

Inteligência Artificial ao alcance de todos



Aula 16/03/2021: Estatística Descritiva - Parte I
Professor: Eng. Rodolfo Magliari de Paiva

Objetivo da Aula

- Compreender a diferença entre Matemática e Estatística;
- Conhecer a História da Estatística;
- Áreas de Aplicação da Estatística;
- Efetuar uma boa Análise Exploratória de Dados, entendendo:
Tipos de Dados, Gráficos Estatísticos, Coleta e Organização de Dados, e Tabela de Distribuição de Frequências.

Um pouco da História da Matemática e da Estatística

- A **Matemática** é uma ciência bem mais antiga do que a **Estatística** formal que conhecemos hoje.
- É fato que só houveram (na origem), descobertas e avanços matemáticos em regiões que tinham a necessidade da **agricultura**.
- Aos poucos, os profissionais das ciências exatas foram surgindo (corda, balança, astros...), e com eles novas necessidades de descobertas junto com o desenvolvimento das cidades.
- A Matemática Primitiva teve quatro alicerces: Clima, Agricultura, Comércio e Registro.

A Ciência Estatal

Esteve presente entre os hebreus, chineses, egípcios, maias, romanos, hindus, persas e babilônios, povos que se organizaram ao redor de um Estado, que necessitavam de informações censitárias, informações essas que eram colhidas pelo **estadista**.

A Estatística

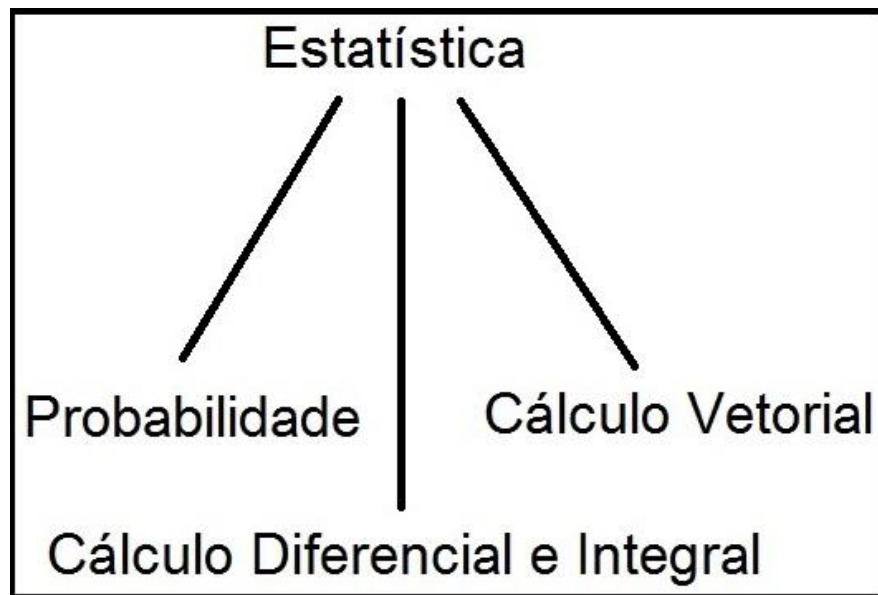
1858 – Florence Nightingale, com o trabalho:

“Notas sobre os Assuntos que Afetam a Eficiência da Saúde e Administração Hospitalar do Exército Britânico”

Nascida em 12 de Maio de 1820, Florence Nightingale era enfermeira britânica.

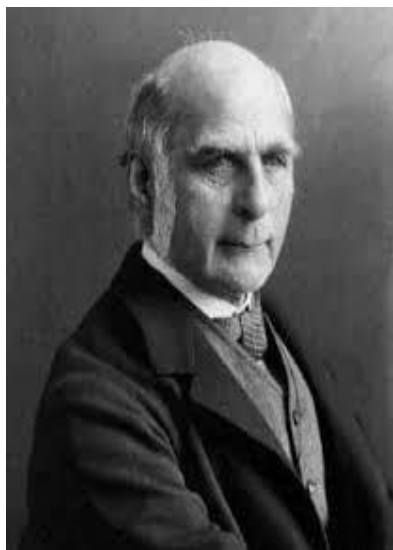


Com o passar do tempo a Estatística foi tomando forma, com os avanços de outros conteúdos matemáticos:



Tendo a atenção de estudiosos como: Christian Huygens, Pierre Fermat, Blaise Pascal, John Graunt, Jacques Bernoulli, Thomas Bayes, Poisson, Mary Somerville, entre outros...

No final do século XIX, com Francis Galton, Francis Ysidro Edgeworth, Karl Pearson e George Udny Yule, a Estatística ganhou sua aplicabilidade e operacionalidade, além do visual moderno.





Embora a Estatística Clássica se fez por conta de conteúdos da Matemática, hoje em dia sabemos que são **ciências distintas**, com alguns conteúdos em intersecção.



