In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
```

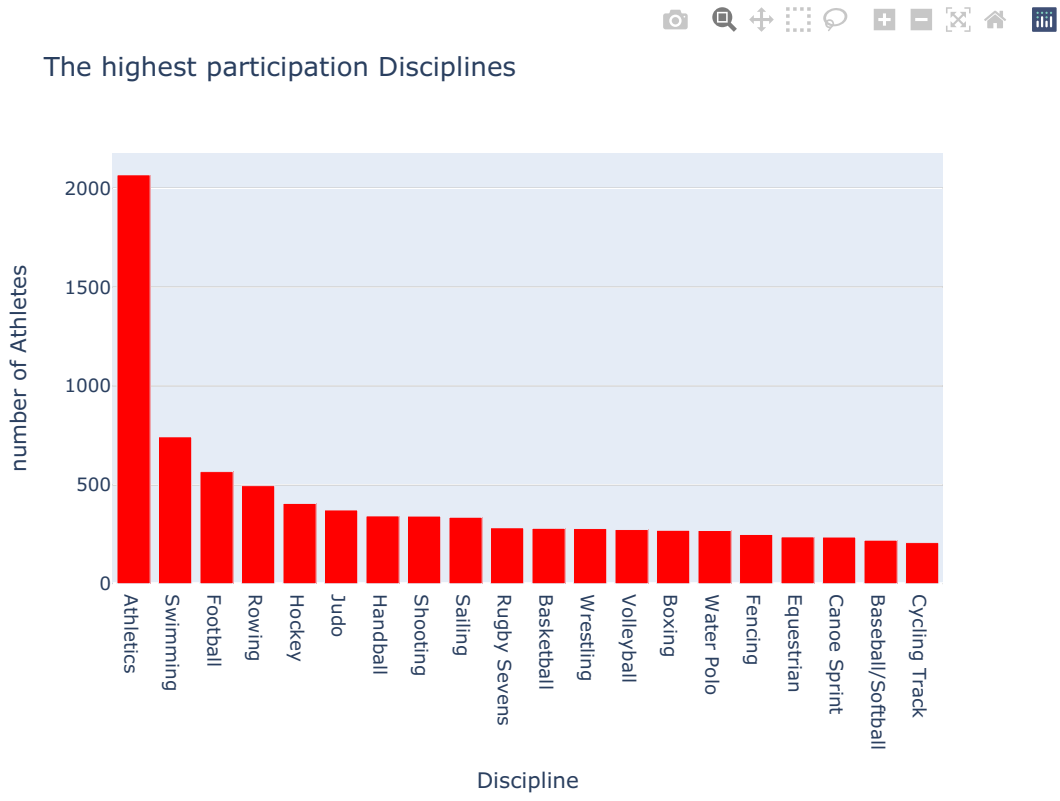
In [2]:
```
df=pd.read_excel("Athletes.xlsx")
```

In [3]:
```
df.head()
```

Out[3]:

	Name	NOC	Discipline
0	AALERUD Katrine	Norway	Cycling Road
1	ABAD Nestor	Spain	Artistic Gymnastics
2	ABAGNALE Giovanni	Italy	Rowing
3	ABALDE Alberto	Spain	Basketball
4	ABALDE Tamara	Spain	Basketball

In [4]:
```
s=df["Discipline"].value_counts().sort_values(ascending=False).nlargest(20)
px.bar(x=s.index,y=s.values,title="The highest participation Disciplines",
       labels={"x":"Discipline","y":"number of Athletes"},
        color_discrete_sequence=["red"])
```

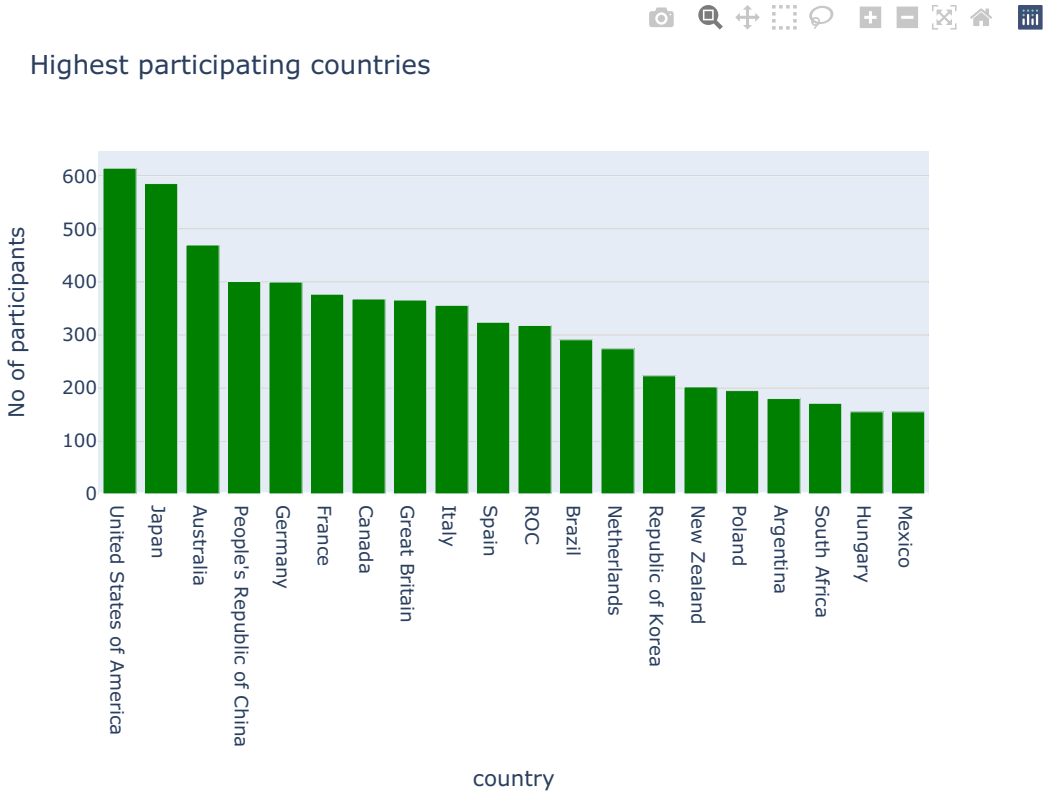


In [8]:
```
series=df['NOC'].value_counts().sort_values(ascending=False).nlargest(20)
series
```

Out[8]:
```
United States of America      615
Japan                         586
Australia                     470
People's Republic of China    401
Germany                       400
France                        377
Canada                        368
Great Britain                 366
Italy                         356
Spain                         324
ROC                           318
Brazil                        291
Netherlands                   274
```

```
Republic of Korea             223
New Zealand                   202
Poland                        195
Argentina                     180
South Africa                  171
Hungary                       155
Mexico                        155
Name: NOC, dtype: int64
```

In [11]:
```
px.bar(x=series.index,y=series.values,title='Highest participating countries',labels={"x":"country","y":"No of pa
```



In [12]:
```
df1=pd.read_excel('EntriesGender.xlsx')
```

In [13]:
```
df1.head()
```

Out[13]:

	Discipline	Female	Male	Total
0	3x3 Basketball	32	32	64
1	Archery	64	64	128
2	Artistic Gymnastics	98	98	196
3	Artistic Swimming	105	0	105
4	Athletics	969	1072	2041

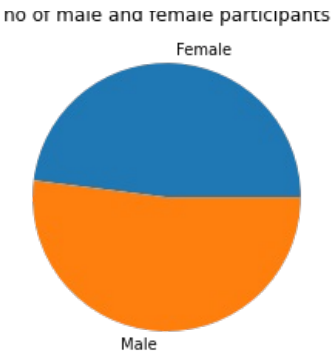
In [ ]:
```

```

In [46]:
```
%matplotlib inline
```

In [47]:
```
total=[df1['Female'].sum(),df1['Male'].sum()]
labels=['Female','Male']
fig,ax=plt.subplots()
print(total)
ax.pie(total,
    labels=labels)
ax.set_title('no of male and female participants')
```

```
[5432, 5884]
```

Out[47]:
```
Text(0.5, 1.0, 'no of male and female participants')
```



In [88]:
```
df2=pd.read_excel("Medals.xlsx")
```

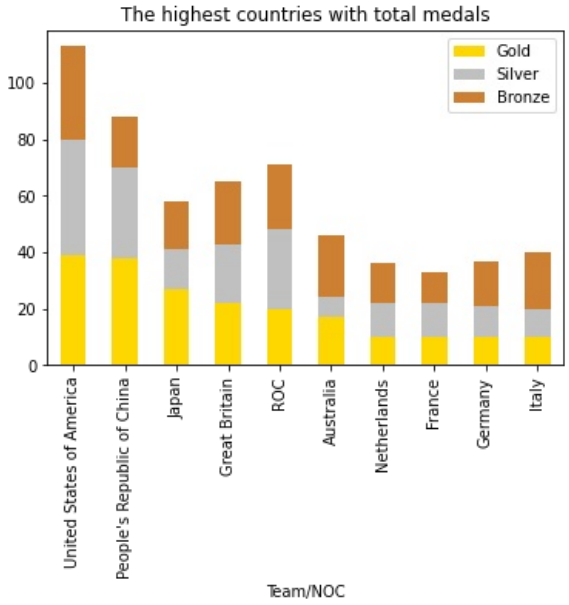
```
C:\Users\kishore\Anaconda3\anaconda\lib\site-packages\openpyxl\styles\stylesheet.py:221: UserWarning:

Workbook contains no default style, apply openpyxl's default
```

In [100…
```
max_medals=df2.loc[0:10:,'Team/NOC':'Total']
px.pie(max_medals,names="Team/NOC",values="Total",title="Highest 10 countries with total medals",
      )
```

In [98]:
```
max_golds=df2.loc[0:10:,'Team/NOC':'Total']
px.pie(max_golds,names="Team/NOC",values="Gold",title="Highest 10 countries with gold medals",
      )
```

In [74]:

```
df3=df2.head(10)
cols=df3.drop(['Rank','Total','Rank by Total'],axis=1)

cols.plot(x='Team/NOC',kind='bar',stacked=True,color=["#FFD700","#C0C0C0","#CD7F32"],title="The highest countries
         )
```



In [ ]: