1. Why MapReduce program is needed in Pig Programming?

Ans. MapReduce requires programmers:

• Programmers must think in terms of map and reduce functions.

• Most probably Java programmers are required.

We cannot reduce programing to zero level. There will be some programing required that is why

Mapreduce is used in pig.

2)What are advantages of pig over MapReduce?

Ans. MapReduce has upto 200 lines code which is very hectic to code, whereas Pig compresses the code in just 8-9 line.

Also pig provides high level language for data analysts and data programmer.

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

Ans. To analyze data using Apache Pig, programmers need to write scripts using Pig Latin language. All these scripts are internally converted to Map and Reduce tasks. Apache Pig has a component known as Pig Engine that accepts the Pig Latin scripts as input and converts those scripts into MapReduce jobs.

It converts the 200 line code of mapreduce to 8-9 lines that is the great importance of mapreduce.

4.What are the modes of Pig execution?

Ans. Pig has two execution modes or exectypes:

Local Mode - To run Pig in local mode, you need access to a single machine; all files are installed and run using your local host and file system. Specify local mode using the -x flag (pig -x local). Note that local mode does not support parallel mapper execution with Hadoop 0.20.x and 1.0.0. This is because the LocalJobRunner of these Hadoop versions is not thread-safe.

Mapreduce Mode - To run Pig in mapreduce mode, you need access to a Hadoop cluster and HDFS installation. Mapreduce mode is the default mode; you can, *but don't need to*, specify it using the -x flag (pig OR pig -x mapreduce).

5. What is grunt shell in Pig?

**Ans .Grunt Shell**: Enter Pig commands manually using Pig’s interactive shell, Grunt.

6. What are the features of Pig Latin language?

Ans. Apache Pig is a high-level platform for creating programs that run on [Apache Hadoop](https://en.wikipedia.org/wiki/Hadoop). The language for this platform is called Pig Latin.[[1]](https://en.wikipedia.org/wiki/Pig_(programming_tool)#cite_note-mainpage-1) Pig can execute its Hadoop jobs in MapReduce, Apache Tez, or [Apache Spark](https://en.wikipedia.org/wiki/Apache_Spark). Pig Latin abstracts the programming from the [Java](https://en.wikipedia.org/wiki/Java_(programming_language)) MapReduce idiom into a notation which makes MapReduce programming high level, similar to that of [SQL](https://en.wikipedia.org/wiki/SQL) for [RDBMSs](https://en.wikipedia.org/wiki/RDBMS).

7. Is Pig latin commands case sensitive?

Ans .The names (aliases) of relations and fields are case sensitive. The names of Pig Latin functions are case sensitive. The names of parameters (see Parameter Substitution) and all other Pig Latin keywords are case insensitive.

In the example below, note the following:

1. The names (aliases) of relations A, B, and C are case sensitive.
2. The names (aliases) of fields f1, f2, and f3 are case sensitive.
3. Function names PigStorage and COUNT are case sensitive.
4. Keywords LOAD, USING, AS, GROUP, BY, FOREACH, GENERATE, and DUMP are case insensitive. They can also be written as load, using, as, group, by, etc.
5. In the FOREACH statement, the field in relation B is referred to by positional notation ($0).

8. What is a data flow language?

Ans. In [computer programming](https://en.wikipedia.org/wiki/Computer_programming), dataflow programming is a [programming paradigm](https://en.wikipedia.org/wiki/Programming_paradigm) that models a program as a [directed graph](https://en.wikipedia.org/wiki/Directed_graph) of the data flowing between operations, thus implementing [dataflow](https://en.wikipedia.org/wiki/Dataflow) principles and architecture.