

```
In [1]: import pandas as pd
```

```
In [2]: df = pd.read_csv("IPL IMB381IPL2013.csv")
```

```
In [3]: df
```

Out[3]:

	SI.NO.	PLAYER NAME	AGE	COUNTRY	TEAM	PLAYING ROLE	T- RUNS	T- WKTS	ODI- RUNS- S	ODI- SR-B	...
0	1	Abdulla, YA	2	SA	KXIP	Allrounder	0	0	0	0.00	...
1	2	Abdur Razzak	2	BAN	RCB	Bowler	214	18	657	71.41	...
2	3	Agarkar, AB	2	IND	KKR	Bowler	571	58	1269	80.62	...
3	4	Ashwin, R	1	IND	CSK	Bowler	284	31	241	84.56	...
4	5	Badrinath, S	2	IND	CSK	Batsman	63	0	79	45.93	...
...
125	126	Yadav, AS	2	IND	DC	Batsman	0	0	0	0.00	...
126	127	Younis Khan	2	PAK	RR	Batsman	6398	7	6814	75.78	...
127	128	Yuvraj Singh	2	IND	KXIP+	Batsman	1775	9	8051	87.58	...
128	129	Zaheer Khan	2	IND	MI+	Bowler	1114	288	790	73.55	...
129	130	Zoysa, DNT	2	SL	DC	Bowler	288	64	343	95.81	...

130 rows × 26 columns



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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 130 entries, 0 to 129
Data columns (total 26 columns):
#   Column                Non-Null Count  Dtype
---  -
0   SL.NO.                130 non-null    int64
1   PLAYER NAME          130 non-null    object
2   AGE                  130 non-null    int64
3   COUNTRY              130 non-null    object
4   TEAM                 130 non-null    object
5   PLAYING ROLE         130 non-null    object
6   T-RUNS               130 non-null    int64
7   T-WKTS              130 non-null    int64
8   ODI-RUNS-S           130 non-null    int64
9   ODI-SR-B             130 non-null    float64
10  ODI-WKTS             130 non-null    int64
11  ODI-SR-BL            130 non-null    float64
12  CAPTAINCY EXP        130 non-null    int64
13  RUNS-S               130 non-null    int64
14  HS                   130 non-null    int64
15  AVE                  130 non-null    float64
16  SR-B                 130 non-null    float64
17  SIXERS               130 non-null    int64
18  RUNS-C               130 non-null    int64
19  WKTS                 130 non-null    int64
20  AVE-BL               130 non-null    float64
21  ECON                 130 non-null    float64
22  SR-BL                130 non-null    float64
23  AUCTION YEAR         130 non-null    int64
24  BASE PRICE           130 non-null    int64
25  SOLD PRICE           130 non-null    int64
dtypes: float64(7), int64(15), object(4)
memory usage: 26.5+ KB
```

1)No null value in dataset


2)No need to change data types

