Session 7

Assignment 1

|  |  |
| --- | --- |
| **Prepared For:** | AcadGild |
|  |  |
| **Document Approval:** | **AcadGild** |
|  |  |
|  |  |
|  |  |
|  |  |
| **Project Title:** | Session 7 – Assignment 1 |
|  |  |
| **Prepared By:** | Duncan Burgess |
|  |  |
|  | dburgess@duncb.com |
|  |  |
| **Primary Engineer:** | Duncan Burgess |
|  |  |
| **Document Reference:** | **Session 7 – Assignment 1** |
|  |  |
| **Start Date:** | 21/09/2017 |
|  |  |
|  |  |

# 

# Contents

[Contents 2](#_Toc493766785)

[Change History 3](#_Toc493766786)

[1. Problem Statement 4](#_Toc493766787)

[2. Solution 5](#_Toc493766788)

[2.1. Create the database 5](#_Toc493766789)

[2.2. Create the tables required 5](#_Toc493766790)

[2.3. Import Data 5](#_Toc493766791)

[2.4. Get required result no employees with each skills 6](#_Toc493766792)

# Change History

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Document Revision** | **Date** | **Authored By** | **Authorised By** | **Sections Affected** | **Reason for Change** |
| Rev 01 | 21/09/2017 | Duncan Burgess |  | All | Initial release. |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

# Problem Statement

Calculate the number of employees corresponding to each skill from the table 'employee' which is loaded into Demo.

**Emp\_details.txt'**

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

# Solution

To display the database working in I set

***Set hive.cli.print.current.db = true***

## Create the database

Create the database Demo as required.

*create database Demo;*

**Use the database created**

*use Demo;*

## Create the tables required

*hive (Demo)> create table if not exists employee*

*> (*

*> emp\_name string,*

*> skill string,*

*> exp int,*

*> location string*

*> )*

*> row format delimited*

*> fields terminated by ',';*

*OK*

*Time taken: 0.056 seconds*

**Check tables**

*hive (Demo)> describe employee;*

*OK*

*emp\_name                  string*

*skill            string*

*exp            int*

*location string*

*Time taken: 0.054 seconds, Fetched: 4 row(s)*

## Import Data

Data is loaded form the file emp\_details.txt.

*load data local inpath '/home/acadgild/Datasets/emp\_details.txt'*

*into table employee;*

**Check results**

*hive (Demo)> select \* from employee;*

*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: 0.064 seconds, Fetched: 8 row(s)*

## Get required result no employees with each skills

hive> (Demo)> select skill, count(\*) from employee group by skill;

*OK*

*ASP 1*

*Big Data 2*

*C# 1*

*DBA 1*

*Java 2*

*Web Technology 1*

*Time taken: 20.549 seconds, Fetched: 6 row(s)*