*Spark*

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

Session 9: Application Develop-ment – Demo

Assignment 1*Spark*

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

*Session 9: Application Development – Demo*

*Assignment 1*

**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 write the code for the given questions.

**2. Objective**

This assignment will help you to understand Big Data basics.

**3. Prerequisites**

APIs in Spark

**4. Associated Data Files**

N/A

**5. Problem Statement**

1.Implement an intersection of the below list

List((1, "panda"), (2, "happy"))

List((2, "pandas")

val inputFirst = sc.parallelize(List((1,”panda”),(2,”happy”)))

val inputSecond = sc.parallelize(List((2,”pandas”)))

val result = inputFirst.cogroup(inputSecond)

result.collect()

**6. Approximate Time to Complete Task**

30 min