

Netflix Data Analytics

In [2]:

```
## Import the necessary library

import pandas as pd
import seaborn as sns
```

In [19]:

```
# Import the dataset

Netflix = pd.read_csv('C:/Users/User/Downloads/Netflix.csv')
```

In [20]:

```
Netflix.head(2)
```

Out[20]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &... In a future where the elite inhabit an island ...
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies After a devastating earthquake hits Mexico Cit...

In [21]:

```
Netflix.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7789 entries, 0 to 7788
Data columns (total 11 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Show_Id         7789 non-null   object
1   Category        7789 non-null   object
2   Title           7789 non-null   object
3   Director        5401 non-null   object
4   Cast            7071 non-null   object
5   Country         7282 non-null   object
6   Release_Date    7779 non-null   object
7   Rating          7782 non-null   object
8   Duration        7789 non-null   object
9   Type            7789 non-null   object
10  Description      7789 non-null   object
dtypes: object(11)
memory usage: 669.5+ KB
```

In [22]:

```
Netflix.shape
```

Out[22]:

(7789, 11)

Task 1. Is there any Duplicate Record in this dataset? If yes, then

Task.1. Is there any Duplicate record in this dataset ? If yes, then remove the duplicate records.

In [26]:

```
## using duplicate function

Netflix.duplicated().sum()
```

Out[26]:

2

In [29]:

```
Netflix = Netflix.drop_duplicates()
```

In [30]:

```
Netflix.duplicated().sum()
```

Out[30]:

0

Previously we have 7789 rows no we have 7787 rows that means we have sucessfully removed the duplicate values.

Task.2. Is there any Null Value present in any column ? Show with Heat-map.

In [32]:

```
## Checking for null value

Netflix.isnull().sum()
```

Out[32]:

```
Show_Id          0
Category         0
Title            0
Director        2388
Cast            718
Country         507
Release_Date     10
Rating           7
Duration         0
Type            0
Description      0
dtype: int64
```

In [36]:

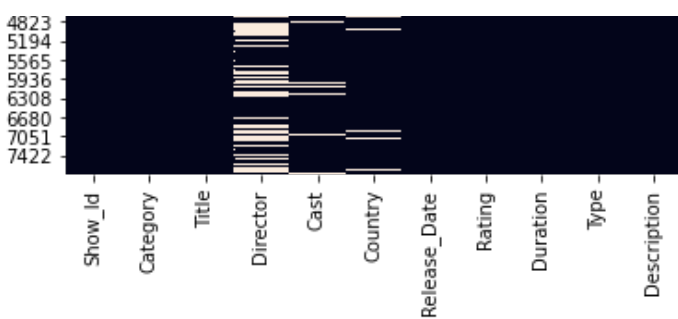
```
## Drawing a heat map

sns.heatmap(Netflix.isnull(), cbar= False)
```

Out[36]:

<matplotlib.axes._subplots.AxesSubplot at 0x289817c8a00>





Q.1. For 'House of Cards', what is the Show Id and Who is the Director of this show ?

In [37]:

```
Netflix.head(2)
```

Out[37]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...

In [39]:

```
Netflix[Netflix['Title'].isin(['House of Cards'])]
```

Out[39]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	
2832	s2833	TV Show	House of Cards	Robin Wright, David Fincher, Gerald McRaney, J...	Kevin Spacey, Robin Wright, Kate Mara, Corey S...	United States	November 2, 2018	TV-MA	6 Seasons	TV Dramas, TV Thrillers	A ruthless politician will stop at nothing to ...

Q.2. In which year highest number of the TV Shows & Movies were released ? Show with Bar Graph.

In [40]:

```
## Let's type the dtypes of all the columns
```

```
Netflix.dtypes
```

Out[40]:

```
Show_Id      object
Category     object
Title        object
Director     object
Cast         object
Country      object
Release_Date object
Rating       object
```

Duration object
Type object
Description object
dtype: object

In [41]:

```
## As we see that Release_Data is in string form so we need to convert the string into da  
tetime.
```

In [43]:

```
## We first create a new column  
  
Netflix['Date_N'] = pd.to_datetime(Netflix['Release_Date'])
```

In [44]:

```
Netflix.dtypes
```

Out[44]:

Show_Id object
Category object
Title object
Director object
Cast object
Country object
Release_Date object
Rating object
Duration object
Type object
Description object
Date_N datetime64[ns]
dtype: object

In [45]:

```
Netflix.head()
```

Out[45]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connelly...	United States	November 16, 2017	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...
					Jim Sturgess, Kevin						A brilliant group

Show_ID	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Dr... Type	Description
4			Robert Luketic	Spa... Kate Bosworth, Aar...	United States	January 1, 2020	Rating	Duration	Dr... Type	of students become card-

In [46]:

```
## Les count the year using value_counts()
```

In [51]:

```
Netflix['Date_N'].dt.year.value_counts()
```

Out[51]:

```
2019.0    2153
2020.0    2009
2018.0    1685
2017.0    1225
2016.0     443
2021.0     117
2015.0      88
2014.0      25
2011.0      13
2013.0      11
2012.0       3
2009.0       2
2008.0       2
2010.0       1
Name: Date_N, dtype: int64
```

In [52]:

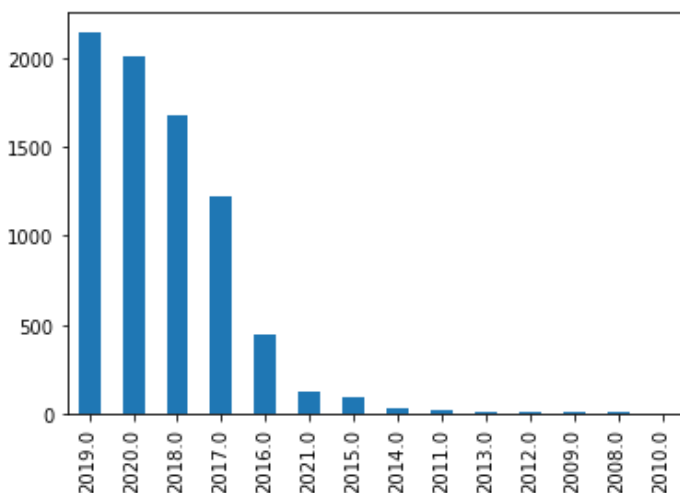
```
## Let's draw a barchart
```

In [53]:

```
Netflix['Date_N'].dt.year.value_counts().plot(kind = 'bar')
```

Out[53]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x289825774c0>
```



Q.3. How many Movies & TV Shows are in the dataset ? Show with Bar Graph.

In [54]:

```
Netflix['Category'].value_counts()
```

Out[54]:

```
Movie    5077
```

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	De
				Jim						

Show_Id	Category	Title	Director	Starring	Country	Release_Date	Rating	Duration	Type	Details	
4	s5	Movie	21	Robert Luketic	Kevin Spacey, Kate Bosworth, Aaron...	United States	January 1, 2020	PG-13	123 min	Dramas	car
6	s7	Movie	122	Yasir Al Yasiri	Amina Khalil, Ahmed Dawood, Tarek Lotfy, Ahmed...	Egypt	June 1, 2020	TV-MA	95 min	Horror Movies, International Movies	ac ad
14	s15	Movie	3022	John Suits	Omar Epps, Kate Walsh, Miranda Cosgrove, Angus...	United States	March 19, 2020	R	91 min	Independent Movies, Sci-Fi & Fantasy, Thrillers	S v : des
27	s28	Movie	#Alive	Cho Il	Yoo Ah-in, Park Shin-hye	South Korea	September 8, 2020	TV-MA	99 min	Horror Movies, International Movies, Thrillers	A ran cit
28	s29	Movie	#AnneFrank - Parallel Stories	Sabina Fedeli, Anna Migotto	Helen Mirren, Gengher Gatti	Italy	July 1, 2020	TV-14	95 min	Documentaries, International Movies	Thr die Fran
...	
7762	s7761	Movie	Zaki Chan	Wael Ihsan	Ahmed Helmy, Yasmin Abdulaziz, Hassan Hosny, H...	Egypt	May 19, 2020	TV-PG	109 min	Comedies, International Movies, Romantic Movies	un yo has c
7783	s7782	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	PG	88 min	Children & Family Movies, Comedies	fron life, sup
7784	s7783	Movie	Zozo	Josef Fares	Imad Creidi, Antoinette Turk, Elias Gergi, Car...	Sweden, Czech Republic, United Kingdom, Denmar...	October 19, 2020	TV-MA	99 min	Dramas, International Movies	Le Zozo
7786	s7785	Movie	Zulu Man in Japan	NaN	Nasty C	NaN	September 25, 2020	TV-MA	44 min	Documentaries, International Movies, Music & M...	docu Sout rapp
7788	s7787	Movie	ZZ TOP: THAT LITTLE OL' BAND FROM TEXAS	Sam Dunn	NaN	United Kingdom, Canada, United States	March 1, 2020	TV-MA	90 min	Documentaries, Music & Musicals	docu de the r

1312 rows × 13 columns



Q.5. Show only the Titles of all TV Shows that were released in India only.

In [70]:

```
Netflix['Title'][(Netflix.Category == 'TV Show') & (Netflix.Country == 'India')]
```

Out[70]:

```
86          21 Sarfarosh: Saragarhi 1897
132          7 (Seven)
340          Agent Raghav
364          Akbar Birbal
533          Anjaan: Rural Myths
...
6249          The Creative Indians
6400  The Golden Years with Javed Akhtar
6469          The House That Made Me
7294          Typewriter
7705          Yeh Meri Family
Name: Title, Length: 71, dtype: object
```

Q.6. Show Top 10 Directors, who gave the highest number of TV Shows & Movies to Netflix ?

In [69]:

```
Netflix['Director'].value_counts().head(10)
```

Out[69]:

```
Raúl Campos, Jan Suter      18
Marcus Raboy                16
Jay Karas                   14
Cathy Garcia-Molina         13
Martin Scorsese              12
Youssef Chahine              12
Jay Chapman                  12
Steven Spielberg             10
David Dhawan                  9
Robert Rodriguez             8
Name: Director, dtype: int64
```

Q.7. Show all the Records, where "Category is Movie and Type is Comedies" or "Country is United Kingdom".

In [75]:

```
Netflix[(Netflix.Category == "Movie") & (Netflix.Type == "Comedies") | (Netflix.Country == 'United Kingdom')]
```

Out[75]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Dr
19	s20	Movie	'89	NaN	Lee Dixon, Ian Wright, Paul Merson	United Kingdom	May 16, 2018	TV-PG	87 min	Sports Movies
33	s34	Movie	#realityhigh	Fernando Lebrija	Nesta Cooper, Kate Walsh, John Michael Higgins...	United States	September 8, 2017	TV-14	99 min	Comedies
58	s59	TV Show	100% Hotter	NaN	Daniel Palmer, Melissa Sophia,	United Kingdom	November 1, 2019	TV-14	1 Season	British TV Shows, International TV

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Shows, Real Type	Description
72	s73	Movie	17 Again	Burr Steers	Zac Efron, Leslie Mann, Matthew Perry, Thomas ...	United States	January 1, 2021	PG-13	102 min	Comedies
82	s83	Movie	2036 Origin Unknown	Hasraf Dulull	Katee Sackhoff, Ray Fearon, Julie Cox, Steven ...	United Kingdom	December 20, 2018	TV-14	95 min	Sci-Fi & Fantasy
...
7670	s7669	TV Show	World War II in Colour	NaN	Robert Powell	United Kingdom	August 1, 2017	TV-MA	1 Season	British TV Shows, Docuseries, International TV...
7671	s7670	TV Show	World's Busiest Cities	NaN	Anita Rani, Ade Adepitan, Dan Snow	United Kingdom	February 1, 2019	TV-PG	1 Season	British TV Shows, Docuseries
7688	s7687	Movie	XV: Beyond the Tryline	Pierre Deschamps	NaN	United Kingdom	March 18, 2020	TV-14	91 min	Documentaries, Sports Movies
7725	s7724	Movie	You Can Tutu	James Brown	Lily O'Regan, Jeannettsy Enriquez Borges, Joel...	United Kingdom	December 31, 2017	TV-G	87 min	Children & Family Movies
7740	s7739	TV Show	Young Wallander	NaN	Adam Pålsson, Richard Dillane, Leanne Best, El...	United Kingdom	September 3, 2020	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Dramas

485 rows x 13 columns

◀												▶
---	--	--	--	--	--	--	--	--	--	--	--	---

Q.8. In how many movies/shows, Tom Cruise was cast ?

In [76]:

```
Netflix.head(2)
```

Out[76]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Date	
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...	2020-08-
1	s2	Movie	07:19	Jorge Michel	Demián Bichir, Héctor ...	Mexico	December 23, 2020	TV-MA	93 min	Dramas, International	After a devastating earthquake	2010-

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Movie Type	Description	Box Office	Release Date
				Bonilla, Oscar		2016	MA					12-
				Serrano, ...								

In [79]:

```
## To find whether or not a particular string exist in the given column we need to use str.contains()

Netflix[Netflix['Cast'].str.contains('Tom Cruise')]
```

```
-----
ValueError                                Traceback (most recent call last)
<ipython-input-79-bb013ea8792a> in <module>
      1 ## To find whether or not a particular string exist in the given column we need to use str.contains()
      2
----> 3 Netflix[Netflix['Cast'].str.contains('Tom Cruise')]

~\anaconda3\lib\site-packages\pandas\core\frame.py in __getitem__(self, key)
    2788
    2789         # Do we have a (boolean) 1d indexer?
-> 2790         if com.is_bool_indexer(key):
    2791             return self._getitem_bool_array(key)
    2792

~\anaconda3\lib\site-packages\pandas\core\common.py in is_bool_indexer(key)
    134         na_msg = "Cannot mask with non-boolean array containing NA / NaN values"
    135         if isna(key).any():
--> 136             raise ValueError(na_msg)
    137         return False
    138         return True
```

ValueError: Cannot mask with non-boolean array containing NA / NaN values

In [80]:

```
## To fix this error we have to Just a friendly reminder thatove the null values in the cast column

#Checking for null values

Netflix['Cast'].isnull().sum()
```

Out[80]:

718

In [81]:

```
## Let's create a new dataframe by dropping all the null values in the older dataframe

New_netflix = Netflix.dropna()
```

In [83]:

```
New_netflix.isnull().sum()
```

Out[83]:

```
Show_Id      0
Category     0
Title        0
Director     0
Cast         0
Country      0
Release_Date 0
Rating       0
Duration     0
Type         0
Description  0
```

```
Description      0
Date_N           0
year            0
dtype: int64
```

Their is no single null values in the data

In [84]:

```
## Lets repeat the previous command

New_netflix[New_netflix['Cast'].str.contains('Tom Cruise')]
```

Out[84]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	
3860	s3861	Movie	Magnolia	Paul Thomas Anderson	Jeremy Blackman, Tom Cruise, Melinda Dillon, A...	United States	January 1, 2020	R	189 min	Dramas, Independent Movies	Through chance, human action, past history and...
5071	s5071	Movie	Rain Man	Barry Levinson	Dustin Hoffman, Tom Cruise, Valeria Golino, Ge...	United States	July 1, 2019	R	134 min	Classic Movies, Dramas	A fast-talking yuppie is forced to slow down w...

Q.9. What are the different Ratings defined by Netflix ?

In [92]:

```
Netflix['Rating'].value_counts()
```

Out[92]:

```
TV-MA      2863
TV-14      1931
TV-PG       806
R           665
PG-13       386
TV-Y        280
TV-Y7       271
PG          247
TV-G        194
NR           84
G           39
TV-Y7-FV     6
UR           5
NC-17        3
Name: Rating, dtype: int64
```

Q.9.1. How many Movies got the 'TV-14' rating, in Canada ?

In [103]:

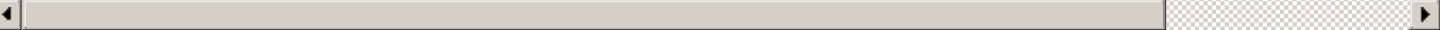
```
Netflix[Netflix.Rating == 'TV-14']
```

Out[103]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Desc
				Divya						

Show_Id	Category	Title	Director	Dutta, Atul Cast Kulkarni,	Country	Release_Date	Rating	Duration	Horror Type Movies,	Desc	
8	s9	Movie	706	Kumar	Mohan Agashe, Anup...	India	April 1, 2019	TV-14	118 min	International Movies	Basu misst psych
15	s16	Movie	Oct-01	Kunle Afolayan	Sadiq Daba, David Bailie, Kayode Olaiya, Kehin...	Nigeria	September 1, 2019	TV-14	149 min	Dramas, International Movies, Thrillers	Again backd Ni lo
16	s17	TV Show	Feb-09	NaN	Shahd El Yaseen, Shaila Sabt, Hala, Hanadi Al- ...	NaN	March 20, 2019	TV-14	1 Season	International TV Shows, TV Dramas	psycl pro Alzhei
18	s19	Movie	15-Aug	Swapnaneel Jayakar	Rahul Pethe, Mrunmayee Deshpande, Adinath Koth...	India	March 29, 2019	TV-14	124 min	Comedies, Dramas, Independent Movies	On l Indepen Day, mish
20	s21	Movie	Kuch Bheege Alfaaz	Onir	Geetanjali Thapa, Zain Khan Durrani, Shray Rai...	India	September 1, 2018	TV-14	110 min	Dramas, Independent Movies, International Movies	accide conn ov Int
...
7765	s7764	Movie	Zenda	Avadhoot Gupte	Santosh Juvekar, Siddharth Chandekar, Sachit P...	India	February 15, 2018	TV-14	120 min	Dramas, International Movies	A cha the lead of a p
7766	s7765	Movie	Zero	Aanand Rai	Shah Rukh Khan, Anushka Sharma, Katrina Kaif, ...	India	May 21, 2019	TV-14	159 min	Comedies, Dramas, International Movies	Throu relation wi wildl
7770	s7769	Movie	Zindagi Kitni Haseen Hay	Anjum Shahzad	Feroze Khan, Sajal Ali, Jibrayl Ahmed Rajput, ...	Pakistan	October 1, 2018	TV-14	126 min	Dramas, International Movies, Romantic Movies	Two p strug kee r
7771	s7770	Movie	Zindagi Na Milegi Dobara	Zoya Akhtar	Hrithik Roshan, Abhay Deol, Farhan Akhtar, Kat...	India	December 15, 2019	TV-14	154 min	Comedies, Dramas, International Movies	Three f adven trip/l
7785	s7784	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrap po wor way into

1931 rows x 13 columns



In [105]:

```
Netflix['Duration'].max()
```

Out[105]:

'99 min'

How can we sort the dataset by Year ?

In [106]:

```
Netflix.sort_values(by='year',ascending = False)
```

Out[106]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	
5564	s5564	Movie	Sherlock Holmes	Guy Ritchie	Robert Downey Jr., Jude Law, Rachel McAdams, M...	United States, Germany, United Kingdom, Australia	January 1, 2021	PG-13	128 min	Action & Adventure, Comedies	The is afc ecc det
5919	s5919	Movie	Surf's Up	Ash Brannon, Chris Buck	Shia LaBeouf, Jeff Bridges, Zooey Deschanel, J...	United States	January 1, 2021	PG	86 min	Children & Family Movies, Comedies, Sports Movies	This C nomi anim cc goes
980	s981	TV Show	Bling Empire	NaN	NaN	NaN	January 15, 2021	TV-MA	1 Season	Reality TV	Follow w Asia
5955	s5955	Movie	Taare Zameen Par	Aamir Khan	Aamir Khan, Darsheel Safary, Tanay Chheda, Tis...	India	January 10, 2021	PG	162 min	Dramas, International Movies	daydre Ish s boa
366	s367	Movie	Al acecho	Francisco D'Eufemia	Rodrigo de la Serna, Belen Blanco, Walter Jako...	NaN	January 12, 2021	TV-MA	81 min	International Movies, Thrillers	Looki a start, i range
...	
3374	s3375	TV Show	Kikoriki	NaN	Igor Dmitriev	NaN	NaN	TV-Y	2 Seasons	Kids' TV	A v rabb his g anima
3492	s3493	TV Show	La Familia P. Luche	NaN	Eugenio Derbez, Consuelo Duval, Luis Manuel Áv...	United States	NaN	TV-14	3 Seasons	International TV Shows, Spanish-Language TV Sh...	irrev s fe Lud Fi
3946	s3947	TV Show	Maron	NaN	Marc Maron, Judd Hirsch, Josh Brener, Nora Zeh...	United States	NaN	TV-MA	4 Seasons	TV Comedies	Maron as Maror intervi
5138	s5138	TV Show	Red vs. Blue	NaN	Burnie Burns, Jason Saldaña, Gustavo	United States	NaN	NR	13 Seasons	TV Action & Adventure, TV Comedies, TV Sci-Fi ...	This p o p sh g

Show_Id	Category	Title	Director	Sorola, G... Cast	Country	Release_Date	Rating	Duration	Type	Descr	
6066	s6066	TV Show	The Adventures of Figaro Pho	NaN	Jurevicius, Craig Behenna, Charlotte Haml...	Australia	NaN	TV-Y7	2 Seasons	Kids' TV, TV Comedies	Im your fears m th

7787 rows × 13 columns



In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []: