



# Final Report: YouTube Trending Video Analytics

---



## Project Objective

**The goal of this project is to uncover regional patterns in YouTube trending videos using data from different countries. The analysis focuses on:**

- Most popular video categories
- Sentiment of video titles
- Regional comparison of trends
- Duration of videos in trending



## Tools & Technologies Used

- Python: Data processing, sentiment analysis, and data visualization
- Libraries: pandas, plotly, sqlite3, textblob / VADER
- SQLite: For querying and ranking categories by average views
- Plotly: Interactive and time-series visualizations
- Datasets: YouTube trending videos (multi-country)


## Data Preprocessing & Cleaning

- Merged and cleaned datasets from multiple countries (e.g., India, USA).
- Converted trending date formats and added country codes.
- Removed nulls and handled duplicates.
- Normalized category names and mapped IDs to readable labels.

## • Exploratory Data Analysis (EDA)

-  **Top Categories by Average Views (via SQLite)**

```
SELECT category, ROUND(AVG(views), 2) as avg_views  
FROM trending_data  
GROUP BY category  
ORDER BY avg_views DESC;
```

 **Result:** *Music, Entertainment, and News* emerged as top categories across most regions.

## ◆ Sentiment Analysis of Titles

- Used TextBlob to analyze polarity of video titles.
- Titles were classified as Positive, Neutral, or Negative.
- Majority were positive, with Music and Entertainment videos showing highest positive sentiment.

## ◆ Duration in Trending (Plotly Time-Series)

- Tracked how many days videos stayed on trending list.
- Created **interactive line graphs** to visualize daily presence.
- Viral videos showed spikes lasting up to **5–7 days**, while others faded quickly.



## Regional Insights & Comparison



### Category Popularity by Region

- Used SQLite joins and grouping to compare category trends.
- **India** had more news and film-related content.
- **US** showed strong interest in vlogs and tech reviews.
- **Music** was universally dominant.



### Sentiment Distribution Across Regions

- Visualized sentiment using **Plotly pie charts** per region.
- **US titles** had more emotionally expressive tones.
- **India and Canada** skewed towards neutral tones.

## Key Insights

### Insight

 Popular Genres

 Sentiment

 Regional  
Variation

 Trending Duration

### Observation

*Music* consistently ranks top in avg views

Most trending titles are positive or neutral

Trends vary: India = News/Film, US = Vlogs/Tech

Most videos stay trending **2–4 days**, viral ones longer

## Interactive Dashboard Highlights

- Plotly Bar Charts: Most viewed categories by country
- Pie Charts: Sentiment split across regions
- Time-Series Line Graphs: Views and trending duration

## Conclusion

This project reveals how YouTube trends differ across regions, categories, and viewer sentiment. The combination of **SQLite for querying** and **Plotly for visuals** makes the analysis interactive and insightful.

### Future Enhancements

- Add **real-time YouTube API data**
- Include **likes/comments for deeper engagement analysis**
- Expand to **more countries** and finer time resolutions

---

## Appendix

- **Data Source:** Kaggle YouTube Trending Dataset
- **Libraries Used:** pandas, sqlite3, plotly, textblob
- **Analysis Date:** [17/05/2025]
- **Author:** [Anubhav Yadav]