1. **Why Mapreduce program is needed in Pig Programming?**

Ans. Pig – platform for analyzing large data sets. It consist of high level programming language (Pig Latin) for expressing data analysis programs, and its compiler which produces MapReduce programs ready to be executed using Hadoop

**2.** **What are advantages of pig over MapReduce?**

**Ans.** The following are the advantages of Pig over MapReduce:

* Pig is a high-level data flow platform, whereas MapReduce is a low-level data processing paradigm.
* Without writing complex Java implementations in MapReduce, programmers can achieve the same implementations very easily using Pig Latin.
* Pig provides nested data types like tuples, bags, and maps that are missing from MapReduce.
* Pig provides many built-in operators to support data operations like joins, filters, ordering, sorting etc. Whereas to perform the same function in MapReduce is a humongous task.

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

Ans. Pig engine acts as interpreter between pig Latin script and MapReduce jobs. It creates environment to execute pig scripts into series of MapReduce jobs in parallel manner.

**4. What are the modes of Pig execution?**

Ans. Pig has two execution modes or types. They are:

* **Local Mode:** In Local Mode of Pig execution, all the input data will be taken from local file system. After execution it provides output on top of local file system.  In local mode, Pig runs in a single JVM and accesses the local file system . This mode of suitable only for small datasets and when trying out Pig.
* **MapReduce Mode**: In this mode Pig will take the input form HDFS paths only, and after processing data it will put output files on top of HDFS. In MapReduce mode of execution, Pig translates queries into mapreduce jobs and runs them on a Hadoop Cluster.

5. **What is grunt shell in Pig?**

Ans. The Grunt shell of Pig is mainly used to write Pig Latin scripts. Prior to that, we can invoke any shell commands using **sh** and **fs**. After invoking the Grunt shell, we can run Pig scripts in the shell. In addition to that, there are certain useful shell and utility commands provided by the Grunt shell.

6. **What are the features of Pig Latin language?**

Ans.Pig Latin language comes with the following features :

* **Rich set of operators** − It provides many operators to perform operations like join, sort, filer, etc.
* **Ease of programming** − Pig Latin is similar to SQL and it is easy to write a Pig script if good at SQL.
* **Optimization opportunities** − The tasks in Pig optimize their execution automatically, so the programmers need to focus only on semantics of the language.
* **Extensibility** − Using the existing operators, users can develop their own functions to read, process, and write data.
* **Handles all kinds of data** − Pig analyzes all kinds of data, both structured as well as unstructured. It stores the results in HDFS.
* Pig Latin script describes a directed acyclic graph (DAG) where the edges are data flows and the nodes are operators that process the data.

7. **Is Pig Latin commands case sensitive?**

**Ans.** Pig Latin commandsare both case sensitive and case insensitive. User defined functions, the field name, and relations are case sensitive i.e. EMPLOYEE is not same as employee.

Whereas Pig script keywords are case insensitive i.e. LOAD is same as load.

It is difficult to say whether Pig is case sensitive or case insensitive. For instance, user defined functions, relations and field names in Pig are case sensitive. On the other hand, keywords in Apache Pig are case insensitive.

8. **What is a data flow language?**

Ans. To access the external data, every language must follow many rules and regulations. The instructions are flowing through data by executing different control statements, but don’t get moved. Dataflow language can get a stream of data which passes from one instruction to another instruction to be processed. Pig can easily process those conditions, jumps, loops and process the data in efficient manner.