FORMAÇÃO CIENTISTA DE DADOS

REGRESSÃO LINEAR MÚLTIPLA: FÓRMULAS



FÓRMULAS

Correlação Inclinação Interceptação Previsão

ALCOHOLD TO SERVICE TO

CORRELAÇÃO DE PEARSON

ldade		Custo
	18	871
	23	1100
	25	1393
	33	1654
	34	1915
	43	2100
	48	2356
	51	2698
	58	2959
	63	3000
	67	3100

$$r = \frac{cov(X, Y)}{\sqrt{var(x). var(y)}}$$

cov: covariação

var: variância

$$r = \frac{11869,71}{\sqrt{(255,5371).(564932,6942)}}$$

$$r = \frac{11869,71}{\sqrt{144361313,3}}$$

$$r = \frac{11869,71}{12015,04529}$$

$$r = 0,9879$$

INCLINAÇÃO

ldade		Custo
	18	871
	23	1100
	25	1393
;	33	1654
	34	1915
	43	2100
	48	2356
	51	2698
!	58	2959
	63	3000
	67	3100

$$m = r \left(\frac{S_y}{S_x} \right)$$

r: correlação (0,9879)

s: desvio padrão

$$m = 0.9879 \left(\frac{751,6200}{15,9855} \right)$$

$$m = 46, 45$$

INTERCEPTAÇÃO

ldade		Custo
	18	871
	23	1100
	25	1393
	33	1654
	34	1915
	43	2100
	48	2356
	51	2698
	58	2959
	63	3000
	67	3100

$$b = \bar{y} - m\bar{x}$$

 \bar{y} : média de y

 \bar{x} : média de x

m: Inclinação (46,45)

$$b = 2104,182 - 46,45 * 42,09$$

$$b = 149,0577$$

PREVISÃO

ldade	Custo
18	871
23	1100
25	1393
33	1654
34	1915
43	2100
48	2356
51	2698
58	2959
63	3000
67	3100

$$P = b + (m * v)$$

b: interceptação (149,0577)
m: inclinação (46,45)
v: variável dependente

$$v = 54 \ anos$$
 $p = 149,0577 + (46,45 * 54)$
 $p = 2657,355$