netflix-dataset

July 20, 2023

```python
[1]: import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
     import seaborn as sns
```

```python
[2]: netf=pd.read_csv('netflix_titles.csv',encoding='Latin1')
     netf
```

```
[2]:      show_id    type                  title          director  \
     0         s1   Movie   Dick Johnson Is Dead  Kirsten Johnson
     1         s2  TV Show        Blood & Water              NaN
     2         s3  TV Show            Ganglands  Julien Leclercq
     3         s4  TV Show  Jailbirds New Orleans             NaN
     4         s5  TV Show         Kota Factory              NaN
     …        …      …                   …                  …
     8802   s8803   Movie               Zodiac    David Fincher
     8803   s8804  TV Show          Zombie Dumb              NaN
     8804   s8805   Movie           Zombieland  Ruben Fleischer
     8805   s8806   Movie                 Zoom     Peter Hewitt
     8806   s8807   Movie               Zubaan      Mozez Singh

                                                    cast        country  \
     0                                               NaN  United States
     1       Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
     2       Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…            NaN
     3                                               NaN            NaN
     4       Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…          India
     …                                                 …              …
     8802    Mark Ruffalo, Jake Gyllenhaal, Robert Downey J…  United States
     8803                                            NaN            NaN
     8804    Jesse Eisenberg, Woody Harrelson, Emma Stone, …  United States
     8805    Tim Allen, Courteney Cox, Chevy Chase, Kate Ma…  United States
     8806    Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan…          India

                  date_added  release_year rating   duration  \
     0     September 25, 2021          2020  PG-13     90 min
     1     September 24, 2021          2021  TV-MA  2 Seasons
```

```
2      September 24, 2021            2021   TV-MA   1 Season
3      September 24, 2021            2021   TV-MA   1 Season
4      September 24, 2021            2021   TV-MA   2 Seasons
...            ...                    ...     ...        ...
8802   November 20, 2019            2007       R    158 min
8803        July 1, 2019            2018   TV-Y7   2 Seasons
8804    November 1, 2019            2009       R     88 min
8805    January 11, 2020            2006      PG     88 min
8806       March 2, 2019            2015   TV-14    111 min

                                                    listed_in  \
0                                              Documentaries
1           International TV Shows, TV Dramas, TV Mysteries
2        Crime TV Shows, International TV Shows, TV Act…
3                                   Docuseries, Reality TV
4        International TV Shows, Romantic TV Shows, TV …
...                                                        …
8802                        Cult Movies, Dramas, Thrillers
8803                Kids' TV, Korean TV Shows, TV Comedies
8804                               Comedies, Horror Movies
8805                    Children & Family Movies, Comedies
8806     Dramas, International Movies, Music & Musicals

                                                  description
0       As her father nears the end of his life, filmm…
1       After crossing paths at a party, a Cape Town t…
2       To protect his family from a powerful drug lor…
3       Feuds, flirtations and toilet talk go down amo…
4       In a city of coaching centers known to train I…
...                                                        …
8802  A political cartoonist, a crime reporter and a…
8803  While living alone in a spooky town, a young g…
8804  Looking to survive in a world taken over by zo…
8805  Dragged from civilian life, a former superhero…
8806  A scrappy but poor boy worms his way into a ty…

[8807 rows x 12 columns]
```

[3]: `netf.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
```

[4]: `netf.isnull().sum()`

[4]:
```
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
```

[5]: `len(netf)`

[5]: 8807

[6]: `netf.fillna({'rating':'null','cast':'unavailable','country':'null','country':`
`↪'unavailable'},inplace=True)`

[7]: `netf`

[7]:
```
        show_id     type                    title          director  \
0           s1    Movie    Dick Johnson Is Dead    Kirsten Johnson
1           s2  TV Show           Blood & Water                NaN
2           s3  TV Show               Ganglands    Julien Leclercq
3           s4  TV Show    Jailbirds New Orleans               NaN
4           s5  TV Show             Kota Factory               NaN
...        ...      ...                     ...                ...
8802     s8803    Movie                  Zodiac      David Fincher
8803     s8804  TV Show             Zombie Dumb                NaN
```

```
8804   s8805   Movie              Zombieland   Ruben Fleischer
8805   s8806   Movie                    Zoom     Peter Hewitt
8806   s8807   Movie                  Zubaan     Mozez Singh

                                            cast        country  \
0                                    unavailable  United States
1      Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
2      Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…    unavailable
3                                    unavailable    unavailable
4      Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…          India
…                                            …              …
8802   Mark Ruffalo, Jake Gyllenhaal, Robert Downey J…  United States
8803                                 unavailable    unavailable
8804   Jesse Eisenberg, Woody Harrelson, Emma Stone, …  United States
8805   Tim Allen, Courteney Cox, Chevy Chase, Kate Ma…  United States
8806   Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan…          India

              date_added  release_year rating    duration  \
0      September 25, 2021          2020  PG-13      90 min
1      September 24, 2021          2021  TV-MA   2 Seasons
2      September 24, 2021          2021  TV-MA    1 Season
3      September 24, 2021          2021  TV-MA    1 Season
4      September 24, 2021          2021  TV-MA   2 Seasons
…              …             …      …          …
8802    November 20, 2019          2007      R     158 min
8803         July 1, 2019          2018  TV-Y7   2 Seasons
8804     November 1, 2019          2009      R      88 min
8805     January 11, 2020          2006     PG      88 min
8806        March 2, 2019          2015  TV-14     111 min

                                          listed_in  \
0                                     Documentaries
1         International TV Shows, TV Dramas, TV Mysteries
2       Crime TV Shows, International TV Shows, TV Act…
3                              Docuseries, Reality TV
4       International TV Shows, Romantic TV Shows, TV …
…                                             …
8802                       Cult Movies, Dramas, Thrillers
8803              Kids' TV, Korean TV Shows, TV Comedies
8804                              Comedies, Horror Movies
8805                   Children & Family Movies, Comedies
8806       Dramas, International Movies, Music & Musicals

                                        description
0      As her father nears the end of his life, filmm…
1      After crossing paths at a party, a Cape Town t…
2      To protect his family from a powerful drug lor…
```

```
3      Feuds, flirtations and toilet talk go down amo…
4      In a city of coaching centers known to train I…
…                                                   …
8802   A political cartoonist, a crime reporter and a…
8803   While living alone in a spooky town, a young g…
8804   Looking to survive in a world taken over by zo…
8805   Dragged from civilian life, a former superhero…
8806   A scrappy but poor boy worms his way into a ty…

[8807 rows x 12 columns]
```

[8]: `netf.isnull().sum()`

```
[8]: show_id          0
     type             0
     title            0
     director      2634
     cast             0
     country          0
     date_added      10
     release_year     0
     rating           0
     duration         3
     listed_in        0
     description      0
     dtype: int64
```

[9]: `netf.head()`

```
[9]:   show_id     type                  title          director  \
     0      s1    Movie   Dick Johnson Is Dead   Kirsten Johnson
     1      s2  TV Show          Blood & Water               NaN
     2      s3  TV Show              Ganglands   Julien Leclercq
     3      s4  TV Show   Jailbirds New Orleans              NaN
     4      s5  TV Show            Kota Factory              NaN

                                            cast        country  \
     0                                 unavailable  United States
     1   Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
     2   Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…    unavailable
     3                                 unavailable     unavailable
     4   Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…          India

              date_added  release_year rating   duration  \
     0  September 25, 2021          2020  PG-13     90 min
     1  September 24, 2021          2021  TV-MA  2 Seasons
     2  September 24, 2021          2021  TV-MA   1 Season
```

```
3  September 24, 2021         2021  TV-MA   1 Season
4  September 24, 2021         2021  TV-MA   2 Seasons

                                        listed_in  \
0                               Documentaries
1     International TV Shows, TV Dramas, TV Mysteries
2  Crime TV Shows, International TV Shows, TV Act…
3                          Docuseries, Reality TV
4  International TV Shows, Romantic TV Shows, TV …

                                        description
0  As her father nears the end of his life, filmm…
1  After crossing paths at a party, a Cape Town t…
2  To protect his family from a powerful drug lor…
3  Feuds, flirtations and toilet talk go down amo…
4  In a city of coaching centers known to train I…
```

```python
[10]: netf['type'].value_counts() \
        .plot(kind='bar')
      plt.show()
```
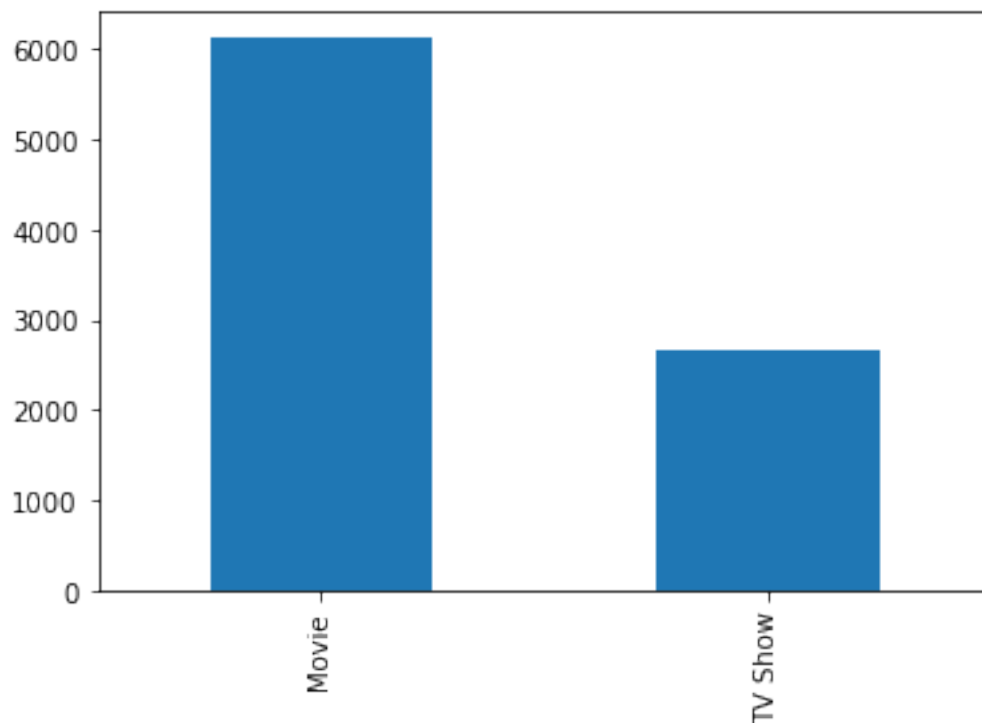


```python
[11]: netf['release_year'].value_counts() \
        .head(15) \
```

```
    .plot(kind='bar')
plt.show()
```
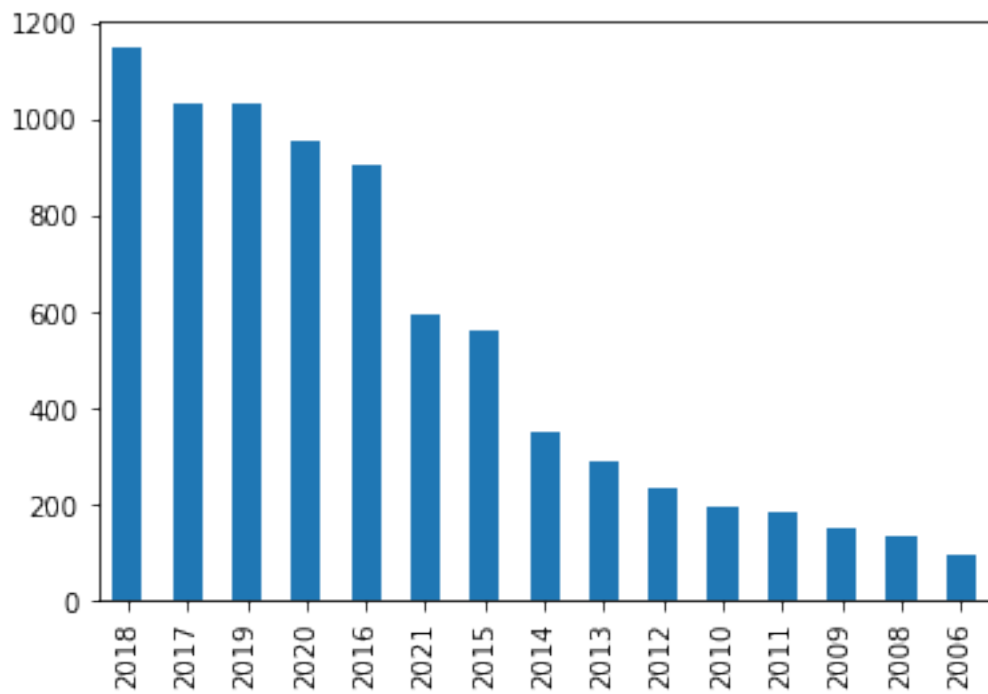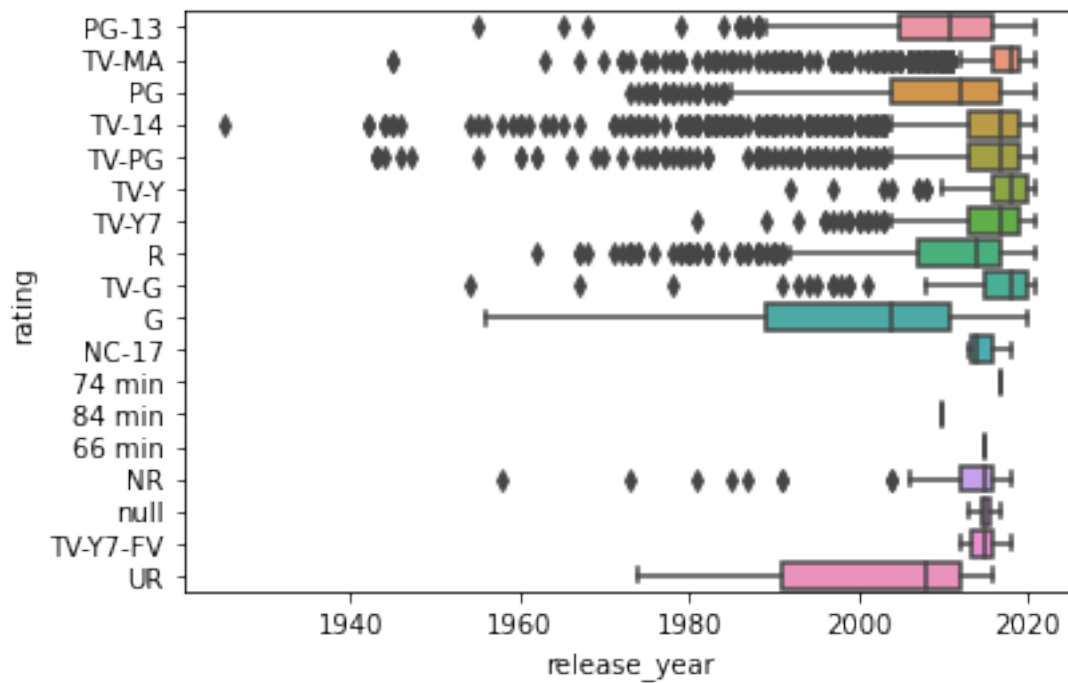