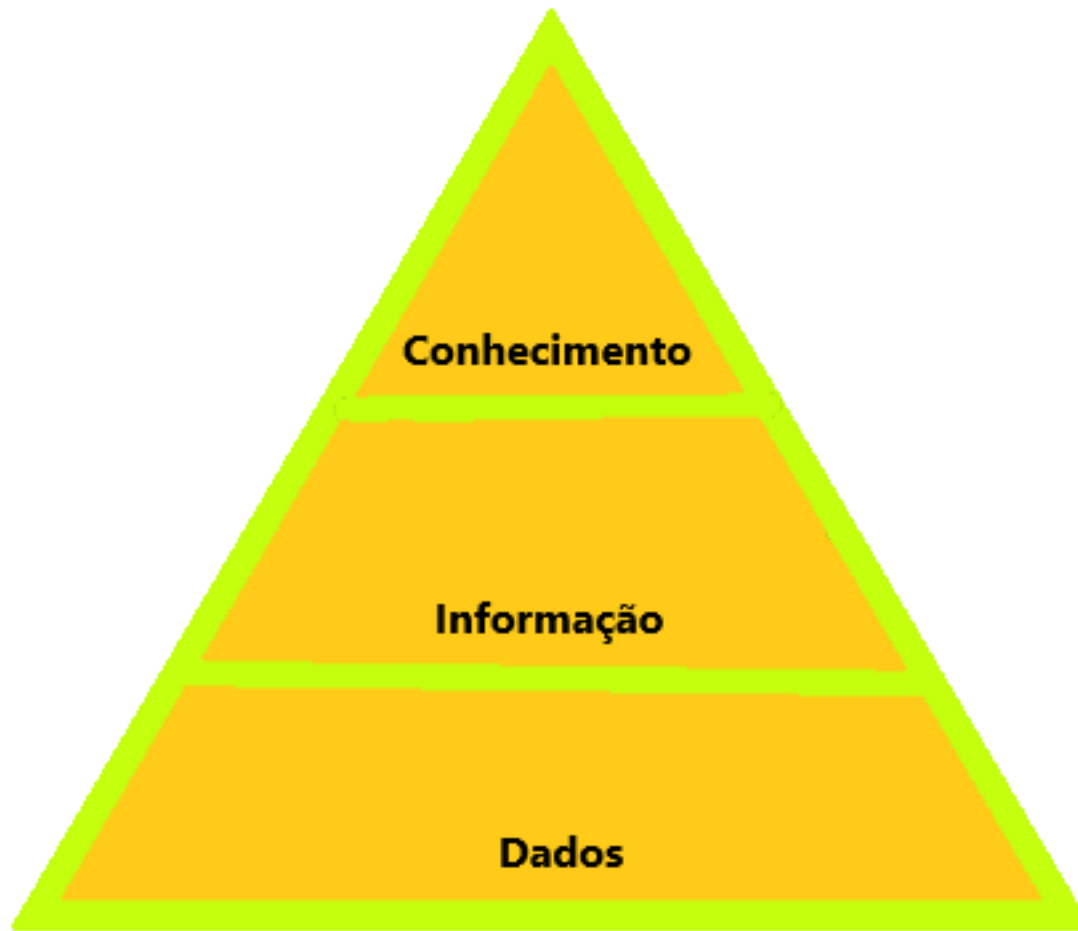
Importância

Utilizar Estatística se faz importante para:

- Tomada de decisão;
- Quantificar incerteza;
- Levantamento de dados;
- Entendimento de um fenômeno;
- Etc...



Vale destacar:



Áreas de Aplicação

- Bioestatística;
- Ciência Atuarial;
- Demografia;
- Econometria;
- Epidemiologia;
- Geoestatística;
- Controle da Qualidade;
- Data Science;
- B.I.;
- Inteligência Artificial;
- Jurimetria;
- Etc...



Análise Exploratória de Dados

As ferramentas da **Estatística Descritiva** permitem uma boa Análise Exploratória de Dados de forma rápida e sem muita complicação:

- Tipos de Dados;
- Gráficos Estatísticos;
- Coleta e Organização de Dados;
- Tabela de Distribuição de Frequências;
- Medidas de Tendência Central;
- Medidas de Dispersão;
- Medidas Separatrizes.

Tipos de Dados

Os dados (variáveis), são classificados em dois grupos:

Quantitativo X Qualitativo



Quantitativo: Discreto = Números inteiros, contagem, é mensurável
Exemplo: número de filhos, idade, ...

Contínuo = Conjunto dos Reais, também mensurável
Exemplo: altura, peso, ...

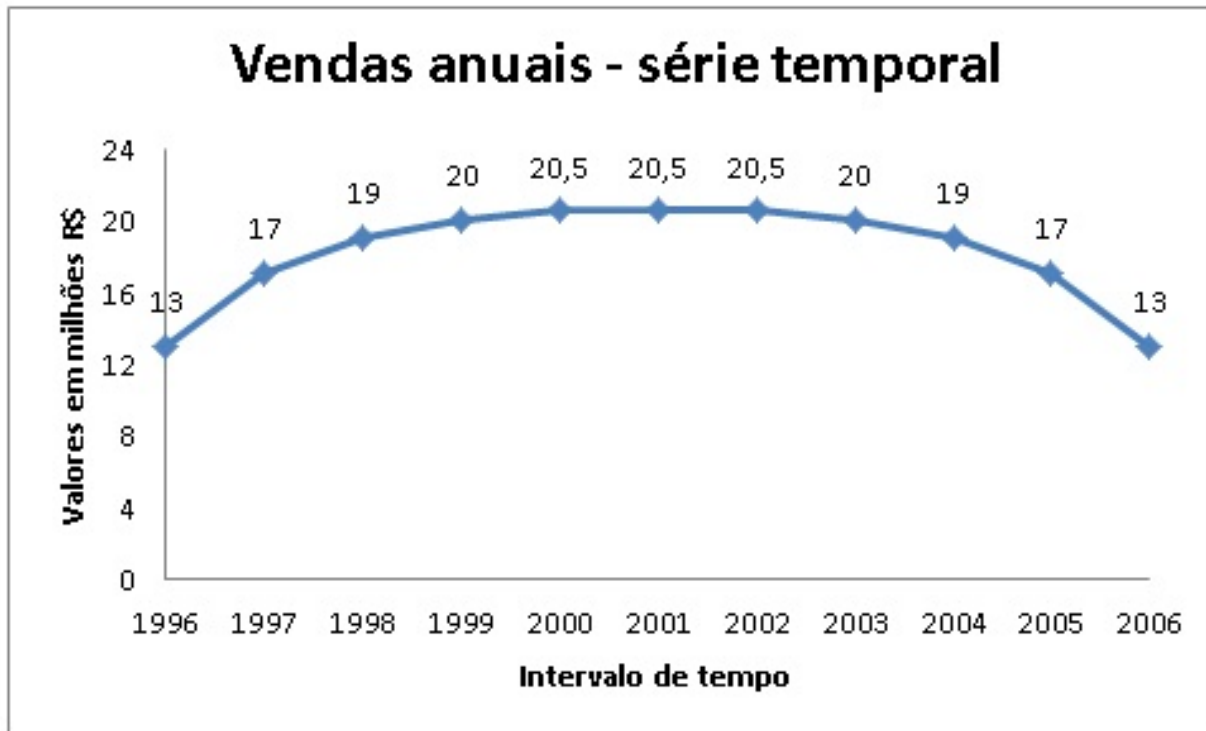
Qualitativo: Ordinal = Apresenta hierarquia
Exemplo: cargo, escolaridade, ...

Nominal = Não apresenta hierarquia
Exemplo: cor, sexo, ...

Gráficos Estatísticos

Gráficos auxiliam de uma forma rápida e prática que entendamos o **comportamento de um fenômeno**.

Exemplo:



Porém, devemos ter cuidado ao representar dados em gráficos, pois existem gráficos específicos para Variáveis Qualitativas e específicos para Variáveis Quantitativas!

- **Gráficos para Variáveis Qualitativas:**

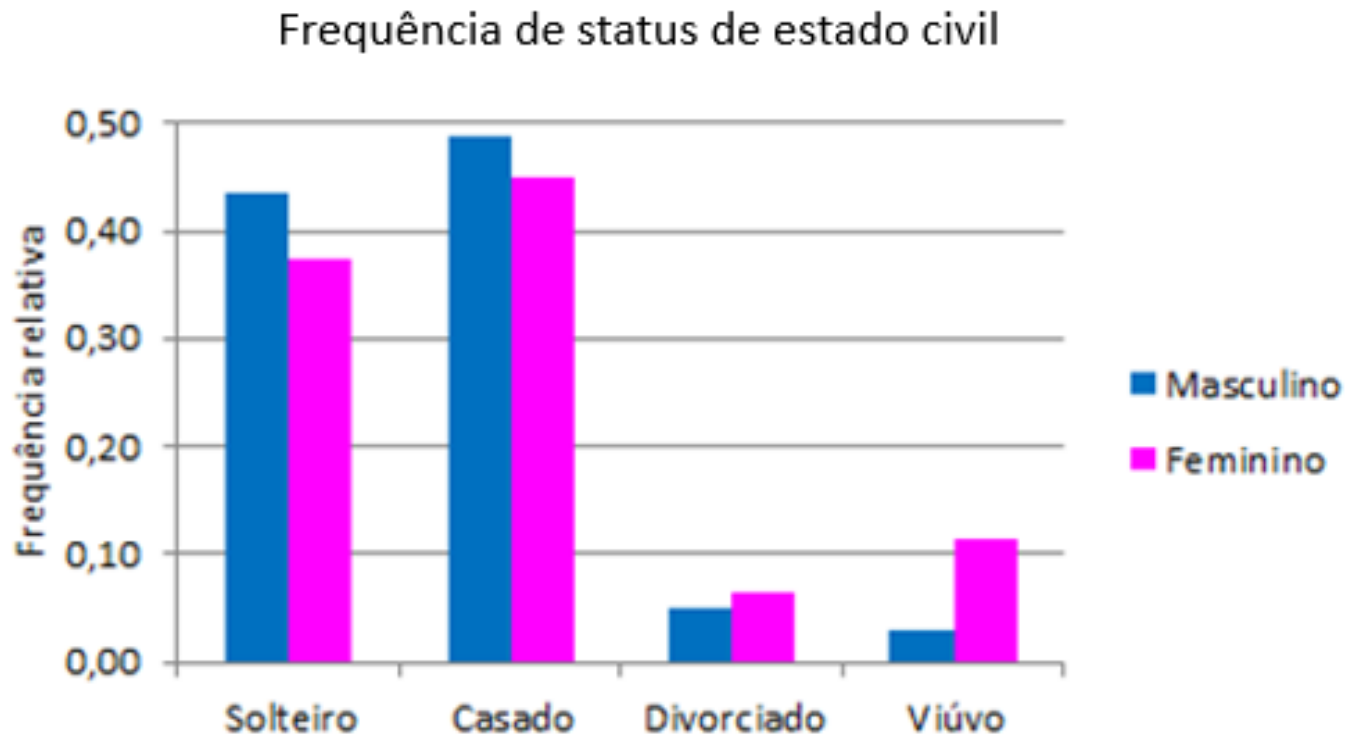


Gráfico de Barras
Verticais

Número de Nascimentos em 3 Hospitais

