Netflix Data Analysis

```
In [1]:   import pandas as pd
```

```
In [73]:  netdata= pd.read_csv(r'C:\Users\lenovo\Downloads\Netflix Dataset.csv')
          netdata.head()
```

Out[73]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	Int T TV
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Int
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Int
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connelly...	United States	November 16, 2017	PG-13	80 min	Inc M
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 min	

```
In [4]:   netdata.tail()
```

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
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 mir
7785	s7784	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	TV-14	111 mir
7786	s7785	Movie	Zulu Man in Japan	NaN	Nasty C	NaN	September 25, 2020	TV-MA	44 mir
7787	s7786	TV Show	Zumbo's Just Desserts	NaN	Adriano Zumbo, Rachel Khoo	Australia	October 31, 2020	TV-PG	1 Season
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 mir

In [5]: `netdata.shape`

Out[5]: (7789, 11)

In [6]: `netdata.size`

Out[6]: 85679

In [7]: `netdata.columns`

Out[7]: Index(['Show_Id', 'Category', 'Title', 'Director', 'Cast', 'Country',
       'Release_Date', 'Rating', 'Duration', 'Type', 'Description'],
      dtype='object')

In [8]: `netdata.dtypes`

```
Out[8]: 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 [9]: `netdata.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 [10]: `#to check duplicate data set value , if yes then remove the duplicate records.`
`netdata.duplicated()`

```
Out[10]: 0        False
         1        False
         2        False
         3        False
         4        False
                  ...
         7784     False
         7785     False
         7786     False
         7787     False
         7788     False
         Length: 7789, dtype: bool
```

In [11]: `#to check duplicates`
`netdata[netdata.duplicated()]`
```

Out[11]:

| | Show_Id | Category | Title | Director | Cast | Country | Release_Date | Rating | Duration |
|---|---|---|---|---|---|---|---|---|---|
| **6300** | s684 | Movie | Backfire | Dave Patten | Black Deniro, Byron "Squally" Vinson, Dominic ... | United States | April 5, 2019 | TV-MA | 97 min |
| **6622** | s6621 | Movie | The Lost Okoroshi | Abba T. Makama | Seun Ajayi, Judith Audu, Tope Tedela, Ifu Enna... | Nigeria | September 4, 2020 | TV-MA | 94 min |

In [12]:
```python
# to remove the duplicates
netdata.drop_duplicates()
```

Out[12]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duratio
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	Seaso
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 m
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 m
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connelly...	United States	November 16, 2017	PG-13	80 m
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 m
...	...	...	...	...	...	...	...	...	
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 m
7785	s7784	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	TV-14	111 m
7786	s7785	Movie	Zulu Man in Japan	NaN	Nasty C	NaN	September 25, 2020	TV-MA	44 m
7787	s7786	TV Show	Zumbo's Just Desserts	NaN	Adriano Zumbo, Rachel Khoo	Australia	October 31, 2020	TV-PG	1 Seas
7788	s7787	Movie	ZZ TOP: THAT LITTLE OL'	Sam Dunn	NaN	United Kingdom, Canada,	March 1, 2020	TV-MA	90 m

	Show_Id	Category	Title	Director		Cast	Country	Release_Date	Rating	Duratic
			BAND FROM TEXAS				United States			

7787 rows × 11 columns

```
In [13]:  # to remove duplicates permanently
          netdata.drop_duplicates(inplace= True)
          netdata[netdata.duplicated()]
```

Out[13]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Descr

◀ ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ ▶

```
In [14]:  netdata.shape
```

Out[14]: (7787, 11)

```
In [15]:  # to show null values
          netdata.isnull()
```

Out[15]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type
0	False	False	False	True	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...
7784	False	False	False	False	False	False	False	False	False	False
7785	False	False	False	False	False	False	False	False	False	False
7786	False	False	False	True	False	True	False	False	False	False
7787	False	False	False	True	False	False	False	False	False	False
7788	False	False	False	False	True	False	False	False	False	False

7787 rows × 11 columns

◀ ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ ▶

```
In [16]:  # to count null vallues
          netdata.isnull().sum()
```

```
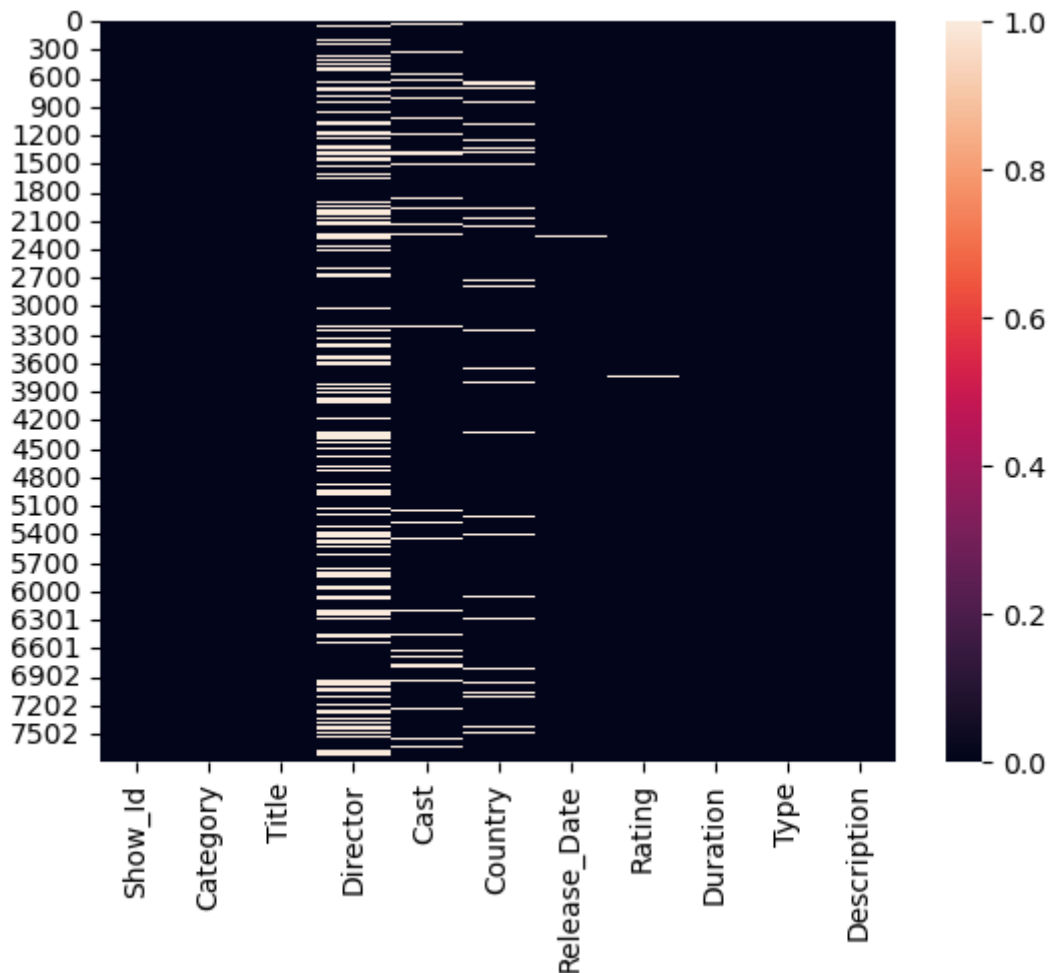Out[16]:  Show_Id              0
          Category             0
          Title                0
          Director          2388
          Cast               718
          Country            507
          Release_Date        10
          Rating               7
          Duration             0
          Type                 0
          Description          0
          dtype: int64
```

Seaborn Library (Heat Map)

```
In [17]:  import seaborn as sns
```

```
In [18]:  sns.heatmap(netdata.isnull())
```

```
Out[18]:  <AxesSubplot:>
```



```
In [19]:  # for 'house of card' , what is the show Id and who is the director of this Show?

          netdata['Title']
```

```
Out[19]: 0                                      3%
         1                                  07:19
         2                                  23:59
         3                                      9
         4                                     21
                          ...
         7784                                Zozo
         7785                              Zubaan
         7786                    Zulu Man in Japan
         7787                 Zumbo's Just Desserts
         7788    ZZ TOP: THAT LITTLE OL' BAND FROM TEXAS
         Name: Title, Length: 7787, dtype: object
```

In [20]: `netdata[netdata['Title'].isin(['House of Cards'])]`

Out[20]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	
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	Dr Th

In [21]: `netdata[netdata['Title'].str.contains('House of Cards')]`

Out[21]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	
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	Dr Th

In [22]:
```python
# in which year highest number of the TV shows & Movies were released ? show with k

netdata['Date_N']= pd.to_datetime(netdata['Release_Date'])
```

In [80]:
```python
netdata.head(3)
```

Out[80]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	Int T TV
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Int
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Int

```
In [24]: netdata['Date_N'].dt.year.value_counts()  #it counts the occurance of all individu
```

```
Out[24]: 2019.0    2153
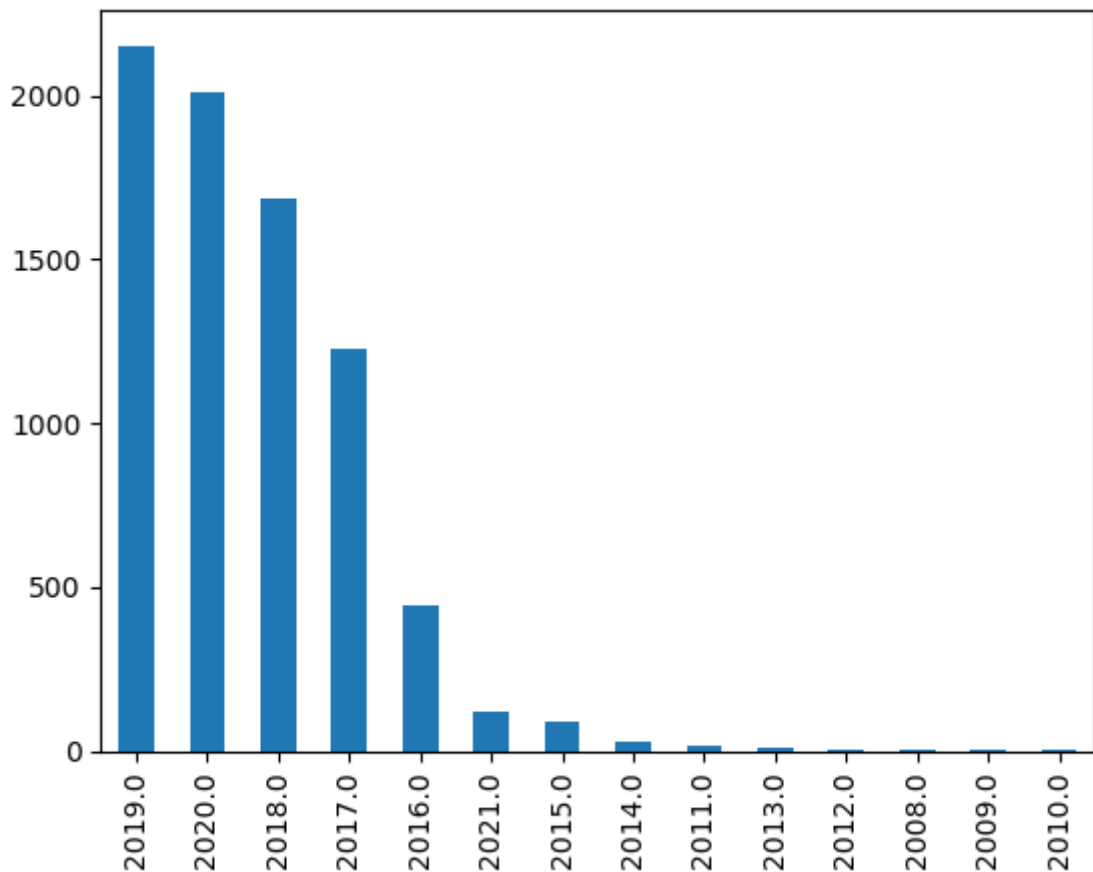         2020.0    2009
         2018.0    1685
         2017.0    1225
         2016.0     443
         2021.0     117
         2015.0      88
         2014.0      25
         2011.0      13
         2013.0      11
         2012.0       3
         2008.0       2
         2009.0       2
         2010.0       1
         Name: Date_N, dtype: int64
```

With Bar Graph

```
In [25]: netdata['Date_N'].dt.year.value_counts().plot(kind='bar')
```

Out[25]: <AxesSubplot:>

2000 — 1500 — 1000 — 500 — 0

2019.0 2020.0 2018.0 2017.0 2016.0 2021.0 2015.0 2014.0 2011.0 2013.0 2012.0 2008.0 2009.0 2010.0

How many Moovies & TV shows are in the dataset? show with Bar Graoph

In [26]:
```python
# By group by
netdata.groupby('Category').Category.count()  #to group all unique items of a colo
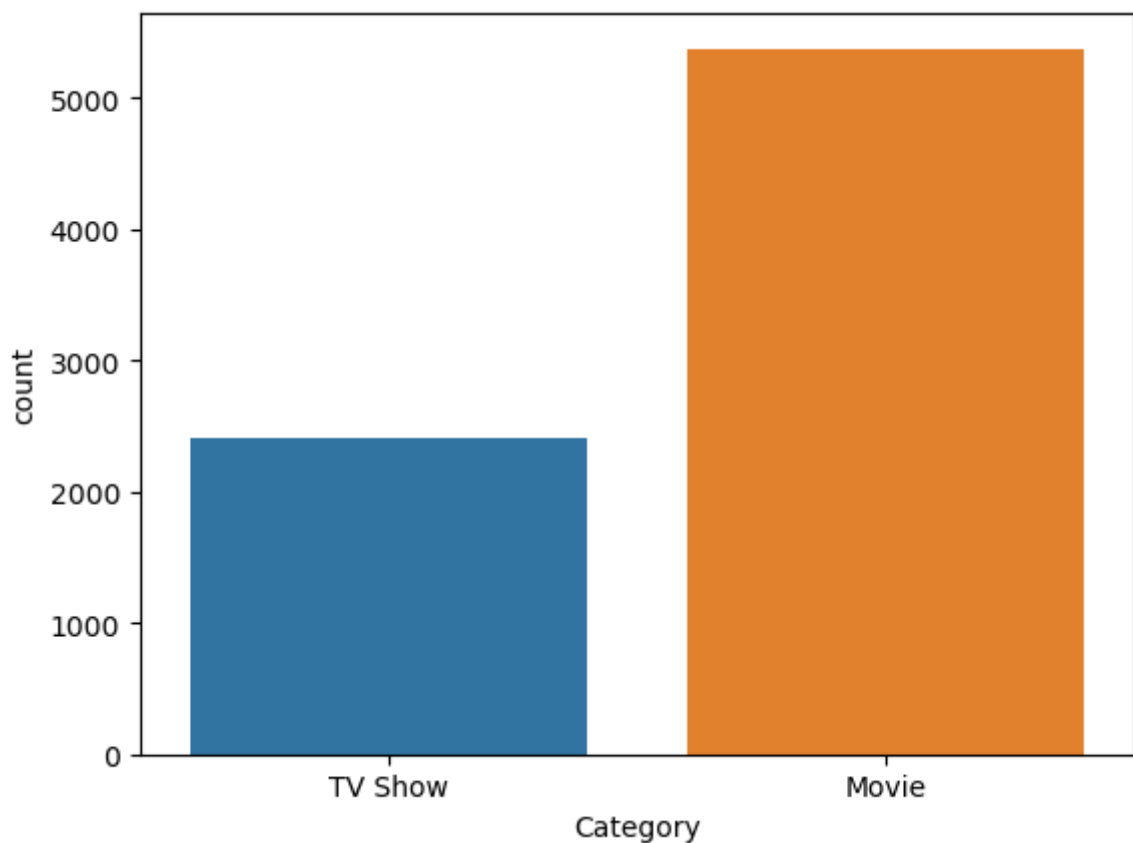```

Out[26]:
```
Category
Movie   5377
TV Show   2410
Name: Category, dtype: int64
```

In [27]:
```python
sns.countplot(netdata['Category'])
```

```
C:\Users\lenovo\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarni
ng: Pass the following variable as a keyword arg: x. From version 0.12, the only v
alid positional argument will be `data`, and passing other arguments without an ex
plicit keyword will result in an error or misinterpretation.
  warnings.warn(
```

Out[27]:
```
<AxesSubplot:xlabel='Category', ylabel='count'>
```

Show all the moovies released in year 2000

creating new coloumn

```
In [28]:  netdata.head(2)
```

Out[28]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	Inte TV TV TV S
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Inte

```
In [29]:  netdata['Year'] = netdata['Date_N'].dt.year   #to create a new Year coloumn ; it wil
```

```
In [30]:  netdata.head(2)
```

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min

In [31]: `netdata[(netdata['Category'] == 'Movie') & (netdata['Year']==2000)]`

Out[31]:

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Descr

show only the titles of all TV Shows that were released in india only

Filtering

In [32]: `netdata.head(2)`

Out[32]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min

In [33]: `netdata [(netdata['Category']== 'TV Show') & (netdata['Country']== 'India')] ['Tit`

```
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
```

show top 10 directors , who have the highest number of TV shows & movie to netflix ?

In [34]:
```python
netdata['Director'].value_counts().head(10)
```

Out[34]:
```
Raúl Campos, Jan Suter    18
Marcus Raboy              16
Jay Karas                 14
Cathy Garcia-Molina       13
Jay Chapman               12
Youssef Chahine           12
Martin Scorsese           12
Steven Spielberg          10
David Dhawan               9
Hakan Algül                8
Name: Director, dtype: int64
```

Show all the records , where category is Movie and Type is "Comedies" or "country is UK"

Filtering (And , Or Operators)

In [79]:
```python
netdata[(netdata['Category']=='Movie') & (netdata['Type']=='Comedies')].head(3)
```

Out[79]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
33	s34	Movie	#realityhigh	Fernando Lebrija	Nesta Cooper, Kate Walsh, John Michael Higgins...	United States	September 8, 2017	TV-14	99 min
72	s73	Movie	17 Again	Burr Steers	Zac Efron, Leslie Mann, Matthew Perry, Thomas ...	United States	January 1, 2021	PG-13	102 min
84	s85	Movie	21 & Over	Jon Lucas, Scott Moore	Miles Teller, Skylar Astin, Justin Chon, Sarah...	United States	April 16, 2019	R	93 min

In [78]: `netdata[(netdata['Category']=='Movie') & (netdata['Type']=='Comedies') | (netdata[`

Out[78]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duratio
19	s20	Movie	'89	NaN	Lee Dixon, Ian Wright, Paul Merson	United Kingdom	May 16, 2018	TV-PG	87 mi
33	s34	Movie	#realityhigh	Fernando Lebrija	Nesta Cooper, Kate Walsh, John Michael Higgins...	United States	September 8, 2017	TV-14	99 mi
58	s59	TV Show	100% Hotter	NaN	Daniel Palmer, Melissa Sophia, Karen Williams,...	United Kingdom	November 1, 2019	TV-14	1 Seaso

In how many Movies/shows Tom Cruise was cast ?

In [37]: `netdata[netdata['Cast']=='Tom Cruise']`

Out[37]: | **Show_Id** | **Category** | **Title** | **Director** | **Cast** | **Country** | **Release_Date** | **Rating** | **Duration** | **Type** | **Descr** |
|---|---|---|---|---|---|---|---|---|---|---|

◄ ━━ ►

By str.contains()

In [38]:
```python
netdata[netdata['Cast'].str.contains('Tom Cruise')]  #there is a nan value so the
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_6212\4257250678.py in <module>
----> 1 netdata[netdata['Cast'].str.contains('Tom Cruise')]  #there is a nan value
so the error showing

~\anaconda3\lib\site-packages\pandas\core\frame.py in __getitem__(self, key)
   3493
   3494            # Do we have a (boolean) 1d indexer?
-> 3495            if com.is_bool_indexer(key):
   3496                return self._getitem_bool_array(key)
   3497

~\anaconda3\lib\site-packages\pandas\core\common.py in is_bool_indexer(key)
    142                    # Don't raise on e.g. ["A", "B", np.nan], see
    143                    #  test_loc_getitem_list_of_labels_categoricalindex_wi
th_na
--> 144                    raise ValueError(na_msg)
    145                return False
    146            return True

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

By creating new data frame /// drop the rows that contains all or any missing values .

In [39]:
```python
new_data = netdata.dropna()
```

In [40]:
```python
new_data.head(2)     # all the  missimg values has been removed
```

Out[40]:
	Show_Id	**Category**	**Title**	**Director**	**Cast**	**Country**	**Release_Date**	**Rating**	**Duration**
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Inter
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Inter

◄ ━━ ►

In [41]:
```python
new_data[new_data['Cast'].str.contains('Tom Cruise')]       # solved
```

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
3860	s3861	Movie	Magnolia	Paul Thomas Anderson	Jeremy Blackman, Tom Cruise, Melinda Dillon, A...	United States	January 1, 2020	R	189 mi
5071	s5071	Movie	Rain Man	Barry Levinson	Dustin Hoffman, Tom Cruise, Valeria Golino, Ge...	United States	July 1, 2019	R	134 mi

What are the different Ratings defined by Netflix ?

By nunique() function

In [42]:
```python
netdata['Rating'].nunique()      # To count the ratings or values
```

Out[42]: 14

In [43]:
```python
netdata.Rating.nunique()
```

Out[43]: 14

By unique()

In [44]:
```python
netdata['Rating'].unique()     # To show the exact values
```

Out[44]:
```
array(['TV-MA', 'R', 'PG-13', 'TV-14', 'TV-PG', 'NR', 'TV-G', 'TV-Y', nan,
       'TV-Y7', 'PG', 'G', 'NC-17', 'TV-Y7-FV', 'UR'], dtype=object)
```

How many Movies got the 'TV-14' ratings in canada ?

In [45]:
```python
netdata[(netdata['Category'] == 'Movie') & (netdata['Rating'] == 'TV-14')].head(5)
```

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Durat
8	s9	Movie	706	Shravan Kumar	Divya Dutta, Atul Kulkarni, Mohan Agashe, Anup...	India	April 1, 2019	TV-14	118
15	s16	Movie	Oct-01	Kunle Afolayan	Sadiq Daba, David Bailie, Kayode Olaiya, Kehin...	Nigeria	September 1, 2019	TV-14	149
18	s19	Movie	15-Aug	Swapnaneel Jayakar	Rahul Pethe, Mrunmayee Deshpande, Adinath Koth...	India	March 29, 2019	TV-14	124
20	s21	Movie	Kuch Bheege Alfaaz	Onir	Geetanjali Thapa, Zain Khan Durrani, Shray Rai...	India	September 1, 2018	TV-14	110
21	s22	Movie	Goli Soda 2	Vijay Milton	Samuthirakani, Bharath Seeni, Vinoth, Esakki B...	India	September 15, 2018	TV-14	128

In [46]: `netdata[(netdata['Category'] == 'Movie') & (netdata['Rating'] == 'TV-14')].shape`

Out[46]: `(1272, 13)`

In [47]: `netdata[(netdata['Category'] == 'Movie') & (netdata['Rating'] == 'TV-14') & (netda`

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating
30	s31	Movie	#cats_the_mewvie	Michael Margolis	NaN	Canada	February 5, 2020	TV-14
34	s35	Movie	#Roxy	Michael Kennedy	Jake Short, Sarah Fisher, Booboo Stewart, Dann...	Canada	April 10, 2019	TV-14
608	s609	Movie	At First Light	Jason Stone	Stefanie Scott, Theodore Pellerin, Percy Hynes...	Canada	December 27, 2018	TV-14
1506	s1507	Movie	Counterfeiting in Suburbia	Jason Bourque	Sarah Butler, Larissa Albuquerque, Kayla Walla...	Canada	July 1, 2018	TV-14
3713	s3714	Movie	Long Time Running	Jennifer Baichwal, Nicholas de Pencier	Gord Downie, Paul Langlois, Gord Sinclair, Rob...	Canada	November 12, 2017	TV-14

```
In [48]: netdata[(netdata['Category']=='Movie')& (netdata['Rating']=='TV-14')&(netdata['Cou
```

Out[48]: `(11, 13)`

How many TV show got the 'R' rating after year 2018 ?

```
In [49]: netdata.head(3)
```

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	Int T TV
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Int
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Int

In [50]:
```python
netdata[(netdata['Category']=='TV Show')&(netdata['Rating']=='R')&(netdata['Year']
```

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
6437	s6436	TV Show	The Hateful Eight: Extended Version	Quentin Tarantino	Samuel L. Jackson, Kurt Russell, Jennifer Jaso...	NaN	April 25, 2019	R	1 Season

What is the maximum duration of a Movie/Show on Netflix ?

In [51]:
```python
netdata.head(2)
```

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	Inte T TV TV S
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Inte

In [52]:
```python
netdata.Duration.unique()    #By extracting all duratiopns
```

```
Out[52]:  array(['4 Seasons', '93 min', '78 min', '80 min', '123 min', '1 Season',
                 '95 min', '119 min', '118 min', '143 min', '103 min', '89 min',
                 '91 min', '149 min', '144 min', '124 min', '87 min', '110 min',
                 '128 min', '117 min', '100 min', '2 Seasons', '84 min', '99 min',
                 '90 min', '102 min', '104 min', '105 min', '56 min', '125 min',
                 '81 min', '97 min', '106 min', '107 min', '109 min', '44 min',
                 '75 min', '101 min', '3 Seasons', '37 min', '113 min', '114 min',
                 '130 min', '94 min', '140 min', '135 min', '82 min', '70 min',
                 '121 min', '92 min', '164 min', '53 min', '83 min', '116 min',
                 '86 min', '120 min', '96 min', '126 min', '129 min', '77 min',
                 '137 min', '148 min', '28 min', '122 min', '176 min', '85 min',
                 '22 min', '68 min', '111 min', '29 min', '142 min', '168 min',
                 '21 min', '59 min', '20 min', '98 min', '108 min', '76 min',
                 '26 min', '156 min', '30 min', '57 min', '150 min', '133 min',
                 '115 min', '154 min', '127 min', '146 min', '136 min', '88 min',
                 '131 min', '24 min', '112 min', '74 min', '63 min', '38 min',
                 '25 min', '174 min', '60 min', '153 min', '158 min', '151 min',
                 '162 min', '54 min', '51 min', '69 min', '64 min', '147 min',
                 '42 min', '79 min', '5 Seasons', '40 min', '45 min', '172 min',
                 '10 min', '163 min', '9 Seasons', '55 min', '72 min', '61 min',
                 '71 min', '160 min', '171 min', '48 min', '139 min', '157 min',
                 '15 min', '65 min', '134 min', '161 min', '62 min', '8 Seasons',
                 '186 min', '49 min', '73 min', '58 min', '165 min', '166 min',
                 '138 min', '159 min', '141 min', '132 min', '52 min', '67 min',
                 '34 min', '66 min', '312 min', '180 min', '47 min', '6 Seasons',
                 '155 min', '14 min', '177 min', '11 min', '9 min', '46 min',
                 '145 min', '11 Seasons', '7 Seasons', '13 Seasons', '8 min',
                 '12 min', '12 Seasons', '10 Seasons', '43 min', '50 min', '23 min',
                 '185 min', '200 min', '169 min', '27 min', '170 min', '196 min',
                 '33 min', '181 min', '204 min', '32 min', '35 min', '167 min',
                 '16 Seasons', '179 min', '193 min', '13 min', '214 min', '17 min',
                 '173 min', '192 min', '209 min', '187 min', '41 min', '182 min',
                 '224 min', '233 min', '189 min', '152 min', '19 min', '15 Seasons',
                 '208 min', '237 min', '31 min', '178 min', '230 min', '194 min',
                 '228 min', '195 min', '3 min', '16 min', '5 min', '18 min',
                 '205 min', '190 min', '36 min', '201 min', '253 min', '203 min',
                 '191 min'], dtype=object)
```

In [53]: `netdata.Duration.dtypes` *#this is in object format so we have to convert it into n*

Out[53]: `dtype('O')`

By using str.split()

In [54]: `netdata[['Minutes', 'Unit']] = netdata['Duration'].str.split(' ', expand = True)`

In [55]: `netdata.head(2)`

Out[55]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	Inte TV TV S
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Inte

Using Max()

```
In [56]: netdata['Minutes'].max()
```

```
Out[56]: '99'
```

```
In [57]: netdata['Duration'].min()
```

```
Out[57]: '1 Season'
```

```
In [58]: netdata['Duration'].max()
```

```
Out[58]: '99 min'
```

```
In [59]: netdata['Minutes'].mean()        #infinite
```

```
Out[59]: inf
```

Which individual country has the Highest No. of TV Shows ?

```
In [60]: netdata.head(2)
```

Out[60]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	Inte T' TV TV S
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Inte

```
In [61]:   # creating new data frame which contains only TV shows

           data_tvshow = netdata[netdata['Category']== 'TV Show']
```

```
In [62]:   data_tvshow.head(2)
```

Out[62]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	Inter TV TV S
5	s6	TV Show	46	Serdar Akar	Erdal Beşikçioğlu, Yasemin Allen, Melis Birkan...	Turkey	July 1, 2017	TV-MA	1 Season	Inter TV TV N

```
In [63]:   data_tvshow.Country.value_counts()
```

```
Out[63]:   United States                                              705
           United Kingdom                                             204
           Japan                                                      157
           South Korea                                                147
           India                                                       71
                                                                      ...
           Canada, United States, United Kingdom, France, Luxembourg    1
           United States, Italy                                         1
           Chile, Italy                                                 1
           Canada, United Kingdom                                       1
           United States, France, South Korea, Indonesia               1
           Name: Country, Length: 183, dtype: int64
```

```
In [64]:   data_tvshow.Country.value_counts().head(1)
```

```
Out[64]:   United States    705
           Name: Country, dtype: int64
```

How can we sort the dataset by Year ?

```
In [65]:   netdata.head(2)
```

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min

```
netdata.sort_values(by='Year').head(2)     #By using sort_value function
```

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duratio
7115	s7114	Movie	To and From New York	Sorin Dan Mihalcescu	Barbara King, Shaana Diya, John Krisiukenas, Y...	United States	January 1, 2008	TV-MA	81 mi
1765	s1766	TV Show	Dinner for Five	NaN	NaN	United States	February 4, 2008	TV-MA	1 Seaso

```
netdata.sort_values(by='Year', ascending = False).head()
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_6212\664028296.py in <module>
----> 1 netdata.sort_values(by='Year', ascending = False).head()

~\anaconda3\lib\site-packages\pandas\util\_decorators.py in wrapper(*args, **kwarg
s)
    309                         stacklevel=stacklevel,
    310                     )
--> 311                 return func(*args, **kwargs)
    312
    313         return wrapper

~\anaconda3\lib\site-packages\pandas\core\frame.py in sort_values(self, by, axis,
ascending, inplace, kind, na_position, ignore_index, key)
    6320
    6321             by = by[0]
-> 6322         k = self._get_label_or_level_values(by, axis=axis)
    6323
    6324         # need to rewrap column in Series to apply key function

~\anaconda3\lib\site-packages\pandas\core\generic.py in _get_label_or_level_values
(self, key, axis)
    1838             values = self.axes[axis].get_level_values(key)._values
    1839         else:
-> 1840             raise KeyError(key)
    1841
    1842         # Check for duplicates

KeyError: 'Year'
```

Find all the instance where : Category is 'Movie' and Type is 'Dramas' or Category is 'TV Show' and Type is 'Kids' TV'

In [68]: `netdata.head(2)`

Out[68]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min

In [69]: `netdata[(netdata['Category']=='Movie')&(netdata['Type']=='Dramas')].head(2)`

Out[69]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Typ
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 min	Drama
7	s8	Movie	187	Kevin Reynolds	Samuel L. Jackson, John Heard, Kelly Rowan, Cl...	United States	November 1, 2019	R	119 min	Drama

In [70]: `netdata[(netdata['Category']=='TV Show')&(netdata['Type']=="Kids' TV")].head(2)`

Out[70]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
111	s112	TV Show	44 Cats	NaN	Sarah Natochenny, Suzy Myers, Simona Berman, E...	Italy	October 1, 2020	TV-Y7	2 Seasons
276	s277	TV Show	Abby Hatcher	Kyran Kelly	Macy Drouin, Wyatt White, Paul Sun-Hyung Lee, ...	United States, Canada	July 1, 2020	TV-Y	1 Season

In [71]: `netdata[(netdata['Category']=='Movie')&(netdata['Type']=='Dramas') | (netdata['Cat`

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duratio
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 m
7	s8	Movie	187	Kevin Reynolds	Samuel L. Jackson, John Heard, Kelly Rowan, Cl...	United States	November 1, 2019	R	119 m
111	s112	TV Show	44 Cats	NaN	Sarah Natochenny, Suzy Myers, Simona Berman, E...	Italy	October 1, 2020	TV-Y7	Seaso
170	s171	Movie	A Family Man	Mark Williams	Gerard Butler, Gretchen Mol, Alison Brie, Will...	Canada, United States	December 15, 2019	R	110 m
232	s233	Movie	A Stoning in Fulham County	Larry Elikann	Ken Olin, Jill Eikenberry, Maureen Mueller, Gr...	United States	October 1, 2011	TV-14	95 m
...	...	...	...	...	...	...	...	...
7668	s7667	TV Show	World of Winx	Iginio Straffi	Haven Paschall, Alysha Deslorieux, Jessica Paq...	Italy, United States	June 16, 2017	TV-Y7	Seaso
7717	s7716	TV Show	Yoko	NaN	Eileen Stevens, Alyson Leigh Rosenfeld, Sarah ...	NaN	June 23, 2018	TV-Y	1 Seas
7719	s7718	TV Show	YOM	NaN	Sairaj, Devyani Dagaonkar, Ketan Singh, Mayur ...	NaN	June 7, 2018	TV-Y7	1 Seas
7758	s7757	TV Show	Z4	NaN	Apollo Costa, Gabriel Santana, Matheus Lustosa...	Brazil	February 22, 2019	TV-PG	Seaso
7761	s7760	TV Show	Zak Storm	NaN	Michael Johnston, Jessica Gee-George, Christin...	United States, France, South	September 13, 2018	TV-Y7	Seaso

Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration
					Korea, Indonesia			

322 rows × 15 columns