King Saud University

College of Computer and Information Sciences

Department of Information Technology

كلية علوم الحاسب والمعلومات
قسم تقنية المعلومات

IT 362 Course Project
Semester-2, 1446H

Analyzing the popularity of gaming videos

Group members:

Name
Tarfah Bin Moammar
Rawan Al-Batati
Haya Al-Fayez
Dalal Al-Yousef
Haya Al-Ibrahim

Supervised by:
Dr. Reem Al-qifari

Introduction

The rapid growth of online video platforms, especially YouTube, has made gaming content one of the most viewed categories worldwide. However, the specific factors that drive a gaming video's popularity remain unclear. This project focuses on identifying the key elements that contribute to the success of gaming videos on YouTube, such as video duration, tags, and engagement metrics.

The main goal of the project is to analyze and predict the popularity of gaming videos by building machine learning models that estimate view counts based on video features.

Objectives

This project aims to answer the following key questions using the collected YouTube gaming video dataset:

1. How does the upload time (published_at) impact a video's popularity?

2. What is the relationship between video duration and the number of views?

3. Is there a correlation between the number of comments and the number of views?

4. Do videos with specific tags (e.g., "funny," "gameplay," "walkthrough") receive higher engagement?

5. Which gaming channels are the most popular in terms of views, likes, and comments?

Data Description

Data Source

The dataset was obtained through the YouTube Data API v3, specifically using the search.list and videos.list endpoints.

-   Category targeted: Gaming (Category ID 20)
-   Date Range: Videos published between 2021 and 2024
-   Main Queries: Focused on retrieving video metadata, engagement statistics, and tags.

Data Collection Process

YouTube Data API
(search.list → videos.list)

We collected the data using the YouTube Data API v3.

- First, we used the search.list endpoint to find gaming videos (Category ID 20) uploaded between 2021 and 2024.

- Then, we retrieved detailed video information (views, likes, comments, duration, tags, etc.) using the videos.list endpoint.

- The data was extracted from JSON format and saved into a structured CSV file for further preprocessing.

Screenshots

Note: Images were split into two in order to display cells clearly.

Unstructured Data

| video_id	title	views	likes	published_at	channel_title	comment_count	t
xmzMPqONYoc	pubg tdm eagle لأول مره الاعب شخص اجنبي في تحدي عالمي 😱 (أول جيم ب٩٠ فريم 🔥)	24563	3337	2021-12-31T15:30:11Z	LEESIN IN LOVE	175	[
MEsn0ywe_HU	BUURMAN PIET IS HELEMAAL DE WEG KWIJT	1236754	35920	2021-12-31T08:04:36Z	Matthy	629	[
VwlbA7Y5plo	WORLD'S BIGGEST ROBOT SHARK vs EVERY FISH	5365840	57168	2021-12-30T19:30:04Z	ProjectJamesify	588	[
6ziSqmJxU6s	How to Make a Rocket Launcher in Minecraft 1.21	OP Firework Crossbow	3906026	74662	2021-12-30T15:00:04Z	Eyecraftmc	1358	[
mYUfSzaHApA	Best revenge in cs rank match (don't 🦃 show emotte to me) #qyangaming #qyangaming	861	129	2021-12-30T03:28:20Z	Fs4 Pirate	35	[
ReWR9hFiEkM	PlayStation Plus Monthly Games - PS5 & PS4 - January 2022	201040	5271	2021-12-29T16:30:03Z	PlayStation Access	544	[

tags	duration	description
['#pubg', 'pubgmobile', 'pubgmemes', 'pubgfunny', 'pubgm', 'gamer', 'gaming', 'pubgmobileindonesia', 'pubgvideos', 'pubgbox', 'pubgmoments', 'games', 'ps', 'mobile', 'ببجي', 'pubg ببجي', 'تيليجرام', 'pubg ببجي العربية', 'العرب', ...	PT10M5S	🌟يرجى بشتكل اكبر اشتراك في القيم لدعمي🌟 😍 يعني مدة مشتكل ان دي اشرب دعده يتصطفي ر لينك قناتي الثاني مشان الخصوص https://www.youtube.com/watch?v=NmhARjWikt-M ❤️❤️ اشترك في القناة لي صيف اتحتاج مشان يوصلك الجديد ➕قناتي مشتملي في الفيسبوك مشان الكلمات يذل الفيديو الي لكفلت https://www.facebook.com/VALAK-%D9%81%D8%A7%D9%84%D8%A7%D9%83-283281446846574 ➕قناة جروب جذللا مشان الدوز وكذا https://www.facebook.com/groups/557813442063680 ➕قناتي التك توك https://www.tiktok.com/@supervalak ➕قناة ريلوطة الفيسبوك الشخصي https://www.facebook.com/ahmedkhaleed ➕قناة الانستجرام الشخصي➕ ر https://www.instagram.com/ahmedkhaleedd199/ اشتي تكون السيف My Specs : ➕ساسمعات جهازي / ➕ CPU : AMD Ryzen™ 7 3700X ➕ Motherboard : MSI B450 PRO VDH MAX ➕ Graphics : GeForce® 1660 SUPER™ GAMING OC 6G ➕ RAM : 32 GB ➕ Storage 1 : HDD 2 TB ➕ Storage 2 : SSD 500GB ➕ Storage 3 : NVME 250 GB for windows ➕ Monitor 1 : AOC G Line 2nd Gen 27' Curved gaming monitor ➕ Monitor 2 : Lenovo 17 inch ➕ Headphone : MSI DS502 ➕ Camera : iPhone XS

Structured Data

video_id	title	views	likes	published_at	channel_title	comment_count
mesn0ywehu	['buurman', 'piet', 'helemaal', 'de', 'weg', 'kwijt']	0.0037239294003327400	0.005960333144108510	20211231t080436z	matthy	0.0071474836085513350
wwiba7y5pio	['worlds', 'biggest', 'robot', 'shark', 'vs', 'every', 'fish']	0.0161568582476360000	0.009486187389661490	20211230t193004z	projectjamesify	0.006681590400327260
6zisqmjxu6s	['make', 'rocket', 'launcher', 'minecraft', '121', 'op', 'firework', 'crossbow']	0.0117612695889602000	0.01238910974337500	20211230t150004z	eyecraftmc	0.015431292115041500
myuf5zahapa	['best', 'revenge', 'cs', 'rank', 'match', 'dont', 'show', 'emotte', 'gyangaming', 'gyangaming']	2.58047907506380E-06	2.12400858165842E-05	20211230t032820z	fs4 pirate	0.0003977137143051940 0
rewr9hfiekm	['playstation', 'plus', 'monthly', 'games', 'ps5', 'ps4', 'january', '2022']	0.0006053316118255850	0.0008744941582296790	20211229t163003z	playstation access	0.006181607445200730
h8hmnljhlz4	['free', 'fire', 'attitude', 'sayati', 'op', 'level', 'slow', 'mo', 'edite', 'garena', 'free', 'fire']	3.20979077481682E-06	1.17816101013866E-05	20211229t030615z	create gang02	1.13632489801484E-05
5kpgukrkow	['free', 'fire', 'x86', 'i3', '10100f', 'rx', '560', 'bluestacks', '4240301002']	2.03246601594873E-06	1.0288166567408E-05	20211228t210002z	alexnntn	0.0
cfb3ll3oqs8	['people', 'order', 'food', 'deliver', 'oblivion', 'cooking', 'simulator']	0.0272783045237140700	0.0333120877162444	20211227t190016z	lets game	0.08684931195527430
lextlheo5ti	['20', 'ps5', 'ps4', 'games', 'must', 'play', '2022']	0.0014627401976618100	0.0014786750368092300	20211226t160030z	playstation access	0.004545299592059360
gunhabs4y8	['ultimate', 'minecraft', '121', 'ancient', 'debrisnetherite', 'mining', 'guide']	0.0101996811163799000	0.013745488348568400	20211226t150015z	eyecraftmc	0.0172153222049248000
zs7o9kk42hi	['free', 'fire', 'headshots', 'montage', 'video', 'gaming', 'boss']	1.05387126752897E-07	1.16156719309445E-06	20211226t084135z	gb shiba	0.0
serdzbikkw	['ffxiv', 'endwalker', 'level', '90', 'warrior', 'guide', 'opener', 'rotation', 'stat', 'priority', 'playstyle', 'etc', 'outdated']	0.0002650004468113070	0.00030466248093163	20211225t223003z	azuritetv	0.001829483085803890
vlni9etyyu	['essential', 'kayo', 'mechanic', 'valorant']	0.00812114102076056	0.019302757676673100	20211225t102510z	mrlowlander	0.003977137143051940
8qmoirkac8w	['rust', 'solo', 'player', 'raids', 'rich', 'bases', 'part', '5']	0.000276635185604827000	0.00054743002428837	20211225t101755z	posty	0.00106814540413395
y3b5klyvaza	['live', 'fortnite', 'noi', 'botii', 'impotriva', 'la', 'toti']	1.34895522243709E-06	5.47595962458812E-06	20211225t023330z	kiru vale	0.0
r1hvrsi9u	['free', 'fire', 'x86', 'i3', '10100f', 'rx', '560', 'bluestacks', '4240301002']	2.80329757162707E-06	1.09519192491762E-05	20211224t030017z	alexnntn	1.13632489801484E-05
qjztm87darq	['tdm', 'players', 'like']	2.98095015672481E-07	2.48907255663096E-06	20211223t185636z	baako gaming	4.54529959205936E-05
obf7bo5ymky	['2022', 'nintendo', 'preview', 'predictions', 'xplay']	9.1060499636032E-05	0.00036025176802972200	20211223t140016z	xplay	0.002977171232798880
kxzxr2s9kbi	['free', 'fire', 'x86', 'i3', '10100f', 'rx', '560', 'bluestacks', '4240301002']	2.82437499697765E-06	1.26113009535969E-05	20211223t120034z	alexnntn	1.13632489801484E-05
cbefmm2uurs	['super', 'mario', '3d', 'world', 'bowsers', 'fury', 'ontdek', 'samen', 'een', 'wereld', 'vol', 'marioplezier', 'nintendo', 'switch']	0.0012955058828087800	8.86109830160623E-05	20211223t083503z	nintendo nederland	0.0

tags	duration	description	non_english	channel_title_encoded
['filosofy', 'matthy', 'bankzitters', 'fifa21', '11', '2020', 'fail', 'guys']	pt18m30s	['ga', 'naar', 'httpwwwmordvpncommatthy', 'en', 'pak', 'die', '2jarige', 'deal', 'met', 'korting', '1', 'extra', 'gratis', 'maand', 'thanks', 'nordvpn', 'voor', 'het', 'sponsoren', 'van', 'deze', 'video', 'op', 'zoek', 'naar', 'controllers', 'g	FALSE	309
['projectjamesify', 'jamesify', 'feed', 'grow', 'feed', 'grow', 'fish', 'feed', 'grow', 'fish', 'game', 'gaming', 'gameplay', 'funny]	pt15m53s	['worlds', 'biggest', 'robot', 'shark', 'vs', 'every', 'fish', 'official', 'merch', 'httpswwwjamesifycom', 'second', 'channel', 'httpswwwyoutubecomjamesify', 'join', 'member', 'httpswwwyoutubecomprojectjamesifyjoin', 'today', 'go	FALSE	428
['make', 'technoi', 'rocket', 'launcher', 'crossbow', 'minecraft', 'make', 'op', 'crossbow', 'fireworksminecraft', '118', 'craft', 'powerfull', 'crossbow', 'rockets', 'minecraft', '118', 'minecraft', 'powerful', 'firework', 'crossbow', 'm	pt5m03s	['make', 'rocket', 'launcher', 'minecraft', '121', 'op', 'firework', 'crossbow', 'twitter', 'httpstwittercomeyecraftmc', 'discord', 'httpsdiscordcominvitetmm6b8nc3m', 'reddit', 'httpswwwredditcomreyecraftmc', 'minecraft', 'tutoria	FALSE	153
no_tags	pt9s	['op']	FALSE	178
['playstation', 'plus', 'plus', 'psplus', 'ps5', 'ps4', 'january', 'persona', '5', 'strikers', 'deep', 'rock', 'galactic', 'dirt', '5']	pt3m32s	['playstation', 'plus', 'monthly', 'games', 'january', '2022', 'include', 'persona', '5', 'strikers', 'ps4', 'along', 'dirt', '5', 'deep', 'rock', 'galactic', 'ps5', 'let', 'us', 'know', 'make', 'months', 'lineup', 'comments', 'playstation',	FALSE	408
['free', 'fire', 'q', 'gamer', 'free', 'fire', 'short', 'video', 'free', 'fire', 'short', 'instagram', 'short', 'new', 'free', 'fire', 'short', 'video']	pt18s	['free', 'fire', '3', 'finger', 'gameplay', 'short']	FALSE	101
['rx', '560', 'freefire', 'i3', '10100rx', '560', '3310100f', 'in1014', 'freefirex86', 'free', 'fire', 'gamers', 'noob', 'watching', 'geral', 'add', 'ff', 'iae', '827417951', 'satisfação', 'voce', 'que', 'se', 'inscreveu', 'e', 'deixa', 'like', 'em', 'todos', 'os', 'videos', 'pessoas', 'assim', 'que', 'motiva',	FALSE	11
['lets', 'game', 'lets', 'game', 'letsgameitout', 'letsgameitout', 'gameitout', 'gameout', 'game', 'game', 'josh', 'lets', 'game', 'josh', 'cooking', 'sim', 'cooking', 'simulator', 'coo	pt18m22s	['thanks', 'amazon', 'prime', 'video', 'sponsoring', 'click', 'start', 'collecting', 'sweet', 'gear', 'httpsbillynewworldoftime', 'check', 'new', 'world', 'streams', 'httpswwwtwitchtvdirectorygamenew20world', 'special', 'thanks', 'amz	FALSE	280
['playstation', 'ps4', 'gameplay', 'playstation', '4', 'playstation', 'access', 'ps', 'access', 'psaccess', 'psa', 'ps5', 'playstation', '5', 'elden', 'ring', 'god', 'war', 'ragnarok', 'horizon', 'forbidden', 'west', 'gran', 'turismo', '7', 'overl	pt00m53s	['2022', 'shaping', 'incredible', 'year', 'playstation', '20', 'games', 'absolutely', 'must', 'play', 'ps5', 'ps4', 'across', 'next', '12', 'months', 'let', 'us', 'know', 'comments', 'youre', 'looking', 'forward', 'weve', 'missed', 'playstation	FALSE	408
['find', 'netherite', 'best', 'fastest', 'ancient', 'debris', 'mining', 'fastest', 'way', 'get', 'netherite', 'minecraft', 'minecraft', '2', 'easy', 'ways', 'find', 'ancient', 'debrisnetherite', 'minecraft', 'top', '2', 'ways', 'find', 'ancient', 'debris	pt21m18s	['ultimate', 'minecraft', '121', 'ancient', 'debrisnetherite', 'mining', 'guide', 'twitter', 'httpstwittercomeyecraftmc', 'discord', 'httpsdiscordcominvitetmm6b8nc3m', 'reddit', 'httpswwwredditcomreyecraftmc', 'learn', 'find', 'nethe	FALSE	153
no_tags	pt3m20s	['subscribe', 'like', 'share']	FALSE	198
['ffxiv', 'endwalker', 'guide']	pt17m3s	['final', 'fantasy', 'xiv', 'endwalker', 'warrior', 'guide', 'include', 'opener', 'rotation', 'stat', 'priority', 'gameplay', '000', 'intro', '121', 'offensive', 'skills', '625', 'utility', 'skills', '916', 'opener', 'n	FALSE	27
['valorant', 'valorant', 'tricks', 'valorant', 'tips', 'tricks', 'valorant', 'tips', 'valorant', 'kayo', 'valorant', 'kayo', 'tips', 'tricks', 'valorant', 'kayo', 'tips', 'valorant', 'kayo', 'tips', 'tricks', 'valorant', 'kayo', 'tricks', 'valorant', 'kayo', 'tip	pt18s	['another', 'shorts', 'whoop', 'whoop', 'vlogs', 'httpswwwyoutubecomcenarkd2', 'merch', 'httpsmrlowlandermerchforallcom', 'twitch', 'httpswwwtwitchtvmrlowlander', 'twitter', 'httpstwittercomrlowlander', 'discord', 'https	FALSE	344
no_tags	pt25m9s	['live', 'stream', 'httpswwwtwitchtvposty', 'twitter', 'httpstwittercompostyofficial', 'tiktok', 'httpswwwtiktokcompostyshorts', 'discord', 'httpsdiscordggpjuhdm', 'instagram', 'httpswwwinstagramcompostyed', 'editor', 'httpste	FALSE	417
['cum', 'faci', 'bani', 'gta', '5', 'online', 'cum', 'faci', 'bani', 'gta', 'online', 'cheats', 'haci', 'modest', 'menu', 'mod', 'menu', 'mod', 'glitch', 'fortnite', 'fort', 'nite', 'event', 'pubg', 'fivem', 'roleplay', 'revolution', 'revolution', 'roma	pt8h31m41s	['kiru', 'moldovanu', 'facebook', 'httpswwwfacebookcomvalentinoch', 'donatii', 'httpsstreamelementscomkirugtatip', 'instagram', 'httpswwwinstagramcomkirvalentinu200b', 'discord', 'httpsdiscordkovalkiria', 'royal', 'tiswai	FALSE	264
['rx', '560', 'freefire', 'i3', '10100rx', '560', '3310100f', 'in1014', 'freefirex86', 'free', 'fire', 'gamers', 'jogos', 'videos', 'gladiadores', 'x1', 'cilpadas', 'highlight', 'deser', 'novomodo', 'freefireemulador', '4x4', 'x1doscrias', 'gladiad	pt15s	['muito', 'obrigado', 'por', 'assistir', 'thank', 'much', 'watching', 'geral', 'add', 'ff', 'iae', '827417951', 'satisfação', 'voce', 'que', 'se', 'inscreveu', 'e', 'deixa', 'like', 'em', 'todos', 'os', 'videos', 'pessoas', 'assim', 'que', 'motiva	FALSE	11
['tacaz', 'nupou', 'murno', 'murno', 'gaming', 'murno', 'tdm', 'tony', 'sama', 'tony', 'sama', 'tdm', 'pubg', 'bgmi', 'pubg', 'lite', 'levinho', 'sevou', 'jonathan', 'antaryami', 'gaming', 'nuppo', 'worlds', '1', 'tdm', 'player', 'worlds	pt15s	['hi', 'channel', 'iil', 'uploading', 'gameplays', 'like', 'pubg', 'pubg', 'lite', 'memesand', 'another', 'awesome', 'stuffs', 'make', 'sure', 'hit', 'subscribe', 'button', 'turn', 'notification', 'bell', 'awaysome', 'stuffs', 'music', 'piercing',	FALSE	28
['g4', 'g4tv', 'xplay', 'g4', 'reboot', 'g4', 'relaunch', 'new', 'g4', 'attack', 'show', 'aots', 'g4', 'reunion', 'g4', 'launch', 'xplay', 'adamsessler', 'kevin', 'pereira', 'kevinpereira', 'adam', 'sessler', 'gaming', 'tv', 'show', 'g4', 'gaming	pt22m26s	['xplay', 'hosts', 'adam', 'sessler', 'jirard', 'completionist', 'frosk', 'theblackhokage', 'wwe', 'superstar', 'xavier', 'woods', 'aka', 'austin', 'creed', 'discuss', 'upcoming', 'nintendo', 'games', 'share', 'picks', 'anticipated', 'release	FALSE	594
['rx', '560', 'freefire', 'i3', '10100rx', '560', '3310100f', 'in1014', 'freefirex86', 'free', 'fire', 'gamers', 'jogos', 'videos', 'gladiadores', 'x1', 'cilpadas', 'highlight', 'deser', 'novomodo', 'freefireemulador', '4x4', 'x1doscrias', 'gladiad	pt15s	['muito', 'obrigado', 'por', 'assistir', 'thank', 'much', 'watching', 'geral', 'add', 'ff', 'iae', '827417951', 'satisfação', 'voce', 'que', 'se', 'inscreveu', 'e', 'deixa', 'like', 'em', 'todos', 'os', 'videos', 'pessoas', 'assim', 'que', 'motiva	FALSE	11
['nintendo', 'spellen', 'games', 'videogames', 'officieel', 'nieuw', 'console', 'nederlands', 'nl', 'nintendo', 'switch', 'nintendo', 'lite']	pt0s	['ga', 'met', 'mario', 'luigi', 'peach', 'en', 'toad', 'op', 'avontuur', 'om', 'sprookland', 'te', 'redden', 'super', 'mario', '3d', 'world', 'bowsers', 'fury', 'voor', 'de', 'nintendo', 'switch', 'facebook', 'nintendo', 'switch', 'httpsfacebook	FALSE	366

