



Prof^a.

Elaine Marques

Monitora:

Gabriela Maia

Regressão linear simples com uso do software R



Aplicações no R



Indústria



Economia



Saúde



Educação



Biologia



Engenharia



Por onde começar?

- ✓ Conhecimentos de Matemática;
- ✓ Estatística;
- ✓ Lógica de Programação;
- ✓ Banco de Dados;
- ✓ Linguagens: R



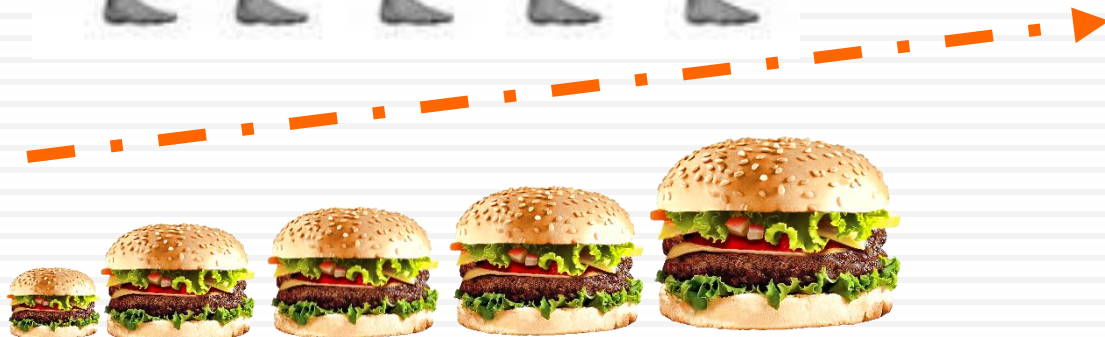
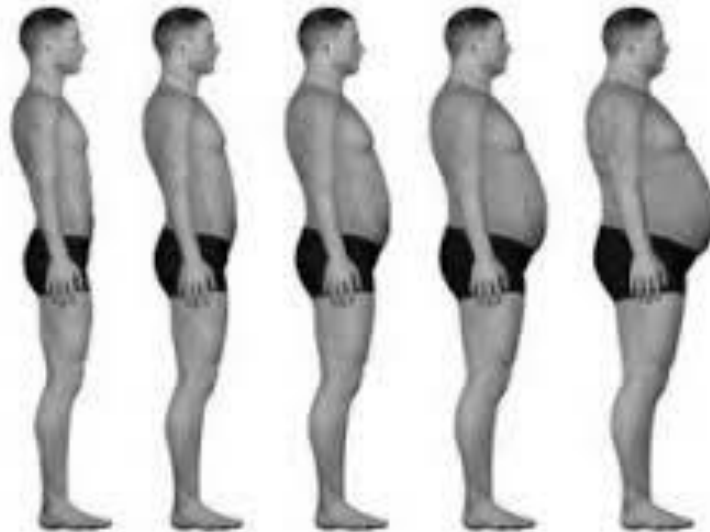
Vamos relembrar!

...O que você entende
por **Correlação** ?





...olhe pra essa imagem e pense **um pouco mais!**





Correlação

“É uma técnica estatística capaz de avaliar a existência de **relação entre duas variáveis**... Essa relação pode ser expressa através de uma **força e direção**”





Correlação

- Mede o grau da correlação (positiva ou negativa) entre duas variáveis de escala métrica;
- ✓ Se $r = 1$, dizemos que a correlação é **perfeita positiva** entre as duas variáveis.
- ✓ Se $r = -1$, dizemos que a correlação é **perfeita negativa** entre as duas variáveis.
- ✓ Se $r = 0$, dizemos que a correlação é **nula**.
- ✓ Se $0 < r < 1$, dizemos que a correlação é **positiva**.
- ✓ Se $-1 < r < 0$, dizemos que a correlação é **negativa**.

Coeficiente de correlação de Pearson

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \times \sum_{i=1}^n (y_i - \bar{y})^2}}$$



Aplicações no R

A tabela abaixo mostra duas variáveis relacionadas a profissionais de Engenharia:

Profissional	Tempo de Experiência (anos)	Salário
1	0,1	R\$ 1.290,00
2	0,8	R\$ 2.330,00
3	0,6	R\$ 1.800,00
4	3	R\$ 3.452,00
5	5,1	R\$ 5.890,00
6	7,4	R\$ 6.730,00
7	8,6	R\$ 8.600,00

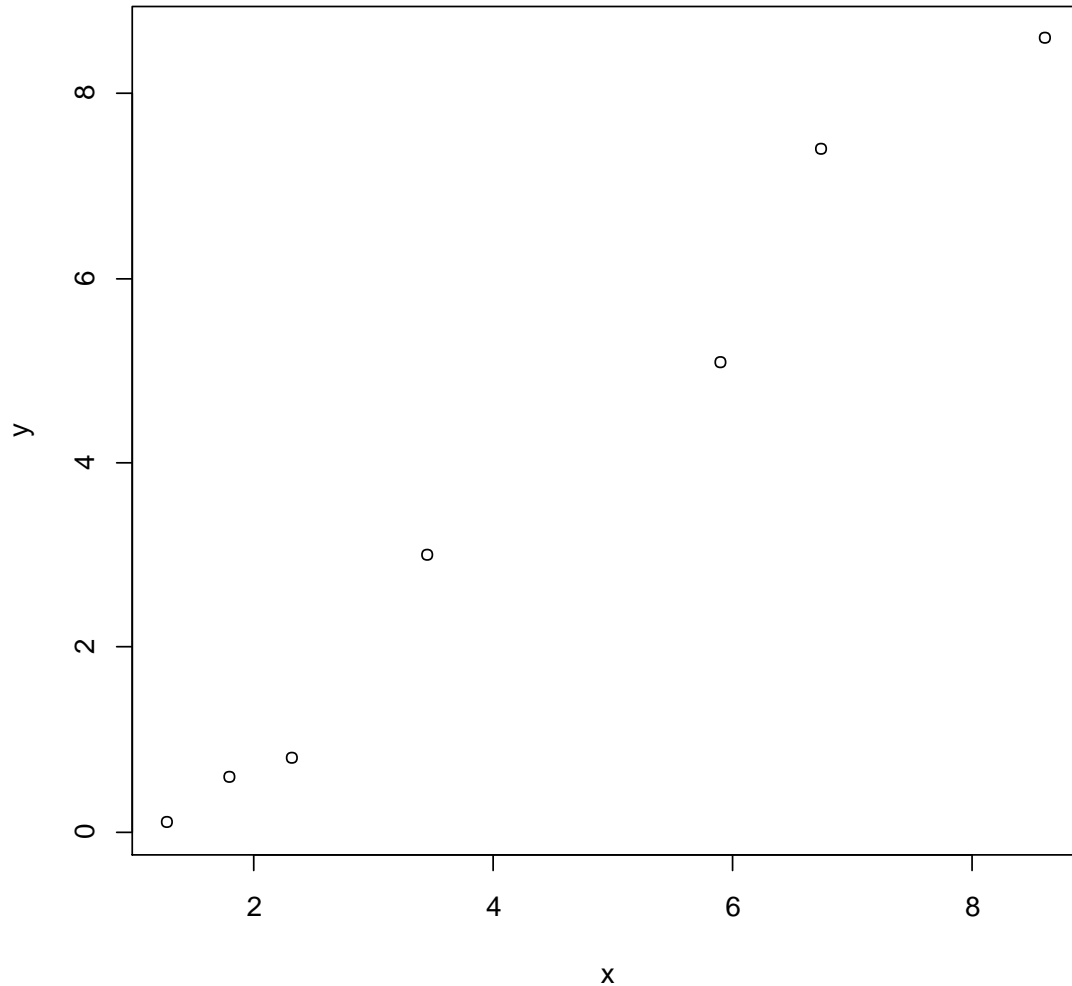


Como realizo a correlação no R?

```
1 ##### Curso R #####  
2 ## Correlacao  
3  
4 x<-c(1.290, 2.330, 1.800, 3.452, 5.890, 6.730, 8.600)  
5 y<-c(0.1, 0.8, 0.6, 3, 5.1, 7.4, 8.6)  
6  
7 cor(x,y)  
8  
9 library(stats)  
10 cor.test(x,y, method = "pearson", alternative="two.sided")  
11 cor.test(x,y, method = "pearson", alternative="less")
```



Como realizo a correlação no R?





Análise de Regressão

A **análise de Regressão** é um procedimento estatístico que tem como objetivo descrever, por meio de um modelo matemático, a relação entre duas variáveis, partindo de **n** observações das mesmas.





Análise de Regressão

A análise de regressão também parte de um conjunto de observações pareadas $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, relativas às variáveis X e Y . Suponha que podemos escrever a relação entre as duas variáveis, da seguinte maneira:

$$y_i = \alpha + \beta x_i + \epsilon_i,$$

y_i é a variável aleatória associada à i -ésima observação de Y ;

x_i é a i -ésima observação do valor fixado para a variável independente (e não aleatória) x ;

ϵ_i é o **erro aleatório** da i -ésima observação;

α é chamado de intercepto (ou coeficiente linear);

β é chamado de coeficiente angular;

α e β são parâmetros que precisam ser estimados a partir de n conjuntos de pontos (x_i, y_i) .



Aplicações no R

Um farmacêutico está interessado em medir o efeito da temperatura sobre a potência de um antibiótico. Dez amostras de 50 gramas cada foram guardadas a diferentes temperaturas, e pós 15 dias mediu-se a potência. Os resultados estão na tabela abaixo:

Temperatura (x)	30	35	50	54	57	70	76	78	90	92
Potência (y)	38	43	32	26	33	19	27	23	14	21

- Calcule os coeficientes da reta de regressão.
- Faça o gráfico dos pontos e da reta ajustada.
- Você acha que o modelo adotado é razoável? Justifique sua resposta.



