

Finlatics Data Science Case Project

Domain:

“Media and Technology”

Analysing Youtube content creators dataset

About the Case Project

- A collection of YouTube giants, this dataset offers a perfect avenue to analyze and gain valuable insights from the luminaries of the platform.
- With comprehensive details on top creators' subscriber counts, video views, upload frequency, country of origin, earnings, and more.
- This treasure trove of information is a must-explore for aspiring content creators, data enthusiasts, and anyone intrigued by the ever-evolving online content landscape.
- Immerse yourself in the world of YouTube success and unlock a wealth of knowledge with this extraordinary dataset.

About the Dataset

Key Features:

- **rank:** Position of the YouTube channel based on the number of subscribers
- **Youtuber:** Name of the YouTube channel
- **subscribers:** Number of subscribers to the channel
- **video views:** Total views across all videos on the channel
- **category:** Category or niche of the channel
- **Title:** Title of the YouTube channel
- **uploads:** Total number of videos uploaded on the channel
- **Country:** Country where the YouTube channel originates
- **Abbreviation:** Abbreviation of the country
- **created_year:** Year when the YouTube channel was created
- **created_month:** Month when the YouTube channel was created
- **created_date:** Exact date of the YouTube channel's creation
- **Gross tertiary education enrollment (%):** Percentage of the population enrolled in tertiary education in the country
- **Population:** Total population of the country
- **Unemployment rate:** Unemployment rate in the country
- **Urban_population:** Percentage of the population living in urban areas
- **Latitude:** Latitude coordinate of the country's location
- **Longitude:** Longitude coordinate of the country's location
- **channel_type:** Type of the YouTube channel
- **video_views_rank:** Ranking of the channel based on total video views
- **country_rank:** Ranking of the channel based on the number of subscribers within its country
- **channel_type_rank:** Ranking of the channel based on its type
- **video_views_for_the_last_30_days:** Total video views in the last 30 days
- **lowest_monthly_earnings:** Lowest estimated monthly earnings from the channel
- **highest_monthly_earnings:** Highest estimated monthly earnings from the channel
- **lowest_yearly_earnings:** Lowest estimated yearly earnings from the channel
- **highest_yearly_earnings:** Highest estimated yearly earnings from the channel
- **subscribers_for_last_30_days:** Number of new subscribers gained in the last 30 days

Analysis to do :

NOTE: Preprocess the dataset prior to analysis for accurate insights

1. What are the top 10 YouTube channels based on the number of subscribers?
2. Which category has the highest average number of subscribers?
3. How many videos, on average, are uploaded by YouTube channels in each category?
4. What are the top 5 countries with the highest number of YouTube channels?
5. What is the distribution of channel types across different categories?
6. Is there a correlation between the number of subscribers and total video views for YouTube channels?
7. How do the monthly earnings vary throughout different categories?
8. What is the overall trend in subscribers gained in the last 30 days across all channels?
9. Are there any outliers in terms of yearly earnings from YouTube channels?
10. What is the distribution of channel creation dates? Is there any trend over time?
11. Is there a relationship between gross tertiary education enrollment and the number of YouTube channels in a country?
12. How does the unemployment rate vary among the top 10 countries with the highest number of YouTube channels?
13. What is the average urban population percentage in countries with YouTube channels?
14. Are there any patterns in the distribution of YouTube channels based on latitude and longitude coordinates?
15. What is the correlation between the number of subscribers and the population of a country?
16. How do the top 10 countries with the highest number of YouTube channels compare in terms of their total population?
17. Is there a correlation between the number of subscribers gained in the last 30 days and the unemployment rate in a country?
18. How does the distribution of video views for the last 30 days vary across different channel types?
19. Are there any seasonal trends in the number of videos uploaded by YouTube channels?
20. What is the average number of subscribers gained per month since the creation of YouTube channels till now?

Data Preprocessing

Dataset Before Preprocessing

	rank	Youtuber	subscribers	video views	category	Title	uploads	Country of origin	Country	Abbreviation	...	subscribers_for_last_30_days	created_year	created_month	created_date	Gross tertiary education enrollment (%)	Population	Unemployment rate	Urban
0	1	T-Series	245000000.0	2.280000e+11	Music	T-Series	20082	India	india	IN	...	2000000.0	2006.0	Mar	13.0	28.1	1.366418e+09	5.36	
1	2	YouTube Movies	170000000.0	0.000000e+00	Film & Animation	youtubemovies	1	United States	United States	US	...	NaN	2006.0	NaN	5.0	88.2	3.282395e+08	14.70	
2	3	MrBeast	166000000.0	2.836884e+10	Entertainment	MrBeast	741	United States	United States	US	...	8000000.0	2012.0	Feb	20.0	88.2	3.282395e+08	14.70	
3	4	Cocomelon - Nursery Rhymes	162000000.0	1.640000e+11	Education	Cocomelon - Nursery Rhymes	966	United States	United States	US	...	1000000.0	2006.0	Sep	1.0	88.2	3.282395e+08	14.70	
4	5	SET India	159000000.0	1.480000e+11	Shows	SET India	116536	India	India	IN	...	1000000.0	2006.0	Sep	20.0	28.1	1.366418e+09	5.36	
5	6	Music	119000000.0	0.000000e+00	NaN	Music	0	NaN	NaN	NaN	...	NaN	2013.0	Sep	24.0	NaN	NaN	NaN	
6	7	yyy Kids Diana Show	112000000.0	9.324704e+10	People & Blogs	yyy Kids Diana Show	1111	United States	United States	US	...	NaN	2015.0	May	12.0	88.2	3.282395e+08	14.70	
7	8	PewDiePie	111000000.0	2.905804e+10	Gaming	PewDiePie	4716	Japan	Japan	JP	...	NaN	2010.0	Apr	29.0	63.2	1.262266e+08	2.29	
8	9	Like Nastya	106000000.0	9.047906e+10	People & Blogs	Like Nastya Vlog	493	Russia	Russia	RU	...	100000.0	2016.0	Jan	14.0	81.9	1.443735e+08	4.59	
9	10	Vlad and Niki	98900000.0	7.718017e+10	Entertainment	Vlad and Niki	574	United States	United States	US	...	600000.0	2018.0	Apr	23.0	88.2	3.282395e+08	14.70	
10	11	Zee Music Company	96700000.0	5.785629e+10	Music	Zee Music Company	8548	India	India	IN	...	1100000.0	2014.0	Mar	12.0	28.1	1.366418e+09	5.36	
11	12	WWE	96000000.0	7.742847e+10	Sports	WWE	70127	United States	United States	US	...	600000.0	2007.0	May	11.0	88.2	3.282395e+08	14.70	
12	13	Gaming	93600000.0	0.000000e+00	NaN	Gaming	0	NaN	NaN	NaN	...	NaN	2013.0	Dec	15.0	NaN	NaN	NaN	
13	14	BLACKPINK	89800000.0	3.214460e+10	NaN	BLACKPINK	543	South Korea	South Korea	KR	...	700000.0	2016.0	Jun	29.0	94.3	5.170910e+07	4.15	
14	15	Goldmines	86900000.0	2.411823e+10	NaN	goldmines	1	NaN	NaN	NaN	...	NaN	2006.0	Aug	15.0	NaN	NaN	NaN	
15	16	Sony SAB	83000000.0	1.010000e+11	Shows	Sony SAB	71270	India	India	IN	...	1100000.0	2007.0	Aug	4.0	28.1	1.366418e+09	5.36	
16	17	5-Minute Crafts	80100000.0	2.623679e+10	Howto & Style	5-Minute Crafts 2.0	1	United Kingdom	United Kingdom	GB	...	NaN	2020.0	Jul	27.0	60.0	6.683440e+07	3.85	
17	18	BANGTANTV	75600000.0	2.082699e+10	Music	BANGTANTV	2281	South Korea	South Korea	KR	...	400000.0	2012.0	Dec	17.0	94.3	5.170910e+07	4.15	
18	19	Sports	75000000.0	0.000000e+00	NaN	sports	3	United States	United States	US	...	NaN	2006.0	Jan	30.0	88.2	3.282395e+08	14.70	
19	20	Justin Bieber	71600000.0	3.060812e+10	Music	Justin Bieber	249	Canada	Canada	CA	...	100000.0	2007.0	Jan	15.0	68.9	3.699198e+07	5.56	

```

...   Number of missing values in the dataset:

...   rank                                0
      Youtuber                            0
      subscribers                         3
      video_views                         0
      category                           55
      Title                              0
      uploads                            0
      Country of origin                  125
      Country                           125
      Abbreviation                       125
      channel_type                       32
      video_views_rank                   1
      country_rank                       119
      channel_type_rank                  35
      video_views_for_the_last_30_days   57
      lowest_monthly_earnings            0
      highest_monthly_earnings           0
      lowest_yearly_earnings             0
      highest_yearly_earnings            0
      subscribers_for_last_30_days       340
      created_year                       5
      created_month                      12
      created_date                       5
      Gross tertiary education enrollment (%) 126
      Population                         126
      Unemployment rate                  126
      Urban_population                   126
      Latitude                           126
      Longitude                          126
      dtype: int64

```

missing values in the Dataset

Dataset After Preprocessing

	rank	Youtuber	subscribers	video views	category	Title	uploads	Country of origin	Country	Abbreviation	...	subscribers_for_last_30_days	created_year	created_month	created_date	Gross tertiary education enrollment (%)	Population	Unemployment rate	Urban
0	1	T-Series	245000000.0	2.280000e+11	Music	T-Series	20082	India	india	IN	...	2.000000e+06	2006.0	Mar	13.0	28.100000	1.366418e+09	5.360000	4.
1	2	YouTube Movies	170000000.0	0.000000e+00	Film & Animation	youtubemovies	1	United States	United States	US	...	3.495419e+05	2006.0	Jan	5.0	88.200000	3.282395e+08	14.700000	2.
2	3	MrBeast	166000000.0	2.836884e+10	Entertainment	MrBeast	741	United States	United States	US	...	8.000000e+06	2012.0	Feb	20.0	88.200000	3.282395e+08	14.700000	2.
3	4	Cocomelon - Nursery Rhymes	162000000.0	1.640000e+11	Education	Cocomelon - Nursery Rhymes	966	United States	United States	US	...	1.000000e+06	2006.0	Sep	1.0	88.200000	3.282395e+08	14.700000	2.
4	5	SET India	159000000.0	1.480000e+11	Shows	SET India	116536	India	India	IN	...	1.000000e+06	2006.0	Sep	20.0	28.100000	1.366418e+09	5.360000	4.
5	6	Music	119000000.0	0.000000e+00	Entertainment	Music	0	United States	United States	US	...	3.495419e+05	2013.0	Sep	24.0	63.597273	4.304586e+08	9.258966	2.
6	7	ýýý Kids Diana Show	112000000.0	9.324704e+10	People & Blogs	ýýý Kids Diana Show	1111	United States	United States	US	...	3.495419e+05	2015.0	May	12.0	88.200000	3.282395e+08	14.700000	2.
7	8	PewDiePie	111000000.0	2.905804e+10	Gaming	PewDiePie	4716	Japan	Japan	JP	...	3.495419e+05	2010.0	Apr	29.0	63.200000	1.262266e+08	2.290000	1.
8	9	Like Nastya	106000000.0	9.047906e+10	People & Blogs	Like Nastya Vlog	493	Russia	Russia	RU	...	1.000000e+05	2016.0	Jan	14.0	81.900000	1.443735e+08	4.590000	1.
9	10	Vlad and Niki	98900000.0	7.718017e+10	Entertainment	Vlad and Niki	574	United States	United States	US	...	6.000000e+05	2018.0	Apr	23.0	88.200000	3.282395e+08	14.700000	2.
10	11	Zee Music Company	96700000.0	5.785629e+10	Music	Zee Music Company	8548	India	India	IN	...	1.100000e+06	2014.0	Mar	12.0	28.100000	1.366418e+09	5.360000	4.
11	12	WWE	96000000.0	7.742847e+10	Sports	WWE	70127	United States	United States	US	...	6.000000e+05	2007.0	May	11.0	88.200000	3.282395e+08	14.700000	2.
12	13	Gaming	93600000.0	0.000000e+00	Entertainment	Gaming	0	United States	United States	US	...	3.495419e+05	2013.0	Dec	15.0	63.597273	4.304586e+08	9.258966	2.
13	14	BLACKPINK	89800000.0	3.214460e+10	Entertainment	BLACKPINK	543	South Korea	South Korea	KR	...	7.000000e+05	2016.0	Jun	29.0	94.300000	5.170910e+07	4.150000	4.
14	15	Goldmines	86900000.0	2.411823e+10	Entertainment	goldmines	1	United States	United States	US	...	3.495419e+05	2006.0	Aug	15.0	63.597273	4.304586e+08	9.258966	2.
15	16	Sony SAB	83000000.0	7.010000e+11	Shows	Sony SAB	71270	India	India	IN	...	1.100000e+06	2007.0	Aug	4.0	28.100000	1.366418e+09	5.360000	4.
16	17	5-Minute Crafts	80100000.0	2.623679e+10	Howto & Style	5-Minute Crafts 2.0	1	United Kingdom	United Kingdom	GB	...	3.495419e+05	2020.0	Jul	27.0	60.000000	6.683440e+07	3.850000	5.
17	18	BANGTANTV	75600000.0	2.082699e+10	Music	BANGTANTV	2281	South Korea	South Korea	KR	...	4.000000e+05	2012.0	Dec	17.0	94.300000	5.170910e+07	4.150000	4.
18	19	Sports	75000000.0	0.000000e+00	Entertainment	sports	3	United States	United States	US	...	3.495419e+05	2006.0	Jan	30.0	88.200000	3.282395e+08	14.700000	2.
19	20	Justin Bieber	71600000.0	3.060812e+10	Music	Justin Bieber	249	Canada	Canada	CA	...	1.000000e+05	2007.0	Jan	15.0	68.900000	3.699198e+07	5.560000	3.

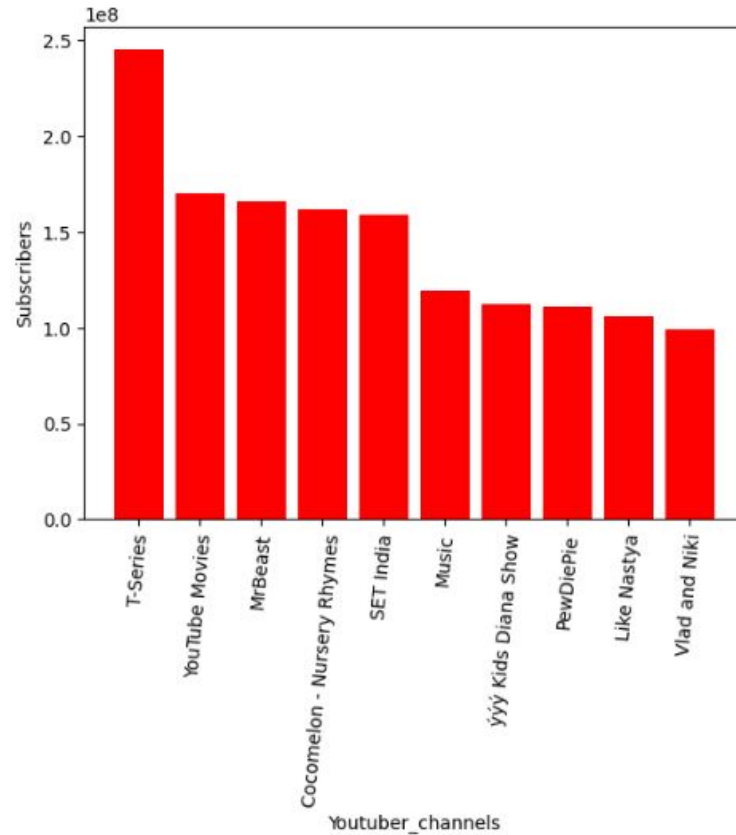
20 rows × 29 columns

Insights

Top 10 Youtube Channels based on no. of Subscribers

T-series has the most no. of subscribers .

