**Project Report: YouTube Trending Video Analytics**

A logo on a black background

AI-generated content may be incorrect.

**Objective:**

This project aims to uncover hidden patterns in trending YouTube videos by analysing datasets from different countries.  
The goal is to understand viewer preferences, sentiment, and content trends regionally and globally in India and the USA.

**Tools Used:**

1. Python: Data cleaning, sentiment analysis (Pandas, NLTK, Matplotlib, Seaborn)
2. SQL: Analytical querying and ranking
3. Power BI/Tableau: Interactive dashboard and data visualization

**Dataset Overview:**

* Data sourced from YouTube trending datasets across multiple countries.
* Each record includes fields such as video\_id, title, channel\_title, publish\_time, trending\_date, views, likes, dislikes, comment\_count, and tags.
* Merged and standardized into a single Excel file: YT\_combined.xlsx

**Data Cleaning and Standardization**

* Combined country-specific datasets into a unified Data Frame.
* Standardized trending\_date and publish\_time into proper datetime formats.
* Removed duplicates and rows with missing critical fields.
* Mapped category\_id into human-readable category names (optional step suggested).

**Sentiment Analysis**

* **NLTK’s VADER** (Valence Aware Dictionary and sEntiment Reasoner) was used.
* Sentiment scores were calculated for:
  + **Video Titles**
  + **Tags**
* Sentiment classification into Positive, Neutral, and Negative categories based on compound scores.

**SQL-Based Analysis**

**Query:** Ranking categories by **Average Views**.

Code:  
SELECT category\_id, AVG(views) AS avg\_views

FROM youtube\_trending

GROUP BY category\_id

ORDER BY avg\_views DESC;

Major Findings: Some content categories (e.g., Music, Entertainment, Gaming) consistently outperformed others in average views across regions.

**Key Visualizations:**

* Channel Dominance byViews (Treemap)
* Top KPIs- Total Views, Avg Likes, Avg Dislikes, Total Trending Videos
* Bar Chart- Trending Category vs Time Duration
* Area Chart- Relation between Title\_sentiment and Tag\_sentiment

**Dashboard Highlights**

* KPIs: Total Views, Total Trending Videos, Avg Likes
* Genre-wise and Region-wise popularity insights
* Sentiment heatmaps and comparisons
* Trending timeline evolution

**Conclusion**

* Music and Entertainment remain dominant across all countries.
* Positive sentiment in titles correlates slightly with higher average views.
* Region-specific preferences affect category rankings (e.g., Gaming is huge in the US, less so in India).
* Trending duration often clusters around 1–5 days for most videos.