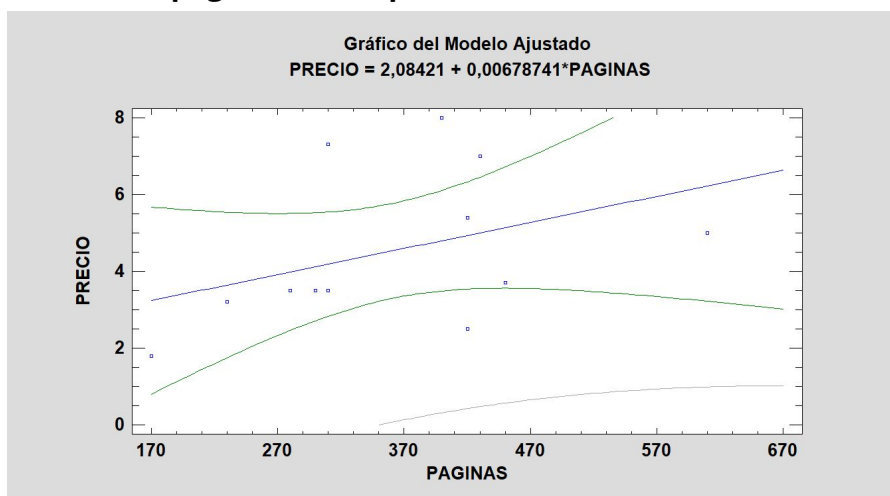


# Ejercicio 1

En la tabla adjunta se presentan el número de páginas y el precio de doce libros técnicos (...).

- a. Ajustar una recta de regresión que explique el precio en función del número de páginas e interpretar los resultados.



- b. Construir la tabla de ANOVA asociada, ¿es el ajuste adecuado?

## Análisis de Varianza

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	7,01592	1	7,01592	1,90	0,1985
Residuo	36,9907	10	3,69907		
Total (Corr.)	44,0067	11			

Como Valor-P = 0,1985 >  $\alpha = 0,05$  (por defecto) no existen evidencias suficientes como para rechazar la hipótesis nula, puesto que  $\beta = 0,00678741 \approx 0$ .

El estadístico R-Cuadrada indica que el modelo ajustado explica 15.9429 % de la variabilidad en PRECIO. El coeficiente de correlación es igual a 0.399285, indicando una relación relativamente débil entre las variables.

- c. Calcular intervalos de confianza al 90 % para los parámetros del modelo.

## Valores Predichos

	Pronosticado	Inferior 90%	Superior 90%	Inferior 90%	Superior 90%
X	Y	Límite Pred.	Límite Pred.	Límite Conf.	Límite Conf.
170,0	3,23807	-0,770672	7,24681	1,25857	5,21757
610,0	6,22453	1,96801	10,4811	3,78191	8,66715

Los intervalos serían [1.25857, 5.57] y [3.78191, 8.66715].

- d. Calcular un intervalo de confianza al 90 % para el precio de un libro de 500 páginas.

Opciones de Ventana > Meter una fila más

Valores Predichos					
	<i>Pronosticado</i>	<i>Inferior 90%</i>	<i>Superior 90%</i>	<i>Inferior 90%</i>	<i>Superior 90%</i>
X	Y	<i>Límite Pred.</i>	<i>Límite Pred.</i>	<i>Límite Conf.</i>	<i>Límite Conf.</i>
170,0	3,23807	-0,770672	7,24681	1,25857	5,21757
610,0	6,22453	1,96801	10,4811	3,78191	8,66715
500,0	5,47792	1,64261	9,31322	3,87855	7,07728

## Ejercicio 2

La resistencia del cemento depende del tiempo de secado del cemento. En un experimento se obtuvo la resistencia de bloques de cemento con diferente tiempo de secado de los resultados fueron los siguientes *Tiempo* (días) *Resistencia* (Kg / cm<sup>2</sup>) (...).

**a. Analiza la posible existencia de una relación entre estas dos variables.**

El estadístico R-Cuadrada indica que el modelo ajustado explica un 61,818 % de la variabilidad en Resistencia.

El coeficiente de correlación es igual a 0,786244, indicando una relación moderadamente fuerte entre las variables. El error estándar del estimado indica que la desviación estándar de los residuos es 5,73925.

**b. Dar las conclusiones que se deducen del contraste de regresión y linealidad.**

Contraste de regresión.

Como Valor-P =  $0 < \alpha = 0,05$  (por defecto) existen evidencias suficientes como para rechazar la hipótesis nula.

Por tanto, **existe regresión lineal.**

Contraste de linealidad.

Como Valor-P =  $0 < \alpha = 0,05$  (por defecto) existen evidencias suficientes como para rechazar la hipótesis nula de linealidad.

Por tanto, el modelo es skr.

**c. ¿Qué podemos decir si se utiliza un ajuste cuadrático? ¿Se obtienen mejores resultados?**

Si utilizamos «Cuadrado de Y», podemos observar que se dan mejores resultados con este modelo, ya que se explica un 72,76 % de variabilidad frente al 61,818 % que se explica con el modelo lineal.

## Ejercicio 3

La variable Y representa (en miles), el número de *Asnos* en España y la X el % del presupuesto del Estado dedicado a *Educación*:

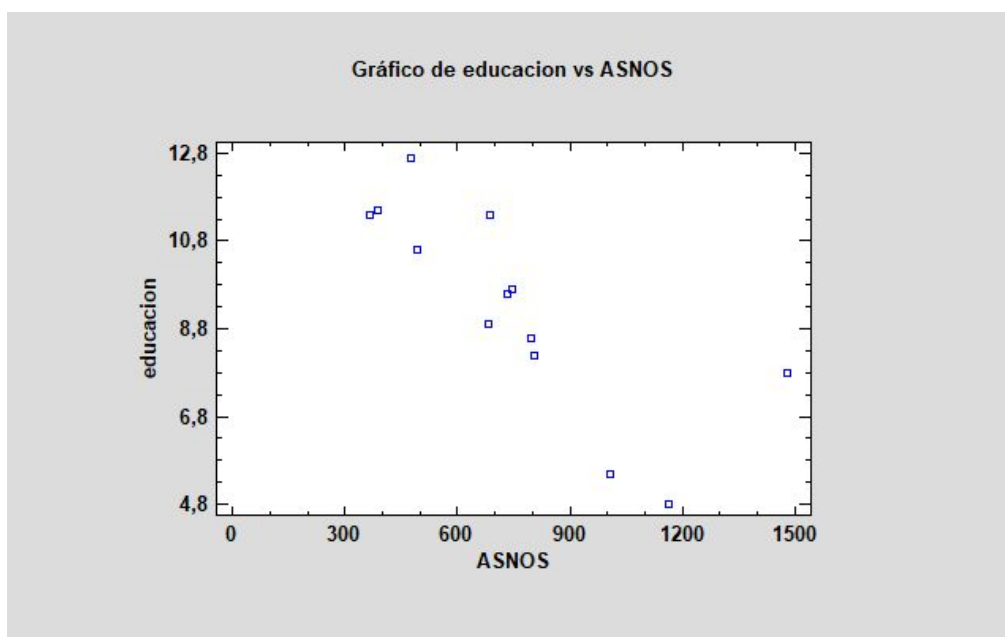
Año	Y	X	Año	Y	X	Año	Y	X
1920	1,006	5.5	1945	747	9.7	1970	476	12.7
1925	1,162	4.8	1950	732	9.6	1975	386	11.5
1930	1,479	7.8	1955	683	8.9	1980	368	11.4
1935	805	8.2	1960	686	11.4			
1940	795	8.6	1965	493	10.6			

Nota: Este problema es del libro de Daniel Peña. Entre estas variables existe una alta correlación estadística pero es claro que no existe relación entre ellas. La relación que hay entre ambas es debido a la que ambas tiene respecto a una tercera (el tiempo) y que no la tenemos en cuenta en el estudio. A este tipo de correlación se le denomina correlación espúrea.

### a) Representar gráficamente los datos.

Graficar > Gráficos de Dispersión > Gráficos X-Y

X = Asnos, Y = Educación



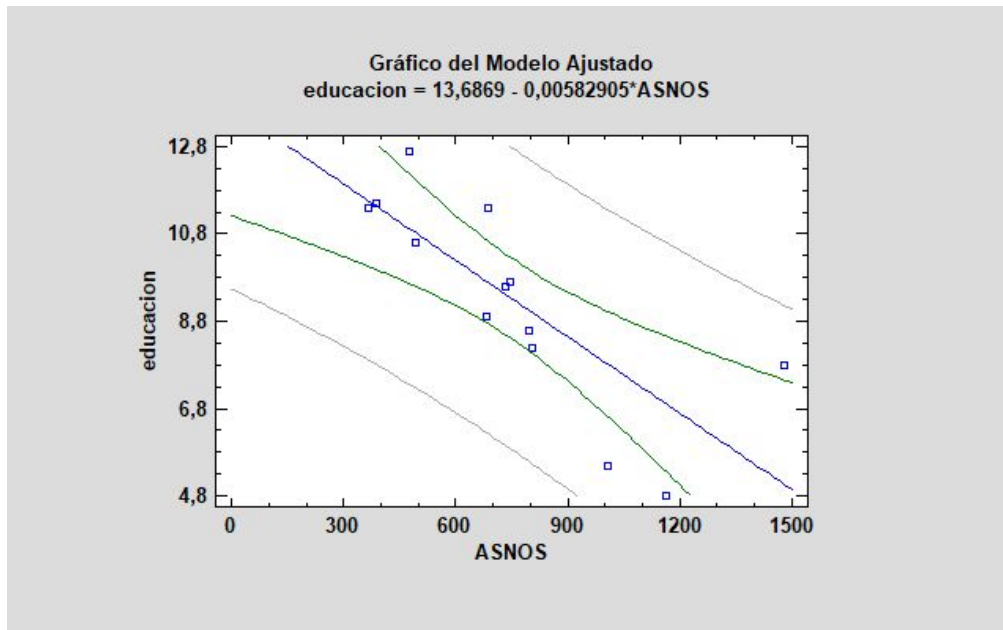
### b) Dar la recta de regresión que explique el tanto por ciento del presupuesto de Estado dedicado a Educación en función del número de asnos en España.

Relacionar > Un Factor > Regresión simple

X = Educación, Y = Asnos

La recta de regresión viene definida como:

$$\text{Educación} = 13,6869 - 0,00582905 \cdot \text{Asnos}$$

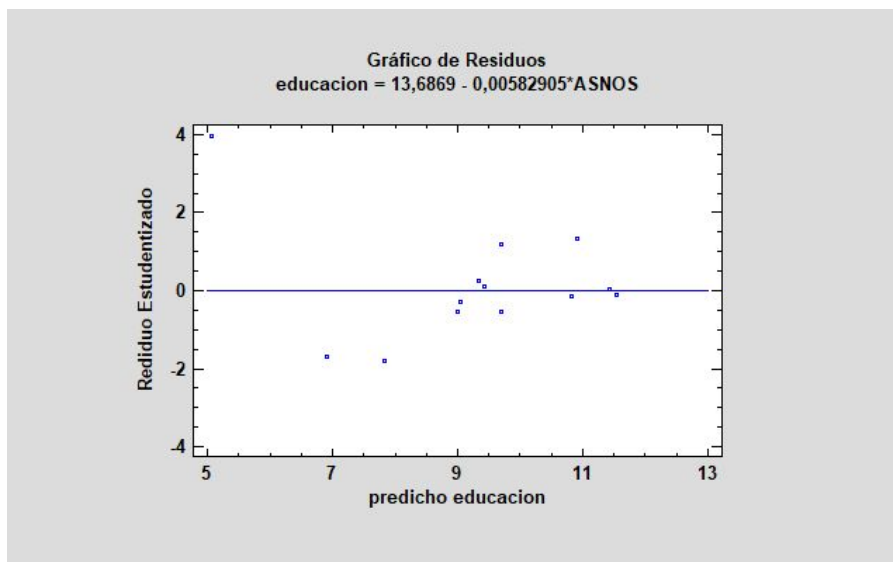
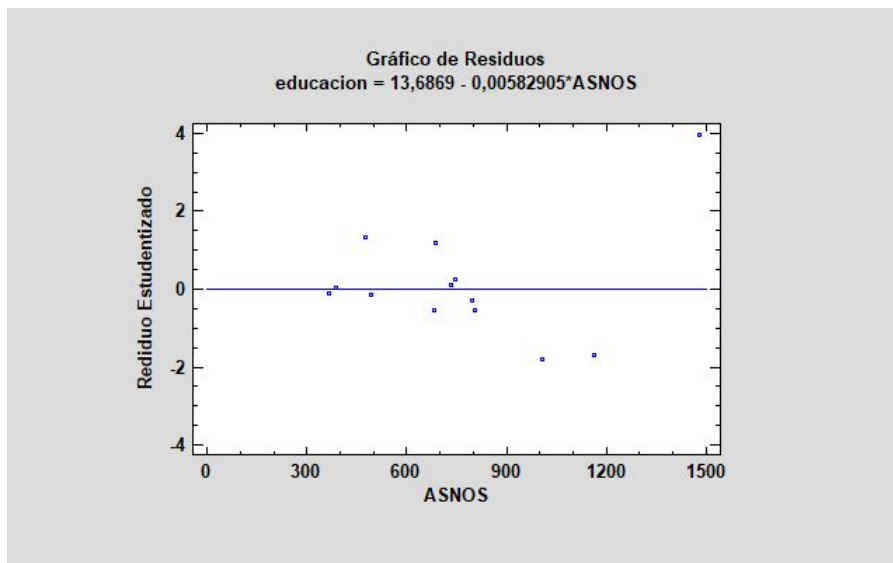


c) **¿Es significativo el coeficiente de correlación entre estas dos variables?**

