HIVE CASE STUDY

Data Transfer

`hadoop fs -ls /` Command that lists the contents in the Hadoop Distributed File System

`hadoop fs -mkdir /user/hive/demo` command is used to make a new directory in the present working directory

```
[hadoop@ip-172-31-46-21 ~]$ hadoop fs -ls /user/hive
Found 2 items
drwxr-xr-x   - hive hadoop          0 2021-05-30 13:57 /user/hive/.hiveJars
drwxrwxrwt   - hdfs hadoop          0 2021-05-30 13:55 /user/hive/warehouse
[hadoop@ip-172-31-46-21 ~]$ hadoop fs -mkdir /user/hive/demo
[hadoop@ip-172-31-46-21 ~]$ hadoop fs -mkdir /user/hive/
mkdir: `/user/hive': File exists
[hadoop@ip-172-31-46-21 ~]$ hadoop fs -ls /user/hive/
Found 3 items
drwxr-xr-x   - hive   hadoop          0 2021-05-30 13:57 /user/hive/.hiveJars
drwxr-xr-x   - hadoop hadoop          0 2021-05-30 15:16 /user/hive/demo
drwxrwxrwt   - hdfs   hadoop          0 2021-05-30 13:55 /user/hive/warehouse
[hadoop@ip-172-31-46-21 ~]$
```

`hadoop distcp 's3://e-commerce-events-ml/*' '/user/hive/demo/'` command is used to copy files in the S3 bucket to the Hadoop file system.

```
[hadoop@ip-172-31-46-21 ~]$ hadoop distcp 's3://e-commerce-events-ml/*' '/user/hive/demo/'
ERROR: Tools helper ///usr/lib/hadoop/libexec/tools/hadoop-distcp.sh was not found.
2021-05-30 15:19:14,550 INFO tools.DistCp: Input Options: DistCpOptions{atomicCommit=false, syncFolder=false, deleteMissing=false, ignoreFailures=false, overwrite=false, append=false, useDiff=false, useRdiff=false, fromSnapshot=null, toS
napshot=null, skipCRC=false, blocking=true, numListstatusThreads=0, maxMaps=20, mapBandwidth=0.0, copyStrategy='uniformsize', preserveStatus=[BLOCKSIZE], atomicWorkPath=null, logPath=null, sourceFileListing=null, sourcePaths=[s3://e-comm
erce-events-ml/*], targetPath=/user/hive/demo, filtersFile='null', blocksPerChunk=0, copyBufferSize=8192, verboseLog=false, directWrite=false}, sourcePaths=[s3://e-commerce-events-ml/*], targetPathExists=true, preserveRawXattrsfalse
2021-05-30 15:19:14,889 INFO client.RMProxy: Connecting to ResourceManager at ip-172-31-46-21.ec2.internal/172.31.46.21:8032
```

Checking the Dataset

hadoop fs -ls /user/hive/ command to check the files in the directory

```
[hadoop@ip-172-31-46-21 ~]$ hadoop fs -ls /user/hive/demo/
Found 2 items
-rw-r--r--   1 hadoop hadoop  545839412 2021-05-30 15:19 /user/hive/demo/2019-Nov.csv
-rw-r--r--   1 hadoop hadoop  482542278 2021-05-30 15:19 /user/hive/demo/2019-Oct.csv
[hadoop@ip-172-31-46-21 ~]$
```

hadoop fs -cat /user/hive/demo/2019-Nov.csv | head command to check the first 10 entries in Nov dataset

```
[hadoop@ip-172-31-46-21 ~]$ hadoop fs -cat /user/hive/demo/2019-Nov.csv | head
event_time,event_type,product_id,category_id,category_code,brand,price,user_id,user_session
2019-11-01 00:00:02 UTC,view,5802432,1487580009286598681,,,0.32,562076640,09fafd6c-6c99-46b1-834f-33527f4de241
2019-11-01 00:00:09 UTC,cart,5844397,1487580006317032337,,,2.38,553329724,2067216c-31b5-455d-a1cc-af0575a34ffb
2019-11-01 00:00:10 UTC,view,5837166,1783999064103190764,,pnb,22.22,556138645,57ed222e-a54a-4907-9944-5a875c2d7f4f
2019-11-01 00:00:11 UTC,cart,5876812,1487580010100293687,,jessnail,3.16,564506666,186c1951-8052-4b37-adce-dd9644b1d5f7
2019-11-01 00:00:24 UTC,remove_from_cart,5826182,1487580007483048900,,,3.33,553329724,2067216c-31b5-455d-a1cc-af0575a34ffb
2019-11-01 00:00:24 UTC,remove_from_cart,5826182,1487580007483048900,,,3.33,553329724,2067216c-31b5-455d-a1cc-af0575a34ffb
2019-11-01 00:00:25 UTC,view,5856189,1487580009026551821,,runail,15.71,562076640,09fafd6c-6c99-46b1-834f-33527f4de241
2019-11-01 00:00:32 UTC,view,5837835,1933472286753424063,,,3.49,514649199,432a4e95-375c-4b40-bd36-0fc039e77580
2019-11-01 00:00:34 UTC,remove_from_cart,5870838,1487580007675986893,,milv,0.79,429913900,2f0bff3c-252f-4fe6-afcd-5d8a6a92839a
cat: Unable to write to output stream.
[hadoop@ip-172-31-46-21 ~]$
```

Hive Database Creation

Jump to hive coding console

```
[hadoop@ip-172-31-46-21 ~]$ hive
SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".
SLF4J: Defaulting to no-operation (NOP) logger implementation
SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further details.
Hive Session ID = 56742b48-750b-4e9b-9c7e-d042a5c9e239

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j2.properties Async: false
Hive Session ID = 63c0525f-92e5-40d4-974c-56554cc22388
2021-05-30 15:27:23,208 INFO  [Tez session start thread] client.RMProxy: Connecting to ResourceManager at ip-172-31-46-21.ec2.internal/172.31.46.21:8032
hive> 2021-05-30 15:27:24,137 INFO  [pool-6-thread-1] client.RMProxy: Connecting to ResourceManager at ip-172-31-46-21.ec2.internal/172.31.46.21:8032
2021-05-30 15:27:24,143 INFO  [pool-6-thread-1] client.AHSProxy: Connecting to Application History server at ip-172-31-46-21.ec2.internal/172.31.46.21:10200
2021-05-30 15:27:24,143 INFO  [Tez session start thread] client.AHSProxy: Connecting to Application History server at ip-172-31-46-21.ec2.internal/172.31.46.21:10200
```

create database if not exists demo ; query to create a database demo
show demo ;
use demo ; query to use demo database

```
hive> create database if not exists demo ;
OK
Time taken: 0.225 seconds
```

```
hive> show databases ;
OK
default
demo
Time taken: 0.027 seconds, Fetched: 2 row(s)
```

```
hive> use demo ;
OK
Time taken: 0.057 seconds
```

Creating Base Table

create external table if not exists case_study (event_time timestamp, event_type string, product_id string, category_id string, category_code string, brand string, price float, user_id bigint, user_session string) row format serde 'org.apache.hadoop.hive.serde2.OpenCSVSerde' stored as textfile location "/user/hive/demo/" tblproperties ("skip.header.line.count" = "1") ; query to create base table. For memory optimisation we are merging the October and November datasets using the location of the directory.

```
hive> create external table if not exists case_study ( event_time timestamp, event_type string, product_id string, category_id string, category_code string, brand string, price float, user_id bigint, user_session string ) row format serd
e 'org.apache.hadoop.hive.serde2.OpenCSVSerde' stored as textfile location "/user/hive/demo/"  tblproperties ( "skip.header.line.count" = "1" ) ;
OK
Time taken: 0.323 seconds
hive>
```

select * from case_study limit 5 ; query to view the dataset

```
hive> set hive.cli.print.header = true ;
hive> select * from case_study limit 5 ;
OK
case_study.event_time   case_study.event_type   case_study.product_id   case_study.category_id  case_study.category_code        case_study.brand        case_study.price        case_study.user_id      case_study.user_session
2019-11-01 00:00:02 UTC view    5802432 1487580009286598681                     0.32    562076640       09fafd6c-6c99-46b1-834f-33527f4de241
2019-11-01 00:00:09 UTC cart    5844397 1487580006317032337                     2.38    553329724       2067216c-31b5-455d-a1cc-af0575a34ffb
2019-11-01 00:00:10 UTC view    5837166 1783999064103190764             pnb     22.22   556138645       57ed222e-a54a-4907-9944-5a875c2d7f4f
2019-11-01 00:00:11 UTC cart    5876812 1487580010100293687             jessnail        3.16    564506666       186c1951-8052-4b37-adce-dd9644b1d5f7
2019-11-01 00:00:24 UTC remove_from_cart        5826182 1487580007483048900                     3.33    553329724       2067216c-31b5-455d-a1cc-af0575a34ffb
Time taken: 0.34 seconds, Fetched: 5 row(s)
hive>
```

1. Find the total revenue generated due to purchases made in October.

```
select month(event_time) as month, sum(price) as toatl_price from case_study where event_type = 'purchase' and month(event_time) = 10 group by month(event_time) ;
```

```
hive> select month(event_time) as month, sum(price) as toatl_price from case_study where event_type = 'purchase' and month(event_time) = 10 group by month(event_time) ;
Query ID = hadoop_20210530153621_55812b0d-76a5-4573-afba-e196b52463a3
Total jobs = 1
Launching Job 1 out of 1
Tez session was closed. Reopening...
2021-05-30 15:36:25,313 INFO  [56742b48-750b-4e9b-9c7e-d042a5c9e239 main] client.RMProxy: Connecting to ResourceManager at ip-172-31-46-21.ec2.internal/172.31.46.21:8032
2021-05-30 15:36:25,314 INFO  [56742b48-750b-4e9b-9c7e-d042a5c9e239 main] client.AHSProxy: Connecting to Application History server at ip-172-31-46-21.ec2.internal/172.31.46.21:10200
Session re-established.
Session re-established.
Status: Running (Executing on YARN cluster with App id application_1622383007600_0005)

--------------------------------------------------------------------------------
        VERTICES      MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
--------------------------------------------------------------------------------
Map 1 .......... container     SUCCEEDED      2          2        0        0       0       0
Reducer 2 ...... container     SUCCEEDED      2          2        0        0       0       0
--------------------------------------------------------------------------------
VERTICES: 02/02  [==========================>>] 100%  ELAPSED TIME: 53.60 s
--------------------------------------------------------------------------------
OK
10      1211538.4299997438
Time taken: 155.396 seconds, Fetched: 1 row(s)
```

Total revenue generated due to purchases made in October is 1211538.429

Time taken to execute query from base table is 155.396 sec

Creating Partitioned Table

For query optimisation we do partitioning

```
hive> set hive.vectorized.execution.enabled = true ;
hive> set hive.vectorized.execution.enabled ;
hive.vectorized.execution.enabled=true
```

