Assignment-12-CSP-554

Exercise 1

Read the article "A Big Data Modeling Methodology for Apache Cassandra" available on the blackboard in the 'Articles' section. Provide a ½ page summary including your comments and impressions.

Summary:

The paper covers traditional data modeling, Cassandra data modeling, conceptual and logical data modeling, application workflow, query-driven mapping from a conceptual to a logical data model, and physical data modeling.

Cassandra Data Model:

A CQL table can be considered a grouping of divisions containing rows with similar structures. A partition key is distinct from each partition in a table, whereas a clustering key is distinct from each row within a partition. A primary key is a combination of a partition key and a clustering key that uniquely identifies a database row. A table schema is a collection of columns that contains a primary key. Each column's data type is either primitive (int, text, etc.), complex (set, list, or map), or counter. CQL, which has a syntax similar to SQL, is used to express queries over tables. CQL does not support binary operations like joins and instead relies on a set of query predicates rules to ensure efficiency and scalability.

Conceptual data modeling and application workflow:

Understanding the data to be maintained and how a data-driven application needs to access it is required when designing a Cassandra database schema. The ER diagram depicts the former Application workflow diagrams, which define data access patterns for application tasks and capture the latter.

Query driven mapping Data Modeling Principles:

The four data modeling principles listed below serve as a foundation for translating conceptual data models into logical data models.

DMP1 (Know your data): The first step in successful database design is to understand the data, which is recorded using a conceptual data model.

DMP2 (Know your Questions): Knowing your queries captured by an application process is the second key to a successful database design.

DMP3 (Data Nesting): Data nesting is the third key to a successful database design.

DMP4 (Data Duplication): Data duplication is the fourth key to a successful database design. Mapping Rule: - The following are five mapping rules that facilitate a query-driven transition from a conceptual data model to a logical data model.

MR1 -> (Entities and Relationships): In MR1, entities and relationships map to table rows, whereas entity and relationship types of the map to tables

MR2 -> (Equality Search Attributes): In a query predicate, equality search attributes correspond to the prefix columns of a table's primary key.

MR3 -> (Inequality Search Attributes): A key column in a table clustering corresponds to an inequality search attribute used in a query predicate.

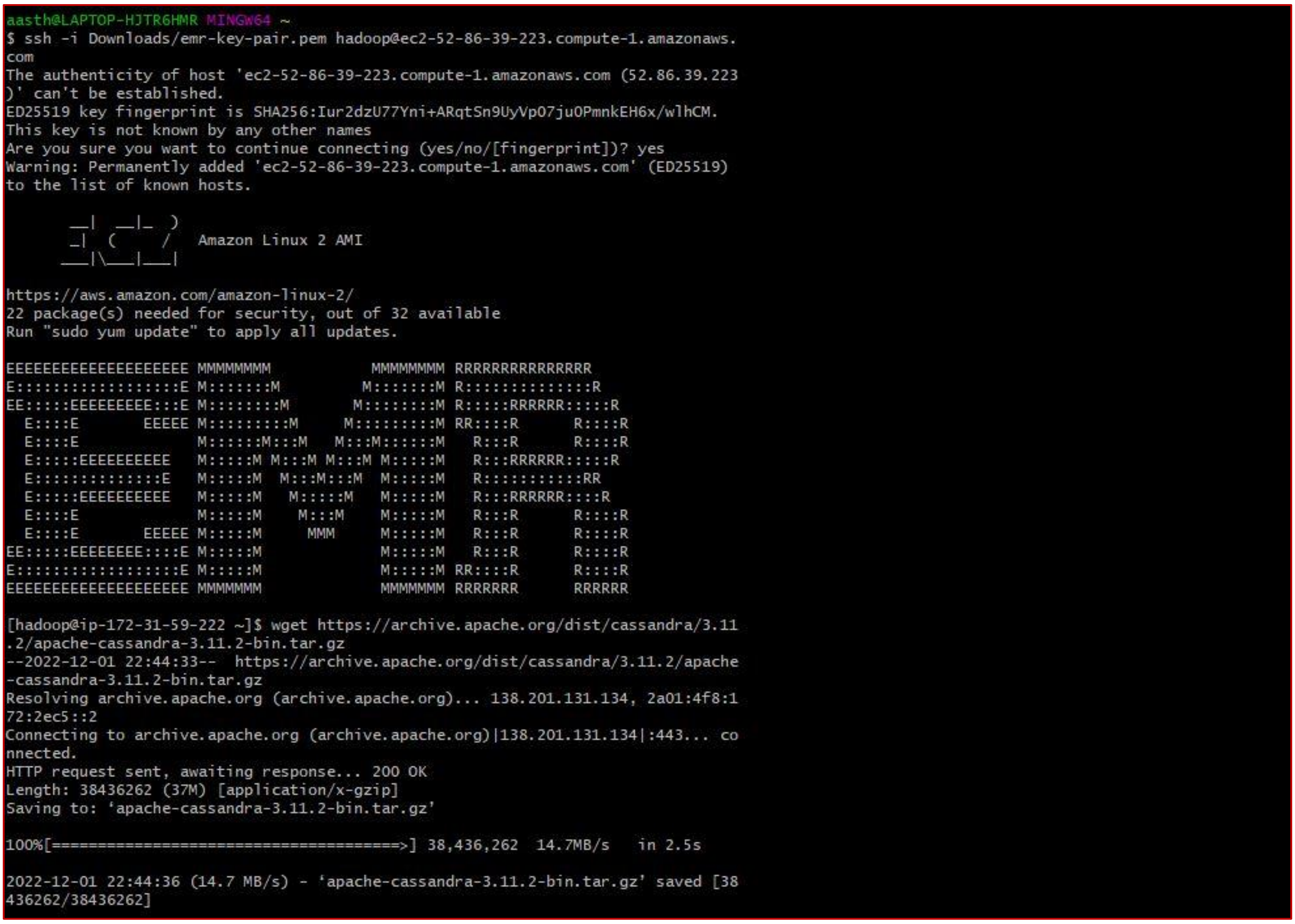
MR4 -> (Ordering Attributes): Ordering attributes, which are supplied in a query, map to clustering key columns in the query's chosen ascending or descending clustering order. MR5 -> (Key Attributes): Primary key columns are mapped to key attribute types.

Mapping Patterns: Mapping Patterns are used to automate Cassandra database schema design. Physical Data Modeling: The final step is to analyze and optimize a logical data model in a physical data model.

Exercise 2

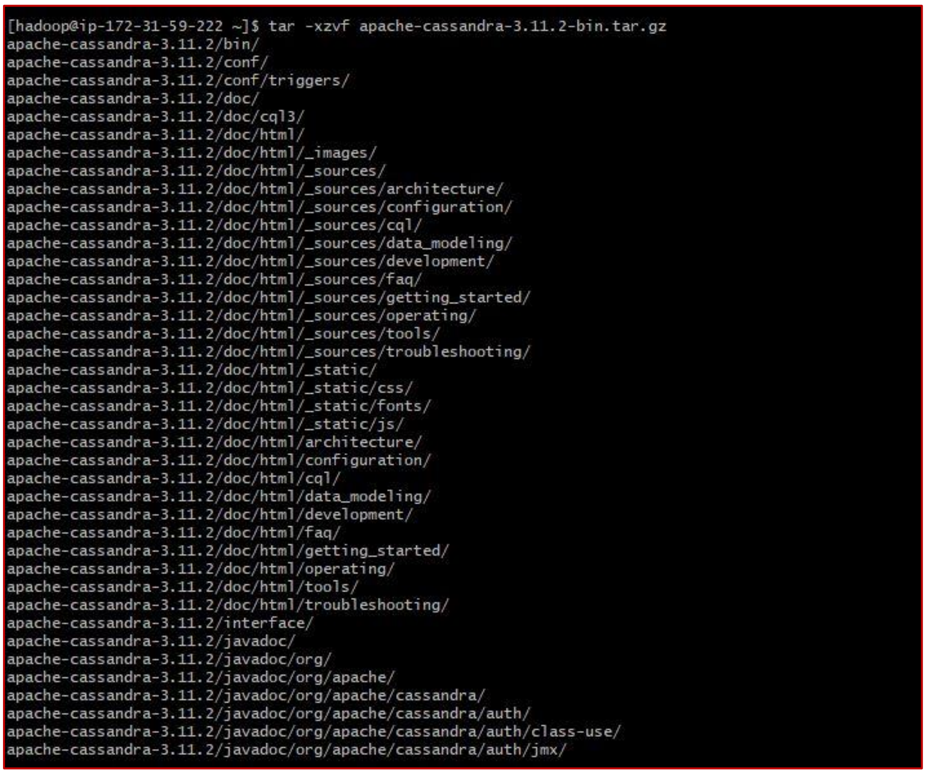
Using command

wget https://archive.apache.org/dist/cassandra/3.11.2/apache-cassandra-3.11.2-bin.tar.gz



```
aasth@LAPTOP-HJTR6HMR MINGW64 ~
$ ssh -i Downloads/emr-key-pair.pem hadoop@ec2-52-86-39-223.compute-1.amazonaws.
com
The authenticity of host 'ec2-52-86-39-223.compute-1.amazonaws.com (52.86.39.223
)' can't be established.
ED25519 key fingerprint is SHA256:Iur2dzU77Yni+ARqtSn9UyVp07ju0PmnkEH6x/wlhCM.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'ec2-52-86-39-223.compute-1.amazonaws.com' (ED25519)
to the list of known hosts.

       __|  __|_  )
       _|  (     /   Amazon Linux 2 AMI
      ___|\___|___|

https://aws.amazon.com/amazon-linux-2/
22 package(s) needed for security, out of 32 available
Run "sudo yum update" to apply all updates.

EEEEEEEEEEEEEEEEEEEE MMMMMMMM           MMMMMMMM RRRRRRRRRRRRRRR
E::::::::::::::::::E M:::::::M         M:::::::M R::::::::::::::R
EE:::::EEEEEEEEE:::E M::::::::M       M::::::::M R:::::RRRRRR:::::R
  E::::E       EEEEE M:::::::::M     M:::::::::M RR::::R      R::::R
  E::::E             M::::::M:::M   M:::M::::::M   R:::R      R::::R
  E:::::EEEEEEEEEE   M:::::M M:::M M:::M M:::::M   R:::RRRRRR:::::R
  E::::::::::::::E   M:::::M  M:::M:::M  M:::::M   R:::::::::::RR
  E:::::EEEEEEEEEE   M:::::M   M:::::M   M:::::M   R:::RRRRRR::::R
  E::::E             M:::::M    M:::M    M:::::M   R:::R      R::::R
  E::::E       EEEEE M:::::M     MMM     M:::::M   R:::R      R::::R
EE:::::EEEEEEEE::::E M:::::M             M:::::M   R:::R      R::::R
E::::::::::::::::::E M:::::M             M:::::M RR::::R      R::::R
EEEEEEEEEEEEEEEEEEEE MMMMMMM             MMMMMMM RRRRRRR      RRRRRR

[hadoop@ip-172-31-59-222 ~]$ wget https://archive.apache.org/dist/cassandra/3.11
.2/apache-cassandra-3.11.2-bin.tar.gz
--2022-12-01 22:44:33--  https://archive.apache.org/dist/cassandra/3.11.2/apache
-cassandra-3.11.2-bin.tar.gz
Resolving archive.apache.org (archive.apache.org)... 138.201.131.134, 2a01:4f8:1
72:2ec5::2
Connecting to archive.apache.org (archive.apache.org)|138.201.131.134|:443... co
nnected.
HTTP request sent, awaiting response... 200 OK
Length: 38436262 (37M) [application/x-gzip]
Saving to: ‘apache-cassandra-3.11.2-bin.tar.gz’

100%[======================================>] 38,436,262  14.7MB/s   in 2.5s

2022-12-01 22:44:36 (14.7 MB/s) - ‘apache-cassandra-3.11.2-bin.tar.gz’ saved [38
436262/38436262]
```

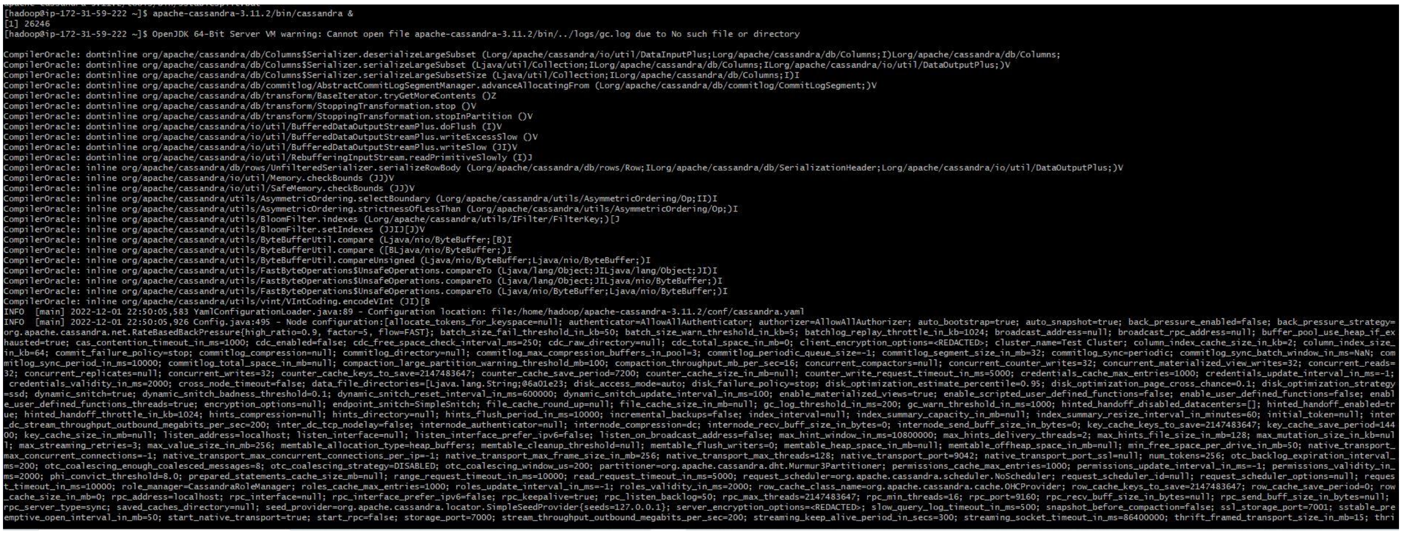
tar -xzvf apache-cassandra-3.11.2-bin.tar.gz



```
[hadoop@ip-172-31-59-222 ~]$ tar -xzvf apache-cassandra-3.11.2-bin.tar.gz
apache-cassandra-3.11.2/bin/
apache-cassandra-3.11.2/conf/
apache-cassandra-3.11.2/conf/triggers/
apache-cassandra-3.11.2/doc/
apache-cassandra-3.11.2/doc/cql3/
apache-cassandra-3.11.2/doc/html/
apache-cassandra-3.11.2/doc/html/_images/
apache-cassandra-3.11.2/doc/html/_sources/
apache-cassandra-3.11.2/doc/html/_sources/architecture/
apache-cassandra-3.11.2/doc/html/_sources/configuration/
apache-cassandra-3.11.2/doc/html/_sources/cql/
apache-cassandra-3.11.2/doc/html/_sources/data_modeling/
apache-cassandra-3.11.2/doc/html/_sources/development/
apache-cassandra-3.11.2/doc/html/_sources/faq/
apache-cassandra-3.11.2/doc/html/_sources/getting_started/
apache-cassandra-3.11.2/doc/html/_sources/operating/
apache-cassandra-3.11.2/doc/html/_sources/tools/
apache-cassandra-3.11.2/doc/html/_sources/troubleshooting/
apache-cassandra-3.11.2/doc/html/_static/
apache-cassandra-3.11.2/doc/html/_static/css/
apache-cassandra-3.11.2/doc/html/_static/fonts/
apache-cassandra-3.11.2/doc/html/_static/js/
apache-cassandra-3.11.2/doc/html/architecture/
apache-cassandra-3.11.2/doc/html/configuration/
apache-cassandra-3.11.2/doc/html/cql/
apache-cassandra-3.11.2/doc/html/data_modeling/
apache-cassandra-3.11.2/doc/html/development/
apache-cassandra-3.11.2/doc/html/faq/
apache-cassandra-3.11.2/doc/html/getting_started/
apache-cassandra-3.11.2/doc/html/operating/
apache-cassandra-3.11.2/doc/html/tools/
apache-cassandra-3.11.2/doc/html/troubleshooting/
apache-cassandra-3.11.2/interface/
apache-cassandra-3.11.2/javadoc/
apache-cassandra-3.11.2/javadoc/org/
apache-cassandra-3.11.2/javadoc/org/apache/
apache-cassandra-3.11.2/javadoc/org/apache/cassandra/
apache-cassandra-3.11.2/javadoc/org/apache/cassandra/auth/
apache-cassandra-3.11.2/javadoc/org/apache/cassandra/auth/class-use/
apache-cassandra-3.11.2/javadoc/org/apache/cassandra/auth/jmx/
```

Note, this will create a new directory (apache-cassandra-3.11.2) holding the Cassandra software release. Then enter this command to start Cassandra (lots of diagnostic messages will appear):

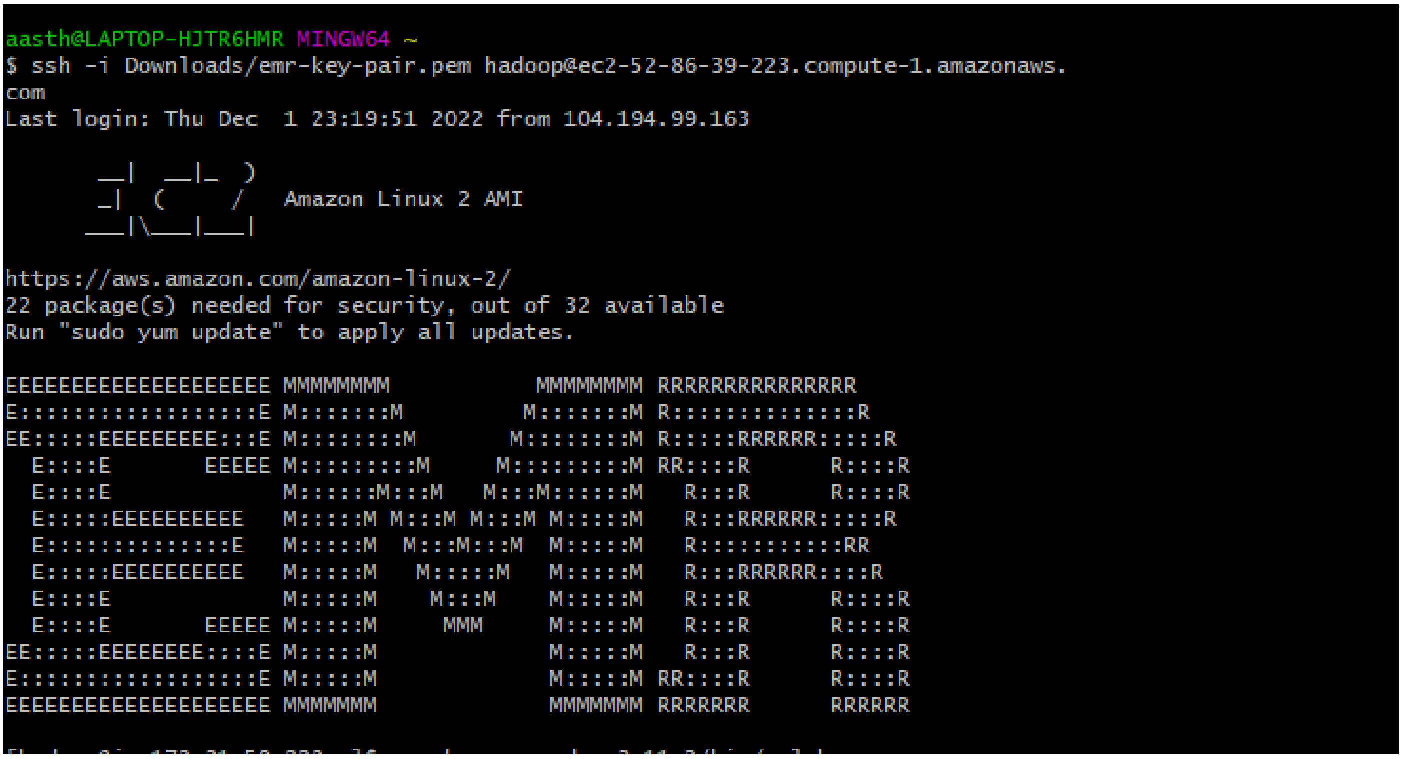
apache-cassandra-3.11.2/bin/cassandra &



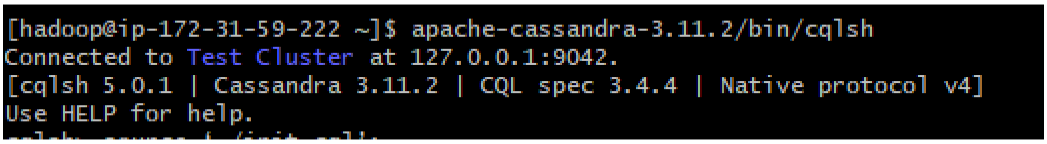
```
[hadoop@ip-172-31-59-222 ~]$ apache-cassandra-3.11.2/bin/cassandra &
[1] 26246
[hadoop@ip-172-31-59-222 ~]$ OpenJDK 64-Bit Server VM warning: Cannot open file apache-cassandra-3.11.2/bin/../logs/gc.log due to No such file or directory

CompilerOracle: dontinline org/apache/cassandra/db/Columns$Serializer.deserializeLargeSubset (Lorg/apache/cassandra/io/util/DataInputPlus;Lorg/apache/cassandra/db/Columns;I)Lorg/apache/cassandra/db/Columns;
CompilerOracle: dontinline org/apache/cassandra/db/Columns$Serializer.serializeLargeSubset (Ljava/util/Collection;ILorg/apache/cassandra/db/Columns;ILorg/apache/cassandra/io/util/DataOutputPlus;)V
CompilerOracle: dontinline org/apache/cassandra/db/Columns$Serializer.serializeLargeSubsetSize (Ljava/util/Collection;ILorg/apache/cassandra/db/Columns;I)I
CompilerOracle: dontinline org/apache/cassandra/db/commitlog/AbstractCommitLogSegmentManager.advanceAllocatingFrom (Lorg/apache/cassandra/db/commitlog/CommitLogSegment;)V
CompilerOracle: dontinline org/apache/cassandra/db/transform/BaseIterator.tryGetMoreContents ()Z
CompilerOracle: dontinline org/apache/cassandra/db/transform/StoppingTransformation.stop ()V
CompilerOracle: dontinline org/apache/cassandra/db/transform/StoppingTransformation.stopInPartition ()V
CompilerOracle: dontinline org/apache/cassandra/io/util/BufferedDataOutputStreamPlus.doFlush (I)V
CompilerOracle: dontinline org/apache/cassandra/io/util/BufferedDataOutputStreamPlus.writeExcessSlow ()V
CompilerOracle: dontinline org/apache/cassandra/io/util/BufferedDataOutputStreamPlus.writeSlow (JI)V
CompilerOracle: dontinline org/apache/cassandra/io/util/RebufferingInputStream.readPrimitiveSlowly (I)J
CompilerOracle: inline org/apache/cassandra/db/rows/UnfilteredSerializer.serializeRowBody (Lorg/apache/cassandra/db/rows/Row;ILorg/apache/cassandra/db/SerializationHeader;Lorg/apache/cassandra/io/util/DataOutputPlus;)V
CompilerOracle: inline org/apache/cassandra/io/util/Memory.checkBounds (JJ)V
CompilerOracle: inline org/apache/cassandra/io/util/SafeMemory.checkBounds (JJ)V
CompilerOracle: inline org/apache/cassandra/utils/AsymmetricOrdering.selectBoundary (Lorg/apache/cassandra/utils/AsymmetricOrdering/Op;II)I
CompilerOracle: inline org/apache/cassandra/utils/AsymmetricOrdering.strictnessOfLessThan (Lorg/apache/cassandra/utils/AsymmetricOrdering/Op;)I
CompilerOracle: inline org/apache/cassandra/utils/BloomFilter.indexes (Lorg/apache/cassandra/utils/IFilter/FilterKey;)[J
CompilerOracle: inline org/apache/cassandra/utils/BloomFilter.setIndexes (JJIJ[J)V
CompilerOracle: inline org/apache/cassandra/utils/ByteBufferUtil.compare (Ljava/nio/ByteBuffer;[B)I
CompilerOracle: inline org/apache/cassandra/utils/ByteBufferUtil.compare ([BLjava/nio/ByteBuffer;)I
CompilerOracle: inline org/apache/cassandra/utils/ByteBufferUtil.compareUnsigned (Ljava/nio/ByteBuffer;Ljava/nio/ByteBuffer;)I
CompilerOracle: inline org/apache/cassandra/utils/FastByteOperations$UnsafeOperations.compareTo (Ljava/lang/Object;JILjava/lang/Object;JI)I
CompilerOracle: inline org/apache/cassandra/utils/FastByteOperations$UnsafeOperations.compareTo (Ljava/lang/Object;JILjava/nio/ByteBuffer;)I
CompilerOracle: inline org/apache/cassandra/utils/FastByteOperations$UnsafeOperations.compareTo (Ljava/nio/ByteBuffer;Ljava/nio/ByteBuffer;)I
CompilerOracle: inline org/apache/cassandra/utils/vint/VIntCoding.encodeVInt (JI)[B
INFO  [main] 2022-12-01 22:50:05,583 YamlConfigurationLoader.java:89 - Configuration location: file:/home/hadoop/apache-cassandra-3.11.2/conf/cassandra.yaml
INFO  [main] 2022-12-01 22:50:05,926 Config.java:495 - Node configuration:[allocate_tokens_for_keyspace=null; authenticator=AllowAllAuthenticator; authorizer=AllowAllAuthorizer; auto_bootstrap=true; auto_snapshot=true; back_pressure_enabled=false; back_pressure_strategy=
org.apache.cassandra.net.RateBasedBackPressure{high_ratio=0.9, factor=5, flow=FAST}; batch_size_fail_threshold_in_kb=50; batch_size_warn_threshold_in_kb=5; batchlog_replay_throttle_in_kb=1024; broadcast_address=null; broadcast_rpc_address=null; buffer_pool_use_heap_if_ex
hausted=true; cas_contention_timeout_in_ms=1000; cdc_enabled=false; cdc_free_space_check_interval_ms=250; cdc_raw_directory=null; cdc_total_space_in_mb=0; client_encryption_options=<REDACTED>; cluster_name=Test Cluster; column_index_cache_size_in_kb=2; column_index_size_
in_kb=64; commit_failure_policy=stop; commitlog_compression=null; commitlog_directory=null; commitlog_max_compression_buffers_in_pool=3; commitlog_periodic_queue_size=-1; commitlog_segment_size_in_mb=32; commitlog_sync=periodic; commitlog_sync_batch_window_in_ms=NaN; com
