Give a brief answers to the questions below:

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

Ans. Pig is application that runs on top of MapReduce and abstracts Java MapReduce jobs away from developers. Pig Latin queries are converted to Map and Reduce jobs and hence they take advantage of parallel processing.

Q2. What are advantages of pig over MapReduce?

Ans.1. Pig is application that runs on top of MapReduce and abstracts Java MapReduce jobs away from developers.

2. Pig Latin uses a lot fewer lines of code than the Java MapReduce script.

3. The Pig Latin script was is easier to read for someone without a Java background.

4. MapReduce jobs can written in Pig Latin.

5.Java is a great and powerful language, but it has a higher learning curve than something like Pig Latin. Therefore, using a higher-level language, like Pig Latin, enables many more developers/analysts to write MapReduce jobs.

Q3. What is pig engine and what is its importance?

Ans .Execute engine execute the MapReduce jobs which are submitted to Hadoop in a sorted order. Finally, these MapReduce jobs are executed on Hadoop producing the desired results by pig execution engine.

Q4. What are the modes of Pig execution?

Ans. Apache pig has two modes, namely, Local Mode and HDFS mode.

Local Mode Pig-x local

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.

MapReduce Mode (default)

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.

Q5. What is grunt shell in Pig?

Ans. Interactive Mode (**Grunt shell**) – We can run Apache Pig in interactive mode using the Grunt shell. In this shell, you can enter the Pig Latin statements and get the output (using Dump operator).After invoking the Grunt shell, we can run your Pig scripts in the shell. In addition to that, there are certain useful shell and utility commands provided by the Grunt shell

Q6. What are the features of Pig Latin language?

Ans .Apache Pig comes with the following features –

1.**Rich set of operators** – It provides many operators to perform operations like join, sort, filer, etc.

2.**Ease of programming** – Pig Latin is similar to SQL and it is easy to write a Pig script if you are good at SQL.

3.**Optimization opportunities** – The tasks in Apache Pig optimize their execution automatically, so the programmers need to focus only on semantics of the language.

4.**Extensibility** – Using the existing operators, users can develop their own functions to read, process, and write data.

5.**UDF's** – Pig provides the facility to create User-defined Functions in other programming languages such as Java and invoke or embed them in Pig Scripts.

6.**Handles all kinds of data** – Apache Pig analyzes all kinds of data, both structured as well as unstructured. It stores the results in HDFS.

7. Is Pig latin commands case sensitive?

Ans. **Keywords** in Pig Latin are not case-sensitive; for example, LOAD is equivalent to load . But relation and field names are. So A = load 'foo'; is not equivalent to a = load 'foo'; . UDF names are also case-sensitive, thus COUNT is not the same UDF as count .

8. What is a data flow language?

Ans.In computer programming, 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. A language like Pig which enables us to do data flow programming instead of control flow programming is called data flow language.