

## Assignment 12.1

### 1. What is in-memory processing in Spark?

- a) Processing data in each node
- b) Storing data in RDBMS during processing
- c) Maximum effective utilization of RAM during the processing
- d) Using more no of CPU threads

**Answer: c) Maximum effective utilization of RAM during the processing**

### 2. What are the features of Apache Spark?

- a) In-memory processing
- b) Ease of use APIs
- c) Unified high level tools
- d) Runs Everywhere (Hadoop, Mesos, standalone, or in the cloud. It can access diverse data sources including HDFS, Cassandra, HBase, S3.)
- e) All the above
- f) None of the above

**Answer: e) All the above**

### 3. Apache Hadoop is more active project than Apache Spark in open source community in the last year.

- a) true
- b) false

**Answer: a) true**

### 4. Driver program will be launched in every node of the worker.

- a) true
- b) false

**Answer: b) false, Tasks will be launched in each node under executor.**

### 5. Spark only supports Stream processing.

- a. Yes
- b. No

**Answer: b. No, Spark supports multiple processing including batch and stream.**