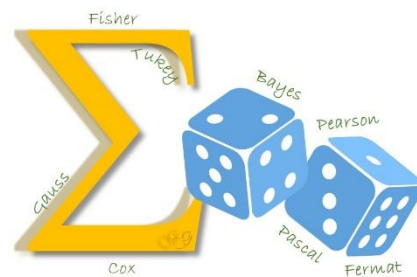


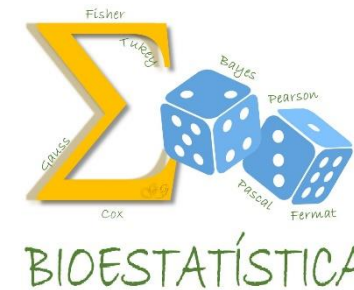
BIOESTATÍSTICA



BIOESTATÍSTICA

HEP 0175 - BIOESTATÍSTICA I

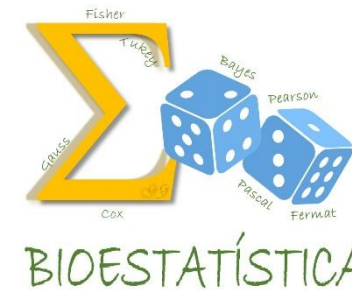
GLEICE M S CONCEIÇÃO
MARIA DO ROSÁRIO D D LATORRE
MARCELO GABRIEL
FSP USP



INTRODUÇÃO

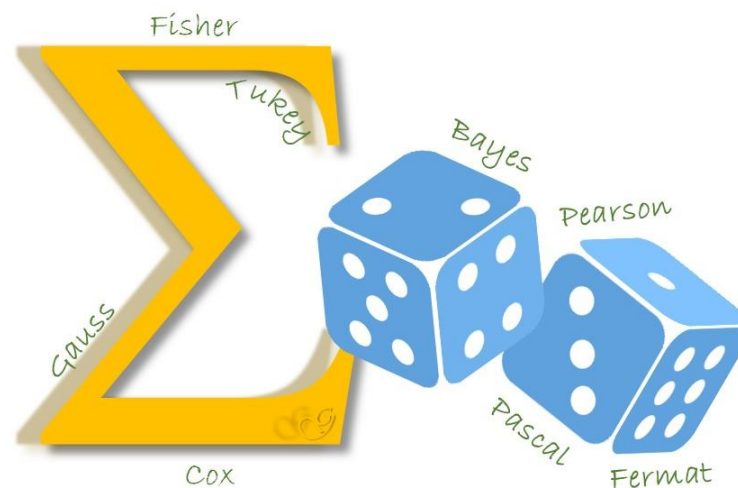
O que é estatística?

(veja o vídeo 1 no Moodle)



VÍDEO 1

O que é estatística ?

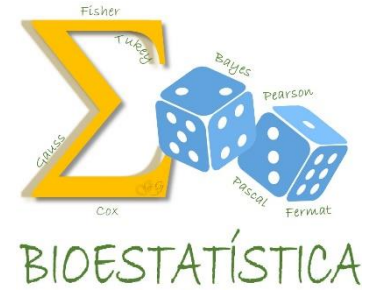


BIOESTATÍSTICA

ANÁLISE DESCRITIVA

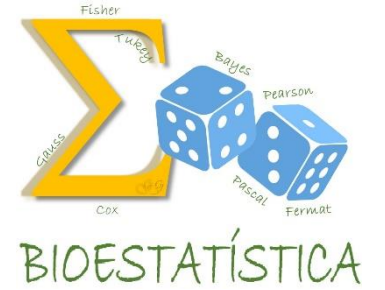
GLEICE M S CONCEIÇÃO
MARIA DO ROSÁRIO D D LATORRE
FSP USP

Pesquisa Descritiva

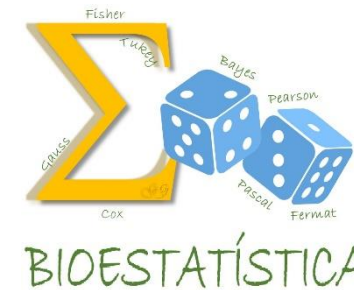


- Descreve as características de determinada população ou fenômeno.
- Permite estabelecer correlações entre variáveis e definir a sua natureza.
- Não explica o fenômeno que descreve, mas serve de base para possíveis explicações.
- Pesquisas de opinião são descritivas

Unidades de Análise (exemplos)



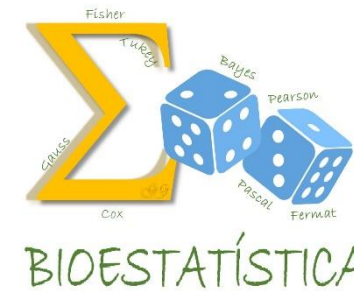
- Indivíduos
 - 60% da amostra era composta por mulheres.
 - 10% era casada.
- Famílias
 - 20% das famílias tem um único filho.
 - 15% das famílias tem apenas um dos pais.
- Lares
 - 5% dos lares tem 2 quartos.



ANÁLISE DESCRITIVA

Tipos de Variáveis

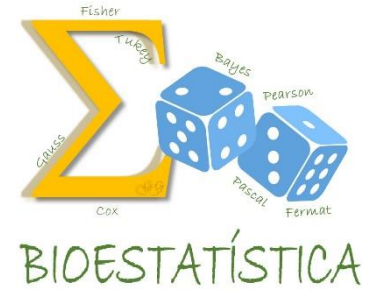
(veja o vídeo 2 no Moodle)



VÍDEO 2

Tipos de Variáveis

Análise descritiva

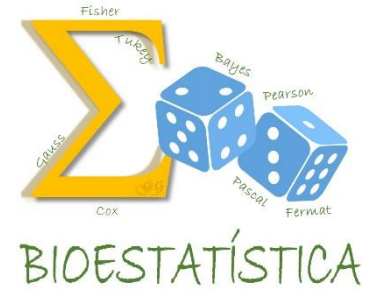


A análise descritiva consiste em "resumir" o conjunto de dados, por meio de

- ✓ tabelas
- ✓ gráficos
- ✓ medidas resumo numéricas

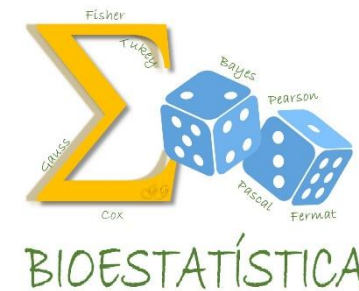
para que a informação contida nos dados seja visualizada com facilidade.

Análise descritiva



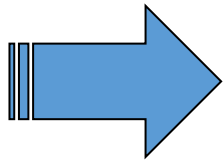
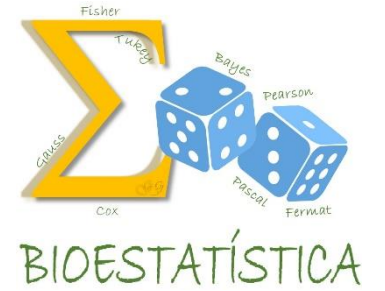
Essas ferramentas de visualização, se forem aplicadas adequadamente, revelam com detalhes o comportamento dos dados e as relações entre as variáveis de um modo que não seria possível apenas observando o conjunto de dados brutos.

O que vamos aprender



- ✓ Construir e interpretar algumas medidas resumo numéricas e gráficos comumente utilizados em uma análise descritiva
- ✓ Entender quais medidas e gráficos são adequados a cada situação
- ✓ Tais medidas e gráficos devem possibilitar a apreensão da informação contida nos dados de forma simples e direta.

Tipos de Variáveis



Para escolher a medida ou o gráfico mais adequado devemos levar em conta o tipo de variável que está sendo analisada.

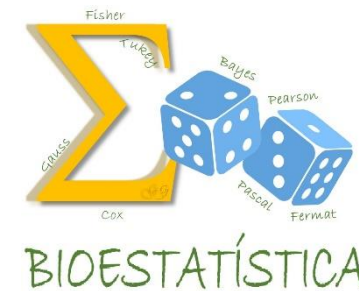
- **variáveis qualitativas ou categóricas**
 - qualitativa nominal (sexo, tipo de doença)
 - qualitativa ordinal (escolaridade)
- **variáveis quantitativas ou numéricas**
 - quantitativa discreta (número de filhos)
 - quantitativa contínua (peso, altura, anos de estudo)

O conjunto de dados “**DC.pdf**”, disponível no Moodle, contém informações sobre a presença de doença coronariana, idade (anos), sexo, prática regular de exercícios (em vezes/semana), estado nutricional, IMC (kg/m^2) e pressão arterial sistólica (em mmHg) após teste ergométrico em 40 pacientes da Clínica TAL, na cidade de Araraquara, em 2015.



Paciente	Doença Coronariana	Idade	Sexo	Exercício	IMC	Estado nutricional	Pressão Sistólica
1	N	42.0	F	2	19.1	eutrófico	90.2
2	N	34.6	F	0	22.3	eutrófico	104.9
3	N	45.8	F	0	18.7	eutrófico	120.2
4	N	44.8	F	0	19.8	eutrófico	124.5
5	N	40.4	M	0	22.9	eutrófico	131.3
.
.
.
36	S	64.6	F	0	30.7	grau II	202.6
37	N	29.7	M	0	36.1	grau II	207.5
38	N	82.2	M	0	36.9	grau II	209.1
39	S	85.6	F	0	35.7	grau II	224.8
40	N	61.7	F	0	38.1	grau II	239.0

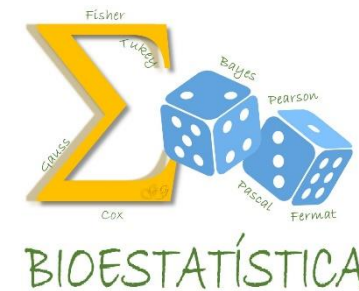
Exercício 1



Utilizando o conjunto de dados "DC.pdf", classifique, quanto à natureza, cada uma das variáveis do estudo envolvendo os 40 pacientes da Clínica TAL.

Depois, confira a solução no próximo vídeo.

Exercício 1 - solução



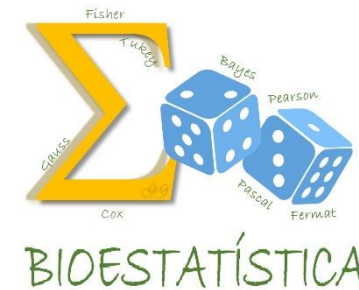
No conjunto de dados “DC.pdf”

- ✓ Doença coronariana – Qualitativa nominal
- ✓ Idade – Quantitativa contínua
- ✓ Sexo – Qualitativa nominal
- ✓ Exercício – Quantitativa discreta
- ✓ IMC – Quantitativa contínua
- ✓ Obesidade – Qualitativa ordinal
- ✓ Pressão sistólica – Quantitativa contínua

Análise descritiva



1. Descrever uma variável de cada vez
2. Descrever duas variáveis conjuntamente



1. Descrevendo / visualizando uma variável de cada vez

Variáveis Qualitativas

Medidas resumo

- ✓ Tabela de frequência
contendo frequências e %

Gráficos

- ✓ Gráfico em barras
- ✓ Diagrama circular ou “pizza”

Variáveis Quantitativas

Medidas resumo

- ✓ Medidas de posição (média, mediana, moda)
- ✓ Medidas de dispersão (variância, desvio-padrão, amplitude, quartis)

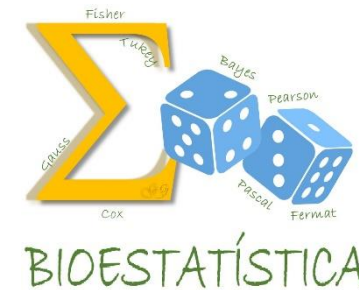
Gráficos (variáveis quantitativas contínuas)

- ✓ Box plot
- ✓ Histograma
- ✓ Polígono de frequências
- ✓ Ogiva de frequências

Para as variáveis quantitativas

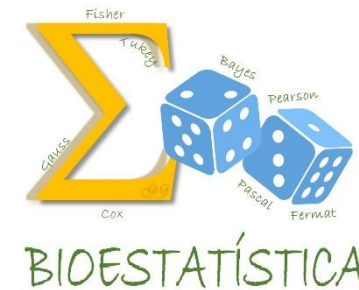
discretas que assumem poucos valores:

- ✓ Além destas ferramentas, podem ser utilizadas aquelas descritas para variáveis qualitativas
- ✓ De fato, em algumas situações, aquelas podem ser mais interessantes



**Descrevendo / visualizando
uma variável de cada vez**

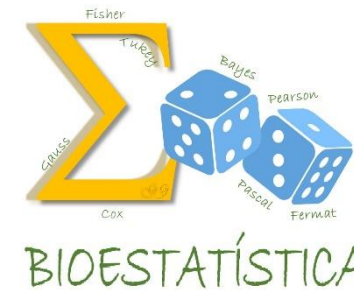
AS VARIÁVEIS QUALITATIVAS



ANÁLISE DESCRITIVA

Aprendendo a construir uma Tabela de Frequências

(veja o vídeo 3 no Moodle)



VÍDEO 3

Aprendendo a construir uma Tabela de Frequências

Variáveis Qualitativas



No conjunto de dados “DC.pdf”

- ✓ Doença coronariana
- ✓ Sexo
- ✓ Obesidade

Variáveis Qualitativas



Medidas resumo

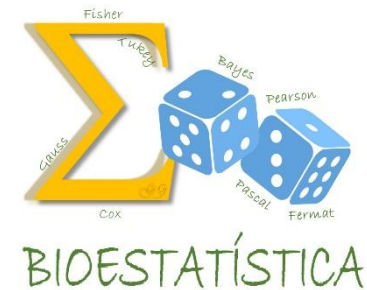
- ✓ Tabela de frequências

Gráficos

- ✓ Gráfico em barras

- ✓ Diagrama circular ou “pizza”

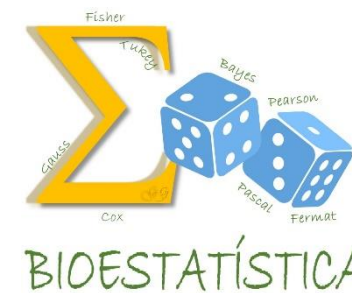
Tabela : é uma distribuição de frequências. É diferente de Quadro



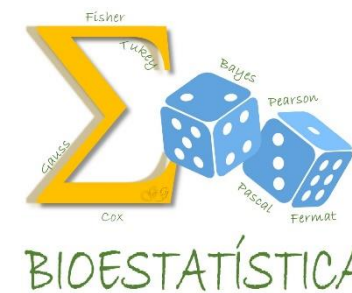
Titulo: o que (natureza do fato estudado)? como (variáveis)? onde? quando?

FONTE

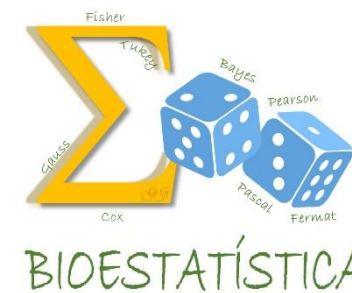
notas, chamadas



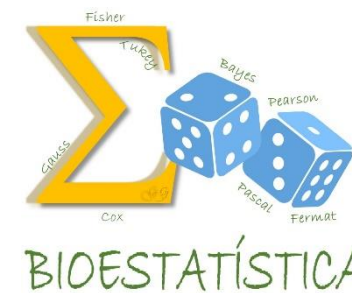
Doença coronariana	Frequência n_i	Proporção p_i	Porcentagem $100 \times p_i (\%)$
Sim			
Não			
Total	40		



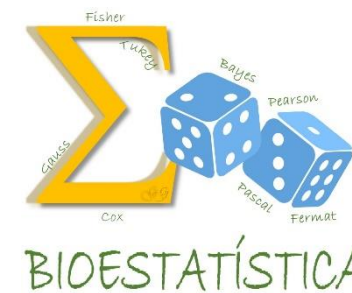
Doença coronariana	Frequência n_i	Proporção p_i	Porcentagem $100 \times p_i (\%)$
Sim	13		
Não			
Total	40		



Doença coronariana	Frequência n_i	Proporção p_i	Porcentagem $100 \times p_i (\%)$
Sim	13		
Não	27		
Total	40		

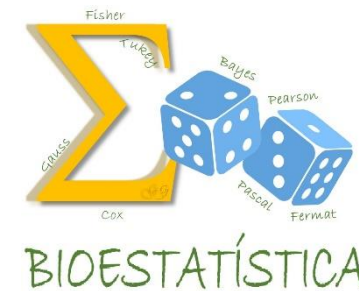


Doença coronariana	Frequência n_i	Proporção p_i	Porcentagem $100 \times p_i (\%)$
Sim	13	0,325	
Não	27		
Total	40		



Doença coronariana	Frequência n_i	Proporção p_i	Porcentagem $100 \times p_i (\%)$
Sim	13	0,325	
Não	27	0,675	
Total	40	1	

Distribuição do número e porcentagem dos pacientes segundo a presença de doença coronariana. Araraquara, 2015.



Doença coronariana	Frequência n_i	Proporção p_i	Porcentagem $100 \times p_i (\%)$
Sim	13	0,325	32,5
Não	27	0,675	67,5
Total	40	1	100

Em um artigo científico...

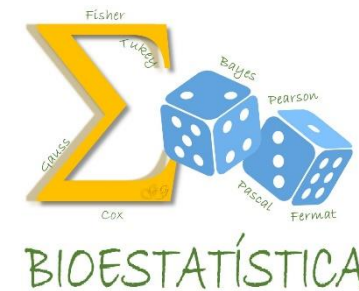


Tabela 1. Distribuição dos pacientes segundo a presença de doença. Araraquara, 2015.

Doença coronariana	n	%
Presente	13	32,5
Ausente	27	67,5
Total	40	100

Exemplo

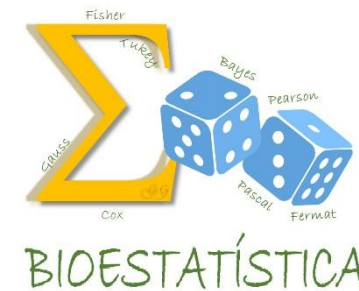


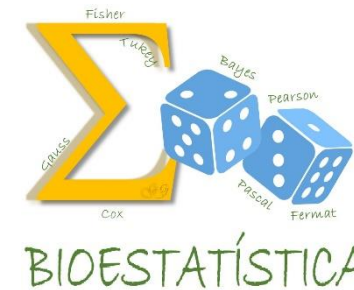
Tabela 1. Distribuição de crianças segundo níveis séricos de retinol. Cansação – Bahia, 1992.

Nível de retinol sérico ⁽¹⁾	N	%
Aceitável	89	55,3
Baixo	65	40,4
Deficiente	7	4,3
Total	161	100

⁽¹⁾ aceitável: 20,0 – 49,9 $\mu\text{g/dl}$; baixo: 10,0 – 19,9 $\mu\text{g/dl}$; deficiente: <10,0 $\mu\text{g/dl}$

Fonte: Prado MS et al., 1995.

Apresentação Tabular



Exemplo

Tabela 1. Distribuição de crianças segundo níveis séricos de retinol. Cansação – Bahia, 1992.

