```
[cloudera@quickstart ~]$ hive

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.p
roperties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> show databases;
OK
default
Time taken: 0.627 seconds, Fetched: 1 row(s)
```

```
hive> create database rjc;
OK
Time taken: 2.833 seconds
hive> use rjc;
OK
Time taken: 0.186 seconds
```

```
hive> show databases;
OK
default
rjc
Time taken: 0.022 seconds, Fetched: 2 row(s)
```

```
hive> show tables;
OK
Time taken: 0.077 seconds
```

```
hive> create table employee(ID int, name string ,salary float,age int)
    > row format delimited
    > fields terminated by',';
OK
Time taken: 0.46 seconds
```

```
hive> describe employee;
OK
id                      int
name                    string
salary                  float
age                     int
Time taken: 0.161 seconds, Fetched: 4 row(s)
```

```
hive> describe formatted employee;
OK
# col_name              data_type               comment

id                      int
name                    string
salary                  float
age                     int

# Detailed Table Information
Database:               rjc
Owner:                  cloudera
CreateTime:             Thu Mar 10 19:31:57 PST 2022
LastAccessTime:         UNKNOWN
Protect Mode:           None
Retention:              0
Location:               hdfs://quickstart.cloudera:8020/user/hive/warehouse/rjc.
db/employee
Table Type:             MANAGED_TABLE
Table Parameters:
        transient_lastDdlTime   1646969517

# Storage Information
SerDe Library:          org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe
InputFormat:            org.apache.hadoop.mapred.TextInputFormat
OutputFormat:           org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputForm
at
Compressed:             No
Num Buckets:            -1
Bucket Columns:         []
Sort Columns:           []
Storage Desc Params:
        field.delim             ,
        serialization.format    ,
Time taken: 0.093 seconds, Fetched: 30 row(s)
```

```
hive> create external table emloyee2 (ID int, name string, salary float, age int
)
    > row format delimited
    > fields terminated by','
    > stored as textfile;
OK
Time taken: 0.053 seconds
```

```
hive> describe emloyee2;
OK
id                      int
name                    string
salary                  float
age                     int
Time taken: 0.092 seconds, Fetched: 4 row(s)
```

```
hive> create external table emloyee3 (ID int, name string, salary float, age int)
    > row format delimited
    > fields terminated by ','
    > location '/user/cloudera/vj';
OK
Time taken: 0.084 seconds
hive>
```

```
hive> describe emloyee3;
OK
id                      int
name                    string
salary                  float
age                     int
Time taken: 0.086 seconds, Fetched: 4 row(s)
hive>
```

```
hive> alter table emloyee3 RENAME TO emptable;
OK
Time taken: 0.204 seconds
```

```
hive> show tables;
OK
emloyee2
employee
emptable
Time taken: 0.011 seconds, Fetched: 3 row(s)
```

```
hive> describe emptable;
OK
id                      int
name                    string
salary                  float
age                     int
Time taken: 0.081 seconds, Fetched: 4 row(s)
hive> Alter table emptable add columns (surname string);
OK
Time taken: 0.252 seconds
hive> describe emptable;
OK
id                      int
name                    string
salary                  float
age                     int
surname                 string
Time taken: 0.076 seconds, Fetched: 5 row(s)
hive>
```

```
hive> Alter table emptable change name first_name string;
OK
Time taken: 0.177 seconds
hive> describe emptable;
OK
id                      int
first_name              string
salary                  float
age                     int
surname                 string
Time taken: 0.088 seconds, Fetched: 5 row(s)
hive>
```

```
hive>
    >
    > create database rjcstudent;
OK
Time taken: 0.035 seconds
hive> show databases;
OK
default
rjc
rjcstudent
Time taken: 0.017 seconds, Fetched: 3 row(s)
```

```
hive> create table student (ID int, Name string, Age int)
    > partitioned by(Course string)
    > row format delimited
    > fields terminated by ',';
OK
Time taken: 0.152 seconds
hive>
```

```
hive> describe student;
OK
id                      int
name                    string
age                     int
course                  string

# Partition Information
# col_name              data_type               comment

course                  string
Time taken: 0.071 seconds, Fetched: 9 row(s)
hive>
```

```
hive> load data local inpath '/home/cloudera/Documents/Student.csv' into table student
    > partition(Course ='HADOOP')
    > ;
Loading data to table rjcstudent.student partition (course=HADOOP)
Partition rjcstudent.student{course=HADOOP} stats: [numFiles=1, numRows=0, totalSize=99, rawDataSize=0]
OK
Time taken: 0.945 seconds
```

```
hive> SELECT * FROM STUDENT;
OK
NULL    NAME    NULL    HADOOP
1       REHAN   NULL    HADOOP
2       RISHI   NULL    HADOOP
3       SHIVAM  NULL    HADOOP
4       ANAND   NULL    HADOOP
5       PRINCE  NULL    HADOOP
Time taken: 0.532 seconds, Fetched: 6 row(s)
hive>
```

```
hive> load data local inpath '/home/cloudera/Documents/Student.csv' into table student
    > partition(Course ='PYTHON')
    > ;
Loading data to table rjcstudent.student partition (course=PYTHON)
Partition rjcstudent.student{course=PYTHON} stats: [numFiles=1, numRows=0, totalSize=99, rawDataSize=0]
OK
Time taken: 0.367 seconds
hive> SELECT * FROM STUDENT;
OK
NULL    NAME    NULL    HADOOP
1       REHAN   NULL    HADOOP
2       RISHI   NULL    HADOOP
3       SHIVAM  NULL    HADOOP
4       ANAND   NULL    HADOOP
5       PRINCE  NULL    HADOOP
NULL    NAME    NULL    PYTHON
1       REHAN   NULL    PYTHON
2       RISHI   NULL    PYTHON
3       SHIVAM  NULL    PYTHON
4       ANAND   NULL    PYTHON
5       PRINCE  NULL    PYTHON
Time taken: 0.087 seconds, Fetched: 12 row(s)
```

```
hive> load data local inpath '/home/cloudera/Documents/Student.csv' into table student
    > partition(Course ='ML')
    > ;
Loading data to table rjcstudent.student partition (course=ML)
Partition rjcstudent.student{course=ML} stats: [numFiles=1, numRows=0, totalSize=99, rawDataSize=0]
OK
Time taken: 0.374 seconds
hive> SELECT * FROM STUDENT;
OK
NULL    NAME    NULL    HADOOP
1       REHAN   NULL    HADOOP
2       RISHI   NULL    HADOOP
3       SHIVAM  NULL    HADOOP
4       ANAND   NULL    HADOOP
5       PRINCE  NULL    HADOOP
NULL    NAME    NULL    ML
1       REHAN   NULL    ML
2       RISHI   NULL    ML
3       SHIVAM  NULL    ML
4       ANAND   NULL    ML
5       PRINCE  NULL    ML
NULL    NAME    NULL    PYTHON
1       REHAN   NULL    PYTHON
2       RISHI   NULL    PYTHON
3       SHIVAM  NULL    PYTHON
4       ANAND   NULL    PYTHON
5       PRINCE  NULL    PYTHON
Time taken: 0.078 seconds, Fetched: 18 row(s)
hive>
```

```
hive> DROP table student;
OK
Time taken: 0.774 seconds
hive> create table student (ID int, Name string, Age int)
    > row format delimited
    > fields terminated by ','
    > tblproperties("skip.header.line.count" ="1");
```

```
hive> create table student (ID int, Name string, Age int)
    > row format delimited
    > fields terminated by ','
    > tblproperties("skip.header.line.count" ='1');
OK
Time taken: 0.06 seconds
hive> SELECT * FROM STUDENT;
OK
Time taken: 0.061 seconds
```

```
hive> load data local inpath '/home/cloudera/Documents/Student.csv' into table s
tudent
    > partition(Course="Hadoop")
    > ;
Loading data to table default.student partition (course=Hadoop)
Partition default.student{course=Hadoop} stats: [numFiles=1, numRows=0, totalSiz
e=127, rawDataSize=0]
OK
Time taken: 1.007 seconds
hive> select * from student;
OK
NULL    Name    NULL    Hadoop
1       Akshata NULL    Hadoop
2       Sarita  NULL    Hadoop
3       Priti   NULL    Hadoop
4       Shivani NULL    Hadoop
5       Kajal   NULL    Hadoop
6       Ajay    NULL    Hadoop
Time t
```

```
hive> load data local inpath '/home/cloudera/Documents/Student.csv' into table s
tudent
    > partition(Course="Java");;
Loading data to table default.student partition (course=Java)
Partition default.student{course=Java} stats: [numFiles=1, numRows=0, totalSize=
127, rawDataSize=0]
OK
Time taken: 0.42 seconds
hive> load data local inpath '/home/cloudera/Documents/Student.csv' into table s
tudentpartition(Course="Python");
Loading data to table default.student partition (course=Python)
Partition default.student{course=Python} stats: [numFiles=1, numRows=0, totalSiz
e=127, rawDataSize=0]
OK
Time taken: 0.558 seconds
```

```
hive> drop table student;
OK
Time t;
OK
Time taken: 0.053 seconds
```

```
hive> create table student(ID int,Name string,Course string,Age int)
    > row format delimited
    > fields terminated by','
    > tblproperties("skip.header.line.count"="1");
OK
Time taken: 0.096 seconds
```

```
hive> load data local inpath '/home/cloudera/Documents/Student.csv' into table s
tudent;
Loading data to table default.student
Table default.student stats: [numFiles=1, totalSize=127]
OK
Time taken: 0.209 seconds
hive> select * from student;
OK
1       Akshata Hadoop  21
2       Sarita  Java    22
3       Priti   Java    23
4       Shivani Python  25
5       Kajal   Hadoop  21
6       Ajay    Python  23
Time taken: 0.076 seconds, Fetched: 6 row(s)
```

```
hive> create database hiveql;
OK
Time taken: 0.154 seconds
```

```
hive> create table employee(ID int,Name string,Department string,YOJ int,Salary
float)
    > ROW FORMAT DELIMITED
    >  FIELDS TERMINATED BY ','
    > tblproperties("skip.header.line.count"="1");
OK
Time taken: 0.299 seconds
hive> describe employee;
OK
id                      int
name                    string
department              string
yoj                     int
salary                  float
Time taken: 0.182 seconds, Fetched: 5 row(s)
```

```
hive> load data local inpath '/home/cloudera/Documents/employee.csv' into table
employee;
Loading data to table default.employee
Table default.employee stats: [numFiles=1, totalSize=121]
OK
Time taken: 0.538 seconds
hive> select * from employee;
OK
1       Akshata IT      2018    50000.0
2       Sarita  Sales   2022    23000.0
3       Priti   HR      2012    34000.0
4       Shivani SC      2015    45000.0
NULL    NULL    NULL    NULL    NULL
Time taken: 0.264 seconds, Fetched: 5 row(s)
```

```
hive> select * from employee where salary >=25000;
OK
1       Akshata IT      2018    50000.0
3       Priti   HR      2012    34000.0
4       Shivani SC      2015    45000.0
Time taken: 0.246 seconds, Fetched: 3 row(s)
```

```
hive> select * from employee where salary <25000;
OK
2       Sarita  Sales   2022    23000.0
Time taken: 0.109 seconds, Fetched: 1 row(s)
```

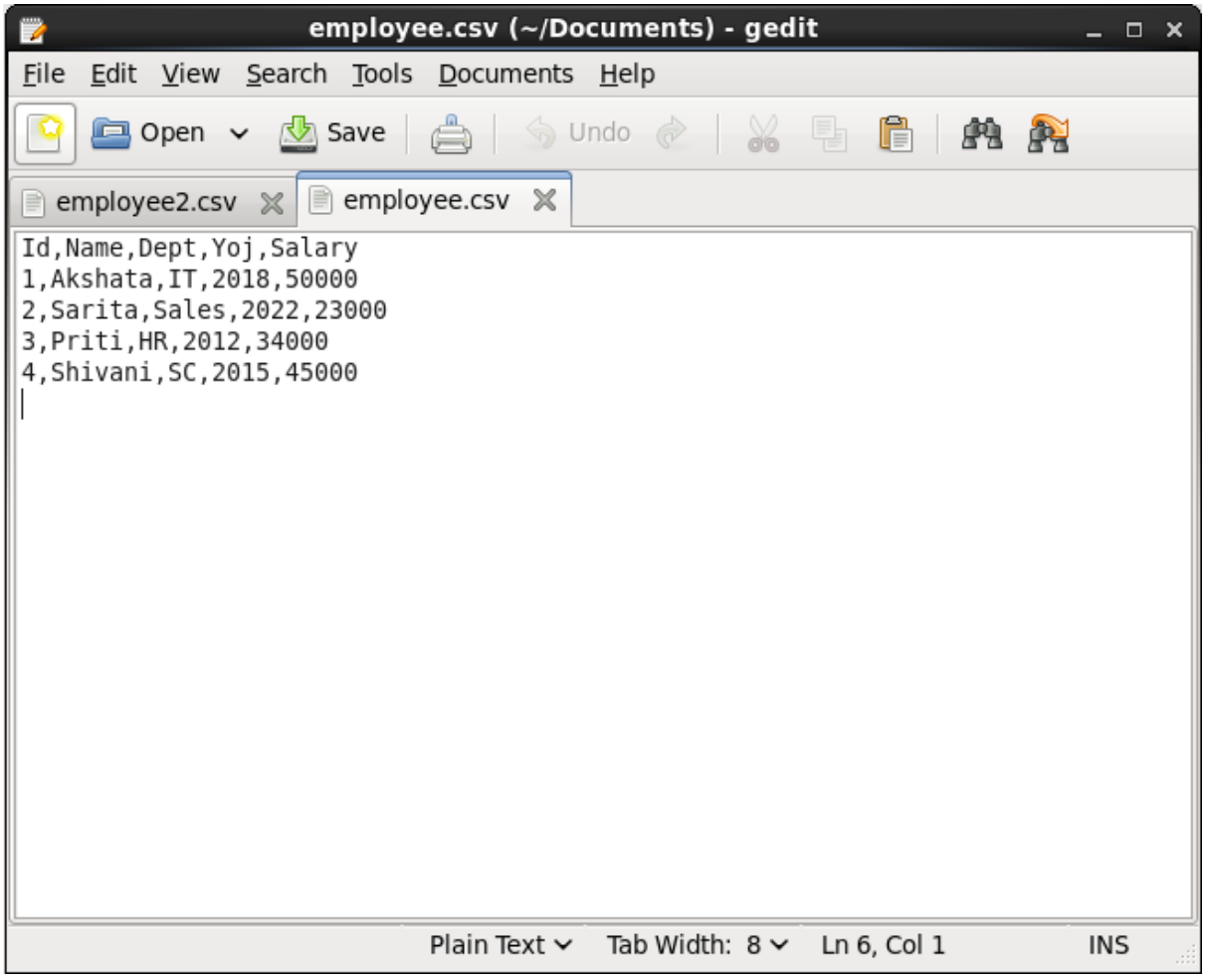
```
hive> select ID ,name,salary+5000 from employee;
OK
1       Akshata 55000.0
2       Sarita  28000.0
3       Priti   39000.0
4       Shivani 50000.0
NULL    NULL    NULL
Time taken: 0.07 seconds, Fetched: 5 row(s)
```

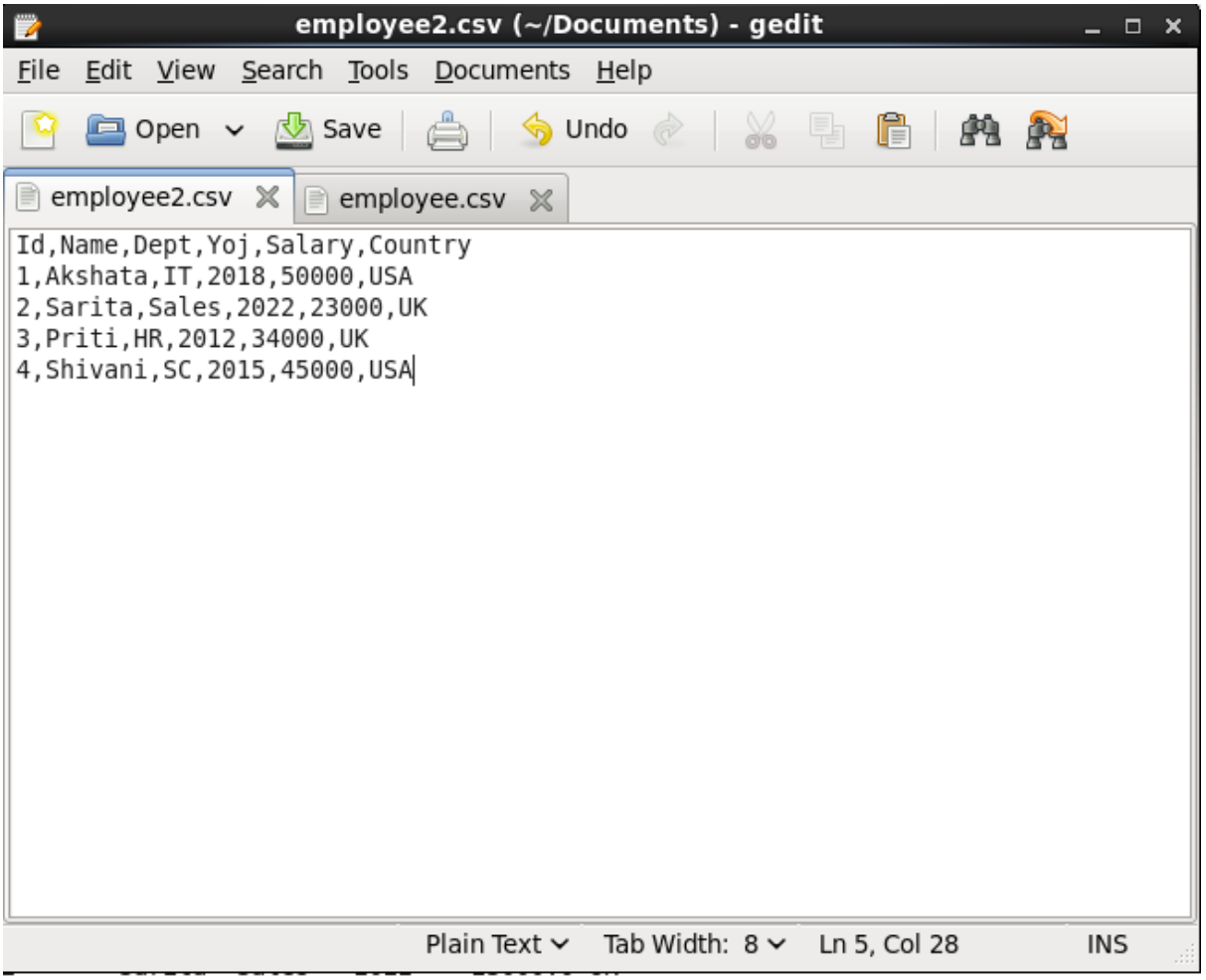
```
hive> select max(Salary)from employee;
Query ID = cloudera_20220314210707_1f8c8a5d-d609-44de-aa5e-7437d774bbb3
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<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>
Starting Job = job_1644894610889_0012, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1644894610889_0012/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1644894610889_0012
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-03-14 21:08:09,884 Stage-1 map = 0%,  reduce = 0%
2022-03-14 21:08:18,811 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.86 se
c
2022-03-14 21:08:29,591 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.96
sec
MapReduce Total cumulative CPU time: 1 seconds 960 msec
Ended Job = job_1644894610889_0012
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.96 sec   HDFS Read: 6880 HD
FS Write: 8 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 960 msec
OK
50000.0
Time taken: 34.969 seconds, Fetched: 1 row(s)
```

```
hive> select min(Salary)from employee;
Query ID = cloudera_20220314210808_a7f2dda9-084c-4d61-8f70-f52d55661d33
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 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>
Starting Job = job_1644894610889_0013, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1644894610889_0013/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1644894610889_0013
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-03-14 21:09:01,568 Stage-1 map = 0%,  reduce = 0%
2022-03-14 21:09:09,159 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.77 se
c
2022-03-14 21:09:17,760 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.84
sec
MapReduce Total cumulative CPU time: 1 seconds 840 msec
Ended Job = job_1644894610889_0013
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.84 sec   HDFS Read: 6903 HD
FS Write: 8 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 840 msec
OK
23000.0
Time taken: 27.823 seconds, Fetched: 1 row(s)
```

```
hive> select ID,Name,sqrt(Salary)from employee;
OK
1       Akshata 223.60679774997897
2       Sarita  151.65750888103102
3       Priti   184.39088914585776
4       Shivani 212.13203435596427
NULL    NULL    NULL
Time taken: 0.082 seconds, Fetched: 5 row(s)
```

```
hive> select ID,upper(Name)from employee;
OK
1       AKSHATA
2       SARITA
3       PRITI
4       SHIVANI
NULL    NULL
Time taken: 0.072 seconds, Fetched: 5 row(s)
```






```
hive> create table empgrp(ID int,Name string,Department string,YOJ int,Salary fl
oat,Country string)
    > ROW FORMAT DELIMITED
    >  FIELDS TERMINATED BY ','
    > tblproperties("skip.header.line.count"="1");
OK
Time taken: 0.071 seconds
```

```
hive> load data local inpath '/home/cloudera/Documents/employee2.csv' into table
 empgrp;
Loading data to table default.empgrp
Table default.empgrp stats: [numFiles=1, totalSize=142]
OK
Time taken: 0.372 seconds
hive> select * from empgrp;
OK
1       Akshata IT      2018    50000.0 USA
2       Sarita  Sales   2022    23000.0 UK
3       Priti   HR      2012    34000.0 UK
4       Shivani SC      2015    45000.0 USA
Time taken: 0.044 seconds, Fetched: 4 row(s)
```

```
hive> select Country,sum(Salary) from empgrp group by Country;
Query ID = cloudera_20220314211717_43a54ee2-d4d8-4df0-9fea-1cbc1a8e97bb
Total jobs = 1
Launching Job 1 out of 1
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>
Starting Job = job_1644894610889_0014, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1644894610889_0014/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1644894610889_0014
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-03-14 21:17:56,866 Stage-1 map = 0%,  reduce = 0%
2022-03-14 21:18:04,436 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.81 se
c
2022-03-14 21:18:12,942 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.88
sec
MapReduce Total cumulative CPU time: 1 seconds 880 msec
Ended Job = job_1644894610889_0014
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.88 sec   HDFS Read: 7321 HD
FS Write: 23 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 880 msec
OK
UK      57000.0
USA     95000.0
Time taken: 28.81 seconds, Fetched: 2 row(s)
```

```
hive> select Country,sum(Salary) from empgrp group by Country having sum(Salary)
>50000;
Query ID = cloudera_20220314211818_e90ab897-5891-4345-b25d-700a092f3ecb
Total jobs = 1
Launching Job 1 out of 1
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>
Starting Job = job_1644894610889_0015, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1644894610889_0015/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1644894610889_0015
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-03-14 21:18:55,704 Stage-1 map = 0%,  reduce = 0%
2022-03-14 21:19:03,196 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.8 sec
2022-03-14 21:19:12,920 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.3 s
ec
MapReduce Total cumulative CPU time: 2 seconds 300 msec
Ended Job = job_1644894610889_0015
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.3 sec   HDFS Read: 7759 HDF
S Write: 23 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 300 msec
OK
UK      57000.0
USA     95000.0
Time taken: 28.218 seconds, Fetched: 2 row(s)
```

```
hive> select * from empgrp order by Salary desc;
Query ID = cloudera_20220314211919_ff2ca655-14b8-4c7f-9a6f-0cda9ca4fa59
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 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>
Starting Job = job_1644894610889_0016, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1644894610889_0016/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1644894610889_0016
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-03-14 21:19:51,589 Stage-1 map = 0%,  reduce = 0%
2022-03-14 21:20:00,259 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.74 se
c
2022-03-14 21:20:08,901 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.83
sec
MapReduce Total cumulative CPU time: 1 seconds 830 msec
Ended Job = job_1644894610889_0016
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.83 sec   HDFS Read: 7186 HD
FS Write: 118 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 830 msec
OK
1       Akshata IT      2018    50000.0 USA
4       Shivani SC      2015    45000.0 USA
3       Priti   HR      2012    34000.0 UK
2       Sarita  Sales   2022    23000.0 UK
Time taken: 28.749 seconds, Fetched: 4 row(s)
```

```
Time taken: 28.749 seconds, Fetched: 4 row(s)
hive> select * from empgrp sort by Salary desc;
Query ID = cloudera_20220314212020_0570cc84-f98b-4239-99cf-dd239654abb6
Total jobs = 1
Launching Job 1 out of 1
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>
Starting Job = job_1644894610889_0017, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1644894610889_0017/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1644894610889_0017
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-03-14 21:20:49,709 Stage-1 map = 0%,  reduce = 0%
2022-03-14 21:20:57,245 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.74 se
c
2022-03-14 21:21:05,800 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.81
sec
MapReduce Total cumulative CPU time: 1 seconds 810 msec
Ended Job = job_1644894610889_0017
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.81 sec   HDFS Read: 7186 HD
FS Write: 118 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 810 msec
OK
1       Akshata IT      2018    50000.0 USA
4       Shivani SC      2015    45000.0 USA
3       Priti   HR      2012    34000.0 UK
2       Sarita  Sales   2022    23000.0 UK
Time taken: 27.599 seconds, Fetched: 4 row(s)
hive>
```