*Spark*

**A C A D G I L D Page 1**

Session 5: Spark Core Part 2

Assignment 3*Spark*

**A C A D G I L D Page 2**

*Session 5: Spark Core Part 2*

*Assignment 3*

**Table of Contents**

1. Introduction .......................................................................................................................................... 3

2. Objective ............................................................................................................................................... 3

3. Prerequisites ......................................................................................................................................... 3

4. Associated Data Files ............................................................................................................................ 3

5. Problem Statement ............................................................................................................................... 3

6. Approximate Time to Complete Task ................................................................................................... 3 *Spark*

**A C A D G I L D Page 3**

**1. Introduction**

In this assignment you need to answer True/False for the given questions

**2. Objective**

This assignment will help you to understand Big Data basics.

**3. Prerequisites**

None

**4. Associated Data Files**

N/A

**5. Problem Statement**

1. Network shuffle happens for map operation.

Ans :No

2. Network shuffle happens for groupByKey operation.

Ans:Yes

3. Lower level scheduler will schedule data based on locality.

Ans :Yes

4. Partial aggregation increases data movement.

Ans :Yes

5. Transformations are executed on the Spark Workers.

ANS :Yes

**6. Approximate Time to Complete Task**

10 min