**Lab2**

**The purpose of this Lab is to become more familiar with pyspark**

Supply solution for following tasks using pyspark on your local machine. Use RDD where possible.

1. Create some text file.
   1. Count a number of words in this file
   2. Count a number of different words in this file
   3. Write to the file pairs (word, number of its occurrences), sorted by number of occurrences of a word in the file
2. Define a set of K random points (x,y) with coordinates

-1 < x < 1 and -1 < y < 1

Let N be a number of points from this set that are inside a circle with center in (0, 0) and radius 1.

Calculate approximate value of **** as (4\*N)/K.