**Analysis of Global YouTube Statistics using Power BI**

**Introduction:**

Global YouTube Statistics project, an in-depth exploration of the world's leading YouTube channels. This comprehensive analysis leverages a meticulously curated dataset to uncover key insights into subscriber counts, video views, content genres, and channel performance across various countries.

Our objective is to provide a detailed and interactive view of YouTube’s dynamic landscape, enabling users to gain a deeper understanding of trends and patterns within the platform. By utilizing advanced visualizations and interactive features, this project aims to empower content creators, data enthusiasts, and industry professionals with actionable intelligence and strategic insights to navigate and excel in the evolving world of online video content.

**Scenario 1: YouTube Content Strategy:**

A content creator notices that videos in the Entertainment genre receive significantly higher engagement and views compared to other genres. Based on this insight, the creator decides to focus more on producing entertainment-related content and optimize their video titles, thumbnails, and descriptions to align with viewer preferences. This strategy aims to increase viewer retention, attract a larger audience, and boost overall channel performance.

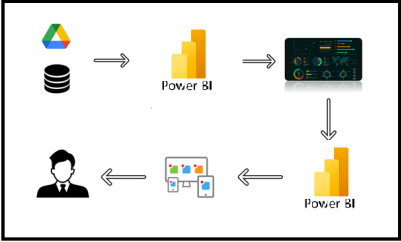
**Scenario 2: Region wise Analysis of YouTube viewership**

Analysis reveals that a YouTube channel’s viewership is concentrated in specific countries, while other regions show lower engagement. The channel’s management decides to tailor content to appeal to audiences in underrepresented regions by incorporating local languages, cultural references, and region-specific topics. This targeted approach aims to increase global reach and diversify the channel’s audience base.

**Scenario 3: Targeted Content Development and Viewer Interaction**

Analysis reveals that viewer retention rates are higher for personalized content tailored to specific interests and preferences. Based on this insight, a YouTube channel implements a strategy to segment its audience and create customized playlists, targeted video recommendations, and interactive content that aligns with viewer preferences. This approach aims to enhance viewer engagement, boost watch time, and foster a stronger connection with the audience.

**Technical Architecture:**

****

**Project Flow**

To accomplish this, we have to complete all the activities listed below,

* Data Collection
  + Collect the dataset,
  + Connect Data with Power BI
* Data Preparation
  + Prepare the Data for Visualization
* Data Visualizations
  + Visualizations
* Dashboard
  + Responsive and Design of Dashboard
* Report
* Report Creation
* Performance Testing
  + Amount of Data Rendered to DB
  + Utilization of Data Filters
  + No. of Calculation fields
  + No. of Visualizations/Graphs
* Project Demonstration & Documentation
  + Record explanation Video for project end to end solution
  + Project Documentation-Step by step project development procedure

**Milestone 1: Data Collection & Extraction from Database**

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, evaluate outcomes and generate insights from the data.

**Activity 1: Downloading the dataset**

Please use the link to download the dataset: [Link](https://www.kaggle.com/datasets/nelgiriyewithana/global-youtube-statistics-2023/data)

**Activity 1.1: Understand the data**

Data contains all the meta information regarding the columns described in the CSV files

**Column Description of the Dataset:**

**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

**Channel type**: Type of the YouTube channel (e.g., individual, brand)

**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 (individual or brand)

**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

**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

**Milestone 2: Data Preparation**

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete.

**Activity 1: Prepare the Data for Visualization**

This process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency. Since the data is already cleaned, we can move to visualization.

1.1: Data Loading

[**Click Here**](https://drive.google.com/file/d/1A8PdBvONeN9oYxyBgRpVV0AI4fP6sleq/view?usp=sharing)

1.2: Data Cleaning

[**Click Here**](https://drive.google.com/file/d/1bGkUxB9tUxLOFFW8boUzlCTW5B7MP7j6/view?usp=sharing)

**Milestone 3: Data Visualization**

Data visualization is the process of creating graphical representations of data to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

**Activity 1: Visualizations**

**Activity 1.1: Ranking**

****

**Insight:** This visualization provides valuable insight into the ranking of YouTube channels based on their subscriber count. It highlights the channels that have successfully garnered the most subscribers globally through their engaging content. This allows us to identify the most influential and popular YouTube channels on the platform.

