Import Libraries and Dataset

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

sns.set_style('whitegrid')
plt.style.use('fivethirtyeight')

deliveries = pd.read_csv('deliveries.csv')
matches = pd.read_csv('matches.csv')

deliveries.head()

   match_id  inning           batting_team
bowling_team  over  \
0    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0
1    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0
2    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0
3    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0
4    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0

   ball       batter   bowler  non_striker  batsman_runs
extra_runs  \
0     1   SC Ganguly  P Kumar  BB McCullum             0           1

1     2  BB McCullum  P Kumar   SC Ganguly             0           0

2     3  BB McCullum  P Kumar   SC Ganguly             0           1

3     4  BB McCullum  P Kumar   SC Ganguly             0           0

4     5  BB McCullum  P Kumar   SC Ganguly             0           0


   total_runs extras_type  is_wicket player_dismissed dismissal_kind
fielder
0           1     legbyes          0              NaN            NaN
NaN
1           0         NaN          0              NaN            NaN
NaN
2           1       wides          0              NaN            NaN
NaN
3           0         NaN          0              NaN            NaN
NaN
```

```
4           0         NaN          0              NaN            NaN
NaN

matches.head()

       id   season        city        date match_type player_of_match
\
0  335982  2007/08   Bangalore  2008-04-18     League     BB McCullum

1  335983  2007/08  Chandigarh  2008-04-19     League      MEK Hussey

2  335984  2007/08       Delhi  2008-04-19     League     MF Maharoof

3  335985  2007/08      Mumbai  2008-04-20     League      MV Boucher

4  335986  2007/08     Kolkata  2008-04-20     League       DJ Hussey


                                        venue
team1  \
0                       M Chinnaswamy Stadium  Royal Challengers
Bangalore
1  Punjab Cricket Association Stadium, Mohali              Kings XI
Punjab
2                            Feroz Shah Kotla             Delhi
Daredevils
3                            Wankhede Stadium               Mumbai
Indians
4                                Eden Gardens        Kolkata Knight
Riders

                         team2                  toss_winner
toss_decision  \
0        Kolkata Knight Riders  Royal Challengers Bangalore
field
1          Chennai Super Kings          Chennai Super Kings
bat
2             Rajasthan Royals             Rajasthan Royals
bat
3  Royal Challengers Bangalore               Mumbai Indians
bat
4              Deccan Chargers              Deccan Chargers
bat

                        winner   result  result_margin  target_runs  \
0        Kolkata Knight Riders     runs          140.0        223.0
1          Chennai Super Kings     runs           33.0        241.0
2             Delhi Daredevils  wickets            9.0        130.0
3  Royal Challengers Bangalore  wickets            5.0        166.0
4        Kolkata Knight Riders  wickets            5.0        111.0
```

```
   target_overs super_over method    umpire1        umpire2
0          20.0          N    NaN  Asad Rauf     RE Koertzen
1          20.0          N    NaN  MR Benson      SL Shastri
2          20.0          N    NaN  Aleem Dar  GA Pratapkumar
3          20.0          N    NaN   SJ Davis       DJ Harper
4          20.0          N    NaN  BF Bowden     K Hariharan
```

```
matches.shape
```

```
(1095, 20)
```

```
deliveries.shape
```

```
(260920, 17)
```

```
matches['team1'].value_counts()
```

```
team1
Royal Challengers Bangalore    135
Chennai Super Kings            128
Mumbai Indians                 123
Kolkata Knight Riders          121
Rajasthan Royals               101
Kings XI Punjab                 92
Sunrisers Hyderabad             86
Delhi Daredevils                85
Delhi Capitals                  41
Deccan Chargers                 39
Punjab Kings                    31
Lucknow Super Giants            23
Pune Warriors                   23
Gujarat Titans                  21
Gujarat Lions                   16
Royal Challengers Bengaluru      9
Rising Pune Supergiant           7
Rising Pune Supergiants          7
Kochi Tuskers Kerala             7
Name: count, dtype: int64
```

```
matches['team2'].value_counts()
```

```
team2
Mumbai Indians                 138
Kolkata Knight Riders          130
Rajasthan Royals               120
Chennai Super Kings            110
Royal Challengers Bangalore    105
Kings XI Punjab                 98
Sunrisers Hyderabad             96
Delhi Daredevils                76
Delhi Capitals                  50
```

```
Deccan Chargers                 36
Punjab Kings                    25
Gujarat Titans                  24
Pune Warriors                   23
Lucknow Super Giants            21
Gujarat Lions                   14
Rising Pune Supergiant           9
Rising Pune Supergiants          7
Kochi Tuskers Kerala             7
Royal Challengers Bengaluru      6
Name: count, dtype: int64
```

```
deliveries['batting_team'].value_counts()
```

```
batting_team
Mumbai Indians                 31437
Kolkata Knight Riders          29514
Chennai Super Kings            28651
Royal Challengers Bangalore    28205
Rajasthan Royals               26242
Kings XI Punjab                22646
Sunrisers Hyderabad            21843
Delhi Daredevils               18786
Delhi Capitals                 10946
Deccan Chargers                 9034
Punjab Kings                    6833
Gujarat Titans                  5494
Pune Warriors                   5443
Lucknow Super Giants            5400
Gujarat Lions                   3566
Rising Pune Supergiant          1900
Royal Challengers Bengaluru     1818
Kochi Tuskers Kerala            1582
Rising Pune Supergiants         1580
Name: count, dtype: int64
```

Data Cleaning

```
matches['team1'].replace({'Rising Pune Supergiants':'Rising Pune Supergiant'}, inplace = True)
matches['team2'].replace({'Rising Pune Supergiants':'Rising Pune Supergiant'}, inplace = True)
matches['toss_winner'].replace({'Rising Pune Supergiants':'Rising Pune Supergiant'}, inplace = True)
matches['winner'].replace({'Rising Pune Supergiants':'Rising Pune Supergiant'}, inplace = True)

matches['team2'].value_counts()
```

```
team2
Mumbai Indians                 138
Kolkata Knight Riders          130
Rajasthan Royals               120
Chennai Super Kings            110
Royal Challengers Bangalore    105
Kings XI Punjab                 98
Sunrisers Hyderabad             96
Delhi Daredevils                76
Delhi Capitals                  50
Deccan Chargers                 36
Punjab Kings                    25
Gujarat Titans                  24
Pune Warriors                   23
Lucknow Super Giants            21
Rising Pune Supergiant          16
Gujarat Lions                   14
Kochi Tuskers Kerala             7
Royal Challengers Bengaluru      6
Name: count, dtype: int64
```

