```
import numpy as np
import pandas as pd
!pip install seaborn
import seaborn as sns
import matplotlib.pyplot as plt
```

```
Requirement already satisfied: seaborn in /usr/local/lib/python3.10/dist-packages (0.13.1)
Requirement already satisfied: numpy!=1.24.0,>=1.20 in /usr/local/lib/python3.10/dist-packages (from seaborn) (1.25.2)
Requirement already satisfied: pandas>=1.2 in /usr/local/lib/python3.10/dist-packages (from seaborn) (2.0.3)
Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in /usr/local/lib/python3.10/dist-packages (from seaborn) (3.7.1)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.2.
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (4.5
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.4
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (24.0)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (9.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (3.1.
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas>=1.2->seaborn) (2023.4)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas>=1.2->seaborn) (2024.1)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.7->matplotlib!=3.6.1,>=3.4->
```

```
! gdown https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/000/940/original/netflix.csv
```

```
Downloading...
From: https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/000/940/original/netflix.csv
To: /content/netflix.csv
100% 3.40M/3.40M [00:00<00:00, 44.7MB/s]
```

```
netflix = pd.read_csv('/content/netflix.csv')
netflix
```

	show_id	type	title	director	cast	country	date_added	release_year
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan	India	September 24, 2021	2021

Next steps: Generate code with netflix | View recommended plots

```
netflix.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 12 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   show_id       8807 non-null   object
 1   type          8807 non-null   object
```

```
 2   title         8807 non-null   object
 3   director      6173 non-null   object
 4   cast          7982 non-null   object
 5   country       7976 non-null   object
 6   date_added    8797 non-null   object
 7   release_year  8807 non-null   int64
 8   rating        8803 non-null   object
 9   duration      8804 non-null   object
 10  listed_in     8807 non-null   object
 11  description   8807 non-null   object
dtypes: int64(1), object(11)
memory usage: 825.8+ KB
```

```
print(type(netflix))
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
netflix.columns
```

```
Index(['show_id', 'type', 'title', 'director', 'cast', 'country', 'date_added',
       'release_year', 'rating', 'duration', 'listed_in', 'description'],
      dtype='object')
```

```
netflix.keys()
```

```
Index(['show_id', 'type', 'title', 'director', 'cast', 'country', 'date_added',
       'release_year', 'rating', 'duration', 'listed_in', 'description'],
      dtype='object')
```

Un-nesting the columns

```
netflix['director'] = netflix['director'].apply(lambda x: x.split(', ') if isinstance(x, str) else x)
netflix['cast'] = netflix['cast'].apply(lambda x: x.split(', ') if isinstance(x, str) else x)
netflix['country'] = netflix['country'].apply(lambda x: x.split(', ') if isinstance(x, str) else x)
```

```
a = netflix.explode('director')[['title', 'director']].reset_index().drop(columns =['index'])
```

```
b = netflix.explode('cast')[['title', 'cast']].reset_index().drop(columns =['index'])
```

```
c = netflix.explode('country')[['title', 'country']].reset_index().drop(columns =['index'])
```

```
merged_ab = netflix.merge(a, on='title', how='outer')

merged_abc = merged_ab.merge(b, on='title', how='outer')

netflix_merged = merged_abc.merge(c, on='title', how='outer')
```

```
netflix_1a = netflix_merged.drop(columns=['director_x', 'cast_x', 'country_x']).rename(columns={'director_y':'director',    'cast_y':'cast',
netflix_1a
```

	show_id	type	title	date_added	release_year	rating	duration	listed_in	description	director	cast	country
0	s1	Movie	Dick Johnson Is Dead	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...	Kirsten Johnson	NaN	United States
1	s2	TV Show	Blood & Water	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	NaN	Ama Qamata	South Africa
2	s2	TV Show	Blood & Water	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	NaN	Khosi Ngema	South Africa
3	s2	TV Show	Blood & Water	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape	NaN	Gail Mabalane	South Africa

Next steps: Generate code with netflix_1a View recommended plots

Treating duration columns

```
netflix_1a['movies_duration'] = netflix_1a[netflix_1a['type']=='Movie']['duration'].apply(lambda x:int(x.split(' ')[0]) if isinstance(x, str
```

```
netflix_1a['tv_seasons'] = netflix_1a[netflix_1a['type']=='TV Show']['duration'].apply(lambda x:int(x.split(' ')[0])if isinstance(x, str)els
```

```
netflix_1a = netflix_1a.drop(columns =['duration'] )
netflix_1a
```

	show_id	type	title	date_added	release_year	rating	listed_in	description	director	cast	country	movies_durati
0	s1	Movie	Dick Johnson Is Dead	September 25, 2021	2020	PG-13	Documentaries	As her father nears the end of his life, filmm...	Kirsten Johnson	NaN	United States	90
1	s2	TV Show	Blood & Water	September 24, 2021	2021	TV-MA	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	NaN	Ama Qamata	South Africa	Na
2	s2	TV Show	Blood & Water	September 24, 2021	2021	TV-MA	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	NaN	Khosi Ngema	South Africa	Na
3	s2	TV Show	Blood & Water	September 24, 2021	2021	TV-MA	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	NaN	Gail Mabalane	South Africa	Na
4	s2	TV Show	Blood & Water	September 24, 2021	2021	TV-MA	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	NaN	Thabang Molaba	South Africa	Na
...	...	...	...	...	...	...	...	...	...	...	...	

Next steps: Generate code with netflix_1a View recommended plots

Treating date added column

```
unique_dates = netflix_1a['date_added'].unique()
```

```
print(unique_dates)

netflix_1a['date_added'] = pd.to_datetime(netflix_1a['date_added'], errors='coerce', infer_datetime_format=True)

invalid_dates = netflix_1a[netflix_1a['date_added'].isnull()]['date_added']
print("Invalid Dates:")
print(invalid_dates)
```

```
['September 25, 2021' 'September 24, 2021' 'September 23, 2021' ...
 'December 6, 2018' 'March 9, 2016' 'January 11, 2020']
Invalid Dates:
59259   NaT
59260   NaT
59261   NaT
59262   NaT
59263   NaT
         ..
88099   NaT
88100   NaT
88101   NaT
88374   NaT
88765   NaT
Name: date_added, Length: 829, dtype: datetime64[ns]
<ipython-input-195-b0620c784bca>:4: UserWarning: The argument 'infer_datetime_format' is deprecated and will be removed in a future vers
  netflix_1a['date_added'] = pd.to_datetime(netflix_1a['date_added'], errors='coerce', infer_datetime_format=True)
```

Start coding or generate with AI.

Treating the null values in the categorical columns and continous variables.

```
netflix_1a[netflix_1a['country'].isna()]
```

	show_id	type	title	date_added	release_year	rating	listed_in	description	director	cast	country	movies_duration
20	s3	TV Show	Ganglands	2021-09-24	2021	TV-MA	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...	Julien Leclercq	Sami Bouajila	NaN	NaN
21	s3	TV Show	Ganglands	2021-09-24	2021	TV-MA	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...	Julien Leclercq	Tracy Gotoas	NaN	NaN
22	s3	TV Show	Ganglands	2021-09-24	2021	TV-MA	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...	Julien Leclercq	Samuel Jouy	NaN	NaN
23	s3	TV Show	Ganglands	2021-09-24	2021	TV-MA	Crime TV Shows, International	To protect his family from a	Julien Leclercq	Nabiha Akkari	NaN	NaN

```
netflix_1a['country'].fillna('Unknown Country', inplace = True)
netflix_1a['director'].fillna('Unknown Director', inplace = True)
netflix_1a['cast'].fillna('Unknown Cast', inplace = True)
netflix_1a['movies_duration'].fillna(0, inplace = True)

netflix_1a.head()
```

	show_id	type	title	date_added	release_year	rating	listed_in	description	director	cast	country	movies_duration	tv
0	s1	Movie	Dick Johnson Is Dead	2021-09-25	2020	PG-13	Documentaries	As her father nears the end of his life, filmm...	Kirsten Johnson	Unknown Cast	United States	90.0	
1	s2	TV Show	Blood & Water	2021-09-24	2021	TV-MA	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	Unknown Director	Ama Qamata	South Africa	0.0	
		TV	Blood &				International TV Shows, TV	After crossing	Unknown	Khosi	South		

