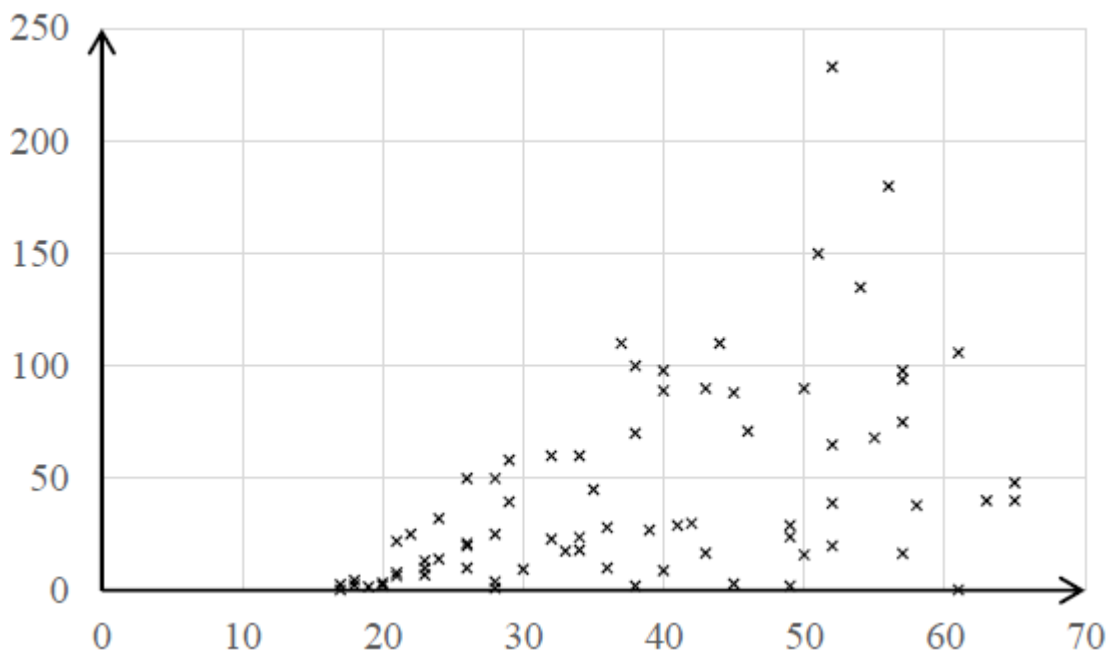# Correlation Overview (27/2/23)

Q1 [Practice/E]

Data has been collected for two variables (as shown below), and possible association between the variables is to be investigated. What issues should be considered, to ensure that the test is valid?

Q2 [Problem/H]

Show that the formula $r = \dfrac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$ can be written as

$r_s = 1 - \dfrac{6\sum d_i^2}{n(n^2-1)}$ when the data items are ranks.

[In other words, instead of using the formula for Spearman's coefficient, it is theoretically possible to use the standard formula for r.]

Q3 [Practice/E]

x_i	1	2	3	4	5
y_i	1	4	9	16	25

Find the PMCC for the above data, using formulae.

Comment on the suitability of the PMCC in this case.