

PROGRAM CODE

IN PYTHON

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
from wordcloud import WordCloud, STOPWORDS
from IPython.display import Image
import warnings
warnings.filterwarnings("ignore")
```

```
data = pd.read_csv("/content/netflix_titles.csv")
netflix_df = pd.read_csv("/content/netflix_titles_original.csv")
netflix_df.head(10)
```

→

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	25-Sep-21	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	24-Sep-21	2021	TV-MA	2 Seasons	Dramas	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabl...	NaN	24-Sep-21	2021	TV-MA	1 Season	Crime	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	24-Sep-21	2021	TV-MA	1 Season	Documentaries	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	24-Sep-21	2021	TV-MA	2 Seasons	Romantic	In a city of coaching centers known to train l...
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	NaN	24-Sep-21	2021	TV-MA	1 Season	Horror	The arrival of a charismatic young priest brin...
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	24-Sep-21	2021	PG	91 min	Family	Equestria's divided. But a bright-eyed hero be...
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	24-Sep-21	1993	TV-MA	125 min	Dramas	On a photo shoot in Ghana, an American model s...
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...	United Kingdom	24-Sep-21	2021	TV-14	9 Seasons	Documentaries	A talented batch of amateur bakers face off in...
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	24-Sep-21	2021	PG-13	104 min	Dramas	A woman adjusting to life after a loss contend...

```
netflix_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       8789 non-null   object
11  description     8807 non-null   object
dtypes: int64(1), object(11)
memory usage: 825.8+ KB
```

```
netflix_df.shape
```

➔
(8807, 12)

```
netflix_df.columns
```

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

```
netflix_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	8789 non-null	object
11	description	8807 non-null	object

dtypes: int64(1), object(11)
memory usage: 825.8+ KB

```
netflix_df.describe()
```

➔

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

▾ Checking if there is any Null Value is present in the dataset or not ?

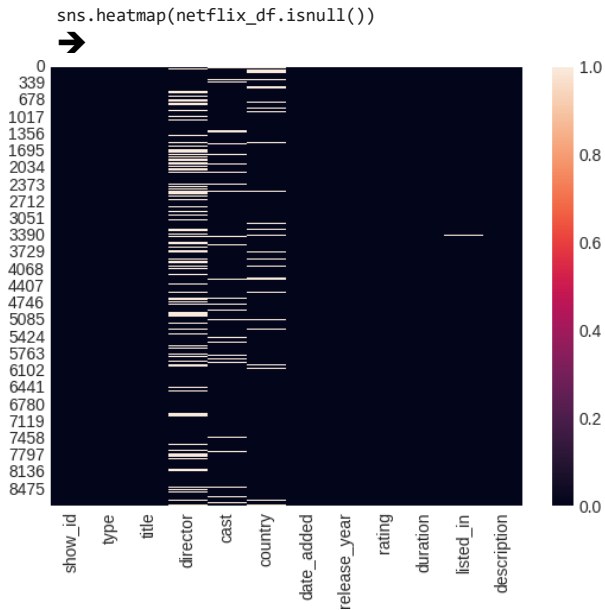
```
netflix_df.isnull().sum()
```

➔

show_id	0
type	0
title	0
director	2634
cast	825
country	831
date_added	10
release_year	0
rating	4
duration	3
listed_in	18
description	0

dtype: int64

Plotting Heatmap for null values



▾ Checking if there any duplicate value present in the dataset or not?

```
netflix_df.duplicated().sum()
```

→ 0

▾ Checking unique values

```
netflix_df["type"].unique()
```

→

```
array(['Movie', 'TV Show'], dtype=object)
```

```
netflix_df["rating"].unique()
```

→

```
array(['PG-13', 'TV-MA', 'PG', 'TV-14', 'TV-PG', 'TV-Y', 'TV-Y7', 'R',  
      'TV-G', 'G', 'NC-17', '74 min', '84 min', '66 min', 'NR', nan,  
      'TV-Y7-FV', 'UR'], dtype=object)
```

```
netflix_df["release_year"].unique()
```

→

```
array([2020, 2021, 1993, 2018, 1996, 1998, 1997, 2010, 2013, 2017, 1975,  
      1978, 1983, 1987, 2012, 2001, 2014, 2002, 2003, 2004, 2011, 2008,  
      2009, 2007, 2005, 2006, 1994, 2015, 2019, 2016, 1982, 1989, 1990,  
      1991, 1999, 1986, 1992, 1984, 1980, 1961, 2000, 1995, 1985, 1976,  
      1959, 1988, 1981, 1972, 1964, 1945, 1954, 1979, 1958, 1956, 1963,  
      1970, 1973, 1925, 1974, 1960, 1966, 1971, 1962, 1969, 1977, 1967,  
      1968, 1965, 1946, 1942, 1955, 1944, 1947, 1943])
```

```
data["listed_in"].nunique()
```

→

35

```
netflix_df["type"].value_counts()
```

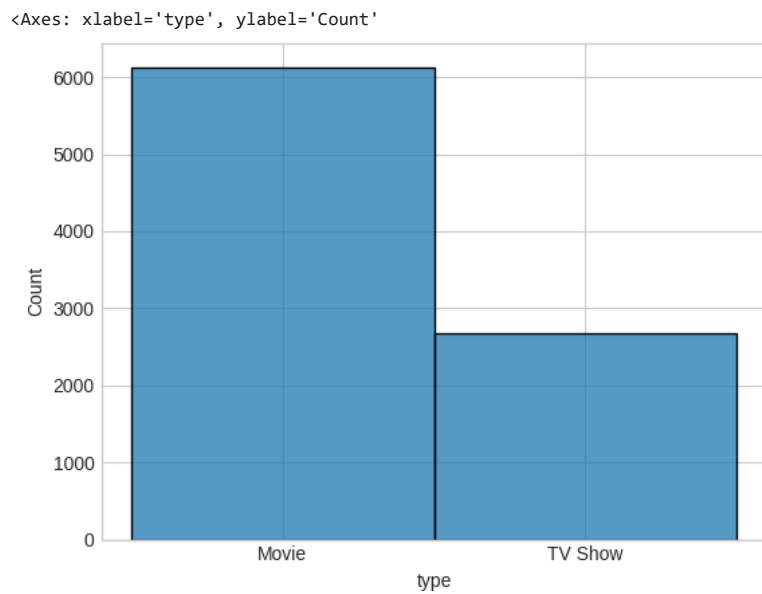
→

```
Movie      6131  
TV Show    2676  
Name: type, dtype: int64
```

▼ Checking no. of movies and no. of tv shows

```
sns.histplot(netflix_df["type"])
```

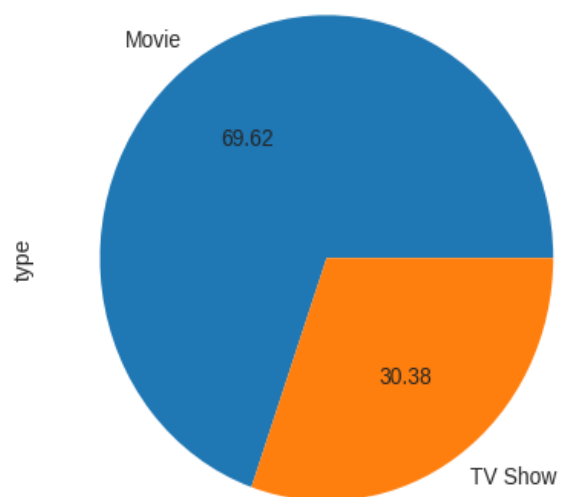
→



```
netflix_df["type"].value_counts().plot(kind = "pie", autopct = "%.2f")
```

→

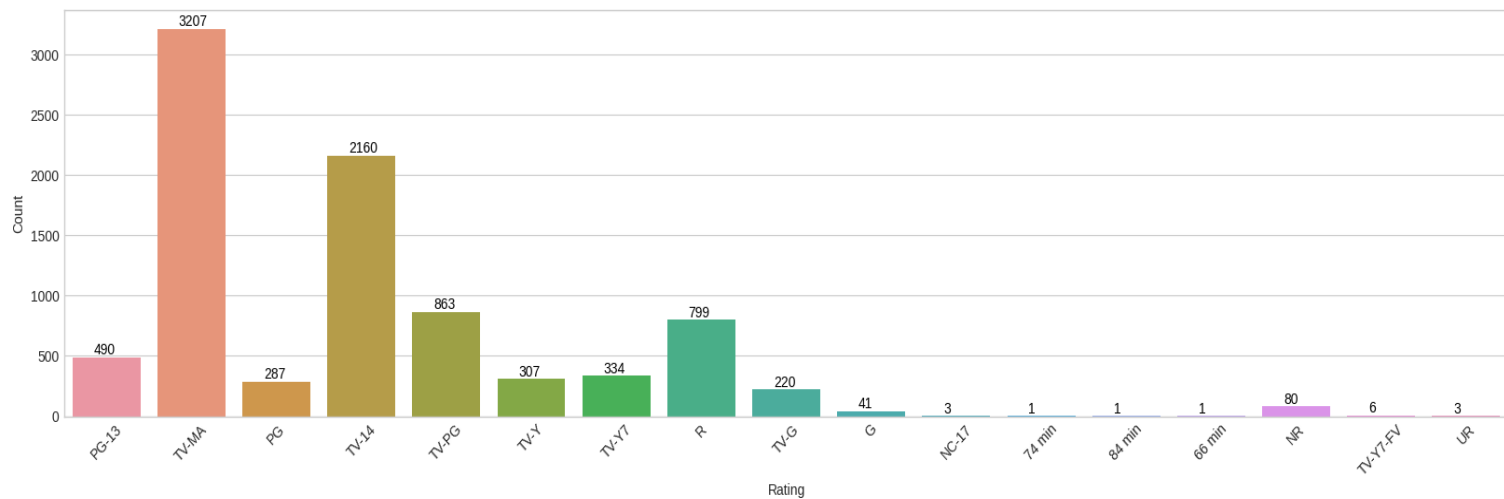
<Axes: ylabel='type'>



▼ Checking no. of ratings in each category for movies and tv shows

```
plt.figure(figsize = (20, 5))
plt.style.use('seaborn-whitegrid')
ax = sns.countplot(x = 'rating', data = netflix_df)
plt.xlabel('Rating')
plt.ylabel('Count')
plt.xticks(rotation = 45)
for p in ax.patches:
    ax.annotate(int(p.get_height()), (p.get_x()+0.25, p.get_height()+1), va = 'bottom', color = 'black')
```

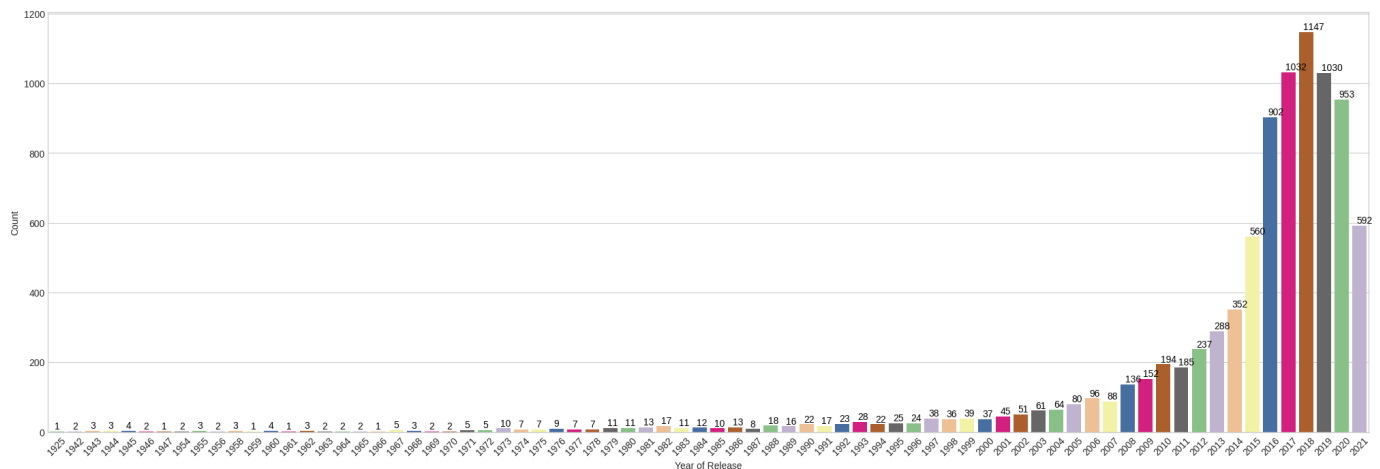
→



Checking how many Content release each year

```
plt.figure(figsize = (25, 8))
plt.style.use('seaborn-whitegrid')
ax = sns.countplot(x = 'release_year', data = netflix_df, palette = "Accent")
plt.xlabel('Year of Release')
plt.ylabel('Count')
plt.xticks(rotation = 45)
for p in ax.patches:
    ax.annotate(int(p.get_height()), (p.get_x()+0.25, p.get_height()+1), va = 'bottom', color = 'black')
```

→



▼ changing "date_added" column type from object to datetime64[ns]

```
netflix_df['date_added'] = pd.to_datetime(netflix_df['date_added'])
netflix_df.dtypes
```

```
→
show_id      object
type         object
title        object
director     object
cast         object
country      object
date_added   datetime64[ns]
release_year  int64
rating       object
duration     object
listed_in    object
description  object
dtype: object
```

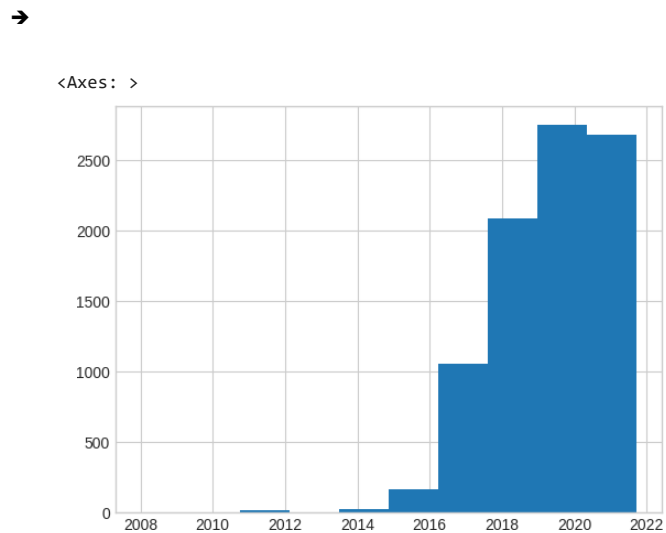
```
netflix_df.sample()
```

```
→
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
141	s142	Movie	Extraction	Steven C. Miller	Bruce Willis, Kellan Lutz, Gina Carano, D.B. S...	United States, United Kingdom, Canada	2021-09-01	2015	R	82 min	Action & Adventure	When a retired CIA agent is kidnapped, his son...

▼ Plotting histogram of how many Content added in netflix each year

```
netflix_df["date_added"].hist()
```



For example :- How many TV Shows released in 2001

```
netflix_df[(netflix_df["type"] == "TV Show") & (netflix_df["release_year"] == 2001)]
```



	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
657	s658	TV Show	The Mole	NaN	Anderson Cooper	United States	2021-06-21	2001	TV-14	2 Seasons	Documentaries	In this competition show, contestants try to e...
5672	s5673	TV Show	InuYasha	NaN	Kappei Yamaguchi, Satsuki Yukino, Kumiko Watan...	Japan	2016-12-15	2001	TV-14	2 Seasons	Anime	Yanked down a well, schoolgirl Kagome Higurash...
5811	s5812	TV Show	Behind Enemy Lines	NaN	NaN	United States	2016-08-15	2001	TV-14	1 Season	Documentaries	After dire setbacks in 1940, Winston Churchill...
7782	s7783	TV Show	Power Rangers Time Force	NaN	Jason Faunt, Erin Cahill, Kevin Kleinberg, Deb...	United States	2016-01-01	2001	TV-Y7	1 Season	Children	The team launches a new campaign against evil ...
8214	s8215	TV Show	The Blue Planet: A Natural History of the Oceans	Alastair Fothergill	David Attenborough	United Kingdom	2015-10-10	2001	TV-G	1 Season	Documentaries	David Attenborough narrates this definitive ex...

Checking the availability of Content on Netflix by their name

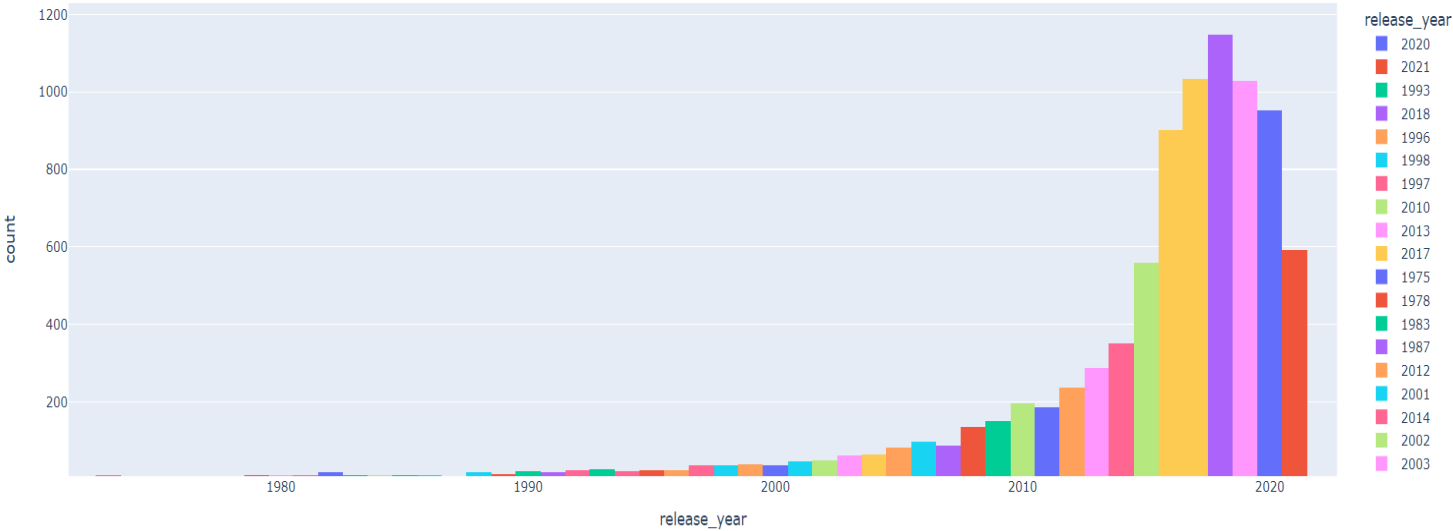
```
netflix_df[netflix_df["title"].isin(["Dark"])]
```



	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
2327	s2328	TV Show	Dark	NaN	Louis Hofmann, Oliver Masucci, Jördis Triebel,...	Germany, United States	2020-06-27	2020	TV-MA	3 Seasons	Crime	A missing child sets four families on a frantic...

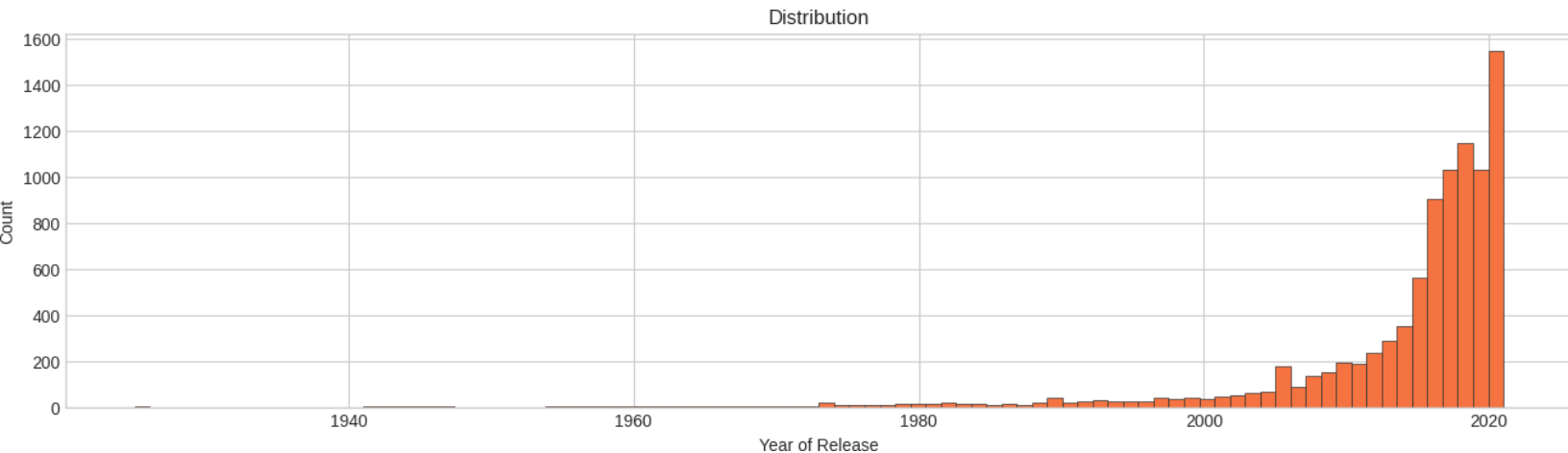
Plotting histogram on release of Content

```
px.histogram(netflix_df, x = "release_year", color = "release_year")
```



Plotting histogram on how many Content release each year by using - Matplotlib

```
plt.figure(figsize=(16, 4))
plt.style.use('seaborn-whitegrid')
plt.hist(netflix_df["release_year"], bins = 90, facecolor = '#F47340', edgecolor='#323232', linewidth=0.5)
plt.title('Distribution')
plt.xlabel('Year of Release')plt.ylabel('Count')
plt.show()
```



Number of Movies made only in Japan

```
movies_only_in_japan = netflix_df[(netflix_df["type"] == "Movie") & (netflix_df["country"] == "Japan")]["title"]
movies_only_in_japan.count()
```

76

```
movies_only_in_japan.unique()
```

```
array(['InuYasha the Movie 2: The Castle Beyond the Looking Glass', 'InuYasha the Movie  
3: Swords of an Honorable Ruler',  
'InuYasha the Movie 4: Fire on the Mystic Island',  
'InuYasha the Movie: Affections Touching Across Time',  
'Naruto Shippuden the Movie: Blood Prison', 'Naruto Shippûden  
the Movie: Bonds',  
'Naruto Shippûden the Movie: The Will of Fire', 'Naruto Shippuden:  
The Movie',  
'Naruto Shippuden: The Movie: The Lost Tower',  
'Naruto the Movie 2: Legend of the Stone of Gelel',  
'Naruto the Movie 3: Guardians of the Crescent Moon Kingdom',  
'Naruto the Movie: Ninja Clash in the Land of Snow', 'HOMUNCULUS', 'GANTZ:0', 'Ride or  
Die',  
'Seven Souls in the Skull Castle: Season Flower',  
'Any Crybabies Around?', 'Blue, Painful, Fragile', 'MOTHER',  
'DTC Yukemuri Junjo Hen From High & Low', 'High & Low The Movie', 'High & Low The  
Movie 2 / End of Sky',  
'High & Low The Movie 3 / Final Mission',  
'High & Low The Red Rain', 'Road To High & Low',  
'Fate/Grand Order -First Order-', 'Romance Doll', 'A Whisker Away', 'Okko's Inn',  
'Modest Heroes: Ponoc Short Films Theatre',  
'Sol Levante',  
'Violet Evergarden: Eternity and the Auto Memory Doll',  
'Lu Over the Wall', 'Pokémon: Mewtwo Strikes Back - Evolution', 'Dragon Quest Your  
Story', '37 Seconds', 'NiNoKuni',  
'Pokémon the Movie: Power of Us', 'Hot Gimmick: Girl Meets Boy',  
'Fireworks', 'No Game No Life: Zero', 'Girls und Panzer der Film', 'K-on! the movie',  
'The Garden of Words', 'The Forest of Love',  
'TUNA GIRL', 'Hikaru Utada Laughter in the Dark Tour 2018', 'EVANGELION: DEATH  
(TRUE)²', 'The End of Evangelion',
```



```
'A Silent Voice', 'River's Edge', 'GODZILLA The Planet Eater','Pokémon the Movie:
I Choose You!','
The Seven Deadly Sins the Movie: Prisoners of the Sky',
'Aggretsuko: We Wish You a Metal Christmas',
'GODZILLA City on the Edge of Battle', 'Expelled from Paradise','Lupin the 3rd: The
Castle of Cagliostro: Special Edition',
'In This Corner of the World', 'FullMetal Alchemist',
'Berserk: The Golden Age Arc II - The Battle for Doldrey','Berserk: The
Golden Age Arc III - The Advent', 'BLAME!', 'Berserk: The Golden Age Arc I -
The Egg of the King',
'Bleach The Movie: Fade to Black', 'Bleach the Movie: Hell Verse','Ghost Pain',
'Ghost Tears', 'Ghost Whispers', 'Golden Time',
'Inuyasha the Movie - L'isola del fuoco scarlatto",
'Inuyasha the Movie - La spada del dominatore del mondo',
'InuYasha: The Movie 2: The Castle Beyond the Looking Glass',"Mary and the
Witch's Flower", 'Miss Hokusai',
'Naruto Shippuden : Blood Prison'], dtype=object)
```

Working with Month Wise Content Added

```
netflix_df["date_added"].dt.month
```

→

```
0      9.0
1      9.0
2      9.0
3      9.0
4      9.0
```

...

```
8802    11.0
8803     7.0
8804    11.0
8805     1.0
8806     3.0
```

```
Name: date_added, Length: 8807, dtype: float64
```

```
netflix_df["date_added_month"] = netflix_df["date_added"].dt.month.fillna(-1)
netflix_df["date_added_month"].unique()
```

→

```
array([ 9.,  8.,  7.,  6.,  5.,  4.,  3.,  2.,  1., 12., 11., 10., -1.])
```

```
netflix_df["date_added_month"] = netflix_df["date_added_month"].astype("Int64")netflix_df.info()
```

→

```
<class 'pandas.core.frame.DataFrame'>RangeIndex: 8807
entries, 0 to 8806
```

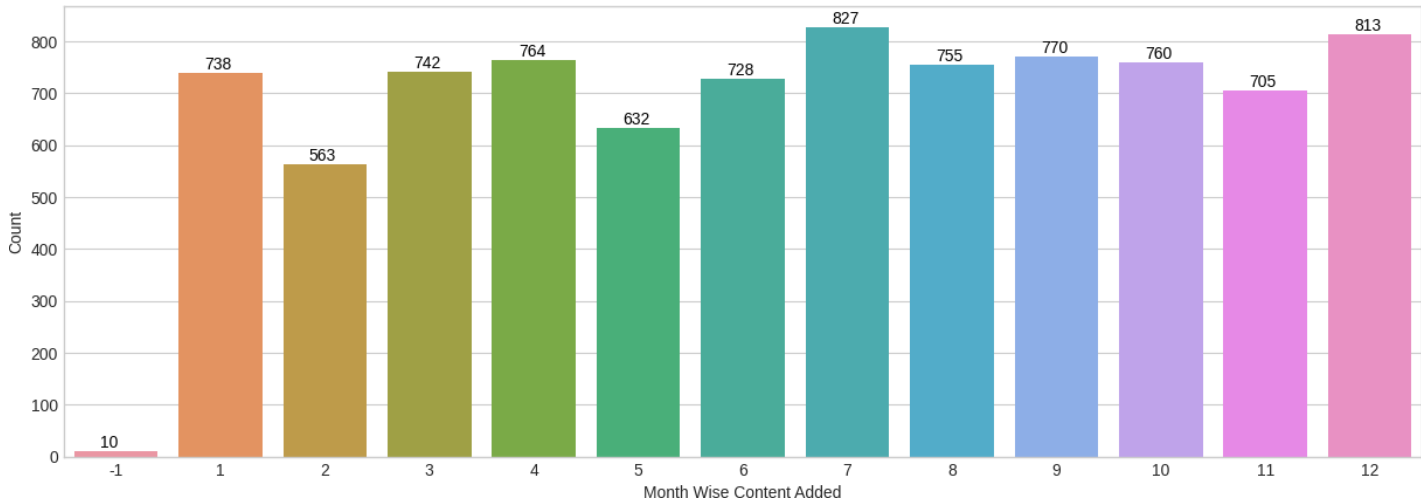
Data columns (total 13 columns):

#	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	datetime64[ns]
7	release_year	8807 non-null	int64
8	rating	8803 non-null	object
9	duration	8804 non-null	object
10	listed_in	8789 non-null	object
11	description	8807 non-null	object
12	date_added_month	8807 non-null	Int64

```
dtypes: Int64(1), datetime64[ns](1), int64(1), object(10)
memory usage: 903.2+ KB
```

```
plt.figure(figsize = (15, 5))
ax = sns.countplot(x = 'date_added_month', data = netflix_df)
plt.xlabel('Month Wise Content Added')
plt.ylabel('Count')
for p in ax.patches:
    ax.annotate(int(p.get_height()), (p.get_x()+0.25, p.get_height()+1), va = 'bottom', color = 'black')
```

➔



▼ Checking the number of Content added Day-wise

```
netflix_df["date_added_day"] = netflix_df["date_added"].dt.day.fillna(-1)
netflix_df["date_added_day"].unique()

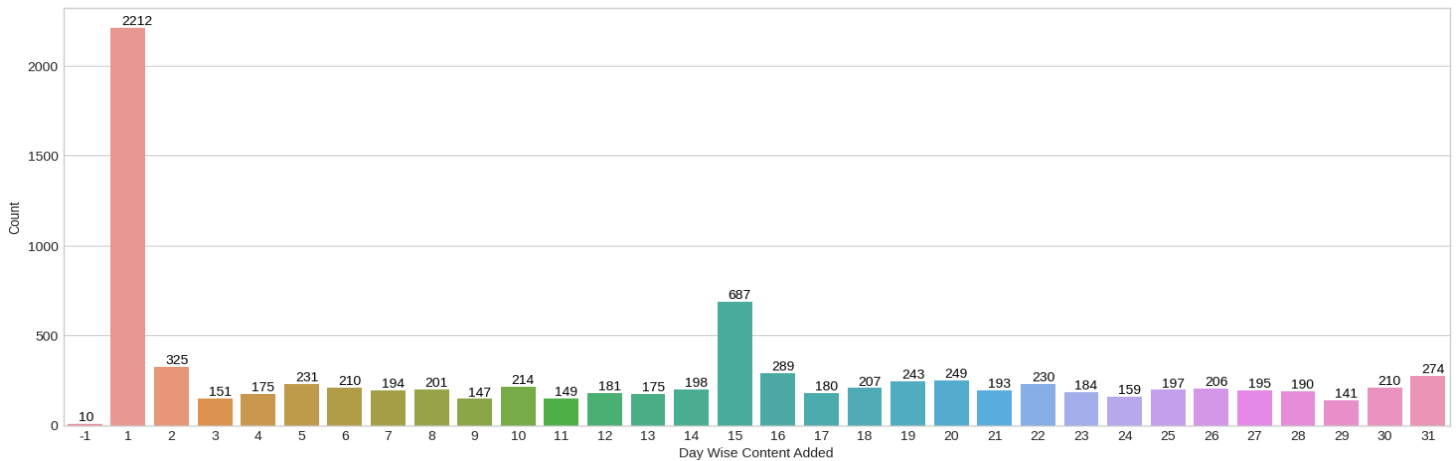
array([25., 24., 23., 22., 21., 20., 19., 17., 16., 15., 14., 11., 10.,
       9., 8., 7., 6., 5., 4., 3., 2., 1., 31., 29., 28., 27.,
       26., 18., 13., 12., 30., -1.])
```

```
netflix_df["date_added_day"] = netflix_df["date_added_day"].astype("Int64")
netflix_df.info()
```

➔

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 14 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   datetime64[ns]
7   release_year          8807 non-null   int64
8   rating                8803 non-null   object
9   duration              8804 non-null   object
10  listed_in             8789 non-null   object
11  description            8807 non-null   object
12  date_added_month      8807 non-null   Int64
13  date_added_day        8807 non-null   Int64
dtypes: Int64(2), datetime64[ns](1), int64(1), object(10)
memory usage: 980.6+ KB
```

```
plt.figure(figsize = (18, 6))
ax = sns.countplot(x = 'date_added_day', data = netflix_df)
plt.xlabel('Day Wise Content Added')
plt.ylabel('Count')
for p in ax.patches:
    ax.annotate(int(p.get_height()), (p.get_x()+0.25, p.get_height()+1), va = 'bottom', color = 'black')
```



- ▼ Working with country

```
country_count = netflix_df.copy()
country_count = pd.concat([country_count, netflix_df["country"].str.split(",", expand = True)], axis = 1)
country_count
```

[illegible]

```
country_count = country_count.melt(id_vars = ["type", "title"], value_vars = range(12), value_name = "country")
country_count = country_count[country_count["country"].notna()]
country_count
```



	type	title	variable	country
0	Movie	Dick Johnson Is Dead	0	United States
1	TV Show	Blood & Water	0	South Africa
4	TV Show	Kota Factory	0	India
7	Movie	Sankofa	0	United States
8	TV Show	The Great British Baking Show	0	United Kingdom
...
78859	Movie	The Look of Silence	8	Germany
85496	Movie	Barbecue	9	Sweden
87666	Movie	The Look of Silence	9	Netherlands

```
country_count["country"].nunique()
```

197

10019 rows x 4 columns

```
country_count["country"].value_counts()[:10]
```



United States	3211
India	1008
United Kingdom	628
United States	479
Canada	271
Japan	259
France	212
South Korea	211
Spain	181
France	181

Name: country, dtype: int64

- Top 10 country where most of the contents are produced
- There are 2 different United States entry in the Dataset, so we will Merge the Duplicate value with Original to Make a single entry

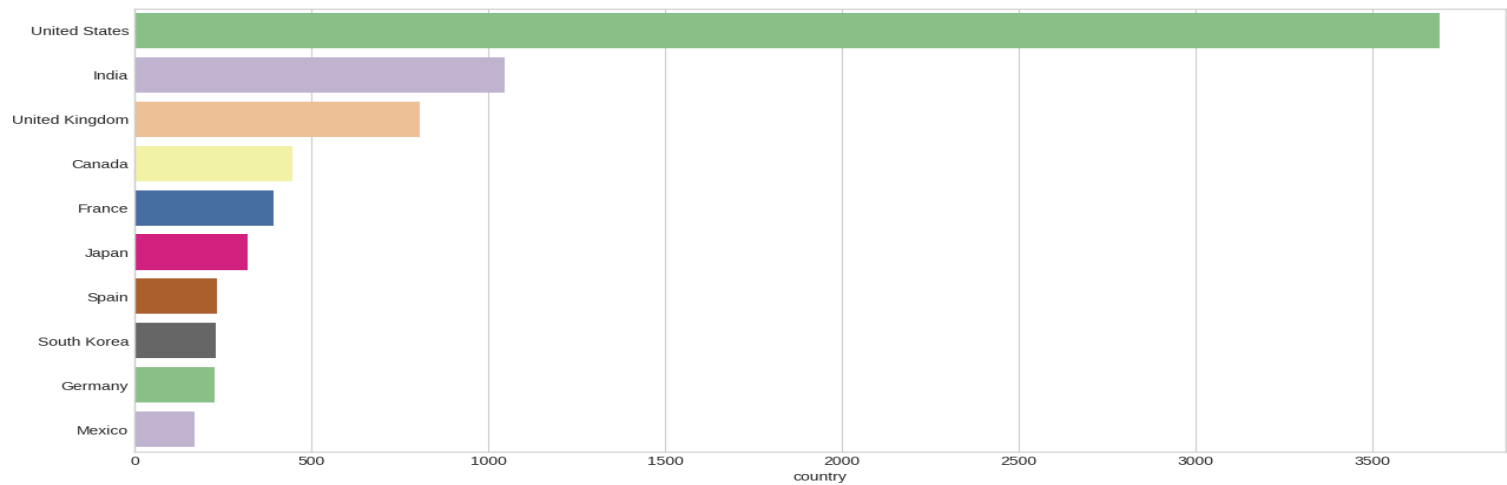
```
country_count["country"] = country_count["country"].str.strip()
country_count["country"].value_counts()[:10]
```

United States	3690
India	1046
United Kingdom	806
Canada	445
France	393
Japan	318
Spain	232
South Korea	231
Germany	226
Mexico	169

Name: country, dtype: int64

```
plt.figure(figsize = (15, 7))
country = country_count["country"].value_counts()[:10]
sns.barplot(x = country, y = country.index, palette = "Accent")
```

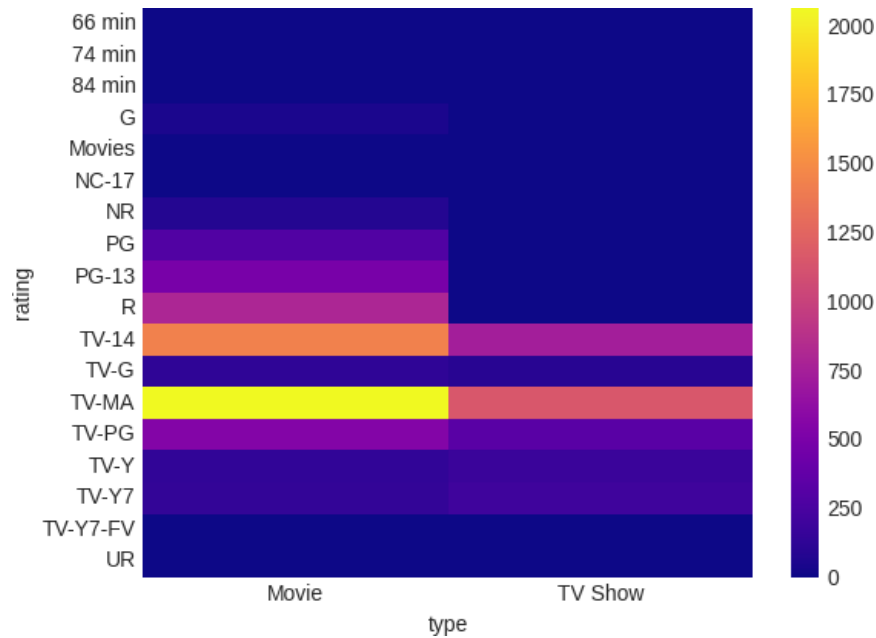
<Axes: xlabel='country'>



Plotting Heatmap b/w Content Rating vs Content Type

```
colormap = plt.cm.plasma
sns.heatmap(pd.crosstab(data["rating"], data["type"]), cmap = colormap)
```