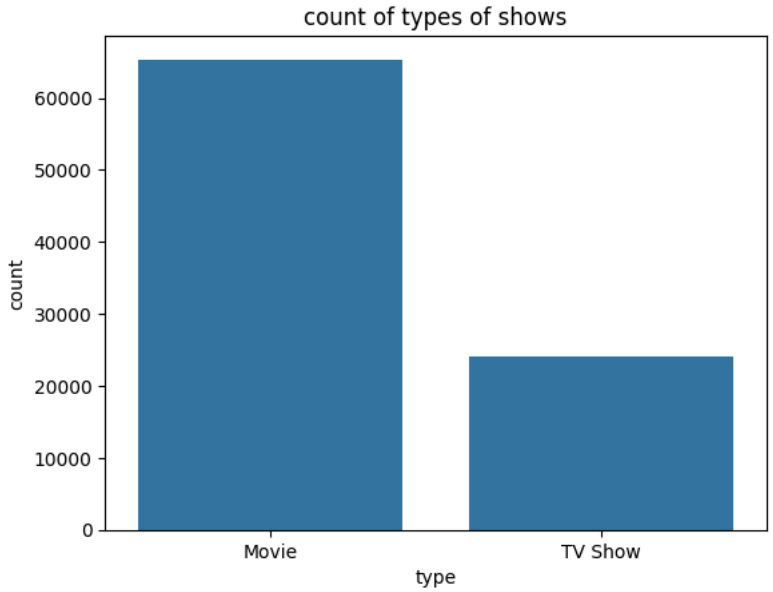
Next steps: Generate code with netflix_1a | View recommended plots

#1. Finding the counts of each categorical variable both using graphical and non-graphical analysis.

```
netflix_1a['type'].value_counts()
```

```
type
Movie      65346
TV Show    24036
Name: count, dtype: int64
```

```
sns.countplot(x = 'type', data=netflix_1a)
plt.title('count of types of shows')
plt.xlabel('type')
plt.ylabel('count')
plt.show()
```



Looking at the graphical representation of count of movies and tv shows we see that netflix has produced more movies that the tv shows over the past years

```
netflix_1a['cast'].value_counts()
```

```
cast
Unknown Cast              1190
Alfred Molina               85
Liam Neeson                 82
John Krasinski              67
Frank Langella              66
                          ...
Arun Vijay                   1
Griffin Robert Faulkner      1
David Pendleton              1
Ahmed Adel                   1
```

```
Chittaranjan Tripathy         1
Name: count, Length: 36440, dtype: int64
```

```
netflix_1a['director'].value_counts()
```

```
director
Unknown Director         21937
Martin Scorsese            217
Steven Spielberg           205
Martin Campbell            154
Raja Gosnell               154
                         ...
Alexx Media                  1
Thomas Toby Parkinson        1
Bruce Gowers                 1
Emma Hatherley               1
Kirsten Johnson              1
Name: count, Length: 4994, dtype: int64
```

```
netflix_1a['country'].value_counts()
```

```
country
United States      30471
India               8537
United Kingdom      5713
Unknown Country     5421
Canada              3946
                   ...
Mongolia               1
Armenia                1
Panama                 1
Uganda                 1
Palestine              1
Name: count, Length: 128, dtype: int64
```

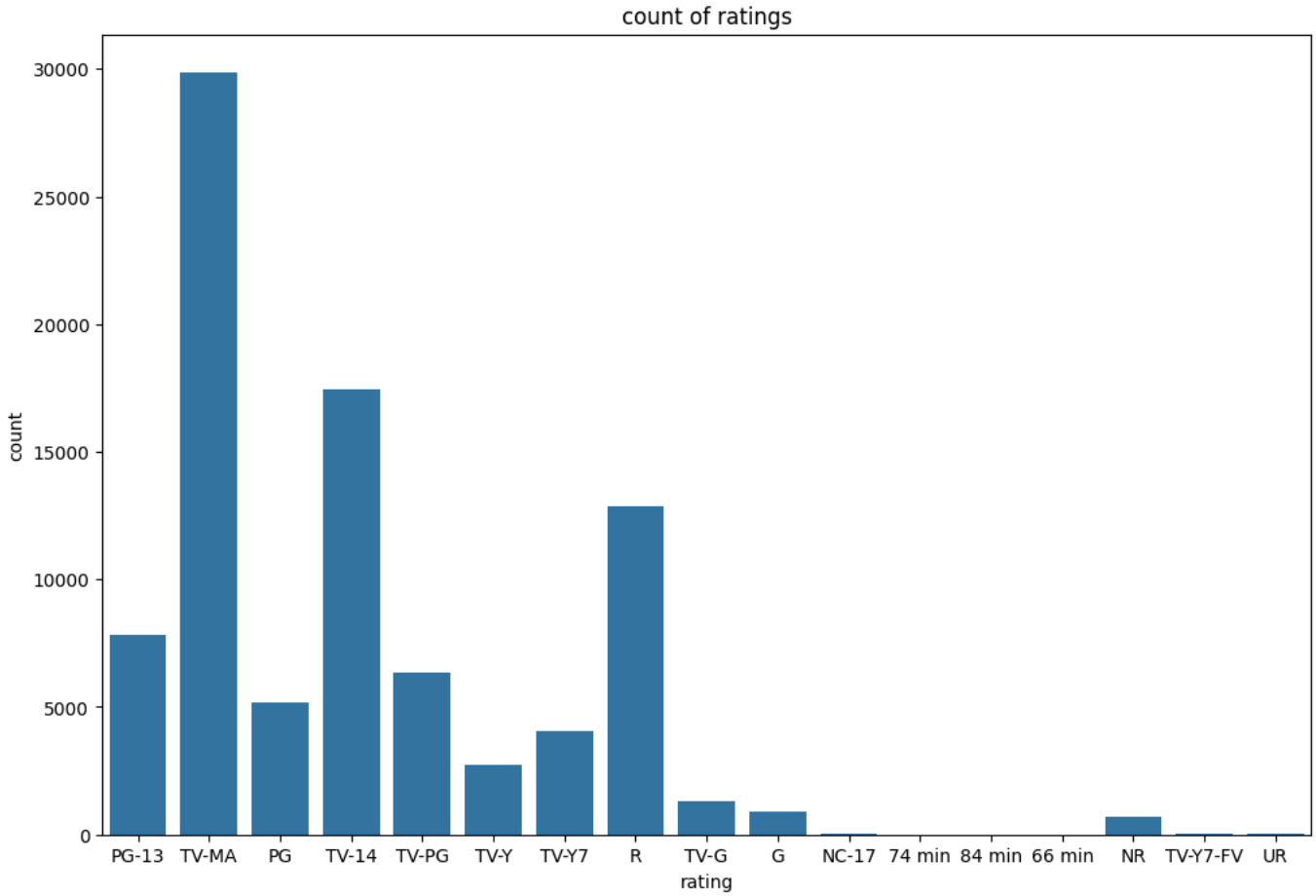
```
netflix_1a['release_year'].value_counts()
```

```
release_year
2018    10470
2019     9418
2017     9274
2020     8639
2016     8322
        ...
1946        5
1947        4
1943        4
1942        3
1925        1
Name: count, Length: 74, dtype: int64
```

```
netflix_1a['rating'].value_counts()
```

```
rating
TV-MA       29860
TV-14       17437
R           12865
PG-13        7814
TV-PG        6356
PG           5182
TV-Y7        4073
TV-Y         2745
TV-G         1329
G             878
NR            676
NC-17          55
TV-Y7-FV       39
UR             32
74 min          1
84 min          1
66 min          1
Name: count, dtype: int64
```

```
plt.figure(figsize=(12, 8))
sns.countplot(x = 'rating', data=netflix_1a)
plt.title('count of ratings')
plt.xlabel('rating')
plt.ylabel('count')
plt.show()
```



In the above graphical representation we clearly see that TV-MA is the highest used rating platform. Netflix could prioritize acquiring or producing more mature-themed content to cater to the audience that enjoys TV-MA rated shows. This could involve investing in adult-oriented series, movies, and documentaries.

```
netflix_1a['listed_in'].value_counts()
```

```
listed_in
Dramas, International Movies                            4255
Children & Family Movies, Comedies                      3578
Dramas, Independent Movies, International Movies        3465
Children & Family Movies                                2912
Comedies, Dramas, International Movies                  2841
                                                        ...
Crime TV Shows, International TV Shows, Reality TV         1
Docuseries, Reality TV, Teen TV Shows                      1
Reality TV, Science & Nature TV, TV Action & Adventure     1
Docuseries, Science & Nature TV, TV Comedies               1
British TV Shows, Docuseries, Reality TV                   1
Name: count, Length: 514, dtype: int64
```

comparision of tv shows vs movies

```
movies_df = netflix_1a[netflix_1a['type'] == 'Movie']

movie_each_country = movies_df.groupby('country')['title'].nunique().reset_index()


top_10_countries_movies = movie_each_country.sort_values(by='title', ascending=False).head(10)

top_10_countries_movies.reset_index().drop(columns=['index'])
```

	country	title
0	United States	2751
1	India	962
2	United Kingdom	532
3	Unknown Country	440
4	Canada	319
5	France	303
6	Germany	182
7	Spain	171
8	Japan	119
9	China	114

Since united states have the most movies/tv-shows netflix should partner up with more american actors/directors.

Emphasize the creation of original content from U.S.-based studios and filmmakers.

Taking feedback and opinion from the u.s community would help to understand what the users are looking for

```
tvshows_df = netflix_1a[netflix_1a['type'] == 'TV Show']

tvshows_each_country = tvshows_df.groupby('country')['title'].nunique().reset_index()

top_10_countries_tvshows = tvshows_each_country.sort_values(by='title', ascending=False).head(10)

top_10_countries_tvshows.reset_index().drop(columns=['index'])
```

	country	title
0	United States	938
1	Unknown Country	391
2	United Kingdom	272
3	Japan	199
4	South Korea	170
5	Canada	126
6	France	90
7	India	84
8	Taiwan	70
9	Australia	66

What is the best time to launch a TV show?

```
netflix_1a['week_added'] = netflix_1a['date_added'].dt.isocalendar().week
netflix_1a.head()
```

	show_id	type	title	date_added	release_year	rating	listed_in	description	director	cast	country	movies_duration	tv
0	s1	Movie	Dick Johnson Is Dead	2021-09-25	2020	PG-13	Documentaries	As her father nears the end of his life, filmm...	Kirsten Johnson	Unknown Cast	United States	90.0	
1	s2	TV Show	Blood & Water	2021-09-24	2021	TV-MA	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	Unknown Director	Ama Qamata	South Africa	0.0	
		TV	Blood &				International TV Shows, TV	After crossing	Unknown	Khosi	South		

Next steps: Generate code with netflix_1a View recommended plots

```
tvshows_week = netflix_1a[netflix_1a['type'] == 'TV Show']
tvshows_each_week=tvshows_week.groupby('week_added')['type'].count()
tvshows_each_week.sort_values(ascending=False)
```

```
week_added
35    862
31    812
27    794
53    726
24    703
26    636
13    626
18    603
5     600
44    582
48    580
40    542
19    528
15    499
22    486
46    480
50    477
51    471
37    469
33    448
1     444
49    429
52    418
38    414
7     411
12    402
36    400
32    399
21    398
14    395
23    392
42    383
20    370
9     362
11    356
8     354
25    350
29    349
45    345
39    342
17    335
30    321
34    313
47    299
4     289
10    289
2     285
41    276
3     243
6     240
28    236
43    224
16    220
Name: type, dtype: int64
```

In the above data we can look at the top 5 weeks when the tvshows was most released

I recommend focusing on releasing new content during weeks 35, 31, 27, 53, and 24, as these periods appear to be popular for launching TV shows and movies. By aligning releases with peak weeks, Netflix can maximize viewership and engagement.

```
movies_week = netflix_1a[netflix_1a['type'] == 'Movie']
movies_each_week=movies_week.groupby('week_added')['type'].count()
movies_each_week.sort_values(ascending=False)
```

```
week_added
1     3706
44    2561
35    2278
9     2192
40    2185
26    2131
31    1837
39    1672
27    1641
13    1635
18    1597
15    1581
48    1539
30    1462
22    1453
23    1435
5     1337
17    1254
28    1239
14    1200
29    1159
7     1158
43    1157
50    1156
36    1154
37    1133
25    1128
51    1099
33    1069
49    1058
10    1057
16    1054
11    1001
38     994
34     973
3      961
42     932
52     859
47     832
24     796
20     779
41     767
19     749
53     731
2      714
6      702
46     687
8      670
21     633
45     615
32     584
12     576
4      474
Name: type, dtype: int64
```

Since most of the movies are released in the first week of every year i recommend to drop more content on the 1st week of every year

```
week_summary = pd.merge(movies_each_week, tvshows_each_week, on='week_added', how='outer').fillna(0)
week_summary.rename(columns={'type_x' : 'movies', 'type_y':'tvshows'}).sort_values(by = 'movies', ascending = False)
```

week_added	movies	tvshows
1	3706	444
44	2561	582
35	2278	862
9	2192	362
40	2185	542
26	2131	636
31	1837	812
39	1672	342
27	1641	794
13	1635	626
18	1597	603
15	1581	499
48	1539	580
30	1462	321
22	1453	486
23	1435	392
5	1337	600
17	1254	335
28	1239	236
14	1200	395
29	1159	349
7	1158	411
43	1157	224
50	1156	477
36	1154	400
37	1133	469
25	1128	350
51	1099	471
33	1069	448
49	1058	429
10	1057	289
16	1054	220
11	1001	356
38	994	414
34	973	313
3	961	243
42	932	383
52	859	418
47	832	299
24	796	703
20	779	370
41	767	276
19	749	528
53	731	726
2	714	285

6	702	240
46	687	480
8	670	354
21	633	398
45	615	345
32	584	399
12	576	402
4	474	289

All over i would recommend droping new content on the most popular weeks since the user will be attentive and will be expecting something new to watch

```
netflix_1a['date_added'] = pd.to_datetime(netflix_1a['date_added'], errors='coerce')

netflix_1a['month_added'] = netflix_1a['date_added'].dt.month
netflix_1a.head()
```

	show_id	type	title	date_added	release_year	rating	listed_in	description	director	cast	country	movies_duration	tv
0	s1	Movie	Dick Johnson Is Dead	2021-09-25	2020	PG-13	Documentaries	As her father nears the end of his life, filmm...	Kirsten Johnson	Unknown Cast	United States	90.0	
1	s2	TV Show	Blood & Water	2021-09-24	2021	TV-MA	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	Unknown Director	Ama Qamata	South Africa	0.0	
		TV	Blood &				International TV Shows, TV	After crossing	Unknown	Khosi	South		

Next steps: Generate code with `netflix_1a` View recommended plots

```
tvshows_month = netflix_1a[netflix_1a['type'] == 'TV Show']
tvshows_each_month=tvshows_month.groupby('month_added')['type'].count()
tvshows_each_month.sort_values(ascending=False)
```

```
month_added
8.0     2260
12.0    2155
7.0     2140
6.0     2042
9.0     2017
1.0     1960
11.0    1877
4.0     1862
5.0     1827
3.0     1736
10.0    1709
2.0     1622
Name: type, dtype: int64
```

```
movies_month = netflix_1a[netflix_1a['type'] == 'Movie']
movies_each_month=movies_month.groupby('month_added')['type'].count()
movies_each_month.sort_values(ascending=False)
```

```
month_added
7.0     6749
1.0     6320
10.0    6196
9.0     5909
4.0     5879
12.0    5754
8.0     5211
6.0     5105
11.0    5073
3.0     4989
5.0     4207
```

```
2.0     3954
Name: type, dtype: int64
```

```
month_summary = pd.merge(movies_each_month, tvshows_each_month, on='month_added', how='outer').fillna(0)
month_summary.rename(columns={'type_x' : 'movies', 'type_y':'tvshows'}).sort_values(by = 'movies', ascending = False)
```

month_added	movies	tvshows
7.0	6749	2140
1.0	6320	1960
10.0	6196	1709
9.0	5909	2017
4.0	5879	1862
12.0	5754	2155
8.0	5211	2260
6.0	5105	2042
11.0	5073	1877
3.0	4989	1736
5.0	4207	1827
2.0	3954	1622

The 8th month of every year and the 7th month of every year has the highest tvshows, movies release .Hence it is recommend to drop new content during those months

```
#top 10 actors with most movies and tvshows
```

```
netflix_1a.head()
```

	show_id	type	title	date_added	release_year	rating	listed_in	description	director	cast	country	movies_duration	tv
0	s1	Movie	Dick Johnson Is Dead	2021-09-25	2020	PG-13	Documentaries	As her father nears the end of his life, filmm...	Kirsten Johnson	Unknown Cast	United States	90.0	
1	s2	TV Show	Blood & Water	2021-09-24	2021	TV-MA	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	Unknown Director	Ama Qamata	South Africa	0.0	
		TV	Blood &				International TV Shows, TV	After crossing	Unknown	Khosi	South		

Next steps: Generate code with `netflix_1a` View recommended plots

```
tv_actors= netflix_1a[(netflix_1a['type'] == 'TV Show')&(netflix_1a['cast'] != 'Unknown Cast')]
top_tv_actors=tv_actors.groupby('cast')['title'].nunique().reset_index()
top_tv_actors.sort_values(by = 'title', ascending = False).reset_index().drop(columns =['index']).head(10)
```

	cast	title
0	Takahiro Sakurai	25
1	Yuki Kaji	19
2	Junichi Suwabe	17
3	Daisuke Ono	17
4	Ai Kayano	17
5	Yuichi Nakamura	16
6	Yoshimasa Hosoya	15
7	Jun Fukuyama	15
8	David Attenborough	14
9	Kana Hanazawa	13

```
mv_actors= netflix_1a[(netflix_1a['type'] == 'Movie')&(netflix_1a['cast'] != 'Unknown Cast')]
top_mv_actors=mv_actors.groupby('cast')['title'].nunique().reset_index()
top_mv_actors.sort_values(by = 'title', ascending = False).reset_index().drop(columns =['index']).head(10)
```

	cast	title
0	Anupam Kher	42
1	Shah Rukh Khan	35
2	Naseeruddin Shah	32
3	Akshay Kumar	30
4	Om Puri	30
5	Paresh Rawal	28
6	Amitabh Bachchan	28
7	Julie Tejwani	28
8	Boman Irani	27
9	Rupa Bhimani	27

```
tv_directors= netflix_1a[(netflix_1a['type'] == 'TV Show')&(netflix_1a['director'] != 'Unknown Director')]
top_tv_directors=tv_directors.groupby('director')['title'].nunique().reset_index()
top_tv_directors.sort_values(by = 'title', ascending = False).reset_index().drop(columns =['index']).head(10)
```

	director	title
0	Ken Burns	3
1	Alastair Fothergill	3
2	Stan Lathan	2
3	Jung-ah Im	2
4	Joe Berlinger	2
5	Hsu Fu-chun	2
6	Gautham Vasudev Menon	2
7	Lynn Novick	2
8	Iginio Straffi	2
9	Shin Won-ho	2

```
mv_directors= netflix_1a[(netflix_1a['type'] == 'Movie')&(netflix_1a['director'] != 'Unknown Director')]
top_mv_directors=mv_directors.groupby('director')['title'].nunique().reset_index()
top_mv_directors.sort_values(by = 'title', ascending = False).reset_index().drop(columns =['index']).head(10)
```

	director	title
0	Rajiv Chilaka	22
1	Jan Suter	21
2	Raúl Campos	19
3	Suhas Kadav	16
4	Marcus Raboy	15
5	Jay Karas	15
6	Cathy Garcia-Molina	13
7	Jay Chapman	12
8	Martin Scorsese	12
9	Youssef Chahine	12

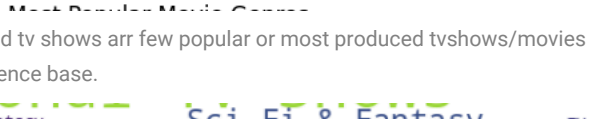
In the above datas we can clearly see all the top actors and directors

collaborate with top actors for exclusive projects. High-profile actors can attract a larger audience and generate buzz around new releases. top actors will have huge fan bases

directors Top directors bring unique storytelling skills and cinematic expertise, elevating the overall quality and appeal of Netflix originals.

Many top actors and directors have international appeal, helping Netflix expand its global presence and reach diverse audiences.

```
#55. Which genre movies are more popular or produced more

genres = netflix_1a['listed_in'].dropna().str.split(', ')

from collections import Counter

genre_counts = Counter([genre for sublist in genres for genre in sublist])


from wordcloud import WordCloud
import matplotlib.pyplot as plt


wordcloud = WordCloud(width=800, height=400, background_color='white').generate_from_frequencies(genre_counts)


plt.figure(figsize=(10, 6))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.title('Most Popular Movie Genres')
plt.show()
```



We see that Dramas, international movies and tv shows arr few popular or most produced tvshows/movies Produce more content in genres that are highly popular and have a large audience base.

#6. Find After how many days the movie will be added to Netflix after the release of the movie

```
netflix_1a['date_added'] = pd.to_datetime(netflix_1a['date_added'])
netflix_1a['release_year'] = pd.to_numeric(netflix_1a['release_year'], errors='coerce')

netflix_1a['days_diff'] = (netflix_1a['date_added'] - pd.to_datetime(netflix_1a['release_year'], format='%Y')).dt.days

netflix_1a.head()
```

	show_id	type	title	date_added	release_year	rating	listed_in	description	director	cast	country	movies_duration	tv
0	s1	Movie	Dick Johnson Is Dead	2021-09-25	2020	PG-13	Documentaries	As her father nears the end of his life, filmm...	Kirsten Johnson	Unknown Cast	United States	90.0	
1	s2	TV Show	Blood & Water	2021-09-24	2021	TV-MA	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...	Unknown Director	Ama Qamata	South Africa	0.0	
		TV	Blood &				International TV Shows, TV	After crossing	Unknown	Khosi	South		

Next steps: Generate code with netflix_1a View recommended plots

```
netflix_1a[['title', 'release_year', 'date_added', 'days_diff']]
```

	title	release_year	date_added	days_diff
0	Dick Johnson Is Dead	2020	2021-09-25	633.0
1	Blood & Water	2021	2021-09-24	266.0
2	Blood & Water	2021	2021-09-24	266.0