**Advanced Topics in Data Intensive Computing**

**Programming Assignment 1**

The goal of this asignment is to get you some hands on experience with map-reduce programming.   
  
Assume that you have a large data set comprising of temperature readings of various zip codes (a sample data set will be provided shortly). Assume that the zip codes range from 000 to 999. For any zip code there can be multiple readings (you may assume that there will be at least one reading for each zip code).  
  
This assignment will be comprised of three parts.  
  
(1) Write map-reduce programs for computing the mean (average) of temperatures for each zip code. The output should be a list of zip codes and mean temperatures for each zip code.  
  
(2) Write map-reduce-style programs for computing the standard deviation for each zip code. The output should be a list of zip codes and the standard deviation for that zip code. For example, if the temperature values for the zip code 111 are 40, 60, 70, 80, 100 the standard deviation for 111 should be 20.  
  
(3) Write map-reduce-style programs for computing the zip codes whose mean temperatures are higher than the mean temperatures of both of its neighbors. For the sake of this assignment, the neighboring zip codes of the zip code 'x' are (x+1) mod 1000 and (x-1) mode 1000. The output should be a list of zip codes satisfying this constraint.  
  
Due Date: October 10, 2014  
  
[Sample Dataset-1](http://www.cs.uga.edu/~laks/courses/adic-2012f/MRTest-1.txt) (2000 records)  
[Sample Dataset-2](http://www.cs.uga.edu/~laks/courses/adic-2012f/MRTest-2.txt) (10002000 records)