Introduction

- As you all know the craze for football is increasing day by day, being a football fan, I have decided to implement a novel visualization on a dataset which I would love working with.
- I have tried to visualize the goals scored by the top 10 football clubs in UCL (UEFA Champions League) in each season since 1994.
- I have used 2 datasets to visualize this idea i.e. firstly https://www.kaggle.com/datasets/bakar31/ucl-statistics/code?select=ucl_stats.csv dataset and the other dataset "Standings.csv" was web-scraped using python from https://www.transfermarkt.co.uk/ website which has the standing of each club from 2000 to 2022.

Tool used:

- Tableau was used to visualize our data by combing worksheets into a single dashboard.
- Python was used to web scrap the data and perform preprocessing on it (such as to consider only the top 10 teams)

Dataset Type/Data Type

- Dataset 1 contains the UCL stats of all the teams from 1994-2022.
- The dataset type here is TABLES and the attribute type of every attribute is shown in Table 1.
- Dataset 1 is preprocessed ie only the top 10 teams were selected in Europe now like Real Madrid, Bayern Munich, Barcelona, Manchester United, Chelsea, Paris Saint-Germain, Arsenal, Juventus, Liverpool and Manchester City.

Attributes	Attribute Type
Year	Categorical (ordinal)
Team	Categorical
Goal scored	Quantitative
Goal Conceded	Quantitative
Wins	Quantitative
Loses	Quantitative

Table 1

- Dataset 2 (standings.csv) contains the standing of each of these teams from 2000 to 2022
- Here, the dataset type is TABLES and the attribute type is given in table 2.
- To visualize the data in the required way, I made a derived attribute from standing as Numerical Standing.

Attributes	Attribute Type
Year	Categorical (ordinal)
Team	Categorical
Standings	Categorical (ordinal)

Table 2

Visual encoding channel and idioms

- Tasks are basically comprised of action for a target, where we must ask the question **'Why'** we must visualize.
- Action can be defined as the verb to define the goals of the user.
- The target is on what action is done.

- We tried to **Discover** the **Distribution** of the goals scored by a team in a year in UCL and also **Compare** the **Association** of standing in a particular year in the League.

1) To **Discover** the **Distribution** of the goals scored by a team in a year in UCL

- I have used **Sankey** based approach to visualize the distribution of the number of goals scored by each team over the years.
- The Sankey chart gives us the perfect flow of goals scored by one team over these years.
- This Sankey idiom is divided into 3 parts, teams, the flow and then the year.
- The teams have been arranged in ascending order of the total number of goals scored by them from 1994-2022 i.e. Real Madrid who have scored the most number of goals are placed at the top whereas, Manchester City who has scored the least are been placed at the bottom
- Teams are been encoded into size ie the size of the team depends on the number of goals scored.
- The team is also encoded into colour ie each team have a different colour.
- The team is related to the year attribute with the help of flow(line) if and only if the team has scored in that particular year.
- The colour of the line of the flow is the same as the colour of the team it represents.
- The size of the flow is encoded into goals scored in that year ie for more goals scored the flow line is thick and thin for less goals scored.
- The third part is the year, The years have been encoded into size by the total number of goals scored by all the teams. We can see that the size of 2018 is the highest as the goals scored are 422 which is the highest amongst other years.

2. **Compare** the **Association** of standing in a particular year in the League.

- I have used Radial Bar Chart to compare the standing of the team over years.
- As I wanted to visualize data that changes over time (attribute "year") and also the length of the bar defines the standing of the team during the year.
- Here, the derived attribute, numerical standing was used to represent the height of the bar.
- This attribute "numerical standing" is encoded into colour where red colour means that the team has an early exit from the tournament and blue means it had a decent finish in the tournament.
- Also, it was encoded into size, which indicates that the higher the number, more is the thickness of the bar.

Interactive Elements

- I have used 2 idioms here which Sankey Diagram and Radial Bar Chart.
- We have a team and year in common in these two idioms.
- I have made the Visualization interactive as

1) When we hover the mouse over the team name in the Sankey chart,

- It highlights the Sankey flows in all the years that the team has scored
- It shows the details such as the number of goals scored, goals conceded, wins and losses in the tooltip.
- It highlights the radial bar chart wrt that hovered team.

2) When we hover the mouse over the flow line in the Sankey chart,

- It highlights the team and the year to which the flow belongs.
- It also highlights the column of the radial chart based on the year of the highlighted team.
- It shows the team and year involved in the tooltip

3) When we hover the mouse over the year in the Sankey chart,

- It highlights the flows that each team that scored in that year.
- It highlights the bar of each team that scored goals in that particular year.
- The tooltips show the total number of goals scored in the selected year.

4) When we select the bar on the radial chart.

- It highlights the corresponding team, flow line and year of the selected year in the Sankey chart.
- The tooltip shows the standing of that team in the selected year.

The strength of the visualization is that it lets us know the legacy of the football team based on how it has performed in the UCL tournament and at what position it has secured each year.

The weakness is that it should have more interaction by modifying the visualization based on clicking some parts, but I feel that what I have done is enough to visualize the dataset. Also, the visualization is not too fancy and messy so the user gets confused by having a first look at the dashboard.

Link to the Video : https://youtu.be/uJQwtVdz3lM

References

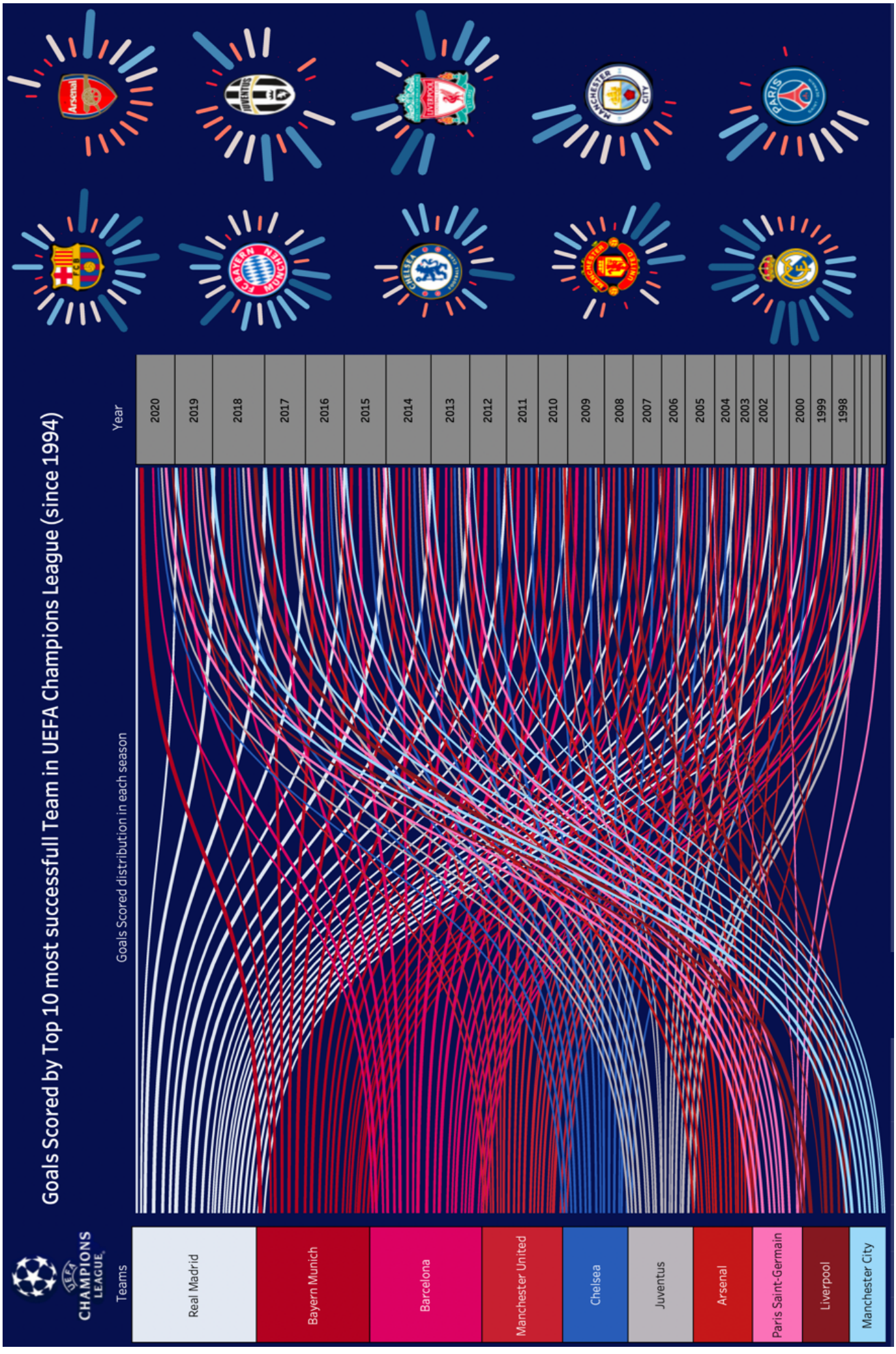
1) Sleeper Ryan “Visual Analytics and Tablaue Blog” 3 Ways to Make Beautiful Bar Charts in Tableau https://playfairdata.com/3-ways-to-make-beautiful-bar-charts-in-tableau

2) Bima Putra Pratama “**How to Make Sankey Diagram in Tableau**” 27th May 2020 https://towardsdatascience.com/how-to-make-sankey-diagram-in-tableau-f5f8730e5962

3) UCL-stats.csv https://www.kaggle.com/datasets/bakar31/ucl-statistics/code?select=ucl_stats.csv

4) UCL Team Standing https://www.transfermarkt.co.uk/

5) Parul Kapoor “The Data School” Jun 30 , 2019 https://www.thedataschool.com.au/mipadmin/creating-sankey-chart-in-tableau/


Goals Scored by Top 10 most successfull Team in UEFA Champions League (since 1994)
Goals Scored distribution in each season
Teams
Year
Real Madrid
Bayern Munich
Barcelona
Manchester United
Chelsea
Juventus
Arsenal
Paris Saint-Germain
Liverpool
Manchester City
2020
2019
2018
2017
2016
2015
2014
2013
2012
2011
2010
2009
2008
2007
2006
2005
2004
2003
2002
2000
1999
1998