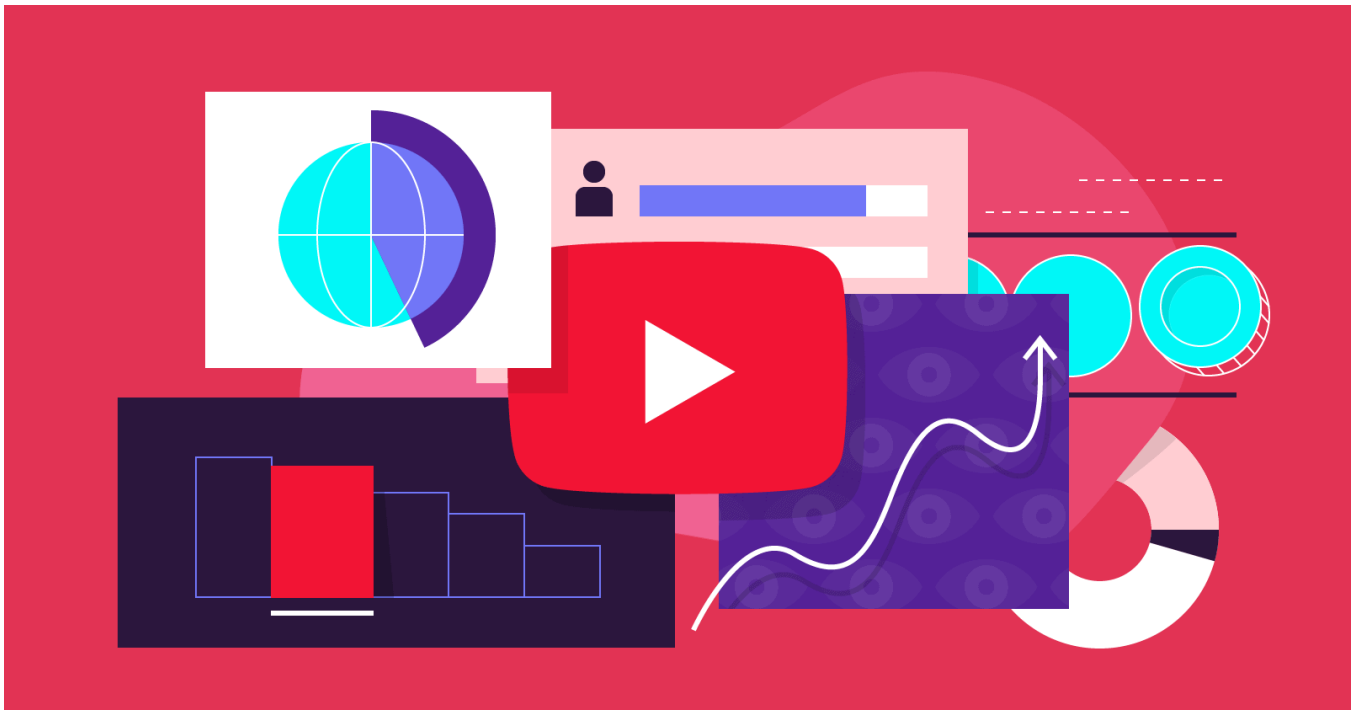
Data set on Global YouTube Statistics 2023

A collection of YouTube giants.



GROUP 34

M.A.S.NETHMINI 16659

H.S.S.PERERA- 16661

M.O.P.PEIRIS -16953

B.S.R.Mendis -16658

M.P.T.T.PATHIRANA-16952

Table of Contents

Acknowledgement

This is to express our warm gratitude to Dr. Dilhari Atigala for feeding us with knowledge and guiding us through this project. Also our thanks to all of our demostrators of ST1009 for guidance and pointing out any deficiencies. We are grateful to the University of Colombo for providing this valuable opportunity. Last but not the least we thank our parents and friends for supporting us directly and indirectly throughout the project

Executive summary

Our data set consists of around 900 individuals, 9 Qualitative variables and 8 Quantitative variables. Therefore, this is a quite large data set. We created a Histogram, a Boxplot and Scatter plots to analyze monthly earning (mean of monthly highest earning and monthly lowest earning). Then some Bar graphs, Dot plots, a Pie chart are created to analyze about video views, subscribers, categories, created year and their relationships.

Since this is a project of identifying how much YouTube pays for youtube views, we had to extract the behaviour of monthly earnings and video views with other variables. Then we examined monthly earnings for YouTube channels, relationships about subscribers and video views between other variables such as category, country.

As the results of our analysis, we found that there is a good relationship between video views and monthly earnings. Furthermore we could find the specific categories of youtube channels that would be more beneficial for youtubers to get a high number of video views and subscribers in order for them to earn more and also the most efficient countries to create these youtube channels with the resources available.

Finally we think that this project suggests the most beneficial ways for new youtubers to create their videos to get more video views and finally a good a monthly earning.

Introduction

YouTube's global impact is undeniable, as it continues to shape how we consume, create, and engage with video content. Its influence spans entertainment, education, culture, and beyond, making it a central player in the digital landscape. As YouTube evolves and grows, it remains a fascinating subject of study and a powerful medium for communication and expression on a global scale.

YouTube is an American online video sharing and social media platform headquartered in San Bruno, California, United States. Accessible worldwide, it was launched on February 14, 2005, by Steve Chen, Chad Hurley, and Jawed Karim. It is owned by Google and is the second most visited website, after Google Search. YouTube has more than 2.5 billion monthly users, who collectively watch more than one billion hours of videos each day. As of May 2019, videos were being uploaded at a rate of more than 500 hours of content per minute.

YouTube include music videos, video clips, news, short films, feature films, songs, documentaries, movie trailers, teasers, live streams, vlogs, and more. Most content is generated by individuals, including collaborations between YouTubers and corporate sponsors.

