In [2]:
```python
#The objective of this project is to see the boxplot distribution for the different countries winning the medal.
#The following data is from the website https://www.kaggle.com/arjunprasadsarkhel/2021-olympics-in-tokyo
```

In [97]:
```python
import pandas as pd
import os
print(os.getcwd())
```

C:\Users\spark\Desktop\Data Science Projects\Olympics in Tokyo 2021

In [98]:
```python
import matplotlib.pyplot as plt
import seaborn as sns #Uses matplotlib but works as an upgrade of matplotlib.
%matplotlib inline
```

In [99]:
```python
os.chdir('C:\\Users\\spark\\Desktop\\Data Science Projects\\Olympics in Tokyo 2021') #This is changing the working d
Athletes = pd.read_csv('Athletes.csv')
```

In [100]:
```python
1  Athletes.drop(columns= ["Name"],axis=1,inplace=False) # "Name" Column from the data frame is not necessary so droppe
```

Out[100]:

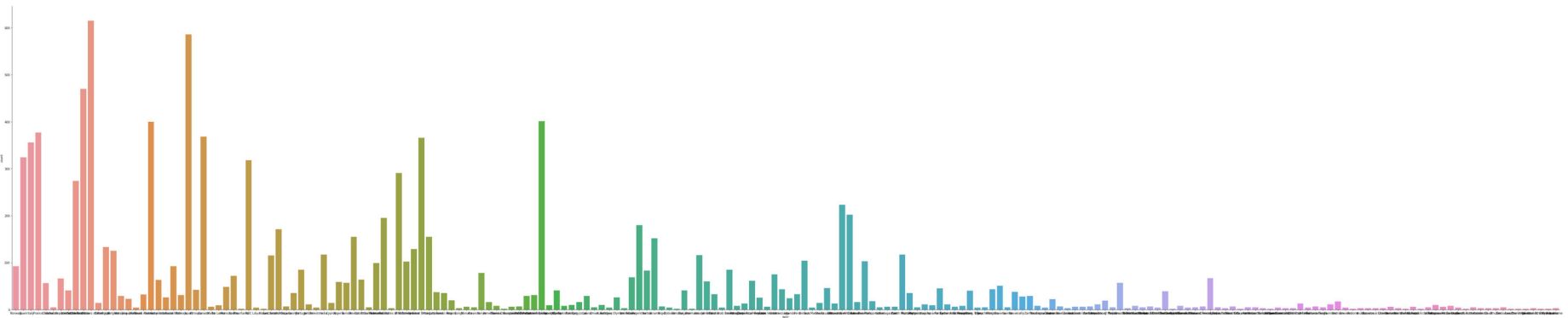|	NOC	Discipline
0	Norway	Cycling Road
1	Spain	Artistic Gymnastics
2	Italy	Rowing
3	Spain	Basketball
4	Spain	Basketball
...	...	...
11080	Germany	Hockey
11081	Poland	Canoe Slalom
11082	ROC	Shooting
11083	ROC	Sailing
11084	Poland	Archery

11085 rows × 2 columns

In [101]:
```python
1  Athletes.head()
```

Out[101]:

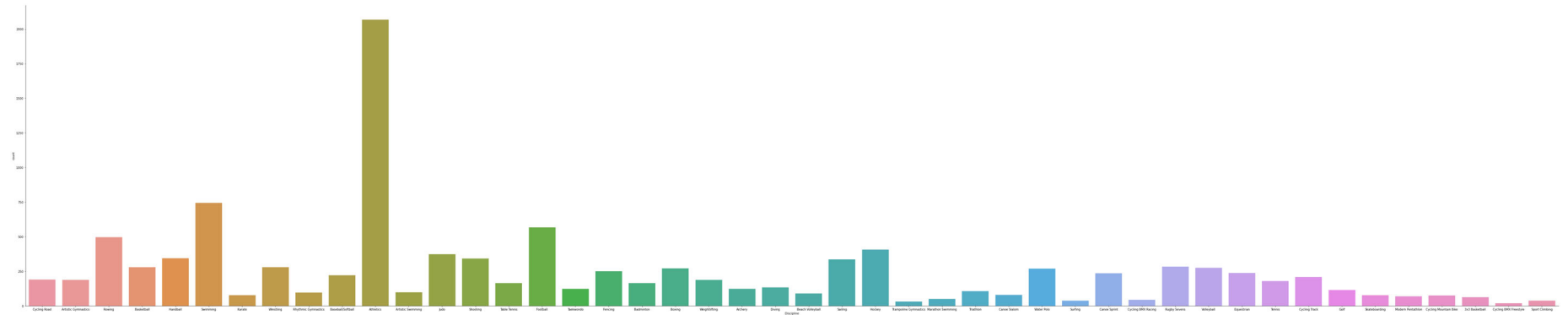|	Name	NOC	Discipline
0	AALERUD Katrine	Norway	Cycling Road
1	ABAD Nestor	Spain	Artistic Gymnastics
2	ABAGNALE Giovanni	Italy	Rowing
3	ABALDE Alberto	Spain	Basketball
4	ABALDE Tamara	Spain	Basketball

In [102]:
```python
sns.catplot(x='NOC',
            kind="count",
            height=15,
            aspect=5,
            data=Athletes)

#In this case, since there are so many bars, I only want to see few bars from here.
```

Out[102]: <seaborn.axisgrid.FacetGrid at 0x1dc23cc1a30>

In [103]:
```python
sns.catplot(x='Discipline',
            kind="count",
            height=15,
            aspect=5,
            data=Athletes)
```

Out[103]: <seaborn.axisgrid.FacetGrid at 0x1dc23089970>

In [155]:

```python
#Looking at the graph above, we realize that Athletics is not part of the types of sports.
#to start, for the Discipline, we want to make sure to drop the following.
#Cycling Road, Artistic Gymnatstics, Rowing, Basketball,

indexNames = Athletes[Athletes['Discipline'] == 'Sport Climbing'].index
indexNames2 = Athletes[Athletes['Discipline'] == 'Cycling BMX Freestyle'].index
indexNames3 = Athletes[Athletes['Discipline'] == '3x3 Basketball'].index
indexNames4 = Athletes[Athletes['Discipline'] == 'Cycling Mountain Bike'].index
indexNames5 = Athletes[Athletes['Discipline'] == 'Modern Pentathlon'].index
indexNames6 = Athletes[Athletes['Discipline'] == 'Skateboarding'].index
indexNames7 = Athletes[Athletes['Discipline'] == 'Golf'].index
indexNames8 = Athletes[Athletes['Discipline'] == 'Cycling Track'].index
indexNames9 = Athletes[Athletes['Discipline'] == 'Tennis'].index
indexNames10 = Athletes[Athletes['Discipline'] == 'Equestrian'].index
indexNames11 = Athletes[Athletes['Discipline'] == 'Volleyball'].index
indexNames12 = Athletes[Athletes['Discipline'] == 'Rugby Sevens'].index
indexNames13 = Athletes[Athletes['Discipline'] == 'Equ'].index
indexNames14 = Athletes[Athletes['Discipline'] == 'PromoRepublic'].index
indexNames15 = Athletes[Athletes['Discipline'] == 'Buffer'].index
indexNames16 = Athletes[Athletes['Discipline'] == 'Twitter for Ipad'].index
indexNames17 = Athletes[Athletes['Discipline'] == 'Taekwondo'].index
indexNames18 = Athletes[Athletes['Discipline'] == 'Artistic Swimming'].index
indexNames19 = Athletes[Athletes['Discipline'] == 'Cycling Road'].index
indexNames20 = Athletes[Athletes['Discipline'] == 'Artistic Gymnastics'].index
indexNames21 = Athletes[Athletes['Discipline'] == 'Karate'].index
indexNames22 = Athletes[Athletes['Discipline'] == 'Swimming'].index
indexNames23 = Athletes[Athletes['Discipline'] == 'Handball'].index
indexNames24 = Athletes[Athletes['Discipline'] == 'Rowing'].index
indexNames25 = Athletes[Athletes['Discipline'] == 'Basketball'].index
indexNames26 = Athletes[Athletes['Discipline'] == 'Wrestling'].index
indexNames27 = Athletes[Athletes['Discipline'] == 'Rhythmic Gymnastics'].index
indexNames28 = Athletes[Athletes['Discipline'] == 'Baseball/Softball'].index
indexNames29 = Athletes[Athletes['Discipline'] == 'Judo'].index
indexNames30 = Athletes[Athletes['Discipline'] == 'Shooting'].index
indexNames31 = Athletes[Athletes['Discipline'] == 'Judo'].index
indexNames32 = Athletes[Athletes['Discipline'] == 'Table Tennis'].index
indexNames33 = Athletes[Athletes['Discipline'] == 'Fencing'].index
indexNames34 = Athletes[Athletes['Discipline'] == 'Badminton'].index
indexNames35 = Athletes[Athletes['Discipline'] == 'Trampoline Gymnastics'].index
indexNames36 = Athletes[Athletes['Discipline'] == 'Marathon Swimming'].index
indexNames37 = Athletes[Athletes['Discipline'] == 'Triathlon'].index
```

```
42  indexNames38 = Athletes[Athletes['Discipline'] == 'Canoe Slalom'].index
43  indexNames39 = Athletes[Athletes['Discipline'] == 'Surfing'].index
44  indexNames40 = Athletes[Athletes['Discipline'] == 'Cycling BMX Racing'].index
45  indexNames41 = Athletes[Athletes['Discipline'] == 'Beach Volleyball'].index
46  indexNames42 = Athletes[Athletes['Discipline'] == 'Cycling BMX Racing'].index
47  indexNames43 = Athletes[Athletes['Discipline'] == 'Cycling Mountain Bike'].index
48  indexNames44 = Athletes[Athletes['Discipline'] == 'Archery'].index
49  indexNames45 = Athletes[Athletes['Discipline'] == 'Athletics'].index
```

In [156]:

```
1  Athletes.drop(indexNames, inplace=True)
2  Athletes.drop(indexNames2, inplace=True)
3  Athletes.drop(indexNames3, inplace=True)
4  Athletes.drop(indexNames5, inplace=True)
5  Athletes.drop(indexNames6, inplace=True)
6  Athletes.drop(indexNames7, inplace=True)
7  Athletes.drop(indexNames8, inplace=True)
8  Athletes.drop(indexNames9, inplace=True)
9  Athletes.drop(indexNames10, inplace=True)
10 Athletes.drop(indexNames11, inplace=True)
11 Athletes.drop(indexNames12, inplace=True)
12 Athletes.drop(indexNames13, inplace=True)
13 Athletes.drop(indexNames14, inplace=True)
14 Athletes.drop(indexNames15, inplace=True)
15 Athletes.drop(indexNames16, inplace=True)
16 Athletes.drop(indexNames17, inplace=True)
17 Athletes.drop(indexNames18, inplace=True)
18 Athletes.drop(indexNames19, inplace=True)
19 Athletes.drop(indexNames20, inplace=True)
20 Athletes.drop(indexNames21, inplace=True)
21 Athletes.drop(indexNames22, inplace=True)
22 Athletes.drop(indexNames23, inplace=True)
23 Athletes.drop(indexNames24, inplace=True)
24 Athletes.drop(indexNames25, inplace=True)
25 Athletes.drop(indexNames26, inplace=True)
26 Athletes.drop(indexNames27, inplace=True)
27 Athletes.drop(indexNames28, inplace=True)
28 Athletes.drop(indexNames29, inplace=True)
29 Athletes.drop(indexNames30, inplace=True)
30 Athletes.drop(indexNames31, inplace=True)
31 Athletes.drop(indexNames32, inplace=True)
32 Athletes.drop(indexNames33, inplace=True)
33 Athletes.drop(indexNames34, inplace=True)
34 Athletes.drop(indexNames35, inplace=True)
35 Athletes.drop(indexNames36, inplace=True)
36 Athletes.drop(indexNames37, inplace=True)
37 Athletes.drop(indexNames38, inplace=True)
38 Athletes.drop(indexNames39, inplace=True)
39 Athletes.drop(indexNames40, inplace=True)
40 Athletes.drop(indexNames41, inplace=True)
41 Athletes.drop(indexNames42, inplace=True)
```

```
42  Athletes.drop(indexNames43, inplace=True)
43  Athletes.drop(indexNames44, inplace=True)
44  Athletes.drop(indexNames45, inplace=True)
```
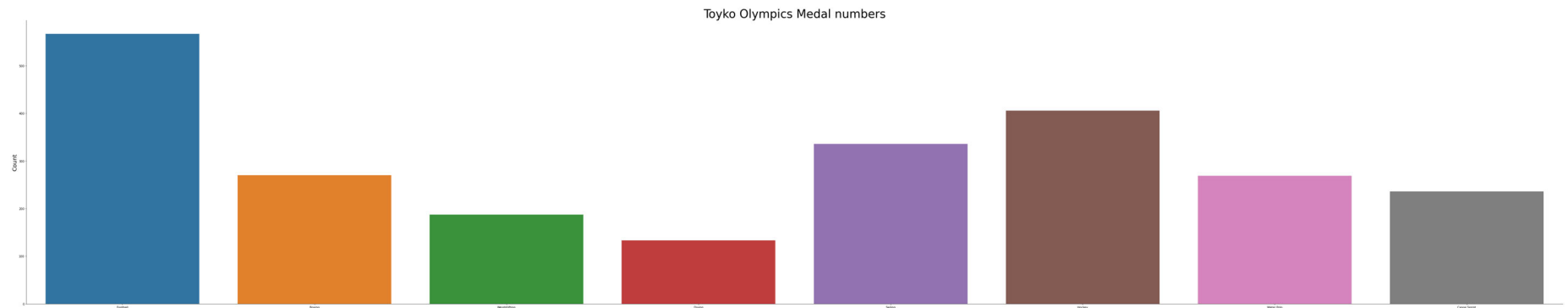
In [169]:
```
 1  sns.catplot( x='Discipline',
 2               kind="count",
 3               height=15,
 4               aspect=5,
 5               data=Athletes)
 6  plt.xlabel("Discipline", size=20,color='white')
 7  plt.ylabel("Count", size=20)
 8  plt.title("Toyko Olympics Medal numbers",size = 40)
 9  plt.tight_layout()
10  plt.savefig("TokoOlympicsMedalnumber.png")
```


Toyko Olympics Medal numbers

In [168]:
```
1  #Analysis: The result indicates that from the 11,000 atheltes with 47 disciplines,
2  #participants received the most medals in disciplines of Football, Boxing,
3  # Weightlifting, Diving, Sailing, Hockey, Water Polo, Canoe Sprint. The common feature
4  # of these disciplines is that these require group works. For example, soccer can always
5  # be done with a team of 11.
```