King Saud University
College of Computer and Information Sciences
Information Technology Department

IT362- Data Science

CrashCourse Channel Analysis

Prepared by:

<Bashair Alsadhan, 443200668>
<Noura Alwohaibi, 443200415>
<Waref ALyousef, 442200377>
<Rama Alshebel, 443200929>
<Rana AlSayyari, 443200565>

Supervised by:
I.Reem Alqifari

Contents

Introduction

In the competitive realm of YouTube, educational channels like CrashCourse face challenges in maximizing audience reach. A key concern for CrashCourse is the inconsistency in viewership across its videos, worsened by YouTube's evolving algorithms. This project aims to analyze CrashCourse's video data, comprising 1514 observations extracted by the YouTube API, to uncover insights and strategies aimed at enhancing the channel's impact. By examining videos features such as title, views, likes, and comments, and answering five key questions, we aim to provide actionable recommendations to optimize content strategy, increase viewer engagement, and amplify CrashCourse's presence on YouTube.

This report will outline the problem we seek to address, present our data source, describe the data and its attributes, and formulate five questions to guide us in resolving the problem. These questions will be addressed in subsequent phases.

Problem

Any channel on Youtube faces the challenge of maximizing its viewership, and educational channels have that twice as hard. CrashCourse channel specifically faces a problem of fluctuating views between videos, which raises issues with Youtube's constantly new algorithms and recommendation system. It is important to understand why certain videos have higher views than others, and what patterns lie within in order to formulate effective and actionable strategies to enhance videos performances and expand audience reach, thus improving CrashCourse's overall impact on the platform.

Data source

The data for our project is sourced from the YouTube API, which provides access to various metrics and information regarding videos, channels, and user interactions on the YouTube platform, including data related to the CrashCourse channel. We can obtain important insights into the channel's performance, audience engagement, and general dynamics by utilizing the extensive data made available by the YouTube API. This will help with analysis and decision-making.

About The Data

Our dataset contains data about the CrashCourse channel on YouTube, a renowned platform known for its high-quality educational videos. We collected 1514 observations, with each row representing a video from the channel. Each video is characterized by nine features:

1. **Title**: The title of the video.
2. **ID**: A unique identifier assigned to each video for identification purposes.
3. **Published_date**: The date and time when the video was published on the channel.
4. **Tags**: Any tags associated with the video, providing insight into its content.
5. **Views**: The number of views the video has garnered.
6. **Likes**: The count of likes received by the video.
7. **Comments**: The number of comments posted on the video.
8. **Duration**: The duration of the video.
9. **Captions**: A binary indicator representing whether captions are available for the video.

"Title" and "ID" offer qualitative nominal data. "Published_date" provides quantitative interval data. Tags offer qualitative nominal insights. Metrics like "Views," "Likes," and "Comments" give quantitative ratio data on viewer engagement. "Duration" also provides quantitative ratio data, indicating video lengths. Lastly, "Captions" offers qualitative binary data, signaling caption availability.

This dataset offers a comprehensive overview of the CrashCourse channel's video content, enabling analysis of viewership metrics, engagement levels, and content characteristics. It provides valuable insights for understanding audience preferences, engagement trends, and the overall impact of educational content on the platform.

GitHub link: YouTube API

Our Questions

1. Does the timing of a video's posting affect its viewership on CrachCourse Channel?

2. Are likes and comments correlated with views on CrachCourse Channel?

3. Does the duration of a video affect its viewership on CrachCourse Channel?

4. Does the length of a video's title influence its views on CrachCourse Channel?

5. Do the number of tags affect the viewership of a CrachCourse Channel video?