You tube is a great way to earn money. Earning money from YouTube requires dedication, consistency, and a strong understanding of your audience and niche. To earn money from YouTube basically the channel must have at least 1,000 subscribers and 4,000 watch hours in the last 12 months. That way we can say that videos can be monetized with ads as the number of views increases

Objectives of the study

study, is to identify the income that can be earned as a YouTuber from views.

Conclusions about

- monthly earning
- relationships of video views
- relationships of subscribers
- relationships of category
- relationships of created year by studying

Methodology

The methods used to discover the behavior of the collected data and organize them effectively are very important. Minitab and Microsoft Excel statistical software are used for this analysis. Analysis is done using graphical methods such as bar graphs, pie charts, histograms, scatter plots and box plots. Mean, standard deviation, and quartiles are some of the numerical statistical summaries used here

- Minitab 19 Software

  https://www.minitab.com/en-us/products/minitab/free-trial/
- Microsoft Excel Software

  https://support.microsoft.com/en-us/excel

Types of chart

- Bar chart – Bar chart is a chart or graph that presets categorical data with rectangular bars with heights or lengths proportional to the values that they represent.
- Pie chart – A pie chart is a circular statistical graphic, which is divided into slices to illustrate numerical proportion.
- Dot plot – A dot plot is a simple from of data visualization that consists of data points plotted as dots on a graph with an x- and y- axis.
- Histogram – This is similar appearance into bar graph. A histogram is a graphical representation that organizes a group of data points into user-specified ranges.
- Boxplot – In descriptive statistics, a boxplot is a method for graphically depicting groups of numerical data through their quartiles.
- Scatter Plot – A scatter plot is a type of data display that shows the relationship between two numerical variables.

Data

Minitab - Minitab (1).mpx

File Edit Data Calc Stat Graph View Help Assistant Additional Tools

	C6-T	C7-T	C8	C9	C10	C11	C12	C13	C14	C15-D	C16	C1
	Country	channel_type	country_rank	video_views_for_the_last_30_day	lowest_monthly_earnings	highest_monthly_earnings	lowest_yearly_earnings	highest_yearly_earnings	created_year	created_month	Unemployment rate	monthly
1	India	Music	1	2258000000	564600	9000000	6800000	108400000	2006	Mar	5.36	4
2	United States	Games	7670	12	0	0	0	1	2006	Mar	14.70	
3	United States	Entertainment	1	1348000000	337000	5400000	4000000	64700000	2012	Feb	14.70	2
4	United States	Education	2	1975000000	493800	7900000	5900000	94800000	2006	Sep	14.70	4
5	India	Entertainment	2	1824000000	455900	7300000	5500000	87500000	2006	Sep	5.36	3
6	United States	Entertainment	3	731674000	182900	2900000	2200000	35100000	2015	May	14.70	1
7	other	Entertainment	1	39184000	9800	156700	117600	1900000	2010	Apr	2.29	
8	other	People	5	48947000	12200	195800	146800	2300000	2016	Jan	4.59	
9	United States	Entertainment	5	580574000	145100	2300000	1700000	27900000	2018	Apr	14.70	1
10	India	Music	3	803613000	200900	3200000	2400000	38600000	2014	Mar	5.36	1
11	United States	Sports	6	714614000	178700	2900000	2100000	34300000	2007	May	14.70	1
12	South Korea	Music	1	498930000	124700	2000000	1500000	23900000	2016	Jun	4.15	1
13	India	Entertainment	5	1657000000	414300	6600000	5000000	79600000	2007	Aug	5.36	3
14	United Kingdom	Entertainment	4797	1	0	0	0	0	2020	Jul	3.85	
15	South Korea	Music	2	168290000	42100	673200	504900	8100000	2012	Dec	4.15	
16	United States	Entertainment	6266	16	0	0	0	1	2006	Jan	14.70	
17	Canada	Music	1	176326000	44100	705300	529000	8500000	2007	Jan	5.56	
18	South Korea	Music	3	598173000	149500	2400000	1800000	28700000	2008	Jun	4.15	1
19	India	Entertainment	6	1707000000	426800	6800000	5100000	81900000	2005	Dec	5.36	3
20	United States	Education	8	473387000	118300	1900000	1400000	22700000	2011	Dec	14.70	1
21	Brazil	Music	1	447223000	0	0	0	0	2012	Mar	12.08	
22	India	Education	7	420292000	105100	1700000	1300000	20200000	2013	Feb	5.36	
23	India	Music	8	254961000	63700	1000000	764900	12200000	2010	Jun	5.36	
24	India	Entertainment	9	1188000000	296900	4800000	3600000	57000000	2008	Jun	5.36	2
25	India	Music	5803	10	0	0	0	0	2018	Aug	5.36	
26	United States	Sports	9	141200000	35300	564800	423600	6800000	2009	Mar	14.70	
27	United States	Film	9	114668000	28700	458700	344000	5500000	2006	Apr	14.70	
28	India	Music	11	422634000	105700	1700000	1300000	20300000	2007	May	5.36	
29	other	Music	1	611828000	153000	2400000	1800000	29400000	2011	Jun	9.79	1
30	India	Music	12	232025000	58000	928100	696100	11100000	2014	Oct	5.36	
31	India	News	13	461148000	115300	1800000	1400000	22100000	2009	Aug	5.36	

You Tube Statistics

You Tube Statistics 100%

This data set is abstracted from a well- known website specially designed for the purpose of providing data set samples ready for analyses, called KAGGLE. Kaggle is an online community platform for data scientists and machine learning enthusiasts. Kaggle allows users to collaborate with other users, find and publish datasets, use GPU integrated notebooks, and compete with other data scientists to solve data science challenges. The aim of this online platform (founded in 2010 by Anthony Goldbloom and Jeremy Howard and acquired by Google in 2017) is to help professionals and learners reach their goals in their data science journey with the powerful tools and resources it provides. As of today (2021), there are over 8 million registered users on Kaggle.

This dataset contains **869** observations.

