# Executive Summary

Tiktok claims classification Project

## Overview

The data team at tiktok is trying to make a machine learning model to classify videos as claims or opinions. After the approval of the project proposal, now exploration and Transformation of data needs to be done for future analysis.

## Objective

For further analysis, understanding the data, strong and weak indicators and the distribution of data points is important.

## Results

Initial exploration of data revealed a relation between the claim status and ban status of the creator. Videos with claims are more thoroughly scrutinized compared to ones with opinions.

Also the video engagement stats are higher for claim videos than opinion videos.

The data is divided 50-50 making it suitable to be used for training.

New columns were created based on existing ones since they proved more valuable.

## Next Steps

The EDA process can now begin since transformation of data was performed by creating new necessary columns. The categories are almost 50% distributed making the dataset suitable for training purposes.