Correlation Q3 [Practice/E](27/2/23)

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.

Solution

						$\sum$
x_i	1	2	3	4	5	15
y_i	1	4	9	16	25	55
${x_i}^2$	1	4	9	16	25	55
${y_i}^2$	1	16	81	256	625	979
$x_i y_i$	1	8	27	64	125	225

$$S_{xx} = \sum {x_i}^2 - n\bar{x}^2 = 55 - 5 \times \left(\frac{15}{5}\right)^2 = 10$$

$$S_{yy} = \sum {y_i}^2 - n\bar{y}^2 = 979 - 5 \times \left(\frac{55}{5}\right)^2 = 374$$

$$S_{xy} = \sum x_i y_i - n\bar{x}\bar{y} = 225 - 5 \times \left(\frac{15}{5}\right)\left(\frac{55}{5}\right) = 60$$

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{60}{\sqrt{10 \times 374}} = 0.981$$

Suitability?

- values taken by x are non-random, so PMCC not suitable (also, the y values are obviously artificial, and clearly non-random)

- plotted values form a curve; linear relation probably not appropriate

- plotted values don't give elliptical pattern