IPL Auction Price Data

- The Indian Premier League (IPL) is a professional League for Twenty20(T20) cricket championship that was started in 2008 in India.
- The first IPL auction was held in 2008 for ownership of the team for 10 years with a base price of USD 50 million.
- The franchise acquires players, through an English auction that is conducted every year. However, there are several rules imposed by the IPL for example there is a maximum cap on the money a franchise can spend on buying players.
- The performance of the players could be measured through several metrics. Although the IPL follows the T20 format of the game, it is possible that the performance of the players in the other formats of the game such as Test matches, One-Day matches can influence player pricing. Few players have excellent records in the Test Matches, but their records in T20 matches are not very impressive.
- The Dataset consists of the performance of 130 Players measured through various performance metrics such as Batting strike rate, economy rate, wicket in Tets matches, Bowling average..etc in the year 2013

Dataset Link:

https://raw.githubusercontent.com/sitmbadept/sitmbadept.github.io/main/BDTM/R/IPL.csv

```
In [29]:   #1. Import pandas library in Python

           # Download & Install pandas library if you have not installed
           # !pip install pandas
```

```
In [30]:   import pandas as pd
```

```
In [31]:   ipl = pd.read_csv("C:/dataset/IPL.csv")
```

```
In [32]:   # View DataFrame Structure
           ipl.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 132 entries, 0 to 131
Data columns (total 26 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Sl.NO.         132 non-null    int64
 1   PLAYER NAME    132 non-null    object
 2   AGE            130 non-null    float64
 3   COUNTRY        132 non-null    object
 4   TEAM           132 non-null    object
 5   PLAYING ROLE   132 non-null    object
 6   T-RUNS         130 non-null    float64
 7   T-WKTS         130 non-null    float64
 8   ODI-RUNS-S     130 non-null    float64
```

```
 9   ODI-SR-B       130 non-null    float64
10   ODI-WKTS       130 non-null    float64
11   ODI-SR-BL      130 non-null    float64
12   CAPTAINCY EXP  130 non-null    float64
13   RUNS-S         130 non-null    float64
14   HS             130 non-null    float64
15   AVE            130 non-null    float64
16   SR-B           130 non-null    float64
17   SIXERS         130 non-null    float64
18   RUNS-C         130 non-null    float64
19   WKTS           130 non-null    float64
20   AVE-BL         130 non-null    float64
21   ECON           130 non-null    float64
22   SR-BL          130 non-null    float64
23   AUCTION YEAR   131 non-null    float64
24   BASE PRICE     131 non-null    float64
25   SOLD PRICE     131 non-null    float64
dtypes: float64(21), int64(1), object(4)
memory usage: 26.9+ KB
```

In [33]:
```python
# View No of Rows & Columns
ipl.shape

# 132 rows & 26 columns
```

Out[33]: (132, 26)

In [34]:
```python
# Display First & Last few records
ipl.head() #Default is 5
```

Out[34]:

	Sl.NO.	PLAYER NAME	AGE	COUNTRY	TEAM	PLAYING ROLE	T-RUNS	T-WKTS	ODI-RUNS-S	ODI-SR-B	...	SR-B	SIXER
0	1	Abdulla, YA	2.0	SA	KXIP	Allrounder	0.0	0.0	0.0	0.00	...	0.00	0
1	2	Abdur Razzak	2.0	BAN	RCB	Bowler	214.0	18.0	657.0	71.41	...	0.00	0
2	3	Agarkar, AB	2.0	IND	KKR	Bowler	571.0	58.0	1269.0	80.62	...	121.01	5
3	4	Ashwin, R	1.0	IND	CSK	Bowler	284.0	31.0	241.0	84.56	...	76.32	0
4	5	Badrinath, S	2.0	IND	CSK	Batsman	63.0	0.0	79.0	45.93	...	120.71	28

5 rows × 26 columns

In [35]:
```python
ipl.head(2) #Custom number of records
```

Out[35]:

	SI.NO.	PLAYER NAME	AGE	COUNTRY	TEAM	PLAYING ROLE	T-RUNS	T-WKTS	ODI-RUNS-S	ODI-SR-B	...	SR-B	SIXERS	F
0	1	Abdulla, YA	2.0	SA	KXIP	Allrounder	0.0	0.0	0.0	0.00	...	0.0	0.0	
1	2	Abdur Razzak	2.0	BAN	RCB	Bowler	214.0	18.0	657.0	71.41	...	0.0	0.0	

2 rows × 26 columns

```
In [36]:    ipl.tail() #Default is 5
```

Out[36]:

	SI.NO.	PLAYER NAME	AGE	COUNTRY	TEAM	PLAYING ROLE	T-RUNS	T-WKTS	ODI-RUNS-S	ODI-SR-B	...	SR-B	SIXE
127	128	Yuvraj Singh	2.0	IND	KXIP+	Batsman	1775.0	9.0	8051.0	87.58	...	131.88	6
128	129	Zaheer Khan	2.0	IND	MI+	Bowler	1114.0	288.0	790.0	73.55	...	91.67	
129	130	Zoysa, DNT	2.0	SL	DC	Bowler	288.0	64.0	343.0	95.81	...	122.22	
130	131	Dummy	NaN	Dummy	Dummy	Dummy	NaN	NaN	NaN	NaN	...	NaN	N
131	132	Dummy	NaN	Dummy	Dummy	Dummy	NaN	NaN	NaN	NaN	...	NaN	N

5 rows × 26 columns

```
In [37]:    # Frequency count of unique value present in Columns
            ipl['COUNTRY'].value_counts()
```

```
Out[37]:    IND       53
            AUS       22
            SA        16
            SL        12
            PAK        9
            NZ         7
            WI         6
            ENG        3
            Dummy      2
            BAN        1
            ZIM        1
            Name: COUNTRY, dtype: int64
```

```
In [38]:    # Cross tab for multiple columns
            pd.crosstab(ipl['COUNTRY'], ipl['PLAYING ROLE'])
```

Out[38]:

PLAYING ROLE	Allrounder	Batsman	Bowler	Dummy	W. Keeper
COUNTRY					
AUS	6	9	6	0	1
BAN	0	0	1	0	0
Dummy	0	0	0	2	0
ENG	2	1	0	0	0
IND	7	18	23	0	5
NZ	4	2	0	0	1
PAK	2	2	4	0	1
SA	7	3	4	0	2
SL	4	2	5	0	1
WI	3	2	1	0	0
ZIM	0	0	0	0	1

In [39]:
```python
ipl=ipl[ipl['AGE'].isnull()==False]
```

In [40]:
```python
# Converting Data From one data type to another
ipl['AUCTION YEAR'] = ipl['AUCTION YEAR'].astype('int')
```

In [41]:
```python
ipl[['COUNTRY', 'TEAM', 'PLAYING ROLE','AUCTION YEAR', 'BASE PRICE', 'SOLD PRICE']].sor
```

Out[41]:

	COUNTRY	TEAM	PLAYING ROLE	AUCTION YEAR	BASE PRICE	SOLD PRICE
73	AUS	RCB	Allrounder	2010	20000.0	20000.0
46	IND	RR+	Bowler	2009	20000.0	24000.0
0	SA	KXIP	Allrounder	2009	50000.0	50000.0
1	BAN	RCB	Bowler	2008	50000.0	50000.0
118	SA	RCB+	Allrounder	2011	50000.0	50000.0
...	...	...	...	...	...	...
113	IND	MI+	Batsman	2011	100000.0	1600000.0
111	IND	MI	Batsman	2011	400000.0	1800000.0
50	IND	RCB	Batsman	2011	150000.0	1800000.0
93	IND	DD	Batsman	2011	400000.0	1800000.0
127	IND	KXIP+	Batsman	2011	400000.0	1800000.0

130 rows × 6 columns

```python
ipl[['COUNTRY', 'TEAM', 'PLAYING ROLE','AUCTION YEAR', 'BASE PRICE', 'SOLD PRICE']]\
    .sort_values(['SOLD PRICE'], ascending=False)
```

Out[42]:

	COUNTRY	TEAM	PLAYING ROLE	AUCTION YEAR	BASE PRICE	SOLD PRICE
93	IND	DD	Batsman	2011	400000.0	1800000.0
127	IND	KXIP+	Batsman	2011	400000.0	1800000.0
50	IND	RCB	Batsman	2011	150000.0	1800000.0
111	IND	MI	Batsman	2011	400000.0	1800000.0
113	IND	MI+	Batsman	2011	100000.0	1600000.0
...	...	...	...	...	...	...
34	AUS	KKR+	Allrounder	2011	50000.0	50000.0
5	AUS	CSK	Batsman	2009	50000.0	50000.0
0	SA	KXIP	Allrounder	2009	50000.0	50000.0
46	IND	RR+	Bowler	2009	20000.0	24000.0
73	AUS	RCB	Allrounder	2010	20000.0	20000.0

130 rows × 6 columns

In [43]:

```python
# Creating new columns
ipl['AUCTION_BY'] = "BCCI" #"IPL board"
```

In [44]:

```python
## Grouping & Aggregation(Summary)

ipl['BASE PRICE'].sum() #2.49 Crore
```

Out[44]: 24990000.0

In [45]:

```python
# Base Price summary by Team
ipl.groupby(['TEAM']).agg({"BASE PRICE":"sum"})
```

Out[45]:

	BASE PRICE
TEAM	
CSK	3325000.0
CSK+	750000.0
DC	1075000.0
DC+	1825000.0
DD	1625000.0
DD+	1670000.0
KKR	1060000.0

	BASE PRICE
TEAM	
KKR+	1925000.0
KXI+	300000.0
KXIP	800000.0
KXIP+	1470000.0
MI	1450000.0
MI+	950000.0
RCB	1245000.0
RCB+	3300000.0
RR	1225000.0
RR+	995000.0

In [46]:
```python
# Base Price summary by Team
ipl.groupby(['TEAM']).agg({"BASE PRICE":["sum","mean", "max","min"]})
```

Out[46]:

	BASE PRICE			
	sum	**mean**	**max**	**min**
TEAM				
CSK	3325000.0	237500.000000	950000.0	50000.0
CSK+	750000.0	150000.000000	250000.0	50000.0
DC	1075000.0	153571.428571	250000.0	50000.0
DC+	1825000.0	182500.000000	300000.0	100000.0
DD	1625000.0	270833.333333	400000.0	150000.0
DD+	1670000.0	167000.000000	250000.0	50000.0
KKR	1060000.0	212000.000000	335000.0	125000.0
KKR+	1925000.0	160416.666667	300000.0	50000.0
KXI+	300000.0	300000.000000	300000.0	300000.0
KXIP	800000.0	160000.000000	225000.0	50000.0
KXIP+	1470000.0	210000.000000	400000.0	20000.0
MI	1450000.0	241666.666667	400000.0	150000.0
MI+	950000.0	158333.333333	300000.0	50000.0
RCB	1245000.0	138333.333333	250000.0	20000.0
RCB+	3300000.0	275000.000000	1350000.0	50000.0
RR	1225000.0	204166.666667	450000.0	100000.0
RR+	995000.0	110555.555556	250000.0	20000.0

```
In [47]:   # Base Price summary by Team
           ipl.groupby(['TEAM']).agg({"BASE PRICE":["sum","mean", "max","min"],
                                      "SOLD PRICE":["sum","mean", "max","min"]})
```

Out[47]:

	BASE PRICE				SOLD PRICE			
	sum	mean	max	min	sum	mean	max	min
TEAM								
CSK	3325000.0	237500.000000	950000.0	50000.0	8415000.0	601071.428571	1550000.0	50000.0
CSK+	750000.0	150000.000000	250000.0	50000.0	2390000.0	478000.000000	675000.0	290000.0
DC	1075000.0	153571.428571	250000.0	50000.0	2560000.0	365714.285714	750000.0	100000.0
DC+	1825000.0	182500.000000	300000.0	100000.0	5850000.0	585000.000000	1350000.0	100000.0
DD	1625000.0	270833.333333	400000.0	150000.0	4275000.0	712500.000000	1800000.0	225000.0
DD+	1670000.0	167000.000000	250000.0	50000.0	4980000.0	498000.000000	850000.0	80000.0
KKR	1060000.0	212000.000000	335000.0	125000.0	1450000.0	290000.000000	425000.0	125000.0
KKR+	1925000.0	160416.666667	300000.0	50000.0	5940000.0	495000.000000	950000.0	50000.0
KXI+	300000.0	300000.000000	300000.0	300000.0	900000.0	900000.000000	900000.0	900000.0
KXIP	800000.0	160000.000000	225000.0	50000.0	1225000.0	245000.000000	400000.0	50000.0
KXIP+	1470000.0	210000.000000	400000.0	20000.0	4745000.0	677857.142857	1800000.0	50000.0
MI	1450000.0	241666.666667	400000.0	150000.0	4675000.0	779166.666667	1800000.0	150000.0
MI+	950000.0	158333.333333	300000.0	50000.0	3650000.0	608333.333333	1600000.0	200000.0
RCB	1245000.0	138333.333333	250000.0	20000.0	3980000.0	442222.222222	1800000.0	20000.0
RCB+	3300000.0	275000.000000	1350000.0	50000.0	6705000.0	558750.000000	1550000.0	50000.0
RR	1225000.0	204166.666667	450000.0	100000.0	2525000.0	420833.333333	950000.0	100000.0
RR+	995000.0	110555.555556	250000.0	20000.0	3494000.0	388222.222222	950000.0	24000.0

```
In [48]:   # Renaming columns
           ipl.rename(columns={'PLAYER NAME':'PLAYER',
                               'AGE': 'AgeGroup',
                               'PLAYING ROLE': 'ROLE'},
                      inplace=True
                      )
```

```
In [49]:   ipl.drop(["ROLE", "PLAYER", "AgeGroup"],
                    inplace=True,
                    axis=1)
```

```
In [50]:   # Filtering Data based on Condition
           ipl[ipl['COUNTRY']=='SL']
```

Out[50]:

	SI.NO.	COUNTRY	TEAM	T-RUNS	T-WKTS	ODI-RUNS-S	ODI-SR-B	ODI-WKTS	ODI-SR-BL	CAPTAINCY EXP	...	SIXERS	R
16	17	SL	DD+	4722.0	32.0	6455.0	86.80	67.0	58.3	1.0	...	24.0	
21	22	SL	MI	249.0	97.0	239.0	60.96	187.0	34.7	0.0	...	0.0	
41	42	SL	MI	6973.0	98.0	13430.0	91.21	323.0	46.0	1.0	...	39.0	
42	43	SL	KXIP+	10440.0	6.0	10596.0	78.08	7.0	83.1	1.0	...	33.0	
56	57	SL	DD	556.0	25.0	1042.0	84.44	133.0	33.3	0.0	...	9.0	
57	58	SL	MI	275.0	101.0	327.0	73.81	185.0	31.1	0.0	...	4.0	1
59	60	SL	KKR+	1219.0	7.0	1447.0	82.59	42.0	43.0	0.0	...	12.0	
69	70	SL	CSK+	1261.0	800.0	674.0	77.56	534.0	35.2	0.0	...	0.0	1
91	92	SL	KXIP+	9382.0	0.0	10472.0	75.75	0.0	0.0	1.0	...	27.0	
100	101	SL	DC	537.0	1.0	1587.0	70.40	1.0	42.0	0.0	...	1.0	
117	118	SL	DC	3089.0	355.0	2025.0	72.52	400.0	39.4	0.0	...	3.0	
129	130	SL	DC	288.0	64.0	343.0	95.81	108.0	39.4	0.0	...	0.0	

12 rows × 24 columns

In [51]:
```python
# Filtering Data based on Conditions AND(&), OR(|)
ipl[
    (ipl['COUNTRY']=='SL') & (ipl['TEAM']=='MI')
]
```

Out[51]:

	SI.NO.	COUNTRY	TEAM	T-RUNS	T-WKTS	ODI-RUNS-S	ODI-SR-B	ODI-WKTS	ODI-SR-BL	CAPTAINCY EXP	...	SIXERS	RUN
21	22	SL	MI	249.0	97.0	239.0	60.96	187.0	34.7	0.0	...	0.0	29
41	42	SL	MI	6973.0	98.0	13430.0	91.21	323.0	46.0	1.0	...	39.0	39
57	58	SL	MI	275.0	101.0	327.0	73.81	185.0	31.1	0.0	...	4.0	138

3 rows × 24 columns