

NETFLIX-Case Study

```
!gdown https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/000/940/original/netflix.csv
```

Downloading...

From: https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/000/940/original/netflix.csv

To: /content/netflix.csv

100% 3.40M/3.40M [00:00<00:00, 27.5MB/s]

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

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

	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 Mababane, 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 l...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabl...	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	Koh-Ekshara	NaN	Mayur More, Jitendra Kumar, Dhanraj Rai	India	September	2021	TV-MA	2	International TV Shows, Reality TV Shows	In a city of coaching centers, how to train

```
constraint=df['cast'].apply(lambda x: str(x).split(',')).tolist()
df_t1=pd.DataFrame(constraint,index=df['title'])
df_t1=df_t1.stack()
df_t1=pd.DataFrame(df_t1)
df_t1.reset_index(inplace=True)
df_t1=df_t1[['title',0]]
df_t1.columns=['title','cast']
df_t1[df_t1['cast']=='Rupa Bhimani']
```

	title	cast
313	Chhota Bheem	Rupa Bhimani
708	Mighty Raju	Rupa Bhimani
3373	Chhota Bheem - Dinosaur World	Rupa Bhimani
3379	Chhota Bheem - Neeli Pahaadi	Rupa Bhimani
3386	Chhota Bheem & Ganesh	Rupa Bhimani

```

constraint=df['country'].apply(lambda x: str(x).split(',')).tolist()
df_t2=pd.DataFrame(constraint,index=df['title'])
df_t2=df_t2.stack()
df_t2=pd.DataFrame(df_t2)
df_t2.reset_index(inplace=True)
df_t2=df_t2[['title',0]]
df_t2.columns=['title','country']
df_t2[df_t2['title']=='Sankofa']

```

	title	country
7	Sankofa	United States
8	Sankofa	Ghana
9	Sankofa	Burkina Faso

```

constraint=df['director'].apply(lambda x: str(x).split(',')).tolist()
df_t3=pd.DataFrame(constraint,index=df['title'])
df_t3=df_t3.stack()
df_t3=pd.DataFrame(df_t3)
df_t3.reset_index(inplace=True)
df_t3=df_t3[['title',0]]
df_t3.columns=['title','director']
df_t3[df_t3['director']=='Robert Cullen']

```

	title	director
6	My Little Pony: A New Generation	Robert Cullen

```

constraint=df['listed_in'].apply(lambda x: str(x).split(',')).tolist()
df_t4=pd.DataFrame(constraint,index=df['title'])
df_t4=df_t4.stack()
df_t4=pd.DataFrame(df_t4)
df_t4.reset_index(inplace=True)
df_t4=df_t4[['title',0]]
df_t4.columns=['title','listed_in']
df_t4[df_t4['listed_in']=='Docuseries']

```

	title	listed_in
7	Jailbirds New Orleans	Docuseries

```
table1 = df_t1.merge(df_t2,left_on='title', right_on='title',how='inner')
table1
```

	title	cast	country
0	Dick Johnson Is Dead	nan	United States
1	Blood & Water	Ama Qamata	South Africa
2	Blood & Water	Khosi Ngema	South Africa
3	Blood & Water	Gail Mabalane	South Africa
4	Blood & Water	Thabang Molaba	South Africa
...
81703	Zubaan	Manish Chaudhary	India
81704	Zubaan	Meghna Malik	India



```
table2=table1.merge(df_t3,left_on='title',right_on='title',how='inner')
table2
```

	title	cast	country	director
0	Dick Johnson Is Dead	nan	United States	Kirsten Johnson
1	Blood & Water	Ama Qamata	South Africa	nan
2	Blood & Water	Khosi Ngema	South Africa	nan
3	Blood & Water	Gail Mabalane	South Africa	nan
4	Blood & Water	Thabang Molaba	South Africa	nan
...
89377	Zubaan	Manish Chaudhary	India	Mozez Singh

```
table3=table2.merge(df_t4,left_on='title',right_on='title',how='inner')
table3
```

	title	cast	country	director	listed_in
0	Dick Johnson Is Dead	nan	United States	Kirsten Johnson	Documentaries
1	Blood & Water	Ama Qamata	South Africa	nan	International TV Shows
2	Blood & Water	Ama Qamata	South Africa	nan	TV Dramas
3	Blood & Water	Ama Qamata	South Africa	nan	TV Mysteries
4	Blood & Water	Khosi Ngema	South Africa	nan	International TV Shows
...
201986	Zubaan	Anita Shabdish	India	Mozez Singh	International Movies
201987	Zubaan	Anita Shabdish	India	Mozez Singh	Music & Musicals
201988	Zubaan	Chittaranjan Tripathy	India	Mozez Singh	Dramas
201989	Zubaan	Chittaranjan Tripathy	India	Mozez Singh	International Movies

```
df_half=table3.merge(df1,left_on='title',right_on='title',how='inner')
df_half
```

	title	cast	country	director	listed_in	show_id	type	date_added	release_year	rating	duration
0	Dick Johnson Is Dead	nan	United States	Kirsten Johnson	Documentaries	s1	Movie	September 25, 2021	2020	PG-13	90 min
1	Blood & Water	Ama Qamata	South Africa	nan	International TV Shows	s2	TV Show	September 24, 2021	2021	TV-MA	2 Seasons
2	Blood & Water	Ama Qamata	South Africa	nan	TV Dramas	s2	TV Show	September 24, 2021	2021	TV-MA	2 Seasons
3	Blood & Water	Ama Qamata	South Africa	nan	TV Mysteries	s2	TV	September 24,	2021	TV-MA	2

```
df1=df.drop(['cast','country','director','listed_in','description'],axis=1)
df1
```

	show_id	type	title	date_added	release_year	rating	duration
0	s1	Movie	Dick Johnson Is Dead	September 25, 2021	2020	PG-13	90 min
1	s2	TV Show	Blood & Water	September 24, 2021	2021	TV-MA	2 Seasons
2	s3	TV Show	Ganglands	September 24, 2021	2021	TV-MA	1 Season
3	s4	TV Show	Jailbirds New Orleans	September 24, 2021	2021	TV-MA	1 Season
4	s5	TV Show	Kota Factory	September 24, 2021	2021	TV-MA	2 Seasons
...
8802	s8803	Movie	Zodiac	November 20, 2019	2007	R	158 min

```
df_half.loc[(df_half['cast']=='nan'),'cast']='unknown'
df_half.loc[(df_half['director']=='nan'),'director']='unknown'
df_half.loc[(df_half['country']=='nan'),'country']='unknown'
df_half.loc[(df_half['listed_in']=='nan'),'listed_in']='unknown'
df_half.loc[(df_half['listed_in']=='nan'),'listed_in']='unknown'
df_half.loc[(df_half['date_added']=='nan'),'date_added']='unknown'
```

```
df_half['new_duration'] = df_half['duration'].str.split(' ',expand=True)[0]
```

```
df_half['date_added']=pd.to_datetime(df_half['date_added'])
df_half
```

	title	cast	country	director	listed_in	show_id	type	date_added	release_year	rating	duration	new_duration
0	Dick Johnson Is Dead	unknown	United States	Kirsten Johnson	Documentaries	s1	Movie	2021-09-25	2020	PG-13	90 min	90
1	Blood & Water	Ama Qamata	South Africa	unknown	International TV Shows	s2	TV Show	2021-09-24	2021	TV-MA	2 Seasons	2
2	Blood & Water	Ama Qamata	South Africa	unknown	TV Dramas	s2	TV Show	2021-09-24	2021	TV-MA	2 Seasons	2
3	Blood & Water	Ama Qamata	South Africa	unknown	TV Mysteries	s2	TV Show	2021-09-24	2021	TV-MA	2 Seasons	2

```
df_half['month_of_date'] = df_half['date_added'].dt.month
df_half['month_of_date'].dropna().isnull().sum()
df_half.dropna(inplace=True)
df_half
```

	title	cast	country	director	listed_in	show_id	type	date_added	release_year	rating	duration	new_duration	month_of_date
0	Dick Johnson Is Dead	unknown	United States	Kirsten Johnson	Documentaries	s1	Movie	2021-09-25	2020	PG-13	90 min	90	9
1	Blood & Water	Ama Qamata	South Africa	unknown	International TV Shows	s2	TV Show	2021-09-24	2021	TV-MA	2 Seasons	2	9
2	Blood & Water	Ama Qamata	South Africa	unknown	TV Dramas	s2	TV Show	2021-09-24	2021	TV-MA	2 Seasons	2	9
3	Blood & Water	Ama Qamata	South Africa	unknown	TV Mysteries	s2	TV Show	2021-09-24	2021	TV-MA	2 Seasons	2	9
4	Blood & Water	Khosi Ngema	South Africa	unknown	International TV Shows	s2	TV Show	2021-09-24	2021	TV-MA	2 Seasons	2	9

types of content in diff countries?

INSIGHT: HERE WE CAN SEE GENRE AVAILABLE IN DIFFERENT COUNTRIES

RECOMMENDATIONS: WE CAN PRODUCE MORE NO OF MOVIES/TV SHOWS OF GENRES PERTICULAR TO EACH COUNTRY TO GET MORE REVENUE

```
[248] contry=df_half.groupby(['country'])['listed_in'].unique().reset_index()
contry
```

	country	listed_in
0		[International TV Shows, TV Dramas, Dramas, In...
1	Afghanistan	[Documentaries, International Movies]
2	Albania	[Dramas, International Movies]
3	Algeria	[Dramas, Independent Movies, International Mov...
4	Angola	[Action & Adventure, International Movies]
...
123	Venezuela	[Dramas, International Movies, Thrillers, Docu...
124	Vietnam	[Dramas, International Movies, Romantic Movies...

1.How has the number of movies released per year changed over the last 20-30 years?

INSIGHTS: Here we can see no of movies and TV shows produced each year increasing gradually

RECOMMENDATIONS: We can produce more no of Movies and TV shows.

[249]

```
info= df_half.groupby(['release_year','type'])[['title']].count().reset_index()
info
```

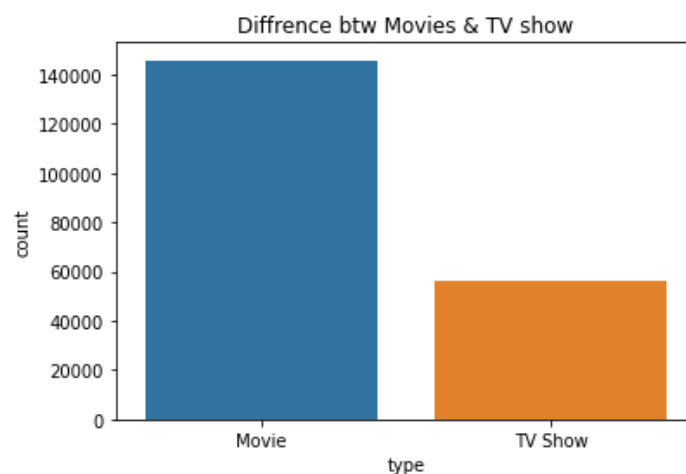
	release_year	type	title
0	1925	TV Show	1
1	1942	Movie	6
2	1943	Movie	5
3	1944	Movie	24
4	1945	Movie	12
...
114	2019	TV Show	10005
115	2020	Movie	9590
...

2.Comparison of tv shows vs. movies

INSIGHT: In the below piechart we can see that TV shows are one third of the Movies

RECOMMENDATION: Produce more no of TV shows

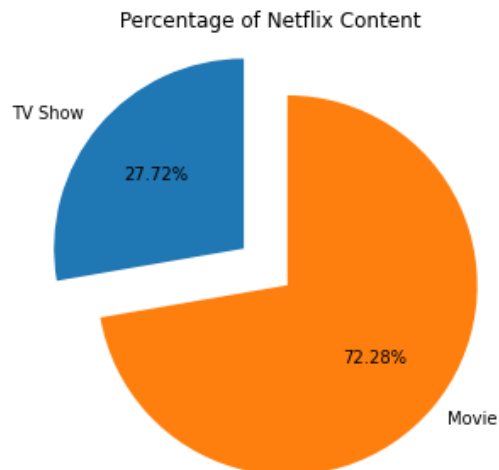
```
[313] sns.countplot(data=df_half,x='type')
plt.title('Diffrence btw Movies & TV show')
plt.show()
```



```

plt.figure(figsize=(9,5))
label=['TV Show', 'Movie']
plt.pie(df_half['type'].value_counts().sort_values(), labels=label, explode=[0.15,0.15],
        autopct='%1.2f%%', startangle=90)
plt.title(' Percentage of Netflix Content')
plt.axis('equal')
plt.show()

```



3.What is the best time to launch a TV show?

INSIGHT: We can see that more no of TV shows released in 12th month(December) since it's holiday season

RECOMMENDATIONS: WE can produce more no of TV shoes as well as movies on that month to get more BOX OFFICE

```

df_half.loc[(df_half['type']=='TV Show'),'month_of_date'].value_counts()

```

```

12    5454
7      5227
8      5162
6      5043
9      4900
4      4543
11     4532
3      4352
1      4307
10     4255
5      4248
2      3909
Name: month_of_date, dtype: int64

```

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

INSIGHT: In the below dataframes we can see the relations btw (a). No of movies and TV shoes produced by directors and their respective genres (b). No of movies and TV shoes acted in by cast and their respective genres

RECOMMENDATIONS: we can find new talented (cast members and directors) to keep the audience more interested.

```
director_directed=df_half.groupby(['director','type'])['title'].count()  
director_directed
```

```
director      type  
Aaron Moorhead  Movie    24  
Aaron Woolf    Movie     2  
Abbas Alibhai Burmawalla  Movie    18  
Abdullah Al Noor  Movie    30  
Abhinav Shiv Tiwari  Movie    24  
..             ..  
Çagan Irmak      Movie    20  
Ísöld Uggadóttir  Movie    27  
Óskar Thór Axelsson  Movie    20  
Ömer Faruk Sorak  Movie    54  
Şenol Sönmez     Movie    48  
Name: title, Length: 5188, dtype: int64
```

```
director_genre=df_half.groupby(['director','listed_in'])['title'].count()  
director_genre
```

```
director      listed_in  
Aaron Moorhead  Independent Movies    6  
                Sci-Fi & Fantasy      6  
                Thrillers             12  
Aaron Woolf     Documentaries         2  
Abbas Alibhai Burmawalla  Action & Adventure    6  
..             ..  
Ömer Faruk Sorak  Romantic Movies    8  
Şenol Sönmez     Comedies            16  
                Dramas               8  
                International Movies  16  
                Romantic Movies      8  
Name: title, Length: 12290, dtype: int64
```

```
cast_acted=df_half.groupby(['cast','type'])['title'].count()  
cast_acted
```

```
cast      type  
Jr.       Movie    1  
          TV Show   1  
"Riley" Lakdhar Dridi  Movie    3  
'Najite Dede  Movie    3  
              TV Show   2  
..       ..  
Şinasi Yurtsever  Movie    9  
Şükran Ovalı     Movie    2  
Şükrü Özyıldız   Movie    2  
              TV Show   3  
Şöpe DİRİSÜ     Movie    2  
Name: title, Length: 40816, dtype: int64
```



```
cast_genre=df_half.groupby(['cast','listed_in'])['title'].count()
cast_genre
```

cast	listed_in	
Jr.	Comedies	1
	TV Dramas	1
"Riley" Lakdhar Dridi	Comedies	1
	Dramas	1
	International Movies	1
	..	
Şükrü Özyıldız	International TV Shows	1
	TV Dramas	1
	TV Thrillers	1
Şopê Dirîsû	Horror Movies	1
	Independent Movies	1

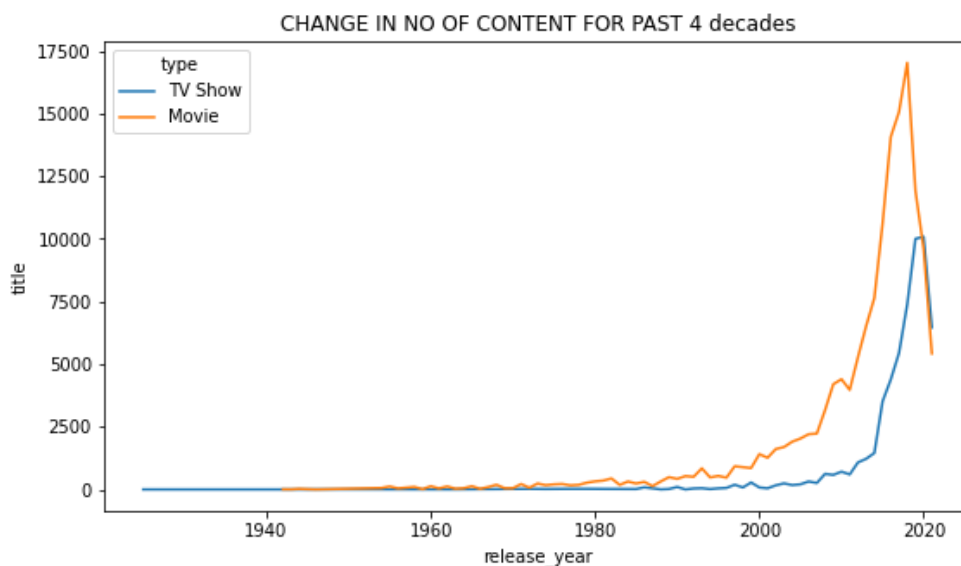
Name: title, Length: 113151, dtype: int64

5.Does Netflix has more focus on TV Shows than movies in recent years?

INSIGHT: According to the graph the no of movies is more than the TV shows in recent years

RECOMMENDATIONS: we can produce more no of TV shows as well.

```
plt.figure(figsize=(9,5))
sns.lineplot(data=info, x="release_year", y="title",hue='type')
plt.title('CHANGE IN NO OF CONTENT FOR PAST 4 decades')
plt.show()
```



6. Understanding what content is available in different countries

INSIGHT: In the below graph we can see some of the countries with most no. of movies and TV shows

RECOMMENDATION: We can target the countries with less no. of movies and TV shows and offer some popular recommendations to watch.

```
info1= df_half.groupby(['country','type'])[['title']].count().reset_index()  
info1.rename(columns={'title':'count'},inplace=True)  
info1
```

