

# Data dictionary

This project uses a dataset called `tiktok_dataset.csv`. It contains synthetic data created for this project in partnership with TikTok. Examine each data variable gathered.

**19,383 rows** – Each row represents a different published TikTok video in which a claim/opinion has been made.

## 12 columns

Column name	Description
#	TikTok assigned number for video with claim/opinion.
claim_status	Whether the published video has been identified as an “opinion” or a “claim.” In this dataset, an “opinion” refers to an individual’s or group’s personal belief or thought. A “claim” refers to information that is either unsourced or from an unverified source.
video_id	Random identifying number assigned to video upon publication on TikTok.
video_duration_sec	How long the published video is measured in seconds.
video_transcription_text	Transcribed text of the words spoken in the published video.
verified_status	Indicates the status of the TikTok user who published the video in terms of their verification, either “verified” or “not verified.”
author_ban_status	Indicates the status of the TikTok user who published the video in terms of their permissions: “active,” “under scrutiny,” or “banned.”
video_view_count	The total number of times the published video has been viewed.
video_like_count	The total number of times the published video has been liked by other users.
video_share_count	The total number of times the published video has been shared by other users.
video_download_count	The total number of times the published video has been downloaded by other users.
video_comment_count	The total number of comments on the published video.