Nível de retinol sérico ⁽¹⁾	N	%
Aceitável	89	55,3
Baixo	65	40,4
Deficiente	7	4,3
Total	161	100

⁽¹⁾ aceitável: 20,0 – 49,9 µg/dl; baixo: 10,0 – 19,9 µg/dl;
deficiente: <10,0 µg/dl

Fonte: Prado MS et al., 1995.

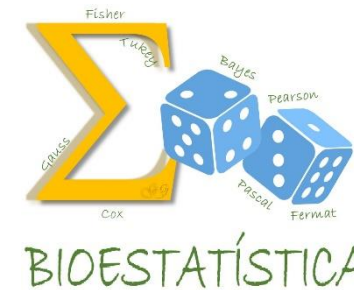
Elementos essenciais

- ✓ Título autoexplicativo (o quê, quem, onde, quando), mas não repetitivo
- ✓ Cabeçalho
- ✓ Coluna indicadora
- ✓ Corpo
- ✓ Inclua as unidades de medida

Eventualmente, no rodapé

- ✓ Notas, chamadas
- ✓ Fonte

Apresentação Tabular



Exemplo

Tabela 1. Distribuição de crianças segundo níveis séricos de retinol. Cansação – Bahia, 1992.

Nível de retinol sérico ⁽¹⁾	N	%
Aceitável	89	55,3
Baixo	65	40,4
Deficiente	7	4,3
Total	161	100

⁽¹⁾ aceitável: 20,0 – 49,9 µg/dl; baixo: 10,0 – 19,9 µg/dl;
deficiente: <10,0 µg/dl

Fonte: Prado MS et al., 1995.

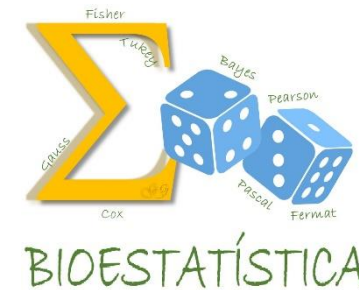
Bordas

- ✓ Para destacar a coluna indicadora, os totais e onde termina a tabela Não coloque mais bordas do que o necessário.
- ✓ Evite bordas nas laterais

No corpo da tabela

- ✓ Nenhuma casela (intersecção entre linha e coluna) deve ficar em branco
- ✓ 0 (zero) quando o valor numérico é nulo ou
- ✓ --- quando não se dispõe do dado
- ✓ Inclua totais de linhas e/ou colunas para facilitar as comparações.

Apresentação Tabular



Exemplo

Tabela 1. Distribuição de crianças segundo níveis séricos de retinol. Cansação – Bahia, 1992.

Nível de retinol sérico ⁽¹⁾	N	%
Aceitável	89	55,3
Baixo	65	40,4
Deficiente	7	4,3
Total	161	100

⁽¹⁾ aceitável: 20,0 – 49,9 µg/dl; baixo: 10,0 – 19,9 µg/dl;
deficiente: <10,0 µg/dl

Fonte: Prado MS et al., 1995.

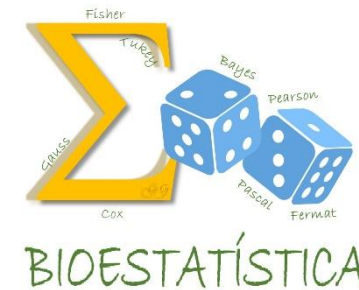
Casas decimais

- ✓ A tabela deve ser uniforme quanto ao número de casas decimais;
- ✓ Não utilize mais casas decimais do que o necessário

Ao analisar uma tabela

- ✓ Evite ler o seu conteúdo, mas aponte os padrões e as principais tendências sugeridas pelos dados
- ✓ Resuma as similaridades e destaque as diferenças.

Apresentação Tabular



Elementos essenciais

- ✓ Título autoexplicativo (o quê, quem, onde, quando), mas não repetitivo
- ✓ Corpo
- ✓ Cabeçalho
- ✓ Coluna indicadora
- ✓ Inclua as unidades de medida

Eventualmente, no rodapé

- ✓ Fonte
- ✓ Notas, chamadas

Bordas

- ✓ Para destacar a coluna indicadora, os totais e onde termina a tabela Não coloque mais bordas do que o necessário.
- ✓ Evite bordas nas laterais

Apresentação Tabular

No corpo da tabela

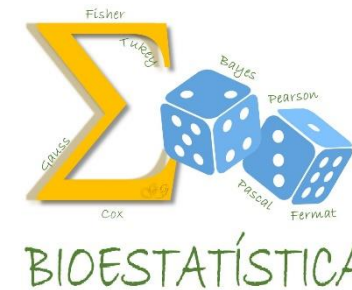
- ✓ Nenhuma casela (intersecção entre linha e coluna) deve ficar em branco
 - - :valor numérico é nulo
 - ... :quando não se dispõe do dado
 - 0 / 0,0/ 0,00 : valor numérico é muito pequeno
 - ? :dúvidas quanto à exatidão da frequência
 - § : quando o dado retifica informação já publicada
- ✓ Inclua totais de linhas e/ou colunas para facilitar as comparações.

Casas decimais

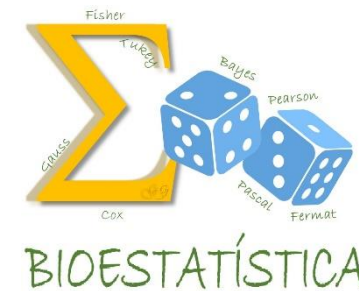
- ✓ A tabela deve ser uniforme quanto ao número de casas decimais;
não utilize mais casas decimais do que o necessário

Ao analisar uma tabela

- ✓ Evite ler o seu conteúdo, mas aponte os padrões e as principais tendências sugeridas pelos dados
- ✓ Resuma as similaridades e destaque as diferenças.



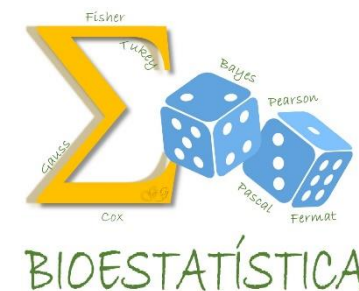
Exercício 2



- a) Utilizando o conjunto de dados "DC.pdf", construa uma tabela contendo a distribuição de frequências da variável "Sexo".
- b) Como você apresentaria as distribuições de frequências de ambas as variáveis, "Doença Coronariana" e "Sexo", em um artigo científico? Dê especial atenção aos cuidados com a apresentação tabular.
- c) Seria razoável apresentar a distribuição de frequências da variável "Obesidade"? Sim? Não? Por quê?

Depois, confira os resultados no Moodle.

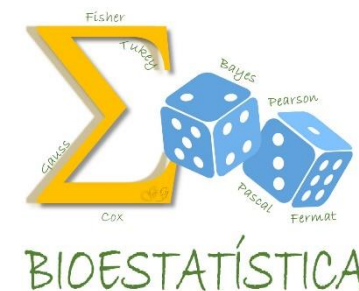
Exercício 2 – a)



Distribuição dos pacientes segundo sexo.
Araraquara, 2015.

Sexo	Frequência n_i	Proporção p_i	Porcentagem $100 \times p_i(\%)$
Masculino	25	0,625	62,5
Feminino	15	0,375	37,5
Total	40	1	100%

Exercício 2 – b)



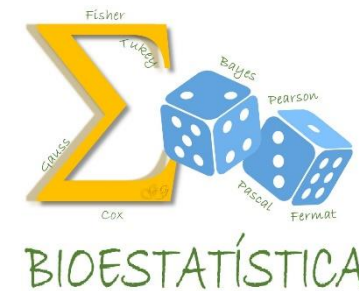
Em um artigo científico...

Tabela 1. Distribuição dos pacientes segundo a presença de doença coronariana e sexo. Araraquara, 2015.

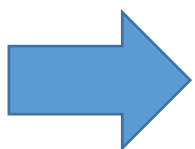
Variável	n	%
Doença coronariana		
Presente	13	32,5
Ausente	27	67,5
Sexo		
Masculino	25	62,5
Feminino	15	37,5

Exercício 2 – c)

Distribuição dos pacientes segundo o grau de obesidade.
Araraquara, 2015.



Obesidade	Frequência n_i	Proporção p_i	Porcentagem $100 \times p_i(\%)$
Ausente	10	0,25	25
Sobrepeso	10	0,25	25
Grau I	10	0,25	25
Grau II	10	0,25	25
Total	40	1	100



Como o pesquisador decidiu amostrar 10 pacientes em cada categoria de obesidade, não faz sentido apresentar esta tabela como um resultado do estudo!

exemplo

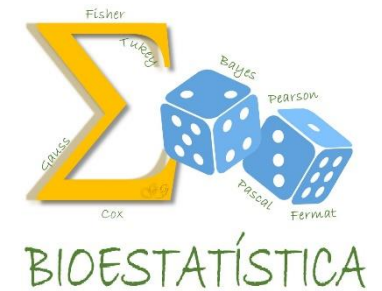


Tabela 1.2 - Distribuição do número e porcentagem de nascidos vivos segundo sexo. Local A, 1999.

Sexo	Nº	%
Masculino	26	52
Feminino	24	48
Total	50	100

FONTE: Ruiz F, *Conceitos básicos de estatística, demografia e mortalidade*. Ministério da Saúde , Brasília, 1976 (modificado).

exemplo



Tabela 1.9 - Distribuição do número e percentagem de nascidos vivos, segundo idade da mãe (anos): Local A, 1999.