**Activity 1.2: No: of YouTubers**

****

**Insight:** This visualization provides information on the number of top YouTubers with channels across the globe. It highlights the global distribution of leading content creators, showcasing the widespread popularity and influence of these YouTube channels.

**Activity 1.3: Sum Of Uploads**

****

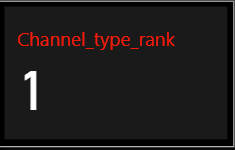
**Insight:** This card visualization provides information on the total number of videos uploaded by YouTubers on their channels. It highlights the content volume produced by top creators, giving an idea of their activity and engagement levels on the platform.

**Activity 1.4: Sum of Video Views**

****

**Insight:** This card visualization provides information on the total view count of YouTube videos, expressed in thousands, millions, or billions. It highlights the reach and popularity of the content created by top YouTubers, showcasing their ability to attract and engage a large audience.

**Activity 1.5: Channel type rank**

****

**Insight:** This card visualization provides information about the rank of a YouTube channel within its specific genre, such as Entertainment, Music, Education, News, and more.

For example, the T-Series YouTube channel holds the "Channel type rank - 1" in the Music genre. This insight highlights the channel's leading position and influence within its content category.

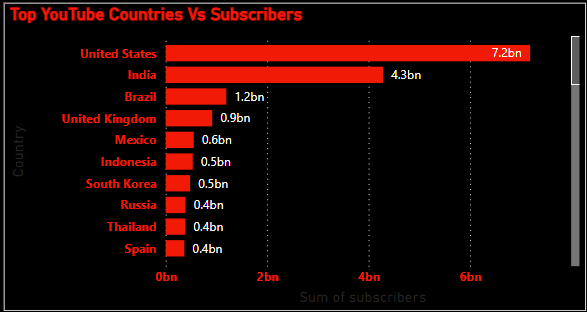
**Activity 1.6: No: of Channels**

****

**Insight:** This card visualization provides information about the number of YouTube channels within each genre worldwide.

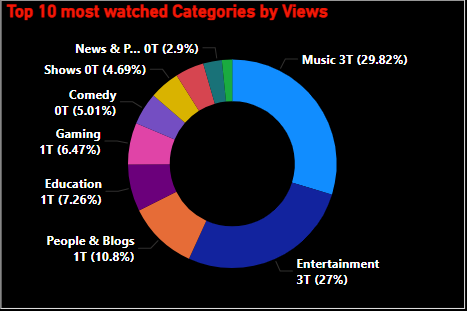
For example, the Entertainment genre has "284" YouTube channels globally. This insight showcases the diversity and distribution of content creators across different categories on the platform.

**Activity 1.7: Top YouTube Countries Vs Subscribers**

****

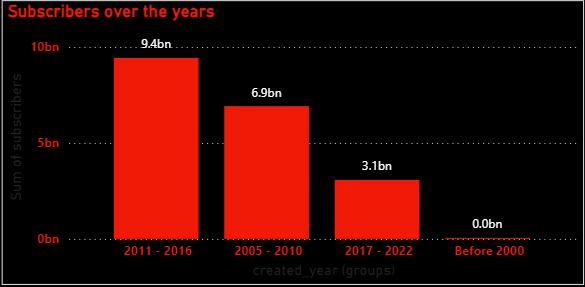
**Insight:** This bar chart visualization provides information about countries with the highest number of subscribers and video views on YouTube. It highlights the nations where YouTube channels are most popular, reflecting the global distribution of viewership and subscriber engagement across different regions.

**Activity 1.8: Top 10 most watched Categories by Views**

****

**Insight:** This donut chart visualization displays the proportion of YouTube channel genres based on the number of viewers. For example, the Music genre holds the largest share, accounting for 29.82% of the total viewership across various genres. This insight highlights the dominance of specific content categories in terms of audience engagement.

**Activity 1.9: Subscribers Over the Years**

****

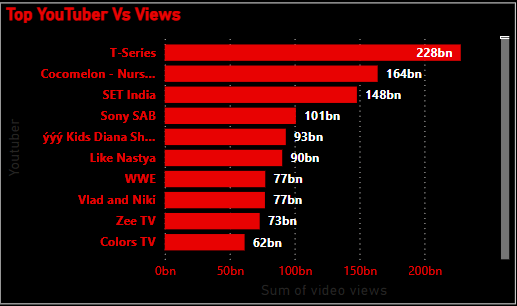
**Insight**: This column chart visualization provides information on the number of subscribers gained by YouTube channels over a period of time. It illustrates the growth trends in subscriber counts, showcasing how different channels have evolved and attracted audiences throughout various time intervals.

**Activity 1.10: Countries**

****

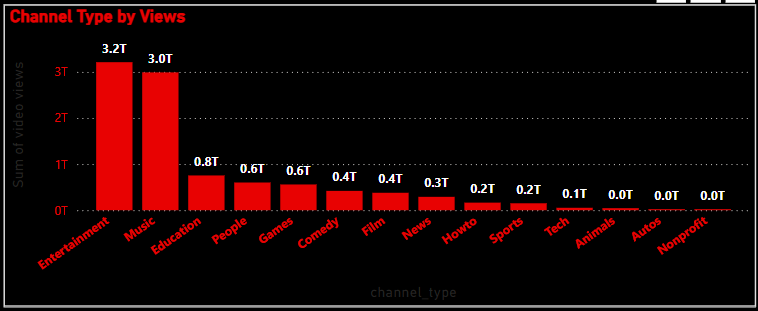
**Insight**: This map chart visualization displays the distribution of YouTube users across different countries worldwide. The size of the bubbles represents the number of users, with larger bubbles indicating countries with higher YouTube user populations. This insight highlights the global reach and user distribution of the platform.

**Activity 1.11: Top YouTuber Vs views**

****

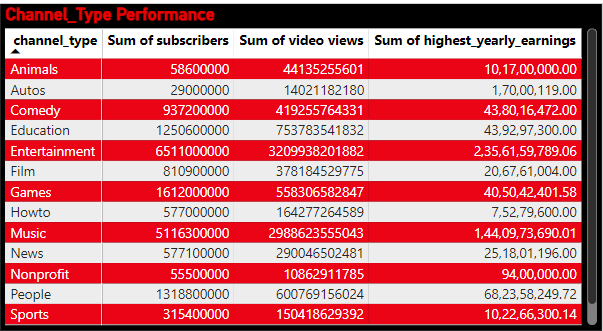
**Insight:** This bar graph provides information about the top YouTubers with the highest total views on their channels worldwide. It highlights the leading content creators based on their overall viewership, showcasing which channels attract the most attention and engagement globally.

**Activity 1.12: Channel Type by Views**

****

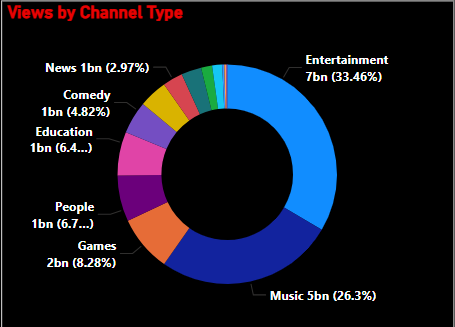
**Insight:** This column chart displays the distribution of views across different YouTube genres, such as Entertainment, Music, People, and others, worldwide. It illustrates how viewership is spread among various content categories, highlighting the popularity and engagement levels of different genres on the platform.

**Activity 1.13: Channel Type Performance**

****

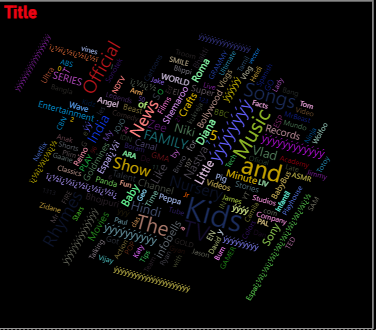
**Insight:** This table provides detailed information on channel performance, including metrics such as subscriber base, total video views, and highest annual earnings. It allows for a comprehensive comparison of top YouTube channels, showcasing their reach, engagement, and revenue potential**.**

**Activity 1.14: Views by Channel Type**

****

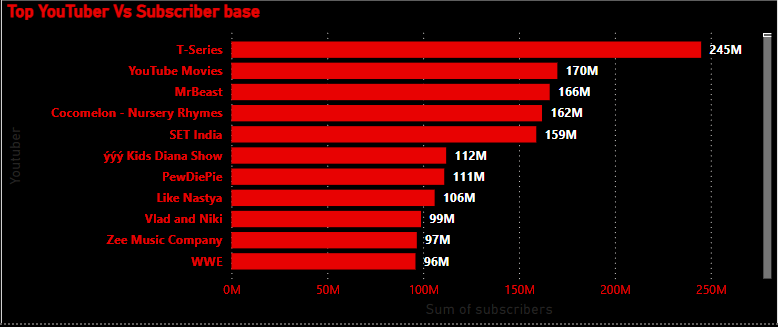
**Insight:** This donut chart illustrates the proportion of different YouTube channel types based on user preference for leisure time. It shows how viewers allocate their time among various channel types, providing insights into which genres are most favored for entertainment and engagement.

**Activity 1.15: Title**



**Insight:** This word cloud chart visually represents the titles of YouTube channels in a dynamic format. The size of each title reflects its frequency or prominence, offering an engaging way to identify the most common or notable channel names on the platform.

**Activity 1.16: Top YouTuber Vs Subscriber Base**



**Insight:** This bar graph highlights YouTube channels with a strong subscriber base and billions of views. It visually represents the leading channels that not only have substantial subscriber counts but also achieve significant viewership, showcasing their widespread popularity and impact.

**Activity 1.17: Slicers**

****

****



**Insight:** Slicers provide the flexibility to filter and select specific countries, years, and YouTubers. This interactive feature allows users to customize their view and analyze data dynamically, with the dashboard updating to reflect the chosen criteria.

**Milestone 4: Dashboard**

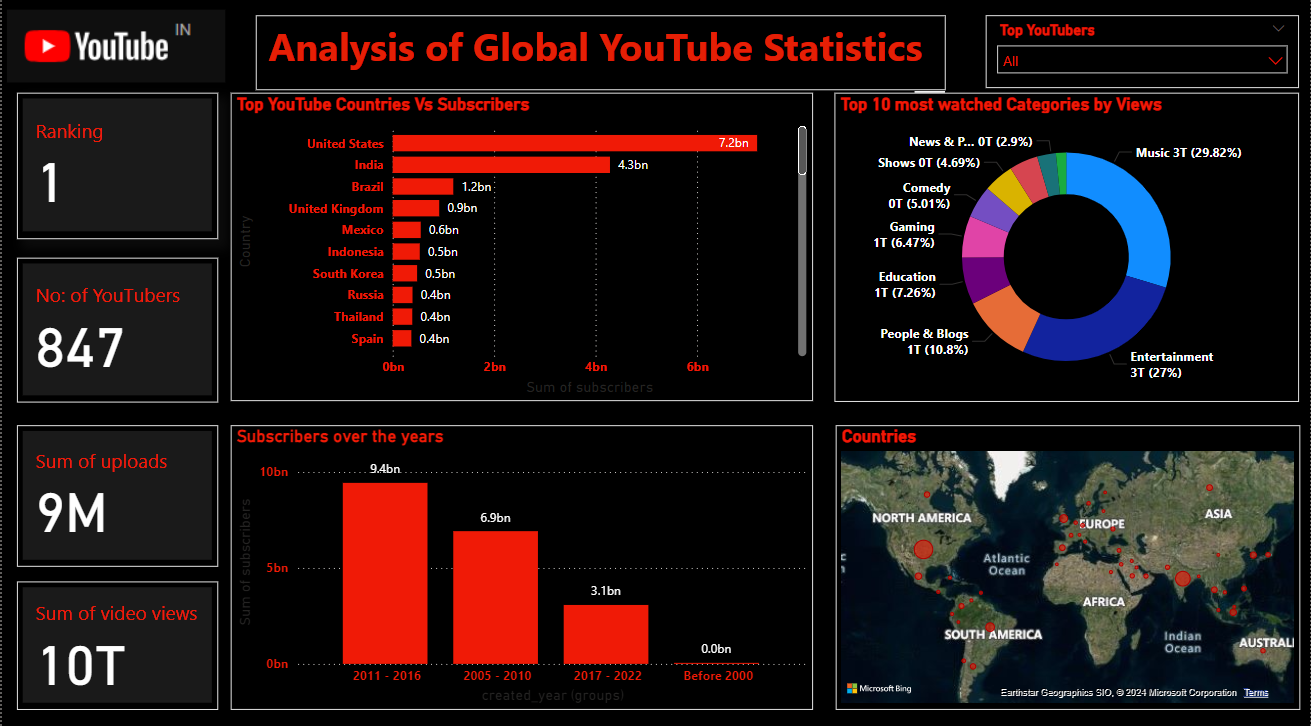
A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

