Student's Name: Rajatha Raju

Student Id: RR24805N

Course: Research Project

Date: 05/12/2024

Unraveling the Influence of Thumbnail Background Colors on YouTube Viewer Engagement

Abstract:

This research paper investigates the influence of thumbnail background colors on viewer engagement metrics, such as likes, views, and comments, on YouTube. Utilizing data obtained from the YouTube Data API[5][6] and employing regression modeling techniques, the study explores the relationship between thumbnail background colors and engagement metrics across various content categories. Findings suggest that thumbnail background colors indeed impact engagement, with red backgrounds correlating with higher likes and blue backgrounds correlating with increased views. Moreover, sentiment analysis reveals potential associations between background colors and the sentiment expressed in video descriptions and titles. However, while significant, the effect sizes of thumbnail background colors on engagement metrics are relatively modest, highlighting the multifaceted nature of viewer engagement on

YouTube. The paper concludes by emphasizing the importance of considering factors beyond thumbnail colors, such as content relevance and viewer preferences, in optimizing engagement strategies on the platform.

Introduction:

With the exponential growth of digital video consumption, platforms like YouTube are increasingly focusing on optimizing engagement metrics to attract and retain audiences. Thumbnails play a pivotal role in this ecosystem, serving as visual cues that entice viewers to click and engage with content. Understanding the impact of thumbnail features, particularly background colors, on viewer engagement is crucial for content creators and marketers seeking to maximize their reach and impact. This paper aims to investigate the influence of thumbnail background colors on viewer engagement metrics, offering insights into the nuanced dynamics of audience interaction on YouTube. By analyzing data from the YouTube Data API[5][6] and employing regression modeling techniques, the study seeks to uncover patterns and associations between thumbnail background colors and engagement metrics across different content categories and regions. The findings are expected to provide valuable guidance for content creators and marketers in optimizing their thumbnail designs and overall engagement strategies on the platform.

Literature review:

Impact of Thumbnail Attributes on Video Views

In their study titled "Country by Country Comparison of Thumbnail Features Contributing to Views Using AIME for YouTube," Rintaro Fukui, Ryotaro Okada, Ayako Minematsu, and Takafumi Nakanishi introduce a thumbnail feature extraction method utilizing AIME to determine which features contribute to the number of views on YouTube. Analyzing thumbnails from seven countries, including the U.S., Japan, and Germany, the authors observed variations in influential features. For instance, Japan prioritizes text area in thumbnails, while Germany emphasizes the richness of facial expressions. This method aids in understanding cultural differences in viewer preferences across different regions, providing valuable insights for content creators aiming to maximize viewership.[1]

In their study titled "How Digitally Extractable Attributes of YouTube Video Thumbnails and Titles Affect Video Views," Nathan Jacky Lee, Muhammad Devin Nayottama A.P., Anderies, and Andry Chowanda investigate the impact of content-independent visual attributes of YouTube video thumbnails and titles on video views. Utilizing Python libraries and AI models, the researchers extract data such as image complexity, thumbnail text, sentiment, and faces from a dataset of 1600 videos. Contrary to expectations, the study finds that these attributes have a negligible effect on a video's ability to attract views. Despite some minor findings, the research concludes that none of the studied attributes reliably affect video performance. Limitations include the dataset being limited to "Science and Technology" channels and the possibility that viewers are more attracted to video content rather than thumbnail and title design quality.

Future research directions include exploring attributes in other channel themes/categories and testing the proposed possibilities further. (Lee et al., 2023, pp. 82-85)[2]

Non-Content Features Influencing Video Popularity

In their study titled "Identifying content unaware features influencing popularity of videos on YouTube: A study based on seven regions," Zahid Halim, Sajjad Hussain, and Raja Hashim Ali present an AI-based framework to identify the non-content features influencing a video's trend potential on YouTube. Analyzing data from seven regions, including Canada, France, Germany, India, Pakistan, United Arab Emirates, and the United States, the research identifies key attributes affecting video popularity. By utilizing feature selection methods and AI classifiers, the study offers insights into factors contributing to a video's trending status, aiding content creators in optimizing their videos for increased visibility. The work contributes to understanding YouTube content dynamics and provides practical implications for content creators to enhance their video's appeal and reach.[3]

Contextualization: Growth of IP Video Traffic

According to 2018 Cisco predictions, IP video will account for 82% of all IP traffic globally by 2022, which is an increase from 75% in 2017. So digital platforms most certainly are thinking about optimizing thumbnail performance. That is evident by the pervasiveness of thumbnails on video platform home screens and increasingly rich, interactive thumbnail experiences that resemble micro-video portals.[4]

Hypothesis Development: Thumbnail Background Colors and Viewer Engagement

“Hypothesis: Thumbnail background colors influence viewer engagement metrics such as likes, views, and comments on YouTube videos. I posit that certain background colors, such as red, green, and blue, may evoke different emotional responses and attract varying levels of attention from viewers. By analyzing a dataset of YouTube videos across different content categories, I aim to uncover patterns that demonstrate the impact of thumbnail background colors on viewer interaction and engagement. Additionally, I hypothesize that factors such as the sentiment expressed in video descriptions and titles may further influence viewer engagement, interacting with the effect of thumbnail colors to shape overall engagement metrics. Through comprehensive data analysis and regression modeling, I seek to provide empirical evidence supporting the hypothesis that thumbnail background colors play a significant role in driving viewer engagement on YouTube."

Methodology

Data Collection and Preprocessing

YouTube Data API[5][6] data provides insights into video popularity and characteristics, useful for research on trends, engagement, and content preferences. Obtained an API key from YouTube Data API[5][6].Sent API requests to fetch video data, specifying parts (snippet, statistics, contentDetails), region codes, and parameters like chart type. Processed JSON

responses to extract video information like ID, title, description, stats, duration, and category. Merged data from multiple API requests into a single DataFrame using Pandas.

It has 5650 rows with 2096 unique records. With columns such as country, video id, title, description, thumnail_url, likes, views, comments, duration, definition, category name. additional columns were added such as background_color with rbg values, description_length, title_length, views_to_likes_ratio likes_to_views_ratio likes_to_comments_ratio, comments_to_likes_ratio for in detail analysis.

Thumbnail Feature Extraction

The code imports necessary libraries such as Pandas, NumPy, PIL, and requests to process YouTube video data stored in a CSV file named 'youtube_data.csv' into a DataFrame called all_data. It defines a function, extract_background_color(image_url), to derive the dominant color from thumbnail images. This function downloads images using requests, opens them with PIL, resizes them for efficiency, converts them to NumPy arrays, and computes the mean RGB value as the background color. This function is then applied to each 'thumbnail_url' cell in all_data using the apply() function, updating a new column named 'background_color' with the extracted colors. Finally, the code displays the DataFrame's head with the added column for review.

Summary Statistics of YouTube Video Metrics

The mean number of likes, views, and comments are approximately 361,619, 12,490,200, and 9523, respectively. The background colors of the video thumbnails are represented by three

categories: red, green, and blue. On average, there are 82.87 video thumbnails with a red background, 73.62 with a green background, and 69.95 with a blue background.

Regarding the sentiment analysis, the mean description subjectivity and polarity are approximately 0.24 and 0.10, respectively. Similarly, the mean title subjectivity and polarity are approximately 0.15 and 0.05, respectively. The mean description length is 850.65 characters, and the mean title length is 50.91 characters.

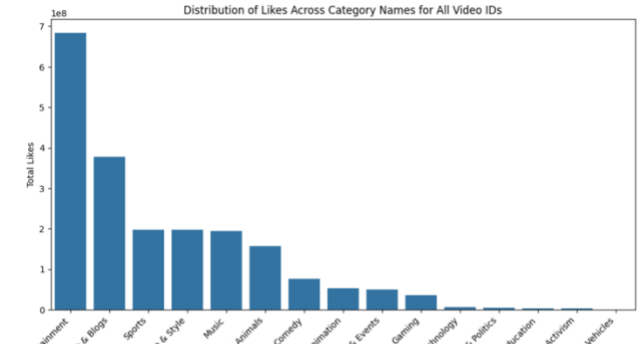
Additionally, there are calculated ratios such as views-to-likes, likes-to-views, and likes-to-comments ratios. The mean views-to-likes ratio is infinite, indicating that some videos have zero likes. Similarly, the mean likes-to-views ratio is approximately 0.043, and the mean likes-to-comments ratio is also infinite, suggesting some videos have zero comments. Finally, the mean comments-to-likes ratio is also infinite, indicating some videos have zero comments.

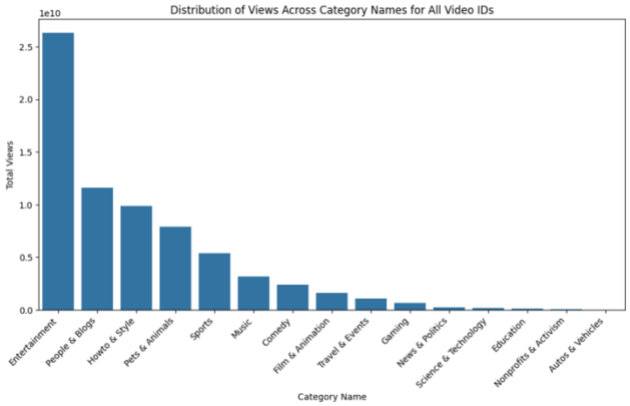
Results

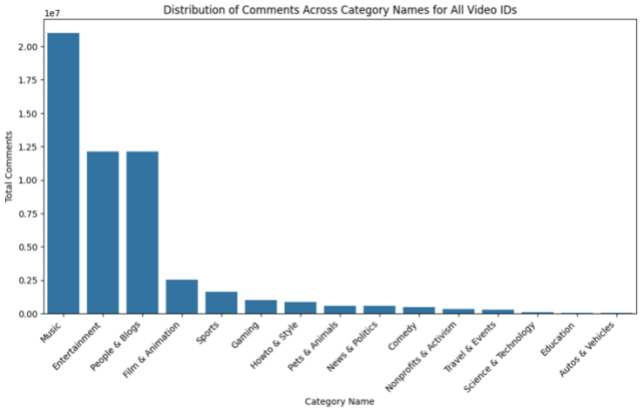
Distribution of Thumbnail Background Colors

In the analyzed dataset of YouTube videos, those featuring red-background thumbnails (468214) are the most abundant, closely followed by videos with green-background thumbnails (415978). Thumbnails with blue backgrounds (395192) are slightly less prevalent compared to the others

Engagement Across Content Categories



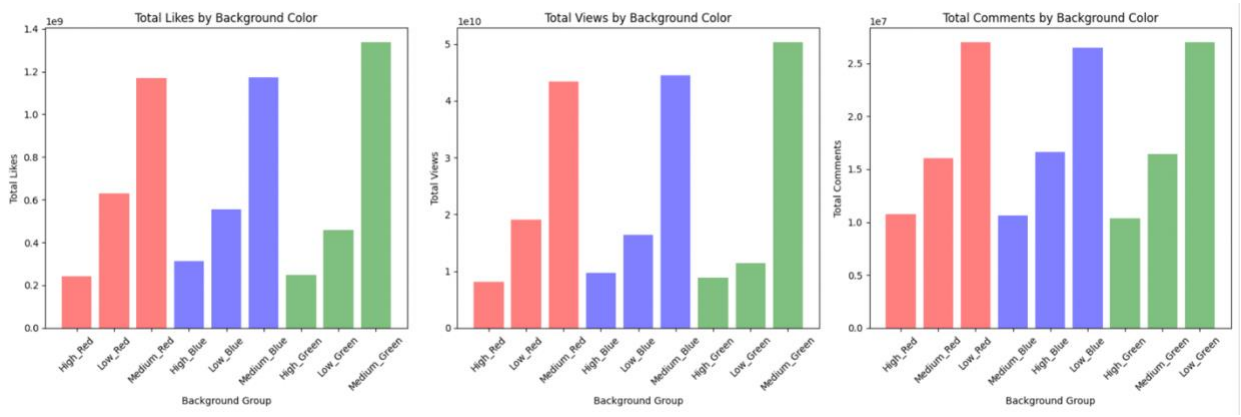




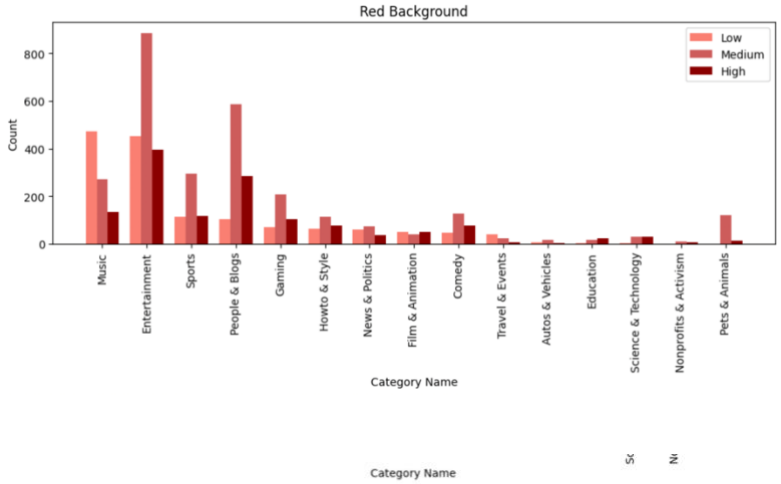
The data reveals significant insights into viewer engagement on YouTube across various content categories, considering total views, likes, and comments for all video IDs. Entertainment emerges as the most viewed and liked category, accumulating a remarkable 26,301,990,818 total views and 683,474,336 total likes, respectively, indicating substantial viewer interest and appreciation for entertainment content. People & Blogs closely follows, with 11,581,944,593 total views and 377,625,106 total likes, highlighting considerable engagement with personal and lifestyle-related content. Moreover, Music stands out with the highest total comments at 21,000,610, indicative of active viewer interaction and discussion around music videos.

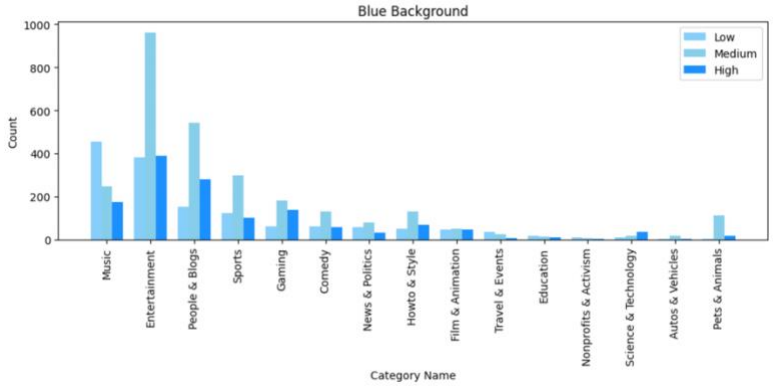
Overall, these findings underscore the diverse preferences and engagement levels across content categories, offering valuable insights for content creators and marketers to tailor their strategies effectively.

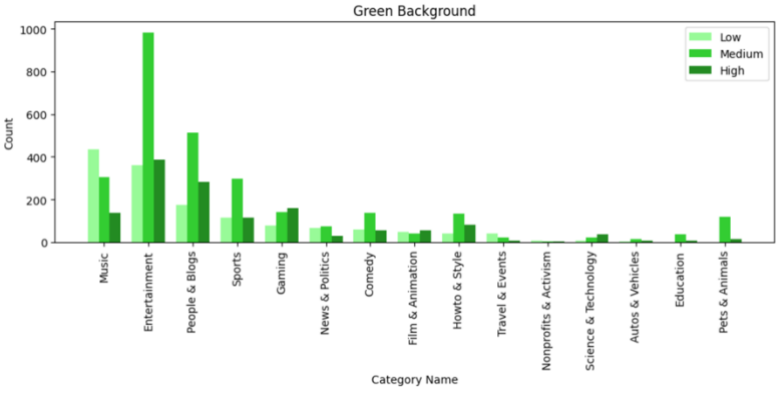
Insights from Thumbnail Color Analysis



Red-backgrounded video thumbnails amassed a total of 1.77 billion likes, predominantly in the medium intensity category with 1.17 billion likes, followed by blue-backgrounded video thumbnails with 2.31 billion likes, primarily medium intensity at 1.17 billion. Green-backgrounded video thumbnails received 2.01 billion likes, with 1.33 billion likes in the medium intensity category. In terms of views, red-backgrounded video thumbnails garnered 30.57 billion views, with 43.36 billion in the medium intensity. Blue-backgrounded video thumbnails accumulated 70.70 billion views, mainly medium intensity at 44.46 billion. Green-backgrounded video thumbnails received 70.79 billion views, with 50.14 billion in the medium intensity. For comments, red-backgrounded videos obtained 53.70 million comments, mostly from low intensity video thumbnails at 26.97 million. Blue-backgrounded video thumbnails received 53.21 million comments, primarily from high intensity video thumbnails at 16.65 million. Green-backgrounded video thumbnails garnered 53.81 million comments, with 27.00 million from low-ranked video thumbnails. Overall, medium-intensity video thumbnails showed the highest engagement levels across background colors, reflecting the dynamic interaction patterns of YouTube viewers.



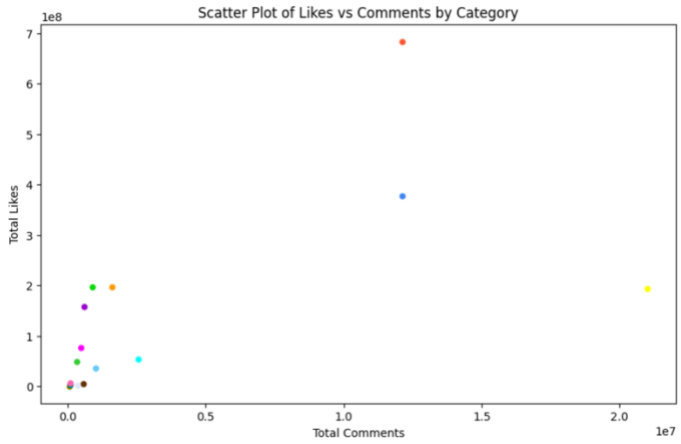




The analysis highlights the impact of thumbnail background colors (Red, Blue, Green) on viewer engagement across various YouTube content categories. Video thumbnails with red backgrounds generally received low engagement, notably in categories like Autos & Vehicles and Comedy, with counts ranging from 12 to 77 across intensity levels. In contrast, the Entertainment category showed high engagement, particularly in the medium intensity level (480). For blue backgrounds, engagement patterns mirrored those of red backgrounds, albeit with slight variations in intensity levels. Meanwhile, green background videos exhibited mixed engagement, with some categories, such as Entertainment and Music, seeing high engagement in the medium intensity range, while others varied across low and high intensity levels.

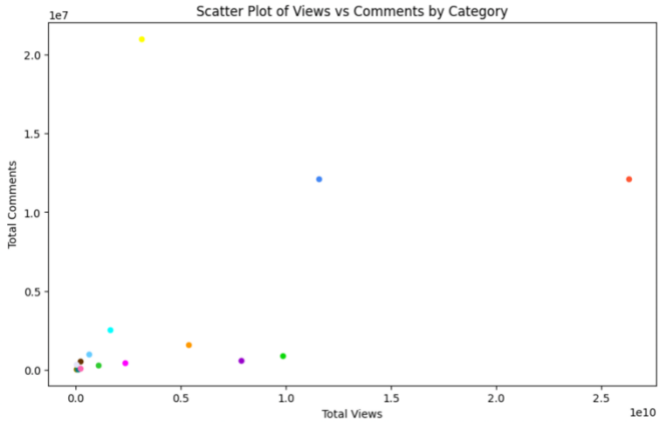
Overall, these insights underscore the significance of thumbnail colors in driving viewer interaction, offering valuable guidance to content creators and marketers for optimizing engagement strategies.

Correlation Analysis: Likes, Views, and Comments



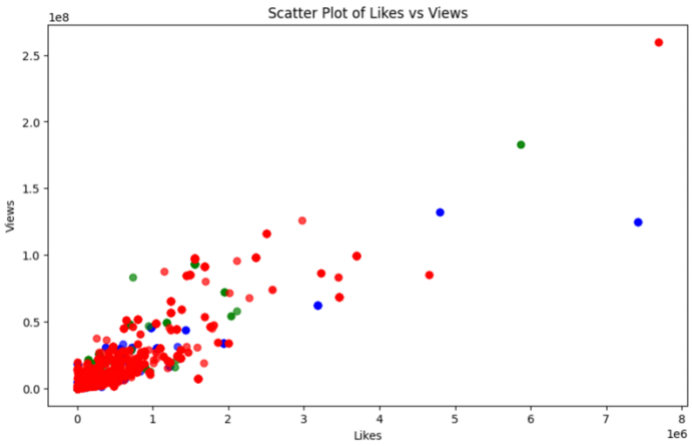
There is a weak positive correlation between the number of likes and the number of comments for videos with positively rated comments. This means that as the number of likes increases, the number of comments tends to increase as well, but this trend is not very strong. It appears that the categories with high likes also have high comments, namely Entertainment, People & Blogs, and Music. Further analyzing the top 10 outputs of likes versus comments, the data suggests that the choice of thumbnail background color can influence viewer engagement, as indicated by the likes and comments received across various content categories. Different colors appear to have varying effects across categories; for example, red backgrounds seem effective in categories like "Howto & Style," while blue and green backgrounds may be less engaging in "Travel & Events." However, within specific categories like "Pets & Animals" and "Music," there's consistency in engagement levels across different background colors. Factors beyond color, such as video content and viewer preferences, also play crucial roles in engagement.

Overall, the data reflects a diverse range of engagement levels, highlighting the complexity of viewer behavior and the need for content creators to experiment with thumbnail designs while focusing on creating compelling content.



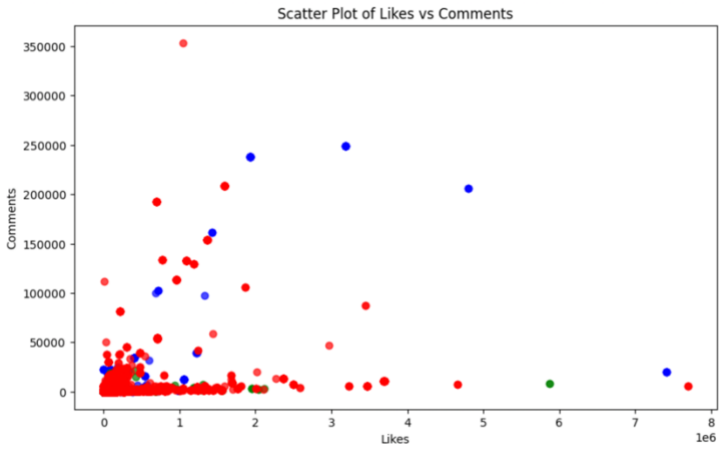
The scatter plot analysis reveals that there is no clear overall correlation between the number of views and the number of comments. This suggests that videos with a high number of views don't necessarily have a high number of comments, and vice versa. However, upon further analysis of the top 10 outputs for views versus comments ratio, thumbnail background colors seem to influence engagement differently across categories. For instance, in the "Entertainment" category, the video titled "Ages 1 - 100 Decide Who Wins $250,000" received a high number of views (132,394,511) and comments (206,275) despite having a blue background (106). This indicates that the content itself may be more impactful than the thumbnail color. Interestingly, some categories exhibit consistency in engagement levels across different thumbnail background colors, while others show variability. For example, videos in the "Pets & Animals" category received similar engagement levels regardless of the background color, suggesting that content relevance might outweigh color choice in this category. Moreover, the data reflects a diverse range of engagement levels across categories, with some videos garnering millions of views and thousands of comments, while

others receive comparatively fewer engagements. This diversity underscores the complex nature of viewer behavior and preferences. Additionally, the titles and themes of the videos likely play a crucial role in attracting viewer engagement. For instance, videos with intriguing titles like "THE AMAZING DIGITAL CIRCUS - Ep 2: Candy Carrier Chaos!" in the "Film & Animation" category received significant engagement despite variations in thumbnail background colors. This suggests that compelling content and titles can sometimes override the influence of thumbnail colors on engagement.



The scatter plot analysis reveals several key insights into viewer engagement on YouTube. Firstly, there's a faint upward trend from the bottom left corner to the top right corner, indicating that videos with more views tend to also have more likes. However, the spread of the data points suggests that this correlation is weak. Thumbnail colors vary across the chart, with shades of blue and red being more prevalent. While thumbnail color may influence engagement to some extent, there's significant variability in the data. Many videos with a high number of views have a low number of likes, and vice versa, suggesting that factors other than thumbnail color likely play a more significant role in determining engagement levels. Additionally, delving into the top 10 videos with the highest views-to-likes ratio reveals intriguing patterns. These videos, such as

"#خَلْطي وطَيْبي ناجح مليار بالمئة👌بالسميد فقط..." and "По ту сторону потолка 🐶 #visp #ремонт #натяжны...", showcase exceptional engagement despite receiving relatively fewer likes compared to their views. This suggests a strong interest from viewers in the content, potentially due to its broad appeal. Conversely, analyzing the top 10 videos with the highest likes-to-views ratio unveils a different trend. Videos like "[n월의 석진] Message from Jin : May 2024 💌" and "RM 'Come back to me' Official Teaser" exhibit a high ratio of likes to views, indicating significant engagement and appreciation from the audience. These insights underscore the multifaceted nature of viewer engagement on YouTube, emphasizing the importance for content creators to understand audience preferences and tailor their content strategy accordingly.

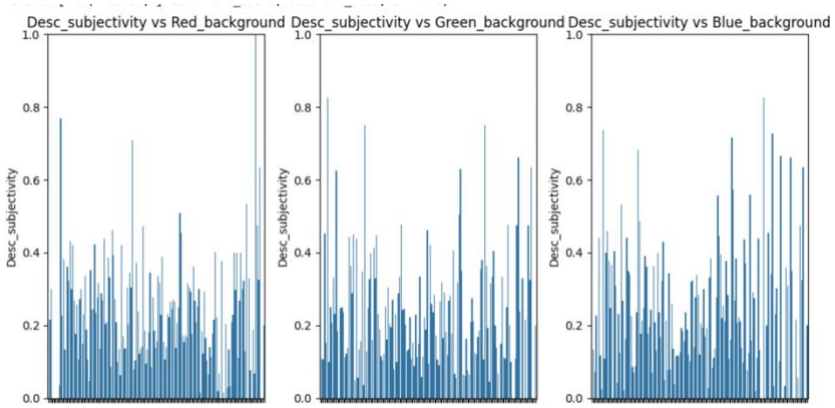


The scatter plot analysis reveals several key insights into viewer engagement on YouTube. Firstly, there's a faint upward trend from the bottom left corner to the top right corner, indicating that videos with more likes tend to also have more comments. However, the spread of the data points suggests that this correlation is weak. Additionally, while the colors of the dots represent the dominant color in the video thumbnail, there's no single dominant color across videos with the most likes and comments. Some colors appear more frequently throughout the chart than others, but it's

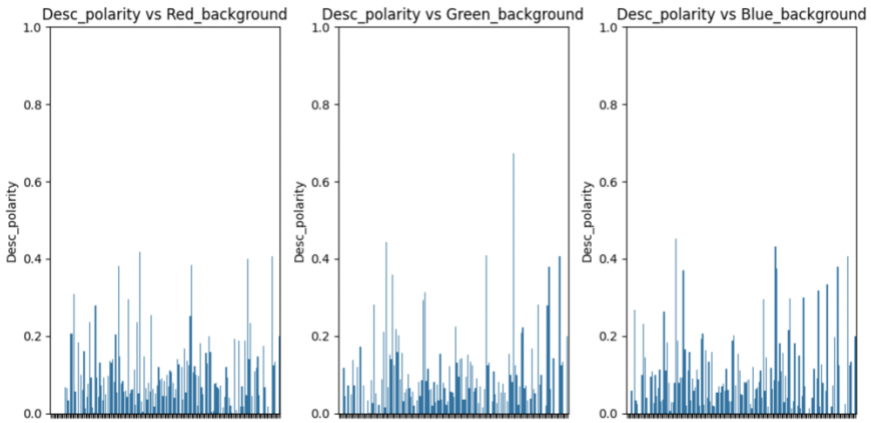
difficult to definitively associate a specific color with a higher likelihood of receiving likes and comments. Moreover, there's significant variability in the data, with many videos having a high number of likes but a low number of comments, and vice versa. This variability suggests that factors other than thumbnail color likely play a bigger role in determining how many comments a video receives. Turning to the top 10 analysis, examining the videos with the highest likes-to-comments ratio reveals a notable disparity between likes and comments. Videos like "#بالسميد فقط🤌خَلْطي وطَيْبي ناجح مليار بالمئة..." and "По ту сторону потолка 😱 #visp #ремонт #натяжны..." showcase a significant number of likes relative to their comments, suggesting a strong positive reception from viewers without generating a proportional amount of comments. Conversely, exploring the videos with the highest comments-to-likes ratio unveils a contrasting trend. Videos such as "Papa Feeding Baby ♡|" and "Mom Meets Peter Dinklage..💀," exhibit an exceptionally high number of comments relative to their likes, indicating extensive discussion and interaction among viewers despite a lower number of likes. Interestingly, emotionally charged or thought-provoking content appears to drive increased engagement through comments.

In summary, these analyses highlight the multifaceted nature of viewer engagement on YouTube. While some videos may excel in generating likes without generating a proportional amount of comments, others may foster extensive discussion and interaction among viewers despite a lower number of likes. This underscores the importance of considering both likes and comments as valuable metrics for assessing audience engagement and tailoring content strategies accordingly.

Sentiment Analysis of Titles and Descriptions

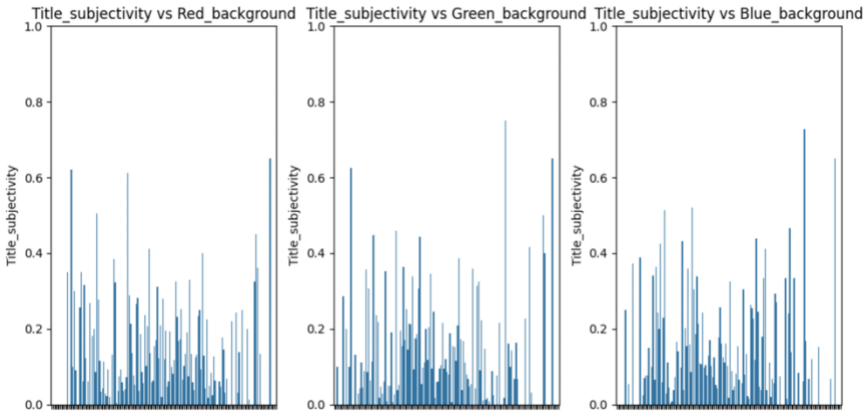


The analysis of the chart suggests a potential association between the background thumbnail color of a video and the average subjectivity of its description. On average, descriptions for video thumbnails with green backgrounds exhibit the highest level of subjectivity, followed by video thumbnails with blue backgrounds and then red backgrounds. Subjectivity in writing refers to the expression of personal opinions or beliefs, implying that videos with green backgrounds might employ more subjective language in their descriptions compared to those with other background colors. For instance, descriptions for videos with green backgrounds might include phrases like "This is the best product ever" or "This movie is terrible," reflecting the subjective viewpoints of the content creators or reviewers.



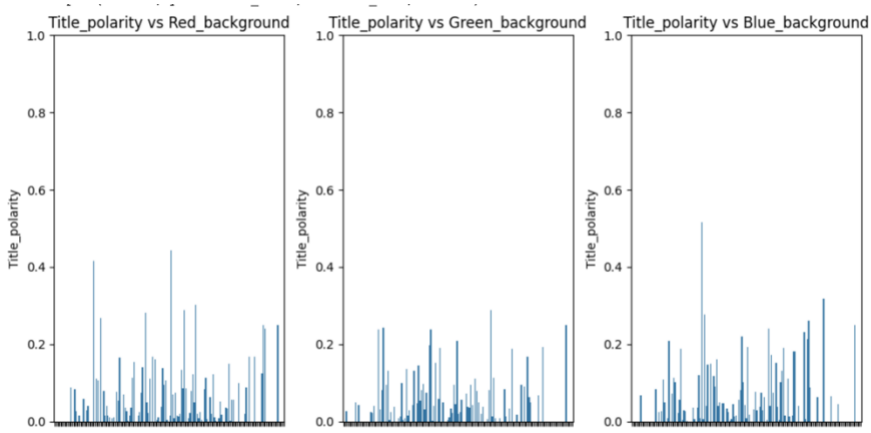
The analysis of the chart reveals a potential correlation between the thumbnail background color of a video and the average polarity of its description. It suggests that video thumbnails with

different background colors may elicit varying sentiments in their descriptions. Specifically, the average description polarity appears to be slightly negative for video thumbnails with red backgrounds, slightly positive for those with green backgrounds, and around zero for video thumbnails with blue backgrounds. Description polarity likely refers to whether the description leans more positive or negative in sentiment. For instance, a description that says "This product is amazing" would likely be considered positive, while a description that says "This product is awful" would be considered negative.



The analysis of the data suggests a potential link between thumbnail background color and title subjectivity, although the limited number of data points makes it challenging to draw definitive conclusions. Title subjectivity appears relatively similar for videos with green and blue backgrounds, slightly higher than for video thumbnails with red backgrounds. Subjectivity, in the context of writing, denotes the expression of personal opinions or beliefs, implying that titles for videos with green and blue backgrounds might utilize more subjective language compared to those with red backgrounds. For instance, a video titled "The Best Way to Make Chocolate Chip Cookies" on a

green background may reflect a more subjective tone than a video titled "How to Make Chocolate Chip Cookies" on a red background.



The chart's limited scope, focusing on just three thumbnail background colors, makes it challenging to discern a definitive overall trend. When considering individual colors, video thumbnails with red backgrounds tend to exhibit slightly negative title polarity, while those with green backgrounds lean somewhat positive. For video thumbnails with blue backgrounds, the title polarity falls between the extremes of red and green, though it remains uncertain whether it leans towards positivity or negativity.

In summary, the analyses highlight the potential influence of thumbnail background color on the language and sentiment expressed in video descriptions and titles. While there are discernible trends, the relationship appears to be complex and nuanced, suggesting that background color may play a role in shaping the textual attributes of YouTube videos, albeit with variations across different colors.

Regression Models: Thumbnail Color Impact on Engagement Metrics

Regression analyses were performed separately for likes, views, and comments, with two models constructed for each target variable. One model excluded thumbnail colors, while the

other included thumbnail colors to assess their potential significance. The findings are outlined below.

Regression Analysis of **Likes:**

The regression results for both models indicate a high level of explanatory power, with R-squared values of 0.851 and 0.852 respectively. When considering the impact of adding red, green, and blue thumbnail background variables, it's notable that the coefficients for red background and blue background become statistically significant, suggesting their influence on the dependent variable, likes. In the first model without background color variables, the coefficient for description polarity is statistically significant, indicating that description polarity has a significant effect on likes. However, in the second model, the coefficient for description polarity remains significant, along with additional significant coefficients for red background and blue background. The coefficient for green background, however, is not statistically significant in either model.

Specifically, in the second model, the coefficient for red background is negative (-477.2501) with a p-value of 0.022, indicating that video thumbnails with red backgrounds tend to have slightly fewer likes compared to those without. Conversely, the coefficient for blue background is positive (561.8866) with a p-value of 0.024, suggesting that video thumbnails with blue backgrounds tend to receive slightly more likes. The coefficient for green background is near zero (-19.9471) and not statistically significant (p-value: 0.949), implying that green background has little to no discernible impact on likes.

Overall, while the addition of background color variables does show some influence on likes, particularly for red and blue thumbnail backgrounds, the magnitude of this effect may not be substantial. The differences observed may be considered relatively minor, with the influence of thumbnail background color on likes being nuanced rather than highly pronounced. However, the statistical significance of red background and blue_background coefficients suggests that these variables do contribute to variations in likes, albeit to a modest degree.

Regression Analysis of **Views:**

The regression analyses reveal that both models exhibit a relatively high level of explanatory power, with R-squared values of 0.834 and 0.835 respectively. When examining the impact of including red, green, and blue thumbnail background variables, it's evident that these variables contribute to the models' explanatory capabilities.

In the first model without thumbnail background color variables, the coefficients for hd (high-definition), sd (standard-definition), and description length are not statistically significant, suggesting they have little influence on views. However, in the second model with thumbnail background color variables, the coefficients for red background, blue background, and green background become statistically significant.

The significant coefficients indicate that background color does play a role in influencing views. Specifically, video thumbnails with blue backgrounds tend to attract significantly fewer views compared to video thumbnails without background color variables, as indicated by the coefficient of -50370 for blue background. Conversely, video thumbnails with green backgrounds tend to receive significantly more views, as indicated by the coefficient of 27030

for green background. The coefficient for red background is positive (10660), but it is not statistically significant at the 0.05 level, suggesting a less pronounced effect compared to blue and green backgrounds.

Overall, the addition of red, green, and blue thumbnail background variables enhances the models' ability to capture variations in views, highlighting the importance of considering visual elements in understanding viewer engagement. While the impact of red background appears less significant, the contrasting effects of blue and green backgrounds suggest that background color can indeed influence viewer behavior.

Regression Analysis of **Comments:**

In the first model without thumbnail background color variables, the overall explanatory power, as indicated by the R-squared value, is 0.405, suggesting that approximately 40.5% of the variance in comments can be explained by the included variables. Among the predictors, description polarity, title polarity, views, likes, and description length appear to be statistically significant ($p < 0.05$), indicating that they have a significant impact on the number of comments. However, variables such as hd, sd, description subjectivity, title subjectivity, and duration minutes do not seem to have a significant effect on comments.

In the second model including red, green, and blue thumbnail background variables, the overall explanatory power increases slightly to an R-squared value of 0.411. However, the addition of these background color variables does not seem to have a substantial impact on the model's performance, as their coefficients are not statistically significant at the 0.05 level. This suggests

that the inclusion of red, green, and blue background variables does not significantly improve the model's ability to predict the number of comments.

Overall, while the model with thumbail background color variables shows a slight improvement in explanatory power, the addition of red, green, and blue background variables does not appear to make a significant difference in predicting the number of comments. The key predictors of comments remain consistent across both models, including description polarity, title_polarity, views, likes, and description length.

In summary, the regression analyses highlight the nuanced influence of thumbnail colors, particularly red and blue backgrounds, on engagement metrics such as likes and views. While these colors show statistically significant effects, their impact may not be substantial compared to other predictors such as description polarity. Overall, the findings underscore the importance of considering visual elements, including thumbnail colors, in understanding viewer engagement on YouTube.

Discussion

The analysis delves into the potential influence of thumbnail background color on viewer engagement metrics like likes, views, and comments on YouTube. It reveals a distribution where red-background thumbnails are most prevalent, followed closely by green and blue backgrounds, setting the stage for understanding their impact on engagement. Across various content categories, such as Entertainment and People & Blogs, substantial engagement is observed, highlighting diverse viewer preferences. The study suggests that thumbnail background colors may indeed influence engagement, with red backgrounds attracting more

likes and blue backgrounds garnering higher views. Additionally, sentiment analysis hints at associations between background colors and the sentiment expressed in descriptions and titles. Regression analysis further quantifies these relationships, showing modest effects of background colors on engagement metrics. Overall, the analysis underscores the multifaceted nature of viewer engagement on YouTube, emphasizing the importance of considering factors beyond just thumbnail colors, such as content relevance and viewer preferences, in optimizing engagement strategies.

While the analysis provides evidence supporting the hypothesis that thumbnail background colors influence viewer engagement metrics on YouTube, the effect sizes are relatively small. Other factors, such as content relevance, video quality, and viewer preferences, also play significant roles in determining engagement levels. Therefore, while thumbnail design is important, it should be considered as part of a broader strategy aimed at optimizing engagement on the platform.

Limitations and Future Directions

While this research offers valuable insights into the relationship between thumbnail background colors and viewer engagement metrics on YouTube, several limitations should be acknowledged. Firstly, the study primarily focuses on a subset of features extracted from thumbnails, such as background colors, and their impact on engagement metrics. Other potentially influential factors, such as thumbnail imagery and text, were not extensively explored in this analysis. Secondly, the dataset used in this study may not fully represent the diverse range of content and viewer preferences on YouTube, as it was limited to specific

content categories and regions. Moreover, the analysis relies on correlation-based approaches, which do not establish causation and may overlook underlying complexities in viewer behavior. Additionally, the study's findings may be subject to biases inherent in data collection and processing methods, as well as limitations in the accuracy of sentiment analysis algorithms. Therefore, while the results offer valuable insights, they should be interpreted with caution, and further research incorporating a broader range of factors and methodologies is warranted to deepen our understanding of viewer engagement on YouTube.

Conclusion:

In conclusion, this study illuminates the complex interplay between thumbnail background colors and viewer engagement metrics on YouTube. Our analysis reveals that thumbnail colors indeed influence engagement, with red backgrounds correlating with higher likes and blue backgrounds with increased views. Additionally, sentiment analysis hints at associations between background colors and the sentiment expressed in video descriptions and titles. However, while thumbnail design is significant, it is just one aspect of a broader engagement strategy on the platform. Factors such as content relevance, video quality, and viewer preferences also play substantial roles. Acknowledging the limitations of our study, including its focus on a subset of thumbnail features and the need for further research, we emphasize the importance of a holistic approach to engagement optimization. Content creators and marketers should leverage a nuanced understanding of viewer behavior and preferences, considering various factors beyond thumbnail colors to create compelling content that resonates with their audience.

Appendix

1. https://ieeexplore-ieee-org.rlib.pace.edu/document/10488191

R. Fukui, R. Okada, A. Minematsu and T. Nakanishi, "Country by Country Comparison of Thumbnail Features Contributing to Views Using AIME for YouTube," 2023 15th International Congress on Advanced Applied Informatics Winter (IIAI-AAI-Winter), Bali, Indonesia, 2023, pp. 342-347, doi: 10.1109/IIAI-AAI-Winter61682.2023.00069. keywords: {Video on demand;Machine learning;Media;Feature extraction;Internet;Web sites;Australia;AIME;XAI;Thumbnail features contribute;Number of views prediction},

2. https://ieeexplore-ieee-org.rlib.pace.edu/document/10405243

N. J. Lee, M. D. Nayottama A.P., Anderies and A. Chowanda, "How Digitally Extractable Attributes of YouTube Video Thumbnails and Titles Affect Video Views," 2023 IEEE 7th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Purwokerto, Indonesia, 2023, pp. 82-87, doi: 10.1109/ICITISEE58992.2023.10405243. keywords: {Visualization;Video on demand;Codes;Web sites;Data mining;Reliability;Artificial intelligence;You Tube;thumbnail;title;data mining;feature extraction},

3. https://www.sciencedirect.com/science/article/abs/pii/S0957417422010946

Zahid Halim, Sajjad Hussain, Raja Hashim Ali,Identifying content unaware features influencing popularity of videos on YouTube: A study based on seven regions,Expert Systems with Applications,Volume 206,2022,117836,ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2022.117836.

4. https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_2022_Forecast_Highlights.pdf

5. YouTube Data API. (2024, May 12). Retrieved from https://developers.google.com/youtube/v3/docs/videos/list

6. YouTube Data API. (2024, May 12). Retrieved from https://developers.google.com/youtube/v3/docs/videoCategories/list