1.Introdution

Cristiano Ronaldo dos Santos Averiro

Cristiano Ronaldo dos santos Averio is a Portuguese profiessional footballer who plays as a forward for premier League club Manhcester United and captaons the Portugal national team...

- Current team:Portugal national football team(#7/Forwad)Trending
- Born:Februray 5,1985(age 39 years).Hospital Dr.Nelio Mendonca,Funchal,Portugal
- Height: 1.87 m
- Partner: Georgina Rodriguez(2017_)
- Salary: 26.52 million GBP (2022)
- Total Net Worth: $260 millions
- Children:Cristiano Ronaldo Jr.,Alana MArtina dos Santos Averio,Eva Maria Dos Santos,Mateo Ronaldo

Content

- data.csv file containing Goal_no, Season, Competition, Matchday, Venue, Team, Opponent, Result, Position, Minute, At_score, Type_of_goal

Carrier Statics

- Appearances: Total number of matches played.
- Goals: Total number of goals scored.
- Assists: Total number of assists provided.
- Minutes Played: Total number of minutes played.
- Goals Per Game: Average number of goals scored per game.

International Career Statistics

1.Portugal National Team (2003-present)

- Appearances
- Goals
- Assits
- Major Tournaments (e.g., Euro Cup, World Cup): Performance and achievements.

Key Achievements

1.Champions League Titles: Number of UEFA Champions League wins.

2.Domestic League Titles: Number of league titles won in each country.

3.International Trophies: Major international titles with the national team (e.g., Euro 2016, Nations League 2019).

4.Individual Awards: Ballon d'Or wins, FIFA World Player of the Year, etc.

Example Statistics

As of the end of the 2023-2024 season, here are some general statistics (subject to updates):

- Total Career Goals: Over 850 goals.
- Total Career Appearances: Over 1200 appearances.
- Total Career Assists: Around 250+ assists.
- International Goals: Over 120 goals for Portugal.

Tools & Liabrary

- Pandas:Data manipulation and data cleaning
- Numpy:Numerical Compitation
- Seaborn:Data visualization
- Matplotlib:Additional Visualization

In [8]:
```
import pandas as pd
import numpy as np
```

In [9]:
```
ronaldo_7 = pd.read_csv("data (1).csv")
ronaldo_7.head()
```

Out[9]:

	Season	Competition	Matchday	Date	Venue	Club	Opponent	Result	Playing_Position	Minute	At_score	Type	Goa
0	02/03	Liga Portugal	6	10-07-02	H	Sporting CP	Moreirense FC	3:00	LW	34	2:00	Solo run	
1	02/03	Liga Portugal	6	10-07-02	H	Sporting CP	Moreirense FC	3:00	LW	90+5	3:00	Header	R
2	02/03	Liga Portugal	8	10/26/02	A	Sporting CP	Boavista FC	1:02	NaN	88	1:02	Right-footed shot	
3	02/03	Taca de Portugal Placard	Fourth Round	11/24/02	H	Sporting CP	CD Estarreja	4:01	NaN	67	3:00	Left-footed shot	
4	02/03	Taca de Portugal Placard	Fifth Round	12/18/02	H	Sporting CP	FC Oliveira do Hospital	8:01	NaN	13	3:00	NaN	

In [12]:
```
ronaldo_7 ['Season'].value_counts()
```

Out[12]:
```
Season
14/15     61
11/12     60
Dec-13    55
10/11     53
13/14     51
15/16     51
17/18     44
16/17     42
07/08     42
19/20     37
20/21     36
09/10     33
18/19     28
08/09     26
21/22     24
06/07     23
05/06     12
22/23     12
04/05      9
03/04      6
02/03      5
Name: count, dtype: int64
```

In [14]:
```
ronaldo_7 ['Competition'].value_counts()
```

Out[14]:
```
Competition
LaLiga                             311
UEFA Champions League              140
Premier League                     103
Serie A                             81
Copa del Rey                        22
FA Cup                              13
Saudi Pro League                     9
FIFA Club World Cup                  7
Italy Cup                            4
EFL Cup                              4
Supercopa                            4
Liga Portugal                        3
Taca de Portugal Placard             2
Supercoppa Italiana                  2
UEFA Super Cup                       2
Europa League                        2
UEFA Champions League Qualifying     1
Name: count, dtype: int64
```

In [15]:
```
ronaldo_7 ['Club'].value_counts()
```

Out[15]:
```
Club
Real Madrid          450
Manchester United    145
Juventus FC          101
Al-Nassr FC            9
Sporting CP            5
Name: count, dtype: int64
```

In [17]:
```
ronaldo_7 ['Venue'].value_counts()
```

Out[17]:
```
Venue
H    404
A    306
Name: count, dtype: int64
```

In [18]:
```
ronaldo_7 ['Type'].value_counts()
```

Out[18]:
```
Type
Right-footed shot         253
Penalty                   132
Left-footed shot          113
Header                    112
Direct free kick           49
Tap-in                     14
Long distance kick          9
Counter attack goal         6
Penalty rebound             3
Solo run                    2
Deflected shot on goal      2
Name: count, dtype: int64
```

In [19]:
```
ronaldo_7 ['Playing_Position'].value_counts()
```

Out[19]:
```
Playing_Position
LW     356
CF     209
RW      79
LW       5
CF       3
Name: count, dtype: int64
```

