In [61]:
```
import pandas as pd
import numpy as np
```

In [62]:
```
df = pd.read_csv('iplAuction2024.csv')
```

In [63]:
```
df
```

Out[63]:

	Unnamed: 0	PLAYER	NATIONALITY	TYPE	PRICE PAID	TEAM
0	1	Avanish Rao Aravelly	Indian	Wicket-Keeper	20,00,000	Chennai Super Kings
1	2	Mustafizur Rahman	Overseas	Bowler	2,00,00,000	Chennai Super Kings
2	3	Daryl Mitchell	Overseas	All-Rounder	14,00,00,000	Chennai Super Kings
3	4	Sameer Rizvi	Indian	Batter	8,40,00,000	Chennai Super Kings
4	5	Rachin Ravindra	Overseas	All-Rounder	1,80,00,000	Chennai Super Kings
...	...	...	...	...	...	...
67	68	Akash Singh	Indian	Bowler	20,00,000	Sunrisers Hyderabad
68	69	Jaydev Unadkat	Indian	Bowler	1,60,00,000	Sunrisers Hyderabad
69	70	Wanindu Hasaranga	Overseas	All-Rounder	1,50,00,000	Sunrisers Hyderabad
70	71	Pat Cummins	Overseas	All-Rounder	20,50,00,000	Sunrisers Hyderabad
71	72	Travis Head	Overseas	Batter	6,80,00,000	Sunrisers Hyderabad

72 rows × 6 columns

In [64]:
```
df.head()
```

Out[64]:

	Unnamed: 0	PLAYER	NATIONALITY	TYPE	PRICE PAID	TEAM
0	1	Avanish Rao Aravelly	Indian	Wicket-Keeper	20,00,000	Chennai Super Kings
1	2	Mustafizur Rahman	Overseas	Bowler	2,00,00,000	Chennai Super Kings
2	3	Daryl Mitchell	Overseas	All-Rounder	14,00,00,000	Chennai Super Kings
3	4	Sameer Rizvi	Indian	Batter	8,40,00,000	Chennai Super Kings
4	5	Rachin Ravindra	Overseas	All-Rounder	1,80,00,000	Chennai Super Kings

In [65]:
```
df.tail()
```

Out[65]:

	Unnamed: 0	PLAYER	NATIONALITY	TYPE	PRICE PAID	TEAM
67	68	Akash Singh	Indian	Bowler	20,00,000	Sunrisers Hyderabad
68	69	Jaydev Unadkat	Indian	Bowler	1,60,00,000	Sunrisers Hyderabad
69	70	Wanindu Hasaranga	Overseas	All-Rounder	1,50,00,000	Sunrisers Hyderabad
70	71	Pat Cummins	Overseas	All-Rounder	20,50,00,000	Sunrisers Hyderabad
71	72	Travis Head	Overseas	Batter	6,80,00,000	Sunrisers Hyderabad

In [66]:
```
df.shape
```

Out[66]:
```
(72, 6)
```

In [67]:
```
df.isnull().sum()
```

Out[67]:
```
Unnamed: 0     0
PLAYER         0
NATIONALITY    0
TYPE           0
PRICE PAID     0
TEAM           0
dtype: int64
```

In [68]:
```
df.describe()
```

Out[68]:

	Unnamed: 0
count	72.00000
mean	36.50000
std	20.92845
min	1.00000
25%	18.75000
50%	36.50000
75%	54.25000
max	72.00000

In [69]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 72 entries, 0 to 71
Data columns (total 6 columns):
 #   Column       Non-Null Count  Dtype 
---  ------       --------------  ----- 
 0   Unnamed: 0   72 non-null     int64 
 1   PLAYER       72 non-null     object
 2   NATIONALITY  72 non-null     object
 3   TYPE         72 non-null     object
 4   PRICE PAID   72 non-null     object
 5   TEAM         72 non-null     object
dtypes: int64(1), object(5)
memory usage: 3.5+ KB
```

In [70]:
```
df.columns
```

Out[70]:
```
Index(['Unnamed: 0', 'PLAYER', 'NATIONALITY', 'TYPE', 'PRICE PAID', 'TEAM'], dtype='object')
```

Total Spending per team

In [71]:
```
# Clean the 'PRICE PAID' column by removing commas and converting to numerical values for calculations
df['PRICE PAID'] = df['PRICE PAID'].replace({',': ''}, regex=True).astype(float)
```

In [72]:
```
total_spending = df.groupby('TEAM')['PRICE PAID'].sum().reset_index()
```

In [73]:
```
total_spending.head()
```

Out[73]:

	TEAM	PRICE PAID
0	Chennai Super Kings	304000000.0
1	Delhi Capitals	190500000.0
2	Gujarat Titans	303000000.0
3	Kolkata Knight Riders	313500000.0
4	Lucknow Super Giants	122000000.0

Average price per player role

In [74]:
```
average_price_per_player = df.groupby('TYPE')['PRICE PAID'].mean().reset_index()
```

In [75]:
```
average_price_per_player.head()
```

Out[75]:

	TYPE	PRICE PAID
0	All-Rounder	3.154000e+07
1	Batter	3.400000e+07
2	Bowler	3.617308e+07
3	Wicket-Keeper	1.668750e+07

Total spending by Nationality

In [76]:
```
total_spending_by_nationality = df.groupby('NATIONALITY')['PRICE PAID'].sum().reset_index()
```

In [77]:
```
total_spending_by_nationality.head()
```

Out[77]:

	NATIONALITY	PRICE PAID
0	Indian	7.945000e+08
1	Overseas	1.510000e+09

Number of players per team

In [78]:
```
players_per_team = df.groupby('TEAM')['PLAYER'].size().reset_index()
```

In [79]:
```
players_per_team.head()
```

Out[79]:

	TEAM	PLAYER
0	Chennai Super Kings	6
1	Delhi Capitals	9
2	Gujarat Titans	8
