```
[cloudera@quickstart ~]$ hive
hive> create database flight_dbs;
hive> show databases;
```

a) Creating, Dropping, and altering Database tables.

a2)creating table

```
hive> use flight_dbs;
hive> create table flight_info(
    > flight_no INT,
    > day_of_week INT,
    > dep_time INT,
    > origin STRING,
    > dest STRING
    > )
    > row format delimited
    > fields terminated by ','
    > stored as textfile;

hive> show table
hive> desc flight_info;
```

a3)altering table by changing name or add column

```
hive> alter table flight_info ADD columns(distance INT) ;
hive> desc flight_info;
```

a4)dropping table

```
hive> drop table flight_info;
hive> show tables;

[cloudera@quickstart ~]$ hdfs dfs -ls /
[cloudera@quickstart ~]$ hdfs dfs -mkdir -p /flight_data
[cloudera@quickstart ~]$ hdfs dfs -ls /

[cloudera@quickstart ~]$ hdfs dfs -put /home/cloudera/Desktop/flight_info.csv /flight_data
[cloudera@quickstart ~]$ hdfs dfs -ls /flight_data
```

b) Creating an external Hive table.(flight_ext)

```
[cloudera@quickstart ~]$ hive
hive> create external table flight_ext(
    > year INT,
    >  month INT,
    >  day INT,
    >  day_of_week INT,
    >  dep_time INT,
    >  crs_dep_time INT,
    >  arr_time INT,
    >  crs_arr_time INT,
    >  unique_carrier STRING,
    >  flight_num INT,
    >  tail_num STRING,
    >  actual_elapsed_time INT,
```

```
    >    crs_elapsed_time INT,
    >    air_time INT,
    >    arr_delay INT,
    >    dep_delay INT,
    >    origin STRING,
    >    dest STRING,
    >    distance INT,
    >    taxi_in INT,
    >    taxi_out INT,
    >    cancelled INT,
    >    cancellation_code STRING,
    >    diverted INT,
    >    carrier_delay STRING,
    >    weather_delay STRING,
    >    nas_delay STRING,
    >    late_aircraft_delay STRING,
    >    security_delay STRING
    > )
    > row format delimited
    > fields terminated by ','
    > stored as textfile
    > location '/flight_data';

hive> desc flight_ext;
```

c) Load table with data, insert new values and field in the table, Join tables with Hive

c1)create internal table(flight_int)

```
hive> create table flight_int AS
    > select
    > year,month,day,flight_num,dep_delay,origin,dest
    > FROM flight_ext;

hive> select * from flight_int LIMIT 10;
```

c3)insert new values in table

```
hive> insert into flight_int values(2010,2,21,505,6,'ENG','IND');
hive> select * from flight_int WHERE flight_num = 505;
```

c4) join tables

```
hive> create table f_price(flight_num int,price float);
hive> insert into f_price values (505,5000.0);
hive> select * from f_price;

hive> select a.flight_num,a.day,a.month,a.year,b.price
    > from flight_int a JOIN f_price b ON (a.flight_num = b.flight_num);
```

d)Create index on Flight Information Table

```
hive> create index flight_index on table flight_int(flight_num)
    > AS 'COMPACT'
```

```
    > with DEFERRED REBUILD;
hive> show index on flight_int;
```

e1)find avg departure delay

```
hive> select AVG(dep_delay) FROM flight_int;
```

e2)Find the average departure delay per day in 2008.

```
hive> select day,month,year,AVG(dep_delay) AS avg_dep_del
    > FROM flight_int
    > WHERE dep_delay IS NOT NULL
    > GROUP BY day,month,year;
```

e3) find monthly departure delay

```
hive> select month,AVG(dep_delay)
    > FROM flight_int
    > WHERE dep_delay IS NOT NULL
    > GROUP BY month;
```