

Practical-19

1. Table Creation Branch A

create table BranchA (Name string,Branch string)row format delimited fields terminated by ',' lines terminated by '\n';

load data local inpath 'Bev_BranchA.txt' overwrite into table BranchA;

```
hive> load data local inpath '/home/cloudera/Coffee Shop Analysis/Bev_BranchA.txt'
overwrite into table BranchA;
Loading data to table default.brancha
Table default.brancha stats: [numFiles=1, numRows=0, totalSize=2168, rawDataSize=0]
OK
Time taken: 1.267 seconds
hive> describe BranchA;
OK
name                string
branch              string
Time taken: 0.513 seconds, Fetched: 2 row(s)
```

2. Table Creation Branch B

create table BranchB (Name string,Branch string)row format delimited fields terminated by ',' lines terminated by '\n';

load data local inpath 'Bev_BranchB.txt' overwrite into table BranchB;

```
hive> load data local inpath '/home/cloudera/Coffee Shop Analysis/Bev_BranchB.txt'
overwrite into table BranchB;
Loading data to table default.branchb
Table default.branchb stats: [numFiles=1, numRows=0, totalSize=4335, rawDataSize=0]
OK
Time taken: 0.455 seconds
```

3. Table Creation Branch C

create table BranchC (Name string,Branch string)row format delimited fields terminated by ',' lines terminated by '\n';

load data local inpath 'Bev_BranchC.txt' overwrite into table BranchC;

```
hive> create table BranchC (Name string,Branch string)row format delimited
> fields terminated by ',' lines terminated by '\n';
OK
Time taken: 0.101 seconds
hive> load data local inpath '/home/cloudera/Coffee Shop Analysis/Bev_BranchC.txt'
overwrite into table BranchC;
Loading data to table default.branchc
Table default.branchc stats: [numFiles=1, numRows=0, totalSize=6494, rawDataSize=0]
OK
Time taken: 0.706 seconds
```

4. Table Creation Branch_Con_Count A

create table Branch_ConA (Name string,Con_count int)row format delimited fields terminated by ',' lines terminated by '\n';

load data local inpath 'Bev_ConcountA.txt' overwrite into table Branch_ConA;

```
hive> create table Branch_ConA (Name string,Con_count int)row format delimited
> fields terminated by ',' lines terminated by '\n';
OK
Time taken: 0.08 seconds
hive> load data local inpath '/home/cloudera/Coffee Shop Analysis/Bev_ConcountA.txt'
overwrite into table Branch_ConA;
Loading data to table default.branch_cona
Table default.branch_cona stats: [numFiles=1, numRows=0, totalSize=17592, rawDataSize=0]
OK
```

5. Table Creation Branch_Con_Count B

create table Branch_ConB (Name string,Con_count int)row format delimited fields terminated by ','
lines terminated by '\n';

load data local inpath 'Bev_ConcountB.txt' overwrite into table Branch_ConB;

```
hive> create table Branch_ConB (Name string,Con_count int)row format delimited  
> fields terminated by ',' lines terminated by '\n';  
OK  
Time taken: 0.153 seconds  
hive> load data local inpath '/home/cloudera/Coffee Shop Analysis/Bev_ConcountB  
.txt' overwrite into table Branch_ConB;  
Loading data to table default.branch_conb  
Table default.branch_conb stats: [numFiles=1, numRows=0, totalSize=35191, rawDataSize=0]  
OK  
Time taken: 0.593 seconds
```

6. Table Creation Branch_Con_Count C

create table Branch_ConC (Name string,Con_count int)row format delimited fields terminated by ','
lines terminated by '\n';

load data local inpath 'Bev_ConcountC.txt' overwrite into table Branch_ConC;

```
hive> load data local inpath '/home/cloudera/Coffee Shop Analysis/Bev_ConcountC  
.txt' overwrite into table Branch_ConC;  
Loading data to table default.branch_conc  
Table default.branch_conc stats: [numFiles=1, numRows=0, totalSize=52827, rawDataSize=0]  
OK  
Time taken: 0.538 seconds
```

7. What is the total number of consumers for Branch1?

select sum(con_count) from Branch_ConA;

```
hive> select sum(con_count) from Branch_ConA;  
Query ID = cloudera_20210420013838_7e907ba5-801c-4501-baf1-d658dd7903be  
Total jobs = 1  
Launching Job 1 out of 1  
Number of reduce tasks determined at compile time: 1  
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>
```

```

MapReduce Total cumulative CPU time: 4 seconds 720 msec
Ended Job = job_1618905448492_0002
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.72 sec HDFS Read: 25012 H
DFS Write: 7 SUCCESS
Total MapReduce CPU Time Spent: 4 seconds 720 msec
OK
535837
Time taken: 55.669 seconds, Fetched: 1 row(s)
hive> █

```

8. What is the number of consumers for the Branch2?

```
select sum(con_count) from Branch_ConB;
```

```

hive> select sum(con_count) from Branch_ConB;
Query ID = cloudera_20210420014444_0faf1958-970c-4717-ae0-186b1cf24590
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
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>

```

```

MapReduce Total cumulative CPU time: 5 seconds 270 msec
Ended Job = job_1618905448492_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 5.27 sec
DFS Write: 8 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 270 msec
OK
1019262
Time taken: 50.905 seconds, Fetched: 1 row(s)
hive> █

```

9. What is the most consumed beverage on Branch1?

select name,sum(con_count) tot_count from Branch_ConA group by name order by tot_count desc limit 1;

```
hive> select name,sum(con_count) tot_count from Branch_ConA group by name order by tot_count desc limit 1;
Query ID = cloudera_20210420014747_ee167403-d69a-4175-af45-cb16ba291d17
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
```

```
DFS Write: 1884 SUCCESS
Stage-Stage-2: Map: 1 Reduce: 1 Cumulative CPU: 4.35 sec HDFS
FS Write: 25 SUCCESS
Total MapReduce CPU Time Spent: 9 seconds 110 msec
OK
Special_cappuccino      19704
Time taken: 91.599 seconds, Fetched: 1 row(s)
hive>
```

10. What are the beverages available on Branch10, Branch8, and Branch1?

Command:

Create table newbranch as select name,branch from (
select * from brancha where branch = 'Branch10' or branch = 'Branch8' or branch = 'Branch1' union
all
select * from branchb where branch = 'Branch10' or branch = 'Branch8' or branch = 'Branch1'
union all
select * from branchc where branch = 'Branch10' or branch = 'Branch8' or branch = 'Branch1'
)UnionResult;

```
hive> Create table newbranch as select name,branch from (
  > select * from brancha where branch = 'Branch10' or branch = 'Branch8' or branch = 'Branch1' union all
  > select * from branchb where branch = 'Branch10' or branch = 'Branch8' or branch = 'Branch1' union all
  > select * from branchc where branch = 'Branch10' or branch = 'Branch8' or branch = 'Branch1' )UnionResult ;
Query ID = cloudera_20210420015858_7f30400d-4911-4224-946f-c85ea223422b
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
```

select distinct(name),branch from newbranch order by name;

```
hive> select distinct(name),branch from newbranch
> order by name;
Query ID = cloudera_20210420020101_2ba8abf9-ee04-4b6e-912a-75e8299858ca
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
```

```
OK
Cold_Coffee      Branch8
Cold_LATTE       Branch1
Cold_LATTE       Branch8
Cold_Lite        Branch1
Cold_Lite        Branch8
Cold_cappuccino  Branch8
Double_Coffee    Branch8
Double_Espresso  Branch1
Double_Espresso  Branch8
Double_LATTE     Branch8
Double_MOCHA     Branch1
```