

Write an application using HBase and HiveQL for flight information system which will include

a. Create Flight Info Hbase Table(with Flight information, schedule, and delay)

hbase shell

create 'flight_info', 'info', 'schedule', 'delay'

describe 'flight_info'

put 'flight_info', 'FL001', 'info:airline', 'IndiGo'

put 'flight_info', 'FL001', 'schedule:departure', '2024-12-01T10:00'

put 'flight_info', 'FL001', 'delay:departure', '15'

put 'flight_info', 'FL002', 'info:airline', 'SpiceJet'

put 'flight_info', 'FL002', 'schedule:departure', '2025-01-01T9:00'

put 'flight_info', 'FL002', 'delay:departure', '45'

b. Demonstrate Creating, Dropping, and altering Database tables in Hbase

create 'sample_table', 'data'

alter 'sample_table', NAME => 'extra' (add another column called "extra")

describe 'sample_table'

list

disable 'sample_table'

drop 'sample_table'

list

c. Creating an external Hive table to connect to the HBase for Flight Information Table

hive

CREATE EXTERNAL TABLE flight_hive (

key STRING,

airline STRING,

departure STRING,

departure_delay INT

)

```
STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'  
WITH SERDEPROPERTIES (  
    "hbase.columns.mapping" = ":key,info:airline,schedule:departure,delay:departure"  
)  
TBLPROPERTIES ("hbase.table.name" = "flight_info");
```

d. Find the total departure delay in Hive

```
SELECT SUM(departure_delay) AS total_departure_delay FROM flight_hive;
```

e. Find the average departure delay in Hive

```
SELECT AVG(departure_delay) AS avg_departure_delay FROM flight_hive;
```

f. Create index on Flight information Table

```
CREATE INDEX idx_departure_delay  
ON TABLE flight_hive (departure_delay)  
AS 'COMPACT'  
WITH DEFERRED REBUILD;  
  
ALTER INDEX idx_departure_delay ON flight_hive REBUILD;  
  
SHOW INDEX ON flight_hive;
```