(Part II) Lab-session 4

1) Consider the following ordered pairs of data:

x-values:	2	4	8	11	16	19	21	27	29	31
y-values:	6	13	24	33	48	53	60	77	83	87

Show that these data are linearly correlated, that is the best fit function follows

$$y = \alpha x + \beta . \tag{1}$$

    a) Use the SLOPE and INTERCEPT functions to determine the coefficients α and β.

    b) Use the LINEST array function to determine the coefficients α, β and the square of the regression coefficient.

    c) Produce an XY chart with trendline to determine the same quantities.

    d) Compute the residues $z_i = \alpha x_i + \beta - y_i$. Plot the z_i-values against the x_i-values. Are the values randomly distributed? Does this plot confirm that the data are linearly correlated?

2) For the following ordered pairs of data

x-values:	2	4	8	11	16	19	21	27	29	31
y-values:	6	18	60	109	235	340	421	738	840	1000

produce an XY chart with trendline. Show that these data are correlated via a cubic equation, that is the best fit function follows

$$y = \alpha x^3 + \beta x^2 + \gamma x + \delta . \tag{2}$$

Determine the coefficients α, β, γ, δ.

3) Write a subroutine called "mean" which given a set of points (x_1, y_1), (x_2, y_2),...,(x_n, y_n) computes the averages:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i,$$

and writes the values into cells C1 and C2 of the Excel Worksheet. In order to try your programme take the set of points of question 1). The main challenge in writing this subroutine is to have a systematic way of reading the values of x_i and y_i from the Excel Worksheet. Once this is achieved the sums above should be computed by means of a loop structure.

Write the values x_i in column A and the values y_i in column B. One way of reading these values from the Excel Worksheet is to use the command line *ActiveCell.Offset(i,j).Value* as was done in the last example of Lecture 4. This command line gives the value of the cell located in row i and column j of the Excel Worksheet. However when using *ActiveCell.Offset(i,j).Value* your must be careful to notice that rows and columns are labelled starting at the active cell, which is then associated to row number 0 and column number 0. Write *Range("A1").Select* at the beginning of your subroutine so that the programme starts reading the values of x_i at cell A1 (cell A1 is now row 0, column 0).

4) If you have managed to solve question 3) successfully you may now be able to modify your subroutine so that given the same set of points it computes the *slope*, *intercept* and *regression coefficient* according to the formulae that were given in page 2 of Lecture 4.