MODULE 2: QUIZ QUESTIONS

1. Which of the following are the Big Data Solutions Candidates?

    a) Processing 1.5 TB data everyday
    b) Processing 30 minutes Flight sensor data
    c) Interconnecting 50K data points (approx. 1 MB input file)
    d) Processing User clicks on a website
    e) All of the above

2. Which of the following could be the data sources for Spark?

    a) Auto-populated data
    b) Files in HDFS
    c) Files in a local File System
    d) Kafka
    e) All of the above

3. Spark in Standalone mode means it is using no Cluster Manager

    a) True
    b) False

4. Entire Spark calculations in cluster are co-ordinated by:

    a) Master
    b) Executer
    c) Zookeeper
    d) SparkContext

5. Which of them is NOT the part of Apache Spark ecosystem:

    a) SparkSQL
    b) Shark and Impala
    c) Bagel
    d) GraphX

6. Spark can replace an existing enterprise OLTP system:

    a) True
    b) False

7. Spark allows different machines with different JAVA_HOME/SCALA_HOME variables in cluster:

    a) True
    b) False

8. In Spark, data can be loaded only from local file system?

    a) True
    b) False

Solution

1. (e) All of the above
2. (e) All of the above
3. (b) False
4. (d) SparkContext
5. (b) Shark and Impala
6. (b) False
7. (a) True
8. (b) False