

# Data Analysis using Python - Task6

May 5, 2024

```
[2]: # Importing all the libraries that we need
```

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

```
[3]: # Importing our dataset
```

```
df = pd.read_csv("C:\\Users\\SANDEEP\\Downloads\\disney_plus_titles.csv")
```

```
[4]: # Checking first five rows by calling df.head()
```

```
df.head()
```

```
[4]:  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  Apton 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...

```

```
[5]: df.tail()
```

```

[5]:   show_id  type  title  director \
1363  s1364  Movie  The Sword in the Stone  Wolfgang Reitherman
1364  s1365  Movie    Those Callows         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...

```

```
[6]: # Take a look at the column names
df.columns.values
```

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

```
[8]: df.shape
```

```
[8]: (1368, 12)
```

```
[9]: df = df.drop_duplicates()
df.shape
```

```
[9]: (1368, 12)
```

```
[10]: 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

```

```
[11]: columns_to_remove = ['description']
df = df.drop(columns = columns_to_remove)
df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1368 entries, 0 to 1367
Data columns (total 11 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
dtypes: int64(1), object(10)
memory usage: 117.7+ KB

```

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

```
[12]: 794
```

```

[13]: # Dropping the duplicate values from our data
df = df.drop_duplicates()
df

```

```

[13]:
   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
...      ...      ...
1363  s1364  Movie  The Sword in the Stone
1364  s1365  Movie  Those Calloways
1365  s1366  TV Show  Disney Kirby Buckets
1366  s1367  TV Show  Disney Mech-X4
1367  s1368  TV Show  Imagination Movers

   director \
0  Jason Sterman, Leanne Dare
1          Robert Vince
2          Josh Boone
3              NaN
4              NaN
...
1363  Wolfgang Reitherman

```

1364	Norman Tokar
1365	NaN
1366	NaN
1367	NaN

		cast \
0	Apton Corbin, Louis Gonzales	
1	Tucker Albrizzi, Diedrich Bader, Ameko Eks Mas...	
2	Shailene Woodley, Ansel Elgort, Laura Dern, Sa...	
3	Matt Beisner	
4	Benjamin Valic, Lily Sanfelippo, Jakari Fraser...	
...	...	
1363	Sebastian Cabot, Karl Swenson, Rickie Sorensen...	
1364	Brian Keith, Vera Miles, Brandon de Wilde, Wal...	
1365	Jacob Bertrand, Mekai Curtis, Cade Sutton, Oli...	
1366	Nathaniel Potvin, Raymond Cham, Kamran Lucas, ...	
1367	Rich Collins, Dave Poche, Scott Durbin, Scott ...	

	country	date_added	release_year	rating \
0	NaN	September 24, 2021	2021	TV-PG
1	United States, Canada	September 24, 2021	2011	G
2	United States	September 24, 2021	2014	PG-13
3	United States	September 22, 2021	2019	TV-PG
4	United States	September 22, 2021	2021	TV-Y
...	...	...	...	...
1363	United States	October 1, 2019	1963	G
1364	United States	October 1, 2019	1965	PG
1365	United States	NaN	2014	TV-Y7
1366	Canada	NaN	2016	TV-Y7
1367	United States	NaN	2008	TV-Y

	duration	listed_in
0	88 min	Documentary
1	93 min	Comedy, Fantasy, Kids
2	127 min	Coming of Age, Drama, Romance
3	2 Seasons	Animals & Nature, Docuseries, Family
4	1 Season	Action-Adventure, Animation, Kids
...	...	...
1363	80 min	Action-Adventure, Animation, Comedy
1364	132 min	Animals & Nature, Drama, Family
1365	3 Seasons	Action-Adventure, Comedy, Coming of Age
1366	2 Seasons	Action-Adventure, Comedy, Science Fiction
1367	3 Seasons	Kids, Music

[1368 rows x 11 columns]

```
[14]: # Replacing the null values with 0
df1 = df.fillna(value = 0)
df1
```

```
[14]:
```

	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
...	...	...	...
1363	s1364	Movie	The Sword in the Stone
1364	s1365	Movie	Those Calloways
1365	s1366	TV Show	Disney Kirby Buckets
1366	s1367	TV Show	Disney Mech-X4
1367	s1368	TV Show	Imagination Movers

	director \
0	Jason Sterman, Leanne Dare
1	Robert Vince
2	Josh Boone
3	0
4	0
...	...
1363	Wolfgang Reitherman
1364	Norman Tokar
1365	0
1366	0
1367	0

	cast \
0	Apthon Corbin, Louis Gonzales
1	Tucker Albrizzi, Diedrich Bader, Ameko Eks Mas...
2	Shailene Woodley, Ansel Elgort, Laura Dern, Sa...
3	Matt Beisner
4	Benjamin Valic, Lily Sanfelippo, Jakari Fraser...
...	...
1363	Sebastian Cabot, Karl Swenson, Rickie Sorensen...
1364	Brian Keith, Vera Miles, Brandon de Wilde, Wal...
1365	Jacob Bertrand, Mekai Curtis, Cade Sutton, Oli...
1366	Nathaniel Potvin, Raymond Cham, Kamran Lucas, ...
1367	Rich Collins, Dave Poche, Scott Durbin, Scott ...

	country	date_added	release_year	rating \
0	0	September 24, 2021	2021	TV-PG
1	United States, Canada	September 24, 2021	2011	G
2	United States	September 24, 2021	2014	PG-13

3	United States	September 22, 2021	2019	TV-PG
4	United States	September 22, 2021	2021	TV-Y
...	...	...	...	...
1363	United States	October 1, 2019	1963	G
1364	United States	October 1, 2019	1965	PG
1365	United States	0	2014	TV-Y7
1366	Canada	0	2016	TV-Y7
1367	United States	0	2008	TV-Y

	duration	listed_in
0	88 min	Documentary
1	93 min	Comedy, Fantasy, Kids
2	127 min	Coming of Age, Drama, Romance
3	2 Seasons	Animals & Nature, Docuseries, Family
4	1 Season	Action-Adventure, Animation, Kids
...	...	...
1363	80 min	Action-Adventure, Animation, Comedy
1364	132 min	Animals & Nature, Drama, Family
1365	3 Seasons	Action-Adventure, Comedy, Coming of Age
1366	2 Seasons	Action-Adventure, Comedy, Science Fiction
1367	3 Seasons	Kids, Music

[1368 rows x 11 columns]

```
[15]: # We can use dropna() to remove all the rows with missing data.
df2 = df.dropna()
df2
```

```
[15]: show_id  type  title \
1      s2  Movie  Spooky Buddies
2      s3  Movie  The Fault in Our Stars
6      s7  Movie  Confessions of a Shopaholic
18     s19  Movie  Pirates of the Caribbean: On Stranger Tides
19     s20  Movie  Twenty Something
...     ...    ...    ...
1359  s1360  Movie  Lady and the Tramp II: Scamp's Adventure
1361  s1362  Movie  The Cat from Outer Space
1362  s1363  Movie  The Great Mouse Detective
1363  s1364  Movie  The Sword in the Stone
1364  s1365  Movie  Those Calloways

      director \
1      Robert Vince
2      Josh Boone
6      P.J. Hogan
18     Rob Marshall
19     Apton Corbin
```

```

...
1359          Darrell Rooney
1361          Norman Tokar
1362 John Musker, Ron Clements, Dave Michener
1363          Wolfgang Reitherman
1364          Norman Tokar

```

```

cast \
1   Tucker Albrizzi, Diedrich Bader, Ameko Eks Mas...
2   Shailene Woodley, Ansel Elgort, Laura Dern, Sa...
6   Isla Fisher, Hugh Dancy, Joan Cusack, John Goo...
18  Johnny Depp, Penélope Cruz, Ian McShane, Geoff...
19  Kaylin Price, Ariana Brown, Aliyah Taylor, Jan...

```

```

...
1359 Scott Wolf, Alyssa Milano, Chazz Palminteri, J...
1361 Ken Berry, Sandy Duncan, Harry Morgan, Roddy M...
1362 Vincent Price, Barrie Ingham, Val Bettin, Susa...
1363 Sebastian Cabot, Karl Swenson, Rickie Sorensen...
1364 Brian Keith, Vera Miles, Brandon de Wilde, Wal...

```

```

country      date_added  release_year  rating \
1   United States, Canada  September 24, 2021      2011      G
2   United States          September 24, 2021      2014     PG-13
6   United States          September 17, 2021      2009      PG
18  United States, United Kingdom  September 10, 2021      2011     PG-13
19  United States          September 10, 2021      2021      PG

...
1359 United States, Australia  October 1, 2019      2001      G
1361 United States          October 1, 2019      1978      G
1362 United States          October 1, 2019      1986      G
1363 United States          October 1, 2019      1963      G
1364 United States          October 1, 2019      1965     PG

```

```

duration      listed_in
1   93 min      Comedy, Fantasy, Kids
2   127 min     Coming of Age, Drama, Romance
6   106 min     Comedy, Romance, Romantic Comedy
18  140 min     Action-Adventure, Fantasy
19  11 min      Animation, Family

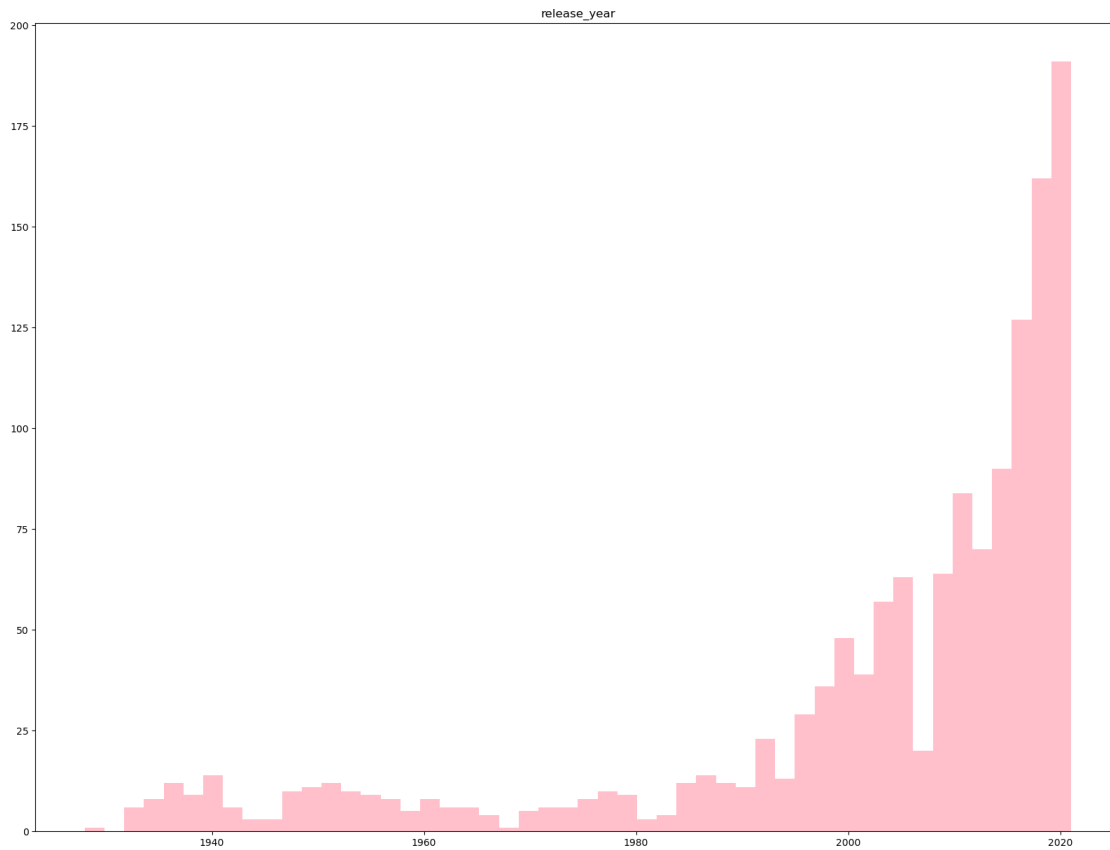
...
1359 73 min     Action-Adventure, Animation, Coming of Age
1361 104 min    Comedy, Family, Science Fiction
1362 75 min     Animation, Family, Mystery
1363 80 min     Action-Adventure, Animation, Comedy
1364 132 min    Animals & Nature, Drama, Family

```

[788 rows x 11 columns]



```
[14]: # plotting histogram of all numeric values.
df.hist(bins = 50, grid = False, figsize = (20, 15), color = ['pink']);
```



```
[17]: # Generating descriptive statistics.
df.describe()
```

```
[17]:      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
```

```
[21]: questions = ["1. How many Movie type in all release_year and How many TV Show_
↳type in all release_year?",
                  "2. How many Movie, TV Show types released by their respective_
↳directors?",
```

```
"3. Type of which dates added in disney plus?",
"4. Which type of country release in disney plus?",
"5. Type of which rating in disney plus?",
"6. How much duration for each type in disney plus?",
"7. What are different type of genres present in disney plus?"]
```

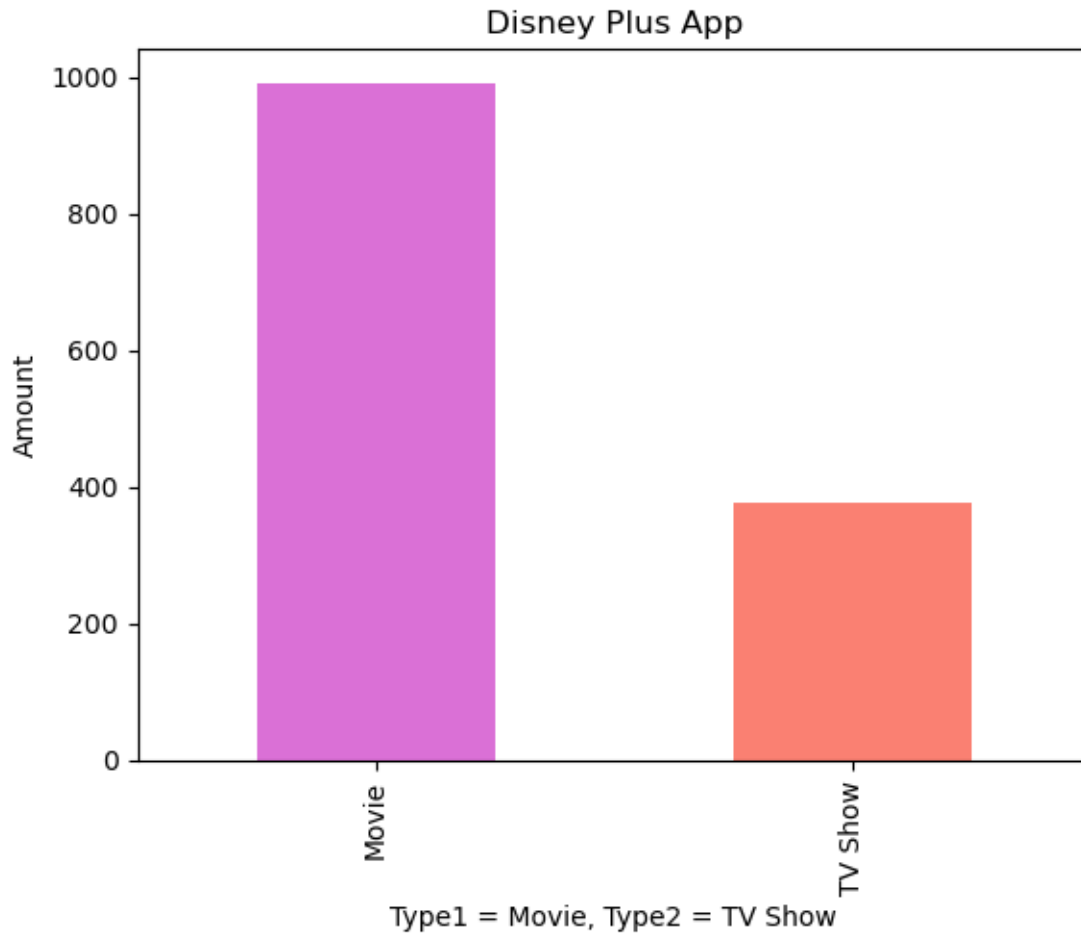
questions

```
[21]: ['1. How many Movie type in all release_year and How many TV Show type in all
release_year?',
'2. How many Movie, TV Show types released by their respective directors?',
'3. Type of which dates added in disney plus?',
'4. Which type of country release in disney plus?',
'5. Type of which rating in disney plus?',
'6. How much duration for each type in disney plus?',
'7. What are different type of genres present in disney plus?']
```

```
[8]: # Let's find the answer of first question.
# 1. How many Movie type in all release_year and How many TV Show type in all
↳release_year?
# Getting the values
df.type.value_counts()
```

```
[8]: type
Movie      991
TV Show    377
Name: count, dtype: int64
```

```
[18]: # Plotting bar chart
df.type.value_counts().plot(kind = 'bar', color = ["orchid", "salmon"])
plt.title("Disney Plus App")
plt.xlabel("Type1 = Movie, Type2 = TV Show")
plt.ylabel("Amount");
```



```
[29]: pd.crosstab(df.type, df.release_year)
```

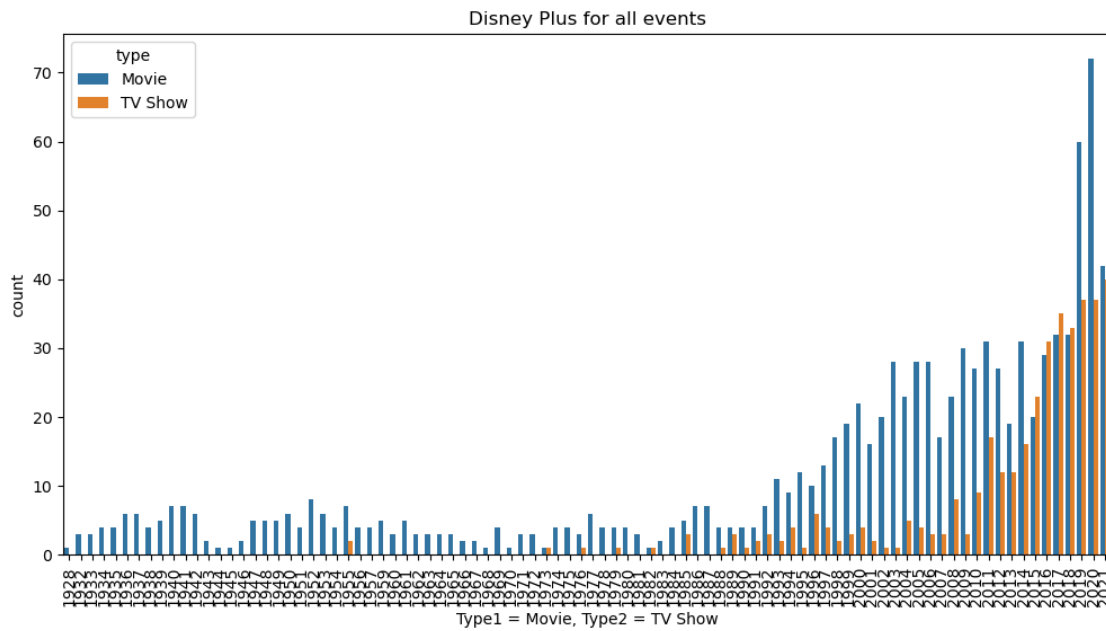
```
[29]: release_year  1928  1932  1933  1934  1935  1936  1937  1938  1939  1940  ... \
type
Movie           1    3    3    4    4    6    6    4    5    7  ...
TV Show         0    0    0    0    0    0    0    0    0    0  ...

release_year  2012  2013  2014  2015  2016  2017  2018  2019  2020  2021
type
Movie          27   19   31   20   29   32   32   60   72   42
TV Show         12   12   16   23   31   35   33   37   37   40
```

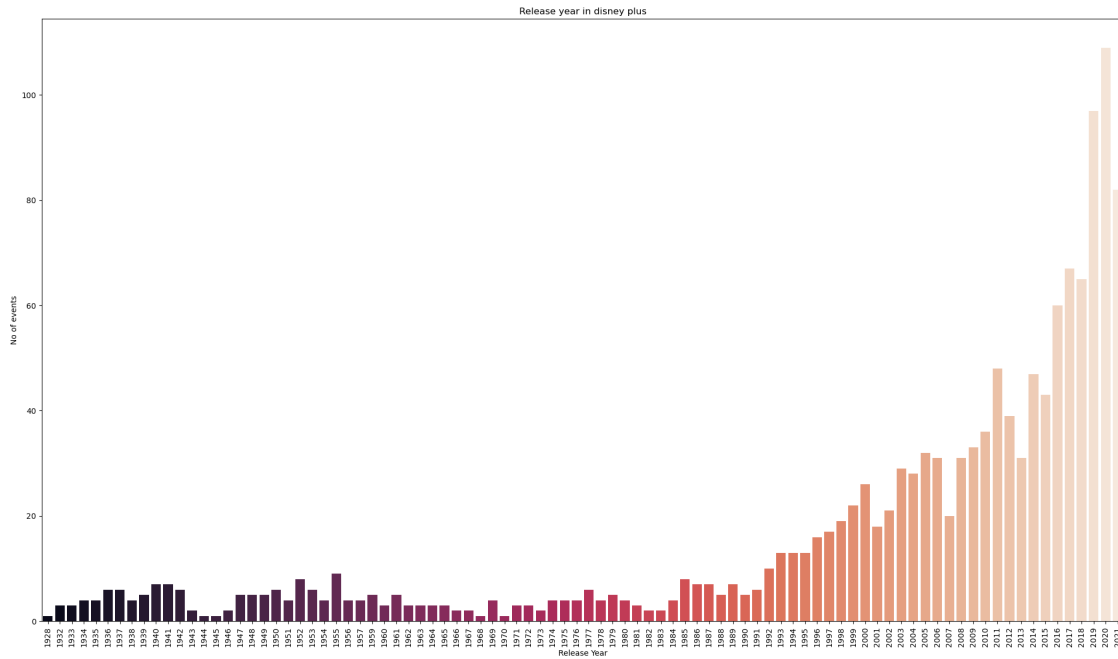
[2 rows x 90 columns]

```
[4]: plt.figure(figsize = (12, 6))
sns.countplot(x = 'release_year', data = df, hue = "type")
plt.title("Disney Plus for all events")
```

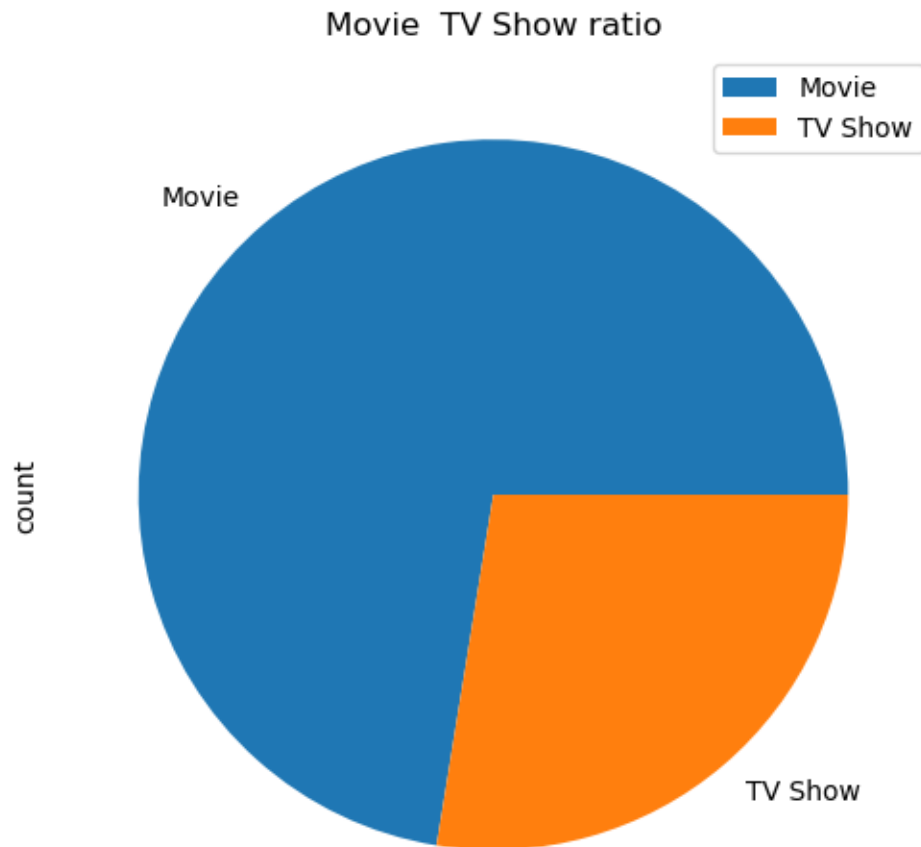
```
plt.xlabel("Type1 = Movie, Type2 = TV Show")
plt.xticks(rotation = 90);
```



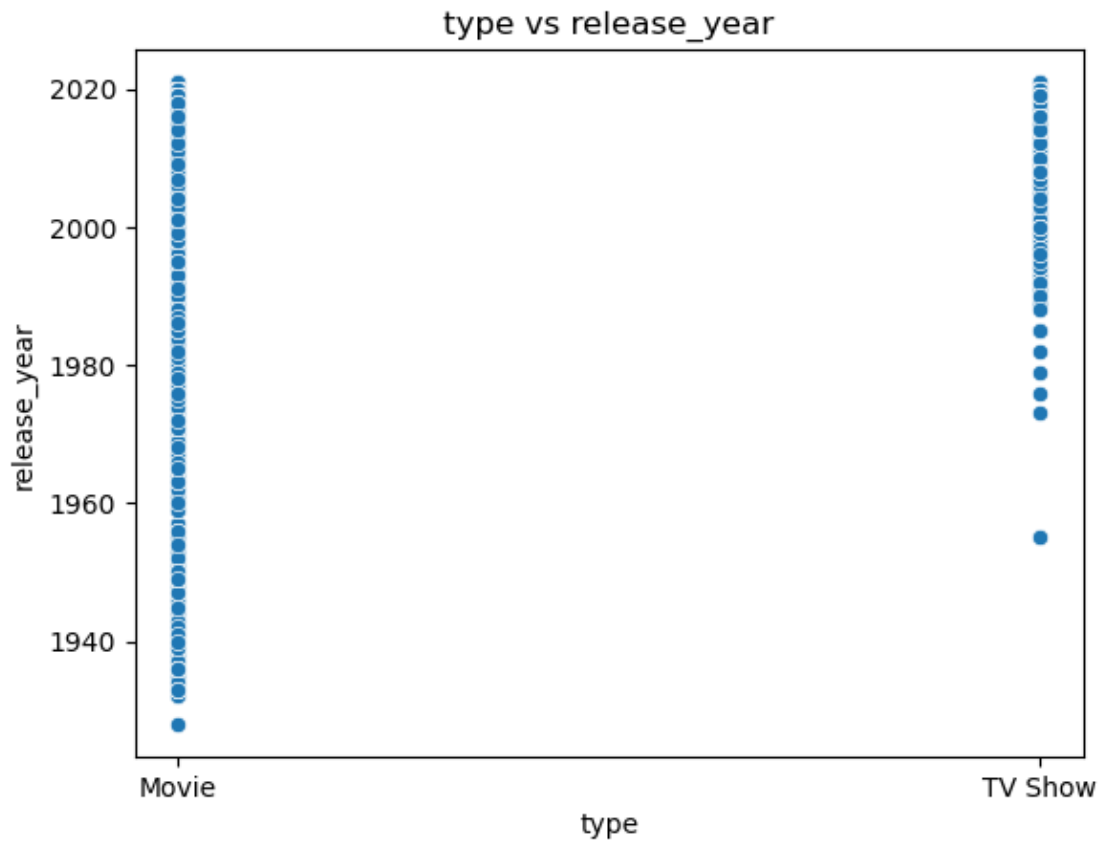
```
[5]: # Count the no of events released in year
release_per_year = df['release_year'].value_counts().sort_index()
# Create a bar plot
plt.figure(figsize = (25, 14))
sns.barplot(x = release_per_year.index, y = release_per_year.values, palette = 'rocket')
plt.title('Release year in disney plus')
plt.xlabel('Release Year')
plt.ylabel('No of events')
plt.xticks(rotation = 90)
plt.show()
```



```
[41]: # Plotting a pie chart
df.type.value_counts().plot(kind = 'pie', figsize = (8, 6))
plt.title("Movie TV Show ratio")
plt.legend(['Movie', 'TV Show']);
```



```
[42]: # Scatter plot between 'type' and 'release_year'.
sns.scatterplot(data = df, x = 'type', y = 'release_year')
plt.title('type vs release_year')
plt.xlabel('type')
plt.ylabel('release_year')
plt.show()
```



```
[38]: # Let's find the answer of our 2nd question.
# 2. How many Movie, TV Show Types released by their respective directors?
pd.crosstab(df.type, df.director)
```

```
[38]: director  Aaron Blaise, Robert Walker  Adam Shankman  \
type
Movie                                1                3

director  Adam Stein, Zach Lipovsky  Alan Barillaro  Alan Shapiro  \
type
Movie                                1                1                1

director  Alan Taylor  Alastair Fothergill, Andy Byatt  \
type
Movie                                1                1

director  Alastair Fothergill, Jeff Wilson  \
type
Movie                                2
```

```

director  Alastair Fothergill, Keith Scholey  \
type
Movie                                          1

director  Alastair Fothergill, Mark Linfield  ...  Wolfgang Reitherman  \
type                                          ...
Movie                                      1  ...                                3

director  Wolfgang Reitherman, Hamilton Luske, Clyde Geronimi  \
type
Movie                                          1

director  Wolfgang Reitherman, John Lounsbery  \
type
Movie                                          1

director  Wolfgang Reitherman, John Lounsbery, Art Stevens  Xavier Koller  \
type
Movie                                          1                                1

director  Zach A. Parrish  Zak Knutson  Zhong Yu  Érik Canuel      ,
type
Movie          1          1          1          1          1

[1 rows x 578 columns]

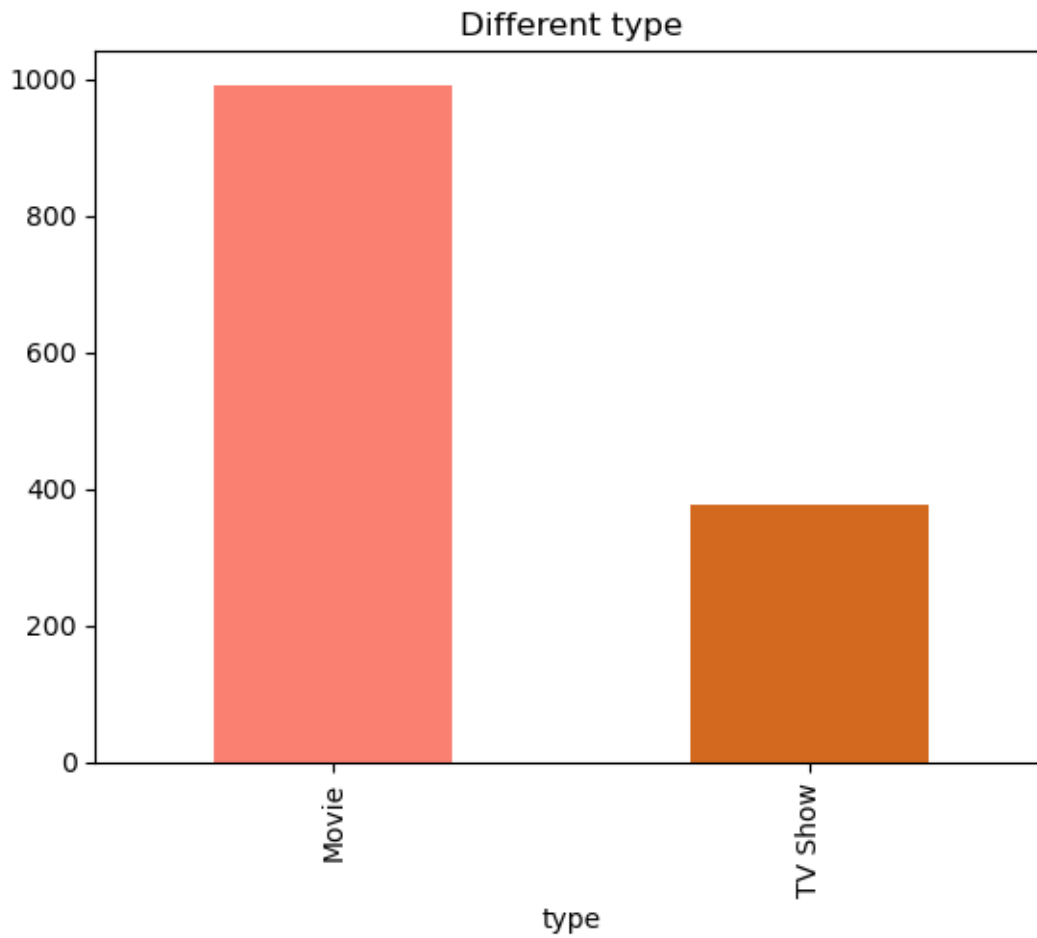
```

```

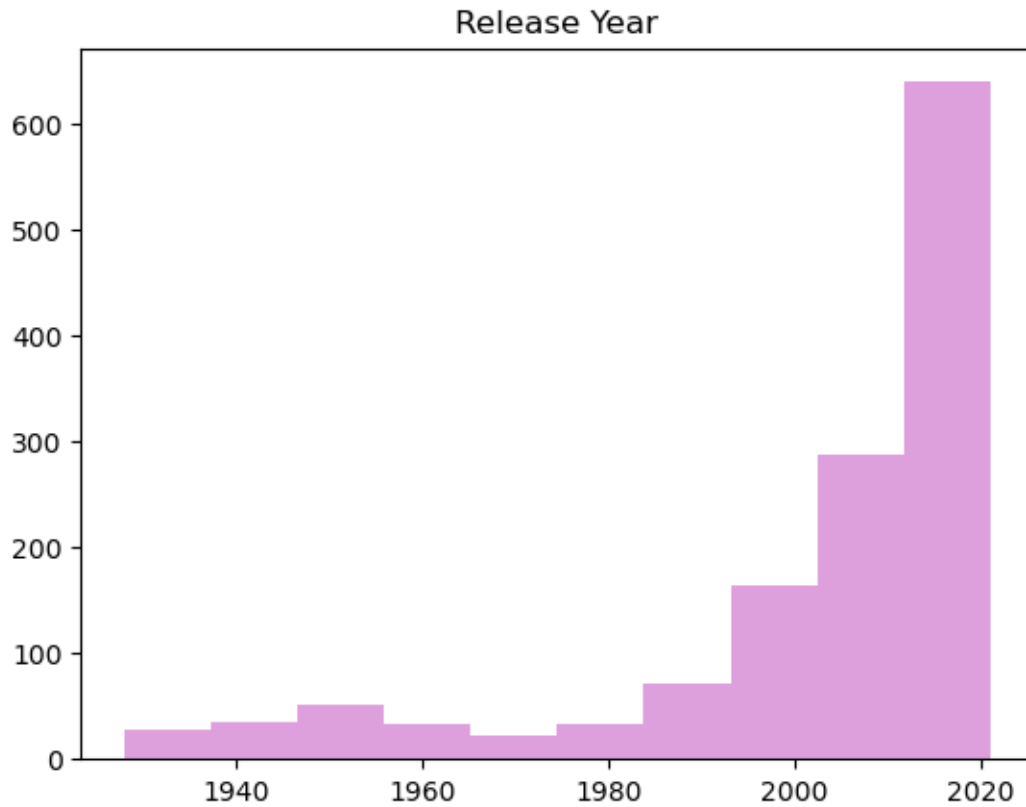
[23]: df.type.value_counts().plot(kind = 'bar', color = ['salmon', 'chocolate'])
      plt.title('Different type');

```





```
[16]: # Histogram of release_year
plt.hist(df['release_year'], color = ['plum'])
# Adding Title to the histogram
plt.title("Release Year")
# Adding the legends
plt.show()
```



```
[34]: # Let's move to question 3
      # 3. Type of which dates added in disney plus?
      # Counting the different dates added
      df.date_added.value_counts()
```

```
[34]: date_added
      November 12, 2019      723
      January 1, 2020        28
      April 3, 2020         26
      May 1, 2020           20
      April 2, 2021         16
      ...
      March 13, 2020         1
      March 5, 2020          1
      July 14, 2021          1
      February 25, 2020      1
      September 11, 2020     1
      Name: count, Length: 150, dtype: int64
```

```
[36]: from datetime import datetime
      import datetime
```

```
[37]: df['date_added'] = pd.to_datetime(df['date_added'])
df.head(2)
```

```
[37]: show_id  type          title          director \
0      s1  Movie  A Spark Story  Jason Stermann, Leanne Dare
1      s2  Movie  Spooky Buddies          Robert Vince

                                cast          country \
0                                Apton Corbin, Louis Gonzales          NaN
1  Tucker Albrizzi, Diedrich Bader, Ameko Eks Mas...  United States, Canada

date_added  release_year  rating  duration          listed_in \
0  2021-09-24          2021  TV-PG    88 min          Documentary
1  2021-09-24          2011      G    93 min  Comedy, Fantasy, Kids

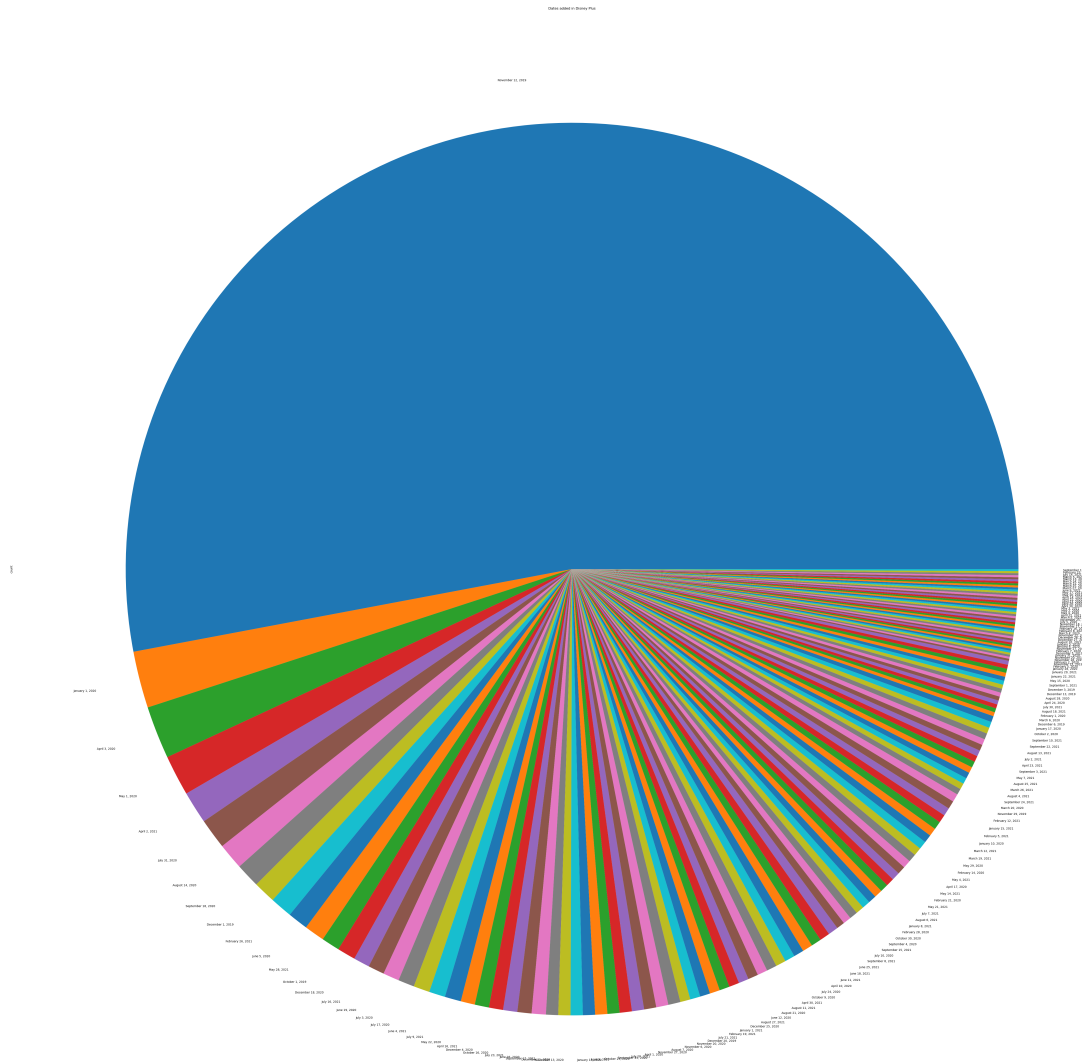
                                description
0  Two Pixar filmmakers strive to bring their uni...
1  The puppies go on a spooky adventure through a...
```

```
[5]: pd.crosstab(df.date_added, df.type)
```

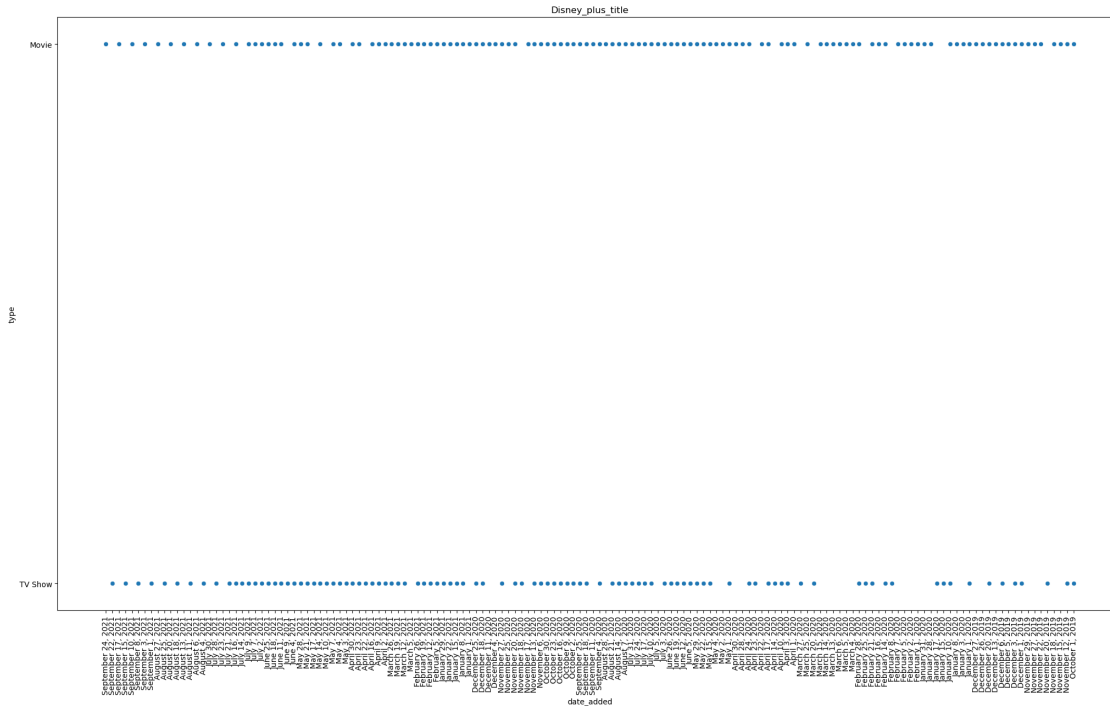
```
[5]: type          Movie  TV Show
date_added
April 1, 2020          6          0
April 10, 2020         2          3
April 14, 2020         0          1
April 16, 2021         4          3
April 17, 2020         2          2
...
September 24, 2021     3          0
September 25, 2020     4          2
September 3, 2021      3          0
September 4, 2020      3          1
September 8, 2021      0          4
```

[150 rows x 2 columns]

```
[15]: df.date_added.value_counts().plot(kind = 'pie', figsize = (100, 70))
plt.title('Dates added in Disney Plus');
```



```
[7]: plt.figure(figsize = (25, 14))
sns.scatterplot(data = df, x = 'date_added', y = 'type')
plt.title('Disney_plus_title')
plt.xlabel('date_added')
plt.ylabel('type')
plt.xticks(rotation = 90)
plt.show()
```



```
[9]: # Let's move to question 4.
      # 4. Which type of country release in disney plus?
      # Counting how many countries are present in disney plus
      df.country.value_counts()
```

```
[9]: country
United States          976
United States, Canada   28
United Kingdom         23
United States, United Kingdom  19
Canada, United States   17
...
United Kingdom, United States, Canada  1
United States, Germany, Ireland, United Kingdom  1
United States, United Kingdom, South Africa  1
United Kingdom, United States, Ireland  1
New Zealand, United States  1
Name: count, Length: 87, dtype: int64
```

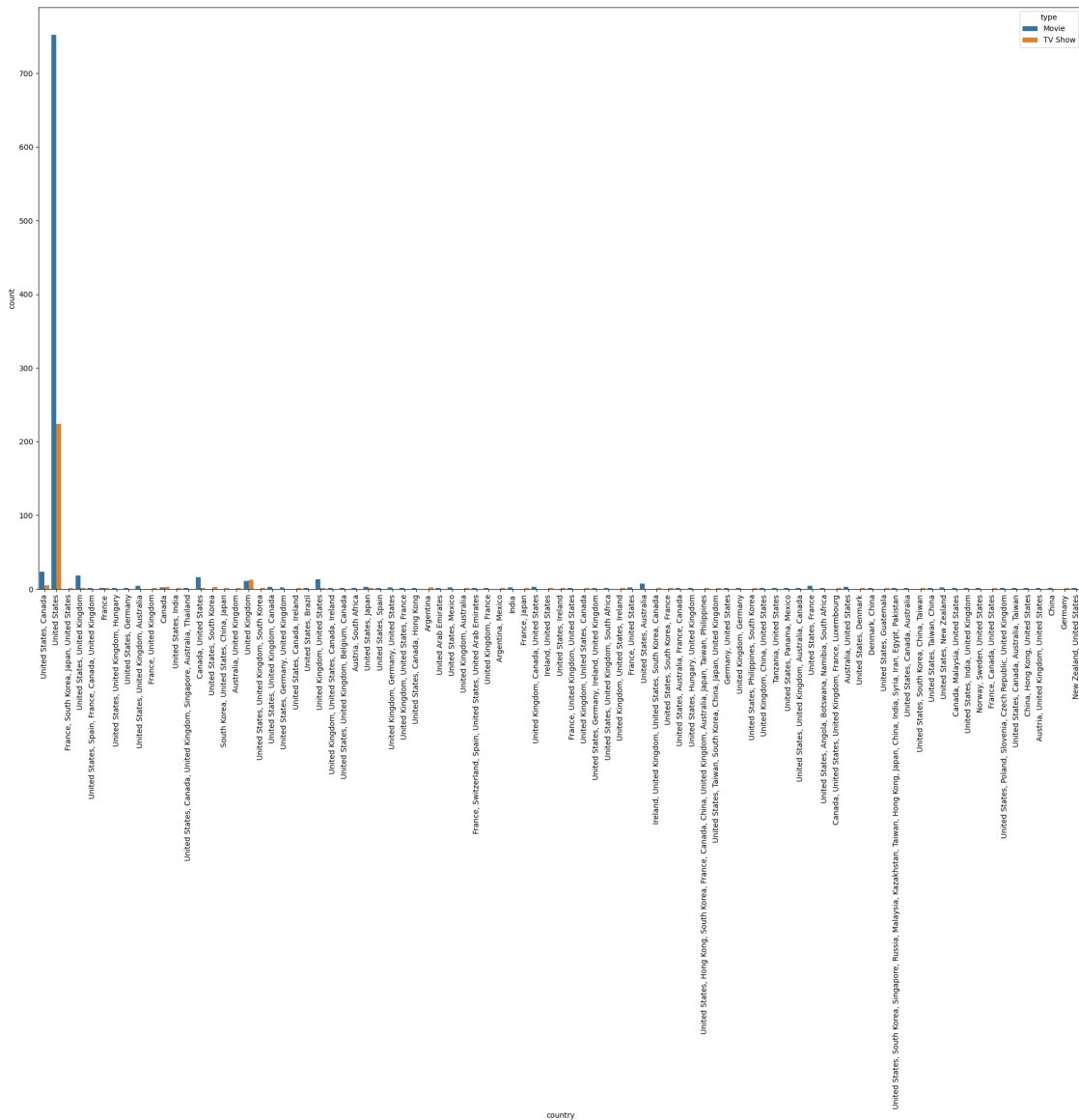
```
[13]: # Plotting a bar chart
      plt.figure(figsize = (25, 14))
      df.country .value_counts().plot(kind = 'bar', color = ['khaki'])
      plt.title('disney_plus vs count');
```



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]

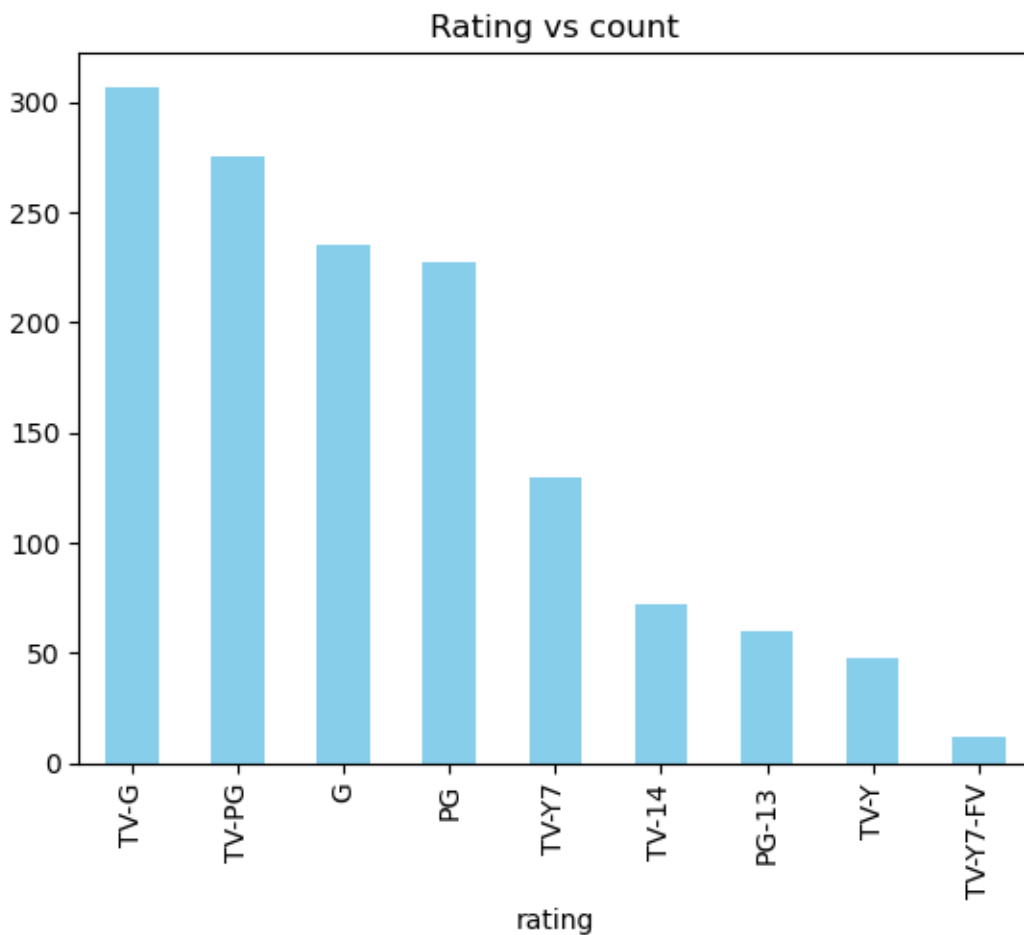
```
[6]: plt.figure(figsize = (25, 14))
sns.countplot(x = 'country', data = df, hue = 'type')
plt.xticks(rotation = 90);
```



```
[21]: # Now question 5?  
# 5. Type of which rating in disney plus?  
df.rating.value_counts()
```

```
[21]: rating  
TV-G      307  
TV-PG     275  
G         235  
PG        227  
TV-Y7     130  
TV-14      72  
PG-13      60  
TV-Y       48  
TV-Y7-FV   12  
Name: count, dtype: int64
```

```
[28]: df.rating.value_counts().plot(kind = 'bar', color = ['skyblue'])  
plt.title('Rating vs count');
```

