

Project Name: YouTube Trending Video Analytics

Country Focus: United States

Dataset Source: Kaggle - Trending YouTube Video Statistics

1. Introduction - Project Goals:

- Understand trends in YouTube videos trending in the US.
- Analyze content categories, user engagement, and sentiment.
- Measure how long videos stay trending and what influences popularity.
- Deliver insights through SQL queries, visualizations, and Tableau dashboard

2. Dataset Overview:

<u>Column Name</u>	<u>Description</u>
video_id	Unique YouTube video ID
trending_date	Date when video was trending (YY.DD.MM)
title	Title of the video
channel_title	Name of the channel
category_id	ID of the video category
publish_time	Timestamp of video publication
tags	Tags associated with the video
views	Total views
likes	Number of likes
dislikes	Number of dislikes
comment_count	Number of comments
thumbnail_link	Thumbnail URL
comments_disabled	Comments enabled or not

<u>Column Name</u>	<u>Description</u>
ratings_disabled	Ratings enabled or not
video_error_or_removed	If video has error/removed
description	Description of the video

3. Data Cleaning :

- Removed duplicates using drop_duplicates().
- Converted trending_date and publish_time to datetime format.
- Created publish_hour and trending_duration columns.
- Removed rows with missing titles or views.
- Mapped category_id to readable category names.

4. EDA Insights - Category Trends & Correlations :

- **Most Viewed Categories (Average Views)**
Music, Entertainment, and Sports are the top categories by average views.
- **Likes vs Views**
Strong positive correlation (0.84+), indicating more likes = more views.
- **Comments & Dislikes**
Comments also correlate positively with views.
Dislikes vary less but show engagement.

5. Sentiment Analysis - Title Impact :

- Used VADER Sentiment Analyzer to classify title text.
- Sentiment Score (compound) split into:
 - Positive: > 0.05
 - Neutral: between -0.05 and 0.05
 - Negative: < -0.05

Sentiment Distribution:

- **Positive titles** dominated trending videos (~60%).
- Neutral: ~30%, Negative: ~10%

Insight:

- Videos with **positive or neutral titles** tend to receive more views.

6. Dashboard Visuals (from Power Bi)

- Bar Chart: Top 10 categories by average views
- Pie Chart: Title sentiment distribution
- Time-Series Line: Views over publish date
- Filter by Category, Channel, Sentiment
- Map View (if multi-country data added later)

7. Conclusion & Insights

- **Music and Entertainment** dominate trending content in the US.
- **Positive sentiment** titles are more likely to trend.
- **Views, likes, comments** are tightly correlated.
- Shorter trending duration is common, with few viral exceptions.

GitHub Link: