

Experiment 08

Querying ,Sorting ,Aggregating data using Hiveql

```
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: 2.327 seconds, Fetched: 1 row(s)
hive> create database RJC;
OK
Time taken: 9.281 seconds
hive> show databases;
OK
default
rjc
Time taken: 0.056 seconds, Fetched: 2 row(s)
hive> drop database RJC;
OK
Time taken: 2.987 seconds
hive> create database RJC;
OK
Time taken: 0.142 seconds
hive> use RJC;
OK
Time taken: 0.901 seconds
hive> show tables;
OK
Time taken: 0.509 seconds
hive> drop database RJC;
OK
Time taken: 0.109 seconds
hive>
```

```
cloudera@quickstart:~$ hive
hive> create database RJC;
OK
Time taken: 2.875 seconds
hive> show databases;
OK
default
rjc
Time taken: 0.11 seconds, Fetched: 2 row(s)
hive> use rjc;
OK
Time taken: 0.129 seconds
hive> create table employee(ID int,Name string,Salary float,Age int)
> row format delimited
> fields terminated by ',';
OK
Time taken: 0.35 seconds
hive> describe employee;
OK
id                int
name              string
salary           float
age              int
Time taken: 0.708 seconds, Fetched: 4 row(s)
hive> describe formatted employee;
FAILED: SemanticException [Error 10001]: Table not found formatted
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:       Mon Jul 26 10:59:09 PDT 2021
LastAccessTime:   UNKNOWN
Protect Mode:     None
```

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
hive> create external table employee2(ID int,Name string,Salary float,Age int)  
> row format delimited  
> fields terminated by ','  
> stored as textfile;  
OK  
Time taken: 0.45 seconds  
hive> describe employee2;  
OK  
id                int  
name              string  
salary            float  
age               int  
Time taken: 0.555 seconds, Fetched: 4 row(s)  
hive> describe formatted emplyee2;  
FAILED: SemanticException [Error 10001]: Table not found emplyee2  
hive> describe formatted employee2;  
OK  
# col_name          data_type          comment  
  
id                  int  
name                string  
salary              float  
age                 int  
  
# Detailed Table Information  
Database:           rjc  
Owner:              cloudera  
CreateTime:         Mon Jul 26 11:04:56 PDT 2021  
LastAccessTime:     UNKNOWN  
Protect Mode:       None  
Retention:          0  
Location:           hdfs://quickstart.cloudera:8020/user/hive/warehouse/rjc.db/employee2  
Table Type:         EXTERNAL_TABLE  
Table Parameters:  
    EXTERNAL                TRUE  
    transient_lastDdlTime   1627322696  
  
# Storage Information  
SerDe Library:      org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe  
InputFormat:        org.apache.hadoop.mapred.TextInputFormat  
OutputFormat:       org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat  
Compressed:         No  
Num Buckets:        -1  
Bucket Columns:     []  
Sort Columns:       []  
Storage Desc Params:  
    field.delim        ,  
    serialization.format  
Time taken: 0.56 seconds, Fetched: 31 row(s)
```

cloudera-quickstart-vm-5.4.2-0-virtualbox [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

Applications Places System

Hue - Hive Editor - Query - Mozilla Firefox

Hue - Hive Editor - Q... x

quickstart.cloudera:8888/ beeswax/execute/query/1#query/results

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security

Hive Editor Query Editor My Queries Saved Queries History

Assist Settings

DATABASE

rjc

Table name...

employee

- id (int)
- name (string)
- salary (float)
- age (int)

employee2

- id (int)
- name (string)
- salary (float)
- age (int)

```
1 show databases;  
2
```

Execute Save as... Explain or create a New query

Recent queries Query Log Columns Results Chart

	database_name
0	default
1	rjc

```

hive> create external table employee3(Id int,Name string,Salary float,Age int)
> row format delimited
> fields terminated by ','
> location '/user/cloudera/manoj';
OK
Time taken: 9.842 seconds
hive> describe employee3;
OK
id                int
name              string
salary            float
age               int
Time taken: 1.706 seconds, Fetched: 4 row(s)
hive> █

```

Hue - Hive Editor - Query x

quickstart.cloudera:8888/beeswax/execute/query/2#query

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security

Hive Editor Query Editor My Queries Saved Queries History

Assist Settings

DATABASE rjc

Table name...

- employee
 - id (int)
 - name (string)
 - salary (float)
 - age (int)
- employee2
 - id (int)
 - name (string)
 - salary (float)
 - age (int)
- employee3
 - id (int)
 - name (string)
 - salary (float)
 - age (int)

```

1 show databases;
2

```

Execute Save Save as... Explain or create a New query

Recent queries Query Log Columns Results Chart

	database_name
0	default
1	rjc

```

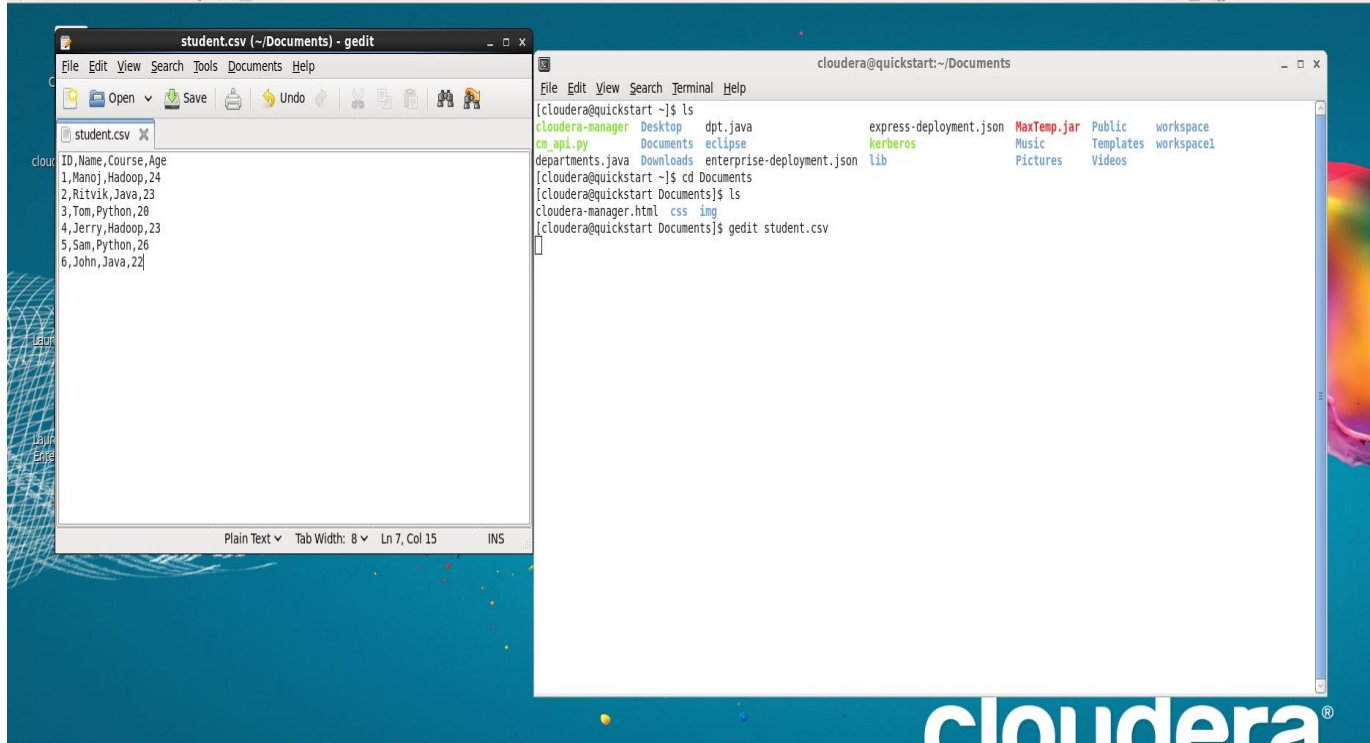
hive> create external table employee3(Id int,Name string,Salary float,Age int)
> row format delimited
> fields terminated by ','
> location '/user/cloudera/manoj';
OK
Time taken: 9.842 seconds
hive> describe employee3;
OK
id                int
name              string
salary            float
age               int
Time taken: 1.706 seconds, Fetched: 4 row(s)
hive> show tables;
OK
employee
employee2
employee3
Time taken: 3.211 seconds, Fetched: 3 row(s)
hive> alter table employee3 RENAME TO emptble;
OK
Time taken: 76.774 seconds
hive> alter table emptble add columns(surname string);
OK
Time taken: 1.155 seconds
hive> describe emptble;
OK
id                int
name              string
salary            float
age               int
surname           string
Time taken: 0.541 seconds, Fetched: 5 row(s)
hive> █

```

