Лабораторная работа по Hive.
Выполнила Бабушкина Татьяна

1. Запустим контейнер с установленным hive и командой docker ps смотрим на наши контейнеры.

```
(base) tatyana@tatyana-Inspiron-5559:~/Desktop/laba2$ docker ps
CONTAINER ID   IMAGE                                              COMMAND                  CREATED         STATUS                  PORTS                                              NAMES
6dd28edec209   cloudera/quickstart:latest                         "/usr/bin/docker-qui…"   10 minutes ago  Up 10 minutes           0.0.0.0:8888->8888/tcp                             compass
ionate_wing
ed3f36f6a2e3   bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8     "/entrypoint.sh /run…"   7 months ago    Up 43 minutes (healthy) 9864/tcp                                           datanod
e
4c877fb07c44   bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8  "/entrypoint.sh /run…"   7 months ago    Up 43 minutes (healthy) 8042/tcp                                           nodeman
ager
93c33bd129ce   bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8     "/entrypoint.sh /run…"   7 months ago    Up 43 minutes (healthy) 0.0.0.0:9000->9000/tcp, 0.0.0.0:9870->9870/tcp    namenod
e
0b1d3b23e33a   bde2020/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8  "/entrypoint.sh /run…"  7 months ago  Up 41 minutes (healthy) 8088/tcp                                       resourc
emanager
f2ac16187b00   bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8  "/entrypoint.sh /run…"   7 months ago   Up 43 minutes (healthy) 8188/tcp                                        history
server
```

2. Поместим данные в контейнер.

```
(base) tatyana@tatyana-Inspiron-5559:~/Desktop/laba2$ docker cp movies.dat 6dd28edec209:/volume/
(base) tatyana@tatyana-Inspiron-5559:~/Desktop/laba2$ docker cp ratings.dat 6dd28edec209:/volume/
(base) tatyana@tatyana-Inspiron-5559:~/Desktop/laba2$ docker cp users.dat 6dd28edec209:/volume/
```

3. Находим нужный контейнер (compassionate_wing) и переходим в него:

```
(base) tatyana@tatyana-Inspiron-5559:~/Desktop/laba2$ sudo docker exec -it compassionate_wing bash
[root@quickstart /]# ls
bin    etc   lib64       mnt           packer-files  sbin     sys  var
boot   home  lost+found  movies.t      proc          selinux  tmp  volume
dev    lib   media       opt           root          srv      usr
```

4. Конвертируем все данные.

```
[root@quickstart /]# cd volume
[root@quickstart volume]# ls
movies.dat  ratings.dat  users.dat
[root@quickstart volume]# sed 's/::/,/g' movies.dat > movies.t
[root@quickstart volume]# sed 's/::/,/g' ratings.dat > ratings.t
[root@quickstart volume]# sed 's/::/,/g' users.dat > users.t
[root@quickstart volume]# ls
movies.dat  movies.t  ratings.dat  ratings.t  users.dat  users.t
```

5. Создаем директории под наши данные.

```
[root@quickstart volume]# hdfs dfs -mkdir -p /moviedata/ratings
[root@quickstart volume]# hdfs dfs -mkdir -p /moviedata/users
[root@quickstart volume]# hdfs dfs -mkdir -p /moviedata/movies
```

6. Переносим данные в hdfs.

```
[root@quickstart volume]# hdfs dfs -put movies.t /moviedata/movies
[root@quickstart volume]# hdfs dfs -put users.t /moviedata/users
[root@quickstart volume]# hdfs dfs -put ratings.t /moviedata/ratings
```

7. Используем Beeline:

```
[root@quickstart volume]# beeline -u jdbc:hive2://quickstart:10000/default -n admin -d org.apache.hive.jdbc.HiveDriver
2020-10-30 17:52:36,009 WARN  [main] mapreduce.TableMapReduceUtil: The hbase-prefix-tree module jar containing PrefixTreeCodec is not present.  Continuing without it.
Connecting to jdbc:hive2://quickstart:10000/default
Connected to: Apache Hive (version 1.1.0-cdh5.7.0)
Driver: Hive JDBC (version 1.1.0-cdh5.7.0)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 1.1.0-cdh5.7.0 by Apache Hive
```

8. Создаем БД.

```
Beeline version 1.1.0-cdh5.7.0 by Apache Hive
0: jdbc:hive2://quickstart:10000/default> create database moviedata;
INFO  : Compiling command(queryId=hive_20201030175656_a0f9e3b9-f67d-4638-aa1a-05de80f55c4a): create database moviedata
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO  : Completed compiling command(queryId=hive_20201030175656_a0f9e3b9-f67d-4638-aa1a-05de80f55c4a); Time taken: 0.736 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20201030175656_a0f9e3b9-f67d-4638-aa1a-05de80f55c4a): create database moviedata
INFO  : Starting task [Stage-0:DDL] in serial mode
INFO  : Completed executing command(queryId=hive_20201030175656_a0f9e3b9-f67d-4638-aa1a-05de80f55c4a); Time taken: 0.65 seconds
INFO  : OK
No rows affected (1.674 seconds)
0: jdbc:hive2://quickstart:10000/default> use moviedata;
INFO  : Compiling command(queryId=hive_20201030175757_99caa96d-5962-4f3c-a78f-c858eaa60600): use moviedata
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO  : Completed compiling command(queryId=hive_20201030175757_99caa96d-5962-4f3c-a78f-c858eaa60600); Time taken: 0.214 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20201030175757_99caa96d-5962-4f3c-a78f-c858eaa60600): use moviedata
INFO  : Starting task [Stage-0:DDL] in serial mode
INFO  : Completed executing command(queryId=hive_20201030175757_99caa96d-5962-4f3c-a78f-c858eaa60600); Time taken: 0.006 seconds
INFO  : OK
No rows affected (0.247 seconds)
```

10. Создаем таблицы для наших данных.

```
3 rows selected (0.562 seconds)
0: jdbc:hive2://quickstart:10000/default> CREATE EXTERNAL TABLE movies (userid INT, movie STRING, category STRING) ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS TEXTFILE LOCATION '/moviedata/mov
ies';
INFO  : Compiling command(queryId=hive_20200406112121_9c8ee76b-136a-4eab-aef2-355861407c85): CREATE EXTERNAL TABLE movies (userid INT, movie STRING, category STRING) ROW FORMAT DELIMITED FIELDS TERMINATED
 BY ',' STORED AS TEXTFILE LOCATION '/moviedata/movies'
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO  : Completed compiling command(queryId=hive_20200406112121_9c8ee76b-136a-4eab-aef2-355861407c85); Time taken: 0.028 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20200406112121_9c8ee76b-136a-4eab-aef2-355861407c85): CREATE EXTERNAL TABLE movies (userid INT, movie STRING, category STRING) ROW FORMAT DELIMITED FIELDS TERMINATED
 BY ',' STORED AS TEXTFILE LOCATION '/moviedata/movies'
INFO  : Starting task [Stage-0:DDL] in serial mode
INFO  : Completed executing command(queryId=hive_20200406112121_9c8ee76b-136a-4eab-aef2-355861407c85); Time taken: 0.077 seconds
INFO  : OK
No rows affected (0.127 seconds)
0: jdbc:hive2://quickstart:10000/default> select * from movies limit 5;
INFO  : Compiling command(queryId=hive_20200406112121_456dd395-f237-4a4b-ba6d-890fb834eb01): select * from movies limit 5
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:movies.userid, type:int, comment:null), FieldSchema(name:movies.movie, type:string, comment:null), FieldSchema(name:movies.category, ty
pe:string, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20200406112121_456dd395-f237-4a4b-ba6d-890fb834eb01); Time taken: 0.136 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20200406112121_456dd395-f237-4a4b-ba6d-890fb834eb01): select * from movies limit 5
INFO  : Completed executing command(queryId=hive_20200406112121_456dd395-f237-4a4b-ba6d-890fb834eb01); Time taken: 0.0 seconds
INFO  : OK
+---------------+------------------------+-------------------------------+--+
| movies.userid |       movies.movie      |        movies.category        |
+---------------+------------------------+-------------------------------+--+
| 1             | Toy Story (1995)       | Animation|Children's|Comedy   |
| 2             | Jumanji (1995)         | Adventure|Children's|Fantasy  |
| 3             | Grumpier Old Men (1995)| Comedy|Romance                |
| 4             | Waiting to Exhale (1995)| Comedy|Drama                 |
```

```
0: jdbc:hive2://quickstart:10000/default> CREATE EXTERNAL TABLE ratings (userid INT, movieid INT, rating INT, tstamp STRING) ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS TEXTFILE LOCATION '/mov
iedata/ratings';
INFO  : Compiling command(queryId=hive_20200406111717_e4d97f6b-2310-4474-9bb0-dfc39a356f2c): CREATE EXTERNAL TABLE ratings (userid INT, movieid INT, rating INT, tstamp STRING) ROW FORMAT DELIMITED FIELDS
TERMINATED BY ',' STORED AS TEXTFILE LOCATION '/moviedata/ratings'
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO  : Completed compiling command(queryId=hive_20200406111717_e4d97f6b-2310-4474-9bb0-dfc39a356f2c); Time taken: 0.023 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20200406111717_e4d97f6b-2310-4474-9bb0-dfc39a356f2c): CREATE EXTERNAL TABLE ratings (userid INT, movieid INT, rating INT, tstamp STRING) ROW FORMAT DELIMITED FIELDS
TERMINATED BY ',' STORED AS TEXTFILE LOCATION '/moviedata/ratings'
INFO  : Starting task [Stage-0:DDL] in serial mode
INFO  : Completed executing command(queryId=hive_20200406111717_e4d97f6b-2310-4474-9bb0-dfc39a356f2c); Time taken: 0.767 seconds
INFO  : OK
No rows affected (0.818 seconds)
0: jdbc:hive2://quickstart:10000/default> select * from ratings limit 5;
INFO  : Compiling command(queryId=hive_20200406111717_0de20edf-9b37-4d1e-8ada-c670f7b31c4b): select * from ratings limit 5
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:ratings.userid, type:int, comment:null), FieldSchema(name:ratings.movieid, type:int, comment:null), FieldSchema(name:ratings.rating, ty
pe:int, comment:null), FieldSchema(name:ratings.tstamp, type:string, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20200406111717_0de20edf-9b37-4d1e-8ada-c670f7b31c4b); Time taken: 0.307 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20200406111717_0de20edf-9b37-4d1e-8ada-c670f7b31c4b): select * from ratings limit 5
INFO  : Completed executing command(queryId=hive_20200406111717_0de20edf-9b37-4d1e-8ada-c670f7b31c4b); Time taken: 0.001 seconds
INFO  : OK
+----------------+-----------------+----------------+----------------+--+
| ratings.userid | ratings.movieid | ratings.rating | ratings.tstamp |
+----------------+-----------------+----------------+----------------+--+
| 1              | 1193            | 5              | 978300760      |
| 1              | 661             | 3              | 978302109      |
| 1              | 914             | 3              | 978301968      |
| 1              | 3408            | 4              | 978300275      |
| 1              | 2355            | 5              | 978824291      |
+----------------+-----------------+----------------+----------------+--+
5 rows selected (0.562 seconds)
```

11. Пишем сами запросы:

```
0: jdbc:hive2://quickstart:10000/default> select * from movies m where m.category = 'Comedy' limit 10;
INFO  : Compiling command(queryId=hive_20200406120606_f9ddcd7a-01ed-4735-a7b7-be487719ed73): select * from movies m where m.category = 'Comedy' limit 10
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:m.userid, type:int, comment:null), FieldSchema(name:m.movie, type:string, comment:null), FieldSchema(name:m.category, type:string, comm
ent:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20200406120606_f9ddcd7a-01ed-4735-a7b7-be487719ed73); Time taken: 0.311 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20200406120606_f9ddcd7a-01ed-4735-a7b7-be487719ed73): select * from movies m where m.category = 'Comedy' limit 10
INFO  : Completed executing command(queryId=hive_20200406120606_f9ddcd7a-01ed-4735-a7b7-be487719ed73); Time taken: 0.0 seconds
INFO  : OK
+-----------+-----------------------------------------------------------------------------------+---------------+--+
| m.userid  |                                      m.movie                                      |  m.category   |
+-----------+-----------------------------------------------------------------------------------+---------------+--+
| 5         | Father of the Bride Part II (1995)                                                 | Comedy        |
| 19        | Ace Ventura: When Nature Calls (1995)                                              | Comedy        |
| 38        | It Takes Two (1995)                                                                | Comedy        |
| 52        | Mighty Aphrodite (1995)                                                            | Comedy        |
| 63        | Don't Be a Menace to South Central While Drinking Your Juice in the Hood (1996)    | Comedy        |
| 65        | Bio-Dome (1996)                                                                    | Comedy        |
| 69        | Friday (1995)                                                                      | Comedy        |
| 88        | Black Sheep (1996)                                                                 | Comedy        |
| 96        | In the Bleak Midwinter (1995)                                                      | Comedy        |
| 101       | Bottle Rocket (1996)                                                              | Comedy        |
+-----------+-----------------------------------------------------------------------------------+---------------+--+
10 rows selected (0.404 seconds)
```

```
0: jdbc:hive2://quickstart:10000/default> select ratings.movieid,max(ratings.rating) from ratings group by ratings.movieid limit 10;
INFO  : Compiling command(queryId=hive_20200406120303_c0a24f6e-f2d5-4e92-a905-0a839e747f99): select ratings.movieid,max(ratings.rating) from ratings group by ratings.movieid limit 10
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:ratings.movieid, type:int, comment:null), FieldSchema(name:_c1, type:int, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20200406120303_c0a24f6e-f2d5-4e92-a905-0a839e747f99); Time taken: 0.161 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20200406120303_c0a24f6e-f2d5-4e92-a905-0a839e747f99): select ratings.movieid,max(ratings.rating) from ratings group by ratings.movieid limit 10
INFO  : Query ID = hive_20200406120303_c0a24f6e-f2d5-4e92-a905-0a839e747f99
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks not specified. Estimated from input data size: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : Starting Job = job_1586167015334_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1586167015334_0008/
INFO  : Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1586167015334_0008
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2020-04-06 12:03:26,650 Stage-1 map = 0%,  reduce = 0%
INFO  : 2020-04-06 12:03:51,326 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 13.98 sec
INFO  : 2020-04-06 12:04:52,204 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 13.98 sec
INFO  : 2020-04-06 12:04:54,281 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 16.74 sec
INFO  : MapReduce Total cumulative CPU time: 16 seconds 740 msec
INFO  : Ended Job = job_1586167015334_0008
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 16.74 sec   HDFS Read: 21600885 HDFS Write: 41 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 16 seconds 740 msec
INFO  : Completed executing command(queryId=hive_20200406120303_c0a24f6e-f2d5-4e92-a905-0a839e747f99); Time taken: 97.459 seconds
INFO  : OK
+-----------------+------+--+
| ratings.movieid | _c1 |
+-----------------+------+--+
| 1               | 5   |
| 2               | 5   |
| 3               | 5   |
| 4               | 5   |
| 5               | 5   |
| 6               | 5   |
| 7               | 5   |
| 8               | 5   |
| 9               | 5   |
| 10              | 5   |
+-----------------+------+--+
10 rows selected (97.674 seconds)
```

```
0: jdbc:hive2://quickstart:10000/default> select movies.category, count(movies.category) from movies group by movies.category limit 10;
INFO  : Compiling command(queryId=hive_20200406114747_da588977-86c8-416f-a220-b9e31c867610): select movies.category, count(movies.category) from movies group by movies.category limit 10
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:movies.category, type:string, comment:null), FieldSchema(name:_c1, type:bigint, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20200406114747_da588977-86c8-416f-a220-b9e31c867610); Time taken: 0.283 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20200406114747_da588977-86c8-416f-a220-b9e31c867610): select movies.category, count(movies.category) from movies group by movies.category limit 10
INFO  : Query ID = hive_20200406114747_da588977-86c8-416f-a220-b9e31c867610
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks not specified. Estimated from input data size: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : Starting Job = job_1586167015334_0004, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1586167015334_0004/
INFO  : Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1586167015334_0004
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2020-04-06 11:47:20,210 Stage-1 map = 0%,  reduce = 0%
INFO  : 2020-04-06 11:47:45,935 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 8.55 sec
INFO  : 2020-04-06 11:48:36,405 Stage-1 map = 100%,  reduce = 67%, Cumulative CPU 10.89 sec
INFO  : 2020-04-06 11:48:38,564 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 12.0 sec
INFO  : MapReduce Total cumulative CPU time: 12 seconds 0 msec
INFO  : Ended Job = job_1586167015334_0004
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 12.0 sec   HDFS Read: 170861 HDFS Write: 161 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 12 seconds 0 msec
INFO  : Completed executing command(queryId=hive_20200406114747_da588977-86c8-416f-a220-b9e31c867610); Time taken: 91.583 seconds
INFO  : OK
+----------------------------------------------+------+--+
|                movies.category               | _c1  |
+----------------------------------------------+------+--+
| 20th Century (Zbogum na dvadesetiot vek) (1998) | 1  |
| 302 (1995)                                   | 1    |
| A (1920)                                     | 1    |
| A (1923)                                     | 1    |
| A (1932)                                     | 1    |
| A (1937)                                     | 2    |
| A (1938)                                     | 1    |
| A (1945)                                     | 1    |
| A (1951)                                     | 2    |
| A (1957)                                     | 2    |
+----------------------------------------------+------+--+
10 rows selected (91.93 seconds)
```

```
0: jdbc:hive2://quickstart:10000/default> select count(*) from users where users.age > 50;
INFO  : Compiling command(queryId=hive_20200406115353_07e7439e-fe5b-427c-8918-56661aaf269a): select count(*) from users where users.age > 50
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:_c0, type:bigint, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20200406115353_07e7439e-fe5b-427c-8918-56661aaf269a); Time taken: 0.211 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20200406115353_07e7439e-fe5b-427c-8918-56661aaf269a): select count(*) from users where users.age > 50
INFO  : Query ID = hive_20200406115353_07e7439e-fe5b-427c-8918-56661aaf269a
INFO  : Total jobs = 1
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-1:MAPRED] in serial mode
INFO  : Number of reduce tasks determined at compile time: 1
INFO  : In order to change the average load for a reducer (in bytes):
INFO  :   set hive.exec.reducers.bytes.per.reducer=<number>
INFO  : In order to limit the maximum number of reducers:
INFO  :   set hive.exec.reducers.max=<number>
INFO  : In order to set a constant number of reducers:
INFO  :   set mapreduce.job.reduces=<number>
INFO  : Starting Job = job_1586167015334_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1586167015334_0006/
INFO  : Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1586167015334_0006
INFO  : Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
INFO  : 2020-04-06 11:53:41,311 Stage-1 map = 0%,  reduce = 0%
INFO  : 2020-04-06 11:53:55,603 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.3 sec
INFO  : 2020-04-06 11:54:04,381 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.49 sec
INFO  : MapReduce Total cumulative CPU time: 4 seconds 490 msec
INFO  : Ended Job = job_1586167015334_0006
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.49 sec   HDFS Read: 117617 HDFS Write: 4 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 4 seconds 490 msec
INFO  : Completed executing command(queryId=hive_20200406115353_07e7439e-fe5b-427c-8918-56661aaf269a); Time taken: 32.973 seconds
INFO  : OK
+------+--+
| _c0  |
+------+--+
| 380  |
+------+--+
1 row selected (33.239 seconds)
```

```
0: jdbc:hive2://quickstart:10000/default> select u.*,m.* from users u inner join movies m on u.movieid=m.movieid limit 10 ;
Error: Error while compiling statement: FAILED: SemanticException [Error 10002]: Line 1:53 Invalid column reference 'movieid' (state=42000,code=10002)
0: jdbc:hive2://quickstart:10000/default> select u.*,m.* from users u inner join movies m on u.movieid = m.movieid limit 10 ;
Error: Error while compiling statement: FAILED: SemanticException [Error 10002]: Line 1:53 Invalid column reference 'movieid' (state=42000,code=10002)
0: jdbc:hive2://quickstart:10000/default> select u.*,m.* from users u inner join movies m on u.userid = m.userid limit 10 ;
INFO  : Compiling command(queryId=hive_20200406120000_45fa77e2-1093-4946-b2a3-54ac76e41461): select u.*,m.* from users u inner join movies m on u.userid = m.userid limit 10
INFO  : Semantic Analysis Completed
INFO  : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:u.userid, type:int, comment:null), FieldSchema(name:u.gender, type:string, comment:null), FieldSchema(name:u.age, type:int, comment:nul
l), FieldSchema(name:u.unk1, type:int, comment:null), FieldSchema(name:u.unk2, type:int, comment:null), FieldSchema(name:m.userid, type:int, comment:null), FieldSchema(name:m.movie, type:string, comment:n
ull), FieldSchema(name:m.category, type:string, comment:null)], properties:null)
INFO  : Completed compiling command(queryId=hive_20200406120000_45fa77e2-1093-4946-b2a3-54ac76e41461); Time taken: 0.211 seconds
INFO  : Concurrency mode is disabled, not creating a lock manager
INFO  : Executing command(queryId=hive_20200406120000_45fa77e2-1093-4946-b2a3-54ac76e41461): select u.*,m.* from users u inner join movies m on u.userid = m.userid limit 10
INFO  : Query ID = hive_20200406120000_45fa77e2-1093-4946-b2a3-54ac76e41461
INFO  : Total jobs = 1
INFO  : Starting task [Stage-4:MAPREDLOCAL] in serial mode
INFO  : Execution completed successfully
INFO  : MapredLocal task succeeded
INFO  : Launching Job 1 out of 1
INFO  : Starting task [Stage-3:MAPRED] in serial mode
INFO  : Number of reduce tasks is set to 0 since there's no reduce operator
INFO  : Starting Job = job_1586167015334_0007, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1586167015334_0007/
INFO  : Kill Command = /usr/lib/hadoop/bin/hadoop job  -kill job_1586167015334_0007
INFO  : Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0
INFO  : 2020-04-06 12:00:34,677 Stage-3 map = 0%,  reduce = 0%
INFO  : 2020-04-06 12:00:52,338 Stage-3 map = 100%,  reduce = 0%, Cumulative CPU 14.04 sec
INFO  : MapReduce Total cumulative CPU time: 14 seconds 40 msec
INFO  : Ended Job = job_1586167015334_0007
INFO  : MapReduce Jobs Launched:
INFO  : Stage-Stage-3: Map: 1   Cumulative CPU: 14.04 sec   HDFS Read: 10611 HDFS Write: 558 SUCCESS
INFO  : Total MapReduce CPU Time Spent: 14 seconds 40 msec
INFO  : Completed executing command(queryId=hive_20200406120000_45fa77e2-1093-4946-b2a3-54ac76e41461); Time taken: 37.59 seconds
INFO  : OK
+----------+----------+-------+---------+---------+----------+------------------------------------+----------------------------+--+
| u.userid | u.gender | u.age | u.unk1  | u.unk2  | m.userid |              m.movie               |         m.category         |
+----------+----------+-------+---------+---------+----------+------------------------------------+----------------------------+--+
| 1        | F        | 1     | 10      | 48067   | 1        | Toy Story (1995)                   | Animation|Children's|Comedy |
| 2        | M        | 56    | 16      | 70072   | 2        | Jumanji (1995)                     | Adventure|Children's|Fantasy |
| 3        | M        | 25    | 15      | 55117   | 3        | Grumpier Old Men (1995)            | Comedy|Romance             |
| 4        | M        | 45    | 7       | 2460    | 4        | Waiting to Exhale (1995)           | Comedy|Drama               |
| 5        | M        | 25    | 20      | 55455   | 5        | Father of the Bride Part II (1995) | Comedy                     |
| 6        | F        | 50    | 9       | 55117   | 6        | Heat (1995)                        | Action|Crime|Thriller      |
| 7        | M        | 35    | 1       | 6810    | 7        | Sabrina (1995)                     | Comedy|Romance             |
| 8        | M        | 25    | 12      | 11413   | 8        | Tom and Huck (1995)                | Adventure|Children's       |
| 9        | M        | 25    | 17      | 61614   | 9        | Sudden Death (1995)                | Action                     |
| 10       | F        | 35    | 1       | 95370   | 10       | GoldenEye (1995)                   | Action|Adventure|Thriller  |
+----------+----------+-------+---------+---------+----------+------------------------------------+----------------------------+--+
10 rows selected (38.01 seconds)
```

:)