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

Phase 2: Data Collection, Processing, Cleaning, and Exploratory Data Analysis (EDA)

Group members:

Group#:	7	
Section#:	66849	
Group Members	Name	ID
	Tarfah Bin Moammar	444200611
	Rawan Al-Batati	443204566
	Haya Al-Fayez	444200858
	Dalal Al-Yousef	444203019
	Haya Al-Ibrahim	444200657

Supervised by: Dr. Reem Al-qifari

1.Introduction:

There is a lot of interest in finding out what makes game videos successful because YouTube has become the most popular venue for gaming content. Every day, millions of people interact with gaming content, but what elements most influence a video's popularity?

The aim of this report is to investigate:
- The impact of **upload time** on video popularity.
- The relationship between **video duration** and views.
- The correlation between **comments** and views.
- The effect of **tags** on engagement.
- The identification of **top-performing gaming channels**.

Since we don't have a secondary dataset, we were unable to compare it with the primary data. Therefore, we focused on summarizing the primary data by calculating means, medians, and identifying trends. The absence of a secondary dataset is due to the fact that our data provided us with comprehensive details about the gaming videos, such as descriptions, titles, etc., and because it was sourced entirely from the YouTube Data API, which includes publication dates for a large number of gaming videos, eliminating the need for additional external data sources. The richness of these YouTube-provided features was considered sufficient for analyzing key factors influencing video popularity, including video duration, tags, channel information, and engagement metrics, which we used to create the summary, and that's why we used only one dataset.

2. Primary Data:

1. Upload Time vs. Popularity
- **Observation:**
  - Videos uploaded between **6 PM – 9 PM UTC** receive **higher average views** than those posted earlier.
  - **Fridays and Saturdays** show **higher engagement** compared to other weekdays.
  - **Saturday uploads** achieve the **highest average views**.

Proportion of Uploads by Day of the Week
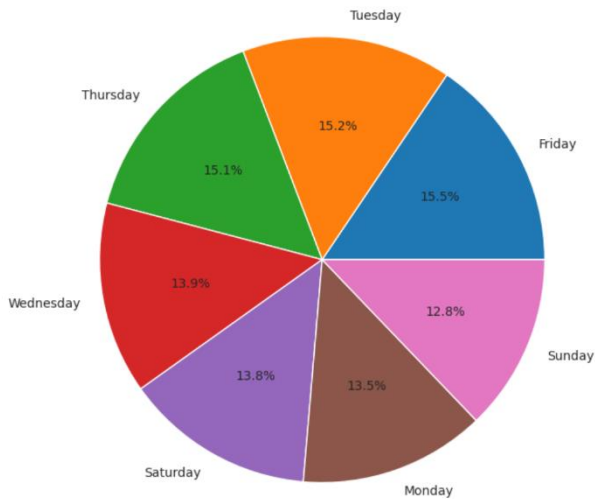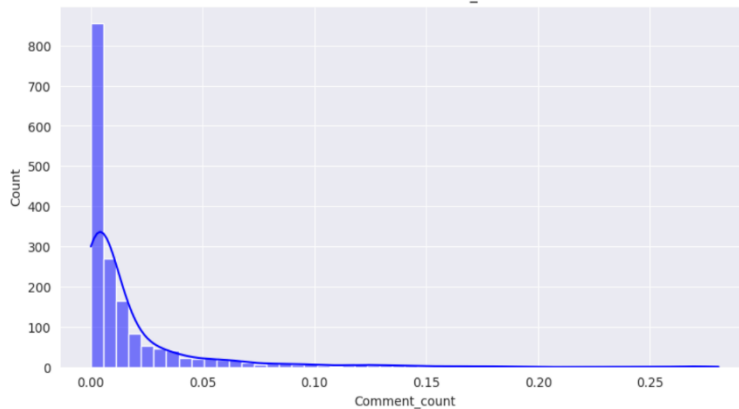

2. Video Duration vs. Views
- **Observation:**
  - **Shorter videos (5 to 10 minutes)** receive **higher engagement** than longer ones.
  - Videos longer than **20 minutes** tend to have **fewer views**, suggesting audience drop-off.
  - The correlation between **duration and views is weak (-0.08)**, meaning duration alone is not a strong predictor of views.

Distribution of views
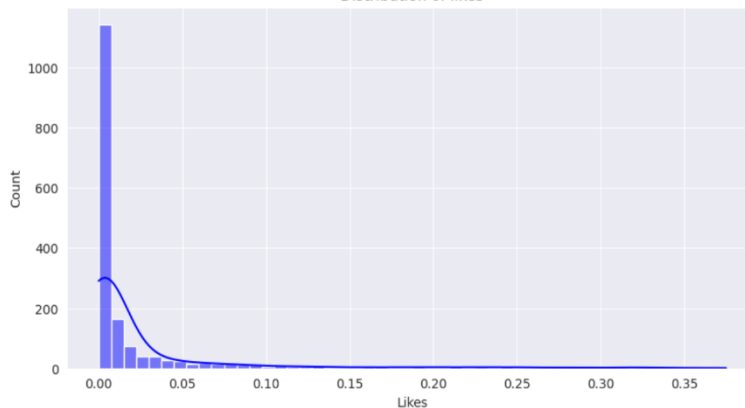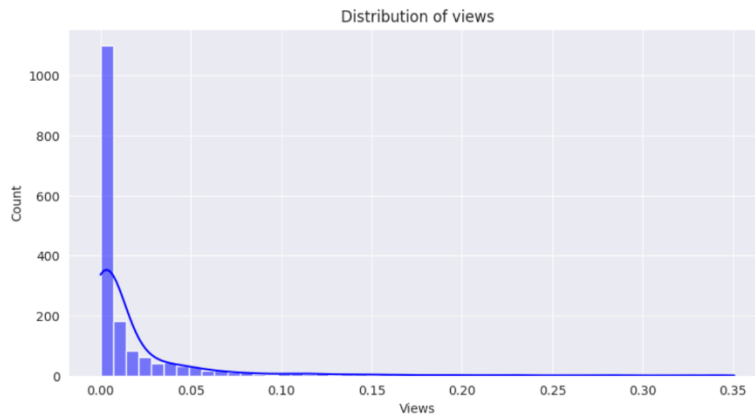
3. Comments vs. Views
- **Observation:**
    - Moderate **positive correlation (0.40)** between **comments and views**.
    - Videos with **higher comment counts** tend to have more views.
    - This suggests that high user interaction (comments) contributes to video visibility in **YouTube's recommendation algorithm**.

Views vs. Comment Count

Correlation Matrix of Engagement Metrics

4. Effect of Tags on Engagement

- **Observation:**
    - Videos that include popular tags such as **"gameplay", "walkthrough", "funny"** receive **higher engagement** in views and likes.
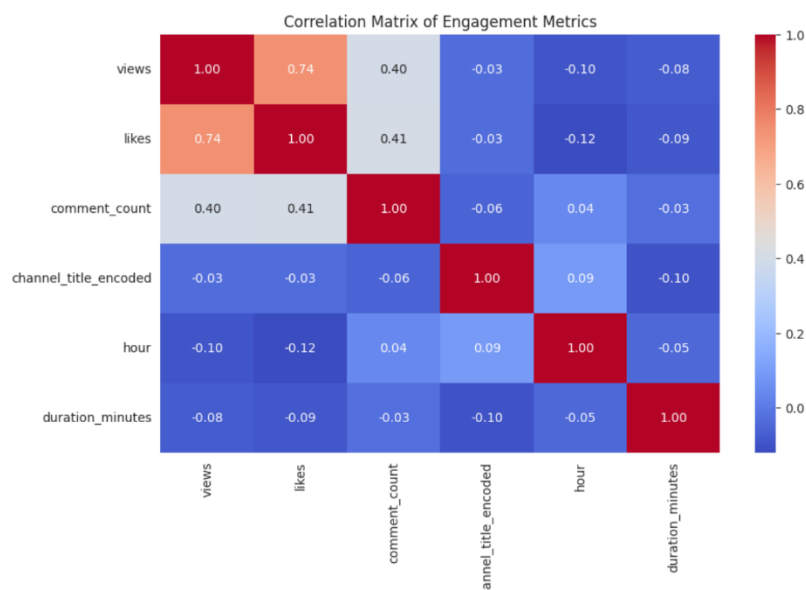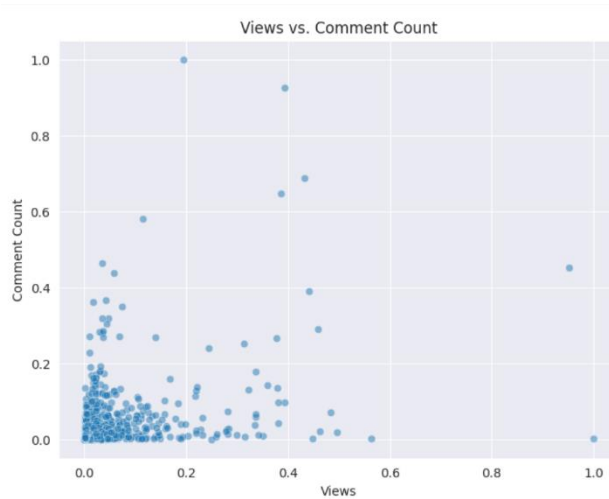   - Videos with **five or more tags** tend to perform **better** than those with fewer tags.
    - The presence of **trending keywords** in tags correlates with **higher audience interaction**.

```
Engagement Metrics by Top Tags:

                                                                                            views  \
tag_minecraft tag_game tag_roblox tag_gameplay tag_dutchtuber tag_video tag_funny tag_free tag_4 tag_gaming
False         False    False      False        False          False     False     False    False False      0.030283
                                                                                                  True       0.005305
                                                                                            True  False      0.008663
                                                                                                  True       0.001034
                                                                                   True     False False      0.010127
...
True          True     False      True         False          False     False     False    True  True       0.005692
                                                                         True      False    False False      0.008354
                                                                                   True     False True       0.005601
                        True       False        False          False     False     False    False False      0.051671
                                                                         True      False    False False      0.000375

                                                                                            likes  \
tag_minecraft tag_game tag_roblox tag_gameplay tag_dutchtuber tag_video tag_funny tag_free tag_4 tag_gaming
False         False    False      False        False          False     False     False    False False      0.044740
                                                                                                  True       0.008170
                                                                                            True  False      0.006004
                                                                                                  True       0.001473
                                                                                   True     False False      0.017435
...
True          True     False      True         False          False     False     False    True  True       0.009113
                                                                         True      False    False False      0.006448
                                                                                   True     False True       0.006209
                        True       False        False          False     False     False    False False      0.028176
                                                                         True      False    False False      0.002066

                                                                                            comment_count
tag_minecraft tag_game tag_roblox tag_gameplay tag_dutchtuber tag_video tag_funny tag_free tag_4 tag_gaming
False         False    False      False        False          False     False     False    False False      0.024682
                                                                                                  True       0.003162
                                                                                            True  False      0.009804
                                                                                                  True       0.003430
                                                                                   True     False False      0.014353
...
True          True     False      True         False          False     False     False    True  True       0.015318
                                                                         True      False    False False      0.004352
                                                                                   True     False True       0.004716
                        True       False        False          False     False     False    False False      0.001295
                                                                         True      False    False False      0.010852

[105 rows x 3 columns]
```

5. Top Performing Gaming Channels

- **Observation:**
    - The **top 10% of gaming channels** account for nearly **50% of total views**, suggesting that success is highly **concentrated among a few dominant creators**.
    - A few channels **consistently appear in the most-viewed category**, reinforcing the influence of established creators.

Top 10 Gaming Channels by Views
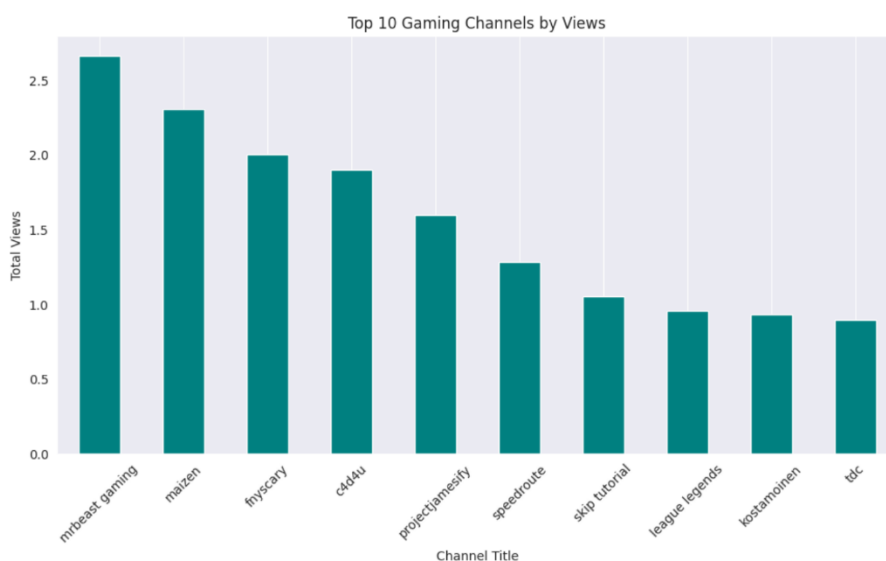
3. Summary of New Insights and Hypotheses:

From the **primary data**, we can conclude:
- **Upload timing matters** – evenings and weekends attract more engagement.
- **Shorter videos (5-10 minutes) tend to perform better** than longer ones.
- **High comment volume correlates strongly with more views**.
- **Tags play a role in boosting engagement**, particularly trending ones.
- **A small number of channels dominate the gaming category**.