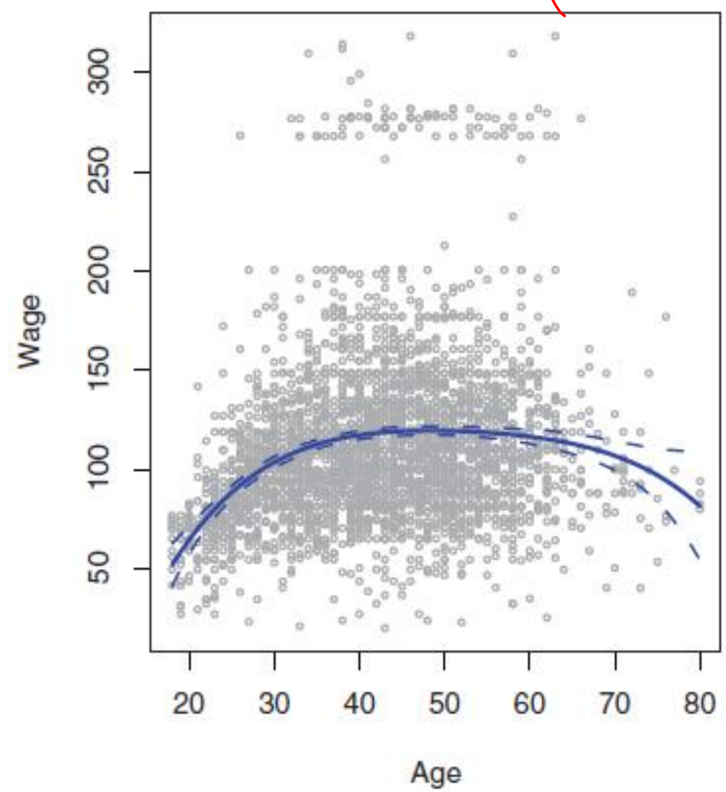
Regresión No lineal

Regresión Polinomial

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \ldots + \beta_d x_i^d + \epsilon_i \quad \textbf{(1)}$$

Los coeficientes de (1) son calculados mediante mínimos cuadrados pues es un modelo lineal con p $x_i, x_i^2, x_i^3, \ldots, x_i^d$

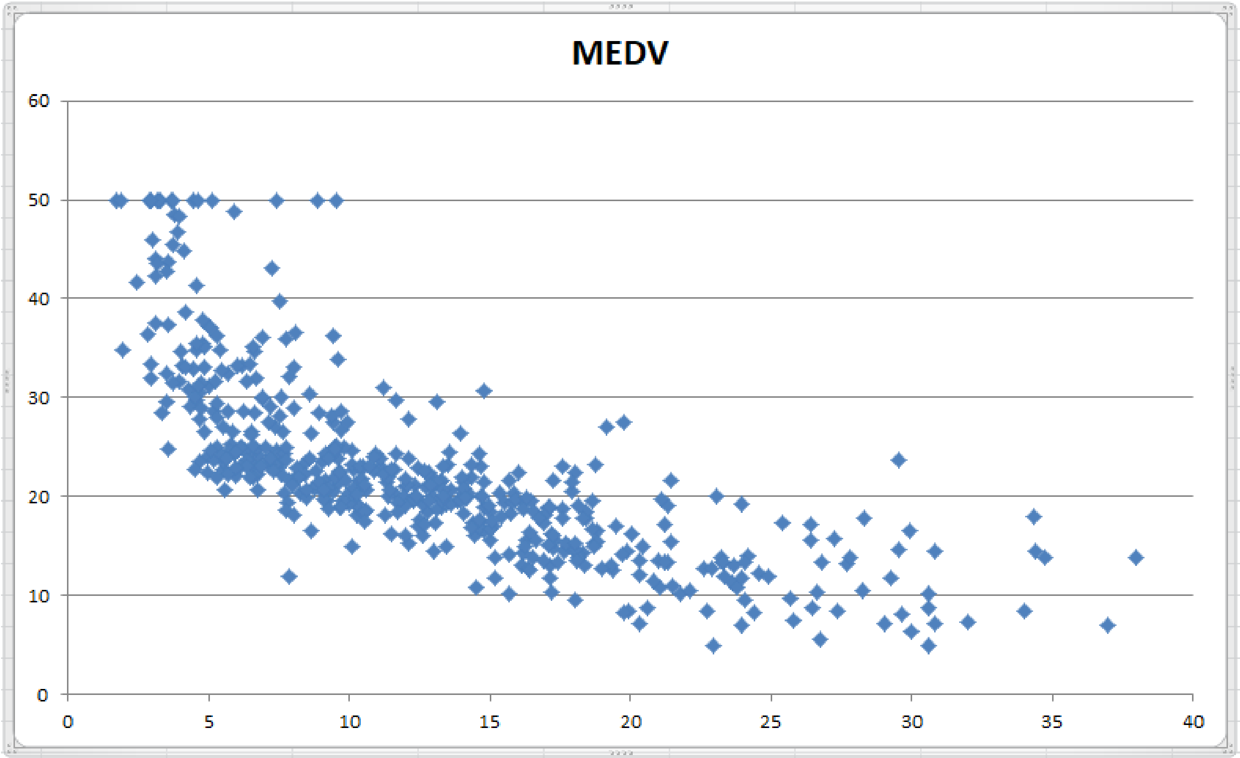
$$\hat{f}(x_0) = \hat{\beta}_0 + \hat{\beta}_1 x_0 + \hat{\beta}_2 x_0^2 + \hat{\beta}_3 x_0^3 + \hat{\beta}_4 x_0^4$$



Regresión utilizando otras ecuaciones No Lineales

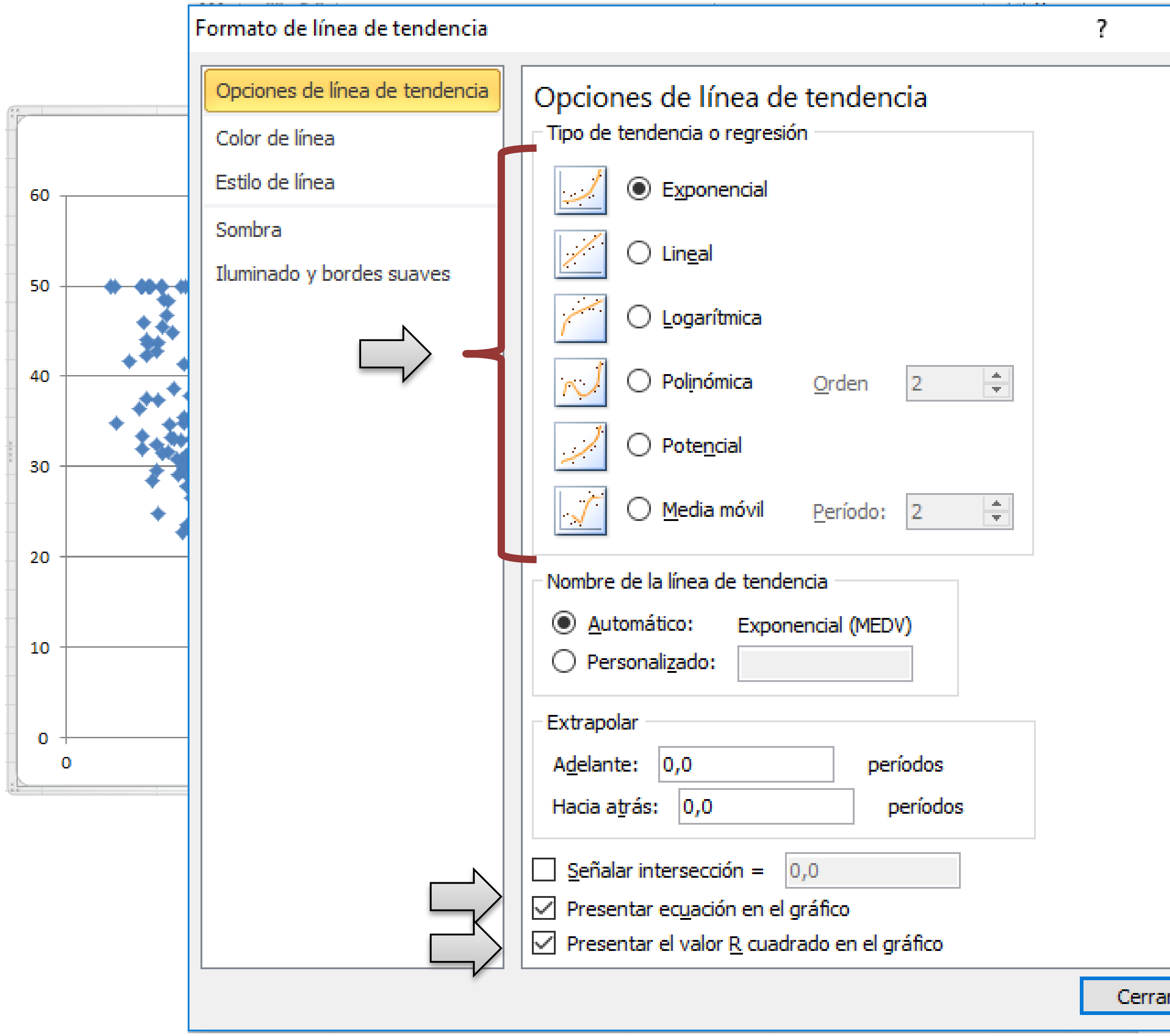
En Excel

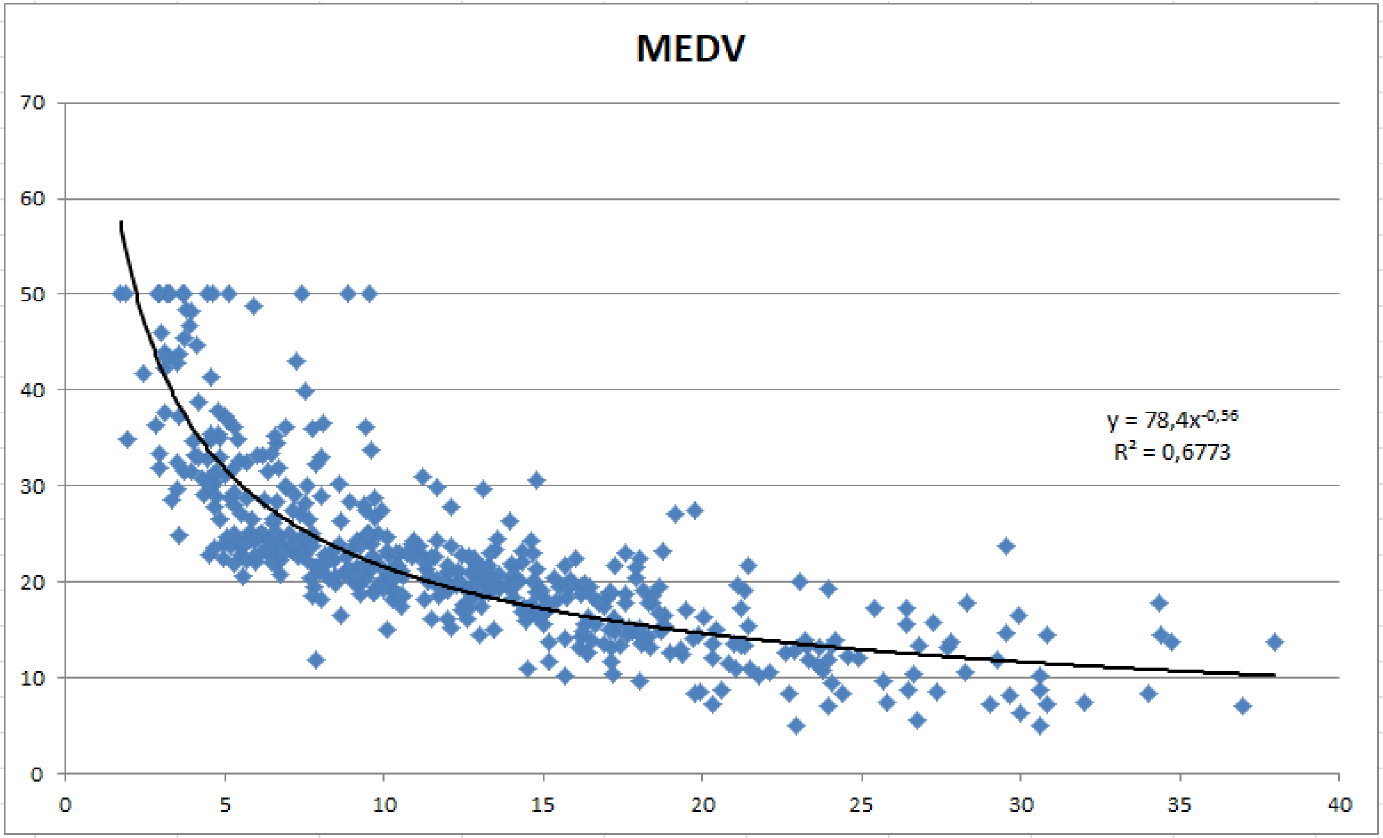
1) Seleccionar todos los puntos del gráfica pulsando el mouse sobre uno de ellos.
2) Pulsar botón derecho y seleccionar "Agregar línea de tendencia"



En Excel

3) Seleccionar la ecuación de ajuste deseada (exponencial, polinómica, logarítmica o potencial). Tildar "Presentar ecuación el gráfico" y "Presentar el valor R cuadrado en el gráfico"





Esta es la ecuación de Regresión No Lineal Potencial : **$y = 78{,}4x^{-0{,}56}$ (2)**
Y este es su $R^2 = 0{,}6773$

¿Pero no habíamos dicho que el R2 indicaba el grado de dependencia LINEAL entre las dos variables? Sí, esto es correcto. Lo que se hace aquí es realizar transformaciones de las variables para que la forma final sí sea lineal. Veamos el ejemplo con ecuación POTENCIAL cuyo R2 es 0,6773:

Asumimos que la función que relaciona a ***x*** con ***y*** es no lineal, siendo su forma correspondiente una función “potencial”. La ecuación que la representa es:

$$y = ax^b \qquad (3)$$

Estos coeficientes *a* y *b* no son los β_0 y β_1 , si no que serán funciones de ellos (ya lo veremos a continuación)

Si aplicamos logaritmos a cada término de la ecuación (3) obtendremos lo siguiente:

$$\log(y) = \log(a) + b.\log x$$

Y como el formato final tiene que ser lineal del tipo $P = \beta_0 + \beta_1 z \Rightarrow P = \log(y)$

Y como

$$\log(a) + b.\log x = \beta_0 + \beta_1 z$$

$$\Rightarrow z = \log(x)\ ,\ \beta_0 = \log(a)\ \text{ y }\ \beta_1 = b$$

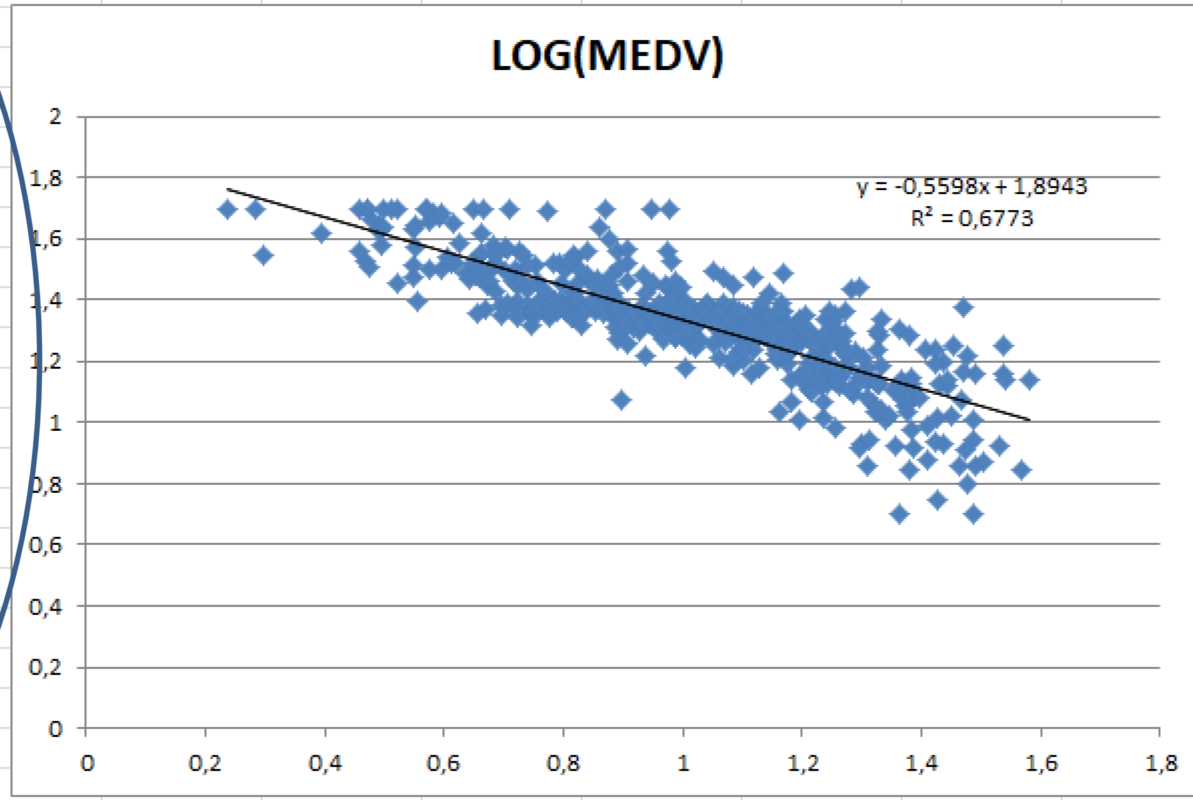
Entonces los valores finales de a y b de la ecuación (1) serán:

$$a = 10^{\beta 0} \qquad (4)$$
$$b = \beta_1 \qquad (5)$$

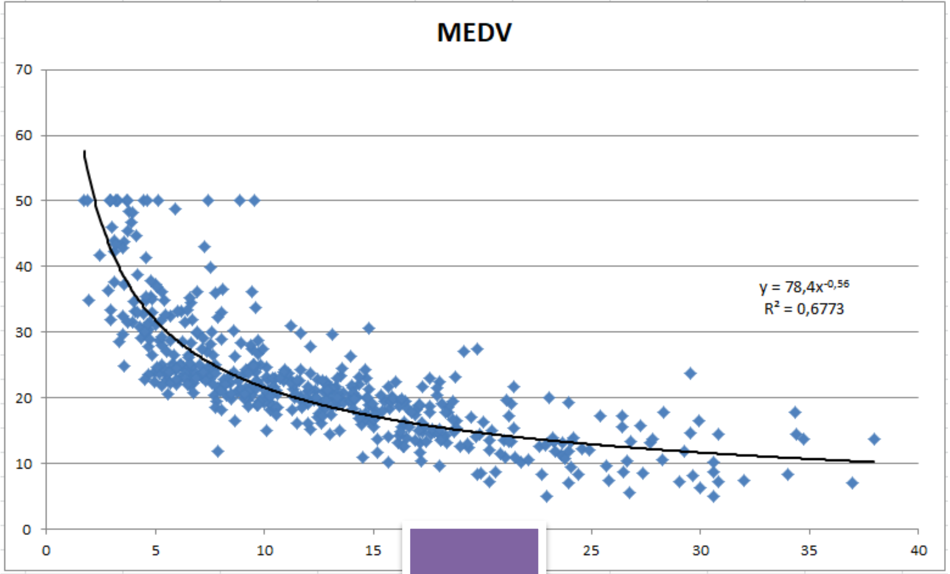
Y de esta manera podremos determinar los parámetros de la función potencial de la ecuación (3)

D2 f_x =LOG(B2;10)

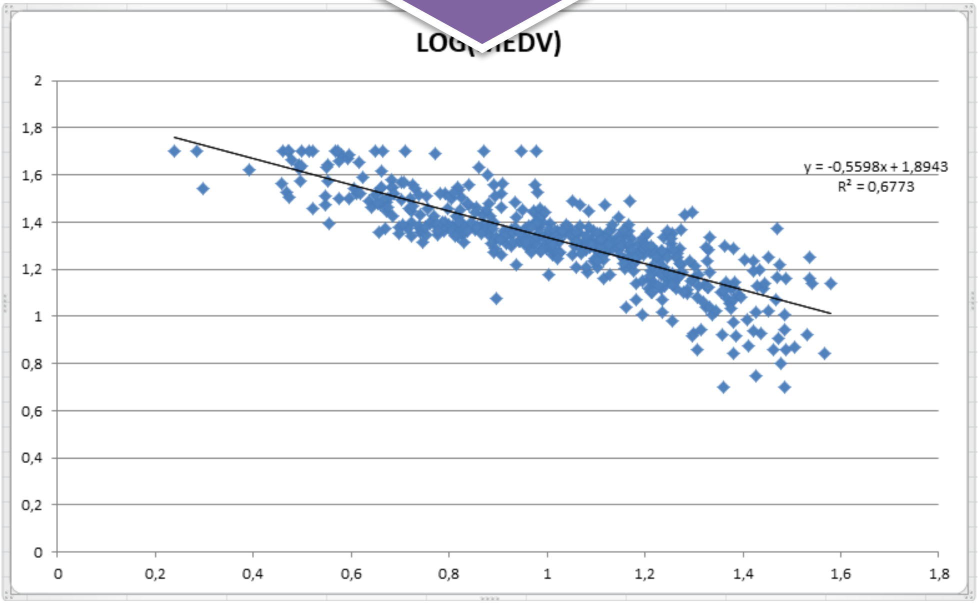
	A	B	C	D
1	LSTAT	MEDV	log(LSTAT)	LOG(MEDV)
2	4,98	24	0,69722934	1,38021124
3	9,14	21,6	0,9609462	1,33445375
4	4,03	34,7	0,60530505	1,54032947
5	2,94	33,4	0,46834733	1,52374647
6	5,33	36,2	0,72672721	1,55870857
7	5,21	28,7	0,71683772	1,4578819
8	12,43	22,9	1,09447113	1,35983548
9	19,15	27,1	1,28216878	1,43296929
10	29,93	16,5	1,47610672	1,21748394
11	17,1	18,9	1,23299611	1,2764618
12	20,45	15	1,31069331	1,17609126
13	13,27	18,9	1,12287092	1,2764618
14	15,71	21,7	1,19617619	1,33645973
15	8,26	20,4	0,91698005	1,30963017
16	10,26	18,2	1,01114736	1,26007139
17	8,47	19,9	0,92788341	1,29885308
18	6,58	23,1	0,81822589	1,36361198
19	14,67	17,5	1,16643011	1,24303805
20	11,69	20,2	1,06781451	1,30535137
21	11,28	18,2	1,0523091	1,26007139



No Lineal (POTENCIAL)



Transf de POTENCIAL a LINEAL



Observar que el R2 es **el MISMO** !

$R^2 = 0,6773$

Determinación de los parámetros de la función de regresión potencial

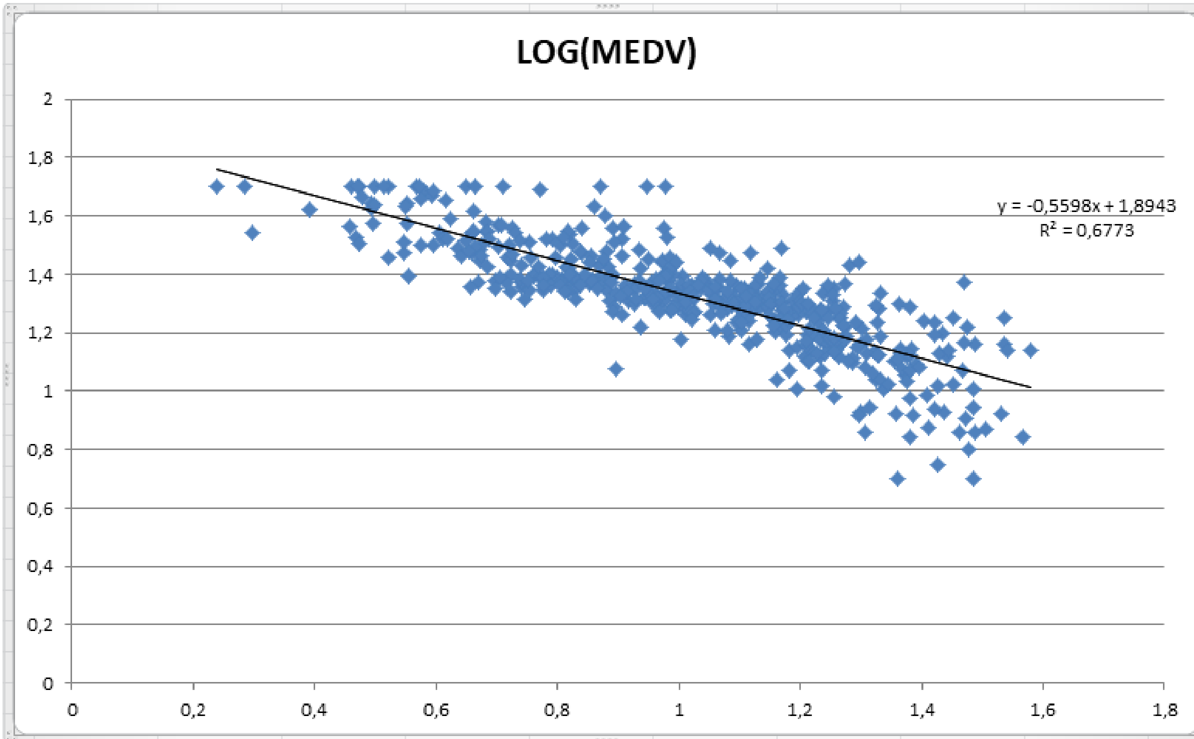
A partir de esta ecuación lineal obtenida a partir de las transformaciones de las variables originales x e y , ajustamos esta función lineal:

Y = 1,8943 - 0,5598X

Y la ecuación potencial final será

$$y = ax^b$$

con $a = 10^{\beta_0}$ y $b = \beta_1$



Corroborémoslo: a=78,4 b= 0,56 , por ende la ecuación potencial que relaciona MEDV con LSTAT nos queda:

MEDV = 78,4 LSTAT $^{-0,56}$

Que es la misma calculada en la ecuación (2) del slide 4

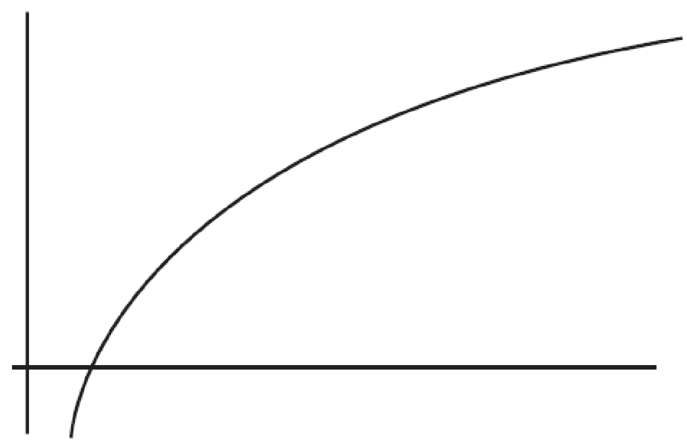
Transformación Logarítmica

$$y = A\ln(x) + B \qquad (3)$$

Y como el formato final tiene que ser lineal del tipo $P = \beta_0 + \beta_1 z \;\Rightarrow P = y$

Y como

$$\Rightarrow z = \log(x)\,,\; \beta_0 = B \text{ y } \beta_1 = A$$



$$y = A\ln(x) + B$$

Transformación Exponencial

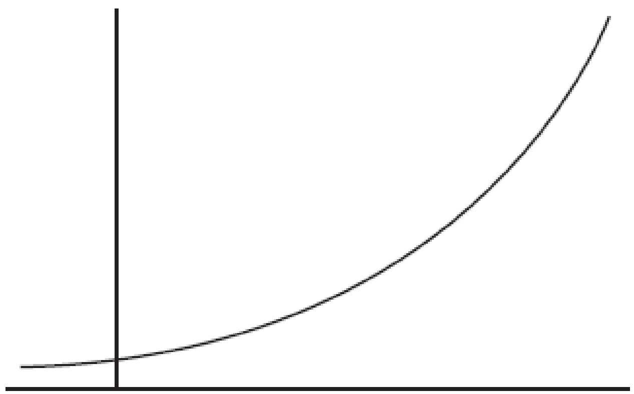
$$y = Ce^{Ax}$$

Y como el formato final tiene que ser lineal del tipo $P = \beta_0 + \beta_1 z$, calculemos las transformaciones

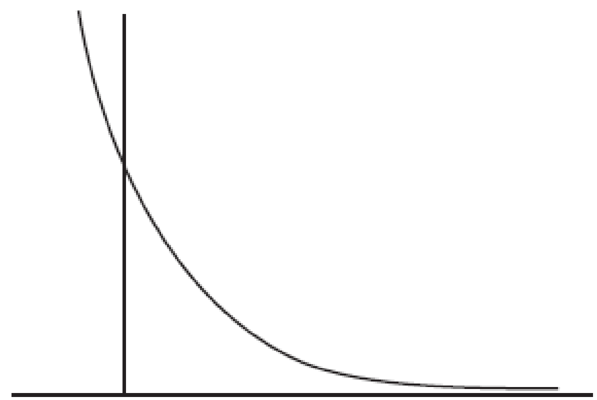
$$\ln(y) = \ln(C) + Ax$$

$$\Rightarrow \mathrm{P} = \ln(y)\,, \quad z = x, \beta_0 = \ln(\mathrm{C}) \quad \text{y} \quad \beta_1 = A$$

Para luego de ajustar a los datos y obtener los coeficientes de la regresión lineal, calcular: $C = e^{\beta 0}$



$y = Ce^{Ax}$



$y = Ce^{Ax}$