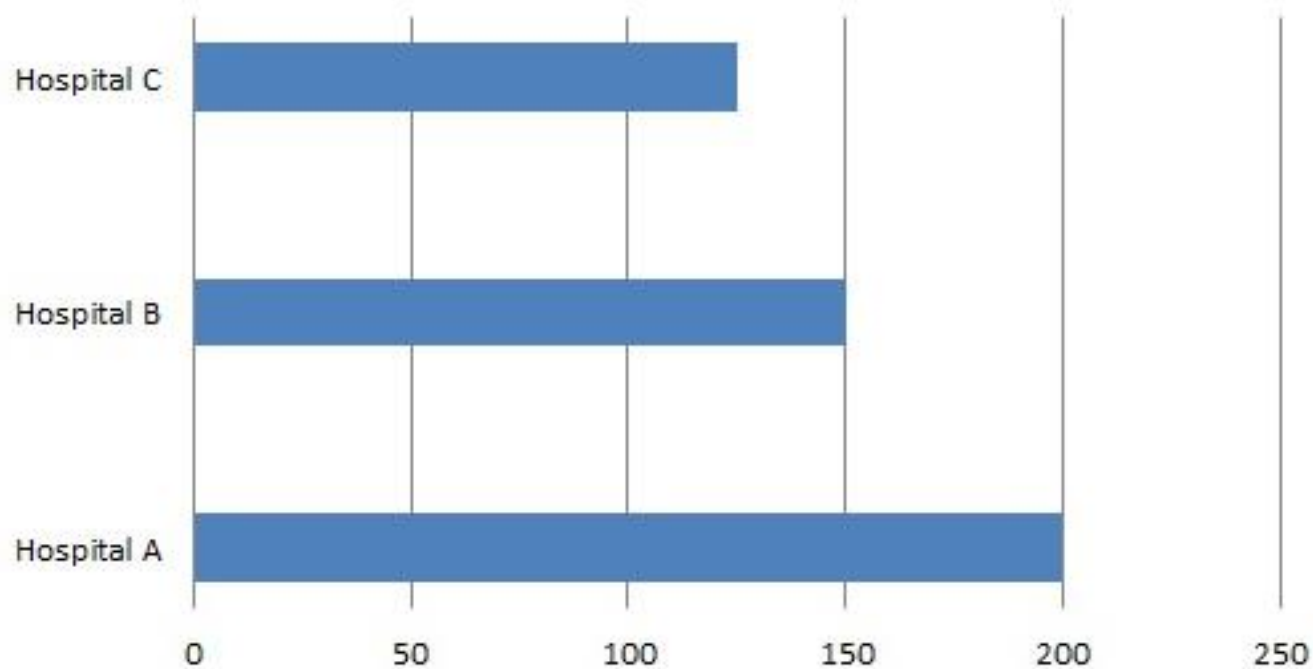
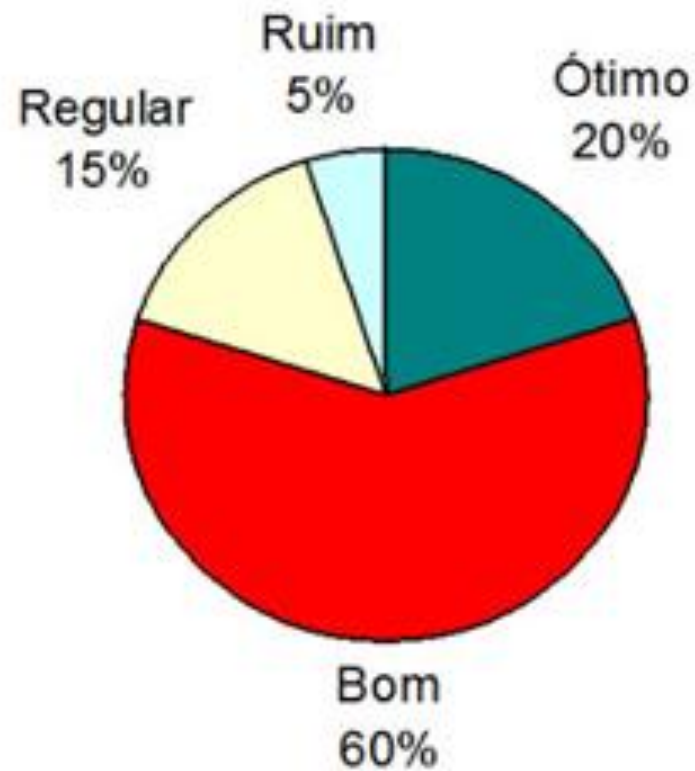
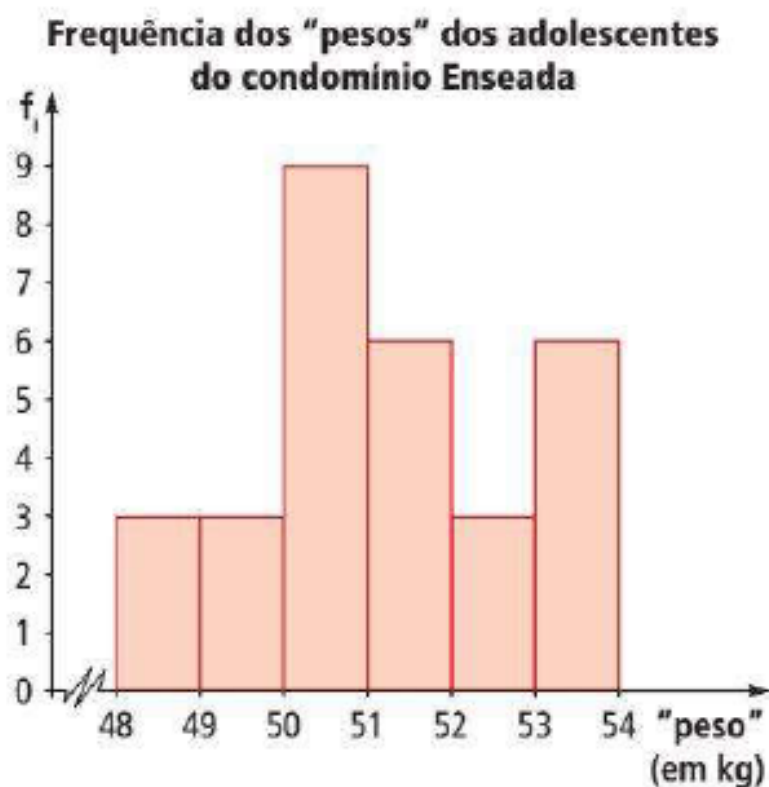


Gráfico de Barras
Horizontais

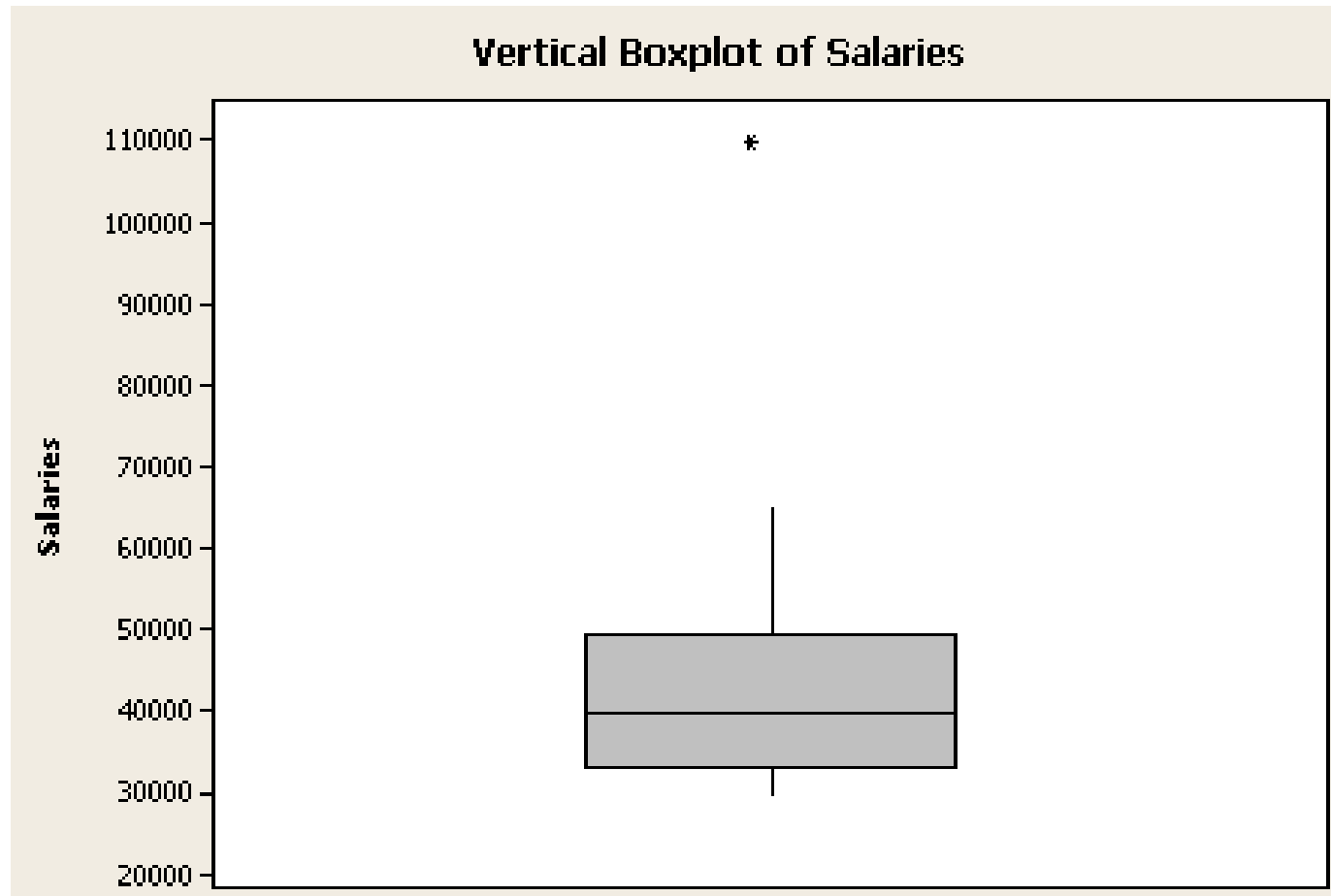


**Gráfico de Setores
ou Circular**

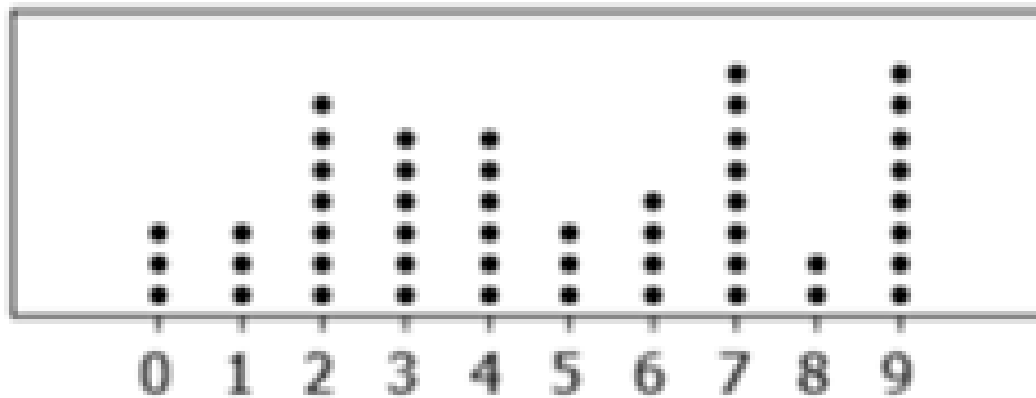
- Gráficos para Variáveis Quantitativas:



Histograma



Boxplot



Dotplot ou Strip Chart

Outros tipos de Representação de Dados

O quanto apanhamos?

Como os brasileiros foram castigados durante a infância.



70%

apanharam



1 em cada 5 apanhavam pelo menos uma vez por semana.

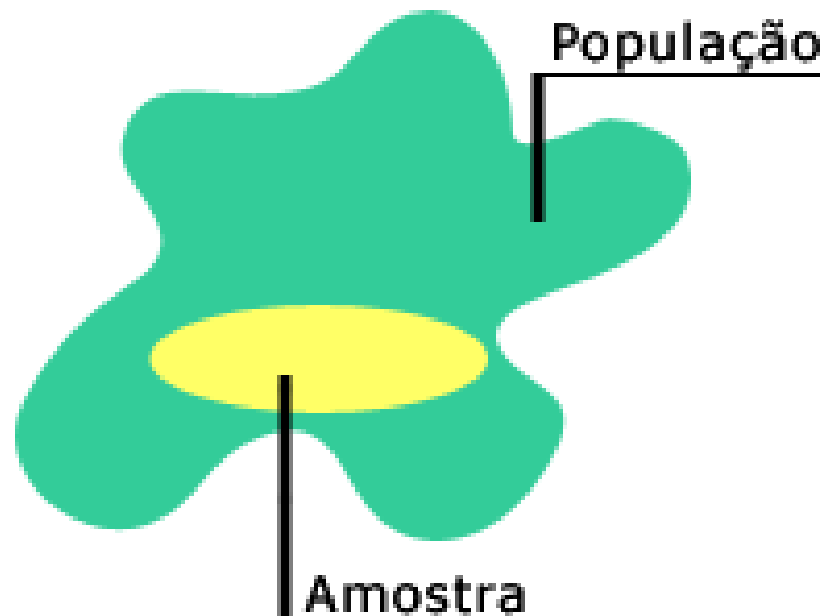
Ilustrações: Editora de Arte

Infográfico: Pictograma

Coleta e Organização de Dados

A coleta de dados é uma parte extremamente importante, pois sem dados não existe a possibilidade de ser feito qualquer tipo de estudo estatístico.

Para obter dados, a Estatística utiliza amostragens, que podem ser **probabilísticas** e **não probabilísticas**.



Tipos de Amostragem

- **Probabilística:**
 - Aleatória Simples;
 - Sistemática;
 - Estratificada;
 - Conglomerados.
- **Não Probabilística:**
 - Intencional;
 - Cotas;
 - Conveniência.

Uma vez os dados coletados (dados brutos), a forma mais prática de organizá-los é por meio de **tabelas**.

Aparelho	Potência (KW)	Tempo de uso diário (horas)
Ar condicionado	1,5	8
Chuveiro elétrico	3,3	1/3
Freezer	0,2	10
Geladeira	0,35	10
Lâmpadas	0,10	6

Nº de pessoas que doaram	Valor Doador
12	R\$ 5,00
10	R\$ 7,00
8	R\$ 10,00
7	R\$ 12,00
3	R\$ 15,00

Com os dados observados, é possível extrair vários tipos de informações!

Tabela de Distribuição de Frequências

Dentre todos os tipos de tabelas, é muito útil utilizarmos a **Tabela de Distribuição de Frequências!**

As utilizamos para representar **Variáveis Quantitativas** (Discretas e Contínuas);

Vejamos um exemplo com uma Variável Quantitativa Discreta!

Devemos sempre ter (tanto para Variáveis Quantitativas Discretas como para as Contínuas):

- **Variável (x_i)**: Representa a característica de interesse, com os respectivos valores (dados) a respeito dessa característica.
- **Frequência Absoluta (f_i)**: Representa a frequência que cada elemento se repetiu da variável.
- **Frequência Absoluta Acumulada (f_{ia})**: Representa a soma acumulativa da frequência absoluta.
- **Frequência Relativa (fr)**: Representa o percentual de vezes que a frequência absoluta ocorreu.
- **Frequência Relativa Acumulada (fra)**: Representa a soma acumulativa da frequência relativa.

Exemplo

No condomínio Enseada, o síndico decidiu montar um Salão de Jogos, mas para isso, resolveu verificar se existem adolescentes o suficiente no prédio para fazer a obra.

Ele verificou que existem 30 adolescentes no condomínio sendo:

6 adolescentes com 14 anos;

12 adolescentes com 15 anos;

9 adolescentes com 16 anos;

3 adolescentes com 17 anos.

Monte uma tabela de distribuição de frequências para representar os dados coletados pelo síndico.

Idade dos adolescentes do condomínio Enseada

Idade (x_i)	Frequência absoluta (f_i)	Frequência absoluta acumulada (f_{ia})	Frequência relativa (f_r)	Frequência relativa acumulada (f_{ra}) (%)
14	6	6	$\frac{6}{30}$ ou 20%	20
15	12	18	$\frac{12}{30}$ ou 40%	60
16	9	27	$\frac{9}{30}$ ou 30%	90
17	3	30	$\frac{3}{30}$ ou 10%	100
	n = 30		Total = 1 ou 100%	

OBS: Colocar os dados numéricos em **rol** (ordem crescente ou decrescente), é muito importante e facilita a leitura da tabela!

Conclusão

Com essas ferramentas de análise exploratória de dados da **Estatística Descritiva** é possível entender o comportamento de diversos fenômenos do dia a dia, basta **ter** ou **iniciar** a coleta de dados e na sequência:



APLICAR!



Exercícios

- 1-) Defina com suas palavras, de acordo com o que foi falado em sala de aula, o que é Estatística e para que serve.
- 2-) Qual a diferença entre dados brutos e dados em rol?
- 3-) Como se dividem as Variáveis Quantitativas e Qualitativas? Dê um exemplo para cada.
- 4-) O que é população e amostra dentro da Estatística?

5-) Marque a alternativa correta.

- a-) A apresentação de dados estatísticos pode ser feita através de gráficos, como um histograma, setograma, entre outros.
- b-) Apenas observando as tabelas de frequências é possível dizer qual o Desvio Padrão.
- c-) São exemplos de Medidas de Tendência Central a moda, a mediana e a variância.
- d-) São exemplos de Medidas de Dispersão o desvio padrão, desvio médio e a média.
- e-) Se uma amostra possui um elemento que se repete várias vezes, dizemos que é uma amostra amodal.

6-) Uma loja de roupas decidiu registrar o número de solicitações para a troca de peças compradas ao longo de 40 dias. Os resultados foram os da tabela a seguir:

14	15	15	16	14	13	16	17	14	16
15	17	17	15	16	15	14	14	13	15
16	14	15	14	16	14	14	15	16	14
13	14	15	17	16	15	16	14	15	17

Monte uma tabela de distribuição de frequências dos dados coletados.

7-) Uma empresa de telefonia solicitou a um analista que coletasse o número de reclamações ao longo de 20 dias e os resultados foram:

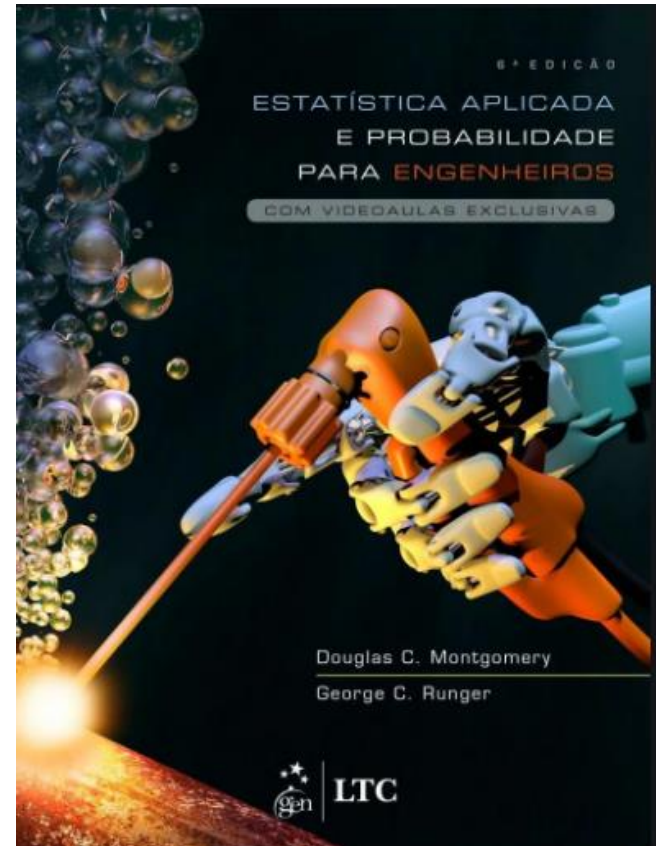
1	5	6	5	2	2	2	4	6	5
2	3	3	1	6	6	5	5	4	2

Monte uma tabela com a distribuição das frequências dos dados.

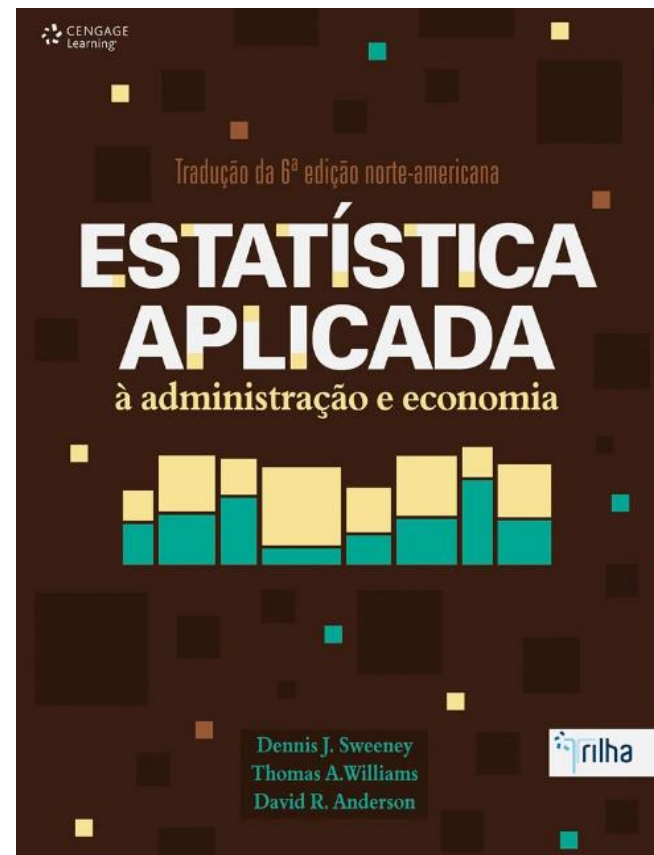
Bibliografia

MONTGOMERY, Douglas
C. e RUNGER, George C.
***Estatística Aplicada e
Probabilidade para
Engenheiros.***

6ª Edição. Rio de Janeiro:
Editora GEN | LTC, 2016



SWEENEY, Dennis J; WILLIAMS, Thomas A. e ANDERSON, David R. ***Estatística Aplicada à Administração e Economia.***
6ª Edição. São Paulo:
Editora Cengage Learning, 2013.



Contatos

Prof. Eng. Rodolfo Magliari de Paiva



Cel.: (11) 9-6866-5501



E-mail: rodolfomagliari@gmail.com



LinkedIn: Rodolfo Magliari de Paiva