```
[12]: netf['listed_in'].str.split(', ') \
    .explode() \
    .value_counts() \
    .head(30) \
    .plot(kind='barh', figsize=(14, 13))
plt.show()
```
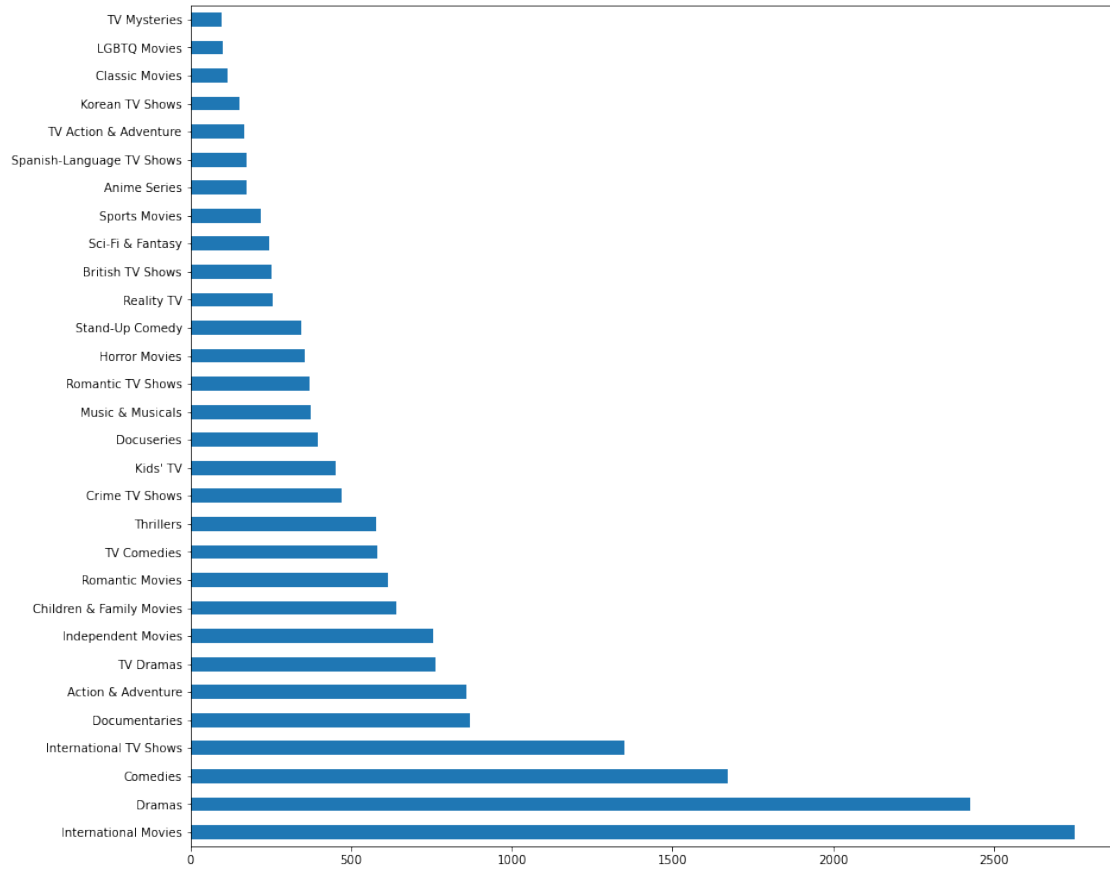
```
[13]: sns.barplot(data=netf,x='release_year',y='rating',hue='type')
      plt.grid()
```
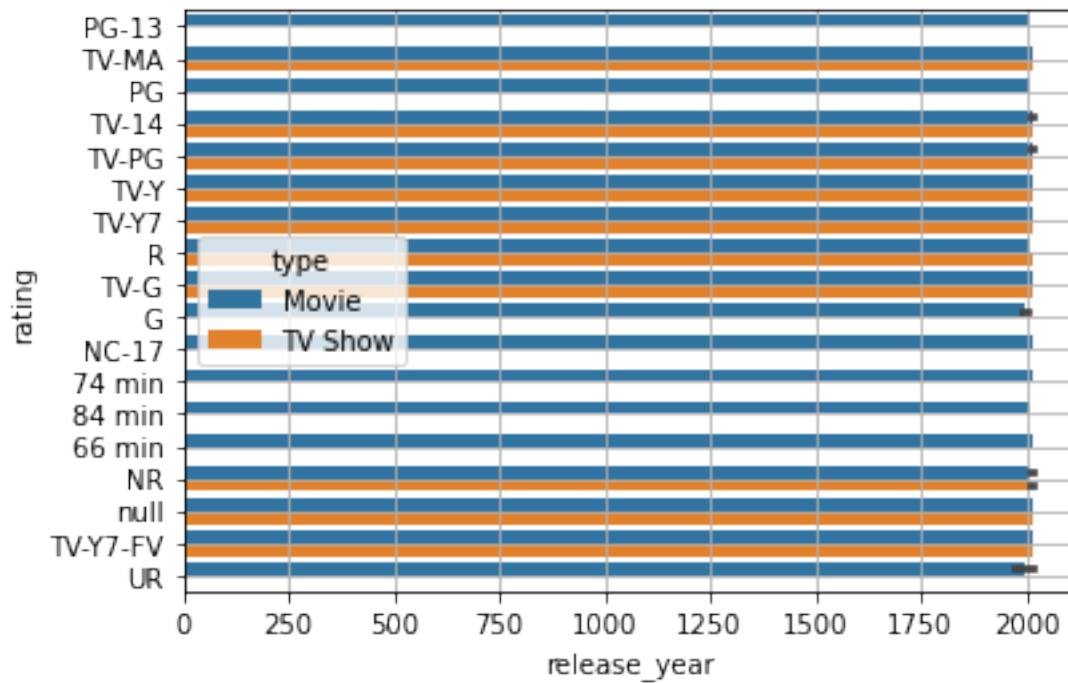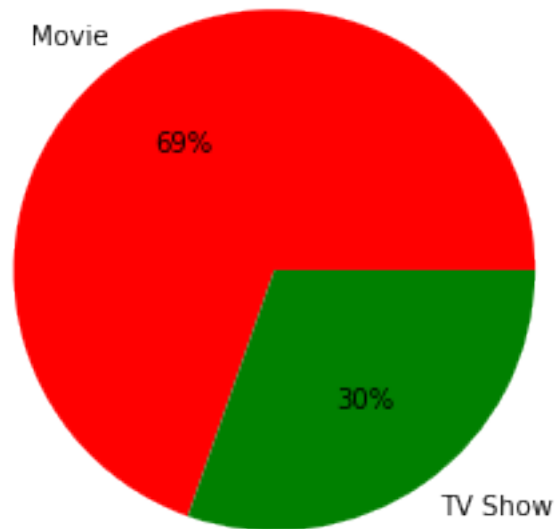
```
[14]: sns.boxplot(data=netf,x='release_year',y='rating')
      plt.show()
```

