

PROFESSIONAL TRAINING REPORT

AT

**Sathyabama Institute of Science and Technology
(Deemed to be University)**

Submitted in partial fulfillment of the requirements for the award
Of Bachelor of Technology Degree in Information Technology

By

RAJA REGAN A

40120071



**DEPARTMENT OF INFORMATION TECHNOLOGY
SCHOOL OF COMPUTING**

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

Accredited with Grade "A" by NAAC

JEPPIAAR NAGAR, RAJIV GANDHI SALAI,

CHENNAI - 600 119.TAMILNADU

APRIL-2023



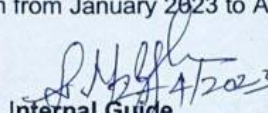
SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

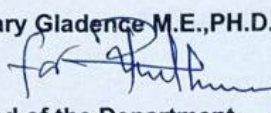
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the Bonafide work of **RAJA REGAN A, 40120071** who carried out the project entitled "**CHANNEL VIEWS PREDICTION AND ANALYSIS**" under my supervision from January 2023 to April 2023.

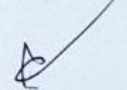

Internal Guide

(DR.L.Mary Gladence M.E.,PH.D.,)


Head of the Department

(DR.R.SUBHASHINI M.E.,PH.D.,)

Submitted for Viva voce Examination held on 28/4/23


Internal Examiner


External Examiner

DECLARATION

I, RAJA REGAN A, hereby declare that the Project Report entitled "**CHANNEL VIEWS PREDICTION AND ANALYSIS**" in power bi done by me under the guidance of Dr.L.MARY GLADENCE M.E.,Ph.D, (At Sathyabama Institute of Science and Technology (Deemed to be University),Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600119 is submitted in partial fulfillment of the requirements for the award of Bachelor of Technology degree in Information Technology

DATE: 28/4/23

PLACE: Chennai

A. Regan
SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board of Management of SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr.T.SASIKALA M.E.,Ph.D, Dean**, School of Computing **Dr.R.Subhashini M.E.,Ph.D**, Head of the Department of Information Technology for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr.L.Mary Gladence M.E.,Ph.D**, for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **Department of Information Technology** who were helpful in many ways for the completion of the project

TRAINING CERTIFICATE

No. 34351


एन एस आई सी
N S I C

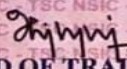
राष्ट्रीय लघु उद्योग निगम लिमिटेड
NATIONAL SMALL INDUSTRIES CORPORATION LIMITED
(A Govt. of India Enterprise)

NSIC - TECHNICAL SERVICES CENTRE
Sector B-24, Guindy Industrial Estate, Ekkaduthangal, Chennai - 600 032.

CERTIFICATE

This is to certify that Mr. **RAJA REGAN A.** S/o.
Shri. **R. AMUTHAN**, student of III year B.Tech. (IT),
Sathyabama Institute of Science and Technology has
undergone "**SKILL DEVELOPMENT PROGRAM ON**
DATA ANALYTICS" conducted by us for a period of
one week (30 Hours) from 09.03.2023 to
14.03.2023.




HEAD OF TRAINING


HEAD OF CENTRE

ABSTRACT

This report analyzes YouTube, the top video-sharing website. This paper examines the video-sharing site to better understand features and trends of the top YouTube channels. Statistics from Social Blade were used to acquire data. This data includes information such as total view counts and number of subscribers, among other statistics. The report compiles the top 100 YouTube channels based on its number of subscribers. In addition, this report analyzes the top 100 channels to understand what the most frequent genres are as well as to see which genres make the most money. The data on “YouTuber” earnings is well documented on sites such as Celebrity Net Worth. A question this report addresses is what views or subscriptions mean for making money. Number of subscribers and total view count are the two indicators of a video’s popularity and key to understanding what advertisers might value more.

Keywords: subscribers, views, likes, notification, regions, types of videos

TABLE OF CONTENT

CHAPTER	TITLE	PAGE NO
	ABSTRACT	
	LIST OF FIGURES	
	LIST OF ABBREVIATIONS	
1	INTRODUCTION	1
2	AIM AND SCOPE	2
	2.1 AIM	2
	2.2 SCOPE	3
3	SYSTEM REQUIREMENTS	4
4	METHODOLOGY	5
	4.1 Data collection	5
	4.2 Data mining	6
	4.3 Data analysis	7
	4.4 Types of analysis	8
	4.4.1 Descriptive analysis	8
	4.4.2 Diagnostic analysis	9
	4.4.3 Predictive analysis	10
	4.4.4 Prescriptive analysis	11
	4.5 Power bi	12
	4.5.1 Power bi uses	12
	4.5.2 Power bi features	13
	4.6 Architecture	14
	4.7 Project implementation	15
5	RESULT AND DISCUSSION	17
	5.1 Prediction	17

	5.2 Q&A	18
	5.3 Smart narrative	19
	5.4 Views by date	20
	5.5 Language by month	21
	5.6 Shared videos per month	21
	5.7 Subscriber count by month	22
	5.8 Trending videos by month	23
	5.9 Likes by month	23
	5.10 Views by month	24
	5.11 Number of videos per month	25
	5.12 Watch time by month	25
	5.13 Subscriber count by region	26
	5.14 Categories of videos	26
	5.15 Cards (likes, views, subscribers)	28
6	CONCLUSION	29
7	REFERENCE	30

LIST OF FIGURES

FIGURE NO	FIGURE NAME	PAGE NO
4.1	Data mining	6
4.2	System Architecture	14
4.3	Dataset 1	16
4.4	Dataset 2	16
5.1	Prediction table	17
5.2	Q&A	18
5.3	Smart narrative	19
5.4	Views by date	20
5.5	Language by month	21
5.6	Shared videos per month	22
5.7	Subscriber count by month	22
5.8	Trending videos by month	23
5.9	Likes by month	24
5.10	Views by month	25
5.11	Number of videos per month	25
5.12	Watch time by month	26
5.13	Subscriber count by region	26
5.14	Categories of videos	27
5.15	cards	28

LIST OF ABBREVIATION AND EXPANSION

ABBREVIATIONS

EXPANSION

LK	LIKED VIDEOS
SV	SHARED VIDEOS
SUBS	SUBSCRIBERS
YT	YOUTUBE
CMT	COMMENT
FF	FANS FICTION
ETA	ESTIMATED TIME OF ARRIVAL
TBA	TO BE ANNONCED
CPY	COPYRIGHTS

1.INTRODUCTION

YouTube is one of the most popular websites on the Internet and generates an immense amount of traffic everyday. YouTube and Google Video, which owns YouTube, accounted for four times the amount of unique users as the closest competing online video distribution network (Cool, Seitz, & Metrist, 2010, 1). In addition YouTube ranks third on a list of the top online brands, falling behind only Google and Apple (Ibid). Another amazing aspect of YouTube is its rapid growth. In 2006, only one year after its establishment, 70% of YouTube viewers were American. Only two years later the amount of YouTube users who were American had dwindled down to 30%, which demonstrates the global popularity of the site. Further research indicates by 2009 over 152 million Americans had viewed online videos 16.8 billion times (Strangelove, 2010, 11). These statistics demonstrate YouTube's immense scope on our social and demographic landscape. Users who upload numerous videos under a single account on YouTube have what is called "a channel". On a YouTube channel, an individual or an organization can upload a collection of videos, which are available together on a single page. Users' channels vary greatly according to genre, spanning from video game analysis to world politics.

Viewers who take a liking towards a particular channel can subscribe to it in order to stay up to date on news and additional videos uploaded to the channel's page. This report focuses on the top 100 most subscribed YouTube channels to understand what the most popular genres are. Another key area on YouTube is the money making process. While there are few YouTube personalities that make millions off their videos through advertising, the majority of content producers make very little money. For the users that do make a decent earning, there are two main statistics that can dictate the value of a YouTube channel. These two statistics are views and subscribers. This report examines the importance and flaws of both statistics.

