



INTRODUCTION TO DATA MANAGEMENT PROJECT REPORT

(Project Semester August-December 2021)

PROJECT REPORT ON ANALYSING UEFA CHAMPIONS LEAGUE STATS FROM 2013 to 2021

Submitted by

Aditya Prasad

Reg No:11904374

Programme: Bachelor of Technology

Section: KM008

Course Code: INT217

Under the Guidance of

Komal Arora: 17783

Assistant Professor

Discipline of CSE/IT

Lovely School of Computer Science & Engineering

Lovely Professional University, Phagwara

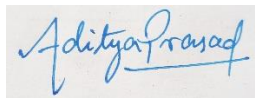
DECLARATION

I, Aditya Prasad, student of Computer Science & Engineering under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 12/12/2021

Aditya Prasad

Registration No: 11809930



Signature

ACKNOWLEDGEMENT

Primarily I'd thank my teacher/mentor Ms. Komal Arora for being able to complete my project with success. I'd like to thank my mentor, whose valuable guidance has been the ones that helped me patch this project and make it full proof success in contribution towards the completion of this project.

Last but not least I'd rather thanks to Lovely Professional University, and my parent's inspiration, who gave me this golden opportunity to learn many new things, to learn another aspect of life.

Aditya Prasad

CONTENTS

Sr No	Title	Page No
1	Introduction	1
2	Objectives/Scope of the Analysis	2
3	Source of Dataset	3-4
4	ETL Process	4-6
5	Analysis of Dataset	6-18
7	Final Dashboard	19
8	Bibliography	20

INTRODUCTION

- Data management is important because the data your organization creates is a very valuable resource.
- The last thing you want to do is spend time and resources collecting data and business intelligence, only to lose or misplace that information.
- In that case, you would then have to spend time and resources again to get that same business intelligence you already had and on that data analysis is carried out which show visualization of us problems in efficient way.
- Data Analysis is a process of inspecting, cleansing, transforming, and modelling data with the goal of discovering useful information, informing conclusions, and supporting decision- making.
- This project is based on such data analysis on UEFA CHAMPIONS LEAGUE stats from 2013 to 2021
- The UEFA Champions League is an annual club football competition organised by the Union of European Football Associations and contested by top-division European clubs, deciding the competition winners through a round robin group stage to qualify for a double-legged knockout format, and a single leg final.
- This UCL dataset contains 5000+ rows and 100+ columns.

OBJECTIVES/SCOPE OF ANALYSIS

After analysis of the dataset, the aim of this project is to give answer of given objectives in easy way:

- Objective 1-Top players with highest rating and reason of their factors determining their ratings
- Objective 2-Analysing clean sheets to determine whether it leads to winning championship
- Objective 3-Top players performance graph over the years
- Objective 4-Giving ratings to goalkeeper by determining various factors
- Objective 5-Analysing various stats of football (UEFA Champions League)

SOURCE OF DATASET:

Source of dataset: <https://www.kaggle.com/sarangpurandare/uefa-champions-league-player-statistics>

The dataset is based on all players and their performance played in UEFA Champions League from 2013 to 2021.

The rows contain player name every time he is played a season. There are almost 5000+rows with player

The columns which are used mostly are-

- Goals
- Name
- Season Year
- Ratings
- Assists
- Appearances
- Total minutes played
- Penalty Goals
- Free kick Goals

Sample of Data With Given Data fields

	A	B	C	D	E	F	G	H	I	J	K	
1	Name	Season	Season Year	Team	Rating	Goals	Bigchancescreated	Bigchancesmissed	Assists	Goalsassistsum	Accuratepasses	
2	Andrej Kramarić	Champions League 13/14	13/14	1899-hoffenheim		0	1	0	0	0	1	
3	Joshua Brenet	Champions League 13/14	13/14	1899-hoffenheim		0	0	0	0	0	0	
4	Christian Poulsen	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
5	Klaas-Jan Huntelaar	Champions League 13/14	13/14	ajax		0	1	0	0	0	0	
6	Lasse Schöne	Champions League 13/14	13/14	ajax		0	2	0	0	1	3	
7	Bojan Krkić	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
8	Siem de Jong	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
9	Niklas Moisander	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
10	Daley Blind	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
11	Lerin Duarte	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
12	Kolbeinn Sigthorsson	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
13	Viktor Fischer	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
14	Nicolai Boilesen	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
15	Davy Klaassen	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
16	Lesley De Sa	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
17	Danny Hoesen	Champions League 13/14	13/14	ajax		0	1	0	0	0	1	
18	Lucas Andersen	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
19	Mike Van der Hoorn	Champions League 13/14	13/14	ajax		0	0	0	0	0	0	
20	Thulani Serero	Champions League 13/14	13/14	ajax		0	1	0	0	2	3	