Idade da mãe (anos)		Nº	%
15	— 20	14	28
20	— 25	17	34
25	— 30	8	16
30	— 35	7	14
35	— 40	2	4
40	— 45	2	4
Total		50	100

FONTE: Ruiz F, *Conceitos básicos de estatística, demografia e mortalidade*. Ministério da Saúde, Brasília, 1976 (modificado).

exemplo

Tabela 1.13 - Distribuição do número e percentagem de nascidos vivos segundo sexo e peso ao nascer. Local A, 1999.

Peso ao nascer*	Sexo				Total	
	Masculino		Feminino			
	Nº	%	Nº	%	Nº	%
baixo peso	3	11,5	2	8,3	5	10,0
peso normal	23	88,5	22	91,7	45	90,0
Total	26	100	24	100	50	100

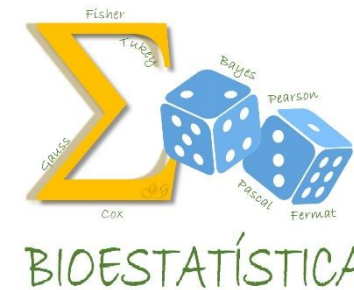
* critérios: baixo peso = abaixo 2.500g; peso normal = 2.500g ou mais

FONTE: Ruiz F, *Conceitos básicos de estatística, demografia e mortalidade*. Ministério da Saúde, Brasília, 1976 (modificado).

Exercício

Para a próxima aula, ler a parte descritiva dos artigos:

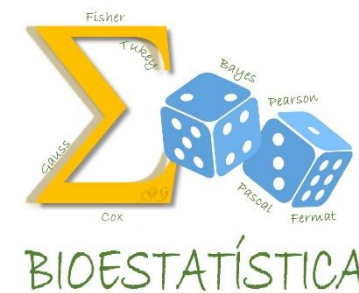
- 1. Correlação entre consumo alimentar e nível de atividade física habitual de praticantes de exercícios físicos em academia (Duran et al, 2004).**
- 2. Caminhoneiros de rota curta e sua vulnerabilidade ao HIV, Santos, SP (Villaninho et al, 2002).**



ANÁLISE DESCRITIVA

Aprendendo a construir um Gráfico em Barras

(veja o vídeo 4 no Moodle)



VÍDEO 4

Aprendendo a construir um Gráfico em Barras

Distribuição dos pacientes segundo a presença de doença coronariana. Araraquara, 2015.



Doença coronariana	Frequência n_i	Proporção p_i	Porcentagem $100 \times p_i (\%)$
Sim	13	0,325	32,5
Não	27	0,675	67,5
Total	40	1	100%

Gráfico em barras

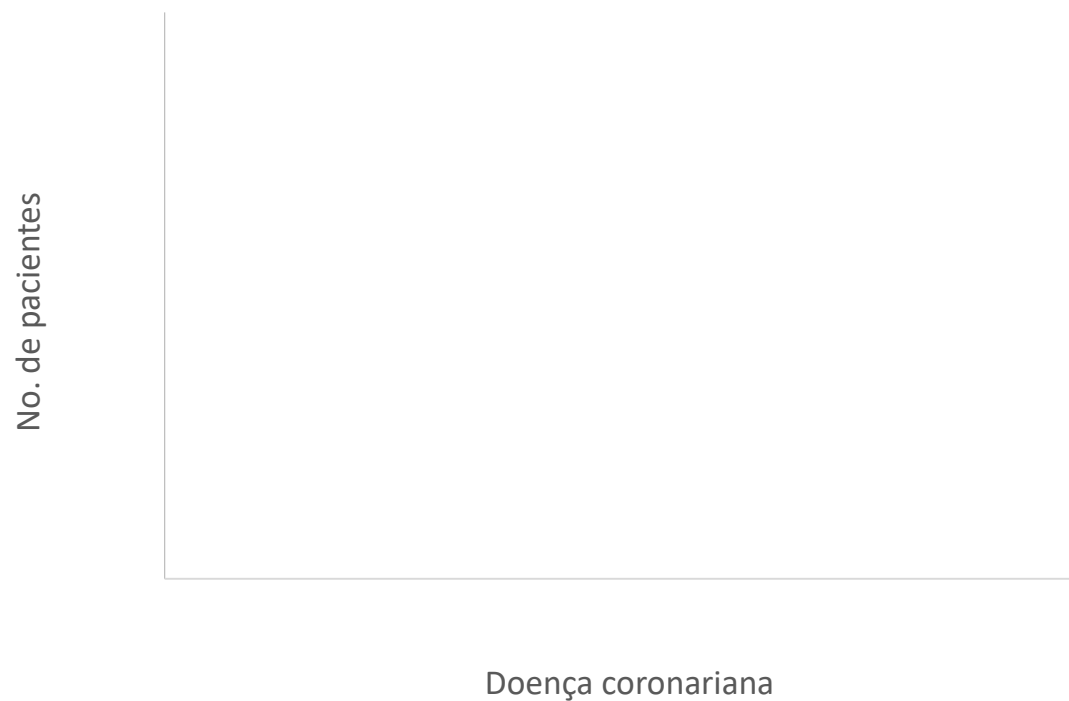
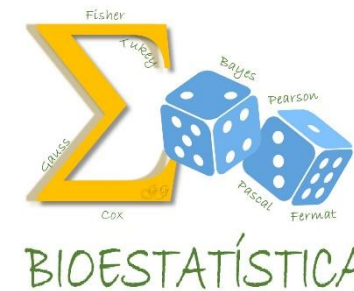


Gráfico em barras

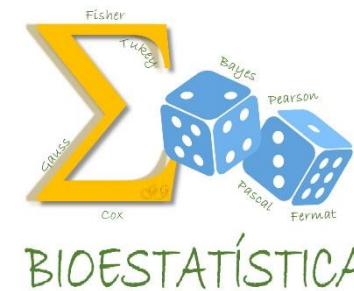


Gráfico em barras

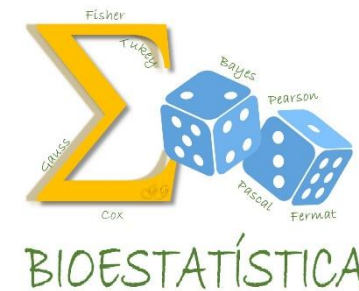


Gráfico em barras

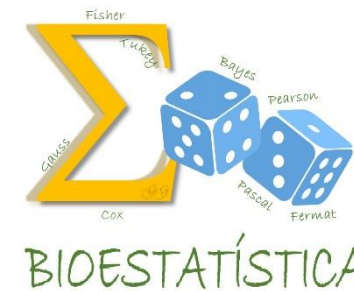


Gráfico em barras

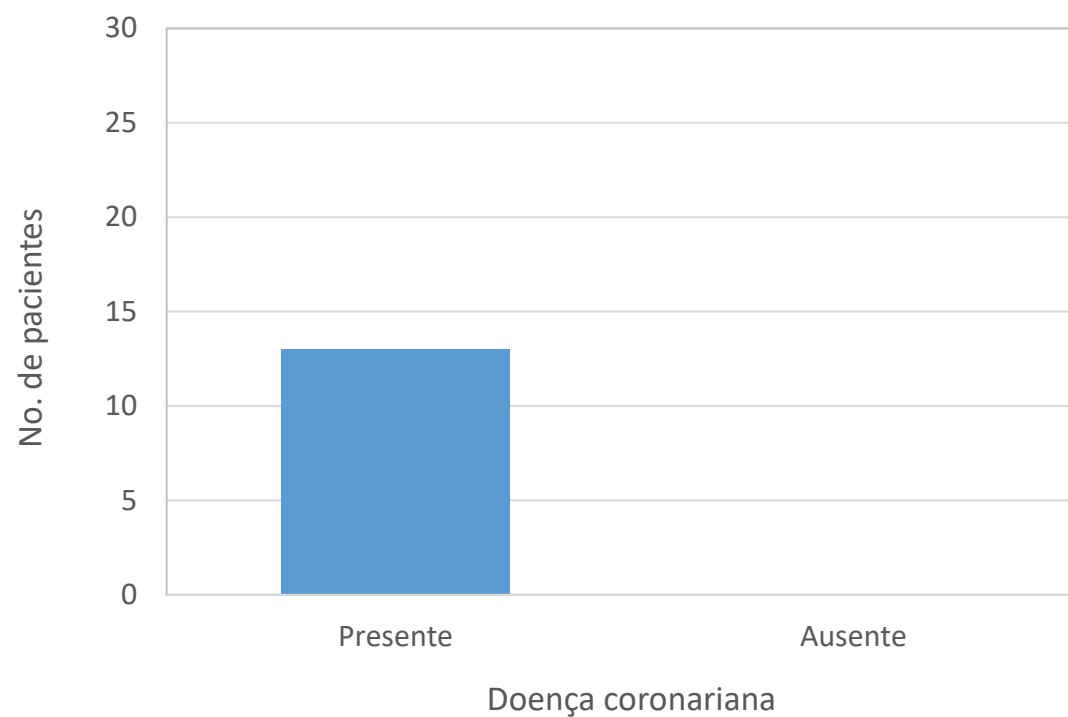
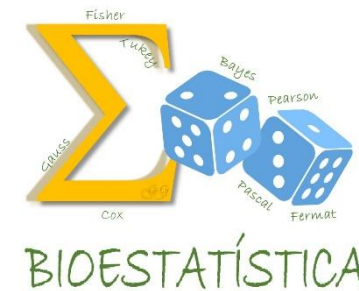


Gráfico em barras

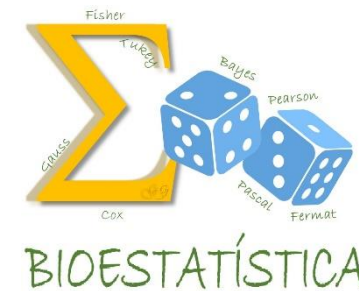


Gráfico 1. Distribuição dos pacientes segundo a presença de doença coronariana . Araraquara, 2015.

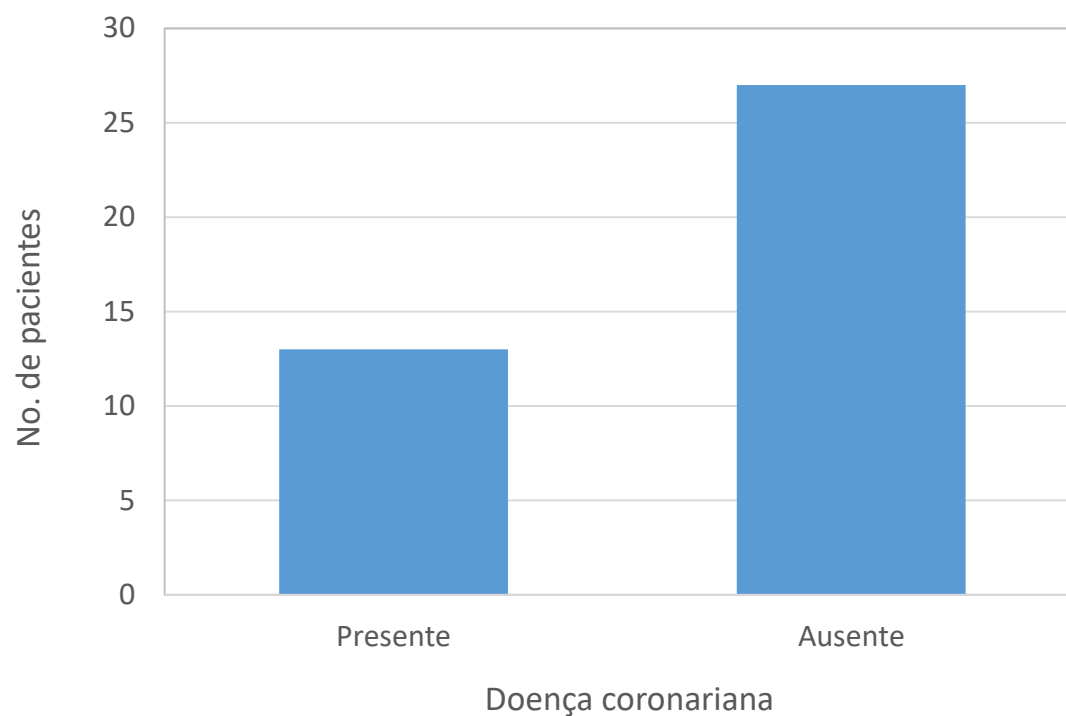


Gráfico em barras



Gráfico 1. Distribuição percentual dos pacientes segundo a presença de doença coronariana . Araraquara, 2015.

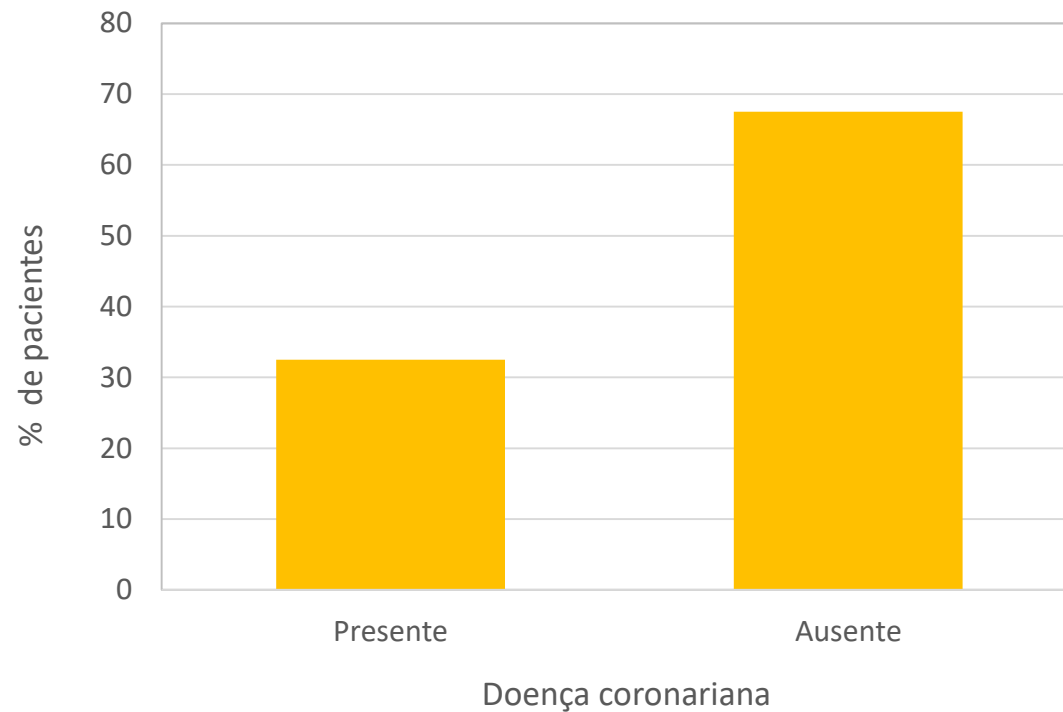


Gráfico em barras

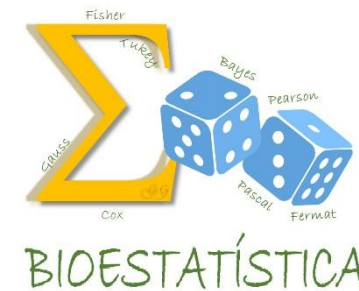


Gráfico 2. Distribuição dos pacientes segundo o sexo.
Araraquara, 2015.

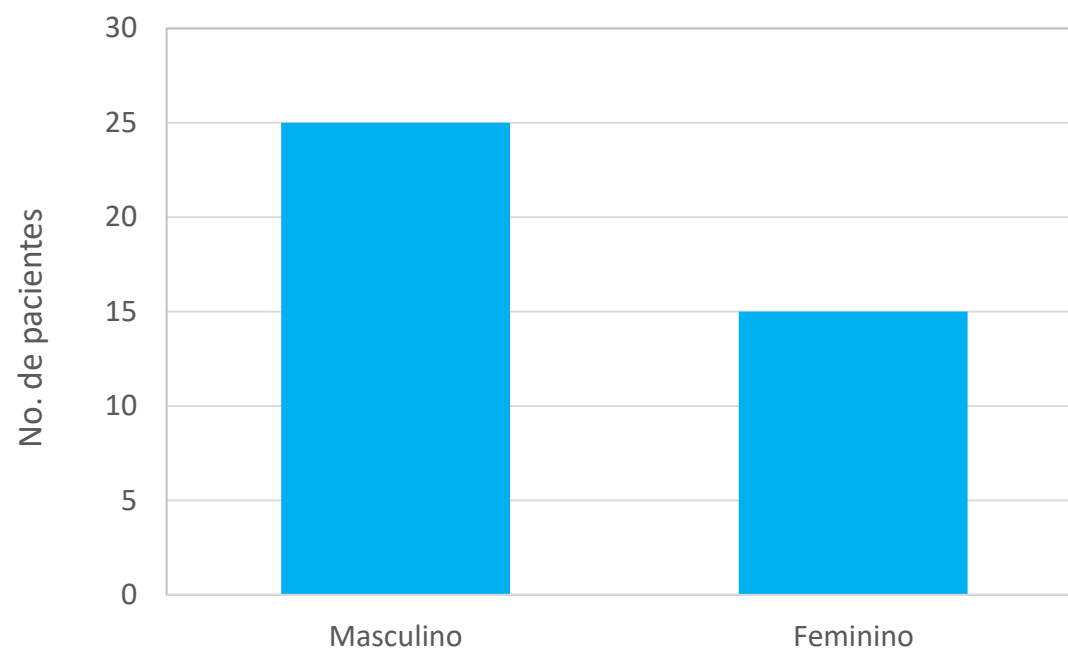


Gráfico em barras

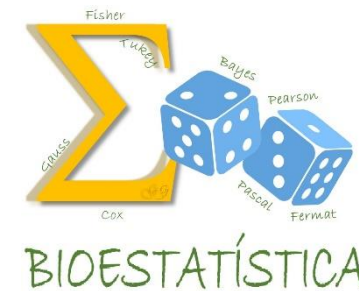
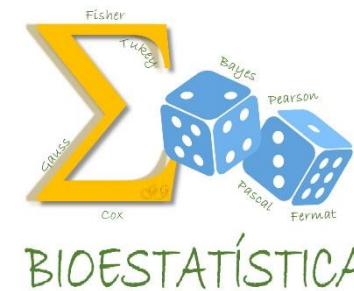


Gráfico 2. Distribuição dos pacientes segundo o sexo.
Araraquara, 2015.



Apresentação Gráfica



Elementos essenciais

- ✓ Título do gráfico (o quê , como, onde, quando)
- ✓ Escala e numeração dos eixos
- ✓ Títulos dos eixos, rótulos das categorias
- ✓ Unidade de medida da variável

Eventualmente, no rodapé

- ✓ Notas, chamadas
- ✓ Fonte

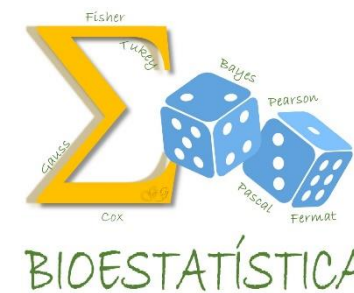
Além disso

- ✓ Não utilizar mais casas decimais do que o necessário
- ✓ Não utilizar mais cores do que o necessário

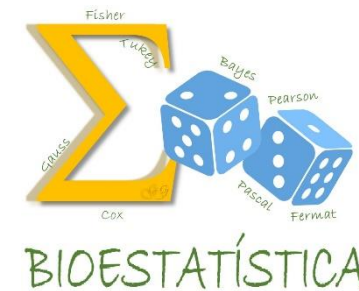
Apresentação Gráfica

Ao analisar um gráfico

- ✓ Evite ler o seu conteúdo, mas aponte os padrões e as principais tendências sugeridas pelos dados
- ✓ Resuma as similaridades e destaque as diferenças



Exercício 3



Segundo o Ministério da Saúde, durante o mês de março de 2015, foram registradas 377 internações por dengue em hospitais do Sistema Único de Saúde (SUS) na região Norte, 1783 no Nordeste, 3600 no Sudeste, 371 no Sul e 1349 no Centro-Oeste.

- a) Construa uma tabela e um gráfico que permita visualizar essas informações adequadamente. Não se esqueça dos cuidados com a apresentação.
- b) Interprete a tabela e o gráfico que você construiu.

Depois, confira os resultados no Moodle.

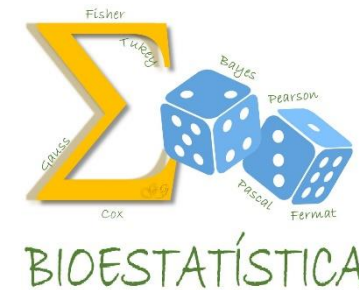
Exercício 3 – a)



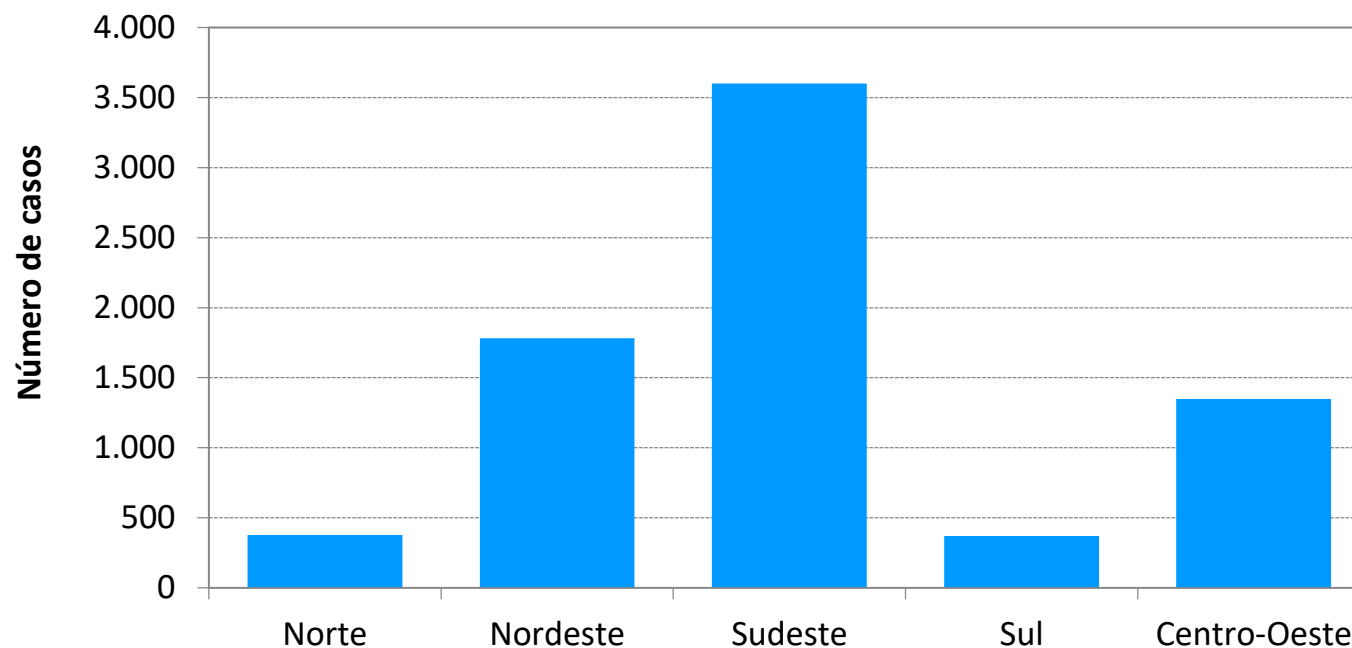
Distribuição dos casos de internação por dengue em hospitais do Sistema Único de Saúde (SUS). Brasil e Regiões, de março de 2015.

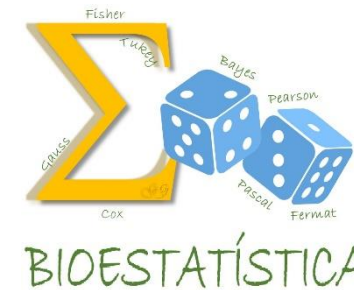
Região	n	%
Norte	377	5,0
Nordeste	1783	23,8
Sudeste	3600	48,1
Sul	371	5,0
Centro-Oeste	1349	18,0
Total	7480	100

Exercício 3 – a)



Distribuição dos casos de internação por dengue em hospitais do Sistema Único de Saúde (SUS). Brasil e Regiões, de março de 2015.

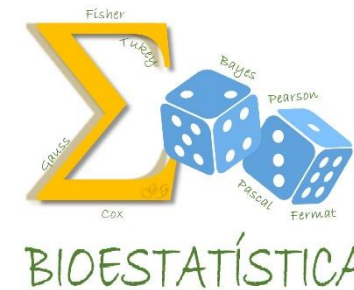




ANÁLISE DESCRITIVA

Apresentando o Diagrama Circular

(veja o vídeo 5 no Moodle)



VÍDEO 5

Apresentando o Diagrama Circular

Diagrama circular

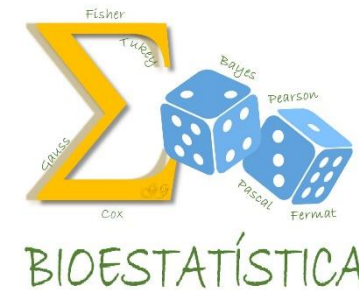


Gráfico 3. Distribuição percentual dos pacientes segundo a presença de doença coronariana . Araraquara, 2015.

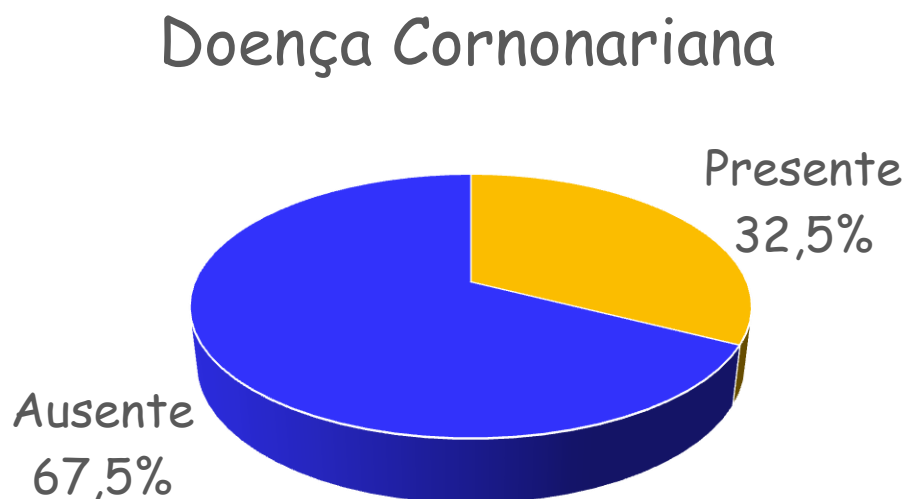


Diagrama circular

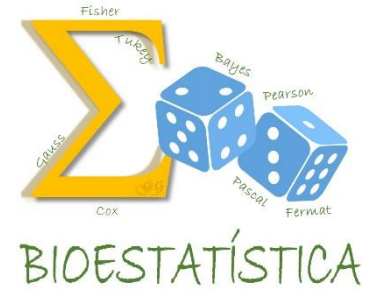


Gráfico 3. Distribuição percentual dos pacientes segundo a presença de doença coronariana . Araraquara, 2015.

