//Assignment 1

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

A:- Pig is a high-level platform that makes many Hadoop data analysis issues easier to execute. The language we use for this platform is: Pig Latin. A program written in Pig Latin is like a query written in SQL, where we need an execution engine to execute the query. So, when a program is written in Pig Latin, Pig compiler will convert the program into MapReduce jobs. Here, MapReduce acts as the execution engine.

2) What are advantages of pig over MapReduce?

A:-

* Pig Latin uses a lot fewer lines of code than the Java MapReduce .
* No Knowledge of Java background is required.
* Performing a Join operation in Apache Pig is pretty simple.

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

A:-It is an interpreter between Pig latin script and map reduce jobs. It create environment to execute Pig scripts to map reduce jobs in parallel manner.

1. What are the modes of Pig execution?

A:- There are two types:-

### Local Mode

In this mode, all the files are installed and run from your local host and local file system. There is no need of Hadoop or HDFS. This mode is generally used for testing purpose.

**Command −**

pig –x local

### MapReduce Mode

MapReduce mode is where we load or process the data that exists in the Hadoop File System (HDFS) using Apache Pig. In this mode, whenever we execute the Pig Latin statements to process the data, a MapReduce job is invoked in the back-end to perform a particular operation on the data that exists in the HDFS.

**Command −**

pig -x mapreduce

1. What is grunt shell in Pig?

A:- It is an interactive shell to run Pig commands.

Invoke the Grunt shell in a desired mode (local/MapReduce) using the **−x** option

Example:-

pig –x local

pig -x mapreduce

1. What are the features of Pig Latin language?

A:-

Pig Latin script describes a directed acyclic graph (DAG), where the edges are data flows and the nodes are operators that process the data.

Pig Latin provides all of the standard data-processing operations, such as join, filter, group by, order by, union, etc.

1. Is Pig latin commands case sensitive ?

A:-No

1. What is a data flow language?

A:-**dataflow** programming is a programming paradigm that models a program as a directed graph of the **data** flowing between operations, thus implementing **dataflow** principles and architecture.