

# Module-5: Pig and Pig Latin

---

Assignment Solution

edureka!

**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
1. Problem solution .....	2
1.1 Copy datasets to HDFS .....	2
1.2 Execute Pig commands .....	2

edureka!

## 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

```
user@ubuntu:~$ hadoop fs -put hadoop/Pig/Assignment_4/student hdfs:/
user@ubuntu:~$ hadoop fs -put hadoop/Pig/Assignment_4/results hdfs:/
user@ubuntu:~$ hadoop fs -ls hdfs:/
Found 3 items
-rw-r--r--  1 user supergroup      151 2014-04-19 09:17 hdfs:///results
-rw-r--r--  1 user supergroup     185 2014-04-19 09:17 hdfs:///student
drwxrwx---  - user supergroup      0 2014-04-19 09:14 hdfs:///tmp
```

#### 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;
```

ՃԼՈՍԷ՝ զՈՒՆ ԿԶԶԶԳ՝

Name of the successful students  
in the class.

### FIGURE 1-3 PIG SCRIPT

[2665C, 79bndf879pgoJcJ76f2\jqm\[\[A7f2\q10\!76\!2,7,1-51[AgoJ-[A7f2\p