**Session 6: HIVE Introduction**

Assignment 6.1

Student Name: Abarajithan SA

Course: Big Data Hadoop & Spark Training

Start Date:  2017-09-09

End Date:  2017-11-26

**Assignment 6.1**– Create a database named **'custom'**.

* Create a table named **temperature\_data** inside custom having below fields:

1. date (mm-dd-yyyy) format

2. zip code

3. temperature

* The table will be loaded from comma-delimited file.
* Load the dataset.txt (which is ',' delimited) in the table.

Contents

[Introduction: 2](#_Toc495337200)

[Associated Data files: 2](#_Toc495337201)

[Problem Statement: 2](#_Toc495337202)

[HIVE Queries: 2](#_Toc495337203)

[Expected Output: 3](#_Toc495337204)

# Introduction:

In this assignment, I’m going to create a table called **‘custom’** with a table name called **‘temperature\_data’** which contains the columns date, Zip code and the temperature. The given dataset file will be loaded in to the database ‘custom’ into the table temperature\_data.

The given dataset file is placed at **home/acadgild/hadoop.**

# Associated Data files:

<https://drive.google.com/file/d/0Bxr27gVaXO5sa0JBamZXdkpYUFk/view?usp=sharing>

# Problem Statement:

Create a database named **'custom'**.

* Create a table named **temperature\_data** inside custom having below fields:

1. date (mm-dd-yyyy) format

2. zip code

3. temperature

* The table will be loaded from comma-delimited file.
* Load the dataset.txt (which is ',' delimited) in the table.

# HIVE Queries:

*CREATE DATABASE custom;*

*CREATE TABLE temperature\_data*

*(*

*full\_date STRING,*

*zip INT,*

*temperature INT*

*)*

*ROW FORMAT DELIMITED*

*FIELDS TERMINATED BY ',';*

*LOAD DATA LOCAL INPATH '/home/acadgild/hadoop/dataset\_Session 14.txt'*

*INTO TABLE custom.temperature\_data;*

# Expected Output:



