

Fundamentos de Mineração de Dados

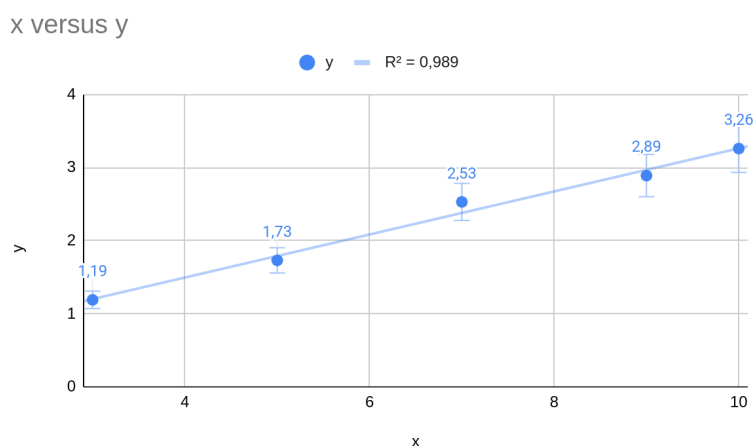
Exercício teórico sobre regressão linear

1- O tempo de execução de uma consulta para várias palavras é:

x	Palavras	3	5	7	9	10
y	Tempo	1.19	1.73	2.53	2.89	3.26

De posse desses dados, calcule:

a) Calcule os parâmetros de uma regressão simples utilizando esse conjunto de dados. b) Calcule o R^2 do modelo criado.



$$y = 0,317 + 0,295x$$

$$R^2 = 0,989$$

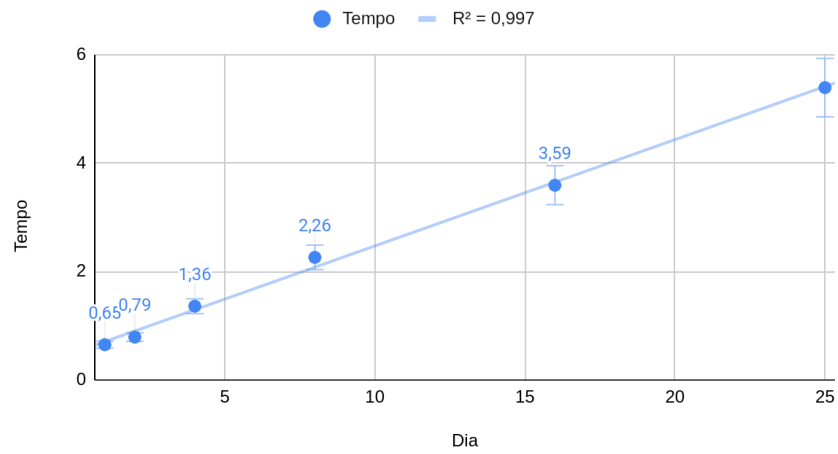
2 - Os projetistas de um banco de dados para um sistema de informação querem desenvolver uma fórmula para prever o tempo gasto em uma operação de busca nos dados dos últimos X dias. O tempo real gasto foi medido para diferentes valores de X, conforme mostrado abaixo.

Dias	1	2	4	8	16	25
Tempo	0,65	0,79	1,36	2,26	3,59	5,39

Faça o que é pedido abaixo:

a) apresente uma fórmula e a sua avaliação (R^2) para essa previsão.
b) qual é o tempo estimado para 1000 dias?

Dia versus Tempo



$$y = 0,511 + 0,196x$$

$$R^2 = 0,997$$

$$y(1000) = 0,511 + 0,196 \cdot 1000 = 196,526$$