To set dynamic partitioning to true

```
hive> set hive.exec.dynamic.partition.mode = nonstrict ;
hive> set hive.exec.dynamic.partition = true ;
hive>
```

create table if not exists part_case_study (event_time timestamp, event_type string, product_id string, category_id string, category_code string, brand string, price float, user_id bigint, user_session string) partitioned by (month int) clustered by (event_type) into 4 buckets row format serde 'org.apache.hadoop.hive.serde2.OpenCSVSerde' stored as textfile ; query to create partitioned table

```
hive> create table if not exists part_case_study ( event_time timestamp, event_type string, product_id string, category_id string, category_code string, brand string, price float,
user_id bigint, user_session string ) partitioned by (month int) clustered by (event_type) into 4 buckets row format serde 'org.apache.hadoop.hive.serde2.OpenCSVSerde' stored as te
xtfile ;
OK
Time taken: 0.174 seconds
```

insert into table part_case_study partition (month) select cast(replace(event_time, 'UTC', '') as timestamp), event_type, product_id, category_id, category_code, brand, price, user_id, user_session, month(cast(replace(event_time, 'UTC', '') as timestamp)) from case_study ; query to insert the Base Table data to the Partitioned Table.

```
hive> insert into table part_case_study partition (month) select cast(replace(event_time, 'UTC', '') as timestamp), event_type, product_id, category_id, category_code, brand, price
, user_id, user_session, month(cast(replace(event_time, 'UTC', '') as timestamp)) from case_study ;
2021-05-30 15:46:23,913 INFO  [56742b48-750b-4e9b-9c7e-d042a5c9e239 main] reducesink.VectorReduceSinkObjectHashOperator: VectorReduceSinkObjectHashOperator constructor vectorReduce
SinkInfo org.apache.hadoop.hive.ql.plan.VectorReduceSinkInfo@7282f7b1
Query ID = hadoop_20210530154623_b1c173a1-9410-4e9f-9d07-0c788d98f1f6
Total jobs = 1
Launching Job 1 out of 1
Tez session was closed. Reopening...
2021-05-30 15:46:24,117 INFO  [56742b48-750b-4e9b-9c7e-d042a5c9e239 main] client.RMProxy: Connecting to ResourceManager at ip-172-31-46-21.ec2.internal/172.31.46.21:8032
2021-05-30 15:46:24,117 INFO  [56742b48-750b-4e9b-9c7e-d042a5c9e239 main] client.AHSProxy: Connecting to Application History server at ip-172-31-46-21.ec2.internal/172.31.46.21:102
00
Session re-established.
Session re-established.
Status: Running (Executing on YARN cluster with App id application_1622383007600_0006)

--------------------------------------------------------------------------------
        VERTICES      MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
--------------------------------------------------------------------------------
Map 1 ......... container     SUCCEEDED      2          2        0        0       0       0
Reducer 2 ...... container     SUCCEEDED     11         11        0        0       0       0
--------------------------------------------------------------------------------
VERTICES: 02/02  [==========================>>] 100%  ELAPSED TIME: 204.21 s
--------------------------------------------------------------------------------
Loading data to table demo.part_case_study partition (month=null)

Loaded : 2/2 partitions.
         Time taken to load dynamic partitions: 0.411 seconds
         Time taken for adding to write entity : 0.003 seconds
OK
_col0   _col1   _col2   _col3   _col4   _col5   _col6   _col7   _col8   _col9
Time taken: 215.182 seconds
```

select * from part_case_study limit 5 ; query to view the dataset

```
hive> select * from part_case_study limit 5 ;
OK
part_case_study.event_time      part_case_study.event_type      part_case_study.product_id      part_case_study.category_id     part_case_study.category_code   part_case_study.brand   part_case_study.price   part_case_study.user_id part_
case_study.user_session part_case_study.month
2019-10-08 10:40:20     cart    5700070 1487580009362096156             runail  0.95    492960693       2f9e5e6e-b156-4a67-8726-3ae4c2f3f5a3    10
2019-10-07 20:53:08     cart    5635080 1487580005754995573                     4.44    527827629       b5f0f964-9457-4dfd-bade-239a9cde9c5d    10
2019-10-08 10:40:20     cart    5751423 1511892746070131099             uno     9.37    482861168       2557df11-378c-4c54-b596-50b6fd33cd04    10
2019-10-01 00:00:03     cart    5773353 1487580005134238553             runail  2.62    463240011       26dd6e6e-4dac-4778-8d2c-92e149dab885    10
2019-10-10 11:52:49     cart    5867044 1783999064136745198             de.lux  2.86    558783199       8e9004fe-a437-4bda-aled-a93cbd45d01b    10
Time taken: 0.246 seconds, Fetched: 5 row(s)
hive>
```

