**Project**

**AirLines Analysis**

**Problem Statement:**

1. 1. Find list of Airports operating in the Country India **140**

select airportname from airlines where country ='India';

1. 2. Find the list of Airlines having zero stops **540**

select distinct(a.airlinesname) from airlines a join routes r on (a.airlinesid=r.aid) where r.stops=o;

1. 3. List of Airlines operating with code share **146**

select distinct(a.airlinesname) from airlines a join routes r on (a.airlinesid=r.aid) where r.cshare='Y';

1. 4. Which country (or) territory having highest Airports
2. **United States 1697**

select country,count(\*) as noofap from airports group by country order by noofap desc limit 1;

1. 5. Find the list of Active Airlines in United state **141**

select distinct (airlinesname) from airlines where status='Y' and country="United States";

**Dataset Description:**

In this use case there are 3 data sets. **Final\_airlines, routes.dat, airports\_mod.dat**

**Air Ports data set i.e airports\_mod.dat**

**Air Ports data set i.e airports\_mod.dat**

|  |  |
| --- | --- |
| It contains the following fields Airport ID | Unique OpenFlights identifier for this airport. |
| Name | Name of airport. May or may not contain the City name. |
| City | Main city served by airport. May be spelled differently from Name. |
| Country | Country or territory where airport is located. |
| IATA/FAA | 3-letter FAA code, for airports located in Country "United States of America". |
| 3-letter IATA code, for all other airports. | |

Blank if not assigned.

|  |  |
| --- | --- |
| ICAO | 4-letter ICAO code. |
| Blank if not assigned. | |
| Latitude | Decimal degrees, usually to six significant digits. Negative is South, positive is North. |
| Longitude | Decimal degrees, usually to six significant digits. Negative is West, positive is East. |
| Altitude | In feet. |
| Timezone | Hours offset from UTC. Fractional hours are expressed as decimals, eg. India is 5.5. |
| DST | Daylight savings time. One of E (Europe), A (US/Canada), S (South America), O (Australia), Z (New |

Tz database time Timezone in "tz" (Olson) format, eg. "America/Los\_Angeles". Zone

**Air Lines Data set:**

It contains the following fields:

Airline Unique OpenFlights identifier for this airline. ID

Name Name of the airline.

Alias Alias of the airline. For example, All Nippon Airways is commonly known as "ANA".

IATA 2-letter IATA code, if available.

ICAO 3-letter ICAO code, if available.

Callsign Airline callsign.

Country Country or territory where airline is incorporated.

Active "Y" if the airline is or has until recently been operational, "N" if it is defunct. This field is not reliable: in particular, major airlines that stopped flying long ago, but have not had their IATA code reassigned (eg. Ansett/AN), will incorrectly show as "Y".

**Routes Data set i.e routes.dat**

It contains the following fields

|  |  |
| --- | --- |
| Airline | 2-letter (IATA) or 3-letter (ICAO) code of the airline. |
| Airline ID | Unique OpenFlights identifier for airline (see Airline). |
| Source airport | 3-letter (IATA) or 4-letter (ICAO) code of the source airport. |
| Source airport ID | Unique OpenFlights identifier for source airport (see Airport) |
| Destination airport | 3-letter (IATA) or 4-letter (ICAO) code of the destination airport. |
| Destination airport ID Unique OpenFlights identifier for destination airport (see Airport) | |
| Codeshare | "Y" if this flight is a codeshare (that is, not operated by Airline, but another carrier), |
| empty otherwise. | |
| Stops | Number of stops on this flight ("0" for direct) |

|  |  |
| --- | --- |
| Equipment | 3-letter codes for plane type(s) generally used on this flight, separated by spaces |

create table airports(airportId int,airportName string,city string,country string,iata string,iaco string,latitude float,longitude float,altitude float,timezone double,dst string,tz string) row format delimited fields terminated by ',';

load data local inpath 'hivedata/airports\_mod.dat' overwrite into table airports;

create table airlines(airlinesId int,airlinesName string,airlineAlias string,iata int,icao int,callsign string,country string,status string)row format delimited fields terminated by ',';

load data local inpath 'hivedata/Final\_airlines' overwrite into table airlines;