ETL PROCESS

- ETL is a process that extracts the data from different source systems, then transforms the data (like applying calculations, concatenations, etc.) and finally loads the data into the Data Warehouse system.
- Full form of ETL is Extract, Transform and Load.
- The triple combination of ETL provides crucial functions that are many times combined into a single application or suite of tools that help in the following areas:
 - Enhances Business Intelligence solutions for decision making.
 - Allows verification of data transformation, aggregation and calculations rules.
 - Allows sample data comparison between source and target system.
 - Helps to improve productivity as it codifies and reuses without additional technical skills.

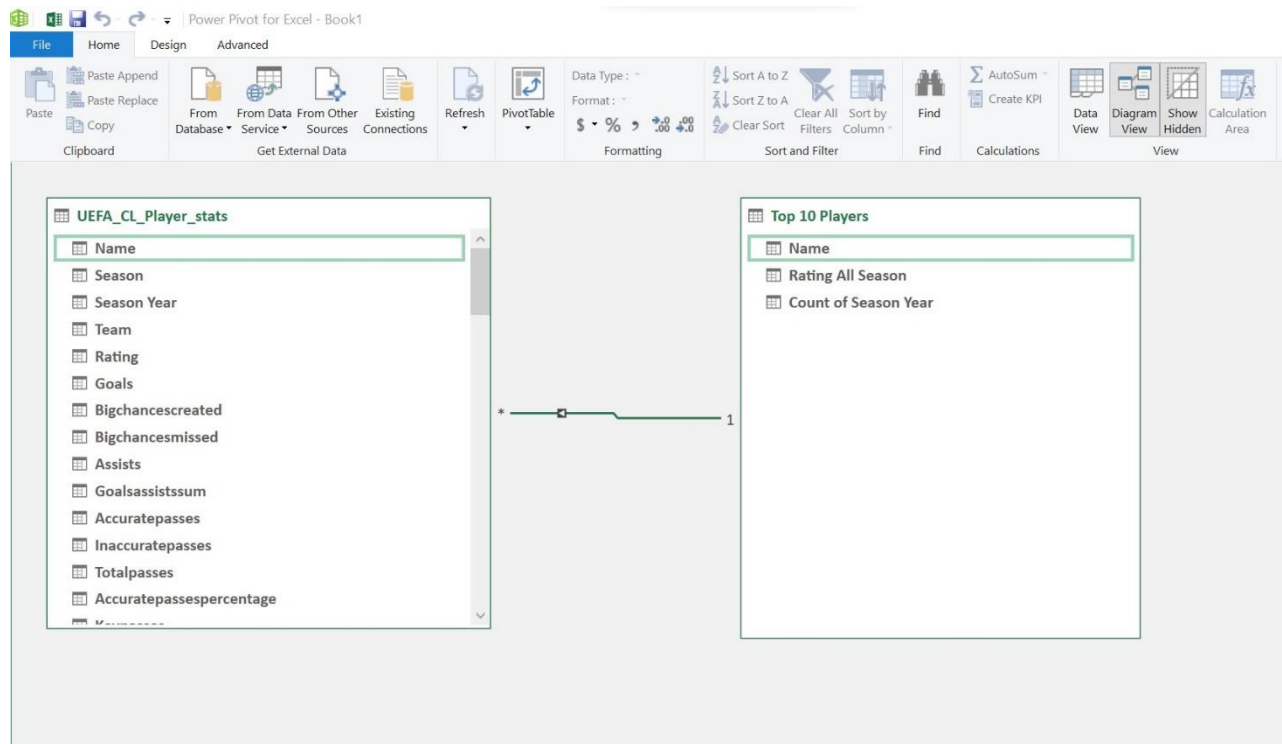
Initially the data had many unnecessary columns and null values-

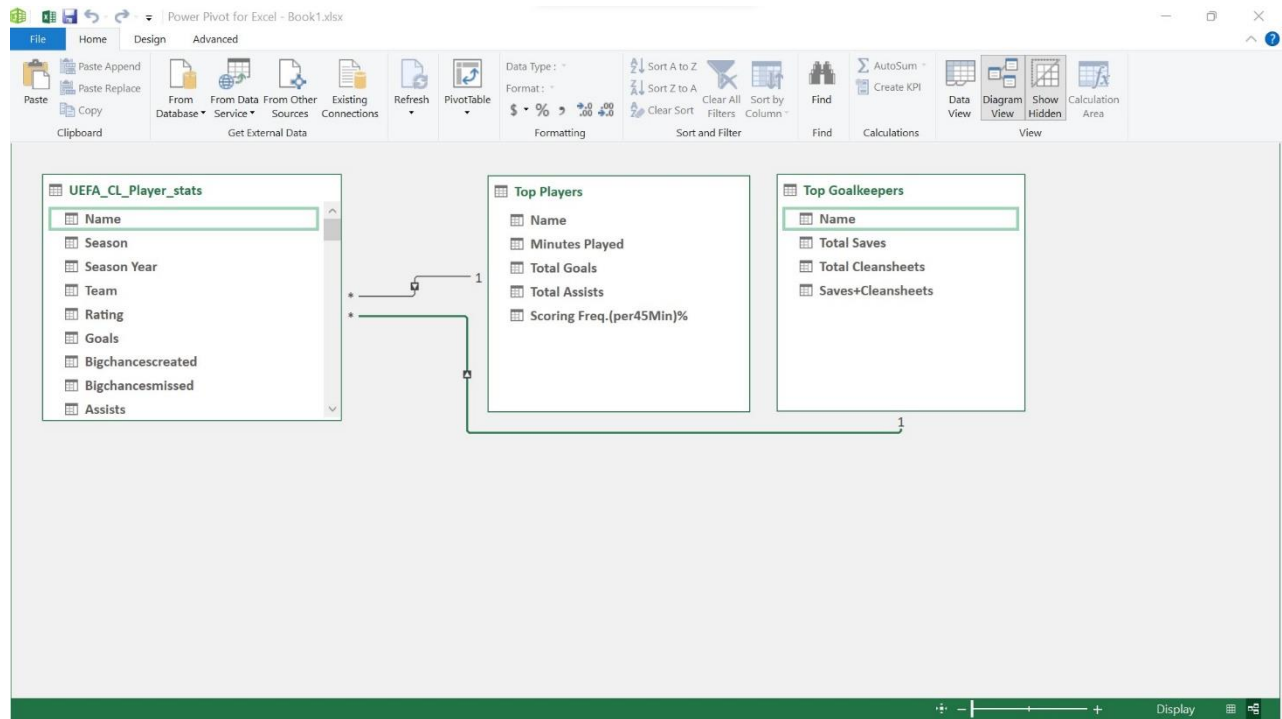
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S		
1		name	season	season ye	team	rating	goals	bigChance	bigChance	assists	goalsAssis	accurateP	inaccurate	total	Pass	accurateP	accurateO	accurateO	accurateFi	keyPasses	su
2		0 Robin van	Champion 13/14		manchester-united		4														
3		1 Ashley Col	Champion 13/14		chelsea		0			0	0										
4		2 JosÃ© Aní	UEFA Chai 15/16		sevilla	7.25	0	0	1	0	0	58	26	84	69.04762	18	42	27	2		
5		3 Roman W	Champion 13/14		borussia-dortmund		0			0	0										
6		4 Roman W	Champion 14/15		borussia-dortmund		0			0	0										
7		5 Roman W	UEFA Chai 16/17		borussia-c	6.6	0	0	0	0	0	43	12	55	78.18182	39	4	0	0		
8		6 Roman W	UEFA Chai 17/18		borussia-c	0	0	0	0	0	0	2	1	3	66.66667	2	0	0	0		
9		7 Sebastian	Champion 13/14		borussia-dortmund	1				0	1										
10		8 Sebastian	Champion 14/15		borussia-dortmund		0			1	1										
11		9 Daniel Bra	Champion 13/14		fc-kobenhavn	1				0	1										
12		10 Oliver Kirc	Champion 13/14		borussia-dortmund		0			0	0										
13		11 Oliver Kirc	Champion 14/15		borussia-dortmund		0			0	0										
14		12 Manuel Fr	Champion 13/14		borussia-dortmund		0			0	0										
15		13 Daniel var	Champion 13/14		bayern-munchen		0			0	0										
16		14 Tim Hoogl	Champion 13/14		fc-schalke-04	1															
17		15 Hamit Altı	Champion 14/15		galatasaray		0			0	0										
18		16 Christian F	Champion 13/14		ajax		0			0	0										
19		17 Gonzalo C	Champion 13/14		borussia-dortmund		0			0	0										
20		18 Gonzalo C	Champion 14/15		borussia-dortmund		0			1	1										

After the data has been cleaned the sheet looked like this-

	A	B	C	D	E	F	G	H	I	J	K	
1	Name	Season	Season Year	Team	Rating	Goals	Bigchancescreated	Bigchancesmissed	Assists	Goalsassistsum	Accuratepasse	
2	Andrej Kramarić	Champions League 13/14	13/14	1899-hoffenheim		0	1	0	0	1		
3	Joshua Brenet	Champions League 13/14	13/14	1899-hoffenheim		0	0	0	0	0		
4	Christian Poulsen	Champions League 13/14	13/14	ajax		0	0	0	0	0		
5	Klaas-Jan Huntelaar	Champions League 13/14	13/14	ajax		0	1	0	0	0		
6	Lasse Schöne	Champions League 13/14	13/14	ajax		0	2	0	1	3		
7	Bojan Krkić	Champions League 13/14	13/14	ajax		0	0	0	0	0		
8	Siem de Jong	Champions League 13/14	13/14	ajax		0	0	0	0	0		
9	Niklas Moisander	Champions League 13/14	13/14	ajax		0	0	0	0	0		
10	Daley Blind	Champions League 13/14	13/14	ajax		0	0	0	0	0		
11	Lerin Duarte	Champions League 13/14	13/14	ajax		0	0	0	0	0		
12	Kolbeinn Sigthorsson	Champions League 13/14	13/14	ajax		0	0	0	0	0		
13	Viktor Fischer	Champions League 13/14	13/14	ajax		0	0	0	0	0		
14	Nicolai Boilesen	Champions League 13/14	13/14	ajax		0	0	0	0	0		
15	Davy Klaassen	Champions League 13/14	13/14	ajax		0	0	0	0	0		
16	Lesley De Sa	Champions League 13/14	13/14	ajax		0	0	0	0	0		
17	Danny Hoesen	Champions League 13/14	13/14	ajax		0	1	0	0	1		
18	Lucas Andersen	Champions League 13/14	13/14	ajax		0	0	0	0	0		
19	Mike Van der Hoorn	Champions League 13/14	13/14	ajax		0	0	0	0	0		
20	Thulani Serero	Champions League 13/14	13/14	ajax		0	1	0	2	3		

Data Model Created for all objectives-





Analysis on Dataset

1. Top players with highest rating and reason of their factors determining their ratings

Introduction

The analysis is done on name, ratings of all season, assists, scoring frequency and no of season played. The players that have played all the seasons have been preferred since it will show their full compatibility.

Specific functions and procedure

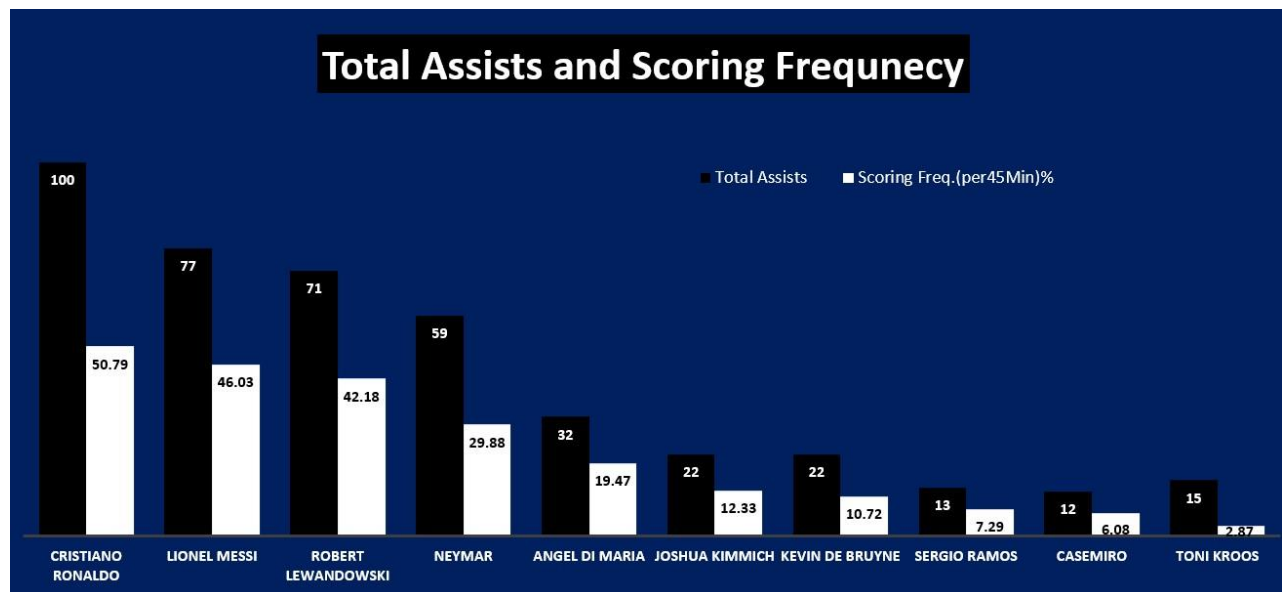
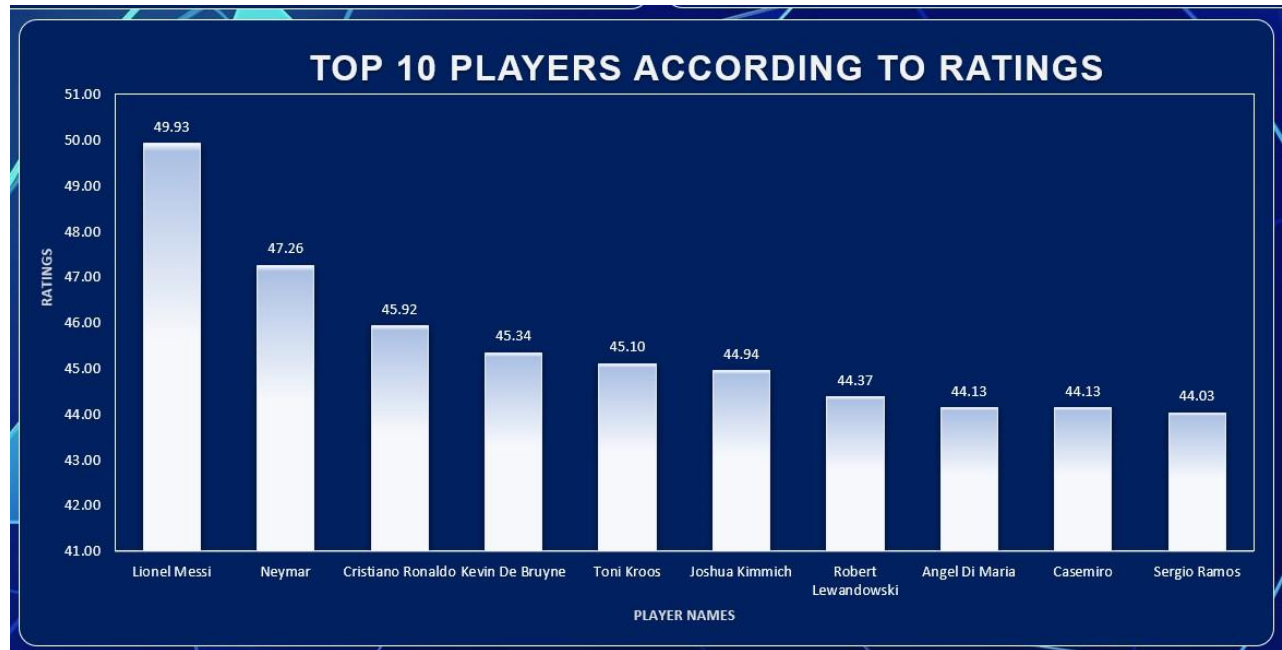
Pivot has been used to get the list of all players with their ratings and the top 10 has been extracted from that table based on both seasons played and ratings. Formula used for scoring frequency is $=[(TOTAL\ GOALS/(MINUTES\ PLAYED/45))*100]$

Tables

TOP 10 PLAYERS		
Name	Rating All Season	Count of Season Year
Lionel Messi	49.93	8
Neymar	47.26	8
Cristiano Ronaldo	45.92	8
Kevin De Bruyne	45.34	7
Toni Kroos	45.10	8
Joshua Kimmich	44.94	6
Robert Lewandowski	44.37	8
Angel Di Maria	44.13	7
Casemiro	44.13	8
Sergio Ramos	44.03	8

Factors For Being In Top 10				
Name	Minutes Played	Total Goals	Total Assists	Scoring Freq.(per45Min)%
Cristiano Ronaldo	7088	80	100	50.79
Lionel Messi	5768	59	77	46.03
Robert Lewandowski	6294	59	71	42.18
Neymar	5271	35	59	29.88
Angel Di Maria	3930	17	32	19.47
Joshua Kimmich	3650	10	22	12.33
Kevin De Bruyne	2938	7	22	10.72
Sergio Ramos	5552	9	13	7.29
Casemiro	5179	7	12	6.08
Toni Kroos	6282	4	15	2.87

Visualization



2. Analysing clean sheets to determine whether it leads to winning championship

Introduction

The analysis is done on name of goalkeepers, total saves, Total saves and cleansheets and team name.

Specific functions and procedure

Pivot table has been used to get the total saves, teamname and cleansheets, after this the data has been extracted of top 10 goalkeepers by determining the factor of saves+cleansheets.

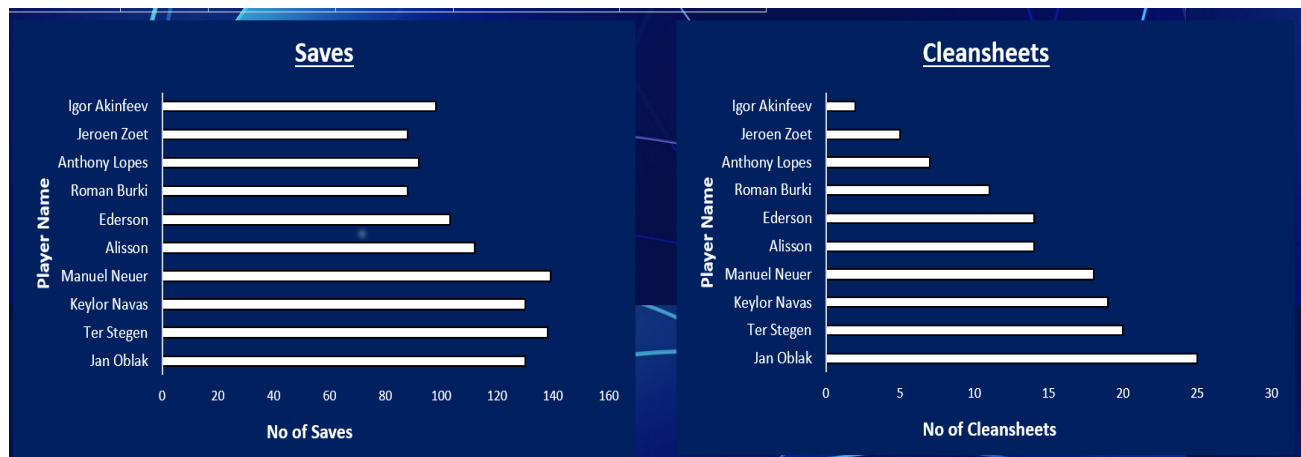
The observation made through the data is that out of the top goalkeepers 6 had won a championship or had been runner-up. Thus, it can be conclude that saves+cleansheets made during the championship leads to the path of winning a championship.

Tables

TOP 10 GOALKEEPERS					
Name	Total Saves	Total Cleansheets	Saves+Clean sheets	Team Name	Winner/Runner Up
Jan Oblak	130	25	155	atletico-madrid	YES
Ter Stegen	138	20	158	barcelona	YES
Keylor Navas	130	19	149	real-madrid	YES
Manuel Neuer	139	18	157	bayern-munchen	YES
Alisson	112	14	126	liverpool	YES
Ederson	103	14	117	manchester-city	YES
Roman Burki	88	11	99	borussia-dortmund	NO
Anthony Lopes	92	7	99	olympique-lyonnais	NO
Jeroen Zoet	88	5	93	psv-eindhoven	NO
Igor Akinfeev	98	2	100	cspa-moscow	NO

WINNERS		
Year	Winner	Runner Up
13/14	13/15	atletico-madrid
14/15	barcelona	juventus
15/16	real-madrid	atletico-madrid
16/17	real-madrid	juventus
17/18	real-madrid	liverpool
18/19	liverpool	tootenham-hotspur
19/20	bayern-munchen	paris-saint-germain
20/21	chelsea	manchester-city

Visualization



3. Top players performance graph over the years

Introduction

The analysis is done on year wise players performance which includes the stats of keypasses, goals scored, assists and ratings.

Specific functions and procedure

The analysis is done using pivot table the data had been taken from the data model that had been created earlier. Year-wise performance of all players

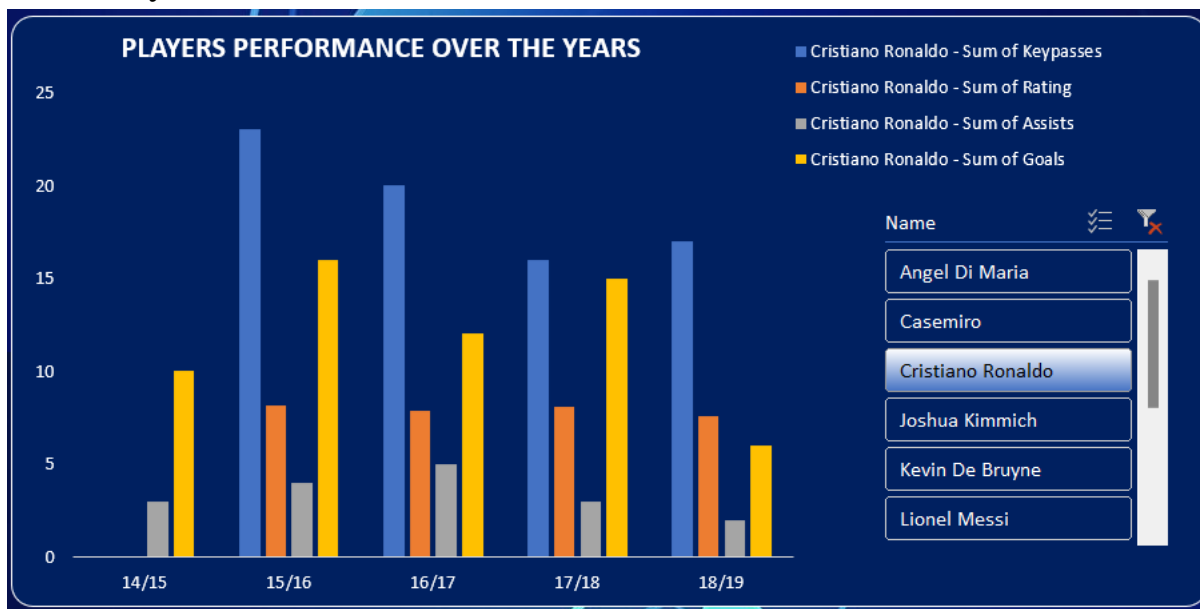
determining the above factors has been converted into the table further which a dynamic chart is created for visualization.

Table

Year	Player Name	Sum of Keypasses	Sum of Rating	Sum of Assists	Sum of Goals
	Cristiano Ronaldo				
14/15		0	0	3	10
15/16		23	8.12	4	16
16/17		20	7.85	5	12
17/18		16	8.09	3	15
18/19		17	7.54	2	6

Visualization

It is a dynamic chart, players name can be selected to see their performance over the years.



4. Giving ratings to goalkeeper by determining various factors

Introduction

The analysis is done on name, total minutes played, total cleansheets, total appearances and total saves.

Specific functions and procedure

All of the data was extracted by the use of pivot table and a new ratings column was created. Ratings were given by using the formula

$$=(\text{SAVES}+\text{CLEANSHEETS})/((\text{MINUTES PLAYED}/\text{APPEARANCES})/45))$$

OBSERVATIONS	
CRITERIA	COUNT
PLAYERS HAVE RATING >40	19
PLAYERS HAVE RATING BETWEEN 20-40	17
PLAYERS HAVE RATING <20	110
AVERAGE RATING 15.92 AND 50 GOALKEEPERS HAVE RATING GREATER THAN THE AVERAGE RATING	
Formula Used- (Total Clansheets+Total saves)/Per 45mins played	

Tables

Rating column has been added.

Name	Total Mins Played	Total Appreances	Total Cleansheets	Total Saves	Ratings
Ter Stegen	4320	48	20	138	79.00
Manuel Neuer	3930	43	18	139	77.30
Jan Oblak	4590	50	25	130	75.98
Keylor Navas	4459	49	19	130	73.68
Alisson	2880	32	14	112	63.00
Ederson	3555	40	14	103	59.24
Anthony Lopes	2644	30	7	92	50.55
Igor Akinfeev	2069	23	2	98	50.02
Roman Burki	2789	31	11	88	49.52
Wojciech Szczęsny	2571	30	10	81	47.78
Hugo Lloris	2509	28	7	87	47.21
Gianluigi Buffon	3179	35	15	78	46.08
Jeroen Zoet	1740	19	5	88	45.70
Andriy Pyatov	2160	24	0	89	44.50
Fernando Muslera	1620	18	3	86	44.50

5. Analysing various stats of football (UEFA Champions League)

Introduction

The analysis is done on total redcards, penaltygoals, owngoals, all goals scored, freekicks goal and assists of all years from 2013-2021.

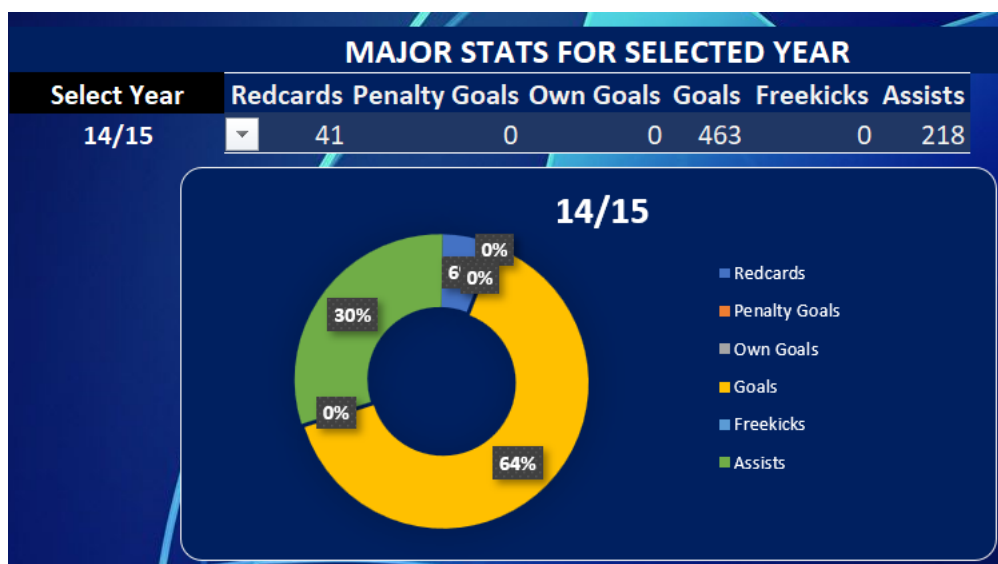
Specific functions and procedure

The data has been extracted from by the use of pivot table and it has been shown in the dashboard with the help of dynamic pie chart. All the stats related to all seasons have been compiled in one place. There is a option to select the season year which shows the stats year-wise.

Tables

Row Labels	Sum of Redcards	Sum of Penaltygoals	Sum of Owngoals	Sum of Goals	Sum of Freekickgoal	Sum of Assists
13/14	25	0	0	463	0	218
14/15	41	0	0	465	0	206
15/16	19	25	14	333	12	244
16/17	24	33	14	366	13	248
17/18	22	28	18	383	7	272
18/19	21	34	6	360	11	257
19/20	20	37	10	376	1	270
20/21	7	22	9	147	2	98

Visualisation



FINAL DASHBOARD



BIBLIOGRAPHY

Dataset Link- <https://www.kaggle.com/sarangpurandare/uefa-champions-league-player-statistics>

Background Image Link- <https://wallpaperaccess.com/full/1348815.jpg>