3	Kolkata Knight Riders	10
4	Lucknow Super Giants	6

In [80]:
```
df.columns
```

Out[80]:
```
Index(['Unnamed: 0', 'PLAYER', 'NATIONALITY', 'TYPE', 'PRICE PAID', 'TEAM'], dtype='object')
```

Most Expensive player

In [81]:
```
# Identify the most expensive player(s)
most_expensive_player = df[df['PRICE PAID'] == df['PRICE PAID'].max()]
```

In [82]:
```
most_expensive_player.head()
```

Out[82]:

	Unnamed: 0	PLAYER	NATIONALITY	TYPE	PRICE PAID	TEAM
27	28	Mitchell Starc	Overseas	Bowler	247500000.0	Kolkata Knight Riders

Role wise expensive player

In [95]:
```
# Calculate the most expensive player purchase for each team
role_wise = df.loc[df.groupby('TEAM')['PRICE PAID'].idxmax()]
```

In [97]:
```
role_wise = df.loc[df.groupby('TYPE')['PRICE PAID'].idxmax()]
```

In [98]:
```
role_wise_expensive_player = role_wise[['TEAM', 'PLAYER', 'TYPE', 'PRICE PAID']]
```

In [99]:
```
role_wise_expensive_player.head()
```

Out[99]:

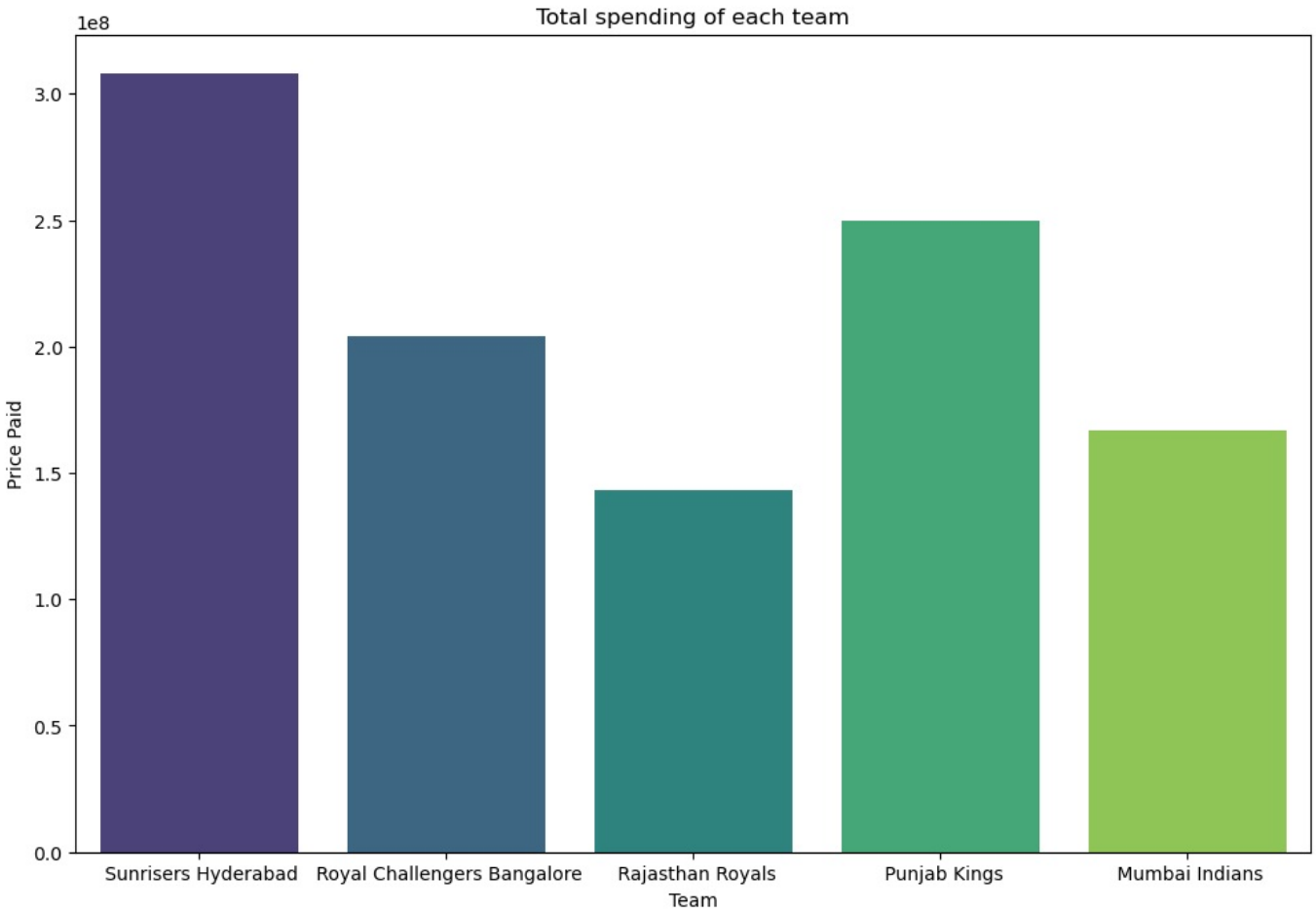
	TEAM	PLAYER	TYPE	PRICE PAID
70	Sunrisers Hyderabad	Pat Cummins	All-Rounder	205000000.0
3	Chennai Super Kings	Sameer Rizvi	Batter	84000000.0
27	Kolkata Knight Riders	Mitchell Starc	Bowler	247500000.0
12	Delhi Capitals	Kumar Kushagra	Wicket-Keeper	72000000.0

Visualizations

In [100…
```
import matplotlib.pyplot as plt
import seaborn as sns
```

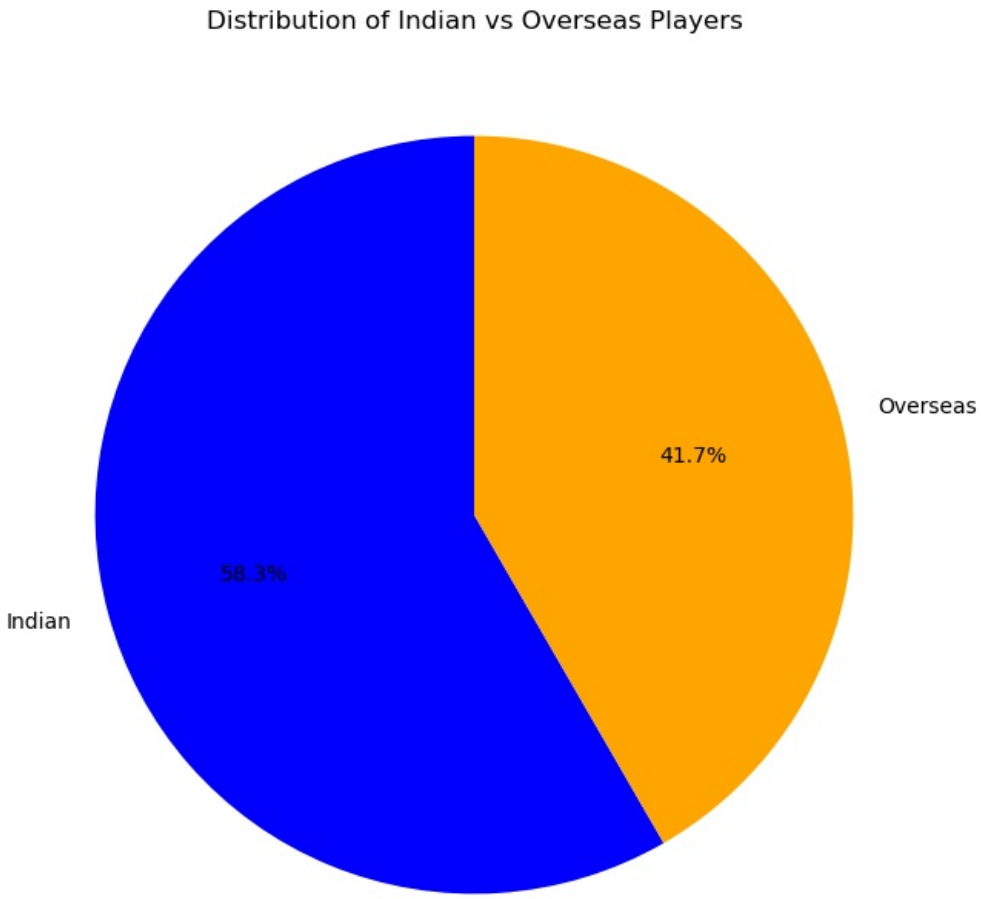
In [101…
```
top_team = total_spending.sort_values(by='TEAM', ascending=False).head(5)
```

In [102…
```
plt.figure(figsize=(12, 8))
sns.barplot(x='TEAM', y='PRICE PAID', data=top_team, palette='viridis')
plt.title("Total spending of each team")
plt.xlabel('Team')
plt.ylabel('Price Paid')
plt.show()
```



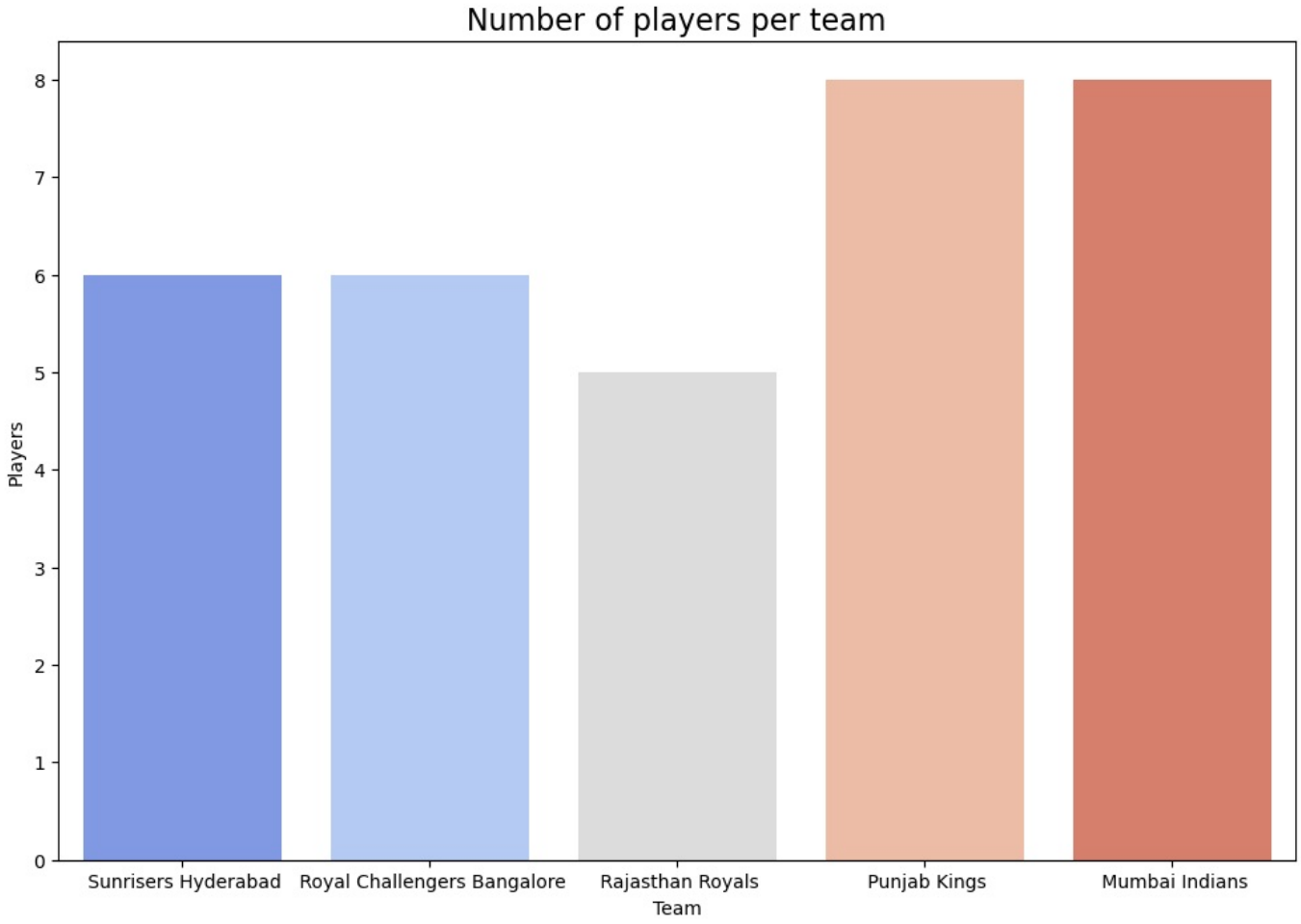
In [103…
```
nationality_counts = df['NATIONALITY'].value_counts()
```

In [104…
```
# Step 4: Plot the pie chart
plt.figure(figsize=(10, 8))  # To control the size of the pie chart
plt.pie(nationality_counts, labels=nationality_counts.index, autopct='%1.1f%%', startangle=90, colors=['blue',

# Adding a title
plt.title('Distribution of Indian vs Overseas Players')

# Display the chart
plt.show()
```



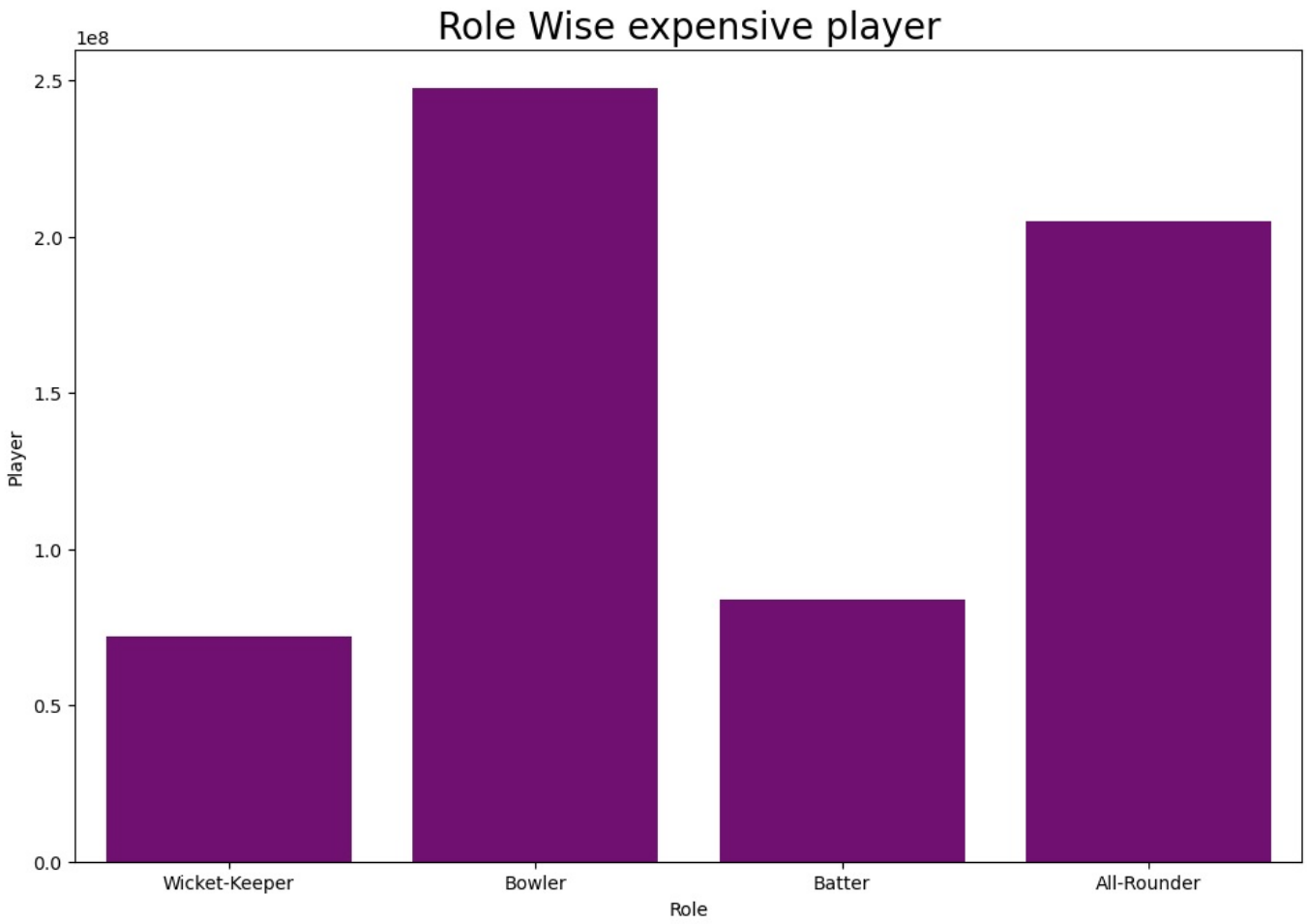
In [105…
```
top_players = players_per_team.sort_values(by='TEAM', ascending=False).head()
```

In [106…
```
plt.figure(figsize=(12, 8))
sns.barplot(x='TEAM', y='PLAYER', data=top_players, palette='coolwarm')
plt.title("Number of players per team", fontsize=16)
plt.xlabel('Team')
plt.ylabel('Players')
plt.show()
```



In [108…
```
top_role_wise = role_wise_expensive_player.sort_values(by='TYPE', ascending=False).head()
```

In [111…
```
plt.figure(figsize=(12, 8))
sns.barplot(x='TYPE', y='PRICE PAID', data=top_role_wise, color='purple')
plt.title("Role Wise expensive player", fontsize=20)
plt.xlabel('Role')
plt.ylabel('Player')
plt.show()
```



In [112…

```
plt.figure(figsize=(12, 8))
sns.barplot(x='PLAYER', y='PRICE PAID', data=top_role_wise, color='red')
plt.title("Role Wise expensive player", fontsize=20)
plt.xlabel('Role')
plt.ylabel('Player')
plt.show()
```