

# Project Report

## Project Title

### Subscribers Galore : Exploring World's Top Youtube Channels

#### 1) INTRODUCTION:

##### 1.1 Project Overview:

The "Subscribers Galore" data analytics project aims to explore and analyze the world's top YouTube channels to gain insights into their subscribers, content, and growth strategies. With over 2 billion logged-in monthly users, YouTube is one of the most influential social media platforms, making it essential for content creators and businesses to understand the dynamics of successful channels. This project will use data analytics techniques to uncover patterns, trends, and correlations among top YouTube channels, helping creators and marketers make informed decisions for their own channels.

##### Key Objectives:

- 1. Data Collection:** Gather data from various sources, including the YouTube API, web scraping, and publicly available datasets, to create a comprehensive dataset of top YouTube channels, including their video metrics, demographics, and historical data.
- 2. Data Cleaning and Preprocessing:** Clean and preprocess the collected data to ensure its quality, consistency, and usability for analysis. This involves handling missing data, removing outliers, and standardizing data formats.
- 3. Exploratory Data Analysis (EDA):** Conduct EDA to understand the characteristics and distributions of variables. Analyze the correlation between factors such as video frequency, content categories, and subscriber growth.
- 4. Content Analysis:** Analyze the content types, topics, and keywords that attract the most subscribers. Determine which types of videos (e.g., tutorials, vlogs, reviews) perform best for different channel categories.
- 5. Subscriber Growth Patterns:** Identify patterns in subscriber growth, such as seasonal trends, sudden spikes, or gradual increases, and determine the factors influencing growth.
- 6. Demographic Analysis:** Investigate the demographics of subscribers, including age, gender, and location. Understand if the audience profile varies between different channel categories.
- 7. Competitive Analysis:** Compare and contrast different channels within the same category and identify strategies that set successful channels apart from others.
- 8. Predictive Modeling:** Develop predictive models to forecast future subscriber counts based on historical data, allowing creators and businesses to set growth targets.

**9. Recommendations and Insights:** Provide actionable insights and recommendations for content creators, businesses, and marketers to optimize their YouTube channel growth, content strategies, and audience engagement.

**10. Visualization and Reporting:** Create visualizations, dashboards, and a comprehensive report to communicate the findings and insights effectively.

**Expected Outcomes:**

1. A detailed dataset of top YouTube channels.
2. Insights into the factors influencing subscriber growth.
3. Recommendations for optimizing content and channel strategies.
4. Predictive models for subscriber count forecasting.
5. Visualizations and a comprehensive report for stakeholders.

**Benefits:**

- Content creators can optimize their content strategies to attract more subscribers.
- Businesses can make data-driven decisions for influencer marketing and advertising.
- Marketers can better understand audience demographics and preferences.
- Enthusiasts and researchers gain valuable insights into the YouTube ecosystem.

The "Subscribers Galore" project will enable a deeper understanding of YouTube channel success, helping individuals and organizations thrive in the competitive world of online video content creation and promotion.

**1.2 Purpose:**

The purpose of the "Subscribers Galore: Exploring World's Top YouTube Channels" project is to provide valuable insights and actionable information for content creators, businesses, marketers, and researchers within the YouTube ecosystem. The project serves the following key purposes:

**1. Audience Growth Optimization:** To help content creators understand the dynamics behind subscriber growth and provide recommendations on how to optimize their content and channel strategies for increased subscribership.

**2. Business Decision Support:** To assist businesses in making data-driven decisions for influencer marketing and advertising by identifying high-performing YouTube channels and understanding the characteristics that attract audiences.

**3. Audience Understanding:** To offer marketers a deeper understanding of YouTube audience demographics, preferences, and behavior, enabling them to tailor their strategies to engage and target their desired audience more effectively.

**4. Research and Insights:** To provide enthusiasts and researchers with valuable insights into the world of YouTube, contributing to a broader understanding of online content creation, social media, and digital marketing trends.

**5. Predictive Modeling:** To develop models that can forecast future subscriber counts, allowing content creators and businesses to set growth targets and plan their strategies accordingly.

In summary, the project aims to serve the diverse needs of stakeholders in the YouTube community by extracting meaningful information from data, enabling them to improve their channels, make informed business decisions, and gain a deeper understanding of the YouTube platform and its users.

## **2) LITERATURE SURVEY**

### **2.1 Existing Problem:**

- 1. Content Saturation:** High competition among content creators on YouTube makes it challenging to stand out and attract subscribers.
- 2. Algorithmic Changes:** Frequent updates to YouTube's algorithms affect channel visibility and subscriber growth, often without clear documentation.
- 3. Audience Engagement:** Maintaining viewer engagement and understanding effective engagement strategies is difficult in a distracting digital landscape.
- 4. Monetization Challenges:** Creators depend on YouTube for income, and changes in monetization policies and ad revenue can impact their financial stability.
- 5. Data Fragmentation:** Creators and marketers rely on fragmented and inconsistent data sources, hindering their ability to analyze and optimize channel performance effectively.

## 2.2 References:

<https://www.linkedin.com/pulse/exploring-worlds-top-10-most-subscribed-youtube-channels-winston/>  
<https://nm.smartinternz.com/saas-guided-project/1/subscribers-galore-exploring-world-s-top-youtube-channels>  
<https://www.forbesindia.com/article/explainers/most-subscribed-youtube-channels-in-the-world/87475/1>  
[https://en.wikipedia.org/wiki/List\\_of\\_most-subscribed\\_YouTube\\_channels](https://en.wikipedia.org/wiki/List_of_most-subscribed_YouTube_channels)

## 2.3 Problem Statement Definition

The problem statement for the "Subscribers Galore: Exploring World's Top YouTube Channels" project can be defined as follows:

"In the highly competitive landscape of YouTube content creation, content creators, businesses, and marketers face significant challenges related to standing out, growing their subscriber base, and optimizing content strategies. The problem at hand is the need to gain a comprehensive understanding of the factors influencing subscriber growth, audience engagement, and content performance on YouTube. This includes tackling issues such as content saturation, algorithmic changes, demographic shifts, and the dilemma of quality vs. quantity. Additionally, the fragmented nature of available data sources poses a challenge in deriving actionable insights."

The project aims to address these challenges by leveraging data analytics techniques to gather, analyze, and interpret data related to top YouTube channels. By doing so, it seeks to provide insights, recommendations, and predictive models to help content creators and businesses navigate the complexities of the YouTube platform, make informed decisions, and enhance their channel growth and success."

## 3) IDEATION & PROPOSED SOLUTION:

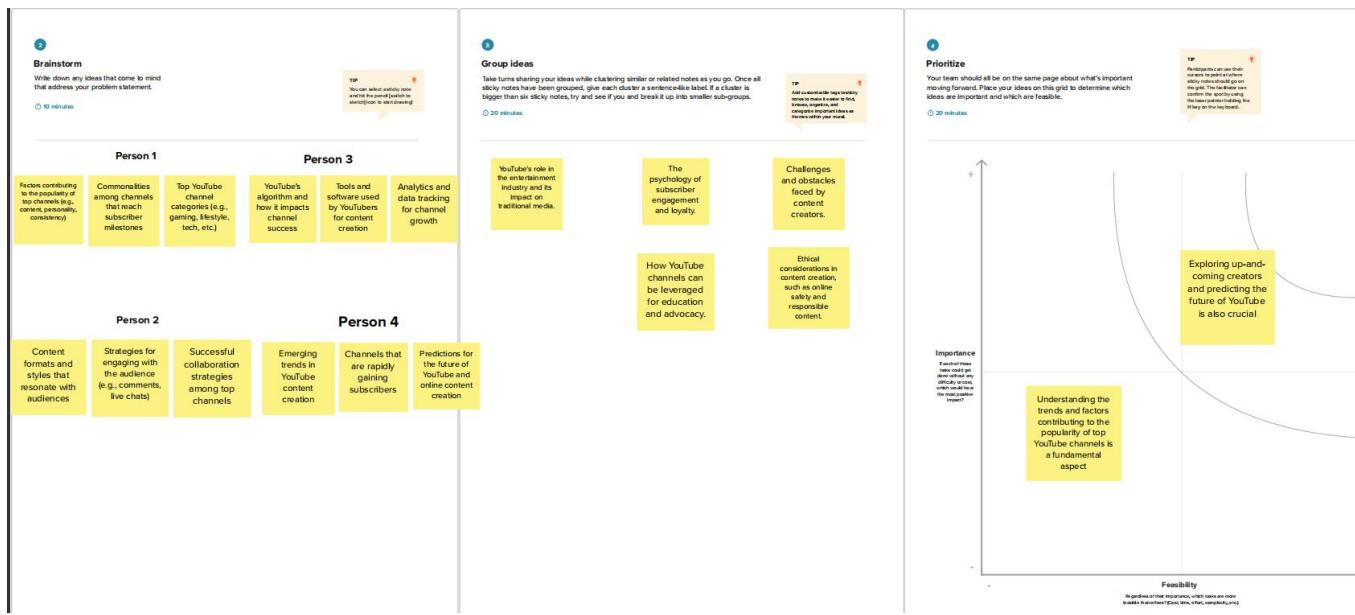
### 3.1 Empathy Map Canvas

#### **Develop shared understanding and empathy**

Summarize the data you have gathered related to the people that are impacted by your work. It will help you generate ideas, prioritize features, or discuss decisions.



## 3.2 Ideation & Brainstroming:



## 4) REQUIREMENT ANALYSIS:

### 4.1 Functional Requirement:

#### 1. Data Collection:

- The system must be able to collect data from various sources, including the YouTube API, web scraping, and publicly available datasets, to create a comprehensive dataset of top YouTube channels.
- Data collection should be automated and scheduled for regular updates.

#### 2. Data Cleaning and Preprocessing:

- The system should clean and preprocess the collected data to ensure its quality, consistency, and usability for analysis. This includes handling missing data, removing outliers, and standardizing data formats.

#### 3. Exploratory Data Analysis (EDA):

- The system must conduct EDA to analyze the characteristics and distributions of variables within the dataset.
- It should generate descriptive statistics and visualizations to provide an overview of the data.

#### 4. Content Analysis:

- The system should analyze the types, topics, and keywords of videos on top YouTube channels to determine what content attracts the most subscribers.
- It should identify and categorize content types (e.g., tutorials, vlogs, reviews) that perform well for different channel categories.

#### 5. Subscriber Growth Analysis:

- The system must identify patterns in subscriber growth, including seasonal trends, sudden spikes, and gradual increases.
- It should determine the factors that influence subscriber growth.

#### 6. Demographic Analysis:

- The system should investigate the demographics of subscribers, including age, gender, and location.
- It should analyze whether the audience profile varies among different channel categories.

#### 7. Competitive Analysis:

- The system must compare and contrast different channels within the same category.

- It should identify strategies that set successful channels apart from others.

## **8. Predictive Modeling:**

- The system should develop predictive models to forecast future subscriber counts based on historical data.
- It should allow users to input parameters for forecasting and provide predictions.

## **9. Recommendations and Insights:**

- The system should provide actionable insights and recommendations for content creators, businesses, and marketers to optimize their YouTube channel growth and content strategies.

## **10. Visualization and Reporting:**

- The system should create visualizations, dashboards, and a comprehensive report to communicate the findings and insights effectively.
- It should support the export of reports and visualizations in various formats (e.g., PDF, Excel).

### **● User Authentication and Access Control:**

- The system should have user authentication and role-based access control to ensure data security and restrict access to authorized personnel.

### **● Scalability and Performance:**

- The system should be scalable to handle a growing dataset and increasing user demand.
- It should be optimized for performance to provide fast and responsive analytics.

### **● Data Backup and Recovery:**

- The system must regularly back up the collected data to prevent data loss in case of failures.
- It should have a data recovery plan in case of unexpected data corruption or loss.

### **● Data Privacy and Compliance:**

- The system should adhere to data privacy regulations and guidelines, ensuring the protection of user and subscriber data.

### **● Documentation:**

- The project documentation should be thorough and include user manuals, code documentation, and data dictionaries to facilitate understanding and usage.

## **4.2 Non Functional Requirements:**

**1. Performance:** The system should provide fast and responsive access to data and analysis results, even when dealing with large datasets. Users should experience minimal delays in retrieving information.

**2. Security:** Data security is paramount. The system must implement encryption, user authentication, and access controls to protect sensitive user and subscriber data from unauthorized access or breaches.

**3. Scalability:** The system should be designed to handle growing data volume and user traffic. It should seamlessly adapt to increased demands without sacrificing performance.

**4. Reliability:** The system should have minimal downtime and provide high availability. Scheduled maintenance or updates should be communicated in advance to minimize disruption.

**5. Data Privacy:** Adherence to data privacy regulations (e.g., GDPR, CCPA) and best practices is crucial. All data collected or stored should respect legal and ethical standards, ensuring user privacy.

**6. Usability:** The user interface should be intuitive and easy to navigate, catering to users with varying levels of technical expertise. A user-friendly design fosters effective usage.

**7. Documentation:** Comprehensive system documentation should be available, including user manuals, code documentation, and data dictionaries. Clear instructions support users and administrators.

**8. Training and Support:** Training resources and user support should be accessible to help users effectively utilize the system. A responsive support mechanism should be in place for addressing user questions and issues.

