

# Exploratory Data Analysis: Taylor Swift YouTube Channel

August 6, 2024

Tran Kien Quoc, Le Hoang Trang Nha

## Introduction

This report presents an exploratory data analysis of the Taylor Swift YouTube channel. The analysis covers video metrics including views, likes, and comments, as well as the distribution of regular videos versus shorts.

## Data Preparation

We began by importing the necessary libraries and loading the data from ‘data.csv’. The following steps were taken to prepare the data:

- Converted date strings to datetime objects
- Converted numeric strings to appropriate numeric types
- Replaced null values in the description column with empty strings
- Filled missing ‘likes’ values with the median

## Dataset Overview

The dataset contains [Number of rows] rows and [Number of columns] columns. Key attributes include:

- id: Unique identifier for each video
- published\_date: The date the video was published
- title: The title of the video
- description: The video description
- views: Number of views
- likes: Number of likes
- comments: Number of comments
- is\_short: Boolean indicating if the video is a short

## Univariate Analysis

### Categorical Variables

Distribution of short videos

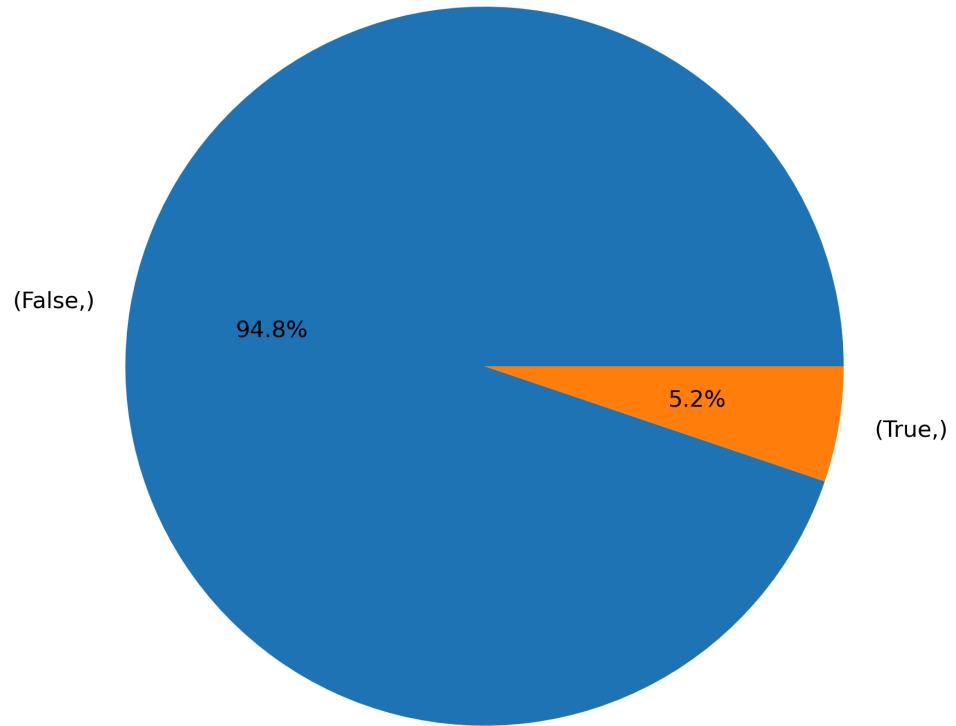


Figure 1: Distribution of Short Videos

## Numerical Variables

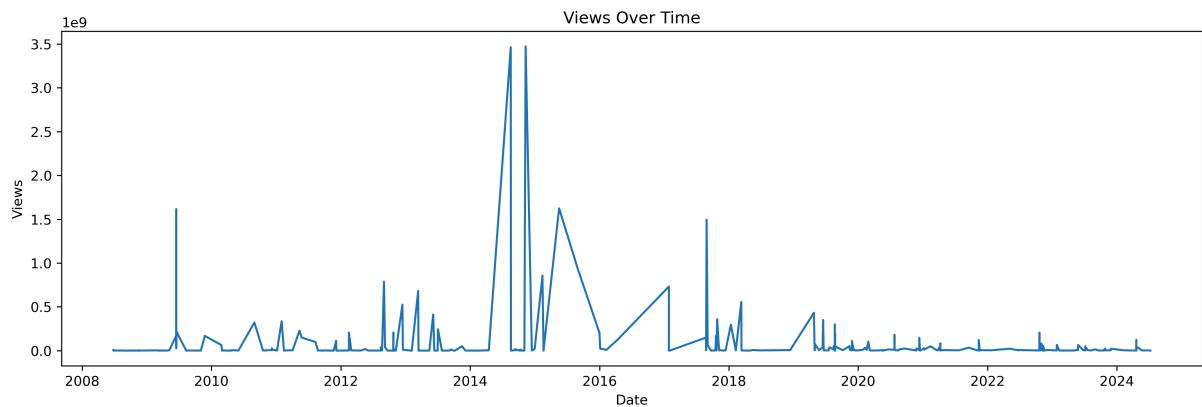


Figure 2: Views Over Time

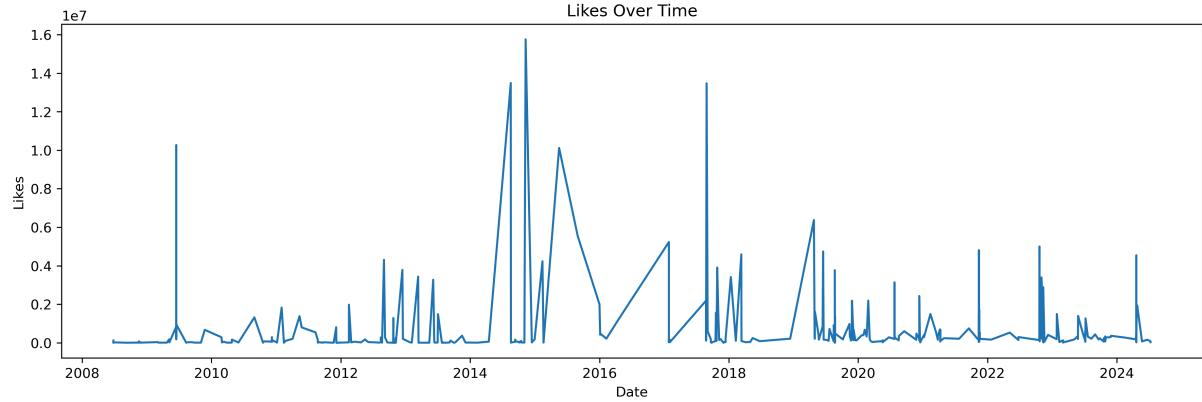


Figure 3: Likes Over Time

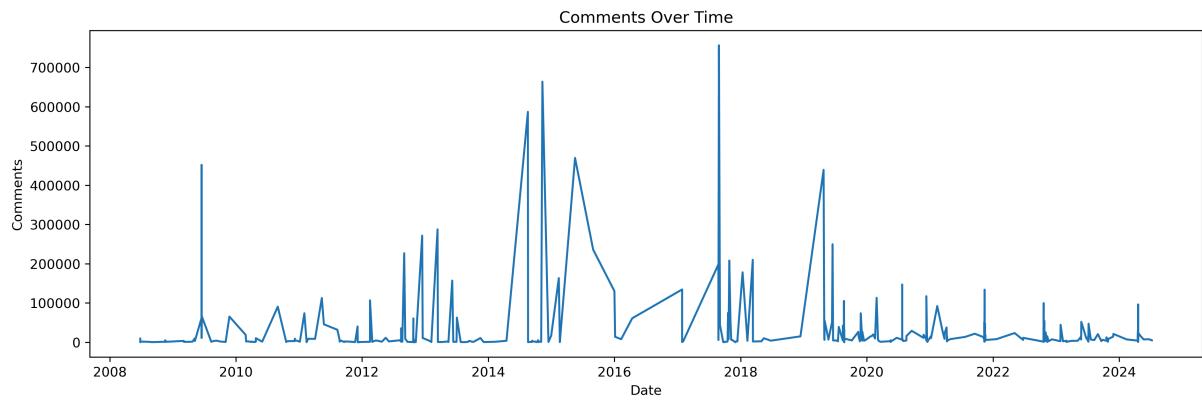


Figure 4: Comments Over Time

## Key Statistics

- Total videos: 578
- Total views: 30,244,783,942
- Total likes: 289,633,946
- Total comments: 11,638,854
- Average views per video: 52,326,616
- Average likes per video: 501,097
- Average comments per video: 20,136

## Top Performing Videos By Views

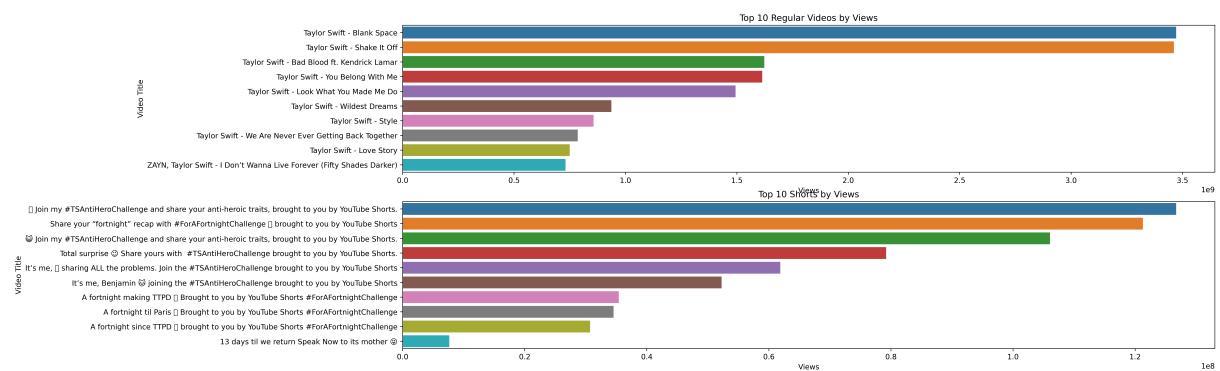


Figure 5: Top 10 regular and shorts videos by Views

## By Likes

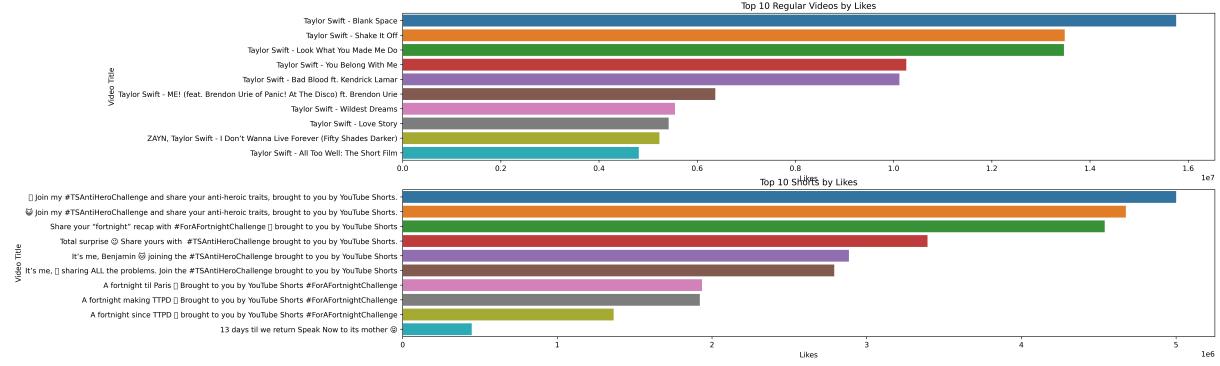


Figure 6: Top 10 regular and shorts videos by Likes

## By Comments

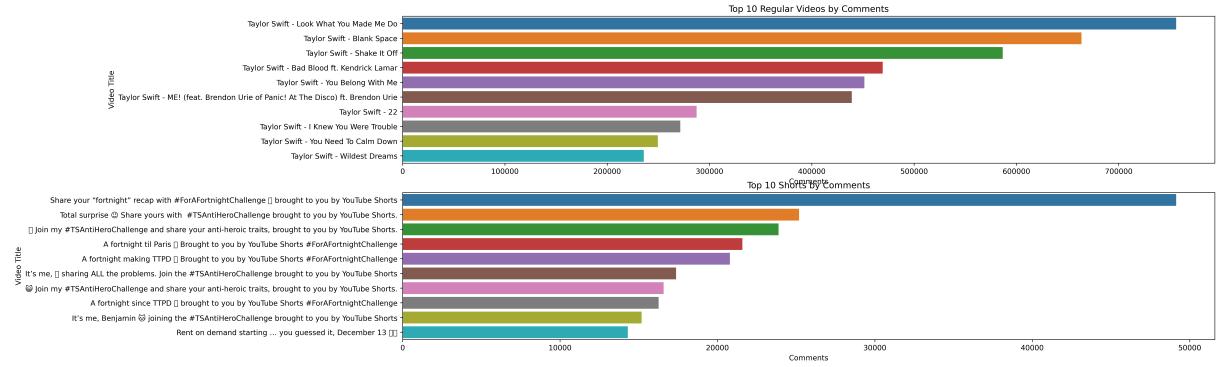


Figure 7: Top 10 regular and shorts videos by Comments

## Shorts vs Regular Videos Analysis

### Performance Comparison

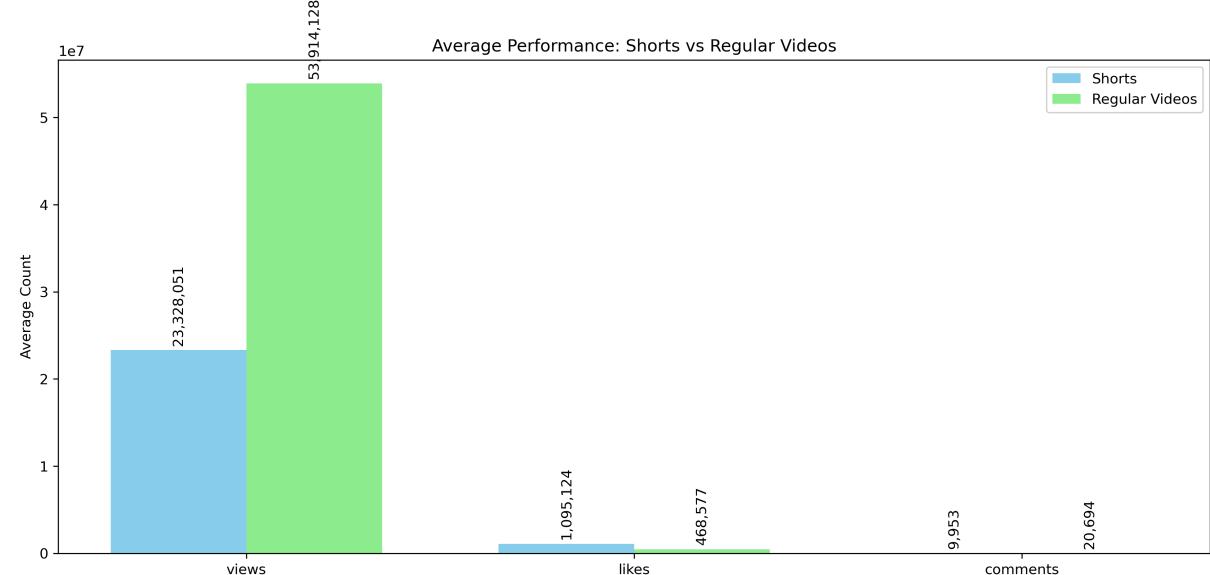


Figure 8: Average Performance: Shorts vs Regular Videos

## Posting Frequency

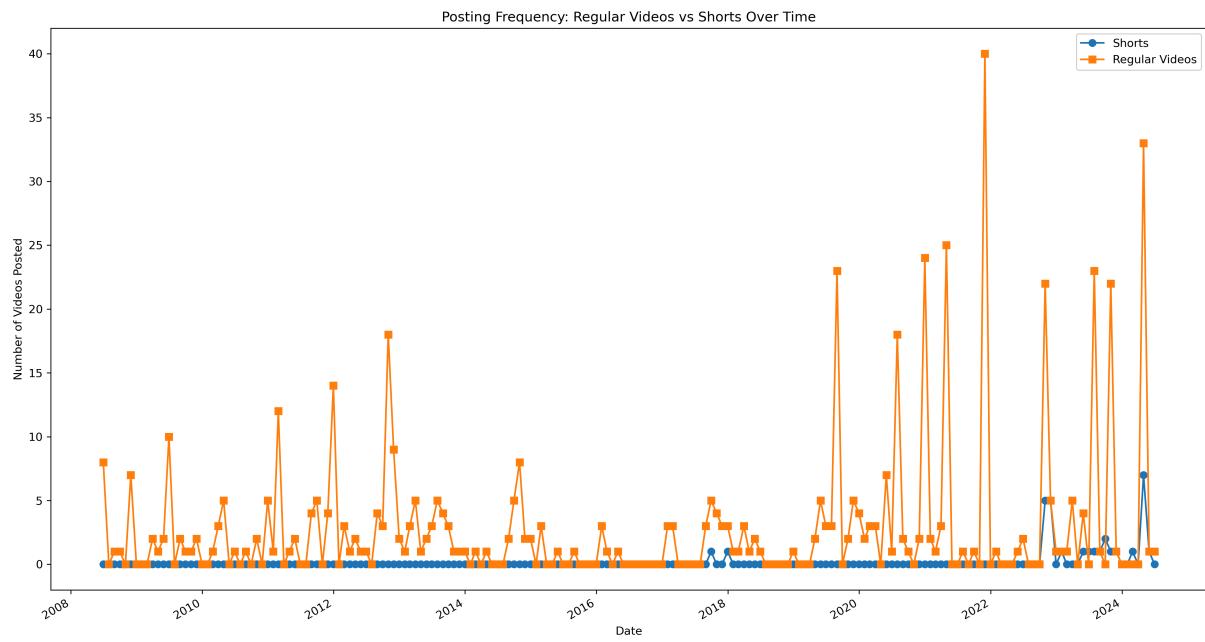


Figure 9: Posting Frequency: Regular Videos vs Shorts Over Time

## Correlation Analysis

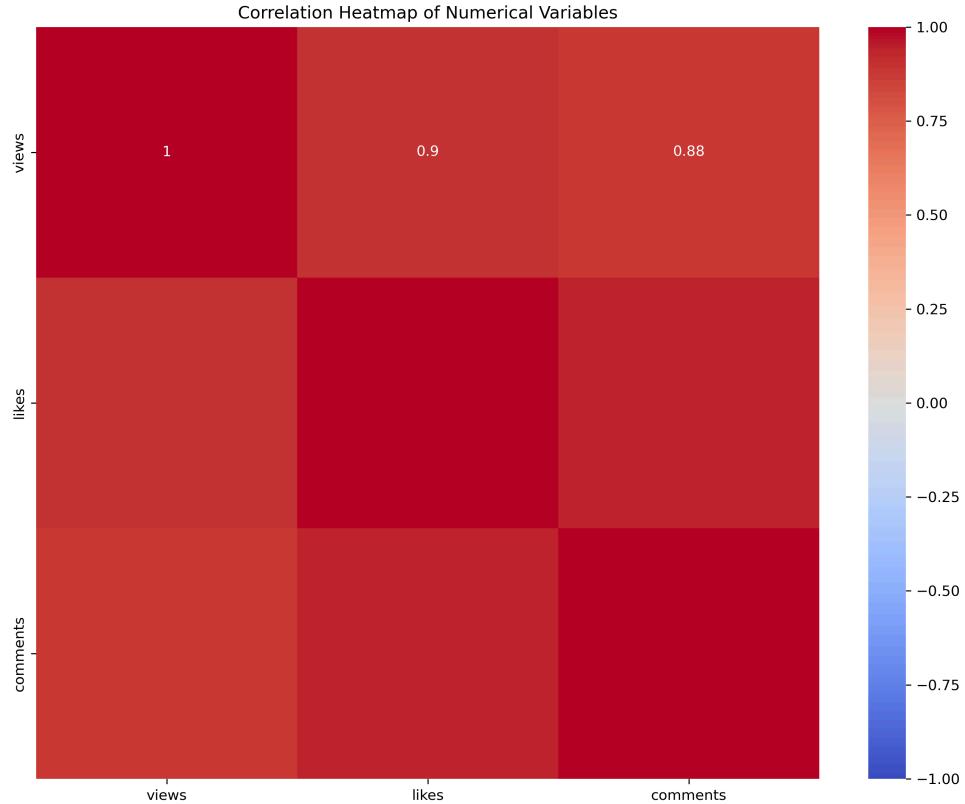


Figure 10: Correlation Heatmap of Numerical Variables