|  |  |  |
| --- | --- | --- |
| A | Id | Row id |
| B | CANDIDATE | Candidate mentioned |
| C | CANDIDATE\_CONFIDENCE | Confidence of the candidate mentioned |
| D | RELEVANT\_YN | "no" means that the tweet was meant to be part of the dataset but was not available when contributors went to judge it |
| E | RELEVANT\_YN\_CONFIDENCE | confidence in the existence/non-existence of the tweet |
| F | SENTIMENT | Tweet Sentiment |
| G | SENTIMENT\_CONFIDENCE | Confidence of the sentiment |
| H | SUBJECT\_MATTER | Tweet subject |
| I | SUBJECT\_MATTER\_CONFIDENCE | Confidence of the subject matter |
| J | CANDIDATE\_GOLD | whether the candidate was included in the gold standard for the model |
| K | NAME | the user who tweeted |
| L | RELEVANT\_YN\_GOLD | whether the tweet yn value is golden |
| M | RETWEET\_COUNT | number of times the user has retweeted |
| N | SENTIMENT\_GOLD | if the profile is golden, what is the sentiment |
| O | SUBJECT\_MATTER\_GOLD | whether the subject matter was included in the gold standard for the model |
| P | TEXT | the text of the tweet |
| Q | TWEET\_COORD | if the user in column K has location turned on, the coordinates as a string with the format "[*latitude*, *longitude*]" |
| R | TWEET\_CREATED | When tweet was created |
| S | TWEET\_ID | Tweet identification number |
| T | TWEET\_LOCATION | User’s country, city, state |
| U | USER\_TIMEZONE | User time zone |