1) Creating a Database Hive_Project1 and Using It

```
hive (default)> create database Hive_Project1;
OK
Time taken: 0.341 seconds
hive (default)> use Hive_Project1;
OK
Time taken: 0.028 seconds
hive (Hive_Project1)>
```

2) Creating Table Booking in The Dataset and showing columns in the Booking Table

```
hive (Hive_Project1)> create table Booking(Booking_id int,No_of_Seats int,Seat_T
ype string);
OK
Time taken: 0.233 seconds
hive (Hive_Project1)> desc Booking;
OK
booking_id              int
no_of_seats             int
seat_type               string
Time taken: 0.133 seconds, Fetched: 3 row(s)
hive (Hive_Project1)>
```

3) Creating Table Customer_Details in the Database and showing columns in the Customer_Details Table

```
hive (Hive_Project1)> create table customer_details(booking_id int,Customer_Name string,Age int,Contact_No int);
OK
Time taken: 0.106 seconds
hive (Hive_Project1)> desc customer_details;
OK
booking_id              int
customer_name           string
age                     int
contact_no              int
Time taken: 0.07 seconds, Fetched: 4 row(s)
hive (Hive_Project1)>
```

4) Creating Table Journey_Details in the Database and showing columns in the Journey_Details Table

```
hive (Hive_Project1)> create table Journey_Details(booking_id int,Driver_Name string,Starting_Point string,Destination string,Fare_Amount int);
OK
Time taken: 0.237 seconds
hive (Hive_Project1)> desc Journey_Details;
OK
booking_id              int
driver_name             string
starting_point          string
destination             string
fare_amount             int
Time taken: 0.153 seconds, Fetched: 5 row(s)
hive (Hive_Project1)>
```

5) Booking Table After Inserting Values

```
hive (Hive_Project1)> select * from Booking;
OK
1       2       Sleeper
2       3       Sleeper
3       5       Seater
4       2       Sleeper
5       9       Seater
6       1       Seater
7       3       Sleeper
8       6       Seater
9       2       Sleeper
10      5       Sleeper
11      1       Seater
12      6       Sleeper
Time taken: 0.097 seconds, Fetched: 12 row(s)
hive (Hive_Project1)>
```

6) Customer_Details Table After Inserting Values

```
hive (Hive_Project1)> select * from customer_details;
OK
1       Sumit   23      987654321
2       Shubham 21      986754321
3       Shiva   26      877643211
4       Sakthes 22      878777271
5       Rohan   19      878876098
6       Rihan   21      234543678
7       Rajesh  26      234542588
8       Ramya   25      234132567
9       Harsh   29      985432567
10      Bharat  28      985776655
11      Bhavna  25      985754431
12      Bhumi   21      985750551
Time taken: 0.074 seconds, Fetched: 12 row(s)
```

7) Journey_Details Table After Inserting Values

```
hive (Hive_Project1)> select * from Journey_Details;
OK
1       Raghu   Mumbai  Pune    1000
2       Gokul   Chennai Coimbatore      600
3       Vipul   Palakkad        Banglore        780
4       Divakar Tirunelveli     Palani  460
5       Deepak  Tuticorin       Kanyakumari     800
6       Amir    Jaipur  Delhi   1600
7       Raghu   Mumbai  Pune    1000
8       Gokul   Chennai Coimbatore      600
9       Vipul   Palakkad        Banglore        780
10      Divakar Tirunelveli     Palani  460
11      Deepak  Tuticorin       Kanyakumari     800
12      Amir    Jaipur  Delhi   1600
Time taken: 0.051 seconds, Fetched: 12 row(s)
hive (Hive_Project1)>
```

8) Creating a View V1 to show Customer Name & Contact Number of Customer

```
hive (Hive_Project1)> create view V1 AS select customer_details.customer_name,customer_details.contact_no from customer_details;
OK
Time taken: 0.069 seconds
hive (Hive_Project1)> select * from v1;
OK
Sumit   987654321
Shubham 986754321
Shiva   877643211
Sakthes 878777271
Rohan   878876098
Rihan   234543678
Rajesh  234542588
Ramya   234132567
Harsh   985432567
Bharat  985776655
Bhavna  985754431
Bhumi   985750551
Time taken: 0.075 seconds, Fetched: 12 row(s)
hive (Hive_Project1)>
```

9) Using Order by to make the customer_details table in descending order by booking_id

```
hive (Hive_Project1)> select * from customer_details order by booking_id  desc;
Query ID = cloudera_20231103054343_841f723a-0f73-4115-b8ff-49fc8a07ed01
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1699006550186_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1699006550186_0006/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1699006550186_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-11-03 05:43:50,076 Stage-1 map = 0%,  reduce = 0%
2023-11-03 05:43:58,707 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.18 sec
2023-11-03 05:44:09,636 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.92 sec
MapReduce Total cumulative CPU time: 2 seconds 920 msec
Ended Job = job_1699006550186_0006
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.92 sec   HDFS Read: 6823 HDFS Write: 262 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 920 msec
OK
12      Bhumi   21      985750551
11      Bhavna  25      985754431
10      Bharat  28      985776655
9       Harsh   29      985432567
8       Ramya   25      234132567
7       Rajesh  26      234542588
6       Rihan   21      234543678
5       Rohan   19      878876098
4       Sakthes 22      878777271
3       Shiva   26      877643211
2       Shubham 21      986754321
1       Sumit   23      987654321
Time taken: 31.645 seconds, Fetched: 12 row(s)
hive (Hive_Project1)>
```

10) Performed Right Join to show customer_name from Customer Details table and seat_type from booking table

```
hive (Hive_Project1)> select b.customer_name,c.seat_type from customer_details b right join booking c on b.booking_id=c.booking_id;
Query ID = cloudera_20231103061616_1af618ad-ee07-4400-bc33-5e1a4fe1a5e4
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20231103061616_1af618ad-ee07-4400-bc33-5e1a4fe1a5e4.log
2023-11-03 06:17:06     Starting to launch local task to process map join;      maximum memory = 1013645312
2023-11-03 06:17:07     Dump the side-table for tag: 0 with group count: 12 into file: file:/tmp/cloudera/7e5529d0-892f-45b0-9117-514
3_06-16-58_949_1344602178890336626-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable
2023-11-03 06:17:07     Uploaded 1 File to: file:/tmp/cloudera/7e5529d0-892f-45b0-9117-51473ac828db/hive_2023-11-03_06-16-58_949_1344
-10003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable (570 bytes)
2023-11-03 06:17:07     End of local task; Time Taken: 1.593 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1699006550186_0007, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1699006550186_0007/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1699006550186_0007
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2023-11-03 06:17:20,513 Stage-3 map = 0%,  reduce = 0%
2023-11-03 06:17:30,426 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 1.14 sec
MapReduce Total cumulative CPU time: 1 seconds 140 msec
Ended Job = job_1699006550186_0007
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 1.14 sec   HDFS Read: 5984 HDFS Write: 170 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 140 msec
OK
Sumit   Sleeper
Shubham Sleeper
Shiva   Seater
Sakthes Sleeper
Rohan   Seater
Rihan   Seater
Rajesh  Sleeper
Ramya   Seater
Harsh   Sleeper
Bharat  Sleeper
Bhavna  Seater
Bhumi   Sleeper
Time taken: 32.593 seconds, Fetched: 12 row(s)
hive (Hive_Project1)>
```

11) Performed Inner Join to show Booking_id , Customer_Name, Driver_name and Fare Amount

```
hive (hive_project1)> select c.booking_id,c.customer_name,j.driver_name,j.fare_a
mount from customer_details c join journey_details j  on c.booking_id=j.booking_
id;
Query ID = cloudera_20231104011111_7c6018e9-1219-49fe-a26a-27db7c8741a6
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20231104011111_7c6018e9-1219-49fe-a26a-
27db7c8741a6.log
2023-11-04 01:11:15     Starting to launch local task to process map join;     m
aximum memory = 1013645312
2023-11-04 01:11:16     Dump the side-table for tag: 0 with group count: 12 into
 file: file:/tmp/cloudera/f06d1429-a867-4eb9-8405-929953b71fcb/hive_2023-11-04_0
1-11-07_580_1627817847420958081-1/-local-10003/HashTable-Stage-3/MapJoin-mapfile
00--.hashtable
2023-11-04 01:11:16     Uploaded 1 File to: file:/tmp/cloudera/f06d1429-a867-4eb
9-8405-929953b71fcb/hive_2023-11-04_01-11-07_580_1627817847420958081-1/-local-10
003/HashTable-Stage-3/MapJoin-mapfile00--.hashtable (570 bytes)
2023-11-04 01:11:16     End of local task; Time Taken: 1.842 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1699083302397_0002, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1699083302397_0002/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1699083302397_0002
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2023-11-04 01:11:31,081 Stage-3 map = 0%,  reduce = 0%
2023-11-04 01:11:39,870 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 1.34 se
c
MapReduce Total cumulative CPU time: 1 seconds 340 msec
Ended Job = job_1699083302397_0002
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 1.34 sec   HDFS Read: 6835 HDFS Write: 2
34 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 340 msec
OK
1       Sumit   Raghu   1000
2       Shubham Gokul   600
3       Shiva   Vipul   780
4       Sakthes Divakar 460
5       Rohan   Deepak  800
6       Rihan   Amir    1600
7       Rajesh  Raghu   1000
8       Ramya   Gokul   600
9       Harsh   Vipul   780
10      Bharat  Divakar 460
11      Bhavna  Deepak  800
12      Bhumi   Amir    1600
Time taken: 33.421 seconds, Fetched: 12 row(s)
hive (hive_project1)>
```

12) To show Average Fare Amount of Customers

```
hive (Hive_Project1)> select avg(fare_amount) from journey_details;
Query ID = cloudera_20231104040202_c95eb3d8-3ca3-49b9-a851-32c8ef875ade
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1699083302397_0010, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1699083302397_0010/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1699083302397_0010
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-11-04 04:03:11,075 Stage-1 map = 0%,  reduce = 0%
2023-11-04 04:03:18,897 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.25 sec
2023-11-04 04:03:29,030 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.86 sec
MapReduce Total cumulative CPU time: 2 seconds 860 msec
Ended Job = job_1699083302397_0010
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.86 sec   HDFS Read: 7899 HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 860 msec
OK
873.3333333333334
Time taken: 33.876 seconds, Fetched: 1 row(s)
hive (Hive_Project1)>
```