# YouTube Trending Video Analytics Report

## Introduction

The rise of social media platforms, especially YouTube, has revolutionized the way content is consumed worldwide. This project aims to analyse trending YouTube videos across multiple countries to uncover meaningful patterns and insights related to viewership trends, user sentiment, and content category popularity. It explores regional differences, sentiment impacts, and temporal trends using a combination of data processing, sentiment analysis, and interactive dashboarding.

## Abstract

This project investigates patterns in trending YouTube videos by analysing datasets from different from countries like the USA, India, Great Britain, and more. By merging, cleaning, and enriching the data (e.g., via sentiment analysis), key insights were derived and visualized. A Power BI dashboard displays metrics like total views, comments, likes, and sentiment distribution, offering stakeholders a clear understanding of viewer preferences and regional content behaviour.

## Tools Used

- **Python**: Data cleaning, merging, sentiment analysis
- **SQL (SQLite)**: Category-level aggregation and ranking
- **TextBlob**: Sentiment analysis
- **Power BI**: Dashboard and visual storytelling

## Steps Involved in Building the Project

### 1. Data Collection

- Merged trending YouTube video datasets from multiple countries (US, GB, IN, etc.)
- **Data Source:** Trending YouTube Video Statistics from Kaggle

  https://www.kaggle.com/datasets/datasnaek/youtube-new?resource=download

2. **Data Cleaning and Preprocessing**
    - Removed unnecessary columns
    - Standardized column names and values (e.g., country codes)
    - Converted date formats (trending_date and publish_time)
3. **Sentiment Analysis**
    - Combined video title and tags
    - Used **TextBlob** to generate sentiment scores
    - Labelled each video as **Positive**, **Negative**, or **Neutral**
4. **SQL Analysis**
    - Loaded cleaned data into SQLite
    - Ranked category_id by average views using SQL window functions
5. **Dashboard Creation (Power BI)**
    - Created interactive visuals for:
        - Region-wise views, likes, comments
        - Sentiment distribution
        - Time-series of views and trends
        - Category-wise engagement

### Key Insights

- Great Britain had the highest views and likes
- Most videos had **neutral sentiment**, followed by **positive**
- Categories with IDs between 10–15 had highest views
- **May 2018** saw a peak in video views

## Conclusion

This project demonstrates the power of combining Python, SQL, and BI tools to end-to-end YouTube trending video analytics pipeline. The results highlighted geographic viewership trends, sentiment polarity of content, and category popularity. The dashboard offers interactive exploration for data-driven decision-making in content strategy and marketing, which can help content creators, marketers, and analysts to better understand viewer preferences globally.