

YouTube Trending Videos Analysis – Final Report

Introduction

This project analyzes YouTube trending videos to uncover insights into user engagement, popular categories, and the speed at which videos reach trending status. The objective was to clean raw data, explore patterns, and build an interactive dashboard for storytelling and decision-making.

Abstract

The dataset contained multiple attributes of trending YouTube videos such as views, likes, dislikes, comments, categories, and publishing dates. Through systematic cleaning, exploratory data analysis, and visualization, the project highlights viewing trends, engagement ratios, and performance by category, channel, and country. The final dashboard provides an interactive platform for analyzing video performance and audience behavior.

Tools Used

Power BI – Data cleaning, transformation (Power Query), and interactive dashboards. **Python (Jupyter Notebook)** – Data preprocessing with Pandas; visual analysis with Matplotlib and Seaborn. **SQL** – Data extraction, aggregation, and validation of trends.

Steps Involved in Building the Project

Data Cleaning: Removed duplicates, standardized broken dates, converted numeric fields, and trimmed text fields. **Feature Engineering:** Created calculated columns such as *days_to_trend*, *like_ratio*, and *engagement_rate*. **Exploratory Data Analysis (EDA):** Used Python (Matplotlib, Seaborn) for time-series plots, category comparisons, and correlation studies. **SQL Queries:** Performed aggregations (views, likes, comments per category/country) to cross-check accuracy. **Dashboard Development in Power BI:** Built two-page dashboard — Page 1: Overview (trends over time, top categories, country map). Page 2: Engagement & Performance (like vs dislike ratios, scatter plots, top channels).

Conclusion

The analysis revealed that Music and Entertainment dominate trending lists, while Comedy and Gaming trend faster. Engagement rates vary by category, with News often polarizing audiences. The project demonstrates how combining SQL, Python, and Power BI provides a comprehensive workflow for cleaning, analyzing, and visualizing large datasets effectively.