1. Find the total revenue generated due to purchases made in October.

```
select month, sum(price) as total_price from part_case_study where event_type = 'purchase' and month = 10 group by month ;
```

```
hive> select month, sum(price) as total_price from part_case_study where event_type = 'purchase' and month = 10 group by month ;
Query ID = hadoop_20210530155304_dd1be198-a20e-47d8-b426-7df78e64de99
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1622383007600_0006)

--------------------------------------------------------------------------------
        VERTICES      MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
--------------------------------------------------------------------------------
Map 1 .......... container     SUCCEEDED      4          4        0        0       0       0
Reducer 2 ...... container     SUCCEEDED      2          2        0        0       0       0
--------------------------------------------------------------------------------
VERTICES: 02/02  [==========================>>] 100%  ELAPSED TIME: 33.92 s
--------------------------------------------------------------------------------
OK
month   total_price
10      1211538.4299997778
Time taken: 35.294 seconds, Fetched: 1 row(s)
hive>
```

Total revenue generated due to purchases made in October is 1211538.429

Time taken to execute query from base table is 35.396 sec

2. Write a query to yield the total sum of purchases per month in a single output.

```
select month, sum(price) as total_price from part_case_study where event_type = 'purchase' group by month ;
```

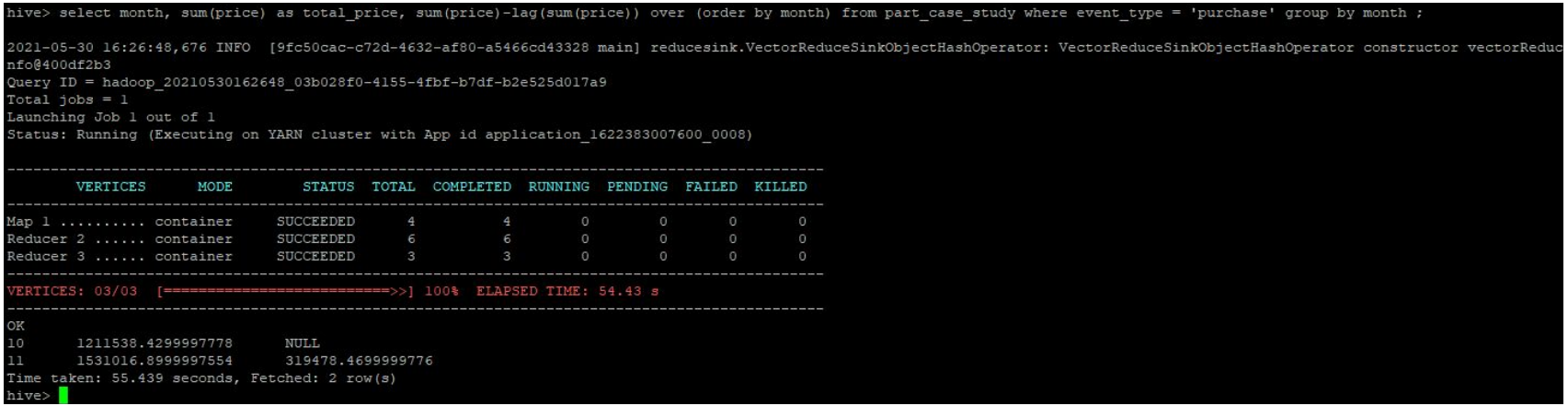
```
hive> select month, sum(price) as total_price from part_case_study where event_type = 'purchase' group by month ;
Query ID = hadoop_20210530162034_411d3efe-a983-49c4-a737-50024lee9c61
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1622383007600_0008)

--------------------------------------------------------------------------------
        VERTICES      MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
--------------------------------------------------------------------------------
Map 1 .......... container     SUCCEEDED      4          4        0        0       0       0
Reducer 2 ...... container     SUCCEEDED      6          6        0        0       0       0
--------------------------------------------------------------------------------
VERTICES: 02/02  [==========================>>] 100%  ELAPSED TIME: 60.74 s
--------------------------------------------------------------------------------
OK
11      1531016.8999997554
10      1211538.4299997778
Time taken: 61.71 seconds, Fetched: 2 row(s)
hive>
```

Total revenue generated due to purchases made in October is 1211538.429

Total revenue generated due to purchases made in November is 1531016.899

3. Write a query to find the change in revenue generated due to purchases from October to November

```
select month, sum(price) as total_price, sum(price)-lag(sum(price)) over (order by month) from part_case_study where event_type = 'purchase' group by month ;
```



Total revenue generated to purchases in November is 319470.769 more than October

4. Find distinct categories of products. Categories with null category code can be ignored

```
select distinct(category_code) from part_case_study where category_code is not null ;
```

```
hive> select distinct(category_code) from part_case_study where category_code is not null ;
Query ID = hadoop_20210530164343_9a2f8682-3f5d-4f3b-9ff0-d2b3347f3af0
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1622383007600_0008)

----------------------------------------------------------------------------------------------
        VERTICES      MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
----------------------------------------------------------------------------------------------
Map 1 .......... container     SUCCEEDED      4          4        0        0       0       0
Reducer 2 ...... container     SUCCEEDED     12         12        0        0       0       0
----------------------------------------------------------------------------------------------
VERTICES: 02/02  [==========================>>] 100%  ELAPSED TIME: 61.03 s
----------------------------------------------------------------------------------------------
OK
accessories.cosmetic_bag
furniture.living_room.cabinet
furniture.living_room.chair
appliances.personal.hair_cutter
apparel.glove
appliances.environment.air_conditioner
sport.diving

furniture.bathroom.bath
accessories.bag
appliances.environment.vacuum
stationery.cartrige
Time taken: 61.878 seconds, Fetched: 12 row(s)
```

There are 11 distinct categories from both October and November dataset

5. Find the total number of products available under each category

```
select category_code, count(product_id) as Total_Product_Count from part_case_study where category_code is not null group by category_code ;
```

```
hive> select category_code, count(product_id) as Total_Product_Count from part_case_study group by category_code ;
Query ID = hadoop_20210530163513_dc26ee8d-cf41-4e1f-b455-d47b27b5ae10
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1622383007600_0008)

--------------------------------------------------------------------------------
        VERTICES      MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
--------------------------------------------------------------------------------
Map 1 .......... container     SUCCEEDED      4          4        0        0       0       0
Reducer 2 ...... container     SUCCEEDED     12         12        0        0       0       0
--------------------------------------------------------------------------------
VERTICES: 02/02  [==========================>>] 100%  ELAPSED TIME: 61.19 s
--------------------------------------------------------------------------------
OK
accessories.cosmetic_bag        1248
furniture.living_room.cabinet   13439
furniture.living_room.chair     308
appliances.personal.hair_cutter 1643
apparel.glove   18232
appliances.environment.air_conditioner  332
sport.diving    2
        8594895
furniture.bathroom.bath 9857
accessories.bag 11681
appliances.environment.vacuum   59761
stationery.cartrige     26722
Time taken: 62.249 seconds, Fetched: 12 row(s)
```

