Fifa_Project using matplotlib, seaborn

In [3]:
```
import matplotlib.pyplot as plt # For creating plots and visualiz
import pandas as pd  # For data manipulation and analysis
import seaborn as sns          # For advanced visualizations and
```

Exploratory Data Analysis(EDA)

In [5]:
```
df=pd.read_csv('C://Users//EliteBook//Desktop//ds//fifa_project.c
```

In [8]:
```
df.head()
```

Out[8]:

	Unnamed: 0	ID	Name	Age	Photo	Nat
0	0	158023	L. Messi	31	https://cdn.sofifa.org/players/4/19/158023.png	A
1	1	20801	Cristiano Ronaldo	33	https://cdn.sofifa.org/players/4/19/20801.png	
2	2	190871	Neymar Jr	26	https://cdn.sofifa.org/players/4/19/190871.png	
3	3	193080	De Gea	27	https://cdn.sofifa.org/players/4/19/193080.png	
4	4	192985	K. De Bruyne	27	https://cdn.sofifa.org/players/4/19/192985.png	

5 rows × 89 columns

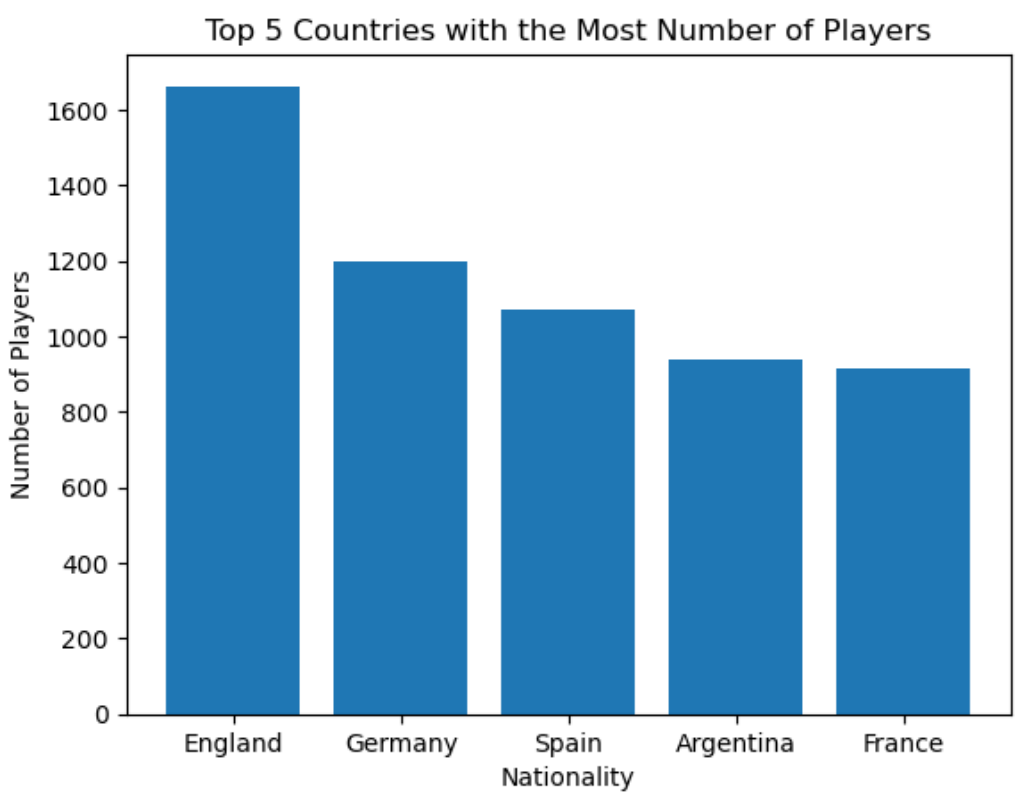
1. Which country has the most number of players

In [9]:
```
country_with_most_players = df['Nationality'].value_counts().idxma
print(f"The country with the most number of players is: {country_
```

```
The country with the most number of players is: England
```

2. Plot a bar chart of 5 top countries with the most number of players

In [10]:
```
top_countries = df['Nationality'].value_counts().nlargest(5)
plt.bar(top_countries.index, top_countries.values)
plt.xlabel('Nationality')
plt.ylabel('Number of Players')
plt.title('Top 5 Countries with the Most Number of Players')
plt.show()
```



Bar chart showing the 5 countries with the highest number of players. England is the 1st position

Column Names

In [16]:
```
df.columns
```

Out[16]:
```
Index(['Unnamed: 0', 'ID', 'Name', 'Age', 'Photo', 'Nationality',
'Flag',
       'Overall', 'Potential', 'Club', 'Club Logo', 'Value', 'Wag
e', 'Special',
       'Preferred Foot', 'International Reputation', 'Weak Foot',
       'Skill Moves', 'Work Rate', 'Body Type', 'Real Face', 'Pos
ition',
       'Jersey Number', 'Joined', 'Loaned From', 'Contract Valid
Until',
       'Height', 'Weight', 'LS', 'ST', 'RS', 'LW', 'LF', 'CF', 'R
F', 'RW',
       'LAM', 'CAM', 'RAM', 'LM', 'LCM', 'CM', 'RCM', 'RM', 'LW
B', 'LDM',
       'CDM', 'RDM', 'RWB', 'LB', 'LCB', 'CB', 'RCB', 'RB', 'Cros
sing',
       'Finishing', 'HeadingAccuracy', 'ShortPassing', 'Volleys',
'Dribbling',
       'Curve', 'FKAccuracy', 'LongPassing', 'BallControl', 'Acce
leration',
       'SprintSpeed', 'Agility', 'Reactions', 'Balance', 'ShotPow
er',
       'Jumping', 'Stamina', 'Strength', 'LongShots', 'Aggressio
n',
       'Interceptions', 'Positioning', 'Vision', 'Penalties', 'Co
mposure',
       'Marking', 'StandingTackle', 'SlidingTackle', 'GKDiving',
'GKHandling',
       'GKKicking', 'GKPositioning', 'GKReflexes', 'Release Claus
e'],
      dtype='object')
```

3. Display the player with the highest salary

In [26]:
```
#Converte
df_sorted = df.sort_values(by='Wage', ascending=False)


highest_salary_player = df_sorted.iloc[0]['Name']
print("Player with the highest salary:", highest_salary_player)
```

```
Player with the highest salary: L. Messi
```

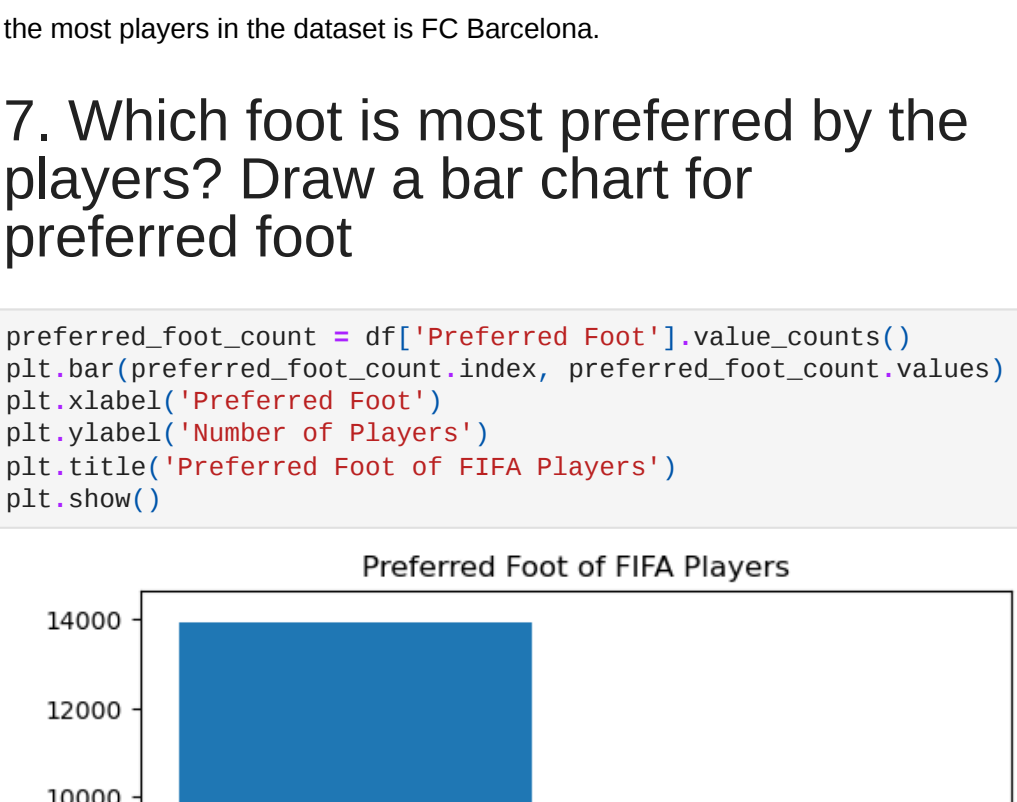
The player with the highest salary is "L. Messi" with a salary of "€565K".

4. Plot a histogram to get the salary range of the players

In [6]:
```

plt.hist(df['Wage'], bins=20)
plt.xlabel('Salary')
plt.ylabel('Number of Players')
plt.title('Salary Range of FIFA Players')
plt.show()
```

```
C:\Users\EliteBook\anaconda3\lib\site-packages\IPython\core\pylab
tools.py:151: UserWarning: Glyph 128 (\x80) missing from current
font.
  fig.canvas.print_figure(bytes_io, **kw)
```



Histogram depicting the distribution of player salaries.

5.Who is the tallest player in the fifa players?

In [34]:
```
#creating a smaller data for easy calculation
df_ht = pd.read_csv('C://Users//EliteBook//Desktop//ds//fifa_proj

#converting type of height to float for comparison
df_ht['Height'] = df_ht['Height'].replace({"'": "."}, regex=True)

# sorting based on height
print("The tallest player with height in feet:")
df_ht.sort_values(by=['Height'], ascending = False).iloc[0]
```

Out[34]:
```
The tallest player with height in feet:
Name      T. Holý
Height        6.9
Name: 11614, dtype: object
```

The tallest player in the FIFA dataset is "T.Holy" with a height of 6.9 feet.

6. Which club has the most number of players?

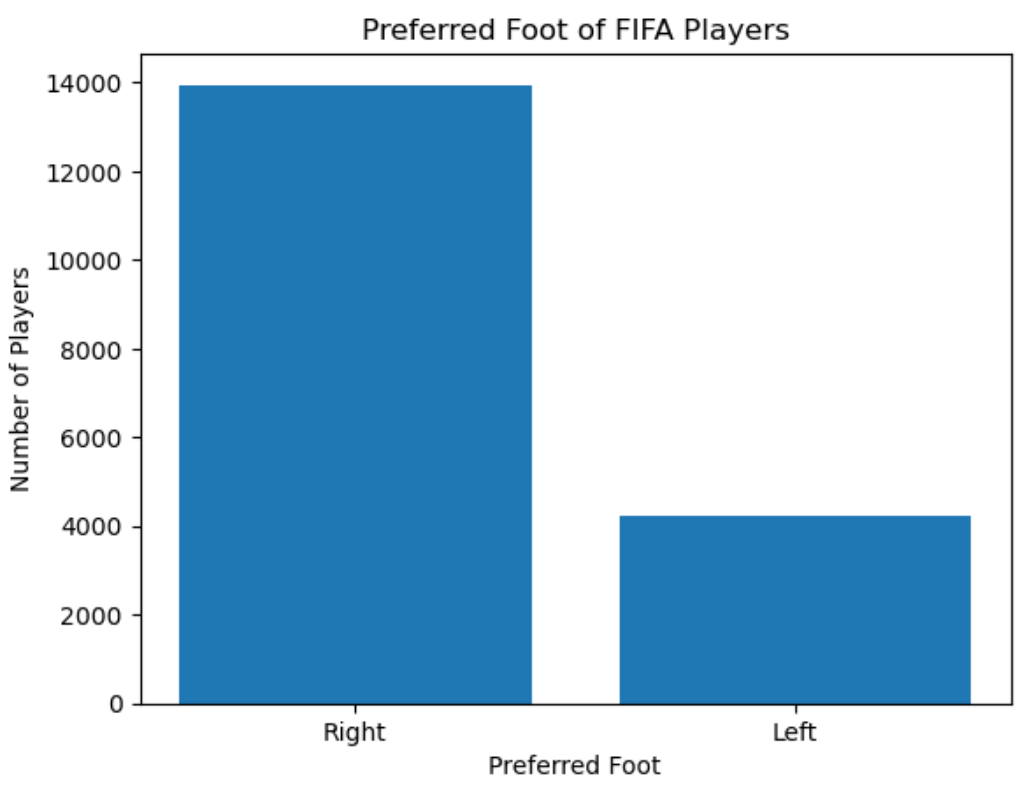
In [36]:
```
club_with_most_players = df['Club'].value_counts().idxmax()
print(f"The club with the most number of players is: {club_with_m
```

```
The club with the most number of players is: FC Barcelona
```

The club with the most players in the dataset is FC Barcelona.

7. Which foot is most preferred by the players? Draw a bar chart for preferred foot

In [38]:
```
preferred_foot_count = df['Preferred Foot'].value_counts()
plt.bar(preferred_foot_count.index, preferred_foot_count.values)
plt.xlabel('Preferred Foot')
plt.ylabel('Number of Players')
plt.title('Preferred Foot of FIFA Players')
plt.show()
```



Bar chart displaying the preferred foot right of players

Conclusion:

The exploratory data analysis of the FIFA Players Dataset provided valuable insights into the demographics, salaries, physical attributes, and club affiliations of football players. It revealed the top countries with the highest number of players, the highest-paid player, the distribution of player salaries, the tallest player, the club with the most players, and the preferred foot among players.