Data source : https://www.kaggle.com/datasets/nelgiriyewithana/global-youtube-statistics-2023

Qualitative Variables

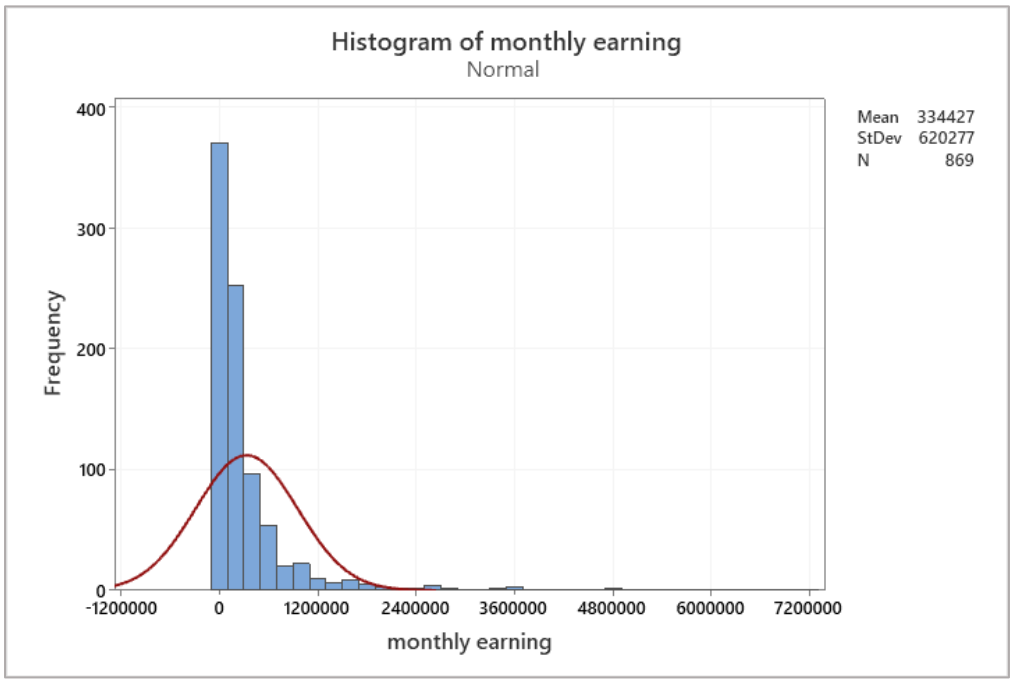
- Rank - Depending on the number of subscribers, the rating of the channel on YouTube
- Youtuber – Name of the YouTube channel
- Category – “Music”, “Film & Animation”, “Entertainment”, “Education”, ”Shows”, “People & Blogs”, “Comedy”, “News & Politics”, “Howto & Style”, “News & Politics”
- Country – Country where YouTube channel originates Country rank
- Channel_type - “Music”, “Games”, “Entertainment”, “Education”, “Music”, “People”, “Sports”Created year
- Created year– Year which is YouTube channel created
- Created month - Month which is YouTube channel created
- Unemployment rate - Ranking countries according to the you tubers who haven't any job

Quantitative Variables

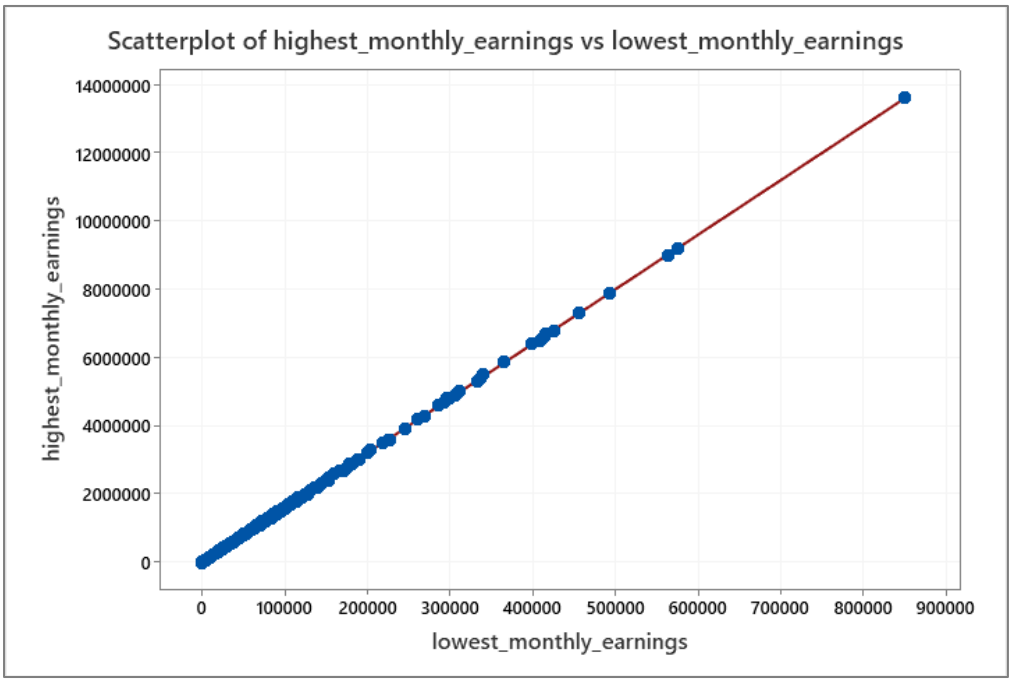
- Subscribers – Number of subscribers of the channel
- Uploads – Total number of uploaded videos on the channel
- Country rank - Rank of the channel based on the total amount of subscribers within its country
- video_views_for_the_last_30_days - Views for all videos in the past 30 days
- lowest_monthly_earnings - Maximum monthly earnings estimates for the channel
- highest_monthly_earnings - Minimum monthly earnings estimates for the channel
- lowest_yearly_earnings - Maximum yearly earnings estimates for the channel
- highest_yearly_earnings - Minimum yearly earnings estimates for the channel
- monthly earning – Mean of the lowest and highest monthly earning

