YouTube Trending Video Analytics

Introduction

YouTube is one of the largest platforms for video sharing and content consumption globally. Millions of videos are uploaded daily, with content spanning categories like entertainment, education, music, and more. Understanding what drives a video to trend - be it likes, views, comments, or sentiment - is valuable for creators and analysts. This project focuses on analyzing trending YouTube video data across countries (e.g., India and the US) to extract meaningful patterns and insights.

Abstract

The goal of this project is to perform a comprehensive analysis of YouTube trending videos using a combination of SQL, Python, and Power BI. The dataset contains attributes such as video titles, categories, likes, comments, and sentiment labels. Key objectives include:

- Identifying the most liked and viewed content categories.
- Analyzing user sentiment distribution across countries.
- Visualizing trends through a dynamic dashboard.
- Comparing engagement metrics (likes, comments) based on sentiment.

Tools Used

- MySQL: Used to write and execute SQL queries to extract and manipulate data.
- Power BI: Created data visualizations and the final interactive dashboard.
- Python (Jupyter Notebook): For preprocessing, initial exploration, and possibly sentiment tagging oranalysis.
- Pandas & Matplotlib: Used in Python for data handling and early visualization.

Steps Involved in Building the Project

1. Data Loading and Cleaning

- Loaded a cleaned version of YouTube trending video data.
- Ensured there were no duplicates or missing values.
- Added readable category names using category ID mappings.

2. Sentiment Labeling

- Each video or associated comments were labeled as positive, neutral, or negative.
- This data was used to analyze how sentiment impacts likes and comment counts.

3. SQL-Based Data Analysis

- Extracted insights using SQL queries (as seen in Youtube_Analytics.sql):
- Ranked video categories by average views across countries.
- Counted number of videos under each sentiment per country.
- Retrieved top trending videos by view count.

4. Visual Dashboard Creation (Power BI)

- Designed a multi-chart dashboard (Youtube_Trend_Analysis.png) including:
- Pie chart of likes by sentiment.
- Line chart showing likes by country.
- Donut chart of comment count by country.
- Bar chart comparing likes and comment count by sentiment.
- Card and gauge visuals to display total comment count.

5. Interpretation of Insights

- Neutral content received the highest likes.
- The US dominated in total likes and comment counts compared to India.
- Videos with neutral and positive sentiments generally had better engagement.

Conclusion

This project offered a data-driven perspective on what drives YouTube videos to trend. By leveraging SQL for querying, Python for preprocessing, and Power BI for visualization, the project effectively uncovers viewer behavior and engagement patterns across sentiments and geographies. The final dashboard enables quick, insightful decision-making for stakeholders such as content creators, marketers, and analysts aiming to understand and capitalize on YouTube trends.