

```
In [40]: import pandas as pd
import warnings
warnings.filterwarnings('ignore')

import matplotlib.pyplot as plt
```

```
In [12]: data=pd.read_csv("C:\\Users\\HP\\Downloads\\amazon.csv",encoding= 'unicode_escape')
```

```
In [15]: data.head(10)
```

Out[15]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	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...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train l...
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	NaN	September 24, 2021	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries	The arrival of a charismatic young priest brin...
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2021	PG	91 min	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	September 24, 2021	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
8	s9	TV Show	The Great British Baking Show	Andy Devonshire	Mel Giedroyc, Sue Perkins, Mary Berry, Paul Ho...	United Kingdom	September 24, 2021	2021	TV-14	9 Seasons	British TV Shows, Reality TV	A talented batch of amateur bakers face off in...
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	September 24, 2021	2021	PG-13	104 min	Comedies, Dramas	A woman adjusting to life after a loss contend...

In [17]: `data.shape`

Out[17]: (8807, 12)

In [18]: `data.info()`

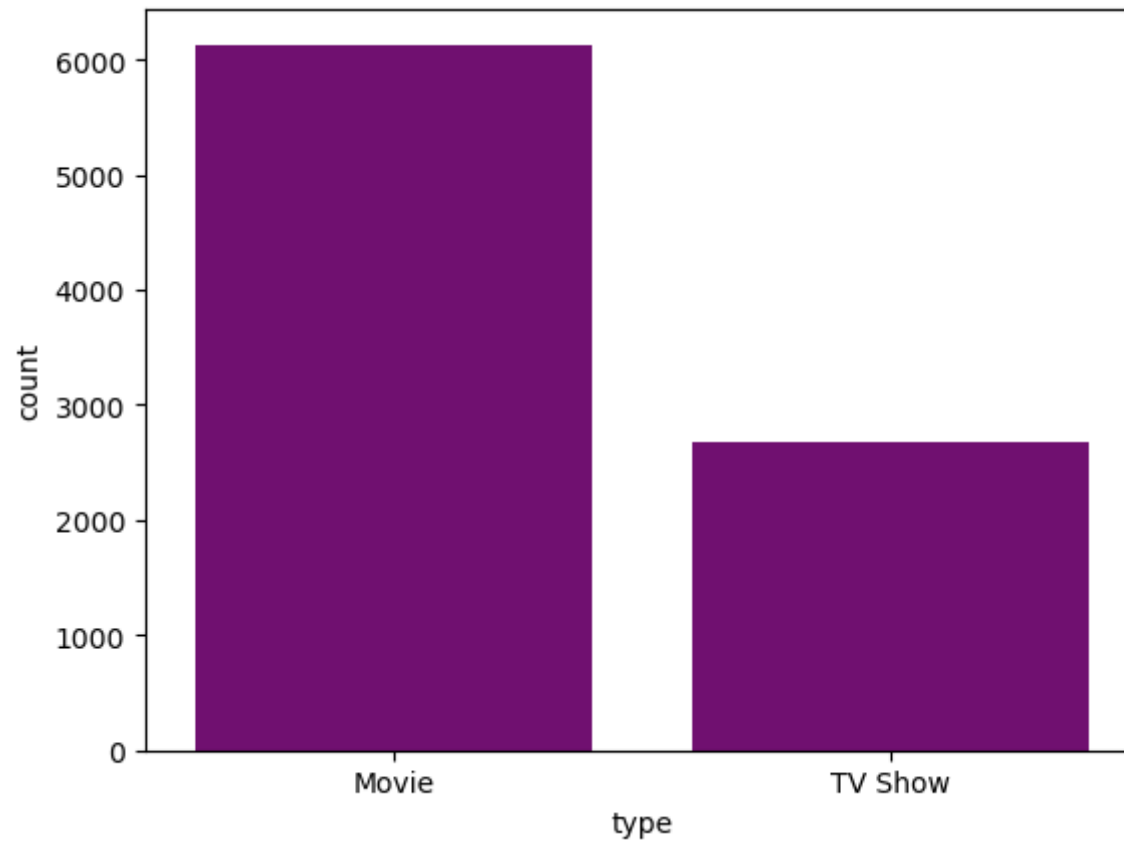
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 12 columns):
#   Column          Non-Null Count  Dtype
---  -
0   show_id         8807 non-null   object
1   type            8807 non-null   object
2   title           8807 non-null   object
3   director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

