

**Spark SQL Error**  
**Query 2, Query 4**

16/01/01 21:08:32 ERROR Utils: Uncaught exception in thread Thread-ScriptTransformation-Feed  
java.io.IOException: Stream closed  
    at java.lang.ProcessBuilder\$NullOutputStream.write(ProcessBuilder.java:453)  
    at java.io.OutputStream.write(OutputStream.java:127)  
    at java.io.BufferedOutputStream.flushBuffer(BufferedOutputStream.java:93)  
    at java.io.BufferedOutputStream.write(BufferedOutputStream.java:137)  
    at java.io.DataOutputStream.write(DataOutputStream.java:118)  
    at org.apache.hadoop.hive.ql.exec.TextRecordWriter.write(TextRecordWriter.java:53)  
org.apache.spark.sql.hive.execution.ScriptTransformationWriterThread\$\$anonfun\$run\$1\$\$anonfun\$apply\$mcV\$sp\$  
1.apply(ScriptTransformation.scala:255)  
    at scala.collection.Iterator\$class.foreach(Iterator.scala:727)  
    at scala.collection.AbstractIterator.foreach(Iterator.scala:1157)  
    at org.apache.spark.util.Utils\$.logUncaughtExceptions(Utils.scala:1699)  
org.apache.spark.sql.hive.execution.ScriptTransformationWriterThread.run(ScriptTransformation.scala:244)  
Exception in thread "Thread-ScriptTransformation-Feed" java.io.IOException: Stream closed  
    at java.lang.ProcessBuilder\$NullOutputStream.write(ProcessBuilder.java:453)  
    at java.io.OutputStream.write(OutputStream.java:127)  
    at java.io.BufferedOutputStream.flushBuffer(BufferedOutputStream.java:93)  
    at java.io.BufferedOutputStream.write(BufferedOutputStream.java:137)  
    at java.io.DataOutputStream.write(DataOutputStream.java:118)  
    at org.apache.hadoop.hive.ql.exec.TextRecordWriter.write(TextRecordWriter.java:53)Exception in thread  
"Thread-ScriptTransformation-Feed" java.io.IOException: Stream closed  
    at java.lang.ProcessBuilder\$NullOutputStream.write(ProcessBuilder.java:453)  
    at java.io.OutputStream.write(OutputStream.java:127)  
    at java.io.BufferedOutputStream.flushBuffer(BufferedOutputStream.java:93)  
    at java.io.BufferedOutputStream.write(BufferedOutputStream.java:137)  
    at java.io.DataOutputStream.write(DataOutputStream.java:118)  
    at org.apache.hadoop.hive.ql.exec.TextRecordWriter.write(TextRecordWriter.java:53)  
    at  
org.apache.spark.sql.hive.execution.ScriptTransformationWriterThread\$\$anonfun\$run\$1\$\$anonfun\$apply\$mcV\$sp\$  
1.apply(ScriptTransformation.scala:277)

**Query 5**

16/01/01 21:28:12 ERROR Executor: Exception in task 0.0 in stage 150.0 (TID 150)  
org.apache.spark.SparkException: Kryo serialization failed: Buffer overflow. Available: 0, required: 6. To avoid this,  
increase spark.kryoserializer.buffer.max value.  
    at org.apache.spark.serializer.KryoSerializerInstance.serialize(KryoSerializer.scala:263)  
    at org.apache.spark.executor.Executor\$TaskRunner.run(Executor.scala:240)  
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1176)  
    at java.util.concurrent.ThreadPoolExecutor\$Worker.run(ThreadPoolExecutor.java:641)  
    at java.lang.Thread.run(Thread.java:853)  
16/01/01 21:28:12 WARN TaskSetManager: Lost task 0.0 in stage 150.0 (TID 150, localhost):  
org.apache.spark.SparkException: Kryo serialization failed: Buffer overflow. Available: 0, required: 6. To avoid this,  
increase spark.kryoserialize  
r.buffer.max value.  
    at org.apache.spark.serializer.KryoSerializerInstance.serialize(KryoSerializer.scala:263)  
    at org.apache.spark.executor.Executor\$TaskRunner.run(Executor.scala:240)  
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1176)

```
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:641)
at java.lang.Thread.run(Thread.java:853)
```

```
16/01/01 21:28:12 ERROR TaskSetManager: Task 0 in stage 150.0 failed 1 times; aborting job
16/01/01 21:28:12 INFO TaskSchedulerImpl: Removed TaskSet 150.0, whose tasks have all completed, from pool
16/01/01 21:28:12 INFO TaskSchedulerImpl: Cancelling stage 150
16/01/01 21:28:12 INFO DAGScheduler: ResultStage 150 (run at ThreadPoolExecutor.java:1176) failed in 9.771 s
16/01/01 21:28:12 INFO DAGScheduler: Job 150 failed: run at ThreadPoolExecutor.java:1176, took 9.790863 s
```

### Query 7

Error in query:

UUnsupported language features in query: INSERT INTO TABLE q07\_spark\_run\_query\_0\_result

SELECT

ca\_state,

COUNT(\*) AS cnt

FROM

customer\_address a,

customer c,

store\_sales s,

q07\_spark\_run\_query\_0\_temp highPriceltems

WHERE a.ca\_address\_sk = c.c\_current\_addr\_sk

AND c.c\_customer\_sk = s.ss\_customer\_sk

AND ca\_state IS NOT NULL

AND ss\_item\_sk = highPriceltems.i\_item\_sk --cannot use "ss\_item\_sk IN ()". Hive only supports a single "IN" subquery

per SQL statement.

AND s.ss\_sold\_date\_sk IN

( --during a given month

SELECT d\_date\_sk

FROM date\_dim

WHERE d\_year = 2004

AND d\_moy = 7

)

GROUP BY ca\_state

HAVING cnt >= 10 --at least 10 customers

ORDER BY cnt DESC, ca\_state --top 10 states in descending order

LIMIT 10

TOK\_QUERY 6, 0,178, 2

TOK\_FROM 6, 26,53, 2

TOK\_JOIN 6, 30,53, 2

TOK\_JOIN 6, 40,40, 2

TOK\_JOIN 6, 33,33, 2

TOK\_TABREF 6, 30,32, 2

TOK\_TABNAME 6, 30,30, 2

customer\_address 6, 30,30, 2

a 6, 32,32, 19

TOK\_TABREF 7, 37,39, 2

TOK\_TABNAME 7, 37,37, 2

customer 7, 37,37, 2

c 7, 39,39, 11

## Query 8

```
16/01/01 21:54:12 ERROR Utils: Uncaught exception in thread Thread-ScriptTransformation-Feed
java.lang.NullPointerException
    at org.apache.spark.unsafe.memory.TaskMemoryManager.getPage(TaskMemoryManager.java:235)
    at org.apache.spark.util.collection.unsafe.sort.
UnsafeInMemorySorter$SortedIterator.loadNext(UnsafeInMemorySorter.java:165)
    at org.apache.spark.sql.execution.UnsafeExternalRowSorter$1.next(UnsafeExternalRowSorter.java:142)
    at org.apache.spark.sql.execution.UnsafeExternalRowSorter$1.next(UnsafeExternalRowSorter.java:129)
    at scala.collection.Iterator$$anon$11.next(Iterator.scala:328)
    at scala.collection.Iterator$$anon$11.next(Iterator.scala:328)
    at scala.collection.Iterator$class.foreach(Iterator.scala:727)
    at scala.collection.AbstractIterator.foreach(Iterator.scala:1157)
    at
org.apache.spark.sql.hive.execution.ScriptTransformationWriterThread$$anonfun$run$1.apply$mcV$sp(ScriptTransf
ormation.scala:255)
    at org.apache.spark.sql.hive.execution.
ScriptTransformationWriterThread$$anonfun$run$1.apply(ScriptTransformation.scala:244)
    at org.apache.spark.sql.hive.execution.
ScriptTransformationWriterThread$$anonfun$run$1.apply(ScriptTransformation.scala:244)
    at org.apache.spark.util.Utils$.logUncaughtExceptions(Utils.scala:1699)
    at org.apache.spark.sql.hive.execution.
ScriptTransformationWriterThread.run(ScriptTransformation.scala:244)
16/01/01 21:54:12 ERROR ScriptTransformationWriterThread:
16/01/01 21:54:12 INFO Executor: Finished task 10.0 in stage 155.0 (TID 220). 2088 bytes result sent to driver
16/01/01 21:54:12 ERROR Utils: Uncaught exception in thread Thread-ScriptTransformation-Feed
java.lang.NullPointerException
    at org.apache.spark.unsafe.memory.TaskMemoryManager.getPage(TaskMemoryManager.java:235)
    at org.apache.spark.util.collection.unsafe.sort.
UnsafeInMemorySorter$SortedIterator.loadNext(UnsafeInMemorySorter.java:165)
    at org.apache.spark.sql.execution.UnsafeExternalRowSorter$1.next(UnsafeExternalRowSorter.java:142)
    at org.apache.spark.sql.execution.UnsafeExternalRowSorter$1.next(UnsafeExternalRowSorter.java:129)
    at scala.collection.Iterator$$anon$11.next(Iterator.scala:328)
    at scala.collection.Iterator$$anon$11.next(Iterator.scala:328)
    at scala.collection.Iterator$class.foreach(Iterator.scala:727)
    at scala.collection.AbstractIterator.foreach(Iterator.scala:1157)
    at org.apache.spark.sql.hive.execution.
ScriptTransformationWriterThread$$anonfun$run$1.apply$mcV$sp(ScriptTransformation.scala:255)
    at org.apache.spark.sql.hive.execution.
ScriptTransformationWriterThread$$anonfun$run$1.apply(ScriptTransformation.scala:244)
    at org.apache.spark.sql.hive.execution.
ScriptTransformationWriterThread$$anonfun$run$1.apply(ScriptTransformation.scala:244)
    at org.apache.spark.util.Utils$.logUncaughtExceptions(Utils.scala:1699)
    at org.apache.spark.sql.hive.execution.
ScriptTransformationWriterThread.run(ScriptTransformation.scala:244)
16/01/01 21:54:12 ERROR ScriptTransformationWriterThread:
16/01/01 21:54:12 INFO Executor: Finished task 44.0 in stage 155.0 (TID 254). 2088 bytes result sent to driver
```

