

NETFLIX DATA ANALYSIS

IMPORT LIBRARIES

In [191...]

```
import numpy as np
import pandas as pd
import plotly.express as px
import plotly.io as pio
pio.renderers.default = "notebook"
```

IMPORT DATA

In [192...]

```
df = pd.read_csv("data/file.csv")
df.head()
```

Out[192]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connelly...	United States	November 16, 2017	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 min	Dramas	A brilliant group of students become card-coun...

In [193...]

```
df.shape
```

Out[193]:

```
(7789, 11)
```

In [194...]

```
df.describe()
```

Out[194]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description
count	7789	7789	7789	5401	7071	7282	7779	7782	7789	7789	7789
unique	7787	2	7787	4050	6831	681	1565	14	216	492	7769
top	s6621	Movie	The Lost Okoroshi	Raúl Campos, Jan Suter	David Attenborough	United States	January 1, 2020	TV-MA	1 Season	Documentaries	Multiple women report their husbands as missin...
freq	2	5379	2	18	18	2556	118	2865	1608	334	3

In [195...]

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7789 entries, 0 to 7788
Data columns (total 11 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Show_Id     7789 non-null    object  
 1   Category    7789 non-null    object  
 2   Title       7789 non-null    object  
 3   Director    5401 non-null    object  
 4   Cast        7071 non-null    object  
 5   Country     7282 non-null    object  
 6   Release_Date 7779 non-null    object  
 7   Rating      7782 non-null    object  
 8   Duration    7789 non-null    object  
 9   Type        7789 non-null    object  
 10  Description 7789 non-null    object  
dtypes: object(11)
memory usage: 669.5+ KB
```

Distribution by Rating

In [196...]

```
x = df['Rating'].value_counts()  
x
```

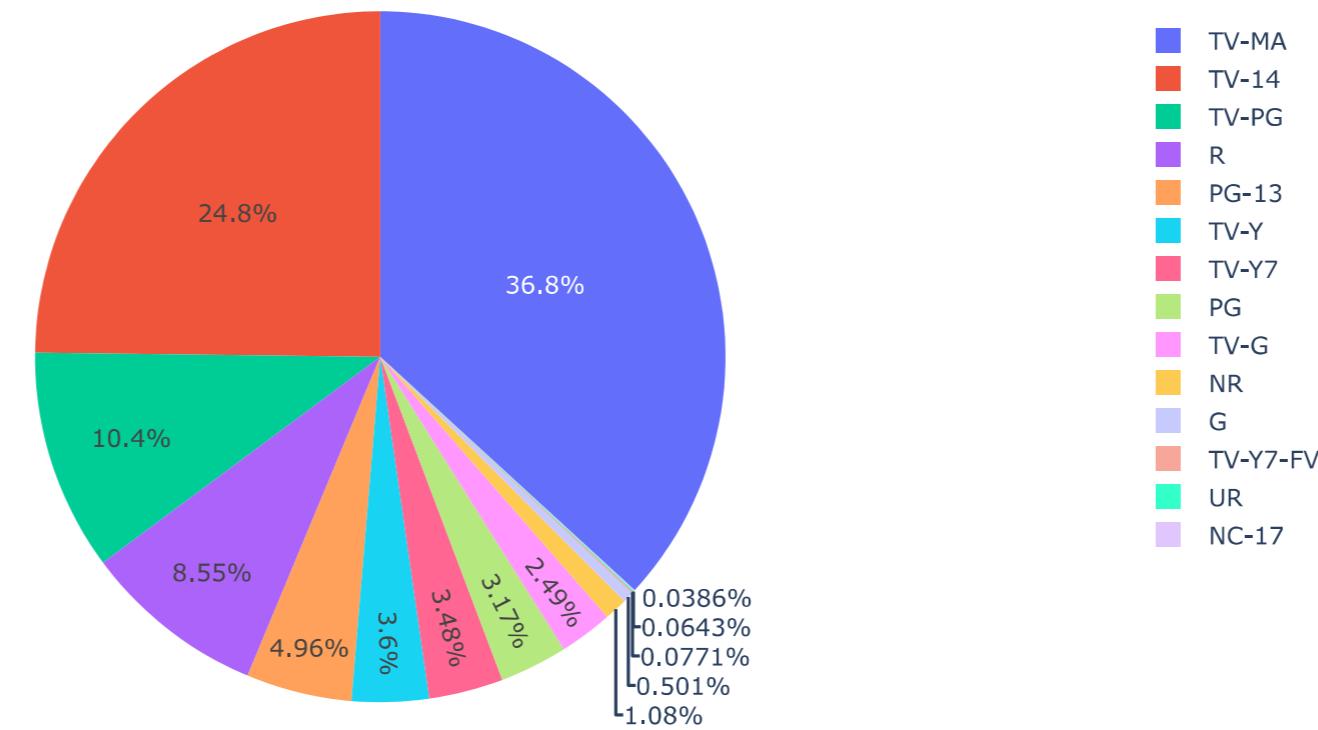
Out[196]:

```
TV-MA      2865  
TV-14      1931  
TV-PG      806  
R          665  
PG-13      386  
TV-Y       280  
TV-Y7      271  
PG          247  
TV-G       194  
NR          84  
G           39  
TV-Y7-FV    6  
UR          5  
NC-17      3  
Name: Rating, dtype: int64
```

In [197...]

```
px.pie(x, names=x.index, values=x.values, title='Distribution by Rating').show()
```

Distribution by Rating



Top 5 Directors

```
In [198]: df['Director'].value_counts()
```

```
Out[198]:
```

Director	Count
Raúl Campos, Jan Suter	18
Marcus Raboy	16
Jay Karas	14
Cathy Garcia-Molina	13
Jay Chapman	12
..	..
Lucien Jean-Baptiste	1
Jason Krawczyk	1
Quinn Lasher	1
Elle Callahan	1
Sam Dunn	1

Name: Director, Length: 4050, dtype: int64

```
In [199]: df['Director'].isna().sum()
```

```
Out[199]: 2388
```

```
In [200]: # filling Nan values
df['Director'].fillna('No Director Specified', inplace = True)
```

```
In [201]: df.head()
```

Out[201]:	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description
0	s1	TV Show	3%	No Director Specified	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connolly...	United States	November 16, 2017	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 min	Dramas	A brilliant group of students become card-coun...

In [202...]

```
# Top 5 Directors
director_df = df['Director'].to_frame()
director_df = director_df[director_df['Director'] != 'No Director Specified']
top5_directors = director_df['Director'].value_counts().head(5)
top5_directors
```

Out[202]:

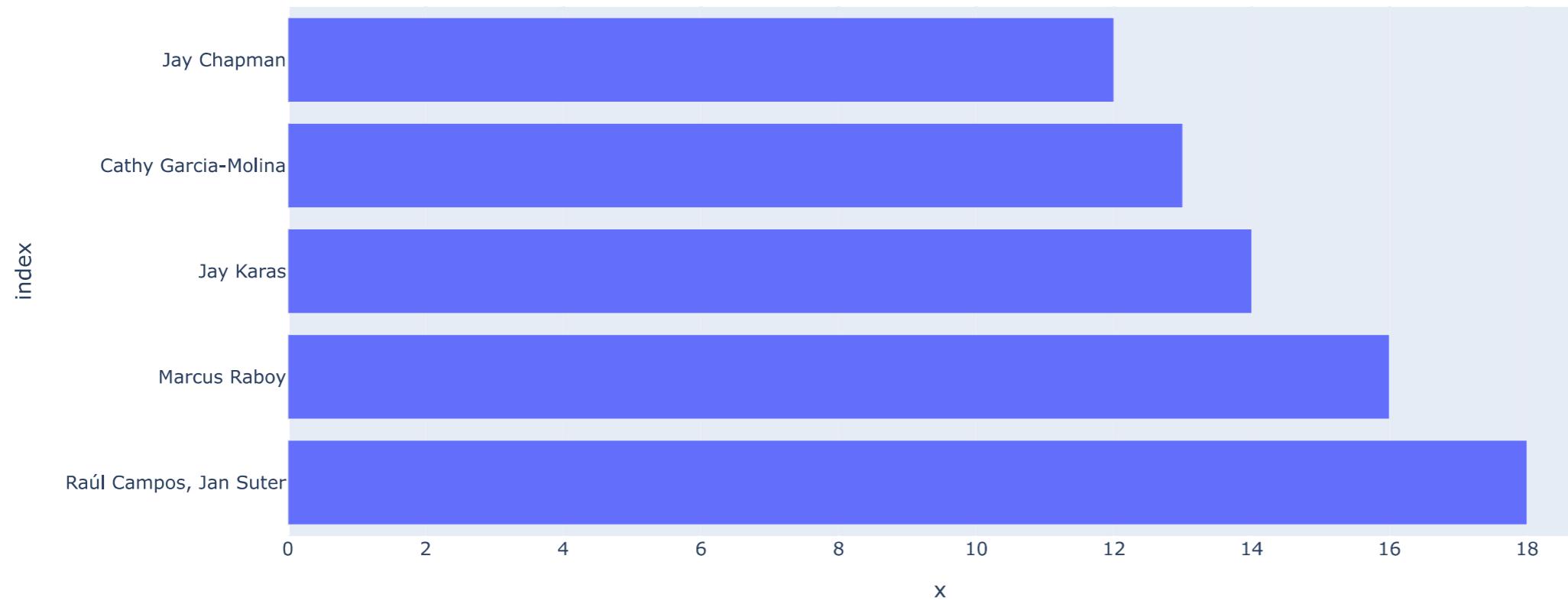
Raúl Campos, Jan Suter	18
Marcus Raboy	16
Jay Karas	14
Cathy Garcia-Molina	13
Jay Chapman	12

Name: Director, dtype: int64

In [203...]

