



Finlatics Data Science

**Indian Institute of
Technology Jodhpur**

Project Report by Chaitanya Patil

Table of Contents

- 01 Introduction to Project
- 02 Data Preprocessing
- 03 Analysis of the Topic
- 04 Keyfinding and Conclusion

Introduction to Project



Overview

- This project report focuses on Media and Technology within data science.
- This Project Analyzing global YouTube statistics using data science.
- Objective of this project is to understand trends, patterns, and insights in the YouTube ecosystem .
- The Dataset used in this project was Global Youtube Statistics.csv

Using Google Colab



Cloud-Based

- Google Colab facilitated programming by eliminating local software installation.
- This cloud platform allowed for easy collaboration on the project.
- Google Colab supports various essential data analysis libraries.
- Streamlined development across various devices was achieved via Google Colab.

Data Preprocessing



Data Preprocessing

- Pandas, Matplotlib, NumPy, and Seaborn were essential libraries used.
- Removed duplicate data for data integrity
- Pre-written code modules were incorporated for efficiency.
- Filled missing values with '0' to handle missing data.

Dataset Headers and Missing Values

Data Headers and missing values

rank	0
Youtuber	0
subscribers	3
video_views	0
category	53
Title	0
uploads	0
Country of origin	122
Country	122
Abbreviation	122
channel_type	30
video_views_rank	1
country_rank	116
channel_type_rank	33
video_views_for_the_last_30_days	56
lowest_monthly_earnings	0
highest_monthly_earnings	0
lowest_yearly_earnings	0
highest_yearly_earnings	0
subscribers_for_last_30_days	337
created_year	5
created_month	12
created_date	5
Gross tertiary education enrollment (%)	123
Population	123
Unemployment rate	123
Urban_population	123
Latitude	123
Longitude	123
dtype: int64	

Total number of missing values in the data
1755

Total number of missing values in the data ,After fillin the '0'
0

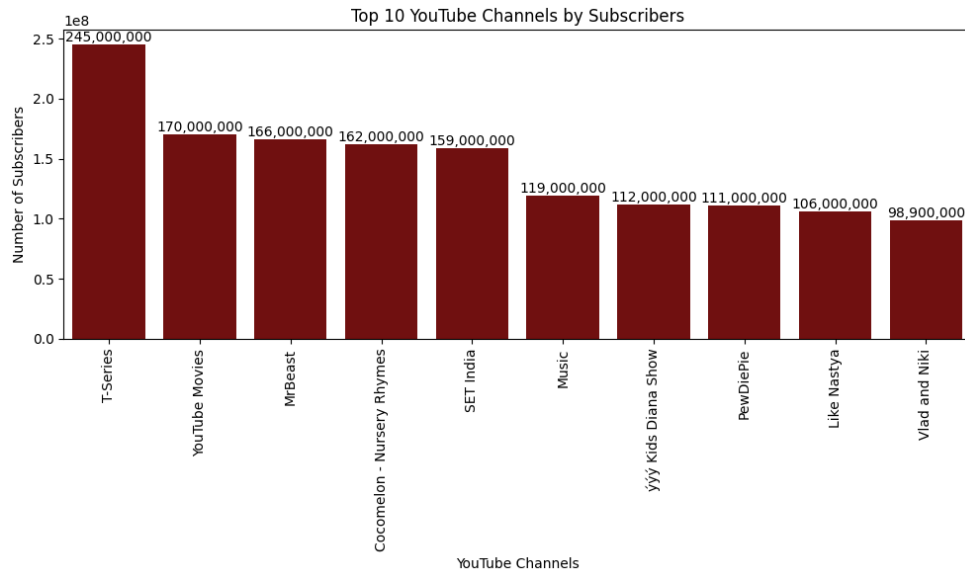
Top 10 YouTube Channels by Subscribers



- Ranked channels based on subscriber count.
- Large subscriber base indicates reach and influence.

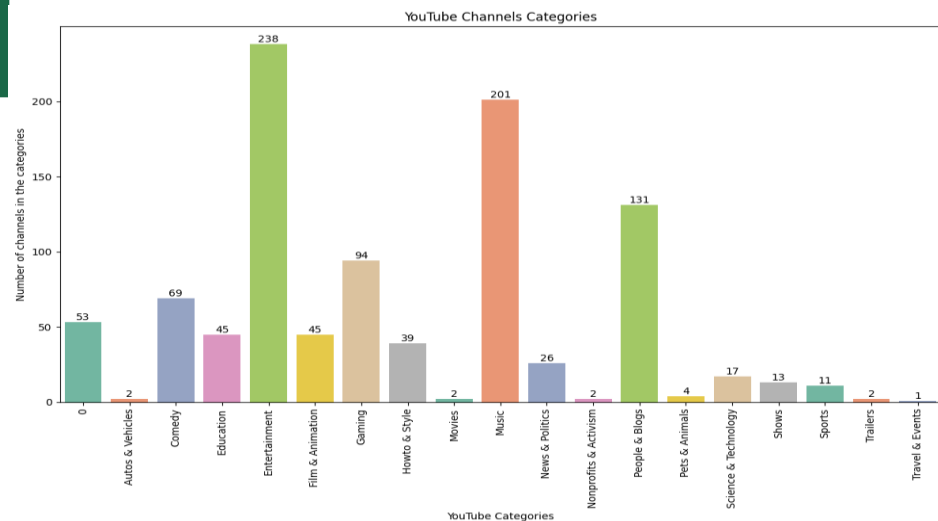
Top 10 Youtube Channels by Comparing the Subscribers:

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



Category with Highest Average Subscribers

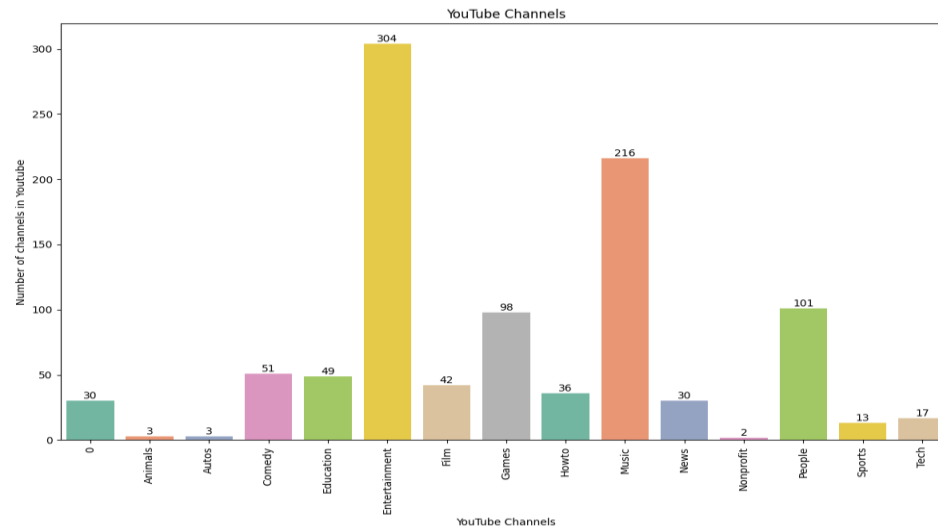
5



- "Shows" category has the highest average subscriber.
- Discuss factors contributing to the category's popularity.

Category with highest average subscribers: Shows

Different types of categories		Different types of channel types	
category	count	channel_type	count
0	0	0	30
1	Autos & Vehicles	1	Animals
2	Comedy	2	Autos
3	Education	3	Comedy
4	Entertainment	4	Education
5	Film & Animation	5	Entertainment
6	Gaming	6	Film
7	Howto & Style	7	Games
8	Movies	8	Howto
9	Music	9	Music
10	News & Politics	10	News
11	Nonprofits & Activism	11	Nonprofit
12	People & Blogs	12	People
13	Pets & Animals	13	Sports
14	Science & Technology	14	Tech
15	Shows		
16	Sports		
17	Trailers		
18	Travel & Events		

