

Hive code 4:

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System cloudera@quickstart:/tmp/hive_prac
File Edit View Search Terminal Help
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 ~]$ cd /tmp/hive_prac
[cloudera@quickstart ~]$ vim users.csv
[cloudera@quickstart ~]$ vim locations.csv
[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;
OK
Time taken: 0.901 seconds
hive> create table users
> (
>   id int,
>   name string,
>   salary int,
>   unit string
> )
> row format delimited
> fields terminated by ',';
OK
Time taken: 0.111 seconds
hive> load data local inpath
> 'file:///tmp/hive_prac/users.csv' into table users;
Loading data to table hive_1.users
Table hive_1.users stats: [numFiles=1, totalSize=115]
OK
Time taken: 2.634 seconds
hive> select * from users;
OK
1  Vighnesh 100 DNA
2  Raj 200 DNA
3  Sunil 300 DNA
4  Rohit 500 FCS
5  Rajni 100 FCS
6  Chandra 200 FCS
7  Rahul 300 DNA
Time taken: 0.975 seconds, Fetched: 7 row(s)
hive> create table locations
> (
>   id int,
>   location string
> )
> row format delimited
> fields terminated by ',';
OK
```

```
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
rm: remove regular file 'many_column.udf.py'? y
[cloudera@quickstart ~]$ cd /tmp/hive_prac
[cloudera@quickstart ~]$ mv many_column.udf.py /tmp/hive_prac
[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;
OK
Time taken: 0.363 seconds
hive> add file /tmp/hive_prac/many_column.udf.py;
Added resources: /tmp/hive_prac/many_column.udf.py
hive> select transform(country,ordernumber,quantityordered)
> using 'python many_column.udf.py' as (country string, ordernumber int,
multiplied quantity int)
> from sales_order_data orc limit 10;
Query ID = cloudera_20230324050707_2722dcf-344b-456c-abeb-81c19292e0ed
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1679653899842_0010, Tracking URL = http://quickstart.cloud
era:0088/proc/application_1679653899842_0010/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679653899842_0010
Hadoop job information for Stage-1: number of mappers: 1; number of reducers:
0
2023-03-24 05:07:27,731 Stage-1 map = 0%, reduce = 0%
2023-03-24 05:07:34,276 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.37
sec
MapReduce Total cumulative CPU time: 2 seconds 370 msec
Ended Job = job_1679653899842_0010
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 2.37 sec HDFS Read: 27942 HDFS Writ
e: 175 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 370 msec
OK
USA 10107 30000
France 10121 34000
France 10134 41000
USA 10145 45000
USA 10159 49000
USA 10168 36000
France 10180 29000
Norway 10180 48000
USA 10201 22000
France 10211 41000
Time taken: 28.015 seconds, Fetched: 10 row(s)
hive>
```

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System cloudera@quickstart:tmp/hive_prac
File Edit View Search Terminal Help
Time taken: 0.238 seconds
hive> load data inpath 'file:///tmp/hive_prac/locations.csv' into table lacti
ons;
FAILED: SemanticException [Error 10001]: Line 1:66 Table not found 'lactions'
hive> load data inpath 'file:///tmp/hive_prac/locations.csv' into table locat
ions;
Loading data to table hive.1.locations
Table hive.1.locations stats: [numFiles=1, totalSize=55]
OK
Time taken: 1.534 seconds
hive> select * from locations;
OK
1      UP
2      BIHAR
3      MP
4      AP
5      MAHARASHTRA
6      GOA
7      JHARKHAND
Time taken: 0.232 seconds, Fetched: 7 row(s)
hive> create table buck_users
> (
>   id int,
>   name string,
>   salary int,
>   unit string,
> );
Nov 16 10:30:11 2023: at org.apache.hadoop.hive.ql.parse.HiveParser$IdentifiersParser.ident
ifier(HiveParser$IdentifiersParser.java:11566)
at org.apache.hadoop.hive.ql.parse.HiveParser$Identifier(HiveParser.j
ava:44913)
at org.apache.hadoop.hive.ql.parse.HiveParser$columnNameType(HivePars
er.java:38559)
at org.apache.hadoop.hive.ql.parse.HiveParser$columnNameTypeList(Hive
Parser.java:38789)
at org.apache.hadoop.hive.ql.parse.HiveParser$createTableStatement(Hi
veParser.java:5857)
at org.apache.hadoop.hive.ql.parse.HiveParser$ddlStatement(HiveParse
r.java:2557)
at org.apache.hadoop.hive.ql.parse.HiveParser$execStatement(HiveParse
r.java:1589)
at org.apache.hadoop.hive.ql.parse.HiveParser$statement(HiveParser.ja
va: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.compile(Driver.java:522)
Cloudera Live: Welco... cloudera@quickstart:it...
12°C Rain showers
Search
ENG IN 16:49 24-03-2023
```

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System cloudera@quickstart:tmp/hive_prac
File Edit View Search Terminal Help
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:226)
at org.apache.hadoop.hive.cli.CliDriver.processCmd(CliDriver.java:175)
)
at org.apache.hadoop.hive.cli.CliDriver.processLine(CliDriver.java:38)
9)
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:666)
at org.apache.hadoop.util.RunJar.run(RunJar.java:221)
at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
FAILED: ParseException line 7:0 cannot recognize input near '<EOF>' '<EOF
>' in column specification
hive> create table buck_users
> (
>   id int,
>   name string,
>   salary int,
>   unit string
> )
> clustered by (id)
> sorted by (id)
> into 2 buckets;
OK
Time taken: 0.36 seconds
hive> create table buck_locations
> (
>   id int,
>   location string
> )
> clustered by (id)
> sorted by (id)
> into 2 buckets;
OK
Time taken: 0.298 seconds
hive> set hive.enforce.bucketing=true;
hive> insert overwrite table buck_users select * from users;
Query ID = cloudera_28238324894040_0c8138fc-7e34-498d-acd7-d9e8663b8375
Cloudera Live: Welco... cloudera@quickstart:it...
12°C Rain showers
Search
ENG IN 16:50 24-03-2023
```

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:/tmp/hive_prac
File Edit View Search Terminal Help
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 2
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_1679653899842_0011, Tracking URL = http://quickstart.cloud
era:8088/proxy/application_1679653899842_0011/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679653899842_0011
Hadoop job information for Stage-1: number of mappers: 1; number of reducers:
2
2023-03-24 09:40:50,034 Stage-1 map = 0%, reduce = 0%
2023-03-24 09:41:10,254 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 6.27
sec
2023-03-24 09:41:27,947 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 17
.2 sec
MapReduce Total cumulative CPU time: 17 seconds 200 msec
Ended Job = job_1679653899842_0011
Loading data to table hive.1.buck.users
Table hive.1.buck.users stats: [numFiles=2, numRows=7, totalSize=115, rawData
Size=108]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 2 Cumulative CPU: 17.2 sec HDFS Read: 1117
1 HDFS Write: 261 SUCCESS
Total MapReduce CPU Time Spent: 17 seconds 200 msec
OK
Time taken: 70.558 seconds
hive> insert overwrite table buck_locations select * from locations;
Query ID = cloudera.20230324094242_0a5e4c3a-a130-4c08-9909-5241df127bfa
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 2
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_1679653899842_0012, Tracking URL = http://quickstart.cloud
era:8088/proxy/application_1679653899842_0012/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679653899842_0012
Hadoop job information for Stage-1: number of mappers: 1; number of reducers:
2
2023-03-24 09:42:39,800 Stage-1 map = 0%, reduce = 0%
2023-03-24 09:42:54,524 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 5.02
sec
Cloudera Live: Welco... cloudera@quickstart:/t...
```

```
sec
2023-03-24 09:43:15,526 Stage-1 map = 100%, reduce = 50%, Cumulative CPU 11.
35 sec
2023-03-24 09:43:16,831 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 11
.35 sec
MapReduce Total cumulative CPU time: 17 seconds 10 msec
Ended Job = job_1679653899842_0012
Loading data to table hive.1.buck.locations
Table hive.1.buck.locations stats: [numFiles=2, numRows=7, totalSize=55, rawD
ataSize=48]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 2 Cumulative CPU: 17.01 sec HDFS Read: 103
53 HDFS Write: 209 SUCCESS
Total MapReduce CPU Time Spent: 17 seconds 10 msec
OK
Time taken: 61.385 seconds
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 hive_prac]$ hadoop fs -ls /user/hive/warehouse/hive.1.db
/buck/users/
Found 2 items
-rwxrwxrwx 1 cloudera supergroup 48 2023-03-24 09:41 /user/hive/war
ehouse/hive.1.db/buck/users/000000_0
-rwxrwxrwx 1 cloudera supergroup 67 2023-03-24 09:41 /user/hive/war
ehouse/hive.1.db/buck/users/000001_0
[cloudera@quickstart hive_prac]$ hadoop fs -ls /user/hive/warehouse/hive.1.db
/location_users/
ls: '/user/hive/warehouse/hive.1.db/location_users/': No such file or directo
ry
[cloudera@quickstart hive_prac]$ hadoop fs -ls /user/hive/warehouse/hive.1.db
/locations_users/
ls: '/user/hive/warehouse/hive.1.db/locations_users/': No such file or direct
ory
[cloudera@quickstart hive_prac]$ hadoop fs -ls /user/hive/warehouse/hive.1.db
/buck/locations/
Found 2 items
-rwxrwxrwx 1 cloudera supergroup 19 2023-03-24 09:43 /user/hive/war
ehouse/hive.1.db/buck/locations/000000_0
-rwxrwxrwx 1 cloudera supergroup 36 2023-03-24 09:43 /user/hive/war
ehouse/hive.1.db/buck/locations/000001_0
[cloudera@quickstart hive_prac]$
```

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:/tmp/hive_prac
File Edit View Search Terminal Help
hive> select * from buck users u INNER JOIN buck locations l ON u.id = l.id;
Query ID = cloudera_20230324121313_02ede0d-ba30-408c-9088-fa51e5c2f043
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 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_1679653899842_0013, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679653899842_0013/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679653899842_0013
Hadoop job information for Stage-1: number of Mappers: 2; number of reducers: 1
2023-03-24 12:13:11,666 Stage-1 map = 0%, reduce = 0%
2023-03-24 12:13:23,047 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 4.21 sec
2023-03-24 12:13:25,183 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 9.62 sec
2023-03-24 12:13:31,628 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 11.06 sec
MapReduce Total cumulative CPU time: 11 seconds 860 msec
Ended Job = job_1679653899842_0013
MapReduce Jobs Launched:
Stage-Stage-1: Map: 2 Reduce: 1 Cumulative CPU: 11.06 sec HDFS Read: 152
54 HDFS Write: 170 SUCCESS
Total MapReduce CPU Time Spent: 11 seconds 860 msec
OK
1  Vighnesh 100 DNA 1 UP
2  Raj 200 DNA 2 BIHAR
3  Sunil 300 DNA 3 MP
4  Rohit 500 FCS 4 AP
5  Rajni 100 FCS 5 MAHARASHTRA
6  Chandra 200 FCS 6 GOA
7  Rahul 300 DNA 7 JHARKHAND
Time taken: 12.532 seconds, Fetched: 7 row(s)
hive> set hive.auto.convert.join=true;
hive> select * from buck users u INNER JOIN buck locations l ON u.id = l.id;
Query ID = cloudera_20230324121414_8f7bec87-9c63-45f7-b724-e912a7340095
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20230324121414_8f7bec87-9c63-45f7-b724-e912a7340095.log
2023-03-24 12:14:46 Starting to launch local task to process map join; m
aximum memory = 932184064
2023-03-24 12:14:46 Dump the side-table for tag: 1 with group count: 7 in
to file: file:/tmp/cloudera/ef3a5137-df22-4454-ac24-2379fb434231/hive_2023-03-24-12-14-42_118_5851416760145352125-1/-local-10003/HashTable-Stage-3/MapJoin
-24_12-14-42_118_5851416760145352125-1/-local-10003/HashTable-Stage-3/MapJoin
2023-03-24 12:14:46 End of local task; Time Taken: 1.59 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1679653899842_0014, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679653899842_0014/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679653899842_0014
Hadoop job information for Stage-3: number of Mappers: 1; number of reducers: 0
2023-03-24 12:14:56,676 Stage-3 map = 0%, reduce = 0%
2023-03-24 12:15:04,207 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 2.38 sec
MapReduce Total cumulative CPU time: 2 seconds 380 msec
Ended Job = job_1679653899842_0014
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1 Cumulative CPU: 2.38 sec HDFS Read: 7398 HDFS Write
: 170 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 380 msec
OK
6  Chandra 200 FCS 6 GOA
4  Rohit 500 FCS 4 AP
2  Raj 200 DNA 2 BIHAR
7  Rahul 300 DNA 7 JHARKHAND
5  Rajni 100 FCS 5 MAHARASHTRA
3  Sunil 300 DNA 3 MP
1  Vighnesh 100 DNA 1 UP
Time taken: 23.231 seconds, Fetched: 7 row(s)
hive> set hive.optimize.bucketmapjoin=true;
hive> set hive.auto.convert.join=true;
hive> select * from buck users u INNER JOIN buck locations l ON u.id = l.id;
Query ID = cloudera_20230324121616_76383a68-98c-4753-8c15-4863ba749957
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20230324121616_76383a68-98c-4753-8c15-4863ba749957.log
2023-03-24 12:16:46 Starting to launch local task to process map join; m
aximum memory = 932184064
2023-03-24 12:16:47 Dump the side-table for tag: 1 with group count: 7 in
to file: file:/tmp/cloudera/ef3a5137-df22-4454-ac24-2379fb434231/hive_2023-03-24-12-16-42_381_665628371807180592-1/-local-10003/HashTable-Stage-3/MapJoin
-mapfile1--hashtable
2023-03-24 12:16:47 Uploaded 1 file to: file:/tmp/cloudera/ef3a5137-df22-4454-ac24-2379fb434231/hive_2023-03-24-12-16-42_381_665628371807180592-1/-local-10003/HashTable-Stage-3/MapJoin
-mapfile1--hashtable (434 bytes)
2023-03-24 12:16:47 End of local task; Time Taken: 1.453 sec.
```

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:/tmp/hive_prac
File Edit View Search Terminal Help
-mapfile1--hashtable
2023-03-24 12:16:47 Uploaded 1 file to: file:/tmp/cloudera/ef3a5137-df22-4454-ac24-2379fb434231/hive_2023-03-24-12-16-42_381_665628371807180592-1/-local-10003/HashTable-Stage-3/MapJoin
-mapfile1--hashtable (434 bytes)
2023-03-24 12:16:47 End of local task; Time Taken: 1.59 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1679653899842_0014, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679653899842_0014/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679653899842_0014
Hadoop job information for Stage-3: number of Mappers: 1; number of reducers: 0
2023-03-24 12:14:56,676 Stage-3 map = 0%, reduce = 0%
2023-03-24 12:15:04,207 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 2.38 sec
MapReduce Total cumulative CPU time: 2 seconds 380 msec
Ended Job = job_1679653899842_0014
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1 Cumulative CPU: 2.38 sec HDFS Read: 7398 HDFS Write
: 170 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 380 msec
OK
6  Chandra 200 FCS 6 GOA
4  Rohit 500 FCS 4 AP
2  Raj 200 DNA 2 BIHAR
7  Rahul 300 DNA 7 JHARKHAND
5  Rajni 100 FCS 5 MAHARASHTRA
3  Sunil 300 DNA 3 MP
1  Vighnesh 100 DNA 1 UP
Time taken: 23.231 seconds, Fetched: 7 row(s)
hive> set hive.optimize.bucketmapjoin=true;
hive> set hive.auto.convert.join=true;
hive> select * from buck users u INNER JOIN buck locations l ON u.id = l.id;
Query ID = cloudera_20230324121616_76383a68-98c-4753-8c15-4863ba749957
Total jobs = 1
Execution log at: /tmp/cloudera/cloudera_20230324121616_76383a68-98c-4753-8c15-4863ba749957.log
2023-03-24 12:16:46 Starting to launch local task to process map join; m
aximum memory = 932184064
2023-03-24 12:16:47 Dump the side-table for tag: 1 with group count: 7 in
to file: file:/tmp/cloudera/ef3a5137-df22-4454-ac24-2379fb434231/hive_2023-03-24-12-16-42_381_665628371807180592-1/-local-10003/HashTable-Stage-3/MapJoin
-mapfile1--hashtable
2023-03-24 12:16:47 Uploaded 1 file to: file:/tmp/cloudera/ef3a5137-df22-4454-ac24-2379fb434231/hive_2023-03-24-12-16-42_381_665628371807180592-1/-local-10003/HashTable-Stage-3/MapJoin
-mapfile1--hashtable (434 bytes)
2023-03-24 12:16:47 End of local task; Time Taken: 1.453 sec.
```

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System cloudera@quickstart:tmp/hive_prac
File Edit View Search Terminal Help
Execution completed successfully
MapreduceLocal task succeeded
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1679653899842_0015, Tracking URL = http://quickstart.cloud
era:8088/proxy/application_1679653899842_0015/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679653899842_0015
Hadoop job information for Stage-3: number of mappers: 1; number of reducers:
0
2023-03-24 12:16:57,509 Stage-3 map = 0%, reduce = 0%
2023-03-24 12:17:03,959 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 2.16
sec
MapReduce Total cumulative CPU time: 2 seconds 160 msec
Ended Job = job_1679653899842_0015
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1 Cumulative CPU: 2.16 sec HDFS Read: 7389 HDFS Write
: 170 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 160 msec
OK
6 Chandra 200 FCS 6 GOA
4 Rohit 500 FCS 4 AP
2 Raj 200 DNA 2 BIHAR
7 Rahul 300 DNA 7 JHARKHAND
5 Rajni 100 FCS 5 MAHARASHTRA
3 Sunil 300 DNA 3 MP
1 Vighnesh 100 DNA 1 UP
Time taken: 22.735 seconds, Fetched: 7 row(s)
hive> set hive.enforce.sortmergebucketmapjoin=false;
hive> set hive.auto.convert.sortmerge.join=true;
hive> set hive.optimize.bucketmapjoin = true;
hive> set hive.optimize.bucketmapjoin.sortedmerge = true;
hive> set hive.auto.convert.join = false;
hive> select * from buck users u INNER JOIN buck locations l ON u.id = l.id;
Query ID = cloudera_20230324121919_90662349-908c-4839-ae78-9252bdf771c47
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1679653899842_0016, Tracking URL = http://quickstart.cloud
era:8088/proxy/application_1679653899842_0016/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679653899842_0016
Hadoop job information for Stage-1: number of mappers: 2; number of reducers:
0
2023-03-24 12:19:32,155 Stage-1 map = 0%, reduce = 0%
2023-03-24 12:20:00,733 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 3.09
sec
2023-03-24 12:20:01,817 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 6.07
sec
MapReduce Total cumulative CPU time: 6 seconds 70 msec
Ended Job = job_1679653899842_0016
Cloudera Live: Welco... cloudera@quickstart:lt...
```

```
MapReduce Jobs Launched:
Stage-Stage-1: Map: 2 Cumulative CPU: 6.07 sec HDFS Read: 14352 HDFS Writ
e: 170 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 70 msec
OK
6 Chandra 200 FCS 6 GOA
4 Rohit 500 FCS 4 AP
2 Raj 200 DNA 2 BIHAR
7 Rahul 300 DNA 7 JHARKHAND
5 Rajni 100 FCS 5 MAHARASHTRA
3 Sunil 300 DNA 3 MP
1 Vighnesh 100 DNA 1 UP
Time taken: 20.119 seconds, Fetched: 7 row(s)
hive>
```