**Problem Statement**

==================================================================================

Calculate the number of employees corresponding to each skill from the table 'employee' which is

loaded in the Demo.

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

===================================================================

**Hive Input Commands**

hive> CREATE DATABASE IF NOT EXISTS skillsAssessment ;

hive> set hive.cli.print.current.db = true;

hive (default)> SHOW DATABASES LIKE 's%' ;

hive (default)> USE skillsAssessment ;

hive (skillsAssessment)> CREATE TABLE IF NOT EXISTS employee

> (

> empname string,

> skill string,

> experience int,

> location string

> )

> row format delimited fields terminated by ','

> ;

hive (skillsAssessment)> SHOW TABLES ;

hive (skillsAssessment)> DESCRIBE FORMATTED employee;

hive (skillsAssessment)> LOAD DATA

> LOCAL INPATH "/home/cloudera/chhaya/hive/employee.txt"

> INTO TABLE employee

> ;

hive (skillsAssessment)> select skill , COUNT(skill) from employee group by skill;

hive (skillsAssessment)> select skill , count(skill) from employee group by skill having skill is not null ;

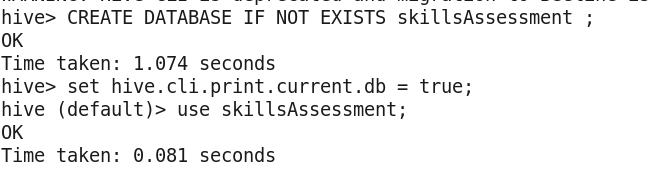
hive (skillsAssessment)> quit;

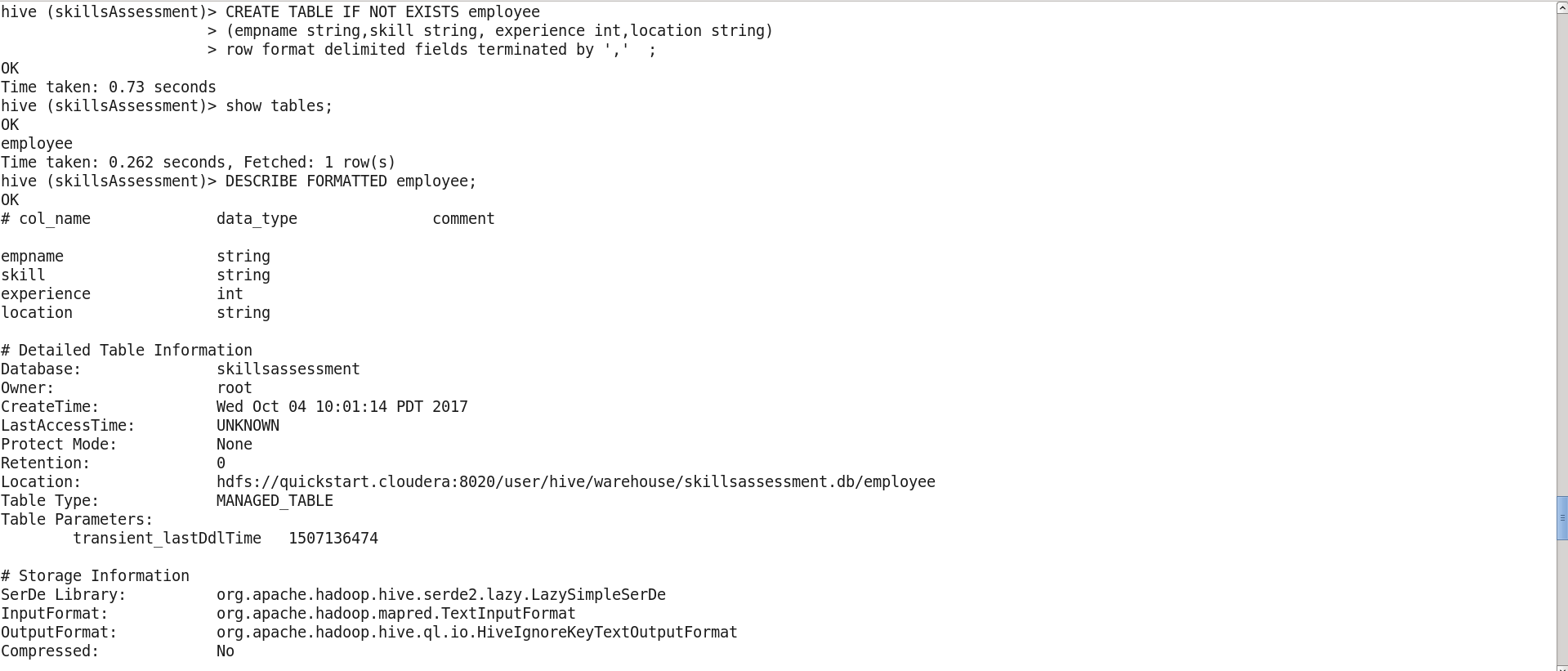
[cloudera@quickstart ~]$ hadoop fs -ls /user/hive/warehouse/

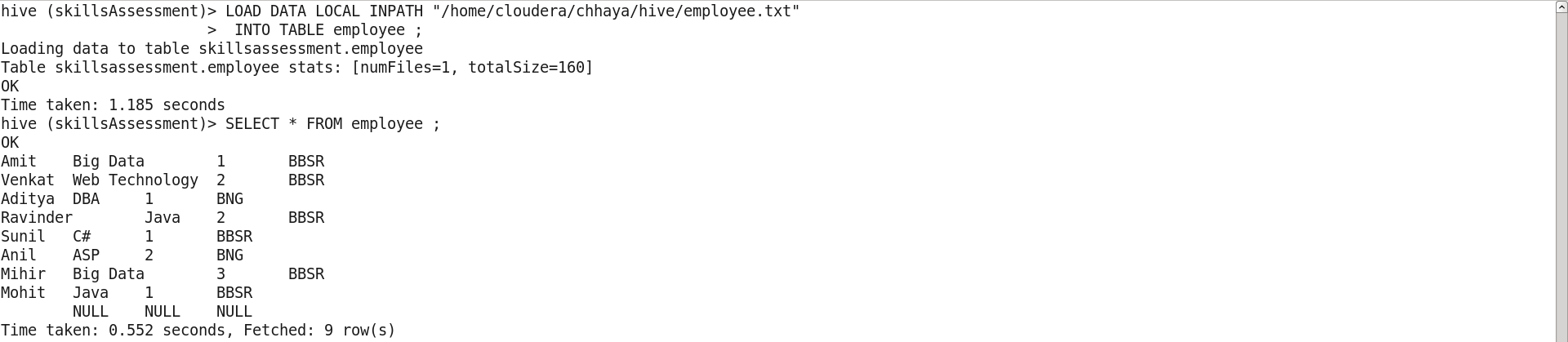
[cloudera@quickstart ~]$ hadoop fs -ls /user/hive/warehouse/skillsAssessment.db ;

[cloudera@quickstart ~]$ hadoop fs -ls /user/hive/warehouse/skillsAssessment.db/employee ;

**Output :**

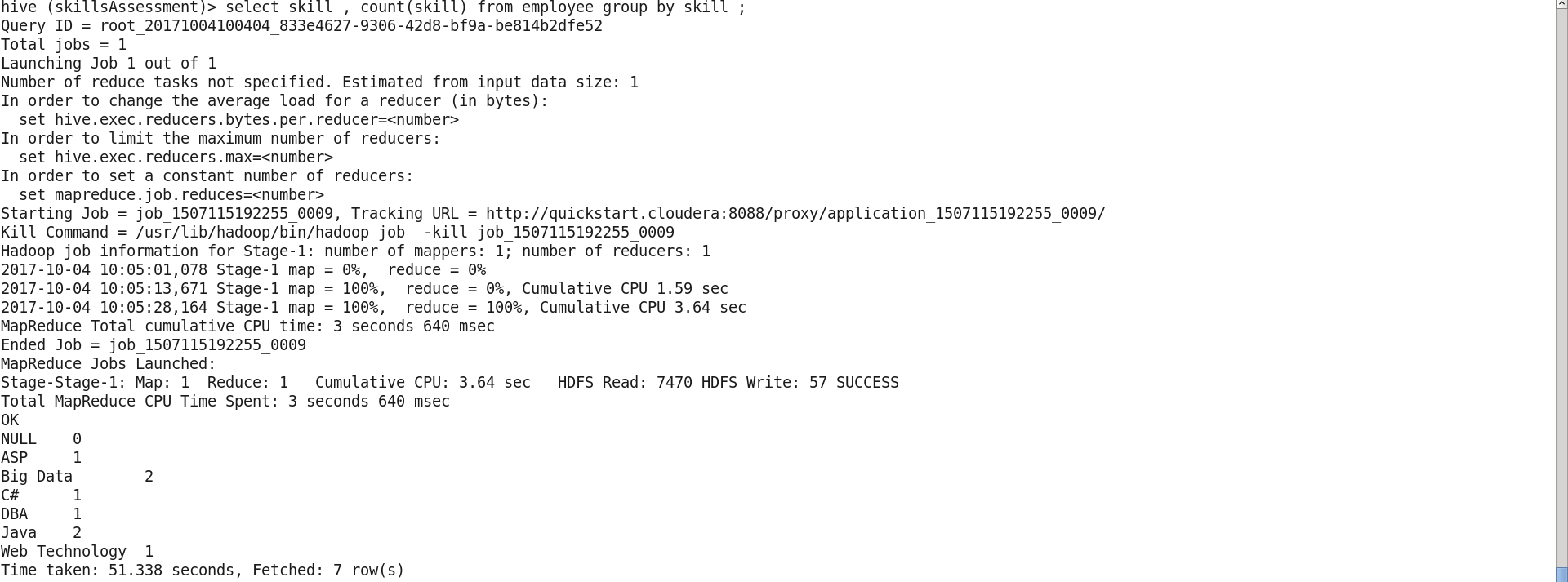






**Final Output:**

**Calculating the total nr of employees in each skill without filtering null values:**



Calculating total number of employees after filtering out null values:

