Practical No.4

Data Science and Visualization (Honors Course)

Name:Preeti Shantilal Chaudhari

PRN: 72017986G

Class: TE ENTC 'B'

In this practical we will perform Data Visualization.

In [1]:

```
import pandas as pd
```

In [6]:

```
df = pd.read_csv('netflix_titles.csv')
df.head(8807)
```

Out[6]:

	show_id	type	title	director	cast	country	date_added	release_year	rat
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	F
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	
...	...	...	...	...	...	...	...	...	
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	

	show_id	type	title	director	cast	country	date_added	release_year	rat
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV

8807 rows × 12 columns

In [5]:

```
df.shape
```

Out[5]:

```
(8807, 12)
```

In [8]:

```
categories=df['listed_in']
```

In [9]:

```
total_child=sum(df['listed_in'].str.contains('Child'))
```

In [10]:

```
total_child
```

Out[10]:

```
641
```

In [11]:

```
Standup_Comedies=sum(df['listed_in'].str.contains('Stand'))
```

In [13]:

```
Standup_Comedies
```

Out[13]:

```
399
```

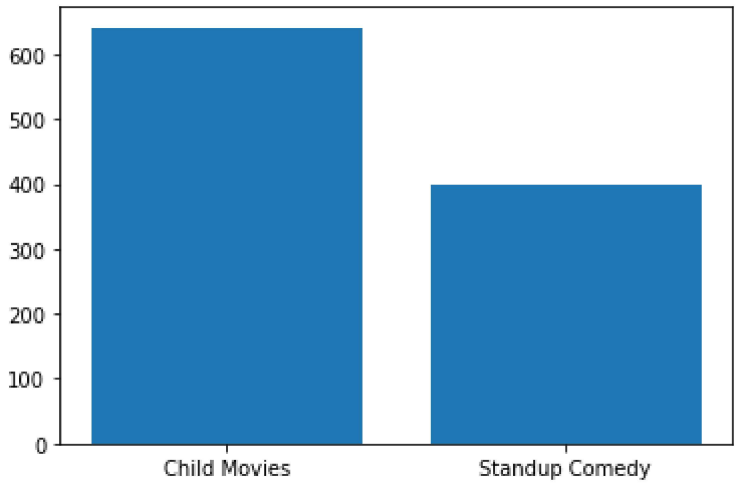
We determined the number of child movies/shows and standup comedies.We will visualize this number using plot.

In [14]:

```
import matplotlib.pyplot as plt
```

In [19]:

```
plt.bar(['Child Movies','Standup Comedy'],
        [total_child, Standup_Comedies])
plt.show()
```



In [20]:

```
set(df['type'])
```

Out[20]:

```
{'Movie', 'TV Show'}
```

In [21]:

```
tv_shows = df[df['type'] == 'TV Show'] #Boolean Filtering
```

In [30]:

```
seasons13 = tv_shows [tv_shows [ 'duration'] == '13 Seasons']

seasons15 = tv_shows [tv_shows['duration'] == '15 Seasons']

seasons16= tv_shows [tv_shows['duration'] == '16 Seasons']

seasons12 = tv_shows [tv_shows['duration'] == '12 Seasons']

seasons11= tv_shows [tv_shows['duration'] == '11 Seasons']
```

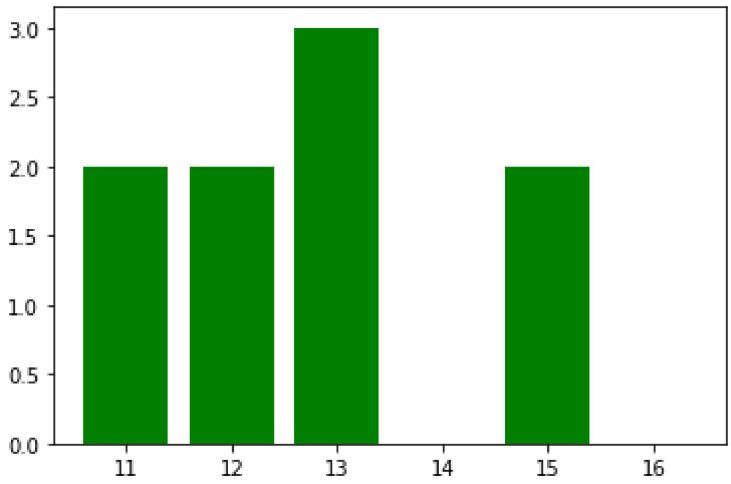
In [31]:

```
plt.bar ([11, 12, 13, 15, 16],
[len(seasons11), len(seasons12), len(seasons13), len (seasons15), len(seasons16)],

color='green')
```

Out[31]:

```
<BarContainer object of 5 artists>
```



In [32]:

```
from collections import Counter
ratings = Counter(df['rating'])
```

In [33]:

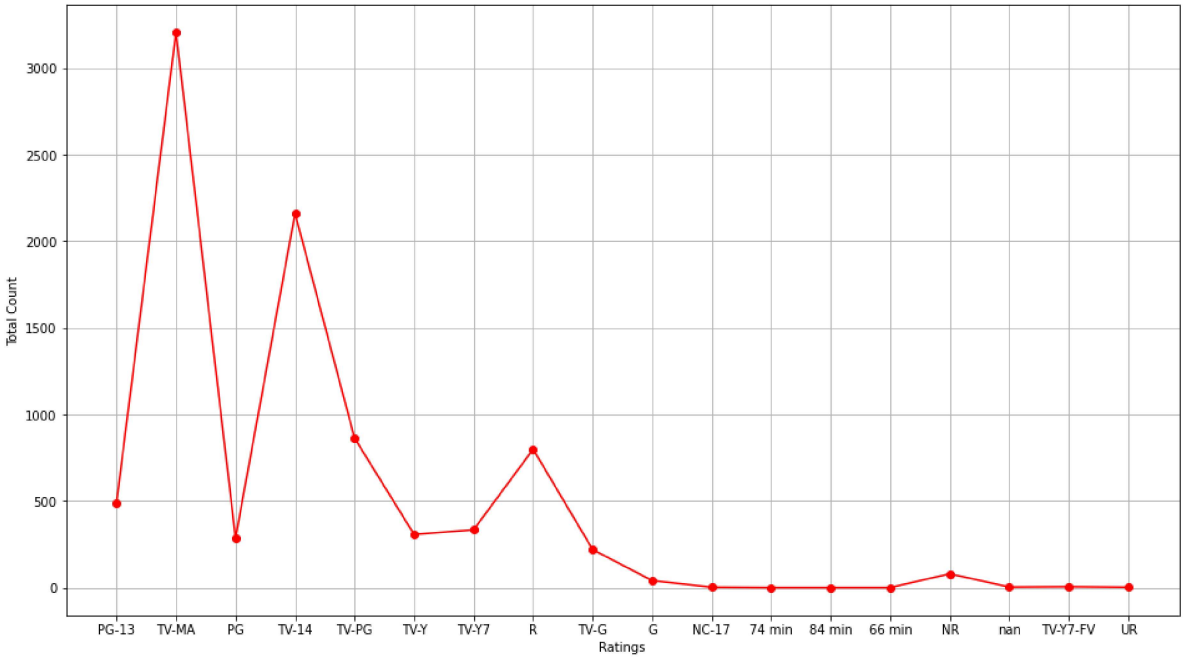
```
ratings
```

Out[33]:

```
Counter({'PG-13': 490,
         'TV-MA': 3207,
         'PG': 287,
         'TV-14': 2160,
         'TV-PG': 863,
         'TV-Y': 307,
         'TV-Y7': 334,
         'R': 799,
         'TV-G': 220,
         'G': 41,
         'NC-17': 3,
         '74 min': 1,
         '84 min': 1,
         '66 min': 1,
         'NR': 80,
         nan: 4,
         'TV-Y7-FV': 6,
         'UR': 3})
```

In [36]:

```
plt. figure(figsize=(16,9))
plt.plot(ratings.keys(), ratings.values(), color = 'red', marker='o')
plt.xlabel('Ratings'); plt.ylabel('Total Count')
plt.grid()
```



If we wish to plot all these plots in the same plot we can use subplot.

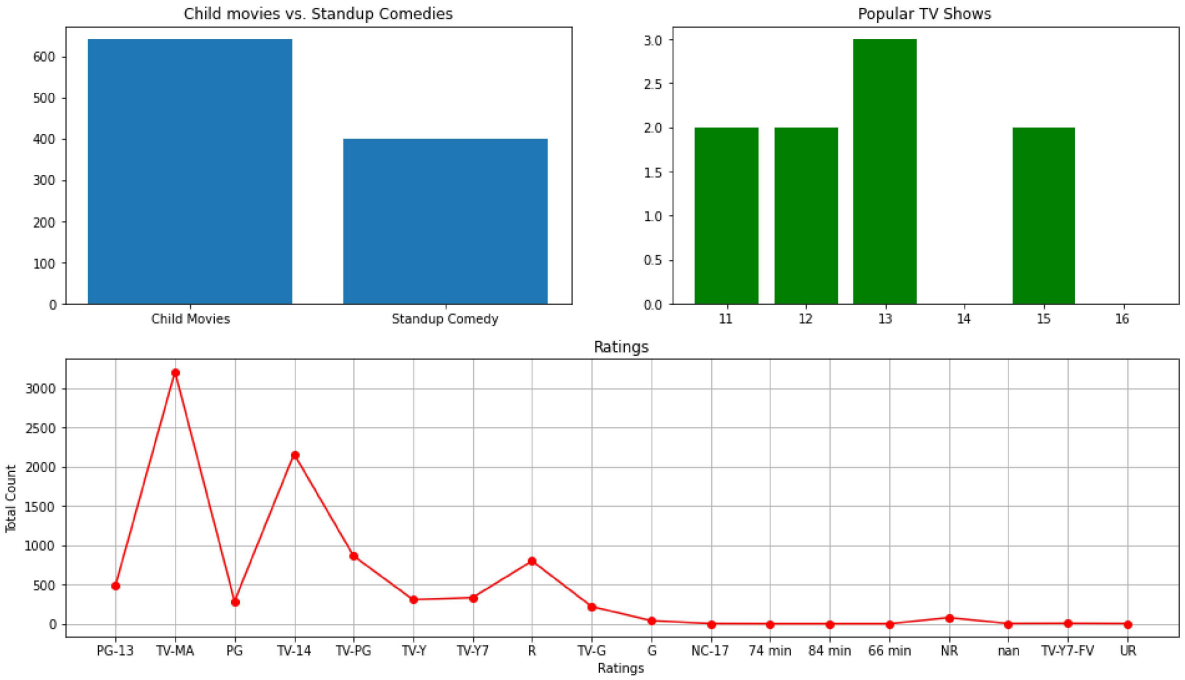
In [40]:

```
plt. figure (figsize=(16,9))

#plot1
plt.subplot (2,2,1)
plt.title ("Child movies vs. Standup Comedies")
plt.bar(['Child Movies', 'Standup Comedy'], [total_child, Standup_Comedies])

#plot2
plt. subplot (2,2,2)
plt.title('Popular TV Shows')
plt.bar([11, 12, 13, 15, 16],
[len (seasons11), len (seasons12), len(seasons13),
len (seasons15), len (seasons16)],
color='green')

#plot3
plt.subplot (2,1,2)
plt.title('Ratings')
plt.plot(ratings.keys (), ratings.values(), color='red', marker='o')
plt.xlabel('Ratings'); plt.ylabel('Total Count')
plt.grid()
```



In [ ]: