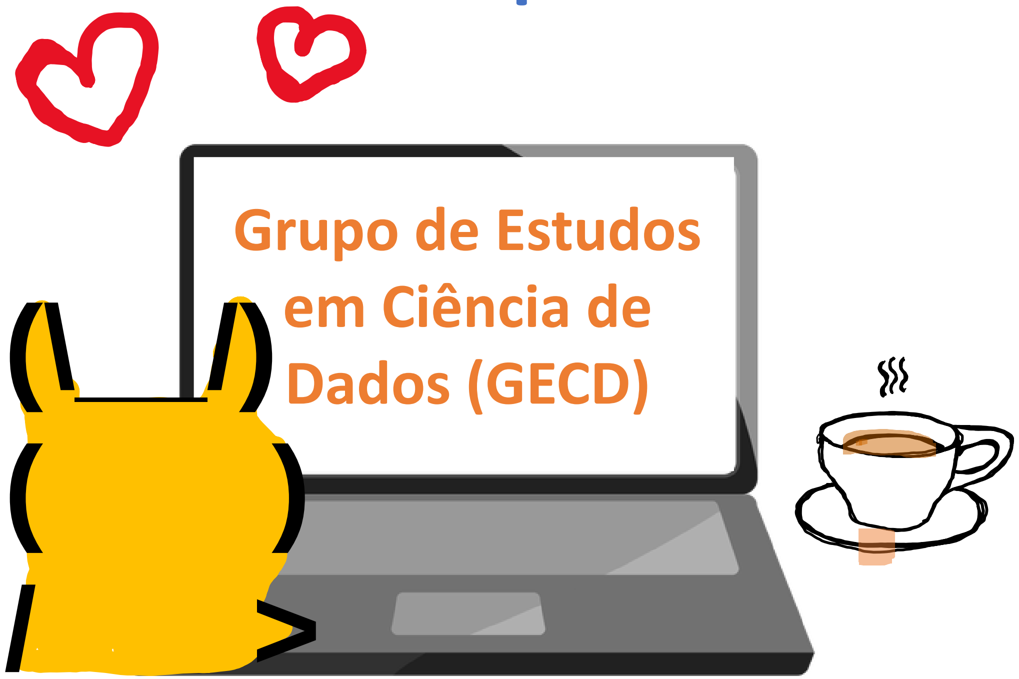
Bem vindo!

歡迎您！

¡Bienvenido!

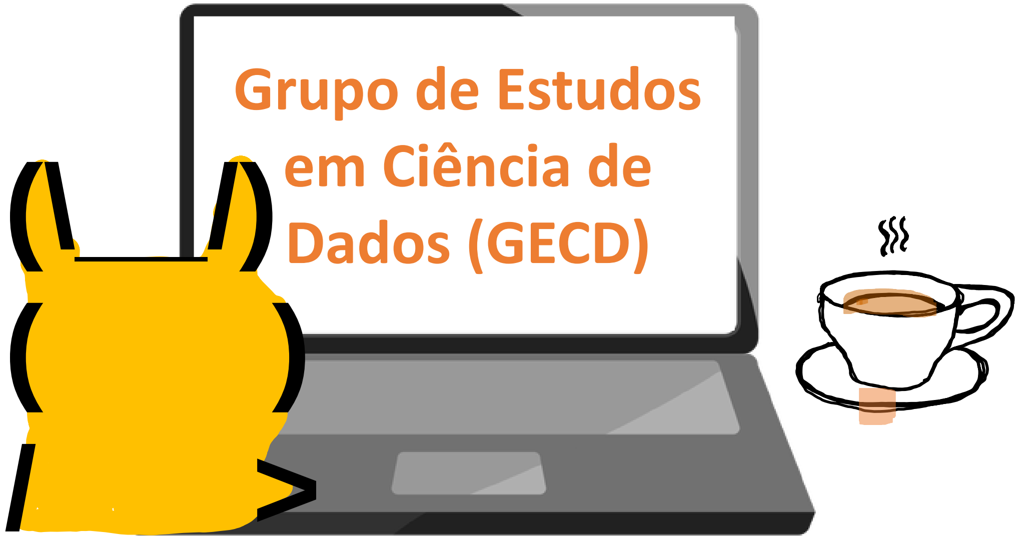
Welcome!



Começaremos em breve

Estatística Descritiva

Prof. Marília Favalesso
E-mail: mariliabioufpr@gmail.com



ESTATÍSTICA, a base da ciência de dados

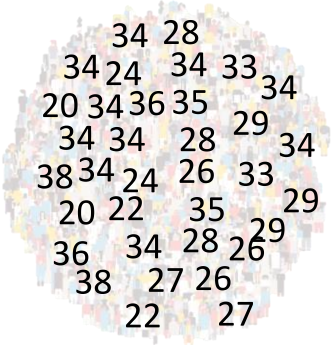
Oferece <u>ferramentas</u> para:

Coleta → Organização → Descrição → Análise → Interpretação de dados

ESTATÍSTICA, a base da ciência de dados

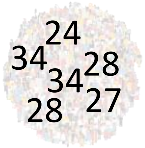
Oferece ferramentas para:

Coleta → Organização → Descrição → Análise → Interpretação de dados



População

Técnica de amostragem



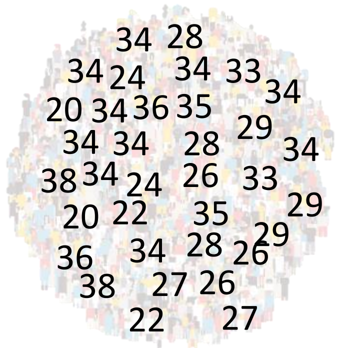
Amostra

Análise descritiva dos dados

- Retirar informações contidas nos dados
- Criar hipóteses para testes/modelos estatísticos

Inferência estatística

Conclusões sobre os nossos dados!
(pode ou não ser extrapolado para a pop)



População

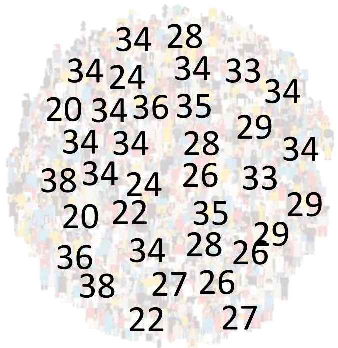
<u>População</u> é um conjunto de elementos que possuem a mesma característica. A população pode ser ampla ou restrita.

Termo meio abstrato!

<u>Depende da dimensão que queremos alcançar</u>

Ex.

Idade média da:
População de cientistas de dados da Usina de Itaipu
População de cientistas de dados da cidade de Foz do Iguaçu
População de cientistas de dados do Paraná
População de cientistas de dados do Brasil
População de cientistas de dados da América Latina
População de cientistas de dados do Mundo!



População

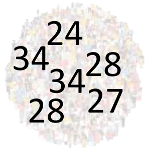
População é um conjunto de elementos que possuem a mesma característica. A população pode ser ampla ou restrita.

Termo meio abstrato!

Depende da dimensão que queremos alcançar

Ex.

Idade média da:
População de cientistas de dados da Usina de Itaipu
População de cientistas de dados da cidade de Foz do Iguaçu
População de cientistas de dados do Paraná
População de cientistas de dados do Brasil
População de cientistas de dados da América Latina
População de cientistas de dados do Mundo!



Amostra

Subconjunto da população, considerando as características essenciais desta.

*Diferentes método para estimar o **n amostral**

Idade média da:
População de cientistas de dados da Usina de Itaipu
População de cientistas de dados da cidade de Foz do Iguaçu
População de cientistas de dados do Paraná
População de cientistas de dados do Brasil
População de cientistas de dados da América Latina
População de cientistas de dados do Mundo!

A *idade* é a **variável** em estudo

Variável:
Atributo, mensurável ou não, sujeito à variação quantitativa ou qualitativa, no interior de um conjunto.

Ex.

A variável idade pode **variar** entre gênero feminino e masculino na população de cientistas de dados da cidade de Foz do Iguaçu

♀	♂
25±6 anos	35±10 anos

Idade média da:
População de cientistas de dados da Usina de Itaipu
População de cientistas de dados da cidade de Foz do Iguaçu
População de cientistas de dados do Paraná
População de cientistas de dados do Brasil
População de cientistas de dados da América Latina
População de cientistas de dados do Mundo!

A *idade* é a **variável** em estudo

Variável:
Atributo, mensurável ou não, sujeito à variação quantitativa ou qualitativa, no interior de um conjunto.

Ex.

A variável idade média dos cientistas de dados pode **variar** entre as cidades de Foz do Iguaçu, Puerto Iguazú e Ciudad del Este.

Foz	**Iguazú**	**Ciudad del Este**
25±6 anos	30±10 anos	35±5 anos

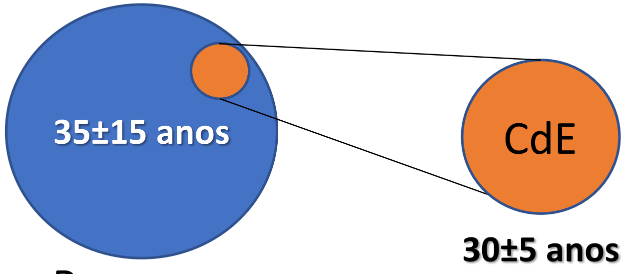
Idade média da:
População de cientistas de dados da Usina de Itaipu
População de cientistas de dados da cidade de Foz do Iguaçu
População de cientistas de dados do Paraná
População de cientistas de dados do Brasil
População de cientistas de dados da América Latina
População de cientistas de dados do Mundo!

A *idade* é a **variável** em estudo

Variável:
Atributo, mensurável ou não, sujeito à variação quantitativa ou qualitativa, no interior de um conjunto de dados.

Ex.

A variável idade média dos cientistas de dados pode variar entre quando comparamos a população de Ciudad del Este com a população de todo o Paraguay.



Paraguay

Lembrando que a **variável sempre é medida em cada unidade experimental** que temos, ou seja, nesse caso eu mensurei a idade de cada um dos cientistas de dados do PY e então calculei a média.

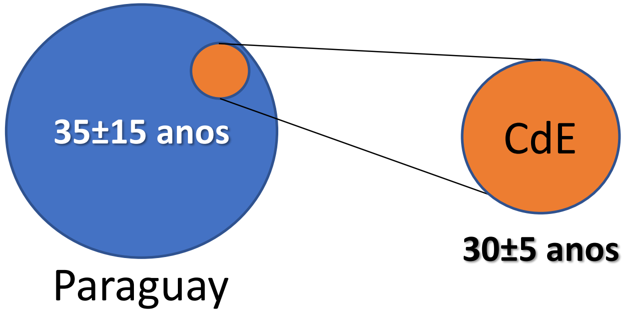
Idade média da:
População de cientistas de dados da Usina de Itaipu
População de cientistas de dados da cidade de Foz do Iguaçu
População de cientistas de dados do Paraná
População de cientistas de dados do Brasil
População de cientistas de dados da América Latina
População de cientistas de dados do Mundo!

A *idade* é a **variável** em estudo

Variável:
Atributo, mensurável ou não, sujeito à variação quantitativa ou qualitativa, no interior de um conjunto.

Ex.

A variável idade média dos cientistas de dados pode variar entre quando comparamos a população de Ciudad del Este com a população de todo o Paraguay.



Tipos de variáveis

Variáveis Quantitativas

Tipos de variáveis

<u>Quantitativa discreta</u>

Os números são inteiros.

Ex. Número de filhos, número de alunos.

Variáveis Quantitativas

Tipos de variáveis

Tipos de variáveis

Variáveis Quantitativas

Quantitativa discreta
Os números são inteiros.
Ex. Número de filhos, número de alunos.

Quantitativa Contínua
Medidas em escalas contínuas; números com casas decimais.
Ex. Idade, altura, peso.

Tipos de variáveis

Variáveis Quantitativas

Quantitativa discreta
Os números são inteiros.
Ex. Número de filhos, número de alunos.

Quantitativa Contínua
Medidas em escalas contínuas; números com casas decimais.
Ex. Idade, altura, peso.

Variáveis Qualitativas

Tipos de variáveis

Variáveis Quantitativas

Quantitativa discreta

Os números são inteiros.
Ex. Número de filhos, número de alunos.

Quantitativa Contínua

Medidas em escalas contínuas; números com casas decimais.
Ex. Idade, altura, peso.

Variáveis Qualitativas

Qualitativa nominal

Dicotômica: Possui apenas duas respostas.
Ex. Sim/Não; Presença/Ausência.

Categórica: Mais de duas respostas; não existe ordem entre categorias.
Ex. Estado conjugal (solteiro, casado divorciado); país (AR, BR, PY, etc.).

Tipos de variáveis

Variáveis Quantitativas

Quantitativa discreta

Os números são inteiros.
Ex. Número de filhos, número de alunos.

Quantitativa Contínua

Medidas em escalas contínuas; números com casas decimais.
Ex. Idade, altura, peso.

Variáveis Qualitativas

Qualitativa nominal

Dicotômica: Possui apenas duas respostas.
Ex. Sim/Não; Presença/Ausência.

Categórica: Mais de duas respostas; não existe ordem entre categorias.
Ex. Estado conjugal (solteiro, casado divorciado); país (AR, BR, PY, etc.).

Qualitativa ordinal

Quando existe uma ordem ou grau nos possíveis resultados.
Ex.: Série escolar (1ª série, 2ª série etc).

Como tabulamos nossas variáveis para uso em R e também em Python?

Uma variável por coluna!!

Cada unidade experimental em uma linha (ex. cada indivíduo em uma linha).

	A	B	C	D
1	**ID**	**Variável A**	**Variável B**	**Variável C**
2	Indivíduo 1	Valor para indivíduo 1	Valor para indivíduo 1	Valor para indivíduo 1
3	Indivíduo 2	Valor para indivíduo 2	Valor para indivíduo 2	Valor para indivíduo 2
4	Indivíduo 3	Valor para indivíduo 3	Valor para indivíduo 3	Valor para indivíduo 3
5	Indivíduo 4	Valor para indivíduo 4	Valor para indivíduo 4	Valor para indivíduo 4
6	Indivíduo 5	Valor para indivíduo 5	Valor para indivíduo 5	Valor para indivíduo 5
7	Indivíduo 6	Valor para indivíduo 6	Valor para indivíduo 6	Valor para indivíduo 6
8	...	...	...	...
9				

Como tabulamos nossas variáveis para uso em R e também em Python?

Uma variável por coluna!!

Cada <u>unidade experimental</u> em uma linha (ex. cada indivíduo em uma linha).

	A	B	C	D
1	**ID**	**Gênero**	**Idade**	**Escolaridade**
2	Amanda	Feminino	25	Ensino médio completo
3	Felipe	Masculino	30	Ensino superior incompleto
4	Fernando	Masculino	29	Ensino superior completo
5	Alicia	Feminino	31	Ensino médio incompleto
6	Jeferson	Masculino	27	Pós-graduação incompleto
7	Alejandro	Masculino	26	Ensino superior incompleto
8	...	...	...	...

Introduzindo a **estatística descritiva**

- Apenas o início!

- Dados nunca são apresentados de maneira integral, ao invés disso, resumimos nossos dados utilizado a **estatística descritiva**.

- Para as variáveis qualitativas, usamos as **distribuições de frequência**.

- Para as variáveis quantitativas, podemos descrever os dados pelas **Medidas de tendência central** e pelas **Medidas de variabilidade**.

Distribuição de frequências

Distribuição de frequências

- Em geral, <u>dados qualitativos</u> são representados por sua frequência absoluta e sua frequência relativa percentual.

