


```
In [1]: import requests
import pandas as pd
from bs4 import BeautifulSoup
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px
import numpy as np
import re
from datetime import datetime
```

```
In [16]: url='https://www.themoviedb.org/movie'
```

```
In [17]: url = 'https://www.themoviedb.org'

header = {'User-Agent': 'Mozilla/5.0 (Windows NT 6.3; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/99.0.4844.82 Safari/537.36'}

response = requests.get(url+'/movie', headers = header)
```



```
In [18]: response
```

```
Out[18]: <Response [200]>
```

```
In [19]: content=response.text
```

```
In [20]: soup = BeautifulSoup(content,'lxml')
```

```
In [21]: card_style = soup.find('div',class_='card style_1')
movie_name= card_style.find('h2').text
movie_rating= soup.find('div',class_='user_score_chart')['data-percent']
movie_link='https://www.themoviedb.org'+card_style.find('a')['href']
url2=requests.get(movie_link,headers=header)
movie_url=BeautifulSoup(url2.text,'lxml')
movie_release=movie_url.find('span',class_='release').text
movie_genres=movie_url.find('span',class_='genres').text.strip().replace('\xa0','')
movie_director=movie_url.find('li',class_='profile').text
movie_genres
```

Out[21]: 'Action,Science Fiction,Comedy,Family,Adventure'

```
In [22]: card_style
```

```
Out[22]: <div class="card style_1">
<div class="image">
<div class="wrapper">
<a class="image" href="/movie/675353" title="Sonic the Hedgehog 2">

</a>
</div>
<div class="options" data-id="675353" data-media-type="movie" data-object-id="5e4fc5e635811d0013522699">
<a class="no_click" href="#"><div class="glyphicons_v2 circle-more white"></div></a>
</div>
</div>
<div class="content">
<div class="consensus tight">
<div class="outer_ring">
<div class="user_score_chart 5e4fc5e635811d0013522699" data-bar-color="#21d07a" data-percent="77.0" data-track-color
="#204529">
<div class="percent">
<span class="icon icon-r77"></span>
</div>
</div>
</div>
</div>
```

```
In [23]: page_url='https://www.themoviedb.org/movie?page='
```



```

In [75]: all_movieinfo=[]
def get_all_movies_data():
    count=0
    for link in lst_url:
        count+=1
        print(count)
        response = requests.get(link,headers={'User-Agent': 'Mozilla/5.0 '})
        html_text = response.text
        movie_soup = BeautifulSoup(html_text,'lxml')
        all_movie_pages = movie_soup.find_all('div',class_='card style_1')
        for item in all_movie_pages:
            movie_name = item.find('div',class_='content').h2.text
            movie_ratings= item.find('div',class_='user_score_chart')['data-percent']
            movie_link= item.find('div',class_='content').a['href']
            movie_link= "https://www.themoviedb.org"+movie_link
            movie_specificurl = requests.get(movie_link,headers = {'User-Agent': 'Mozilla/5.0 '} )
            movie_specificurl2= BeautifulSoup(movie_specificurl.text,'lxml')
            movie_specificurl2
            movie_release = movie_specificurl2.find('span',class_='release').text.strip()
            movie_genere = movie_specificurl2.find('span',class_='genres').text.strip().replace(',','')
            runtime = movie_specificurl2.find('span',class_='runtime')
            if runtime is not None:
                runtime=(runtime.text.strip())
            director=movie_specificurl2.find('li',class_='profile')
            if director is not None:
                director=(director.p.text)
            my_movies = {
                'Name':movie_name,
                'Ratings':movie_ratings,
                'Genre':movie_genere,
                'Release date':movie_release,
                'Runtime' :runtime,
                'Director' :director,
                'Url':movie_link
            }

            all_movieinfo.append(my_movies)
get_all_movies_data()

```

1

2

3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19

In [76]: `df= pd.DataFrame(all_movieinfo)`

In [77]: df

Out[77]:

	Name	Ratings	Genre	Release date	Runtime	Director	Url
0	Sonic the Hedgehog 2	77.0	Action, Science Fiction, Comedy, Family, Adven...	04/08/2022 (US)	2h 2m	Josh Miller	<a href="https://www.themoviedb.org/movie/675353">https://www.themoviedb.org/movie/675353</a>
1	The Batman	78.0	Crime, Mystery, Thriller	03/04/2022 (IN)	2h 56m	Matt Reeves	<a href="https://www.themoviedb.org/movie/414906">https://www.themoviedb.org/movie/414906</a>
2	Doctor Strange in the Multiverse of Madness	75.0	Fantasy, Action, Adventure	05/06/2022 (US)	2h 6m	Stan Lee	<a href="https://www.themoviedb.org/movie/453395">https://www.themoviedb.org/movie/453395</a>
3	The Bad Guys	77.0	Animation, Comedy, Action, Family, Crime	04/22/2022 (US)	1h 40m	Pierre Perifel	<a href="https://www.themoviedb.org/movie/629542">https://www.themoviedb.org/movie/629542</a>
4	Uncharted	72.0	Action, Adventure	02/18/2022 (US)	1h 56m	Ruben Fleischer	<a href="https://www.themoviedb.org/movie/335787">https://www.themoviedb.org/movie/335787</a>
...	...	...	...	...	...	...	...
3995	Twister	64.0	Action, Adventure, Drama	05/10/1996 (US)	1h 53m	Jan de Bont	<a href="https://www.themoviedb.org/movie/664">https://www.themoviedb.org/movie/664</a>
3996	Superman/Shazam!: The Return of Black Adam	73.0	Animation, Action, Science Fiction, Fantasy	11/16/2010 (ES)	25m	C.C. Beck	<a href="https://www.themoviedb.org/movie/43641">https://www.themoviedb.org/movie/43641</a>
3997	27 Dresses	64.0	Comedy, Romance	01/18/2008 (US)	1h 51m	Anne Fletcher	<a href="https://www.themoviedb.org/movie/6557">https://www.themoviedb.org/movie/6557</a>
3998	Tangled: Before Ever After	68.0	Adventure, Animation, Comedy, Family, Fantasy,...	10/26/2017 (DE)	55m	Tom Caulfield	<a href="https://www.themoviedb.org/movie/438747">https://www.themoviedb.org/movie/438747</a>
3999	The Intruder	61.0	Thriller	05/03/2019 (US)	1h 42m	Deon Taylor	<a href="https://www.themoviedb.org/movie/524247">https://www.themoviedb.org/movie/524247</a>

4000 rows × 7 columns



In [78]: df.to\_csv('all\_movieinfo.csv')

```
In [2]: data= pd.read_csv('all_movieinfo.csv')
```

```
In [3]: data
```

```
Out[3]:
```

	Unnamed: 0	Name	Ratings	Genre	Release date	Runtime	Director	
0	0	Sonic the Hedgehog 2	77.0	Action, Science Fiction, Comedy, Family, Adven...	04/08/2022 (US)	2h 2m	Josh Miller	<a href="https://www.themoviedb.org/n">https://www.themoviedb.org/n</a>
1	1	The Batman	78.0	Crime, Mystery, Thriller	03/04/2022 (IN)	2h 56m	Matt Reeves	<a href="https://www.themoviedb.org/n">https://www.themoviedb.org/n</a>
2	2	Doctor Strange in the Multiverse of Madness	75.0	Fantasy, Action, Adventure	05/06/2022 (US)	2h 6m	Stan Lee	<a href="https://www.themoviedb.org/n">https://www.themoviedb.org/n</a>
3	3	The Bad Guys	77.0	Animation, Comedy, Action, Family, Crime	04/22/2022 (US)	1h 40m	Pierre Perifel	<a href="https://www.themoviedb.org/n">https://www.themoviedb.org/n</a>
4	4	Uncharted	72.0	Action, Adventure	02/18/2022 (US)	1h 56m	Ruben Fleischer	<a href="https://www.themoviedb.org/n">https://www.themoviedb.org/n</a>
...	...	...	...	...	...	...	...	...
3995	3995	Twister	64.0	Action, Adventure, Drama	05/10/1996 (US)	1h 53m	Jan de Bont	<a href="https://www.themoviedb.o">https://www.themoviedb.o</a>
3996	3996	Superman/Shazam!: The Return of Black Adam	73.0	Animation, Action, Science Fiction, Fantasy	11/16/2010 (ES)	25m	C.C. Beck	<a href="https://www.themoviedb.org/">https://www.themoviedb.org/</a>
3997	3997	27 Dresses	64.0	Comedy, Romance	01/18/2008 (US)	1h 51m	Anne Fletcher	<a href="https://www.themoviedb.org/">https://www.themoviedb.org/</a>
3998	3998	Tangled: Before Ever After	68.0	Adventure, Animation, Comedy, Family, Fantasy,...	10/26/2017 (DE)	55m	Tom Caulfield	<a href="https://www.themoviedb.org/n">https://www.themoviedb.org/n</a>
3999	3999	The Intruder	61.0	Thriller	05/03/2019 (US)	1h 42m	Deon Taylor	<a href="https://www.themoviedb.org/n">https://www.themoviedb.org/n</a>

4000 rows × 8 columns

```
In [4]: data.drop(['Unnamed: 0'],axis=1,inplace=True)
```

In [5]: data

Out[5]:

	Name	Ratings	Genre	Release date	Runtime	Director	Url
0	Sonic the Hedgehog 2	77.0	Action, Science Fiction, Comedy, Family, Adven...	04/08/2022 (US)	2h 2m	Josh Miller	<a href="https://www.themoviedb.org/movie/675353">https://www.themoviedb.org/movie/675353</a>
1	The Batman	78.0	Crime, Mystery, Thriller	03/04/2022 (IN)	2h 56m	Matt Reeves	<a href="https://www.themoviedb.org/movie/414906">https://www.themoviedb.org/movie/414906</a>
2	Doctor Strange in the Multiverse of Madness	75.0	Fantasy, Action, Adventure	05/06/2022 (US)	2h 6m	Stan Lee	<a href="https://www.themoviedb.org/movie/453395">https://www.themoviedb.org/movie/453395</a>
3	The Bad Guys	77.0	Animation, Comedy, Action, Family, Crime	04/22/2022 (US)	1h 40m	Pierre Perifel	<a href="https://www.themoviedb.org/movie/629542">https://www.themoviedb.org/movie/629542</a>
4	Uncharted	72.0	Action, Adventure	02/18/2022 (US)	1h 56m	Ruben Fleischer	<a href="https://www.themoviedb.org/movie/335787">https://www.themoviedb.org/movie/335787</a>
...	...	...	...	...	...	...	...
3995	Twister	64.0	Action, Adventure, Drama	05/10/1996 (US)	1h 53m	Jan de Bont	<a href="https://www.themoviedb.org/movie/664">https://www.themoviedb.org/movie/664</a>
3996	Superman/Shazam!: The Return of Black Adam	73.0	Animation, Action, Science Fiction, Fantasy	11/16/2010 (ES)	25m	C.C. Beck	<a href="https://www.themoviedb.org/movie/43641">https://www.themoviedb.org/movie/43641</a>
3997	27 Dresses	64.0	Comedy, Romance	01/18/2008 (US)	1h 51m	Anne Fletcher	<a href="https://www.themoviedb.org/movie/6557">https://www.themoviedb.org/movie/6557</a>
3998	Tangled: Before Ever After	68.0	Adventure, Animation, Comedy, Family, Fantasy,...	10/26/2017 (DE)	55m	Tom Caulfield	<a href="https://www.themoviedb.org/movie/438747">https://www.themoviedb.org/movie/438747</a>
3999	The Intruder	61.0	Thriller	05/03/2019 (US)	1h 42m	Deon Taylor	<a href="https://www.themoviedb.org/movie/524247">https://www.themoviedb.org/movie/524247</a>

4000 rows × 7 columns





```
In [6]: data['Release date']=data['Release date'].str[0:10]
data
```

Out[6]:

	Name	Ratings		Genre	Release date	Runtime	Director	Url
0	Sonic the Hedgehog 2	77.0	Action, Science Fiction, Comedy, Family, Adven...		04/08/2022	2h 2m	Josh Miller	<a href="https://www.themoviedb.org/movie/675353">https://www.themoviedb.org/movie/675353</a>
1	The Batman	78.0		Crime, Mystery, Thriller	03/04/2022	2h 56m	Matt Reeves	<a href="https://www.themoviedb.org/movie/414906">https://www.themoviedb.org/movie/414906</a>
2	Doctor Strange in the Multiverse of Madness	75.0		Fantasy, Action, Adventure	05/06/2022	2h 6m	Stan Lee	<a href="https://www.themoviedb.org/movie/453395">https://www.themoviedb.org/movie/453395</a>
3	The Bad Guys	77.0		Animation, Comedy, Action, Family, Crime	04/22/2022	1h 40m	Pierre Perifel	<a href="https://www.themoviedb.org/movie/629542">https://www.themoviedb.org/movie/629542</a>
4	Uncharted	72.0		Action, Adventure	02/18/2022	1h 56m	Ruben Fleischer	<a href="https://www.themoviedb.org/movie/335787">https://www.themoviedb.org/movie/335787</a>
...	...	...		...	...	...	...	...
3995	Twister	64.0		Action, Adventure, Drama	05/10/1996	1h 53m	Jan de Bont	<a href="https://www.themoviedb.org/movie/664">https://www.themoviedb.org/movie/664</a>
3996	Superman/Shazam!: The Return of Black Adam	73.0		Animation, Action, Science Fiction, Fantasy	11/16/2010	25m	C.C. Beck	<a href="https://www.themoviedb.org/movie/43641">https://www.themoviedb.org/movie/43641</a>
3997	27 Dresses	64.0		Comedy, Romance	01/18/2008	1h 51m	Anne Fletcher	<a href="https://www.themoviedb.org/movie/6557">https://www.themoviedb.org/movie/6557</a>
3998	Tangled: Before Ever After	68.0	Adventure, Animation, Comedy, Family, Fantasy,...		10/26/2017	55m	Tom Caulfield	<a href="https://www.themoviedb.org/movie/438747">https://www.themoviedb.org/movie/438747</a>
3999	The Intruder	61.0		Thriller	05/03/2019	1h 42m	Deon Taylor	<a href="https://www.themoviedb.org/movie/524247">https://www.themoviedb.org/movie/524247</a>

4000 rows × 7 columns



```
In [7]: data.columns = ['Name', 'Rating', 'Genres', 'Date', 'Runtime', 'Director', 'link']
data
```

Out[7]:

	Name	Rating	Genres	Date	Runtime	Director	link
0	Sonic the Hedgehog 2	77.0	Action, Science Fiction, Comedy, Family, Adven...	04/08/2022	2h 2m	Josh Miller	<a href="https://www.themoviedb.org/movie/675353">https://www.themoviedb.org/movie/675353</a>
1	The Batman	78.0	Crime, Mystery, Thriller	03/04/2022	2h 56m	Matt Reeves	<a href="https://www.themoviedb.org/movie/414906">https://www.themoviedb.org/movie/414906</a>
2	Doctor Strange in the Multiverse of Madness	75.0	Fantasy, Action, Adventure	05/06/2022	2h 6m	Stan Lee	<a href="https://www.themoviedb.org/movie/453395">https://www.themoviedb.org/movie/453395</a>
3	The Bad Guys	77.0	Animation, Comedy, Action, Family, Crime	04/22/2022	1h 40m	Pierre Perifel	<a href="https://www.themoviedb.org/movie/629542">https://www.themoviedb.org/movie/629542</a>
4	Uncharted	72.0	Action, Adventure	02/18/2022	1h 56m	Ruben Fleischer	<a href="https://www.themoviedb.org/movie/335787">https://www.themoviedb.org/movie/335787</a>
...	...	...	...	...	...	...	...
3995	Twister	64.0	Action, Adventure, Drama	05/10/1996	1h 53m	Jan de Bont	<a href="https://www.themoviedb.org/movie/664">https://www.themoviedb.org/movie/664</a>
3996	Superman/Shazam!: The Return of Black Adam	73.0	Animation, Action, Science Fiction, Fantasy	11/16/2010	25m	C.C. Beck	<a href="https://www.themoviedb.org/movie/43641">https://www.themoviedb.org/movie/43641</a>
3997	27 Dresses	64.0	Comedy, Romance	01/18/2008	1h 51m	Anne Fletcher	<a href="https://www.themoviedb.org/movie/6557">https://www.themoviedb.org/movie/6557</a>
3998	Tangled: Before Ever After	68.0	Adventure, Animation, Comedy, Family, Fantasy,...	10/26/2017	55m	Tom Caulfield	<a href="https://www.themoviedb.org/movie/438747">https://www.themoviedb.org/movie/438747</a>
3999	The Intruder	61.0	Thriller	05/03/2019	1h 42m	Deon Taylor	<a href="https://www.themoviedb.org/movie/524247">https://www.themoviedb.org/movie/524247</a>

4000 rows × 7 columns

```
In [8]: data.isnull().sum()
```

```
Out[8]: Name          0
        Rating        0
        Genres        24
        Date          0
        Runtime       73
        Director      37
        link           0
        dtype: int64
```

```
In [9]: data['Genres'].fillna("0", inplace = True)
```

```
In [10]: data['Runtime'].fillna("0", inplace = True)
```

```
In [11]: data['Director'].fillna("0", inplace = True)
```

```
In [12]: data.isnull().sum()
```

```
Out[12]: Name          0
        Rating        0
        Genres        0
        Date          0
        Runtime       0
        Director      0
        link          0
        dtype: int64
```

```
In [13]: data.max()
```

```
Out[13]: Name          어부바
        Rating          100.0
        Genres          Western, Drama, Adventure
        Date            12/31/2006
        Runtime          9m
        Director        薛凌
        link            https://www.themoviedb.org/movie/99942 (https://www.themoviedb.org/movie/99942)
        dtype: object
```

```
In [14]: data.min()
```

```
Out[14]: Name                #Alive  
Rating                0.0  
Genres                0  
Date                01/01/1981  
Runtime                0  
Director                0  
link    https://www.themoviedb.org/movie/100042 (https://www.themoviedb.org/movie/100042)  
dtype: object
```

```
In [15]: sum(data.duplicated())
```

```
Out[15]: 0
```

```
In [16]: data.columns
```

```
Out[16]: Index(['Name', 'Rating', 'Genres', 'Date', 'Runtime', 'Director', 'link'], dtype='object')
```

## calculating average number of rating

```
In [17]: data.groupby('Genres')['Rating'].agg(['count', 'mean'])
```

Out[17]:

	count	mean
Genres		
0	24	37.37500
Action	32	61.40625
Action, Adventure	8	67.00000
Action, Adventure, Animation	4	72.50000
Action, Adventure, Animation, Comedy, Family	2	68.00000
...	...	...
Western, Adventure, Action, Romance	1	74.00000
Western, Animation, Adventure, Comedy, Family, Drama	1	77.00000
Western, Comedy	1	52.00000
Western, Drama	1	62.00000
Western, Drama, Adventure	1	75.00000

1286 rows × 2 columns

```
In [18]: data.groupby(['Name'])['Rating', 'Genres', 'Date', 'Runtime', 'Director'].min()
```

C:\Users\Admin\AppData\Local\Temp\ipykernel\_26436\1553414436.py:1: FutureWarning: Indexing with multiple keys (implicitly converted to a tuple of keys) will be deprecated, use a list instead.

```
data.groupby(['Name'])['Rating', 'Genres', 'Date', 'Runtime', 'Director'].min()
```

Out[18]:

	Rating	Genres	Date	Runtime	Director
Name					
#Alive	73.0	Action, Horror, Thriller	06/24/2020	1h 38m	Cho Il
(500) Days of Summer	73.0	Comedy, Drama, Romance	08/07/2009	1h 35m	Marc Webb
10 Cloverfield Lane	70.0	Thriller, Science Fiction, Drama, Horror	03/11/2016	1h 44m	Josh Campbell
10 Things I Hate About You	75.0	Comedy, Romance, Drama	03/31/1999	1h 37m	Gil Junger
10,000 BC	54.0	Adventure, Action, Drama, Fantasy	03/05/2008	1h 49m	Roland Emmerich
...	...	...	...	...	...
Ženy a život	0.0	Comedy	05/12/2022	0	Petr Zahrádka
Ботан и супербаба	0.0	Comedy	05/12/2022	0	Dmitry Menyaylo
Тайна амулета	0.0	Family	05/12/2022	1h 34m	Elena Bychkova
灵契	30.0	Fantasy	03/11/2018	0	薛凌
어부바	0.0	Family	05/11/2022	0	Choi Jong-hak

3918 rows × 5 columns

```
In [19]: data.groupby(['Name'])['Rating', 'Genres', 'Date', 'Runtime', 'Director'].max()
```

C:\Users\Admin\AppData\Local\Temp\ipykernel\_26436\1984316264.py:1: FutureWarning: Indexing with multiple keys (implicitly converted to a tuple of keys) will be deprecated, use a list instead.

```
data.groupby(['Name'])['Rating', 'Genres', 'Date', 'Runtime', 'Director'].max()
```

Out[19]:

	Rating	Genres	Date	Runtime	Director
Name					
#Alive	73.0	Action, Horror, Thriller	06/24/2020	1h 38m	Cho Il
(500) Days of Summer	73.0	Comedy, Drama, Romance	08/07/2009	1h 35m	Marc Webb
10 Cloverfield Lane	70.0	Thriller, Science Fiction, Drama, Horror	03/11/2016	1h 44m	Josh Campbell
10 Things I Hate About You	75.0	Comedy, Romance, Drama	03/31/1999	1h 37m	Gil Junger
10,000 BC	54.0	Adventure, Action, Drama, Fantasy	03/05/2008	1h 49m	Roland Emmerich
...	...	...	...	...	...
Ženy a život	0.0	Comedy	05/12/2022	0	Petr Zahrádka
Ботан и супербаба	0.0	Comedy	05/12/2022	0	Dmitry Menyaylo
Тайна амулета	0.0	Family	05/12/2022	1h 34m	Elena Bychkova
灵契	30.0	Fantasy	03/11/2018	0	薛凌
어부바	0.0	Family	05/11/2022	0	Choi Jong-hak

3918 rows × 5 columns

```
In [20]: data.describe()
```

```
Out[20]:
```

	Rating
count	4000.000000
mean	64.561000
std	13.229192
min	0.000000
25%	60.000000
50%	66.000000
75%	72.000000
max	100.000000

```
In [21]: data['Name']
```

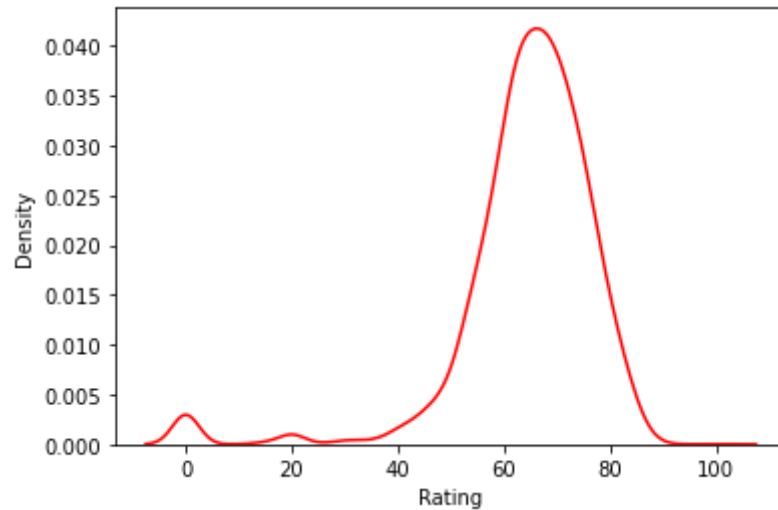
```
Out[21]: 0          Sonic the Hedgehog 2
1          The Batman
2    Doctor Strange in the Multiverse of Madness
3          The Bad Guys
4          Uncharted
...
3995          Twister
3996    Superman/Shazam!: The Return of Black Adam
3997          27 Dresses
3998    Tangled: Before Ever After
3999          The Intruder
Name: Name, Length: 4000, dtype: object
```



```
In [22]: sns.distplot(data['Rating'],hist=False,color='red')
```

C:\Users\Admin\anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `kdeplot` (an axes-level function for kernel density plots).  
warnings.warn(msg, FutureWarning)

```
Out[22]: <AxesSubplot:xlabel='Rating', ylabel='Density'>
```



```
In [ ]:
```

```
In [162]: df1=data.groupby(['Name'])[['Rating']].max()  
df1
```

Out[162]:

	Rating
Name	
#Alive	73.0
(500) Days of Summer	73.0
10 Cloverfield Lane	70.0
10 Things I Hate About You	75.0
10,000 BC	54.0
...	...
Ženy a život	0.0
Ботан и супербаба	0.0
Тайна амулета	0.0
灵契	30.0
어부바	0.0

3918 rows × 1 columns

```
In [163]: s1 = df1.sort_values(by='Rating',ascending=False).head()  
s1
```

```
Out[163]:
```

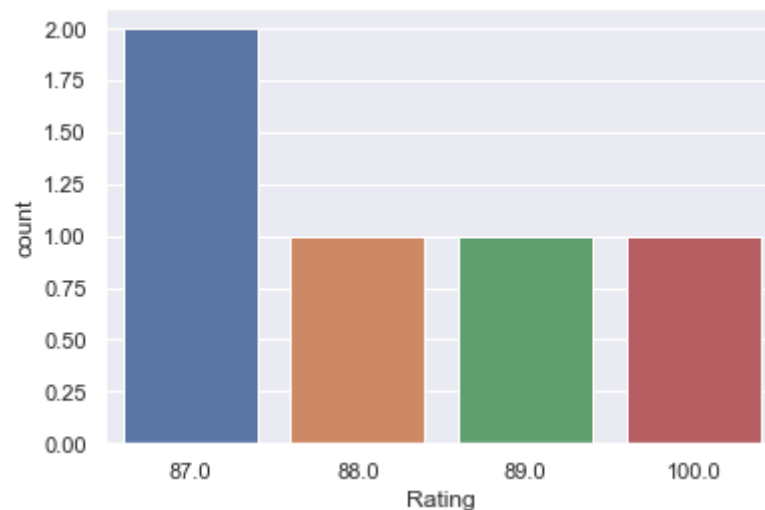
	Rating
Name	
Onsen porno chitai	100.0
Everything Everywhere All at Once	89.0
Demon Slayer: Kimetsu no Yaiba Sibling's Bond	88.0
Avatar Spirits	87.0
The Shawshank Redemption	87.0

```
In [165]: sns.countplot(s1['Rating'])
```

C:\Users\Admin\anaconda3\lib\site-packages\seaborn\\_decorators.py:36: FutureWarning:

Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
Out[165]: <AxesSubplot:xlabel='Rating', ylabel='count'>
```



```
In [169]: df2=data.groupby(['Director'])[['Runtime']].max()  
df2
```

Out[169]:

Runtime	
Director	
<hr/>	
0	5m
A.B. Stone	1h 32m
Aakash Bhatia	2h 11m
Aaron Fjellman	1h 21m
Aaron Horvath	1h 25m
...	...
Álvaro Curiel	0
Álvaro Fernández Armero	1h 42m
Ángel Gómez Hernández	1h 38m
Éric Toledano	1h 53m
薛凌	0

2616 rows × 1 columns

```
In [173]: s2=df2.sort_values(by='Runtime',ascending=False).head()  
s2
```

```
Out[173]:
```

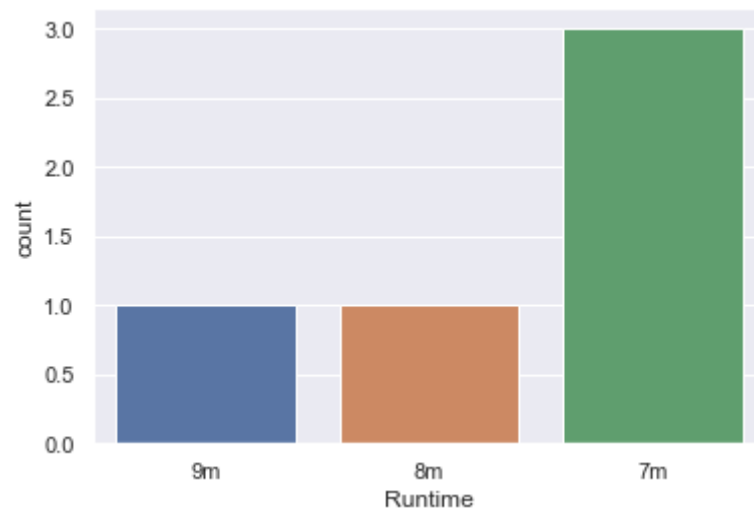
Runtime	
Director	
Kevin Nolting	9m
Jennifer Lee	8m
Lalo Valdivia	7m
Neill Blomkamp	7m
Natalie Nourigat	7m

```
In [174]: sns.countplot(s2['Runtime'])
```

C:\Users\Admin\anaconda3\lib\site-packages\seaborn\\_decorators.py:36: FutureWarning:

Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
Out[174]: <AxesSubplot:xlabel='Runtime', ylabel='count'>
```



```
In [191]: df4=data.groupby(['Director'])[['Rating']].max()  
df4
```

Out[191]:

Rating	
Director	
0	85.0
A.B. Stone	54.0
Aakash Bhatia	63.0
Aaron Fjellman	62.0
Aaron Horvath	74.0
...	...
Álvaro Curiel	76.0
Álvaro Fernández Armero	63.0
Ángel Gómez Hernández	65.0
Éric Toledano	83.0
薛凌	30.0

2616 rows × 1 columns

Type *Markdown* and LaTeX:  $\alpha^2$

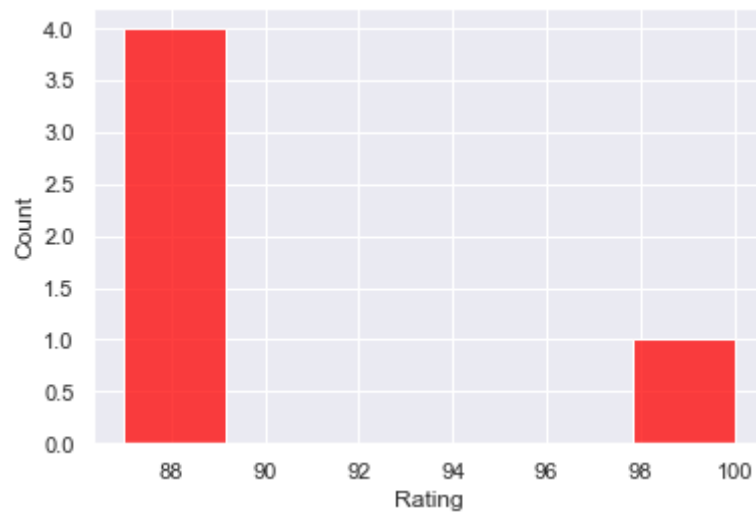
```
In [196]: s4=df4.sort_values(by='Rating',ascending=False).head()  
s4
```

```
Out[196]:
```

	Rating
Director	
Wataru Nakagawa	100.0
Daniel Kwan	89.0
Haruo Sotozaki	88.0
Francis Ford Coppola	87.0
Ernesto Contreras	87.0

```
In [197]: #sns.barplot(x = 'Name', y = 'Runtime',data = s4)  
#sns.relplot(x='Name',y='Runtime',data=s4,hue='size',palette='brg')  
sns.histplot(data=s4,x='Rating',color='red')
```

```
Out[197]: <AxesSubplot:xlabel='Rating', ylabel='Count'>
```



```
In [200]: df5=data.groupby(['Director'])[['Rating']].max()  
df5
```

Out[200]:

	Rating
Director	
0	85.0
A.B. Stone	54.0
Aakash Bhatia	63.0
Aaron Fjellman	62.0
Aaron Horvath	74.0
...	...
Álvaro Curiel	76.0
Álvaro Fernández Armero	63.0
Ángel Gómez Hernández	65.0
Éric Toledano	83.0
薛凌	30.0

2616 rows × 1 columns



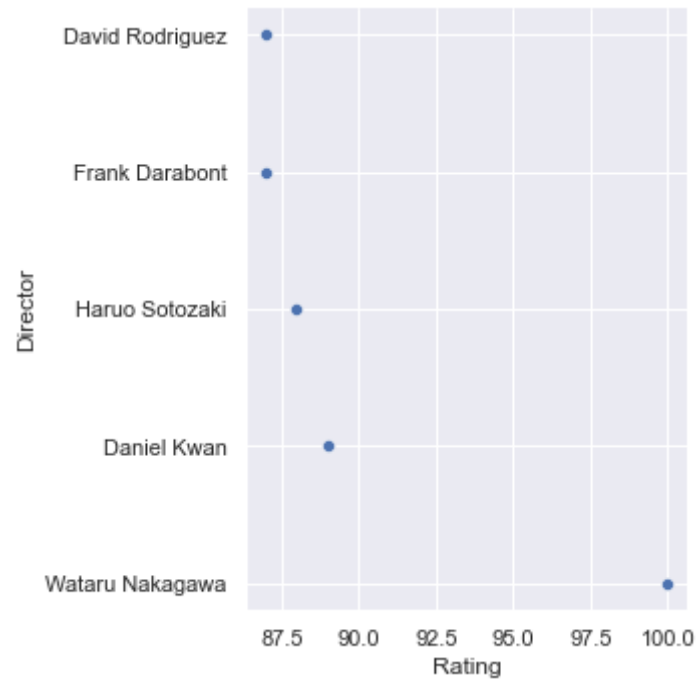
```
In [201]: s5=df5.sort_values(by='Rating',ascending=True).tail()  
s5
```

Out[201]:

	Rating
Director	
David Rodriguez	87.0
Frank Darabont	87.0
Haruo Sotozaki	88.0
Daniel Kwan	89.0
Wataru Nakagawa	100.0

```
In [202]: plt.figure(figsize=(25,25))  
sns.relplot(x='Rating',y='Director',data=s5)  
plt.show()
```

<Figure size 1800x1800 with 0 Axes>



```
In [ ]: data.columns
```

```
In [203]: df6=data.groupby(['Date'])[['Genres']].max()  
df6
```

Out[203]:

Genres	
Date	
1932-04-09	Action, Crime, Drama, Thriller
1939-08-25	Adventure, Fantasy, Family
1940-02-23	Animation, Family
1940-04-12	Drama
1940-06-17	Drama, War, Romance
...	...
2022-08-06	Animation, Adventure, Action, Fantasy
2022-08-25	Romance
2022-10-21	Action, Fantasy, Adventure
2022-11-02	0
2022-11-11	Action, Adventure, Science Fiction

2509 rows × 1 columns

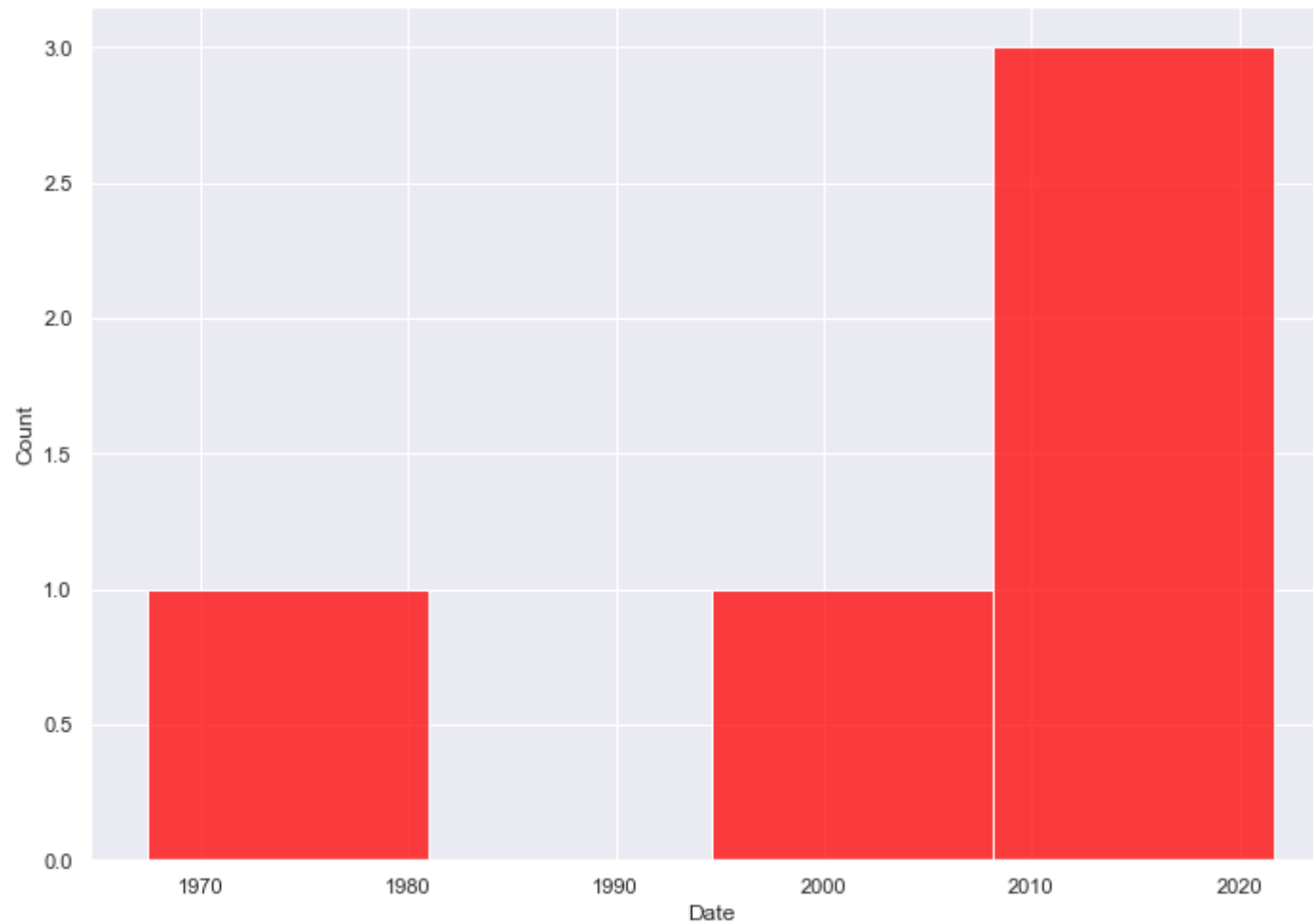
```
In [210]: s6=df6.sort_values(by='Genres',ascending=True).tail()  
s6
```

Out[210]:

Genres	
Date	
1967-06-07	Western, Adventure, Action, Romance
2002-05-24	Western, Animation, Adventure, Comedy, Family,...
2015-12-11	Western, Comedy
2021-09-17	Western, Drama
2016-01-29	Western, Drama, Adventure

```
In [211]: sns.set(rc={'figure.figsize':(11.7,8.27)})  
sns.histplot(data=s6,x='Date',color='red')
```

```
Out[211]: <AxesSubplot:xlabel='Date', ylabel='Count'>
```



In [ ]:

```
In [137]: data['Date'] = pd.to_datetime(data['Date'])
data['Date']
```

```
Out[137]: 0      2022-04-08
1      2022-03-04
2      2022-05-06
3      2022-04-22
4      2022-02-18
...
3995   1996-05-10
3996   2010-11-16
3997   2008-01-18
3998   2017-10-26
3999   2019-05-03
Name: Date, Length: 4000, dtype: datetime64[ns]
```

```
In [138]: data['year'] = data['Date'].dt.year  
y_data = data['year'].value_counts().to_frame()  
y_data  
#data['year'] = data['Date'].dt.year  
#r_data = data['Rating'].value_counts().to_frame()  
#r_data
```

Out[138]:

	year
<b>2021</b>	543
<b>2022</b>	271
<b>2017</b>	215
<b>2020</b>	214
<b>2019</b>	213
...	...
<b>1956</b>	1
<b>1959</b>	1
<b>1961</b>	1
<b>1970</b>	1
<b>1949</b>	1

78 rows × 1 columns

```
In [143]: def total_count(column):
            total_count = data[column].str.cat(sep = ',')
            total_count = pd.Series(total_count.split(','))
            total_v = total_count.value_counts(ascending = False).rename_axis('Name').reset_index(name='Counts')
            return total_v
total_v = total_count('Genres')
total_v
```

Out[143]:

	Name	Counts
0	Thriller	817
1	Action	713
2	Drama	676
3	Adventure	645
4	Comedy	621
5	Action	572
6	Fantasy	541
7	Drama	540
8	Family	502
9	Comedy	484
10	Science Fiction	455
11	Horror	444
12	Animation	427
13	Romance	365
14	Animation	345
15	Crime	316
16	Mystery	277
17	Adventure	268
18	Horror	235
19	Thriller	201



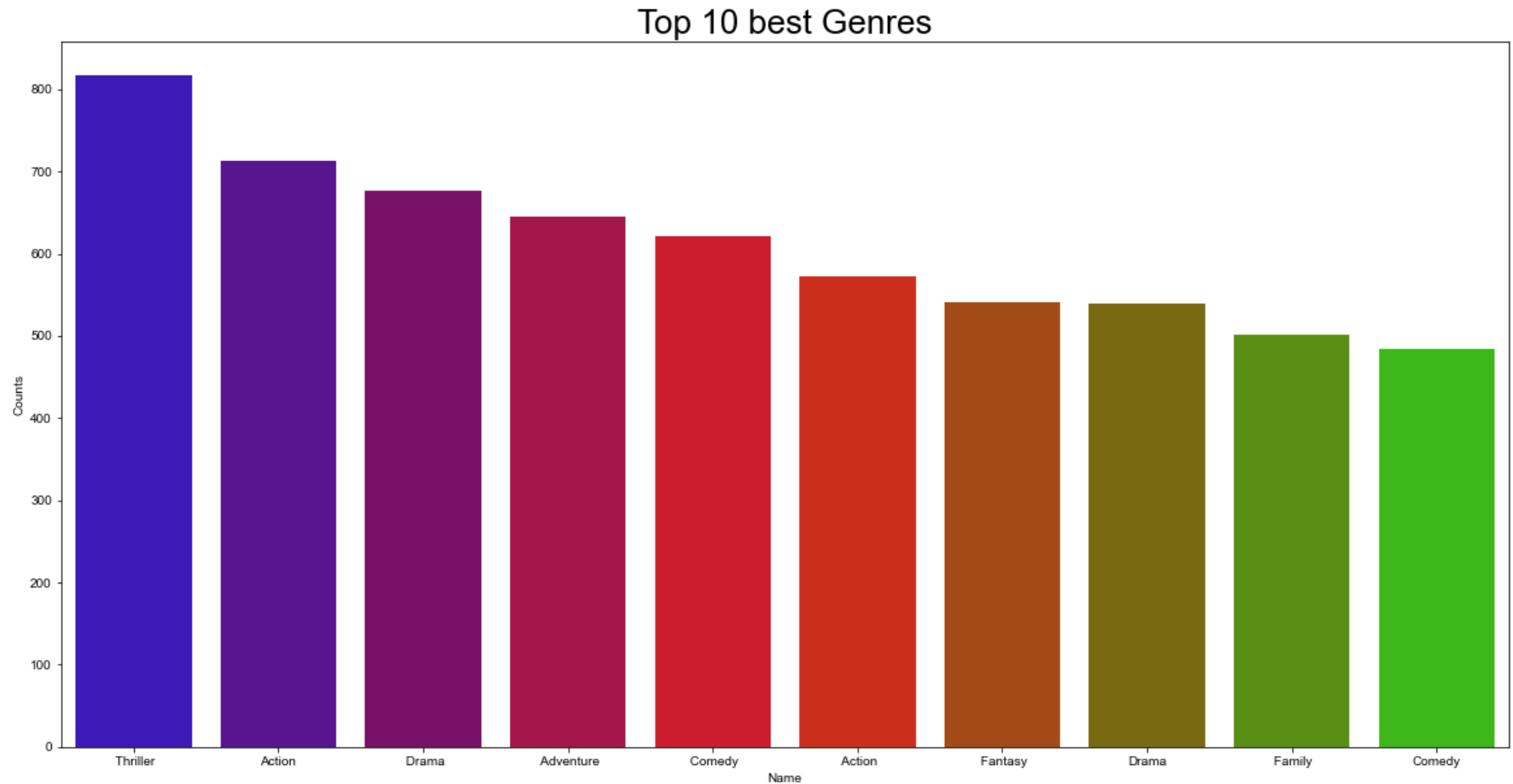
	<b>Name</b>	<b>Counts</b>
<b>20</b>	Family	162
<b>21</b>	Science Fiction	159
<b>22</b>	Crime	128
<b>23</b>	Romance	123
<b>24</b>	Fantasy	123
<b>25</b>	History	112
<b>26</b>	Documentary	81
<b>27</b>	Music	79
<b>28</b>	War	66
<b>29</b>	TV Movie	57
<b>30</b>	Mystery	40
<b>31</b>	War	29
<b>32</b>	Western	24
<b>33</b>	Music	24
<b>34</b>	0	24
<b>35</b>	Documentary	13
<b>36</b>	TV Movie	12
<b>37</b>	Western	10
<b>38</b>	History	8

```
In [144]: populargenres = total_v.head(10)
populargenres
```

```
Out[144]:
```

	<b>Name</b>	<b>Counts</b>
<b>0</b>	Thriller	817
<b>1</b>	Action	713
<b>2</b>	Drama	676
<b>3</b>	Adventure	645
<b>4</b>	Comedy	621
<b>5</b>	Action	572
<b>6</b>	Fantasy	541
<b>7</b>	Drama	540
<b>8</b>	Family	502
<b>9</b>	Comedy	484

```
In [145]: plt.figure(figsize=(20,10))  
plt.title('Top 10 best Genres',size=26)  
  
sns.barplot(data=populargenres, x='Name', y='Counts', palette='brg')  
sns.set()  
plt.show()
```



In this senario thriller movie get highest counting

```
In [33]: data['Date']= pd.to_datetime(data['Date'])
```

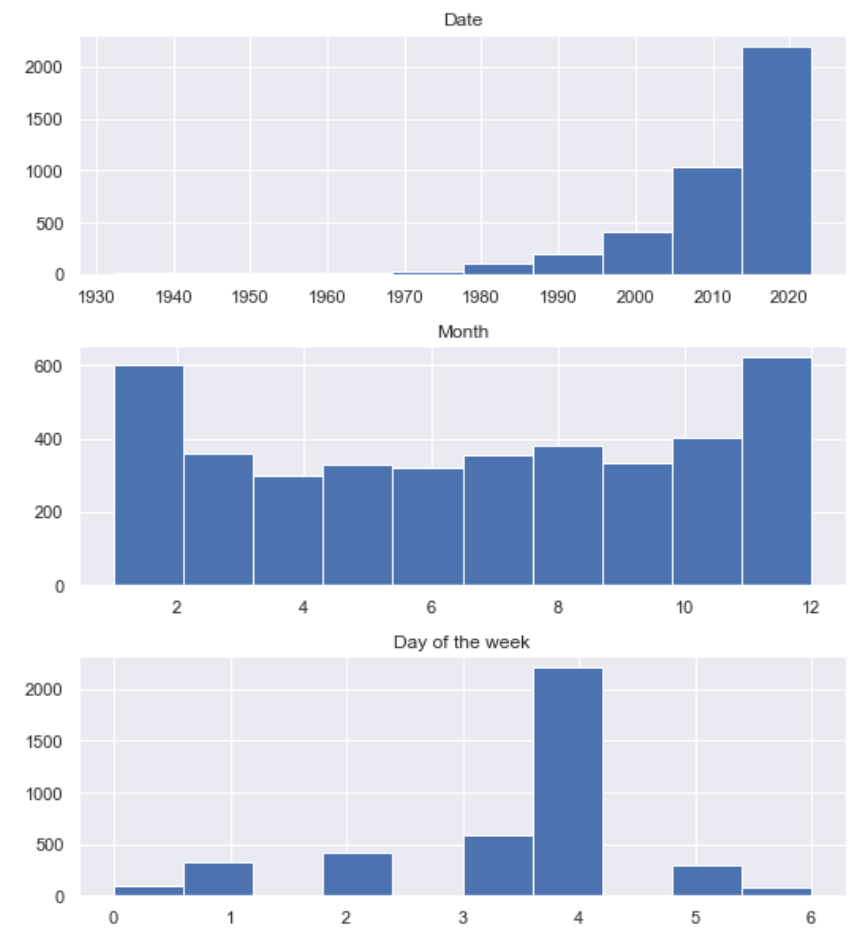
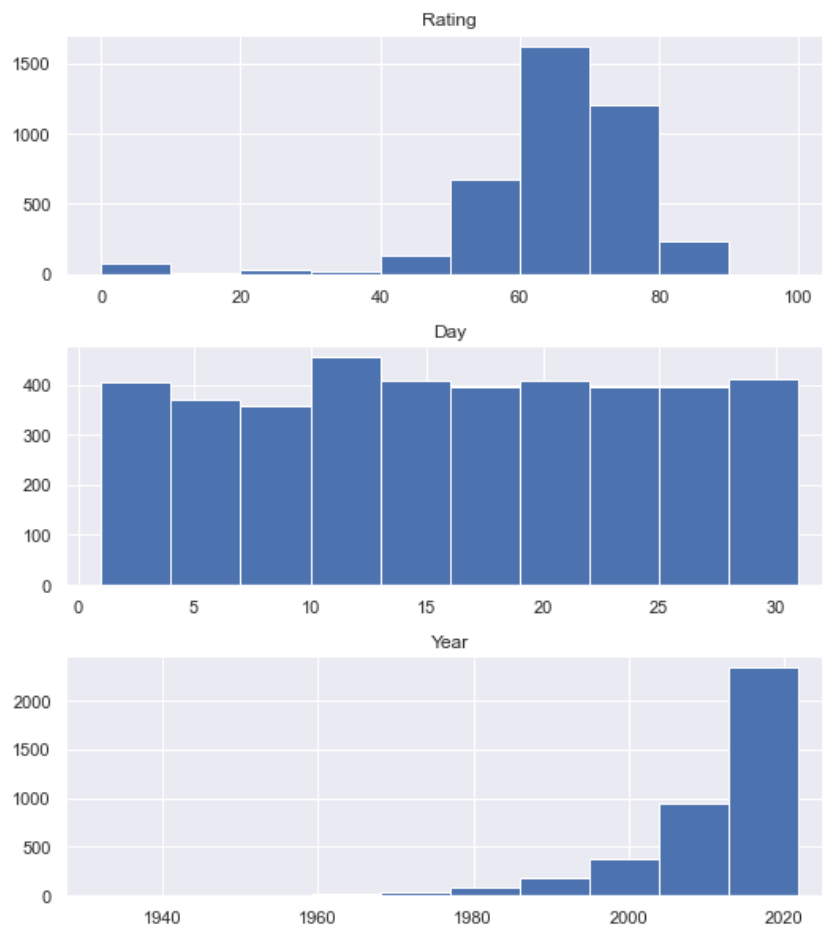
```
In [34]: data['Day'] = data['Date'].dt.day
data['Month'] = data['Date'].dt.month
data['Year'] = data['Date'].dt.year
data['Day of the week'] = data['Date'].dt.weekday #It uses a datetime class object and returns the day of the week as an data
```

Out[34]:

	Name	Rating	Genres	Date	Runtime	Director	link	Day
0	Sonic the Hedgehog 2	77.0	Action, Science Fiction, Comedy, Family, Adven...	2022-04-08	2h 2m	Josh Miller	<a href="https://www.themoviedb.org/movie/675353">https://www.themoviedb.org/movie/675353</a>	8
1	The Batman	78.0	Crime, Mystery, Thriller	2022-03-04	2h 56m	Matt Reeves	<a href="https://www.themoviedb.org/movie/414906">https://www.themoviedb.org/movie/414906</a>	4
2	Doctor Strange in the Multiverse of Madness	75.0	Fantasy, Action, Adventure	2022-05-06	2h 6m	Stan Lee	<a href="https://www.themoviedb.org/movie/453395">https://www.themoviedb.org/movie/453395</a>	6
3	The Bad Guys	77.0	Animation, Comedy, Action, Family, Crime	2022-04-22	1h 40m	Pierre Perifel	<a href="https://www.themoviedb.org/movie/629542">https://www.themoviedb.org/movie/629542</a>	22
4	Uncharted	72.0	Action, Adventure	2022-02-18	1h 56m	Ruben Fleischer	<a href="https://www.themoviedb.org/movie/335787">https://www.themoviedb.org/movie/335787</a>	18
...	...	...	...	...	...	...	...	...
3995	Twister	64.0	Action, Adventure, Drama	1996-05-10	1h 53m	Jan de Bont	<a href="https://www.themoviedb.org/movie/664">https://www.themoviedb.org/movie/664</a>	10
3996	Superman/Shazam!: The Return of Black Adam	73.0	Animation, Action, Science Fiction, Fantasy	2010-11-16	25m	C.C. Beck	<a href="https://www.themoviedb.org/movie/43641">https://www.themoviedb.org/movie/43641</a>	16
3997	27 Dresses	64.0	Comedy, Romance	2008-01-18	1h 51m	Anne Fletcher	<a href="https://www.themoviedb.org/movie/6557">https://www.themoviedb.org/movie/6557</a>	18
3998	Tangled: Before Ever After	68.0	Adventure, Animation, Comedy, Family, Fantasy,...	2017-10-26	55m	Tom Caulfield	<a href="https://www.themoviedb.org/movie/438747">https://www.themoviedb.org/movie/438747</a>	26
3999	The Intruder	61.0	Thriller	2019-05-03	1h 42m	Deon Taylor	<a href="https://www.themoviedb.org/movie/524247">https://www.themoviedb.org/movie/524247</a>	3

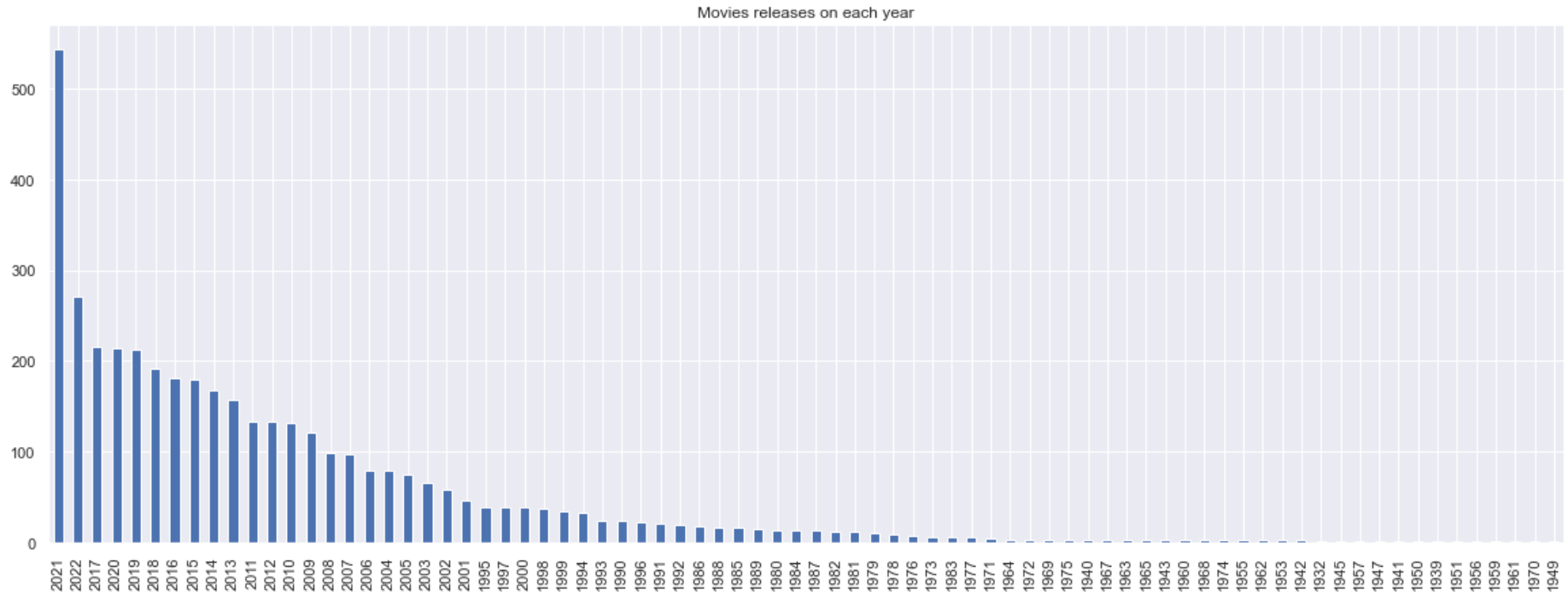
4000 rows × 11 columns

```
In [38]: data.hist(figsize = (20,10));
```



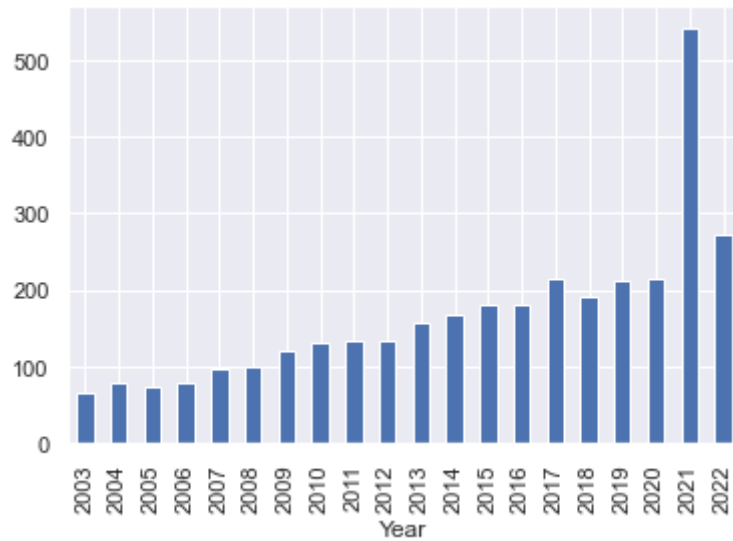
```
In [222]: data['Year'].value_counts().plot(x='Year', kind='bar', figsize=(20,7), title='Movies releases on each year')
```

```
Out[222]: <AxesSubplot:title={'center': 'Movies releases on each year'}>
```



```
In [49]: data[data.Year >= 2003].groupby('Year').size().plot(kind = 'bar')
```

```
Out[49]: <AxesSubplot: xlabel='Year'>
```



Here we can see in 2021 movies produced on large scale



```
In [35]: Monthly_data= data.groupby(['Month', 'Name'])[['Date', "Rating"]].sum().sort_values(by=['Month']).reset_index()  
Monthly_data
```

Out[35]:

	Month	Name	Rating
0	1	12 Strong	62.0
1	1	Riverdance: The Animated Adventure	53.0
2	1	Room	81.0
3	1	Rumble in the Bronx	68.0
4	1	Run Gun	51.0
...	...	...	...
3985	12	Hook	68.0
3986	12	Home Alone 3	53.0
3987	12	Hilda and the Mountain King	76.0
3988	12	Hunter x Hunter: The Last Mission	71.0
3989	12	ariana grande: excuse me, i love you	84.0

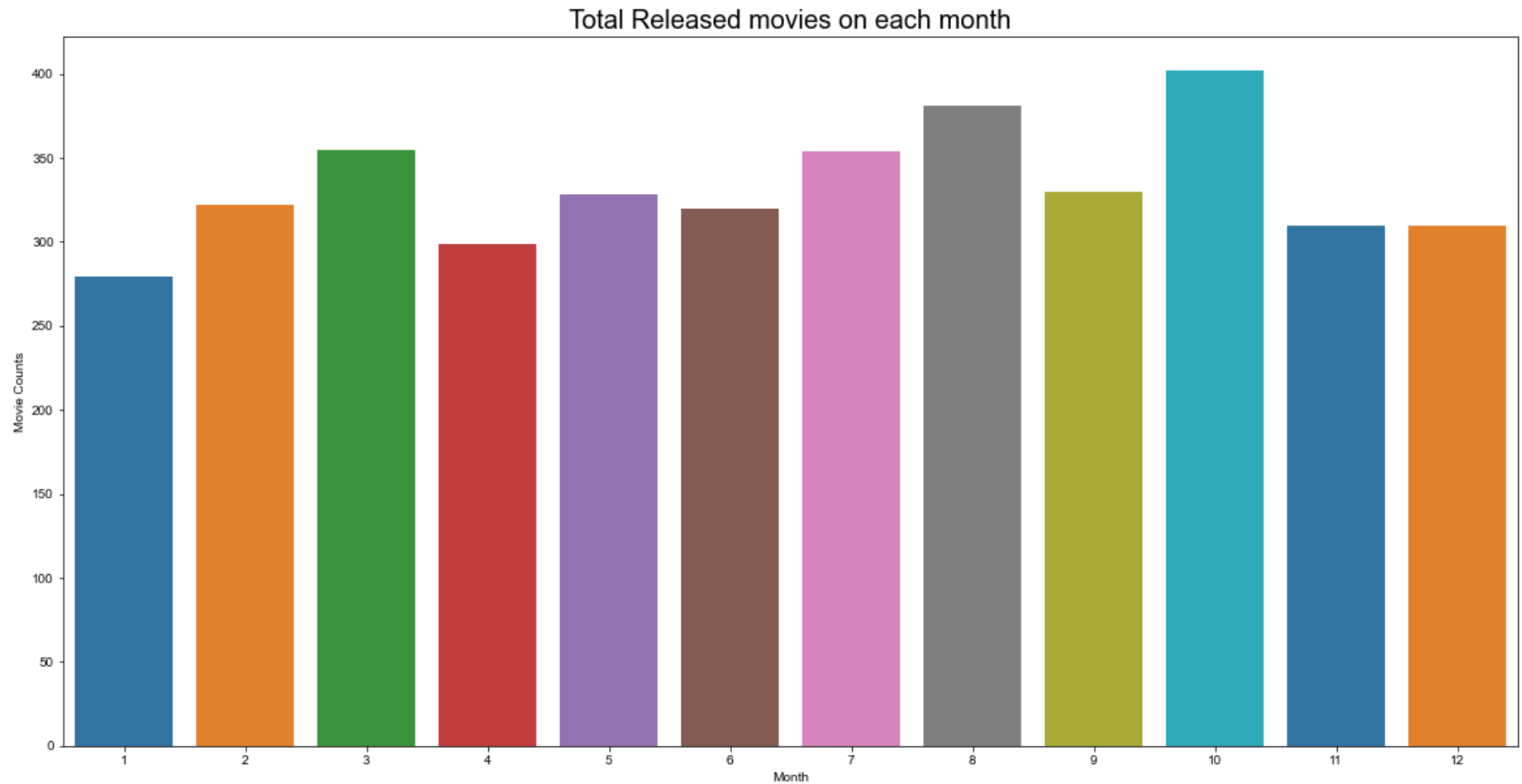
3990 rows × 3 columns

```
In [36]: Month_count =Monthly_data['Month'].value_counts()  
Month_count =Month_count.rename_axis('Month').reset_index(name='Movie Counts')  
Month_count
```

Out[36]:

	Month	Movie Counts
0	10	402
1	8	381
2	3	355
3	7	354
4	9	330
5	5	328
6	2	322
7	6	320
8	11	310
9	12	310
10	4	299
11	1	279

```
In [37]: plt.figure(figsize=(20,10))  
plt.title('Total Released movies on each month',size=20)  
  
sns.barplot(data=Month_count, x= 'Month', y='Movie Counts', palette='tab10')  
sns.set()  
plt.show()
```

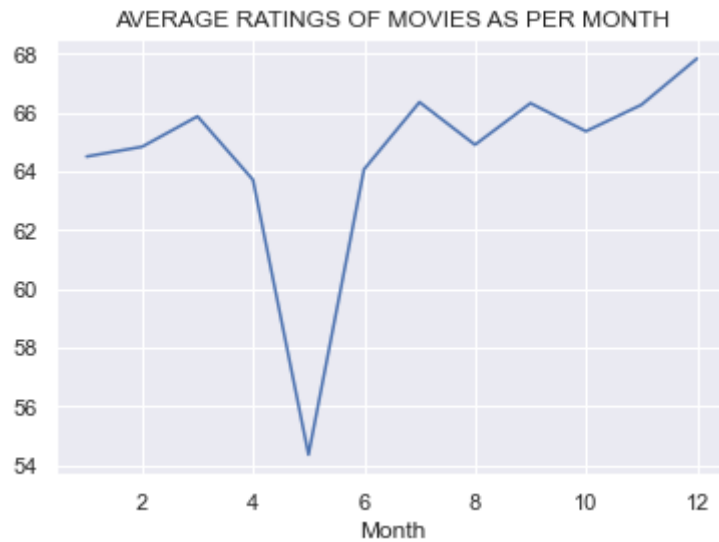


```
In [40]: data.groupby('Month')['Rating'].mean()
```

```
Out[40]: Month
1      64.508961
2      64.841615
3      65.879552
4      63.713333
5      54.354545
6      64.071651
7      66.361582
8      64.910761
9      66.326284
10     65.369727
11     66.272436
12     67.845161
Name: Rating, dtype: float64
```

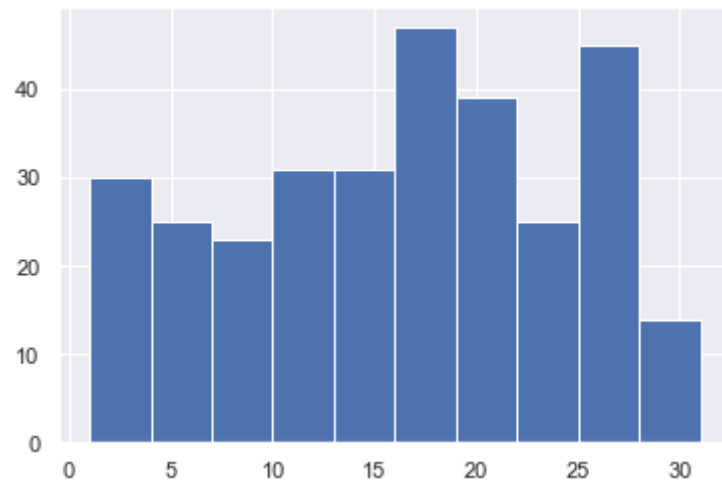
```
In [41]: data.groupby('Month')['Rating'].mean().plot(title = 'AVERAGE RATINGS OF MOVIES AS PER MONTH')
```

```
Out[41]: <AxesSubplot:title={'center': 'AVERAGE RATINGS OF MOVIES AS PER MONTH'}, xlabel='Month'>
```



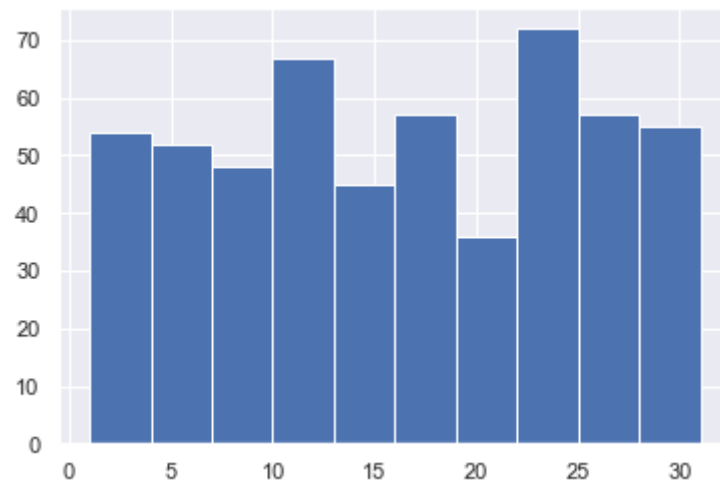
```
In [42]: data[data.Month == 12].Day.hist()
```

Out[42]: <AxesSubplot:>



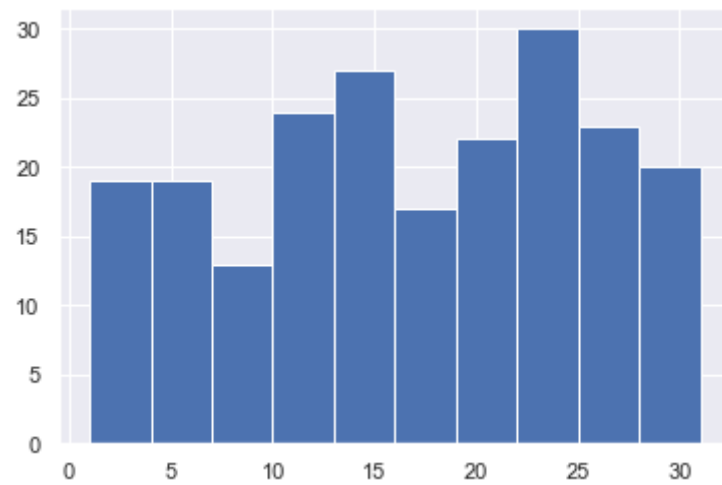
```
In [43]: data[data.Year == 2021].Day.hist()
```

```
Out[43]: <AxesSubplot:>
```



```
In [44]: data[data.Year == 2020].Day.hist()
```

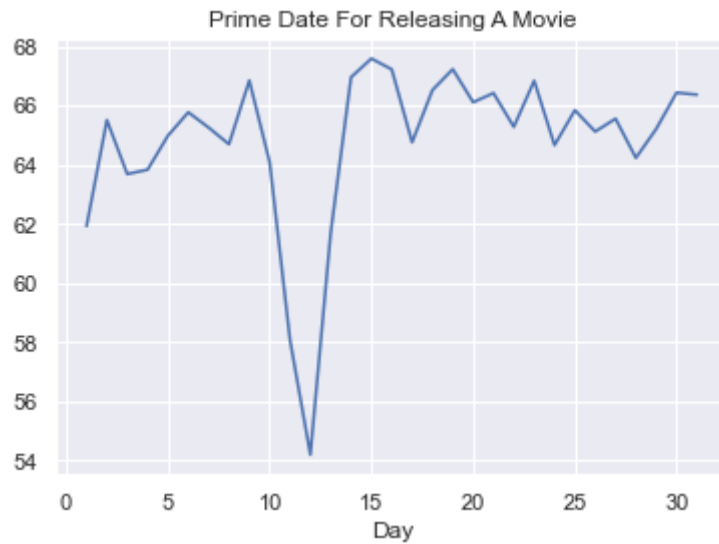
```
Out[44]: <AxesSubplot:>
```





```
In [45]: data.groupby('Day')['Rating'].mean().plot(title = 'Prime Date For Releasing A Movie')
```

```
Out[45]: <AxesSubplot:title={'center':'Prime Date For Releasing A Movie'}, xlabel='Day'>
```



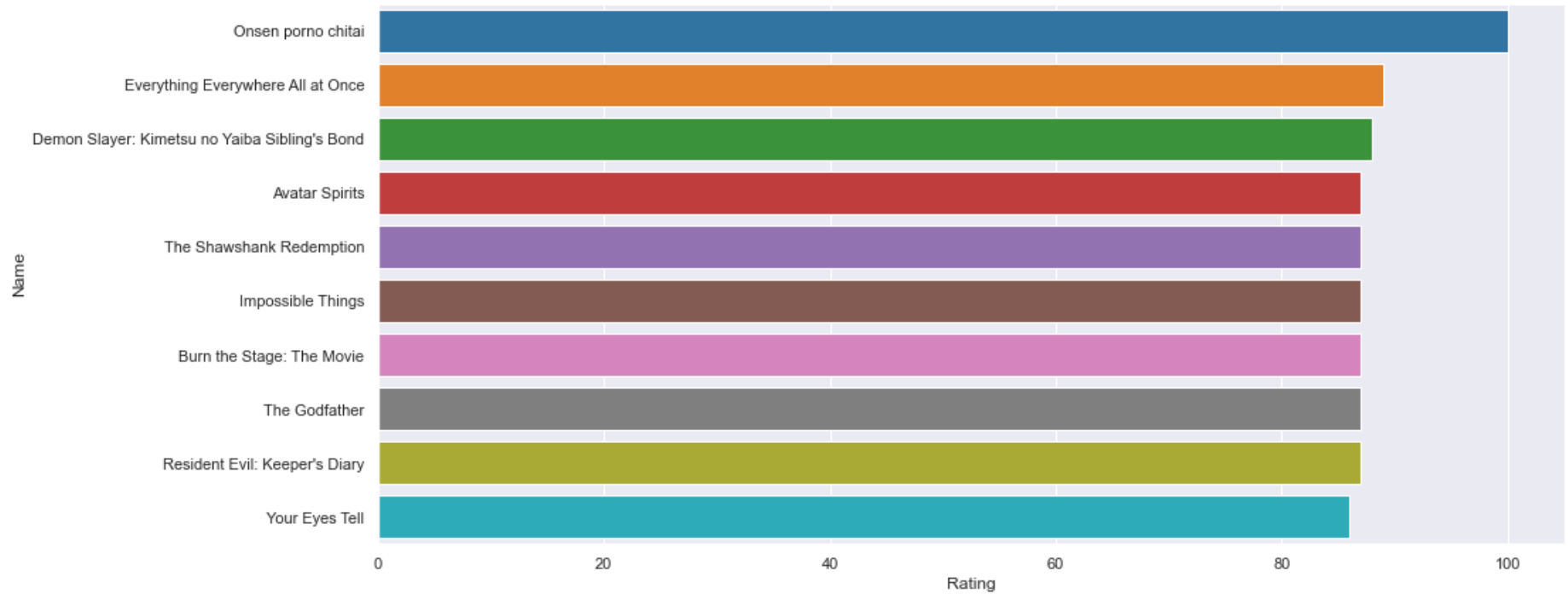
**Top movies mostly in theator with highest Rating**

```
In [50]: Fav_movies=data.sort_values(ascending= False, by=["Rating"])
top_10_movies= Fav_movies.head(10)#listing top 10 movies between them.
top_10_movies
```

Out[50]:

	Name	Rating	Genres	Date	Runtime	Director	link	Day	Month	Year	Day of the week
3776	Onsen porno chitai	100.0		1972-01-08	1h 4m	Wataru Nakagawa	<a href="https://www.themoviedb.org/movie/946030">https://www.themoviedb.org/movie/946030</a>	8	1	1972	5
570	Everything Everywhere All at Once	89.0	Action, Science Fiction, Comedy, Adventure	2022-04-08	2h 19m	Daniel Kwan	<a href="https://www.themoviedb.org/movie/545611">https://www.themoviedb.org/movie/545611</a>	8	4	2022	4
833	Demon Slayer: Kimetsu no Yaiba Sibling's Bond	88.0	Action, Animation, Fantasy	2019-03-29	1h 45m	Haruo Sotozaki	<a href="https://www.themoviedb.org/movie/820232">https://www.themoviedb.org/movie/820232</a>	29	3	2019	4
2923	Avatar Spirits	87.0	Documentary	2010-06-22	32m	Kurt Mattila	<a href="https://www.themoviedb.org/movie/278698">https://www.themoviedb.org/movie/278698</a>	22	6	2010	1
785	The Shawshank Redemption	87.0	Drama, Crime	1994-10-14	2h 22m	Frank Darabont	<a href="https://www.themoviedb.org/movie/278">https://www.themoviedb.org/movie/278</a>	14	10	1994	4
2839	Impossible Things	87.0	Family, Drama	2021-06-17	1h 29m	Ernesto Contreras	<a href="https://www.themoviedb.org/movie/667257">https://www.themoviedb.org/movie/667257</a>	17	6	2021	3
811	Burn the Stage: The Movie	87.0	Documentary, Music	2018-11-15	1h 25m	Park Jun-soo	<a href="https://www.themoviedb.org/movie/553512">https://www.themoviedb.org/movie/553512</a>	15	11	2018	3
851	The Godfather	87.0	Drama, Crime	1972-03-14	2h 55m	Francis Ford Coppola	<a href="https://www.themoviedb.org/movie/238">https://www.themoviedb.org/movie/238</a>	14	3	1972	1
3318	Resident Evil: Keeper's Diary	87.0	Horror	2014-01-06	6m	David Rodriguez	<a href="https://www.themoviedb.org/movie/471666">https://www.themoviedb.org/movie/471666</a>	6	1	2014	0
952	Your Eyes Tell	86.0	Romance, Drama	2020-10-23	2h 3m	Takahiro Miki	<a href="https://www.themoviedb.org/movie/730154">https://www.themoviedb.org/movie/730154</a>	23	10	2020	4

```
In [52]: fig = plt.figure(figsize=(15,7))  
sns.barplot(data=top_10_movies,x= 'Rating', y ='Name', palette='tab10')  
plt.show()
```



**Most Favourite movie in last two decade**

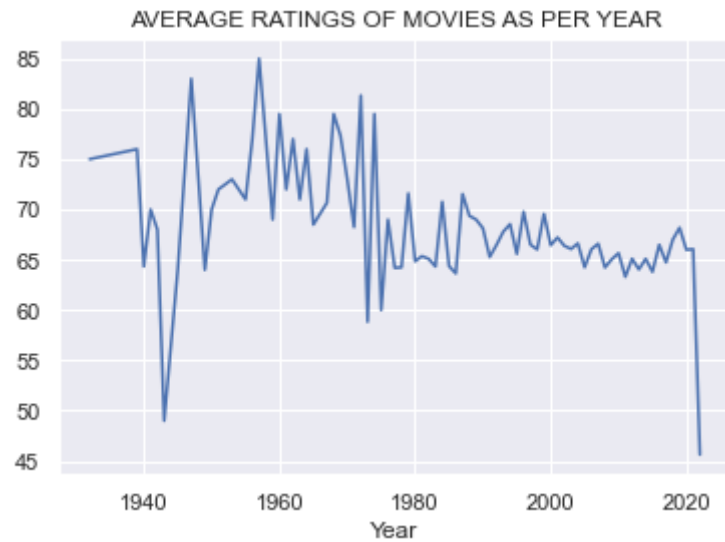
```
In [53]: Fav_movies[Fav_movies.Year >= 2003 ].sort_values(ascending= False, by=["Rating"])
```

	Name	Rating	Genre	Date	Runtime	Director	Link	Box Office	
3147	My Hero Academia - Save! Rescue Training	0.0		0	2017-04-04	26m	0	<a href="https://www.themoviedb.org/movie/969003">https://www.themoviedb.org/movie/969003</a>	4
1163	Miraculous Ladybug & Cat Noir: The Movie	0.0	Action, Animation, Adventure, Fantasy, Music	2022-08-03	0	Jeremy Zag	<a href="https://www.themoviedb.org/movie/496450">https://www.themoviedb.org/movie/496450</a>	3	
309	Thor: Love and Thunder	0.0	Action, Adventure, Fantasy	2022-07-08	0	Taika Waititi	<a href="https://www.themoviedb.org/movie/616037">https://www.themoviedb.org/movie/616037</a>	8	
2799	Bullet Train	0.0	Action, Thriller, Mystery	2022-08-05	0	David Leitch	<a href="https://www.themoviedb.org/movie/718930">https://www.themoviedb.org/movie/718930</a>	5	

**Best movie decade**

```
In [54]: Fav_movies.groupby('Year')['Rating'].mean().plot(title = 'AVERAGE RATINGS OF MOVIES AS PER YEAR')
```

```
Out[54]: <AxesSubplot:title={'center':'AVERAGE RATINGS OF MOVIES AS PER YEAR'}, xlabel='Year'>
```



**Highest movies with Ratin and Runtime**

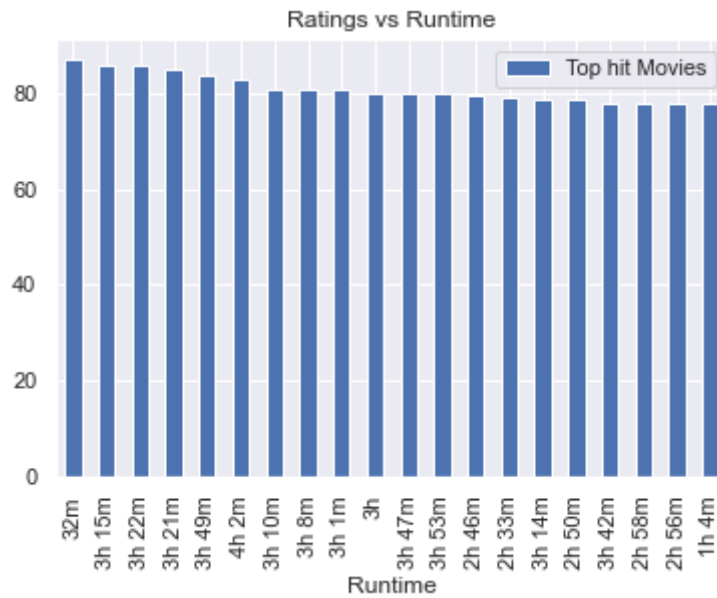
```
In [60]: Rt=data['Runtime'].value_counts().rename_axis('Common Runtime').reset_index(name='counts')
Rt
```

Out[60]:

	Common Runtime	counts
0	1h 30m	146
1	1h 40m	110
2	1h 34m	105
3	1h 37m	99
4	1h 35m	98
...	...	...
182	1m	1
183	27m	1
184	21m	1
185	3h 3m	1
186	3h 53m	1

187 rows × 2 columns

```
In [62]: rtc=data.groupby(['Runtime'])['Rating'].mean().sort_values(ascending=False).reset_index().head(20)
rtc.plot(x='Runtime', y='Rating', kind='bar')
plt.title("Ratings vs Runtime")
plt.legend(['Top hit Movies']);
```



**Mostly Audience watch to love movies**

```
In [64]: m_genres=data.copy()
m_genres.head(10)
```

Out[64]:

	Name	Rating	Genres	Date	Runtime	Director	link	Day	Month	Year	w
0	Sonic the Hedgehog 2	77.0	Action, Science Fiction, Comedy, Family, Adven...	2022-04-08	2h 2m	Josh Miller	<a href="https://www.themoviedb.org/movie/675353">https://www.themoviedb.org/movie/675353</a>	8	4	2022	
1	The Batman	78.0	Crime, Mystery, Thriller	2022-03-04	2h 56m	Matt Reeves	<a href="https://www.themoviedb.org/movie/414906">https://www.themoviedb.org/movie/414906</a>	4	3	2022	
2	Doctor Strange in the Multiverse of Madness	75.0	Fantasy, Action, Adventure	2022-05-06	2h 6m	Stan Lee	<a href="https://www.themoviedb.org/movie/453395">https://www.themoviedb.org/movie/453395</a>	6	5	2022	
3	The Bad Guys	77.0	Animation, Comedy, Action, Family, Crime	2022-04-22	1h 40m	Pierre Perifel	<a href="https://www.themoviedb.org/movie/629542">https://www.themoviedb.org/movie/629542</a>	22	4	2022	
4	Uncharted	72.0	Action, Adventure	2022-02-18	1h 56m	Ruben Fleischer	<a href="https://www.themoviedb.org/movie/335787">https://www.themoviedb.org/movie/335787</a>	18	2	2022	
5	Spider-Man: No Way Home	81.0	Action, Adventure, Science Fiction	2021-12-17	2h 28m	Steve Ditko	<a href="https://www.themoviedb.org/movie/634649">https://www.themoviedb.org/movie/634649</a>	17	12	2021	
6	The eighth clause	41.0	Thriller	2022-04-29	0	Frank Ariza	<a href="https://www.themoviedb.org/movie/956101">https://www.themoviedb.org/movie/956101</a>	29	4	2022	
7	Ambulance	70.0	Action, Thriller, Crime	2022-04-08	2h 16m	Michael Bay	<a href="https://www.themoviedb.org/movie/763285">https://www.themoviedb.org/movie/763285</a>	8	4	2022	
8	Turning Red	74.0	Animation, Family, Comedy, Fantasy	2022-03-10	1h 40m	Domee Shi	<a href="https://www.themoviedb.org/movie/508947">https://www.themoviedb.org/movie/508947</a>	10	3	2022	
9	Doctor Strange	74.0	Action, Adventure, Fantasy, Science Fiction	2016-11-04	1h 55m	Scott Derrickson	<a href="https://www.themoviedb.org/movie/284052">https://www.themoviedb.org/movie/284052</a>	4	11	2016	



```
In [65]: gnrs=m_genres.Genres.str.split(",", expand=True)
gnrs.rename(columns={0:'Genre 1', 1:'Genre 2', 2:"Genre 3", 3:'Genre 4', 4:"Genre 5" , 5:"Genre 6" , 6:"Genre 7"}, inplace=True)
gnrs.head(10)
```

Out[65]:

	Genre 1	Genre 2	Genre 3	Genre 4	Genre 5	Genre 6	Genre 7	7
0	Action	Science Fiction	Comedy	Family	Adventure	None	None	None
1	Crime	Mystery	Thriller	None	None	None	None	None
2	Fantasy	Action	Adventure	None	None	None	None	None
3	Animation	Comedy	Action	Family	Crime	None	None	None
4	Action	Adventure	None	None	None	None	None	None
5	Action	Adventure	Science Fiction	None	None	None	None	None
6	Thriller	None	None	None	None	None	None	None
7	Action	Thriller	Crime	None	None	None	None	None
8	Animation	Family	Comedy	Fantasy	None	None	None	None
9	Action	Adventure	Fantasy	Science Fiction	None	None	None	None

```
In [66]: m_genres=pd.merge(m_genres, gnrs, left_index=True, right_index=True, how="inner")
m_genres.head(5)
```

Out[66]:

	Name	Rating	Genres	Date	Runtime	Director	link	Day	Month	Year	Date th week
0	Sonic the Hedgehog 2	77.0	Action, Science Fiction, Comedy, Family, Adven...	2022-04-08	2h 2m	Josh Miller	<a href="https://www.themoviedb.org/movie/675353">https://www.themoviedb.org/movie/675353</a>	8	4	2022	
1	The Batman	78.0	Crime, Mystery, Thriller	2022-03-04	2h 56m	Matt Reeves	<a href="https://www.themoviedb.org/movie/414906">https://www.themoviedb.org/movie/414906</a>	4	3	2022	
2	Doctor Strange in the Multiverse of Madness	75.0	Fantasy, Action, Adventure	2022-05-06	2h 6m	Stan Lee	<a href="https://www.themoviedb.org/movie/453395">https://www.themoviedb.org/movie/453395</a>	6	5	2022	
3	The Bad Guys	77.0	Animation, Comedy, Action, Family, Crime	2022-04-22	1h 40m	Pierre Perifel	<a href="https://www.themoviedb.org/movie/629542">https://www.themoviedb.org/movie/629542</a>	22	4	2022	
4	Uncharted	72.0	Action, Adventure	2022-02-18	1h 56m	Ruben Fleischer	<a href="https://www.themoviedb.org/movie/335787">https://www.themoviedb.org/movie/335787</a>	18	2	2022	

```
In [67]: fav_genres_alltime=m_genres.groupby(["Genre 1","Genre 2","Genre 3","Genre 4","Genre 5",'Genre 6','Genre 7'])['Rating'].me
```

In [68]: fav\_genres\_alltime

Out[68]:

	Genre 1	Genre 2	Genre 3	Genre 4	Genre 5	Genre 6	Genre 7	Rating
0	Animation	Action	Adventure	Comedy	Drama	Fantasy	Romance	80.0
1	Family	Fantasy	TV Movie	Music	Adventure	Action	Romance	74.0
2	Animation	Action	Adventure	Comedy	Family	Fantasy	Science Fiction	68.0
3	Science Fiction	Adventure	Mystery	Action	Thriller	Fantasy	Comedy	62.0
4	Adventure	Fantasy	Animation	Action	Comedy	Thriller	Science Fiction	60.0

All time Favourite genres are: Animation, Action, Adventure, Comedy, Drama, Fantasy, Romance.

**Most Popular one from each and every year**

```
In [69]: pop_genre_of_the_year=m_genres.groupby('Year')['Genres'].describe()
pop_genre_of_the_year
```

Out[69]:

	count	unique	top	freq
Year				
1932	1	1	Action, Crime, Drama, Thriller	1
1939	1	1	Adventure, Fantasy, Family	1
1940	3	3	Animation, Family	1
1941	1	1	Animation, Family	1
1942	2	2	Animation, Drama, Family	1
...	...	...	...	...
2018	192	139	Drama	7
2019	213	151	Drama	11
2020	214	151	Horror	17
2021	543	273	Drama	38
2022	271	130	Documentary	28

78 rows × 4 columns

```
In [70]: pop_genre_of_the_year.sort_values(by="freq", ascending=False).head(10)
```

Out[70]:

	count	unique	top	freq
Year				
2021	543	273	Drama	38
2022	271	130	Documentary	28
2020	214	151	Horror	17
2019	213	151	Drama	11
2017	215	156	Drama	10
2013	157	111	Comedy	9
2014	167	121	Comedy	9
2015	180	136	Action	8
2009	121	89	Comedy, Romance	8
2018	192	139	Drama	7

In [ ]:

```
In [75]: M_Directors=m_genres.copy()
M_Directors.head(5)
```

Out[75]:

	Name	Rating	Genres	Date	Runtime	Director	link	Day	Month	Year	Da c th wee
0	Sonic the Hedgehog 2	77.0	Action, Science Fiction, Comedy, Family, Adven...	2022-04-08	2h 2m	Josh Miller	<a href="https://www.themoviedb.org/movie/675353">https://www.themoviedb.org/movie/675353</a>	8	4	2022	
1	The Batman	78.0	Crime, Mystery, Thriller	2022-03-04	2h 56m	Matt Reeves	<a href="https://www.themoviedb.org/movie/414906">https://www.themoviedb.org/movie/414906</a>	4	3	2022	
2	Doctor Strange in the Multiverse of Madness	75.0	Fantasy, Action, Adventure	2022-05-06	2h 6m	Stan Lee	<a href="https://www.themoviedb.org/movie/453395">https://www.themoviedb.org/movie/453395</a>	6	5	2022	
3	The Bad Guys	77.0	Animation, Comedy, Action, Family, Crime	2022-04-22	1h 40m	Pierre Perifel	<a href="https://www.themoviedb.org/movie/629542">https://www.themoviedb.org/movie/629542</a>	22	4	2022	
4	Uncharted	72.0	Action, Adventure	2022-02-18	1h 56m	Ruben Fleischer	<a href="https://www.themoviedb.org/movie/335787">https://www.themoviedb.org/movie/335787</a>	18	2	2022	

```
In [76]: Top_best_Director=M_Directors.groupby(["Director" ])[ 'Rating' ].mean().sort_values(ascending=False).reset_index()  
Top_best_Director
```

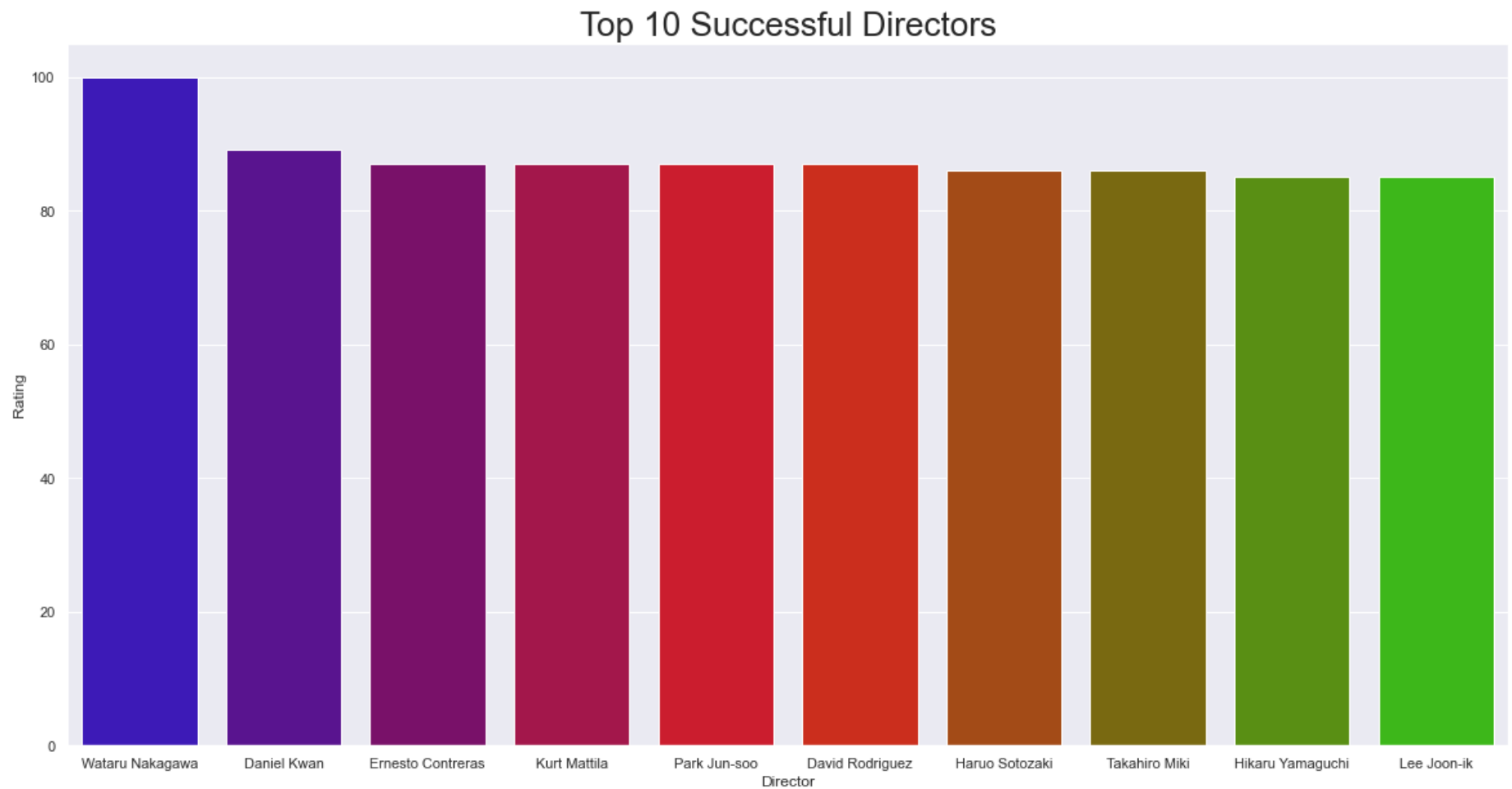
Out[76]:

	Director	Rating
0	Wataru Nakagawa	100.0
1	Daniel Kwan	89.0
2	Ernesto Contreras	87.0
3	Kurt Mattila	87.0
4	Park Jun-soo	87.0
...	...	...
2611	Harald Aue	0.0
2612	Choi Jong-hak	0.0
2613	Charlotte Silvera	0.0
2614	Chadd Harbold	0.0
2615	Otto Binder	0.0

2616 rows × 2 columns

```
In [77]: Top_10_Director=Top_best_Director.head(10)
plt.figure(figsize=(20,10))
plt.title('Top 10 Successful Directors',size=26)

sns.barplot(data=Top_10_Director, x='Director', y='Rating', palette='brg')
sns.set()
plt.show()
```





**Web Series**

```
In [95]: web_series=data[data['Runtime'] == '40m']
web_series
```

Out[95]:

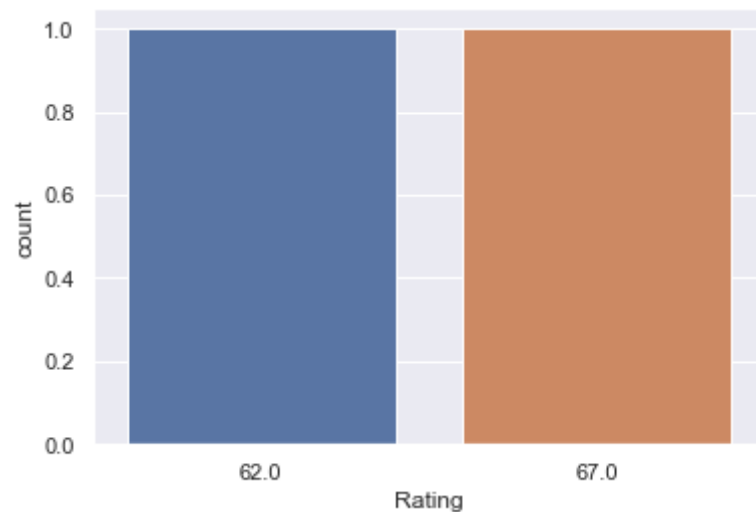
	Name	Rating	Genres	Date	Runtime	Director	link	Day	Month	Year	Day of the week
<b>436</b>	Hold Your Breath: The Ice Dive	62.0	Documentary	2022-05-03	40m	Ian Derry	<a href="https://www.themoviedb.org/movie/958080">https://www.themoviedb.org/movie/958080</a>	3	5	2022	1
<b>2273</b>	Naruto: The Lost Story - Mission: Protect the ...	67.0	Animation, Action, Adventure	2003-12-20	40m	Hayato Date	<a href="https://www.themoviedb.org/movie/410685">https://www.themoviedb.org/movie/410685</a>	20	12	2003	5

```
In [96]: sns.countplot(web_series['Rating'])
```

C:\Users\Admin\anaconda3\lib\site-packages\seaborn\\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

```
Out[96]: <AxesSubplot:xlabel='Rating', ylabel='count'>
```



## Short movies

```
In [97]: short_movie=data[data['Runtime']=='1h']
```

In [99]: short\_movie

Out[99]:

	Name	Rating	Genres	Date	Runtime	Director	link	Day	Month	Year
310	Given	85.0	Animation, Drama, Music, Romance	2020-08-22	1h	Hikaru Yamaguchi	<a href="https://www.themoviedb.org/movie/632632">https://www.themoviedb.org/movie/632632</a>	22	8	2020
478	Dragon Ball Z: The World's Strongest	64.0	Action, Animation, Science Fiction	1990-03-10	1h	Daisuke Nishio	<a href="https://www.themoviedb.org/movie/39100">https://www.themoviedb.org/movie/39100</a>	10	3	1990
1076	Doctor Who: The Time of the Doctor	81.0	Drama, Science Fiction, TV Movie	2013-12-25	1h	Jamie Payne	<a href="https://www.themoviedb.org/movie/282848">https://www.themoviedb.org/movie/282848</a>	25	12	2013
1175	Barbie & Chelsea: The Lost Birthday	72.0	Animation	2021-05-14	1h	Cassi Simonds	<a href="https://www.themoviedb.org/movie/812456">https://www.themoviedb.org/movie/812456</a>	14	5	2021
1628	Re:ZERO -Starting Life in Another World- Memor...	72.0	Animation, Adventure, Fantasy	2018-10-06	1h	Shinichi Otsuka	<a href="https://www.themoviedb.org/movie/532321">https://www.themoviedb.org/movie/532321</a>	6	10	2018
2450	Ao Oni The Animation	52.0	Horror, Animation	2017-02-11	1h	Toshirou Hamamura	<a href="https://www.themoviedb.org/movie/432130">https://www.themoviedb.org/movie/432130</a>	11	2	2017
3172	Stitch! The Movie	65.0	Comedy, Adventure, Family, Animation	2003-08-26	1h	Roberts Gannaway	<a href="https://www.themoviedb.org/movie/15567">https://www.themoviedb.org/movie/15567</a>	26	8	2003
3902	Ver es un acto	0.0	Documentary	2022-05-12	1h	Bárbara Pestan Florás	<a href="https://www.themoviedb.org/movie/965772">https://www.themoviedb.org/movie/965772</a>	12	5	2022

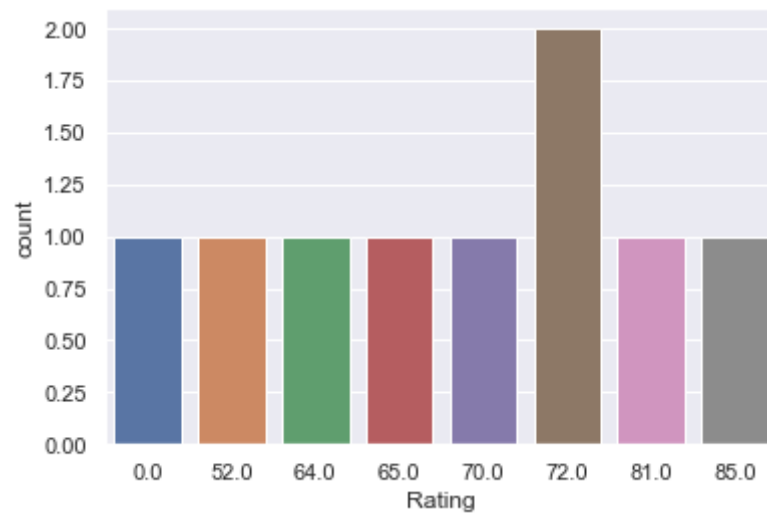
	Name	Rating	Genres	Date	Runtime	Director	link	Day	Month	Year
3941	Harry Potter: A History Of Magic	70.0	Documentary	2017-10-28	1h	Jude Ho	<a href="https://www.themoviedb.org/movie/482408">https://www.themoviedb.org/movie/482408</a>	28	10	2017

In [101]: `sns.countplot(short_movie['Rating'])`

C:\Users\Admin\anaconda3\lib\site-packages\seaborn\\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

Out[101]: `<AxesSubplot:xlabel='Rating', ylabel='count'>`



In [ ]:

