Everton has 450 million worth players and yet remained at the 12th position in the table, obviously not qualifying to play in the Champions League.

 The team that is top and to the most left is very good at managing performance even with very low value players. For example, we can see that Rennes ended up in third place in the table with a team value of around 220 million. They are even qualified to play in the Champions League.

 Real Madrid has a team worth 1250 million, and they won the league, qualified for the Champions League, and even won the Champions League. Even though they spent huge amounts, they had excellent results.

In the case of Manchester City, the team is worth 1300 million. They did not win the league and never won the Champions League.

A picture containing graphical user interface

Description automatically generated

**Managers who made their teams reach UEFA**

Here is a chart with the names of managers who sent their teams to the Champions League. The size of the bubble represents the value of the team. Lower-sized bubbles are managers managing the team with low value. These managers are very good at what they are doing. Given a very good team, their probabilities of winning the Champions League will increase.

Chart, bubble chart

Description automatically generated

***Value, wage, and potential to buy the player***

This is a simple visualization of the potential, value, wage, and team a player is playing for. This is utilized in the dashboard 2.

Table

Description automatically generated

**Player attributes from parameter**

This visualization shows different attributes of a player like their overall rating, value, wage, crossing ability, finishing ability, heading, pass, volley, etc. I created a calculated field and a parameter called "player attributes" with these attributes in it. In the parameter drop-down menu, if we select Wages, it shows the wages of players in a horizontal bar chart.

The first picture below is for the visualization, and the second one is a parameter used as a filter.

Chart, bar chart

Description automatically generatedGraphical user interface, text, application

Description automatically generated

**Avg of different Attributes according to player position**

This graph shows the average of attributes passed through the parameter for different player positions. With this, you get information on which player positions are being highly paid and which player positions are highly valued, among so many other attributes according to positions.

Chart, bar chart

Description automatically generated

**Number of players in League and clubs (hierarchy)**

So many teams buy a large number of players and keep paying them huge amounts even though they are supposed to play only 14–16 players, including substitutes, in a game. This costs teams a lot, ends up spoiling the form and value of the player, and increases unnecessary spending. The teams in red are the ones with more than 35 players.

This is a drill down visualization that shows the number of players in each league, and when drilled down to one level, it shows team names in the leagues and their total players.

Chart, bar chart

Description automatically generated

**Country, stadium, and Capacities**

This visualization shows the seating capacity of stadiums. Stadium capacity is the main source of income for the team, along with jersey sales. This is a drill-down visualization that shows seating capacity country-wise and, when drilled down to one level, shows the names of stadiums in those countries and their capacities.

Chart, bar chart

Description automatically generated

**Count of players generated by country**

This is a simple map visualization that shows the number of players generated from each country. The darker the color, the higher the number of players from that country. We can see that it is highly concentrated in countries like Spain, France, Germany, Brazil, and Argentina.

Map

Description automatically generated

**Players percentage by each country**

This is a map visualization that is similar to the above one. It compares all the countries’ players percentages. ((count of players/ total population) \*100)

Map

Description automatically generated

**Average rating of players from different countries**

This is also a map visualization. This shows average ratings of players in country. The darker the color, higher is the average rating.

Map

Description automatically generated

**Team's performance vs money attributes – Dashboard 1**

This is a dashboard that answers the research questions. This is a mixture of the first two visualizations. The dashboard is interactive. As clearly shown, winning is not directly proportional to wage or value.

Graphical user interface, application

Description automatically generated

**All attributes for buying a player in the market - Dashboard 2**

This is a dashboard that combines a few of the worksheets created above. I have made this dashboard to make transfers easy. The first insight addresses nationality. If you want to buy a player belonging to any country, you can click on the country, and the rest of the two visualizations filter themselves up and display the players of that country. This means they are interactive. Now, you can see the players and their current club’s name, potential, value to buy, and salary to pay in the bottom right insight. If you select the name of any player out of those, you can see his attributes in the bottom left chart of the dashboard.

 This way, you will have all the information, his attributes, and his value and wage. You can sort the values and see the highest-paid, most-valued, or most-potential players.

 This can be done in a reverse manner as well. If you select any player, the map will show you which country he is from.

A picture containing timeline

Description automatically generated

**Conclusion**

Did huge money help big teams win their leagues?

Huge amounts of money really helped many teams win leagues and the Champions League. But there are also other factors that impact the outcome. Teams like Arsenal, Everton, and West Ham United have big players getting huge salaries and are still not able to win their league or enter the top four positions. Teams like Rennes, Atlanta, and Borussia Dortmund have teams with lesser wages and yet managed to be in the top four and enter the Champions League. The graph is not straight line in the plot (wage vs rankings). Huge money may bring them closer to cups, but it does not necessarily win them any.

Is money really a big factor in winning cups?

Money is, of course, a huge factor in winning cups. Teams with lesser wages and value managed to enter the top 4 league positions and managed to enter the Champions League but won neither of those. A team like Real Madrid, which has the highest wages and highest value, has won both the league and the Champions League. With this, we can surely say that money is one of the biggest factors in winning cups. But it is not directly proportional. Winning cups depends on so many other factors, like team strategies and players' mentalities, which completely depend on the manager. There are teams that reached UEFA even with the smallest teams. I would say that manager is the key to winning cups.

Additional Research Questions:

* What can be done additionally for teams like Everton, who spend a lot of money and still do not get results?
* So many teams are still winning leagues without spending huge amount. What can be done for the teams’ cost cutting? For example, PSG and REAL MADRID spend a lot for almost same results.

**References:**

* <https://www.kaggle.com/code/ekrembayar/fifa-data-analysis-visualization>
* <https://rstudio-pubs-static.s3.amazonaws.com/730030_c88737c416c743aaa6f85e032a853928.html>
* <https://pianalytix.com/data-visualisation-and-analysis-on-fifa-2020/>