In [21]:
```
ronaldo_7 ['Opponent'].value_counts()
```

Out[21]:
```
Opponent
Sevilla FC                27
Atletico de Madrid        25
Getafe CF                 23
Celta de Vigo             20
FC Barcelona              20
                          ..
Burnley FC                 1
Brentford FC               1
Brighton & Hove Albion     1
Al-Fateh                   1
Abha Club                  1
Name: count, Length: 129, dtype: int64
```

In [24]:
```
ronaldo_7 ['Goal_assist'].value_counts()
```

Out[24]:
```
Goal_assist
Karim Benzema          44
Gareth Bale            29
Mesut Ozil             26
Marcelo                23
Angel Di Maria         22
                       ..
Abdulrahman Ghareeb     1
Sami Al-Najei           1
\t                      1
Sultan Al-Ghannam       1
Ayman Yahya             1
Name: count, Length: 91, dtype: int64
```

In [25]:
```
ronaldo_7 ['Matchday'].value_counts()
```

Out[25]:
```
Matchday
Group Stage       75
last 16           34
Quarter-Finals    31
Semi-Finals       26
29                20
18                19
6                 18
23                18
8                 18
20                17
37                16
Final             16
13                16
12                15
25                15
35                15
21                15
22                15
27                15
16                14
15                14
28                14
14                14
3                 14
17                13
10                13
36                13
9                 13
31                13
24                13
38                13
11                12
34                12
30                11
4                 11
26                10
33                10
19                10
5                  9
7                  8
32                 8
2                  7
1                  6
Fifth Round        4
final              4
Fourth Round       3
Third Round        3
Sixth Round        2
4th round          2
Round of 16        1
3rd round          1
6th round          1
Name: count, dtype: int64
```

In [40]:
```
ronaldo_7.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 710 entries, 0 to 709
Data columns (total 13 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Season            710 non-null    object
 1   Competition       710 non-null    object
 2   Matchday          710 non-null    object
 3   Date              710 non-null    object
 4   Venue             710 non-null    object
 5   Club              710 non-null    object
 6   Opponent          710 non-null    object
 7   Result            710 non-null    object
 8   Playing_Position  652 non-null    object
 9   Minute            710 non-null    object
 10  At_score          710 non-null    object
 11  Type              695 non-null    object
 12  Goal_assist       464 non-null    object
dtypes: object(13)
memory usage: 72.2+ KB
```

- The info() method is particularly useful for getting a quick overview of the DataFrame's structure, checking for missing values, and understanding the types of data you're working with.

In [41]:
```
ronaldo_7.describe(include=['object']).T
```

Out[41]:

	count	unique	top	freq
Season	710	21	14/15	61
Competition	710	17	LaLiga	311
Matchday	710	52	Group Stage	75
Date	710	468	04-05-15	5
Venue	710	2	H	404
Club	710	5	Real Madrid	450
Opponent	710	129	Sevilla FC	27
Result	710	57	3:00	49
Playing_Position	652	5	LW	356
Minute	710	106	90	17
At_score	710	43	1:00	111
Type	695	11	Right-footed shot	253
Goal_assist	464	91	Karim Benzema	44

- The describe part calculates summary statistics specifically for columns with object data types (like strings or categorical variables).
- The .T at the end transposes the result of the describe method, making it easier to view or interpret the statistics, especially when dealing with many columns.

In [35]:
```
pd.DataFrame(ronaldo_7.apply(lambda col: len(col.unique())),columns=["Unique Values Count"])
```

Out[35]:

	Unique Values Count
Season	21
Competition	17
Matchday	52
Date	468
Venue	2
Club	5
Opponent	129
Result	57
Playing_Position	6
Minute	106
At_score	43
Type	12
Goal_assist	92

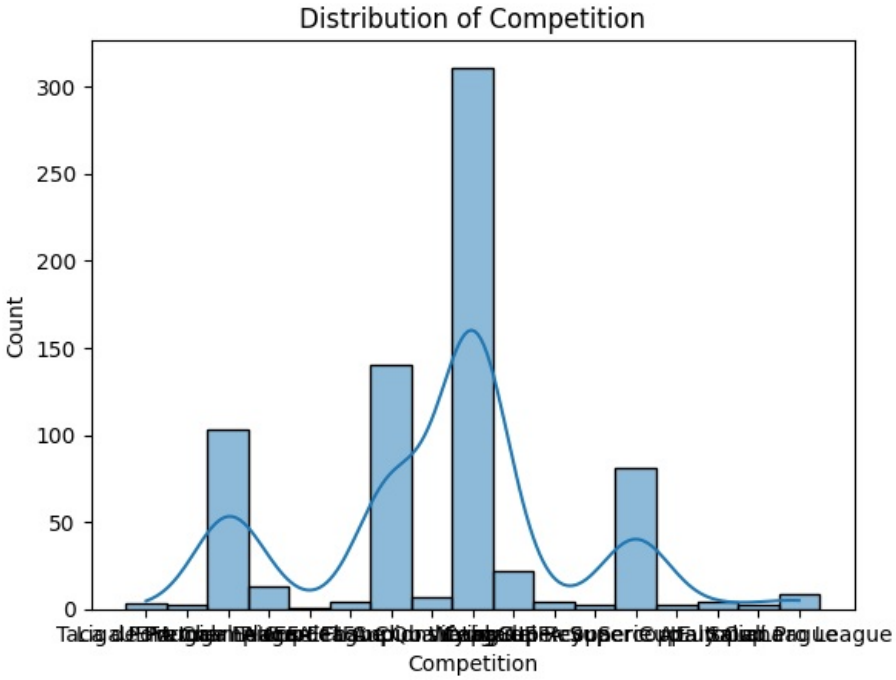
- It calculates the number of unique values for each column in the ronaldo_7 DataFrame.It converts this result into a new DataFrame where each row represents a column from the original DataFrame and contains the count of unique values in that column.The new DataFrame has a single column named "Unique Values Count" that holds these counts.

