```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline

df = pd.read_csv('C:\\Users\\JAISON ABISHEK\\Downloads\\
disney_plus_titles.csv')

df.head()
```

```
  show_id     type                              title  \
0      s1    Movie                    A Spark Story
1      s2    Movie                   Spooky Buddies
2      s3    Movie          The Fault in Our Stars
3      s4  TV Show                 Dog: Impossible
4      s5  TV Show  Spidey And His Amazing Friends

                      director  \
0  Jason Sterman, Leanne Dare
1                Robert Vince
2                  Josh Boone
3                         NaN
4                         NaN

                                                 cast
country  \
0                      Apthon Corbin, Louis Gonzales
NaN
1  Tucker Albrizzi, Diedrich Bader, Ameko Eks Mas...  United States,
Canada
2  Shailene Woodley, Ansel Elgort, Laura Dern, Sa...           United
States
3                                     Matt Beisner           United
States
4  Benjamin Valic, Lily Sanfelippo, Jakari Fraser...           United
States

           date_added  release_year rating   duration  \
0  September 24, 2021          2021  TV-PG     88 min
1  September 24, 2021          2011      G     93 min
2  September 24, 2021          2014  PG-13    127 min
3  September 22, 2021          2019  TV-PG  2 Seasons
4  September 22, 2021          2021   TV-Y   1 Season

                              listed_in  \
0                          Documentary
1                 Comedy, Fantasy, Kids
2         Coming of Age, Drama, Romance
3  Animals & Nature, Docuseries, Family
```

```
4     Action-Adventure, Animation, Kids

                                          description
0  Two Pixar filmmakers strive to bring their uni...
1  The puppies go on a spooky adventure through a...
2  Hazel and Gus share a love that sweeps them on...
3  Matt Beisner uses unique approaches to modifyi...
4  Spidey teams up with pals to become The Spidey...

df.tail()

     show_id     type                    title             director  \
1363   s1364    Movie  The Sword in the Stone  Wolfgang Reitherman
1364   s1365    Movie         Those Calloways         Norman Tokar
1365   s1366  TV Show    Disney Kirby Buckets                  NaN
1366   s1367  TV Show          Disney Mech-X4                  NaN
1367   s1368  TV Show      Imagination Movers                  NaN

                                                  cast        country
\
1363  Sebastian Cabot, Karl Swenson, Rickie Sorensen...  United States

1364  Brian Keith, Vera Miles, Brandon de Wilde, Wal...  United States

1365  Jacob Bertrand, Mekai Curtis, Cade Sutton, Oli...  United States

1366  Nathaniel Potvin, Raymond Cham, Kamran Lucas, ...         Canada

1367  Rich Collins, Dave Poche, Scott Durbin, Scott ...  United States


          date_added  release_year rating   duration  \
1363  October 1, 2019          1963      G     80 min
1364  October 1, 2019          1965     PG    132 min
1365              NaN          2014  TV-Y7  3 Seasons
1366              NaN          2016  TV-Y7  2 Seasons
1367              NaN          2008   TV-Y  3 Seasons

                                     listed_in  \
1363       Action-Adventure, Animation, Comedy
1364           Animals & Nature, Drama, Family
1365    Action-Adventure, Comedy, Coming of Age
1366  Action-Adventure, Comedy, Science Fiction
1367                              Kids, Music

