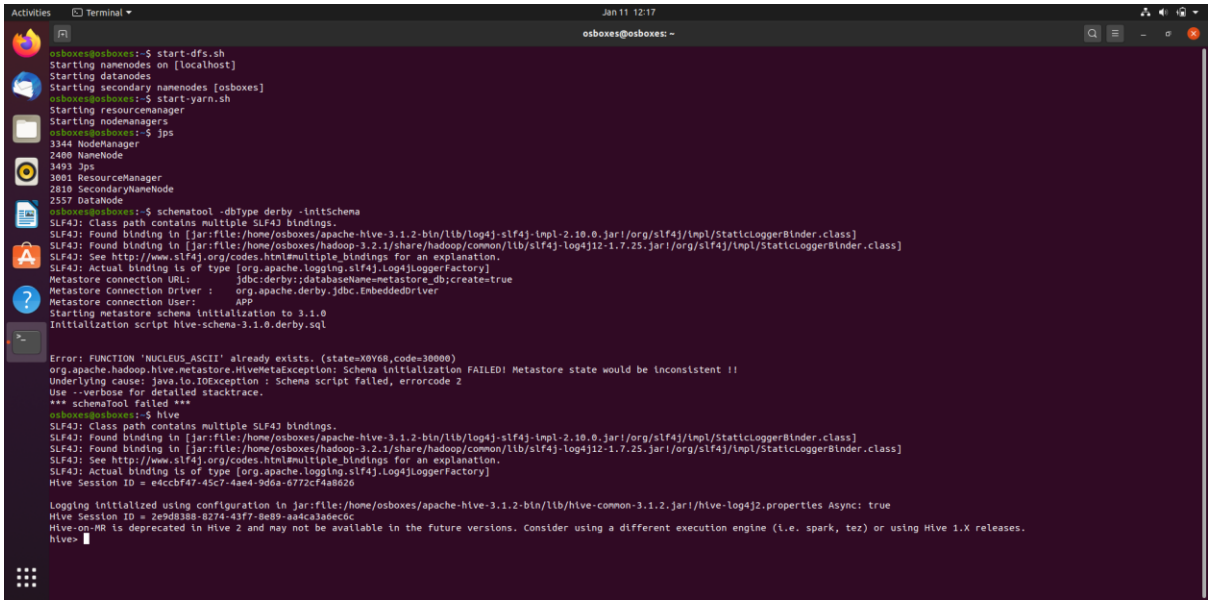
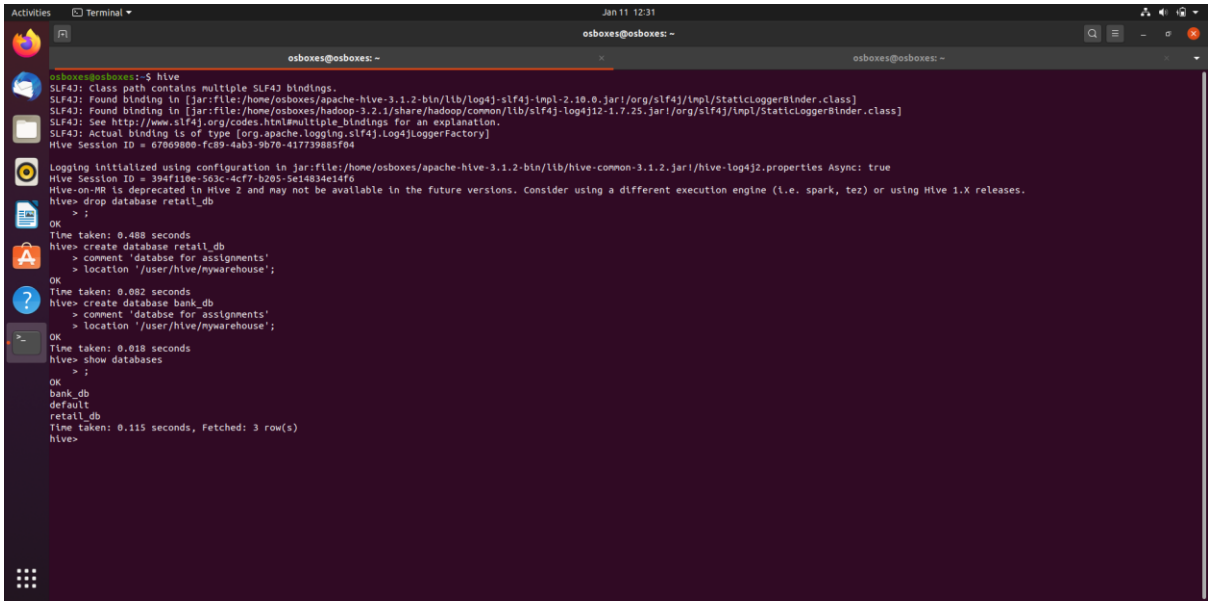
1.Install Hive on the system using the distribution provided.
Initialize schema:-
schematool -dbType derby -initSchema
>hive

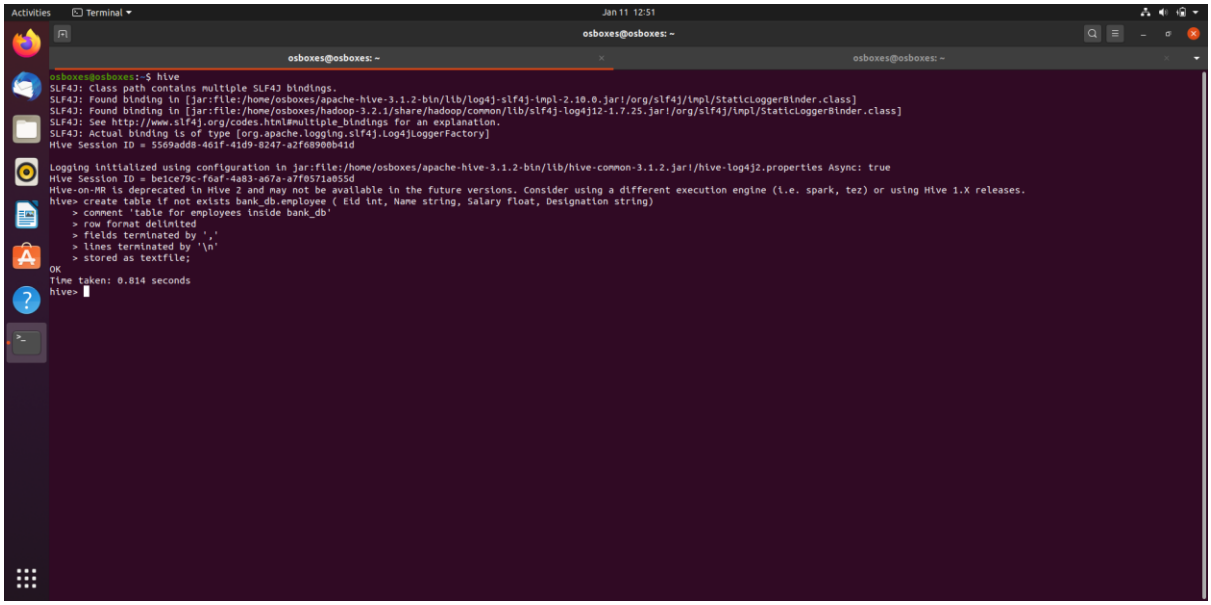


2. Create two databases with the names ‘retail_db’ and ‘bank_db’ having comments and located at /user/hive/mywarehouse/.
Create database retail_db
Create database bnk_db



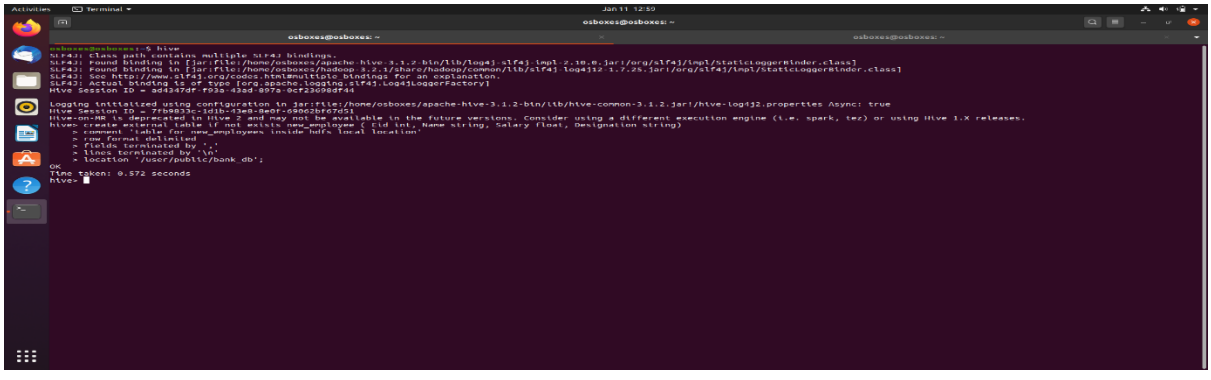
3. Create a hive managed table ‘employee’ in ‘bank_db’ database with the following structure:
Create table if not exists bank_db.employee (eid int, name string, salary float, designation string)

Comment 'table for employees inside bank_db'
Row format delimited
Fields terminated by ','
Lines terminated by '\n'
Stored as textfile;



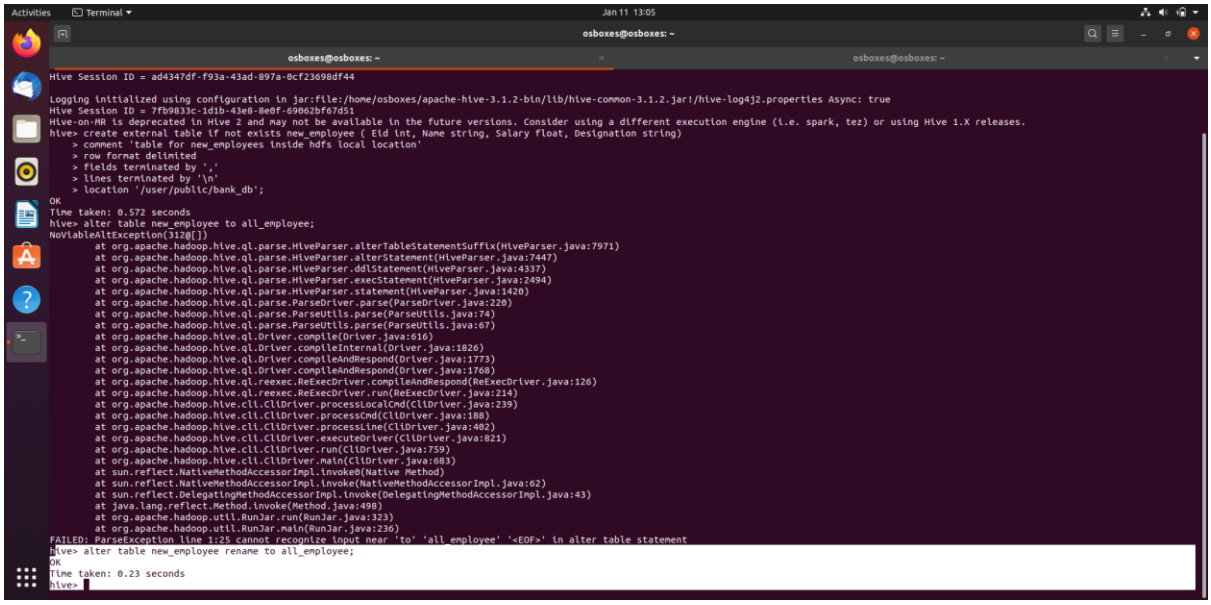
4. Create external table 'new_employee' with the same structure at HDFS location /public/bank_db

Create external table if not exists new_employee (eid int, name string, salary float, designmation string)
Comment ' table '
Row format delimited
Fields terminated by ','
Lines terminated by '\n'
Location '/user/public/bank_db'

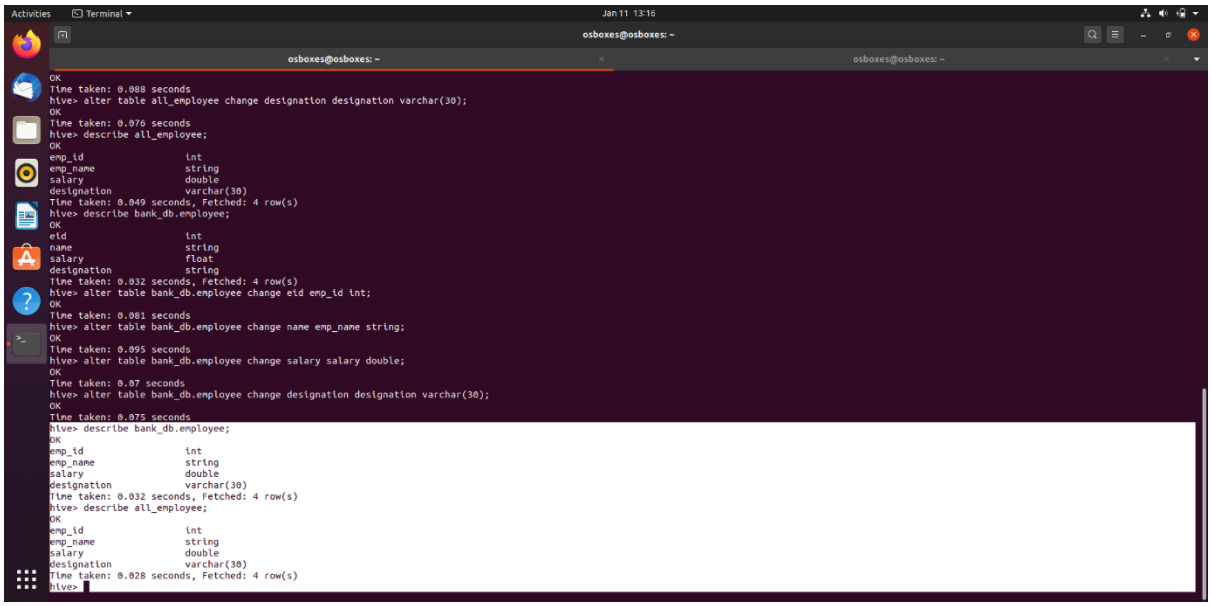


5. Rename table 'new_employee' to 'all_employee'.

Alter table new_employee rename to all_employee;

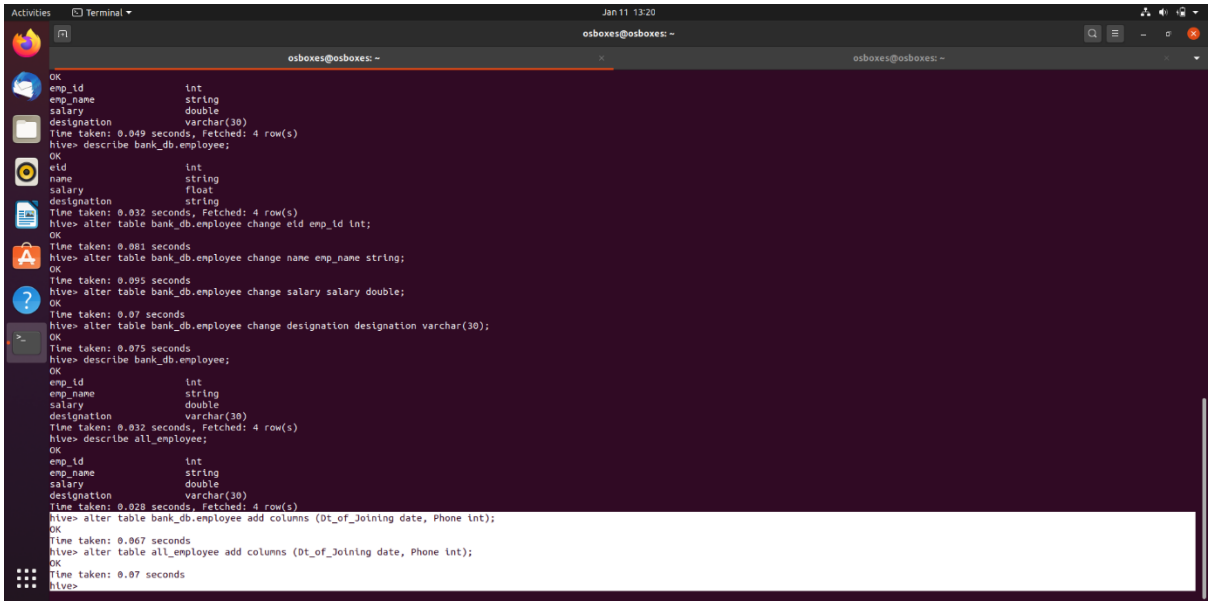


6. Apply the following changes to ‘employee’ and ‘new_employee’ tables:
Describe bank_db.employee;
Alter table all_employee change designation designation varchar(30);
Alter table bank_db.employee change name emp_name string;
Alter table bank_db.employee change salary salary double;
Alter table bank_db.employee change designation designation varchar(30);



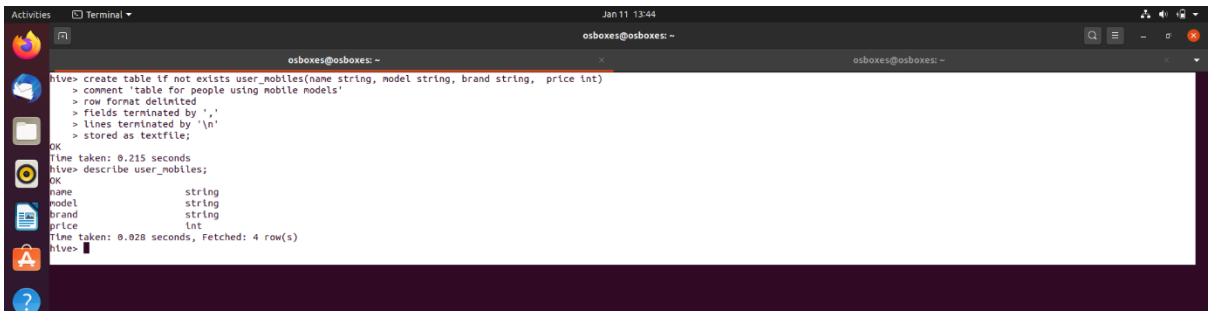
7. Add two column to the above table:
Column Name: Dt_of_Joining

Data Type: DATE
Column Name: Phone
Data Type: INT
Alter table bank_db.employee add columns (dt_of_joining date, phone int);
Alter table all_employee add cloumns (dt_of_joining date, phone int);

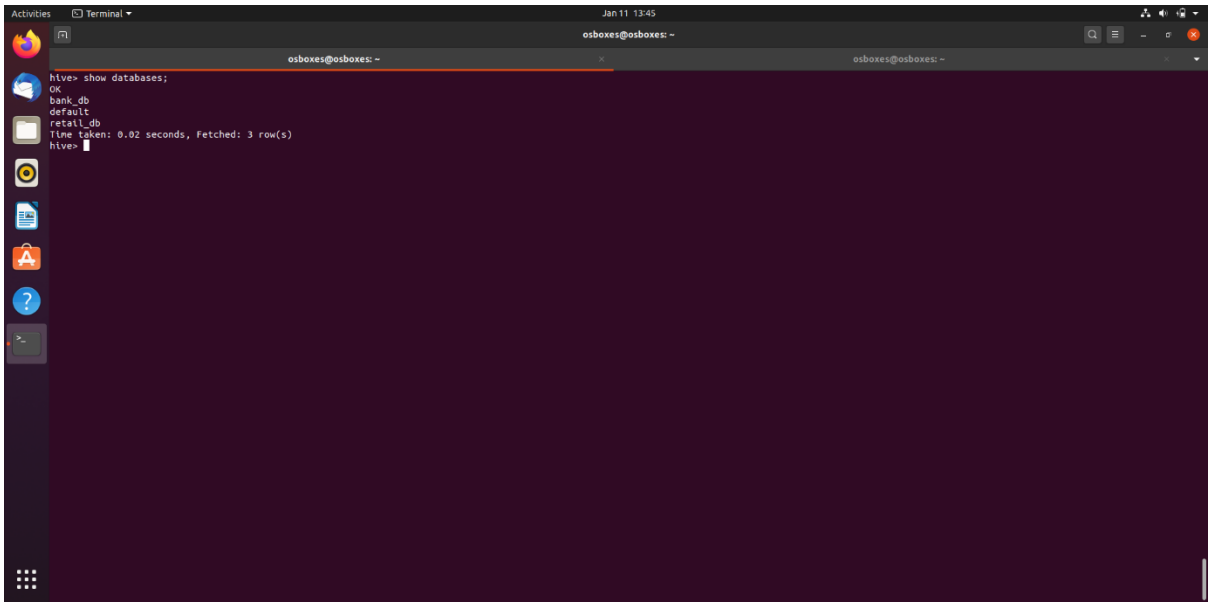


9. Create table in hive that can hold the following records. Choose appropriate field names and types.
Ajay, Lumia 1020, Nokia, 10000
Shiva, iphone6, Apple, 34000
Srejeeth, Galaxy 4, Samsung, 20000

Create table if not exists (name string. Model string, brand string, price int)
Comment 'table'
Row format delimited
Fields terminated by ','
Lines terminated by '\n'
Stored as textfile;

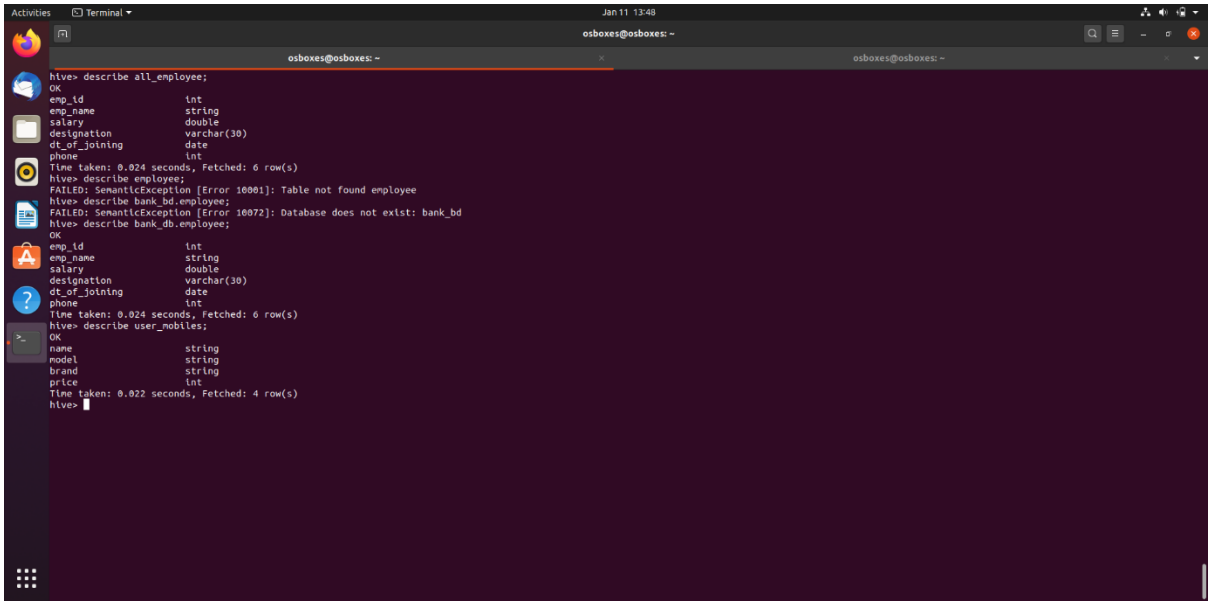


10. Apply 'show' DDL command for the databases and tables created.
Show databases;



11. Apply ‘describe’ DDL command for the databases and tables created.

Describe all_employees;



12. Create all 6 tables listed in figure in hive as managed tables, delimiter is "|", in database 'retail_db'.

```
hive> create table if not exists categories(category_id int, category_departmen
t_id int, category_name varchar(45))
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
OK
Time taken: 0.22 seconds
hive> create table if not exists products1(products_id int, product_category_id
 int, product_name varchar(45),product_description varchar(255), product_price
float, product_image varchar(255))
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
OK
Time taken: 0.188 seconds
hive> create table if not exists order_items(order_item_id int, order_item_orde
r_id int, order_item_product_id int ,order_item_quantity int, order_item_subtot
al float, order_item_product_price float)
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
OK
Time taken: 0.158 seconds
hive> create table if not exists orders(order_id int, order_date datetime,order
_customer_id int, order_status varchar(45))
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
```

```
hive> create table if not exists departments1(department_id int, department_nam
e varchar(45))
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
OK
```

```
hive> create table if not exists orders(order_id int, order_date datetime,order
_customer_id int, order_status varchar(45))
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
FAILED: SemanticException [Error 10099]: DATETIME type isn't supported yet. Ple
ase use DATE or TIMESTAMP instead
hive> create table if not exists orders(order_id int, order_date date,order_cus
tomer_id int, order_status varchar(45))
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
OK
Time taken: 0.116 seconds
hive> create table if not exists customers(customer_id int, customer_fname varc
har(45), customer_lname varchar(45), customer_email varchar(45), customer_passw
ord varchar(45), customer_street varchar(45), customer_city varchar(45), custom
er_state varchar(45), customer_zipcode varchar(45))
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
OK
Time taken: 0.124 seconds
```

```
hive> describe order_items;
OK
order_item_id           int
order_item_order_id     int
order_item_product_id   int
order_item_quantity     int
order_item_subtotal     float
order_item_product_price        float
Time taken: 0.107 seconds, Fetched: 6 row(s)
hive> describe orders;
OK
order_id                int
order_date              date
order_customer_id       int
order_status            varchar(45)
Time taken: 0.108 seconds, Fetched: 4 row(s)
hive> describe customers;
OK
customer_id             int
customer_fname          varchar(45)
customer_lname          varchar(45)
customer_email          varchar(45)
customer_password       varchar(45)
customer_street         varchar(45)
customer_city           varchar(45)
customer_state          varchar(45)
customer_zipcode        varchar(45)
Time taken: 0.109 seconds, Fetched: 9 row(s)
hive>
```

```
hive> describe departments1;
OK
department_id           int
department_name         varchar(45)
Time taken: 0.124 seconds, Fetched: 2 row(s)
hive> describe categories;
OK
category_id             int
category_department_id  int
category_name           varchar(45)
Time taken: 0.107 seconds, Fetched: 3 row(s)
hive> describe products;
FAILED: SemanticException [Error 10001]: Table not found products
hive> describe products1;
OK
products_id             int
product_category_id     int
product_name            varchar(45)
product_description     varchar(255)
product_price           float
product_image           varchar(255)
Time taken: 0.071 seconds, Fetched: 6 row(s)
```

13. Drop tables ‘order_items’ and ‘orders’.

```
hive> Drop table order_items;
OK
Time taken: 2.329 seconds
hive> Drop table orders;
OK
Time taken: 0.309 seconds
hive>
```

14. Create the above two tables again as external tables at a location different from rest of the tables.

```
hive> create external table if not exists customers(customer_id int, customer_f
name varchar(45), customer_lname varchar(45), customer_email varchar(45), custo
mer_password varchar(45), customer_street varchar(45), customer_city varchar(45
), customer_state varchar(45), customer_zipcode varchar(45))
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
OK
Time taken: 0.071 seconds
hive> create external table if not exists orders(order_id int, order_date date,
order_customer_id int, order_status varchar(45))
    > row format delimited
    > fields terminated by '|'
    > lines terminated by '\n';
OK
```

15. Drop database ‘bank_db’ along with all the tables in it in one command.

Drop database if exixts bank_db cascade;