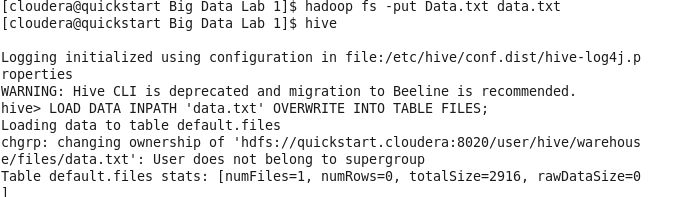
**Big Data Lab 3**

**MapReduce In Hive**

**Gaurav Ankalagi J004**

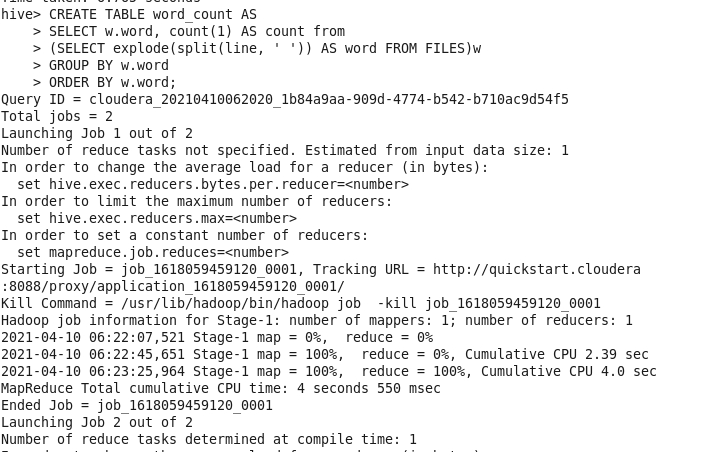
**Aim: Word Count Using Hive**

Step 1: Create a text file with some words in it. Make sure it has some words repeating it so that the word count program is able to show up some repeated words and not frequency for each word as 1.

Step 2: Move the created file to the Hadoop file system using the command hadoop fs -put data.txt data.txt Launch hive and put the file in in a table using the command > LOAD DATA INPATH 'data1.txt' OVERWRITE INTO TABLE FILES; 

Step 3: Create a table word word\_count to store the frequency of each word in the table.

> CREATE TABLE word\_count AS

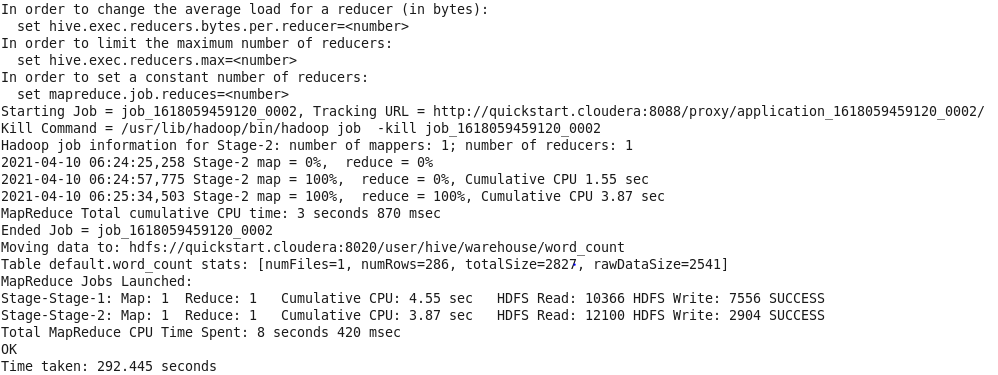


> SELECT w.word, count(1) AS count from

> (SELECT explode(split(line, ' ')) as word from FILES) w

> GROUP BY w.word

> ORDER BY w.word;



Step 4: Displaying the word counts stored in the word\_count table with command.

> select \* from word\_count;