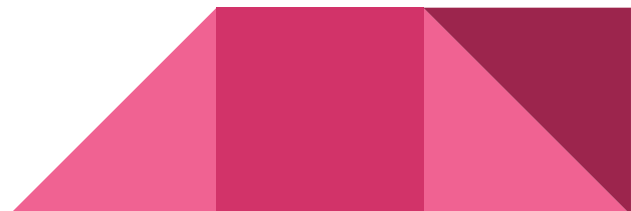
	Youtuber	subscribers	Country
0	T-Series	245000000.0	india
1	YouTube Movies	170000000.0	United States
2	MrBeast	166000000.0	United States
3	Cocomelon - Nursery Rhymes	162000000.0	United States
4	SET India	159000000.0	India
5	Music	119000000.0	United States
6	ýýý Kids Diana Show	112000000.0	United States
7	PewDiePie	111000000.0	Japan
8	Like Nastya	106000000.0	Russia
9	Vlad and Niki	98900000.0	United States



```
Category wise average subscribers: category
Autos & Vehicles      1.660000e+07
Comedy                2.002571e+07
Education             2.712391e+07
Entertainment         2.248462e+07
Film & Animation       2.752878e+07
Gaming                2.088626e+07
Howto & Style         1.935750e+07
Movies                2.565000e+07
Music                 2.604926e+07
News & Politics       2.063077e+07
Nonprofits & Activism 2.775000e+07
People & Blogs        2.053130e+07
Pets & Animals        1.732000e+07
Science & Technology  1.861765e+07
Shows                 4.161538e+07
Sports                3.285000e+07
Trailers              3.900000e+07
Travel & Events       1.250000e+07
Name: subscribers, dtype: float64
```

```
category : Shows , has max avg subscribers with an average of :41615384.615384616 subscribers
```

Category “shows” has the highest no. of subscribers across all the categories .

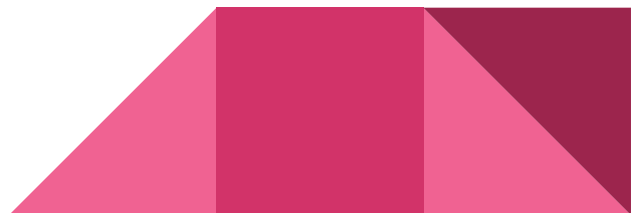


Average number of videos uploaded by YouTube channels in each category:

category	
Autos & Vehicles	1550.666667
Comedy	1202.557143
Education	3087.086957
Entertainment	9938.399317
Film & Animation	2861.844444
Gaming	4285.273684
Howto & Style	1695.500000
Movies	3553.000000
Music	2325.945813
News & Politics	112484.384615
Nonprofits & Activism	102912.000000
People & Blogs	9256.793893
Pets & Animals	3562.800000
Science & Technology	2114.058824
Shows	27443.692308
Sports	19129.833333
Trailers	6839.000000
Travel & Events	766.000000

Name: uploads, dtype: float64

avg highest no. of videos
uploaded belong to “News
and Politics” category



Top 5 countries with the highest number of YouTube channels:

Country	count
United States	440
India	169
Brazil	62
United Kingdom	44
Mexico	33

Name: count, dtype: int64

US has the max no. of Youtube channels around the world about 440

Distribution of channel types across different categories:

category	count
Entertainment	293
Music	203
People & Blogs	131
Gaming	95
Comedy	70
Education	46
Film & Animation	45
Howto & Style	40
News & Politics	26
Science & Technology	17
Shows	13
Sports	12
Pets & Animals	5
Autos & Vehicles	3
Trailers	2
Nonprofits & Activism	2
Movies	2
Travel & Events	1

Name: count, dtype: int64

most of the Youtube channels make content for the "Entertainment" of people



Outliers in terms of yearly earnings:

	rank	Youtuber	subscribers	video views	category	Title	uploads	Country of origin	Country
0	1	T-Series	245000000.0	2.280000e+11	Music	T-Series	20082	India	India
3	4	Cocomelon - Nursery Rhymes	162000000.0	1.640000e+11	Education	Cocomelon - Nursery Rhymes	966	United States	United States
4	5	SET India	159000000.0	1.480000e+11	Shows	SET India	116536	India	India
15	16	Sony SAB	83000000.0	1.010000e+11	Shows	Sony SAB	71270	India	India
21	22	Zee TV	70500000.0	7.313905e+10	Entertainment	Zee TV	129204	India	India
83	84	GR6 EXPLODE	38900000.0	2.515423e+10	Music	GR6 EXPLODE	3043	Brazil	Brazil
139	140	StarPlus	32000000.0	2.680067e+10	Entertainment	StarPlus	44892	India	India
302	303	KL BRO Biju Rithvik	22800000.0	1.798835e+10	Entertainment	KL BRO Biju Rithvik	1841	United States	United States
332	333	ZAMZAM ELECTRONICS TRADING	21700000.0	8.507819e+09	Entertainment	ZAMZAM ELECTRONICS TRADING	15126	United Arab Emirates	United Arab Emirates
417	418	DaFuq!? Boom!	19600000.0	7.906182e+09	Film & Animation	DaFuq!? Boom!	214	United States	United States
495	496	ய்யய்யய்ய KIMPRO	17700000.0	1.920670e+10	Entertainment	ய்யய்யய்ய KIMPRO	1646	South Korea	South Korea

11 rows × 29 columns

Around 11 Youtube channels behave as has very big deviation in yearly income

created_year

count

Missing: 0 (0%)
Distinct: 15 (75%)



1970	1
2005	24
2006	93
2007	52
2008	46
2009	52
2010	48
2011	82
2012	68
2012.6133866134	5
2013	78
2014	100
2015	73
2016	78
2017	68
2018	46
2019	34
2020	30
2021	23
2022	5

The Highest no. of Youtube channel were created in the Year “2014” - 100

while 2nd highest was year 2013 - 78

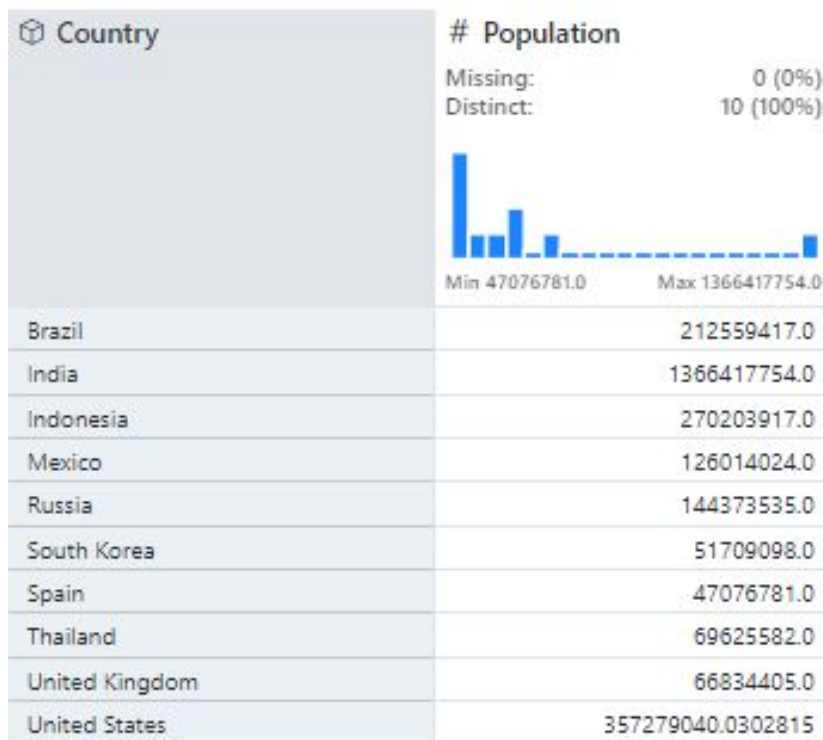
- Correlation b/w the No. of subscribers and the total number of video views is approximately 74 % , i.e around 74 % of the subscribers watch the uploaded video content after subscribing.
- The overall subscribers in the last 30 days across all channels is 351,639,103.
- Around 77.5 % of the of the the population across the tertiary education enrollment use Youtube channels for education purposes.
- T-series Channel has the highest no. of avg subscribers added per month since its creation date



Spain has the most unemployment rate of its total population approx 14 % followed by US around 13.15 % Brazil around 12 % India around 5.3 %

Population of top 10 Countries with highest no. of Youtube Channels:

India has the most population



Thank You

