**Advanced Hive**

**Assignment 2**

**Problem Statement**

Write a hive UDF that implements functionality of string concat\_ws(string SEP,

array<string>).

This UDF will accept two arguments, one string and one array of string.

It will return a single string where all the elements of the array are separated by the

SEP.

**Step1**-**Creating the table and loading the data**

The command used to create table is as below.

**CREATE TABLE assignment2(**

**skill array<string>**

**)ROW FORMAT DELIMITED FIELDS TERMINATED by '\t' collection items TERMINATED by ',';**

****

**Step 2-Load sample data into table**

**LOAD DATA LOCAL INPATH '/home/acadgild/hive/assignment2\_input.txt'**

**INTO TABLE assignment2**

****

****

**Step 3-Write UDF in Java Eclipse**

****

**Step4-Add the jar to Hive shell and use the jar**

Add the jar using the below command

ADD JAR /home/acadgild/hive/hive\_udf.jar;

**Step5-CREATE TEMPORARY FUNCTION**

CREATE TEMPORARY FUNCTION concat AS 'udf.String\_Concat\_Ws';

****

**Step6-Call the concat function by passing the arguments**

**Query:**

SELECT skill,concat("\*",skill) from assignment2;

**Output :**

["hadoop","test","hive"] hadoop\*test\*hive\*

["This","is","an","apple"] This\*is\*an\*apple\*

["Good","Morning"] Good\*Morning\*

["Time","is","Gold"] Time\*is\*Gold\*

****