YOUTUBE TRENDING VIDEO ANALYSIS

1.INTODUCTION

YouTube, the preeminent video-sharing platform, curates a daily list of trending videos, showcasing content that captures the zeitgeist of global and local cultures. The trending algorithm, as reported by Variety magazine, relies not just on views but also on user interactions, such as shares, comments, and likes. These metrics indicate the videos' virality rather than sheer viewership. Our project conducts a comprehensive analysis of YouTube's trending videos data, aiming to unearth insights into what propels a video to trend, the characteristics of trending content, and the dynamics of user engagement.

- **Objectives:**
  The project aims to:
  -Analyze patterns and trends in video performance across different categories.
  -Investigate the correlation between views, likes, comments, and trending duration.
  -Understand the impact of various factors on a video's likelihood to trend.
  -Visualize data distributions and relationships using advanced graphical techniques.
  -Provide insights that can help content creators strategize their content for better engagement and trending potential.

2.ABSTRACT

The dataset is a meticulous aggregation of several months' worth of data on daily trending YouTube videos, specifically tailored for the US region, featuring up to 200 trending videos per day. It includes the following attributes for each video:

- Video ID and Title

- Channel Title

- Publish Time

- Tags

- Views, Likes, Dislikes, Comment Count

- Thumbnail Link

- Comments Disabled Indicator

- Ratings Disabled Indicator

- Video Error or Removed Indicator

- Description

Additionally, the dataset encompasses a category_id attribute, correlating to YouTube's video categories. These categories can be decoded using the accompanying JSON files provided for each region included in the dataset.

This enriched dataset is sourced from Kaggle and is a structurally enhanced version of a previous dataset, ensuring better data integrity and ease of analysis.

3.TOOLS USED

- **Data Manipulation:** Pandas, NumPy

- **Data Visualization:** Matplotlib, Seaborn, Plotly

- **Statistical Analysis:** SciPy

- **Machine Learning:** scikit-learn (for predictive modeling, if applicable)

- **Database Management:** SQL (for larger datasets or integration with web applications)

- **Version Control:** Git, GitHub

4.STEPS INVOLVED IN BULDING THE PROJECT

The project will follow a structured approach:

1. **Data Cleaning:** Handling missing values, removing duplicates, and ensuring data quality.

2. **Data Exploration:** Understanding the basic statistics of the dataset and exploring each variable individually.

3. **Data Preprocessing:** Preparing data for analysis by encoding categorical variables and normalizing numerical values.

4. **Exploratory Data Analysis (EDA):** Identifying trends, patterns, and anomalies in the data.

5. **In-Depth Analysis:** Performing statistical tests and building models to analyze the impact of different variables on trending potential.

6. **Visualization:** Creating interactive charts and graphs to represent findings effectively.

7. **Insights and Recommendations:** Concluding with actionable insights and strategies for content creators.

5.CONCLUSION

The YouTube Trending Video Analysis project provided valuable insights into the factors that influence a video's popularity and engagement on the platform. By leveraging Python for data analysis and visualization, and SQL for efficient data extraction and management, the study successfully identified patterns and correlations between key performance indicators such as views, likes, comments, and video duration.

The analysis revealed that audience engagement is not solely dependent on view count but is also significantly affected by factors like video category, publishing time, and regional preferences. Certain categories—such as music, entertainment, and news—tend to dominate the trending list, highlighting user interest trends across demographics and time periods.