**EXPERIMENT-3**

**Om Hedau**

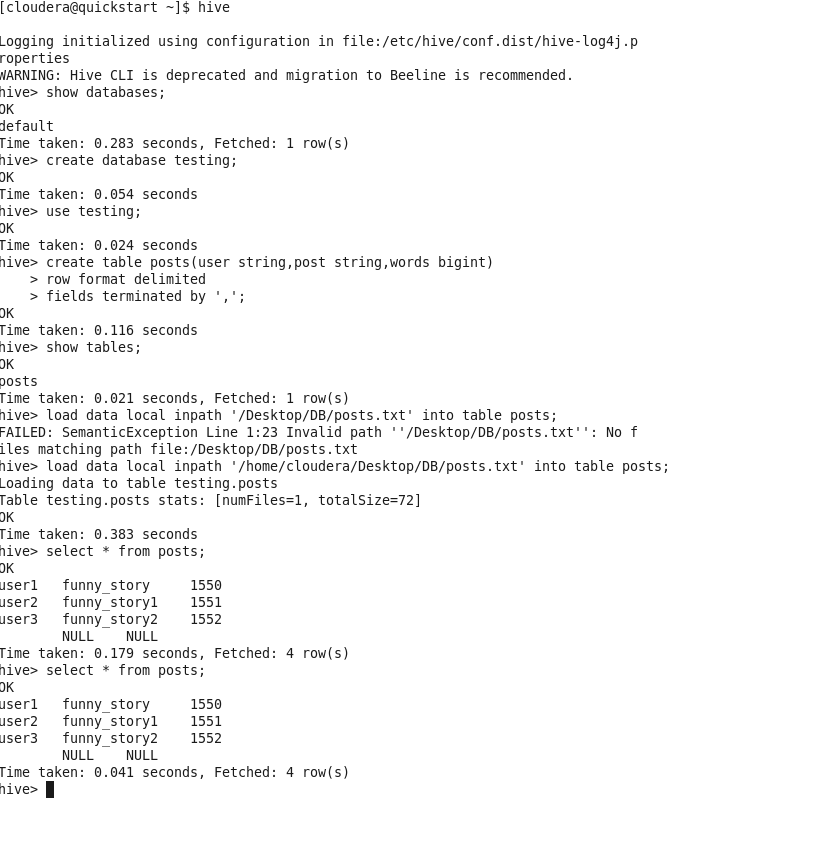
**BE-3-25**

* hive
* show databases
* create database testing;
* use testing;
* create table posts (user string, post string, words bigint)

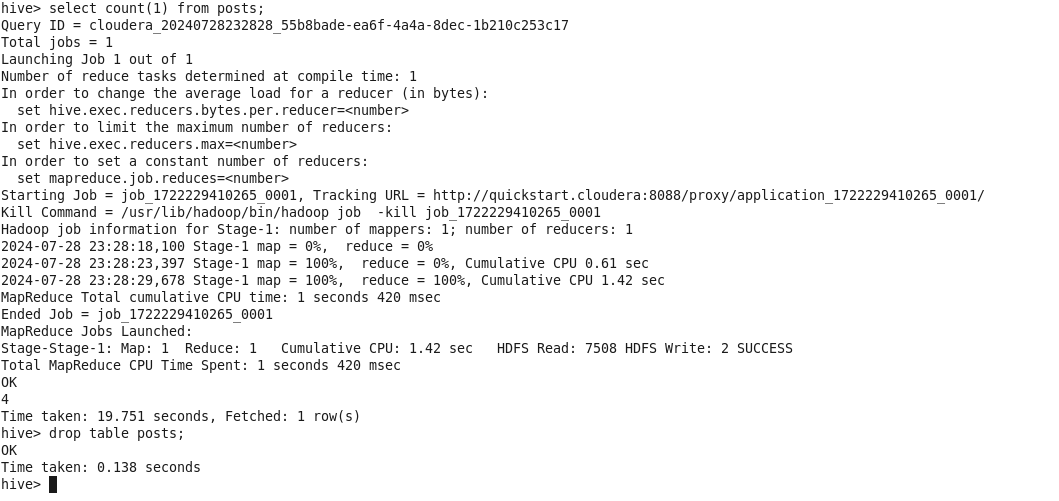
row format delimited

fields terminated by ‘,’;

* show tables;
* load data local inpath ‘/home/cloudera/Desktop/DB/posts.txt’ into table posts;
* select \* from posts;



* select count (1) from posts;
* drop table posts;



* CREATE EXTERNAL TABLE posts

(user STRING, post STRING, time BIGINT)

ROW FORMAT DELIMITED

FIELDS TERMINATED BT ‘,’

STORED AS TEXTFILE;

* CREATE TABLE posts1 (user STRING, post STRING, time BIGINT)

PARTITIONED BY (country STRING)

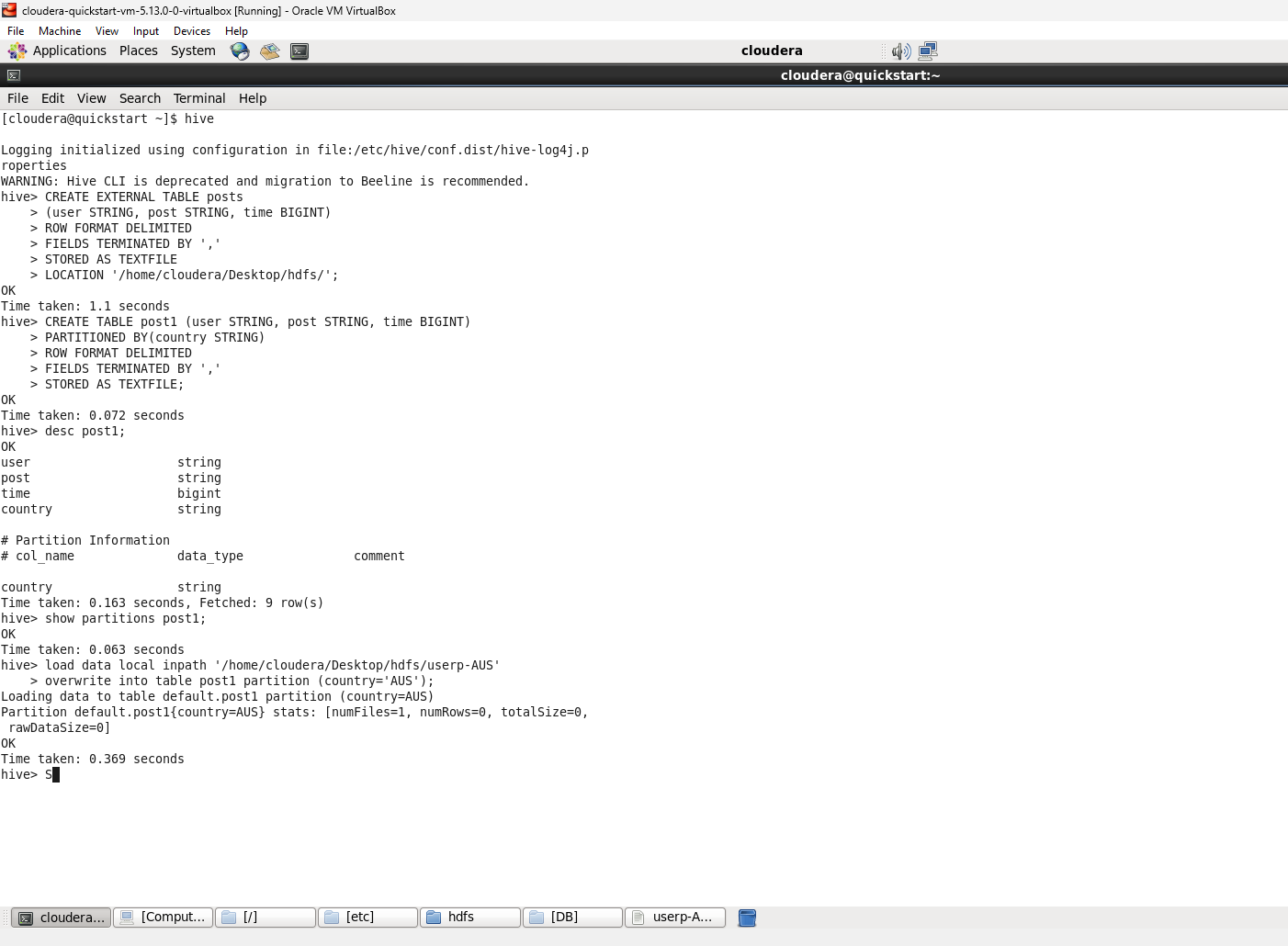
ROW FORMAT DELIMITED

FIELDS TERMINATED BY ‘,’

STORED AS TEXTFILE;

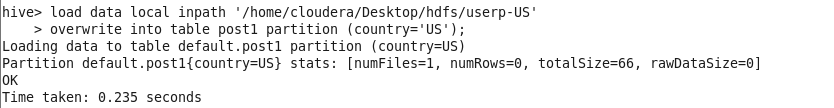
* desc post1;
* show partitions post1;
* load data local inpath ‘/home/cloudera/Desktop/hdfs/userp-AUS’

overwrite into table post1 partition (country=’AUS’)



* load data local inpath ‘/home/cloudera/Desktop/hdfs/userp-US’

overwrite into table post1 partition (country=’US’)



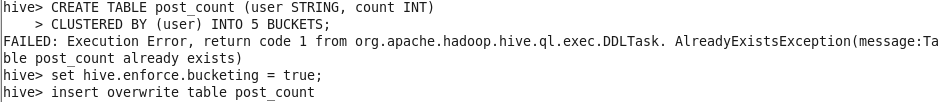
* **BUCKETING**
* CREATE TABLE post\_count (user STRING, count INT)

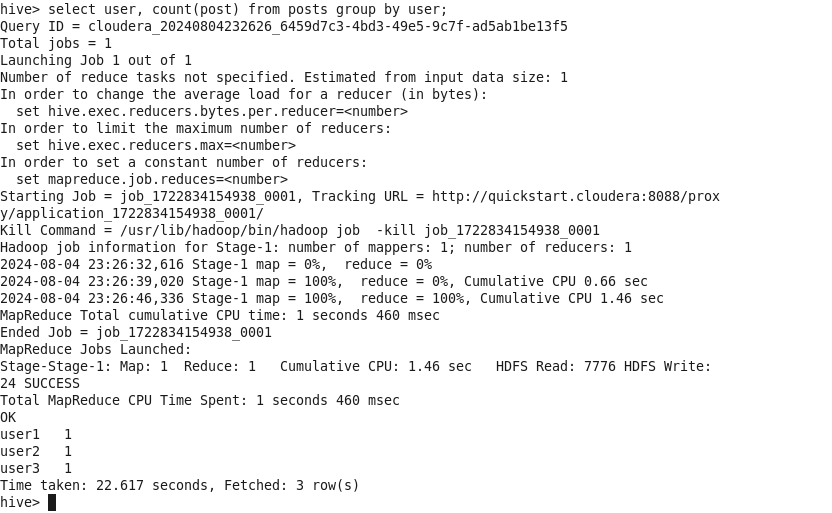
CLUSTERED BY (user) INTO 5 BUCKETS;

set hive.enforce.bucketing=true;

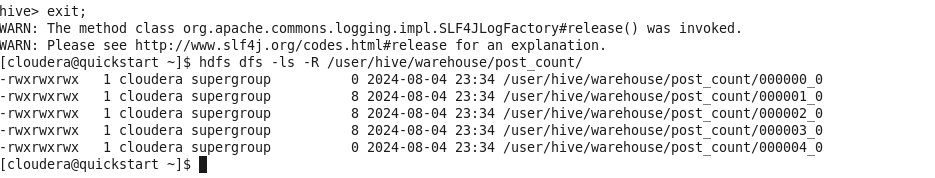
insert overwrute table post\_cpunt

select user, count(post) from posts group by user;





* exit



* **JOINS**

