YouTube Songs Data Analysis Report

Objective

The primary objective of this analysis was to explore song-related metadata from YouTube to uncover patterns in viewer engagement, such as views, likes, comments, and video duration, using a dataset of 19,345 entries.

Dataset Overview

- Source: songs.xlsx
- Total Records: 19,345
- Key Features:
  - title, description, tags: Metadata of each song
  - viewCount, likeCount, commentCount: Engagement metrics
  - duration: Length of the video
  - definition, caption: Video quality and caption availability
  - publishedAt: Publishing date and time

Initial Data Insights

- The dataset includes only videos uploaded by "T-Series".
- Most songs are labeled as Lo-Fi Mixes.
- Video quality is consistently high (hd).
- All caption values are False.

Exploratory Data Analysis (EDA)

- Views vs Likes: Positive correlation.
- Comments: Active audience engagement.
- Duration: Most between 3 to 6 minutes.
- Publishing Trend: Spikes on Aug 20, 2023.

Visualizations Used

YouTube Songs Data Analysis Report

- Histogram of video durations

- Scatter plot of viewCount vs likeCount

- Line graph of viewCount by publishedAt

- Bar chart showing top songs by engagement

Observations

- Top-performing song: "Har Har Mahadev" (100,000+ views, 6,230 likes)

- Religious songs received more engagement.

- Lo-Fi mixes had consistent viewer interest.

Conclusion

The dataset gives a detailed view of how song types perform on YouTube. Key insights include:

- Correlation between likes and views

- Standardized song durations

- Influence of upload timing