Math 3300 Programming Assignment 5

Instructions: Work on the following programs and submit your source code to me via Blackboard. Send me 2 total cpp files.
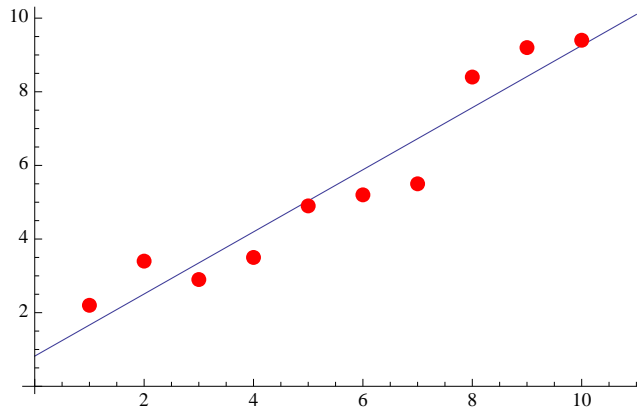
1. You will write a program which calculates the "best fit" line for a number of user-input data points and the "correlation constant".

    Here's an example:

    Suppose an experiment has been performed and the following points have been measured:

x	1	2	3	4	5	6	7	8	9	10
y	2.2	3.4	2.9	3.5	4.9	5.2	5.5	8.4	9.2	9.4

    Here's a graph of those data points along with the best fit line:

    The correlation constant is a real number between -1 and 1 which indicates how closely the data can be fit with a line. If r is close to 1, the data is very close to lying on a line with positive slope. If r is close to -1, the data is very close to lying on a line with negative slope. If r is close to 0, the data cannot be fit very well with a line.

    In our example, the equation for the best fit line is: $y = .843636 \cdot x + .82$ and the correlation constant is $r = .959309$.

    Your program will:

    (a) Get the points from a user-specified file (the user should be able to choose any filename). Each line of the file will contain an x-coordinate

and a y-coordinate, but you do not know how many lines are in the file.

(b) Your program will then calculate the equation for the best-fit line and also the correlation constant and display them to the monitor.

(c) Save your program as: **bestfit.cpp**

Here are the necessary formulas (n = the number of data points):

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} \quad \text{Average of x values}$$

$$\overline{y} = \frac{y_1 + y_2 + \cdots + y_n}{n} \quad \text{Average of y values}$$

$$s_x = x_1^2 + x_2^2 + \cdots + x_n^2 \quad \text{Sum of squares of x values}$$

$$s_y = y_1^2 + y_2^2 + \cdots + y_n^2 \quad \text{Sum of squares of y values}$$

$$\overline{s_x} = \frac{s_x}{n} \quad \text{Average square of x values}$$

$$\overline{s_y} = \frac{s_y}{n} \quad \text{Average square of y values}$$

$$s_{xy} = x_1y_1 + x_2y_2 + \cdots + x_ny_n \quad \text{Sum of product of x and y values}$$

$$\overline{s_{xy}} = \frac{s_{xy}}{n} \quad \text{Average product of x and y values}$$

Now, the best fit line will have equation:

$$y = ax + b$$

where

$$a = \frac{\overline{s_{xy}} - \overline{x} \cdot \overline{y}}{\overline{s_x} - \overline{x}^2}$$

$$b = \overline{y} - a\overline{x}$$

and the correlation constant is given by:

$$r = \frac{s_{xy} - n\overline{x} \cdot \overline{y}}{\sqrt{(s_x - n\overline{x}^2)(s_y - n\overline{y}^2)}}$$

So given the data points, you will need to determine: $n, \overline{x}, \overline{y}, s_x, s_y, \overline{s_x}, s_{xy}, \overline{s_{xy}}$ in order to calculate a, b, r.

2. Write a program which allows the user to input a single integer. The program will then calculate and output all factors of this number. If the number is positive, and the only 2 positive factors are the number 1 and itself, the program should indicate that the number is prime.

For example: The user enters 10. The output is:
-2 and -5

-1 and -10
1 and 10
2 and 5
10 is not a prime number

or if the user enters 7, the output is:
-1 and -7
1 and 7
7 is a prime number

You can choose the order you want the factors to be output, but they must come in pairs, and for full credit, the pairs shouldn't be repeated. Save your program as **factor.cpp**.