

Netflix DATA Analysis

```
In [1]: import pandas as pd

In [4]: data=pd.read_csv('C:\Users\VAHITA\Downloads\Netflix_Data_Analysis_project\Netflix_titles.csv')

In [5]: data

Out[5]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kristen Johnson	N/A	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, firm...
1	s2	TV Show	Blood & Water	N/A	Anna Qamata, Khosi Ngema, Gal Mababane, Thab...	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 Gotsis, Samuel Josy, Nahi...	N/A	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lo...
3	s4	TV Show	Jabard Neen Okaans	N/A	N/A	India	September 24, 2021	2021	TV-MA	1 Season	Documentaries, Reality TV	Foals, firations and toilet talk go down amo...
4	s5	TV Show	Kala Fashy	N/A	Mayur More, Jitendra Kumar, Rajan Raj, Alim 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	s8802	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 30, 2019	2007	R	152 min	Cult Movies, Dramas, Thrillers	A political conspiracy, a crime reporter and a...
8803	s8803	TV Show	Zombe Dumb	N/A	N/A	N/A	July 1, 2019	2018	TV-17	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 Hamilton, 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, Courtney 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	Zubeen	Moose Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chan...	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 [ ]:
```

Getting Some Basic Information about Dataset

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

In [16]: # to show first 5 records from dataset

Out[16]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kristen Johnson	N/A	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, firm...
1	s2	TV Show	Blood & Water	N/A	Anna Qamata, Khosi Ngema, Gal Mababane, Thab...	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 Gotsis, Samuel Josy, Nahi...	N/A	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lo...
3	s4	TV Show	Jabard Neen Okaans	N/A	N/A	India	September 24, 2021	2021	TV-MA	1 Season	Documentaries, Reality TV	Foals, firations and toilet talk go down amo...
4	s5	TV Show	Kala Fashy	N/A	Mayur More, Jitendra Kumar, Rajan Raj, Alim 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 [15]: data.tail()

In [16]: # to show last 5 records from dataset

Out[16]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
8802	s8802	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 30, 2019	2007	R	150 min	Cult Movies, Dramas, Thrillers	A political conspiracy, a crime reporter and a...
8803	s8803	TV Show	Zombe Dumb	N/A	N/A	N/A	July 1, 2019	2018	TV-17	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 Hamilton, 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, Courtney 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	Zubeen	Moose Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chan...	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 [16]: data.shape

Out[16]: (8807, 12)

In [17]: data.size

In [17]: # this attribute will show the total elements present in the dataframe

Out[17]: 105684

In [18]: data.columns

In [18]: # this will show the columns names of the dataset

Out[18]: Index(['show_id', 'type', 'title', 'director', 'cast', 'country', 'date_added', 'release_year', 'rating', 'duration', 'listed_in', 'description'], dtype='object')

In [19]: data.info()

In [19]: # to show information about dataframe like indexes, columns, datatype of each column, memory size of dataframe

Out[19]:
```

```
Out[19]:
Out[19]:
In [19]: data.dtypes

In [19]: # to return data types of each column in the dataframe

Out[19]:
```

Distribution of Content:

To begin the task of analyzing Netflix data, I'll start by looking at the distribution of content ratings on Netflix:

```
In [15]: data.head(2)