Coeficiente de correlación = - 0,786205

Como el coeficiente de correlación de Pearson pertenece al intervalo  $[-1, 1]$  y es muy cercano a la correlación negativa perfecta (- 1), sí es significativo.

d) ¿Los residuos son independientes?

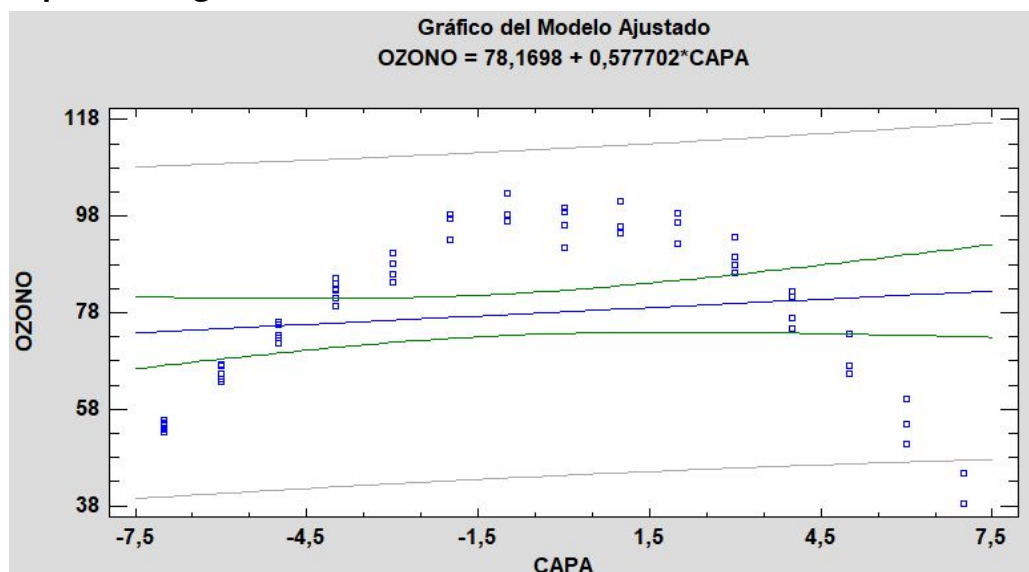


Puede observarse que ambos gráficos tienen un punto influyente menor a 3, pero aparte de ese punto influyente, no hay ninguna forma o estructura determinada, por lo que no se cumplirán las hipótesis del modelo.

## Ejercicio 4

Los siguientes datos corresponden a la cantidad de ozono registrado,  $Y$ , y su presión parcial para cada capa de altitud. Cada capa tiene aproximadamente un kilómetro de altura. Las capas se han escalado a un intervalo de -7 a +7 (...).

### a. Representar gráficamente los datos.



Los datos en este caso, son la parábola de puntos.

### b. Ajustar una función de regresión lineal del ozono frente a la capa.

$$OZONO = 78,1698 + 0,577702 \cdot CAPA$$

### c. Calcular la tabla del ANOVA y los contrastes de regresión y de linealidad.

Análisis de Varianza con Carencia de Ajuste					
Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	392,566	1	392,566	1,39	0,2424
Residuo	16608,3	59	281,496		
Carencia de Ajuste	16245,7	13	1249,67	158,56	0,0000
Error Puro	362,54	46	7,8813		
Total (Corr.)	17000,9	60			

#### Contraste de regresión

Como el P-Valor del modelo = 0,2424 > 0,05 no rechazamos la hipótesis nula, por lo que existen evidencias suficientes como para decir que no hay regresión lineal.

#### Contraste de linealidad

Como el P-Valor del ajuste es 0 < 0.05, rechazamos  $H_0$ , por lo que no existen evidencias suficientes como para decir que el modelo se ajusta a los datos.

**d. ¿Existe un modelo no lineal que mejore el ajuste lineal.**

Sí, el modelo que arroja el valor más alto es R-Cuadrada con 3,21173 %.  
Este es un 0,902639 % mayor que el modelo seleccionado.