Data Visualization

Goals per Competition

In [62]:
```
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
import warnings
warnings.filterwarnings('ignore')
```

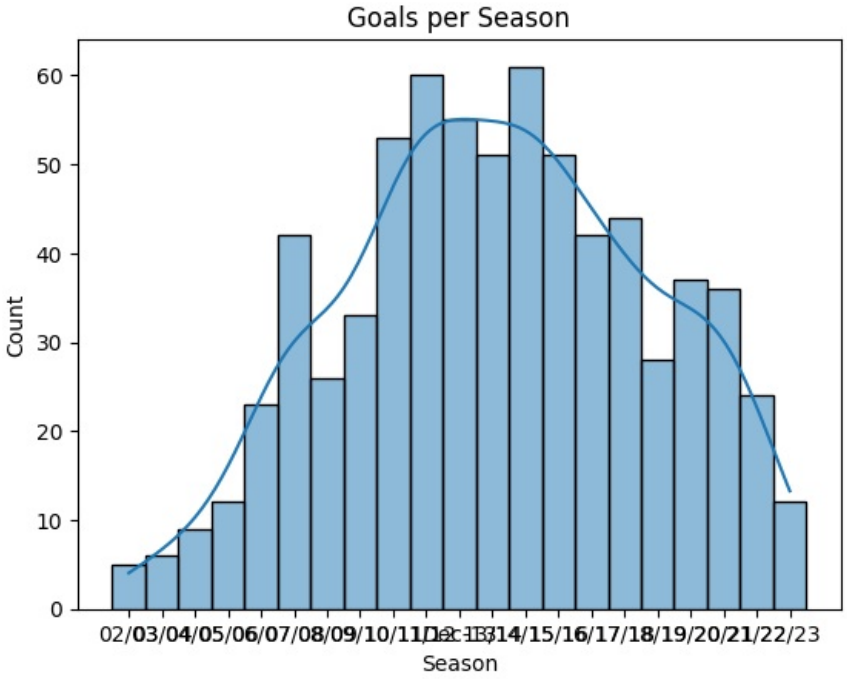
In [67]:
```
sns.histplot(ronaldo_7['Competition'], bins=30,kde=True)
plt.title('Distribution of Competition')
plt.show()
```



In [9]:
```
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import matplotlib.pyplot as plt
import seaborn as sns
from datetime import timedelta
import warnings
import os
```

Goals per Season

In [37]:
```
sns.histplot(ronaldo_7['Season'], bins=30,kde=True)
plt.title('Goals per Season')
plt.show()
```



- This visualization is useful for understanding the distribution of 'Season' values and the overall pattern of the data. If 'Season' is a categorical variable rather than numerical, you might want to adjust your plotting approach or preprocess the data accordingly.

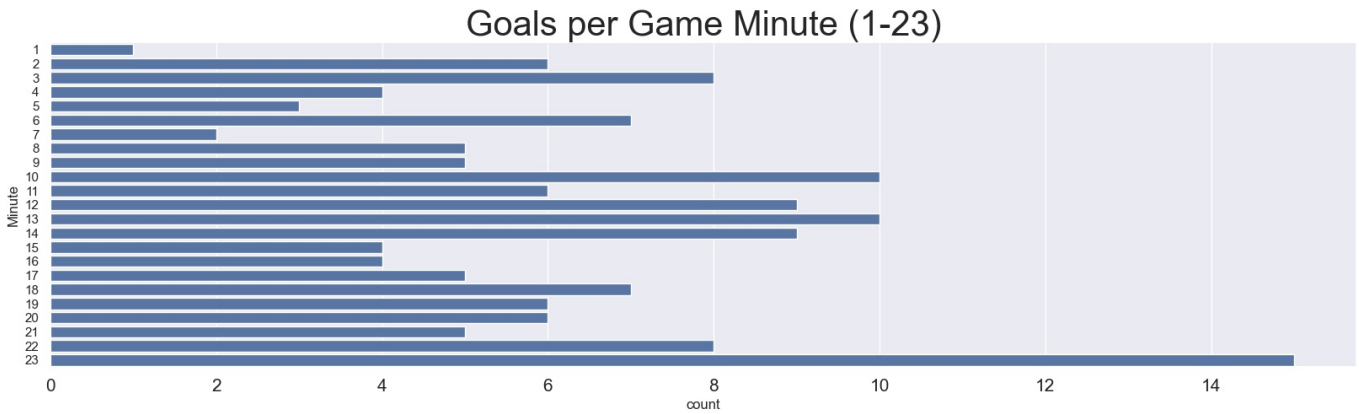
Goals Per Game Minute

In [41]:
```
mins=list(map(str, range(1,121)))
mins.insert(45,"1stE")
mins.insert(91,"2ndE")
mins.insert(122,"ExtE")
ronaldo_7.loc[ronaldo_7.Minute.str[:3]=='45+', 'Minute'] = '1stE'
ronaldo_7.loc[ronaldo_7.Minute.str[:3]=='90+', 'Minute'] = '2ndE'
```

```
ronaldo_7.loc[ronaldo_7.Minute.str[:4]=='120+', 'Minute'] = 'ExtE'
mins1=mins[:23]
mins2=mins[23:46]
mins3=mins[46:69]
mins4=mins[69:92]
mins5=mins[92:]
```

