Starting the Hive Service:

```
[hduser@localhost -]$ hive --service metastore
ls: cannot access /usr/local/spark/lib/spark-assembly-*.jar: No such file or directory
Starting Hive Metastore Server
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.mapjoin.optimized.keys does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.mapjoin.lazy.hashtable does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.thrift.http.min.worker.threads does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.thrift.http.min.worker.threads does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.thrift.http.max.worker.threads does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.torgint.http.max.worker.threads does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:16:03 WARN conf.HiveConf of name hive.server2.logging.ope
```

Connecting to beeline:

```
[Induser@localhost -]$ beeline
ls: cannot access /usr/local/spark/lib/spark-assembly-*.jar: No such file or directory
Beeline version 1.2.2 by Apache Hive
beeline> !connect jdbc:hive2://
SLF43: Class path contains multiple SLF4J bindings.
SLF43: Found binding in [jar:file:/usr/local/hbase/lib/phoenix-4.11.0-HBase-0.98-client.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF41: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF41: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF41: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF41: See http://www.slf4j.org/codes.htmi#multiple bindings for an explanation.
SLF41: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
Connecting to jdbc:hive2://
Enter username for jdbc:hive2://: scott
Enter password for jdbc:hive2://: scott
Enter 18:138:12 [main]: wARN conf. HiveConf: HiveConf of name hive.mapjoin.lazy.hashtable does not exist
22/07/04 15:188:12 [main]: wARN conf. HiveConf: HiveConf of name hive.mapjoin.lazy.hashtable does not exist
22/07/04 15:188:12 [main]: wARN conf. HiveConf: HiveConf of name hive.server2.thrift.http.min.worker.threads does not exist
22/07/04 15:188:12 [main]: wARN conf. HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:188:12 [main]: wARN conf. HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:188:12 [main]: wARN conf. HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:188:12 [main]: wARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.verbose does not exist
22/07/04 15:188:12 [main]: wARN conf.HiveConf: HiveConf of name hive.server2.logging.operation.ver
```

Static Partitioning

Creating Files:

```
[hduser@localhost ~]$ gedit Hyd.txt
[hduser@localhost ~]$ gedit Ban.txt
[hduser@localhost ~]$ cat Hyd.txt
1,Rahul,20,14000
2,Ramesh,30,14500
3,Guru,35,17850
[hduser@localhost ~]$ gedit Ban.txt
[hduser@localhost ~]$ cat Ban.txt
4,Giri,23,13890
5,Mohan,45,65000
5,Venu,40,45000
[hduser@localhost ~]$ gedit Delhi.txt
[hduser@localhost ~]$ cat Delhi.txt
7,Aravind,45,80000
8,Mukesh,50,45600
```

Moving to HDFS:

Table Creation:

```
0: jdbc:hive2://> create table Static_Partition_Demo(eid int,ename string,eage int) partitioned by(eloc string) row format delimited fields terminated by ','
lines terminated by 'n';
OK
No rows affected (7.046 seconds)
```

Loading the Data using partitions:

```
0: jdbc:hive2://> load data inpath "/user/hduser/hive_practice/Hyd.txt" into table Static_Partition_Demo partition(eloc = "Hyd");
Loading data to table practice.static_partition_demo partition (eloc=Hyd)
Partition practice.static_partition_demo(eloc=Hyd) stats: [numFiles=1, numRows=0, totalSize=51, rawDataSize=0]
OK
No rows affected (3.609 seconds)
O: jdbc:hive2://> load data inpath "/user/hduser/hive_practice/Ban.txt" into table Static_Partition_Demo partition(eloc = "Ban");
Loading data to table practice.static_partition_demo partition (eloc=Ban)
Partition practice.static_partition_demo(eloc=Ban) stats: [numFiles=1, numRows=0, totalSize=49, rawDataSize=0]
OK
No rows affected (1.322 seconds)
O: jdbc:hive2://> load data inpath "/user/hduser/hive_practice/Delhi.txt" into table Static_Partition_Demo partition(eloc = "Delhi");
Loading data to table practice.static_partition_demo partition (eloc=Delhi)
Partition practice.static_partition_demo(eloc=Delhi) stats: [numFiles=1, numRows=0, totalSize=37, rawDataSize=0]
OK
No rows affected (1.473 seconds)
O: jdbc:hive2://>
O: jdbc:hive2://>
O: jdbc:hive2://>
```

```
[hduser@localhost ~]$ hdfs dfs -ls /user/hive/warehouse/practice.db/static_partition_demo
22/07/04 16:09:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 3 items
drwxr-xr-x - hduser supergroup 0 2022-07-04 16:05 /user/hive/warehouse/practice.db/static_partition_demo/eloc=Ban
drwxr-xr-x - hduser supergroup 0 2022-07-04 16:05 /user/hive/warehouse/practice.db/static_partition_demo/eloc=Delhi
drwxr-xr-x - hduser supergroup 0 2022-07-04 16:05 /user/hive/warehouse/practice.db/static_partition_demo/eloc=Hyd
```

Dynamic Partitioning

```
[hduser@localhost -]$ gedit Customers.txt
[hduser@localhost -]$ cat Customers.txt
1,Kumar,23,Hyderabad
2,Mahesh,34,Mumbai
3,Suresh,24,Hyderabad
4,Gopi,45,Pune
5,Surender,45,Pune
6,Krishna,56,Hyderabad
[hduser@localhost -]$ hdfs dfs -ls /user/hduser/hive_practice
22/07/04 16:17:29 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[hduser@localhost -]$ hdfs dfs -put Customers.txt /user/hduser/hive_practice
22/07/04 16:17:33 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[hduser@localhost -]$ hdfs dfs -ls /user/hduser/hive_practice
22/07/04 16:17:42 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[hduser@localhost -]$ hdfs dfs -ls /user/hduser/hive_practice
22/07/04 16:17:42 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r- 1 hduser supergroup
119 2022-07-04 16:17 /user/hduser/hive_practice/Customers.txt
[hduser@localhost -]$
```

Creating table without Partitions and loading the data:

Creating a table with Partition Column:

Inserting data from previous table to partition containing table:

```
hive> set hive.exec.dynamic.partition = true;
hive> set hive.exec.dynamic.partition.mode = nonstrict;
hive> set hive.exec.dynamic.partition.mode = nonstrict;
hive> insert into customer_part partition(cloc) select * from customer_nopart;
Query ID = hduser_20220704162917_405137640-7557-4092-a446-c9f844008a63
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Loading class `com.mysql.jdbc.Driver'. This is deprecated. The new driver class is `com.mysql.cj.jdbc.Driver'. The driver is automatically registered via the
SFI and manual loading of the driver class is generally unnecessary.
Starting Job = job_1636927935583_0003, Tracking URL = http://Inceptez:8088/proxy/application_1656927935583_0003/
Kill Command = /usr/local/hadoop/pib.n/hadoop job -kill job_1656927935583_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
```

```
hive> select * from customer_part;
OK

1 Kumar 23 Hyderabad
3 Suresh 24 Hyderabad
6 Krishna 56 Hyderabad
2 Mahesh 34 Humbai
4 Gopi 45 Pune
5 Surender 45 Pune
5 Time taken: 0.578 seconds, Fetched: 6 row(s)
hive>
```

```
hive> show partitions customer_part;
OK
cloc=Hyderabad
cloc=Pune
Time taken: 0.403 seconds, Fetched: 3 row(s)
hive>
```

```
[hduser@localhost -]$ hdfs dfs -ls /user/hive/warehouse/practice.db/customer_part
22/07/04 16:32:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 3 items
drwxr-xr-x - hduser supergroup
0 2022-07-04 16:30 /user/hive/warehouse/practice.db/customer_part/cloc=Hyderabad
drwxr-xr-x - hduser supergroup
0 2022-07-04 16:30 /user/hive/warehouse/practice.db/customer_part/cloc=Humbai
drwxr-xr-x - hduser supergroup
0 2022-07-04 16:30 /user/hive/warehouse/practice.db/customer_part/cloc=Pune
[hduser@localhost -]$
```

```
[hduser@localhost -]$ hdfs dfs -ls /user/hive/warehouse/practice.db/customer part/cloc=Pune
22/07/04 16:32:39 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r- 1 hduser supergroup 24 2022-07-04 16:30 /user/hive/warehouse/practice.db/customer_part/cloc=Pune/000000_0
[hduser@localhost -]$ hdfs dfs -ls /user/hive/warehouse/practice.db/customer_part/cloc=Pune/000000_0
22/07/04 16:32:48 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r- 1 hduser supergroup 36 2022-07-04 16:30 /user/hive/warehouse/practice.db/customer_part/cloc=Hyderabad/000000_0
[Induser@localhost -]$ hdfs dfs -ls /user/hive/warehouse/practice.db/customer part/cloc=Humbai
22/07/04 16:32:56 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r- 1 hduser supergroup 12 2022-07-04 16:30 /user/hive/warehouse/practice.db/customer_part/cloc=Humbai/000000_0
[hduser@localhost -]$ 12 2022-07-04 16:30 /user/hive/warehouse/practice.db/customer_part/cloc=Humbai/000000_0
[hduser@localhost -]$
```

Bucketing

Creating table with some data without buckets:

Table creation with buckets:

```
hive> create table customer_buck(cid int,cname string,cage int,cloc string) clustered by(cloc) into 3 buckets row format delimited fields terminated by ',' li nes terminated by '\n'; 0K

Time taken: 0.635 seconds
hive> desc customer_buck;
0K
cid int
cname string
cage int
cloc string
Time taken: 0.25 seconds, Fetched: 4 row(s)
hive> desc customer_buck; 0K
```

Loading the data into new table with buckets:

```
nive insert into customer_buck select * from customer_nopart;

Query ID = hduser_20220704163606_d8ff9093-9bb0-4a70-8442-bf8120d08727

Total jobs = 1

Launching Job 1 out of 1

Number of reduce tasks determined at compile time: 3

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 hive.exec.reducers.max=<number>

Starting Job = job_1656927935583_0004, Tracking URL = http://Inceptez:8088/proxy/application_1656927935583_0004/

Kill Command = /usr/Local/hadoop/bin/hadoop job -kill job_1656927935583_0004

Kill Command = /usr/Local/hadoop/bin/hadoop job -kill job_1656927935583_0004

Kall Command = /usr/Local/hadoop/bin/hadoop job -kill job_1656927935583_0004

Kill Command = /usr/Local/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hadoop/bin/hado
```

```
hive> select * from customer_buck;
OK

Surender 45 Pune
4 Gopi 45 Pune
2 Mahesh 34 Mumbai
6 Krishna 56 Hyderabad
3 Suresh 24 Hyderabad
1 Kumar 23 Hyderabad
1 Time taken: 0.483 seconds, Fetched: 6 row(s)
hive>
```

```
[hduser@localhost -]$ hdfs dfs -ls /user/hive/warehouse/practice.db/customer_buck
22/07/04 16:38:46 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 3 items
--W------ 1 hduser supergroup 34 2022-07-04 16:37 /user/hive/warehouse/practice.db/customer_buck/0000000_0
--W------- 1 hduser supergroup 19 2022-07-04 16:37 /user/hive/warehouse/practice.db/customer_buck/0000001_0
--W------- 1 hduser supergroup 66 2022-07-04 16:37 /user/hive/warehouse/practice.db/customer_buck/0000002_0
[hduser@localhost -]$
```

Partition with Bucketing:

Creating table with some data without buckets and partitions:

Creation of table with bucket and partition:

```
hive> create table customer_partbuck(cid int,cname string,cage int) partitioned by(cloc string) clustered by(cid) into 3 buckets row format delimited fields to erminated by '\n';

OK

Time taken: 0.423 seconds
hive> desc customer_partbuck;

OK

cid int

cname string

cage int

cloc string

# Partition Information
# col_name data_type comment

cloc string

Time taken: 0.245 seconds, Fetched: 9 row(s)
hive>
```

Loading the data:

```
hive> insert into customer_partbuck partition(cloc) select * from customer_nopart;

Query ID = hduser_20220704164419_7e874635-fc66-480d-94b5-2elblbee4lac

Total jobs = 1

Launching Job | out of 1

Number of reduce tasks determined at compile time: 3

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>
```

```
hive> show partitions customer_partbuck;

OK

cloc=Hyderabad

cloc=Pune

Time taken: 0.471 seconds, Fetched: 3 row(s)

hive>

Image: Partition of the partitio
```