

```
In [3]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
import re
import plotly.express as px
import plotly.graph_objs as go
import plotly.figure_factory as ff
warnings.filterwarnings('ignore')
import plotlib

In [4]: df=pd.read_csv('E:/Data Analyst/Project/netflix_titles.csv')
```

```
In [8]: df.head()
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, film...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mababane, Thabani...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town l...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabil...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lot...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, firations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jilendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train l...

```
In [12]: df.tail()
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-V7	2 Seasons	Kids TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanai...	India	March 2, 2019	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...

```
In [15]: df.shape
Out[15]: (8807, 12)
In [18]: df.columns
Out[18]: Index(['show_id', 'type', 'title', 'director', 'cast', 'country', 'date_added', 'release_year', 'rating', 'duration', 'listed_in', 'description'], dtype='object')
In [19]: 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        7862 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 [20]: df.describe()
```

	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

inspect misssing values in the Dataset

```
In [26]: df.isnull().sum().sort_values(ascending=False)
```

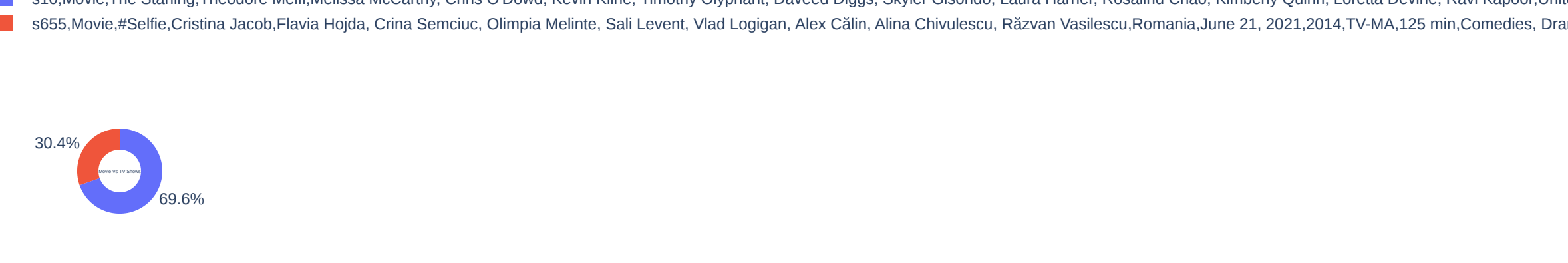
```
Out[26]: director      2634
country         831
cast            825
date_added      10
rating           4
duration         3
show_id          0
type            0
title           0
release_year     0
listed_in        0
description      0
dtype: int64
```

```
In [28]: df['director'].value_counts().head(10)
```

```
Out[28]: director
Rajiv Chilaka      19
Raul Campos, Jan Suter      18
Marcus Raboy       16
Suhaz Kadav        16
Jay Naras          14
Cathy Garcia-Molina 13
Martin Scorsese     12
Youssef Chahine     12
Jay Chapman         12
Steven Spielberg    11
Name: count, dtype: int64
```

Movies vs Tv Shows

```
In [17]: go.Figure(data=[go.Pie(labels=df.value_counts(normalize=True).index,
values=df.type.value_counts(normalize=True).values, hole=.5,
title='Movie Vs Tv Shows')])
```



```
In [41]: df.type.value_counts()
```

```
Out[41]: type
Movie      6181
TV Show    2676
Name: count, dtype: int64
```

```
In [42]: df.rating.value_counts()
```

```
Out[42]: rating
TV-MA      3287
TV-14      2168
TV-PG      863
R           799
PG-13      499
TV-Y7      334
TV-Y       387
PG         287
TV-G       229
NR          86
G           41
TV-Y7-FV   6
NC-17      3
UR          3
74 min     1
84 min     1
66 min     1
Name: count, dtype: int64
```

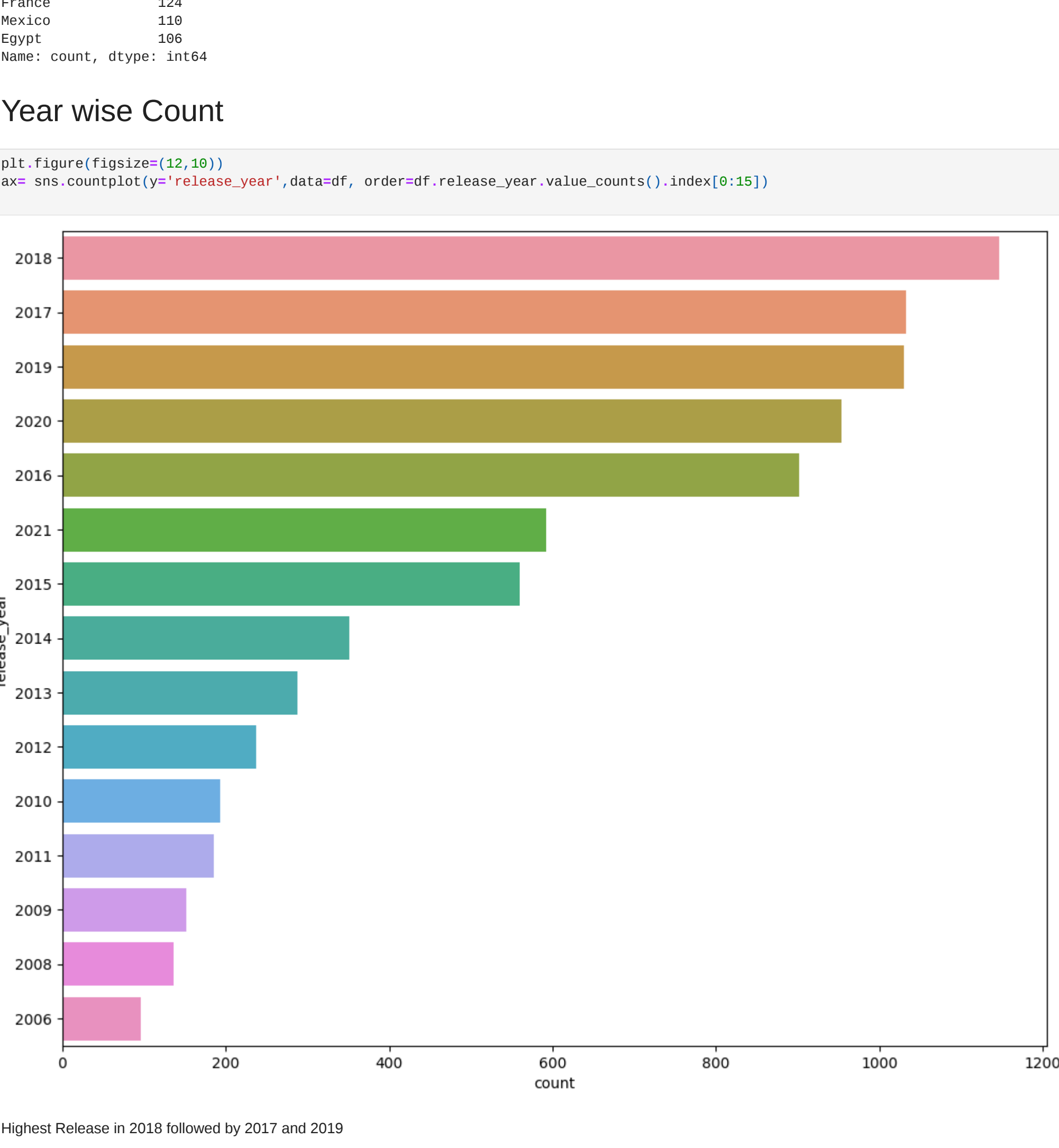


```
In [50]: df.country.value_counts().head(10)
```

```
Out[50]: country
United States      2818
India              972
United Kingdom     419
Japan              245
South Korea        199
Canada            181
Spain              145
France            124
Mexico            119
Egypt             106
Name: count, dtype: int64
```

Year wise Count

```
In [53]: plt.figure(figsize=(12,10))
ax=sns.countplot(y='release_year',data=df, order=df.release_year.value_counts().index[0:15])
```

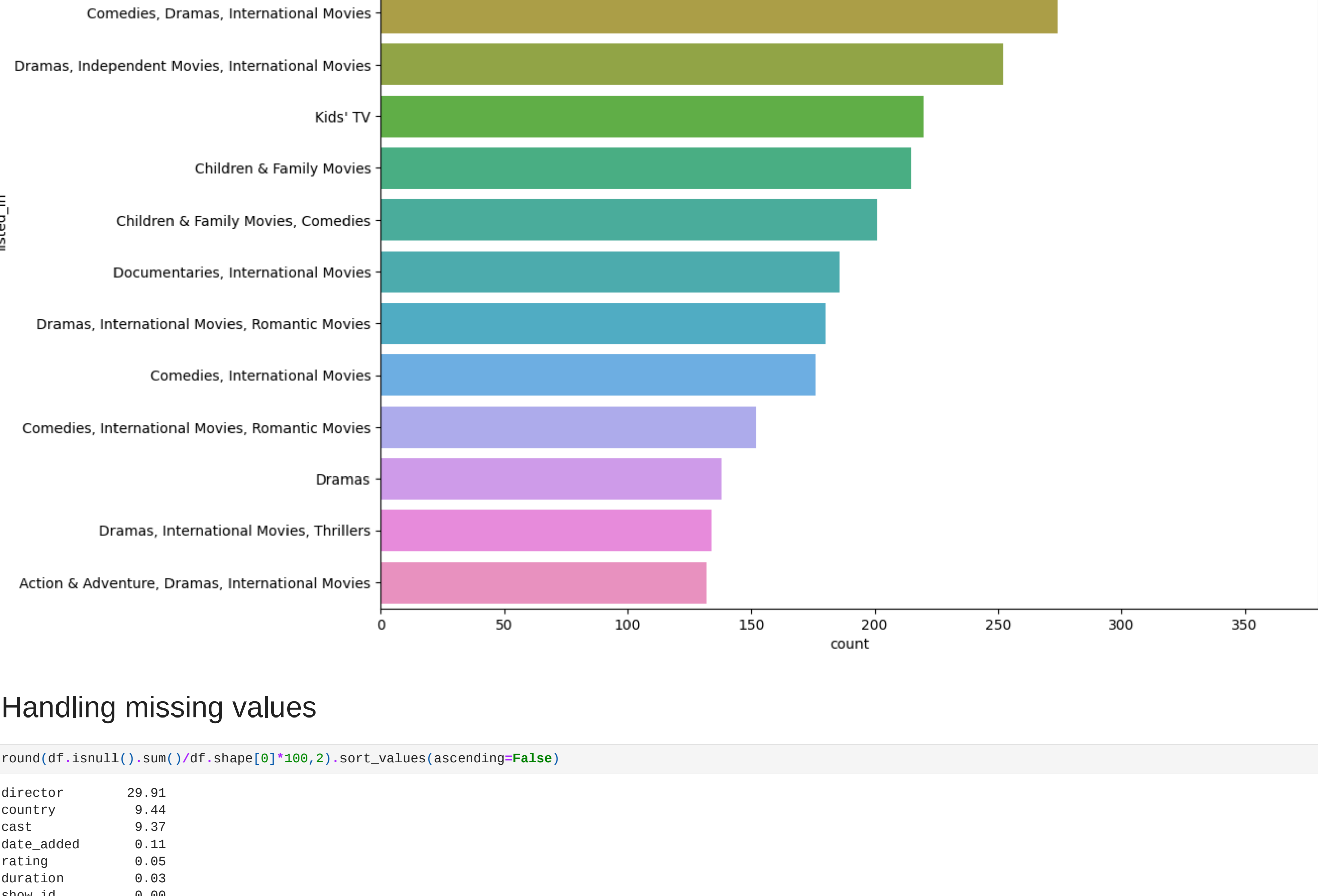


Highest Release in 2018 followed by 2017 and 2019

```
In [55]: df.listed_in.value_counts().head(10)
```

```
Out[55]: listed_in
Dramas, International Movies      362
Documentaries                    359
Stand-Up Comedy                  334
Comedies, Dramas, International Movies 274
Dramas, Independent Movies, International Movies 252
Kids' TV                          220
Children & Family Movies           215
Children & Family Movies, Comedies 191
Documentaries, International Movies 186
Dramas, International Movies, Romantic Movies 169
Name: count, dtype: int64
```

```
In [60]: plt.figure(figsize=(12,10))
ax=sns.countplot(y='listed_in',data=df,order=df.listed_in.value_counts().index[0:15])
```



Handling missing values

```
In [65]: round(df.isnull().sum())/df.shape[0]*100,2).sort_values(ascending=False)
```

```
Out[65]: director      29.91
country         9.44
cast            9.37
date_added      0.11
rating          0.05
duration        0.03
show_id         0.00
type            0.00
title           0.00
release_year    0.00
listed_in       0.00
description     0.00
dtype: float64
```

```
In [67]: round(df.isnull().sum())
```

```
Out[67]: show_id      0
type            0
title           0
director      2634
cast          825
country       831
date_added    10
rating        4
duration      3
listed_in     0
description   0
dtype: int64
```

```
In [69]: df.cast.value_counts().head(10)
```

```
Out[69]: cast
David Attenborough      19
Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna Bhardwaj, Rajesh Kava, Mousam, Swapnil Samuel West      10
Jeff Dunham              7
David Spade, London Hughes, Fortune Feimster      6
Kevin Hart                6
Craig Sechler             6
Michela Luci, Jamie Watson, Eric Peterson, Anna Claire Bartlam, Nicolas Agui, Cory Doran, Julie Lemieux, Derek McGrath      6
Bill Burr                 5
Iliza Shlesinger          5
Name: count, dtype: int64
```

```
In [71]: ## replace missing values in a cast with "No Cast"
df['cast'].replace(np.NaN,'No Cast',inplace=True)
```

```
In [76]: ## replace missing values Director with "No Director"
df['director'].replace(np.NaN,'No Director',inplace=True)
```

```
In [77]: round(df.isnull().sum())/df.shape[0]*100,2).sort_values(ascending=False)
```

```
Out[77]: country      9.44
date_added      0.11
rating          0.05
duration        0.03
show_id         0.00
type            0.00
title           0.00
director        0.00
cast            0.00
release_year    0.00
listed_in       0.00
description     0.00
dtype: float64
```

```
In [78]: df['title']
```

```
Out[78]: 0      Dick Johnson Is Dead
1      Blood & Water
2      Ganglands
3      Jailbirds New Orleans
4      Kota Factory
...
8802     Zodiac
8803     Zombie Dumb
8804     Zombieland
8805     Zoom
8806     Zubaan
Name: title, Length: 8807, dtype: object
```

```
In [88]: cast_shows=df[df.cast!='No Cast'].set_index('title').cast.str.split(', ', expand=True).stack().reset_index(level=1,drop=True)
plt.figure(figsize=(13,7))
plt.title('Top 10 Actor Movies based on the no of titles')
sns.countplot(y=filtered_cast_show, order=cast_show.value_counts().index[:10],palette='pastel')
plt.show()
```

```
Cell In[88], line 1
cast_shows=df[df.cast!='No Cast'].set_index('title').cast.str.split(', ', expand=True).stack().reset_index(level=1,drop=True)
```

```
SyntaxError: invalid syntax. Maybe you meant '=' or ':' instead of '!='?
```

```
In [94]: movies_df= df.loc[(df['type']=='Movie')]
movies_df
```

```
Out[94]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	No Cast	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, film...
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Lloba	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2021	PG	81 min	Children & Family Movies	Equestria's divided. But a high-eyed hero be...
7	s8	Movie	Sarkofa	Haile Gerima	Kofi Ghahabla, Oyulunike Ogundimu, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	September 24, 2021	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...
9	s10	Movie	The Starling	Therodora Melli	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	September 24, 2021	2021	PG-13	104 min	Comedies, Dramas	A woman adjusting to life after a loss contend...
12	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Weder, Janis Hawndner, Milan Peschel, ...	Germany, Czech Republic	September 23, 2021	2021	TV-MA	127 min	Dramas, International Movies	After most of her family is murdered in a ter...
...
8801	s8802	Movie	Zirzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yassa, Ali Al-Jabri, ...	United Arab Emirates, Jordan	March 9, 2016	2015	TV-MA	96 min	Dramas, International Movies, Thrillers	Recovering alcoholic Taha wakes up inside a s...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanai...	India	March 2, 2019	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...

6131 rows × 12 columns

```
In [10]: shows_df= df.loc[(df['type']=='TV Show')]
shows_df
```

```
Out[10]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mababane, Thabani...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town l...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabil...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lot...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, firations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jilendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train l...
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	NaN	September 24, 2021	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries	The arrival of a charismatic young priest b...
...
8795	s8796	TV Show	Yu-Gi-Oh! Arc-V	NaN	Mike Liscio, Emily Bauer, Billy Bob Thompson, ...	Japan, Canada	May 1, 2018	2015	TV-Y7	2 Seasons	Anime Series, Kids' TV	Now that he's discovered the Pendulum Summonin...
8796	s8797	TV Show	Yusuf Erme	NaN	Gökhan Altay, Payidar Tokelçioğlu, Baran Akbu...	Turkey	January 17, 2017	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas	During the Mongol invasions, Yurus Erme leaves...
8797	s8798	TV Show	Zak Storm	NaN	Michael Johnston, Jessica Gee, George Chisholm, ...	United States, France, South Korea, Indonesia	September 13, 2019	2016	TV-Y7	3 Seasons	Kids' TV	Teen surfer Zak Storm is mysteriously transpor...
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saad, Fawad Khan, Ayesha Omar, Mehreen ...	Pakistan	December 15, 2016	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Karachi and cafe...
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...

2676 rows × 12 columns

```
In [ ]:
In [ ]:
```