In [45]:
```
sns.set(rc={'figure.figsize':(20,5)})
plt.xticks(fontsize=15)
p=sns.countplot(ronaldo_7['Minute'],order=mins1)
p.axes.set_title("Goals per Game Minute (1-23)",fontsize=30)
```

Out[45]:
```
Text(0.5, 1.0, 'Goals per Game Minute (1-23)')
```



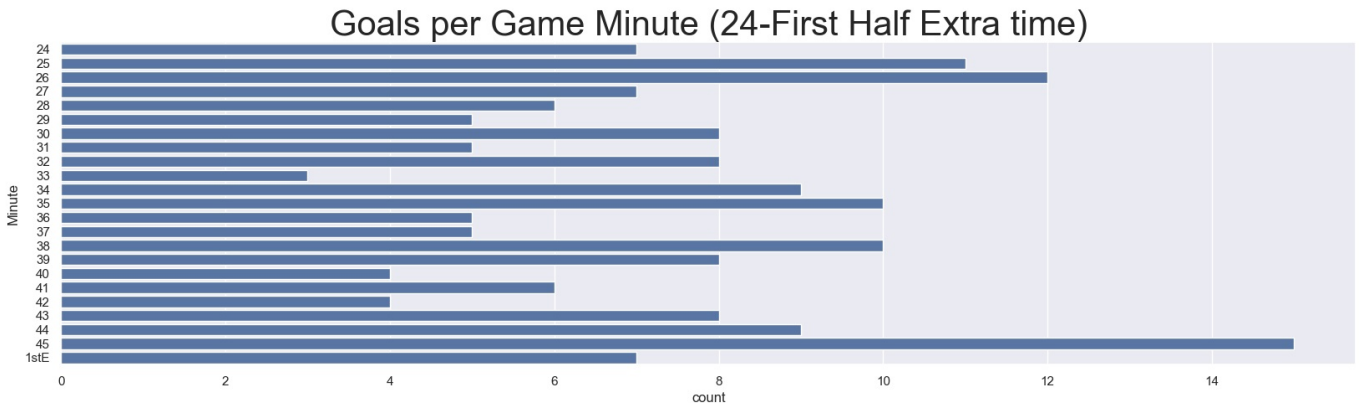
- sns.set(rc={'figure.figsize':(20,5)}): Sets the figure size for the plot to 20x5 inches.

plt.xticks(fontsize=15): Adjusts the font size of x-axis tick labels to 15 points. sns.countplot(x=ronaldo_7['Minute'], order=mins1): Creates a bar plot of the frequency of goals scored in different minutes, with bars ordered according to mins1. p.axes.set_title("Goals per Game Minute (1-23)", fontsize=30): Sets a title for the plot with a font size of 30 points.

- This visualization will show the number of goals scored during each minute of the game from 1 to 23, with a clear and large title and appropriately sized tick labels for readability. Make sure mins1 is defined to properly order the x-axis categories.

In [46]:
```
p=sns.countplot(ronaldo_7 ['Minute'],order=mins2)
p.axes.set_title("Goals per Game Minute (24-First Half Extra time)",fontsize=30)
```

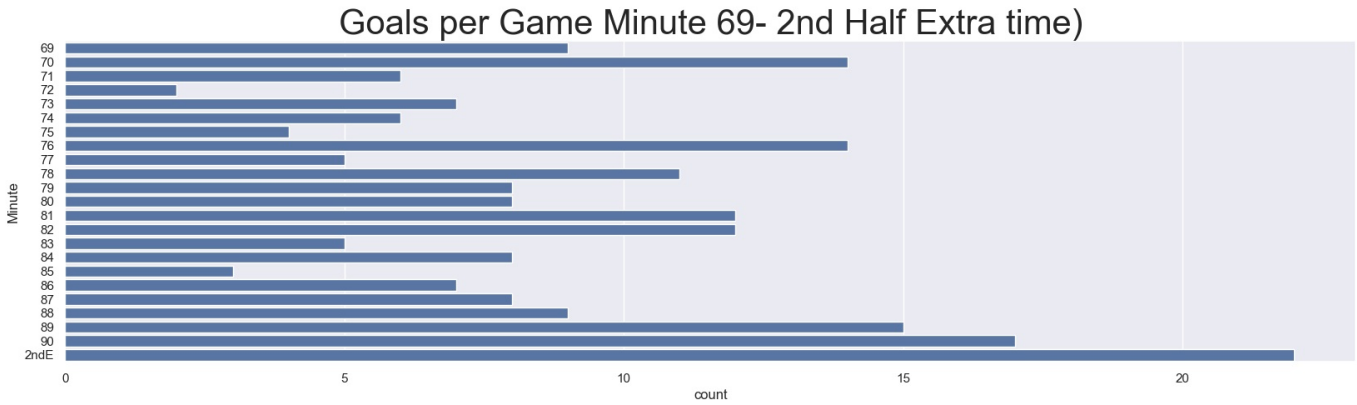
Out[46]:
```
Text(0.5, 1.0, 'Goals per Game Minute (24-First Half Extra time)')
```



- The code creates a count plot displaying the distribution of goals scored in specific minutes from 24 to the end of the first half's extra time, with bars ordered according to mins2 and a prominent title explaining the focus of the plot.