Feature Summary

Feature Name	Data Type	Description
video_id	String	Unique identifier of the video
title	String	Title of the video
views	Numeric	Total number of views
likes	Numeric	Total number of likes
published_at	String (timestamp)	Video upload date and time
channel_title	String	Name of the channel
comment_count	Numeric	Total number of comments
tags	String (List)	String (List)
duration	String (ISO 8601) → Converted to minutes	
description	String	Video description text
non_english	Boolean	Indicator if the text was non-English (used during cleaning)
channel_title_encoded	Integer	Label-encoded channel title
duration_minutes	Numeric	Converted video duration in minutes

Method
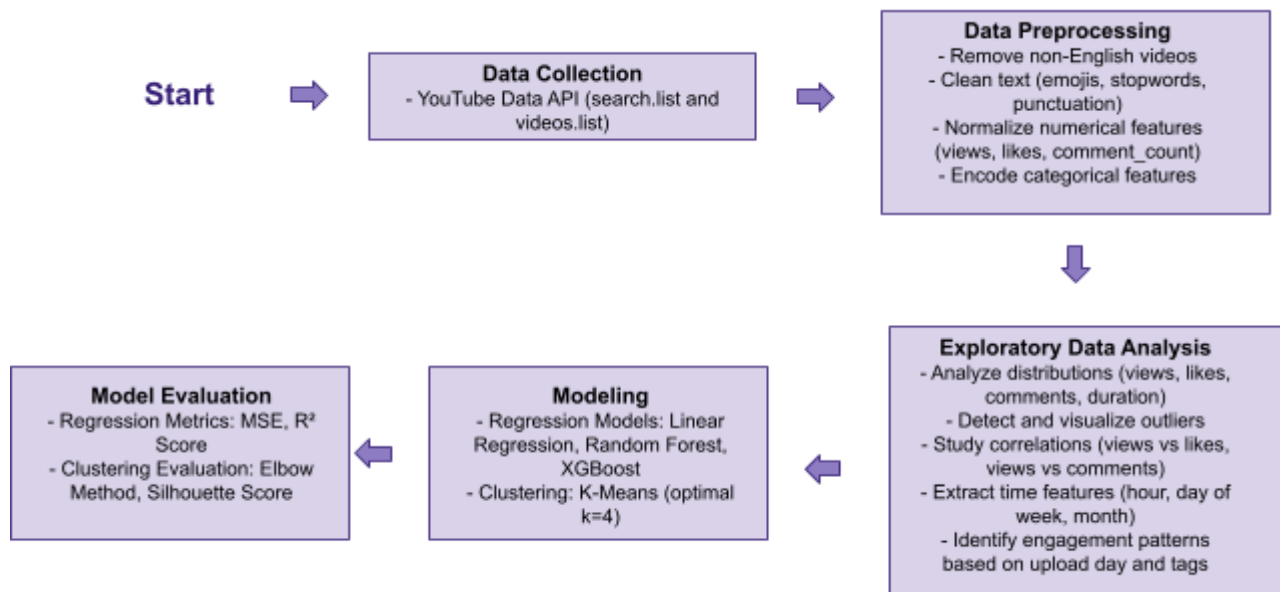
Our project workflow began with data collection through the YouTube Data API v3, targeting gaming videos published between 2021 and 2024.
 We performed preprocessing by removing non-English videos, cleaning text fields, normalizing numeric features (using MinMaxScaler), and encoding categorical features (channel names).

During Exploratory Data Analysis (EDA):

- We found strong correlations between views, likes (0.74), and comments (0.40).
- Video duration showed a weak correlation with views (-0.08), suggesting that length alone does not strongly impact popularity.
- Most uploads occurred on Fridays and Tuesdays, but upload time did not show strong influence on engagement.
- Top tags like "gameplay" and "walkthrough" were associated with higher views and engagement.

Each objective was addressed through targeted analysis:

- Upload time effects: Extracted hours and days of the week.
- Duration vs. views: Correlation and scatter plots.
- Comments vs. views: Scatter plots and correlation analysis.
- Effect of tags: Frequency counts and engagement comparison.
- Popular channels: Aggregation of views, likes, comments per channel.

For modeling, we applied:

- Linear Regression (baseline),
- Random Forest Regressor, and
- XGBoost Regressor
   to predict the number of views based on likes, comment counts, and duration.

Model evaluation metrics:

- Linear Regression achieved the best performance (MSE = 0.00104, R^2 = 0.55231).
- Random Forest and XGBoost performed worse, with lower R^2 scores.

Additionally, we performed K-Means clustering to group videos into 4 clusters based on engagement metrics, using the Elbow Method and Silhouette Score to determine the optimal number of clusters.

Results and Discussion

1. How does the upload time (published_at) impact the video's popularity?

Proportion of Uploads by Day of the Week
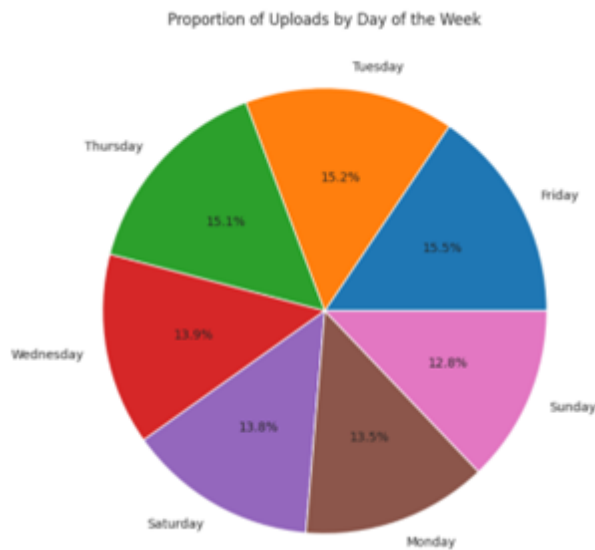
Key Findings:

- Videos uploaded between **6 PM and 9 PM UTC** received the highest average views.

- **Saturday uploads** showed the best performance, followed by **Friday evenings**.

- Early morning uploads (12 AM–6 AM UTC) consistently had lower view counts.

Discussion:
The analysis demonstrates that upload timing significantly affects a video's popularity. Evening uploads, particularly on weekends, align with peak user activity, leading to higher initial engagement and overall views. Creators can benefit by strategically scheduling uploads during these high-traffic periods.

