BIG DATA ANALYSIS ON YOUTUBE WITH TABLEAU

**[1]JOHANES FERNANDES ANDRY, [2]*HENDY TANNADY, [3]ISABELLE IVANA LIMAWAL,
[4]GLISINA DWINOOR REMBULAN, [5]RUSTONO FARADY MARTA**

[1]Department of Information System, Universitas Bunda Mulia, Jakarta, Indonesia 14430
[2]*Department of Management Graduate Program, Institut Teknologi dan Bisnis Kalbis, Jakarta, Indonesia 13210
[3]Department of Information System, Universitas Bunda Mulia, Jakarta, Indonesia 14430
[4]Department of Industrial Engineering, Universitas Bunda Mulia, Jakarta, Indonesia 14430
[5]Master's Degree of Communication Science Department, Post-Graduate Studies, Universitas Bunda Mulia, Jakarta, Indonesia, 14430
E-mail: [1]jandry@bundamulia.ac.id, [2]*hendytannady@gmail.com, [3]isabelleivana@gmail.com, [4]grembulan@bundamulia.ac.id, [5]rmarta@bundamulia.ac.id

ABSTRACT

YouTube is the second biggest search engine on the web and a platform with features where users can post, view, comment, and link to videos. YouTube has more than a billion active users where users can see Recommended Channels, which are based on videos that users watch frequently. Furthermore, Trending is based on video trending on the number of clicks the video gets every day. This study aims to provide insight into how understanding and implementation of Big Data on YouTube. In the current era, Big Data is gaining much interest because of the opportunities and benefits felt to be unprecedented. Big data analytics can leverage business change, enhance decision-making by applying advanced analytic techniques on big data, and reveal hidden insights and valuable knowledge. To better understand the changes brought about by big data, this paper focuses on data analysis using data visualization. In analyzing YouTube trending videos, we use tableau software to help visualize the data that has been obtained. The following are some of the charts used in this analysis, such as pie charts, area charts, horizontal bars, highlight tables, treemaps, and mapping. Data Visualization can help YouTube, and its users identify which areas need to be improved, which factors affect users' satisfaction and dissatisfaction, and what to do and know to develop parts lacking and pay attention to today's users' needs. Visualized data can give a better prediction of current trending topics or videos and future growth. It can be used to influence our decision-making.

Keywords: *Big Data, YouTube, Data Analytics, Data Visualization*

1. INTRODUCTION

Social network data, audio, video in social media, and social networking sites are ample data examples. One example is YouTube. This platform was founded in 2005 and developed rapidly to become the most crucial video-sharing website on the web [1]. Following its acquisition by Google, the location has continued to grow in popularity. It is now one of the top-visited sites globally and attracts 800 million unique visitors per month [2]. Where users can upload videos, watch various videos available on the platform with various topics. This platform provides various facilities to interact with one another and answer certain video content. Using big data, users can use the facilities provided, like users can see Recommended Channels, where users can view videos recommended supported videos that users often watch. Also, there is Trending, where videos that are being watched by many users are often seen

in Trending. There are also Recommended Videos, which are videos that are recommended, supported videos that users watch. YouTube has great potential to succeed in potential consumers, media, and interest groups. With numerous users using the platform, indirectly creating new data are often processed to supply a replacement analysis result. In Internet usage, each user automatically collects large amounts of knowledge a day. This data is stored directly by the physical device. It is more feasible for giant amounts of storage to store within the cloud than a locally available hard disc. Because storing on the hard disc, the danger of losing data is bigger than storing within the cloud. Around 2016, active internet users reached 3,424,971,237 worldwide. Users of social sites or social media are increasing a day and leading to data collection. Google and YouTube have about 15 exabytes of knowledge, and Facebook and Twitter have tons of knowledge collection. The info collected is getting bigger and

larger because it has increasingly collected everywhere - sensing mobile information devices, antennas, sensory technology, software logs, cameras, microphones, frequency identification readers, and wireless sensor networks. Supported world technology, information storage capacity has doubled every 40 months since the 1980s; as of 2012. A day the info has been generated about 2.5 trillion [3].

Big data provides organizations with insights into their business processes and enables them to market their products better by employing analytical techniques like data mining [4]. The method of extracting valuable information and data from large ones is named "Data Mining." The platform uses big data technology to take care of its filing system. The quantity of knowledge generated from the start of our time until the last decade is perhaps generated every few days today [5]. Altogether probability, this trend is about to continue because of the increase in social media applications. Simple access to data through mobile devices exponentially increases the quantity of knowledge captured every second [6].

The platform has billions of individuals who connect, inform, and encourage the planet by making videos a day. Therefore, it is not surprising that YouTube now features a significant influence on the web but has severe problems. Archiving, processing, and efficient analysis of sizeable data is extremely difficult. The info generated by the billions of YouTube videos is appropriate. Consistent with statistics published by Google, YouTube has quite one billion customers. Also, as many as 33% of individuals spend their time watching YouTube because the platform has around 300 hours of video a day and may even exceed the amount published by Google [7]. Social media's exponential growth in contemporary society makes them necessary tools for communication, content creation, sharing, and business growth [8].

The authors will analyze daily statistics for trending videos where the dataset comes from Kaggle.com. To work out which videos are trending and are on the highest list this year, YouTube uses various factors, including measuring user engagement (trending dates, total videos, region, and category). The info to be analyzed is within the USA, France, Russia, India, and Japan, with up to 200 trending videos registered per day. Each region data is during a separate file. Data includes trending dates, total videos, region, and category. The authors will visualize data that has been taken from Kaggle.com.

Data visualization is the use of natural human skills to enhance processing and organization efficiency. Data visualization represents data or information during a graphic, chart, or other visual formats. Visualization can help us affect more complex information and enhance memory [9]. It communicates the relationships of the info with images. This is often vital because it allows trends and patterns to be more easily seen. With the increase of massive data, we would like to interpret increasingly larger batches of knowledge. Machine learning makes it easier to conduct analyses like predictive analysis, which may then function helpful visualizations to present. In this sense, data visualization is not only crucial for data scientists and data analysts, but all people in any career got to understand data visualization.

It is often used interchangeably with information graphics, statistical charts, and knowledge visualization. We want data visualization because that visual summary information makes it easier to spot patterns and trends than to look through the thousands of rows on a spreadsheet. That is how the human brain works. Since the aim of knowledge analysis is to realize insight, data is far more valuable when visualized. Albeit a knowledge analyst can pull insights from data without visualization, it will be tougher to speak the meaning without visualization [10]. Data visualization may be a big field with many disciplines. It is precisely due to this interdisciplinary nature that the visualization field is filled with vitality and opportunities. To research and visualize the info, the authors are going to be using Tableau Software.

Tableau is an end-to-end data analytics platform that permits us to prep, analyze, collaborate, and share our significant data insights. Tableau excels in self-service visual analysis, allowing people to ask new questions of governed big data and share those insights across the organization [11]. Tableau works closely with the leaders during this space to support whatever platform the customer chooses. Tableau allows users to seek out this value in company data and investments within the technology to get the most out of their data. Tableau can help businesses see and understand Big Data. During this time, efficient analysis of business data and their storage on appropriate devices is tedious. The YouTube videos and, therefore, the multimedia data, which are generated from it, are usually unstructured. Perfect analysis of those semi-structured and unstructured multimedia data may be a big challenge. This study aims to supply insight into how understanding and implementation of massive Data on YouTube

2.   LITERATURE REVIEW

This section will provide the foundation and support for a new insight that authors contribute. However, a literature review's main target is to summarize and synthesize others' arguments and ideas without adding original contributions.

2.1 Big Data

The term "Big Data" appeared for the primary time in 1998 on the slide deck Silicon Graphics (SGI) by John Mashey under the title "Big Data and therefore the NextWave of InfraStress." The word Big Data is very relevant from the start, as the first book to mention "Big Data" is a mining data book that appeared in 1998 by Weiss and Indrukya. For the first time, an academic paper with the inscription "Big data" in the title appeared sometime later in 2000 in a Newspaper by Diebold. The origin of the term "Big Data" is that we are currently creating a file of data every day. Big data can also be interpreted as data sets of sizes beyond ordinary software capabilities used to capture, curate, manage, and process data in considerable time [12]. Big data is somewhat unusual in that it was broadly accepted in the commercial and public space before the academic discourse has had time to catch up [13]. This may explain why most of the literature on big data has increased during the past few years.

Big data is not new, but the rapid rate of adoption in recent times may make it appear so [14]. Sensors, mobiles, and social media networks are just a couple of samples of modern digital technologies that have penetrated our daily lives. An excessive amount of digital data is being generated a day. Nowadays, an increasing number of knowledge silos are created globally, suggesting that this growth will never stop. These data are not only voluminous but also continuous, streaming, real-time, dynamic, and volatile [15]. Data have grown exponentially in "volume", variety' and "velocity" [16], [17]. This is generally described by the "3 V" of big data (volume, variety, and velocity), which are used to describe this phenomenon. Literature indicates that big data can unlock plenty of new opportunities and deliver operational and financial value [18], [19]. Internet advent has triggered a boom in information research. Companies are flooded by the wealth of knowledge that results from simple internet browsing.

2.2 YouTube

YouTube was founded in 2005 and developed the platform rapidly to become the largest video-sharing website on the internet [1]. Following its acquisition by Google, the location has continued to grow in popularity and now attracts almost half a billion unique users per month. The location allows users to upload a vast number of video clips, which may be viewed and linked to anyone. YouTube provides various facilities to permit registered users to interact with one another and share their content responses. One option is for users to post direct video responses to what they need to be viewed. Other options are afforded by the user profiles, or "channels." There are personal messaging service and a more visible commenting option on the profile, and users can prefer to store and display lists of their favorite videos.

They can also befriend other users on the site or become "fans" by subscribing to receive alerts once they post new material. The location provides sharing, rating, and commenting options on individual videos. The one that has uploaded the video can change the comment settings. Only certain users can comment and need that comments are subject to moderation before they become publicly visible or prefer to disable the comments facility altogether. Users can "rate" comments using thumbs-up and thumbs-down buttons, as they will for videos. The comments that gather the foremost significant number of positive votes are displayed above the most bulk of comments. The comments facility was intended to be "a section of text for users to supply information associated with a video" and where users could express their opinion on the video. However, in practice, this is often not always what happens. Technology provides affordances but is not prescriptive; they have often put to uses aside from that that it had been initially intended. This is often the case with the comments facility: a user can write things within the comments box that are not relevant to the video [2].

2.3 Data Visualization

With the advent of technology, YouTube ought to deal with numerous rows of data. To make a better judgment, the data must be analyzed. When a large amount of data is displayed to an audience dealing with it, daunting and minor trends might always skip the analyst's eye. Data visualization is a process wherein the analyzed data is represented visually in graphs and pictures, both attractive and functional. Data is one among the foremost powerful and – often overlooked – tools to speak a message. However, those numbers alone would not necessarily make the desired impact if the audience cannot interpret them. That is why data visualization is the key [20]. Interactive charts and figures help understand the varied outcomes possible in each situation. This type of visualization helps people add an extra dimension

to their presentations, and cohesively display essential statistics.

Tableau is a business intelligence and data visualization tool that helps chart out the data onto graphs, maps, etc. and break it down into several components on a dashboard for real-time analysis. With the help of Tableau, YouTube can explore and analyses data visually. The program also offers the option to create a dynamic presentation on the app itself; post data visualization sheets have been generated. By presenting the data in a visually engaging way, one can enhance one's own and the audience's experience. Tableau quickly creates graphs and visualizations that include generic filters, using graphs as filters; tooltips (hover over pop-ups); and embedded webpages. A Tableau data extract will take a snapshot of our data and put it into the Tableau format that makes it easy for the software to compute. This comes in handy when one has a large data set. Data visualization is not just the process of creating a few charts. It is about helping the audience understand the significance of the data that is being presented. Tableau makes the data easier to comprehend, making statistical relationships, patterns, and critical trends more exposed and recognized [21].

3. METHODOLOGY

In this section, we will discuss the step-by-step process, which helped achieve this paper's completion.
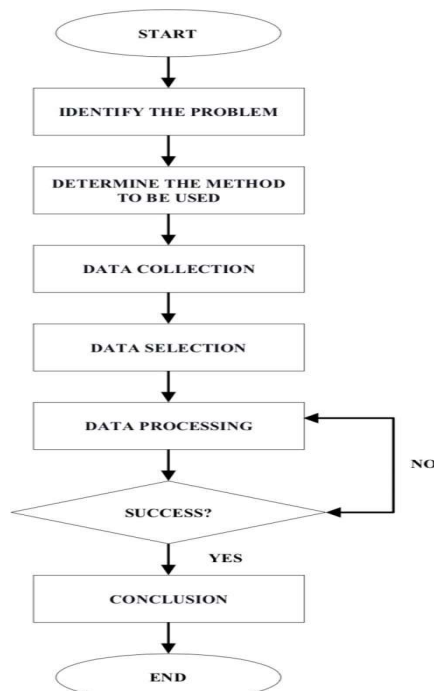

Figure 1: Research Flowchart[22]

a. Identify The Problem

The initial stage of research is to identify existing problems. The authors want to analyze and visualize the Trending YouTube Videos dataset, where the data are taken in the USA, Great Britain, South Korea, and Japan.

b. Determine The Method To Be Used

After the problem has been identified, the authors then determine the method that will be used. The method that will be used is data visualization. Data visualization visually represents quantitative data with or without axes in schematic or diagrammatic forms, e.g., Table, Line chart, Pie chart, Histogram, Scatter plot, etc. [23].

c. Data Collection

After determining the appropriate method, the authors then collected the data. The authors used the Trending YouTube Videos dataset, where the dataset is taken from Kaggle.com.

d. Data Selection

Next is the data selection stage. At this stage, the data will be selected. The input variables include video titles, channel titles, published times, tags, views, likes and dislikes, descriptions, and the number of comments.

e. Data processing

After all the necessary data have been selected, the next stage is data processing. At this stage, the data will be processed using Tableau, where the data will be transformed or change the data attribute values into the appropriate data so that the data can be processed using data visualization. The complete dataset will be obtained, which is used to process to the next stage [24].

f. Testing Data

At this stage, the data will be tested where data testing is a stage that determines whether the test made is suitable for use or not. If the test has produced an output that suits the authors' needs, the authors can proceed to the next stage. However, if the test has not produced an output that matches the needs, we have to return to the data processing stage; if it has entered the repeat or failure stage, it is necessary to solve it by analyzing data processing [25].

g. Conclusion

This stage is the last stage where the authors draw conclusions that refer to the formulation of the problem and research objectives based on the test results. Suggestions are used to develop further research and are included to improve the quality of research.

4. RESULT AND ANALYSIS

Before using the Tableau application, make sure the data to be processed first. The data chosen is data from YouTube Trending Videos in the USA, France, Russia, India, and Japan from 2017 to 2018. When all data is valid, data processing can be done using the Tableau application. The first step is to connect the data. Open Tableau Desktop, and after that, we will choose a select connector, which is what data we will use. Because the data taken from Kaggle is CSV, we choose a text file; after that, we will be directed to choose what data to use. After the data has been selected, we will enter the Data Source. In the data source, we can see the dataset that has been taken from Kaggle because we are using multiple data sources; we can see them all listed in the data source. After that, we can combine the data we want to retrieve, namely the USA, France, Russia, India, and Japan.

After combining the data, use a data interpreter to help clean up the CSV data that has been retrieved. After that, we can select some of the data needed for data visualization. Data can be classified into two parts, namely dimension or measure. These data are data that have been selected and processed to produce the data visualization previously described. In Table 1, we can see an explanation of the data that we used. After the data has been selected, the data will be processed.

Table 1: Type of Data

| Category	A category is a group, a group of various types of videos found on YouTube. These categories consist of Autos & Vehicles, Comedy, Entertainment, Education, Film & Animation, Gaming, How to & Style, Movies, Music, Nonprofits & Activism, News & Politics, People & Blogs, Pets & Animals, Science & Technology, Trailers, Shows, Sports, Travel & Events.
Region	The region is a part/territory of a specific area such as the USA, France, Russia, India, and Japan.
Trending Date	Trending Date is the date a video is being considered attractive by many YouTube viewers/users.
Total Videos	The total number of videos uploaded to YouTube by users.
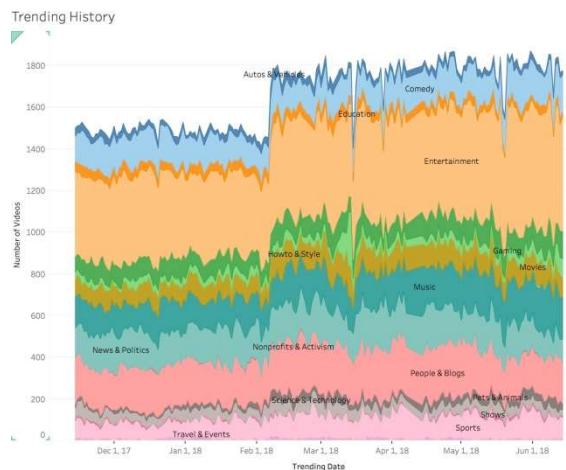
Figure 2: Trending History

Next, we will go to sheet 1, where we will visualize the existing data. The dataset taken from Kaggle starts from 14 November 2017 to 14 June 2018. After that, we will explain the results of the data visualization research that we have examined. As can be seen in Figure 2, the chart is an area chart. An area chart is a line chart in which the area between the lines and axes is shaded according to the part being represented. These charts are typically wont to represent accumulated totals over time and are the traditional thanks to displaying stacked lines. The area chart may be a combination between a line graph and a stacked bar graph. By stacking the quantity beneath the

road, the chart shows the entire fields as their relative size to every other.

Based on the results of the data visualization in Figure 2, it can be seen that the Autos & Vehicles category is the category with the highest number of videos, and the Travel & Events category is the category with the least number of videos. This shows that YouTube users are more interested in videos in the Autos & Vehicles category than other categories.
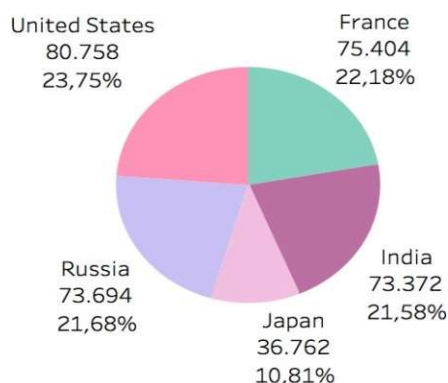

Figure 3: Trending by country

Next in Figure 3 is a pie chart. The pie chart represents data as slices of a circle with different sizes and colors. The slices are labeled, and therefore the numbers like each slice are additionally represented within the chart. We will select the chart option from the Marks card to make a chart. It can be seen in Figure 3 that the highest number of trending videos came from the USA region, amounting to 80,758. A total percent of 23.75%, followed by France with a total number of trending videos of 75,404 and a total percent of 22.18%, Russia with a total number of trending video 73,694 and a total percent of 21.68% India with a total trending video of 73,372 and a total percent of 21.58% Japan with a total number of video trending 36,762 and a total percent of 10.81%.

Based on the horizontal bars in Figure 4, it can be seen that total trending videos in 2017 came from the USA with 18,918, followed by France with 17,760 videos, India with 17,706 videos, Russia with 17,276 videos. Furthermore, it can be seen on the trending date for 2018 that the USA still occupies the first position for the highest number of trending videos with a total number of videos of 61,840, followed by France with a total of 57,644 videos, Russia with a total of 56,418 videos, India with a total of 55,666 videos, and finally Japan with a total of 36,762 videos.
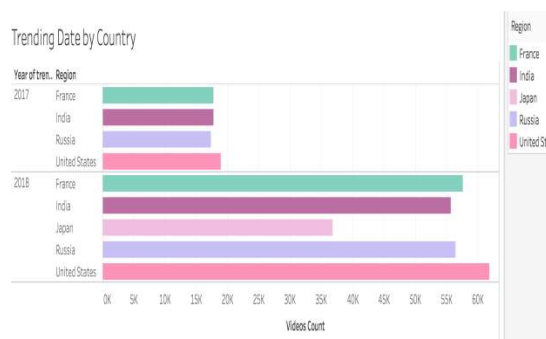

Figure 4: Trending date by country

Figure 5 shows the highlight tables that show the total trending videos by category and country. The highest number of trending videos uploaded comes from India and is in the Entertainment category, amounting to 32,924 videos. Based on the category, it can be seen that the total videos from the five regions that attracted the most users' interest were the Entertainment category. Whereas previously mentioned, India had 32,924 videos, followed by USA 19,638 videos, France 19,020 videos, Japan 11,734 videos, and Russia 11,692 videos. Furthermore, it can be seen that the Trailers category, with a total of 4 trending videos, is the category that attracts YouTube users the least.

Trending by Country and Category -

Category Title	France	India	Japan	Russia	United States
Autos & Vehicles	1.220	138	538	3.116	758
Comedy	8.446	6.814	1.372	5.968	6.870
Education	1.480	2.360	212	1.326	3.284
Entertainment	19.020	32.924	11.734	11.692	19.638
Film & Animation	3.768	3.298	2.140	5.676	4.680
Gaming	2.786	132	1.834	2.050	1.606
Howto & Style	4.668	1.674	1.574	3.928	8.280
Movies	22	32		2	
Music	7.658	7.714	2.480	3.664	12.874
News & Politics	6.526	10.346	2.654	9.858	4.818
Nonprofits & Activism					106
People & Blogs	9.346	4.988	5.792	18.452	6.122
Pets & Animals	468	6	2.250	1.154	1.832
Science & Technology	1.588	1.096	300	2.226	4.722
Shows	198	410		388	114
Sports	8.002	1.424	3.606	3.684	4.250
Trailers	4				
Travel & Events	204	16	276	510	804

Figure 5: Trending by Country and Category

Next, we will analyze our data geographically. With map visualization, it is easier for us to see our data. To create map visualization, we first need to change Region to Latitude and Longitude geographic roles. After that, we will enter Longitude in columns and Latitude in rows. To make it easier to view the map, each region's colors can be changed so that each region's colors differ. In Figure 6, we can see the mapping of the
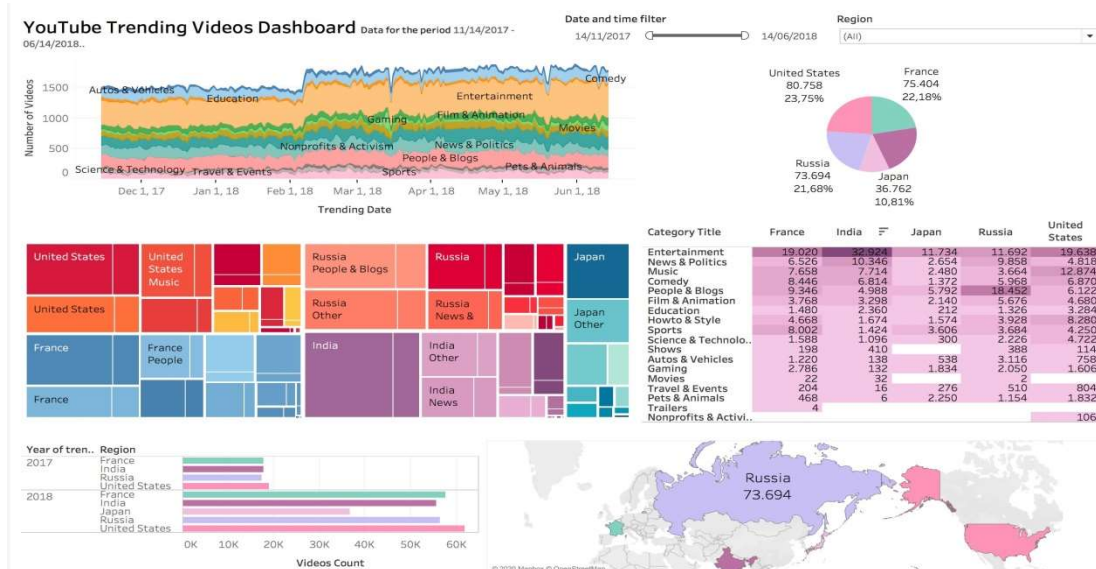
Figure 8. YouTube Trending Videos Dashboard

USA, France, Russia, India, and Japan regions with the total number of trending videos from 2017 to 2018.
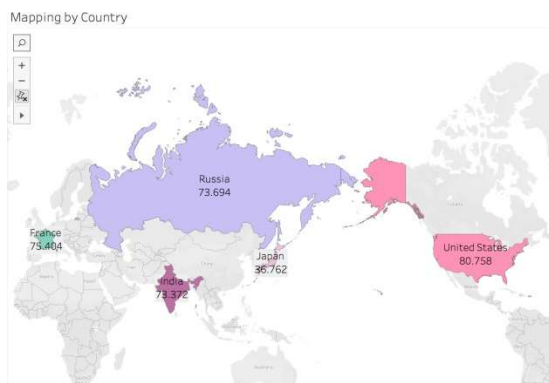

Figure 6: Mapping by Country

Figure 7 shows a treemap chart. Treemap in Tableau may be a primary chart type that's represented by nested rectangular boxes. This chart is often useful for giant datasets for visualization. This graph will mark the hierarchical data for comparative analysis and a crucial chart to research the data set's anomalies. In addition to the scatter plot, treemaps are only visualization types that reasonably allow us to communicate and consume many marks on one view. This makes it easier to identify patterns and relationships that we would not rather be ready to see. Colors and gradients are wont to group items while still identifying individual items. This chart works rather well for a massive amount of knowledge. When the info increases, it also increases the matter of understanding the data.

Treemap charts are effective practices once we want to speak and consume a fair number of marks on one view. With this significant behavior, a user can easily spot patterns and relationships between them, otherwise impossible.
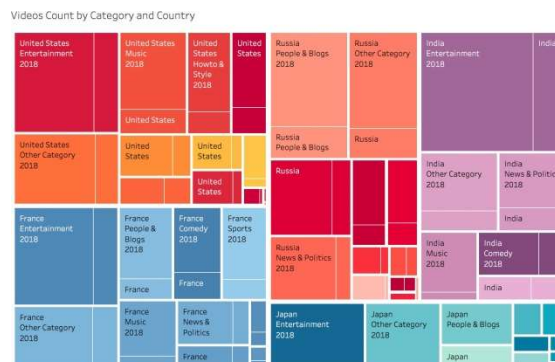

Figure 7: Videos Count by Category and Country

As shown in figure 7, the largest square size is India, with the Entertainment category in 2018 with 25,156 videos trending. Meanwhile, for 2017 with the same country and category, there were 7,768 videos. Next is the USA in 2018 with the same category as before, with 15,010 trending videos. It can be seen that from 2017 to 2018, the Entertainment category attracted the most interest from YouTube users, and in 2018, the second most popular category was the people & blogs category. In 2017, the second most popular category was the comedy category, followed by the people & blogs category.

A dashboard is a collection of sheets, which we can compare with various data simultaneously.

After creating several data visualization sheets on Tableau, we can combine the data visualizations made on these sheets. As we can see in Figure 8, we can see the full data visualization result simultaneously.

CONCLUSION

In this paper, we have analyzed trending YouTube videos data using data visualization. As the "Age of Big Data" kicks into high gear, visualization is an increasingly crucial tool to make sense of the trillions of rows of data generated every day. To better understand the changes brought about by big data, this paper focuses on data analysis using data visualization, where this data visualization we are easier to understand through the visual world, where pictures speak louder than words. Data visualization is especially important for big data and data analysis projects because it allows trends and patterns to be more easily seen. Good visuals tell a story, remove any distractions from data, and highlight useful information. Furthermore, since visualization is so prolific, it is also one of the most useful professional skills to develop.

With the rise of big data upon us, we need to interpret increasingly larger batches of data. Nevertheless, data visualization is not only crucial for data scientists and data analysts; it is necessary to understand data visualization in any career. Whether we work in finance, marketing, tech, design, or anything else, we need to visualize data. That fact showcases the importance of data visualization. Complex algorithms are easier to understand in a data visualization format than lines and text lines and numbers. With so much information being collected through data analysis in the business world today, we must have a way to paint a picture of that data to interpret it. If we are a YouTuber, we will indirectly think about how our videos can be trending and what type of videos trending among users. For any YouTubers who question this, we can use data visualization in answering these questions.

In analyzing YouTube trending videos, we use tableau software to help visualize the data that has been obtained. The following are some of the charts used in this analysis, such as pie charts, area charts, horizontal bars, highlight tables, treemaps, and mapping. Based on the results of the analysis that has been done, we can conclude that the Entertainment category is the category that most attracts YouTube users. This can be proven in Figures 5 and 7, where from 2017 to 2018, the entertainment category was the category with the most trending videos compared to other categories. Furthermore, it can also be seen that the number of trending videos in 2017 was significantly less than in 2018. In 2018 trending videos increased dramatically, and with this comparison, it can be concluded that YouTube users in 2018 also increased.

In 2017 the categories that received much attention from users were entertainment, people & blogs, news & politics, comedy, music, and film & animation. Simultaneously, the less attractive categories to users are the categories of movies, trailers, nonprofits & activism, shows, travels & events. In 2018 the categories that received much attention from users were entertainment, people & blogs, music, news & politics, comedy, and sports. As we can see in Figure 3 Trending by Country, it can be concluded that the most trending videos are from the USA, followed by France and Russia.

With this analysis results, Data Visualization can help YouTube and its users identify which areas need to be improved, which factors affect users' satisfaction and dissatisfaction, and what to do and know to develop parts lacking and pay attention to today's users' needs. Visualized data can give YouTube and its users a better prediction of current trending topics or videos and future growth. It can be used to influence our decision-making. It is especially important when analyzing and implementing strategies to improve YouTube; visualization makes it easy to ensure any platform is as optimized as possible. We need data visualization because the human brain is not well equipped to devour such a lot of raw, unorganized information and switch it into something usable and understandable. We want graphs and charts to speak data findings to spot patterns and trends to realize insight and make better decisions faster.

The problem lies in too much data that is difficult to analyze. In this paper, we suggest using data visualization in analyzing data to make it easier to read and understand. By applying such analytics to big data, valuable information can be extracted and exploited to improve decision-making and support informed decisions. Finally, if appropriately implemented, any new technology can bring with it some potential benefits and innovations,

let alone big data, which is an extraordinary field with a bright future if approached correctly. We believe that big data analytics is essential in this era of data glut and can provide unexpected insights and benefits for decision-makers in various fields. If adequately exploited and applied, big data analytics can provide a basis for progress at the scientific, technological, and human levels.

REFERENCES

[1] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, "YouTube traffic characterization: A view from the edge," 2007, doi: 10.1145/1298306.1298310.

[2] A. Madden, I. Ruthven, and D. McMenemy, "A classification scheme for content analyses of YouTube video comments," *J. Doc.*, 2013, doi: 10.1108/JD-06-2012-0078.

[3] F. E. Gunawan, J. F. Andry, H. Tannady and R. Meylovsky, "Designing Enterprise Architecture Using TOGAF Framework in Meteorological, Climatological, and Geophysical Agency," *Journal of Theoretical and Applied Information Technology*, vol. 97, no. 20, pp. 2376-2385. 2019.

[4] E. W. T. Ngai, L. Xiu, and D. C. K. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification," *Expert Systems with Applications*. 2009, doi:10.1016/j.eswa.2008.02.021.

[5] J. F. Andry, H. Tannady and F. E. Gunawan, "Purchase Order Information System Using Feature Driven Development Methodology," *International Journal of Advanced Trends in Computer Science and Engineering,* vol. 9, no. 2, pp. 1107-1112.

[6] D. Korte, "Business Intelligence in the Hospitality Industry," *Int. J. Innov. Manag. Technol.*, 2013, doi: 10.7763/ijimt.2013.v4.435.

[7] L. C. S. Reddy and Dr. D. Murali, "YouTube: big data analytics using Hadoop and map reduce.," *Int. J. Eng. Technol.*, vol. 7, pp. 12–15, 2018.\

[8] H. Tannady, J. F. Andry, F. E. Gunawan and J. Mayseleste, "Enterprise Architecture Artifacts Enablers for IT Strategy and Business Alignment in Forwarding Services," *International Journal of Advanced Trends in Computer Science and Engineering,* vol. 9, no. 2, pp. 1465-1472.

[9] E. D. Madyatmadja, L. Liliana, J. F. Andry and H. Tannady, "Risk Analysis of Human Resource Information Systems Using Cobit 5," *Journal of Theoretical and Applied Information Technology*, vol. 98, no. 21, pp. 3357-3367.

[10] Nikhat Akhtar, Nazia Tabassum, Asif Perwej, and Yusuf Perwej, "Data analytics and visualization using Tableau utilitarian for COVID-19 (Coronavirus)," *Glob. J. Eng. Technol. Adv.*, 2020, doi: 10.30574/gjeta.2020.3.2.0029.

[11] "Big Data Analytics: Transform terabytes into insights." https://www.tableau.com/learn/articles/big-data-anal ytics#:~:text=Tableau is an end-to,those insights across the organization.

[12] J. R. Mashey, *Big Data and the Next Wave of InfraStress Problems, Solutions, Opportunities*. 1998.

[13] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *Int. J. Inf. Manage.*, 2015,doi:10.1016/j.ijinfomgt.2014.10.007.

[14] J. F. Andry, H. Tannady and F. Nurprihatin, "*Eliciting Requirements of Order Fulfillment in A Company*," The 2nd International Conference on Engineering and Applied Sciences 2019 (InCEAS 2019), IOP Conference Series: Materials Science and Engineering.

[15] J. L. Monino and S. Sedkaoui, *Big Data, Open Data and Data Development*. 2016.

[16] S. V. K. Jagadish *et al.*, "Genetic Advances in Adapting Rice to a Rapidly Changing Climate," *Journal of Agronomy and Crop Science*. 2012, doi: 10.1111/j.1439-037X.2012.00525.x.

[17] A. McAfee and E. Brynjolfsson, "Big data: The management revolution," *Harv. Bus. Rev.*, 2012.

[18] F. J. Ohlhorst, *Big Data Analytics: Turning Big Data into Big Money*. 2012.

[19] Resdiansyah, J. Darmawan, A. H. Wijaya, L. Hakim and H. Tannady, "*Comparing Freeman Chain Code 4 Adjacency Algorithm and LZMA Algorithm in Binary Image Compression,*" Annual Conference on Science and Technology Research (ACOSTER) 2020, Journal of Physics: Conference Series.

[20] Hubspot & Visage, "How To Use Data Visualization To Win Over Your Audience."

[21] S. Ashish, Varma & Sakshi, "Data

www.jatit.org

Visualization: Using Tableau for Decision Making," *Res. Rev. Int. J. Multidiscip.*, vol. 4, no. 4, 2019.

[22] P. Pathak, S. Gupta, and G. S. Dangayach, "Sustainable Waste Management: A Case From Indian Cement Industry," Brazilian J. Oper. Prod. Manag., 2015, doi: 10.14488/bjopm.2015.v12.n2.a7.

[23] S. Sara Banu, D. Uma, M. Waseem Ashfaque, andQ. S. Ali Ahmed, "A review on Visualization Approaches of Data mining in heavy spatial databases," *IOSR J. Comput. Eng. Ver. V*, 2015, doi:10.9790/0661-171595105.

[24] M. Khan and S. Shah Khan, "Data and Information Visualization Methods, and Interactive Mechanisms: A Survey," 2011.

[25] A. S. Nagdive and D. R. M. Tugnayat, "Overview on Performance Testing Approach in Big Data," *Int. J. Adv. Res. Comput. Sci.*, 2014.