Out[15]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kristen Johnson	N/A	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, firm...
1	s2	TV Show	Blood & Water	N/A	Anna Qamata, Khosi Ngema, Gal Mababane, Thab...	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...

```
In [16]: #data.groupby(['rating']).size().reset_index(name='count')

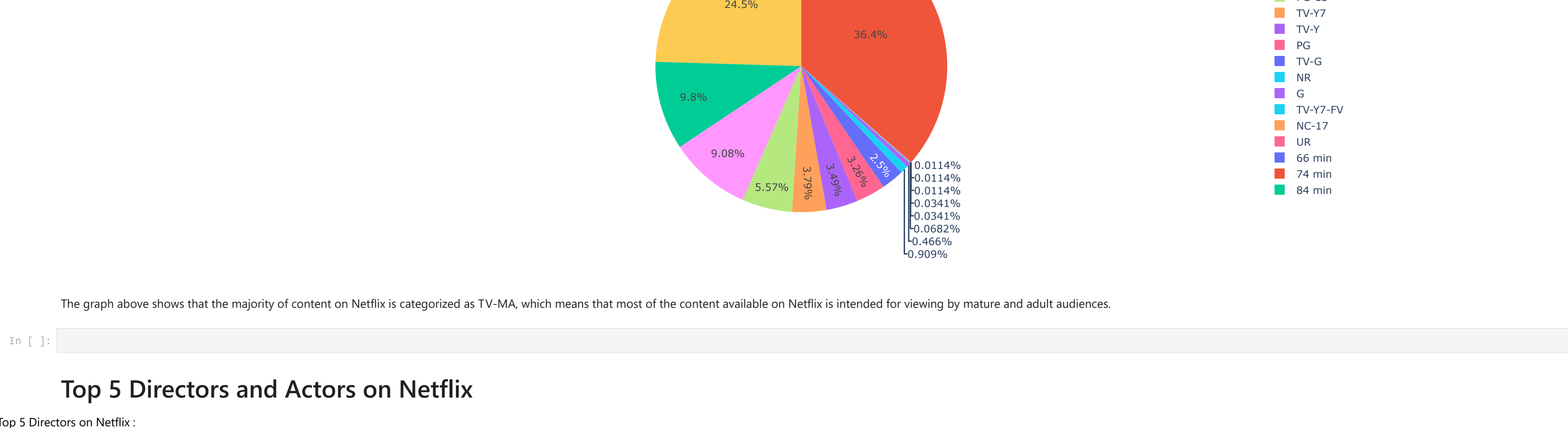
In [16]: z

Out[16]:
```

	rating	count
0	GS min	1
1	74 min	1
2	84 min	1
3	G	41
4	NC-17	3
5	NR	80
6	PG	287
7	PG-13	492
8	R	779
9	TV-14	2160
10	TV-G	230
11	TV-MA	3307
12	TV-PG	863
13	TV-Y	307
14	TV-Y7	334
15	TV-Y7-FV	6
16	UR	3

```
In [18]: import plotly.express as px

In [19]: px.plot(z.values('count'),names='rating',title='Distribution of content ratings on Netflix',color='rating')
```



The graph above shows that the majority of content on Netflix is categorized as TV-MA, which means that most of the content available on Netflix is intended for mature and adult audiences.

Top 5 Directors and Actors on Netflix

```
In [15]: data.head(2)

Out[15]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kristen Johnson	N/A	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, firm...
1	s2	TV Show	Blood & Water	N/A	Anna Qamata, Khosi Ngema, Gal Mababane, Thab...	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...

```
In [21]: data.director.fillna('no director specified',inplace=True)

In [22]: data.director.isnull().sum()

Out[22]: 0

In [23]: filtered_directors=pd.DataFrame()

In [24]: filtered_directors=data['director'].str.split(',').expand=True).stack()

In [25]: filtered_directors

Out[25]:
```

0	0	Kristen Johnson
1	0	no director specified
2	0	Julien Leclercq
3	0	no director specified
4	0	no director specified
...
8802	0	David Fincher
8803	0	no director specified
8804	0	Ruben Fleischer
8805	0	Peter Hewitt
8806	0	Moose Singh

Length: 6495, dtype: object

```
In [26]: filtered_directors=filtered_directors.to_frame()

In [27]: filtered_directors

Out[27]:
```

0	0	Kristen Johnson
1	0	no director specified
2	0	Julien Leclercq
3	0	no director specified
4	0	no director specified
...
8802	0	David Fincher
8803	0	no director specified
8804	0	Ruben Fleischer
8805	0	Peter Hewitt
8806	0	Moose Singh

9612 rows x 1 columns

```
In [28]: filtered_directors.columns=['director']

In [29]: filtered_directors

Out[29]:
```

	Director	Total Content
0	Aaron Moorhead	2
1	Aaron Woolf	1
2	Abbas Akkha Burnawalla	1
3	Abdullah Al Noor	1
4	Abhinav Shiv Tiwari	1
...
5115	Cagay Imrak	1
5117	Luigi Liguori	1
5118	Osar Thor Awolom	1
5119	Omer Faruk Sorak	2
5120	Serol Sonmez	2

5121 rows x 2 columns

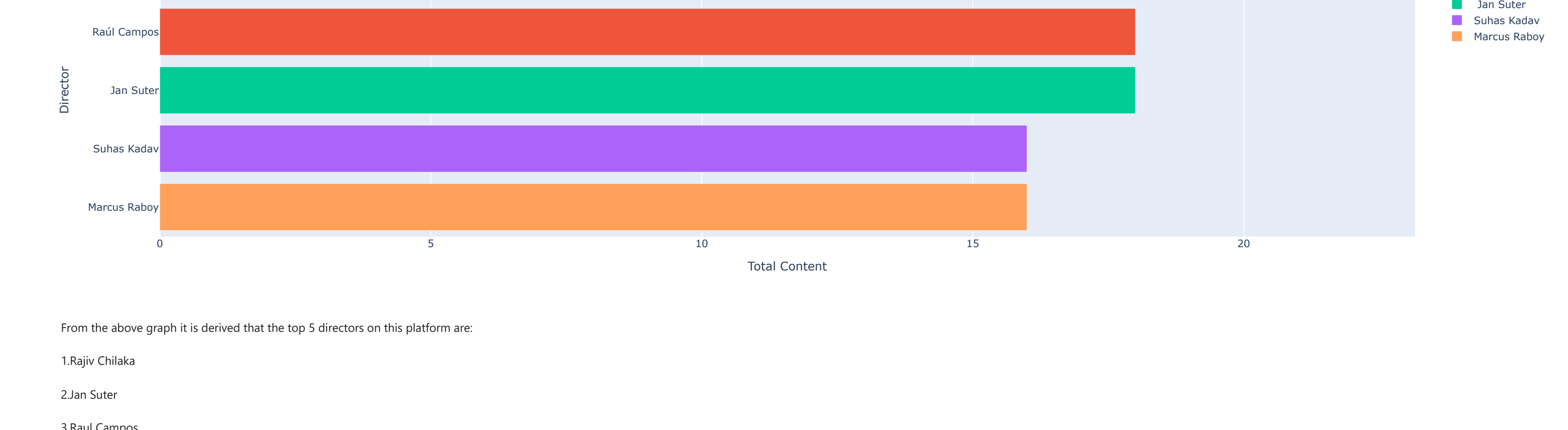
```
In [32]: directors_top5=directors[directors.director!='no director specified'].sort_values(by='Total Content',ascending=False).head()

In [33]: directors_top5

Out[33]:
```

	Director	Total Content
4020	Rajiv Chhila	22
4067	Raul Campos	18
261	Jan Suter	16
4051	Suhak Kadav	16
3235	Marcus Raboy	16

```
In [34]: px.bar(directors_top5,x=directors_top5['Total Content'],y=directors_top5['director'],title='Top 5 Directors on Netflix',color='director')
```



From the above graph it is derived that the top 5 directors on this platform are:

- 1.Rajiv Chhila
- 2.Jan Suter
- 3.Raul Campos
- 4.Suhak Kadav
- 5.Marcus Raboy

2.Top 5 Actors on Netflix:

```
In [15]: data.head(2)

Out[15]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kristen Johnson	N/A	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, firm...
1	s2	TV Show	Blood & Water	no director specified	Anna Qamata, Khosi Ngema, Gal Mababane, Thab...	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...

```
In [16]: data.cast.isnull().sum()

In [17]: data.cast.fillna('no cast specified',inplace=True)

In [18]: data.cast.isnull().sum()

Out[18]: 0

In [19]: filtered_cast=pd.DataFrame()

In [20]: filtered_cast=data.cast.str.split(',').expand=True).stack()

In [21]: filtered_cast

Out[21]:
```

0	0	no cast specified
1	0	Anna Qamata
2	0	Khosi Ngema
3	0	Gal Mababane
...
8806	3	Manish Chaudhary
8807	4	Meghna Malik
8808	5	Mahesh Rauti
8809	6	Armita Shasthi
8810	7	Chitraranjan Tripathy

Length: 6495, dtype: object

```
In [22]: filtered_cast=filtered_cast.to_frame()

In [23]: filtered_cast

Out[23]:
```

0	0	no cast specified
1	0	Anna Qamata
2	0	Khosi Ngema
3	0	Gal Mababane
...
8806	3	Manish Chaudhary
8807	4	Meghna Malik
8808	5	Mahesh Rauti
8809	6	Armita Shasthi
8810	7	Chitraranjan Tripathy

6495 rows x 1 columns

```
In [24]: filtered_cast.columns=['Actor']

In [25]: actors=filtered_cast.groupby(['Actor']).size().reset_index(name='Total Content')

In [26]: actors

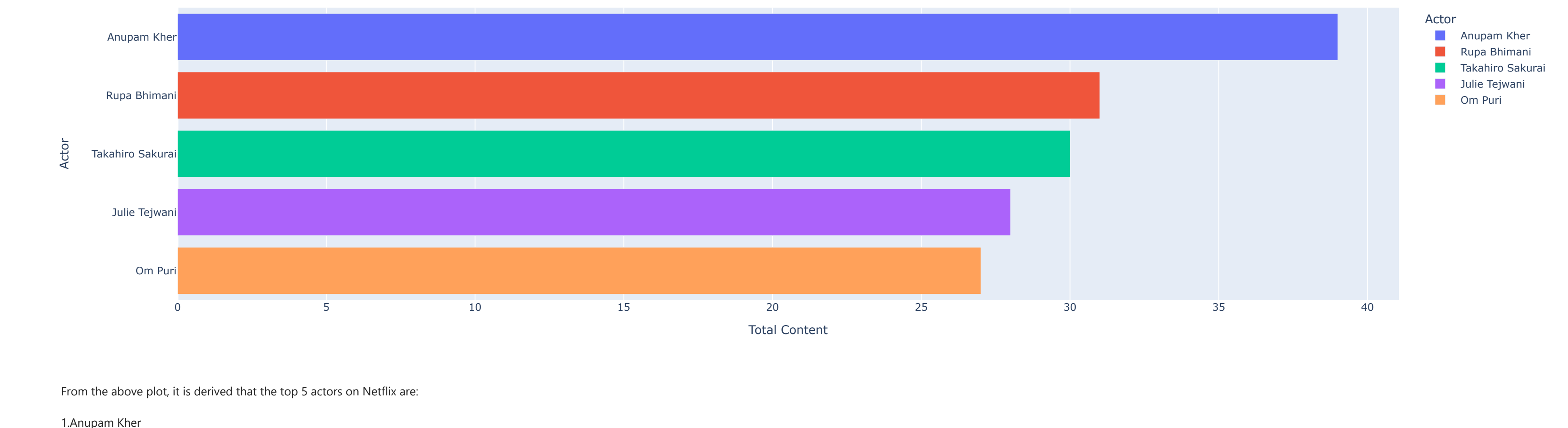
Out[26]:
```

	Actor	Total Content
0	Jr.	2
1	Wiley Jaffer Dadi	1
2	Naghi Dadi	1
3	2 Chang	1
4	2Mex	1
...
39292	Ibrahim Baydak	1
39293	Ibrahim Cekkai	1
39294	Sahin Imrak	1
39295	Sani Olayide	1
39296	Sope Olayide	1

39297 rows x 2 columns

```
In [29]: actors_top5=actors[actors.Actor!='no cast specified'].sort_values(by='Total Content',ascending=False).head()

In [30]: px.bar(actors_top5,x=actors_top5['Total Content'],y=actors_top5['Actor'],title='Top 5 Actors on Netflix',color='Actor')
```



From the above plot, it is derived that the top 5 actors on Netflix are:

- 1.Anupam Kher
- 2.Rupa Bhimani
- 3.Takahiro Sakurai
- 4.Julie Tejwani
- 5.Om Puri

The trend of production over the years on Netflix :

```
In [49]: data.head(2)

Out[49]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kristen Johnson	no cast specified	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, firm...
1	s2	TV Show	Blood & Water	no director specified	Anna Qamata, Khosi Ngema, Gal Mababane, Thab...	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...

```
In [50]: data.rename(columns={'release_year':'Release Year','type':'Type'},inplace=True)

In [51]: data[Release Year].unique()

In [51]: # shows the all years present in the data

Out[51]: array([1920, 2021, 1993, 2018, 1996, 1998, 1999, 2007, 2010, 2012, 2017, 1975, 1978, 1981, 1987, 2012, 2001, 2014, 2002, 2005, 2015, 2016, 2017, 2018, 2019, 2009, 2007, 2005, 2006, 1994, 2013, 2015, 2016, 2017, 1982, 1989, 1990, 1991, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379,
```