2. AIM AND SCOPE

2.1 AIM

The majority of videos uploaded to YouTube are by individuals unaffiliated with companies or organizations. A conducted survey of the 4,320 most popular videos on YouTube indicates that 2,177 of them were coming from beyond the broadcast and established media (Burgess & Green, 2009, 43). This trend holds true when compared to present statistics on the YouTube landscape.

The resource used to acquire the YouTube data is Social Blade, an online site that tracks YouTube analytics. The sole purpose of the site is to provide data to better understand progress and growth on YouTube. Social Blade is useful in understanding what the most popular YouTube channels are and reliable since it is considered a YouTube “certified company”.

Based on the data the music genre leads followed closely behind by the comedy genre. A little further below that are channels focused on video games. There were a lot of other genres that didn’t quite amount to have substantial meaning in the larger context of the data. Therefore these data points are grouped together in the category “other”. Based on the data it is clear music, comedy, and video games are the dominant genres on YouTube.

In addition the site VidStatsX also tracks YouTube analytics. VidStatsX doesn’t seem to have any stake on this topic, unlike Social Blade, there is no official certification between VidStatsX and YouTube. VidStatsX is useful in analyzing the top trending YouTube channels in recent timespans. Site data is updated on a daily basis, so the most up to date information is obtained. Since VidStatsX data provides up to the minute results on YouTube rankings

2.2 SCOPE

The question of whether views or subscribers matters more for making money is complex. Multiple sources have slightly varying answers to this question. From a certain perspective it appears views matter more. In other cases subscribers seem to be the best indicator of prosperity. Jinn Kim offers insight on the matter in his article *The institutionalization of YouTube: From user-generated content to professionally generated content* (Kim, 2012). This resource is an online journal database of peer-reviewed articles involving media. The article was published to explore the transformation of YouTube from a user generated content site to a professionally generated content site. Important concepts are explained such as advertising on YouTube as well as the growing interest of large corporations investing in online video. Kim offers a simple explanation of the relationship between advertising, YouTube, and the content provider. The process is when an advertisement is played before the featured video, the revenue from the ad is split between YouTube and the content provider. While this information is useful, it doesn't go into the breakdown of earnings into numerical figures.

A case study by the Harvard Business Review gives additional insight into advertising dollars on YouTube. This resource says advertising rates apply for every 1000 impressions of a video, also known as a CPM. The estimated range of earnings YouTube receives is between \$10 and \$15 for every 1000 ad impressions (Elberse & Gupta, 2010, 5). It is unclear how much of this money the content provider receives. Contrary to the previous two studies, Celebrity Net Worth offers a more in depth analysis of the role views and subscriptions play in earning money. Figure 4 breaks down this information into a flow chart. The article explains for every 1000 impressions for an ad preceding the video (CPM), YouTube receives between \$20 and \$25. In addition once an individual gets a certain number of subscribers for the channel, the owner can apply to be a member of the Partner Program.

3 SYSTEM REQUIREMENTS

Windows

- Microsoft® Windows® 7/8/10 64-bit
- 4 GB RAM minimum, 8 GB RAM recommended
- 2 GB of available disk space minimum,
- 4 GB Recommended 500 MB for IDE + 1.5 GB for Android SDK and emulator system image
- 1280 x 800 minimum screen resolution

Mac

- Mac® OS X® 10.10 Yosemite or higher, up to 10.14 macOS Mojave
- 4 GB RAM minimum, 8 GB RAM recommended
- 2 GB of available disk space minimum,
- 4 GB Recommended 500 MB for IDE + 1.5 GB for Android SDK and emulator system image
- 1280 x 800 minimum screen resolution

Linux

- GNOME or KDE desktop - Tested on gLinux based on Debian.
- 64-bit distribution capable of running 32-bit applications
- GNU C Library glibc 2.19 or later 17
- 4 GB RAM minimum, 8 GB RAM recommended
- 2 GB of available disk space minimum

4 METHODOLOGY

4.1 DATA COLLECTION

Once the project undertaking is completely comprehended, our subsequent stage is to gather the information that is required for future model building. The information accumulation was an issue as data was not situated at a single source. The data was kept in unique information design. To achieve the end goal, it requires a clear understanding of the correct location of the data. As we can see the currency exchange bureau Statistics gives detailed information on every single currency exchange rates, which incorporates their current high, current low, current price, current exchange rates at the particular time interval usually months. We consolidated a portion of the information properties with currency exchange bureaus to shape a join data set.

Since the datasets for every year are very massive, we decrease our concentration to one-year, i.e., 2018, which as of now contains 365 days records for the most significant currencies in the world. In this venture, we have taken 2022 as our preparation set and 2023 as our test set. There are also various factors affecting the exchange rates of the currencies such as inflation, interest rates, government debt/public, potential stability, economic recession, terms of trade, current account deficit, confidence and speculation, government intervention, the stock markets and etc., the datasets are taken based on the above factors that affect the exchange rates.

Regardless of the field of study or preference for defining data (quantitative, qualitative), accurate data collection is essential to maintaining the integrity of research. Both the selection of appropriate data collection instruments (existing, modified, or newly developed) and clearly delineated instructions for their correct use reduce the likelihood of errors occurring. inability to answer research questions accurately, inability to repeat and validate the study, distorted findings resulting in wasted resources, misleading other researchers to pursue fruitless avenues of investigation, compromising decisions for public policy causing harm to human participants and animal subjects

4.2 DATA MINING

In this section, there is an overview of the process of data mining and data modeling, collecting the data, through the data preparation and finally the data modeling. Data cleaning and formatting can be considered as the most critical part of the whole project according to several data scientists. Figure shows how the process of data mining works to extract knowledge using algorithms.

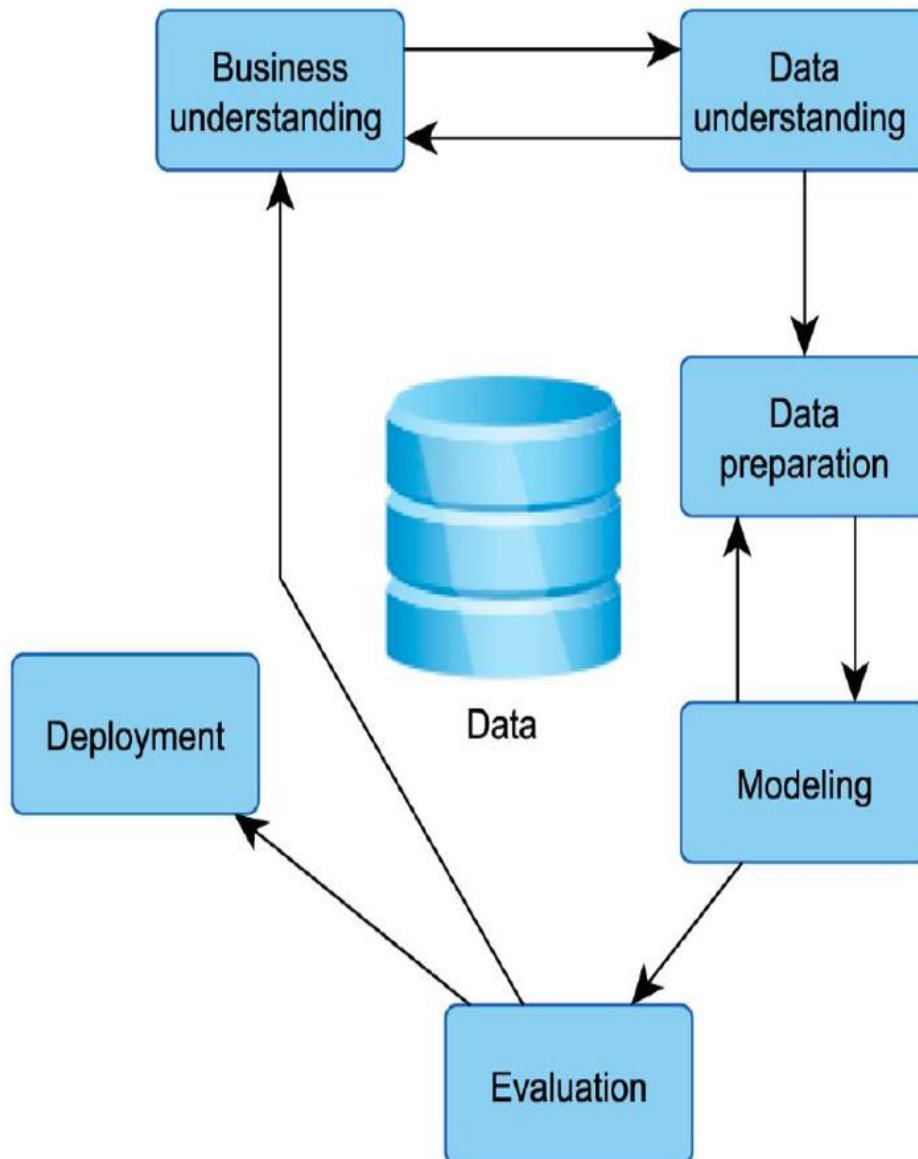


FIGURE 4.1 : DATA MINING

4.3 DATA ANALYSIS

The process of data analysis, or alternately, data analysis steps, involves gathering all the information, processing it, exploring the data, and using it to find patterns and other insights. The process of data analysis consists of: Data Requirement Gathering: Ask yourself why you're doing this analysis, what type of data you want to use, and what data you plan to analyze.

Data Collection: Guided by your identified requirements, it's time to collect the data from your sources. Sources include case studies, surveys, interviews, questionnaires, direct observation, and focus groups. Make sure to organize the collected data for analysis. Data Cleaning: Not all the data you collect will be useful, so it's time to clean it up. This process is where you remove white spaces, duplicate records, and basic errors. Data cleaning is mandatory before sending the information on for analysis.

Data Analysis: Here is where you use data analysis software and other tools to help you interpret and understand the data and arrive at conclusions. Data analysis tools include Excel, Python, R, Looker, Rapid Miner, Charito, Meta base, Rehash, and Microsoft Power BI. Data Interpretation: Now that you have your results, you need to interpret them and come up with the best courses of action based on your findings.

Data Visualization: Data visualization is a fancy way of saying, "graphically show your information in a way that people can read and understand it." You can use charts, graphs, maps, bullet points, or a host of other methods. Visualization helps you derive valuable insights by helping you compare datasets and observe relationships.

If you want to make informed decisions, you need data, but there's more to it. The data in question must be accurate. Data analysis helps businesses acquire relevant, accurate information, suitable for developing future marketing strategies, business plans, and realigning the company's vision or mission. A simple example of data analysis can be seen whenever we make a decision in our daily lives by evaluating what has happened in the past or what will happen if we make that decision. Basically, this is the process of analyzing the past or future and making a decision based on that analysis.

4.4 TYPES OF DATA ANALYSIS

The four types of data analysis are:

- Descriptive Analysis
- Diagnostic Analysis
- Predictive Analysis
- Prescriptive Analysis

4.4.1 DESCRIPTIVE ANALYSIS

Descriptive analytics answer the question, “What happened?”. This type of analytics is by far the most commonly used by customers, providing reporting and analysis centered on past events. It helps companies understand things such as:

- How much did we sell as a company?
- What was our overall productivity?
- How many customers churned in the last quarter?

Descriptive analytics is used to understand the overall performance at an aggregate level and is by far the easiest place for a company to start as data tends to be readily available to build reports and applications. It’s extremely important to build core competencies first in descriptive analytics before attempting to advance upward in the data analytics maturity model. Core competencies include things such as data modeling fundamentals, communicating data with right visualization and basic dashboard skills.

It’s likely you’ve adopted some form of descriptive analytics internally, whether that be static P&L statements, PDF reports, or reporting within an analytics tool. For a true descriptive analytics program to be implemented, the concepts of repeatability and automation of tasks must be top of mind. Repeatability in that a data process is standardized and can be regularly applied with minimal effort (think a weekly sales report), and automation in that complex tasks (VLOOKUPS, merging of excel spreadsheets, etc.) are automated—requiring little to no manual intervention.

4.4.2 DIAGNOSTIC ANALYSIS

Diagnostic analytics, just like descriptive analytics, uses historical data to answer a question. But instead of focusing on “what”, diagnostic analytics addresses the critical question of why an occurrence or anomaly occurred within your data. Diagnostic analytics also happens to be the most overlooked and skipped step within the analytics maturity model. Anecdotally, I see most customers attempting to go from “what happened” to “what will happen” without ever taking the time to address the “why did it happen” step. This type of analytics helps companies answer questions such as:

- Why did our company sales decrease in the previous quarter?
- Why are we seeing an increase in customer churn?
- Why is a specific basket of products vastly outperforming their prior year sales figures?

Diagnostic analytics tends to be more accessible and fit a wider range of use cases than machine learning/predictive analytics. You might even find that it solves some business problems you earmarked for predictive analytics use cases. Being at the diagnostic analytics phase likely means you’ve adopted a modern analytics tool. Most modern analytics tools contain a variety of search-based, or lightweight artificial intelligence capabilities.

These features allow for detailed insights a layer deeper (for example: the Key Drivers visualization in Power BI, or Qlik’s search-based insight functionality). To be clear, these are an effective lightweight means to address diagnostic analytics use cases but are not a means to a full-scale implementation. Software vendors like Sisu have built their core business around addressing diagnostic analytics use cases (what they call “augmented analytics”) and are a great bet. Diagnostic analytics is an important step in the maturity model that unfortunately tends to get skipped or obscured. If you cannot infer why your sales decreased 20% in 2020, then jumping to predictive analytics and trying to answer “what will happen to sales in 2021” is a stretch in advancing upward in the analytics maturity model.

4.4.3 PREDICTIVE ANALYSIS

Predictive analytics is a form of advanced analytics that determines what is likely to happen based on historical data using machine learning. Historical data that comprises the bulk of descriptive and diagnostic analytics is used as the basis of building predictive analytics models. Predictive analytics helps companies address use cases such as:

- Predicting maintenance issues and part breakdown in machines.
- Determining credit risk and identifying potential fraud.
- Predict and avoid customer churn by identifying signs of customer dissatisfaction.

At the outset of any predictive analytics build, three core elements need to be established:

- Identify a problem to solve,
- Define what is you want to predict, and
- State what you will achieve by doing so.

To start you should collect existing data, organize data in a useful way to allow for data modeling, cleanse your data and review overall quality, and finally determine your modeling objective.

While modeling takes up the spotlight in predictive analytics, data prep is a crucial step that needs to happen first. This is why organizations with a rock-solid foundation in descriptive and diagnostic analytics are better equipped to handle predictive analytics. Simply put, the time and effort to prep, transform, and ensure data quality for retrospective reporting has already taken place. The groundwork should be relatively well laid to quickly identify and leverage data for the modeling phase. I always encourage customers with well-defined KPIs and business logic in a specific business reporting area (think sales reporting for example) to use that as the first predictive analytics use case. The goal is to derive value quickly, and there is no better place to start than an area where you know data is well defined and of high quality.

4.4.4 PRESCRIPTIVE ANALYSIS

Prescriptive analytics is the fourth and final pillar of modern analytics. Prescriptive analytics pertains to true guided analytics where your analytics is prescribing or guiding you toward a specific action to take. It is effectively the merging of descriptive, diagnostic, and predictive analytics to drive decision making. Existing scenarios or conditions (think your current fleet of freight trains) and the ramifications of a decision or occurrence (parts breakdown on the freight trains) are applied to create a guided decision or action for the user to take (proactively buy more parts for preventative maintenance).

Prescriptive analytics requires strong competencies in descriptive, diagnostic, and predictive analytics which is why it tends to be found in highly specialized industries (oil and gas, clinical healthcare, finance, and insurance to name a few) where use cases are well defined. Prescriptive analytics help to address use cases such as:

- Automatic adjustment of product pricing based on anticipated customer demand and external factors.
- Flagging select employees for additional training based on incident reports in the field.

Prescriptive analytics' primary aim is to take the educated guess or assessment out of data analytics and streamline the decision-making process. Prescriptive analytics is commonly considered the merging of descriptive, diagnostic, and predictive analytics. Getting started isn't so much a step-by-step list but rather the time and effort up front to build your competencies within the analytics maturity curve.

Simply put, there is no starting point in prescriptive analytics without the requisite first three pillars of modern analytics being established first. If you're ready for prescriptive analytics, then quantifying your call to action and the underlying criteria will be the first requirement. For example: if the use case is to call corrective action for an employee (i.e. – additional training based on poor performance) then the factors that necessitate this action must be firmly established and the action itself must be clearly defined. Knowing how each kind of analytics helps you better understand your data and how to use it.

4.5 POWER BI

Microsoft Power BI is a business intelligence (BI) platform that provides nontechnical business users with tools for aggregating, analyzing, visualizing and sharing data. Power BI's user interface is fairly intuitive for users familiar with Excel, and its deep integration with other Microsoft products makes it a versatile self-service tool that requires little upfront training.

Users can download an application for Windows 10, called Power BI Desktop, and native mobile apps for Windows, Android and iOS devices. There is also Power BI Report Server for companies that must maintain their data and reports on premises. That version of Power BI requires a special version of the desktop app -- aptly called Power BI Desktop for Power BI Report Server.

4.5.1 USES OF POWER BI

Microsoft Power BI is used to find insights within an organization's data. Power BI can help connect disparate data sets, transform and clean the data into a data model and create charts or graphs to provide visuals of the data. All of this can be shared with other Power BI users within the organization.

The data models created from Power BI can be used in several ways for organizations, including the following:

- telling stories through charts and data visualizations;
- examining "what if" scenarios within the data; and
- creating reports that can answer questions in real time and help with forecasting to make sure departments meet business metrics.

Power BI can also provide executive dashboards for administrators or managers, giving management more insight into how departments are doing.

4.5.2 FEATURES OF POWER BI

Microsoft has added a number of data analytics features to Power BI since its inception and continues to do so. Some of the most important features are the following:

Artificial intelligence. Users can access image recognition and text analytics in Power BI, create machine learning models using automated ML capabilities and integrate with Azure Machine Learning.

Hybrid deployment support. This feature provides built-in connectors that allow Power BI tools to connect with a number of different data sources from Microsoft, Salesforce and other vendors.

Quick Insights. This feature allows users to create subsets of data and automatically apply analytics to that information.

Common data model support. Power BI's support for the common data model allows the use of a standardized and extensible collection of data schemas (entities, attributes and relationships).

Cortana integration. This feature, which is especially popular on mobile devices, allows users to verbally query data using natural language and access results using Cortana, Microsoft's digital assistant.

Customization. This feature allows developers to change the appearance of default visualization and reporting tools and import new tools into the platform.

APIs for integration. This feature provides developers with sample code and application program interfaces (APIs) for embedding the Power BI dashboard in other software products.

Self-service data prep. Using Power Query, business analysts can ingest, transform, integrate and enrich big data into the Power BI web service. Ingested data can be shared across multiple Power BI models, reports and dashboards.

Modeling view. This allows users to divide complex data models by subject area into separate diagrams, multiselect objects and set common properties, view and modify properties in the properties pane, and set display folders for simpler consumption of complex data models.

4.6 ARCHITECTURE

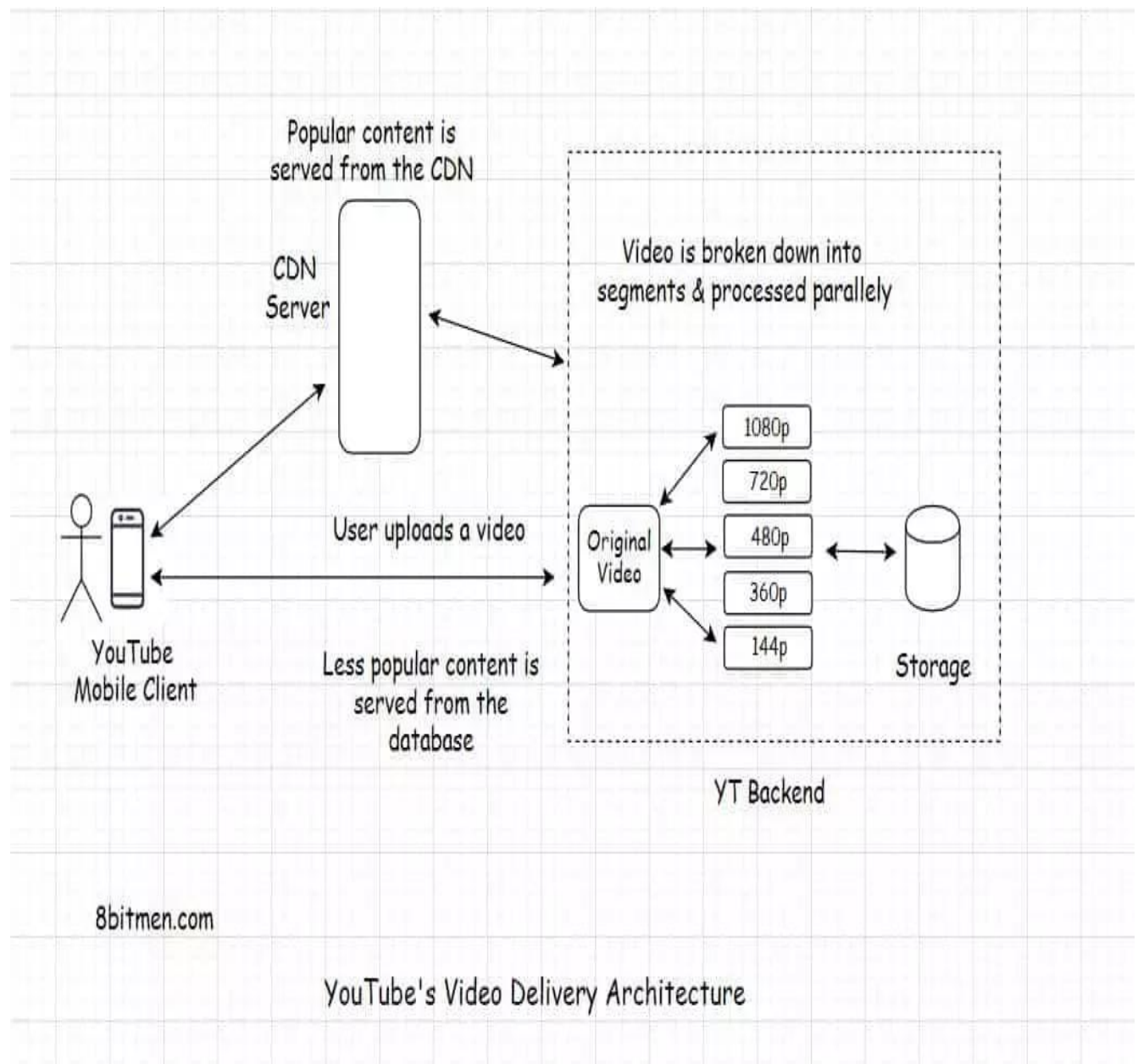


FIGURE 4.2 : ARCHITECTURE

4.7 PROJECT IMPLEMENTATION

The project is implemented in Microsoft excel and Microsoft power Bi. The Microsoft excel is used to collect the data and form the data's into the table. Microsoft power bi is used to analyze and transform the raw data into an analyzable format.

We need to collect the datasets for youtube channel stats. We need to collect the stats such as views count, subscribers count, likes etc.,

The dataset includes;

- date(January 2022-december 2022)
- number of views every day

Now enter the dataset in Microsoft excel sheet as per the date and month. Next one is the main objective of the project that is outcome of views analysis.

The value of views count depends on various factor namely:

- subscriber increased
- shared videos
- views
- watch time
- trending
- number of languages
- number of likes per video

Now collect the dataset of these values for all the currencies. Import to the excel sheet. These data's are taken for the currencies such as us dollar, European currency, Indian rupee, pound, Singapore dollar, Sri Lankan rupee, Australian dollar, Chinese yuan etc.,

Then import the excel sheet into the Microsoft power bi.

The following diagram shows the datasets taken

	January	February	March	April	May	June	July	August	September	October	November	December
subscribers count	35000	56000	77000	93000	115000	136000	143000	178000	212000	226000	243000	264000
shared videos	350	560	770	930	1150	1360	1430	1780	2120	2260	2430	2640
views	139949	206132	255138	308229	370907	519883	680020	620126	933997	975442	1678471	2595514
watch time	22987	41533	53784	67057	82726	119970	160005	600126	176799	234860	400617	628878
trending video	4	5	7	3	6	5	2	4	8	6	5	6
number of videos	18	20	22	24	26	28	30	15	13	10	8	6
language	4	4	4	4	4	4	4	4	4	4	4	4
likes	13752	20245	31204	37245	51412	68297	62987	93369	96753	97157	167851	254562

FIGURE 4.3 : DATASET 1

	January	February	March	April	May	June	July	August	September	October	November	December
01-01-2020	35000	56000	77000	93000	115000	136000	143000	178000	212000	226000	243000	264000
02-01-2020	350	560	770	930	1150	1360	1430	1780	2120	2260	2430	2640
03-01-2020	139949	206132	255138	308229	370907	519883	680020	620126	933997	975442	1678471	2595514
04-01-2020	22987	41533	53784	67057	82726	119970	160005	600126	176799	234860	400617	628878
05-01-2020	4	5	7	3	6	5	2	4	8	6	5	6
06-01-2020	18	20	22	24	26	28	30	15	13	10	8	6
07-01-2020	4	4	4	4	4	4	4	4	4	4	4	4
08-01-2020	13752	20245	31204	37245	51412	68297	62987	93369	96753	97157	167851	254562
09-01-2020												
10-01-2020												
11-01-2020												
12-01-2020												
13-01-2020												
14-01-2020												
15-01-2020												
16-01-2020												
17-01-2020												
18-01-2020												
19-01-2020												
20-01-2020												
21-01-2020												
22-01-2020												
23-01-2020												
24-01-2020												

FIGURE 4.4 : DATASET 2

5 RESULTS AND DISCUSSION

As a first exploratory examination, we consider the watched likelihood of deferral in minutes on the whole dataset. The best path is through a histogram and graphs.

5.1 PREDICTION

Although its primary purpose is data visualization, Power BI contains surprisingly powerful predictive forecasting options, specifically in the area of time series forecasting. Time series forecasting is defined as taking a model that is fit to historical data and using it to predict future observations. That is, predicting what a certain numerical value (e.g., sales or revenue) will be at a specific point in the future.

There are many different statistical models and methods that can be applied in this use case, but with Power BI, the end user doesn't have to worry about the behind-the-scenes math going on. Instead, they can use one of Power BI's preexisting visualization tools to quickly build a forecast model and visualize the results. If you'd like to learn more about the math behind the model, we recommend reading this blog post, which goes into detail on the math behind numerous different types of time series models.

- We need to upload the data of views by date(For minimum one year)
- Draw a line chart graph for the above data and enable forecast
- Using this graph we can predict the view count for upcoming days

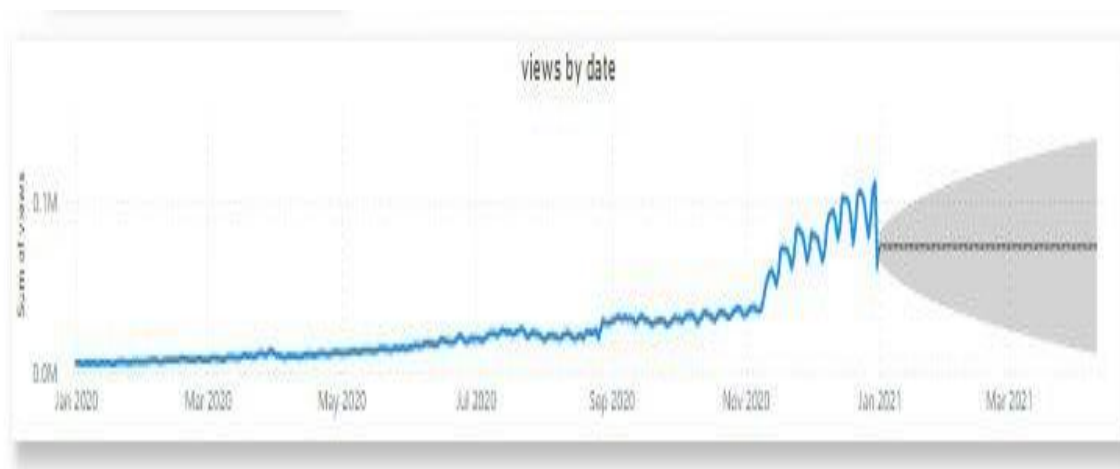


FIGURE 5.1 : PREDICTION TABLE

5.2 Q & A:

A report designer can add a specific type of interactive visual called the Q&A visual. Power BI report end users can ask natural language questions and create a new report visual that represents the answer to the question. The Q&A visual behaves like any other visual in a report. Cross-filter, cross-highlight, bookmarks, and comments are enabled.

Q&A visuals have a question box across the top of the visual. You type your question into the question box and Q&A responds with a visual. The Q&A visual can be used over and over again to ask questions about your data. When you leave the report, the Q&A visual resets to its default.