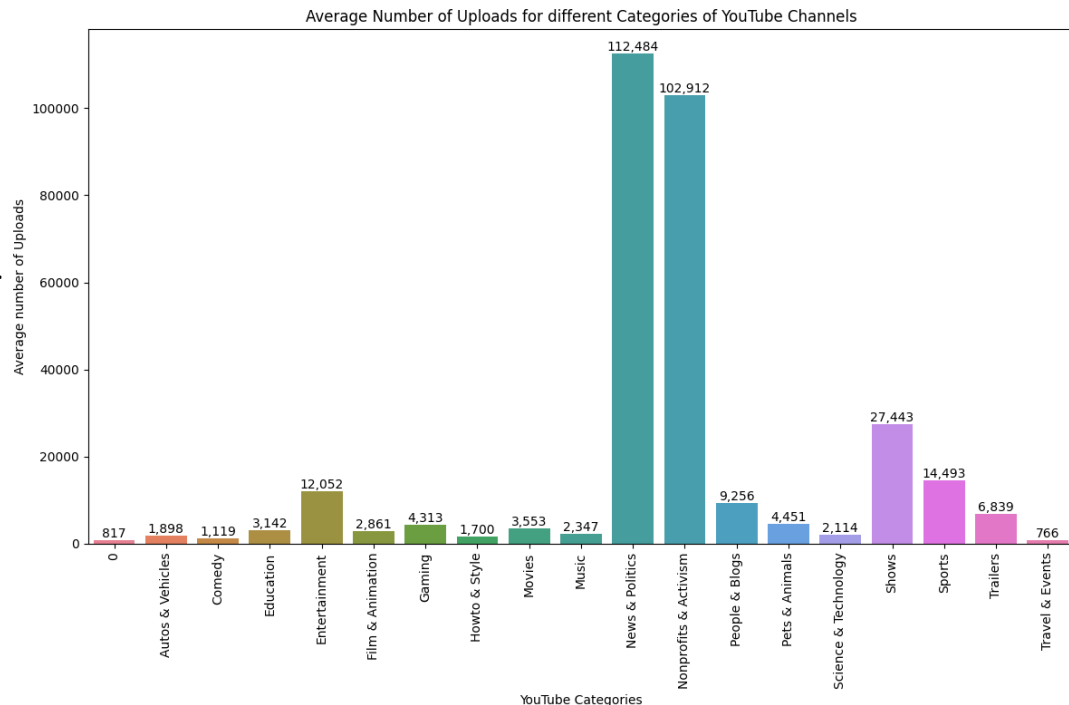


Average Uploads by Category

- Differences in upload frequencies across categories.
- Factors: content type, production schedules.

Average uploads by category:

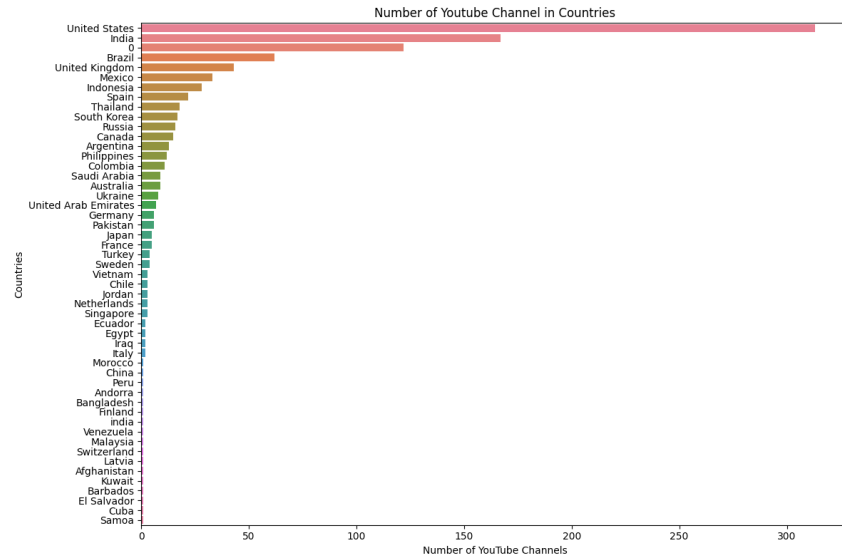
	category	avg_upload
0		817.339623
1	Autos & Vehicles	1898.500000
2	Comedy	1119.753623
3	Education	3142.866667
4	Entertainment	12052.445378
5	Film & Animation	2861.844444
6	Gaming	4313.414894
7	Howto & Style	1700.794872
8	Movies	3553.000000
9	Music	2347.129353
10	News & Politics	112484.384615
11	Nonprofits & Activism	102912.000000
12	People & Blogs	9256.793893
13	Pets & Animals	4451.500000
14	Science & Technology	2114.058824
15	Shows	27443.692308
16	Sports	14493.727273
17	Trailers	6839.000000
18	Travel & Events	766.000000



Top Countries with Most YouTube Channels

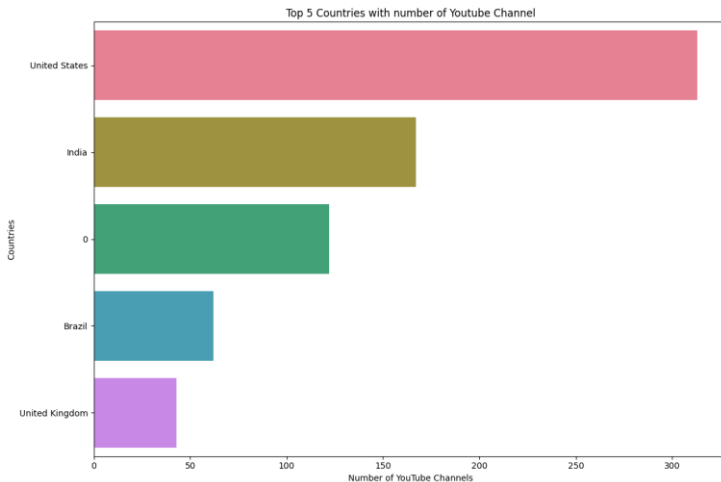
7

- List of top countries .
- Factors: population, internet penetration, content creation culture.



Top 5 Countries by Number of YouTube Channels:

	Country	top_5_country
0	United States	313
1	India	167
2	Brazil	122
3	United Kingdom	62
4		43

