Give a brief elaboration of the below questions.

**Swapna\_Assignment**

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

**Pig:**

* Pig is a high level data flow language
* It is execution framework for parallel computation for MapReduce operations.
* It is used to process structured and semi structured data

**Map-Reduce:**

* MapReduce is a framework using which we can write applications to process huge amounts of data, in parallel, on large clusters of commodity hardware in a reliable manner.
* Can implement MapReduce jobs in three ways (Java, Pig and Hive)
  + Java is used to handle structured, unstructured and Semi-Structured data
  + Pig is used to handle Structured and Semi-Structured data
  + Hive is used to handle structured data
* Pig is a top level project of apache to process data with less complexity and less Development effort.
* Programmers can write Pig Latin scripts and execute on top of MapReduce.
* Without MapReduce can't execute Pig Script to process data.

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

* Pig is High level of Abstraction
* Decrease development time. Code Efficiency is relatively less
* Joins are easily implemented in Pig
* Easy to control the execution of code in every step.
* Process the data quickly

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

* Pig is an engine to execute pig Latin scripts
* It acts interpreter between Pig Latin Script and Map reduce Jobs.
* Creating environment to execute Pig Latin Scripts into MapReduce jobs in parallel.

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

* Local Mode
  + In Local Mode, all input data taken from local file system.
  + After execution it provides output on top of local file system.
  + To start the local mode, use command :
    - # pig -x local
* Map-Reduce Mode
  + In MapReduce mode, all input data taken form HDFS paths only,
  + After processing data it will put output files on top of HDFS.
  + To Start Map-Reduce mode, Use command :
    - # pig or pig -x mr or pig -x mapreduce
* Tez Mode
  + Tez provides an alternative execution engine than MapReduce focusing on performance
  + To Start Tez mode, use command:
    - # pig -x tez

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

* Grunt shell used to write Pig Latin scripts.
* Can invoke any shell commands using sh and fs.
  + sh Command
    - Using sh command, can invoke any shell commands from the Grunt shell.
  + fs Command
    - Using the fs command, we can invoke any FsShell commands from the Grunt shell.

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

* Process the data quickly
* Programmer can transform it by sorting, grouping, joining, projecting, and filtering data
* User Defined Functions

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

* The aliases names and fields are case sensitive.
* The names of Pig Latin functions are case sensitive.
* The names of parameters and all other Pig Latin keywords are case insensitive.

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

* Dataflow language can get the steam of data which passes from one instruction to another instruction to be processed.
* Pig can easily process those conditions, jumps, loops.