```
In [5]: df.describe()
```

Out[5]:

	SI.NO.	AGE	T-RUNS	T-WKTS	ODI-RUNS-S	ODI-SR-B	ODI-WI
count	130.000000	130.000000	130.000000	130.000000	130.000000	130.000000	130.000
mean	65.500000	2.092308	2166.715385	66.530769	2508.738462	71.164385	76.076
std	37.671829	0.576627	3305.646757	142.676855	3582.205625	25.898440	111.205
min	1.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000
25%	33.250000	2.000000	25.500000	0.000000	73.250000	65.650000	0.000
50%	65.500000	2.000000	542.500000	7.000000	835.000000	78.225000	18.500
75%	97.750000	2.000000	3002.250000	47.500000	3523.500000	86.790000	106.000
max	130.000000	3.000000	15470.000000	800.000000	18426.000000	116.660000	534.000

8 rows × 22 columns

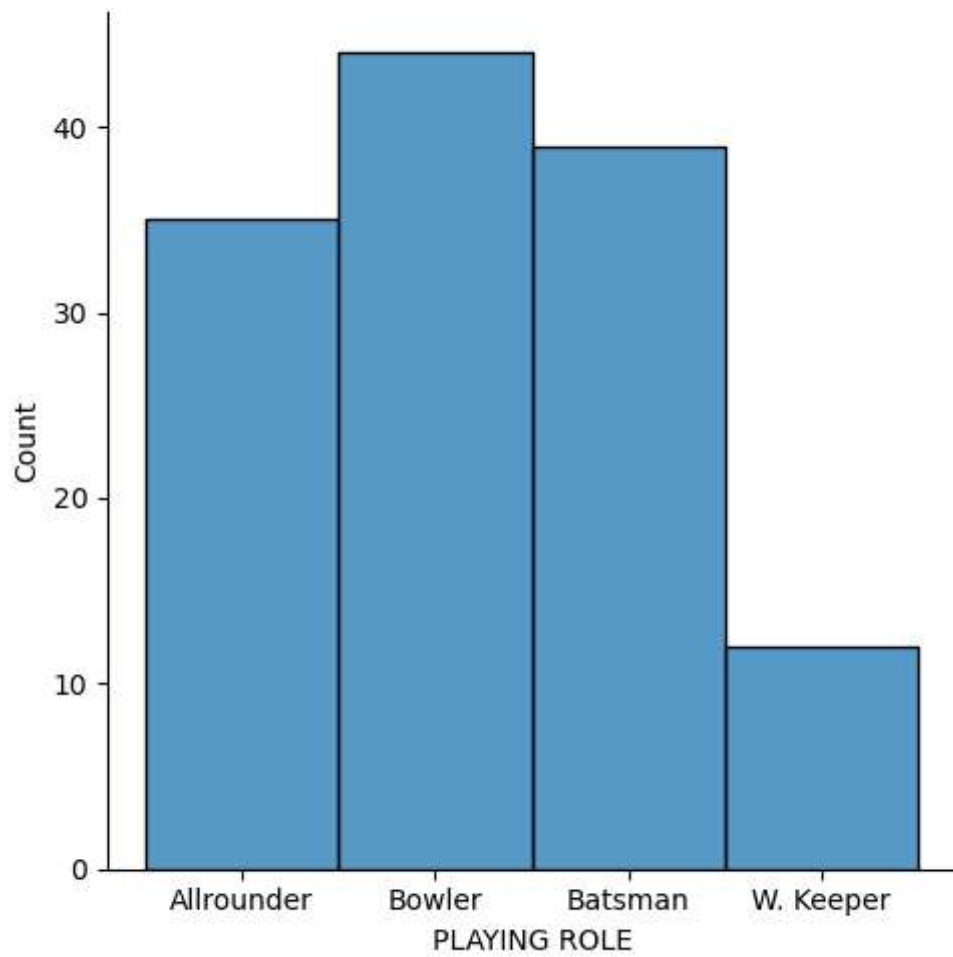


Statistical calculation

