Analysis of Football Match Events

Vinicius Jodar S. Costa

INTRODUCTION

- This project used analyzes match events from the biggest 5 European football league: England, Spain, Germany, Italy, France from 2011/2012 season to 2016/2017 season as of 25.01.2017.

- Purpose: Use this data to practice real-world data preparation and analysis with pandas, seaborn and matplotlib.

4 Topics projects

1

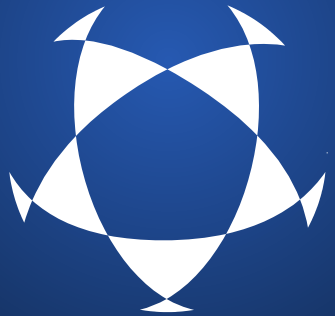
Dataset Structure

Understanding the dataset

2

Data Cleaning

Remove irrelevant informations



3

EDA

Initial investigation of data

4

Data Visualization

Create Graphics and summaries to explore

DataSet Structure

- Datasets used: Events.csv
- More than 900.000 rows of match activity.

- Key columns used:
- event_type: type of event (goal, foul, pass, etc.)
- Player: player responsible for the action.
- event_team: team that performed the event.
- assist_method, is_goal, etc..

Data Cleaning

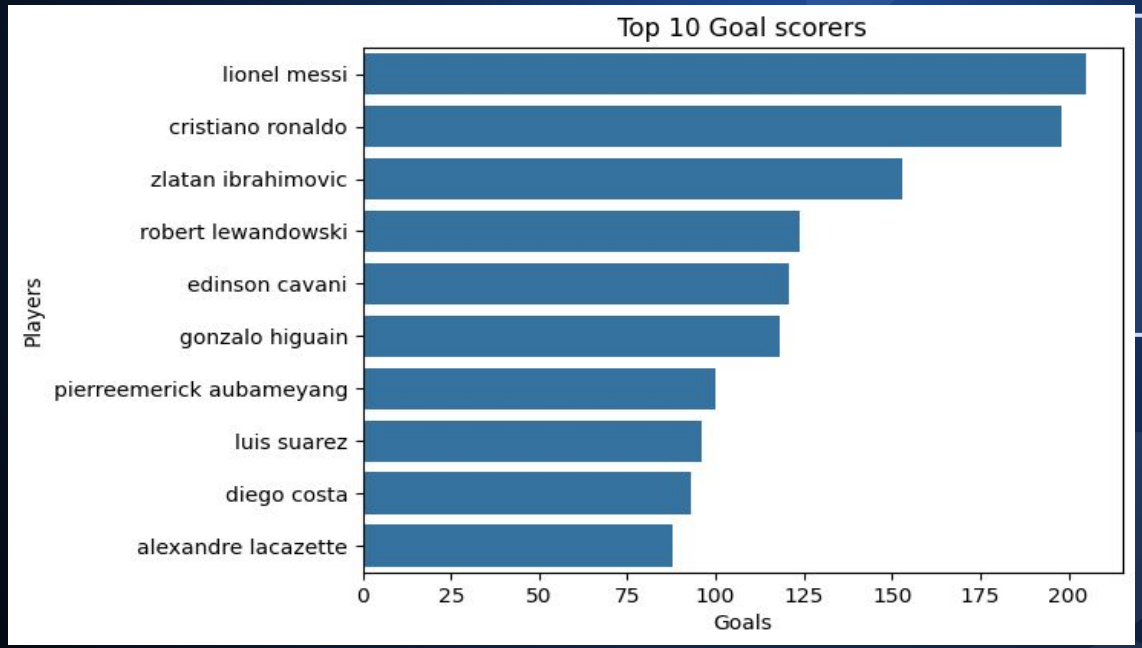
Before analysis, the event data set was cleaned to ensure data quality and relevance.

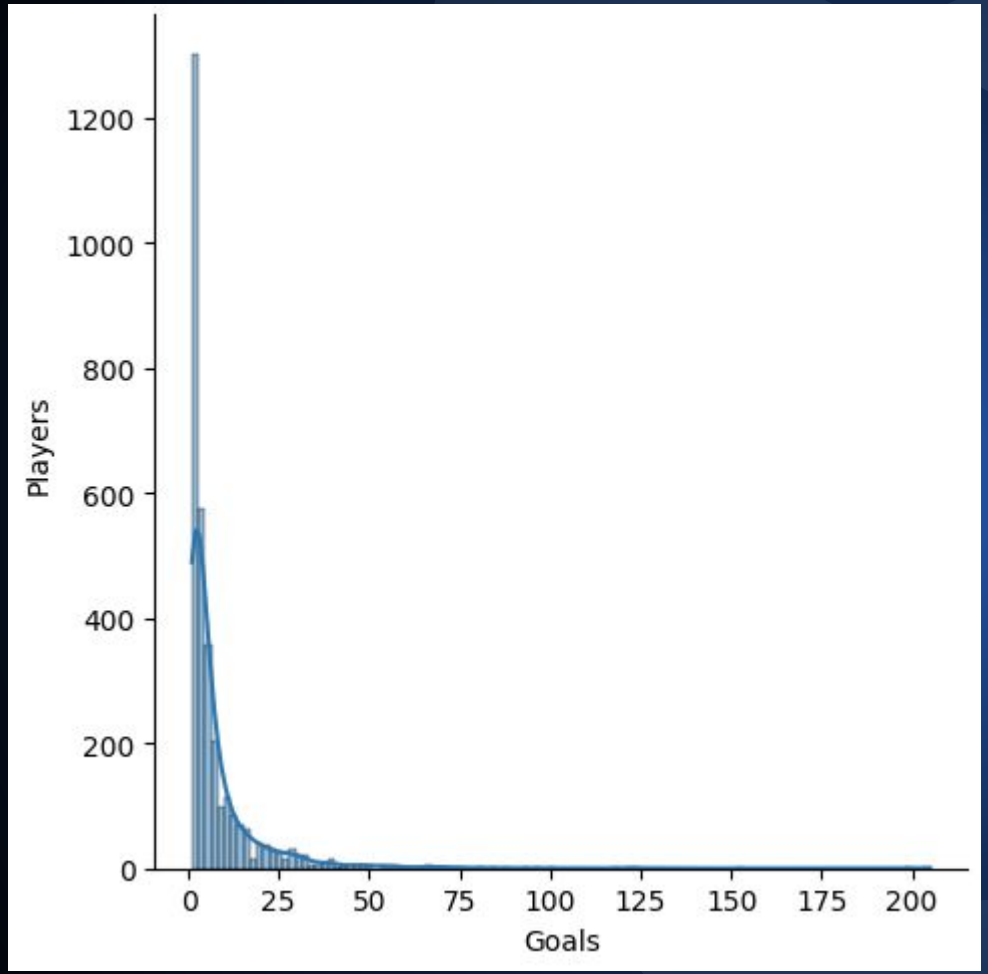
- Removed irrelevant columns.
- Removed duplicate rows.
- Checked and handled missing values.
- Used .info(), .isnull(), and .nunique() to guide decisions.

	id_event	event_type	side	event_team	player	is_goal	assist_method
0	UFot0hit1	1	2	Hamburg SV	mladen petric	0	1
1	UFot0hit2	2	1	Borussia Dortmund	dennis diekmeier	0	0
2	UFot0hit3	2	1	Borussia Dortmund	heiko westermann	0	0
3	UFot0hit4	3	1	Borussia Dortmund	sven bender	0	0
4	UFot0hit5	8	2	Hamburg SV	gokhan tore	0	0
...	...	...	...	...	...	...	...
941004	z5L2OT5E123	8	2	Sampdoria	lucas torreira	0	0
941005	z5L2OT5E124	2	2	Sampdoria	NaN	0	0
941006	z5L2OT5E125	1	2	Sampdoria	fabio quagliarella	0	1
941007	z5L2OT5E126	8	1	Atalanta	alberto grassi	0	0
941008	z5L2OT5E127	3	2	Sampdoria	pedro pereira	0	0

