Abstract

The goal of this project is to use YouTube API data to examine videos features (i.e. duration, likes, dislikes) in order to increase YouTube user engagement. After scraping and cleaning the data, I built an interactive dashboard in Tableau and communicated the results in a presentation.

Design

The purpose of the project will be to identify the video features (i.e. duration, genre, likes, dislikes, number of comments, view count) of the most successful videos created by YouTubers. The data science model will predict the likes of videos based on a feature (i.e. duration) in order to increase user engagement.

Data

The YouTube dataset was built by scraping video features from over 1000 videos from 26 popular videos in 8 genres. The genres were Music, Tech, Education, Entertainment, Sports, How-To, News and Nonprofit, and Gaming.

Algorithms

- 1. Dropping rows with null values for likes.
- 2. Applying the YouTube API to scrape video features.
- 3. Converting duration from ISO 8601 format to minutes.
- 4. Converting date from ISO 8601 format to standard date format.
- 5. Visualizing the data and interpreting the results.

Tools

- 1. Google Sheets for scraping the data with YouTube API and cleaning.
- 2. Tableau for building an interactive dashboard and linear model.

Communication

The results were communicated in a presentation.





