National University of Computer and Emerging Sciences, Lahore Campus

	Course:	BIGDATA	**Course Code:**	**CS**
	Program:	MS(CS)	**Semester:**	**Spring 2018**
	Duration:	**15 Minutes**	**Total Marks:**	
	Paper Date:	**15-3-2018**	**NAME:**	
	Exam:	**Quiz 2**	**ROLLNO:**	

Instruction/Notes: Solve the exam in the space provided.

Question 1: [10 marks] Develop an <u>efficient</u> Map reduce algorithm to compute standard deviation of a **large set** of integers stored in a file. The standard deviation can be calculated using the following formula:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

Where N is the total number of integers, and $\bar{x}$ is the mean of N numbers.

Input file

```
12.4
13.9
5.6
2.9
30
23.2
45
```

Output should be the standard deviation = 13.74762