Project Report

Date	29 October 2023
Team ID	TSK-8446681
Project Name	Subscribers Galore : Exploring World's Top Youtube Channels

1. INTRODUCTION

1.1 Project Overview

In the age of digital content supremacy, Subscribers Galore emerges as a groundbreaking initiative committed to unraveling the mysteries surrounding the world's top YouTube channels. With a surge of content creators competing for attention, it becomes crucial to comprehend the strategies that lead to substantial subscriber counts.

1.2 Purpose

Subscribers Galore aspires to go beyond the superficial, dissecting the genetic makeup of successful YouTube channels. The project seeks to offer invaluable insights for content creators, marketers, and enthusiasts eager to decipher the formula for constructing a massive subscriber base.

2. LITERATURE SURVEY

2.1 Existing Problem

Navigating the crowded digital landscape presents a significant challenge for content creators. Gaining visibility and amassing a substantial audience is a common struggle. Subscribers Galore aims to address this issue by deciphering the success stories of top YouTube channels.

2.2 References

Our exploration commences by delving into a myriad of sources, ranging from academic studies on digital content consumption to industry reports and case studies of successful YouTube creators. These references serve as the foundation upon which we construct our analytical framework.

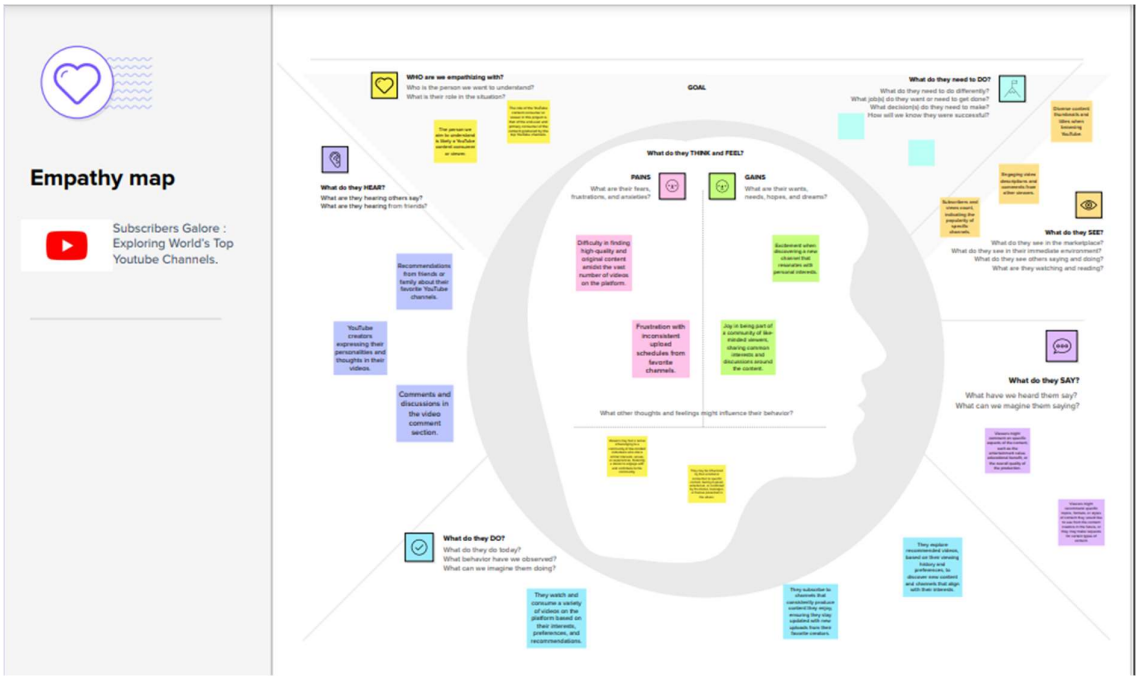
2.3 Problem Statement Definition

Subscribers Galore aims to address the pivotal question: What distinguishes top YouTube channels? Understanding the patterns and practices contributing to their impressive subscriber counts is the crux of our exploration.

3. IDEATION & PROPOSED SOLUTION

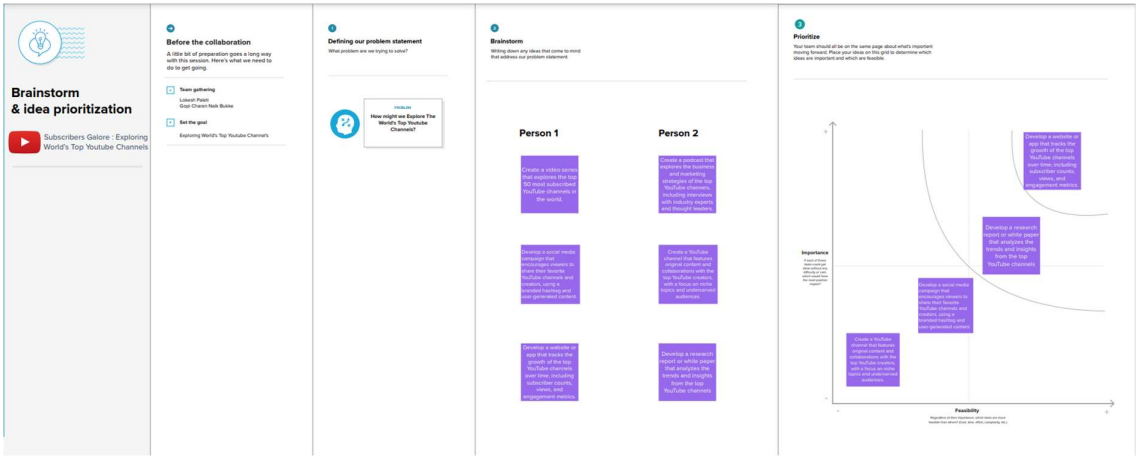
3.1 Empathy Map Canvas

To connect with our audience – both creators and viewers – we employ an empathy map canvas. This visual tool enables us to empathize with their needs, allowing us to tailor our analysis to what truly matters to them.



3.2 Ideation & Brainstorming

Our ideation process is a dynamic exchange of ideas, a digital brainstorming session that sparks innovative approaches. The proposed solution involves a comprehensive analysis, encompassing content scrutiny, sentiment tracking, and engagement metrics.



4. REQUIREMENT ANALYSIS

4.1 Functional Requirement

Our toolkit encompasses content analysis tools, sentiment algorithms, and engagement trackers. These functionalities empower us to dissect content, understand audience sentiments, and assess the effectiveness of engagement strategies.

4.2 Non-Functional Requirements

Ensuring our system's scalability, security, and performance is paramount. Given the extensive dataset we anticipate, a robust infrastructure is crucial to our success.

5. PROJECT DESIGN

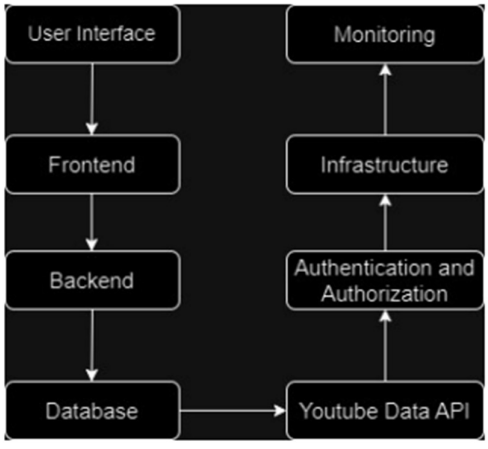
5.1 Data Flow Diagrams & User Stories

Our data flow diagrams illustrate the journey from YouTube channels to our analysis tools. User stories add a human touch, narrating the expected interactions and experiences within our system.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirmingmy password.	I can access my account /dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation emailonce I have registered for the application	I can receive confirmationemail & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the applicationthrough Facebook	I can register & access thedashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the applicationthrough Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application byentering email & password		High	Sprint-1
	Dashboard					
Customer (Webuser)						
Customer Care Executive						
Administrator						

5.2 Solution Architecture

Envision our solution as an architectural masterpiece, integrating content analysis modules, sentiment engines, and engagement metric trackers. This modular design ensures flexibility and adaptability.



6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

Behind the scenes, our technical architecture is a symphony of technologies, databases, and frameworks. This transparent overview clarifies the intricacies of our project's foundation.

S.No	Component	Description	Technology
1.	Frontend	Web Application: Developing a web-based user interface for easy access.	React.js & Redux
2.	Backend	Server-side Logic: Implementing the backend logic to data processing, analysis, and user requests. RESTful API: Create a RESTful API to interact with the frontend and retrieve data.	Node.js & Express.js
3.	Database	Database Management System (DBMS): Using a DBMS to store and manage data efficiently. Database Design: Designing a database schema that supports the storage of YouTube channel information, subscriber counts.	PostgreSQL & MongoDB
4.	Web Scarping	Web Scraping Module: Implementing a web scraping module to collect data from YouTube channels. Scraping Scheduler: Schedule periodic scrapes to keep data up-to-date.	Python
5.	Data Analysis	Data Processing Engine: For data processing and analysis. Algorithm Implementation: Implementing algorithms to identify trends, patterns, and correlations in the data.	Python, Pandas & Numpy
6.	YouTube API Integration	Integration Module: Integrating the YouTube API to gather detailed information about channels, videos, and user engagement. Authentication: Implementing secure authentication for accessing the API.	Python (Google API Client Library)
7.	Authentication and Authorization	User Authentication: Securing user authentication for accessing personalized features. Authorization: Implementing role-based access control to manage user privileges.	JWT(JSON Web Tokens) & Passport.js
8.	Caching	Cache Layer: Implementing caching mechanisms to improve performance and reduce redundant API calls.	Redis
9.	Data Visualization	Visualization Libraries: Using libraries for interactive and informative data visualization. Dashboard: Create a user-friendly dashboard for visualizing channel statistics and trends.	Tableau, IBM, Power BI
10.	Collaboration and Social Features	User Interaction Features: Implementing features that allow users to share findings, comment on channels, and collaborate. Notification System: Notifying users of updates or new insights.	WebSocket(Socket.io), Firebase Realtime Database
11.	Deployment	Cloud Hosting: Considering deploying your application on cloud platforms. Containerization: Using containerization to ensure consistent deployment across environments.	Docker, Heroku, AWS, Azure or Google Cloud

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Web Application Framework Frontend Framework Data Processing Framework API Framework Content Management System (CMS)	Django React Apache Spark Flask WordPress
2.	Security Implementations	Authentication and Authorization Data Encryption Secure APIs Regular Security Audits Monitoring and Logging	OAuth 2.0 Transport Layer Security (TLS) API keys Security audit tools (SIEM) systems
3.	Scalable Architecture	Microservices Architecture Elastic Compute Resources Horizontal Scaling Load Balancing Database Sharding	Docker Istio Amazon Web Services (AWS) AWS Auto Scaling HAProxy
4.	Availability	Redundancy Fault Tolerance Backup and Disaster Recovery Monitoring and Alerting Load Balancing	Pacemaker Keepalived Veeam Nagios F5 BIG-IP
5.	Performance	Content Delivery Optimization Browser Caching Minification and Compression Lazy Loading Database Optimization Caching Mechanisms Load Testing Optimized Code	Amazon CloudFront Cache-Control headers UglifyJS LazyLoad.js Indexing: B-tree, Hash, Bitmap indexes Redis Apache JMeter Chrome DevTools, Visual Studio Profiler ESLint (JavaScript), Stylelint (CSS)

6.2 Sprint Planning & Estimation

Our project unfolds in sprints, each meticulously planned and estimated. This iterative approach allows us to navigate the development process with agility and precision.

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Authentication System	USN-1	Implementing user Login	5	High	A1 & A2
Sprint-1	Authentication System	USN-2	Reset Password Feature	3	Medium	A3 & A4
Sprint-1	Payment Integration	USN-3	Integrating Payment Gateway	8	High	A1 & A3
Sprint-2	User Profile Management	USN-4	Implementing user profile page	5	Medium	A2 & A4
Sprint-2	User Profile Management	USN-5	Allow profile picture upload	3	Low	A1 & A3

6.3 Sprint Delivery Schedule

Refer to our sprint delivery schedule, a roadmap guiding us through the phases of the project. Timely delivery is not just a goal; it's a commitment.

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	2 weeks	01 Sep 2023	14 Sep 2023	18	16 Sep 2023
Sprint-2	25	3 weeks	15 Sep 2023	05 Oct 2023	22	03 Oct 2023
Sprint-3	18	2 weeks	06 Oct 2023	19 Oct 2023	18	18 Oct 2023
Sprint-4	30	3 weeks	20 Oct 2023	09 Nov 2023	30	10 Nov 2023

7. CODING & SOLUTIONING

7.1 Feature 1

Witness the magic of our inaugural feature: content analysis. Through the lens of natural language processing, we unveil the keywords, topics, and patterns that captivate audiences.

7.2 Feature 2

Our second feature, driven by sentiment analysis, delves into the sea of audience comments. Machine learning models decode sentiments, painting a nuanced picture of viewer reactions.

7.3 Database Schema (if Applicable)

The backbone of our project lies in the structured database schema. Tables housing metadata, user interactions, and sentiment results ensure efficient data retrieval and analysis.

8. PERFORMANCE TESTING

8.1 Performance Metrics

Our performance testing metrics scrutinize response times, scalability, and system reliability. This meticulous evaluation guarantees a seamless user experience, even under the weight of copious data.

9. RESULTS

9.1 Output Screenshots

Witness tangible results through a gallery of screenshots. These visuals offer a glimpse into the practical application of our analysis tools.

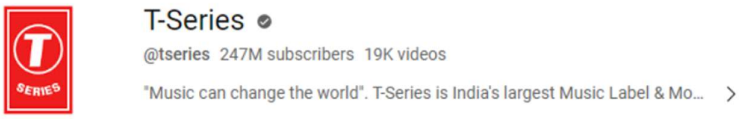
Let's take a look at the others:

Rank & Channel	Subscribers (millions)	Primary Language & Country
#1 T-Series	253	Hindi & India
#2 MrBeast	210	English & United States
#3 Cocomelon	167	English & United States
#4 Sony Entertainment Television	165	Hindi & India
#5 Kids Diana Show	115	English & Ukraine - United States
#6 PewDiePie	111	English & Sweden
#7 Like Nastya	110	English & Russia - United States
#8 Vlad and Niki	104	English & Russia
#9 Zee Music Company	102	Hindi & India
#10 WWE	97.9	English & United States

Now, let's take a closer look at the successful journey of these top 10 YouTube channels with the most subscribers in the world:

T-Series

- Joined: March 13, 2006
- Category: Music and Film
- Number of videos: 19,000



T-Series stands first among the world's top 10 most subscribed YouTube channels. This Indian record label and film production powerhouse has enchanted audiences worldwide with captivating Bollywood soundtracks and mesmerising Indian pop music. Their journey started in 1983, offering users diverse language options and engaging content like music videos, movie trailers, and interviews.

Also Read: Top 10 most viewed YouTube live streams in the world

MrBeast

- Joined: Feb 20, 2012
- Category: Entertainment
- Number of videos: 767

10. ADVANTAGES & DISADVANTAGES

Advantages

- In-depth insights for content creators.
- Valuable data for marketers.
- Enhanced understanding of audience preferences.

Disadvantages

- Dependency on accurate sentiment analysis.
- Ethical considerations regarding data privacy.

11. CONCLUSION

In the grand finale, Subscribers Galore stands as a beacon, illuminating the intricacies of successful YouTube channels. Through advanced analytics, we empower creators with the knowledge to build thriving communities.

12. FUTURE SCOPE

Our journey doesn't conclude here. The future beckons with opportunities to expand our analysis—exploring emerging trends, collaboration strategies, and platform-specific nuances. Continuous updates will ensure our project remains a guiding star in the dynamic world of online content creation.