In [2]:	import pandas as pd import numpy as np import seaborn as sns import matplotlib.pyplot as plt  %matplotlib inline import warnings
<pre>In [4]: In [5]: Out[5]:</pre>	<pre>warnings.filterwarnings('ignore')  ipl = pd.read_csv('C:/Users/VIKAS/Desktop/drashti dextop/IPL DATASET/ipl_2022_dataset.csv')  ipl.head()  Unnamed: 0</pre>
In [6]:	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 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 Shubman Gill Draft Pick BATTER 8.0 1040.0 KKR Gujarat Titans
Out[6]: In [12]: Out[12]:	<pre></pre>
	3
	1 15.0 1950.0 MI Gujarat Titans 2 10.0 1300.0 KKR Gujarat Titans 3 9.0 1170.0 RR Gujarat Titans 4 8.0 1040.0 KKR Gujarat Titans 628 NAN NAN NAN NAN Unsold 629 NAN NAN NAN NAN Unsold 630 NAN NAN NAN NAN Unsold 631 NAN NAN NAN NAN Unsold 632 NAN NAN NAN NAN Unsold
In [13]:	<pre>[633 rows x 8 columns]&gt;  ipl.info()  <class 'pandas.core.frame.dataframe'=""> RangeIndex: 633 entries, 0 to 632 Data columns (total 8 columns): # Column Non-Null Count Dtype</class></pre>
In [14]:	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
Out[14]:  In [15]:  In [16]:	<pre>ipl.columns Index(['Unnamed: 0', 'Player', 'Base Price', 'TYPE', 'COST IN ₹ (CR.)',</pre>
Out[16]:	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
In [17]: Out[17]:	<pre>ipl.isnull().sum()  Player     0 Base Price     0 TYPE</pre>
In [18]: Out[18]:	ipl[ipl['Cost IN \$ (000)'].isnull()]    Player   Base Price   TYPE   COST IN ₹ (CR.)   Cost IN \$ (000)   2021 Squad   Team
	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
In [19]:	632 Aaron Hardie 20 Lakh ALL-ROUNDER NaN NaN NaN Unsold  396 rows × 7 columns  ipl['COST IN ₹ (CR.)'] = ipl['COST IN ₹ (CR.)'].fillna(0) ipl['Cost IN \$ (000)'] = ipl['Cost IN \$ (000)'].fillna(0)  ipl[ipl['2021 Squad'].isnull()]
Out[20]:	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         20 Lakh         BATTER         2.6         338.0         NaN         Gujarat Titans           10         Matthew Wade         2 Cr         WICKETKEEPER         2.4         312.0         NaN         Gujarat Titans           11         Alzarri Joseph         75 Lakh         BOWLER         2.4         312.0         NaN         Gujarat Titans           17         Varun Aaron         50 Lakh         BOWLER         0.5         65.0         NaN         Gujarat Titans
In [21]: In [22]: Out[22]:	<pre>ipl['2021 Squad'] = ipl['2021 Squad'].fillna('Not Participated')  ipl.isnull().sum()  Player     0 Base Price     0 TYPE     0 COOT, THE # (CD.)     0</pre>
<pre>In [23]: Out[23]:</pre>	COST IN ₹ (CR.) 0 COST IN \$ (000) 0 2021 Squad 0 Team 0 dtype: int64  teams = ipl[ipl['COST IN ₹ (CR.)']>0]['Team'].unique() teams  array(['Gujarat Titans', 'Chennai Super Kings', 'Delhi Capitals',
In [24]: In [25]: Out[25]:	'Rajasthan Royals', 'Sunrisers Hyderabad'], dtype=object)  ipl['status'] = ipl['Team'].replace(teams, 'sold')  ipl.head()  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
In [26]:	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
Out[26]:	0Rashid KhanDraft PickBOWLER15.01950.0SRHGujarat Titanssold1Hardik PandyaDraft PickALL-ROUNDER15.01950.0MIGujarat Titanssold2Lockie Ferguson2 CrBOWLER10.01300.0KKRGujarat Titanssold3Rahul Tewatia40 LakhALL-ROUNDER9.01170.0RRGujarat Titanssold4Shubman GillDraft PickBATTER8.01040.0KKRGujarat Titanssold
	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 633 rows × 8 columns
In [27]: Out[27]:	ip1[ip1['Player'].duplicated(keep=False)]           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           240         Amit Mishra         1.5 Cr         BOWLER         0.00         0.0         DC         Unsold         Unsold           499         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
In [28]:	Lalit Yadav 20 Lakh BOWLER 0.00 0.0 Not Participated Unsold Unsold Unsold Unsold State Shubham Singh 20 Lakh ALL-ROUNDER 0.00 0.0 Not Participated Unsold Unsold Unsold Unsold Shivam Sharma 20 Lakh BOWLER 0.00 0.0 Not Participated Unsold Uns
Out[28]: In [29]: Out[29]:	How many types of players have participated?  types = ipl['TYPE'].value_counts() types.reset_index()  index TYPE
In [31]:	<pre>0 ALL-ROUNDER 242 1 BOWLER 215 2 BATTER 112 3 WICKETKEEPER 64  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)</pre>
Out[31]:	Role of Players Participated  ALL-ROUNDER  38.23%  MICKETKEEPER
In [35]:	BOWLER  # Players sold and unsold using a bar graph.
	<pre>plt.figure(figsize=(10,5)) fig = sns.countplot(ip1['status'], palette=['Orange', 'Pink']) plt.xlabel('Sold or Unsold') plt.ylabel('Number of Players') plt.title('Sold vs Unsold', fontsize = 15) plt.plot()  for p in fig.patches:     fig.annotate(format(p.get_height(), '.0f'), (p.get_x() +         p.get_width()/2., p.get_height()), ha = 'center', va = 'center',         xytext = (0, 4), textcoords = 'offset points')</pre>
	Sold vs Unsold  400 - 396  300 - 58 - 237  50 - 237
	150 - 100 -
<pre>In [38]: Out[38]: In [40]:</pre>	<pre>ipl.groupby('status')['Player'].count()  status Unsold    396 sold    237 Name: Player, dtype: int64  # Total number of players bought by each team. plt.figure(figsize=(20,10))</pre>
	<pre>fig = sns.countplot(ip1[ip1['Team']!='Unsold']['Team']) plt.xlabel('Team Names') plt.ylabel('Number of Players') plt.title('Players Bought by Each Team', fontsize=12) plt.xticks(rotation=70) plt.plot()  for p in fig.patches:     fig.annotate(format(p.get_height(), '.0f'), (p.get_x() +         p.get_width()/2., p.get_height()), ha = 'center', va = 'center',         xytext = (0, 4), textcoords = 'offset points')</pre>
	Players Bought by Each Team  25 - 25 - 25 - 25 - 25 - 25 - 25 - 25
	Number of Players  15 -
	5 -
	Colliana Titana (Chenna) Super Kings (Super Kings) (Super
In [41]: In [42]: In [47]:	<pre>ipl['retention'] = ipl['Base Price']  ipl['retention'].replace(['2 Cr', '40 Lakh', '1 Cr', '75 Lakh',</pre>
In [48]: In [49]: In [50]:	<pre>ipl['Base Price'].replace('Draft Pick',0, inplace = True)  ipl['base_price_unit'] = ipl['Base Price'].apply(lambda x: str(x).split(' ')[-1]) ipl['base_price'] = ipl['Base Price'].apply(lambda x: str(x).split(' ')[0])  ipl['base_price'].replace('Retained',0,inplace=True)  ipl.head()</pre>
Out[50]:	
In [51]: Out[51]:	<pre># Total Players retained and bought ipl.groupby(['Team', 'retention']).count()[:-1]  Team</pre>
	Kolkata Knight Riders From Auction 21 Retained 4 Lucknow Super Giants Draft Pick 3 From Auction 19 Mumbai Indians From Auction Retained 4 Punjab Kings From Auction Retained 2 Rajasthan Royals From Auction Retained 3 Royal Challengers Bangalore From Auction 19
In [53]:	Retained 3 Sunrisers Hyderabad From Auction 20 Retained 3 Name: retention, dtype: int64  plt.figure(figsize=(20,10)) fig = sns.countplot(ipl[ipl['Team']!='Unsold']['Team'], hue=ipl['TYPE']) plt.title('Players in Each Team') plt.xlabel('Team Names') plt.ylabel('Number of Player') plt.xticks(rotation=50)
Out[53]:	(array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9]), [Text(0, 0, 'Gujarat Titans'), Text(1, 0, 'Chennai Super Kings'), Text(2, 0, 'Delhi Capitals'), Text(3, 0, 'Kolkata Knight Riders'), Text(4, 0, 'Punjab Kings'), Text(5, 0, 'Lucknow Super Giants'), Text(6, 0, 'Mumbai Indians'), Text(7, 0, 'Royal Challengers Bangalore'), Text(8, 0, 'Rajasthan Royals'),
	Type    Description   Players in Each Team   Type   Barter
	Number of Player  By The Company of Player
	O Light Light Centre Spath Light Grant Gra
<pre>In [66]: Out[66]: In [58]:</pre>	# Highest amount spent on a single player by each team  ipl[ipl['retention']=='From Aucton'].groupby(['Team'])['COST IN ₹ (CR.)'].max()[:-1].sort_values(ascending = False)  Sorios([] Normal COST IN ₹ (CR.)   dtypes float64)
Out[58]:  In [70]:	Player Base Price TYPE COST IN ₹ (CR.) Cost IN \$ (000) 2021 Squad Team status retention base_price_unit base_price  22 Ravindra Jadeja Retained ALL-ROUNDER 16.0 2080.0 CSK Chennai Super Kings sold Retained 0  # Top 5 Bowlers  ipl[(ipl['retention']=='From Aucton') & (ipl['TYPE']=='BOWLER')].sort_values(by = 'COST IN ₹ (CR.)', ascending = False)
Out[70]: In [69]: Out[69]:	ipl.head()           Player         Base Price         TYPE         COST IN ₹ (CR.)         Cost IN \$ (000)         2021 Squad         Team status         retention base_price_unit base_price           0         Rashid Khan         0         BOWLER         15.0         1950.0         SRH Gujarat Titans         sold Draft Pick         0         0           1         Hardik Pandya         0         ALL-ROUNDER         15.0         1950.0         MI Gujarat Titans         sold Draft Pick         0         0
In [ ]:	2 Lockie Ferguson From Auction BOWLER 10.0 1300.0 KKR Gujarat Titans sold From Auction Auction From 3 Rahul Tewatia From Auction ALL-ROUNDER 9.0 1170.0 RR Gujarat Titans sold From Auction From 4 Shubman Gill 0 BATTER 8.0 1040.0 KKR Gujarat Titans sold Draft Pick 0 0