```
CREATE EXTERNAL TABLE IF NOT EXISTS twitter_joey.twitter_sent (
  `id` string,
  `location` string,
  `sentiment` int,
  `vader_label` int,
  `party` string
)
ROW FORMAT SERDE 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
WITH SERDEPROPERTIES (
  'serialization.format' = '     ',
  'field.delim' = ' ',
  'collection.delim' = '•',
  'mapkey.delim' = '•'
) LOCATION 's3://wcd-athena/twitter_sent/twitter_sent/'
TBLPROPERTIES ('has_encrypted_data'='false');

select * from twitter_sent

create table twitter_predict as
select * from twitter_sent
where id != 'id';

select * from twitter_predict limit 5;
```

```
CREATE EXTERNAL TABLE IF NOT EXISTS twitter_joey.twitter_df (
  `id` string,
  `vader_label` int,
  `hour` int,
  `dayofweek` int,
  `followers_cnt` int,
  `party` string,
  `tweet` string,
  `location` string
)
ROW FORMAT SERDE 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
WITH SERDEPROPERTIES (
  'serialization.format' = '     ',
  'field.delim' = ' ',
  'collection.delim' = '•',
  'mapkey.delim' = '•'
) LOCATION 's3://wcd-athena/twitter_df/twitter_df/'
TBLPROPERTIES ('has_encrypted_data'='false');

create table twitter_vader as
select * from twitter_df
where id != 'id';
```

```sql
select * from twitter_vader limit 50

create table twitter_vader_word as
(WITH sub AS (
  SELECT id, split(tweet, ' ') as split_string FROM twitter_vader
  )
SELECT id, word
FROM sub
CROSS JOIN UNNEST(split_string) as t(word));

select * from twitter_vader_word limit 10;
```

Run query | Save as | Create | (Run time: 0.44 seconds, Data scanned: 582.29 KB)

Use Ctrl + Enter to run query, Ctrl + Space to autocomplete | Athena engine versi

Results

	id	word
1	1430965557738840074	years
2	1430965557738840074	tax
3	1430965557738840074	sale
4	1430965557738840074	home
5	1430965557738840074	liberal
6	1430965557738840074	platform
7	1430965571076509696	joined

```sql
CREATE TABLE twitter_noparty as(SELECT * FROM twitter_vader
WHERE party = 'No Party');

CREATE TABLE noparty_word as (WITH sub AS (
  SELECT id, split(tweet, ' ') as split_string FROM twitter_noparty
  )
SELECT id, word
FROM sub
CROSS JOIN UNNEST(split_string) as t(word));

CREATE TABLE twitter_liberal as(SELECT * FROM twitter_vader
WHERE party = 'Liberal');

CREATE TABLE liberal_word as (WITH sub AS (
  SELECT id, split(tweet, ' ') as split_string FROM twitter_liberal
  )
SELECT id, word
FROM sub
CROSS JOIN UNNEST(split_string) as t(word));
```

```
CREATE TABLE twitter_conservative as (SELECT * FROM twitter_vader
                                      WHERE party = 'Conservative')

CREATE TABLE conservative_word as (WITH sub AS (
  SELECT id, split(tweet, ' ') as split_string FROM twitter_conservative
  )
SELECT id, word
FROM sub
CROSS JOIN UNNEST(split_string) as t(word));


CREATE TABLE twitter_ndp as (SELECT * FROM twitter_vader
                                      WHERE party = 'NDP')

CREATE TABLE ndp_word as (WITH sub AS (
  SELECT id, split(tweet, ' ') as split_string FROM twitter_ndp
  )
SELECT id, word
FROM sub
CROSS JOIN UNNEST(split_string) as t(word));
```