```

cloudera@quickstart:~
File Edit View Search Terminal Help
at org.apache.hadoop.hive.cli.CliDriver.main(CliDriver.java:615)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:606)
at org.apache.hadoop.util.RunJar.run(RunJar.java:221)
at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
FAILED: ParseException line 1:5 cannot recognize input near 'show' 'rjc' '<EOF>' in ddl statement
hive> show databases;
OK
default
rjc
Time taken: 0.025 seconds, Fetched: 2 row(s)
hive> use rjc;
OK
Time taken: 0.058 seconds
hive> show tables;
OK
employee
employee2
emptble
Time taken: 0.081 seconds, Fetched: 3 row(s)
hive> alter table emptble change name first_name string;
OK
Time taken: 14.39 seconds
hive> describe emptble;
OK
id                int
first_name        string
salary            float
age               int
surname           string
Time taken: 0.483 seconds, Fetched: 5 row(s)
hive> █

```



```
cloudera@quickstart:~
File Edit View Search Terminal Help
at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
FAILED: ParseException line 1:7 cannot recognize input near 'create' 'databases' 'rjcStudent' in ddl statement
hive> create database rjcStudent;
OK
Time taken: 2.822 seconds
hive> show databases;
OK
default
rjc
rjcstudent
Time taken: 0.204 seconds, Fetched: 3 row(s)
hive> use rjcstudent
> ;
OK
Time taken: 0.107 seconds
hive> create table student(ID int,Name string,Age int)
> partitioned by(Course string)
> row format delimited
> fields terminated by ',';
OK
Time taken: 2.294 seconds
hive> describe student;
OK
id                int
name              string
age              int
course           string
# Partition Information
# col_name        data_type        comment
course           string
Time taken: 0.471 seconds, Fetched: 9 row(s)
hive>
```

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
Time taken: 0.471 seconds, Fetched: 9 row(s)  
hive> load data local inpath '/home/cloudera/Documents/Student.csv' into table student  
  > partition(Course="Hadoop");  
FAILED: SemanticException Line 1:23 Invalid path ''/home/cloudera/Documents/Student.csv': No files matching path file:/home/  
cloudera/Documents/Student.csv  
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=119, rawDataSize=0]  
OK  
Time taken: 4.521 seconds  
hive> select *from student;  
OK  
NULL    Name    NULL    Hadoop  
1       Manoj   NULL    Hadoop  
2       Ritvik  NULL    Hadoop  
3       Tom    NULL    Hadoop  
4       Jerry  NULL    Hadoop  
5       Sam    NULL    Hadoop  
6       John   NULL    Hadoop  
Time taken: 8.009 seconds, Fetched: 7 row(s)  
hive> load data local inpath '/home/cloudera/Documents/student.csv' into table student  
  > partition(Course="Java");  
Loading data to table rjcstudent.student partition (course=Java)  
Partition rjcstudent.student{course=Java} stats: [numFiles=1, numRows=0, totalSize=119, rawDataSize=0]  
OK  
Time taken: 2.44 seconds  
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=119, rawDataSize=0]  
OK  
Time taken: 1.308 seconds  
hive> █
```

```
Time taken: 1.308 seconds  
hive> create table student(ID int,Name string,Course string,Age int)  
  > row format delimited  
  > fields terminated by ',';  
FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.DDLTask. AlreadyExistsException(message:Table stu  
dent already exists)  
hive> create table students(ID int,Name string,Course string,Age int)  
  > row format delimited  
  > fields terminated by ',';  
OK  
Time taken: 0.513 seconds  
hive> load data local inpath '/home/cloudera/Documents/student.csv' into table students;  
Loading data to table rjcstudent.students  
Table rjcstudent.students stats: [numFiles=1, totalSize=119]  
OK  
Time taken: 0.778 seconds  
hive> select *from students;  
OK  
NULL    Name    Course  NULL  
1       Manoj   Hadoop  24  
2       Ritvik  Java    23  
3       Tom    Python  20  
4       Jerry  Hadoop  23  
5       Sam    Python  26  
6       John   Java    22  
Time taken: 0.879 seconds, Fetched: 7 row(s)  
hive> █
```

```

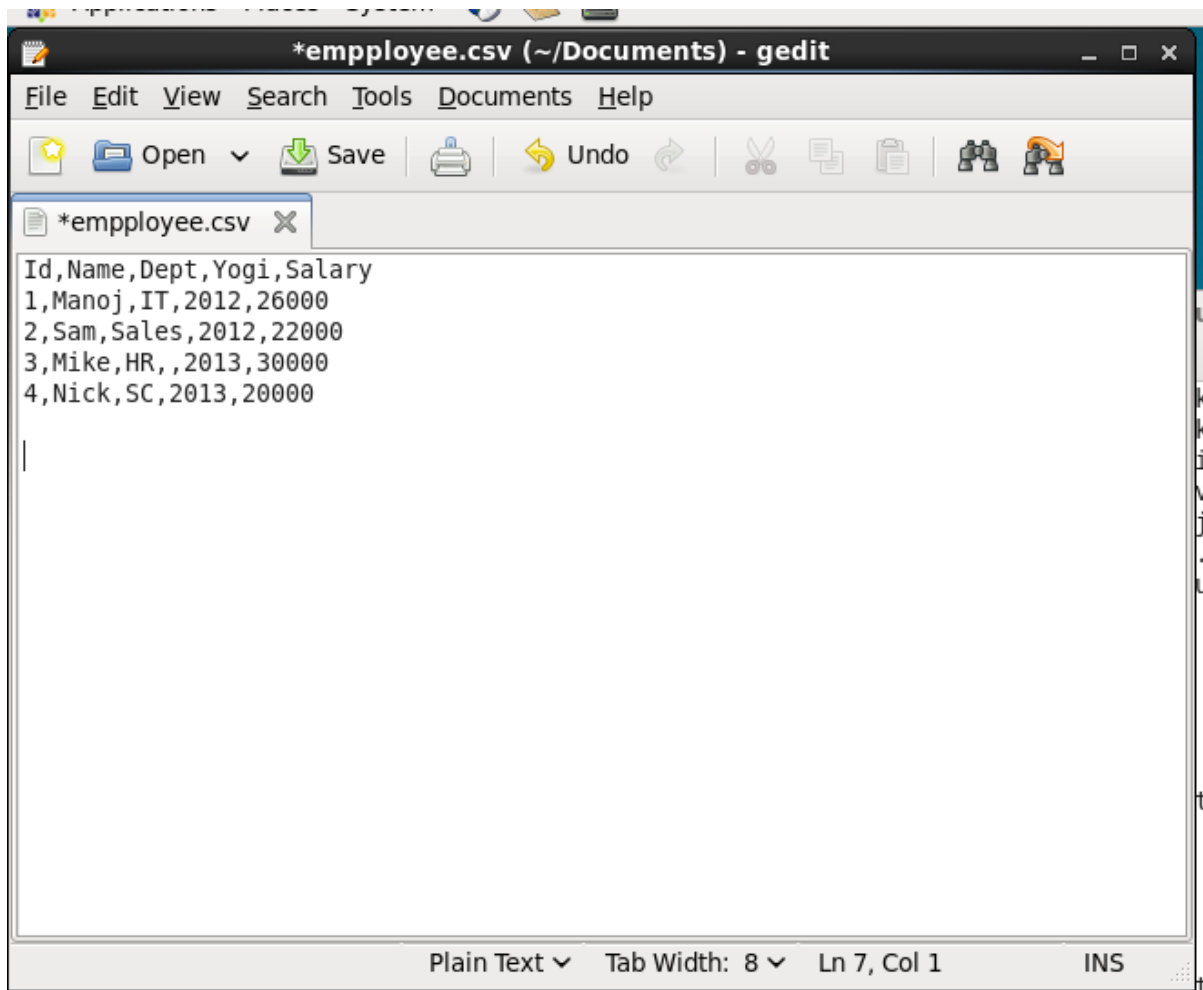
hive> create table students(ID int,Name string,Course string,Age int)
> row format delimited
> fields terminated by ','
> tblproperties("skip.header.line.count"="1");
OK
Time taken: 0.371 seconds
hive> select *from students;
OK
Time taken: 0.348 seconds
hive> load data local inpath '/home/cloudera/Documents/student.csv' into table students;
Loading data to table rjcstudent.students
Table rjcstudent.students stats: [numFiles=1, totalSize=119]
OK
Time taken: 1.826 seconds
hive> select *from students;
OK
1      Manoj   Hadoop  24
2      Ritvik  Java    23
3      Tom     Python  20
4      Jerry   Hadoop  23
5      Sam     Python  26
6      John    Java    22
Time taken: 0.185 seconds, Fetched: 6 row(s)
hive>

```

```

cloudera@quickstart:~/Documents
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ ls
cloudera-manager Downloads kerberos Public
cm_api.py dpt.java lib Templates
departments.java eclipse MaxTemp.jar Videos
Desktop enterprise-deployment.json Music workspace
Documents express-deployment.json Pictures workspace1
[cloudera@quickstart ~]$ cd Documents
[cloudera@quickstart Documents]$ ls
cloudera-manager.html css img student.csv student.csv~
[cloudera@quickstart Documents]$ gedit employee.csv

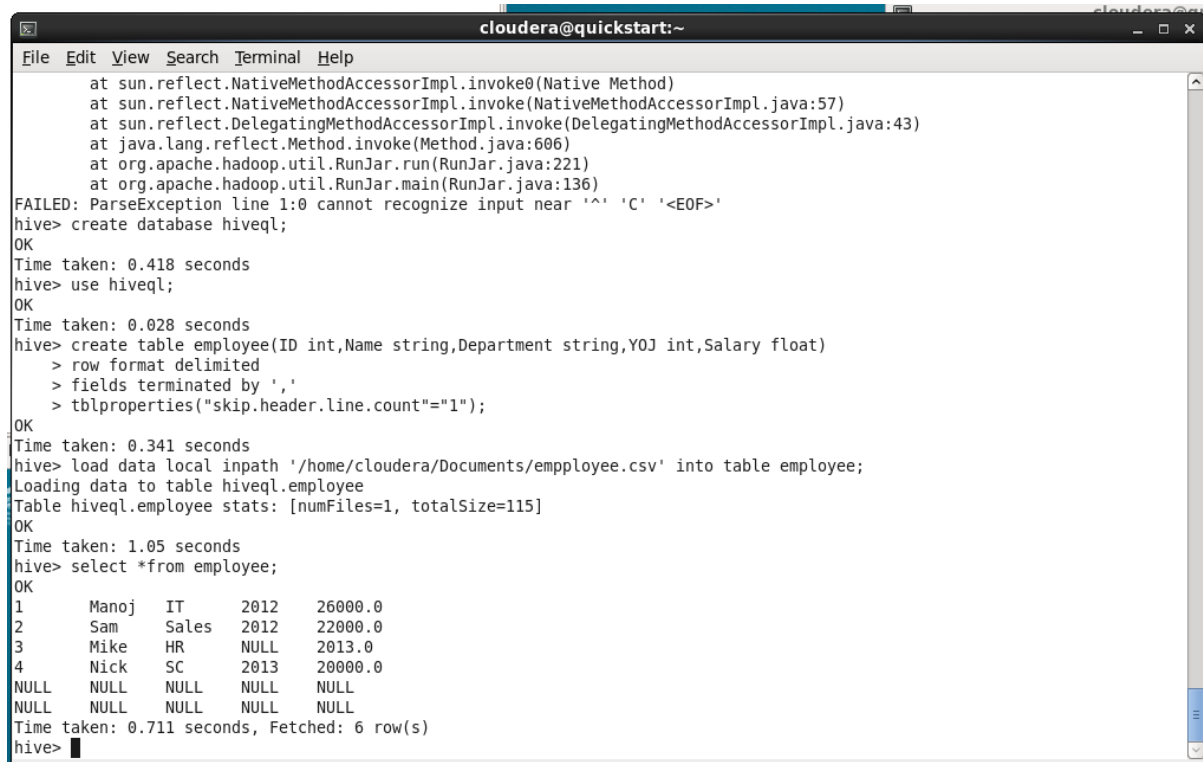
```



The screenshot shows a gedit window titled '*employee.csv (~/.Documents) - gedit'. The window contains the following text:

```
Id,Name,Dept,Yogi,Salary
1,Manoj,IT,2012,26000
2,Sam,Sales,2012,22000
3,Mike,HR,,2013,30000
4,Nick,SC,2013,20000
```

The status bar at the bottom indicates 'Plain Text', 'Tab Width: 8', 'Ln 7, Col 1', and 'INS'.



The screenshot shows a terminal window titled 'cloudera@quickstart:~'. The terminal displays the following commands and output:

```
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:606)
at org.apache.hadoop.util.RunJar.run(RunJar.java:221)
at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
FAILED: ParseException line 1:0 cannot recognize input near '^' 'C' '<EOF>'
hive> create database hiveql;
OK
Time taken: 0.418 seconds
hive> use hiveql;
OK
Time taken: 0.028 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: 0.341 seconds
hive> load data local inpath '/home/cloudera/Documents/employee.csv' into table employee;
Loading data to table hiveql.employee
Table hiveql.employee stats: [numFiles=1, totalSize=115]
OK
Time taken: 1.05 seconds
hive> select *from employee;
OK
1      Manoj      IT      2012      26000.0
2      Sam        Sales    2012      22000.0
3      Mike        HR       NULL      2013.0
4      Nick        SC       2013      20000.0
NULL   NULL         NULL     NULL      NULL
NULL   NULL         NULL     NULL      NULL
Time taken: 0.711 seconds, Fetched: 6 row(s)
hive>
```



```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
3      Mike      HR      2013      30000.0  
Time taken: 0.344 seconds, Fetched: 3 row(s)  
hive> drop table employee;  
OK  
Time taken: 1.36 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: 0.271 seconds  
hive> load data local inpath '/home/cloudera/Documents/employee.csv' into table employee;  
Loading data to table hiveql.employee  
Table hiveql.employee stats: [numFiles=1, totalSize=112]  
OK  
Time taken: 1.309 seconds  
hive> select *from employee where Salary>=25000;;  
OK  
1      Manoj      IT      2012      26000.0  
3      Mike      HR      2013      30000.0  
Time taken: 0.297 seconds, Fetched: 2 row(s)  
hive> select *from employee where Salary<25000;;  
OK  
2      Sam      Sales  2012      22000.0  
4      Nick      SC      2013      20000.0  
Time taken: 0.401 seconds, Fetched: 2 row(s)  
hive> select ID,Name,Salary+5000 from employee;  
OK  
1      Manoj      31000.0  
2      Sam      27000.0  
3      Mike      35000.0  
4      Nick      25000.0  
Time taken: 0.257 seconds, Fetched: 4 row(s)  
hive>
```

```
hive> select max(salary) from employee;  
Query ID = cloudera_20210728212727_c36fbde7-1bb3-4383-8c86-3e1a012efb28  
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_1627528366682_0001, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1627528366682_0001/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1627528366682_0001  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2021-07-28 21:29:06,429 Stage-1 map = 0%, reduce = 0%  
2021-07-28 21:29:51,375 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.27 sec  
2021-07-28 21:30:47,166 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 4.03 sec  
2021-07-28 21:30:48,565 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.19 sec  
MapReduce Total cumulative CPU time: 5 seconds 190 msec  
Ended Job = job_1627528366682_0001  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.19 sec HDFS Read: 6900 HDFS Write: 8 SUCCESS  
Total MapReduce CPU Time Spent: 5 seconds 190 msec  
OK  
30000.0  
Time taken: 199.197 seconds, Fetched: 1 row(s)  
hive>
```

```
hive> select min(salary) from employee;  
Query ID = cloudera_20210728213131_11381397-1930-4bab-8457-f4190f638d21  
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_1627528366682_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1627528366682_0002/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1627528366682_0002  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2021-07-28 21:32:29,936 Stage-1 map = 0%, reduce = 0%  
2021-07-28 21:33:00,271 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.39 sec  
2021-07-28 21:33:29,688 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.67 sec  
MapReduce Total cumulative CPU time: 5 seconds 670 msec  
Ended Job = job_1627528366682_0002  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.67 sec HDFS Read: 6931 HDFS Write: 8 SUCCESS  
Total MapReduce CPU Time Spent: 5 seconds 670 msec  
OK  
20000.0  
Time taken: 92.726 seconds, Fetched: 1 row(s)  
hive>
```

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2021-07-28 21:32:29,936 Stage-1 map = 0%, reduce = 0%  
2021-07-28 21:33:00,271 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.39 sec  
2021-07-28 21:33:29,688 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.67 sec  
MapReduce Total cumulative CPU time: 5 seconds 670 msec  
Ended Job = job_1627528366682_0002  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.67 sec HDFS Read: 6931 HDFS Write: 8 SUCCESS  
Total MapReduce CPU Time Spent: 5 seconds 670 msec  
OK  
20000.0  
Time taken: 92.726 seconds, Fetched: 1 row(s)  
hive> select ID,Name,sqrt(Salary) from employee;  
OK  
1      Manoj      161.24515496597098  
2      Sam        148.32396974191326  
3      Mike       173.20508075688772  
4      Nick       141.4213562373095  
Time taken: 1.479 seconds, Fetched: 4 row(s)  
hive> select ID,upper(Name) from employee;  
OK  
1      MANOJ  
2      SAM  
3      MIKE  
4      NICK  
Time taken: 0.682 seconds, Fetched: 4 row(s)  
hive> select ID,lower(Name) from employee;  
OK  
1      manoj  
2      sam  
3      mike  
4      nick  
Time taken: 0.806 seconds, Fetched: 4 row(s)  
hive>
```

```
employee2 (~/.Documents) - gedit  
File Edit View Search Tools Documents Help  
Open Save Undo  
employee.csv employee2  
Id,Name,Dept,YoJ,Salary,Country  
1,Manoj,IT,2012,26000,India  
2,Sam,Sales,2012,22000,USA  
3,Mike,HR,2013,30000,UK  
4,Nick,SC,2013,20000,UK  
Plain Text Tab Width: 8 Ln 5, Col 24 INS
```

```

hive> create table employee3(ID int,Name string,Department string,YOJ int,Salary float,Country string)
> row format delimited
> fields terminated by ','
> tblproperties("skip.header.line.count"="1");
OK
Time taken: 3.675 seconds
hive> load data local inpath '/home/cloudera/Documents/employee2.csv' into table employee3;
FAILED: SemanticException Line 1:23 Invalid path ''/home/cloudera/Documents/employee2.csv': No files matching path file:/home/clo
udera/Documents/employee2.csv
hive> load data local inpath '/home/cloudera/Documents/employee2' into table employee3;
FAILED: SemanticException Line 1:23 Invalid path ''/home/cloudera/Documents/employee2': No files matching path file:/home/clouder
a/Documents/employee2
hive> load data local inpath '/home/cloudera/Documents/employee2.csv' into table employee3;
FAILED: SemanticException Line 1:23 Invalid path ''/home/cloudera/Documents/employee2.csv': No files matching path file:/home/clo
udera/Documents/employee2.csv
hive> load data local inpath '/home/cloudera/Documents/employee2.csv' into table employee3;
Loading data to table hiveql.employee3
Table hiveql.employee3 stats: [numFiles=1, totalSize=135]
OK
Time taken: 1.845 seconds
hive> select *from employee3;
OK
1      Manoj  IT      2012    26000.0 India
2      Sam    Sales  2012    22000.0 USA
3      Mike   HR     2013    30000.0 UK
4      Nick   SC     2013    20000.0 UK
Time taken: 0.298 seconds, Fetched: 4 row(s)

hive> select country,sum(salary) from employee3 group by Country;
Query ID = cloudera_20210728214444_5cf8c322-786a-48ca-b63d-f64836429f6a
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_1627528366682_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1627528366682_0003/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1627528366682_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-28 21:45:13,525 Stage-1 map = 0%, reduce = 0%
2021-07-28 21:45:49,886 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.96 sec
2021-07-28 21:46:09,788 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.88 sec
MapReduce Total cumulative CPU time: 4 seconds 880 msec
Ended Job = job_1627528366682_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.88 sec HDFS Read: 7376 HDFS Write: 37 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 880 msec
OK
India    26000.0
UK       50000.0
USA      22000.0
Time taken: 95.778 seconds, Fetched: 3 row(s)
hive>

```

```

cloudera@quickstart:~
File Edit View Search Terminal Help
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.88 sec HDFS Read: 7376 HDFS Write: 37 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 880 msec
OK
India    26000.0
UK       50000.0
USA      22000.0
Time taken: 95.778 seconds, Fetched: 3 row(s)
hive> select country,sum(salary) from employee3 group by Country having sum(salary)>50000;
Query ID = cloudera_20210728214747_1b8481d8-eb26-4bab-87c0-f712edffa4a0
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_1627528366682_0004, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1627528366682_0004/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1627528366682_0004
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-28 21:48:22,788 Stage-1 map = 0%, reduce = 0%
2021-07-28 21:48:52,136 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.36 sec
2021-07-28 21:49:31,643 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 5.69 sec
2021-07-28 21:49:32,739 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.36 sec
MapReduce Total cumulative CPU time: 6 seconds 360 msec
Ended Job = job_1627528366682_0004
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.36 sec HDFS Read: 7814 HDFS Write: 0 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 360 msec
OK
Time taken: 108.02 seconds
hive>

```

```

hive> Select *from employee3 order by salary desc;
Query ID = cloudera_20210728215151_7e431daf-83cf-4e6b-b78f-dcef4eecfe80
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_1627528366682_0005, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1627528366682_0005/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1627528366682_0005
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-28 21:52:07,837 Stage-1 map = 0%, reduce = 0%
2021-07-28 21:52:27,713 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.26 sec
2021-07-28 21:52:54,269 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.37 sec
MapReduce Total cumulative CPU time: 5 seconds 370 msec
Ended Job = job_1627528366682_0005
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.37 sec HDFS Read: 7241 HDFS Write: 111 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 370 msec
OK
3      Mike      HR      2013      30000.0 UK
1      Manoj     IT      2012      26000.0 India
2      Sam       Sales   2012      22000.0 USA
4      Nick      SC      2013      20000.0 UK
Time taken: 72.532 seconds, Fetched: 4 row(s)
hive> █

```

```

hive> Select *from employee3 sort by salary desc;
Query ID = cloudera_20210728215353_14e4795f-cb4c-4a8d-ba55-8ab46d94596d
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_1627528366682_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1627528366682_0006/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1627528366682_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2021-07-28 21:54:20,425 Stage-1 map = 0%, reduce = 0%
2021-07-28 21:54:38,159 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.33 sec
2021-07-28 21:55:05,667 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 5.42 sec
MapReduce Total cumulative CPU time: 5 seconds 420 msec
Ended Job = job_1627528366682_0006
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.42 sec HDFS Read: 7241 HDFS Write: 111 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 420 msec
OK
3      Mike      HR      2013      30000.0 UK
1      Manoj     IT      2012      26000.0 India
2      Sam       Sales   2012      22000.0 USA
4      Nick      SC      2013      20000.0 UK
Time taken: 72.148 seconds, Fetched: 4 row(s)
hive> █

```