Simple Linear Regression

Ejercicio 2: En una aula de instituto, fueron medidos y pesados todos los alumnos obteniendo los siguientes resultados:

	Alumno	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Σ
(x)	Altura (cm)	159	178	150	178	142	172	154	145	141	161	160	149	173	165	167	152	140	163	2849
(y)	Peso (kg)	63.64	83.12	42.5	80.95	37.09	70.85	70.85	48.21	26.11	40.91	47.75	42.46	71.19	57.71	67.03	43.71	39.12	61.43	994,6
	x^2	25.281	31.684	22.500	31.684	20.164	29.584	23.716	21.025	19.881	25.921	25.600	22.201	29.929	27.225	27.889	23.104	19.600	26.569	453.55
	y^2	4050,0496	6.908,9344	1.806,25	6552,9025	1375,6681	5019,7225	5019,7225	2324,2041	681,7321	1673,6281	2280,0625	1802,8516	5068,0161	3330,4441	4493,0209	1910,5641	1530,3744	3773,6449	59.601,7
	$x \cdot y$	10.118,76	14795,36	6375	14409,1	5266,78	12186,2	10.910,9	6990,45	3.681,51	6.586,51	7640	6.326,54	12315,87	9522,15	11194,01	6.643,92	5.476,8	10.013,09	160.452,

a) Obtén la ecuación de la recta (***y* = *mx* + *b***)

b) Realiza el diagrama de dispersión con la ecuación de la recta obtenida

c) Realiza estas 2 estimaciones según la ecuación de la recta obtenida:

- Un alumno que mida 168 cm, su peso estimado es: 66,4673 kg.
- Un alumno que pese 74.36 kg, su altura estimada es: 174,845 cm.

d) Obtén el *Error estándar de la estimación* (***Sxy***)

e) Obtén el *Coeficiente de correlación de Pearson* (***R***)

f) Obtén el *Coeficiente de determinación* (***R²***)

* Variable independiente (x) = Altura
* Variable dependiente (y) = Peso
* n = 18
* $\bar{x} = \frac{\Sigma x}{n} = \frac{2849}{18} = 158,2\hat{7}$
* $\bar{y} = \frac{\Sigma y}{n} = \frac{994,63}{18} = 55,257\hat{2}$

a)

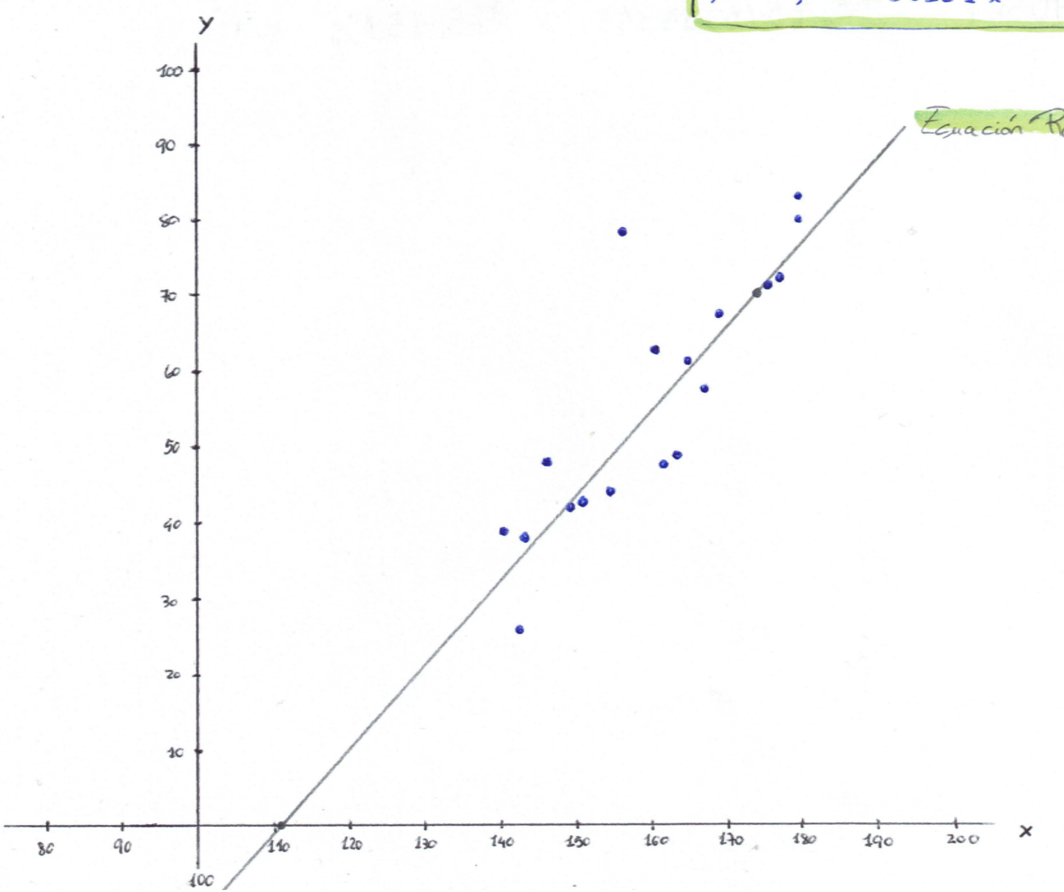
* Pendiente (m) $= \frac{\Sigma xy - n \cdot \bar{x} \cdot \bar{y}}{\Sigma x^2 - n \cdot \bar{x}^2} = \frac{160.452,95 - (18 \cdot 158,2\hat{7} \cdot 55,257\hat{2})}{453.557 - [18 \cdot (158,2\hat{7})^2]} = \boxed{1,153038221 = m}$

* Intercepto (b): $y = mx + b \rightarrow b = \bar{y} - m\bar{x} \rightarrow b = 55,257\hat{2} - 1,153038221 \cdot 158,2\hat{7} \rightarrow b = -127,24310$

* Ecuación de la recta : $y = mx + b \rightarrow \boxed{y = 1,153038221\,x - 127,2431051}$

$y=0 \rightarrow x = 110,3$

$y=70 \rightarrow x = 171,0$



c.1) Si $x = 168 \rightarrow y = 66,4673$ Kg

c.2) Si $y = 74,36 \rightarrow x = 174,845$ cm

- Error estándar de la estimación (S_{xy}) : $S_{xy} = \sqrt{\dfrac{\Sigma y^2 - b \cdot \Sigma y - m \cdot \Sigma xy}{n-2}}$

$\hookrightarrow S_{xy} = \sqrt{\dfrac{59.601,7925 - (-127,24531051 \cdot 994,63) - (1,153038221 \cdot 160.452,95)}{18 - 2}} =$

$= 8,489766241 = S_{xy}$

- Coeficiente de correlación para la recta de regresión (Coeficiente de Pearson) : r

$\hookrightarrow r = \dfrac{n \cdot \Sigma xy - (\Sigma x \cdot \Sigma y)}{\sqrt{[(n \cdot \Sigma x^2) - (\Sigma x)^2] \cdot [(n \cdot \Sigma y^2) \cdot (\Sigma y)^2]}} = \dfrac{18 \cdot 160452,95 - (2849 \cdot 994,63)}{\sqrt{[(18 \cdot 453.557) - (2849)^2] \cdot [(18 \cdot 59601,7925) - (994,63)^2}}$

$= 0,8669090584 = r$

- Coeficiente de determinación (r^2) :

$\hookrightarrow r^2 = 0,8669090584^2 = 0,7515313155 \simeq 75,15\ \% = r^2$