

```
In [1]: #Import the libraries

In [3]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')

In [5]: #Loading Dataset

In [6]: df = pd.read_csv(r"C:\Users\Lenovo\Downloads\DATASET\netflix_titles.csv")

In [7]: df

Out[7]:
```

	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, film...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mababane, Thabani...	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 L...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotsias, Samuel Jouy, Nabil...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Acti...	To protect his family from a powerful drug or...
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, Jyendra 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...
...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Saraah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...

8807 rows x 12 columns

```
In [8]: df.head()

Out[8]:
```

	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, film...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mababane, Thabani...	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 L...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotsias, Samuel Jouy, Nabil...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Acti...	To protect his family from a powerful drug or...
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, Jyendra 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...

```
In [9]: df.tail()

Out[9]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies	While living alone in a spooky town, a young g...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Saraah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...

```
In [10]: #Exploratory data analysis

In [11]: df.shape

Out[11]: (8807, 12)

In [12]: df.describe()

Out[12]:
```

	release_year
count	8807.000000
mean	2014.180198
std	8.819312
min	1925.000000
25%	2013.000000
50%	2017.000000
75%	2018.000000
max	2021.000000

```
In [13]: #checking the missing values
df.isnull().sum()

Out[13]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
show_id	0	0	0	2634	825	831	0	0	4	3	0	0
type	0	0	0	0	0	0	0	0	0	0	0	0
title	0	0	0	0	0	0	0	0	0	0	0	0
director	2634	0	0	0	0	0	0	0	0	0	0	0
cast	825	0	0	0	0	0	0	0	0	0	0	0
country	831	0	0	0	0	0	0	0	0	0	0	0
date_added	0	0	0	0	0	0	0	0	0	0	0	0
release_year	0	0	0	0	0	0	0	0	0	0	0	0
rating	4	0	0	0	0	0	0	0	0	0	0	0
duration	3	0	0	0	0	0	0	0	0	0	0	0
listed_in	0	0	0	0	0	0	0	0	0	0	0	0
description	0	0	0	0	0	0	0	0	0	0	0	0
dtype:	int64	object	object	object	object	object	object	object	object	object	object	object

```
In [14]: #basic information of the data
df.info()

Out[14]:
```

```
<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 [15]: #unique values of the data
df.nunique()

Out[15]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
show_id	8807	2	8807	4528	7982	748	1767	74	17	229	514	8775
type	2	2	2	2	2	2	2	2	2	2	2	2
title	8807	8807	8807	4528	7982	748	1767	74	17	229	514	8775
director	6173	6173	6173	6173	6173	6173	6173	6173	6173	6173	6173	6173
cast	7982	7982	7982	7982	7982	7982	7982	7982	7982	7982	7982	7982
country	7976	7976	7976	7976	7976	7976	7976	7976	7976	7976	7976	7976
date_added	8797	8797	8797	8797	8797	8797	8797	8797	8797	8797	8797	8797
release_year	8807	8807	8807	8807	8807	8807	8807	8807	8807	8807	8807	8807
rating	8803	8803	8803	8803	8803	8803	8803	8803	8803	8803	8803	8803
duration	8804	8804	8804	8804	8804	8804	8804	8804	8804	8804	8804	8804
listed_in	8807	8807	8807	8807	8807	8807	8807	8807	8807	8807	8807	8807
description	8807	8807	8807	8807	8807	8807	8807	8807	8807	8807	8807	8807
dtype:	int64	object	object	object	object	object	object	object	object	object	object	object

```
In [16]: #drop missing values
df = df.dropna()
df.shape

Out[16]: (5332, 12)

In [17]: #convert date_added column to datetime format
df['date_added'] = pd.to_datetime(df['date_added'])
df['day_added'] = df['date_added'].dt.day
df['month_added'] = df['date_added'].dt.month
df['year_added'] = df['date_added'].dt.year

