

1. Create table for airports

create table airports (ID int,airport_name string,city string,country string,faa_code string,icao string,latitude double,longitude double,altitude double,timezone double,dst string,tz string) row format delimited fields terminated by ',';

```
Time taken: 0.796 seconds
hive> create table airports (ID int,airport_name string,city string,country string,faa_code string,icao string,latitude double,longitude double,altitude double,timezone double,dst string,tz string) row format delimited fields terminated by ',';
OK
Time taken: 1.32 seconds
hive> describe airports;
OK
id                                int
airport_name                      string
city                             string
country                          string
faa_code                         string
icao                              string
latitude                         double
longitude                        double
altitude                         double
timezone                         double
dst                              string
tz                              string
Time taken: 0.67 seconds, Fetched: 12 row(s)
hive>
```

2. Load airport dataset to table airports

load data local inpath 'Desktop/airports.txt' into table airports;

```
hive> load data local inpath 'Desktop/airports.txt' into table airports;
Copying data from file:/home/edureka/Desktop/airports.txt
Copying file: file:/home/edureka/Desktop/airports.txt
Loading data to table deepika.airports
Table deepika.airports stats: [numFiles=1, numRows=0, totalSize=739515, rawDataSize=0]
OK
Time taken: 1.573 seconds
hive>
```

If needed, we can check whether the data loaded correctly using the below command:

Select * from airports limit 5;

```
Time taken: 1.573 seconds
hive> select * from airports limit 5;
OK
1      Goroka      Goroka      Papua New Guinea      GKA      AYGA      -6.081689      1
45.391881      5282.0      10.0      U      Pacific/Port_Moresby
2      Madang      Madang      Papua New Guinea      MAG      AYMD      -5.207083      1
45.7887      20.0      10.0      U      Pacific/Port_Moresby
3      Mount Hagen      Mount Hagen      Papua New Guinea      HGU      AYMH      -
5.826789      144.295861      5388.0      10.0      U      Pacific/Port_Moresby
4      Nadzab      Nadzab      Papua New Guinea      LAE      AYNZ      -6.569828      1
46.726242      239.0      10.0      U      Pacific/Port_Moresby
5      Port Moresby Jacksons Intl      Port Moresby      Papua New Guinea      P
OM      AYPY      -9.443383      147.22005      146.0      10.0      U      Pacific/
Port_Moresby
Time taken: 0.512 seconds, Fetched: 5 row(s)
hive>
```

3. Create table for airlines

create table airlines (airline_ID string,airline_name string, airline_alias string, airline_iata string, airline_icao string,callsign string,territory string, active string) row format delimited fields terminated by ',';

```
hive> create table airlines (airline_ID string,airline_name string, airline_alias string, airline_iata string, airline_icao string,callsign string,territory string, active string) row format delimited fields terminated by ',';
OK
Time taken: 0.145 seconds
hive> describe airlines;
OK
airline_id          string
airline_name        string
airline_alias       string
airline_iata        string
airline_icao         string
callsign            string
territory           string
active              string
Time taken: 0.188 seconds, Fetched: 8 row(s)
hive> █
```

4. Load airlines dataset to table airlines

load data local inpath 'Desktop/airlines.txt' into table airlines;

5. Create table for route:

create table routes (route_iata string, route_airlineid int, source_airport string, source_airportid int, destination_airport string,destination_airportid int, codeshare string,stops int,equip string) row format delimited fields terminated by ',';

6. Load route dataset to table routes:

load data local inpath 'Desktop/route.txt' into table routes;

```
hive> select * from routes limit 5;
OK
2B      410      AER      2965      KZN      2990      0      CR2
2B      410      ASF      2966      KZN      2990      0      CR2
2B      410      ASF      2966      MRV      2962      0      CR2
2B      410      CEK      2968      KZN      2990      0      CR2
2B      410      CEK      2968      OVB      4078      0      CR2
Time taken: 0.489 seconds, Fetched: 5 row(s)
hive> █
```

Problem Statement :

A. Find list of Airports operating in the Country India

I have created a separate table for airports in India.

```
create table indianairports as select airport_name from airports where country='India';
```

```
Diu Airport
Hubli Airport
Shillong Airport
Lokpriya Gopinath Bordoloi International Airport
Dimapur Airport
Tezpur Airport
Dibrugarh Airport
Shimla Airport
Nanded Airport
Kangra Airport
Gorakhpur Airport
Latur Airport
Mysore Airport
Nagaur Airport
New Delhi Train Station
Agra Cantonment Railway Station
Gangtok Helipad
Havelock Island Seaport
Yelahanka AFB
Mormugao-Goa Port
Cochin Port
Helipad of Viraj Group
Time taken: 0.12 seconds, Fetched: 140 row(s)
hive> █
```

B. Find the list of Airlines having zero stops

I have created a separate table for zero stops by inner joining routes table and airlines table

```
create table zerostops as select a.airline_name as airline, a.airline_id as airlineid, b.stops
from routes b inner join airlines a on b.route_airlineid = a.airline_id;
```

```
Select * from zerostops where stops=0;
```

```
hive> 
```

```
hive> 
```

D. Which country (or) territory having highest Airports

Select country, count(*) as cnt from airports group by country order by cnt desc;

```
c
2022-07-24 12:11:49,286 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 3.13
sec
MapReduce Total cumulative CPU time: 3 seconds 130 msec
Ended Job = job_1658643766569_0004
MapReduce Jobs Launched:
Job 0: Map: 1 Reduce: 1 Cumulative CPU: 3.39 sec HDFS Read: 739743 HDFS Write: 7092 SUCCESS
Job 1: Map: 1 Reduce: 1 Cumulative CPU: 3.13 sec HDFS Read: 7458 HDFS Write: 3151 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 520 msec
OK
United States 1697
Canada 435
Germany 321
Australia 263
Russia 249
France 233
China 219
Brazil 213
United Kingdom 209
India 140
Japan 131
Indonesia 114
```

E. Find the list of Active Airlines in United states

select airline_name from airlines where territory='United States' and active='Y';

```
Southjet
Southjet connect
Southjet cargo
Patriot Airways
Vision Airlines (V2)
Yellowtail
Royal Airways
Executive AirShare
T.J. Air
Maryland Air
Apache Air
National Air Cargo
Eastern Atlantic Virtual Airlines
Comfort Express Virtual Charters Albany
Comfort Express Virtual Charters
XAIR USA
Envoy Air
Rainbow Air (RAI)
Rainbow Air Polynesia
Rainbow Air US
Spike Airlines
All America
Time taken: 24.021 seconds, Fetched: 141 row(s)
hive> 
```