```
In [16]: import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

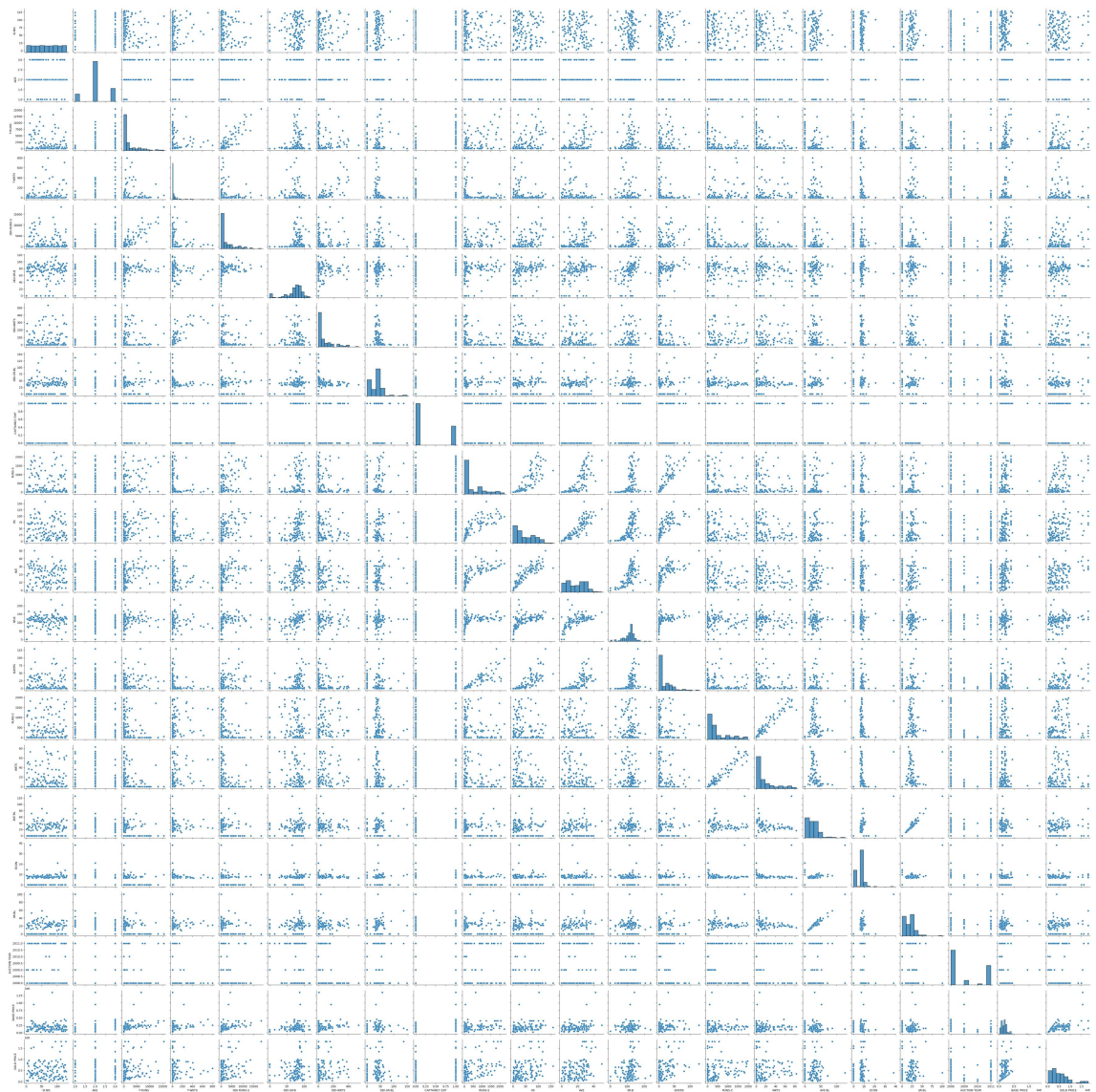
```
In [17]: sns.displot(df["PLAYING ROLE"])
```

```
Out[17]: <seaborn.axisgrid.FacetGrid at 0x21e48f0cb50>
```



