

STATIC PARTITION:

```
hive> create table orders_w_partition(id string,customer_id string,product_id string,quantity int,amount double,
zipcode char(5))
> partitioned by (state char(2))
> row format delimited
> fields terminated by ',';
OK
Time taken: 32.13 seconds
hive> load data local inpath '/home/hduser/order_ca.csv' into table orders_w_partition partition(state='CA');
Loading data to table hive1.orders_w_partition partition (state=CA)
Partition hive1.orders_w_partition{state=CA} stats: [numFiles=1, numRows=0, totalSize=88, rawDataSize=0]
OK
Time taken: 13.967 seconds
```

order_ca.csv

id	customer_id	product_id	quantity	amount	zipcode
o1	c1	p1	1	1.11	90111
o2	c2	p2	1	2.22	90222
o3	c3	p3	1	3.33	90333
o4	c4	p4	1	4.44	90444

order_ct.csv

id	customer_id	product_id	quantity	amount	zipcode
o1	c1	p1	1	1.11	90555
o2	c2	p2	1	2.22	90666
o3	c3	p3	1	3.33	90777

```
hive> load data local inpath '/home/hduser/order_ct.csv' into table orders_w_partition partition(state='CT');
Loading data to table hive1.orders_w_partition partition (state=CT)
Partition hive1.orders_w_partition{state=CT} stats: [numFiles=1, numRows=0, totalSize=66, rawDataSize=0]
OK
Time taken: 2.484 seconds
hive> show partitions orders_w_partition
> ;
OK
state=CA
state=CT
Time taken: 0.662 seconds, Fetched: 2 row(s)
hive>
```

```
hduser@localhost ~]$ hdfs dfs -ls /user/hive/warehouse/hive1.db/orders_w_partition
2/07/04 16:27:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using
builtin-java classes where applicable
Found 2 items
-rwxr-xr-x - hduser supergroup          0 2022-07-04 16:13 /user/hive/warehouse/hive1.db/orders_w_partition
state=CA
-rwxr-xr-x - hduser supergroup          0 2022-07-04 16:23 /user/hive/warehouse/hive1.db/orders_w_partition
state=CT
hduser@localhost ~]$
```


```
hive> select id,state,customer_id,product_id from orders_w_partition where state="CA";
OK
o1      CA      c1      p1
o2      CA      c2      p2
o3      CA      c3      p3
o4      CA      c4      p4
Time taken: 15.443 seconds, Fetched: 4 row(s)
hive>
```

DYNAMIC PARTITION:

orders_CT_with_state.csv

orders_CA_with_state.csv

o1,c1,p1,1,1.11,90111,CA
o2,c2,p2,1,2.22,90222,CA
o3,c3,p3,1,3.33,90333,CA
o4,c4,p4,1,4.44,90444,CA

Open 

orders_CT_with_state.csv

```
hive> create table orders_no_partition(id string,customer_id string,product_id string,quantity int,amount double,zipcode char(5),state char(2))
> row format delimited
> fields terminated by ',';
```

```
hive> load data local inpath '/home/hduser/orders_CA_with_state.csv' into table orders_no_partition;
Loading data to table default.orders_no_partition
Table default.orders_no_partition stats: [numFiles=1, totalSize=100]
OK
Time taken: 2.188 seconds
hive> load data local inpath '/home/hduser/orders_CT_with_state.csv' into table orders_no_partition;
Loading data to table default.orders_no_partition
Table default.orders_no_partition stats: [numFiles=2, totalSize=175]
OK
Time taken: 1.196 seconds
hive> select * from orders_no_partition;
OK
o1      c1      p1      1      1.11    90111   CA
o2      c2      p2      1      2.22    90222   CA
o3      c3      p3      1      3.33    90333   CA
o4      c4      p4      1      4.44    90444   CA
o1      c1      p1      1      1.11    90555   CT
o2      c2      p2      1      2.22    90666   CT
o3      c3      p3      1      3.33    90777   CT
Time taken: 1.414 seconds, Fetched: 7 row(s)
hive>
```

```
hive> create table orders(id string,customer_id string,product_id string,quantity int,amount double,zipcode char(5)) partitioned by (state char(2))
> row format delimited
> fields terminated by ',';
OK
Time taken: 0.369 seconds
hive> insert into table orders partition(state)
> select * from orders no partition;
```

```
hive> show partitions orders;
OK
state=CA
state=CT
Time taken: 0.446 seconds, Fetched: 2 row(s)
hive>
```

```
[hduser@localhost ~]$ hdfs dfs -ls /user/hive/warehouse/hive1.db/orders
22/07/04 19:39:42 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
drwxr-xr-x - hduser supergroup          0 2022-07-04 19:35 /user/hive/warehouse/hive1.db/orders/state=CA
drwxr-xr-x - hduser supergroup          0 2022-07-04 19:35 /user/hive/warehouse/hive1.db/orders/state=CT
[hduser@localhost ~]$
```

BUCKETING:

```
Open [icon] *newproducts.csv
~/
1,iphon379.99,mobiles
2,doll,8.99,toys
3,Galaxy X,100.0,mobile
5,Nokia Y,39.99,mobile
6,truck,7.99,toys
7,makeup,100.0,fashion
8,earrings,69.0,fashion
9,chair,129.0,furniture
10,table,269.0,furniture
11,waterpistol,9.0,toys|
```

```
hive> create table products_no_bucket(id int,name string,cost double,category string)
> row format delimited
> fields terminated by ",";
OK
Time taken: 0.327 seconds
hive> load data local inpath '/home/hduser/newproducts.csv' into table products_no_bucket;
Loading data to table hive1.products_no_bucket
Table hive1.products_no_bucket stats: [numFiles=1, totalSize=224]
OK
Time taken: 0.969 seconds
hive> █
```

```
[hduser@localhost ~]$ hdfs dfs -ls /user/hive/warehouse/hive1.db/products_no_bucket
22/07/04 20:04:51 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 224 2022-07-04 20:03 /user/hive/warehouse/hive1.db/products_no_bucket/newproducts.csv
[hduser@localhost ~]$ █
```

```
hive> from products_no_bucket insert into table products_w_bucket select id,name,cost,category;
```


```
[hduser@localhost ~]$ hdfs dfs -ls /user/hive/warehouse/hive1.db/products_w_bucket
22/07/04 20:11:34 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 4 items
-rw-r--r-- 1 hduser supergroup 24 2022-07-04 20:11 /user/hive/warehouse/hive1.db/products_w_bucket/000000_0
-rw-r--r-- 1 hduser supergroup 67 2022-07-04 20:11 /user/hive/warehouse/hive1.db/products_w_bucket/000001_0
-rw-r--r-- 1 hduser supergroup 60 2022-07-04 20:11 /user/hive/warehouse/hive1.db/products_w_bucket/000002_0
-rw-r--r-- 1 hduser supergroup 71 2022-07-04 20:11 /user/hive/warehouse/hive1.db/products_w_bucket/000003_0
[hduser@localhost ~]$ █
```

```

hive> select * from products_w_bucket tablesample(bucket 1 out of 4);
OK
8      earrings      69.0    fashion
Time taken: 0.297 seconds, Fetched: 1 row(s)
hive> select * from products_w_bucket tablesample(bucket 2 out of 4);
OK
9      chair    129.0    furniture
5      Nokia Y 39.99    mobile
1      iphon379.99      NULL    NULL
Time taken: 0.172 seconds, Fetched: 3 row(s)
hive> select * from products_w_bucket tablesample(bucket 3 out of 4);
OK
10     table    269.0    furniture
6      truck    7.99     toys
2      doll     8.99     toys
Time taken: 0.159 seconds, Fetched: 3 row(s)
hive> select * from products_w_bucket tablesample(bucket 4 out of 4);
OK
11     waterpistol    9.0     toys
7      makeup    100.0    fashion
3      Galaxy X    100.0    mobile
Time taken: 0.159 seconds, Fetched: 3 row(s)
hive>

```

Partitioning with 2 columns:

Open  orders_country_w_states.csv

```

o1,c1,p1,1,1.11,90111,WA,S1
o2,c2,p2,1,2.22,90222,CA,S2
o3,c3,p3,1,3.33,90333,WA,S1
o4,c4,p4,1,4.44,90444,CA,S2

```

```

hive> create table orders_no_partition1(id string,customer_id string,product_id string,quantity int,amount double,zipcode char(5),country char(2),state char(2))
> row format delimited
> fields terminated by ",";
OK
Time taken: 0.38 seconds
hive> load data local inpath '/home/hduser/orders_country_w_states.csv' into table orders_no_partition1;
Loading data to table hive1.orders_no_partition1
Table hive1.orders_no_partition1 stats: [numFiles=1, totalSize=112]
OK
Time taken: 1.104 seconds

```

```

hive> load data local inpath '/home/hduser/orders_country_w_states.csv' into table orders_no_partition1;
Loading data to table hive1.orders_no_partition1
Table hive1.orders_no_partition1 stats: [numFiles=1, totalSize=112]
OK
Time taken: 1.104 seconds
hive> create table all_orders(id string,customer_id string,product_id string,quantity int,amount double,postalcode string)
> partitioned by(country string,state string)
> row format delimited
> fields terminated by ",";
OK
Time taken: 0.313 seconds
hive> insert into table all_orders partition(country,state)
> select * from orders_no_partition1;

```

```

hive> show partitions all_orders;
OK
country=CA/state=S2
country=WA/state=S1

```

```
[hduser@localhost ~]$ hdfs dfs -ls -R /user/hive/warehouse/hive1.db/all_orders;
22/07/04 20:35:33 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
drwxr-xr-x  - hduser supergroup          0 2022-07-04 20:33 /user/hive/warehouse/hive1.db/all_orders/country=CA
drwxr-xr-x  - hduser supergroup          0 2022-07-04 20:33 /user/hive/warehouse/hive1.db/all_orders/country=CA/state=S2
-rw-r--r--  1 hduser supergroup         44 2022-07-04 20:33 /user/hive/warehouse/hive1.db/all_orders/country=CA/state=S2/000000_0
drwxr-xr-x  - hduser supergroup          0 2022-07-04 20:33 /user/hive/warehouse/hive1.db/all_orders/country=WA
drwxr-xr-x  - hduser supergroup          0 2022-07-04 20:33 /user/hive/warehouse/hive1.db/all_orders/country=WA/state=S1
-rw-r--r--  1 hduser supergroup         44 2022-07-04 20:33 /user/hive/warehouse/hive1.db/all_orders/country=WA/state=S1/000000_0
```

Partitioning with Bucketing:

```
hive> set hive.enforce.bucketing=true;
hive> create table products_partitioned_buckets(
  > id int,name string,cost double) partitioned by (category string)
  > clustered by (id) into 4 buckets
  > row format delimited
  > fields terminated by ",";
OK
Time taken: 0.395 seconds
hive> insert into table products_partitioned_buckets partition(category) select id,name,cost,category from products_no_bucket;
```

```
ser@localhost ~]$ hdfs dfs -ls -R /user/hive/warehouse/hive1.db/products_partitioned_buckets
```

```
hive> show partitions products_partitioned_buckets;
OK
category= __HIVE_DEFAULT_PARTITION__
category=fashion
category=furniture
category=mobile
category=toys
Time taken: 0.438 seconds, Fetched: 5 row(s)
hive> select * from products_partitioned_buckets tablesample(bucket 1 out of 4);
```