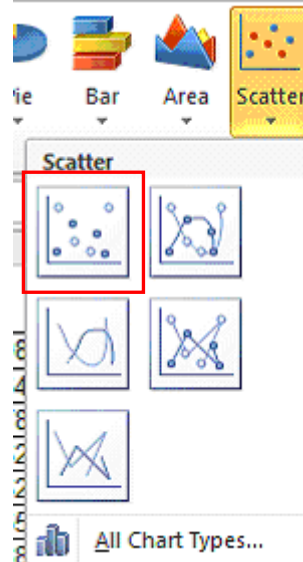


STATISTICAL ANALYSIS OF MEDICAL DATA: CORRELATION AND REGRESSION ANALYSIS - HINTS

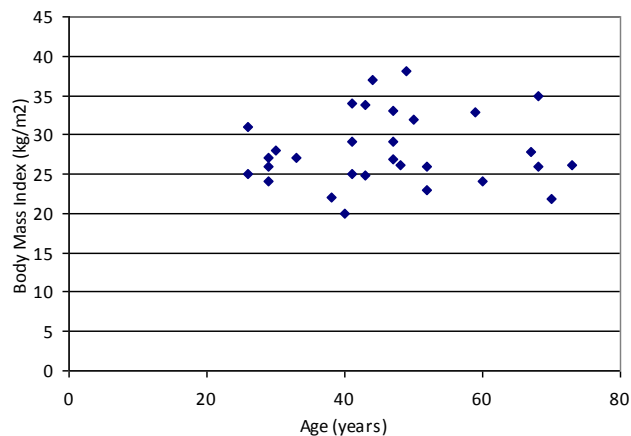
This document contains general hints! You have other database on which you are working on!

Scatter chart

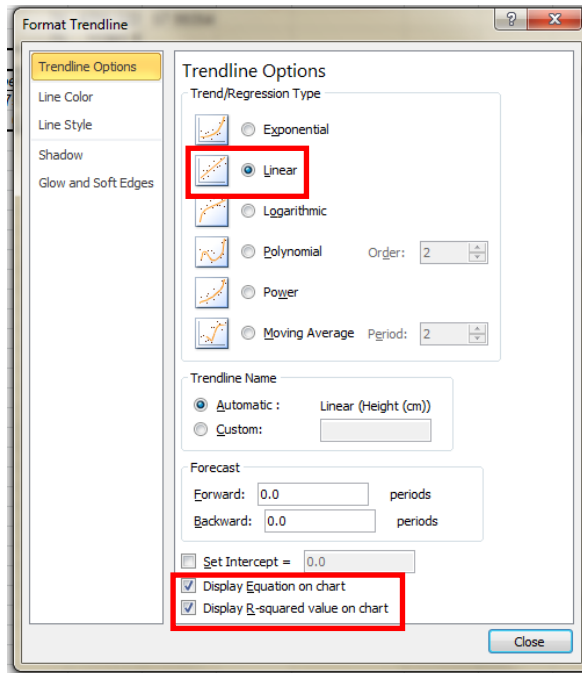
- To create a scatter plot: [Insert – Charts – Scatter] (choose the first type of Scatter)



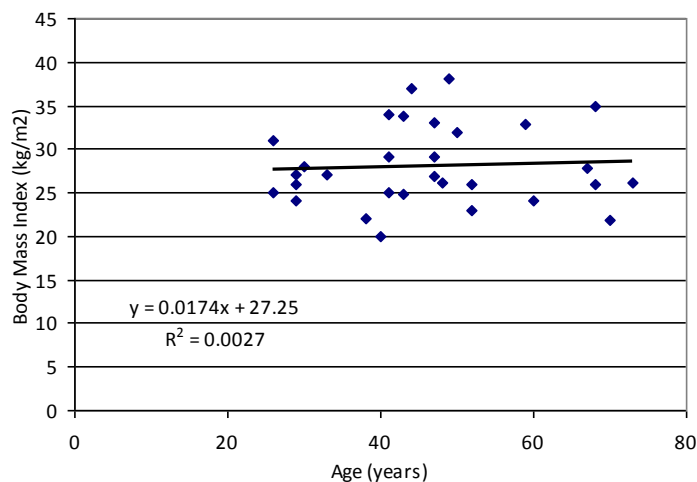
- Work on the graphical representation to look like the one in the image bellow:



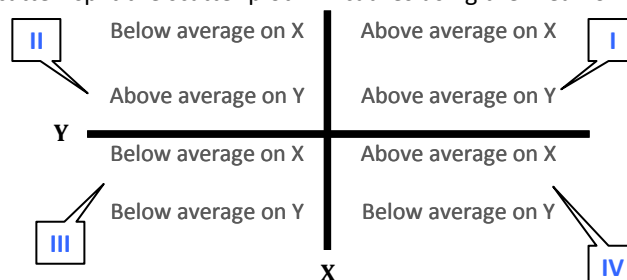
- Add Trendline on chart:**
 - Select the data series for the trendline by clicking one of its markers;
 - Right-click and choose **Add Trendline** from the shortcut menu;
 - In the Add Trendline dialog box, pick a trend/regression type as **Linear**. Click also on **'Display Equation on chart'** and 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, $R^2 = 0.0027 \rightarrow 0.27\%$ 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.