

Video Content User Claim Prediction Project Preliminary Data Summary

Executive summary report for TikTok

Overview

TikTok's data team needs help in preparing the data needed for the claims classification project. The dataframe for the claims data will be built, the data organized for the process of exploratory data analysis, and the data team will be updated on the data analysis insights.

Objective

- Build a dataframe for the TikTok dataset
 - Read in data from TikTok csv file
 - Display rows within dataframe
 - Examine data type of each column
 - Gather descriptive statistics
 - Visualize the TikTok data in Python
-

Results

- Used a Jupyter notebook.
 - Read data from the provided tiktok_dataset.csv file into a dataframe.
 - Conducted exploration of the dataframe.
 - Used code to count the number of outliers in numeric columns.
 - Counted the number of videos with different claim_status.
 - Grouped data by claim_status and author_ban_status then calculated mean and median values.
 - Created three new columns amount of likes/comments/shares per view for a video.
-

Next Steps

- Conduct a complete exploratory data analysis.
 - Perform any data cleaning and data analysis steps to understand anomalies such as outliers.
 - Use descriptive statistics to learn more about the data.
- 