

```
In [1]: ┌─ 1 import numpy as np
```

```
In [2]: ┌─ 1 import pandas as pd  
      2 import matplotlib.pyplot as plt  
      3 import seaborn as sns
```

```
In [3]: ┌─ 1 n_data=pd.read_csv('netflix_titles.csv')
```

In [4]: ┌ 1 n_data

Out[4]:

| | show_id | type | title | director | cast | country | date_added | release_year |
|------|---------|---------|-----------------------|-----------------|---|---------------|--------------------|--------------|
| 0 | s1 | Movie | Dick Johnson Is Dead | Kirsten Johnson | NaN | United States | September 25, 2021 | 2020 |
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | South Africa | September 24, 2021 | 2021 |
| 2 | s3 | TV Show | Ganglands | Julien Leclercq | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... | NaN | September 24, 2021 | 2021 |
| 3 | s4 | TV Show | Jailbirds New Orleans | NaN | NaN | NaN | September 24, 2021 | 2021 |
| 4 | s5 | TV Show | Kota Factory | NaN | Mayur More, Jitendra Kumar, Ranjan Raj, Alam K... | India | September 24, 2021 | 2021 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8802 | s8803 | Movie | Zodiac | David Fincher | Mark Ruffalo, Jake Gyllenhaal, Robert Downey J... | United States | November 20, 2019 | 2007 |
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 | 2018 |
| 8804 | s8805 | Movie | Zombieland | Ruben Fleischer | Jesse Eisenberg, Woody Harrelson, Emma Stone, ... | United States | November 1, 2019 | 2009 |
| 8805 | s8806 | Movie | Zoom | Peter Hewitt | Tim Allen, Courteney Cox, Chevy Chase, Kate Ma... | United States | January 11, 2020 | 2006 |

| show_id | type | title | director | cast | country | date_added | release_year | |
|---------|-------|-------|----------|-------------|---|------------|---------------|------|
| 8806 | s8807 | Movie | Zubaan | Mozez Singh | Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan... | India | March 2, 2019 | 2015 |

8807 rows × 12 columns

```
In [5]: ┌─ 1 ─┐ n_data[n_data['type']=='Movie']
```

Out[5]:

| | show_id | type | title | director | cast | country | date_added | release_1 |
|------|---------|-------|----------------------------------|-------------------------------|---|---|--------------------|-----------|
| 0 | s1 | Movie | Dick Johnson Is Dead | Kirsten Johnson | NaN | United States | September 25, 2021 | 2 |
| 6 | s7 | Movie | My Little Pony: A New Generation | Robert Cullen, José Luis Ucha | Vanessa Hudgens, Kimiko Glenn, James Marsden, ... | NaN | September 24, 2021 | 2 |
| 7 | s8 | Movie | Sankofa | Haile Gerima | Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D... | United States, Ghana, Burkina Faso, United Kin... | September 24, 2021 | 1 |
| 9 | s10 | Movie | The Starling | Theodore Melfi | Melissa McCarthy, Chris O'Dowd, Kevin Kline, T... | United States | September 24, 2021 | 2 |
| 12 | s13 | Movie | Je Suis Karl | Christian Schwochow | Luna Wedler, Jannis Niewöhner, Milan Peschel, ... | Germany, Czech Republic | September 23, 2021 | 2 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8801 | s8802 | Movie | Zinzana | Majid Al Ansari | Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ... | United Arab Emirates, Jordan | March 9, 2016 | 2 |
| 8802 | s8803 | Movie | Zodiac | David Fincher | Mark Ruffalo, Jake Gyllenhaal, Robert Downey J... | United States | November 20, 2019 | 2 |
| 8804 | s8805 | Movie | Zombieland | Ruben Fleischer | Jesse Eisenberg, Woody Harrelson, Emma Stone, ... | United States | November 1, 2019 | 2 |
| 8805 | s8806 | Movie | Zoom | Peter Hewitt | Tim Allen, Courteney Cox, Chevy Chase, Kate Ma... | United States | January 11, 2020 | 2 |

| show_id | type | title | director | cast | country | date_added | release_> |
|---------|-------|-------|----------|-------------|--|------------|---------------|
| 8806 | s8807 | Movie | Zubaan | Mozez Singh | Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanani... | India | March 2, 2019 |

6131 rows × 12 columns

In [6]: ┌ 1 movies=n_data[n_data['type']=='Movie']

In [7]:  1 movies

Out[7]:

| | show_id | type | title | director | cast | country | date_added | release_1 |
|------|---------|-------|----------------------------------|-------------------------------|---|---|--------------------|-----------|
| 0 | s1 | Movie | Dick Johnson Is Dead | Kirsten Johnson | Nan | United States | September 25, 2021 | 2 |
| 6 | s7 | Movie | My Little Pony: A New Generation | Robert Cullen, José Luis Ucha | Vanessa Hudgens, Kimiko Glenn, James Marsden, ... | Nan | September 24, 2021 | 2 |
| 7 | s8 | Movie | Sankofa | Haile Gerima | Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D... | United States, Ghana, Burkina Faso, United Kin... | September 24, 2021 | 1 |
| 9 | s10 | Movie | The Starling | Theodore Melfi | Melissa McCarthy, Chris O'Dowd, Kevin Kline, T... | United States | September 24, 2021 | 2 |
| 12 | s13 | Movie | Je Suis Karl | Christian Schwochow | Luna Wedler, Jannis Niewöhner, Milan Peschel, ... | Germany, Czech Republic | September 23, 2021 | 2 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8801 | s8802 | Movie | Zinzana | Majid Al Ansari | Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ... | United Arab Emirates, Jordan | March 9, 2016 | 2 |
| 8802 | s8803 | Movie | Zodiac | David Fincher | Mark Ruffalo, Jake Gyllenhaal, Robert Downey J... | United States | November 20, 2019 | 2 |
| 8804 | s8805 | Movie | Zombieland | Ruben Fleischer | Jesse Eisenberg, Woody Harrelson, Emma Stone, ... | United States | November 1, 2019 | 2 |
| 8805 | s8806 | Movie | Zoom | Peter Hewitt | Tim Allen, Courteney Cox, Chevy Chase, Kate Ma... | United States | January 11, 2020 | 2 |

| show_id | type | title | director | cast | country | date_added | release_> |
|---------|-------|-------|----------|-------------|--|------------|---------------|
| 8806 | s8807 | Movie | Zubaan | Mozez Singh | Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanani... | India | March 2, 2019 |

6131 rows × 12 columns

In [8]: ┌ 1 TV_shows=n_data[n_data['type']=='TV Show']

