//Table to Store Twitter JSON Data.

CREATE EXTERNAL TABLE covid\_raw\_tweets(id BIGINT,created\_at STRING,source STRING,favorited BOOLEAN, retweeted\_status STRUCT<text:STRING, `user` :STRUCT<screen\_name:STRING,name:STRING>,retweet\_count:INT>,text STRING,entities STRUCT<hashtags:ARRAY<STRUCT<text:STRING>>>,`user` STRUCT<screen\_name:STRING,friends\_count:INT,followers\_count:INT,statuses\_count:INT,verified:BOOLEAN,utc\_offset:INT,time\_zone:STRING>,in\_reply\_to\_screen\_name STRING)

ROW FORMAT SERDE 'org.apache.hive.hcatalog.data.JsonSerDe'

// Load the data from HDFS Path to the above-created Table.

load data inpath '/user/maria\_dev/covid\_data\_day (1)' into TABLE covid\_raw\_tweets;

//Divide the text into words.

create view temp\_1 as select id,covid\_raw\_tweets.text, words from covid\_raw\_tweets lateral view explode(sentences(lower(text))) dummy as words;

//Divide words into a single word row.

create view temp\_2 as select id,temp\_1.text, word from temp\_1 lateral view explode( words ) dummy as word ;

//Now we import our Dicyinary file as well.

CREATE EXTERNAL TABLE dictionary (

type string,

length int,

word string,

pos string,

stemmed string,

polarity string

)

ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t'

STORED AS TEXTFILE

LOCATION '/user/maria\_dev/dictionary';

//load dictionary file into the table:

load data inpath '/user/maria\_dev/dictinary.tsv' into TABLE covid\_raw\_tweets;

//Calculate Polarity by joining with dictionary.

create view temp\_3 as select

id,temp\_2.text,

temp\_2.word,

case s\_d.polarity

when 'negative' then -1

when 'positive' then 1

else 0 end as polarity

from temp\_2 left outer join dictionary s\_d on temp\_2.word = s\_d.word;

//Sum single word polarity value for every single user based on UsedId and assign sentiment.

create table tweets\_sent\_final stored as orc as select

id,

case

when sum(polarity) > 0 then 'positive'

when sum(polarity) < 0 then 'negative'

else 'neutral' end as sentiment,text

from temp\_3 group by id, text;