12/4/2023

Final Project

Report

**A black and grey text

Description automatically generated**

**Group 4 – Team Members**

Leela Prasanna, Akkala

Hemanth Kumar, Ramanadham

Mabel, Ogonna

Priyanka, Chahande

Karan Manish, Bhosale

**Table Of Contents**

Executive Summary ---------------------------------------------------------------------------------- 1

1. Description of YouTube
   1. Overview of YouTube ---------------------------------------------------------------------- (3 – 4)
   2. History of YouTube ------------------------------------------------------------------------- 3
   3. Growth and Expansion --------------------------------------------------------------------- 3
   4. Products and Policies ----------------------------------------------------------------------- 3
   5. Revenue -------------------------------------------------------------------------------------- 4
   6. Challenges and Progress-------------------------------------------------------------------- 4
2. Business Strategy of YouTube----------------------------------------------------------------------- (6 – 9)
   1. SWOT Analysis ----------------------------------------------------------------------------- 6
   2. Five Forces Model -------------------------------------------------------------------------- 8
   3. Value Chain Analysis ----------------------------------------------------------------------- 9
3. Big Data Strategy -------------------------------------------------------------------------------------- (10-12)
   1. Data Monetization-----------------------------------------------------------------------------10
   2. Strategy1: Keeping data proprietary--------------------------------------------------------10
   3. Strategy2: Sharing data with partners ----------------------------------------------------- 10
   4. Strategy3: Selling data products ----------------------------------------------------------- 11
   5. Strategy4: Making the data freely available ----------------------------------------------12
   6. You tube’s Monetization policies ----------------------------------------------------------12
4. Big Data Management Strategies ------------------------------------------------------------------- 13
   1. Data Measures --------------------------------------------------------------------------------13
   2. Data Capture ----------------------------------------------------------------------------------13
   3. Data Fusion------------------------------------------------------------------------------------13
5. Privacy and security risks -----------------------------------------------------------------------------(14-15)
   1. Privacy and security risks--------------------------------------------------------------------14
   2. Privacy Issues ---------------------------------------------------------------------------------14
   3. Security Risks --------------------------------------------------------------------------------14
   4. Mitigating Risks ------------------------------------------------------------------------------15
6. Data Visualization -------------------------------------------------------------------------------------(16-18)

6.1. Data Visualization ---------------------------------------------------------------------------16

6.2. Tableau for Data Visualization -------------------------------------------------------------16

6.3. Visualization ----------------------------------------------------------------------------------16

7. Conclusion ------------------------------------------------------------------------------------------------19

8. Appendix ------------------------------------------------------------------------------------------------- 20

1. Fig 2.1: Evolution of business strategies over time------------------------------------------------ 6
2. Fig 2.2: SWOT Analysis: YOUTUBE -------------------------------------------------------------- 7
3. Fig 2.3: Five Forces Model: YOUTUBE ----------------------------------------------------------- 8
4. Fig 2.4: Value Chain Analysis: YOUTUBE -------------------------------------------------------- 9
5. Fig 6.1: Data Visualization – Top Youtubers based on total number of subscribers---------- 16
6. Fig 6.2: Data Visualization – Top Youtubers based on total number of videos views-------- 17
7. Fig 6.3: Data Visualization –Youtubers geographical locations--------------------------------- 17
8. Fig 6.4: Data Visualization – Top 10 Youtubers based total videos views and subscribers---18
9. Fig 6.4: Data Visualization – Most viewed You Tube channel types---------------------------- 18

**Executive Summary**

YouTube has become the leading online video platform globally by leveraging a vast data analytics infrastructure and implementing strategic big data management practices. As a subsidiary of Google, YouTube captures petabytes of video viewing data daily from its billions of users, which is stored and analyzed at massive scale using Google Cloud.

Sophisticated algorithms then provide personalized recommendations that drive most user engagement, while precise audience profiling optimizes targeted ads to generate substantial revenue. Seamless interfaces ensure continuous data ingestion, and metadata taxonomy coupled with redundant infrastructure guarantee scalability and reliability.

Nonetheless, the inherent privacy and security risks arising from amassing users' personal information necessitate stringent access controls, consent mechanisms and ongoing security improvements. In aggregate, YouTube's utilization of big data analytics at every step, from content generation to monetization, exemplifies how extracting value from voluminous and variable digital information fuels innovation and industry leadership.

This report analyzes the business strategy and big data analytics practices of YouTube, the world's leading online video sharing platform. Through leveraging extensive data collection and sophisticated analytics capabilities, YouTube has established itself as the de facto online destination for video content.

The report begins with an overview of YouTube, describing its history, growth, products, policies, and revenue model. This sets the context for understanding YouTube's evolution and global dominance in the online video space.

The business strategy section then examines YouTube's strategic priorities and competitive positioning using diagnostic tools like SWOT analysis, Five Forces model and Value Chain analysis. Key findings show that YouTube utilizes its massive network effects, brand recognition and data-driven recommendations to maximize engagement and monetization opportunities.

The major focus of the report is on YouTube's big data strategy. Details are provided on how YouTube collects immense volumes of unstructured video data daily and applies advanced analytics to power features like personalized recommendations and targeted advertising. Comprehensive big data management strategies such as scalable storage, metadata tagging, and privacy controls are also analyzed.

The report further explores how YouTube's utilization of big data analytics supports strategic objectives at every stage, from content ingestion to monetization. Insights from big data fuel product innovations that deepen user experiences while precision audience profiling optimizes revenue generation.

In conclusion, the report finds that YouTube exemplifies how extracting value from voluminous digital information through balanced data-driven decision making and privacy safeguards underpins business leadership in the digital age. Continued optimization across technical, operational and policy dimensions will be vital for YouTube to address emerging challenges and opportunities.

**1. Introduction**

YouTube, a subsidiary of Google (Alphabet Inc.), is a global online video-sharing and social media platform. There are several facets to YouTube's appeal as an organization, mostly due to the platform's unmatched impact and reach. Being the world's second-biggest search engine, after its parent company, Google, YouTube has morphed into a virtual gold mine of a variety of contents, including news, entertainment, instructional and educational learning, and user-generated content. YouTube embodies innovation and adaptability through its use of data.

Our interest in YouTube is intrinsically linked to its effective and successful utilization of big data. YouTube is a great example of how volume, velocity, and variety of data have been utilized to improve user experiences, personalize content recommendations, and ultimately, efficiently monetize its platform through big data analytics. As Data Science students, relating our future to a company like YouTube seems like a natural fit, given the profound impact of big data analytics on YouTube's operations and its ability to shape the future of the platform.

**2. Business Strategy**

YouTube's main business goal is to create a constantly changing content environment through user-generated videos, which can be vlogs, tutorials, music videos, or gaming videos. The site puts a lot of emphasis on getting users involved, using tools like comments, likes, and subscriptions to build a strong community. Advertising is a big source of income. Different types of ads are shown on videos, and the people who make the material get a cut of the money that is made. By working with media companies and creators,

YouTube adds movies, documentaries, and live events to its list of material, in addition to user-generated content. With the launch of premium services like YouTube Premium and channel memberships, users can enjoy experiences without ads and access to exclusive content, while artists have more ways to make money. YouTube's complex strategy includes going global, improving technology, enforcing rules, and getting insights from data. All these things work together to keep YouTube as the world's biggest online video site.

**3. Big Data Strategy**

YouTube has built one of the largest big data platforms in the world . Immense amounts of video viewing data it collects each day from its billions of users. The company leverages Google's massive cloud infrastructure to store hundreds of petabytes of video content and metadata generated. YouTube analyzes this unparalleled trove of data through sophisticated analytics and machine learning algorithms to gain valuable insights into trending topics, popular creators, and individual user preferences.

Big data fuels YouTube's highly personalized recommendation engines that drive most of the time spent on the platform by surfacing new videos users might enjoy. The data also optimizes YouTube's targeted digital ad platforms through precise audience profiling and segmentation. In addition, YouTube's big data capabilities underpin features like automatic copyright enforcement through Content ID and provide creators with audience analytics to improve video performance. Overall, big data is at the core of YouTube's strategy, empowering every step from content to monetization.

**4. Big Data Management Strategies**

YouTube has developed robust big data management capabilities to handle the immense volumes of unstructured video data it collects daily. Advanced interfaces allow for the seamless capture and ingestion of petabytes of new video uploads and metadata into their data platforms. Google's cloud infrastructure provides massive scalable storage across global data centers to archive YouTube's historic video treasure trove. Strict governance policies and security controls protect user privacy and confidentiality. Sophisticated analytics engines extract valuable insights by analyzing patterns in viewing behaviors.

Creators and viewers can access customized analytics reports while maintaining anonymity. Real-time pipelines continuously refresh metadata to power personalized recommendations. Geographic redundancy and disaster recovery precautions protect against data loss. Metadata is organized using taxonomies and knowledge graphs for improved searchability. AI infrastructure trains self-learning models on historical data to enhance features. Through these comprehensive strategies, YouTube can effectively manage the scale of big data underpinning the world's largest online video sharing platform.

**5. Privacy and security risks**

As the world's largest hosting platform for video content and personal data, YouTube faces significant privacy and security challenges at its massive scale of operations. Billions of users share highly sensitive personal information which, if exposed in a data breach, could cause serious harm. Without robust access controls and authentication mechanisms, unauthorized actors may access private videos, comments, or backend systems. Malicious uploads pose risks of malware spreading or users falling for phishing scams. The detailed tracking of viewer behaviors raises privacy concerns if not with user consent.

Copyrighted material uploaded without permission invites legal issues. Misinformation spreads rapidly requiring moderation. Cross-border data transfers are at risk due to new regulations. Downtime or poorly designed algorithms damage user experience and trust. Insider threats also loom over vast troves of personal profiles and activity logs. Therefore, YouTube must continually fortify defenses, implement strategic controls and safeguards, and uphold transparency to adequately address these privacy and security vulnerabilities facing its colossal data ecosystem.

**1.1. Overview of YouTube**

YouTube is an online video sharing and social media platform launched on February 14th, 2005 . Its Headquarters is in San Bruno , California , United States[1]. You Tube is the second most visited website in the world after google search and is owned by Google. It has more than 2.5 billion monthly users and the visitors have a watch time of around 6 billion hours every month. The mission of YouTube is to give everyone a voice and show them the world. “We believe that everyone deserves to have a voice, and that the world is a better place when we listen, share and build community through our stories.”[2]

YouTube has grown beyond its main website since being acquired by Google, adding mobile applications, network television, and cross-platform connectivity. YouTube offers a wide variety of video content, including music videos, news, vlogs, short and feature films, songs, documentaries, teasers, live streams, and movie trailers. Most videos are produced by one person, even in cases where YouTubers and corporate sponsors work together. To reach a wider audience for their advertisements, well-known media companies including Disney, Paramount, NBCUniversal, and Warner Bros. Discovery have also launched and grown their corporate YouTube pages.

**1.2. History of YouTube**

YouTube started as a technology firm backed by venture financing, raising capital from various investors between November 2005 and April 2006. The two largest investors were Sequoia Capital and Artis Capital Management. Before YouTube, Vimeo, the first video-sharing platform, was launched in 2004 as a side project by developers from College Humor. However, YouTube gained prominence as a video-sharing platform.

YouTube's original headquarters were situated above a Japanese restaurant and a pizza place in San Mateo, California. The website www.youtube.com went live in February 2005, and the first video titled "Me at the zoo" was posted on April 23, 2005. This video, with its caption, still appears on the website today.

The startup released a public beta in May, and by November of the same year, a Nike commercial featuring Ronaldinho had accumulated one million views. By the time the site was officially launched on December 15, 2005, it was receiving eight million views daily. It's worth noting that, during this period, video clips were limited to 100 megabytes or 30 seconds in duration.

**1.3. Growth and Expansion**

Since 2005, YouTube has expanded its workforce to over 2,000 employees in offices worldwide. International growth has enabled it to serve users across 100 countries and in over 80 languages. It continues launching new offerings and acquiring other companies to fuel further growth. In recent years, YouTube has strengthened its policies and teams around safety, well-being, moderation, copyright, elections and more to better manage the platform at its massive scale. While the dominant player, YouTube faces competition from Facebook, TikTok and other emerging rivals. Maintaining an open platform while mitigating harm also presents ongoing challenges.

**1.4. You Tube’s products and policies**

**1.4.1.Products:**

The products of Tou Tubes include “You Tube Search”, which uses relevance, engagement, and quality to determine the most useful search results out of over 500 hours of new uploads every minute. Individual search results may differ based on a user's watch history and previous interactions with the platform. “YouTube recommendations” provide personalized suggestions of other videos users might enjoy, whether on the homepage or Up Next panel during viewing, to connect people with uniquely inspiring, informative, or entertaining content. The goal is to accurately predict videos viewers want to watch from billions of possibilities. “You Tube’s live streaming enables creators to host real-time livestreams and premieres that bring global viewers together to experience cultural moments, discuss trending topics, and build social communities.

**1.4.2.Policies:**

You Tube’s goal is for people to enjoy the greatest possible YouTube experience. Thus, people may remain in control of what they see and find what they are looking for with the aid of You Tube’s goods. To ensure that everyone follows the same regulations, You Tube’s policies outline what the people are allowed and are not allowed to do while they are using the site. There are certain community guidelines that needs to be followed while using YouTube .These guidelines ensure that the community can be protected from things such as harassment, spam and harmful content applied to all the content such as videos, comments, links, and thumbnails. Every content uploaded on YouTube has a copyright owned by the content creator. Every content that is posted on the site will ensure that it will follow the community guidelines and will have to align with the local Law based on the country it is being viewed.

**1.5. Revenue of YouTube**

Google purchased YouTube for $1.65 billion in October 2006.Because of Google's ownership of YouTube, the platform's revenue strategy has evolved from only selling ads to including paid content like movies and YouTube-only material. A premium subscription alternative called YouTube Premium is also available for ad-free content viewing. Additionally, YouTube gave producers permission to take part in Google's AdSense program, which aims to increase revenue for both companies. YouTube made $28.8 billion in revenue from advertising in 2021, a $9 billion increase over the previous year. In 2022, YouTube announced $29.2 billion in revenue.[1]

**1.6. Challenges and Progress**

YouTube has faced immense challenges in scaling its platform responsibly as it grew to host hundreds of hours of uploads per minute. Early on, rampant copyright infringement threatened its business model until it developed the Content ID system. Policing hate speech, extremism, and other harmful or inappropriate content across all areas also proved tremendously difficult given YouTube's massive volume.

The company has invested heavily in both AI technologies and human review teams to detect policy violations better and more quickly. Other issues like privacy concerns, platform manipulation attempts, child safety risks, and monetizing content in sensitive areas challenged YouTube's role and controls. It has addressed these through new community guidelines, privacy tools, advertising policies, and specialized features or spaces like YouTube Kids. Continuous adaptation is still needed as bad actors seek new ways to abuse open platforms and as YouTube's global influence evolves. Overall, proactive policy work, technological solutions, and open dialog have helped YouTube make progress oversight coming challenges of scale.

With unparalleled reach and resources, YouTube is well-positioned for continued leadership in online video. Its next stages of innovation and policy reform will further impact global media and digital communities.

**2.1 Business Strategy of YouTube**

YouTube has pursued an aggressive growth strategy to establish itself as the largest online video platform globally. It focused initially on accumulating a vast user base and content library that fuel each other through strong network effects. This allowed YouTube to monetize its reach through a primary revenue model of placing advertisements against free content videos. Additionally, YouTube has expanded its offerings with premium subscription services like YouTube Premium, Music, and TV to further diversify monetization.

A key strategic priority for YouTube has been investing heavily in sophisticated recommendation algorithms and personalization techniques to maximize viewer engagement and time spent on the site. This also benefits YouTube's creator community, as popular uploads incentivized by YouTube's revenue sharing help drive users back to the platform. Looking ahead, YouTube will continue adapting to evolving consumer trends through new content formats while also making safety and moderation improvements required to sustain its open and influential platform amid international expansion.

The Business strategies of YouTube can be further analyzed and explained by diagnostic tools. These are introduced by Michael Porter .There are three best-known frameworks for analyzing the competitiveness of firms:

* 1. **SWOT Analysis:**
* Introduced in 1960s .
* SWOT analysis is a strategic planning technique that is useful to identify strengths, weaknesses, opportunities, and threats regarding competition. This analysis has been widely used since the 1960s.[4]
  1. **Five Forces model:**
* Introduced in 1980.
* It analyzes five forces around a firm, such as Supplier, Buyer, New Entrants, Substitutes, and Industry Rivalry. Compared to SWOT analysis, the model focuses on external environment.[4]
  1. **Value Chain Analysis.**
* Introduced in 1985
* A Value Chain consists of a set of activities, such as primary and support activities that a firm operates. By evaluating a firm’s competitiveness in the value chain, the firm can develop a strategic plan.[4]

Value Chain   
Analysis



1985

Five Forces  
Model



1980

SWOT   
Analysis

1960s

Fig 2.1: Evolution of business strategies over time.

