

IPL real-time data visualization

```
In [31]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
```

load data

```
In [2]: df=pd.read_csv('deliveries.csv')
```

Basic Exploration of Data

```
In [3]: df.head()
```

	match_id	inning	batting_team	bowling_team	over	ball	batter	bowler	non_striker	batstype
0	335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	1	SC Ganguly	P Kumar	BB McCullum	BB
1	335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	2	BB McCullum	P Kumar	SC Ganguly	
2	335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	3	BB McCullum	P Kumar	SC Ganguly	
3	335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	4	BB McCullum	P Kumar	SC Ganguly	
4	335982	1	Kolkata Knight Riders	Royal Challengers Bangalore	0	5	BB McCullum	P Kumar	SC Ganguly	

```
In [4]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 260920 entries, 0 to 260919
Data columns (total 17 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   match_id         260920 non-null   int64  
 1   inning            260920 non-null   int64  
 2   batting_team     260920 non-null   object  
 3   bowling_team     260920 non-null   object  
 4   over              260920 non-null   int64  
 5   ball              260920 non-null   int64  
 6   batter            260920 non-null   object  
 7   bowler             260920 non-null   object  
 8   non_striker       260920 non-null   object  
 9   batsman_runs      260920 non-null   int64  
 10  extra_runs        260920 non-null   int64  
 11  total_runs        260920 non-null   int64  
 12  extras_type       14125 non-null    object  
 13  is_wicket          260920 non-null   int64  
 14  player_dismissed 12950 non-null    object  
 15  dismissal_kind    12950 non-null    object  
 16  fielder            9354 non-null    object  
dtypes: int64(8), object(9)
memory usage: 33.8+ MB
```

In [32]: `df.describe()`

	match_id	inning	over	ball	batsman_runs	extra_runs	
count	2.609200e+05	260920.000000	260920.000000	260920.000000	260920.000000	260920.000000	260
mean	9.070665e+05	1.483531	9.197677	3.624486	1.265001	0.067806	
std	3.679913e+05	0.502643	5.683484	1.814920	1.639298	0.343265	
min	3.359820e+05	1.000000	0.000000	1.000000	0.000000	0.000000	
25%	5.483340e+05	1.000000	4.000000	2.000000	0.000000	0.000000	
50%	9.809670e+05	1.000000	9.000000	4.000000	1.000000	0.000000	
75%	1.254066e+06	2.000000	14.000000	5.000000	1.000000	0.000000	
max	1.426312e+06	6.000000	19.000000	11.000000	6.000000	7.000000	

In [33]: `df.iloc[0]`

```
Out[33]:
```

match_id	335982
inning	1
batting_team	Kolkata Knight Riders
bowling_team	Royal Challengers Bangalore
over	0
ball	1
batter	SC Ganguly
bowler	P Kumar
non_striker	BB McCullum
batsman_runs	0
extra_runs	1
total_runs	1
extras_type	legbyes
is_wicket	0
player_dismissed	NaN
dismissal_kind	NaN
fielder	NaN

Name: 0, dtype: object

```
In [7]: df.iloc[1]
```

```
Out[7]:
```

match_id	335982
inning	1
batting_team	Kolkata Knight Riders
bowling_team	Royal Challengers Bangalore
over	0
ball	2
batter	BB McCullum
bowler	P Kumar
non_striker	SC Ganguly
batsman_runs	0
extra_runs	0
total_runs	0
extras_type	NaN
is_wicket	0
player_dismissed	NaN
dismissal_kind	NaN
fielder	NaN

Name: 1, dtype: object

```
In [8]: print(df.isnull())
```

```

match_id    inning    batting_team    bowling_team    over    ball    batter    \
0           False     False          False          False   False   False   False
1           False     False          False          False   False   False   False
2           False     False          False          False   False   False   False
3           False     False          False          False   False   False   False
4           False     False          False          False   False   False   False
...
260915     False     False          False          False   False   False   False
260916     False     False          False          False   False   False   False
260917     False     False          False          False   False   False   False
260918     False     False          False          False   False   False   False
260919     False     False          False          False   False   False   False

bowler      non_striker    batsman_runs    extra_runs    total_runs    \
0           False        False          False          False       False
1           False        False          False          False       False
2           False        False          False          False       False
3           False        False          False          False       False
4           False        False          False          False       False
...
260915     False        False          False          False       False
260916     False        False          False          False       False
260917     False        False          False          False       False
260918     False        False          False          False       False
260919     False        False          False          False       False

extras_type  is_wicket    player_dismissed    dismissal_kind    fielder
0           False        False            True           True       True
1           True         False            True           True       True
2           False        False            True           True       True
3           True         False            True           True       True
4           True         False            True           True       True
...
260915     True         False            True           True       True
260916     True         False            True           True       True
260917     True         False            True           True       True
260918     True         False            True           True       True
260919     True         False            True           True       True

```

[260920 rows x 17 columns]

In [9]: `df.isnull().sum()`

```
Out[9]:
```

match_id	0
inning	0
batting_team	0
bowling_team	0
over	0
ball	0
batter	0
bowler	0
non_striker	0
batsman_runs	0
extra_runs	0
total_runs	0
extras_type	246795
is_wicket	0
player_dismissed	247970
dismissal_kind	247970
fielder	251566

dtype: int64

```
In [34]: df.dropna()
```

Out[34]:

	match_id	inning	batting_team	bowling_team	over	ball	batter	bowler	non_striker
6670	336010	1	Kolkata Knight Riders	Royal Challengers Bangalore	8	6	DJ Hussey	A Kumble	T Taibu
9022	336020	1	Delhi Daredevils	Deccan Chargers	15	2	G Gambhir	PP Ojha	S Dhawan
21707	392217	2	Chennai Super Kings	Rajasthan Royals	16	6	ML Hayden	SK Warne	S Badrinath
21927	392218	2	Royal Challengers Bangalore	Mumbai Indians	13	2	RE van der Merwe	Harbhajan Singh	MV Boucher
22808	392222	1	Kings XI Punjab	Mumbai Indians	17	6	K Goel	SL Malinga	PP Chawla
24387	392228	2	Deccan Chargers	Kolkata Knight Riders	17	1	A Symonds	Mashrafe Mortaza	RG Sharma
29246	419114	2	Delhi Daredevils	Mumbai Indians	8	5	KD Karthik	Harbhajan Singh	MF Maharoff
38560	419153	1	Kolkata Knight Riders	Chennai Super Kings	4	1	DJ Hussey	R Ashwin	MK Tiwary
40074	419159	2	Chennai Super Kings	Kings XI Punjab	4	1	M Vijay	RR Powar	SK Raina
40814	419162	2	Royal Challengers Bangalore	Mumbai Indians	6	2	KP Pietersen	Harbhajan Singh	R Dravid
42421	501201	1	Delhi Daredevils	Mumbai Indians	16	4	UT Yadav	Harbhajan Singh	RE van der Merwe
46496	501219	2	Kolkata Knight Riders	Kochi Tuskers Kerala	16	6	MK Tiwary	RA Jadeja	R Bhatia
54075	501251	2	Mumbai Indians	Kings XI Punjab	3	6	RG Sharma	Bipul Sharma	AC Blizzard
74829	548375	2	Pune Warriors	Kolkata Knight Riders	7	5	MJ Clarke	YK Pathan	JD Ryder
79791	598012	1	Royal Challengers Bangalore	Chennai Super Kings	8	5	MA Agarwal	R Ashwin	V Kohli
99050	729315	3	Kolkata Knight Riders	Rajasthan Royals	0	1	SA Yadav	JP Faulkner	Shakib Al Hasan
100461	733979	2	Kolkata Knight Riders	Rajasthan Royals	15	1	MK Pandey	PV Tambe	Shakib Al Hasan
101590	733989	1	Sunrisers Hyderabad	Rajasthan Royals	19	3	IK Pathan	JP Faulkner	DW Steyn
102297	733995	2	Chennai Super Kings	Mumbai Indians	7	7	SK Raina	KA Pollard	DR Smith
106634	734031	1	Mumbai Indians	Delhi Daredevils	19	4	JJ Bumrah	WD Parnell	S Gopal

	match_id	inning	batting_team	bowling_team	over	ball	batter	bowler	non_striker
109324	829707	2	Delhi Daredevils	Chennai Super Kings	18	5	A Mishra	A Nehra	JA Morkel
109961	829713	1	Kolkata Knight Riders	Royal Challengers Bangalore	19	6	Shakib Al Hasan	VR Aaron	AD Russell
115702	829763	1	Royal Challengers Bangalore	Rajasthan Royals	16	3	SN Khan	SR Watson	KD Karthik
116439	829769	2	Rajasthan Royals	Mumbai Indians	19	5	STR Binny	R Vinay Kumar	TG Southee
117053	829775	2	Kings XI Punjab	Mumbai Indians	18	5	WP Saha	SL Malinga	GJ Bailey
120237	829803	1	Kings XI Punjab	Royal Challengers Bangalore	5	5	DA Miller	YS Chahal	GJ Bailey
122130	829821	1	Royal Challengers Bangalore	Chennai Super Kings	19	4	HV Patel	DJ Bravo	SN Khan
124384	980917	1	Mumbai Indians	Gujarat Lions	8	6	JC Buttler	SB Jakati	PA Patel
130652	980969	1	Rising Pune Supergiants	Royal Challengers Bangalore	14	2	SS Tiwary	YS Chahal	AM Rahane
133061	980989	2	Kolkata Knight Riders	Rising Pune Supergiants	0	2	RV Uthappa	R Ashwin	G Gambhir
136969	1082592	1	Mumbai Indians	Rising Pune Supergiant	19	6	HH Pandya	AB Dinda	TG Southee
142909	1082617	1	Kolkata Knight Riders	Royal Challengers Bangalore	9	2	YK Pathan	YS Chahal	MK Pandey
144492	1082625	1	Gujarat Lions	Mumbai Indians	8	2	KD Karthik	KH Pandya	Ishan Kishan
147815	1082639	1	Kings XI Punjab	Kolkata Knight Riders	17	6	WP Saha	Kuldeep Yadav	AR Patel
153403	1136573	2	Delhi Daredevils	Kolkata Knight Riders	0	6	JJ Roy	PP Chawla	G Gambhir
153675	1136574	2	Royal Challengers Bangalore	Mumbai Indians	9	4	Mandeep Singh	KH Pandya	V Kohli
159860	1136600	2	Kings XI Punjab	Rajasthan Royals	2	1	CH Gayle	K Gowtham	KL Rahul
165867	1175360	2	Chennai Super Kings	Delhi Capitals	6	5	SR Watson	A Mishra	SK Raina
178494	1181766	2	Delhi Capitals	Sunrisers Hyderabad	7	4	S Dhawan	DJ Hooda	PP Shaw

	match_id	inning	batting_team	bowling_team	over	ball	batter	bowler	non_striker
201184	1254089	1	Chennai Super Kings	Rajasthan Royals	14	4	MM Ali	R Tewatia	RD Gaikwad
208092	1304047	1	Chennai Super Kings	Kolkata Knight Riders	7	6	RV Uthappa	CV Varun	AT Rayudu
210216	1304055	2	Mumbai Indians	Rajasthan Royals	18	9	KA Pollard	M Prasidh Krishna	M Ashwin
220309	1304097	1	Mumbai Indians	Gujarat Titans	18	3	Tilak Varma	LH Ferguson	TH David
224880	1304116	1	Sunrisers Hyderabad	Punjab Kings	19	6	R Shepherd	NT Ellis	B Kumar
225101	1312197	1	Rajasthan Royals	Gujarat Titans	19	6	JC Buttler	Yash Dayal	R Parag
225102	1312197	1	Rajasthan Royals	Gujarat Titans	19	7	R Ashwin	Yash Dayal	R Parag
233329	1359505	1	Punjab Kings	Mumbai Indians	9	2	LS Livingstone	PP Chawla	Atharva Taide
233639	1359506	1	Royal Challengers Bangalore	Rajasthan Royals	19	1	PWH de Silva	Sandeep Sharma	KD Karthik
234879	1359511	1	Rajasthan Royals	Chennai Super Kings	19	4	D Padikkal	M Pathirana	Dhruv Jurel
236509	1359518	1	Delhi Capitals	Gujarat Titans	1	2	PK Garg	HH Pandya	DA Warner
242686	1359544	1	Royal Challengers Bangalore	Gujarat Titans	9	3	MK Lomror	Noor Ahmad	V Kohli
243398	1370352	1	Gujarat Titans	Mumbai Indians	6	3	WP Saha	PP Chawla	Shubman Gill
249678	1426262	2	Gujarat Titans	Rajasthan Royals	15	3	Shubman Gill	YS Chahal	R Tewatia
251484	1426270	1	Gujarat Titans	Delhi Capitals	8	5	M Shahrukh Khan	T Stubbs	R Tewatia
252273	1426273	2	Delhi Capitals	Sunrisers Hyderabad	8	5	Abishek Porel	M Markande	T Stubbs
260426	1426310	2	Rajasthan Royals	Royal Challengers Bengaluru	10	1	SV Samson	KV Sharma	R Parag

In [11]: `#df=df.fillna(method='tfill')`

In [12]: `df.columns`

```
Out[12]: Index(['match_id', 'inning', 'batting_team', 'bowling_team', 'over', 'ball',
   'batter', 'bowler', 'non_striker', 'batsman_runs', 'extra_runs',
   'total_runs', 'extras_type', 'is_wicket', 'player_dismissed',
   'dismissal_kind', 'fielder'],
  dtype='object')
```

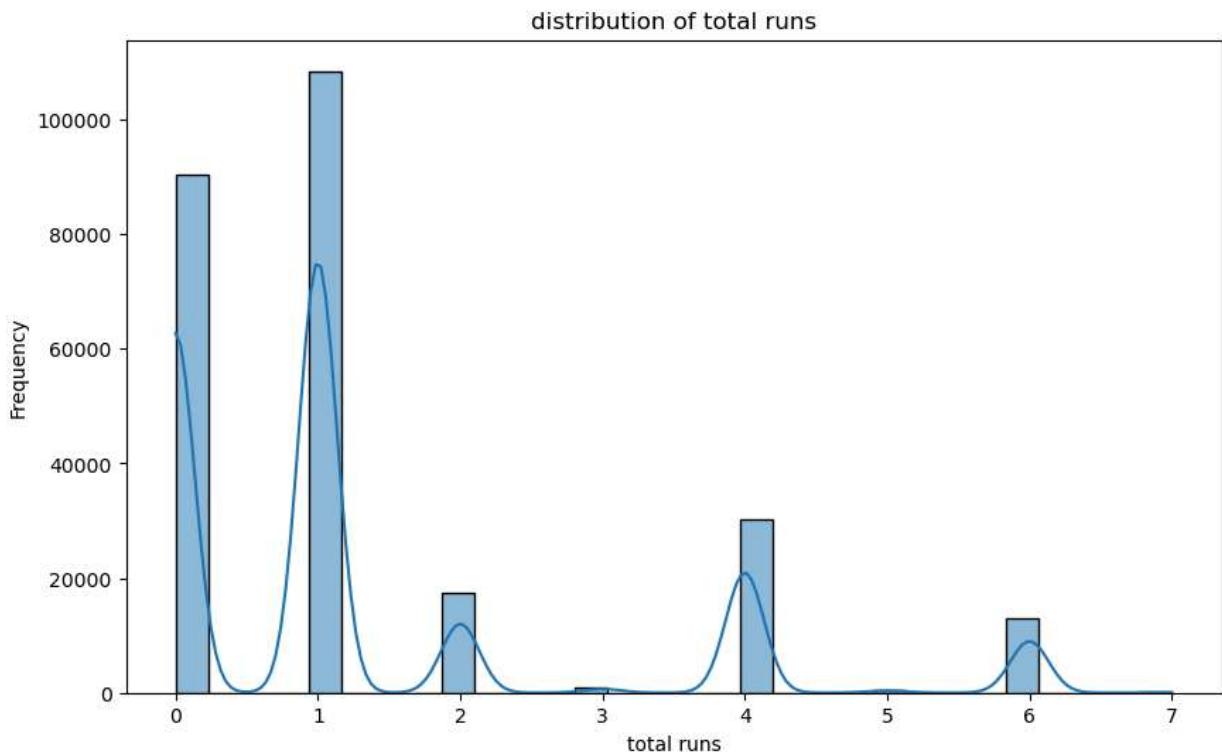
Visualizing Key Columns

Total runs distribution across overs

```
In [35]: plt.figure(figsize=(10,6))
sns.histplot(df['total_runs'], kde=True, bins=30)
plt.title('distribution of total runs')
plt.xlabel("total runs")
plt.ylabel('Frequency')
plt.show()
```

d:\Users\hi\anaconda3\Lib\site-packages\seaborn_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.

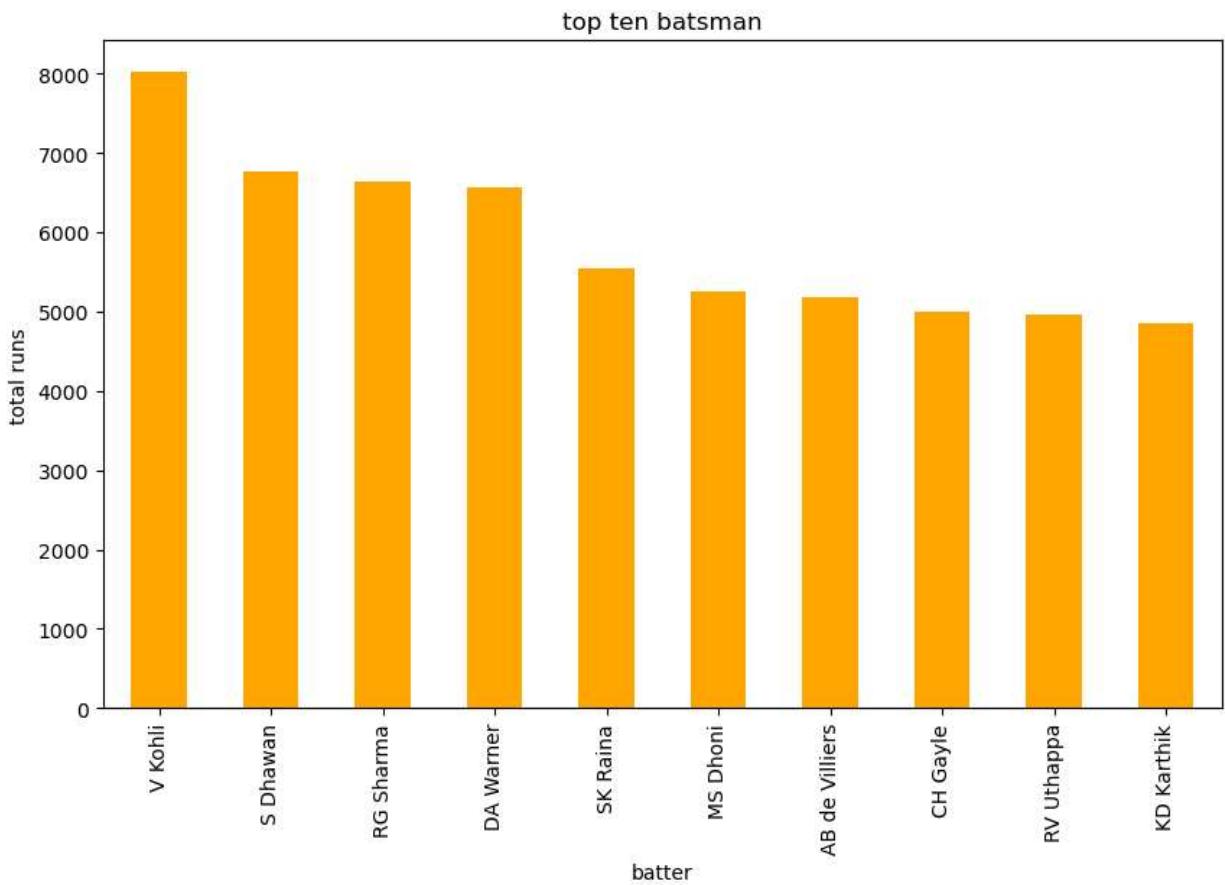
```
with pd.option_context('mode.use_inf_as_na', True):
```