mitlog_sync_period_in_ms=10000; commitlog_total_space_in_mb=null; compaction_large_partition_warning_threshold_mb=100; compaction_throughput_mb_per_sec=16; concurrent_compactors=null; concurrent_counter_writes=32; concurrent_materialized_view_writes=32; concurrent_reads=
32; concurrent_replicates=null; concurrent_writes=32; counter_cache_keys_to_save=2147483647; counter_cache_save_period=7200; counter_cache_size_in_mb=null; counter_write_request_timeout_in_ms=5000; credentials_cache_max_entries=1000; credentials_update_interval_in_ms=-1;
 credentials_validity_in_ms=2000; cross_node_timeout=false; data_file_directories=[Ljava.lang.String;@6a01e23; disk_access_mode=auto; disk_failure_policy=stop; disk_optimization_estimate_percentile=0.95; disk_optimization_page_cross_chance=0.1; disk_optimization_strategy
=ssd; dynamic_snitch=true; dynamic_snitch_badness_threshold=0.1; dynamic_snitch_reset_interval_in_ms=600000; dynamic_snitch_update_interval_in_ms=100; enable_materialized_views=true; enable_scripted_user_defined_functions=false; enable_user_defined_functions=false; enabl
e_user_defined_functions_threads=true; encryption_options=null; endpoint_snitch=SimpleSnitch; file_cache_round_up=null; file_cache_size_in_mb=null; gc_log_threshold_in_ms=200; gc_warn_threshold_in_ms=1000; hinted_handoff_disabled_datacenters=[]; hinted_handoff_enabled=tr
ue; hinted_handoff_throttle_in_kb=1024; hints_compression=null; hints_directory=null; hints_flush_period_in_ms=10000; incremental_backups=false; index_interval=null; index_summary_capacity_in_mb=null; index_summary_resize_interval_in_minutes=60; initial_token=null; inter
_dc_stream_throughput_outbound_megabits_per_sec=200; inter_dc_tcp_nodelay=false; internode_authenticator=null; internode_compression=dc; internode_recv_buff_size_in_bytes=0; internode_send_buff_size_in_bytes=0; key_cache_keys_to_save=2147483647; key_cache_save_period=144
00; key_cache_size_in_mb=null; listen_address=localhost; listen_interface=null; listen_interface_prefer_ipv6=false; listen_on_broadcast_address=false; max_hint_window_in_ms=10800000; max_hints_delivery_threads=2; max_hints_file_size_in_mb=128; max_mutation_size_in_kb=nul
l; max_streaming_retries=3; max_value_size_in_mb=256; memtable_allocation_type=heap_buffers; memtable_cleanup_threshold=null; memtable_flush_writers=0; memtable_heap_space_in_mb=null; memtable_offheap_space_in_mb=null; min_free_space_per_drive_in_mb=50; native_transport_
max_concurrent_connections=-1; native_transport_max_concurrent_connections_per_ip=-1; native_transport_max_frame_size_in_mb=256; native_transport_max_threads=128; native_transport_port=9042; native_transport_port_ssl=null; num_tokens=256; otc_backlog_expiration_interval_
ms=200; otc_coalescing_enough_coalesced_messages=8; otc_coalescing_strategy=DISABLED; otc_coalescing_window_us=200; partitioner=org.apache.cassandra.dht.Murmur3Partitioner; permissions_cache_max_entries=1000; permissions_update_interval_in_ms=-1; permissions_validity_in_
ms=2000; phi_convict_threshold=8.0; prepared_statements_cache_size_mb=null; range_request_timeout_in_ms=10000; read_request_timeout_in_ms=5000; request_scheduler=org.apache.cassandra.scheduler.NoScheduler; request_scheduler_id=null; request_scheduler_options=null; reques
t_timeout_in_ms=10000; role_manager=CassandraRoleManager; roles_cache_max_entries=1000; roles_update_interval_in_ms=-1; roles_validity_in_ms=2000; row_cache_class_name=org.apache.cassandra.cache.OHCProvider; row_cache_keys_to_save=2147483647; row_cache_save_period=0; row
_cache_size_in_mb=0; rpc_address=localhost; rpc_interface=null; rpc_interface_prefer_ipv6=false; rpc_keepalive=true; rpc_listen_backlog=50; rpc_max_threads=2147483647; rpc_min_threads=16; rpc_port=9160; rpc_recv_buff_size_in_bytes=null; rpc_send_buff_size_in_bytes=null;
rpc_server_type=sync; saved_caches_directory=null; seed_provider=org.apache.cassandra.locator.SimpleSeedProvider{seeds=127.0.0.1}; server_encryption_options=<REDACTED>; slow_query_log_timeout_in_ms=500; snapshot_before_compaction=false; ssl_storage_port=7001; sstable_pre
emptive_open_interval_in_mb=50; start_native_transport=true; start_rpc=false; storage_port=7000; stream_throughput_outbound_megabits_per_sec=200; streaming_keep_alive_period_in_secs=300; streaming_socket_timeout_in_ms=86400000; thrift_framed_transport_size_in_mb=15; thri
```

Open a second terminal connection to the EMR master node. Going forward we will call this terminal connection: Cqlsh-Term. Enter the following into this terminal to start the command line interface csqlsh:

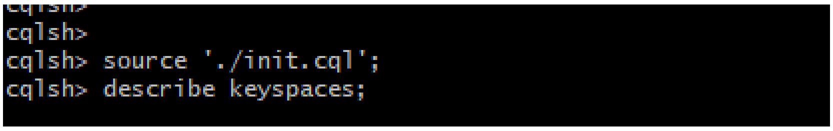
apache-cassandra-3.11.2/bin/cqlsh



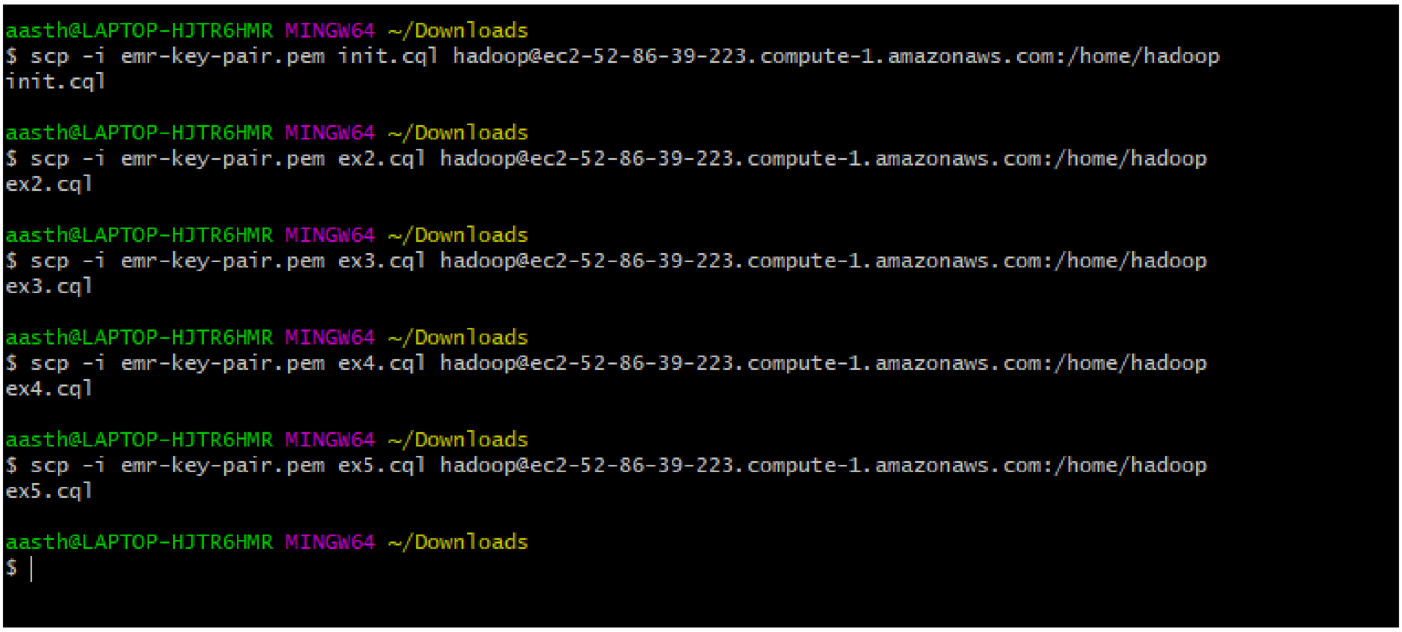
```
aasth@LAPTOP-HJTR6HMR MINGW64 ~
$ ssh -i Downloads/emr-key-pair.pem hadoop@ec2-52-86-39-223.compute-1.amazonaws.
com
Last login: Thu Dec  1 23:19:51 2022 from 104.194.99.163

       __|  __|_  )
       _|  (     /   Amazon Linux 2 AMI
      ___|\___|___|

https://aws.amazon.com/amazon-linux-2/
22 package(s) needed for security, out of 32 available
Run "sudo yum update" to apply all updates.

EEEEEEEEEEEEEEEEEEEE MMMMMMMM           MMMMMMMM RRRRRRRRRRRRRRR
E::::::::::::::::::E M:::::::M         M:::::::M R::::::::::::::R
EE:::::EEEEEEEEE:::E M::::::::M       M::::::::M R:::::RRRRRR:::::R
  E::::E       EEEEE M:::::::::M     M:::::::::M RR::::R      R::::R
  E::::E             M::::::M:::M   M:::M::::::M   R:::R      R::::R
  E:::::EEEEEEEEEE   M:::::M M:::M M:::M M:::::M   R:::RRRRRR:::::R
  E::::::::::::::E   M:::::M  M:::M:::M  M:::::M   R:::::::::::RR
  E:::::EEEEEEEEEE   M:::::M   M:::::M   M:::::M   R:::RRRRRR::::R
  E::::E             M:::::M    M:::M    M:::::M   R:::R      R::::R
  E::::E       EEEEE M:::::M     MMM     M:::::M   R:::R      R::::R
EE:::::EEEEEEEE::::E M:::::M             M:::::M   R:::R      R::::R
E::::::::::::::::::E M:::::M             M:::::M RR::::R      R::::R
EEEEEEEEEEEEEEEEEEEE MMMMMMM             MMMMMMM RRRRRRR      RRRRRR
```



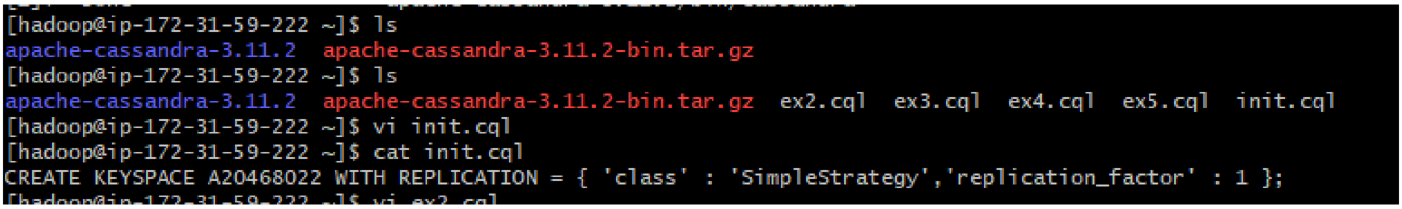
```
[hadoop@ip-172-31-59-222 ~]$ apache-cassandra-3.11.2/bin/cqlsh
Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.2 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
```



```
cqlsh>
cqlsh> source './init.cql';
cqlsh> describe keyspaces;
```



```
aasth@LAPTOP-HJTR6HMR MINGW64 ~/Downloads
$ scp -i emr-key-pair.pem init.cql hadoop@ec2-52-86-39-223.compute-1.amazonaws.com:/home/hadoop
init.cql

aasth@LAPTOP-HJTR6HMR MINGW64 ~/Downloads
$ scp -i emr-key-pair.pem ex2.cql hadoop@ec2-52-86-39-223.compute-1.amazonaws.com:/home/hadoop
ex2.cql

aasth@LAPTOP-HJTR6HMR MINGW64 ~/Downloads
$ scp -i emr-key-pair.pem ex3.cql hadoop@ec2-52-86-39-223.compute-1.amazonaws.com:/home/hadoop
ex3.cql

aasth@LAPTOP-HJTR6HMR MINGW64 ~/Downloads
$ scp -i emr-key-pair.pem ex4.cql hadoop@ec2-52-86-39-223.compute-1.amazonaws.com:/home/hadoop
ex4.cql

aasth@LAPTOP-HJTR6HMR MINGW64 ~/Downloads
$ scp -i emr-key-pair.pem ex5.cql hadoop@ec2-52-86-39-223.compute-1.amazonaws.com:/home/hadoop
ex5.cql

aasth@LAPTOP-HJTR6HMR MINGW64 ~/Downloads
$ |
```

