Executive Summary: Hypothesis Testing Results

TikTok Claims Classification Project

Project Overview

As the TikTok data team works on developing a machine learning model to assist in the classification of claims for user submissions, they are interested in learning if there a relationship between view count and the verification status of the video's author. They will conduct a hypothesis test.

Key Insights

- The analysis found that there is a significantly significant difference in view counts between TikToks posted by verified and unverified users.
- This suggests that there might be only behavioral differences between the 2 groups

Details

The team looked into the relationship between the verified_status field and the video_view_count field. They did so by first calculating the mean view count for each status as shown in the graphic below. It is clear that videos posted by verified users have, on average, a lot less views.

```
verified_status
not verified    265663.785339
verified         91439.164167
Name: video_view_count, dtype: float64
```

The team then conducted a two-sample t test to determine the significance of the hypothesis. The team found that there is indeed a significant difference.

Next Steps

The TikTok data team suggests that the next step be building a regression model on verified status. This model can continue to help the team analyze user behavior in the verified status groups.