

FORMAÇÃO CIENTISTA DE DADOS

INTRODUÇÃO A CORRELAÇÃO E REGRESSÃO LINEAR



VARIÁVEIS

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

- **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

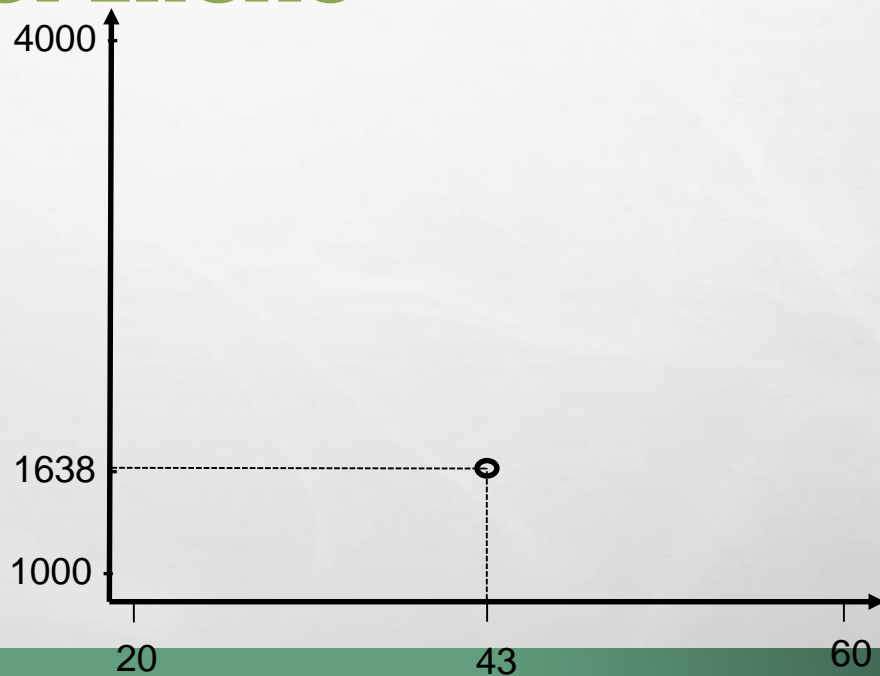
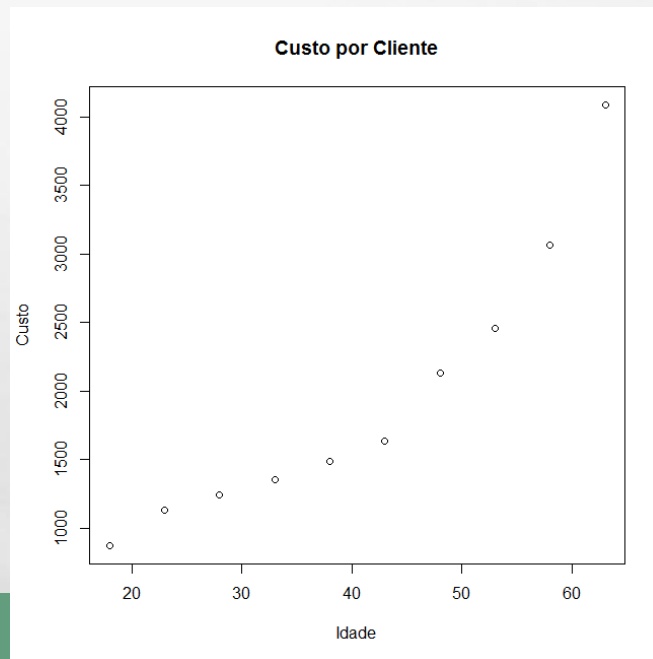


GRÁFICO DE DISPERSÃO

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



PLANO CARTESIANO

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

Y
Custo

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

Idade **X**

Na regressão é o que explica,
ou usamos para prever

REGRESSÃO LINEAR