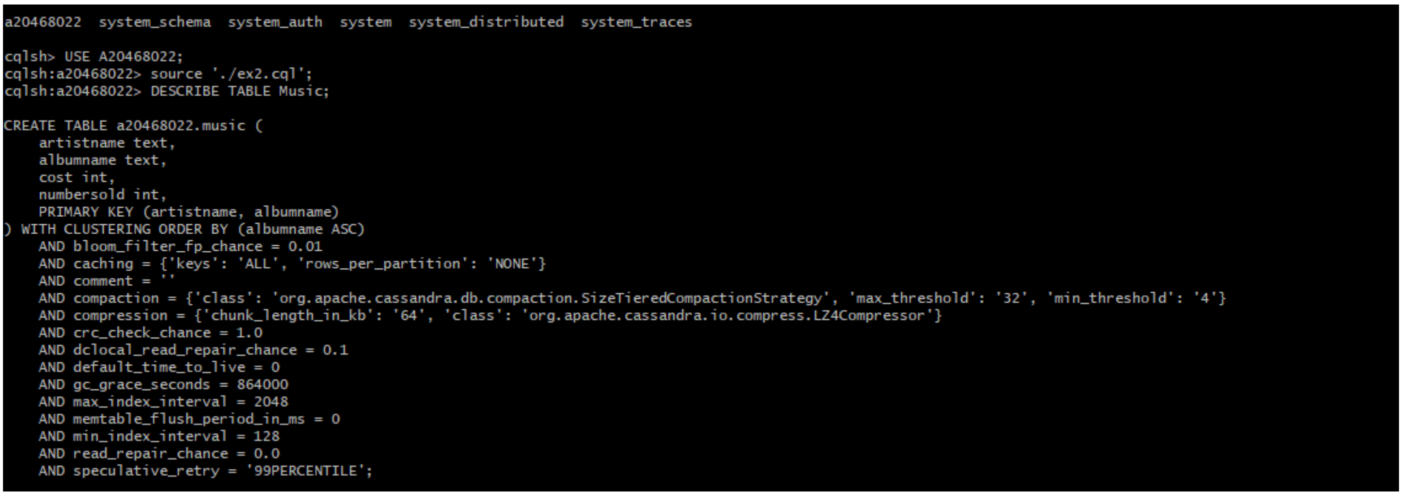


```
[hadoop@ip-172-31-59-222 ~]$ ls
apache-cassandra-3.11.2  apache-cassandra-3.11.2-bin.tar.gz
[hadoop@ip-172-31-59-222 ~]$ ls
apache-cassandra-3.11.2  apache-cassandra-3.11.2-bin.tar.gz  ex2.cql  ex3.cql  ex4.cql  ex5.cql  init.cql
[hadoop@ip-172-31-59-222 ~]$ vi init.cql
[hadoop@ip-172-31-59-222 ~]$ cat init.cql
CREATE KEYSPACE A20468022 WITH REPLICATION = { 'class' : 'SimpleStrategy','replication_factor' : 1 };
[hadoop@ip-172-31-59-222 ~]$ vi ex2.cql
```

Execute the below command:

USE A20468022;

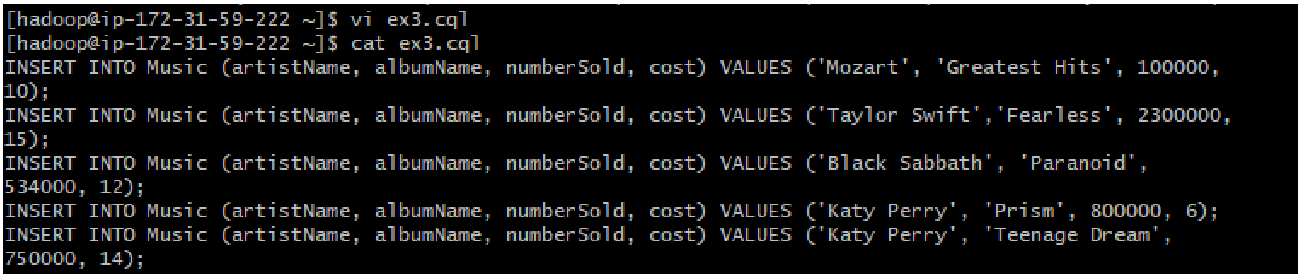
source './ex2.cql'; DESCRIBE TABLE Music;



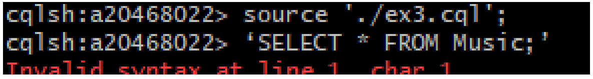
```
a20468022  system_schema  system_auth  system  system_distributed  system_traces

cqlsh> USE A20468022;
cqlsh:a20468022> source './ex2.cql';
cqlsh:a20468022> DESCRIBE TABLE Music;

CREATE TABLE a20468022.music (
    artistname text,
    albumname text,
    cost int,
    numbersold int,
    PRIMARY KEY (artistname, albumname)
) WITH CLUSTERING ORDER BY (albumname ASC)
    AND bloom_filter_fp_chance = 0.01
    AND caching = {'keys': 'ALL', 'rows_per_partition': 'NONE'}
    AND comment = ''
    AND compaction = {'class': 'org.apache.cassandra.db.compaction.SizeTieredCompactionStrategy', 'max_threshold': '32', 'min_threshold': '4'}
    AND compression = {'chunk_length_in_kb': '64', 'class': 'org.apache.cassandra.io.compress.LZ4Compressor'}
    AND crc_check_chance = 1.0
    AND dclocal_read_repair_chance = 0.1
    AND default_time_to_live = 0
    AND gc_grace_seconds = 864000
    AND max_index_interval = 2048
    AND memtable_flush_period_in_ms = 0
    AND min_index_interval = 128
    AND read_repair_chance = 0.0
    AND speculative_retry = '99PERCENTILE';
```

Exercise 3) (3 points)

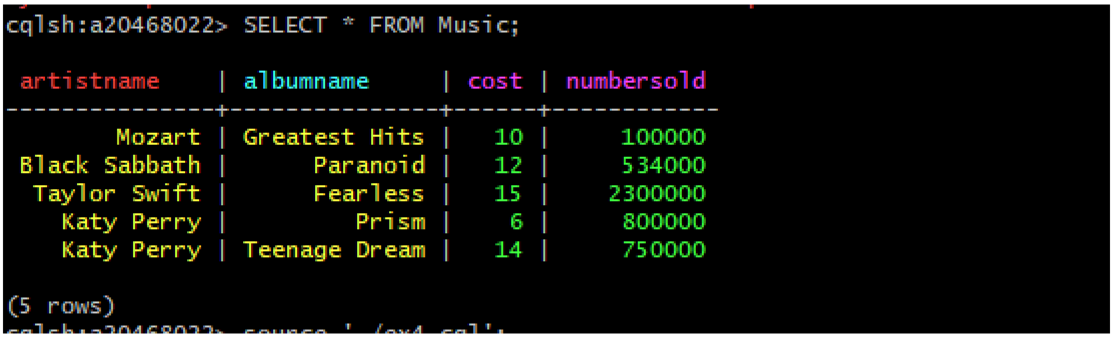
a) Execute ex3.cql. Provide the content of this file as the result of this exercise.



```
[hadoop@ip-172-31-59-222 ~]$ vi ex3.cql
[hadoop@ip-172-31-59-222 ~]$ cat ex3.cql
INSERT INTO Music (artistName, albumName, numberSold, cost) VALUES ('Mozart', 'Greatest Hits', 100000,
10);
INSERT INTO Music (artistName, albumName, numberSold, cost) VALUES ('Taylor Swift','Fearless', 2300000,
15);
INSERT INTO Music (artistName, albumName, numberSold, cost) VALUES ('Black Sabbath', 'Paranoid',
534000, 12);
INSERT INTO Music (artistName, albumName, numberSold, cost) VALUES ('Katy Perry', 'Prism', 800000, 6);
INSERT INTO Music (artistName, albumName, numberSold, cost) VALUES ('Katy Perry', 'Teenage Dream',
750000, 14);
```

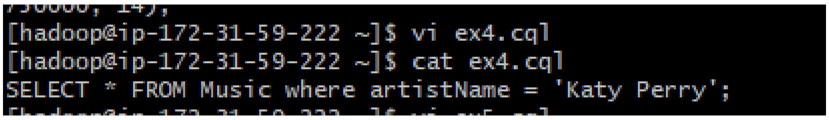


```
cqlsh:a20468022> source './ex3.cql';
cqlsh:a20468022> 'SELECT * FROM Music;'
Invalid syntax at line 1, char 1
```

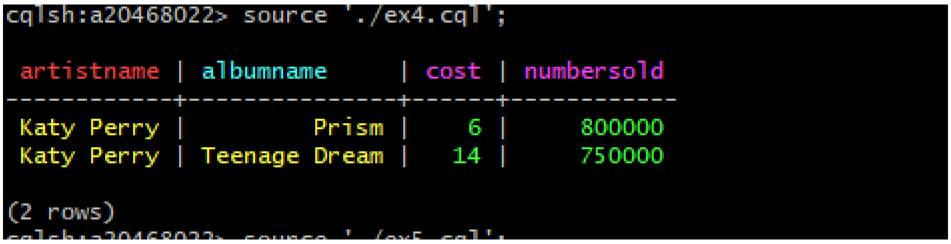


```
cqlsh:a20468022> SELECT * FROM Music;

 artistname    | albumname     | cost | numbersold
---------------+---------------+------+------------
        Mozart | Greatest Hits |   10 |     100000
 Black Sabbath |      Paranoid |   12 |     534000
  Taylor Swift |      Fearless |   15 |    2300000
    Katy Perry |         Prism |    6 |     800000
    Katy Perry | Teenage Dream |   14 |     750000

(5 rows)
```

Exercise 4) (2 points)

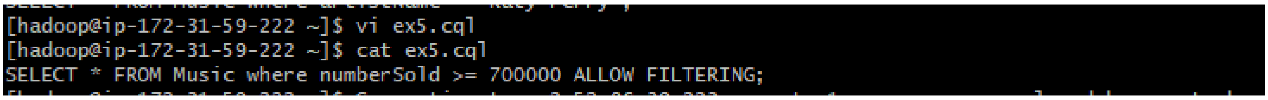


```
[hadoop@ip-172-31-59-222 ~]$ vi ex4.cql
[hadoop@ip-172-31-59-222 ~]$ cat ex4.cql
SELECT * FROM Music where artistName = 'Katy Perry';
```

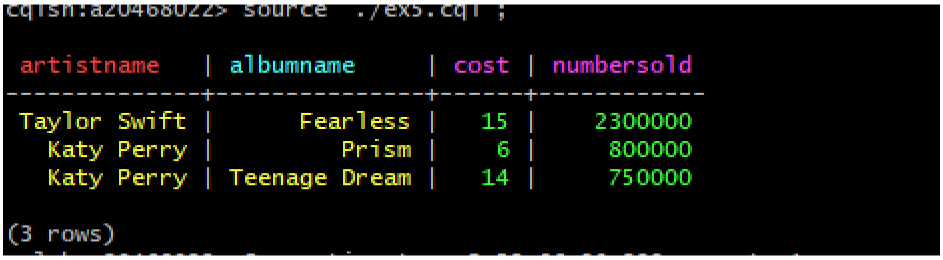


```
cqlsh:a20468022> source './ex4.cql';

 artistname | albumname     | cost | numbersold
------------+---------------+------+------------
 Katy Perry |         Prism |    6 |     800000
 Katy Perry | Teenage Dream |   14 |     750000

(2 rows)
```

Exercise 5) (2 points)



```
[hadoop@ip-172-31-59-222 ~]$ vi ex5.cql
[hadoop@ip-172-31-59-222 ~]$ cat ex5.cql
SELECT * FROM Music where numberSold >= 700000 ALLOW FILTERING;
```



```
cqlsh:a20468022> source './ex5.cql';

 artistname   | albumname     | cost | numbersold
--------------+---------------+------+------------
 Taylor Swift |      Fearless |   15 |    2300000
   Katy Perry |         Prism |    6 |     800000
   Katy Perry | Teenage Dream |   14 |     750000

(3 rows)
```

Submitted By: -

Aastha Dhir

CWID-A20468022

adhir2@hawk.iit.edu