

# Partitions in hive

We will use below dataset to perform the operations:

- user\_info1.txt
- user\_info2.txt
- user\_info3.txt

## Understanding Data

user\_info1.txt, user\_info2.txt, user\_info3.txt files has following columns.

first\_name

last\_name

id

```
hive> create table user1(first_name string,last_name string,id string)
> partitioned by (country string, region string)
> row format delimited fields terminated by ','
> stored as textfile;
```

```
hive> describe formatted user1;
```

OK

col_name	data_type	comment
# col_name	data_type	comment
first_name	string	
last_name	string	
id	string	

### # Partition Information

# col_name	data_type	comment
country	string	
region	string	

### # Detailed Table Information

Database: retail  
Owner: cloudera  
CreateTime: Wed Jan 27 04:24:48 PST 2021  
LastAccessTime: UNKNOWN  
Protect Mode: None  
Retention: 0  
Location: hdfs://quickstart.cloudera:8020/user/hive/warehouse/retail.db/user1  
Table Type: MANAGED\_TABLE  
Table Parameters:  
transient\_lastDdlTime 1611750288

#### # Storage Information

SerDe Library: org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe  
InputFormat: org.apache.hadoop.mapred.TextInputFormat  
OutputFormat: org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat  
Compressed: No  
Num Buckets: -1  
Bucket Columns: []  
Sort Columns: []  
Storage Desc Params:

field.delim ,  
serialization.format ,

Time taken: 0.085 seconds, Fetched: 35 row(s)

- **Static Partitioning**
- **Dynamic Partitioning**

### Static Partitioning

Static partitioning needs to be applied when we know data belongs to which partition.

```
hive> load data local inpath 'file:///home/cloudera/khasimbabu/HIVE_Excercise/user_info1.txt' into  
table user1
```

```
> partition(country='New Zealand',region = 'Quebec');
```

```
hive> load data local inpath 'file:///home/cloudera/khasimbabu/HIVE_Excercise/user_info2.txt' into  
table user1
```

```
> partition (country = 'Nigeria',region = 'Pomorskie');
```

So, it will only go to the required partition and fetches the result. No, need to scan whole table

```
hive> select * from user1 where country='New Zealand' and region='Quebec';
```

```
hive> select * from user1 where country='Nigeria' and region='Pomorskie';
```

```
[cloudera@quickstart /]$ hdfs dfs -ls /user/hive/warehouse/retail.db/user1/
```

Found 2 items

```
drwxrwxrwx - cloudera supergroup    0 2021-01-27 07:58  
/user/hive/warehouse/retail.db/user1/country=New Zealand  
drwxrwxrwx - cloudera supergroup    0 2021-01-27 08:08  
/user/hive/warehouse/retail.db/user1/country=Nigeria
```

Writing individual queries for every partition is not possible for large number of partitions. In this case we use Dynamic partition.

## **Dynamic Partitioning**

```
hive> create table user2(first_name string,last_name string,id string)
> partitioned by (country string, region string)
> row format delimited fields terminated by ',' stored as textfile;
```

```
hive> create table user3(first_name string,last_name string,id string,country string,region string)
> row format delimited fields terminated by ',' stored as textfile;
```

```
hive> load data local inpath 'file:///home/cloudera/khasimbabu/HIVE_Excercise/user_info3.txt' into
table user3;
```

```
Hive>set hive.exec.dynamic.partition=true;
Hive>set hive.exec.dynamic.partition.mode=nonstrict;
Hive>set hive.exec.max.dynamic.partitions=100;
Hive>set hive.exec.max.dynamic.partitions.pernode=100;
```

## **Inserting data into table user2 dynamically**

```
hive> insert into table user2 partition(country,region)
> select first_name,last_name,id,country,region from user3;
```

```
[cloudera@quickstart /]$ hdfs dfs -ls /user/hive/warehouse/retail.db/user2/
Found 25 items
```

```
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=%22Korea
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Angola
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Austria
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Belize
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Burundi
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Christmas Island
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Gabon
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Guam
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Kazakhstan
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Kiribati
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Latvia
drwxrwxrwx - cloudera supergroup      0 2021-01-27 08:43
/user/hive/warehouse/retail.db/user2/country=Liberia
```

drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Luxembourg  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Myanmar  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Namibia  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Nicaragua  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Nigeria  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Oman  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Saint Kitts and Nevis  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Suriname  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Svalbard and Jan Mayen Islands  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Tanzania  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Togo  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Venezuela  
drwxrwxrwx - cloudera supergroup 0 2021-01-27 08:43  
/user/hive/warehouse/retail.db/user2/country=Viet Nam