**Qn1 Why MapReduce program is needed in Pig Programming?**

Pig Latin is Pig's language that allows developers to sort, join, parse, transform and calculate unstructured and semi-structured data in MapReduce all while using a language similar to SQL versus Java. Pig is an application that works on top of MapReduce, Yarn or Tez. Pig is written in Java and compiles Pig Latin scripts into to MapReduce jobs. Pig is a compiler that takes Pig Latin scripts and transforms them into Java. This is why Mapreduce is needed in Pig Programming.

**Qn2 What are advantages of pig over MapReduce?**

1.Development time is very less

2.Learning curve is not steep, anyone who does not know how to write map-reduce or SQL can write map-reduce job using pig.

3.Procedural, not declarative unlike SQL, so easier to follow the commands. it is much more like an english language. It is easier to read for someone with a little SQL background.

4.Since it is procedural, you could control of the execution of every step. If you want to write your own UDF(User Defined Function) and inject in one specific part in the pipeline, it is straightforward.

5.It could optimize the program beginning to end and optimizer could produce an efficient plan to execute.

6.Enjoys everything that Hadoop offers, parallelization, fault-tolerancy with many relational database features.

7.It is quite effective for unstructured and messy large datasets. Actually, Pig is one of the best tool to make the large unstructured data to structured.

8.It's faster to write in Pig Latin because of the less number of lines.

9.No need to import libraries

**Qn3 What is pig engine and what is its importance?**

**Pig Engine :**

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

**Pig Engine Importance :**

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.

**Qn4 What are the modes of Pig execution?**

We can run Apache Pig in two modes, namely, Local Mode and HDFS mode.

**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.

**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.

**Qn5 What is grunt shell in Pig?**

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. The Grunt shell of Apache Pig is mainly used to write Pig Latin scripts.

**Qn6.What are the features of Pig Latin language?**

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.

**Qn7.Is Pig latin commands case sensitive?**

Pig is case sensitive as well as case insensitive.

case sensitive : For alias, in-built functions, UDFs

case in-sensitive :For operators like foreach,as etc

**Qn8.What is a data flow language?**

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 data doesn'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, process the data in efficient manner.