                                          description
1363  Merlin trains a young orphan who's destined to...
1364  A strong-willed family struggles to establish ...
1365   Welcome to Kirby's world! It's rude and sketchy.
1366  Ryan discovers his ability to control a giant ...
1367  Rock out with the Imagination Movers, Disney J...
```

```
df.columns.values
```

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

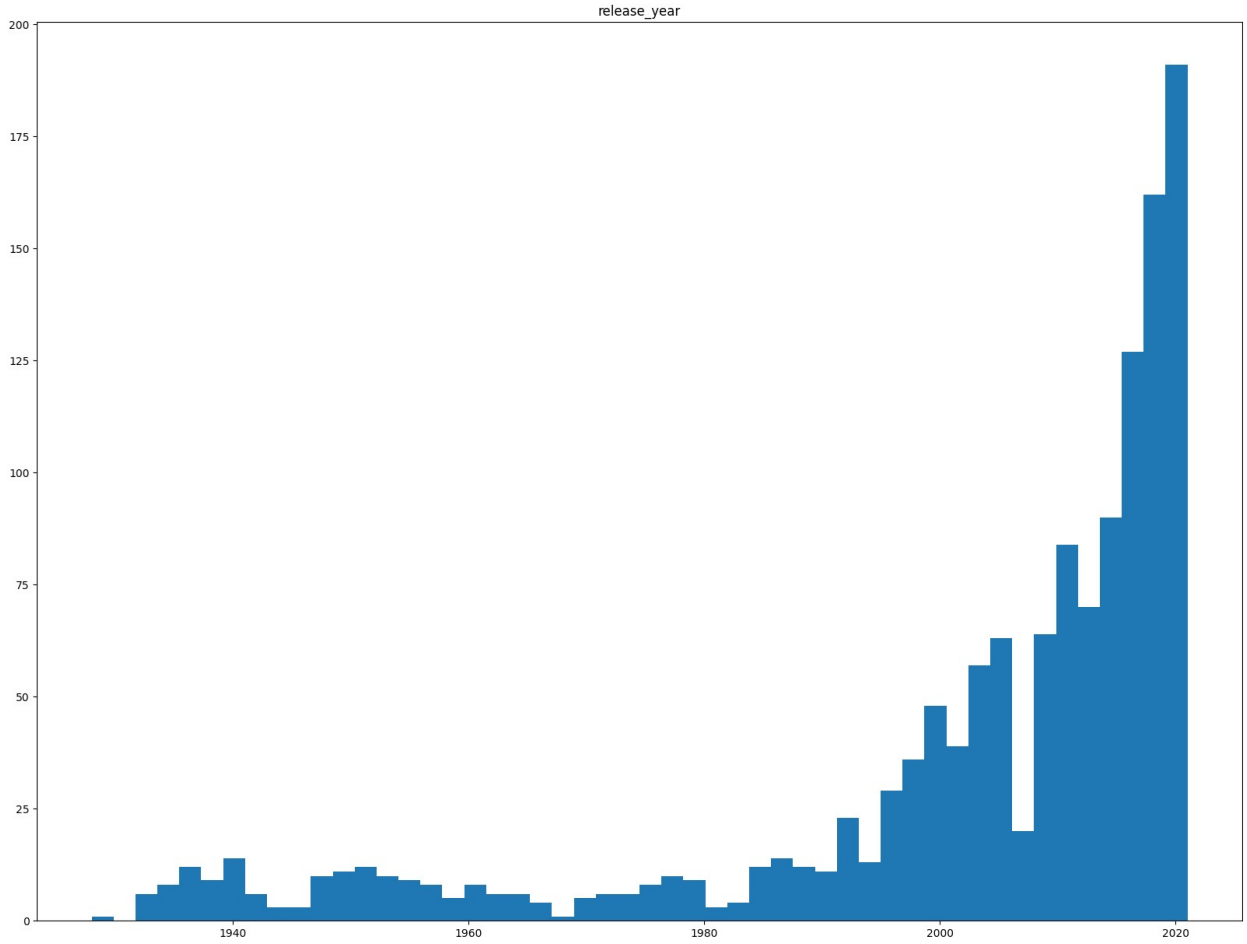
```
df.isna().sum()
```

```
show_id           0
type              0
title             0
director        440
cast            174
country         175
date_added        3
release_year      0
rating            2
duration          0
listed_in         0
description       0
dtype: int64
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1368 entries, 0 to 1367
Data columns (total 12 columns):
 #   Column        Non-Null Count  Dtype 
---  ------        --------------  ----- 
 0   show_id       1368 non-null   object
 1   type          1368 non-null   object
 2   title         1368 non-null   object
 3   director      928 non-null    object
 4   cast          1194 non-null   object
 5   country       1193 non-null   object
 6   date_added    1365 non-null   object
 7   release_year  1368 non-null   int64 
 8   rating        1366 non-null   object
 9   duration      1368 non-null   object
 10  listed_in     1368 non-null   object
 11  description   1368 non-null   object
dtypes: int64(1), object(11)
memory usage: 128.4+ KB
```

```
df.hist(bins = 50, grid = False, figsize=(20,15));
```

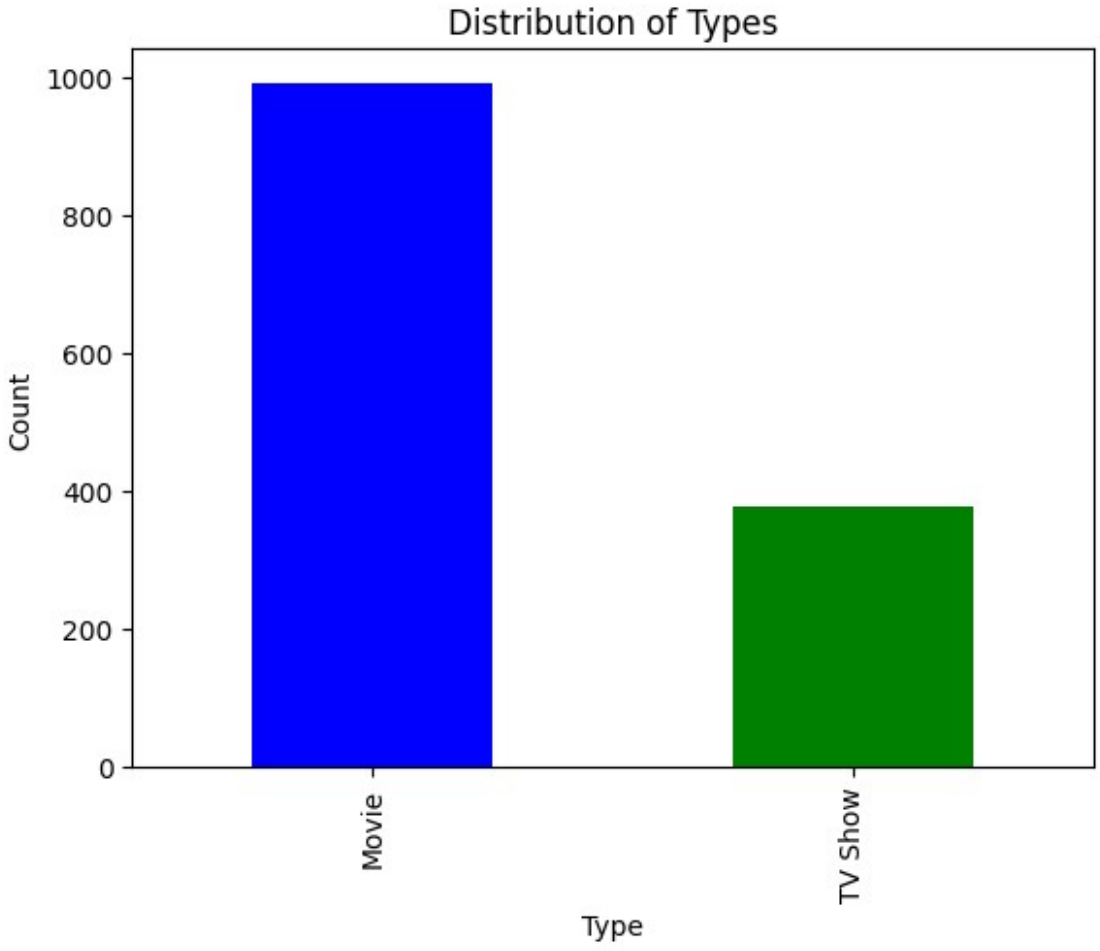



```
df.describe()
```

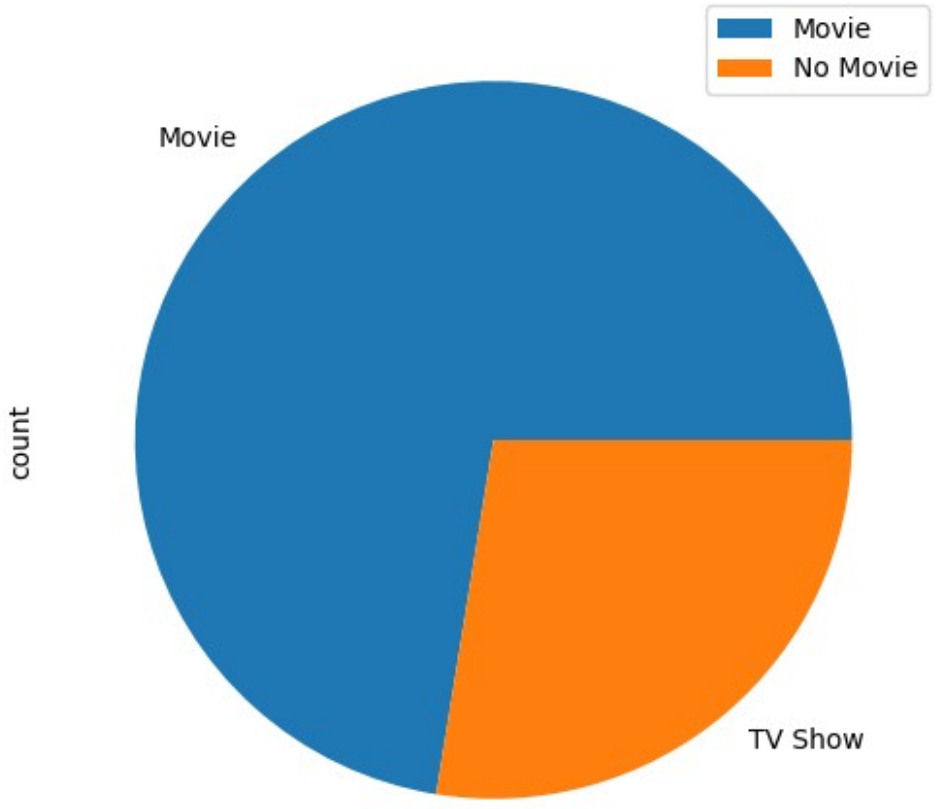
	release_year
count	1368.000000
mean	2002.348684
std	22.127559
min	1928.000000
25%	1998.000000
50%	2011.000000
75%	2018.000000
max	2021.000000

```
questions = [
    "1. How many shows were added in each year?",
    "2. Which country has the most shows in the dataset?",
    "3. Which director has directed the most shows?",
    "4. What are the different ratings given to shows, and how many
shows have each rating?",
]
questions
```

```
['1. How many shows were added in each year?',
 '2. Which country has the most shows in the dataset?',
 '3. Which director has directed the most shows?',
 '4. What are the different ratings given to shows, and how many shows
have each rating?']

df.type.value_counts()

type
Movie      991
TV Show    377
Name: count, dtype: int64

if 'type' in df.columns:
    df['type'].value_counts().plot(kind='bar', color=["blue",
"green"])
    plt.title("Distribution of Types")
    plt.xlabel("Type")
    plt.ylabel("Count")
    plt.show()
```



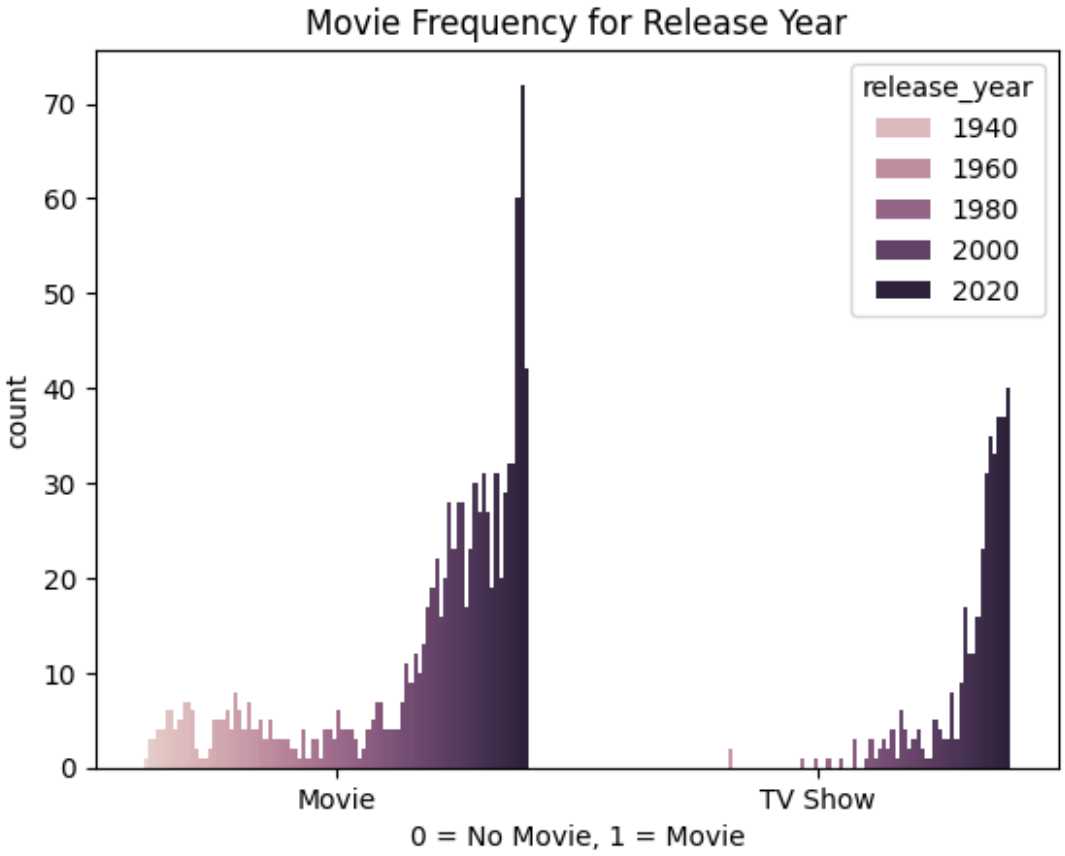
```
df.type.value_counts().plot(kind = 'pie', figsize = (8,6))
plt.legend(["Movie", "No Movie"]);
```




```
df.type.value_counts()

type
Movie      991
TV Show    377
Name: count, dtype: int64

pd.crosstab(df.title, df.type)

type                                          Movie  TV Show
title
10 Things I Hate About You                        1        0
101 Dalmatian Street                              0        1
101 Dalmatians                                    1        0
101 Dalmatians (Series)                           0        1
101 Dalmatians II: Patch's London Adventure       1        0
...                                             ...      ...