In [18]: df.dtypes

Out[18]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
show_id	int64	object	object	object	object	object	datetime64[ns]	int64	object	object	object	object
type	object	object	object	object	object	object	object	object	object	object	object	object
title	object	object	object	object	object	object	object	object	object	object	object	object
director	object	object	object	object	object	object	object	object	object	object	object	object
cast	object	object	object	object	object	object	object	object	object	object	object	object
country	object	object	object	object	object	object	object	object	object	object	object	object
date_added	datetime64[ns]	object	object	object	object	object	object	object	object	object	object	object
release_year	int64	object	object	object	object	object	object	object	object	object	object	object
rating	object	object	object	object	object	object	object	object	object	object	object	object
duration	object	object	object	object	object	object	object	object	object	object	object	object
listed_in	object	object	object	object	object	object	object	object	object	object	object	object
description	object	object	object	object	object	object	object	object	object	object	object	object
day_added	int64	object	object	object	object	object	object	object	object	object	object	object
month_added	int64	object	object	object	object	object	object	object	object	object	object	object
year_added	int64	object	object	object	object	object	object	object	object	object	object	object
dtype:	object	object	object	object	object	object	object	object	object	object	object	object

```
In [19]: df_movies=(df["type"]=="Movie")
df_tvshows=(df["type"]=="TV Shows")

In [20]: #Data visualization

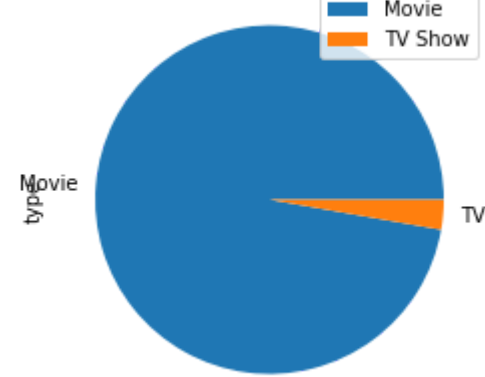
In [22]: #mostly watched type on netflix
df.type.value_counts()

Out[22]:
```

	Movie	TV Show
Movie	5185	147
TV Show	147	179
Name:	type, dtype: int64	

```
In [23]: df.type.value_counts().plot(kind='pie')
plt.title('Mostly watched show type on netflix')
plt.legend()
plt.show()

mostly watched show type on netflix
```



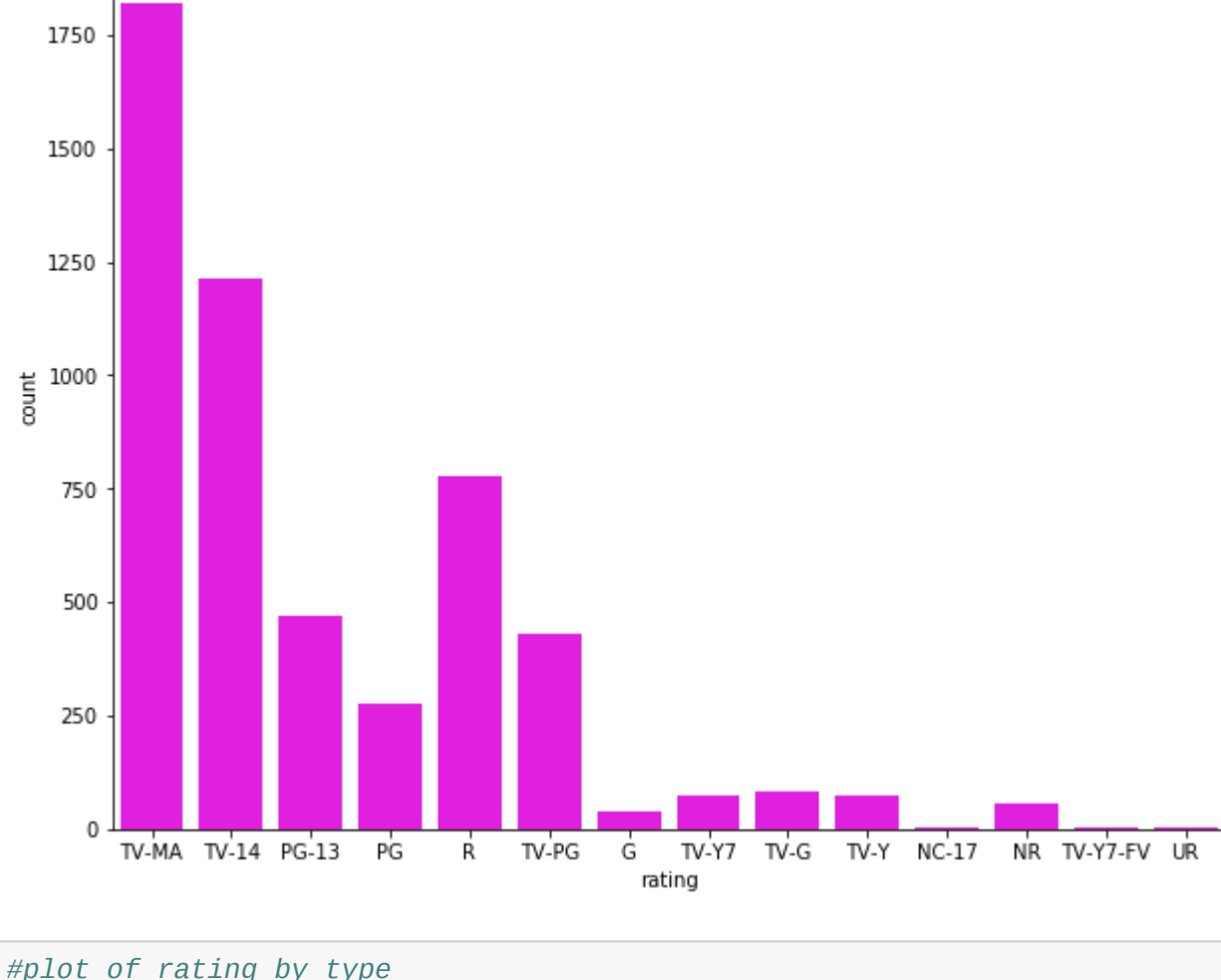
```
In [24]: #top 10 directors of netflix
df.director.value_counts().head(10)

Out[24]:
```

	Raúl Campos, Jan Suter	Marcus Raboy	Jay Karas	Cathy Garcia-Molina	Martin Scorsese	Youssef Chahine	Jay Chapman	Steven Spielberg	Don Michael Paul	David Dhawan
director	18	15	14	13	12	12	12	11	10	9
Name:	director, dtype: int64									

