

Lab Exercise 06: Analysis the Supermarket Sales data using Hive Query Language

This exercise try to Analysis the Supermarket Sales data using Hive Query Language.

Step 01: Display Available Database

>show databases;

```
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> show databases;
OK
default
Time taken: 8.582 seconds, Fetched: 1 row(s)
```

Step 02: Create Database as “SuperMarket”

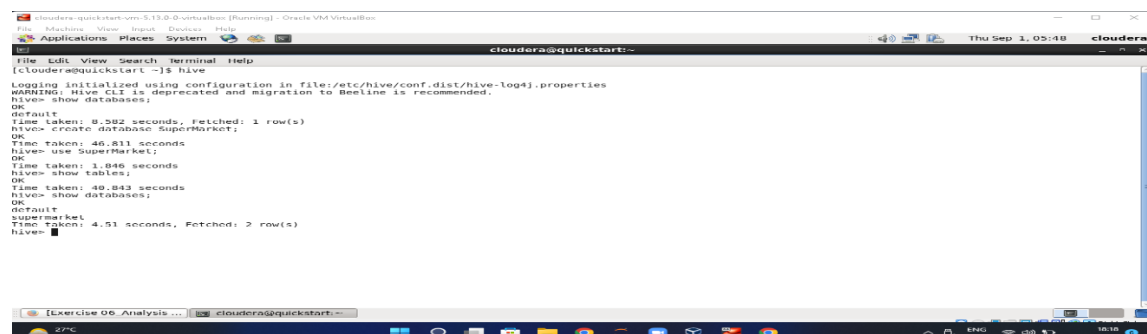
>create database SuperMarket;

A screenshot of a terminal window titled 'cloudera@quickstart:~'. The terminal shows the execution of several Hive commands. It starts with 'show databases;' which returns 'default'. Then, 'create database SuperMarket;' is executed, followed by 'use SuperMarket;'. Finally, 'show tables;' is executed, returning an empty result. The terminal also shows the execution of 'show databases;' again, which now returns both 'default' and 'supermarket'. The output for 'show tables;' is as follows:

```
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> show databases;
OK
default
Time taken: 8.582 seconds, Fetched: 1 row(s)
hive> create database SuperMarket;
OK
Time taken: 46.811 seconds
hive> use SuperMarket;
OK
Time taken: 1.846 seconds
hive> show tables;
OK
Time taken: 40.843 seconds
hive> show databases;
OK
default
supermarket
Time taken: 4.51 seconds, Fetched: 2 row(s)
hive>
```

Step 03: Use “SuperMarket” Database

>use SuperMarket;

A screenshot of a terminal window titled 'cloudera@quickstart:~'. The terminal shows the execution of several Hive commands. It starts with 'show databases;' which returns 'default'. Then, 'create database SuperMarket;' is executed, followed by 'use SuperMarket;'. Finally, 'show tables;' is executed, returning an empty result. The terminal also shows the execution of 'show databases;' again, which now returns both 'default' and 'supermarket'. The output for 'show tables;' is as follows:

```
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> show databases;
OK
default
Time taken: 8.582 seconds, Fetched: 1 row(s)
hive> create database SuperMarket;
OK
Time taken: 46.811 seconds
hive> use SuperMarket;
OK
Time taken: 1.846 seconds
hive> show tables;
OK
Time taken: 40.843 seconds
hive> show databases;
OK
default
supermarket
Time taken: 4.51 seconds, Fetched: 2 row(s)
hive>
```

Step 04: Display available Tables

>describe Sales;

```
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: 34.143 seconds, Fetched: 14 row(s)
hive>
```

Step 05: Create Table “Sales” with following scheme

**>create table Sales(InvoiceID STRING,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)row format delimited fields terminated by','
tblproperties("skip.header.line.count"="1");**

```
hive> create table Sales(InvoiceID STRING,Branch STRING,City STRING,Customertype STRING,GENDER STRING,Productline STRING,Unitprice FLOAT,Quantity INT,Tax5 FLOAT,Total F
LOAT,Date DATE,Payment STRING,Grossincome FLOAT,Rating FLOAT)row format delimited fields terminated by',' tblproperties("skip.header.line.count"="1");
OK
Time taken: 1.958 seconds
```

Step 06: Copy ‘supermarket_sales.csv’ into Hadoop local (/home/cloudera/)

**>LOAD DATA LOCAL INPATH
‘/home/cloudera/supermarket_sales.csv’ INTO TABLE Sales;**

```
hive> create table Sales(InvoiceID INT,Branch STRING,City STRING,Customertype STRING,GENDER STRING,Productline STRING,Unitprice FLOAT,Quantity INT,Tax5 FLOAT,Total FLOA
T,Date DATE,Payment STRING,Grossincome FLOAT,Rating FLOAT)row format delimited fields terminated by',' tblproperties("skip.header.line.count"="1");
OK
Time taken: 7.503 seconds
hive> LOAD DATA LOCAL INPATH '/home/cloudera/supermarket_sales.csv' INTO TABLE Sales;
Loading data to table supermarket.Sales
Table supermarket.Sales stats: [numFiles=1, totalSize=107805]
OK
Time taken: 41.243 seconds
hive>
```

Big Data Management and Analytics Lab

Step 07: Load 'supermarket_sales.csv' data into table 'Sales'

>select * from Sales;

cloudera@quickstart:~\$ select * from Sales;

invoiceid	branch	city	customertype	gender	productline	unitprice	quantity	tax5	total	date	payment	grossincome	rating	product details
746-04-1077	B	Mandalay	Member	Female	Food and beverages	84.63	10	42.315	888.615	NULL	Credit card	42.315	9.0	Food and beverages
448-34-8700	B	Mandalay	Member	Male	Home and lifestyle	36.91	7	12.9185	271.2885	NULL	Ewallet	12.9185	6.7	Home and lifestyle
452-04-8808	B	Mandalay	Member	Male	Electronic accessories	87.08	7	30.478	640.038	NULL	Cash	30.478	5.5	Electronic accessories
531-56-4728	A	Yangon	Normal	Male	Home and lifestyle	80.08	3	12.012	252.252	NULL	Cash	12.012	5.4	Home and lifestyle
744-82-9138	C	Naypyitaw	Normal	Male	Fashion accessories	86.13	2	8.613	180.873	NULL	Cash	8.613	8.2	Fashion accessories
883-69-1285	B	Mandalay	Member	Male	Fashion accessories	49.92	2	4.992	104.832	NULL	Credit card	4.992	7.0	Fashion accessories
221-25-5073	A	Yangon	Normal	Female	Food and beverages	74.66	4	14.932	313.572	NULL	Cash	14.932	8.5	Food and beverages
518-71-0847	B	Mandalay	Member	Male	Food and beverages	26.6	6	7.98	167.58	NULL	Ewallet	7.98	4.9	Food and beverages
156-20-0370	B	Mandalay	Normal	Female	Electronic accessories	25.45	1	1.2725	26.7225	NULL	Credit card	1.2725	5.1	Electronic accessories
151-33-7434	B	Mandalay	Normal	Female	Food and beverages	67.77	1	3.3085	71.1585	NULL	Credit card	3.3085	6.5	Food and beverages
728-47-9078	C	Naypyitaw	Member	Male	Food and beverages	59.59	4	11.918	250.278	NULL	Cash	11.918	9.8	Food and beverages
809-46-1866	A	Yangon	Normal	Male	Health and beauty	58.15	4	11.63	244.23	NULL	Cash	11.63	8.4	Health and beauty
139-32-4183	A	Yangon	Member	Female	Sports and travel	97.48	9	43.866	921.186	NULL	Ewallet	43.866	7.4	Sports and travel
148-41-7930	C	Naypyitaw	Normal	Male	Health and beauty	99.96	7	34.986	734.706	NULL	Cash	34.986	6.1	Health and beauty
189-40-5216	C	Naypyitaw	Normal	Male	Electronic accessories	96.37	7	33.7295	708.3195	NULL	Cash	33.7295	6.0	Electronic accessories
374-38-5555	B	Mandalay	Normal	Female	Fashion accessories	63.71	5	15.9275	334.4775	NULL	Ewallet	15.9275	8.5	Fashion accessories
764-44-8999	B	Mandalay	Normal	Female	Health and beauty	14.76	2	1.476	30.996	NULL	Ewallet	1.476	4.3	Health and beauty
552-44-5977	B	Mandalay	Member	Male	Health and beauty	62.0	8	24.8	520.8	NULL	Credit card	24.8	6.2	Health and beauty
267-62-7380	C	Naypyitaw	Member	Male	Electronic accessories	82.34	10	41.17	864.57	NULL	Ewallet	41.17	4.3	Electronic accessories
430-53-4718	B	Mandalay	Member	Male	Health and beauty	75.37	8	30.148	633.108	NULL	Credit card	30.148	8.4	Health and beauty
886-18-2897	A	Yangon	Normal	Female	Food and beverages	56.56	5	14.14	296.94	NULL	Credit card	14.14	4.5	Food and beverages
602-10-6955	B	Mandalay	Normal	Female	Sports and travel	76.6	10	38.3	804.3	NULL	Ewallet	38.3	6.0	Sports and travel
745-74-0715	A	Yangon	Normal	Male	Electronic accessories	58.03	2	5.803	121.863	NULL	Ewallet	5.803	8.8	Electronic accessories
690-01-6631	B	Mandalay	Normal	Male	Fashion accessories	17.49	10	8.745	183.645	NULL	Ewallet	8.745	6.6	Fashion accessories
652-49-6720	C	Naypyitaw	Member	Female	Electronic accessories	60.95	1	3.0475	63.9975	NULL	Ewallet	3.0475	5.9	Electronic accessories
233-67-5758	C	Naypyitaw	Normal	Male	Health and beauty	40.35	1	2.0175	42.3675	NULL	Ewallet	2.0175	6.2	Health and beauty
303-96-2227	B	Mandalay	Normal	Female	Home and lifestyle	97.38	10	48.69	1022.49	NULL	Ewallet	48.69	4.4	Home and lifestyle
727-02-1313	A	Yangon	Member	Male	Food and beverages	31.84	1	1.592	33.432	NULL	Cash	1.592	7.7	Food and beverages
347-56-2442	A	Yangon	Normal	Male	Home and lifestyle	65.82	1	3.291	69.111	NULL	Cash	3.291	4.1	Home and lifestyle
849-09-3807	A	Yangon	Member	Female	Fashion accessories	88.34	7	30.919	649.299	NULL	Cash	30.919	6.6	Fashion accessories

Time taken: 0.2 seconds, Fetched: 1000 row(s)

hive>

Step 08: Display the content in table 'Sales'

>Describe Sales;

cloudera@quickstart:~\$ describe Sales;

Field	Type
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: 34.143 seconds, Fetched: 14 row(s)

hive>

Q01: Display records who paid money through credit card?

>select * from Sales where Payment='Credit card';

```

cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
hive> select * from Sales where payment='Credit card';
OK
NULL A Yangon Normal Male Home and lifestyle 46.33 7 16.2155 340.5255 NULL Credit card 16.2155 7.4
NULL A Yangon Member Female Health and beauty 36.26 2 3.626 76.146 NULL Credit card 3.626 7.2
NULL B Mandalay Member Female Food and beverages 54.84 3 8.226 172.746 NULL Credit card 8.226 5.9
NULL A Yangon Member Female Health and beauty 68.93 7 24.1255 506.6355 NULL Credit card 24.1255 4.6
NULL A Yangon Normal Male Sports and travel 72.61 6 21.783 457.443 NULL Credit card 21.783 6.9
NULL A Yangon Normal Male Food and beverages 54.67 3 8.2005 172.2105 NULL Credit card 8.2005 8.6
NULL B Mandalay Normal Male Home and lifestyle 33.2 2 3.32 69.72 NULL Credit card 3.32 4.4
NULL A Yangon Member Female Home and lifestyle 52.59 8 21.036 441.756 NULL Credit card 21.036 8.5
NULL A Yangon Normal Female Fashion accessories 87.67 2 8.767 184.187 NULL Credit card 8.767 7.7
NULL B Mandalay Normal Male Fashion accessories 94.13 5 23.5325 494.1825 NULL Credit card 23.5325 4.8
NULL A Yangon Normal Male Health and beauty 96.58 2 9.658 202.818 NULL Credit card 9.658 5.1
NULL B Mandalay Member Male Health and beauty 56.69 9 25.5105 535.7205 NULL Credit card 25.5105 8.4
NULL B Mandalay Member Male Electronic accessories 18.93 6 5.679 119.259 NULL Credit card 5.679 8.1
NULL C Naypyitaw Member Male Fashion accessories 15.43 1 0.7715 16.2015 NULL Credit card 0.7715 6.1
NULL C Naypyitaw Member Female Sports and travel 24.74 3 3.711 77.931 NULL Credit card 3.711 10.0
NULL A Yangon Member Normal Sports and travel 15.81 10 7.905 166.005 NULL Credit card 7.905 8.6
NULL C Naypyitaw Normal Male Food and beverages 89.48 10 44.74 939.54 NULL Credit card 44.74 9.6
NULL C Naypyitaw Normal Female Electronic accessories 41.65 10 20.825 437.325 NULL Credit card 20.825 5.4
NULL C Naypyitaw Member Male Fashion accessories 49.04 9 22.068 463.428 NULL Credit card 22.068 8.6
NULL A Yangon Member Female Fashion accessories 20.01 9 9.0045 189.0945 NULL Credit card 9.0045 5.7
NULL C Naypyitaw Normal Female Health and beauty 99.19 6 29.757 624.897 NULL Credit card 29.757 5.5
NULL C Naypyitaw Member Female Food and beverages 80.36 4 16.072 337.512 NULL Credit card 16.072 8.3
NULL A Yangon Member Male Food and beverages 49.38 17.283 362.943 NULL Credit card 17.283 7.3
NULL B Mandalay Member Male Health and beauty 64.36 9 28.962 608.202 NULL Credit card 28.962 8.6
NULL C Naypyitaw Normal Male Health and beauty 89.75 1 4.4875 94.2375 NULL Credit card 4.4875 6.6
NULL C Naypyitaw Normal Male Fashion accessories 27.38 6 8.214 172.494 NULL Credit card 8.214 7.9
NULL C Naypyitaw Member Female Health and beauty 98.21 3 14.7315 309.3615 NULL Credit card 14.7315 7.8
NULL C Naypyitaw Member Female Home and lifestyle 80.79 9 36.3555 763.4655 NULL Credit card 36.3555 9.5
NULL C Naypyitaw Normal Female Fashion accessories 27.02 3 4.053 85.113 NULL Credit card 4.053 7.1
NULL A Yangon Normal Female Electronic accessories 99.56 8 39.824 836.304 NULL Credit card 39.824 5.2