```
deliveries.head()
```

```
   match_id  inning           batting_team
bowling_team  over  \
0    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0
1    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0
2    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0
3    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0
4    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0

   ball       batter   bowler  non_striker  batsman_runs
extra_runs  \
0     1   SC Ganguly  P Kumar  BB McCullum             0           1

1     2  BB McCullum  P Kumar   SC Ganguly             0           0

2     3  BB McCullum  P Kumar   SC Ganguly             0           1

3     4  BB McCullum  P Kumar   SC Ganguly             0           0

4     5  BB McCullum  P Kumar   SC Ganguly             0           0


   total_runs extras_type  is_wicket player_dismissed dismissal_kind
```

```
fielder
0           1     legbyes          0              NaN          NaN
NaN
1           0         NaN          0              NaN          NaN
NaN
2           1       wides          0              NaN          NaN
NaN
3           0         NaN          0              NaN          NaN
NaN
4           0         NaN          0              NaN          NaN
NaN
```

```
deliveries['batting_team'].replace({'Rising Pune Supergiants':'Rising
Pune Supergiant'}, inplace = True)
deliveries['bowling_team'].replace({'Rising Pune Supergiants':'Rising
Pune Supergiant'}, inplace = True)
```

```
deliveries['bowling_team'].value_counts()
```

```
bowling_team
Mumbai Indians                 31505
Kolkata Knight Riders          29663
Chennai Super Kings            28576
Royal Challengers Bangalore    28358
Rajasthan Royals               26432
Kings XI Punjab                22483
Sunrisers Hyderabad            21717
Delhi Daredevils               18725
Delhi Capitals                 11216
Deccan Chargers                 9039
Punjab Kings                    6719
Pune Warriors                   5457
Gujarat Titans                  5301
Lucknow Super Giants            5226
Gujarat Lions                   3545
Rising Pune Supergiant          3543
Royal Challengers Bengaluru     1801
Kochi Tuskers Kerala            1614
Name: count, dtype: int64
```

```
matches['team1'].replace({'Royal Challengers Bengaluru':'Royal
Challengers Bangalore'}, inplace = True)
matches['team2'].replace({'Royal Challengers Bengaluru':'Royal
Challengers Bangalore'}, inplace = True)
matches['toss_winner'].replace({'Royal Challengers Bengaluru':'Royal
Challengers Bangalore'}, inplace = True)
matches['winner'].replace({'Royal Challengers Bengaluru':'Royal
Challengers Bangalore'}, inplace = True)
```

```
deliveries['batting_team'].replace({'Royal Challengers
Bengaluru':'Royal Challengers Bangalore'}, inplace = True)
deliveries['bowling_team'].replace({'Royal Challengers
Bengaluru':'Royal Challengers Bangalore'}, inplace = True)
```

```
matches['city'].isna().sum()
```

```
51
```

```
matches['method'].value_counts()
```

```
method
D/L    21
Name: count, dtype: int64
```

```
matches['city'].fillna('Unknown', inplace=True)
matches['method'].fillna('Non D/L', inplace=True)
```

Success Rate of Teams

```
def random_colors(no_of_colors):
    return['#'+''.join([np.random.choice(list('0123456789ABCDEF'))for
_ in range(6)]) for _ in range(no_of_colors)]
```

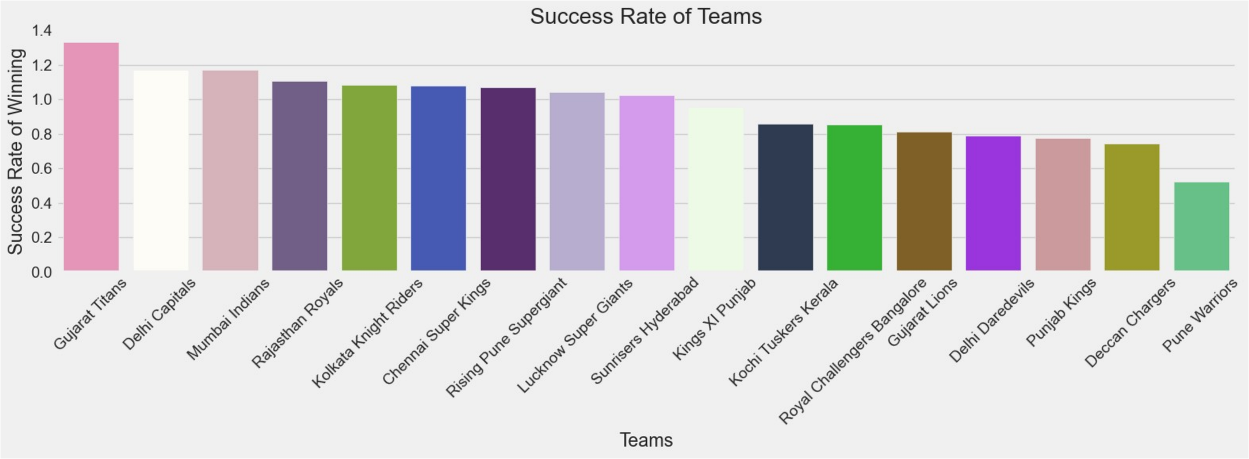
```
random_colors(6)
```

```
['#8FA63A', '#14CE40', '#FB08CC', '#533EEB', '#89ABB2', '#682982']
```

```
wins_by_team = matches.groupby('winner').count()['id']
total_matches_by_team = matches.groupby('team1').count()['id']
success_rate = wins_by_team / total_matches_by_team
success_rate_desc = success_rate.sort_values(ascending=False)
plt.figure(figsize=(16,6))
sns.barplot(
        x=success_rate_desc.index,
        y=success_rate_desc.values,
        palette=random_colors(len(success_rate_desc)),
        legend=False
    )
plt.title('Success Rate of Teams')
plt.xlabel('Teams')
plt.ylabel('Success Rate of Winning')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

```
C:\Users\RHUSHI\AppData\Local\Temp\ipykernel_12404\1006154206.py:6:
FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be
removed in v0.14.0. Assign the `x` variable to `hue` and set
`legend=False` for the same effect.
```

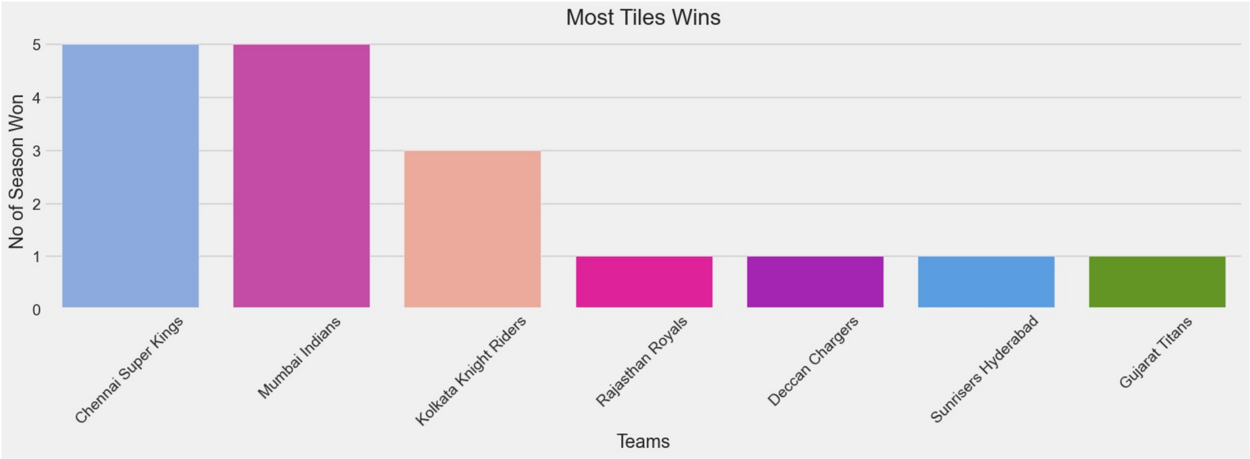
```
  sns.barplot(
```



Most Titles Wins

```python
winner_counts = matches.groupby('season')[['season',
'winner']].tail(1)['winner'].value_counts()
plt.figure(figsize=(16,6))
sns.barplot(
        x=winner_counts.index,
        y=winner_counts.values,
        palette=random_colors(len(winner_counts)),
        legend=False
    )
plt.title('Most Tiles Wins')
plt.xlabel('Teams')
plt.ylabel('No of Season Won')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

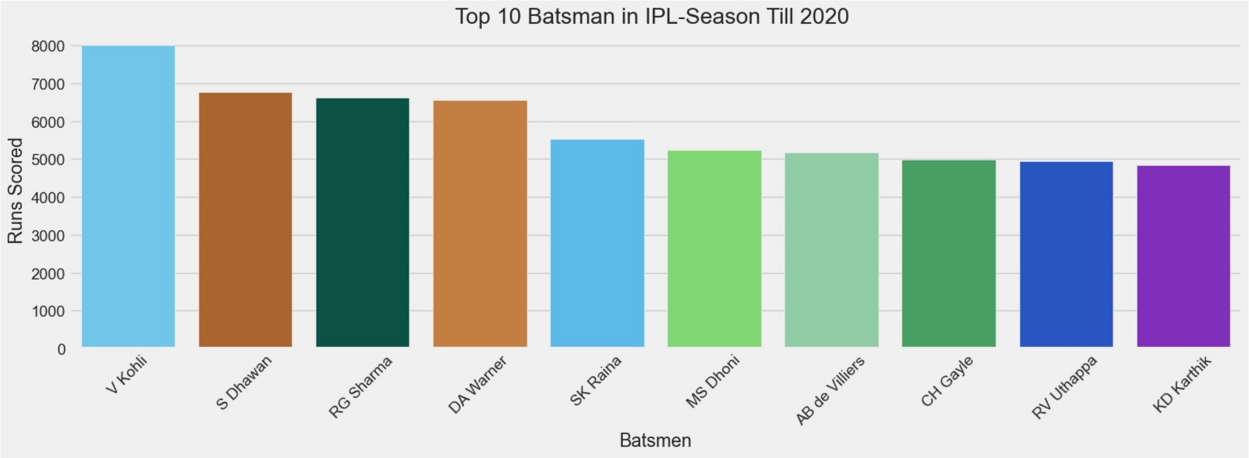
```
C:\Users\RHUSHI\AppData\Local\Temp\ipykernel_12404\3067784886.py:3:
FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be
removed in v0.14.0. Assign the `x` variable to `hue` and set
`legend=False` for the same effect.

  sns.barplot(
```



Top 10 Players with Most Runs

```
batting_total = deliveries.groupby('batter')
['batsman_runs'].sum().reset_index(name='Runs')
top_batsman = batting_total.sort_values(by='Runs', ascending=False)
[:10]
plt.figure(figsize=(16,6))
sns.barplot(
        x='batter',
        y='Runs',
        data = top_batsman,
        palette=random_colors(10),
        legend=False
    )
plt.title('Top 10 Batsman in IPL-Season Till 2020 ')
plt.xlabel('Batsmen')
plt.ylabel('Runs Scored')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

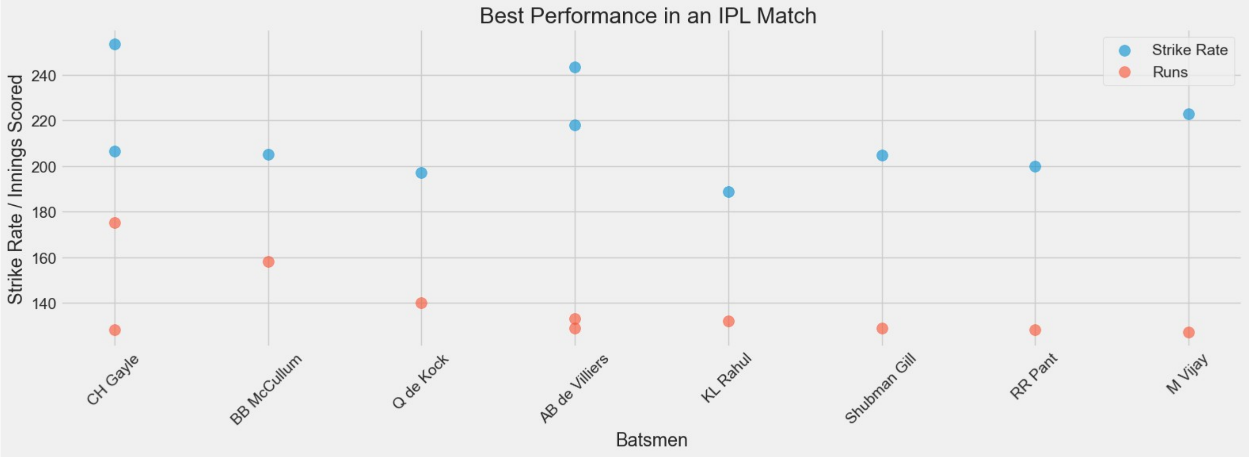
```
C:\Users\RHUSHI\AppData\Local\Temp\ipykernel_12404\2694915974.py:4:
FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be
removed in v0.14.0. Assign the `x` variable to `hue` and set
`legend=False` for the same effect.

  sns.barplot(
```



Top 10 Best Performance in a Match

```
deliveries.head(1)

   match_id  inning            batting_team
bowling_team  over  \
0    335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore     0

   ball      batter   bowler  non_striker  batsman_runs  extra_runs  \
0     1  SC Ganguly  P Kumar  BB McCullum             0           1

   total_runs extras_type  is_wicket player_dismissed dismissal_kind
fielder
0           1     legbyes          0              NaN            NaN
NaN

batting_ings = deliveries.groupby(['match_id', 'batter'])
['batsman_runs'].sum().reset_index(name='Innings Runs')
top_batsman_scores = batting_ings.sort_values(by='Innings Runs',
ascending=False)[:10]

batsman_ball_faces = deliveries.groupby(['match_id', 'batter'])
['batsman_runs'].count().reset_index(name='Balls Faced')

batsman_performance = pd.merge(top_batsman_scores, batsman_ball_faces,
how = 'inner', on=['match_id', 'batter'])
batsman_performance['Strike Rate for Match'] =
np.round(batsman_performance['Innings
Runs']*100/batsman_performance['Balls Faced'],2)

plt.figure(figsize=(16,6))
plt.scatter(batsman_performance['batter'], batsman_performance['Strike
Rate for Match'], label = 'Strike Rate', s=100, alpha=0.6)
plt.scatter(batsman_performance['batter'],
batsman_performance['Innings Runs'], label = 'Runs', s=100, alpha=0.6)
```

```
plt.title('Best Performance in an IPL Match ')
plt.xlabel('Batsmen')
plt.ylabel('Strike Rate / Innings Scored')
plt.xticks(rotation=45)
plt.legend()
plt.tight_layout()
plt.show()
```



Top 10 Bowlers till 2020

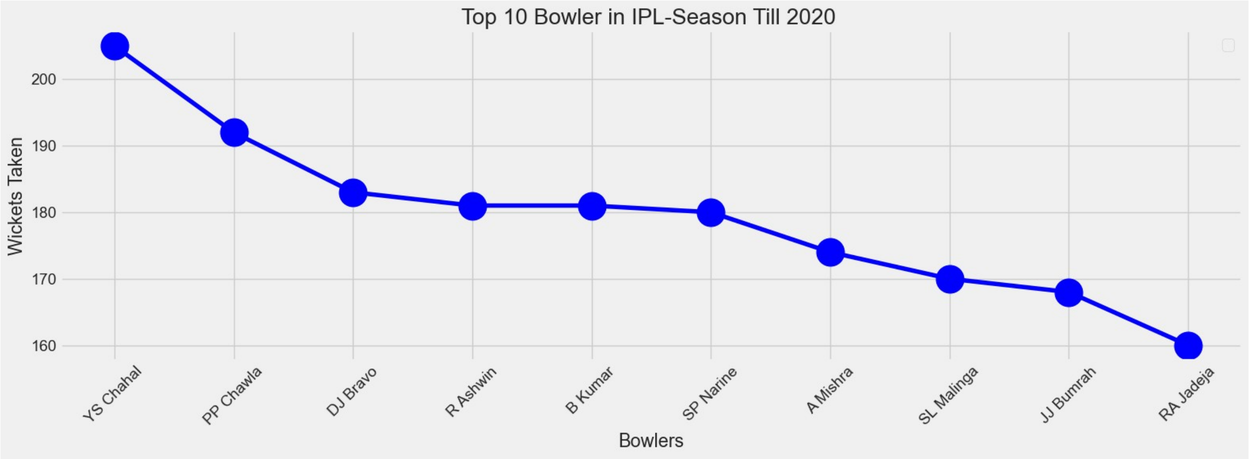
```
deliveries

        match_id  inning           batting_team
bowling_team  \
0         335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore
1         335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore
2         335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore
3         335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore
4         335982       1  Kolkata Knight Riders  Royal Challengers
Bangalore
...          ...     ...                    ...
...
260915   1426312       2  Kolkata Knight Riders        Sunrisers
Hyderabad
260916   1426312       2  Kolkata Knight Riders        Sunrisers
Hyderabad
260917   1426312       2  Kolkata Knight Riders        Sunrisers
Hyderabad
260918   1426312       2  Kolkata Knight Riders        Sunrisers
Hyderabad
260919   1426312       2  Kolkata Knight Riders        Sunrisers
Hyderabad
```

```
        over  ball       batter         bowler  non_striker
batsman_runs  \
0          0     1   SC Ganguly        P Kumar  BB McCullum
0
1          0     2  BB McCullum        P Kumar   SC Ganguly
0
2          0     3  BB McCullum        P Kumar   SC Ganguly
0
3          0     4  BB McCullum        P Kumar   SC Ganguly
0
4          0     5  BB McCullum        P Kumar   SC Ganguly
0
...      ...   ...          ...            ...          ...
...
260915     9     5      SS Iyer     AK Markram      VR Iyer
1
260916     9     6      VR Iyer     AK Markram      SS Iyer
1
260917    10     1      VR Iyer  Shahbaz Ahmed      SS Iyer
1
260918    10     2      SS Iyer  Shahbaz Ahmed      VR Iyer
1
260919    10     3      VR Iyer  Shahbaz Ahmed      SS Iyer
1

        extra_runs  total_runs extras_type  is_wicket player_dismissed
\
0                1           1     legbyes          0              NaN

1                0           0         NaN          0              NaN

2                1           1       wides          0              NaN

3                0           0         NaN          0              NaN

4                0           0         NaN          0              NaN

...            ...         ...         ...        ...              ...

260915           0           1         NaN          0              NaN

260916           0           1         NaN          0              NaN

260917           0           1         NaN          0              NaN

260918           0           1         NaN          0              NaN

260919           0           1         NaN          0              NaN
```

```
       dismissal_kind fielder
0                 NaN     NaN
1                 NaN     NaN
2                 NaN     NaN
3                 NaN     NaN
4                 NaN     NaN
...               ...     ...
260915            NaN     NaN
260916            NaN     NaN
260917            NaN     NaN
260918            NaN     NaN
260919            NaN     NaN

[260920 rows x 17 columns]
```

