

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

```
In [2]: yt_data=pd.read_csv("Top YouTube Channels Data .csv")
yt_data
```

Out[2]:

	rank	youtuber	subscribers	video views	video count	category	started
0	1	T-Series	213000000	188,073,919,029	16708.0	Music	2006
1	2	YouTube Movies	150000000	167,122,746,349	NaN	Film & Animation	2015
2	3	Cocomelon - Nursery Rhymes	133000000	126,822,520,940	751.0	Education	2006
3	4	SET India	131000000	101,541,977,714	78334.0	Shows	2006
4	5	Music	116000000	78,437,871,689	NaN	Music	2013
...
95	96	Markiplier	32600000	18,011,837,263	5129.0	Gaming	2012
96	97	Like Nastya ESP	32600000	15,144,858,210	584.0	Entertainment	2017
97	98	Ryan's World	32400000	51,312,603,726	2155.0	Entertainment	2015
98	99	ABP News	32300000	9,850,740,503	209351.0	People & Blogs	2012
99	100	Desi Music Factory	32200000	9,115,577,588	122.0	Music	2014

100 rows × 7 columns

```
In [3]: yt_data.shape
```

Out[3]: (100, 7)

```
In [4]: yt_data.columns.tolist()
```

```
Out[4]: ['rank ',  
         'youtuber',  
         'subscribers ',  
         'video views ',  
         'video count ',  
         'category ',  
         'started ']
```

```
In [5]: yt_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 100 entries, 0 to 99  
Data columns (total 7 columns):  
#   Column          Non-Null Count  Dtype  
---  ---  
0   rank            100 non-null   int64  
1   youtuber        100 non-null   object  
2   subscribers     100 non-null   int64  
3   video views     100 non-null   object  
4   video count     95 non-null    float64  
5   category        100 non-null   object  
6   started         100 non-null   int64  
dtypes: float64(1), int64(3), object(3)  
memory usage: 5.6+ KB
```

```
In [6]: yt_data.describe()
```

Out[6]:

	rank	subscribers	video count	started
count	100.000000	1.000000e+02	95.000000	100.000000
mean	50.500000	5.336300e+07	15847.221053	2010.800000
std	29.011492	2.869713e+07	40955.200388	5.504819
min	1.000000	3.220000e+07	45.000000	1970.000000
25%	25.750000	3.620000e+07	393.500000	2007.750000
50%	50.500000	4.320000e+07	1139.000000	2012.000000
75%	75.250000	5.710000e+07	4986.000000	2014.000000
max	100.000000	2.130000e+08	209351.000000	2018.000000

```
In [7]: yt_data.isnull()
```

```
Out[7]:
```

	rank	youtuber	subscribers	video views	video count	category	started
0	False	False	False	False	False	False	False
1	False	False	False	False	True	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	True	False	False
...
95	False	False	False	False	False	False	False
96	False	False	False	False	False	False	False
97	False	False	False	False	False	False	False
98	False	False	False	False	False	False	False
99	False	False	False	False	False	False	False

100 rows × 7 columns

```
In [8]: yt_data.dropna(inplace=True) #drop the entries where Nan data is present
```

In [9]: yt_data

Out[9]:

	rank	youtuber	subscribers	video views	video count	category	started
0	1	T-Series	213000000	188,073,919,029	16708.0	Music	2006
2	3	Cocomelon - Nursery Rhymes	133000000	126,822,520,940	751.0	Education	2006
3	4	SET India	131000000	101,541,977,714	78334.0	Shows	2006
5	6	PewDiePie	111000000	28,260,779,633	4472.0	Gaming	2010
6	7	MrBeast	93900000	15,417,304,461	721.0	Entertainment	2012
...
95	96	Markiplier	32600000	18,011,837,263	5129.0	Gaming	2012
96	97	Like Nastya ESP	32600000	15,144,858,210	584.0	Entertainment	2017
97	98	Ryan's World	32400000	51,312,603,726	2155.0	Entertainment	2015
98	99	ABP News	32300000	9,850,740,503	209351.0	People & Blogs	2012
99	100	Desi Music Factory	32200000	9,115,577,588	122.0	Music	2014

95 rows × 7 columns

In [10]: yt_data.head()

Out[10]:

	rank	youtuber	subscribers	video views	video count	category	started
0	1	T-Series	213000000	188,073,919,029	16708.0	Music	2006
2	3	Cocomelon - Nursery Rhymes	133000000	126,822,520,940	751.0	Education	2006
3	4	SET India	131000000	101,541,977,714	78334.0	Shows	2006
5	6	PewDiePie	111000000	28,260,779,633	4472.0	Gaming	2010
6	7	MrBeast	93900000	15,417,304,461	721.0	Entertainment	2012

```
In [11]: yt_data.tail()
```

```
Out[11]:
```

	rank	youtuber	subscribers	video views	video count	category	started
95	96	Markiplier	32600000	18,011,837,263	5129.0	Gaming	2012
96	97	Like Nastya ESP	32600000	15,144,858,210	584.0	Entertainment	2017
97	98	Ryan's World	32400000	51,312,603,726	2155.0	Entertainment	2015
98	99	ABP News	32300000	9,850,740,503	209351.0	People & Blogs	2012
99	100	Desi Music Factory	32200000	9,115,577,588	122.0	Music	2014

```
In [13]: yt_data.describe().T
```

```
Out[13]:
```

	count	mean	std	min	25%	50%	75%	max
rank	95.0	5.202105e+01	2.824625e+01	1.0	28.5	52.0	75.5	100.0
subscribers	95.0	5.123263e+07	2.640947e+07	32200000.0	36200000.0	42700000.0	55450000.0	213000000.0
video count	95.0	1.584722e+04	4.095520e+04	45.0	393.5	1139.0	4986.0	209351.0
started	95.0	2.010663e+03	5.612696e+00	1970.0	2007.0	2011.0	2014.0	2018.0

```
In [22]: #top 10 of 95
yt_data.youtuber.head(10)
```

```
Out[22]: 0          T-Series
2    Cocomelon - Nursery Rhymes
3          SET India
5        PewDiePie
6         MrBeast
7    Kids Diana Show
9    Like Nastya
10         WWE
11    Zee Music Company
12    Vlad and Niki
Name: youtuber, dtype: object
```

```
In [26]: #bottom 10 of top 95
yt_data['youtuber'].tail(10)
```

```
Out[26]: 90                J Balvin
91          Masha and The Bear
92                VEGETTA777
93                YouTube
94  Goldmines Gaane Sune Ansune
95                Markiplier
96          Like Nastya ESP
97          Ryan's World
98                ABP News
99          Desi Music Factory
Name: youtuber, dtype: object
```

```
In [27]: #correlation between different metrics
yt_data.corr()
```

```
Out[27]:
```

	rank	subscribers	video count	started
rank	1.000000	-0.751911	0.062303	0.009507
subscribers	-0.751911	1.000000	0.001657	-0.065214
video count	0.062303	0.001657	1.000000	-0.088179
started	0.009507	-0.065214	-0.088179	1.000000

```
In [31]: yt_data['youtuber'].unique()
```

```
Out[31]: array(['T-Series ', 'Cocomelon - Nursery Rhymes ', 'SET India ',
                'PewDiePie ', 'MrBeast ', 'Kids Diana Show ', 'Like Nastya',
                'WWE ', 'Zee Music Company ', 'Vlad and Niki ', '5-Minute Crafts',
                'BLACKPINK ', 'Goldmines ', 'Justin Bieber ', 'Sony SAB ',
                'Canal KondZilla', 'BANGTANTV ', 'HYBE LABELS ', 'Zee TV ',
                'Shemaroo Filmi Gaane ', 'Dude Perfect ',
                "Pinkfong Baby Shark - Kids' Songs & Stories ", 'Movieclips ',
                'ChuChu TV Nursery Rhymes & Kids Songs ', 'Marshmello ',
                'Colors TV ', 'Wave Music ', 'EminemMusic ', 'Ed Sheeran ',
                'Aaj Tak ', 'Ariana Grande ', 'Sony Music India ',
                'Tips Official ', 'T-Series Bhakti Sagar ', 'El Reino Infantil ',
                "LooLoo Kids - Nursery Rhymes and Children's Songs ",
                'Taylor Swift ', 'YRF ', 'JuegaGerman', 'Badabun ',
                'Billie Eilish ', 'Fernanfloo', 'Felipe Neto ',
                'Whinderssonnunes ', 'HolaSoyGerman ', 'BRIGHT SIDE ',
                'Infobells - Hindi ', 'Katy Perry ', 'Voce Sabia?', 'Alan Walker ',
                'SonyMusicIndiaVEVO', 'Like Nastya Show ', 'elrubiusOMG',
                'Speed Records ', 'A4', 'Get Movies ', 'ABS-CBN Entertainment ',
                'Rihanna ', 'Little Baby Bum - Nursery Rhymes & Kids Songs ',
                'Bad Bunny ', 'Luisito Comunica ', 'TheEllenShow',
                'Маша и Медведь', 'GR6 EXPLODE ', 'LUCCAS NETO - LUCCAS TOON ',
                'One Direction ', 'XXXTENTACION ', 'Ishtar Music ', 'Voot Kids ',
                'Daddy Yankee ', 'Kimberly Loaiza', 'WorkpointOfficial ',
                'Shemaroo', 'Maroon 5', 'Shakira ', 'Toys and Colors ',
                'CarryMinati ', 'Shfa ', 'Ozuna ', 'shfa2 - شفا', 'Bruno Mars ',
                'TEDx Talks ', 'Jkk Entertainment ',
                'Super Simple Songs - Kids Songs ', 'one31', 'J Balvin ',
                'Masha and The Bear ', 'VEGETTA777', 'YouTube ',
                'Goldmines Gaane Sune Ansune ', 'Markiplier ', 'Like Nastya ESP ',
                "Ryan's World ", 'ABP News ', 'Desi Music Factory '], dtype=object)
```

```
In [32]: yt_data['youtuber'].nunique()
```

```
Out[32]: 95
```



```
In [38]: yt_data.loc[7]
```

```
Out[38]: rank                8
youtuber      Kids Diana Show
subscribers    93800000
video views    73,296,417,664
video count     977.0
category       People & Blogs
started        2015
Name: 7, dtype: object
```

```
In [41]: yt_data['category ']
```

```
Out[41]: 0          Music
2      Education
3          Shows
5          Gaming
6      Entertainment
      ...
95         Gaming
96      Entertainment
97      Entertainment
98      People & Blogs
99         Music
Name: category , Length: 95, dtype: object
```

```
In [42]: yt_data['category'].value_counts()
```

```
Out[42]: Music                33
Entertainment                19
People & Blogs              10
Gaming                      7
Education                   6
Shows                       4
Film & Animation            4
Comedy                     4
Sports                     2
How to & Style              2
News & Politics            2
Trailers                   1
Nonprofits & Activism      1
Name: category, dtype: int64
```

```
In [43]: #Top 10 youtubers based on video count
yt_data.sort_values(by=['video count'],ascending=False).head(10)['youtuber']
```

```
Out[43]: 98          ABP News
86          TEDx Talks
33          Aaj Tak
60  ABS-CBN Entertainment
89          one31
29          Colors TV
3           SET India
10          WWE
22          Zee TV
75  WorkpointOfficial
Name: youtuber, dtype: object
```

```
In [44]: #Top 10 youtubers based on video views
yt_data.sort_values(by=['video views'],ascending=False).head(10)['youtuber']
```

```
Out[44]: 98          ABP News
45          Fernanfloo
49          BRIGHT SIDE
70          XXXTENTACION
99    Desi Music Factory
18          Sony SAB
9          Like Nastya
7          Kids Diana Show
56          elrubiusOMG
64    Luisito Comunica
Name: youtuber, dtype: object
```

```
In [45]: #Top 10 youtubers based on subscribers
yt_data.sort_values(by=['subscribers'],ascending=False).head(10)['youtuber']
```

```
Out[45]: 0          T-Series
2    Cocomelon - Nursery Rhymes
3          SET India
5          PewDiePie
6          MrBeast
7          Kids Diana Show
9          Like Nastya
10         WWE
11    Zee Music Company
12    Vlad and Niki
Name: youtuber, dtype: object
```

```
In [49]: yt_data[['youtuber', 'subscribers']].head(7)
```

Out[49]:

	youtuber	subscribers
0	T-Series	213000000
2	Cocomelon - Nursery Rhymes	133000000
3	SET India	131000000
5	PewDiePie	111000000
6	MrBeast	93900000
7	Kids Diana Show	93800000
9	Like Nastya	91000000

```
In [50]: yt_data[yt_data['started'] == 2017]
```

Out[50]:

	rank	youtuber	subscribers	video views	video count	category	started
49	50	BRIGHT SIDE	43200000	9,762,131,005	4843.0	How to & Style	2017
55	56	Like Nastya Show	41000000	18,474,229,002	551.0	Entertainment	2017
84	85	shfa2 - شفا	34700000	19,676,656,583	1224.0	People & Blogs	2017
87	88	Jkk Entertainment	34500000	14,403,267,858	186.0	Comedy	2017
96	97	Like Nastya ESP	32600000	15,144,858,210	584.0	Entertainment	2017

```
In [52]: yt_data['started'].nunique()
```

Out[52]: 15

