Session 08: 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.

**Solution--**

**JAVA code for UDF:-**

**package** com.hive.UDF.example;

**import** java.util.ArrayList;

**import** org.apache.hadoop.hive.ql.exec.UDF;

/\*\*

\*

\* **@author** Nihit Rai

\* HIVE UDF for Assignment 8.2

\* 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

\*

\*/

**public** **class** concat\_ws **extends** UDF{

**public** String evaluate(String SEP, ArrayList<String> strArray)

{

//using string builder

StringBuilder sb = **new** StringBuilder();

//enhanced for loop for iterating over string array and appending array element and SEP to StringBuilder

**for** (String str : strArray)

{

sb.append(str);

sb.append(SEP);

}

//Removing extra SEP from end of string builder

sb.setLength(sb.length() - SEP.length());

//Converting string builder to String and returning it

**return** sb.toString();

}

}

**Project file-**



**JAR file-**

****

**Running the UDF:-**

**Creating table and inserting data:-**

**Commands :-**

Hive> CREATE TABLE Products

(

id INT,

ProductName STRING,

ProductColorOptions ARRAY<STRING>

);

Hive> INSERT INTO TABLE Products

SELECT 1, 'Widgets', array('Red', 'Blue', 'Green')

UNION ALL

SELECT 2, 'Cogs', array('Blue', 'Green', 'Yellow');

****

Now running the UDF

**Commands :-**

hive> ADD JAR Assignment8\_2.jar;

hive> CREATE TEMPORARY FUNCTION concat\_ws AS 'com.hive.UDF.example.concat\_ws';

hive> select id, ProductName, concat\_ws("|",ProductColorOptions) from Products;