In [50]:
```
p=sns.countplot(ronaldo_7 ['Minute'],order=mins4)
p.axes.set_title("Goals per Game Minute 69- 2nd Half Extra time)",fontsize=30)
```

Out[50]:
```
Text(0.5, 1.0, 'Goals per Game Minute 69- 2nd Half Extra time)')
```



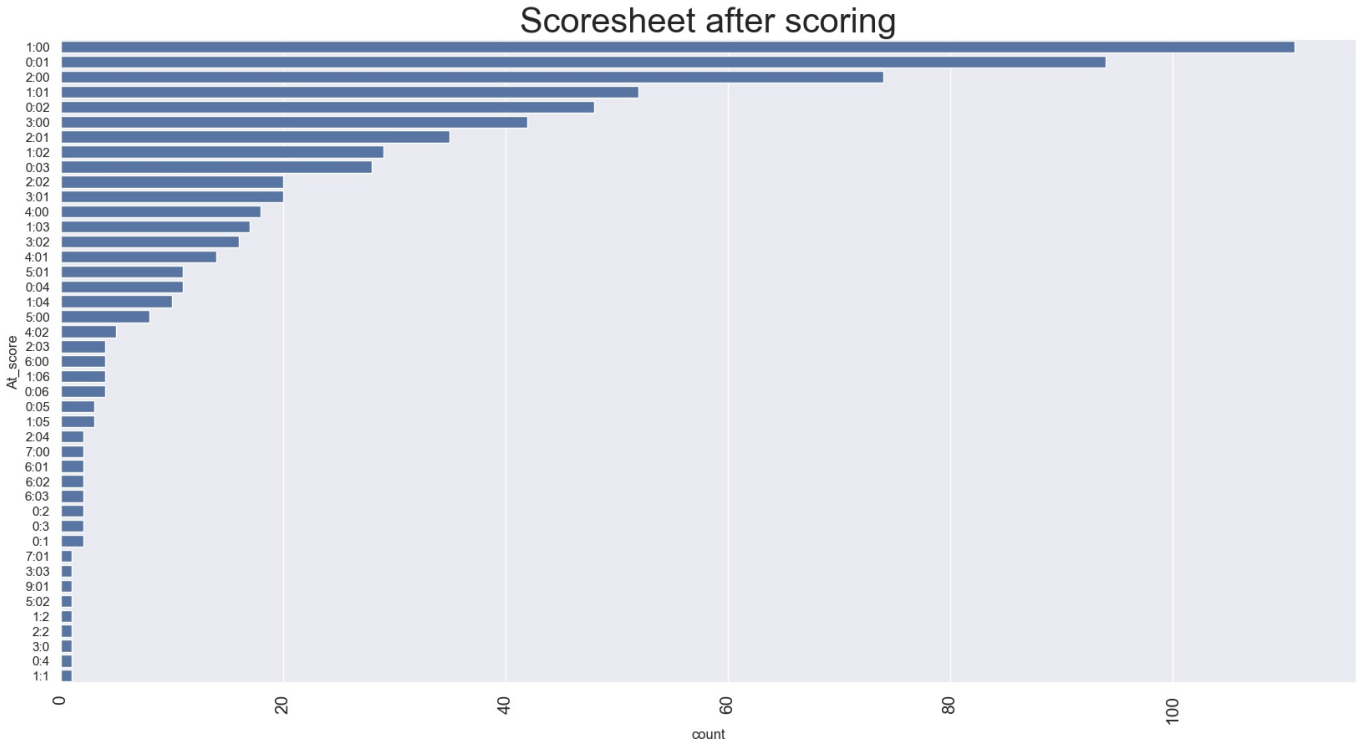
The code produces a count plot for goals scored from the 69th minute through the 2nd half extra time, using mins4 to order the x-axis, and sets a title describing this time range with a font size of 30.

- He scored in all 90 game minutes!

Scorline after goals

In [64]:
```
sns.set(rc={'figure.figsize':(20,10)})
plt.xticks(fontsize=15,rotation='vertical')
p=sns.countplot(ronaldo_7['At_score'],hue_order=ronaldo_7.groupby('Competition'),order=ronaldo_7.At_score.value_
p.axes.set_title("Scoresheet after scoring",fontsize=30)
```

Out[64]:
```
Text(0.5, 1.0, 'Scoresheet after scoring')
```



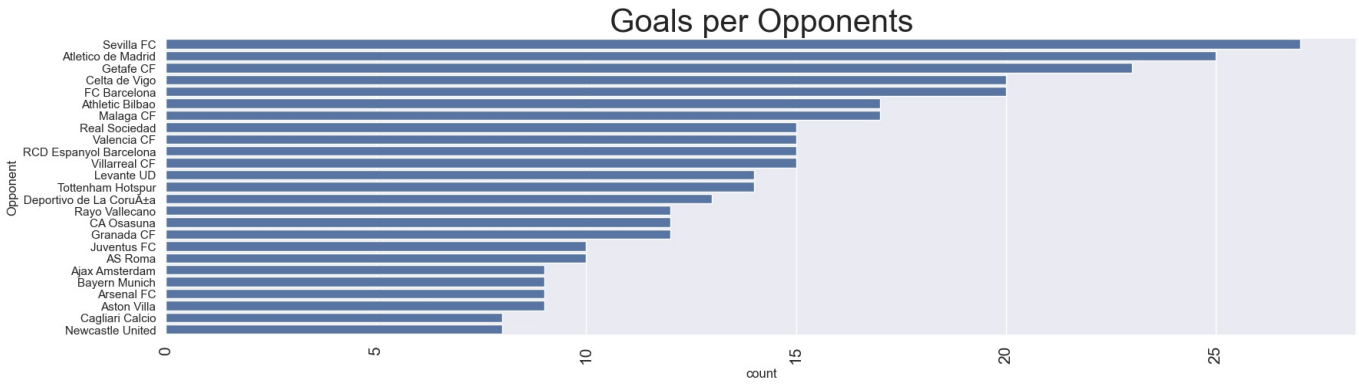
- The code configures the plot size to 20x10 inches and rotates x-axis labels vertically for better readability. It creates a count plot of 'At_score' with bars ordered by the frequency of scores and hues representing different competitions, and sets a title with a font size of 30 to describe the scoresheet after scoring.

Goals Per Opponents

In [71]:
```
mins=list(map(str, ronaldo_7.Opponent.value_counts().sort_values(ascending=False).index))
for min in ronaldo_7['Opponent']:
    if min not in mins:
        mins.append(min)
mins1=mins[:int(len(mins)/5)]
mins2=mins[int(len(mins)/5):int(2*len(mins)/5)]
mins3=mins[2*int(len(mins)/5):int(3*len(mins)/5)]
mins4=mins[3*int(len(mins)/5):int(4*len(mins)/5)]
mins5=mins[int(4*len(mins)/5):]
```

In [72]:
```
sns.set(rc={'figure.figsize':(20,5)})
plt.xticks(fontsize=15,rotation='vertical')
p=sns.countplot(ronaldo_7['Opponent'],order=mins1)
p.axes.set_title("Goals per Opponents",fontsize=30)
```

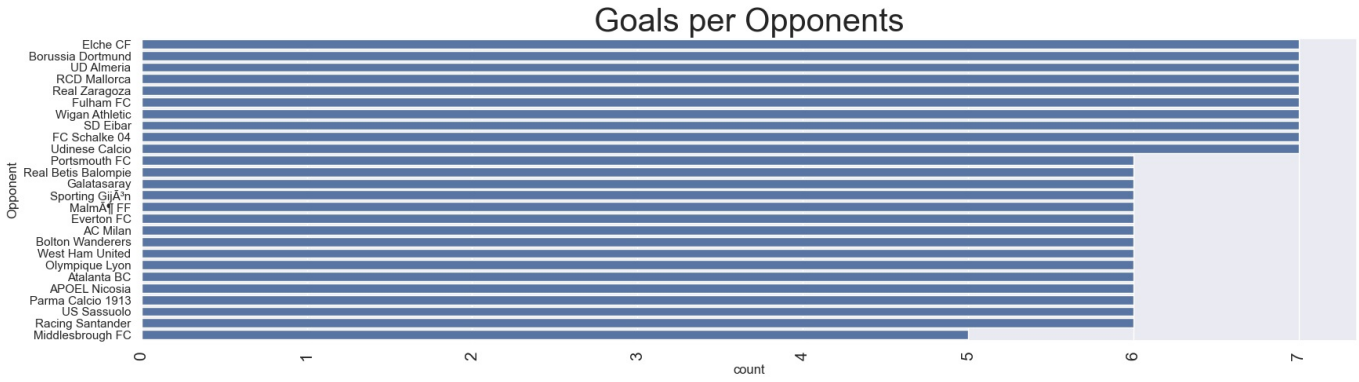
Out[72]:
```
Text(0.5, 1.0, 'Goals per Opponents')
```



- The code sets the plot size to 20x5 inches and rotates x-axis labels vertically for readability. It generates a count plot of 'Opponent', ordering bars by mins1, and sets a title with a font size of 30 to illustrate the number of goals scored against each opponent.

In [74]:
```
sns.set(rc={'figure.figsize':(20,5)})
plt.xticks(fontsize=15,rotation='vertical')
p=sns.countplot(ronaldo_7['Opponent'],order=mins2)
p.axes.set_title("Goals per Opponents",fontsize=30)
```

