

Hive code 2 :

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~$ hive> select * from sales_data;
OK
clothes NULL
toys NULL
clocks NULL
Time taken: 1.059 seconds, Fetched: 3 row(s)
hive> exit;
WARN: The method class org.apache.commons.logging.impl.SLF4JLogFactory@release
e() was invoked.
WARN: Please see http://www.slf4j.org/codes.html#release for an explanation.
[cloudera@quickstart data]$ head /tmp/sales_prac/sales_data_raw.csv
head: cannot open '/tmp/sales_prac/sales_data_raw.csv' for reading: No such f
ile or directory
[cloudera@quickstart data]$ cd /tmp/sales_data_prac
[cloudera@quickstart sales_data_prac]$ ls
sales_data_raw.csv
[cloudera@quickstart sales_data_prac]$ head sales_data_raw.csv
clothes,1000
toys,2000
clocks,3000
[cloudera@quickstart sales_data_prac]$ cd data
bash: cd: data: No such file or directory
[cloudera@quickstart sales_data_prac]$ ls
sales_data_raw.csv
[cloudera@quickstart sales_data_prac]$ head sales_data_raw.csv
clothes,1000
toys,2000
clocks,3000
[cloudera@quickstart sales_data_prac]$ 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> use hive;
OK
Time taken: 0.829 seconds
hive> drop table sales_data;
OK
Time taken: 0.733 seconds
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 sales_data_prac]$ cd
[cloudera@quickstart ~]$ ls
cloudera-manager cs_apl.py create data Desktop Documents Downloads eclipse enterprise-deployment.json express-deployment.json herberos lib Music parcels Pictures Public Templates test2.txt Videos vigya workspace
[cloudera@quickstart ~]$ cd /tmp
```

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~$ hive> select * from sales_data;
OK
clothes NULL
toys NULL
clocks NULL
Time taken: 1.059 seconds, Fetched: 3 row(s)
hive> exit;
WARN: The method class org.apache.commons.logging.impl.SLF4JLogFactory@release
e() was invoked.
WARN: Please see http://www.slf4j.org/codes.html#release for an explanation.
[cloudera@quickstart data]$ head /tmp/sales_prac/sales_data_raw.csv
head: cannot open '/tmp/sales_prac/sales_data_raw.csv' for reading: No such f
ile or directory
[cloudera@quickstart data]$ cd /tmp/sales_data_prac
[cloudera@quickstart sales_data_prac]$ ls
sales_data_raw.csv
[cloudera@quickstart sales_data_prac]$ head sales_data_raw.csv
clothes,1000
toys,2000
clocks,3000
[cloudera@quickstart sales_data_prac]$ cd data
bash: cd: data: No such file or directory
[cloudera@quickstart sales_data_prac]$ ls
sales_data_raw.csv
[cloudera@quickstart sales_data_prac]$ head sales_data_raw.csv
clothes,1000
toys,2000
clocks,3000
[cloudera@quickstart sales_data_prac]$ 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> use hive;
OK
Time taken: 0.829 seconds
hive> drop table sales_data;
OK
Time taken: 0.733 seconds
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 sales_data_prac]$ cd
[cloudera@quickstart ~]$ ls
cloudera-manager cs_apl.py create data Desktop Documents Downloads eclipse enterprise-deployment.json express-deployment.json herberos lib Music parcels Pictures Public Templates test2.txt Videos vigya workspace
[cloudera@quickstart ~]$ cd /tmp
```

```
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
at org.apache.hadoop.hive.cli.CliDriver.run(CliDriver.java:699)
at org.apache.hadoop.hive.cli.CliDriver.main(CliDriver.java:634)
at sun.reflect.NativeMethodAccessorImpl.invoke(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 3:13 mismatched input '<EOF>' expecting ')' near 'string' in create table statement
hive> create table sales_data
> (
> p type string,
> total_sales int
> row format delimited
> fields terminated by ',';
FAILED: ParseException line 3:8 missing ')' at 'row' near '<EOF>'
hive> create table sales_data
> (
> p type string,
> total_sales int
> )
> row format delimited
> fields terminated by ',';
OK
Time taken: 0.813 seconds
hive> load data local inpath 'file:///tmp/hive prac/sales_data_raw.csv' into
table sales_data;
Loading data to table hive 1.sales_data
Table hive 1.sales_data stats: [numFiles=1, totalSize=29]
OK
Time taken: 1.787 seconds
hive> select*from sales_data;
OK
"clothes" 10000
"toys" 20000
Time taken: 1.063 seconds, Fetched: 2 row(s)
hive> drop table sales_data;
OK
Time taken: 0.606 seconds
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 ~]$ clear

[cloudera@quickstart ~]$ 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> use hive 1;
```

```
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
OK
Time taken: 1.244 seconds
hive> create table sales_data
