# 1NT19IS108 CHAITANYA P C1 BATCH

```
hdoop@admin1-HP-280-G4-NT-Business-PC:- cd SHADDOP_HOME
bdoop@admin1-HP-280-G4-NT-Business-PC:- cd SHADDOP_HOME
bdoop@admin1-HP-280-G4-NT-Business-PC:- (hadoop-3.2.1) so ds bin/
bdoop@admin1-HP-280-
```

# Dropping database Rm -rf metastore db

```
parquet-logging.properties
hdoopgadmin1-HP-280-G4-MT-Business-PC:-/apache-hive-3.1.2-bin/conf$ rm -rf metastore_db
hdoopgadmin1-HP-280-G4-MT-Business-PC:-/apache-hive-3.1.2-bin/conf$ schematool -initSchema -dbType derby
SLF43: Class path contains multiple SLF4J bindings.
SLF43: 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]
```

```
hdoop@admin1-HP-280-G4-MT-Business-PC: ~/apache-hive-3.1.2-bin/conf
ndoopgadMin1-HP-280-G4-M1-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.c
lass]
                                                                   PC:~/apache-hive-3.1.2-bin/confS hive
CUSF):
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 = a36d5c8e-d0d8-41a2-9832-9af9457d4436
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: t
THE
HIVE SESSION ID = 3e60ae10-704c-470e-acf4-2b988cbcd39a
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;
default
Time taken: 0.446 seconds, Fetched: 1 row(s) hive> create database if not exists salesdb;
Time taken: 0.076 seconds
hive> use salesdb;
uk
Time taken: 0.03 seconds
hive> create table sales(product string,price int,payment_type string,name string,city string,state string,country string);
itme taken: 0.502 seconds
hive> load data local inpath '/home/hdoop/sales.csv' into table sales;
FAILED: SemanticException Line 1:23 Invalid path ''/home/hdoop/sales.csv'': No files matching path file:/home/hdoop/sales.csv
hive> load data local inpath '/home/hdoop/Desktop/sales.csv' into table sales;
Loading data to table salesdb.sales
Time taken: 0.859 seconds hive> select * from sales;
,,,,,, NULL NULL NULL NULL NULL NOLL NOLL
Product1,1200,Mastercard,carolina,Basildon,England,United Kingdom
Product1,1200,Visa,Betina,Parkville,MO,United States NULL NU
                                           NULL
                                                           NULL
                                                                          NULL
                                                                                                                                    NULL
                                                                                                                                                      NULL
                                                                                                                                                                     NULL
                                                                                                                                                                                    NULL
                                                                                                                                                                                                   NULL
                                                                                                                                                                                                                   NULL
```

# show databases; create database if not exists chaitanyadb; use chaitanyadb;

```
hive> show databases;
OK

default
salesdb
Time taken: 0.039 seconds, Fetched: 2 row(s)
hive> create database if not exists chaitanyadb;
OK
Time taken: 0.102 seconds
hive> use chaitanyadb;
OK
Time taken: 0.33 seconds
hive> use chaitanyadb;
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: 0.094 seconds
hive> 's insert into sales values('product1,5000,'Mastercard','harry,'Astoria','California','USA');
MismatchedTokenException(24!=374)
    at org.antlr.runtime.BaseRecognizer.recoverFromMismatchedToken(BaseRecognizer.java:617)
    at org.antlr.runtime.BaseRecognizer.match(BaseRecognizer.java:115)
    at org.apache.hadoop.hive.ql.parse.HiveParser_IdentifiersParser.expressionsInParenthesis(HiveParser_IdentifiersParser.java:2316)
    at org.apache.hadoop.hive.ql.parse.HiveParser_expressionsInParenthesis(HiveParser_FromClauseParser.java:6214)
    at org.apache.hadoop.hive.ql.parse.HiveParser_FromClauseParser.valueRowConstructor(HiveParser_FromClauseParser.java:6131)
```

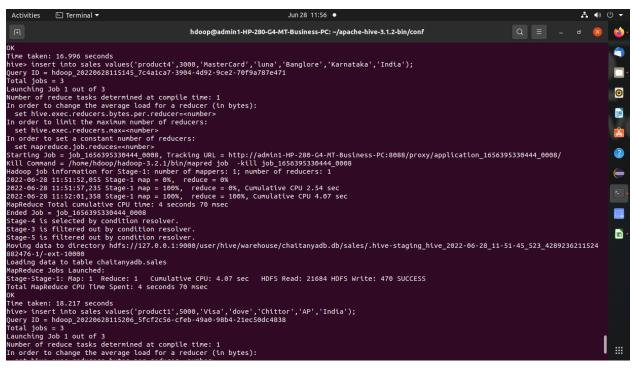
### Creating table:

create table sales(product string,price int,payment\_type string,name string,city string,state string,country string);

insert into sales values('product1,5000,'Mastercard','harry,'Astoria','California','USA'); insert into sales values('product2',1000,'Visa','ron','Banglore','Karnataka','India');

insert into sales values('product3',3000,'Diners','draco','Banglore','Karnataka','India'); insert into sales values('product4',3000,'MasterCard','luna','Banglore','Karnataka','India'); insert into sales values('product1',5000,'Visa','dove','Chittor','AP','India'); insert into sales values('product2',6000,'Rupay','elena','ATP','AP','India');

```
hive> select * from sales;
OK
               5000
                                               Astoria California
product1
                       Mastercard
                                       harry
product2
                       Visa
                                       Banglore
                                                       Karnataka
                                                                       India
               1000
                               ron
                       Diners draco
                                                                       India
product3
                                       Banglore
                                                       Karnataka
               3000
product4
                                                                               India
               3000
                       MasterCard
                                       luna
                                               Banglore
                                                               Karnataka
                                                       India
product1
               5000
                       Visa
                               dove
                                       Chittor AP
product2
               6000
                       Rupay
                               elena
                                       ATP
                                               AP
                                                       India
Time taken: 0.07 seconds, Fetched: 6 row(s)
```



#### Dropping the table:

create table Salestb(Product string,Price int,Payment\_type string,Name string,City string,State string,Country string)
drop table Salestb;

```
hive> create table Salestb(Product string,Price int,Payment_type string,Name string,City string,State string,Country string);
OK
Time taken: 0.056 seconds
hive> drop table Salestb;
OK
Time taken: 0.841 seconds
hive>
```

#### Alter table:

alter table sales rename to chaitanya\_sales;

```
hive> alter table sales rename to chaitanya_sales;

OK

Time taken: 0.137 seconds

hive> show tables;

OK

chaitanya_sales

Time taken: 0.014 seconds, Fetched: 1 row(s)

hive>
```

## 3. Built in operators:

Arithmetic Operator in Hive

select product,name,price+1100 from sales;

```
chaitanya_sales
Time taken: 0.014 seconds, Fetched: 1 row(s)
hive> select product,name,price+1100 from chaitanya_sales;
OK
product1
                harry
                        6100
product2
                ron
                        2100
product3
                draco
                       4100
product4
                luna
                       4100
product1
                dove
                        6100
product2
                elena
                        7100
Time taken: 0.097 seconds, Fetched: 6 row(s)
hive>
```

```
hive> select product,name,price+1100 from sales;
OK
product1
                harry
                        6100
product2
                ron
                        2100
product3
                draco
                        4100
product4
                luna
                       4100
product1
                dove
                        6100
product2
                elena
                        7100
Time taken: 0.187 seconds, Fetched: 6 row(s)
hive>
```

select product,name,(price\*10)/2 from sales;

```
hive> select product,name,(price*10)/2 from sales;
OK
product1
                harry
                         25000.0
product2
                ron
                         5000.0
product3
                draco
                         15000.0
product4
                luna
                         15000.0
product1
                dove
                         25000.0
product2
                         30000.0
                elena
Time taken: 0.1 seconds, Fetched: 6 row(s)
```

Relational Operators in Hive select \* from sales where price>4000;

```
hive> select * from 1nt19is107_sales where price>4000;
FAILED: SemanticException [Error 10001]: Line 1:14 Table not found '1nt19is107_sales'
hive> select * from sales where price>4000;
OK
                5000
                                                astoria california
                                                                         USA
product1
                        mastercard
                                        harry
product1
                5000
                        Visa
                                dove
                                        Chittor AP
                                                         India
product2
                6000
                        Rupay
                                elena
                                        ATP
                                                AP
                                                         India
Time taken: 0.099 seconds, Fetched: 3 row(s)
```

#### 4. Built in functions:

Mathematical Functions in Hive select product.name.sqrt(price) from sales:

```
hive> select product,name,sqrt(price) from sales;
OK
product1
                        70.71067811865476
                harry
product2
                ron
                        31.622776601683793
product3
                draco
                        54.772255750516614
product4
                luna
                        54.772255750516614
product1
                        70.71067811865476
                dove
product2
                elena
                        77.45966692414834
Time taken: 0.103 seconds, Fetched: 6 row(s)
```

select product,name,floor(price) from sales;

```
Time taken: 0.103 seconds, Fetched: 6 row(s)
hive> select product,name,floor(price) from sales;
OK
product1
                harrv
                        5000
product2
                ron
                        1000
product3
                draco
                        3000
product4
                luna
                        3000
product1
                dove
                        5000
product2
                elena
                        6000
Time taken: 0.107 seconds, Fetched: 6 row(s)
hive> select product,name,cos(price) from sales;
```

select product,name,cos(price) from sales;

```
hive> select product,name,cos(price) from sales;
product1
                harry 0.15466840618074712
product2
                        0.562379076290703
                ron
product3
                draco -0.9756821998857504
product4
                luna
                       -0.9756821998857504
                       0.15466840618074712
product1
                dove
product2
                       0.9039115103477952
                elena
Time taken: 0.117 seconds, Fetched: 6 row(s)
hive> select max(price) from sales where state=Karn
Aggregate Functions in Hive
```

select max(price) from sales where state=Karnataka;

```
hive> select max(price) from sales where state='Karnataka';
Query ID = hdoop_20220628134704_026eb5a1-00fc-477f-b8ef-56377f113a08
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=cnumber>
In order to limit the maximum number of reducers:
    set hive.exec.reducers.max=cnumber>
In order to set a constant number of reducers:
    set mapreduce.job.reduces=<number>
Starting Job = job_1656401979434_0007, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656401979434_0007/
Kill Command = /home/hdoop/hadopo-3.2.1/bin/mapred job -kill job_1656401979434_0007
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-66-28 13:47:10,006 Stage-1 map = 0%, reduce = 0%
2022-66-28 13:47:10,106 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.3 sec
2022-06-28 13:47:19,232 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.89 sec
MapReduce Total cumulative CPU time: 3 seconds 890 msec
Ended Job = job_1656401979434_0007
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.89 sec HDFS Read: 15461 HDFS Write: 104 SUCCESS
```

Other built-in functions select product, upper(name) from sales;

select product, reverse(name) from sales;

```
Time taken: 0.075 seconds, Fetched: 6 row(s)
hive> select product, reverse(name) from sales;
OK
product1
                yrrah
product2
                nor
product3
                ocard
product4
                anul
product1
                evod
product2
                anele
Time taken: 0.098 seconds, Fetched: 6 row(s)
```

select product, concat(name, state) from sales;

```
Time taken: 0.098 seconds, Fetched: 6 row(s)
hive> select product, concat(name,state) from sales;
OK
product1 harrycalifornia
product2 ronKarnataka
product3 dracoKarnataka
product4 lunaKarnataka
product1 doveAP
product2 elenaAP
```

select product, trim(name) from sales;

```
Time taken: 0.096 seconds, Fetched: 6 row(s)
hive> select product, trim(name) from sales;
OK
product1
                harry
product2
                ron
product3
                draco
product4
                luna
product1
                dove
product2
                elena
Time taken: 0.089 seconds, Fetched: 6 row(s)
```

#### 5.VIEWS:

create view visa as select \* from sales where payment\_type='Visa'; select \* from visa;

```
nive> create view visa as select * from sales where payment_type='Visa';
DK
Time taken: 0.1 seconds
nive> select * from visa;
hive> select * from visa;
OK
product2
                1000
                        Visa
                                гоп
                                        Banglore
                                                        Karnataka
                                                                        India
product1
                                        Chittor AP
                5000
                        Visa
                                dove
                                                        India
Time taken: 0.1 seconds, Fetched: 2 row(s)
hive>
```

create view mastercard as select \* from sales where payment\_type=MasterCard; select \* from mastercar;

```
hive> create view mastercar as select * from sales where payment_type='MasterCard';
OK
Time taken: 0.124 seconds
hive> select * from mastercar;
OK
product4 3000 MasterCard luna Banglore Karnataka India
Time taken: 0.102 seconds, Fetched: 1 row(s)
```

#### 6.HIVEQL

#### **GROUPBY:**

select count(\*)as count, country from sales group by country;

```
hive> select count(*)as count,country from sales group by country;
Query ID = hdoop_20220628140240_56837ef9-a2da-4c70-b264-587855bba5de
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 a constant number of reducers:
    set maximum produce; job.preduces=senumber>
Starting Job = job_1656401979434_0008, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656401979434_0008/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job_1656401979434_0008
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job_1656401979434_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 14:02:45,939 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 1.83 sec
2022-06-28 14:02:55,169 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.04 sec
MapReduce Total cumulative CPU time: 3 seconds 40 msec
Ended Job = job_1656401979434_0008
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.04 sec HDFS Read: 14708 HDFS Write: 125 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 40 msec

S India
1 USA

Lime taken: 16 082 seconds Eatched: 2 source)
```

select count(\*) as count, state from sales group by state;

```
hive> select count(*) as count,state from sales group by state;
Query ID = hdoop_20220628140302_df4ddbf8-df09-4de8-8ee2-9efad3ed66b8
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 htve.exec.reducers.max==number>
In order to set a constant number of reducers:
    set mapreduce.job.reduces=<number>
Starting Job = job_1656401979434_0009, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1656401979434_0009/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job_1656401979434_0009
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 14:03:08,797 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 1.83 sec
2022-06-28 14:03:17,060 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.29 sec
MapReduce Total cumulative CPU time: 3 seconds 290 msec
Ended Job = job_1656401979434_0009
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.29 sec HDFS Read: 14800 HDFS Write: 153 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 290 msec
OK
```

```
Total MapReduce CPU Time Spent: 3 seconds 290 msec
OK
2 AP
3 Karnataka
1 california
Time taken: 16.216 seconds, Fetched: 3 row(s)
```

select product,name from sales group by product,name;

```
hive> select product,name from sales group by product,name;
Query ID = hdoop_20220628140325_b651616c-8164-4b3c-a616-dbec8ef8bd3b
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_1656401979434_0010, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job_1656401979434_0010
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 14:03:31,074 Stage-1 map = 0%, reduce = 0%
2022-06-28 14:03:35,226 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.55 sec
2022-06-28 14:03:40,378 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.02 sec
MapReduce Total cumulative CPU time: 3 seconds 20 msec
Ended Job = job 1656401979434 0010
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.02 sec HDFS Read: 13723 HDFS Write: 245 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 20 msec
OK
product1
                    dove
product1
                    harry
product2
                    elena
product2
                    ron
product3
                    draco
product4
                    luna
```

#### WHERE:

select \* from sales where state=Karnataka:

```
hive> select * from sales where state='Karnataka';
oĸ
product2
                1000
                        Visa
                                ron
                                         Banglore
                                                         Karnataka
                                                                          India
product3
                3000
                        Diners draco
                                         Banglore
                                                         Karnataka
                                                                          India
product4
                3000
                        MasterCard
                                         luna
                                                 Banglore
                                                                 Karnataka
                                                                                  India
Time taken: 0.063 seconds, Fetched: 3 row(s)
hive> select * from sales order by price;
```

## **ORDERBY**:

select \* from sales order by price;

```
hive> select * from sales order by price;
Query ID = hdoop_20220628140413_0b0554a2-7ff2-4621-b870-681e45d24e4d
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>
set mapreduce.job.reduces=<number>
Starting Job = job_1656401979434_0011, Tracking URL = http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_165
Kill Command = /home/hdoop/hadoop-3.2.1/bin/mapred job -kill job_1656401979434_0011
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-06-28 14:04:19,375 Stage-1 map = 0%, reduce = 0%
2022-06-28 14:04:23,492 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.41 sec
2022-06-28 14:04:28,629 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.75 sec
MapReduce Total cumulative CPU time: 2 seconds 750 msec
Ended Job = job_1656401979434_0011
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 2.75 sec HDFS Read: 13315 HDFS Write: 448 SUCCESS
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 2.75 sec HDFS Read: 13315 HDFS Write: 448 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 750 msec
OK
                                                Visa ron Banglore Karnataka
MasterCard luna Banglore Kar
Diners draco Banglore Karnataka
Visa dove Chittor AP India
product2
                                 1000
                                                                                                                                                   India
                                 3000
                                                                                                                                                                     India
product4
                                                                                                                            Karnataka
product3
                                                                                                                                                    India
Marnataka chittor AP India harry astoria california Time taken: 15.922 seconds, Fetched: 6 row(s)
                                 3000
                                                                                                                                                    USA
```