

Problem Statement

1. Understanding what content is available on netflix
2. Based on observations analysing which shows to produce and how to grow business
3. Understanding how number of movies release in the recent years changed
4. Comparison of TV shows and Movies
5. Analysis of casts and Directors
6. Content analysis in different countries

```
In [1748]: import numpy as np  
import pandas as pd  
import matplotlib.pyplot as plt  
import seaborn as sns
```

```
In [1749]: data = pd.read_csv(r"E:\DSML\Projects\1\netflix.csv")
```

```
In [1750]: data.head()
```

Out[1750]:

	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, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	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 t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	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 lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra 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 I...

```
In [1751]: data.shape
```

```
Out[1751]: (8807, 12)
```

In the given Netflix data there are 8807 rows and 12 columns

```
In [1752]: data.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
```

There are 12 columns, out of them 11 columns are of 'Object' datatype.

We need to convert date_added column to dateTime format.

Here we can observe columns --> director, cast, country, date_added, rating, duration contains null values

```
In [1753]: data.describe(include= 'all')
```

Out[1753]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
count	8807	8807	8807	6173	7982	7976	8797	8807.000000	8803	8804	8807	8807
unique	8807	2	8807	4528	7692	748	1767	NaN	17	220	514	8775
top	s1	Movie	Dick Johnson Is Dead	Rajiv Chilaka	David Attenborough	United States	January 1, 2020	NaN	TV-MA	1 Season	Dramas, International Movies	Paranormal activity at a lush, abandoned prop...
freq	1	6131		19	19	2818	109	NaN	3207	1793	362	4
mean	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2014.180198	NaN	NaN	NaN	NaN
std	NaN	NaN	NaN	NaN	NaN	NaN	NaN	8.819312	NaN	NaN	NaN	NaN
min	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1925.000000	NaN	NaN	NaN	NaN
25%	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2013.000000	NaN	NaN	NaN	NaN
50%	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2017.000000	NaN	NaN	NaN	NaN
75%	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2019.000000	NaN	NaN	NaN	NaN
max	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2021.000000	NaN	NaN	NaN	NaN

From the above description We can see that basic Overview of the data.

Here we can see Movie is the most frequently coming data with almost 6131 times. According to release year of data in the given dataset 1925 is least year which we can see. TV MA which is Mature content has appeared most of the times in data. And Country which occurs very certainly is United States.

```
In [1754]: data.columns
```

Out[1754]: Index(['show_id', 'type', 'title', 'director', 'cast', 'country', 'date_added', 'release_year', 'rating', 'duration', 'listed_in', 'description'], dtype='object')

Here total 12 columns are present for given dataset which are 'show_id', 'type', 'title', 'director', 'cast', 'country', 'date_added', 'release_year', 'rating', 'duration', 'listed_in', 'description' respectively.

```
In [ ]:
```

```
In [1755]: data.nunique()
```

```
Out[1755]: show_id      8807  
type          2  
title         8807  
director      4528  
cast          7692  
country        748  
date_added    1767  
release_year   74  
rating         17  
duration       220  
listed_in      514  
description    8775  
dtype: int64
```

Here all the show id's are unique.

And 2 types of content is present in the dataset. So we can convert it to type category.

All the titles in the dataset are Unique. And for Director, cast, country and listed in we can observe data in list format so we need to split it as per the requirement

```
In [1756]: data.isna().sum()
```

```
Out[1756]: show_id      0  
type          0  
title         0  
director     2634  
cast          825  
country       831  
date_added    10  
release_year   0  
rating         4  
duration       3  
listed_in      0  
description    0  
dtype: int64
```

Here we are getting missing values in director, cast, country, date_added, rating and duration columns.

```
In [1757]: data.isna().sum() / data.shape[0] * 100
```

```
Out[1757]: show_id      0.000000
type          0.000000
title         0.000000
director      29.908028
cast          9.367549
country        9.435676
date_added    0.113546
release_year   0.000000
rating         0.045418
duration       0.034064
listed_in      0.000000
description    0.000000
dtype: float64
```

1. Here we can see about 30 % of Director cells are null --> As this is greater amount we will deal with it by using Mode value based on other columns. For cast also we will do the same.
2. For country we will replace NA values with mode value of country column.

In [1758]: `data[data['country'].isna()]`

Out[1758]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	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 lord...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down among...
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 brings...
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	91 min	Children & Family Movies	Equestria's divided. But a bright-eyed hero becomes...
10	s11	TV Show	Vendetta: Truth, Lies and The Mafia	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, Docuseries, International TV Shows...	Sicily boasts a bold "Anti-Mafia" coalition. But...
...
8718	s8719	Movie	Westside vs. the World	Michael Fahey	Ron Perlman, Louie Simmons	NaN	August 9, 2019	2019	TV-MA	96 min	Documentaries, Sports Movies	A look into the journey of influential strength...
8759	s8760	Movie	World's Weirdest Homes	Storm Theunissen	Charlie Luxton	NaN	February 1, 2019	2015	TV-PG	49 min	Movies	From a bubble-shaped palace to an island built...

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	
8783	s8784	TV Show	Yoko	NaN	Eileen Stevens, Alyson Leigh Rosenfeld, Sarah ...	NaN	June 23, 2018	2016	TV-Y	1 Season	Kids' TV	Friends Mai, Oto and Vik's games at the park b...
8785	s8786	TV Show	YOM	NaN	Sairaj, Devyani Dagaonkar, Ketan Singh, Mayur ...	NaN	June 7, 2018	2016	TV-Y7	1 Season	Kids' TV	With the mind of a human being, and the body o...
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...

831 rows × 12 columns

In [1759]: `data['country'].mode()`

Out[1759]: 0 United States
dtype: object

Here we can see that mode value for country is United States so we can use that for na values

In [1760]: `data['country'].fillna("United States", inplace= True)`

In []:

In [1761]: `data[data['date_added'].isna()]`

Out[1761]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
6066	s6067	TV Show	A Young Doctor's Notebook and Other Stories	NaN	Daniel Radcliffe, Jon Hamm, Adam Godley, Chris...	United Kingdom	NaN	2013	TV-MA	2 Seasons	British TV Shows, TV Comedies, TV Dramas	Set during the Russian Revolution, this comic ...
6174	s6175	TV Show	Anthony Bourdain: Parts Unknown	NaN	Anthony Bourdain	United States	NaN	2018	TV-PG	5 Seasons	Docuseries	This CNN original series has chef Anthony Bour...
6795	s6796	TV Show	Frasier	NaN	Kelsey Grammer, Jane Leeves, David Hyde Pierce...	United States	NaN	2003	TV-PG	11 Seasons	Classic & Cult TV, TV Comedies	Frasier Crane is a snooty but lovable Seattle ...
6806	s6807	TV Show	Friends	NaN	Jennifer Aniston, Courteney Cox, Lisa Kudrow, ...	United States	NaN	2003	TV-14	10 Seasons	Classic & Cult TV, TV Comedies	This hit sitcom follows the merry misadventure...
6901	s6902	TV Show	Gunslinger Girl	NaN	Yuuka Nanri, Kanako Mitsuhashi, Eri Sendai, Am...	Japan	NaN	2008	TV-14	2 Seasons	Anime Series, Crime TV Shows	On the surface, the Social Welfare Agency appe...
7196	s7197	TV Show	Kikoriki	NaN	Igor Dmitriev	United States	NaN	2010	TV-Y	2 Seasons	Kids' TV	A wacky rabbit and his gang of animal pals hav...
7254	s7255	TV Show	La Familia P. Luche	NaN	Eugenio Derbez, Consuelo Duval, Luis Manuel Áv...	United States	NaN	2012	TV-14	3 Seasons	International TV Shows, Spanish-Language TV Sh...	This irreverent sitcom features Ludovico, Feder...
7406	s7407	TV Show	Maron	NaN	Marc Maron, Judd Hirsch, Josh Brener, Nora Zeh...	United States	NaN	2016	TV-MA	4 Seasons	TV Comedies	Marc Maron stars as Marc Maron, who interviews...
7847	s7848	TV Show	Red vs. Blue	NaN	Burnie Burns, Jason Saldaña, Gustavo Sorola, G...	United States	NaN	2015	NR	13 Seasons	TV Action & Adventure, TV Comedies, TV Sci-Fi ...	This parody of first-person shooter games, mil...

show_id	type		title	director		cast	country	date_added	release_year	rating	duration	listed_in	description
8182	s8183	TV Show	The Adventures of Figaro Pho	Nan	Luke Jurevicius, Craig Behenna, Charlotte Ham...	Australia	Nan	2015	TV-Y7	2 Seasons	Kids' TV, TV Comedies	Imagine your worst fears, then multiply them: ...	

In [1762]: `data['date_added'] = pd.DatetimeIndex(data['date_added'])`

Here we will convert date which is in object format to Datetime format and replace all the na values with mean value

In [1763]: `data['date_added'].mean()`

Out[1763]: `Timestamp('2019-05-17 05:59:08.436967168')`

In [1764]: `data['date_added'].fillna("2019-05-17", inplace= True)`

In [1765]: `data['year_added'] = data['date_added'].dt.year
data['month_added'] = data['date_added'].dt.month
data['day_added'] = data['date_added'].dt.day`

In [1766]: `data[data['duration'].isnull()]`

Out[1766]:

show_id	type		title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_
5541	s5542	Movie	Louis C.K. 2017	Louis C.K.	Louis C.K.	United States	2017-04-04	2017	74 min	Nan	Movies	Louis C.K. muses on religion, eternal love, gi...	2017		4
5794	s5795	Movie	Louis C.K.: Hilarious	Louis C.K.	Louis C.K.	United States	2016-09-16	2010	84 min	Nan	Movies	Emmy-winning comedy writer Louis C.K. brings h...	2016		9
5813	s5814	Movie	Live at the Comedy Store	Louis C.K.	Louis C.K.	United States	2016-08-15	2015	66 min	Nan	Movies	The comic puts his trademark hilarious/thought...	2016		8

Here we can clearly see that rating values assigned with the duration values only for the columns where duration value is NA. So we need to do swap between this columns

```
In [1767]: # if(data['duration'].isna()):
#     data[['duration','rating']] = data[['rating','duration']]
# Filter NAN Data selection column of strings by not operator.
data[data.duration.isna()].head()
```

Out[1767]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_
5541	s5542	Movie	Louis C.K. 2017	Louis C.K.	Louis C.K.	United States	2017-04-04	2017	74 min	NaN	Movies	Louis C.K. muses on religion, eternal love, gi...	2017		4
5794	s5795	Movie	Louis C.K.: Hilarious	Louis C.K.	Louis C.K.	United States	2016-09-16	2010	84 min	NaN	Movies	Emmy-winning comedy writer Louis C.K. brings h...	2016		9
5813	s5814	Movie	Live at the Comedy Store	Louis C.K.	Louis C.K.	United States	2016-08-15	2015	66 min	NaN	Movies	The comic puts his trademark hilarious/thought...	2016		8

```
In [1768]: data.loc[data['director'] == 'Louis C.K.',['duration'] = data['rating']
data[data.director == 'Louis C.K.'].head()
```

Out[1768]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_
5541	s5542	Movie	Louis C.K. 2017	Louis C.K.	Louis C.K.	United States	2017-04-04	2017	74 min	74 min	Movies	Louis C.K. muses on religion, eternal love, gi...	2017		4
5794	s5795	Movie	Louis C.K.: Hilarious	Louis C.K.	Louis C.K.	United States	2016-09-16	2010	84 min	84 min	Movies	Emmy-winning comedy writer Louis C.K. brings h...	2016		9
5813	s5814	Movie	Louis C.K.: Live at the Comedy Store	Louis C.K.	Louis C.K.	United States	2016-08-15	2015	66 min	66 min	Movies	The comic puts his trademark hilarious/thought...	2016		8

```
In [1769]: data.loc[data['director'] == 'Louis C.K.', 'rating'] = 'Unavailable'
data[data.director == 'Louis C.K.'].head()
```

Out[1769]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
5541	s5542	Movie	Louis C.K. 2017	Louis C.K.	Louis C.K.	United States	2017-04-04	2017	Unavailable	74 min	Movies	Louis C.K. muses on religion, eternal love, gi...	2017	4
5794	s5795	Movie	Louis C.K.: Hilarious	Louis C.K.	Louis C.K.	United States	2016-09-16	2010	Unavailable	84 min	Movies	Emmy-winning comedy writer Louis C.K. brings h...	2016	9
5813	s5814	Movie	Louis C.K.: Live at the Comedy Store	Louis C.K.	Louis C.K.	United States	2016-08-15	2015	Unavailable	66 min	Movies	The comic puts his trademark hilarious/thought...	2016	8

```
In [1770]: data[data['duration'].isnull()]
```

Out[1770]:

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
---------	------	-------	----------	------	---------	------------	--------------	--------	----------	-----------	-------------	------------	-------------	-----------

Type *Markdown* and *LaTeX*: α^2

```
In [1771]: data['rating'].mode()
```

Out[1771]: 0 TV-MA
dtype: object

```
In [1772]: data['rating'].fillna("TV-MA", inplace=True)
```

```
In [1773]: data[data['rating'] =='Unavailable']
```

Out[1773]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
5541	s5542	Movie	Louis C.K. 2017	Louis C.K.	Louis C.K.	United States	2017-04-04	2017	Unavailable	74 min	Movies	Louis C.K. muses on religion, eternal love, gi...	2017	4
5794	s5795	Movie	Louis C.K.: Hilarious	Louis C.K.	Louis C.K.	United States	2016-09-16	2010	Unavailable	84 min	Movies	Emmy-winning comedy writer Louis C.K. brings h...	2016	9
5813	s5814	Movie	Louis C.K.: Live at the Comedy Store	Louis C.K.	Louis C.K.	United States	2016-08-15	2015	Unavailable	66 min	Movies	The comic puts his trademark hilarious/thought...	2016	8



```
In [1774]: data.loc[data['rating'] == 'Unavailable', 'rating'] = data['rating'].mode()
```

```
In [1775]: data[data['rating'] =='Unavailable']
```

Out[1775]:

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
---------	------	-------	----------	------	---------	------------	--------------	--------	----------	-----------	-------------	------------	-------------	-----------

```
In [1776]: data[data['rating'].isna()]
```

Out[1776]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_
5541	s5542	Movie	Louis C.K. 2017	Louis C.K.	Louis C.K.	United States	2017-04-04	2017	NaN	74 min	Movies	Louis C.K. muses on religion, eternal love, gi...	2017		4
5794	s5795	Movie	Louis C.K.: Hilarious	Louis C.K.	Louis C.K.	United States	2016-09-16	2010	NaN	84 min	Movies	Emmy-winning comedy writer Louis C.K. brings h...	2016		9
5813	s5814	Movie	Louis C.K.: Live at the Comedy Store	Louis C.K.	Louis C.K.	United States	2016-08-15	2015	NaN	66 min	Movies	The comic puts his trademark hilarious/thought...	2016		8



```
In [ ]:
```

Replaced all the NA values for ratings column

In [1777]: `data.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 15 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     8807 non-null   object  
 6   date_added  8807 non-null   datetime64[ns]
 7   release_year 8807 non-null   int64   
 8   rating      8804 non-null   object  
 9   duration    8807 non-null   object  
 10  listed_in   8807 non-null   object  
 11  description 8807 non-null   object  
 12  year_added  8807 non-null   int64   
 13  month_added 8807 non-null   int64   
 14  day_added   8807 non-null   int64   
dtypes: datetime64[ns](1), int64(4), object(10)
memory usage: 1.0+ MB
```

In [1778]: *#Except director and cast all null values handled*

```
In [1779]: data.nunique()
```

```
Out[1779]: show_id      8807
type          2
title        8807
director     4528
cast         7692
country       748
date_added   1714
release_year  74
rating        14
duration      220
listed_in     514
description   8775
year_added    14
month_added   12
day_added     31
dtype: int64
```

Conversion of type and rating column to category

```
In [1780]: to_include = ['type', 'rating']
for x in to_include:
    data[x] = data[x].astype('category')
```

In [1781]: `data.head()`

Out[1781]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Nan	United States	2021-09-25	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, film...	2021	9	1
1	s2	TV Show	Blood & Water	Nan	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9	1
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...	2021	9	1
3	s4	TV Show	Jailbirds New Orleans	Nan	Nan	United States	2021-09-24	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...	2021	9	1
4	s5	TV Show	Kota Factory	Nan	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train I...	2021	9	1

```
In [1782]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 15 columns):
 #   Column      Non-Null Count  Dtype  
 --- 
 0   show_id     8807 non-null   object  
 1   type        8807 non-null   category
 2   title       8807 non-null   object  
 3   director    6173 non-null   object  
 4   cast        7982 non-null   object  
 5   country     8807 non-null   object  
 6   date_added  8807 non-null   datetime64[ns]
 7   release_year 8807 non-null   int64  
 8   rating      8804 non-null   category
 9   duration    8807 non-null   object  
 10  listed_in   8807 non-null   object  
 11  description 8807 non-null   object  
 12  year_added  8807 non-null   int64  
 13  month_added 8807 non-null   int64  
 14  day_added   8807 non-null   int64  
dtypes: category(2), datetime64[ns](1), int64(4), object(8)
memory usage: 912.6+ KB
```

```
In [1783]: data['type'].value_counts()
```

```
Out[1783]: Movie      6131
TV Show    2676
Name: type, dtype: int64
```

For type we have two unique values which are Movie and TV Show respectively

```
In [1784]: data['rating'].value_counts()
```

```
Out[1784]: TV-MA      3211  
TV-14       2160  
TV-PG        863  
R            799  
PG-13        490  
TV-Y7        334  
TV-Y         307  
PG           287  
TV-G          220  
NR            80  
G             41  
TV-Y7-FV      6  
NC-17         3  
UR            3  
Name: rating, dtype: int64
```

For rating we have -- >

1. TV-MA Mature content
2. TV-14 content for age 14 above
3. TV-PG Parental Guidance required content
4. R Restricted content(above 17)
5. PG-13 Parental Guidance required content
6. TV-Y7 Young Audience content
7. TV-Y Young Audience content
8. PG Parental Guidance required conten
9. TV-G For all ages
10. NR Not rated
11. G all ages
12. TV-Y7-FV 7 and older
13. NC-17 Adult
14. UR

```
In [1785]: len(data['rating'].value_counts())
```

```
Out[1785]: 14
```

We have around 14 unique values for rating

In []:

In []:

Analysis

Movie Vs TV show Analysis

```
In [1786]: typeDf = pd.DataFrame([data['type'].value_counts()])
typeDf
```

Out[1786]:

	Movie	TV Show
type	6131	2676

```
In [1787]: typeDf = pd.DataFrame(typeDf.stack()).reset_index()
typeDf
```

Out[1787]:

	level_0	level_1	0
0	type	Movie	6131
1	type	TV Show	2676

```
In [1788]: typeDf = typeDf[['level_1' , 0]]
```

```
In [1789]: typeDf.columns = ['type' , 'count']
typeDf
```

Out[1789]:

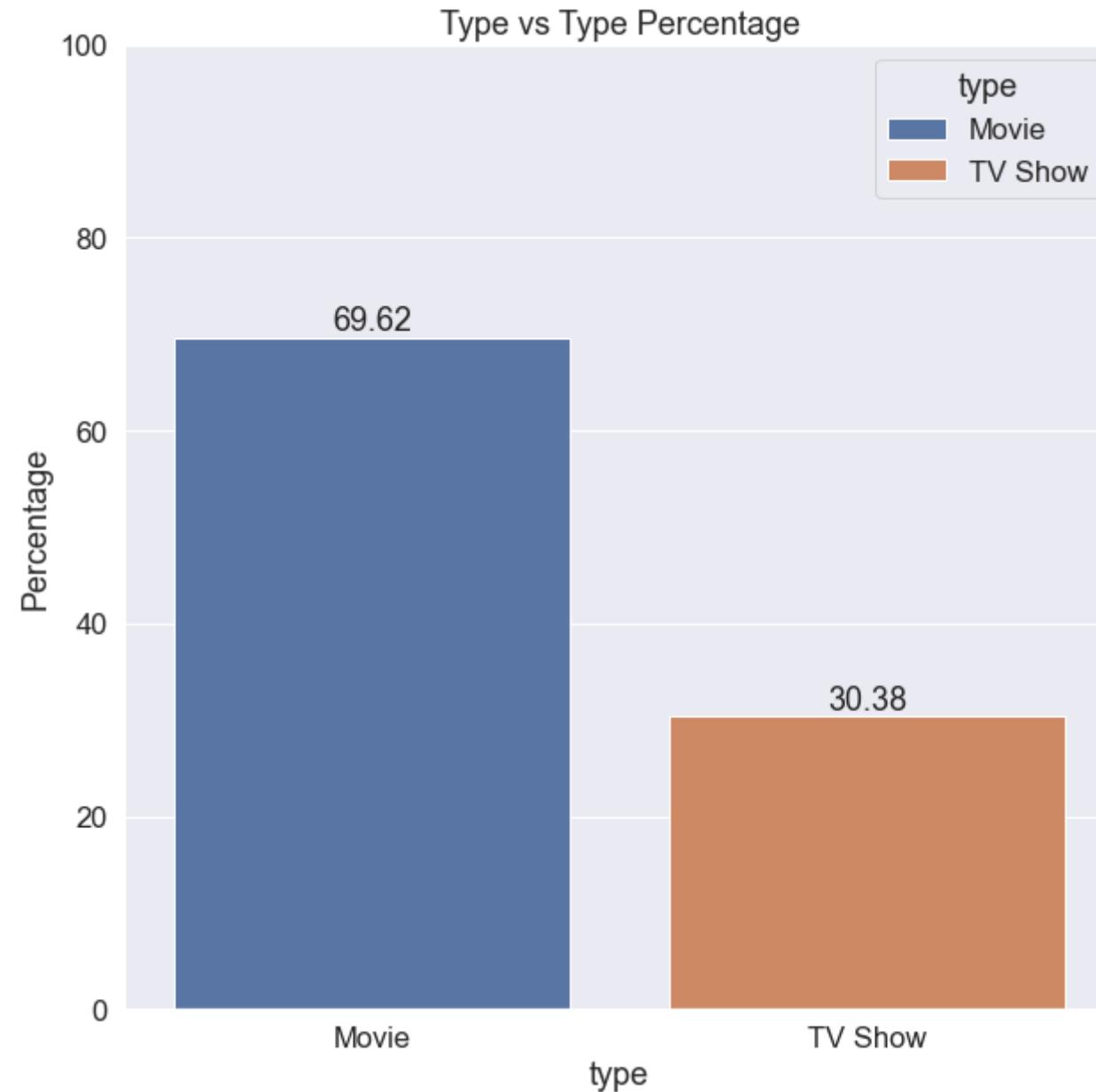
	type	count
0	Movie	6131
1	TV Show	2676

```
In [1790]: typeDf['Percentage'] = np.round(typeDf['count'] * 100 /typeDf['count'].sum() , 2)  
typeDf
```

Out[1790]:

	type	count	Percentage
0	Movie	6131	69.62
1	TV Show	2676	30.38

```
In [1791]: plt.figure(figsize = (10,10))
sns.set(font_scale = 1.5)
graph = sns.barplot(x = typeDf['type'] , y = typeDf['Percentage'], hue = typeDf['type'],dodge=False)
for i in graph.containers:
    graph.bar_label(i,)
plt.ylim(0,100)
plt.title("Type vs Type Percentage")
plt.show()
```



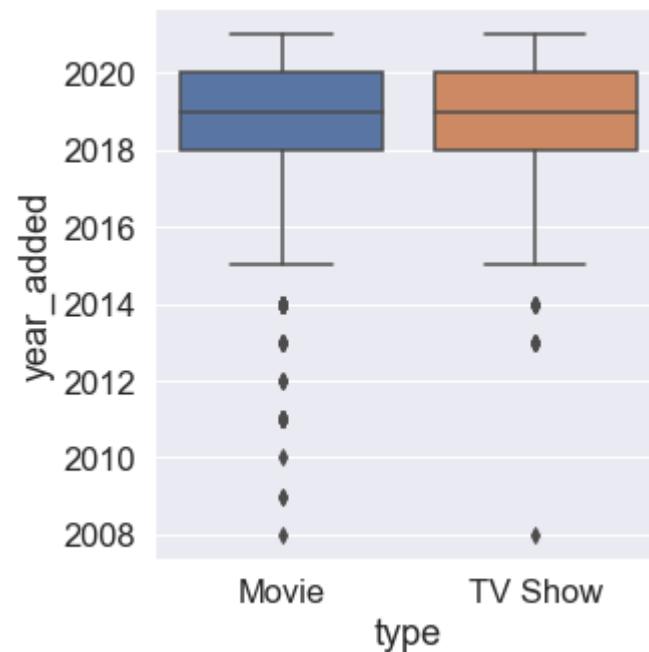
From the above graphical and Non Graphical data we can say that about 69.62 i.e., around 70 % are movies and remaining 30% are TV shows present on Netflix

In []:

Content type Analysis wrt to Year added on netflix

In [1792]:

```
plot = sns.catplot(x="type", y="year_added", kind="box", data=data)
```



From the above boxplot we can say that between 2018 to 2020 most of the content for movies and TV shows added on netflix

Let's see outliers for them

In [1793]:

```
def box_outliers(series):
    Q1 = series.quantile(0.25)
    Q3 = series.quantile(0.75)
    IQR = Q3 - Q1
    mask = (series < Q1 - 1.5 * IQR) | (series > Q3 + 1.5 * IQR)
    series = series[mask]
    return series
```

In [1794]:

```
data_outliers = pd.DataFrame([])
```

```
In [1795]: for i in data.copy().type.unique():
    df_outlier = data.copy().sort_values(['type'])
    df_outlier = df_outlier.loc[df_outlier['type'] == i, :].reset_index(drop=True)
    df_outlier = df_outlier[['type', 'year_added']].set_index('type').apply(box_outliers)
    df_outlier = df_outlier.reset_index()
    data_outliers = data_outliers.append(df_outlier)
```

```
In [1796]: data_outliers.reset_index(drop=True)
```

Out[1796]:

	type	year_added
0	Movie	2012
1	Movie	2008
2	Movie	2009
3	Movie	2009
4	Movie	2011
5	Movie	2011
6	Movie	2011
7	Movie	2011
8	Movie	2011
9	Movie	2011
10	Movie	2011
11	Movie	2011
12	Movie	2011
13	Movie	2011
14	Movie	2011
15	Movie	2011
16	Movie	2011
17	Movie	2012
18	Movie	2013
19	Movie	2013
20	Movie	2013
21	Movie	2014
22	Movie	2014
23	Movie	2014
24	Movie	2014
25	Movie	2014
26	Movie	2014

	type	year_added
27	Movie	2014
28	Movie	2014
29	Movie	2014
30	Movie	2014
31	Movie	2014
32	Movie	2014
33	Movie	2014
34	Movie	2013
35	Movie	2014
36	Movie	2013
37	Movie	2014
38	Movie	2014
39	Movie	2010
40	Movie	2014
41	Movie	2014
42	Movie	2013
43	Movie	2014
44	Movie	2012
45	TV Show	2014
46	TV Show	2013
47	TV Show	2014
48	TV Show	2013
49	TV Show	2013
50	TV Show	2008
51	TV Show	2014
52	TV Show	2014
53	TV Show	2013
54	TV Show	2013
55	TV Show	2014

From the above data we got outlier values for data added --> Movies and TV shows added before 2015 are Outliers

In []:

Content type analysis wrt to it's release year

```
In [1797]: data_releaseYear = pd.DataFrame(data['release_year'].value_counts())
data_releaseYear.reset_index(inplace= True)
data_releaseYear
```

Out[1797]:

	index	release_year
0	2018	1147
1	2017	1032
2	2019	1030
3	2020	953
4	2016	902
...
69	1959	1
70	1925	1
71	1961	1
72	1947	1
73	1966	1

74 rows × 2 columns

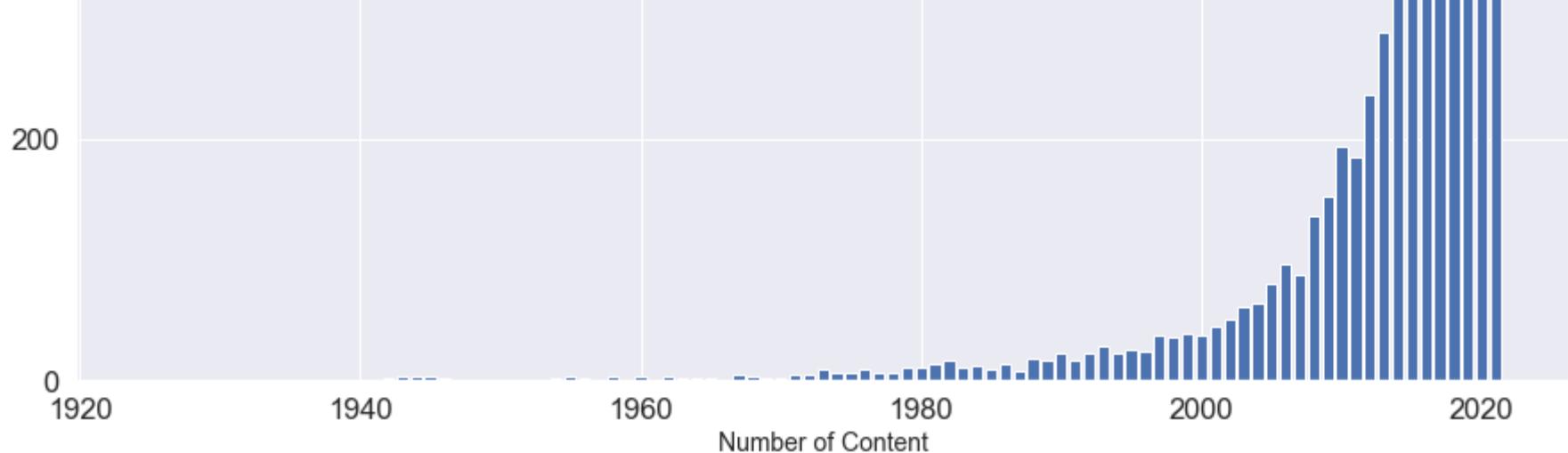
```
In [1798]: data_releaseYear.columns = ['release_year', 'count_content']
data_releaseYear.head()
```

Out[1798]:

	release_year	count_content
0	2018	1147
1	2017	1032
2	2019	1030
3	2020	953
4	2016	902

```
In [1799]: plt.figure(figsize=(15,15))
plt.bar(data_releaseYear['release_year'] ,data_releaseYear['count_content'])
plt.xlabel("Number of Content", fontsize=14)
plt.ylabel("Year", fontsize=14)
plt.title("Release Year vs Content Count ", fontsize=16)
plt.legend(data_releaseYear['count_content'])
plt.autoscale()
plt.show()
```





Here we can clearly see the content released after 2009 is growing

```
In [1800]: data_releaseYearAfter2009 = data_releaseYear[data_releaseYear['release_year'].between(2009,2021)]
```

```
In [1801]: data_releaseYearAfter2009
```

```
Out[1801]:
```

	release_year	count_content
0	2018	1147
1	2017	1032
2	2019	1030
3	2020	953
4	2016	902
5	2021	592
6	2015	560
7	2014	352
8	2013	288
9	2012	237
10	2010	194
11	2011	185
12	2009	152

```
In [1802]: TotalContent = data_releaseYear['count_content'].sum()
```

```
In [1803]: contentAfter2009 = data_releaseYearAfter2009['count_content'].sum()
```

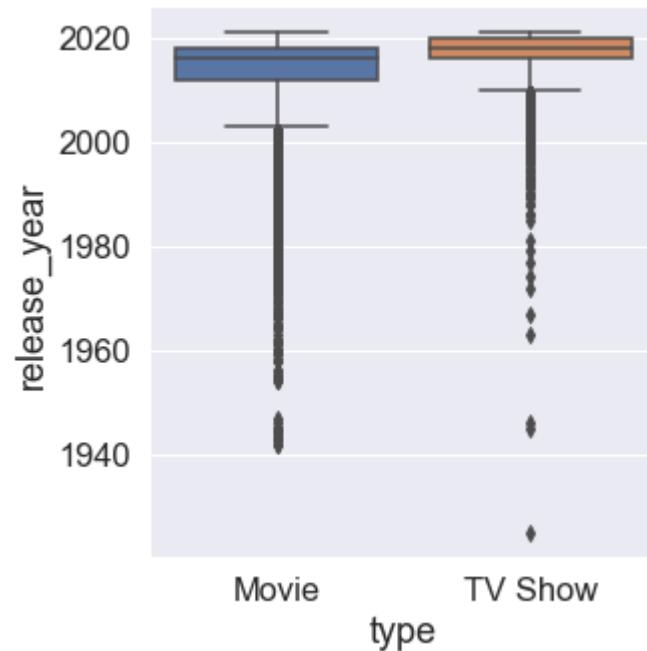
```
In [1804]: change = contentAfter2009 * 100 / TotalContent  
change
```

```
Out[1804]: 86.56750312251619
```

From above calculations we can say that around 87 % content on netflix is content released after 2009

In [1805]:

```
plot = sns.catplot(x="type", y="release_year", kind="box", data=data)
plt.show()
```

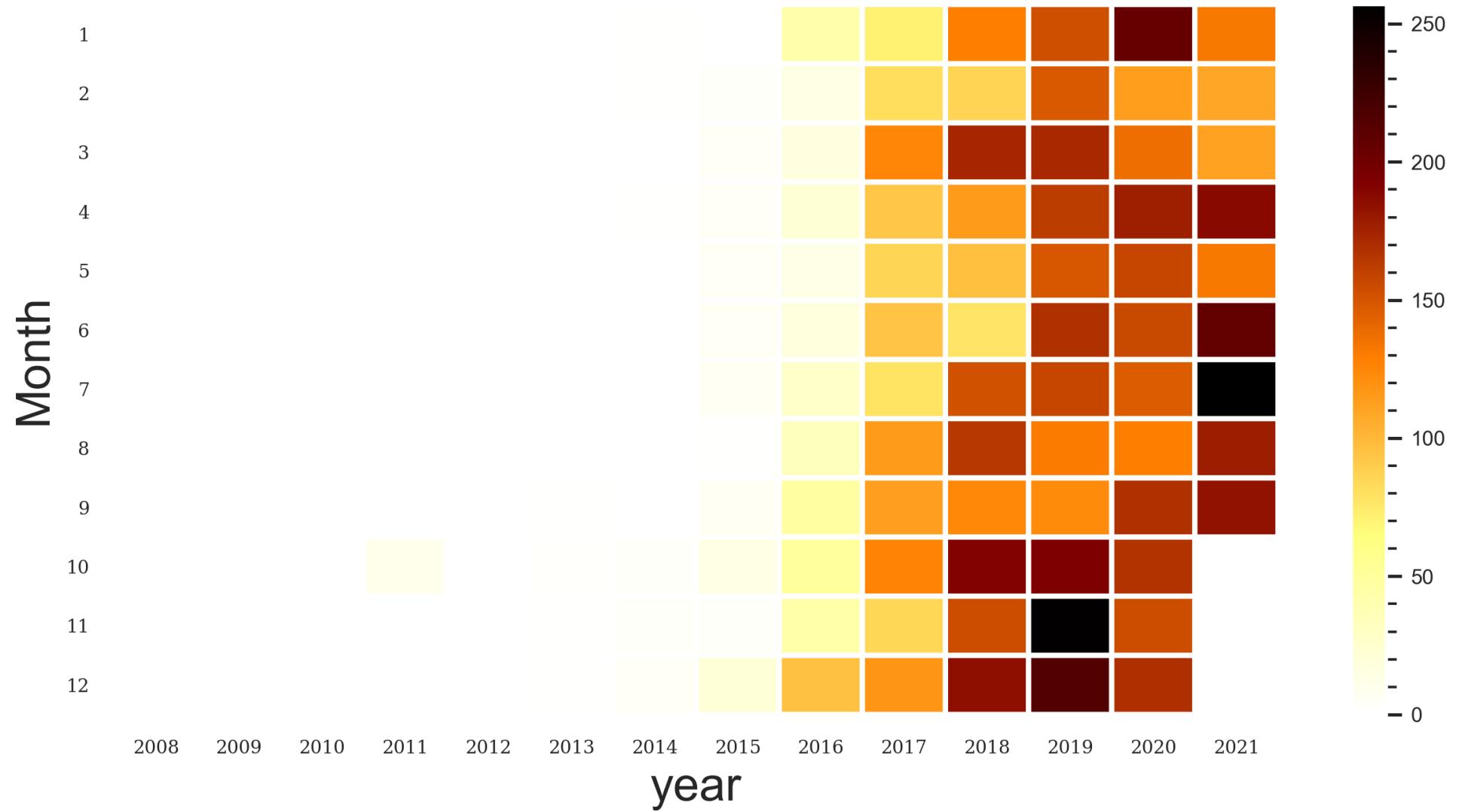


From the above boxplot we can say that content released before 2003/2004 is an outlier

Month Wise Content Added On netflix

```
In [1806]: month_order = [1,2,3,4,5,6,7,8,9,10,11,12][::-1]
month_HM = data.groupby('year_added')['month_added'].value_counts().unstack().fillna(0)[month_order].T
plt.figure(figsize=(10, 5), dpi=300)
plt.pcolor(month_HM, cmap='afmhot_r', edgecolors='white', linewidths=2) # heatmap
plt.xlabel("year")
plt.ylabel("Month")
plt.xticks(np.arange(0.5, len(month_HM.columns), 1), month_HM.columns, fontsize=7, fontfamily='serif')
plt.yticks(np.arange(0.5, len(month_HM.index), 1), month_HM.index, fontsize=7, fontfamily='serif')
plt.title('Netflix Monthly Contents Update', fontsize=12, fontweight='bold', position=(0.40, 1.0+0.02))
cbar = plt.colorbar()
cbar.ax.tick_params(labelsize=8)
cbar.ax.minorticks_on()
plt.show()
```

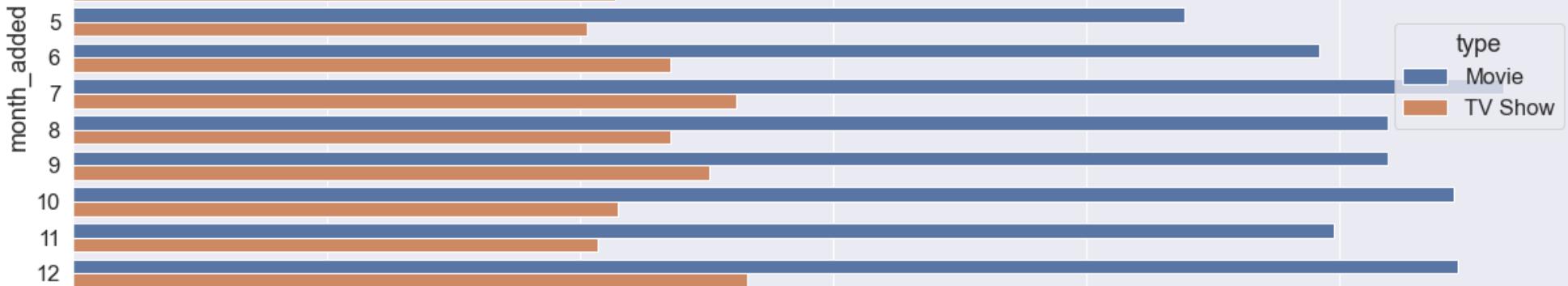
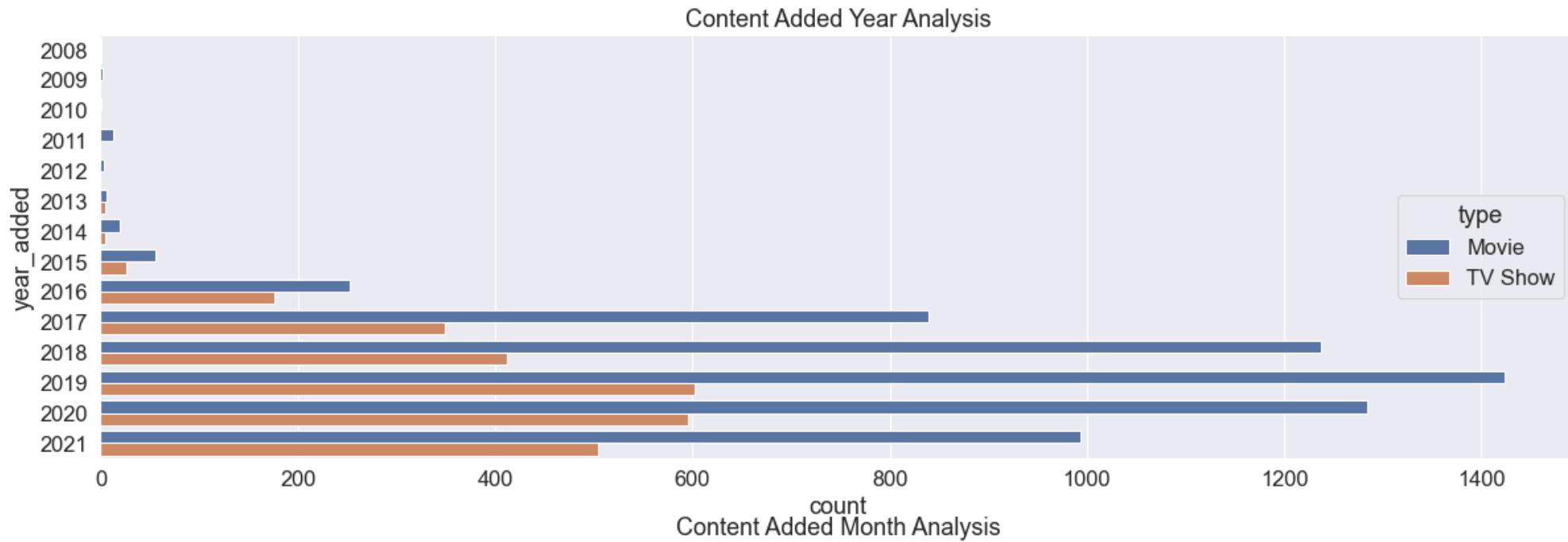
Netflix Monthly Contents Update

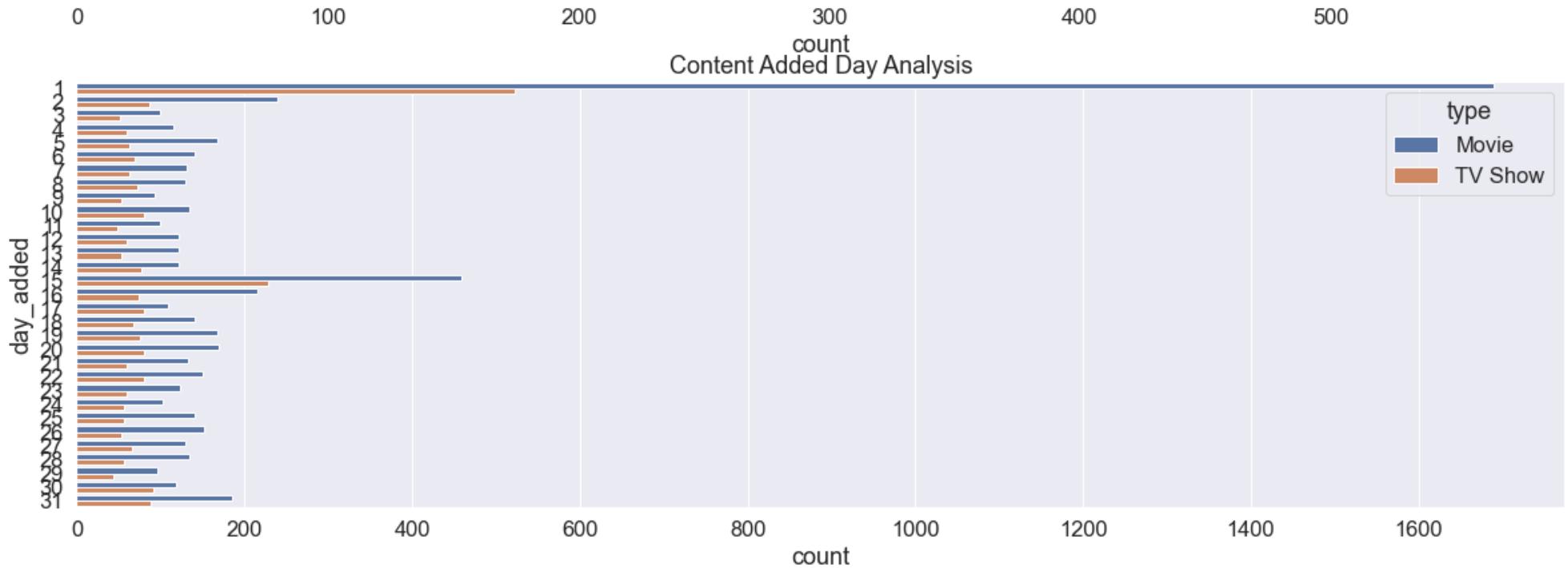


From the above heatmap we can say that in the year 2021 in July and year 2019 December month Netflix added highest number of content

This heatmap shows that there are more contents being added during the last few months compare to other months.

```
In [1807]: fig, axs = plt.subplots(nrows=3, figsize=(20,20))
# yearAddedContent = pd.DataFrame(dataframe.groupby(['year_added', 'type']).size().reset_index())
first = sns.countplot(y = 'year_added', hue = 'type', data = data, ax = axs[0])
first.set_title("Content Added Year Analysis")
second = sns.countplot(y = 'month_added', hue = 'type', data = data, ax = axs[1])
second.set_title("Content Added Month Analysis")
third = sns.countplot(y = 'day_added', hue = 'type', data = data, ax = axs[2])
third.set_title("Content Added Day Analysis")
plt.show()
```





Here we can do analysis for when to add content on Netflix

Observations:

1. In comparison with other years in the year 2019 around 1400 + movies i.e., 23.60 % percent of movies added and 600 TV shows added i.e., 22.5 % of TV shows added. We can say that 21% of total content Added alone in the year 2019.
2. After 2019 in the year 2020 we can see rate to add content is decreased for both TV shows and Movies For movies it is decreased by 10% and for TV Shows it is decreased by 1 %
In 2020, there is a slight increase in the rate of addition of TV shows and a slight reduction in rate of addition of movies which may be a factor of the COVID-19 pandemic.
3. In comparison with 2018 both the contents increased by almost 50% in 2019
4. For Movies as well as TV shows Content added in January, April, October and December is slightly greater as compared to others
5. And for the day comparison we can see content gets added on Netflix either on 1st or 15th of the month

Suggestions:

Based on the observation for the maximum profit, Netflix can add content in January, April, October and December Months either on 1 st or 15 th of the month

In []:

```
In [1808]: dataframe = data[['year_added' , 'type']]  
yearAddedContent = pd.DataFrame(dataframe.groupby(['year_added', 'type']).size().reset_index())  
yearAddedContent.rename({0 : "Total Content"}, axis = 1, inplace= True)  
yearAddedContent
```

Out[1808]:

	year_added	type	Total Content
0	2008	Movie	1
1	2008	TV Show	1
2	2009	Movie	2
3	2009	TV Show	0
4	2010	Movie	1
5	2010	TV Show	0
6	2011	Movie	13
7	2011	TV Show	0
8	2012	Movie	3
9	2012	TV Show	0
10	2013	Movie	6
11	2013	TV Show	5
12	2014	Movie	19
13	2014	TV Show	5
14	2015	Movie	56
15	2015	TV Show	26
16	2016	Movie	253
17	2016	TV Show	176
18	2017	Movie	839
19	2017	TV Show	349
20	2018	Movie	1237
21	2018	TV Show	412
22	2019	Movie	1424
23	2019	TV Show	602
24	2020	Movie	1284

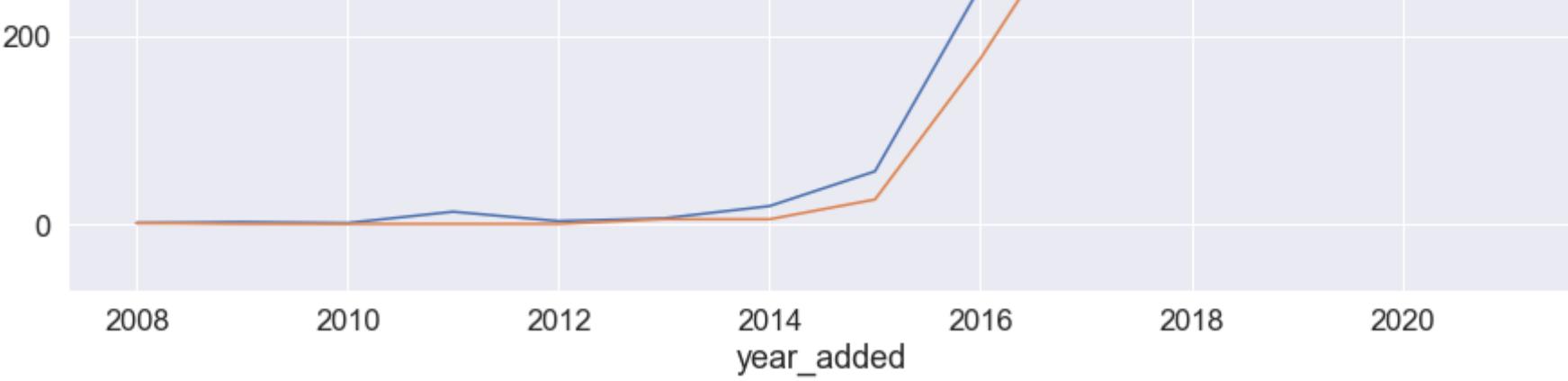
	year_added	type	Total Content
25	2020	TV Show	595
26	2021	Movie	993
27	2021	TV Show	505

```
In [1809]: totalCount = yearAddedContent['Total Content'].sum()  
totalCount
```

```
Out[1809]: 8807
```

```
In [1810]: plt.figure(figsize=(15,15))
sns.lineplot(x = 'year_added', y = 'Total Content', data = yearAddedContent, hue = 'type')
plt.legend(loc='upper left')
plt.title("TV Show and Movies wrt Year Added comparison", fontsize = 14)
plt.show()
```





Initially, the rate of addition of TV shows and movies seems to be at the same pace. From 2016 to 2018, a total of 2377 movies have been added while only 976 TV shows have been added. In 2020, there is a slight increase in the rate of addition of TV shows and a slight reduction in rate of addition of movies which may be a factor of the COVID-19 pandemic.

In []:

Release year wise content

```
In [1811]: df1 = data[['release_year' , 'type']]  
yearWiseContent = pd.DataFrame(df1.groupby(['release_year', 'type']).size().reset_index())  
yearWiseContent.rename({0 : "Total Content"}, axis = 1, inplace= True)  
yearWiseContent
```

Out[1811]:

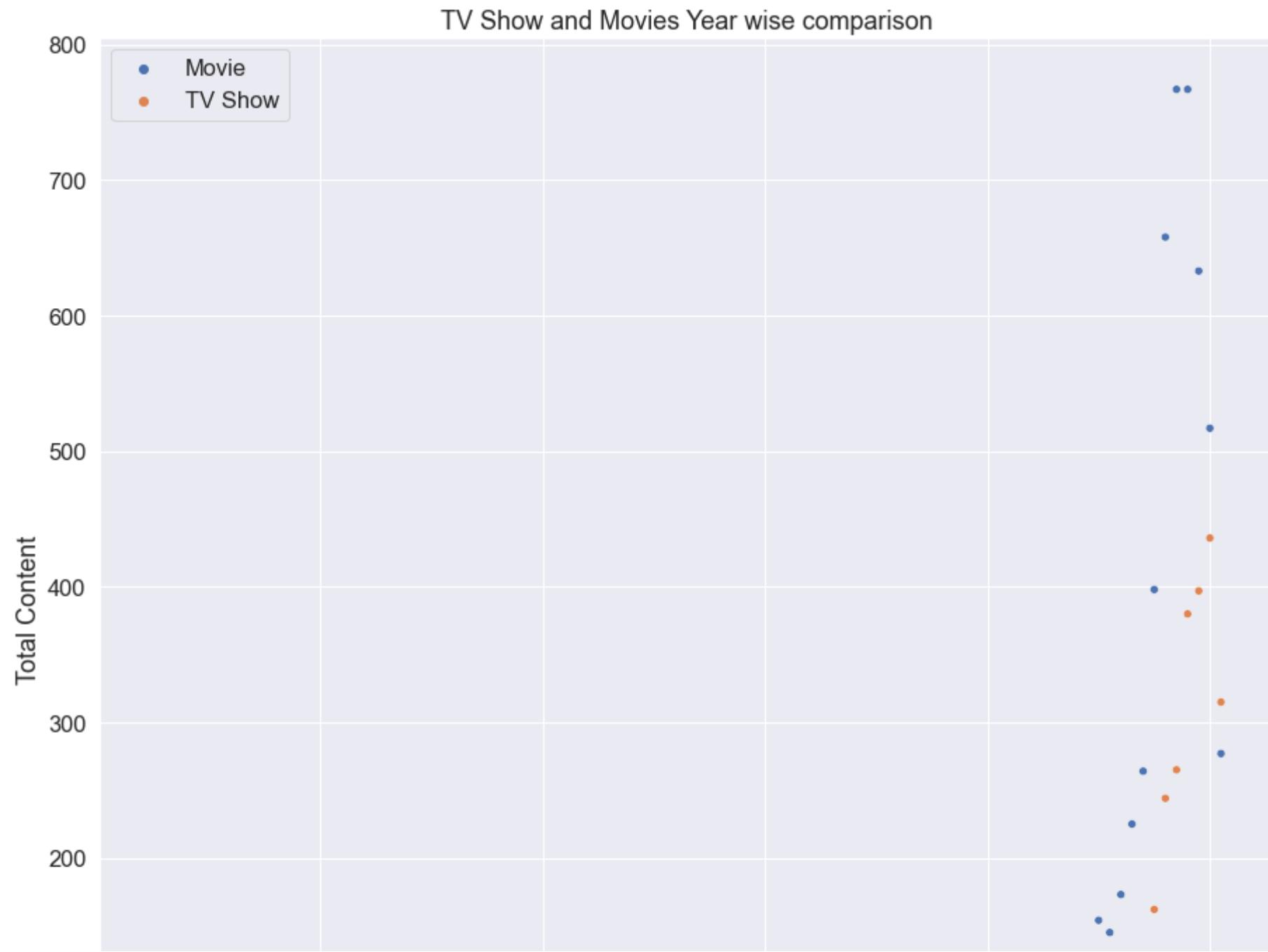
	release_year	type	Total Content
0	1925	Movie	0
1	1925	TV Show	1
2	1942	Movie	2
3	1942	TV Show	0
4	1943	Movie	3
...
143	2019	TV Show	397
144	2020	Movie	517
145	2020	TV Show	436
146	2021	Movie	277
147	2021	TV Show	315

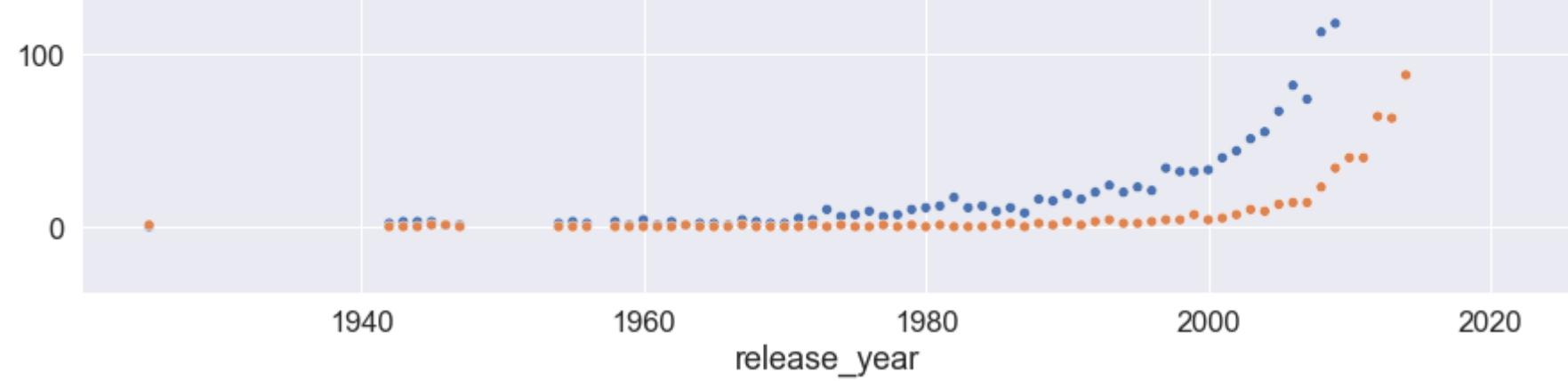
148 rows × 3 columns

```
In [1812]: totalCount = yearWiseContent['Total Content'].sum()  
totalCount
```

Out[1812]: 8807

```
In [1813]: plt.figure(figsize=(15,15))
sns.scatterplot(x = 'release_year', y = 'Total Content', data = yearWiseContent, hue = 'type')
plt.legend(loc='upper left')
plt.title("TV Show and Movies Year wise comparison")
plt.show()
```

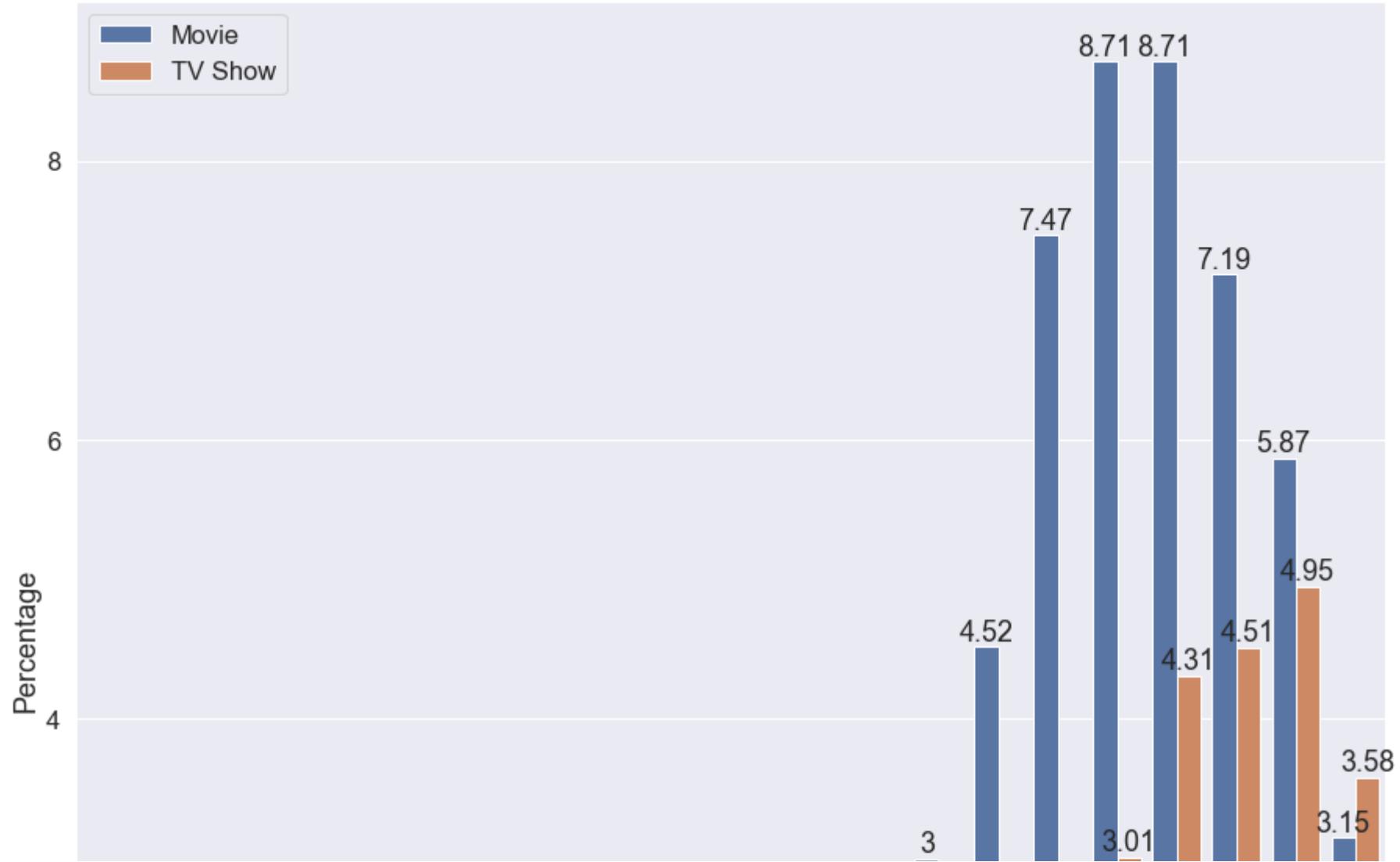


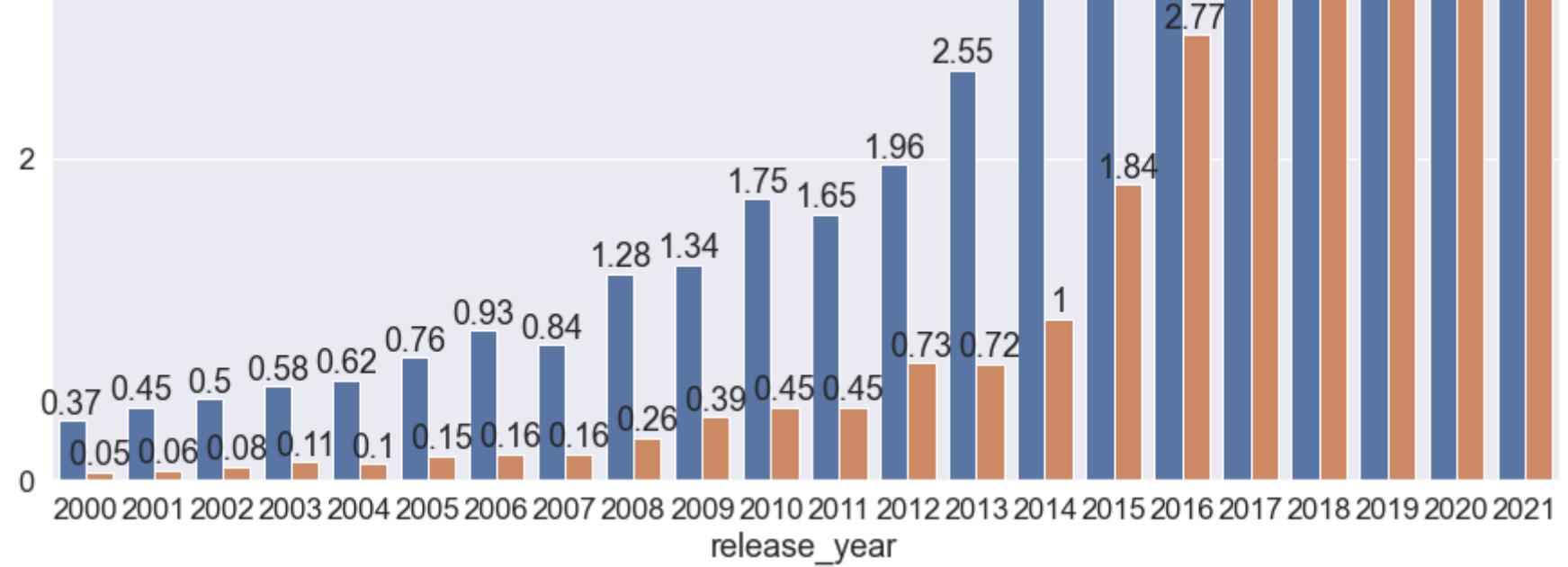


In [1814]: # Percentag plot for yearwise comparison

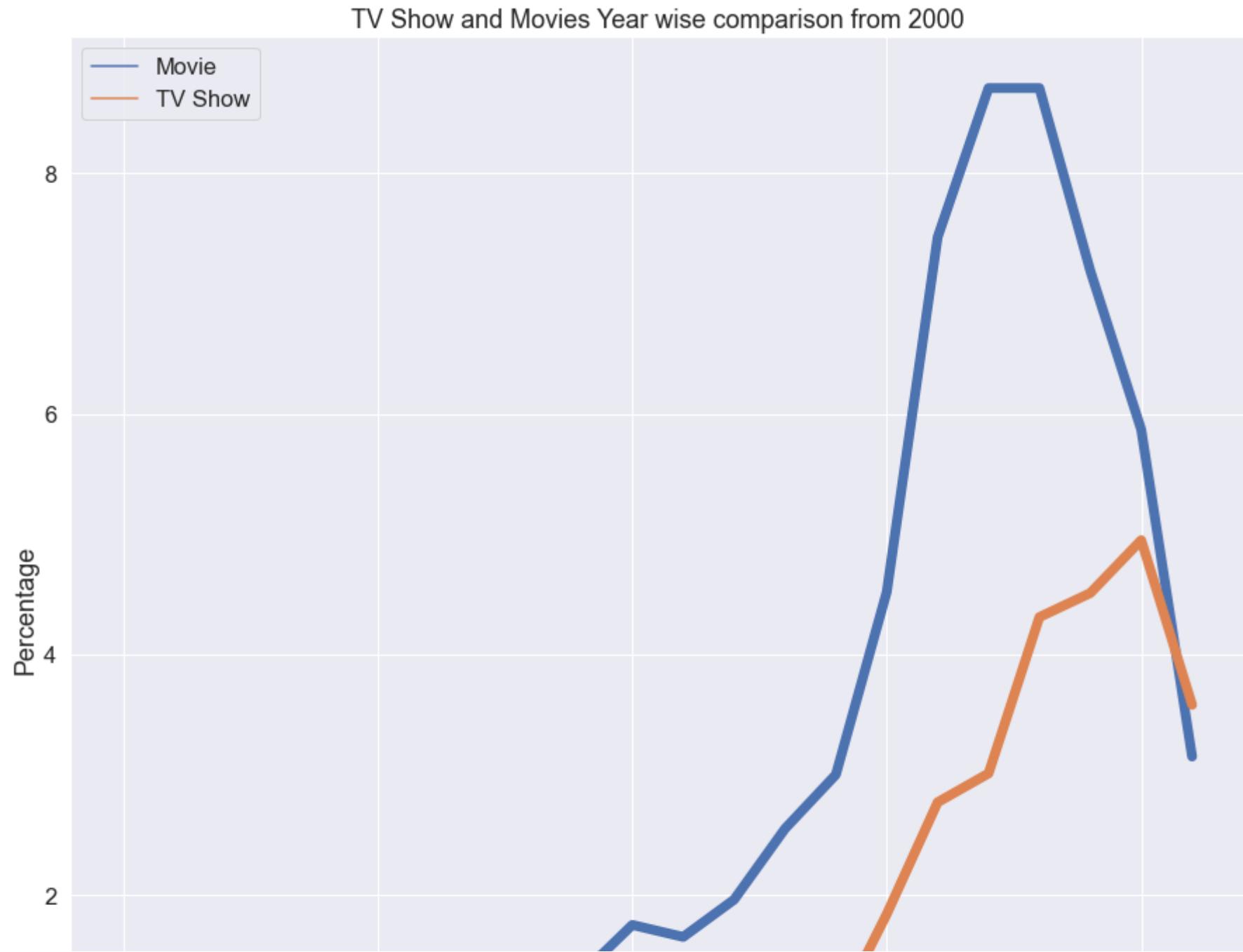
```
yearWiseContent['Percentage'] = np.round((yearWiseContent['Total Content'] * 100 / yearWiseContent['Total Content'].sum()) , 2)
yearWiseContent2 = yearWiseContent[yearWiseContent['release_year'] >= 2000]
plt.figure(figsize=(15,15))
graph = sns.barplot(x = 'release_year', y = 'Percentage', data = yearWiseContent2, hue = 'type')
for i in graph.containers:
    graph.bar_label(i)
plt.legend(loc='upper left')
plt.title("TV Show and Movies Year wise comparison from 2000")
# plt.text(x, y-150, '{:.0f}'.format(y), color='white').set_backgroundcolor('#965786')
plt.show()
```

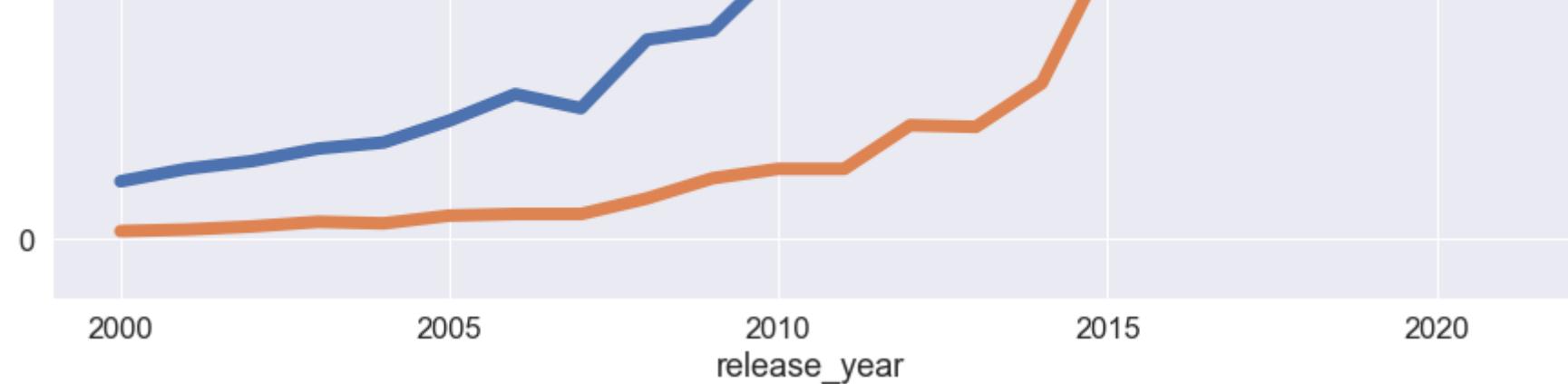
TV Show and Movies Year wise comparison from 2000





```
In [1815]: plt.figure(figsize=(15,15))
graph = sns.lineplot(x = 'release_year', y = 'Percentage', data = yearWiseContent2, hue = 'type', lw=7)
plt.legend(loc='upper left')
plt.title("TV Show and Movies Year wise comparison from 2000")
plt.show()
```





Most of the content present on Netflix is the content released after 2008

Rating Analysis

```
In [1816]: ratingDf = pd.DataFrame(data['rating'].value_counts(ascending= False))
ratingDf
```

Out[1816]:

rating	rating
TV-MA	3211
TV-14	2160
TV-PG	863
R	799
PG-13	490
TV-Y7	334
TV-Y	307
PG	287
TV-G	220
NR	80
G	41
TV-Y7-FV	6
NC-17	3
UR	3

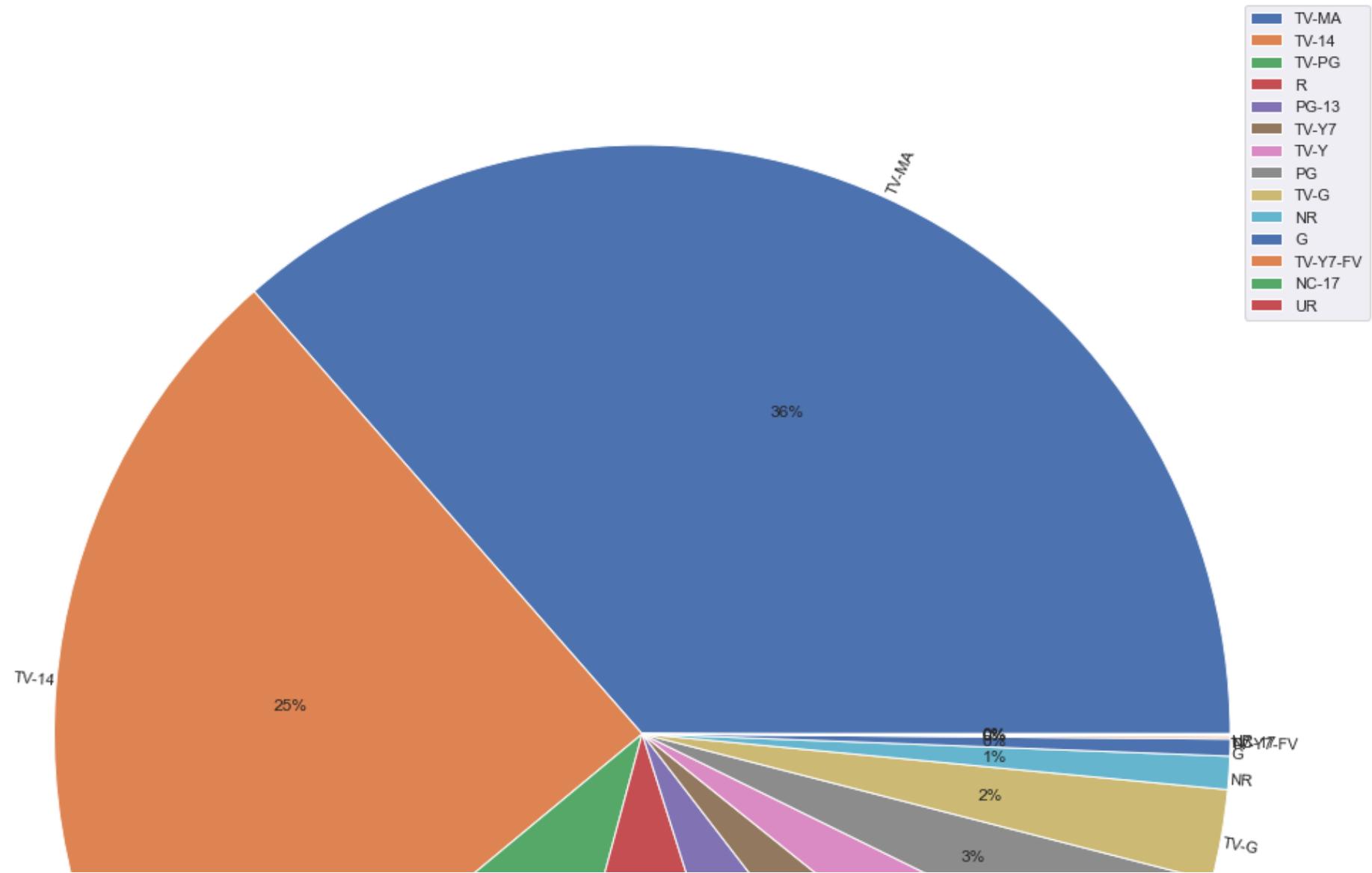
```
In [1817]: ratingDf.reset_index(inplace=True)  
ratingDf.columns = ['rating', 'rating_count']  
ratingDf
```

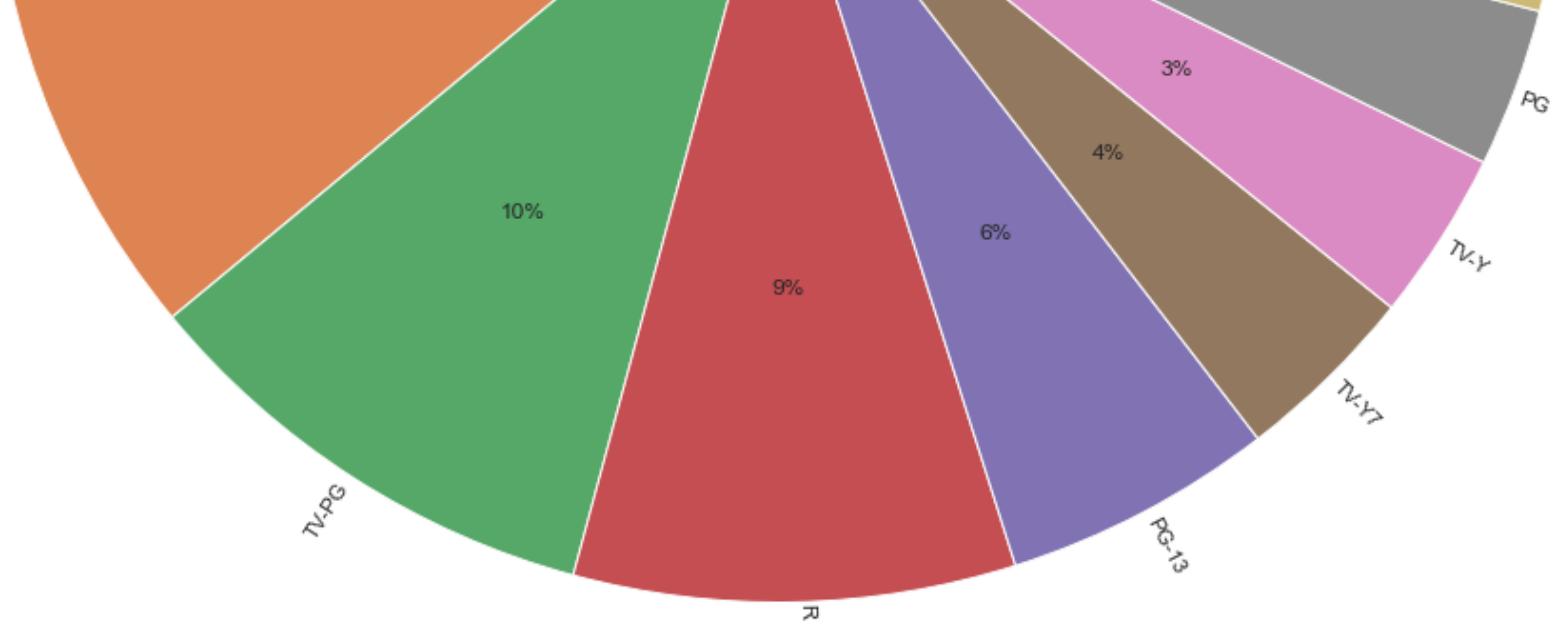
Out[1817]:

	rating	rating_count
0	TV-MA	3211
1	TV-14	2160
2	TV-PG	863
3	R	799
4	PG-13	490
5	TV-Y7	334
6	TV-Y	307
7	PG	287
8	TV-G	220
9	NR	80
10	G	41
11	TV-Y7-FV	6
12	NC-17	3
13	UR	3

```
In [1818]: values = list(ratingDf['rating_count'])
labels = list(ratingDf['rating'])
sns.set(font_scale = 1)
plt.figure(figsize=(15,15))
plt.pie(values ,labels = labels, autopct='%.0f%%', labeldistance=1 , rotate_labels = 270) # To show the portions in %ages
plt.title("Rating Percentage Pie Chart" , fontsize = 20)
plt.legend(labels, loc = 'best' )
plt.tight_layout()
plt.show()
```

Rating Percentage Pie Chart





From the above univariate analysis for Ratings we can say that about 36% of content is "TV-MA" i.e., Adults content. Around 25 % content is 'TV-14' i.e., Teens content, Around 10 % is 'TV-PG' i.e., Parental Guidance Suggested content, and 9% content is "R" type is Adult content as well So these are top 4 ratings on Netflix

In []:

```
In [1819]: df = data[['rating', 'type']]  
ratingVsType = pd.DataFrame(df.groupby(['rating', 'type']).size().reset_index())  
ratingVsType.rename({0: 'Content Count'}, axis = 1, inplace= True)  
ratingVsType
```

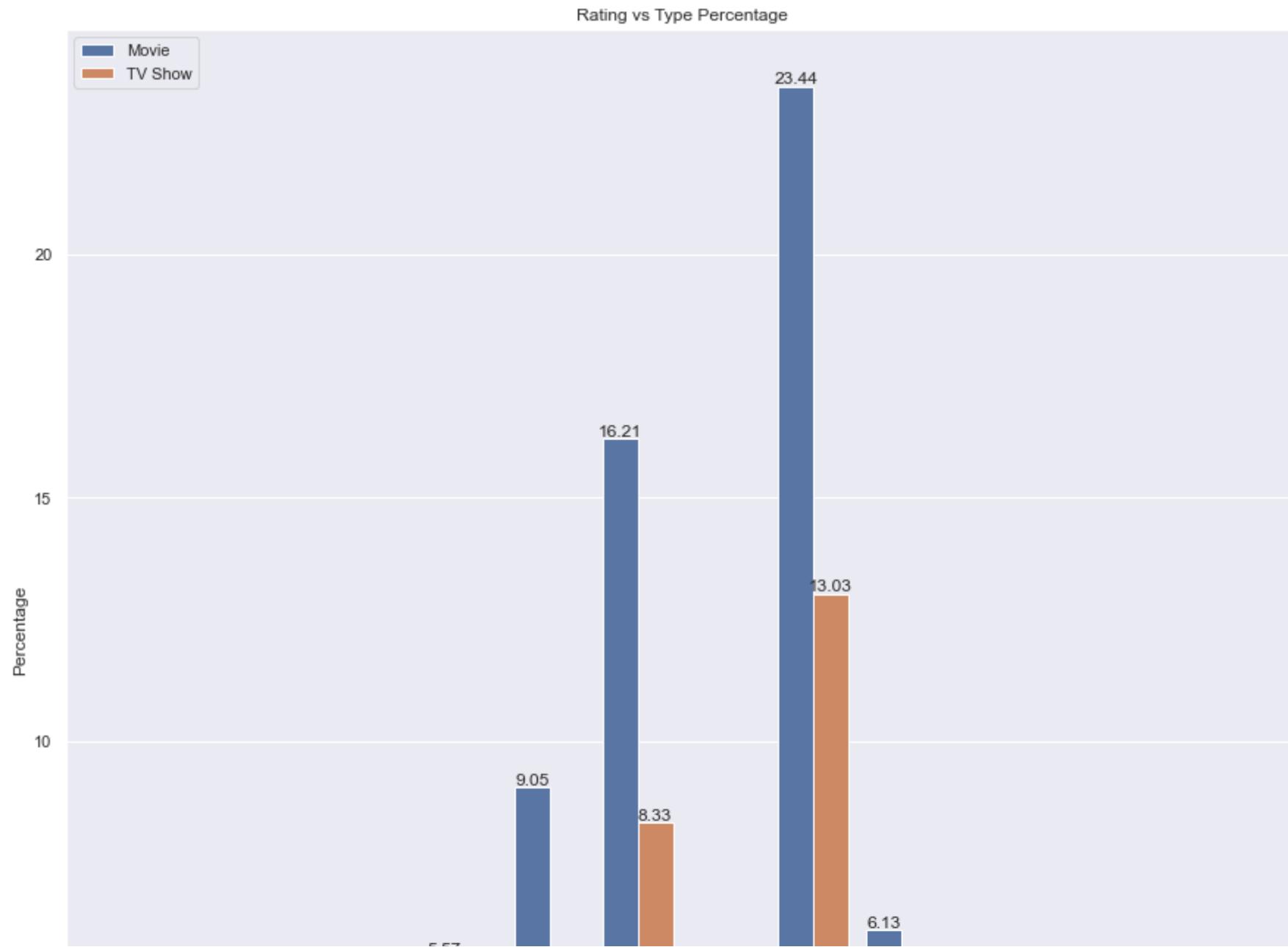
Out[1819]:

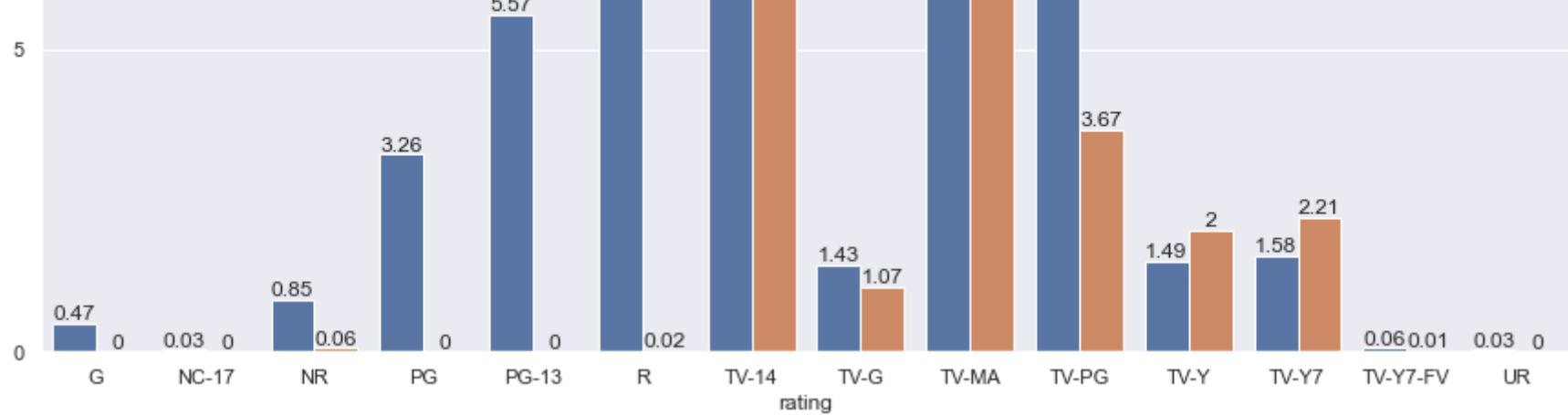
	rating	type	Content Count
0	G	Movie	41
1	G	TV Show	0
2	NC-17	Movie	3
3	NC-17	TV Show	0
4	NR	Movie	75
5	NR	TV Show	5
6	PG	Movie	287
7	PG	TV Show	0
8	PG-13	Movie	490
9	PG-13	TV Show	0
10	R	Movie	797
11	R	TV Show	2
12	TV-14	Movie	1427
13	TV-14	TV Show	733
14	TV-G	Movie	126
15	TV-G	TV Show	94
16	TV-MA	Movie	2064
17	TV-MA	TV Show	1147
18	TV-PG	Movie	540
19	TV-PG	TV Show	323
20	TV-Y	Movie	131
21	TV-Y	TV Show	176
22	TV-Y7	Movie	139
23	TV-Y7	TV Show	195
24	TV-Y7-FV	Movie	5

	rating	type	Content Count
25	TV-Y7-FV	TV Show	1
26	UR	Movie	3
27	UR	TV Show	0

```
In [1820]: ratingVsType['Percentage'] = np.round(ratingVsType['Content Count'] * 100 / ratingVsType['Content Count'].sum(),2)
```

```
In [1821]: plt.figure(figsize=(15,15))
graph = sns.barplot(x = 'rating', y = 'Percentage', data = ratingVsType, hue = 'type')
for i in graph.containers:
    graph.bar_label(i)
plt.legend(loc='upper left')
plt.title("Rating vs Type Percentage")
plt.show()
```





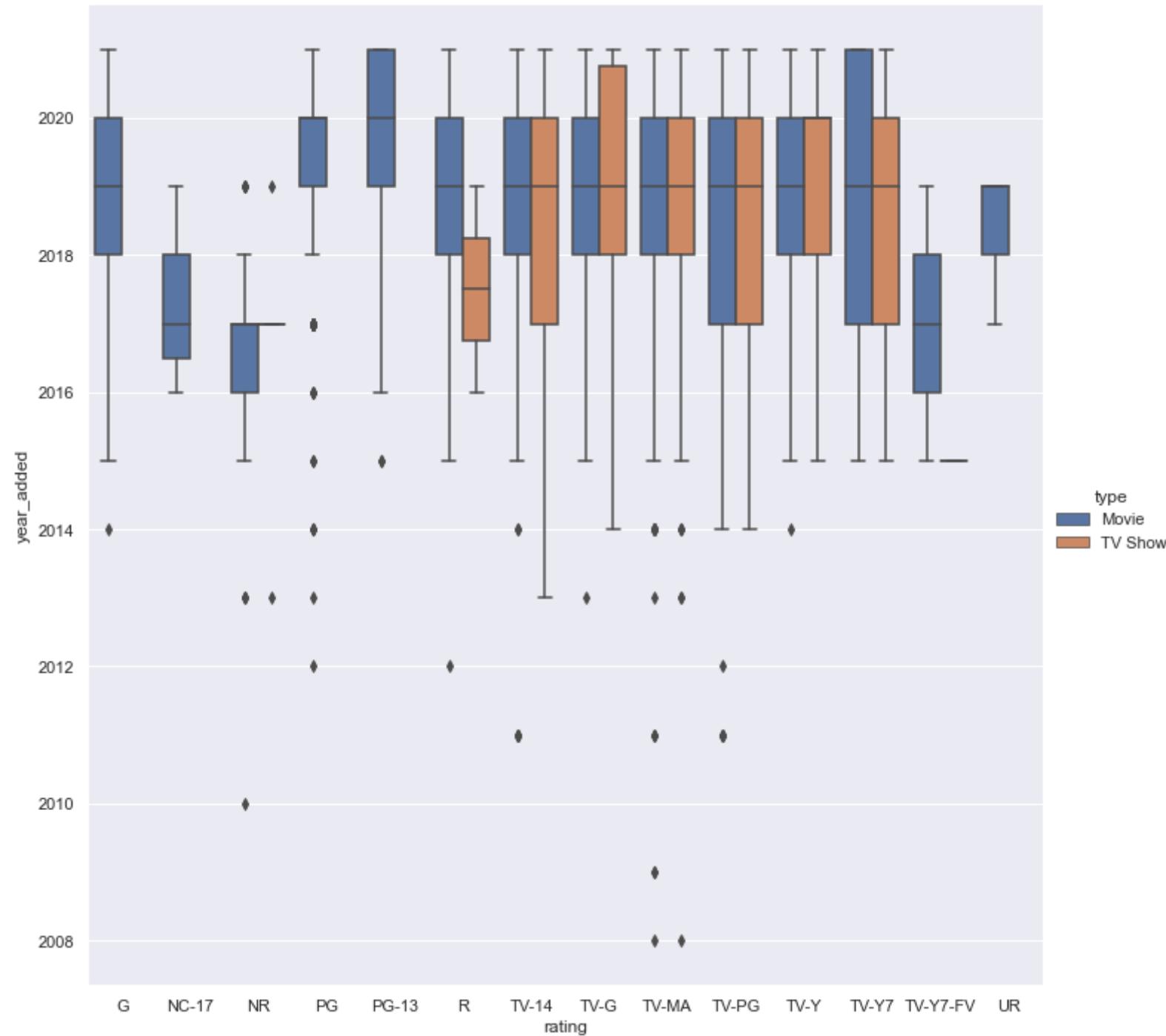
From the above analysis for rating of the content present on netflix TV-MA is the content which is present in largest amount for Movies as well as TV shows.

Which is about 36% of the total content. Which indicates that people are more attracted towards adult content as compared to other.

For Kids Content i.e., G NR PG PG13, R we can see movies and No Tv shows. And In comparison with Mature or Adult content This content in less amount which indicates that Netflix is mostly watched by adults

In []:

```
In [1822]: sns.catplot(x="rating", y="year_added", hue="type", kind="box", data= data, height = 10, aspect = 1)  
plt.show()
```



From the above box plot we can see that kids content is started to show as outlier for years 2018, 2020.

In []:

Movie Data Analysis

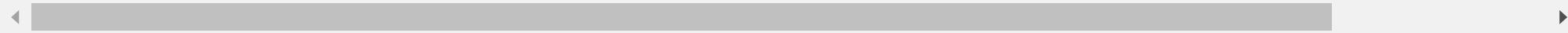
```
In [1823]: moviesData = pd.DataFrame(data[data['type'] == 'Movie'])
moviesData.reset_index()
```

Out[1823]:

	index	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added
0	0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmmaker... filmm...	2021
1	6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	United States	2021-09-24	2021	PG	91 min	Children & Family Movies	Equestria's divided. But a bright-eyed hero be... be...	2021
2	7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	2021-09-24	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s... s...	2021
3	9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	2021-09-24	2021	PG-13	104 min	Comedies, Dramas	A woman adjusting to life after a loss contend... contend...	2021
4	12	s13	Movie	Je Suis Karl	Christian Schwuchow	Luna Wedler, Jannis Niewöhner, Milan Peschel, ...	Germany, Czech Republic	2021-09-23	2021	TV-MA	127 min	Dramas, International Movies	After most of her family is murdered in a terr... terr...	2021
...
6126	8801	s8802	Movie	Zinzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ...	United Arab Emirates, Jordan	2016-03-09	2015	TV-MA	96 min	Dramas, International Movies, Thrillers	Recovering alcoholic Talal wakes up inside a s... s...	2016
6127	8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey Jr...	United States	2019-11-20	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a... a...	2019

index	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	
6128	8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...	2019
6129	8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020
6130	8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019

6131 rows × 16 columns



```
In [1824]: moviesData['duration'] = moviesData['duration'].apply(lambda x: str(x).split(' ')[0])
moviesData
```

Out[1824]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, filmm...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	2021-09-24	1993	TV-MA	125	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...	2021	
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	2021-09-24	2021	PG-13	104	Comedies, Dramas	A woman adjusting to life after a loss contend...	2021	
12	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Wedler, Jannis Niewöhner, Milan Peschel, ...	Germany, Czech Republic	2021-09-23	2021	TV-MA	127	Dramas, International Movies	After most of her family is murdered in a terr...	2021	
...
8801	s8802	Movie	Zinzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ...	United Arab Emirates, Jordan	2016-03-09	2015	TV-MA	96	Dramas, International Movies, Thrillers	Recovering alcoholic Talal wakes up inside a s...	2016	
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	2019-11-20	2007	R	158	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...	2019	

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01	2009	R	88	Comedies, Horror Movies	Looking to survive in a world taken over by zo...	2019
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019

6131 rows × 15 columns

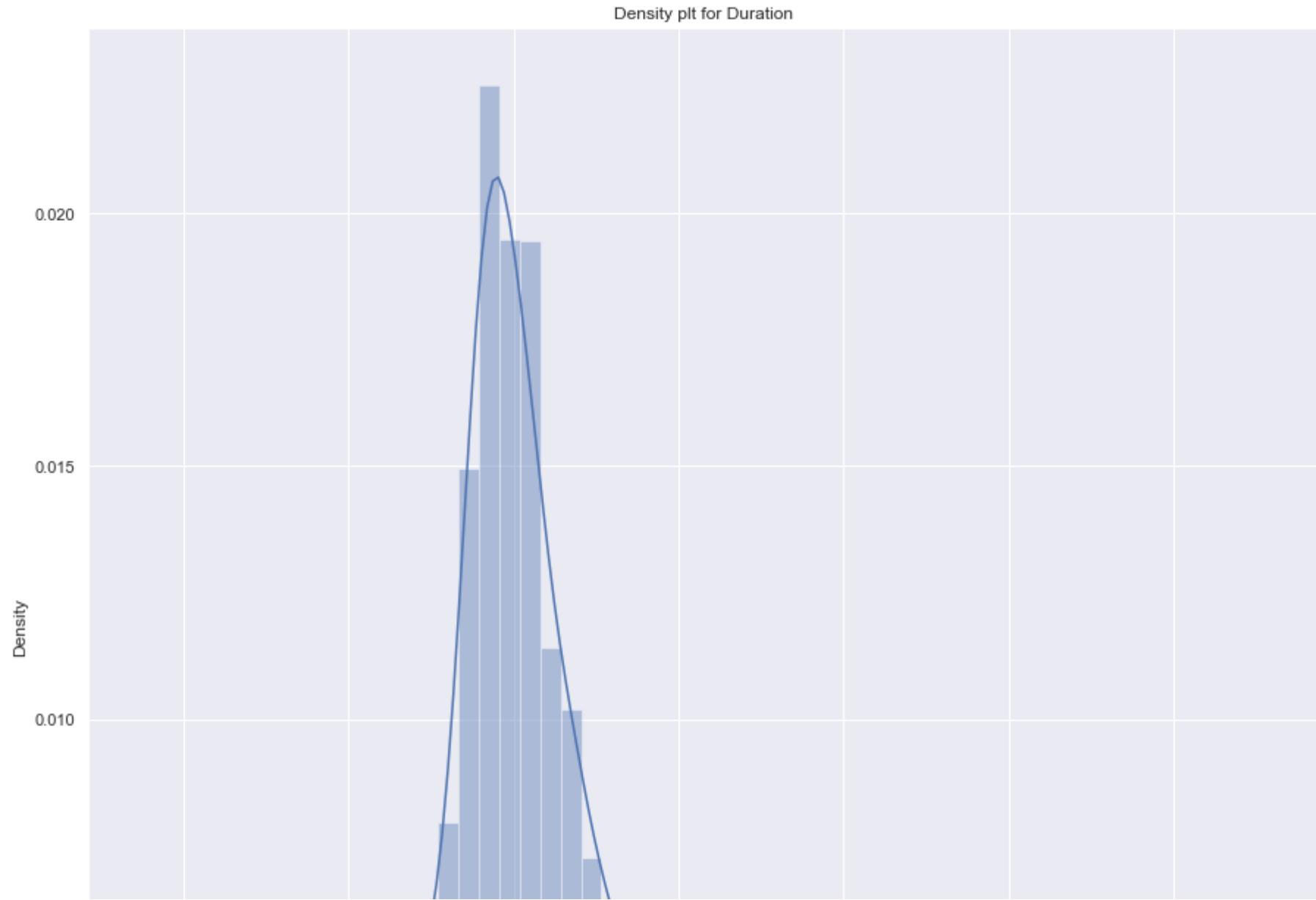


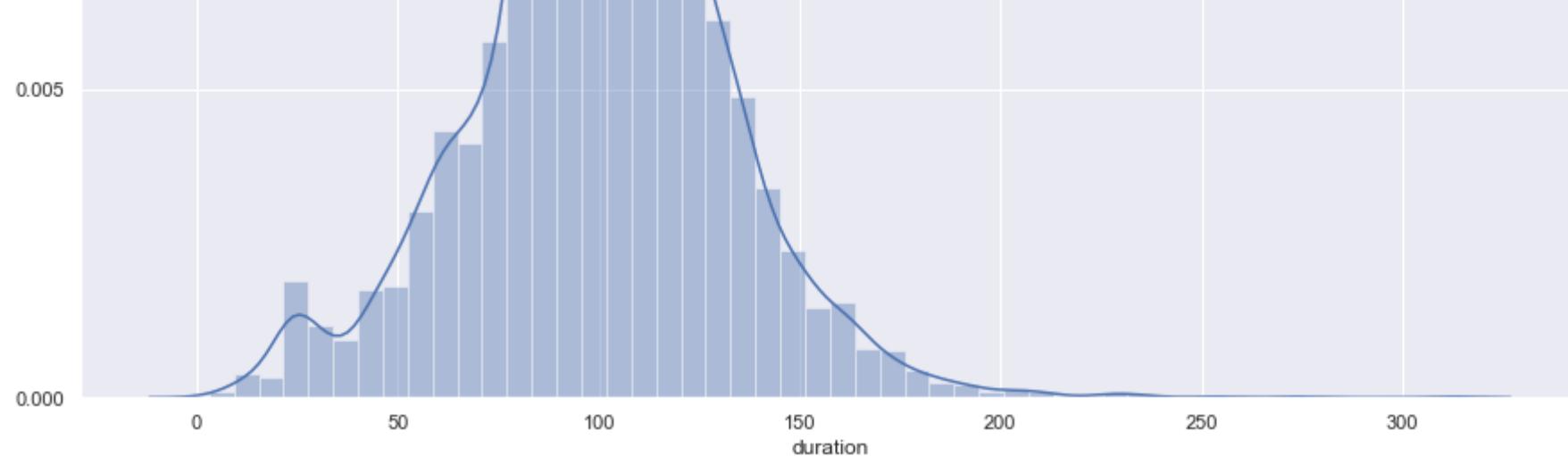
In [1825]:

```
plt.figure(figsize=(15,15))
sns.distplot(moviesData['duration'])
plt.title("Density plt for Duration")
plt.show()
```

C:\Program Files\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 `histplot` (an axes-level function for histograms).

```
warnings.warn(msg, FutureWarning)
```





```
In [1826]: moviesDataCountry = moviesData.assign(country= moviesData['country'].str.split(', ')).explode('country')
moviesDataCountry
```

Out[1826]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, film...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States	2021-09-24	1993	TV-MA	125	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	Ghana	2021-09-24	1993	TV-MA	125	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	Burkina Faso	2021-09-24	1993	TV-MA	125	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...	2021	
...
8801	s8802	Movie	Zinzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ...	Jordan	2016-03-09	2015	TV-MA	96	Dramas, International Movies, Thrillers	Recovering alcoholic Talal wakes up inside a s...	2016	
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	2019-11-20	2007	R	158	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...	2019	

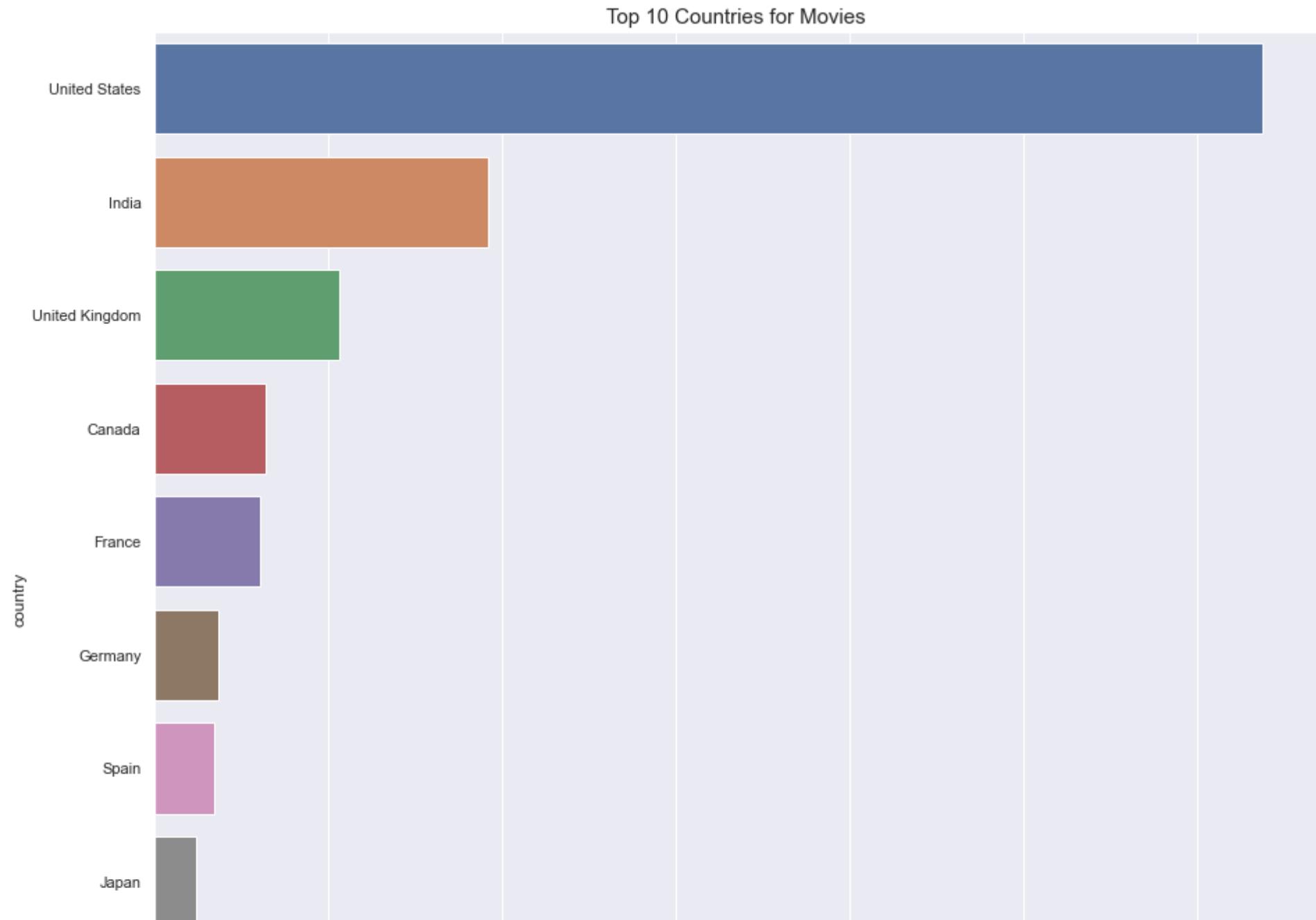
show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_adc
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01	2009	R	88	Comedies, Horror Movies	Looking to survive in a world taken over by zo...	2019
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019

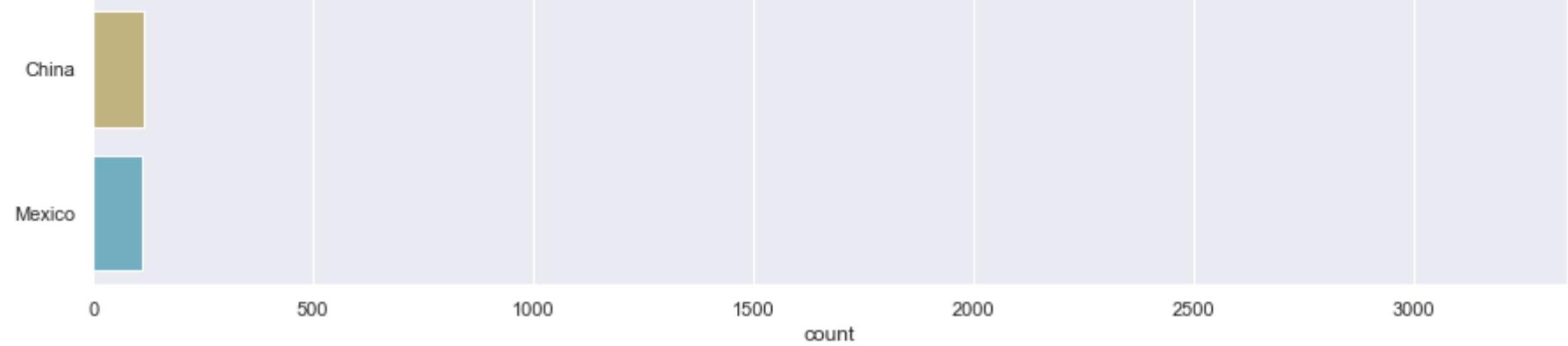
7814 rows × 15 columns



In [1827]: # Country wise analysis for movies

```
plt.figure(figsize=(15, 15))
sns.countplot(y = moviesDataCountry['country'],order = moviesDataCountry['country'].value_counts().index[0:10])
plt.title("Top 10 Countries for Movies" , fontsize=15)
plt.show()
```





For movies we can see here most of the movies are originated in United States, India, United Kingdom, Canada and France

```
In [1828]: moviesDataCast = moviesData.assign(cast= moviesData['cast'].str.split(', ')).explode('cast')
moviesDataCast
```

Out[1828]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, filmm...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Kimiko Glenn	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	James Marsden	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Sofia Carson	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
...
8806	s8807	Movie	Zubaan	Mozez Singh	Manish Chaudhary	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	
8806	s8807	Movie	Zubaan	Mozez Singh	Meghna Malik	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	
8806	s8807	Movie	Zubaan	Mozez Singh	Malkeet Rauni	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8806	s8807	Movie	Zubaan	Mozez Singh	Anita Shabdish	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019
8806	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019

44950 rows × 15 columns



In [1829]: moviesDataCast[moviesDataCast['cast'].isna()]

Out[1829]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	mont
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, filmmaker... ...	2021	
16	s17	Movie	Europe's Most Dangerous Man: Otto Skorzeny in ...	Pedro de Echave García, Pablo Azorín Williams	NaN	United States	2021-09-22	2020	TV-MA	67	Documentaries, International Movies	Declassified documents reveal the post-WWII li... ...	2021	
45	s46	Movie	My Heroes Were Cowboys	Tyler Greco	NaN	United States	2021-09-16	2021	PG	23	Documentaries	Robin Wiltshire's painful childhood was rescue... ...	2021	
91	s92	Movie	The Women and the Murderer	Mona Achache, Patricia Tourancheau	NaN	France	2021-09-09	2021	TV-14	92	Documentaries, International Movies	This documentary traces the capture of serial	2021	
101	s102	Movie	Untold: Breaking Point	Chapman Way, Maclain Way	NaN	United States	2021-09-07	2021	TV-MA	80	Documentaries, Sports Movies	Under pressure to continue a winning tradition... ...	2021	
...	
8739	s8740	Movie	Why We Fight: The Battle of Russia	Frank Capra, Anatole Litvak	NaN	United States	2017-03-31	1943	TV-PG	82	Documentaries	This installment of Frank Capra's acclaimed do... ...	2017	
8744	s8745	Movie	Williams	Morgan Matthews	NaN	United Kingdom	2017-11-04	2017	TV-MA	105	Documentaries, International Movies, Sports Mo... ...	This inspiring documentary charts the birth of... ...	2017	
8746	s8747	Movie	Winnie	Pascale Lamche	NaN	France, Netherlands, South Africa, Finland	2018-02-26	2017	TV-14	85	Documentaries, International Movies	Winnie Mandela speaks about her extraordinary	2018	

show_id	type	title		director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month
8756	s8757	Movie	Woodstock		Barak Goodman	Nan	United States	2019-08-13	2019	TV-MA	97	Documentaries, Music & Musicals	For the 50th anniversary of the legendary Wood...	2019
8763	s8764	Movie	WWII: Report from the Aleutians		John Huston	Nan	United States	2017-03-31	1943	TV-PG	45	Documentaries	Filmmaker John Huston narrates this Oscar-nomi...	2017

475 rows × 15 columns

◀	▶
---	---

In [1830]: moviesDataCast.describe(include=['object', 'int64'])

Out[1830]:

	show_id	title	director	cast	country	release_year	duration	listed_in	description	year_added	month_added	day_added
count	44950	44950	44211	44475	44950	44950.000000	44950	44950	44950	44950.000000	44950.000000	44950.000000
unique	6131	6131	4354	25951	651	Nan	205	278	6105	Nan	Nan	Nan
top	s6187	Arthur Christmas	Martin Scorsese	Anupam Kher	United States	Nan	94	Dramas, International Movies	At Santa Claus's super-secret toy-making facil...	Nan	Nan	Nan
freq	44	44	139	42	17557	Nan	1255	3022	44	Nan	Nan	Nan
mean	Nan	Nan	Nan	Nan	Nan	2012.300178	Nan	Nan	Nan	2019.013014	6.637597	11.661935
std	Nan	Nan	Nan	Nan	Nan	9.833517	Nan	Nan	Nan	1.526362	3.466023	9.909213
min	Nan	Nan	Nan	Nan	Nan	1942.000000	Nan	Nan	Nan	2008.000000	1.000000	1.000000
25%	Nan	Nan	Nan	Nan	Nan	2010.000000	Nan	Nan	Nan	2018.000000	4.000000	1.000000
50%	Nan	Nan	Nan	Nan	Nan	2016.000000	Nan	Nan	Nan	2019.000000	7.000000	11.000000
75%	Nan	Nan	Nan	Nan	Nan	2018.000000	Nan	Nan	Nan	2020.000000	10.000000	20.000000
max	Nan	Nan	Nan	Nan	Nan	2021.000000	Nan	Nan	Nan	2021.000000	12.000000	31.000000

```
In [1831]: moviesDataCastUS = moviesDataCast[moviesDataCast['country'] == 'United States']
moviesDataCastUS
```

Out[1831]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, film...	2021	9
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	9
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Kimiko Glenn	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	9
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	James Marsden	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	9
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Sofia Carson	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	9
...
8805	s8806	Movie	Zoom	Peter Hewitt	Ryan Newman	United States	2020-01-11	2006	PG	88	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020	1
8805	s8806	Movie	Zoom	Peter Hewitt	Michael Cassidy	United States	2020-01-11	2006	PG	88	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020	1
8805	s8806	Movie	Zoom	Peter Hewitt	Spencer Breslin	United States	2020-01-11	2006	PG	88	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020	1

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	
8805	s8806	Movie	Zoom	Peter Hewitt	Rip Torn	United States	2020-01-11	2006	PG	88	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020	1
8805	s8806	Movie	Zoom	Peter Hewitt	Kevin Zegers	United States	2020-01-11	2006	PG	88	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020	1

17557 rows × 15 columns



In [1832]: `moviesDataCastUS.describe(include=['object', 'int64'])`

Out[1832]:

	show_id	title	director	cast	country	release_year	duration	listed_in	description	year_added	month_added	day_added
count	17557	17557	17043	17298	17557	17557.000000	17557	17557	17557	17557.000000	17557.000000	17557.000000
unique	2498	2498	1801	10803	1	NaN	172	205	2490	NaN	NaN	NaN
top	s5614	Michael Bolton's Big, Sexy Valentine's Day Spe...	Rajiv Chilaka	Rupa Bhimani	United States	NaN	97	Children & Family Movies	After Santa tells him he needs 75,000 new babi...	NaN	NaN	NaN
freq	41	41	98	23	17557	NaN	562	1202	41	NaN	NaN	NaN
mean	NaN	NaN	NaN	NaN	NaN	2011.691462	NaN	NaN	NaN	2019.160563	6.563707	10.938942
std	NaN	NaN	NaN	NaN	NaN	10.742395	NaN	NaN	NaN	1.660499	3.461914	9.727181
min	NaN	NaN	NaN	NaN	NaN	1942.000000	NaN	NaN	NaN	2008.000000	1.000000	1.000000
25%	NaN	NaN	NaN	NaN	NaN	2009.000000	NaN	NaN	NaN	2018.000000	4.000000	1.000000
50%	NaN	NaN	NaN	NaN	NaN	2016.000000	NaN	NaN	NaN	2019.000000	7.000000	9.000000
75%	NaN	NaN	NaN	NaN	NaN	2019.000000	NaN	NaN	NaN	2020.000000	9.000000	19.000000
max	NaN	NaN	NaN	NaN	NaN	2021.000000	NaN	NaN	NaN	2021.000000	12.000000	31.000000

```
In [1833]: moviesDataCastUS['cast'].mode()
```

```
Out[1833]: 0    Julie Tejwani  
1    Rupa Bhimani  
dtype: object
```

```
In [1834]: moviesDataCastUS['cast'].fillna("Julie Tejwani", inplace=True)
```

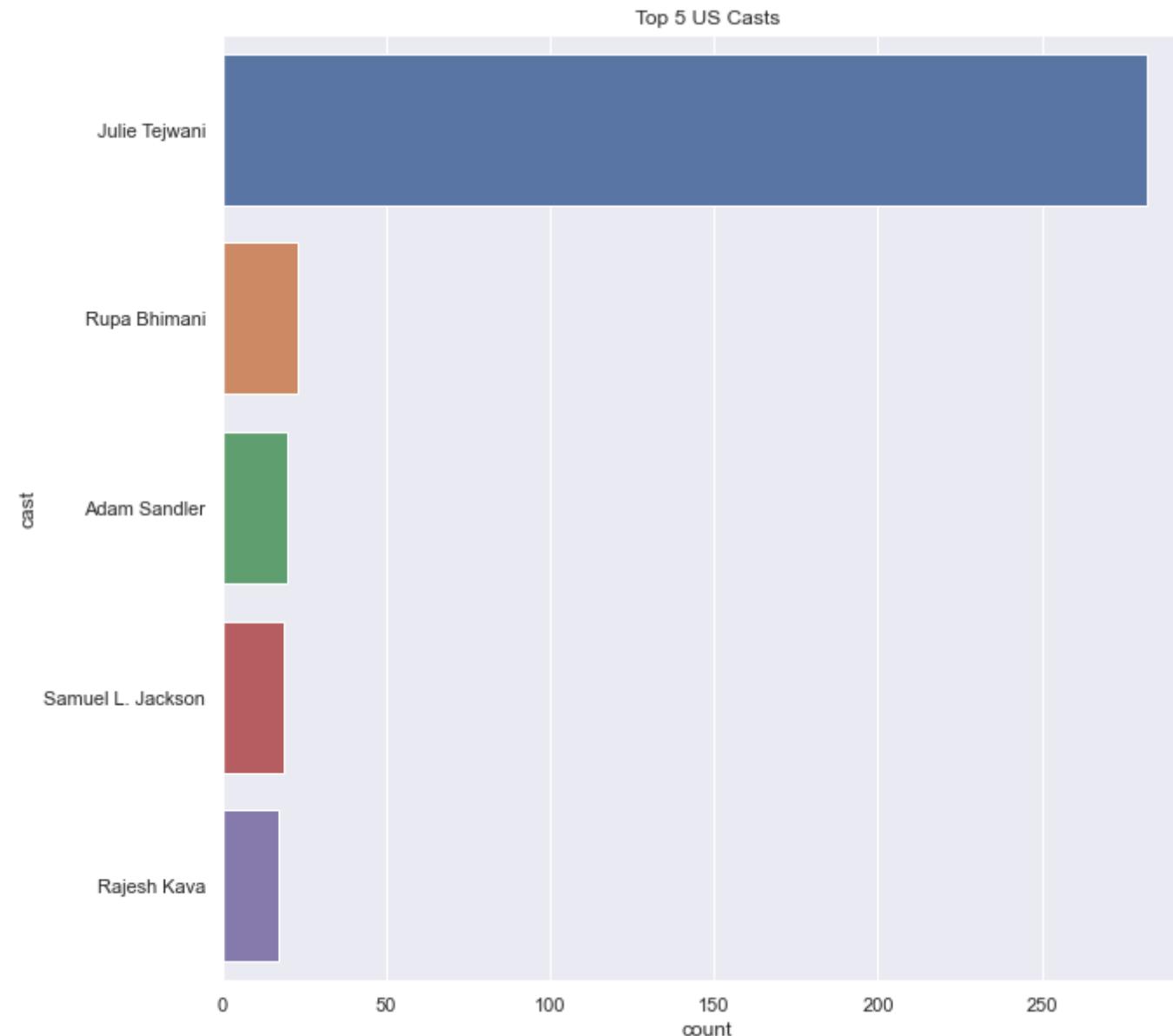
```
C:\Program Files\Anaconda3\lib\site-packages\pandas\core\generic.py:6392: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy  
return self._update_inplace(result)
```

```
In [1835]: moviesDataCastUS['cast'].value_counts()
```

```
Out[1835]: Julie Tejwani      282  
Rupa Bhimani       23  
Adam Sandler        20  
Samuel L. Jackson     19  
Rajesh Kava         17  
...  
Hong-Chi Lee          1  
Christine Ko          1  
Kofi Kingston         1  
Babatunde Aiyegbusi    1  
Ryan Newman           1  
Name: cast, Length: 10803, dtype: int64
```

```
In [1836]: plt.figure(figsize=(10,10))
sns.countplot(y = moviesDataCastUS['cast'] , order = moviesDataCastUS['cast'].value_counts().index[0:5])
plt.title("Top 5 US Casts", fontsize = 12)
plt.show()
```



Top 5 Casts in United States for movies are :

1. Julie Tejwani
2. Rupa Bhimani
3. Adam Sandler
4. Samuel L. Jackson
5. Rajesh Kava

```
In [1837]: moviesDataCastIndia = moviesDataCast[moviesDataCast['country'] == 'India']
moviesDataCastIndia
```

Out[1837]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
24	s25	Movie	Jeans	S. Shankar	Prashanth	India	2021-09-21	1998	TV-14	166	Comedies, International Movies, Romantic Movies	When the father of the man she loves insists t...	2021	9
24	s25	Movie	Jeans	S. Shankar	Aishwarya Rai Bachchan	India	2021-09-21	1998	TV-14	166	Comedies, International Movies, Romantic Movies	When the father of the man she loves insists t...	2021	9
24	s25	Movie	Jeans	S. Shankar	Sri Lakshmi	India	2021-09-21	1998	TV-14	166	Comedies, International Movies, Romantic Movies	When the father of the man she loves insists t...	2021	9
24	s25	Movie	Jeans	S. Shankar	Nassar	India	2021-09-21	1998	TV-14	166	Comedies, International Movies, Romantic Movies	When the father of the man she loves insists t...	2021	9
105	s106	Movie	Angamaly Diaries	Lijo Jose Pellissery	Antony Varghese	India	2021-09-05	2017	TV-14	128	Action & Adventure, Comedies, Dramas	After growing up amidst the gang wars of his h...	2021	9
...
8806	s8807	Movie	Zubaan	Mozez Singh	Manish Chaudhary	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3
8806	s8807	Movie	Zubaan	Mozez Singh	Meghna Malik	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3
8806	s8807	Movie	Zubaan	Mozez Singh	Malkeet Rauni	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	
8806	s8807	Movie	Zubaan	Mozez Singh	Anita Shabdish	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3
8806	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3

6835 rows × 15 columns



In [1838]: moviesDataCastIndia[moviesDataCastIndia['cast'].isna()]

Out[1838]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_
1853	s1854	Movie	Rooting for Roona	Pavitra Chalam, Akshay Shankar	NaN	India	2020-10-15	2020	TV-PG	42	Documentaries, International Movies	In rural India, a child with hydrocephalus get...	2020	
2334	s2335	Movie	Seven (Tamil)		NaN NaN	India	2020-06-26	2019	TV-MA	111	Dramas, International Movies, Romantic Movies	Multiple women report their husbands as missin...	2020	
2737	s2738	Movie	Mighty Raju Rio Calling	Rajiv Chilaka, Anirban Majumder, Alka Amarkant...	NaN	India	2020-04-01	2014	TV-Y7	97	Children & Family Movies, Sports Movies	Mighty Raju preps for a new adventure in Rio d...	2020	
3468	s3469	Movie	Limitless	Vrinda Samartha	NaN	India	2019-10-01	2017	TV-PG	59	Documentaries, International Movies, Sports Mo...	This documentary follows eight women in India ...	2019	
3714	s3715	Movie	Unbroken	Svati Chakravarty Bhatkal	NaN	India	2019-06-26	2019	TV-14	109	Documentaries, International Movies	After the brutal murders of their loved ones, ...	2019	
4187	s4188	Movie	Adhugo		NaN NaN	India	2019-01-15	2018	TV-14	105	Comedies, Independent Movies, International Mo...	A fast-footed piglet named Bunty becomes an ob...	2019	
4393	s4394	Movie	Muzaffarnagar Baqi Hai	Nakul Singh Sawhney	NaN	India	2018-11-15	2015	TV-MA	136	Documentaries, International Movies	Through interviews with locals, this documenta...	2018	
4995	s4996	Movie	Ladies First	Uraaz Bahl	NaN	India	2018-03-08	2018	TV-G	40	Documentaries, International Movies, Sports Mo...	Born amid poverty and limited women's rights i...	2018	
6150	s6151	Movie	An American in Madras	Karan Bali	NaN	India	2017-01-01	2013	TV-PG	80	Documentaries, International Movies	Extensive film clips and interviews tell the s...	2017	

show_id	type	title		director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_
6206	s6207	Movie	Bachelor Girls		Shikha Makan	NaN	India	2018-03-01	2016	TV-MA	61	Documentaries, International Movies	Through intimate interviews, this documentary ...	2018
6737	s6738	Movie	Famous in Ahmedabad		Hardik Mehta	NaN	India	2016-12-15	2015	TV-14	29	Documentaries, International Movies	Follow an 11-year-old boy as he goes from ordi...	2016
7135	s7136	Movie	John & Jane		Ashim Ahluwalia	NaN	India	2016-08-15	2005	TV-14	79	Documentaries, International Movies	Truth and fiction blend in this quasi-document...	2016
7413	s7414	Movie	Martyrs of Marriage		Deepika Narayan Bhardwaj	NaN	India	2018-03-15	2016	TV-14	83	Documentaries, International Movies	This documentary reveals the cracks in India's...	2018
8350	s8351	Movie	The Human Factor: The Untold Story of the Bomb...		Rudradeep Bhattacharjee	NaN	India	2017-04-01	2013	TV-PG	53	Documentaries, International Movies, Music & M...	A family of Parsi musicians collectively works...	2017
8475	s8476	Movie	The Rat Race		Miriam Chandy Menacherry	NaN	India	2017-01-01	2010	TV-PG	52	Documentaries, International Movies	Examine the troublesome coexistence of humans ...	2017



In [1839]: moviesDataCastIndia['cast'].mode()

Out[1839]: 0 Anupam Kher
dtype: object

In [1840]: moviesDataCastIndia['cast'].fillna("Anupam Kher", inplace= True)

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\generic.py:6392: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

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

```
return self._update_inplace(result)
```

In [1841]: moviesDataCastIndia

Out[1841]:

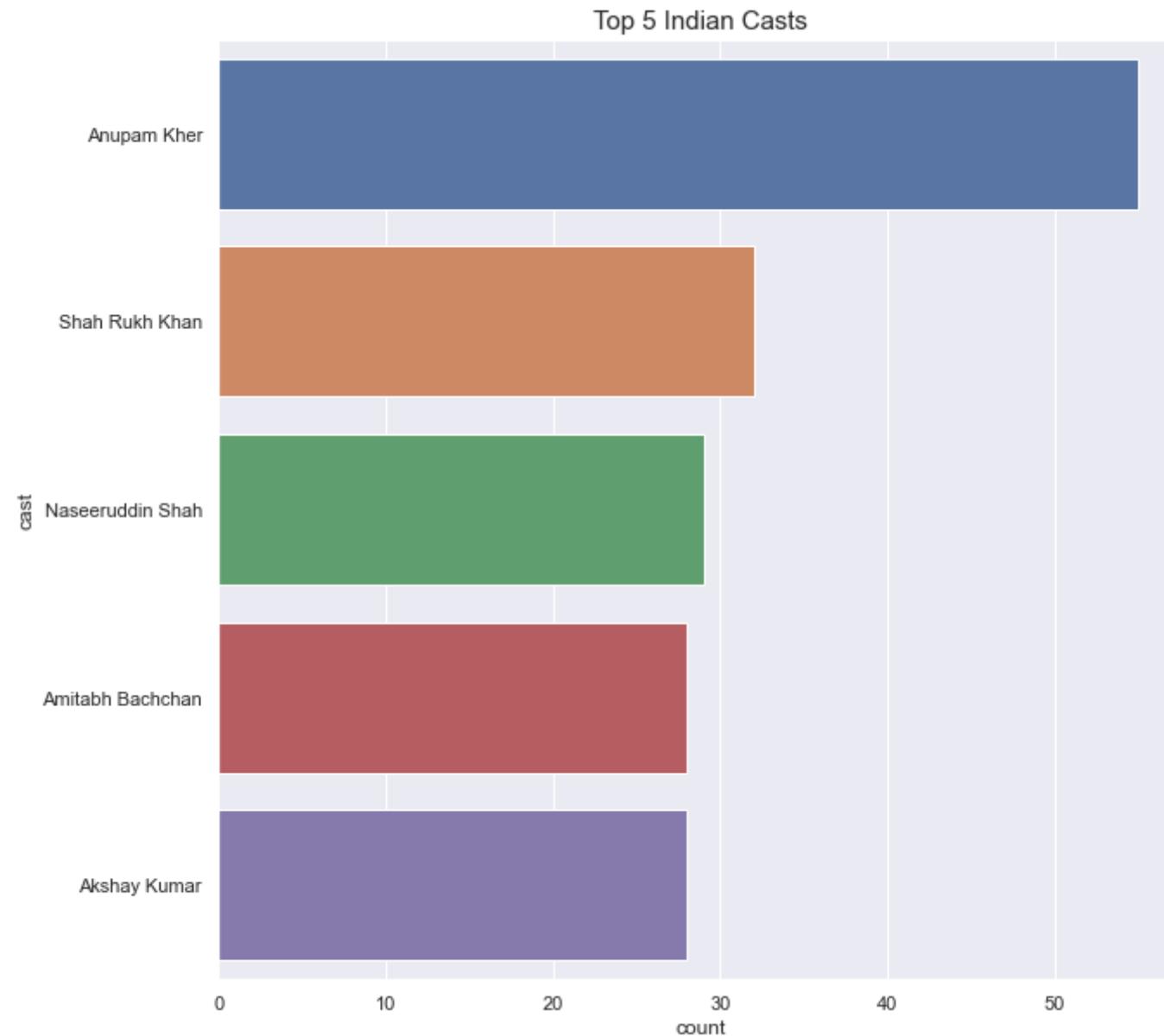
	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
24	s25	Movie	Jeans	S. Shankar	Prashanth	India	2021-09-21	1998	TV-14	166	Comedies, International Movies, Romantic Movies	When the father of the man she loves insists t...	2021	9
24	s25	Movie	Jeans	S. Shankar	Aishwarya Rai Bachchan	India	2021-09-21	1998	TV-14	166	Comedies, International Movies, Romantic Movies	When the father of the man she loves insists t...	2021	9
24	s25	Movie	Jeans	S. Shankar	Sri Lakshmi	India	2021-09-21	1998	TV-14	166	Comedies, International Movies, Romantic Movies	When the father of the man she loves insists t...	2021	9
24	s25	Movie	Jeans	S. Shankar	Nassar	India	2021-09-21	1998	TV-14	166	Comedies, International Movies, Romantic Movies	When the father of the man she loves insists t...	2021	9
105	s106	Movie	Angamaly Diaries	Lijo Jose Pellissery	Antony Varghese	India	2021-09-05	2017	TV-14	128	Action & Adventure, Comedies, Dramas	After growing up amidst the gang wars of his h...	2021	9
...
8806	s8807	Movie	Zubaan	Mozez Singh	Manish Chaudhary	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3
8806	s8807	Movie	Zubaan	Mozez Singh	Meghna Malik	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3
8806	s8807	Movie	Zubaan	Mozez Singh	Malkeet Rauni	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	
8806	s8807	Movie	Zubaan	Mozez Singh	Anita Shabdish	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3
8806	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	3

6835 rows × 15 columns



```
In [1842]: plt.figure(figsize=(10,10))
sns.countplot(y = moviesDataCastIndia['cast'] , order = moviesDataCastIndia['cast'].value_counts().index[0:5])
plt.title("Top 5 Indian Casts", fontsize = 15)
plt.show()
```



Top 5 casts from India on Netflix:

1. Anupam Kher
2. Shah Rukh Khan
3. Naseeruddin Shah
4. Amitabh bacchan
5. Akshay Kumar

In [1843]: moviesDataCast[moviesDataCast['cast'].isna()]

Out[1843]:

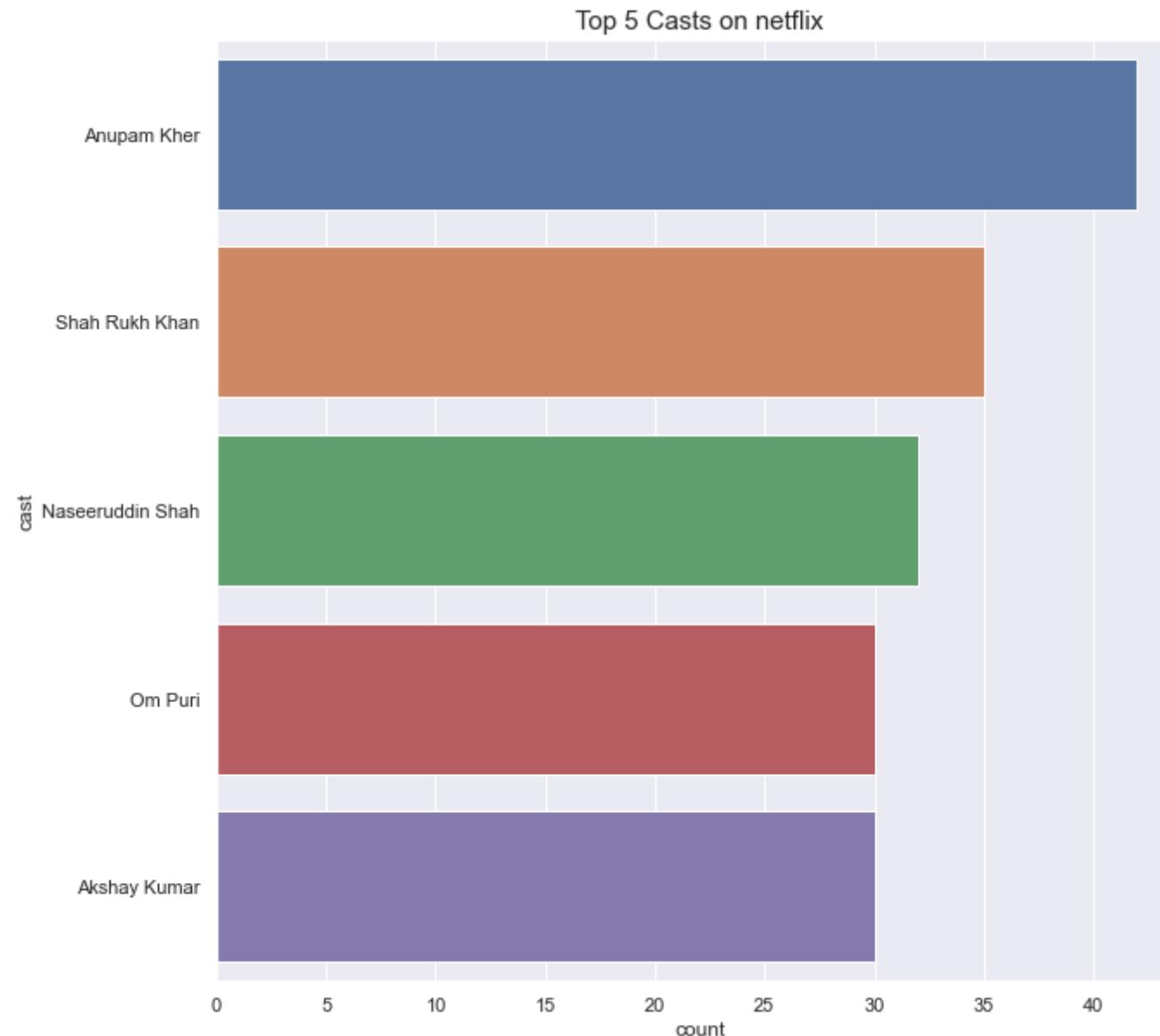
	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_a
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, film...	2021	
16	s17	Movie	Europe's Most Dangerous Man: Otto Skorzeny in ...	Pedro de Echave García, Pablo Azorín Williams	NaN	United States	2021-09-22	2020	TV-MA	67	Documentaries, International Movies	Declassified documents reveal the post-WWII li...	2021	
45	s46	Movie	My Heroes Were Cowboys	Tyler Greco	NaN	United States	2021-09-16	2021	PG	23	Documentaries	Robin Wiltshire's painful childhood was rescue...	2021	
91	s92	Movie	The Women and the Murderer	Mona Achache, Patricia Tourancheau	NaN	France	2021-09-09	2021	TV-14	92	Documentaries, International Movies	This documentary traces the capture of serial ...	2021	
101	s102	Movie	Untold: Breaking Point	Chapman Way, Maclain Way	NaN	United States	2021-09-07	2021	TV-MA	80	Documentaries, Sports Movies	Under pressure to continue a winning tradition...	2021	
...
8739	s8740	Movie	Why We Fight: The Battle of Russia	Frank Capra, Anatole Litvak	NaN	United States	2017-03-31	1943	TV-PG	82	Documentaries	This installment of Frank Capra's acclaimed do...	2017	
8744	s8745	Movie	Williams	Morgan Matthews	NaN	United Kingdom	2017-11-04	2017	TV-MA	105	Documentaries, International Movies, Sports Mo...	This inspiring documentary charts the birth of...	2017	
8746	s8747	Movie	Winnie	Pascale Lamche	NaN	France, Netherlands, South Africa, Finland	2018-02-26	2017	TV-14	85	Documentaries, International Movies	Winnie Mandela speaks about her extraordinary	2018	

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_a
8756	s8757	Movie	Woodstock	Barak Goodman	NaN	United States	2019-08-13	2019	TV-MA	97	Documentaries, Music & Musicals	For the 50th anniversary of the legendary Wood...	2019	
8763	s8764	Movie	WWII: Report from the Aleutians	John Huston	NaN	United States	2017-03-31	1943	TV-PG	45	Documentaries	Filmmaker John Huston narrates this Oscar-nomi...	2017	

475 rows × 15 columns



```
In [1844]: moviesDataCast['cast'].fillna("Others", inplace = True)
plt.figure(figsize=(10,10))
sns.countplot(y = moviesDataCast['cast'] , order = moviesDataCast['cast'].value_counts().index[1:6])
plt.title("Top 5 Casts on netflix", fontsize = 15)
plt.show()
```



Overall Top 5 casts on Netflix:

1. Anupam Kher
2. Shah Rukh Khan
3. Naseeruddin Shah
4. Om Puri
5. Akshay Kumar

In [1845]: `moviesDataCast[moviesDataCast['cast'].isna()]`

Out[1845]:

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
---------	------	-------	----------	------	---------	------------	--------------	--------	----------	-----------	-------------	------------	-------------	-----------

```
In [1846]: moviesDataDirector = moviesData.assign(director= moviesData['director'].str.split(', ')).explode('director')
moviesDataDirector
```

Out[1846]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, film...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
6	s7	Movie	My Little Pony: A New Generation	José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmi Ogunlana, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	2021-09-24	1993	TV-MA	125	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...	2021	
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	2021-09-24	2021	PG-13	104	Comedies, Dramas	A woman adjusting to life after a loss contend...	2021	
...
8801	s8802	Movie	Zinzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ...	United Arab Emirates, Jordan	2016-03-09	2015	TV-MA	96	Dramas, International Movies, Thrillers	Recovering alcoholic Talal wakes up inside a s...	2016	
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	2019-11-20	2007	R	158	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...	2019	

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01	2009	R	88	Comedies, Horror Movies	Looking to survive in a world taken over by zo...	2019	
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020	
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	

6854 rows × 15 columns



In [1847]: moviesDataDirectorIndia[moviesDataDirectorIndia['director'].isna()]

Out[1847]:

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
---------	------	-------	----------	------	---------	------------	--------------	--------	----------	-----------	-------------	------------	-------------	-----------

```
In [1848]: moviesDataDirectorUS = moviesDataDirector[moviesDataDirector['country'] == 'United States']
moviesDataDirector
```

Out[1848]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, film...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
6	s7	Movie	My Little Pony: A New Generation	José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmi Ogunlana, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	2021-09-24	1993	TV-MA	125	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...	2021	
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	2021-09-24	2021	PG-13	104	Comedies, Dramas	A woman adjusting to life after a loss contend...	2021	
...
8801	s8802	Movie	Zinzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ...	United Arab Emirates, Jordan	2016-03-09	2015	TV-MA	96	Dramas, International Movies, Thrillers	Recovering alcoholic Talal wakes up inside a s...	2016	
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	2019-11-20	2007	R	158	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...	2019	

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01	2009	R	88	Comedies, Horror Movies	Looking to survive in a world taken over by zo...	2019	
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...	2020	
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	

6854 rows × 15 columns



In [1849]: moviesDataDirectorUS[moviesDataDirectorUS['director'].isna()]

Out[1849]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	mon
404	s405	Movie	9to5: The Story of a Movement	NaN	NaN	United States	2021-07-22	2021	TV-MA	85	Documentaries	In this documentary, female office workers in ...	2021	
470	s471	Movie	Bridgerton - The Afterparty	NaN	David Spade, London Hughes, Fortune Feimster	United States	2021-07-13	2021	TV-14	39	Movies	"Bridgerton" cast members share behind-the-sc...	2021	
483	s484	Movie	Last Summer	NaN	Fatih Şahin, Ece Çeşmioğlu, Halit Özgür Sarı, ...	United States	2021-07-09	2021	TV-MA	102	Dramas, International Movies, Romantic Movies	During summer vacation in a beachside town, 16...	2021	
641	s642	Movie	Sisters on Track	NaN	NaN	United States	2021-06-24	2021	PG	97	Documentaries, Sports Movies	Three track star sisters face obstacles in lif...	2021	
717	s718	Movie	Headspace: Unwind Your Mind	NaN	Andy Puddicombe, Evelyn Lewis Prieto, Ginger D...	United States	2021-06-15	2021	TV-G	273	Documentaries	Do you want to relax, meditate or sleep deeply...	2021	
...
7780	s7781	Movie	Power Rangers Super Samurai: Stuck on Christmas	NaN	Alex Heartman, Najee De-Tiege, Hector David Jr...	United States	2016-01-01	2012	TV-Y7	24	Children & Family Movies, Sci-Fi & Fantasy	The Samurai Rangers get stuck in the Megazord'...	2016	
7781	s7782	Movie	Power Rangers Super Samurai: Trickster Treat	NaN	Alex Heartman, Najee De-Tiege, Hector David Jr...	United States	2016-01-01	2012	TV-Y7	23	Movies	When a Nighlok invades the Samurai Rangers' dr...	2016	

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	mon
7787	s7788	Movie	Power Rangers: Megaforce: Raising Spirits	NaN	Andrew M. Gray, John Mark Loudermilk, Ciara Ha...	United States	2016-01-01	2013	TV-Y7	24	Children & Family Movies	On Halloween, the scariest night of the year, ...	2016	
8000	s8001	Movie	She Did That	NaN	Luvvie Ajayi, Renae Bluitt, Melissa Butler, Li...	United States	2020-02-04	2019	TV-PG	71	Documentaries	Go inside the lives of extraordinary, black fe...	2020	
8044	s8045	Movie	Smash: Motorized Mayhem	NaN	NaN	United States	2017-05-15	2017	TV-MA	66	Documentaries, Sports Movies	Buckle up for a look inside a biannual school ...	2017	

117 rows × 15 columns



In [1850]: `moviesDataDirectorUS['director'].mode()`

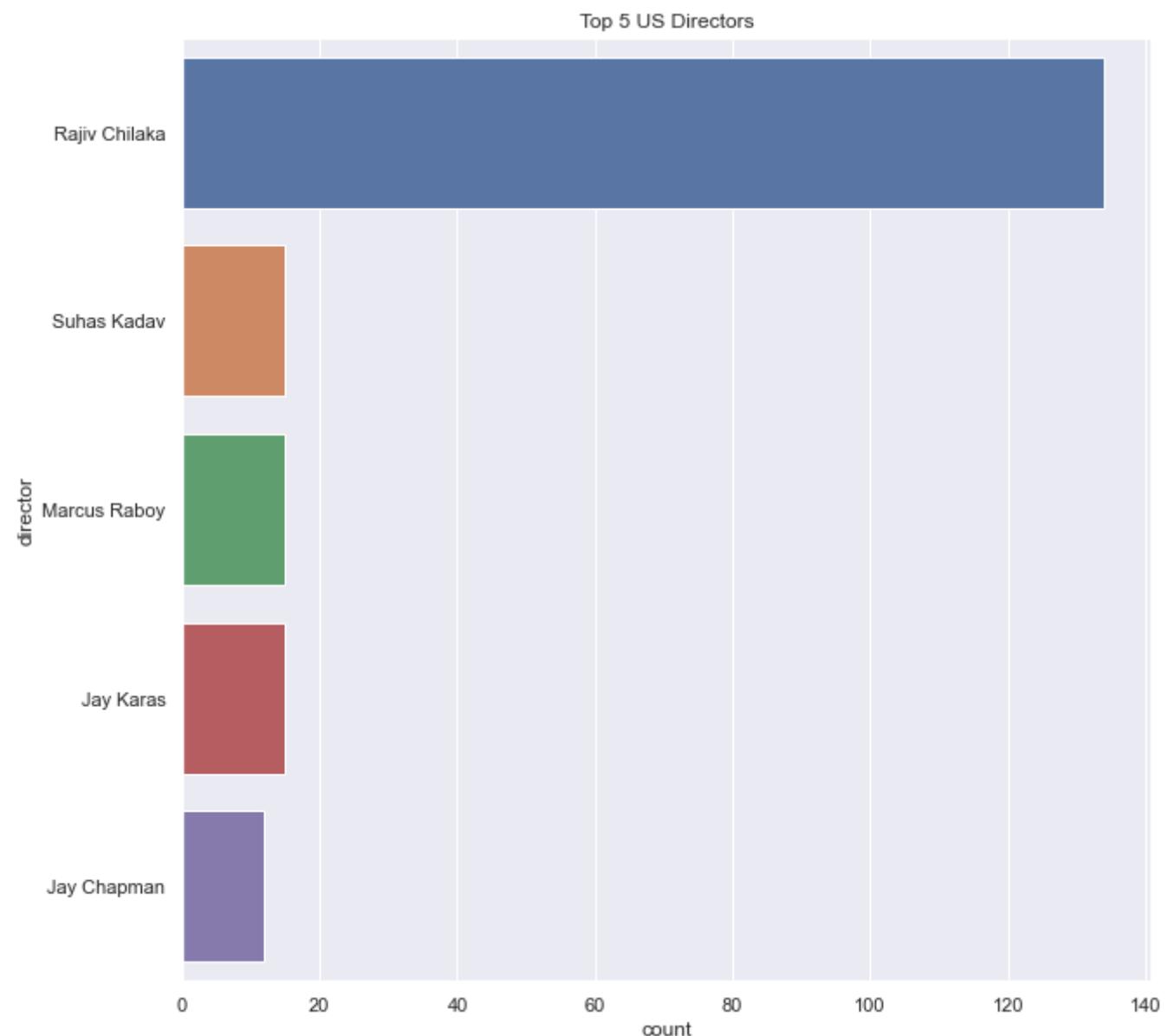
Out[1850]: 0 Rajiv Chilaka
dtype: object

In [1851]: `moviesDataDirectorUS['director'].fillna("Rajiv Chilaka", inplace=True)`

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\generic.py:6392: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

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)
return self._update_inplace(result)

```
In [1852]: plt.figure(figsize=(10,10))
sns.countplot(y = moviesDataDirectorUS['director'] , order = moviesDataDirectorUS['director'].value_counts().index[0:5])
plt.title("Top 5 US Directors", fontsize = 12)
plt.show()
```



Top 5 directors in US:

1. Rajiv Chilaka
2. Suhas Kadav
3. Marcus Raboy
4. Jay Karas
5. Jay Chapman

```
In [1853]: moviesDataDirectorIndia = moviesDataDirector[moviesDataDirector['country'] == 'India']
moviesDataDirectorIndia
```

Out[1853]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_adde
24	s25	Movie	Jeans	S. Shankar	Prashanth, Aishwarya Rai Bachchan, Sri Lakshmi...	India	2021-09-21	1998	TV-14	166	Comedies, International Movies, Romantic Movies	When the father of the man she loves insistst...	2021	
105	s106	Movie	Angamaly Diaries	Lijo Jose Pellissery	Antony Varghese, Reshma Rajan, Binny Rinky Ben...	India	2021-09-05	2017	TV-14	128	Action & Adventure, Comedies, Dramas	After growing up amidst the gang wars of his h...	2021	
114	s115	Movie	Anjaam	Rahul Rawail	Madhuri Dixit, Shah Rukh Khan, Tinnu Anand, Jo...	India	2021-09-02	1994	TV-14	143	Dramas, International Movies, Thrillers	A wealthy industrialist's dangerous obsession ...	2021	
116	s117	Movie	Dhanak	Nagesh Kukunoor	Krrish Chhabria, Hetal Gada, Vipin Sharma, Gul...	India	2021-09-02	2015	TV-PG	114	Comedies, Dramas, Independent Movies	A movie-loving 10-year-old and her blind littl...	2021	
118	s119	Movie	Gurgaon	Shanker Raman	Akshay Oberoi, Pankaj Tripathi, Ragini Khanna,...	India	2021-09-02	2017	TV-14	106	Dramas, International Movies, Thrillers	When the daughter of a wealthy family returns ...	2021	
...
8772	s8773	Movie	Yamla Pagla Deewana 2	Sangeeth Sivan	Dharmendra, Sunny Deol, Bobby Deol, Neha Sharm...	India	2017-05-01	2013	TV-14	147	Action & Adventure, Comedies, International Mo...	Up to his old tricks, con man Dharam poses as ...	2017	
8773	s8774	Movie	Yanda Kartavya Aahe	Kedar Shinde	Ankush Choudhary, Smita Shewale, Mohan Joshi, ...	India	2018-01-01	2006	TV-PG	151	Comedies, Dramas, International Movies	Thanks to an arranged marriage that was design...	2018	

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_adde	
8798	s8799	Movie	Zed Plus	Chandra Prakash Dwivedi	Adil Hussain, Mona Singh, K.K. Raina, Sanjay M...	India	2019-12-31	2014	TV-MA	131	Comedies, Dramas, International Movies	A philandering small-town mechanic's political...	2019	1
8799	s8800	Movie	Zenda	Avadhoot Gupte	Santosh Juvekar, Siddharth Chandekar, Sachit P...	India	2018-02-15	2009	TV-14	120	Dramas, International Movies	A change in the leadership of a political part...	2018	
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	

950 rows × 15 columns



In [1854]: moviesDataDirectorIndia[moviesDataDirectorIndia['director'].isna()]

Out[1854]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month
998	s999	Movie	Searching For Sheela	NaN	Ma Anand Sheela	India	2021-04-22	2021	TV-14	58	Documentaries, International Movies	Journalists and fans await Ma Anand Sheela as ...	2021	
2334	s2335	Movie	Seven (Tamil)	NaN	NaN	India	2020-06-26	2019	TV-MA	111	Dramas, International Movies, Romantic Movies	Multiple women report their husbands as missin...	2020	
2869	s2870	Movie	Amit Tandon: Family Tandoncies	NaN	Amit Tandon	India	2020-02-28	2019	TV-14	72	Stand-Up Comedy	From the death of romance in marriage to the i...	2020	
3196	s3197	Movie	Lakeeran	NaN	Harman Virk, Yuvika Chaudhry, Nirmal Rishi, Sh...	India	2019-12-01	2016	TV-PG	111	Dramas, International Movies, Romantic Movies	Guilt-ridden over his involvement in a fatal c...	2019	
4187	s4188	Movie	Adhugo	NaN	NaN	India	2019-01-15	2018	TV-14	105	Comedies, Independent Movies, International Mo...	A fast-footed piglet named Bunty becomes an ob...	2019	
4465	s4466	Movie	Yeh Hain Bakrapur	NaN	Amit Sial, Yoshika Verma, Anshuman Jha, Asif B...	India	2018-11-01	2014	TV-PG	95	Comedies, Dramas, Independent Movies	A poor family's goat suddenly acquires divine ...	2018	
5295	s5296	Movie	Chadi Jawani Budhe Nu	NaN	Anand	India	2017-09-01	1976	TV-PG	135	Comedies, International Movies, Romantic Movies	After reaching his 40th birthday, an egotistic...	2017	
6707	s6708	Movie	Ethirmarai	NaN	Ramanathan, Rakesh	India	2018-04-01	2018	TV-14	91	International Movies, Music & Musicals, Thrillers	While investigating a brutal string of deaths,...	2018	

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month
6907	s6908	Movie	Haapus	NaN	Makarand Anaspure, Madhura Velankar, Subodh Bhave	India	2018-03-15	2010	TV-14	138	Comedies, Dramas, International Movies	The son of an astrologer faces opposition from...	2018
7168	s7169	Movie	Kanika	NaN	Sharad Ponkshie, Smita Shewale	India	2018-03-01	2017	TV-14	110	Horror Movies, International Movies, Thrillers	A doctor and his colleagues who use sex-determ...	2018
7662	s7663	Movie	One Heart: The A.R. Rahman Concert Film	NaN	A.R. Rahman	India	2017-11-01	2017	TV-G	87	Documentaries, International Movies, Music & M...	In this concert film, Indian musical maestro A...	2017
8647	s8648	Movie	Twisted Trunk, Big Fat Body	NaN	Vijay Maurya, Naman Jain, Usha Nadkarni, Mukes...	India	2017-01-15	2015	TV-14	89	Dramas, International Movies	After terrorists place a bomb inside a toy Lor...	2017

In [1855]: moviesDataDirectorIndia['director'].mode()

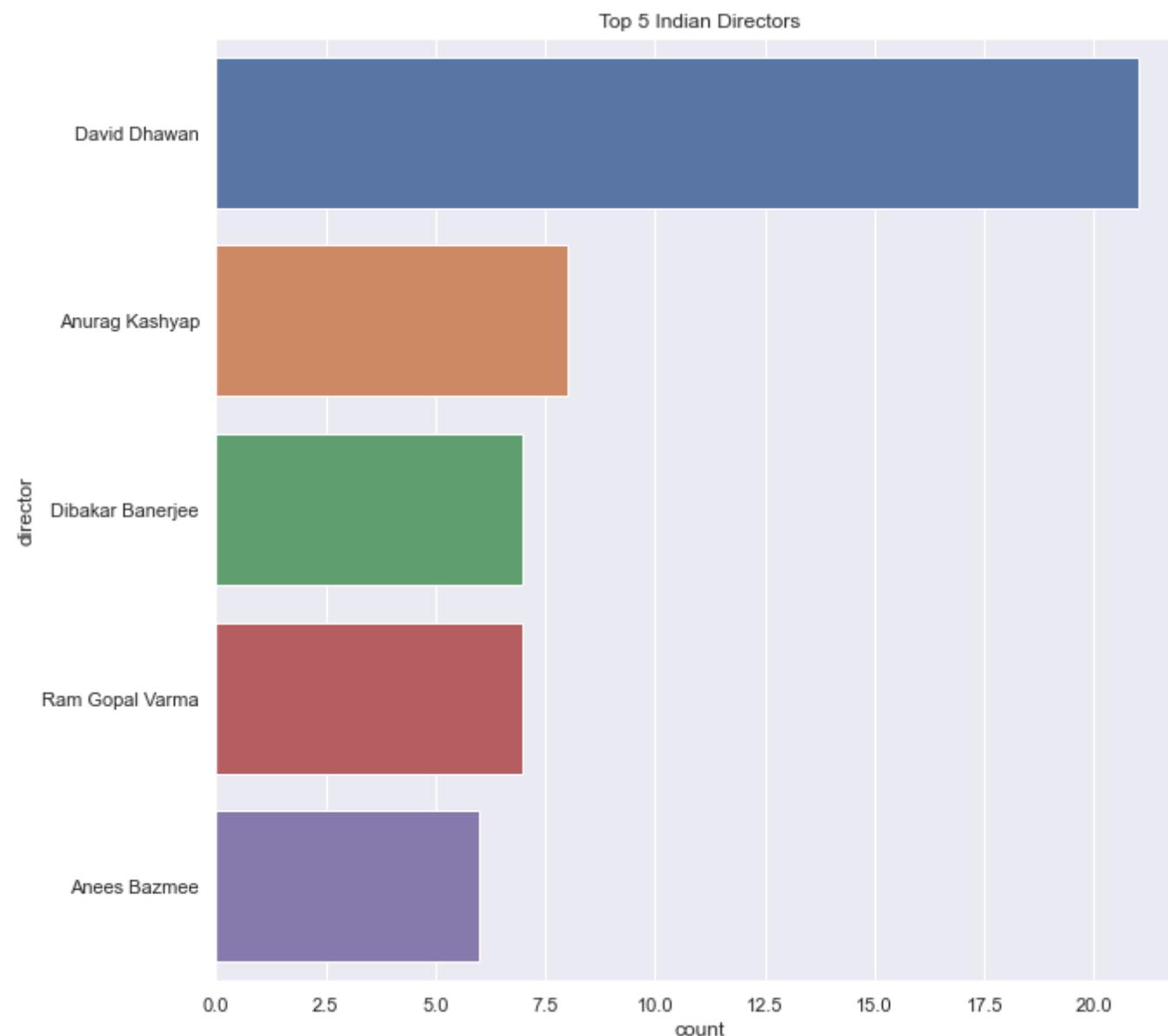
Out[1855]: 0 David Dhawan
dtype: object

In [1856]: moviesDataDirectorIndia['director'].fillna("David Dhawan", inplace= True)

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\generic.py:6392: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

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)
return self._update_inplace(result)

```
In [1857]: plt.figure(figsize=(10,10))
sns.countplot(y = moviesDataDirectorIndia['director'] , order = moviesDataDirectorIndia['director'].value_counts().index[0:5])
plt.title("Top 5 Indian Directors", fontsize = 12)
plt.show()
```

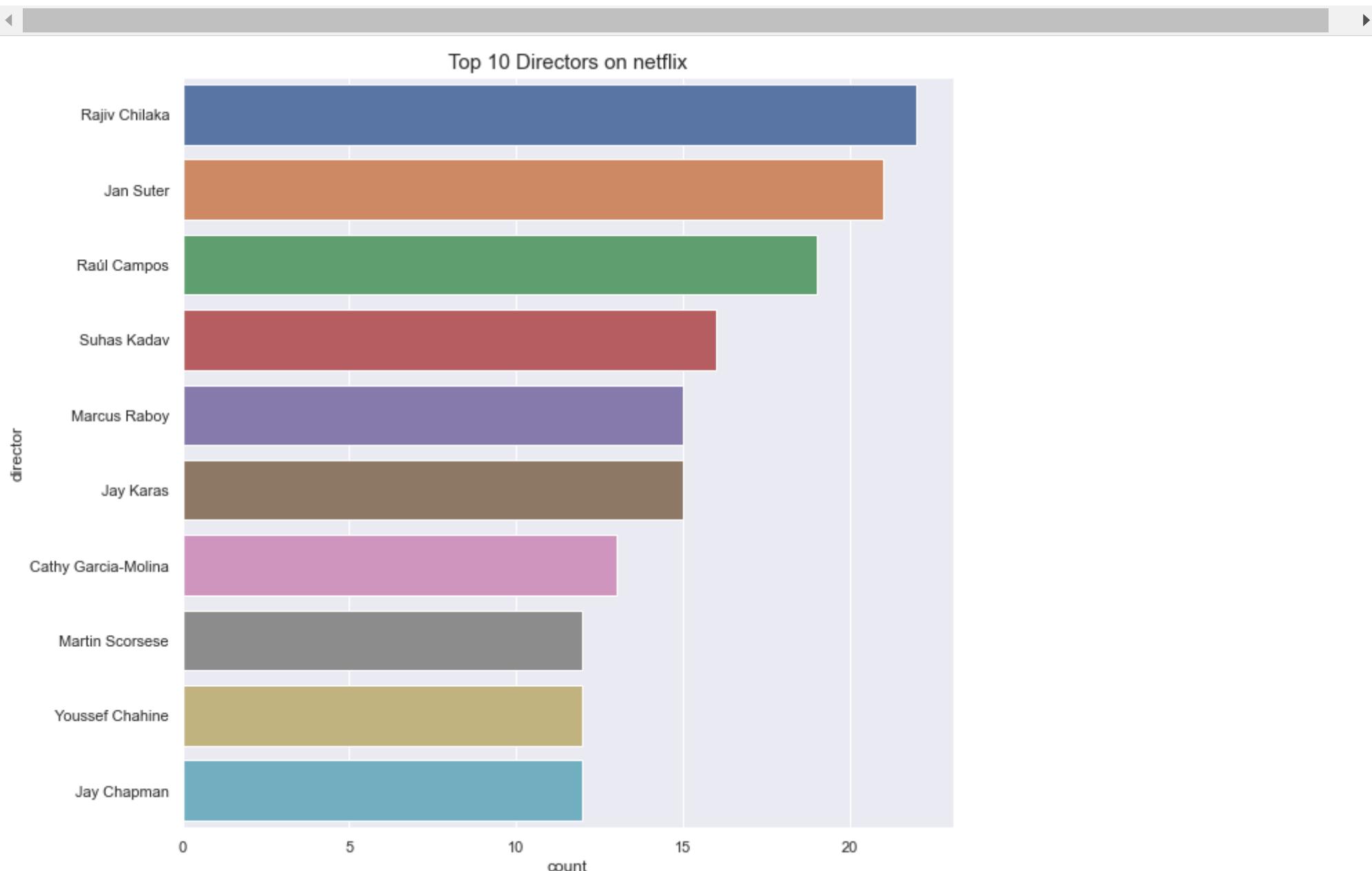


Top 5 Directors in India:

1. David Dhawan
2. Anurag Kashyap
3. Dibakar Banerjee
4. Ram Gopal Varma
5. Anees Bazmee

In [1858]: # Without considering Na values Overall top 5 actors

```
plt.figure(figsize=(10,10))
sns.countplot(y = moviesDataDirector['director'] , order = moviesDataDirector['director'].value_counts(ascending=False).index[0:10]
plt.title("Top 10 Directors on netflix", fontsize = 15)
plt.show()
```



Top 10 Directors on Netflix:

1. Rajiv Chilaka
2. Jan Suter
3. Raul Campos
4. Suhas Kadav
5. Mercus Raboy
6. Jay karas
7. Cathy Garcia-Molina
8. Martin Scorsese
9. Youssef Chahine
10. Jay Chapman

In []:

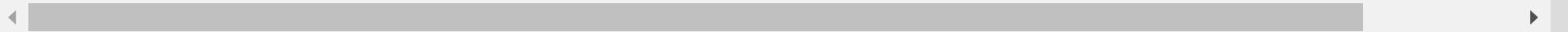
```
In [1859]: moviesDataListedIn = moviesData.assign(listed_in= moviesData['listed_in'].str.split(', ')).explode('listed_in')  
moviesDataListedIn
```

Out[1859]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, film...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmi Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	2021-09-24	1993	TV-MA	125	Dramas	On a photo shoot in Ghana, an American model s...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmi Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	2021-09-24	1993	TV-MA	125	Independent Movies	On a photo shoot in Ghana, an American model s...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmi Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	2021-09-24	1993	TV-MA	125	International Movies	On a photo shoot in Ghana, an American model s...	2021	
...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88	Children & Family Movies	Dragged from civilian life, a former superhero...	2020	

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88	Comedies	Dragged from civilian life, a former superhero...	2020	
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	Dramas	A scrappy but poor boy worms his way into a ty...	2019	
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	International Movies	A scrappy but poor boy worms his way into a ty...	2019	
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	

13190 rows × 15 columns



In [1860]: moviesDataListedIn[moviesDataListedIn['listed_in'].isna()]

Out[1860]:

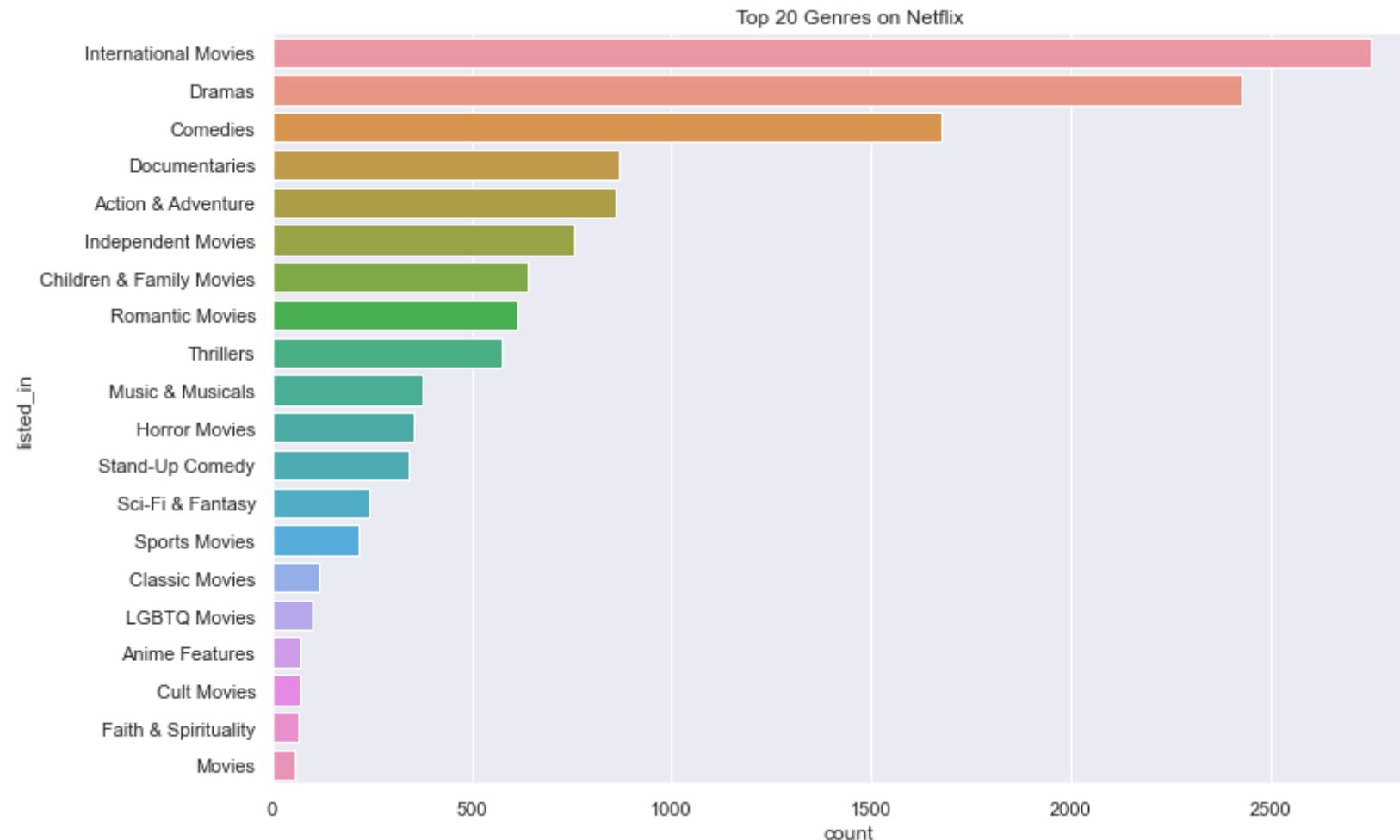
show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
---------	------	-------	----------	------	---------	------------	--------------	--------	----------	-----------	-------------	------------	-------------	-----------

```
In [1861]: dff = pd.DataFrame(moviesDataListedIn['listed_in'].value_counts())
dff.reset_index(inplace=True)
dff.columns = ['listed_in', 'count']
dff
dff['percentage'] = dff['count'] * 100 / dff['count'].sum()
dff
```

Out[1861]:

	listed_in	count	percentage
0	International Movies	2752	20.864291
1	Dramas	2427	18.400303
2	Comedies	1674	12.691433
3	Documentaries	869	6.588324
4	Action & Adventure	859	6.512509
5	Independent Movies	756	5.731615
6	Children & Family Movies	641	4.859742
7	Romantic Movies	616	4.670205
8	Thrillers	577	4.374526
9	Music & Musicals	375	2.843063
10	Horror Movies	357	2.706596
11	Stand-Up Comedy	343	2.600455
12	Sci-Fi & Fantasy	243	1.842305
13	Sports Movies	219	1.660349
14	Classic Movies	116	0.879454
15	LGBTQ Movies	102	0.773313
16	Anime Features	71	0.538287
17	Cult Movies	71	0.538287
18	Faith & Spirituality	65	0.492798
19	Movies	57	0.432146

```
In [1862]: plt.figure(figsize = (12,8))
sns.countplot(y=moviesDataListedIn['listed_in'],order = moviesDataListedIn['listed_in'].value_counts().index[0:20],data = data)
plt.title('Top 20 Genres on Netflix', fontsize = 12)
plt.show()
```



For movies most watched genre on Netflix are following:

1. International Movies 20%
2. Dramas 18%
3. Comedies 11%
4. Documentries 7%
5. Action and Adventure 6%

In [1863]:

```
df = moviesDataListedIn[['release_year', 'listed_in']]  
listedIn = pd.DataFrame(df.groupby(['release_year', 'listed_in']).size().reset_index())  
listedIn .rename({0 : "Total Content"}, axis = 1, inplace= True)  
listedIn
```

Out[1863]:

	release_year	listed_in	Total Content
0	1942	Classic Movies	2
1	1942	Documentaries	2
2	1943	Classic Movies	1
3	1943	Documentaries	3
4	1944	Classic Movies	3
...
768	2021	Music & Musicals	19
769	2021	Romantic Movies	25
770	2021	Sports Movies	15
771	2021	Stand-Up Comedy	12
772	2021	Thrillers	33

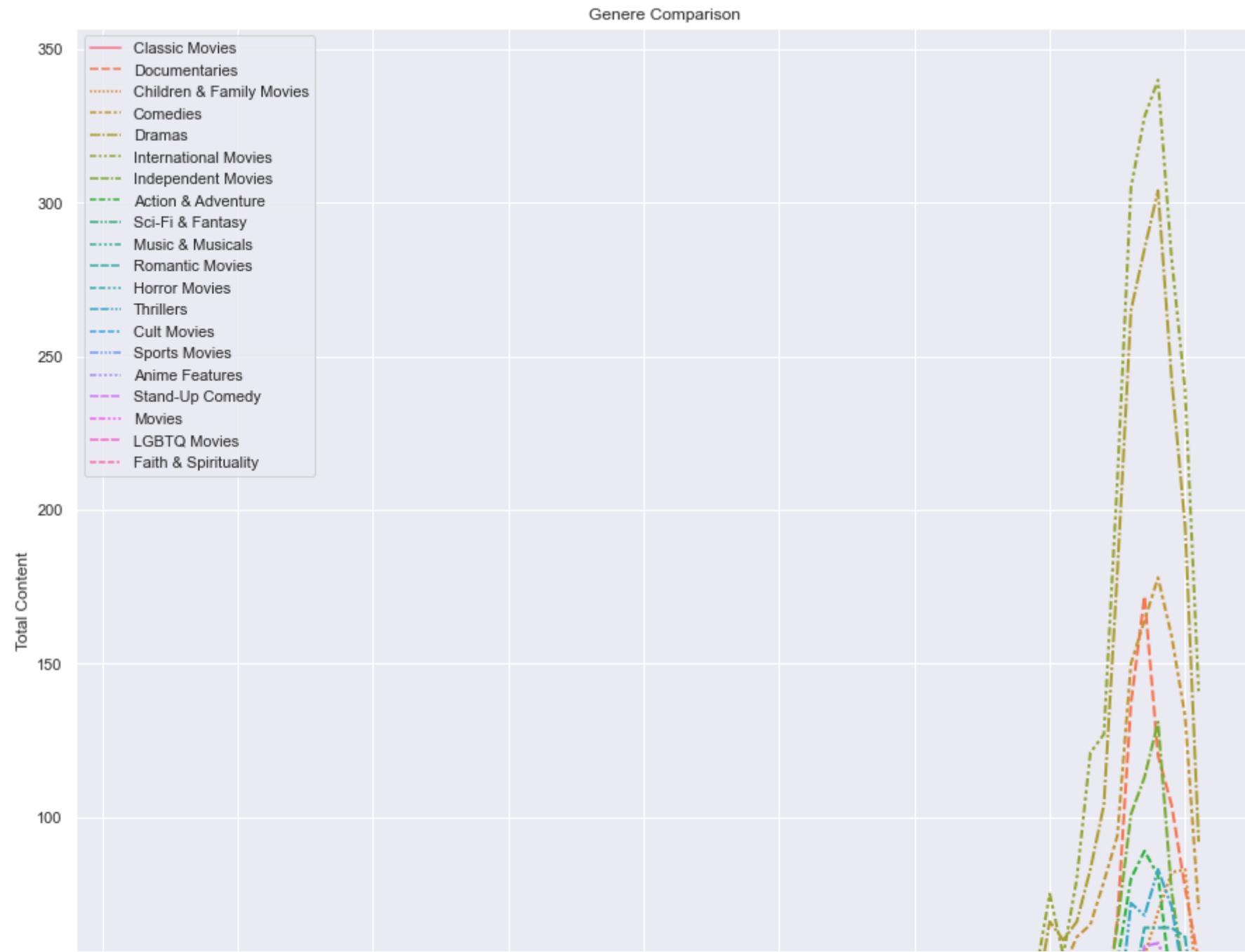
773 rows × 3 columns

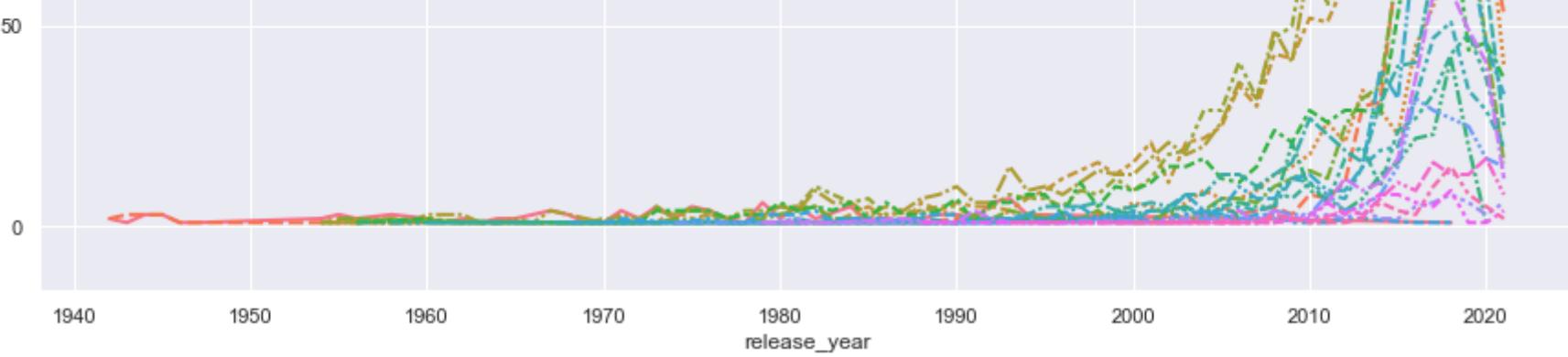
```
In [1864]: listedIn['listed_in'].value_counts()
```

```
Out[1864]: Dramas          65  
International Movies    62  
Comedies             58  
Action & Adventure    57  
Music & Musicals       45  
Classic Movies         44  
Romantic Movies        43  
Children & Family Movies 42  
Independent Movies     41  
Horror Movies          38  
Cult Movies            37  
Documentaries          36  
Thrillers              36  
Sports Movies           33  
Sci-Fi & Fantasy        32  
Stand-Up Comedy         29  
Anime Features          27  
Faith & Spirituality    17  
LGBTQ Movies            16  
Movies                  15  
Name: listed_in, dtype: int64
```

```
In [ ]:
```

```
In [1865]: plt.figure(figsize=(15,15))
sns.lineplot(x = 'release_year', y = 'Total Content', data = listedIn, hue = 'listed_in', style= 'listed_in' , lw =2)
plt.legend(loc='upper left')
plt.title("Genre Comparison")
plt.show()
```





From above graph we can see in International movies movies released after 2014/2015 are popular as compared to ovies released before 2010 and after 2020

In [1866]:

```
moviesDataListedIn = moviesDataListedIn.assign(country= moviesData['country'].str.split(', ')).explode('country')
moviesDataListedIn
```

Out[1866]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Nan	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, filmm...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States	2021-09-24	1993	TV-MA	125	Dramas	On a photo shoot in Ghana, an American model s...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	Ghana	2021-09-24	1993	TV-MA	125	Dramas	On a photo shoot in Ghana, an American model s...	2021	
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	Burkina Faso	2021-09-24	1993	TV-MA	125	Dramas	On a photo shoot in Ghana, an American model s...	2021	
...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88	Children & Family Movies	Dragged from civilian life, a former superhero...	2020	
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88	Comedies	Dragged from civilian life, a former superhero...	2020	

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	Dramas	A scrappy but poor boy worms his way into a ty...	2019
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	International Movies	A scrappy but poor boy worms his way into a ty...	2019
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111	Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019

16873 rows × 15 columns



In []:

In [1867]: `top10MovieCountries = dict(moviesDataListedIn['country'].value_counts(ascending = False)[:10])`

In [1868]: `moviesDataListedIn[moviesDataListedIn['country'].isna()]`

Out[1868]:

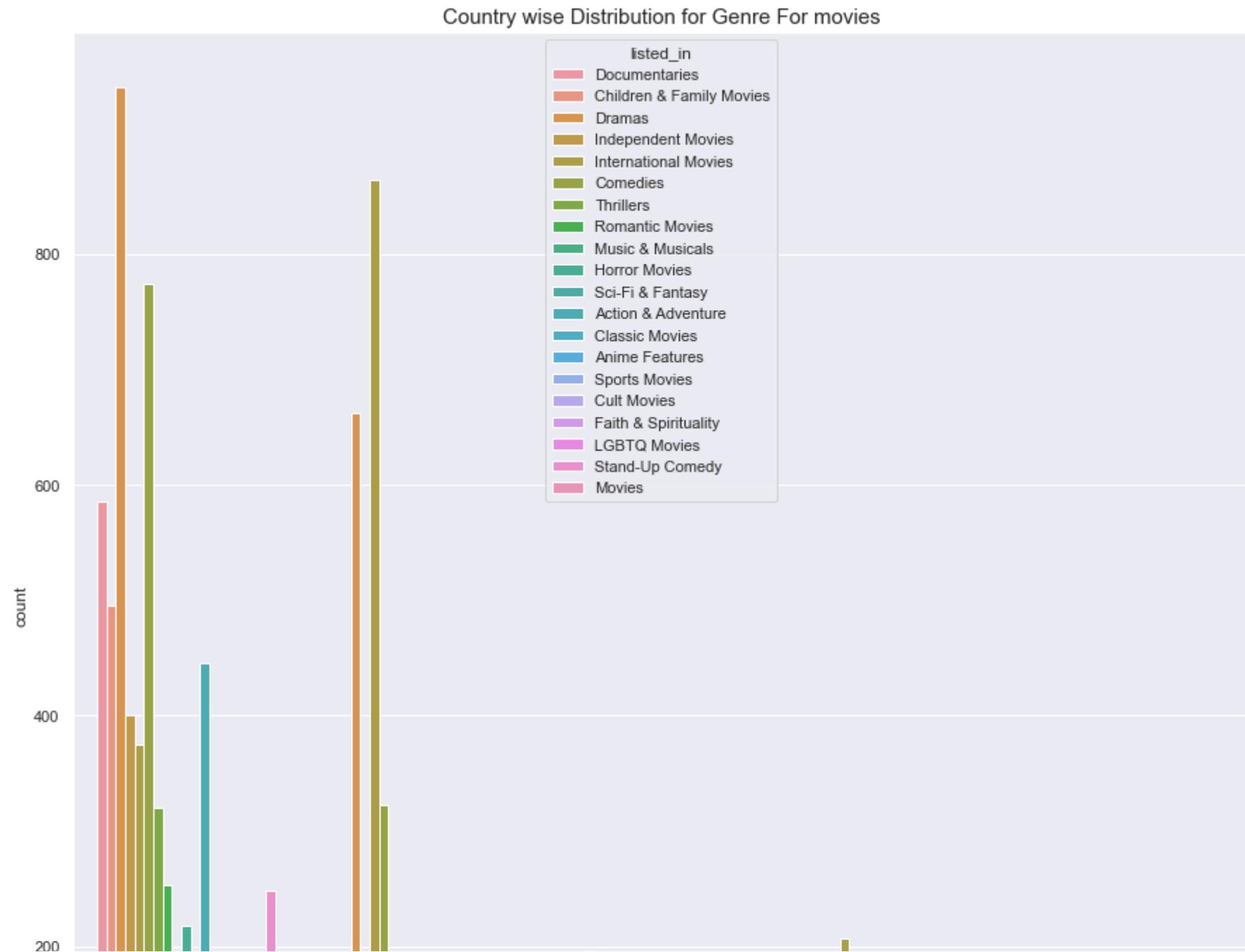
show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
---------	------	-------	----------	------	---------	------------	--------------	--------	----------	-----------	-------------	------------	-------------	-----------

In [1869]: `Top10Countries = list(top10MovieCountries.keys())`

In [1870]: Top10Countries

Out[1870]: ['United States',
 'India',
 'United Kingdom',
 'France',
 'Canada',
 'Spain',
 'Germany',
 'Japan',
 'China',
 'Egypt']

```
In [1871]: plt.figure(figsize=(15,15))
sns.countplot(x = 'country', data = moviesDataListedIn, order = moviesDataListedIn['country'].value_counts(ascending= False).index
plt.title("Country wise Distribution for Genre For movies", fontsize = 15)
plt.show()
```





Above graph is analysis for top 5 countries for Movies:

1. For United States Dramas, Comedies, Documentries are Most popular Genre while Anime feature is least watched genre
2. In India International Movies, Drama and Comedy movies are popular
3. In united Kingdom drama is most watched content
4. In france and canada International Movies

In []:

```
In [1872]: moviesData = moviesData.assign(director= moviesData['director'].str.split(', ')).explode('director')
moviesData = moviesData.assign(cast= moviesData['cast'].str.split(', ')).explode('cast')
moviesData = moviesData.assign(listed_in= moviesData['listed_in'].str.split(', ')).explode('listed_in')
moviesData = moviesData.assign(cast= moviesData['cast'].str.split(', ')).explode('cast')
```

In [1873]: moviesData

Out[1873]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90	Documentaries	As her father nears the end of his life, filmm...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen	Vanessa Hudgens	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen	Kimiko Glenn	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen	James Marsden	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen	Sofia Carson	United States	2021-09-24	2021	PG	91	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...	2021	
...	
8806	s8807	Movie	Zubaan	Mozez Singh	Anita Shabdish	India	2019-03-02	2015	TV-14	111	International Movies	A scrappy but poor boy worms his way into a ty...	2019	
8806	s8807	Movie	Zubaan	Mozez Singh	Anita Shabdish	India	2019-03-02	2015	TV-14	111	Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019	
8806	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	2019-03-02	2015	TV-14	111	Dramas	A scrappy but poor boy worms his way into a ty...	2019	

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8806	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	2019-03-02	2015	TV-14	111	International Movies	A scrappy but poor boy worms his way into a ty...	2019
8806	s8807	Movie	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	2019-03-02	2015	TV-14	111	Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019

111989 rows × 15 columns

20 oldest movies on Netflix

```
In [1874]: movie_asc = moviesData.sort_values("release_year", ascending = True)
movie_asc[['title', "release_year"]][:20]
```

Out[1874]:

	title	release_year
8205	The Battle of Midway	1942
8205	The Battle of Midway	1942
7790	Prelude to War	1942
7790	Prelude to War	1942
8205	The Battle of Midway	1942
8205	The Battle of Midway	1942
8763	WWII: Report from the Aleutians	1943
8739	Why We Fight: The Battle of Russia	1943
8660	Undercover: How to Operate Behind Enemy Lines	1943
8739	Why We Fight: The Battle of Russia	1943
8660	Undercover: How to Operate Behind Enemy Lines	1943
8640	Tunisian Victory	1944
8640	Tunisian Victory	1944
8419	The Memphis Belle: A Story of a\nFlying Fortress	1944
8640	Tunisian Victory	1944

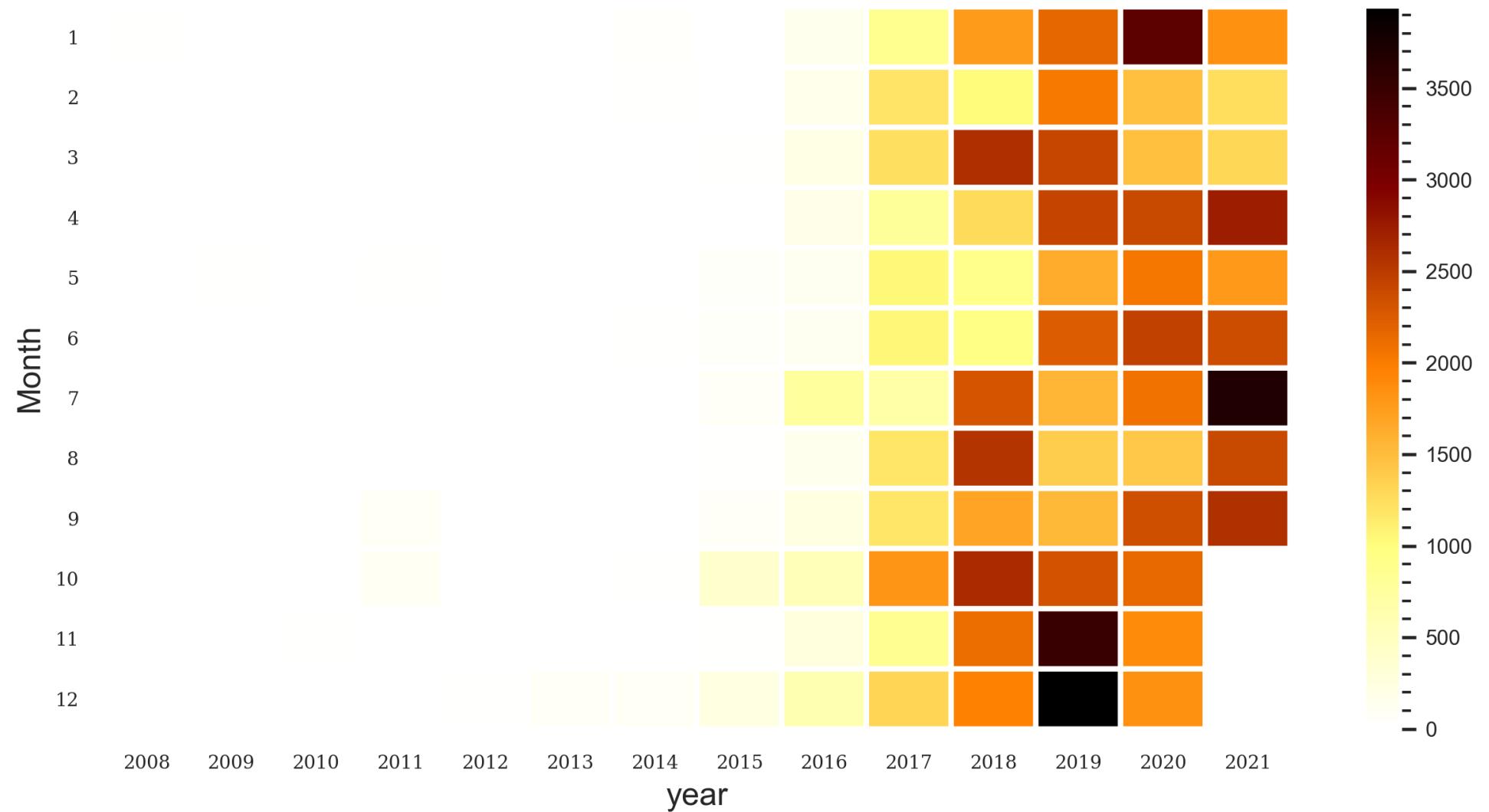
In []:

In []:

In []:

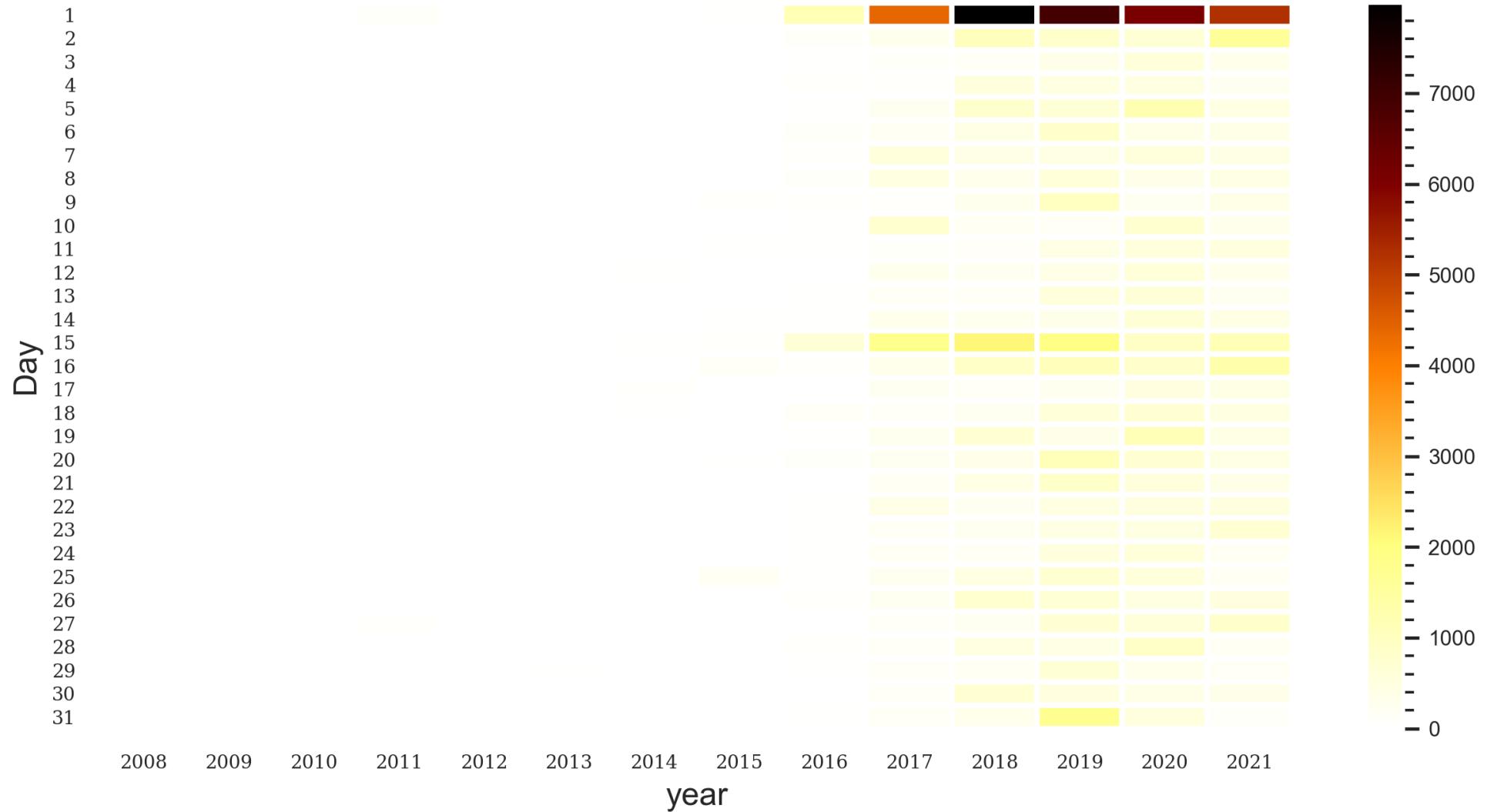
```
In [1875]: month_order = [1,2,3,4,5,6,7,8,9,10,11,12][::-1]
month_HM = moviesData.groupby('year_added')['month_added'].value_counts().unstack().fillna(0)[month_order].T
plt.figure(figsize=(10, 5), dpi=300)
plt.pcolor(month_HM, cmap='afmhot_r', edgecolors='white', linewidths=2) # heatmap
plt.xlabel("year")
plt.ylabel("Month")
plt.xticks(np.arange(0.5, len(month_HM.columns), 1), month_HM.columns, fontsize=7, fontfamily='serif')
plt.yticks(np.arange(0.5, len(month_HM.index), 1), month_HM.index, fontsize=7, fontfamily='serif')
plt.title('Netflix Monthly Contents Update For movies', fontsize=12, fontweight='bold', position=(0.40, 1.0+0.02))
cbar = plt.colorbar()
cbar.ax.tick_params(labelsize=8)
cbar.ax.minorticks_on()
plt.show()
```

Netflix Monthly Contents Update For movies



```
In [1876]: month_order = [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31][::-1]
month_HM = moviesData.groupby('year_added')['day_added'].value_counts().unstack().fillna(0)[month_order].T
plt.figure(figsize=(10, 5), dpi=300)
plt.pcolor(month_HM, cmap='afmhot_r', edgecolors='white', linewidths=2) # heatmap
plt.xlabel("year")
plt.ylabel("Day")
plt.xticks(np.arange(0.5, len(month_HM.columns), 1), month_HM.columns, fontsize=7, fontfamily='serif')
plt.yticks(np.arange(0.5, len(month_HM.index), 1), month_HM.index, fontsize=7, fontfamily='serif')
plt.title('Netflix Day wise Contents Update For movies', fontsize=12, fontweight='bold', position=(0.40, 1.0+0.02))
cbar = plt.colorbar()
cbar.ax.tick_params(labelsize=8)
cbar.ax.minorticks_on()
plt.show()
```

Netflix Day wise Contents Update For movies



For movies here we can see that In 2019 in the last few months Movie's added in large quantity and in Jan 2020 --> Covid 19 Started in December 2019
We can say that as movie theaters were closed so It was the Growth period for netflix

And from day analysis we can see that most of the movies gets added on netflix on 1 st of the month

In []:

In []:

TV Shows Analysis

```
In [1877]: tvSDData = pd.DataFrame(data[data['type'] == 'TV Show'])
tvSDData.reset_index()
```

Out[1877]:

0	1	s2	TV Show	Blood & Water	Nan	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021						
1	2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...	2021						
2	3	s4	TV Show	Jailbirds New Orleans	Nan	Nan	United States	2021-09-24	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...	2021						
3	4	s5	TV Show	Kota Factory	Nan	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train I...	2021						
4	5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	United States	2021-09-24	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries	The arrival of a charismatic young priest brin...	2021						
...	
2671	8795	s8796	TV Show	Yu-Gi-Oh! Arc-V	Nan	Mike Liscio, Emily Bauer, Billy Bob Thompson, ...	Japan, Canada	2018-05-01	2015	TV-Y7	2 Seasons	Anime Series, Kids' TV	Now that he's discovered the Pendulum Summonin...	2018						

index	show_id	type		title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	r
2672	8796	s8797	TV Show	Yunus Emre	NaN	Gökhan Atalay, Payidar Tüfekçioğlu, Baran Akbu...	Turkey	2017-01-17	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas	During the Mongol invasions, Yunus Emre leaves...	2017	
2673	8797	s8798	TV Show	Zak Storm	NaN	Michael Johnston, Jessica Gee-George, Christin...	United States, France, South Korea, Indonesia	2018-09-13	2016	TV-Y7	3 Seasons	Kids' TV	Teen surfer Zak Storm is mysteriously transpor...	2018	
2674	8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ...	Pakistan	2016-12-15	2012	TV-PG	1 Season	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	
2675	8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...	2019	

2676 rows × 16 columns



```
In [1878]: tvShowListedIn = tvSData.assign(listed_in= tvSData['listed_in'].str.split(', ')).explode('listed_in')
tvShowListedIn
```

Out[1878]:

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
1	s2	TV Show	Blood & Water	Nan Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows	After crossing paths at a party, a Cape Town t...	2021	9
1	s2	TV Show	Blood & Water	Nan Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	TV Dramas	After crossing paths at a party, a Cape Town t...	2021	9
1	s2	TV Show	Blood & Water	Nan Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9
2	s3	TV Show	Ganglands	Julien Leclercq Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1 Season	Crime TV Shows	To protect his family from a powerful drug lor...	2021	9
2	s3	TV Show	Ganglands	Julien Leclercq Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1 Season	International TV Shows	To protect his family from a powerful drug lor...	2021	9

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ...	Pakistan	2016-12-15	2012	TV-PG	1 Season	Romantic TV Shows	Strong-willed, middle-class Kashaf and carefree...	2016	12
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ...	Pakistan	2016-12-15	2012	TV-PG	1 Season	TV Dramas	Strong-willed, middle-class Kashaf and carefree...	2016	12
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	Kids' TV	While living alone in a spooky town, a young g...	2019	7
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	Korean TV Shows	While living alone in a spooky town, a young g...	2019	7
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	TV Comedies	While living alone in a spooky town, a young g...	2019	7

6133 rows × 15 columns



```
In [1879]: tvSDData['duration'] = tvSDData['duration'].apply(lambda x: str(x).split(' ')[0])  
tvSDData
```

Out[1879]:

		show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021		
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...	2021		
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	United States	2021-09-24	2021	TV-MA	1	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...	2021		
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021	TV-MA	2	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train l...	2021		
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	United States	2021-09-24	2021	TV-MA	1	TV Dramas, TV Horror, TV Mysteries	The arrival of a charismatic young priest brin...	2021		
...
8795	s8796	TV Show	Yu-Gi-Oh! Arc-V	NaN	Mike Liscio, Emily Bauer, Billy Bob Thompson, ...	Japan, Canada	2018-05-01	2015	TV-Y7	2	Anime Series, Kids' TV	Now that he's discovered the Pendulum Summonin...	2018		

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8796	s8797	TV Show	Yunus Emre	NaN	Gökhan Atalay, Payidar Tüfekçioğlu, Baran Akbu...	Turkey	2017-01-17	2016	TV-PG	2	International TV Shows, TV Dramas	During the Mongol invasions, Yunus Emre leaves...	2017	
8797	s8798	TV Show	Zak Storm	NaN	Michael Johnston, Jessica Gee-George, Christin...	United States, France, South Korea, Indonesia	2018-09-13	2016	TV-Y7	3	Kids' TV	Teen surfer Zak Storm is mysteriously transpor...	2018	
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ...	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	1
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...	2019	

2676 rows × 15 columns



```
In [1880]: tvSDataCountry = tvSData.assign(country=tvSData['country'].str.split(',', ')).explode('country'))
tvSDataCountry
```

Out[1880]:

		show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9	
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...	2021	9	
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	United States	2021-09-24	2021	TV-MA	1	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...	2021	9	
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021	TV-MA	2	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train I...	2021	9	
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	United States	2021-09-24	2021	TV-MA	1	TV Dramas, TV Horror, TV Mysteries	The arrival of a charismatic young priest brin...	2021	9	
...	
8797	s8798	TV Show	Zak Storm	NaN	Michael Johnston, Jessica Gee-George, Christin...	France	2018-09-13	2016	TV-Y7	3	Kids' TV	Teen surfer Zak Storm is mysteriously transpor...	2018	9	

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8797	s8798	TV Show	Zak Storm	NaN	Michael Johnston, Jessica Gee-George, Christin...	South Korea	2018-09-13	2016	TV-Y7	3	Kids' TV	Teen surfer Zak Storm is mysteriously transpor...	2018	9
8797	s8798	TV Show	Zak Storm	NaN	Michael Johnston, Jessica Gee-George, Christin...	Indonesia	2018-09-13	2016	TV-Y7	3	Kids' TV	Teen surfer Zak Storm is mysteriously transpor...	2018	9
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ...	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	12
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...	2019	7

3031 rows × 15 columns



In [1881]: tvSDataCountry.info()

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 3031 entries, 1 to 8803
Data columns (total 15 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   show_id     3031 non-null   object  
 1   type        3031 non-null   category
 2   title       3031 non-null   object  
 3   director    272 non-null   object  
 4   cast        2655 non-null   object  
 5   country     3031 non-null   object  
 6   date_added  3031 non-null   datetime64[ns]
 7   release_year 3031 non-null   int64   
 8   rating      3031 non-null   category
 9   duration    3031 non-null   object  
 10  listed_in   3031 non-null   object  
 11  description 3031 non-null   object  
 12  year_added  3031 non-null   int64   
 13  month_added 3031 non-null   int64   
 14  day_added   3031 non-null   int64   
dtypes: category(2), datetime64[ns](1), int64(4), object(8)
memory usage: 338.2+ KB
```

In [1882]: # Country wise analysis for TV Shows

```
plt.figure(figsize=(15, 15))
sns.countplot(y = tvSDataCountry['country'],order = tvSDataCountry['country'].value_counts().index[0:10])

plt.title("Top 10 Countries for TV Shows" , fontsize=15)
plt.show()
```



Top 5 countries for TV shows:

1. United States
2. Japan
3. United Kingdom
4. Sout Korea
5. Canada

```
In [1883]: tvSDataCountry['country'].value_counts(ascending = False)
```

```
Out[1883]: United States      1329  
United Kingdom       272  
Japan                199  
South Korea          170  
Canada               126  
...  
Malta                 1  
Belarus               1  
United Arab Emirates  1  
Uruguay               1  
Switzerland           1  
Name: country, Length: 66, dtype: int64
```

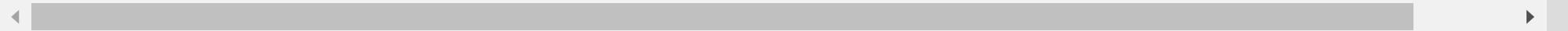
```
In [1884]: tvSDataCountry = tvSDataCountry.assign(cast= tvSDataCountry['cast'].str.split(', ')).explode('cast')
tvSDataCountry
```

Out[1884]:

		show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
1	s2	TV Show	Blood & Water		NaN	Ama Qamata	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9
1	s2	TV Show	Blood & Water		NaN	Khosi Ngema	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9
1	s2	TV Show	Blood & Water		NaN	Gail Mabalane	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9
1	s2	TV Show	Blood & Water		NaN	Thabang Molaba	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9
1	s2	TV Show	Blood & Water		NaN	Dillon Windvogel	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9
...
8800	s8801	TV Show	Zindagi Gulzar Hai		NaN	Samina Peerzada	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	12

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	
8800	s8801	TV Show	Zindagi Gulzar Hai	Nan	Waseem Abbas	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	12
8800	s8801	TV Show	Zindagi Gulzar Hai	Nan	Javed Sheikh	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	12
8800	s8801	TV Show	Zindagi Gulzar Hai	Nan	Hina Khawaja Bayat	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	12
8803	s8804	TV Show	Zombie Dumb	Nan	Nan	United States	2019-07-01	2018	TV-Y7	2	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...	2019	7

23177 rows × 15 columns



```
In [1885]: tvSDataCountry = tvSDataCountry.assign(director= tvSDataCountry['director'].str.split(', ')).explode('director')
tvSDataCountry
```

Out[1885]:

		show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day
1	s2	TV Show	Blood & Water	NaN	Ama Qamata	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9		
1	s2	TV Show	Blood & Water	NaN	Khosi Ngema	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9		
1	s2	TV Show	Blood & Water	NaN	Gail Mabalane	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9		
1	s2	TV Show	Blood & Water	NaN	Thabang Molaba	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9		
1	s2	TV Show	Blood & Water	NaN	Dillon Windvogel	South Africa	2021-09-24	2021	TV-MA	2	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9		
...
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Samina Peerzada	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	12		

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Waseem Abbas	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	12	
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Javed Sheikh	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	12	
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Hina Khawaja Bayat	Pakistan	2016-12-15	2012	TV-PG	1	International TV Shows, Romantic TV Shows, TV ...	Strong-willed, middle-class Kashaf and carefre...	2016	12	
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...	2019	7	

24036 rows × 15 columns



```
In [1886]: tvSDDataCountryUs = tvSDDataCountry[tvSDDataCountry['country'] == 'United States']
```

```
In [ ]:
```

```
In [1887]: modCastUs = pd.DataFrame(tvSDataCountry.groupby('country').apply(lambda x : x['cast'].mode())))
modCastUs.loc[['United States', 'United Kingdom', 'Japan', 'South Korea', 'Canada']]
```

Out[1887]:

	0	
	country	
United States	0	Grey Griffin
	1	Vincent Tong
United Kingdom	0	David Attenborough
Japan	0	Takahiro Sakurai
South Korea	0	Bae Doona
	1	Cho Seong-ha
	2	Kim Tae-hun
	3	Kim Won-hae
	4	Lee Ji-eun (IU)
	5	Sung Dong-il
Canada	0	John Paul Tremblay
	1	Robb Wells

```
In [1888]: modDirectors = pd.DataFrame(tvSDataCountry.groupby('country').apply(lambda x : x['director'].mode()))
```

```
In [1889]: modDirectors.loc[['United States', 'United Kingdom', 'Japan', 'South Korea', 'Canada']]
```

```
Out[1889]:
```

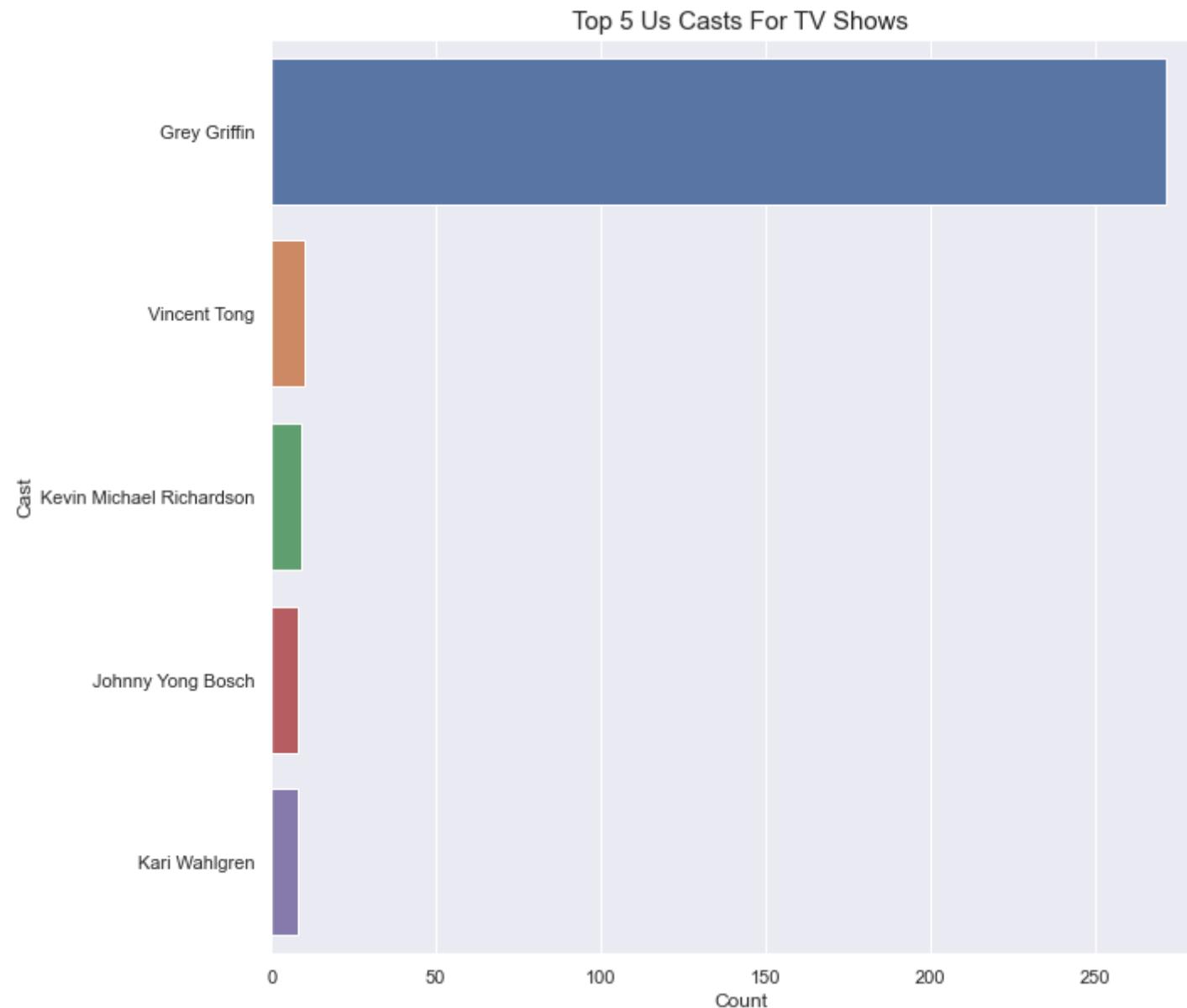
country	0	
United States	0	Rob Seidenglanz
United Kingdom	0	Noam Murro
Japan	0	Yasuhiro Irie
South Korea	0	Thomas Astruc
Canada	0	Stefan Brogren

```
In [1890]: tvSDataCountryUs['cast'].fillna("Grey Griffin", inplace = True)
tvSDataCountryUs['director'].fillna("Rob Seidenglanz", inplace = True)
```

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\generic.py:6392: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

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

```
In [1891]: plt.figure(figsize=(10,10))
sns.countplot(y = tvSDataCountryUs['cast'] , order = tvSDataCountryUs['cast'].value_counts().index[0:5])
plt.xlabel("Count" ,fontsize = 12)
plt.ylabel("Cast" ,fontsize = 12)
plt.title("Top 5 Us Casts For TV Shows" , fontsize = 15)
plt.show()
```



Top 5 US Casts for TV Shows:

1. Grey Griffin
2. Vincent Tong
3. Kevin Michael Richardson
4. Johnny Yong Bosch
5. Kari Wahlgren

```
In [1892]: tvSDataCountryUs.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 9224 entries, 2 to 8803
Data columns (total 15 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   show_id     9224 non-null    object  
 1   type        9224 non-null    category
 2   title       9224 non-null    object  
 3   director    9224 non-null    object  
 4   cast        9224 non-null    object  
 5   country     9224 non-null    object  
 6   date_added  9224 non-null    datetime64[ns]
 7   release_year 9224 non-null    int64   
 8   rating      9224 non-null    category
 9   duration    9224 non-null    object  
 10  listed_in   9224 non-null    object  
 11  description 9224 non-null    object  
 12  year_added  9224 non-null    int64   
 13  month_added 9224 non-null    int64   
 14  day_added   9224 non-null    int64   
dtypes: category(2), datetime64[ns](1), int64(4), object(8)
memory usage: 1.3+ MB
```

```
In [ ]:
```

```
In [1893]: tvSDataCountryUk = tvSDataCountry[tvSDataCountry['country'] == 'United Kingdom']
tvSDataCountryJapan = tvSDataCountry[tvSDataCountry['country'] == 'Japan']
```

In [1894]: tvSDataCountryJapan

Out[1894]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
76	s77	TV Show	Yowamushi Pedal	NaN	Daiki Yamashita	Japan	2021-09-14	2013	TV-14	1	Anime Series, International TV Shows	A timid, anime-loving teen gets drawn into a s...	2021	9
76	s77	TV Show	Yowamushi Pedal	NaN	Kohsuke Toriumi	Japan	2021-09-14	2013	TV-14	1	Anime Series, International TV Shows	A timid, anime-loving teen gets drawn into a s...	2021	9
76	s77	TV Show	Yowamushi Pedal	NaN	Jun Fukushima	Japan	2021-09-14	2013	TV-14	1	Anime Series, International TV Shows	A timid, anime-loving teen gets drawn into a s...	2021	9
76	s77	TV Show	Yowamushi Pedal	NaN	Hiroki Yasumoto	Japan	2021-09-14	2013	TV-14	1	Anime Series, International TV Shows	A timid, anime-loving teen gets drawn into a s...	2021	9
76	s77	TV Show	Yowamushi Pedal	NaN	Showtaro Morikubo	Japan	2021-09-14	2013	TV-14	1	Anime Series, International TV Shows	A timid, anime-loving teen gets drawn into a s...	2021	9
...
8795	s8796	TV Show	Yu-Gi-Oh! Arc-V	NaN	Mike Liscio	Japan	2018-05-01	2015	TV-Y7	2	Anime Series, Kids' TV	Now that he's discovered the Pendulum Summonin...	2018	5
8795	s8796	TV Show	Yu-Gi-Oh! Arc-V	NaN	Emily Bauer	Japan	2018-05-01	2015	TV-Y7	2	Anime Series, Kids' TV	Now that he's discovered the Pendulum Summonin...	2018	5

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	
8795	s8796	TV Show	Yu-Gi-Oh! Arc-V	NaN	Billy Bob Thompson	Japan	2018-05-01	2015	TV-Y7	2	Anime Series, Kids' TV	Now that he's discovered the Pendulum Summonin...	2018	5
8795	s8796	TV Show	Yu-Gi-Oh! Arc-V	NaN	Alyson Leigh Rosenfeld	Japan	2018-05-01	2015	TV-Y7	2	Anime Series, Kids' TV	Now that he's discovered the Pendulum Summonin...	2018	5
8795	s8796	TV Show	Yu-Gi-Oh! Arc-V	NaN	Michael Crouch	Japan	2018-05-01	2015	TV-Y7	2	Anime Series, Kids' TV	Now that he's discovered the Pendulum Summonin...	2018	5

2350 rows × 15 columns



```
In [1895]: tvSDDataCountryUk
```

```
Out[1895]:
```

		show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc	United Kingdom	2021-09-24	2021	TV-14	9	British TV Shows, Reality TV	A talented batch of amateur bakers face off in...	2021		
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Sue Perkins	United Kingdom	2021-09-24	2021	TV-14	9	British TV Shows, Reality TV	A talented batch of amateur bakers face off in...	2021		
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mary Berry	United Kingdom	2021-09-24	2021	TV-14	9	British TV Shows, Reality TV	A talented batch of amateur bakers face off in...	2021		

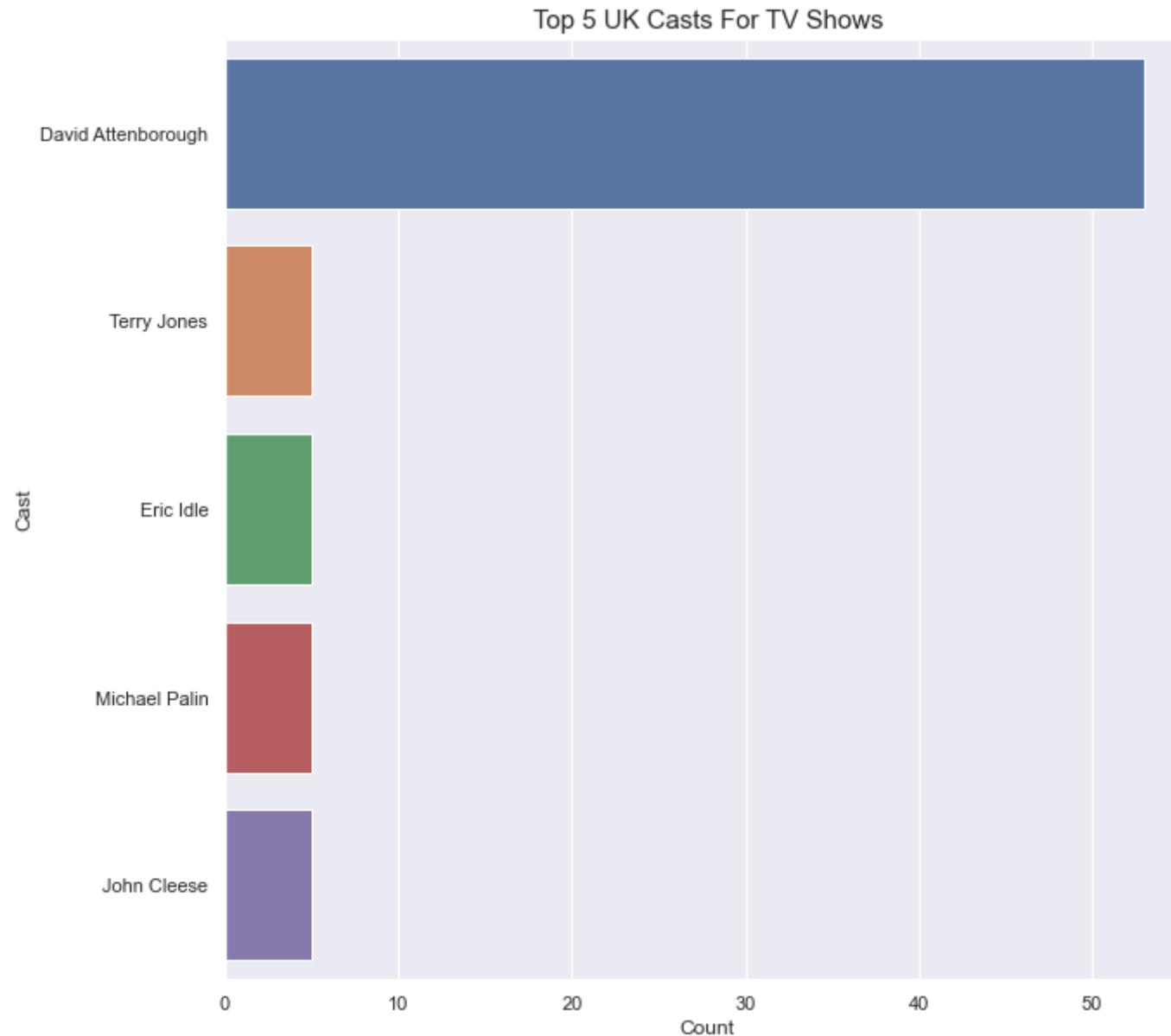
```
In [1896]: tvSDDataCountryUk['cast'].fillna("David Attenborough", inplace = True)  
tvSDDataCountryUk['director'].fillna("Noam Murro", inplace = True)
```

C:\Program Files\Anaconda3\lib\site-packages\pandas\core\generic.py:6392: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

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)
return self._update_inplace(result)

```
In [1897]: tvSDDataCountryJapan['cast'].fillna("Takahiro Sakurai", inplace = True)  
tvSDDataCountryJapan['director'].fillna("Yasuhiro Irie", inplace = True)
```

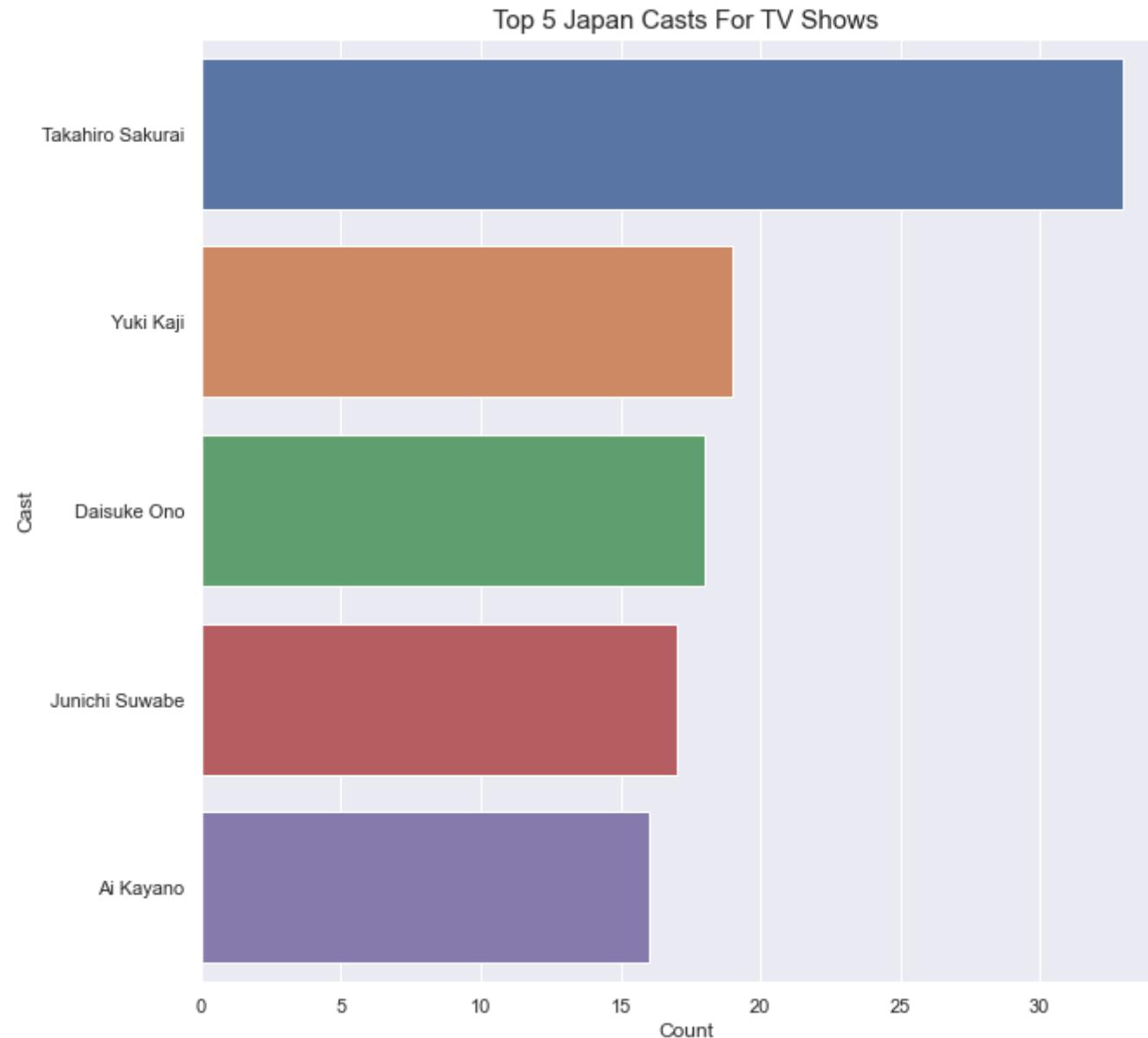
```
In [1898]: plt.figure(figsize=(10,10))
sns.countplot(y = tvSDataCountryUk['cast'] , order = tvSDataCountryUk['cast'].value_counts().index[0:5])
plt.xlabel("Count" ,fontsize = 12)
plt.ylabel("Cast" ,fontsize = 12)
plt.title("Top 5 UK Casts For TV Shows" , fontsize = 15)
plt.show()
```



Top 5 UK Casts for TV Shows:

1. David Attenborough
2. Terry Jones
3. Eric Idle
4. Michael Palin
5. John Cleese

```
In [1899]: plt.figure(figsize=(10,10))
sns.countplot(y = tvSDataCountryJapan['cast'] , order = tvSDataCountryJapan['cast'].value_counts().index[0:5])
plt.xlabel("Count" ,fontsize = 12)
plt.ylabel("Cast" ,fontsize = 12)
plt.title("Top 5 Japan Casts For TV Shows" , fontsize = 15)
plt.show()
```



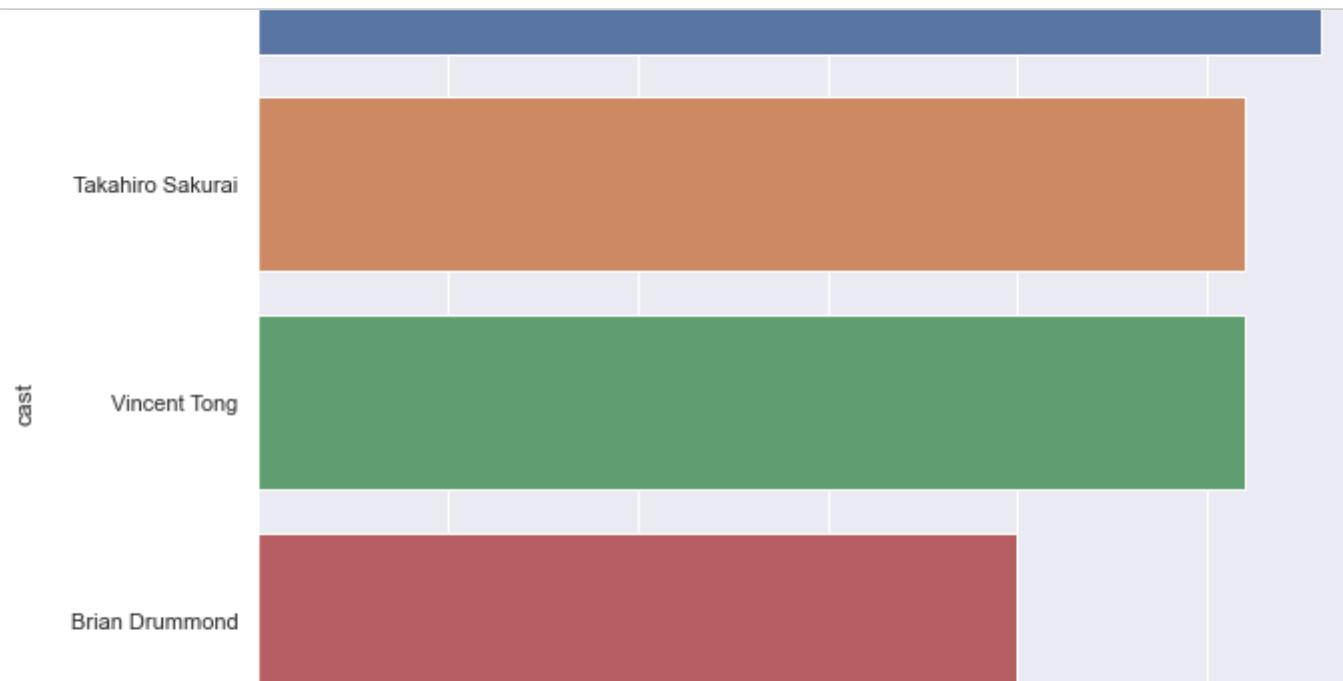
Top 5 UK Casts for TV Shows:

1. Takahiro Sakurai
2. Yuki Kaji
3. Daisuke Ono
4. Junichi Suwabe
5. Ai Kayno

In []:

In [1900]: # Without considering NA values

```
plt.figure(figsize=(10,10))
sns.countplot(y = tvSDataCountry['cast'] , order = tvSDataCountry['cast'].value_counts().index[0:5])
plt.title("Top 5 Casts on netflix for TV Shows", fontsize = 15)
plt.show()
```

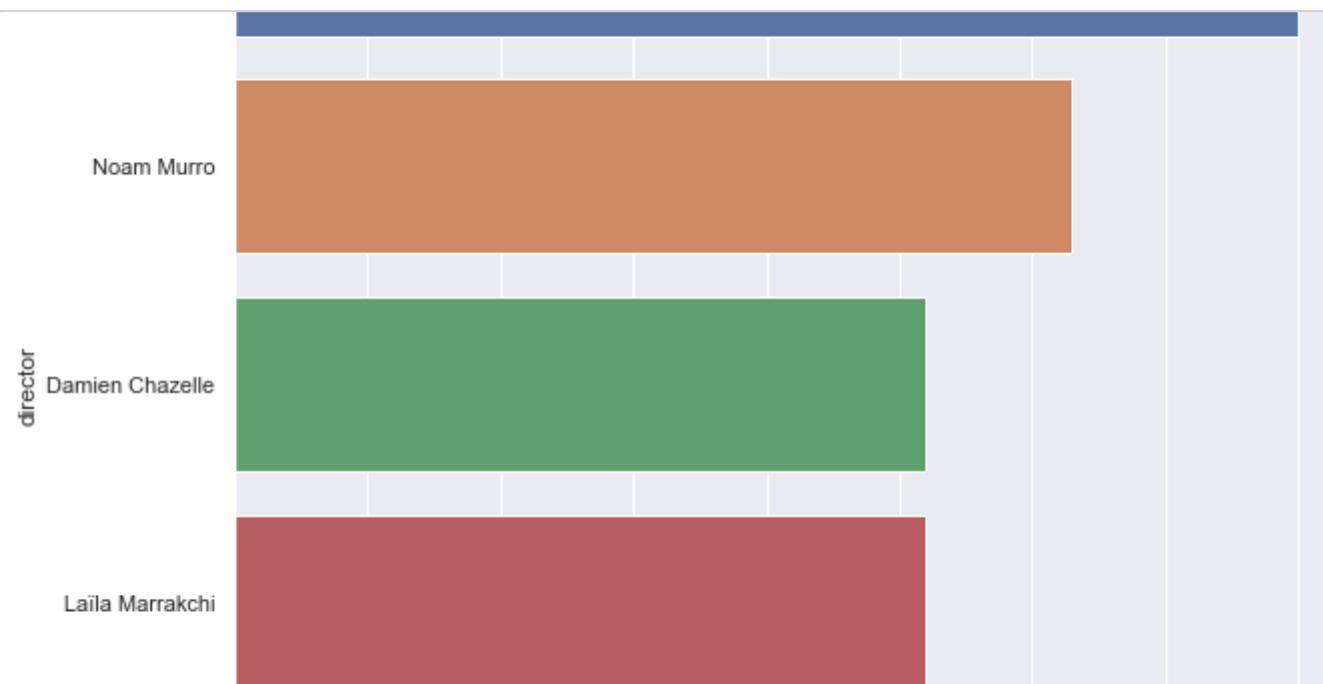


Top 5 Casts for TV Shows:

1. David Attenborough
2. Takahiro Sakurai
3. Vincent Tong
4. Brian Drummond
5. Yuki Kaji

In [1901]: # Without considering NA values

```
plt.figure(figsize=(10,10))
sns.countplot(y = tvSDataCountry['director'] , order = tvSDataCountry['director'].value_counts().index[0:5])
plt.title("Top 5 Directors on netflix for TV shows", fontsize = 15)
plt.show()
```



Top 5 Casts for TV Shows:

1. Thomas Astruc
2. Noam Murro
3. Damien Chazzelle
4. Laila Marrakchi
5. Alan Paul

In [1902]: tvShowListedIn

Out[1902]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows	After crossing paths at a party, a Cape Town t...	2021	9
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	TV Dramas	After crossing paths at a party, a Cape Town t...	2021	9
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	9
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1 Season	Crime TV Shows	To protect his family from a powerful drug lor...	2021	9
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1 Season	International TV Shows	To protect his family from a powerful drug lor...	2021	9
...
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen	Pakistan	2016-12-15	2012	TV-PG	1 Season	Romantic TV Shows	Strong-willed, middle-class Kashaf and carefre...	2016	12

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ...	Pakistan	2016-12-15	2012	TV-PG	1 Season	TV Dramas	Strong-willed, middle-class Kashaf and carefre...	2016	12
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	Kids' TV	While living alone in a spooky town, a young g...	2019	7
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	Korean TV Shows	While living alone in a spooky town, a young g...	2019	7
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	TV Comedies	While living alone in a spooky town, a young g...	2019	7

6133 rows × 15 columns

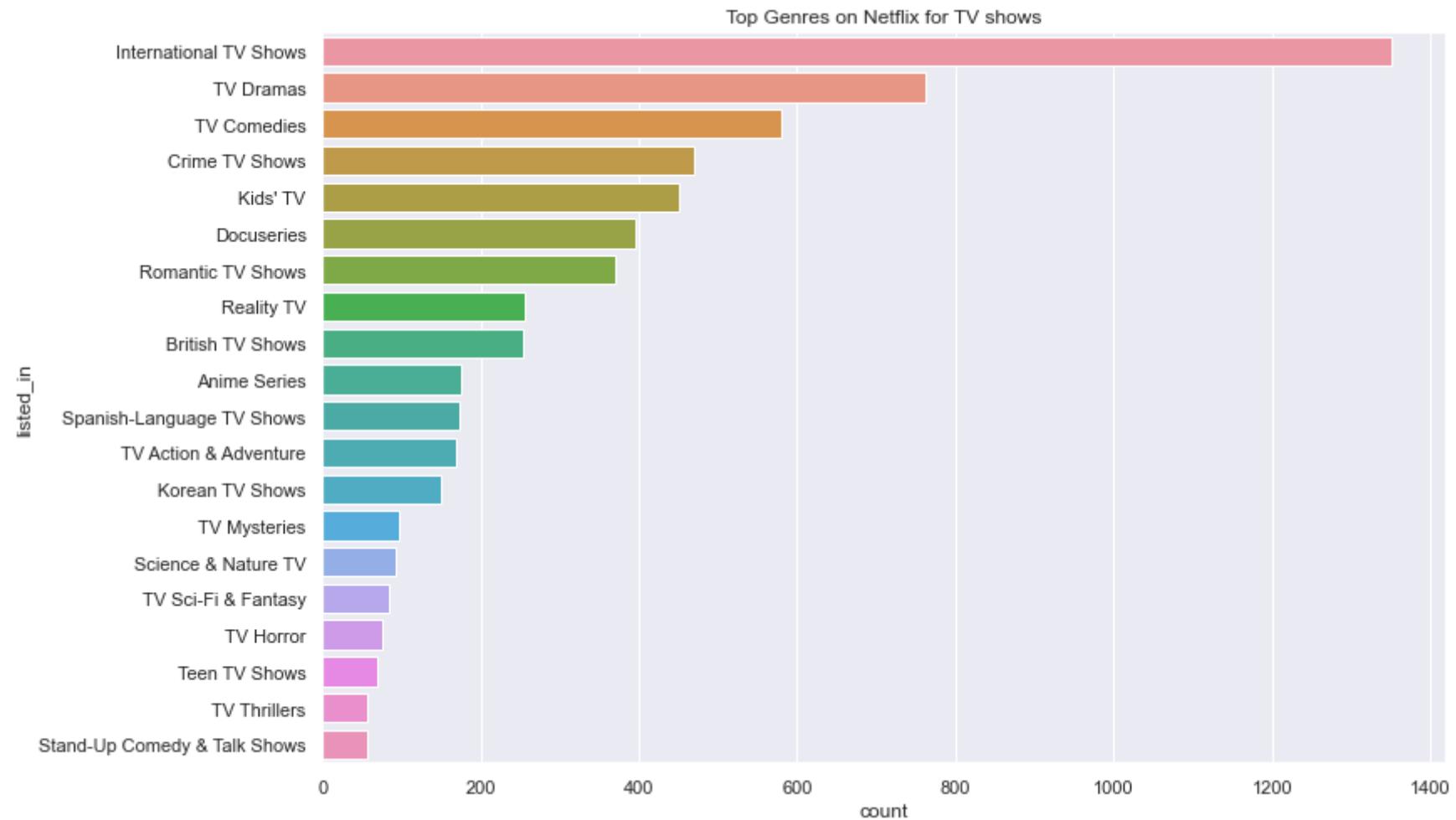


```
In [1903]: dff = pd.DataFrame(tvShowListedIn['listed_in'].value_counts())
dff.reset_index(inplace=True)
dff.columns = ['listed_in', 'count']
dff
dff['percentage'] = dff['count'] * 100 / dff['count'].sum()
dff
```

Out[1903]:

	listed_in	count	percentage
0	International TV Shows	1351	22.028371
1	TV Dramas	763	12.440894
2	TV Comedies	581	9.473341
3	Crime TV Shows	470	7.663460
4	Kids' TV	451	7.353661
5	Docuseries	395	6.440567
6	Romantic TV Shows	370	6.032937
7	Reality TV	255	4.157835
8	British TV Shows	253	4.125224
9	Anime Series	176	2.869721
10	Spanish-Language TV Shows	174	2.837111
11	TV Action & Adventure	168	2.739279
12	Korean TV Shows	151	2.462090
13	TV Mysteries	98	1.597913
14	Science & Nature TV	92	1.500082
15	TV Sci-Fi & Fantasy	84	1.369640
16	TV Horror	75	1.222893
17	Teen TV Shows	69	1.125061
18	TV Thrillers	57	0.929398
19	Stand-Up Comedy & Talk Shows	56	0.913093
20	Classic & Cult TV	28	0.456547
21	TV Shows	16	0.260884

```
In [1904]: plt.figure(figsize = (12,8))
sns.countplot(y=tvShowListedIn['listed_in'],order = tvShowListedIn['listed_in'].value_counts().index[0:20])
plt.title('Top Genres on Netflix for TV shows', fontsize = 12)
plt.show()
```



Popular Genre for TV shows:

1. International TV shows 21 %
2. TV Dramas 12 %
3. TV Comedies 9%
4. Crime TV Show 8%
5. Kid's Tv 7%

```
In [1905]: tvShowListedIn = tvShowListedIn.assign(country= tvSData['country'].str.split(', ')).explode('country')
tvShowListedIn
```

Out[1905]:

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_adde
1	s2	TV Show	Blood & Water	Nan Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	International TV Shows	After crossing paths at a party, a Cape Town t...	2021	
1	s2	TV Show	Blood & Water	Nan Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	TV Dramas	After crossing paths at a party, a Cape Town t...	2021	
1	s2	TV Show	Blood & Water	Nan Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	TV Mysteries	After crossing paths at a party, a Cape Town t...	2021	
2	s3	TV Show	Ganglands	Julien Leclercq Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1 Season	Crime TV Shows	To protect his family from a powerful drug lor...	2021	
2	s3	TV Show	Ganglands	Julien Leclercq Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	United States	2021-09-24	2021	TV-MA	1 Season	International TV Shows	To protect his family from a powerful drug lor...	2021	

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ...	Pakistan	2016-12-15	2012	TV-PG	1 Season	Romantic TV Shows	Strong-willed, middle-class Kashaf and carefree...	2016	1
8800	s8801	TV Show	Zindagi Gulzar Hai	NaN	Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ...	Pakistan	2016-12-15	2012	TV-PG	1 Season	TV Dramas	Strong-willed, middle-class Kashaf and carefree...	2016	1
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	Kids' TV	While living alone in a spooky town, a young g...	2019	
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	Korean TV Shows	While living alone in a spooky town, a young g...	2019	
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	United States	2019-07-01	2018	TV-Y7	2 Seasons	TV Comedies	While living alone in a spooky town, a young g...	2019	

6881 rows × 15 columns

In [1906]:

```
top10TVSCountries = dict(tvShowListedIn['country'].value_counts(ascending = False)[:10])
```

In [1907]:

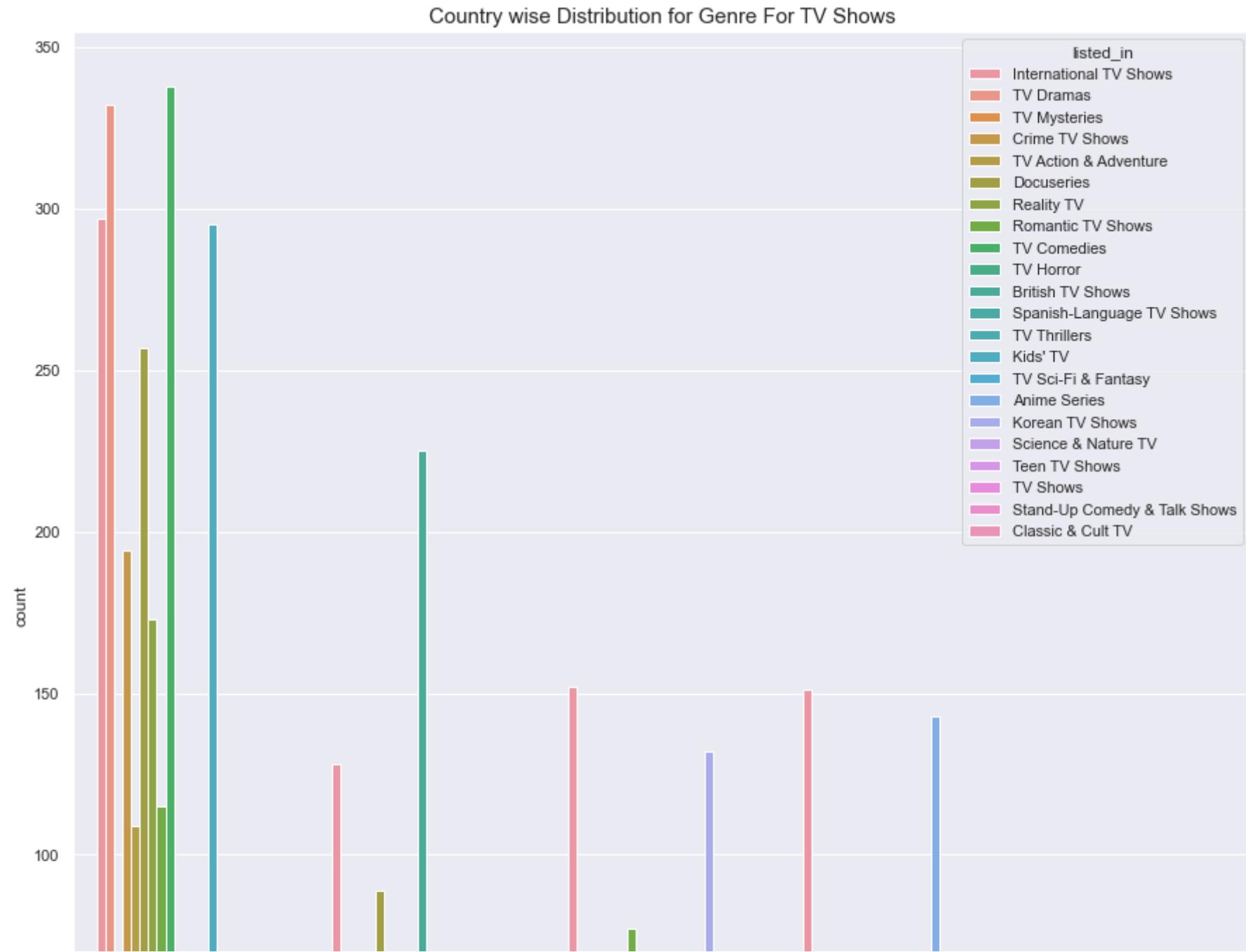
```
tvShowListedIn[tvShowListedIn['country'].isna()]
```

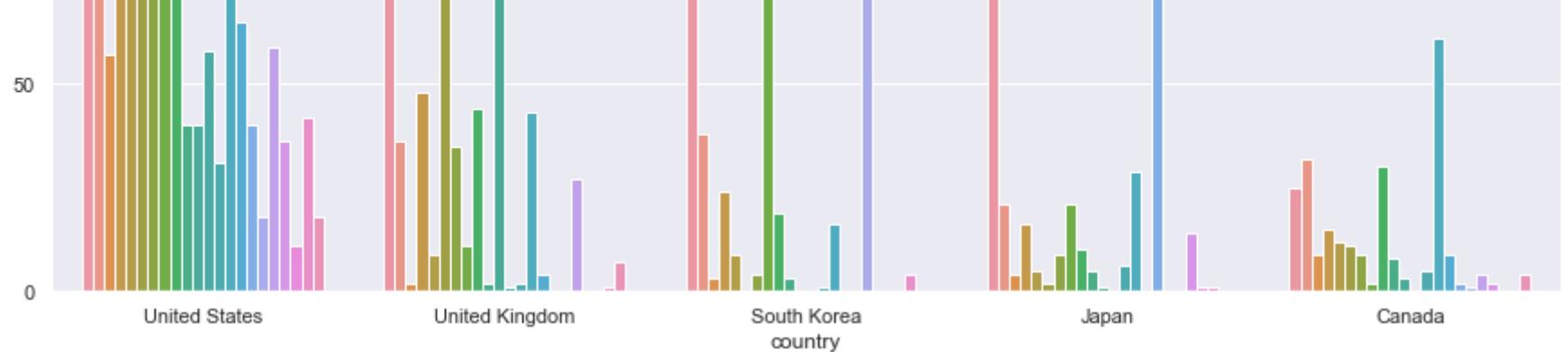
Out[1907]:

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added	month_added	day_added
---------	------	-------	----------	------	---------	------------	--------------	--------	----------	-----------	-------------	------------	-------------	-----------

```
In [1908]: Top10Countries = list(top10MovieCountries.keys())
```

```
In [1909]: plt.figure(figsize=(15,15))
sns.countplot(x = 'country', data = tvShowListedIn, order = tvShowListedIn['country'].value_counts(ascending= False).index[0:5], h
plt.title("Country wise Distribution for Genre For TV Shows", fontsize = 15)
plt.show()
```





Country wise Distribution for Genre:

1. In United states TV Comedies and TV dramas are very popular
2. In UK British TV Shows are popular
3. In South Korea and Japan International TV shows are most watched

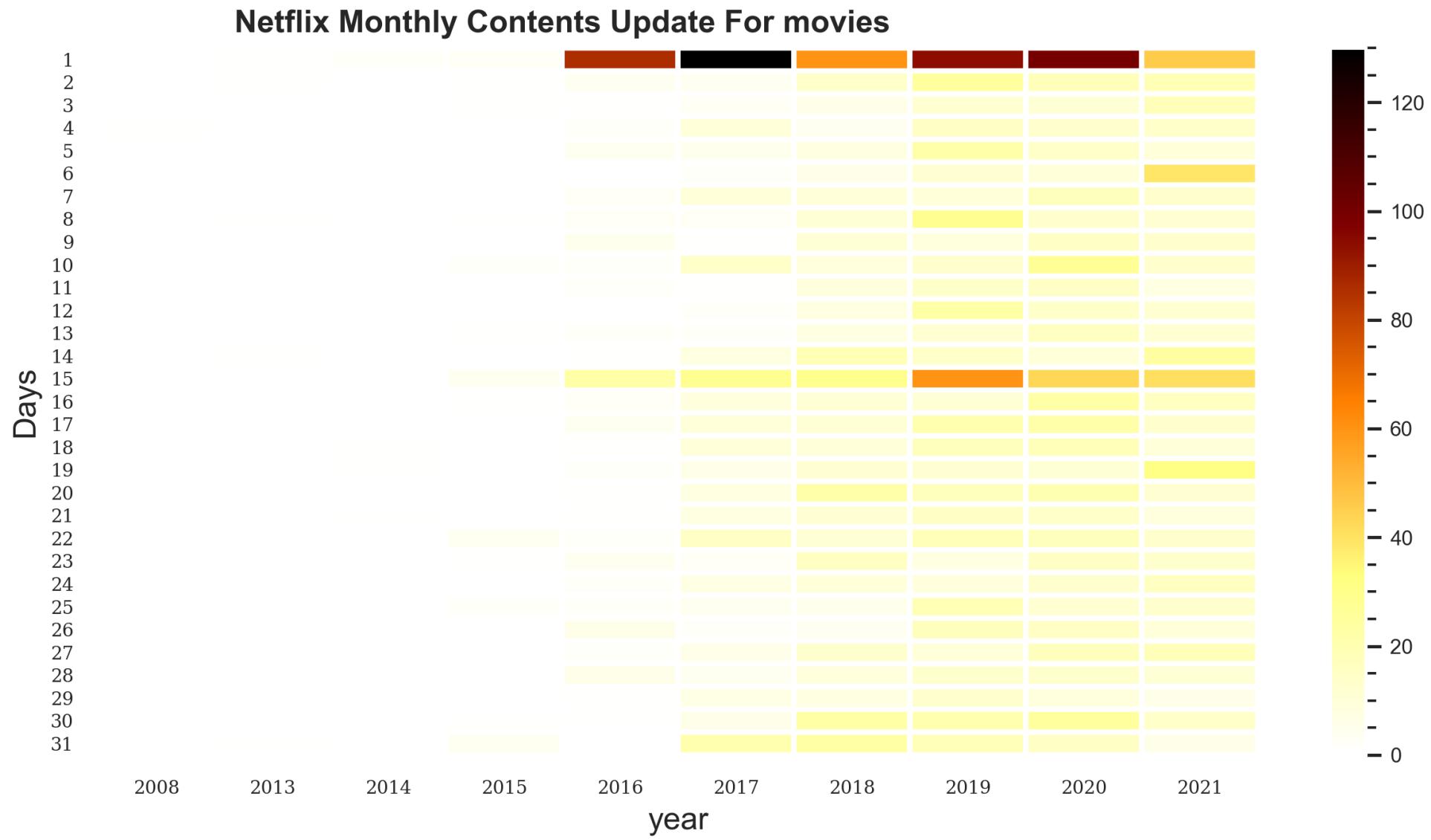
In []:

```
In [1910]: month_order = [1,2,3,4,5,6,7,8,9,10,11,12][::-1]
month_HM = tvSData.groupby('year_added')['month_added'].value_counts().unstack().fillna(0)[month_order].T
plt.figure(figsize=(10, 5), dpi=300)
plt.pcolor(month_HM, cmap='afmhot_r', edgecolors='white', linewidths=2) # heatmap
plt.xlabel("year")
plt.ylabel("Month")
plt.xticks(np.arange(0.5, len(month_HM.columns), 1), month_HM.columns, fontsize=7, fontfamily='serif')
plt.yticks(np.arange(0.5, len(month_HM.index), 1), month_HM.index, fontsize=7, fontfamily='serif')
plt.title('Netflix Monthly Contents Update For movies', fontsize=12, fontweight='bold', position=(0.40, 1.0+0.02))
cbar = plt.colorbar()
cbar.ax.tick_params(labelsize=8)
cbar.ax.minorticks_on()
plt.show()
```

```
In [ ]:
```

In 2021, we can see maximum content gets added after may month for TV shows

```
In [1911]: month_order = [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31][::-1]
month_HM = tvSData.groupby('year_added')[ 'day_added'].value_counts().unstack().fillna(0)[month_order].T
plt.figure(figsize=(10, 5), dpi=300)
plt.pcolor(month_HM, cmap='afmhot_r', edgecolors='white', linewidths=2) # heatmap
plt.xlabel("year")
plt.ylabel("Days")
plt.xticks(np.arange(0.5, len(month_HM.columns), 1), month_HM.columns, fontsize=7, fontfamily='serif')
plt.yticks(np.arange(0.5, len(month_HM.index), 1), month_HM.index, fontsize=7, fontfamily='serif')
plt.title('Netflix Monthly Contents Update For movies', fontsize=12, fontweight='bold', position=(0.40, 1.0+0.02))
cbar = plt.colorbar()
cbar.ax.tick_params(labelsize=8)
cbar.ax.minorticks_on()
plt.show()
```



We can clearly observe here most of the content gets added on 1st of the month and 15th at some extent

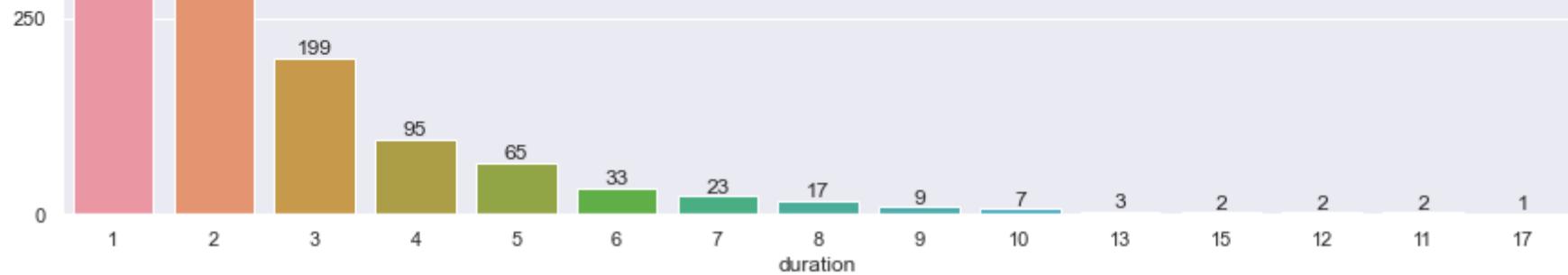
```
In [1912]: # TV shows number od seasons analysis
```

In []:

In []:

```
In [1913]: plt.figure(figsize=(15,15))
graph = sns.countplot(x = 'duration', data = tvSData, order = tvSData['duration'].value_counts(ascending = False).index)
for i in graph.containers:
    graph.bar_label(i)
plt.title("Seasons Analysis For TV Shows", fontsize = 15)
plt.show()
```





Most of the TV show consist of 1 season only.

In []:

Problem Statement

1. Understanding what content is available on netflix
2. Based on observations analysing which shows to produce and how to grow business
3. Understanding how number of movies release in the recent years changed
4. Comparison of TV shows and Movies
5. Analysis of casts and Directors
6. Content analysis in different countries

Shape of Dataset = (8807, 12) Except year all the columns were of object type --> Conversion made on date_added column to convert it into dateTime type and rating consists of 2 and 14 unique values respectively which can work as category s we can convert it to type category. Found out missing value percentage for each column to analyze it.

Buisness Insight: After 2018 Growth of adding content on netflix is increased at massive level. TV-MA is the most watched content present on Netflix USA, India, United Kingdom are the top Countries for content International Movie shows, TV dramas, comedy shows are top genre for TV shows while International movies, Dramas, comedies, Documentries are popular in MOvies section But in 2021 we can see growth of adding content on netflix has been decreased Netflix started focusing on TV shows as compared to Movies

