ESTADÍSTICA

Práctica 6

Grupo 6 - Práctica con SPSS.
Alumno: Elvi Mihai Sabau Sabau.
DNI: 51254875L

6.6 Ejercicios.

1. Se está estudiando la relación entre el número de años que una persona está suscrita a una cierta publicación y el nivel de satisfacción con los contenidos de dicha publicación. Para ello se parte de los datos de 10 individuos tomados aleatoriamente de personas suscritas:

Años	8	7	10	3	6	13	4	9	10	5
Satisfacción (de 0 a 10)	7	5	8	5	9	9	3	8	8	7

Creamos las variables y añadimos los datos:

Nombre	Tipo	Anchura	Decimales	Etiqueta
Años	Numérico	2	0	X
Satisfaccion	Numérico	2	0	Y

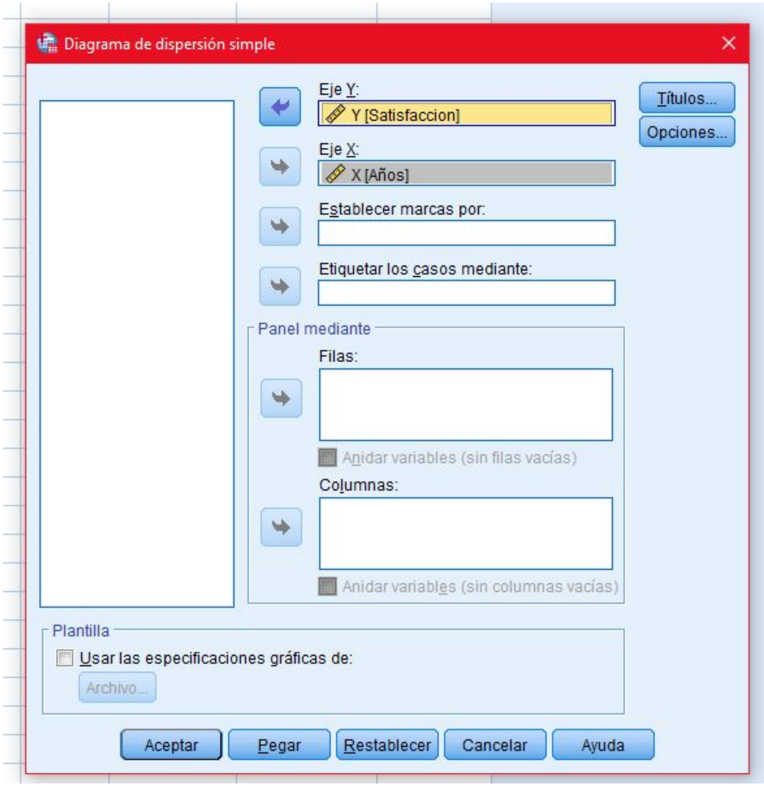
	Años	Satisfaccion
1	3	5
2	4	3
3	5	7
4	6	9
5	7	5
6	8	7
7	9	8
8	10	8
9	10	8
10	13	9

a) Construye un diagrama de dispersión para estos datos.

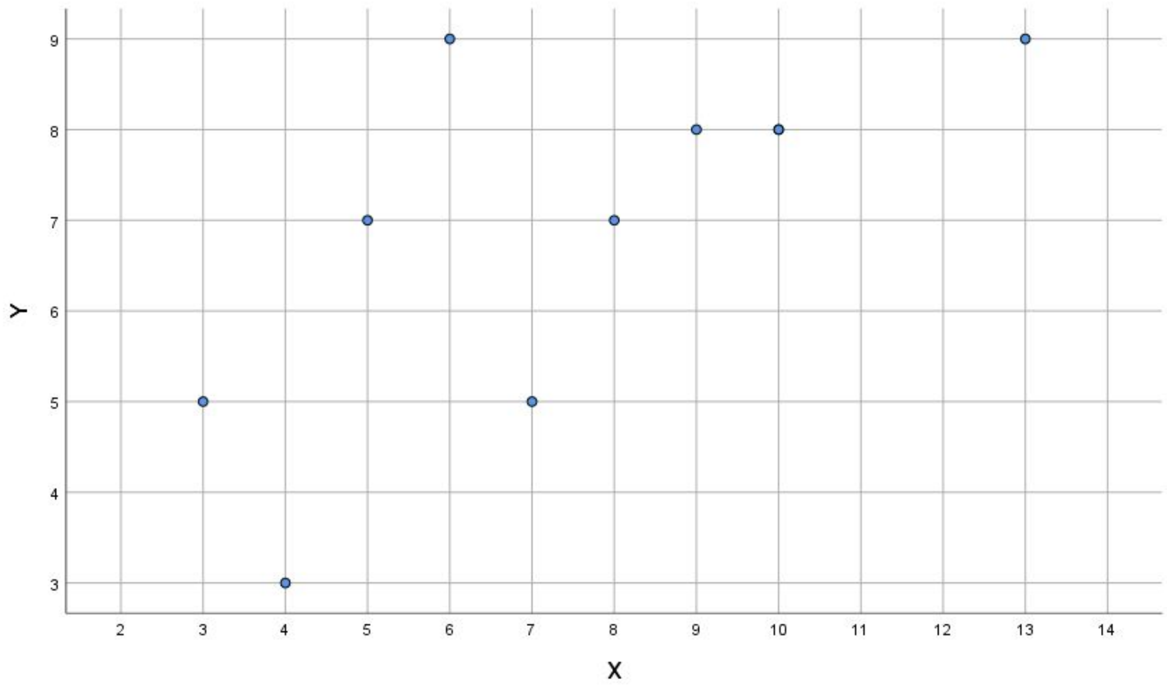
Para ello vamos a Gráficos > Cuadro de diálogo antiguos > Dispersión/Puntos

Tras lo cual nos aparecerá un pop up que nos dará a elegir entre los distintos tipos de diagramas existentes; sin embargo, en este caso únicamente emplearemos 'Dispersión simple'.

Continuamos definiendo qué variable será el eje Y y cual el eje X, en este caso, como solo queremos el dibujo, daría igual cual es cual, de todas maneras, para mantener consistencia, nuestra variable X (años) estará en el eje X, y la variable Y (Satisfacción) en el eje Y.



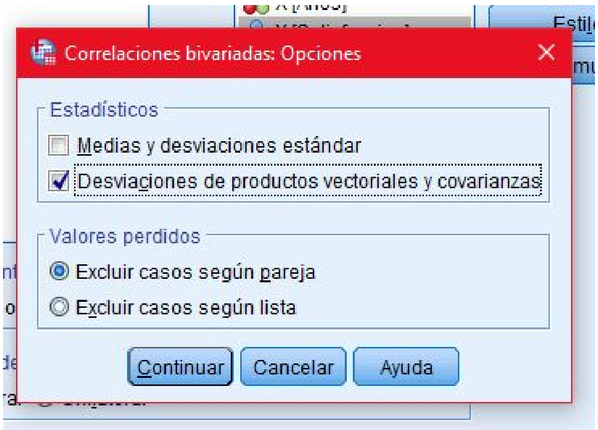
Le damos a “Aceptar”, y se nos generará el diagrama (el diagrama de dispersión que se muestra a continuación ha sido ajustado en ambos ejes).



b) Calcula el coeficiente de correlación y comenta el resultado obtenido

Para esto, vamos a ir a Analizar > Correlaciones > Bivariadas,

Una vez en la pestaña de correlaciones, seleccionamos ambas variables y nos aseguramos de que la casilla de “Pearson” esté marcada, además, vamos a ir al menú de opciones para marcar la casilla “Desviaciones de productos vectoriales y covarianzas.



Una vez hecho esto, le damos “Continuar” al popup, y a “Aceptar” a la ventana de correlaciones, ahora obtendremos los datos solicitados.

Correlaciones

		X	Y
X	Correlación de Pearson	1	,701*
	Sig. (bilateral)		,024
	Suma de cuadrados y productos vectoriales	86,500	38,500
	Covarianza	9,611	4,278
	N	10	10
Y	Correlación de Pearson	,701*	1
	Sig. (bilateral)	,024	
	Suma de cuadrados y productos vectoriales	38,500	34,900
	Covarianza	4,278	3,878
	N	10	10

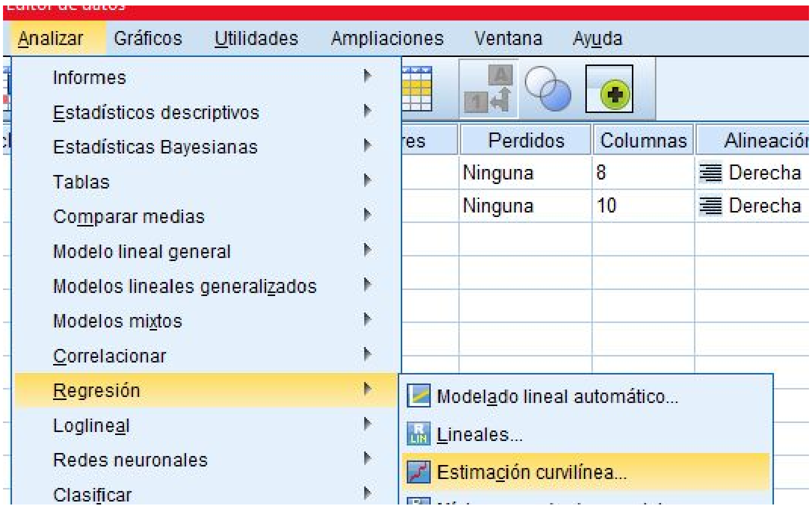
*. La correlación es significativa en el nivel 0,05 (bilateral).

Coeficiente de correlación.

El coeficiente de correlación es de 0.701, esto significa que nuestras variables de Años y Satisfacción tienen una relación bastante notable.

c) Predecir el índice de satisfacción de una persona que lleva 11 años suscrita a la publicación

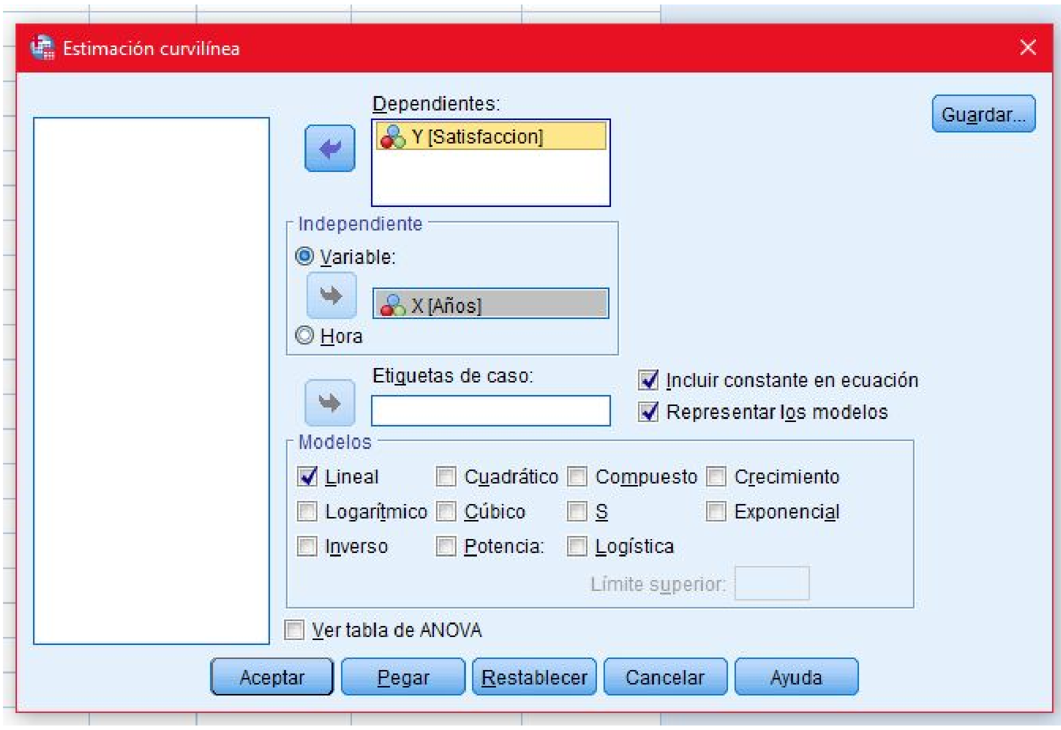
Para esto, tendremos que sacar las estimaciones, para ello, vamos a Analizar > Regresión > Estimación curvilínea.



Una vez aquí, como lo que nos pide es predecir la satisfacción (Nuestra variable Y) dado 11 años de subscripción (Nuestra variable X), lo que tendremos que hacer es la función de Y sobre X:

Y = a * X + b.

La variable dependiente es Y (grado de satisfacción), y la variable independiente será X (Años de subscripción), sabiendo esto, rellenamos los campos de la ventana de Estimación curvilínea, esto nos permitirá hallar nuestra “a” y “b”.



Resumen de modelo y estimaciones de parámetro

Variable dependiente: Y

	Resumen del modelo					Estimaciones de parámetro	
Ecuación	R cuadrado	F	gl1	gl2	Sig.	Constante	b1
Lineal	,491	7,717	1	8	,024	3,562	,445
						b	a

La variable independiente es X.

Nuestra a = 0,445 y nuestra b = 3,562.

Entonces: Y = 0,445 * 11 + 3,562 = 8,457 Índice de satisfacción.

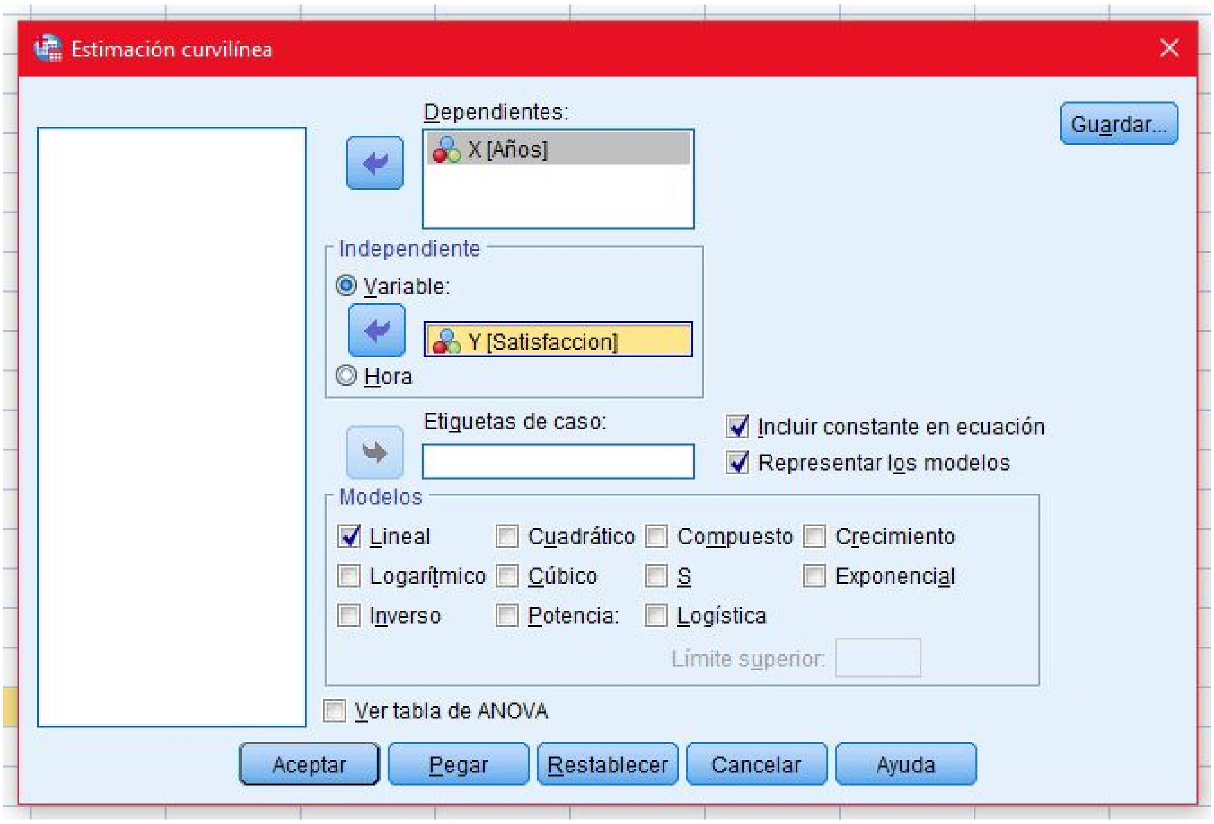
d) Conociendo que el índice de satisfacción es de 6 predecir, los años que lleva suscrita a la publicación

El proceso es similar al apartado anterior, en este caso lo que se diferencia es que nos pide la función de X sobre Y:

X = a * Y + b.

Para hallar nuestras a y b, tendremos que realizar la misma operación, pero esta vez X (Años) será la variable dependiente, e Y (Grado de satisfacción) la variable independiente.

Volvemos a Analizar > Regresión > Estimación curvilínea, y rellenamos, y le damos a aceptar.



Resumen de modelo y estimaciones de parámetro

Variable dependiente: X

Ecuación	Resumen del modelo					Estimaciones de parámetro	
	R cuadrado	F	gl1	gl2	Sig.	Constante	b1
Lineal	,491	7,717	1	8	,024	-,112	1,103
						b	a

La variable independiente es Y.

Ahora que ya tenemos los datos, podemos operar.
X = -0,112 * 6 + 1,103 = 0,431 Años.

2. En la tabla adjunta se recogen el gasto mensual en publicidad (X) y las ventas mensuales (Y) de una empresa (ambas en miles de euros):

X	15,2	14,9	15	14,9	14,2	14,6	15,5	15,1	15,4	14,7	14,3	15,7
Y	715	705	704	715	654	698	758	708	714	703	676	771

Declaramos las variables, e insertamos los datos:

Nombre	Tipo	Anchura	Decimales	Etiqueta	
Gasto	Numérico	4	1	X	Ni
Ventas	Numérico	4	0	Y	Ni

	Gasto	Ventas
1	14,2	654
2	14,3	878
3	14,6	698
4	14,7	703
5	14,9	705
6	14,9	715
7	15,0	704
8	15,1	708
9	15,2	715
10	15,4	714
11	15,5	758
12	15,7	771

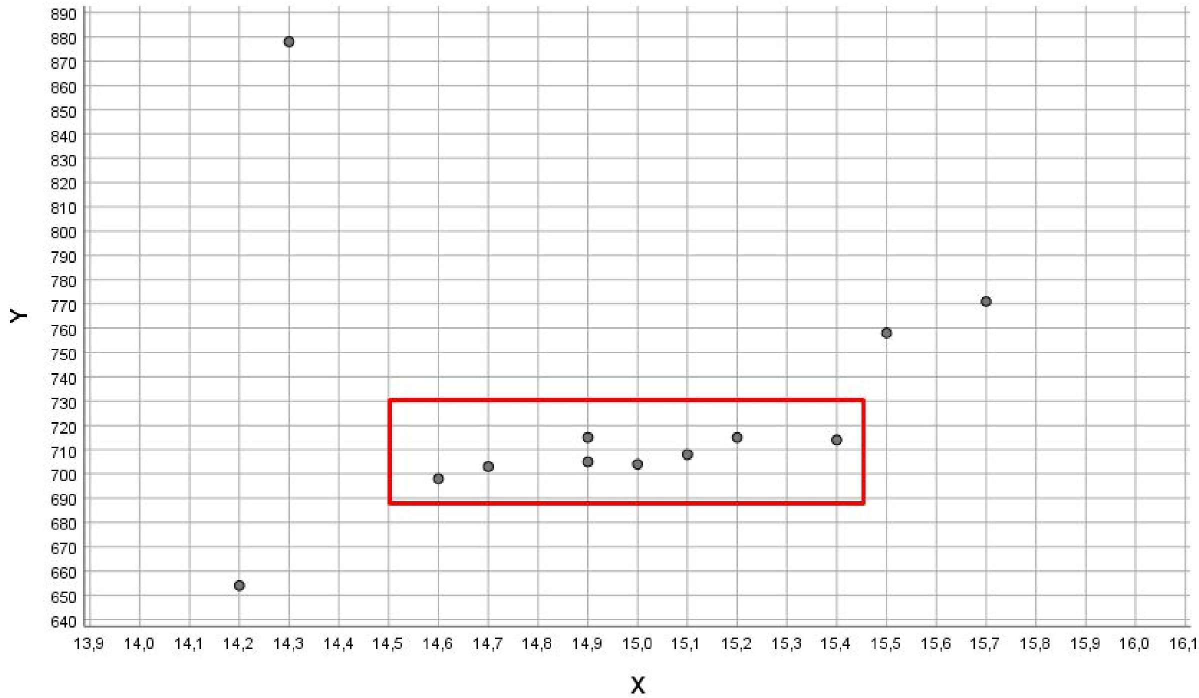
a) Construye el gráfico de dispersión (nube de puntos) de los datos.

Para ello vamos a Gráficos > Cuadro de diálogo antiguos > Dispersión/Puntos

Tras lo cual nos aparecerá un pop up que nos dará a elegir entre los distintos tipos de diagramas existentes; sin embargo, en este caso únicamente emplearemos 'Dispersión simple'.

Insertamos las variables

Y ya tendremos el gráfico.



¿Parece plausible ajustar una recta de regresión?

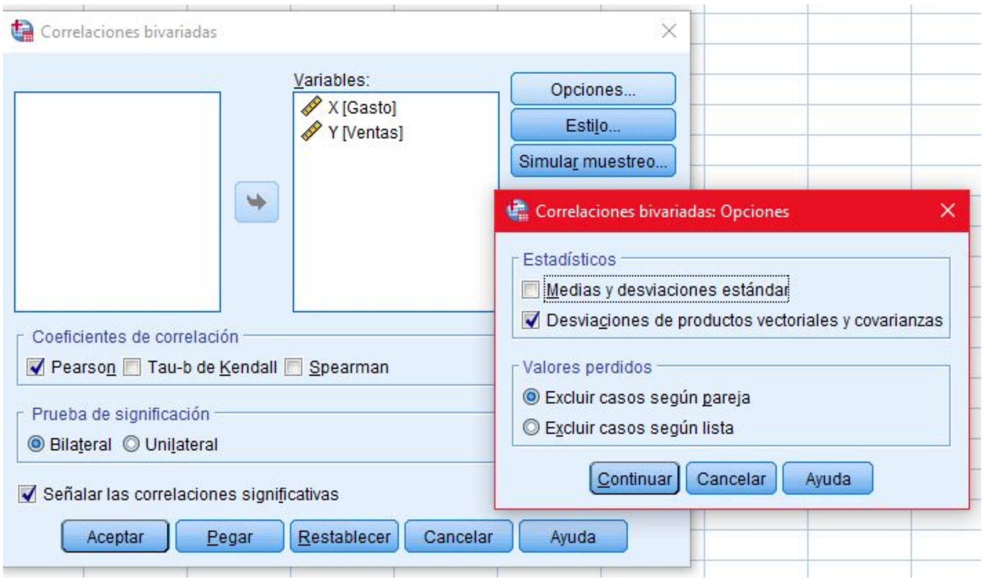
Difícilmente, **no**, parece que hay cierta relación entre las variables, pero es minúscula, casi imperceptible. A ojo diría que la recta de regresión se ajustaría a una línea trazada horizontalmente usando como referencia los puntos centrales, pero, **no sería adecuada para representar todos los datos**.

¿Cómo debe salir el coeficiente de correlación? Razona la respuesta.

Viendo el gráfico de dispersión, y como he mencionado anteriormente, **el coeficiente será extremadamente minúsculo**, ya que parece que apenas tiene relación una con otra.

b) Calcula la covarianza existente entre ambas variables así como el coeficiente de correlación.

Para esto vamos a Analizar > Correlación > Bivariadas, y seleccionamos las variables, y chequeamos la casilla de “desviaciones de productos…”



Así obtendremos los datos:

Correlaciones

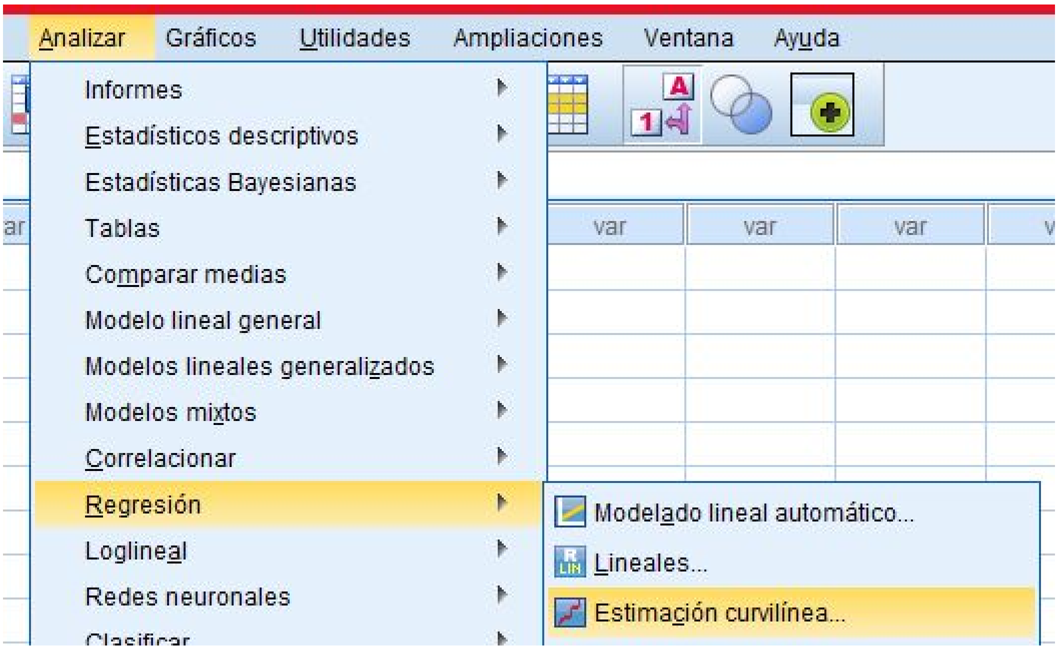
		X	Y
X	Correlación de Pearson	1	,041
	Sig. (bilateral)		,899
	Suma de cuadrados y productos vectoriales	2,329	11,658
	Covarianza	,212	1,060
	N	12	12
Y	Correlación de Pearson	,041	1
	Sig. (bilateral)	,899	
	Suma de cuadrados y productos vectoriales	11,658	34274,917
	Covarianza	1,060	3115,902
	N	12	12

Covarianza Coeficiente

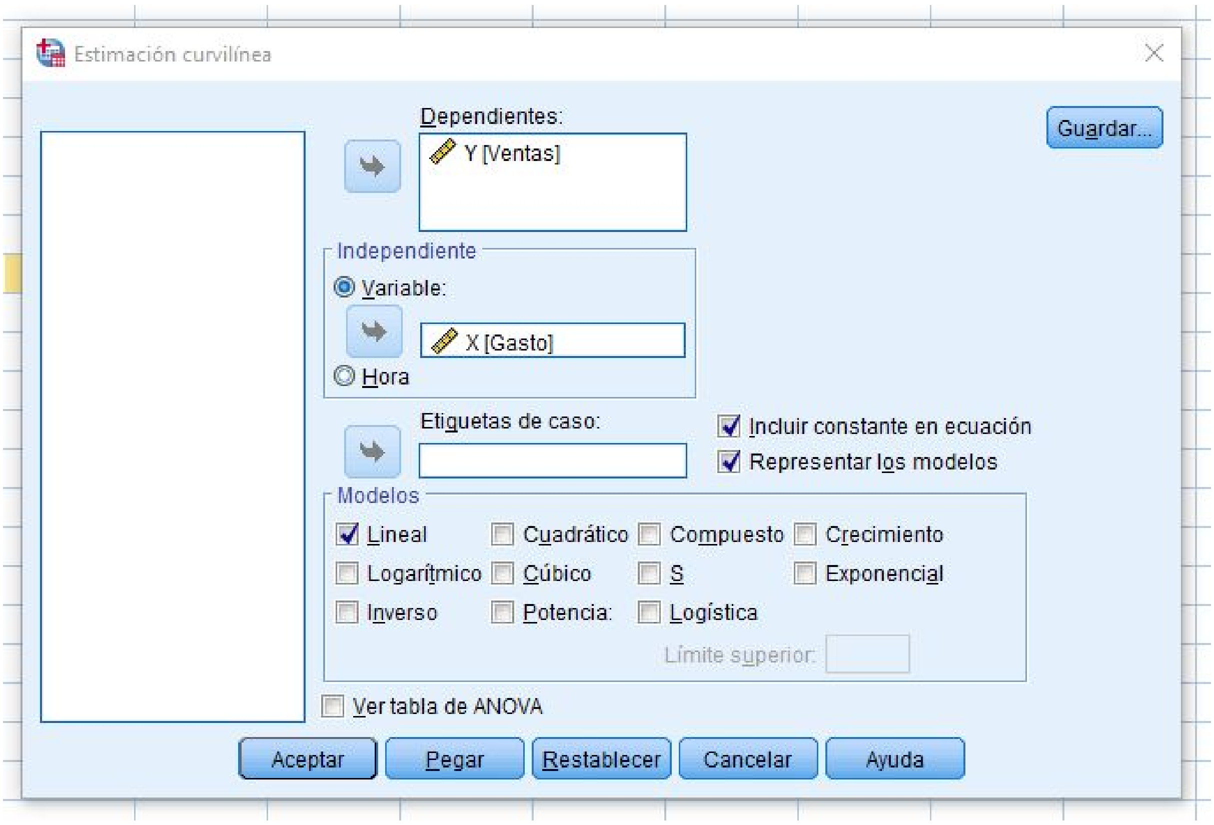
La covarianza entre las variables es de 1,060, y el coeficiente es de 0,041, teniendo una dependencia lineal extremadamente pobre, los gastos a penas dependen de las ventas y viceversa.

c) Calcula la recta de regresión que explique las ventas en función del gasto en publicidad.

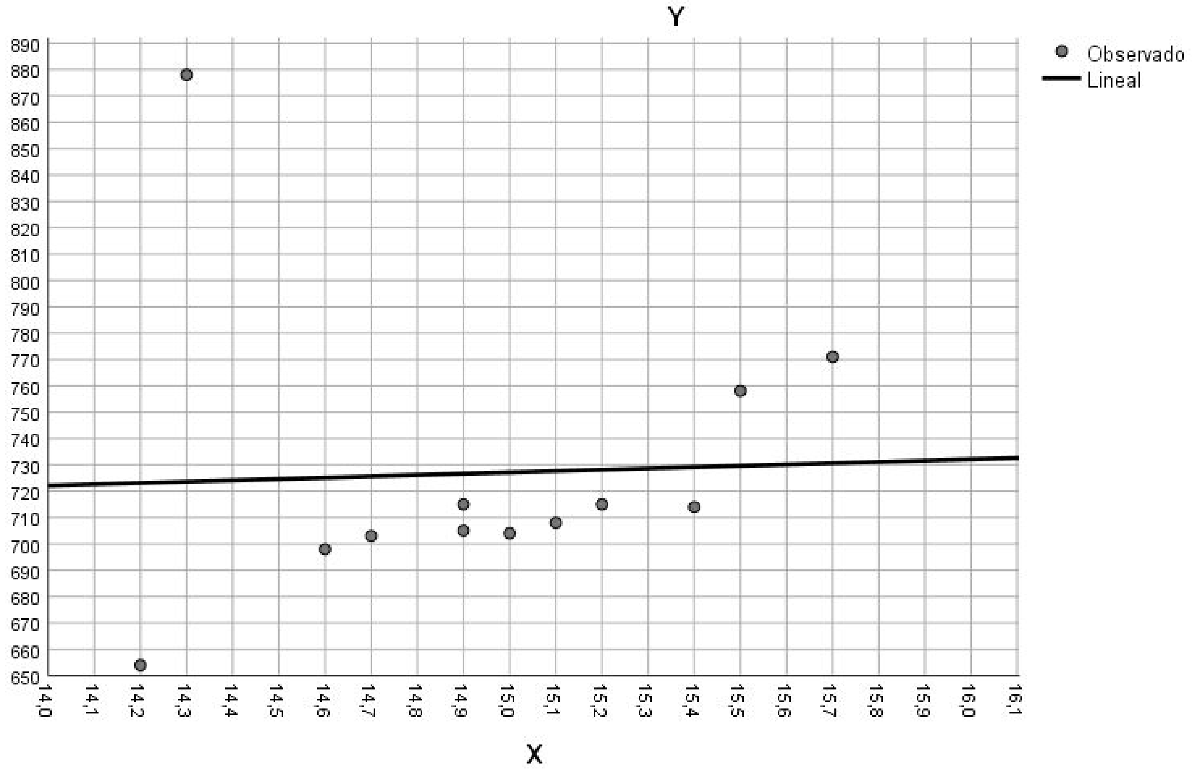
Para esto, vamos a Analizar > Regresión > Estimación curvilínea.



Como lo que nos pide es la recta representando las ventas (Y) en función del gasto (X), la variable dependiente será Y y la independiente X.



Obtenemos el gráfico de la recta:



d) ¿Qué ventas cabe esperar para un gasto en publicidad de 16000€?

Calculado del apartado anterior, Y siendo las ventas, y X los gastos, nos pide Y en función de X.

Y = a * X + b.

Obtenemos a y b del apartado anterior.

Resumen de modelo y estimaciones de parámetro

Variable dependiente: Y

	Resumen del modelo					Estimaciones de parámetro	
Ecuación	R cuadrado	F	gl1	gl2	Sig.	Constante	b1
Lineal	,002	,017	1	10	,899	$\boxed{652,045}$	$\boxed{5,005}$
						b	a

La variable independiente es X.

Y hacemos el cálculo:

16000 eur, los datos están representados en miles de euros, X = 16.

Y = 5 * 16 + 652 = 732 miles de euros en ventas.