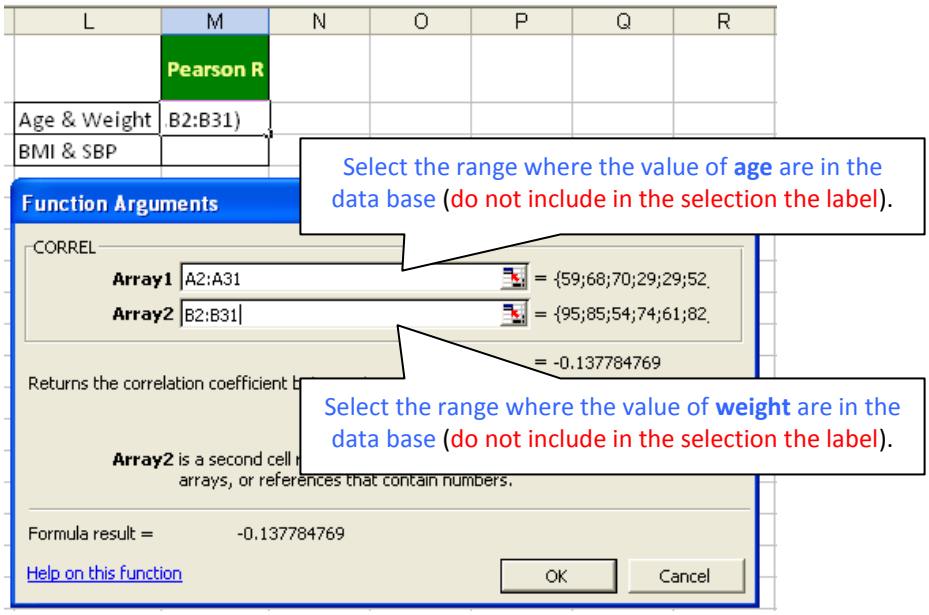
STATISTICAL ANALYSIS OF MEDICAL DATA: CORRELATION AND REGRESSION ANALYSIS - HINTS

To compute the Pearson correlation coefficient:

- Create next to the data the following table

	L	M
		Pearson R
Age & Weight		
BMI & SBP		

- Select the cell where you want the results (in this example M2) and identify the CORREL predefine function [Insert function – Statistical – Correl]
- Fill the window of CORREL function as in the image bellow:



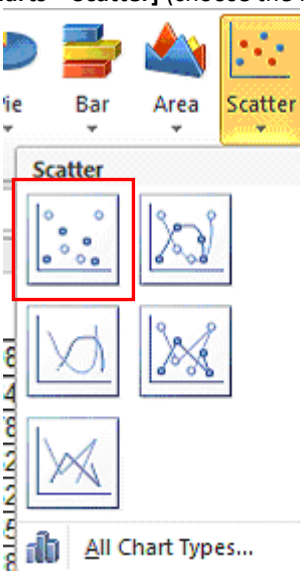
- Write the interpretation next to the result:

	L	M	N	O	P	Q	R	S	T
		Pearson R							
Age & Weight		-0.137785		Degree of association: no relation & Direction: negative					
BMI & SBP									

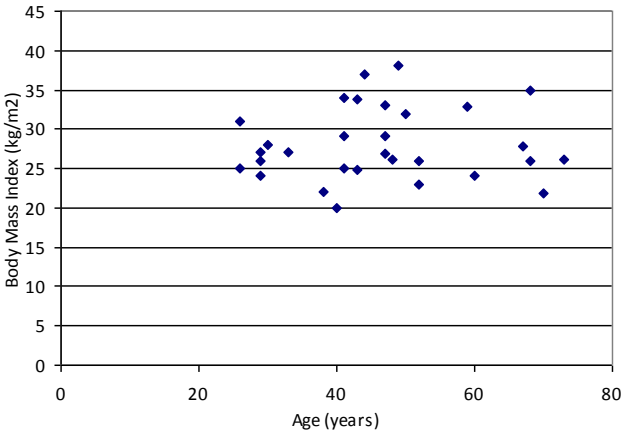
- Do the same steps for the second correlation coefficient (Request 3).

Scatter chart & Regression equation

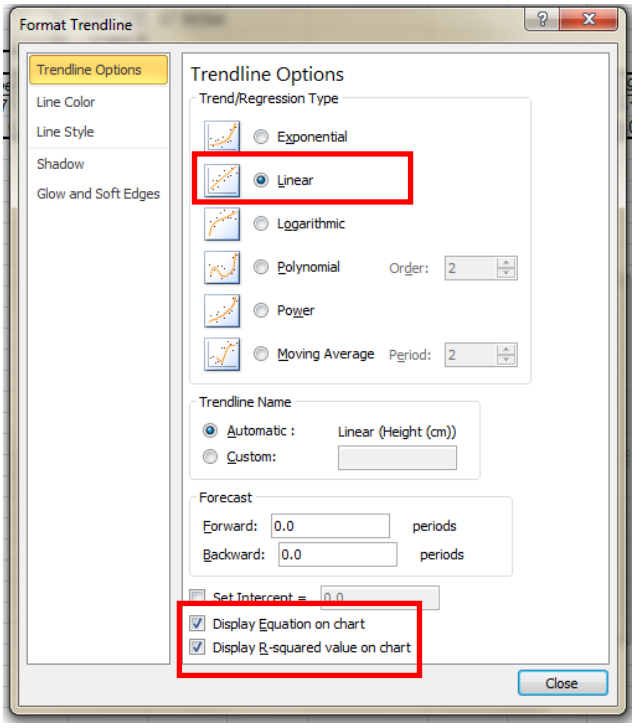
- To create a scatter plot: **[Insert – Charts – Scatter]** (choose the first type of Scatter)



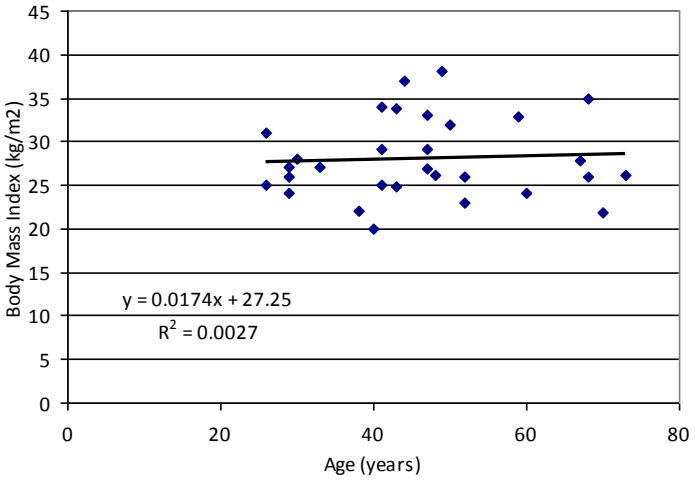
- Work on the graphical representation to look like the one in the image bellow:



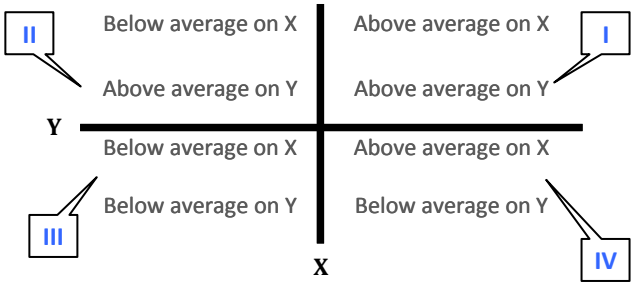
- **Add Trendline on chart**:
  - Select the data series for the trendline by clicking one of its markers;
  - Right-click and choose **Add Trendline** from the shortcut menu;
  - In the Add Trendline dialog box, pick a trend/regression type as **Linear**. Click also on **'Display Equation on chart'** and on **'Display r-squared value on chart'**:



- Your chart will be as in the image bellow:



- Interpretations of determination coefficient (R^2) and scatter and coefficients of regression:
  - R^2 answer to the following question: how much of the percentage of variation in Y can be explained by the linear relationship between Y and X? For Request 5, R^2 = 0.0027 → 0.27% from variation in BMI could be explained by the linear relation between BMI and Age.
  - Interpretation of Scatter: split the scatter plot in 4 cadres using the mean of X and the mean of Y:



If a linear relationship exists between X and Y, the markers of the plot will be in cadres II and IV (negative direction – descendant trend) or I and III (positive direction – ascendant trend). If the

markers are uniformly dispersed in all four cadres, the scatter indicates a null relationship between X and Y.

- Interpretation of coefficients of regression equation by example: Request 5.
  y = 0.0174x + 27.25: The equation shows that the coefficient for BMI in kg/m^2 is 0.0174 years. The coefficient indicates that for every additional kg/m^2 in BMI you can expect age to increase by an average of 0.0174 years. For age equal zero, the BMI is expected to be 27.25.