

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
```

```
df= pd.read_csv("IPL 2022.CSV")
```

```
df
```

	match_id	date	
venue \			
0	1	March 26,2022	Wankhede Stadium,
Mumbai			
1	2	March 27,2022	Brabourne Stadium,
Mumbai			
2	3	March 27,2022	Dr DY Patil Sports Academy,
Mumbai			
3	4	March 28,2022	Wankhede Stadium,
Mumbai			
4	5	March 29,2022	Maharashtra Cricket Association
Stadium,Pune			
..	
...			
69	70	May 22,2022	Wankhede Stadium,
Mumbai			
70	71	May 24,2022	Eden Gardens,
Kolkata			
71	72	May 25,2022	Eden Gardens,
Kolkata			
72	73	May 27,2022	Narendra Modi Stadium,
Ahmedabad			
73	74	May 29,2022	Narendra Modi Stadium,
Ahmedabad			

	team1	team2	stage	toss_winner	toss_decision
first_ings_score \					
0	Chennai	Kolkata	Group	Kolkata	Field
131					
1	Delhi	Mumbai	Group	Delhi	Field
177					
2	Banglore	Punjab	Group	Punjab	Field
205					
3	Gujarat	Lucknow	Group	Gujarat	Field
158					
4	Hyderabad	Rajasthan	Group	Hyderabad	Field
210					
..
...					
69	Hyderabad	Punjab	Group	Hyderabad	Bat

157						
70	Gujarat	Rajasthan	Playoff	Gujarat	Field	
188						
71	Banglore	Lucknow	Playoff	Lucknow	Field	
207						
72	Banglore	Rajasthan	Playoff	Rajasthan	Field	
157						
73	Gujarat	Rajasthan	Final	Rajasthan	Bat	
130						
	first_ings_wkts	second_ings_score	second_ings_wkts	match_winner		
\						
0	5	133	4	Kolkata		
1	5	179	6	Delhi		
2	2	208	5	Punjab		
3	6	161	5	Gujarat		
4	6	149	7	Rajasthan		
..		
69	8	160	5	Punjab		
70	6	191	3	Gujarat		
71	4	193	6	Banglore		
72	8	161	3	Rajasthan		
73	9	133	3	Gujarat		
	won_by	margin	player_of_the_match	top_scorer		
highscore \						
0	Wickets	6	Umesh Yadav	MS Dhoni	50	
1	Wickets	4	Kuldeep Yadav	Ishan Kishan	81	
2	Wickets	5	Odean Smith	Faf du Plessis	88	
3	Wickets	5	Mohammed Shami	Deepak Hooda	55	
4	Runs	61	Sanju Samson	Aiden Markram	57	
..	
69	Wickets	5	Harpreet Brar	Liam Livingstone	49	

70	Wickets	7	David Miller	Jos Buttler	89
71	Runs	14	Rajat Patidar	Rajat Patidar	112
72	Wickets	7	Jos Buttler	Jos Buttler	106
73	Wickets	7	Hardik Pandya	Shubman Gill	45

	best_bowling	best_bowling_figure
0	Dwayne Bravo	3--20
1	Kuldeep Yadav	3--18
2	Mohammed Siraj	2--59
3	Mohammed Shami	3--25
4	Yuzvendra Chahal	3--22
..
69	Harpreet Brar	3--26
70	Hardik Pandya	1--14
71	Josh Hazlewood	3--43
72	Prasidh Krishna	3--22
73	Hardik Pandya	3--17

[74 rows x 20 columns]

df.head()

	match_id	date	
venue \			
0	1	March 26,2022	Wankhede Stadium,
Mumbai			
1	2	March 27,2022	Brabourne Stadium,
Mumbai			
2	3	March 27,2022	Dr DY Patil Sports Academy,
Mumbai			
3	4	March 28,2022	Wankhede Stadium,
Mumbai			
4	5	March 29,2022	Maharashtra Cricket Association
Stadium,Pune			

	team1	team2	stage	toss_winner	toss_decision
first_ings_score \					
0	Chennai	Kolkata	Group	Kolkata	Field
131					
1	Delhi	Mumbai	Group	Delhi	Field
177					
2	Banglore	Punjab	Group	Punjab	Field
205					
3	Gujarat	Lucknow	Group	Gujarat	Field
158					
4	Hyderabad	Rajasthan	Group	Hyderabad	Field

210

	first_ings_wkts	second_ings_score	second_ings_wkts	match_winner
won_by \				
0	5	133	4	Kolkata
Wickets				
1	5	179	6	Delhi
Wickets				
2	2	208	5	Punjab
Wickets				
3	6	161	5	Gujarat
Wickets				
4	6	149	7	Rajasthan
Runs				

	margin	player_of_the_match	top_scorer	highscore	
best_bowling \					
0	6	Umesh Yadav	MS Dhoni	50	Dwayne
Bravo					
1	4	Kuldeep Yadav	Ishan Kishan	81	Kuldeep
Yadav					
2	5	Odean Smith	Faf du Plessis	88	Mohammed
Siraj					
3	5	Mohammed Shami	Deepak Hooda	55	Mohammed
Shami					
4	61	Sanju Samson	Aiden Markram	57	Yuzvendra
Chahal					

	best_bowling_figure
0	3--20
1	3--18
2	2--59
3	3--25
4	3--22

df.shape

(74, 20)

df.info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 74 entries, 0 to 73

Data columns (total 20 columns):

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	match_id	74 non-null	int64
1	date	74 non-null	object
2	venue	74 non-null	object
3	team1	74 non-null	object

4	team2	74	non-null	object
5	stage	74	non-null	object
6	toss_winner	74	non-null	object
7	toss_decision	74	non-null	object
8	first_ings_score	74	non-null	int64
9	first_ings_wkts	74	non-null	int64
10	second_ings_score	74	non-null	int64
11	second_ings_wkts	74	non-null	int64
12	match_winner	74	non-null	object
13	won_by	74	non-null	object
14	margin	74	non-null	int64
15	player_of_the_match	74	non-null	object
16	top_scorer	74	non-null	object
17	highscore	74	non-null	int64
18	best_bowling	74	non-null	object
19	best_bowling_figure	74	non-null	object

dtypes: int64(7), object(13)

memory usage: 11.7+ KB

df.columns

```
Index(['match_id', 'date', 'venue', 'team1', 'team2', 'stage',
      'toss_winner',
      'toss_decision', 'first_ings_score', 'first_ings_wkts',
      'second_ings_score', 'second_ings_wkts', 'match_winner',
      'won_by',
      'margin', 'player_of_the_match', 'top_scorer', 'highscore',
      'best_bowling', 'best_bowling_figure'],
      dtype='object')
```

df.isnull().sum()

match_id	0
date	0
venue	0
team1	0
team2	0
stage	0
toss_winner	0
toss_decision	0
first_ings_score	0
first_ings_wkts	0
second_ings_score	0
second_ings_wkts	0
match_winner	0
won_by	0
margin	0
player_of_the_match	0
top_scorer	0
highscore	0

```

best_bowling      0
best_bowling_figure  0
dtype: int64

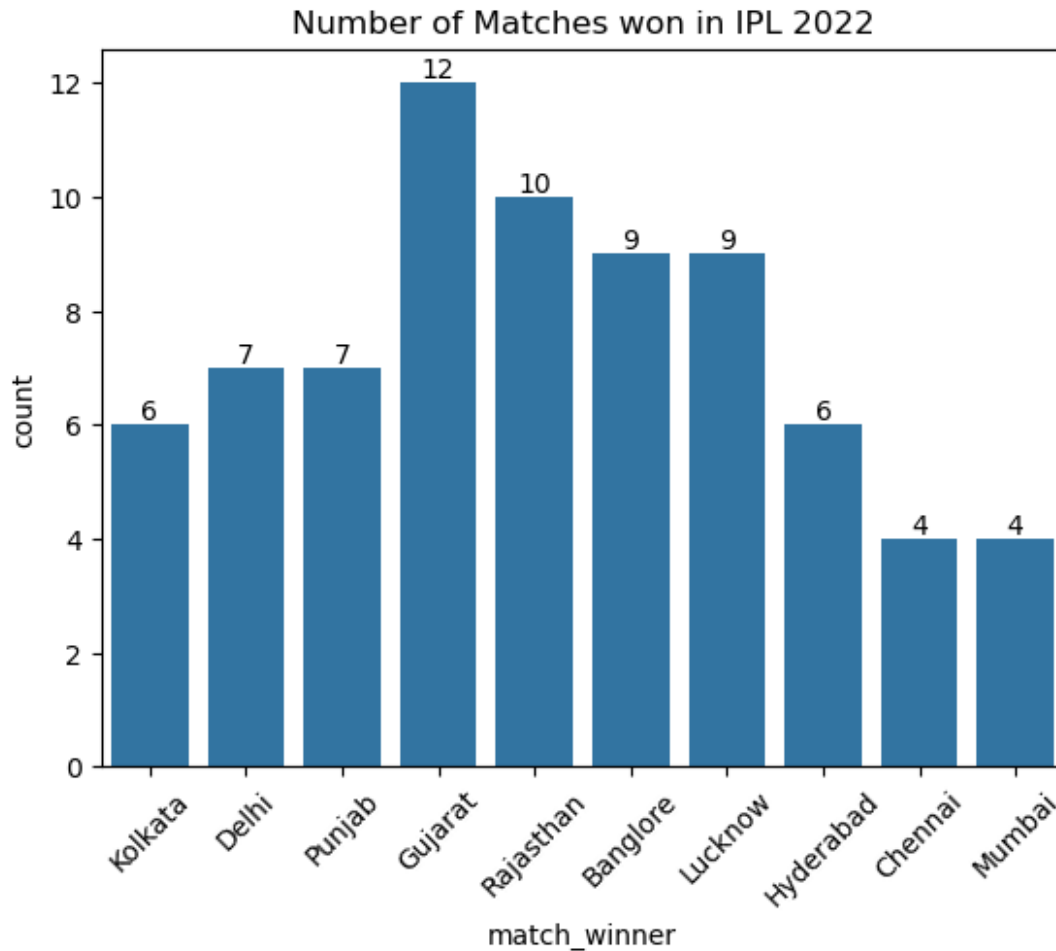
df.columns
Index(['match_id', 'date', 'venue', 'team1', 'team2', 'stage',
      'toss_winner',
      'toss_decision', 'first_ings_score', 'first_ings_wkts',
      'second_ings_score', 'second_ings_wkts', 'match_winner',
      'won_by',
      'margin', 'player_of_the_match', 'top_scorer', 'highscore',
      'best_bowling', 'best_bowling_figure'],
      dtype='object')

