

DATA ANALYSE

Netflix TV Shows and Movies

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Load The File

```
data_df=pd.read_csv("/content/netflix_titles.csv")
```

First 5 Rows

```
data_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...
					Sami Bouajila							To protect his

Double-click (or enter) to edit

Last 5 Rows

```
data_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...
					Jesse Eisenberg, Woody	United States	November 1, 2019				Comedies,	Looking to survive in a

Finding null Values

```
data_df.isnull()
```



	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	False	False	False	False	True	False	False	False	False	False	False	False
1	False	False	False	True	False	False	False	False	False	False	False	False
2	False	False	False	False	False	True	False	False	False	False	False	False
3	False	False	False	True	True	True	False	False	False	False	False	False
4	False	False	False	True	False	False	False	False	False	False	False	False
...
8802	False	False	False	False	False	False	False	False	False	False	False	False
8803	False	False	False	True	True	True	False	False	False	False	False	False
8804	False	False	False	False	False	False	False	False	False	False	False	False
8805	False	False	False	False	False	False	False	False	False	False	False	False
8806	False	False	False	False	False	False	False	False	False	False	False	False

8807 rows × 12 columns

Replace null Values

```
data_df.dropna()
```



	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-	United Arab Emirates,	March 9, 2016	2015	TV-MA	96 min	Dramas, International Movies, ...	Recovering alcoholic Talal wakes up inside a

Info On The Dataset

```
data_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

```

Get some description of the data - ONLY release_year is NUMBER(S) VALUE(S)

```
data_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

```

Columns in the dataset

```
data_df.columns
```

```

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

```

Shape of The data frame

```
data_df.shape
```


```
(8807, 12)
```

How many years of Data (1966-2021) (release_year)

```
data_df['release_year'].count()
```

```
8807
```

```
data_df.release_year.value_counts()
```




	count
release_year	
2018	1147
2017	1032
2019	1030
2020	953
2016	902
...	...
1959	1
1925	1
1961	1
1947	1
1966	1

74 rows × 1 columns

data: int64

✓ MOVIES vs TV SHOWS (type)

```
#movies_df=data_df[data_df['type']=='Movie']
#tv_shows_df=data_df[data_df['type']=='TV show']
#print(movies_df.dtypes)
#print(tv_shows_df.dtypes)
data_df['type'].value_counts()
```



	count
type	
Movie	6131
TV Show	2676

data: int64

✓ Check for null values in all columns of dataset

```
data_df.isnull().sum()
```




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


data: int64

▼ Replace NAN value(s) with TV-MA

```
data_df.fillna('TV-MA')
```




	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	TV-MA	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	TV-MA	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...	TV-MA	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	TV-MA	TV-MA	TV-MA	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	TV-MA	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...




▼ Replace NAN value(s) with United States

```
data_df.fillna('United States')
```




	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	United States	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	United States	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...	United States	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	United States	United States	United States	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	United States	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...




▼ Rename the column "listed_in" to "genre"

```
data_df.rename(columns={"listed_in": "genre"}, inplace=True)
data_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	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train l...

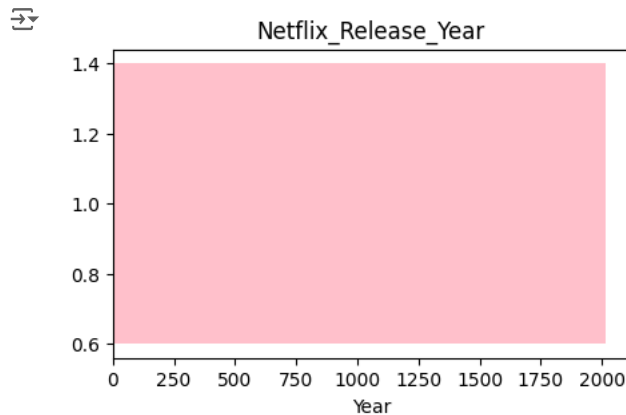


VISUALIZATION

✓ Calculating MEAN of release_year

```
mean_release_year=data_df['release_year'].mean()