```
[15]: counts = netf['type'].value_counts()
```

```
[16]: plt.pie(counts.values, labels=counts.index,autopct='%i%%' ,colors=['red',
      ↪'green'])
      plt.title("Movies vs TV Shows", fontsize=20)
      plt.axis('equal')
      plt.show()
```
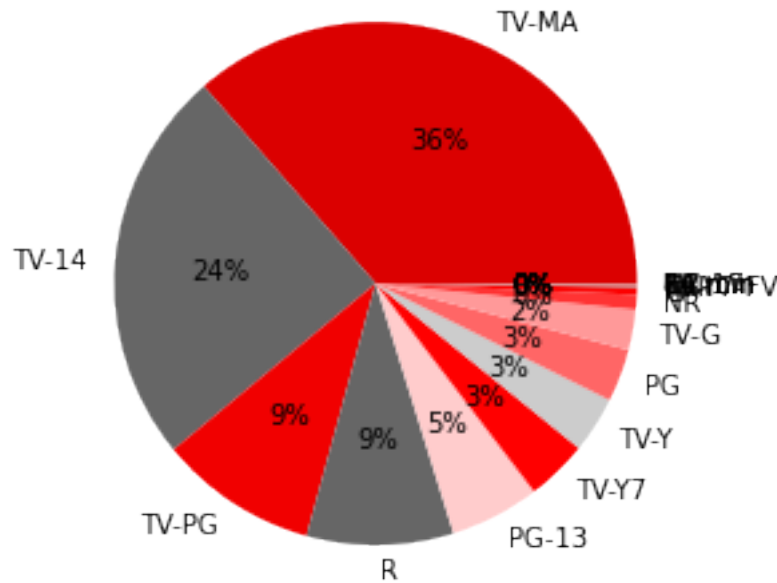
Movies vs TV Shows


```
[17]: counts = netf['rating'].value_counts()
```

```
[18]: plt.pie(counts.values, labels=counts.index,autopct='%i%%',colors=['#db0000',
      ↪'#666666', '#f10000','#666666' ,'#ffcccc', '#ff0000', '#cccccc', '#ff6666',
      '#ff9999', '#ff3333', '#fc0000','#999999', '#e60000','#ffffff'])
      plt.title("Ratings on Netflix Content", fontsize=20)
      plt.axis('equal')
      plt.show()
```

Ratings on Netflix Content


