Importing required librabries

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

data = pd.read_csv('netflix.csv')

data.head()

$\overline{\Rightarrow}$		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	11.
	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	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV	To protect his family from a powerful drug lor	
	4													>

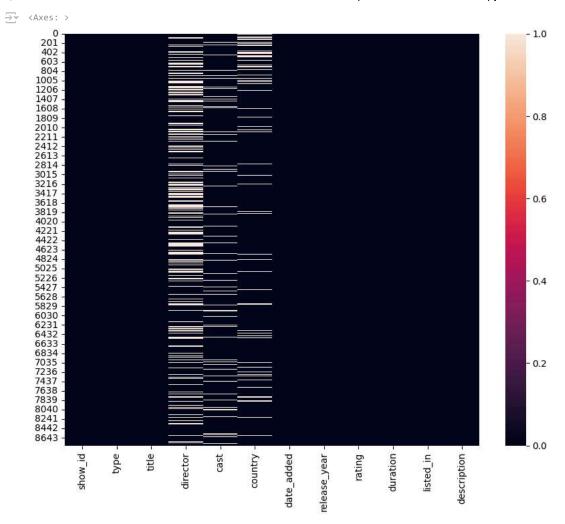
New interactive sheet

View recommended plots

Generate code with data Next steps:

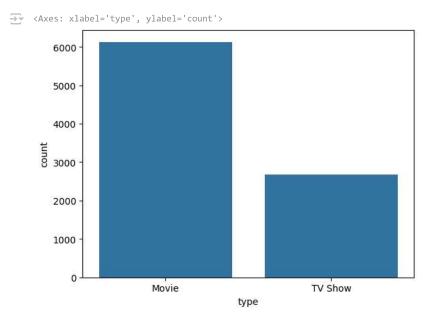
#Checking the null values plt.figure(figsize=(10,8))

sns.heatmap(data.isna())



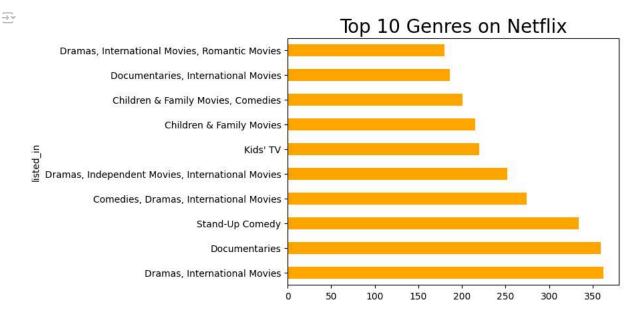
The Director, Cast, and Country fields in many of the data records are incomplete. The primary objective is to identify which television series and films performed the best, with the aim of increasing Netflix's revenue. The analysis begins by examining the distribution of the number of TV series and films available on Netflix. Subsequently, further details about those that received the highest ratings will be explored.

```
sns.countplot(x= 'type', data=data)
```



plt.figure(figsize=(22,10))
data["listed_in"].value_counts()[:10].plot(kind="barh", color="orange")

plt.title("Top 10 Genres on Netflix",size=20);



There are significantly more movies compared to TV shows in the dataset. Therefore, the data will be split to ensure that the analysis remains unbiased. Prior to this, the data needs to be cleaned and preprocessed to obtain better insights.

ANALYSIS FOR MOVIES

```
movies = data.copy()
movies = movies[movies['type'] == 'Movie']
```

movies.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
	6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden,	NaN	September 24, 2021	2021	PG	91 min	Children & Family Movies	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	September 24, 2021	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s
4						Malieea							

Next steps: Generate code with movies View recommended plots New interactive sheet

As the Director and Cast columns contain limited values, they will be dropped from the dataset. However, the Country column will be retained, as it is one of the primary attributes for this case study.

```
movies.drop(['director', 'cast'], axis=1, inplace=True)
movies.head()
```

_												
<u> </u>	sł	now_id	type	title	COI	untry	date_added	release_year	rating	duration	listed_i	n description
	0	s1	Movie	Dick Johnson Is Dead	United S	States	September 25, 2021	2020	PG-13	90 min	Documentarie	As her father nears s the end of his life, filmm
	6	s7	Movie	My Little Pony: A New Generation		NaN	September 24, 2021	2021	PG	91 min	Children & Famil Movie	
	7	s8	Movie	Sankofa	United S Ghana, B Faso, U	urkina	September 24, 2021	1993	TV-MA	125 min	Dramas Independent Movies International Movie	Ghana, an
 ext	9 steps:			<u>The Starling</u> e with movies			September	\neg			Comedies, Drama	A woman adjusting sto life after a loss
ies	[movie	es['cou	ntry'].	isnull()]								
,		show_i	d type	e	title cou	intry (date_added	release_year	rating	duration	listed_i	n description
_	6	S	7 Movie	e My Little F New Gen		NaN	September 24, 2021	2021	PG	91 min	Children & Family Movies	Bur a hridht-eved
	13	s1	4 Movie	e Confession Invisi	ns of an ble Girl	NaN	September 22, 2021	2021	TV-PG	91 min	Children & Family Movies, Comedies	socialiv-awkward
	16	s1 [°]	7 Movie	Europe e Dangerou Otto Skorze	ıs Man:	NaN	September 22, 2021	2020	TV-MA	67 min	Documentaries International Movies	documents reveal
	18	s1	9 Movie	e Ir	trusion	NaN	September 22, 2021	2021	TV-14	94 min	Thrillers	After a deadly home invasion at a couple's new
	22	s2	3 Movie	e Avvai Sha	nmughi	NaN	September 21, 2021	1996	TV-PG	161 min	Comedies International Movies	denied visitation
	8585	s858	6 Movie	e Three-Q	uarters Decent	NaN	June 20, 2019	2010	TV-14	96 min	Comedies, Dramas International Movies	COTTINION IN NIS
	8602	s860	3 Movie	Tom and Jer Mag	ry: The ic Ring	NaN	December 15, 2019	2001	TV-Y7	60 min	Children & Family Movies, Comedies	wizard leaves Tom
	8622	s862	3 Movie		mors 2: shocks	NaN	January 1, 2020	1995	PG-13	100 min	Comedies, Horro Movies, Sci-Fi & Fantas	survivalists and
				es['country']. isnull()]	ffill(axis	=0)						
	show	w_id ty	/pe ti	tle country	date_added	releas	e_year ra	ting duration	listed	_in descr	iption	
	_			movies								
,		show i	d type	e	title	count	rv date a	dded release	vear ra	nting dura	tion listed_in	description
-	5989		O Movie	13TH: A C	onversation h Winfrey & Ava	Unite	ed Januar	v 26	2017		7 min Movies	Oprah Winfrey sits down with director Ava DuVe
	7537	s753	3 Movie	e My Honor V		Ita		ch 1,	2015	NaN 11	5 min Dramas I	Amid the chaos and

Given the small number of missing values, filling them manually using information available online is a practical and efficient approach.

```
# Defining the missing values manually
movies.loc[5989, 'rating'] = 'TV-PG'
movies.loc[7537, 'rating'] = 'TV-MA'

movies[movies['rating'].isnull()]

show_id type title country date_added release_year rating duration listed_in description
```

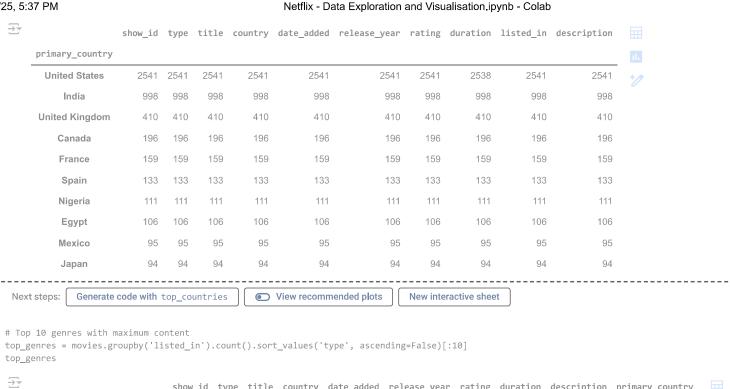
As observed, the Country column contains multiple values representing the countries where the content was displayed. Assuming that the first country listed is the primary source of the content, it will be used for further analysis.

 $movies['primary_country'] = movies['country'].apply(lambda x: x.split(',')[0]) \\ movies$

\Rightarrow		show_id	type	title	country	date_added	release_year	rating	duration	listed_in	description	primary_country	
_	0	s1	Movie	Dick Johnson Is Dead	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm	United States	11.
	6	s7	Movie	My Little Pony: A New Generation	United States	September 24, 2021	2021	PG	91 min	Children & Family Movies	Equestria's divided. But a bright-eyed hero be	United States	
	7	s8	Movie	Sankofa	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	United States	
	9	s10	Movie	The Starling	United States	September 24, 2021	2021	PG-13	104 min	Comedies, Dramas	A woman adjusting to life after a loss contend	United States	
	12	s13	Movie	Je Suis Karl	Germany, Czech Republic	September 23, 2021	2021	TV-MA	127 min	Dramas, International Movies	After most of her family is murdered in a terr	Germany	
	8801	s8802	Movie	Zinzana	United Arab Emirates, Jordan	March 9, 2016	2015	TV-MA	96 min	Dramas, International Movies, Thrillers	Recovering alcoholic Talal wakes up inside a s	United Arab Emirates	
					J Lnited.	_ November				Cult Movies,	A political cartoonist, a		
Next:	steps:	Generate code with movies			O Vi	ew recommend	ded plots N	ew interac	tive sheet				

The analysis for the problem statement will begin by dividing the data based on the top 10 main countries identified. The data will then be grouped by genres and ratings. This approach will provide insights into which types of content received the best reviews in each country, based on the top genres. With this information, the development of similar content can be promoted in the future.

```
# Top 10 countries with maximum content
top_countries = movies.groupby('primary_country').count().sort_values('type', ascending=False)[:10]
top_countries
```



<u> </u>		show_id	type	title	country	date_added	release_year	rating	duration	description	primary_country	
	listed_in											11.
I	Oramas, International Movies	362	362	362	362	362	362	362	362	362	362	7
	Documentaries	359	359	359	359	359	359	359	359	359	359	
	Stand-Up Comedy	334	334	334	334	334	334	334	334	334	334	
	Comedies, Dramas, International Movies	274	274	274	274	274	274	274	274	274	274	
	Oramas, Independent ies, International Movies	252	252	252	252	252	252	252	252	252	252	
Ch	ildren & Family Movies	215	215	215	215	215	215	215	215	215	215	
Ch	Children & Family Movies, Comedies		201	201	201	201	201	201	201	201	201	
	Documentaries, Interpational Wovies	186	186	186	186	186	186	186	186	186	186	
lext step	s: Generate code with t	op_genres		View	/ recomme	nded plots	New interactive	e sheet				

Filtering the data for top 10 genres

Next steps:

movies = movies[movies['listed_in'].isin(list(top_genres.index))]

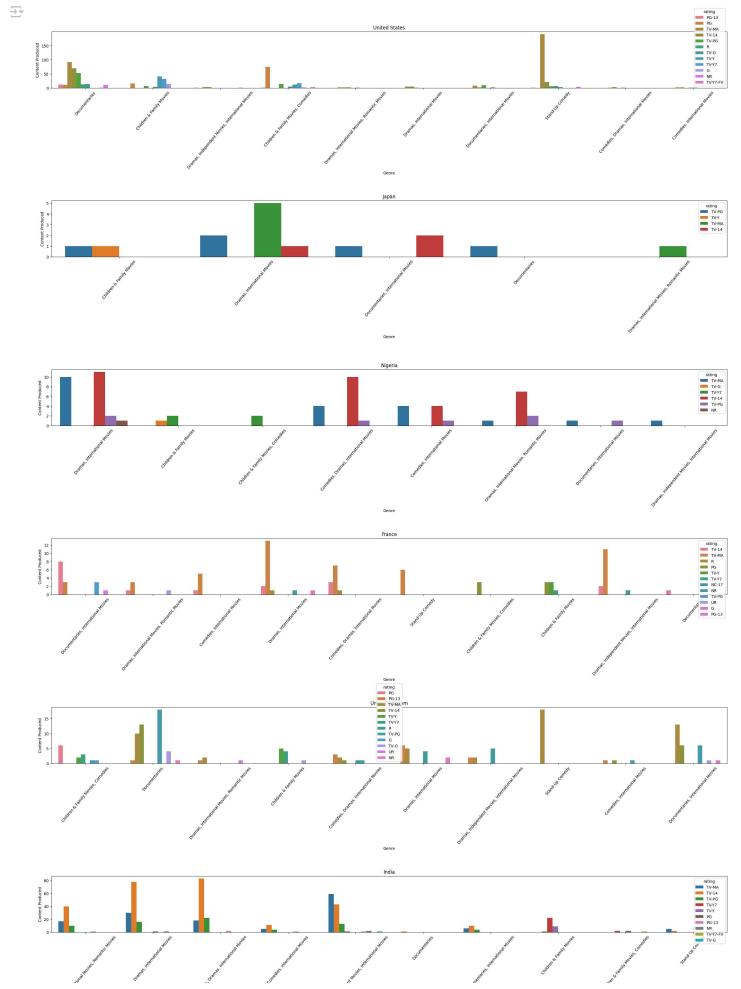
Filtering the data for top 10 countries

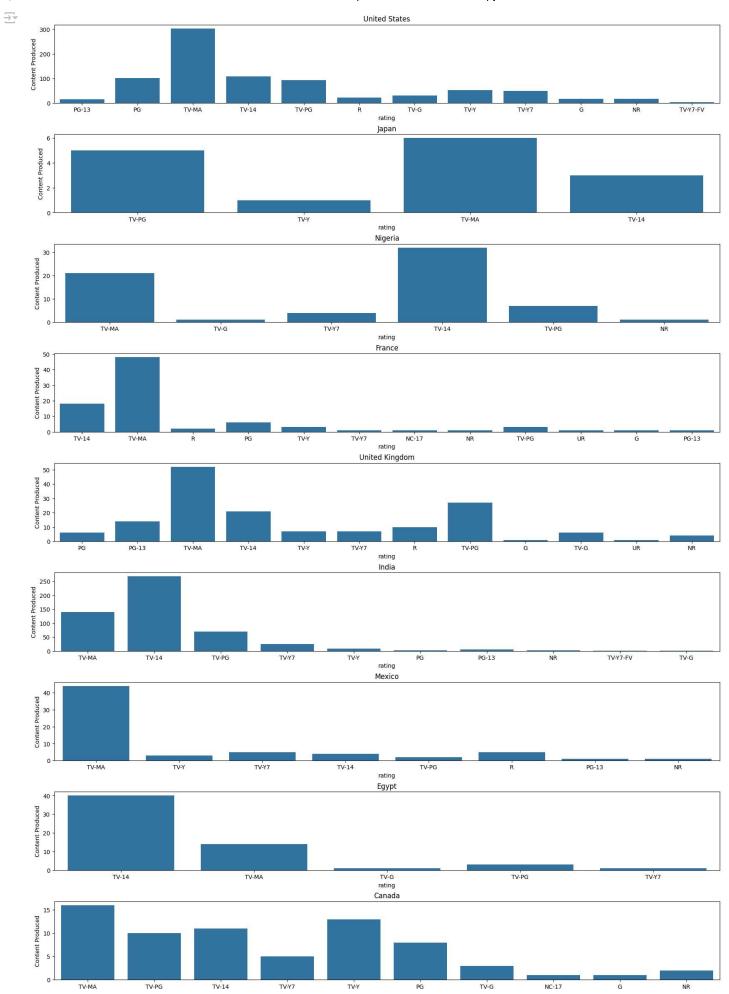
movies = movies[movies['primary_country'].isin(list(top_countries.index))] movies

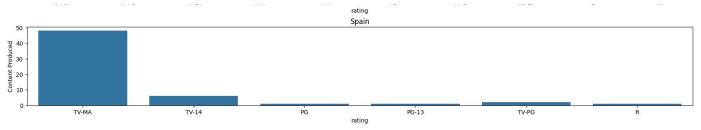
\Rightarrow		show_id	type	title	country	date_added	release_yea	r ratin	g duration	listed_in	description	primary_country	
	0	s1	Movie	Dick Johnson Is Dead	United States	September 25, 2021	202	0 PG-1	3 90 min	Documentaries	As her father nears the end of his life, filmm	United States	11. */
	6	s7	Movie	My Little Pony: A New Generation	United States	September 24, 2021	202	!1 P(91 min	Children & Family Movies	Equestria's divided. But a bright-eyed hero be	United States	
	7	s8	Movie	Sankofa	United States, Ghana, Burkina Faso, United Kin	September 24, 2021	199	93 TV-M	A 125 min	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s	United States	
	30	s31	Movie	Ankahi Kahaniya	United States, India, France	September 17, 2021	202	?1 TV-1	4 111 min	Dramas, Independent Movies, International Movies	As big city life buzzes around them, lonely so	United States	
	45	s46	Movie	My Heroes Were Cowboys	United States	September 16, 2021	202	:1 P(3 23 min	Documentaries	Robin Wiltshire's painful childhood was rescue	United States	
	8793	s8794	Movie	Yours, Mine and Ours	United States	November 20, 2019	200)5 P(3 88 min	Children & Family Movies, Comedies	When a father of eight and a mother of 10 prep	United States	
	8794	s8795	Movie	اشتباك	Egypt,	October 11,	201	6 TV-1	4 98 min	Dramas, Independent Movies,	Amid the tumult following	Egypt	
Next	steps:	Genera	te code	with movies	O V	iew recommer	ided plots	New inte	ractive sheet				
<pre>ratings=[] for rate in movies['rating'].unique(): ratings.append(rate)</pre>													
<pre>countries = movies['primary_country'].unique()</pre>													
listin	g = m	ovies['li	.sted_i	n'].unique()									

The top content based on genres and the highest ratings within each genre across different countries.

```
fig = plt.figure(figsize=(30,40))
for i, name in enumerate(countries):
 frame = movies[movies['primary_country'] == str(name)]
 ax = fig.add_subplot(len(countries),1,i+1)
 topic = name
 sns.countplot(x='listed_in', data= frame[frame['listed_in'].isin(listing)], hue='rating')
 ax.set_title(topic)
 plt.subplots_adjust(left=0.1,
                   bottom=0.1,
                   right=0.9,
                   top=1.5,
                   wspace=0.5,
                   hspace=2.0)
 plt.xlabel('Genre')
 plt.xticks(rotation = 50)
 ax.set(ylabel='Content Produced')
```







The graph above clearly illustrates the types of content and ratings that are most appealing to audiences in the top 10 countries.

Note: The top 10 genres were selected for displaying this data.

All the steps and analysis done for Movies is repeated for TV Shows as well and conclusions are drawn from them on a similar basis

ANALYSIS ON TV SHOWS

```
tv = data.copy()
tv = tv[tv['type'] == 'TV Show']
tv.head()
```

₹	show	_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	
-	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	11
	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	
Next	steps:	Gen	erate co	ode with tv	● Vie	w recommend	led plots	New intera	ctive sheet					

tv.drop(['director', 'cast'], axis=1, inplace=True)
tv.head()