

## 1.Import required Libraies

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

## 2.Loading Data

```
df = pd.read_csv("netflix.csv")
```

```
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... Sami	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...

## Type

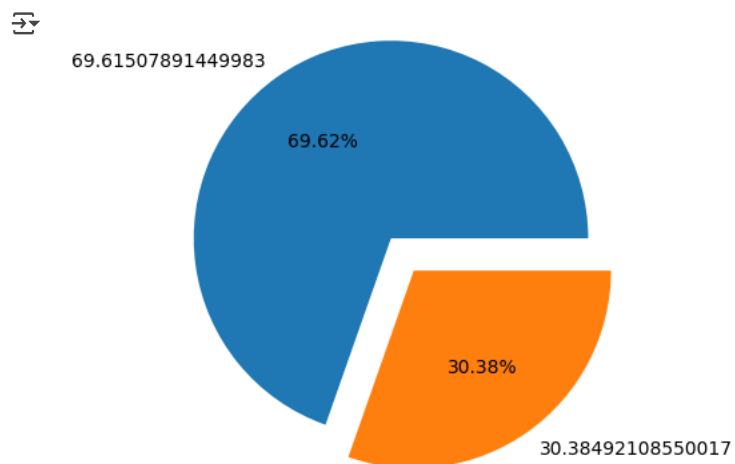
```
df["type"].value_counts()
```

	count
type	
Movie	6131
TV Show	2676

```
dtype: int64
```

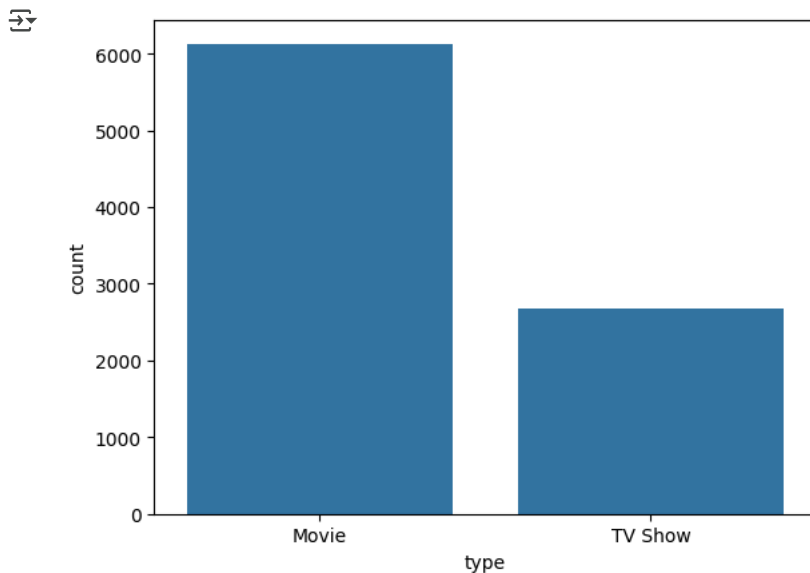
```
movies=df["type"].value_counts()
TV_Shows=df["type"].value_counts()
```

```
plt.pie(movies,labels=movies,autopct="%.2f%",explode=(0.2,0))
plt.show()
```



Double-click (or enter) to edit

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



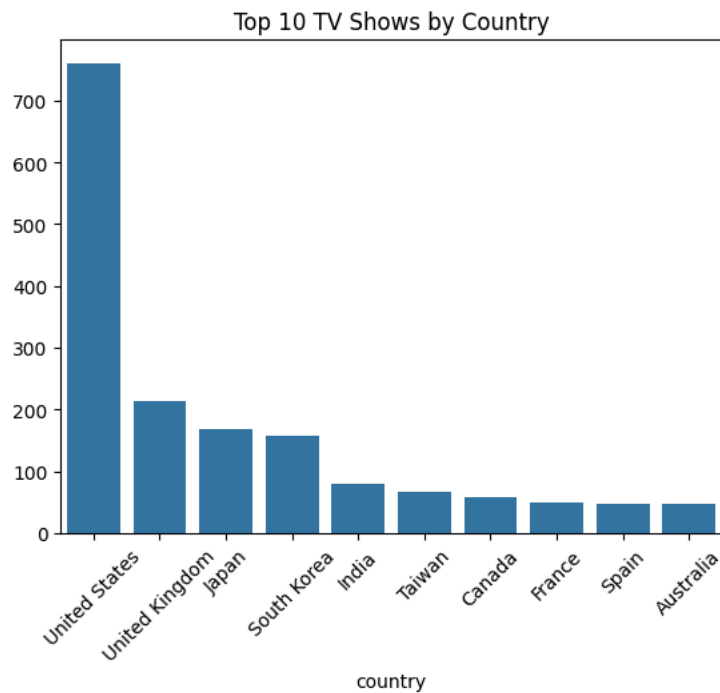
#Movies are 69 % of the total collection and shows are only 30%

#### ▼ Release Year

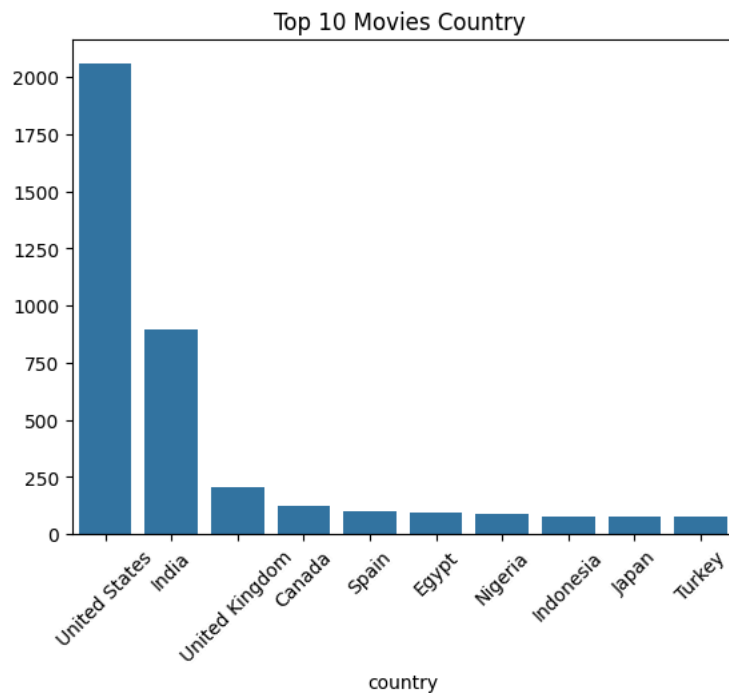
```
top_10_con_movie=df[df["type"] == "Movie"].groupby("country")["title"].nunique().sort_values(ascending=False).head(10)
```

```
top_10_con_TV=df[df["type"] == "TV Show"].groupby("country")["title"].nunique().sort_values(ascending=False).head(10)
```

```
sns.barplot(x=top_10_con_TV.index,y=top_10_con_TV.values)
plt.title("Top 10 TV Shows by Country")
plt.xticks(rotation=45)
plt.figure(figsize=(15,10))
plt.show()
sns.barplot(x=top_10_con_movie.index,y=top_10_con_movie.values)
plt.title("Top 10 Movies Country ")
plt.xticks(rotation=45)
plt.figure(figsize=(15,10))
plt.show()
```



<Figure size 1500x1000 with 0 Axes>




<Figure size 1500x1000 with 0 Axes>

- us United States dominates with over 700 top TV shows—more than triple the next country.
- gb United Kingdom holds a strong second place (~200), showing consistent global appeal.
- jpkr Japan and South Korea are emerging powerhouses in TV content, likely driven by anime and K-dramas.
- in India and tw Taiwan show growing contributions, signaling regional expansion.
- au Australia has the fewest among the top 10, but still notable

USA Dominates with 2000 movies. India is Second in movie production, but global reach is limited. Egypt, Nigeria, Turkey High movie output, low global visibility. Recommendation: Support streaming-first releases and festival circuits. Partner with local creators to produce globally resonant stories.

```
df['date_added'] = pd.to_datetime(df['date_added'], errors='coerce')
```

```
df[df["type"] == "Movie"]["date_added"].dt.weekday.value_counts().sort_values(ascending=False).head(10)
```




	count
date_added	
4	1566
3	1053
2	906
1	852
0	628
6	569
5	557

dtype: int64

- Friday emerges as the most common day for movie releases.

```
df[df["type"] == "TV Show"]["date_added"].dt.weekday.value_counts().sort_values(ascending=False).head(10)
```



	count
date_added	
4.0	910
2.0	370
3.0	334
1.0	330
5.0	246
0.0	217
6.0	171

dtype: int64

- Friday emerges as the most common day for TV show releases.

```
df_clean = df.dropna(subset=["director", "title"])
```

```
top_directors = (
    df_clean.groupby("director")["title"]
    .nunique()
    .sort_values(ascending=False)
    .head(10)
)
```

```
top_directors.groupby("director").value_counts()
```

- Rajiv Chilaka and Suhas Kadav dominate Indian animation, showing how regional content creators can scale across dozens of titles.

```
tv_df = df[df["type"] == "TV Show"].dropna(subset=["cast"])
```

```
tv_df["cast"] = tv_df["cast"].str.split(",")
```

```
tv_df_exploded = tv_df.explode("cast")
```

```
tv_df_exploded["cast"] = tv_df_exploded["cast"].str.strip()
```

```
cast_counts = tv_df_exploded["cast"].value_counts().sort_values(ascending=False)
```

```
cast_counts.head(10)
```



	count
cast	
<b>Takahiro Sakurai</b>	25
<b>Yuki Kaji</b>	19
<b>Ai Kayano</b>	17
<b>Daisuke Ono</b>	17
<b>Junichi Suwabe</b>	17
<b>Yuichi Nakamura</b>	16
<b>Jun Fukuyama</b>	15
<b>Yoshimasa Hosoya</b>	15
<b>David Attenborough</b>	14
<b>Takehito Koyasu</b>	13

dtype: int64

The top names — Takahiro Sakurai, Yuki Kaji, Ai Kayano, Daisuke Ono, etc. — are all prolific Japanese voice actors, strongly suggesting that anime content forms a major chunk of your TV Show data. High Recurrence: Takahiro Sakurai appears in 25 different TV shows, which is a strong indicator of either: His popularity across genres The platform's preference for featuring shows with recurring voice talent

```
movie_df = df[df["type"] == "Movie"].dropna(subset=["cast"])
```

```
movie_df["cast"] = movie_df["cast"].str.split(",")
```

```
movie_df_explode = movie_df.explode("cast")
```

```
movie_df_explode["cast"] = movie_df_explode["cast"].str.strip()
```

```
cast_counts = movie_df_explode["cast"].value_counts().sort_values(ascending=False)
```

```
cast_counts.head(10)
```



	count
cast	
<b>Anupam Kher</b>	42
<b>Shah Rukh Khan</b>	35
<b>Naseeruddin Shah</b>	32
<b>Om Puri</b>	30
<b>Akshay Kumar</b>	30
<b>Julie Teiwani</b>	28
<b>Paresh Rawal</b>	28
<b>Amitabh Bachchan</b>	28
<b>Boman Irani</b>	27
<b>Rupa Bhimani</b>	27

dtype: int64

Frequent actors like Anupam Kher, Shah Rukh Khan, and Akshay Kumar span genres and decades, appealing to diverse viewer segments. Voice talents like Julie Teiwani and Rupa Bhimani suggest a strong presence of dubbed or animated content — ideal for family or regional audiences.

```
movie_df = df[df["type"] == "Movie"].dropna(subset=["listed_in"])
movie_df
```



Show hidden output

```
movie_df["listed_in"].sort_values(ascending=False).head(10)
```

	listed_in
8520	Thrillers
7728	Thrillers
8566	Thrillers
3734	Thrillers
7701	Thrillers
788	Thrillers
358	Thrillers
1358	Thrillers
8537	Thrillers
8535	Thrillers

dtype: object

Thrillers are among the most frequently produced or listed genres in your dataset.

df


	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	2021-09-25	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	2021-09-24	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	2021-09-24	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	2021-09-24	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	2021-09-24	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,	United States	2019-11-20	2007	R	158 min	Cult Movies, Dramas, ...	A political cartoonist, a crime

```
df[df["type"]=="Movie"]["release_year"].value_counts()
```



	count
release_year	
2017	767
2018	767
2016	658
2019	633
2020	517

df["release\_year"]



	1959	release_year	1
0		2020	
1	1966	2021	1
2		2021	
3	1947	2021	1
4		2021	
dtype: int64		...	
8802		2007	
8803		2018	
8804		2009	
8805		2006	
8806		2015	
8807 rows × 1 columns			
dtype: int64			