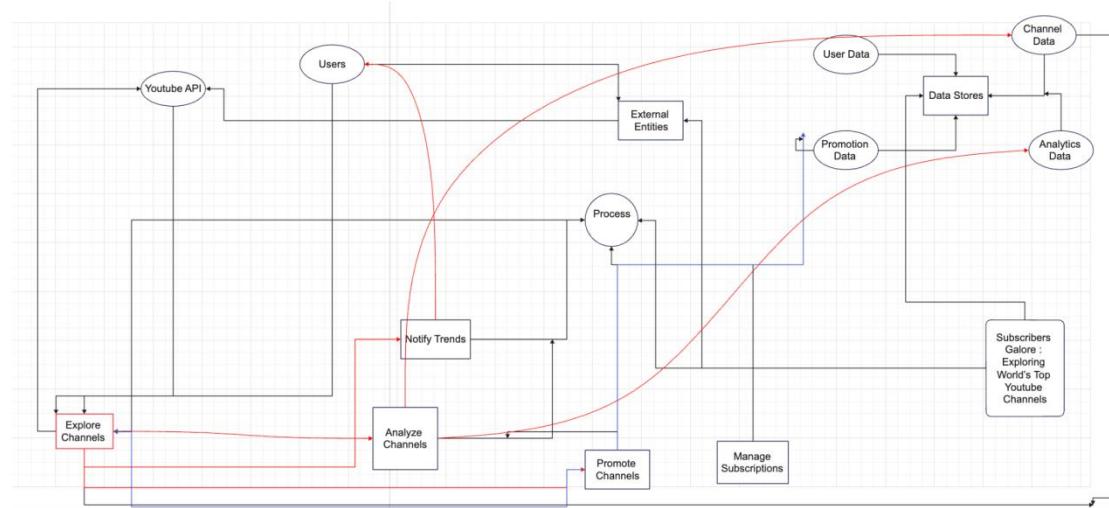
**9. Load Balancing:** Implement load balancing mechanisms to evenly distribute incoming requests, ensuring that the system maintains efficient performance during peak usage periods.

**10. Accessibility:** The system should be designed with accessibility in mind, conforming to WCAG standards. It should be usable by individuals with disabilities to promote inclusivity.

## 5) PROJECT DESIGN:

### 5.1 Data Flow Diagrams & User Stories:

#### Data Flow Diagram:



#### USER STORIES:

| User Type | Functional Requirement (Epic) | User Story Number | User Story / Task | Acceptance criteria | Priority | Release |
|-----------|-------------------------------|-------------------|-------------------|---------------------|----------|---------|
|-----------|-------------------------------|-------------------|-------------------|---------------------|----------|---------|

|                             |                                   |        |   |   |        |             |
|-----------------------------|-----------------------------------|--------|---|---|--------|-------------|
| User A - YouTube Enthusiast | Discover Popular YouTube Channels | USG-01 | <p>As User A, I want to explore and discover popular YouTube channels in various categories, so I can find new content to subscribe to.</p> | <p>The system should provide a homepage with recommended channels based on my viewing history.</p> <p>I can filter channels by categories like gaming, music, vlogs, etc.</p> <p>I can see a brief description and statistics of each channel, including the number of subscribers.</p> | High   | Version 1.0 |
| User B - Trend Follower     | Stay Updated on Trending Channels | USG-02 | <p>As User B, I want to receive notifications about trending YouTube channels, so I can stay up-to-date with the latest content.</p>        | <p>The system should provide a "Trending Now" section on the homepage.</p> <p>It should send me push notifications when a new channel is trending.</p> <p>I can click on a notification to view the channel and subscribe if interested.</p>  | Medium | Version 1.1 |
| User C - Data Enthusiast    | Detailed Channel Analytics        | USG-03 | <p>As User C, I want detailed analytics for YouTube channels, so I can analyze their growth and performance.</p>                            | <p>I can search for any YouTube channel and access detailed analytics.</p>  | High   | Version 1.2 |

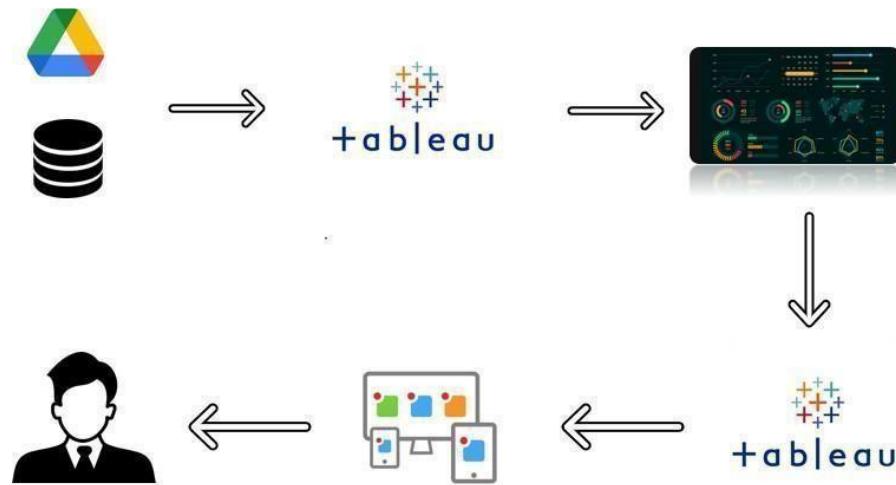
|                          |                    |        |  |   |        |             |
|--------------------------|--------------------|--------|--|---|--------|-------------|
| User D - Content Creator | Promote My Channel | USG-04 | <p>As User D, I want to promote my own YouTube channel within the platform, so I can attract more subscribers.</p> | <p>Analytics should include historical subscriber counts, daily views, and engagement metrics.</p> <p>I can export this data for further analysis.</p> <p>The system should provide visualizations for easy interpretation.</p> | Medium | Version 1.2 |
|--------------------------|--------------------|--------|--|---|--------|-------------|

## 5.2 Solution Architecture:

- **Data Collection:** Utilize the YouTube Data API and web scraping to gather information about top YouTube channels, including channel details, video statistics, and subscriber counts.
- **Data Storage:** Store the collected data in a database (relational or NoSQL) for efficient management and consider using a data warehouse for in-depth analytics.
- **Data Processing:** Implement ETL processes for data cleaning and transformation. Use batch and real-time processing for handling large datasets and immediate updates.
- **User Interface:** Develop a web or mobile application with data visualization for users to explore and interact with the collected data.
- **Search and Recommendations:** Implement search functionality and recommendation systems for user personalization.
- **Security and Compliance:** Prioritize data security, user privacy, and legal compliance. Ensure scalability, monitoring, and continuous improvement for long-term success.

## 6) PROJECT PLANNING & SCHEDULING:

### 6.1 Technical Architecture:



**Technology Stack:** Choose suitable technologies for web development, including programming languages, frameworks, and databases, considering the project's requirements and scalability.

**Data Collection and Storage:** Create a system to collect and store data about top YouTube channels, using a combination of web scraping and API integration, then design an efficient database schema.

**User-Friendly Interface:** Develop an intuitive, responsive, and user-friendly front-end for displaying channel information, along with data analysis and visualization tools.

**Security and Compliance:** Implement robust security measures, user authentication, and comply with legal and ethical considerations related to data usage and user privacy.

**Monitoring and Optimization:** Continuously monitor system performance, apply performance optimizations, and maintain clear documentation to facilitate collaboration and troubleshooting.

**Table-1 : Components & Technologies:**

| S.No | Component             | Description   | Technology   |
|------|-----------------------|---|--|
| 1.   | Web Scraping Engine   | Responsible for collecting data on top YouTube channels, such as subscriber counts, video counts, and more, from the web.                                   | Python (Scrapy, BeautifulSoup), Web scraping libraries                             |
| 2.   | YouTube Data API      | Interface with the YouTube Data API to access channel information and retrieve data such as video details, comments, and channel statistics.                | YouTube Data API, OAuth 2.0  |
| 3.   | Database              | Justify the scalability of architecture (3-tier, Micro-services)  | PostgreSQL (or other relational databases), MongoDB (or other NoSQL databases)     |
| 4.   | User Authentication   | Provide secure user registration and login processes, including password hashing and account management.  | OAuth 2.0, JWT (JSON Web Tokens)   |
| 5.   | Front-End Application | Create a user-friendly web interface for users to search and explore top YouTube channels, including features like search, filters, and data visualization. | HTML, CSS, JavaScript, React (or other front-end libraries), Responsive Web Design |
| 6.   | Data Visualization    | Implement data visualization tools such as charts and graphs to display channel statistics and insights in an understandable format.                        | D3.js, Chart.js, or similar data visualization libraries                           |

|     |  |  |   |
|-----|--|--|---|
| 7.  | Notification System  | Enable users to subscribe to channels and receive notifications for changes in subscriber counts, video uploads, or other relevant updates.                                    | WebSocket, Push Notifications                           |
| 8.  | WebSocket, Push Notifications                                  | Monitor system performance and user behavior, gather insights for optimization, and track application usage.   | Google Analytics, Elasticsearch, Kibana                 |
| 9.  | Security Measures  | Ensure security by implementing measures against common web vulnerabilities (e.g., SQL injection, XSS), user data protection, and compliance with legal and ethical standards. | SSL/TLS, Security Headers, OWASP Top Ten Best Practices |
| 10. | Continuous Integration/ Continuous Deployment (CI/CD) Pipeline | Automate the deployment process for updates and enhancements, ensuring smooth development, testing, and release cycles.  | Jenkins, Travis CI, CircleCI                            |

**Table-2: Application Characteristics:**

| S.N o | Characteristics       | Description   | Technology   |
|-------|-----------------------|---|--|
| 1.    | Scalability           | The ability of the application to handle a growing user base and data load. Scalability ensures the system can expand as needed without compromising performance.               | Load balancers, horizontal scaling, cloud platforms (e.g., AWS, Azure, Google Cloud) |
| 2.    | Real-Time Updates     | The application should provide real-time or near-real-time updates on subscriber counts, new video uploads, and other channel statistics, enhancing user engagement..           | WebSocket, Server-Sent Events (SSE)  |
| 3.    | Mobile Responsiveness | The application should be optimized for mobile devices, providing a seamless user experience on smartphones and tablets.  | Responsive web design, mobile app development (e.g., React Native)                   |
| 4.    | Data Security         | Ensuring the privacy and security of user data and compliance with relevant data protection regulations (e.g., GDPR). Protecting against common web vulnerabilities is crucial. | Encryption, security headers, access controls, penetration testing                   |
| 5     | Multilingual Support  | To cater to a global audience, the application should support multiple languages and provide content in different languages.  | Internationalization (i18n), localization (l10n), language libraries and frameworks  |

## 6.2 SPRINTING:

### Product Backlog, Sprint Schedule, and Estimation (4 Marks)

Use the below template to create product backlog and sprint schedule

| Sprint   | Functional Requirement (Epic)   | User Story Number | User Story / Task                                  | Story Points | Priority |
|----------|---------------------------------|-------------------|--|--------------|----------|
| Sprint-1 | User Authentication             | USN-1             | As a user, I can register for an account           | 5            | High     |
| Sprint-1 |                                 | USN-2             | As a user, I can log in to my account              | 5            | High     |
| Sprint-1 |                                 | USN-3             | As a user, I can view my profile                   | 10           | Medium   |
| +        |                                 |                   |  |              |          |
| Sprint-2 | Search and Filter Functionality | USN-4             | As a user, I can search for YouTube channels       | 10           | High     |
| Sprint-2 |                                 | USN-5             | As a user, I can filter channels by category       | 10           | Medium   |
| Sprint-3 | Channel Details Page            | USN-6             | As a user, I can view detailed channel information | 20           | High     |

### Project Tracker, Velocity & Burndown Chart: (4 Marks)

| Sprint   | Total Story Points | Duration | Sprint Start Date | Sprint End Date (Planned) | Story Points Completed (as on Planned End Date) | Sprint Release Date (Actual) |
|----------|--------------------|----------|-------------------|---------------------------|---|------------------------------|
| Sprint-1 | 20                 | 6 Days   | 23 Oct 2023       | 26 Oct 2023               | 20  | 26 Oct 2023                  |
| Sprint-2 | 20                 | 6 Days   | 26 Oct 2023       | 30 Oct 2023               | 20  | 30 Oct 2023                  |
| Sprint-3 | 20                 | 6 Days   | 01 Nov 2023       | 04 Nov 2023               | 20  | 04 Nov 2023                  |

#### Velocity:

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

## Jira Software Software:

Smartbridge data analytics - X +

jayarammova9.atlassian.net/jira/software/projects/SDA/boards/1/timeline

Jira Software Your work Projects Filters Dashboards Teams Apps Create Search

Smartbridge data anal... Software project You're on the Free plan UPGRADE

PLANNING Timeline Backlog Board + Add view DEVELOPMENT Code Project pages Add shortcut Project settings You're in a team-managed project Learn more

Timeline

Give feedback Share Export ...

Q Status category Epic

|                                       | SEP  | OCT | NOV | DEC |
|---------------------------------------|------|-----|-----|-----|
| Sprints                               |      |     |     |     |
| SDA-1 User Authentication             |      |     |     |     |
| SDA-2 As a user, I can ...            | DONE |     |     |     |
| SDA-3 As a user, I can ...            | DONE |     |     |     |
| SDA-4 As a user, I can ...            | DONE |     |     |     |
| SDA-6 Search and Filter Functionality |      |     |     |     |
| SDA-9 As a user... IN PROGRESS        |      |     |     |     |
| SDA-10 As a user... IN PROGRESS       |      |     |     |     |
| SDA-7 Channel Details Page            |      |     |     |     |
| SDA-12 As a user... IN PROGRESS       |      |     |     |     |
| + Create Epic                         |      |     |     |     |

Today Weeks Months Quarters Insights

Smartbridge data analytics - X +

jayarammova9.atlassian.net/jira/software/projects/SDA/boards/1/backlog?epics=visible

Jira Software Your work Projects Filters Dashboards Teams Apps Create Search

Smartbridge data anal... Software project You're on the Free plan UPGRADE

PLANNING Backlog Timeline Board + Add view DEVELOPMENT Code Project pages Add shortcut Project settings You're in a team-managed project Learn more

Backlog

Import work Insights

Epic

Issues without epic

User Authentication

Search and Filter Functionality

Channel Details Page

+ Create issue

Sprint 2 26 Oct – 30 Oct (5 issues)

The Functional requirement is Search and Filter Functionality

SDA-2 As a user, I can register for an account

SDA-3 As a user, I can log in to my account

SDA-4 As a user, I can view my profile

SDA-9 As a user, I can search for YouTube channels

SDA-10 As a user, I can filter channels by category

0 20 20 Complete sprint ...

Sprint 3 1 Nov – 4 Nov (1 issue)

0 20 0 Complete sprint ...

Backlog (0 issues)

Your backlog is empty.

+ Create issue

Smartbridge data analytics - x +

jayarammova9.atlassian.net/jira/software/projects/SDA/boards/1/backlog?epics=visible

Jira Software Your work Projects Filters Dashboards Teams Apps Create Search ...

Smartbridge data anal... Software project You're on the Free plan UPGRADE

PLANNING Timeline Backlog Board + Add view DEVELOPMENT Code Project pages Add shortcut Project settings

You're in a team-managed project Learn more

Backlog

Epic Issues without epic User Authentication Search and Filter Functionality Channel Details Page Backlog (2 issues) + Create issue + Create epic

Sprint 3 1 Nov – 4 Nov (1 issue) Functional Requirement is Channel Details Page 0 20 0 Complete sprint ...

SDA-12 As a user, I can view detailed channel information CHANNEL D... IN PROGRESS... 20 ...

Import work Insights

This screenshot shows the Jira Software Backlog page for the 'Smartbridge data analytics' project. On the left, a sidebar lists project details, current plan (Timeline), and backlog items like 'Backlog'. The main area displays an epic titled 'User Authentication' with three child issues: 'SDA-12 As a user, I can view detailed channel information', 'SDA-10 As a user, I can filter channels...', and 'SDA-9 As a user, I can search for You...'. A 'Create sprint' button is visible at the bottom right of the epic card.

Smartbridge data analytics - x +

jayarammova9.atlassian.net/jira/software/projects/SDA/boards/1/backlog?epics=visible&selectedIssue=SDA-1

Jira Software Your work Projects Filters Dashboards Teams Apps Create Search ...

Smartbridge data anal... Software project You're on the Free plan UPGRADE

PLANNING Timeline Backlog Board + Add view DEVELOPMENT Code Project pages Add shortcut Project settings

You're in a team-managed project Learn more

Backlog

Epic Issues without epic User Authentication Start date None Due date None View all details + Create epic

Backlog (3 issues)

SDA-1 SDA-9 As a user, I can search for You... DONE SDA-10 As a user, I can filter channels... DONE SDA-12 As a user, I can view detailed ... DONE

Import work Insights

User Authentication

To Do Actions Description Add a description...

Child issues Order by ... 100% Done

SDA-2 As a user, I can re... 5 K DONE SDA-3 As a user, I can lo... 5 M DONE SDA-4 As a user, I can v... 10 R DONE

Link goals

Add a comment... Pro tip: press M to comment

This screenshot shows the Jira Software Backlog page for the same project, but with the epic 'User Authentication' expanded. The expanded view shows the epic title and three child issues: 'SDA-12 As a user, I can view detailed channel information', 'SDA-10 As a user, I can filter channels...', and 'SDA-9 As a user, I can search for You...'. Each child issue has a status of 'DONE'. To the right of the epic, there are sections for 'Actions' (To Do, Description, Child issues), 'Link goals', and a comment input field.

Smartbridge data analytics - x +

jayarammovva9.atlassian.net/jira/software/projects/SDA/boards/1/timeline?selectedIssue=SDA-7

Jira Software Your work Projects Filters Dashboards Teams Apps Create Search Give feedback Share Export ...

Smartbridge data anal... Software project You're on the Free plan UPGRADE

PLANNING Timeline Backlog Board + Add view DEVELOPMENT Code Project pages Add shortcut Project settings You're in a team-managed project Learn more

Timeline

Sprints

SDA-1 User Authentication

- SDA-2 As a user, I can ... DONE
- SDA-3 As a user, I can ... DONE
- SDA-4 As a user, I can ... DONE

SDA-6 Search and Filter Functi... DONE

- SDA-9 As a user, I can ... DONE
- SDA-10 As a user, I can ... DONE

SDA-7 Channel Details Page DONE

- SDA-12 As a user, I can... DONE

+ Create Epic

EP OCT Spr... S...

Today Weeks Months Quarters i

SDA-7 Channel Details Page

Done Actions Description Add a description... Child issues Order by ...

- SDA-12 As a user, I ... = 20 DONE

RECOMMENDED FOR EPICS Project plan TEMPLATE Use this template to map out project logistics, like milestones and scope, to ensure a project runs smoothly. A

Add a comment... Pro tip: press M to comment

| SDA-1 User Authentication       | SDA-6 Search and Filter Functi... | SDA-7 Channel Details Page      |
|---------------------------------|-----------------------------------|---------------------------------|
| SDA-2 As a user, I can ... DONE | SDA-9 As a user, I can ... DONE   | SDA-12 As a user, I can... DONE |
| SDA-3 As a user, I can ... DONE | SDA-10 As a user, I can ... DONE  |                                 |
| SDA-4 As a user, I can ... DONE |                                   |                                 |

Smartbridge data analytics - x +

jayarammovva9.atlassian.net/jira/software/projects/SDA/boards/1/timeline?assignee=712020%3Acdc9bcc1-f36e-4eec-8375-3bd8f14a3f5c%2C712...

Jira Software Your work Projects Filters Dashboards Teams Apps Create Search Give feedback Share Export ...

Smartbridge data anal... Software project You're on the Free plan UPGRADE

PLANNING Timeline Backlog Board + Add view DEVELOPMENT Code Project pages Add shortcut Project settings You're in a team-managed project Learn more

Timeline

Sprints

SDA-1 User Authentication

- SDA-2 As a user, I can ... DONE
- SDA-4 As a user, I can ... DONE

SDA-6 Search and Filter Functi... DONE

- SDA-10 As a user, I can... DONE

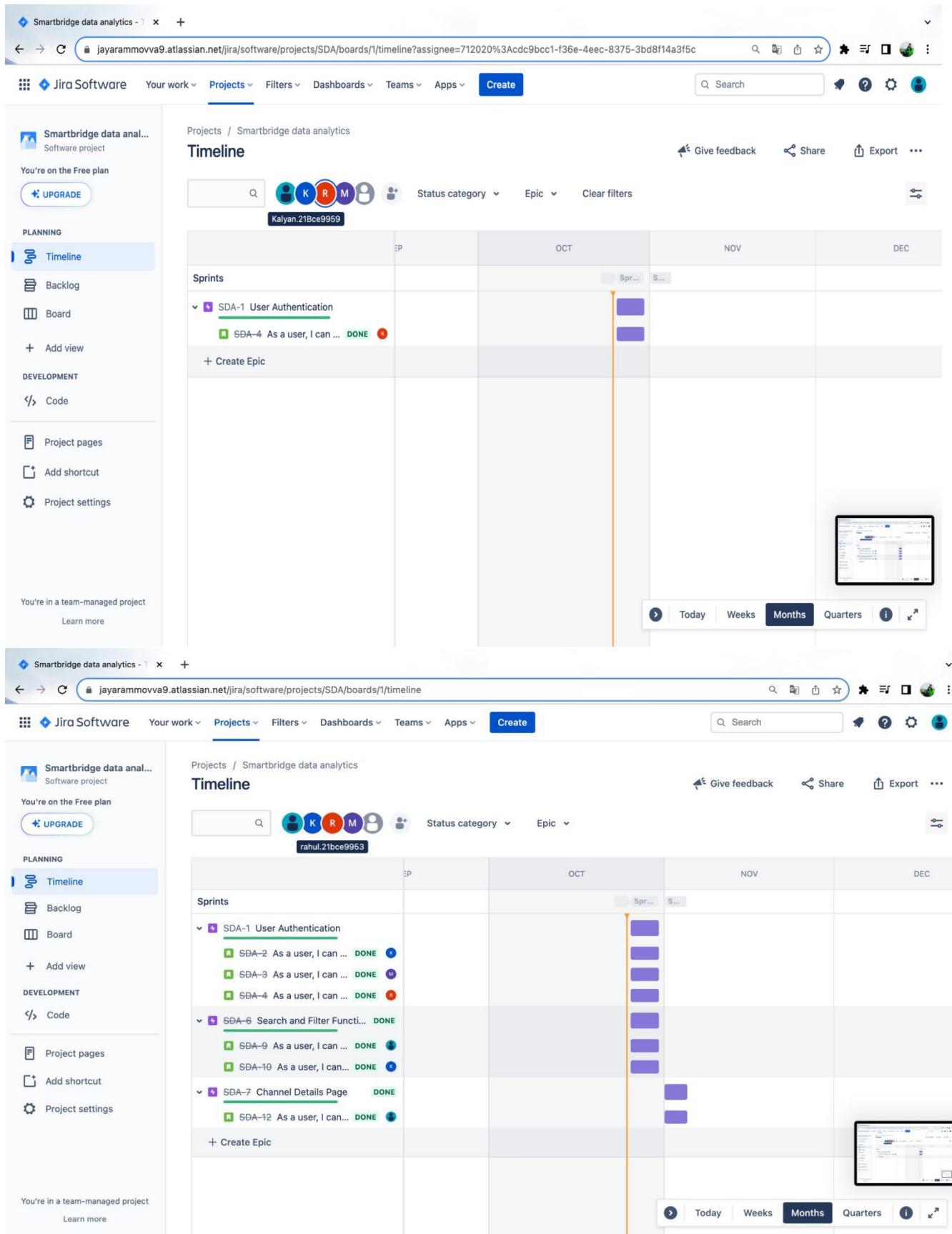
+ Create Epic

EP OCT NOV DEC Spr... S...

Today Weeks Months Quarters i

MOVVA JAYARAM 21BCE8265

| SDA-1 User Authentication       | SDA-6 Search and Filter Functi... |
|---------------------------------|-----------------------------------|
| SDA-2 As a user, I can ... DONE | SDA-10 As a user, I can... DONE   |
| SDA-4 As a user, I can ... DONE |                                   |



## 7) PERFORMANCE TESTING:

### 7.1 Performance Metrics

### **MODEL PERFORMANCE TESTING:(Tableau)**

| S.NO | Parameter                            | Screenshot/Values  |
|------|--------------------------------------|--|
| 1    | Dashboard Design                     | No of Visualisations/Graphs -- 9 Visualisations            |
| 2    | Data Responsiveness                  |  |
| 3    | Amount Data to Rendered(DB2 Metrics) | No Database is used, Directly connected through Excel file |
| 4    | Utilization of Data Filters          | Yes-2 Filters used   |
| 5    | Effective User Story                 | No of Scene Added-4 Scenes                                 |
| 6    | Descriptive Reports                  | No of Visualisations/Graphs-No Report is created           |

### **8) RESULTS:**

#### **8.1 Output Screenshots**

##### **SPRINT-1**

#### **DATA EXTRACTION:**

WPS Office | List of most-subscribed YouTube channels | + | Free trial | Share |

Menu Insert Page Layout Formulas Data Review View Special features Unynchronized Share

Cut Copy Format Painter Calibri 11 A A Merge and Center Wrap Text General Conditional Formatting

Format as Table AutoSum AutoFilter Sort Fill Format Row Col

A1 Q fx

| Rank | Name                          | Link | Brand channel | Subscribers       | Primary language | Category       | Country |
|------|-------------------------------|------|---------------|-------------------|------------------|----------------|---------|
| 0    | T-Series                      | Link | Yes           | 238 Hindi[7][8]   | Music            | India          |         |
| 1    | Cocomelon                     | Link | Yes           | 155 English       | Education        | United States  |         |
| 2    | Sony Entertainment Television | Link | Yes           | 153 Hindi[9]      | Entertainment    | India          |         |
| 3    | MrBeast                       | Link | No            | 137 English       | Entertainment    | United States  |         |
| 4    | PewDiePie                     | Link | No            | 111 English       | Games            | Sweden         |         |
| 5    | Kids Diana Show               | Link | Yes           | 109 English[10]   | Entertainment    | Ukraine        |         |
| 6    | Like Nasty                    | Link | No            | 105 English       | Entertainment    | United States  |         |
| 7    | Vlad and Niki                 | Link | No            | 94.9 English      | Entertainment    | Russia         |         |
| 8    | WWE                           | Link | Yes           | 93.8 English      | Sports           | United States  |         |
| 9    | Zee Music                     | Link | Yes           | 93.4 Hindi[13][1] | Music            | India          |         |
| 10   | Blackpink                     | Link | Yes           | 84.8 Korean       | Music            | South Korea    |         |
| 11   | Goldmines                     | Link | Yes           | 83.3 Hindi        | Film             | India          |         |
| 12   | 5-Minute                      | Link | Yes           | 79.2 English      | How-to           | Cyprus[a]      |         |
| 13   | Sony SAB                      | Link | Yes           | 78.2 Hindi        | Entertainment    | India          |         |
| 14   | BangtanTV                     | Link | No            | 73.9 Korean       | Music            | South Korea    |         |
| 15   | Justin Bieber                 | Link | No            | 71.1 English      | Music            | Canada         |         |
| 16   | Hybe Lab                      | Link | Yes           | 69.5 Korean       | Music            | South Korea    |         |
| 17   | Canal Kon                     | Link | Yes           | 66.4 Portuguese   | Music            | Brazil         |         |
| 18   | Zee TV                        | Link | Yes           | 66.3 Hindi        | Entertainment    | India          |         |
| 19   | Pinkfong                      | Link | Yes           | 66.1 English      | Education        | South Korea    |         |
| 20   | Shemaroo                      | Link | Yes           | 64.2 Hindi        | Music            | India          |         |
| 21   | ChuChu TV                     | Link | Yes           | 63.2 Hindi[16]    | Education        | India          |         |
| 22   | Colors TV                     | Link | Yes           | 60.6 Hindi        | Entertainment    | India          |         |
| 23   | Dude Perfect                  | Link | No            | 59 English        | Sports           | United States  |         |
| 24   | Movieflicks                   | Link | Yes           | 58.8 English      | Film             | United States  |         |
| 25   | T-Series                      | Link | Yes           | 58.4 Hindi        | Music            | India          |         |
| 26   | Tips Indus                    | Link | Yes           | 57.3 Hindi        | Entertainment    | India          |         |
| 27   | Wave Music                    | Link | Yes           | 56.6 Bhojpuri     | Music            | India          |         |
| 28   | Marshmello                    | Link | No            | 56.2 English      | Music            | United States  |         |
| 29   | Sony Music                    | Link | Yes           | 56.1 Hindi        | Music            | India          |         |
| 30   | El Reino In                   | Link | Yes           | 56 Spanish        | Music            | Argentina      |         |
| 31   | Aaj Tak                       | Link | Yes           | 55.9 Hindi        | News             | India          |         |
| 32   | Eminem                        | Link | No            | 55.6 English      | Music            | United States  |         |
| 33   | LooLoo KiKi                   | Link | Yes           | 53.1 English      | Music            | Romania        |         |
| 34   | Ed Sheerha                    | Link | No            | 53 English        | Music            | United Kingdom |         |
| 35   | Yash Raj                      | Link | Yes           | 52.7 Hindi        | Music            | India          |         |
| 36   | Ariana Grande                 | Link | No            | 52.4 English      | Music            | United States  |         |
| 37   | Twinkl                        | Link | No            | 51.4 English      | Education        | United States  |         |
| 38   | YouTube Kids                  | Link | No            | 51.4 English      | Education        | United States  |         |

List of most-subscribed YouTube (youtube)

Connections Add

youtube Microsoft Excel

Sheets

List of most-subscribed YouTube

New Union

New Table Extension

Need more data?

Drag tables here to relate them. [Learn more](#)

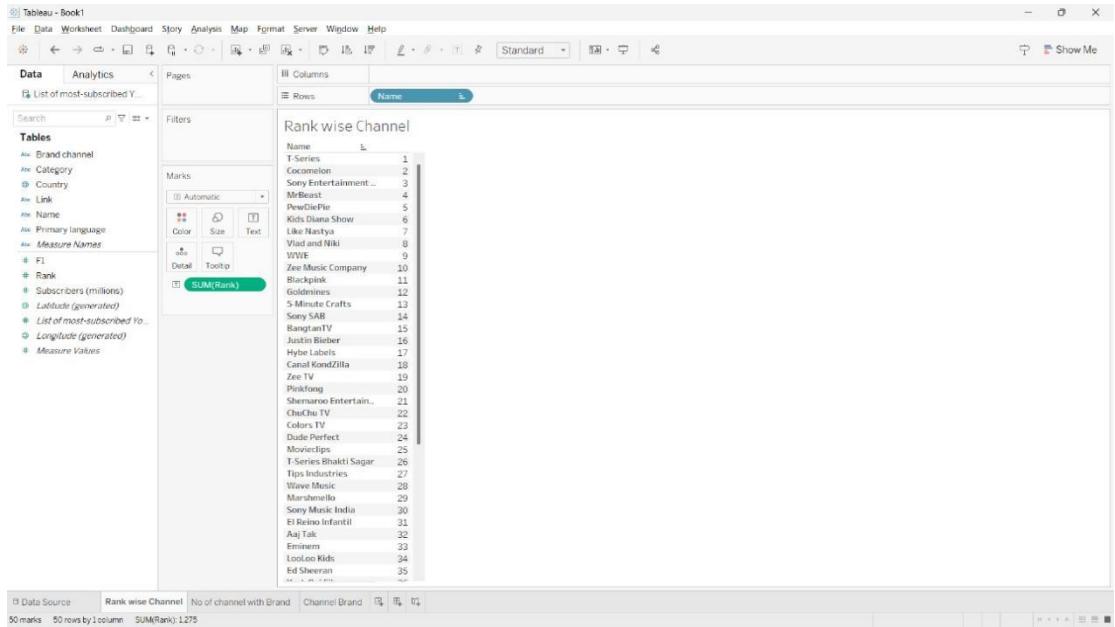
List of most-subscribed YouTube (youtube) 9 fields 50 rows

| # | F1 | Rank | Name                          | Link |
|---|----|------|-------------------------------|------|
| 0 |    | 1    | T-Series                      | Link |
| 1 |    | 2    | Cocomelon                     | Link |
| 2 |    | 3    | Sony Entertainment Television | Link |
| 3 |    | 4    | MrBeast                       | Link |
| 4 |    | 5    | PewDiePie                     | Link |
| 5 |    | 6    | Kids Diana Show               | Link |
| 6 |    | 7    | Like Nasty                    | Link |
| 7 |    | 8    | Vlad and Niki                 | Link |

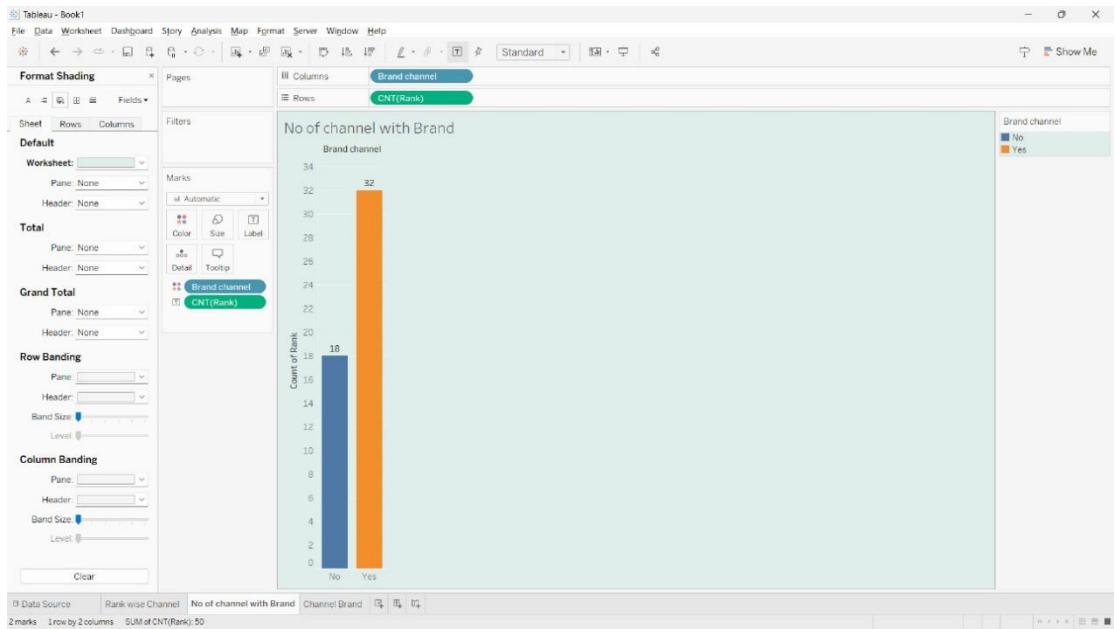
Rows: 50 | Go to Worksheet | Data Source | Sheet 1 |

## VISUALISATION:

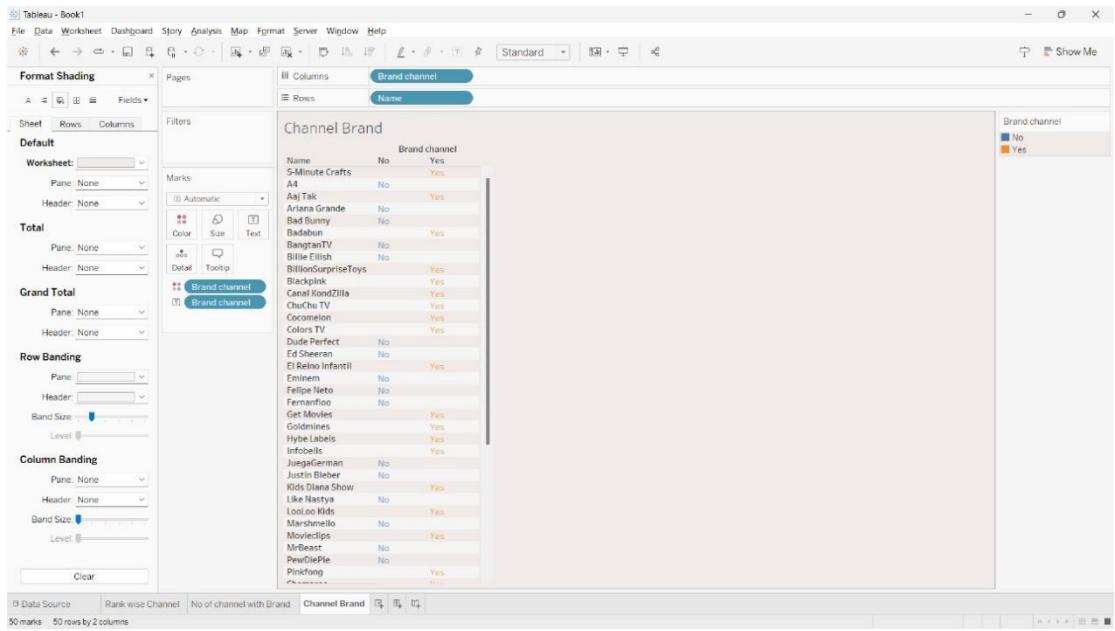
## 1) Rank wise channel:



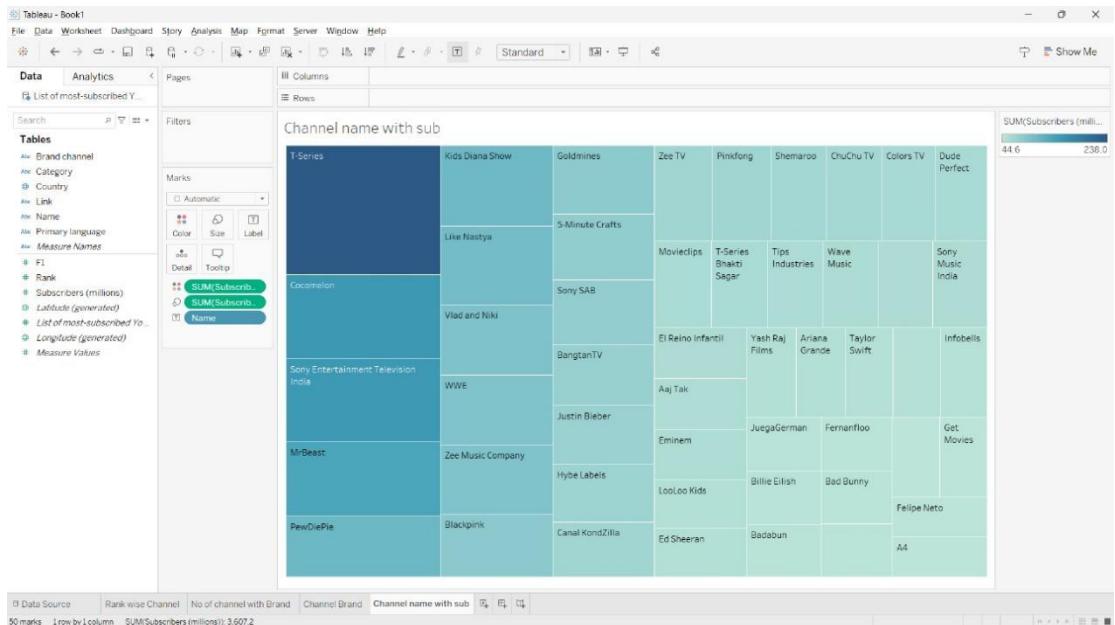
## 2) No of channel with Brand:



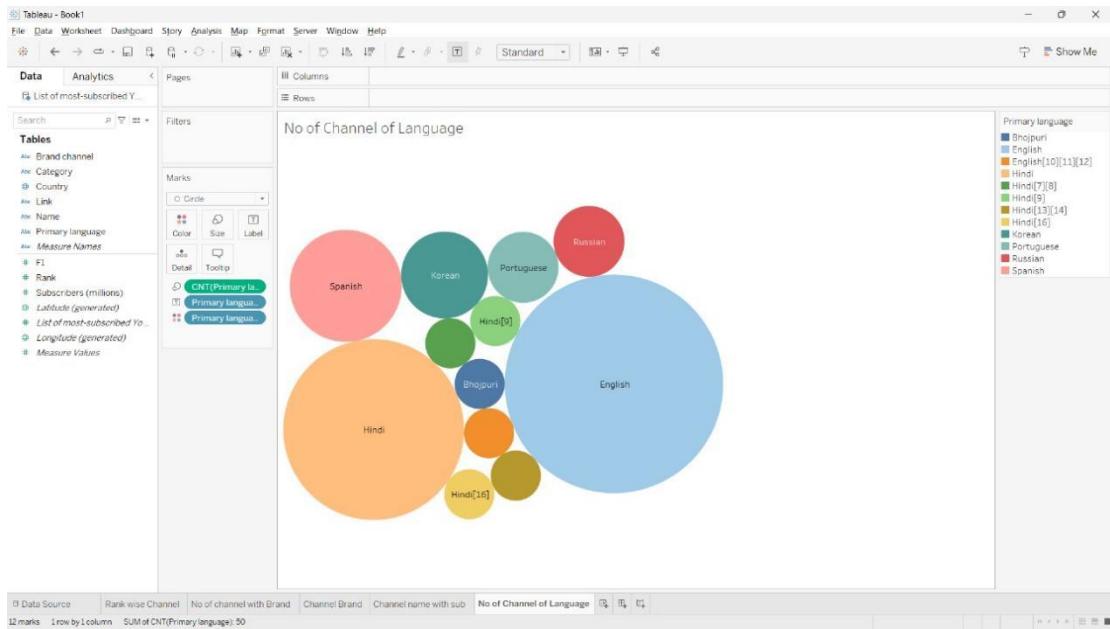
## 3) Channel Brand:



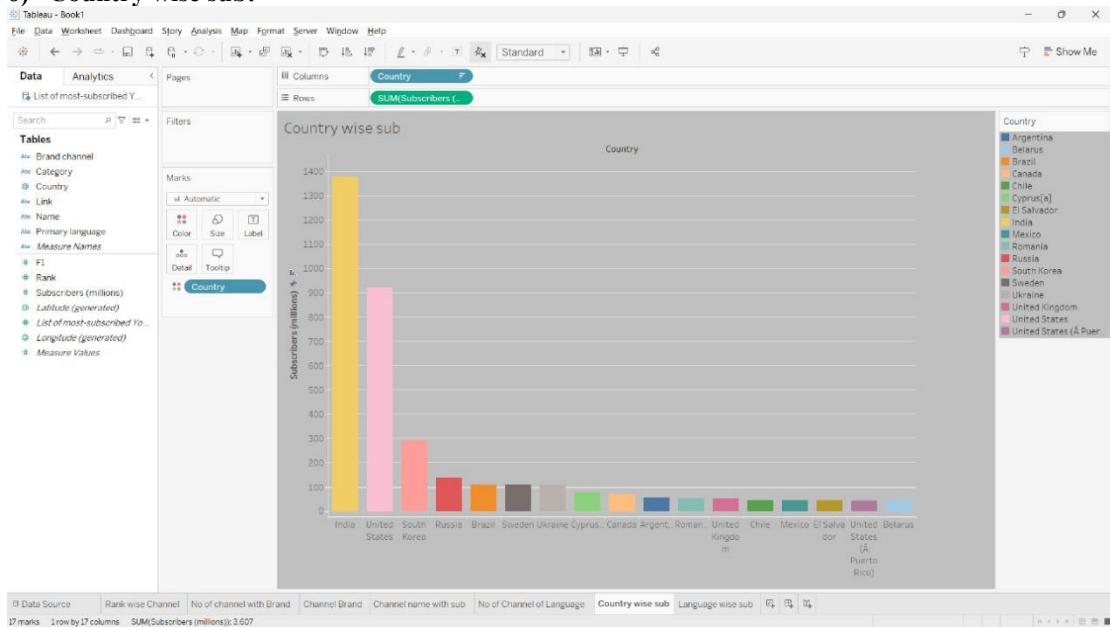
#### 4) Channel name with Sub:



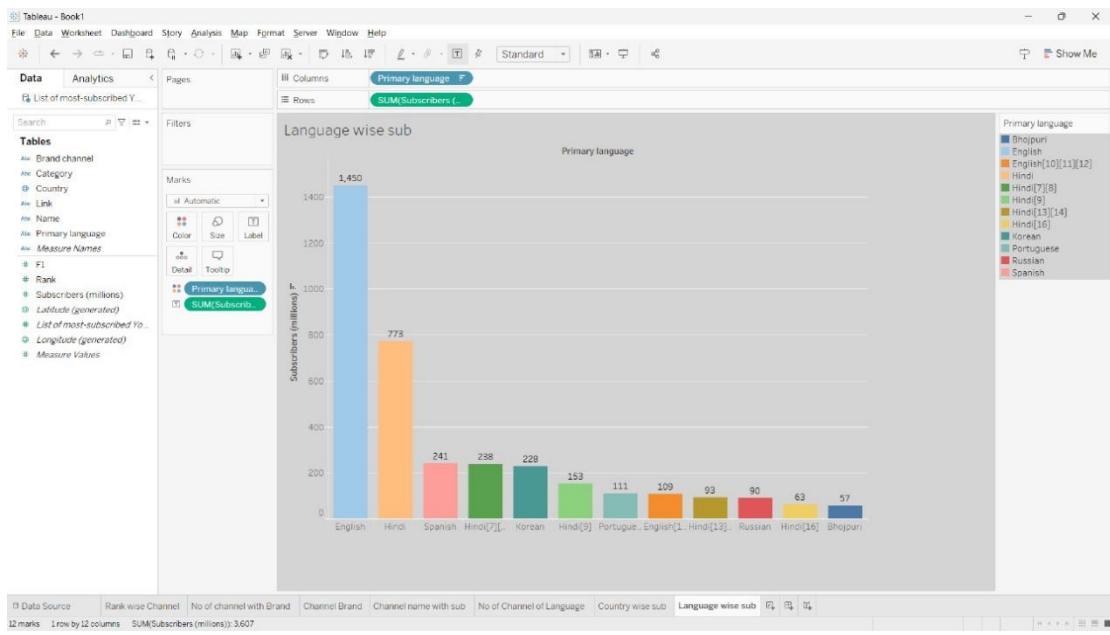
#### 5) No of Channel of Language:



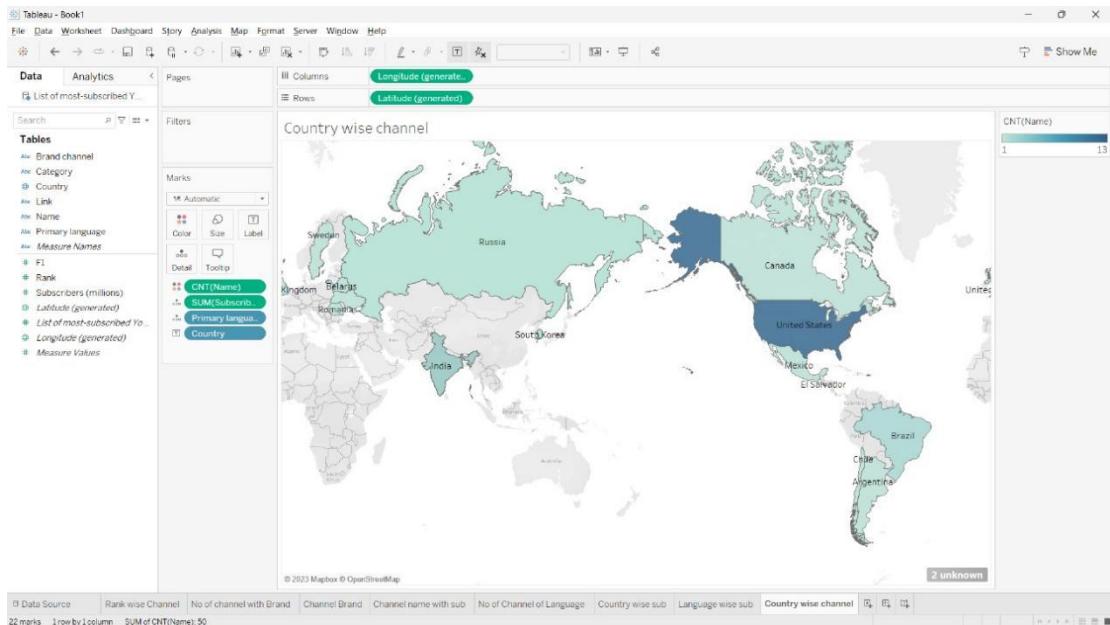
## 6) Country wise sub:



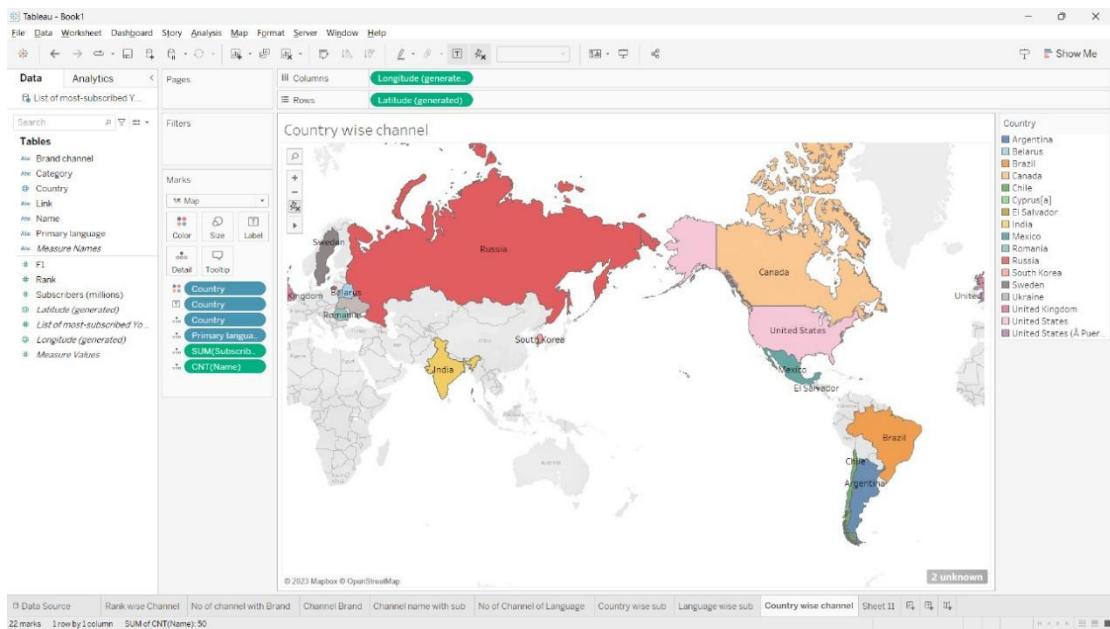
## 7) Language wise sub:



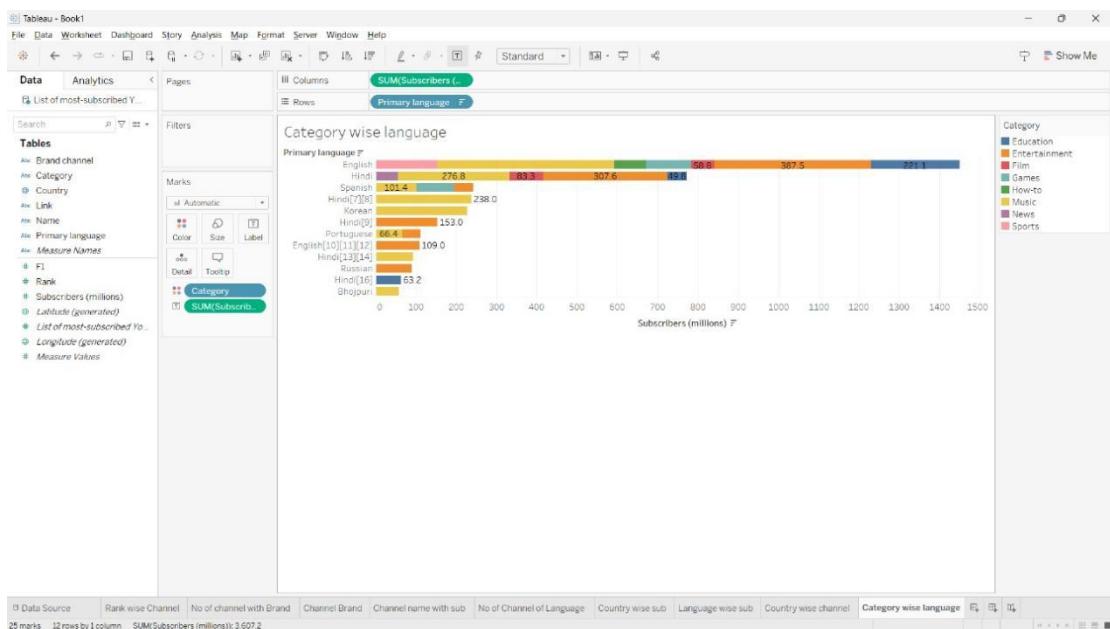
#### **8) Country wise Channel:**



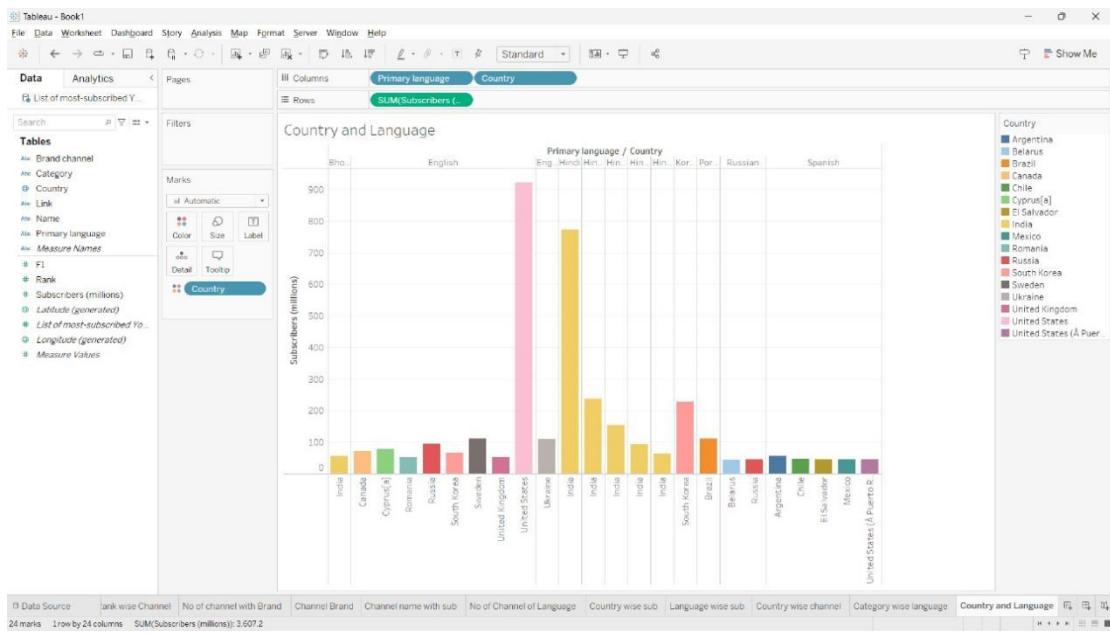
### **9) Country wise Channel:**



## 10) Category wise Language:



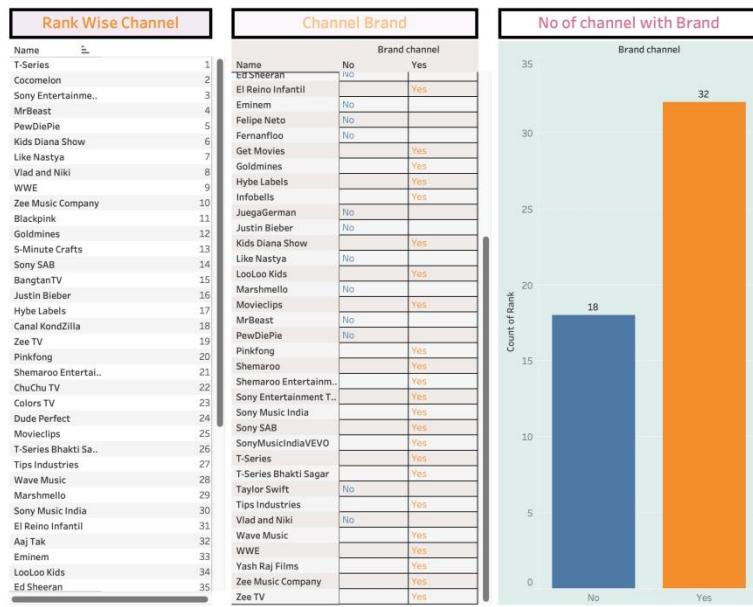
## 11) Country and Language:



SPRINT-2

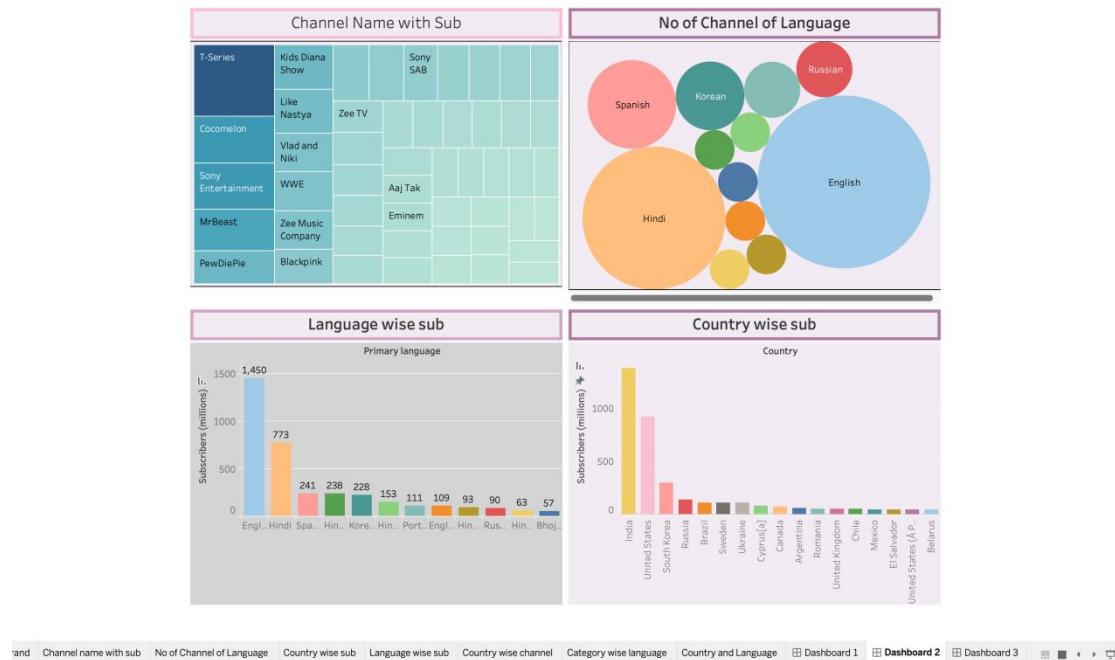
## DASHBOARD:

1)



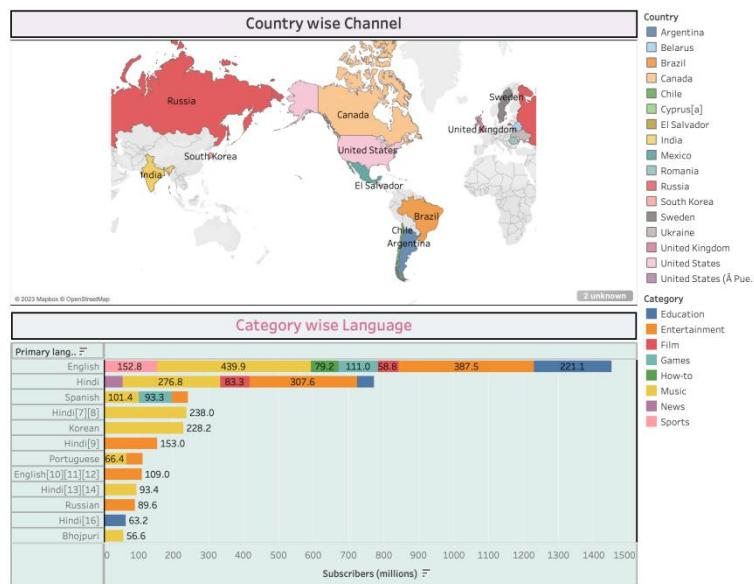
rand Channel Brand Channel name with sub No of Channel of Language Country wise sub Language wise sub Country wise channel Category wise language Country and language Dashboard 1 Dashboard 2

2)



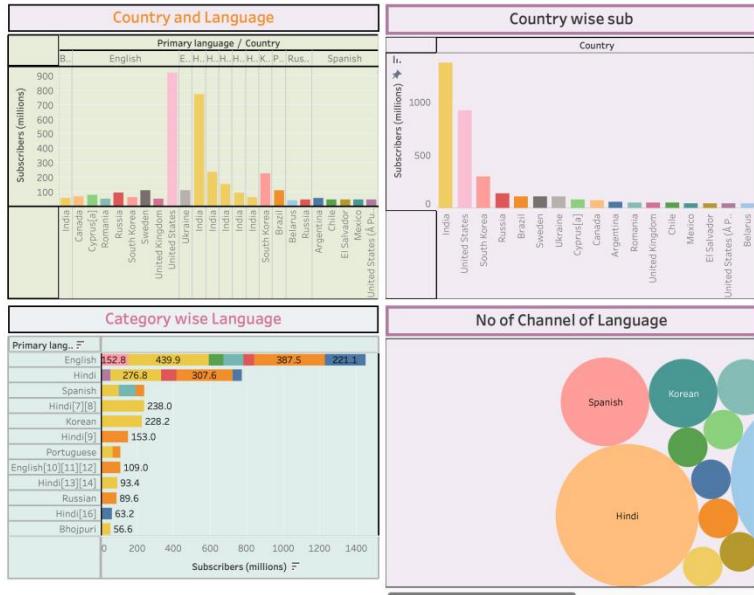
and Channel name with sub No of Channel of Language Country wise sub Language wise sub Country wise channel Category wise language Country and Language Dashboard 1 Dashboard 2 Dashboard 3

3)



of Channel of Language Country wise sub Language wise sub Country wise channel Category wise language Country and Language Dashboard 1 Dashboard 2 Dashboard 3 Dashboard 4 Dashboard 5

4)

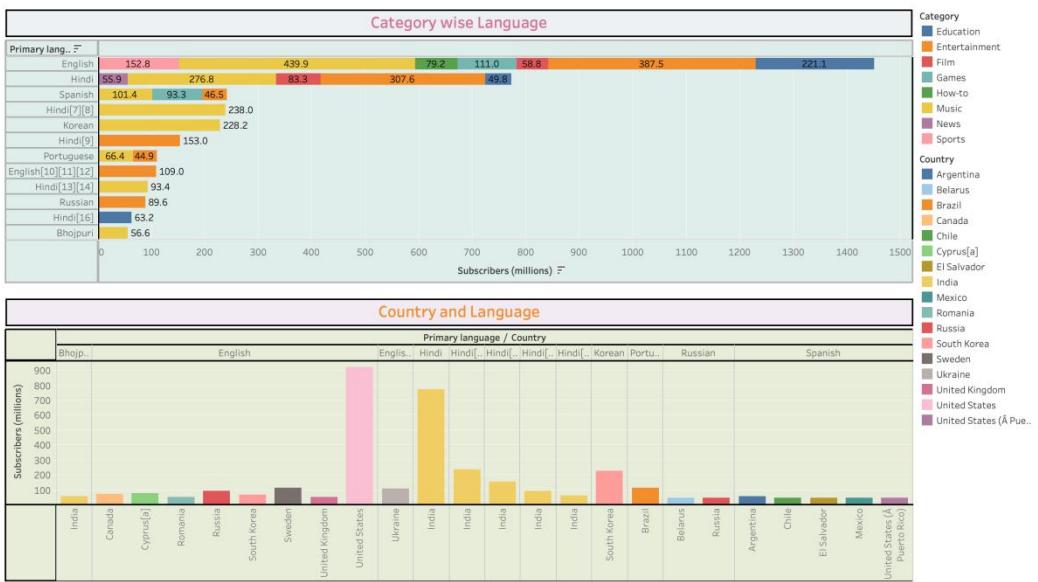


Channel of Language Country wise sub Language wise sub Country wise channel Category wise language Country and Language  Dashboard 1  Dashboard 2  Dashboard 3  Dashboard 4  Dashboard 5 

5)

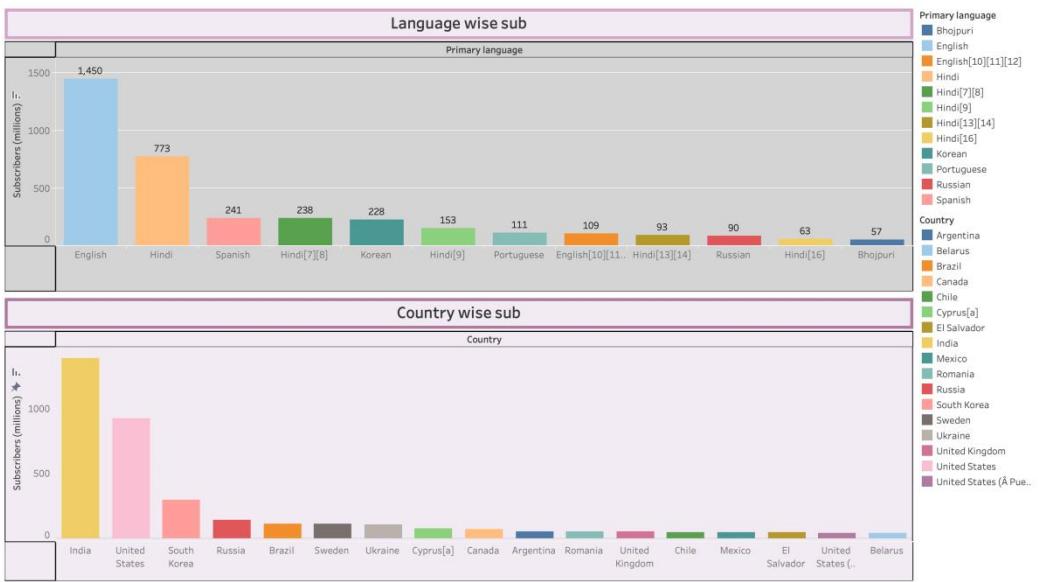


intel Category wise language Country and Language Dashboard 1 Dashboard 2 Dashboard 3 Dashboard 4 Dashboard 5 Dashboard 6 Dashboard 7 Dashboard 8 Story 1 Sheet 12 Sheet 11 Sheet 13



inlet Category wise language Country and Language Dashboard 1 Dashboard 2 Dashboard 3 Dashboard 4 Dashboard 6 Dashboard 7 Dashboard 8 Story 1 Sheet 12 Sheet 11 Sheet 13