```

```

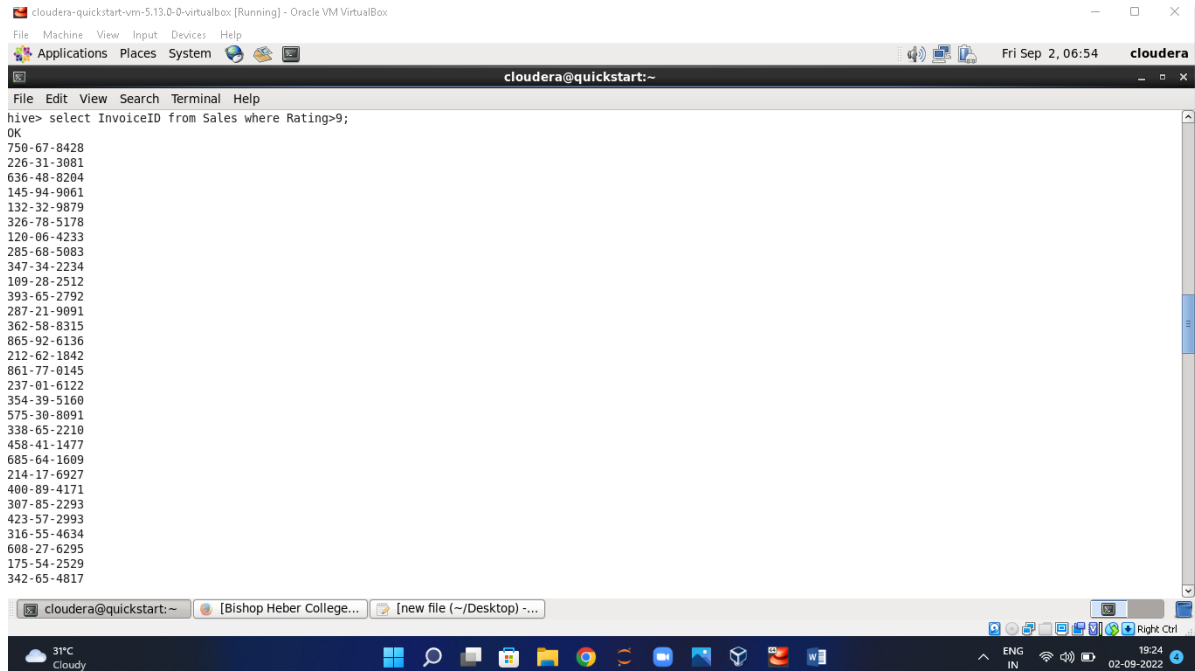
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
Time taken: 4.588 seconds, Fetched: 311 row(s)
hive>

```

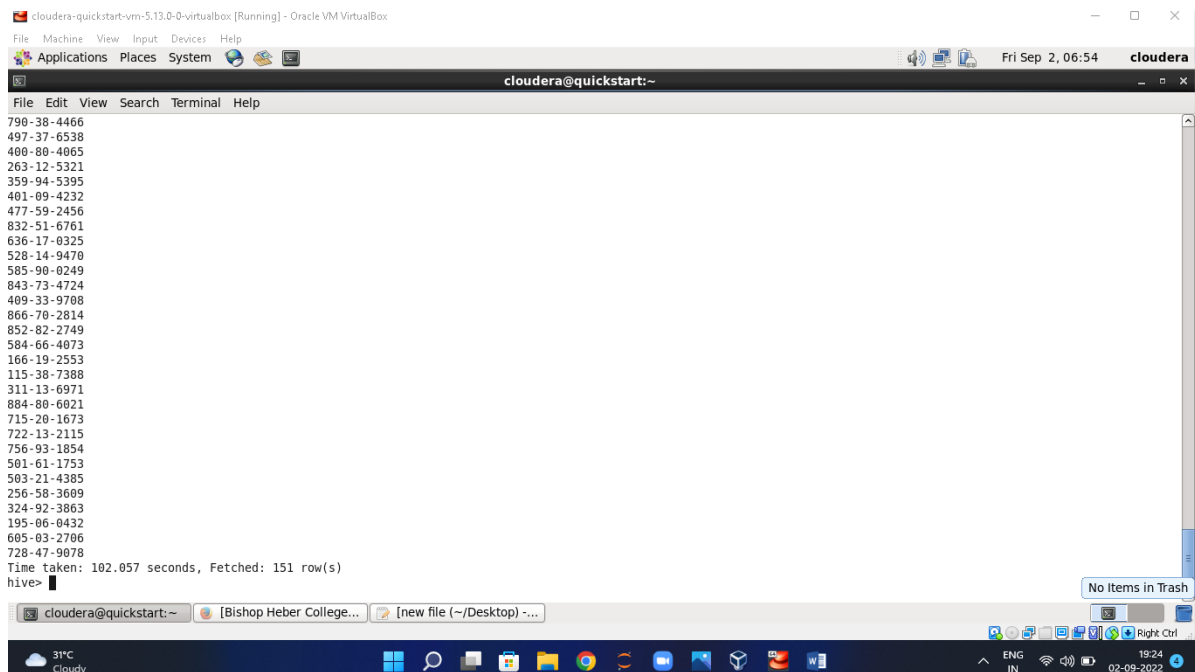
Big Data Management and Analytics Lab

Q02: Show all Invoice ID which table entry in has more than 9 rating.

>select InvoiceID from Sales where Rating>9;



```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
hive> select InvoiceID from Sales where Rating>9;
OK
750-67-8428
226-31-3081
636-48-8204
145-94-9061
132-32-9879
326-78-5178
120-06-4233
285-68-5083
347-34-2234
109-28-2512
393-65-2792
287-21-9091
362-58-8315
865-92-6136
212-62-1842
661-77-0145
237-01-6122
354-39-5160
575-30-8091
338-65-2210
458-41-1477
685-64-1609
214-17-6927
400-89-4171
307-85-2293
423-57-2993
316-55-4634
608-27-6295
175-54-2529
342-65-4817
```

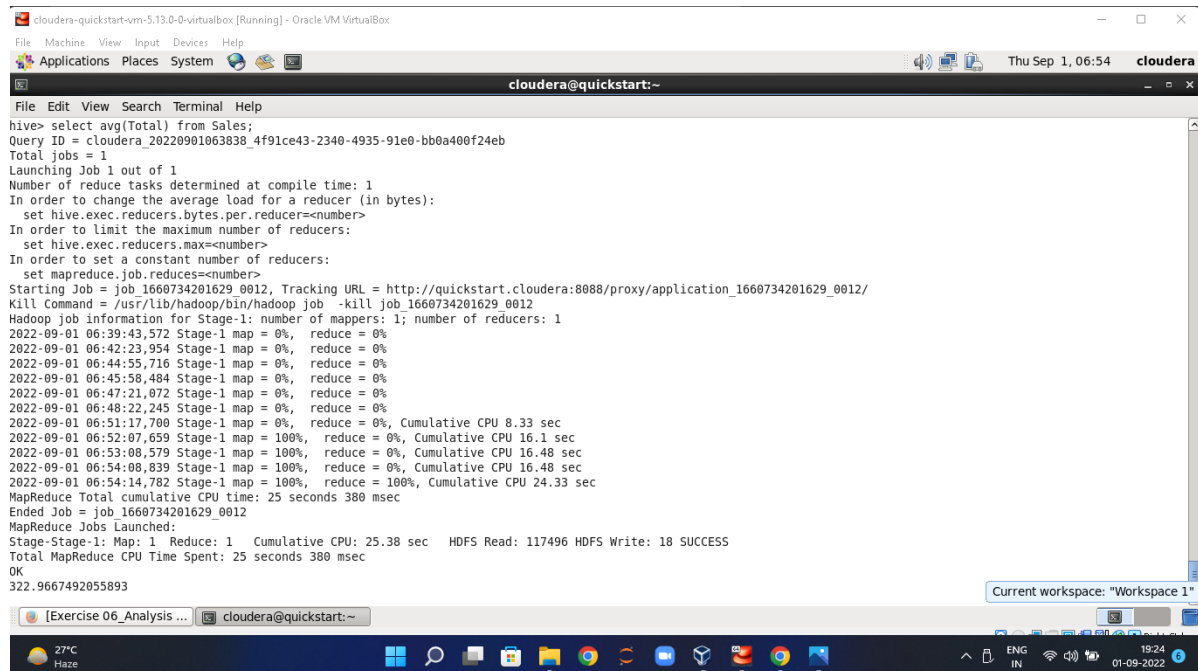


```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
790-38-4466
497-37-6538
400-80-4065
263-12-5321
359-94-5395
401-09-4232
477-59-2456
832-51-6761
636-17-0325
528-14-9470
585-90-0249
843-73-4724
409-33-9708
866-70-2814
852-82-2749
584-66-4073
166-19-2553
115-38-7388
311-13-6971
884-80-6021
715-20-1673
722-13-2115
756-93-1854
501-61-1753
503-21-4385
256-58-3609
324-92-3863
195-06-0432
605-03-2706
728-47-9078
Time taken: 102.057 seconds, Fetched: 151 row(s)
hive>
```

Big Data Management and Analytics Lab

Q03: Find Average sale amount in the data set.

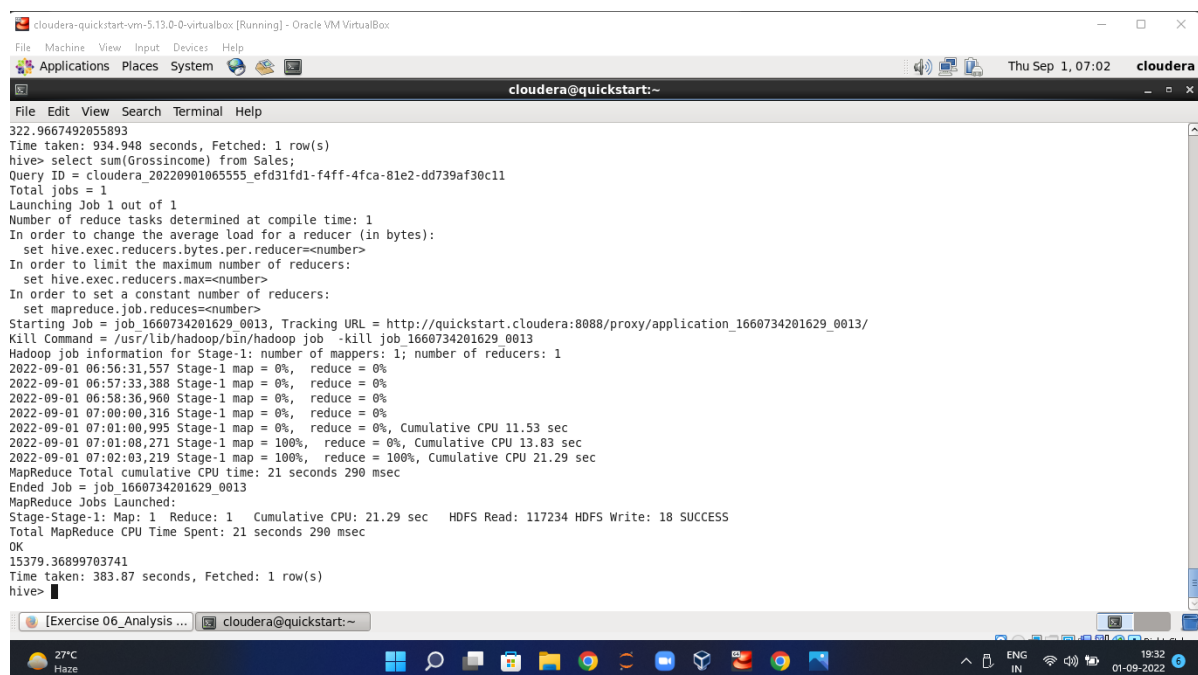
>select avg(Total) from Sales;



```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
hive> select avg(Total) from Sales;
Query ID = cloudera_20220901063838_4f91ce43-2340-4935-91e0-bb0a400f24eb
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_1660734201629_0012, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1660734201629_0012/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1660734201629_0012
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-01 06:39:43,572 Stage-1 map = 0%, reduce = 0%
2022-09-01 06:42:23,954 Stage-1 map = 0%, reduce = 0%
2022-09-01 06:44:55,716 Stage-1 map = 0%, reduce = 0%
2022-09-01 06:45:58,484 Stage-1 map = 0%, reduce = 0%
2022-09-01 06:47:21,072 Stage-1 map = 0%, reduce = 0%
2022-09-01 06:48:22,245 Stage-1 map = 0%, reduce = 0%
2022-09-01 06:51:17,700 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 8.33 sec
2022-09-01 06:52:07,659 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 16.1 sec
2022-09-01 06:53:08,579 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 16.48 sec
2022-09-01 06:54:08,539 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 16.48 sec
2022-09-01 06:54:14,782 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 24.33 sec
MapReduce Total cumulative CPU time: 25 seconds 380 msec
Ended Job = job_1660734201629_0012
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 25.38 sec HDFS Read: 117496 HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 25 seconds 380 msec
OK
322.9667492055893
```

Q04: Calculate total gross income in the data set.

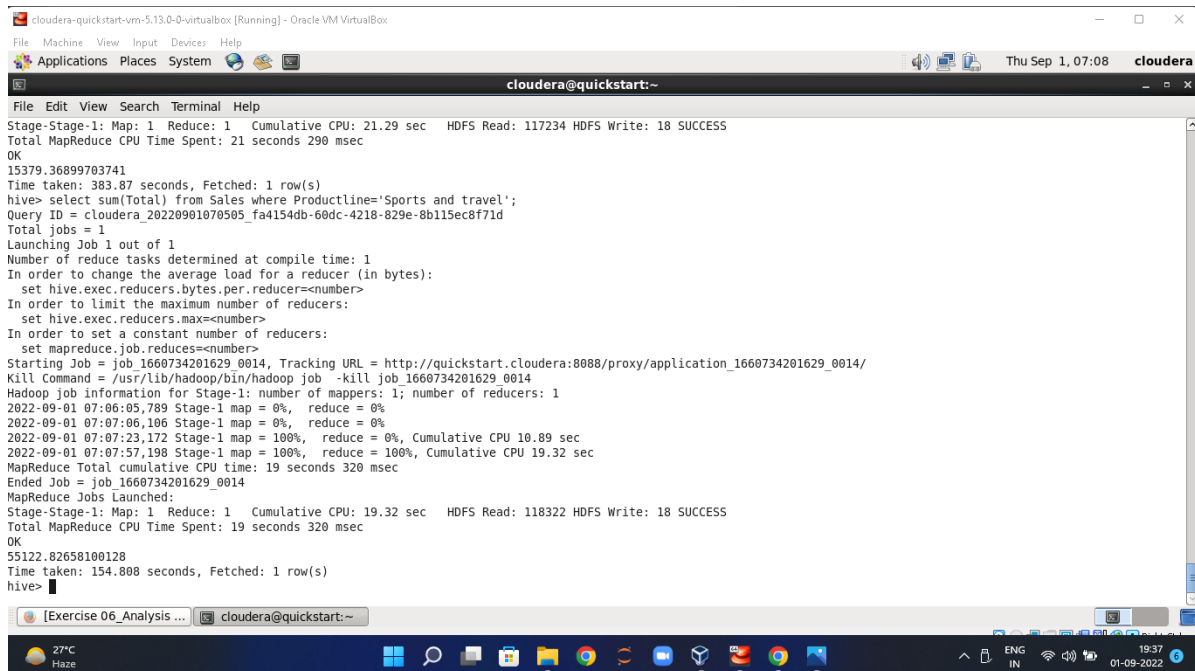
>select sum(Grossincome) from Sales;



```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
322.9667492055893
Time taken: 934.948 seconds, Fetched: 1 row(s)
hive> select sum(Grossincome) from Sales;
Query ID = cloudera_20220901065555_efd31fd1-f4ff-4fca-81e2-dd739af30c11
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_1660734201629_0013, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1660734201629_0013/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1660734201629_0013
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-01 06:56:31,557 Stage-1 map = 0%, reduce = 0%
2022-09-01 06:57:33,388 Stage-1 map = 0%, reduce = 0%
2022-09-01 06:58:36,960 Stage-1 map = 0%, reduce = 0%
2022-09-01 07:00:00,316 Stage-1 map = 0%, reduce = 0%
2022-09-01 07:01:00,995 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 11.53 sec
2022-09-01 07:01:08,271 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 13.83 sec
2022-09-01 07:02:03,219 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 21.29 sec
MapReduce Total cumulative CPU time: 21 seconds 290 msec
Ended Job = job_1660734201629_0013
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 21.29 sec HDFS Read: 117234 HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 21 seconds 290 msec
OK
15379.36899703741
Time taken: 383.87 seconds, Fetched: 1 row(s)
hive>
```


Q05: Find total sale amount under 'Sports and travel' product line.

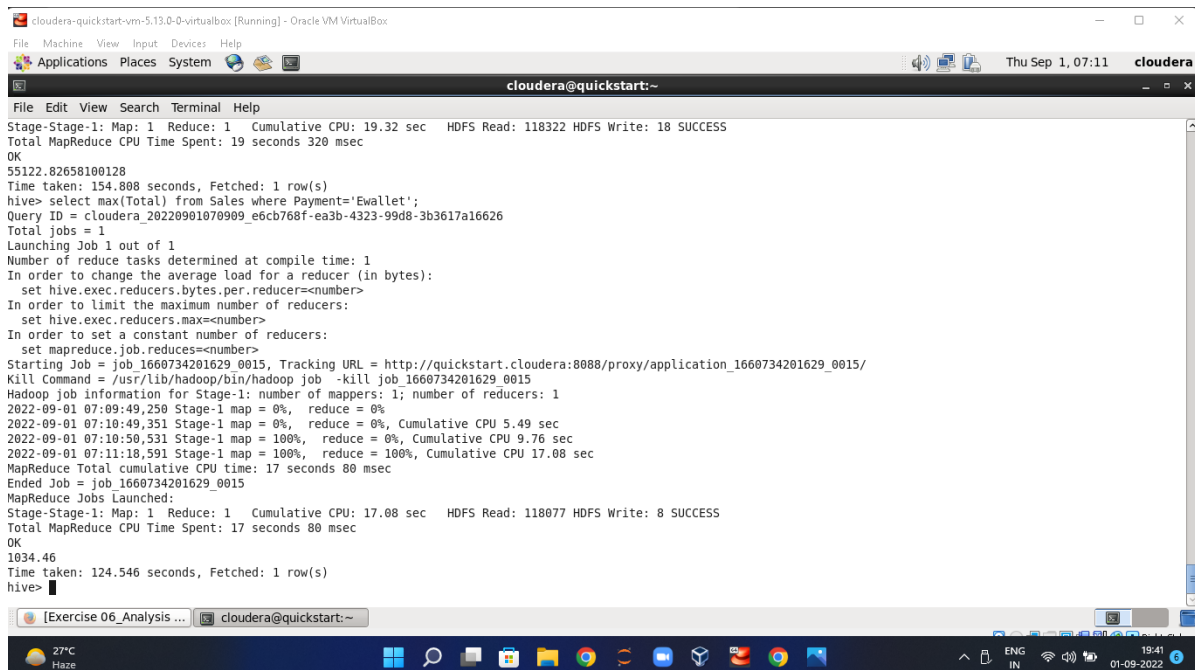
>select sum(Total) from Sales where Productline="Sports and travel";



```
cloudera@quickstart:~$ hiveshell
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 21.29 sec HDFS Read: 117234 HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 21 seconds 290 msec
OK
15379.36899703741
Time taken: 383.87 seconds, Fetched: 1 row(s)
hive> select sum(Total) from Sales where Productline='Sports and travel';
Query ID = cloudera_20220901070505_fa4154db-60dc-4218-829e-8b115ec8f71d
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_1660734201629_0014, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1660734201629_0014/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1660734201629_0014
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-01 07:06:05,789 Stage-1 map = 0%, reduce = 0%
2022-09-01 07:07:06,106 Stage-1 map = 0%, reduce = 0%
2022-09-01 07:07:23,172 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 10.89 sec
2022-09-01 07:07:57,198 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 19.32 sec
MapReduce Total cumulative CPU time: 19 seconds 320 msec
Ended Job = job_1660734201629_0014
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 19.32 sec HDFS Read: 118322 HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 19 seconds 320 msec
OK
55122.82658100128
Time taken: 154.808 seconds, Fetched: 1 row(s)
hive>
```

Q06: Find maximum sales amount in which payment in 'Ewallet' payment

>select max(Total) from Sales where Payment="EWallet";

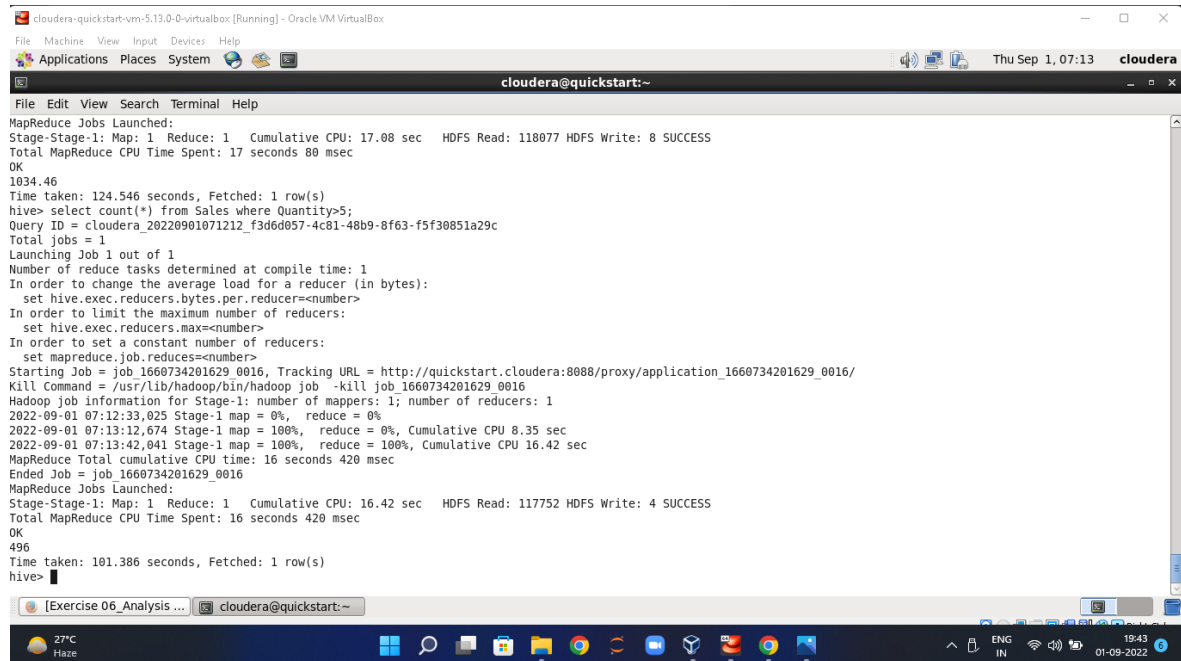


```
cloudera@quickstart:~$ hiveshell
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 19.32 sec HDFS Read: 118322 HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 19 seconds 320 msec
OK
55122.82658100128
Time taken: 154.808 seconds, Fetched: 1 row(s)
hive> select max(Total) from Sales where Payment='EWallet';
Query ID = cloudera_20220901070909_e6cb768f-ea3b-4323-99d8-3b3617a16626
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_1660734201629_0015, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1660734201629_0015/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1660734201629_0015
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-01 07:09:49,250 Stage-1 map = 0%, reduce = 0%
2022-09-01 07:10:49,351 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 5.49 sec
2022-09-01 07:10:50,531 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 9.76 sec
2022-09-01 07:11:18,591 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 17.08 sec
MapReduce Total cumulative CPU time: 17 seconds 80 msec
Ended Job = job_1660734201629_0015
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 17.08 sec HDFS Read: 118077 HDFS Write: 8 SUCCESS
Total MapReduce CPU Time Spent: 17 seconds 80 msec
OK
1034.46
Time taken: 124.546 seconds, Fetched: 1 row(s)
hive>
```

Big Data Management and Analytics Lab

Q07: Count number of sales which is has >5 quantity.

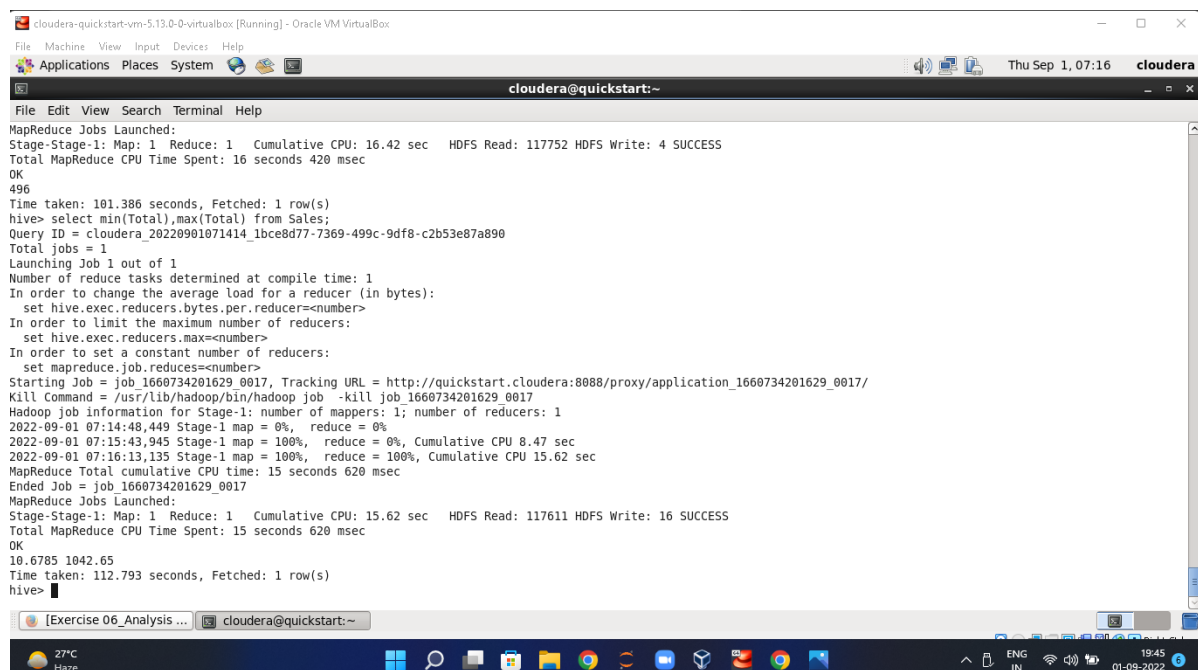
>select count(*) from Sales where Quantity>5;



```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 17.08 sec HDFS Read: 118077 HDFS Write: 8 SUCCESS
Total MapReduce CPU Time Spent: 17 seconds 80 msec
OK
1034.46
Time taken: 124.546 seconds, Fetched: 1 row(s)
hive> select count(*) from Sales where Quantity>5;
Query ID = cloudera_20220901071212_f3d6d057-4c81-48b9-8f63-f5f30851a29c
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_1660734201629_0016, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1660734201629_0016/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1660734201629_0016
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-01 07:12:33,025 Stage-1 map = 0%, reduce = 0%
2022-09-01 07:12:33,025 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 8.35 sec
2022-09-01 07:13:42,041 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 16.42 sec
MapReduce Total cumulative CPU time: 16 seconds 420 msec
Ended Job = job_1660734201629_0016
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 16.42 sec HDFS Read: 117752 HDFS Write: 4 SUCCESS
Total MapReduce CPU Time Spent: 16 seconds 420 msec
OK
496
Time taken: 101.386 seconds, Fetched: 1 row(s)
hive>
```

Q08: Find minimum and maximum sale total amount

>select min(Total),max(Total) from Sales;

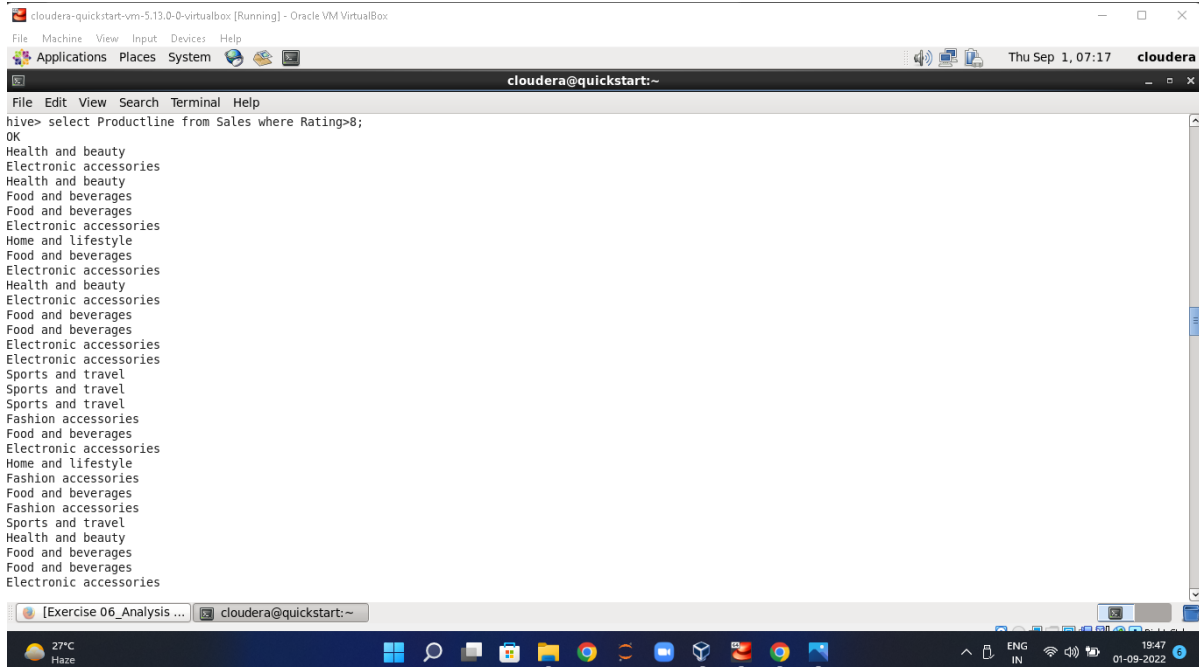


```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 16.42 sec HDFS Read: 117752 HDFS Write: 4 SUCCESS
Total MapReduce CPU Time Spent: 16 seconds 420 msec
OK
496
Time taken: 101.386 seconds, Fetched: 1 row(s)
hive> select min(Total),max(Total) from Sales;
Query ID = cloudera_20220901071414_1bce8d77-7369-499c-9df8-c2b53e87a890
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_1660734201629_0017, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1660734201629_0017/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1660734201629_0017
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-09-01 07:14:48,449 Stage-1 map = 0%, reduce = 0%
2022-09-01 07:15:43,945 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 8.47 sec
2022-09-01 07:16:13,135 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 15.62 sec
MapReduce Total cumulative CPU time: 15 seconds 620 msec
Ended Job = job_1660734201629_0017
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 15.62 sec HDFS Read: 117611 HDFS Write: 16 SUCCESS
Total MapReduce CPU Time Spent: 15 seconds 620 msec
OK
10.6785 1042.65
Time taken: 112.793 seconds, Fetched: 1 row(s)
hive>
```

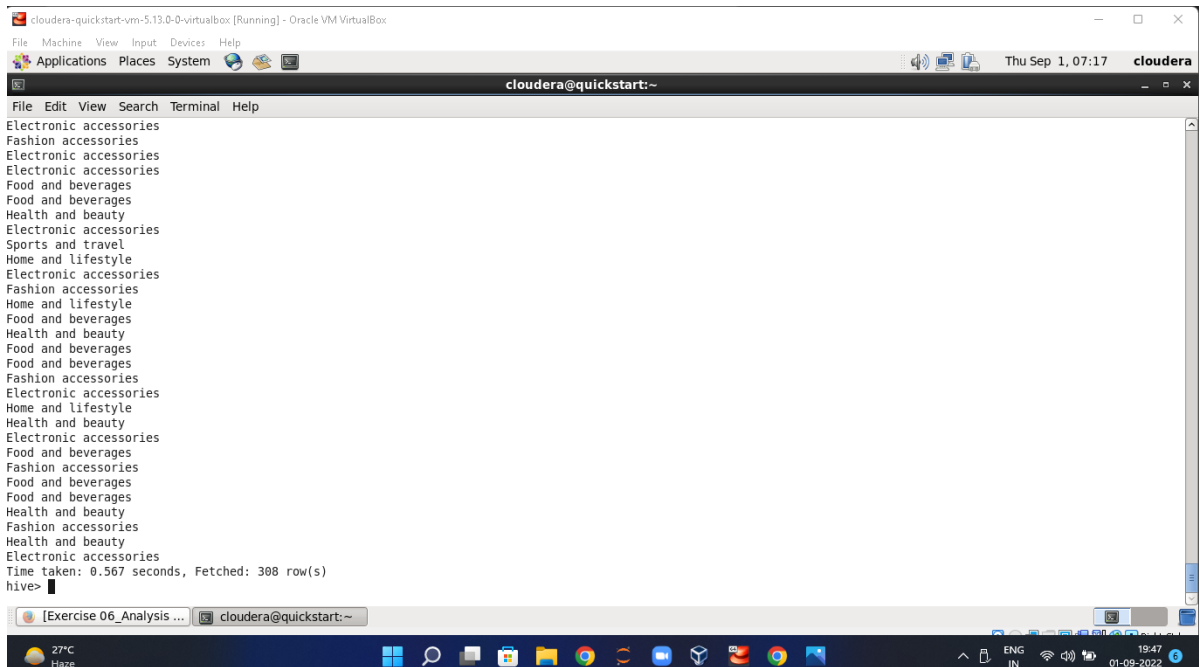

Big Data Management and Analytics Lab

Q09: List the product line which has more than 8 rating

>select Productline from Sales where Rating>8;



```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System cloudera@quickstart:~
File Edit View Search Terminal Help
hive> select Productline from Sales where Rating>8;
OK
Health and beauty
Electronic accessories
Health and beauty
Food and beverages
Food and beverages
Electronic accessories
Home and lifestyle
Food and beverages
Electronic accessories
Health and beauty
Electronic accessories
Food and beverages
Food and beverages
Electronic accessories
Electronic accessories
Sports and travel
Sports and travel
Sports and travel
Fashion accessories
Food and beverages
Electronic accessories
Home and lifestyle
Fashion accessories
Food and beverages
Fashion accessories
Sports and travel
Health and beauty
Food and beverages
Food and beverages
Electronic accessories
```



```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System cloudera@quickstart:~
File Edit View Search Terminal Help
Electronic accessories
Fashion accessories
Electronic accessories
Electronic accessories
Food and beverages
Food and beverages
Health and beauty
Electronic accessories
Sports and travel
Home and lifestyle
Electronic accessories
Fashion accessories
Home and lifestyle
Food and beverages
Health and beauty
Food and beverages
Food and beverages
Fashion accessories
Electronic accessories
Home and lifestyle
Health and beauty
Electronic accessories
Food and beverages
Fashion accessories
Food and beverages
Food and beverages
Health and beauty
Fashion accessories
Health and beauty
Electronic accessories
Time taken: 0.567 seconds, Fetched: 308 row(s)
hive>
```

Big Data Management and Analytics Lab

Q10: Display the records who has paid through cash and female customer.

>select * from Sales where Payment="Cash" and Gender="Female";

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System cloudera@quickstart:~
Thu Sep 1, 07:19 cloudera

File Edit View Search Terminal Help
hive> select * from Sales where Payment='Cash' and Gender='Female';
OK
NULL C Naypyitaw Normal Female Electronic accessories 15.28 5 3.82 88.22 NULL Cash 3.82 9.6
NULL A Yangon Normal Female Health and beauty 71.38 10 35.69 749.49 NULL Cash 35.69 5.7
NULL B Mandalay Member Female Sports and travel 93.72 6 28.116 590.436 NULL Cash 28.116 4.5
NULL B Mandalay Normal Female Food and beverages 88.36 5 22.89 463.89 NULL Cash 22.89 9.6
NULL B Mandalay Member Female Sports and travel 69.12 6 20.736 435.456 NULL Cash 20.736 5.6
NULL C Naypyitaw Member Female Food and beverages 98.7 8 39.48 829.08 NULL Cash 39.48 7.6
NULL B Mandalay Member Female Electronic accessories 93.96 4 18.792 394.632 NULL Cash 18.792 9.5
NULL A Yangon Member Female Food and beverages 44.59 5 11.1475 234.0975 NULL Cash 11.1475 8.5
NULL C Naypyitaw Normal Female Electronic accessories 85.98 8 34.392 722.232 NULL Cash 34.392 8.2
NULL A Yangon Member Female Home and lifestyle 72.35 10 36.175 759.675 NULL Cash 36.175 5.4
NULL A Yangon Member Female Health and beauty 18.33 1 0.9165 19.2465 NULL Cash 0.9165 4.3
NULL C Naypyitaw Normal Female Electronic accessories 75.91 6 22.773 478.233 NULL Cash 22.773 8.7
NULL C Naypyitaw Normal Female Health and beauty 20.38 5 5.095 106.995 NULL Cash 5.095 6.0
NULL B Mandalay Normal Female Health and beauty 76.99 6 23.097 485.037 NULL Cash 23.097 6.1
NULL C Naypyitaw Member Female Home and lifestyle 47.38 4 9.476 198.996 NULL Cash 9.476 7.1
NULL C Naypyitaw Normal Female Electronic accessories 12.45 6 3.735 78.435 NULL Cash 3.735 4.1
NULL C Naypyitaw Normal Female Electronic accessories 23.07 9 10.3815 218.0115 NULL Cash 10.3815 4.9
NULL C Naypyitaw Normal Female Food and beverages 33.98 9 15.291 321.111 NULL Cash 15.291 4.2
NULL B Mandalay Normal Female Fashion accessories 72.84 7 25.494 535.374 NULL Cash 25.494 8.4
NULL A Yangon Normal Member Sports and travel 32.25 5 8.0625 169.3125 NULL Cash 8.0625 9.0
NULL B Mandalay Normal Female Fashion accessories 39.62 7 13.867 291.207 NULL Cash 13.867 7.5
NULL A Yangon Member Sports and travel 92.13 6 27.639 580.419 NULL Cash 27.639 8.3
NULL B Mandalay Normal Female Sports and travel 34.84 4 6.968 146.328 NULL Cash 6.968 7.4
NULL A Yangon Member Home and lifestyle 34.42 6 10.326 216.846 NULL Cash 10.326 9.8
NULL C Naypyitaw Member Female Health and beauty 68.6 10 34.3 720.3 NULL Cash 34.3 9.1
NULL C Naypyitaw Normal Female Sports and travel 80.97 8 32.388 680.148 NULL Cash 32.388 9.3
NULL C Naypyitaw Normal Female Food and beverages 87.8 9 39.51 829.71 NULL Cash 39.51 9.2
NULL C Naypyitaw Member Female Fashion accessories 74.29 1 3.7145 78.0045 NULL Cash 3.7145 5.0
NULL B Mandalay Member Female Electronic accessories 57.49 4 11.498 241.458 NULL Cash 11.498 6.6
NULL C Naypyitaw Normal Female Food and beverages 93.26 9 41.967 881.307 NULL Cash 41.967 8.8
```

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System cloudera@quickstart:~
Thu Sep 1, 07:19 cloudera

File Edit View Search Terminal Help
hive> select * from Sales where Payment='Cash' and Gender='Female';
Time taken: 0.517 seconds, Fetched: 178 row(s)
hive>
```

