



## Session 7: Pig

---

### Assignment 2

## *Session 7: Pig*

### *Assignment –PIG*

#### **Table of Contents**

1. Introduction .....	3
2. Objective .....	3
3. Prerequisites .....	3
4. Associated Data Files .....	3
5. Problem Statement .....	3
6. Expected output .....	5
7. Approximate Time to Complete Task .....	5

# Big Data and Hadoop Development

## 1. Introduction

In this assignment you will learn to do the wordcount program using Pig.

## 2. Objective

This assignment will help you to understand basic pig commands and its implementation.

## 3. Prerequisites

Acadgild's VM , or Linux operating system with Hadoop and Pig installed in it.

## 4. Associated Data Files

N/A

## 5. Problem Statement

1. Pig can be run in batch mode using \_\_\_\_\_ .

a) Pig shell command

b) Pig scripts **ANSWER : b**

c) Pig options

d) All of the mentioned

2. Which of the following is the default mode?

a) Mapreduce **ANSWER : a**

b) Tez

c) Local

d) All of the mentioned

3. Which of the following command can be used for debugging?

a) exec

b) execute

c) error

d) throw

4. Which of the following function is used to read data in PIG?

## Big Data and Hadoop Development

- a) WRITE
- b) READ
- c) LOAD **ANSWER : c**
- d) None of the mentioned

**5. Which of the following will run pig in local mode?**

- a) \$ pig -x local ... **ANSWER : a**
- b) \$ pig -x tez\_local ...
- c) \$ pig ...
- d) None of the mentioned

**6. PigUnit runs in Pig's \_\_\_\_\_ mode by default.**

- a) local
- b) tez
- c) mapreduce **ANSWER : c**
- d) None of the mentioned

**7. Point out the wrong statement:**

- a) Pig can invoke code in language like Java Only. **----\***
- b) Pig enables data workers to write complex data transformations without knowing Java.
- c) Pig's simple SQL-like scripting language is called Pig Latin, and appeals to developers already familiar with scripting languages and SQL.
- d) Pig is complete, so you can do all required data manipulations in Apache Hadoop with Pig.

**8. Pig Latin is \_\_\_\_\_ and fits very naturally in the pipeline paradigm while SQL is instead declarative.**

- a) functional
- b) procedural **ANSWER : b**
- c) declarative
- d) All of the mentioned

**6. Expected output**

**N/A**

**7. Approximate Time to Complete Task**

30 min