Out[74]:
```
Text(0.5, 1.0, 'Goals per Opponents')
```

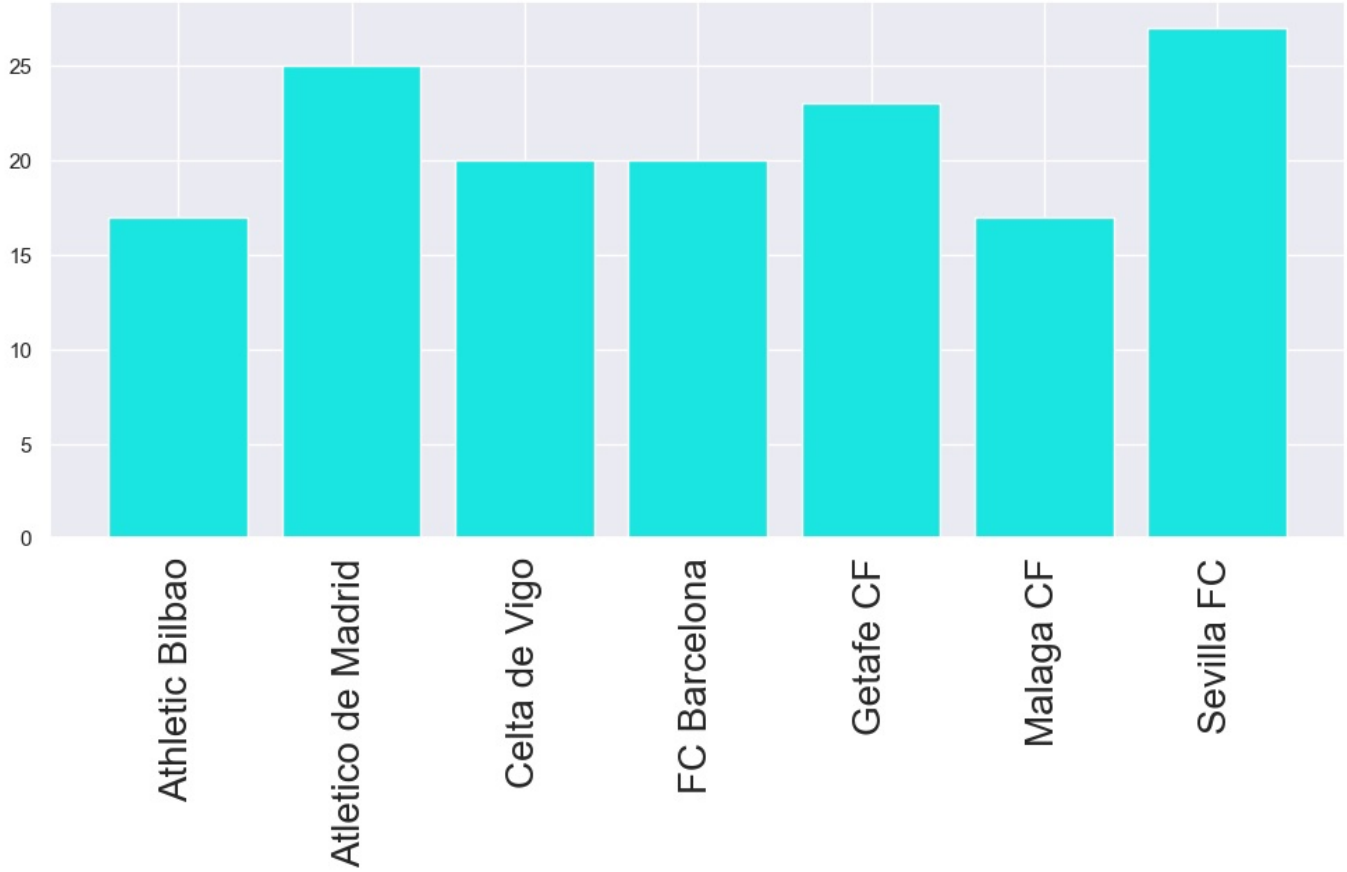


- The code sets the plot size to 20x5 inches and rotates x-axis labels vertically for improved readability. It creates a count plot of 'Opponent', ordering bars according to mins2, and adds a title with a font size of 30 to display the number of goals scored against each opponent.

Favourite Opponents

In [79]:
```
sns.set(rc={'figure.figsize':(12,5)})
opponents_ronaldo_7=ronaldo_7.groupby('Opponent').size().reset_index(name='count')
fav_opponents_ronaldo_7=opponents_ronaldo_7[opponents_ronaldo_7["count"]>15]
plt.xticks(fontsize=20,rotation='vertical')
plt.bar(x=fav_opponents_ronaldo_7['Opponent'],height=fav_opponents_ronaldo_7['count'],color="#1ae5e1")
```

Out[79]:
```
<BarContainer object of 7 artists>
```



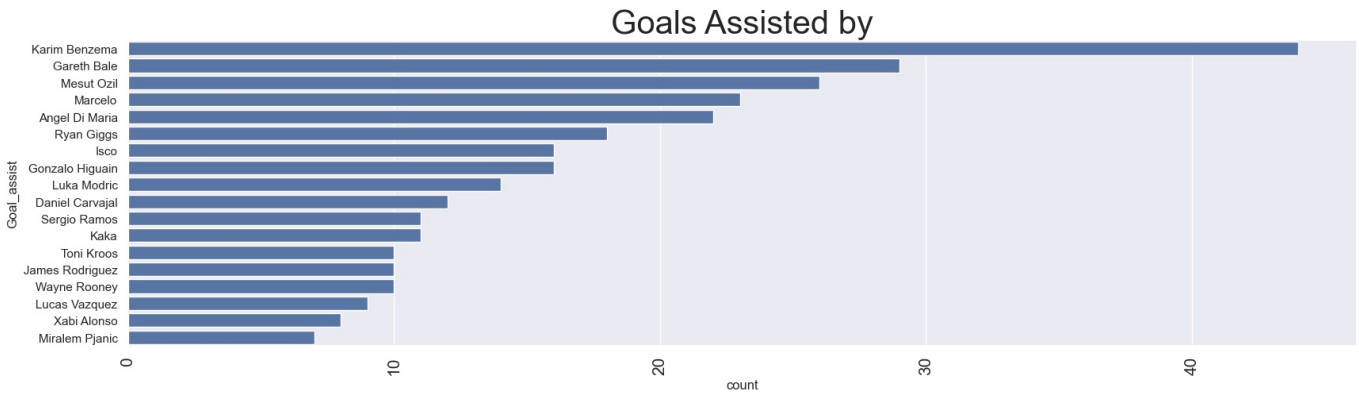
- The code sets the plot size to 12x5 inches and prepares data by grouping the ronaldo_7 DataFrame by 'Opponent' and filtering for opponents with more than 15 matches. It then creates a vertical bar plot showing the count of these frequent opponents, with bars colored in #1ae5e1 and x-axis labels rotated for better readability.

