Session 3

# Assignment 1

1. HDFS is built around the idea that data is written \_\_\_\_\_but read many times.
   1. Many
   2. Twice
   3. Data already exist
   4. Once

Answer – Once

1. Hadoop divides input into fixed size pieces called what?
   1. output result
   2. input splits
   3. input data
   4. input blogs

Answer – b. input splits

1. All the blocks are replicated in other nodes for \_\_\_\_\_\_
   1. Security
   2. Big Data
   3. Pool
   4. Fault Tolerance

Answer – d. Fault Tolerance

1. Block size can be changed using the properties in \_\_\_\_\_\_
   1. core-site.xml
   2. Hadoop-env.sh
   3. hdfs-site.xml
   4. yarn-site.xml

Answer – c. hdfs-site.xml

1. Hadoop uses the \_\_\_\_\_\_representation of the data stored in the file blocks known as Input splits.
   1. Physical
   2. Logical
   3. Mechanical
   4. None

Answer – b. Logical

1. DFS calls NameNode to create file in file system’s\_\_\_\_\_
   1. Dataspace
   2. Resourcespace
   3. Namespace
   4. Nodespace

Answer – c. Namespace

1. Data packets are streamed to first DataNode in the \_\_\_\_\_\_\_\_
   1. Handshake
   2. Pipeline
   3. hard disk
   4. hdfs

Answer – b. Pipeline

1. The client has finished writing data, it calls \_\_\_\_\_\_\_on the stream.
   1. close()
   2. read()
   3. open()
   4. check()

Answer – a. close()

1. Blocks are read in order, with the \_\_\_\_\_\_\_\_\_ opening new connections to datanodes as the client reads through the stream.
   1. DFSoutputstream
   2. DFSInputStream
   3. DFStrackManager
   4. DFSStringConcatination

Answer – a. DFSoutputstream

1. If I have 100 input splits, how many maps will run?
   1. 200
   2. 50
   3. 100
   4. 1

Answer - 100