```
[19]: plt.figure(figsize=(15,8))
      country_names = netf['country'].value_counts().index[:10]
      country_values = netf['country'].value_counts().values[:10]
      bar = sns.barplot(x= country_names, y = country_values)
      bar.set_xlabel("Countries", fontdict = {'size' : 'xx-large', 'fontweight' :␣
       ↪'bold'})
      bar.set_ylabel("No of Movies and TV Shows", fontdict = {'size' : 'xx-large',␣
       ↪'fontweight' : 'bold'})
      bar.set_title("Total Number of Movies and TV Shows", fontdict = {'size' :␣
       ↪'xx-large', 'fontweight' : 'bold'})
      bar.bar_label(bar.containers[0])
```

```
[19]: [Text(0, 0, '2818'),
       Text(0, 0, '972'),
       Text(0, 0, '831'),
       Text(0, 0, '419'),
       Text(0, 0, '245'),
       Text(0, 0, '199'),
       Text(0, 0, '181'),
       Text(0, 0, '145'),
       Text(0, 0, '124'),
       Text(0, 0, '110')]
```

Total Number of Movies and TV Shows
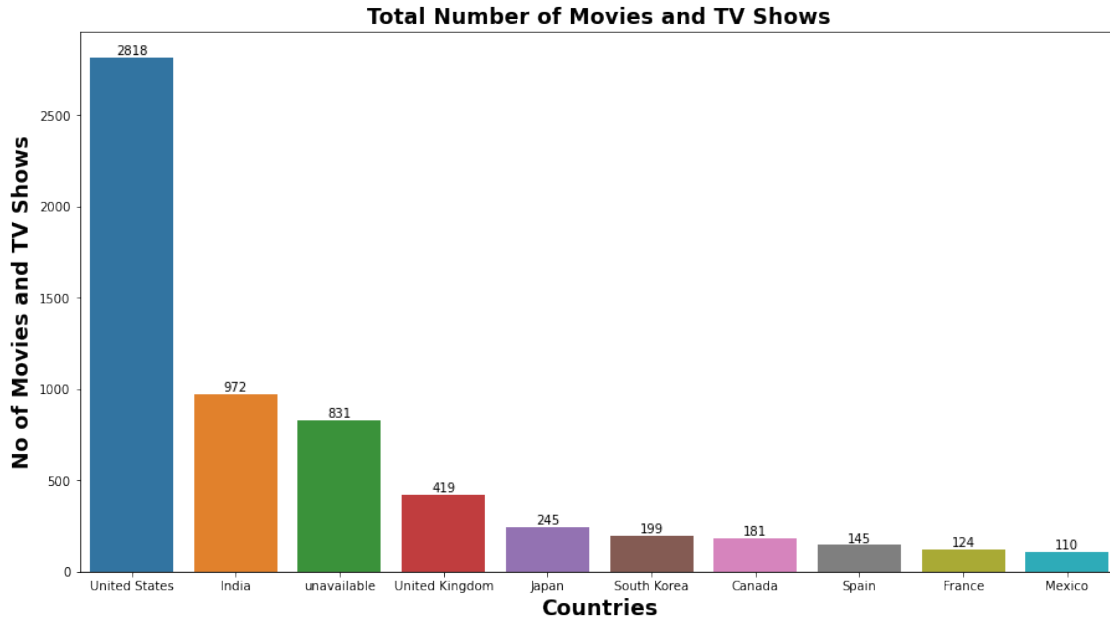
```
[20]: most_viewed=netf['title'].mode()[0]
```

```
[21]: most_viewed
```

```
[21]: '#Alive'
```

```
[22]: type=netf['type'].mode()[0]
      type
```

```
[22]: 'Movie'
```

```
[23]: country=netf['country'].mode()[0]
      country
```

```
[23]: 'United States'
```

```
[24]: listed=netf['listed_in'].mode()[0]
      listed
```

```
[24]: 'Dramas, International Movies'
```

```
[25]: in_director=netf['director'].mode()[0]
      in_director
```

```
[25]: 'Rajiv Chilaka'
```

```
[ ]:
```