In [9]:  1 TV_shows

Out[9]:

| | | show_id | type | title | director | cast | country | date_added | release_year |
|------|-------|---------|-----------------------|-----------------|---|---|--------------------|------------|--------------|
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | South Africa | September 24, 2021 | 2021 | |
| 2 | s3 | TV Show | Ganglands | Julien Leclercq | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... | NaN | September 24, 2021 | 2021 | |
| 3 | s4 | TV Show | Jailbirds New Orleans | NaN | NaN | NaN | September 24, 2021 | 2021 | |
| 4 | s5 | TV Show | Kota Factory | NaN | Mayur More, Jitendra Kumar, Ranjan Raj, Alam K... | India | September 24, 2021 | 2021 | |
| 5 | s6 | TV Show | Midnight Mass | Mike Flanagan | Kate Siegel, Zach Gilford, Hamish Linklater, H... | NaN | September 24, 2021 | 2021 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8795 | s8796 | TV Show | Yu-Gi-Oh! Arc-V | NaN | Mike Liscio, Emily Bauer, Billy Bob Thompson, ... | Japan, Canada | May 1, 2018 | 2015 | |
| 8796 | s8797 | TV Show | Yunus Emre | NaN | Gökhan Atalay, Payidar Tüfekçioğlu, Baran Akbu... | Turkey | January 17, 2017 | 2016 | |
| 8797 | s8798 | TV Show | Zak Storm | NaN | Michael Johnston, Jessica Gee-George, Christin... | United States, France, South Korea, Indonesia | September 13, 2018 | 2016 | |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | NaN | Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ... | Pakistan | December 15, 2016 | 2012 | |

| show_id | type | title | director | cast | country | date_added | release_year |
|---------|-------|---------|-------------|------|---------|------------|--------------|
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 |

2676 rows × 12 columns

In [10]: 1 TV_shows.info()

```
<class 'pandas.core.frame.DataFrame'>
Index: 2676 entries, 1 to 8803
Data columns (total 12 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   show_id     2676 non-null    object 
 1   type        2676 non-null    object 
 2   title       2676 non-null    object 
 3   director    230 non-null    object 
 4   cast         2326 non-null   object 
 5   country     2285 non-null   object 
 6   date_added  2666 non-null   object 
 7   release_year 2676 non-null   int64  
 8   rating      2674 non-null   object 
 9   duration    2676 non-null   object 
 10  listed_in   2676 non-null   object 
 11  description 2676 non-null   object 
dtypes: int64(1), object(11)
memory usage: 271.8+ KB
```

In [11]: 1 TV_shows.isna().sum()

```
Out[11]: show_id      0
type          0
title         0
director     2446
cast          350
country       391
date_added   10
release_year  0
rating        2
duration      0
listed_in    0
description   0
dtype: int64
```

missing values

In [91]: 1 TV_shows.info()

```
<class 'pandas.core.frame.DataFrame'>
Index: 2676 entries, 1 to 8803
Data columns (total 12 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   show_id     2676 non-null   object 
 1   type        2676 non-null   object 
 2   title       2676 non-null   object 
 3   director    230 non-null   object 
 4   cast         2326 non-null   object 
 5   country     2285 non-null   object 
 6   date_added  2666 non-null   object 
 7   release year 2676 non-null   int64  
 8   rating      2674 non-null   object 
 9   duration    2676 non-null   object 
 10  listed_in   2676 non-null   object 
 11  description 2676 non-null   object 
dtypes: int64(1), object(11)
memory usage: 271.8+ KB
```

missing values

In [92]: 1 TV_shows.isna().sum()

```
Out[92]: show_id      0
          type        0
          title       0
          director   2446
          cast        350
          ...
```

In [12]: ┌── 1 TV_shows
 2

Out[12]:

| | | show_id | type | title | director | cast | country | date_added | release_year |
|------|-------|---------|-----------------------|-----------------|---|---|--------------------|------------|--------------|
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | South Africa | September 24, 2021 | 2021 | |
| 2 | s3 | TV Show | Ganglands | Julien Leclercq | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... | NaN | September 24, 2021 | 2021 | |
| 3 | s4 | TV Show | Jailbirds New Orleans | NaN | NaN | NaN | September 24, 2021 | 2021 | |
| 4 | s5 | TV Show | Kota Factory | NaN | Mayur More, Jitendra Kumar, Ranjan Raj, Alam K... | India | September 24, 2021 | 2021 | |
| 5 | s6 | TV Show | Midnight Mass | Mike Flanagan | Kate Siegel, Zach Gilford, Hamish Linklater, H... | NaN | September 24, 2021 | 2021 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8795 | s8796 | TV Show | Yu-Gi-Oh! Arc-V | NaN | Mike Liscio, Emily Bauer, Billy Bob Thompson, ... | Japan, Canada | May 1, 2018 | 2015 | |
| 8796 | s8797 | TV Show | Yunus Emre | NaN | Gökhan Atalay, Payidar Tüfekçioğlu, Baran Akbu... | Turkey | January 17, 2017 | 2016 | |
| 8797 | s8798 | TV Show | Zak Storm | NaN | Michael Johnston, Jessica Gee-George, Christin... | United States, France, South Korea, Indonesia | September 13, 2018 | 2016 | |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | NaN | Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ... | Pakistan | December 15, 2016 | 2012 | |

| show_id | type | title | director | cast | country | date_added | release_year |
|---------|-------|---------|-------------|------|---------|------------|--------------|
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 |

2676 rows × 12 columns

In [13]: 1 t1=TV_shows[['title','director']]

In [14]: 1 t1

Out[14]:

| | title | director |
|------|-----------------------|-----------------|
| 1 | Blood & Water | NaN |
| 2 | Ganglands | Julien Leclercq |
| 3 | Jailbirds New Orleans | NaN |
| 4 | Kota Factory | NaN |
| 5 | Midnight Mass | Mike Flanagan |
| ... | ... | ... |
| 8795 | Yu-Gi-Oh! Arc-V | NaN |
| 8796 | Yunus Emre | NaN |
| 8797 | Zak Storm | NaN |
| 8800 | Zindagi Gulzar Hai | NaN |
| 8803 | Zombie Dumb | NaN |

2676 rows × 2 columns

In [15]: 1 TV_shows['director'] = TV_shows['director'].str.split(', ')

C:\Users\Dell\AppData\Local\Temp\ipykernel_12272\2177744900.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

TV_shows['director'] = TV_shows['director'].str.split(', ')

```
In [16]: 1 TV_shows['country'] = TV_shows['country'].str.split(', ')
```

C:\Users\DELL\AppData\Local\Temp\ipykernel_12272\1403327742.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy ([http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

```
TV_shows['country'] = TV_shows['country'].str.split(', ')
```

```
In [17]: 1 TV_shows['listed_in'] = TV_shows['listed_in'].str.split(', ')
```

C:\Users\DELL\AppData\Local\Temp\ipykernel_12272\3621196347.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy ([http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

```
TV_shows['listed_in'] = TV_shows['listed_in'].str.split(', ')
```

In [18]:  1 TV_shows

Out[18]:

| | show_id | type | title | director | cast | country | date_added | release_year |
|------|---------|---------|-----------------------|-------------------|--|---|--------------------|--------------|
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamatia, Khosi Ngema, Gail Mabalane, Thaban... | [South Africa] | September 24, 2021 | 202 |
| 2 | s3 | TV Show | Ganglands | [Julien Leclercq] | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... | NaN | September 24, 2021 | 202 |
| 3 | s4 | TV Show | Jailbirds New Orleans | NaN | NaN | NaN | September 24, 2021 | 202 |
| 4 | s5 | TV Show | Kota Factory | NaN | Mayur More, Jitendra Kumar, Ranjan Raj, Alam K... | [India] | September 24, 2021 | 202 |
| 5 | s6 | TV Show | Midnight Mass | [Mike Flanagan] | Kate Siegel, Zach Gilford, Hamish Linklater, H... | NaN | September 24, 2021 | 202 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8795 | s8796 | TV Show | Yu-Gi-Oh! Arc-V | NaN | Mike Liscio, Emily Bauer, Billy Bob Thompson, ... | [Japan, Canada] | May 1, 2018 | 201 |
| 8796 | s8797 | TV Show | Yunus Emre | NaN | Gökhan Atalay, Payidar Tüfekçioglu, Baran Akbu... | [Turkey] | January 17, 2017 | 201 |
| 8797 | s8798 | TV Show | Zak Storm | NaN | Michael Johnston, Jessica Gee-George, Christin... | [United States, France, South Korea, Indonesia] | September 13, 2018 | 201 |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | NaN | Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ... | [Pakistan] | December 15, 2016 | 201 |

| show_id | type | title | director | cast | country | date_added | release_year |
|---------|-------|---------|-------------|------|---------|--------------|--------------|
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | July 1, 2019 | 201 |

2676 rows × 12 columns

In [19]: 1 TV_shows['director'].unique

```
Out[19]: <bound method Series.unique of 1
          2      [Julien Leclercq]
          3      NaN
          4      NaN
          5      [Mike Flanagan]
          ...
8795      NaN
8796      NaN
8797      NaN
8800      NaN
8803      NaN
Name: director, Length: 2676, dtype: object>
```

In [20]: 1 TV_shows['director'].value_counts()

```
Out[20]: director
          [Alastair Fothergill]    3
          [Rob Seidenglanz]        2
          [Hsu Fu-chun]           2
          [Iginio Straffi]         2
          [Shin Won-ho]            2
          ..
          [Juliana Vicente]        1
          [Chang Chin-jung, Chen Rong-hui] 1
          [Thierry Demaizière, Alban Teurlai] 1
          [Manolo Caro]             1
          [Michael Cumming]         1
Name: count, Length: 222, dtype: int64
```

In []: 1

In []: 1

In [21]: 1 t2=TV_shows.explode('director')

In [22]:  1 t2

Out[22]:

| | | show_id | type | title | director | cast | country | date_added | release_yea |
|------|-------|---------|-----------------------|-----------------|---|---|--------------------|------------|-------------|
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | [South Africa] | September 24, 2021 | 2021 | |
| 2 | s3 | TV Show | Ganglands | Julien Leclercq | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... | NaN | September 24, 2021 | 2021 | |
| 3 | s4 | TV Show | Jailbirds New Orleans | NaN | NaN | NaN | September 24, 2021 | 2021 | |
| 4 | s5 | TV Show | Kota Factory | NaN | Mayur More, Jitendra Kumar, Ranjan Raj, Alam K... | [India] | September 24, 2021 | 2021 | |
| 5 | s6 | TV Show | Midnight Mass | Mike Flanagan | Kate Siegel, Zach Gilford, Hamish Linklater, H... | NaN | September 24, 2021 | 2021 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 8795 | s8796 | TV Show | Yu-Gi-Oh! Arc-V | NaN | Mike Liscio, Emily Bauer, Billy Bob Thompson, ... | [Japan, Canada] | May 1, 2018 | 2018 | |
| 8796 | s8797 | TV Show | Yunus Emre | NaN | Gökhan Atalay, Payidar Tüfekçioğlu, Baran Akbu... | [Turkey] | January 17, 2017 | 2016 | |
| 8797 | s8798 | TV Show | Zak Storm | NaN | Michael Johnston, Jessica Gee-George, Christin... | [United States, France, South Korea, Indonesia] | September 13, 2018 | 2016 | |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | NaN | Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ... | [Pakistan] | December 15, 2016 | 2016 | |

| show_id | type | title | director | cast | country | date_added | release_yea |
|---------|-------|---------|-------------|------|---------|------------|--------------|
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 |

2758 rows × 12 columns

In [23]: ┌ 1 t3=t2.explode('country')

In [24]:  1 t3

Out[24]:

| | show_id | type | title | director | cast | country | date_added | release_year |
|------|---------|---------|-----------------------|-----------------|---|--------------|--------------------|--------------|
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | South Africa | September 24, 2021 | 2021 |
| 2 | s3 | TV Show | Ganglands | Julien Leclercq | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... | NaN | September 24, 2021 | 2021 |
| 3 | s4 | TV Show | Jailbirds New Orleans | NaN | NaN | NaN | September 24, 2021 | 2021 |
| 4 | s5 | TV Show | Kota Factory | NaN | Mayur More, Jitendra Kumar, Ranjan Raj, Alam K... | India | September 24, 2021 | 2021 |
| 5 | s6 | TV Show | Midnight Mass | Mike Flanagan | Kate Siegel, Zach Gilford, Hamish Linklater, H... | NaN | September 24, 2021 | 2021 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8797 | s8798 | TV Show | Zak Storm | NaN | Michael Johnston, Jessica Gee-George, Christin... | France | September 13, 2018 | 2016 |
| 8797 | s8798 | TV Show | Zak Storm | NaN | Michael Johnston, Jessica Gee-George, Christin... | South Korea | September 13, 2018 | 2016 |
| 8797 | s8798 | TV Show | Zak Storm | NaN | Michael Johnston, Jessica Gee-George, Christin... | Indonesia | September 13, 2018 | 2016 |

| show_id | type | title | director | cast | country | date_added | release_year | |
|---------|-------|---------|--------------------|------|---|------------|-------------------|------|
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | NaN | Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ... | Pakistan | December 15, 2016 | 2012 |
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 | 2018 |

3147 rows × 12 columns

In [25]: ┶ 1 t4=t3.explode('listed_in')

In [26]:  1 t4

Out[26]:

| | show_id | type | title | director | cast | country | date_added | release_year | r |
|------|---------|---------|--------------------|-----------------|---|--------------|--------------------|--------------|-----|
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | South Africa | September 24, 2021 | 2021 | |
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | South Africa | September 24, 2021 | 2021 | |
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban... | South Africa | September 24, 2021 | 2021 | |
| 2 | s3 | TV Show | Ganglands | Julien Leclercq | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... | NaN | September 24, 2021 | 2021 | |
| 2 | s3 | TV Show | Ganglands | Julien Leclercq | Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi... | NaN | September 24, 2021 | 2021 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | NaN | Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ... | Pakistan | December 15, 2016 | 2012 | |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | NaN | Sanam Saeed, Fawad Khan, Ayesha Omer, Mehreen ... | Pakistan | December 15, 2016 | 2012 | |
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 | 2018 | T |

| show_id | type | title | director | cast | country | date_added | release_year | rating |
|---------|-------|---------|-------------|------|---------|------------|--------------|--------|
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 | 2018 T |
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 | 2018 T |

7090 rows × 12 columns

In [27]: 1 t4['cast'] = t4['cast'].str.split(', ')

In [28]: 1 t4.head()

Out[28]:

| show_id | type | title | director | cast | country | date_added | release_year | rating |
|---------|------|---------|---------------|---|--------------|--------------------|--------------|--------|
| 1 | s2 | TV Show | Blood & Water | [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba... | South Africa | September 24, 2021 | 2021 | TV M/ |
| 1 | s2 | TV Show | Blood & Water | [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba... | South Africa | September 24, 2021 | 2021 | TV M/ |
| 1 | s2 | TV Show | Blood & Water | [Ama Qamata, Khosi Ngema, Gail Mabalane, Thaba... | South Africa | September 24, 2021 | 2021 | TV M/ |
| 2 | s3 | TV Show | Ganglands | Julien Leclercq [Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nab... | NaN | September 24, 2021 | 2021 | TV M/ |
| 2 | s3 | TV Show | Ganglands | Julien Leclercq [Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nab... | NaN | September 24, 2021 | 2021 | TV M/ |



```
In [29]: ┌─ 1 ─┐ t5=t4.explode('cast')
```

In [30]:  1 t5

Out[30]:

| | show_id | type | title | director | cast | country | date_added | release_year | rating |
|------|---------|---------|--------------------|----------|--------------------|--------------|--------------------|--------------|--------|
| 1 | s2 | TV Show | Blood & Water | NaN | Ama Qamata | South Africa | September 24, 2021 | 2021 | TV M |
| 1 | s2 | TV Show | Blood & Water | NaN | Khosi Ngema | South Africa | September 24, 2021 | 2021 | TV M |
| 1 | s2 | TV Show | Blood & Water | NaN | Gail Mabalane | South Africa | September 24, 2021 | 2021 | TV M |
| 1 | s2 | TV Show | Blood & Water | NaN | Thabang Molaba | South Africa | September 24, 2021 | 2021 | TV M |
| 1 | s2 | TV Show | Blood & Water | NaN | Dillon Windvogel | South Africa | September 24, 2021 | 2021 | TV M |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | NaN | Javed Sheikh | Pakistan | December 15, 2016 | 2012 | TV P |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | NaN | Hina Khawaja Bayat | Pakistan | December 15, 2016 | 2012 | TV P |
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 | 2018 | TV-Y |
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | NaN | July 1, 2019 | 2018 | TV-Y |

| show_id | type | title | director | cast | country | date_added | release_year | rating |
|---------|-------|---------|-------------|------|---------|--------------|--------------|--------|
| 8803 | s8804 | TV Show | Zombie Dumb | NaN | NaN | July 1, 2019 | 2018 | TV-Y |

56148 rows × 12 columns

In [31]: 1 t5.info()

```
<class 'pandas.core.frame.DataFrame'>
Index: 56148 entries, 1 to 8803
Data columns (total 12 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   show_id     56148 non-null   object 
 1   type        56148 non-null   object 
 2   title       56148 non-null   object 
 3   director    6790 non-null   object 
 4   cast         55330 non-null  object 
 5   country     50450 non-null   object 
 6   date_added  55990 non-null  object 
 7   release_year 56148 non-null  int64  
 8   rating      56090 non-null  object 
 9   duration    56148 non-null  object 
 10  listed_in   56148 non-null  object 
 11  description  56148 non-null  object 
dtypes: int64(1), object(11)
memory usage: 5.6+ MB
```

In [32]: 1 t5.isna().sum()

```
Out[32]: show_id      0
type          0
title         0
director     49358
cast          818
country       5698
date_added   158
release_year  0
rating        58
duration      0
listed_in     0
description   0
dtype: int64
```

In [33]: 1 t5['director'].isna().sum()

Out[33]: 49358

```
In [34]: ┌─ 1 ─┐ t5['director'].fillna('unknown',inplace=True)
```

```
In [35]: ┌─ 1 ─┐ t5['director'].isna().sum()
```

```
Out[35]: 0
```

```
In [36]: ┌─ 1 ─┐ t5['cast'].isna().sum()
```

```
Out[36]: 818
```

```
In [37]: ┌─ 1 ─┐ t5['cast'].fillna('unknown',inplace=True)
```

```
In [38]: ┌─ 1 ─┐ t5['cast'].isna().sum()
```

```
Out[38]: 0
```

```
In [39]: ┌─ 1 ─┐ t5['country'].isna().sum()
```

```
Out[39]: 5698
```

```
In [40]: ┌─ 1 ─┐ t5['country'].fillna('unknown',inplace=True)
```

```
In [41]: ┌─ 1 ─┐ t5['country'].isna().sum()
```

```
Out[41]: 0
```

```
In [42]: ┌─ 1 ─┐ t5.isna().sum()
```

```
Out[42]: show_id          0  
type            0  
title           0  
director        0  
cast             0  
country          0  
date_added     158  
release_year    0  
rating          58  
duration         0  
listed_in        0  
description      0  
dtype: int64
```

```
In [43]: ┌─ 1 ─┐ # t5.dropna(inplace=True)
```

```
In [44]: ┌─ 1 ─┐ t5['date_added']=t5['date_added'].fillna('January 1, 1900')
```

```
In [45]: ┌─ 1 ─┐ t5.isna().sum()
```

```
Out[45]: show_id      0
          type        0
          title       0
          director    0
          cast        0
          country     0
          date_added  0
          release_year 0
          rating      58
          duration    0
          listed_in   0
          description  0
          dtype: int64
```

```
In [46]: ┌─ 1 ─┐ t5['rating']=t5['rating'].fillna(0)
```

```
In [47]: ┌─ 1 ─┐ t5.isna().sum()
```

```
Out[47]: show_id      0
          type        0
          title       0
          director    0
          cast        0
          country     0
          date_added  0
          release_year 0
          rating      0
          duration    0
          listed_in   0
          description  0
          dtype: int64
```

In [48]:  1 t5

Out[48]:

| | show_id | type | title | director | cast | country | date_added | release_year | rating |
|------|---------|---------|--------------------|----------|--------------------|--------------|--------------------|--------------|--------|
| 1 | s2 | TV Show | Blood & Water | unknown | Ama Qamata | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Khosi Ngema | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Gail Mabalane | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Thabang Molaba | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Dillon Windvogel | South Africa | September 24, 2021 | 2021 | T |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Javed Sheikh | Pakistan | December 15, 2016 | 2012 | T |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Hina Khawaja Bayat | Pakistan | December 15, 2016 | 2012 | T |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV- |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV- |

| show_id | type | title | director | cast | country | date_added | release_year | rating |
|---------|-------|---------|-------------|---------|---------|------------|--------------|----------|
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 TV- |

