YouTube Trending Video Analytics –Report

Date: 2025-08-25

# Abstract

This study analyzes trending YouTube videos across multiple regions to uncover genre popularity, sentiment dynamics, and timing effects. After standardizing datasets and enriching them with VADER sentiment on titles and tags, we rank categories by average views using SQL and model trending duration as a time-based signal.

# 1. Introduction

Context, motivation, and objectives.

# 2. Data & Methodology

Datasets, cleaning & standardization, sentiment analysis (VADER), SQL ranking, and time-series methods.

# 3. Results

Key findings with figures and tables.

## 3.1 Genre Popularity

## 3.2 Sentiment vs Performance

## 3.3 Trending Duration & Timing

## 3.4 Regional Comparisons

# 4. Recommendations

Actionable guidance by region/category.

# 5. Limitations & Future Work

Data biases, hidden dislikes post-2021, potential model improvements.

# Appendix

Table dictionary, SQL snippets, feature definitions.

**WORK I HAVE DONE SO FAR**

1. Data model

Staging (raw CSVs per country)

stg\_trending\_<CC>: 1 row per video per day it appears in Trending for country <CC>

Typical columns: video\_id, title, published\_at, channel\_title, category\_id, tags, view\_count, likes, dislikes, comment\_count, trending\_date

Core

trending (union of all countries, cleaned)

tags\_long (one tag per row)

dim\_video\_country (first/last trending snapshot per video × country)

sentiment\_lexicon (word → sentiment score; load a VADER-like CSV)

title\_tokens (title tokenization results per video × country × trending\_date)

sentiment\_title and sentiment\_tags (aggregated sentiment per video × country)

---

1. DDL

BigQuery

-- Core fact

CREATE OR REPLACE TABLE dataset.trending (

video\_id STRING,

country STRING, -- ISO-2

trending\_date DATE,

published\_at\_utc TIMESTAMP,

title STRING,

channel\_title STRING,

category\_id INT64,

category STRING,

view\_count INT64,

likes INT64,

dislikes INT64,

comment\_count INT64

);

-- Tags (long form)

CREATE OR REPLACE TABLE dataset.tags\_long (

video\_id STRING,

country STRING,

trending\_date DATE,

tag STRING

);

-- Distinct video-country lifespan + counters

CREATE OR REPLACE TABLE dataset.dim\_video\_country (

video\_id STRING,

country STRING,

first\_trending\_at DATE,

last\_trending\_at DATE,

trending\_days INT64

);

-- Sentiment lexicon (upload a CSV with columns: token, score)

CREATE OR REPLACE TABLE dataset.sentiment\_lexicon (

token STRING, -- lowercase unigram

score FLOAT64 -- negative..positive

);

-- Tokenized titles per row

CREATE OR REPLACE TABLE dataset.title\_tokens (

video\_id STRING,

country STRING,

trending\_date DATE,

token STRING

);

-- Aggregated sentiment

CREATE OR REPLACE TABLE dataset.sentiment\_title (

video\_id STRING,

country STRING,

trending\_date DATE,

title\_sentiment FLOAT64,

title\_sentiment\_label STRING

);

CREATE OR REPLACE TABLE dataset.sentiment\_tags (

video\_id STRING,

country STRING,

trending\_date DATE,

tag\_sentiment FLOAT64,

tag\_sentiment\_label STRING

);

PostgreSQL

CREATE TABLE trending (

video\_id TEXT,

country CHAR(2),

trending\_date DATE,

published\_at\_utc TIMESTAMPTZ,

title TEXT,

channel\_title TEXT,

category\_id INT,

category TEXT,

view\_count BIGINT,

likes BIGINT,

dislikes BIGINT,

comment\_count BIGINT,

PRIMARY KEY (video\_id, country, trending\_date)

);

CREATE TABLE tags\_long (

video\_id TEXT,

country CHAR(2),

trending\_date DATE,

tag TEXT

);

CREATE TABLE dim\_video\_country (

video\_id TEXT,

country CHAR(2),

first\_trending\_at DATE,

last\_trending\_at DATE,

trending\_days INT

);

CREATE TABLE sentiment\_lexicon (

token TEXT PRIMARY KEY,

score DOUBLE PRECISION

);

CREATE TABLE title\_tokens (

video\_id TEXT,

country CHAR(2),

trending\_date DATE,

token TEXT

);

CREATE TABLE sentiment\_title (

video\_id TEXT,

country CHAR(2),

trending\_date DATE,

title\_sentiment DOUBLE PRECISION,

title\_sentiment\_label TEXT

);

CREATE TABLE sentiment\_tags (

video\_id TEXT,

country CHAR(2),

trending\_date DATE,

tag\_sentiment DOUBLE PRECISION,

tag\_sentiment\_label TEXT

);

---