In [21]: `data.isnull().sum()`

```
Out[21]: show_id      0
         type        0
         title       0
         director    2634
         cast        825
         country     831
         date_added   10
         release_year 0
         rating       4
         duration     3
         listed_in    0
         description  0
         dtype: int64
```

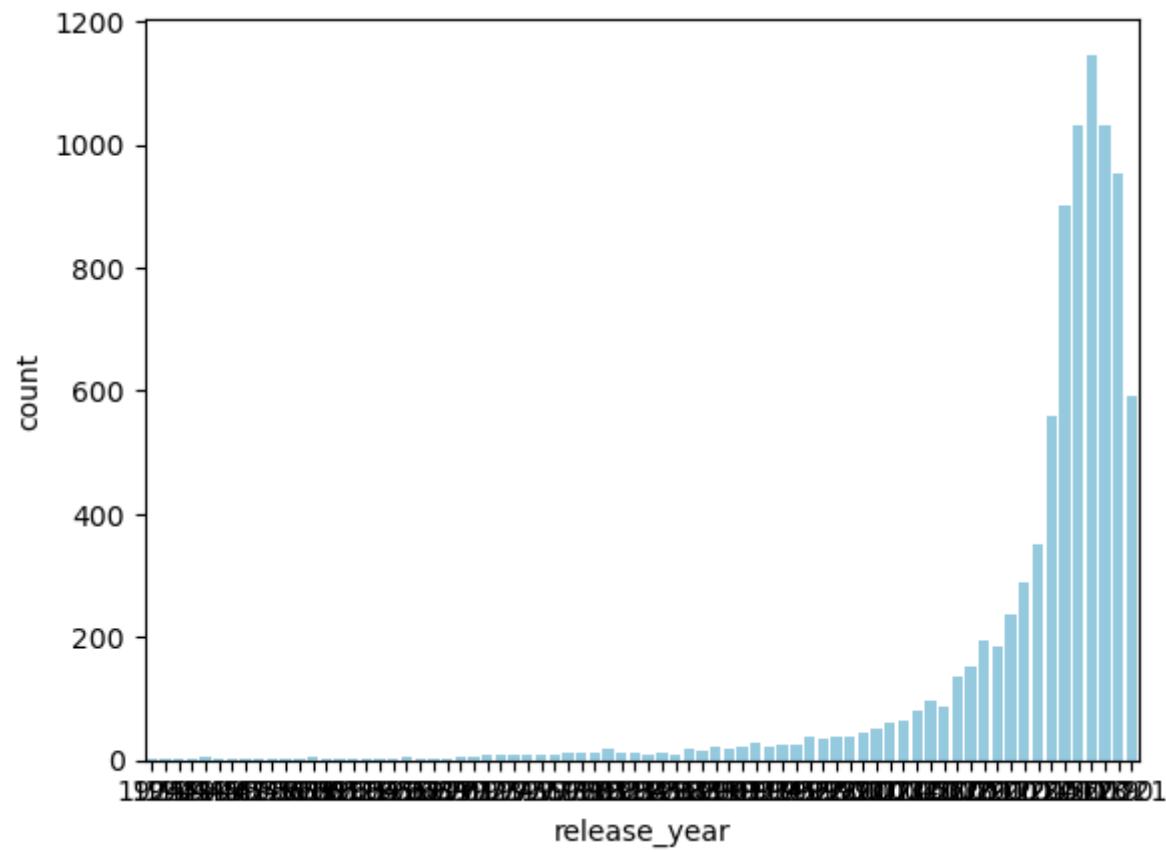
```
In [54]: import seaborn as sns
         sns.countplot(data['type'], label = "Count", color="purple")
```

```
Out[54]: <AxesSubplot:xlabel='type', ylabel='count'>
```



```
In [47]: import seaborn as sns  
sns.countplot(data['release_year'], label = "Count", color="skyblue")
```

```
Out[47]: <AxesSubplot:xlabel='release_year', ylabel='count'>
```



```
In [60]: Country_Count = data.country.value_counts().head(20)
Country_Count
```

```
Out[60]:
```