56148 rows × 12 columns

In [49]: 1 t5.info()

```
<class 'pandas.core.frame.DataFrame'>
Index: 56148 entries, 1 to 8803
Data columns (total 12 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   show_id     56148 non-null   object 
 1   type        56148 non-null   object 
 2   title       56148 non-null   object 
 3   director    56148 non-null   object 
 4   cast        56148 non-null   object 
 5   country     56148 non-null   object 
 6   date_added  56148 non-null   object 
 7   release_year 56148 non-null   int64  
 8   rating      56148 non-null   object 
 9   duration    56148 non-null   object 
 10  listed_in   56148 non-null   object 
 11  description 56148 non-null   object 
dtypes: int64(1), object(11)
memory usage: 5.6+ MB
```

In [50]: 1 t5['duration']=t5['duration'].str.split(' ')

In [51]:  1 t5

Out[51]:

| | show_id | type | title | director | cast | country | date_added | release_year | rating |
|------|---------|---------|--------------------|----------|--------------------|--------------|--------------------|--------------|--------|
| 1 | s2 | TV Show | Blood & Water | unknown | Ama Qamata | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Khosi Ngema | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Gail Mabalane | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Thabang Molaba | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Dillon Windvogel | South Africa | September 24, 2021 | 2021 | T |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Javed Sheikh | Pakistan | December 15, 2016 | 2012 | T |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Hina Khawaja Bayat | Pakistan | December 15, 2016 | 2012 | T |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV- |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV- |

| show_id | type | title | director | cast | country | date_added | release_year | rating |
|---------|-------|---------|-------------|---------|---------|------------|--------------|----------|
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 TV- |

56148 rows × 12 columns

In [52]: ┌ 1 t5['duration']=t5['duration'].apply(lambda x: x[0])

In [53]:  1 t5

Out[53]:

| | show_id | type | title | director | cast | country | date_added | release_year | rating |
|------|---------|---------|--------------------|----------|--------------------|--------------|--------------------|--------------|--------|
| 1 | s2 | TV Show | Blood & Water | unknown | Ama Qamata | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Khosi Ngema | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Gail Mabalane | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Thabang Molaba | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Dillon Windvogel | South Africa | September 24, 2021 | 2021 | T |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Javed Sheikh | Pakistan | December 15, 2016 | 2012 | T F |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Hina Khawaja Bayat | Pakistan | December 15, 2016 | 2012 | T |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV- |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV- |

| show_id | type | title | director | cast | country | date_added | release_year | rating |
|---------|-------|---------|-------------|---------|---------|------------|--------------|----------|
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 TV- |

56148 rows × 12 columns

```
In [54]: ┌─[1]─┐ t5['duration']=t5['duration'].apply(lambda x: int(x))
```

In [55]:  1 t5

Out[55]:

| | show_id | type | title | director | cast | country | date_added | release_year | rating |
|------|---------|---------|--------------------|----------|--------------------|--------------|--------------------|--------------|--------|
| 1 | s2 | TV Show | Blood & Water | unknown | Ama Qamata | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Khosi Ngema | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Gail Mabalane | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Thabang Molaba | South Africa | September 24, 2021 | 2021 | T |
| 1 | s2 | TV Show | Blood & Water | unknown | Dillon Windvogel | South Africa | September 24, 2021 | 2021 | T |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Javed Sheikh | Pakistan | December 15, 2016 | 2012 | T F |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Hina Khawaja Bayat | Pakistan | December 15, 2016 | 2012 | T |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV- |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV- |

| show_id | type | title | director | cast | country | date_added | release_year | rating |
|---------|-------|---------|-------------|---------|---------|------------|--------------|----------|
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 TV- |

56148 rows × 12 columns

In [56]: 1 t5['duration'].mean()

Out[56]: 1.9281007337750231

In [57]: 1 t5.groupby('title')['title'].count()

Out[57]: title

| | |
|------------------------------------|----|
| #blackAF | 7 |
| (Un)Well | 1 |
| 100 Days My Prince | 18 |
| 100 Humans | 6 |
| 100% Hotter | 12 |
| .. | |
| แผนร้ายนายเจ้าเล่ห์ | 12 |
| SAINT SEIYA: Knights of the Zodiac | 24 |
| 忍者ハットリくん | 2 |
| 海的兒子 | 10 |
| 마녀사냥 | 15 |

Name: title, Length: 2676, dtype: int64

In [58]: 1 duplicates=t5[t5.duplicated()]

In [59]: 1 duplicates

Out[59]:

| show_id | type | title | director | cast | country | date_added | release_year | rating | duration |
|---------|------|-------|----------|------|---------|------------|--------------|--------|----------|
| | | | | | | | | | |

In [60]: 1 t6=t5.drop_duplicates()

In [61]: ┌ 1 t6.tail()

Out[61]:

| | show_id | type | title | director | cast | country | date_added | release_year | rating |
|------|---------|---------|--------------------|----------|--------------------|----------|-------------------|--------------|--------|
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Javed Sheikh | Pakistan | December 15, 2016 | 2012 | TV PG |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Hina Khawaja Bayat | Pakistan | December 15, 2016 | 2012 | TV PG |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV-Y |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV-Y |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | July 1, 2019 | 2018 | TV-Y |

◀ ▶

```
In [62]: ┌ 1 # t6['date_added']=pd.to_datetime(t6['date_added'])
  2 t6['date_added']=pd.to_datetime(t6['date_added'],format='mixed')
  3 # t6['date_added']=pd.to_datetime(t6['date_added'],format='ISO8601')
  4 # pd.to_datetime?
```

In [63]:  1 t6

Out[63]:

| | show_id | type | title | director | cast | country | date_added | release_year | rating |
|------|---------|---------|--------------------|----------|--------------------|--------------|------------|--------------|--------|
| 1 | s2 | TV Show | Blood & Water | unknown | Ama Qamata | South Africa | 2021-09-24 | 2021 | T M |
| 1 | s2 | TV Show | Blood & Water | unknown | Khosi Ngema | South Africa | 2021-09-24 | 2021 | T M |
| 1 | s2 | TV Show | Blood & Water | unknown | Gail Mabalane | South Africa | 2021-09-24 | 2021 | T M |
| 1 | s2 | TV Show | Blood & Water | unknown | Thabang Molaba | South Africa | 2021-09-24 | 2021 | T M |
| 1 | s2 | TV Show | Blood & Water | unknown | Dillon Windvogel | South Africa | 2021-09-24 | 2021 | T M |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Javed Sheikh | Pakistan | 2016-12-15 | 2012 | T F |
| 8800 | s8801 | TV Show | Zindagi Gulzar Hai | unknown | Hina Khawaja Bayat | Pakistan | 2016-12-15 | 2012 | T F |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | 2019-07-01 | 2018 | TV- |
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | 2019-07-01 | 2018 | TV- |

| show_id | type | title | director | cast | country | date_added | release_year | rating |
|---------|-------|---------|-------------|---------|---------|------------|--------------|----------|
| 8803 | s8804 | TV Show | Zombie Dumb | unknown | unknown | unknown | 2019-07-01 | 2018 TV- |

