

Dataset:

001	Amit	105	Data Mining
002	Pankaj	85	Data Engineer
003	Kiran	110	Data Scientist
004	Arpitha	95	Data Engineer
005	Viraj	105	Data Mining
006	Smitha	80	Data Analyst
007	Supriya	90	Data Engineer
008	Vihan	120	Data Scientist
009	Emma	100	Data Engineer
010	Siddharath	100	Data Engineer

create table and load data to employee_salary table:

```
[hive> create table employee_salary(emp_id INT,emp_name STRING,emp_sal INT,emp_dep STRING) row format delimited fields terminated by '\t' stored as
textfile;
OK
Time taken: 1.932 seconds
[hive> load data local inpath "/Users/Disha/Downloads/Emp_Sal.txt" into table employee_salary;
Loading data to table custom.employee_salary
OK
Time taken: 2.19 seconds
[hive> select * from employee_salary;
OK
1      Amit      105      Data Mining
2      Pankaj     85      Data Engineer
3      Kiran      110     Data Scientist
4      Arpitha    95      Data Engineer
5      Viraj      105     Data Mining
6      Smitha     80      Data Analyst
7      Supriya    90      Data Engineer
8      Vihan      120     Data Scientist
9      Emma       100     Data Engineer
10     Siddharath 100     Data Engineer
Time taken: 3.089 seconds, Fetched: 10 row(s)
```

Get a list of employees who receive a salary less than 100, compared to their immediate employee with higher salary in the same unit

```
[hive> select * from employee_salary x join (select max(emp_sal) as max_sal, emp_dep from employee_salary group by emp_dep) y where x.emp_sal<100
and x.emp_sal < y.max_sal and x.emp_dep = y.emp_dep;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spa
rk, tez) or using Hive 1.X releases.
Query ID = Disha_20170919233420_c00e2cbc-ad84-4a96-ae34-b6c1136e04f2
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Job running in-process (local Hadoop)
2017-09-19 23:34:23,311 Stage-2 map = 100%,  reduce = 100%
Ended Job = job_local1598940393_0003
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/Cellar/hive/2.1.1/libexec/lib/log4j-slf4j-impl-2.4.1.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/Cellar/hadoop/2.8.1/libexec/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLo
ggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2017-09-19 23:34:36      Starting to launch local task to process map join;          maximum memory = 477626368
2017-09-19 23:34:37      Dump the side-table for tag: 0 with group count: 2 into file: file:/tmp/hive/fb84165e-fb0c-4270-910f-eee41b00646b/hive_201
7-09-19_23-34-20_929_5233119552622490137-1/-local-10005/HashTable-Stage-3/MapJoin-mapfile10---.hashtable
2017-09-19 23:34:37      Uploaded 1 file to: file:/tmp/hive/fb84165e-fb0c-4270-910f-eee41b00646b/hive_2017-09-19_23-34-20_929_5233119552622490137-1
/-local-10005/HashTable-Stage-3/MapJoin-mapfile10---.hashtable (371 bytes)
2017-09-19 23:34:37      End of local task; Time Taken: 1.539 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 2 out of 2
Number of reduce tasks is set to 0 since there's no reduce operator
Job running in-process (local Hadoop)
2017-09-19 23:34:42,262 Stage-3 map = 100%,  reduce = 0%
Ended Job = job_local1690416873_0004
MapReduce Jobs Launched:
Stage-Stage-2:  HDFS Read: 1164 HDFS Write: 0 SUCCESS
Stage-Stage-3:  HDFS Read: 582 HDFS Write: 0 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
OK
2      Pankaj     85      Data Engineer    100     Data Engineer
4      Arpitha    95      Data Engineer    100     Data Engineer
7      Supriya    90      Data Engineer    100     Data Engineer
Time taken: 21.339 seconds, Fetched: 3 row(s)
```

List of all employees who draw higher salary than the average salary of that department

```
hive> select * from employee_salary x join (select avg(emp_sal) as average,emp_dep from employee_salary group by emp_dep) y where x.emp_sal > y.ave
verage and x.emp_dep = y.emp_dep;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spa
rk, tez) or using Hive 1.X releases.
Query ID = Disha_20170919181554_ed1893a8-161a-45a8-a524-39fcb4dcf06e
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Job running in-process (local Hadoop)
2017-09-19 18:15:56,555 Stage-2 map = 100%,  reduce = 100%
Ended Job = job_local398881469_0007
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/Cellar/hive/2.1.1/libexec/lib/log4j-slf4j-impl-2.4.1.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/Cellar/hadoop/2.8.1/libexec/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLo
ggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2017-09-19 18:16:00 Starting to launch local task to process map join; maximum memory = 477626368
2017-09-19 18:16:09 Dump the side-table for tag: 0 with group count: 5 into file: file:/tmp/hive/83eedbaf-345d-44fd-bfd8-71c5c14b05ee/hive_201
7-09-19_18-15-54_562_6402714583794924154-1/-local-10005/HashTable-Stage-3/MapJoin-mapfile20---.hashtable
2017-09-19 18:16:09 Uploaded 1 File to: file:/tmp/hive/83eedbaf-345d-44fd-bfd8-71c5c14b05ee/hive_2017-09-19_18-15-54_562_6402714583794924154-1
/-local-10005/HashTable-Stage-3/MapJoin-mapfile20---.hashtable (534 bytes)
2017-09-19 18:16:09 End of local task; Time Taken: 1.429 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 2 out of 2
Number of reduce tasks is set to 0 since there's no reduce operator
Job running in-process (local Hadoop)
2017-09-19 18:16:14,137 Stage-3 map = 100%,  reduce = 0%
Ended Job = job_local1764014515_0008
MapReduce Jobs Launched:
Stage-Stage-2: HDFS Read: 2910 HDFS Write: 0 SUCCESS
Stage-Stage-3: HDFS Read: 1455 HDFS Write: 0 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
OK
4 Arpitha 95 Data Engineer 94.0 Data Engineer
9 Emma 100 Data Engineer 94.0 Data Engineer
10 Siddharath 100 Data Engineer 94.0 Data Engineer
8 Vihan 120 Data Scientist 115.0 Data Scientist
Time taken: 19.58 seconds, Fetched: 4 row(s)
```