

Formação Cientista de Dados

Introdução a Correlação e Regressão Linear

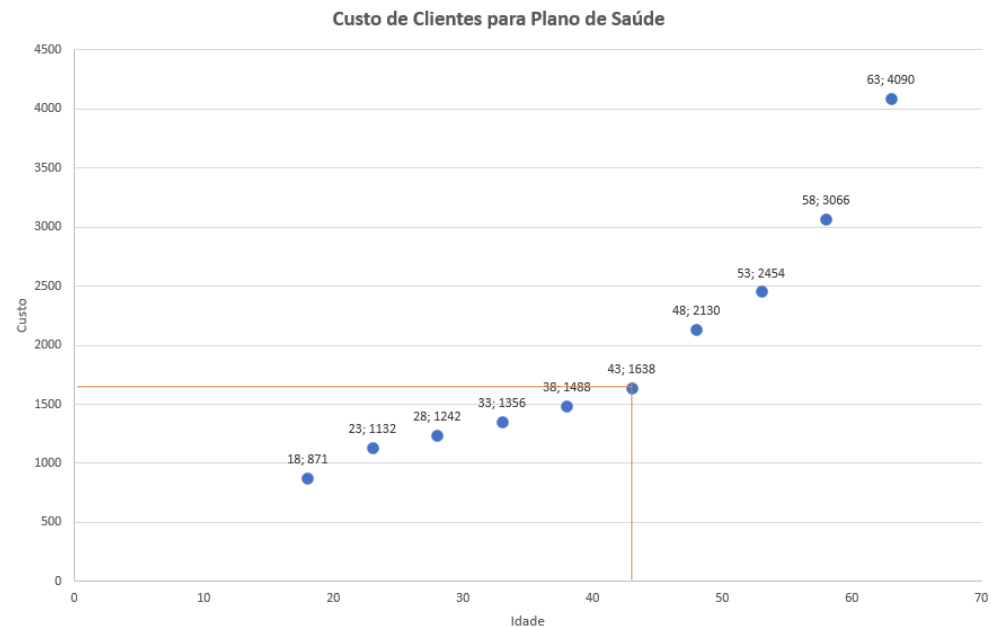
Idade	Custo
18	871
23	1132
28	1242
33	1356
38	1488
43	1638
48	2130
53	2454
58	3066
63	4090

Variáveis

- Existe uma relação matemática entre estas duas variáveis?
- Se existe, como posso medir sua força?
- Poderia usar essa relação para fazer previsões?

Gráfico de Dispersão

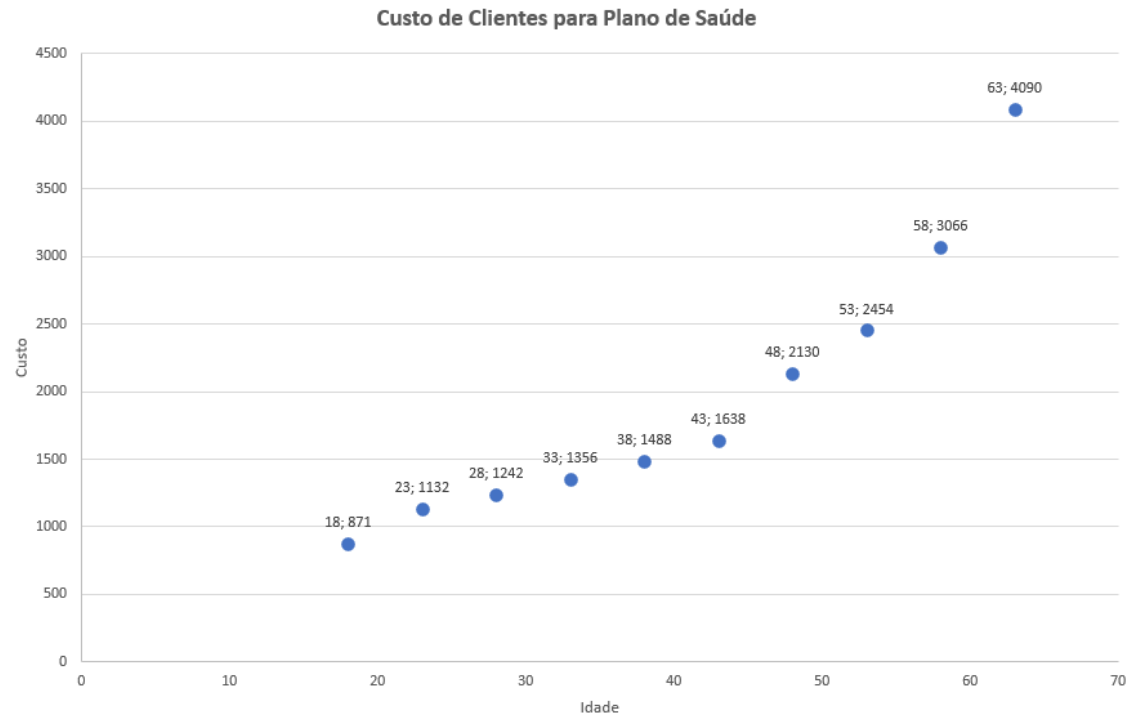
Idade	Custo
18	871
23	1132
28	1242
33	1356
38	1488
43	1638
48	2130
53	2454
58	3066
63	4090



Idade	Custo
18	871
23	1132
28	1242
33	1356
38	1488
43	1638
48	2130
53	2454
58	3066
63	4090

Eixo Y (Vertical)
Variável de Resposta
Ou Dependente
Na regressão é o que
queremos Prever

Qual vai ser o custo para o plano de saúde de um paciente com 45 anos de idade?



Eixo X (Horizontal)
Variável Explanatória
Ou Independente

Na regressão é o que explica,
ou usamos para prever



Regressão Linear

