

## LAB EXERCISE

### Statistics

#### **Background:**

1. Your instructor will provide you with a text file, (*numbers.txt*), containing a large ( $N \leq 1000$ ) number of integers. The integers range in value from 0 to 100. The text file has been created with one value on each line. Due to the potential for the sum of the numbers to be very large, you should use a long integer in your calculation to find the average.
2. The number of integers in the file is unknown. You must read the text file until the EOF marker is encountered.
3. Your program must find the average, standard deviation, and mode of the list of numbers. The mode is defined as the value(s) present with the highest frequency. Calculating the standard deviation consists of the following steps:
  - a. Find the average of the list of numbers.
  - b. Determine the difference of each number from the average, and square each difference. Sum all the differences.
  - c. Divide this sum by (the number of values - 1).
  - d. Take the square root of the above division problem from step c.

Example, given this list of numbers: 7 4 5 9 10

- a. The average = 7
- b. Sum of square of differences:

$$(7 - 7)^2 + (4 - 7)^2 + (5 - 7)^2 + (9 - 7)^2 + (10 - 7)^2 = 26$$

0        +        9        +        4        +        4        +        9

- c. = 6.50
- d. = 2.55

4. For a normal distribution, 68.3% of the data will lie within one standard deviation of the average, while 95.4% will lie within two standard deviations.

#### **Assignment:**

1. Your program should print out the average, standard deviation, and mode of the data in *numbers.txt*. Format the real numbers to print with 2 decimal places.
2. Your program must utilize proper modular design (more than 1 file) and parameter passing.
3. Display run output and call instructor to your workspace for scoring.