



FREDRICK TBB

@Fredricktbb · 923 subscribers · 208 videos

Super music producer 🔥 >

Customize channel

Manage videos

Home

Videos

Shorts

Playlists

Community



Latest

Popular

Oldest



KHAID X REMA TYPE BEAT

12K views · 1 year ago



NAIRA MARLEY X MOHBAD TYPE
BEAT

11K views · 2 years ago



T.I BLAZE TYPE BEAT

10K views · 1 year ago



KHAID X OMAH LAY TYPE BEAT

7K views · 2 months ago

Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion
- Appendix

Executive Summary

- The YouTube Channel Analysis project aimed to assess performance metrics and engagement strategies for a specific YouTube channel. Key findings include:
- Strong positive correlation between view count and like count.
- Weak negative correlation between video duration and engagement metrics.
- Genre analysis highlighted a strong interest in street/hip-hop content.
- Effective engagement strategies identified include CTAs and social media sharing.
- SVM identified as the best-performing method for predictive analysis.
- Overall, the analysis underscores the importance of engagement strategies and genre focus for channel growth.

Introduction

- The YouTube Channel Analysis project delves into the performance metrics and engagement strategies of my YouTube channel. With the proliferation of online video content, understanding how to optimize viewer engagement and maximize channel growth is paramount for content creators like myself. By leveraging the YouTube API, this project aims to gain insights into the factors that contribute to the success of my channel, with a focus on metrics such as view count, like count, comment count, and video duration.
- Problems to Address
- The project seeks to address several key questions:
 - What are the main performance metrics of my YouTube channel, and how do they correlate with each other?
 - Are there specific genres or content types that resonate more with my audience?
 - What are effective engagement strategies for increasing viewer interaction and channel growth?
 - Can predictive analysis help identify trends and patterns to inform future content strategies?
 - By answering these questions, the project aims to provide valuable insights and recommendations for optimizing the performance and growth of my YouTube channel.

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
 - Data collection involved utilizing the YouTube API to gather comprehensive information about video performance metrics, including view count, like count, comment count, and video duration.
- Perform data wrangling
 - Data wrangling techniques were applied to clean and preprocess the collected data, ensuring consistency and accuracy. This involved dropping unwanted values, removing duplicates, and standardizing data formats.
 - Perform exploratory data analysis (EDA) using visualization and python
- Perform interactive visual analytics using Plotly Dash
- Perform predictive analysis using classification models

Data Collection

- YouTube API Integration
- Integration of the YouTube API for data retrieval.
- Formulation of queries to extract relevant video metrics.
- Iterative data retrieval process for comprehensive coverage.
- Storage of collected data in AWS DBS.

```
import pandas as pd
from googleapiclient.discovery import build
from datetime import datetime
```

```
[ ] # Define your API key
API_KEY = "XXXXXXXXXX"

# Define your YouTube channel ID
CHANNEL_ID = "XXXXXXXXXX"
```

```
[ ] # Create a YouTube API client
youtube = build('youtube', 'v3', developerKey=API_KEY)

from dateutil import parser

def fetch_video_data(video_id):
    video_response = youtube.videos().list(
        part='snippet,statistics,contentDetails',
        id=video_id
    ).execute()

    if video_response['items']:
        video = video_response['items'][0]
        snippet = video['snippet']
```

```
[ ] #Main

#database credentials
host_name = 'XXXXXXXXXX'
dbname = 'XXXXXXXXXX'
port = 'XXXXXXXXXX'
username = 'XXXXXXXXXX'
password = 'XXXXXXXXXX'
conn = None

#establish a connection to db
conn = connect_to_db(host_name, dbname, port, username, password)
curr = conn.cursor()
```

Connected!

```
[ ] #create table

create_table(curr)
```

```
[ ] #update data for existing videos

new_vid_df = update_db(curr,df)
conn.commit()
```

Data Wrangling

Data Cleaning

- Removal of Youtube shorts videos to ensure data integrity.
- Standardization of data formats for consistency.

Data Transformation

- Conversion of categorical variables into numerical representations.
- Normalization or scaling of numerical features to facilitate analysis.
- Creation of derived features to enhance predictive power.

Data Aggregation

- Grouping of data by relevant attributes for aggregation.
- Calculation of summary statistics such as means, medians, and totals.
- Aggregation of data into meaningful subsets for analysis. variables into numerical representations.

```
[ ] #dropping youtube shorts videos to avoid skewing the data

sorted_data = df_from_db.sort_values(by="video_duration", ascending=False)

# Get the indices of the last 34 rows (sorted by 'video_duration')
indices_to_delete = sorted_data.tail(34).index

# Delete the rows with the identified indices
df_from_db = df_from_db.drop(indices_to_delete)

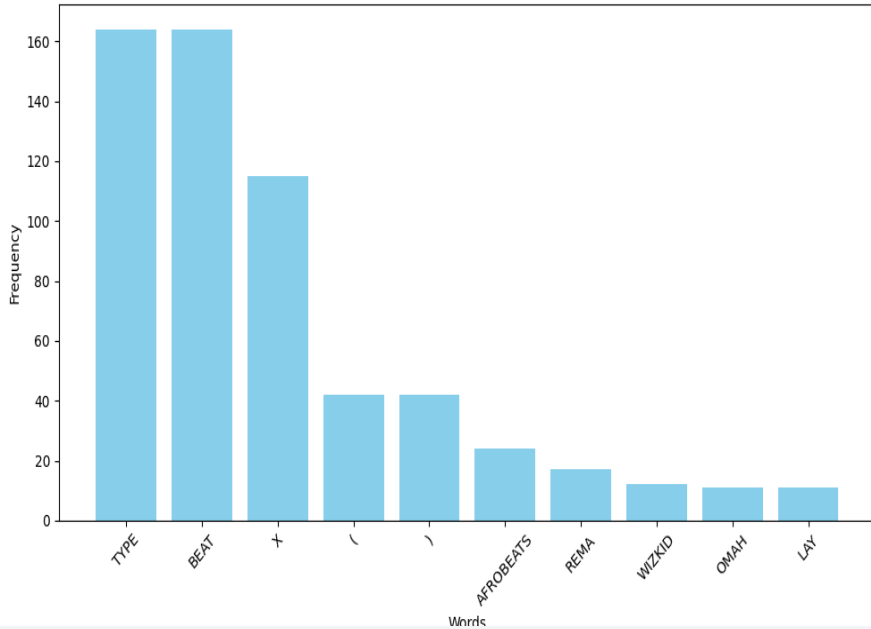
# Verify that the rows have been deleted
print(df_from_db)
```

	video_id	video_title	upload_date	\
0	p_PH5kD6CyM	SHALLIPOPI X COSTA TITCH TYPE BEAT	2023-12-15	
1	s9v_hMQSR8Y	DOUBLE (AFROBEATS TYPE BEAT)	2022-12-16	
2	a0MlTHduXkU	FOLD (AFROBEATS TYPE BEAT)	2021-12-10	
3	hnZNffTECms	ASA X PPRIME TYPE BEAT	2022-06-21	
4	o6TgwAE2CwU	KHAID X OMAH LAY TYPE BEAT	2023-12-09	
..	
200	x505x5sRoGQ	FAVE X TENI TYPE BEAT	2022-02-11	
201	MpbweU3RN9M	(SMOOTH) FIREBOY TYPE BEAT	2021-04-17	
202	Xb5QSouAtQU	TRENT (AFROBEATS TYPE BEAT)	2021-11-12	
203	-g_9TzgD6VI	JOEBOY TYPE BEAT	2020-10-23	
204	CH01aBSK2UI	YOUNG JONN TYPE BEAT	2022-12-02	

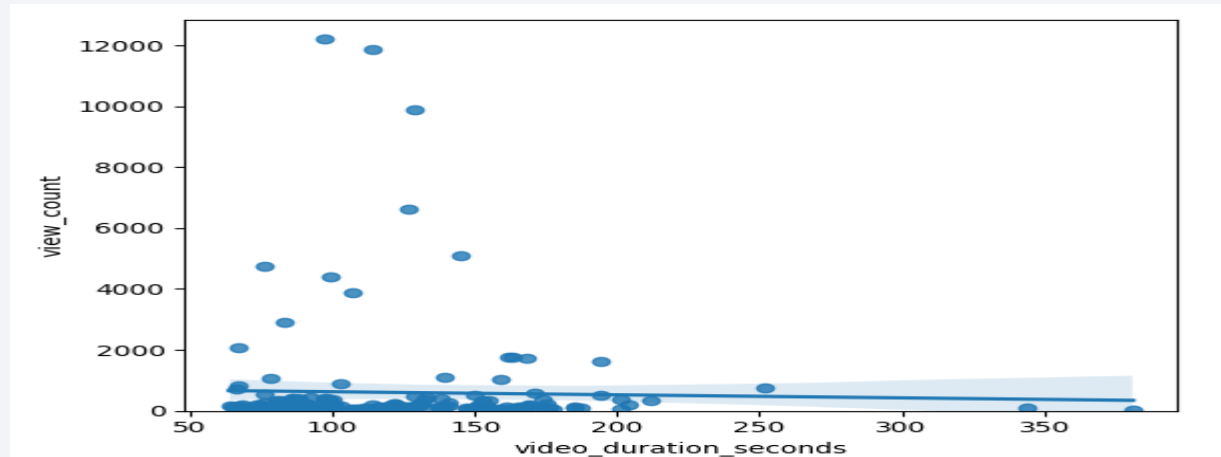
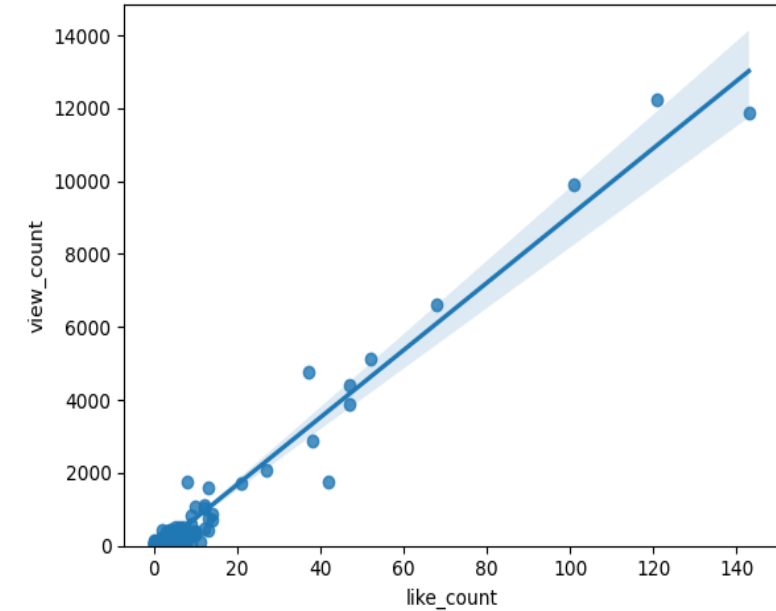
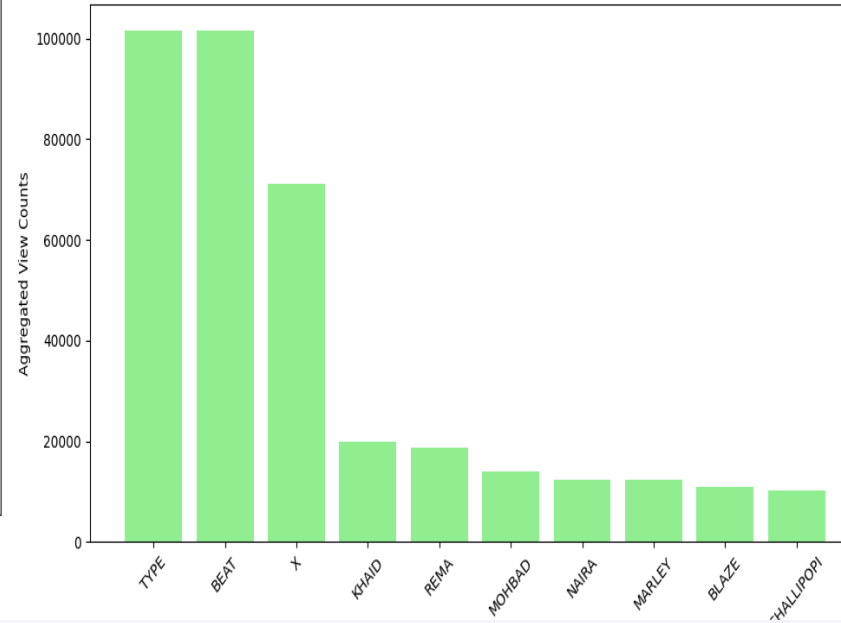
	video_duration	publishing_time	view_count	like_count	comment_count
0	76	05:48:04	4745	37	0

EDA with Data Visualization

Top 10 Most Common Words in Video Titles



Top 10 Words with the Most View Counts in Video Titles



EDA with Data Visualization Continuation

- During the exploratory data analysis (EDA) phase, I utilized data visualization techniques to uncover insights within the dataset. Specifically, I plotted the frequency of the top 10 most common words found in video titles, allowing for a visual representation of their occurrence. Additionally, I created bar charts to showcase the top 10 words with the highest view counts, providing a clear comparison of their popularity. Furthermore, scatter plots were employed to examine the correlation between view count and like count, view count and video duration, revealing any underlying relationships between these metrics. These visualizations offered valuable insights into the characteristics and trends of the data, guiding subsequent analysis and decision-making.

Predictive Analysis (Classification)

- Building the Classification Model
- Selection of classification algorithms such as Random Forest, Logistic Regression, Support Vector Machine, Decision Tree, and K Nearest Neighbors.
- Splitting the dataset into training and testing sets for model training and evaluation.
- Evaluating Model Performance
- Utilization of evaluation metrics such as accuracy score to assess the performance of each classification model.
- Cross-validation techniques to ensure robustness and generalization of the models.
- Improving Model Performance
- Hyperparameter tuning using techniques like grid search or random search to optimize model parameters.
- Feature engineering to enhance model predictive power and reduce overfitting.
- Finding the Best Performing Model
- Comparison of performance metrics across different classification algorithms.
- Selection of the model with the highest accuracy score or best performance on validation data.

Results

- The analysis indicates a need for improvement in my channel's success rate. Successful videos exhibit exceptional performance, suggesting the potential for exponential growth with proper solutions. Encouraging viewers to like and comment on videos, along with engaging with comments and actively sharing content on social media platforms, are effective strategies to increase engagement and ultimately boost view counts. Additionally, collaborating with influencers and creators in my niche can help expand my reach and encourage cross-promotion, further enhancing engagement and visibility. Furthermore, the analysis reveals a weak correlation between shorter video duration and title length with higher view counts. Although the correlation is not very strong, it suggests that viewers may prefer concise and to-the-point content, leading to increased engagement. Identifying the keywords related to the street/hip-hop genre with the highest views suggests a strong interest from my audience in this particular genre. Capitalizing on this insight by creating more content within this genre aligns with my audience's preferences, potentially leading to increased engagement and higher view counts.
- SVM is identified as the most effective method for predictive analysis. Leveraging these insights can contribute to the growth and success of my YouTube channel.



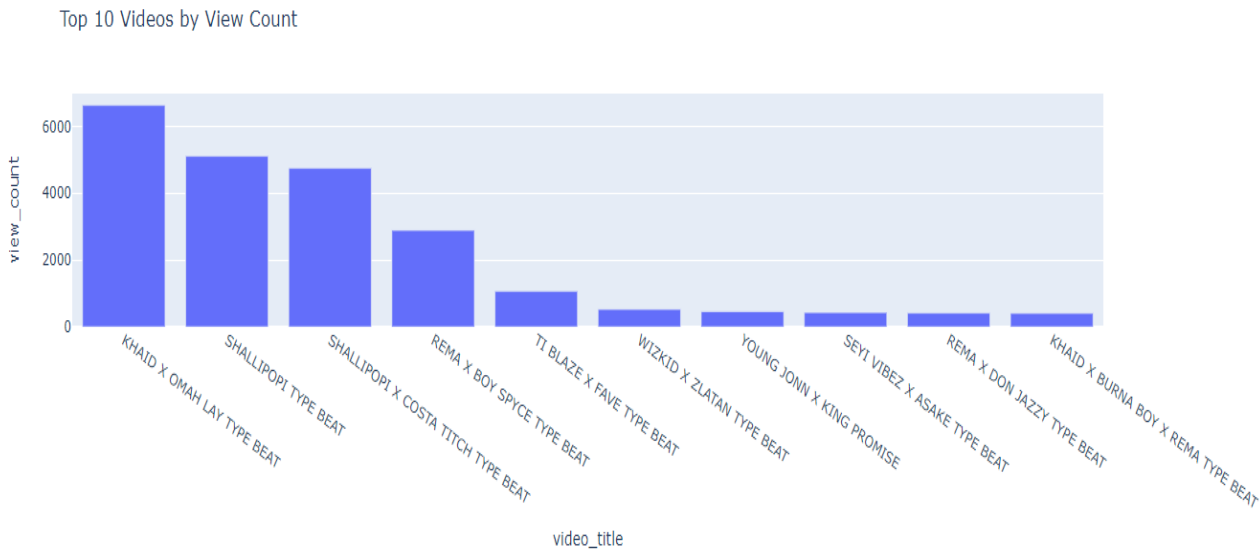
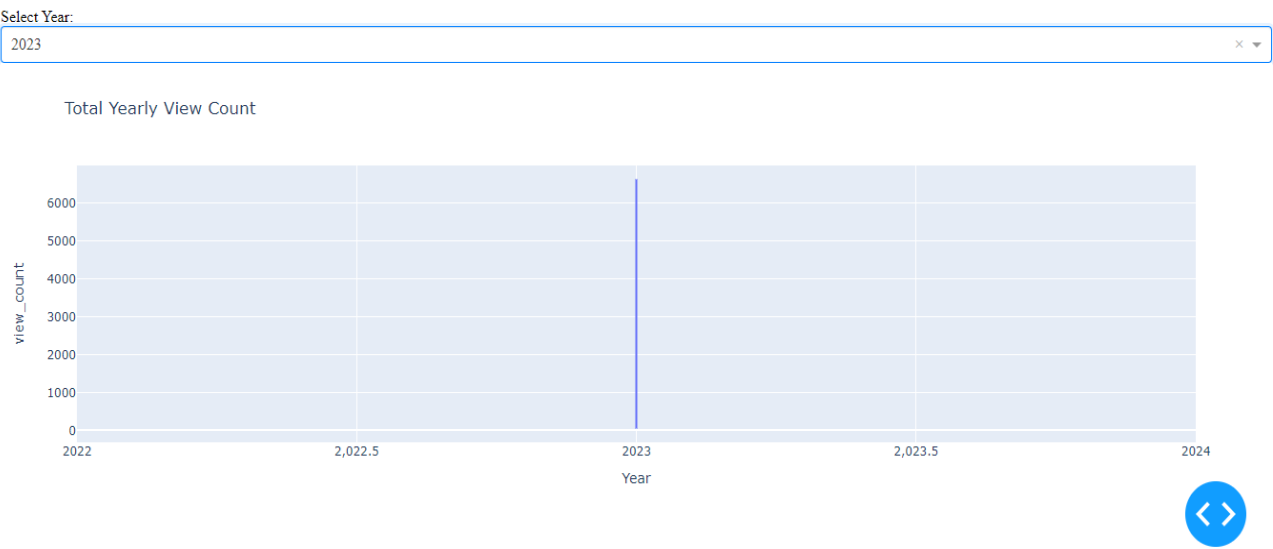
Section 4

Build a Dashboard with Plotly Dash

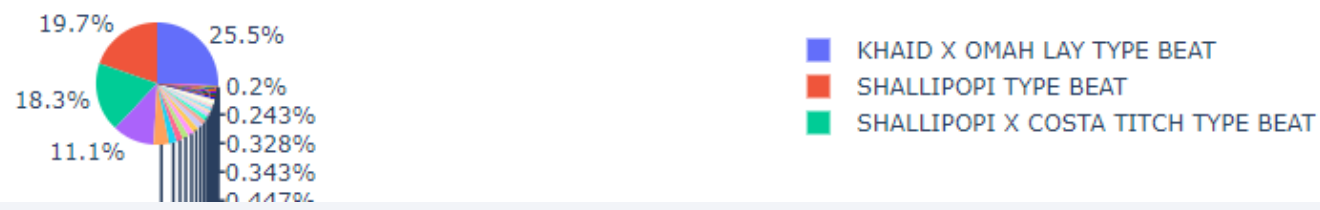
Dashboard Screenshots

Total yearly view count, top 10 videos by view count and view count share by video title for selected year is depicted by line chart, bar chart and pie chart respectively

YouTube Video Analytics Dashboard



View Count Share by Video Title



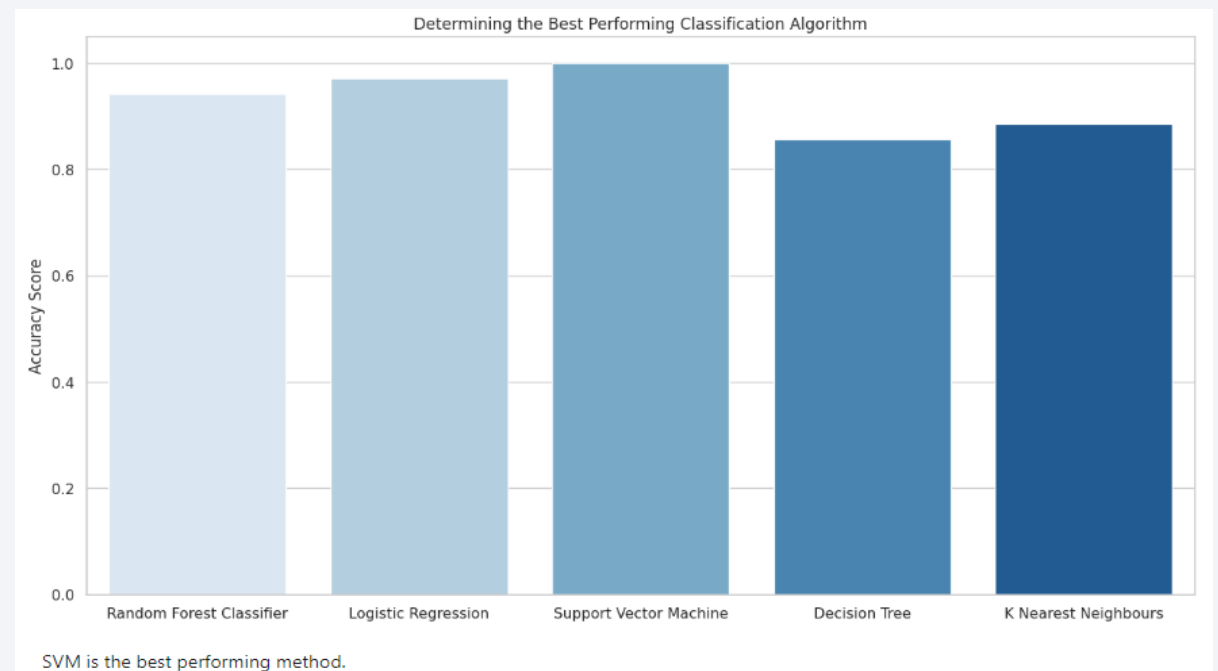
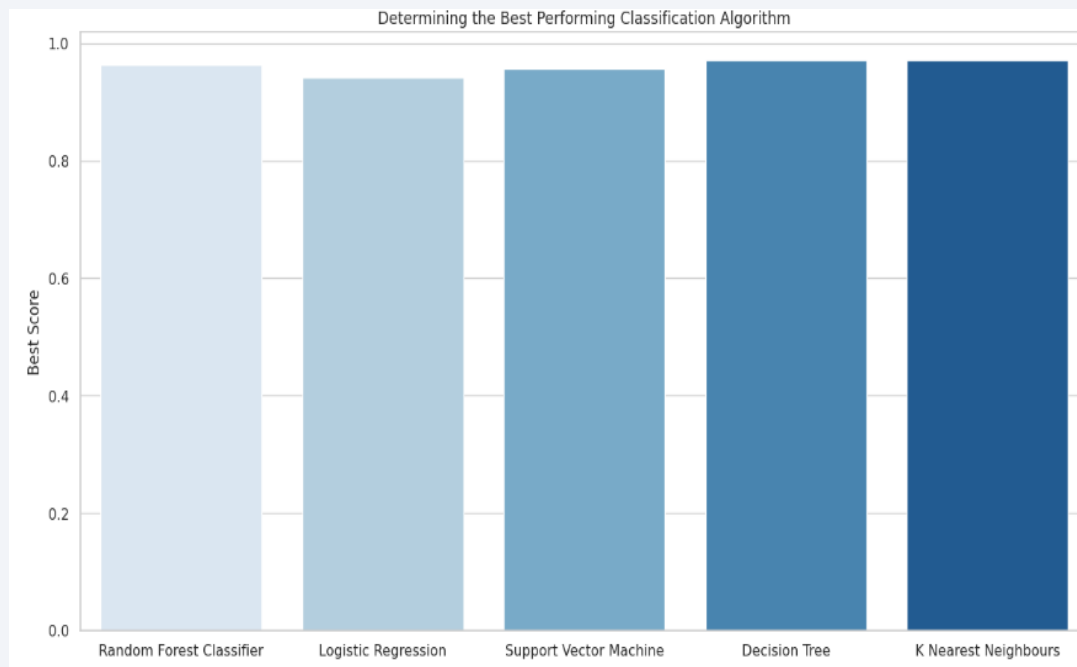


Section 5

Predictive Analysis (Classification)

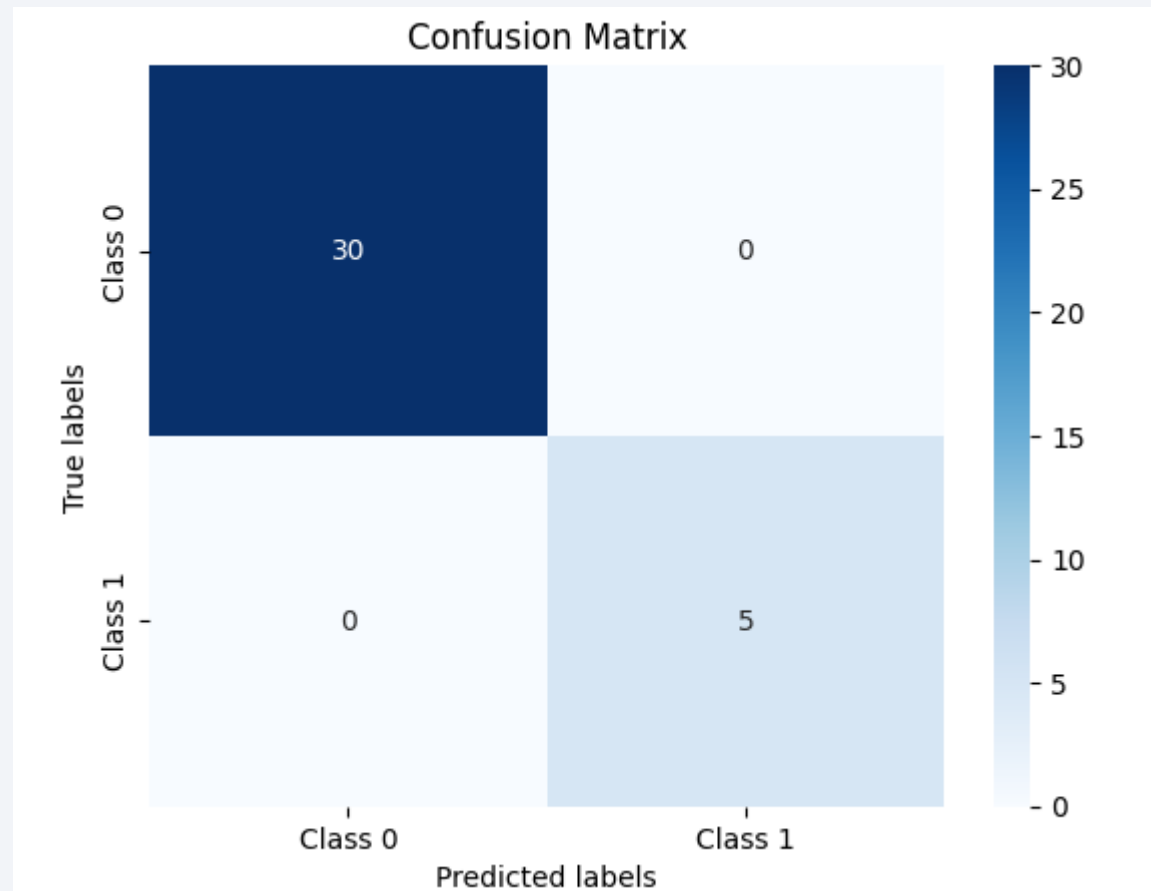
Classification Accuracy

- SVM is the best performing method



Confusion Matrix

- confusion matrix of the best performing model.



Conclusions

- 1. Encourage Engagement: Foster viewer interaction by prompting likes, comments, and social media shares to boost engagement and increase view counts.
- 2. Collaboration Opportunities: Explore partnerships with influencers and creators in the niche to expand reach, encourage cross-promotion, and enhance visibility.
- 3. Concise Content Creation: Consider creating shorter, more focused videos as viewers may prefer concise content, potentially leading to increased engagement and higher view counts.
- 4. Genre-Specific Focus: Capitalize on audience preferences by producing more content within the street/hop genre, aligning content with viewer interests to maximize engagement.
- 5. Utilize SVM Insights: Leverage insights from Support Vector Machine (SVM) analysis for predictive modeling to inform content strategies and drive channel growth effectively.

Appendix

- Sorting and dropping

```
[ ]  
#sorting out YouTube shorts videos by video_duration  
df_from_db.sort_values(by="video_duration", ascending=False).tail(34)
```

```
[ ]  
#dropping youtube shorts videos to avoid skewing the data  
  
sorted_data = df_from_db.sort_values(by="video_duration", ascending=False)  
  
# Get the indices of the last 34 rows (sorted by 'video_duration')  
indices_to_delete = sorted_data.tail(34).index  
  
# Delete the rows with the identified indices  
df_from_db = df_from_db.drop(indices_to_delete)  
  
# Verify that the rows have been deleted  
print(df_from_db)
```

Thank you!