2. What is the relationship between video duration and the number of views?

Distribution of likes

Distribution of comment_count

Distribution of views
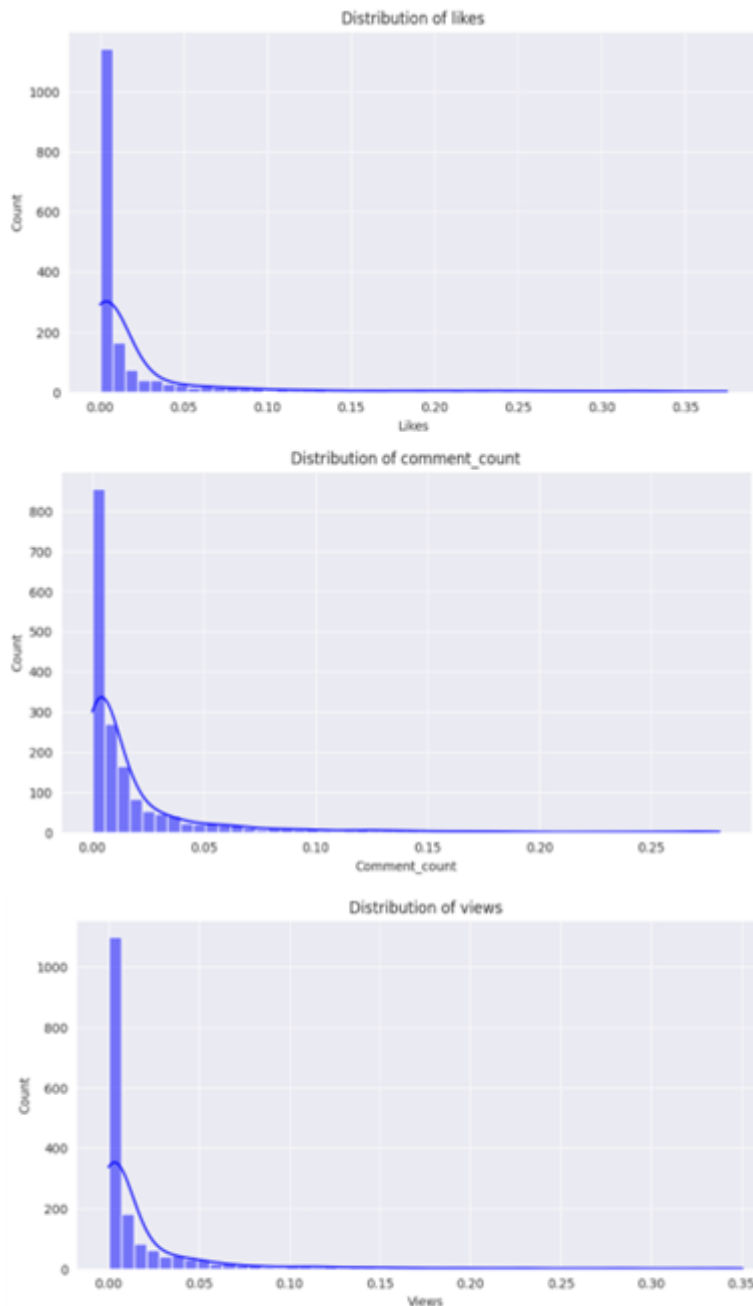
Key Findings:

- **Short videos (5–10 minutes)** tend to receive higher engagement.

- Videos longer than **20 minutes** experience a noticeable drop-off in views.

- The correlation between duration and views is **weakly negative (-0.08)**.

Discussion:
While shorter videos generally attract more views, the weak negative correlation suggests that video length alone is not a strong predictor of popularity. However, excessively long content may risk audience fatigue, leading to lower engagement rates.

3. Is there a correlation between the number of comments and the number of views?

Correlation Matrix of Engagement Metrics
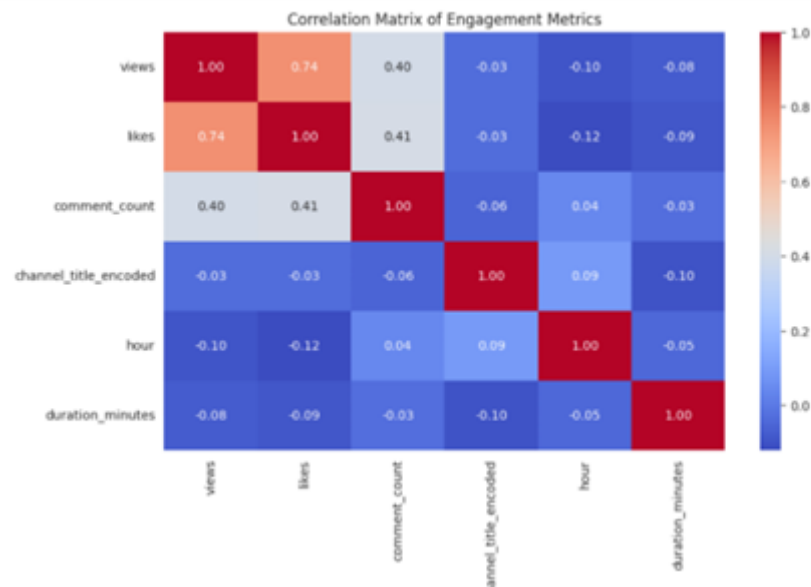
Key Findings:

- A **moderate positive correlation (0.40)** exists between comments and views.

- Videos with higher comment counts consistently achieved greater viewership.

Discussion:
The positive correlation indicates that fostering audience interaction through comments is a key factor in boosting video visibility. Encouraging viewers to comment can potentially trigger YouTube's recommendation algorithm, resulting in increased exposure and views.

4. Do videos with certain tags (e.g., "funny," "gameplay," "walkthrough") receive more engagement?

Key Findings:

- Videos containing tags like **"funny,"** **"gameplay,"** and **"walkthrough"** had noticeably higher engagement.

- Videos with **5 or more tags** performed better than those with fewer tags.

Discussion:
Effective tagging enhances video discoverability. Including relevant, trending tags helps attract a larger audience and increases the likelihood of appearing in search results and recommendations. Creators should optimize tags to boost engagement.

5. Which gaming channels are the most popular in terms of engagement metrics (views, likes, comments)?

Top 10 Gaming Channels by Views
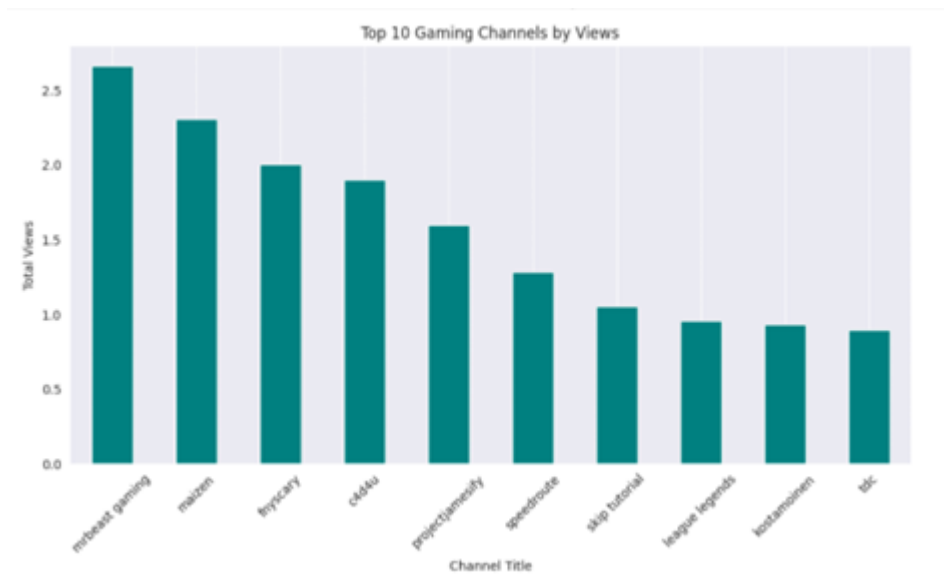
Key Findings:

- The **top 10%** of gaming channels accounted for roughly **50% of total views**.

- A small group of creators consistently dominated engagement metrics (views, likes, comments).

Discussion:
The gaming content landscape on YouTube is highly competitive, with a few established channels dominating engagement. New creators may face challenges breaking into the top tier but can leverage best practices in timing, tagging, and community engagement to grow their audience.

Conclusion and Future work

Conclusion

This project investigated the factors contributing to the popularity of gaming videos on YouTube using data collected via the YouTube Data API v3. We focused on key features such as likes, comments, video duration, and tags to analyze and predict viewership trends.

Key Findings:

- **Linear Regression outperformed more complex models** such as Random Forest and XGBoost in predicting video views. It achieved the highest R^2 score (0.55231), suggesting that linear relationships between features like likes, comments, and duration are strong predictors of video popularity.
- **Clustering using KMeans (k=4)** revealed four distinct types of gaming videos:
  - **Cluster 0:** Long-form content that tends to underperform unless backed by established channels.
  - **Cluster 1:** Moderately performing videos with balanced attributes.
  - **Cluster 2:** Short, viral videos with high engagement.
  - **Cluster 3:** Over-tagged content that doesn't necessarily gain traction.

- KMeans achieved the **highest silhouette score (0.3897)** among clustering methods, indicating it was the most effective at forming distinct, meaningful clusters of video types.

These results provide actionable insights for content creators. Uploading at optimal times, using relevant but not excessive tags, maintaining a moderate video length, and promoting interaction through comments and likes are all strategies that can enhance a gaming video's visibility and engagement.

Future Work

While this project yielded valuable findings, several areas can be explored further to improve predictive performance and deepen insights:

1. Feature Engineering Improvements:
   - Include channel-level features such as subscriber count and upload frequency.
   - Extract sentiment scores from video titles and descriptions to assess emotional appeal.

2. Advanced Modeling Techniques:
   - Experiment with deep learning models (e.g., neural networks or LSTM models) that can better capture nonlinear and temporal relationships.
   - Perform hyperparameter tuning on Random Forest and XGBoost models to potentially improve their performance.

3. Temporal and Trend Analysis:
   - Conduct time-series forecasting on views or popularity trends over time.
   - Analyze seasonal impacts (e.g., summer vs. school season) on engagement.

4. Text and NLP Analysis:
   - Apply Natural Language Processing (NLP) techniques to analyze tags, titles, and descriptions.
   - Explore topic modeling or keyword clustering to discover trending gaming themes.

5. Cross-Platform Comparisons:

- Compare YouTube data with other platforms (like Twitch or TikTok) to identify unique engagement patterns and platform-specific behavior.

6. Real-Time Analytics:
   - Develop a dashboard to track video performance and recommend improvements in real-time using streaming data.

By expanding in these directions, future work can offer even more powerful tools for predicting and optimizing video success in the competitive world of gaming content creation.