**Activity 1- Responsive and Design of Dashboard**

**Explanation video link:**

[**Click here**](https://drive.google.com/file/d/1DNwbVOVLlTLJ2Y9wjMn9RR0uJWR1rWHf/view?usp=sharing)

**Dashboard:**

****

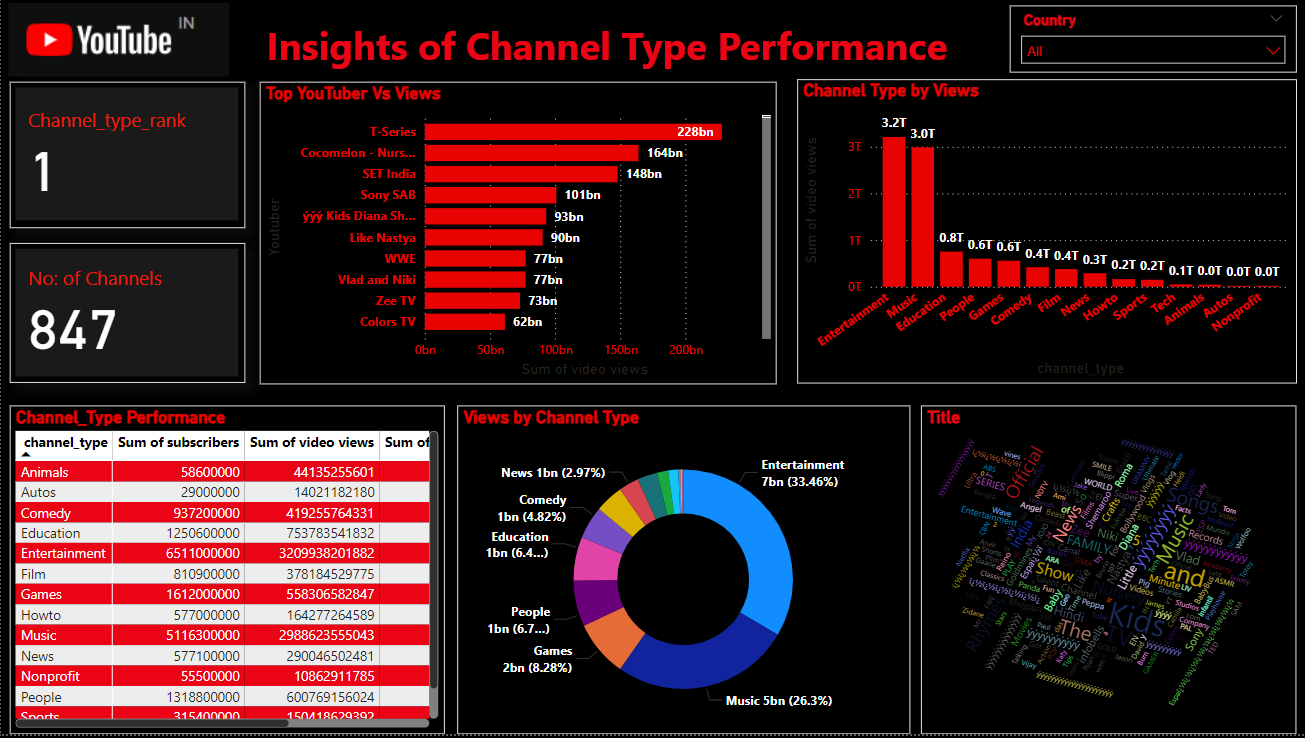
**Insights:** The providedPower BI Dashboard offers a Analysis of Global YouTube Statistics across the world. here is a breakdown of the key Insights:

The United States leads with the highest number of subscribers, significantly surpassing Samoa, which has the lowest subscriber count at 13.1 million.US accounting for 36.93% of the total subscribers, it underscores the country's dominant influence and reach on the YouTube platform.

YouTube subscribers have grown exponentially over the years. Between 2005 and 2010, the subscriber base expanded to 6.9 billion. This growth continued between 2011 and 2016, with subscribers reaching 9.4 billion, highlighting the platform's increasing popularity and user engagement.

The Music genre dominates YouTube with 29.82% of the total viewership, while Nonprofits and Activism account for the lowest viewership at 0.11%. This insight underscores the significant audience engagement disparities across different content categories on the platform.

**Activity 2 - Responsive and Design of Dashboard**

****

**Insights:** The providedPower BI Dashboard offers a insights of Channel type Performance across the world. here is a breakdown of the key Insights:

The T-Series YouTube channel ranks first in the Music genre, holds the highest views at nearly 228 billion worldwide, showcasing its dominance and unparalleled influence in the music industry on the platform.

The Entertainment channel type boasts 3.2 trillion views across 284 YouTube channels, capturing a substantial 33.46% share of overall genres globally. This highlights the genre's widespread appeal and significant influence in the YouTube ecosystem.

The Nonprofit channel type has the lowest earnings, totaling 94 lakhs and accounting for just 0.29% of the overall genre revenue. This underscores the limited financial impact and lower monetization potential of nonprofit content on YouTube.

**Milestone 5: Report**

A report is a comprehensive document that provides a detailed and structured account of data analysis, findings, and insights. It is typically used for in-depth analysis, documentation, and communication of results. Reports are suitable for a diverse audience, including decision-makers, analysts, and stakeholders who need a comprehensive understanding of the data.

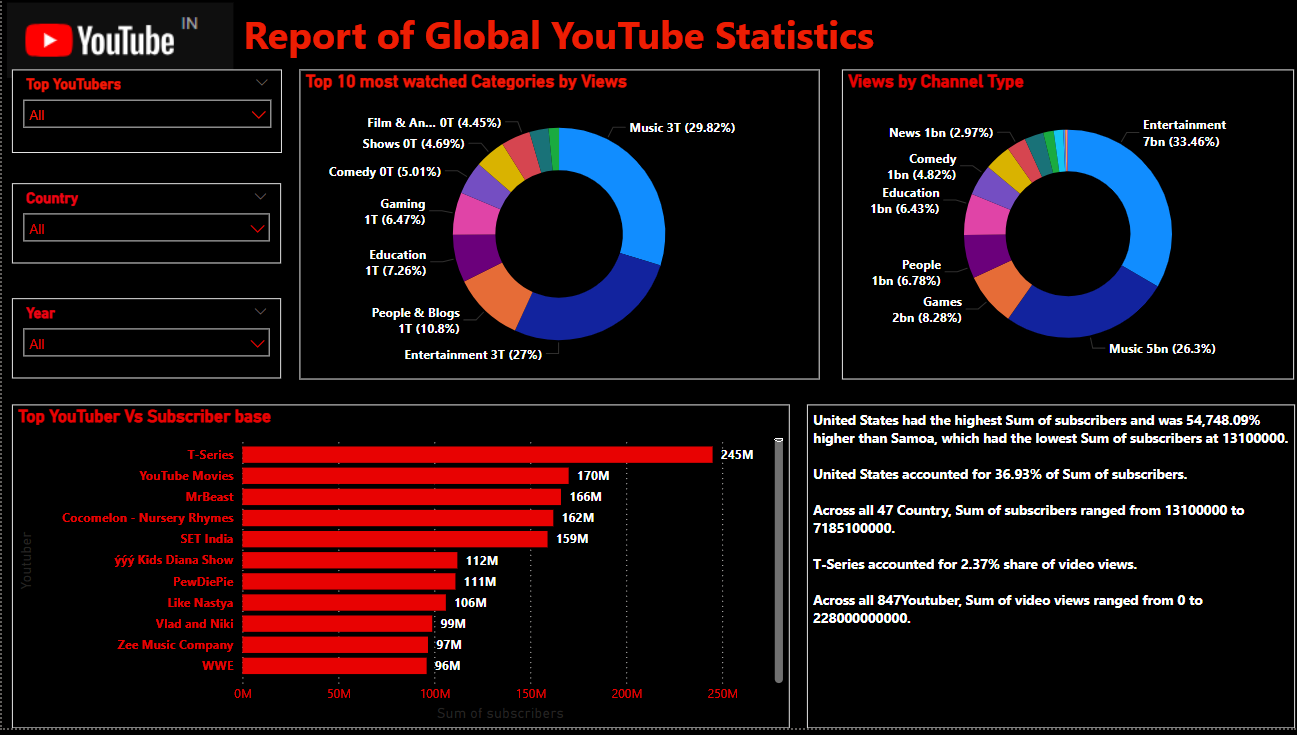
**Activity 1: Design of Report**

Designing a report in Power BI involves connecting to data sources, creating visualizations like charts and graphs, customizing their appearance and interactivity, organizing them logically on the canvas, formatting elements for consistency and clarity, and optionally creating dashboards for a summarized view. Throughout the process, it's essential to consider the audience's needs and ensure the report effectively communicates insights from the data. Finally, iterate based on feedback to continually improve the report's design and usefulness.

**Explanation video link:**

[**Click here**](https://drive.google.com/file/d/1kOUIuNB8r8oMv7Be1ks20jl3_sEvmi2j/view?usp=sharing)

**Report:**

****

**Insights:**

With a strong subscriber base of 245 million, T-Series commands 2.37% of the total video views on YouTube, showcasing its significant presence and influence in the online video landscape.

YouTube subscriber growth has been exponential, with the community expanding from 6.9 billion subscribers between 2005 and 2010 to 9.4 billion from 2011 to 2016, emphasizing the platform's rapid and continuous growth.

The Entertainment genre commands an impressive 3.2 trillion views across 284 channels, representing 33.46% of the global viewership, showcasing its widespread popularity and significant influence on the platform.

**Milestone 6: Performance Testing**

Performance testing is a crucial aspect of software development aimed at evaluating the speed, responsiveness, stability, and scalability of an application under various workload conditions. It involves simulating real-world usage scenarios to assess how the system behaves and performs under stress, peak loads, or normal conditions.

**Activity 1: Amount of Data Loaded**

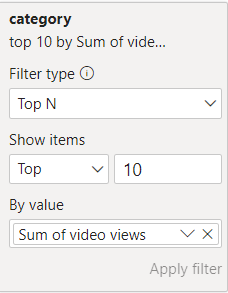
"Amount of Data Loaded" refers to the quantity or volume of data that has been imported, retrieved, or loaded into a system, software application, database, or any other data storage or processing environment. It's a measure of how much data has been successfully processed and made available for analysis, manipulation, or use within the system.



**Activity 2: Utilization of Filters**

"Utilization of Filters" refers to the application or use of filters within a system, software application, or data processing pipeline to selectively extract, manipulate, or analyze data based on specified criteria or conditions.

**Activity 2.1: “Top 10 most watched Category Channel”**

****

**Activity 3: No of Visualizations/ Graphs**

1. Ranking: Card Visual
2. No: of YouTubers Card Visual: Card Visual
3. Sum Of Uploads; Card Visual: Card Visual
4. Sum of Video Views: Card Visual
5. Channel type rank: Card Visual
6. No: of Channels: Card Visual
7. Top YouTube Countries Vs Subscribers: Bar Graph
8. Top 10 most watched Categories by Views: Donut Chart
9. Subscribers Over the Years: Column Chart
10. Countries: Map Visual
11. Top YouTuber Vs views: Bar Graph
12. Channel Type by Views: Column Chart
13. Channel Type Performance: Matrix Table
14. Views by Channel Type: Donut Chart
15. Title: Word Cloud
16. Top YouTuber Vs Subscriber Base: Bar Graph
17. Top YouTubers, Country, Year: Slicers

**Milestone 7: Project Demonstration & Documentation**

Below mentioned deliverables to be submitted along with other deliverables

**Activity 1: - Record explanation Video for the project's end-to-end solution**

**Activity 2: - Project Documentation-Step by step project development procedure**