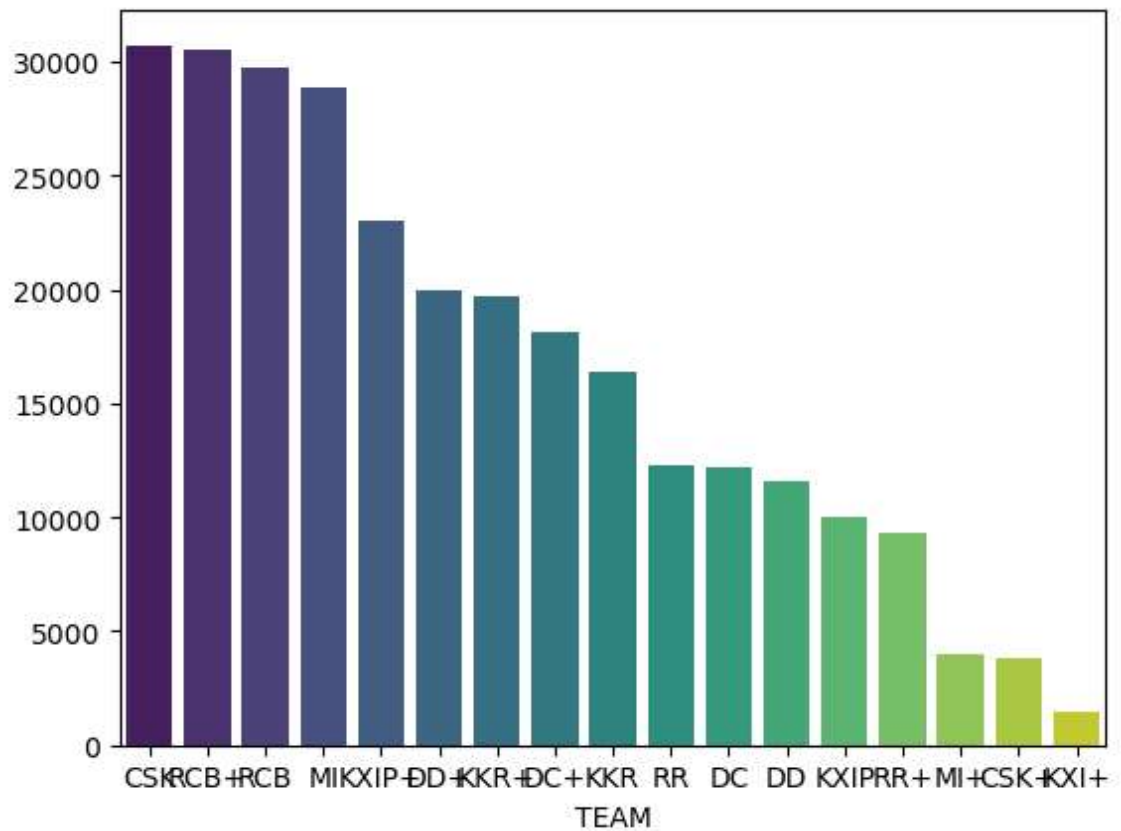
```
In [32]: sns.pairplot(df)
```

```
Out[32]: <seaborn.axisgrid.PairGrid at 0x21e4e873210>
```



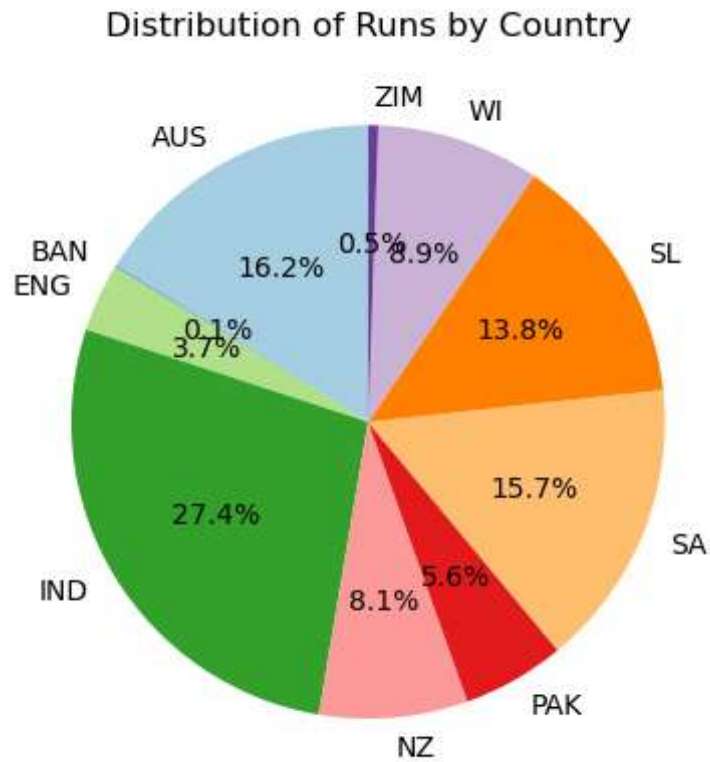
```
In [36]: total_runs_by_team = df.groupby('TEAM')['T-RUNS'].sum().sort_values(ascending=True)
sns.barplot(x=total_runs_by_team.index, y=total_runs_by_team.values, palette='magma')
```

Out[36]: <Axes: xlabel='TEAM'>



```
In [43]: ▶ runs_by_country = df.groupby('COUNTRY')['T-RUNS'].sum()
plt.pie(runs_by_country, labels=runs_by_country.index, autopct='%1.1f%%', s
plt.title('Distribution of Runs by Country')
```

Out[43]: Text(0.5, 1.0, 'Distribution of Runs by Country')



In []: ▶