```
px.bar(top5_directors, x = top5_directors.values, y = top5_directors.index, title = 'Top 5 Directors').show()
```

Top 5 Directors



TOP 5 Actors

In [204]:	<pre>df['Cast'].isna().sum()</pre>																																																																							
Out[204]:	718																																																																							
In [205]:	<pre>df['Cast'].fillna('No Actor Specified', inplace = True) df.head()</pre>																																																																							
Out[205]:	<table border="1"> <thead> <tr> <th>Show_Id</th> <th>Category</th> <th>Title</th> <th>Director</th> <th>Cast</th> <th>Country</th> <th>Release_Date</th> <th>Rating</th> <th>Duration</th> <th>Type</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>s1</td> <td>TV Show</td> <td>3%</td> <td>No Director Specified</td> <td>João Miguel, Bianca Comparato, Michel Gomes, R...</td> <td>Brazil</td> <td>August 14, 2020</td> <td>TV-MA</td> <td>4 Seasons</td> <td>International TV Shows, TV Dramas, TV Sci-Fi &...</td> <td>In a future where the elite inhabit an island ...</td> </tr> <tr> <td>1</td> <td>s2</td> <td>Movie</td> <td>07:19</td> <td>Jorge Michel Grau</td> <td>Demián Bichir, Héctor Bonilla, Oscar Serrano, ...</td> <td>Mexico</td> <td>December 23, 2016</td> <td>TV-MA</td> <td>93 min</td> <td>Dramas, International Movies</td> <td>After a devastating earthquake hits Mexico Cit...</td> </tr> <tr> <td>2</td> <td>s3</td> <td>Movie</td> <td>23:59</td> <td>Gilbert Chan</td> <td>Tedd Chan, Stella Chung, Henley Hii, Lawrence ...</td> <td>Singapore</td> <td>December 20, 2018</td> <td>R</td> <td>78 min</td> <td>Horror Movies, International Movies</td> <td>When an army recruit is found dead, his fellow...</td> </tr> <tr> <td>3</td> <td>s4</td> <td>Movie</td> <td>9</td> <td>Shane Acker</td> <td>Elijah Wood, John C. Reilly, Jennifer Connolly...</td> <td>United States</td> <td>November 16, 2017</td> <td>PG-13</td> <td>80 min</td> <td>Action & Adventure, Independent Movies, Sci-Fi...</td> <td>In a postapocalyptic world, rag-doll robots hi...</td> </tr> <tr> <td>4</td> <td>s5</td> <td>Movie</td> <td>21</td> <td>Robert Luketic</td> <td>Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...</td> <td>United States</td> <td>January 1, 2020</td> <td>PG-13</td> <td>123 min</td> <td>Dramas</td> <td>A brilliant group of students become card-coun...</td> </tr> </tbody> </table>	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	0	s1	TV Show	3%	No Director Specified	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...	1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...	2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...	3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connolly...	United States	November 16, 2017	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...	4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 min	Dramas	A brilliant group of students become card-coun...
Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description																																																														
0	s1	TV Show	3%	No Director Specified	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...																																																													
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...																																																													
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...																																																													
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connolly...	United States	November 16, 2017	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...																																																													
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 min	Dramas	A brilliant group of students become card-coun...																																																													

In [206...]

```
actors_df = pd.DataFrame()
actors_df = df['Cast'].str.split(',', expand = True).stack()
actors_df = actors_df.to_frame()
actors_df
```

Out[206]:

		0
0	0	João Miguel
1		Bianca Comparato
2		Michel Gomes
3		Rodolfo Valente
4		Vaneza Oliveira
...
7785	7	Chittaranjan Tripathy
7786	0	Nasty C
7787	0	Adriano Zumbo
	1	Rachel Khoo
7788	0	No Actor Specified

56689 rows × 1 columns

In [207...]

```
actors_df.columns = ['Actor']
actors_df
```

Out[207]:

		Actor
0	0	João Miguel
1		Bianca Comparato
2		Michel Gomes
3		Rodolfo Valente
4		Vaneza Oliveira
...
7785	7	Chittaranjan Tripathy
7786	0	Nasty C
7787	0	Adriano Zumbo
	1	Rachel Khoo
7788	0	No Actor Specified

56689 rows × 1 columns

In [208...]

```
actors= actors_df.groupby('Actor').size().reset_index(name = 'Count')
actors= actors[actors['Actor'] != 'No Actor Specified']
actors.sort_values(by = 'Count', ascending = False,inplace = True)
```

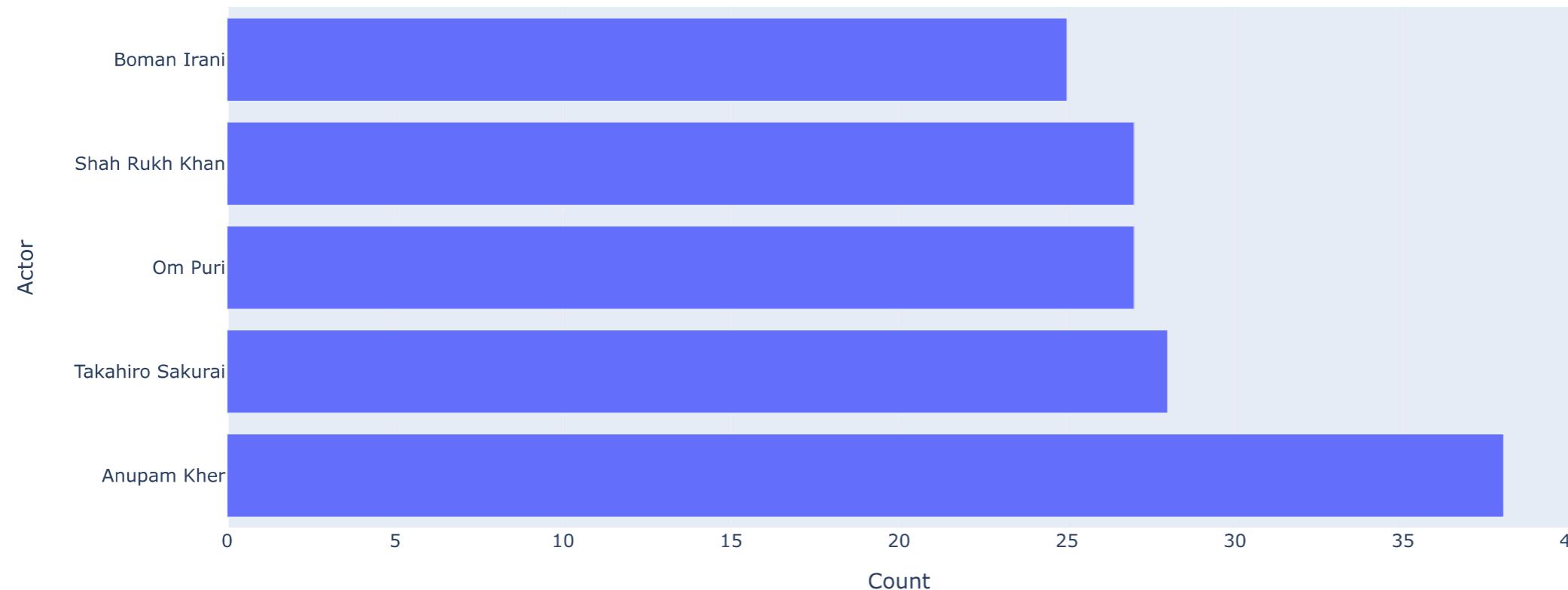
```
In [209]: top5actors = actors.head(5)
top5actors
```