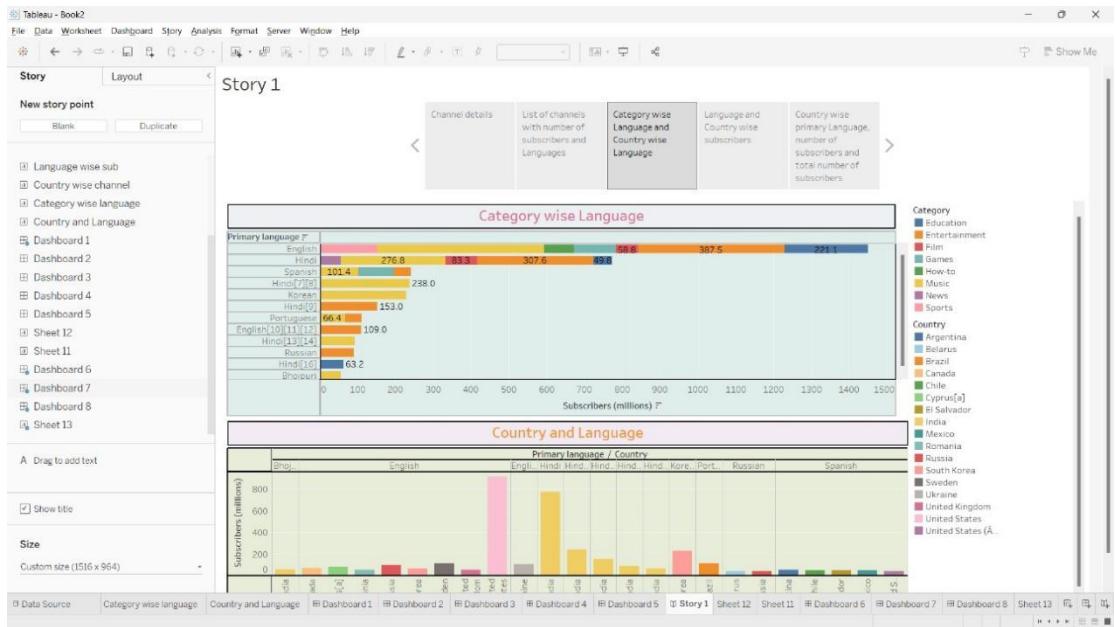
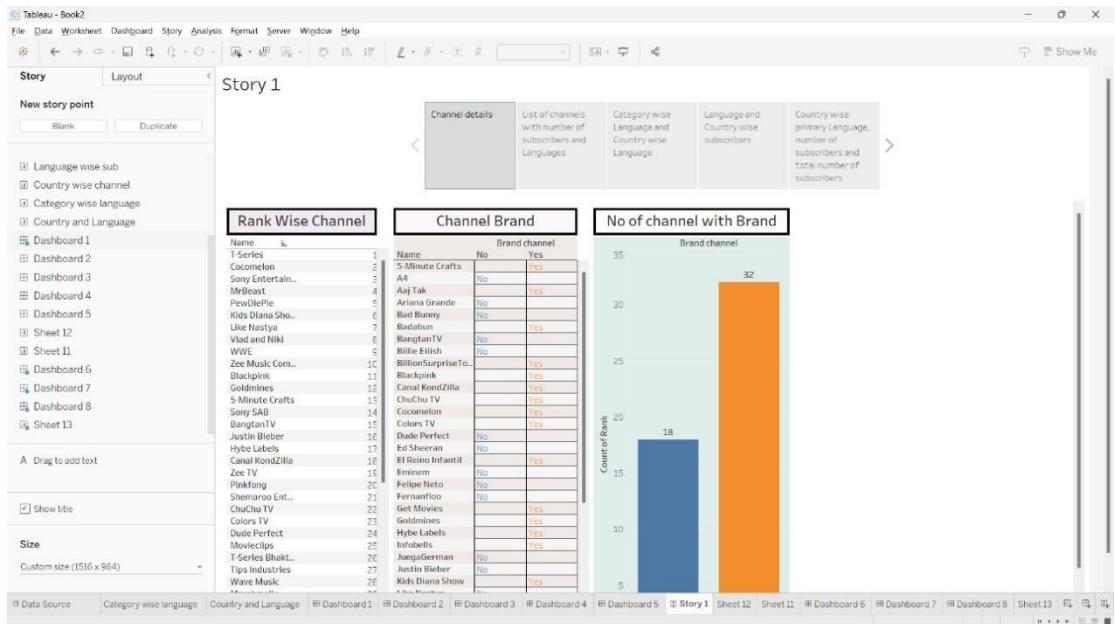
7)

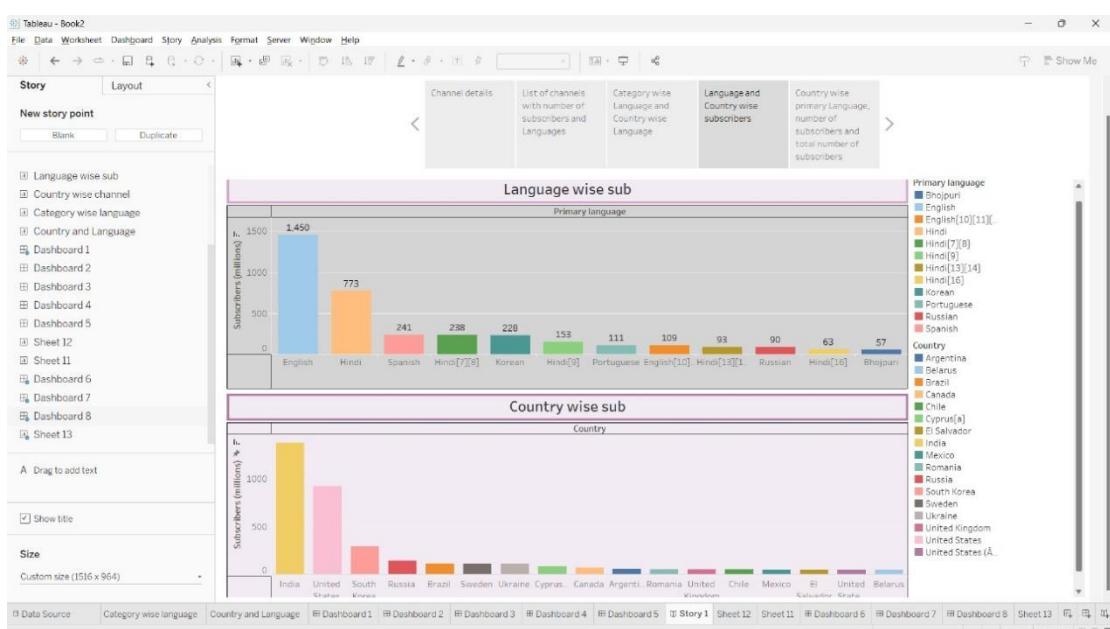
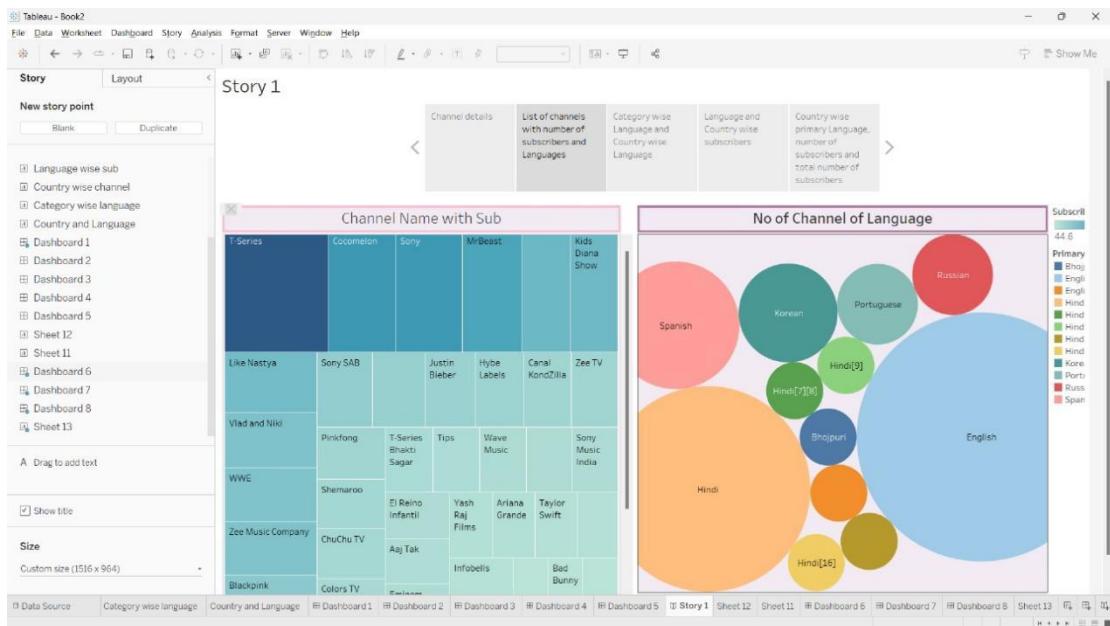


inlet Category wise language Country and Language Dashboard 1 Dashboard 2 Dashboard 3 Dashboard 4 Dashboard 6 Dashboard 7 Dashboard 8 Story 1 Sheet 12 Sheet 11 Sheet 13

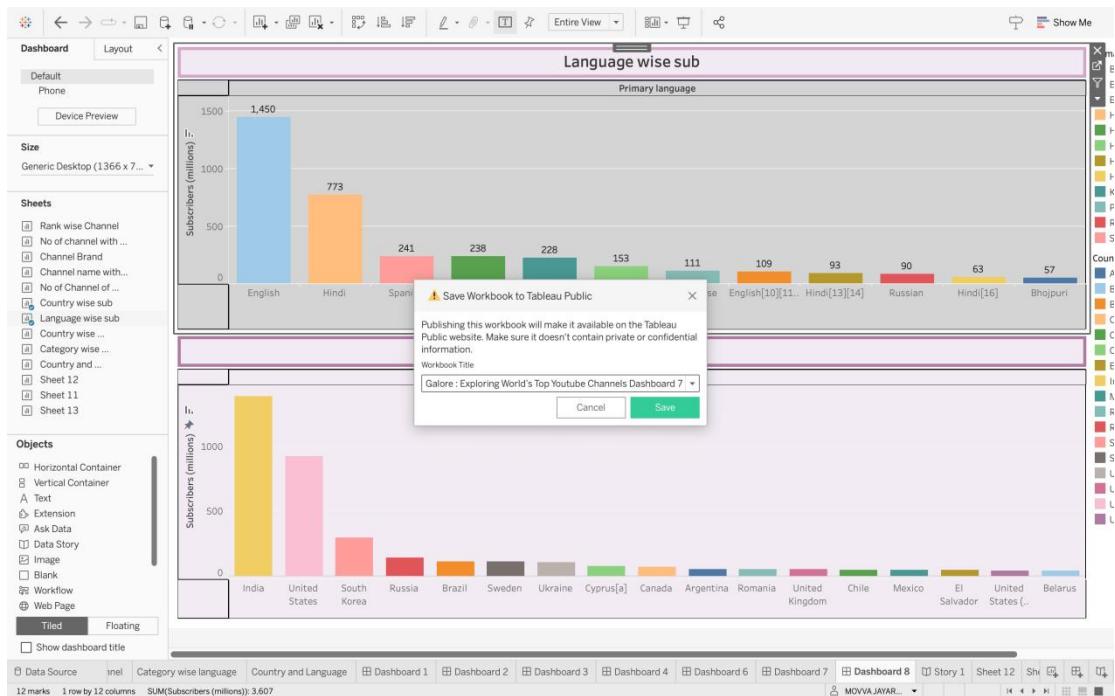
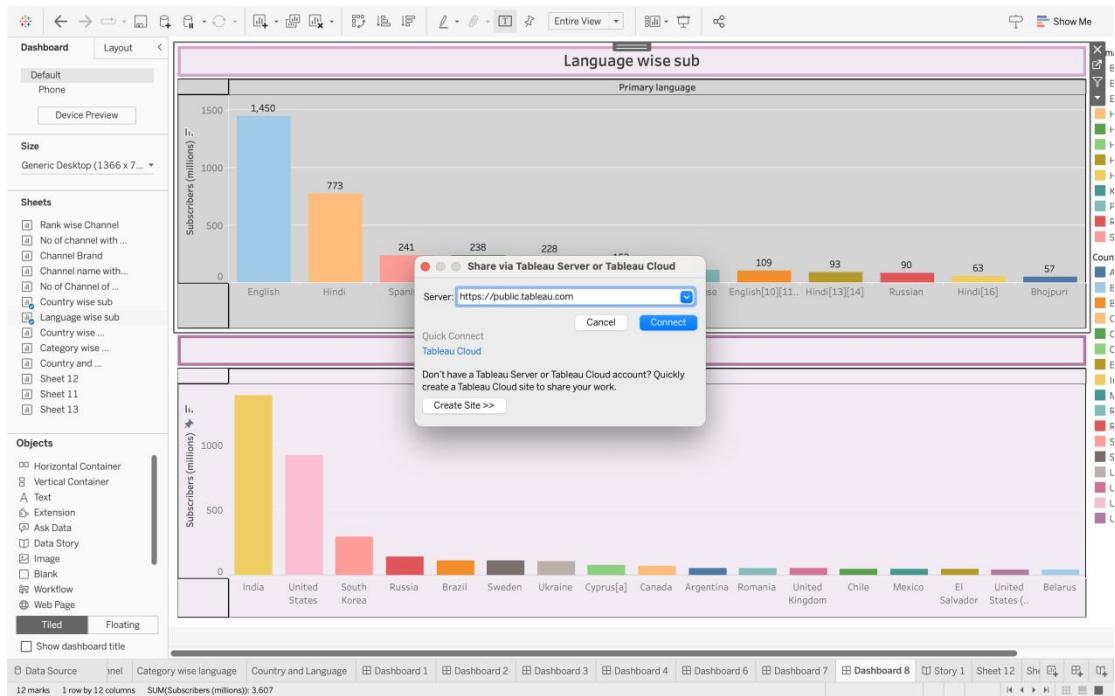
## SPRINT-3

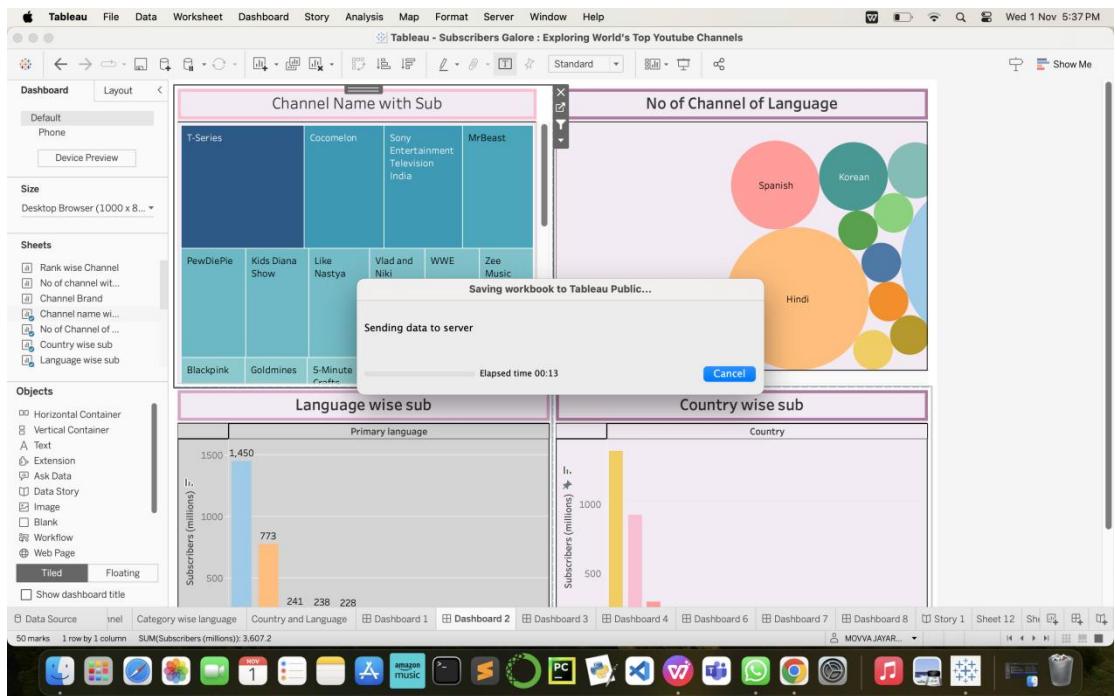
### STORY:



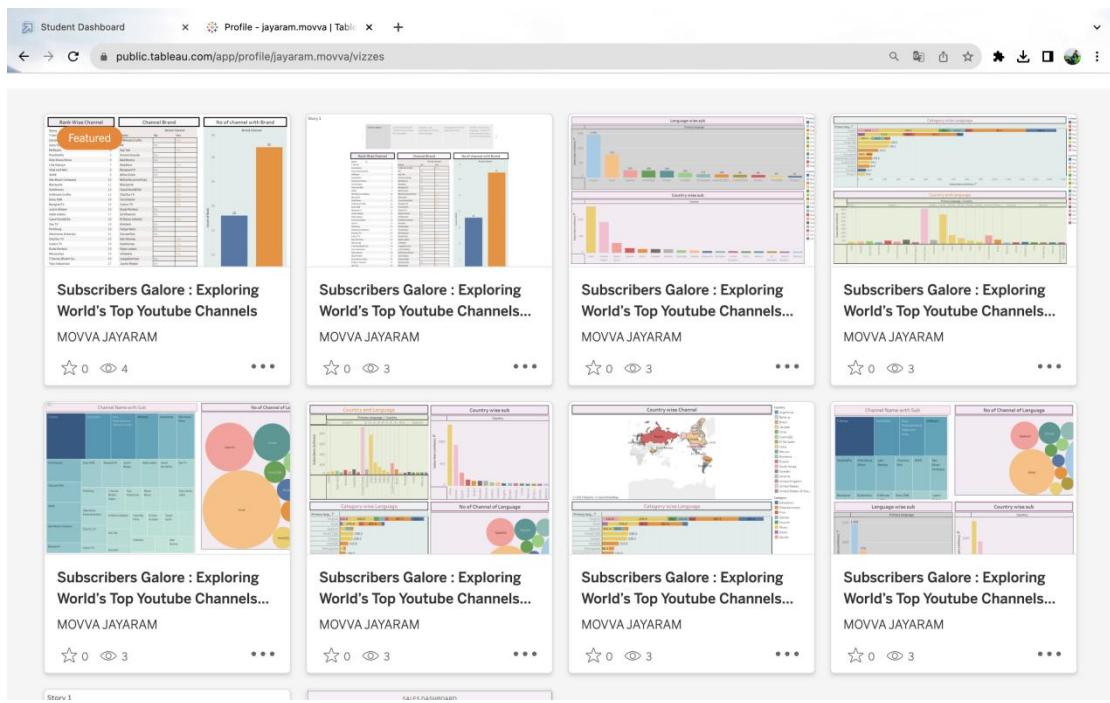


## PUBLISHING:





This screenshot shows a Tableau Public profile page. At the top, there's a navigation bar with links for 'Student Dashboard', 'Profile - jayaram.movva | Table', and 'Create' and 'Learn' buttons. Below the navigation is a banner with a green gradient background and a circular profile picture of a man. The banner text reads: 'Want to take your data skills to the next level? Connect with the Tableau Community to accelerate your learning. [Show me →](#)' and 'Customize Banner'. To the right of the profile picture are social media icons for Twitter, Facebook, and LinkedIn. The main profile area features the name 'MOVVA JAYARAM' in large, bold, black capital letters, followed by the subtitle 'Vijayawada, Andhra Pradesh, India'. Below this are two buttons: 'Edit Profile' and 'Hire Me'. Underneath the name are four metrics: 'Vizzes 10', 'Favorites 0', 'Following 0', and 'Followers 0'. A blue button labeled 'Create a Viz' is located on the far right. At the bottom, there are four preview cards for different data visualizations, each with a 'Featured' badge.



## 9) ADVANTAGES & DIS ADVANTAGES

### ADVANTAGES:

- 1. Informed Decision-Making:** Enables data-driven decisions for content creators, businesses, and marketers.
- 2. Subscriber Growth Insights:** Provides a deeper understanding of factors influencing subscriber growth.
- 3. Content Strategy Optimization:** Helps creators refine content strategies by analyzing what attracts subscribers.
- 4. Demographic Understanding:** Offers insights into subscriber demographics for more targeted marketing.
- 5. Competitive Analysis:** Identifies successful strategies used by top YouTube channels.
- 6. Predictive Modeling:** Allows for forecasting future subscriber counts for goal setting and planning.
- 7. User Engagement Improvement:** Aids creators in improving audience engagement and retention.
- 8. Data-Driven Marketing:** Assists businesses in making informed decisions for influencer marketing and advertising on YouTube.
- 9. Research Opportunities:** Provides valuable data and insights for enthusiasts and researchers studying online trends.
- 10. Clear Communication:** Visualizations and reporting features facilitate effective communication of findings to stakeholders.

### DIS ADVANTAGES:

- 1. Data Privacy Concerns:** Collecting and analyzing data on YouTube channels may raise privacy issues if the project involves scraping or using publicly available but sensitive information.
- 2. Incomplete Data:** The project may face limitations in accessing comprehensive and up-to-date data for all top YouTube channels, which could lead to biased or incomplete analyses.
- 3. Data Quality Issues:** Data collected from various sources may contain errors, inconsistencies, or missing values, affecting the accuracy of the analysis.
- 4. Ethical Considerations:** Analyzing YouTube channels may raise ethical concerns, particularly if it involves potentially controversial or inappropriate content.
- 5. Sampling Bias:** The data collected may not be representative of the entire population of YouTube channels, leading to skewed results.
- 6. Changing Algorithmic Factors:** YouTube's recommendation and ranking algorithms frequently change, which can impact the accuracy and relevance of the analysis over time.
- 7. Biased User Behavior:** User behavior on YouTube, such as liking, commenting, and subscribing, can be influenced by factors that are not easily measurable, leading to biased results.
- 8. Competition and Secrecy:** Some YouTube channels may employ strategies to manipulate their subscriber counts or hide certain information, making it challenging to obtain accurate data.
- 9. Data Security:** Storing and handling large datasets with potentially sensitive information may pose security risks, such as data breaches or unauthorized access.
- 10. Limited Predictive Power:** While data analytics can provide insights, it may have limitations in accurately predicting the future success or trends of YouTube channels, as many factors are beyond the scope of the project.

## 10) CONCLUSION

In conclusion, the project "Subscribers Galore: Exploring World's Top YouTube Channels" offers the potential for valuable insights into the performance and trends of YouTube channels. However, it is essential to be aware of the numerous disadvantages and challenges associated with such a project. These include concerns about data privacy, the potential for incomplete or low-quality data, ethical considerations, and the ever-changing nature of YouTube's algorithms.

Furthermore, issues related to sampling bias, biased user behavior, competition, and data security should be carefully addressed to ensure the project's credibility and usefulness. It's important to recognize that while data analytics can provide valuable insights, it may have limitations in accurately predicting the future success or trends of YouTube channels, as there are many complex and external factors involved.

To mitigate these disadvantages and maximize the project's potential, a well-planned and ethical approach to data collection, analysis, and interpretation is crucial. Additionally, staying up to date with YouTube's policies and algorithmic changes is essential to maintain the project's relevance and accuracy over time. Despite these challenges, a thoughtful and rigorous data analytics project can still provide valuable insights into the world of YouTube channels and online content creation.

## 11) FUTURE SCOPE:

- 1. Predictive Analytics:** The future scope includes the development of predictive models using historical data and machine learning to forecast the success and trends of YouTube channels, aiding content creators, and marketers in making data-driven decisions.

**2. Advanced Recommendation Algorithms:** Enhancing content recommendation algorithms based on user behavior and channel attributes to improve user engagement and satisfaction on YouTube and other video platforms.

**3. Audience Segmentation:** Segmenting YouTube channel audiences by their preferences and behavior to enable personalized content creation and targeted marketing, facilitating channel growth and monetization.

**4. Content Moderation and Ethics:** Future projects can focus on the development of AI-driven tools for content moderation and ethical assessments, helping address concerns regarding the appropriateness and legality of YouTube content.

**5. Influencer Marketing and Collaboration:** Analyzing the impact of influencer marketing on channel growth and exploring collaboration opportunities for content creators and brands as influencer marketing continues to evolve.

These future scope points highlight the potential for more sophisticated data analytics and insights, the improvement of user experiences, and the alignment of YouTube content with evolving ethical and regulatory standards.

## **12) APPENDIX:**

### **Tableau Public Link:**

<https://public.tableau.com/app/profile/jayaram.movva/vizzes>

### **GitHub Repository Link:**

<https://github.com/smartinternz02/SI-GuidedProject-587483-1697199874>

### **Video Demonstration Link:**

[https://drive.google.com/file/d/1R7u1euwScCv8ToVhStjQdILJJ5WCr6fH/view?usp=drive\\_link](https://drive.google.com/file/d/1R7u1euwScCv8ToVhStjQdILJJ5WCr6fH/view?usp=drive_link)