Assignment 7.3

1. Hive Data Definitions

Hive DDL statements are here, including:

- CREATE DATABASE/SCHEMA, TABLE, VIEW, FUNCTION, INDEX
- DROP DATABASE/SCHEMA, TABLE, VIEW, INDEX
- TRUNCATE TABLE
- ALTER DATABASE/SCHEMA, TABLE, VIEW
- MSCK REPAIR TABLE (or ALTER TABLE RECOVER PARTITIONS)
- SHOW DATABASES/SCHEMAS, TABLES, TBLPROPERTIES, VIEWS, PARTITIONS, FUNCTIONS, INDEX[ES], COLUMNS, CREATE TABLE
- DESCRIBE DATABASE/SCHEMA, table_name, view_name.

Example:

- hive> CREATE SCHEMA userdb;
- hive> SHOW DATABASES;
- hive> DROP DATABASE IF EXISTS userdb;
- hive> CREATE TABLE emp(emp_id int,emp_name String);

2. Hive Data Manipulations

Hive allows for the manipulation of data in HDFS using a variant of SQL.DML refers to "Data Manipulation Language", a subset of SQL statements that modify the data stored in tables.

To understand the Hive DML commands, let's see the demo table.

```
hive> select * from demo;
OK
Amit     Big Data        1       BBSR
Venkat   Web Technology  2       BBSR
Aditya   DBA     1       BNG
Ravinder         Java    2       BBSR
Sunil    C#      1       BBSR
Anil     ASP     2       BNG
Mihir    Big Data        3       BBSR
Mohit    Java    1       BBSR
Time taken: 0.095 seconds, Fetched: 8 row(s)
hive> █
```

Load data local inpath '/home/acadgild/emp_details.txt' overwrite into table demo ;

select skill,count(emp_name) from demo group by skill;

```
ASP      1
Big Data         2
C#       1
DBA      1
Java     2
Web Technology  1
Time taken: 27.589 seconds, Fetched: 6 row(s)
hive> █
```

3. HiveQL Manipulations

HiveQL manipulations are same as data manipulatons instead of writng queries in hive shell, hiveql is file which contains executable queries which can be directly executed in the linux shell using

hive –f /pathOfTheHiveqlFile