STUDENT TABLES IN HIVE:

Create student1.txt and student2.txt as follows:

student1.txt:

1,John Smith,Computer Science
2,Jane Doe,Electrical Engineering
3,Michael Johnson,Mechanical Engineering
4,Sarah Williams,Chemistry
5,David Lee,Physics
6,Emily Brown,Mathematics
7,Robert Davis,English Literature
8,Lisa Wilson,Business Administration
9,James Anderson,Psychology
10,Amy Taylor,History

student2.txt:

1,Programming,90
2,Electrical Circuits,75
3,Thermodynamics,88
4,Organic Chemistry,70
5,Quantum Mechanics,85
6,Calculus,82
7,English Literature,92
8,Management,78
9,Psychological Disorders,77
10,World History,83

CREATE TABLES :

```
hive> CREATE TABLE student1(
    >   STUDENT_ID INT,
    >   STUDENT_NAME STRING,
    >   STUDENT_BRANCH STRING
    > )
    > ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';
OK
Time taken: 0.07 seconds
hive> LOAD DATA LOCAL INPATH
    > '/home/akansha/student1.txt'
    > INTO TABLE student1;
Loading data to table default.student1
```

```
OK
Time taken: 0.22 seconds
hive> select * from student1;
OK
1    John Smith    Computer Science
2    Jane Doe    Electrical Engineering
3    Michael Johnson    Mechanical Engineering
4    Sarah Williams    Chemistry
5    David Lee    Physics
6    Emily Brown    Mathematics
7    Robert Davis    English Literature
8    Lisa Wilson    Business Administration
9    James Anderson    Psychology
10    Amy Taylor    History
Time taken: 0.157 seconds, Fetched: 10 row(s)
hive>

hive> CREATE TABLE student2 (
    >    STUDENT_ID INT,
    >    SUBJECT STRING,
    >    MARKS INT
    > )
    > ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';
OK
Time taken: 0.109 seconds
hive> LOAD DATA LOCAL INPATH
    > '/home/akansha/student2.txt'
    > INTO TABLE student2;
Loading data to table default.student2
OK
Time taken: 0.656 seconds
hive> select * from student2;
OK
1    Programming    90
2    Electrical Circuits    75
3    Thermodynamics    88
4    Organic Chemistry    70
5    Quantum Mechanics    85
6    Calculus    82
7    English Literature    92
8    Management    78
9    Psychological Disorders    77
10    World History    83
Time taken: 0.143 seconds, Fetched: 10 row(s)
hive>
```

Perform JOIN:

Query ID = akansha_20230528042141_536876ae-3dca-443d-85d5-a44a9a157cdc
Total jobs = 1
2023-05-28 04:21:48   Dump the side-table for tag: 1 with group count: 10 into file: file:/tmp/hive/java/akansha/1ce77038-0521-49ca-8952-374c7ab05d69/hive_2023-05-28_04-21-41_267_2807766107383456637-1/-local-10004/HashTable-Stage-3/MapJoin-mapfile11--.hashtable
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_0008, Tracking URL = http://akansha-VirtualBox:8088/proxy/application_1685219240726_0008/
Kill Command = /home/akansha/hadoop-3.2.4/bin/mapred job -kill job_1685219240726_0008
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2023-05-28 04:21:56,130 Stage-3 map = 0%, reduce = 0%
2023-05-28 04:22:05,654 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 4.54 sec
MapReduce Total cumulative CPU time: 4 seconds 540 msec
Ended Job = job_1685219240726_0008
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 4.54 sec   HDFS Read: 9828 HDFS Write: 697 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 540 msec
OK
1   John Smith   Computer Science   Programming   90
2   Jane Doe   Electrical Engineering   Electrical Circuits   75
3   Michael Johnson   Mechanical Engineering   Thermodynamics   88
4   Sarah Williams   Chemistry   Organic Chemistry   70
5   David Lee   Physics   Quantum Mechanics   85
6   Emily Brown   Mathematics   Calculus   82
7   Robert Davis   English Literature   English Literature   92
8   Lisa Wilson   Business Administration   Management   78
9   James Anderson   Psychology   Psychological Disorders   77
10   Amy Taylor   History   World History   83
Time taken: 25.54 seconds, Fetched: 10 row(s)
hive>

Query ID = akansha_20230528042552_9e8bedb8-3738-4a04-b983-4770d3f034f7
Total jobs = 1
2023-05-28 04:25:59   End of local task; Time Taken: 1.547 sec.
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_0009, Tracking URL =
http://akansha-VirtualBox:8088/proxy/application_1685219240726_0009/
Kill Command = /home/akansha/hadoop-3.2.4/bin/mapred job -kill
job_1685219240726_0009
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
2023-05-28 04:26:08,288 Stage-3 map = 0%, reduce = 0%
2023-05-28 04:26:14,572 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 2.81 sec
MapReduce Total cumulative CPU time: 2 seconds 810 msec
Ended Job = job_1685219240726_0009
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1   Cumulative CPU: 2.81 sec   HDFS Read: 9837 HDFS Write: 437
SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 810 msec
OK
1   John Smith   Computer Science   Programming   90
3   Michael Johnson   Mechanical Engineering   Thermodynamics   88
5   David Lee   Physics   Quantum Mechanics   85
6   Emily Brown   Mathematics   Calculus   82
7   Robert Davis   English Literature   English Literature   92
10   Amy Taylor   History   World History   83
Time taken: 22.88 seconds, Fetched: 6 row(s)
hive>

hive> select * from student2 where MARKS>80;
OK
1   Programming   90
3   Thermodynamics   88
5   Quantum Mechanics   85
6   Calculus   82
7   English Literature   92
10   World History   83
Time taken: 0.159 seconds, Fetched: 6 row(s)
hive>