Runs Scored by Each Batsman

```
In [36]: batsman_runs=df.groupby('batter')['batsman_runs'].sum().sort_values(ascending=False)
plt.figure(figsize=(10,6))
batsman_runs.head(10).plot(kind='bar',color='orange')
plt.title('top ten batsman')
plt.ylabel("total runs")
```

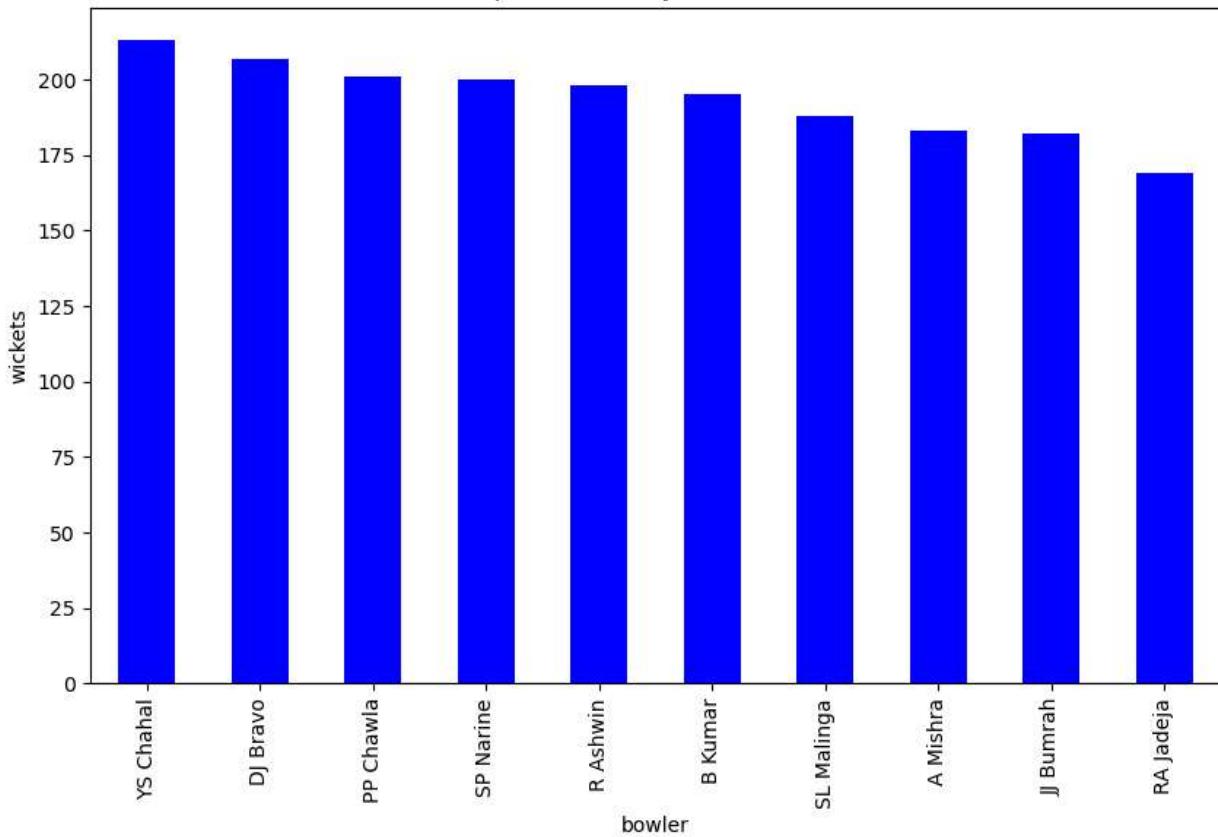
Out[36]: Text(0, 0.5, 'total runs')



