1.Why MapReduce program is needed in Pig Programming?

* It requires Mapreduce program to perform mapper and reducer class functions.
* Pig's infrastructure layer consists of a compiler that produces sequences of Map-Reduce programs, for which large-scale parallel implementations already exist.
* Pig runs on Hadoop needs maprecude to perform mapping and reducing operations.

2. What are advantages of Pig over MapReduce?

* Pig simplifies the use of Hadoop by allowing SQL-like queries to a distributed dataset.
* It allows users to describe how data from one or more inputs should be read, processed, and then stored to one or more outputs in parallel.
* User need not necessarily think in terms of map and reduce, because based on the task pig automatically convert into map or reduce.
* In mapreduce it’s too difficult to join multiple data sets.
* Pig Latin language, aims at DataFlow operations instead of control flow operations, hence more focus is on data analysis.

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

* Pig has a component known as Pig Engine that accepts the Pig Latin scripts as input and converts those scripts into MapReduce jobs.
* It acts as a interpreter between pig Latin scripts and map reduce jobs.
* It allows pig script to execute into a series of map reduce job in parallel.

4.What are the modes of Pig execution?

* It has two modes namely
* Local mode
* Mapreduce mode

Local Mode - To run pig in local mode we need not install or start hadoop, all we need is acess to single file system.

* All files are installed and run using your local host and file system
* Specify local mode using the -x flag (pig -x local).

MapReduce Mode – To run in this mode we need to start hadoop

* Mostly Mapreduce mode is the default mode.
* Here pig script runs and stored in HDFS.
* Pig and pig –x mapreduce to execute pig in mapreduce mode

5. What is Grunt Shell in Pig?

* The Grunt shell of Apache Pig is mainly used to write Pig Latin scripts
* Grunt is an interactive shell for the Pig queries.
* We can run Pig programs in three ways via Script, Grunt, or embedding the script into Java code.
* Used when script file is not provided.

6. What are the features of Pig Latin language?

* Rich set of operators.
* Ease of programming.
* Optimization opportunities.
* Extensibility.
* UDF’s(User Defined Functions).

7. Is Pig Latin commands case sensitive?

* Pig Latin cannot decide whether it is case-sensitive.
* Keywords in Pig Latin are not case-sensitive.for example load is equal to LOAD
* But relation and field names are case-sensitive that is r1= load ... is not equal to R1=load...
* UDF names are also case-sensitive.

8. What is a data flow language?

* The instructions are flowing through data by executing different control statements, but data does not get moved. Data Flow Language can get a stream of data which passes from one instruction to another instruction to be processed. Pig being a data flow language can easily process those conditions ,jumps ,loops and process data in efficient manner.