**2.1. SWOT Analysis:**

|  |  |
| --- | --- |
| **Strengths 👍**   * Market leader position with over 2 billion users * Deep pockets from parent company Alphabet * Powerful brand recognition worldwide * Massive repository of user-uploaded content * Strong AI capabilities for recommendations | Weakness 👎   * Dependency on advertising revenue model * Difficulty in monetizing long-form content * Scalability challenges for content moderation * Potential over-reliance on Google services * Privacy and data usage concerns persist   SWOT ANALYSIS |
| * Expand into new geographies and   regions.   * Grow new monetization streams like premium services. * Advance video formats like live streaming * Diversify into new content verticals. * Improve connected experiences across devices.   **Opportunities ✌** | * Increased competition from rivals like TikTok, Vimeo, Dailymotion, Flicker, IGTV * Stricter regulations around content and privacy * Declines in advertising spending during downturns * Dependence on retaining top creators. * Potential for brand or reputation damage * Technology and IP theft risks remain.   **Threats🖐** |

Fig 2.2: SWOT Analysis: YOUTUBE

**2.1.1. Strategic Actions:**

YouTube is well positioned to build on its strengths by further leveraging its massive user base, brand recognition, and AI capabilities. Increasing investment in machine learning can help supercharge its core recommendation and personalization engines. Expanding integrations with Google's products and services capitalizes on existing synergies in areas like search, advertising, and cloud infrastructure. To address weaknesses, YouTube should diversify its revenue sources beyond just advertising. Growing subscription offerings like YouTube Premium provides more stable income. Automating content review using AI can help scale moderation to the platform's immense size. Data privacy also requires ongoing reform to build greater trust.

Pursuing strategic opportunities, targeted acquisitions in adjacent spaces like livestreaming can accelerate YouTube's capabilities. Growing subscribers internationally through localized initiatives in emerging markets sets the stage for further uptake. Monetizing emerging video formats better incentivizes top creator talent. YouTube faces threats from competitors seeking to capture younger demographics or those prioritizing privacy. Proactive partnerships with policymakers can help shape regulation. Transparency into processes may also mitigate brand risks from copyright issues or harmful uploads not detected. A diversified business model could make YouTube less dependent on any single risk factor.

Overall, leveraging existing advantages around content, users and technology paired with initiatives to address weaknesses and new opportunities positions YouTube favorably for continued leadership. Proactively mitigating emerging threats will be important to sustaining its platform role and business model over time.

**2.2. Five Force Model:**

|  |  |  |
| --- | --- | --- |
| **Five forces**  **Analysis** |  | **Threat of new entry**   * High barrier to entry due to huge infrastructure and moderation costs. * Strong network effects make it difficult for new players to compete. * However, niche players can enter specific content categories. |
|  | **Threat of substitution**   * Services like TikTok capturing short video, Facebook Reels for entertainment. * Streaming platforms for long-form videos and movies. * Gaming/social platforms for livestreaming. * Substitutes increasing for certain use cases. |
|  | **Supplier power**   * Large, fragmented base of millions of creators globally. * Some large creators can negotiate better revenue terms. * Overall bargaining power is moderate due to options like monetizing elsewhere. |
|  | **Buyer power**   * Viewers have high bargaining power due to easily switching platforms. * Advertisers also shift budgets to platforms optimizing ROI. * YouTube needs to continue satisfying both for business model. |
|  | **Competitive rivalry**   * Direct competition from Facebook, TikTok, Netflix is intense. * Indirect competition also from substitutes. * Scale advantages provide some protection for now. |

Fig 2.3: Five Forces Model: YOUTUBE

|  |  |
| --- | --- |
| * Research groups in AI/ML, cloud infrastructure, etc. * Platform experimentation via A/B and multivariate testing * App development and maintenance across platforms * Security research for threats like malware, hacking   Secondary Activities  Procurement    Human resource Management    Technology Development    Firm Infrastructure     * Strategic supply chain partners for hardware/software * Outsourced services for content review at scale * Technology licenses, subscriptions, professional services * Facility management of offices, data centers worldwide * Global talent acquisition strategy and employer branding * Learning and development programs for all employees * Performance management and career growth processes * Compensation, benefits, and retention best practices | Operations   * Global content delivery infrastructure of servers and CDNs * Ad placement systems for pre-roll, display, overlay ads * Revenue calculation and allocation to eligible creators * Systems for content ID, strikes, and moderation workflows.   Customer Service   * Technical support via phone, email for account/streaming issues * Copyright dispute management and resolution workflows * Feedback systems to report policy violations. * Live chat and community forums for creator collaboration   Inbound Logistics   * User interface for video uploads and channel management * Metadata extraction and analysis from uploaded videos * Initial content review and filtering processes * Cataloguing and organizing videos in internal library.   Marketing and Sales   * Search engine promotion and optimization of YouTube * Display advertising on YouTube and across Google * Direct sales team and YouTube Partners program * Programmatic real-time bidding sales through DV360   Outbound Logistics   * Live streaming and on-demand video delivery capabilities * Video playback across web, mobile, smart TV platforms * Embedding and integration tools for publishers/website * APIs for third party integrations |

* 1. **Value Chain Analysis:**

Margin

Margin

* Legal, finance, policy and other centralized teams
* owned data centers and third-party cloud infrastructure
* Systems for analytics, metrics, and business intelligence

Primary Activities

Fig 2.4: Value Chain Analysis: YOUTUBE

**3.1. Big Data Strategy**

**3.1.1. Data Monetization**

The method of using data to make money is called "data monetization." Direct or external data monetization includes selling data to third parties directly or through a broker, sharing data to get better business terms and conditions, and providing information services or goods.[5]

Investing in a company's data gathering can help it make more money. Good data monetization plans make sure that businesses get the most value from their data, both inside and outside the company. They can sell the data to outside parties to make more money, cut costs within the company, and make the most of possibilities for the company.

There are several strategies that can be implemented to achieve data monetization.

**3.2. Strategy 1: Keeping data Proprietary.**

There are three ways of keeping the data proprietary based on the strengths of organizational capabilities, whether to use the data for themselves or not.

3.2.1. YouTube Strategies for keeping data proprietary:

YouTube, being a part of alphabet Inc has strict data privacy and protection. Being an owner of data always comes with a greater responsibility on how the data is being collected , how it is used and how secure the data is. Strong data management strategies are required to ensure data governance , data access controls and secured data storage. Continuous monitoring of any data breaches is required to minimize any impact on proprietary data.

YouTube likely considers its recommendation algorithms and other proprietary models as valuable intellectual property. Protecting these algorithms from reverse engineering or unauthorized access is crucial to maintaining a competitive edge.

Additionally, YouTube uses its proprietary data for internal operations to enhance its platform, improve user experience, and personalize content recommendations. For example, the platform can leverage user behavior data to refine its recommendation algorithms, ensuring users are more likely to engage with suggested content.

**3.3. Strategy 2: Sharing data with partners.**

The concept of "sharing data with partners" as it pertains to big data strategies for monetization involves the act of working in conjunction with external entities, including other organizations or businesses, to exchange or grant access to data to generate revenue and achieve mutual benefits. This may constitute a pivotal element of an organization's strategy for monetizing big data. Several methods exist for incorporating data exchange with partners into a broader monetization strategy.

There are several ways in which You tube shares data with partners.

3.3.1. YouTube Partner Program (YPP):

YouTube Partner Program participants are granted access to a multitude of analytics and data pertaining to the videos they produce. This consists of data pertaining to viewer demographics, duration of viewing, and metrics of user engagement. This information enables creators to gain a deeper understanding of their audience and optimize their content.[6]

3.3.2. Ad Revenue Sharing:

YouTube distributes a portion of advertising revenue to content creators via the Partner Program. A portion of the revenue generated from advertisements that are displayed on their recordings is distributed to the creators. This model of content distribution empowers creators to generate captivating and superior material, which is advantageous for creators as well as YouTube.

3.3.3. YouTube Analytics:

YouTube provides channel proprietors and content creators with an extensive analytics platform. This tool provides evaluations of video efficacy, traffic sources, and viewer demographics. This information can be utilized by creators to customize their content strategy and increase audience engagement.

3.3.4. YouTube API Services:

Application Programming Interfaces (APIs) provided by YouTube enable authorized third-party developers to programmatically access specific YouTube data. Typically, this access is authorized via OAuth tokens and API keys, and it is governed by the terms of service for YouTube's API.

3.3.5. Content ID System

Copyright holders can manage and govern their content on YouTube via the Content ID system. Content ID enables copyright holders, including movie studios and record labels, to identify, administer, and monetize their YouTube content.

3.3.6. YouTube for Content Owners:

YouTube for Content Owners is a suite of features and tools intended for significant content creators and media organizations. It offers sophisticated capabilities for content management, monetization, and reporting. This initiative entails the exchange of information and cooperation between YouTube and the content proprietors who are participating.

3.3.7. Analytics for YouTube Artists:

YouTube analytics for artists may grant access to data and insights to musicians and artists. This software application offers customized analytics for musicians, enabling them to gain insights into their fan base and enhance the quality of their music-related materials.

**3.4. Strategy 3: Selling Data Products**

YouTube does not directly sell user data for monetization purposes. YouTube generates revenue primarily through advertising, where advertisers can target specific audiences based on demographic and user engagement data.

YouTube collects user data to personalize user experiences, improve content recommendations, and serve relevant ads. This data includes information such as user demographics, viewing history, interactions with videos, and engagement metrics. However, YouTube does not sell personal information to third parties for their marketing purposes.

YouTube is subject to various privacy laws and regulations, such as the General Data Protection Regulation (GDPR) in the European Union and the California Consumer Privacy Act (CCPA) in the United States. These regulations impose restrictions on the collection, use, and sharing of user data and require companies to provide transparency and control to users regarding their data.[7]

**3.5. Strategy 4: Making the data freely available.**

You Tube does not make user data freely available to the public. YouTube has privacy policies and terms of service in place that outline how user data is collected, used, and shared. While YouTube may use user data for internal purposes, such as improving their platform and personalizing user experiences, they do not openly provide unrestricted access to user data.

YouTube is committed to protecting user privacy and complying with applicable privacy laws and regulations. They have implemented measures to safeguard user data and provide transparency to users regarding the collection and use of their information. Users can manage their privacy settings and control the visibility of their personal information on the platform.

It's important to note that YouTube is subject to privacy laws and regulations, which impose restrictions on the collection and disclosure of user data. These laws prioritize user privacy and require companies to obtain user consent and provide clear information about how data is used.

**3.6. YouTube’s Monetization Policies:[8]**

It is critical that any YouTube channel that generates revenue adheres to the platform's monetization policies. The following policies are included:

3.6.1. Community Guidelines: YouTube's Community Guidelines, which govern both individual videos and the channel, must be followed by content creators. Infractions of these guidelines render content ineligible for monetization and potentially subject to removal.

3.6.2. AdSense Program Policies: Creators are required to adhere to YouTube's Terms of Service and the AdSense program policies. Advertisers can generate revenue through AdSense. Monetization on channels with repetitive content, where the videos are extremely similar, may be removed. Monetized content ought to provide an intriguing and captivating experience for viewers.

3.6.3. Reused Content: Monetization may be revoked from channels that repurpose the content of others without substantially augmenting it with original commentary or educational merit. However, it may be permitted if the content is altered in a humorous or reflective manner.

3.6.4. Quality Principles for Kids and Family Content: YouTube's quality principles for family and children's content stipulate that channel featuring "made for kids" material are obligated to adhere to these principles. Channels that exclusively feature substandard "made for kids" material run the risk of being removed from the YouTube Partner Program.

3.6.5. Advertiser-Friendly Content Guidelines: The advertiser-friendly content guidelines must be adhered to by all content that is monetized with advertisements. These principles guarantee that the material is appropriate for both advertisers and viewers.

3.6.6. Fan Funding: To generate income from fan funding functionalities, initial users are required to approve the Commerce Product Module (CPM) and adhere to the monetization policies of Commerce Products.

* 1. **Big Data Management Strategies**

The measure, capture, and fusion of big data for YouTube involves gathering and integrating data from various sources to gain insights and improve decision-making. Measurement entails directly monitoring key metrics through YouTube Analytics and indirectly collecting qualitative feedback through surveys and interviews.

Data capture involves both passive methods, such as extracting metadata from YouTube's API and analyzing log files, and active methods like engaging with users for specific data points. Data fusion focuses on integrating data from multiple sources, cleansing, and transforming it for consistency, and applying advanced analytics techniques to uncover patterns and trends.

* 1. **Data Measures**

Measurement involves tracking and analyzing key metrics to assess the performance and effectiveness of YouTube channels and videos.

4.2.1 Direct Measure

YouTube provides its own analytics platform called YouTube Analytics, which offers valuable insights into metrics like views, watch time, engagement, and subscriber growth. This direct measurement helps content creators and organizations understand the reach and impact of their content on the platform.

4.2.2. Indirect Measure

Indirect measurement methods, such as surveys, interviews, and focus groups, can provide qualitative feedback from YouTube users, viewers, and content creators. This qualitative data offers deeper insights into user preferences, satisfaction, and behavior.

* 1. **Data Capture**

Data capture involves collecting relevant data from various sources.

4.3.1. Passive data capture

Passive data capture refers to automatically gathering data without user intervention. This includes extracting metadata and performance metrics from YouTube's API, analyzing log files, and utilizing web scraping techniques to gather information from public sources. Data can be collected on video views, likes, comments, user demographics, and content categories.

* + 1. Active data capture

Active data capture , on the other hand, involves engaging with users through interactive elements such as polls, comments, and surveys. This enables organizations to collect specific data points and gather insights directly from users.

* 1. **Data Fusion**

Data fusion is the process of integrating and combining data from multiple sources to create a comprehensive and unified view. YouTube involves integrating data from various channels such as YouTube itself, social media platforms, website analytics, customer relationship management (CRM) systems, and third-party data providers.

Data fusion allows organizations to have a holistic understanding of their audience, content performance, and user behavior. By combining and analyzing data from different sources, patterns, trends, and correlations can be identified, enabling organizations to make informed decisions and optimize their YouTube strategies.

**5.1. Privacy and security risks**

The rise of big data has revolutionized the way organizations operate, enabling them to collect, store, and analyze vast amounts of information to gain insights, improve decision-making, and enhance customer experiences. However, the increasing volume and complexity of big data have also introduced significant privacy and security challenges.

For an organization like YouTube which manages extensive volumes of user data encompassing personal details, viewing history, and search behaviors, these challenges become especially pronounced. YouTube, as a leading video-sharing platform, recognizes the significance of safeguarding user data. This section outlines six key guidelines to minimize data breaches and enhance user trust.

**5.2. Privacy Issues**

* + 1. Data Collection

YouTube collects a wide range of user data, including personally identifiable information (PII) such as names, email addresses, and IP addresses. This raises concerns about the potential for data misuse and unauthorized access. In 2019, YouTube faced a child privacy case for illegally collecting personal information from children under 13 years of age without parents’ consent.

* + 1. Data Sharing

YouTube shares user data with third-party partners for advertising and analytics purposes. This raises concerns about the transparency of data sharing practices and the potential for data misuse by third parties.

* + 1. Data Retention

YouTube retains user data for extended periods, even after users have deleted their accounts. This raises concerns about the potential for data erosion and the difficulty for users to control their digital footprints.

* 1. **Security Risks[9]**
     1. Data Breaches

YouTube alongside its parent company, Google has faced several high-profile data breaches in the past, exposing millions of user accounts to unauthorized access. This highlights the vulnerability of big data systems to cyberattacks.

* + 1. Data Misuse

Insider threats and unauthorized access to user data pose significant risks. This raises concerns about the potential for data misuse, such as identity theft, fraud, and targeted advertising.

* + 1. Data Exploitation

The extensive data gathered by YouTube could be exploited to disseminate misinformation, sway public opinion, or disrupt elections. This gives rise to concerns about the possible misuse of big data for malicious purposes.

* 1. **Mitigating Risks: Minimizing Risks for Data Monetization**

  To minimize the risks associated with big data breaches, YouTube can implement several strategies:

* + 1. Affirm Ownership and Control of User Data:

YouTube should maintain clear and transparent privacy policies that explicitly detail how user data is collected, processed, and shared. Users should be informed about the types of data being collected and the purposes for which it is being used. By reinforcing users' control over their data, this establishes a foundational principle of respecting and protecting user information.

* + 1. Aggressive Anonymization Approach

Before using big data for monetization, YouTube should employ techniques such as anonymization and aggregation when sharing or transferring aggregate user data to third parties for purposes like advertising or analytics, to minimize the risk of exposing individual user identities. This helps in ensuring valuable insights can be derived from data without compromising individual user information.

* + 1. Clarify User Participation and Control Options

YouTube needs to communicate transparently about how user data powers platform features and offerings. It should also provide simple, conspicuous tools for users to view, delete or correct their data and control how it is used via granular consent options. This upholds user understanding and participation.

* + 1. Best Practices in Data Storage and Transmission

Strong encryption, access controls, auditing, and proactive measures to identify and mitigate potential vulnerabilities should protect any sensitive user data in transit and at rest (data in storage), whether stored on YouTube's own systems or transferred to others. Technologies like hashing and tokenization can also be utilized to reduce risks. By ensuring the highest standards in data security,  regular security reviews and testing. YouTube ensures data protections stay robust, minimizes the risk of unauthorized access,  and data breaches.

* + 1. Community Sentiment Tracking and Policy Adaptation

User forums, comments and surveys, community sentiment can provide early warnings about data handling concerns. YouTube must demonstrate responsiveness by quickly addressing issues that contradict reasonable user expectations regarding their personal information. YouTube should be able to promptly retract policies that violate user expectations. By being responsive to user concerns, YouTube will be able to maintain a positive and trusted relationship with its user base.

* + 1. Establish a Chief Data Officer(CDO)

A dedicated Chief Data Officer (CDO) for YouTube would assume comprehensive responsibility for overseeing the entirety of YouTube's data governance program with a primary focus on user interests. This pivotal role encompasses operational oversight, policy development, compliance monitoring, risk analysis, and optimization of data usage while prioritizing the minimization of privacy exposures and preventing unauthorized access. Although Google, as the parent company of YouTube, has a CDO (who recently resigned in September 2023 yet to be replaced), we advocate for YouTube to have its own dedicated CDO to specifically manage and safeguard the data governance of the platform.

Adopting these comprehensive strategies guided by clear ownership, transparency, controls, and accountability principles can help YouTube minimize data security risks and breaches, protect user privacy on its platform and foster a secure digital environment. Continuous improvement is also needed to stay ahead of emerging threats.[10]

**6.1. Data Visualizations**

The utilization of data visualization is vital for the effective comprehension and communication of complex information. Data visualization utilizes various visual elements such as charts, diagrams, maps, and more to represent data. This functionality empowers users to comprehend patterns, trends, and insights that might be challenging to ascertain solely from raw data. By improving the visibility and comprehension of data, it facilitates the detection of correlations, outliers, and relationships.

In addition to facilitating the presentation of data in a compelling and intuitive fashion, data visualization facilitates storytelling by providing stakeholders with actionable insights and the ability to make informed decisions. Furthermore, the utilization of dynamic and interactive data visualizations grants users the ability to investigate and engage with the data, thereby facilitating the revelation of more profound understandings and the introduction of fresh viewpoints. In general, data visualization serves as a potent instrument that augments comprehension of data, promotes efficient discourse, and propels decision-making processes grounded in data.

**6.2. Tableau for Data Visualizations**

Tableau is a powerful and versatile approach to explore, analyze, and present data in a visually appealing and interactive manner. Tableau provides a user-friendly interface that enables users to easily connect to multiple data sources, transform raw data into meaningful visualizations, and create interactive dashboards and reports. With its drag-and-drop functionality and extensive library of visualization options, Tableau allows users to customize and design visual representations that best convey their insights and narratives.

**6.3. Visualizations**

Data set : Global YouTube Statistics[11]

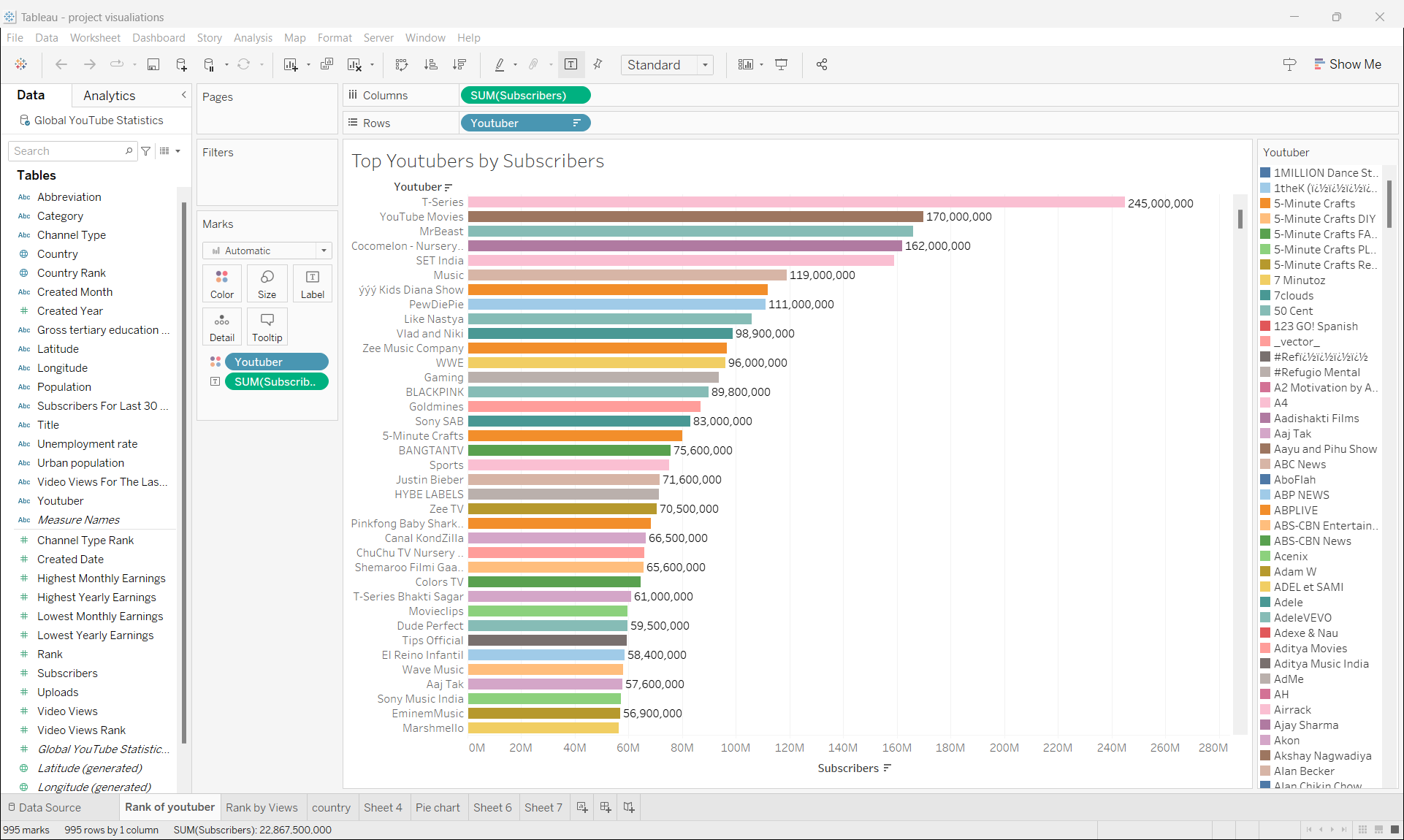


Fig 6.1: Data Visualization – Top Youtubers based on total number of subscribers.

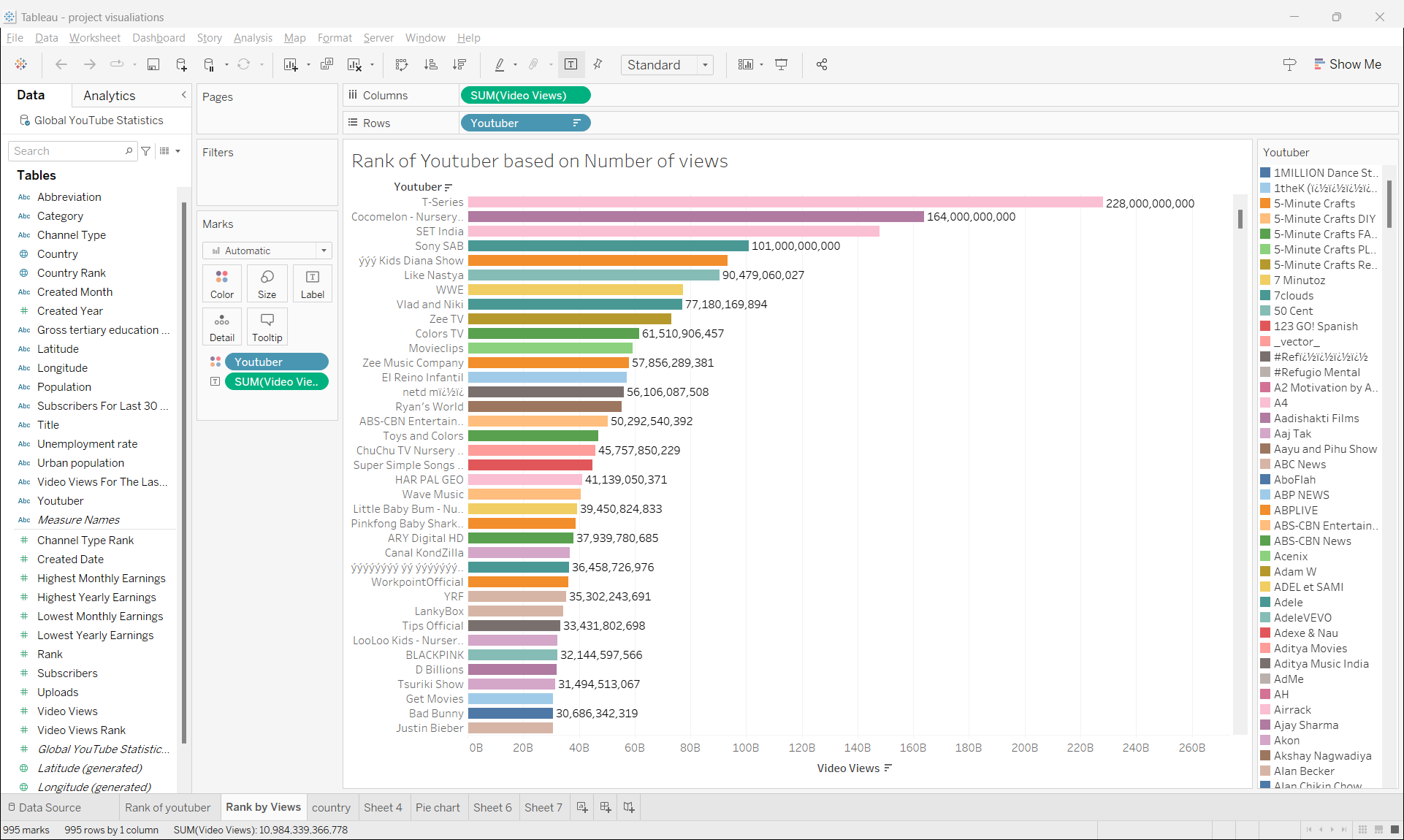


Fig 6.2: Data Visualization – Top Youtubers based on total number of videos views

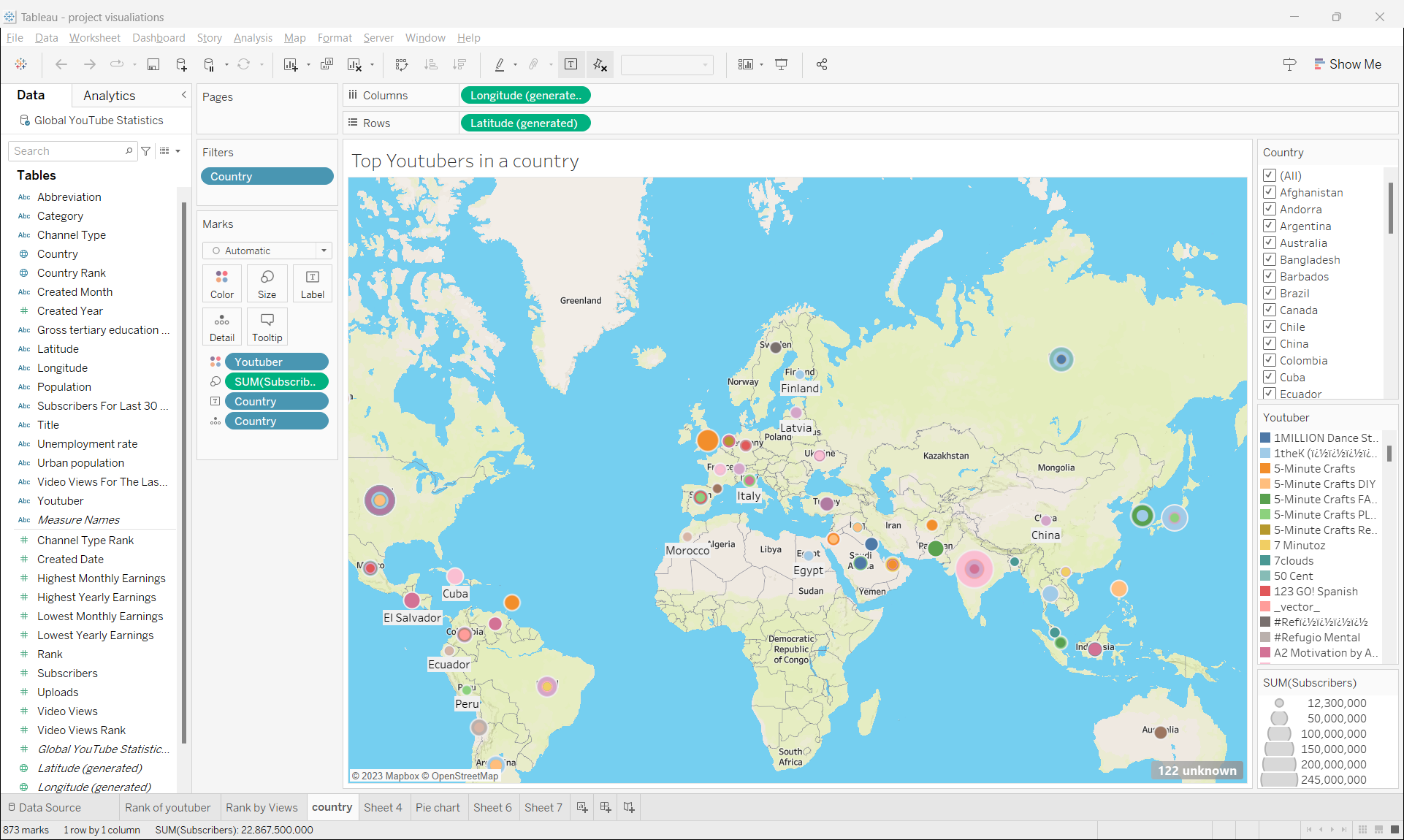


Fig 6.3: Data Visualization –Youtubers geographical locations

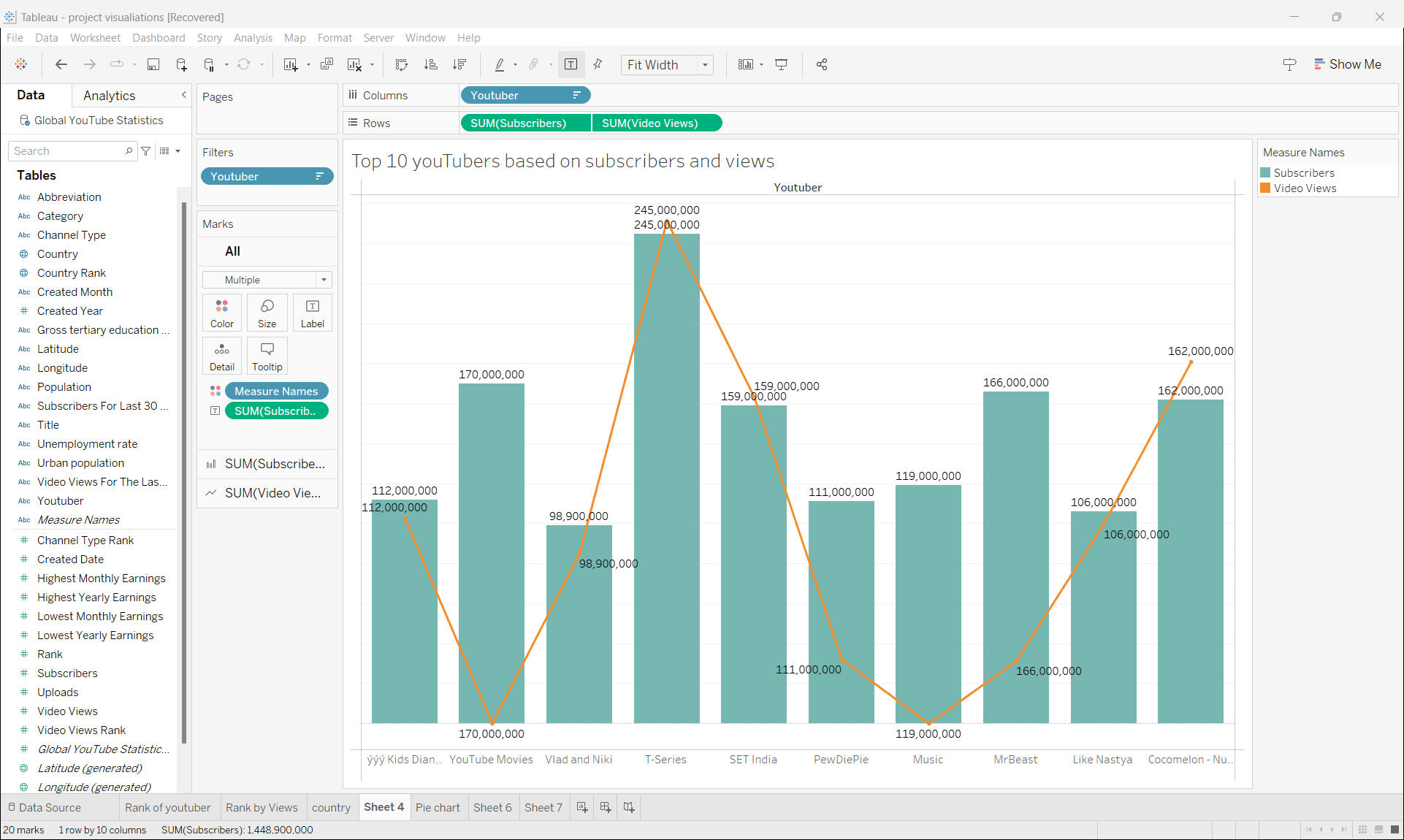


Fig 6.4: Data Visualization – Top 10 Youtubers based total videos views and subscribers.

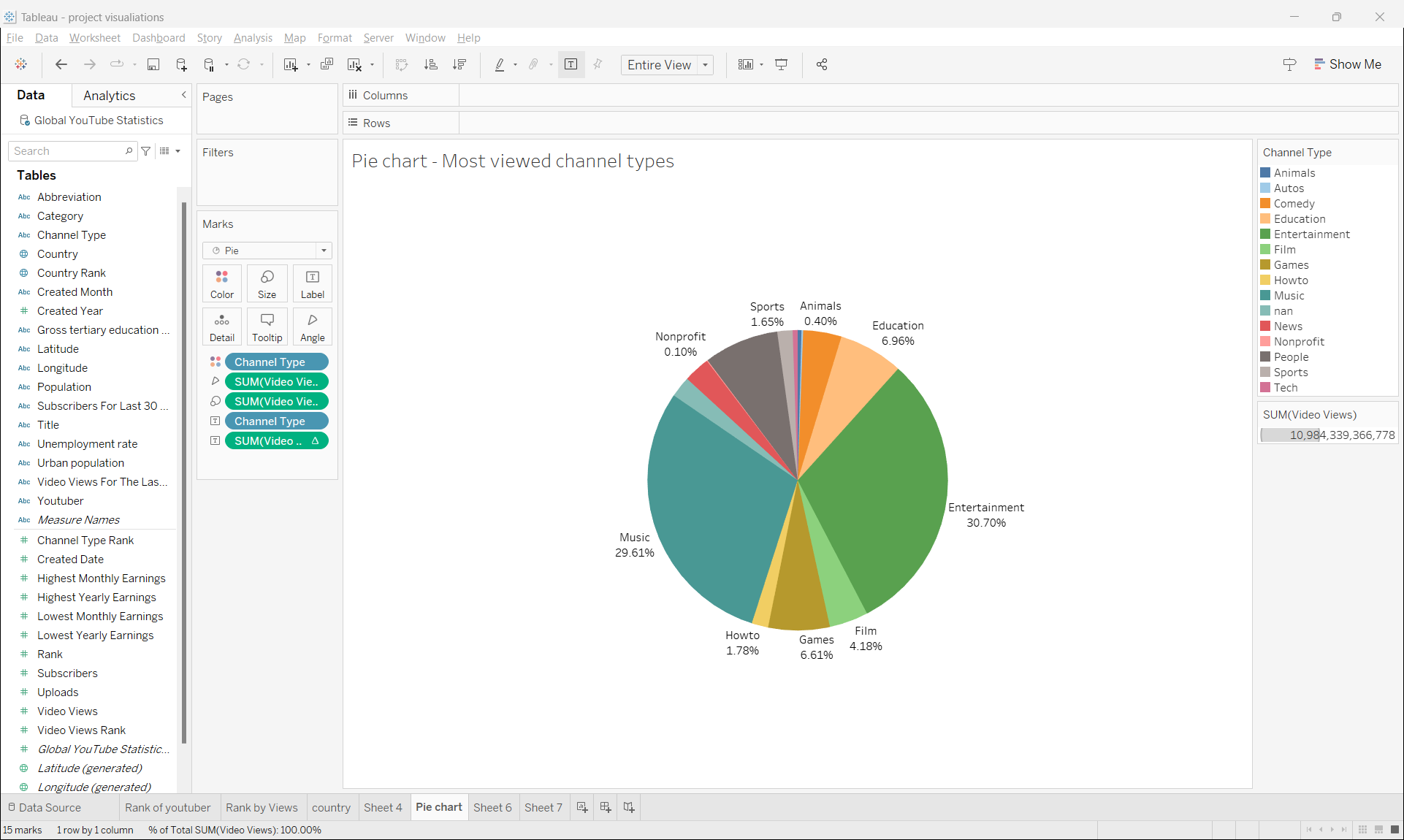


Fig 6.4: Data Visualization – Most viewed You Tube channel types

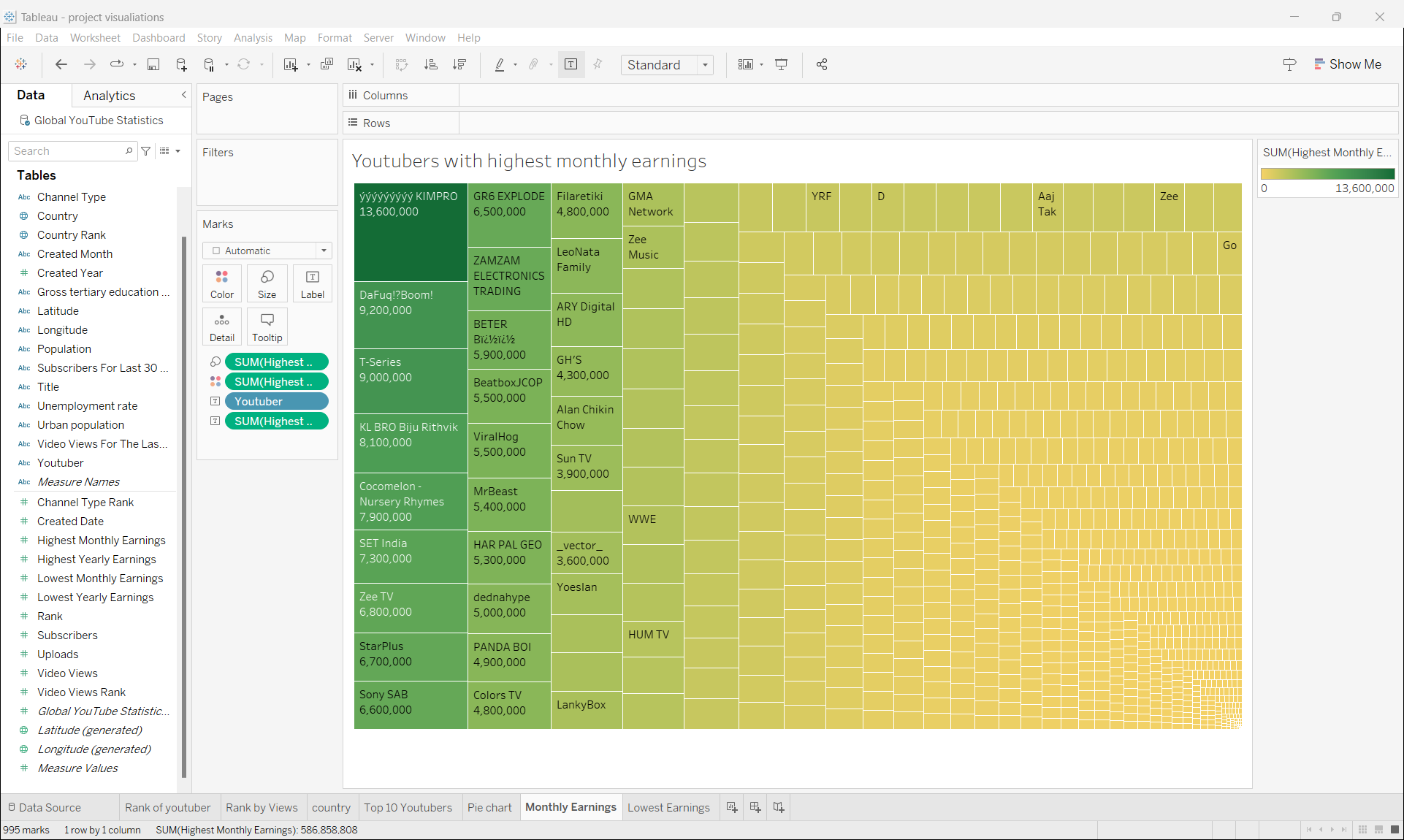
****

Fig 6.4: Data Visualization – Youtubers with highest monthly earnings

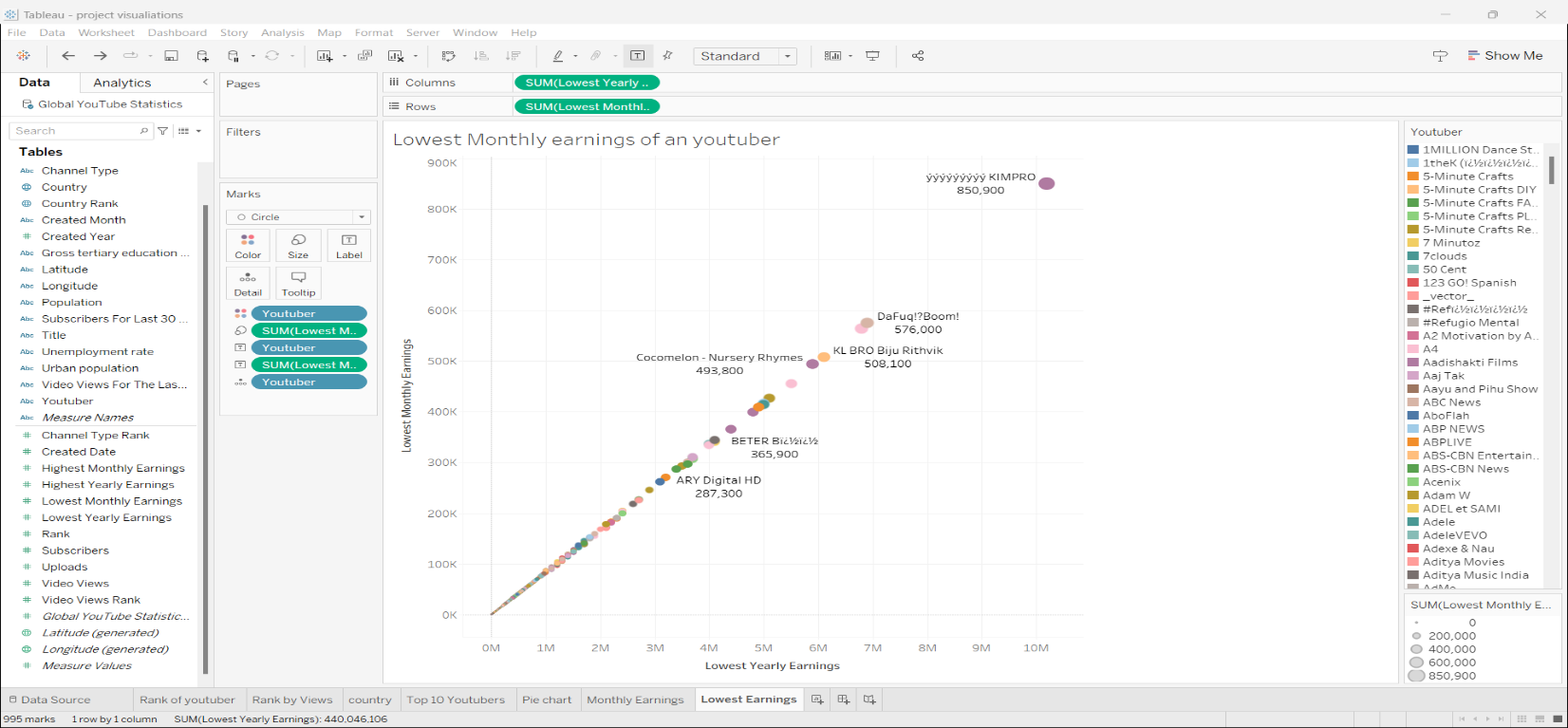


Fig 6.4: Data Visualization – Youtubers with lowest monthly earnings

**7.0. Conclusion**

YouTube has established itself as the leading online video platform globally, leveraging its vast data analytics infrastructure and implementing strategic big data management practices. With billions of users and petabytes of video viewing data generated daily, YouTube has harnessed the power of big data to enhance user experiences, personalize content recommendations, and drive substantial revenue through targeted advertising.

YouTube's business strategy revolves around creating a dynamic content environment through user-generated videos, fostering user engagement, and building a strong community. Advertising plays a significant role in generating income, and YouTube has expanded its offerings by collaborating with media companies and creators, adding movies, documentaries, and live events to its platform.

The implementation of a robust big data strategy has been instrumental in YouTube's success. The platform collects massive amounts of video viewing data and analyzes it using sophisticated analytics and machine learning algorithms. This enables YouTube to gain valuable insights into user preferences, trending topics, and popular creators, which in turn fuels highly personalized recommendation engines and optimizes targeted digital advertising.

To effectively manage the scale of big data, YouTube has developed comprehensive big data management strategies. These strategies include seamless data capture and ingestion, scalable storage infrastructure, governance policies, security controls, and real-time analytics. YouTube's big data capabilities also support features such as automatic copyright enforcement and audience analytics for creators.

The report also highlights the privacy and security risks associated with YouTube's vast data ecosystem. As a platform hosting personal data, YouTube faces challenges in protecting user privacy, preventing data breaches, combating malicious uploads, addressing copyright issues, and ensuring transparency. It is crucial for YouTube to continually strengthen its defenses, implement strategic controls, and uphold user trust.

In conclusion, YouTube's utilization of big data analytics at every step, from content generation to monetization, has propelled its innovation and industry leadership. By leveraging its data analytics infrastructure, YouTube has enhanced user experiences, personalized content recommendations, and effectively monetized its platform. While YouTube's big data capabilities have driven its success, it is imperative for the platform to address privacy and security risks to maintain user trust and continue its growth as the world's largest online video sharing platform.

**Appendix:**

[1] "YouTube." Wikipedia: The Free Encyclopedia, Wikimedia Foundation, 24 Nov. 2023, 12:00 UTC, https://en.wikipedia.org/wiki/YouTube.

[2] "YouTube": Youtube.co. https://www.youtube.com/howyoutubeworks/

[3] "5 Force Analysis template". Study Corgi: . https://studycorgi.com/free-writing-tools/porters-five-forces-template/

[4] "Immyung choi" Module 3 . Business Strategy

[5] "Data Monetization". Jaspersoft. https://www.jaspersoft.com/articles/what-is-data-monetization#:~:text=Direct%20or%20external%20data%20monetization,offering%20information%20services%20or%20products.

[6] "Monetization Policies", YouTube. /https://www.youtube.com/howyoutubeworks/policies/monetization-policies/

[7] "Policies Overview", YouTube. https://www.youtube.com/howyoutubeworks/policies/overview/

[8] "Monetization Policies", YouTube. /https://www.youtube.com/howyoutubeworks/policies/monetization-policies/

[9] Heiligenschein, M. X. (2023, October 5). "Google data breaches: Full timeline through 2023". Firewall Times. https://firewalltimes.com/google-data-breach-timeline/

[10] Walker, R. (n.d.). "From Big Data to Big Profits: Success with Data and Analytics".

[11] "Global YouTube statistics 2023". Nidula Elgiritewithana. https://www.kaggle.com/datasets/nelgiriyewithana/global-youtube-statistics-2023