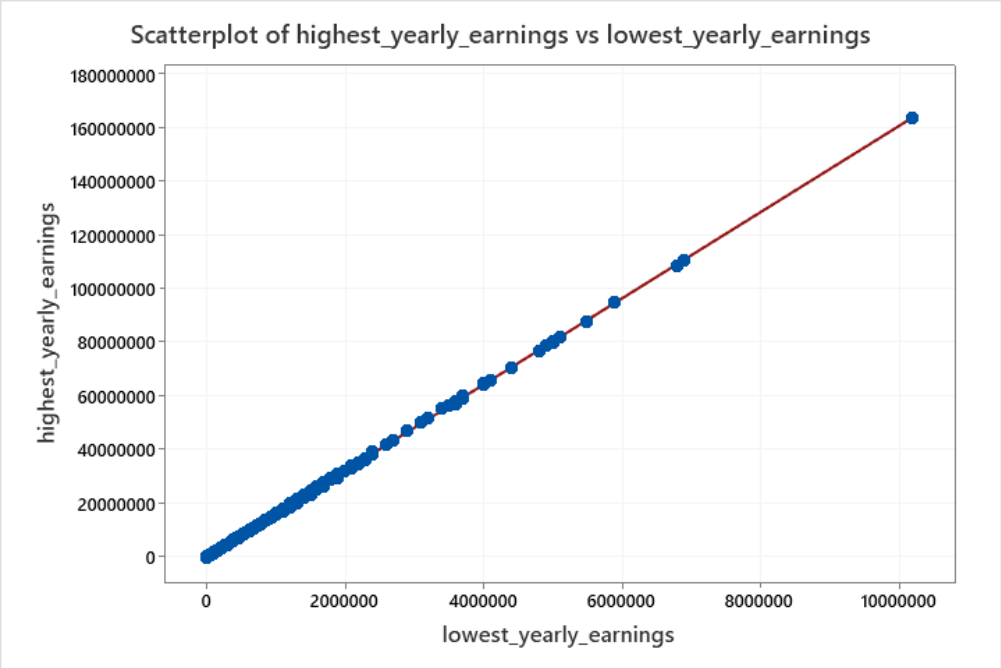
Analysis

Earning of youtube channels

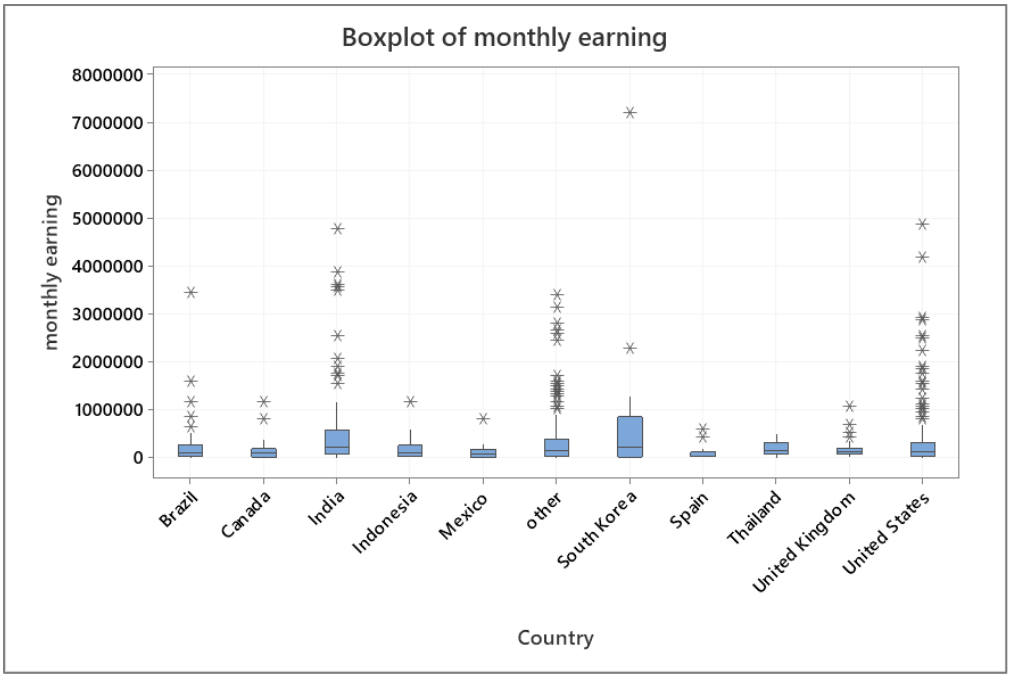


We used this histogram to represent variable of monthly earnings. It is created as a mean of the two variables “highest monthly earning” and “lowest monthly earning”. In this histogram has long right tail then it is a positively skewed distribution because the mean of the above variable is 334427 while the Median is 134500. The mode of monthly earning is represent in first bar and it range is 0 to 100000.



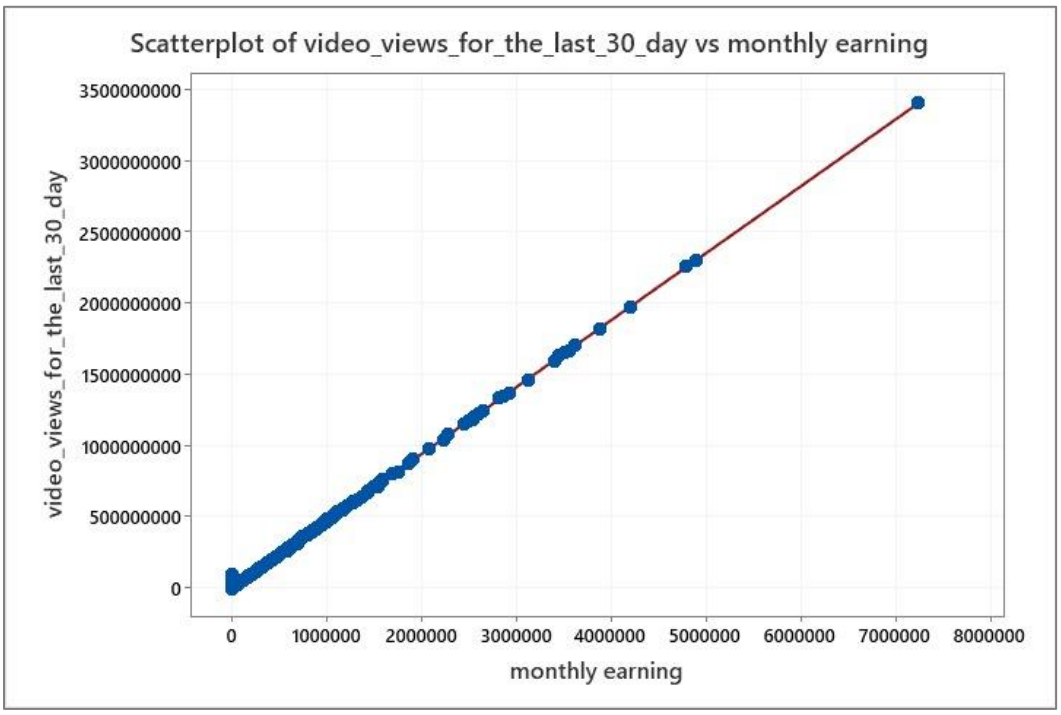


Scatter plot is used to show a relationship between two variables. In these scatter plots is shown relationships about ("highest monthly earning" and "lowest monthly earning") and ("highest yearly earning" and "lowest yearly earning"). We can say, in this all of the observations are on the line therefore this is strong positive correlation in these scatter plot.

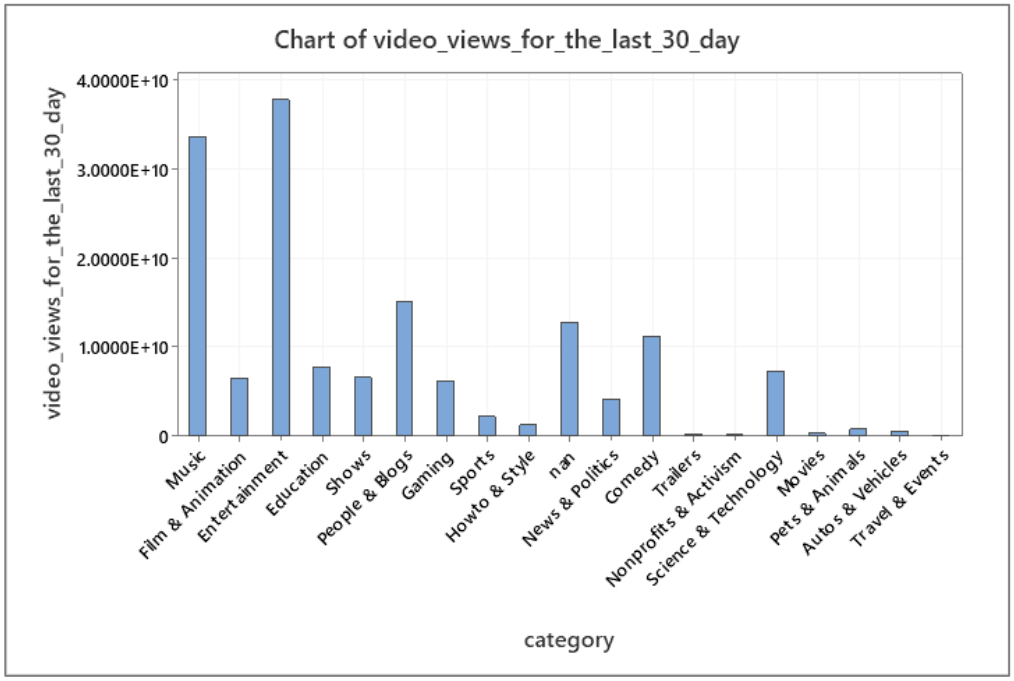


In this boxplot shows relationship between monthly earning and countries. In this described median of monthly earning in multiple countries. The X axis represent countries and there are 11 categories as South Korea, India, Brazil, United status etc. South Korea has most IQrange of monthly earning and it is 833350. Outliers in the graph can be important to identify because they may indicate errors in data collection or be genuinely unusual data points.

Video views



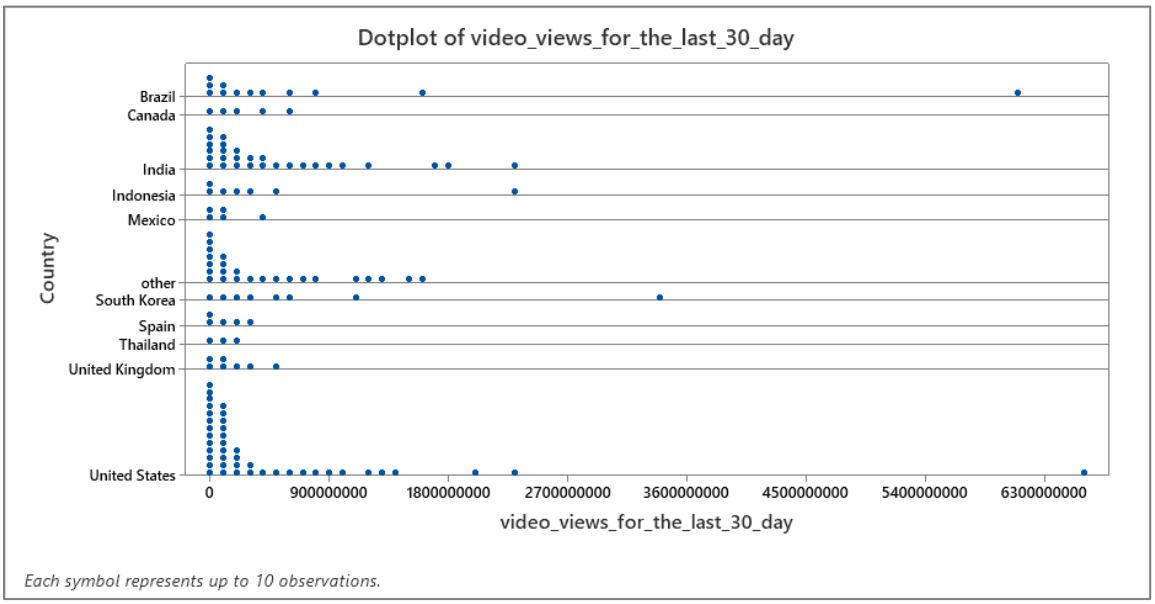
In this scatter plot show a relationship about monthly earning and video views. The “x axis” is represented about monthly earning and the “Y axis “ is represented about video views for the last 30 days. In this scatter plot we can see a strong relationship between these two variables. So we can get an idea, increasing video views is a reason to get good monthly income.



By studying the above bar graph drawn to observe the behavior of the video views for the last 30 days with the categories of the YouTube channels, we can conclude that the category of entertainment has the highest number of video views in the last 30 days.

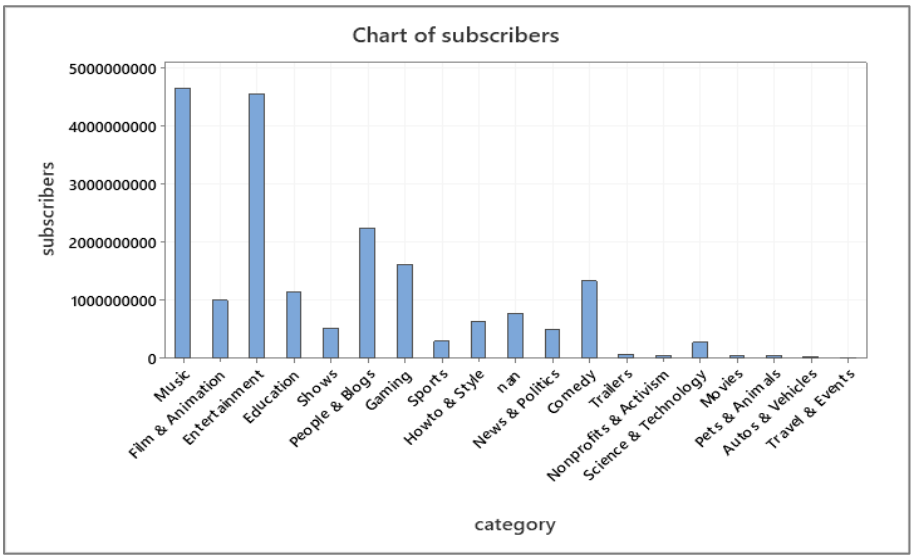
It can be observed that the categories of entertainment and music have exceptionally higher values of video views compared to the other categories like education, news and politics, shows etc.

Categories such as trailers and movies, even though are part of entertainment, have gotten the least amount of video views as it can be observed in the above graph. Therefore, it can be stated that the views not only depend on the content of the video but also on factors such as the price the viewers have to bear to view that particular video on You Tube, the existence of better social media platforms to view them more efficiently, limitation of time for the viewers with their busy lives, etc.



This dot plot shows relationship about video views for last 30 days between country. It can be observe that all of the countries had the most amount of video views less than 5 billion. We can see that united states, Brazil and South Korea have exceptionally high video views for last 30 days.

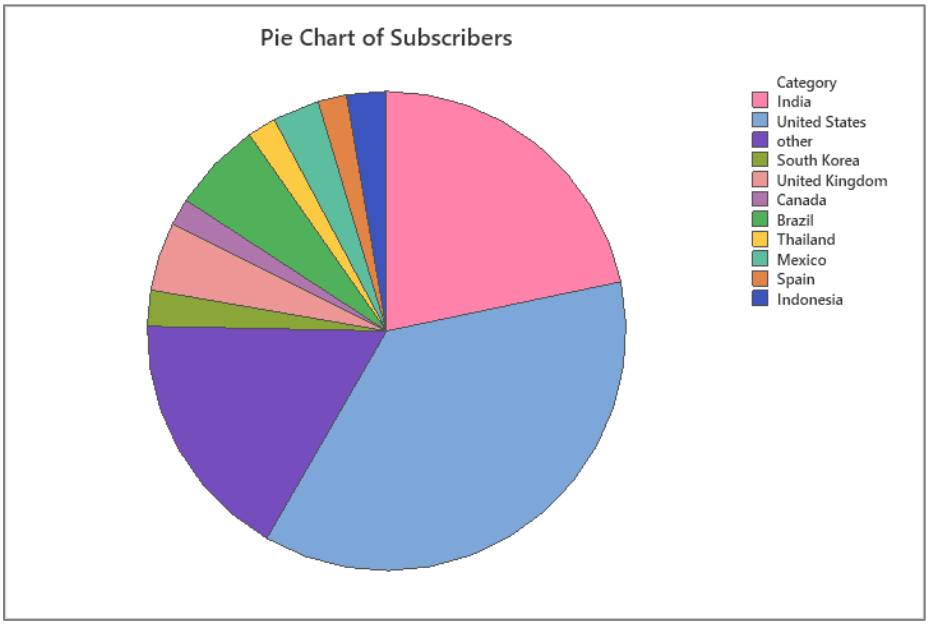
Subscribers



By studying the above bar graph drawn to observe the variation of the number of subscribers with the categories of the You Tube channels, we can observe that the categories of music and entertainment have an extremely higher number of subscribers and music has a slight number of subscribers more than entertainment.

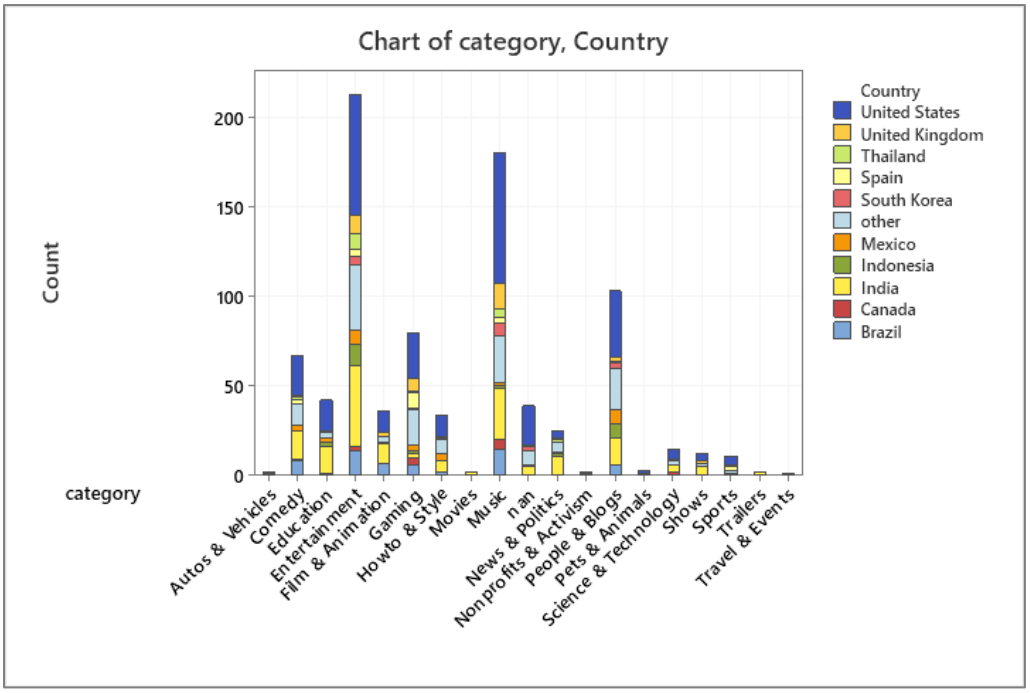
It can be stated that the viewers of the music and entertainment channels have subscribed to these channels in large numbers compared to the others as videos of these categories bring happiness and relaxation to the very busy and hectic lives people lead.

It also can be observed that the categories of travel and events, autos and vehicles, pets and animals, have the least number of subscribers.

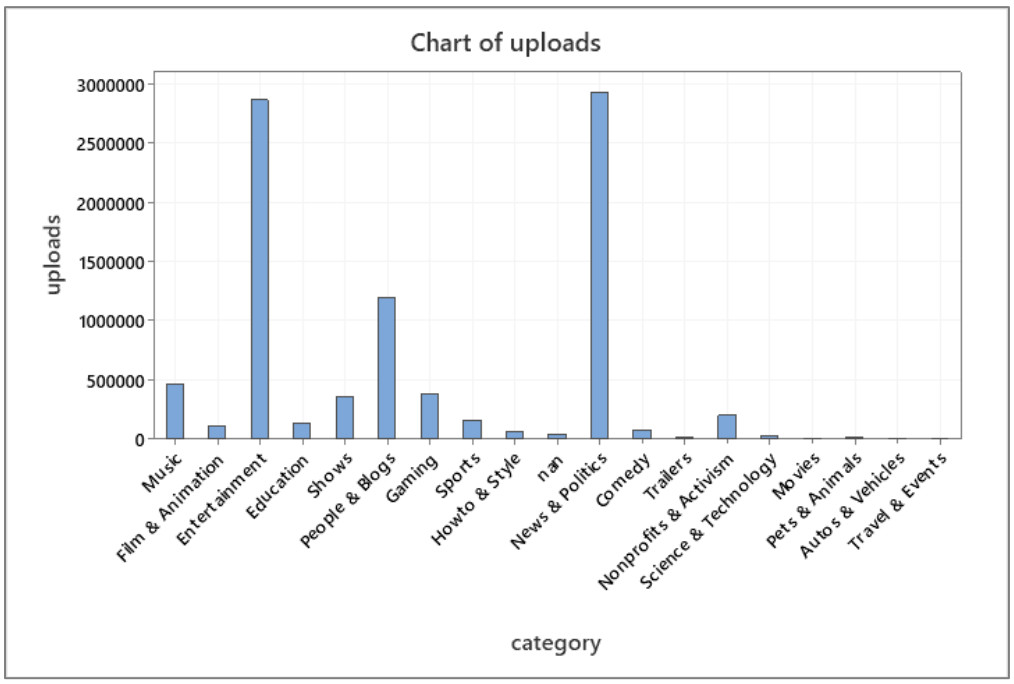


The above graph shows the relationship between the number of subscribers of different countries. It can be observed that United States has the highest number of percentage and the second highest number of percentage is from India. Countries such as Thailand, Spain, Canada have very less amount of subscribers in them.

Categories

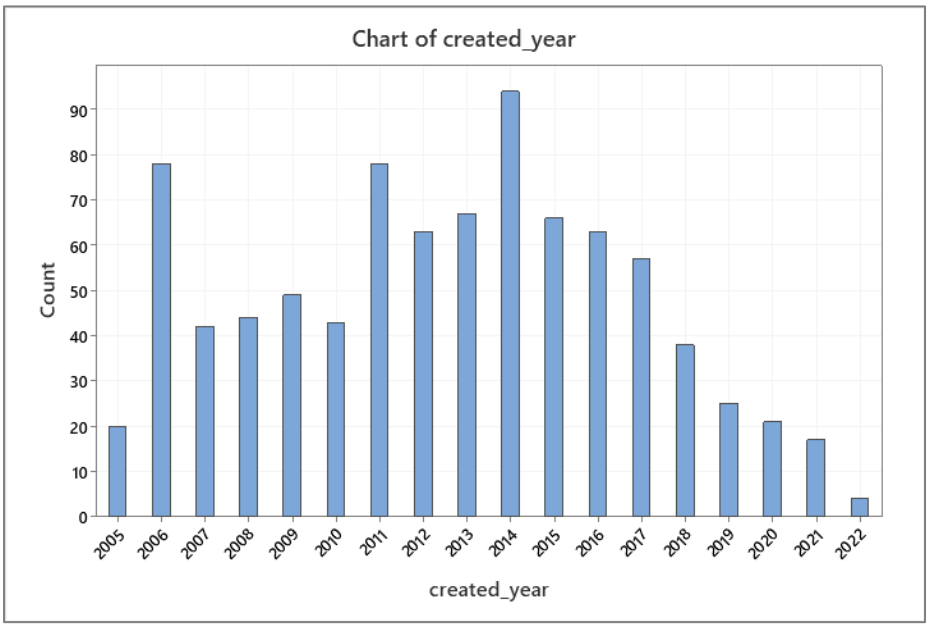


This bar chart shows that the relationship between category and the country. There are more than 10 countries including United States, United Kingdom, Thailand, Spain. There are 19 categories including Entertainment, Music, People & Blogs. The most popular category among countries is Entertainment and least popular category is Travel & Events category. When we consider the 'Entertainment' category, United States is the most viewed country in that category and Canada is the least viewed country. In the music category also the most viewed country in the 'Music' category is the United States. Briefly we can say that United States is the country which is mostly using YouTube.

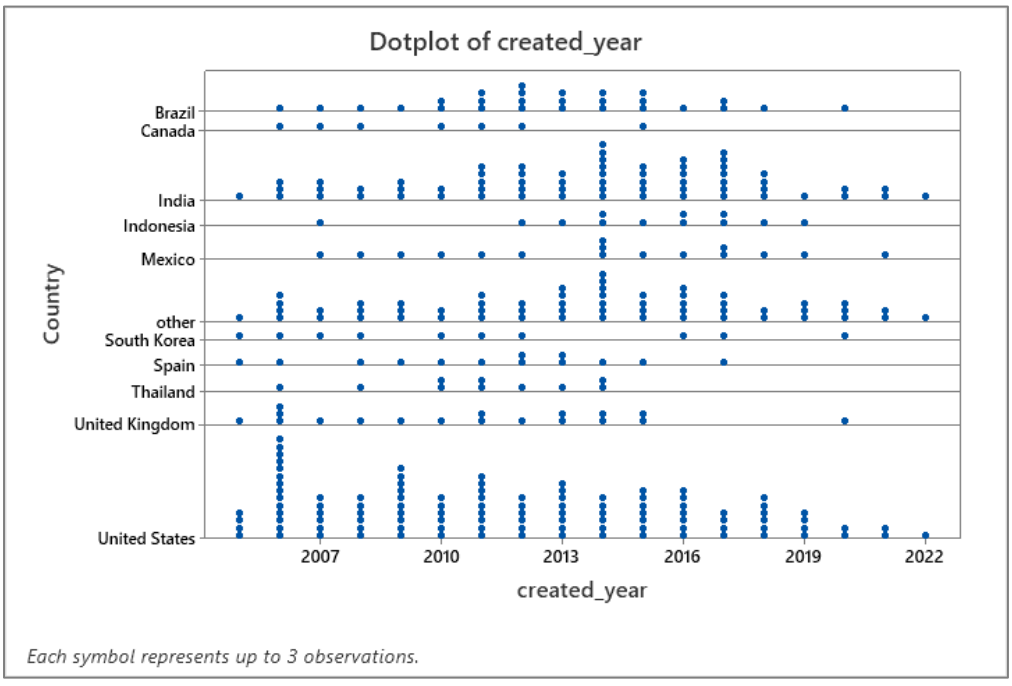


This bar graph shows the number of uploads are in various categories. The large proportion of uploads are in News and Politics category and the small proportion of uploads are in Travel and Events category. The distance between the Entertainment category and the News and Politics category is the few number of uploads. This shows there is a relationship between the category and the uploads.

Created year



This bar graph describes the variable created_year when the YouTube channel was created. The x-axis represents the created_year while the y-axis of the graph represents the count for the each year. The variable created year represents the years from 2005 to 2022. The year 1970, entered incorrectly. So we removed it as youtube starts from the year 2005. In the year 2014, the most number of youtube channels were created and least were created in the year 2022. In the years 2006 and 2011, the number of channels created are same



This dot plot illustrates frequency of created year according to countries . United states have highest number of you tube channels and 2006 is highest frequency. Also, it is highest frequency in the whole plot. Approximately in year 2014, all countries have started to create lots of you tube channels. In year 2022 has lowest frequency. Also, lots of countries not started channel in year 2022. By considering this dot plot we can say 2011-2017 period all countries started to create a considerable amount of channels.

Conclusions & Recommendations

- Good YouTube creator can earn sufficient money for his life without job because the histogram of monthly earnings represent mode of monthly earnings 0- 100000 range
- Lowest video views is the reason we can see zero earning in scatter plot about monthly and yearly earnings.
- We can choose most video views and subscribers of categories as music, entertainment, peoples & blogs. It can be recommended for a new you tuber to create a new you tube channel on the above categories.
- If you create entertainment, people & blogs and news & politics you have to always upload videos necessarily.
- The highest number of subscribers and video views can be observed in united states, spain and united kingdom so it can be concluded that is easier for you tubers in these countries than other countries to earn money because they have better resources.
- After you tube was created in 2005 the number of you tube channels created has increased in 2006 because you had become popular those days. after 2014 making of you tube channels have decreased because new social media platform such as instergram, tiktok, snapchat, threads are popular than you tube in these days.

References

Global YouTube Statistics 2023. (2023, July 28). Kaggle.

https://www.kaggle.com/datasets/nelgiriyewithana/global-youtube-statistics-2023

Wikipedia contributors. (2023). YouTube.

https://en.wikipedia.org/wiki/YouTube

Make Money Matt. (2023, March 1). *How Much YouTube Pays You For 1,000 Views In 2023*

https://www.youtube.com/watch?v=KA5FmevAzTc