

2. Realice la **estimación de una recta** de regresión para cada uno de los siguientes cuatro conjuntos de datos.

Calcule también los coeficientes de correlación respectivos.

Realice las **gráficas** que considere pertinentes.

¿Qué se puede **concluir** de este ejercicio?

i	X <sub>1</sub>	Y <sub>1</sub>	X <sub>2</sub>	Y <sub>2</sub>	X <sub>3</sub>	Y <sub>3</sub>	X <sub>4</sub>	Y <sub>4</sub>
1	10	8.04	10	9.14	10	7.46	8	6.58
2	8	6.95	8	8.14	8	6.77	8	5.76
3	13	7.58	13	8.74	13	12.74	8	7.71
4	9	8.81	9	8.77	9	7.11	8	8.84
5	11	8.33	11	9.26	11	7.81	8	8.47
6	14	9.96	14	8.10	14	8.84	8	7.04
7	6	7.24	6	6.13	6	6.08	8	5.25
8	4	4.26	4	3.10	4	5.39	19	12.50
9	12	10.84	12	9.13	12	8.15	8	5.56
10	7	4.82	7	7.26	7	6.42	8	7.91
11	5	5.68	5	4.74	5	5.73	8	6.89

Fuente: Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician* 27, pp. 17 – 21.

Coeficientes de correlación:

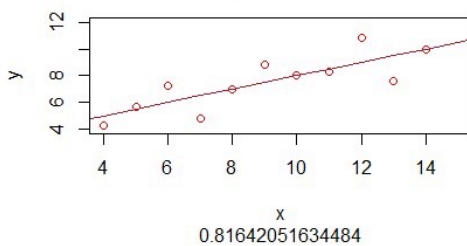
$$\text{cor}(X_1, Y_1) = 0.81642052$$

$$\text{cor}(X_2, Y_2) = 0.8162365$$

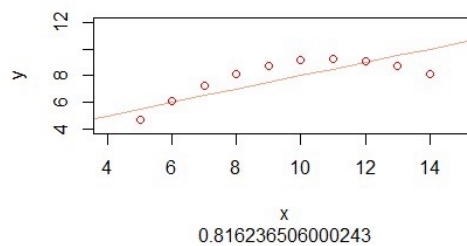
$$\text{cor}(X_3, Y_3) = 0.8162867$$

$$\text{cor}(X_4, Y_4) = 0.8165214$$

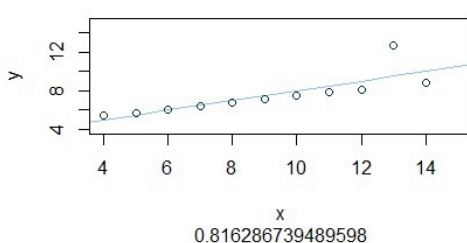
Regresión X1 Y1



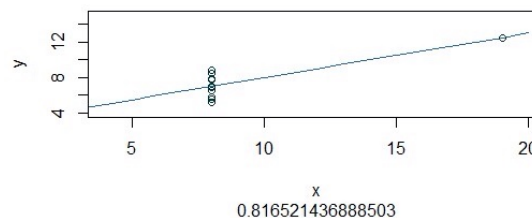
Regresión X2 Y2



Regresión X3 Y3



Regresión X4 Y4



Se puede concluir tanto los parámetros como sus correlaciones del modelo son casi iguales. Pero los puntos son distintos, por lo que esto nos enseña que puedes tener puntos diferentes y sin embargo, una correlación y parámetros parecidos. Diferentes datos pueden darnos una misma estimación de recta de regresión. Los datos de la gráfica 1 y 3 se adaptan a una MRLS, los de la 2 parecen ser más cuadráticos.

En la gráfica 4 los datos de  $x$  parece que no tiene variabilidad, más que por un punto el cual no nos ayuda para hacer un pronóstico sobre el comportamiento de  $y$  a diferentes valores de  $x$ .