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/hiveData';
OK
Time taken: 0.259 seconds
hive> select * from products;
OK
1       Phone   20000.0
2       Camera  10000.0
3       Watch   5000.0
4       Laptop  70000.0
Time taken: 0.211 seconds, Fetched: 4 row(s)
```

Sub-Queries:

From Clause:

```
hive> create table freshProducts(
    > id string,
    > title string,
    > cost float)
    > row format delimited
    > fields terminated by ','
    > stored as textfile;
OK
Time taken: 0.295 seconds
hive> load data local inpath '/home/hduser/freshP.csv' into table freshProducts;
Loading data to table demo_db.freshproducts
Table demo_db.freshproducts stats: [numFiles=1, totalSize=57]
OK
Time taken: 0.782 seconds
```

Union:

```
hive> select * from (
    > select * from products
    > union all
    > select * from freshProducts) t;
Query ID = hduser_20220701103749_87444f94-c393-41fe-b74b-87063b885569
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_1656650050438_0002, Tracking URL = http://Inceptez:8088/proxy/application_1656650050438_0002/
Kill Command = /usr/local/hadoop/bin/hadoop job  -kill job_1656650050438_0002
Hadoop job information for Stage-1: number of mappers: 2; number of reducers: 0
2022-07-01 10:38:11,787 Stage-1 map = 0%,   reduce = 0%
2022-07-01 10:38:27,634 Stage-1 map = 100%,   reduce = 0%, Cumulative CPU 2.75 sec
MapReduce Total cumulative CPU time: 2 seconds 750 msec
Ended Job = job_1656650050438_0002
MapReduce Jobs Launched:
Stage-Stage-1: Map: 2   Cumulative CPU: 2.75 sec   HDFS Read: 9850 HDFS Write: 130 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 750 msec
OK
1       Phone   20000.0
2       Camera  10000.0
3       Watch   5000.0
4       Laptop  70000.0
1       Dell    1000000.0
2       mivi    5000.0
3       Realme  15000.0
4       Fossil  10000.0
Time taken: 40.151 seconds, Fetched: 8 row(s)
```

Join:

```
hive> select distinct(t.product_id)
    > from (select product_id from customers
    > join orders
    > where customers.id=orders.customer_id) t;
Query ID = hduser_20220701105230_12af0751-b90b-4644-8b61-5d093132c321
Total jobs = 1
MapReduce Jobs Launched:
Stage-Stage-2: Map: 1  Reduce: 1   Cumulative CPU: 6.22 sec   HDFS Read: 12260 HDFS Write: 73 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 220 msec
OK
Camera
Camera1
Desktop
Desktop1
Laptop
Laptop1
Phone
Phone1
Watch
Watch1
Time taken: 61.572 seconds, Fetched: 10 row(s)
```

2.Where Clause

In:

```
hive> select name from customers
    > where id IN(
    > select customer_id from orders);
Query ID = hduser_20220701110033_f815be3c-c9b4-412a-8d21-e0a256f15a91
Total jobs = 1

Deepti
Alekhya
Likhith
Keerthana
Time taken: 53.999 seconds, Fetched: 4 row(s)
```

Not In:

```
hive> select name from customers
    > where id NOT IN(
    > select customer_id from orders);
```

```
MapReduce Jobs Launched:
Stage-Stage-4: Map: 1  Reduce: 1   Cumulative CPU: 5.78 sec   HDFS Read: 7568 HDFS Write: 114 SUCCESS
Stage-Stage-6: Map: 1   Cumulative CPU: 2.16 sec   HDFS Read: 5008 HDFS Write: 201 SUCCESS
Stage-Stage-5: Map: 1   Cumulative CPU: 2.5 sec   HDFS Read: 6314 HDFS Write: 0 SUCCESS
Total MapReduce CPU Time Spent: 10 seconds 440 msec
OK
Time taken: 146.987 seconds
```

Exists:

```
hive> select id from customers
    > where exists
    > (select customer_id from orders where customers.id=orders.customer_id);
Query ID = hduser_20220701110731_7826c8b2-2711-4db3-8a16-ab5bcb96a631
Total jobs = 1
```

```
OK
1
2
3
4
Time taken: 49.611 seconds, Fetched: 4 row(s)
```

Not exists:

```
hive> select id from customers
    > where not exists
    > (select customer_id from orders where customers.id=orders.customer_id);
Query ID = hduser_20220701110911_fe60fc1f-531c-4b17-bb40-890485250152
Total jobs = 1
OK
Time taken: 46.676 seconds
```

Views:

```
hive> create view customer_purchase as
    > select customer_id,product_id,address
    > from customers join orders
    > where customers.id=orders.customer_id;
OK
Time taken: 0.285 seconds
hive> show tables;
OK
customer_purchase
customers
freshproducts
orders
products
Time taken: 0.088 seconds, Fetched: 5 row(s)
hive> select * from customer_purchase;
Query ID = hduser_20220701111613_696d3413-53e8-47b1-81c9-1e91d833f47f
Total jobs = 1
OK
1       Phone   Hyderabad
3       Camera  Pune
4       Laptop  Vizag
2       Watch   Nizamabad
1       Desktop Hyderabad
1       Phone1  Hyderabad
3       Camera1 Pune
4       Laptop1 Vizag
2       Watch1  Nizamabad
1       Desktop1        Hyderabad
Time taken: 41.391 seconds, Fetched: 10 row(s)
```

Explode and Lateral Views

```
hive> create table author_details(
    > author_name string,
    > book_name array<string>)
    > row format delimited
    > fields terminated by ','
    > collection items terminated by ':';
OK
Time taken: 0.229 seconds
hive> load data local inpath '/home/hduser/authors.csv' into table author_details;
Loading data to table demo_db.author_details
Table demo_db.author_details stats: [numFiles=1, totalSize=154]
OK
Time taken: 0.642 seconds
hive> select * from author_details;
OK
JK Rowling      ["harrypotter","half-blood-prince","deathly-hallows"]
Amish Tripati   ["Ram"]
Ruskin Bond     ["Flowers","Fruits","Heaven"]
Dan Brown       ["DaVinchi-Code"]
Time taken: 0.11 seconds, Fetched: 4 row(s)
```

Explode:

```
hive> select explode(book_name) from author_details;
OK
harrypotter
half-blood-prince
deathly-hallows
Ram
Flowers
Fruits
Heaven
DaVinchi-Code
Time taken: 0.121 seconds, Fetched: 8 row(s)
```

Lateral View:

```
hive> select author_name,b_name
    > from author_details
    > lateral view
    > explode(book_name) book_table as b_name;
OK
JK Rowling      harrypotter
JK Rowling      half-blood-prince
JK Rowling      deathly-hallows
Amish Tripati   Ram
Ruskin Bond     Flowers
Ruskin Bond     Fruits
Ruskin Bond     Heaven
Dan Brown       DaVinchi-Code
Time taken: 0.213 seconds, Fetched: 8 row(s)
```

Complex datatypes:

1.Arrays

```
hive> create table mobilePhones(
    > id string,
    > title string,
    > coat float,
    > colors array<string>,
    > screensize array<float>)
    > ;
OK
Time taken: 0.274 seconds
hive> insert into mobilePhones
    > select "redmimote7","Redmi note 7",30000,array("red","blue","silver"),array(float(4.5))
    > union all
    > select "sam7","Samsung note 7",80000,array("black","blue","green"),array(float(5.5));
Query ID = hduser_20220701114238_1abbbfaf-d0b8-40fd-b2f0-92a24686316a
Total jobs = 3
hive> select colors from mobilePhones;
OK
["red","blue","silver"]
["black","blue","green"]
Time taken: 0.103 seconds, Fetched: 2 row(s)
hive> select colors[0] from mobilePhones;
OK
red
black
Time taken: 0.13 seconds, Fetched: 2 row(s)
```

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: 0.856 seconds
hive> load data local inpath '/home/hduser/mobile.csv' into table mobilephones_new;
Loading data to table default.mobilephones_new
Table default.mobilephones_new stats: [numFiles=1, totalSize=153]
OK
Time taken: 0.669 seconds
hive> select * from mobilephones_new;
OK
samsungj7       Samsung J7     250.0   ["red","blue","black"]  [5.5]   {"camera":true,"dualsim-false":null}
onplus7 One Plus 7      450.0   ["gold","silver"]       [4.5,5.5]       {"camera":true,"dualsim":true,"autofocus":true}
Time taken: 0.607 seconds, Fetched: 2 row(s)
hive> select id,features["camera"] from mobilephones_new;
OK
samsungj7       true
onplus7 true
Time taken: 0.153 seconds, Fetched: 2 row(s)
hive>
```

3.Struct

```
hive>
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: 0.302 seconds
hive> load data local inpath '/home/hduser/mobile_new1.csv' into table mobilephones_new1;
Loading data to table default.mobilephones_new1
Table default.mobilephones_new1 stats: [numFiles=1, totalSize=175]
OK
Time taken: 0.498 seconds
hive> select * from mobilephones_new1;
OK
samsungj7       Samsung J7     250.0   ["red","blue","black"]  [5.5]   {"camera":true,"dualsim-false":null}    {"battery":"24 hours","camera":"2MP"}
onplus7 One Plus 7      450.0   ["gold","silver"]       [4.5,5.5]       {"camera":true,"dualsim":true,"autofocus":true} {"battery":"12 hours","camera":null}
Time taken: 0.078 seconds, Fetched: 2 row(s)
hive>
```

```
hive> select information.battery from mobilePhones_new1;
OK
24 hours
12 hours
Time taken: 0.294 seconds, Fetched: 2 row(s)
hive>
```