56148 rows × 12 columns

In [64]: 1 t6.groupby(['country', 'title'])['title'].count()

```
Out[64]: country      title
          D.P.                  8
Argentina  Almost Happy           9
          Apache: The Life of Carlos Tevez    27
          Carmel: Who Killed Maria Marta?     3
          Creators                   6
          ..
unknown    You Are My Spring      63
          Zombie Dumb                 3
          12                      دفعه القاهرة
          แผนร้ายนายเจ้าเล่ห์           12
          海的儿子                     10
Name: title, Length: 3031, dtype: int64
```

In [65]: 1 t7_country=t6.groupby(['country', 'title'])['country'].count()

In [66]: 1 # t7_country.to_excel('t7_countries.xlsx')

In [67]: 1 directors=t6[['show_id', 'director']].drop_duplicates(keep = 'first')[::10]

In [68]: 1 directors

```
Out[68]: director
          unknown            2446
          Alastair Fothergill      3
          Ken Burns                3
          Iginio Straffi           2
          Gautham Vasudev Menon     2
          ...
          Jesse Vile               1
          Ellena Wood              1
          Picky Talarico            1
          Pedro Waddington          1
          Michael Cumming            1
Name: count, Length: 300, dtype: int64
```

In []: 1

```
In [ ]: ┌ 1
```

```
In [ ]: ┌ 1
```

```
In [69]: ┌ 1 t6.groupby('cast')['show_id'].nunique().sort_values(ascending=False).h
```

```
Out[69]: cast
unknown           350
Takahiro Sakurai    25
Yuki Kaji            19
Daisuke Ono           17
Ai Kayano            17
Junichi Suwabe         17
Yuichi Nakamura        16
Yoshimasa Hosoya         15
Jun Fukuyama           15
David Attenborough       14
Name: show_id, dtype: int64
```

```
In [70]: ┌ 1 t6.groupby('country')['show_id'].nunique().sort_values(ascending=False).h
```

```
Out[70]: country
United States      938
unknown             391
United Kingdom      272
Japan                199
South Korea          170
Canada                126
France                 90
India                  84
Taiwan                 70
Australia               66
Name: show_id, dtype: int64
```

```
In [71]: ┌ 1 t6.groupby('listed_in')['show_id'].nunique().sort_values(ascending=False).h
```

```
Out[71]: listed_in
International TV Shows   1351
TV Dramas                 763
TV Comedies                 581
Crime TV Shows              470
Kids' TV                   451
Docuseries                  395
Romantic TV Shows            370
Reality TV                   255
British TV Shows              253
Anime Series                  176
Name: show_id, dtype: int64
```

```
In [72]: ┌─ 1 t6.groupby('director')['show_id'].nunique().sort_values(ascending=False)
```

Out[72]: director
unknown 2446
Ken Burns 3
Alastair Fothergill 3
Jung-ah Im 2
Joe Berlinger 2
Hsu Fu-chun 2
Stan Lathan 2
Gautham Vasudev Menon 2
Lynn Novick 2
Shin Won-ho 2
Name: show_id, dtype: int64

```
In [73]: ┌─ 1 # t6.to_excel('t6B_clean.xlsx')
```

```
In [74]: ┌─ 1 cast_data=t6.groupby('cast')['show_id'].nunique().sort_values(ascending=False)
```

```
In [75]: ┌─ 1 country_data=t6.groupby('country')['show_id'].nunique().sort_values(ascending=False)
```

```
In [76]: ┌─ 1 genre_data=t6.groupby('listed_in')['show_id'].nunique().sort_values(ascending=False)
```

```
In [77]: ┌─ 1 director_data=t6.groupby('director')['show_id'].nunique().sort_values(ascending=False)
```

```
In [78]: ┌─ 1 cast_data.head(10)
```

Out[78]: cast
unknown 350
Takahiro Sakurai 25
Yuki Kaji 19
Daisuke Ono 17
Ai Kayano 17
Junichi Suwabe 17
Yuichi Nakamura 16
Yoshimasa Hosoya 15
Jun Fukuyama 15
David Attenborough 14
Name: show_id, dtype: int64

In [79]: 1 country_data.head(10)

Out[79]: country

| | |
|----------------|-----|
| United States | 938 |
| unknown | 391 |
| United Kingdom | 272 |
| Japan | 199 |
| South Korea | 170 |
| Canada | 126 |
| France | 90 |
| India | 84 |
| Taiwan | 70 |
| Australia | 66 |

Name: show_id, dtype: int64

In [80]: 1 genre_data.head(10)

Out[80]: listed_in

| | |
|------------------------|------|
| International TV Shows | 1351 |
| TV Dramas | 763 |
| TV Comedies | 581 |
| Crime TV Shows | 470 |
| Kids' TV | 451 |
| Docuseries | 395 |
| Romantic TV Shows | 370 |
| Reality TV | 255 |
| British TV Shows | 253 |
| Anime Series | 176 |

Name: show_id, dtype: int64

In [81]: 1 director_data.head(10)

Out[81]: director

| | |
|-----------------------|------|
| unknown | 2446 |
| Ken Burns | 3 |
| Alastair Fothergill | 3 |
| Jung-ah Im | 2 |
| Joe Berlinger | 2 |
| Hsu Fu-chun | 2 |
| Stan Lathan | 2 |
| Gautham Vasudev Menon | 2 |
| Lynn Novick | 2 |
| Shin Won-ho | 2 |

Name: show_id, dtype: int64

In [82]: 1 # cast_data.to_excel('t6B_cast_data_01.xlsx')
2 # country_data.to_excel('t6B_country_data_01.xlsx')
3 # genre_data.to_excel('t6B_genre_data_01.xlsx')
4 # director_data.to_excel('t6B_director_data_01.xlsx')
5

In [83]: 1 t_cast=cast_data.head(10)

```
In [84]: ┌─ 1 t_country=country_data.head(10)
```

```
In [85]: ┌─ 1 t_genre=genre_data.head(10)
```

```
In [86]: ┌─ 1 t_director=director_data.head(10)
```

```
In [87]: ┌─ 1 t_cast
```

```
Out[87]: cast
unknown           350
Takahiro Sakurai    25
Yuki Kaji            19
Daisuke Ono           17
Ai Kayano             17
Junichi Suwabe          17
Yuichi Nakamura          16
Yoshimasa Hosoya          15
Jun Fukuyama             15
David Attenborough          14
Name: show_id, dtype: int64
```

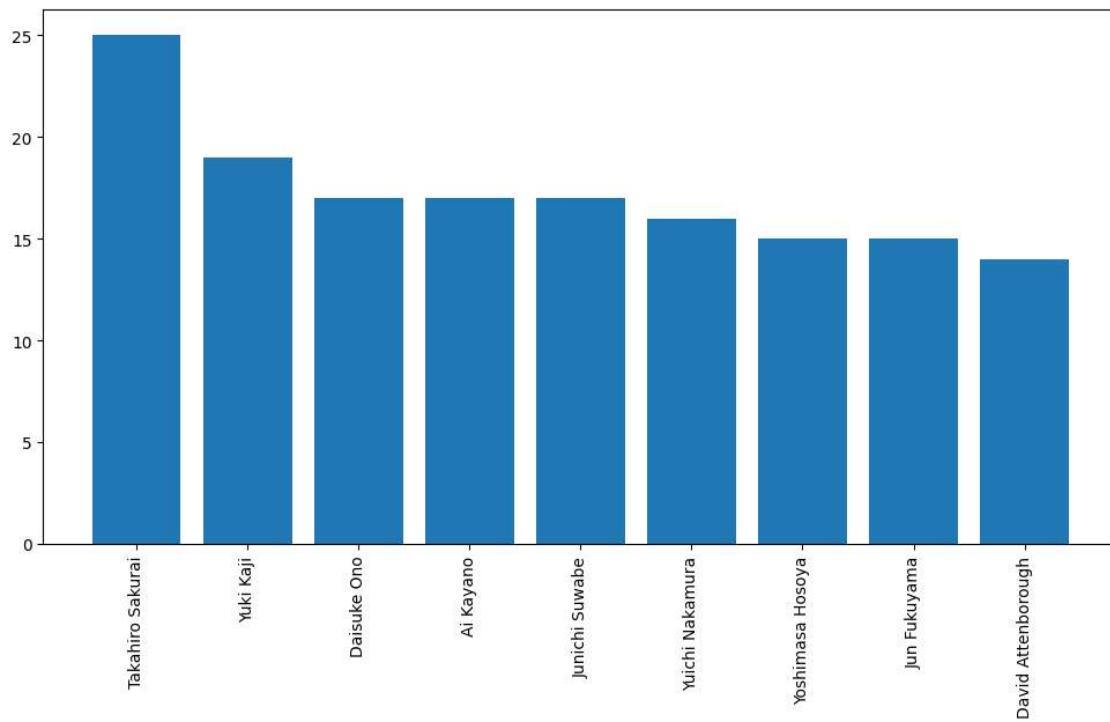
```
In [88]: ┌─ 1 t_cast_01=t_cast[1:]
```

```
In [89]: ┌─ 1 t_cast_01
```

```
Out[89]: cast
Takahiro Sakurai      25
Yuki Kaji                19
Daisuke Ono               17
Ai Kayano                  17
Junichi Suwabe                 17
Yuichi Nakamura                 16
Yoshimasa Hosoya                 15
Jun Fukuyama                   15
David Attenborough                 14
Name: show_id, dtype: int64
```

In [90]:

```
1 plt.figure(figsize=(12,6))
2 plt.bar(t_cast_01.index,t_cast_01)
3 plt.xticks(rotation=90)
4 plt.show()
```



In [91]:

```
1 t_country
```

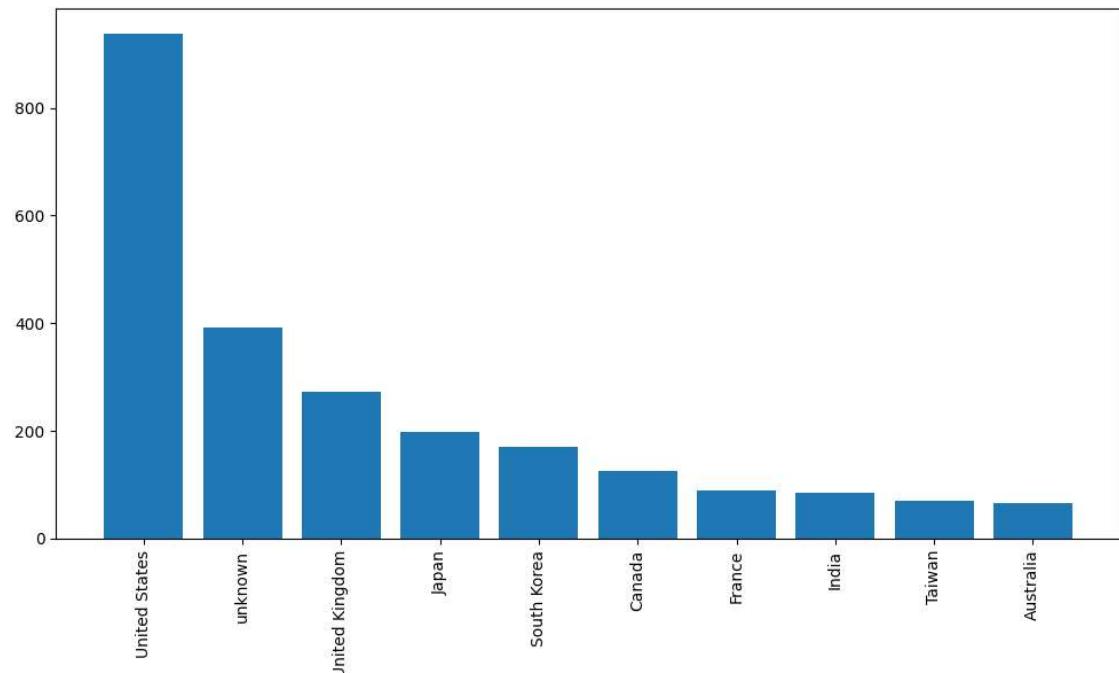
Out[91]:

| country | count |
|----------------|-------|
| United States | 938 |
| unknown | 391 |
| United Kingdom | 272 |
| Japan | 199 |
| South Korea | 170 |
| Canada | 126 |
| France | 90 |
| India | 84 |
| Taiwan | 70 |
| Australia | 66 |

Name: show_id, dtype: int64

In [92]:

```
1 plt.figure(figsize=(12,6))
2 plt.bar(t_country.index,t_country)
3 plt.xticks(rotation=90)
4 plt.show()
```



In [93]:

```
1 t_genre
```

Out[93]:

| listed_in | Count |
|------------------------|-------|
| International TV Shows | 1351 |
| TV Dramas | 763 |
| TV Comedies | 581 |
| Crime TV Shows | 470 |
| Kids' TV | 451 |
| Docuseries | 395 |
| Romantic TV Shows | 370 |
| Reality TV | 255 |
| British TV Shows | 253 |
| Anime Series | 176 |

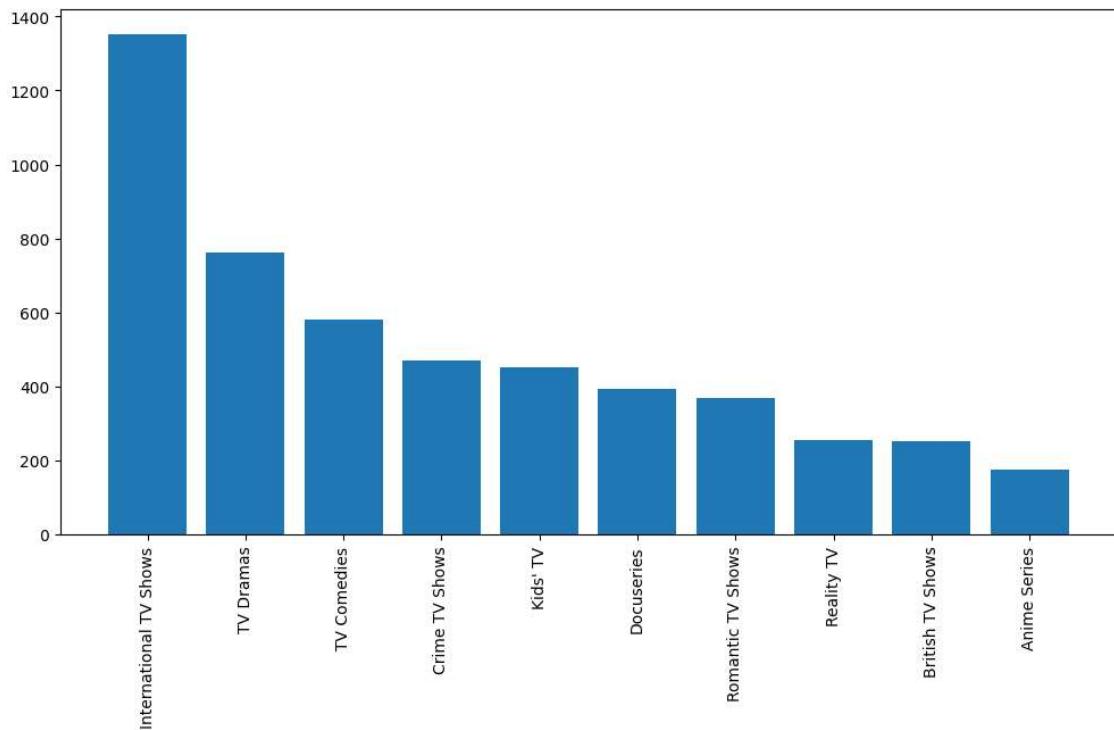
Name: show_id, dtype: int64

In [94]:

```

1 plt.figure(figsize=(12,6))
2 plt.bar(t_genre.index,t_genre)
3 plt.xticks(rotation=90)
4 plt.show()

```



In [95]:

1 t_director

Out[95]: director

| director | count |
|-----------------------|-------|
| unknown | 2446 |
| Ken Burns | 3 |
| Alastair Fothergill | 3 |
| Jung-ah Im | 2 |
| Joe Berlinger | 2 |
| Hsu Fu-chun | 2 |
| Stan Lathan | 2 |
| Gautham Vasudev Menon | 2 |
| Lynn Novick | 2 |
| Shin Won-ho | 2 |

Name: show_id, dtype: int64

In [95]:

1 t_director

Out[95]: director

| director | count |
|-----------------------|-------|
| unknown | 2446 |
| Ken Burns | 3 |
| Alastair Fothergill | 3 |
| Jung-ah Im | 2 |
| Joe Berlinger | 2 |
| Hsu Fu-chun | 2 |
| Stan Lathan | 2 |
| Gautham Vasudev Menon | 2 |
| Lynn Novick | 2 |
| Shin Won-ho | 2 |

Name: show_id, dtype: int64

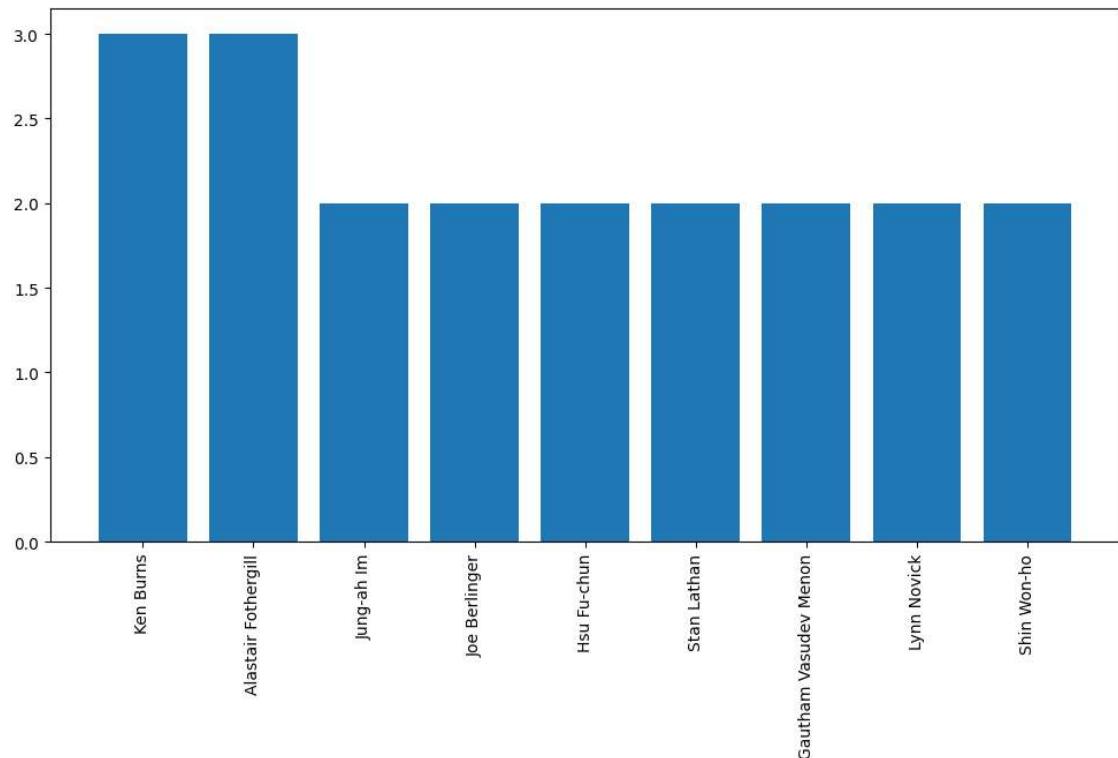
tv shows director names are not available
for major chunk of the data

```
In [96]: ┌ 1 t_director_01=t_director[1:]
```

```
In [97]: ┌ 1 t_director_01
```

```
Out[97]: director
Ken Burns           3
Alastair Fothergill 3
Jung-ah Im          2
Joe Berlinger        2
Hsu Fu-chun         2
Stan Lathan          2
Gautham Vasudev Menon 2
Lynn Novick          2
Shin Won-ho          2
Name: show_id, dtype: int64
```

```
In [98]: ┌ 1 plt.figure(figsize=(12,6))
2 plt.bar(t_director_01.index,t_director_01)
3 plt.xticks(rotation=90)
4 plt.show()
```



```
In [ ]: ┌ 1
```