

```
[ ]: # DATASET IPL 2022 DATA ANALYSIS USING PYTHON

In [2]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import matplotlib inline
import warnings
warnings.filterwarnings('ignore')

In [3]: ipl = pd.read_csv('C:\\Users\\Majid Aslam\\Downloads\\IPL DATA SET.csv')
ipl
```

	Unnamed: 0	Player	Base Price	TYPE	COST IN ₹ (CR.)	Cost IN \$ (000)	2021 Squad	Team
0	0	Rashid Khan	Draft Pick	BOWLER	15.0	1950.0	SRH	Gujarat Titans
1	1	Hardik Pandya	Draft Pick	ALL-ROUNDER	15.0	1950.0	MI	Gujarat Titans
2	2	Lockie Ferguson	2 Cr	BOWLER	10.0	1300.0	KKR	Gujarat Titans
3	3	Rahul Tewatia	40 Lakh	ALL-ROUNDER	9.0	1170.0	RR	Gujarat Titans
4	4	Shubman Gill	Draft Pick	BATTER	8.0	1040.0	KKR	Gujarat Titans
...
628	628	Sairaj Patil	20 Lakh	BATTER	NaN	NaN	NaN	Unsold
629	629	Monu Singh	20 Lakh	BOWLER	NaN	NaN	NaN	Unsold
630	630	Nivethan Radhakrishnan	20 Lakh	BOWLER	NaN	NaN	NaN	Unsold
631	631	Lance Morris	20 Lakh	BOWLER	NaN	NaN	NaN	Unsold
632	632	Aaron Hardie	20 Lakh	ALL-ROUNDER	NaN	NaN	NaN	Unsold

633 rows × 8 columns

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

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 633 entries, 0 to 632
Data columns (total 8 columns):
 #   Column              Non-Null Count  Dtype
---  --
 0   Unnamed: 0          633 non-null   int64
 1   Player              633 non-null   object
 2   Base Price          633 non-null   object
 3   TYPE                633 non-null   object
 4   COST IN ₹ (CR.)     237 non-null   float64
 5   Cost IN $ (000)     237 non-null   float64
 6   2021 Squad          199 non-null   object
 7   Team                633 non-null   object
dtypes: float64(2), int64(1), object(5)
memory usage: 39.7+ KB

In [5]: ipl.shape

Out[5]: (633, 8)

In [7]: ipl.columns

Out[7]: Index(['Unnamed: 0', 'Player', 'Base Price', 'TYPE', 'COST IN ₹ (CR.)',
          'Cost IN $ (000)', '2021 Squad', 'Team'],
          dtype='object')
```

```
In [8]: ipl.drop('Unnamed: 0', axis = 1, inplace = True)
```

	Player	Base Price	TYPE	COST IN ₹ (CR.)	Cost IN \$ (000)	2021 Squad	Team
0	Rashid Khan	Draft Pick	BOWLER	15.0	1950.0	SRH	Gujarat Titans
1	Hardik Pandya	Draft Pick	ALL-ROUNDER	15.0	1950.0	MI	Gujarat Titans
2	Lockie Ferguson	2 Cr	BOWLER	10.0	1300.0	KKR	Gujarat Titans
3	Rahul Tewatia	40 Lakh	ALL-ROUNDER	9.0	1170.0	RR	Gujarat Titans
4	Shubman Gill	Draft Pick	BATTER	8.0	1040.0	KKR	Gujarat Titans
...
628	Sairaj Patil	20 Lakh	BATTER	NaN	NaN	NaN	Unsold
629	Monu Singh	20 Lakh	BOWLER	NaN	NaN	NaN	Unsold
630	Nivethan Radhakrishnan	20 Lakh	BOWLER	NaN	NaN	NaN	Unsold
631	Lance Morris	20 Lakh	BOWLER	NaN	NaN	NaN	Unsold
632	Aaron Hardie	20 Lakh	ALL-ROUNDER	NaN	NaN	NaN	Unsold

633 rows × 7 columns

```
In [10]: ipl.isnull().sum()

Out[10]: Player      0
Base Price    0
TYPE          0
COST IN ₹ (CR.)    396
Cost IN $ (000)    396
2021 Squad      434
Team           0
dtype: int64

In [11]: ipl[ipl['Cost IN $ (000)'].isnull()]

Out[11]:
```

	Player	Base Price	TYPE	COST IN ₹ (CR.)	Cost IN \$ (000)	2021 Squad	Team
237	Suresh Raina	2 Cr	BATTER	NaN	NaN	CSK	Unsold
238	Steve Smith	2 Cr	BATTER	NaN	NaN	DC	Unsold
239	Shakib Al Hasan	2 Cr	ALL-ROUNDER	NaN	NaN	KKR	Unsold
240	Amit Mishra	1.5 Cr	BOWLER	NaN	NaN	DC	Unsold
241	Adil Rashid	2 Cr	BOWLER	NaN	NaN	PBKS	Unsold
...
628	Sairaj Patil	20 Lakh	BATTER	NaN	NaN	NaN	Unsold
629	Monu Singh	20 Lakh	BOWLER	NaN	NaN	NaN	Unsold
630	Nivethan Radhakrishnan	20 Lakh	BOWLER	NaN	NaN	NaN	Unsold
631	Lance Morris	20 Lakh	BOWLER	NaN	NaN	NaN	Unsold
632	Aaron Hardie	20 Lakh	ALL-ROUNDER	NaN	NaN	NaN	Unsold

396 rows × 7 columns

```
In [15]: ipl['COST IN ₹ (CR.)'] = ipl['COST IN ₹ (CR.)'].fillna(0)
ipl['Cost IN $ (000)'] = ipl['Cost IN $ (000)'].fillna(0)
```

	Player	Base Price	TYPE	COST IN ₹ (CR.)	Cost IN \$ (000)	2021 Squad	Team
0	Rashid Khan	Draft Pick	BOWLER	15.0	1950.0	SRH	Gujarat Titans
1	Hardik Pandya	Draft Pick	ALL-ROUNDER	15.0	1950.0	MI	Gujarat Titans
2	Lockie Ferguson	2 Cr	BOWLER	10.0	1300.0	KKR	Gujarat Titans
3	Rahul Tewatia	40 Lakh	ALL-ROUNDER	9.0	1170.0	RR	Gujarat Titans
4	Shubman Gill	Draft Pick	BATTER	8.0	1040.0	KKR	Gujarat Titans
...
628	Sairaj Patil	20 Lakh	BATTER	0.0	0.0	NaN	Unsold
629	Monu Singh	20 Lakh	BOWLER	0.0	0.0	NaN	Unsold
630	Nivethan Radhakrishnan	20 Lakh	BOWLER	0.0	0.0	NaN	Unsold
631	Lance Morris	20 Lakh	BOWLER	0.0	0.0	NaN	Unsold
632	Aaron Hardie	20 Lakh	ALL-ROUNDER	0.0	0.0	NaN	Unsold

633 rows × 7 columns

	Player	Base Price	TYPE	COST IN ₹ (CR.)	Cost IN \$ (000)	2021 Squad	Team
6	Yash Dayal	20 Lakh	BOWLER	3.2	416.0	NaN	Gujarat Titans
9	Abhinav Sadarangani	2 Cr	BATTER	2.6	338.0	NaN	Gujarat Titans
10	Matthew Wade	2 Cr	WICKETKEEPER	2.4	312.0	NaN	Gujarat Titans
11	Alzani Joseph	75 Lakh	BOWLER	2.4	312.0	NaN	Gujarat Titans
17	Varun Aaron	50 Lakh	BOWLER	0.5	65.0	NaN	Gujarat Titans
...
628	Sairaj Patil	20 Lakh	BATTER	0.0	0.0	NaN	Unsold
629	Monu Singh	20 Lakh	BOWLER	0.0	0.0	NaN	Unsold
630	Nivethan Radhakrishnan	20 Lakh	BOWLER	0.0	0.0	NaN	Unsold
631	Lance Morris	20 Lakh	BOWLER	0.0	0.0	NaN	Unsold
632	Aaron Hardie	20 Lakh	ALL-ROUNDER	0.0	0.0	NaN	Unsold

434 rows × 7 columns

```
In [18]: ipl['2021 Squad'] = ipl['2021 Squad'].fillna('not participated')
```

	Player	Base Price	TYPE	COST IN ₹ (CR.)	Cost IN \$ (000)	2021 Squad	Team
0	Rashid Khan	Draft Pick	BOWLER	15.0	1950.0	SRH	Gujarat Titans
1	Hardik Pandya	Draft Pick	ALL-ROUNDER	15.0	1950.0	MI	Gujarat Titans
2	Lockie Ferguson	2 Cr	BOWLER	10.0	1300.0	KKR	Gujarat Titans
3	Rahul Tewatia	40 Lakh	ALL-ROUNDER	9.0	1170.0	RR	Gujarat Titans
4	Shubman Gill	Draft Pick	BATTER	8.0	1040.0	KKR	Gujarat Titans
...
628	Sairaj Patil	20 Lakh	BATTER	0.0	0.0	not participated	Unsold
629	Monu Singh	20 Lakh	BOWLER	0.0	0.0	not participated	Unsold
630	Nivethan Radhakrishnan	20 Lakh	BOWLER	0.0	0.0	not participated	Unsold
631	Lance Morris	20 Lakh	BOWLER	0.0	0.0	not participated	Unsold
632	Aaron Hardie	20 Lakh	ALL-ROUNDER	0.0	0.0	not participated	Unsold

633 rows × 7 columns

```
In [20]: ipl.isnull().sum()

Out[20]: Player      0
Base Price    0
TYPE          0
COST IN ₹ (CR.)    0
Cost IN $ (000)    0
2021 Squad      0
Team           0
dtype: int64

In [25]: teams = ipl[ipl['COST IN ₹ (CR.)']>0]['Team'].unique()
teams

Out[25]: array(['Gujarat Titans', 'Chennai Super Kings', 'Delhi Capitals',
                'Kolkaata Knight Riders', 'Punjab Kings', 'Lucknow Super Giants',
                'Mumbai Indians', 'Royal Challengers Bangalore',
                'Rajasthan Royals', 'Sunrisers Hyderabad'], dtype=object)
```

	Player	Base Price	TYPE	COST IN ₹ (CR.)	Cost IN \$ (000)	2021 Squad	Team	status
0	Rashid Khan	Draft Pick	BOWLER	15.0	1950.0	SRH	Gujarat Titans	sold
1	Hardik Pandya	Draft Pick	ALL-ROUNDER	15.0	1950.0	MI	Gujarat Titans	sold
2	Lockie Ferguson	2 Cr	BOWLER	10.0	1300.0	KKR	Gujarat Titans	sold
3	Rahul Tewatia	40 Lakh	ALL-ROUNDER	9.0	1170.0	RR	Gujarat Titans	sold
4	Shubman Gill	Draft Pick	BATTER	8.0	1040.0	KKR	Gujarat Titans	sold
...
628	Sairaj Patil	20 Lakh	BATTER	0.0	0.0	not participated	Unsold	Unsold
629	Monu Singh	20 Lakh	BOWLER	0.0	0.0	not participated	Unsold	Unsold
630	Nivethan Radhakrishnan	20 Lakh	BOWLER	0.0	0.0	not participated	Unsold	Unsold
631	Lance Morris	20 Lakh	BOWLER	0.0	0.0	not participated	Unsold	Unsold
632	Aaron Hardie	20 Lakh	ALL-ROUNDER	0.0	0.0	not participated	Unsold	Unsold

633 rows × 8 columns

	Player	Base Price	TYPE	COST IN ₹ (CR.)	Cost IN \$ (000)	2021 Squad	Team	status
62	Lalit Yadav	20 Lakh	ALL-ROUNDER	0.65	84.5	DC	Delhi Capitals	sold
440	Amit Mishra	1.5 Cr	BOWLER	0.00	0.0	DC	Unsold	Unsold
299	Amit Mishra	20 Lakh	BOWLER	0.00	0.0	not participated	Unsold	Unsold
518	Shivam Sharma	20 Lakh	ALL-ROUNDER	0.00	0.0	not participated	Unsold	Unsold
530	Lalit Yadav	20 Lakh	BOWLER	0.00	0.0	not participated	Unsold	Unsold
537	Shubham Singh	20 Lakh	ALL-ROUNDER	0.00	0.0	not participated	Unsold	Unsold
544	Shubham Singh	20 Lakh	BOWLER	0.00	0.0	not participated	Unsold	Unsold
569	Shivam Sharma	20 Lakh	ALL-ROUNDER	0.00	0.0	not participated	Unsold	Unsold

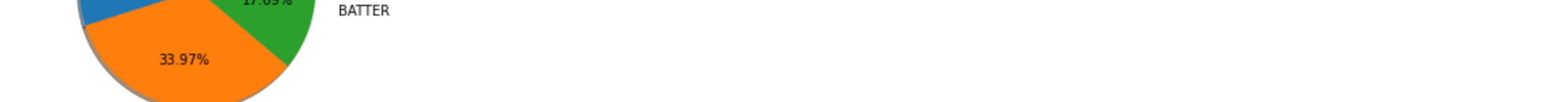
```
In [31]: # how many players participate in 2022 ipl audtion
ipl.shape[0]
```

```
Out[31]: 633
```

```
In [38]: # how many types of players in participated
types = ipl['TYPE'].value_counts()
types.reset_index()
```

	index	TYPE
0	ALL-ROUNDER	242
1	BOWLER	215
2	BATTER	112
3	WICKETKEEPER	64

```
In [42]: # pie chart
plt.pie(types.values, labels = types.index, labeldistance=1.2, autopct = '%1.2f%', shadow = True, startangle =60)
plt.title('Role of players participated', fontsize = 15)
plt.plot()
```



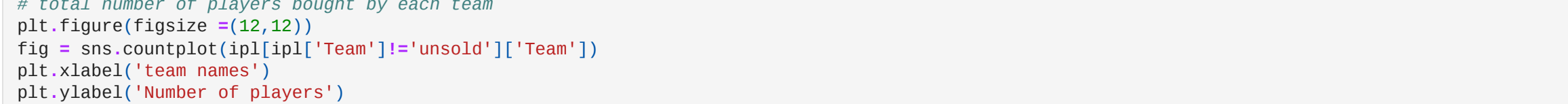
```
In [52]: #how many players sold and unsold using bar graph
plt.figure(figsize =(10,6))
fig = sns.countplot(ipl['status'],palette = ['pink','yellow'])
plt.xlabel('sold or unsold')
plt.ylabel('Number of players')
plt.title('sold or unsold',fontsize =15)
plt.plot()
```



```
In [54]: ipl.groupby('status')['Player'].count()
```

status	396
unsold	237
sold	237
Name: Player, dtype: int64	

```
In [61]: # total number of players bought by each team
plt.figure(figsize =(12,12))
fig = sns.countplot(ipl[ipl['Team']!='unsold']['Team'])
plt.xlabel('team names')
plt.ylabel('number of players')
plt.title('players bought by each team',fontsize =10)
plt.xticks(rotation=60)
plt.plot()
```



```
In [62]: # first ball of ipl match
ipl.iloc[0]
```

Player	Rashid Khan
Base Price	Draft Pick
TYPE	BOWLER
COST IN ₹ (CR.)	15.0
Cost IN \$ (000)	1950.0
2021 Squad	SRH
Team	Gujarat Titans
status	sold
Name: 0, dtype: object	

```
In [63]: ipl.groupby(['Player', 'TYPE']).count().droplevel(level=1).index.value_counts()
```

Shubham Singh	2
Imran Tahir	2
Dev Lakra	1
Dushmanta Chameera	1
...	1
Jason Sangha	1
B. P. Sharath	1
Dhruv Jurel	1
Brandon King	1
Ansh Patel	1
Name: Player, Length: 629, dtype: int64	

```
In [ ]:
```

```
In [ ]:
```