Session 10: Oozie and Sqoop

Assignment 2 Question

**Table of Contents**

1. Introduction

2. Problem Statement

3. Output

**1. Introduction**

This assignment will help you to consolidate the concepts learnt in the session.

**2. Problem Statement**

Implement the concept given in below blog link and share the complete steps along with

screenshots.

https://acadgild.com/blog/loading-data-into-hbase-using-pig-scripts/

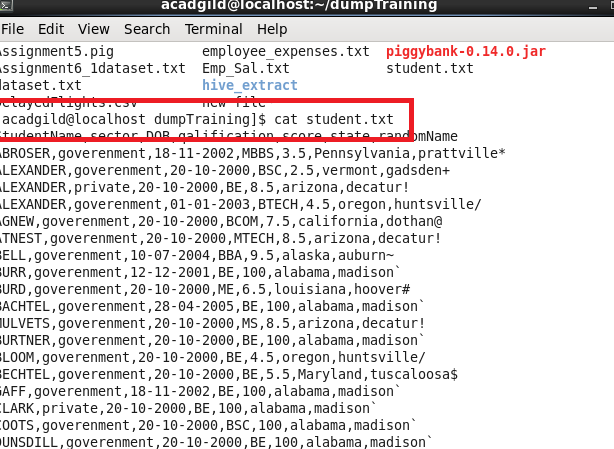
NA

**3. Output**

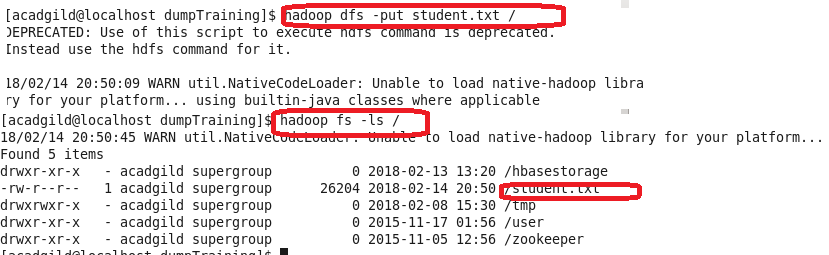
﻿

he description for the above data set containing  seven columns named as:

StudentName, sector, DOB, qualification, score, state, randomName.



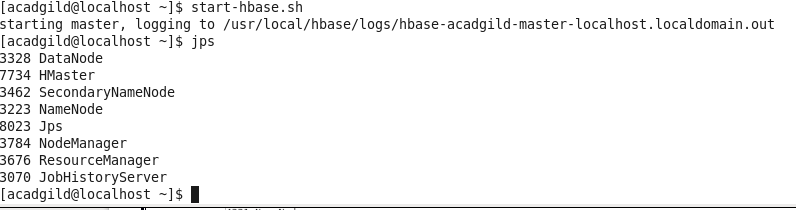
﻿We will be copying the data set in to HDFS which will be further loaded into HBase.

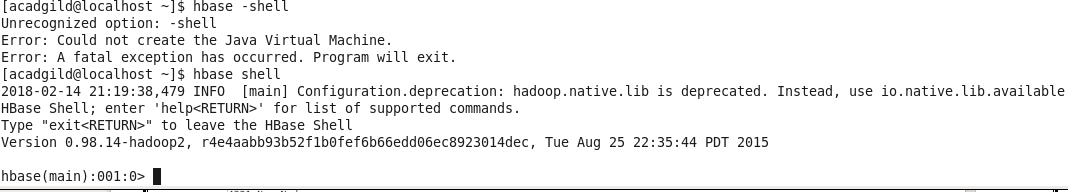


We will be including few jar files of HBase to the Pig classpath.

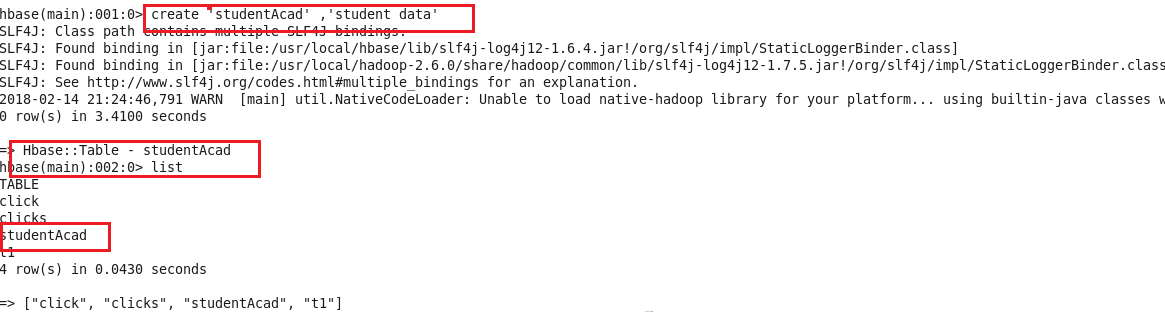


We will now start HBase shell and create a table.





We only need this table as skeleton so PIG can Store data inside this by referring the table name.



We can come out from HBase by typing exit and switch to PIG grunt shell.

Once we are inside PIG mode we can load data from HDFS to Alias relation.

