

$$S = \sum_{i=1}^n (y_i - \bar{y})^2$$

Entendendo Estatística Divertidamente

Profa. Adriana Silva

Seja bem vindX!!!

Câmera ligada e

Microfone mutado sempre
que não estiver falando

História (Pré História)

- No meio do século XVII a teoria das probabilidades, base matemática da Estatística, foi concebida.
- As cartas entre Blaise Pascal e Pierre de Fermat (1650), além dos estudos dos Bernoullis e Moivre, foram essenciais para a dedução da teoria.

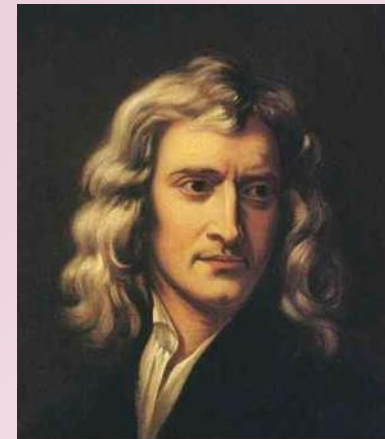
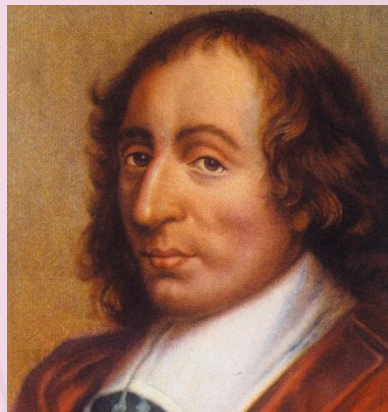




Imagem retirada do filme: Prediction by the numbers - Netflix

Assistir para a próxima aula: [Prediction by the numbers – Netflix](#)

História (Primeiras Aplicações)

- Apenas um século depois, por volta de 1749 que o termo Estatística apareceu.
- A Etimologia da palavra mostra bem o quanto as primeiras aplicações estavam ligadas ao Estado, em cada idioma a palavra aparece com um significado ligeiramente diferente, porém o ponto comum é a idéia de Ciência do Estado.

STATISTICUM COLLEGIUM

STATISTA

STATISTIK

STATISTIC

História (Primeiras Aplicações)

- As primeiras aplicações eram voltadas para as necessidades dos estados. Basicamente as aplicações se preocupavam com a coleta e classificação de dados.
- Apenas no começo do século XIX que a idéia de erro apareceu pela primeira. Laplace escreveu um livro definitivo sobre o cálculo das posição dos astros. Entre os cálculos existia a função erro, contrapondo a idéia determinista da época.



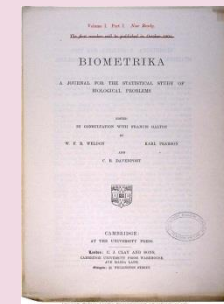
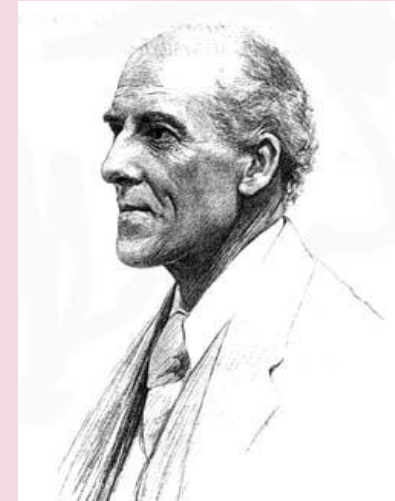
Modelo determinista versus modelo estatístico

- Modelo determinista → Existe uma relação exata entre as características
- Modelo estatístico → Existe uma relação genérica entre as características, o objetivo é chegar o mais perto possível

No início do século XX, uma revolução gradativa na ciência começou a acontecer, nos próximos slides iremos mostrar essa mudança do ponto de vista dos maiores envolvidos.

Karl Pearson (1857-1936)

- Cientista Político (nascido Carl Pearson)
- Influenciado por Sir Galton, se dedicou a Estatística
- Editor da Revista Biométrie
- Idealizou as Distribuições Assimétricas
- Deu nome ao Coeficiente de Correlação
- Idealizou o Teste Qui Quadrado
- Patrocinado pelo governo Britânico
- Buscou provar por medições a Teoria da Evolução
- Muitas das suas descobertas foram corrigidas no século XX



William Gosset (1876-1937)

- Trabalhava na Cervejaria Guinness
- Fez a ponte entre a teoria e a prática
- Utilizava o codinome Student em suas publicações
- Idealizou a distribuição “T de Student”
- Construiu a base para os Testes de Hipótese
- Descobriu a primeira aplicação da distribuição de Poisson



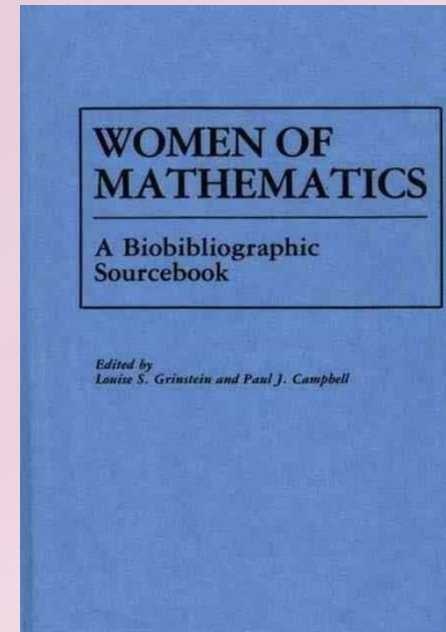
Ronald Fisher (1890-1962)

- Considerado por muitos o pai da Estatística
- Trabalhou na Estação Experimental Agrícola de Rothamsted
- De lá, sozinho, lançou uma série de artigos que serviram como base para muitas das técnicas disponíveis hoje
- Entre os estudos de Fisher, podemos destacar: ANOVA, ANCOVA, Desenho de Experimentos, EMV, Testes de Hipótese



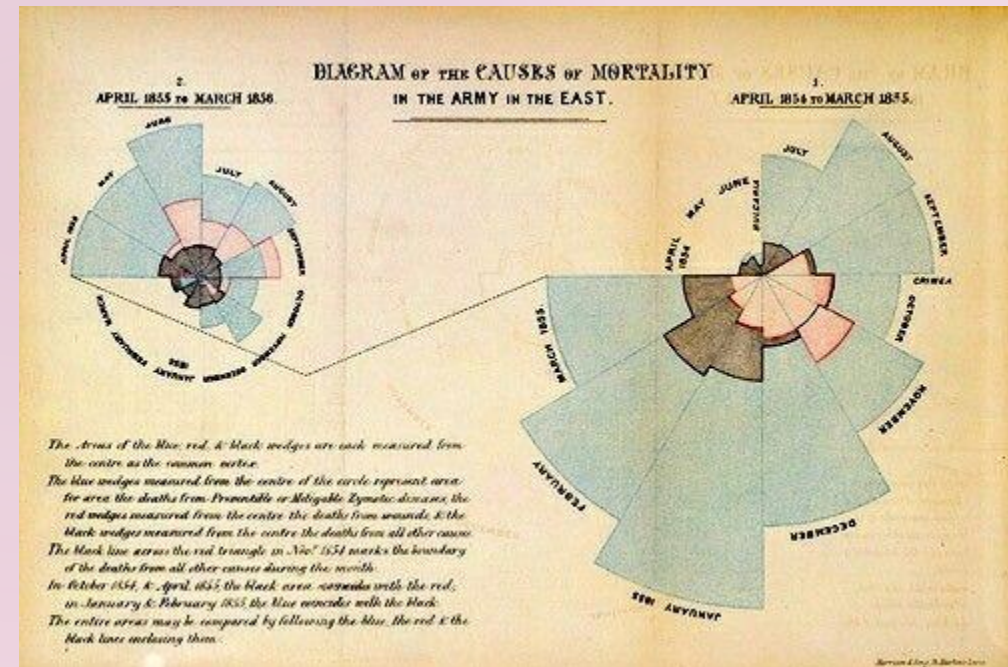
Gertrude Mary Cox (1900-1978)

- Pretendia ser pastora da Igreja Metodista Episcopal e diretora de orfanato, mas ao entrar para a universidade, decidiu que poderia ser mais útil aos fazendeiros de Iowa se estudasse estatística
- fundadora do departamento de Estatística Experimental da North Carolina State University



Florence Nightingale (1820-1910)

- Enfermeira britânica que ficou famosa por ser pioneira no tratamento a feridos de guerra
- apelido de "**A dama da lâmpada**", pelo fato de servir-se deste instrumento para na iluminação ao auxiliar os feridos durante a noite



Muitos Outros

- No decorrer do século XX muitos outros estatísticos contribuíram com os métodos que iremos abordar nesse treinamento:
 - Jerzy Neyman
 - Frank Yates
 - John Tukey
 - George Box
 - Egon Pearson
 - David Cox
 - Jerome Friedman
 - Ada Lovelace
 - Evelyn Fix

Definição Estatística: Hoje

Conjunto de métodos especialmente apropriados à coleta, à apresentação, à análise e à interpretação de dados de observação, tendo como objetivo a compreensão de uma realidade específica para a tomada da decisão.

Estatística Básica

Em geral, manipulamos um conjunto de dados com o objetivo de extrairmos informações sobre o tema em estudo.

A Estatística utiliza a variabilidade presente nos dados para obter tal informação. A variabilidade está presente em todo lugar. Por exemplo, a quantidade vendida de um produto em uma loja não é a mesma ao longo dos dias. Neste caso, a venda apresenta uma variação. Nossa estratégia consiste em avaliar as variações e obter informações através dela.

A aplicação de técnicas estatísticas envolve várias etapas:

- Coleta de dados;
- Entendimento dos dados;
- Modelos Estatísticos.

Estatística Básica

Coleta de Dados

A qualidade da solução está diretamente relacionada com a qualidade dos dados obtidos;

Um bom planejamento para coleta de dados deve considerar as seguintes perguntas:

- Qual a pergunta a ser respondida?
- Qual ferramenta de análise pretendemos usar e como utilizar os resultados?
- Qual tipo de dado é necessário para utilizar as ferramentas desejadas e responder a pergunta?
- Como coletar esses dados com o mínimo de esforço e erro?
- Onde acessar estes dados?
- Quem pode nos fornecer os dados?
- Qual o período em que os dados serão coletados?

Estatística Básica





Termos

- **Dados:** são os fatos e números coletados e sintetizados para apresentação e interpretação.
- **Observações:** são as entidades a respeito das quais se coletam os dados.
- **Variável:** é uma característica das observações que nos interessa.

Empresa	Bolsa de Valores	Mkt Share	Margem de Lucro (%)
DeWolfe Co	Amex	Small	36.7
North Cost Energy	OTC	Big	59.3
York Water Co	OTC	Average	74.2
:	:	:	:

Estatística Básica

Termos

	 Customer_Name	 Customer_Age	 Order_Date	 Total_Retail_Price
1	James Klisurich	38	11JAN2005	\$16.50
2	Sandrina Stephano	28	15JAN2005	\$247.50
3	Dianne Patchin	28	20JAN2005	\$28.30
4	Wendell Summersby	43	28JAN2005	\$32.00
5	Duncan Robertshawe	63	27FEB2005	\$63.60
6	Najma Hicks	21	02MAR20...	\$234.60
7	Tulio Devereaux	58	02MAR20...	\$35.40

**Linhas
(observações)**

Colunas (variáveis)

Estatística Básica

Termos

Um passo muito importante de qualquer análise estatística é o entendimento de cada variável dentro do banco de dados disponível para a análise. Inicialmente, os dados podem ser classificados como "qualitativos" ou "quantitativos".

- Qualitativos

- Dado qualitativo nominal
 - 1: Defeito no velocímetro;
 - 2: Defeito no tanque de gasolina;
 - 0: Outros;
- Dado qualitativo ordinal
 - B: Bom;
 - M: Médio;
 - O: Ruim;

- Quantitativo

- Dado quantitativo Discreto
 - Número de pessoas com um determinado tipo de cartão de crédito, por região
- Dado quantitativo Contínuo
 - Problema automotivo: diâmetro do eixo

Estatística Básica

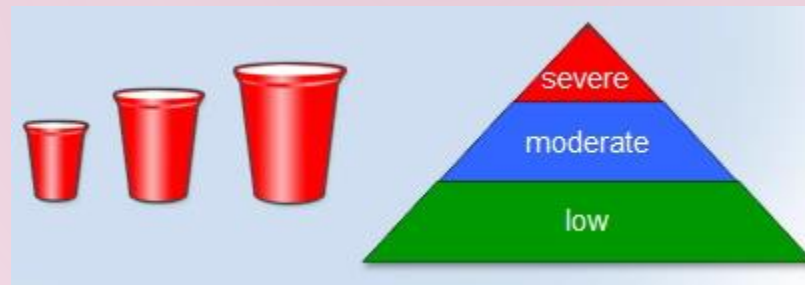
Termos

- Variável Qualitativa

- Nominal



- Ordinal



Estatística Básica

Termos

- Variável Quantitativa

- Contínua

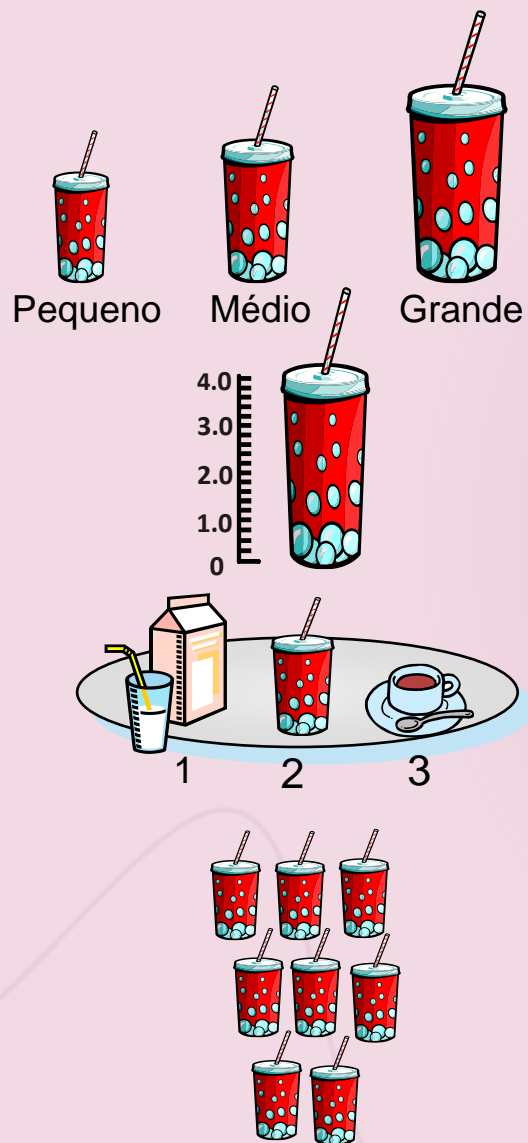


- Discreta



Estatística Básica

Termos - Exemplo tipo variáveis



- Escala Nominal
- Escala Ordinal
- Escala Discreta
- Escala Contínua

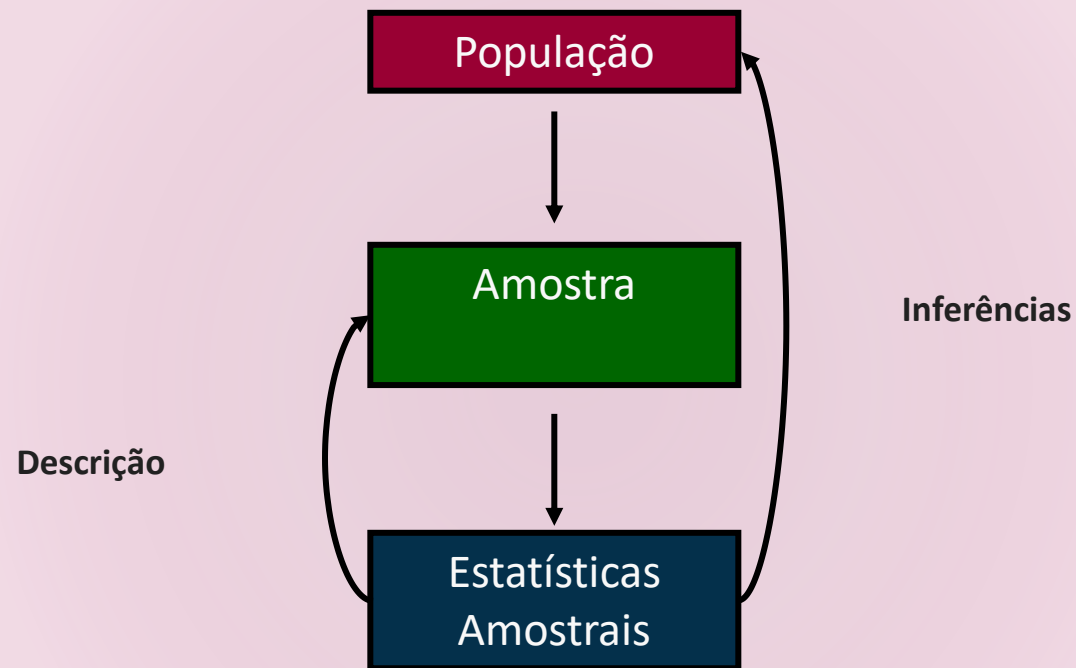
Estatística Básica

Termos

- **População** é o conjunto de indivíduos ou objetos, tendo pelo menos uma variável comum observável.
- **Amostra** é qualquer subconjunto da População.
- **Estatística Descritiva** é o ramo da Estatística responsável por descrever e sumarizar conjuntos de dados.
- **Inferência Estatística** é o ramo da Estatística que, a partir de um subconjunto da população (amostra), faz afirmações sobre esta população. Tais afirmações vêm acompanhadas de uma medida de precisão sobre sua veracidade.

Estatística Básica

Termos



Estatística Básica

Termos

- **Parâmetro** é uma medida usada para descrever uma característica da População. Parâmetros são funções dos valores populacionais.
- **Estatística** é uma característica da Amostra. Estatísticas são funções dos valores amostrais.
- As Estatísticas são utilizadas na tentativa de descrever os Parâmetros Populacionais.

	Parâmetros Populacionais	Estatísticas Amostrais
Média	μ	\bar{x}
Variância	σ^2	s^2
Desvio Padrão	σ	s

Estatística Básica

Termos

- **Dicionário de dados:** Um dicionário de dados é uma coleção de metadados que contém definições e representações de elementos de dados.

Conjunto de dados:

nome	valor	qtde	feedback
Rodrigo	19,75	3	médio
Lucas	67,9	14	ruim
Mariana	37,23	8	ótimo

Dicionário de dados:

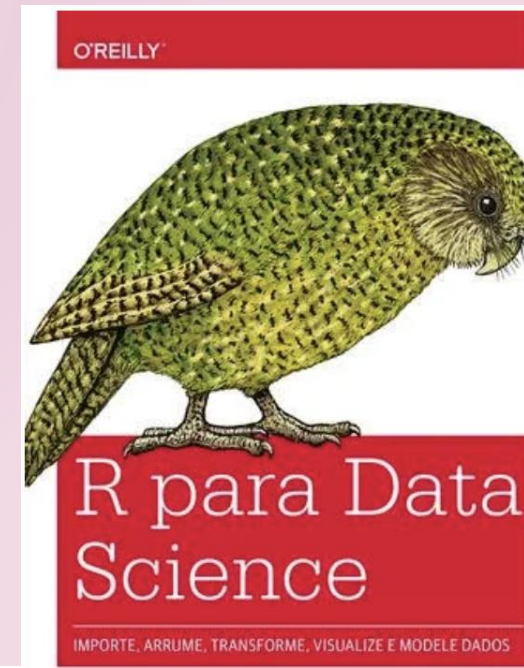
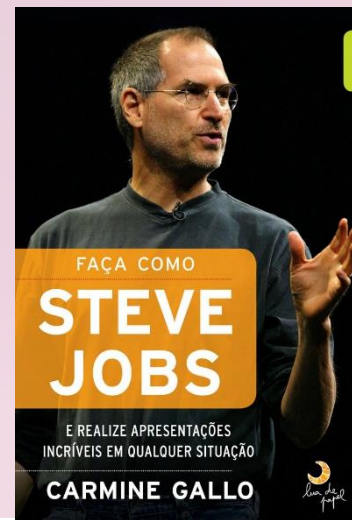
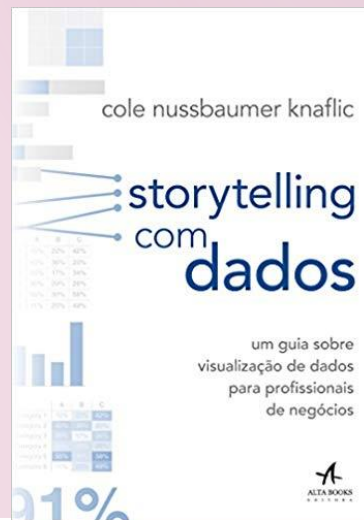
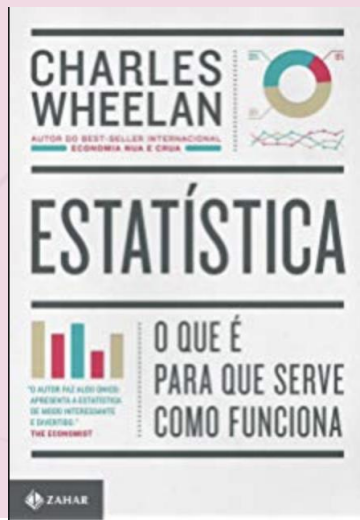
Variavel	Descricao	Tipo
nome	primeiro nome do cliente	nominal
valor	valor da compra	continua
qtde	quantidade de itens na compra	discreta
feedback	grau de satisfação na compra	ordinal

Não esqueça de deixar seu
feedback!

=]

Referência

- Moore, D., McCabe, G., Duckworth, W., Sclove, S. *A prática da Estatística Empresarial*. LTC, Rio de Janeiro, 2006.
- Anderson, D., Sweeney, D., Williams, T. *Estatística Aplicada à Administração e Economia*. Segunda Edição. Cengage Learning, São Paulo, 2011.
- www.asn.rocks
- www.curso-r.com



It's kind of fun to do the IMPOSSIBLE



dri@asn.rocks



/in/adrianamms
/in/asn.rocks



asn.rocks



www.asn.rocks

