YouTube Trending Video Analytics - Internship Project Report

Introduction

This project analyses YouTube trending video datasets from the United States and India. The goal is to uncover regional content patterns, sentiment trends, and performance metrics using data analytics technique.

Abstract

The dataset includes trending YouTube videos from both the US and India. We performed data cleaning, sentiment analysis on video titles, and SQL-based insights to identify trending patterns. The analysis compares regional behaviours, identifies top categories, and visualizes sentiment distributions.

Tools Used

- Python (Pandas, Seaborn, Matplotlib, NLTK)

- SQLite for SQL Analysis- Jupyter Notebook

Steps Involved

1. Loaded and merged US & India datasets.
2. Cleaned and processed timestamp and tag fields.
3. Applied VADER sentiment analysis to video titles.
4. Added sentiment labels (Positive, Neutral, Negative).
5. Loaded data into SQLite and ran SQL queries.
6. Generated visualizations for category trends and sentiment comparisons.

Key Insights

- India had a higher proportion of neutral and positive sentiments in video titles.
- Music and Entertainment were the most trending categories across both regions.
- Time-series data showed clear peaks in trending activity.
- SQL queries confirmed insights from EDA and supported structured aggregation.
- Internship Project Report

Conclusion

The YouTube Trending Video Analytics project successfully highlights the power of combining Python and SQL for data storytelling. The findings can help content creators and marketers better understand viewer behaviour in different regions.