1. Creation of real estate table

create table realest (street string,city string,zip int,state string,beds int,baths int,sqft int,type string,price int) row format delimited fields terminated by ',';

2. Loading data into table

load data local inpath '/home/rachel/Documents/real\_state.csv' into table realest;

Loading data to table states.realest

OK

Time taken: 0.454 seconds

3.Viewing data

hive> select \* from realest limit 10;

OK

street city NULL state NULL NULL NULL type NULL

3526 HIGH ST SACRAMENTO 95838 CA 2 1 836 Residential 59222

51 OMAHA CT SACRAMENTO 95823 CA 3 1 1167 Residential 68212

2796 BRANCH ST SACRAMENTO 95815 CA 2 1 796 Residential 68880

2805 JANETTE WAY SACRAMENTO 95815 CA 2 1 852 Residential 69307

6001 MCMAHON DR SACRAMENTO 95824 CA 2 1 797 Residential 81900

5828 PEPPERMILL CT SACRAMENTO 95841 CA 3 1 1122 Condo 89921

6048 OGDEN NASH WAY SACRAMENTO 95842 CA 3 2 1104 Residential 90895

2561 19TH AVE SACRAMENTO 95820 CA 3 1 1177 Residential 91002

11150 TRINITY RIVER DR Unit 114 RANCHO CORDOVA 95670 CA 2 2 941 Condo 94905

Time taken: 0.322 seconds, Fetched: 10 row(s)

4. setting bucketing

hive> set hive.exec.dynamic.partition=true;

hive> set hive.exec.dynamic.partition.mode=nonstrict;

hive> set hive.exec.max.dynamic.partitions=20000;

hive> set hive.exec.max.dynamic.partitions.pernode=20000;

hive> set hive.enforce.bucketing=true;

5.structure of table

hive> desc realest;

OK

street string

city string

zip int

state string

beds int

baths int

sqft int

type string

price int

Time taken: 0.093 seconds, Fetched: 9 row(s)

6. Creation of bucket table

hive> create table buest(street string,zip int,state string,beds int,baths int,sqft int,type string,price int) partitioned by (city string) clustered by (street) into 4 buckets row format delimited fields terminated by ',';

OK

Time taken: 0.225 seconds

7. Inserting value into bucketing table

hive> insert into table buest partition(city) select street,zip,state,beds,baths,sqft,type,price,city from realest;