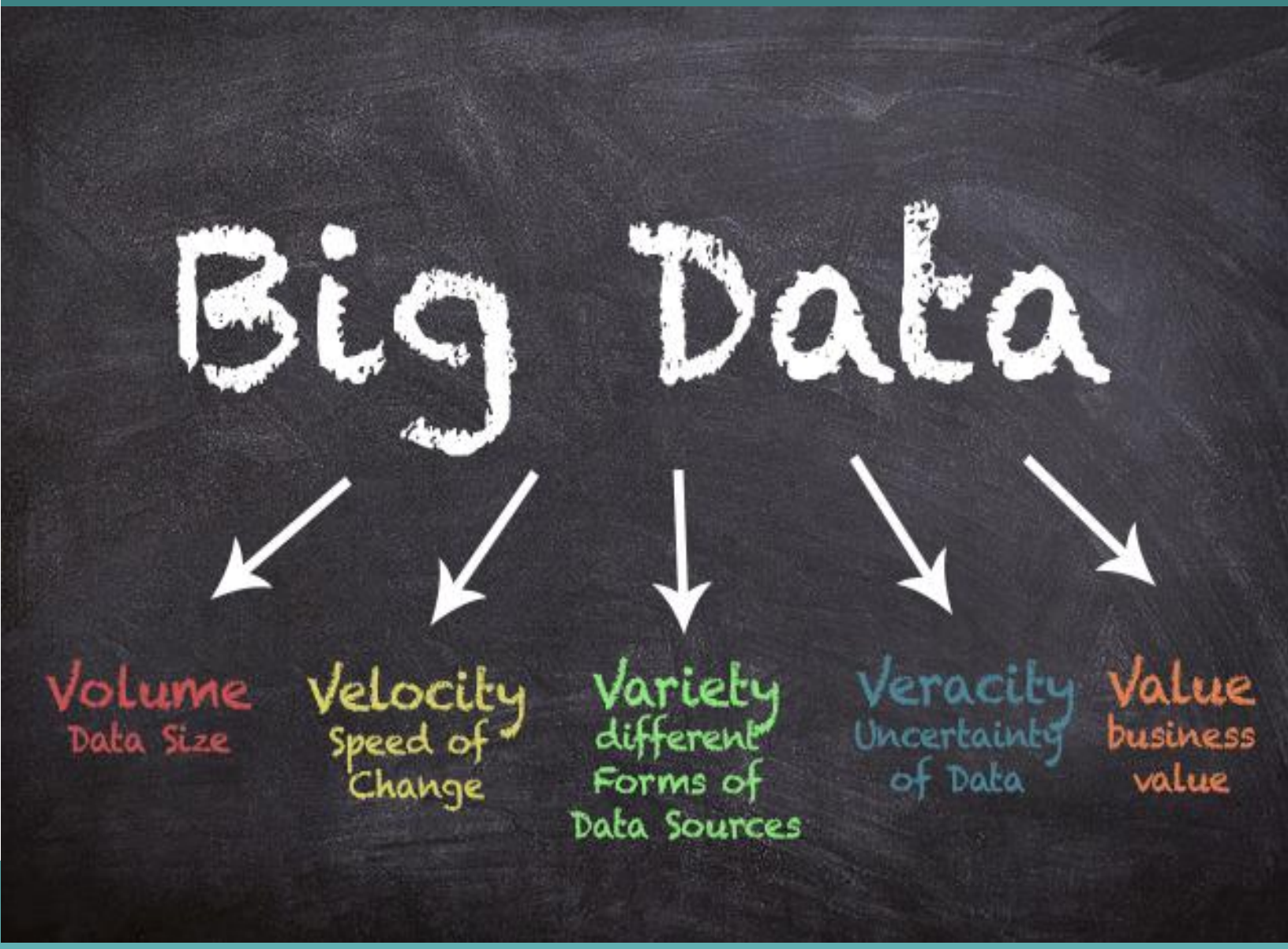
08/15/2018 | By: Ajit K Prasad



Big Data Engineering with Hadoop & Spark

Assignment on Scala Basics

Scala

APACHE Spark™

Acadgild

Session 14: Assignment 14.1

This assignment is aimed at consolidating the concepts that was learnt during the Scala Basics session of the course.

Problem Statement

Task 1:

Given is a list of strings – List [String] (“alpha”, “gamma”, “omega”, “zeta”, “beta”)

scala> val list = List[String] ("alpha", "gamma", "omega", "zeta", "beta")

1. Find count of all strings with length 4.
   scala> list.count(s => s.length == 4)
   scala> list.filter(s => s.length == 4)
2. Convert the list of string to a list of integers, where each string is mapped to its corresponding length.
   scala> list.map(s => s.length)
3. Find count of all strings which contain alphabet ‘m’.
   scala> list.count(s => s.contains("m"))
   scala> list.filter(s => s.contains("m"))
4. Find the count of all strings which start with the alphabet ‘a’.
   scala> list.count(s => s.startsWith("a"))
   scala> list.filter(s => s.startsWith("a"))

```
scala> val list = List[String] ("alpha", "gamma", "omega", "zeta", "beta")
list: List[String] = List(alpha, gamma, omega, zeta, beta)

scala> list.count(s => s.length == 4)
res17: Int = 2

scala> list.filter(s => s.length == 4)
res18: List[String] = List(zeta, beta)

scala> list.map(s => s.length)
res19: List[Int] = List(5, 5, 5, 4, 4)

scala> list.count(s => s.contains("m"))
res20: Int = 2

scala> list.filter(s => s.contains("m"))
res21: List[String] = List(gamma, omega)

scala> list.count(s => s.startsWith("a"))
res22: Int = 1

scala> list.filter(s => s.startsWith("a"))
res23: List[String] = List(alpha)
```

Task 2:

Create a list of tuples, where the 1st element of the tuple is an int and the second element is a string.
Example - ((1, ‘alpha’), (2, ‘beta’), (3, ‘gamma’), (4, ‘zeta’), (5, ‘omega’))

scala> *val listofTuples = List[(Int, String)] ((1, "alpha"), (2, "beta"), (3, "gamma"), (4, "zeta"), (5, "omega"))*

1. For the above list, print the numbers where the corresponding string length is 4.

   scala> *for ((i,s) <-listofTuples if s.size == 4) yield i*

```
scala> val listofTuples = List[(Int, String)] ((1, "alpha"), (2, "beta"), (3, "gamma"), (4, "zeta"), (5, "omega"))
listofTuples: List[(Int, String)] = List((1,alpha), (2,beta), (3,gamma), (4,zeta), (5,omega))

scala> for ((i,s) <-listofTuples if s.size == 4) yield i
res1: List[Int] = List(2, 4)
```

2. Find the average of all numbers, where the corresponding string contains alphabet ‘m’ or alphabet ‘z’.

   scala> *for ((i,s) <-listofTuples if s.contains("m") | s.contains("z")) yield i*

   scala> *val modifiedList = listofTuples.filter(s => s._2.contains('m') ||*
   scala> *s._2.contains('z')).map(x => x._1)*

   scala> *var sum = 0*

   scala> *val ml = modifiedList.foreach(x => sum+=x)*

   scala> *val counts = listofTuples.count(s => s._2.contains('m') || s._2.contains('z'))*

   scala> *val TheAverage = sum /counts*

```
scala> for ((i,s) <-listofTuples if s.contains("m") | s.contains("z")) yield i
res1: List[Int] = List(3, 4, 5)

scala> val modifiedList = listofTuples.filter(s => s._2.contains('m') || s._2.contains('z')).map(x => x._1)
modifiedList: List[Int] = List(3, 4, 5)

scala> var sum = 0
sum: Int = 0

scala> val ml = modifiedList.foreach(x => sum+=x)
ml: Unit = ()

scala> val counts = listofTuples.count(s => s._2.contains('m') || s._2.contains('z'))
counts: Int = 3

scala> val TheAverage = sum / counts
TheAverage: Int = 4
```