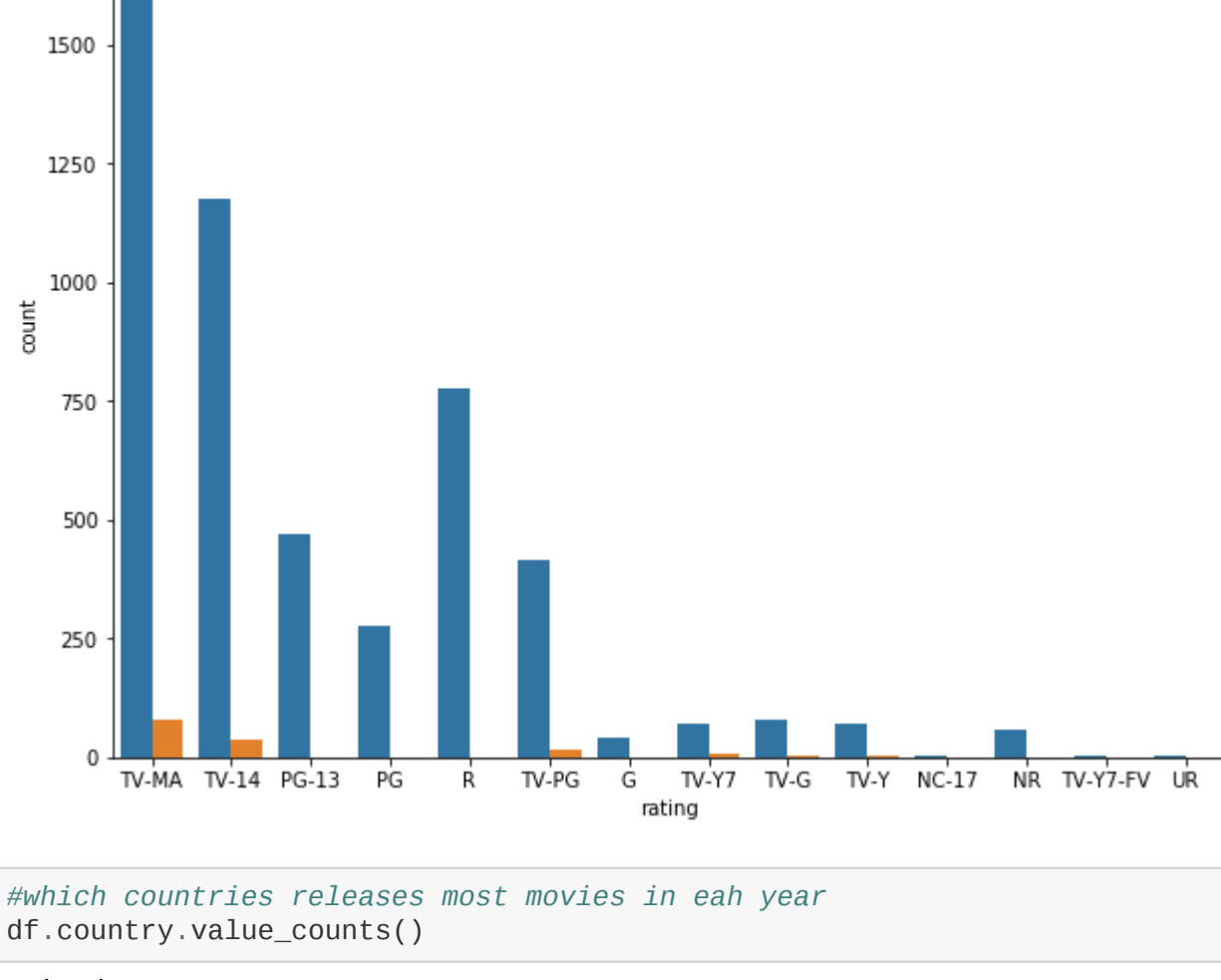
```
In [26]: #plot of rating
plt.figure(figsize=(10,8))
sns.countplot(df.rating,color='magenta')
plt.title("Plot of Rating")

Out[26]: Text(0.5, 1.0, 'Plot of Rating')
```



```
In [27]: #plot of rating by type
plt.figure(figsize=(10,8))
sns.countplot(x='rating',hue='type',data=df)
plt.title("Plot of rating by type")
plt.show()

Plot of rating by type
```



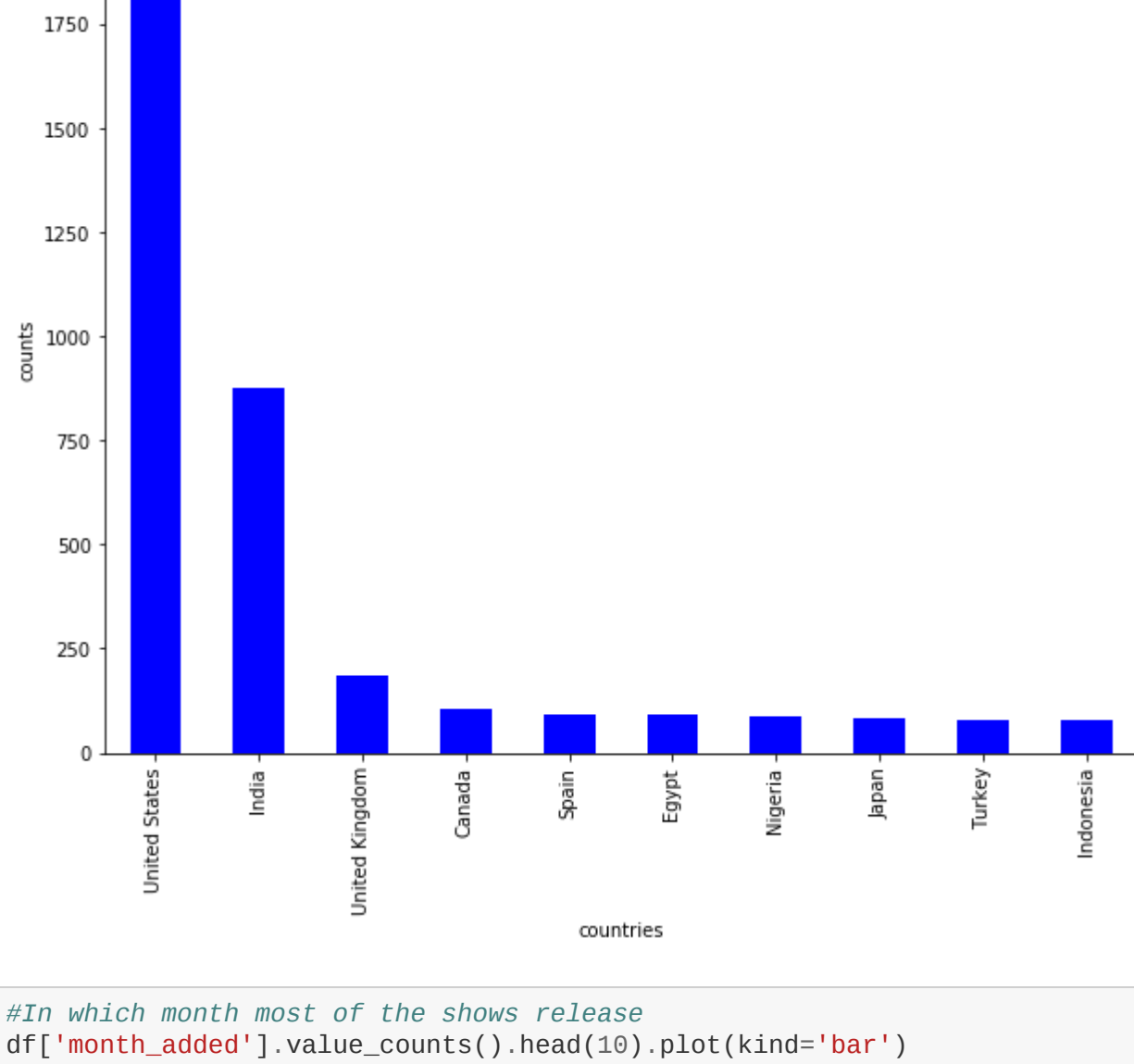
```
In [28]: #which countries releases most movies in eah year
df.country.value_counts()

Out[28]:
```

	United States	India	United Kingdom	Canada	Spain	Uruguay, Guatemala	Romania, Bulgaria, Hungary	Philippines, United States	India, United Kingdom, Canada, United States	United Arab Emirates, Jordan
country	1846	875	183	187	91	1	1	1	1	1
Name:	country, Length: 684, dtype: int64									

