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** | **Type** | **Description** |
| --- | --- | --- |
| # | int | TikTok assigned number for video with claim/opinion. |
| claim\_status | obj | 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 beliefs or thoughts. A “claim” refers to information that is either unsourced or from an unverified source. |
| video\_id | int | Random identifying number assigned to a video upon publication on TikTok. |
| video\_duration\_sec | int | How long the published video is measured in seconds. |
| video\_transcription\_text | obj | Transcribed text of the words spoken in the published video. |
| verified\_status | obj | Indicates the status of the TikTok user who published the video in terms of their verification, either “verified” or “not verified.” |
| author\_ban\_status | obj | Indicates the status of the TikTok user who published the video in terms of their permissions: “active,” “under scrutiny,” or “banned.” |
| video\_view\_count | float | The total number of times the published video has been viewed. |
| video\_like\_count | float | The total number of times the published video has been liked by other users. |
| video\_share\_count | float | The total number of times the published video has been shared by other users. |
| video\_download\_count | float | The total number of times the published video has been downloaded by other users. |
| video\_comment\_count | float | The total number of comments on the published video. |