cloudera@quickstart:~/hiveclass

```
[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> exit
    > ;
WARN: The method class org.apache.commons.logging.impl.SLF4JLogFactory#release()
 was invoked.
WARN: Please see http://www.slf4j.org/codes.html#release for an explanation.
[cloudera@quickstart ~]$ hadoop fs -ls
[cloudera@quickstart ~]$ ls
cloudera-manager  Downloads                 kerberos  Pictures   workspace
cm_api.py         eclipse                   lib       Public
Desktop           enterprise-deployment.json  Music     Templates
Documents         express-deployment.json     parcels   Videos
[cloudera@quickstart ~]$ pwd
/home/cloudera
[cloudera@quickstart ~]$ ls
cloudera-manager  Downloads                 hiveclass  parcels    Videos
cm_api.py         eclipse                   kerberos   Pictures   workspace
Desktop           enterprise-deployment.json  lib        Public
Documents         express-deployment.json     Music      Templates
[cloudera@quickstart ~]$ cd hiveclass
[cloudera@quickstart hiveclass]$ ls
sales_order_data.csv
[cloudera@quickstart hiveclass]$ hive
```

8:14 PM

Store raw data into hdfs location

cloudera@quickstart:~/hiveclass

```
[cloudera@quickstart hiveclass]$ hadoop fs -ls
Found 1 items
drwxr-xr-x   - cloudera cloudera          0 2022-09-11 07:49 hiveclass
[cloudera@quickstart hiveclass]$ ls
sales_order_data.csv
[cloudera@quickstart hiveclass]$ pwd
/home/cloudera/hiveclass
[cloudera@quickstart hiveclass]$ hadoop fs -mkdir hiveclasscsv
[cloudera@quickstart hiveclass]$ hadoop fs -ls
Found 2 items
drwxr-xr-x   - cloudera cloudera          0 2022-09-11 07:49 hiveclass
drwxr-xr-x   - cloudera cloudera          0 2022-09-11 08:08 hiveclasscsv
[cloudera@quickstart hiveclass]$ hadoop fs -copyFromLocal /home/cloudera/hiveclass/sales_order_data.csv hiveclass
copyFromLocal: `hiveclass/sales_order_data.csv': File exists
[cloudera@quickstart hiveclass]$ hadoop fs -copyFromLocal hiveclass/sales_order_data.csv hiveclass
copyFromLocal: `hiveclass/sales_order_data.csv': No such file or directory
[cloudera@quickstart hiveclass]$ hadoop fs -copyFromLocal /hiveclass/sales_order_data.csv hiveclasscsv
copyFromLocal: `/hiveclass/sales_order_data.csv': No such file or directory
[cloudera@quickstart hiveclass]$ hadoop fs -ls /
Found 6 items
drwxrwxrwx   - hdfs  supergroup          0 2017-10-23 09:15 /benchmarks
drwxr-xr-x   - hbase supergroup          0 2022-09-11 07:27 /hbase
drwxr-xr-x   - solr  solr                0 2017-10-23 09:18 /solr
drwxrwxrwt   - hdfs  supergroup          0 2022-08-23 12:19 /tmp
drwxr-xr-x   - hdfs  supergroup          0 2017-10-23 09:17 /user
drwxr-xr-x   - hdfs  supergroup          0 2017-10-23 09:17 /var
[cloudera@quickstart hiveclass]$ cd /home/cloudera/hiveclass
[cloudera@quickstart hiveclass]$ ls
sales_order_data.csv
[cloudera@quickstart hiveclass]$ hadoop fs -copyFromLocal sales_order_data.csv hiveclass
copyFromLocal: `hiveclass/sales_order_data.csv': File exists
[cloudera@quickstart hiveclass]$ pwd
/home/cloudera/hiveclass
[cloudera@quickstart hiveclass]$ ls
sales_order_data.csv
[cloudera@quickstart hiveclass]$ hadoop fs -copyFromLocal /home/cloudera/hiveclass/sales_order_data.csv  /hiveclass
[cloudera@quickstart hiveclass]$ hive

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive>
```

Create a internal hive table "sales_order_csv" which will store csv data sales_order_csv .. make sure to skip header row while creating table

Load data from hdfs path into "sales_order_csv"

cloudera@quickstart:~/hiveclass

```
        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' 'database' '<EOF>' in ddl statement
hive> show databases;
OK
default
monty
Time taken: 1.232 seconds, Fetched: 2 row(s)
hive> load data local inpath '/home/cloudera/hiveclass/sales_order_data.csv' into table sales_order_csv;
FAILED: SemanticException [Error 10001]: Line 1:82 Table not found 'sales_order_csv'
hive> load data local inpath '/hiveclass/sales_order_data.csv' into table sales_order_csv;
FAILED: SemanticException [Error 10001]: Line 1:68 Table not found 'sales_order_csv'
hive> CREATE TABLE IF NOT EXISTS  monty.sales_order_csv
    > (ORDERNUMBER   int,
    > QUANTITYORDERED   int,
    > PRICEEACH   float,
    > ORDERLINENUMBER   int,
    > SALES   int,
    > STATUS   string,
    > QTR_ID   int,
    > MONTH_ID   int,
    > YEAR_ID   int,
    > PRODUCTLINE   string,
    > MSRP   string,
    > PRODUCTCODE   string,
    > PHONE   string,
    > CITY   string,
    > STATE   string,
    > POSTALCODE   string,
    > COUNTRY   string,
    > TERRITORY   string,
    > CONTACTLASTNAME   string,
    > CONTACTFIRSTNAME   string,
    > DEALSIZE   string)
    > ROW FORMAT DELIMITED
    > FIELDS TERMINATED BY ','
    > TBLPROPERTIES("skip.header.line.count"="1");
OK
Time taken: 0.769 seconds
hive> load data local inpath '/hiveclass/sales_order_data.csv' into table   monty.sales_order_csv;
FAILED: SemanticException Line 1:23 Invalid path ''/hiveclass/sales_order_data.csv'': No files matching path file:/hiveclass/sales_order_data.csv
hive> load data local inpath '/home/cloudera/hiveclass/sales_order_data.csv' into table monty.sales_order_csv;
Loading data to table monty.sales_order_csv
Table monty.sales_order_csv stats: [numFiles=1, totalSize=360233]
OK
Time taken: 2.621 seconds
hive>
```

Load data from hdfs path into "sales_order_csv"

Load data from "sales_order_csv" into "sales_order_orc"

```
hive> Insert into table monty.sales_order_orc select * from monty.sales_order_csv;
Query ID = cloudera_20220911082929_e28fe1bd-7d26-4a9a-b9b1-a57048a19ffe
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1662906115514_0001, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662906115514_0001/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662906115514_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2022-09-11 08:30:26,455 Stage-1 map = 0%,  reduce = 0%
2022-09-11 08:30:56,028 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 4.81 sec
MapReduce Total cumulative CPU time: 4 seconds 810 msec
Ended Job = job_1662906115514_0001
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/warehouse/monty.db/sales_order_orc/.hive-staging_hive_2022-09-11_08-29-38_054_2550515385589303670-1/-ext-1
Loading data to table monty.sales_order_orc
Table monty.sales_order_orc stats: [numFiles=1, numRows=2823, totalSize=33441, rawDataSize=3384777]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1   Cumulative CPU: 4.81 sec   HDFS Read: 367268 HDFS Write: 33525 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 810 msec
OK
Time taken: 83.005 seconds
hive>
```

Checked count of "sales_order_csv" "sales_order_orc"

cloudera@quickstart:~/hiveclass

```
Query ID = cloudera_20220911083535_82b89e14-4db3-4cbb-89cf-2ea4cd494dc9
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_1662906115514_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application_
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662906115514_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-11 08:35:39,536 Stage-1 map = 0%,  reduce = 0%
2022-09-11 08:35:59,805 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 3.03 sec
2022-09-11 08:36:11,411 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.6 sec
MapReduce Total cumulative CPU time: 4 seconds 600 msec
Ended Job = job_1662906115514_0002
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.6 sec   HDFS Read: 370242 HDFS Write: 5 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 600 msec
OK
2823
Time taken: 60.907 seconds, Fetched: 1 row(s)
hive> select count(*)from  monty.sales_order_orc;
Query ID = cloudera_20220911083737_706ce679-aac3-45a3-809c-8980f3e23149
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_1662906115514_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662906115514_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-11 08:38:07,789 Stage-1 map = 0%,  reduce = 0%
2022-09-11 08:38:17,173 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.36 sec
2022-09-11 08:38:31,470 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.34 sec
MapReduce Total cumulative CPU time: 3 seconds 340 msec
Ended Job = job_1662906115514_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.34 sec   HDFS Read: 26667 HDFS Write: 5 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 340 msec
OK
2823
Time taken: 37.873 seconds, Fetched: 1 row(s)
hive>
```

Calculate total sales per year

```
Time taken: 37.873 seconds, Fetched: 1 row(s)
hive> select YEAR_ID, sum(SALES) TOTAL_SALES monty.sales_order_orc
    > group  by YEAR_ID;
FAILED: ParseException line 1:39 missing EOF at 'monty' near 'TOTAL_SALES'
hive> select YEAR_ID, sum(SALES) TOTAL_SALES from monty.sales_order_orc
    > group  by YEAR_ID;
Query ID = cloudera_20220911084141_d1a43427-7c31-4063-9062-308d4eb05ac8
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_1662906115514_0004, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662906115514_0004/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662906115514_0004
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-11 08:41:37,865 Stage-1 map = 0%,  reduce = 0%
2022-09-11 08:41:46,732 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.22 sec
2022-09-11 08:41:59,071 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.58 sec
MapReduce Total cumulative CPU time: 3 seconds 580 msec
Ended Job = job_1662906115514_0004
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.58 sec   HDFS Read: 32526 HDFS Write: 39 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 580 msec
OK
2003    3516514
2004    4723531
2005    1791264
```

Find a product for which maximum orders were placed

```
hive> SELECT productline
    > FROM
    >  (SELECT productline,
    >          rank() over (order by cast(quantityordered as int) desc) as r
    >   FROM monty.sales_order_orc) S
    > WHERE S.r = 1;
Query ID = cloudera_20220911090000_a435a294-54c8-4bde-aef0-8809a3e478a7
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Defaulting to jobconf value of: 3
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_1662906115514_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662906115514_0006/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662906115514_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 3
2022-09-11 09:01:02,135 Stage-1 map = 0%,  reduce = 0%
2022-09-11 09:01:13,277 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.71 sec
2022-09-11 09:01:44,681 Stage-1 map = 100%,  reduce = 22%, Cumulative CPU 4.12 sec
2022-09-11 09:01:48,578 Stage-1 map = 100%,  reduce = 33%, Cumulative CPU 5.24 sec
2022-09-11 09:01:50,807 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 10.08 sec
MapReduce Total cumulative CPU time: 10 seconds 80 msec
Ended Job = job_1662906115514_0006
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 3   Cumulative CPU: 10.08 sec   HDFS Read: 46821 HDFS Write: 13 SUCCESS
Total MapReduce CPU Time Spent: 10 seconds 80 msec
OK
Classic Cars
Time taken: 61.61 seconds, Fetched: 1 row(s)
```

Another approach

```
hive> SELECT a.productline,a.quantityordered FROM monty.sales_order_orc a left semi join
    > (SELECT MAX(quantityordered)
    > max_o FROM monty.sales_order_orc) b on (a.quantityordered=b.max_o);
Query ID = cloudera_20220911090606_7360d036-f5aa-4fb5-aac2-cbde353caa3c
Total jobs = 3
Launching Job 1 out of 3
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_1662906115514_0007, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662906115514_0007/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662906115514_0007
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-11 09:06:49,077 Stage-1 map = 0%,  reduce = 0%
2022-09-11 09:06:59,083 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.46 sec
2022-09-11 09:07:11,611 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.56 sec
MapReduce Total cumulative CPU time: 3 seconds 560 msec
Ended Job = job_1662906115514_0007
Stage-6 is selected by condition resolver.
Stage-2 is filtered out by condition resolver.
Execution log at: /tmp/cloudera/cloudera_20220911090606_7360d036-f5aa-4fb5-aac2-cbde353caa3c.log
2022-09-11 09:07:21     Starting to launch local task to process map join;      maximum memory = 1013645312
2022-09-11 09:07:23     Dump the side-table for tag: 1 with group count: 1 into file: file:/tmp/cloudera/4ca224ff-60d6-4290-b213-12eb17290340/hive_2022-09-11_09-06-3
-1/-local-10004/HashTable-Stage-4/MapJoin-mapfile01--.hashtable
2022-09-11 09:07:23     Uploaded 1 File to: file:/tmp/cloudera/4ca224ff-60d6-4290-b213-12eb17290340/hive_2022-09-11_09-06-35_806_601601685006130335-1/-local-10004/Ha
mapfile01--.hashtable (278 bytes)
2022-09-11 09:07:23     End of local task; Time Taken: 1.919 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 3 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1662906115514_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662906115514_0008/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662906115514_0008
Hadoop job information for Stage-4: number of mappers: 1; number of reducers: 0
2022-09-11 09:07:38,132 Stage-4 map = 0%,  reduce = 0%
2022-09-11 09:07:49,352 Stage-4 map = 100%,  reduce = 0%, Cumulative CPU 1.96 sec
MapReduce Total cumulative CPU time: 1 seconds 960 msec
Ended Job = job_1662906115514_0008
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.56 sec   HDFS Read: 28375 HDFS Write: 114 SUCCESS
Stage-Stage-4: Map: 1   Cumulative CPU: 1.96 sec   HDFS Read: 26640 HDFS Write: 16 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 520 msec
OK
Classic Cars    97
Time taken: 74.619 seconds, Fetched: 1 row(s)
```

Calculate the total sales for each quarter

```
hive> select qtr_id, sum(SALES) TOTAL_SALES from monty.sales_order_orc
    > group  by qtr_id;
Query ID = cloudera_20220911104040_398d6fe5-8d30-48ec-8d6f-2014ecf11689
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Defaulting to jobconf value of: 3
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_1662906115514_0010, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1662906115514_0010/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662906115514_0010
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 3
2022-09-11 10:40:58,329 Stage-1 map = 0%,  reduce = 0%
2022-09-11 10:41:18,134 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.99 sec
2022-09-11 10:42:16,487 Stage-1 map = 100%,  reduce = 33%, Cumulative CPU 5.98 sec
2022-09-11 10:42:17,933 Stage-1 map = 100%,  reduce = 67%, Cumulative CPU 9.14 sec
2022-09-11 10:42:19,306 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 12.01 sec
MapReduce Total cumulative CPU time: 12 seconds 650 msec
Ended Job = job_1662906115514_0010
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 3   Cumulative CPU: 12.65 sec   HDFS Read: 42375 HDFS Write: 40 SUCCESS
Total MapReduce CPU Time Spent: 12 seconds 650 msec
OK
3       1758673
1       2350510
4       3874271
2       2047855
Time taken: 109.229 seconds, Fetched: 4 row(s)
hive>
```

d. In which quarter sales was minimum

```
Time taken: 232.632 seconds, Fetched: 1 row(s)
hive> SET mapreduce.job.reduces=3;
hive> SELECT qtr_id, TOTAL_SALES FROM (
    > select qtr_id, sum(SALES) TOTAL_SALES from monty.sales_order_orc
    > group  by qtr_id
    > ) M
    > SORT BY TOTAL_SALES asc
    >  LIMIT 1;
Query ID = cloudera_20220911111616_a2fa7a7c-ee04-4f6a-a3ab-4c00b33dfdb2
Total jobs = 3
Launching Job 1 out of 3
```

```
Starting Job = job_1662906115514_0025, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1662906115514_0025
Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 1
2022-09-11 11:20:01,366 Stage-3 map = 0%,  reduce = 0%
2022-09-11 11:20:18,959 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 1.98 sec
2022-09-11 11:20:39,833 Stage-3 map = 100%,  reduce = 100%, Cumulative CPU 5.7 sec
MapReduce Total cumulative CPU time: 5 seconds 700 msec
Ended Job = job_1662906115514_0025
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 3   Cumulative CPU: 9.92 sec   HDFS Read: 39662 HDFS Write: 376 SUCCESS
Stage-Stage-2: Map: 1  Reduce: 3   Cumulative CPU: 10.1 sec   HDFS Read: 9130 HDFS Write: 332 SUCCESS
Stage-Stage-3: Map: 1  Reduce: 1   Cumulative CPU: 5.7 sec   HDFS Read: 5759 HDFS Write: 10 SUCCESS
Total MapReduce CPU Time Spent: 25 seconds 720 msec
OK
3       1758673
```

```
Total MapReduce CPU Time Spent: 1 minutes 15 seconds 570 msec
OK
2003    1       1357
2003    4       1993
2003    7       1725
2003    10      5515
2004    3       1978
2004    6       2971
2004    9       3171
2004    12      3804
2005    2       3393
2005    5       4357
2003    2       1449
2003    5       2017
2003    8       1974
2003    11      10179
2004    1       3245
2004    4       2077
2004    7       3174
2004    10      5483
2005    3       3852
2003    3       1755
2003    6       1649
2003    9       2510
2003    12      2489
2004    2       3061
2004    5       2618
2004    8       4564
2004    11      10678
2005    1       3395
2005    4       2634
Time taken: 761.147 seconds, Fetched: 29 row(s)
```

Find a month for each year in which maximum number of quantities were sold

```
hive> select distinct a.year_id,a.month_id,a.quantityordered from monty.sales_order_orc a
    > inner join
    > (select  year_id, month_id, max(quantityordered) as quantityordered  from monty.sales_order_orc group  by year_id,month_id) b
    > on (a.year_id=b.year_id and a.month_id=b.month_id and a.quantityordered=b.quantityordered);
Total MapReduce CPU Time Spent: 19 seconds 130 msec
OK
2003    2       50
2003    5       50
2003    11      50
2003    12      49
2004    1       50
2004    7       50
2004    10      50
2004    11      55
2005    3       50
2005    4       97
2003    3       50
2003    6       50
2003    7       49
2003    9       50
2004    2       50
2004    5       50
2004    8       50
2005    1       50
2005    5       70
2003    1       50
2003    4       50
2003    8       49
2003    10      50
2004    3       50
2004    4       49
2004    6       50
2004    9       50
2004    12      50
2005    2       50
```