> (
> p type string,
> total_sales int
> )
> row

exit;
NullPointerException[2228[]]
at org.apache.hadoop.hive ql.parse.HiveParser.tableRowFormat(HiveParser.java:34697)
at org.apache.hadoop.hive ql.parse.HiveParser.createTableStatement(HiveParser.java:5196)
at org.apache.hadoop.hive ql.parse.HiveParser.ddlStatement(HiveParser.java:2557)
at org.apache.hadoop.hive ql.parse.HiveParser.execStatement(HiveParser.java:1589)
at org.apache.hadoop.hive ql.parse.HiveParser.statement(HiveParser.java:1065)
at org.apache.hadoop.hive ql.parse.ParserDriver.parse(ParserDriver.java:281)
at org.apache.hadoop.hive ql.parse.ParserDriver.parse(ParserDriver.java:166)
at org.apache.hadoop.hive ql.Driver.compileInternal(Driver.java:522)
at org.apache.hadoop.hive ql.Driver.compileInternal(Driver.java:1356)
at org.apache.hadoop.hive ql.Driver.runInternal(Driver.java:1473)
at org.apache.hadoop.hive ql.Driver.run(Driver.java:1265)
at org.apache.hadoop.hive ql.Driver.run(Driver.java:1275)
at org.apache.hadoop.hive.cli.CliDriver.processLocalCmd(CliDriver.java:226)
at org.apache.hadoop.hive.cli.CliDriver.processCmd(CliDriver.java:175)
at org.apache.hadoop.hive.cli.CliDriver.processLine(CliDriver.java:38)
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
781) at org.apache.hadoop.hive.cli.CliDriver.executeDriver(CliDriver.java:
at org.apache.hadoop.hive.cli.CliDriver.run(CliDriver.java:699)
at org.apache.hadoop.hive.cli.CliDriver.main(CliDriver.java:634)
at sun.reflect.NativeMethodAccessorImpl.invoke(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorIm
pl.java:57)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAcc
essorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:666)
at org.apache.hadoop.util.RunJar.run(RunJar.java:221)
at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
FAILED: ParseException line 6:9 cannot recognize input near 'exit' '<EO
F>' in table row format specification
hive>
> exit;
WARN: The method class org.apache.commons.logging.impl.SLF4JLogFactory#releas
e() was invoked.
WARN: Please see http://www.slf4j.org/codes.html#release for an explanation.
[cloudera@quickstart ~]$ clear

[cloudera@quickstart ~]$ hive

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4
j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> use hive_1;
OK
Time taken: 0.681 seconds
hive> create table sales_data
> (
> p_type string,
> total_sales int
> )
> row format delimited
> fields terminated by ',';
OK
Time taken: 0.817 seconds
hive> load data local inpath 'file:///tmp/hive-prac/sales_data_raw.csv' into
table sales_data;
Loading data to table hive_1.sales_data
Table hive_1.sales_data stats: [numFiles=1, totalSize=29]
OK
Time taken: 1.985 seconds
hive> select * from sales_data;
OK
"clothes" 10000
"toys" 20000
Time taken: 1.15 seconds, Fetched: 2 row(s)

Cloudera Live: Welco... cloudera@quickstart:~$
```

```
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> create table sales_data backup as select * from sales_data;
Query ID = cloudera_20230317083636_d10dc2a-0fc-4b80-a255-52349857b766
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1679055381392_0001, Tracking URL = http://quickstart.cloud
era:8088/proxy/application_1679055381392_0001/
Kill command = /usr/lib/hadoop/bin/hadoop job -kill job_1679055381392_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers:
0
2023-03-17 08:37:11.112 Stage-1 map = 0%, reduce = 0%
2023-03-17 08:37:36.314 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 6.91
sec
MapReduce Total cumulative CPU time: 6 seconds 910 msec
Ended Job = job_1679055381392_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/hive_1.db
/hive-staging/hive_2023-03-17_08-36-38_832_439404790261000962-1/-ext-100001
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/warehouse/hive_1.db
/sales_data backup
Table hive_1.sales_data backup stats: [numFiles=1, numRows=2, totalSize=29, r
awdataSize=27]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 6.91 sec HDFS Read: 3642 HDFS Write
: 109 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 910 msec
OK
Time taken: 55.077 seconds
hive> describe extended sales_data;
OK
p_type string
total_sales int

Detailed Table Information Table[tableName:sales_data, dbName:hive_1, ow
ner:cloudera, createTime:1679067297, lastAccessTime:0, retention:0, sd:Storag
eDescriptor{cols:[FieldSchema{name:p_type, type:string, comment:null}], FieldS
chema{name:total_sales, type:int, comment:null}], location:hdfs://quickstart.
cloudera:8020/user/hive/warehouse/hive_1.db/sales_data, inputFormat:org.apach
e.hadoop.mapred.TextInputFormat, outputFormat:org.apache.hadoop.hive ql.io.Hi
veNoKeyTextOutputFormat, compressed:false, numBuckets:-1, serdeInfo:SerDe
Info{name:null, serializationLib:org.apache.hadoop.hive.serde2.lazy.LazySimpl
eSerDe, parameters:{serialization.format=, field.delim=}}, bucketCols:[], s
ortCols:[], parameters:{}, skewedInfo:SkewedInfo(skewedColNames:[], skewedCol
Values:[], skewedColValueLocations:()), storedAsSubDirectories:false}, par
titionKeys:[], parameters:{numFiles=1, transient lastDdlTime:1679067338, COU
M_STATS_ACCURATE=true, totalSize=29}, viewOriginalText:null, viewExpandedTex
t:null, tableType:MANAGED TABLE}
```



```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Fri Mar 17, 8:52 AM cloudera
cloudera@quickstart:~
File Edit View Search Terminal Help
eParser.java:41234)
    at org.apache.hadoop.hive ql.parse.HiveParser.fromStatement(HiveParse
r.java:41028)
    at org.apache.hadoop.hive ql.parse.HiveParser.queryStatementExpressio
nBody(HiveParser.java:40695)
    at org.apache.hadoop.hive ql.parse.HiveParser.queryStatementExpressio
n(HiveParser.java:40586)
    at org.apache.hadoop.hive ql.parse.HiveParser.execStatement(HiveParse
r.java:1529)
    at org.apache.hadoop.hive ql.parse.HiveParser.statement(HiveParser.ja
va:1065)
    at org.apache.hadoop.hive ql.parse.ParseDriver.parse(ParseDriver.java
:281)
    at org.apache.hadoop.hive ql.parse.ParseDriver.parse(ParseDriver.java
:166)
    at org.apache.hadoop.hive ql.Driver.compile(Driver.java:522)
    at org.apache.hadoop.hive ql.Driver.compileInternal(Driver.java:1356)
    at org.apache.hadoop.hive ql.Driver.runInternal(Driver.java:1473)
    at org.apache.hadoop.hive ql.Driver.run(Driver.java:1285)
    at org.apache.hadoop.hive ql.Driver.run(Driver.java:1275)
    at org.apache.hadoop.hive cli.CliDriver.processLocalCmd(CliDriver.jav
a:126)
    at org.apache.hadoop.hive cli.CliDriver.processCmd(CliDriver.java:175)
    at org.apache.hadoop.hive cli.CliDriver.processLine(CliDriver.java:38)
    at org.apache.hadoop.hive cli.CliDriver.executeDriver(CliDriver.java:
781)
    at org.apache.hadoop.hive cli.CliDriver.run(CliDriver.java:699)
    at org.apache.hadoop.hive cli.CliDriver.main(CliDriver.java:634)
    at sun.reflect.NativeMethodAccessorImpl.invoke(Native Method)
    at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorIm
pl.java:57)
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAc
cessorImpl.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 1:23 cannot recognize input near 'sales_data back
up' 'overwrite' 'table' in from source
hive> from sales_data backup overwrite table sales_data parquet select;
[1] * Stopped
hive
[cloudera@quickstart ~]$ hive

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4
j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> select * from sales_data parquet;
FAILED: SemanticException [Error 10001]: Line 1:14 Table not found 'sales_dat
a'
```

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Fri Mar 17, 8:58 AM cloudera
cloudera@quickstart:~
File Edit View Search Terminal Help
a parquet'
hive> select * from sales_data backup;
FAILED: SemanticException [Error 10001]: Line 1:14 Table not found 'sales_dat
a backup'
hive> select * from sales_data;
OK
clothes NULL
toys NULL
clocks NULL
Time taken: 1.552 seconds, Fetched: 3 row(s)
hive> drop table sales_data;
OK
Time taken: 0.621 seconds
hive> use hive1;
OK
Time taken: 0.077 seconds
hive> select * from sales_data;
OK
"clothes" 10000
"toys" 20000
Time taken: 0.26 seconds, Fetched: 2 row(s)
hive> select * from sales_data backup;
OK
"clothes" 10000
"toys" 20000
Time taken: 0.245 seconds, Fetched: 2 row(s)
hive> select * from sales_data parquet;
> select * from sales_data parquet;
FAILED: ParseException line 2:0 missing EOF at 'select' near 'sales_data parq
uet'
hive> select * from sales_data parquet;
OK
Time taken: 0.264 seconds
hive> from sales_data insert overwrite table sales_data parquet select;
Query ID = cloudera_20230317084444_76931b6f-e6c8-4468-bb71-0951c58ce45a
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1679055381392_0002, Tracking URL = http://quickstart.cloud
era:8088/proxy/application_1679055381392_0002/
Kill command = /usr/lib/hadoop/bin/hadoop job -kill job_1679055381392_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers:
0
2023-03-17 08:45:17.587 Stage-1 map = 0% reduce = 0%
2023-03-17 08:45:30.945 Stage-1 map = 100% reduce = 0% Cumulative CPU 8.39 sec
MapReduce Total cumulative CPU time: 8 seconds 390 msec
Ended Job = job_1679055381392_0002
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
```



```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Fri Mar 17, 9:07 AM cloudera
cloudera@quickstart:~
File Edit View Search Terminal Help
Stage-5 is filtered out by condition resolver.
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/warehouse/hive_1.db/sales_data.parquet/.hive-staging_hive_2023-03-17_08-44-54_017_4025997569128702125-1/-ext-10000
Loading data to table hive_1.sales_data parquet
Table hive_1.sales_data parquet stats: [numFiles=1, numRows=2, totalSize=363, rawDataSize=4]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 8.39 sec HDFS Read: 4100 HDFS Write: 444 SUCCESS
Total MapReduce CPU Time Spent: 8 seconds 390 msec
OK
Time taken: 52.154 seconds
hive> select * from sales_data parquet;
OK
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/lib/parquet/lib/parquet-format-2.1.0-cdh5.13.0.jar!/shaded/parquet/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/parquet/lib/parquet-hadoop-bundle-1.5.0-cdh5.13.0.jar!/shaded/parquet/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/parquet/lib/parquet-pig-bundle-1.5.0-cdh5.13.0.jar!/shaded/parquet/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/hive/lib/hive-exec-1.1.0-cdh5.13.0.jar!/shaded/parquet/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/hive/lib/hive-jdbc-1.1.0-cdh5.13.0-standalone.jar!/shaded/parquet/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [shaded.parquet.org.slf4j.helpers.NOPLoggerFactory]
"clothes" 10000
"toys" 20000
Time taken: 0.305 seconds, Fetched: 2 row(s)
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.
Mar 17, 2023 8:46:03 AM WARNING: parquet.hadoop.ParquetRecordReader: Can not initialize counter due to context is not a instance of TaskInputOutputContext, but is org.apache.hadoop.mapreduce.task.TaskAttemptContextImpl
Mar 17, 2023 8:46:03 AM INFO: parquet.hadoop.InternalParquetRecordReader: RecordReader initialized will read a total of 2 records.
Mar 17, 2023 8:46:03 AM INFO: parquet.hadoop.InternalParquetRecordReader: At row 0, reading next block
Mar 17, 2023 8:46:03 AM INFO: parquet.hadoop.InternalParquetRecordReader: block read in memory in 31 ms. row count = 2
[cloudera@quickstart ~]$ hadoop fs -ls /user/hive/warehouse/hive_1*
ls: /user/hive/warehouse/hive_1*: Unknown command
[cloudera@quickstart ~]$ hadoop fs -ls /user/hive/warehouse/hive_1*
Found 6 items
drwxrwxrwx - cloudera supergroup 0 2023-03-16 11:30 /user/hive/warehouse/hive_1.db/department_data
drwxrwxrwx - cloudera supergroup 0 2023-03-16 17:03 /user/hive/warehouse/hive_1.db/department_data_from_hdfs
drwxrwxrwx - cloudera supergroup 0 2023-03-16 18:04 /user/hive/warehouse/hive_1.db/employee
drwxrwxrwx - cloudera supergroup 0 2023-03-17 08:35 /user/hive/warehouse/hive_1.db/sales_data
drwxrwxrwx - cloudera supergroup 0 2023-03-17 08:37 /user/hive/warehouse/hive_1.db/sales_data_backup
```

```
drwxrwxrwx - cloudera supergroup 0 2023-03-17 08:45 /user/hive/warehouse/hive_1.db/sales_data parquet
[cloudera@quickstart ~]$ hadoop fs -ls /user/hive/warehouse/hive_1/sales_data
ls: /user/hive/warehouse/hive_1/sales_data: No such file or directory
[cloudera@quickstart ~]$ hadoop fs -ls /user/hive/warehouse/hive_1.db/sales_data
Found 1 items
-rwxrwxrwx 1 cloudera supergroup 29 2023-03-17 08:35 /user/hive/warehouse/hive_1.db/sales_data/sales_data.raw.csv
[cloudera@quickstart ~]$ hadoop fs -ls /user/hive/warehouse/hive_1.db/sales_data parquet
Found 1 items
-rwxrwxrwx 1 cloudera supergroup 363 2023-03-17 08:45 /user/hive/warehouse/hive_1.db/sales_data parquet/000000_0
[cloudera@quickstart ~]$
```