MENTORNESS INTERNSHIP PROGRAM

Comprehensive Analysis of YouTube Songs Data

By PowerBI

Presented by :
Prakhar Purohit

CONTENTS

CONTENTS

Introduction

This project aims to conduct a comprehensive analysis of YouTube songs data using Power BI. The goal is to create insightful visualizations and reports that provide a deeper understanding of YouTube songs' performance, popularity, and user engagement. The analysis aims to uncover trends, preferences, and patterns to aid content creators and stakeholders in optimizing their YouTube song content.

Objective

The main objective of this project is to
1. **Data Cleaning and Preparation**
2. **Exploratory Data Analysis (EDA)**
3. **Content and Channel Analysis**
4. **Temporal Trends Analysis**
5. **User Engagement Insights**

Dataset Overview

The dataset contains key attributes of YouTube songs such as:

1. video_id: Unique identifier for each YouTube video.
2. channelTitle: Title of the YouTube channel publishing the song.
3. title: Title of the YouTube song video.
4. description: Description provided for the YouTube song video.
5. tags: Tags associated with the YouTube song video.
6. publishedAt: Date and time when the YouTube song video was published.
7. viewCount: Number of views received by the YouTube song video.
8. likeCount: Number of likes received by the YouTube song video.
9. favoriteCount: Number of times the YouTube song video has been marked as a favorite.
10. commentCount: Number of comments posted on the YouTube song video.
11. duration: Duration of the YouTube song video.
12. definition: Video definition or quality (e.g., HD, SD).
13. caption: Availability of captions for the YouTube song video

Methodology

1. Extracting basic information about the data, checking for nulls, and duplicates.

2. Data cleaning and preprocessing.

3. Exploratory Data Analysis (EDA) and visualization.

4. Content and Channel Analysis.

5. Temporal Trends Analysis.

6. User Engagement Insights.

7. Preparing reports and recommendations.

Content & Channel Analysis

- Total Songs: 19.34 Thousand
- Average Duration: 4.12 Minutes
- Average Likes: 87.19 Thousand
- Average Comments: 2.64 Thousand
- Average Views: 11.94 Million
- Video Definition Distribution: HD: 16.58 Thousand, SD: 2.76 Thousand
- Video Definition Distribution: Without Captions: 19.14 Thousand, With Captions: 0.2 Thousand
- Channel Performance: T-Series with 19.34 Thousand songs

Youtube Songs Analysis

No. of videos

19.34K

0.00K — 38.68K
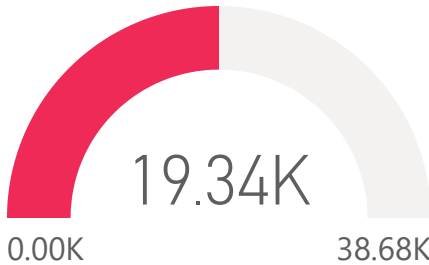
Metric

Comments

Likes

Views

Top

Top 5

Top 20

Top 15

Top 10

Popular Songs

title	SelectedMetric
Vaaste Song: Dhvani Bhanushali, Tanishk Bagchi	Nikhil D	Bhushan Kumar	Radhika Rao, Vinay Sapru	1536719459
Lut Gaye (Full Song) Emraan Hashmi, Yukti	Jubin N, Tanishk B, Manoj M	Bhushan K	Radhika-Vinay	1333673024
DILBAR Lyrical	Satyameva Jayate	John Abraham, Nora Fatehi,Tanishk B, Neha Kakkar,Dhvani, Ikka	1257657363
SIMMBA: Aankh Marey Lyrical	Ranveer Singh, Sara Ali Khan	Tanishk Bagchi, Mika Singh,Neha,Kumar S	1217598096
Guru Randhawa: High Rated Gabru Official Song	DirectorGifty	Bhushan Kumar	T-Series	1183904054
Cham Cham Full Video	BAAGHI	Tiger Shroff, Shraddha Kapoor	Meet Bros, Monali Thakur	Sabbir Khan	1166497013
Guru Randhawa: Lahore (Official Video) Bhushan Kumar	Vee	DirectorGifty	T-Series	1064521703
Bom Diggy Diggy (VIDEO)	Zack Knight	Jasmin Walia	Sonu Ke Titu Ki Sweety	996655164
Leja Re	Dhvani Bhanushali	Tanishk Bagchi	Rashmi Virag |Radhika Rao| Vinay Sapru	Siddharth	958559940
Full Song: KHAIRIYAT (BONUS TRACK)	CHHICHHORE	Sushant, Shraddha	Pritam, Amitabh B	Arijit Singh	941903968
SIMMBA: Aankh Marey	Ranveer Singh, Sara Ali Khan	Tanishk Bagchi,Mika Singh,Neha Kakkar, Kumar S	934317467
Bum Bum Bole (Full Song) Film - Taare Zameen Par	Shaan, Aamir Khan	925908922
'PREM RATAN DHAN PAYO' Title Song (Full VIDEO)	Salman Khan, Sonam Kapoor	Palak Muchhal T-Series	897064936
Official Video: Humnava Mere Song	Jubin Nautiyal	Manoj Muntashir	Rocky - Shiv	Bhushan Kumar	884819674
Official Video: Nikle Currant Song	Jassi Gill	Neha Kakkar	Sukh-E Muzical Doctorz	Jaani	876452469
Badri Ki Dulhania (Title Track) Varun, Alia, Tanishk, Neha, Monali, Ikka	"Badrinath Ki Dulhania"	871077049
Jagga Jasoos: Galti Se Mistake Video Song	Ranbir, Katrina	Pritam, Arijit, Amit	Amitabh B	866956636
Lyrical: Tum Hi Aana	Marjaavaan	Riteish D, Sidharth M, Tara S	Jubin Nautiyal,Payal Dev,Kunaal V	834663356
'Abhi Toh Party Shuru Hui Hai' FULL VIDEO Song	Khoobsurat	Badshah	Aastha	773290899
'Chittiyaan Kalaiyaan' FULL VIDEO SONG	Roy	Meet Bros Anjjan, Kanika Kapoor	T-SERIES	759070440
Total	**20281311641**

Average Duration

4.12

Average Views

11.94M

Average Likes

87.19K

Average Comments

2.64K

Video Caption Distribution
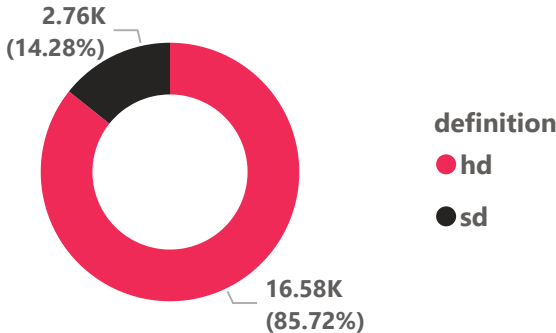
0.2K (1.04%)

19.14K (98.96%)

caption
- False
- True

Video Defination Distribution

2.76K (14.28%)

16.58K (85.72%)

definition
- hd
- sd

Temporal Trends Analysis

- Yearly Song Release Trend: Analysis of the number of songs published every year.
- Monthly Song Release Trend: Analysis of the number of songs published every month.
- Weekly Song Release Trend: Analysis of the number of songs published different days of week.
- Peak Publishing Hours: Analysis of the number of songs published different times of day.

Song Release Trend

Yearly Song Release Trends

- 4.14K (21.41%)
- 1.84K (9.51%)
- 1.62K (8.39%)
- 1.57K (8.14%)
- 1.38K (7.15%)
- 1.32K (6.85%)
- 1.3K (6.74%)
- 1.28K (6.61%)
- 1.08K (5.6%)
- 1.03K (5.35%)
- 0.99K (5.12%)
- 0.98K (5.05%)
- 0.79K (4.08%)

Year
- 2011
- 2022
- 2016
- 2017
- 2019
- 2023
- 2021
- 2018
- 2020
- 2015
- 2014
- 2013
- 2012
- 2010

Monthly Song Release Trends

- 1.25K (6.48%)
- 1.23K (6.37%)
- 1.21K (6.23%)
- 4.81K (24.86%)
- 1.33K (6.89%)
- 1.16K (5.98%)
- 1.48K (7.64%)
- 1.19K (6.17%)
- 1.38K (7.12%)
- 1.25K (6.48%)
- 1.35K (6.98%)
- 1.7K (8.78%)

Month Name
- September
- October
- November
- May
- March
- June
- July
- January
- February
- December
- August
- April

Song Release Trend

Weekly song release trend

Published At

23-12-2010   20-08-2023

1.29K (6.67%)
2.47K (12.78%)
2.67K (13.81%)
3.59K (18.54%)
3.14K (16.24%)
3.1K (16.01%)
3.08K (15.94%)

Day Name
- Monday
- Wednesday
- Friday
- Tuesday
- Thursday
- Saturday
- Sunday

Peak publishing hours

Bar chart: Count of video_id by Hour

Hour	Count
9	0.5K
10	0.5K
11	1.7K
12	1.6K
13	1.7K
14	1.3K
15	1.7K
16	1.8K
17	1.8K
18	2.4K
19	2.0K
20	0.6K
21	0.3K
22	0.2K
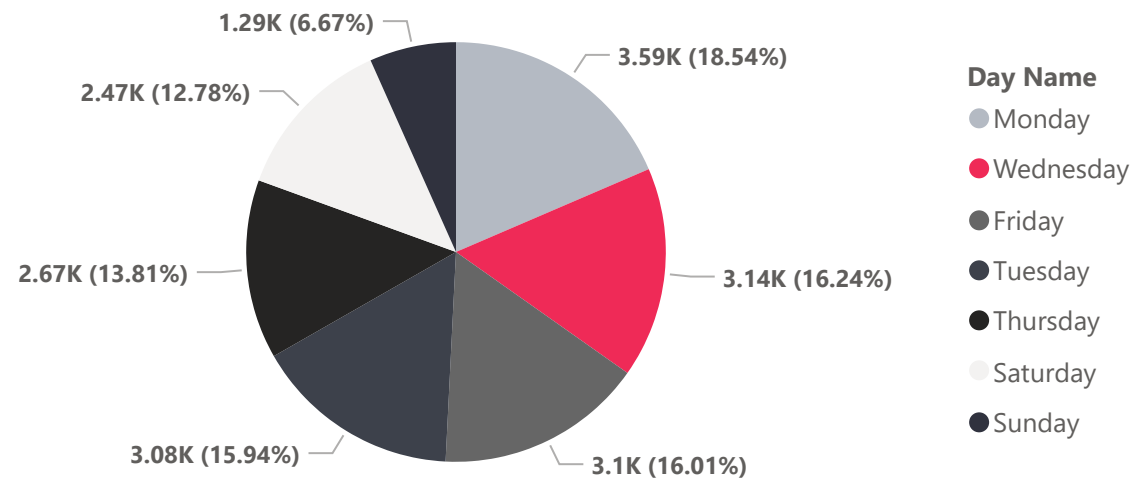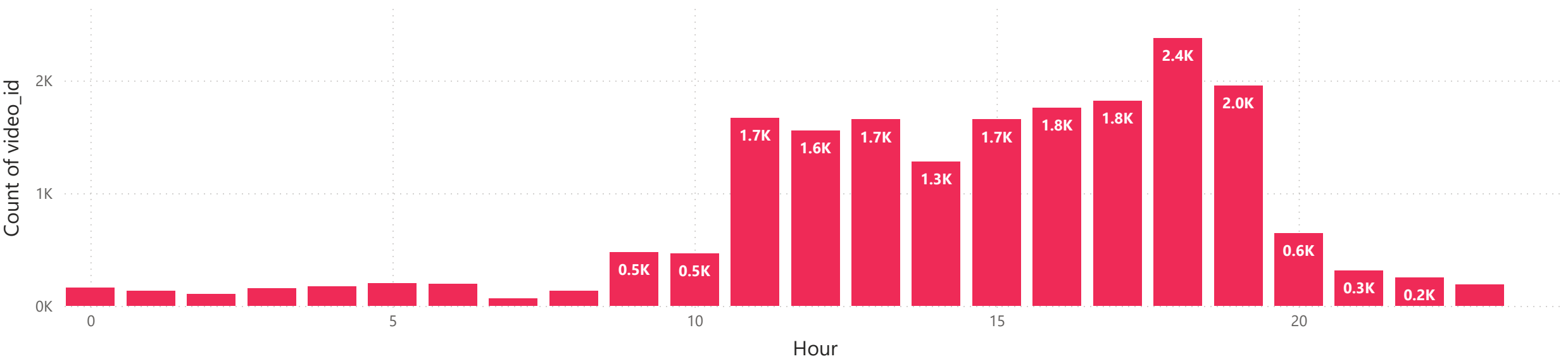
User Engagement Insights

- Tags Analysis: Analysis of tags with user engagement
- Peak engagement times : Analysis of peak engagement years, months ,hours and days of week .
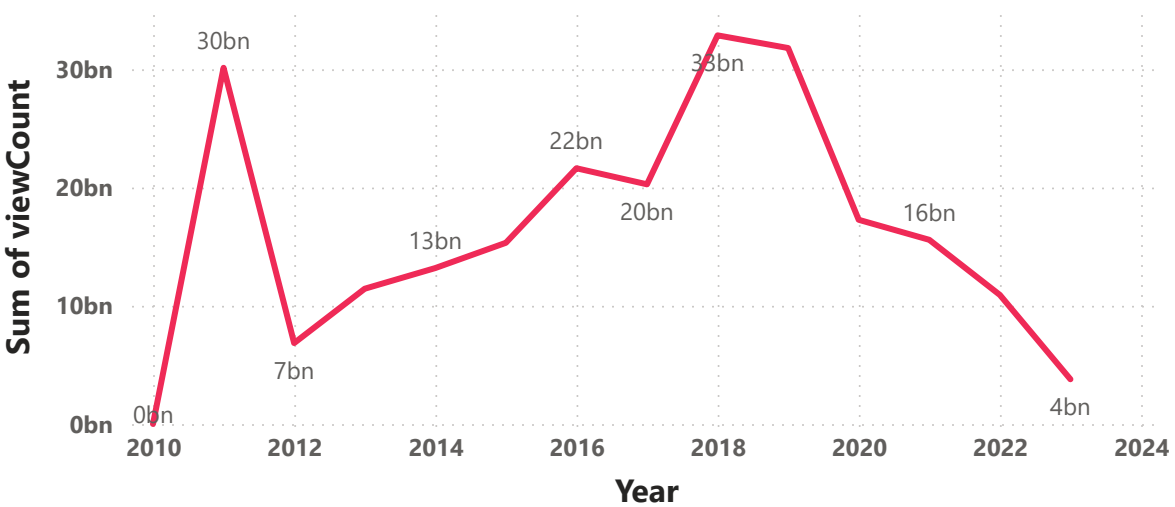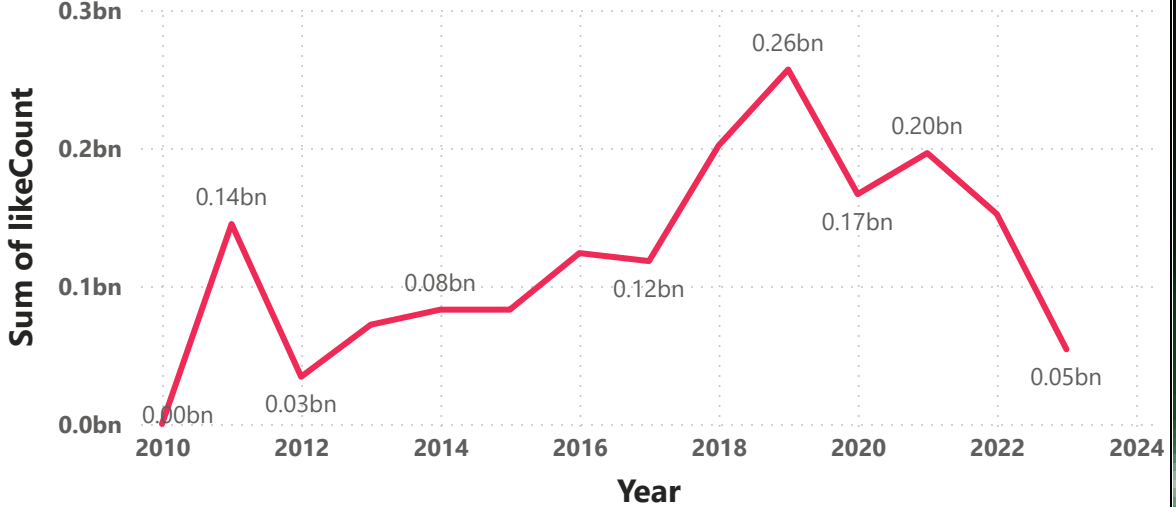- Videos Duration influence on user engagement : Analysis of views, likes ,comments with video duration.
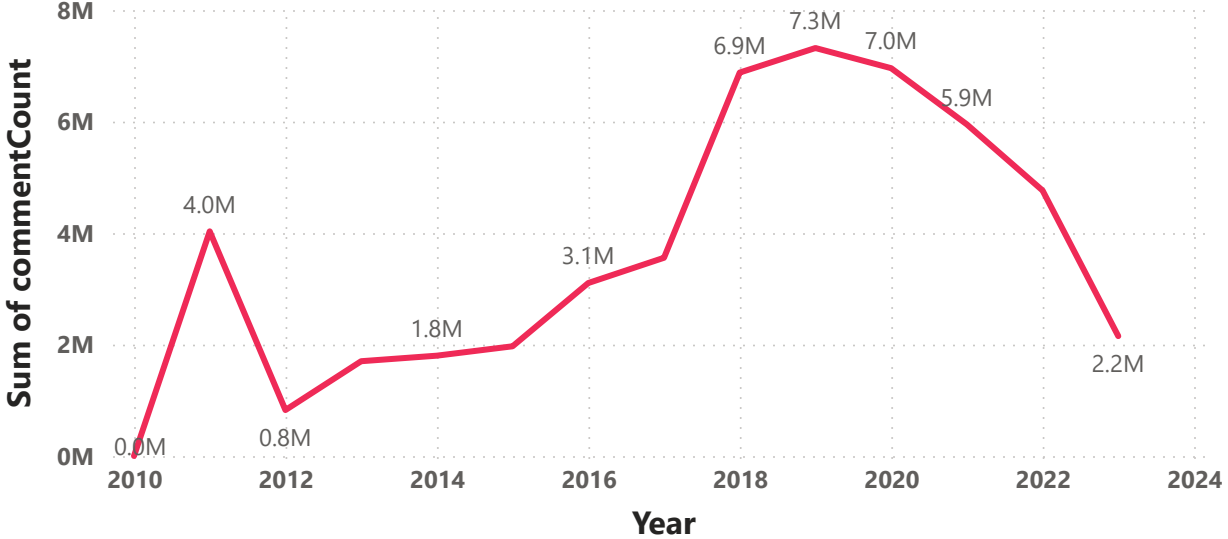
Yearly Engagement Statistics

Yearly View Trends

Yearly Like Trends

Yearly Comment Trends

Monthly Engagement Statistics

Monthly Trend of views

Sum of viewCount vs Month Name

- May: 37bn
- February: 22bn
- December: 21bn
- April: 20bn
- November: 19bn
- October: 19bn
- July: 19bn
- August: 17bn
- September: 15bn
- June: 15bn
- January: 14bn
- March: 13bn

Monthly Trend of likes

Sum of likeCount vs Month Name

- May: 211M
- November: 151M
- April: 150M
- February: 149M
- December: 148M
- October: 146M
- July: 137M
- August: 126M
- January: 125M
- September: 125M
- June: 112M
- March: 105M
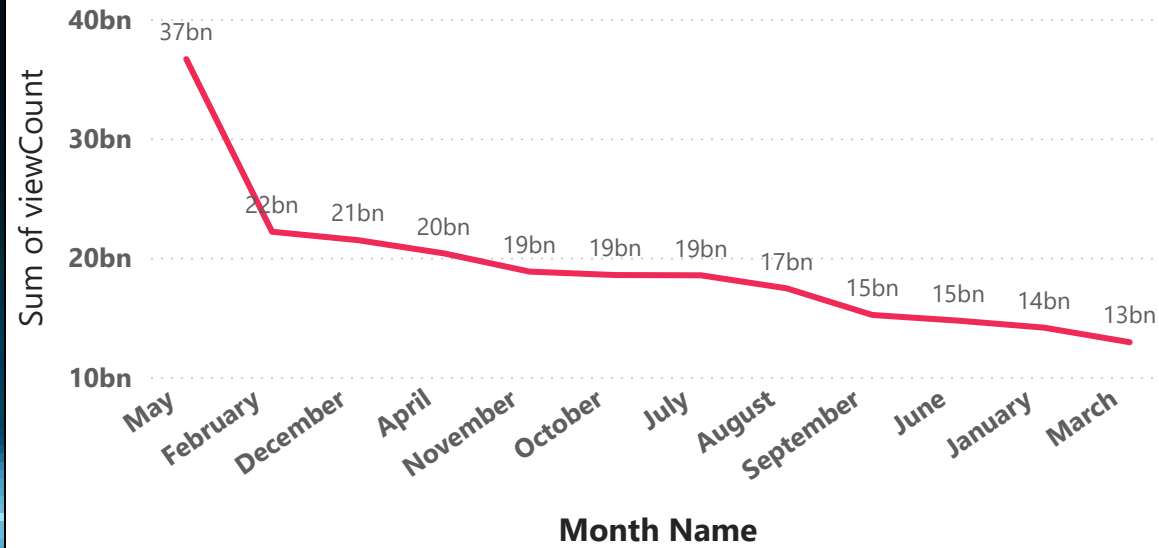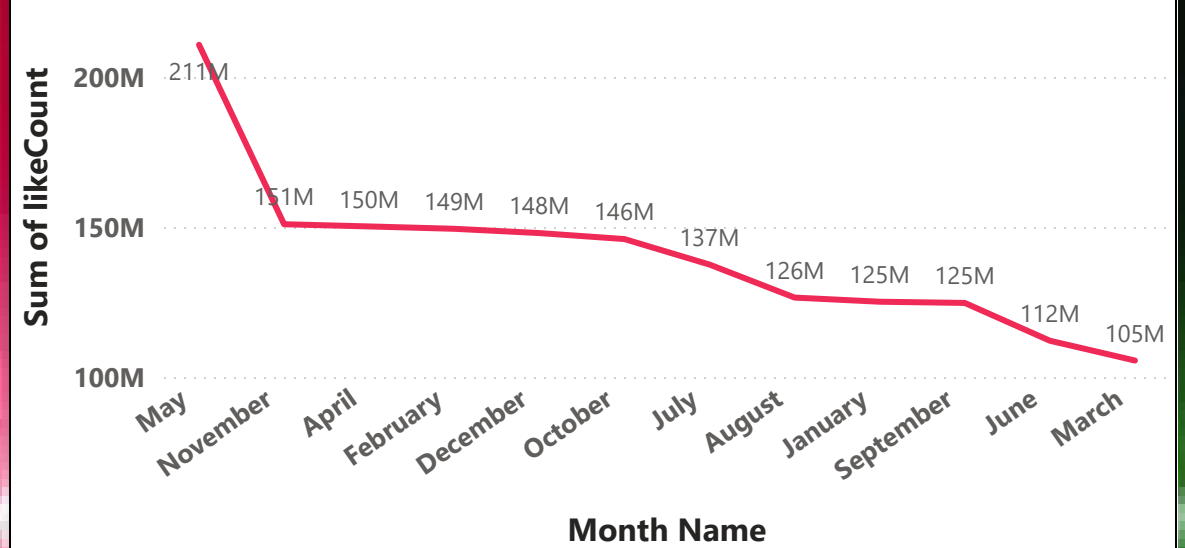
Monthly Trend of comments

Sum of commentCount vs Month Name

- May: 6.2M
- December: 4.9M
- April: 4.6M
- October: 4.6M
- November: 4.5M
- July: 4.3M
- February: 4.1M
- August: 4.0M
- September: 3.6M
- March: 3.6M
- June: 3.3M
- January: 3.3M
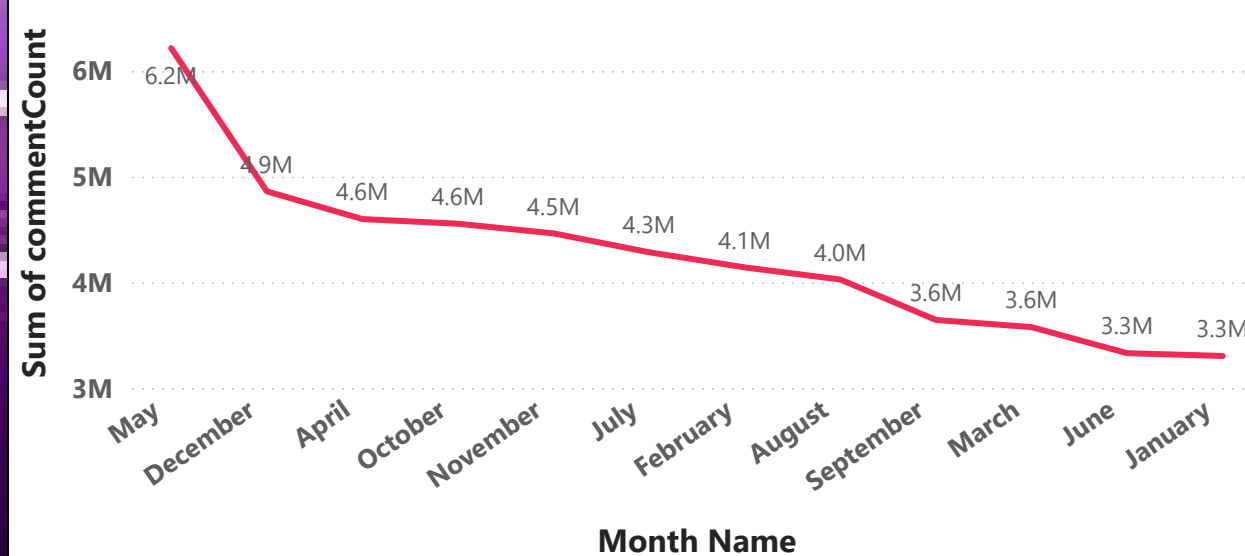
Weekly Engagement Statistics

Views on different days of week

Sum of viewCount vs Day Name

- Wednesday: 46bn
- Monday: 40bn
- Thursday: 39bn
- Friday: 37bn
- Tuesday: 37bn
- Saturday: 24bn
- Sunday: 8bn

Likes on different days of week

Sum of likeCount vs Day Name

- Wednesday: 0.33bn
- Thursday: 0.29bn
- Monday: 0.28bn
- Friday: 0.27bn
- Tuesday: 0.26bn
- Saturday: 0.18bn
- Sunday: 0.07bn

Comments on different days of week

Sum of commentC... vs Day Name

- Wednesday: 9.4M
- Monday: 9.0M
- Thursday: 8.6M
- Tuesday: 8.3M
- Friday: 8.2M
- Saturday: 5.2M
- Sunday: 2.3M
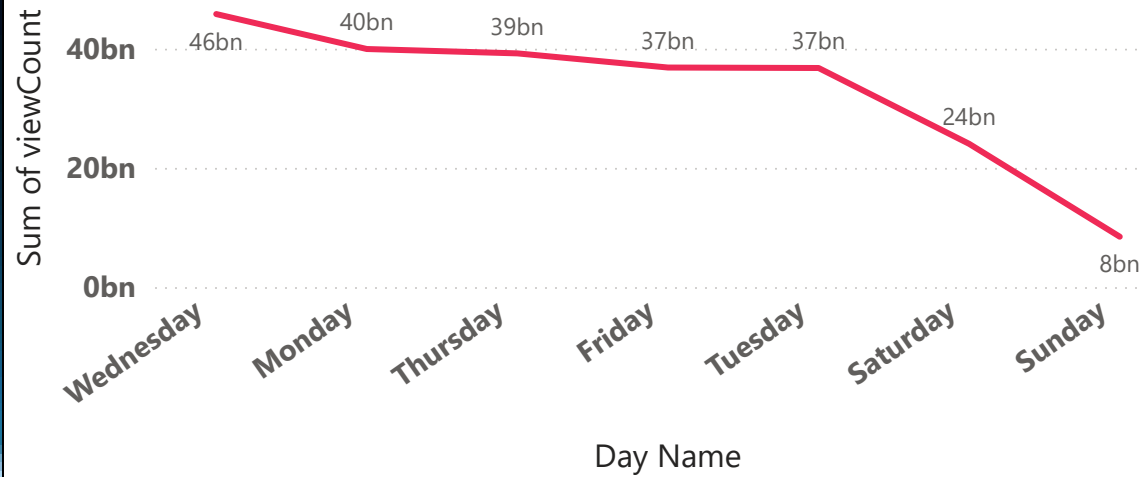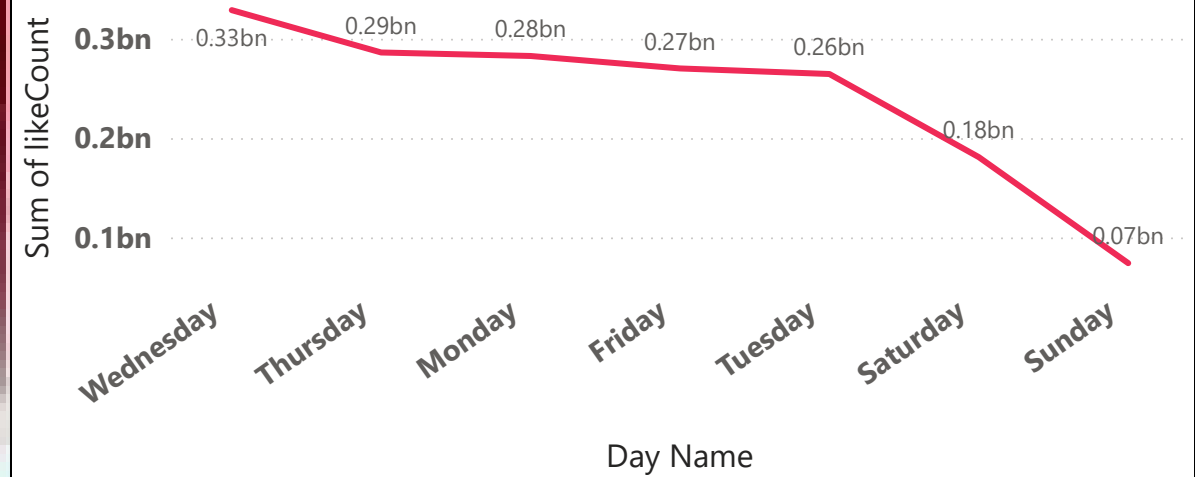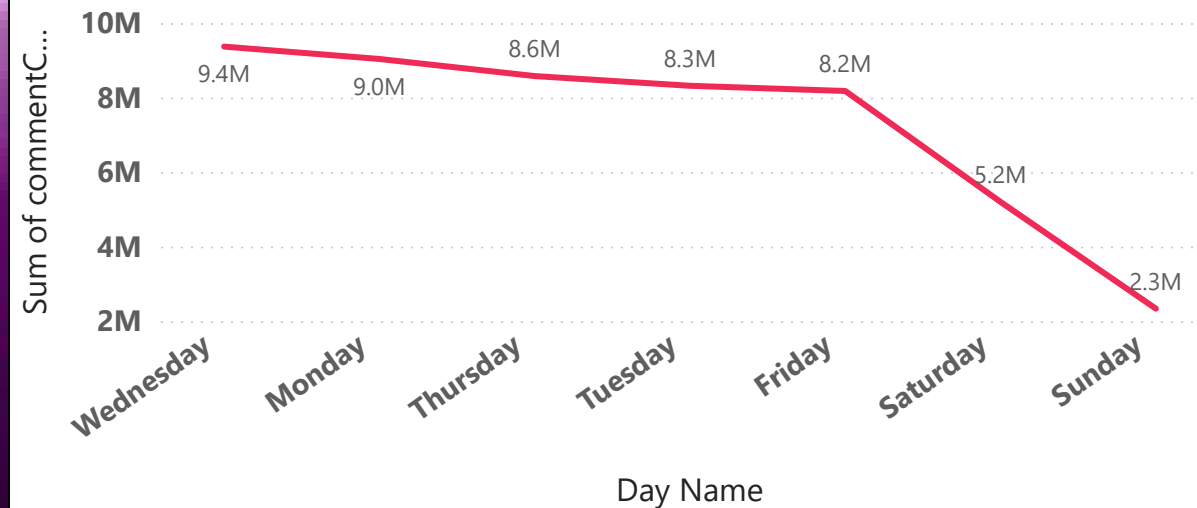
Daily Engagement Statistics

Views on different hours of day

Sum of viewCount vs Hour

- 3bn, 1bn, 0bn, 44bn, 15bn, 15bn, 22bn, 1bn, 3bn

Likes on different hours of day

Sum of likeCount vs Hour

- 0.02bn, 0.00bn, 0.00bn, 0.39bn, 0.11bn, 0.12bn, 0.14bn, 0.01bn, 0.01bn

Comments on different hours of day

Sum of commentCount vs Hour

- 0.4M, 0.1M, 13.8M, 5.1M, 5.3M, 3.1M, 3.5M, 0.3M, 0.4M
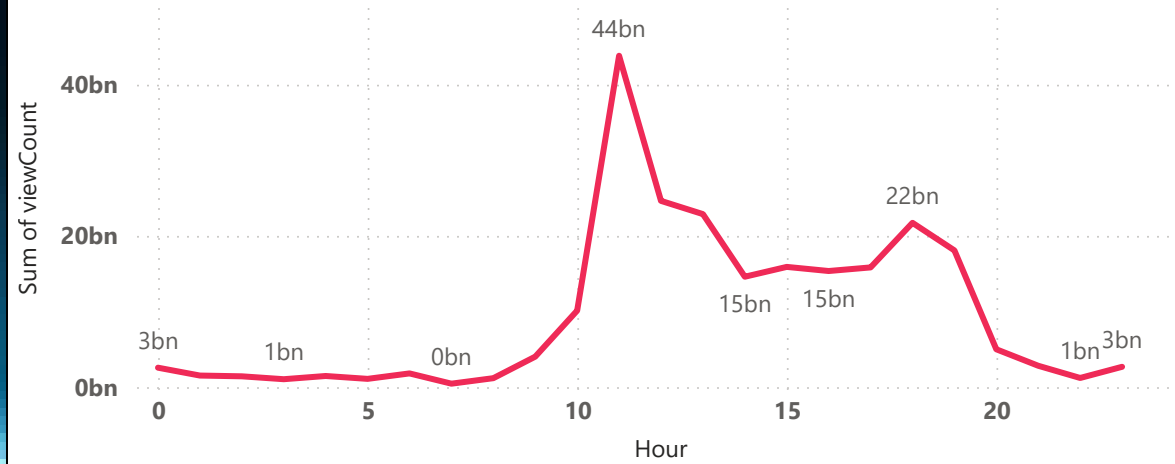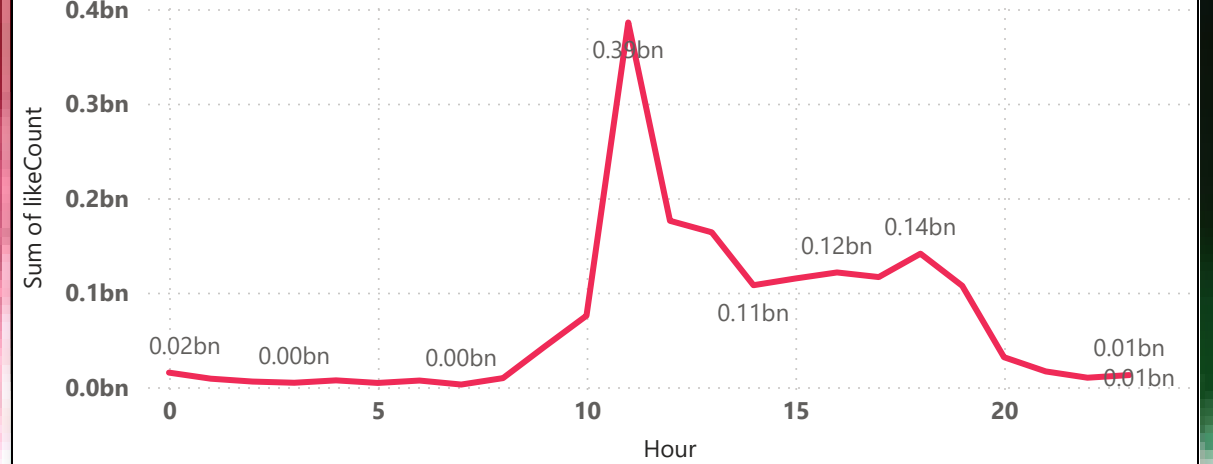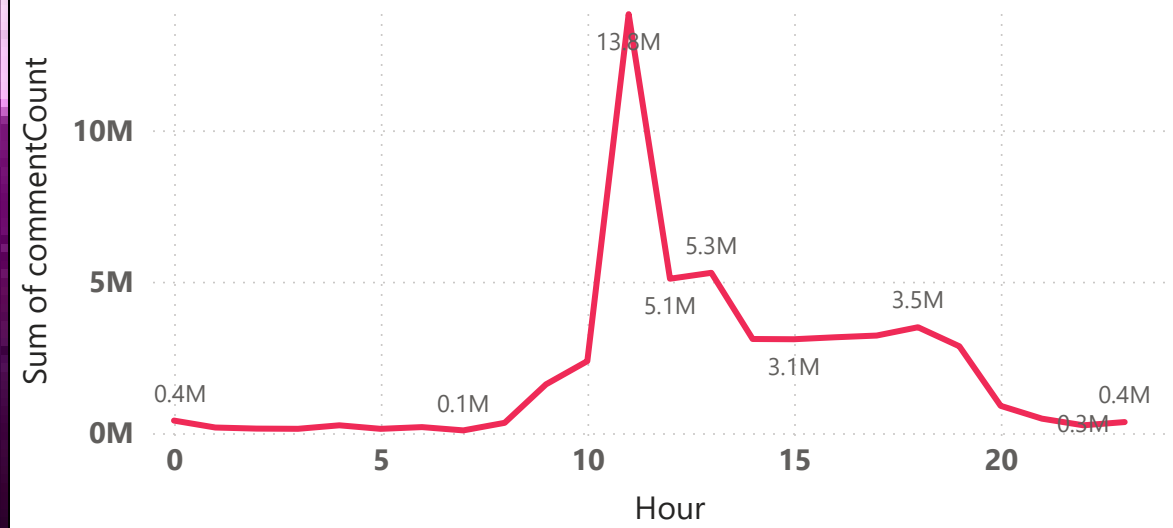
Engagements on dfferent hours on different days of the Week

Day Name ● Wednesday ● Monday ● Thursday ● Friday ● Tuesday ● Saturday ● Sunday

Tags Analysis

Top 30 tags with Total and Average Views

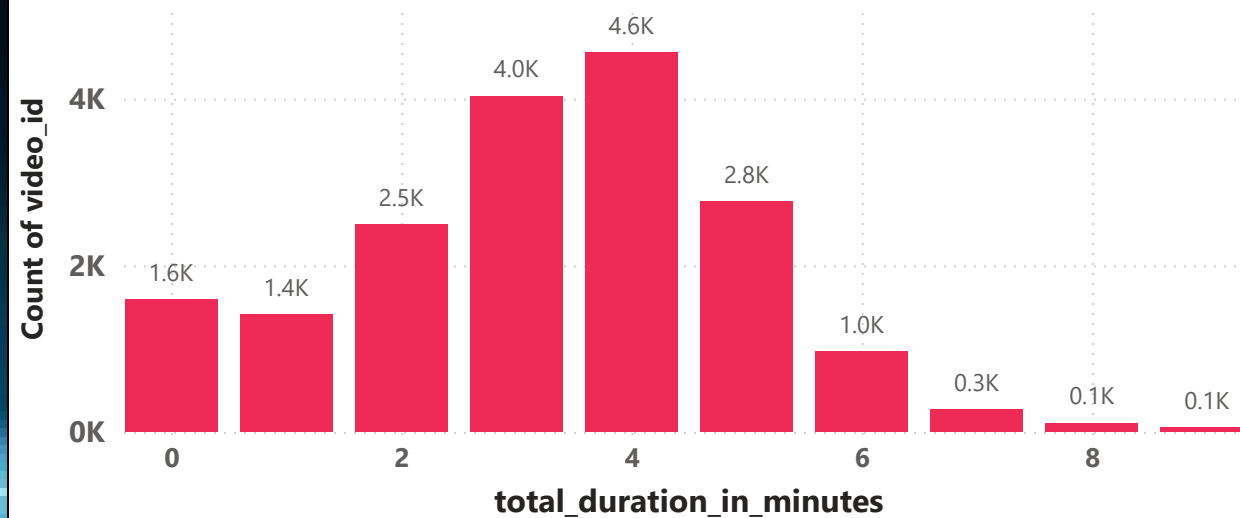| Sum of viewCount	Tag	Average of viewCount
21593887631	new songs 2019	21593887.63
33275935737	Latest Hindi Songs	16184793.65
33022837007	Bollywood	14112323.51
49088908201	New Hindi Songs	13914089.63
38798161439	T-Series	13584790.42
92193697196	SONGS	12760373.31
107310386421	BOLLYWOOD SONGS	12357253.16
26495016212	New Song	12243537.99
24281914324	Hindi Movie Songs	11787337.05
24409329963	Hindi Movies	11774881.80
26260190643	tseries Movies	11184067.57
21946515146	Hindi Songs	10962295.28
24015316224	Hindi Bollywood Song	10749917.74
102649045619	HIndi Songs	9858725.09
60886756553	Tseries Songs	9732537.81
109105770689	TSERIES	9573200.90
43378885713	NEW SONGS	9506659.15
25281459817	Bollywood Song	8224287.51
32920746687	LATEST SONGS	7888987.94
23995669851	Official Song	7785746.22
41872789027	Film Songs	7617389.31
24726324425	movie songs	6731915.17
22740265332	ries	6487950.17
22460403305	tseries offical channel	6335797.83
22112337544	T-Series Music	6281914.08
1054822550706		**10214517.23**

Popular tags by views
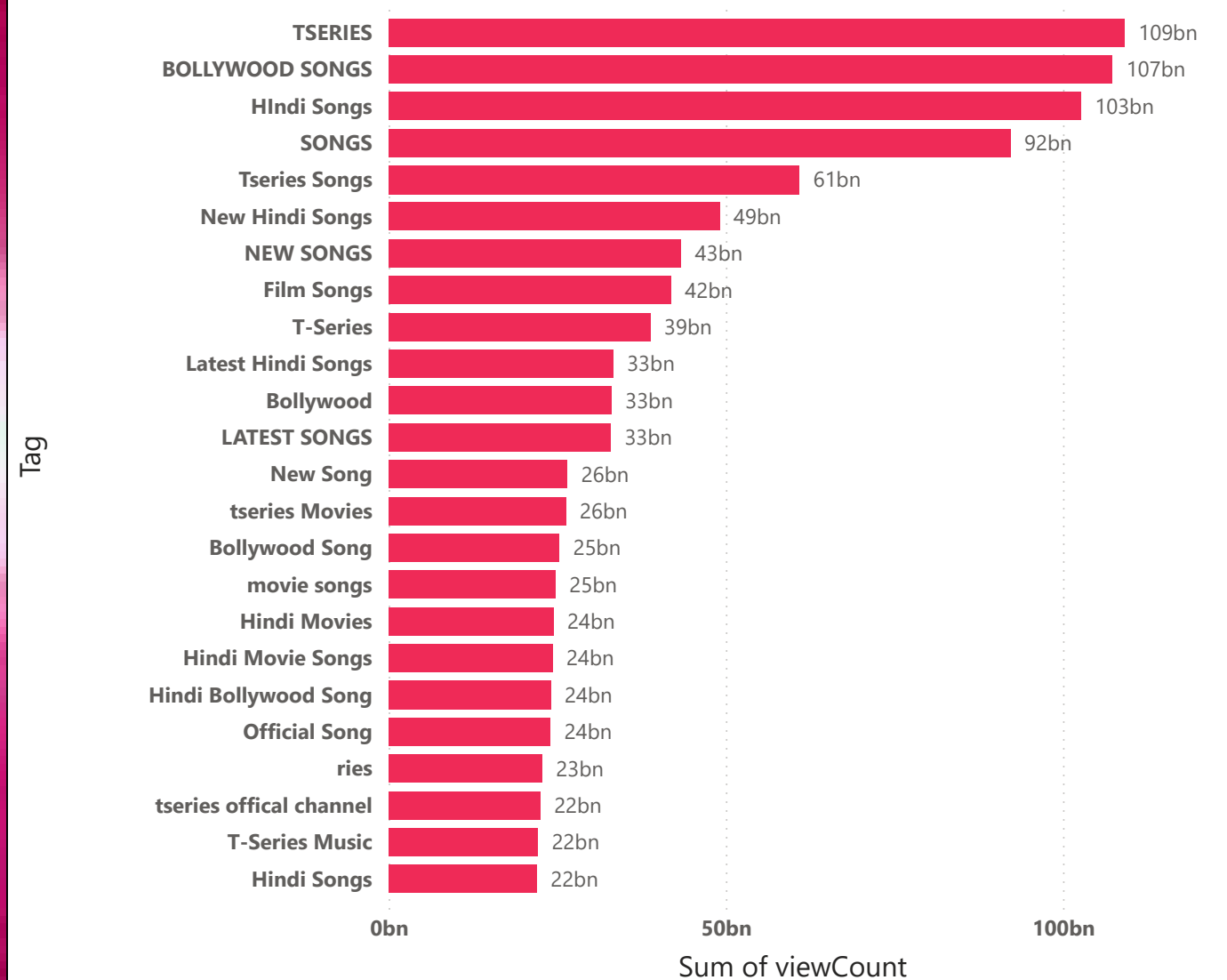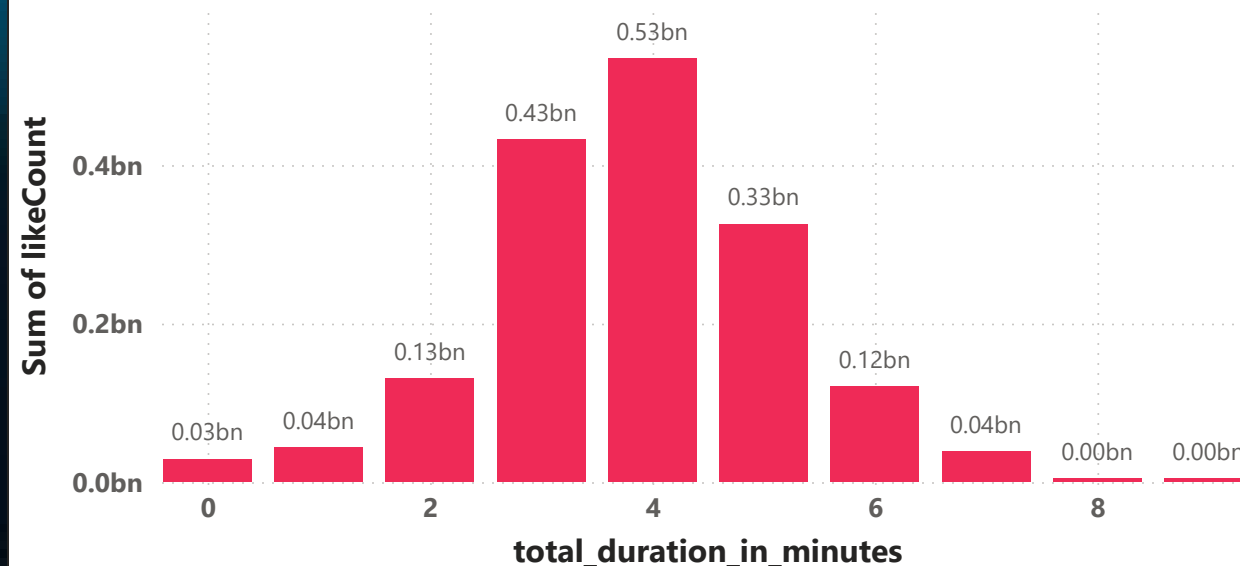
Duration of videos influence on user engagement

Variation of no. of videos by duration

Count of video_id vs total_duration_in_minutes:
- 0: 1.6K
- 1: 1.4K
- 2: 2.5K
- 3: 4.0K
- 4: 4.6K
- 5: 2.8K
- 6: 1.0K
- 7: 0.3K
- 8: 0.1K
- 9: 0.1K

Duration of videos influence on Views

Sum of viewCount vs total_duration_in_minutes:
- 0: 1bn
- 1: 2bn
- 2: 20bn
- 3: 60bn
- 4: 76bn
- 5: 45bn
- 6: 16bn
- 7: 6bn
- 8: 1bn
- 9: 1bn

Duration of videos influence on likes

Sum of likeCount vs total_duration_in_minutes:
- 0: 0.03bn
- 1: 0.04bn
- 2: 0.13bn
- 3: 0.43bn
- 4: 0.53bn
- 5: 0.33bn
- 6: 0.12bn
- 7: 0.04bn
- 8: 0.00bn
- 9: 0.00bn
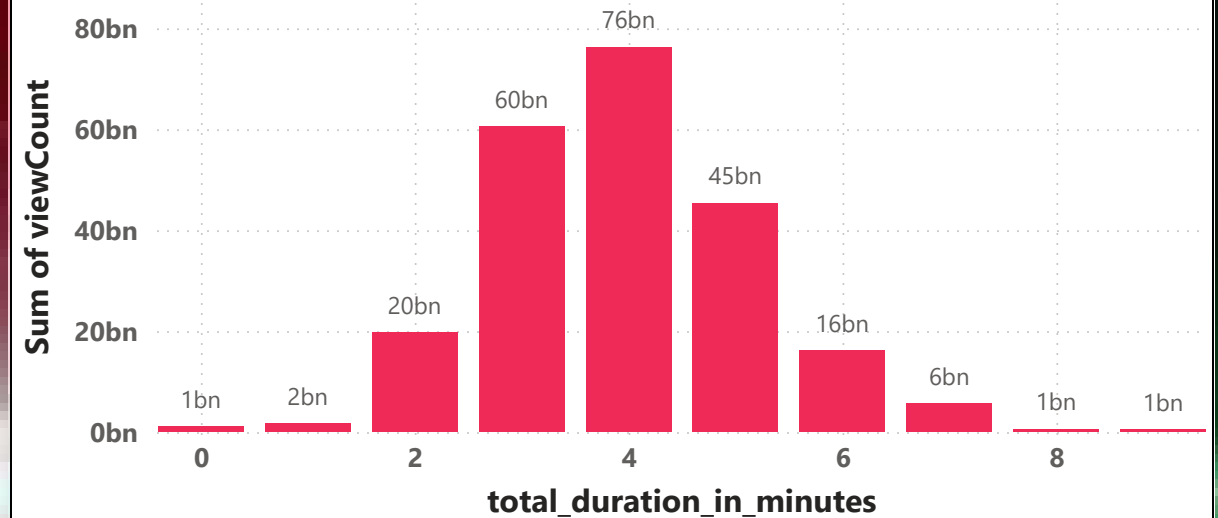
Duration of videos influence on comments

Sum of commentCount vs total_duration_in_minutes:
- 0: 0.8M
- 1: 1.2M
- 2: 4.1M
- 3: 14.0M
- 4: 16.3M
- 5: 9.0M
- 6: 3.3M
- 7: 1.1M
- 8: 0.1M
- 9: 0.1M
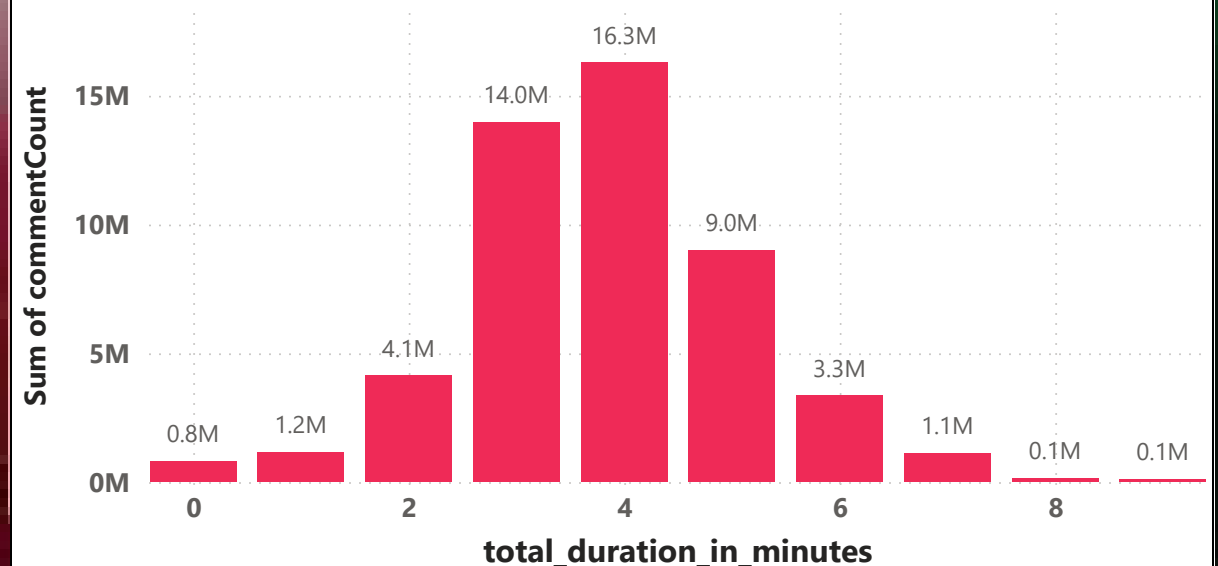
Conclusion

1. Total and Unique Song Data:

- The dataset includes 19,345 unique songs, all from the T-Series channel.
- The data spans from 2010 to 2023, with notable variations in song releases and engagement over these years.

2. Temporal Analysis:

- Peak years for song releases include 2011, 2022, and 2016.
- Monthly analysis shows that May has the highest number of song releases, followed by April and July.
- Daily analysis indicates that Monday is the most active days for publishing songs, while Sunday sees the least activity.
- Hour 18 followed by hour 19 are the most active days for publishing songs.

3. Engagement Metrics:

- Statistical overview shows high variability in view counts, like counts, and comment counts. - Maximum view count reaches over 1.5 billion, with an average of approximately 11.9 million views .
- Likes and comments also exhibit significant variability, with maximums of 12.8 million likes and 420,375 comments.

Conclusion

4. Content Quality and Accessibility:
 - The majority of videos are in high definition (HD), accounting for 16,583 videos, while 2,762 are in standard definition (SD).
 - Most videos do not have captions, with only 201 videos including captions.

5. Engagement Analysis:
-The year 2018 , 2019 see higher engagement in terms of views, likes and comments.
- The month of May stands out with significantly higher engagement metrics, followed by November , December and February.
- Weekdays, especially Wednesday followed Monday and Thursday, see higher engagement in terms of views, likes and comments.
-The hour 11 and 12 see higher engagement in terms of views, likes and comments.
- Hour 11 on Wednesday is the best engagement time.

6.Tags analysis
-TSERIES tag see highest engagement followed by BOLLYWOOD SONGS and Hindi Songs .

7.Duration:
-Duration of 4 minutes followed by 3 and 5 minuteshas higher engagement in terms of views, likes and comments.

Recommendations

1. Optimal Publishing Schedule: - To maximize engagement, content creators should consider publishing songs on Wednesday and Monday, particularly in May, November or December . The Hour 11 and 12 are the best publishing times.

2. Content Quality: - Maintaining high-definition video quality can enhance user experience and engagement.

3. Caption Inclusion: - Increasing the availability of captions can make content more accessible to a broader audience potentially boosting engagement.

4. Duration: - Maintaining a duration around 3 to 5 minutes .Best duration is about 4 minutes.

5. Tags: - Use tags giving higher views while publishing the videos like TSERIES, BOLLYWOOD SONGS and Hindi Songs .