# Module-5: Pig and Pig Latin

**Assignment Solution** 

## edureka!



© 2014 Brain4ce Education Solutions Pvt. Ltd.

### Module-5: Pig and Pig Latin

Assignment Solution - Find out successful students in a class

### **Table of Contents**

Introduction	2
4. Duchlans salution	_
1. Problem solution	4
1.1 Copy datasets to HDFS	2
1.2 Execute Pig commands	



#### Introduction

To solve this, you need to load the datasets into HDFS and join the two datasets using Pig commands to find out the successful students in the class.

#### 1. Problem solution

#### 1.1 Copy datasets to HDFS

Use 'copyFromLocal' Hadoop command to load the datasets into HDFS.

```
$hadoop fs -put <your_data_staging_dir>/student hdfs:/
$hadoop fs -put <your_data_staging_dir>/results hdfs:/
```

#### FIGURE 1-1 DATA LOADING

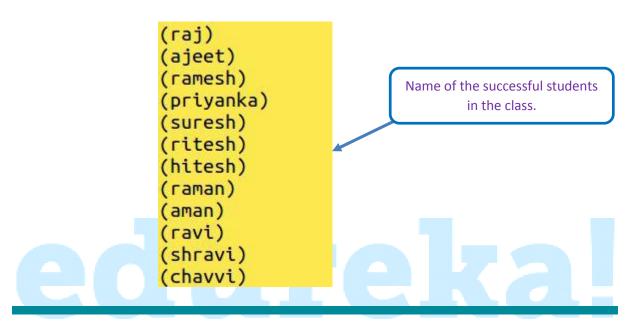
#### 1.2 Execute the Pig commands

Use the following Pig commands in the sequence to join the datasets and find out the result:

```
grunt> Student = load '/student' as (name:chararray, roll:int);
grunt> Results = load '/results' as (roll_:int, status:chararray);
grunt> PassStudents = join A by roll, B by roll_;
grunt> PassStudentNames = foreach C generate name, status;
grunt> Passed = filter D by status matches 'pass';
grunt> dump Passed;
```

#### FIGURE 1-2 PIG COMMANDS

```
grunt> Student = load '/student' as (name: chararray, roll: int);
grunt> Results = load '/results' as (roll: int, grade: chararray);
grunt> PassStudents = filter Results by grade == 'pass';
grunt> PassStudentNames = join Student by roll, PassStudents by roll;
grunt> Passed = foreach PassStudentNames generate name;
grunt> dump Passed;
dunut> qnwb bassed;
```



You can also create a Pig script (say 'script.pig') and execute it from the shell rather than writing each statement individually in grunt shell.

#### FIGURE 1-3 PIG SCRIPT

```
user@ubuntu:~$ pig hadoop/Pig/Solution/script.pig
2014-04-20 03:42:03,479 [main] INFO org.apache.pig.Main - Apache Pig version 0.
12.1-SNAPSHOT (rexported) compiled Mar 22 2014, 07:41:29
2014-04-20 03:42:03,480 [main] INFO org.apache.pig.Main - Logging error message s to: /home/user/pig_1397979723478.log
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/user/hadoop-2.2.0/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]

p\s[4]-rod4][5-1.7.5.jar!\org/slf4]\rubriclessimpler.class]
27-43: Honug prugrud ru []21:176:\ubbriclessimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.classimpler.
```