<Axes: xlabel='type', ylabel='rating'>



▼ Top 10 Actors who involved in most number of Content

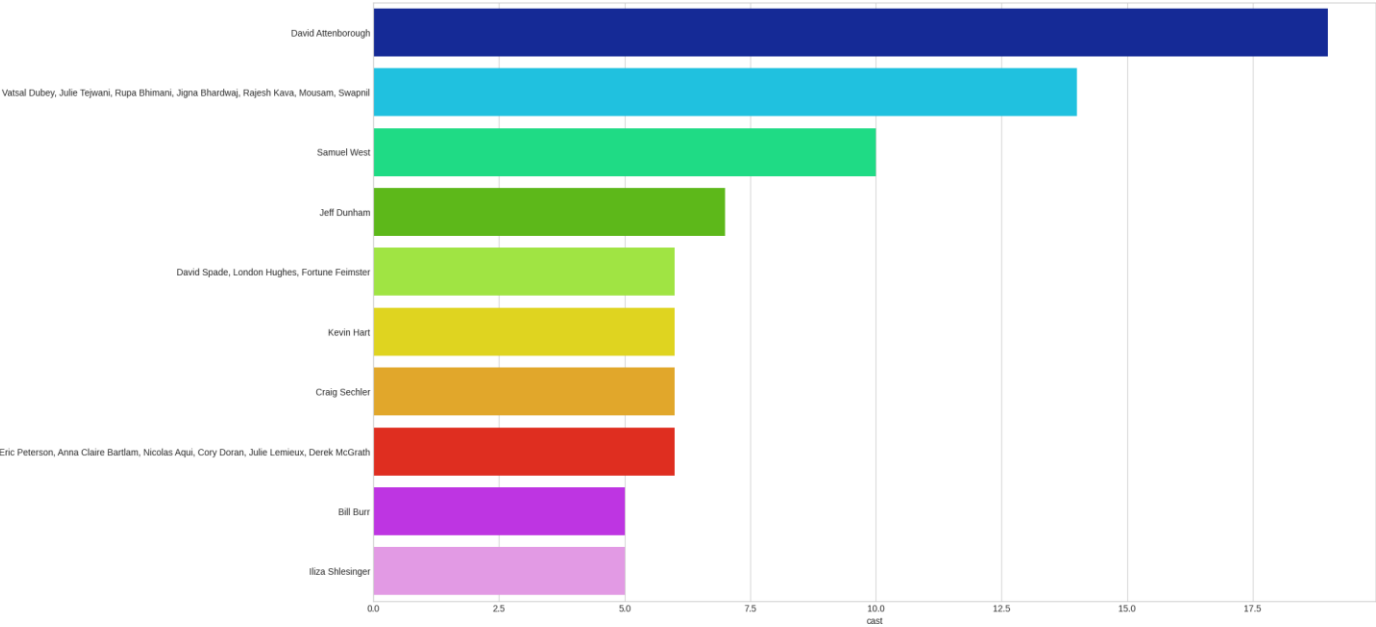
```
cast_count = netflix_df.copy()
cast_count = pd.concat([cast_count, netflix_df["cast"].str.split(",", expand = True)], axis = 1)
cast_count
➡
```

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	...	40	41	42	43	44	45	46	47	48	49
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90 min	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabl...	NaN	2021-09-24	2021	TV-MA	1 Season	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	2021-09-24	2021	TV-MA	1 Season	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021	TV-MA	2 Seasons	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	2019-11-20	2007	R	158 min	...	None	None	None	None	None	None	None	None	None
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	2019-07-01	2018	TV-Y7	2 Seasons	...	None	None	None	None	None	None	None	None	None
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01	2009	R	88 min	...	None	None	None	None	None	None	None	None	None
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88 min	...	None	None	None	None	None	None	None	None	None
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111 min	...	None	None	None	None	None	None	None	None	None

8807 rows × 64 columns

```
plt.figure(figsize = (20, 12))
cast = cast_count["cast"].value_counts()[:10]
sns.barplot(x = cast, y = cast.index, palette = "gist_ncar")
➡
```

<Axes: xlabel='cast'>



How many movies present in Netflix where Bill Murray acted?

```
updated_df = data.dropna()
updated_df[updated_df["cast"].str.contains("Bill Murray")]
```



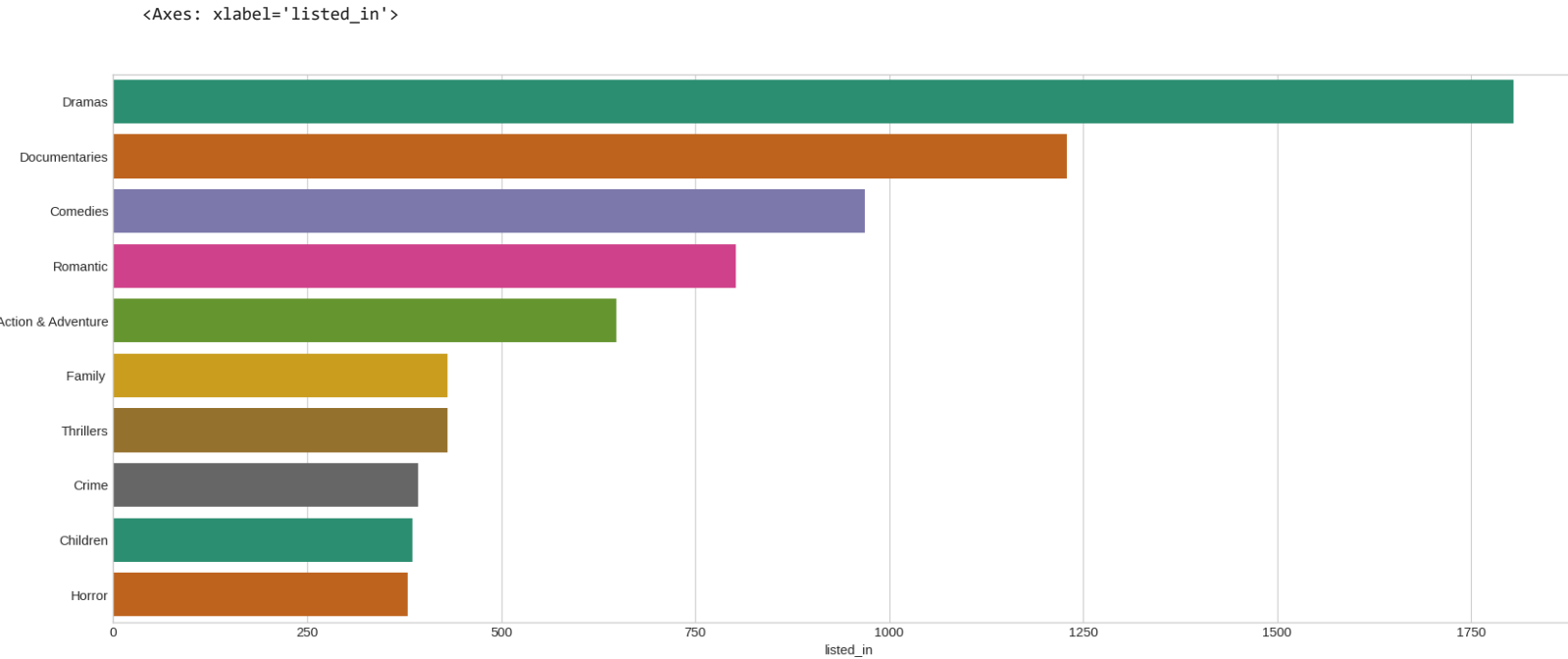
	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
494	s495	Movie	Rock the Kasbah	Barry Levinson	Bill Murray, Kate Hudson, Zooey Deschanel, Dan...	United States	8-Jul-21	2015	R	106 min	Comedies	When a has-been music producer gets stuck in A...
566	s567	Movie	Charlie's Angels	McG	Cameron Diaz, Drew Barrymore, Lucy Liu, Bill M...	United States, Germany	1-Jul-21	2000	PG-13	98 min	Action & Adventure	A tight-knit trio of specially trained agents ...
5881	s5882	Movie	A Very Murray Christmas	Sofia Coppola	Bill Murray, Miley Cyrus, George Clooney, Chri...	United States	4-Dec-15	2015	TV-14	57 min	Comedies	Bill Murray rounds up an all-star cast for an ...
6036	s6037	Movie	A Glimpse Inside the Mind of Charles Swan III	Roman Coppola	Charlie Sheen, Jason Schwartzman, Bill Murray,...	United States	15-Jul-18	2012	R	86 min	Comedies	When his girlfriend walks out and leaves him a...
6842	s6843	Movie	Get Smart	Peter Segal	Steve Carell, Anne Hathaway, Dwayne Johnson, A...	United States	1-Apr-19	2008	PG-13	110 min	Action & Adventure	When the identities of secret agents are compr...
7210	s7211	Movie	Kingpin	Peter Farrelly, Bobby Farrelly	Woody Harrelson, Randy Quaid, Vanessa Angel, B...	United States	1-Jan-20	1996	PG-13	114 min	Comedies	After a bowler loses his hand, his career take...
8105	s8106	Movie	Stripes	Ivan Reitman	Bill Murray, Harold Ramis, Warren Oates, P.J. ...	United States	1-Sep-19	1981	R	106 min	Classic	After losing everything, an indolent sad sack ...
8210	s8211	Movie	The Bill Murray Stories: Life Lessons Learned ...	Tommy Avallone	Tommy Avallone, Bill Murray, Joel Murray, Pete...	United States	31-Dec-18	2018	TV-MA	72 min	Documentaries	This documentary highlights spontaneous encoun...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	1-Nov-19	2009	R	88 min	Horror	Looking to survive in a world taken over by zo...

```
listed_in_count["listed_in"].unique()

array(['Documentaries', 'Dramas', 'Crime ', 'Romantic', 'Horror',
      'Family ', 'Crime', 'Thrillers', 'Action & Adventure', 'Comedies',
      'Sci-Fi', 'Children', 'Classic', 'Anime', 'Documentaries ',
      'Korean', 'Family', 'Documentaries ', 'TV Shows',
      'Action & Adventure ', 'Romantic ', 'Movies', 'Comedies ',
      'Comedies', 'Classic ', 'Family, ', 'Anime ', 'Documentaries',
      'Children ', 'TVHorror ', 'Horror', 'CrimeChildren',
      'Classic Dramas', 'Anime Features', 'ClassicDramas'], dtype=object)
```

Top 10 Genre for the content

```
plt.figure(figsize = (20, 8))
listed_in = listed_in_count["listed_in"].value_counts()[:10]
sns.barplot(x = listed_in, y = listed_in.index, palette = "Dark2")
```



▼ Top 10 Directors who involved in most number of Content

```
director_count = netflix_df.copy()
director_count = pd.concat([director_count, netflix_df["director"].str.split(",", expand = True)], axis = 1)
director_count["director"].nunique()
```

➔

4528

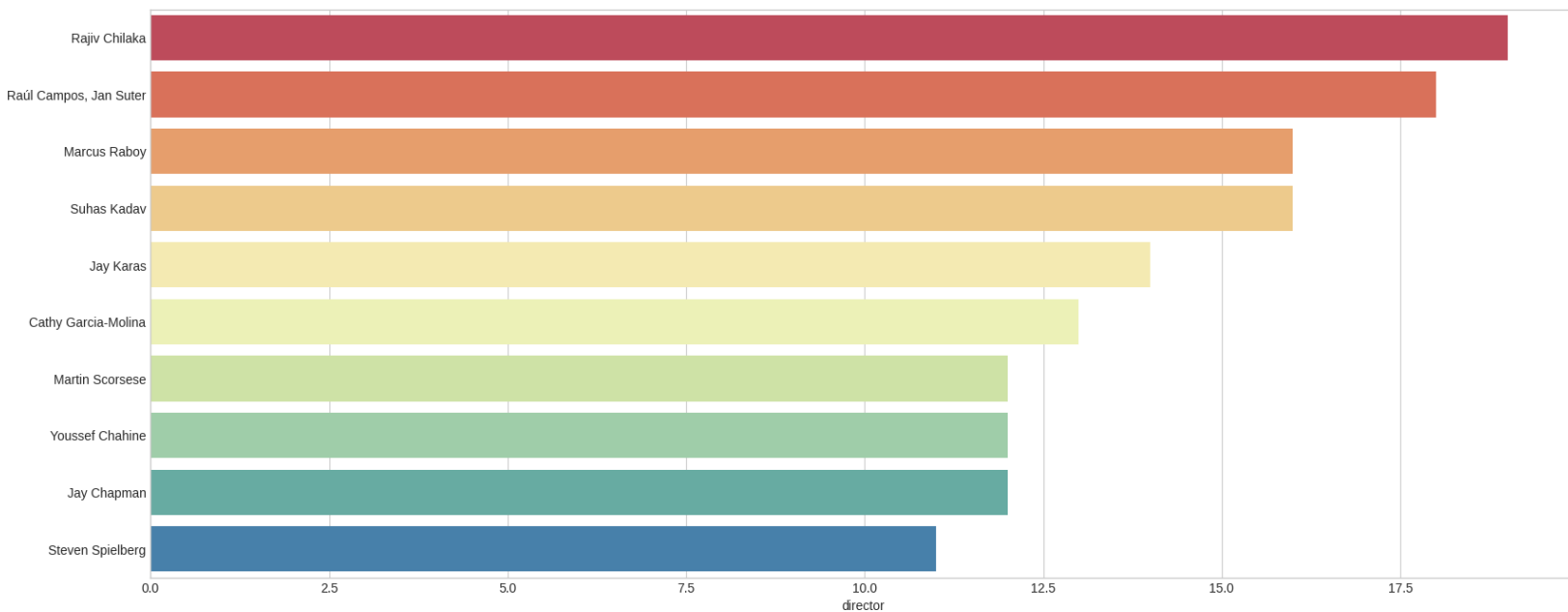
```
director_count["director"].value_counts()
```

➔

```
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
Mozes Singh            1
Name: director, Length: 4528, dtype: int64
```

```
plt.figure(figsize = (20, 8))
director_count = director_count["director"].value_counts()[:10]
sns.barplot(x = director_count, y = director_count.index, palette = "Spectral")<Axes: xlabel='director'>
```

➔



▼ Number of Movies in Rating

```
show_rating = netflix_df.groupby(['rating', 'type'])['rating'].count().reset_index(name='ratings_count')
show_rating = show_rating.set_index('rating')
rating_movie = netflix_df.groupby(['rating', 'type'])['rating'].count().reset_index(name='ratings_count')
rating_movie = rating_movie.set_index('rating')
filter = rating_movie["type"] == "Movie"
```

```
# filtering data
```

```
rating_movie = rating_movie[filter].dropna()
rating_movie
```

	type	ratings_count
rating		
66 min	Movie	1
74 min	Movie	1
84 min	Movie	1
G	Movie	41
NC-17	Movie	3
NR	Movie	75
PG	Movie	287
PG-13	Movie	490
R	Movie	797
TV-14	Movie	1427
TV-G	Movie	126
TV-MA	Movie	2062
TV-PG	Movie	540
TV-Y	Movie	131
TV-Y7	Movie	139
TV-Y7-FV	Movie	5

▼ Number of TV Shows in Rating

```
rating_TV= netflix_df.groupby(['rating', 'type'])['rating'].count().reset_index(name='ratings_count')
rating_TV=rating_TV.set_index('rating')
rating_TV

filter = rating_TV["type"]=="TV Show"
```

```
# filtering data
rating_TV = rating_TV.where(filter).dropna()
rating_TV
```

	type	ratings_count
rating		
NR	TV Show	5.0
R	TV Show	2.0
TV-14	TV Show	733.0
TV-G	TV Show	94.0
TV-MA	TV Show	1145.0
TV-PG	TV Show	323.0
TV-Y	TV Show	176.0
TV-Y7	TV Show	195.0
TV-Y7-FV	TV Show	1.0

PROGRAM CODE IN POWER BI

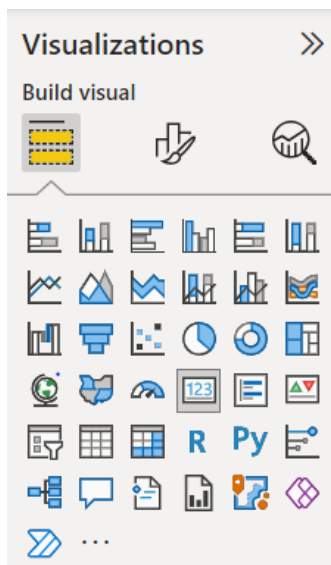
Key Measures:

```
% movies net = DIVIDE([Movies Net],[Shows Netflix],0)
% series net = DIVIDE('All Measures'[Series Net],[Shows Netflix],0)
Movies Net = CALCULATE([Shows Netflix],'netflix-rotten-tomatoes-metacritic-imdb'[Series or Movie] =
    "Movie")
Series Net = CALCULATE([Shows Netflix],'netflix-rotten-tomatoes-metacritic-imdb'[Series or Movie] =
    "Series")
Shows Netflix = COUNT('netflix-rotten-tomatoes-metacritic-imdb'[Title])
```

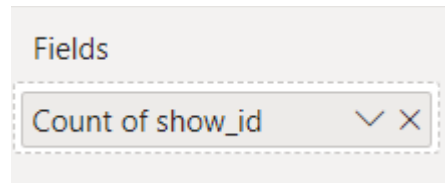
Creating Dashboard by using these Tools and Methods:

1. Card :-

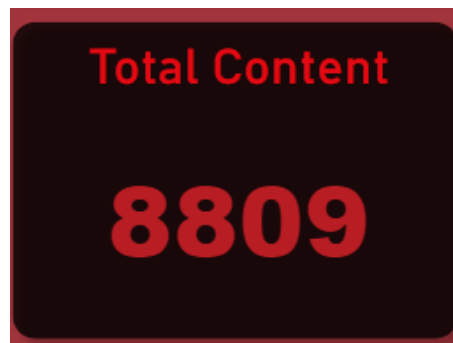
- The first method to create a card is to use the report editor in Power BI Desktop.
- Select the Store > Open Store Count field. Power BI creates a column chart with the one number.
- In the Visualizations pane, select the Card icon.



- We go in Fields option then drop the count of show_id. Due to this we can see Total Content.

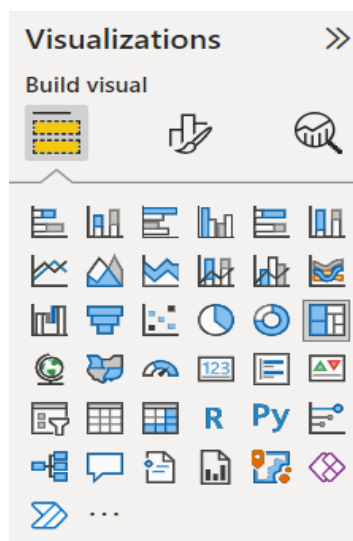


- A Card has been created as shown below.



2. Treemap: -

- Treemaps display hierarchical data as a set of nested rectangles. Each level of the hierarchy is represented by a colored rectangle (branch) containing smaller rectangles (leaves). Power BI bases the size of the space inside each rectangle on the measured value. The rectangles are arranged in size from top left (largest) to bottom right (smallest).
- In the Visualizations pane, select the Treemap icon.

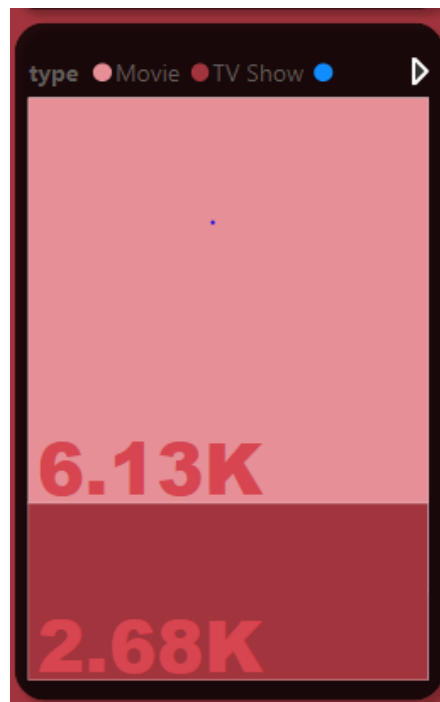


- We go in Category, Details and Value then drop Type, Type and Count of Type Column simultaneously. Due to this we can see Type of Content.

The image shows a configuration interface for a dashboard. It consists of three vertically stacked sections, each with a title and a dropdown menu. The first section is titled 'Category' and has a dropdown menu with 'type' selected. The second section is titled 'Details' and also has a dropdown menu with 'type' selected. The third section is titled 'Values' and has a dropdown menu with 'Count of type' selected. Each dropdown menu has a checkmark icon and an 'X' icon.

Section	Selected Value
Category	type
Details	type
Values	Count of type

- A simple Tree chart has been created as shown below.



- Drag and drops the fields that are to be placed on the Category and Values of the chart respectively.

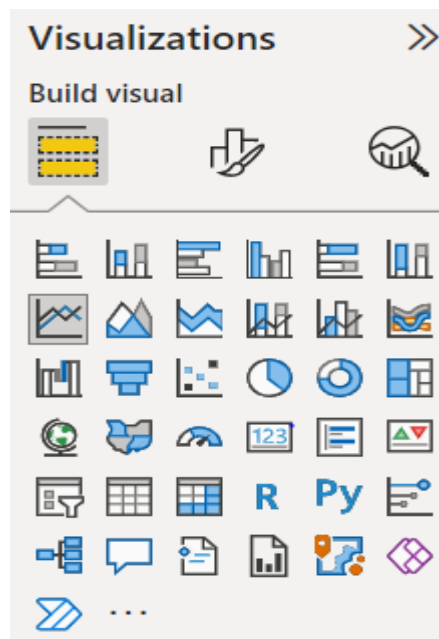
The image shows a configuration panel for a chart. It has three main sections: 'Category', 'Details', and 'Values'. The 'Category' section contains a dropdown menu with 'listed_in' selected. The 'Details' section contains a text box with the placeholder 'Add data fields here'. The 'Values' section contains a dropdown menu with 'Count of listed_in' selected.

- A simple Tree chart has been created as shown below.

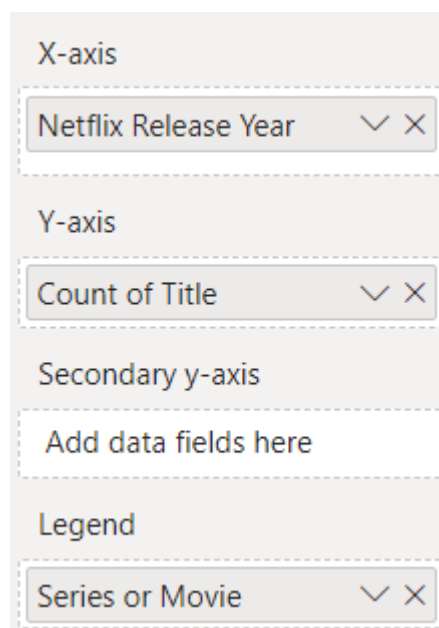


3. Line Chart :-

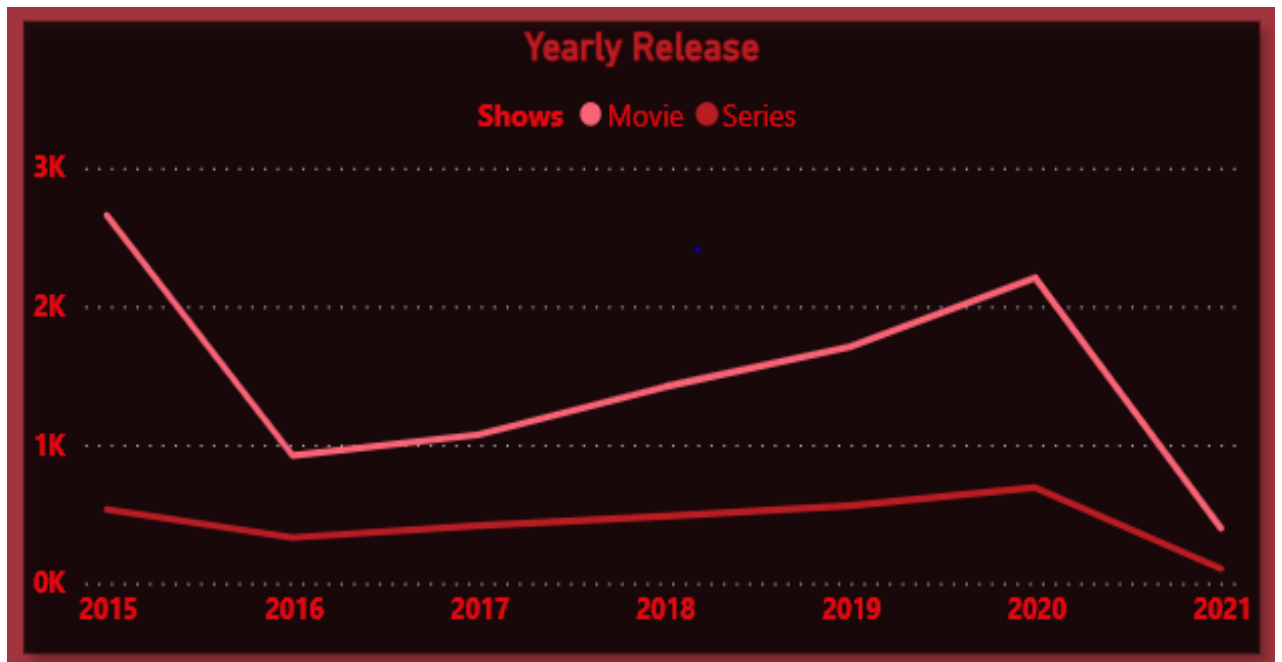
- A line chart is a series of data points that are represented by dots and connected by straight lines. A line chart may have one or many lines. Line charts have an x and a y axis.
- In the Visualizations pane, select the Line Chart icon.



- we go in X-axis, Y-axis and Legend then drop the Netflix Release year, Count of the title and Series or Movies simultaneously. Due to this we can see Yearly Release Movies and Series.



- A simple Line Chart has been created as shown below.



- Drag and drops the fields that are to be placed on the X-axis and Y-axis of the chart respectively.

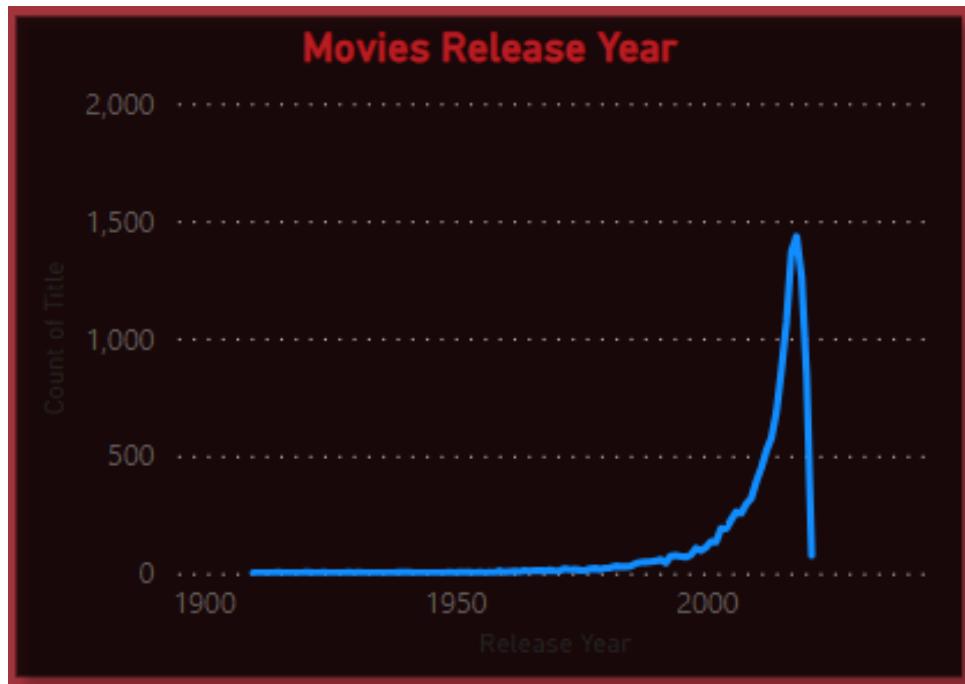
X-axis

Release Year

Y-axis

Count of Title

- A simple Line Chart has been created as shown below.



- Drag and drops the fields that are to be placed on the X-axis and Y-axis of the chart respectively.

X-axis

Netflix Release Year

Y-axis

Count of Title

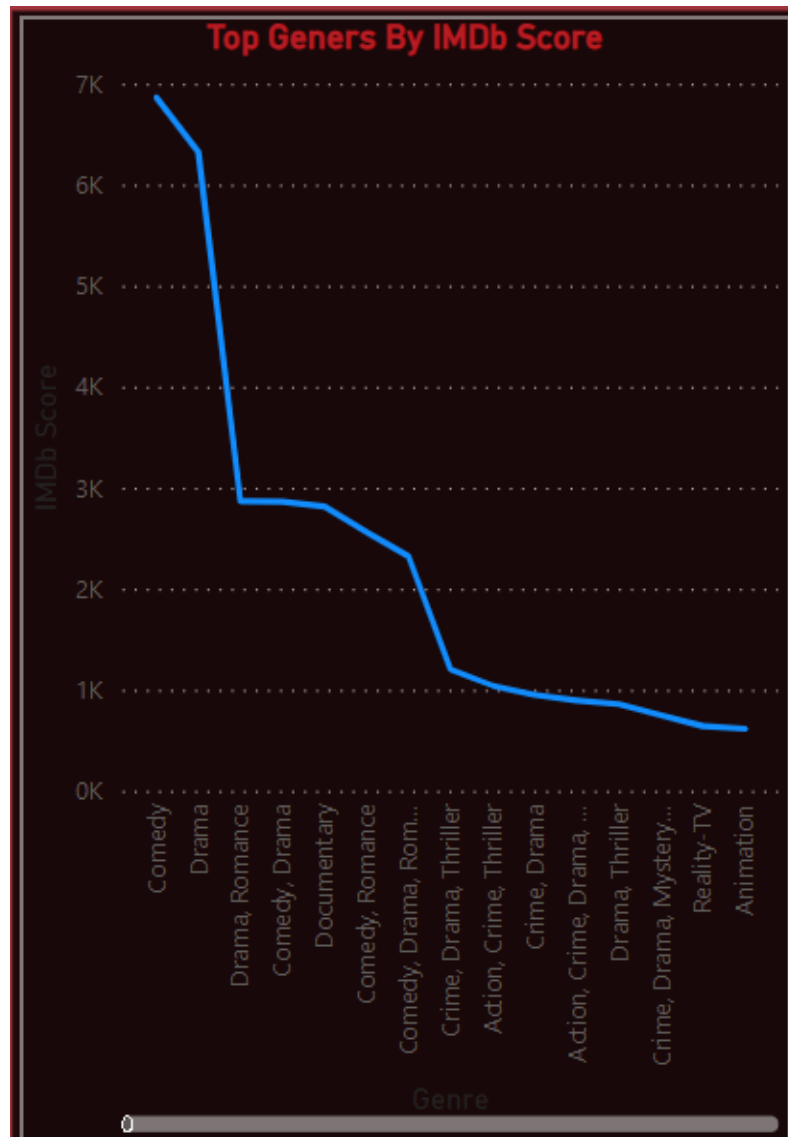
- A simple Line Chart has been created as shown below.



- Drag and drops the fields that are to be placed on the X-axis and Y-axis of the chart respectively.

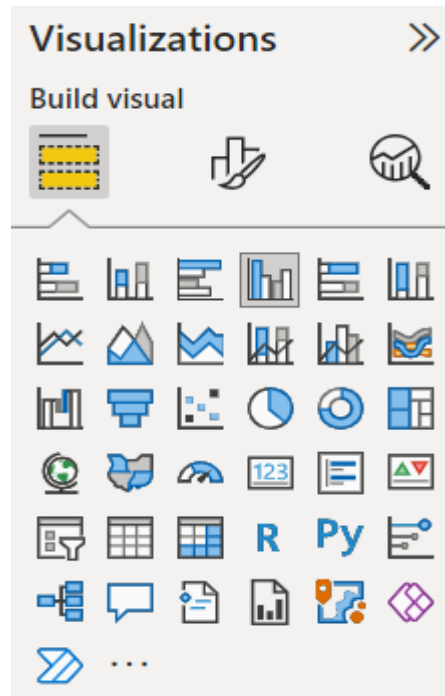
X-axis	
Genre	✓ ✕
Y-axis	
IMDb Score	✓ ✕

- A simple Line Chart has been created as shown below.

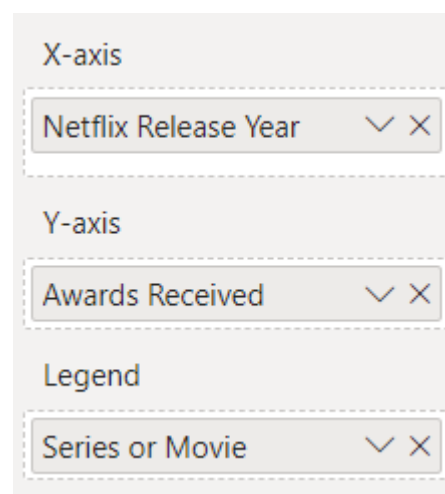


4. **Clustered Column Chart :-** In Power BI, a combo chart is a single visualization that combines a line chart and a column chart. Combining the two charts into one lets you make a quicker comparison of the data.
 - Combo charts are a great choice:
 - a. when you have a line chart and a column chart with the same X axis.
 - b. to compare multiple measures with different value ranges.
 - c. to illustrate the correlation between two measures in one visualization.
 - d. to check whether one measure meets the target defined by another measure
 - e. to conserve canvas space.

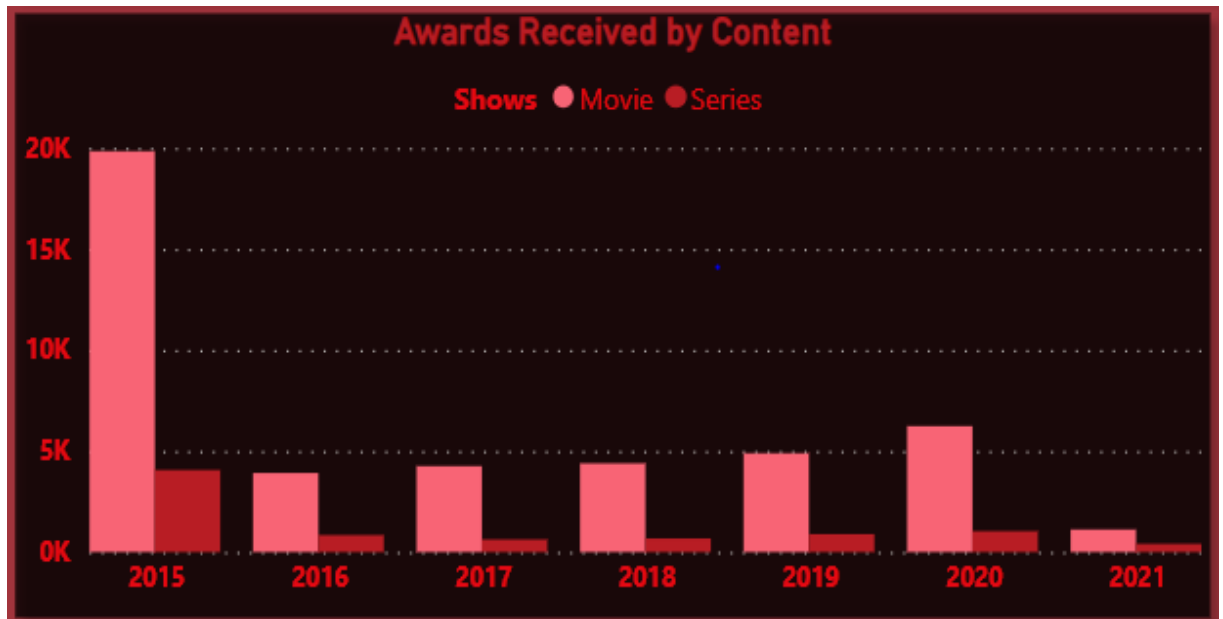
- In the Visualizations pane, select the Clustered Column Chart icon.



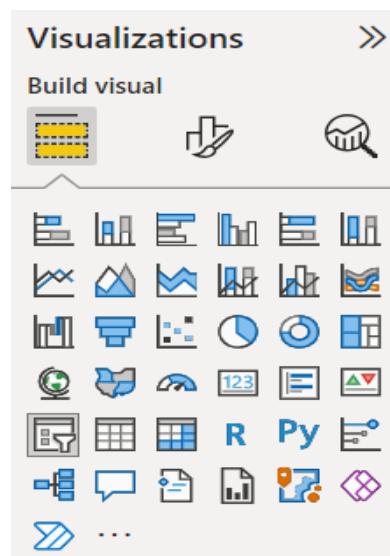
- We go in X-axis, Y-axis and Legend then drop the Netflix Release year, Awards Received and Series or Movies simultaneously. Due to this we can see Yearly Release Movies and Series.



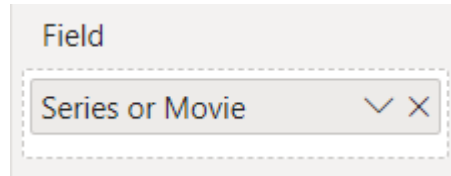
- The Clustered Column Chart has been created as shown below.



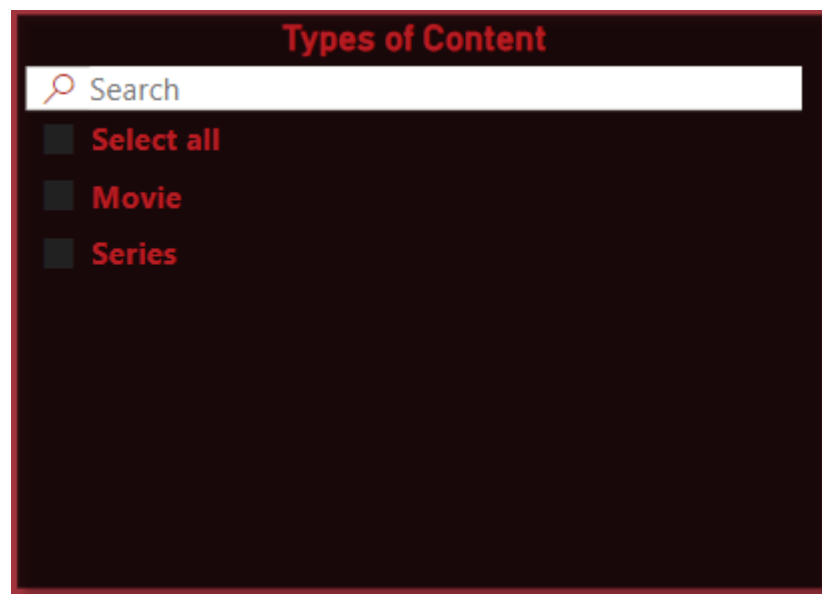
5. Slicer :- Slicers are another way of filtering. They're displayed on the report page, and narrow the portion of the dataset that's shown in the other report visualizations.
 - Create a slicer
 - a. Open Power BI Desktop, and from the menu bar, select File > Open report.
 - b. Browse to the Retail Analysis Sample PBIX.pbix file, then select Open.
 - c. On the left pane, select the Report icon to open the file in report view.
 - d. On the Overview page, with nothing selected on the report canvas, select the Slicer icon in the Visualizations pane to create a new slicer.



- We go in Field then drop Series or Movies.



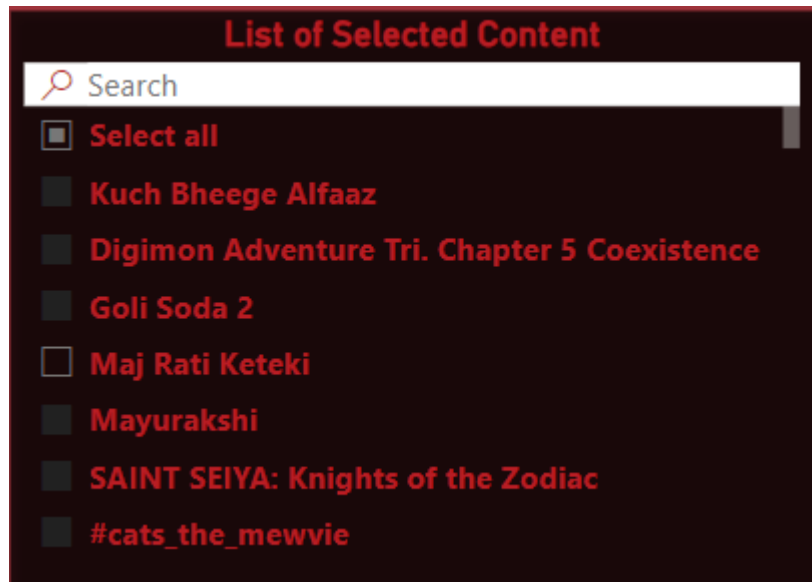
- A Slicer has been created as shown below.



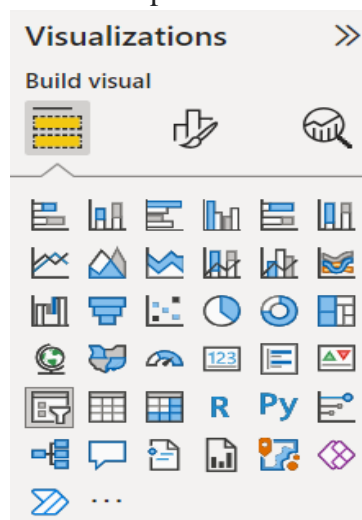
- We go in Field then drop Title.



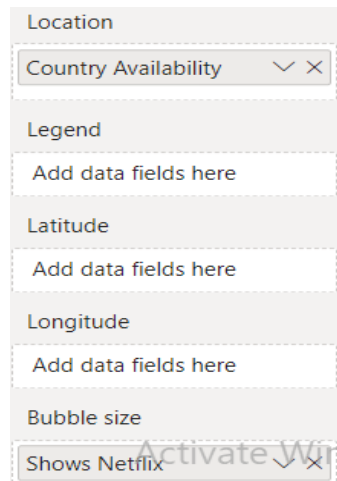
- A Slicer has been created as shown below.



6. **Map :-** Power BI integrates with Bing Maps to provide default map coordinates (a process called geo-coding) so you can create maps. Together they use algorithms to identify the correct location, but sometimes it's a best guess.
- The Power BI service and Power BI Desktop send Bing the geo data it needs to create the map visualization. This may include the data in the Location, Latitude, and Longitude buckets of the visual's field well. Exactly what is sent varies by map type.
 - a. For maps (bubble, scatter, and dot plot maps), if latitude and longitude are provided, then no data is sent to Bing. Otherwise, any data in the Location bucket is sent to Bing.
 - b. Filled maps require a field in the Location bucket; even if latitude and longitude are provided. Whatever data is in the Location, Latitude, or Longitude bucket is sent to Bing.
 - In the Visualizations pane, select the Map icon.



- We go in Location and Bubble Size then drop Country Availability and Shows Netflix.

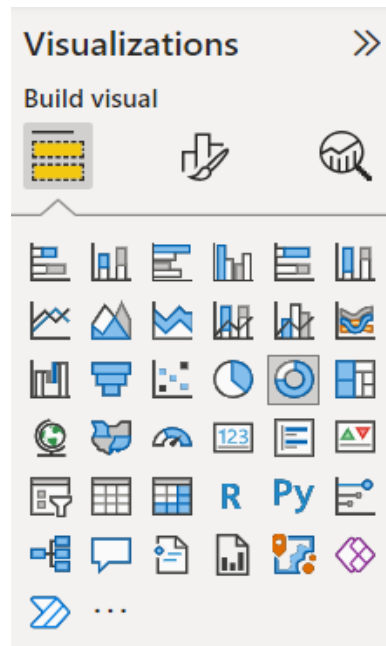


- A Map has been created as shown below.

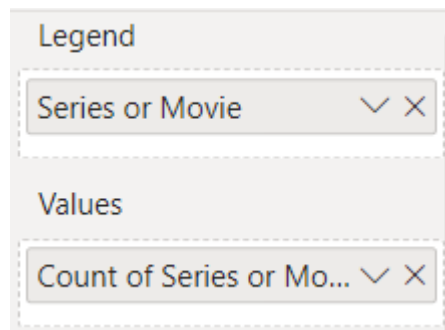


7. **Donut Chart** :- A donut chart is similar to a pie chart in that it shows the relationship of parts to a whole. The only difference is that the center is blank and allows space for a label or icon.

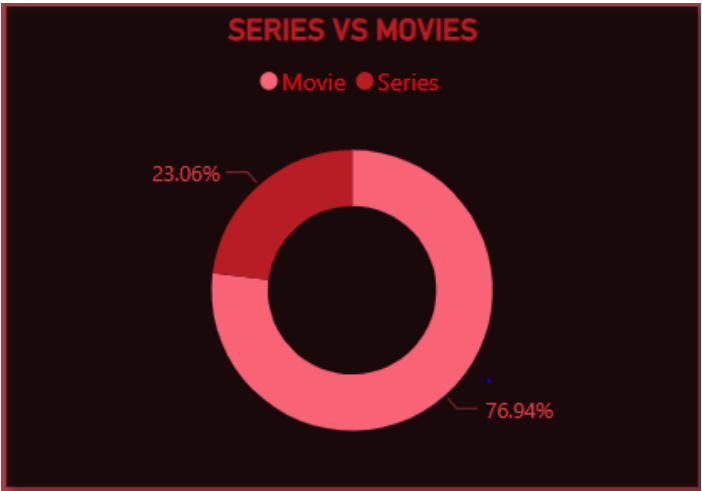
- Create a doughnut chart
 - a. Open Power BI Desktop, and from the menu bar, select File > Open report.
 - b. Browse to the Retail Analysis Sample PBIX.pbix file, then select Open.
 - c. On the left pane, select the Report icon to open the file in report view
 - d. In the Visualizations pane, select the Donut icon.



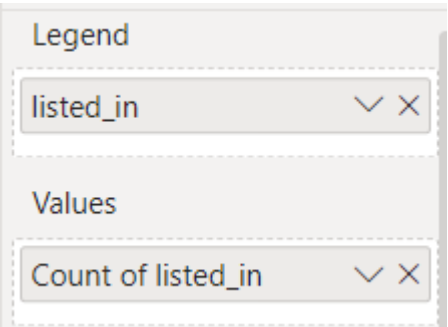
- e. We go in Legend and Value then drop Series or Movies and Count of Series or Movies simultaneously.



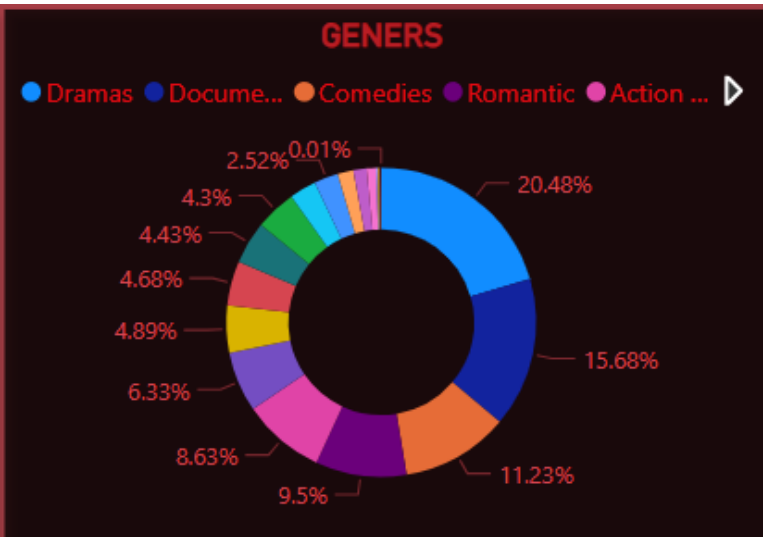
f. A Donut Chart has been created as shown below.



g. We go in Legend and Value then drop listed_in and Count of listed_in simultaneously.

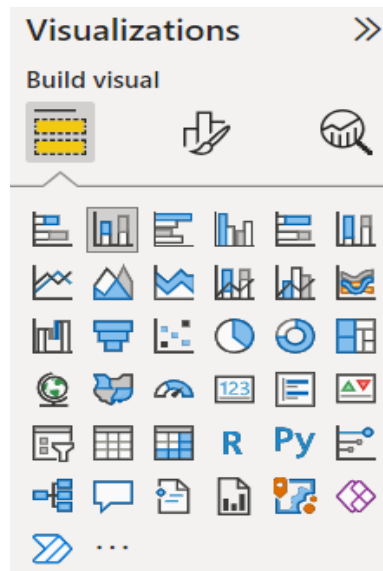


h. A Donut Chart has been created as shown below.

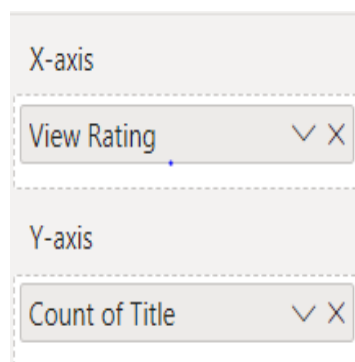


8. **Stacked Column Chart :-** Stacked charts are a sort of bar chart which are multiple-bar in nature where related values are placed at one another. This feature shows the composition and comparison of a few variables, either in absolute or relative terms together with comparing multiple categories simultaneously.

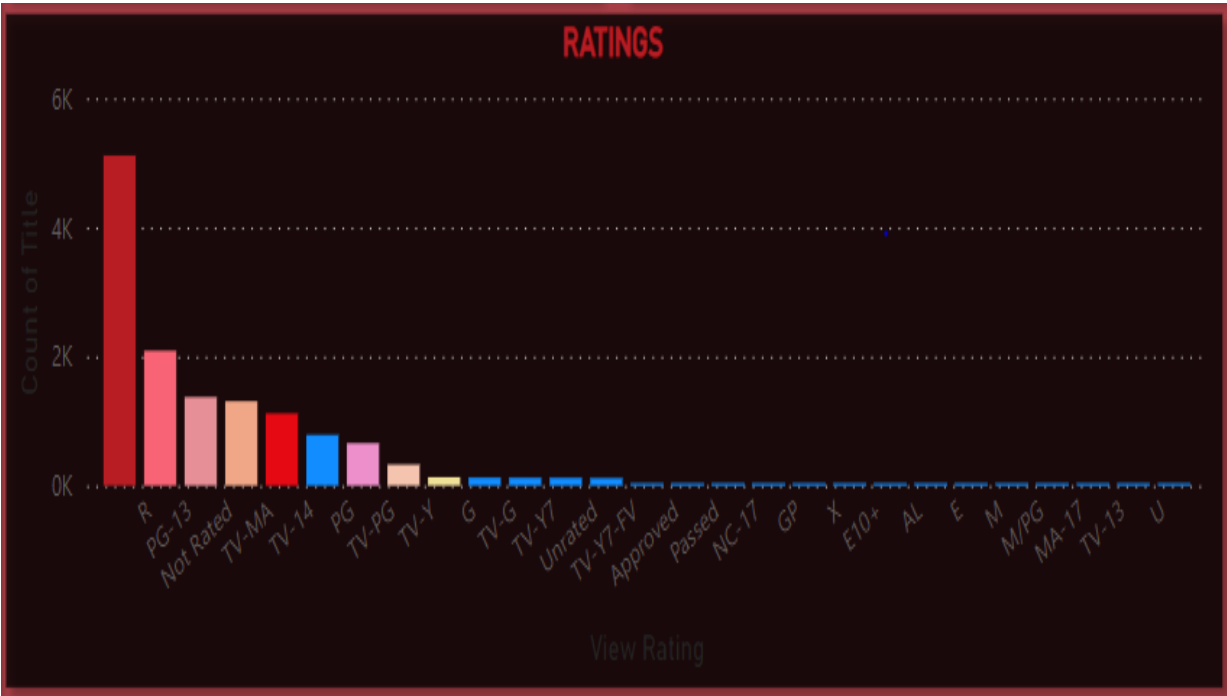
- Creation of Stacked Column Charts
 - a. Import data from Your Excel to Power BI. Go to Home Tab-> Get Data -> Choose Data which you want Example Excel -> Select the file and Open -> Select The sheet and Load.
 - b. Load the column chart under visualization and click on the 'stacked column chart' icon.



- c. The chart will be loaded on the screen and resized if required
- d. To start creating a column chart click on the fields which is to be used in a column chart.
- e. Drag and drops the fields that are to be placed on the X-axis and Y-axis of the chart respectively.

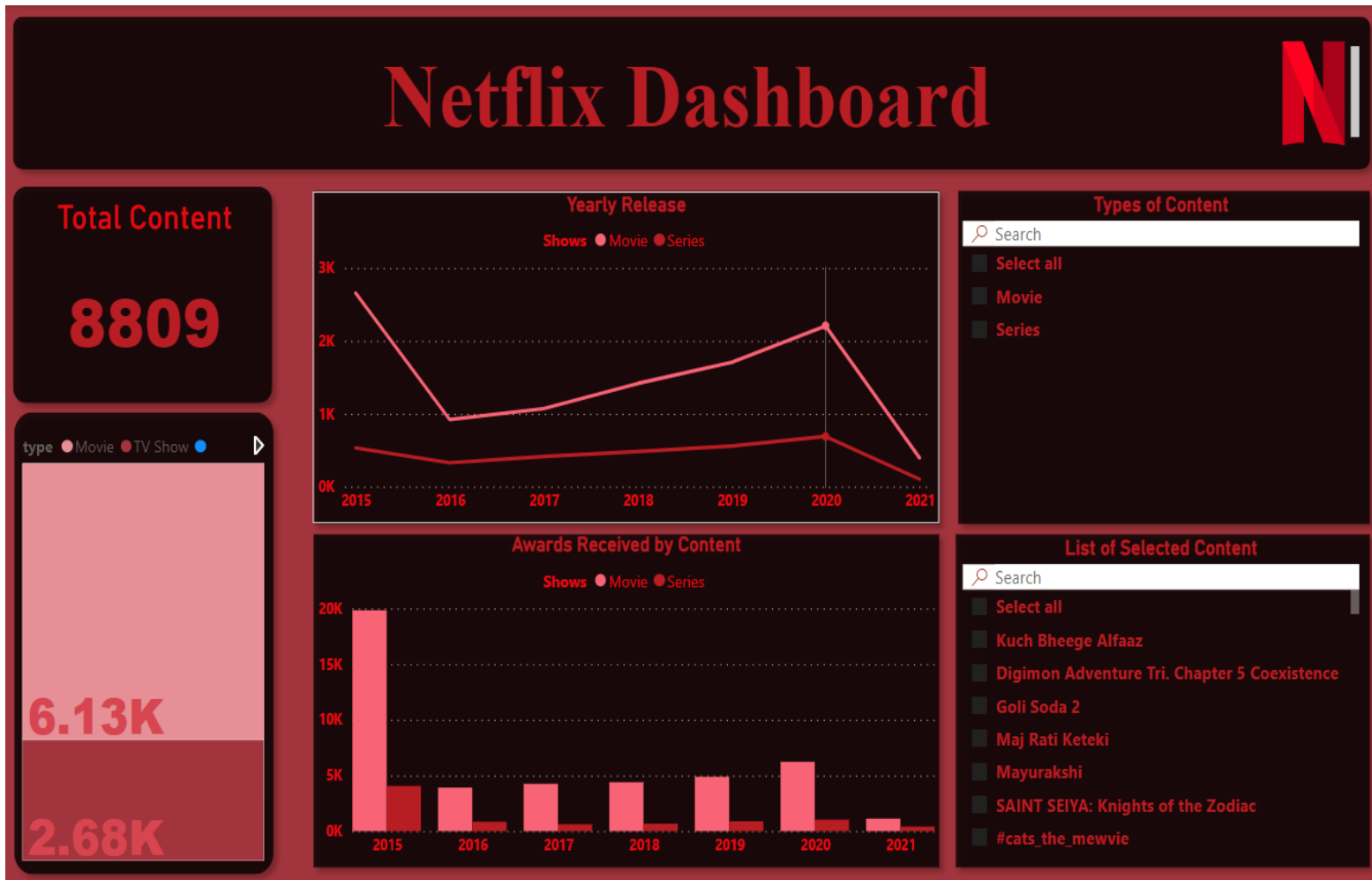


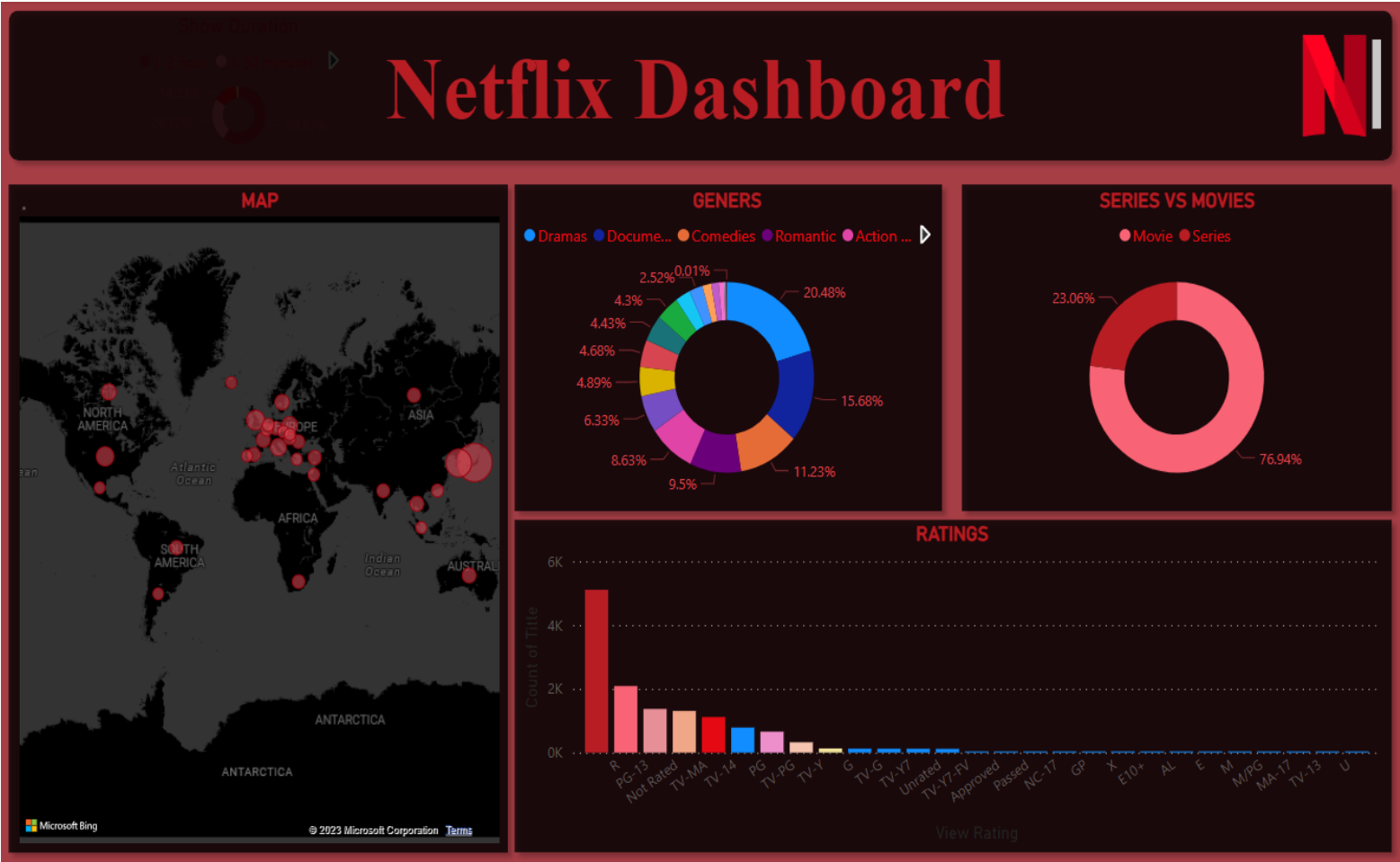
- f. A simple Column chart has been created as shown below.

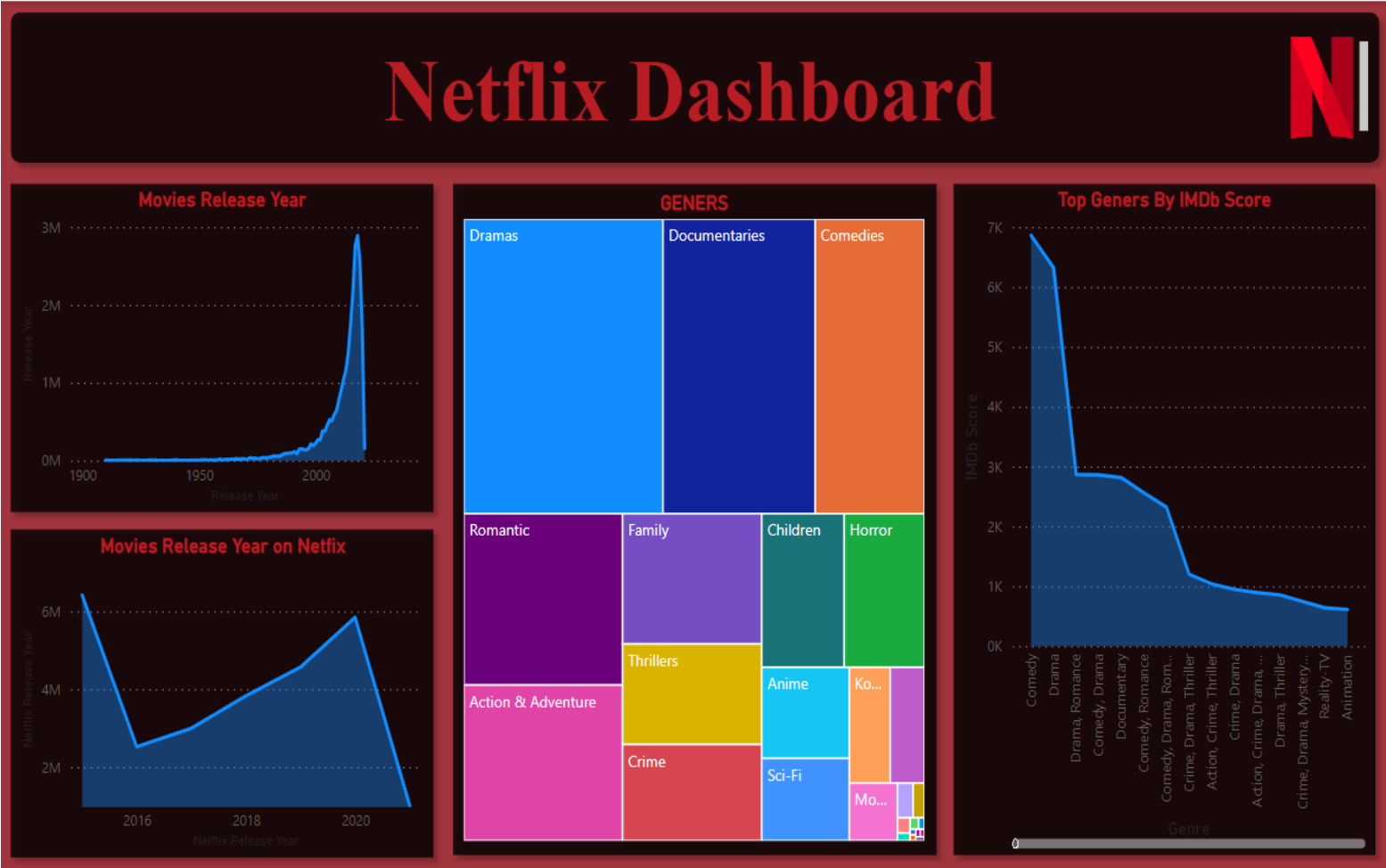


OUTPUT

Page 1







INPUT SCREEN IN PYTHON

```
import pandas as pd
import matplotlib.pyplot as plt import seaborn as sns
import plotly.express as px
from wordcloud import WordCloud, STOPWORDS from IPython.display import Image
import warnings warnings.filterwarnings("ignore")

data = pd.read_csv("/content/netflix_titles.csv")
netflix_df = pd.read_csv("/content/netflix_titles_original.csv")
netflix_df.head(10)
netflix_df.info()
netflix_df.shape
netflix_df.columns

netflix_df.describe()
netflix_df.isnull().sum()
sns.heatmap(netflix_df.isnull())
netflix_df.duplicated().sum()
netflix_df["type"].unique()
netflix_df["rating"].unique()
netflix_df["release_year"].unique()

data["listed_in"].nunique()
netflix_df["type"].value_counts()
sns.histplot(netflix_df["type"])
netflix_df["type"].value_counts().plot(kind = "pie", autopct = "%.2f")
plt.figure(figsize = (20, 5)) plt.style.use('seaborn-whitegrid')
ax = sns.countplot(x = 'rating', data = netflix_df)
plt.xlabel('Rating')
plt.ylabel('Count')
plt.xticks(rotation = 45)
for p in ax.patches:
ax.annotate(int(p.get_height()), (p.get_x()+0.25, p.get_height()+1), va = 'bottom', color = 'black')

plt.figure(figsize = (25, 8))
plt.style.use('seaborn-whitegrid')
ax = sns.countplot(x = 'release_year', data = netflix_df, palette = "Accent")
plt.xlabel('Year of Release')
```



```

plt.ylabel('Count')
plt.xticks(rotation = 45)
for p in ax.patches:
    ax.annotate(int(p.get_height()), (p.get_x()+0.25, p.get_height()+1), va = 'bottom', color = 'black')

```

```

netflix_df['date_added'] = pd.to_datetime(netflix_df['date_added']) netflix_df.dtypes
netflix_df.sample()
netflix_df["date_added"].hist()
netflix_df[(netflix_df["type"] == "TV Show") & (netflix_df["release_year"] == 2001)]
netflix_df[netflix_df["title"].isin(["Dark"])]
px.histogram(netflix_df, x = "release_year", color = "release_year")

```

```

plt.figure(figsize=(16, 4))
plt.style.use('seaborn-whitegrid')
plt.hist(netflix_df["release_year"], bins = 90, facecolor = '#F47340', edgecolor='#323232',
linewidth=0.5)
plt.title('Distribution')
plt.xlabel('Year of Release')
plt.ylabel('Count')
plt.show()

```

```

movies_only_in_japan = netflix_df[(netflix_df["type"] == "Movie") & (netflix_df["country"] ==
"Japan")]["title"]
movies_only_in_japan.count()
movies_only_in_japan.unique()
netflix_df["date_added"].dt.month
netflix_df["date_added_month"] = netflix_df["date_added"].dt.month.fillna(-1)
netflix_df["date_added_month"].unique()
netflix_df["date_added_month"] = netflix_df["date_added_month"].astype("Int64")
netflix_df.info()

```

```

plt.figure(figsize = (15, 5))
ax = sns.countplot(x = 'date_added_month', data = netflix_df)
plt.xlabel('Month Wise Content Added')
plt.ylabel('Count')
for p in ax.patches:
    ax.annotate(int(p.get_height()), (p.get_x()+0.25, p.get_height()+1), va = 'bottom', color = 'black')

```

```

netflix_df["date_added_day"] = netflix_df["date_added"].dt.day.fillna(-1)
netflix_df["date_added_day"].unique()
netflix_df["date_added_day"] = netflix_df["date_added_day"].astype("Int64")
netflix_df.info()

```

```

plt.figure(figsize = (18, 6))
ax = sns.countplot(x = 'date_added_day', data = netflix_df)
plt.xlabel('Day Wise Content Added')
plt.ylabel('Count')
for p in ax.patches:
ax.annotate(int(p.get_height()), (p.get_x()+0.25, p.get_height()+1), va = 'bottom', color = 'black')
country_count = netflix_df.copy()
country_count = pd.concat([country_count, netflix_df["country"].str.split(",", expand = True)], axis = 1)
country_count
country_count = country_count.melt(id_vars = ["type", "title"], value_vars = range(12), value_name = "country")
country_count = country_count[country_count["country"].notna()]
country_count
country_count["country"].nunique()
country_count["country"].value_counts()[:10]

country_count["country"] = country_count["country"].str.strip()
country_count["country"].value_counts()[:10]
plt.figure(figsize = (15, 7))
country = country_count["country"].value_counts()[:10]
sns.barplot(x = country, y = country.index, palette = "Accent")

colormap = plt.cm.plasma
sns.heatmap(pd.crosstab(data["rating"], data["type"]), cmap = colormap)
cast_count = netflix_df.copy()
cast_count = pd.concat([cast_count, netflix_df["cast"].str.split(",", expand = True)], axis = 1)
cast_count
plt.figure(figsize = (20, 12))
cast = cast_count["cast"].value_counts()[:10]
sns.barplot(x = cast, y = cast.index, palette = "gist_ncar")
updated_df = data.dropna()
updated_df[updated_df["cast"].str.contains("Bill Murray")]
listed_in_count["listed_in"].unique()
plt.figure(figsize = (20, 8))
listed_in = listed_in_count["listed_in"].value_counts()[:10]
sns.barplot(x = listed_in, y = listed_in.index, palette = "Dark2")

director_count = netflix_df.copy()
director_count = pd.concat([director_count, netflix_df["director"].str.split(",", expand = True)], axis = 1)
director_count["director"].nunique()

```

```
director_count["director"].value_counts()
plt.figure(figsize = (20, 8))
director_count = director_count["director"].value_counts()[:10]
sns.barplot(x = director_count, y = director_count.index, palette = "Spectral")
```

```
show_rating = netflix_df.groupby(['rating',
'type'])['rating'].count().reset_index(name='ratings_count')
show_rating = show_rating.set_index('rating')
rating_movie = netflix_df.groupby(['rating', 'type'])['rating'].count().reset_in
dex(name='ratings_count')
rating_movie = rating_movie.set_index('rating')
filter = rating_movie["type"] == "Movie"
```

```
# filtering data
rating_movie = rating_movie[filter].dropna()
rating_movie
```

```
rating_TV= netflix_df.groupby(['rating','type'])['rating'].count().reset_index(name='ratings_count')
rating_TV=rating_TV.set_index('rating')
rating_TV
filter = rating_TV["type"]=="TV Show"
# filtering data
rating_TV = rating_TV.where(filter).dropna()
rating_TV
```

OUTPUT SCREEN IN PYTHON



show_id	type		title	director	cast		country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN		United States	25-Sep-21	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabaiane, Thaban...		South Africa	24-Sep-21	2021	TV-MA	2 Seasons	Dramas	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabl...		NaN	24-Sep-21	2021	TV-MA	1 Season	Crime	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN		NaN	24-Sep-21	2021	TV-MA	1 Season	Documentaries	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...		India	24-Sep-21	2021	TV-MA	2 Seasons	Romantic	In a city of coaching centers known to train l...
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...		NaN	24-Sep-21	2021	TV-MA	1 Season	Horror	The arrival of a charismatic young priest brin...
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...		NaN	24-Sep-21	2021	PG	91 min	Family	Equestria's divided. But a bright-eyed hero be...
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...		United States, Ghana, Burkina Faso, United Kin...	24-Sep-21	1993	TV-MA	125 min	Dramas	On a photo shoot in Ghana, an American model s...
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...		United Kingdom	24-Sep-21	2021	TV-14	9 Seasons	Documentaries	A talented batch of amateur bakers face off in...
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...		United States	24-Sep-21	2021	PG-13	104 min	Dramas	A woman adjusting to life after a loss contend...



```
<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   8789 non-null  object
11  description 8807 non-null  object
dtypes: int64(1), object(11)
memory usage: 825.8+ KB
```



(8807,12)



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



```
<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	8789 non-null	object
11	description	8807 non-null	object

```
dtypes: int64(1), object(11)  
memory usage: 825.8+ KB
```

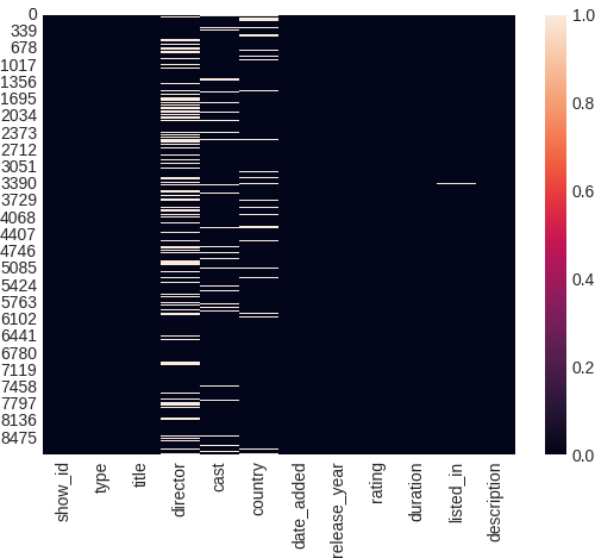


	release_year
count	8807.000000
mean	2014.180198
std	8.819312
min	1925.000000
25%	2013.000000
50%	2017.000000
75%	2019.000000
max	2021.000000



```
show_id      0  
type         0  
title        0  
director    2634  
cast         825  
country     831  
date_added  10  
release_year 0  
rating       4  
duration     3  
listed_in   18  
description  0  
dtype: int64
```

→



→

0

→

```
array(['Movie', 'TV Show'], dtype=object)
array(['PG-13', 'TV-MA', 'PG', 'TV-14', 'TV-PG', 'TV-Y', 'TV-Y7', 'R',
      'TV-G', 'G', 'NC-17', '74 min', '84 min', '66 min', 'NR', nan,
      'TV-Y7-FV', 'UR'], dtype=object)
```

→

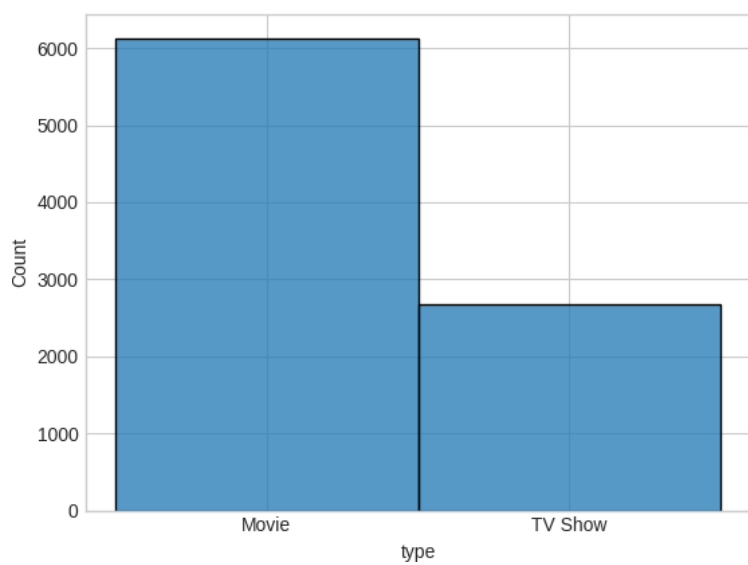
```
array([2020, 2021, 1993, 2018, 1996, 1998, 1997, 2010, 2013, 2017,
      1975,
      1978, 1983, 1987, 2012, 2001, 2014, 2002, 2003, 2004, 2011, 2008,
      2009, 2007, 2005, 2006, 1994, 2015, 2019, 2016, 1982, 1989, 1990,
      1991, 1999, 1986, 1992, 1984, 1980, 1961, 2000, 1995, 1985, 1976,
      1959, 1988, 1981, 1972, 1964, 1945, 1954, 1979, 1958, 1956, 1963,
      1970, 1973, 1925, 1974, 1960, 1966, 1971, 1962, 1969, 1977, 1967,
      1968, 1965, 1946, 1942, 1955, 1944, 1947, 1943])
```

→

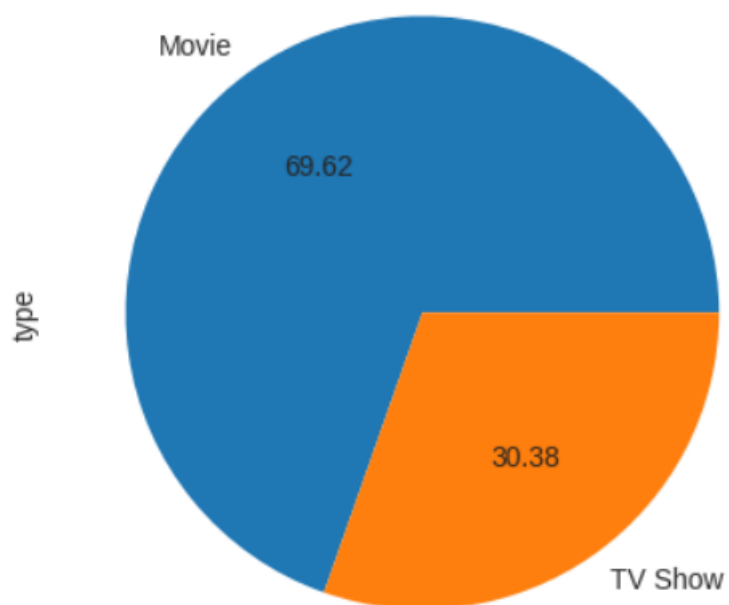
35

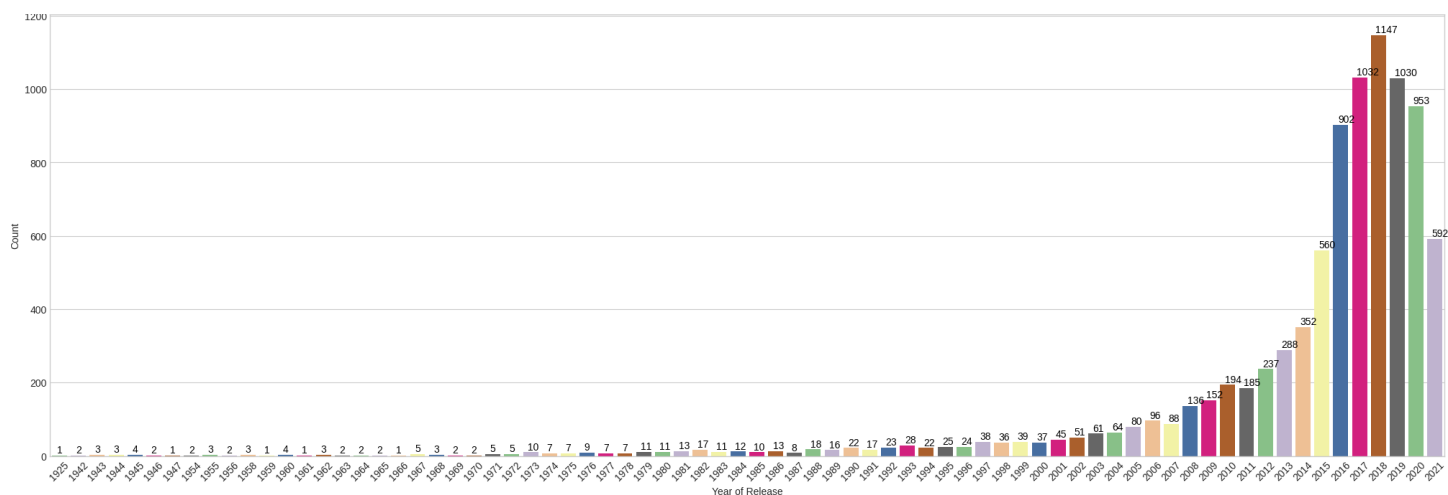
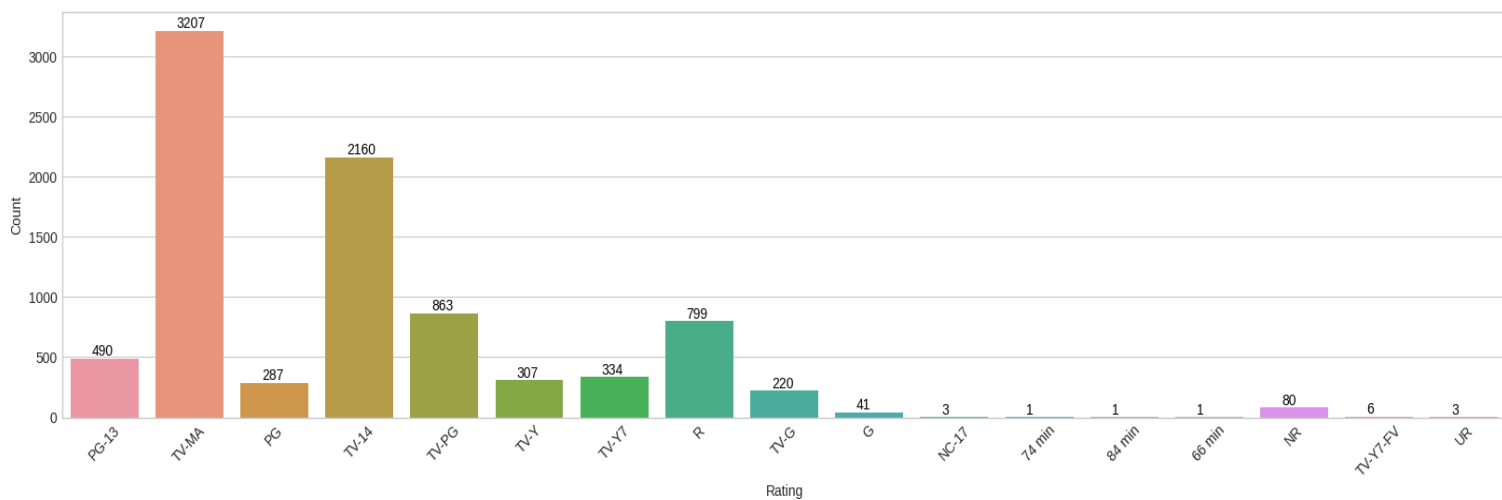
→

```
Movie      6131
TV Show    2676
Name: type, dtype: int64
```



<Axes: ylabel='type'>

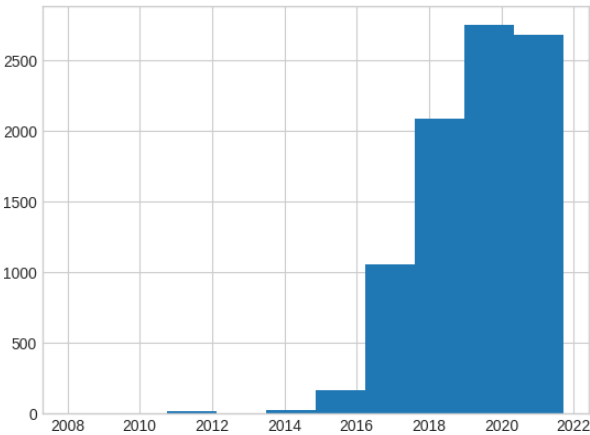




```
show_id      object
type         object
title        object
director     object
cast         object
country      object
date_added   datetime64[ns]
release_year int64
rating       object
duration     object
listed_in    object
description  object
dtype: object
```



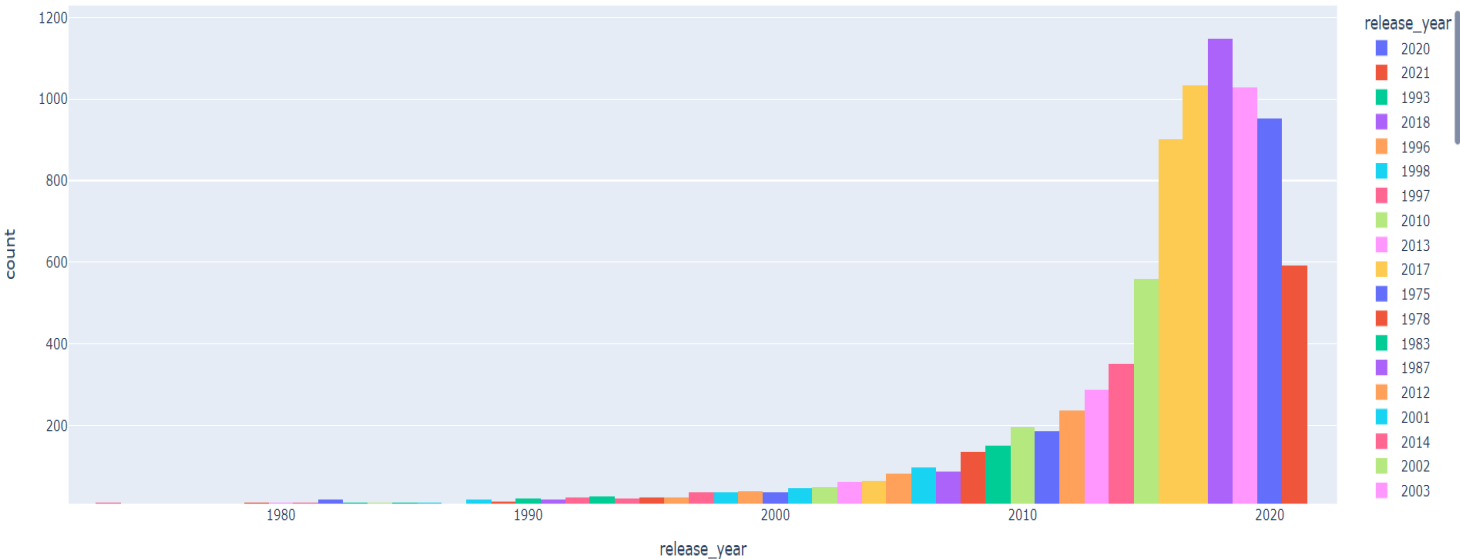

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
141	s142	Movie	Extraction	Steven C. Miller	Bruce Willis, Kellan Lutz, Gina Carano, D.B. S...	United States, United Kingdom, Canada	2021-09-01	2015	R	82 min	Action & Adventure	When a retired CIA agent is kidnapped, his son...

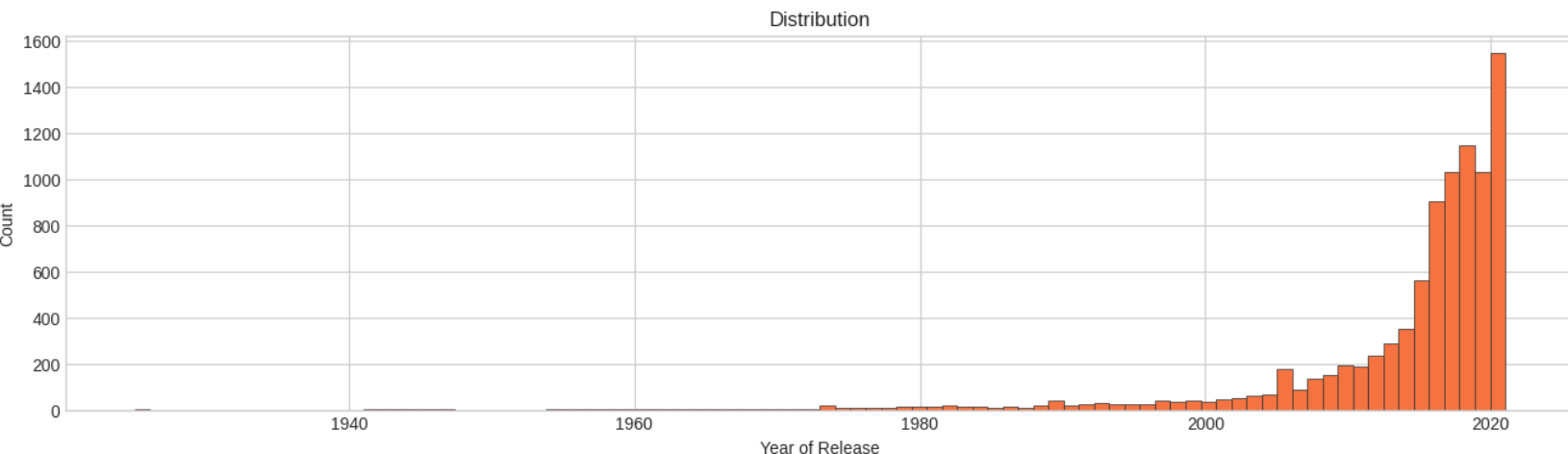


	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	
	657	s658	TV Show	The Mole	NaN	Anderson Cooper	United States	2021-06-21	2001	TV-14	2 Seasons	Documentaries	In this competition show, contestants try to e...
	5672	s5673	TV Show	InuYasha	NaN	Kappei Yamaguchi, Satsuki Yukino, Kumiko Watan...	Japan	2016-12-15	2001	TV-14	2 Seasons	Anime	Yanked down a well, schoolgirl Kagome Higurash...
	5811	s5812	TV Show	Behind Enemy Lines	NaN	NaN	United States	2016-08-15	2001	TV-14	1 Season	Documentaries	After dire setbacks in 1940, Winston Churchill...
	7782	s7783	TV Show	Power Rangers Time Force	NaN	Jason Faunt, Erin Cahill, Kevin Kleinberg, Deb...	United States	2016-01-01	2001	TV-Y7	1 Season	Children	The team launches a new campaign against evil ...
	8214	s8215	TV Show	The Blue Planet: A Natural History of the Oceans	Alastair Fothergill	David Attenborough	United Kingdom	2015-10-10	2001	TV-G	1 Season	Documentaries	David Attenborough narrates this definitive ex...



show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	
2327	s2328	TV Show	Dark	NaN	Louis Hofmann, Oliver Masucci, Jördis Triebel,...	Germany, United States	2020-06-27	2020	TV-MA	3 Seasons	Crime	A missing child sets four families on a franti...





76



```
array(['InuYasha the Movie 2: The Castle Beyond the Looking Glass',  
      'InuYasha the Movie 3: Swords of an Honorable Ruler',  
      'InuYasha the Movie 4: Fire on the Mystic Island',  
      'InuYasha the Movie: Affections Touching Across Time',  
      'Naruto Shippuden the Movie: Blood Prison',  
      'Naruto Shippûden  
  
the Movie: Bonds',  
      'Naruto Shippûden the Movie: The Will of Fire',  
      'Naruto Shippuden: The Movie',  
      'Naruto Shippuden: The Movie: The Lost Tower',  
      'Naruto the Movie 2: Legend of the Stone of Gelel',  
      'Naruto the Movie 3: Guardians of the Crescent Moon Kingdom',  
      'Naruto the Movie: Ninja Clash in the Land of Snow', 'HOMUNCULUS',  
      'GANTZ:0', 'Ride or Die',  
      'Seven Souls in the Skull Castle: Season Flower',  
      'Any Crybabies Around?', 'Blue, Painful, Fragile', 'MOTHER',  
      'DTC Yukemuri Junjo Hen From High & Low', 'High & Low The Movie',  
      'High & Low The Movie 2 / End of Sky',  
      'High & Low The Movie 3 / Final Mission',  
      'High & Low The Red Rain', 'Road To High & Low',  
      'Fate/Grand Order -First Order-', 'Romance Doll', 'A Whisker Away',  
      'Okko's Inn', 'Modest Heroes: Ponoc Short Films Theatre',  
      'Sol Levante',  
      'Violet Evergarden: Eternity and the Auto Memory Doll',  
      'Lu Over the Wall', 'Pokémon: Mewtwo Strikes Back - Evolution',  
      'Dragon Quest Your Story', '37 Seconds', 'NiNoKuni',  
      'Pokémon the Movie: Power of Us', 'Hot Gimmick: Girl Meets Boy',  
      'Fireworks', 'No Game No Life: Zero', 'Girls und Panzer der Film',  
      'K-on! the movie', 'The Garden of Words', 'The Forest of Love',  
      'TUNA GIRL', 'Hikaru Utada Laughter in the Dark Tour 2018',  
      'EVANGELION: DEATH (TRUE)2', 'The End of Evangelion',
```

```
'A Silent Voice', 'River's Edge', 'GODZILLA The Planet Eater',
'Pokémon the Movie: I Choose You!',
'The Seven Deadly Sins the Movie: Prisoners of the Sky',
'Aggretsuko: We Wish You a Metal Christmas',
'GODZILLA City on the Edge of Battle', 'Expelled from Paradise',
'Lupin the 3rd: The Castle of Cagliostro: Special Edition',
'In This Corner of the World', 'FullMetal Alchemist',
'Berserk: The Golden Age Arc II - The Battle for Doldrey',
'Berserk: The Golden Age Arc III - The Advent', 'BLAME!',
'Berserk: The Golden Age Arc I - The Egg of the King',
'Bleach The Movie: Fade to Black', 'Bleach the Movie: Hell Verse',
'Ghost Pain', 'Ghost Tears', 'Ghost Whispers', 'Golden Time',
'Inuyasha the Movie - L'isola del fuoco scarlatto',
'Inuyasha the Movie - La spada del dominatore del mondo',
'InuYasha: The Movie 2: The Castle Beyond the Looking Glass',
'Mary and the Witch's Flower', 'Miss Hokusai',
'Naruto Shippuden : Blood Prison'], dtype=object)
```

➔

```
0      9.0
1      9.0
2      9.0
3      9.0
4       9.
      0

...
8802  11.0
8803   7.0
8804  11.0
8805   1.0
8806   3.0
```

➔

```
Name: date_added, Length: 8807, dtype: float64
array([ 9.,  8.,  7.,  6.,  5.,  4.,  3.,  2.,  1., 12., 11., 10., -1.] )
```

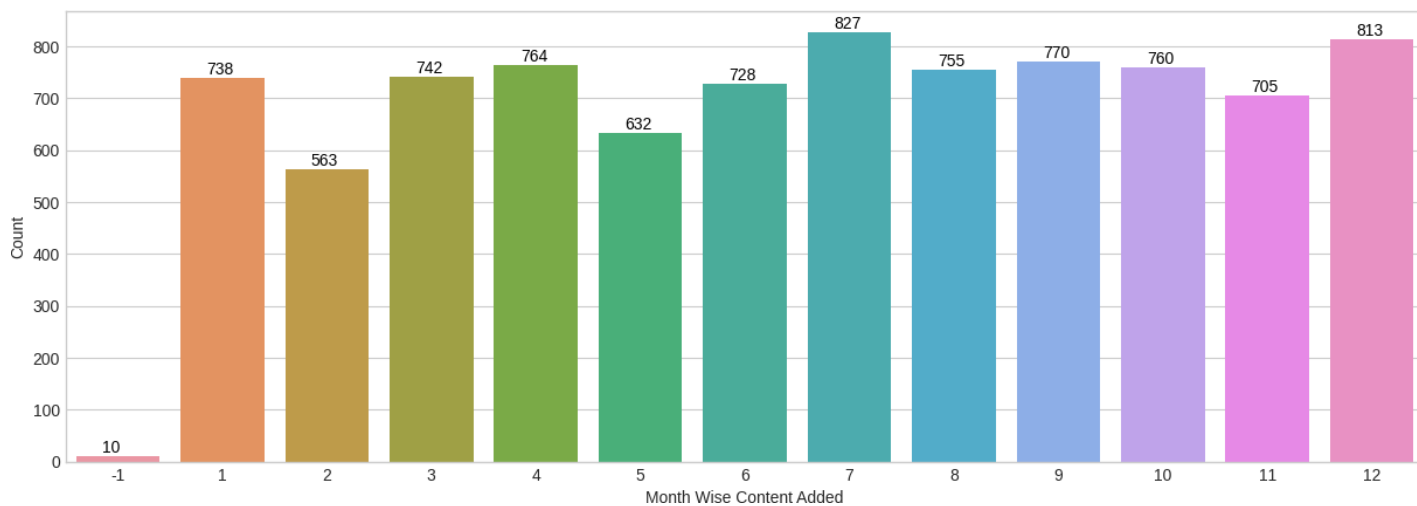
➔

```
<class 'pandas.core.frame.DataFrame'>RangeIndex: 8807
entries, 0 to 8806
```

```
Data columns (total 13 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	datetime64[ns]
7	release_year	8807 non-null	int64
8	rating	8803 non-null	object
9	duration	8804 non-null	object
10	listed_in	8789 non-null	object
11	description	8807 non-null	object
12	date_added_month	8807 non-null	Int64

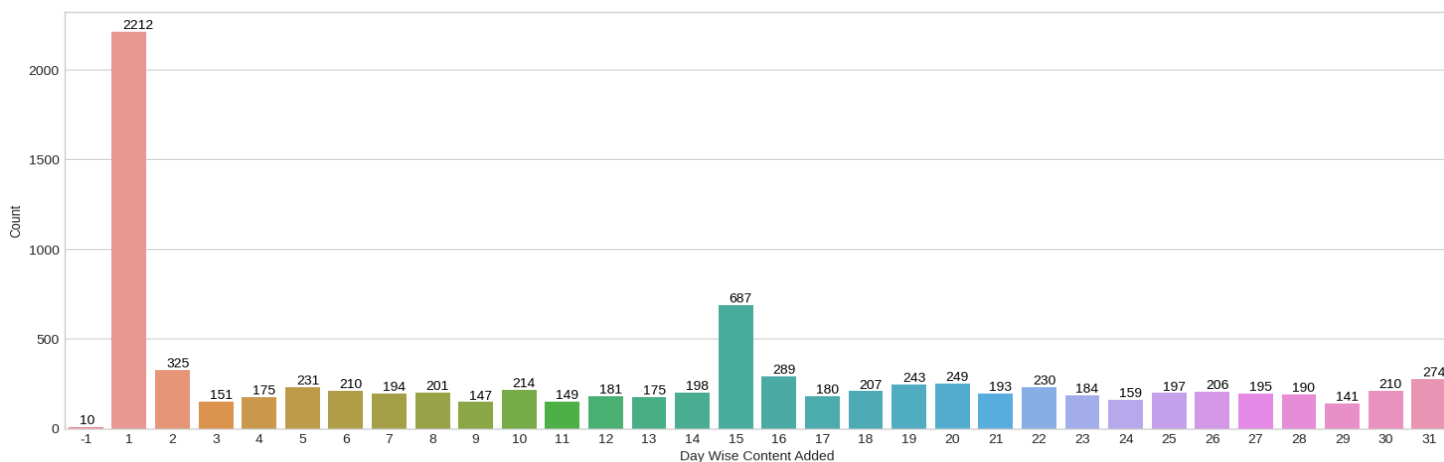
```
dtypes: Int64(1), datetime64[ns](1), int64(1), object(10)
memory usage: 903.2+ KB
```



```
array([25., 24., 23., 22., 21., 20., 19., 17., 16., 15., 14., 11., 10.,
       9., 8., 7., 6., 5., 4., 3., 2., 1., 31., 29., 28., 27.,
       26., 18., 13., 12., 30., -1.])
```



```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 14 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   datetime64[ns]
7   release_year          8807 non-null   int64
8   rating                8803 non-null   object
9   duration              8804 non-null   object
10  listed_in             8789 non-null   object
11  description            8807 non-null   object
12  date_added_month      8807 non-null   Int64
13  date_added_day        8807 non-null   Int64
dtypes: Int64(2), datetime64[ns](1), int64(1), object(10)
memory usage: 980.6+ KB
```





show_id	type	title	director	cast	country	date_added	release_year	rating	duration	...	2	3	4	5	6	7	8	9	10	11
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90 min	...	None	None	None	None	None	None	None	None	None
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	...	None	None	None	None	None	None	None	None	None
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajjila, Tracy Gotoas, Samuel Jouy, Nabl...	NaN	2021-09-24	2021	TV-MA	1 Season	...	None	None	None	None	None	None	None	None	None
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	2021-09-24	2021	TV-MA	1 Season	...	None	None	None	None	None	None	None	None	None
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021	TV-MA	2 Seasons	...	None	None	None	None	None	None	None	None	None
...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	2019-11-20	2007	R	158 min	...	None	None	None	None	None	None	None	None	None
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	2019-07-01	2018	TV-Y7	2 Seasons	...	None	None	None	None	None	None	None	None	None
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01	2009	R	88 min	...	None	None	None	None	None	None	None	None	None
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88 min	...	None	None	None	None	None	None	None	None	None
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111 min	...	None	None	None	None	None	None	None	None	None



	type	title	variable	country
0	Movie	Dick Johnson Is Dead	0	United States
1	TV Show	Blood & Water	0	South Africa
4	TV Show	Kota Factory	0	India
7	Movie	Sankofa	0	United States
8	TV Show	The Great British Baking Show	0	United Kingdom
...
78859	Movie	The Look of Silence	8	Germany
85496	Movie	Barbecue	9	Sweden
87666	Movie	The Look of Silence	9	Netherlands



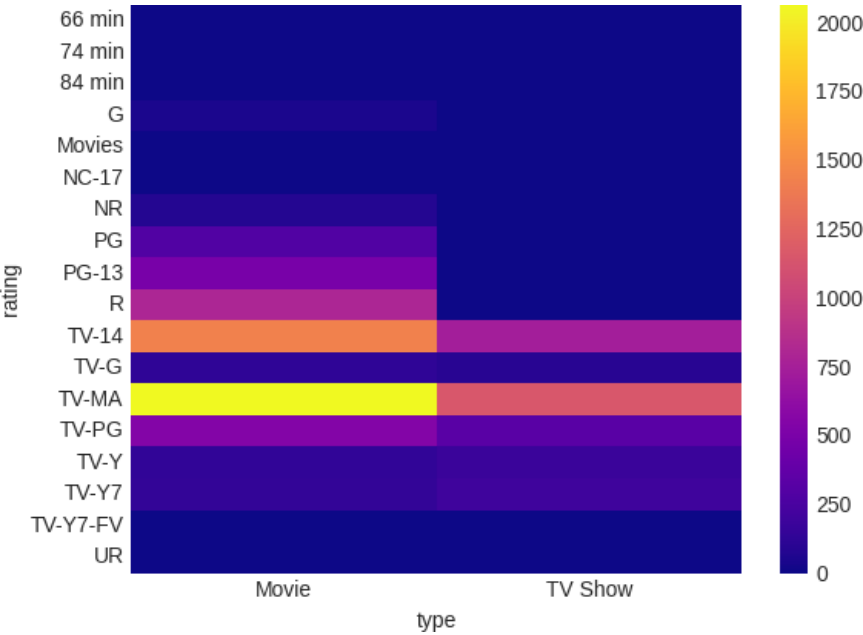
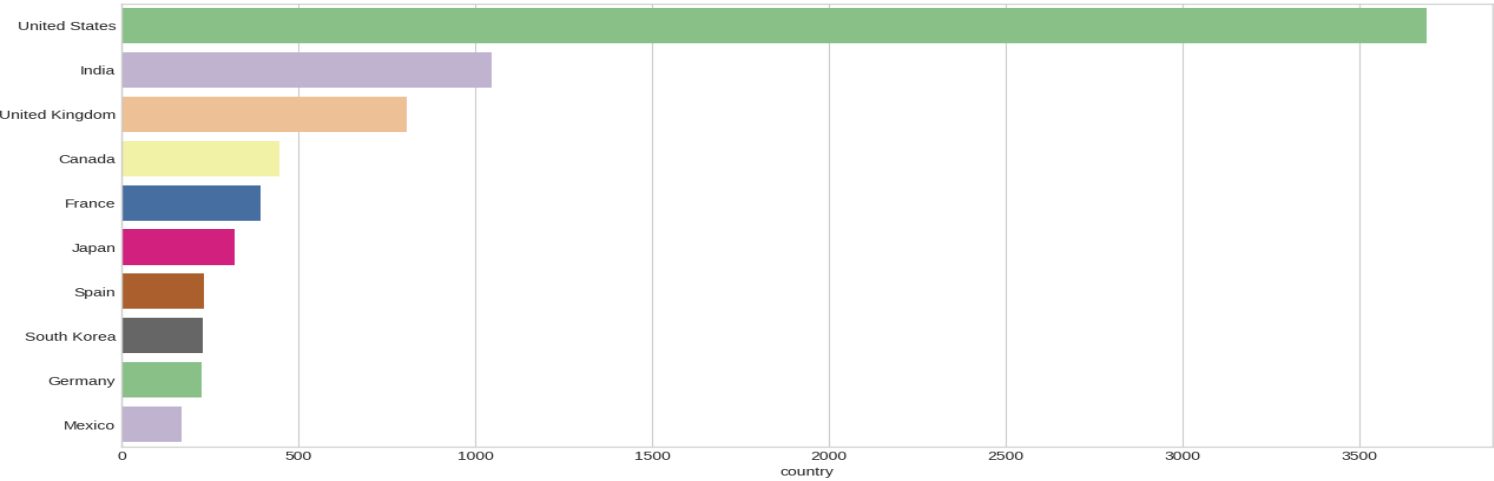
197



United States	3211
India	1008
United Kingdom	628
United States	479
Canada	271
Japan	259
France	212
South Korea	211
Spain	181
France	181



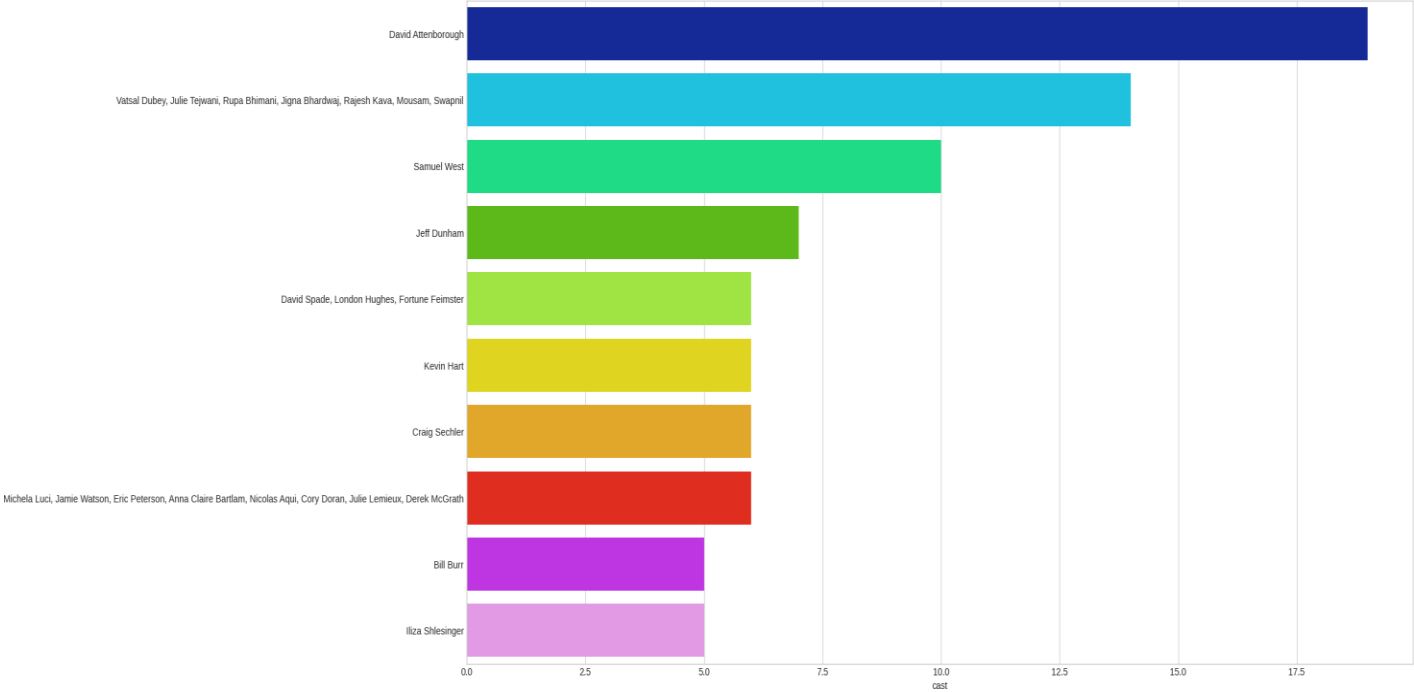
United States	3690
India	1046
United Kingdom	806
Canada	445
France	393
Japan	318
Spain	232
South Korea	231
Germany	226
Mexico	169





	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	...	40	41	42	43	44	45	46	47	48	49
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020	PG-13	90 min	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	s2	TV Show	Blood & Water	NaN	Ama Oamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021	TV-MA	2 Seasons	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabil...	NaN	2021-09-24	2021	TV-MA	1 Season	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	2021-09-24	2021	TV-MA	1 Season	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021	TV-MA	2 Seasons	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	2019-11-20	2007	R	158 min	...	None	None	None	None	None	None	None	None	None	None
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	2019-07-01	2018	TV-Y7	2 Seasons	...	None	None	None	None	None	None	None	None	None	None
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01	2009	R	88 min	...	None	None	None	None	None	None	None	None	None	None
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006	PG	88 min	...	None	None	None	None	None	None	None	None	None	None
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015	TV-14	111 min	...	None	None	None	None	None	None	None	None	None	None

8807 rows × 64 columns

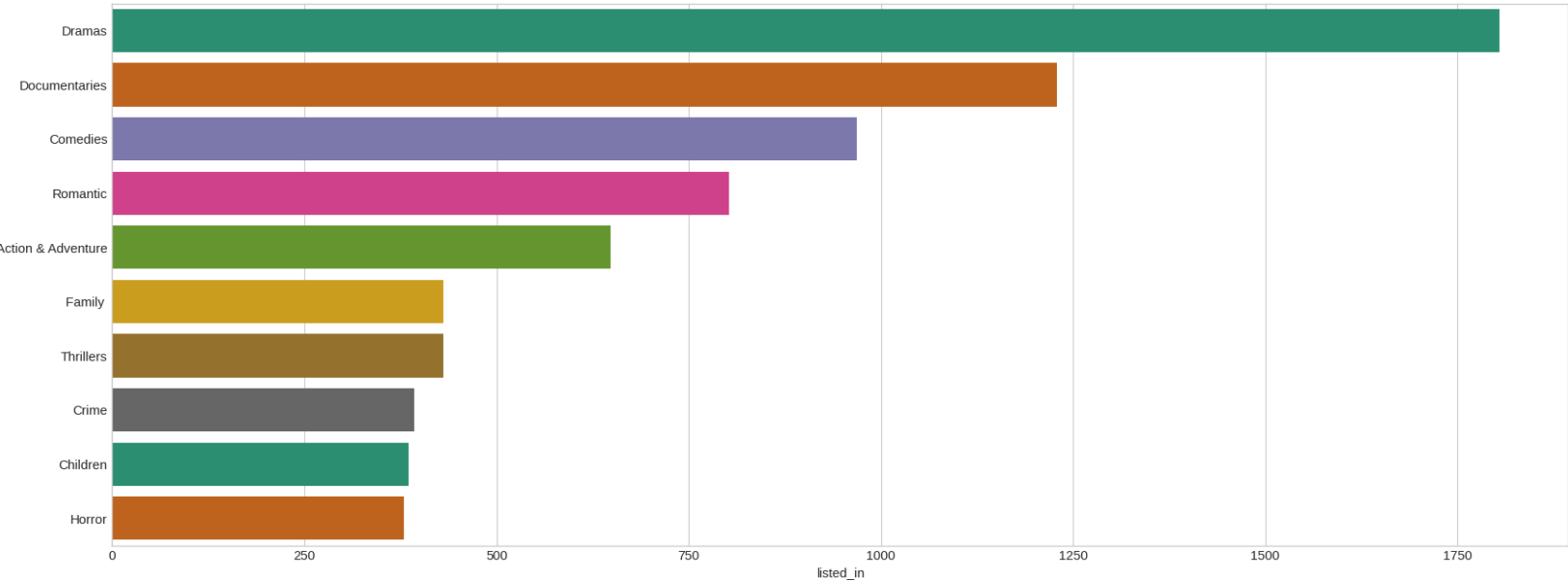




	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
494	s495	Movie	Rock the Kasbah	Barry Levinson	Bill Murray, Kate Hudson, Zoëy Deschanel, Dan...	United States	8-Jul-21	2015	R	106 min	Comedies	When a has-been music producer gets stuck in A...
566	s567	Movie	Charlie's Angels	McG	Cameron Diaz, Drew Barrymore, Lucy Liu, Bill M...	United States, Germany	1-Jul-21	2000	PG-13	98 min	Action & Adventure	A tight-knit trio of specially trained agents ...
5881	s5882	Movie	A Very Murray Christmas	Sofia Coppola	Bill Murray, Miley Cyrus, George Clooney, Chrl...	United States	4-Dec-15	2015	TV-14	57 min	Comedies	Bill Murray rounds up an all-star cast for an ...
6036	s6037	Movie	A Glimpse Inside the Mind of Charles Swan III	Roman Coppola	Charlie Sheen, Jason Schwartzman, Bill Murray,...	United States	15-Jul-18	2012	R	86 min	Comedies	When his girlfriend walks out and leaves him a...
6842	s6843	Movie	Get Smart	Peter Segal	Steve Carell, Anne Hathaway, Dwayne Johnson, A...	United States	1-Apr-19	2008	PG-13	110 min	Action & Adventure	When the identities of secret agents are compr...
7210	s7211	Movie	Kingpin	Peter Farrelly, Bobby Farrelly	Woody Harrelson, Randy Quaid, Vanessa Angel, B...	United States	1-Jan-20	1996	PG-13	114 min	Comedies	After a bowler loses his hand, his career take...
8105	s8106	Movie	Stripes	Ivan Reitman	Bill Murray, Harold Ramis, Warren Oates, P.J. ...	United States	1-Sep-19	1981	R	106 min	Classic	After losing everything, an indolent sad sack ...
8210	s8211	Movie	The Bill Murray Stories: Life Lessons Learned ...	Tommy Avalone	Tommy Avalone, Bill Murray, Joel Murray, Pete...	United States	31-Dec-18	2018	TV-MA	72 min	Documentaries	This documentary highlights spontaneous encoun...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	1-Nov-19	2009	R	88 min	Horror	Looking to survive in a world taken over by zo...



```
array(['Documentaries', 'Dramas', 'Crime ', 'Romantic', 'Horror',  
      'Family ', 'Crime', 'Thrillers', 'Action & Adventure', 'Comedies',  
      'Sci-Fi', 'Children', 'Classic', 'Anime', 'Documentaries ',  
      'Korean', 'Family', 'Documentaries ', 'TV Shows',  
      'Action & Adventure ', 'Romantic ', 'Movies', 'Comedies ',  
      ' Comedies', 'Classic ', 'Family, ', 'Anime ', ' Documentaries',  
      'Children ', 'TVHorror ', ' Horror', 'CrimeChildren',  
      'Classic Dramas', 'Anime Features', 'ClassicDramas'], dtype=object)
```

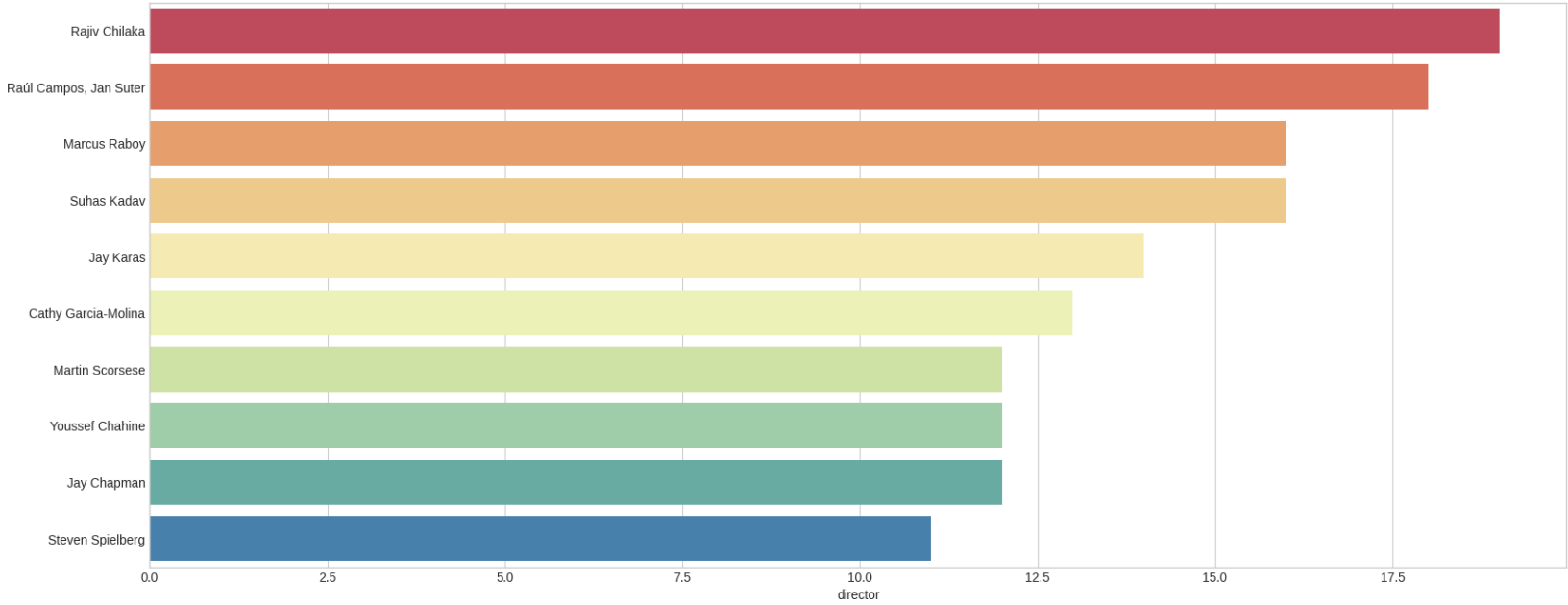




4528



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
Mozez Singh	1





rating	type	ratings_count
66 min	Movie	1
74 min	Movie	1
84 min	Movie	1
G	Movie	41
NC-17	Movie	3
NR	Movie	75
PG	Movie	287
PG-13	Movie	490
R	Movie	797
TV-14	Movie	1427
TV-G	Movie	126
TV-MA	Movie	2062
TV-PG	Movie	540
TV-Y	Movie	131
TV-Y7	Movie	139
TV-Y7-FV	Movie	5



rating	type	ratings_count
NR	TV Show	5.0
R	TV Show	2.0
TV-14	TV Show	733.0
TV-G	TV Show	94.0
TV-MA	TV Show	1145.0
TV-PG	TV Show	323.0
TV-Y	TV Show	176.0
TV-Y7	TV Show	195.0
TV-Y7-FV	TV Show	1.0