Exercise -6

BDMA LAB6: Analysis the supermartket sales data using HQL

Annapoornima S
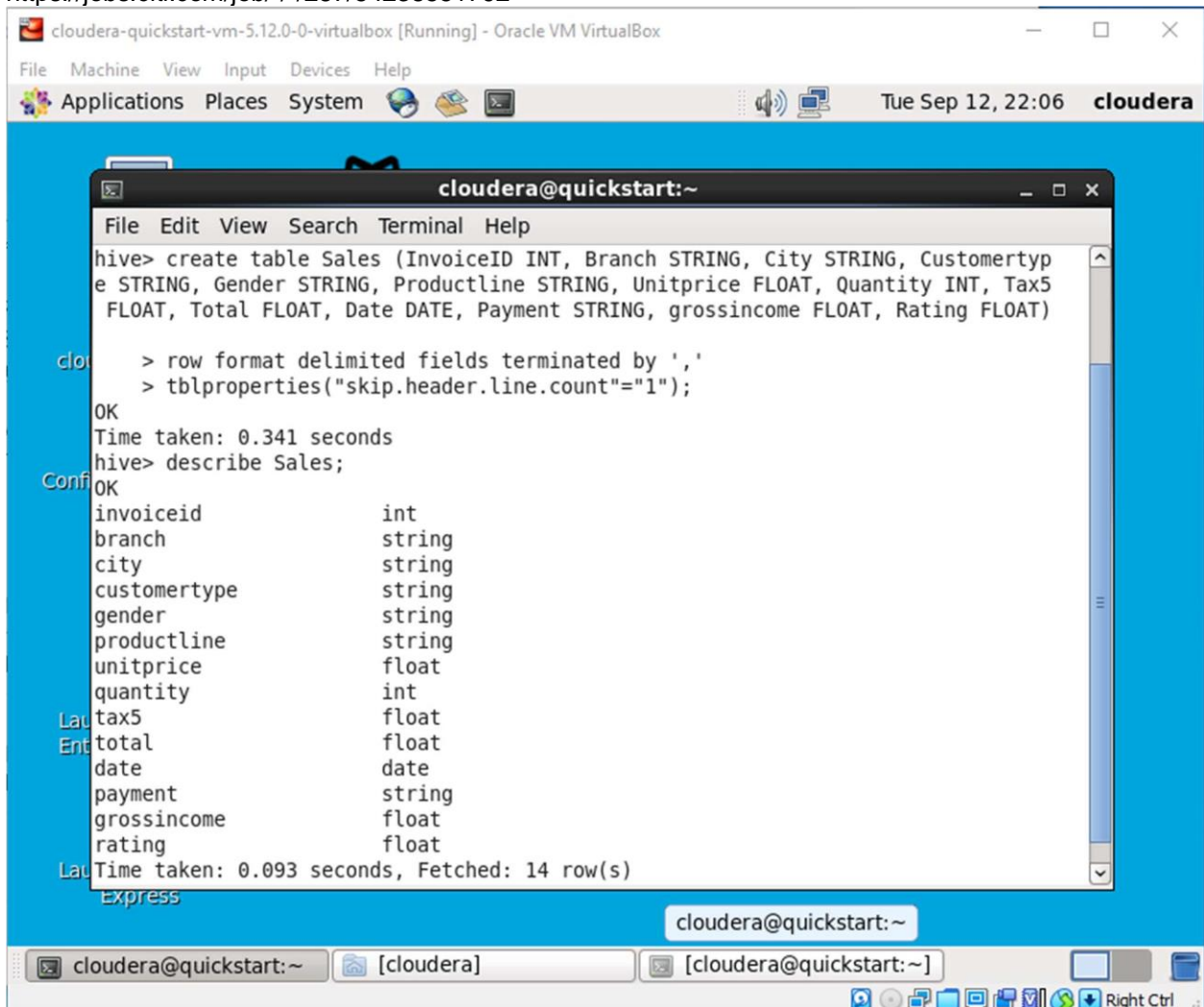## (225229101)

```
cloudera@quickstart:~                                    _  □  ✕
File  Edit  View  Search  Terminal  Help
[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: 0.337 seconds, Fetched: 1 row(s)
hive>
```

```
                              cloudera@quickstart:~                    _  □  ✕

File  Edit  View  Search  Terminal  Help

[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: 0.337 seconds, Fetched: 1 row(s)
hive> show tables;
OK
customer
emp
order
orders
sales
sri
Time taken: 0.036 seconds, Fetched: 6 row(s)
hive>
```

cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

Applications Places System 🌐 ✉ ▣   🔊 💻   Tue Sep 12, 22:06   **cloudera**

cloudera@quickstart:~

File Edit View Search Terminal Help

```
hive> create table Sales (InvoiceID INT, Branch STRING, City STRING, Customertyp
e STRING, Gender STRING, Productline STRING, Unitprice FLOAT, Quantity INT, Tax5
 FLOAT, Total FLOAT, Date DATE, Payment STRING, grossincome FLOAT, Rating FLOAT)

    > row format delimited fields terminated by ','
    > tblproperties("skip.header.line.count"="1");
OK
Time taken: 0.341 seconds
hive> describe Sales;
OK
invoiceid               int
branch                  string
city                    string
customertype            string
gender                  string
productline             string
unitprice               float
quantity                int
tax5                    float
total                   float
date                    date
payment                 string
grossincome             float
rating                  float
Time taken: 0.093 seconds, Fetched: 14 row(s)
```

cloudera@quickstart:~

▣ cloudera@quickstart:~   🏠 [cloudera]   ▣ [cloudera@quickstart:~]

Right Ctrl

6 & 7

File   Edit   View   Search   Terminal   Help

```
hive> describe sales;
OK
invoiceid               int
branch                  string
city                    string
customertype            string
gender                  string
productline             string
unitprice               float
quantity                int
tax5                    float
total                   float
date                    date
payment                 string
grossincome             float
rating                  float
Time taken: 0.046 seconds, Fetched: 14 row(s)
hive> LOAD DATA LOCAL INPATH '/home/cloudera/supermarket_sales.csv' INTO TABLE s
ales;
Loading data to table default.sales
Table default.sales stats: [numFiles=1, totalSize=107063]
OK
Time taken: 1.074 seconds
hive>
```

8

File   Edit   View   Search   Terminal   Help

```
NULL      B        Mandalay          Member    Male       Health and beauty          75.37    8
30.148   633.108  NULL       Credit card       30.148    8.4
NULL      A        Yangon    Normal  Female    Food and beverages        56.56    5         1
4.14     296.94   NULL       Credit card       14.14    4.5
NULL      B        Mandalay          Normal    Female    Sports and travel          76.6     1
0        38.3     804.3      NULL    Ewallet 38.3       6.0
NULL      A        Yangon    Normal  Male       Electronic accessories  58.03    2         5
.803     121.863  NULL       Ewallet 5.803     8.8
NULL      B        Mandalay          Normal    Male       Fashion accessories        17.49    1
0        8.745    183.645 NULL       Ewallet 8.745     6.6
NULL      C        Naypyitaw         Member    Female    Electronic accessories  60.95    1
3.0475   63.9975  NULL       Ewallet 3.0475    5.9
NULL      C        Naypyitaw         Normal    Male       Health and beauty          40.35    1
2.0175   42.3675  NULL       Ewallet 2.0175    6.2
NULL      B        Mandalay          Normal    Female    Home and lifestyle         97.38    1
0        48.69    1022.49 NULL       Ewallet 48.69     4.4
NULL      A        Yangon    Member  Male       Food and beverages        31.84    1         1
.592     33.432   NULL       Cash    1.592     7.7
NULL      A        Yangon    Normal  Male       Home and lifestyle         65.82    1         3
.291     69.111   NULL       Cash    3.291     4.1
NULL      A        Yangon    Member  Female    Fashion accessories        88.34    7         3
0.919    649.299  NULL       Cash    30.919    6.6
Time taken: 0.049 seconds, Fetched: 1000 row(s)
hive>
```

Q01
SELECT *FROMsalesWHERE Payment='Creditcard';

```
cloudera@quickstart:~                                          _ □ ✕

File  Edit  View  Search  Terminal  Help

NULL     B        Mandalay          Normal   Male      Food and beverages      33.33   2
3.333    69.993   NULL      Credit card      3.333    6.4
NULL     B        Mandalay          Normal   Female    Electronic accessories  38.27   2
3.827    80.367   NULL      Credit card      3.827    5.8
NULL     A        Yangon    Member   Male      Home and lifestyle      81.01    3        1
2.1515   255.1815           NULL      Credit card      12.1515 9.3
NULL     B        Mandalay          Member   Female    Electronic accessories  34.49   5
8.6225   181.0725           NULL      Credit card      8.6225  9.0
NULL     B        Mandalay          Member   Female    Food and beverages      84.63   1
0        42.315   888.615 NULL       Credit card      42.315  9.0
NULL     B        Mandalay          Member   Male      Fashion accessories     49.92   2
4.992    104.832 NULL      Credit card      4.992    7.0
NULL     B        Mandalay          Normal   Female    Electronic accessories  25.45   1
1.2725   26.7225 NULL      Credit card      1.2725   5.1
NULL     B        Mandalay          Normal   Female    Food and beverages      67.77   1
3.3885   71.1585 NULL      Credit card      3.3885   6.5
NULL     B        Mandalay          Member   Male      Health and beauty       62.0    8
24.8     520.8    NULL      Credit card      24.8     6.2
NULL     B        Mandalay          Member   Male      Health and beauty       75.37   8
30.148   633.108 NULL      Credit card      30.148   8.4
NULL     A        Yangon    Normal   Female    Food and beverages      56.56    5        1
4.14     296.94   NULL      Credit card      14.14    4.5
Time taken: 0.121 seconds, Fetched: 311 row(s)
hive> █
```
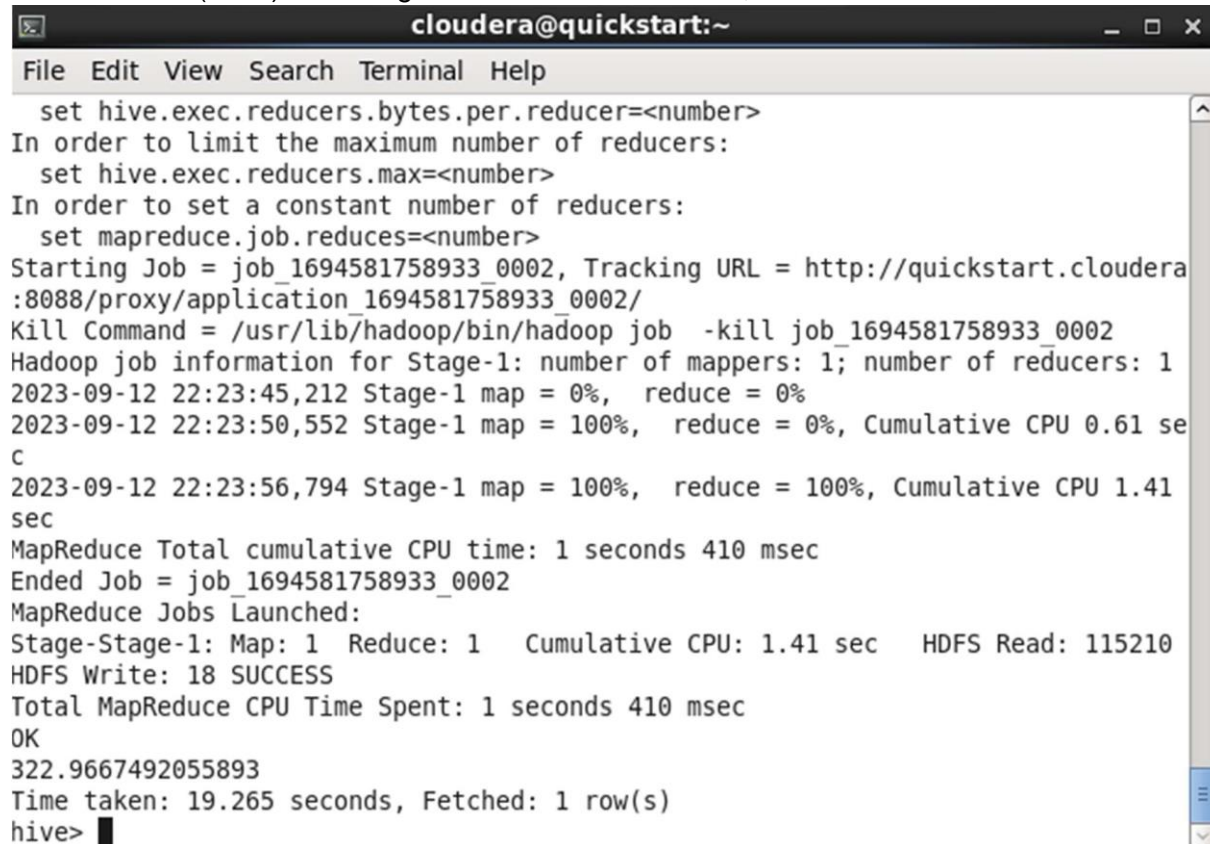
Q02

SELECT InvoiceIDFROMsalesWHERERating>9;

```
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
NULL
Time taken: 0.074 seconds, Fetched: 151 row(s)
hive>
```
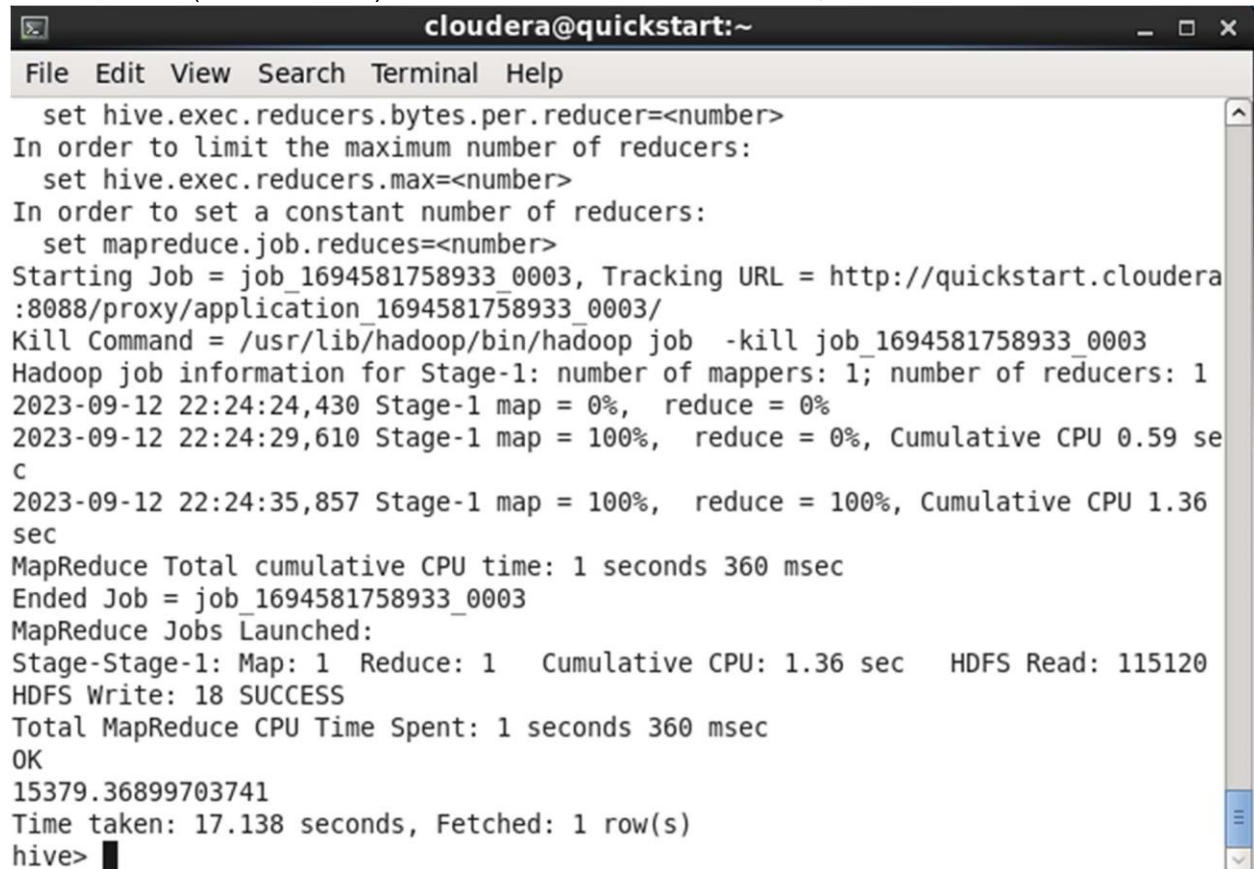
Q03
SELECT AVG(Total)ASAverageSaleAmountFROMsales;

```
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_1694581758933_0002, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1694581758933_0002/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1694581758933_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-09-12 22:23:45,212 Stage-1 map = 0%,  reduce = 0%
2023-09-12 22:23:50,552 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.61 se
c
2023-09-12 22:23:56,794 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.41
sec
MapReduce Total cumulative CPU time: 1 seconds 410 msec
Ended Job = job_1694581758933_0002
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.41 sec   HDFS Read: 115210
HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 410 msec
OK
322.9667492055893
Time taken: 19.265 seconds, Fetched: 1 row(s)
hive>
```
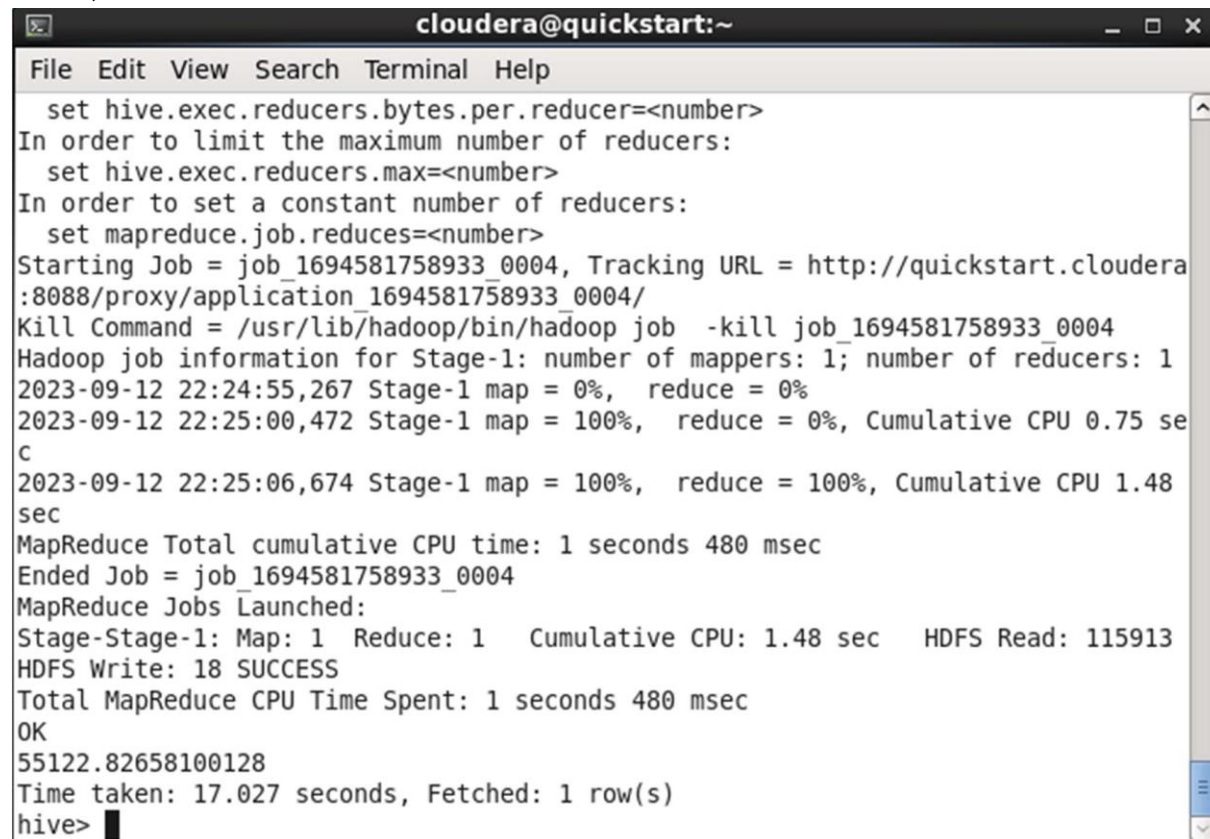
Q04

SELECTSUM(GrossIncome)ASTotalGrossIncomeFROMsales;

```
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_1694581758933_0003, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1694581758933_0003/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1694581758933_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-09-12 22:24:24,430 Stage-1 map = 0%,   reduce = 0%
2023-09-12 22:24:29,610 Stage-1 map = 100%,   reduce = 0%, Cumulative CPU 0.59 se
c
2023-09-12 22:24:35,857 Stage-1 map = 100%,   reduce = 100%, Cumulative CPU 1.36
sec
MapReduce Total cumulative CPU time: 1 seconds 360 msec
Ended Job = job_1694581758933_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.36 sec    HDFS Read: 115120
HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 360 msec
OK
15379.36899703741
Time taken: 17.138 seconds, Fetched: 1 row(s)
hive>
```
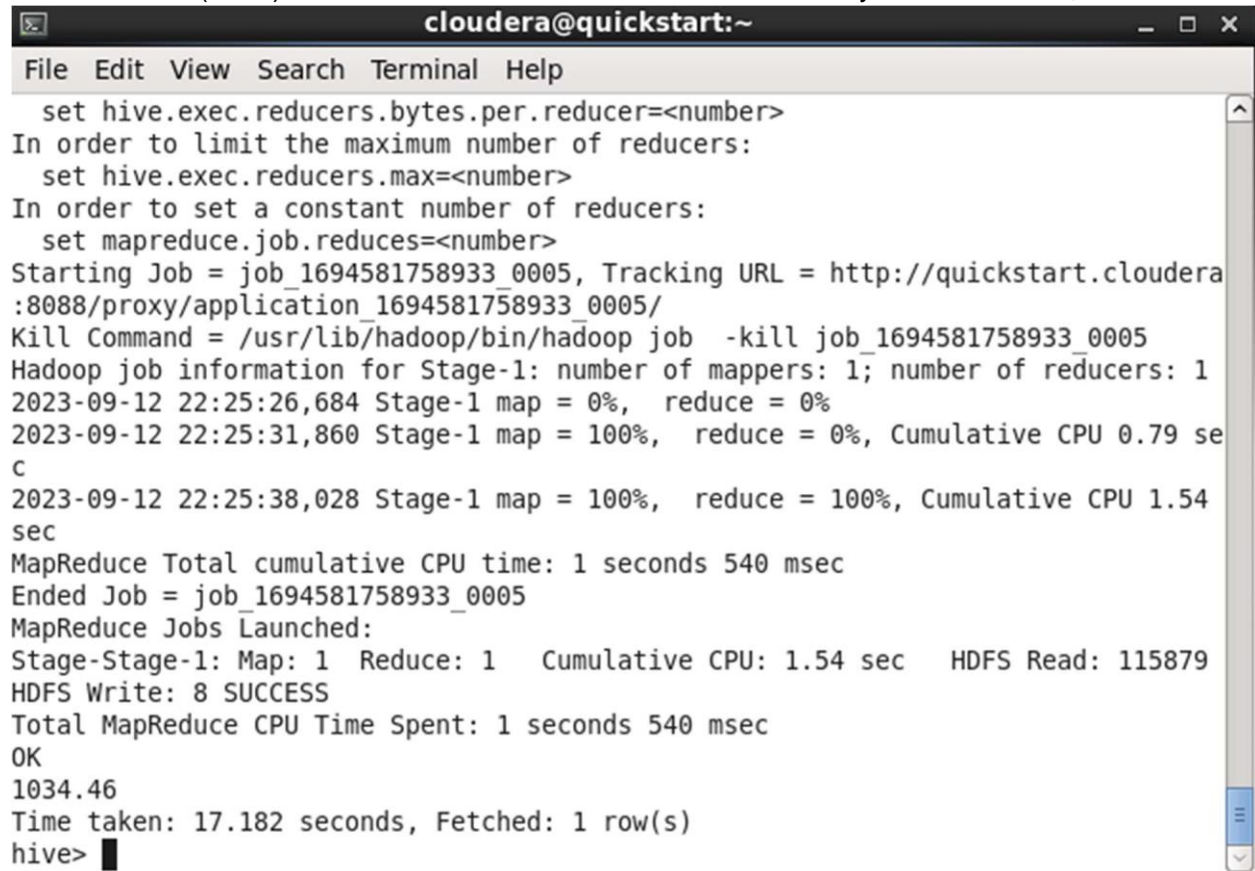
Q05

SELECT SUM(Total)ASTotalSaleAmountFROMsalesWHEREProductLine='Sportsand travel';

```
 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_1694581758933_0004, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1694581758933_0004/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1694581758933_0004
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-09-12 22:24:55,267 Stage-1 map = 0%,  reduce = 0%
2023-09-12 22:25:00,472 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.75 se
c
2023-09-12 22:25:06,674 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.48
sec
MapReduce Total cumulative CPU time: 1 seconds 480 msec
Ended Job = job_1694581758933_0004
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.48 sec   HDFS Read: 115913
HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 480 msec
OK
55122.82658100128
Time taken: 17.027 seconds, Fetched: 1 row(s)
hive>
```

Q06
SELECT MAX(Total)ASMaxSalesAmountFROMsalesWHERE Payment= 'Ewallet';

```
cloudera@quickstart:~                                    _  □  ✕

File  Edit  View  Search  Terminal  Help

  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_1694581758933_0005, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1694581758933_0005/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1694581758933_0005
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-09-12 22:25:26,684 Stage-1 map = 0%,  reduce = 0%
2023-09-12 22:25:31,860 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.79 se
c
2023-09-12 22:25:38,028 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.54
sec
MapReduce Total cumulative CPU time: 1 seconds 540 msec
Ended Job = job_1694581758933_0005
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.54 sec    HDFS Read: 115879
HDFS Write: 8 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 540 msec
OK
1034.46
Time taken: 17.182 seconds, Fetched: 1 row(s)
hive> ▊
```

Q07
SELECT COUNT(*)ASNumberOfSalesFROMsalesWHEREQuantity>5;

```
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_1694581758933_0006, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1694581758933_0006/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1694581758933_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-09-12 22:27:11,719 Stage-1 map = 0%,  reduce = 0%
2023-09-12 22:27:16,900 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.78 se
c
2023-09-12 22:27:22,105 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.53
sec
MapReduce Total cumulative CPU time: 1 seconds 530 msec
Ended Job = job_1694581758933_0006
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.53 sec   HDFS Read: 115817
HDFS Write: 4 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 530 msec
OK
496
Time taken: 17.96 seconds, Fetched: 1 row(s)
hive>
```

Q08
SELECTMIN(Total)ASMinSaleTotal, MAX(Total)ASMaxSaleTotalFROMsales;

```
cloudera@quickstart:~                                          _  □  ✕

 File  Edit  View  Search  Terminal  Help

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_1694581758933_0007, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1694581758933_0007/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1694581758933_0007
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-09-12 22:27:41,681 Stage-1 map = 0%,   reduce = 0%
2023-09-12 22:27:46,832 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.6 sec
2023-09-12 22:27:52,028 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.36
sec
MapReduce Total cumulative CPU time: 1 seconds 360 msec
Ended Job = job_1694581758933_0007
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.36 sec   HDFS Read: 115657
HDFS Write: 16 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 360 msec
OK
10.6785 1042.65
Time taken: 17.005 seconds, Fetched: 1 row(s)
hive> ▮
```

Q09
SELECT DISTINCT ProductLine FROM sales WHERE Rating>8;

```
Starting Job = job_1694581758933_0008, Tracking URL = http://quickstart.cloudera
:8088/proxy/application_1694581758933_0008/
Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1694581758933_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-09-12 22:28:17,619 Stage-1 map = 0%,   reduce = 0%
2023-09-12 22:28:23,021 Stage-1 map = 100%,   reduce = 0%, Cumulative CPU 0.83 se
c
2023-09-12 22:28:29,225 Stage-1 map = 100%,   reduce = 100%, Cumulative CPU 1.59
sec
MapReduce Total cumulative CPU time: 1 seconds 590 msec
Ended Job = job_1694581758933_0008
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 1.59 sec   HDFS Read: 115570
HDFS Write: 117 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 590 msec
OK
Electronic accessories
Fashion accessories
Food and beverages
Health and beauty
Home and lifestyle
Sports and travel
Time taken: 18.222 seconds, Fetched: 6 row(s)
hive> ▮
```

Q10
SELECT *FROMsalesWHEREPayment='Cash'ANDGender='Female';

```
                                     cloudera@quickstart:~            _ □ ✕

 File  Edit  View  Search  Terminal  Help
NULL    A       Yangon   Normal  Female  Fashion accessories       94.67   4       1
8.934   397.614 NULL     Cash    18.934  6.8
NULL    B       Mandalay         Normal  Female  Home and lifestyle        16.37   6
4.911   103.131 NULL     Cash    4.911   7.0
NULL    A       Yangon   Normal  Female  Sports and travel         89.48   5       2
2.37    469.77  NULL     Cash    22.37   7.4
NULL    A       Yangon   Member  Female  Health and beauty         92.09   3       1
3.8135  290.0835         NULL    Cash    13.8135 4.2
NULL    C       Naypyitaw        Member  Female  Food and beverages        21.04   4
4.208   88.368  NULL     Cash    4.208   7.6
NULL    A       Yangon   Normal  Female  Fashion accessories       42.57   7       1
4.8995  312.8895         NULL    Cash    14.8995 6.8
NULL    C       Naypyitaw        Normal  Female  Home and lifestyle        15.8    1
0       7.9     165.9    NULL    Cash    7.9     7.8
NULL    A       Yangon   Member  Female  Food and beverages        98.66   9       4
4.397   932.337 NULL     Cash    44.397  8.4
NULL    A       Yangon   Normal  Female  Health and beauty         15.8    3       2
.37     49.77   NULL     Cash    2.37    9.5
NULL    A       Yangon   Normal  Female  Food and beverages        74.66   4       1
4.932   313.572 NULL     Cash    14.932  8.5
NULL    A       Yangon   Member  Female  Fashion accessories       88.34   7       3
0.919   649.299 NULL     Cash    30.919  6.6
Time taken: 0.04 seconds, Fetched: 178 row(s)
hive> █
```