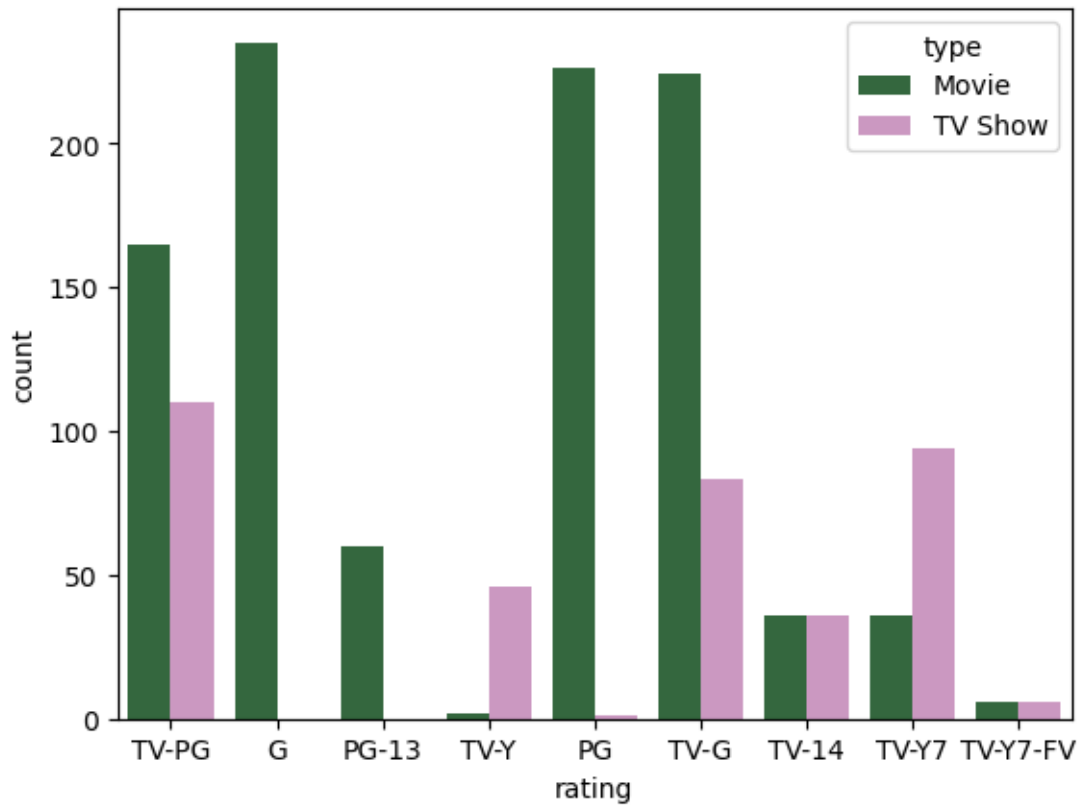




```
[29]: pd.crosstab(df.rating, df.type)
```

```
[29]: type      Movie  TV Show
rating
G           235         0
PG          226         1
PG-13        60         0
TV-14        36        36
TV-G         224        83
TV-PG        165       110
TV-Y          2        46
TV-Y7        36       94
TV-Y7-FV      6         6
```

```
[7]: sns.countplot(x = 'rating', data = df, hue = 'type', palette = 'cubehelix');
```

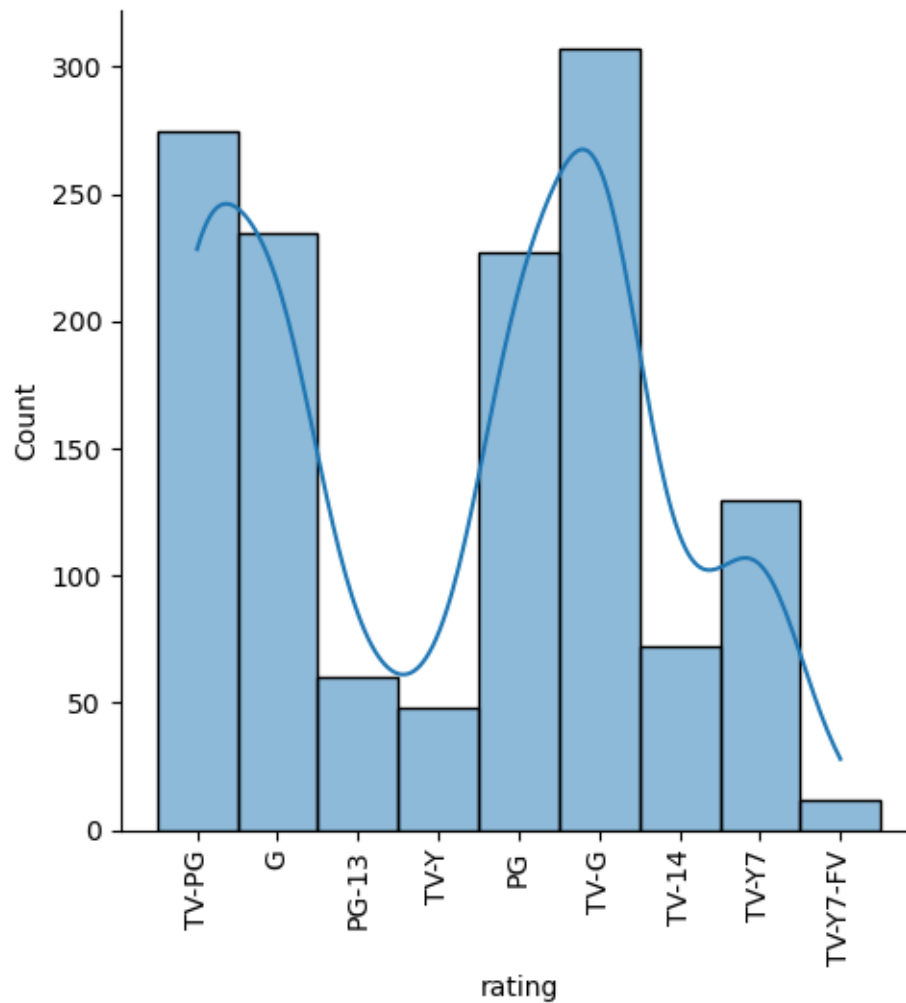


```
[6]: sns.displot(x = 'rating', data = df, bins = 50, kde = True)
plt.xticks(rotation = 90);
```

C:\ProgramData\anaconda3\Lib\site-packages\seaborn\\_oldcore.py:1119:

FutureWarning: use\_inf\_as\_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.

```
with pd.option_context('mode.use_inf_as_na', True):
```



```
[6]: # Let's move to question 6
      # 6. How much duration for each type in disney plus?
      # Getting the values
      df.duration.value_counts()
```

```
[6]: duration
     1 Season      204
     2 Seasons      80
     3 Seasons      43
     7 min         39
    44 min         39
```

...

```

10 Seasons      1
160 min         1
16 Seasons      1
31 min          1
68 min          1
Name: count, Length: 156, dtype: int64

```

```

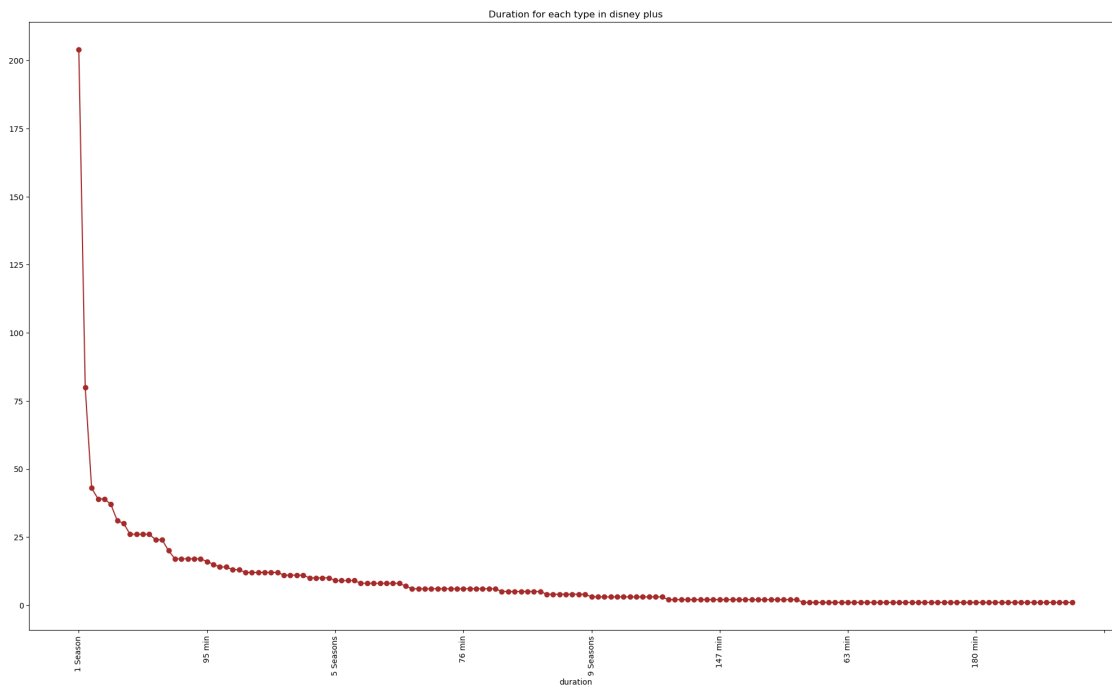
[11]: plt.figure(figsize = (25, 14))
      df.duration.value_counts().plot(kind = 'line',marker = 'o', color = ['brown'])
      plt.title('Duration for each type in disney plus')
      plt.xticks(rotation = 90)

```

```

[11]: (array([-20.,  0., 20., 40., 60., 80., 100., 120., 140., 160., 180.]),
      [Text(-20.0, 0, '148 min'),
        Text(0.0, 0, '1 Season'),
        Text(20.0, 0, '95 min'),
        Text(40.0, 0, '5 Seasons'),
        Text(60.0, 0, '76 min'),
        Text(80.0, 0, '9 Seasons'),
        Text(100.0, 0, '147 min'),
        Text(120.0, 0, '63 min'),
        Text(140.0, 0, '180 min'),
        Text(160.0, 0, ''),
        Text(180.0, 0, '')])

```

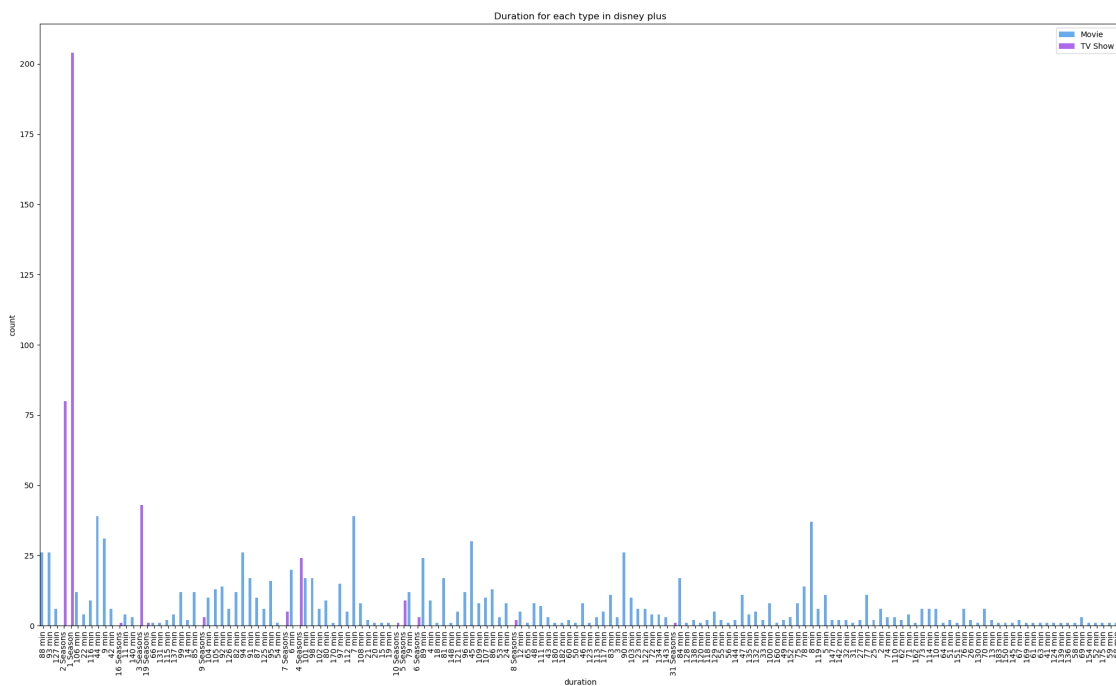


```
[4]: pd.crosstab(df.duration, df.type)
```

```
[4]: type      Movie  TV Show
duration
1 Season      0      204
10 Seasons    0        1
10 min        6        0
100 min       8        0
101 min       17       0
...
95 min       16        0
96 min       12        0
97 min       15        0
98 min       17        0
99 min       12        0
```

[156 rows x 2 columns]

```
[9]: plt.figure(figsize = (25, 14))
sns.countplot(x = 'duration', data = df, hue = 'type', palette = 'cool')
plt.title('Duration for each type in disney plus')
plt.xticks(rotation = 90)
plt.legend(['Movie', 'TV Show']);
```



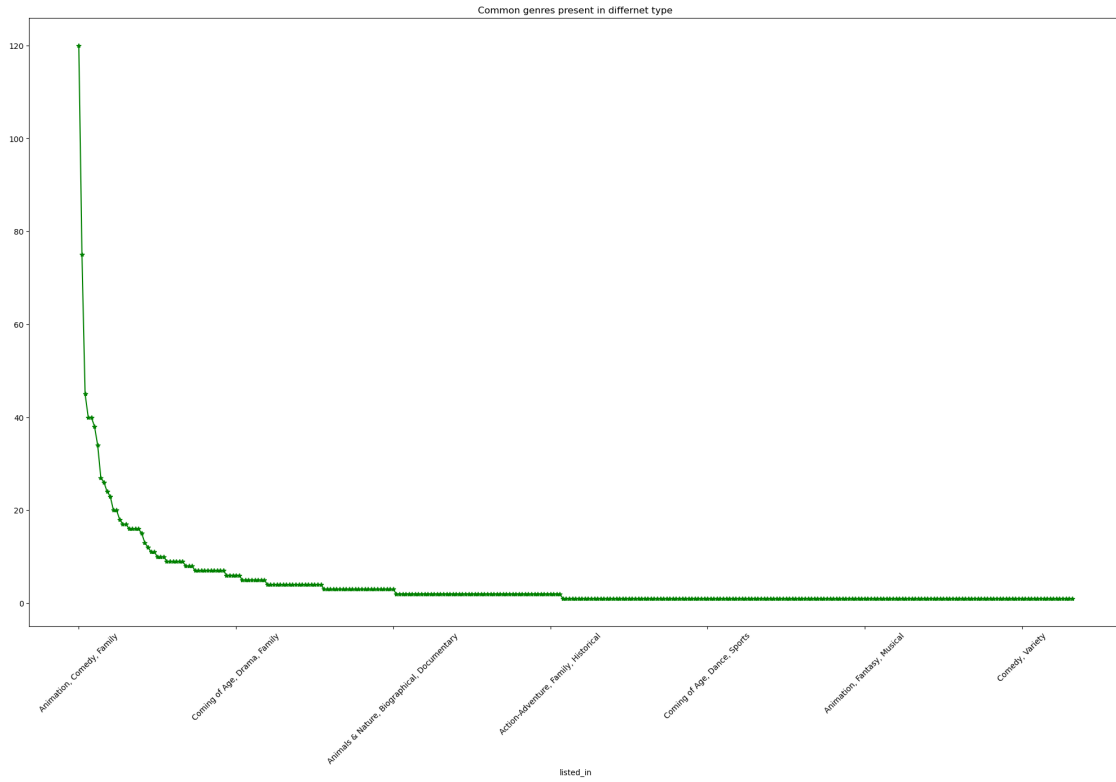
```
[24]: # 7. What are different type of genres present in disney plus?
pd.crosstab(df.listed_in, df.type)
```

```
[24]: type                Movie  TV Show
listed_in
Action-Adventure, Animals & Nature, Animation      1      0
Action-Adventure, Animals & Nature, Comedy          1      0
Action-Adventure, Animals & Nature, Documentary     7      0
Action-Adventure, Animals & Nature, Docuseries      0      9
Action-Adventure, Animals & Nature, Drama          5      0
...
Science Fiction      ...      ...
Series              0      3
Sports              1      0
Talk Show           0      1
Variety             1      0
```

```
[317 rows x 2 columns]
```

```
[23]: plt.figure(figsize = (25, 14))
df.listed_in.value_counts().plot(kind = 'line', marker = '*', color = ['green'])
plt.title('Common genres present in differnet type')
plt.xticks(rotation = 45)
```

```
[23]: (array([-50.,  0.,  50., 100., 150., 200., 250., 300., 350.]),
[Text(-50.0, 0, 'Action-Adventure, Docuseries, Lifestyle'),
Text(0.0, 0, 'Animation, Comedy, Family'),
Text(50.0, 0, 'Coming of Age, Drama, Family'),
Text(100.0, 0, 'Animals & Nature, Biographical, Documentary'),
Text(150.0, 0, 'Action-Adventure, Family, Historical'),
Text(200.0, 0, 'Coming of Age, Dance, Sports'),
Text(250.0, 0, 'Animation, Fantasy, Musical'),
Text(300.0, 0, 'Comedy, Variety'),
Text(350.0, 0, '')])
```



```
[25]: plt.figure(figsize = (30, 16))
sns.countplot(x = 'listed_in', data = df, hue = 'type', palette = 'viridis')
plt.xticks(rotation = 90)
plt.tight_layout();
```

