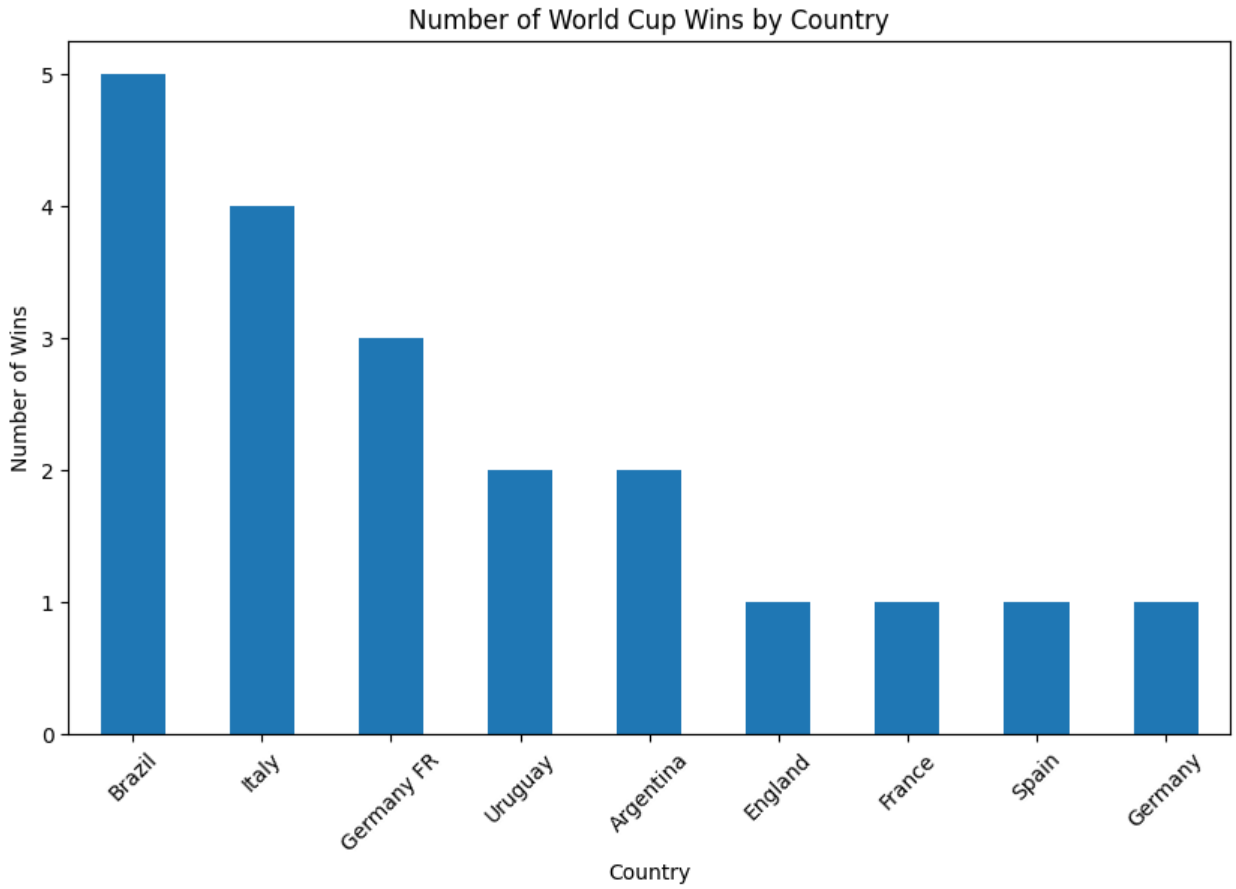
```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import zipfile
import os

# Step 1: Extract the ZIP file
zip_file_path = "FIFA WC data.zip"
extraction_path ="files_fifa"

with zipfile.ZipFile(zip_file_path, 'r') as zip_ref:
    zip_ref.extractall(extraction_path)


# Step 2: Load the datasets
# Assuming the extracted files are in CSV format
# Adjust the file names based on the actual contents of the ZIP file
world_cups_file = os.path.join(extraction_path, 'WorldCups.csv')
world_cup_matches_file = os.path.join(extraction_path,
'WorldCupMatches.csv')
world_cup_players_file = os.path.join(extraction_path,
'WorldCupPlayers.csv')


world_cups = pd.read_csv(world_cups_file)
world_cup_matches = pd.read_csv(world_cup_matches_file)
world_cup_players = pd.read_csv(world_cup_players_file)

# Display the first few rows of the datasets
print("World Cups Data:")
print(world_cups.head())
print("\nWorld Cup Matches Data:")
print(world_cup_matches.head())
print("\nWorld Cup Players Data:")
print(world_cup_players.head())
```

```
World Cups Data:
   Year      Country      Winner      Runners-Up    Third      Fourth
\
0  1930      Uruguay     Uruguay       Argentina      USA  Yugoslavia

1  1934        Italy       Italy  Czechoslovakia  Germany     Austria

2  1938       France       Italy         Hungary   Brazil      Sweden

3  1950       Brazil     Uruguay          Brazil   Sweden       Spain

4  1954  Switzerland  Germany FR         Hungary  Austria     Uruguay
```

```
   GoalsScored  QualifiedTeams  MatchesPlayed Attendance
0           70              13             18    590.549
1           70              16             17    363.000
2           84              15             18    375.700
3           88              13             22  1.045.246
4          140              16             26    768.607

World Cup Matches Data:
     Year              Datetime    Stage         Stadium         City
\
0  1930.0  13 Jul 1930 - 15:00   Group 1         Pocitos  Montevideo

1  1930.0  13 Jul 1930 - 15:00   Group 4  Parque Central  Montevideo

2  1930.0  14 Jul 1930 - 12:45   Group 2  Parque Central  Montevideo

3  1930.0  14 Jul 1930 - 14:50   Group 3         Pocitos  Montevideo

4  1930.0  15 Jul 1930 - 16:00   Group 1  Parque Central  Montevideo


  Home Team Name  Home Team Goals  Away Team Goals Away Team Name  \
0         France              4.0              1.0         Mexico
1            USA              3.0              0.0        Belgium
2     Yugoslavia              2.0              1.0         Brazil
3        Romania              3.0              1.0           Peru
4      Argentina              1.0              0.0         France

  Win conditions  Attendance  Half-time Home Goals  Half-time Away
Goals  \
0                     4444.0                   3.0
0.0
1                    18346.0                   2.0
0.0
2                    24059.0                   2.0
0.0
3                     2549.0                   1.0
0.0
4                    23409.0                   0.0
0.0

                   Referee               Assistant 1  \
0  LOMBARDI Domingo (URU)     CRISTOPHE Henry (BEL)
1       MACIAS Jose (ARG)  MATEUCCI Francisco (URU)
2     TEJADA Anibal (URU)   VALLARINO Ricardo (URU)
3   WARNKEN Alberto (CHI)       LANGENUS Jean (BEL)
4     REGO Gilberto (BRA)      SAUCEDO Ulises (BOL)

                   Assistant 2  RoundID  MatchID Home Team Initials  \
0         REGO Gilberto (BRA)    201.0   1096.0                 FRA
```

```
1       WARNKEN Alberto (CHI)    201.0   1090.0                 USA
2         BALWAY Thomas (FRA)    201.0   1093.0                 YUG
3    MATEUCCI Francisco (URU)    201.0   1098.0                 ROU
4  RADULESCU Constantin (ROU)    201.0   1085.0                 ARG

  Away Team Initials
0                MEX
1                BEL
2                BRA
3                PER
4                FRA

World Cup Players Data:
   RoundID  MatchID Team Initials           Coach Name Line-up  Shirt
Number  \
0      201     1096           FRA  CAUDRON Raoul (FRA)        S
0
1      201     1096           MEX     LUQUE Juan (MEX)        S
0
2      201     1096           FRA  CAUDRON Raoul (FRA)        S
0
3      201     1096           MEX     LUQUE Juan (MEX)        S
0
4      201     1096           FRA  CAUDRON Raoul (FRA)        S
0

        Player Name Position Event
0       Alex THEPOT       GK   NaN
1   Oscar BONFIGLIO       GK   NaN
2  Marcel LANGILLER      NaN  G40'
3      Juan CARRENO      NaN  G70'
4   Ernest LIBERATI      NaN   NaN


world_cups = world_cups['Winner'].value_counts()
print(world_cups)

Winner
Brazil        5
Italy         4
Germany FR    3
Uruguay       2
Argentina     2
England       1
France        1
Spain         1
Germany       1
Name: count, dtype: int64
```

```
plt.figure(figsize=(10,6))
world_cups.plot(kind='bar')
plt.title('Number of World Cup Wins by Country')
plt.xlabel('Country')
plt.ylabel('Number of Wins')
plt.xticks(rotation=45)
plt.show()
```

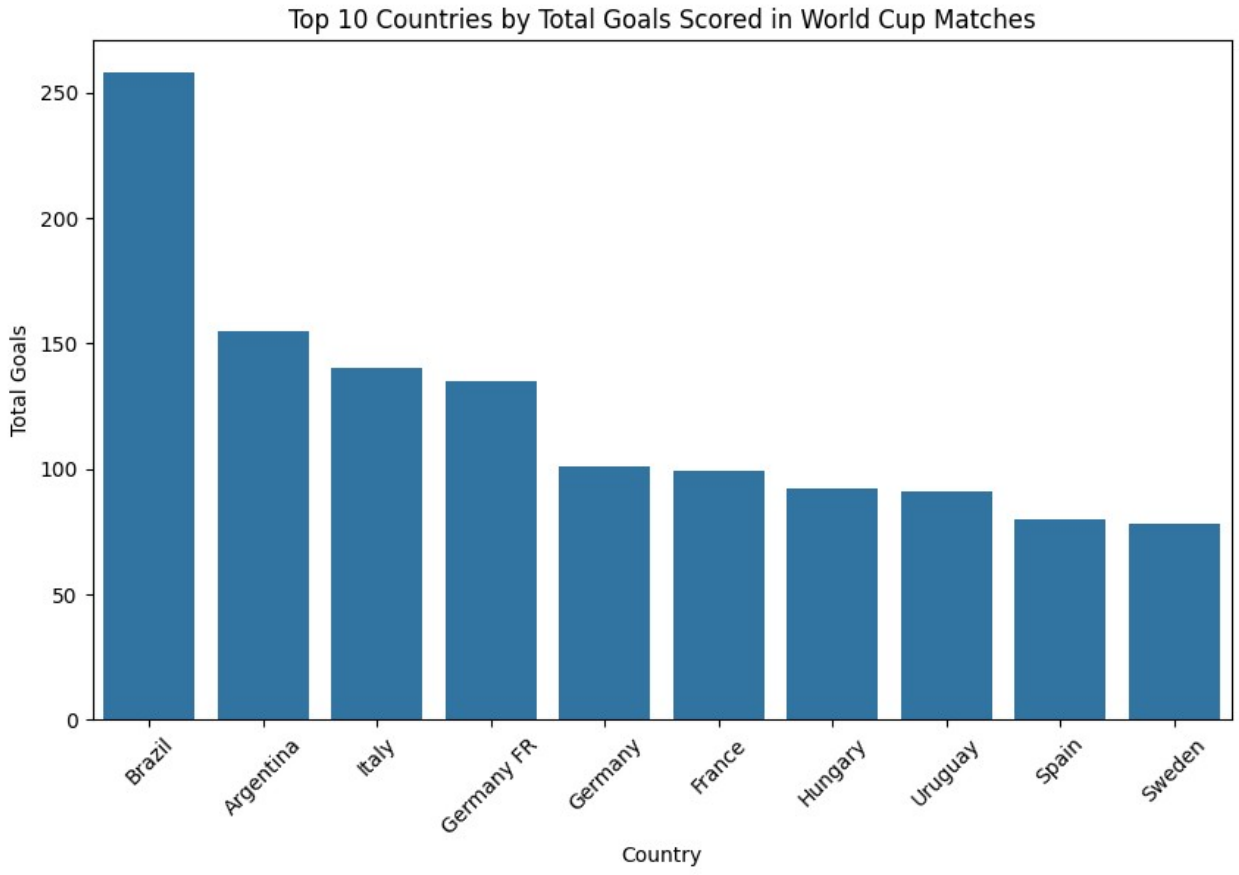



```
world_cup_matches['Total Goals'] = world_cup_matches['Home Team
Goals'] + world_cup_matches['Away Team Goals']
goals_by_country = world_cup_matches.groupby('Home Team Name')['Total
Goals'].sum().reset_index()

goals_by_country = goals_by_country.sort_values(by='Total Goals',
ascending=False)
print(goals_by_country)
```

```
               Home Team Name  Total Goals
7                      Brazil        258.0
2                   Argentina        155.0
35                      Italy        140.0
26                  Germany FR        135.0
```

```
25                      Germany         101.0
..                          ...           ...
46                       Norway           1.0
75  rn">Serbia and Montenegro            1.0
68                      Ukraine           1.0
32                      IR Iran           0.0
76    rn">Trinidad and Tobago             0.0

[78 rows x 2 columns]
```

```
plt.figure(figsize=(10,6))
sns.barplot(x='Home Team Name', y='Total Goals',
data=goals_by_country.head(10))
plt.title('Top 10 Countries by Total Goals Scored in World Cup
Matches')
plt.xlabel('Country')
plt.ylabel('Total Goals')
plt.xticks(rotation=45)
plt.show()
```




```
home_goals = world_cup_matches['Home Team Goals'].sum()
away_goals = world_cup_matches['Away Team Goals'].sum()
```

```
print(f"Total home goals scored: {home_goals}")
print(f"Total away goals scored: {away_goals}")
```

```
Total home goals scored: 1543.0
Total away goals scored: 871.0
```

```
correlation = world_cup_matches[['Home Team Goals', 'Away Team
Goals']].corr()
print(correlation)
```

```
                  Home Team Goals  Away Team Goals
Home Team Goals          1.000000         0.012474
Away Team Goals          0.012474         1.000000
```

```
plt.figure(figsize=(8,6))
sns.heatmap(correlation, annot=True, cmap='coolwarm', fmt='.2f')
plt.title('Correlation Between Home and Away Goals')
plt.show()
```