

Tik-Tok Executive Summary

Engagement Metrics for Claim Status Videos

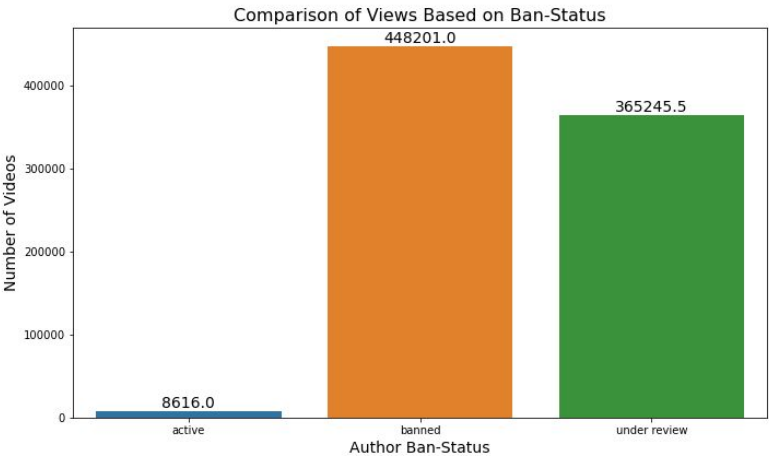
Project Overview

We are performing exploratory data analysis on the engagement metrics for two classes of videos posted on Tik-Tok. We want to clean, structure and validate the data provided to us for further analysis on building a ML model to classify the videos.

Key Insights

- Claim videos are superior to opinion videos in all engagement metrics (views, likes, comment, shares, and downloads)
- Claim video have a higher banned to active author status than videos classified in the opinion category.
- There is a visible trend in the relationship between comment-count and like-count.
- 99% of all views in this dataset come from ‘claim’ videos.
- All engagement metrics are heavily right-skewed with a significant amount of outliers.
 - Outliers are defined as: $\text{median} + 1.5 * \text{IQR}$

Details



Comparison of number of views per author ban status. Blue column indicates the views obtained by all active authors in the platform. Orange represents the views obtained by banned authors, and green represents the views obtained by authors whose status is still under review.

Next Steps

- Manage outliers, by defining a threshold and removing them from the analysis.
- Structure data frames to look only at active authors to understand the present day content available on the platform.
- Look into statistical methods to model the comment to likes relationship that was formerly mentioned.