FOR CREATING EMPLOYEE TABLES:

Create a employee.txt file for first table as:

101,John Smith,Manager,Human Resources
102,Jane Doe,Engineer,Engineering
103,Michael Johnson,Analyst,Finance
104,Sarah Williams,Coordinator,Marketing
105,David Lee,Developer,IT
106,Emily Brown,Assistant,Administration
107,Robert Davis,Supervisor,Operations
108,Lisa Wilson,Designer,Creative
109,James Anderson,Technician,Technical Support
110,Amy Taylor,Accountant,Finance

Create another employee2.txt file for second table as:

101,50000,3
102,65000,5
103,48000,2
104,72000,1
105,55000,4
106,48000,0
107,62000,2
108,53000,3
109,48000,1
110,58000,4

Create Tables with these two files

```
hive> CREATE TABLE emp(
    >   EMPLOYEE_ID INT,
    >   EMPLOYEE_NAME STRING,
    >   EMPLOYEE_DESIGNATION STRING,
    >   DEPARTMENT STRING
    > )
    > ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
    > ;
OK
Time taken: 0.109 seconds
hive> LOAD DATA LOCAL INPATH
```

```
        > '/home/akansha/employee.txt'
        > INTO TABLE emp;
Loading data to table default.emp
OK
Time taken: 0.268 seconds
hive> select * from emp;
OK
101    John Smith    Manager    Human Resources
102    Jane Doe    Engineer    Engineering
103    Michael Johnson    Analyst    Finance
104    Sarah Williams    Coordinator    Marketing
105    David Lee    Developer    IT
106    Emily Brown    Assistant    Administration
107    Robert Davis    Supervisor    Operations
108    Lisa Wilson    Designer    Creative
109    James Anderson    Technician    Technical Support
110    Amy Taylor    Accountant    Finance
Time taken: 0.155 seconds, Fetched: 11 row(s)
hive>

hive> CREATE TABLE emp2 (
        >   EMPLOYEE_ID INT,
        >   SALARY STRING,
        >   LEAVES_TAKEN INT
        > )
        > ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';
OK
Time taken: 0.125 seconds
hive> LOAD DATA LOCAL INPATH
        > '/home/akansha/employee2.txt'
        > INTO TABLE emp2;
Loading data to table default.emp2
OK
Time taken: 0.198 seconds
hive> select * from emp2;
OK
101    50000    3
102    65000    5
103    48000    2
104    72000    1
105    55000    4
106    48000    0
107    62000    2
108    53000    3
109    48000    1
110    58000    4
Time taken: 0.124 seconds, Fetched: 10 row(s)
hive>
```

Perform JOIN:

hive> SELECT e.EMPLOYEE_ID, e.EMPLOYEE_NAME, e.EMPLOYEE_DESIGNATION, e.DEPARTMENT, s.SALARY, s.LEAVES_TAKEN
    > FROM emp e
    > JOIN emp2 s ON e.EMPLOYEE_ID = s.EMPLOYEE_ID;
Query ID = akansha_20230528035317_58e8f370-213b-41fc-a0af-06cb1acf7f90
Total jobs = 1
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1685219240726_0007, Tracking URL =
http://akansha-VirtualBox:8088/proxy/application_1685219240726_0007/
Kill Command = /home/akansha/hadoop-3.2.4/bin/mapred job -kill
job_1685219240726_0007
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2023-05-28 03:53:35,565 Stage-3 map = 0%, reduce = 0%
2023-05-28 03:53:41,936 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 3.43 sec
MapReduce Total cumulative CPU time: 3 seconds 430 msec
Ended Job = job_1685219240726_0007
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 3.43 sec   HDFS Read: 10232 HDFS Write: 660
SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 430 msec
OK
101   John Smith   Manager   Human Resources   50000   3
102   Jane Doe   Engineer   Engineering   65000   5
103   Michael Johnson   Analyst   Finance   48000   2
104   Sarah Williams   Coordinator   Marketing   72000   1
105   David Lee   Developer   IT   55000   4
106   Emily Brown   Assistant   Administration   48000   0
107   Robert Davis   Supervisor   Operations   62000   2
108   Lisa Wilson   Designer   Creative   53000   3
109   James Anderson   Technician   Technical Support   48000   1
110   Amy Taylor   Accountant   Finance   58000   4
Time taken: 26.774 seconds, Fetched: 10 row(s)
hive>

hive> show tables;
OK
emp
emp2
flight_info
Time taken: 0.022 seconds, Fetched: 3 row(s)