*Big Data and Hadoop Development*

**ACADGILD**

**1. Introduction**

You will work on the concepts of Pig.

**2. Objective**

This assignment will help you to understand pig concepts.

**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**

Give a brief answers to the questions below:

1. Why Map-reduce program is needed in Pig Programming?

**Answer: Pig Program leverage Map-Reduce framework to run its query.**

2. What are advantages of pig over MapReduce?

**Answer: It is fast, No Java background knowledge is needed, Simple SQL like syntax, easy to learn, readily available for Data Scientist and Data Analyst**

3. What is pig engine and what is its importance?

**Answer: It works as intermediator between Pig Latin and Map Reduce. It’s very important because it convert simple latin commands in to MapReduce task.**

4. What are the modes of Pig execution?

**Answer: Either Local Mode or Map-Reduce Mode.**

5. What is grunt shell in Pig?

**Answer: The Grunt shell of Pig is mainly used to write Pig Latin scripts.**

6. What are the features of Pig Latin language?

**Answer**

1. **Create simple and complex data flow using command**
2. **Adhoc data analysis**
3. **Iterative Processing**

7. Is Pig latin commands case sensitive?

**Answer: no**

8. What is a data flow language?

**Answer: Language whose execution can be represented by a graph or data flow**