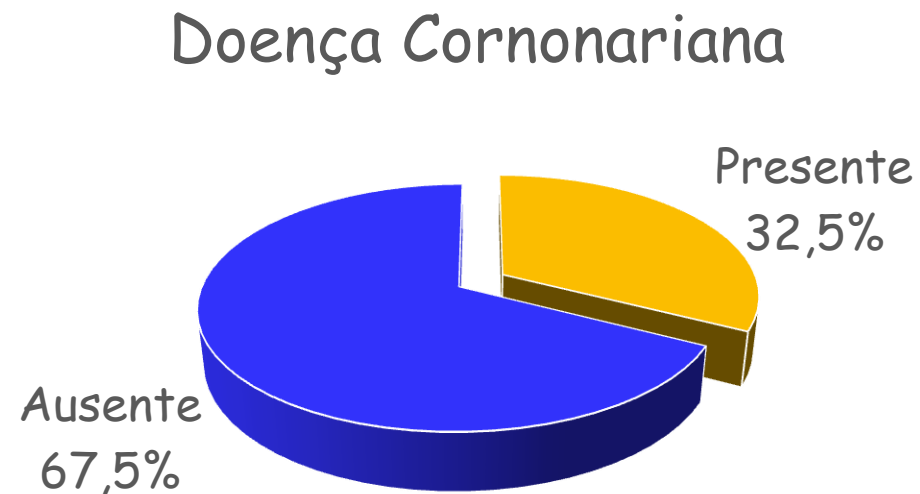


Diagrama circular

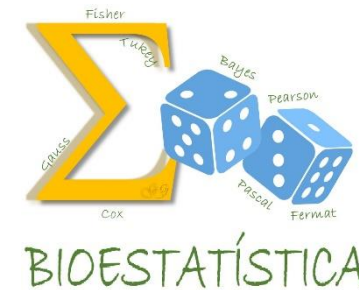
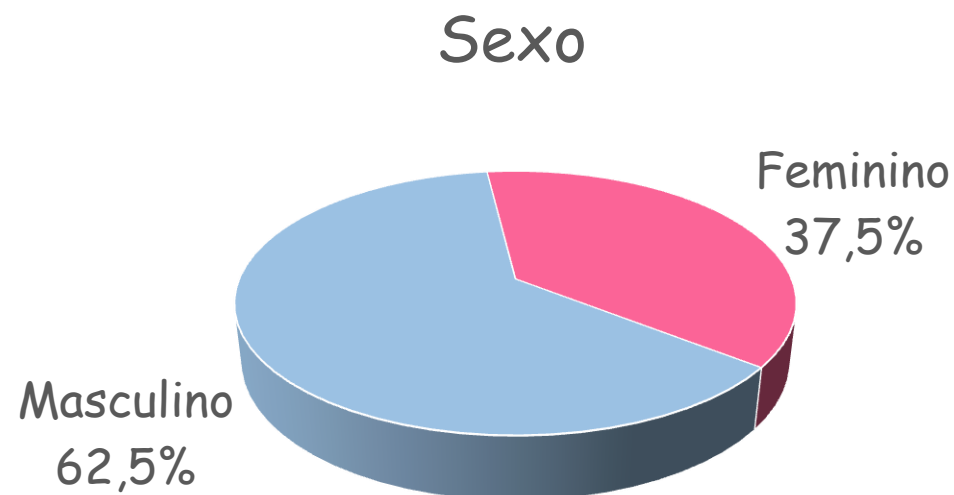
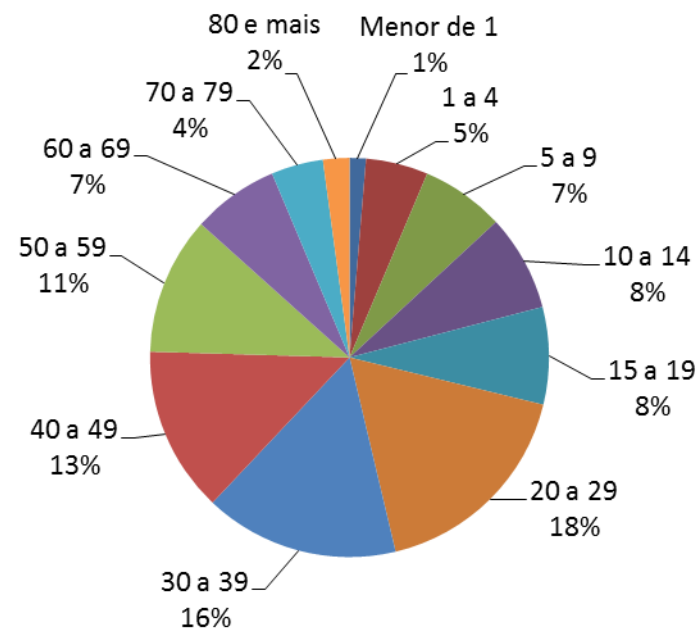


Gráfico 4. Distribuição percentual dos pacientes segundo sexo. Araraquara, 2015.

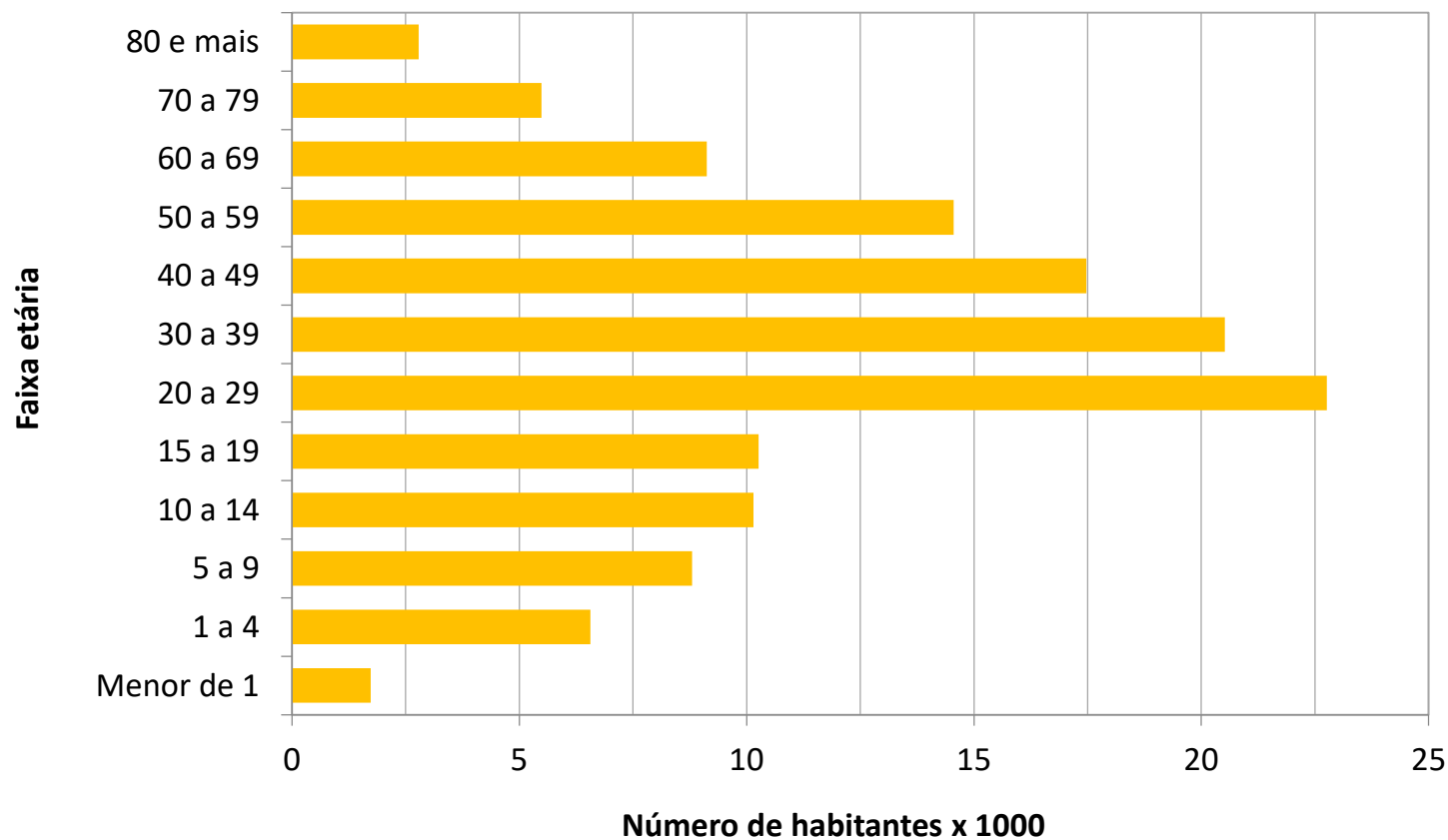


Distribuição percentual de habitantes segundo faixa etária.
Município de Botucatu, 2012.



Difícil de apreender a informação!

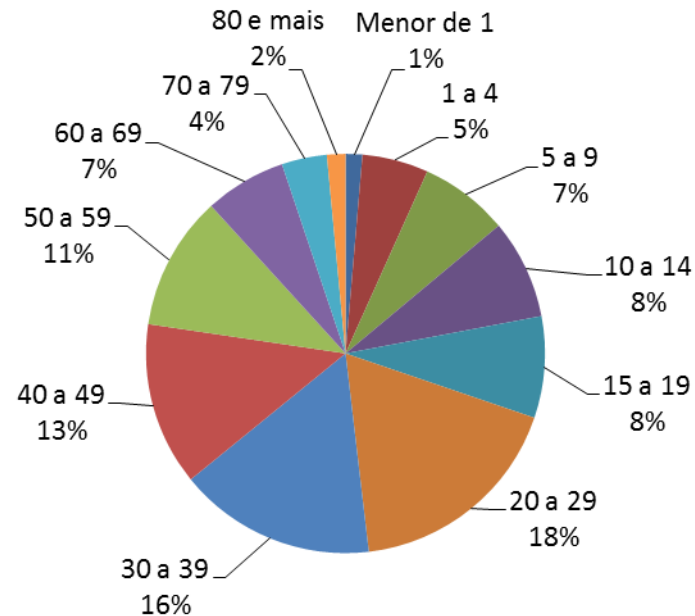
Distribuição percentual de habitantes segundo faixa etária.
Município de Botucatu, 2012.



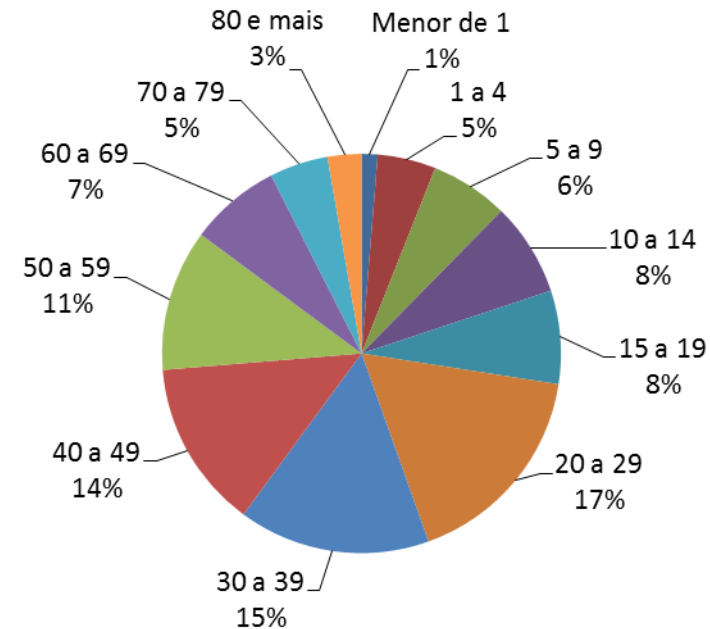
Distribuição percentual de habitantes segundo sexo e faixa etária. Município de Botucatu, 2012.



Masculino

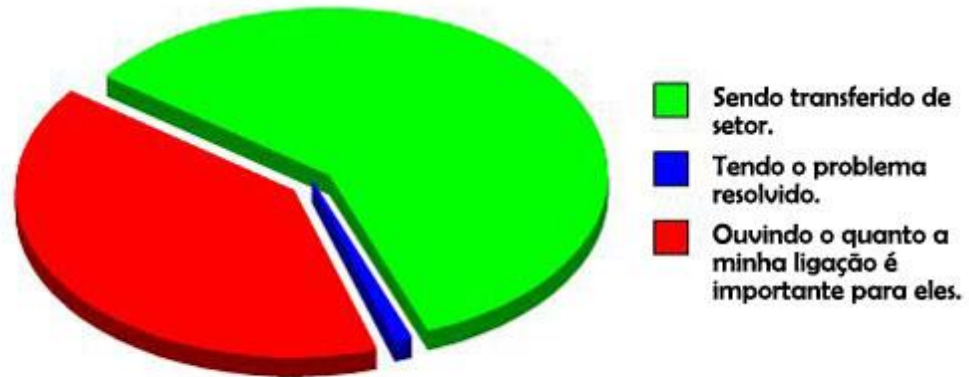


Feminino



Difícil fazer comparações!

Tempo gasto com uma central de atendimento:



GraphJam.com

capinaremos.com



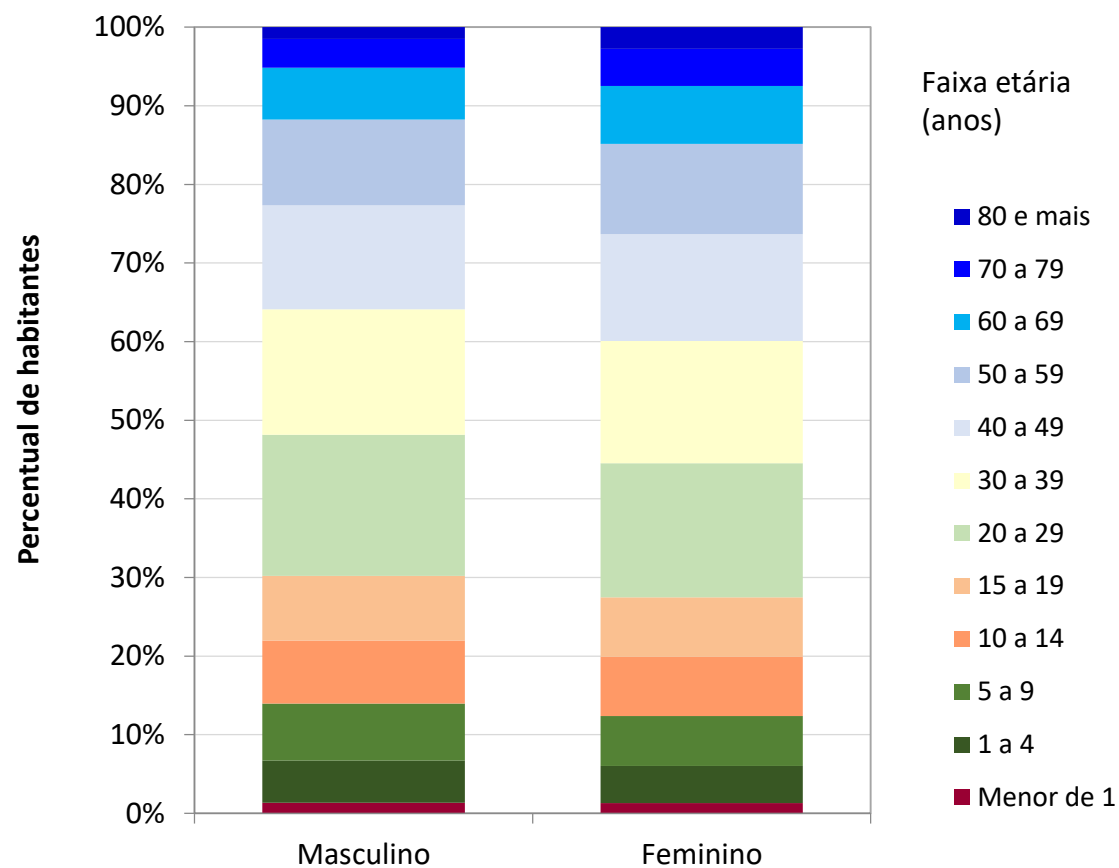
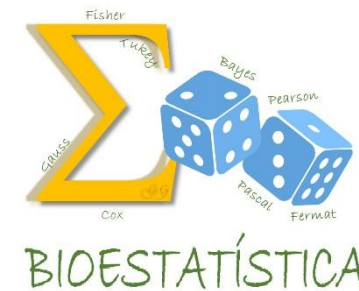
Funções do meu Celular



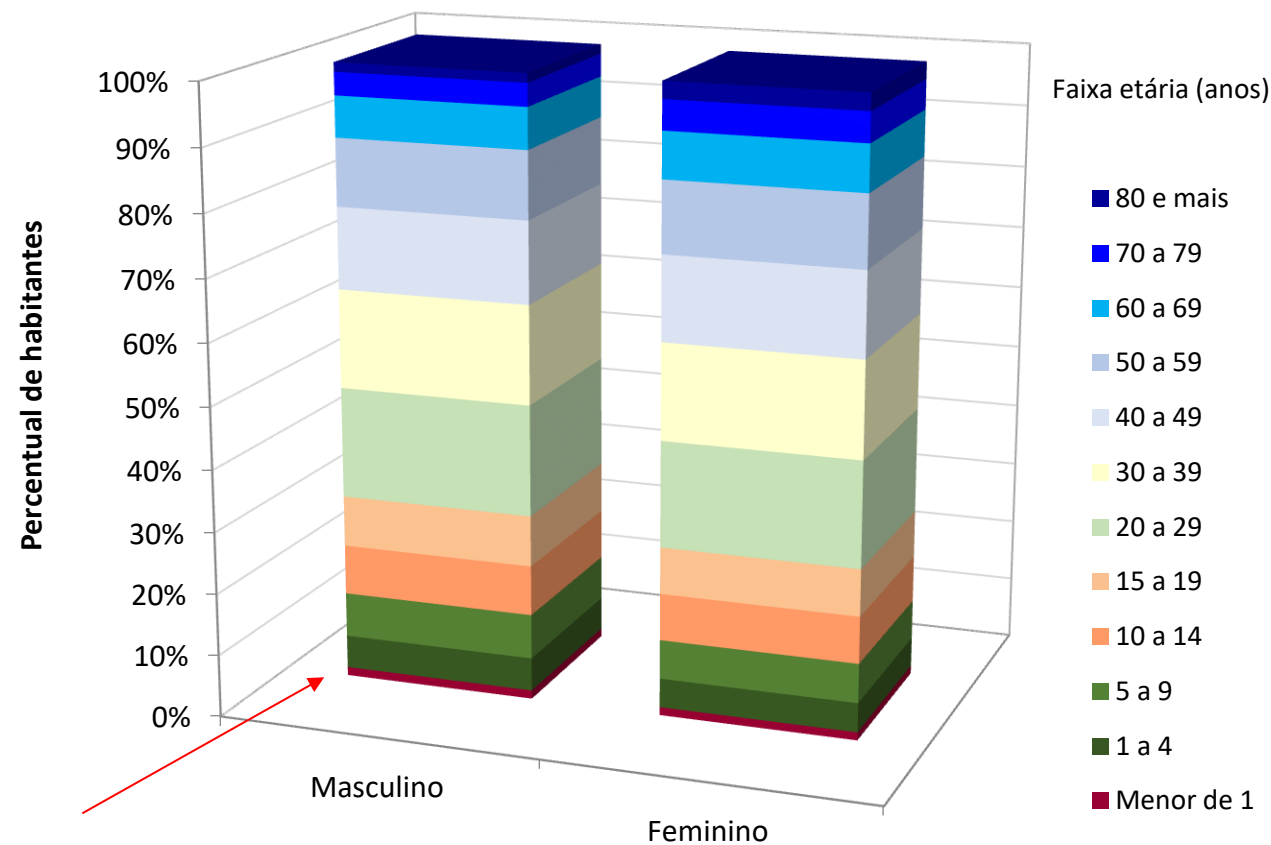
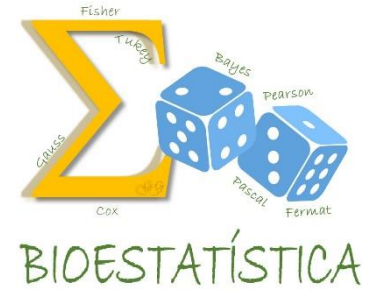
GraphJam.com

capinaremos.com

Distribuição percentual de habitantes segundo sexo e faixa etária. Município de Botucatu, 2012.



Distribuição percentual de habitantes segundo sexo e faixa etária. Município de Botucatu, 2012.



Cuidado com distorções

Cuidado!



BIOESTATÍSTICA

