1.2) Measuring correlation

Worked example

Calculate the product moment correlation coefficient for the following data:

x	y
1	3
2	4
3	5
4	8

Your turn

Calculate the product moment correlation coefficient for the following data:

x	y
1	3
2	6
3	5
4	8

$r = 0.868$ (3 sf)

Worked example

From the large data set, the daily mean temperature, $t\ °C$, and the daily total rainfall, $r\ mm$, were recorded from 27th May to 5th June inclusive 1987 in Leuchars.

Day	1	2	3	4	5	6	7	8	9	10
t	8.5	9.0	10.3	12.8	13.5	12.8	9.8	8.8	10.0	10.4
r	0	2.4	8.1	0.2	0.4	tr	6.1	3.6	tr	31.8

a) State the meaning of tr in the table above.

b) Calculate the product moment correlation coefficient for the ten days, stating clearly how you deal with the 'tr' readings.

c) With reference to your answer to part b, comment on the suitability of a linear regression model for these data.

Your turn

From the large data set, the daily mean windspeed, w knots, and the daily maximum gust, g knots, were recorded for the first 10 days in September in Hurn in 1987.

Day	1	2	3	4	5	6	7	8	9	10
w	4	4	8	7	12	12	3	4	7	10
g	13	12	19	23	33	37	10	n/a	n/a	23

a) State the meaning of n/a in the table above.

b) Calculate the product moment correlation coefficient for the remaining 8 days.

c) With reference to your answer to part b, comment on the suitability of a linear regression model for these data.

a) Data on daily maximum gust is not available for these days

b) $r = 0.9533$ (4 sf)

c) r is close to 1 so there is strong positive correlation between daily mean windspeed and daily maximum gust. This means that the data points lie close to a straight line, so a linear regression model is suitable.