

netflix-business-case-study

May 27, 2024

#About NETFLIX ##### Netflix is one of the most popular media and video streaming platforms. They have over 10000 movies or tv shows available on their platform, as of mid-2021, they have over 222M Subscribers globally. This tabular dataset consists of listings of all the movies and tv shows available on Netflix, along with details such as - cast, directors, ratings, release year, duration, etc.

1 Business Problem

Analyze the data and generate insights that could help Netflix in deciding which type of shows/movies to produce and how they can grow the business in different countries

```
[37]: ### importing all the required libraries
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
```

```
[38]: df = pd.read_csv('https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/000/940/original/netflix.csv')
```

```
[40]: df.head(5)
```

```
[40]:  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

                                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

    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

```

```

                                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 ...

```

```

                                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...

```

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

```

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

```

```
[42]: df.shape
```

```
[42]: (8807, 12)
```

1.1 Observation on Data : Understanding the columns in the dataset, shape, data type of all attributes, missing value detection.

```
[113]: # Filtering Data Column wise so that each row has one director, one genre, one
        # actor, one country (checking for nesting and then un-nesting)
        # 1. Separating Casts in each column

        #finding nesting
        column = 'cast'
        df[df[column].apply(lambda x: "," in str(x))]

        #fixing nesting
        small_df = df[['title', 'cast']]
        small_df["list_of_cast"] = small_df["cast"].apply(lambda x: str(x).split(","))
        small_df=small_df.explode("list_of_cast")
        small_df.head(5)
```

```
[113]:
```

	title	cast \
0	Dick Johnson Is Dead	NaN
1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...

	list_of_cast
0	nan
1	Ama Qamata
1	Khosi Ngema
1	Gail Mabalane
1	Thabang Molaba

```
[114]: #fixing nesting for countries
        new_df = df[['title', 'country']]
        new_df["list_of_country"] = new_df["country"].apply(lambda x: str(x).split(","))
        new_df=new_df.explode("list_of_country")
        new_df.head(5)
```

```
[114]:
```

	title	country	list_of_country
0	Dick Johnson Is Dead	United States	United States
1	Blood & Water	South Africa	South Africa
2	Ganglands	NaN	nan
3	Jailbirds New Orleans	NaN	nan
4	Kota Factory	India	India

```
[115]: #fixing nesting for genre
        new_df1 = df[['title', 'listed_in']]
```

```
new_df1["list_of_genre"] = new_df1["listed_in"].apply(lambda x: str(x).
↳split(", "))
new_df1=new_df1.explode("list_of_genre")
new_df1.sample(5)
```

```
[115]:
```

	title \	listed_in \	list_of_genre
198	King of Boys: The Return of the King		Crime TV Shows
2012	The Barrier		Spanish-Language TV Shows
4027	This Evening		International TV Shows
3955	Pinky Memsaab		Dramas
2491	The School of Mischief		Dramas

```
[116]: new_df2 = df[['title', 'director']]
new_df2["list_of_directors"] = new_df2["director"].apply(lambda x: str(x).
↳split(", "))
new_df2=new_df2.explode("list_of_directors")
new_df2.sample(5)
```

```
[116]:
```

	title	director	list_of_directors
7166	Kaleidoscope	Rupert Jones	Rupert Jones
170	Rhyme & Reason	Peter Spirer	Peter Spirer
4843	Pedal the World	Felix Starck	Felix Starck
1380	Go Dog Go	NaN	nan
1137	The Knight and the Princess	Bashir El Deek	Bashir El Deek

2 Merging the dataframes:

```
[117]: df1 = pd.merge(small_df, new_df, on='title')
df1.head(5)
```

```
[117]:
```

	title	cast \
0	Dick Johnson Is Dead	NaN

1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
2	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
3	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
4	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...

	list_of_cast	country	list_of_country
0	nan	United States	United States
1	Ama Qamata	South Africa	South Africa
2	Khosi Ngema	South Africa	South Africa
3	Gail Mabalane	South Africa	South Africa
4	Thabang Molaba	South Africa	South Africa

```
[118]: df2 = pd.merge(df1,new_df1,on='title')
df2.head(5)
```

```
[118]:
```

	title	cast \
0	Dick Johnson Is Dead	NaN
1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
2	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
3	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
4	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...

	list_of_cast	country	list_of_country \
0	nan	United States	United States
1	Ama Qamata	South Africa	South Africa
2	Ama Qamata	South Africa	South Africa
3	Ama Qamata	South Africa	South Africa
4	Khosi Ngema	South Africa	South Africa

	listed_in	list_of_genre
0	Documentaries	Documentaries
1	International TV Shows, TV Dramas, TV Mysteries	International TV Shows
2	International TV Shows, TV Dramas, TV Mysteries	TV Dramas
3	International TV Shows, TV Dramas, TV Mysteries	TV Mysteries
4	International TV Shows, TV Dramas, TV Mysteries	International TV Shows

```
[119]: df3=pd.merge(df2,new_df2,on='title')
df3.head()
```

```
[119]:
```

	title	cast \
0	Dick Johnson Is Dead	NaN
1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
2	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
3	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
4	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...

	list_of_cast	country	list_of_country \
--	--------------	---------	-------------------

0	nan	United States	United States
1	Ama Qamata	South Africa	South Africa
2	Ama Qamata	South Africa	South Africa
3	Ama Qamata	South Africa	South Africa
4	Khosi Ngema	South Africa	South Africa

	listed_in	list_of_genre \
0	Documentaries	Documentaries
1	International TV Shows, TV Dramas, TV Mysteries	International TV Shows
2	International TV Shows, TV Dramas, TV Mysteries	TV Dramas
3	International TV Shows, TV Dramas, TV Mysteries	TV Mysteries
4	International TV Shows, TV Dramas, TV Mysteries	International TV Shows

	director	list_of_directors
0	Kirsten Johnson	Kirsten Johnson
1	NaN	nan
2	NaN	nan
3	NaN	nan
4	NaN	nan

```
[120]: df4 = pd.merge(df3,df,on='title',how='left')
df4.head()
```

```
[120]:
```

	title	cast_x \
0	Dick Johnson Is Dead	NaN
1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
2	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
3	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
4	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...

	list_of_cast	country_x	list_of_country \
0	nan	United States	United States
1	Ama Qamata	South Africa	South Africa
2	Ama Qamata	South Africa	South Africa
3	Ama Qamata	South Africa	South Africa
4	Khosi Ngema	South Africa	South Africa

	listed_in_x	list_of_genre \
0	Documentaries	Documentaries
1	International TV Shows, TV Dramas, TV Mysteries	International TV Shows
2	International TV Shows, TV Dramas, TV Mysteries	TV Dramas
3	International TV Shows, TV Dramas, TV Mysteries	TV Mysteries
4	International TV Shows, TV Dramas, TV Mysteries	International TV Shows

	director_x	list_of_directors	show_id	type	director_y \
0	Kirsten Johnson	Kirsten Johnson	s1	Movie	Kirsten Johnson
1	NaN	nan	s2	TV Show	NaN

2	NaN	nan	s2	TV Show	NaN
3	NaN	nan	s2	TV Show	NaN
4	NaN	nan	s2	TV Show	NaN

	cast_y	country_y	\
0	NaN	United States	
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	
2	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	
3	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	
4	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	

	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	2 Seasons	
3	September 24, 2021	2021	TV-MA	2 Seasons	
4	September 24, 2021	2021	TV-MA	2 Seasons	

	listed_in_y	\
0	Documentaries	
1	International TV Shows, TV Dramas, TV Mysteries	
2	International TV Shows, TV Dramas, TV Mysteries	
3	International TV Shows, TV Dramas, TV Mysteries	
4	International TV Shows, TV Dramas, TV Mysteries	

	description
0	As her father nears the end of his life, filmm...
1	After crossing paths at a party, a Cape Town t...
2	After crossing paths at a party, a Cape Town t...
3	After crossing paths at a party, a Cape Town t...
4	After crossing paths at a party, a Cape Town t...

```
[52]: df5= df4.
      ↪drop(columns=['cast_x','country_x','listed_in_x','director_x','cast_y','country_y','listed_

#dropping the unnecessary columns.
```

```
[121]: df4["month"]=df4["date_added"].apply(lambda x:str(x).split(', '))
df4["month"]=df4["month"].str[-2]
df4.head()

df4["year"]=df4["date_added"].apply(lambda x:str(x).split(', '))
df4["year"]=df4["year"].str[-1]
df4.head()
```

```
[121]: title cast_x \
0 Dick Johnson Is Dead NaN
```

1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
2	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
3	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
4	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...

	list_of_cast	country_x	list_of_country	\
0	nan	United States	United States	
1	Ama Qamata	South Africa	South Africa	
2	Ama Qamata	South Africa	South Africa	
3	Ama Qamata	South Africa	South Africa	
4	Khosi Ngema	South Africa	South Africa	

	listed_in_x	list_of_genre	\
0	Documentaries	Documentaries	
1	International TV Shows, TV Dramas, TV Mysteries	International TV Shows	
2	International TV Shows, TV Dramas, TV Mysteries	TV Dramas	
3	International TV Shows, TV Dramas, TV Mysteries	TV Mysteries	
4	International TV Shows, TV Dramas, TV Mysteries	International TV Shows	

	director_x	list_of_directors	show_id	...	\
0	Kirsten Johnson	Kirsten Johnson	s1	...	
1	NaN	nan	s2	...	
2	NaN	nan	s2	...	
3	NaN	nan	s2	...	
4	NaN	nan	s2	...	

	cast_y	country_y	\
0	NaN	United States	
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	
2	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	
3	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	
4	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	

	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	2 Seasons	
3	September 24, 2021	2021	TV-MA	2 Seasons	
4	September 24, 2021	2021	TV-MA	2 Seasons	

	listed_in_y	\
0	Documentaries	
1	International TV Shows, TV Dramas, TV Mysteries	
2	International TV Shows, TV Dramas, TV Mysteries	
3	International TV Shows, TV Dramas, TV Mysteries	
4	International TV Shows, TV Dramas, TV Mysteries	

		description	month	year
0	As her father nears the end of his life, filmm...	September 25	2021	
1	After crossing paths at a party, a Cape Town t...	September 24	2021	
2	After crossing paths at a party, a Cape Town t...	September 24	2021	
3	After crossing paths at a party, a Cape Town t...	September 24	2021	
4	After crossing paths at a party, a Cape Town t...	September 24	2021	

[5 rows x 22 columns]

```
[122]: df4.drop(['date_added'],axis =1,inplace=True)
df4.head()
```

```
[122]:
```

	title	cast_x \
0	Dick Johnson Is Dead	NaN
1	Blood & Water Ama Qamata, Khosi Ngema, Gail Mabalan...	
2	Blood & Water Ama Qamata, Khosi Ngema, Gail Mabalan...	
3	Blood & Water Ama Qamata, Khosi Ngema, Gail Mabalan...	
4	Blood & Water Ama Qamata, Khosi Ngema, Gail Mabalan...	

	list_of_cast	country_x	list_of_country \
0	nan	United States	United States
1	Ama Qamata	South Africa	South Africa
2	Ama Qamata	South Africa	South Africa
3	Ama Qamata	South Africa	South Africa
4	Khosi Ngema	South Africa	South Africa

	listed_in_x	list_of_genre \
0	Documentaries	Documentaries
1	International TV Shows, TV Dramas, TV Mysteries	International TV Shows
2	International TV Shows, TV Dramas, TV Mysteries	TV Dramas
3	International TV Shows, TV Dramas, TV Mysteries	TV Mysteries
4	International TV Shows, TV Dramas, TV Mysteries	International TV Shows

	director_x	list_of_directors	show_id ...	director_y \
0	Kirsten Johnson	Kirsten Johnson	s1 ...	Kirsten Johnson
1	NaN	nan	s2 ...	NaN
2	NaN	nan	s2 ...	NaN
3	NaN	nan	s2 ...	NaN
4	NaN	nan	s2 ...	NaN

	cast_y	country_y \
0	NaN	United States
1	Ama Qamata, Khosi Ngema, Gail Mabalan...	South Africa
2	Ama Qamata, Khosi Ngema, Gail Mabalan...	South Africa
3	Ama Qamata, Khosi Ngema, Gail Mabalan...	South Africa
4	Ama Qamata, Khosi Ngema, Gail Mabalan...	South Africa

```

    release_year  rating  duration  \
0         2020    PG-13      90 min
1         2021    TV-MA    2 Seasons
2         2021    TV-MA    2 Seasons
3         2021    TV-MA    2 Seasons
4         2021    TV-MA    2 Seasons

                                listed_in_y  \
0                                Documentaries
1  International TV Shows, TV Dramas, TV Mysteries
2  International TV Shows, TV Dramas, TV Mysteries
3  International TV Shows, TV Dramas, TV Mysteries
4  International TV Shows, TV Dramas, TV Mysteries

                                description      month  year
0  As her father nears the end of his life, filmm...  September 25  2021
1  After crossing paths at a party, a Cape Town t...  September 24  2021
2  After crossing paths at a party, a Cape Town t...  September 24  2021
3  After crossing paths at a party, a Cape Town t...  September 24  2021
4  After crossing paths at a party, a Cape Town t...  September 24  2021

[5 rows x 21 columns]

```

2.1 Handling Null Values

```

[123]: df4["list_of_directors"].replace(["nan"], ["Unknown Director"], inplace=True)
df4["list_of_cast"].replace(["nan"], ["Unknown Cast"], inplace=True)
df4["list_of_genre"].replace(["nan"], ["Unknown Genre"], inplace=True)
df4["list_of_country"].replace(["nan"], ["Unknown Country"], inplace=True)

df4.head()

```

```

[123]:
                                title                                cast_x  \
0  Dick Johnson Is Dead                                             NaN
1      Blood & Water  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
2      Blood & Water  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
3      Blood & Water  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
4      Blood & Water  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...

    list_of_cast  country_x list_of_country  \
0  Unknown Cast  United States  United States
1    Ama Qamata  South Africa  South Africa
2    Ama Qamata  South Africa  South Africa
3    Ama Qamata  South Africa  South Africa
4    Khosi Ngema  South Africa  South Africa

```

	listed_in_x	list_of_genre \
0	Documentaries	Documentaries
1	International TV Shows, TV Dramas, TV Mysteries	International TV Shows
2	International TV Shows, TV Dramas, TV Mysteries	TV Dramas
3	International TV Shows, TV Dramas, TV Mysteries	TV Mysteries
4	International TV Shows, TV Dramas, TV Mysteries	International TV Shows

	director_x	list_of_directors	show_id ...	director_y \
0	Kirsten Johnson	Kirsten Johnson	s1 ...	Kirsten Johnson
1	NaN	UnKnown Director	s2 ...	NaN
2	NaN	UnKnown Director	s2 ...	NaN
3	NaN	UnKnown Director	s2 ...	NaN
4	NaN	UnKnown Director	s2 ...	NaN

	cast_y	country_y \
0	NaN	United States
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
2	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
3	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
4	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa

	release_year	rating	duration \
0	2020	PG-13	90 min
1	2021	TV-MA	2 Seasons
2	2021	TV-MA	2 Seasons
3	2021	TV-MA	2 Seasons
4	2021	TV-MA	2 Seasons

	listed_in_y \
0	Documentaries
1	International TV Shows, TV Dramas, TV Mysteries
2	International TV Shows, TV Dramas, TV Mysteries
3	International TV Shows, TV Dramas, TV Mysteries
4	International TV Shows, TV Dramas, TV Mysteries

	description	month	year
0	As her father nears the end of his life, filmm...	September 25	2021
1	After crossing paths at a party, a Cape Town t...	September 24	2021
2	After crossing paths at a party, a Cape Town t...	September 24	2021
3	After crossing paths at a party, a Cape Town t...	September 24	2021
4	After crossing paths at a party, a Cape Town t...	September 24	2021

[5 rows x 21 columns]

```
[56]: df4.isnull().sum()
```

```
[56]: title          0
      list_of_cast    0
      list_of_country  0
      list_of_genre    0
      list_of_directors 0
      show_id         0
      type            0
      release_year     0
      rating          67
      duration         3
      description      0
      month           158
      year            0
      dtype: int64
```

```
[124]: #converting data_added to datetime.

df4['date_added'] = pd.
      ↳to_datetime(df['date_added'],errors='coerce',format='mixed')
df4.head()
```

```
[124]:
```

	title	cast_x \
0	Dick Johnson Is Dead	NaN
1	Blood & Water Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
2	Blood & Water Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
3	Blood & Water Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
4	Blood & Water Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	

	list_of_cast	country_x	list_of_country \
0	UnKnown Cast	United States	United States
1	Ama Qamata	South Africa	South Africa
2	Ama Qamata	South Africa	South Africa
3	Ama Qamata	South Africa	South Africa
4	Khosi Ngema	South Africa	South Africa

	listed_in_x	list_of_genre \
0	Documentaries	Documentaries
1	International TV Shows, TV Dramas, TV Mysteries	International TV Shows
2	International TV Shows, TV Dramas, TV Mysteries	TV Dramas
3	International TV Shows, TV Dramas, TV Mysteries	TV Mysteries
4	International TV Shows, TV Dramas, TV Mysteries	International TV Shows

	director_x	list_of_directors	show_id ... \
0	Kirsten Johnson	Kirsten Johnson	s1 ...
1	NaN	UnKnown Director	s2 ...
2	NaN	UnKnown Director	s2 ...
3	NaN	UnKnown Director	s2 ...

```
4          NaN  UnKnown Director      s2  ...
```

```

                                cast_y      country_y  \
0                                NaN  United States
1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...  South Africa
2  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...  South Africa
3  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...  South Africa
4  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...  South Africa

```

```

release_year rating      duration  \
0          2020  PG-13         90 min
1          2021  TV-MA      2 Seasons
2          2021  TV-MA      2 Seasons
3          2021  TV-MA      2 Seasons
4          2021  TV-MA      2 Seasons

```

```

                                listed_in_y  \
0                                Documentaries
1  International TV Shows, TV Dramas, TV Mysteries
2  International TV Shows, TV Dramas, TV Mysteries
3  International TV Shows, TV Dramas, TV Mysteries
4  International TV Shows, TV Dramas, TV Mysteries

```

```

                                description      month      year  \
0  As her father nears the end of his life, filmm...  September 25      2021
1  After crossing paths at a party, a Cape Town t...  September 24      2021
2  After crossing paths at a party, a Cape Town t...  September 24      2021
3  After crossing paths at a party, a Cape Town t...  September 24      2021
4  After crossing paths at a party, a Cape Town t...  September 24      2021

```

```

date_added
0 2021-09-25
1 2021-09-24
2 2021-09-24
3 2021-09-24
4 2021-09-24

```

```
[5 rows x 22 columns]
```

```
[125]: df4['date_added'].fillna(0, inplace=True)
df4['rating'].fillna(0,inplace=True)
df4['duration'].fillna(0,inplace=True)
df4.head()
```

```

[125]:          title                                cast_x  \
0  Dick Johnson Is Dead                                NaN
1    Blood & Water  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...

```

2	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
3	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
4	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...

	list_of_cast	country_x	list_of_country	\
0	UnKnown Cast	United States	United States	
1	Ama Qamata	South Africa	South Africa	
2	Ama Qamata	South Africa	South Africa	
3	Ama Qamata	South Africa	South Africa	
4	Khosi Ngema	South Africa	South Africa	

		listed_in_x	list_of_genre	\
0		Documentaries	Documentaries	
1	International TV Shows, TV Dramas, TV Mysteries		International TV Shows	
2	International TV Shows, TV Dramas, TV Mysteries		TV Dramas	
3	International TV Shows, TV Dramas, TV Mysteries		TV Mysteries	
4	International TV Shows, TV Dramas, TV Mysteries		International TV Shows	

	director_x	list_of_directors	show_id	...	\
0	Kirsten Johnson	Kirsten Johnson	s1	...	
1	NaN	UnKnown Director	s2	...	
2	NaN	UnKnown Director	s2	...	
3	NaN	UnKnown Director	s2	...	
4	NaN	UnKnown Director	s2	...	

		cast_y	country_y	\
0		NaN	United States	
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...		South Africa	
2	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...		South Africa	
3	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...		South Africa	
4	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...		South Africa	

	release_year	rating	duration	\
0	2020	PG-13	90 min	
1	2021	TV-MA	2 Seasons	
2	2021	TV-MA	2 Seasons	
3	2021	TV-MA	2 Seasons	
4	2021	TV-MA	2 Seasons	

		listed_in_y	\
0		Documentaries	
1	International TV Shows, TV Dramas, TV Mysteries		
2	International TV Shows, TV Dramas, TV Mysteries		
3	International TV Shows, TV Dramas, TV Mysteries		
4	International TV Shows, TV Dramas, TV Mysteries		

	description	month	year	\
--	-------------	-------	------	---

```

0 As her father nears the end of his life, filmm... September 25 2021
1 After crossing paths at a party, a Cape Town t... September 24 2021
2 After crossing paths at a party, a Cape Town t... September 24 2021
3 After crossing paths at a party, a Cape Town t... September 24 2021
4 After crossing paths at a party, a Cape Town t... September 24 2021

```

```

      date_added
0 2021-09-25 00:00:00
1 2021-09-24 00:00:00
2 2021-09-24 00:00:00
3 2021-09-24 00:00:00
4 2021-09-24 00:00:00

```

[5 rows x 22 columns]

```
[59]: df4.isnull().sum()
```

```

[59]: title                0
      list_of_cast         0
      list_of_country       0
      list_of_genre         0
      list_of_directors     0
      show_id              0
      type                 0
      release_year         0
      rating               0
      duration             0
      description          0
      month                158
      year                 0
      date_added           0
      dtype: int64

```

3 Finding the counts of each categorical variable both using graphical and non- graphical analysis.

Non-Graphical Analysis: Value counts and unique attributes

Unique Attributes:

```

[60]: print("No of Movie title:-", df['title'].unique())
      print("No. of Ratings:-", df['rating'].unique())
      print("No. of Directors:-", df['director'].unique())
      print("No.of Casts:-",df['cast'].unique())
      print("No.of Country:-",df['country'].unique())
      print("Total Genre:-",df['listed_in'].unique())
      print("Total Type:-",df['type'].unique())

```

No of Movie title:- 8807
 No. of Ratings:- 17
 No. of Directors:- 4528
 No.of Casts:- 7692
 No.of Country:- 748
 Total Genre:- 514
 Total Type:- 2

[61]: *#Value Count of casts*

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

```
[61]: cast
David Attenborough
19
Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna Bhardwaj, Rajesh Kava, Mousam,
Swapnil
14
Samuel West
10
Jeff Dunham
7
David Spade, London Hughes, Fortune Feimster
6
..
Michael Peña, Diego Luna, Tenoch Huerta, Joaquin Cosio, José María Yazpik, Matt
Letscher, Alyssa Diaz
1
Nick Lachey, Vanessa Lachey
1
Takeru Sato, Kasumi Arimura, Haru, Kentaro Sakaguchi, Takayuki Yamada, Kendo
Kobayashi, Ken Yasuda, Arata Furuta, Suzuki Matsuo, Koichi Yamadera, Arata Iura,
Chikako Kaku, Kotaro Yoshida      1
Toyin Abraham, Sambasa Nzeribe, Chioma Chukwuka Akpotha, Chioma Omeruah,
Chiwetalu Agu, Dele Odule, Femi Adebayo, Bayray McNwizu, Biodun Stephen
1
Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanana, Manish Chaudhary, Meghna Malik,
Malkeet Rauni, Anita Shabdish, Chittaranjan Tripathy
1
Name: count, Length: 7692, dtype: int64
```

[62]: *#value counts of type*

```
df['type'].value_counts().reset_index()
```

```
[62]:      type  count
0    Movie   6131
1  TV Show   2676
```



```
[63]: #value counts for country
df['country'].value_counts().reset_index()
```

```
[63]:
```

	country	count
0	United States	2818
1	India	972
2	United Kingdom	419
3	Japan	245
4	South Korea	199
..
743	Romania, Bulgaria, Hungary	1
744	Uruguay, Guatemala	1
745	France, Senegal, Belgium	1
746	Mexico, United States, Spain, Colombia	1
747	United Arab Emirates, Jordan	1

[748 rows x 2 columns]

```
[64]: #value counts for genres
df['listed_in'].value_counts().reset_index()
```

```
[64]:
```

	listed_in	count
0	Dramas, International Movies	362
1	Documentaries	359
2	Stand-Up Comedy	334
3	Comedies, Dramas, International Movies	274
4	Dramas, Independent Movies, International Movies	252
..
509	Kids' TV, TV Action & Adventure, TV Dramas	1
510	TV Comedies, TV Dramas, TV Horror	1
511	Children & Family Movies, Comedies, LGBTQ Movies	1
512	Kids' TV, Spanish-Language TV Shows, Teen TV S...	1
513	Cult Movies, Dramas, Thrillers	1

[514 rows x 2 columns]

```
[65]: #value count for casts
df['cast'].value_counts().reset_index()
```

```
[65]:
```

	cast	count
0	David Attenborough	19
1	Vatsal Dubey, Julie Tejawani, Rupa Bhimani, Jig...	14
2	Samuel West	10
3	Jeff Dunham	7
4	David Spade, London Hughes, Fortune Feimster	6
...

7687	Michael Peña, Diego Luna, Tenoch Huerta, Joaqui...	1
7688	Nick Lachey, Vanessa Lachey	1
7689	Takeru Sato, Kasumi Arimura, Haru, Kentaro Sak...	1
7690	Toyin Abraham, Sambasa Nzeribe, Chioma Chukwuk...	1
7691	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	1

[7692 rows x 2 columns]

```
[66]: #value count for directors
df['director'].value_counts().reset_index()
```

```
[66]:
```

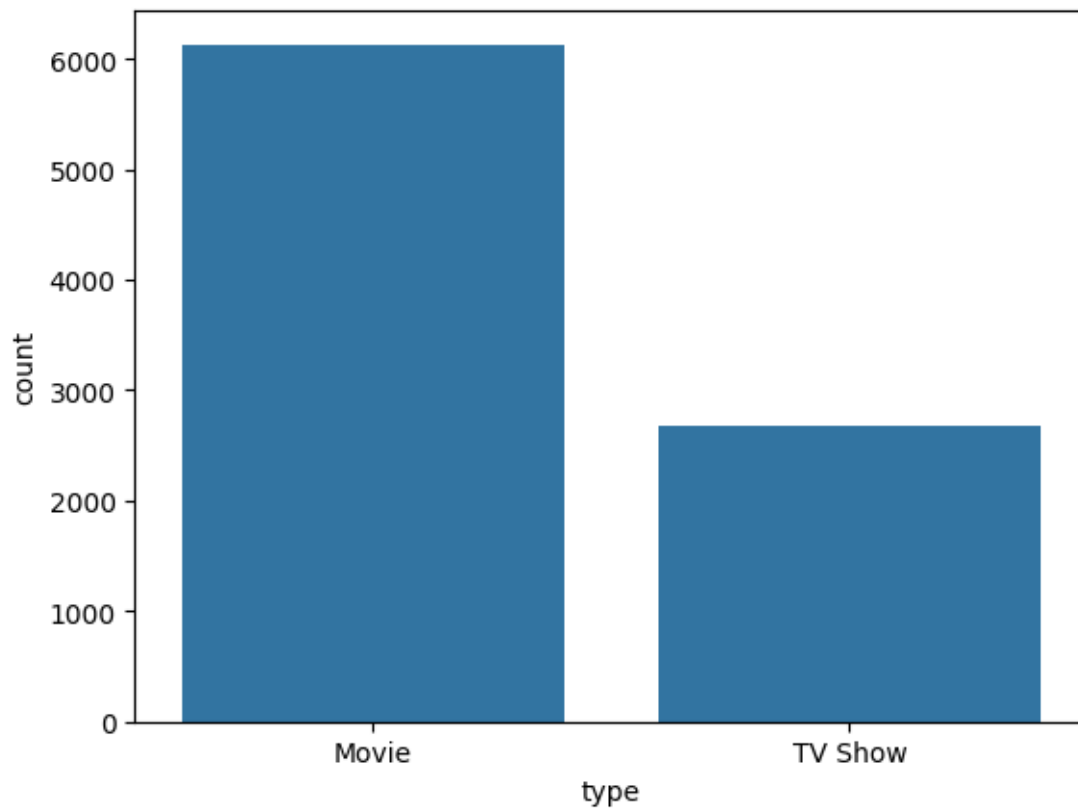
	director	count
0	Rajiv Chilaka	19
1	Raúl Campos, Jan Suter	18
2	Marcus Raboy	16
3	Suhas Kadav	16
4	Jay Karas	14
...
4523	Raymie Muzquiz, Stu Livingston	1
4524	Joe Menendez	1
4525	Eric Bross	1
4526	Will Eisenberg	1
4527	Mozez Singh	1

[4528 rows x 2 columns]

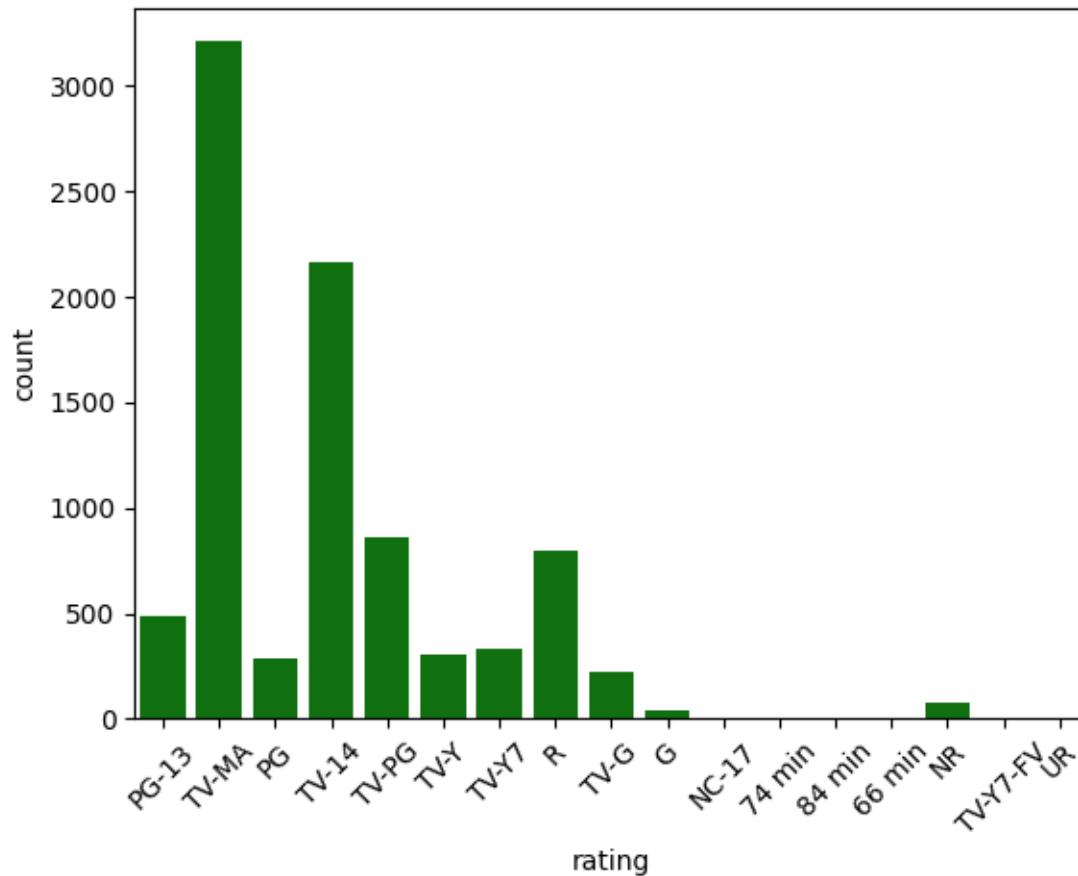
4 Graphical Representation using countplot:

4.0.1 graphical representation of total movies & Tv show(Univariate Analysis)

```
[67]: sns.countplot(x='type',data=df)
plt.show()
```



```
[68]: sns.countplot(x='rating',data=df,color='g')  
plt.xticks(rotation=45)  
plt.show()
```



5 Comparison of tv shows vs. movies.

[126]: *#Seprating the tables for movies and series*

```
def func(X):
    if X=='Movie':
        return 1
    else:
        return 0
df4['new']=df4['type'].apply(func)
df4.head()
```

[126]:

	title	cast_x \
0	Dick Johnson Is Dead	NaN
1	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalone, Thaban...
2	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalone, Thaban...
3	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalone, Thaban...
4	Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalone, Thaban...

	list_of_cast	country_x	list_of_country	\
0	UnKnown Cast	United States	United States	
1	Ama Qamata	South Africa	South Africa	
2	Ama Qamata	South Africa	South Africa	
3	Ama Qamata	South Africa	South Africa	
4	Khosi Ngema	South Africa	South Africa	

	listed_in_x	list_of_genre	\
0	Documentaries	Documentaries	
1	International TV Shows, TV Dramas, TV Mysteries	International TV Shows	
2	International TV Shows, TV Dramas, TV Mysteries	TV Dramas	
3	International TV Shows, TV Dramas, TV Mysteries	TV Mysteries	
4	International TV Shows, TV Dramas, TV Mysteries	International TV Shows	

	director_x	list_of_directors	show_id	...	country_y	release_year	\
0	Kirsten Johnson	Kirsten Johnson	s1	...	United States	2020	
1	NaN	UnKnown Director	s2	...	South Africa	2021	
2	NaN	UnKnown Director	s2	...	South Africa	2021	
3	NaN	UnKnown Director	s2	...	South Africa	2021	
4	NaN	UnKnown Director	s2	...	South Africa	2021	

	rating	duration	listed_in_y	\
0	PG-13	90 min	Documentaries	
1	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	
2	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	
3	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	
4	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	

	description	month	year	\
0	As her father nears the end of his life, filmm...	September 25	2021	
1	After crossing paths at a party, a Cape Town t...	September 24	2021	
2	After crossing paths at a party, a Cape Town t...	September 24	2021	
3	After crossing paths at a party, a Cape Town t...	September 24	2021	
4	After crossing paths at a party, a Cape Town t...	September 24	2021	

	date_added	new
0	2021-09-25 00:00:00	1
1	2021-09-24 00:00:00	0
2	2021-09-24 00:00:00	0
3	2021-09-24 00:00:00	0
4	2021-09-24 00:00:00	0

[5 rows x 23 columns]

```
[70]: #movies :
df_final_movies = df4[df4["new"] == 1].copy()
df_final_movies.drop(['new'],axis=1,inplace=True)
```

```
[127]: df_final_movies.head()
```

```
[127]:
```

		title	list_of_cast	list_of_country	\
0		Dick Johnson Is Dead	UnKnown Cast	United States	
159		My Little Pony: A New Generation	Vanessa Hudgens	Unknown Country	
160		My Little Pony: A New Generation	Vanessa Hudgens	Unknown Country	
161		My Little Pony: A New Generation	Kimiko Glenn	Unknown Country	
162		My Little Pony: A New Generation	Kimiko Glenn	Unknown Country	

		list_of_genre	list_of_directors	show_id	type	release_year	\
0		Documentaries	Kirsten Johnson	s1	Movie	2020	
159		Children & Family Movies	Robert Cullen	s7	Movie	2021	
160		Children & Family Movies	José Luis Ucha	s7	Movie	2021	
161		Children & Family Movies	Robert Cullen	s7	Movie	2021	
162		Children & Family Movies	José Luis Ucha	s7	Movie	2021	

		rating	duration	description	\
0	PG-13	90 min	As her father nears the end of his life, filmm...		
159	PG	91 min	Equestria's divided. But a bright-eyed hero be...		
160	PG	91 min	Equestria's divided. But a bright-eyed hero be...		
161	PG	91 min	Equestria's divided. But a bright-eyed hero be...		
162	PG	91 min	Equestria's divided. But a bright-eyed hero be...		

		month	year
0		September 25	2021
159		September 24	2021
160		September 24	2021
161		September 24	2021
162		September 24	2021

```
[72]: df_final_movies.drop(['date_added'],axis=1,inplace=True)
```

```
[128]: #series
df_final_series = df4[df4["new"] == 0].copy()
df_final_series.drop(['new'],axis=1,inplace=True)
df_final_series.head()
```

```
[128]:
```

		title	cast_x	\
1		Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
2		Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
3		Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
4		Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	
5		Blood & Water	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	

		list_of_cast	country_x	list_of_country	\
1		Ama Qamata	South Africa	South Africa	
2		Ama Qamata	South Africa	South Africa	

3	Ama Qamata	South Africa	South Africa
4	Khosi Ngema	South Africa	South Africa
5	Khosi Ngema	South Africa	South Africa

	listed_in_x	list_of_genre \
1	International TV Shows, TV Dramas, TV Mysteries	International TV Shows
2	International TV Shows, TV Dramas, TV Mysteries	TV Dramas
3	International TV Shows, TV Dramas, TV Mysteries	TV Mysteries
4	International TV Shows, TV Dramas, TV Mysteries	International TV Shows
5	International TV Shows, TV Dramas, TV Mysteries	TV Dramas

	director_x	list_of_directors	show_id	...	\
1	NaN	UnKnown Director	s2	...	
2	NaN	UnKnown Director	s2	...	
3	NaN	UnKnown Director	s2	...	
4	NaN	UnKnown Director	s2	...	
5	NaN	UnKnown Director	s2	...	

	cast_y	country_y \
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
2	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
3	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
4	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa
5	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa

	release_year	rating	duration	\
1	2021	TV-MA	2 Seasons	
2	2021	TV-MA	2 Seasons	
3	2021	TV-MA	2 Seasons	
4	2021	TV-MA	2 Seasons	
5	2021	TV-MA	2 Seasons	

	listed_in_y \
1	International TV Shows, TV Dramas, TV Mysteries
2	International TV Shows, TV Dramas, TV Mysteries
3	International TV Shows, TV Dramas, TV Mysteries
4	International TV Shows, TV Dramas, TV Mysteries
5	International TV Shows, TV Dramas, TV Mysteries

	description	month	year \
1	After crossing paths at a party, a Cape Town t...	September 24	2021
2	After crossing paths at a party, a Cape Town t...	September 24	2021
3	After crossing paths at a party, a Cape Town t...	September 24	2021
4	After crossing paths at a party, a Cape Town t...	September 24	2021
5	After crossing paths at a party, a Cape Town t...	September 24	2021

date_added

```

1 2021-09-24 00:00:00
2 2021-09-24 00:00:00
3 2021-09-24 00:00:00
4 2021-09-24 00:00:00
5 2021-09-24 00:00:00

```

[5 rows x 22 columns]

```
[74]: df_final_series.drop(['date_added'],axis=1,inplace=True)
```

```
[75]: df_final_movies.head(5)
```

```
[75]:
```

	title	list_of_cast	list_of_country	
0	Dick Johnson Is Dead	UnKnown Cast	United States	
159	My Little Pony: A New Generation	Vanessa Hudgens	Unknown Country	
160	My Little Pony: A New Generation	Vanessa Hudgens	Unknown Country	
161	My Little Pony: A New Generation	Kimiko Glenn	Unknown Country	
162	My Little Pony: A New Generation	Kimiko Glenn	Unknown Country	

	list_of_genre	list_of_directors	show_id	type	release_year	
0	Documentaries	Kirsten Johnson	s1	Movie	2020	
159	Children & Family Movies	Robert Cullen	s7	Movie	2021	
160	Children & Family Movies	José Luis Ucha	s7	Movie	2021	
161	Children & Family Movies	Robert Cullen	s7	Movie	2021	
162	Children & Family Movies	José Luis Ucha	s7	Movie	2021	

	rating	duration	description	
0	PG-13	90 min	As her father nears the end of his life, filmm...	
159	PG	91 min	Equestria's divided. But a bright-eyed hero be...	
160	PG	91 min	Equestria's divided. But a bright-eyed hero be...	
161	PG	91 min	Equestria's divided. But a bright-eyed hero be...	
162	PG	91 min	Equestria's divided. But a bright-eyed hero be...	

	month	year
0	September 25	2021
159	September 24	2021
160	September 24	2021
161	September 24	2021
162	September 24	2021

```
[76]: df_final_series.head(5)
```

```
[76]:
```

	title	list_of_cast	list_of_country	list_of_genre	
1	Blood & Water	Ama Qamata	South Africa	International TV Shows	
2	Blood & Water	Ama Qamata	South Africa	TV Dramas	
3	Blood & Water	Ama Qamata	South Africa	TV Mysteries	
4	Blood & Water	Khosi Ngema	South Africa	International TV Shows	

5	Blood & Water	Khosi Ngema	South Africa	TV Dramas
---	---------------	-------------	--------------	-----------

	list_of_directors	show_id	type	release_year	rating	duration	\
1	UnKnown Director	s2	TV Show	2021	TV-MA	2 Seasons	
2	UnKnown Director	s2	TV Show	2021	TV-MA	2 Seasons	
3	UnKnown Director	s2	TV Show	2021	TV-MA	2 Seasons	
4	UnKnown Director	s2	TV Show	2021	TV-MA	2 Seasons	
5	UnKnown Director	s2	TV Show	2021	TV-MA	2 Seasons	

	description	month	year
1	After crossing paths at a party, a Cape Town t...	September 24	2021
2	After crossing paths at a party, a Cape Town t...	September 24	2021
3	After crossing paths at a party, a Cape Town t...	September 24	2021
4	After crossing paths at a party, a Cape Town t...	September 24	2021
5	After crossing paths at a party, a Cape Town t...	September 24	2021

```
[77]: #Find the number of movies produced in each country and pick the top 10
      ↪countries
```

```
df_final_movies.groupby('list_of_country')['title'].nunique().
      ↪sort_values(ascending=False).head(10)
```

```
[77]: list_of_country
United States      2364
India              927
Unknown Country    440
United States      388
United Kingdom     382
Canada             187
France             155
United Kingdom     152
France             148
Canada             132
Name: title, dtype: int64
```

```
[78]: #Find the number of series produced in each country and pick the top 10
      ↪countries
```

```
df_final_series.groupby('list_of_country')['title'].nunique().
      ↪sort_values(ascending=False).head(10)
```

```
[78]: list_of_country
United States      847
Unknown Country    391
United Kingdom     246
Japan              174
South Korea        164
```

United States	91
Canada	84
India	81
Taiwan	70
France	64

Name: title, dtype: int64

6 Analysis of actors/directors of different types of shows/movies.

[79]: *#Identify the top 10 directors who have appeared in most movies or TV shows.*

```
df_final_movies.groupby('list_of_directors')['title'].nunique().
↳sort_values(ascending=False).head(10)
```

[79]: list_of_directors

UnKnown Director	188
Rajiv Chilaka	22
Raúl Campos	18
Jan Suter	18
Suhas Kadav	16
Jay Karas	15
Marcus Raboy	15
Cathy Garcia-Molina	13
Jay Chapman	12
Martin Scorsese	12

Name: title, dtype: int64

[80]: *#Identify the top 10 directors who have appeared in most TV shows.*

```
df_final_series.groupby('list_of_directors')['title'].nunique().
↳sort_values(ascending=False).head(10)
```

[80]: list_of_directors

UnKnown Director	2446
Ken Burns	3
Alastair Fothergill	3
Rob Seidenglanz	2
Shin Won-ho	2
Jung-ah Im	2
Stan Lathan	2
Gautham Vasudev Menon	2
Joe Berlinger	2
Hsu Fu-chun	2

Name: title, dtype: int64

```
[81]: #Which genre movies are more popular or produced more

df_final_movies.groupby('list_of_genre')['title'].nunique().
↳sort_values(ascending=False).head(5)
```

```
[81]: list_of_genre
      International Movies    2624
      Dramas                1600
      Comedies              1210
      Action & Adventure     859
      Documentaries         829
      Name: title, dtype: int64
```

```
[83]: #Find After how many days the movie will be added to Netflix after the release
↳of the movie.

df4['relase_date'] = pd.to_datetime(df4['release_year'], format= '%Y')
```

7 Visual Analysis - Univariate, Bivariate after pre-processing of the data

```
[84]: # considering the top datas from both Movies and TV-SHOWS
#Movies

top3_directors = df_final_movies['list_of_directors'].value_counts().index[:3]
top3_casts = df_final_movies['list_of_cast'].value_counts().index[:3]
top5_duration = df_final_movies['duration'].value_counts().index[:5]
top5_genre = df_final_movies['list_of_genre'].value_counts().index[:5]
top5_country = df_final_movies['list_of_country'].value_counts().index[:5]
top5_rating = df_final_movies['rating'].value_counts().index[:5]
top5_type = df_final_movies['type'].value_counts().index[:5]
top_25_years=df_final_movies["release_year"].value_counts().index[:25]
top5_title = df_final_movies['title'].value_counts().index[:5]
top5_months = df_final_movies['month'].value_counts().index[:5]
```

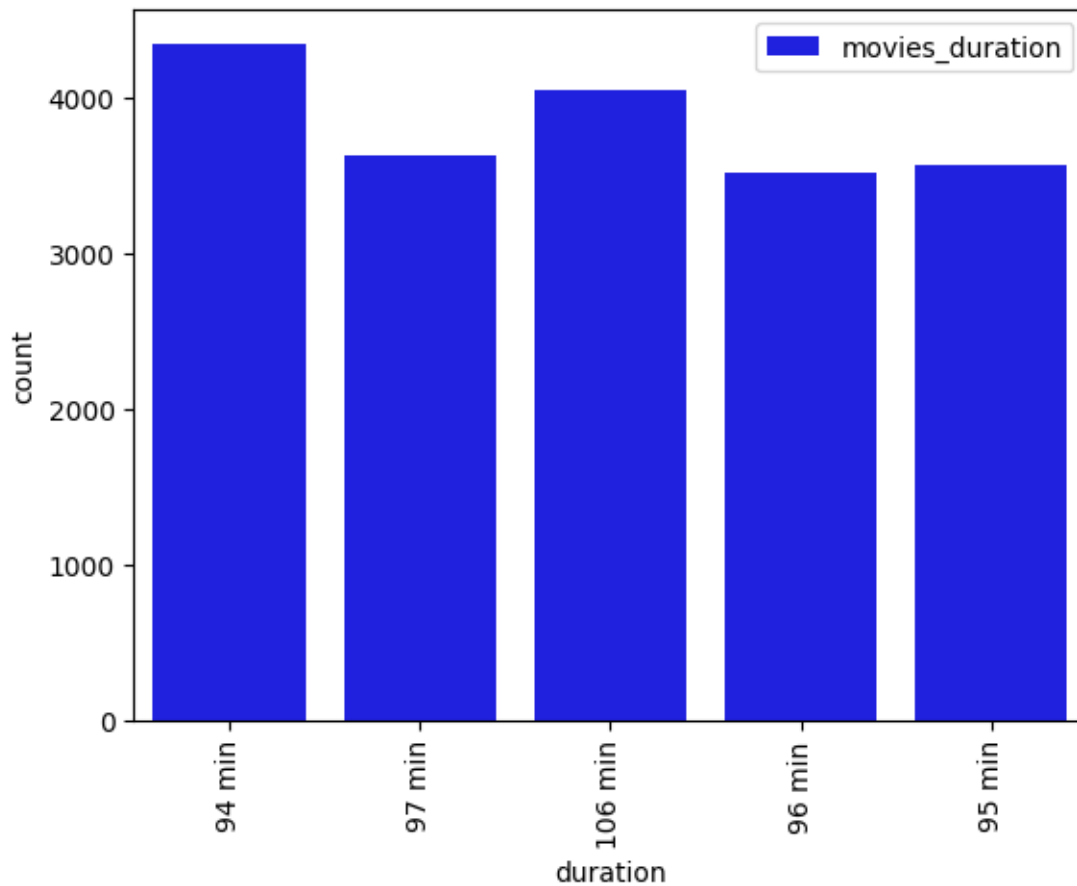
```
[85]: #Series

top3_directors1 = df_final_series['list_of_directors'].value_counts().index[:3]
top3_casts1 = df_final_series['list_of_cast'].value_counts().index[:3]
top5_duration1 = df_final_series['duration'].value_counts().index[:5]
top5_genre1 = df_final_series['list_of_genre'].value_counts().index[:5]
top5_country1 = df_final_series['list_of_country'].value_counts().index[:5]
top5_rating1 = df_final_series['rating'].value_counts().index[:5]
top5_type1 = df_final_series['type'].value_counts().index[:5]
top_25_years1=df_final_series["release_year"].value_counts().index[:25]
top5_title1 = df_final_series['title'].value_counts().index[:5]
```

```
top5_months1 = df_final_series['month'].value_counts().index[:5]
```

```
[86]: #Countplot for duration
#movies

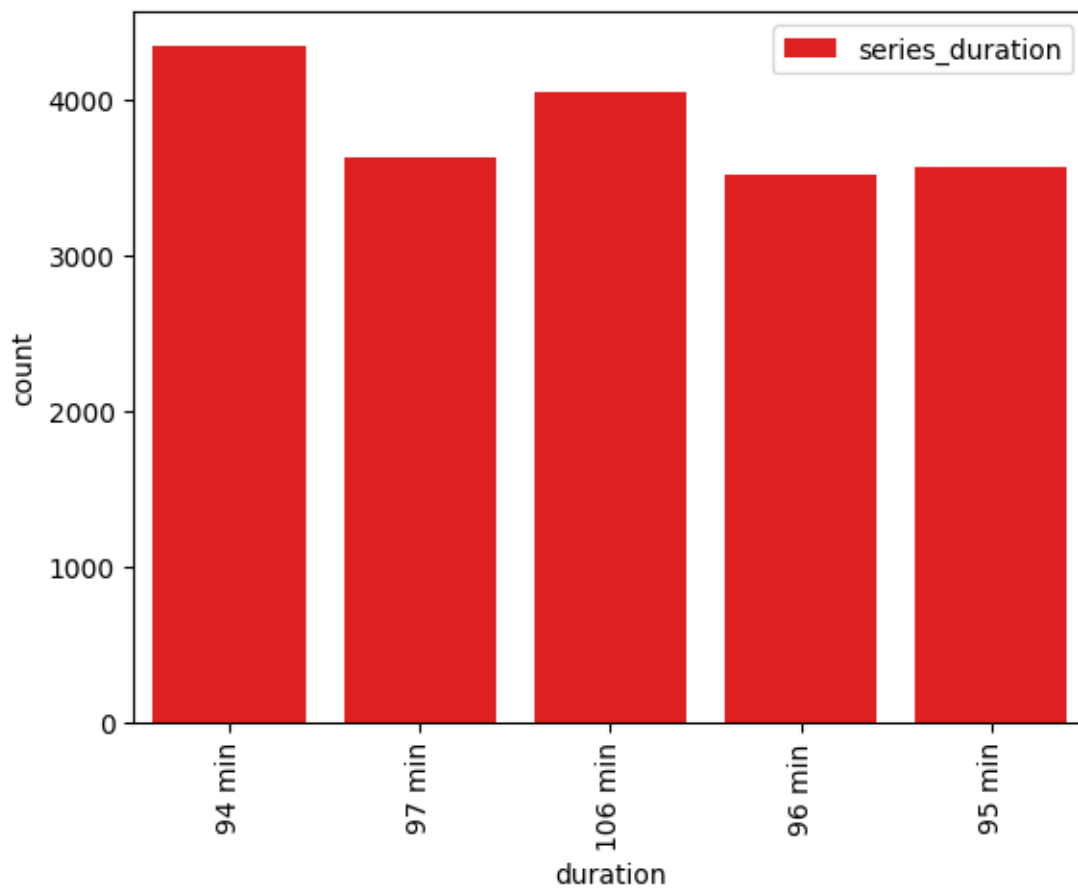
top_5_duration_movies=df_final_movies.loc[(df_final_movies["duration"].
↪isin(top5_duration))]
sns.countplot(data=top_5_duration_movies,x="duration",color="blue")
plt.xticks(rotation=90)
plt.legend(["movies_duration"])
plt.show()
```



```
[87]: #series:

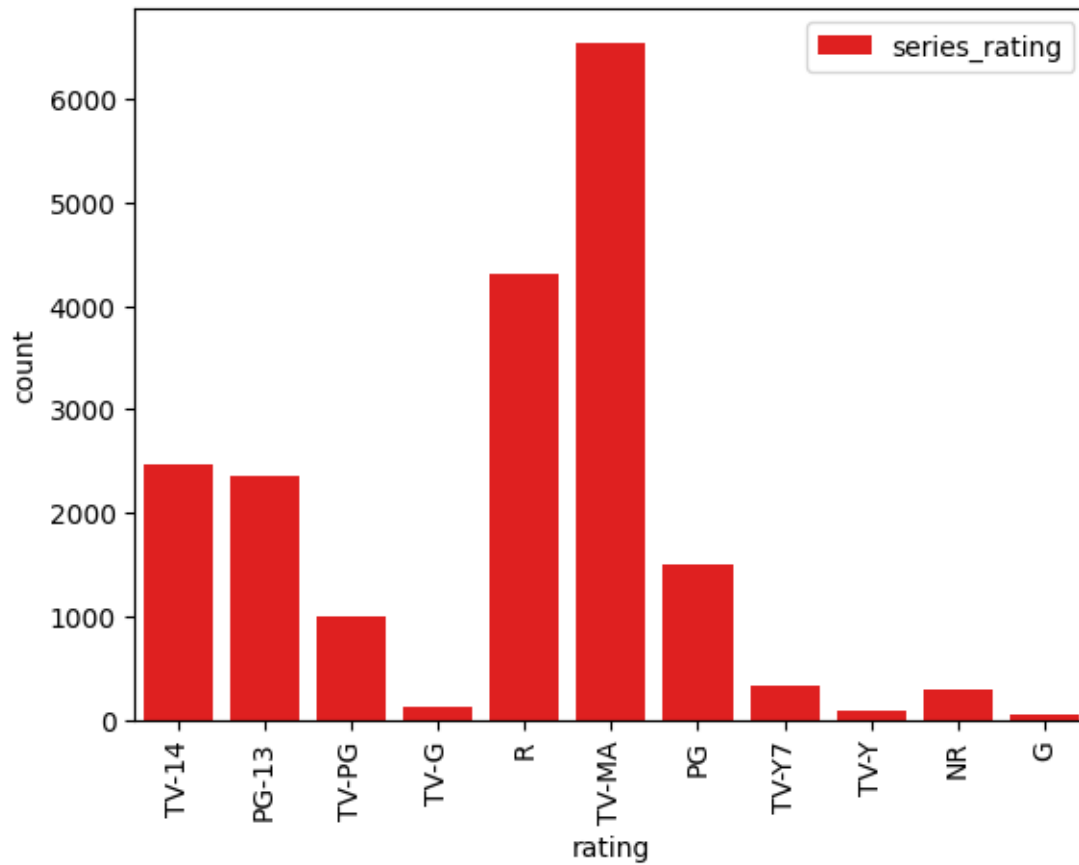
top_5_duration_series=df_final_series.loc[(df_final_series["duration"].
↪isin(top5_duration1))]
sns.countplot(data=top_5_duration_movies,x="duration",color="red")
plt.xticks(rotation=90)
```

```
plt.legend(["series_duration"])
plt.show()
```

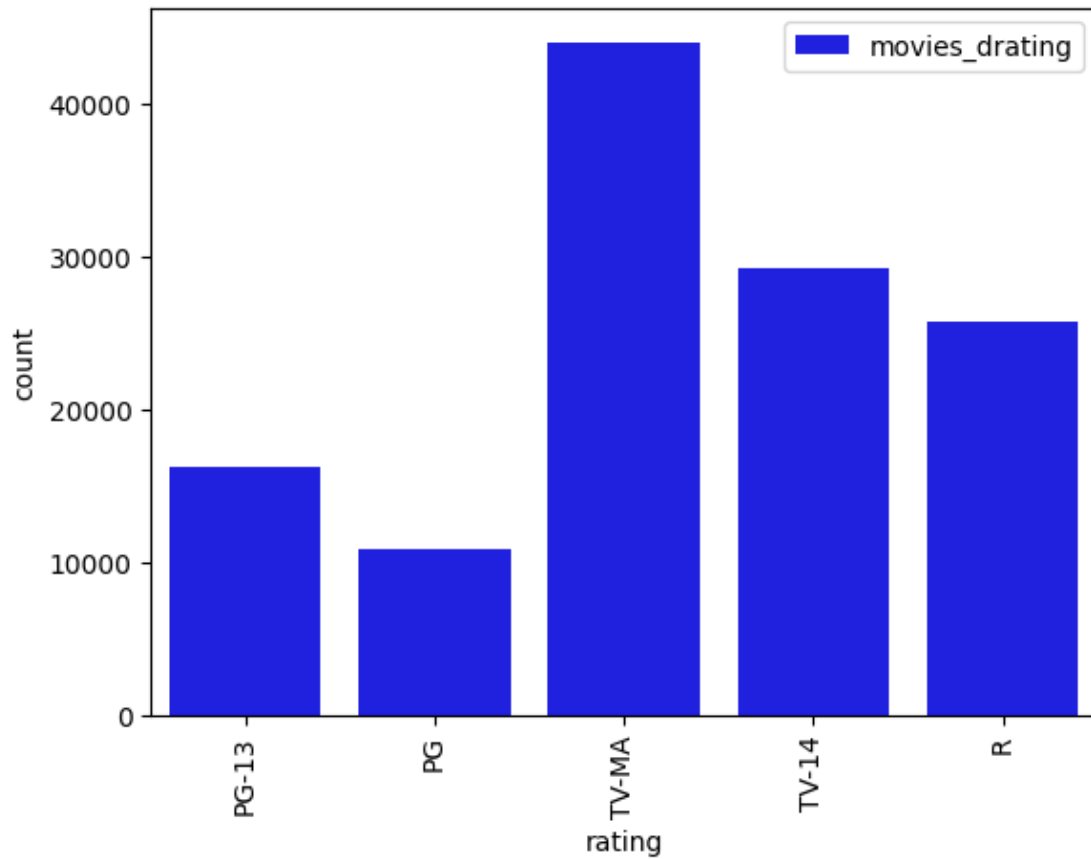


```
[88]: #countplot_for_rating
#series

top_5_rating_series=df_final_series.loc[(df_final_series["rating"].
    ↳isin(top5_rating1))]
sns.countplot(data=top_5_duration_movies,x="rating",color="red")
plt.xticks(rotation=90)
plt.legend(["series_rating"])
plt.show()
```

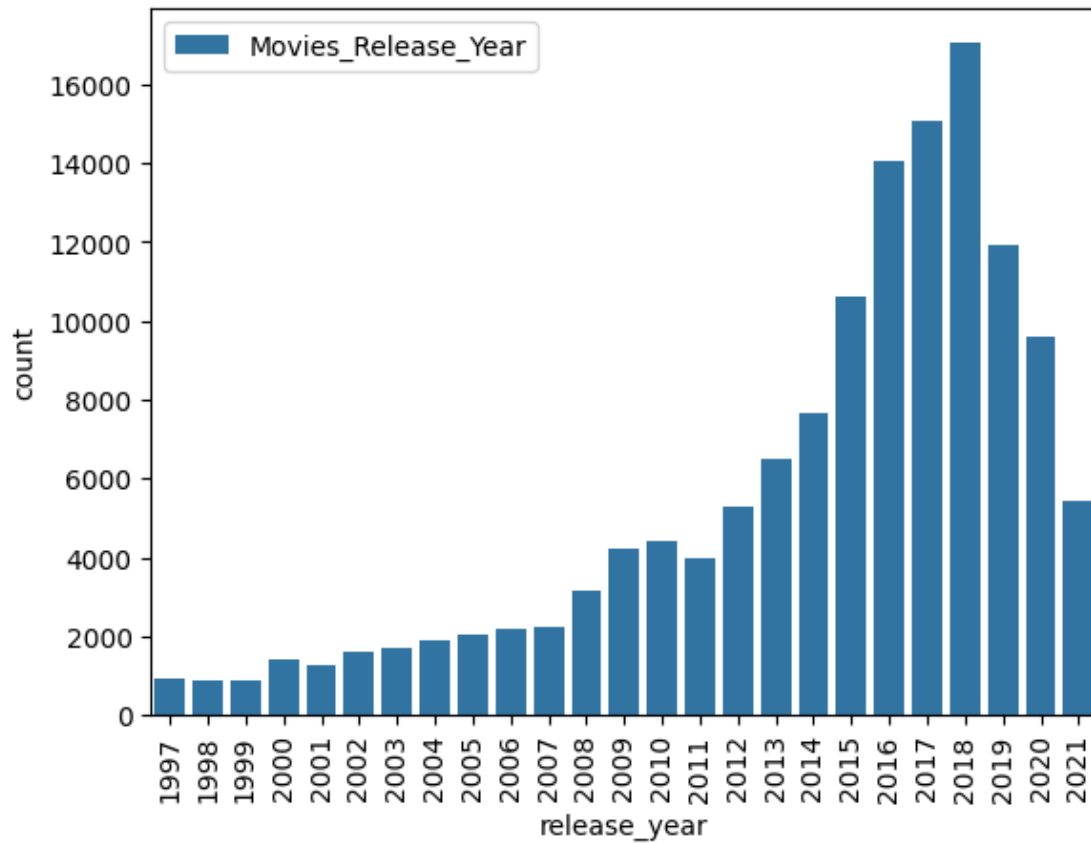


```
[89]: #movies
top_5_rating_movies=df_final_movies.loc[(df_final_movies["rating"].
    ↪isin(top5_rating))]
sns.countplot(data=top_5_rating_movies,x="rating",color="blue")
plt.xticks(rotation=90)
plt.legend(["movies_drating"])
plt.show()
```



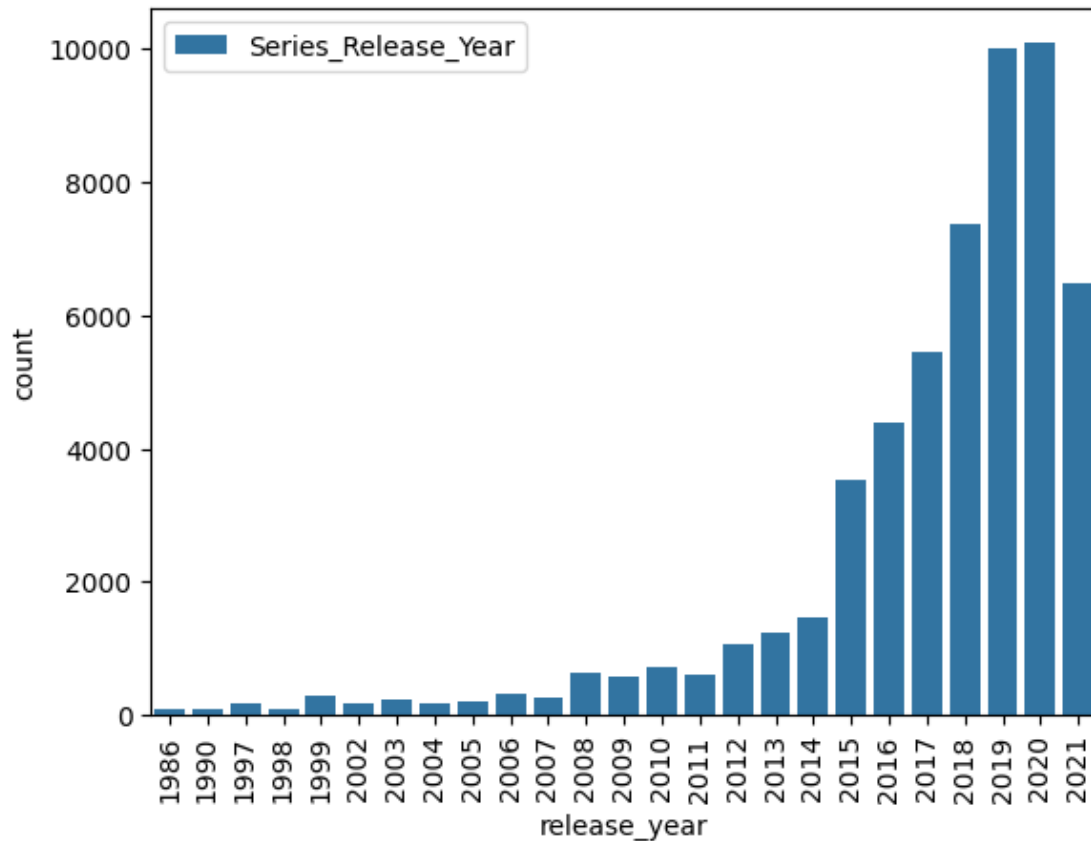
```
[90]: #countplot for release_year
#movies

top_25_releaseyears=df_final_movies.loc[(df_final_movies["release_year"].
    ↪isin(top_25_years))]
sns.countplot(data=top_25_releaseyears,x="release_year")
plt.xticks(rotation=90)
plt.legend(["Movies_Release_Year"])
plt.show()
```



```
[91]: #countplot for release_year
#series

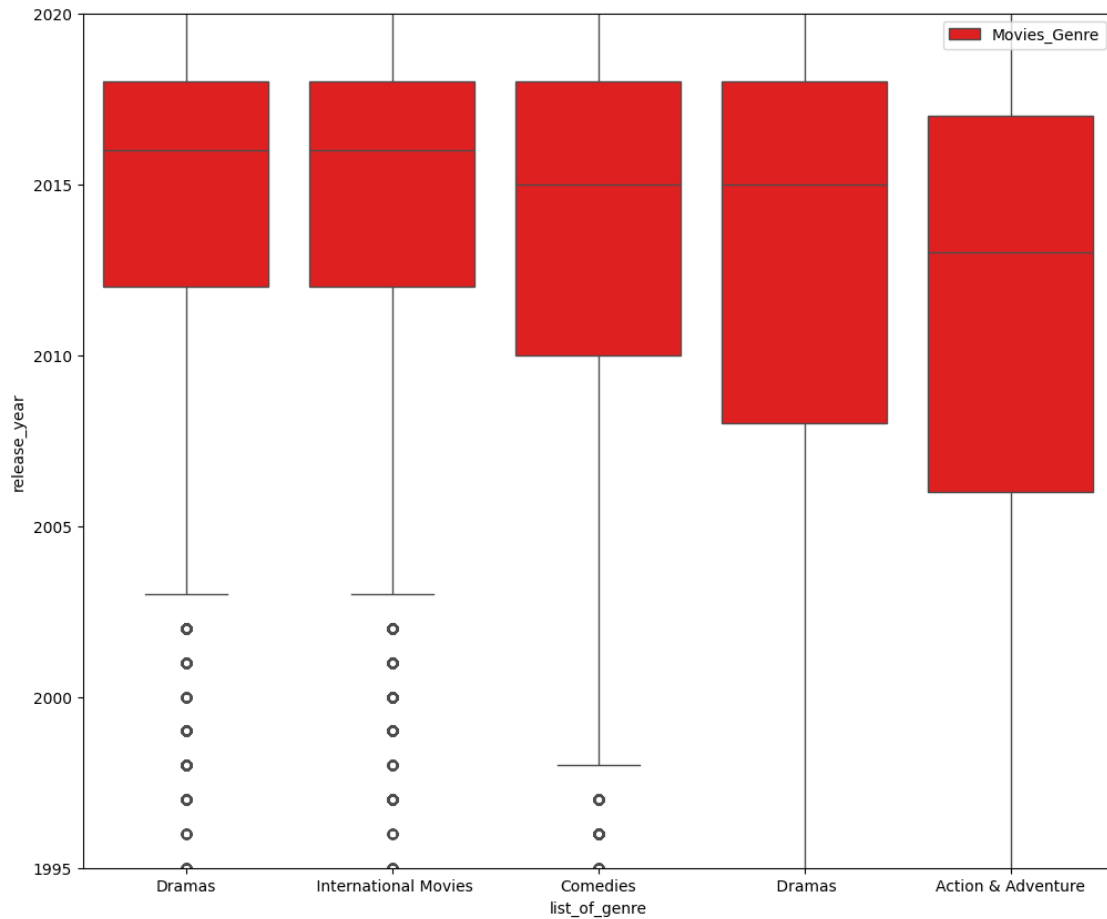
top_25_releaseyears=df_final_series.loc[(df_final_series["release_year"].
    ↪isin(top_25_years1))]
sns.countplot(data=top_25_releaseyears,x="release_year")
plt.xticks(rotation=90)
plt.legend(["Series_Release_Year"])
plt.show()
```

8 For Categorical Variable - Boxplot

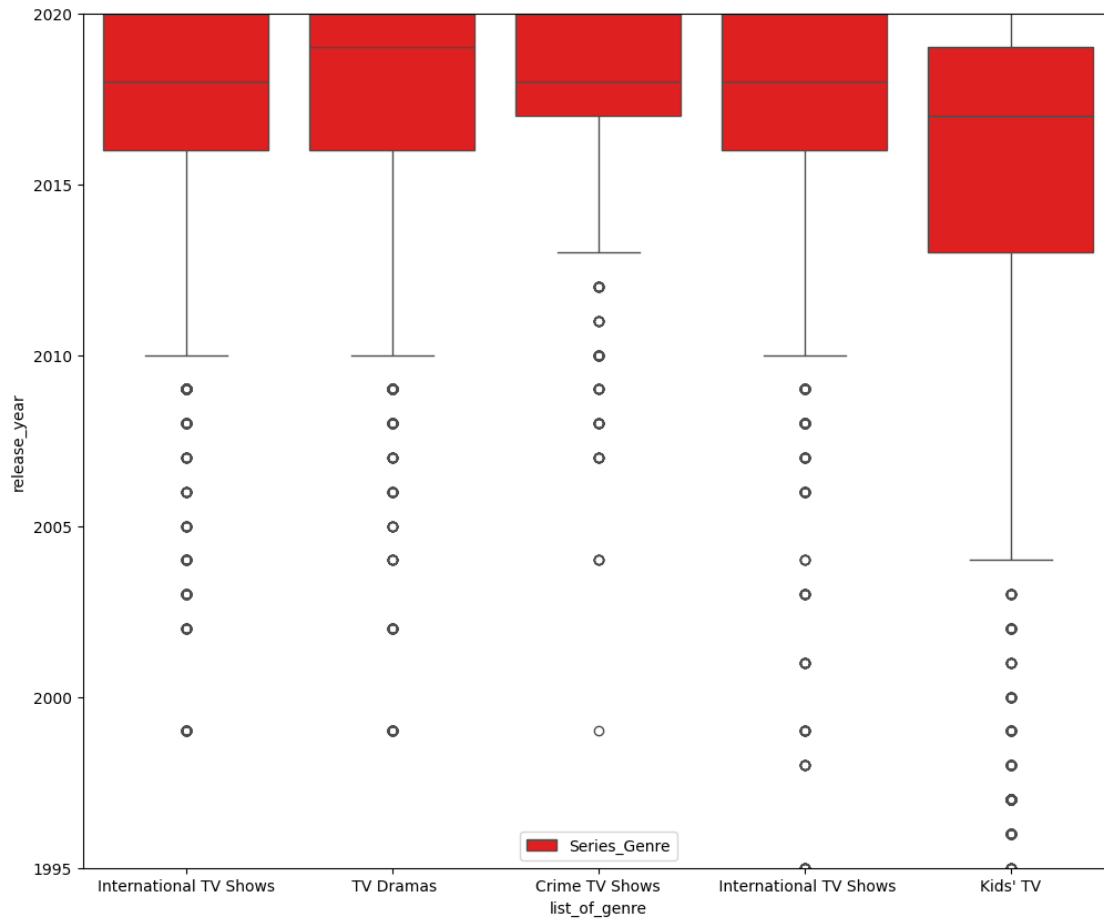
```
[92]: # TOP 5 Genre v/s last 25 years
#Movies

top_5_movies_genre = df_final_movies.loc[(df_final_movies['list_of_genre'].
    ↳isin(top5_genre))]
plt.figure(figsize=(12,10))
sns.
    ↳boxplot(data=top_5_movies_genre,x='list_of_genre',y="release_year",color="r")
plt.ylim(bottom=1995,top=2020)
plt.legend(["Movies_Genre"])
plt.show()
```



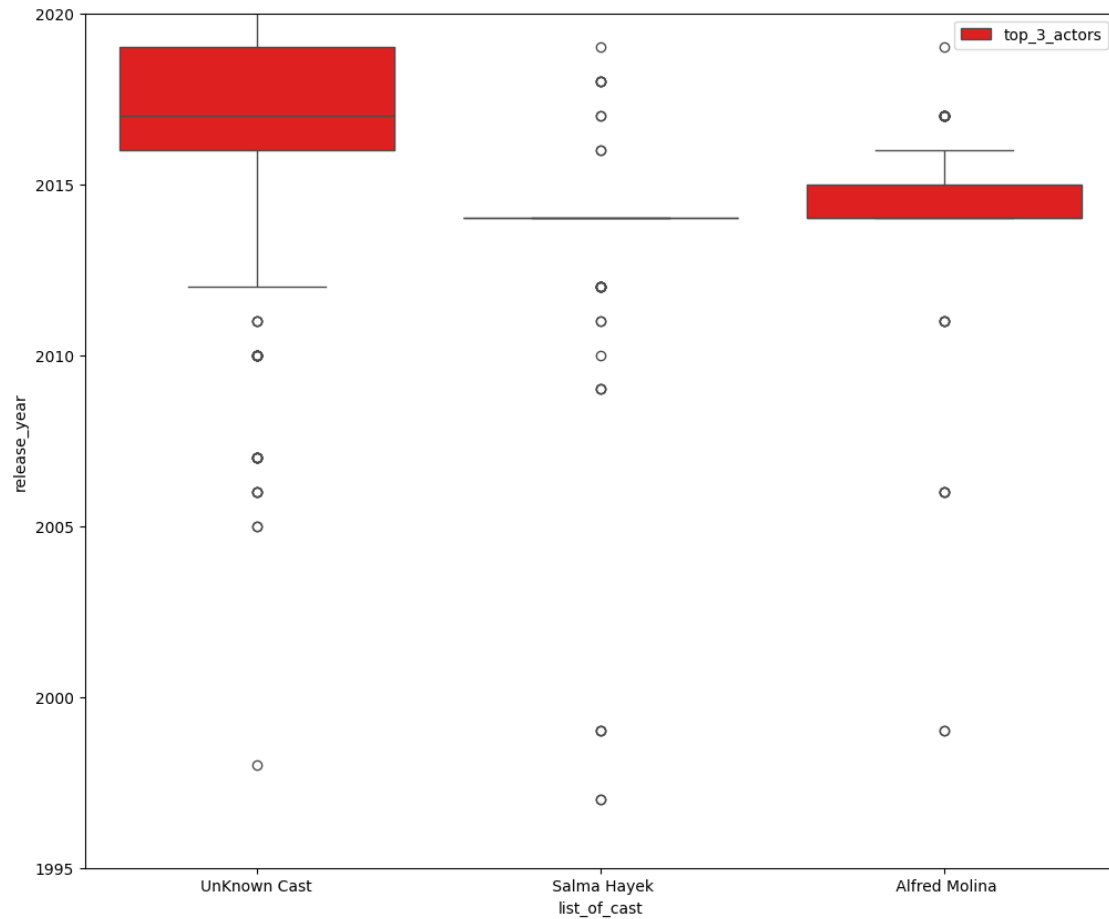
```
[93]: # TOP 5 Genre v/s last 25 years
#Series

top_5_genre_series = df_final_series.loc[(df_final_series['list_of_genre'].
    ↳isin(top5_genre1))]
plt.figure(figsize=(12,10))
sns.
    ↳boxplot(data=top_5_genre_series,x='list_of_genre',y="release_year",color="r")
plt.ylim(bottom=1995,top=2020)
plt.legend(["Series_Genre"])
plt.show()
```



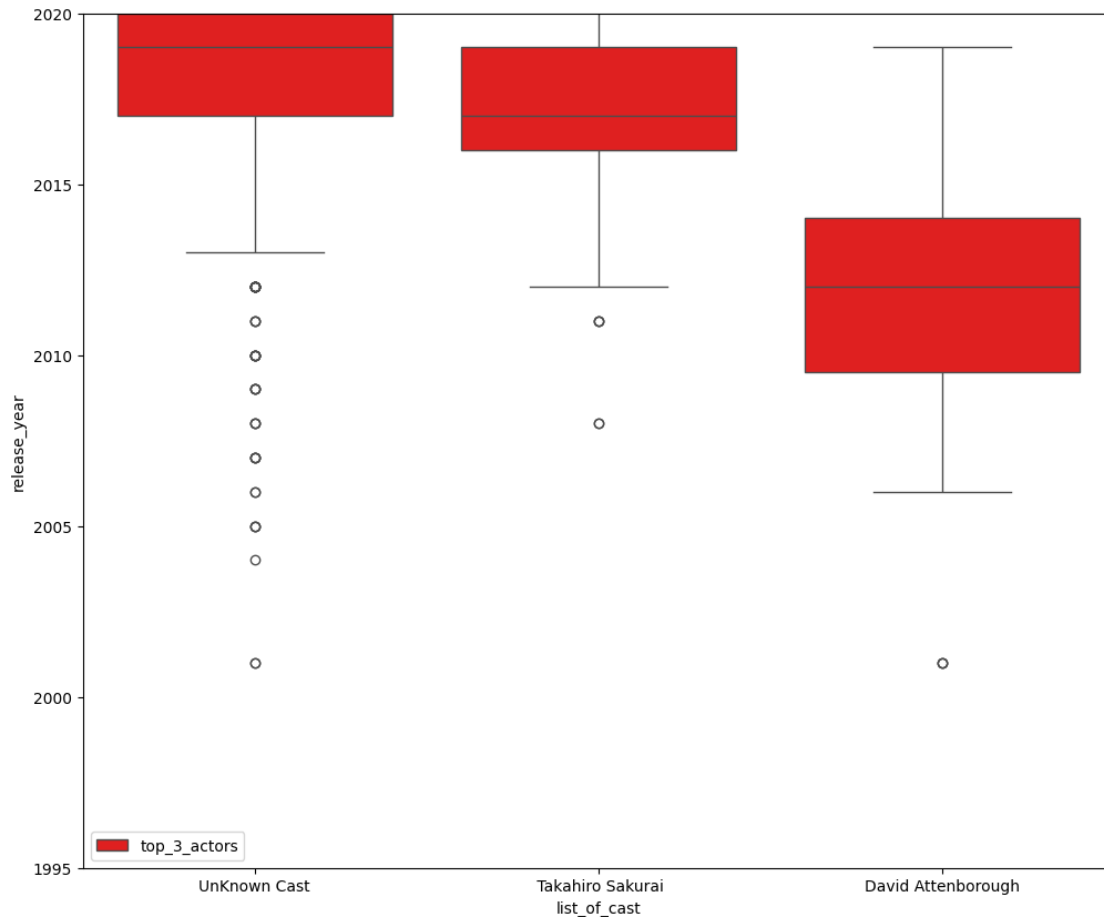
```
[94]: # TOP 3 Actor v/s last 25 years
#Movies

top_3_actors= df_final_movies.loc[(df_final_movies['list_of_cast'].
    ↪isin(top3_casts))]
plt.figure(figsize=(12,10))
sns.boxplot(data=top_3_actors,x='list_of_cast',y="release_year",color="r")
plt.ylim(bottom=1995,top=2020)
plt.legend(["top_3_actors"])
plt.show()
```



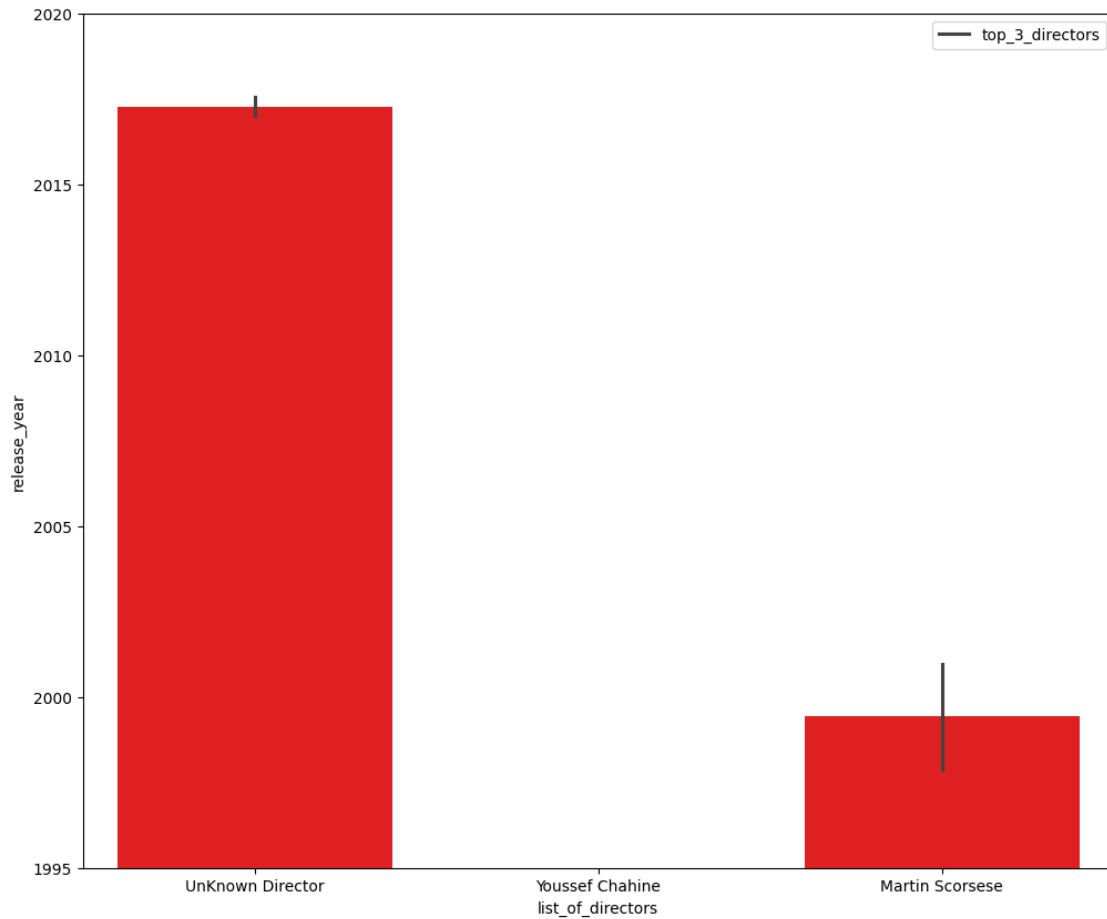
```
[95]: #series

top_3_actors= df_final_series.loc[(df_final_series['list_of_cast'].
    ↪isin(top3_casts1))]
plt.figure(figsize=(12,10))
sns.boxplot(data=top_3_actors,x='list_of_cast',y="release_year",color="r")
plt.ylim(bottom=1995,top=2020)
plt.legend(["top_3_actors"])
plt.show()
```



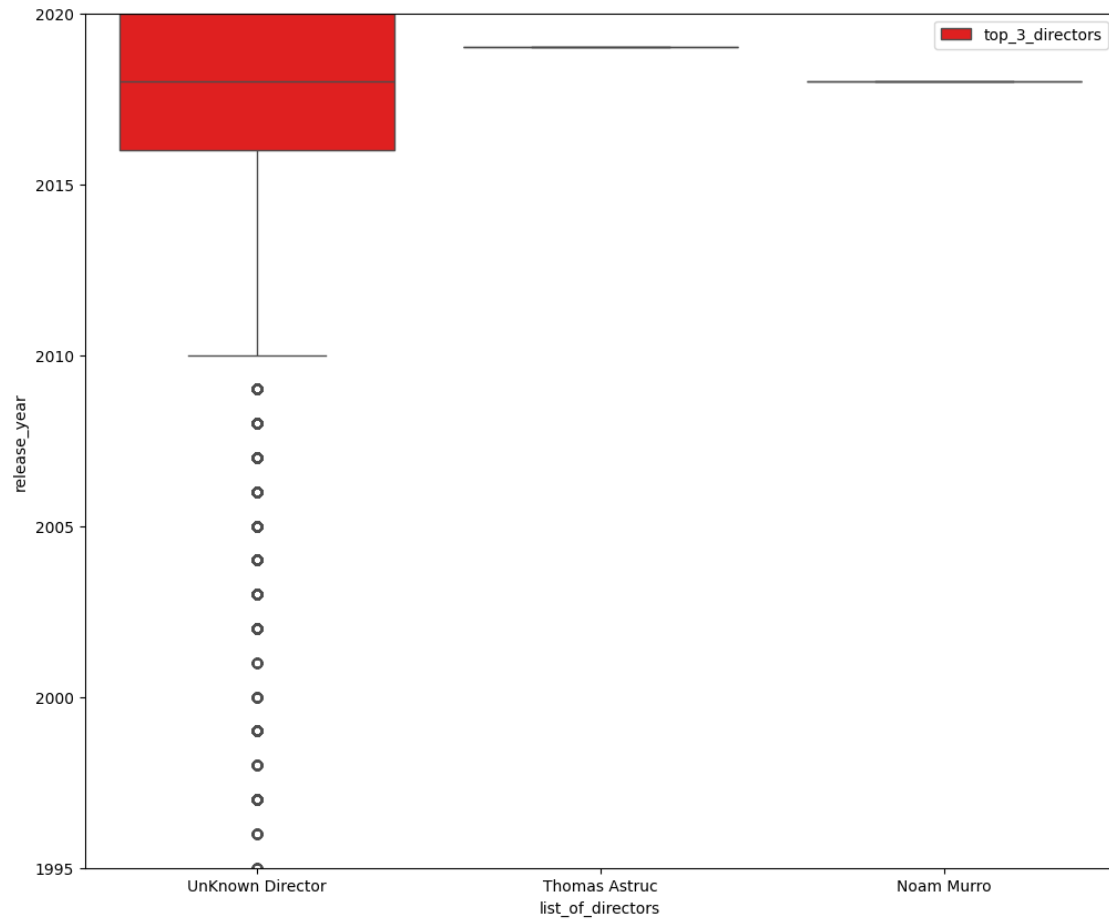
```
[96]: # TOP 3 Director v/s last 25 years
#Movies

top_3_director= df_final_movies.loc[(df_final_movies['list_of_directors'].
    ↳isin(top3_directors))]
plt.figure(figsize=(12,10))
sns.
    ↳barplot(data=top_3_director,x='list_of_directors',y="release_year",color="r")
plt.ylim(bottom=1995,top=2020)
plt.legend(["top_3_directors"])
plt.show()
```



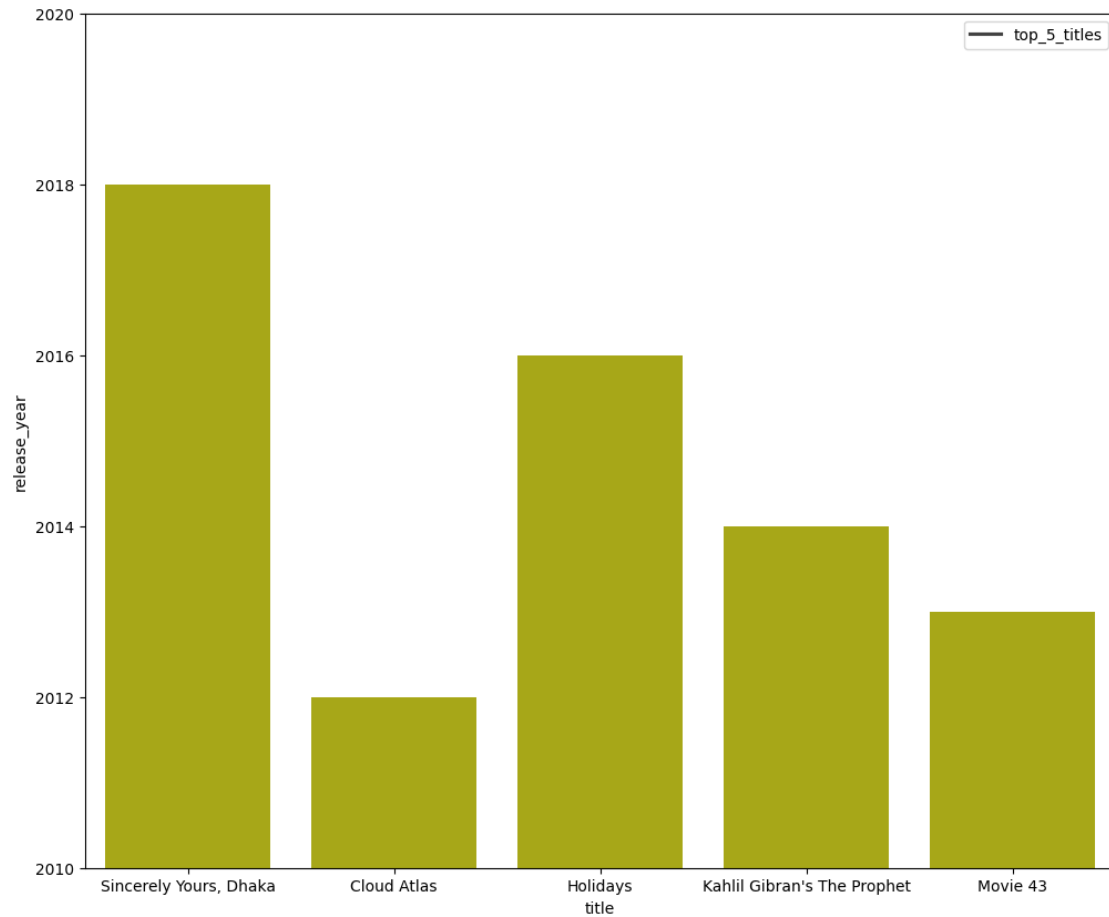
```
[97]: #series

top_3_directors= df_final_series.loc[(df_final_series['list_of_directors'].
    ↳isin(top3_directors1))]
plt.figure(figsize=(12,10))
sns.
    ↳boxplot(data=top_3_directors,x='list_of_directors',y="release_year",color="r")
plt.ylim(bottom=1995,top=2020)
plt.legend(["top_3_directors"])
plt.show()
```

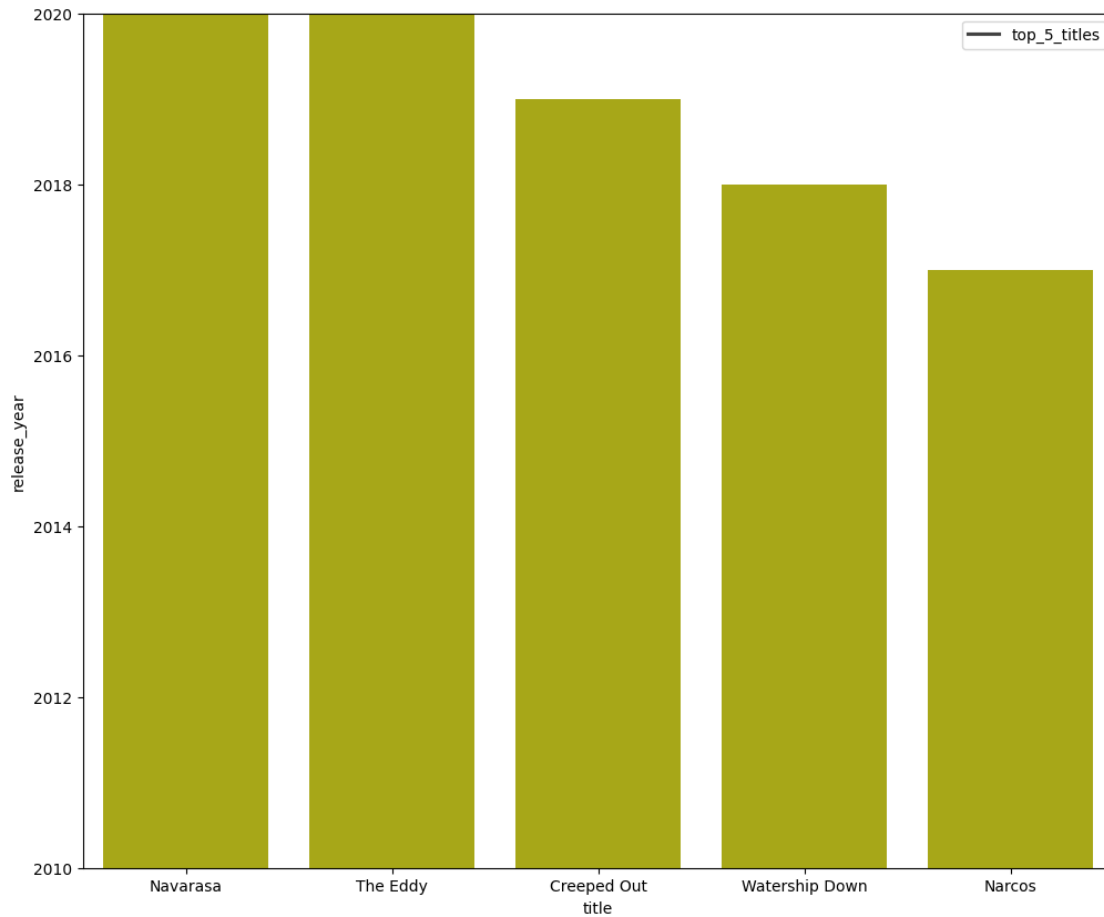


```
[98]: ## top Titles v/s last 10 years
#Movies

top_5_titles = df_final_movies.loc[(df_final_movies['title'].isin(top5_title))]
plt.figure(figsize=(12,10))
sns.barplot(data=top_5_titles,x='title',y="release_year",color="y")
plt.ylim(bottom=2010,top=2020)
plt.legend(["top_5_titles"])
plt.show()
```



```
[99]: top_5_title = df_final_series.loc[(df_final_series['title'].isin(top5_title1))]
plt.figure(figsize=(12,10))
sns.barplot(data=top_5_title,x='title',y="release_year",color="y")
plt.ylim(bottom=2010,top=2020)
plt.legend(["top_5_titles"])
plt.show()
```

9 For correlation: Pairplots

```
[129]: top_data_actors_directors=df_final_movies.loc[(df_final_movies["list_of_cast"].
↳isin(top3_casts)) & (df_final_movies["list_of_directors"].
↳isin(top3_directors))]
top_data_actors_directors.sample(5)
```

```
[129]:
```

	title	list_of_cast	\
18760	Trese After Dark	UnKnown Cast	
53167	Little Singham: Legend of Dugabakka	UnKnown Cast	
133751	Virunga: Gorillas in Peril	UnKnown Cast	
18759	Trese After Dark	UnKnown Cast	
80696	The Road to El Camino: Behind the Scenes of El...	UnKnown Cast	

	list_of_country	list_of_genre	list_of_directors	show_id	\
18760	Unknown Country	Documentaries	UnKnown Director	s739	
53167	Unknown Country	Children & Family Movies	UnKnown Director	s2237	

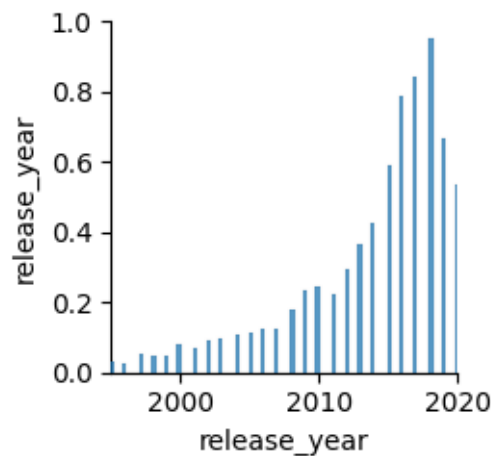
133751	United States	Documentaries	UnKnown Director	s5919
18759	Unknown Country	Anime Features	UnKnown Director	s739
80696	United States	Documentaries	UnKnown Director	s3365

	type	release_year	rating	duration	\
18760	Movie	2021	TV-14	36 min	
53167	Movie	2020	TV-Y7	68 min	
133751	Movie	2015	TV-14	29 min	
18759	Movie	2021	TV-14	36 min	
80696	Movie	2019	TV-14	14 min	

		description	month	year
18760	Stars and creators gather to discuss	"Trese," ...	June 11	2021
53167	As fears rise over a ghost known for nighttime...		July 15	2020
133751	The team behind the searing documentary reveal...		February 13	2015
18759	Stars and creators gather to discuss "Trese," ...		June 11	2021
80696	Aaron Paul, Vince Gilligan and other cast memb...		October 29	2019

```
[101]: plt.figure(figsize=(20,15))
sns.pairplot(data=df_final_movies)
plt.xlim(left=1995,right=2020)
plt.show()
```

<Figure size 2000x1500 with 0 Axes>



10 Insights based on Non-Graphical and Visual Analysis:

- Comments on the range of attributes
1. Based on the given data, we observe that there are wide variety of Genres present nowadays which give many content to users.

2. The platform helps the user to display the most watched shows.

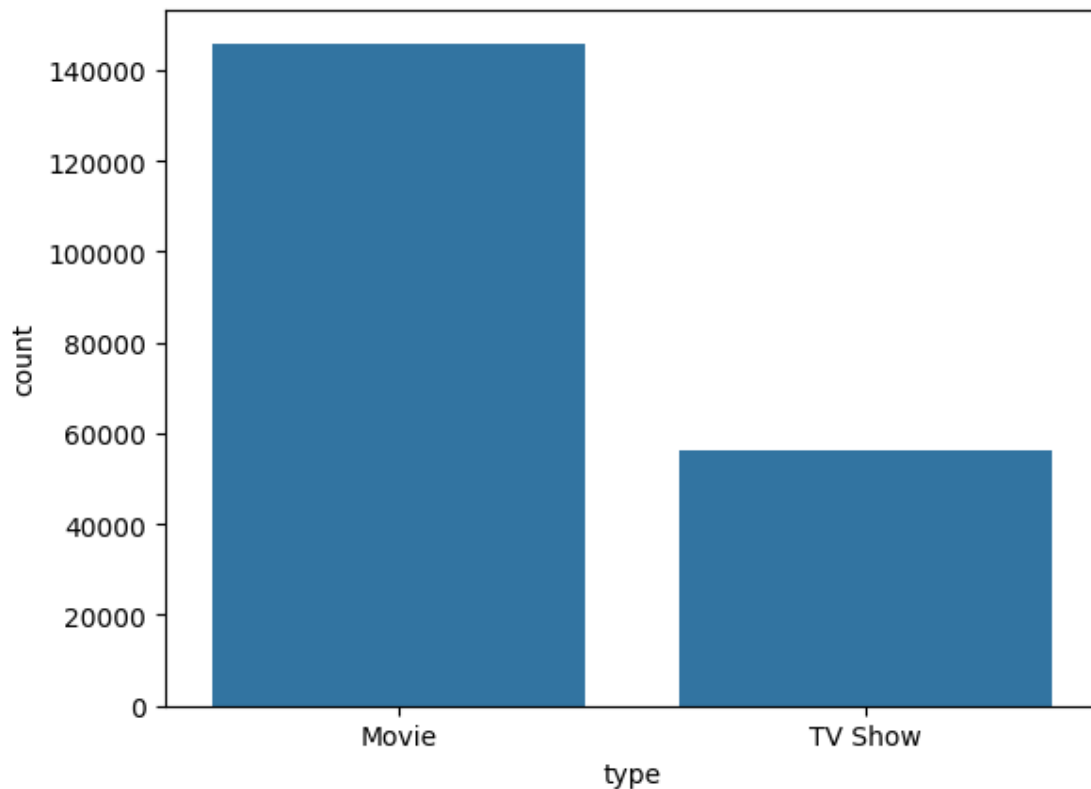
- Comment on Univariate and Bivariate Analysis:

1. The release of movies on OTT platform like Netflix have significantly increased over years, there were downfall during 2020-2022, the factors could be external like covid.

2. For the duration that the movies people like to watch movies duration between 90 to 100 minutes time interval and for tv shows people has high perferance for shows having 1 or 2 seasons.

11 A few major Business Insights:

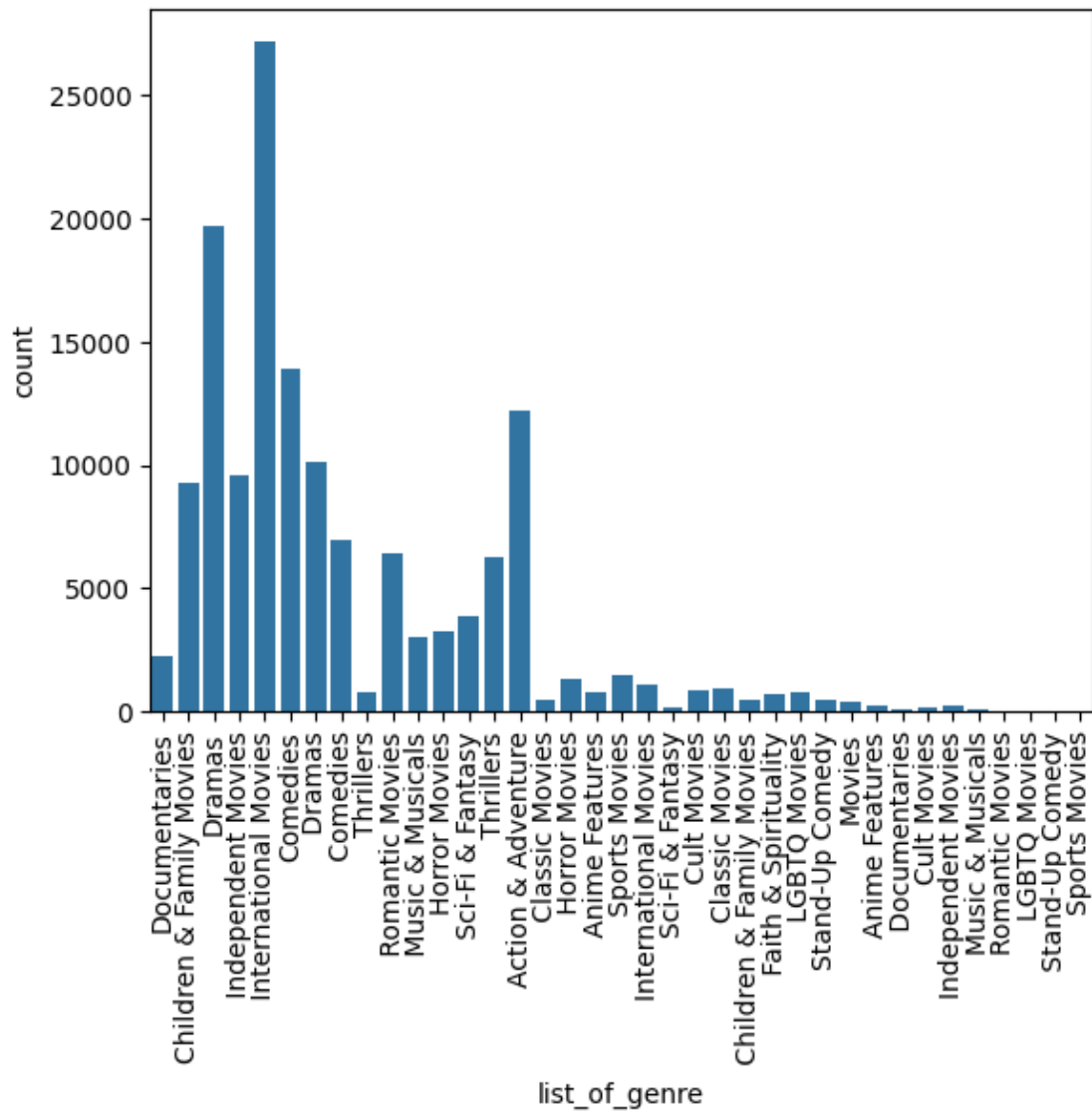
```
[102]: sns.countplot(data=df4,x='type')
plt.show()
### Insight: The graph shows that the demands of movie is higher than the TV_
↳Shows
```

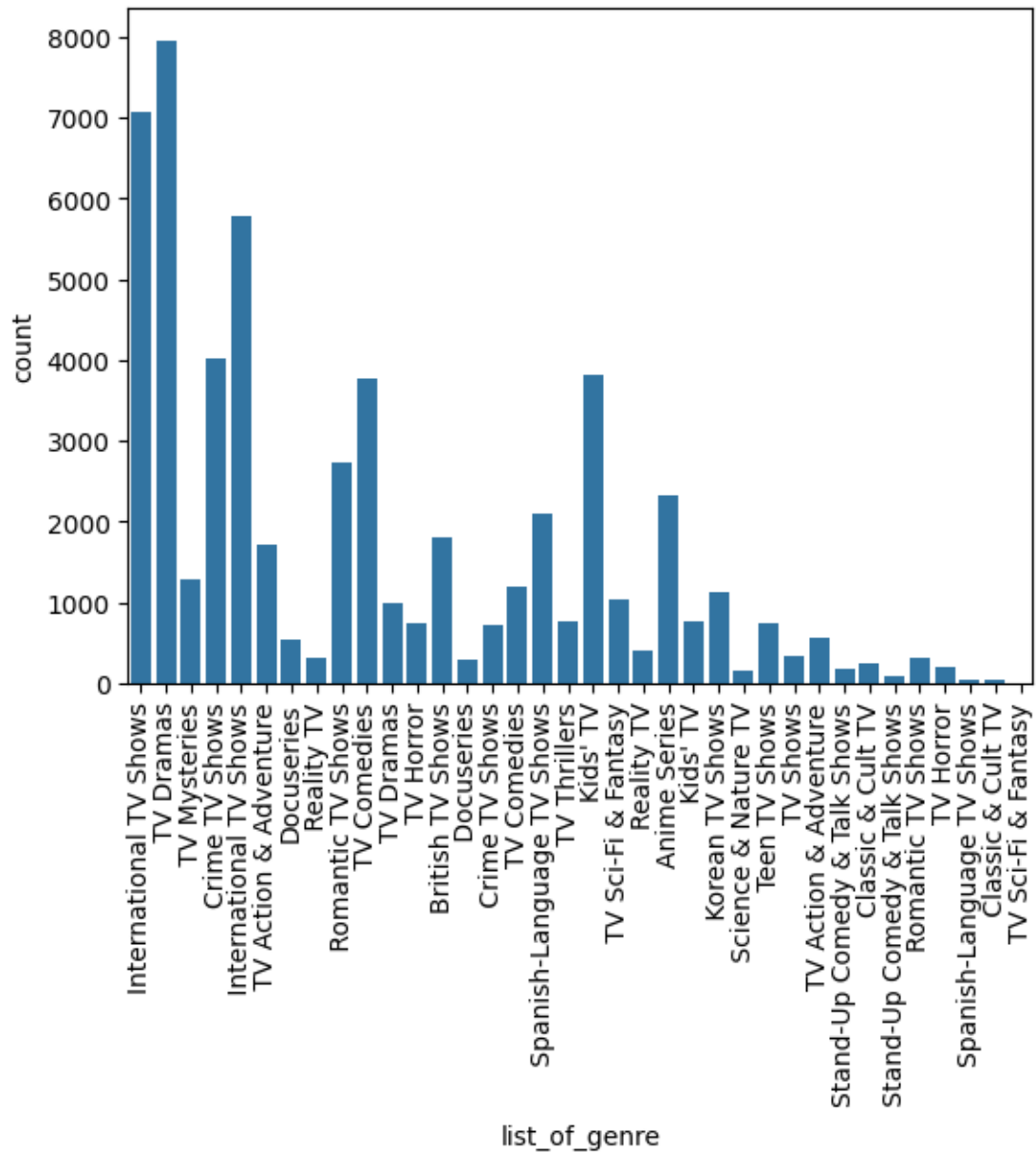


```
[130]: #movies
sns.countplot(data=df_final_movies,x='list_of_genre')
plt.xticks(rotation=90)
plt.show()
```

```
#genre
sns.countplot(data=df_final_series,x='list_of_genre')
plt.xticks(rotation=90)
plt.show()
```

Insights: The most watched genre for TV shows and Movies are, Dramas,
↳ International TV shows, International movies.





```
[107]: #top5genre
#movies
top_5_data_Genres=df_final_movies.loc[(df_final_movies["list_of_genre"].
    ↪isin(top5_genre))]
sns.countplot(data=top_5_data_Genres,x="list_of_genre",color="y")
plt.xticks(rotation=90)
plt.legend(["movies"])
plt.show()

#series
```

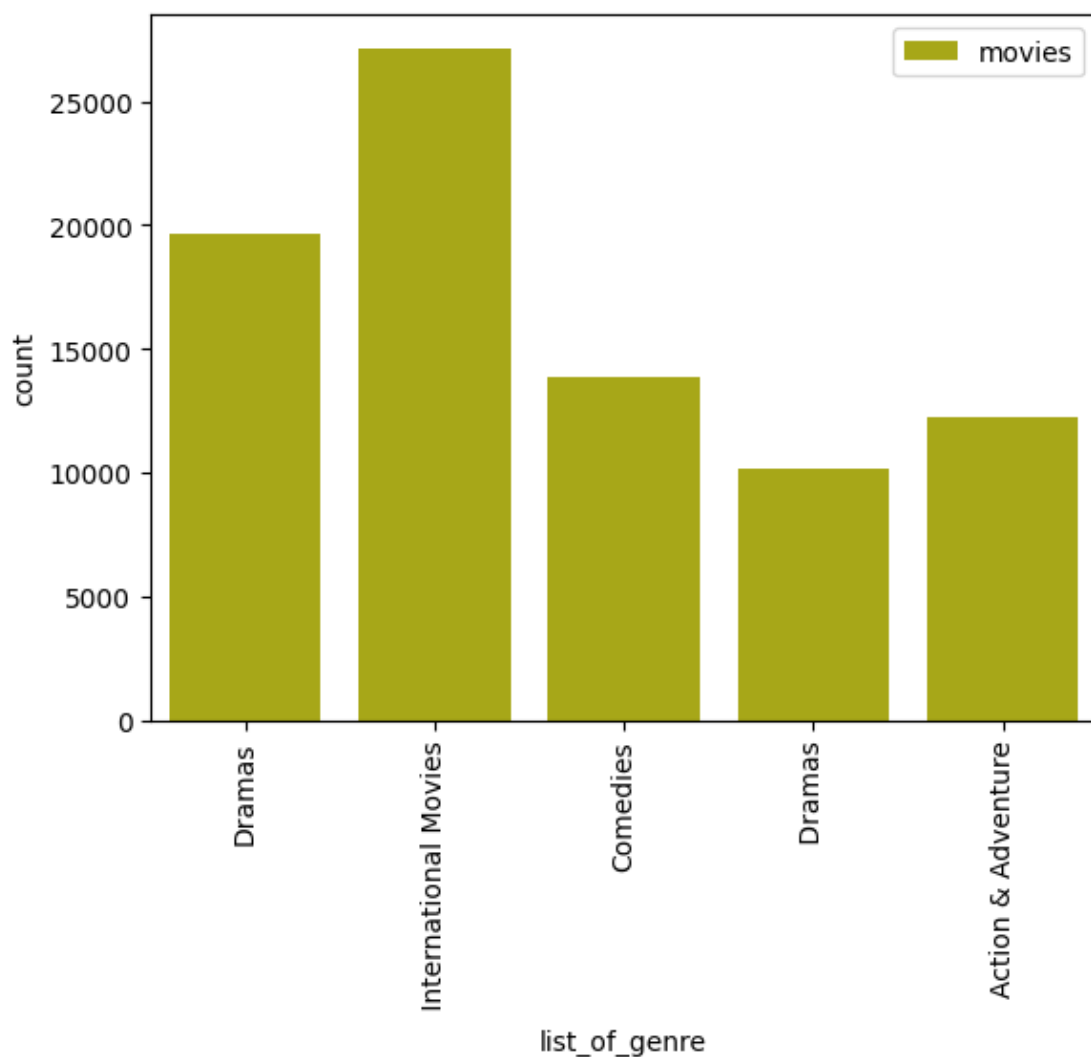
```

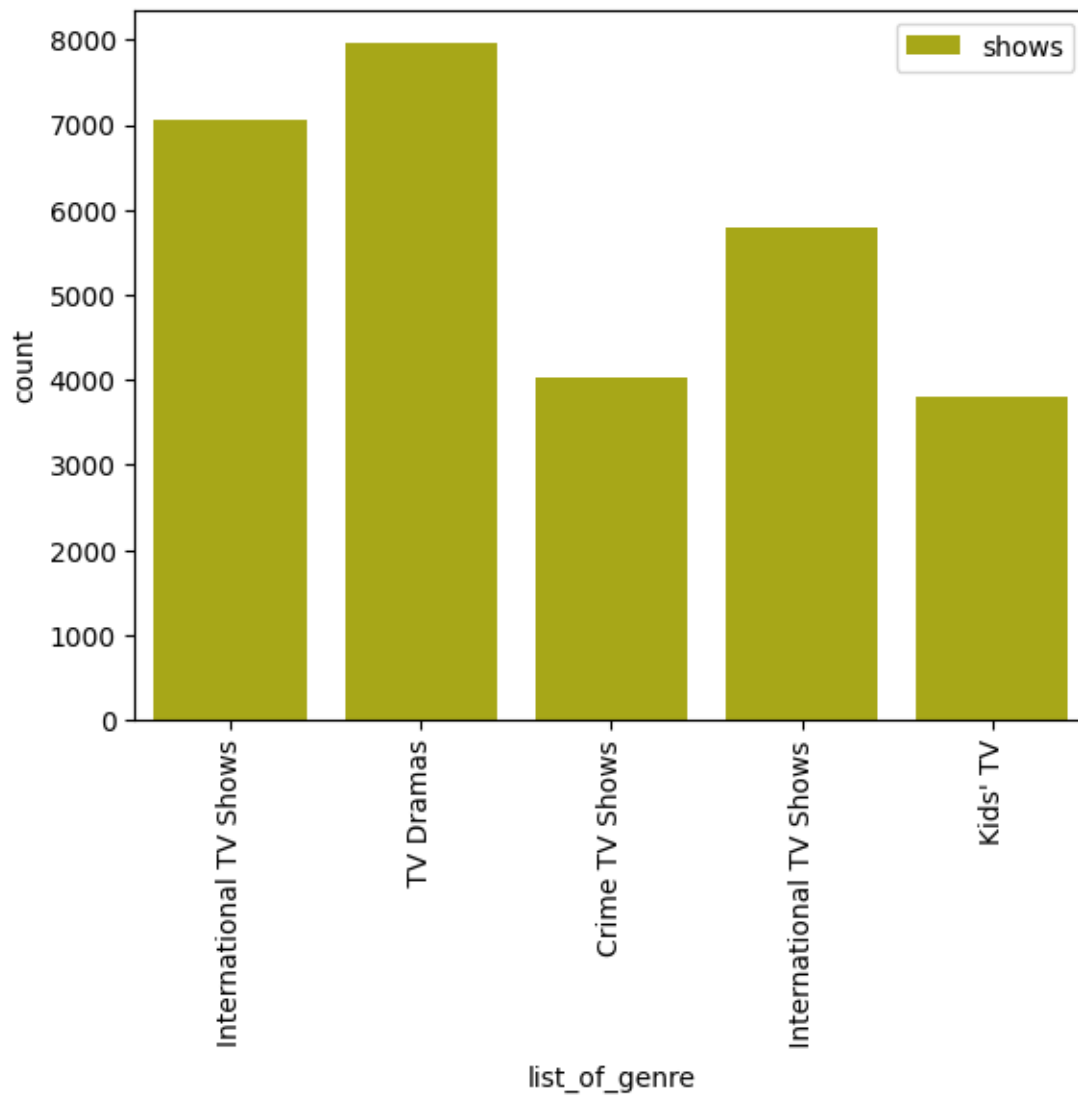
top_5_data_Genres=df_final_series.loc[(df_final_series["list_of_genre"].
    ↳isin(top5_genre1))]
sns.countplot(data=top_5_data_Genres,x="list_of_genre",color="y")
plt.xticks(rotation=90)
plt.legend(["shows"])
plt.show()

```

Insights : From the below graph we can conclude that the Dramas,
 ↳International Movies, Comedies are genre from Movies which are more popular,
 ↳among the

Netflix users and among TV shows we have international TV shows, TV dramas,
 ↳crime Tv shows in the most popular demand.

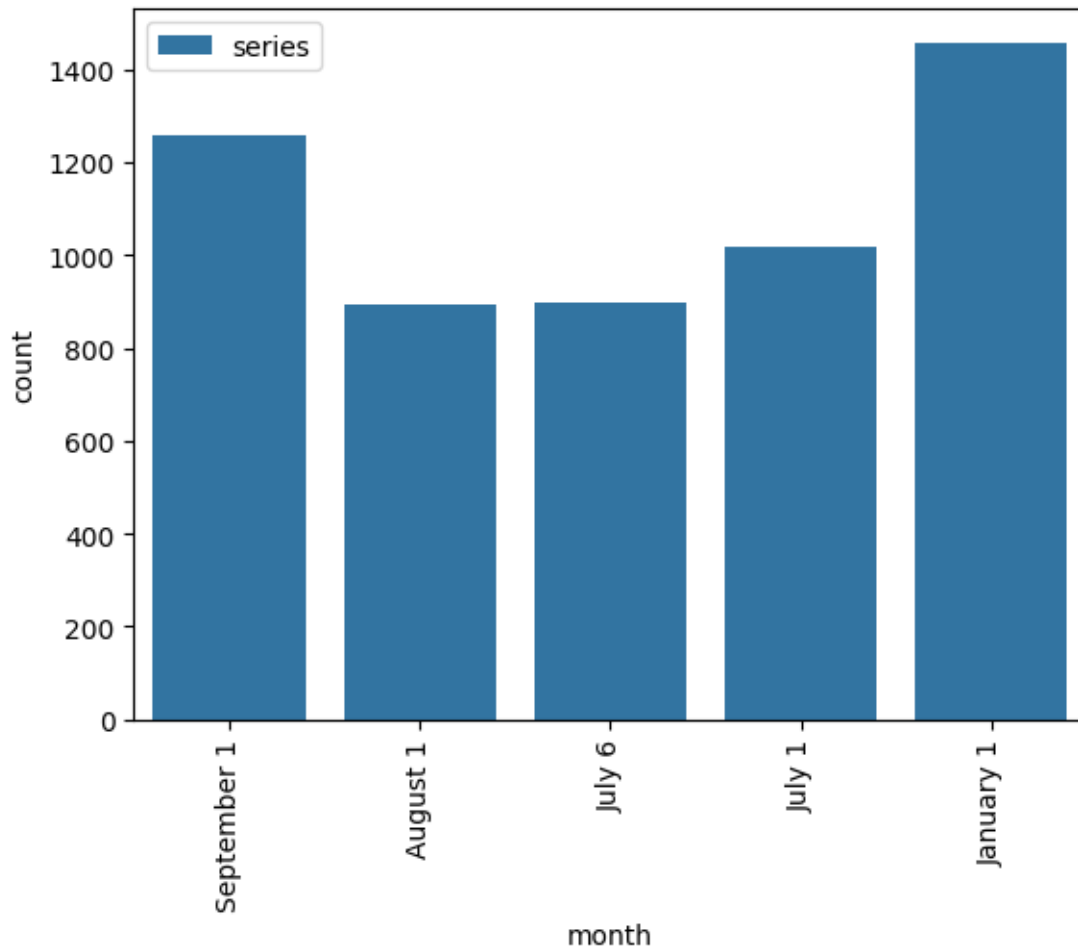




```
[108]: #top5_months

top5_month = df_final_series.loc[(df_final_series['month'].isin(top5_months1))]
sns.countplot(data=top5_month,x="month")
plt.xticks(rotation=90)
plt.legend(["series"])
plt.show()

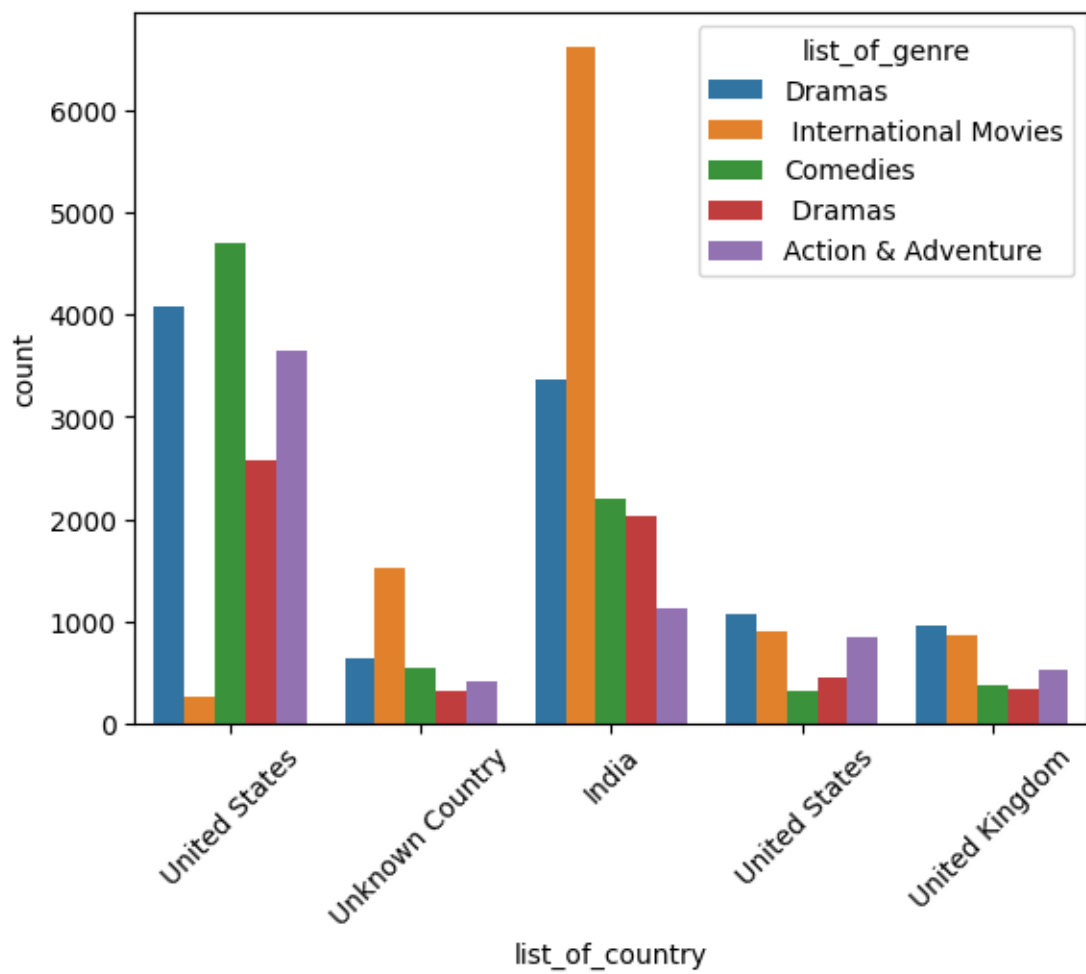
### Insight: It can be inferred from the data that series released in
↳ Sept, January, July are most likely to be watched.
```

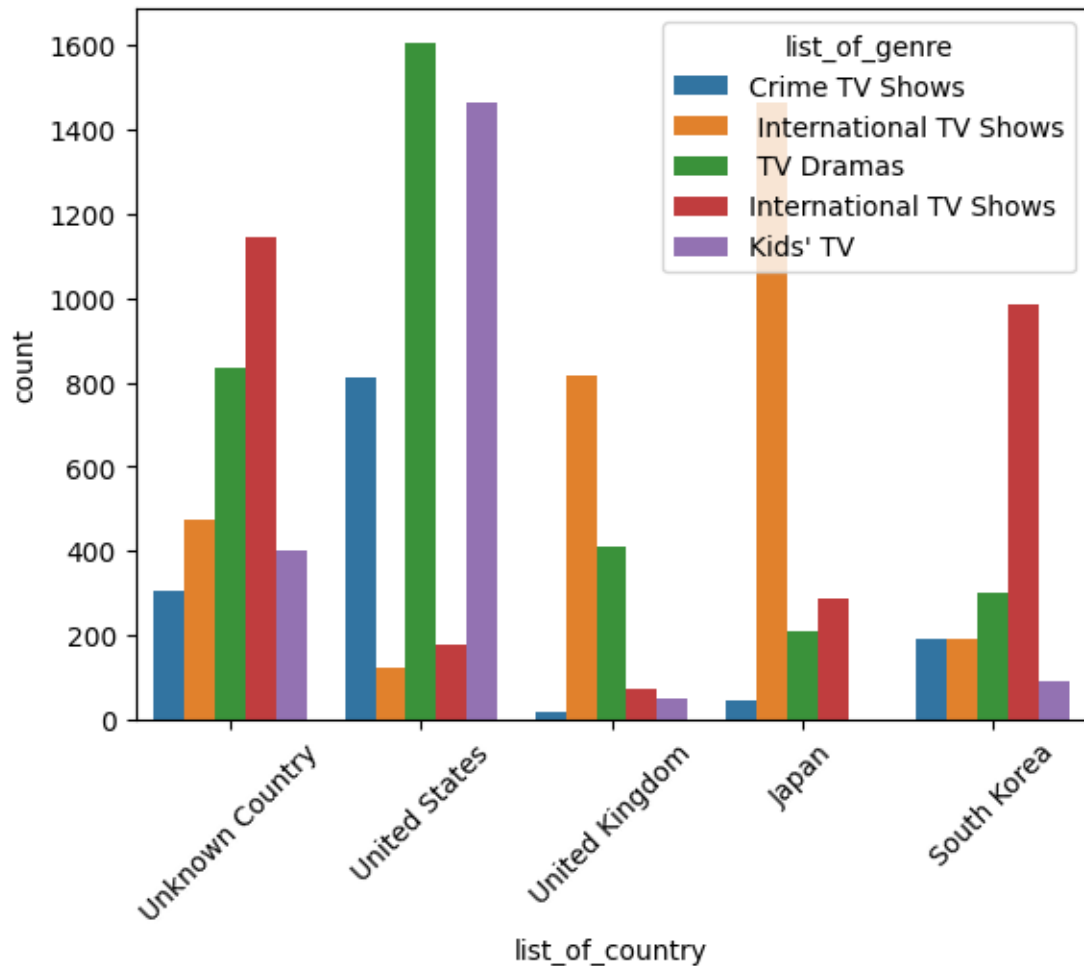


```
[131]: #top5movies
top_5_genre_country = df_final_movies.loc[(df_final_movies['list_of_genre'].
    ↳isin(top5_genre))&(df_final_movies['list_of_country'].isin(top5_country))]
sns.countplot(data=top_5_genre_country,x="list_of_country",hue="list_of_genre")
plt.xticks(rotation=45)
plt.show()

#top5_country
top_5_genre_country = df_final_series.loc[(df_final_series['list_of_genre'].
    ↳isin(top5_genre1))&(df_final_series['list_of_country'].isin(top5_country1))]
sns.countplot(data=top_5_genre_country,x="list_of_country",hue="list_of_genre")
plt.xticks(rotation=45)
plt.show()

### The top5 genre for movies watched in the different countries are : drama,
    ↳comedy action and that in series are International shows.
```



```
[ ]: # Recommendations and actionable Insights:

# 1. The preference is given more towards comedy and drama movies netlix can
    ↳ enhance the business into the
#    same genre by prioritizing the creation and promotion of content

# 2. Collaborating with renowned directors such as Youssef Chahine, Martin
    ↳ Scorsese, and Cathy Garcia for movies,
#    and Houda Benyamina, Thomas Astruc, and Noam Murro for TV shows, could
    ↳ further enhance the quality and diversity of Netflix's content.

#3. To optimize content consumption, Netflix should focus on the duration of
    ↳ movies within the range of 90-110 minutes
#    and limit TV show seasons to 1-2, aligning with observed preferences for
    ↳ shorter durations.
```

#4. Netflix should focus on sourcing and promoting, content from countries such as the United States,
 # the United Kingdom, Japan, and India, recognizing their significance as key markets with consistent viewing preferences.

#5. They could enhance the delivery of content in the peak months like that in January,
 # July, and October, capitalizing on observed trends in user engagement during these periods.

#6. Considering the success factors of similar platforms like Hotstar,jiocinema, Netflix should also analyze the content strategies of competitors to identify regional focuses and tailor its content offerings accordingly
 # for better growth and market penetration, in India, where cutomers still look forward to the regional content.

#7. Netflix should consider prioritizing the production of more movie content over TV shows based on the observation that,
 # over the past 25 years, movie ratings have remained consistently level.

[]:

[]: