#bucketing in hive.

1.createnormal table for storing data:create table products\_no\_buckets(

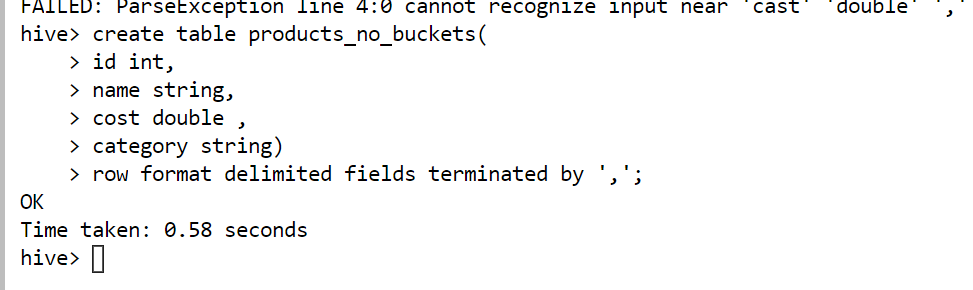
id int,

name string,

cost double ,

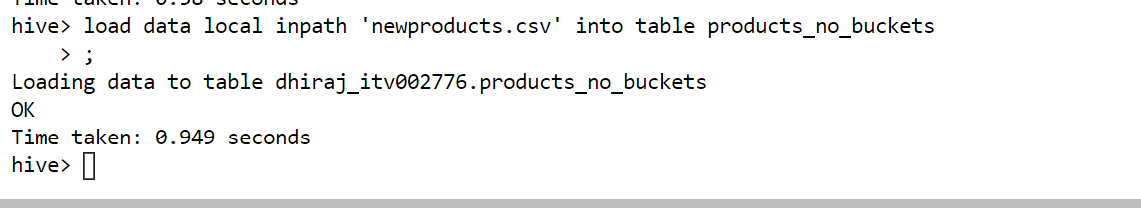
category string)

row format delimited fields terminated by ',';

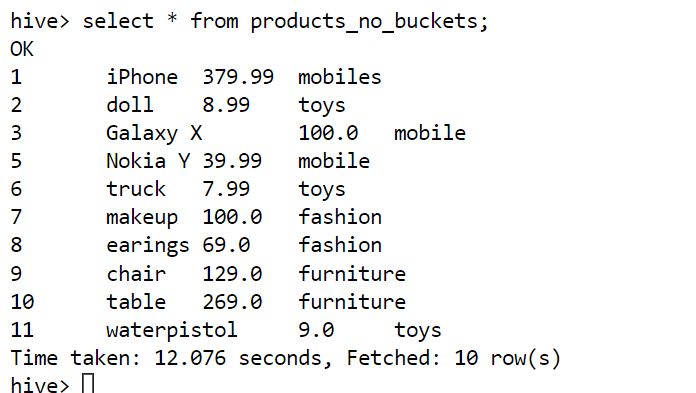


2..load data from local file in normal table

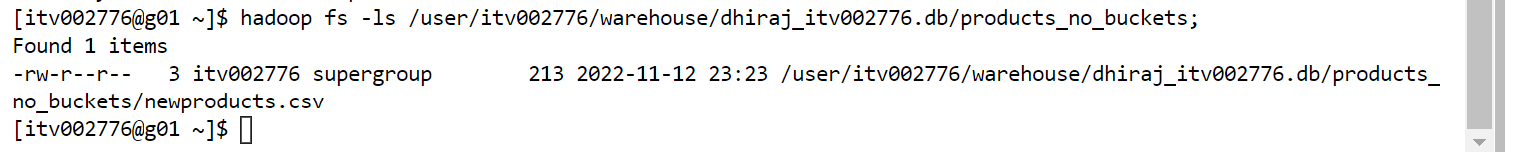
load data local inpath 'newproducts.csv' into table products\_no\_buckets



Records after inserting:



See path in hdfs:



3.set bucketing to true:set hive.enforce.bucketing=true;

4.create bucketed table on column:

create table products\_w\_buckets(

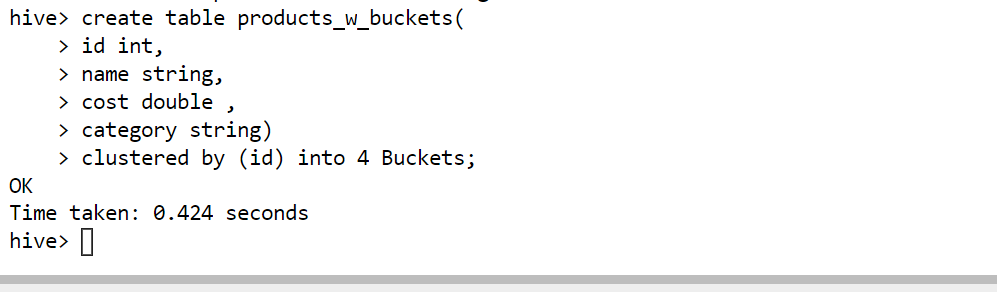
id int,

name string,

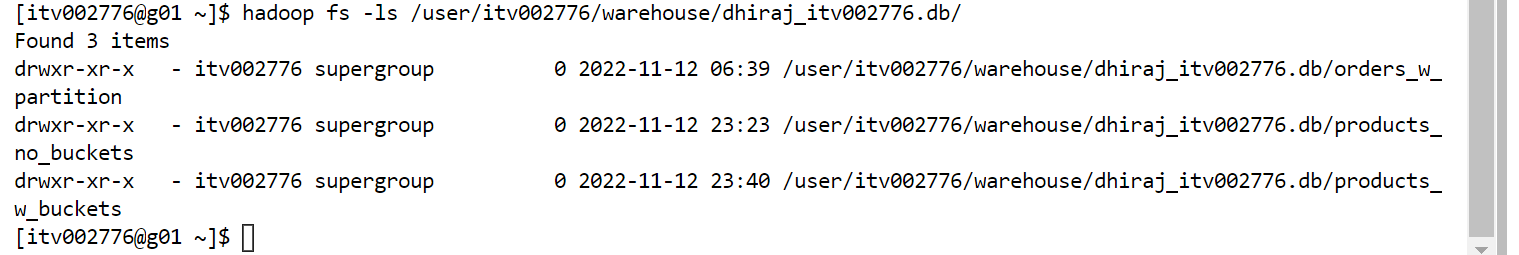
cost double ,

category string)

clustered by (id) into 4 Buckets;

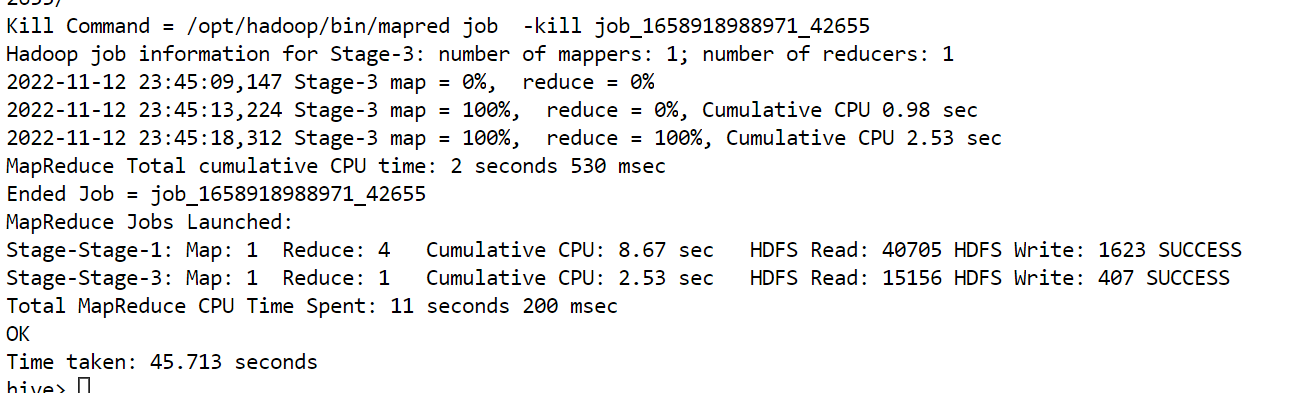


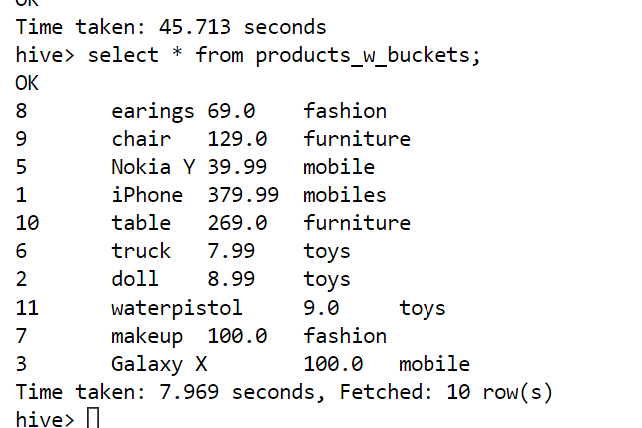
See table in hadoop:



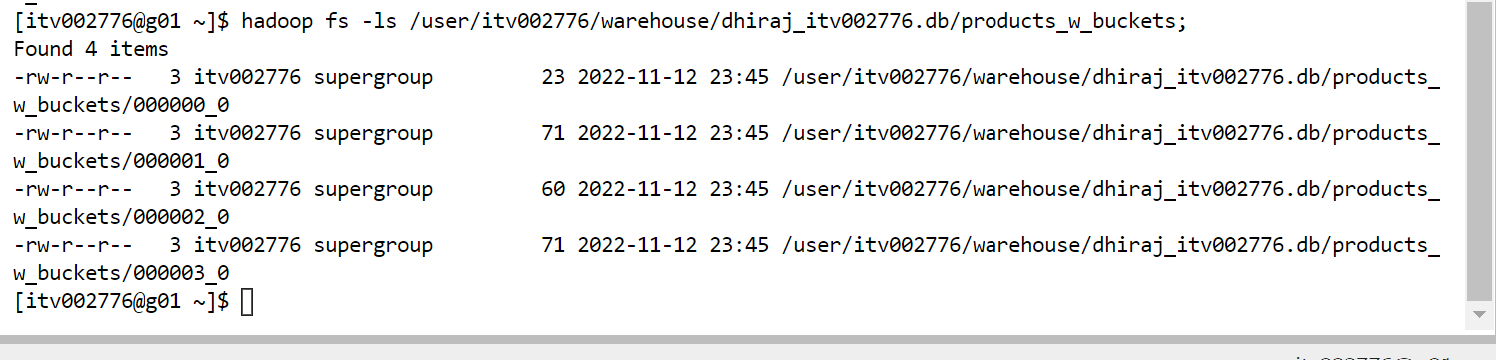
5.load data from non-bucketed to bucketed table:

from products\_no\_buckets insert into products\_w\_buckets select id,name,cost,category;





See file structure in hadoop:



6.To see content bucket wise :

select \* from products\_w\_buckets tablesample(bucket 1 out of 4);

