```
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ cd $HIVE_HOME
hdoop@admin1-HP-280-G4-MT-Business-PC:~/apache-hive-3.1.2-bin$ cd lib/
hdoop@admin1-HP-280-G4-MT-Business-PC:~/apache-hive-3.1.2-bin/lib$ cd ..
hdoop@admin1-HP-280-G4-MT-Business-PC:~/apache-hive-3.1.2-bin$ ls
bin  binary-package-licenses  conf  examples  hcatalog  jdbc  lib  LICENSE  NOTICE  RELEASE_NOTES.txt  scripts
hdoop@admin1-HP-280-G4-MT-Business-PC:~/apache-hive-3.1.2-bin$ cd conf/
```

S

```
hdoop@admin1-HP-280-G4-MT-Business-PC:~/apache-hive-3.1.2-bin/conf$ rm -rf metastore_db/
hdoop@admin1-HP-280-G4-MT-Business-PC:~/apache-hive-3.1.2-bin/conf$ ls
beeline-log4j2.properties.template  hive-env.sh.template                ivysettings.xml                      parquet-logging.properties
derby.log                           hive-exec-log4j2.properties.template  llap-cli-log4j2.properties.template
hive-default.xml.template           hive-log4j2.properties.template       llap-daemon-log4j2.properties.template
hdoop@admin1-HP-280-G4-MT-Business-PC:~/apache-hive-3.1.2-bin/conf$ schematool -dbType derby -initSchema
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/hdoop/hadoop-3.2.1/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Metastore connection URL:        jdbc:derby:;databaseName=metastore_db;create=true
Metastore Connection Driver :    org.apache.derby.jdbc.EmbeddedDriver
Metastore connection User:       APP
Starting metastore schema initialization to 3.1.0
Initialization script hive-schema-3.1.0.derby.sql
```

```
Initialization script completed
schemaTool completed
hdoop@admin1-HP-280-G4-MT-Business-PC:~/apache-hive-3.1.2-bin/conf$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/hdoop/hadoop-3.2.1/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Hive Session ID = 244a7fe9-1535-467c-be3f-f14bdba331d4

Logging initialized using configuration in jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/hive-common-3.1.2.jar!/hive-log4j2.properties Async: true
Hive Session ID = 749f9042-d7fa-4d32-b517-9cecc82c1be5
Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X
releases.
hive> show databases;
OK
default
Time taken: 0.444 seconds, Fetched: 1 row(s)
```

1. Insert 5 records using Insert command in HiveQL

create table sales(product string,price int,payment_type string,name string,city string,state string,country string);

```
hive> create database if not exists salesdb;
OK
Time taken: 0.076 seconds
hive> use salesdb;
OK
Time taken: 0.033 seconds
hive> create table sales(product string,price int,payment_type string,name string,city string,state string,country string);
OK
Time taken: 0.415 seconds
hive> desc sales;
OK
product                 string
price                   int
payment_type            string
name                    string
city                    string
state                   string
country                 string
```

hive> insert into sales values('product201',1000,'Visa','madhuri','Bangalore','Karnataka','India');

```
hive>  insert into sales values('product201',1000,'Visa','madhuri','Bangalore','Karnataka','India');
Query ID = hdoop_20220628104645_abe77fee-7b06-4686-87f0-6362201363ae
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 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>
```

hive> insert into sales values('product202',3000,'Diners','shwetha','seattle','washington','usa');

```
hive>  insert into sales values('product202',3000,'Diners','shwetha','seattle','washington','usa');
Query ID = hdoop_20220628104904_de7408c8-1762-42d6-8e63-86e73d2160e3
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 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_1656388870091_0009, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656388870091_0009/
```

hive> insert into sales values('product203',3000,'Mastercard','bob','ontario','NJ','usa');

```
Time taken: 16.159 seconds
hive> insert into sales values('product203',3000,'Mastercard','bob','ontario','NJ','usa');
Query ID = hdoop_20220628105019_41e8f595-2b54-49a1-8e4e-7a8814f40d4f
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 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>
```

hive> insert into sales values('product204',3540,'Mastercard','bob','Mickleton','kk','usa');

```
Time taken: 15.234 seconds
hive> insert into sales values('product204',3540,'Mastercard','bob','Mickleton','kk','usa');
Query ID = hdoop_20220628105112_4084953d-b1d0-446e-a356-d1d30dbdfed3
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 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_1656388870091_0011, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656388870091_0011/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1656388870091_0011
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 10:51:17,673 Stage-1 map = 0%,  reduce = 0%
2022-06-28 10:51:22,768 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.21 sec
```

hive> insert into sales values('product205',3540,'Mastercard','charlie','hampton','NJ','usa');

```
Time taken: 15.35 seconds
hive>  insert into sales values('product205',3540,'Mastercard','charlie','hampton','NJ','usa');
Query ID = hdoop_20220628105156_62169929-15ad-4ad1-b4b5-b1c3b4353a1e
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 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_1656388870091_0012, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656388870091_0012/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1656388870091_0012
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 10:52:02,835 Stage-1 map = 0%,  reduce = 0%
2022-06-28 10:52:06,964 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.31 sec
2022-06-28 10:52:11,071 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.58 sec
MapReduce Total cumulative CPU time: 3 seconds 580 msec
```

2. Import the dataset from the sales.csv

hive> load data local inpath '/home/hdoop/Desktop/sales.csv' into table sales;
Loading data to table salesdb.sales
OK
Time taken: 0.16 seconds
hive> select *from sales;

```
hive>  load data local inpath '/home/hdoop/Desktop/sales.csv' into table sales;
Loading data to table salesdb.sales
OK
Time taken: 0.16 seconds
hive> select *from sales;
OK
product200      5000    Mastercard      Kumar   Astoria California      USA
product201      1000    Visa    madhuri Bangalore       Karnataka       India
product202      3000    Diners  shwetha seattle Karnataka       India
product203      3000    Mastercard      Krupananda      Mickleton       NJ      US
product204      3540    Mastercard      bob     Mickleton       NJ      US
product205      3540    Mastercard      charlie Mickleton       NJ      US
```

3. Count the number of sales done by each country

hive> select count(*) as count,Country from sales group by Country;

```
hive> select count(*) as count,Country from sales group by Country;
Query ID = hdoop_20220628105224_dc16c886-9a87-49c9-8990-55c159b609c4
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_1656388870091_0013, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656388870091_0013/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1656388870091_0013
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 10:52:31,035 Stage-1 map = 0%,  reduce = 0%
2022-06-28 10:52:34,165 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.31 sec
2022-06-28 10:52:39,302 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.67 sec
MapReduce Total cumulative CPU time: 2 seconds 670 msec
Ended Job = job_1656388870091_0013
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.67 sec   HDFS Read: 14531 HDFS Write: 125 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 670 msec
OK
1       India
4       usa
```

4. Count the number of sales done by each state

hive> select count(*) as count,State from salest group by State;

```
hive> select count(*) as count,State from sales  group by State;
Query ID = hdoop_20220628105508_89239f23-a9de-4525-985a-11f6637bc759
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_1656388870091_0014, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656388870091_0014/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1656388870091_0014
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 10:55:12,926 Stage-1 map = 0%,  reduce = 0%
2022-06-28 10:55:17,032 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.27 sec
2022-06-28 10:55:21,102 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.69 sec
MapReduce Total cumulative CPU time: 2 seconds 690 msec
Ended Job = job_1656388870091_0014
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.69 sec   HDFS Read: 14636 HDFS Write: 170 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 690 msec
OK
1       Karnataka
2       NJ
1       kk
1       washington
Time taken: 13.031 seconds, Fetched: 4 row(s)
```

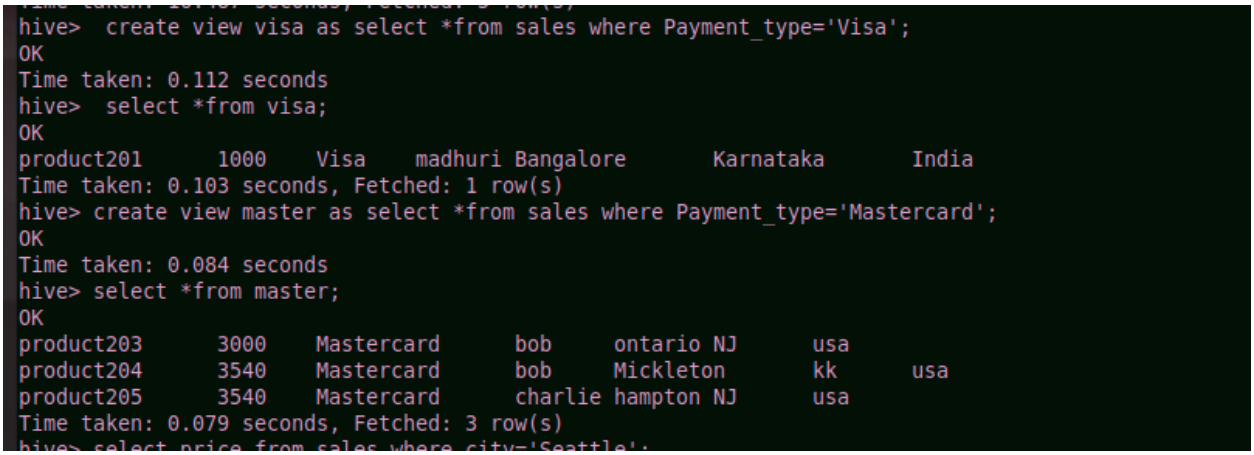
5. Display (Product, name) grouped by product

hive> select Product,name from sales group by Product,name;

```
FAILED: SemanticException [Error 10025]: Line 1:15 Expression not in GROUP BY key 'name'
hive>  select Product,name from sales group by Product,name;
Query ID = hdoop_20220628105628_5b68deea-e3a5-44f2-b1bd-431d2244cbc6
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_1656388870091_0015, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656388870091_0015/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1656388870091_0015
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 10:56:34,642 Stage-1 map = 0%,  reduce = 0%
2022-06-28 10:56:38,737 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.34 sec
2022-06-28 10:56:42,857 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.34 sec
MapReduce Total cumulative CPU time: 2 seconds 760 msec
Ended Job = job_1656388870091_0015
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 2.76 sec   HDFS Read: 13552 HDFS Write: 234 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 760 msec
OK
product201      madhuri
product202      shwetha
product203      bob
product204      bob
product205      charlie
Time taken: 16.487 seconds, Fetched: 5 row(s)
```

6. Create separate views for VISA and Mastercard

```
hive>  create view visa as select *from sales where Payment_type='Visa';
OK
Time taken: 0.112 seconds
hive>  select *from visa;
OK
product201 1000 Visa madhuri Bangalore Karnataka India
Time taken: 0.103 seconds, Fetched: 1 row(s)
hive> create view master as select *from sales where
Payment_type='Mastercard';
OK
Time taken: 0.084 seconds
hive> select *from master;
OK
product203 3000 Mastercard bob ontario NJ usa
product204 3540 Mastercard bob Mickleton kk usa
product205 3540 Mastercard charlie hampton NJ usa
Time taken: 0.079 seconds, Fetched: 3 row(s)
```



7. Show all the transactions done in Seattle

hive> select price from sales where city='seattle';

```
Time taken: 0.07 seconds, Fetched: 3 row(s)
hive> select price from sales where city='seattle';
OK
3000
Time taken: 0.064 seconds, Fetched: 1 row(s)
```

8. Find the max number of transactions done within the state of Ontario

hive> select max(Price) from sales where city='ontario' group by state;

```
Time taken: 15.537 seconds
hive>  select max(Price) from sales where city='ontario' group by state;
Query ID = hdoop_20220628110310_4b536da9-02b7-41a0-a3bf-e4f69790f7a4
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_1656388870091_0017, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656388870091_0017/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1656388870091_0017
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 11:03:15,925 Stage-1 map = 0%,  reduce = 0%
2022-06-28 11:03:20,107 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.22 sec
2022-06-28 11:03:24,198 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.8 sec
MapReduce Total cumulative CPU time: 3 seconds 800 msec
Ended Job = job_1656388870091_0017
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 3.8 sec   HDFS Read: 15812 HDFS Write: 104 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 800 msec
OK
3000
Time taken: 16.011 seconds, Fetched: 1 row(s)
```

9. Find the number of transactions whose price is in between 1500-3600

hive> select count(*) from sales where Price between 1500 and 3600;

```
Time taken: 16.011 seconds, Fetched: 1 row(s)
hive> select count(*) from sales where Price between 1500 and 3600;
Query ID = hdoop_20220628110423_92f96b23-0479-4752-9951-6c28b6a56216
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 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_1656388870091_0018, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656388870091_0018/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job  -kill job_1656388870091_0018
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 11:04:29,422 Stage-1 map = 0%,  reduce = 0%
2022-06-28 11:04:33,525 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.28 sec
2022-06-28 11:04:37,660 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.01 sec
MapReduce Total cumulative CPU time: 4 seconds 10 msec
Ended Job = job_1656388870091_0018
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.01 sec   HDFS Read: 15652 HDFS Write: 101 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 10 msec
OK
4
Time taken: 15.946 seconds, Fetched: 1 row(s)
```

10. List all the transactions done in the United States using Mastercard

hive> select *from sales where Country='usa' and Payment_Type='Mastercard';