```
Out[209]:
```

	Actor	Count
2321	Anupam Kher	38
27238	Takahiro Sakurai	28
21196	Om Puri	27
34593	Shah Rukh Khan	27
3758	Boman Irani	25

```
In [210]: px.bar(top5actors, x = 'Count', y = 'Actor',title = 'Top 5 Actors').show()
```

Top 5 Actors



Content Produced per year

```
In [211]: df['Release_Date'].isna().sum()
```

```
Out[211]: 10
```

```
In [212]: df.dropna(subset = ['Release_Date'], inplace = True)
```

In [213]:

```
df['Release_Date'].isna().sum()
```

Out[213]:

```
df['Release Year'] = df['Release Date'].apply(lambda x: x[-4:])
df
```

In [214]:

	Show_Id	Category	Title	Director	Cast	Country	Release_Date	Rating	Duration	Type	Description	Release Year
0	s1	TV Show	3%	No Director Specified	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	TV-MA	4 Seasons	International TV Shows, TV Dramas, TV Sci-Fi &...	In a future where the elite inhabit an island ...	2020
1	s2	Movie	07:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...	2016
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	December 20, 2018	R	78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...	2018
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connolly...	United States	November 16, 2017	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...	2017
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	January 1, 2020	PG-13	123 min	Dramas	A brilliant group of students become card-coun...	2020
...
7784	s7783	Movie	Zozo	Josef Fares	Imad Creidi, Antoinette Turk, Elias Gergi, Car...	Sweden, Czech Republic, United Kingdom, Denmar...	October 19, 2020	TV-MA	99 min	Dramas, International Movies	When Lebanon's Civil War deprives Zozo of his ...	2020
7785	s7784	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...	2019
7786	s7785	Movie	Zulu Man in Japan	No Director Specified	Nasty C	NaN	September 25, 2020	TV-MA	44 min	Documentaries, International Movies, Music & M...	In this documentary, South African rapper Nast...	2020
7787	s7786	TV Show	Zumbo's Just Desserts	No Director Specified	Adriano Zumbo, Rachel Khoo	Australia	October 31, 2020	TV-PG	1 Season	International TV Shows, Reality TV	Dessert wizard Adriano Zumbo looks for the nex...	2020
7788	s7787	Movie	ZZ TOP: THAT LITTLE OL' BAND FROM TEXAS	Sam Dunn	No Actor Specified	United Kingdom, Canada, United States	March 1, 2020	TV-MA	90 min	Documentaries, Music & Musicals	This documentary delves into the mystique behi...	2020

7779 rows × 12 columns

In [215]:

```
df['Release Year'].value_counts()
```

Out[215]:

2019	2154
2020	2010
2018	1685
2017	1225
2016	443
2021	117
2015	88
2014	25
2011	13
2013	11
2012	3
2008	2
2009	2
2010	1

Name: Release Year, dtype: int64

In [216]:

```
df1 = df[['Release Year', 'Category']]  
df1
```

Out[216]:

	Release Year	Category
0	2020	TV Show
1	2016	Movie
2	2018	Movie
3	2017	Movie
4	2020	Movie
...
7784	2020	Movie
7785	2019	Movie
7786	2020	Movie
7787	2020	TV Show
7788	2020	Movie

7779 rows × 2 columns

In [217]:

```
df2 = df1.groupby(['Release Year', 'Category']).size().reset_index(name = 'Count')  
df2
```

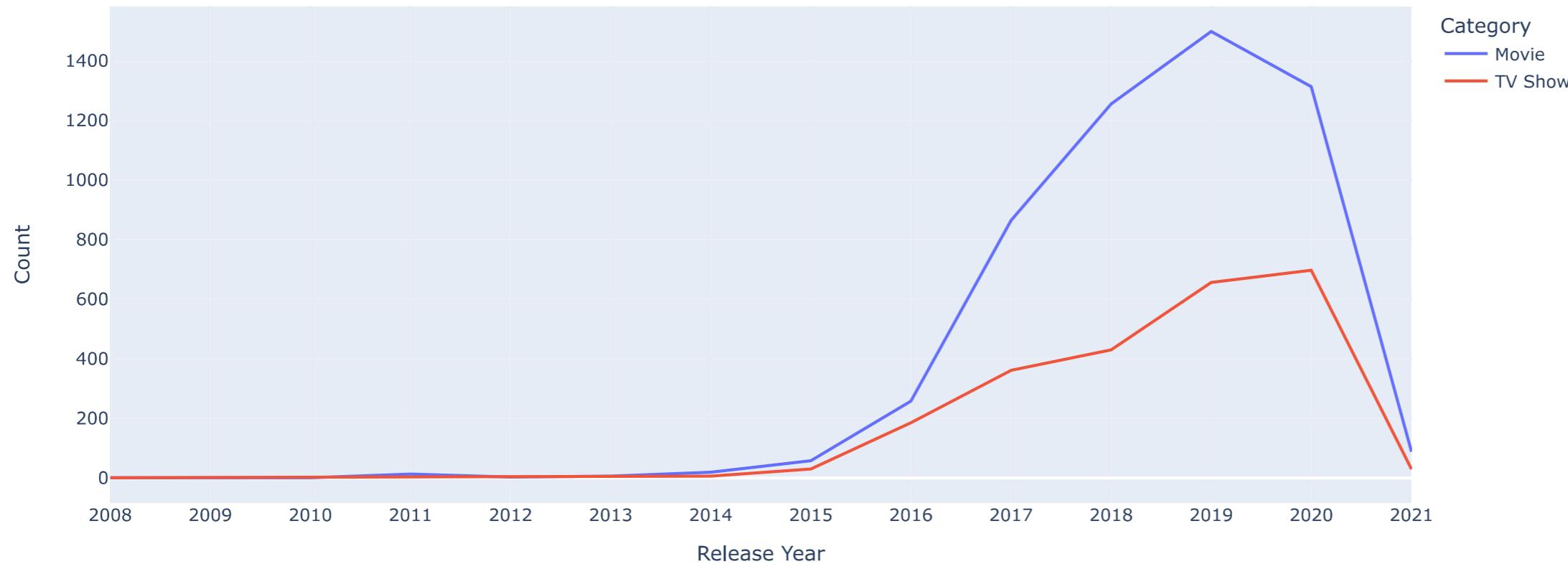
Out[217]:

	Release Year	Category	Count
0	2008	Movie	1
1	2008	TV Show	1
2	2009	Movie	2
3	2010	Movie	1
4	2011	Movie	13
5	2012	Movie	3
6	2013	Movie	6
7	2013	TV Show	5
8	2014	Movie	19
9	2014	TV Show	6
10	2015	Movie	58
11	2015	TV Show	30
12	2016	Movie	258
13	2016	TV Show	185
14	2017	Movie	864
15	2017	TV Show	361
16	2018	Movie	1255
17	2018	TV Show	430
18	2019	Movie	1498
19	2019	TV Show	656
20	2020	Movie	1313
21	2020	TV Show	697
22	2021	Movie	88
23	2021	TV Show	29

In [218...]

```
px.line(df2, x = 'Release Year', y = 'Count', color = 'Category', title = 'Type of Content Produced per year').show()
```

Type of Content Produced per year



TOP 5 Content Producing Countries

In [219...]

```
country_df = df.groupby('Country').size().reset_index(name = 'Count')
country_df.sort_values(by = 'Count', ascending = False, inplace = True)
country_df
```

Out[219]:

	Country	Count
549	United States	2550
229	India	923
462	United Kingdom	396
291	Japan	225
394	South Korea	183
...
252	Indonesia, South Korea, Singapore	1
253	Indonesia, United Kingdom	1
254	Indonesia, United States	1
255	Iran, France	1
680	Zimbabwe	1

681 rows × 2 columns

In [220...]

```
top5_country = country_df.head(5)  
top5_country
```

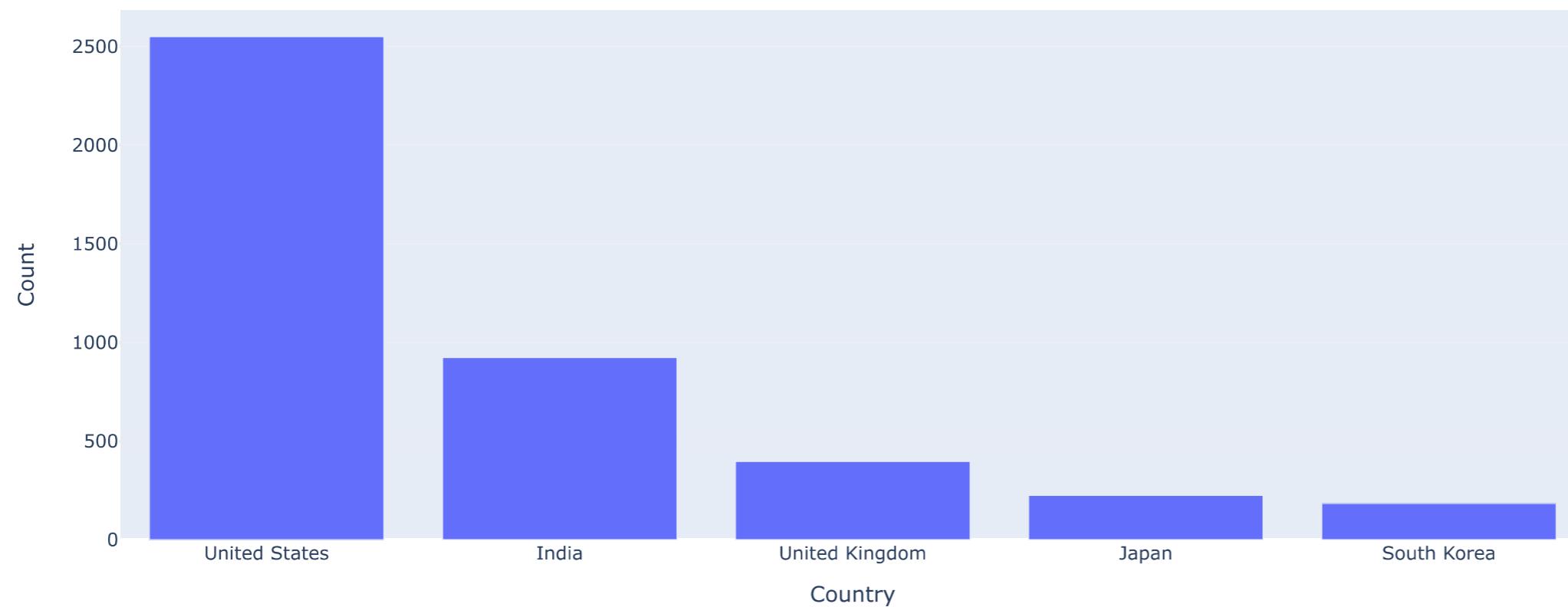
Out[220]:

	Country	Count
549	United States	2550
229	India	923
462	United Kingdom	396
291	Japan	225
394	South Korea	183

In [221...]

```
px.bar(top5_country, y = 'Count', x = 'Country', orientation='v', title = 'Top 5 content producing Countries').show()
```

Top 5 content producing Countries



In []: