01-07-2022

External Table

```
hive> create external table products(
    > id string,
    > title string,
    > cost float)
    > row format delimited
    > fields terminated by ","
    > stored as textfile
    > location '/user/hduser/hivedata1';
OK
Time taken: 0.291 seconds
hive> select * from products;
OK
1       Phone   30000.0
2       Camera  55000.0
3       Laptop  60000.0
4       Watch   12000.0
Time taken: 0.168 seconds, Fetched: 4 row(s)
hive>
```

Sub queries:

1.From clause: **a)UNION ALL**

```
hive> create table freshproducts(
    > id string,
    > title string
    > ,cost float)
    > row format delimited
    > fields terminated by ","
    > stored as textfile;
OK
Time taken: 0.221 seconds
hive> load data local inpath '/home/hduser/freshProducts.csv' into table freshProducts;
Loading data to table sampledb.freshproducts
Table sampledb.freshproducts stats: [numFiles=1, totalSize=61]
OK
Time taken: 0.812 seconds
```

```
hive> select * from (
    > select * from products
    > UNION ALL
    > select * from freshProducts ) t;
Query ID = hduser_20220701103654_ea65ea68-4eec-4025-987c-b765e5520029
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1656645932148_0003, Tracking URL = http://Inceptez:8088/proxy/application_16566459321
_0003/
Kill Command = /usr/local/hadoop/bin/hadoop job  -kill job_1656645932148_0003
Hadoop job information for Stage-1: number of mappers: 2; number of reducers: 0
2022-07-01 10:37:07,849 Stage-1 map = 0%,   reduce = 0%
2022-07-01 10:37:18,630 Stage-1 map = 100%,   reduce = 0%, Cumulative CPU 1.82 sec
MapReduce Total cumulative CPU time: 1 seconds 820 msec
Ended Job = job_1656645932148_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 2   Cumulative CPU: 1.82 sec   HDFS Read: 10008 HDFS Write: 141 SUCCESS
Total MapReduce CPU Time Spent: 1 seconds 820 msec
OK
1       iPhone  130000.0
2       Dell    90000.0
3       Fossil  13000.0
4       Airpods 14000.0
        NULL    NULL
1       Phone   30000.0
2       Camera  55000.0
3       Laptop  60000.0
4       Watch   12000.0
Time taken: 25.292 seconds, Fetched: 9 row(s)
```

b)JOIN

```
hive> select distinct(t.product_id)
    > from (select product_id from customer
    > join
    > orders where customer.id=orders.customer_id) t;
OK
Camera
Laptop
Phone
Watch
Time taken: 36.718 seconds, Fetched: 4 row(s)
```

2. where clause

a)IN

```
hive> select name from customer
    > where id IN
    > (select customer_id from orders );
Query ID = hduser_20220701110139_e0b02509-a77a-4d3c-81d7-902d47216c96
Total jobs = 1
OK
Alekhya
Deepti
Keerthana
Likhith
Time taken: 36.691 seconds, Fetched: 4 row(s)
```

b)NOT IN

```
hive> select name from customer
    > where id NOT IN
    > (select customer_id from orders );
OK
Time taken: 84.436 seconds
hive>
```

c)EXISTS

```
hive> select id from customer
    > where exists
    > (select customer_id from orders where orders.customer_id=customer.id);
Query ID = hduser_20220701110658_27c34a65-0d1a-4562-87cc-0416c91905f3
Total jobs = 1
OK
1
2
3
4
Time taken: 26.916 seconds, Fetched: 4 row(s)
```

d)NOT EXISTS

```
hive> select id from customer
    > where not exists
    > (select customer_id from orders where orders.customer_id=customer.id);
Query ID = hduser_20220701111001_19e61e80-83e7-47be-9b8a-80a2399446e2
Total jobs = 1
OK
Time taken: 26.965 seconds
hive>
```

VIEWS

```
hive> create view customer_purchase as
    > select customer_id,product_id,address from customer
    > join
    > orders where customer.id=orders.customer_id;
OK
Time taken: 0.163 seconds
hive> show tables;
OK
customer
customer_purchase
freshproducts
orders
products
Time taken: 0.05 seconds, Fetched: 5 row(s)
hive> select * from customer_purchase;
Query ID = hduser_20220701111611_27bc8d12-888f-4240-9221-0c1d302754f8
Total jobs = 1
```

```
OK
1        Phone    Hyderabad
2        Camera   Amaravathi
3        Laptop   Vizag
4        Watch    Nalgonda
1        Phone    Hyderabad
2        Camera   Amaravathi
3        Laptop   Vizag
4        Watch    Nalgonda
Time taken: 27.237 seconds, Fetched: 8 row(s)
```

Explode and lateral views

Create a table and load table from local

```
hive> create table authors_details(
    > author_name string,
    > book_name array<string>)
    > row format delimited
    > fields terminated by ","
    > collection items terminated by ":";
OK
Time taken: 0.282 seconds
```

```
hive> load data local inpath '/home/hduser/authors.csv' into table authors_details;
Loading data to table sampledb.authors_details
Table sampledb.authors_details stats: [numFiles=1, totalSize=184]
OK
Time taken: 0.546 seconds
```

```
hive> select * from authors_details;
OK
Salman Rushide  ["Grimus","Shame","Fury"]
Thomas Otway    ["Don Carlos","The Orphan"]
JK Rowling      ["Harrypotter-The chamber of secrets","half-bloodprince"]
Ben Jonson      ["Volpone","Epicene"]
John Milton     ["Arcades","comus"]
Time taken: 0.091 seconds, Fetched: 5 row(s)
hive>
```

Explode:

```
hive> select explode(book_name) from authors_details;
OK
Grimus
Shame
Fury
Don Carlos
The Orphan
Harrypotter-The chamber of secrets
half-bloodprince
Volpone
Epicene
Arcades
comus
Time taken: 0.069 seconds, Fetched: 11 row(s)
hive>
```

Lateral view

```
hive> select author_name,b_name from authors_details lateral view
    > explode(book_name) book_table as b_name;
OK
Salman Rushide   Grimus
Salman Rushide   Shame
Salman Rushide   Fury
Thomas Otway     Don Carlos
Thomas Otway     The Orphan
JK Rowling       Harrypotter-The chamber of secrets
JK Rowling       half-bloodprince
Ben Jonson       Volpone
Ben Jonson       Epicene
John Milton      Arcades
John Milton      comus
Time taken: 0.09 seconds, Fetched: 11 row(s)
hive>
```

Complex datatypes

1. Array

```
hive> create table mobilephones(
    > id string,
    > title string,
    > cost float,
    > colors array<string>,
    > screen_size array<float>);
OK
Time taken: 0.185 seconds
hive> insert into table mobilephones
    > select "redminote7","Redmi Note 7",300,array("white","silver","black"),array(float(4.5))
    > union all
    > select "motoGplus","Moto G Plus",200,array("black","gold"),array(float(4.5),float(5.5))
    > ;
Query ID = hduser_20220701114222_c7274804-8518-428a-b1b6-0adec4e0fe29
Total jobs = 3
```

```
hive> select id,colors from mobilephones;
OK
redminote7      ["white","silver","black"]
motoGplus       ["black","gold"]
Time taken: 0.067 seconds, Fetched: 2 row(s)
hive> select id,colors[0] from mobilephones;
OK
redminote7      white
motoGplus       black
Time taken: 0.092 seconds, Fetched: 2 row(s)
hive> select id,screen_size[1] from mobilephones;
OK
redminote7      NULL
motoGplus       5.5
Time taken: 0.065 seconds, Fetched: 2 row(s)
hive>
```

2. Map

```
hive> create table mobilephones_new(
    >     id string,
    >     title string,
    >     cost float,
    >     colors array<string>,
    >     screen_size array<float>,
    >     features map<string,boolean>)
    >     row format delimited
    >     fields terminated by ","
    >     collection items terminated by "#"
    >     map keys terminated by ":";
OK
Time taken: 0.223 seconds
hive> load data local inpath '/home/hduser/mobilephones.csv' into table mobilephones_new;
Loading data to table sampledb.mobilephones_new
Table sampledb.mobilephones_new stats: [numFiles=1, totalSize=142]
OK
Time taken: 0.516 seconds
hive> select * from mobilephones_new;
OK
samsungj7       Samsung J7      250.0   ["red","blue","black"]  [5.5]   {"camera":true,"dualsim":false}
oneplus7        One Plus 7      450.0   ["gold","silver"]       [4.5,5.5]       {"autofocus":true,"dualsim":true}
Time taken: 0.077 seconds, Fetched: 2 row(s)
hive> select id,features['camera'] from mobilephones_new;
OK
samsungj7       true
oneplus7        NULL
Time taken: 0.083 seconds, Fetched: 2 row(s)
hive>
```

3.Struct

```
hive> create table mobilephones_new1(
    > id string,
    > title string,
    > cost float,
    > colors array<string>,
    > screen_size array<float>,
    > features map<string,boolean>,
    > information struct<battery:string,camera:string>)
    > row format delimited
    > fields terminated by ","
    > collection items terminated by "#"
    > map keys terminated by ":";
OK
Time taken: 0.244 seconds
hive> load data local inpath '/home/hduser/mobile.csv' into table mobilephones_new1;
Loading data to table sampledb.mobilephones_new1
Table sampledb.mobilephones_new1 stats: [numFiles=1, totalSize=164]
OK
Time taken: 0.466 seconds
hive> select * from mobilephones_new1;
OK
samsungj7       Samsung J7      250.0   ["red","blue","black"]  [5.5]   {"camera":true,"dualsim":false} {"battery":"24 hours","camera":"2MP"}
oneplus7        One Plus 7      450.0   ["gold","silver"]       [4.5,5.5]       {"autofocus":true,"dualsim":true}{"battery":"12 hours","camera":null}
Time taken: 0.062 seconds, Fetched: 2 row(s)
hive>
```

```
hive> select information.battery from mobilephones_new1;
OK
24 hours
12 hours
Time taken: 0.103 seconds, Fetched: 2 row(s)
hive> select information.camera from mobilephones_new1;
OK
2MP
NULL
Time taken: 0.217 seconds, Fetched: 2 row(s)
hive>
```