In [1]:
```
# For manupulating
import pandas as pd
import numpy as np

# for data visualization
import seaborn as sns
import matplotlib.pyplot as plt

# for interactivity
from matplotlib.pyplot import figure
import matplotlib.mlab as mlab
import matplotlib
plt.style.use('ggplot')
```

In [4]:
```
# Now we need to read in the data
fifa=pd.read_csv('wcmatches.csv')
```

In [5]:
```
# It will print first 5 row of data
fifa.head()
```

Out[5]:

	year	country	city	stage	home_team	away_team	home_score	away_score	outcom
0	1930	Uruguay	Montevideo	Group 1	France	Mexico	4	1	H
1	1930	Uruguay	Montevideo	Group 4	Belgium	United States	0	3	A
2	1930	Uruguay	Montevideo	Group 2	Brazil	Yugoslavia	1	2	A
3	1930	Uruguay	Montevideo	Group 3	Peru	Romania	1	3	A
4	1930	Uruguay	Montevideo	Group 1	Argentina	France	1	0	H

In [6]:
```
# For showing all the columns we will use for loop
for col in fifa.columns:
    print(col)
```

```
year
country
city
stage
home_team
away_team
home_score
away_score
outcome
win_conditions
winning_team
losing_team
date
month
dayofweek
```

In [7]:
```
# For showing all the row and columns in the dataset
fifa.shape
```

Out[7]:
```
(900, 15)
```

In [8]:
```
# For knowing which country played highest world cup matches
fifa['country'].value_counts()
```

Out[8]:
```
Germany          102
Brazil            86
Mexico            84
France            82
Italy             69
South Africa      64
Russia            64
Spain             52
United States     52
Argentina         38
Sweden            35
England           32
Chile             32
South Korea       32
Japan             32
Switzerland       26
Uruguay           18
Name: country, dtype: int64
```

In [9]:
```
# For knowing Top 10 country who played highest world cup matches
fifa['country'].value_counts()[0:10]
```

Out[9]:
```
Germany          102
Brazil            86
Mexico            84
France            82
Italy             69
South Africa      64
Russia            64
Spain             52
United States     52
Argentina         38
Name: country, dtype: int64
```

In [10]:
```
# For knowing Top 5 country who played highest world cup matches
fifa['country'].value_counts()[0:5]
```

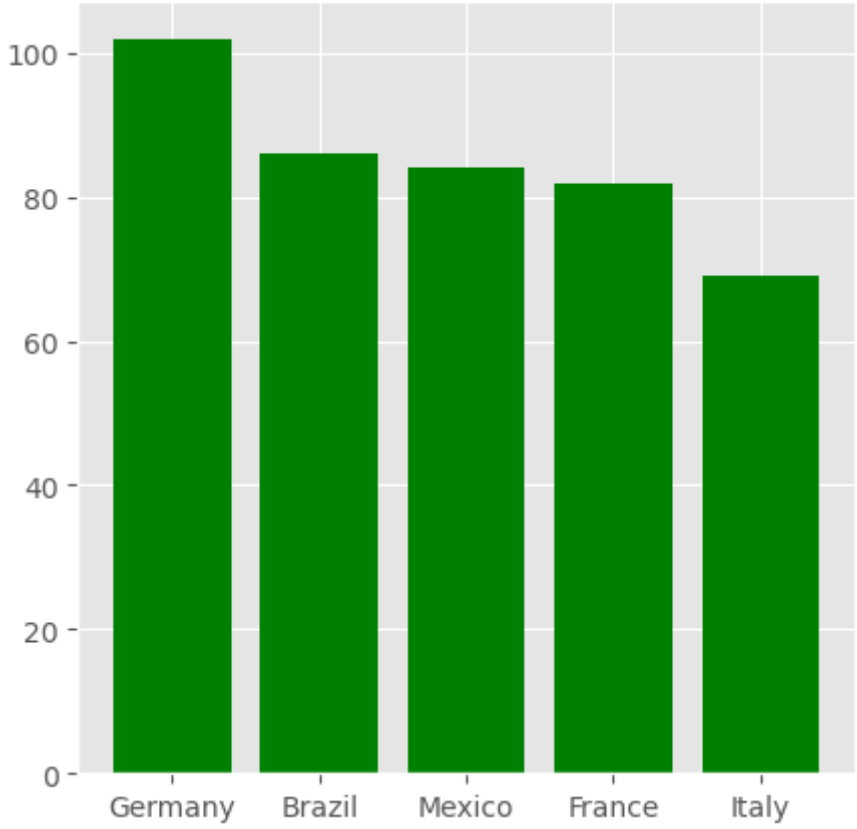
Out[10]:
```
Germany    102
Brazil      86
Mexico      84
France      82
Italy       69
Name: country, dtype: int64
```

In [12]:
```
# Top five countries bar chart for that we need to extract each country value by us
fifa['country'].value_counts()[0:5].keys()
```

Out[12]:
```
Index(['Germany', 'Brazil', 'Mexico', 'France', 'Italy'], dtype='object')
```

In [13]:
```
# Bar chart for top five countries
plt.figure(figsize=(5,5))
plt.bar(list(fifa['country'].value_counts()[0:5].keys()),list(fifa['country'].valu
```

Out[13]:
```
<BarContainer object of 5 artists>
```



In [16]:
```
# by this command we weill find out all wining team and lossing teams in world cup
match_win=fifa[['winning_team','losing_team','home_score','date']]
```

In [17]:
```
match_win.head()
```

Out[17]:

	winning_team	losing_team	home_score	date
0	France	Mexico	4	1930-07-13
1	United States	Belgium	0	1930-07-13
2	Yugoslavia	Brazil	1	1930-07-14
3	Romania	Peru	1	1930-07-14
4	Argentina	France	1	1930-07-15

In [18]:
```
# Highest goal and winning teams from 1930 to 2018
match_win=match_win.sort_values(by=['home_score'],ascending=False)
```

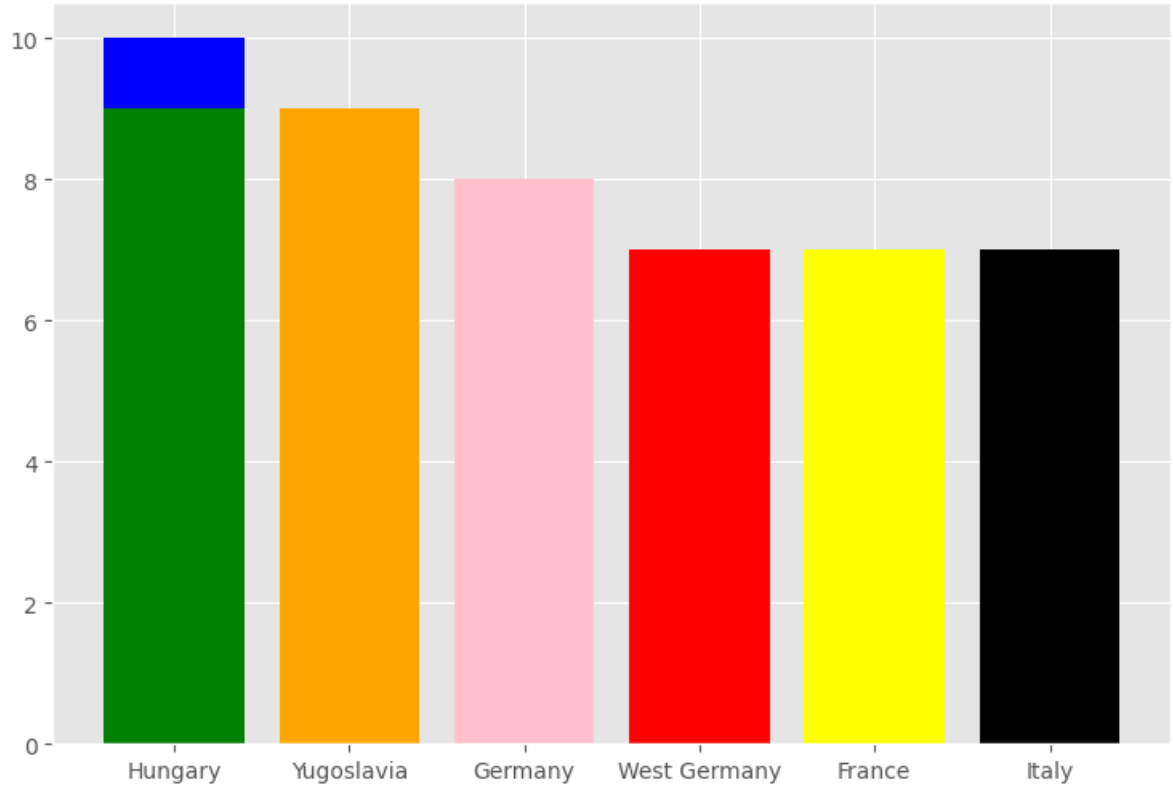
In [19]:
```
match_win.head()
```

Out[19]:

	winning_team	losing_team	home_score	date
311	Hungary	El Salvador	10	1982-06-15
81	Hungary	South Korea	9	1954-06-17
243	Yugoslavia	Zaire	9	1974-06-18
581	Germany	Saudi Arabia	8	2002-06-01
91	West Germany	Turkey	7	1954-06-23

In [29]:
```
# we will make a bar chart by using this top five countries data
plt.figure(figsize=(9,6))
```

```
plt.bar(list(match_win['winning_team'])[0:7],list(match_win['home_score'])[0:7],co
plt.show()
```



In [50]:
```
# For showing all top 10 matches organized country played between 1930 to 2018

Brazil=fifa[fifa['winning_team']=='Brazil']
Brazil=Brazil.sort_values(by=['year'],ascending =False)
Brazil.head(10)
```

Out[50]:

	year	country	city	stage	home_team	away_team	home_score	away_score
888	2018	Russia	Samara	Round of 16	Brazil	Mexico	2	0
878	2018	Russia	Moscow	Group E	Serbia	Brazil	0	2
859	2018	Russia	Saint Petersburg	Group E	Brazil	Costa Rica	2	0
828	2014	Brazil	Fortaleza	Quarterfinals	Brazil	Colombia	2	1
820	2014	Brazil	Belo Horizonte	Round of 16	Brazil	Chile	1	1
805	2014	Brazil	Brasília	Group A	Brazil	Cameroon	4	1
772	2014	Brazil	São Paulo	Group A	Brazil	Croatia	3	1
734	2010	South Africa	Johannesburg	Group G	Brazil	Ivory Coast	3	1
719	2010	South Africa	Johannesburg	Group G	Brazil	North Korea	2	1
760	2010	South Africa	Johannesburg	Round of 16	Brazil	Chile	3	0

In [59]:
```
# Brazill all matches winning summary from 1930 to 2018
Brazil[['city','date','winning_team','losing_team','home_score']].sort_values(by=[
```

Out[59]:

	city	date	winning_team	losing_team	home_score
888	Samara	2018-07-02	Brazil	Mexico	2
878	Moscow	2018-06-27	Brazil	Serbia	0
859	Saint Petersburg	2018-06-22	Brazil	Costa Rica	2
828	Fortaleza	2014-07-04	Brazil	Colombia	2
820	Belo Horizonte	2014-06-28	Brazil	Chile	1
805	Brasília	2014-06-23	Brazil	Cameroon	4
772	São Paulo	2014-06-12	Brazil	Croatia	3
760	Johannesburg	2010-06-28	Brazil	Chile	3
734	Johannesburg	2010-06-20	Brazil	Ivory Coast	3
719	Johannesburg	2010-06-15	Brazil	North Korea	2

In [60]:
```
fifa.head(20)
```

Out[60]:

	year	country	city	stage	home_team	away_team	home_score	away_score	out
0	1930	Uruguay	Montevideo	Group 1	France	Mexico	4	1	
1	1930	Uruguay	Montevideo	Group 4	Belgium	United States	0	3	
2	1930	Uruguay	Montevideo	Group 2	Brazil	Yugoslavia	1	2	
3	1930	Uruguay	Montevideo	Group 3	Peru	Romania	1	3	
4	1930	Uruguay	Montevideo	Group 1	Argentina	France	1	0	
5	1930	Uruguay	Montevideo	Group 1	Chile	Mexico	3	0	
6	1930	Uruguay	Montevideo	Group 2	Bolivia	Yugoslavia	0	4	
7	1930	Uruguay	Montevideo	Group 4	Paraguay	United States	0	3	
8	1930	Uruguay	Montevideo	Group 3	Uruguay	Peru	1	0	
9	1930	Uruguay	Montevideo	Group 1	Argentina	Mexico	6	3	
10	1930	Uruguay	Montevideo	Group 1	Chile	France	1	0	
11	1930	Uruguay	Montevideo	Group 4	Belgium	Paraguay	0	1	
12	1930	Uruguay	Montevideo	Group 2	Bolivia	Brazil	0	4	
13	1930	Uruguay	Montevideo	Group 3	Uruguay	Romania	4	0	
14	1930	Uruguay	Montevideo	Group 1	Argentina	Chile	3	1	
15	1930	Uruguay	Montevideo	Semifinals	Argentina	United States	6	1	
16	1930	Uruguay	Montevideo	Semifinals	Uruguay	Yugoslavia	6	1	
17	1930	Uruguay	Montevideo	Final	Uruguay	Argentina	4	2	
18	1934	Italy	Bologna	Round of 16	Argentina	Sweden	2	3	
19	1934	Italy	Turin	Round of 16	Austria	France	3	2	

In [63]:

```
semi=fifa[fifa['stage']=='Final']
semi[['city','date','winning_team','stage']].sort_values(by=['date'],ascending =Fa
```

Out[63]:

	city	date	winning_team	stage
899	Moscow	2018-07-15	France	Final
835	Rio de Janeiro	2014-07-13	Germany	Final
771	Johannesburg	2010-07-11	Spain	Final
707	Berlin	2006-07-09	Italy	Final
643	Yokohama	2002-06-30	Brazil	Final
579	Saint-Denis	1998-07-12	France	Final
515	Pasadena	1994-07-17	Brazil	Final
463	Rome	1990-07-08	West Germany	Final
411	Mexico City	1986-06-29	Argentina	Final
359	Madrid	1982-07-11	Italy	Final

In [69]:
```
import.os
os.getcwd()
```

```
  File "C:\Users\FC\AppData\Local\Temp\ipykernel_14092\1844270017.py", line 1
    import.os
           ^
SyntaxError: invalid syntax
```

In [70]:
```
pip install nbconveter
```

```
Note: you may need to restart the kernel to use updated packages.
ERROR: Could not find a version that satisfies the requirement nbconveter (from ve
rsions: none)
ERROR: No matching distribution found for nbconveter
```

In [ ]: