

▼ Data Analysis

▼ 1. Load the file

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
import plotly.express as px
netflix_df=pd.read_csv('/content/netflix_titles.csv')
```

▼ 2. Print first 5 rows of data

```
netflix_df.head()
```

|     | show_id | type    | title                | director        | cast  | country       | date_added         | release_year | rating | duration  | listed_in   | description                                       |
|-----|---------|---------|----------------------|-----------------|---|---------------|--------------------|--------------|--------|-----------|---|---|
| 0   | s1      | Movie   | Dick Johnson Is Dead | Kirsten Johnson | NaN   | United States | September 25, 2021 | 2020         | PG-13  | 90 min    | Documentaries                                     | As her father nears the end of his life, filmm... |
| 1   | s2      | TV Show | Blood & Water        | NaN             | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | South Africa  | September 24, 2021 | 2021         | TV-MA  | 2 Seasons | International TV Shows, TV Dramas, TV Mysteries   | After crossing paths at a party, a Cape Town t... |
| 2   | s3      | TV Show | Ganglands            | Julien Leclercq | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... | NaN           | September 24, 2021 | 2021         | TV-MA  | 1 Season  | Crime TV Shows, International TV Shows, TV Act... | To protect his family from a powerful drug lor... |
| ... | ...     | ...     | ...                  | ...             | ...   | ...           | ...                | ...          | ...    | ...       | ...   | Feuds,  |

▼ 3. Print last 5 rows of data

```
netflix_df.tail()
```

|      | show_id | type    | title       | director        | cast  | country       | date_added        | release_year | rating | duration  | listed_in                              | description                                       |
|------|---------|---------|-------------|-----------------|---|---------------|-------------------|--------------|--------|-----------|--|---|
| 8802 | s8803   | Movie   | Zodiac      | David Fincher   | Mark Ruffalo, Jake Gyllenhaal, Robert Downey J... | United States | November 20, 2019 | 2007         | R      | 158 min   | Cult Movies, Dramas, Thrillers         | A political cartoonist, a crime reporter and a... |
| 8803 | s8804   | TV Show | Zombie Dumb | NaN             | NaN   | NaN           | July 1, 2019      | 2018         | TV-Y7  | 2 Seasons | Kids' TV, Korean TV Shows, TV Comedies | While living alone in a spooky town, a young g... |
| 8804 | s8805   | Movie   | Zombieland  | Ruben Fleischer | Jesse Eisenberg, Woody Harrelson, Emma Stone, ... | United States | November 1, 2019  | 2009         | R      | 88 min    | Comedies, Horror Movies                | Looking to survive in a world taken over by zo... |
| ...  | ...     | ...     | ...         | ...             | Tim Allen,  | ...           | ...               | ...          | ...    | ...       | ...                                    | ...   |

4. Print last 5 rows of data

```
netflix_df.tail()
```

|      | show_id | type  | title      | director        | cast  | country                      | date_added        | release_year | rating | duration | listed_in                               | description                                       |
|------|---------|-------|------------|-----------------|---|------------------------------|-------------------|--------------|--------|----------|---|---|
| 8801 | s8802   | Movie | Zinzana    | Majid Al Ansari | Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ... | United Arab Emirates, Jordan | March 9, 2016     | 2015         | TV-MA  | 96 min   | Dramas, International Movies, Thrillers | Recovering alcoholic Talal wakes up inside a s... |
| 8802 | s8803   | Movie | Zodiac     | David Fincher   | Mark Ruffalo, Jake Gyllenhaal, Robert Downey J... | United States                | November 20, 2019 | 2007         | R      | 158 min  | Cult Movies, Dramas, Thrillers          | A political cartoonist, a crime reporter and a... |
| 8804 | s8805   | Movie | Zombieland | Ruben Fleischer | Jesse Eisenberg, Woody Harrelson, Emma Stone, ... | United States                | November 1, 2019  | 2009         | R      | 88 min   | Comedies, Horror Movies                 | Looking to survive in a world taken over by zo... |
|      |         |       |            |                 | Tim Allen,  |                              |                   |              |        |          |   |   |

5. Cleaning the data for missing values, null values etc..

```
netflix_df.dropna(inplace=True)
netflix_df
```

|      | show_id | type    | title                         | director            | cast  | country   | date_added         | release_year | rating | duration  | listed_in  | description                                       |
|------|---------|---------|-------------------------------|---------------------|---|---|--------------------|--------------|--------|-----------|--|---|
| 7    | s8      | Movie   | Sankofa                       | Haile Gerima        | Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D... | United States, Ghana, Burkina Faso, United Kin... | September 24, 2021 | 1993         | TV-MA  | 125 min   | Dramas, Independent Movies, International Movies | On a photo shoot in Ghana, an American model s... |
| 8    | s9      | TV Show | The Great British Baking Show | Andy Devonshire     | Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho... | United Kingdom                                    | September 24, 2021 | 2021         | TV-14  | 9 Seasons | British TV Shows, Reality TV                     | 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                                     | September 24, 2021 | 2021         | PG-13  | 104 min   | Comedies, Dramas                                 | A woman adjusting to life after a loss contend... |
| 12   | s13     | Movie   | Je Suis Karl                  | Christian Schwochow | Luna Wedler, Jannis Niewöhner, Milan Peschel, ... | Germany, Czech Republic                           | September 23, 2021 | 2021         | TV-MA  | 127 min   | Dramas, International Movies                     | After most of her family is murdered in a terr... |
| 24   | s25     | Movie   | Jeans                         | S. Shankar          | Prashanth, Aishwarya Rai Bachchan, Sri Lakshmi... | India   | September 21, 2021 | 1998         | TV-14  | 166 min   | Comedies, International Movies, Romantic Movies  | When the father of the man she loves insists t... |
| ...  | ...     | ...     | ...                           | ...                 | ...   | ...   | ...                | ...          | ...    | ...       | ...  | ...   |
| 8801 | s8802   | Movie   | Zinzana                       | Majid Al Ansari     | Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ... | United Arab Emirates, Jordan                      | March 9, 2016      | 2015         | TV-MA  | 96 min    | Dramas, International Movies, Thrillers          | Recovering alcoholic Talal wakes up inside a s... |
| 8802 | s8803   | Movie   | Zodiac                        | David Fincher       | Mark Ruffalo, Jake Gyllenhaal                     | United States                                     | November 20, 2019  | 2007         | R      | 158 min   | Cult Movies, Dramas, Thrillers                   | A political cartoonist, a crime reporter and      |

## 6. Print info on the dataset

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

## 7. Get some description of the data - ONLY release\_year is NUMBER(S) VALUE(S)

```
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  |

## 8. Columns in the Dataset

```
netflix_df.columns
```

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

## 9. Shape of The data frame

```
netflix_df.shape
```

```
(8807, 12)
```

## 10. How many years of Data (1966-2021) (release\_year)

```
netflix_df.release_year.value_counts()
```

|              | count |
|--------------|-------|
| release_year |       |
| 2017         | 657   |
| 2018         | 648   |
| 2016         | 577   |
| 2019         | 519   |
| 2020         | 442   |
| ...          | ...   |
| 1946         | 1     |
| 1961         | 1     |
| 1942         | 1     |
| 1947         | 1     |
| 1944         | 1     |

72 rows × 1 columns

dtype: int64

## 11. MOVIES vs TV SHOWS(type)

```
netflix_df.type.value_counts()
```



|         | count |
|---------|-------|
| type    |       |
| Movie   | 6131  |
| TV Show | 2676  |

dtype: int64

## 12. Check for null values in all columns of dataset

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



|              | 0 |
|--------------|---|
| show_id      | 0 |
| type         | 0 |
| title        | 0 |
| director     | 0 |
| cast         | 0 |
| country      | 0 |
| date_added   | 0 |
| release_year | 0 |
| rating       | 0 |
| duration     | 0 |
| listed_in    | 0 |
| description  | 0 |

dtype: int64

## 13. Replace NAN value(s) with TV-MA

```
netflix_df.replace(np.nan, 'TV-MA', inplace=True)  
netflix_df
```

|      | show_id | type    | title                         | director            | cast  | country   | date_added         | release_year | rating | duration  | listed_in  | description                                       |
|------|---------|---------|-------------------------------|---------------------|---|---|--------------------|--------------|--------|-----------|--|---|
| 7    | s8      | Movie   | Sankofa                       | Haile Gerima        | Kofi Ghanaba, Oyafulmike Ogunlano, Alexandra D... | United States, Ghana, Burkina Faso, United Kin... | September 24, 2021 | 1993         | TV-MA  | 125 min   | Dramas, Independent Movies, International Movies | On a photo shoot in Ghana, an American model s... |
| 8    | s9      | TV Show | The Great British Baking Show | Andy Devonshire     | Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho... | United Kingdom                                    | September 24, 2021 | 2021         | TV-14  | 9 Seasons | British TV Shows, Reality TV                     | 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                                     | September 24, 2021 | 2021         | PG-13  | 104 min   | Comedies, Dramas                                 | A woman adjusting to life after a loss contend... |
| 12   | s13     | Movie   | Je Suis Karl                  | Christian Schwochow | Luna Wedler, Jannis Niewöhner, Milan Peschel, ... | Germany, Czech Republic                           | September 23, 2021 | 2021         | TV-MA  | 127 min   | Dramas, International Movies                     | After most of her family is murdered in a terr... |
| 24   | s25     | Movie   | Jeans                         | S. Shankar          | Prashanth, Aishwarya Rai Bachchan, Sri Lakshmi... | India   | September 21, 2021 | 1998         | TV-14  | 166 min   | Comedies, International Movies, Romantic Movies  | When the father of the man she loves insists t... |
| ...  | ...     | ...     | ...                           | ...                 | ...   | ...   | ...                | ...          | ...    | ...       | ...  | ...   |
| 8801 | s8802   | Movie   | Zinzana                       | Majid Al Ansari     | Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ... | United Arab Emirates, Jordan                      | March 9, 2016      | 2015         | TV-MA  | 96 min    | Dramas, International Movies, Thrillers          | Recovering alcoholic Talal wakes up inside a s... |
| 8802 | s8803   | Movie   | Zodiac                        | David Fincher       | Mark Ruffalo, Jake Gyllenhaal, ...                | United States                                     | November 20, 2019  | 2007         | R      | 158 min   | Cult Movies, Dramas, Thrillers                   | A political cartoonist, a crime reporter and      |


#### 14. Replace NAN value(s) with United States

```
netflix_df.replace(np.nan, 'United States', inplace=True)
netflix_df
```

|      | show_id | type    | title                         | director            | cast  | country   | date_added         | release_year | rating | duration  | listed_in  | description                                       |
|------|---------|---------|-------------------------------|---------------------|---|---|--------------------|--------------|--------|-----------|--|---|
| 7    | s8      | Movie   | Sankofa                       | Haile Gerima        | Kofi Ghanaba, Oyafulmike Ogunlano, Alexandra D... | United States, Ghana, Burkina Faso, United Kin... | September 24, 2021 | 1993         | TV-MA  | 125 min   | Dramas, Independent Movies, International Movies | On a photo shoot in Ghana, an American model s... |
| 8    | s9      | TV Show | The Great British Baking Show | Andy Devonshire     | Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho... | United Kingdom                                    | September 24, 2021 | 2021         | TV-14  | 9 Seasons | British TV Shows, Reality TV                     | 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                                     | September 24, 2021 | 2021         | PG-13  | 104 min   | Comedies, Dramas                                 | A woman adjusting to life after a loss contend... |
| 12   | s13     | Movie   | Je Suis Karl                  | Christian Schwochow | Luna Wedler, Jannis Niewöhner, Milan Peschel, ... | Germany, Czech Republic                           | September 23, 2021 | 2021         | TV-MA  | 127 min   | Dramas, International Movies                     | After most of her family is murdered in a terr... |
| 24   | s25     | Movie   | Jeans                         | S. Shankar          | Prashanth, Aishwarya Rai Bachchan, Sri Lakshmi... | India   | September 21, 2021 | 1998         | TV-14  | 166 min   | Comedies, International Movies, Romantic Movies  | When the father of the man she loves insists t... |
| ...  | ...     | ...     | ...                           | ...                 | ...   | ...   | ...                | ...          | ...    | ...       | ...  | ...   |
| 8801 | s8802   | Movie   | Zinzana                       | Majid Al Ansari     | Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ... | United Arab Emirates, Jordan                      | March 9, 2016      | 2015         | TV-MA  | 96 min    | Dramas, International Movies, Thrillers          | Recovering alcoholic Talal wakes up inside a s... |
| 8802 | s8803   | Movie   | Zodiac                        | David Fincher       | Mark Ruffalo, Jake Gyllenhaal, ...                | United States                                     | November 20, 2019  | 2007         | R      | 158 min   | Cult Movies, Dramas, Thrillers                   | A political cartoonist, a crime reporter and      |

## 15. Rename the Column "listed\_in" to "genre"

```
netflix_df.rename(columns={'listed_in':'genre'},inplace=True)
netflix_df
```



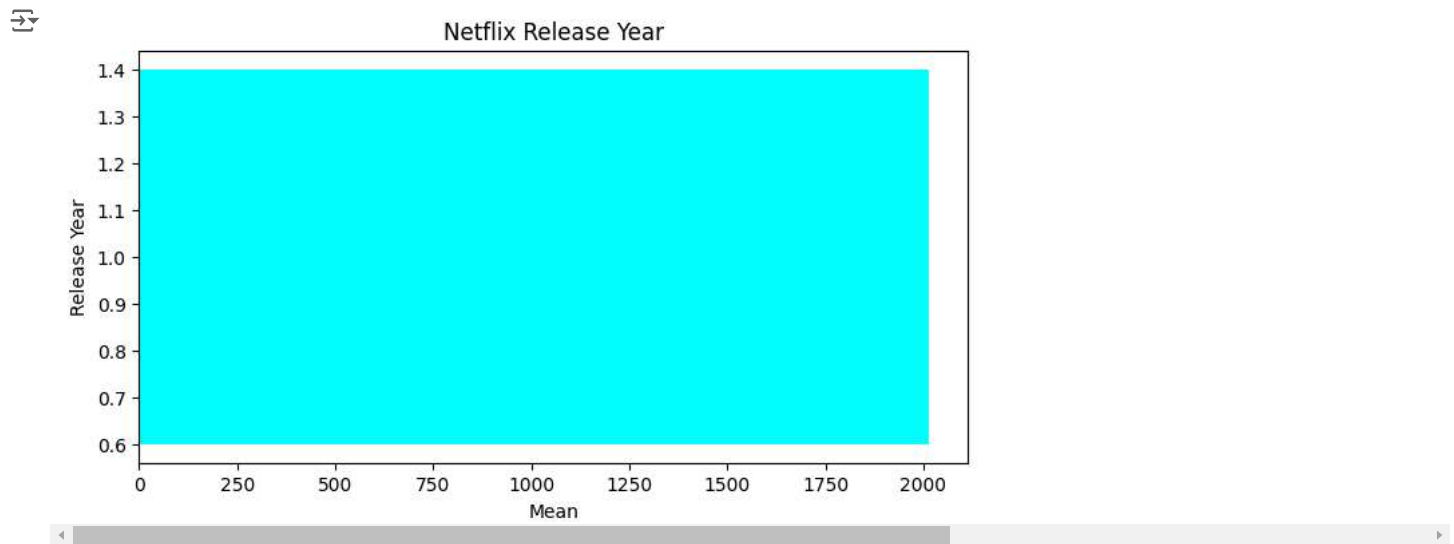
|      | show_id | type    | title                 | director        | cast  | country       | date_added         | release_year | rating | duration  | genre   | description                                       |
|------|---------|---------|-----------------------|-----------------|---|---------------|--------------------|--------------|--------|-----------|---|---|
| 0    | s1      | Movie   | Dick Johnson Is Dead  | Kirsten Johnson |   | United States | September 25, 2021 | 2020         | PG-13  | 90 min    | Documentaries                                     | As her father nears the end of his life, filmm... |
| 1    | s2      | TV Show | Blood & Water         |                 | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | South Africa  | September 24, 2021 | 2021         | TV-MA  | 2 Seasons | International TV Shows, TV Dramas, TV Mysteries   | After crossing paths at a party, a Cape Town t... |
| 2    | s3      | TV Show | Ganglands             | Julien Leclercq | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... |               | September 24, 2021 | 2021         | TV-MA  | 1 Season  | Crime TV Shows, International TV Shows, TV Act... | To protect his family from a powerful drug lor... |
| 3    | s4      | TV Show | Jailbirds New Orleans |                 |   |               | September 24, 2021 | 2021         | TV-MA  | 1 Season  | Docuseries, Reality TV                            | Feuds, flirtations and toilet talk go down amo... |
| 4    | s5      | TV Show | Kota Factory          |                 | Mayur More, Jitendra Kumar, Ranjan Raj, Alam K... | India         | September 24, 2021 | 2021         | TV-MA  | 2 Seasons | International TV Shows, Romantic TV Shows, TV ... | In a city of coaching centers known to train l... |
| ...  | ...     | ...     | ...                   | ...             | ...   | ...           | ...                | ...          | ...    | ...       | ...   | ...   |
| 8802 | s8803   | Movie   | Zodiac                | David Fincher   | Mark Ruffalo, Jake Gyllenhaal, Robert Downey J... | United States | November 20, 2019  | 2007         | R      | 158 min   | Cult Movies, Dramas, Thrillers                    | A political cartoonist, a crime reporter and a... |
|      |         |         |                       |                 |   |               |                    |              |        |           | Kids' TV,   | While living                                      |

✓ DATA VISUALIZATION

✓ 1. Calucating MEAN of release\_year and show it using barh plot (Put title: Netflix release year)

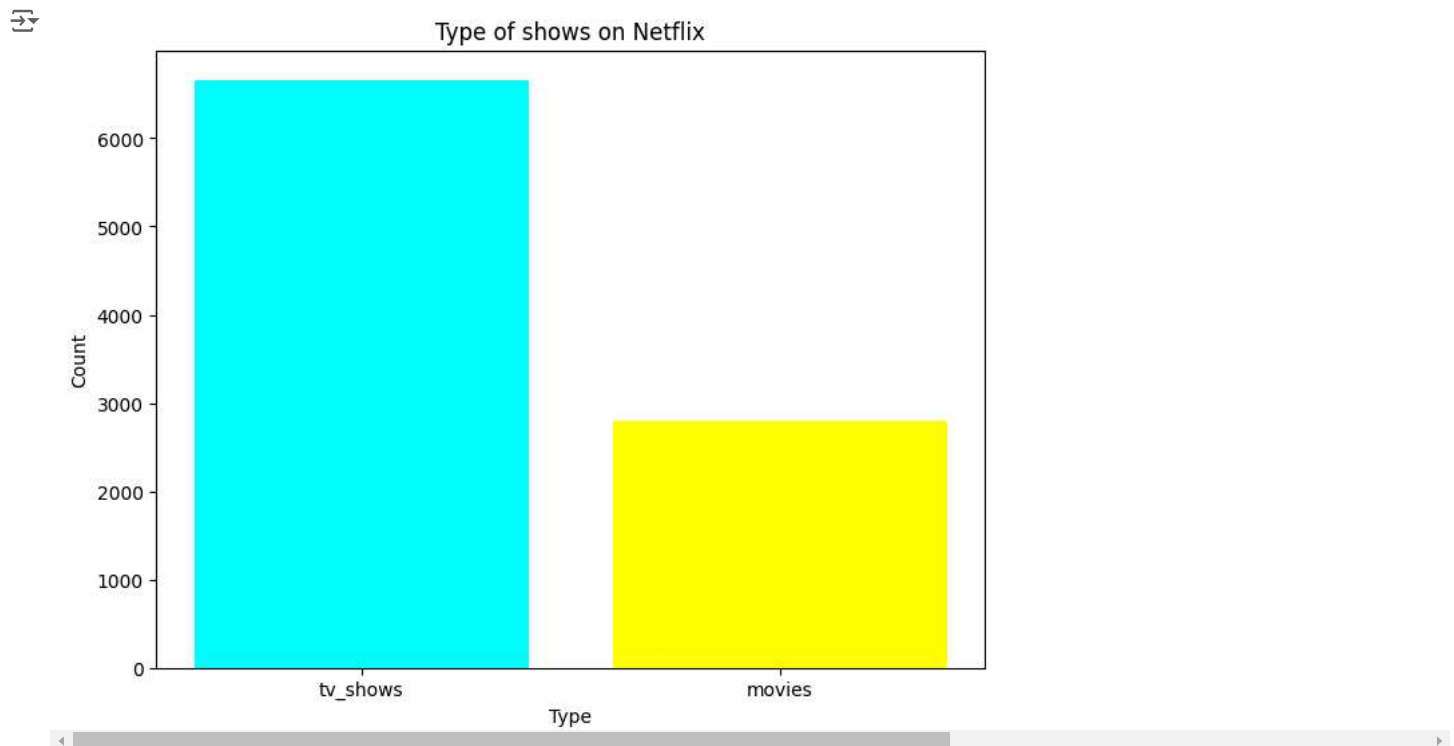
```
mean_release_year = netflix_df['release_year'].mean()
plt.figure(figsize=(8, 4))
plt.barh([1], [mean_release_year], color='cyan')
plt.ylabel('Release Year')
plt.xlabel('Mean')
plt.title('Netflix Release Year')
plt.show()
```





## 2. TV SHOW VS MOVIES On NETFLIX-BAR CHART

```
type=netflix_df['type'].value_counts()
tv_shows=type[0]
movies=type[1]
plt.figure(figsize=(8,6))
plt.bar(['tv_shows','movies'],[tv_shows,movies],color=['cyan','yellow'])
plt.xlabel('Type')
plt.ylabel('Count')
plt.title('Type of shows on Netflix')
plt.show()
```



## 3. Type of shows percentage on Netflix using piechart

```
type=netflix_df['type'].value_counts()
type
```



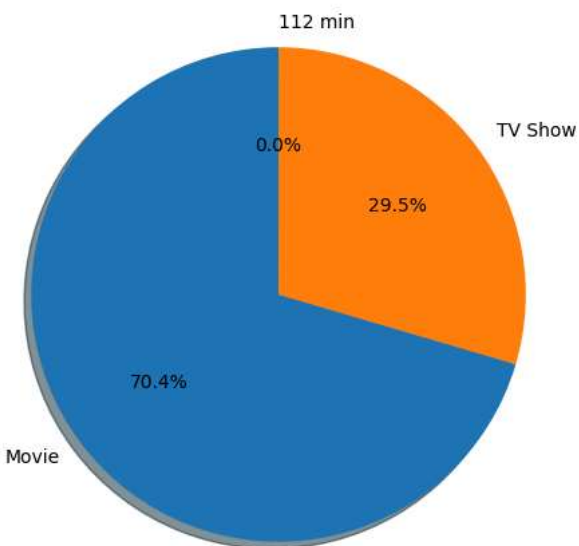
|         | count |
|---------|-------|
| type    |       |
| Movie   | 6665  |
| TV Show | 2795  |
| 112 min | 1     |

**dtype:** int64

```
type=netflix_df['type'].value_counts()
type
fig=plt.figure(figsize=(8,6))
plt.pie(type,labels=type.index,autopct='%1.1f%%',shadow=True,startangle=90)
```



```
(<matplotlib.patches.Wedge at 0x7f82befebf70>,
 <matplotlib.patches.Wedge at 0x7f82befeb50>,
 <matplotlib.patches.Wedge at 0x7f82befdd120>],
 [Text(-0.8807495493146271, -0.6589994168298493, 'Movie'),
 Text(0.880968288893689, 0.6587069712426957, 'TV Show'),
 Text(0.00036513819019486933, 1.0999999393973174, '112 min')],
 [Text(-0.48040884508070564, -0.359454227361736, '70.4%'),
 Text(0.4805281575783758, 0.35929471158692494, '29.5%'),
 Text(0.0001991662855608378, 0.5999999669439913, '0.0%')])
```



#### ✓ 4. Show Netflix Rating Distribution, visualised by a COUNT PLOT.

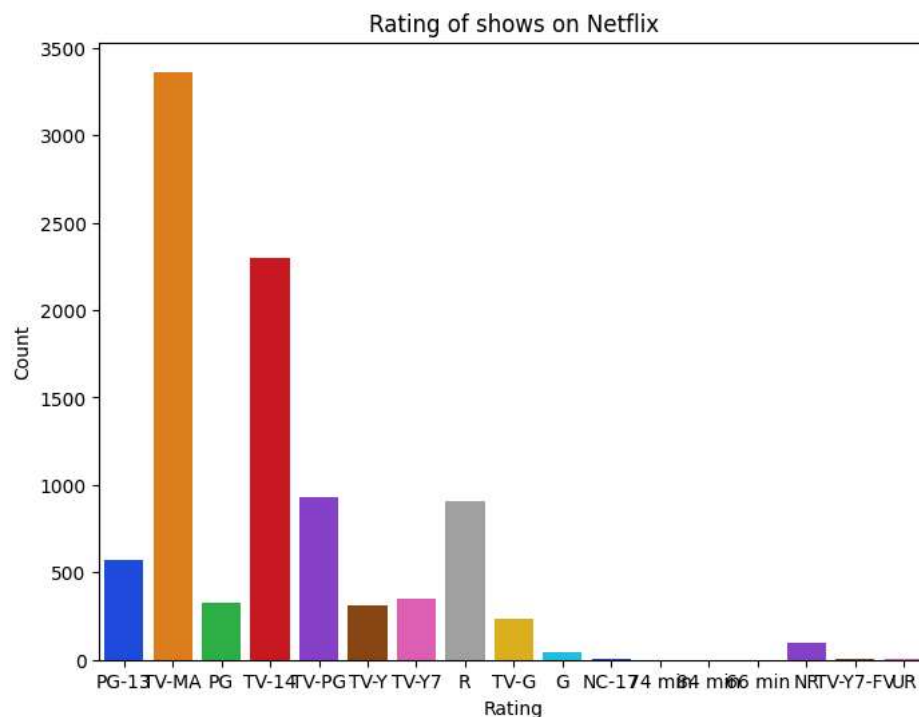
```
rating=netflix_df['rating'].value_counts()
rating
```



|          | count |
|----------|-------|
| rating   |       |
| TV-MA    | 3363  |
| TV-14    | 2295  |
| TV-PG    | 934   |
| R        | 911   |
| PG-13    | 573   |
| TV-Y7    | 353   |
| PG       | 325   |
| TV-Y     | 313   |
| TV-G     | 232   |
| NR       | 94    |
| G        | 47    |
| TV-Y7-FV | 6     |
| UR       | 4     |
| NC-17    | 3     |
| 74 min   | 1     |
| 84 min   | 1     |
| 66 min   | 1     |

dtype: int64

```
plt.figure(figsize=(8,6))
sns.countplot(x='rating',data=netflix_df,palette='bright')
plt.xlabel('Rating')
plt.ylabel('Count')
plt.title('Rating of shows on Netflix')
plt.show()
```



## ✓ 5. Show the Top 5 countries with Highest Movies/TV Shows, visualised by a PIE CHART.

```
country=netflix_df['country'].value_counts()
country
```

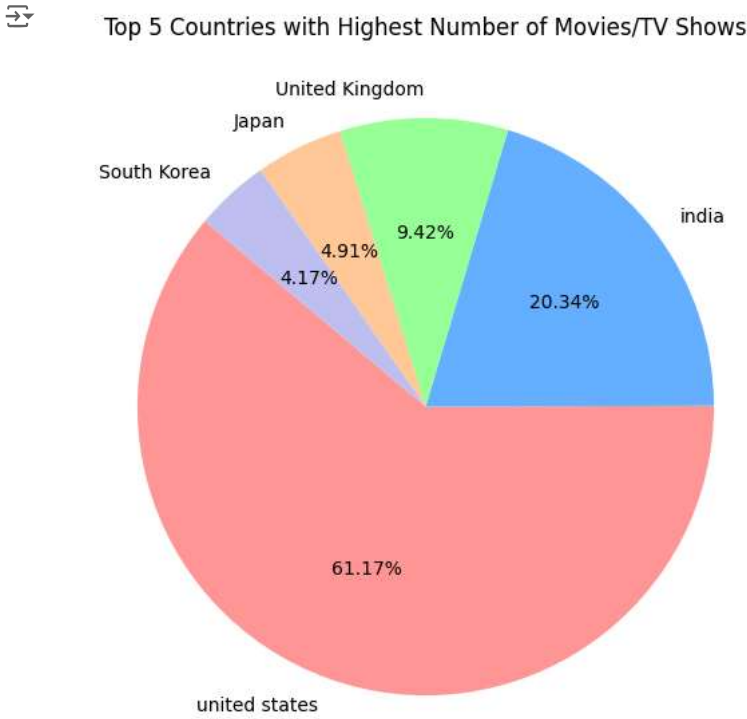
↻

|  | count |
|--|-------|
| country                                |       |
| United States                          | 3065  |
| India                                  | 1019  |
| United Kingdom                         | 472   |
| Japan                                  | 246   |
| South Korea                            | 209   |
| ...                                    | ...   |
| United States, Cambodia, Romania       | 1     |
| Norway, Denmark, Netherlands, Sweden   | 1     |
| United Kingdom, Japan                  | 1     |
| Canada, Germany, France, United States | 1     |
| United Kingdom, Belgium, Sweden        | 1     |

748 rows × 1 columns

dtype: int64

```
countries=['united states','india','United Kingdom','Japan','South Korea']
movies_count=[3065,1019,472,246,209]
plt.figure(figsize=(7, 7))
plt.pie(movies_count, labels=countries, autopct='%1.1f%%', startangle=140, colors=['#ff9999','#66b3ff','#99ff99','#ffcc99', '#c2c2f0'])
plt.title('Top 5 Countries with Highest Number of Movies/TV Shows')
plt.show()
```



```
country=netflix_df['country'].value_counts()
country
countries=['united states','india','United Kingdom','Japan','South Korea']
```



|  | count |
|--|-------|
| country                                |       |
| United States                          | 2818  |
| India                                  | 972   |
| United Kingdom                         | 419   |
| Japan                                  | 245   |
| South Korea                            | 199   |
| ...                                    | ...   |
| Romania, Bulgaria, Hungary             | 1     |
| Uruguay, Guatemala                     | 1     |
| France, Senegal, Belgium               | 1     |
| Mexico, United States, Spain, Colombia | 1     |
| United Arab Emirates, Jordan           | 1     |

748 rows × 1 columns

▼

6. Show the Top 5 Netflix country distribution separated by type of release (ie,.Movies, TV shows), visualizing via Bar Graph.

```
country_type_counts = netflix_df.groupby(['country', 'type'])['show_id'].count().unstack().fillna(0).astype(int)
country_type_counts = country_type_counts.sort_values(['Movie', 'TV Show'], ascending=False)
top_5_countries = country_type_counts.head(5)
top_5_countries.plot(kind='bar', figsize=(12, 6))
plt.title('Top 5 Netflix Country Distribution by Type')
plt.xlabel('Country')
plt.ylabel('Number of Releases')
plt.xticks(rotation=0)
plt.legend(title='Type')
plt.show()
```