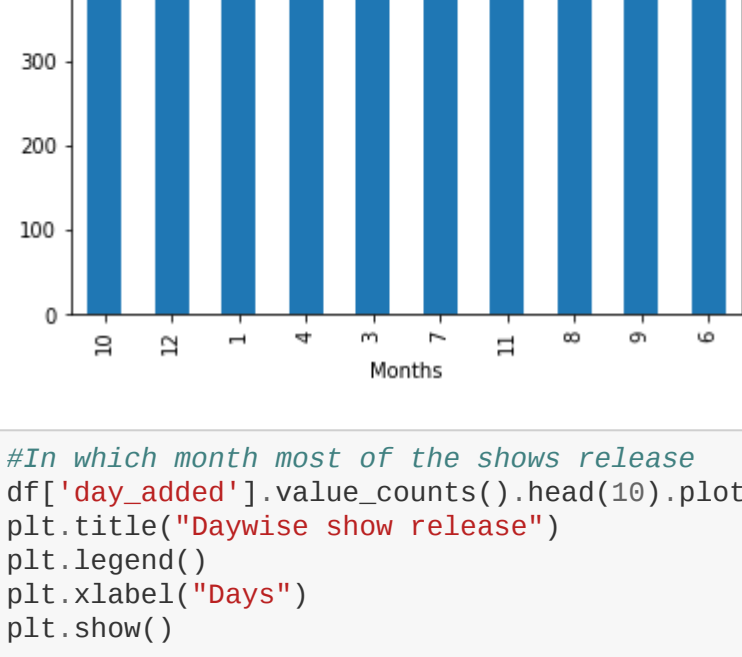
```
In [29]: #top 10 countries releases show on netflix
plt.figure(figsize=(10,8))
df.country.value_counts().head(10).plot(kind='bar',color='blue')
plt.title("top 10 countries releases show on Netflix")
plt.xlabel("countries")
plt.ylabel("counts")
plt.show()

Top 10 countries releases show on Netflix
```



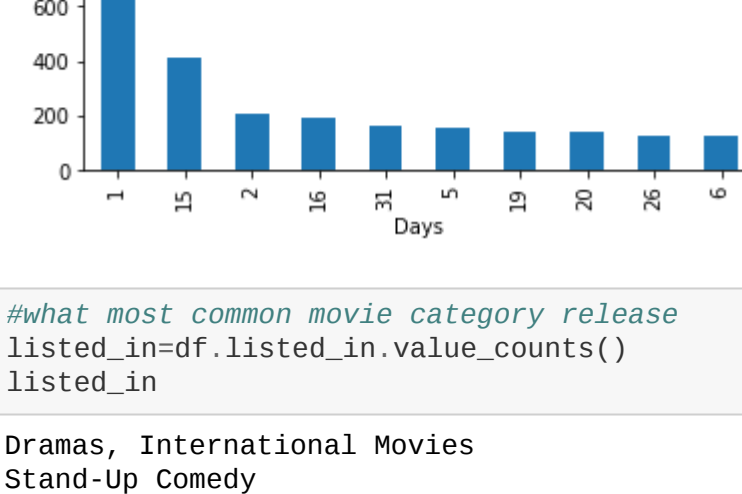
```
In [30]: #In which month most of the shows release
df['month_added'].value_counts().head(10).plot(kind='bar')
plt.title("Monthwise show release")
plt.legend()
plt.xlabel("Months")
plt.show()

Monthwise show release
```



```
In [31]: #In which month most of the shows release
df['day_added'].value_counts().head(10).plot(kind='bar')
plt.title("Daywise show release")
plt.legend()
plt.xlabel("Days")
plt.show()

Daywise show release
```



```
In [32]: #what most common movie category release
listed_in=df.listed_in.value_counts()
listed_in

Out[32]:
```

	Dramas, International Movies	Stand-Up Comedy	Comedies, Dramas, International Movies	Dramas, Independent Movies, International Movies	Children & Family Movies, Comedies	Comedies, Documentaries	International TV Shows, Romantic TV Shows, TV Mysteries	Horror Movies, International Movies, Sci-Fi & Fantasy	Reality TV	Cult Movies, Dramas, Thrillers
listed_in	336	286	257	243	1	1	1	1	1	1
Name:	listed_in, Length: 335, dtype: int64									

```
In [33]: print('Most common category of movie release\nDramas, International Movies:',max(listed_in))

Most common category of movie release
Dramas, International Movies: 336

In [ ]:
```