Definition (Correlation coefficient)

Let x_i and y_i be the values of two variables associated to the ith element in U ($i = 1, 2, \cdots, N$). The *correlation coefficient* (or simply, the correlation) between x and y is defined as

$$r_{xy,U} \equiv \frac{\sum_U (x_i - \bar{x}_U)(y_i - \bar{y}_U)}{\left(\sum_U (x_i - \bar{x}_U)^2 \sum_U (y_i - \bar{y}_U)^2\right)^{1/2}}.$$

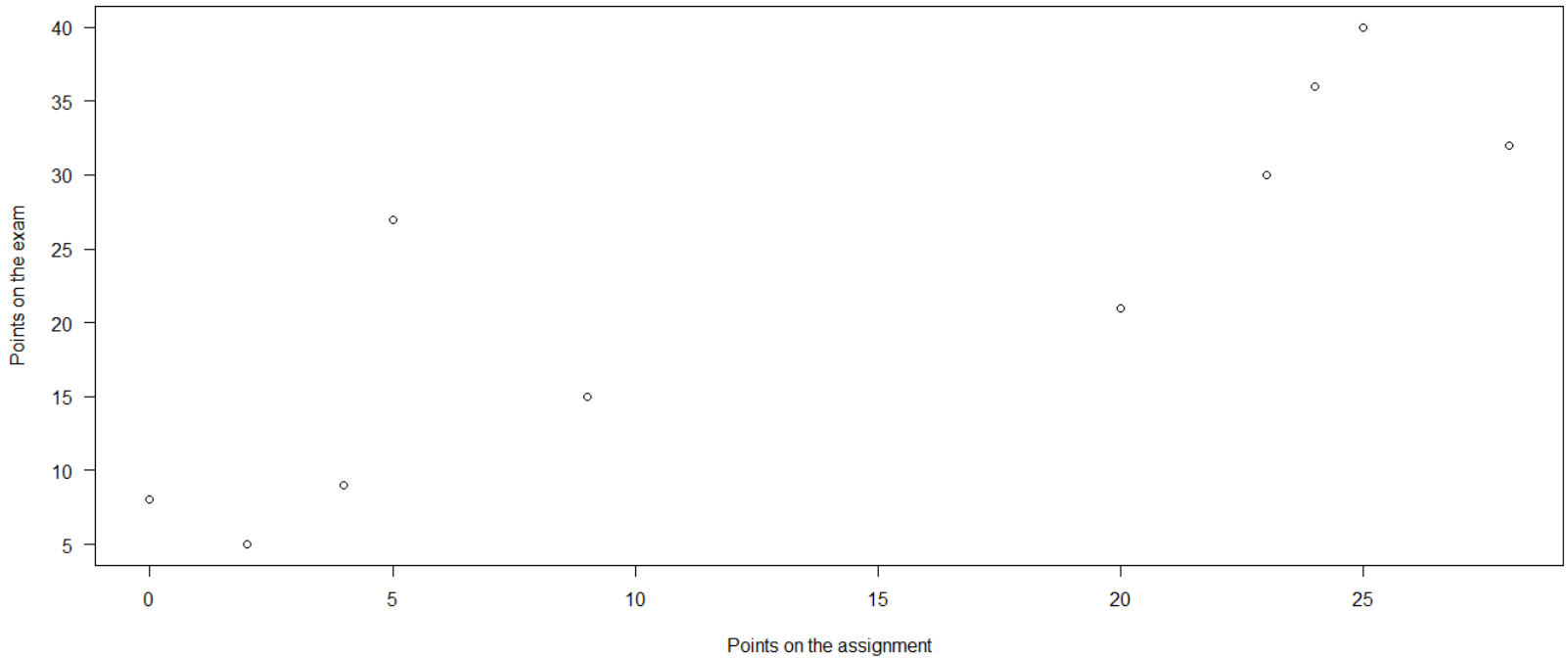
Example

Let U be the population of $N = 10$ students taking a Master course in statistics. Let x_i and y_i be, respectively, the scores in a home assignment and the final exam of the ith student $(i = 1, 2, \cdots, N)$. Table 1 shows the observed values.

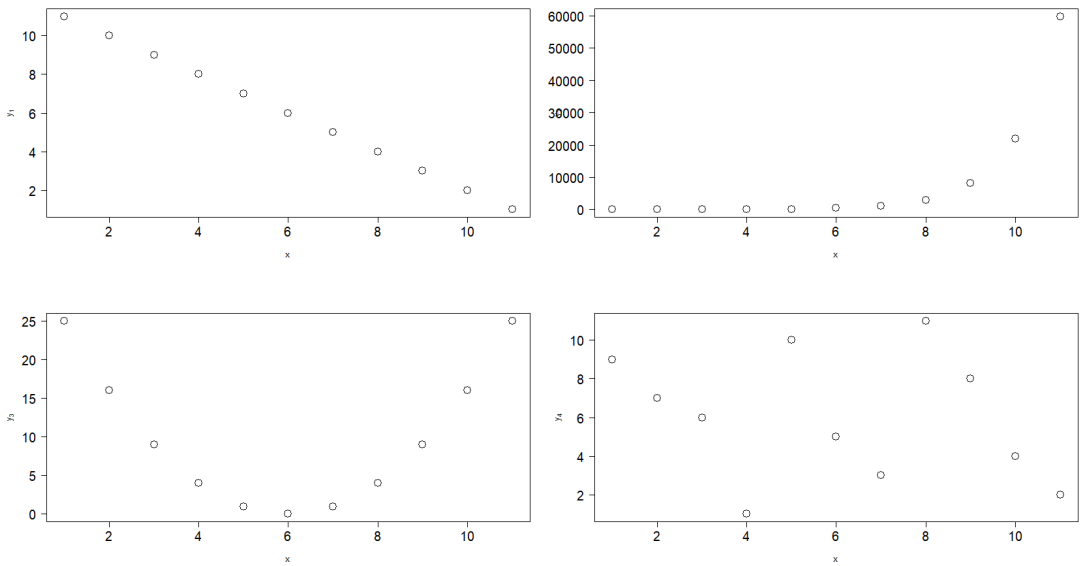
i	1	2	3	4	5	6	7	8	9	10
x_i	0	9	2	24	25	23	4	20	28	5
y_i	8	15	5	36	40	30	9	21	32	27

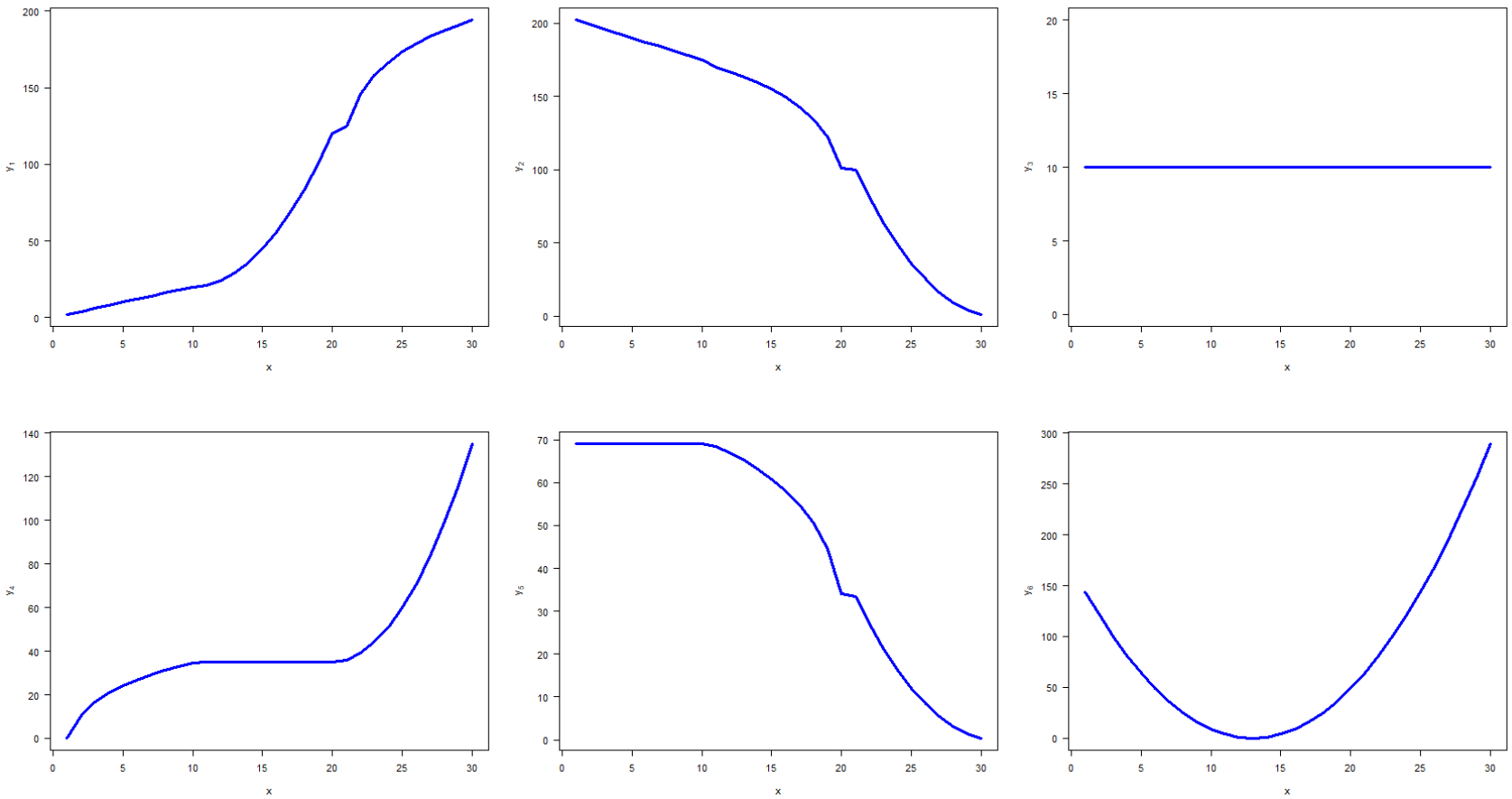
Table: Score of ten students in a home assignment and an exam in statistics

Find the correlation.


Points on the exam
40
35
30
25
20
15
10
5
0
5
10
15
20
25
Points on the assignment

x	y_1	y_2	y_3	y_4
1	11	3	25	9
2	10	7	16	7
3	9	20	9	6
4	8	55	4	1
5	7	148	1	10
6	6	403	0	5
7	5	1097	1	3
8	4	2981	4	11
9	3	8103	9	8
10	2	22026	16	4
11	1	59874	25	2




y1
x
y2
x
y3
x
y4
x
y5
x
y6
x

Definition (Spearman's correlation coefficient)

Let x_i and y_i be the values of two variables associated to the ith element in U $(i = 1, 2, \cdots, N)$. Let also $R(x_i)$ and $R(y_i)$ be their corresponding ranks. *Spearman's correlation coefficient* between x and y is defined as

$$r^s_{xy,U} \equiv \frac{\sum_U (R(x_i) - \bar{R}_U)(R(y_i) - \bar{R}_U)}{\left(\sum_U (R(x_i) - \bar{R}_U)^2 \sum_U (R(y_i) - \bar{R}_U)^2\right)^{1/2}}$$

where $\bar{R}_U = (N + 1)/2$.

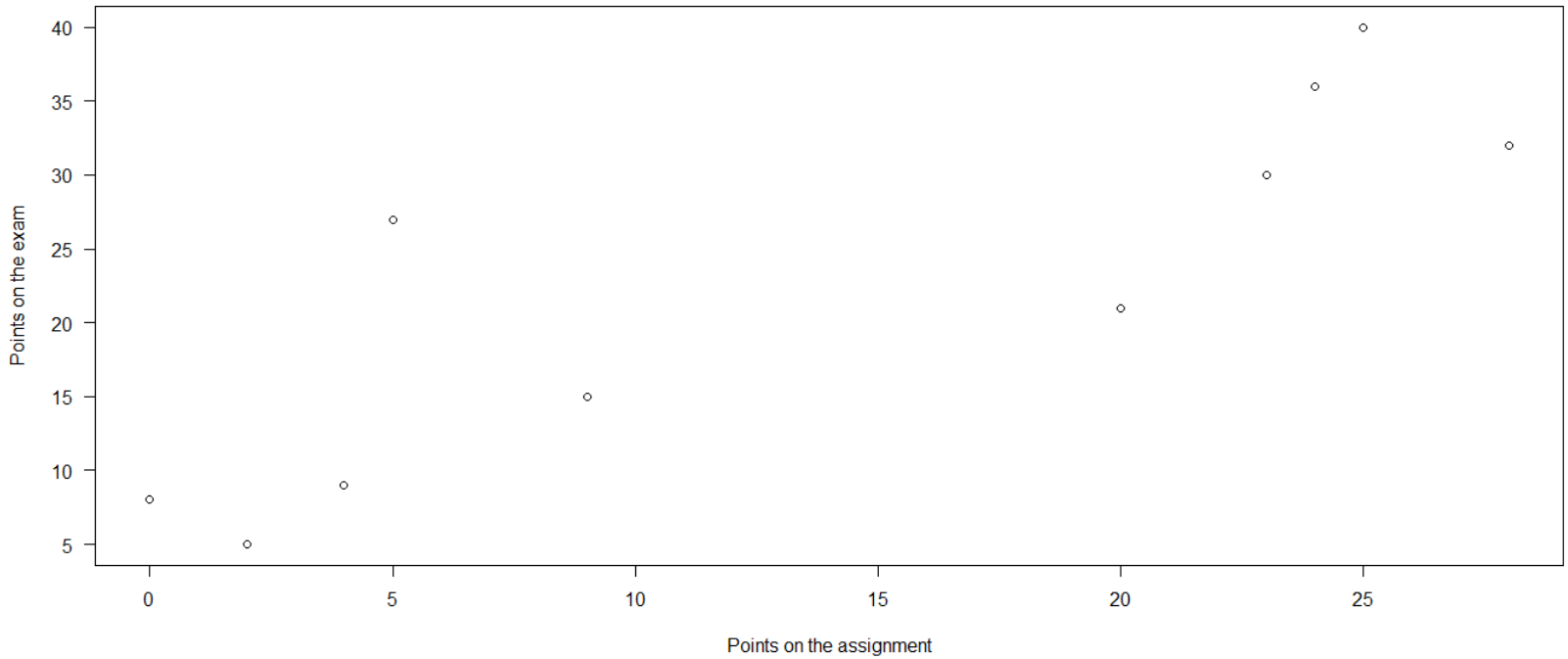
Example

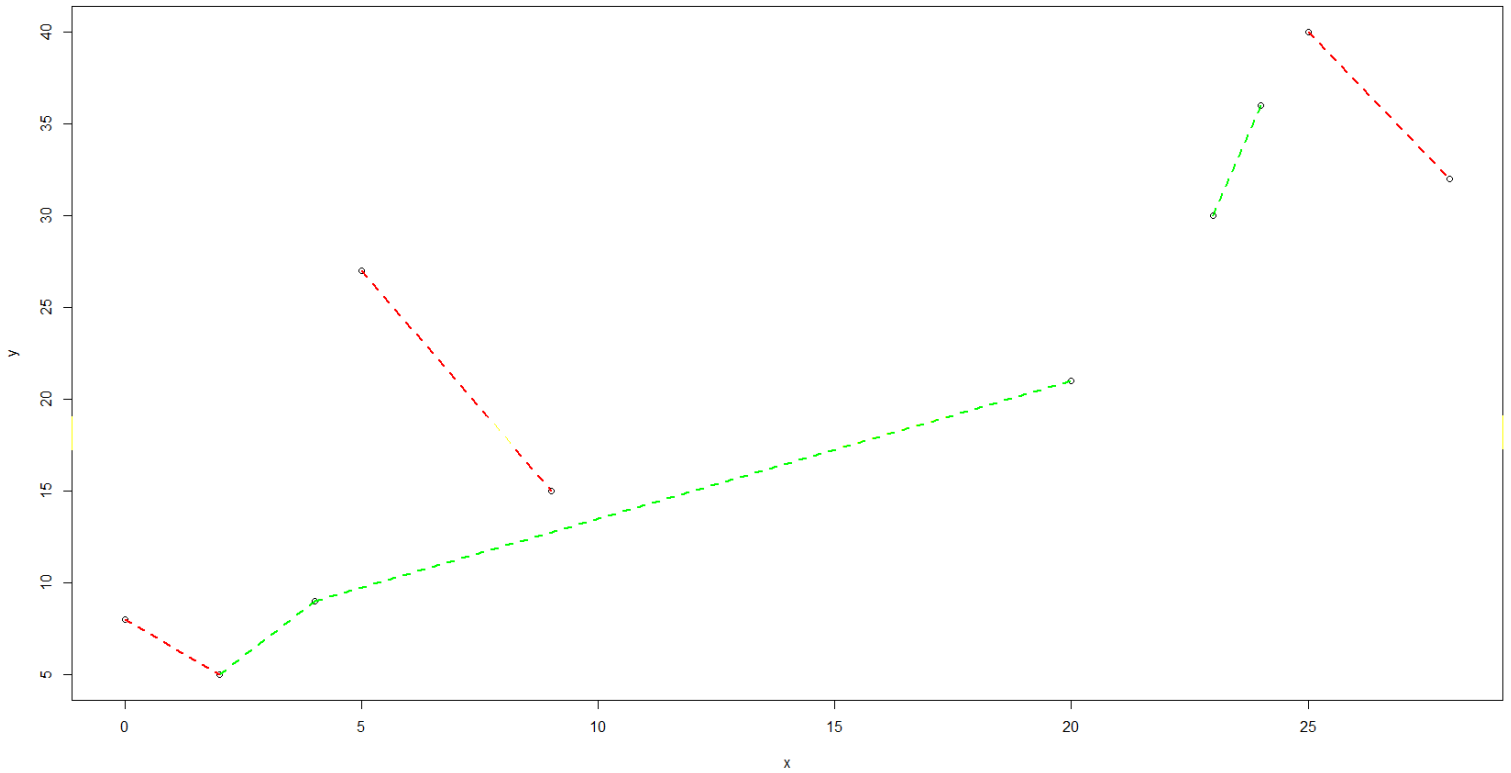
Let U be the population of $N = 10$ students taking a Master course in statistics. Let x_i and y_i be, respectively, the scores in a home assignment and the final exam of the ith student $(i = 1, 2, \cdots, N)$. Table 1 shows the observed values.

i	1	2	3	4	5	6	7	8	9	10
x_i	0	9	2	24	25	23	4	20	28	5
y_i	8	15	5	36	40	30	9	21	32	27

Table: Score of ten students in a home assignment and an exam in statistics

Find the Spearman's correlation coefficient.


Points on the exam
40
35
30
25
20
15
10
5
0
5
10
15
20
25
Points on the assignment


y
x
5
10
15
20
25
30
35
40
0
5
10
15
20
25

Definition (Kendall's correlation coefficient)

Let x_i and y_i be the values of two variables associated to the ith element in U $(i = 1, 2, \cdots, N)$. *Kendall's correlation coefficient* between x and y is defined as

$$r^k_{xy,U} \equiv \frac{2}{n(n-1)} \sum_{i<j} \text{sgn}(x_j - x_i)\text{sgn}(y_j - y_i) \tag{1}$$

where

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \qquad \square$$