United States	2818
India	972
United Kingdom	419
Japan	245
South Korea	199
Canada	181
Spain	145
France	124
Mexico	110
Egypt	106
Turkey	105
Nigeria	95
Australia	87
Taiwan	81
Indonesia	79
Brazil	77
Philippines	75
United Kingdom, United States	75
United States, Canada	73
Germany	67

Name: country, dtype: int64

```
In [67]: Country_Count = data.country.value_counts().tail(500)
Country_Count
```

```
Out[67]:
```

Spain, Switzerland	1
United States, Australia, China	1
South Africa, United States, Germany	1
Spain, Italy, Argentina	1
Chile, Spain, Argentina, Germany	1
..	
Romania, Bulgaria, Hungary	1
Uruguay, Guatemala	1
France, Senegal, Belgium	1
Mexico, United States, Spain, Colombia	1
United Arab Emirates, Jordan	1

Name: country, Length: 500, dtype: int64

```
In [69]: liste = data.listed_in.value_counts().head(20)
liste
```

```
Out[69]: Dramas, International Movies      362
Documentaries                          359
Stand-Up Comedy                        334
Comedies, Dramas, International Movies 274
Dramas, Independent Movies, International Movies 252
Kids' TV                               220
Children & Family Movies               215
Children & Family Movies, Comedies     201
Documentaries, International Movies    186
Dramas, International Movies, Romantic Movies 180
Comedies, International Movies         176
Comedies, International Movies, Romantic Movies 152
Dramas                                 138
Dramas, International Movies, Thrillers 134
Action & Adventure, Dramas, International Movies 132
Action & Adventure                     128
International TV Shows, TV Dramas      121
Comedies, Dramas, Independent Movies   116
Crime TV Shows, International TV Shows, TV Dramas 110
Comedies                               110
Name: listed_in, dtype: int64
```

```
In [73]: liste = data.listed_in.value_counts().tail(120)
liste
```

```
Out[73]: TV Action & Adventure, TV Comedies      1
Classic & Cult TV, Kids' TV, TV Comedies        1
Children & Family Movies, Classic Movies, Dramas 1
Crime TV Shows, Romantic TV Shows, Spanish-Language TV Shows 1
Action & Adventure, Classic Movies, Sci-Fi & Fantasy 1
..
Kids' TV, TV Action & Adventure, TV Dramas      1
TV Comedies, TV Dramas, TV Horror               1
Children & Family Movies, Comedies, LGBTQ Movies 1
Kids' TV, Spanish-Language TV Shows, Teen TV Shows 1
Cult Movies, Dramas, Thrillers                  1
Name: listed_in, Length: 120, dtype: int64
```

```
In [74]: director = data.director.value_counts().head(20)
director
```



```
Out[74]: Rajiv Chilaka      19
         Raúl Campos, Jan Suter 18
         Marcus Raboy         16
         Suhas Kadav          16
         Jay Karas            14
         Cathy Garcia-Molina  13
         Martin Scorsese      12
         Youssef Chahine      12
         Jay Chapman          12
         Steven Spielberg     11
         Don Michael Paul     10
         David Dhawan         9
         Y?lmaz Erdo?an       8
         Lance Bangs          8
         Kunle Afolayan        8
         Quentin Tarantino     8
         Ryan Polito           8
         Troy Miller           8
         Shannon Hartman       8
         Fernando Ayllón       8
         Name: director, dtype: int64
```

```
In [75]: duration = data.duration.value_counts().head(20)
         duration
```

```
Out[75]: 1 Season      1793
          2 Seasons    425
          3 Seasons    199
          90 min       152
          94 min       146
          97 min       146
          93 min       146
          91 min       144
          95 min       137
          96 min       130
          92 min       129
          102 min      122
          98 min       120
          99 min       118
          101 min      116
          88 min       116
          103 min      114
          106 min      111
          100 min      108
          89 min       106
          Name: duration, dtype: int64
```

```
In [76]: ratings= data.rating.value_counts().head(20)
          ratings
```

```
Out[76]: TV-MA      3207
          TV-14     2160
          TV-PG     863
          R         799
          PG-13     490
          TV-Y7     334
          TV-Y      307
          PG        287
          TV-G      220
          NR        80
          G         41
          TV-Y7-FV   6
          NC-17     3
          UR        3
          74 min     1
          84 min     1
          66 min     1
          Name: rating, dtype: int64
```