*Spark*

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

Session 4: Spark Shell

Assignment 2*Spark*

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

*Session 4: Spark Shell*

*Assignment 2*

**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 question.

**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. Create a RDD of prime numbers from 1 to 100.

def isprimes(n: Int): List[Int] = {

require(n >= 2)

val primes = ListBuffer(2)

for (i <- 3 to n) {

if (prime(i, primes.iterator)) {

primes += i

}

}

primes.toList

}

scala> val list= sc.parallelize(List( 1 to 100))

list: org.apache.spark.rdd.RDD[scala.collection.immutable.Range.Inclusive] = ParallelCollectionRDD[2] at parallelize at <console>:27

scala> val primeValue = list.filter(isprimes)

2. Find the square root of even numbers from 1 to 100.

scala> def isEven(number: Int) = number % 2 == 0

isEven: (number: Int)Boolean

scala>val evens = Seq.range(0, 100).filter(isEven(\_))

evens: Seq[Int] = List(0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98)

scala>math.sqrt(evens)

**6. Approximate Time to Complete Task**

30 min