figure = sns.countplot(x='match_winner',data = df)
for bars in figure.containers:
    figure.bar_label(bars)

plt.title("Number of Matches won in IPL 2022")
plt.xticks(rotation=45)
figure

<Axes: title={'center': 'Number of Matches won in IPL 2022'},
xlabel='match_winner', ylabel='count'>

```



From above graphs we can see that Gujarat won the most matches(12), indicating their dominance in IPL 2022.

```
df['won_by'] = df['won_by'].map({"Wickets": "Chasing", "Runs": "Defending"})

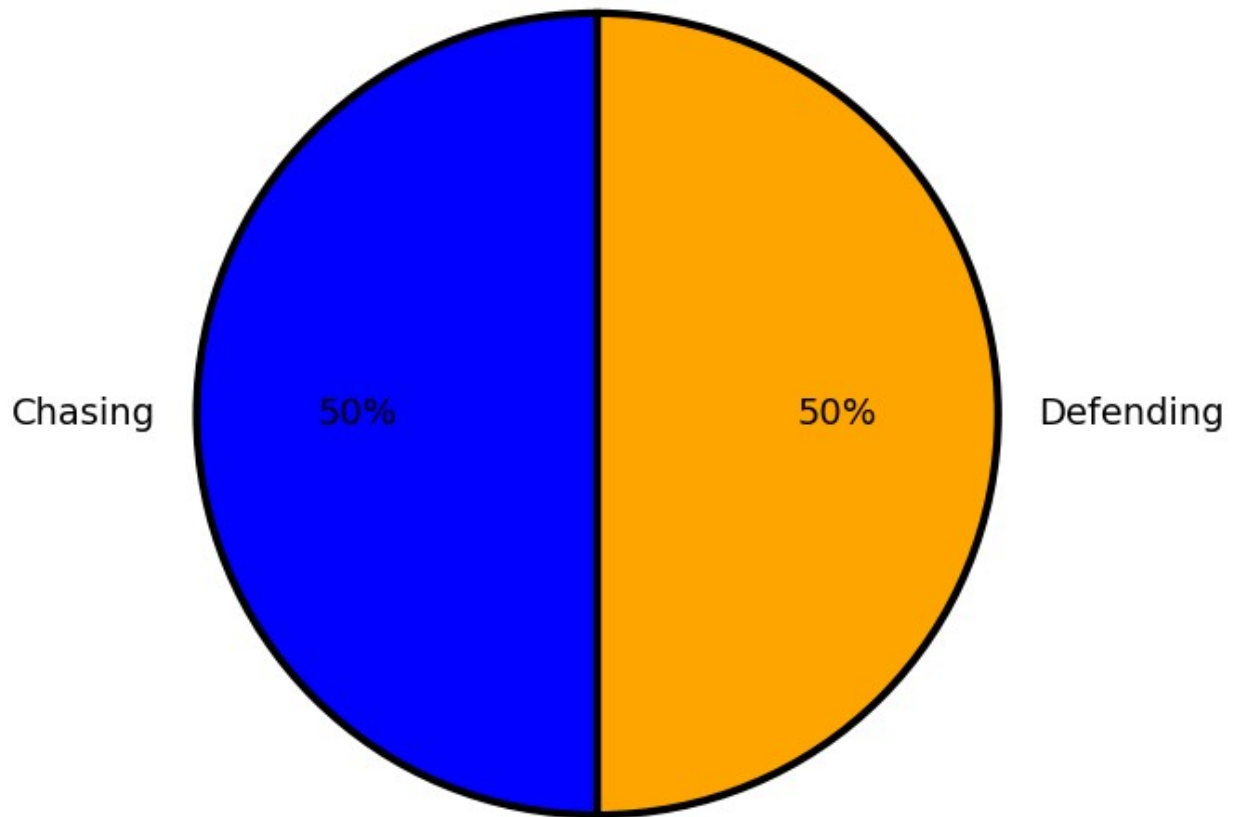
won_by_counts = df['won_by'].value_counts()
labels = won_by_counts.index
counts = won_by_counts.values
colors = ['blue', 'orange']

fig, ax = plt.subplots(figsize=(7, 7))
ax.pie(counts, labels=labels, autopct='%1.0f%%', colors=colors,
startangle=90, textprops={'fontsize': 14},
wedgeprops={'edgecolor': 'black', 'linewidth': 3})

plt.title("Percentage of Matches Won by Defending vs Chasing",
fontsize=16, color='black')

plt.show()
```

Percentage of Matches Won by Defending vs Chasing



From above graphs we can see that In IPL 2022, teams won an equal number of matches while defending and chasing (50% each), indicating a balanced competition with no clear advantage to either strategy.

```
plt.figure(figsize=(16, 8))

df_sorted = df["best_bowling"].value_counts().reset_index()
df_sorted.columns = ["Bowler", "Count"]

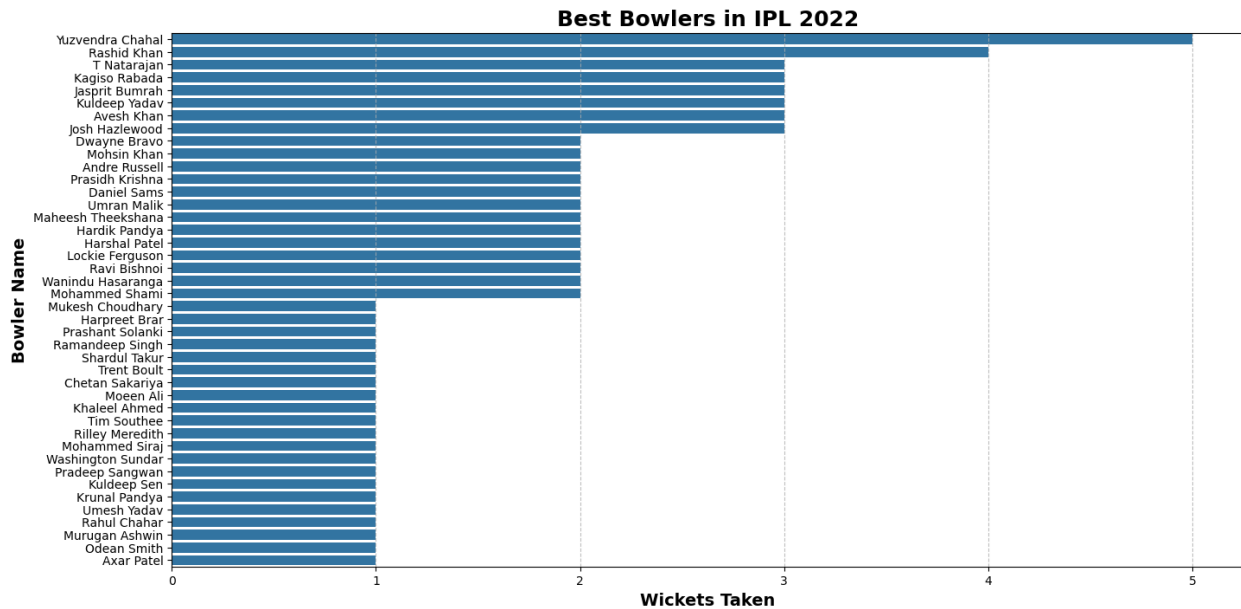
sns.barplot(y=df_sorted["Bowler"], x=df_sorted["Count"])

plt.title("Best Bowlers in IPL 2022", fontsize=18, fontweight='bold',
color="black")
plt.xlabel("Wickets Taken", fontsize=14, fontweight='bold',
color="black")
plt.ylabel("Bowler Name", fontsize=14, fontweight='bold',
color="black")
```



```
plt.grid(axis="x", linestyle="--", alpha=0.8)

plt.show()
```



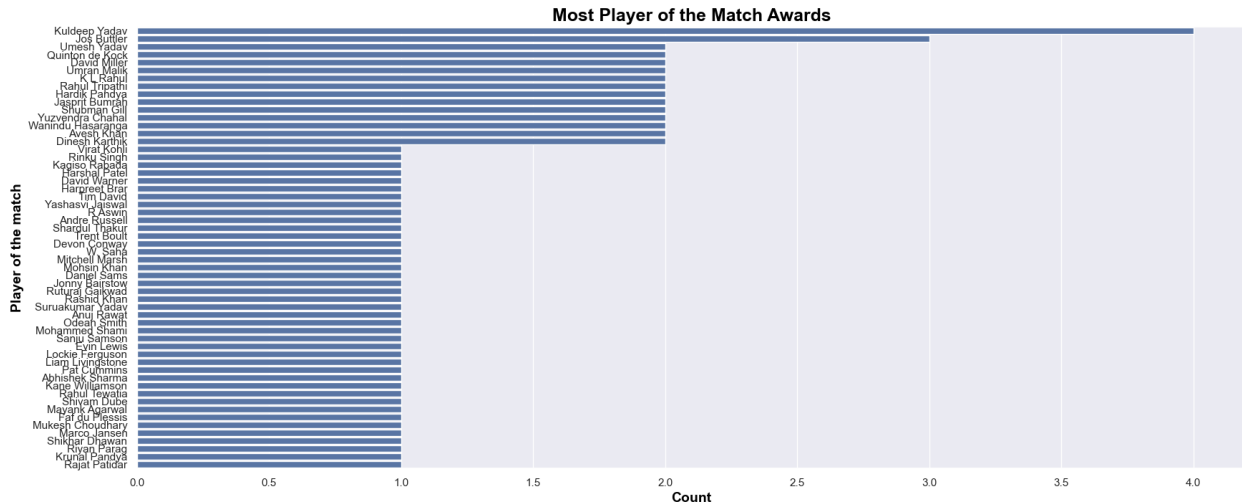
The bar graph highlights the top bowlers in IPL 2022 based on wickets taken, with Yuzvendra Chahal leading the chart, followed by other key performers like Rashid Khan and Kagiso Rabada, indicating the crucial role of spinners and pacers in determining match outcomes.

```
plt.figure(figsize=(20,8))
df_sorted = df["player_of_the_match"].value_counts().reset_index()
df_sorted.columns = ["player_of_the_match", "Count"]

sns.barplot(y=df_sorted["player_of_the_match"], x=df_sorted["Count"])

plt.title("Most Player of the Match Awards", fontsize=18,
fontweight="bold", color="black")
plt.xlabel("Count", fontsize=14, fontweight="bold", color="black")
plt.ylabel("Player of the match", fontsize=14, fontweight="bold",
color="black")

plt.show()
```



The bar graph highlights The top Player of the Match awardees include Jos Buttler, Quinton de Kock, and Kuldeep Yadav, who have received the highest number of awards.

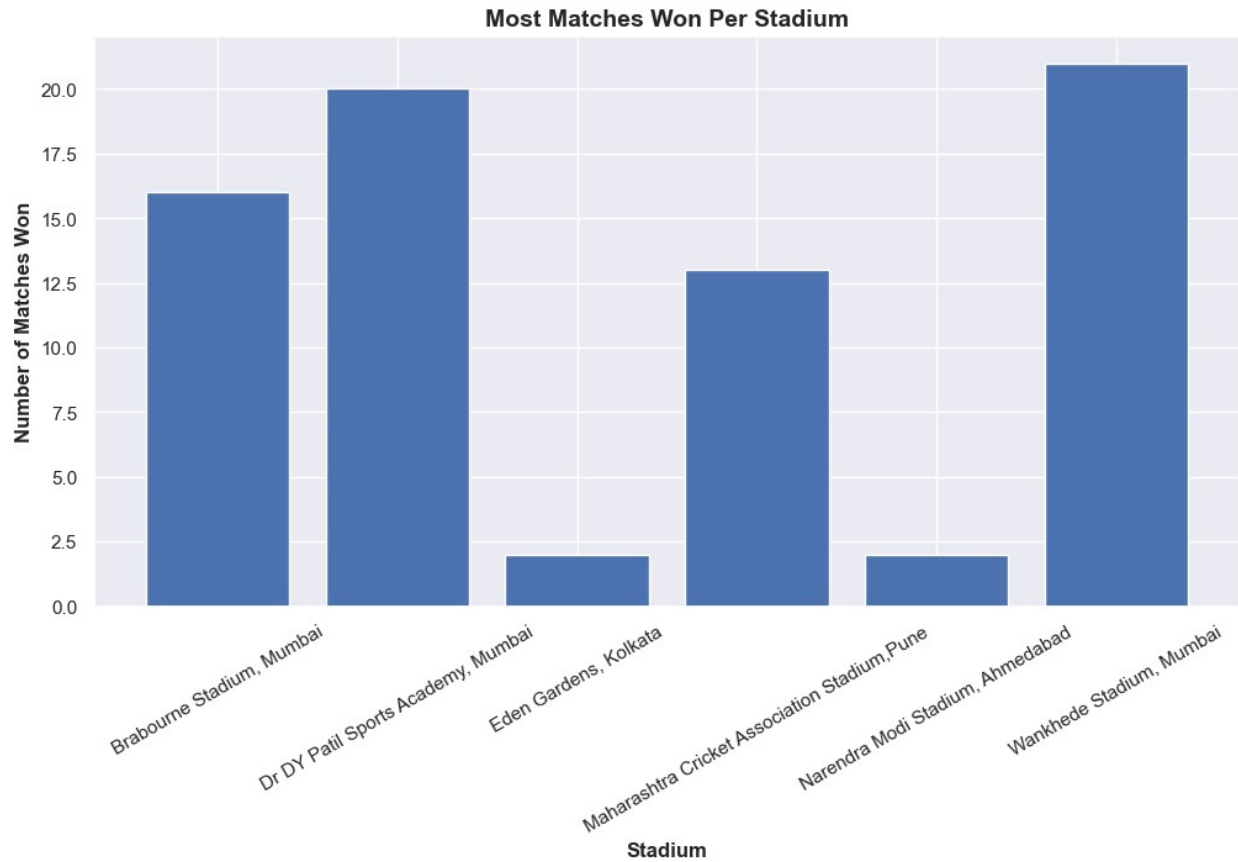
```
venue_wins = df.groupby("venue")["match_winner"].count().reset_index()
venue_wins.columns = ["Stadium", "Total_Wins"]

max_wins_stadium = venue_wins.loc[venue_wins["Total_Wins"].idxmax()]

plt.figure(figsize=(12, 6))
plt.bar(venue_wins["Stadium"], venue_wins["Total_Wins"])

plt.title("Most Matches Won Per Stadium", fontsize=14,
fontweight="bold")
plt.xlabel("Stadium", fontsize=12, fontweight="bold")
plt.ylabel("Number of Matches Won", fontsize=12, fontweight="bold")
plt.xticks(rotation=30)

plt.show()
```



```
venue_wins = df.groupby("venue")["match_winner"].count().reset_index()
venue_wins.columns = ["Stadium", "Total_Wins"]

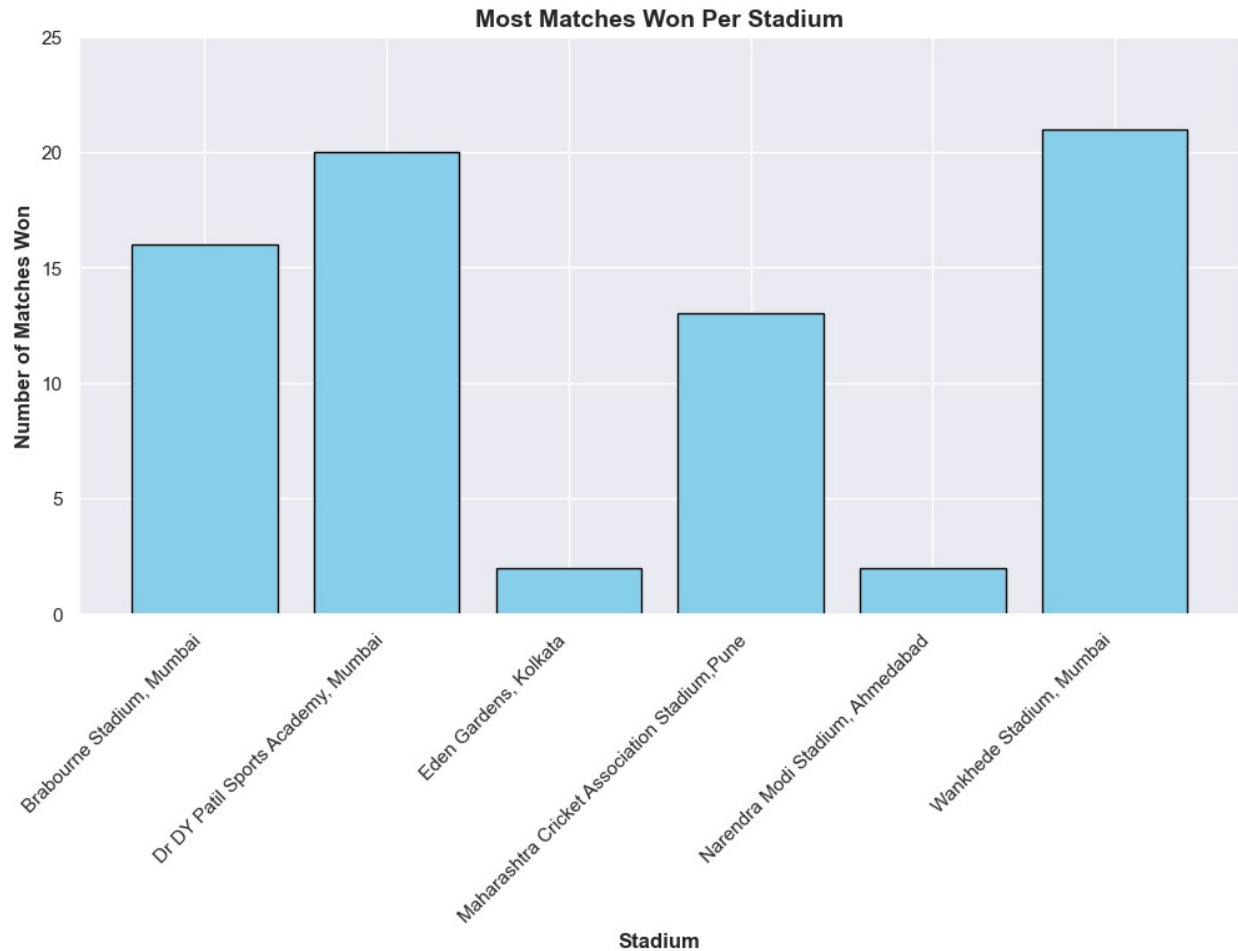
plt.figure(figsize=(12, 6))
plt.bar(venue_wins["Stadium"], venue_wins["Total_Wins"],
color="skyblue", edgecolor="black")

plt.title("Most Matches Won Per Stadium", fontsize=14,
fontweight="bold")
plt.xlabel("Stadium", fontsize=12, fontweight="bold")
plt.ylabel("Number of Matches Won", fontsize=12, fontweight="bold")-

plt.xticks(rotation=45, ha="right")

plt.yticks(range(0, max(venue_wins["Total_Wins"]) + 5, 5))

plt.show()
```



From above graphs we can see that Wankhede Stadium (Mumbai) and Dr. DY Patil Sports Academy (Mumbai) recorded the highest number of match wins, showing their significance as high-impact venues.

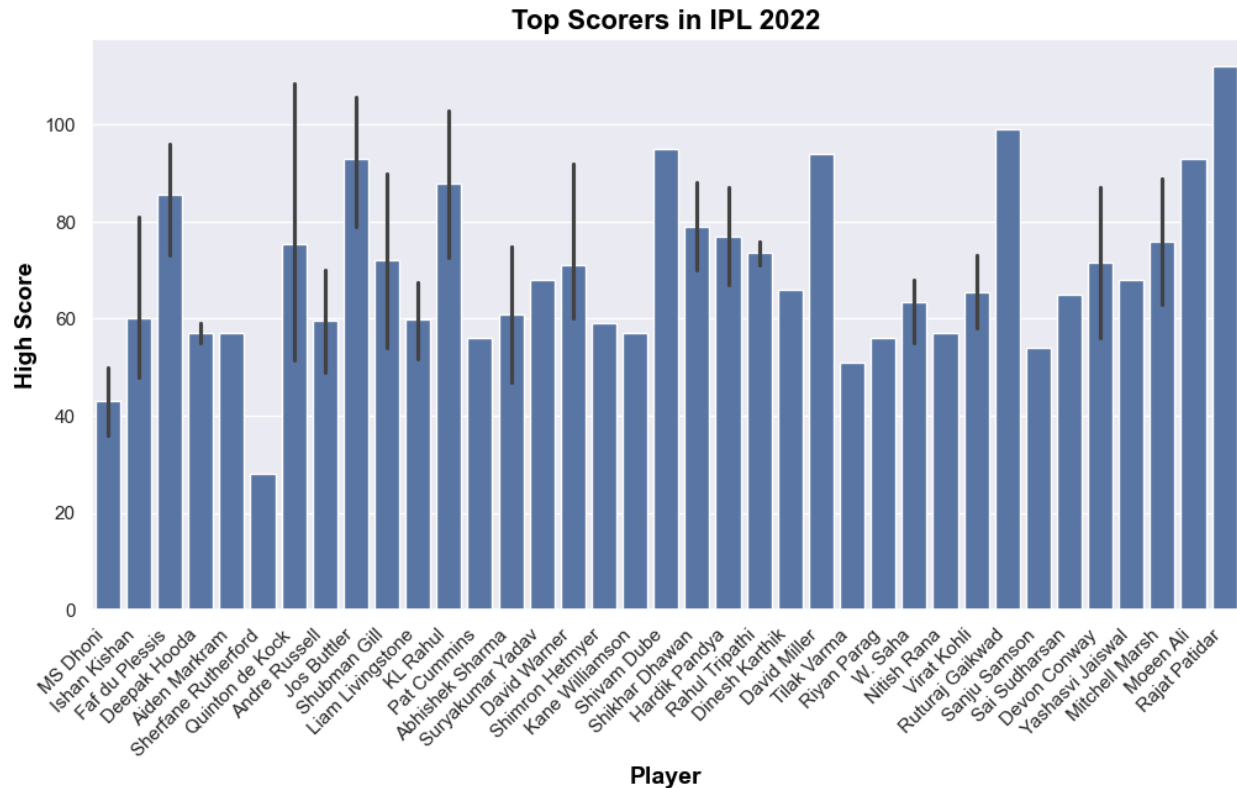
```
plt.figure(figsize=(12, 6))

# Create bar plot
sns.barplot(x=df["top_scorer"], y=df["highscore"])

# Rotate x-axis labels for better visibility
plt.xticks(rotation=45, ha="right")

# Add title and labels
plt.title("Top Scorers in IPL 2022", fontsize=16, fontweight="bold",
color="black")
plt.xlabel("Player", fontsize=14, fontweight="bold", color="black")
plt.ylabel("High Score", fontsize=14, fontweight="bold",
color="black")

plt.show()
```



```

players = df.groupby("top_scorer", as_index=False)["highscore"].sum()
players = players.sort_values(by="highscore",
ascending=False).head(10) # Ensure correct sorting

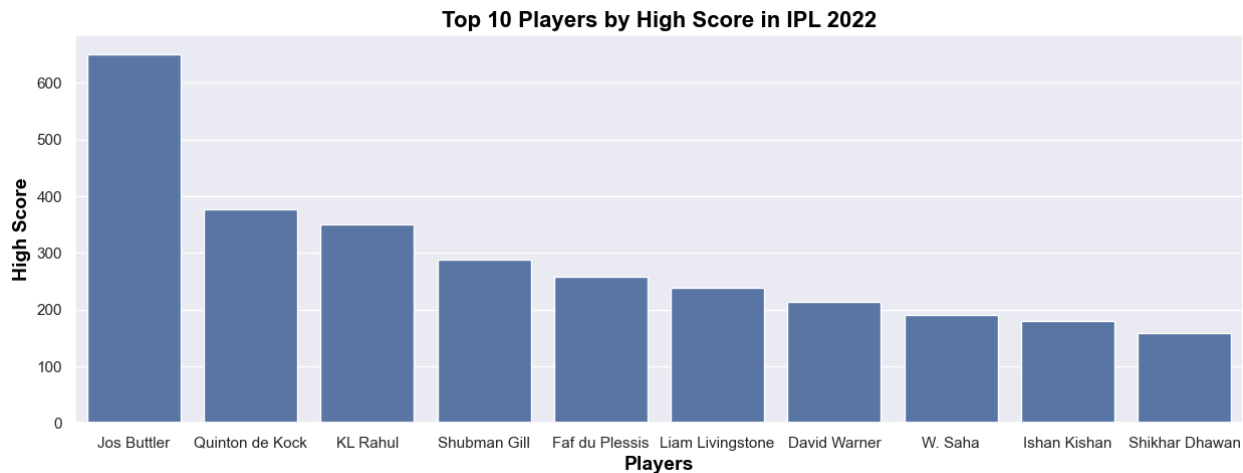
sns.set(rc={'figure.figsize':(15,5)})

sns.barplot(data=players, x="top_scorer", y="highscore")

plt.title("Top 10 Players by High Score in IPL 2022", fontsize=16,
fontweight="bold", color="black")
plt.xlabel("Players", fontsize=14, fontweight="bold", color="black")
plt.ylabel("High Score", fontsize=14, fontweight="bold",
color="black")

plt.show()

```



This graph focuses on analyzing the top 10 players with the highest individual scores in the IPL 2022 season. Jos Buttler recorded the highest individual score, exceeding 600 runs, making him the standout performer of IPL 2022.