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

Pig Convert his script in Map reduce Program in backend.

1. What are advantages of pig over MapReduce?

PIG is a data flow language, the key focus of Pig is manage the flow of data from input source to output store. As part of managing this data flow it moves **data** feeding it to process1, taking output and feeding it to process2. The core features are preventing execution of subsequent stages if previous stage fails, manages temporary storage of data and most importantly compresses and rearranges processing steps for faster processing. While this can be done for any kind of processing tasks Pig is written specifically for managing data flow of Map reduce type of jobs. Most if not all jobs in a Pig are map reduce jobs or data movement jobs. Pig allows for custom functions to be added which can be used for processing in Pig, some default ones are like ordering, grouping, distinct, count etc.  
    Map reduce on the other hand is a data processing paradigm, it is a framework for application developers to write code in so that its easily scaled to PB of tasks, this creates a separation between the developer that writes the application vs the developer that scales the application. Not all applications can be migrated to Map reduce but good few can be including complex ones like k-means to simple ones like counting uniques in a dataset.

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

4. What are the modes of Pig execution?

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.

5. What is grunt shell in Pig?

## Shell Commands

The Grunt shell of Apache Pig is mainly used to write Pig Latin scripts. Prior to that, we can invoke any shell commands using sh and fs.

### sh Command

Using sh command, we can invoke any shell commands from the Grunt shell. Using sh command from the Grunt shell, we cannot execute the commands that are a part of the shell environment (ex − cd).

Syntax

Given below is the syntax of sh command.

grunt> sh shell command parameters

6. What are the features of Pig Latin language?

* + Pig Engine Parses, compiles Pig Latin scripts into MapReduce jobs run on top of Hadoop.
  + Pig Latin is declarative, SQL-like language; the high level language interface for Hadoop
  + Fewer lines of code (Writing map reduce like writing SQL queries)
  + Re-use the code (Pig library, Piggy bank)

7. Is Pig latin commands case sensitive?

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';

8. What is a data flow language?

A programming paradigm in which computation is modelled as a directed graph (which may or may not contain cycles), the nodes of which are either data sources (producers of data), data sinks (consumers), or "processing elements" which compute some function; and the arcs of which represent dataflow between nodes.