EDA (Exploratory Data Analysis)

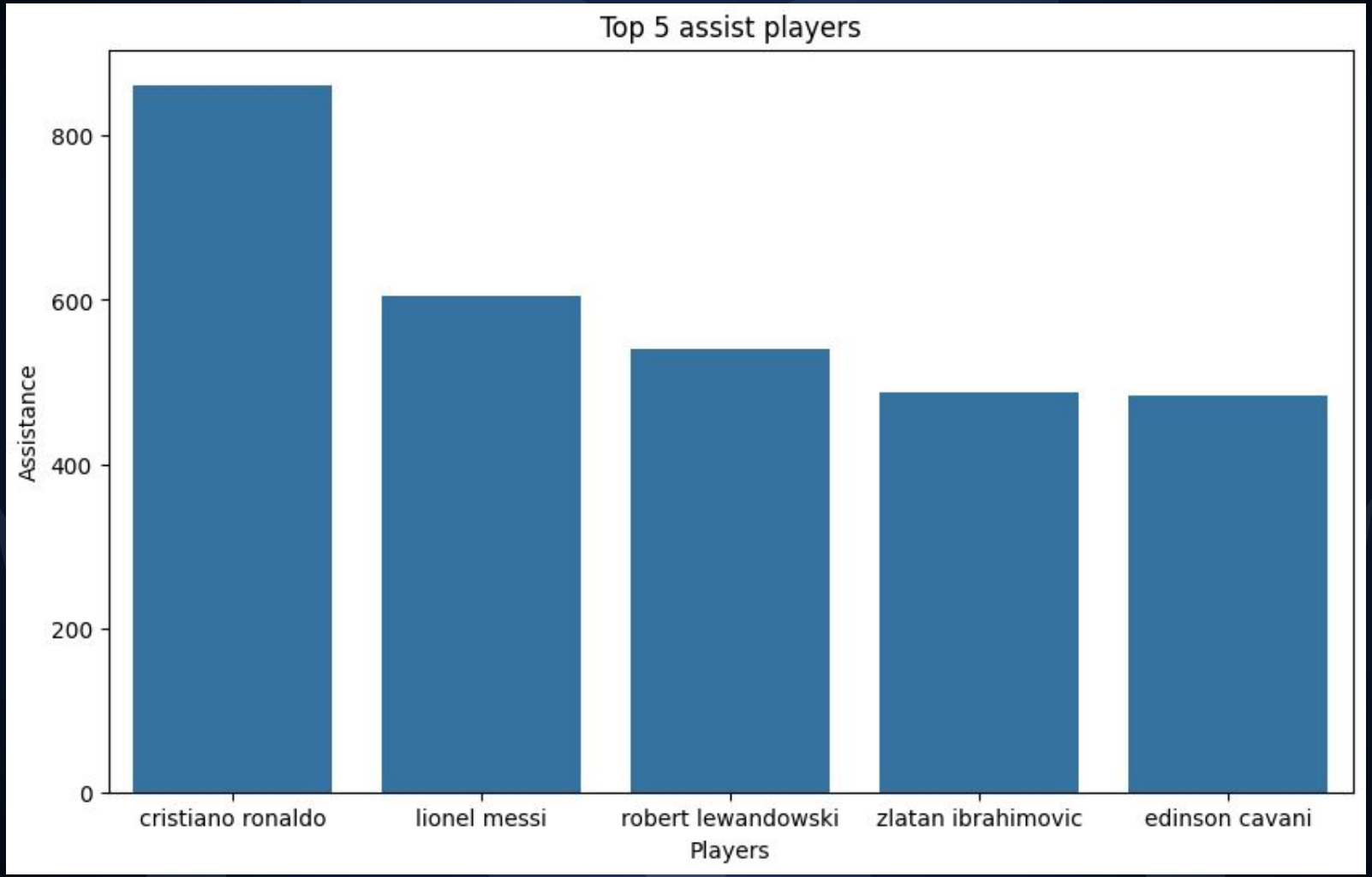
I filtered the data to include only goal events, then counted how many goals each player scored.
The bar chart shows the top 10 goal scorers based on that count.

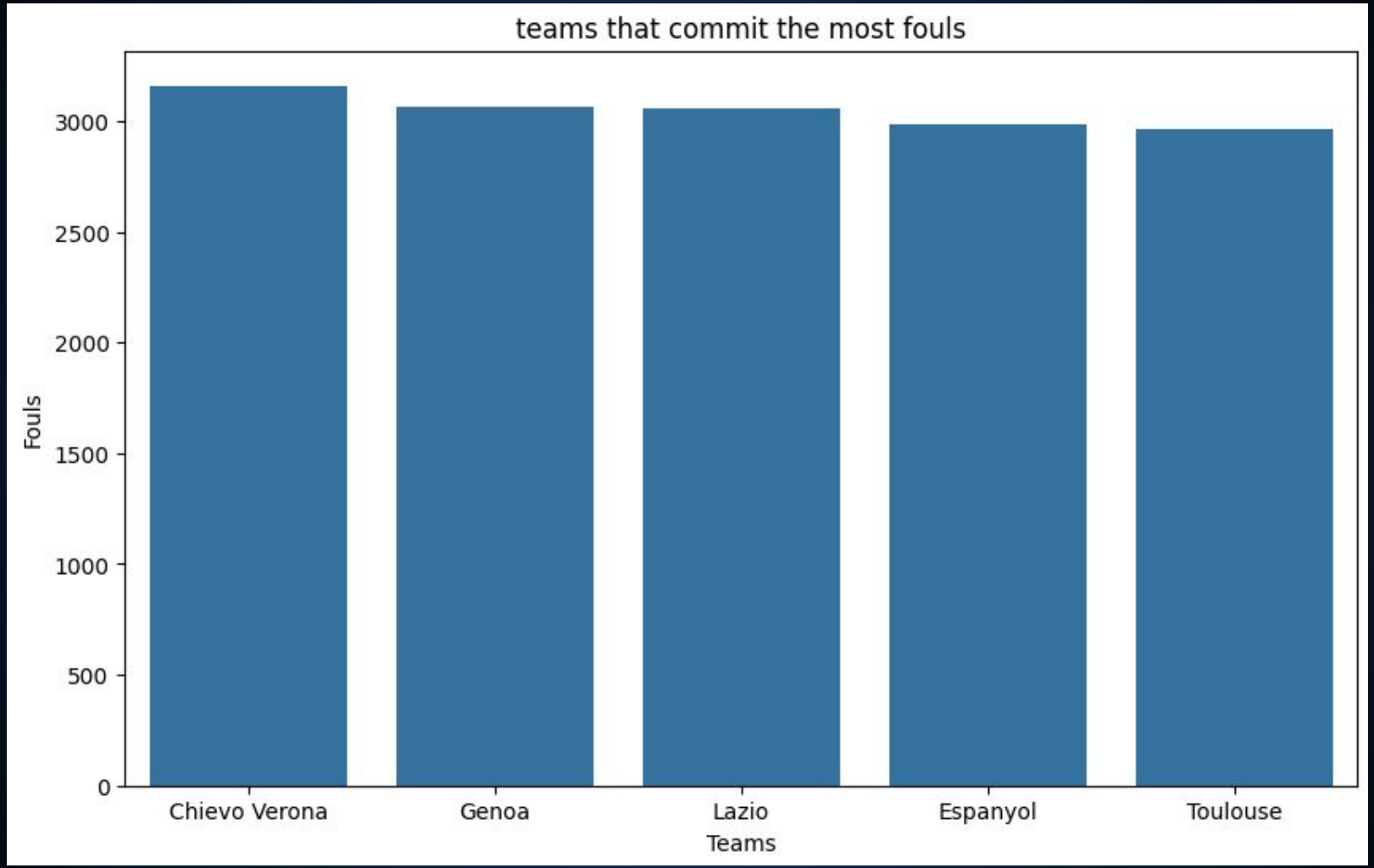


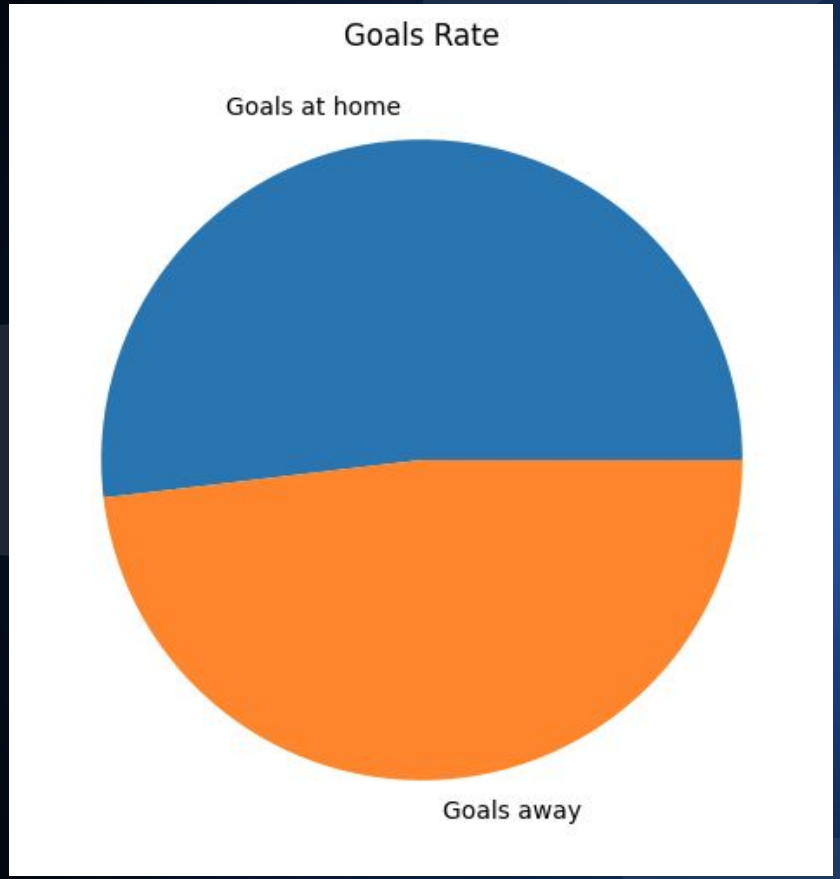


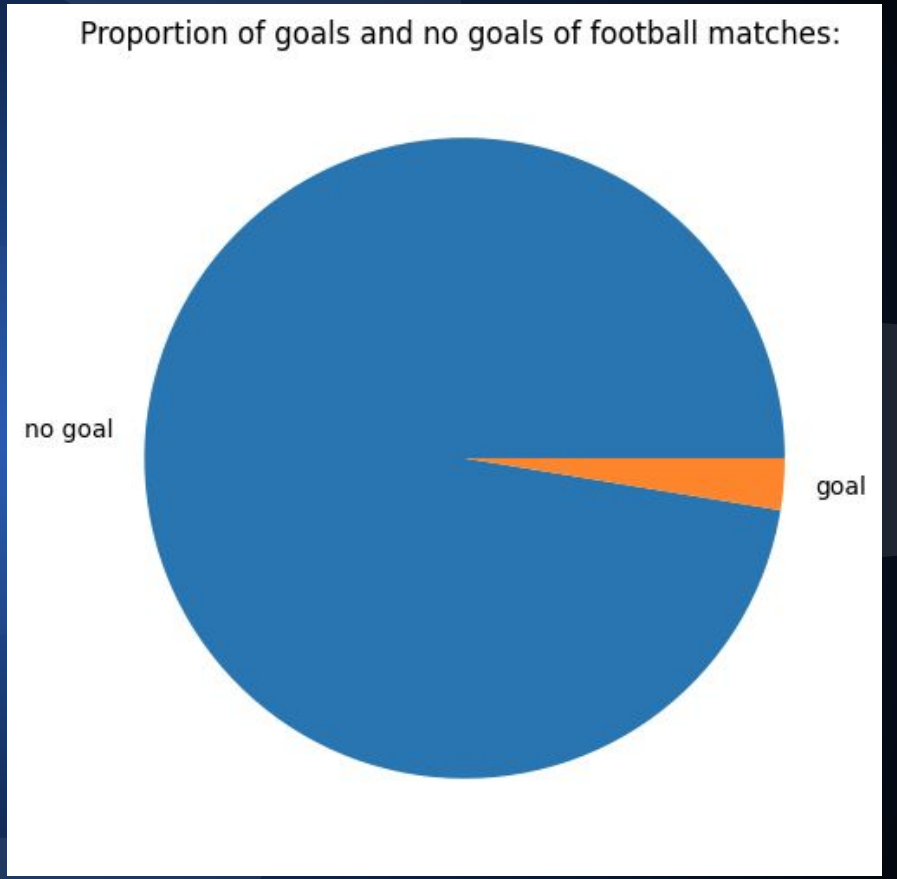
- Insights:

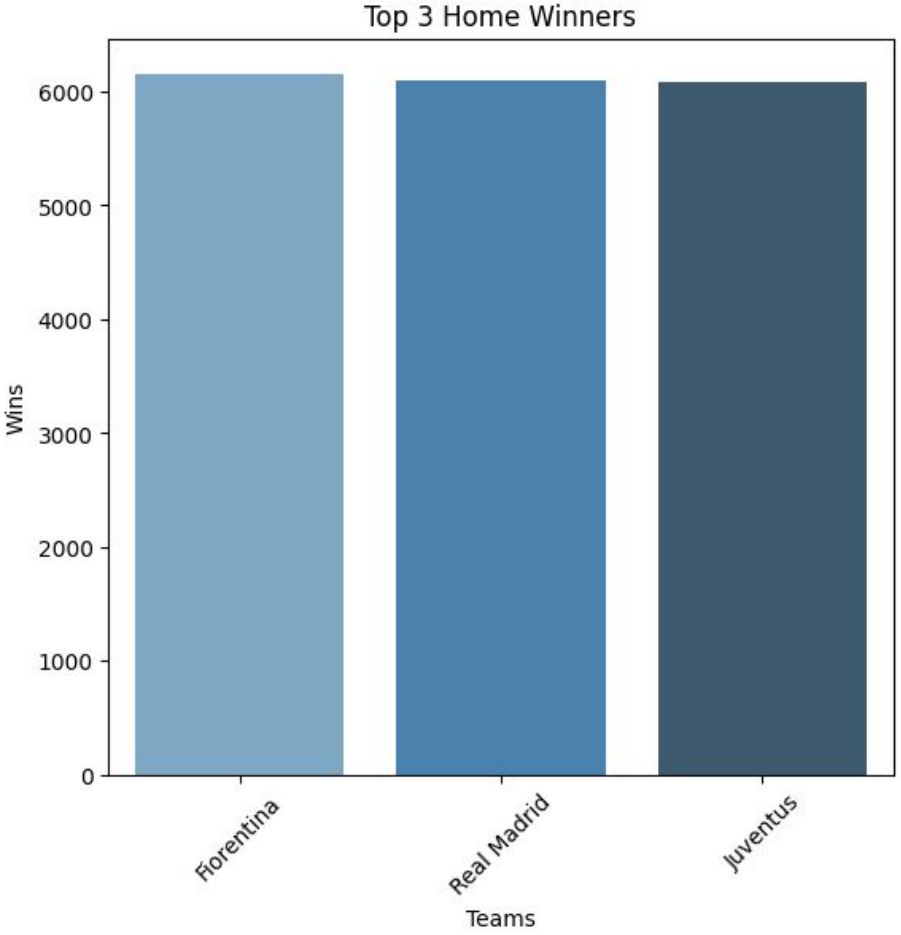
We can see from the graphic that most of the players has less than 25 goals in total.

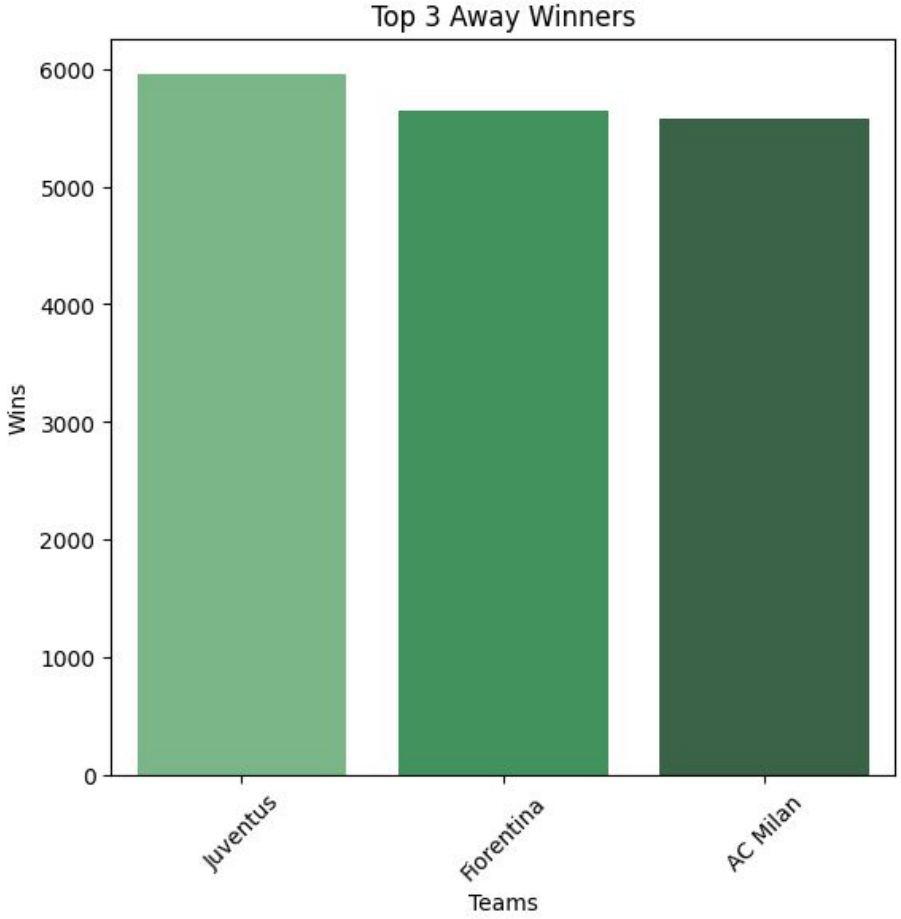

Top 5 assist players
Assistance
800
600
400
200
0
cristiano ronaldo
lionel messi
robert lewandowski
zlatan ibrahimovic
edinson cavani
Players


teams that commit the most fouls
Fouls
3000
2500
2000
1500
1000
500
0
Chievo Verona
Genoa
Lazio
Espanyol
Toulouse
Teams


Goals Rate
Goals at home
Goals away


Proportion of goals and no goals of football matches:
no goal
goal


Top 3 Home Winners
Wins
0
1000
2000
3000
4000
5000
6000
Fiorentina
Real Madrid
Juventus
Teams


Top 3 Away Winners
Wins
0
1000
2000
3000
4000
5000
6000
Juventus
Fiorentina
AC Milan
Teams

THANK YOU !