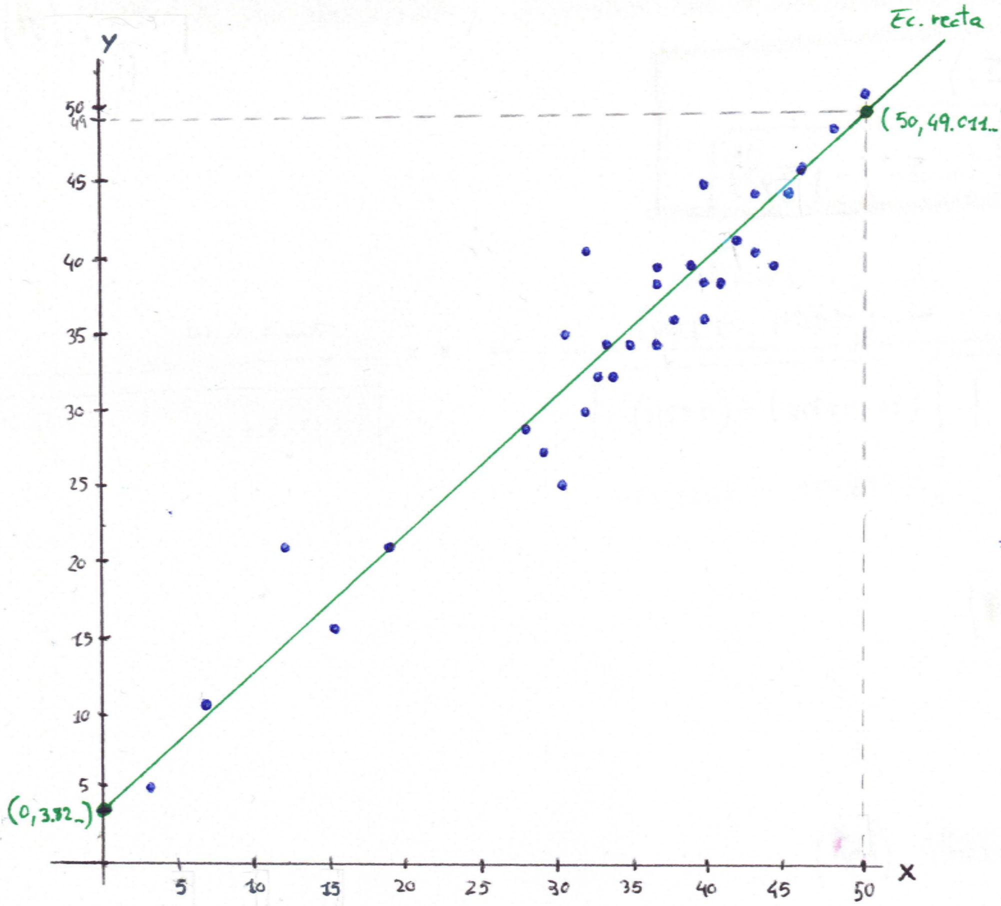
Simple Linear Regression

Ejercicio 1: Se desea realizar un estudio para predecir la reducción de la demanda de oxígeno en cuerpos de agua en función de la concentración de sólidos disueltos en el agua.

Reducción de sólidos (x)	3	7	11	15	18	27	29	30	30	31	31	32	33	33	34	36	36
Reducción de la demanda de oxígeno (y)	5	11	21	16	16	28	27	25	35	30	40	32	34	32	34	37	38
x^2	9	49	121	225	324	729	841	900	900	961	961	1024	1089	1089	1156	1296	1296
y^2	25	121	441	256	256	784	729	625	1225	900	1600	1024	1156	1024	1156	1369	1444
$x \cdot y$	15	77	231	240	288	756	783	750	1050	930	1240	1024	1122	1056	1156	1372	1368

																	Σ
Reducción de sólidos (x)	36	37	38	39	39	39	40	41	42	42	43	44	45	46	47	50	1.10
Reducción de la demanda de oxígeno (y)	34	36	38	37	36	45	39	41	40	44	37	44	46	46	49	51	1.124
x^2	1296	1369	1444	1521	1521	1521	1600	1681	1764	1764	1849	1936	2025	2116	2209	2500	41.086
y^2	1156	1296	1444	1369	1296	2025	1521	1681	1600	1936	1369	1936	2116	2116	2401	2601	41.998
$x \cdot y$	1224	1332	1444	1443	1404	1755	1560	1681	1680	1848	1591	1936	2070	2116	2303	2505	41.355



* Variable independiente (x): Reducción de sólidos

* Variable dependiente (y): Reducción de la demanda de oxígeno

* $n = 33$

* $\bar{x} = \frac{\Sigma x}{n} = \frac{1.104}{33} = \frac{368}{11}$

* $\bar{y} = \frac{\Sigma y}{n} = \frac{1.124}{33} = \frac{1124}{33}$

* Pendiente: $\boxed{m = \frac{\Sigma xy - n \cdot \bar{x} \cdot \bar{y}}{\Sigma x^2 - n \cdot \bar{x}^2}}$

$$m = \frac{41.355 - (33) \cdot \left(\frac{368}{11}\right) \cdot \left(\frac{1.124}{33}\right)}{41.086 - (33) \cdot \left(\frac{368}{11}\right)^2} = \frac{41.273/11}{45.674/11} = 0{,}903643\ldots$$

* Intercepto $\longrightarrow$ $\overbrace{y = b + mx}^{\text{Ecuación recta}}$ (b: intercepto) $\Rightarrow b = \bar{y} - m\bar{x} \rightarrow b = \frac{1124}{33} - 0{,}903643 \cdot \frac{368}{11} = 3{,}8296\ldots$

* Ecuación Recta: $y = b + mx \longrightarrow \boxed{y = 3{,}8296\ldots + 0{,}903643\ldots x}$

- Estimación $(x = 100)$

$$y = 3{,}8296\ldots + 0{,}903643\ldots \underbrace{(x)}_{}^{=100} \rightarrow y = 94{,}19395426$$

- Dibujar ecuación recta $\longrightarrow$ Si $x = 0 \rightarrow y = 3{,}8296\ldots$

Si $x = 50 \rightarrow y = 49{,}01179373$

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

$$S_{xy} = \sqrt{\frac{41.998 - (3{,}8296 \cdot 1124) - (0{,}9036 \cdot 41355)}{33 - 2}} \rightarrow S_{xy} = \sqrt{\frac{37.693{,}49229 - 37.370{,}16497}{31}}$$

$$S_{xy} = \sqrt{\frac{323{,}32\ldots}{31}} = \boxed{3{,}22953763 = S_{xy}}$$

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

$$\boxed{r = \frac{n \cdot \Sigma xy - (\Sigma x \cdot \Sigma y)}{\sqrt{\left[(n \cdot \Sigma x^2) - (\Sigma x)^2\right] \cdot \left[(n \cdot \Sigma y^2) - (\Sigma y)^2\right]}}}$$

$$r = \frac{\overbrace{33 \cdot 41355}^{1.364.715} - \overbrace{(1104 \cdot 1124)}^{1.240.896}}{\sqrt{\left[(33 \cdot 41.086) - (1104)^2\right] \cdot \left[(33 \cdot 41998) - (1124)^2\right]}} \rightarrow r = \frac{123.819}{129.588{,}357} \rightarrow$$

$$\sqrt{\left[1.355.838 - 1218.816\right] \cdot \left[1.385.934 - 1.263.376\right]}$$

$$\rightarrow \boxed{r = 0{,}9554793567}$$

- Coeficiente de determinación (r^2) :

$$0{,}9554793567^2 = 0{,}9129 \simeq \boxed{91{,}29\,\% = r^2}$$