**Session 3**

**Assignment 3.5**

Student Name: Karthik K

Course: Big Data Hadoop & Spark Training

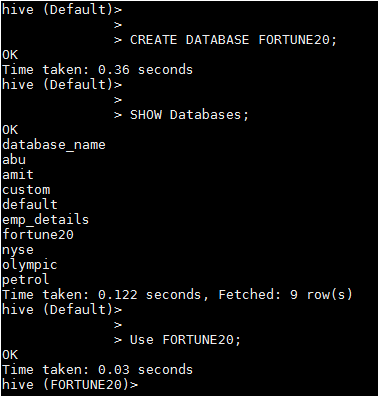
## Create Database and Table

**Create Database FORTUNE20**

### HIVE QL

***CREATE DATABASE FORTUNE20***

***Use FORTUNE20;***



**Create Table Fortune\_company**

### HIVE QL

***CREATE TABLE fortune\_company(rank int, company\_name string,website string, protocal string)***

***ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t';***

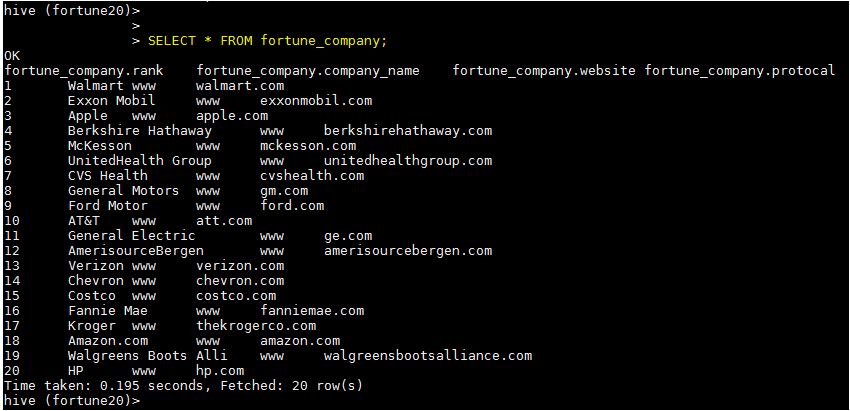
***LOAD DATA LOCAL INPATH '/home/acadgild/hadoop/fortune20.txt’***

***INTO TABLE fortune20.fortune\_company;***



Viewing the data in the table fortune\_company,

***SELECT \* FROM fortune\_company;***



# HIVE UDF Java code

**package** concatws;

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

import org.apache.hadoop.hive.ql.exec.Description;

@Description(name = "concatws", value = "\_FUNC\_(string SEP, array<string>) - RETURN\_TYPE(STRING)\n" + "Description: Concatenate two strings, separated by the seperator",

extended = "Example:\n"

+ " > SELECT CONCAT\_WS (website,’.’,protocal) FROM src;\n"

+ "www.walmart.com")

**public** **class** concatws **extends** UDF

{

**public** String evaluate(String param1, String[] param2)

{

String Output = "";

**if**(param1==**null** && param2==**null**)

{

**return** **null**;

}

**for**(**int** i = 0; i < param2.length; i++)

{

Output+= param2[i];

}

**return**(param1.concat(Output));

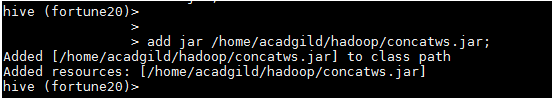
}

}

After that we are adding JAR created from the JAVA class which is defining the UDF using below syntax-

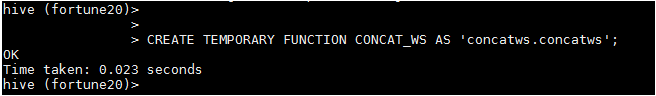
# HIVE UDF CONCAT\_WS function

***add jar /home/acadgild/hadoop/concatws.jar;***



After that we are creating a temporary function “CONCAT\_WS” using below syntax-

***CREATE TEMPORARY FUNCTION CONCAT\_WS AS 'concatws.concatws';***



After that we run below query to take one column (company\_name) input as String and another array(website,’.’,protocal) as Array of Strings and concatenate them,

## HIVE QL

***SELECT rank, company\_name, CONCAT\_WS(website,'.',protocal) from fortune\_company;*** ***SELECT rank, company\_name, CONCAT\_WS(website,'.',protocal) from fortune\_company;*** 

## Required Output

