



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

Open HBase shell:

bash

CopyEdit

```
hbase shell
```

Run the following commands:

hbase

CopyEdit

```
create 'flight_info', 'details', 'schedule', 'delay'
```

- **details**: stores flight number, origin, destination
- **schedule**: stores departure time, arrival time
- **delay**: stores delay in minutes

Example (adding data):

hbase

CopyEdit

```
put 'flight_info', 'FL1001', 'details:flight_no', 'AI101'
```

```
put 'flight_info', 'FL1001', 'details:origin', 'DEL'
```

```
put 'flight_info', 'FL1001', 'details:destination', 'NYC'

put 'flight_info', 'FL1001', 'schedule:departure', '2025-05-07
09:00'

put 'flight_info', 'FL1001', 'schedule:arrival', '2025-05-07
20:00'

put 'flight_info', 'FL1001', 'delay:departure_delay', '45'
```

b. Demonstrate Creating, Dropping, and Altering HBase Tables

Create another table:

hbase

CopyEdit

```
create 'flight_status', 'info'
```

Alter the table:

hbase

CopyEdit

```
alter 'flight_info', NAME => 'status'
```

Drop the table:

hbase

CopyEdit

```
disable 'flight_status'
```

```
drop 'flight_status'
```



c. Create External Hive Table to Connect to HBase Table

Go to Hive shell:

```
bash
```

CopyEdit

```
hive
```

Now create an external Hive table that maps to `flight_info` in HBase:

```
sql
```

CopyEdit

```
CREATE EXTERNAL TABLE flight_info_hive(  
    flight_no STRING,  
    origin STRING,  
    destination STRING,  
    departure_time STRING,  
    arrival_time STRING,
```

```
        departure_delay INT
    )

    STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'

    WITH SERDEPROPERTIES (

        "hbase.columns.mapping" =
        ":key,details:origin,details:destination,schedule:departure,schedule:arrival,delay:departure_delay"

    )

    TBLPROPERTIES ("hbase.table.name" = "flight_info");
```



d. Find Total Departure Delay in Hive

sql

CopyEdit

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



e. Find Average Departure Delay in Hive

sql

CopyEdit

```
SELECT AVG(departure_delay) AS average_departure_delay FROM  
flight_info_hive;
```

f. Create Index on Flight Info Table

You can create an index on a frequently filtered column, like `origin`:

sql

CopyEdit

```
CREATE INDEX idx_origin  
ON TABLE flight_info_hive (origin)  
AS 'COMPACT'  
WITH DEFERRED REBUILD;
```

To rebuild:

sql

CopyEdit

```
ALTER INDEX idx_origin ON flight_info_hive REBUILD;  
SHOW FORMATTED INDEX ON flight_info_hive;
```