

#### 1. CREATE DATABASE

create database if not exists cab\_booking ;

#### 2. CHECK IF CREATED

use cab\_booking ;

#### 3. CREATE TABLE FOR clickstream\_data\_new

create table if not exists clickstream\_data\_new

(customer\_id int ,app\_version string,

os\_version string,lat string ,lon string ,page\_id string,

button\_id string , is\_button\_click varchar(3) ,is\_page\_view varchar(3) ,is\_scroll\_up

varchar(3) ,is\_scroll\_down varchar(3) )

row format delimited fields terminated by "," ;

#### 4. CHECK TABLE SCHEMA

describe clickstream\_data\_new;

#### 5.CREATE TABLE booking\_data

create table if not exists booking\_data

(booking\_id string ,customer\_id int ,driver\_id int , customer\_app\_version string,

customer\_phone\_os\_version string , pickup\_lat double , pickup\_lon double, drop\_lat double,

drop\_lon double, pickup\_timestamp timestamp , drop\_timestamp timestamp ,trip\_fare int,

tip\_amount int, currency\_code string ,cab\_color string, cab\_registration\_no string ,

customer\_rating\_by\_driver int, rating\_by\_customer int ,passenger\_count int )

row format delimited fields terminated by "," ;

#### 6. CHECK TABLE SCHEMA

describe booking\_data;

## 7. CREATE TABLE datewise\_data

create table if not exists datewise\_data

(date string , count int)

row format delimited fields terminated by "," ;

## 8.CHECK TABLE SCHEMA

describe datewise\_data;

Command to load the data into Hive tables

1.

load data inpath

'/user/root/clickstream\_data\_flatten\_new/part-00000-9d710548-36db-4474-a44a-cd09a42fce3e-c000.csv' into table clickstream\_data\_new;

select count(\*) from clickstream\_data\_new; //2984

2.

load data inpath '/user/root/bookings\_data\_new/part-m-00000' into table booking\_data ;

select count(\*) from booking\_data ;//1000

3.

load data inpath

'/user/root/datewise\_aggreation\_new/part-00000-9d97e6f7-e258-41eb-ad5c-1e5cc498064e-c000.csv' into table datewise\_data ;

select count(\*) from datewise\_data ; //290

CHECKING NO OF RECORDS

```
select count(*) from clickstream_data_new;
```

```
select count(*) from booking_data ;
```

```
select count(*) from datewise_data ;
```