

Assignment 7.1

Created the file **emp_details.txt** and loaded into **/user/acadgild**

```
[acadgild@chemlabtest ~]$ hadoop fs -ls /user/acadgild
17/11/08 19:29:18 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform
Found 13 items
drwxr-xr-x - acadgild supergroup          0 2017-10-28 16:55 /user/acadgild/Assignment3.1
drwxr-xr-x - acadgild supergroup          0 2015-11-20 11:46 /user/acadgild/Pictures
drwxr-xr-x - acadgild supergroup          0 2017-10-21 14:02 /user/acadgild/_sqoop
-rw-r--r-- 1 acadgild supergroup        159 2017-11-08 19:29 /user/acadgild/emp_details.txt
drwxr-xr-x - acadgild supergroup          0 2017-10-08 14:38 /user/acadgild/hadoop
drwxr-xr-x - acadgild supergroup          0 2017-11-08 12:43 /user/acadgild/hive
drwxr-xr-x - acadgild supergroup          0 2017-11-07 17:16 /user/acadgild/load
drwxr-xr-x - acadgild supergroup          0 2017-10-07 14:35 /user/acadgild/maxout
drwxr-xr-x - acadgild supergroup          0 2015-11-17 02:03 /user/acadgild/oozie-acad
drwxr-xr-x - acadgild supergroup          0 2015-11-17 02:00 /user/acadgild/share
-rw-r--r-- 1 acadgild supergroup    21007 2017-10-07 12:55 /user/acadgild/sumona-hadoop
drwxr-xr-x - acadgild supergroup          0 2017-11-06 14:14 /user/acadgild/test
drwxr-xr-x - acadgild supergroup          0 2017-11-07 13:22 /user/acadgild/try
[acadgild@chemlabtest ~]$
```

Created the table **emp_details** and loaded the **emp_details.txt** file.

```
hive> create table emp_details
> (
> Name STRING,
> Skill STRING,
> Rank INT,
> Location STRING
> )
> ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';
OK
Time taken: 0.52 seconds
hive> LOAD DATA INPATH '/user/acadgild/emp_details.txt' OVERWRITE INTO TABLE emp_details;
Loading data to table sumona.emp_details
Table sumona.emp_details stats: [numFiles=1, numRows=0, totalSize=159, rawDataSize=0]
OK
```

Command to calculate the number of employee's for each skill:

```
Time taken: 0.155 seconds, Fetched: 8 row(s)
hive> SELECT Skill, COUNT(NAME) FROM emp_details GROUP BY Skill;
Query ID = acadgild_20171108193434_aaa4fc3e5_131a_4c6d_bd4a_07c0f6fc4621
```

Output:

```
Total MapReduce CPU Time  
OK  
ASP      1  
Big Data      2  
C#       1  
DBA      1  
Java     2  
Web Technology 1  
Time taken: 81.685 seconds  
hive> █
```