Zenimation                                        0        1
```

```
Zenon: Girl of the 21st Century                    1         0
Zenon: The Zequel                                  1         0
Zenon: Z3                                          1         0
Zootopia                                           1         0

[1368 rows x 2 columns]
```

```
sns.countplot(x = 'type', data = df, hue = 'release_year')
plt.title("Movie Frequency for Release Year")
plt.xlabel("0 = No Movie, 1 = Movie");
```

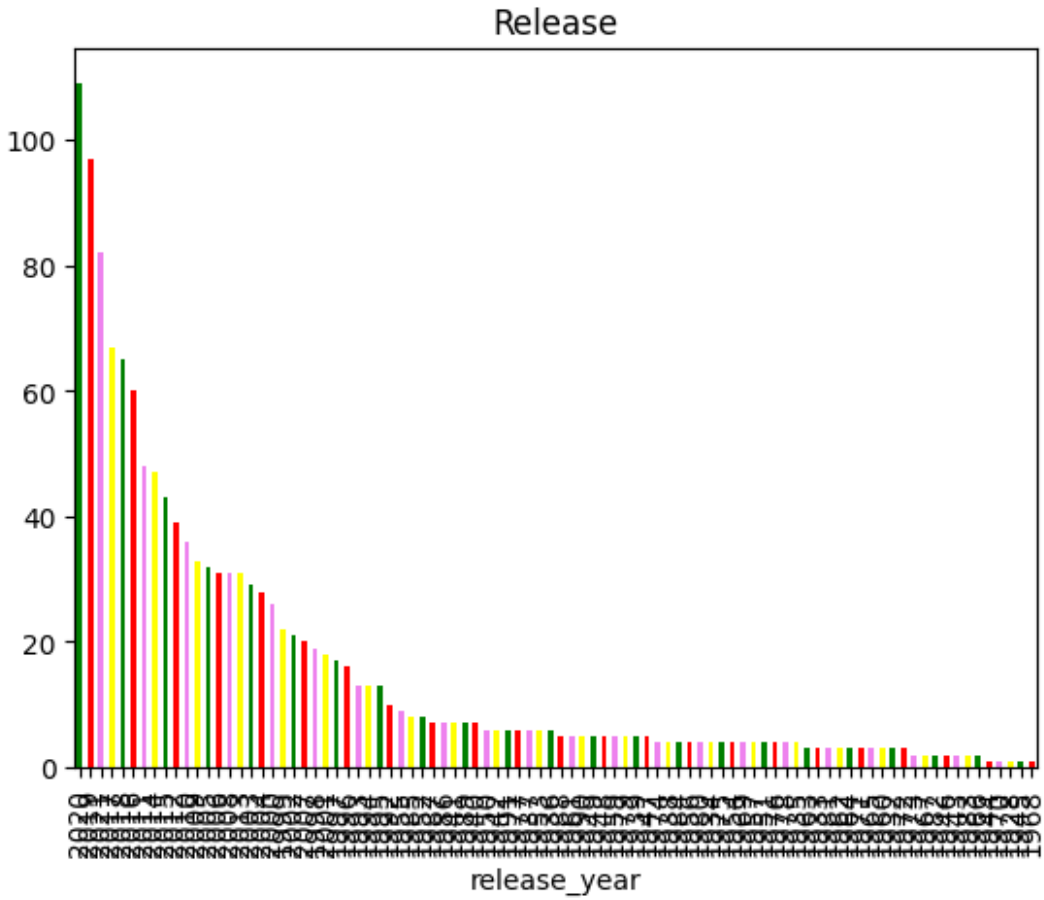



```
df.release_year.value_counts()
```

```
release_year
2020    109
2019     97
2021     82
2017     67
2018     65
       ...
1944      1
1970      1
1928      1
```

```
1945      1
1968      1
Name: count, Length: 90, dtype: int64
```

```
df.release_year.value_counts().plot(kind = 'bar',
color=["green","red","violet","yellow"])
plt.title("Release");
```



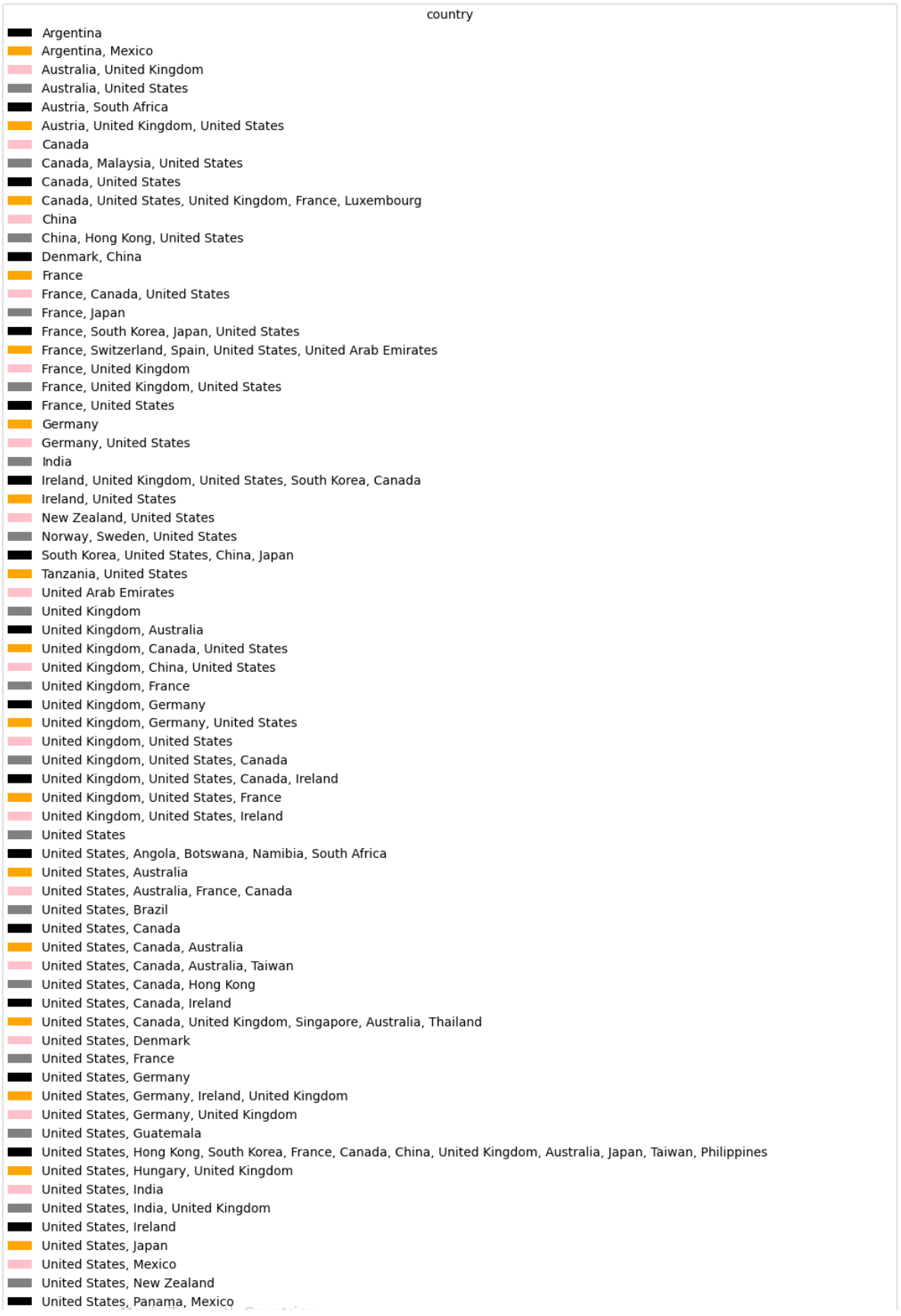

```
pd.crosstab(df.type, df.country)
```

```
country  Argentina  Argentina, Mexico  Australia, United Kingdom  \
type
Movie            0                  0                          0
TV Show          2                  1                          1

country  Australia, United States  Austria, South Africa  \
type
Movie                           3                      1
TV Show                         0                      0

country  Austria, United Kingdom, United States  Canada  \
type
```

```
Movie                                                 1       2
TV Show                                               0       3

country  Canada, Malaysia, United States  Canada, United States  \
type
Movie                                        1                     16
TV Show                                      0                      1

country  Canada, United States, United Kingdom, France,
Luxembourg  ...  \
type                                                                 ..
.
Movie
0           ...
TV Show
1           ...

country  United States, Taiwan, China  \
type
Movie                                  1
TV Show                                0

country  United States, Taiwan, South Korea, China, Japan, United
Kingdom  \
type

Movie                                                          0

TV Show                                                        1


country  United States, United Kingdom  \
type
Movie                                  18
TV Show                                 1

country  United States, United Kingdom, Australia  \
type
Movie                                                    4
TV Show                                                  0

country  United States, United Kingdom, Australia, Canada  \
type
Movie                                                         1
TV Show                                                       0

country  United States, United Kingdom, Belgium, Canada  \
type
Movie                                                       1
TV Show                                                     0
```

```
country  United States, United Kingdom, Canada  \
type
Movie                                          3
TV Show                                        0

country  United States, United Kingdom, Hungary  \
type
Movie                                           1
TV Show                                         0

country  United States, United Kingdom, South Africa  \
type
Movie                                                  1
TV Show                                                0

country  United States, United Kingdom, South Korea
type
Movie                                                 0
TV Show                                               1

[2 rows x 87 columns]
```

```
pd.crosstab(df.type,df.country).plot(kind = 'bar',
color=["black","orange","pink","grey"])
plt.title('Movie Type with Countries')
plt.xlabel('0 = Movie, 1 = Tv Show');
```


country
Argentina
Argentina, Mexico
Australia, United Kingdom
Australia, United States
Austria, South Africa
Austria, United Kingdom, United States
Canada
Canada, Malaysia, United States
Canada, United States
Canada, United States, United Kingdom, France, Luxembourg
China
China, Hong Kong, United States
Denmark, China
France
France, Canada, United States
France, Japan
France, South Korea, Japan, United States
France, Switzerland, Spain, United States, United Arab Emirates
France, United Kingdom
France, United Kingdom, United States
France, United States
Germany
Germany, United States
India
Ireland, United Kingdom, United States, South Korea, Canada
Ireland, United States
New Zealand, United States
Norway, Sweden, United States
South Korea, United States, China, Japan
Tanzania, United States
United Arab Emirates
United Kingdom
United Kingdom, Australia
United Kingdom, Canada, United States
United Kingdom, China, United States
United Kingdom, France
United Kingdom, Germany
United Kingdom, Germany, United States
United Kingdom, United States
United Kingdom, United States, Canada
United Kingdom, United States, Canada, Ireland
United Kingdom, United States, France
United Kingdom, United States, Ireland
United States
United States, Angola, Botswana, Namibia, South Africa
United States, Australia
United States, Australia, France, Canada
United States, Brazil
United States, Canada
United States, Canada, Australia
United States, Canada, Australia, Taiwan
United States, Canada, Hong Kong
United States, Canada, Ireland
United States, Canada, United Kingdom, Singapore, Australia, Thailand
United States, Denmark
United States, France
United States, Germany
United States, Germany, Ireland, United Kingdom
United States, Germany, United Kingdom
United States, Guatemala
United States, Hong Kong, South Korea, France, Canada, China, United Kingdom, Australia, Japan, Taiwan, Philippines
United States, Hungary, United Kingdom
United States, India
United States, India, United Kingdom
United States, Ireland
United States, Japan
United States, Mexico
United States, New Zealand
United States, Panama, Mexico

```
pd.crosstab(df.country, df.type)
```

type	Movie	TV Show
country		
Argentina	0	2
Argentina, Mexico	0	1
Australia, United Kingdom	0	1
Australia, United States	3	0
Austria, South Africa	1	0
...	...	...
United States, United Kingdom, Belgium, Canada	1	0
United States, United Kingdom, Canada	3	0
United States, United Kingdom, Hungary	1	0
United States, United Kingdom, South Africa	1	0
United States, United Kingdom, South Korea	0	1

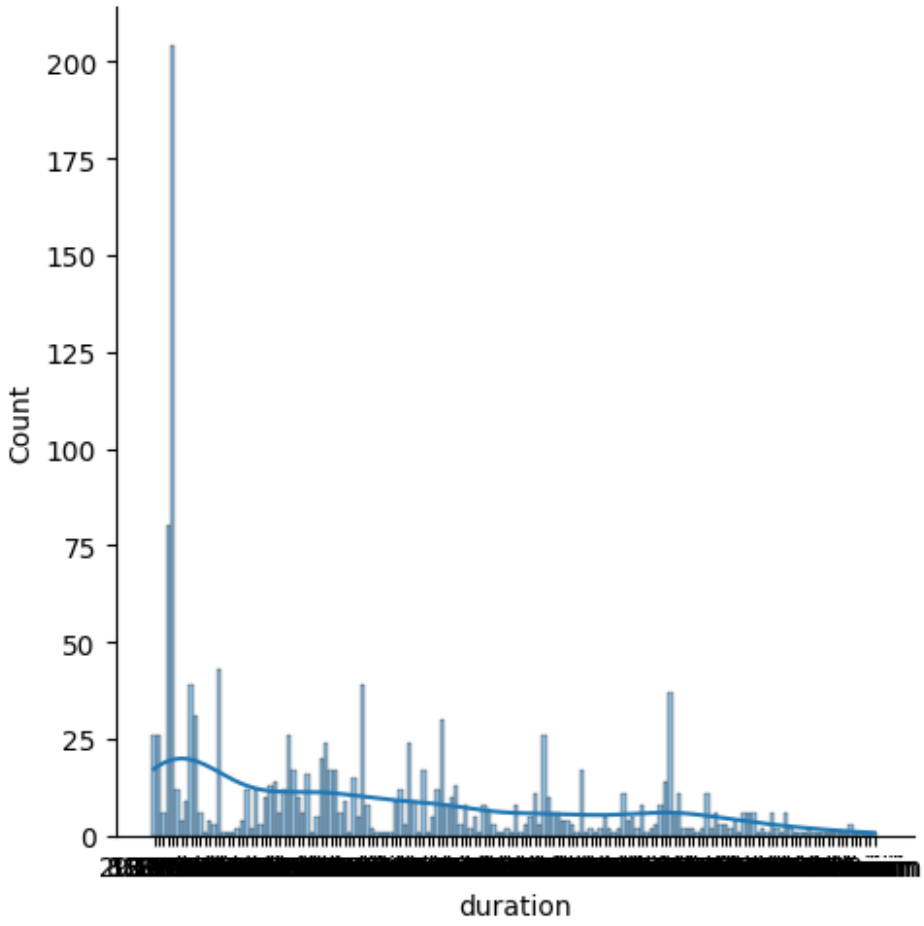
```
[87 rows x 2 columns]
```

```
sns.countplot(x = 'rating', data = df, hue="duration")
```

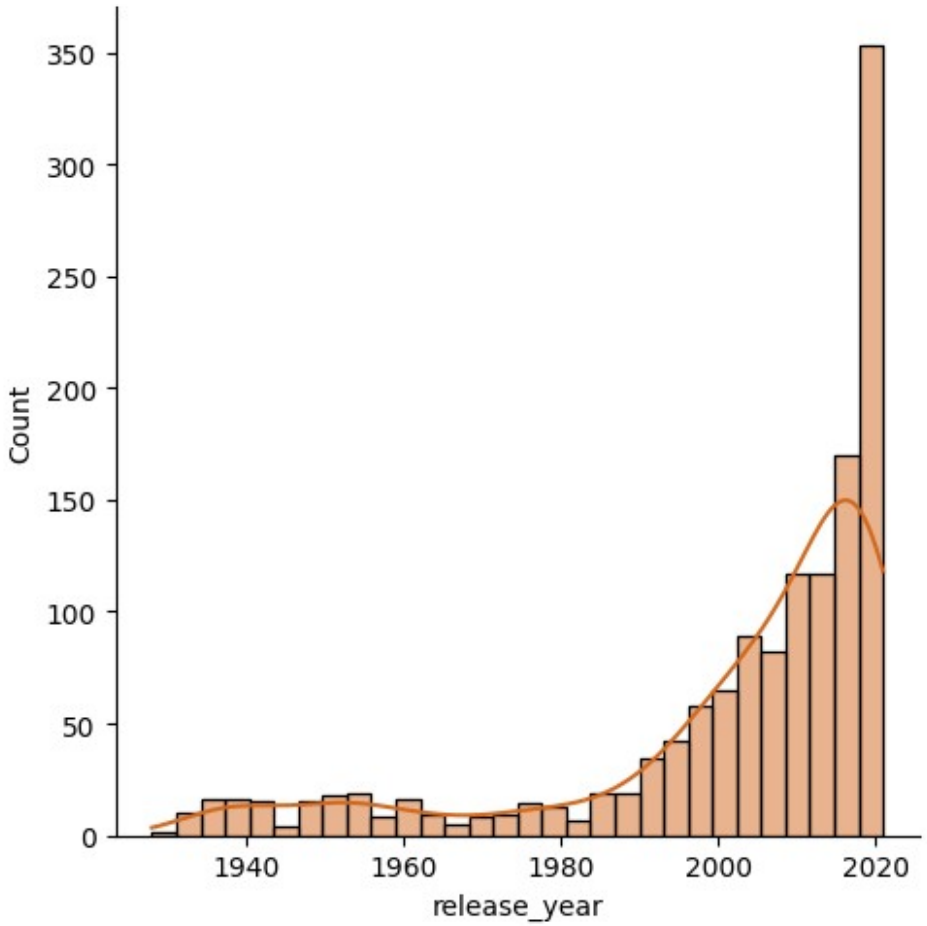
```
<Axes: xlabel='rating', ylabel='count'>
```


duration
88 min
93 min
127 min
2 Seasons
1 Season
106 min
22 min
116 min
44 min
9 min
42 min
16 Seasons
11 min
140 min
3 Seasons
19 Seasons
66 min
131 min
115 min
137 min
99 min
14 min
85 min
9 Seasons
104 min
105 min
92 min
126 min
82 min
94 min
91 min
87 min
125 min
95 min
54 min
7 Seasons
6 min
4 Seasons
101 min
98 min
102 min

```
sns.displot(x = 'duration', data = df, bins = 30, kde = True);
```




```
sns.displot(x = 'release_year', data = df, bins = 30, kde = True,
color = 'chocolate');
```


350
300
250
200
150
100
50
0
Count
1940
1960
1980
2000
2020
release_year