Wickets by Bowler

```
In [37]: wickets=df[df['is_wicket']==1].groupby('bowler')['is_wicket'].count().sort_values(ascending=True)
plt.figure(figsize=(10,6))
wickets.head(10).plot(kind='bar',color='blue')
plt.title('top 10 bowlers by wickets taken')
plt.ylabel('wickets')
plt.show()
```

top 10 bowlers by wickets taken



Visualization of Extras and Dismissals

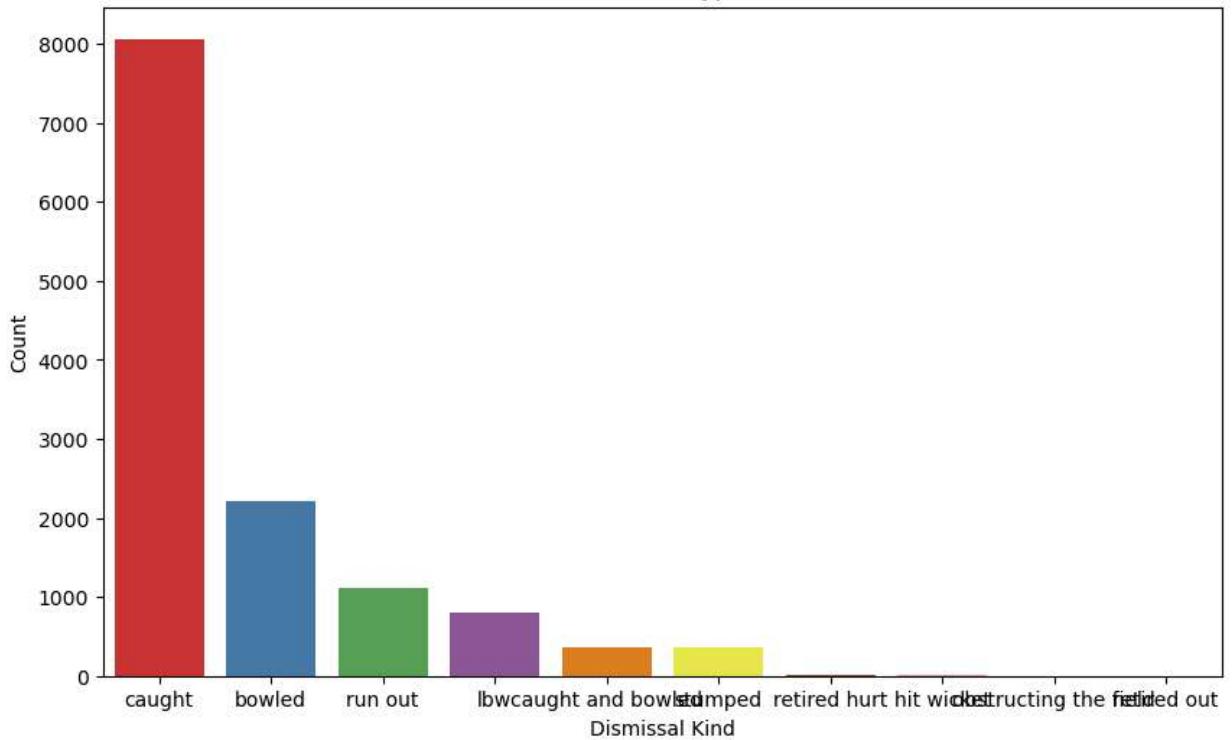
Extras types distribution

```
plt.figure(figsize=(10,6)) sns.countplot(x='extras_type', data=df, palette='Set2',
order=df['extras_type'].value_counts().index) plt.title('Distribution of Extras by Type')
plt.xlabel('Extras Type') plt.ylabel('Count') plt.show()
```

dismissal types distribution

```
In [39]: # Dismissal types distribution
plt.figure(figsize=(10,6))
sns.countplot(x='dismissal_kind', data=df, palette='Set1', order=df['dismissal_kind'].
plt.title('Dismissal Types')
plt.xlabel('Dismissal Kind')
plt.ylabel('Count')
plt.show()
```

Dismissal Types



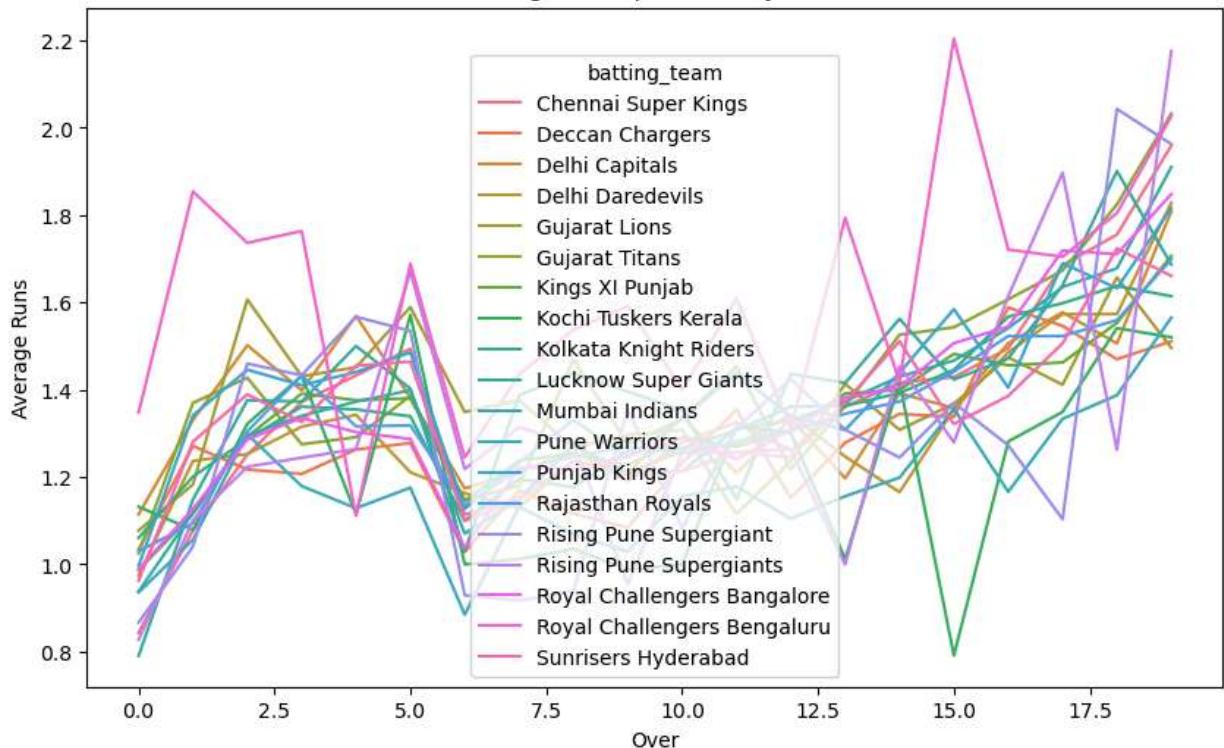
average runs per over by team

```
In [40]: # Average runs per over by team
team_runs_per_over = df.groupby(['batting_team', 'over'])['total_runs'].mean().reset_index()
plt.figure(figsize=(10,6))
sns.lineplot(data=team_runs_per_over, x='over', y='total_runs', hue='batting_team')
plt.title('Average Runs per Over by Team')
plt.xlabel('Over')
plt.ylabel('Average Runs')
plt.show()
import warnings
```

```
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
    with pd.option_context('mode.use_inf_as_na', True):
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119: FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
    with pd.option_context('mode.use_inf_as_na', True):
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_group in a future version of pandas. Pass `(name,)` instead of `name` to silence this warning.
    data_subset = grouped_data.get_group(pd_key)
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_group in a future version of pandas. Pass `(name,)` instead of `name` to silence this warning.
    data_subset = grouped_data.get_group(pd_key)
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_group in a future version of pandas. Pass `(name,)` instead of `name` to silence this warning.
    data_subset = grouped_data.get_group(pd_key)
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_group in a future version of pandas. Pass `(name,)` instead of `name` to silence this warning.
    data_subset = grouped_data.get_group(pd_key)
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_group in a future version of pandas. Pass `(name,)` instead of `name` to silence this warning.
    data_subset = grouped_data.get_group(pd_key)
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_group in a future version of pandas. Pass `(name,)` instead of `name` to silence this warning.
    data_subset = grouped_data.get_group(pd_key)
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_group in a future version of pandas. Pass `(name,)` instead of `name` to silence this warning.
```

```
up in a future version of pandas. Pass `(name,)` instead of `name` to silence this warning.  
    data_subset = grouped_data.get_group(pd_key)  
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When  
grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_gro  
up in a future version of pandas. Pass `(name,)` instead of `name` to silence this wa  
rning.  
    data_subset = grouped_data.get_group(pd_key)  
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When  
grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_gro  
up in a future version of pandas. Pass `(name,)` instead of `name` to silence this wa  
rning.  
    data_subset = grouped_data.get_group(pd_key)  
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When  
grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_gro  
up in a future version of pandas. Pass `(name,)` instead of `name` to silence this wa  
rning.  
    data_subset = grouped_data.get_group(pd_key)  
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When  
grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_gro  
up in a future version of pandas. Pass `(name,)` instead of `name` to silence this wa  
rning.  
    data_subset = grouped_data.get_group(pd_key)  
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When  
grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_gro  
up in a future version of pandas. Pass `(name,)` instead of `name` to silence this wa  
rning.  
    data_subset = grouped_data.get_group(pd_key)  
d:\Users\hi\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1075: FutureWarning: When  
grouping with a length-1 list-like, you will need to pass a length-1 tuple to get_gro  
up in a future version of pandas. Pass `(name,)` instead of `name` to silence this wa  
rning.  
    data_subset = grouped_data.get_group(pd_key)
```

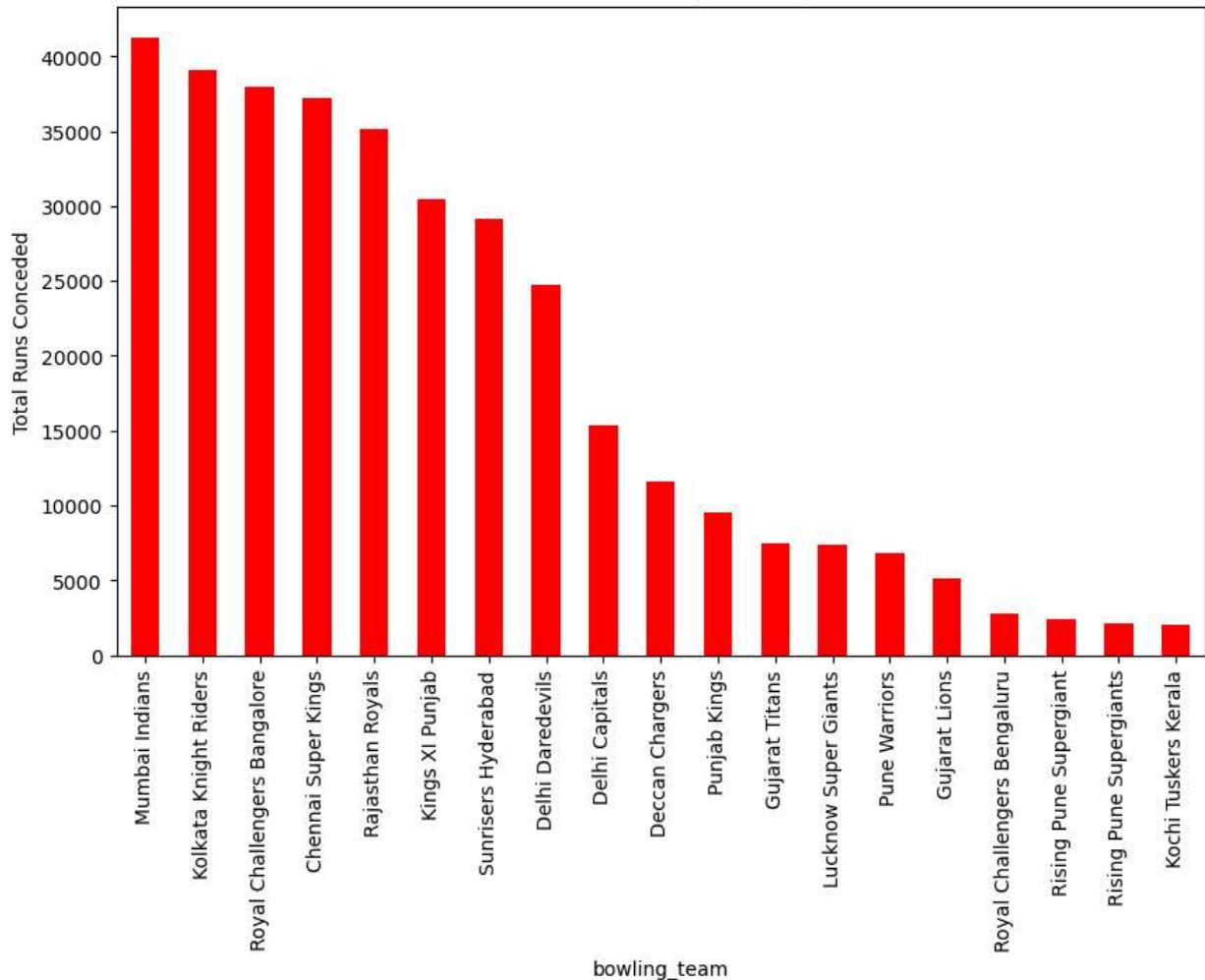
Average Runs per Over by Team



Performance of Bowling Teams (Runs Conceded)

```
In [41]: # Total runs conceded by bowling team
team_bowling_conceded = df.groupby('bowling_team')['total_runs'].sum().sort_values(ascending=True)
plt.figure(figsize=(10,6))
team_bowling_conceded.plot(kind='bar', color='red')
plt.title('Total Runs Conceded by Bowling Teams')
plt.ylabel('Total Runs Conceded')
plt.show()
```

Total Runs Conceded by Bowling Teams



Save the Visualizations

```
In [42]: plt.savefig('visualization_name.png')
```

```
<Figure size 640x480 with 0 Axes>
```

```
In [ ]:
```