

## Create Hive-Managed Tables

### <Command to create the Hive tables>

#### managed table for clickstream\_data

```
create database if not exists cab_booking ;
use cab_booking ;
create table if not exists clickstream_data
(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 "," ;
```

#### managed table for bookings 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 "," ;
```

#### managed table for aggregated data in Task 3

```
create table if not exists datewise_data
(date string , count int)
row format delimited fields terminated by "," ;
```

### <Command to load the data into Hive tables>

1. load data inpath '/user/root/clickstream\_data\_flatten/part-00000-2f2214f7-80c2-4286-90b9-b53bb9e4d91c-c000.csv' into table clickstream\_data ;
2. load data inpath '/user/root/booking\_data\_csv/part-00000-f8248935-24c7-49d5-92c6-f0be2cfe4855-c000.csv' into booking\_data ;
3. load data inpath '/user/root/datewise\_aggregation/part-00000-fd22c33b-ef88-49a5-a2cc-721bd83e426c-c000.csv' into datewise\_data ;