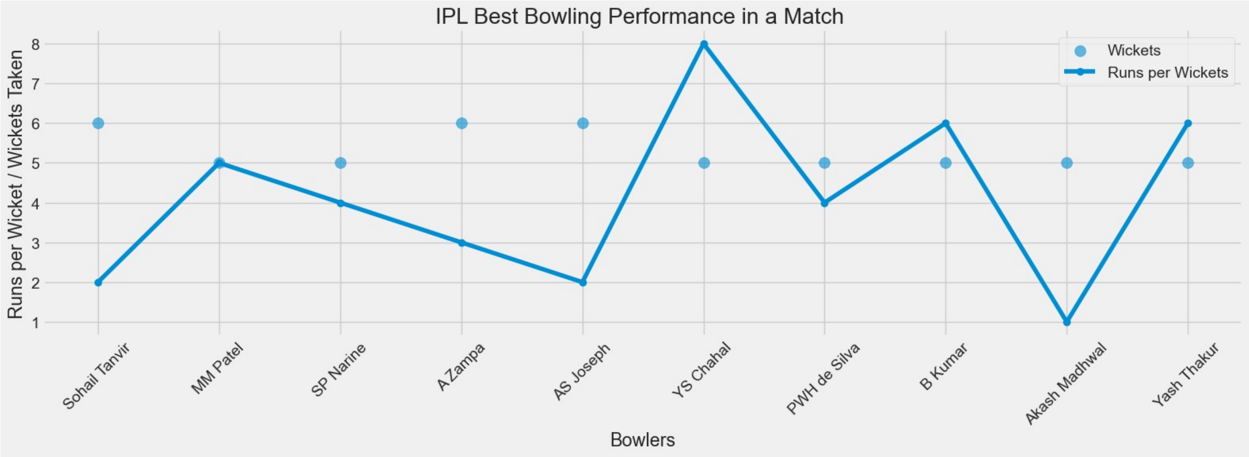
```
bowling_wickets = deliveries[deliveries['dismissal_kind'] != 'run
out']
top_bowlers = bowling_wickets.groupby('bowler')
['dismissal_kind'].count().reset_index(name='Wickets').sort_values(by=
'Wickets', ascending=False)[:10]

plt.figure(figsize=(16,6))
plt.plot(top_bowlers['bowler'], top_bowlers['Wickets'], marker = 'o',
color = 'blue', ms=25)
plt.title('Top 10 Bowler in IPL-Season Till 2020 ')
plt.xlabel('Bowlers')
plt.ylabel('Wickets Taken')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

```
C:\Users\RHUSHI\AppData\Local\Temp\ipykernel_12404\1652592482.py:10:
UserWarning: No artists with labels found to put in legend.  Note that
artists whose label start with an underscore are ignored when legend()
is called with no argument.
  plt.legend()
```



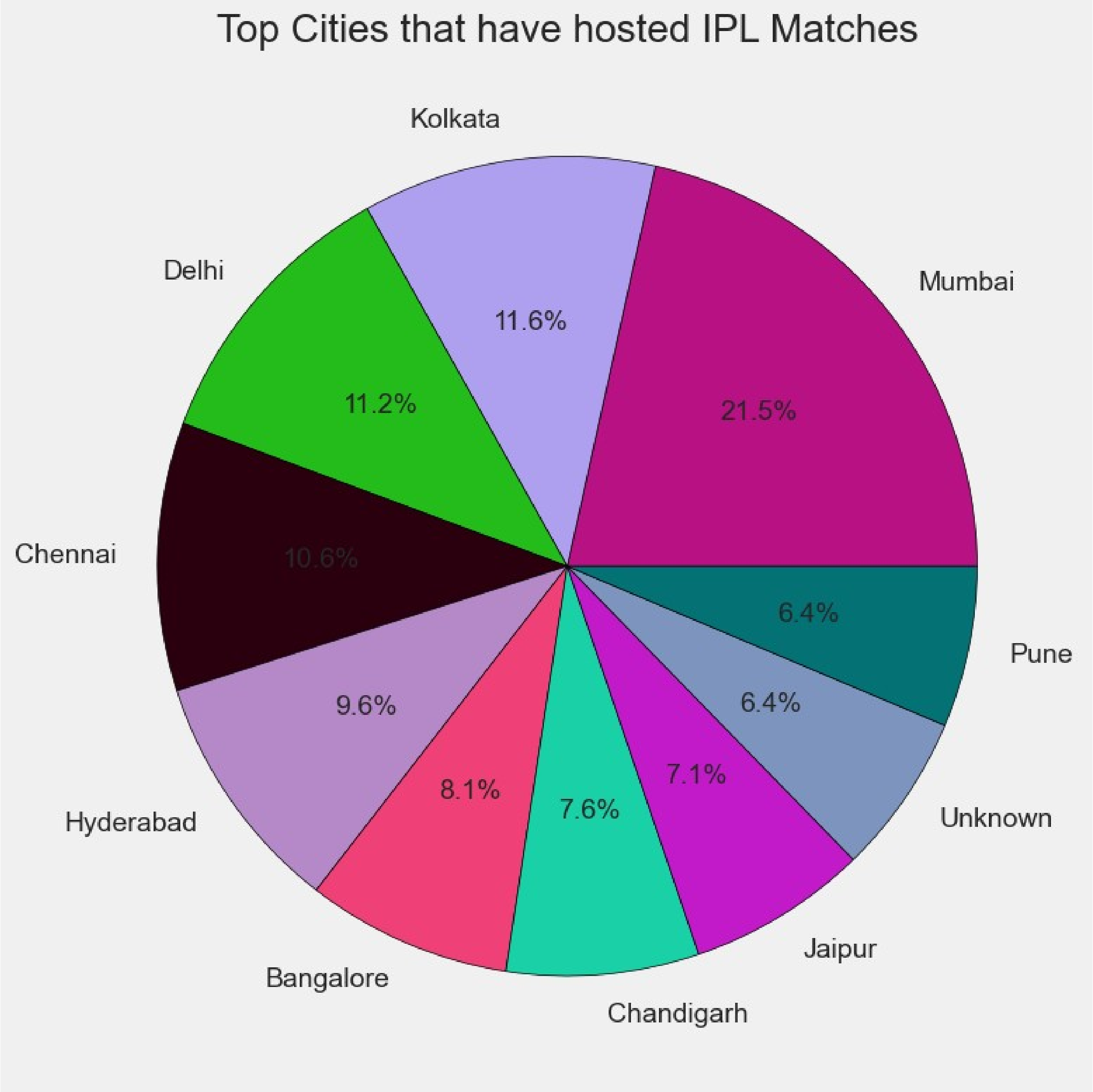
Top 10 Performance till 2020

```
match_bowling_top = bowling_wickets.groupby(['match_id', 'bowler'])
['dismissal_kind'].count().reset_index(name='Wickets').sort_values(by=
'Wickets', ascending=False)[:10]

match_bowler_runs = deliveries.groupby(['match_id', 'bowler'])
['total_runs'].sum().reset_index(name='Runs Conceded')

match_bowler_performance = pd.merge(match_bowler_runs,
match_bowling_top, how='inner', on=['match_id', 'bowler'])

match_bowler_performance['Runs per Wicket'] =
np.round(match_bowler_performance['Runs
Conceded']/match_bowler_performance['Wickets'],0)

plt.figure(figsize=(16,6))
plt.scatter(match_bowler_performance['bowler'],
match_bowler_performance['Wickets'], label = 'Wickets', s = 100, alpha
= 0.6)
plt.plot(match_bowler_performance['bowler'],
match_bowler_performance['Runs per Wicket'], marker='o', label = 'Runs
per Wickets')
plt.title('IPL Best Bowling Performance in a Match ')
plt.xlabel('Bowlers')
plt.ylabel('Runs per Wicket / Wickets Taken')
plt.xticks(rotation=45)
plt.legend()
plt.tight_layout()
plt.show()
```



Top 10 Cities by Number of Matches

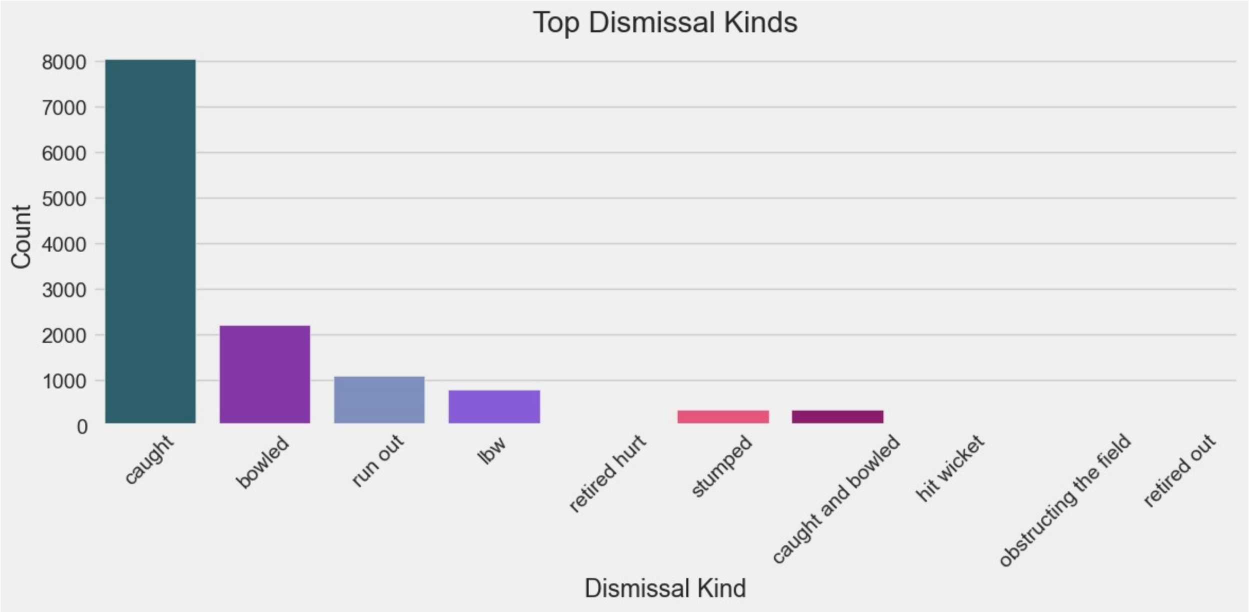
```
top_cities = matches['city'].value_counts().reset_index(name='Match
Count')[:10]
plt.figure(figsize=(8, 8))
plt.pie(top_cities['Match Count'], labels = top_cities['city'],
colors=random_colors(10), autopct='%1.1f%%',
wedgeprops=dict(edgecolor='black'))
plt.title('Top Cities that have hosted IPL Matches')
plt.tight_layout()
plt.show()
```



```
plt.figure(figsize=(12,6))
sns.countplot(x='dismissal_kind', data=deliveries,
palette=random_colors(len(deliveries['dismissal_kind'].value_counts())
))
plt.title('Top Dismissal Kinds')
```

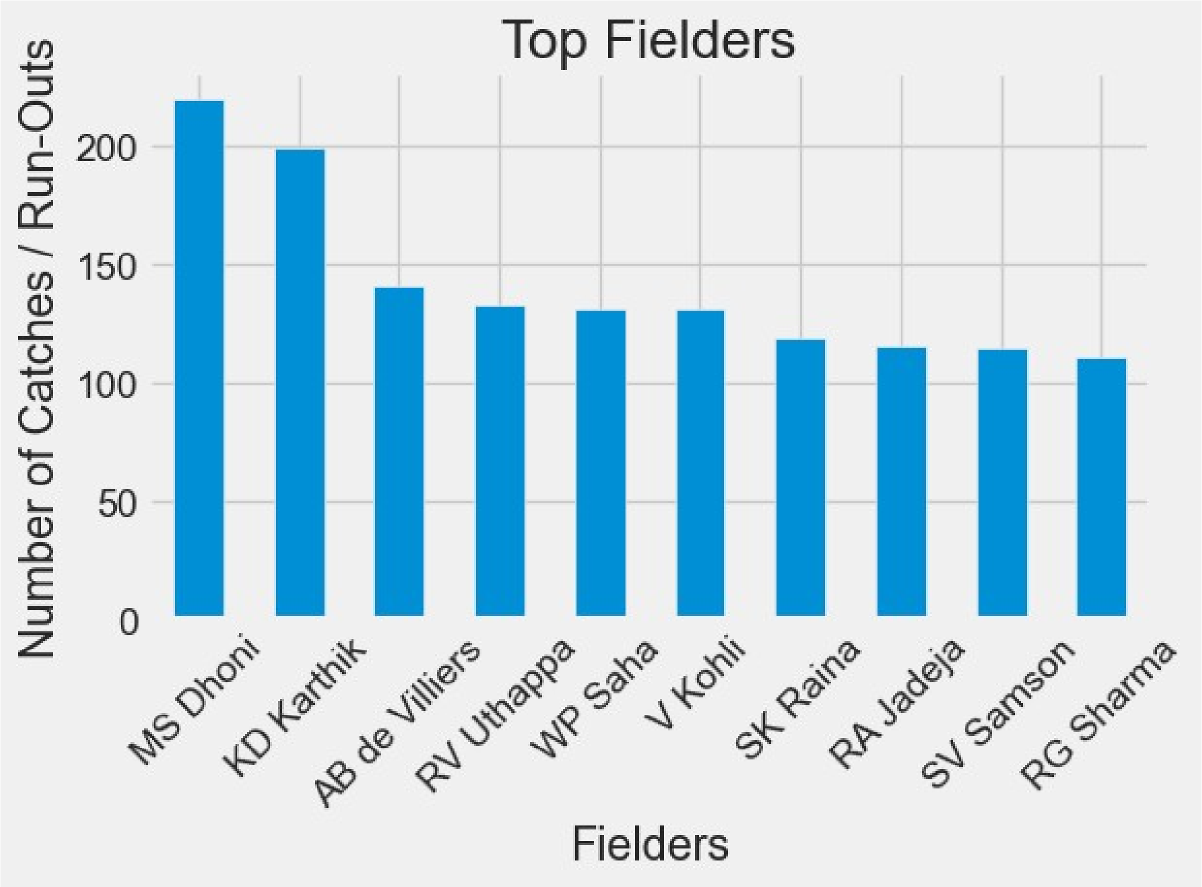
```python
plt.xlabel('Dismissal Kind')
plt.ylabel('Count')
plt.xticks(rotation = 45)
plt.tight_layout()
plt.show()
```

```
C:\Users\RHUSHI\AppData\Local\Temp\ipykernel_12404\2541909958.py:2:
FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be
removed in v0.14.0. Assign the `x` variable to `hue` and set
`legend=False` for the same effect.

  sns.countplot(x='dismissal_kind', data=deliveries,
palette=random_colors(len(deliveries['dismissal_kind'].value_counts())
))
```



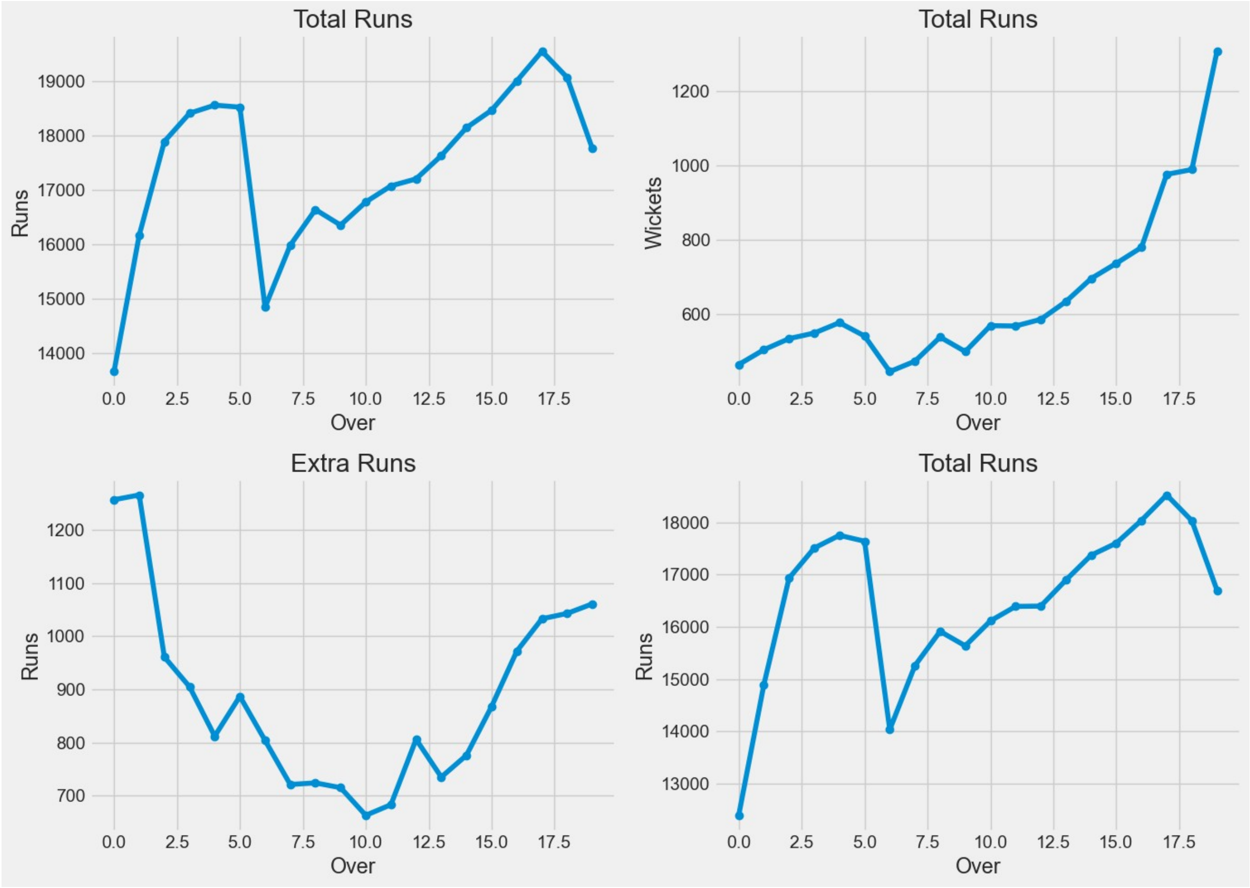
Top 10 Best Fielders

```python
deliveries['fielder'].value_counts()[:10].plot(kind='bar')

plt.title('Top Fielders')
plt.xlabel('Fielders')
plt.ylabel('Number of Catches / Run-Outs')
plt.xticks(rotation = 45)
plt.tight_layout()
plt.show()
```



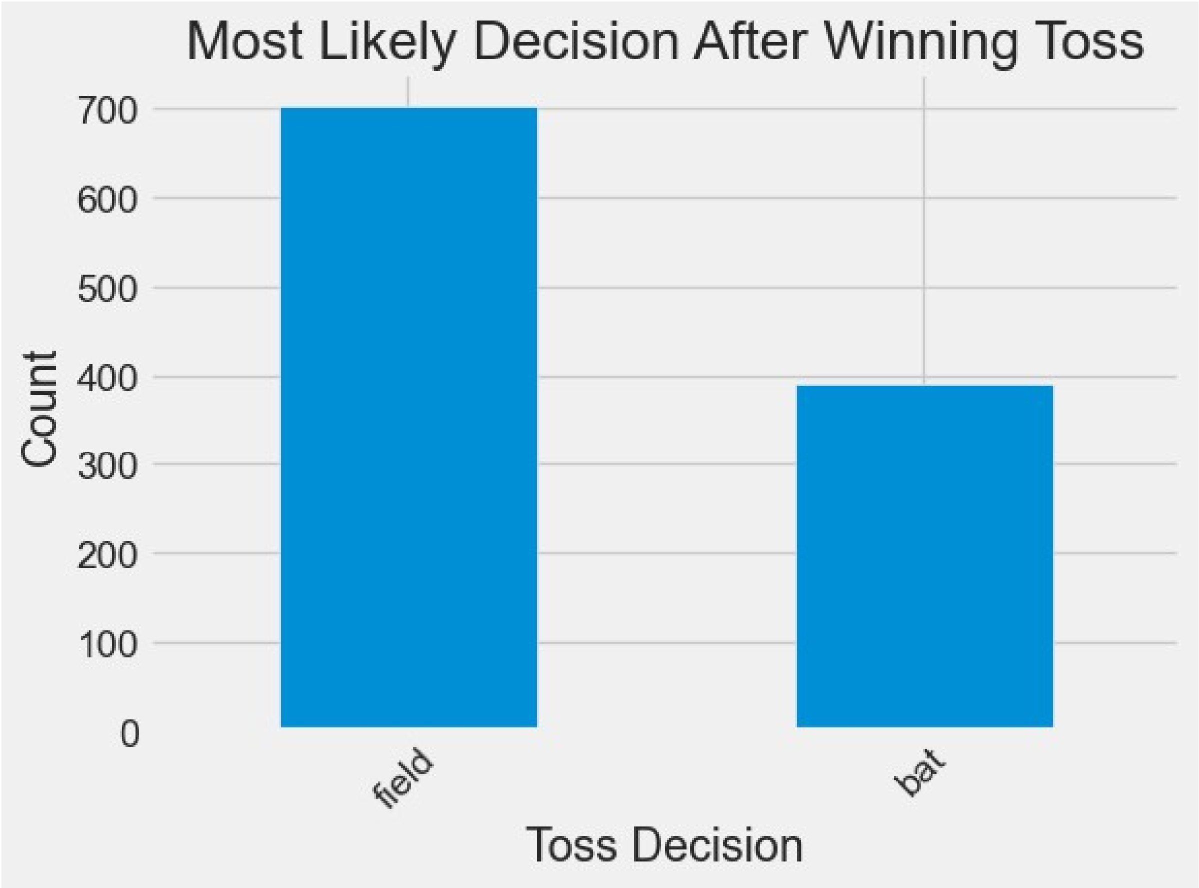
Analysis Over by Over

```
over_summary = deliveries.groupby('over')[['total_runs', 'is_wicket',
'extra_runs', 'batsman_runs']].sum()
fig, axes = plt.subplots(2,2, figsize=(14,10))
axes[0,0].plot(over_summary.index, over_summary['total_runs'],
marker='o')
axes[0,0].set_title('Total Runs')
axes[0,0].set_xlabel('Over')
axes[0,0].set_ylabel('Runs')

axes[0,1].plot(over_summary.index, over_summary['is_wicket'],
marker='o')
axes[0,1].set_title('Total Runs')
axes[0,1].set_xlabel('Over')
axes[0,1].set_ylabel('Wickets')

axes[1,0].plot(over_summary.index, over_summary['extra_runs'],
marker='o')
axes[1,0].set_title('Extra Runs')
axes[1,0].set_xlabel('Over')
axes[1,0].set_ylabel('Runs')

axes[1,1].plot(over_summary.index, over_summary['batsman_runs'],
marker='o')
axes[1,1].set_title('Total Runs')
axes[1,1].set_xlabel('Over')
axes[1,1].set_ylabel('Runs')

plt.tight_layout()
plt.show()
```



Toss Decision

```
matches['toss_decision'].value_counts()[:10].plot(kind='bar')
plt.title('Most Likely Decision After Winning Toss')
plt.xlabel('Toss Decision')
plt.ylabel('Count')
plt.xticks(rotation = 45)
plt.tight_layout()
plt.show()
```


Most Likely Decision After Winning Toss
700
600
500
400
300
200
100
0
Count
field
bat
Toss Decision