

# **Sample Homework**



Paul Watta

Quiz #25

Average and Standard Deviation

September 23, 2013

## 1. Statement of the Problem

The problem here is to create a program that computes the average and standard deviation of a data set. The program will be tested with a data set of random test scores.

## 2. Description of Solution

Given a data set  $x_1, x_2, x_3, \dots, x_n$ , the average value  $\mu$  is given by

$$\mu = \frac{1}{n}[x_1 + x_2 + x_3 + \dots + x_n]$$

The variance, which measures how much the data differs from the average is given by:

$$VAR = \frac{1}{n}[(x_1 - \mu)^2 + (x_2 - \mu)^2 + (x_3 - \mu)^2 + \dots + (x_n - \mu)^2]$$

Finally, the standard deviation is given by the square root of the variance:

$$\sigma = \sqrt{VAR}$$

Since we are interested in test scores, I will assume that the data consist of integers, and hence will use an  $n$ -dimensional `int` array to store the  $n$  values.

To solve this problem, I created 3 functions:

```
void setupTestScores(int N, int x[])

float getAve(int N, int x[])

float getStdDev(int N, int x[])
```

The function `setupTestScores` creates  $N$  random numbers between 50 and 100 and stores the results in the array `x[]`. The functions `getAve` and `getStdDev` compute and return the average and standard deviation, respectively. In both functions, `x[]` is the incoming data array and  $N$  is the dimension of the array. The main programming construct used here is a `FOR` loop to compute the required sums. Finally, a `main()` function was written to create the data and call the functions.

### 3. Output and Testing

To test the program, I tried 2 data sets of test scores with  $N = 10$  and  $N = 20$  students. The results are shown below:

```
Test Scores: 91 55 60 81 94 66 53 83 84 85
              N: 10
              Average: 75.20
Standard Dev: 15.23
```

```
Test Scores: 58 93 59 77 74 84 64 69 54 69 63 68 91 90 53
              N: 15
              Average: 71.07
Standard Dev: 13.40
```

I verified by hand that the results are correct. Note that although I tested on 2 relatively small data sets, the program will work for a data set of any size. In the case of a very large data set, though, I wouldn't display all of the data to the screen. Rather, I would just display the final results.

```

/*-----

Author: Paul Watta
Date: September 23, 2013
Assignment: Quiz #25

-----/

#include <stdio.h>
#include <stdlib.h>

#define N 15    //Number of students

/* SETUP RAND SCORES */

void setupRandScores(int Dim, int x[])
{
    int i;

    for(i=0;i<N;i++)
    {
        x[i] = 50 + rand() % 51;
    }
}

/* GET AVE */

float getAve(int Dim, int x[])
{
    int i;
    float sum, ave;

    sum = 0;
    for(i=0;i<N;i++)
    {
        sum += x[i];
    }

    ave = (float) sum / (float) N;
    return(ave);
}

```

```

/* GET STD DEV */

float getStdDev(int Dim, int x[])
{
    int i;
    float sum, ave, var, std_dev;

    ave = getAve(N, x);

    sum = 0;
    for(i=0;i<N;i++)
    {
        sum += (x[i] - ave) * (x[i] - ave);
    }

    var = (float) sum / (float) N;
    std_dev = sqrt(var);

    return(std_dev);
}

/* MAIN */

int main()
{
    int i;

    float average, stdDev;
    int x[N];

    srand(100);

    setupRandScores(N, x);
    average = getAve(N, x);
    stdDev = getStdDev(N, x);

    printf("\n Test Scores: ");
    for(i=0;i<N;i++)
        printf("%d ", x[i]);

    printf("\n          N: %d", N);
    printf("\n      Average: %.2f", average);
    printf("\nStandard Dev: %.2f\n", stdDev);
    return 0;
}

```