

Exploratory Data Analysis With Netflix Dataset

June 22, 2022

1 Import Basic Libraries

```
[1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

```
[2]: data=pd.read_csv('Downloads//8. Netflix Dataset.csv')
```

```
[3]: data.head()
```

```
[3]:  Show_Id Category  Title      Director \
0      s1  TV Show    3%              NaN
1      s2   Movie  07:19  Jorge Michel Grau
2      s3   Movie  23:59    Gilbert Chan
3      s4   Movie     9      Shane Acker
4      s5   Movie    21    Robert Luketic
```

```
                                Cast      Country \
0  João Miguel, Bianca Comparato, Michel Gomes, R...  Brazil
1  Demián Bichir, Héctor Bonilla, Oscar Serrano, ...  Mexico
2  Tedd Chan, Stella Chung, Henley Hii, Lawrence ...  Singapore
3  Elijah Wood, John C. Reilly, Jennifer Connelly...  United States
4  Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...  United States
```

```
      Release_Date Rating  Duration \
0  August 14, 2020  TV-MA  4 Seasons
1  December 23, 2016  TV-MA    93 min
2  December 20, 2018     R    78 min
3  November 16, 2017  PG-13    80 min
4   January 1, 2020  PG-13   123 min
```

```
                                Type \
0  International TV Shows, TV Dramas, TV Sci-Fi &...
1                                Dramas, International Movies
2                                Horror Movies, International Movies
3  Action & Adventure, Independent Movies, Sci-Fi...
```

```

                                Description
0  In a future where the elite inhabit an island ...
1  After a devastating earthquake hits Mexico Cit...
2  When an army recruit is found dead, his fellow...
3  In a postapocalyptic world, rag-doll robots hi...
4  A brilliant group of students become card-coun...

```

```
[4]: data.tail()
```

```

[4]:      Show_Id Category                                Title      Director \
7784    s7783   Movie                                Zozo    Josef Fares
7785    s7784   Movie                                Zubaan   Mozez Singh
7786    s7785   Movie                                Zulu Man in Japan      NaN
7787    s7786   TV Show                                Zumbo's Just Desserts      NaN
7788    s7787   Movie  ZZ TOP: THAT LITTLE OL' BAND FROM TEXAS    Sam Dunn

```

```

                                Cast \
7784  Imad Creidi, Antoinette Turk, Elias Gergi, Car...
7785  Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...
7786                                     Nasty C
7787                                Adriano Zumbo, Rachel Khoo
7788                                     NaN

```

```

                                Country      Release_Date \
7784  Sweden, Czech Republic, United Kingdom, Denmar...  October 19, 2020
7785                                     India      March 2, 2019
7786                                     NaN    September 25, 2020
7787                                     Australia    October 31, 2020
7788                                United Kingdom, Canada, United States    March 1, 2020

```

```

      Rating  Duration                                Type \
7784  TV-MA      99 min                                Dramas, International Movies
7785  TV-14     111 min    Dramas, International Movies, Music & Musicals
7786  TV-MA      44 min  Documentaries, International Movies, Music & M...
7787  TV-PG     1 Season    International TV Shows, Reality TV
7788  TV-MA      90 min    Documentaries, Music & Musicals

```

```

                                Description
7784  When Lebanon's Civil War deprives Zozo of his ...
7785  A scrappy but poor boy worms his way into a ty...
7786  In this documentary, South African rapper Nast...
7787  Dessert wizard Adriano Zumbo looks for the nex...
7788  This documentary delves into the mystique behi...

```

2 Data Preparation and Cleaning

Now will explore our data in more detail like : number of rows and columns peresnt in Dataset, range of values, check and handle missing, incorrect and invalid data and perform some more analysis as per requirment.

```
[5]: data.shape
```

```
[5]: (7789, 11)
```

```
[6]: data.size
```

```
[6]: 85679
```

```
[7]: data.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
```

```
[8]: data.describe().T
```

```
[8]:
```

| | count | unique | top | \ |
|--------------|-------|--------|---|---|
| Show_Id | 7789 | 7787 | s684 | |
| Category | 7789 | 2 | Movie | |
| Title | 7789 | 7787 | The Lost Okoroshi | |
| Director | 5401 | 4050 | Raúl Campos, Jan Suter | |
| Cast | 7071 | 6831 | David Attenborough | |
| Country | 7282 | 681 | United States | |
| Release_Date | 7779 | 1565 | January 1, 2020 | |
| Rating | 7782 | 14 | TV-MA | |
| Duration | 7789 | 216 | 1 Season | |
| Type | 7789 | 492 | Documentaries | |
| Description | 7789 | 7769 | A surly septuagenarian gets another chance at ... | |

| | freq |
|--------------|------|
| Show_Id | 2 |
| Category | 5379 |
| Title | 2 |
| Director | 18 |
| Cast | 18 |
| Country | 2556 |
| Release_Date | 118 |
| Rating | 2865 |
| Duration | 1608 |
| Type | 334 |
| Description | 3 |

```
[9]: data[data.duplicated()]
```

```
[9]:
```

| | Show_Id | Category | Title | Director | \ |
|------|---------|----------|-------------------|----------------|---|
| 6300 | s684 | Movie | Backfire | Dave Patten | |
| 6622 | s6621 | Movie | The Lost Okoroshi | Abba T. Makama | |

| | | Cast | Country | \ |
|------|---|---------------|---------|---|
| 6300 | Black Deniro, Byron "Squally" Vinson, Dominic ... | United States | | |
| 6622 | Seun Ajayi, Judith Audu, Tope Tedela, Ifu Enna... | Nigeria | | |

| | Release_Date | Rating | Duration | \ |
|------|-------------------|--------|----------|---|
| 6300 | April 5, 2019 | TV-MA | 97 min | |
| 6622 | September 4, 2020 | TV-MA | 94 min | |

| | Type | \ |
|------|---------------------------------------|---|
| 6300 | Dramas, Independent Movies, Thrillers | |
| 6622 | Comedies, Dramas, Independent Movies | |

| | Description |
|------|---|
| 6300 | When two would-be robbers accidentally kill a ... |
| 6622 | A disillusioned security guard transforms into... |

```
[10]: data.drop_duplicates(inplace=True)
```

```
[11]: data.shape
```

```
[11]: (7787, 11)
```

```
[12]: data.isnull().sum().sum()
```

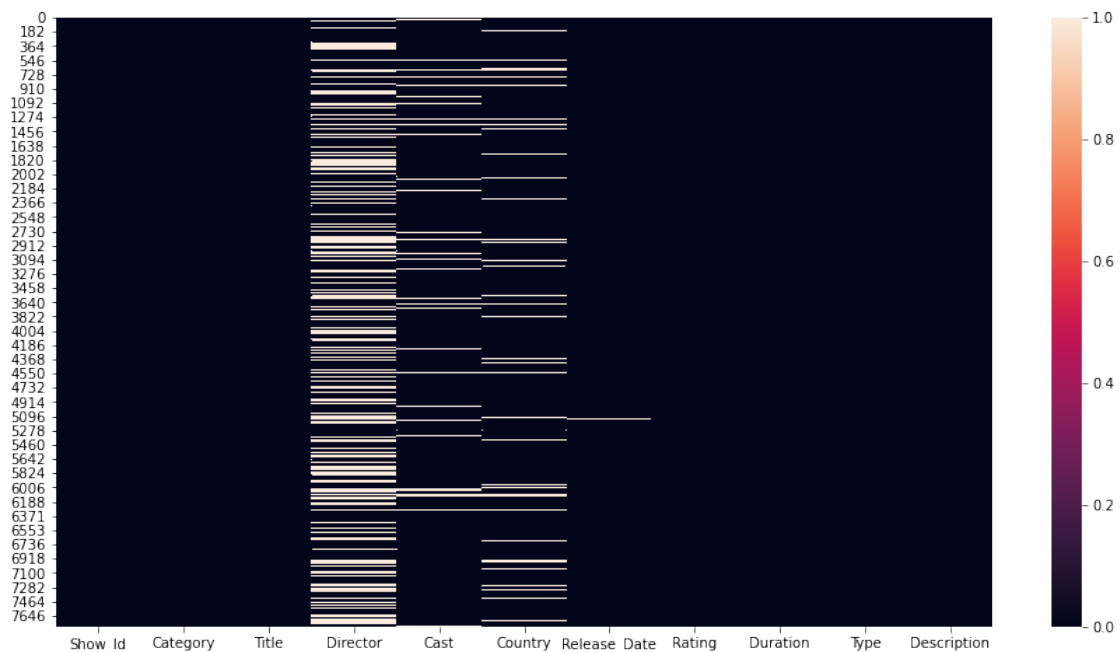
```
[12]: 3630
```

```
[13]: data.isnull().sum()
```

```
[13]: 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
```

```
[14]: plt.figure(figsize=(15,8))
      sns.heatmap(data.isnull())
```

```
[14]: <matplotlib.axes._subplots.AxesSubplot at 0x1a1bbc78af0>
```



Observation:

Director has 2388 missing values
 Cast has 718 missing values
 Country has 507 missing values
 Release_Date has 10 missing values
 Rating has 7 missing values

```
[15]: # creating a copy of dataset so that will not affect our original dataset.  
netflix= data.copy()
```

```
[16]: # dealing with null values  
netflix['Director'].fillna('No Director', inplace=True)  
netflix['Cast'].fillna('No Cast', inplace=True)  
netflix['Country'].fillna('Country Unavailable', inplace=True)  
netflix.dropna(subset=['Release_Date', 'Rating'], inplace=True)
```

```
[17]: #check if there is null value available or not  
netflix.isnull().any()
```

```
[17]: Show_Id      False  
Category      False  
Title         False  
Director      False  
Cast          False  
Country       False  
Release_Date  False  
Rating        False  
Duration      False  
Type          False  
Description   False  
dtype: bool
```

```
[18]: netflix["Release_Date"].value_counts()
```

```
[18]: January 1, 2020      118  
November 1, 2019      94  
December 31, 2019     76  
March 1, 2018         76  
October 1, 2018       72  
  
...  
April 15, 2018         1  
June 8, 2017           1  
December 30, 2019      1  
December 15, 2018      1  
March 16, 2016         1  
Name: Release_Date, Length: 1563, dtype: int64
```

```
[19]: netflix["Release_year"]=netflix["Release_Date"].str.split(" ").str[-1]  
netflix["Release_year"]
```

```
[19]: 0      2020  
1      2016  
2      2018  
3      2017  
4      2020
```

```

...
7784    2020
7785    2019
7786    2020
7787    2020
7788    2020
Name: Release_year, Length: 7770, dtype: object

```

```

[20]: netflix["Release_Day"]=netflix["Release_Date"].str.split(" ").str[-2].str.
      ↪split(",").str[-2]
      netflix["Release_Day"]

```

```

[20]: 0      14
      1      23
      2      20
      3      16
      4       1
      ..
      7784    19
      7785     2
      7786    25
      7787    31
      7788     1
Name: Release_Day, Length: 7770, dtype: object

```

```

[21]: netflix["Release_month"]=netflix["Release_Date"].str.split(" ").str[-3]
      netflix["Release_month"]

```

```

[21]: 0      August
      1    December
      2    December
      3    November
      4     January
      ...
      7784   October
      7785     March
      7786   September
      7787   October
      7788     March
Name: Release_month, Length: 7770, dtype: object

```

3 Exploratory Data Analysis And Visualization

4 For 'house of cards',whats is the show id and who is the director of this show

```
[22]: data[data['Title'].isin(['House of Cards'])]
```

```
[22]:      Show_Id Category      Title \
2832    s2833  TV Show  House of Cards

      Director \
2832  Robin Wright, David Fincher, Gerald McRaney, J...

      Cast      Country \
2832  Kevin Spacey, Robin Wright, Kate Mara, Corey S...  United States

      Release_Date Rating  Duration      Type \
2832  November 2, 2018  TV-MA  6 Seasons  TV Dramas, TV Thrillers

      Description
2832  A ruthless politician will stop at nothing to ...
```

Show id-s2833

Director -Robin Wright, David Fincher, Gerald McRaney

5 Category is movie and type is Dramas

```
[23]: data[(data['Category']=='Movie') & (data['Type']=='Dramas')]
```

```
[23]:      Show_Id Category      Title      Director \
4         s5      Movie          21  Robert Luketic
7         s8      Movie         187  Kevin Reynolds
170       s171      Movie  A Family Man  Mark Williams
232       s233      Movie  A Stoning in Fulham County  Larry Elikann
295       s296      Movie  Across Grace Alley  Ralph Macchio
...      ...      ...      ...      ...
7417     s7416      Movie      Victor  Brandon Dickerson
7422     s7421      Movie      Villain  Philip Barantini
7449     s7448      Movie          VS.      Ed Lilly
7465     s7464      Movie  Wakefield  Robin Swicord
7480     s7479      Movie  War Horse  Steven Spielberg

      Cast \
4  Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...
7  Samuel L. Jackson, John Heard, Kelly Rowan, Cl...
170  Gerard Butler, Gretchen Mol, Alison Brie, Will...
232  Ken Olin, Jill Eikenberry, Maureen Mueller, Gr...
```



```

295 Ben Hyland, Marsha Mason, Karina Smirnoff, Col...
...
7417 Patrick Davis, Lisa Vidal, Josh Pence, José Zú...
7422 Craig Fairbrass, George Russo, Robert Gleniste...
7449 Connor Swindells, Fola Evans-Akingbola, Nichol...
7465 Bryan Cranston, Jennifer Garner, Jason O'Mara,...
7480 Emily Watson, David Thewlis, Peter Mullan, Nie...

```

| | Country | Release_Date | Rating | Duration | Type | \ |
|------|-----------------------|-------------------|--------|----------|--------|---|
| 4 | United States | January 1, 2020 | PG-13 | 123 min | Dramas | |
| 7 | United States | November 1, 2019 | R | 119 min | Dramas | |
| 170 | Canada, United States | December 15, 2019 | R | 110 min | Dramas | |
| 232 | United States | October 1, 2011 | TV-14 | 95 min | Dramas | |
| 295 | United States | December 1, 2016 | TV-PG | 24 min | Dramas | |
| ... | ... | ... | ... | ... | ... | |
| 7417 | United States | July 24, 2017 | PG-13 | 110 min | Dramas | |
| 7422 | United Kingdom | September 5, 2020 | R | 97 min | Dramas | |
| 7449 | United Kingdom | June 19, 2019 | TV-MA | 99 min | Dramas | |
| 7465 | United States | March 2, 2019 | R | 109 min | Dramas | |
| 7480 | United States, India | May 6, 2019 | PG-13 | 147 min | Dramas | |

| | Description |
|------|---|
| 4 | A brilliant group of students become card-coun... |
| 7 | After one of his high school students attacks ... |
| 170 | A ruthless corporate headhunter battles his ri... |
| 232 | After reckless teens kill an Amish child, a pr... |
| 295 | A young boy, upset by his parents' divorce, be... |
| ... | ... |
| 7417 | In 1962 Brooklyn, a Puerto Rican teen who join... |
| 7422 | Fresh out of prison, an ex-con tries to save h... |
| 7449 | A young man in foster care finds his voice in ... |
| 7465 | An unhappy father and lawyer quits his suburba... |
| 7480 | During World War I, the bond between a young E... |

[117 rows x 11 columns]

6 Show all the movies that were released in 2009

```

[24]: #grouping
netflix.groupby(["Category", "Release_year"]).size().reset_index().head(2)

```

```

[24]:  Category Release_year  0
0    Movie           2008  1
1    Movie           2009  2

```

7 Show top 10 directors ,who gives the highest number of TV shows and movies

```
[25]: data['Director'].value_counts().head(10)
```

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

8 Show all the records ,where Category is movie and type is comedy or country is 'United Kingdom'

```
[26]: netflix[(netflix["Category"]=="Movie")&(netflix["Type"]=="Comedy")|(netflix["Country"]=="United Kingdom")]
```

```
[26]:
```

| | Show_Id | Category | Title | Director \ | Cast | Country \ |
|------|---------|----------|------------------------|-------------------|---|----------------|
| 19 | s20 | Movie | '89 | No Director | Lee Dixon, Ian Wright, Paul Merson | United Kingdom |
| 58 | s59 | TV Show | 100% Hotter | No Director | Daniel Palmer, Melissa Sophia, Karen Williams,... | United Kingdom |
| 82 | s83 | Movie | 2036 Origin Unknown | Hasraf Dulull | Katee Sackhoff, Ray Fearon, Julie Cox, Steven ... | United Kingdom |
| 85 | s86 | TV Show | 21 Again | No Director | Laura Morgan | United Kingdom |
| 93 | s94 | Movie | 27: Gone Too Soon | Simon Napier-Bell | Janis Joplin, Jimi Hendrix, Amy Winehouse, Jim... | United Kingdom |
| ... | ... | ... | ... | ... | ... | ... |
| 7670 | s7669 | TV Show | World War II in Colour | No Director | Robert Powell | United Kingdom |
| 7671 | s7670 | TV Show | World's Busiest Cities | No Director | Anita Rani, Ade Adepitan, Dan Snow | United Kingdom |
| 7688 | s7687 | Movie | XV: Beyond the Tryline | Pierre Deschamps | No Cast | United Kingdom |
| 7725 | s7724 | Movie | You Can Tutu | James Brown | | |
| 7740 | s7739 | TV Show | Young Wallander | No Director | | |

| | | |
|------|---|----------------|
| 7725 | Lily O'Regan, Jeannettsy Enriquez Borges, Joel... | United Kingdom |
| 7740 | Adam Pålsson, Richard Dillane, Leanne Best, El... | United Kingdom |

| | Release_Date | Rating | Duration | \ |
|------|-------------------|--------|----------|---|
| 19 | May 16, 2018 | TV-PG | 87 min | |
| 58 | November 1, 2019 | TV-14 | 1 Season | |
| 82 | December 20, 2018 | TV-14 | 95 min | |
| 85 | July 1, 2020 | TV-MA | 1 Season | |
| 93 | May 1, 2018 | TV-MA | 70 min | |
| ... | ... | ... | ... | |
| 7670 | August 1, 2017 | TV-MA | 1 Season | |
| 7671 | February 1, 2019 | TV-PG | 1 Season | |
| 7688 | March 18, 2020 | TV-14 | 91 min | |
| 7725 | December 31, 2017 | TV-G | 87 min | |
| 7740 | September 3, 2020 | TV-MA | 1 Season | |

| | Type | \ |
|------|---|---|
| 19 | Sports Movies | |
| 58 | British TV Shows, International TV Shows, Real... | |
| 82 | Sci-Fi & Fantasy | |
| 85 | British TV Shows, Reality TV | |
| 93 | Documentaries | |
| ... | ... | |
| 7670 | British TV Shows, Docuseries, International TV... | |
| 7671 | British TV Shows, Docuseries | |
| 7688 | Documentaries, Sports Movies | |
| 7725 | Children & Family Movies | |
| 7740 | Crime TV Shows, International TV Shows, TV Dramas | |

| | Description | Release_year | \ |
|------|---|--------------|---|
| 19 | Mixing old footage with interviews, this is th... | 2018 | |
| 58 | A stylist, a hair designer and a makeup artist... | 2019 | |
| 82 | Working with an artificial intelligence to inv... | 2018 | |
| 85 | In a social experiment, a group of daughters s... | 2020 | |
| 93 | Explore the circumstances surrounding the trag... | 2018 | |
| ... | ... | ... | |
| 7670 | Footage of the most dramatic moments from Worl... | 2017 | |
| 7671 | From Moscow to Mexico City, three BBC journali... | 2019 | |
| 7688 | Set against the 2015 Rugby World Cup, this doc... | 2020 | |
| 7725 | A gifted young ballet dancer struggles to find... | 2017 | |
| 7740 | An incendiary hate crime stirs civil unrest, f... | 2020 | |

| | Release_Day | Release_month |
|----|-------------|---------------|
| 19 | 16 | May |
| 58 | 1 | November |
| 82 | 20 | December |
| 85 | 1 | July |

```

93          1          May
...          ...
7670         1          August
7671         1          February
7688        18          March
7725        31          December
7740         3          September

```

```
[396 rows x 14 columns]
```

9 What are different ratings defined by netflix

```
[27]: netflix['Rating'].unique()
```

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

10 How many movies got TV-14 rating in canada

```
[28]: #filtering
netflix[(netflix['Category']=='Movie') & (netflix['Rating']=='TV-14') & (netflix['Country']=='Canada')]
```

```
[28]:
```

| | Show_Id | Category | Title \ |
|------|---------|----------|---|
| 30 | s31 | Movie | #cats_the_mewvie |
| 34 | s35 | Movie | #Roxy |
| 608 | s609 | Movie | At First Light |
| 1506 | s1507 | Movie | Counterfeiting in Suburbia |
| 3713 | s3714 | Movie | Long Time Running |
| 3956 | s3957 | Movie | Martin Matte: La vie, la mort...eh la la..! |
| 5336 | s5336 | Movie | Sadie's Last Days on Earth |
| 5393 | s5393 | Movie | Sardaar ji |
| 6062 | s6062 | Movie | The Accountant of Auschwitz |
| 7025 | s7024 | Movie | The Wrong Babysitter |
| 7041 | s7040 | Movie | There's Something in the Water |

| | Director \ |
|------|--|
| 30 | Michael Margolis |
| 34 | Michael Kennedy |
| 608 | Jason Stone |
| 1506 | Jason Bourque |
| 3713 | Jennifer Baichwal, Nicholas de Pencier |
| 3956 | Jean-François Blais |
| 5336 | Michael Seater |
| 5393 | Rohit Jugraj |
| 6062 | Matthew Shoychet |
| 7025 | George Mendeluk |

7041 Ellen Page, Ian Daniel

| | | Cast Country \ |
|------|---|---------------------|
| 30 | | No Cast Canada |
| 34 | Jake Short, Sarah Fisher, Booboo Stewart, Dann... | Canada |
| 608 | Stefanie Scott, Theodore Pellerin, Percy Hynes... | Canada |
| 1506 | Sarah Butler, Larissa Albuquerque, Kayla Walla... | Canada |
| 3713 | Gord Downie, Paul Langlois, Gord Sinclair, Rob... | Canada |
| 3956 | | Martin Matte Canada |
| 5336 | Morgan Taylor Campbell, Clark Backo, Ricardo H... | Canada |
| 5393 | Diljit Dosanjh, Neeru Bajwa, Mandy Takhar, Jas... | Canada |
| 6062 | | No Cast Canada |
| 7025 | Daphne Zuniga, Seann Gallagher, Lisa Marie Car... | Canada |
| 7041 | | Ellen Page Canada |

| | Release_Date | Rating | Duration \ |
|------|-------------------|--------|------------|
| 30 | February 5, 2020 | TV-14 | 90 min |
| 34 | April 10, 2019 | TV-14 | 105 min |
| 608 | December 27, 2018 | TV-14 | 91 min |
| 1506 | July 1, 2018 | TV-14 | 88 min |
| 3713 | November 12, 2017 | TV-14 | 96 min |
| 3956 | October 14, 2019 | TV-14 | 62 min |
| 5336 | August 22, 2017 | TV-14 | 94 min |
| 5393 | December 1, 2017 | TV-14 | 140 min |
| 6062 | July 1, 2019 | TV-14 | 78 min |
| 7025 | June 18, 2019 | TV-14 | 89 min |
| 7041 | March 27, 2020 | TV-14 | 72 min |

| | Type \ |
|------|---|
| 30 | Documentaries, International Movies |
| 34 | Comedies, Romantic Movies |
| 608 | Dramas, Sci-Fi & Fantasy |
| 1506 | Dramas, Thrillers |
| 3713 | Documentaries, International Movies, Music & M... |
| 3956 | Stand-Up Comedy |
| 5336 | Comedies, Sci-Fi & Fantasy |
| 5393 | Comedies, International Movies, Music & Musicals |
| 6062 | Documentaries, International Movies |
| 7025 | Thrillers |
| 7041 | Documentaries, International Movies |

| | Description | Release_year \ |
|------|---|----------------|
| 30 | This pawesome documentary explores how our fel... | 2020 |
| 34 | A teenage hacker with a huge nose helps a cool... | 2019 |
| 608 | After mysterious lights imbue her with extraor... | 2018 |
| 1506 | Two teenagers print counterfeit money in their... | 2018 |
| 3713 | This documentary follows the 2016 tour of belo... | 2017 |

| | | |
|------|---|------|
| 3956 | Quebecois comedy star Martin Matte serves up e... | 2019 |
| 5336 | High schooler Sadie Mitchell believes the worl... | 2017 |
| 5393 | When a palatial wedding venue turns out to be ... | 2017 |
| 6062 | Decades after WWII, a former SS officer stands... | 2019 |
| 7025 | A widow searches for her missing teenage daugh... | 2019 |
| 7041 | This documentary spotlights the struggle of mi... | 2020 |

| | Release_Day | Release_month |
|------|-------------|---------------|
| 30 | 5 | February |
| 34 | 10 | April |
| 608 | 27 | December |
| 1506 | 1 | July |
| 3713 | 12 | November |
| 3956 | 14 | October |
| 5336 | 22 | August |
| 5393 | 1 | December |
| 6062 | 1 | July |
| 7025 | 18 | June |
| 7041 | 27 | March |

There are total 11 movies ,got rating “TV-14” in Canada

11 What is the maximum duration in the data set

```
[29]: #we need to split the numeric and categorical elements
netflix[['Minutes', 'Unit']] = netflix['Duration'].str.split(' ', expand=True)
netflix.head(2)
```

```
[29]: Show_Id Category Title Director \
0 s1 TV Show 3% No Director
1 s2 Movie 07:19 Jorge Michel Grau
```

```
Cast Country \
0 João Miguel, Bianca Comparato, Michel Gomes, R... Brazil
1 Demián Bichir, Héctor Bonilla, Oscar Serrano, ... Mexico
```

```
Release_Date Rating Duration \
0 August 14, 2020 TV-MA 4 Seasons
1 December 23, 2016 TV-MA 93 min
```

```
Type \
0 International TV Shows, TV Dramas, TV Sci-Fi &...
1 Dramas, International Movies
```

```
Description Release_year Release_Day \
0 In a future where the elite inhabit an island ... 2020 14
1 After a devastating earthquake hits Mexico Cit... 2016 23
```

| | Release_month | Minutes | Unit |
|---|---------------|---------|---------|
| 0 | August | 4 | Seasons |
| 1 | December | 93 | min |

```
[30]: #maximum duration
netflix.Minutes.max()
```

```
[30]: '99'
```

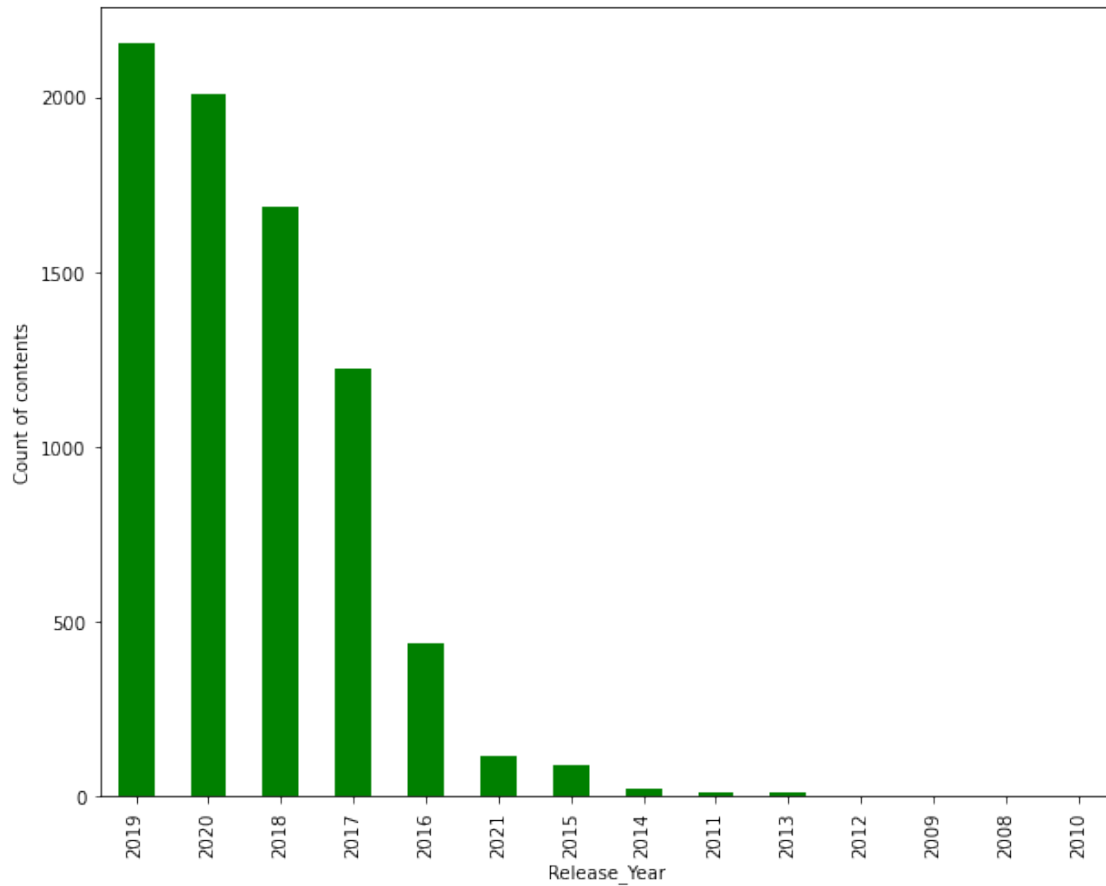
12 In which year Highest number of tv shows and movies were released

```
[31]: #To know the year in which the highest number of TV shows and movies were
↪released We will use value_Counts
netflix['Release_year'].value_counts()
```

```
[31]: 2019    2153
      2020    2009
      2018    1684
      2017    1222
      2016     440
      2021     117
      2015      88
      2014      25
      2011      13
      2013       11
      2012        3
      2009        2
      2008        2
      2010         1
      Name: Release_year, dtype: int64
```

Hence the Answer is year 2019 ,in which highest number of TV shows and movies were released

```
[32]: #With the Graph
plt.figure(figsize=(10,8))
netflix['Release_year'].value_counts().plot(kind='bar',color="g")
plt.xlabel("Release_Year")
plt.ylabel("Count of contents");
```

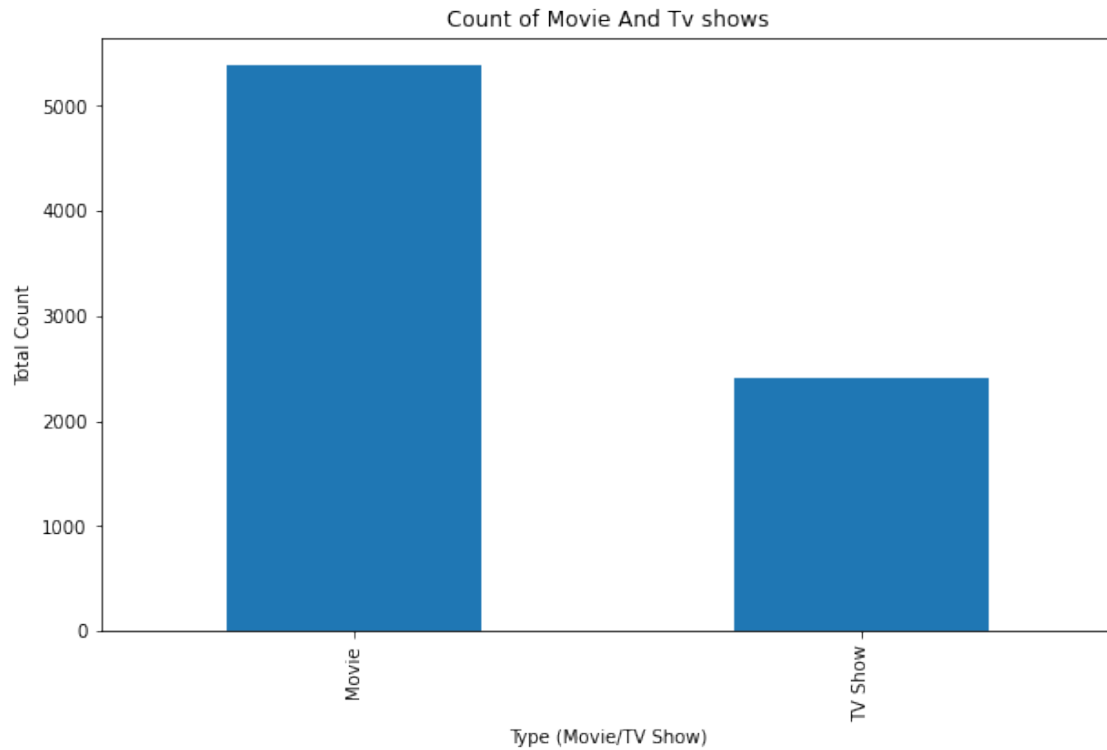


13 How many movies and tv shows are in the dataset

```
[33]: #count the values
netflix.Category.value_counts()
```

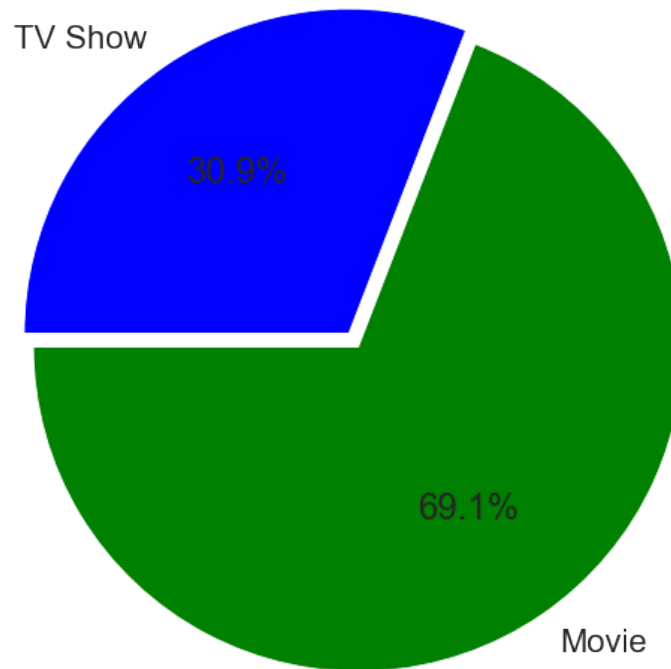
```
[33]: Movie      5372
      TV Show   2398
      Name: Category, dtype: int64
```

```
[34]: #with graph
plt.figure(figsize=(10,6))
data.groupby("Category").Category.count().plot(kind='bar')
plt.title("Count of Movie And Tv shows")
plt.xlabel("Type (Movie/TV Show)")
plt.ylabel("Total Count");
```

```
[35]: #with pie plot
plt.figure(figsize=(15,10))
sns.set(font_scale=2)
plt.title("Percentation of Netflix Categories that are either Movies or TV_
↳Shows")
plt.pie(netflix.Category.value_counts(),labels=netflix.Category.value_counts().
↳index, colors=['green','blue'],autopct='%1.1f%%', startangle=180,explode=(0.
↳025,0.025))
plt.show()
```

Percentation of Netflix Categories that are either Movies or TV Shows



The quantity of Movies is more than TV shows

Only 2 movies released in year 2009

14 Which individual country has highest numbers of tv shows

```
[36]: #filtering data
netflix_tvshow=netflix[netflix['Category']=='TV Show']
```

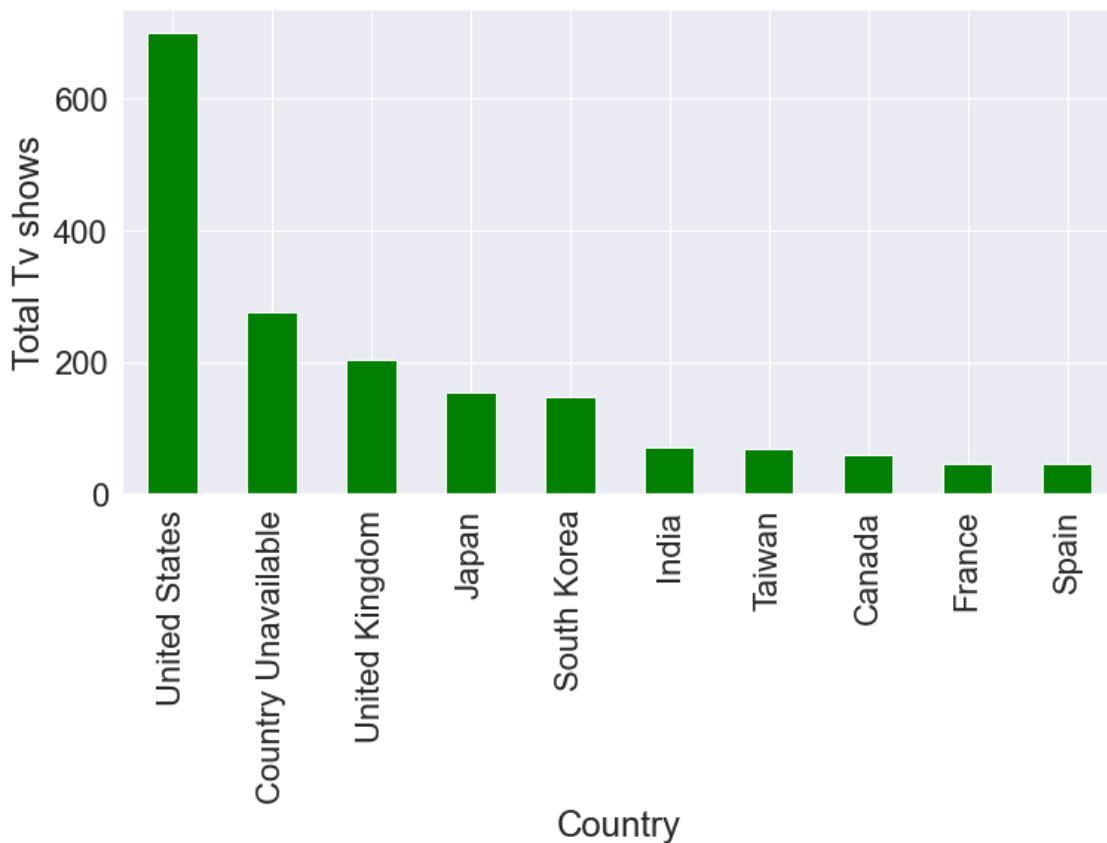
```
[37]: netflix_tvshow.Country.value_counts()
```

```
[37]: United States      699
Country Unavailable  276
United Kingdom      203
Japan               155
South Korea         147
...
France, Canada      1
Jordan              1
United States, Italy 1
```

```
United Kingdom, Spain      1
Philippines                1
Name: Country, Length: 184, dtype: int64
```

United State has highest numbers of TV shows

```
[38]: plt.figure(figsize=(12,6))
netflix_tvshow.Country.value_counts().head(10).plot(kind='bar',color='green')
plt.xlabel("Country")
plt.ylabel("Total Tv shows");
```



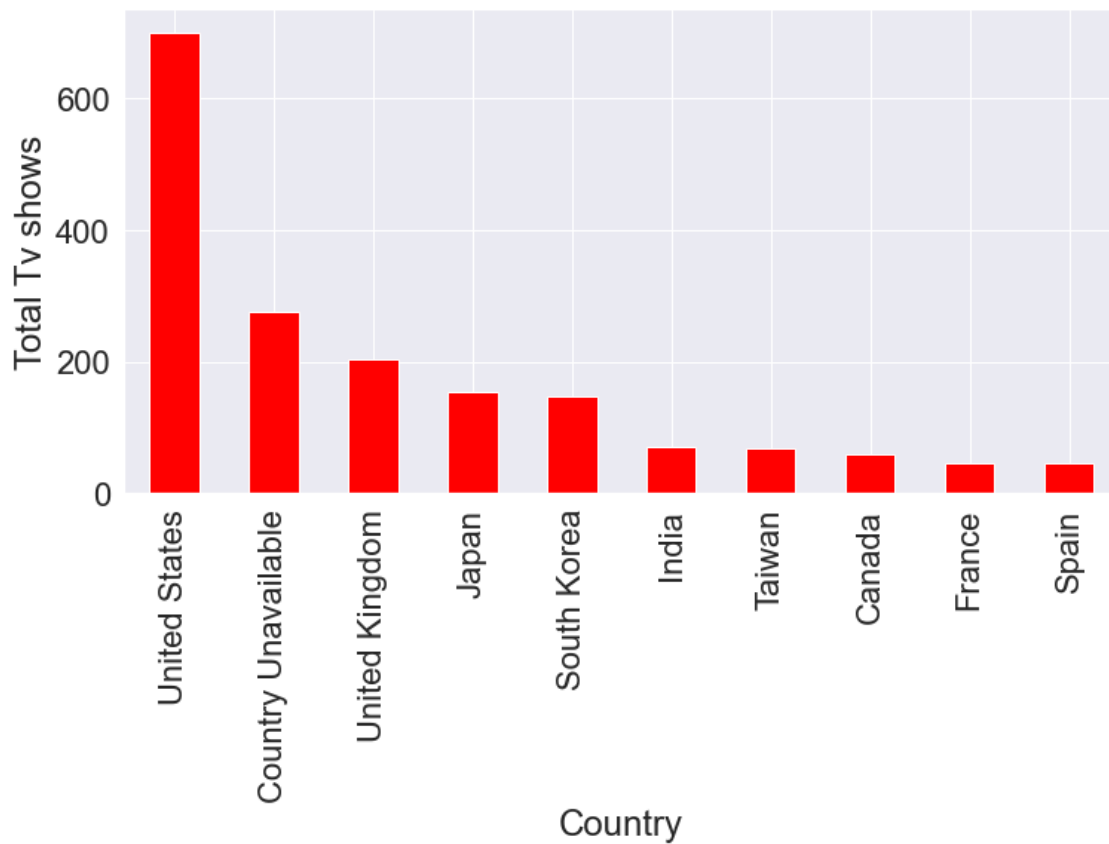
United State has highest numbers of TV shows

15 Which individual country has highest numbers of movies

```
[39]: #Filtering data
netflix_movie=data[data['Category']=='Movie']
```

```
[40]: plt.figure(figsize=(12,6))
netflix_tvshow.Country.value_counts().head(10).plot(kind='bar',color='red')
```

```
plt.xlabel("Country")
plt.ylabel("Total Tv shows");
```

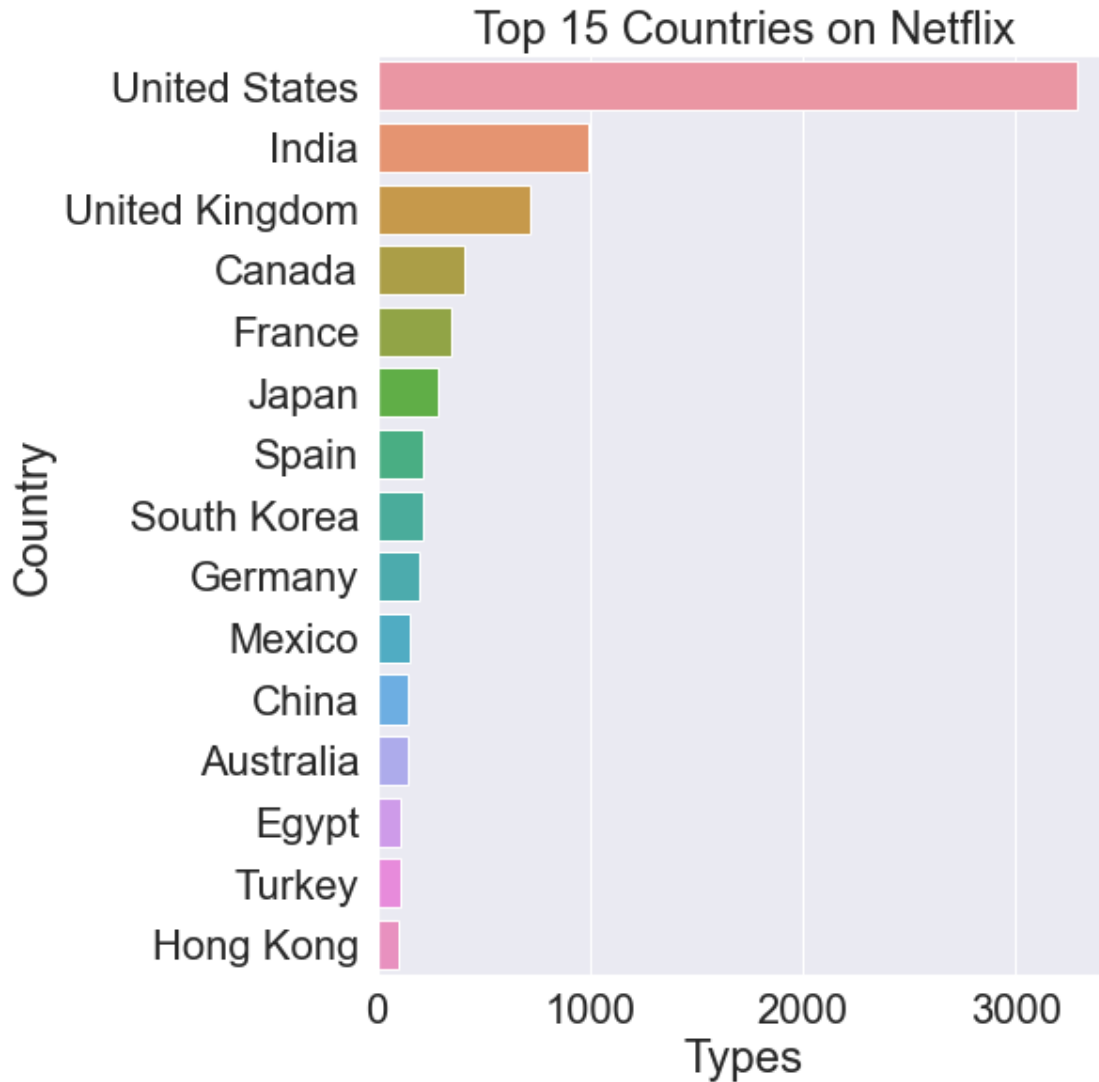


United State has highest numbers of movies

16 Top 15 countries produces most contents

```
[41]: filtered_countries = netflix.set_index('Type').Country.str.split(', ', expand=True).stack().reset_index(level=1, drop=True);
filtered_countries = filtered_countries[filtered_countries != 'Country_
Unavailable']

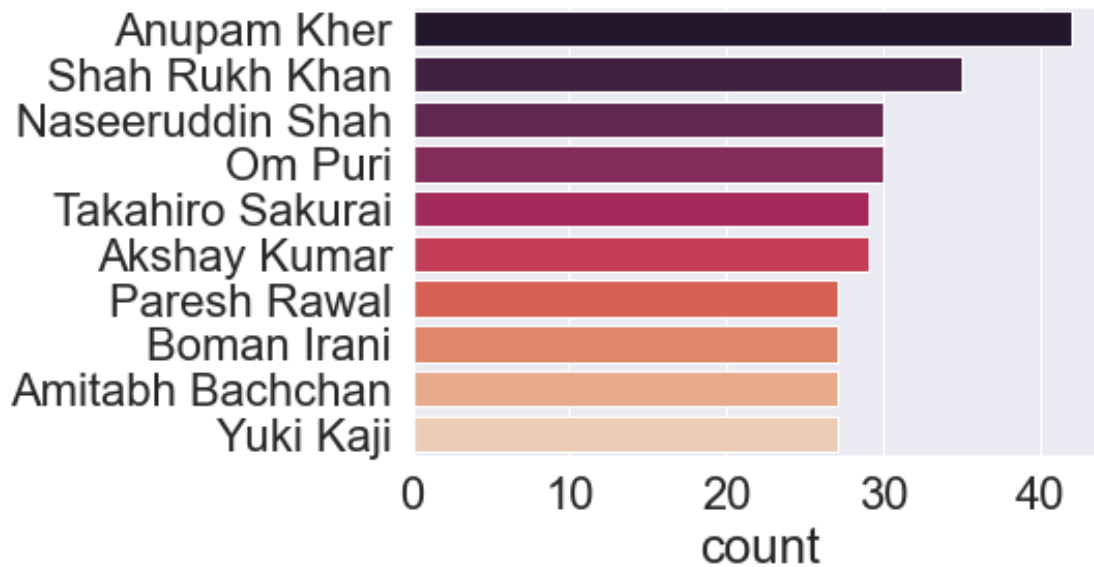
plt.figure(figsize=(7,9))
g = sns.countplot(y = filtered_countries, order=filtered_countries.
    value_counts().index[:15])
plt.title('Top 15 Countries on Netflix')
plt.xlabel('Types')
plt.ylabel('Country')
plt.show()
```



United state make large numbers of contents

17 Who are the top 10 Movies actors on Netflix based on number of titles

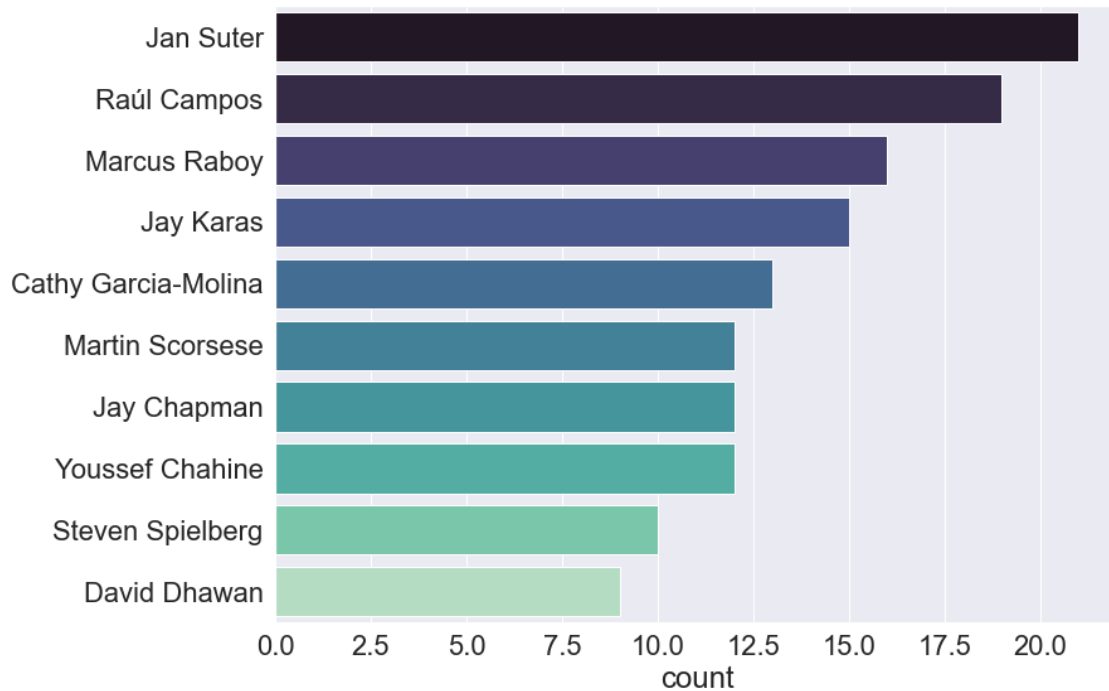
```
[42]: filtered_cast = netflix[netflix.Cast != 'No Cast'].set_index('Title').Cast.str.  
      ↪split(',', expand=True).stack().reset_index(level=1, drop=True)  
      sns.countplot(y = filtered_cast, order=filtered_cast.value_counts().index[:10],  
      ↪palette='rocket')  
      plt.show()
```



The top actor on Netflix Movies, based on the number of titles, is Anupam Kher.

18 Who are the top 10 directors on Netflix with the most releases

```
[43]: plt.figure(figsize=(12,9))
filtered_directors = netflix[netflix.Director != 'No Director'].
    ↳set_index('Title').Director.str.split(', ', expand=True).stack().
    ↳reset_index(level=1, drop=True)
sns.countplot(y = filtered_directors, order=filtered_directors.value_counts().
    ↳index[:10], palette='mako')
plt.show()
```



19 Conclusion

We have drawn many interesting inferences from the dataset Netflix titles; here's a summary of the few of them:

1. The most content type on Netflix is movies.
2. The country by the amount of the produces content is the United States,
3. The most popular director on Netflix , with the most titles, is Jan Suter.
4. International Movies is a genre that is mostly in Netflix.
5. largest count of Netflix content is made with a “TV-14” rating.
6. The most popular actor on Netflix TV Shows based on the number of titles is Takahiro Sakurai.
7. The most popular actor on Netflix movie, based on the number of titles, is Anupam Kher.

It's clear that Netflix has grown over the years. We can see it from the data that the company took certain approaches in their marketing strategy to break into new markets around the world.