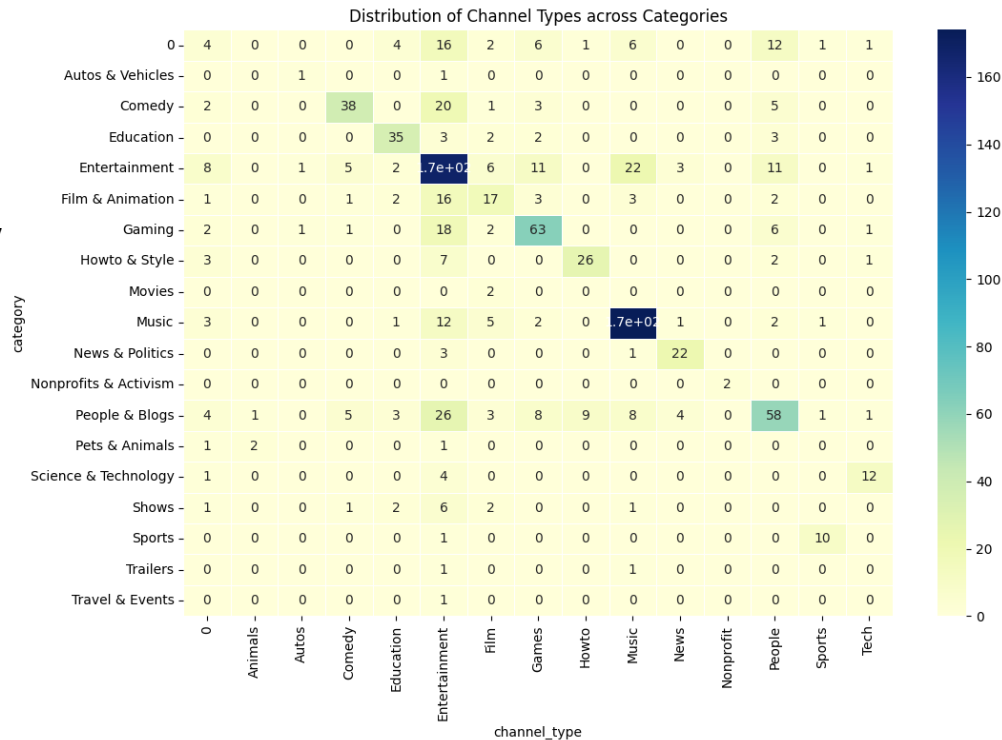


8

- 

```
category    channel_type  group_count
```

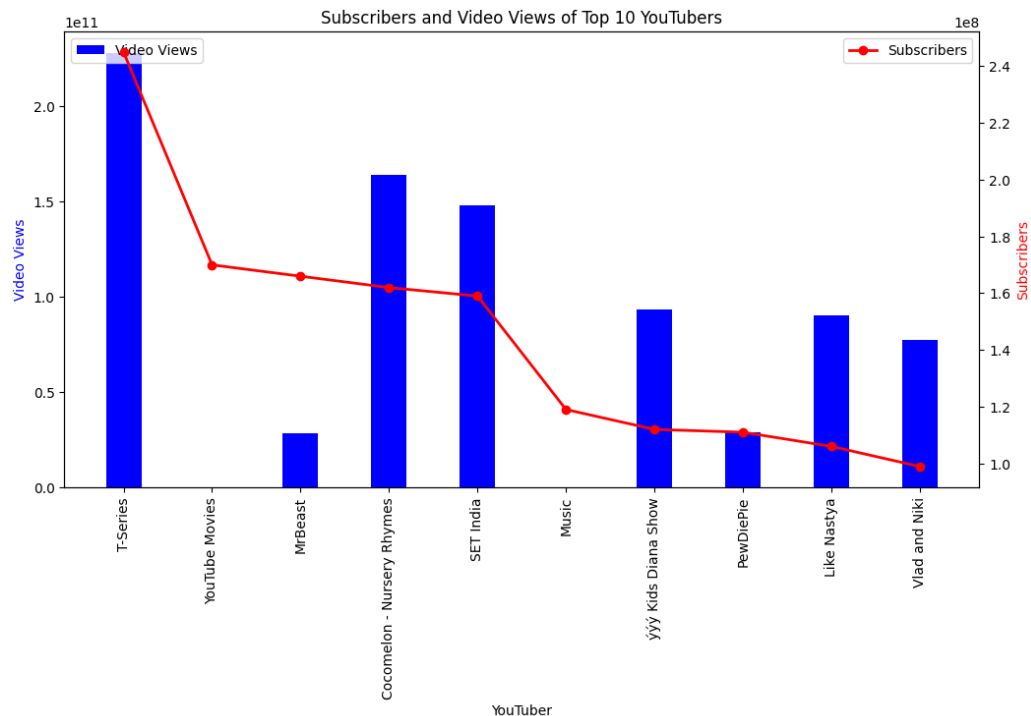
```
[99 rows x 3 columns]
```



Subscribers vs. Video Views for Top YouTubers

- Correlation between subscribers and video views .
- Relationship analysis: subscriber count and video views.

	Youtuber	subscribers	video views
0	T-Series	245000000.0	2.280000e+11
1	YouTube Movies	170000000.0	0.000000e+00
2	MrBeast	166000000.0	2.836884e+10
3	Cocomelon - Nursery Rhymes	162000000.0	1.640000e+11
4	SET India	159000000.0	1.480000e+11
5	Music	119000000.0	0.000000e+00
6	ýýý Kids Diana Show	112000000.0	9.324704e+10
7	PewDiePie	111000000.0	2.905804e+10
8	Like Nastya	106000000.0	9.047906e+10
9	Vlad and Niki	98900000.0	7.718017e+10



Monthly Earnings Distribution by Category

10

- Average monthly earnings (lowest and highest) by category.
- Variation in earnings across categories.

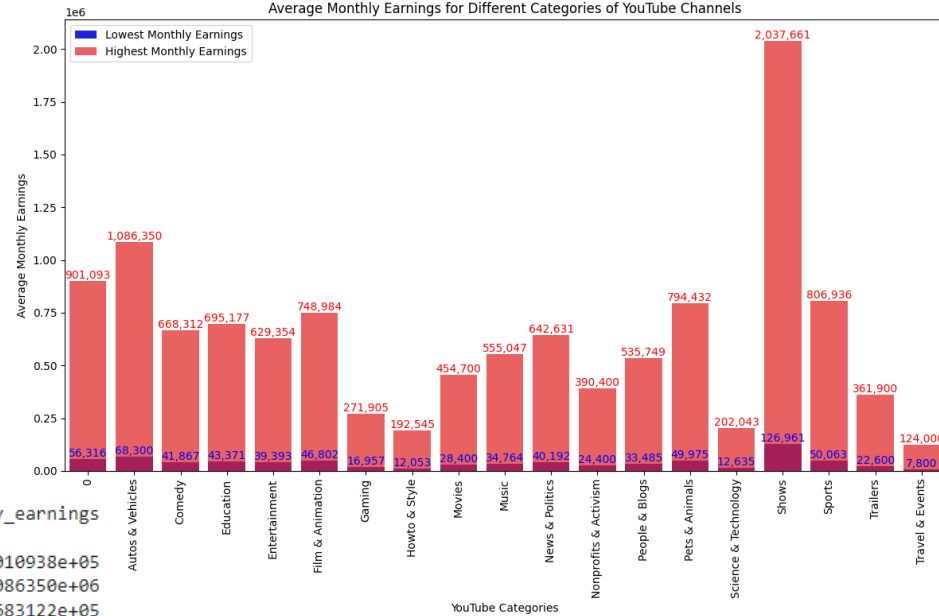
Monthly Earnings by Category:

category	lowest_monthly_earnings	category
0	56316.078302	Autos & Vehicles
Autos & Vehicles	68300.000000	Comedy
Comedy	41867.536232	Education
Education	43371.311333	Entertainment
Entertainment	39393.724370	Film & Animation
Film & Animation	46802.533556	Gaming
Gaming	16957.979574	Howto & Style
Howto & Style	12053.692821	Movies
Movies	28400.000000	Music
Music	34764.449801	News & Politics
News & Politics	40192.625000	Nonprofits & Activism
Nonprofits & Activism	24400.000000	People & Blogs
People & Blogs	33485.993969	Pets & Animals
Pets & Animals	49975.500000	Science & Technology
Science & Technology	12635.411765	Shows
Shows	126961.538462	Sports
Sports	50063.636364	Trailers
Trailers	22600.000000	Travel & Events
Travel & Events	7800.000000	

highest_monthly_earnings

9.010938e+05
1.086350e+06
6.683122e+05
6.951778e+05
6.293549e+05
7.489841e+05
2.719054e+05
1.925452e+05
4.547000e+05
5.550474e+05
6.426320e+05
3.904000e+05
5.357493e+05
7.944322e+05
2.020432e+05
2.037662e+06
8.069364e+05
3.619000e+05
1.240000e+05

Average Monthly Earnings for Different Categories of YouTube Channels



Subscribers Gained in Last 30 Days

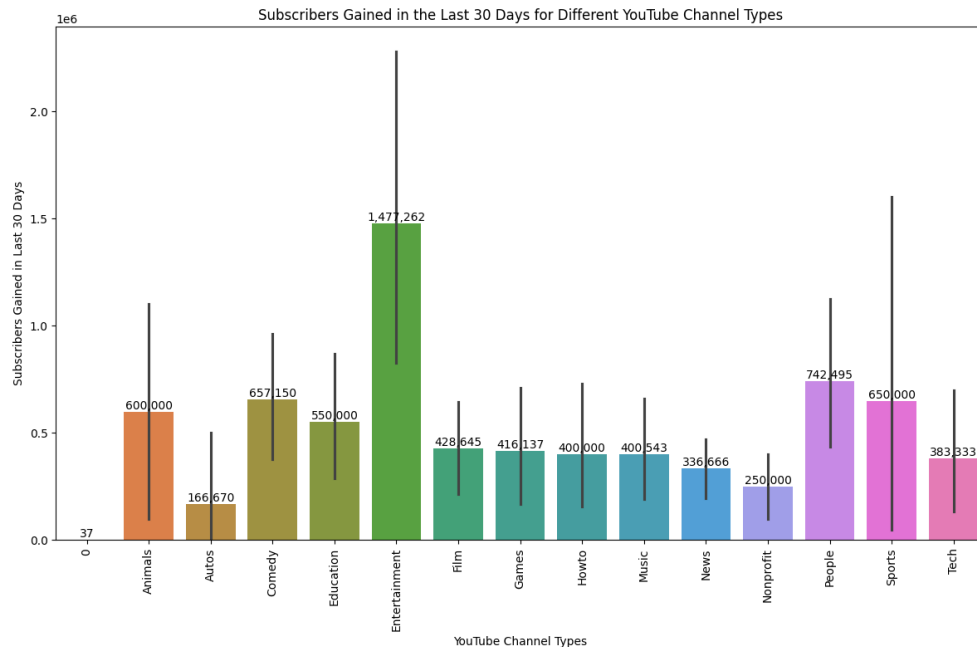


- Subscribers gained in the last 30 days by channel type.
- Patterns: subscriber growth across channel types.

Subscribers gained in the last 30 days across all channels:

	channel_type	subscribers_for_last_30_days	sub_gained
0	0	0.0	17
1	0	1.0	3
2	0	2.0	3
3	0	5.0	2
4	0	6.0	1
..
170	Tech	100000.0	10
171	Tech	200000.0	1
172	Tech	400000.0	1
173	Tech	500000.0	1
174	Tech	1100000.0	1

[175 rows x 3 columns]



Outliers in Highest Yearly Earnings

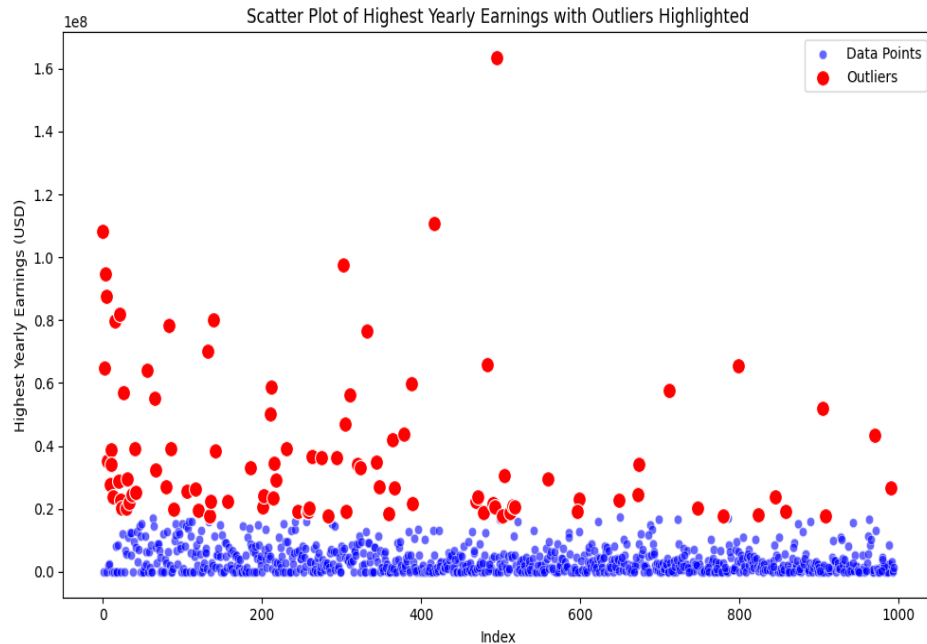


- Scatter plot highlighting outliers in highest yearly earnings.
- Potential reasons: content quality, fan base, monetization strategies.

```

0          T-Series
1          MrBeast
2  Cocomelon - Nursery Rhymes
3          SET India
4    ýýý Kids Diana Show
..          ...
94        Peet Montzingo
95          GH'S
96        Susy Mouriz
97        _vector_
98        Natan por Aİç
  
```

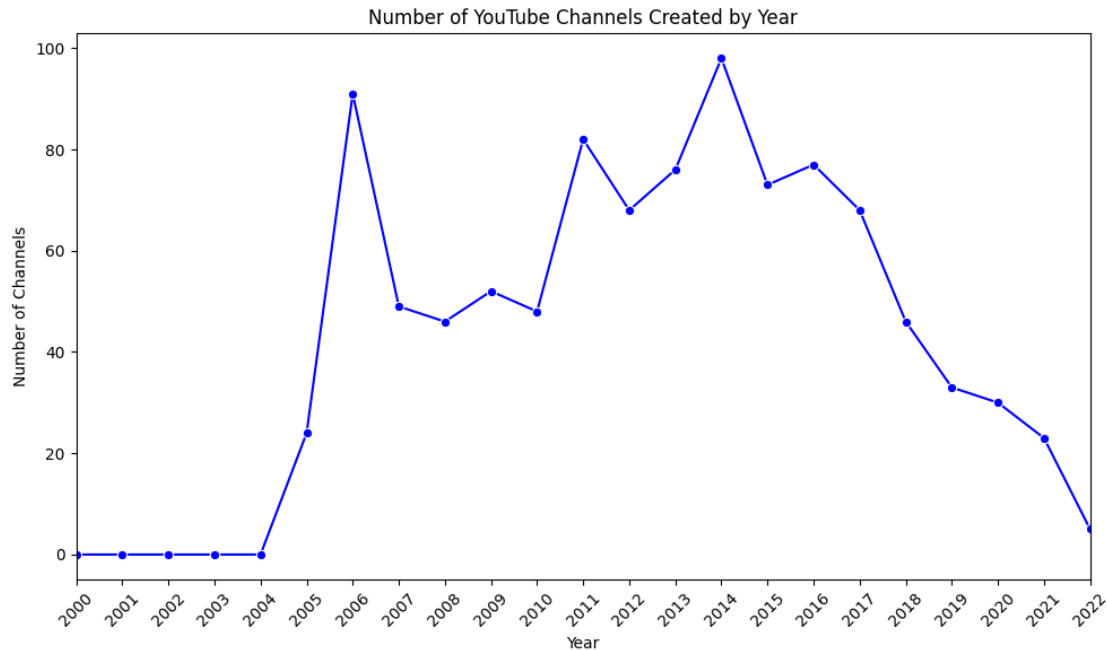
[99 rows x 1 columns]



Channel Creation Trend



- Line plot showing channels created by year.
- Trends and surges in channel creation over time.



Key Findings and Conclusions



Conclusions

This analysis offers valuable insights into the factors driving success on YouTube, from influential creators and content preferences to the impact of geographic and economic influences. By understanding these dynamics, aspiring content creators and data enthusiasts can gain inspiration and guidance for navigating the evolving landscape of online video content, charting their own path to success on the platform.



Summary

- **Influential Creators and Content Preferences:** Discover the leading YouTube channels by subscriber count and video views, highlighting the most influential creators and popular content categories that resonate with global audiences.
- **Geographic and Economic Influences:** Examine how geographic distribution, cultural factors, and economic indicators like unemployment rates and urban populations shape content creation, consumption patterns, and channel growth.
- **Success Strategies and Growth Dynamics:** Uncover the key elements of successful content creation, including engagement metrics, monetization strategies, and pivotal moments that drive the growth of YouTube's top channels..