

data-exploration-and-visualisation

June 4, 2023

```
[477]: import numpy as np, pandas as pd
```

```
[478]: import matplotlib.pyplot as plt, seaborn as sns
```

```
[479]: df = pd.read_csv('Netflix.csv')
```

```
[480]: df
```

```
[480]:
```

	show_id	type	title	director \
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson
1	s2	TV Show	Blood & Water	NaN
2	s3	TV Show	Ganglands	Julien Leclercq
3	s4	TV Show	Jailbirds New Orleans	NaN
4	s5	TV Show	Kota Factory	NaN
...
8804	s8803	Movie	Zodiac	David Fincher
8805	s8804	TV Show	Zombie Dumb	NaN
8806	s8805	Movie	Zombieland	Ruben Fleischer
8807	s8806	Movie	Zoom	Peter Hewitt
8808	s8807	Movie	Zubaan	Mozes Singh

	cast	country \
0	NaN	United States
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
2	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN
3	NaN	NaN
4	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India
...
8804	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States
8805	NaN	NaN
8806	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States
8807	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India

	date_added	release_year	rating	duration \
0	September 25, 2021	2020	PG-13	90 min
1	September 24, 2021	2021	TV-MA	2 Seasons

2	September 24, 2021	2021	TV-MA	1 Season
3	September 24, 2021	2021	TV-MA	1 Season
4	September 24, 2021	2021	TV-MA	2 Seasons
...
8804	November 20, 2019	2007	R	158 min
8805	July 1, 2019	2018	TV-Y7	2 Seasons
8806	November 1, 2019	2009	R	88 min
8807	January 11, 2020	2006	PG	88 min
8808	March 2, 2019	2015	TV-14	111 min

	listed_in \
0	Documentaries
1	International TV Shows, TV Dramas, TV Mysteries
2	Crime TV Shows, International TV Shows, TV Act...
3	Docuseries, Reality TV
4	International TV Shows, Romantic TV Shows, TV ...
...	...
8804	Cult Movies, Dramas, Thrillers
8805	Kids' TV, Korean TV Shows, TV Comedies
8806	Comedies, Horror Movies
8807	Children & Family Movies, Comedies
8808	Dramas, International Movies, Music & Musicals

	description
0	As her father nears the end of his life, filmm...
1	After crossing paths at a party, a Cape Town t...
2	To protect his family from a powerful drug lor...
3	Feuds, flirtations and toilet talk go down amo...
4	In a city of coaching centers known to train I...
...	...
8804	A political cartoonist, a crime reporter and a...
8805	While living alone in a spooky town, a young g...
8806	Looking to survive in a world taken over by zo...
8807	Dragged from civilian life, a former superhero...
8808	A scrappy but poor boy worms his way into a ty...

[8809 rows x 12 columns]

```
[481]: df.shape
```

```
[481]: (8809, 12)
```

```
[482]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8809 entries, 0 to 8808
Data columns (total 12 columns):
```

#	Column	Non-Null Count	Dtype
0	show_id	8809 non-null	object
1	type	8808 non-null	object
2	title	8807 non-null	object
3	director	6173 non-null	object
4	cast	7983 non-null	object
5	country	7976 non-null	object
6	date_added	8797 non-null	object
7	release_year	8807 non-null	object
8	rating	8803 non-null	object
9	duration	8804 non-null	object
10	listed_in	8806 non-null	object
11	description	8806 non-null	object

dtypes: object(12)
memory usage: 826.0+ KB

```
[483]: df.isna().sum(axis=0)
```

```
[483]: show_id      0
       type        1
       title       2
       director    2636
       cast        826
       country     833
       date_added   12
       release_year  2
       rating       6
       duration     5
       listed_in    3
       description  3
       dtype: int64
```

```
[484]: df['director'].value_counts()
```

```
[484]: Rajiv Chilaka      19
       Raúl Campos, Jan Suter  18
       Marcus Raboy      16
       Suhas Kadav       16
       Jay Karas         14
       ..
       Raymie Muzquiz, Stu Livingston  1
       Joe Menendez      1
       Eric Bross        1
       Will Eisenberg    1
       Moez Singh        1
       Name: director, Length: 4528, dtype: int64
```

```
[485]: df['country'].value_counts()
```

```
[485]: United States      2817
      India              972
      United Kingdom    419
      Japan             245
      South Korea       199
      ...
      Romania, Bulgaria, Hungary  1
      Uruguay, Guatemala          1
      France, Senegal, Belgium    1
      Mexico, United States, Spain, Colombia  1
      United Arab Emirates, Jordan  1
      Name: country, Length: 749, dtype: int64
```

```
[486]: df[['type', 'title', 'cast']] = df[['type', 'title', 'cast']].fillna('NA')
```

```
[487]: df
```

```
[487]:   show_id  type      title  director \
0       s1  Movie  Dick Johnson Is Dead  Kirsten Johnson
1       s2  TV Show      Blood & Water          NaN
2       s3  TV Show      Ganglands  Julien Leclercq
3       s4  TV Show  Jailbirds New Orleans          NaN
4       s5  TV Show      Kota Factory          NaN
...     ...   ...
8804   s8803  Movie      Zodiac  David Fincher
8805   s8804  TV Show  Zombie Dumb          NaN
8806   s8805  Movie  Zombieland  Ruben Fleischer
8807   s8806  Movie      Zoom  Peter Hewitt
8808   s8807  Movie      Zubaan  Mozez Singh

      cast      country \
0      NA  United States
1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...  South Africa
2  Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...  NaN
3      NA      NaN
4  Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...  India
...     ...   ...
8804  Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...  United States
8805      NA      NaN
8806  Jesse Eisenberg, Woody Harrelson, Emma Stone, ...  United States
8807  Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...  United States
8808  Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...  India

      date_added  release_year  rating  duration \
0  September 25, 2021      2020  PG-13      90 min
```

1	September 24, 2021	2021	TV-MA	2 Seasons
2	September 24, 2021	2021	TV-MA	1 Season
3	September 24, 2021	2021	TV-MA	1 Season
4	September 24, 2021	2021	TV-MA	2 Seasons
...
8804	November 20, 2019	2007	R	158 min
8805	July 1, 2019	2018	TV-Y7	2 Seasons
8806	November 1, 2019	2009	R	88 min
8807	January 11, 2020	2006	PG	88 min
8808	March 2, 2019	2015	TV-14	111 min

	listed_in \
0	Documentaries
1	International TV Shows, TV Dramas, TV Mysteries
2	Crime TV Shows, International TV Shows, TV Act...
3	Docuseries, Reality TV
4	International TV Shows, Romantic TV Shows, TV ...
...	...
8804	Cult Movies, Dramas, Thrillers
8805	Kids' TV, Korean TV Shows, TV Comedies
8806	Comedies, Horror Movies
8807	Children & Family Movies, Comedies
8808	Dramas, International Movies, Music & Musicals

	description
0	As her father nears the end of his life, filmm...
1	After crossing paths at a party, a Cape Town t...
2	To protect his family from a powerful drug lor...
3	Feuds, flirtations and toilet talk go down amo...
4	In a city of coaching centers known to train I...
...	...
8804	A political cartoonist, a crime reporter and a...
8805	While living alone in a spooky town, a young g...
8806	Looking to survive in a world taken over by zo...
8807	Dragged from civilian life, a former superhero...
8808	A scrappy but poor boy worms his way into a ty...

[8809 rows x 12 columns]

[488]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8809 entries, 0 to 8808
Data columns (total 12 columns):
#   Column      Non-Null Count  Dtype
---  -
0   show_id     8809 non-null   object
```

```

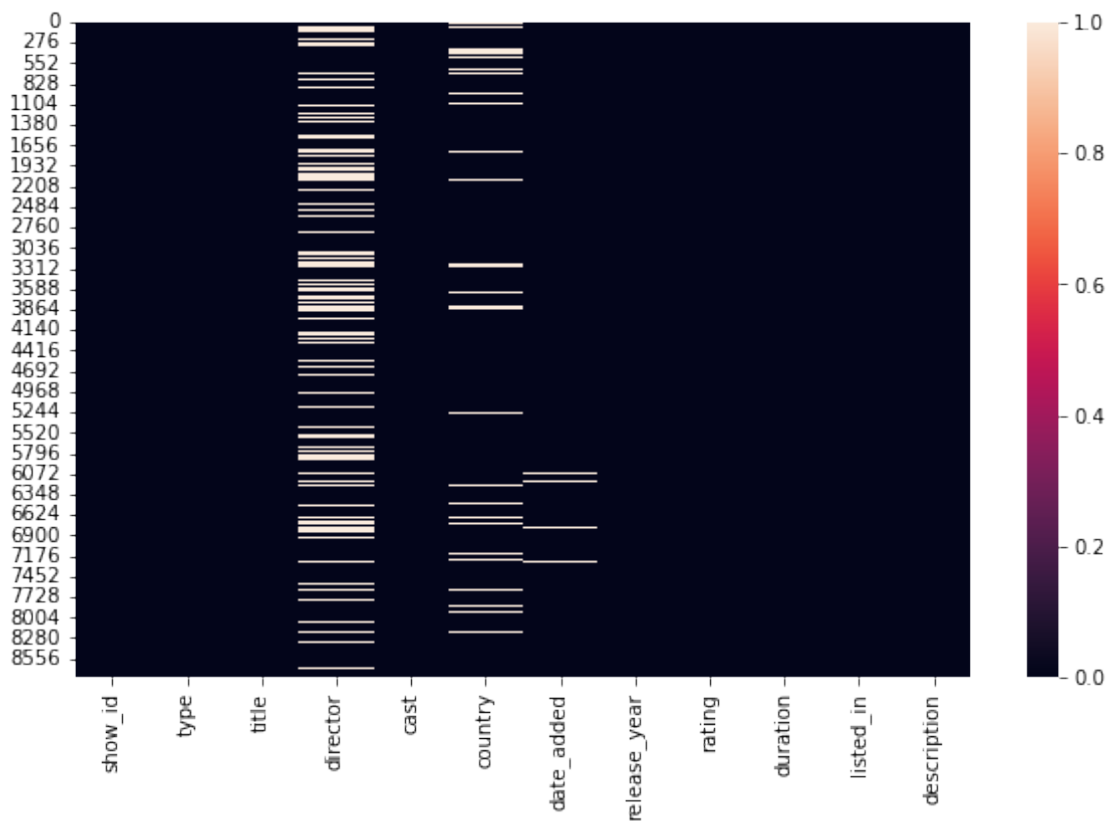
1  type          8809 non-null  object
2  title         8809 non-null  object
3  director      6173 non-null  object
4  cast          8809 non-null  object
5  country       7976 non-null  object
6  date_added    8797 non-null  object
7  release_year  8807 non-null  object
8  rating        8803 non-null  object
9  duration      8804 non-null  object
10 listed_in     8806 non-null  object
11 description   8806 non-null  object
dtypes: object(12)
memory usage: 826.0+ KB

```

```

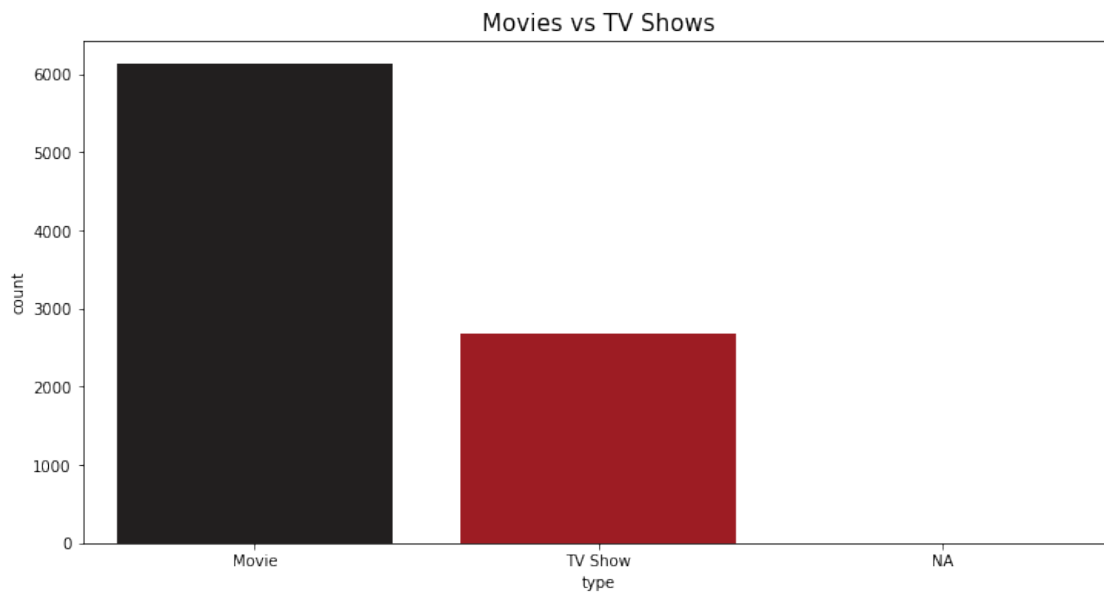
[489]: plt.figure(figsize = (10,6))
sns.heatmap(df.isnull())
plt.show()

```



*Above plot is the heatmap showing the null spaces in all the columns of Netflix dataframe.

```
[490]: df = df[df['type'].str.contains('William')==False]
plt.figure(figsize = (12,6))
sns.countplot(data = df, x = 'type',palette = my_color)
plt.title('Movies vs TV Shows',fontsize = 15)
plt.show()
```



```
[491]: df.reset_index()
```

```
[491]:
```

	index	show_id	type	title	director	\
0	0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	
1	1	s2	TV Show	Blood & Water	NaN	
2	2	s3	TV Show	Ganglands	Julien Leclercq	
3	3	s4	TV Show	Jailbirds New Orleans	NaN	
4	4	s5	TV Show	Kota Factory	NaN	
...	
8803	8804	s8803	Movie	Zodiac	David Fincher	
8804	8805	s8804	TV Show	Zombie Dumb	NaN	
8805	8806	s8805	Movie	Zombieland	Ruben Fleischer	
8806	8807	s8806	Movie	Zoom	Peter Hewitt	
8807	8808	s8807	Movie	Zubaan	Mozez Singh	

	cast	country	\
0	NA	United States	
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	
2	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	
3	NA	NaN	
4	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	

...
8803	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States
8804		NA NaN
8805	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States
8806	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States
8807	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India

	date_added	release_year	rating	duration	\
0	September 25, 2021	2020	PG-13	90 min	
1	September 24, 2021	2021	TV-MA	2 Seasons	
2	September 24, 2021	2021	TV-MA	1 Season	
3	September 24, 2021	2021	TV-MA	1 Season	
4	September 24, 2021	2021	TV-MA	2 Seasons	

...
8803	November 20, 2019	2007	R	158 min
8804	July 1, 2019	2018	TV-Y7	2 Seasons
8805	November 1, 2019	2009	R	88 min
8806	January 11, 2020	2006	PG	88 min
8807	March 2, 2019	2015	TV-14	111 min

	listed_in	\
0	Documentaries	
1	International TV Shows, TV Dramas, TV Mysteries	
2	Crime TV Shows, International TV Shows, TV Act...	
3	Docuseries, Reality TV	
4	International TV Shows, Romantic TV Shows, TV ...	
...	...	
8803	Cult Movies, Dramas, Thrillers	
8804	Kids' TV, Korean TV Shows, TV Comedies	
8805	Comedies, Horror Movies	
8806	Children & Family Movies, Comedies	
8807	Dramas, International Movies, Music & Musicals	

	description
0	As her father nears the end of his life, filmm...
1	After crossing paths at a party, a Cape Town t...
2	To protect his family from a powerful drug lor...
3	Feuds, flirtations and toilet talk go down amo...
4	In a city of coaching centers known to train I...
...	...
8803	A political cartoonist, a crime reporter and a...
8804	While living alone in a spooky town, a young g...
8805	Looking to survive in a world taken over by zo...
8806	Dragged from civilian life, a former superhero...
8807	A scrappy but poor boy worms his way into a ty...

[8808 rows x 13 columns]


```
[492]: df['director'] = df['director'].str.split(',')
```

C:\Users\MYPC~1\AppData\Local\Temp\ipykernel_22080\3162038592.py:1:

SettingWithCopyWarning:

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

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

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

```
df['director'] = df['director'].str.split(',')
```

```
[493]: df = df.explode('director')
```

```
[494]: df
```

```
[494]:
```

	show_id	type	title	director \
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson
1	s2	TV Show	Blood & Water	NaN
2	s3	TV Show	Ganglands	Julien Leclercq
3	s4	TV Show	Jailbirds New Orleans	NaN
4	s5	TV Show	Kota Factory	NaN
...
8804	s8803	Movie	Zodiac	David Fincher
8805	s8804	TV Show	Zombie Dumb	NaN
8806	s8805	Movie	Zombieland	Ruben Fleischer
8807	s8806	Movie	Zoom	Peter Hewitt
8808	s8807	Movie	Zubaan	Mozez Singh

	cast	country \
0	NA	United States
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
2	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN
3	NA	NaN
4	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India
...
8804	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States
8805	NA	NaN
8806	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States
8807	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India

	date_added	release_year	rating	duration \
0	September 25, 2021	2020	PG-13	90 min
1	September 24, 2021	2021	TV-MA	2 Seasons
2	September 24, 2021	2021	TV-MA	1 Season
3	September 24, 2021	2021	TV-MA	1 Season
4	September 24, 2021	2021	TV-MA	2 Seasons

```

...
8804 November 20, 2019      2007      R      158 min
8805      July 1, 2019      2018  TV-Y7  2 Seasons
8806 November 1, 2019      2009      R      88 min
8807 January 11, 2020      2006      PG      88 min
8808      March 2, 2019      2015  TV-14     111 min

```

```

                                listed_in \
0                                Documentaries
1      International TV Shows, TV Dramas, TV Mysteries
2      Crime TV Shows, International TV Shows, TV Act...
3                                Docuseries, Reality TV
4      International TV Shows, Romantic TV Shows, TV ...
...
8804                                Cult Movies, Dramas, Thrillers
8805      Kids' TV, Korean TV Shows, TV Comedies
8806                                Comedies, Horror Movies
8807      Children & Family Movies, Comedies
8808      Dramas, International Movies, Music & Musicals

```

```

                                description
0      As her father nears the end of his life, filmm...
1      After crossing paths at a party, a Cape Town t...
2      To protect his family from a powerful drug lor...
3      Feuds, flirtations and toilet talk go down amo...
4      In a city of coaching centers known to train I...
...
8804 A political cartoonist, a crime reporter and a...
8805 While living alone in a spooky town, a young g...
8806 Looking to survive in a world taken over by zo...
8807 Dragged from civilian life, a former superhero...
8808 A scrappy but poor boy worms his way into a ty...

```

[9613 rows x 12 columns]

```
[495]: df['listed_in'] = df['listed_in'].str.split(',')
```

```
[496]: df = df.explode('listed_in')
```

```
[497]: df
```

```

[497]:   show_id  type      title  director \
0      s1  Movie  Dick Johnson Is Dead  Kirsten Johnson
1      s2  TV Show      Blood & Water           NaN
1      s2  TV Show      Blood & Water           NaN
1      s2  TV Show      Blood & Water           NaN
2      s3  TV Show      Ganglands  Julien Leclercq

```

...
8807	s8806	Movie	Zoom	Peter Hewitt
8807	s8806	Movie	Zoom	Peter Hewitt
8808	s8807	Movie	Zubaan	Mozez Singh
8808	s8807	Movie	Zubaan	Mozez Singh
8808	s8807	Movie	Zubaan	Mozez Singh

		cast	country \
0		NA	United States
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...		South Africa
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...		South Africa
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...		South Africa
2	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...		NaN
...	
8807	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...		United States
8807	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...		United States
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...		India
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...		India
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...		India

	date_added	release_year	rating	duration \
0	September 25, 2021	2020	PG-13	90 min
1	September 24, 2021	2021	TV-MA	2 Seasons
1	September 24, 2021	2021	TV-MA	2 Seasons
1	September 24, 2021	2021	TV-MA	2 Seasons
2	September 24, 2021	2021	TV-MA	1 Season
...
8807	January 11, 2020	2006	PG	88 min
8807	January 11, 2020	2006	PG	88 min
8808	March 2, 2019	2015	TV-14	111 min
8808	March 2, 2019	2015	TV-14	111 min
8808	March 2, 2019	2015	TV-14	111 min

	listed_in \
0	Documentaries
1	International TV Shows
1	TV Dramas
1	TV Mysteries
2	Crime TV Shows
...	...
8807	Children & Family Movies
8807	Comedies
8808	Dramas
8808	International Movies
8808	Music & Musicals

description

```

0    As her father nears the end of his life, filmm...
1    After crossing paths at a party, a Cape Town t...
1    After crossing paths at a party, a Cape Town t...
1    After crossing paths at a party, a Cape Town t...
2    To protect his family from a powerful drug lor...
...
8807 Dragged from civilian life, a former superhero...
8807 Dragged from civilian life, a former superhero...
8808 A scrappy but poor boy worms his way into a ty...
8808 A scrappy but poor boy worms his way into a ty...
8808 A scrappy but poor boy worms his way into a ty...

```

[20914 rows x 12 columns]

```
[498]: df['listed_in'].value_counts()
```

```

[498]: International Movies      2868
      Dramas                    1701
      Comedies                  1318
      Documentaries             1005
      Action & Adventure         939
      ...
      Romantic Movies           3
      Spanish-Language TV Shows 2
      LGBTQ Movies               1
      TV Sci-Fi & Fantasy        1
      Sports Movies              1
      Name: listed_in, Length: 73, dtype: int64

```

```
[499]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 20914 entries, 0 to 8808
Data columns (total 12 columns):
#   Column          Non-Null Count  Dtype
---  -
0   show_id         20914 non-null object
1   type            20914 non-null object
2   title           20914 non-null object
3   director        15028 non-null object
4   cast            20914 non-null object
5   country         19088 non-null object
6   date_added      20892 non-null object
7   release_year    20912 non-null object
8   rating          20906 non-null object
9   duration        20909 non-null object
10  listed_in       20912 non-null object

```

```

11 description    20912 non-null object
dtypes: object(12)
memory usage: 2.1+ MB

```

```
[500]: df['listed_in'].nunique()
```

```
[500]: 73
```

```
[501]: a = df.groupby(['listed_in', 'director'])['director'].count()
a.sort_values(ascending = False)
```

```
[501]: listed_in      director
Children & Family Movies  Rajiv Chilaka      22
Stand-Up Comedy          Jan Suter          18
                        Raúl Campos          18
Children & Family Movies  Suhas Kadav        16
Stand-Up Comedy          Marcus Raboy        15
..
International Movies     Yibrán Asuad         1
                        Yin Chen-hao         1
                        Yorgos Lanthimos        1
                        You-Jeong Chang         1
Thrillers                 Wilson Coneybeare    1
Name: director, Length: 12516, dtype: int64
```

```
[502]: df
```

```
[502]:
```

	show_id	type	title	director \
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson
1	s2	TV Show	Blood & Water	NaN
1	s2	TV Show	Blood & Water	NaN
1	s2	TV Show	Blood & Water	NaN
2	s3	TV Show	Ganglands	Julien Leclercq
...
8807	s8806	Movie	Zoom	Peter Hewitt
8807	s8806	Movie	Zoom	Peter Hewitt
8808	s8807	Movie	Zubaan	Mozez Singh
8808	s8807	Movie	Zubaan	Mozez Singh
8808	s8807	Movie	Zubaan	Mozez Singh

	cast	country \
0	NA	United States
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
2	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN
...

8807	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States
8807	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India

	date_added	release_year	rating	duration	\
0	September 25, 2021	2020	PG-13	90 min	
1	September 24, 2021	2021	TV-MA	2 Seasons	
1	September 24, 2021	2021	TV-MA	2 Seasons	
1	September 24, 2021	2021	TV-MA	2 Seasons	
2	September 24, 2021	2021	TV-MA	1 Season	
...	
8807	January 11, 2020	2006	PG	88 min	
8807	January 11, 2020	2006	PG	88 min	
8808	March 2, 2019	2015	TV-14	111 min	
8808	March 2, 2019	2015	TV-14	111 min	
8808	March 2, 2019	2015	TV-14	111 min	

	listed_in	\
0	Documentaries	
1	International TV Shows	
1	TV Dramas	
1	TV Mysteries	
2	Crime TV Shows	
...	...	
8807	Children & Family Movies	
8807	Comedies	
8808	Dramas	
8808	International Movies	
8808	Music & Musicals	

	description
0	As her father nears the end of his life, filmm...
1	After crossing paths at a party, a Cape Town t...
1	After crossing paths at a party, a Cape Town t...
1	After crossing paths at a party, a Cape Town t...
2	To protect his family from a powerful drug lor...
...	...
8807	Dragged from civilian life, a former superhero...
8807	Dragged from civilian life, a former superhero...
8808	A scrappy but poor boy worms his way into a ty...
8808	A scrappy but poor boy worms his way into a ty...
8808	A scrappy but poor boy worms his way into a ty...

[20914 rows x 12 columns]

```
[503]: df['release_year'] = df['release_year'].fillna(method = 'ffill')
```

```
[504]: df[df['release_year'] == df['release_year'].max()]
```

```
[504]:
```

	show_id	type	title	director	\
1	s2	TV Show	Blood & Water	NaN	
1	s2	TV Show	Blood & Water	NaN	
1	s2	TV Show	Blood & Water	NaN	
2	s3	TV Show	Ganglands	Julien Leclercq	
2	s3	TV Show	Ganglands	Julien Leclercq	
...	
1696	s1697	TV Show	Polly Pocket	NaN	
2920	s2921	TV Show	Love Is Blind	NaN	
2920	s2921	TV Show	Love Is Blind	NaN	
8439	s8438	TV Show	The Netflix Afterparty	NaN	
8439	s8438	TV Show	The Netflix Afterparty	NaN	

	cast	\
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
2	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	
2	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	
...	...	
1696	Emily Tennant, Shannon Chan-Kent, Kazumi Evans...	
2920	Nick Lachey, Vanessa Lachey	
2920	Nick Lachey, Vanessa Lachey	
8439	David Spade, London Hughes, Fortune Feimster	
8439	David Spade, London Hughes, Fortune Feimster	

	country	date_added	release_year	rating	\
1	South Africa	September 24, 2021	2021	TV-MA	
1	South Africa	September 24, 2021	2021	TV-MA	
1	South Africa	September 24, 2021	2021	TV-MA	
2	NaN	September 24, 2021	2021	TV-MA	
2	NaN	September 24, 2021	2021	TV-MA	
...	
1696	Canada, United States, Ireland	November 15, 2020	2021	TV-Y	
2920	United States	February 13, 2020	2021	TV-MA	
2920	United States	February 13, 2020	2021	TV-MA	
8439	United States	January 2, 2021	2021	TV-MA	
8439	United States	January 2, 2021	2021	TV-MA	

	duration	listed_in	\
1	2 Seasons	International TV Shows	
1	2 Seasons	TV Dramas	
1	2 Seasons	TV Mysteries	

2	1 Season	Crime TV Shows
2	1 Season	International TV Shows
...
1696	2 Seasons	Kids' TV
2920	1 Season	Reality TV
2920	1 Season	Romantic TV Shows
8439	1 Season	Stand-Up Comedy & Talk Shows
8439	1 Season	TV Comedies

	description
1	After crossing paths at a party, a Cape Town t...
1	After crossing paths at a party, a Cape Town t...
1	After crossing paths at a party, a Cape Town t...
2	To protect his family from a powerful drug lor...
2	To protect his family from a powerful drug lor...
...	...
1696	After uncovering a magical locket that allows ...
2920	Nick and Vanessa Lachey host this social exper...
2920	Nick and Vanessa Lachey host this social exper...
8439	Hosts David Spade, Fortune Feimster and London...
8439	Hosts David Spade, Fortune Feimster and London...

[1419 rows x 12 columns]

[505]: df

```
[505]:
```

	show_id	type	title	director \
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson
1	s2	TV Show	Blood & Water	NaN
1	s2	TV Show	Blood & Water	NaN
1	s2	TV Show	Blood & Water	NaN
2	s3	TV Show	Ganglands	Julien Leclercq
...
8807	s8806	Movie	Zoom	Peter Hewitt
8807	s8806	Movie	Zoom	Peter Hewitt
8808	s8807	Movie	Zubaan	Mozez Singh
8808	s8807	Movie	Zubaan	Mozez Singh
8808	s8807	Movie	Zubaan	Mozez Singh

	cast	country \
0	NA	United States
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
2	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN
...
8807	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States

8807	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India

	date_added	release_year	rating	duration	\
0	September 25, 2021	2020	PG-13	90 min	
1	September 24, 2021	2021	TV-MA	2 Seasons	
1	September 24, 2021	2021	TV-MA	2 Seasons	
1	September 24, 2021	2021	TV-MA	2 Seasons	
2	September 24, 2021	2021	TV-MA	1 Season	
...	
8807	January 11, 2020	2006	PG	88 min	
8807	January 11, 2020	2006	PG	88 min	
8808	March 2, 2019	2015	TV-14	111 min	
8808	March 2, 2019	2015	TV-14	111 min	
8808	March 2, 2019	2015	TV-14	111 min	

	listed_in	\
0	Documentaries	
1	International TV Shows	
1	TV Dramas	
1	TV Mysteries	
2	Crime TV Shows	
...	...	
8807	Children & Family Movies	
8807	Comedies	
8808	Dramas	
8808	International Movies	
8808	Music & Musicals	

	description
0	As her father nears the end of his life, filmm...
1	After crossing paths at a party, a Cape Town t...
1	After crossing paths at a party, a Cape Town t...
1	After crossing paths at a party, a Cape Town t...
2	To protect his family from a powerful drug lor...
...	...
8807	Dragged from civilian life, a former superhero...
8807	Dragged from civilian life, a former superhero...
8808	A scrappy but poor boy worms his way into a ty...
8808	A scrappy but poor boy worms his way into a ty...
8808	A scrappy but poor boy worms his way into a ty...

[20914 rows x 12 columns]

```
[506]: my_color = ['#221f1f', '#b20710', '#e50914', '#f5f5f1']
sns.palplot(sns.color_palette(my_color))
plt.title("Netflix brand palette",
↵,loc='left',fontfamily='serif',fontsize=15,y=1.2)
plt.show()
```

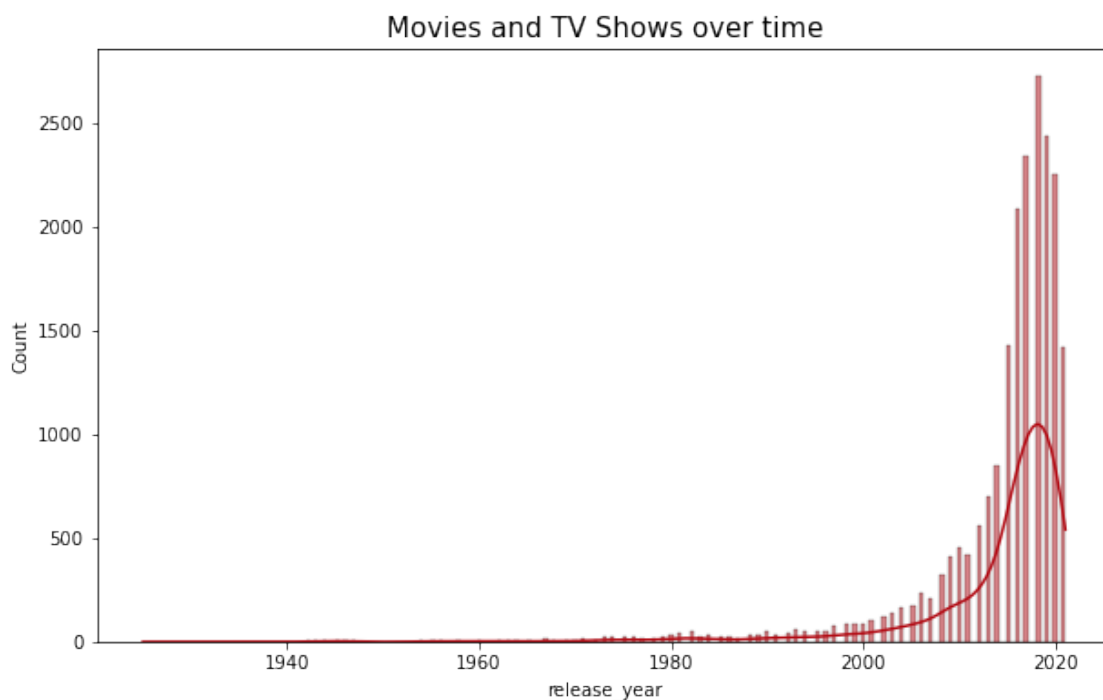
Netflix brand palette



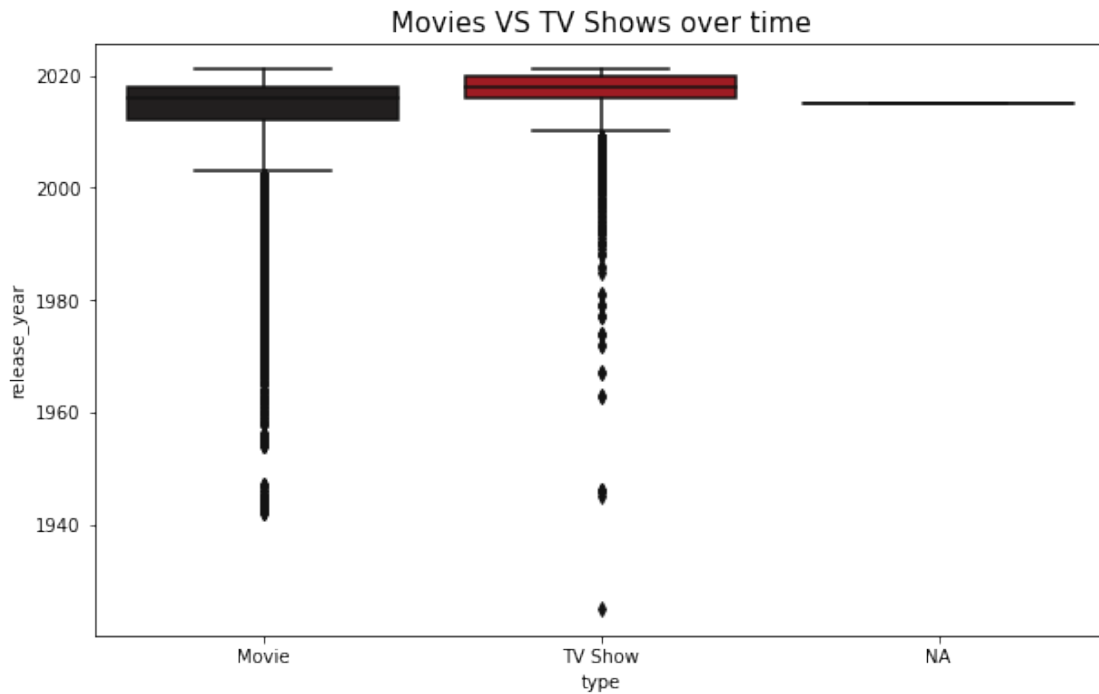
```
[507]: df['release_year'] = df['release_year'].astype(str).astype(int)
```

```
[508]: df['date_added'] = pd.to_datetime(df['date_added'])
```

```
[509]: plt.figure(figsize=(10,6))
sns.histplot(df['release_year'],color = '#b20710',kde= True)
plt.title('Movies and TV Shows over time',fontsize = 15)
plt.show()
```

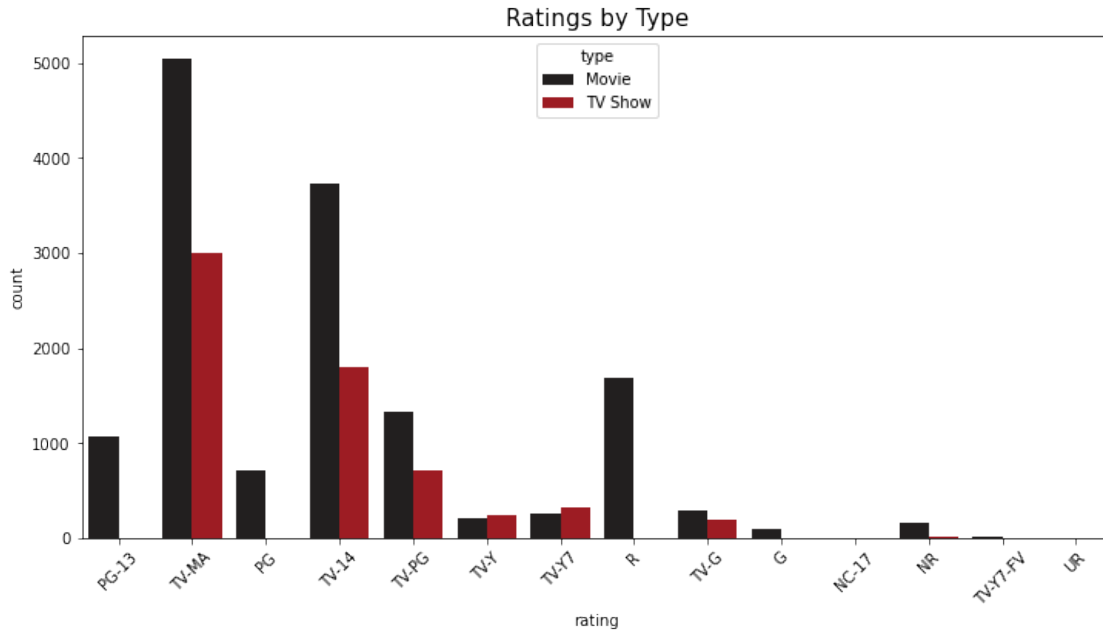


```
[510]: plt.figure(figsize=(10,6))
sns.boxplot(data = df, x = 'type', y = 'release_year',palette = my_color)
plt.title('Movies VS TV Shows over time',fontsize = 15)
plt.show()
```



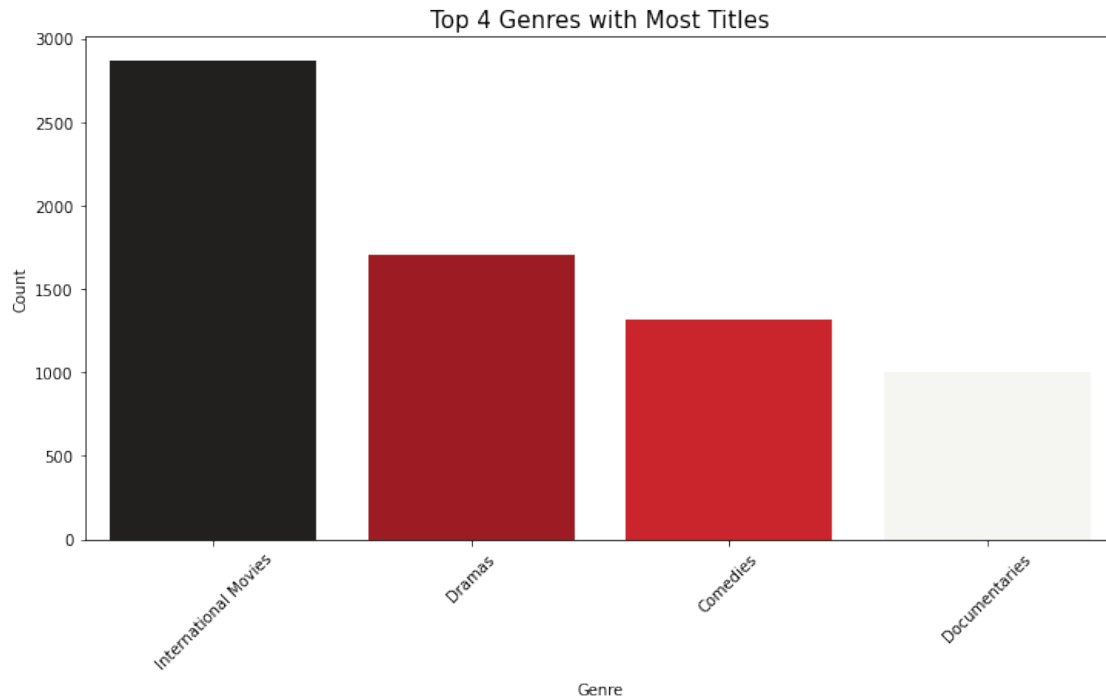
On observing the above plots of Movies VS TV Shows, We could get an insight on type of entertainment Netflix is uploading which are Movies and TV Shows. More of OTT came into our lives in the recent times. We can also infer from the above box plot that more of Movies and TV Shows also got uploaded in the recent years as the demand upsurge for OTT platform. *We could also infer from the plots that more movies got uploaded in Netflix than TV shows. There are a total of 8807 movies and TV shows, out of which 6131 are movies and remaining are TV shows.

```
[511]: df = df[df['rating'].str.contains('min')==False]
plt.figure(figsize = (12,6))
sns.countplot(data = df, x = 'rating', hue = 'type', palette = my_color)
plt.xticks(rotation = 45)
plt.title("Ratings by Type",fontsize = 15)
plt.show()
```



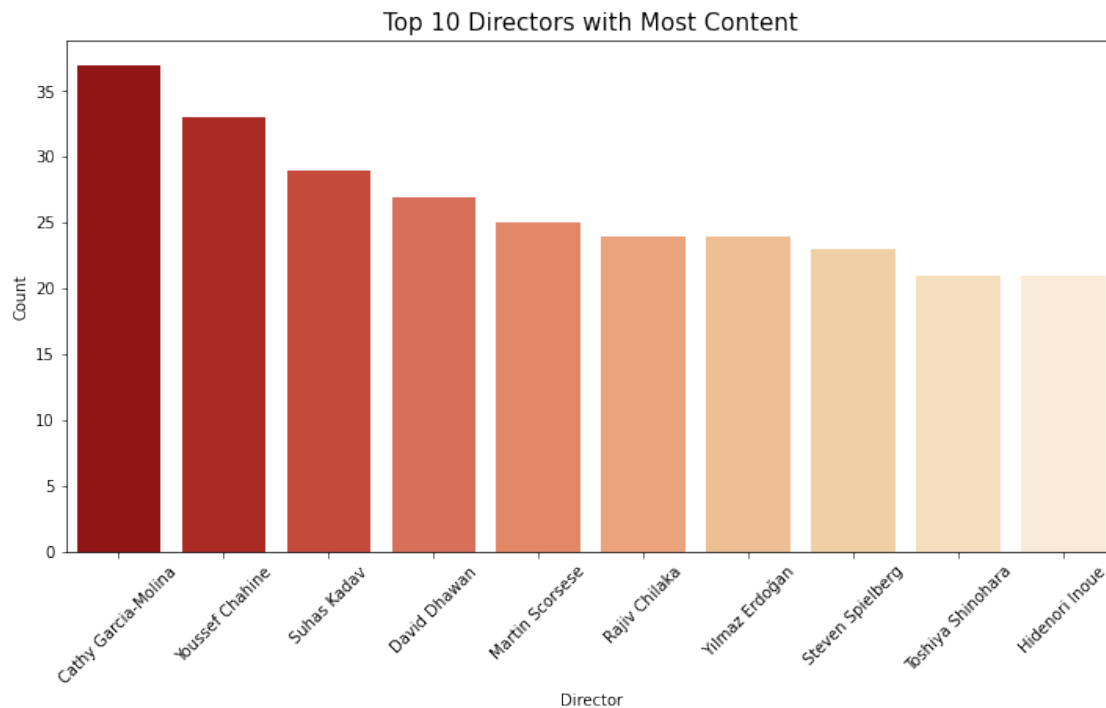
From the ratings plot of Movies and TV shows, *There are 14 kind of ratings given to Movies and TV shows.* Out of 14 kind of ratings, these movies and TV shows acquired pretty good ratings in TV-MA, TV-14 and TV-PG. In addition to these, if we consider only movies, PG-13 and R ratings would also get added to the above listed ratings.

```
[512]: plt.figure(figsize=(12, 6))
sns.countplot(x='listed_in', data=df, order=df['listed_in'].value_counts().
            index[:4], palette = my_color)
plt.title('Top 4 Genres with Most Titles', fontsize = 15)
plt.xticks(rotation = 45)
plt.xlabel('Genre')
plt.ylabel('Count')
plt.show()
```



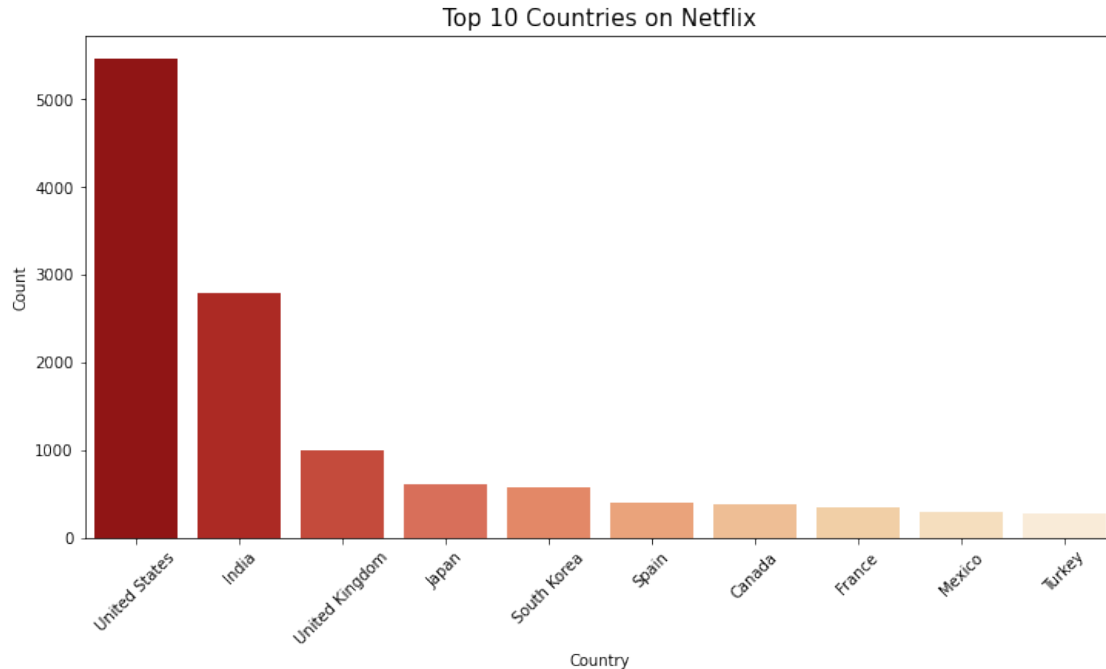
Above plot is the countplot of genres listing the Top 4 genres. For the people who are interested in watching International movies, Dramas and Comedies genre, Netflix would stand as a great option. *Documentaries stands next to the Top 3 Genres in Netflix.

```
[513]: top_directors = df['director'].value_counts().head(10)
plt.figure(figsize=(12, 6))
sns.barplot(data = df, x = top_directors.index, y = top_directors.values,
            palette = 'OrRd_r')
plt.xlabel('Director')
plt.ylabel('Count')
plt.title('Top 10 Directors with Most Content', fontsize = 15)
plt.xticks(rotation=45)
plt.show()
```



Top 10 Directors with most content is the plot listing the Top 10 directors who have highest number of their movies uploaded in Netflix

```
[514]: plt.figure(figsize=(12,6))
top_countries = df['country'].value_counts().head(10)
sns.barplot(data = df, x = top_countries.index,y = top_countries.values,
            palette = 'OrRd_r')
plt.xlabel('Country')
plt.ylabel('Count')
plt.title('Top 10 Countries on Netflix',fontsize = 15)
plt.xticks(rotation=45)
plt.show()
```



```
[515]: df1 = df[(df['country'].isin(['United States','India','United Kingdom'])) &
↳ (df['listed_in'].isin(['International',
↳ 'Movies','Dramas','Comedies','Documentaries','TV Shows']))]
df1
```

```
[515]:
```

	show_id	type	title	director \
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson
9	s10	Movie	The Starling	Theodore Melfi
24	s25	Movie	Jeans	S. Shankar
27	s28	Movie	Grown Ups	Dennis Dugan
42	s43	Movie	Jaws 2	Jeannot Szwarc
...
8793	s8792	Movie	Young Adult	Jason Reitman
8800	s8799	Movie	Zed Plus	Chandra Prakash Dwivedi
8801	s8800	Movie	Zenda	Avadhoot Gupte
8806	s8805	Movie	Zombieland	Ruben Fleischer
8808	s8807	Movie	Zubaan	Mozez Singh

	cast	country \
0	NA	United States
9	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States
24	Prashanth, Aishwarya Rai Bachchan, Sri Lakshmi...	India
27	Adam Sandler, Kevin James, Chris Rock, David S...	United States
42	Roy Scheider, Lorraine Gary, Murray Hamilton, ...	United States
...

8793	Charlize Theron, Patton Oswalt, Patrick Wilson...	United States
8800	Adil Hussain, Mona Singh, K.K. Raina, Sanjay M...	India
8801	Santosh Juvekar, Siddharth Chandekar, Sachit P...	India
8806	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States
8808	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India

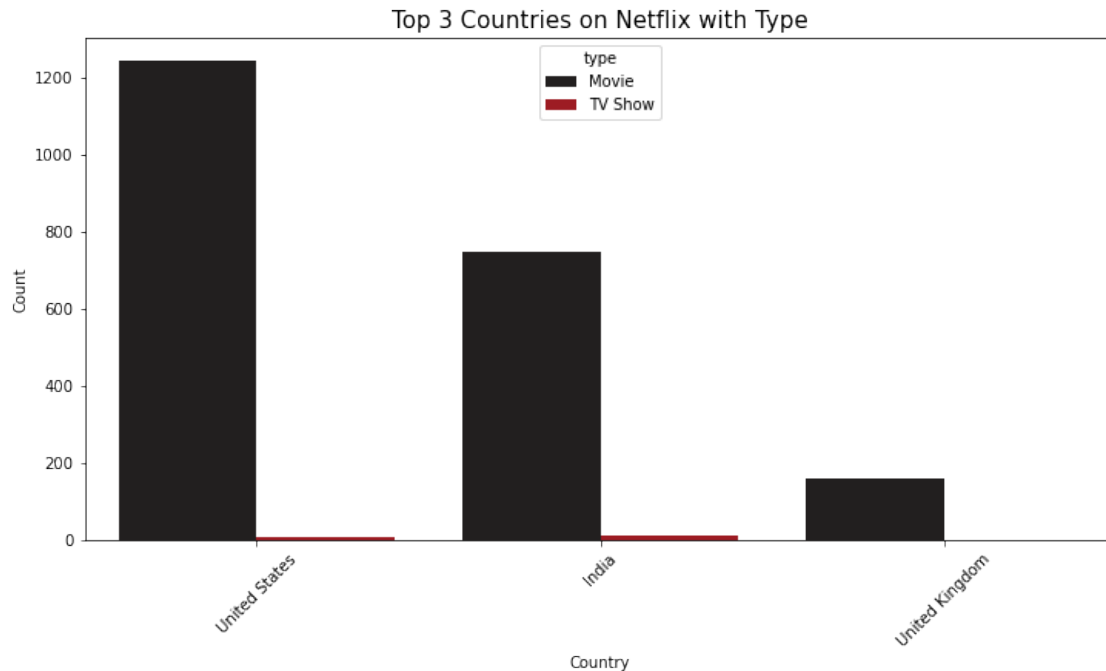
	date_added	release_year	rating	duration	listed_in \
0	2021-09-25	2020	PG-13	90 min	Documentaries
9	2021-09-24	2021	PG-13	104 min	Comedies
24	2021-09-21	1998	TV-14	166 min	Comedies
27	2021-09-20	2010	PG-13	103 min	Comedies
42	2021-09-16	1978	PG	116 min	Dramas
...
8793	2019-11-20	2011	R	94 min	Comedies
8800	2019-12-31	2014	TV-MA	131 min	Comedies
8801	2018-02-15	2009	TV-14	120 min	Dramas
8806	2019-11-01	2009	R	88 min	Comedies
8808	2019-03-02	2015	TV-14	111 min	Dramas

	description
0	As her father nears the end of his life, filmm...
9	A woman adjusting to life after a loss contend...
24	When the father of the man she loves insists t...
27	Mourning the loss of their beloved junior high...
42	Four years after the last deadly shark attacks...
...	...
8793	When a divorced writer gets a letter from an o...
8800	A philandering small-town mechanic's political...
8801	A change in the leadership of a political part...
8806	Looking to survive in a world taken over by zo...
8808	A scrappy but poor boy worms his way into a ty...

[2170 rows x 12 columns]

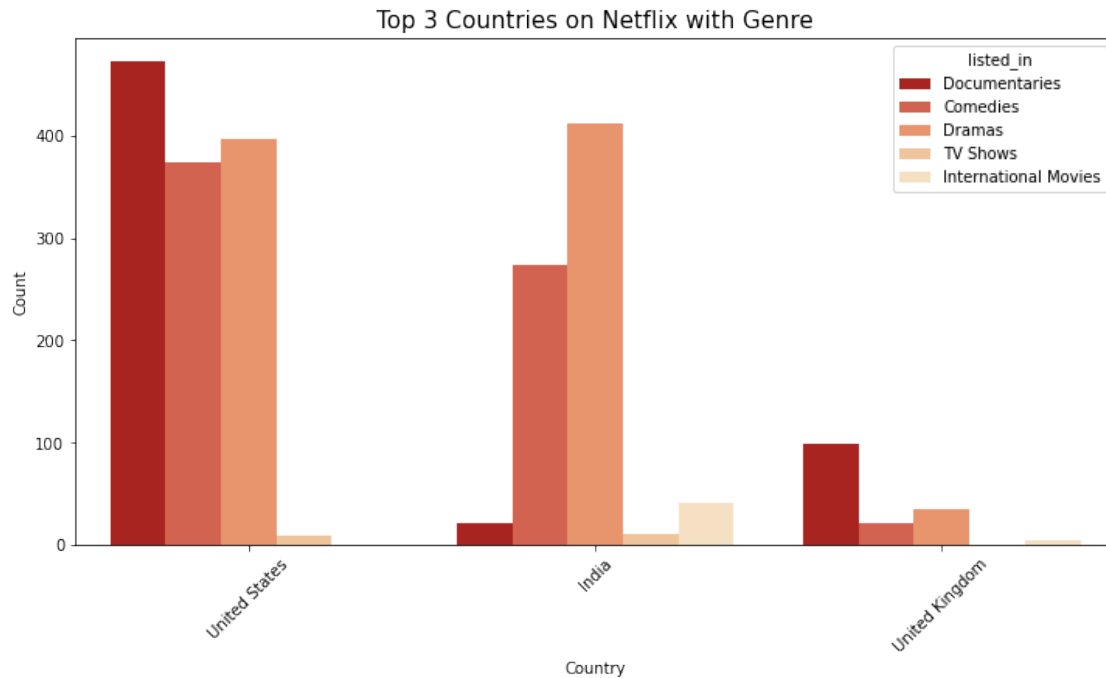
Above plot is the countplot of countries listed with Top 10 Countries which released highest number of movies and TV shows. The United States has released the highest number of movies among the top 10 countries. India and UK are in the Top 2 and 3 positions.

```
[516]: plt.figure(figsize=(12,6))
sns.countplot(data = df1, x = 'country',hue='type',palette = my_color)
plt.xlabel('Country')
plt.ylabel('Count')
plt.title('Top 3 Countries on Netflix with Type',fontsize = 15)
plt.xticks(rotation=45)
plt.show()
```

On observing the above plot, we can infer from the plot that there are more number of movies than TV shows from the Top 3 countries who have their highest number of movies and TV shows uploaded in Netflix.

```
[517]: plt.figure(figsize=(12,6))
sns.countplot(data = df1, x = 'country', hue='listed_in', palette = 'OrRd_r')
plt.xlabel('Country')
plt.ylabel('Count')
plt.title('Top 3 Countries on Netflix with Genre', fontsize = 15)
plt.xticks(rotation=45)
plt.show()
```



Plotting taking into consideration the Top 3 countries and Top 5 genres,US has more number of movies and TV shows in Documentaries and Comedies genre,India listed in more number of movies and TV shows in Dramas and Comedies genre,UK has more number of documentaries.

[]: