

Project 1 - Netflix

April 21, 2022

1 Netflix Project

- The exploration should have a goal. As you explore the data, keep in mind that you want to answer which type of shows to produce and how to grow the business.
 - Ensure each recommendation is backed by data. The company is looking for data-driven insights, not personal opinions or anecdotes.
 - Assume that you are presenting your findings to business executives who have only a basic understanding of data science. Avoid unnecessary technical jargon.
 - Start by exploring a few questions: What type of content is available in different countries?
 - How has the number of movies released per year changed over the last 20-30 years?
 - Comparison of tv shows vs. movies.
 - What is the best time to launch a TV show?
 - Analysis of actors/directors of different types of shows/movies.
-
- Show_id: Unique ID for every Movie / Tv Show
 - Type: Identifier - A Movie or TV Show
 - Title: Title of the Movie / Tv Show
 - Director: Director of the Movie
 - Cast: Actors involved in the movie/show
 - Country: Country where the movie/show was produced
 - Date_added: Date it was added on Netflix
 - Release_year: Actual Release year of the movie/show
 - Rating: TV Rating of the movie/show
 - Duration: Total Duration - in minutes or number of seasons
 - Listed_in: Genre
 - Description: The summary description

[]:

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

[]:

```
[2]: df=pd.read_csv('https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/
↪000/940/original/netflix.csv')
```

```
df.head()
```

```
[2]: show_id    type    title    director \
0      s1      Movie    Dick Johnson Is Dead    Kirsten Johnson
1      s2    TV Show      Blood & Water          NaN
2      s3    TV Show      Ganglands    Julien Leclercq
3      s4    TV Show    Jailbirds New Orleans    NaN
4      s5    TV Show      Kota Factory          NaN

                                cast    country \
0                                NaN    United States
1    Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...    South Africa
2    Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...    NaN
3                                NaN    NaN
4    Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...    India

    date_added    release_year    rating    duration \
0    September 25, 2021    2020    PG-13    90 min
1    September 24, 2021    2021    TV-MA    2 Seasons
2    September 24, 2021    2021    TV-MA    1 Season
3    September 24, 2021    2021    TV-MA    1 Season
4    September 24, 2021    2021    TV-MA    2 Seasons

                                listed_in \
0                                Documentaries
1    International TV Shows, TV Dramas, TV Mysteries
2    Crime TV Shows, International TV Shows, TV Act...
3                                Docuseries, Reality TV
4    International TV Shows, Romantic TV Shows, TV ...

                                description
0    As her father nears the end of his life, filmm...
1    After crossing paths at a party, a Cape Town t...
2    To protect his family from a powerful drug lor...
3    Feuds, flirtations and toilet talk go down amo...
4    In a city of coaching centers known to train I...
```

```
[3]: df.shape
```

```
[3]: (8807, 12)
```

```
[4]: df['show_id'].nunique()
```

```
[4]: 8807
```

```
[5]: df['title'].nunique()
```

```
[5]: 8807
```

```
[6]: df.drop(columns='title',inplace=True)
```

```
[7]: df.head()
```

```
[7]: show_id    type    director \
0      s1    Movie  Kirsten Johnson
1      s2  TV Show           NaN
2      s3  TV Show  Julien Leclercq
3      s4  TV Show           NaN
4      s5  TV Show           NaN

                                cast    country \
0                                NaN  United States
1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...  South Africa
2  Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...    NaN
3                                NaN    NaN
4  Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...    India

    date_added  release_year  rating  duration \
0  September 25, 2021        2020  PG-13    90 min
1  September 24, 2021        2021  TV-MA  2 Seasons
2  September 24, 2021        2021  TV-MA    1 Season
3  September 24, 2021        2021  TV-MA    1 Season
4  September 24, 2021        2021  TV-MA  2 Seasons

                                listed_in \
0                                Documentaries
1  International TV Shows, TV Dramas, TV Mysteries
2  Crime TV Shows, International TV Shows, TV Act...
3                                Docuseries, Reality TV
4  International TV Shows, Romantic TV Shows, TV ...

                                description
0  As her father nears the end of his life, filmm...
1  After crossing paths at a party, a Cape Town t...
2  To protect his family from a powerful drug lor...
3  Feuds, flirtations and toilet talk go down amo...
4  In a city of coaching centers known to train I...
```

```
[8]: df.drop(columns='description',inplace=True)
df.head()
```

```
[8]: show_id    type    director \
0      s1    Movie  Kirsten Johnson
1      s2  TV Show           NaN
```

```

2      s3  TV Show  Julien Leclercq
3      s4  TV Show                NaN
4      s5  TV Show                NaN

```

```

                                cast      country \
0                                NaN  United States
1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...  South Africa
2  Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...      NaN
3                                NaN      NaN
4  Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...      India

```

```

      date_added  release_year  rating  duration \
0  September 25, 2021      2020  PG-13    90 min
1  September 24, 2021      2021  TV-MA  2 Seasons
2  September 24, 2021      2021  TV-MA    1 Season
3  September 24, 2021      2021  TV-MA    1 Season
4  September 24, 2021      2021  TV-MA  2 Seasons

```

```

                                listed_in
0                                Documentaries
1  International TV Shows, TV Dramas, TV Mysteries
2  Crime TV Shows, International TV Shows, TV Act...
3                                Docuseries, Reality TV
4  International TV Shows, Romantic TV Shows, TV ...

```

```
[9]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 10 columns):
#   Column      Non-Null Count  Dtype
---  -
0   show_id     8807 non-null   object
1   type        8807 non-null   object
2   director    6173 non-null   object
3   cast        7982 non-null   object
4   country     7976 non-null   object
5   date_added  8797 non-null   object
6   release_year 8807 non-null   int64
7   rating      8803 non-null   object
8   duration    8804 non-null   object
9   listed_in   8807 non-null   object
dtypes: int64(1), object(9)
memory usage: 688.2+ KB

```

```
[ ]:
```

```
[10]: df.head(10)
```

```
[10]:  show_id      type      director \
0      s1      Movie      Kirsten Johnson
1      s2  TV Show              NaN
2      s3  TV Show      Julien Leclercq
3      s4  TV Show              NaN
4      s5  TV Show              NaN
5      s6  TV Show      Mike Flanagan
6      s7      Movie  Robert Cullen, José Luis Ucha
7      s8      Movie      Haile Gerima
8      s9  TV Show      Andy Devonshire
9     s10      Movie      Theodore Melfi

                                cast \
0                                NaN
1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...
2  Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...
3                                NaN
4  Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...
5  Kate Siegel, Zach Gilford, Hamish Linklater, H...
6  Vanessa Hudgens, Kimiko Glenn, James Marsden, ...
7  Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...
8  Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...
9  Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...

                                country      date_added \
0      United States  September 25, 2021
1      South Africa  September 24, 2021
2              NaN  September 24, 2021
3              NaN  September 24, 2021
4              India  September 24, 2021
5              NaN  September 24, 2021
6              NaN  September 24, 2021
7  United States, Ghana, Burkina Faso, United Kin...  September 24, 2021
8      United Kingdom  September 24, 2021
9      United States  September 24, 2021

  release_year  rating  duration \
0         2020  PG-13    90 min
1         2021  TV-MA  2 Seasons
2         2021  TV-MA    1 Season
3         2021  TV-MA    1 Season
4         2021  TV-MA  2 Seasons
5         2021  TV-MA    1 Season
6         2021    PG    91 min
7         1993  TV-MA   125 min
```

```

8         2021  TV-14  9 Seasons
9         2021  PG-13   104 min

```

```

                                listed_in
0                                Documentaries
1  International TV Shows, TV Dramas, TV Mysteries
2  Crime TV Shows, International TV Shows, TV Act...
3                                Docuseries, Reality TV
4  International TV Shows, Romantic TV Shows, TV ...
5                                TV Dramas, TV Horror, TV Mysteries
6                                Children & Family Movies
7  Dramas, Independent Movies, International Movies
8                                British TV Shows, Reality TV
9                                Comedies, Dramas

```

```
[11]: df['cast'].isna().sum()
```

```
[11]: 825
```

```
[12]: df['cast'].replace(np.nan, 'Others', inplace=True)
```

```
[13]: df['cast'].isna().sum()
```

```
[13]: 0
```

```
[ ]:
```

```
[14]: df['cast']=df['cast'].apply(lambda x : x.split(', '))
```

```
[15]: df['cast'][2]
```

```
[15]: ['Sami Bouajila',
      'Tracy Gotoas',
      'Samuel Jouy',
      'Nabiha Akkari',
      'Sofia Lesaffre',
      'Salim Kechiouche',
      'Noureddine Farihi',
      'Geert Van Rampelberg',
      'Bakary Diombera']
```

```
[16]: df.head()
```

```
[16]:  show_id    type    director \
0      s1  Movie  Kirsten Johnson
1      s2  TV Show             NaN
2      s3  TV Show  Julien Leclercq

```

```

3      s4  TV Show      NaN
4      s5  TV Show      NaN

```

```

                                cast      country \
0                                [Others]  United States
1  [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba...  South Africa
2  [Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nab...      NaN
3                                [Others]      NaN
4  [Mayur More, Jitendra Kumar, Ranjan Raj, Alam ...      India

```

```

      date_added  release_year  rating  duration \
0  September 25, 2021      2020  PG-13    90 min
1  September 24, 2021      2021  TV-MA   2 Seasons
2  September 24, 2021      2021  TV-MA    1 Season
3  September 24, 2021      2021  TV-MA    1 Season
4  September 24, 2021      2021  TV-MA   2 Seasons

```

```

                                listed_in
0                                Documentaries
1  International TV Shows, TV Dramas, TV Mysteries
2  Crime TV Shows, International TV Shows, TV Act...
3                                Docuseries, Reality TV
4  International TV Shows, Romantic TV Shows, TV ...

```

```
[17]: df[(df['director'].isna()) & (df['type']=='TV Show')]
```

```

[17]:   show_id  type director \
1       s2  TV Show      NaN
3       s4  TV Show      NaN
4       s5  TV Show      NaN
10      s11  TV Show      NaN
14      s15  TV Show      NaN
...
8795   s8796  TV Show      NaN
8796   s8797  TV Show      NaN
8797   s8798  TV Show      NaN
8800   s8801  TV Show      NaN
8803   s8804  TV Show      NaN

```

```

                                cast \
1  [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba...
3                                [Others]
4  [Mayur More, Jitendra Kumar, Ranjan Raj, Alam ...
10                                [Others]
14                                [Others]
...
8795 [Mike Liscio, Emily Bauer, Billy Bob Thompson,...

```

```

8796 [Gökhan Atalay, Payidar Tüfekçioğlu, Baran Akb...
8797 [Michael Johnston, Jessica Gee-George, Christi...
8800 [Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen...
8803 [Others]

```

```

country date_added \
1 South Africa September 24, 2021
3 NaN September 24, 2021
4 India September 24, 2021
10 NaN September 24, 2021
14 NaN September 22, 2021
...
8795 Japan, Canada May 1, 2018
8796 Turkey January 17, 2017
8797 United States, France, South Korea, Indonesia September 13, 2018
8800 Pakistan December 15, 2016
8803 NaN July 1, 2019

```

```

release_year rating duration \
1 2021 TV-MA 2 Seasons
3 2021 TV-MA 1 Season
4 2021 TV-MA 2 Seasons
10 2021 TV-MA 1 Season
14 2021 TV-MA 1 Season
...
8795 2015 TV-Y7 2 Seasons
8796 2016 TV-PG 2 Seasons
8797 2016 TV-Y7 3 Seasons
8800 2012 TV-PG 1 Season
8803 2018 TV-Y7 2 Seasons

```

```

listed_in
1 International TV Shows, TV Dramas, TV Mysteries
3 Docuseries, Reality TV
4 International TV Shows, Romantic TV Shows, TV ...
10 Crime TV Shows, Docuseries, International TV S...
14 British TV Shows, Crime TV Shows, Docuseries
...
8795 Anime Series, Kids' TV
8796 International TV Shows, TV Dramas
8797 Kids' TV
8800 International TV Shows, Romantic TV Shows, TV ...
8803 Kids' TV, Korean TV Shows, TV Comedies

```

```
[2446 rows x 10 columns]
```

```
[18]: df[(df['director'].isna()) & (df['type']=='Movie')]
```



```

[18]: show_id  type director \
404      s405  Movie      NaN
470      s471  Movie      NaN
483      s484  Movie      NaN
641      s642  Movie      NaN
717      s718  Movie      NaN
...      ...    ...      ...
8231     s8232  Movie      NaN
8268     s8269  Movie      NaN
8330     s8331  Movie      NaN
8647     s8648  Movie      NaN
8784     s8785  Movie      NaN

                                cast          country \
404                                [Others]          NaN
470          [David Spade, London Hughes, Fortune Feimster]          NaN
483    [Fatih Şahin, Ece Çeşmioğlu, Halit Özgür Sarı,...          NaN
641                                [Others]          NaN
717    [Andy Puddicombe, Evelyn Lewis Prieto, Ginger ...          NaN
...                                ...          ...
8231                                [Chow Yun Fat]          Hong Kong
8268                                [Others]          United Kingdom
8330    [Zo In-sung, Nam Joo-hyuk, Park Sung-woong, Ba...          South Korea
8647    [Vijay Maurya, Naman Jain, Usha Nadkarni, Muke...          India
8784    [Eileen Stevens, Alyson Leigh Rosenfeld, Sarah... Russia, Spain

                                date_added  release_year rating duration \
404          July 22, 2021          2021  TV-MA    85 min
470          July 13, 2021          2021  TV-14    39 min
483          July 9, 2021          2021  TV-MA   102 min
641          June 24, 2021          2021    PG    97 min
717          June 15, 2021          2021  TV-G   273 min
...      ...    ...      ...
8231    September 20, 2018          1983  TV-14   103 min
8268          June 23, 2018          2016  TV-MA    75 min
8330          April 8, 2019          2018  TV-MA   136 min
8647    January 15, 2017          2015  TV-14    89 min
8784          June 23, 2018          2015  TV-Y    78 min

                                listed_in
404                                Documentaries
470                                Movies
483    Dramas, International Movies, Romantic Movies
641                                Documentaries, Sports Movies
717                                Documentaries
...      ...
8231    Action & Adventure, Dramas, International Movies

```

```

8268 Action & Adventure, Independent Movies, Intern...
8330 Action & Adventure, Dramas, International Movies
8647 Dramas, International Movies
8784 Children & Family Movies

```

```
[188 rows x 10 columns]
```

```
[ ]:
```

```
[19]: df['director'].replace(np.nan, 'No Director', inplace=True)
```

```
[20]: df.head(30)
```

```
[20]:
```

	show_id	type	director \
0	s1	Movie	Kirsten Johnson
1	s2	TV Show	No Director
2	s3	TV Show	Julien Leclercq
3	s4	TV Show	No Director
4	s5	TV Show	No Director
5	s6	TV Show	Mike Flanagan
6	s7	Movie	Robert Cullen, José Luis Ucha
7	s8	Movie	Haile Gerima
8	s9	TV Show	Andy Devonshire
9	s10	Movie	Theodore Melfi
10	s11	TV Show	No Director
11	s12	TV Show	Kongkiat Komesiri
12	s13	Movie	Christian Schwochow
13	s14	Movie	Bruno Garotti
14	s15	TV Show	No Director
15	s16	TV Show	No Director
16	s17	Movie	Pedro de Echave García, Pablo Azorín Williams
17	s18	TV Show	No Director
18	s19	Movie	Adam Salky
19	s20	TV Show	No Director
20	s21	TV Show	Olivier Megaton
21	s22	TV Show	No Director
22	s23	Movie	K.S. Ravikumar
23	s24	Movie	Alex Woo, Stanley Moore
24	s25	Movie	S. Shankar
25	s26	TV Show	No Director
26	s27	Movie	Rajiv Menon
27	s28	Movie	Dennis Dugan
28	s29	Movie	Scott Stewart
29	s30	Movie	Robert Luketic

```

                                cast \
0                                [Others]

```

1 [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba...
 2 [Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nab...
 3 [Others]
 4 [Mayur More, Jitendra Kumar, Ranjan Raj, Alam ...
 5 [Kate Siegel, Zach Gilford, Hamish Linklater, ...
 6 [Vanessa Hudgens, Kimiko Glenn, James Marsden,...
 7 [Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra ...
 8 [Mel Giedroyc, Sue Perkins, Mary Berry, Paul H...
 9 [Melissa McCarthy, Chris O'Dowd, Kevin Kline, ...
 10 [Others]
 11 [Sukollawat Kanarot, Sushar Manaying, Pavarit ...
 12 [Luna Wedler, Jannis Niewöhner, Milan Peschel,...
 13 [Klara Castanho, Lucca Picon, Júlia Gomes, Mar...
 14 [Others]
 15 [Logan Browning, Brandon P. Bell, DeRon Horton...
 16 [Others]
 17 [Luis Ernesto Franco, Camila Sodi, Sergio Goyr...
 18 [Freida Pinto, Logan Marshall-Green, Robert Jo...
 19 [Blanca Suárez, Iván Marcos, Óscar Casas, Adri...
 20 [Others]
 21 [Engin Altan Düzyatan, Serdar Gökhan, Hülya Da...
 22 [Kamal Hassan, Meena, Gemini Ganesan, Heera Ra...
 23 [Maisie Benson, Paul Killam, Kerry Gudjohnsen,...
 24 [Prashanth, Aishwarya Rai Bachchan, Sri Lakshm...
 25 [Brooke Satchwell]
 26 [Arvind Swamy, Kajol, Prabhu Deva, Nassar, S.P...
 27 [Adam Sandler, Kevin James, Chris Rock, David ...
 28 [Keri Russell, Josh Hamilton, J.K. Simmons, Da...
 29 [Liam Hemsworth, Gary Oldman, Amber Heard, Har...

	country	date_added \
0	United States	September 25, 2021
1	South Africa	September 24, 2021
2	NaN	September 24, 2021
3	NaN	September 24, 2021
4	India	September 24, 2021
5	NaN	September 24, 2021
6	NaN	September 24, 2021
7	United States, Ghana, Burkina Faso, United Kin...	September 24, 2021
8	United Kingdom	September 24, 2021
9	United States	September 24, 2021
10	NaN	September 24, 2021
11	NaN	September 23, 2021
12	Germany, Czech Republic	September 23, 2021
13	NaN	September 22, 2021
14	NaN	September 22, 2021
15	United States	September 22, 2021

16		NaN	September 22, 2021
17		Mexico	September 22, 2021
18		NaN	September 22, 2021
19		NaN	September 22, 2021
20		NaN	September 22, 2021
21		Turkey	September 22, 2021
22		NaN	September 21, 2021
23		NaN	September 21, 2021
24		India	September 21, 2021
25		Australia	September 21, 2021
26		NaN	September 21, 2021
27		United States	September 20, 2021
28		United States	September 19, 2021
29		United States, India, France	September 19, 2021

	release_year	rating	duration \
0	2020	PG-13	90 min
1	2021	TV-MA	2 Seasons
2	2021	TV-MA	1 Season
3	2021	TV-MA	1 Season
4	2021	TV-MA	2 Seasons
5	2021	TV-MA	1 Season
6	2021	PG	91 min
7	1993	TV-MA	125 min
8	2021	TV-14	9 Seasons
9	2021	PG-13	104 min
10	2021	TV-MA	1 Season
11	2021	TV-MA	1 Season
12	2021	TV-MA	127 min
13	2021	TV-PG	91 min
14	2021	TV-MA	1 Season
15	2021	TV-MA	4 Seasons
16	2020	TV-MA	67 min
17	2020	TV-MA	2 Seasons
18	2021	TV-14	94 min
19	2021	TV-MA	1 Season
20	2021	TV-14	1 Season
21	2018	TV-14	5 Seasons
22	1996	TV-PG	161 min
23	2021	TV-Y	61 min
24	1998	TV-14	166 min
25	2021	TV-14	2 Seasons
26	1997	TV-PG	147 min
27	2010	PG-13	103 min
28	2013	PG-13	97 min
29	2013	PG-13	106 min

```

                                listed_in
0                                Documentaries
1    International TV Shows, TV Dramas, TV Mysteries
2    Crime TV Shows, International TV Shows, TV Act...
3                                Docuseries, Reality TV
4    International TV Shows, Romantic TV Shows, TV ...
5                                TV Dramas, TV Horror, TV Mysteries
6                                Children & Family Movies
7    Dramas, Independent Movies, International Movies
8                                British TV Shows, Reality TV
9                                Comedies, Dramas
10   Crime TV Shows, Docuseries, International TV S...
11   Crime TV Shows, International TV Shows, TV Act...
12                                Dramas, International Movies
13                                Children & Family Movies, Comedies
14    British TV Shows, Crime TV Shows, Docuseries
15                                TV Comedies, TV Dramas
16                                Documentaries, International Movies
17   Crime TV Shows, Spanish-Language TV Shows, TV ...
18                                Thrillers
19   International TV Shows, Spanish-Language TV Sh...
20   Crime TV Shows, Docuseries, International TV S...
21   International TV Shows, TV Action & Adventure,...
22                                Comedies, International Movies
23                                Children & Family Movies
24    Comedies, International Movies, Romantic Movies
25    Docuseries, International TV Shows, Reality TV
26   Comedies, International Movies, Music & Musicals
27                                Comedies
28                                Horror Movies, Sci-Fi & Fantasy
29                                Thrillers

```

```
[21]: df.tail(30)
```

```

[21]:   show_id    type    director \
8777   s8778   Movie  Saratswadee Wongsomphet
8778   s8779   Movie  Saratswadee Wongsomphet
8779   s8780   Movie      Kirati Nakintanon
8780   s8781  TV Show      No Director
8781   s8782   Movie      Mark Risley
8782   s8783   Movie      Kevin Smith
8783   s8784  TV Show      No Director
8784   s8785   Movie      No Director
8785   s8786  TV Show      No Director
8786   s8787   Movie      James Brown
8787   s8788   Movie      Kenny Young
8788   s8789   Movie      Ivona Juka

```

8789	s8790	Movie	Cathy Garcia-Molina
8790	s8791	Movie	Dennis Dugan
8791	s8792	Movie	Jason Reitman
8792	s8793	Movie	Mu Chu
8793	s8794	Movie	Raja Gosnell
8794	s8795	Movie	Mohamed Diab
8795	s8796	TV Show	No Director
8796	s8797	TV Show	No Director
8797	s8798	TV Show	No Director
8798	s8799	Movie	Chandra Prakash Dwivedi
8799	s8800	Movie	Avadhoot Gupte
8800	s8801	TV Show	No Director
8801	s8802	Movie	Majid Al Ansari
8802	s8803	Movie	David Fincher
8803	s8804	TV Show	No Director
8804	s8805	Movie	Ruben Fleischer
8805	s8806	Movie	Peter Hewitt
8806	s8807	Movie	Mozes Singh

cast \

8777	[Sushar Manaying, Supanart Jittaleela, Arisara...
8778	[Supanart Jittaleela, Sushar Manaying, Nisa Bo...
8779	[Supanart Jittaleela, Pimpakan Bangchawong, Ch...
8780	[Johnny Yong Bosch, J.W. Terry, Alicyn Packard...
8781	[Johnny Yong Bosch, Joey D'Auria, Alicyn Packa...
8782	[Harley Quinn Smith, Lily-Rose Depp, Johnny De...
8783	[Eileen Stevens, Alyson Leigh Rosenfeld, Sarah...
8784	[Eileen Stevens, Alyson Leigh Rosenfeld, Sarah...
8785	[Sairaj, Devyani Dagaonkar, Ketan Singh, Mayur...
8786	[Lily O'Regan, Jeannettsy Enriquez Borges, Joe...
8787	[Brely Evans, Andra Fuller, Persia White, Pors...
8788	[Lana Baric, Vojislav Brajovic, Natasa Janjic,...
8789	[John Lloyd Cruz, Sarah Geronimo, Rayver Cruz,...
8790	[Adam Sandler, John Turturro, Emmanuelle Chriq...
8791	[Charlize Theron, Patton Oswalt, Patrick Wilso...
8792	[Qiu Yuen, Charlie Chin, Jackie Chan, Hu Chin,...
8793	[Dennis Quaid, Rene Russo, Sean Faris, Katija ...
8794	[Nelly Karim, Hany Adel, Tarek Abdel Aziz, Ahm...
8795	[Mike Liscio, Emily Bauer, Billy Bob Thompson,...
8796	[Gökhan Atalay, Payidar Tüfekçioğlu, Baran Akb...
8797	[Michael Johnston, Jessica Gee-George, Christi...
8798	[Adil Hussain, Mona Singh, K.K. Raina, Sanjay ...
8799	[Santosh Juvekar, Siddharth Chandekar, Sachit ...
8800	[Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen...
8801	[Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri,...
8802	[Mark Ruffalo, Jake Gyllenhaal, Robert Downey ...
8803	[Others]

8804 [Jesse Eisenberg, Woody Harrelson, Emma Stone,...
 8805 [Tim Allen, Courteney Cox, Chevy Chase, Kate M...
 8806 [Vicky Kaushal, Sarah-Jane Dias, Raaghav Chana...

	country	date_added \
8777	Thailand	December 6, 2018
8778	Thailand	November 8, 2018
8779	Thailand	November 8, 2018
8780	United States	April 1, 2016
8781	Japan, United States	December 1, 2016
8782	United States	December 2, 2016
8783	NaN	June 23, 2018
8784	Russia, Spain	June 23, 2018
8785	NaN	June 7, 2018
8786	United Kingdom	December 31, 2017
8787	United States	December 19, 2017
8788	Croatia, Slovenia, Serbia, Montenegro	July 1, 2016
8789	Philippines	February 27, 2019
8790	United States	September 1, 2019
8791	United States	November 20, 2019
8792	Hong Kong	November 1, 2016
8793	United States	November 20, 2019
8794	Egypt, France	October 11, 2018
8795	Japan, Canada	May 1, 2018
8796	Turkey	January 17, 2017
8797	United States, France, South Korea, Indonesia	September 13, 2018
8798	India	December 31, 2019
8799	India	February 15, 2018
8800	Pakistan	December 15, 2016
8801	United Arab Emirates, Jordan	March 9, 2016
8802	United States	November 20, 2019
8803	NaN	July 1, 2019
8804	United States	November 1, 2019
8805	United States	January 11, 2020
8806	India	March 2, 2019

	release_year	rating	duration \
8777	2010	TV-PG	111 min
8778	2012	TV-PG	112 min
8779	2015	TV-PG	107 min
8780	2015	TV-Y7	1 Season
8781	2016	TV-Y7	96 min
8782	2016	PG-13	88 min
8783	2016	TV-Y	1 Season
8784	2015	TV-Y	78 min
8785	2016	TV-Y7	1 Season
8786	2017	TV-G	87 min

8787	2017	TV-PG	85 min
8788	2015	TV-MA	157 min
8789	2009	TV-PG	116 min
8790	2008	UR	113 min
8791	2011	R	94 min
8792	1973	NR	81 min
8793	2005	PG	88 min
8794	2016	TV-14	98 min
8795	2015	TV-Y7	2 Seasons
8796	2016	TV-PG	2 Seasons
8797	2016	TV-Y7	3 Seasons
8798	2014	TV-MA	131 min
8799	2009	TV-14	120 min
8800	2012	TV-PG	1 Season
8801	2015	TV-MA	96 min
8802	2007	R	158 min
8803	2018	TV-Y7	2 Seasons
8804	2009	R	88 min
8805	2006	PG	88 min
8806	2015	TV-14	111 min

listed_in

8777	International Movies, LGBTQ Movies, Romantic M...
8778	International Movies, LGBTQ Movies, Romantic M...
8779	International Movies, LGBTQ Movies, Romantic M...
8780	Anime Series, Kids' TV
8781	Anime Features, Children & Family Movies
8782	Comedies, Horror Movies, Independent Movies
8783	Kids' TV
8784	Children & Family Movies
8785	Kids' TV
8786	Children & Family Movies
8787	Comedies, Romantic Movies
8788	Dramas, International Movies
8789	Comedies, Dramas, International Movies
8790	Action & Adventure, Comedies
8791	Comedies, Dramas, Independent Movies
8792	Action & Adventure, International Movies
8793	Children & Family Movies, Comedies
8794	Dramas, Independent Movies, International Movies
8795	Anime Series, Kids' TV
8796	International TV Shows, TV Dramas
8797	Kids' TV
8798	Comedies, Dramas, International Movies
8799	Dramas, International Movies
8800	International TV Shows, Romantic TV Shows, TV ...
8801	Dramas, International Movies, Thrillers

8802	Cult Movies, Dramas, Thrillers
8803	Kids' TV, Korean TV Shows, TV Comedies
8804	Comedies, Horror Movies
8805	Children & Family Movies, Comedies
8806	Dramas, International Movies, Music & Musicals

```
[22]: df[5950:6000]
```

```
[22]:
```

	show_id	type	director \
5950	s5951	Movie	Richard Michaels
5951	s5952	Movie	Linda Otto
5952	s5953	Movie	Roger Young
5953	s5954	Movie	Rob LaDuca, Robert C. Ramirez
5954	s5955	Movie	Icíar Bollain
5955	s5956	Movie	Joe Dante
5956	s5957	Movie	Ole Bornedal
5957	s5958	Movie	Sorin Dan Mihalcescu
5958	s5959	Movie	Shane Acker
5959	s5960	Movie	Robert Luketic
5960	s5961	Movie	Kevin Reynolds
5961	s5962	Movie	Shravan Kumar
5962	s5963	Movie	Vikram Bhatt
5963	s5964	Movie	Kunle Afolayan
5964	s5965	TV Show	No Director
5965	s5966	Movie	Paul Greengrass
5966	s5967	Movie	Swapnaneel Jayakar
5967	s5968	Movie	No Director
5968	s5969	Movie	Vijay Milton
5969	s5970	Movie	Santwana Bardoloi
5970	s5971	Movie	Atanu Ghosh
5971	s5972	Movie	Lyric R. Cabral, David Felix Sutcliffe
5972	s5973	Movie	Michael Margolis
5973	s5974	Movie	Michael Kennedy
5974	s5975	TV Show	No Director
5975	s5976	Movie	Óskar Thór Axelsson
5976	s5977	Movie	Ozan Açıktan
5977	s5978	Movie	Karyn Kusama
5978	s5979	TV Show	No Director
5979	s5980	Movie	Adam Deyoe
5980	s5981	Movie	Leif Tilden
5981	s5982	Movie	Roland Emmerich
5982	s5983	Movie	No Director
5983	s5984	Movie	Melinda Janko
5984	s5985	TV Show	No Director
5985	s5986	Movie	Sam Upton
5986	s5987	TV Show	No Director
5987	s5988	Movie	Victor Zarcoff

5988	s5989	Movie	Daniel Stamm
5989	s5990	Movie	No Director
5990	s5991	Movie	Daniel Lee
5991	s5992	Movie	Scott Hussion, Seth Fuller
5992	s5993	Movie	Richard Donner
5993	s5994	Movie	Burr Steers
5994	s5995	Movie	David Marmor
5995	s5996	Movie	Raymond Wood
5996	s5997	Movie	Morgan Neville
5997	s5998	Movie	No Director
5998	s5999	Movie	Hasraf Dulull
5999	s6000	Movie	Mike Mills

cast \

5950	[Mario Van Peebles, Susan Ruttan, Lane R. Davi...
5951	[Jill Clayburgh, Brad Davis, Sam Behrens]
5952	[William Devane, Jeri Ryan, Michele Abrams, Ni...
5953	[Ben Affleck, Mark Hamill, Richard Herd, Maure...
5954	[Luis Tosar, Gael García Bernal, Juan Carlos A...
5955	[Corey Feldman, Tony Todd, Tara Leigh, Erin Wa...
5956	[Anders W. Berthelsen, Rebecka Hemse, Nikolaj ...
5957	[Barbara King, Shaana Diya, John Krisiukenas, ...
5958	[Elijah Wood, John C. Reilly, Jennifer Connell...
5959	[Jim Sturgess, Kevin Spacey, Kate Bosworth, Aa...
5960	[Samuel L. Jackson, John Heard, Kelly Rowan, C...
5961	[Divya Dutta, Atul Kulkarni, Mohan Agashe, Anu...
5962	[Rajneesh Duggal, Adah Sharma, Indraneil Sengu...
5963	[Sadiq Daba, David Bailie, Kayode Olaiya, Kehi...
5964	[Shahd El Yaseen, Shaila Sabt, Hala, Hanadi Al...
5965	[Anders Danielsen Lie, Jon Øigarden, Jonas Str...
5966	[Rahul Pethe, Mrunmayee Deshpande, Adinath Kot...
5967	[Lee Dixon, Ian Wright, Paul Merson]
5968	[Samuthirakani, Bharath Seenii, Vinoth, Esakki ...
5969	[Adil Hussain, Shakil Imtiaz, Mahendra Rabha, ...
5970	[Soumitra Chatterjee, Prasenjit Chatterjee, In...
5971	[Others]
5972	[Others]
5973	[Jake Short, Sarah Fisher, Booboo Stewart, Dan...
5974	[Chutavuth Pattarakampol, Sheranut Yusananda, ...
5975	[Jóhannes Haukur Jóhannesson, Ágústa Eva Erlenn...
5976	[Ayça Erturan, Aydan Taş, Ayşegül Akdemir, Bur...
5977	[Charlize Theron, Marton Csokas, Jonny Lee Mil...
5978	[Alican Yücesoy, Melisa Sözen, Musa Uzunlar, S...
5979	[Lexi Giovagnoli, Justin Ray, Rae Latt, Poonam...
5980	[Billy Crudup, Graham Rogers, Liana Liberato, ...
5981	[Steven Strait, Camilla Belle, Cliff Curtis, J...
5982	[Others]

5983 [Others]
 5984 [Daniel Palmer, Melissa Sophia, Karen Williams...
 5985 [Sam Upton, Jared Abrahamson, Mark Boone Junio...
 5986 [So-yeon Lee, Namkoong Min, Tae-im Lee, So-hui...
 5987 [PJ McCabe, Brianne Moncrief, Sarah Baldwin, J...
 5988 [Mark Webber, Rutina Wesley, Devon Graye, Tom ...
 5989 [Oprah Winfrey, Ava DuVernay]
 5990 [Donnie Yen, Zhao Wei, Wu Chun, Law Kar-Ying, ...
 5991 [Neville Archambault, Amber Midthunder, Brytne...
 5992 [Bruce Willis, Mos Def, David Morse, Jenna Ste...
 5993 [Zac Efron, Leslie Mann, Matthew Perry, Thomas...
 5994 [Nicole Brydon Bloom, Giles Matthey, Taylor Ni...
 5995 [Hayley Lovitt, Teddy Cole, Brook Todd, Ace Ha...
 5996 [Darlene Love, Merry Clayton, Lisa Fischer, Tá...
 5997 [4Minute, B1A4, BtoB, ELSIE, EXID, EXO, Got7, ...
 5998 [Katee Sackhoff, Ray Fearon, Julie Cox, Steven...
 5999 [Annette Bening, Elle Fanning, Greta Gerwig, L...

	country	date_added	release_year \
5950	United States	October 1, 2011	1991
5951	United States	October 1, 2011	1990
5952	United States	October 1, 2011	1991
5953	United States, France, United Kingdom	September 27, 2011	2000
5954	Spain, Mexico, France	May 17, 2011	2010
5955	United States	November 18, 2009	2009
5956	Denmark	May 5, 2009	2007
5957	United States	January 1, 2008	2006
5958	United States	November 16, 2017	2009
5959	United States	January 1, 2020	2008
5960	United States	November 1, 2019	1997
5961	India	April 1, 2019	2019
5962	India	December 15, 2017	2008
5963	Nigeria	September 1, 2019	2014
5964	NaN	March 20, 2019	2018
5965	Norway, Iceland, United States	October 10, 2018	2018
5966	India	March 29, 2019	2019
5967	United Kingdom	May 16, 2018	2017
5968	India	September 15, 2018	2018
5969	India	September 15, 2018	2017
5970	India	September 15, 2018	2017
5971	United States	June 30, 2016	2015
5972	Canada	February 5, 2020	2020
5973	Canada	April 10, 2019	2018
5974	NaN	March 30, 2019	2016
5975	Iceland	March 27, 2018	2017
5976	Turkey	March 10, 2017	2010
5977	United States	February 1, 2018	2005

5978	Turkey	January 17, 2017	2013
5979	United States	July 1, 2017	2014
5980	United States	July 7, 2017	2017
5981	United States, South Africa	June 1, 2019	2008
5982	Spain	July 6, 2018	2018
5983	United States	March 20, 2018	2016
5984	United Kingdom	November 1, 2019	2017
5985	United States	March 14, 2019	2017
5986	South Korea	May 22, 2017	2014
5987	United States	August 13, 2016	2015
5988	United States	January 13, 2019	2014
5989	NaN	January 26, 2017	2017
5990	Hong Kong, China, Singapore	April 3, 2019	2010
5991	United States	October 27, 2018	2018
5992	United States, Germany	November 1, 2019	2006
5993	United States	January 1, 2021	2009
5994	United States	August 23, 2020	2019
5995	United States	July 26, 2019	2018
5996	United States	September 22, 2018	2013
5997	South Korea	April 28, 2017	2015
5998	United Kingdom	December 20, 2018	2018
5999	United States	June 28, 2019	2016

	rating	duration	listed_in
5950	TV-PG	93 min	Dramas, Sports Movies
5951	TV-14	95 min	Dramas
5952	TV-14	93 min	Dramas, Thrillers
5953	TV-PG	75 min	Children & Family Movies, Dramas, Faith & Spir...
5954	TV-MA	103 min	Dramas, International Movies
5955	TV-MA	29 min	Horror Movies
5956	TV-MA	104 min	Dramas, International Movies
5957	TV-MA	81 min	Dramas, Independent Movies, Thrillers
5958	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...
5959	PG-13	123 min	Dramas
5960	R	119 min	Dramas
5961	TV-14	118 min	Horror Movies, International Movies
5962	TV-MA	143 min	Horror Movies, International Movies, Thrillers
5963	TV-14	149 min	Dramas, International Movies, Thrillers
5964	TV-14	1 Season	International TV Shows, TV Dramas
5965	R	144 min	Dramas, Thrillers
5966	TV-14	124 min	Comedies, Dramas, Independent Movies
5967	TV-PG	87 min	Sports Movies
5968	TV-14	128 min	Action & Adventure, Dramas, International Movies
5969	TV-14	117 min	Dramas, International Movies
5970	TV-14	100 min	Dramas, International Movies
5971	NR	84 min	Documentaries
5972	TV-14	90 min	Documentaries, International Movies

5973	TV-14	105 min		Comedies, Romantic Movies
5974	TV-14	1 Season	International TV Shows, Romantic TV Shows, TV ...	
5975	TV-MA	106 min		Horror Movies, International Movies
5976	TV-MA	99 min		Comedies, International Movies
5977	PG-13	93 min		Action & Adventure, Sci-Fi & Fantasy
5978	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Dramas	
5979	TV-PG	89 min		Dramas, Romantic Movies
5980	TV-14	105 min		Dramas, Romantic Movies, Sports Movies
5981	PG-13	109 min		Action & Adventure
5982	TV-MA	93 min		Documentaries, International Movies
5983	TV-14	75 min		Documentaries
5984	TV-14	1 Season	British TV Shows, International TV Shows, Real...	
5985	TV-MA	90 min		Dramas, Independent Movies, Sports Movies
5986	TV-14	1 Season	International TV Shows, Korean TV Shows, Roman...	
5987	NR	90 min		Horror Movies, Independent Movies, Thrillers
5988	R	93 min		Horror Movies, Thrillers
5989	NaN	37 min		Movies
5990	R	113 min		Action & Adventure, International Movies
5991	TV-MA	89 min		Horror Movies, Thrillers
5992	PG-13	102 min		Action & Adventure
5993	PG-13	102 min		Comedies
5994	TV-MA	90 min		Horror Movies, Independent Movies, Thrillers
5995	TV-MA	94 min		Horror Movies, Independent Movies
5996	PG-13	91 min		Documentaries, Music & Musicals
5997	TV-PG	107 min		International Movies, Music & Musicals
5998	TV-14	95 min		Sci-Fi & Fantasy
5999	R	119 min		Dramas, Independent Movies

[23]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 10 columns):
#   Column          Non-Null Count  Dtype
---  -
0   show_id         8807 non-null   object
1   type            8807 non-null   object
2   director        8807 non-null   object
3   cast            8807 non-null   object
4   country         7976 non-null   object
5   date_added      8797 non-null   object
6   release_year    8807 non-null   int64
7   rating          8803 non-null   object
8   duration        8804 non-null   object
9   listed_in       8807 non-null   object
dtypes: int64(1), object(9)
memory usage: 688.2+ KB
```

```
[ ]:
```

```
[24]: df[df['date_added'].isna()]
```

```
[24]:
```

	show_id	type	director \
--	---------	------	------------

6066	s6067	TV Show	No Director
6174	s6175	TV Show	No Director
6795	s6796	TV Show	No Director
6806	s6807	TV Show	No Director
6901	s6902	TV Show	No Director
7196	s7197	TV Show	No Director
7254	s7255	TV Show	No Director
7406	s7407	TV Show	No Director
7847	s7848	TV Show	No Director
8182	s8183	TV Show	No Director

		cast	country \
6066	[Daniel Radcliffe, Jon Hamm, Adam Godley, Chri...		United Kingdom
6174	[Anthony Bourdain]		United States
6795	[Kelsey Grammer, Jane Leeves, David Hyde Pierc...		United States
6806	[Jennifer Aniston, Courteney Cox, Lisa Kudrow,...		United States
6901	[Yuuka Nanri, Kanako Mitsuhashi, Eri Sendai, A...		Japan
7196	[Igor Dmitriev]		NaN
7254	[Eugenio Derbez, Consuelo Duval, Luis Manuel Á...		United States
7406	[Marc Maron, Judd Hirsch, Josh Brener, Nora Ze...		United States
7847	[Burnie Burns, Jason Saldaña, Gustavo Sorola, ...		United States
8182	[Luke Jurevicius, Craig Behenna, Charlotte Ham...		Australia

	date_added	release_year	rating	duration \
6066	NaN	2013	TV-MA	2 Seasons
6174	NaN	2018	TV-PG	5 Seasons
6795	NaN	2003	TV-PG	11 Seasons
6806	NaN	2003	TV-14	10 Seasons
6901	NaN	2008	TV-14	2 Seasons
7196	NaN	2010	TV-Y	2 Seasons
7254	NaN	2012	TV-14	3 Seasons
7406	NaN	2016	TV-MA	4 Seasons
7847	NaN	2015	NR	13 Seasons
8182	NaN	2015	TV-Y7	2 Seasons

	listed_in
6066	British TV Shows, TV Comedies, TV Dramas
6174	Docuseries
6795	Classic & Cult TV, TV Comedies
6806	Classic & Cult TV, TV Comedies
6901	Anime Series, Crime TV Shows
7196	Kids' TV

```

7254 International TV Shows, Spanish-Language TV Sh...
7406 TV Comedies
7847 TV Action & Adventure, TV Comedies, TV Sci-Fi ...
8182 Kids' TV, TV Comedies

```

```
[ ]:
```

```
[25]: df[df['release_year']==2015].tail(31)
```

```
[25]:
```

	show_id	type	director \
8548	s8549	Movie	Mark Neveldine
8550	s8551	Movie	Ossama Abu El Atta, Tarek Al Eryan
8558	s8559	Movie	Robert Eggers
8575	s8576	TV Show	No Director
8577	s8578	Movie	Wim Bonte
8578	s8579	Movie	Raam Reddy
8599	s8600	TV Show	Michael Cumming
8610	s8611	Movie	Daniel Benmayer
8620	s8621	Movie	Jon Reiner, Brad Rothschild
8625	s8626	Movie	Don Michael Paul
8637	s8638	Movie	Jay Roach
8647	s8648	Movie	No Director
8650	s8651	Movie	Susan Lambert, Stefan Moore
8653	s8654	TV Show	No Director
8675	s8676	Movie	Brandon Dickerson
8688	s8689	Movie	Anu Menon
8699	s8700	Movie	Jack Zagha Kababie
8703	s8704	Movie	Max Joseph
8708	s8709	Movie	Alejandro Agresti
8713	s8714	Movie	Ashish R. Mohan
8731	s8732	Movie	Noah Baumbach
8740	s8741	TV Show	No Director
8759	s8760	Movie	Storm Theunissen
8761	s8762	Movie	Yasir Nawaz
8779	s8780	Movie	Kirati Nakintanon
8780	s8781	TV Show	No Director
8784	s8785	Movie	No Director
8788	s8789	Movie	Ivona Juka
8795	s8796	TV Show	No Director
8801	s8802	Movie	Majid Al Ansari
8806	s8807	Movie	Mozez Singh

```
cast \
```

```

8548 [Olivia Taylor Dudley, Michael Peña, Dougray S...
8550 [Mona Zaki, Asser Yassin, Amr Saad, Salwa Moha...
8558 [Anya Taylor-Joy, Ralph Ineson, Kate Dickie, H...
8575 [Jin-mo Joo, Sa-rang Kim, Junior, Ja-in Lee, S...

```

8577 [Others]
 8578 [Thammegowda S., Channegowda, Abhishek H.N., P...
 8599 [Matt Berry, Robert Bathurst, Doon Mackichan, ...
 8610 [Luciano Acuna Jr., Adam Rayner, Marie Avgerop...
 8620 [Others]
 8625 [Michael Gross, Jamie Kennedy, Ernest Ndhlovu,...
 8637 [Bryan Cranston, Diane Lane, Elle Fanning, Joh...
 8647 [Vijay Maurya, Naman Jain, Usha Nadkarni, Muke...
 8650 [Others]
 8653 [Mook Worranit Thawornwong, Puttichai Kasetsin...
 8675 [Patrick Davis, Lisa Vidal, Josh Pence, José Z...
 8688 [Naseeruddin Shah, Kalki Koechlin, Rajat Kapoo...
 8699 [José Carlos Ruiz, Hoze Meléndez]
 8703 [Zac Efron, Wes Bentley, Emily Ratajkowski, Jo...
 8708 [John Cusack, Paul Hipp, Kevin Morris, Alejand...
 8713 [Arshad Warsi, Jacky Bhagnani, Lauren Gottlieb...
 8731 [Ben Stiller, Naomi Watts, Adam Driver, Amanda...
 8740 [Dougray Scott]
 8759 [Charlie Luxton]
 8761 [Javed Sheikh, Danish Taimoor, Janita Asma, Na...
 8779 [Supanart Jittaleela, Pimpakan Bangchawong, Ch...
 8780 [Johnny Yong Bosch, J.W. Terry, Alicyn Packard...
 8784 [Eileen Stevens, Alyson Leigh Rosenfeld, Sarah...
 8788 [Lana Baric, Vojislav Brajovic, Natasa Janjic,...
 8795 [Mike Liscio, Emily Bauer, Billy Bob Thompson,...
 8801 [Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri,...
 8806 [Vicky Kaushal, Sarah-Jane Dias, Raaghav Chana...

	country	date_added	release_year \
8548	United States	April 20, 2020	2015
8550	Egypt	April 18, 2019	2015
8558	United Kingdom, Canada, United States	September 17, 2018	2015
8575	South Korea	May 22, 2017	2015
8577	Belgium	December 15, 2015	2015
8578	India, United States	March 9, 2017	2015
8599	United Kingdom	September 1, 2017	2015
8610	United States	October 22, 2019	2015
8620	Canada	November 1, 2016	2015
8625	South Africa, United States	January 1, 2020	2015
8637	United States	May 19, 2020	2015
8647	India	January 15, 2017	2015
8650	Australia	December 1, 2015	2015
8653	Thailand	February 15, 2019	2015
8675	United States	July 24, 2017	2015
8688	India	May 1, 2018	2015
8699	Mexico	November 1, 2017	2015
8703	United Kingdom, France, United States	January 27, 2020	2015

8708	United States, Argentina	August 15, 2017	2015
8713	India	March 1, 2018	2015
8731	United States	October 23, 2017	2015
8740	United Kingdom	March 31, 2017	2015
8759	NaN	February 1, 2019	2015
8761	Pakistan	July 1, 2017	2015
8779	Thailand	November 8, 2018	2015
8780	United States	April 1, 2016	2015
8784	Russia, Spain	June 23, 2018	2015
8788	Croatia, Slovenia, Serbia, Montenegro	July 1, 2016	2015
8795	Japan, Canada	May 1, 2018	2015
8801	United Arab Emirates, Jordan	March 9, 2016	2015
8806	India	March 2, 2019	2015

	rating	duration	listed_in
8548	PG-13	91 min	Horror Movies, Thrillers
8550	TV-14	98 min	Action & Adventure, Dramas, International Movies
8558	R	92 min	Horror Movies, Independent Movies
8575	TV-14	1 Season	International TV Shows, Korean TV Shows, Roman...
8577	NR	77 min	Documentaries, International Movies, Music & M...
8578	TV-MA	123 min	Dramas, Independent Movies, International Movies
8599	TV-MA	3 Seasons	British TV Shows, Classic & Cult TV, TV Comedies
8610	PG-13	94 min	Action & Adventure
8620	TV-PG	82 min	Documentaries
8625	PG-13	99 min	Comedies, Horror Movies, Sci-Fi & Fantasy
8637	R	124 min	Dramas, Independent Movies
8647	TV-14	89 min	Dramas, International Movies
8650	TV-14	78 min	Documentaries, International Movies
8653	TV-14	4 Seasons	International TV Shows, Romantic TV Shows, TV ...
8675	PG-13	110 min	Dramas
8688	TV-MA	94 min	Dramas, International Movies
8699	TV-14	92 min	Comedies, Dramas, International Movies
8703	R	96 min	Dramas, Independent Movies, Music & Musicals
8708	TV-MA	94 min	Comedies, Dramas, Independent Movies
8713	TV-MA	124 min	Comedies, International Movies
8731	R	97 min	Comedies, Dramas, Independent Movies
8740	TV-PG	1 Season	British TV Shows, Docuseries, Science & Nature TV
8759	TV-PG	49 min	Movies
8761	TV-14	133 min	Comedies, International Movies, Romantic Movies
8779	TV-PG	107 min	International Movies, LGBTQ Movies, Romantic M...
8780	TV-Y7	1 Season	Anime Series, Kids' TV
8784	TV-Y	78 min	Children & Family Movies
8788	TV-MA	157 min	Dramas, International Movies
8795	TV-Y7	2 Seasons	Anime Series, Kids' TV
8801	TV-MA	96 min	Dramas, International Movies, Thrillers
8806	TV-14	111 min	Dramas, International Movies, Music & Musicals

```
[ ]:
```

```
[ ]:
```

```
[26]: df['date_added'].fillna('September 24, 2021',inplace=True)
```

```
[27]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 10 columns):
#   Column          Non-Null Count  Dtype
---  -
0   show_id         8807 non-null   object
1   type            8807 non-null   object
2   director        8807 non-null   object
3   cast            8807 non-null   object
4   country         7976 non-null   object
5   date_added      8807 non-null   object
6   release_year    8807 non-null   int64
7   rating          8803 non-null   object
8   duration        8804 non-null   object
9   listed_in      8807 non-null   object
dtypes: int64(1), object(9)
memory usage: 688.2+ KB
```

```
[ ]:
```

```
[ ]:
```

```
[28]: df[df['rating'].isna()]
```

```
[28]:
```

	show_id	type	director \	cast	country \	date_added	release_year	rating	duration \
5989	s5990	Movie	No Director			January 26, 2017	2017	NaN	37 min
6827	s6828	TV Show	No Director			December 1, 2016	2013	NaN	1 Season
7312	s7313	TV Show	No Director						
7537	s7538	Movie	Alessandro Pepe						

	show_id	type	director \	cast	country \
5989				[Oprah Winfrey, Ava DuVernay]	NaN
6827				[Kaito Ishikawa, Hisako Kanemoto, Ai Kayano, K...	Japan
7312				[Flynn Curry, Olivia Deeble, Madison Lu, Oisín...	Australia
7537				[Leone Frisa, Paolo Vaccarino, Francesco Migli...	Italy

	show_id	type	director \	cast	country \	date_added	release_year	rating	duration \
5989	s5990	Movie	No Director			January 26, 2017	2017	NaN	37 min
6827	s6828	TV Show	No Director			December 1, 2016	2013	NaN	1 Season

7312	February 1, 2018	2015	NaN	1 Season
7537	March 1, 2017	2015	NaN	115 min

	listed_in
5989	Movies
6827	Anime Series, International TV Shows
7312	Kids' TV, TV Comedies
7537	Dramas

```
[29]: df[df['listed_in'].str.contains("Kids' TV")]['rating'].value_counts()
```

```
[29]: TV-Y7      189
TV-Y       176
TV-G       43
TV-PG      41
TV-Y7-FV   1
Name: rating, dtype: int64
```

```
[30]: df.loc[7312, 'rating'] = 'TV-Y7'
```

```
[31]: df[df['rating'].isna()]
```

```
[31]:   show_id  type  director \
5989  s5990  Movie  No Director
6827  s6828  TV Show  No Director
7537  s7538  Movie  Alessandro Pepe
```

		cast	country	\
5989		[Oprah Winfrey, Ava DuVernay]	NaN	
6827		[Kaito Ishikawa, Hisako Kanemoto, Ai Kayano, K...	Japan	
7537		[Leone Frisa, Paolo Vaccarino, Francesco Migli...	Italy	

	date_added	release_year	rating	duration	\
5989	January 26, 2017	2017	NaN	37 min	
6827	December 1, 2016	2013	NaN	1 Season	
7537	March 1, 2017	2015	NaN	115 min	

	listed_in
5989	Movies
6827	Anime Series, International TV Shows
7537	Dramas

```
[ ]:
```

```
[32]: df[df['listed_in'].str.contains("Movies")]['rating'].value_counts()
```

```
[32]: TV-MA      1459
      TV-14     1182
      R         451
      TV-PG     424
      PG        253
      PG-13     252
      TV-Y7     139
      TV-Y      131
      TV-G      104
      NR        54
      G         39
      TV-Y7-FV   5
      NC-17      3
      UR         2
      74 min     1
      84 min     1
      66 min     1
      Name: rating, dtype: int64
```

```
[ ]:
```

```
[33]: df.loc[5989, 'rating'] = 'TV-MA'
```

```
[34]: df[df['rating'].isna()]
```

```
[34]:      show_id      type      director \
6827    s6828  TV Show      No Director
7537    s7538   Movie  Alessandro Pepe

                                     cast country \
6827  [Kaito Ishikawa, Hisako Kanemoto, Ai Kayano, K...  Japan
7537  [Leone Frisa, Paolo Vaccarino, Francesco Migli...  Italy

      date_added  release_year  rating  duration \
6827  December 1, 2016         2013   NaN   1 Season
7537    March 1, 2017         2015   NaN   115 min

                                     listed_in
6827  Anime Series, International TV Shows
7537                                     Dramas
```

```
[35]: df[df['listed_in'].str.contains("Dramas")]['rating'].value_counts()
```

```
[35]: TV-MA      1264
      TV-14     962
      R        376
      TV-PG     249
```

```

PG-13      192
PG          69
TV-G       36
NR          28
G           6
TV-Y        3
UR          2
NC-17       1
TV-Y7       1
Name: rating, dtype: int64

```

```
[36]: df.loc[7537, 'rating'] = 'TV-MA'
```

```
[37]: df[df['rating'].isna()]
```

```

[37]:      show_id      type      director \
6827    s6828  TV Show  No Director

                                         cast country \
6827  [Kaito Ishikawa, Hisako Kanemoto, Ai Kayano, K...  Japan

      date_added  release_year  rating  duration \
6827  December 1, 2016          2013    NaN    1 Season

                                         listed_in
6827  Anime Series, International TV Shows

```

```
[38]: df[df['listed_in'].str.contains("Anime Series, International TV_
↳Shows")]['rating'].value_counts()
```

```

[38]: TV-14      57
      TV-MA      46
      TV-PG       9
      TV-Y7       2
      Name: rating, dtype: int64

```

```
[ ]:
```

```
[39]: df.loc[6827, 'rating'] = 'TV-14'
```

```
[ ]:
```

```
[40]: df['rating'].isna().isna().sum()
```

```
[40]: 0
```

```
[41]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 10 columns):
#   Column          Non-Null Count  Dtype
---  -
0   show_id         8807 non-null   object
1   type            8807 non-null   object
2   director        8807 non-null   object
3   cast            8807 non-null   object
4   country         7976 non-null   object
5   date_added      8807 non-null   object
6   release_year    8807 non-null   int64
7   rating          8807 non-null   object
8   duration        8804 non-null   object
9   listed_in      8807 non-null   object
dtypes: int64(1), object(9)
memory usage: 688.2+ KB

```

```
[42]: df[df['duration'].isna()]
```

```

[42]:      show_id  type  director      cast      country \
5541    s5542  Movie  Louis C.K.  [Louis C.K.]  United States
5794    s5795  Movie  Louis C.K.  [Louis C.K.]  United States
5813    s5814  Movie  Louis C.K.  [Louis C.K.]  United States

      date_added  release_year  rating  duration  listed_in
5541    April 4, 2017         2017  74 min      NaN      Movies
5794  September 16, 2016         2010  84 min      NaN      Movies
5813    August 15, 2016         2015  66 min      NaN      Movies

```

```
[43]: df[df['duration'].isna()]['rating']
```

```

[43]: 5541    74 min
      5794    84 min
      5813    66 min
      Name: rating, dtype: object

```

```
[44]: df.loc[df['duration'].isna(), 'duration'] = df[df['duration'].isna()]['rating']
```

```
[ ]:
```

```
[45]: df[df['duration'].isna()]
```

```

[45]: Empty DataFrame
Columns: [show_id, type, director, cast, country, date_added, release_year,
rating, duration, listed_in]
Index: []

```

```
[ ]:
```

```
[46]: df.loc[df['rating'].str.contains('min'),'rating']='TV-MA'
```

```
[ ]:
```

```
[47]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 10 columns):
#   Column          Non-Null Count  Dtype
---  -
0   show_id         8807 non-null   object
1   type            8807 non-null   object
2   director        8807 non-null   object
3   cast            8807 non-null   object
4   country         7976 non-null   object
5   date_added      8807 non-null   object
6   release_year    8807 non-null   int64
7   rating          8807 non-null   object
8   duration        8807 non-null   object
9   listed_in      8807 non-null   object
dtypes: int64(1), object(9)
memory usage: 688.2+ KB
```

```
[ ]:
```

```
[48]: df.loc[df['country'].isna(),'country']='Country Not Mentioned'
```

```
[49]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 10 columns):
#   Column          Non-Null Count  Dtype
---  -
0   show_id         8807 non-null   object
1   type            8807 non-null   object
2   director        8807 non-null   object
3   cast            8807 non-null   object
4   country         8807 non-null   object
5   date_added      8807 non-null   object
6   release_year    8807 non-null   int64
7   rating          8807 non-null   object
8   duration        8807 non-null   object
9   listed_in      8807 non-null   object
```

```
dtypes: int64(1), object(9)
memory usage: 688.2+ KB
```

```
[ ]:
```

```
[50]: df.head()
```

```
[50]: show_id      type      director \
0      s1      Movie  Kirsten Johnson
1      s2  TV Show      No Director
2      s3  TV Show  Julien Leclercq
3      s4  TV Show      No Director
4      s5  TV Show      No Director

                                cast      country \
0                                [Others]      United States
1  [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba...      South Africa
2  [Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nab...  Country Not Mentioned
3                                [Others]  Country Not Mentioned
4  [Mayur More, Jitendra Kumar, Ranjan Raj, Alam ...      India

      date_added  release_year  rating  duration \
0  September 25, 2021      2020  PG-13      90 min
1  September 24, 2021      2021  TV-MA  2 Seasons
2  September 24, 2021      2021  TV-MA  1 Season
3  September 24, 2021      2021  TV-MA  1 Season
4  September 24, 2021      2021  TV-MA  2 Seasons

                                listed_in
0                                Documentaries
1  International TV Shows, TV Dramas, TV Mysteries
2  Crime TV Shows, International TV Shows, TV Act...
3                                Docuseries, Reality TV
4  International TV Shows, Romantic TV Shows, TV ...
```

```
[ ]:
```

```
[51]: df['listed_in']=df['listed_in'].apply(lambda x : x.split(', '))
```

```
[52]: df['listed_in'][2]
```

```
[52]: ['Crime TV Shows', 'International TV Shows', 'TV Action & Adventure']
```

```
[53]: df.head()
```

```
[53]: show_id      type      director \
0      s1      Movie  Kirsten Johnson
```



```

1      s2  TV Show      No Director
2      s3  TV Show  Julien Leclercq
3      s4  TV Show      No Director
4      s5  TV Show      No Director

```

```

                                cast                country \
0                                [Others]                United States
1  [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba...      South Africa
2  [Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nab...  Country Not Mentioned
3                                [Others]  Country Not Mentioned
4  [Mayur More, Jitendra Kumar, Ranjan Raj, Alam ...      India

```

```

      date_added  release_year  rating  duration \
0  September 25, 2021          2020  PG-13    90 min
1  September 24, 2021          2021  TV-MA    2 Seasons
2  September 24, 2021          2021  TV-MA    1 Season
3  September 24, 2021          2021  TV-MA    1 Season
4  September 24, 2021          2021  TV-MA    2 Seasons

```

```

                                listed_in
0                                [Documentaries]
1  [International TV Shows, TV Dramas, TV Mysteries]
2  [Crime TV Shows, International TV Shows, TV Ac...
3                                [Docuseries, Reality TV]
4  [International TV Shows, Romantic TV Shows, TV...

```

```
[ ]:
```

```
[54]: df[df['director'].apply(lambda x : True if ',' in x else False)]
```

```

[54]:      show_id  type                                director \
6         s7  Movie      Robert Cullen, José Luis Ucha
16        s17  Movie  Pedro de Echave García, Pablo Azorín Williams
23        s24  Movie      Alex Woo, Stanley Moore
30        s31  Movie  Ashwiny Iyer Tiwari, Abhishek Chaubey, Saket C...
68        s69  Movie  Hanns-Bruno Kammertöns, Vanessa Nöcker, Michae...
...      ...      ...
8727    s8728  Movie      Ritu Sarin, Tenzing Sonam
8728    s8729  Movie      Heidi Brandenburg, Mathew Orzel
8737    s8738  Movie      Milla Harrison-Hansley, Alicky Sussman
8739    s8740  Movie      Frank Capra, Anatole Litvak
8765    s8766  Movie  Jovanka Vuckovic, Annie Clark, Roxanne Benjami...

```

```

                                cast \
6  [Vanessa Hudgens, Kimiko Glenn, James Marsden,...
16                                [Others]
23  [Maisie Benson, Paul Killam, Kerry Gudjohnsen,...

```

```

30      [Abhishek Banerjee, Rinku Rajguru, Delzad Hiwa...
68                                     [Michael Schumacher]
...
8727                                     [Others]
8728                                     [Others]
8737                                     [Giles Yeo]
8739                                     [Others]
8765      [Natalie Brown, Jonathan Watton, Melanie Lynsk...

```

	country	date_added	release_year \
6	Country Not Mentioned	September 24, 2021	2021
16	Country Not Mentioned	September 22, 2021	2020
23	Country Not Mentioned	September 21, 2021	2021
30	Country Not Mentioned	September 17, 2021	2021
68	Country Not Mentioned	September 15, 2021	2021
...
8727	United Kingdom, India, United States	December 25, 2016	2013
8728	Peru, United States, United Kingdom	November 30, 2016	2016
8737	United Kingdom	February 1, 2019	2016
8739	United States	March 31, 2017	1943
8765	Canada, United States	June 22, 2017	2017

	rating	duration	listed_in
6	PG	91 min	[Children & Family Movies]
16	TV-MA	67 min	[Documentaries, International Movies]
23	TV-Y	61 min	[Children & Family Movies]
30	TV-14	111 min	[Dramas, Independent Movies, International Mov...]
68	TV-14	113 min	[Documentaries, International Movies, Sports M...]
...
8727	NR	75 min	[Documentaries, International Movies]
8728	TV-14	103 min	[Documentaries, International Movies]
8737	TV-14	50 min	[Documentaries]
8739	TV-PG	82 min	[Documentaries]
8765	R	81 min	[Horror Movies]

[614 rows x 10 columns]

```
[ ]:
```

```
[55]: df['director']=df['director'].apply(lambda x : x.split(', '))
```

```
[56]: df[df['director'].apply(lambda x : True if ',' in x else False)]
```

```
[56]: Empty DataFrame
Columns: [show_id, type, director, cast, country, date_added, release_year,
rating, duration, listed_in]
Index: []
```

```
[ ]:
```

```
[57]: df.head()
```

```
[57]: show_id      type      director \
0      s1      Movie  [Kirsten Johnson]
1      s2  TV Show      [No Director]
2      s3  TV Show  [Julien Leclercq]
3      s4  TV Show      [No Director]
4      s5  TV Show      [No Director]

                                cast      country \
0                                [Others]      United States
1  [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba...      South Africa
2  [Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nab...  Country Not Mentioned
3                                [Others]  Country Not Mentioned
4  [Mayur More, Jitendra Kumar, Ranjan Raj, Alam ...      India

      date_added  release_year  rating  duration \
0  September 25, 2021      2020  PG-13      90 min
1  September 24, 2021      2021  TV-MA  2 Seasons
2  September 24, 2021      2021  TV-MA  1 Season
3  September 24, 2021      2021  TV-MA  1 Season
4  September 24, 2021      2021  TV-MA  2 Seasons

                                listed_in
0                                [Documentaries]
1  [International TV Shows, TV Dramas, TV Mysteries]
2  [Crime TV Shows, International TV Shows, TV Ac...
3                                [Docuseries, Reality TV]
4  [International TV Shows, Romantic TV Shows, TV...
```

```
[ ]:
```

```
[58]: def function(x):
      length=int(x.split())[0])
      if length<=30:
          return "0-30"
      elif length<=60:
          return "31-60"
      elif length<=90:
          return "61-90"
      elif length<=120:
          return "91-120"
      elif length<=150:
          return "121-150"
      elif length<=180:
```

```

        return "151-180"
    else:
        return "Above 180"

```

[]:

```

[59]: df.loc[df['type']=='Movie', 'duration'] = df.loc[df['type']=='Movie', 'duration'].
      ↪ apply(function)

```

[]:

```

[60]: df.loc[df['type']=='Movie', 'duration'].value_counts()

```

```

[60]: 91-120      2996
      61-90      1506
      121-150    897
      31-60      357
      151-180    198
      0-30       130
      Above 180    47
      Name: duration, dtype: int64

```

[]:

[]:

[]:

[]:

[]:

2 Un-Nest

```

[61]: df.info()

```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 10 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   show_id     8807 non-null   object
 1   type        8807 non-null   object
 2   director    8807 non-null   object
 3   cast        8807 non-null   object
 4   country     8807 non-null   object

```

```

5   date_added      8807 non-null   object
6   release_year    8807 non-null   int64
7   rating          8807 non-null   object
8   duration        8807 non-null   object
9   listed_in       8807 non-null   object
dtypes: int64(1), object(9)
memory usage: 688.2+ KB

```

```
[62]: df.head()
```

```

[62]:   show_id      type      director \
0      s1      Movie  [Kirsten Johnson]
1      s2  TV Show    [No Director]
2      s3  TV Show  [Julien Leclercq]
3      s4  TV Show    [No Director]
4      s5  TV Show    [No Director]

                                cast      country \
0                                [Others]      United States
1  [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba...      South Africa
2  [Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nab...  Country Not Mentioned
3                                [Others]  Country Not Mentioned
4  [Mayur More, Jitendra Kumar, Ranjan Raj, Alam ...      India

      date_added  release_year  rating  duration \
0  September 25, 2021      2020  PG-13      61-90
1  September 24, 2021      2021  TV-MA  2 Seasons
2  September 24, 2021      2021  TV-MA   1 Season
3  September 24, 2021      2021  TV-MA   1 Season
4  September 24, 2021      2021  TV-MA  2 Seasons

                                listed_in
0                                [Documentaries]
1  [International TV Shows, TV Dramas, TV Mysteries]
2  [Crime TV Shows, International TV Shows, TV Ac...
3                                [Docuseries, Reality TV]
4  [International TV Shows, Romantic TV Shows, TV...

```

```
[63]: df.shape
```

```
[63]: (8807, 10)
```

```
[ ]:
```

```

[64]: constraint=df['cast'].tolist()
constraint
new_df = pd.DataFrame(constraint,index=df['show_id'])

```

```
new_df.head()
cast_df = pd.DataFrame(new_df.stack(),columns=['Cast']).reset_index().iloc[
    ↪, [0,2]]
cast_df
```

```
[64]:      show_id      Cast
0         s1      Others
1         s2    Ama Qamata
2         s2    Khosi Ngema
3         s2    Gail Mabalane
4         s2    Thabang Molaba
...      ...      ...
64946    s8807    Manish Chaudhary
64947    s8807    Meghna Malik
64948    s8807    Malkeet Rauni
64949    s8807    Anita Shabdish
64950    s8807    Chittaranjan Tripathy
```

[64951 rows x 2 columns]

```
[ ]:
```

```
[ ]:
```

```
[65]: other_df=df[['show_id','type','country','date_added','release_year','rating','duration']]
other_df
```

```
[65]:      show_id      type      country      date_added \
0         s1    Movie    United States    September 25, 2021
1         s2  TV Show    South Africa    September 24, 2021
2         s3  TV Show  Country Not Mentioned    September 24, 2021
3         s4  TV Show  Country Not Mentioned    September 24, 2021
4         s5  TV Show          India    September 24, 2021
...      ...      ...      ...      ...
8802    s8803    Movie    United States    November 20, 2019
8803    s8804  TV Show  Country Not Mentioned    July 1, 2019
8804    s8805    Movie    United States    November 1, 2019
8805    s8806    Movie    United States    January 11, 2020
8806    s8807    Movie          India    March 2, 2019
```

```
      release_year rating  duration
0              2020  PG-13    61-90
1              2021  TV-MA    2 Seasons
2              2021  TV-MA    1 Season
3              2021  TV-MA    1 Season
4              2021  TV-MA    2 Seasons
...              ...      ...      ...
```

8802	2007	R	151-180
8803	2018	TV-Y7	2 Seasons
8804	2009	R	61-90
8805	2006	PG	61-90
8806	2015	TV-14	91-120

[8807 rows x 7 columns]

```
[66]: df1=pd.merge(left=cast_df,right=other_df,on='show_id')
df1
```

```
[66]:
```

	show_id	Cast	type	country \
0	s1	Others	Movie	United States
1	s2	Ama Qamata	TV Show	South Africa
2	s2	Khosi Ngema	TV Show	South Africa
3	s2	Gail Mabalane	TV Show	South Africa
4	s2	Thabang Molaba	TV Show	South Africa
...
64946	s8807	Manish Chaudhary	Movie	India
64947	s8807	Meghna Malik	Movie	India
64948	s8807	Malkeet Rauni	Movie	India
64949	s8807	Anita Shabdish	Movie	India
64950	s8807	Chittaranjan Tripathy	Movie	India

	date_added	release_year	rating	duration
0	September 25, 2021	2020	PG-13	61-90
1	September 24, 2021	2021	TV-MA	2 Seasons
2	September 24, 2021	2021	TV-MA	2 Seasons
3	September 24, 2021	2021	TV-MA	2 Seasons
4	September 24, 2021	2021	TV-MA	2 Seasons
...
64946	March 2, 2019	2015	TV-14	91-120
64947	March 2, 2019	2015	TV-14	91-120
64948	March 2, 2019	2015	TV-14	91-120
64949	March 2, 2019	2015	TV-14	91-120
64950	March 2, 2019	2015	TV-14	91-120

[64951 rows x 8 columns]

```
[ ]:
```

```
[67]: constraint=df['director'].tolist()
constraint
new_df = pd.DataFrame(constraint,index=df['show_id'])
new_df.head()
director_df = pd.DataFrame(new_df.stack(),columns=['Director']).reset_index().
→iloc[:,[0,2]]
```

```
director_df
```

```
[67]:
```

	show_id	Director
0	s1	Kirsten Johnson
1	s2	No Director
2	s3	Julien Leclercq
3	s4	No Director
4	s5	No Director
...
9607	s8803	David Fincher
9608	s8804	No Director
9609	s8805	Ruben Fleischer
9610	s8806	Peter Hewitt
9611	s8807	Mozez Singh

```
[9612 rows x 2 columns]
```

```
[ ]:
```

```
[68]: df1=pd.merge(left=director_df,right=df1,on='show_id')
df1
```

```
[68]:
```

	show_id	Director	Cast	type	country \
0	s1	Kirsten Johnson	Others	Movie	United States
1	s2	No Director	Ama Qamata	TV Show	South Africa
2	s2	No Director	Khosi Ngema	TV Show	South Africa
3	s2	No Director	Gail Mabalane	TV Show	South Africa
4	s2	No Director	Thabang Molaba	TV Show	South Africa
...
70807	s8807	Mozez Singh	Manish Chaudhary	Movie	India
70808	s8807	Mozez Singh	Meghna Malik	Movie	India
70809	s8807	Mozez Singh	Malkeet Rauni	Movie	India
70810	s8807	Mozez Singh	Anita Shabdish	Movie	India
70811	s8807	Mozez Singh	Chittaranjan Tripathy	Movie	India

	date_added	release_year	rating	duration
0	September 25, 2021	2020	PG-13	61-90
1	September 24, 2021	2021	TV-MA	2 Seasons
2	September 24, 2021	2021	TV-MA	2 Seasons
3	September 24, 2021	2021	TV-MA	2 Seasons
4	September 24, 2021	2021	TV-MA	2 Seasons
...
70807	March 2, 2019	2015	TV-14	91-120
70808	March 2, 2019	2015	TV-14	91-120
70809	March 2, 2019	2015	TV-14	91-120
70810	March 2, 2019	2015	TV-14	91-120
70811	March 2, 2019	2015	TV-14	91-120

[70812 rows x 9 columns]

```
[69]: df1.head()
```

```
[69]:   show_id      Director      Cast      type      country \
0      s1  Kirsten Johnson    Others    Movie  United States
1      s2    No Director    Ama Qamata  TV Show   South Africa
2      s2    No Director    Khosi Ngema  TV Show   South Africa
3      s2    No Director    Gail Mabalane  TV Show   South Africa
4      s2    No Director    Thabang Molaba  TV Show   South Africa
```

```
      date_added  release_year  rating  duration
0  September 25, 2021        2020  PG-13      61-90
1  September 24, 2021        2021  TV-MA    2 Seasons
2  September 24, 2021        2021  TV-MA    2 Seasons
3  September 24, 2021        2021  TV-MA    2 Seasons
4  September 24, 2021        2021  TV-MA    2 Seasons
```

```
[ ]:
```

```
[ ]:
```

```
[70]: constraint=df['listed_in'].tolist()
constraint
new_df = pd.DataFrame(constraint,index=df['show_id'])
new_df.head()
listed_in_df = pd.DataFrame(new_df.stack(),columns=['listed_in']).reset_index().
    ↪iloc[:,[0,2]]
listed_in_df
```

```
[70]:   show_id      listed_in
0      s1      Documentaries
1      s2  International TV Shows
2      s2      TV Dramas
3      s2      TV Mysteries
4      s3      Crime TV Shows
...
19318  s8806  Children & Family Movies
19319  s8806      Comedies
19320  s8807      Dramas
19321  s8807  International Movies
19322  s8807      Music & Musicals
```

[19323 rows x 2 columns]

```
[71]: final_df=pd.merge(left=df1,right=listed_in_df,on='show_id')
final_df
```

```
[71]:
```

	show_id	Director	Cast	type	\
0	s1	Kirsten Johnson	Others	Movie	
1	s2	No Director	Ama Qamata	TV Show	
2	s2	No Director	Ama Qamata	TV Show	
3	s2	No Director	Ama Qamata	TV Show	
4	s2	No Director	Khosi Ngema	TV Show	
...	
161211	s8807	Mozez Singh	Anita Shabdish	Movie	
161212	s8807	Mozez Singh	Anita Shabdish	Movie	
161213	s8807	Mozez Singh	Chittaranjan Tripathy	Movie	
161214	s8807	Mozez Singh	Chittaranjan Tripathy	Movie	
161215	s8807	Mozez Singh	Chittaranjan Tripathy	Movie	

	country	date_added	release_year	rating	duration	\
0	United States	September 25, 2021	2020	PG-13	61-90	
1	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	
2	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	
3	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	
4	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	
...	
161211	India	March 2, 2019	2015	TV-14	91-120	
161212	India	March 2, 2019	2015	TV-14	91-120	
161213	India	March 2, 2019	2015	TV-14	91-120	
161214	India	March 2, 2019	2015	TV-14	91-120	
161215	India	March 2, 2019	2015	TV-14	91-120	

	listed_in
0	Documentaries
1	International TV Shows
2	TV Dramas
3	TV Mysteries
4	International TV Shows
...	...
161211	International Movies
161212	Music & Musicals
161213	Dramas
161214	International Movies
161215	Music & Musicals

[161216 rows x 10 columns]

```
[ ]:
```

3 Analysis

```
[72]: final_df.head()
```

```
[72]:  show_id      Director      Cast      type      country \
0      s1  Kirsten Johnson      Others      Movie  United States
1      s2    No Director  Ama Qamata  TV Show   South Africa
2      s2    No Director  Ama Qamata  TV Show   South Africa
3      s2    No Director  Ama Qamata  TV Show   South Africa
4      s2    No Director  Khosi Ngema  TV Show   South Africa

      date_added  release_year  rating  duration      listed_in
0  September 25, 2021      2020  PG-13      61-90      Documentaries
1  September 24, 2021      2021  TV-MA  2 Seasons  International TV Shows
2  September 24, 2021      2021  TV-MA  2 Seasons      TV Dramas
3  September 24, 2021      2021  TV-MA  2 Seasons      TV Mysteries
4  September 24, 2021      2021  TV-MA  2 Seasons  International TV Shows
```

```
[73]: final_df.shape
```

```
[73]: (161216, 10)
```

```
[ ]:
```

```
[74]: final_df['Year Added']=final_df['date_added'].apply(lambda x : int(x.
↪split()[2]))
```

```
[75]: final_df['Month Added']=final_df['date_added'].apply(lambda x : x.split()[0])
```

```
[76]: final_df.drop(columns='date_added',inplace=True)
```

```
[ ]:
```

4 Question 1 :

4.0.1 How has the number of movies and TV Shows released per year changed over the last 10 years?

```
[ ]:
```

```
[77]: final_df.head()
```

```
[77]:  show_id      Director      Cast      type      country  release_year \
0      s1  Kirsten Johnson      Others      Movie  United States      2020
1      s2    No Director  Ama Qamata  TV Show   South Africa      2021
2      s2    No Director  Ama Qamata  TV Show   South Africa      2021
```

3	s2	No Director	Ama Qamata	TV Show	South Africa	2021
4	s2	No Director	Khosi Ngema	TV Show	South Africa	2021

	rating	duration	listed_in	Year Added	Month Added
0	PG-13	61-90	Documentaries	2021	September
1	TV-MA	2 Seasons	International TV Shows	2021	September
2	TV-MA	2 Seasons	TV Dramas	2021	September
3	TV-MA	2 Seasons	TV Mysteries	2021	September
4	TV-MA	2 Seasons	International TV Shows	2021	September

```
[78]: final_df['release_year'].unique()
```

```
[78]: array([2020, 2021, 1993, 2018, 1996, 1998, 1997, 2010, 2013, 2017, 1975,
        1978, 1983, 1987, 2012, 2001, 2014, 2002, 2003, 2004, 2011, 2008,
        2009, 2007, 2005, 2006, 1994, 2015, 2019, 2016, 1982, 1989, 1990,
        1991, 1999, 1986, 1992, 1984, 1980, 1961, 2000, 1995, 1985, 1976,
        1959, 1988, 1981, 1972, 1964, 1945, 1954, 1979, 1958, 1956, 1963,
        1970, 1973, 1925, 1974, 1960, 1966, 1971, 1962, 1969, 1977, 1967,
        1968, 1965, 1946, 1942, 1955, 1944, 1947, 1943], dtype=int64)
```

```
[79]: release_year_plot=final_df.groupby(by=['type', 'release_year']).
      ↪nunique()['show_id']
      release_year_plot
```

```
[79]: type      release_year
      Movie      1942          2
           1943          3
           1944          3
           1945          3
           1946          1
           ...
      TV Show  2017        265
           2018        380
           2019        397
           2020        436
           2021        315
      Name: show_id, Length: 119, dtype: int64
```

```
[80]: movies_year_wise=release_year_plot.loc['Movie'].sort_values(ascending=False)
      movies_year_wise
```

```
[80]: release_year
      2018      767
      2017      767
      2016      658
      2019      633
      2020      517
```

```

...
1961      1
1963      1
1966      1
1947      1
1946      1
Name: show_id, Length: 73, dtype: int64

```

```
[81]: tvshows_year_wise=release_year_plot.loc['TV Show'].sort_values(ascending=False)
      tvshows_year_wise
```

```
[81]: release_year
2020      436
2019      397
2018      380
2021      315
2017      265
2016      244
2015      162
2014       88
2012       64
2013       63
2011       40
2010       40
2009       34
2008       23
2007       14
2006       14
2005       13
2003       10
2004        9
2002        7
1999        7
2001        5
1997        4
1993        4
1998        4
2000        4
1990        3
1992        3
1996        3
1994        2
1995        2
1988        2
1986        2
1977        1
1963        1

```

```

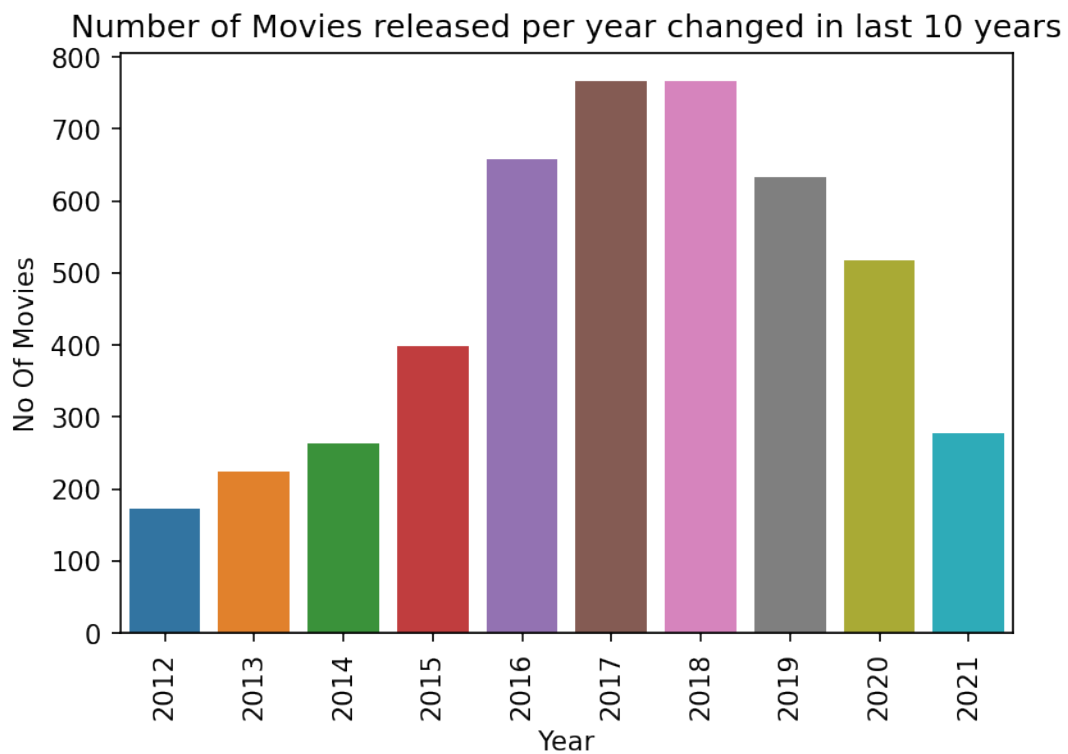
1946      1
1967      1
1972      1
1974      1
1989      1
1979      1
1981      1
1985      1
1991      1
1945      1
1925      1
Name: show_id, dtype: int64

```

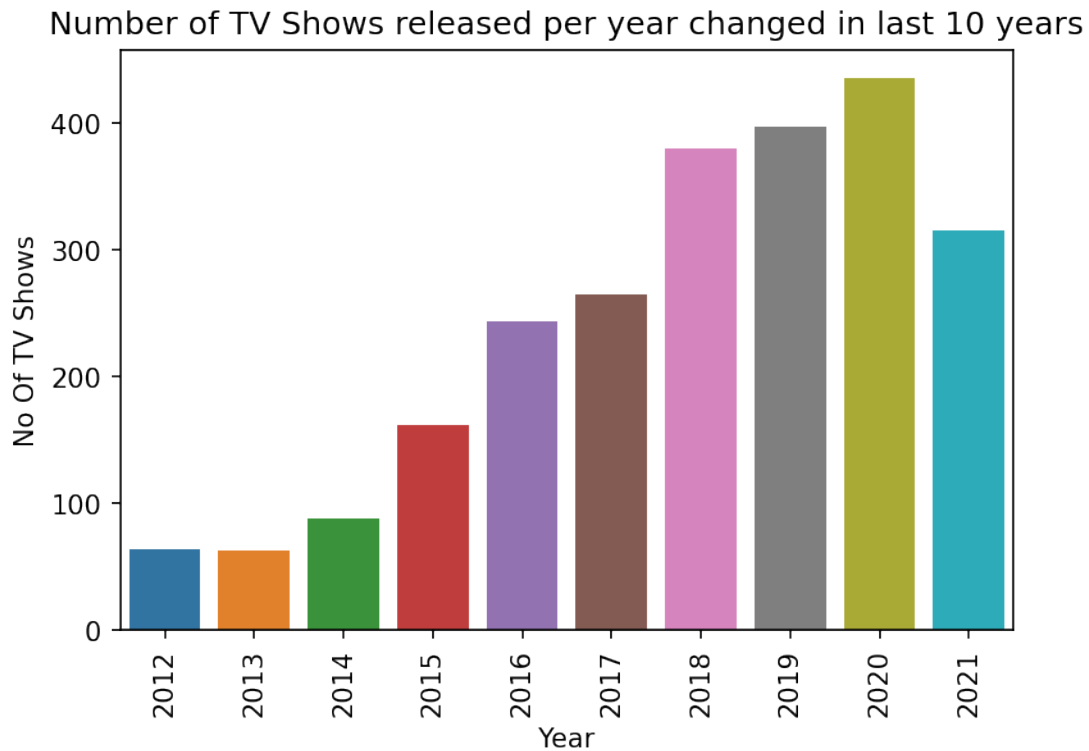
```

[82]: plt.figure(dpi=150)
      sns.barplot(x=movies_year_wise.index[:10],
                  y=movies_year_wise[:10])
      plt.xticks(rotation=90)
      plt.xlabel('Year')
      plt.ylabel('No Of Movies')
      plt.title('Number of Movies released per year changed in last 10 years')
      plt.savefig('Image1.jpg')
      plt.show()

```



```
[83]: plt.figure(dpi=150)
sns.barplot(x=tvshows_year_wise.index[:10],
            y=tvshows_year_wise[:10])
plt.xticks(rotation=90)
plt.xlabel('Year')
plt.ylabel('No Of TV Shows')
plt.title('Number of TV Shows released per year changed in last 10 years')
plt.savefig('Image2.jpg')
plt.show()
```



```
[ ]:
```

```
[ ]:
```

5 Question 2 :

5.0.1 How has the number of Movies and TV Shows added to Netflix per year changed over the last 10 years?

```
[ ]:
```

```
[84]: final_df.head()
```

```
[84]: show_id      Director      Cast      type      country  release_year \
0      s1  Kirsten Johnson      Others      Movie  United States      2020
1      s2      No Director  Ama Qamata  TV Show  South Africa      2021
2      s2      No Director  Ama Qamata  TV Show  South Africa      2021
3      s2      No Director  Ama Qamata  TV Show  South Africa      2021
4      s2      No Director  Khosi Ngema  TV Show  South Africa      2021
```

```

rating  duration      listed_in  Year Added Month Added
0  PG-13      61-90      Documentaries      2021  September
1  TV-MA  2 Seasons  International TV Shows      2021  September
2  TV-MA  2 Seasons      TV Dramas      2021  September
3  TV-MA  2 Seasons      TV Mysteries      2021  September
4  TV-MA  2 Seasons  International TV Shows      2021  September
```

```
[85]: added_year_plot=final_df.groupby(by=['type', 'Year Added']).nunique()['show_id']
added_year_plot
```

```
[85]: type      Year Added
Movie      2008          1
          2009          2
          2010          1
          2011         13
          2012          3
          2013          6
          2014         19
          2015         56
          2016        253
          2017        839
          2018       1237
          2019       1424
          2020       1284
          2021        993
TV Show    2008          1
          2013          5
          2014          5
          2015         26
          2016        176
          2017        349
          2018        412
          2019        592
          2020        595
          2021        515
Name: show_id, dtype: int64
```

```
[86]: movies_added_per_year=added_year_plot.loc['Movie'].sort_values(ascending=False)
movies_added_per_year
```



```
[86]: Year Added
      2019    1424
      2020    1284
      2018    1237
      2021     993
      2017     839
      2016     253
      2015      56
      2014      19
      2011      13
      2013       6
      2012       3
      2009       2
      2008       1
      2010       1
      Name: show_id, dtype: int64
```

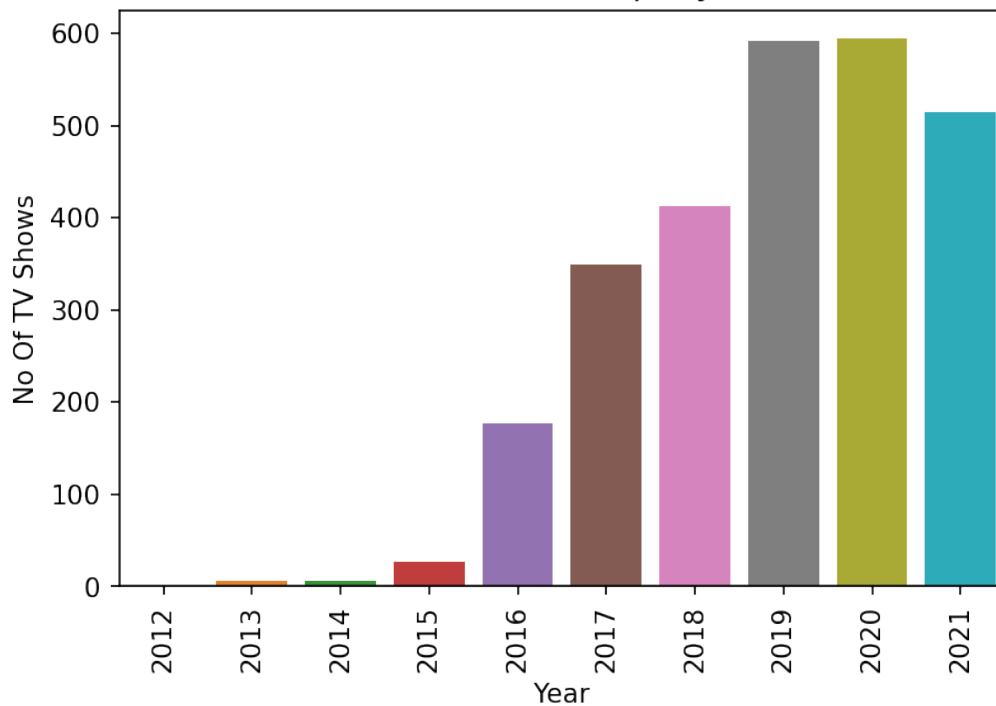
```
[87]: tvshow_added_per_year=added_year_plot.loc['TV Show'].sort_index(ascending=False)
      tvshow_added_per_year
```

```
[87]: Year Added
      2021     515
      2020     595
      2019     592
      2018     412
      2017     349
      2016     176
      2015      26
      2014       5
      2013       5
      2008       1
      Name: show_id, dtype: int64
```

```
[ ]:
```

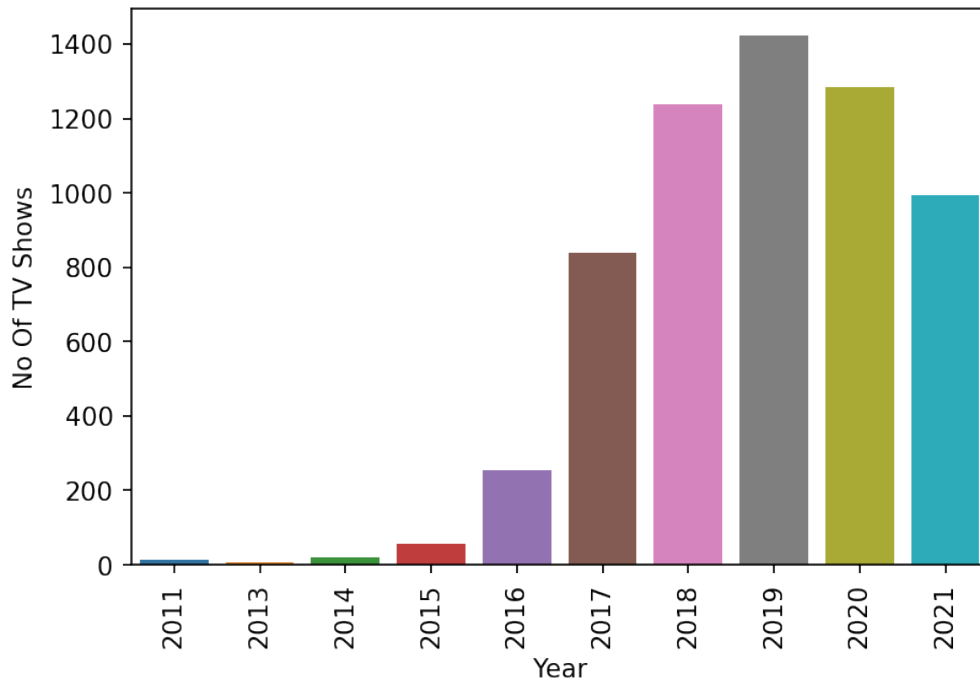
```
[88]: plt.figure(dpi=150)
      sns.barplot(x=list(tvshow_added_per_year.index[:9])+[2012],
                  y=list(tvshow_added_per_year[:9])+[0])
      plt.xticks(rotation=90)
      plt.xlabel('Year')
      plt.ylabel('No Of TV Shows')
      plt.title('Number of TV Shows added to Netflix per year in the last 10 years?')
      plt.savefig('Image3.jpg')
      plt.show()
```

Number of TV Shows added to Netflix per year in the last 10 years?



```
[89]: plt.figure(dpi=150)
sns.barplot(x=movies_added_per_year.index[:10],
            y=movies_added_per_year[:10])
plt.xticks(rotation=90)
plt.xlabel('Year')
plt.ylabel('No Of TV Shows')
plt.title('Number of Movies added to Netflix per year changed in last 10 years?
→')
plt.savefig('Image4.jpg')
plt.show()
```

Number of Movies added to Netflix per year changed in last 10 years?



[]:

[]:

6 Question 3 :

6.0.1 In which months does Netflix add maximum content?

[90]: `final_df.head()`

```
[90]:  show_id      Director      Cast      type      country  release_year \
0      s1  Kirsten Johnson    Others    Movie  United States      2020
1      s2    No Director  Ama Qamata  TV Show  South Africa      2021
2      s2    No Director  Ama Qamata  TV Show  South Africa      2021
3      s2    No Director  Ama Qamata  TV Show  South Africa      2021
4      s2    No Director  Khosi Ngema  TV Show  South Africa      2021
```

```

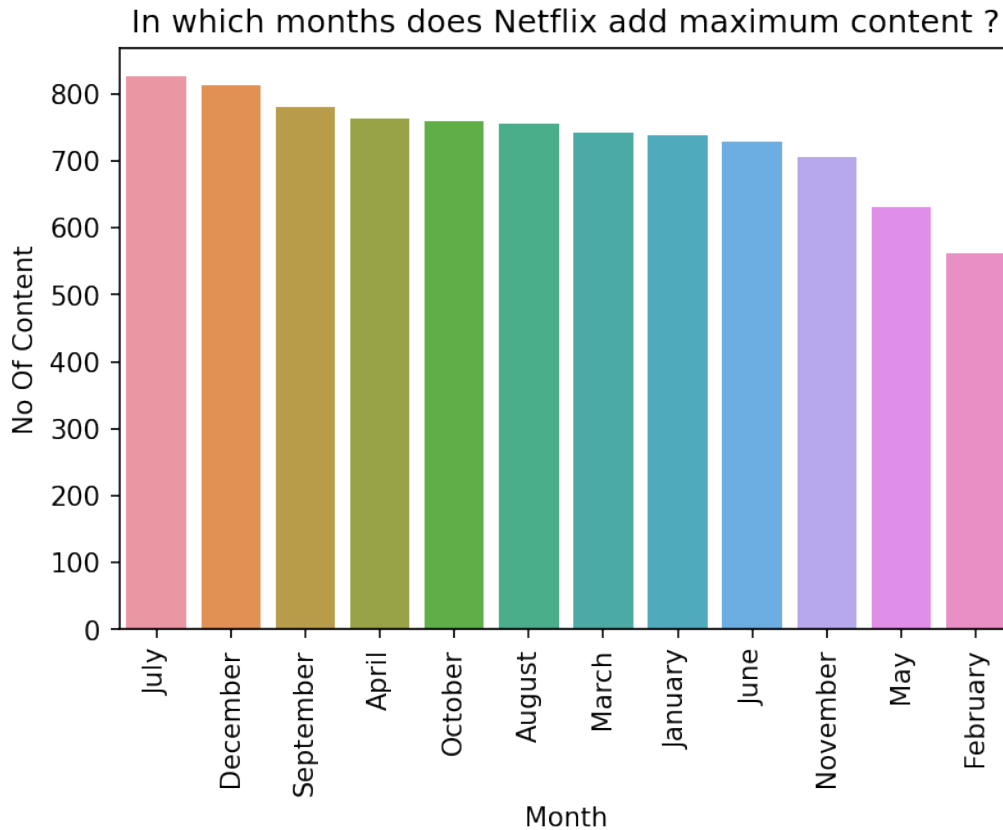
rating  duration      listed_in  Year  Added  Month  Added
0  PG-13    61-90    Documentaries    2021  September
1  TV-MA  2 Seasons  International TV Shows    2021  September
2  TV-MA  2 Seasons      TV Dramas    2021  September
3  TV-MA  2 Seasons    TV Mysteries    2021  September
4  TV-MA  2 Seasons  International TV Shows    2021  September
```

```
[91]: month_year_plot=final_df.groupby(by='Month Added').nunique()['show_id'].
      ↪sort_values(ascending=False)
      month_year_plot
```

```
[91]: Month Added
      July      827
      December  813
      September 780
      April     764
      October   760
      August    755
      March     742
      January   738
      June      728
      November  705
      May       632
      February  563
      Name: show_id, dtype: int64
```

```
[ ]:
```

```
[92]: plt.figure(dpi=150)
      sns.barplot(x=list(month_year_plot.index),y=list(month_year_plot))
      plt.xticks(rotation=90)
      plt.xlabel('Month')
      plt.ylabel('No Of Content')
      plt.title('In which months does Netflix add maximum content ?')
      plt.savefig('Image5.jpg')
      plt.show()
```



```
[ ]:
```

7 Question 4 :

7.0.1 Which directors are the most popular?

```
[93]: final_df.head()
```

```
[93]:
```

	show_id	Director	Cast	type	country	release_year	\
0	s1	Kirsten Johnson	Others	Movie	United States	2020	
1	s2	No Director	Ama Qamata	TV Show	South Africa	2021	
2	s2	No Director	Ama Qamata	TV Show	South Africa	2021	
3	s2	No Director	Ama Qamata	TV Show	South Africa	2021	
4	s2	No Director	Khosi Ngema	TV Show	South Africa	2021	

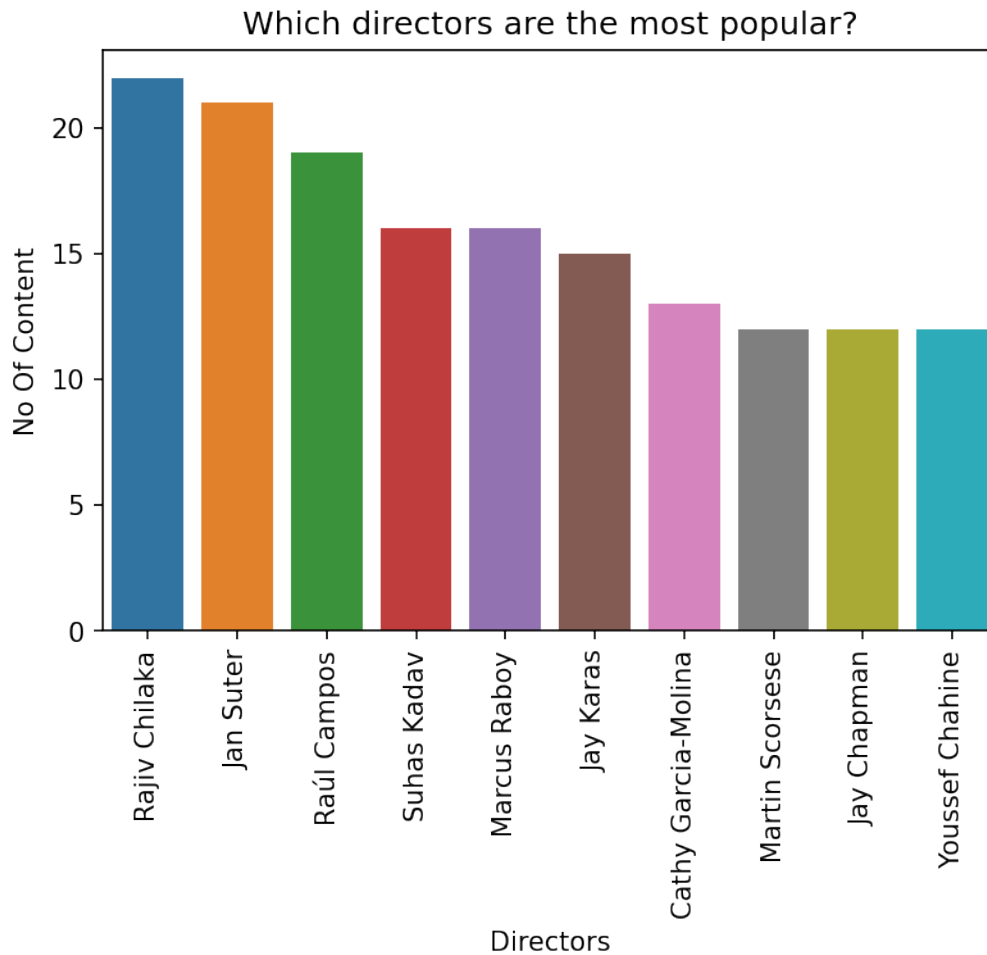
	rating	duration	listed_in	Year	Added	Month	Added
0	PG-13	61-90	Documentaries	2021	September		
1	TV-MA	2 Seasons	International TV Shows	2021	September		
2	TV-MA	2 Seasons	TV Dramas	2021	September		
3	TV-MA	2 Seasons	TV Mysteries	2021	September		

[]:

```
[94]: director_plot=final_df.groupby('Director').nunique()['show_id'].
      ↪sort_values(ascending=False)
      director_plot
```

```
[94]: Director
No Director                2634
Rajiv Chilaka                22
Jan Suter                   21
Raúl Campos                 19
Suhas Kadav                 16
...
Brandon Camp                 1
Juan Antin                   1
Juan Antonio de la Riva      1
Juan Camilo Pinzon           1
María Jose Cuevas            1
Name: show_id, Length: 4994, dtype: int64
```

```
[95]: plt.figure(dpi=150)
sns.barplot(x=list(director_plot.index[1:11]),y=list(director_plot[1:11]))
plt.xticks(rotation=90)
plt.xlabel('Directors')
plt.ylabel('No Of Content')
plt.title('Which directors are the most popular?')
plt.savefig('Image6.jpg')
plt.show()
```



[]:

[]:

[]:

8 Question 5 :

8.0.1 Which cast members are the most popular?

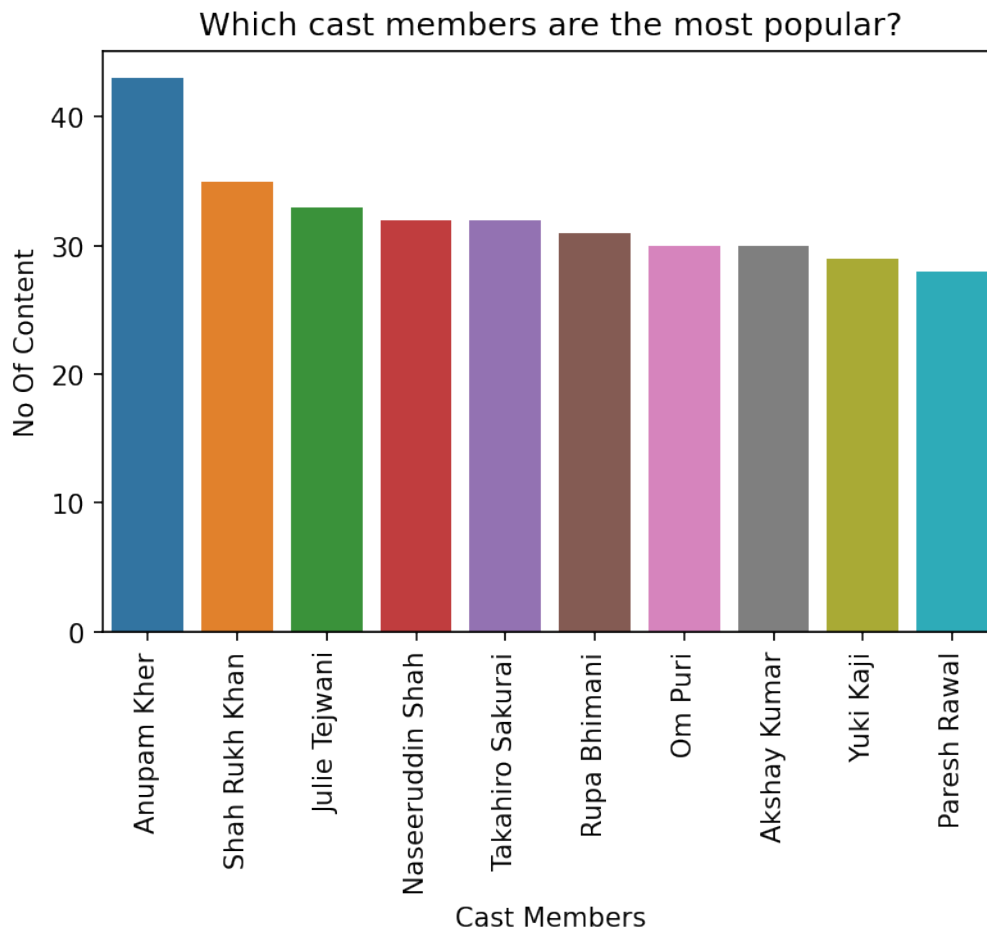
```
[96]: cast_plot=final_df.groupby('Cast').nunique()['show_id'].
      ↪sort_values(ascending=False)
      cast_plot
```

```
[96]: Cast
      Others          825
```

Anupam Kher	43
Shah Rukh Khan	35
Julie Teiwani	33
Naseeruddin Shah	32
...	
Jamie Lee	1
Jamie Kenna	1
Jamie Kaler	1
Jamie Johnston	1
Şopê Dirîsû	1

Name: show_id, Length: 36440, dtype: int64

```
[97]: plt.figure(dpi=150)
sns.barplot(x=list(cast_plot.index[1:11]),y=list(cast_plot[1:11]))
plt.xticks(rotation=90)
plt.xlabel('Cast Members')
plt.ylabel('No Of Content')
plt.title('Which cast members are the most popular?')
plt.savefig('Image7.jpg')
plt.show()
```



9 Question 6 :

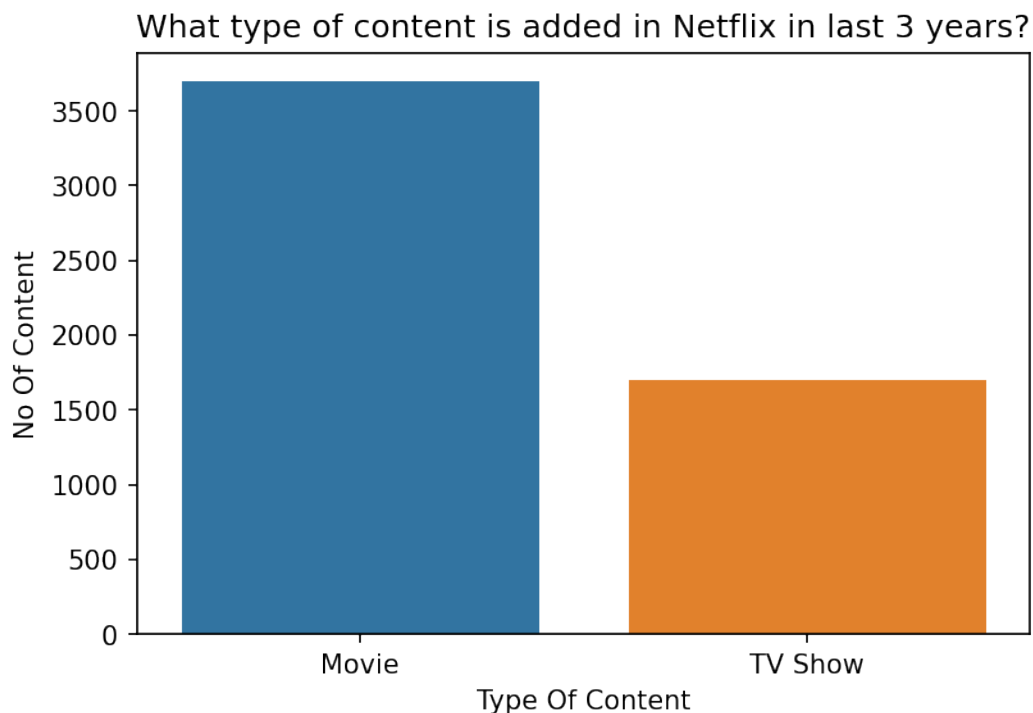
9.0.1 What type of content is mostly added in Netflix in the last 3 years?

```
[ ]:
```

```
[98]: cast_plot=final_df[final_df['Year Added']>=2019].groupby('type').  
      ↪nunique()['show_id'].sort_values(ascending=False)  
      cast_plot
```

```
[98]: type  
      Movie      3701  
      TV Show   1702  
      Name: show_id, dtype: int64
```

```
[99]: plt.figure(dpi=150)  
      sns.barplot(x=list(cast_plot.index),y=list(cast_plot))  
  
      plt.xlabel('Type Of Content')  
      plt.ylabel('No Of Content')  
      plt.title('What type of content is added in Netflix in last 3 years?')  
      plt.savefig('Image8.jpg')  
      plt.show()
```

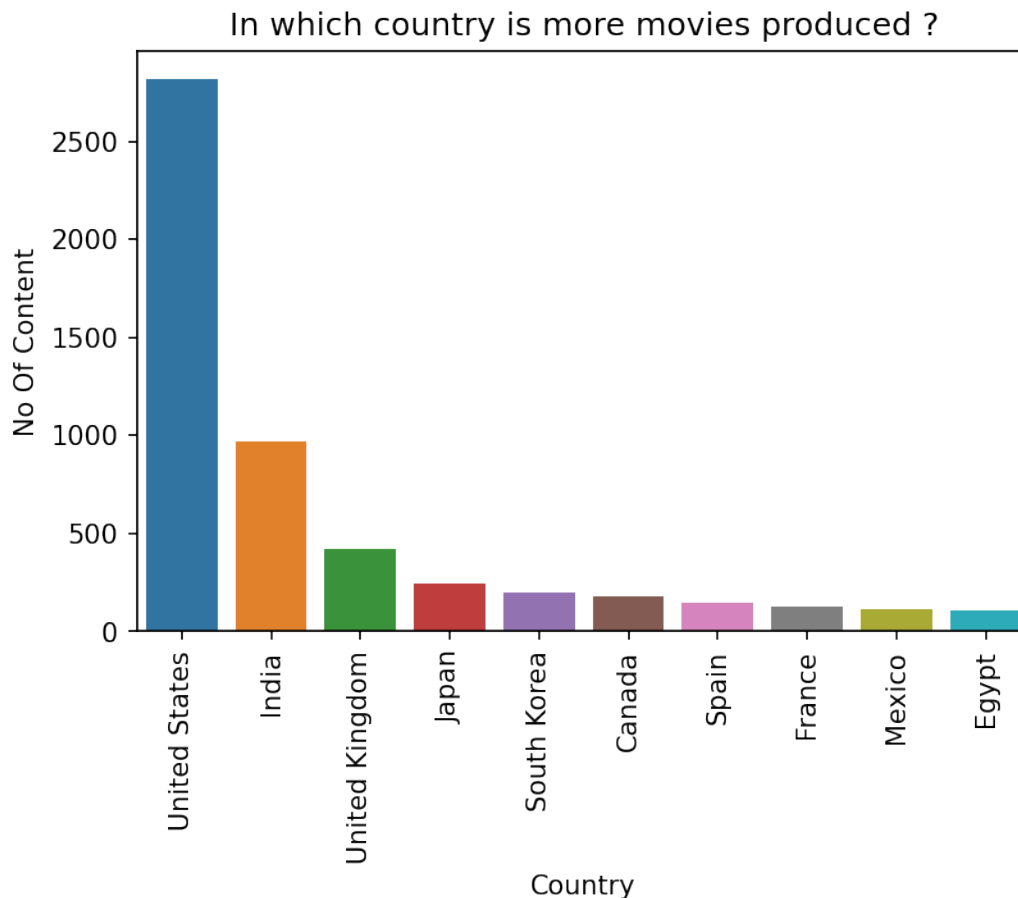


10 Question 7 :

10.0.1 In which country is more movies produced?

[]:

```
[100]: country_plot=final_df.groupby('country').nunique()['show_id'].  
        ↪sort_values(ascending=False)  
country_plot  
plt.figure(dpi=150)  
sns.barplot(x=list(country_plot.iloc[[0, 1, 3, 4, 5, 6, 7, 8, 9, 10]].index),  
            y=list(country_plot.iloc[[0, 1, 3, 4, 5, 6, 7, 8, 9, 10]]))  
plt.xlabel('Country')  
plt.ylabel('No Of Content')  
plt.title('In which country is more movies produced ?')  
plt.xticks(rotation=90)  
plt.savefig('Image9.jpg')  
plt.show()
```



11 Question 8 :

11.0.1 What kind of rated content is most available in Netflix?

```
[101]: final_df.head()
```

```
[101]:
```

	show_id	Director	Cast	type	country	release_year	\
0	s1	Kirsten Johnson	Others	Movie	United States	2020	
1	s2	No Director	Ama Qamata	TV Show	South Africa	2021	
2	s2	No Director	Ama Qamata	TV Show	South Africa	2021	
3	s2	No Director	Ama Qamata	TV Show	South Africa	2021	
4	s2	No Director	Khosi Ngema	TV Show	South Africa	2021	

	rating	duration	listed_in	Year Added	Month Added
0	PG-13	61-90	Documentaries	2021	September
1	TV-MA	2 Seasons	International TV Shows	2021	September
2	TV-MA	2 Seasons	TV Dramas	2021	September
3	TV-MA	2 Seasons	TV Mysteries	2021	September
4	TV-MA	2 Seasons	International TV Shows	2021	September

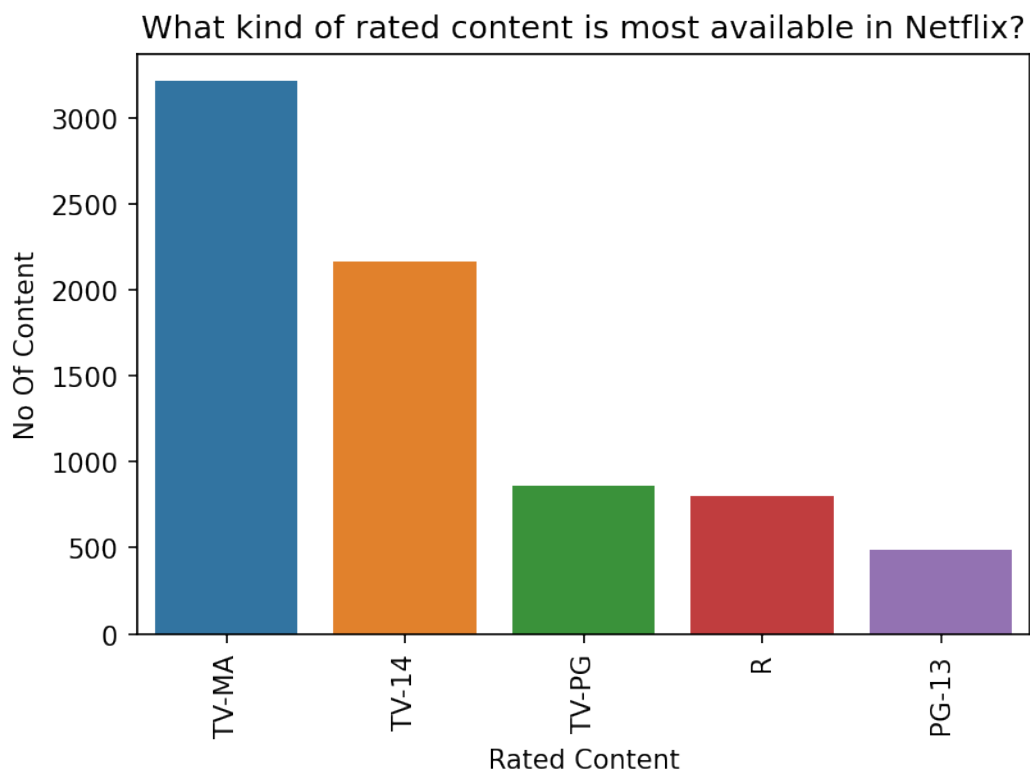
```
[102]: final_df['rating'].unique()
```

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

- TV-Y - Designed to be appropriate for all children
- TV-Y7 - Suitable for ages 7 and up
- G - Suitable for General Audiences
- TV-G - Suitable for General Audiences
- PG - Parental Guidance suggested
- TV-PG - Parental Guidance suggested
- PG-13 - Parents strongly cautioned. May be Inappropriate for ages 12 and under.
- TV-14 - Parents strongly cautioned. May not be suitable for ages 14 and under.
- R - Restricted. May be inappropriate for ages 17 and under.
- TV-MA - For Mature Audiences. May not be suitable for ages 17 and under.
- NC-17 - Inappropriate for ages 17 and under
- NR - Stands for No Result. It occurs when a film is not submitted to the MPAA for rating.
- UR - It means unrated, which means the film hasn't been submitted to the MPAA for a rating.

- TV-Y7-FV - Indicates that the content is designed for children aged 7 and older, and contains “fantasy violence” that may be more intense or combative than other TV-Y7 programs.

```
[103]: rating_plot=final_df.groupby('rating').nunique()['show_id'].
        ↳sort_values(ascending=False)
rating_plot
plt.figure(dpi=150)
sns.barplot(x=list(rating_plot.index[:5]),
            y=list(rating_plot[:5]))
plt.xticks(rotation=90)
plt.xlabel('Rated Content')
plt.ylabel('No Of Content')
plt.title('What kind of rated content is most available in Netflix?')
plt.savefig('Image10.jpg')
plt.show()
```



```
[ ]:
```

```
[ ]:
```

12 Question 9 :

12.0.1 Which is the most popular genre?

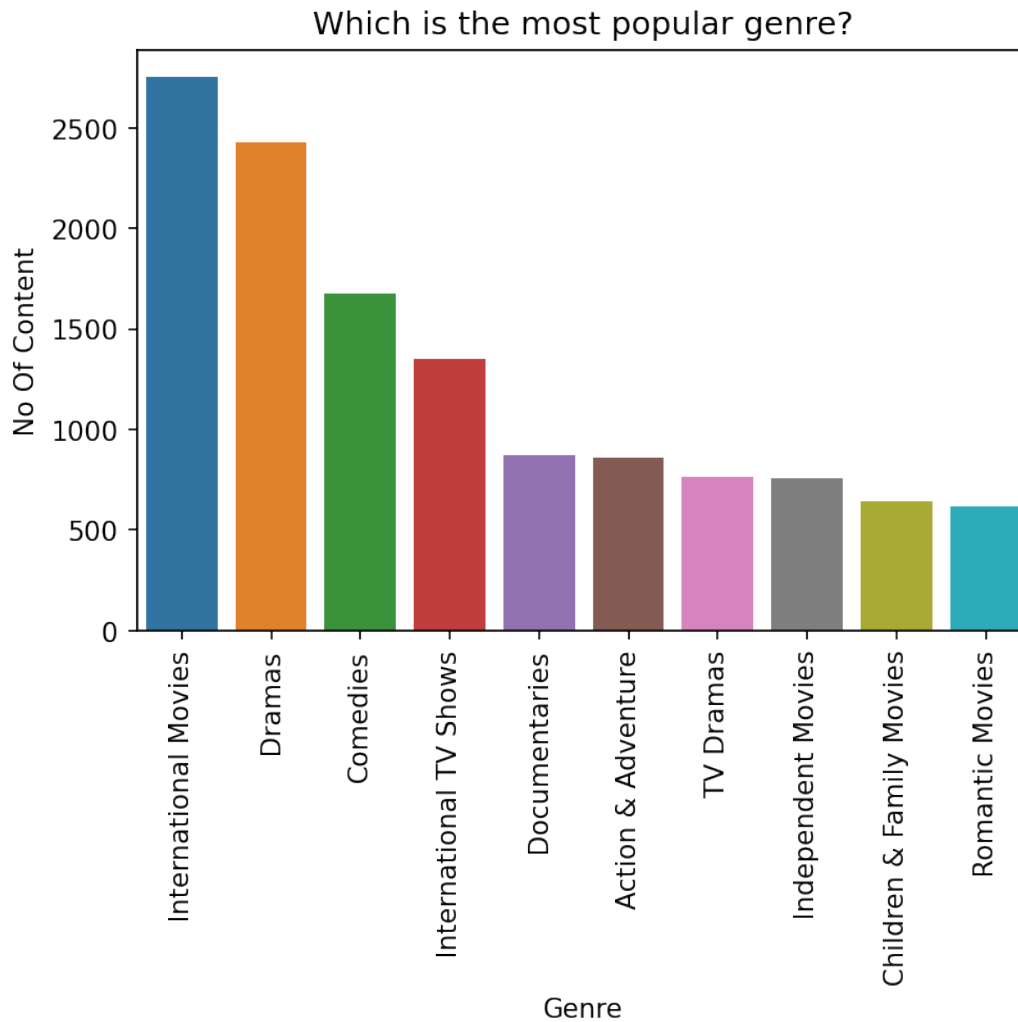
```
[104]: final_df.head()
```

```
[104]:  show_id      Director      Cast      type      country  release_year  \
0      s1  Kirsten Johnson      Others      Movie  United States      2020
1      s2      No Director  Ama Qamata  TV Show  South Africa      2021
2      s2      No Director  Ama Qamata  TV Show  South Africa      2021
3      s2      No Director  Ama Qamata  TV Show  South Africa      2021
4      s2      No Director  Khosi Ngema  TV Show  South Africa      2021

      rating  duration      listed_in  Year Added Month Added
0  PG-13      61-90      Documentaries      2021  September
1  TV-MA  2 Seasons  International TV Shows      2021  September
2  TV-MA  2 Seasons      TV Dramas      2021  September
3  TV-MA  2 Seasons      TV Mysteries      2021  September
4  TV-MA  2 Seasons  International TV Shows      2021  September
```

```
[ ]:
```

```
[105]: listed_in_plot=final_df.groupby('listed_in').nunique()['show_id'].
      ↪sort_values(ascending=False)
listed_in_plot
plt.figure(dpi=150)
sns.barplot(x=list(listed_in_plot.index[:10]),
            y=list(listed_in_plot[:10]))
plt.xlabel('Genre')
plt.ylabel('No Of Content')
plt.title('Which is the most popular genre?')
plt.xticks(rotation=90)
plt.savefig('Image11.jpg')
plt.show()
```



[]:

[]:

[]:

13 Question 10 :

13.0.1 What is the most preferred duration for each type of content?

[106]: `final_df.head()`

```
[106]:   show_id      Director      Cast      type      country  release_year \
0      s1  Kirsten Johnson    Others    Movie  United States        2020
```

1	s2	No Director	Ama Qamata	TV Show	South Africa	2021
2	s2	No Director	Ama Qamata	TV Show	South Africa	2021
3	s2	No Director	Ama Qamata	TV Show	South Africa	2021
4	s2	No Director	Khosi Ngema	TV Show	South Africa	2021

	rating	duration	listed_in	Year Added	Month Added
0	PG-13	61-90	Documentaries	2021	September
1	TV-MA	2 Seasons	International TV Shows	2021	September
2	TV-MA	2 Seasons	TV Dramas	2021	September
3	TV-MA	2 Seasons	TV Mysteries	2021	September
4	TV-MA	2 Seasons	International TV Shows	2021	September

```
[107]: duration_plot=final_df.groupby(by=['type','duration']).nunique()['show_id']
duration_plot
```

```
[107]: type      duration
Movie      0-30          130
          121-150        897
          151-180        198
          31-60          357
          61-90         1506
          91-120        2996
          Above 180         47
TV Show    1 Season      1793
          10 Seasons         7
          11 Seasons         2
          12 Seasons         2
          13 Seasons         3
          15 Seasons         2
          17 Seasons         1
          2 Seasons        425
          3 Seasons        199
          4 Seasons         95
          5 Seasons         65
          6 Seasons         33
          7 Seasons         23
          8 Seasons         17
          9 Seasons          9
Name: show_id, dtype: int64
```

```
[108]: movie_plot=duration_plot.loc['Movie'].sort_values(ascending=False)
movie_plot
```

```
[108]: duration
91-120      2996
61-90       1506
121-150      897
```

```

31-60          357
151-180        198
0-30           130
Above 180       47
Name: show_id, dtype: int64

```

```

[109]: tv_show_plot=duration_plot.loc['TV Show'].sort_values(ascending=False)
tv_show_plot

```

```

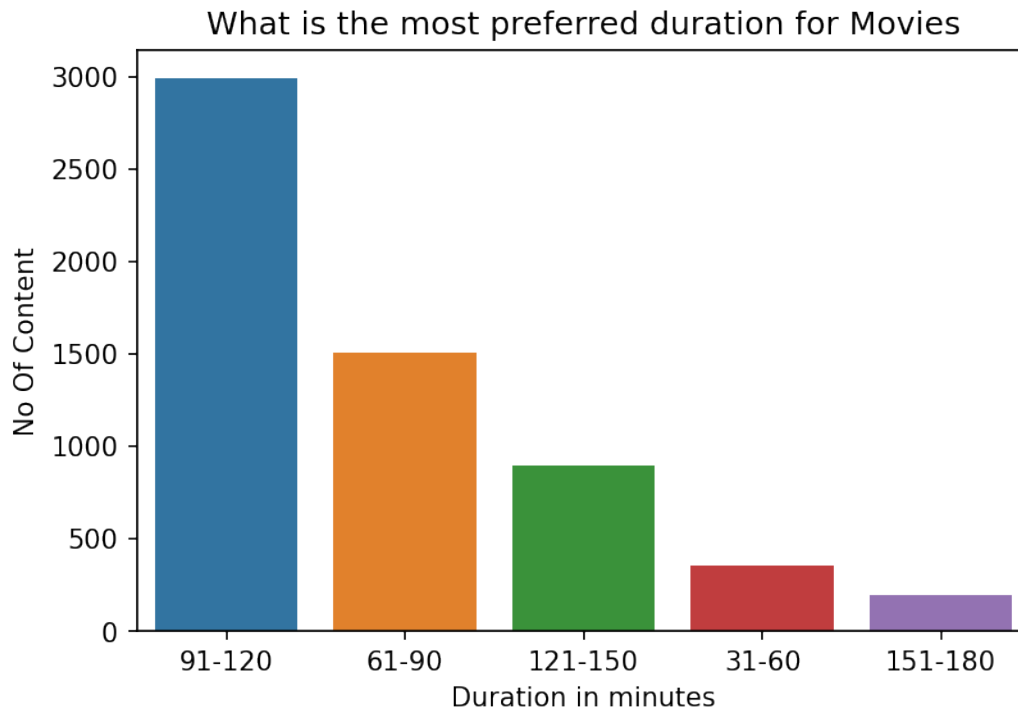
[109]: duration
1 Season      1793
2 Seasons     425
3 Seasons     199
4 Seasons      95
5 Seasons      65
6 Seasons      33
7 Seasons      23
8 Seasons      17
9 Seasons       9
10 Seasons      7
13 Seasons      3
11 Seasons      2
12 Seasons      2
15 Seasons      2
17 Seasons      1
Name: show_id, dtype: int64

```

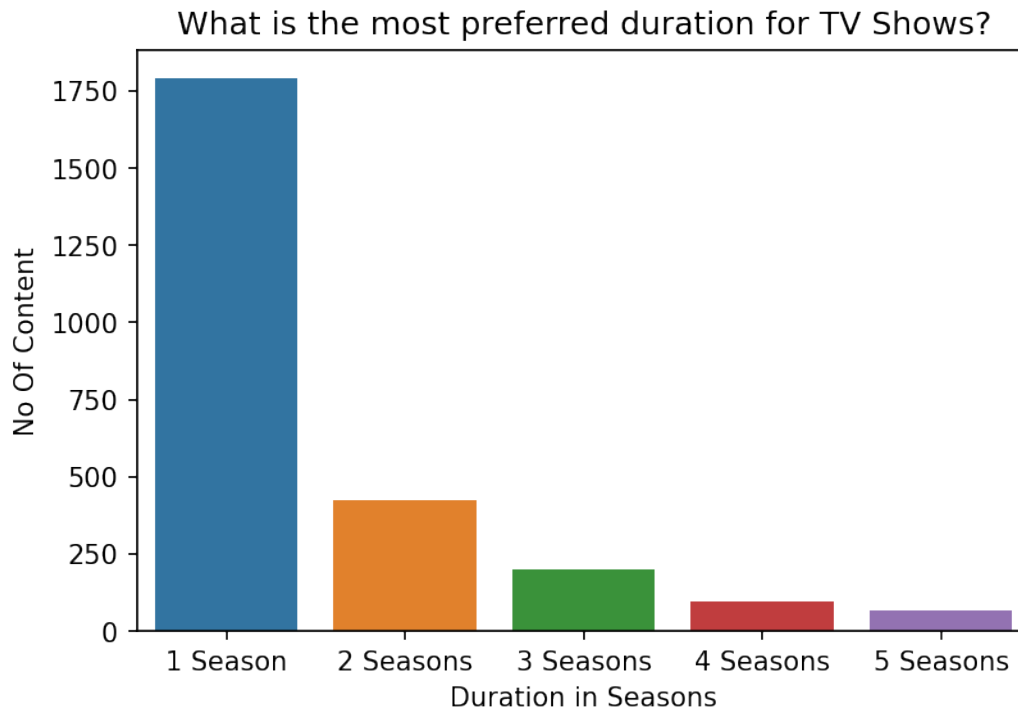
```

[110]: plt.figure(dpi=150)
sns.barplot(x=movie_plot.index[:5],
            y=list(movie_plot[:5]))
# plt.xticks(rotation=90)
plt.xlabel('Duration in minutes')
plt.ylabel('No Of Content')
plt.title('What is the most preferred duration for Movies')
plt.savefig('Image12.jpg')
plt.show()

```

```
[111]: plt.figure(dpi=150)
sns.barplot(x=list(tv_show_plot.index[:5]),
            y=list(tv_show_plot[:5]))
# plt.xticks(rotation=90)
plt.xlabel('Duration in Seasons')
plt.ylabel('No Of Content')
plt.title('What is the most preferred duration for TV Shows?')
plt.savefig('Image13.jpg')
plt.show()
```



[]:

[]:

14 Question 11:

14.0.1 No of content added in Netflix in each month over the last 5 years?

[112]: `final_df.head()`

```
[112]:  show_id      Director      Cast      type      country  release_year  \
0      s1  Kirsten Johnson    Others    Movie  United States      2020
1      s2    No Director  Ama Qamata  TV Show  South Africa      2021
2      s2    No Director  Ama Qamata  TV Show  South Africa      2021
3      s2    No Director  Ama Qamata  TV Show  South Africa      2021
4      s2    No Director  Khosi Ngema  TV Show  South Africa      2021

   rating  duration      listed_in  Year Added Month Added
0  PG-13    61-90    Documentaries    2021  September
1  TV-MA  2 Seasons  International TV Shows    2021  September
2  TV-MA  2 Seasons          TV Dramas    2021  September
3  TV-MA  2 Seasons    TV Mysteries    2021  September
4  TV-MA  2 Seasons  International TV Shows    2021  September
```

[]:

```
[113]: year_month_df = final_df.groupby(['Year Added', 'Month Added']).  
        ↳unique()['show_id']  
year_month_df=year_month_df.loc[[2017,2018,2019,2020,2021]]  
year_month_df
```

```
[113]: Year Added  Month Added  
2017      April      93  
          August     115  
          December   118  
          February    82  
          January     72  
          July        79  
          June        94  
          March       125  
          May         86  
          November    85  
          October     126  
          September   113  
2018      April      115  
          August     164  
          December   185  
          February    87  
          January     129  
          July        152  
          June        78  
          March       173  
          May         97  
          November   154  
          October     191  
          September   124  
2019      April      162  
          August     131  
          December   215  
          February    148  
          January     153  
          July        157  
          June        168  
          March       172  
          May         139  
          November    255  
          October     193  
          September   123  
2020      April      177  
          August     129  
          December   169
```

	February	114
	January	205
	July	146
	June	156
	March	137
	May	157
	November	154
	October	167
	September	168
2021	April	188
	August	178
	February	109
	January	132
	July	257
	June	207
	March	112
	May	132
	September	193

Name: show_id, dtype: int64

[]:

```
[114]: def function(series):
        indexes=set(series.index)
        dict1={}
        for month in ['January','February','March','April','May',
            ↳
            ↳ 'June','July','August','September','October','November','December']:

            if month in indexes:
                dict1[month]=series.loc[month]

            else:
                dict1[month]=0

        return pd.Series(dict1)
```

[]:

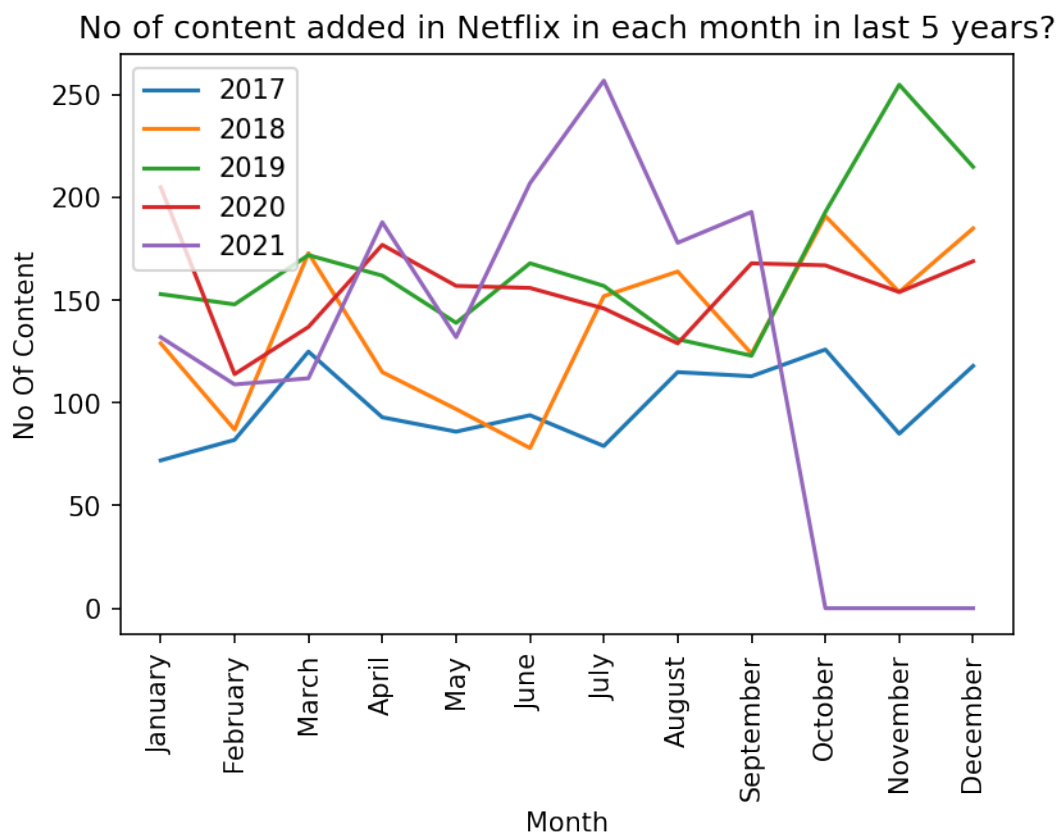
[]:

```
[115]: plt.figure(dpi=150)
        for year in [2017,2018,2019,2020,2021]:
            series=function(year_month_df.loc[year])
            sns.lineplot(x=['January','February','March','April','May',
                ↳
                ↳ 'June','July','August','September','October','November','December'],
```

```

y=series.loc[['January','February','March','April','May',
             'June','July','August','September','October','November','December'])
plt.legend([2017,2018,2019,2020,2021])
plt.xticks(rotation=90)
plt.xlabel('Month')
plt.ylabel('No Of Content')
plt.title('No of content added in Netflix in each month in last 5 years?')
plt.savefig('Image14.jpg')
plt.show()

```



[]:

15 Final Analysis

The given dataset has 8807 rows and 12 columns, out of which we can drop 2 columns - description and title, which are nominal columns and would not provide much scope for deriving insights from the data, which would leave us with 10 columns. The grain of the data is the column show_id which is a unique identifier for every row. All the columns are of object type except release year, which is of the type integer. From the column date_added, we can make 2 new columns - Year Added

of type integer and Month Added of type object and finally drop the original column date_added. We would finally be left with 11 columns in total.

Using the info method, we observed that the columns - director, cast, country, date_added, rating and duration had Nan values. All of the Nan values were filled with a certain value depending on the count of missing values and nature of the column, and none of the rows were dropped.

At the end of the above steps, all the 11 columns had 8807 non-null values.

16 Recommendations :

- Since from the last few years, there has been an increase in the no of TV shows being released and added on Netflix year on year, and decrease in the no of movies being released and added on Netflix year on year, it shows that there is an increasing demand for TV shows, and therefore Netflix should try to add more TV shows going forward.
- Even though there has been an increasing trend for TV shows, but in the last 3 years, the no of movies added compared to TV shows is twice, therefore Netflix should try to add more TV shows on its platform.
- The months May and February are the months in which content added on Netflix is the least. Therefore Netflix can try to add more content in these 2 months to ensure that people have enough content options during May and February.
- Netflix should try to add more content for the most popular directors such as Rajiv Chilaka, Jan Suter and Raul Campos.
- Netflix should try to add more content for the most popular casts such as Anupam Kher and Shahrukh Khan. Also, the presence of casts such as Julie Teiwani, Rupa Bhimani and Takahiro Sakurai in the top 10 tells us that animated movies are very popular and Netflix should try to add more animated movies as well.
- Netflix should try to add more content having ratings “ TV-MA and TV-14.
- Netflix should try to add more content which are produced in the United States, as clearly it is the most popular country where content is produced.
- Netflix should try to add more movies which have a duration in the range of 91 to 120 minutes.
- Netflix should try to add more TV shows which have 1 season.
- Netflix should try to add more content for the genres - International Movies, Dramas and Comedies.

[]:

[]:

[]:

[]:

[]:

[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	
[]:	