

```
In [1]: import numpy as np  
        import pandas as pd  
        import matplotlib.pyplot as plt  
        import seaborn as sns  
        import missingno as msn
```

```
In [2]: # reading the dataset  
df=pd.read_csv("https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/00
```

In [4]: df.head()

Out[4]:

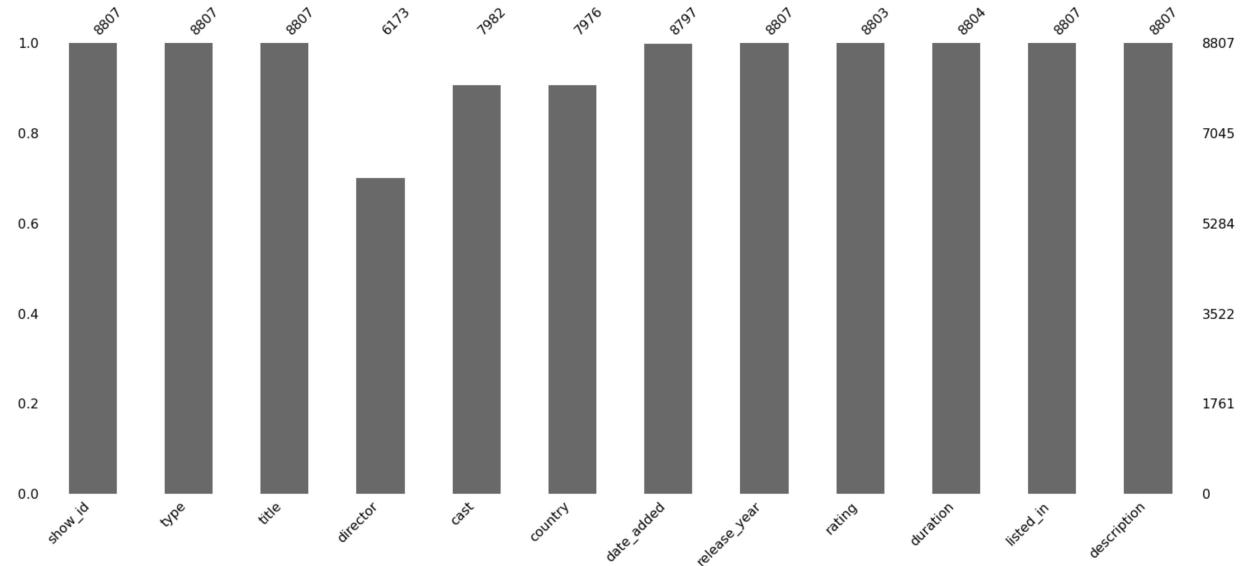
	show_id	type	title	director	cast	country	date_added	release_year	rating	duration
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Nan	United States	September 25, 2021	2020	PG-13	90
1	s2	TV Show	Blood & Water	Nan	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Season 1
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	Nan	September 24, 2021	2021	TV-MA	Season 1
3	s4	TV Show	Jailbirds New Orleans	Nan	Nan	Nan	September 24, 2021	2021	TV-MA	Season 1
4	s5	TV Show	Kota Factory	Nan	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K	India	September 24, 2021	2021	TV-MA	Season 1

In [5]: # Get useful information from the dataset
df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 12 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   show_id     8807 non-null   object  
 1   type        8807 non-null   object  
 2   title       8807 non-null   object  
 3   director    6173 non-null   object  
 4   cast         7982 non-null   object  
 5   country     7976 non-null   object  
 6   date_added  8797 non-null   object  
 7   release_year 8807 non-null   int64  
 8   rating      8803 non-null   object  
 9   duration    8804 non-null   object  
 10  listed_in   8807 non-null   object  
 11  description 8807 non-null   object  
dtypes: int64(1), object(11)
memory usage: 825.8+ KB
```

In [6]: # plotting graph for the missing values
msn.bar(df)

Out[6]: <AxesSubplot:>



Above plot shows the missing values

In [8]: df['type'].unique()

Out[8]: array(['Movie', 'TV Show'], dtype=object)

```
In [9]: df['rating'].unique()
```

```
Out[9]: array(['PG-13', 'TV-MA', 'PG', 'TV-14', 'TV-PG', 'TV-Y', 'TV-Y7', 'R',
   'TV-G', 'G', 'NC-17', '74 min', '84 min', '66 min', 'NR', nan,
   'TV-Y7-FV', 'UR'], dtype=object)
```

```
In [10]: df['release_year'].sort_values().unique()
```

```
Out[10]: array([1925, 1942, 1943, 1944, 1945, 1946, 1947, 1954, 1955, 1956, 1958,
   1959, 1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969,
   1970, 1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980,
   1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991,
   1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002,
   2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013,
   2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021], dtype=int64)
```

```
In [11]: df.shape
```

```
Out[11]: (8807, 12)
```

Dropping rows which are containing time in rating

```
In [12]: df.drop(df.index[df['rating']=='84 min'],inplace=True)
df.drop(df.index[df['rating']=='66 min'],inplace=True)
df.drop(df.index[df['rating']=='74 min'],inplace=True)
df.shape
```

```
Out[12]: (8804, 12)
```

In [13]: df.head()

Out[13]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	dura
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Sea:
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	Se:
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	Se:
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Sea:



In [14]: pd.set_option('display.max_rows', 5000)

In [15]: `df['release_year'].value_counts()`

Out[15]:

2018	1147
2017	1031
2019	1030
2020	953
2016	902
2021	592
2015	559
2014	352
2013	288
2012	237
2010	193
2011	185
2009	152
2008	136
2006	96
2007	88
2005	80
2004	64
2003	61
2002	51
2001	45
1999	39
1997	38
2000	37
1998	36
1993	28
1995	25
1996	24
1992	23
1994	22
1990	22
1988	18
1982	17
1991	17
1989	16
1981	13
1986	13
1984	12
1980	11
1979	11
1983	11
1985	10
1973	10
1976	9
1987	8
1977	7
1974	7
1975	7
1978	7
1972	5
1967	5
1971	5
1960	4
1945	4
1944	3

```
1955      3
1968      3
1962      3
1943      3
1958      3
1964      2
1969      2
1970      2
1963      2
1956      2
1965      2
1946      2
1942      2
1954      2
1959      1
1925      1
1961      1
1947      1
1966      1
Name: release_year, dtype: int64
```

```
In [16]: sum(df[df['release_year']<=2015].value_counts())
```

```
Out[16]: 2328
```

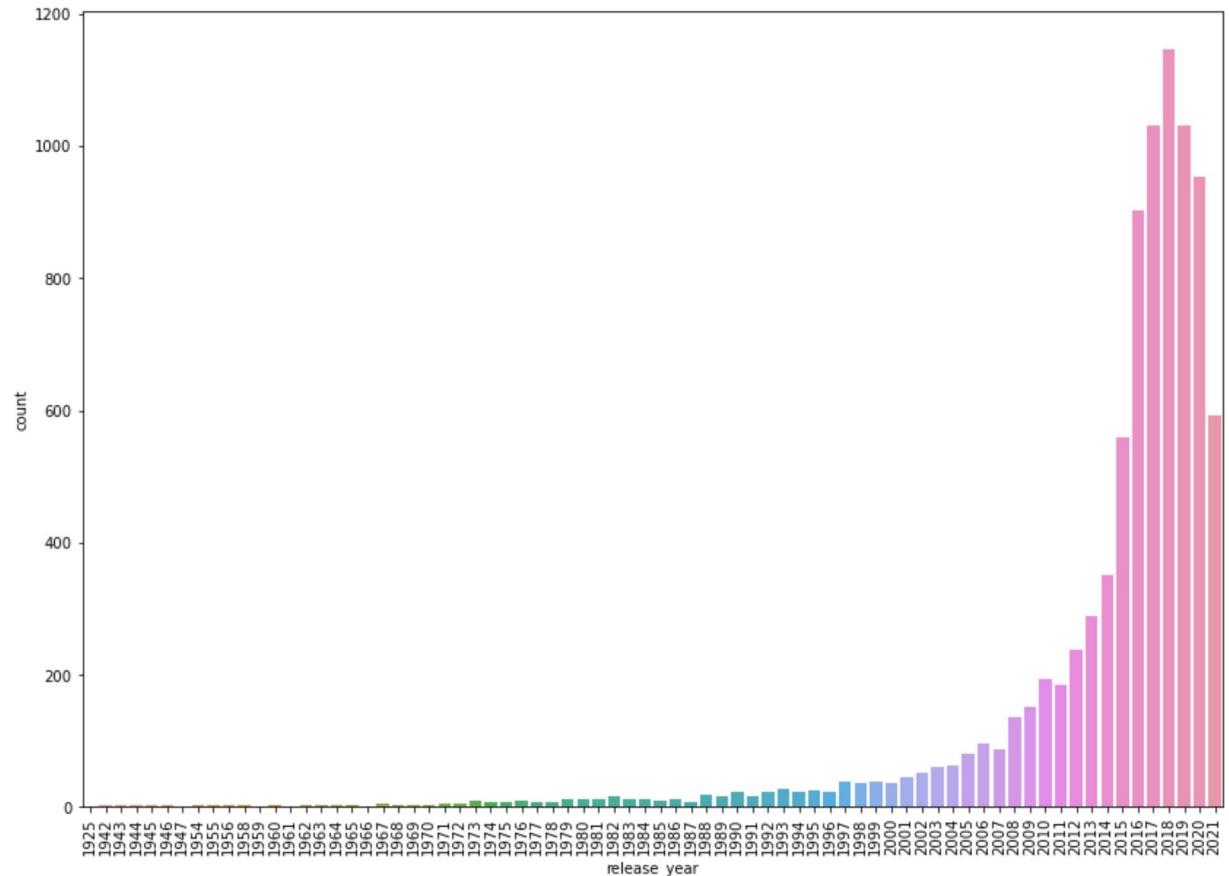
```
In [17]: sum(df[df['release_year']>2015].value_counts())
```

```
Out[17]: 3004
```

```
In [18]: plt.figure(figsize=(14,10))
plt.xticks(rotation=90)
sns.countplot(df['release_year'])
```

C:\Users\hp\anaconda3\lib\site-packages\seaborn_decorators.py:36: FutureWarning:
g: 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 a
n explicit keyword will result in an error or misinterpretation.
warnings.warn(

```
Out[18]: <AxesSubplot:xlabel='release_year', ylabel='count'>
```



```
In [19]: print(df['release_year'].min(), '-', df['release_year'].max())
```

1925 - 2021

```
In [20]: print('Movies released after 2015 : ',sum(df[(df['release_year']>2015) & (df['type']=='Movie')]))
print('TV Show released after 2015 : ',sum(df[(df['release_year']>2015) & (df['type']=='TV Show')]))
```

Movies released after 2015 : 2889
 TV Show released after 2015 : 115

```
In [21]: print('Movies released till 2015 : ',sum(df[(df['release_year']<=2015) & (df['type']=='Movie')]))
print('TV Show released till 2015 : ',sum(df[(df['release_year']<=2015) & (df['type']=='TV Show')]))
```

Movies released till 2015 : 2296
 TV Show released till 2015 : 32

```
In [22]: print('Movie : ',sum(df['type']=='Movie'), '| ', 'TV Show : ',sum(df['type']=='TV Show'))
```

Movie : 6128 | TV Show : 2676

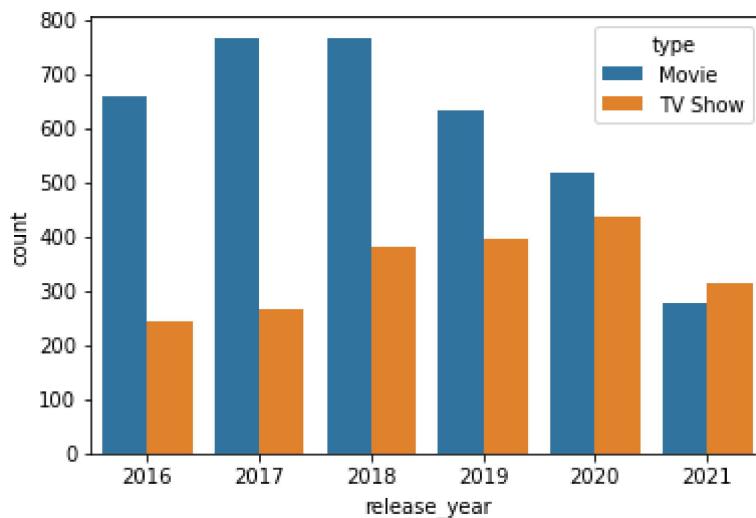
```
In [466]: df_era2=df[df["release_year"]>2015]
df_era2.head()
```

Out[466]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	dura
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Sea...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	Sea...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	Sea...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Sea...

```
In [470]: sns.countplot(data=df_era2,x='release_year',hue='type')
```

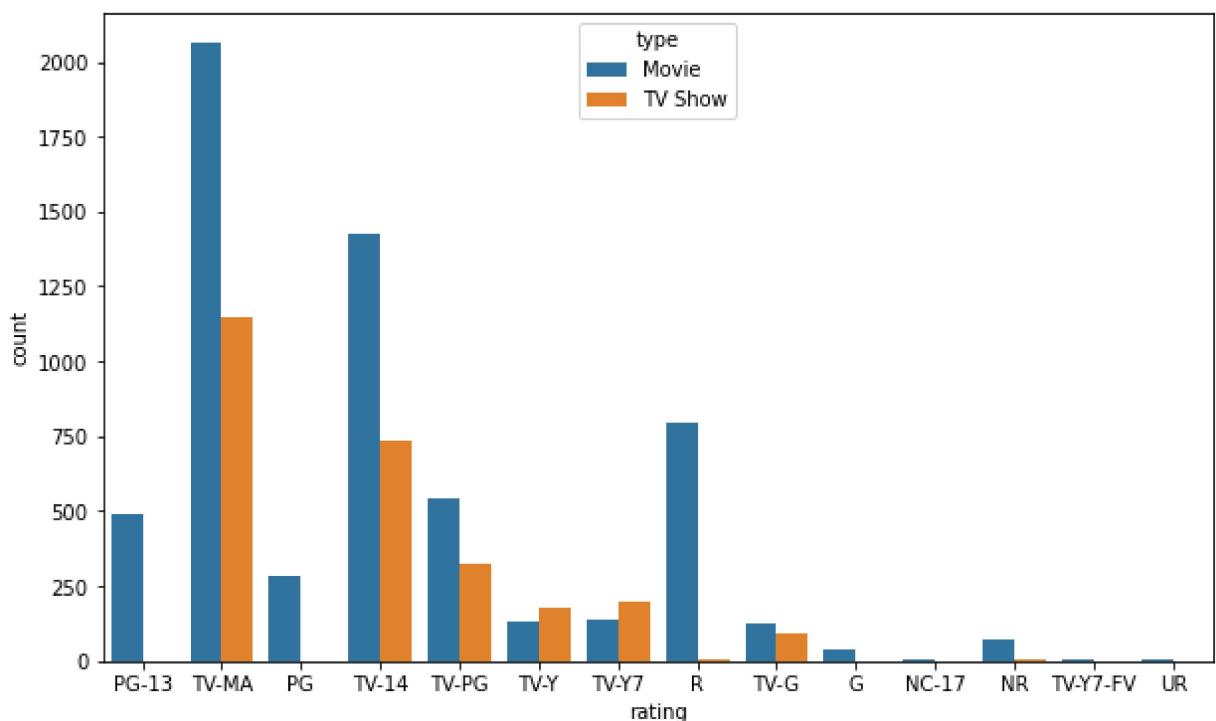
```
Out[470]: <AxesSubplot:xlabel='release_year', ylabel='count'>
```



```
In [23]: plt.figure(figsize=(10,6))
sns.countplot(df['rating'],hue=df['type'])
```

C:\Users\hp\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(

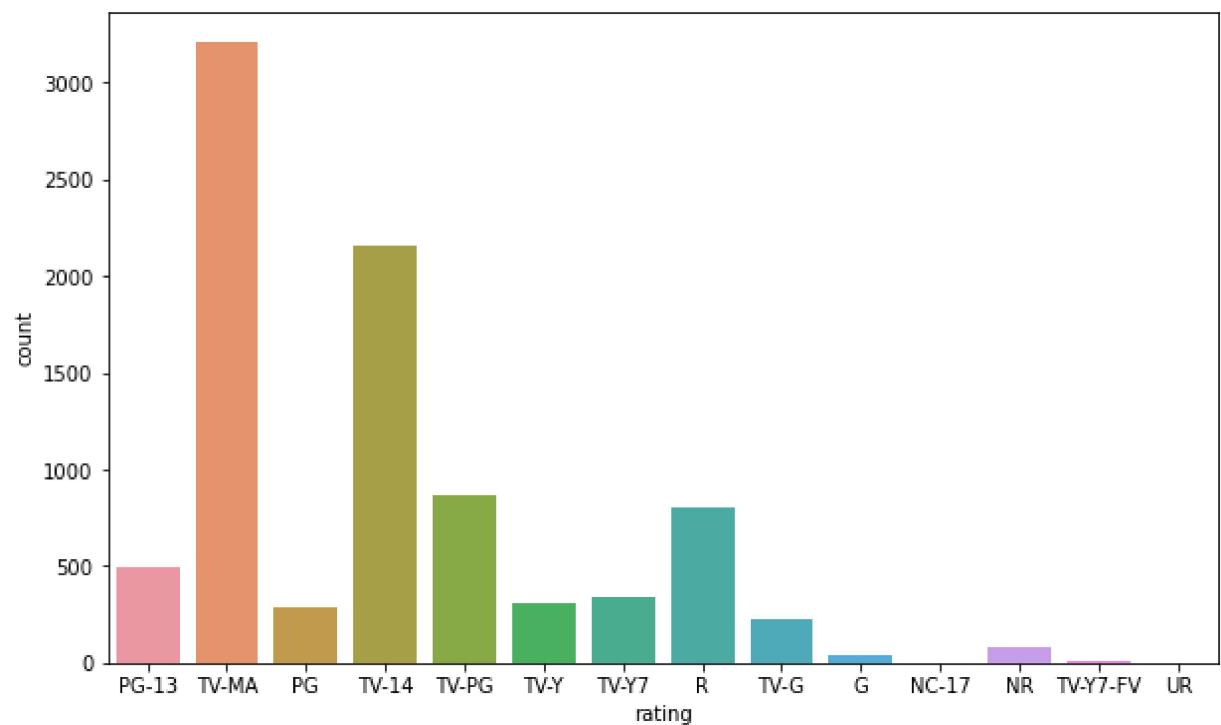
```
Out[23]: <AxesSubplot:xlabel='rating', ylabel='count'>
```



```
In [24]: plt.figure(figsize=(10,6))
sns.countplot(df['rating'])
```

```
C:\Users\hp\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning
g: 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 a
n explicit keyword will result in an error or misinterpretation.
warnings.warn(
```

```
Out[24]: <AxesSubplot:xlabel='rating', ylabel='count'>
```



Finding the country data for Movies and TV Show.

There are movies and TV Shows which are released in multiple countries so we have split and find the data for each country

In [25]: df.head()

Out[25]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	dura
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Sea:
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	Se:
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	Se:
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Sea:

In [77]: df_country=df[["country","type"]].value_counts().to_frame().reset_index()
df_country.head()

Out[77]:

	country	type	0
0	United States	Movie	2055
1	India	Movie	893
2	United States	TV Show	760
3	United Kingdom	TV Show	213
4	United Kingdom	Movie	206

```
In [97]: df_country.rename(columns={0:"count"},inplace=True)  
df_country
```

7		Canada	Movie	122
8		Spain	Movie	97
9		Egypt	Movie	92
10		Nigeria	Movie	86
11		India	TV Show	79
12		Indonesia	Movie	77
13		Turkey	Movie	76
14		Japan	Movie	76
15		France	Movie	75
16		Philippines	Movie	73
17		Mexico	Movie	70
18		Taiwan	TV Show	68
19	United Kingdom, United States		Movie	63

```
In [84]: df_country.pivot(index="country",  
                      columns="type",  
                      values="count"  
)
```

Out[84]:

	type	Movie	TV Show
country			
, France, Algeria	1.0	NaN	
, South Korea	NaN	1.0	
Argentina	38.0	18.0	
Argentina, Brazil, France, Poland, Germany, Denmark	1.0	NaN	
Argentina, Chile	2.0	NaN	
Argentina, Chile, Peru	1.0	NaN	
Argentina, France	1.0	NaN	
Argentina, France, United States, Germany, Qatar	1.0	NaN	
Argentina, Italy	1.0	NaN	
Argentina, Spain	7.0	1.0	

```
In [100]: df_country['country']=df_country['country'].apply(lambda x:str(x).split(','))  
df_country
```

Out[100]:

	country	type	count
0	[United States]	Movie	2055
1	[India]	Movie	893
2	[United States]	TV Show	760
3	[United Kingdom]	TV Show	213
4	[United Kingdom]	Movie	206
5	[Japan]	TV Show	169
6	[South Korea]	TV Show	158
7	[Canada]	Movie	122
8	[Spain]	Movie	97
9	[Egypt]	Movie	92
10	[Nigeria]	Movie	86

```
In [106]: df_country_1=df_country.explode('country')
```

In [131]: df_country_1

16	Philippines	Movie	73
17	Mexico	Movie	70
18	Taiwan	TV Show	68
19	United Kingdom	Movie	63
19	United States	Movie	63
20	Canada	TV Show	59
21	United States	Movie	51
21	Canada	Movie	51
22	Brazil	Movie	50
23	Hong Kong	Movie	50
24	France	TV Show	49

```
In [113]: df_country_2=df_country_1.drop_duplicates()  
df_country_2.reset_index(drop=True,inplace=True)
```

In [186]: df_country_2

Out[186]:

	country	type	count
0	United States	Movie	2055
1	India	Movie	893
2	United States	TV Show	760
3	United Kingdom	TV Show	213
4	United Kingdom	Movie	206
5	Japan	TV Show	169
6	South Korea	TV Show	158
7	Canada	Movie	122
8	Spain	Movie	97
9	Egypt	Movie	92
10	Nigeria	Movie	86

In [189]: df_cwise=df_country_2.groupby(['country','type']).sum()
df_cwise

Indonesia	Movie	80
	TV Show	2
Iran	Movie	1
Ireland	Movie	10
	TV Show	4
Israel	Movie	9
	TV Show	10
Italy	Movie	41
	TV Show	14
Jamaica	Movie	1
Japan	Movie	85
	TV Show	173
Jordan	TV Show	2

From the above tabl we can see how many movies and tv shows released per country

In [190]: df

Out[190]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Nan	United States	September 25, 2021	2020	PG-13
1	s2	TV Show	Blood & Water	Nan	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	Nan	September 24, 2021	2021	TV-MA
3	s4	TV Show	Jailbirds New Orleans	Nan	Nan	Nan	September 24, 2021	2021	TV-MA
4	s5	TV Show	Kota Factory	Nan	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA
...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R
8803	s8804	TV Show	Zombie Dumb	Nan	Nan	Nan	July 1, 2019	2018	TV-Y7
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R

	show_id	type		title	director	cast	country	date_added	release_year	rating
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020		2006	PG
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019		2015	TV-14

8804 rows × 12 columns



In [191]: `df_timewise=df.copy()`

In [192]: df_timewise.head(10)

Out[192]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Nan	United States	September 25, 2021	2020	PG-13
1	s2	TV Show	Blood & Water	Nan	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	Nan	September 24, 2021	2021	TV-MA
3	s4	TV Show	Jailbirds New Orleans	Nan	Nan	Nan	September 24, 2021	2021	TV-MA
4	s5	TV Show	Kota Factory	Nan	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	Nan	September 24, 2021	2021	TV-MA
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	Nan	September 24, 2021	2021	PG
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmi Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	September 24, 2021	1993	TV-MA
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...	United Kingdom	September 24, 2021	2021	TV-14

	show_id	type	title	director	cast	country	date_added	release_year	rating
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	September 24, 2021	2021	PG-13



```
In [199]: df_timewise['month']=pd.to_datetime(df_timewise['date_added']).dt.month_name()
```

In [200]: df_timewise

Out[200]:

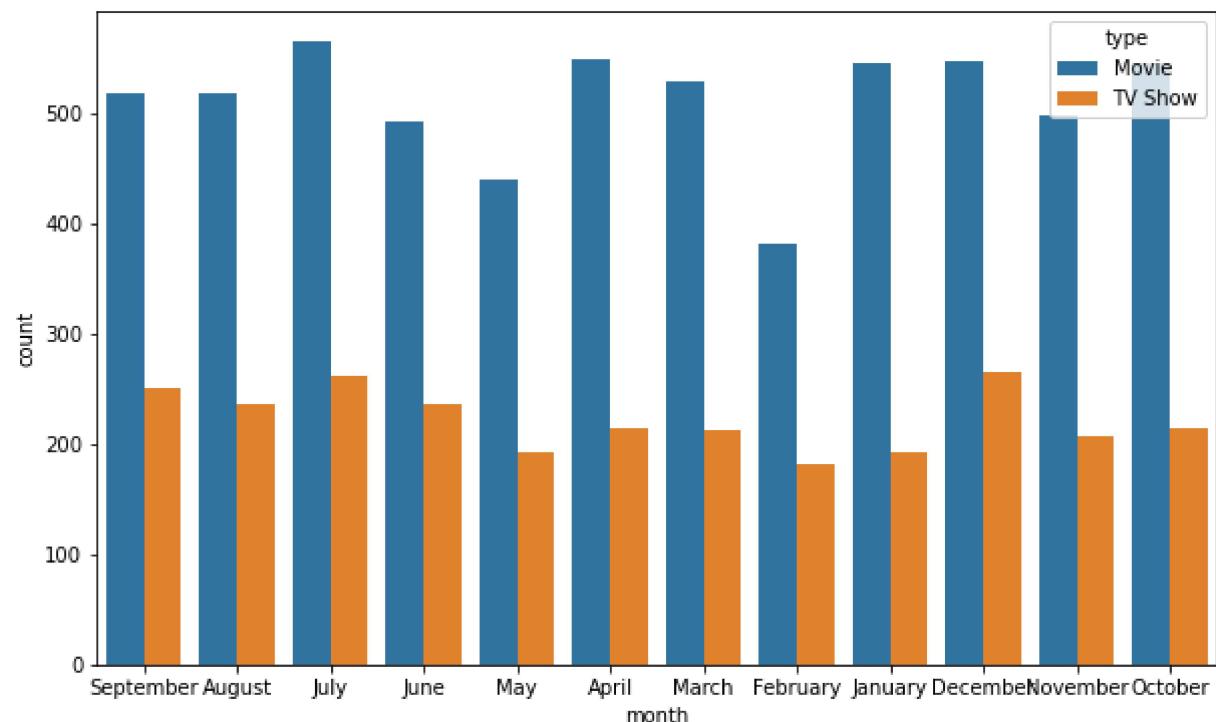
	show_id	type	title	director	cast	country	date_added	release_year	ratio
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Nan	United States	September 25, 2021	2020	P
1	s2	TV Show	Blood & Water	Nan	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	T N
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	Nan	September 24, 2021	2021	T N
3	s4	TV Show	Jailbirds New Orleans	Nan	Nan	Nan	September 24, 2021	2021	T N
4	s5	TV Show	Kota Factory	Nan	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	T N
...
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	rating
8805	s8806	Movie	Zoom	Peter Hewitt Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	F
8806	s8807	Movie	Zubaan	Mozez Singh Vicky Kaushal, Sarah-Jane Dias, Raaghav Chan...	India	March 2, 2019	2015	TV-

8804 rows × 13 columns

```
In [472]: plt.figure(figsize=(10,6))
sns.countplot(data=df_timewise,x="month",hue='type')
```

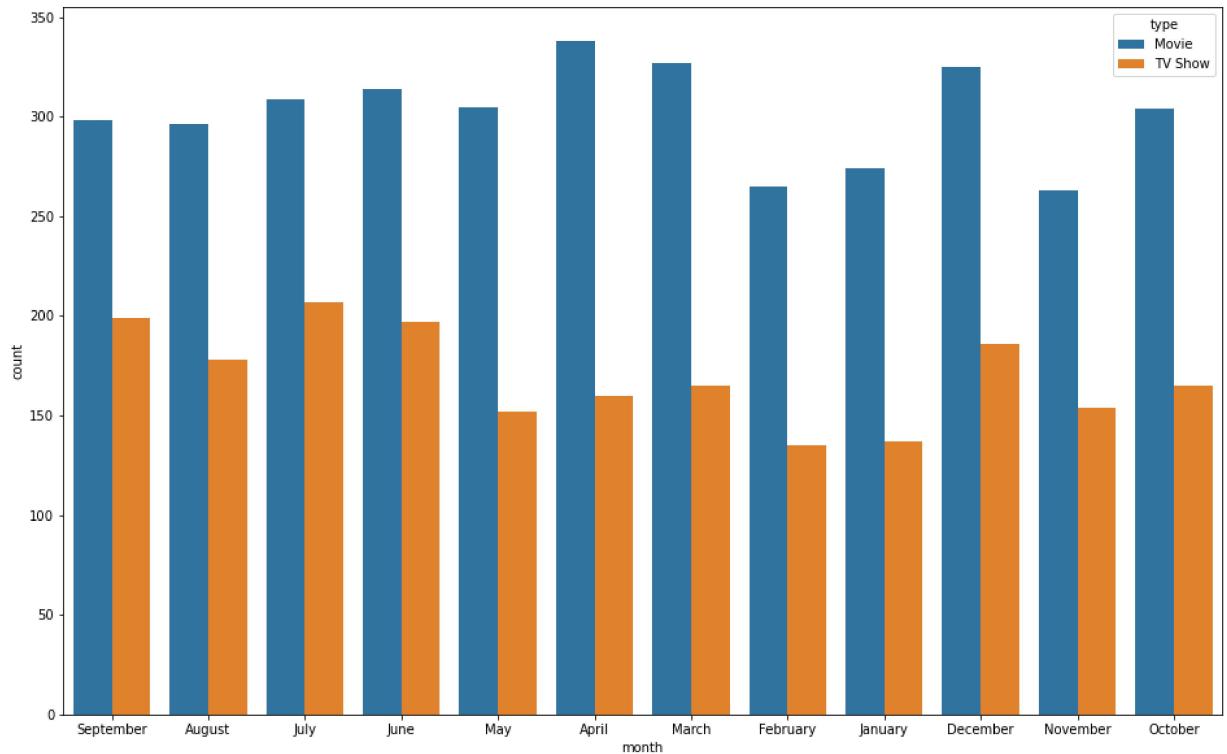
```
Out[472]: <AxesSubplot:xlabel='month', ylabel='count'>
```



```
In [212]: df_2015=df_timewise[df_timewise['release_year']>2015]
```

```
In [218]: plt.figure(figsize=(16,10))
sns.countplot(data=df_2015,x="month",hue='type')
```

```
Out[218]: <AxesSubplot:xlabel='month', ylabel='count'>
```



Plot showing how many movies adn TV Shows released after 2015 month wise.

In [293]: df.head()

Out[293]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	dura
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Sea:
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	Se:
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	Se:
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Sea:



In [294]: df_dir=df.copy()

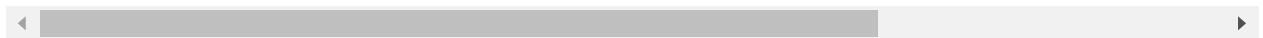
In [296]: # Separate dataset to analyse the director column
df_dir=df_dir.dropna(subset=['director'])

In [297]:

```
df_dir.reset_index(inplace=True)
df_dir.head()
```

Out[297]:

	index	show_id	type	title	director	cast	country	date_added	release_year	ra
0	0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	
1	2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	
2	5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	NaN	September 24, 2021	2021	
3	6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2021	
4	7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	September 24, 2021	1993	



In [306]: `df_dir.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6170 entries, 0 to 6169
Data columns (total 12 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   show_id          6170 non-null    object  
 1   type              6170 non-null    object  
 2   title             6170 non-null    object  
 3   director          6170 non-null    object  
 4   cast               5697 non-null    object  
 5   country            5748 non-null    object  
 6   date_added        6170 non-null    object  
 7   release_year      6170 non-null    int64  
 8   rating             6169 non-null    object  
 9   duration            6170 non-null    object  
 10  listed_in          6170 non-null    object  
 11  description         6170 non-null    object  
dtypes: int64(1), object(11)
memory usage: 578.6+ KB
```

In []: `df_dir.drop('index',axis=1,inplace=True)`

In [302]: df_dir.head()

Out[302]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	100
1	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	80
2	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	NaN	September 24, 2021	2021	TV-MA	80
3	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2021	PG	100
4	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlana, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	September 24, 2021	1993	TV-MA	100



In [309]: df_dir['director']=df_dir['director'].apply(lambda x:str(x).split(','))

In [310]: df_dir

Out[310]:

	show_id	type	title	director	cast	country	date_added	release_year	ratir
0	s1	Movie	Dick Johnson Is Dead	[Kirsten Johnson]	NaN	United States	September 25, 2021	2020	P
1	s3	TV Show	Ganglands	[Julien Leclercq]	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	T M
2	s6	TV Show	Midnight Mass	[Mike Flanagan]	Kate Siegel, Zach Gilford, Hamish Linklater, H...	NaN	September 24, 2021	2021	T M
3	s7	Movie	My Little Pony: A New Generation	[Robert Cullen, José Luis Ucha]	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2021	P
4	s8	Movie	Sankofa	[Haile Gerima]	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	September 24, 2021	1993	T M
...
6165	s8802	Movie	Zinzana	[Majid Al Ansari]	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ...	United Arab Emirates, Jordan	March 9, 2016	2015	T M
6166	s8803	Movie	Zodiac	[David Fincher]	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	
6167	s8805	Movie	Zombieland	[Ruben Fleischer]	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	
6168	s8806	Movie	Zoom	[Peter Hewitt]	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	P

show_id	type	title	director	cast	country	date_added	release_year	ratir
6169	s8807	Movie	Zubaan [Mozez Singh]	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanani...	India	March 2, 2019	2015	TV-

6170 rows × 12 columns



```
In [311]: df_dir1=df_dir.explode('director')
```

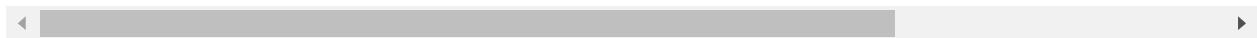
In [312]: df_dir1

Out[312]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Nan	United States	September 25, 2021	2020	PG-13
1	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	Nan	September 24, 2021	2021	TV MA
2	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	Nan	September 24, 2021	2021	TV MA
3	s7	Movie	My Little Pony: A New Generation	Robert Cullen	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	Nan	September 24, 2021	2021	PG
3	s7	Movie	My Little Pony: A New Generation	José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	Nan	September 24, 2021	2021	PG
...
6165	s8802	Movie	Zinzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ...	United Arab Emirates, Jordan	March 9, 2016	2015	TV MA
6166	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	F
6167	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	F

show_id	type	title	director	cast	country	date_added	release_year	rating
6168	s8806	Movie	Zoom	Peter Hewitt Tim Allen, Courtene Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG
6169	s8807	Movie	Zubaan	Mozez Singh Vicky Kaushal, Sarah-Jane Dias, Raaghav Chan...	India	March 2, 2019	2015	TV-14

6975 rows × 12 columns



In [474]: `df_dir=(df_dir1[['director','type','country']].value_counts()).to_frame()`

In [475]: `df_dir.rename(columns={0:'count'},inplace=True)`

In [476]: `df_dir`

Out[476]:

director	type	country	count
Jay Karas	Movie	United States	15
Marcus Raboy	Movie	United States	14
Cathy Garcia-Molina	Movie	Philippines	13
Jay Chapman	Movie	United States	12
Shannon Hartman	Movie	United States	9
...
Mike Smith	TV Show	Canada	1
Mir Mukarram Hossain	Movie	Bangladesh	1
Haissam Hussain	Movie	Pakistan	1
Haile Gerima	Movie	United States, Ghana, Burkina Faso, United Kingdom, Germany, Ethiopia	1
John Dahl	Movie	United States, Australia	1

5293 rows × 1 columns

Above table shows that the no. of movies or TV shows directed by director

In [381]: df.head()

Out[381]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	dura
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Sea:
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	Se:
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	Se:
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Sea:



In [382]: df_list=df.copy()

In [383]: df_list['listed_in']=df_list['listed_in'].apply(lambda x:str(x).split(','))
df_list

Out[383]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Nan	United States	September 25, 2021	2020	PG-13
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA
...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R

show_id	type	title	director	cast	country	date_added	release_year	rating
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006 PG
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015 TV-14

8804 rows × 12 columns



```
In [384]: df_list=df_list.explode('listed_in')
df_list
```

Out[384]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA
...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chan...	India	March 2, 2019	2015	TV-14

show_id	type	title	director	cast	country	date_added	release_year	rating
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chan...	India	March 2, 2019	2015 TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chan...	India	March 2, 2019	2015 TV-14

19320 rows × 12 columns

In [386]: df_list["listed_in"].nunique()

Out[386]: 73

In [387]: df_list["listed_in"].unique()

```
Out[387]: array(['Documentaries', 'International TV Shows', ' TV Dramas',
       ' TV Mysteries', 'Crime TV Shows', ' International TV Shows',
       ' TV Action & Adventure', 'Docuseries', ' Reality TV',
       ' Romantic TV Shows', ' TV Comedies', 'TV Dramas', ' TV Horror',
       'Children & Family Movies', 'Dramas', ' Independent Movies',
       ' International Movies', 'British TV Shows', 'Comedies', ' Dramas',
       ' Docuseries', ' Comedies', ' Crime TV Shows', 'TV Comedies',
       ' Spanish-Language TV Shows', 'Thrillers', ' Romantic Movies',
       ' Music & Musicals', 'Horror Movies', ' Sci-Fi & Fantasy',
       ' TV Thrillers', "Kids' TV", ' Thrillers', 'Action & Adventure',
       ' TV Sci-Fi & Fantasy', ' Classic Movies', ' Horror Movies',
       ' Anime Features', 'Reality TV', ' Sports Movies', 'Anime Series',
       " Kids' TV", 'International Movies', ' Korean TV Shows',
       'Sci-Fi & Fantasy', ' Science & Nature TV', ' Teen TV Shows',
       ' Cult Movies', 'Classic Movies', 'TV Shows',
       ' Children & Family Movies', ' Faith & Spirituality',
       ' LGBTQ Movies', 'Stand-Up Comedy', 'TV Action & Adventure',
       'Movies', 'Stand-Up Comedy & Talk Shows', 'Classic & Cult TV',
       ' Stand-Up Comedy & Talk Shows', 'Anime Features',
       ' Documentaries', 'Romantic TV Shows', 'Cult Movies',
       'Independent Movies', 'TV Horror', 'Spanish-Language TV Shows',
       ' Classic & Cult TV', 'Music & Musicals', 'Romantic Movies',
       'LGBTQ Movies', ' Stand-Up Comedy', 'TV Sci-Fi & Fantasy',
       'Sports Movies'], dtype=object)
```

```
In [388]: df_list["listed_in"].value_counts()
```

```
Out[388]: International Movies           2624
Dramas                           1600
Comedies                          1210
Action & Adventure                 859
Documentaries                      829
Dramas                           827
International TV Shows            774
Independent Movies                736
TV Dramas                          696
Romantic Movies                   613
Children & Family Movies          605
International TV Shows            577
Thrillers                          512
Comedies                          464
TV Comedies                       461
Crime TV Shows                    399
Kids' TV                           388
Music & Musicals                  357
Romantic TV Shows                338
Stand-Up Comedy                   334
Horror Movies                      275
British TV Shows                 253
Sci-Fi & Fantasy                  230
Docuseries                         221
Sports Movies                      218
Anime Series                       176
Docuseries                         174
Spanish-Language TV Shows         172
Korean TV Shows                   151
Reality TV                          135
TV Action & Adventure             128
International Movies               128
TV Comedies                        120
Reality TV                          120
LGBTQ Movies                       101
TV Mysteries                       98
Science & Nature TV                92
TV Sci-Fi & Fantasy                83
Horror Movies                      82
Classic Movies                     80
Crime TV Shows                    71
Teen TV Shows                      69
TV Dramas                          67
Thrillers                          65
Faith & Spirituality              65
TV Horror                           64
Kids' TV                            63
Cult Movies                         59
TV Thrillers                       57
Movies                             54
Anime Features                     50
Documentaries                      40
TV Action & Adventure              40
Classic Movies                      36
Children & Family Movies           36
```

Stand-Up Comedy & Talk Shows	34
Romantic TV Shows	32
Stand-Up Comedy & Talk Shows	22
Classic & Cult TV	22
Anime Features	21
Independent Movies	20
Music & Musicals	18
TV Shows	16
Sci-Fi & Fantasy	13
Cult Movies	12
TV Horror	11
Stand-Up Comedy	9
Classic & Cult TV	6
Romantic Movies	3
Spanish-Language TV Shows	2
LGBTQ Movies	1
TV Sci-Fi & Fantasy	1
Sports Movies	1

Name: listed_in, dtype: int64

About data shows the count of movies according to genre

In [413]: df.head()

Out[413]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	dura
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Sea:
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	Se:
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	Se:
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Sea:

In [414]: df_avg_movie_time=df[df['duration'].str.contains('min')]

In [415]: df_avg_movie_time.head()

Out[415]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG 1
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2021	PG
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmi Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kingdom	September 24, 2021	1993	TV M
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	September 24, 2021	2021	PG 1
12	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Wedler, Jannis Niewöhner, Milan Peschel, ...	Germany, Czech Republic	September 23, 2021	2021	TV M

In [416]: df_avg_movie_time=df_avg_movie_time['duration'].to_frame()

In [417]: df_avg_movie_time['duration']=df_avg_movie_time['duration'].apply(lambda x:str(x))
df_avg_movie_time.head()

Out[417]:

	duration
0	[90, min]
6	[91, min]
7	[125, min]
9	[104, min]
12	[127, min]

In [418]: `df_avg_movie_time['duration']=df_avg_movie_time['duration'].apply(lambda x:int(x))`

Out[418]:

	duration
0	90
6	91
7	125
9	104
12	127
...	...
8801	96
8802	158
8804	88
8805	88
8806	111

6128 rows × 1 columns

In [419]: `df_avg_movie_time.describe()`

Out[419]:

	duration
count	6128.000000
mean	99.577187
std	28.290593
min	3.000000
25%	87.000000
50%	98.000000
75%	114.000000
max	312.000000

Above table shows the information about the timings of the movie.

avg. time for movies is 99 min.

In [429]: df.head()

Out[429]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	dura
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Sea:
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	Se:
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	Se:
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Sea:

In [431]: df_cast=df.copy()

In [432]: df_cast['cast']=df_cast['cast'].apply(lambda x:str(x).split(','))

In [433]: df_cast

Out[433]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	[nan]	United States	September 25, 2021	2020	PG-13
1	s2	TV Show	Blood & Water	NaN	[Ama Qamata, Khosi Ngema, Gail Mabalane, Th...]	South Africa	September 24, 2021	2021	TV-MA
2	s3	TV Show	Ganglands	Julien Leclercq	[Sami Bouajila, Tracy Gotoas, Samuel Jouy, ...]	NaN	September 24, 2021	2021	TV-MA
3	s4	TV Show	Jailbirds New Orleans	NaN	[nan]	NaN	September 24, 2021	2021	TV-MA
4	s5	TV Show	Kota Factory	NaN	[Mayur More, Jitendra Kumar, Ranjan Raj, Al...]	India	September 24, 2021	2021	TV-MA
...
8802	s8803	Movie	Zodiac	David Fincher	[Mark Ruffalo, Jake Gyllenhaal, Robert Downe...]	United States	November 20, 2019	2007	R
8803	s8804	TV Show	Zombie Dumb	NaN	[nan]	NaN	July 1, 2019	2018	TV-Y7
8804	s8805	Movie	Zombieland	Ruben Fleischer	[Jesse Eisenberg, Woody Harrelson, Emma Ston...]	United States	November 1, 2019	2009	R
8805	s8806	Movie	Zoom	Peter Hewitt	[Tim Allen, Courteney Cox, Chevy Chase, Kat...]	United States	January 11, 2020	2006	PG

show_id	type	title	director	cast	country	date_added	release_year	rating
8806	s8807	Movie	Zubaan	Mozez Singh	[Vicky Kaushal, Sarah-Jane Dias, Raaghav Ch...]	India	March 2, 2019	2015 TV-14

8804 rows × 12 columns



```
In [436]: df_cast=df_cast.explode('cast')
```

In [440]: df_cast

Out[440]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	nan	United States	September 25, 2021	2020	PG-13
1	s2	TV Show	Blood & Water	NaN	Ama Qamata	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Khosi Ngema	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Gail Mabalane	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Thabang Molaba	South Africa	September 24, 2021	2021	TV-MA
...
8806	s8807	Movie	Zubaan	Mozez Singh	Manish Chaudhary	India	March 2, 2019	2015	TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Meghna Malik	India	March 2, 2019	2015	TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Malkeet Rauni	India	March 2, 2019	2015	TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Anita Shabdish	India	March 2, 2019	2015	TV-14

show_id	type	title	director	cast	country	date_added	release_year	rating
8806	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	March 2, 2019	2015 TV-14

64948 rows × 12 columns

In [438]: df_cast.info()

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 64948 entries, 0 to 8806
Data columns (total 12 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   show_id          64948 non-null   object 
 1   type              64948 non-null   object 
 2   title             64948 non-null   object 
 3   director          45935 non-null   object 
 4   cast               64948 non-null   object 
 5   country            59895 non-null   object 
 6   date_added        64879 non-null   object 
 7   release_year      64948 non-null   int64  
 8   rating             64910 non-null   object 
 9   duration            64948 non-null   object 
 10  listed_in          64948 non-null   object 
 11  description         64948 non-null   object 
dtypes: int64(1), object(11)
memory usage: 6.4+ MB
```

In [437]: df_cast['cast'].nunique()

Out[437]: 39296

In [439]: df_cast['cast'].value_counts()

```
nan                      825
Anupam Kher                39
Rupa Bhimani                31
Takahiro Sakurai                30
Julie Tejwani                28
...
Tedros Teclebrhan                 1
Maryam Zaree                  1
Melanie Straub                  1
Gabriela Maria Schmeide                1
Chittaranjan Tripathy                1
Name: cast, Length: 39296, dtype: int64
```

In [447]: df_cast['cast'].replace('nan', np.nan, inplace=True)

In [448]: df_cast

Out[448]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13
1	s2	TV Show	Blood & Water	NaN	Ama Qamata	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Khosi Ngema	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Gail Mabalane	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Thabang Molaba	South Africa	September 24, 2021	2021	TV-MA
...
8806	s8807	Movie	Zubaan	Mozez Singh	Manish Chaudhary	India	March 2, 2019	2015	TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Meghna Malik	India	March 2, 2019	2015	TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Malkeet Rauni	India	March 2, 2019	2015	TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Anita Shabdish	India	March 2, 2019	2015	TV-14

show_id	type	title	director	cast	country	date_added	release_year	rating
8806	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	March 2, 2019	2015 TV-14

64948 rows × 12 columns

```
In [449]: df_cast=df_cast.dropna(subset=['cast'])
```

In [450]: df_cast

Out[450]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
1	s2	TV Show	Blood & Water	NaN	Ama Qamata	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Khosi Ngema	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Gail Mabalane	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Thabang Molaba	South Africa	September 24, 2021	2021	TV-MA
1	s2	TV Show	Blood & Water	NaN	Dillon Windvogel	South Africa	September 24, 2021	2021	TV-MA
...
8806	s8807	Movie	Zubaan	Mozez Singh	Manish Chaudhary	India	March 2, 2019	2015	TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Meghna Malik	India	March 2, 2019	2015	TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Malkeet Rauni	India	March 2, 2019	2015	TV-14
8806	s8807	Movie	Zubaan	Mozez Singh	Anita Shabdish	India	March 2, 2019	2015	TV-14

show_id	type	title	director	cast	country	date_added	release_year	rating
8806	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	March 2, 2019	2015 TV-14

64123 rows × 12 columns

In [483]: `df_cas=df_cast.groupby(['country','cast','listed_in'])['cast'].count().to_frame()
df_cas`

Out[483]:

country	cast	listed_in	cast
, France, Algeria	Areen Omari	Dramas, Independent Movies, International Movies	1
	Maisa Abd Elhadi	Dramas, Independent Movies, International Movies	1
	Malak Ermileh	Dramas, Independent Movies, International Movies	1
	Nisreen Faour	Dramas, Independent Movies, International Movies	1
	Souad Massi	Dramas, Independent Movies, International Movies	1
...
Zimbabwe	Kudzai Sevenzo	Comedies, International Movies, Romantic Movies	1
	Michael Kudakwashe	Comedies, International Movies, Romantic Movies	1
	Tendai Nguni	Comedies, International Movies, Romantic Movies	1
	Zihlo	Comedies, International Movies, Romantic Movies	1
	Tendaishe Chitima	Comedies, International Movies, Romantic Movies	1

55319 rows × 1 columns

In [482]: `df_cas.loc['India']`

cast	
Alok Nath	11
Alok Rajwade	1
Alyy Khan	1
Amal Unnithan	1
Amala	1
Amala Paul	2
Aman Attar	1
Aman Hundal	1
Aman Sutdhar	1
Aman Uppal	1
Aman Verma	1
Amaraja Patil	1
Amarjeet Amle	1

Above table shows the list of artist according to country collectively and we can select the artist with their stats.

In []:

In []:

In [508]: df.head()

Out[508]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	dura
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	Sea:
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	Se:
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	Se:
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	Sea:

In [509]: df_tv=df[df['type']=='TV Show']

In [510]: df_tv

Out[510]:

	show_id	type	title	director	cast	count
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	Na
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	Na
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	Ind

In [511]: df_tv['duration']=df_tv['duration'].apply(lambda x:str(x).split(' '))

C:\Users\hp\AppData\Local\Temp\ipykernel_7956\1438972902.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

df_tv['duration']=df_tv['duration'].apply(lambda x:str(x).split(' '))

In [512]: df_tv

		Snow	Snow	Devonsnire	Paul Ho...	Kingac...
10	s11	TV Show	Vendetta: Truth, Lies and The Mafia	NaN	NaN	Na
11	s12	TV Show	Bangkok Breaking	Kongkiat Komesiri	Sukollawat Kanarot, Sushar Manaying, Pavarit M...	Na
14	s15	TV Show	Crime Stories: India Detectives	NaN	NaN	Na
15	s16	TV Show	Dear White People	NaN	Logan Browning, Brandon P. Bell, DeRon Horton,...	United States

In [514]: `df_tv['duration']=df_tv['duration'].apply(lambda x:int(x[0]))`

C:\Users\hp\AppData\Local\Temp\ipykernel_7956\4148963709.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

`df_tv['duration']=df_tv['duration'].apply(lambda x:int(x[0]))`

In [515]: `df_tv`

Out[515]:

	show_id	type	title	director	cast	count
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Afr...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	Na...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	Na...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	Ind...

In [518]: `df_tv["duration"].mean()`

Out[518]: 1.764947683109118

In []:

In []:

In [519]: df

Out[519]:

	show_id	type	title	director	cast	country	date_added	release_year	rating
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Nan	United States	September 25, 2021	2020	P
1	s2	TV Show	Blood & Water	Nan	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	T
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	Nan	September 24, 2021	2021	T
3	s4	TV Show	Jailbirds New Orleans	Nan	Nan	Nan	September 24, 2021	2021	T
4	s5	TV Show	Kota Factory	Nan	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	T
...
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	rating	
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	F	
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chan...	India	March 2, 2019	2015	TV-	

8804 rows × 12 columns

In []: