YouTube Trending Videos Analysis and Dashboard

By: Jangid Diya

Tool Used: Python (Pandas) & Microsoft Power BI, Jupyter Notebook

Dataset: YouTube Trending Videos Dataset (from Kaggle, Multi-country)

1. Introduction

This project analyzes the trends and patterns in YouTube videos that appeared on the trending list across various countries. Using data from multiple CSV and JSON files provided on Kaggle, I cleaned, merged, and visualized the data using Power BI , SQL to uncover insights into the types of content that trend, the channels that dominate the platform, and how user engagement varies across categories and countries.

2. Objectives

- Combine and clean YouTube trending data from multiple countries
- Map video `category_id` to meaningful names using JSON files
- Create a dynamic dashboard in Power BI to:
  - Identify top categories and channels
  - Visualize viewer engagement metrics
  - Explore time trends and regional variations
- Use storytelling to draw conclusions and make insights accessible

3. Data Preprocessing

• Cleaned and merged CSV files for countries such as India (IN), United States (US), etc.
• Handled missing/null values
• Mapped `category_id` using corresponding `*_category_id.json` files to readable `category_name`
• Removed duplicate video entries and retained most recent trending data
• Exported the cleaned dataset as a single CSV for use in Power BI

4. Dashboard Features (Built in Power BI)

Key Visualizations:
• KPI Cards: Total videos, total views, most common category
• Pie Chart: Distribution of trending videos by `category_name`
• Bar Chart: Top 10 channels by number of trending videos
• Bar Chart: Average views per video by category
• Line Chart: Views trend over time (trending_date)
• Donut Chart: Views number of trending videos by a country

5. Key Findings

Most Popular Categories:
• Entertainment and Music were the top two trending categories across most countries.

• These categories alone accounted for over 40% of all trending content.

Channel Dominance:
• Certain channels, such as T-Series (India) and MrBeast (USA), consistently appeared in the top 10 trending lists.
• These channels frequently trend due to large subscriber bases and high engagement.

Viewer Engagement:
• Videos in the Music and Gaming categories tend to get higher likes-to-views ratios.
• Engagement rate was calculated as:
Engagement = (likes + dislikes + comments) / views

Time Trends:
• There was a noticeable spike in views during weekends and during evening upload times (6–9 p.m.).
• This suggests viewers are more active on weekends and during off-work hours.

Regional Variation:
• India had the highest volume of trending videos, followed by the US.
• The type of content trending varied — e.g., news & politics trended more in the U.S. than in India.

6. Conclusion

This project provides a clear and interactive understanding of what makes a video trend on YouTube. The dashboard offers actionable insights for marketers, content creators, and analysts. It shows how content category, timing, and regional preferences shape viewership patterns on the platform.

By leveraging Power BI's interactivity and Python's preprocessing capabilities, the project demonstrates a complete data pipeline from raw files to actionable storytelling.

7. Future Improvements

- Include sentiment analysis of titles or comments to see if positivity influences trending
- Predict trending potential using a machine learning model (e.g., Random Forest)
- Add a machine learning classification dashboard page to estimate trending likelihood

8. Appendix

• Dataset Source: https://www.kaggle.com/datasets/datasnaek/youtube-new
• Tools Used: Python (for preprocessing), Power BI (for dashboarding)