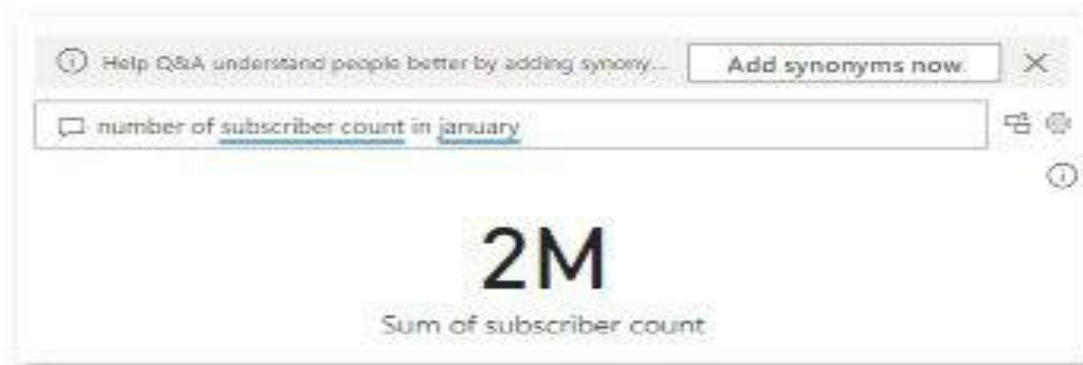


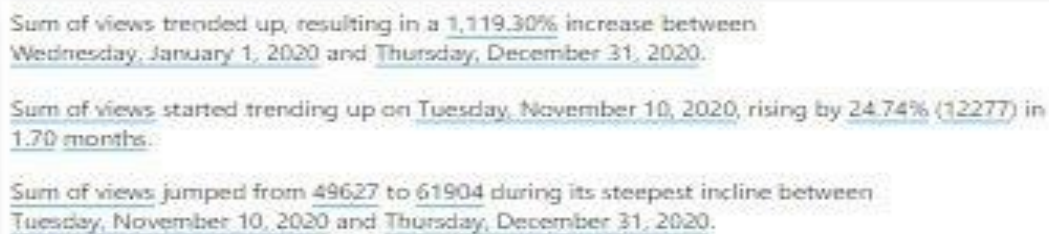
FIGURE 5.2 : Q&A

- Q&A feature lets you explore your data or query your data in your own words
- The smart narrative visualization helps you provide a quick text summary of visuals & reports

The Q&A feature in Power BI lets you explore your data in your own words by using natural language. Q&A is interactive, even fun. Often, one question leads to others as the visualizations reveal interesting paths to pursue. Asking the question is just the beginning.

5.3 SMART NARRATIVE:

The smart narrative visualization helps you provide a quick text summary of visuals and reports. It provides relevant innovative insights that you can customize. Use smart narrative summaries in your reports to address key takeaways, to point out trends, and to edit the language and format for a specific audience



Sum of views trended up, resulting in a 1,119.30% increase between Wednesday, January 1, 2020 and Thursday, December 31, 2020.

Sum of views started trending up on Tuesday, November 10, 2020, rising by 24.74% (12277) in 1.70 months.

Sum of views jumped from 49627 to 61904 during its steepest incline between Tuesday, November 10, 2020 and Thursday, December 31, 2020.

FIGURE 5.3 : SMART NARRATIVE

- Q&A feature lets you explore your data or query your data in your own words
- The smart narrative visualization helps you provide a quick text summary of visuals & reports

You see a narrative that's based on all of the visuals on the page. For example, in the sample file, smart narratives can automatically generate a summary of the report's visuals that address revenue, website visits, and sales. Power BI automatically analyzes trends to show that revenue and visits have both grown. It even calculates growth, which in this case is 72 percent. Sometimes a hidden-summary symbol appears in the smart narrative. It indicates that current data and filters produce no result for the value. A summary is empty when no insights are available.

5.4 VIEWS BY DATE

Yes, YouTube is a social platform, but it's also a search engine. And one of the top strategies for getting more views is YouTube SEO, i.e., optimizing your videos for search.

In other words, when your ideal viewer types in your chosen keywords, you want your video to rank near the top of YouTube's results list. That means you need to know what your audience is looking for—tutorials, inspiration, or entertainment.

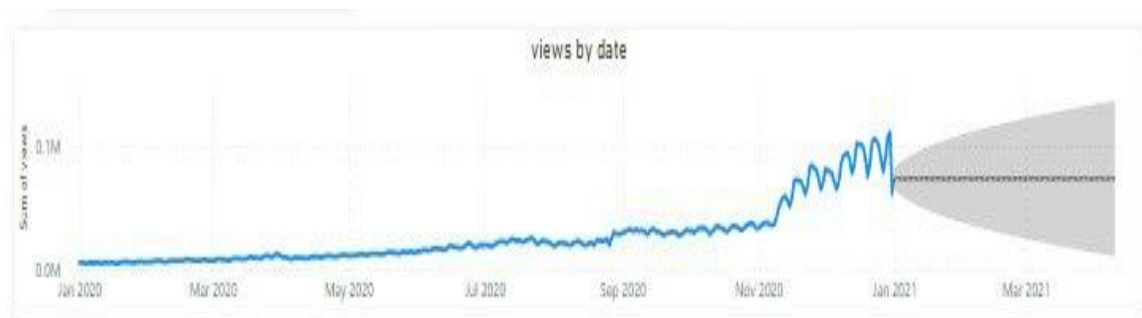


FIGURE 5.4 : views by date

Ranking in search results is the best way to get brand new eyes—not just subscribers and people who are already interested in your channel (although we'll talk more about them later) — on your videos.

But this is easier said than done. So, what can you do to improve your videos' search ranking on YouTube?

Research. You're going to want to use a tool like Google Keyword Planner (note that you'll need to set up a Google Ads account) to do two things:

- Find inspiration for your next video based on what people are already looking for (i.e., take a look at search patterns and see what keywords have a lot of search queries, but few videos, a.k.a. low competition)

5.5 LANGUAGE BY MONTH

This image shows the graph that contains the average languages of videos per month. This is based on the data of videos released in the year of 2022. the average value number of languages for every month is calculated and compared



FIGURE 5.5 : LANGUAGE

5.6 SHARED VIDEOS PER MONTH

Video sharing is when one person shares a video clip with another or many other people. This is usually done over the Internet. There are many websites that offer video sharing. Many people use these sites to share videos they have made with their family and friends. But, sharing videos which are copyrighted, like television programs, without permission from those who own them might be illegal.



FIGURE 5.6 : shared videos

5.7 SUBSCRIBERS COUNT BY MONTH

Your subscriber count reflects how many viewers have subscribed to follow your YouTube channel. You can view your subscriber count in real time and view your growth over time in YouTube Analytics. You'll get an email and find a congratulatory animation within YouTube Studio when you reach a milestone.

The Community Tab is available to channels with at least 500 subscribers. Access: Once you're eligible, you should be able to access the Community Tab from your YouTube dashboard. Features: The Community Tab allows you to post polls, images, videos, and text updates to engage with your subscribers.



FIGURE 5.7 : SUBSCRIBERS COUNT BY MONTH

YouTube is one of the more transparent subscription-based platforms out there. Log in to your account, click on your profile photo, and head to your channel. You'll see your subscriber count underneath your channel's username. You can also see your subscriber count by navigating to YouTube Studio.

5.8 TRENDING VIDEOS BY MONTH

Trending helps viewers see what's happening on YouTube and in the world. Trending aims to surface videos and shorts that a wide range of viewers would find interesting. Some trends are predictable, like a new song from a popular artist or a new movie trailer.

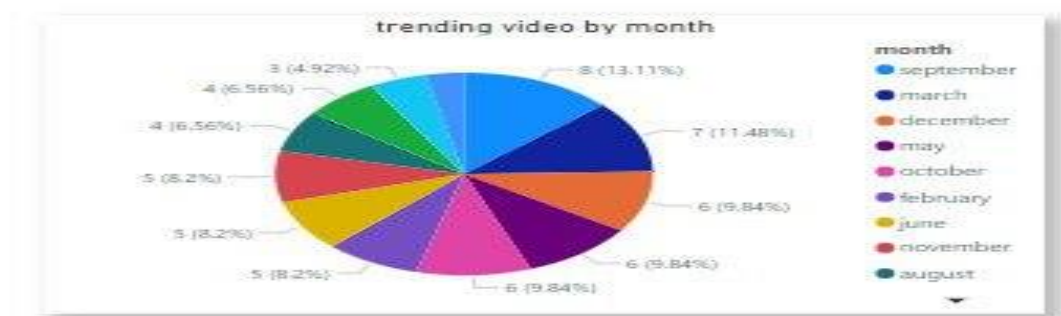


FIGURE 5.8 : TRENDING VIDEOS BY MONTH

5.9 LIKES BY MONTH

Liking a video is a quick way to let a video creator know that you enjoy their work. If you're signed in, liking a video will add it to your "Liked videos" playlist. If you're not the biggest fan of a video, disliking is one way to show your opinion.

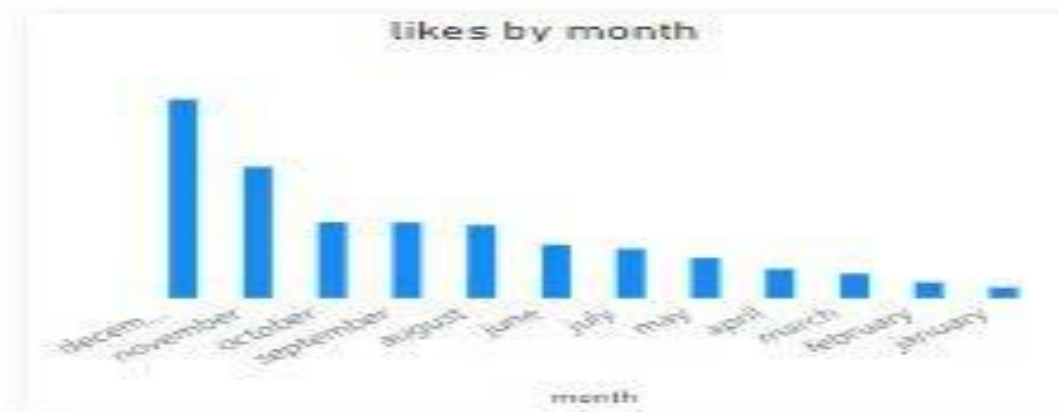


FIGURE 5.9 : LIKES BY MONTH

5.10 VIEWS BY MONTH

The Overview tab shows you a summary of how your channel and videos are performing. The key metrics card shows your views, watch time, subscribers, and estimated revenue (if you're in the YouTube Partner Program).

Note: You may get personalized overview reports that show comparisons to your typical performance. These insights explain why your views may be higher or lower than usual. In this tab, you also get reports for:

- Typical performance: At the channel level, it's a comparison of your channel's typical performance. At the video level, it's a comparison of your video's typical performance.
- Your top content in this period: Your content ranked by views over the last 28 days.
- Realtime: Your performance over the last 48 hours or 60 minutes.
- Stories: Your performance over the last 7 days from your latest Stories.
- Top remixed: Your content that has been used to make Shorts. This report also shows the number of times your content has been remixed and the number of remix views.

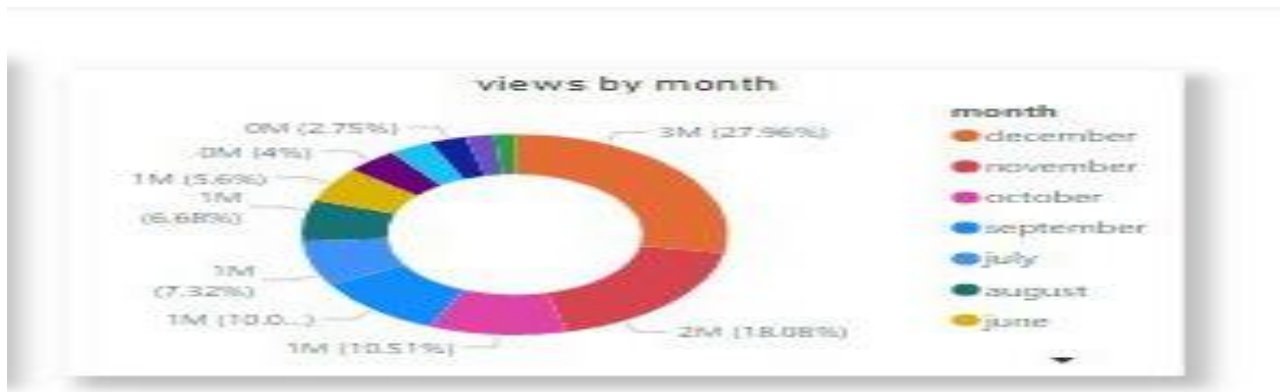


FIGURE 5.10 : VIEWS BY MONTH

The Reach tab gives you a summary of how your audience is discovering your channel. The key metrics card shows your impressions, impressions click-through-rate, views, and unique viewers.

5.11 NUMBER OF VIDEOS BY MONTH

There is no such limit to uploading useful content on YouTube. However, increased frequency of uploading videos help in increasing viewership and and number of followers. Thus it helps in increasing monetized value of the relevant channel and also in return people/followers get useful information.

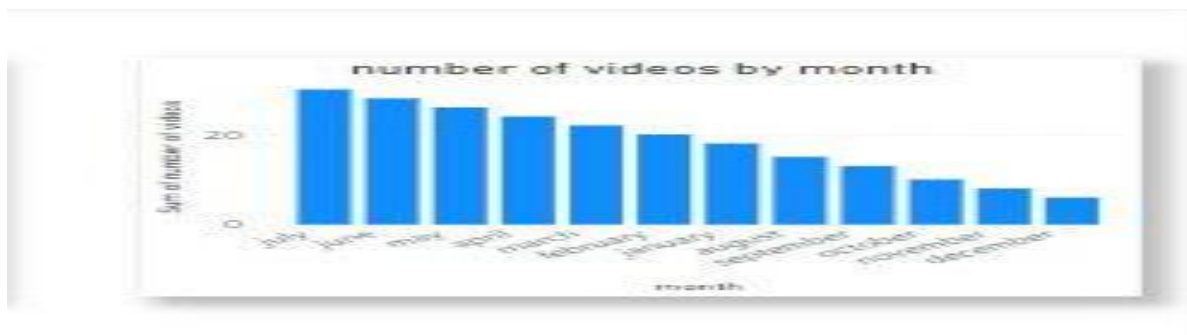


FIGURE 5.11 : NUMBER OF VIDEOS BY MONTH

5.12 WATCHTIME BY MONTH

Watch time is the total accumulated amount of time people have spent watching your video on YouTube. YouTube has confirmed that Watch Time is an important part of their search and discovery algorithm.

It's also important to point out that Watch Time is different than Audience Retention.

Audience Retention measures the percentage of video that people watch.

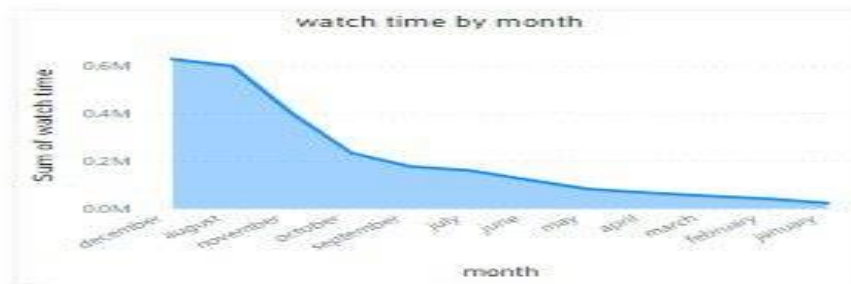


FIGURE 5.12 : WATCHTIME BY MONTH

5.13 SUBSCRIBER COUNT BY REGION

According to the latest statistics, India has the largest YouTube user population, with 467 million active users. With 246 million users, the USA has the next largest user base for the platform. It includes various regions such as india, south africa, north america, austrelia, asia etc.,



FIGURE 5.13 : SUBSCRIBERCOUNT BY REGION

5.14 CATEGORIES OF VIDEOS

There are different kinds of categories for a videos to be uploaded such as animation, stop motion, mixed media, screen recorded video, animated demo, live action, live stream, interactive etc.,

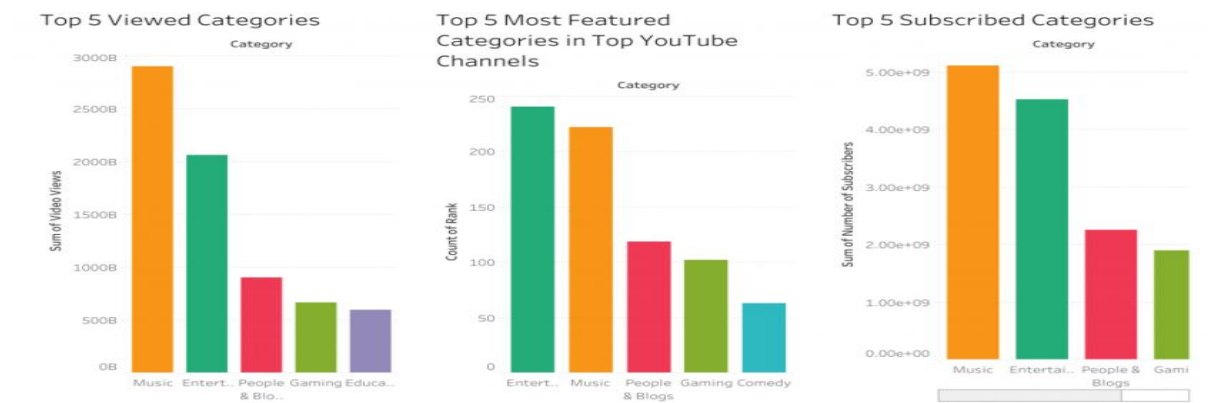


FIGURE 5.14 : videos by categories

5.15 CARDS(VIEWS,SUBSCRIBERS,LIKES)

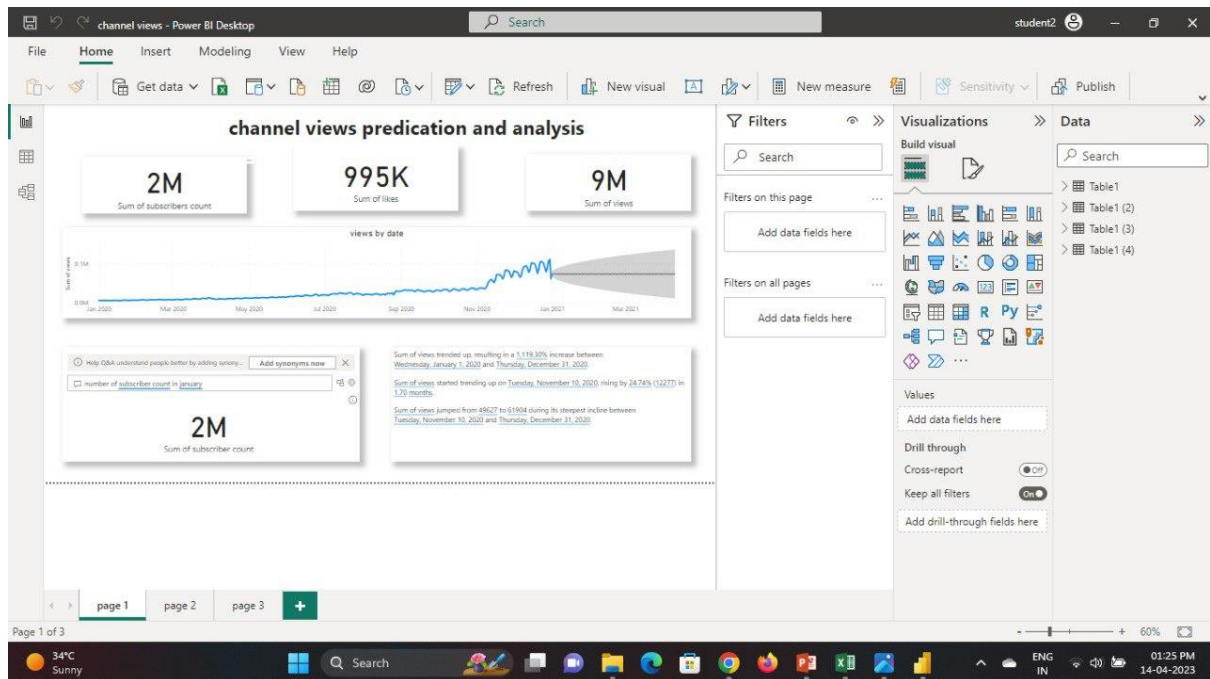


FIGURE 5.15 : CARDS

This image consist of cards. A type of visualization in Power BI called a card may be the best way to view that number. As with almost all of the native Power BI visualizations, cards can be created using the report editor or Q&A.

In the above image there are three types of cards such as likes, subscribers, and views count over the years. The likes represents the total number of views the channel got through various videos, then the subscribers represents the total subscribers gained throughout the year, likes represents the total likes which is used for the growth of the channel.

6.CONCLUSION

Looking at the research, it is clear that YouTube is one of the strongest forces and most influential video-sharing platforms. Data shows based on the top 100 most subscribed channels of all time on YouTube, music is the most subscribed genre. Behind music are the genres comedy and then video games. The top three genres maintain the same order according to the top genres subscribed over the last 30 days. This indicates the current trend will continue to be this way in the near future.

According to the top 40 highest earning YouTube channels, comedy and video games were the most represented genres. One noteworthy finding was the toy genre makes up 10% of the top 40 highest earning channels. This is surprising considering none of the four toy channels exist in the top 100 most subscribed channels. There might be a disconnect however, between certain genres when it comes to accumulating earnings and obtaining subscriptions but there isn't enough evidence to support that claim in this report.

Lastly, there is a question whether views or subscriptions should be valued more in order to earn money. Initially what appears to matter the most for revenue is view count due to the CPM and the divided cost between the content provider and YouTube. Through further analysis it's understood the real way to make money is to acquire as many subscribers as possible. This is because once a producer acquires enough subscribers, a content producer can apply to be a partner and earn money for views that previously were not worth anything. Even though value is based on subscriptions and views, further research demonstrates these statistics may not be reliable ways to measure the value of videos or channels in terms of advertising revenue. In the long run there ought to be foreseeable alternative that is clearer than the current system in place.

Lastly, the argument can be made neither views nor subscriptions are truthful indicators for making money. An article from the tech publishing site Gigaom makes a strong case. The argument of whether views or subscribers on YouTube matters more is a widely debated topic.

7.REFERENCE

Burgess, J., & Green, J. (2009). *YouTube: Online video and participatory culture*. Cambridge: Polity.

Cool, K., Seitz, M., & Mestrits, J. (2010). *YouTube, Google, and the Rise of Internet Video*. Cambridge, MA: Harvard Business Review.

Elberse, A., & Gupta, S. (2010). *YouTube: Time to Charge Users?*. Cambridge, MA: Harvard Business Review.

Kim, J. (2012). The institutionalization of YouTube: From user-generated content to professionally generated content. *Media, Culture & Society*, 34(1), 53–67. Retrieved March 9, 2015, from <http://mcs.sagepub.com/content/34/1/53.full>

Top YouTubers by Subscribed. (n.d.). Retrieved March 4, 2015, from <https://socialblade.com/youtube/top/100/mostsubscribed>

Top 100 Highest Subscriber Gains in the Last 30 Days Ranked by Total Gain. (n.d.). Retrieved March 10, 2015, from <http://vidstatsx.com/top-100-30d-sub-gains>

Warner, B. (2014, March 9). The 25 Highest Earning Youtube Stars. Retrieved March 4, 2015, from <http://www.celebritynetworth.com/articles/celebrity/the-25-highest-earning-youtube-stars/>