6. Which brand had the maximum sales in October and November combined

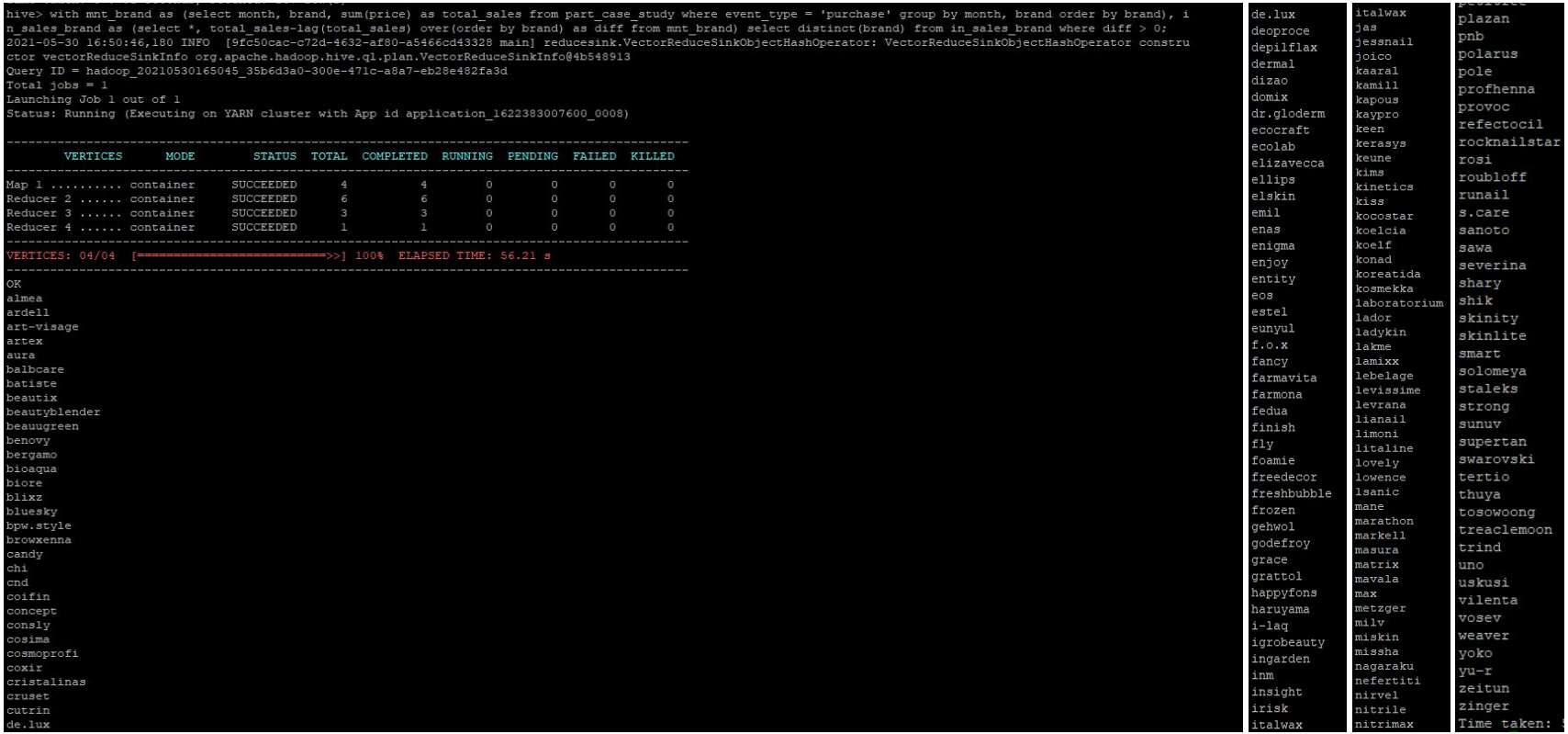
```
select brand, sum(price) as Total_Sales from part_case_study where event_type = 'purchase' group by brand order by sum(price) desc ;
```

```
hive> select brand, sum(price) as Total_Sales from part_case_study where event_type = 'purchase' group by brand order by sum(price) desc ;
2021-05-30 16:45:58,855 INFO  [9fc50cac-c72d-4632-af80-a5466cd43328 main] reducesink.VectorReduceSinkObjectHashOperator: VectorReduceSinkObjectHashOperator constructor vectorReduceSinkInfo org.apache.hadoop.hive.ql.plan.VectorR
nfo@4690a37e
Query ID = hadoop_20210530164558_0bce1ca4-dc40-4dd8-98c4-ef812922b8e1
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1622383007600_0008)

----------------------------------------------------------------------------------------------
        VERTICES      MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
----------------------------------------------------------------------------------------------
Map 1 .......... container     SUCCEEDED      4          4        0        0       0       0
Reducer 2 ...... container     SUCCEEDED      6          6        0        0       0       0
Reducer 3 ...... container     SUCCEEDED      1          1        0        0       0       0
----------------------------------------------------------------------------------------------
VERTICES: 03/03  [==========================>>] 100%  ELAPSED TIME: 55.15 s
----------------------------------------------------------------------------------------------
OK
        1094188.3000000042
runail  148297.93999999674
grattol 106918.25000000343
```

Runail brand have the maximum sales i.e. 148297.939 in October and November combined

7. Which brands increased their sales from October to November?

```
with mnt_brand as (select month, brand, sum(price) as total_sales from part_case_study where event_type = 'purchase' group by month, brand order by brand), in_sales_brand as (select *, total_sales-lag(total_sales) over(order by brand) as diff from mnt_brand) select distinct(brand) from in_sales_brand where diff > 0 ;
```



8. Your company wants to reward the top 10 users of its website with a Golden Customer plan. Write a query to generate a list of top 10 users who spend the most.

```
select user_id, sum(price) as Total_Spent from part_case_study group by user_id order by sum(price) desc limit 10;
```

```
hive> select user_id, sum(price) as Total_Spent from part_case_study group by user_id order by sum(price) desc limit 10;
2021-05-30 17:00:47,789 INFO  [9fc50cac-c72d-4632-af80-a5466cd43328 main] reducesink.VectorReduceSinkObjectHashOperator: VectorReduceSinkObjectHashOperator constructor vectorReduceSinkInfo org.apache.hadoop.hive.ql.plan.V
nfo@2ddb260e
Query ID = hadoop_20210530170047_dac686a7-e571-414e-a6a2-678ce50adfe9
Total jobs = 1
Launching Job 1 out of 1
Tez session was closed. Reopening...
2021-05-30 17:00:47,987 INFO  [9fc50cac-c72d-4632-af80-a5466cd43328 main] client.RMProxy: Connecting to ResourceManager at ip-172-31-46-21.ec2.internal/172.31.46.21:8032
2021-05-30 17:00:47,988 INFO  [9fc50cac-c72d-4632-af80-a5466cd43328 main] client.AHSProxy: Connecting to Application History server at ip-172-31-46-21.ec2.internal/172.31.46.21:10200
Session re-established.
Session re-established.
Status: Running (Executing on YARN cluster with App id application_1622383007600_0009)

--------------------------------------------------------------------------------
        VERTICES      MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
--------------------------------------------------------------------------------
Map 1 ..........  container     SUCCEEDED      4          4        0        0       0       0
Reducer 2 ......  container     SUCCEEDED     12         12        0        0       0       0
Reducer 3 ......  container     SUCCEEDED      1          1        0        0       0       0
--------------------------------------------------------------------------------
VERTICES: 03/03  [==========================>>] 100%  ELAPSED TIME: 74.93 s
--------------------------------------------------------------------------------
OK
557616099       63266.970000000016
557956487       52370.219999999994
550388516       46264.280000000326
531900924       43504.70999999998
352394658       28205.910000000094
550353491       25317.26
443045778       23742.679999999964
479928991       23540.600000000006
554848397       23359.430000000033
526213023       22983.28
Time taken: 161.544 seconds, Fetched: 10 row(s)
```