Assits

In [81]:
```
mins=list(map(str, ronaldo_7.Goal_assist.value_counts().sort_values(ascending=False).index))
for min in ronaldo_7['Goal_assist']:
    if min not in mins:
        mins.append(min)
mins1=mins[:int(len(mins)/5)]
mins2=mins[int(len(mins)/5):int(2*len(mins)/5)]
mins3=mins[2*int(len(mins)/5):int(3*len(mins)/5)]
mins4=mins[3*int(len(mins)/5):int(4*len(mins)/5)]
mins5=mins[int(4*len(mins)/5):]
```

In [82]:
```
sns.set(rc={'figure.figsize':(20,5)})
plt.xticks(fontsize=15,rotation='vertical')
p=sns.countplot(ronaldo_7['Goal_assist'],order=mins1)
p.axes.set_title("Goals Assisted by",fontsize=30)
```

Out[82]:
```
Text(0.5, 1.0, 'Goals Assisted by')
```



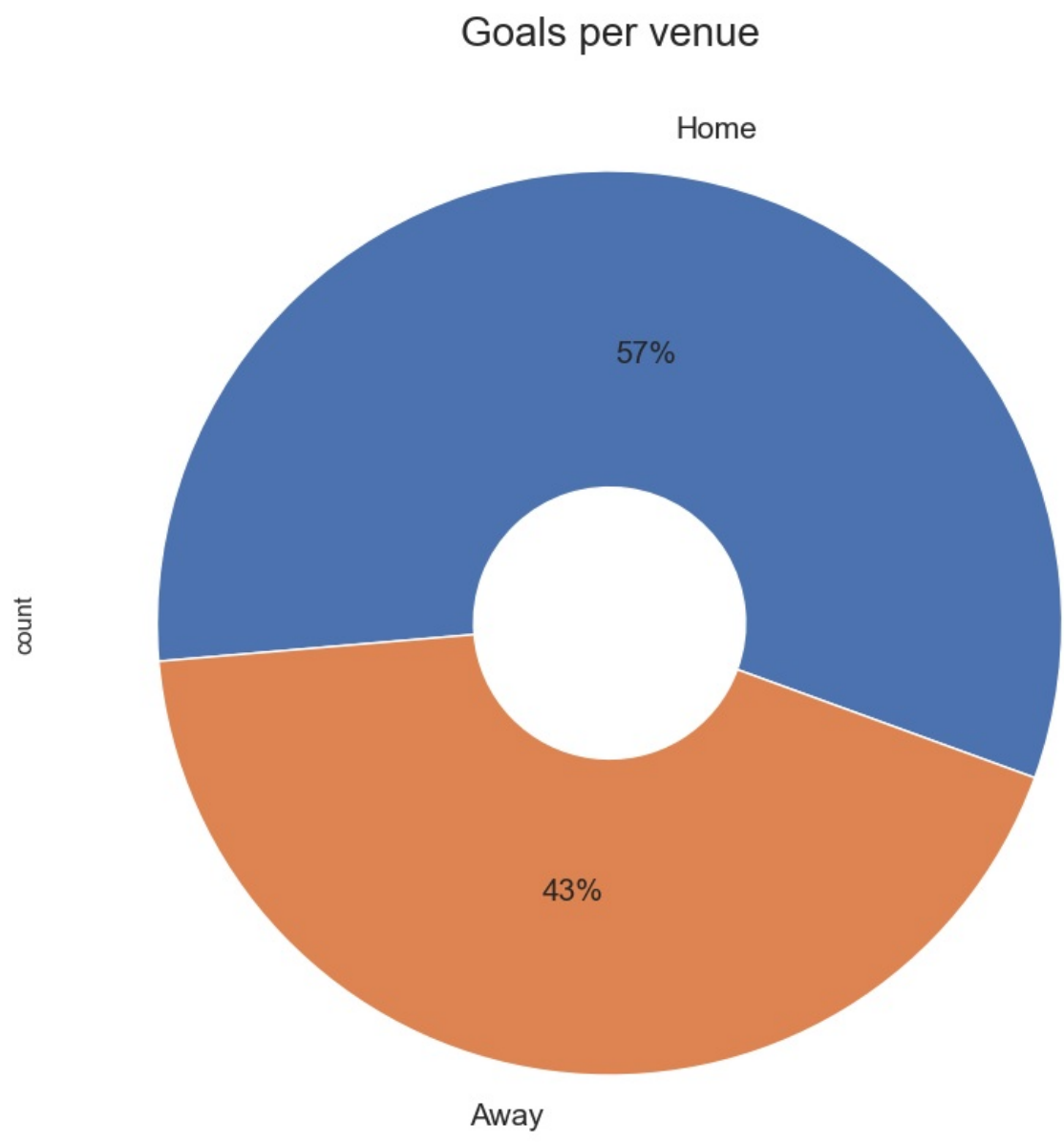
- The code sets the plot size to 20x5 inches and rotates x-axis labels vertically for better readability. It generates a count plot of 'Goal_assist', ordered by mins1, and sets a title with a font size of 30 to show the distribution of assists made.

- The plot visualizes the frequency of assists in various contexts, highlighting how often each type of assist occurs with a clear, large title for easy interpretation.

Home-Away Goals

In [85]:
```
plt.figure(figsize=(30,10))
plt.title('Goals per venue', fontsize=20)
ronaldo_7.Venue.value_counts().plot(kind='pie', labels=['Home', 'Away'], wedgeprops=dict(width=.7), autopct='%1
```

Out[85]:
```
<Axes: title={'center': 'Goals per venue'}, ylabel='count'>
```


Goals per venue
Home
57%
count
43%
Away

- The code sets the figure size to 30x10 inches and creates a pie chart to show the distribution of goals scored at different venues, with labels indicating 'Home' and 'Away' and slice sizes displayed as percentages. The chart is styled with a wide wedge, a starting angle of -20 degrees, and text formatted with a font size of 15.

In [ ]: