```
import pandas as pd
import matplotlib
import matplotlib.pyplot as plt
import seaborn as sns

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

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	list
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documer
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	Interna TV Show Dram Mys
					Sami Bouajila						Cri

```
display(df.describe())
print('-----------------------------')
```

	release_year
count	8807.000000
mean	2014.180198
std	8.819312
min	1925.000000
25%	2013.000000
50%	2017.000000
75%	2019.000000
max	2021.000000

```
-----------------------------

display(df.tail())
print('-----------------------------')
display(df.dtypes)
print('-----------------------------')
display(df.isnull().sum())
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	durati
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 n
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7	Seaso
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 n
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 n
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	111 n

```
-----------------------------
show_id         object
type            object
title           object
director        object
cast            object
country         object
date_added      object
release_year     int64
rating          object
```

```
df['director']=df['director'].fillna('unknown')
df['cast']=df['cast'].fillna('unknown')
df['country']=df['country'].fillna('unknown')
df['rating']=df['rating'].fillna(0)
df.dropna(inplace=True)
df.isnull().sum()
```

```
show_id         0
type            0
title           0
director        0
cast            0
country         0
date_added      0
release_year    0
rating          0
duration        0
listed_in       0
description     0
dtype: int64
```

```python
df.drop_duplicates(inplace=True)
df.shape
```

```
(8794, 12)
```

```python
movie= df[df['type'] == 'Movie']
tvshow= df[df['type'] == 'TV Show']
display(movie.shape)
display(tvshow.shape)
```

```
(6128, 12)
(2666, 12)
```

```python
display(tvshow.head())
print('--------------------------')
display(movie.head())
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration
1	s2	TV Show	Blood & Water	unknown	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	unknown	September 24, 2021	2021	TV-MA	1 Season
3	s4	TV Show	Jailbirds New Orleans	unknown	unknown	unknown	September 24, 2021	2021	TV-MA	1 Season
4	s5	TV Show	Kota Factory	unknown	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater,	unknown	September 24, 2021	2021	TV-MA	1 Season

```python
count=df['type'].value_counts()
display(count)
```

```
Movie      6128
TV Show    2666
Name: type, dtype: int64
```

```python
import plotly.express as px
fig=px.histogram(data_frame=df,x='type',title='No.of Movies and TV shows in netflix')
fig.show()
```

No.of Movies and TV shows in netflix

```
df1=df['release_year'].value_counts().sort_index()
df1
```

```
    1925       1
    1942       2
    1943       3
    1944       3
    1945       4
              ...
    2017    1031
    2018    1146
    2019    1030
    2020     953
    2021     592
    Name: release_year, Length: 74, dtype: int64
```

```
fig_hist=px.histogram(data_frame=df,x='release_year',nbins=8807,title='number of Movies and TV shows released each year')
fig_hist.show()
```

number of Movies and TV shows released each year
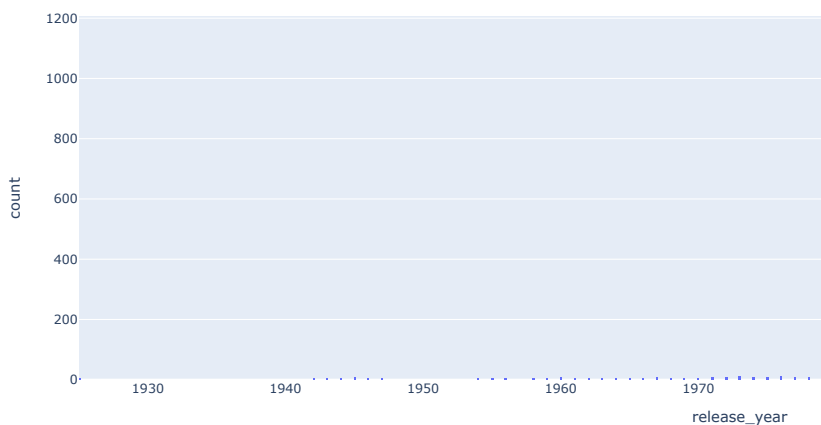

```
country_counts=df['country'].value_counts().head(15).plot(kind='barh',color=['red','yellow'],title='Movies and TV shows available in each country')
display('country_count')
```
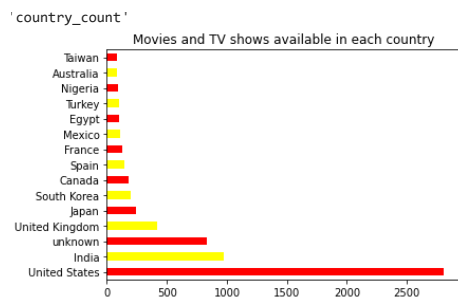
'country_count'


```
movie_count=df['country'].where(df['type']=='Movie')
movie_count.value_counts().head(15).plot(kind='barh',color=['black','purple'],title='Movies available in each country')
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f51d21dff10>
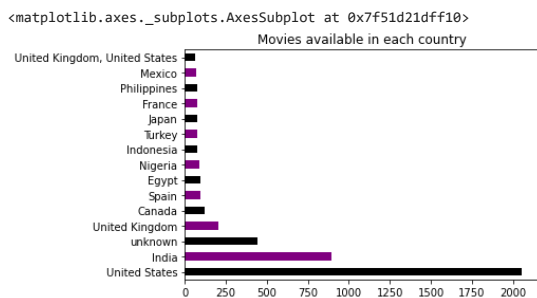

```
TvShows_count=df['country'].where(df['type']=='TV Show')
TvShows_count.value_counts().head(15).plot(kind='barh',color=['orange','blue'],title='TV shows available in each country')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f51d1fd7280>
                 TV shows available in each country
   Brazil
   Turkey
   China
   Mexico
   Australia
```

```
df['cast'].value_counts().head(11)
```

```
unknown                                                                                            825
David Attenborough                                                                                  19
Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna Bhardwaj, Rajesh Kava, Mousam, Swapnil            14
Samuel West                                                                                         10
Jeff Dunham                                                                                          7
Michela Luci, Jamie Watson, Eric Peterson, Anna Claire Bartlam, Nicolas Aqui, Cory Doran, Julie Lemieux, Derek McGrath    6
Craig Sechler                                                                                        6
David Spade, London Hughes, Fortune Feimster                                                         6
Kevin Hart                                                                                           6
Jim Gaffigan                                                                                         5
Bill Burr                                                                                            5
Name: cast, dtype: int64
```

```
df['cast'].value_counts().head(10)
top_ten=pd.DataFrame(df['cast']).value_counts().head(10)
top_ten
```

```
cast
unknown                                                                                            825
David Attenborough                                                                                  19
Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna Bhardwaj, Rajesh Kava, Mousam, Swapnil            14
Samuel West                                                                                         10
Jeff Dunham                                                                                          7
Kevin Hart                                                                                           6
Craig Sechler                                                                                        6
David Spade, London Hughes, Fortune Feimster                                                         6
Michela Luci, Jamie Watson, Eric Peterson, Anna Claire Bartlam, Nicolas Aqui, Cory Doran, Julie Lemieux, Derek McGrath    6
Jim Gaffigan                                                                                         5
dtype: int64
```

```
df['director'].value_counts().head(5)
top_five=pd.DataFrame(df['director']).value_counts().head(5)
top_five
```

```
director
unknown                  2624
Rajiv Chilaka              19
Raúl Campos, Jan Suter     18
Marcus Raboy               16
Suhas Kadav                16
dtype: int64
```