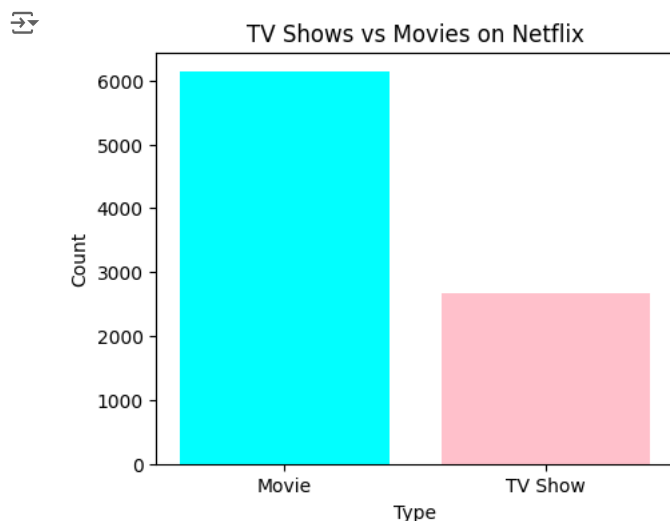
plt.figure(figsize=(5,3))
plt.barh([1],[mean_release_year],color='pink')
plt.xlabel('Year')
plt.title('Netflix_Release_Year')
plt.show()
```



✓ Count the number of TV shows and movies

```
type_counts = data_df['type'].value_counts()

plt.figure(figsize=(5,4))
plt.bar(type_counts.index, type_counts.values, color=['cyan', 'pink'])
plt.xlabel('Type')
plt.ylabel('Count')
plt.title('TV Shows vs Movies on Netflix')
plt.show()
```



✓ Type of shows percentage on Netflix

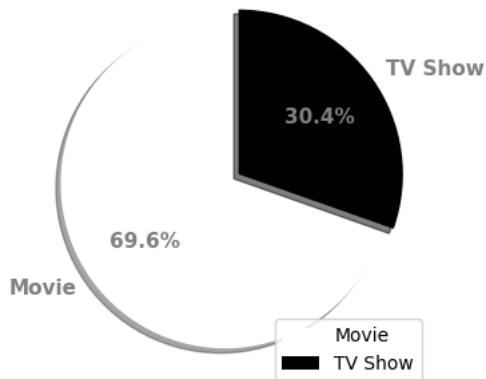
```

labels=["Movie", "TV Show"]
size=data_df["type"].value_counts()
colors = ['white', 'black']
explode=[0,0.1]
plt.figure(figsize=(4, 4))
plt.rcParams["figure.figsize"]=(9,9)
plt.pie(size, labels=labels, colors=colors, autopct='%1.1f%%', explode=explode, shadow=True, textprops=dict(color="gray", weight="bold", font
plt.title("Movies vs. TV Shows on Netflix")
plt.legend()
plt.show()

```



Movies vs. TV Shows on Netflix



Netflix Rating Distribution

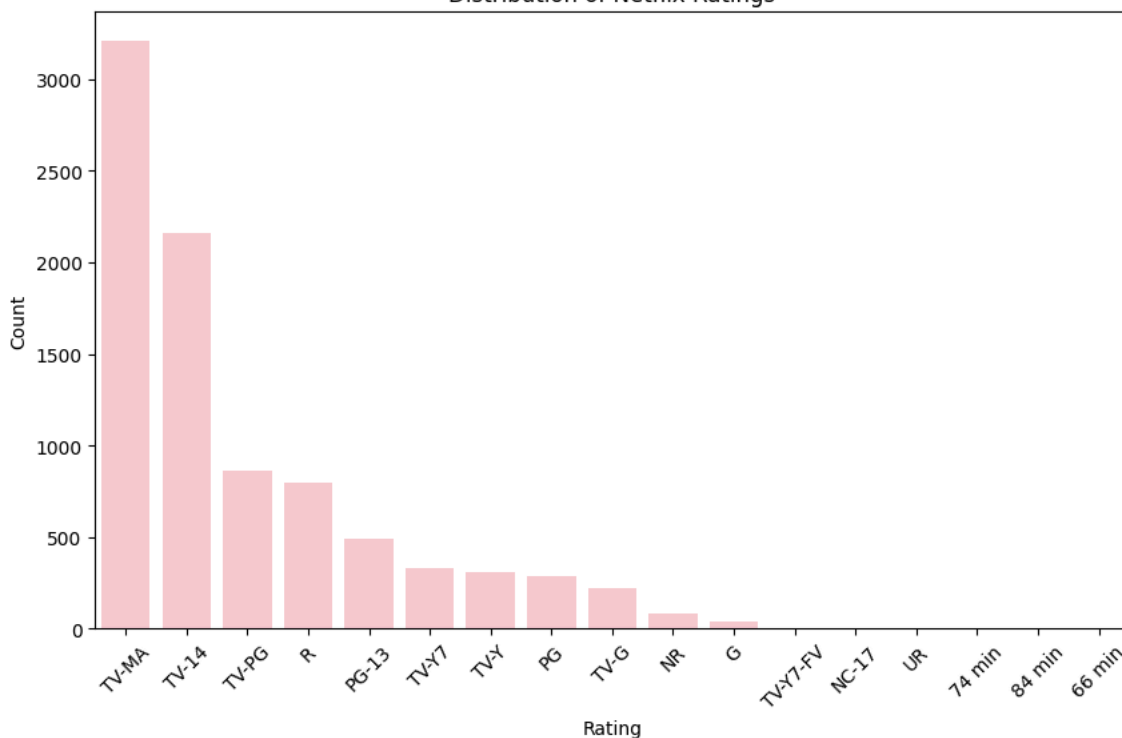
```

plt.figure(figsize=(10, 6))
sns.countplot(data_df, x='rating', order=data_df['rating'].value_counts().index, color='pink')
plt.xlabel('Rating')
plt.ylabel('Count')
plt.title('Distribution of Netflix Ratings')
plt.xticks(rotation=45)
plt.show()

```



Distribution of Netflix Ratings

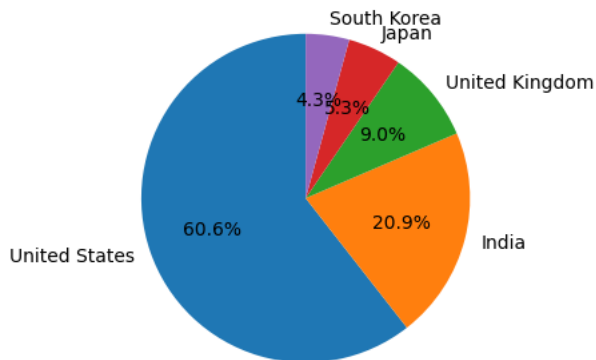


✓ Top 5 countries with the highest Movies/TV Shows

```
top_countries = data_df['country'].value_counts().head(5)
plt.figure(figsize=(4, 4))
plt.pie(top_countries, labels=top_countries.index, autopct='%1.1f%%', startangle=90)
plt.title('Top 5 Countries With Highest Movies/TV Shows')
plt.show()
```



Top 5 Countries with Highest Movies/TV Shows

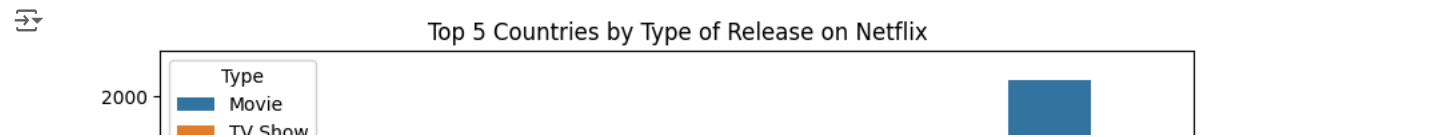


✓ Top 5 Netflix country distribution separated by type

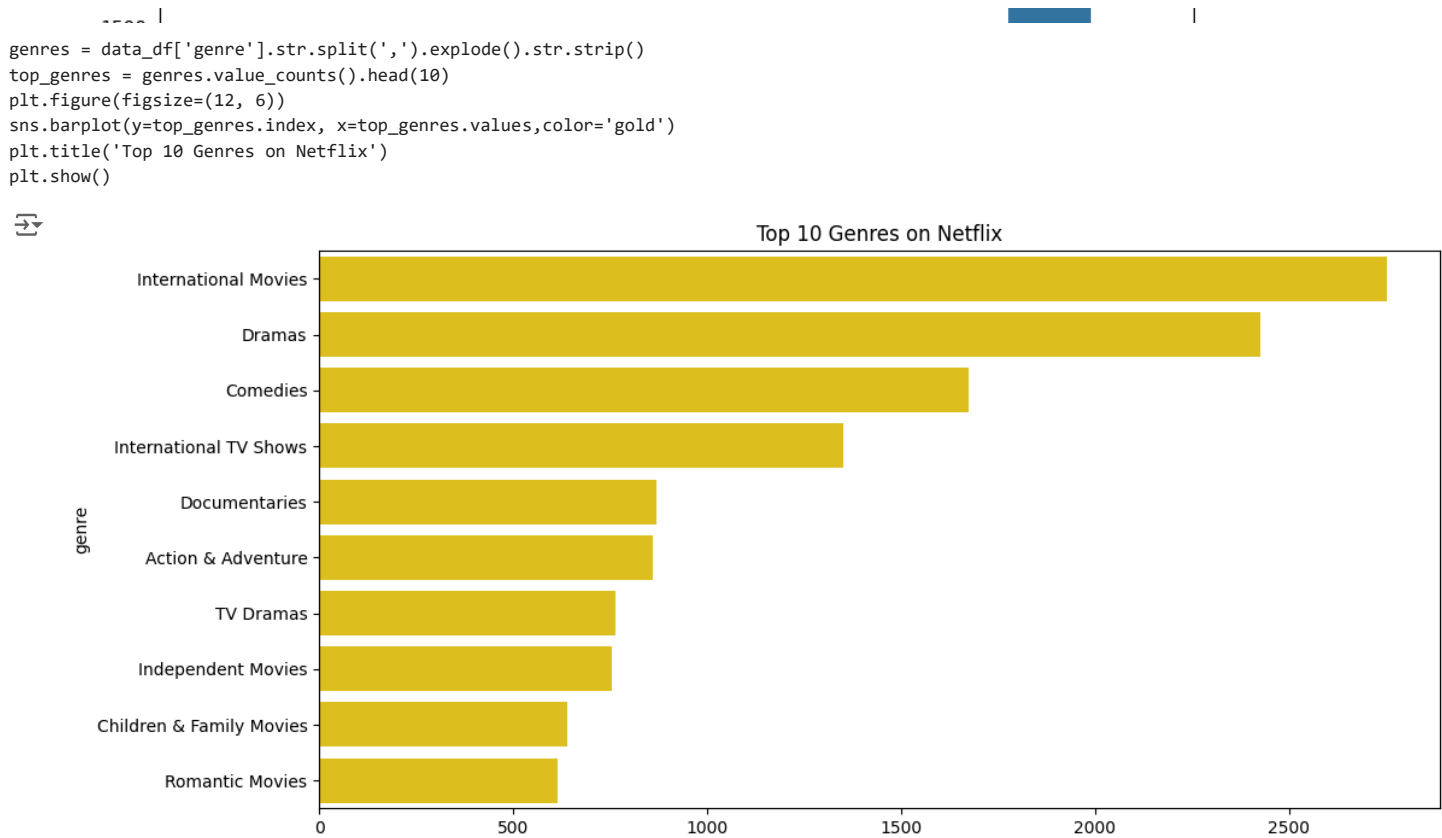
```
country_type_counts = data_df.groupby(['country', 'type']).size().reset_index(name='count')

top_countries = country_type_counts.groupby('country')['count'].sum().nlargest(5).index
filtered_data = country_type_counts[country_type_counts['country'].isin(top_countries)]

plt.figure(figsize=(10, 6))
sns.barplot(data=filtered_data, x='country', y='count', hue='type',)
plt.xlabel('Country')
plt.ylabel('Count')
plt.title('Top 5 Countries by Type of Release on Netflix')
plt.xticks(rotation=45)
plt.legend(title='Type')
plt.show()
```



Top 10 genres of movies and tv shows



Best time to release a movie/TV show

