YouTube Trending Video Analytics

Introduction

With millions of videos uploaded daily, YouTube serves as a powerful platform for content creators and viewers alike. This project aims to uncover key trends in YouTube videos that go viral across different regions. By analyzing trending data from the United States and India, we aim to understand content categories, viewer engagement, and audience sentiment.

---

2. Abstract

This project analyzes publicly available YouTube trending data from Kaggle. We used Python to clean and process the data, perform sentiment analysis, and extract patterns. SQL was used to structure the data for querying, and Power BI was used to visualize insights. We explored the correlation between content categories and popularity, identified top-performing genres, and compared sentiment scores across regions.

---

3. Tools Used

- Python (Pandas, NumPy, Matplotlib, Seaborn, TextBlob)
- SQL (SQLite with SQLAlchemy)
- Power BI
- Jupyter Notebook / VS Code

---

4. Steps Involved in Building the Project

1. Loaded and cleaned US and India trending video datasets
2. Mapped category_id to category_name using JSON metadata
3. Converted trending_date to datetime format for time-series analysis
4. Performed sentiment analysis on video titles using TextBlob
5. Aggregated views, likes, and comments by category and region
6. Created SQL database to store cleaned and structured data
7. Exported summarized data for visualization
8. Built interactive Power BI dashboards for comparison and storytelling

---

5. Conclusion

We found that Entertainment and Music were among the top trending categories in the US, while News and People & Blogs dominated in India. Sentiment analysis showed region-wise variation in viewer engagement tone. The use of Power BI helped visualize the data interactively, enabling actionable insights from the trends observed. This project serves as a blueprint for analyzing digital content performance across demographics.