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Estado civil	**f**	**FR**	**FR%**
Solteiro/a			
Casado/a			
Viúvo/a			

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência absoluta (f): Simplesmente o número de unidades por categoria da minha variável

Estado civil	f	FR	FR%
Solteiro/a			
Casado/a			
Viúvo/a			

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência absoluta (f): Simplesmente o número de unidades por categoria da minha variável

Estado civil	f	FR	FR%
Solteiro/a	10		
Casado/a			
Viúvo/a			

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência absoluta (f): Simplesmente o número de unidades por categoria da minha variável

Estado civil	f	FR	FR%
Solteiro/a	10		
Casado/a	20		
Viúvo/a			

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência absoluta (f): Simplesmente o número de unidades por categoria da minha variável

Estado civil	f	FR	FR%
Solteiro/a	10		
Casado/a	20		
Viúvo/a	5		

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa (FR): É simplesmente a proporção em que cada valor ocorre. Sua fórmula é f/Σ(f).

Estado civil	f	FR	FR%
Solteiro/a	10		
Casado/a	20		
Viúvo/a	5		

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa (FR): É simplesmente a proporção em que cada valor ocorre. Sua fórmula é f/$\boxed{\Sigma(f)}$.

Estado civil	f	FR	FR%
Solteiro/a	10		
Casado/a	20		
Viúvo/a	5		

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa (FR): É simplesmente a proporção em que cada valor ocorre. Sua fórmula é $f/\boxed{\Sigma(f)}$.

Estado civil	f	FR	FR%
Solteiro/a	10	10/35=	
Casado/a	20		
Viúvo/a	5		

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa (FR): É simplesmente a proporção em que cada valor ocorre. Sua fórmula é f/$\boxed{\Sigma(f)}$.

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	
Casado/a	20		
Viúvo/a	5		

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa (FR): É simplesmente a proporção em que cada valor ocorre. Sua fórmula é f/$\boxed{\Sigma(f)}$.

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	
Casado/a	20	20/35=	
Viúvo/a	5		

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa (FR): É simplesmente a proporção em que cada valor ocorre. Sua fórmula é f/$\boxed{\Sigma(f)}$.

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	
Casado/a	20	20/35= **0.5714**	
Viúvo/a	5		

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa (FR): É simplesmente a proporção em que cada valor ocorre. Sua fórmula é f/$\boxed{\Sigma(f)}$.

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	
Casado/a	20	20/35= **0.5714**	
Viúvo/a	5	5/35=	

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa (FR): É simplesmente a proporção em que cada valor ocorre. Sua fórmula é f/$\boxed{\Sigma(f)}$.

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	
Casado/a	20	20/35= **0.5714**	
Viúvo/a	5	5/35= **0.1429**	

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa percentual (FR%): É simplesmente a proporção em que cada valor ocorre multiplicado por 100. Sua fórmula é (f/Σ(f))*100.

Ou FR*100

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	
Casado/a	20	20/35= **0.5714**	
Viúvo/a	5	5/35= **0.1429**	

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa percentual (FR%): É simplesmente a proporção em que cada valor ocorre multiplicado por 100. Sua fórmula é (f/Σ(f))*100.

Ou FR*100

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	0.2851*100=
Casado/a	20	20/35= **0.5714**	
Viúvo/a	5	5/35= **0.1429**	

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa percentual (FR%): É simplesmente a proporção em que cada valor ocorre multiplicado por 100. Sua fórmula é (f/Σ(f))*100.

Ou FR*100

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	0.2851*100= **28.51%**
Casado/a	20	20/35= **0.5714**	
Viúvo/a	5	5/35= **0.1429**	

Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa percentual (FR%): É simplesmente a proporção em que cada valor ocorre multiplicado por 100. Sua fórmula é (f/Σ(f))*100.

Ou FR*100

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	0.2851*100= **28.51%**
Casado/a	20	20/35= **0.5714**	0.5714*100=
Viúvo/a	5	5/35= **0.1429**	

Distribuição de frequências

Ex.

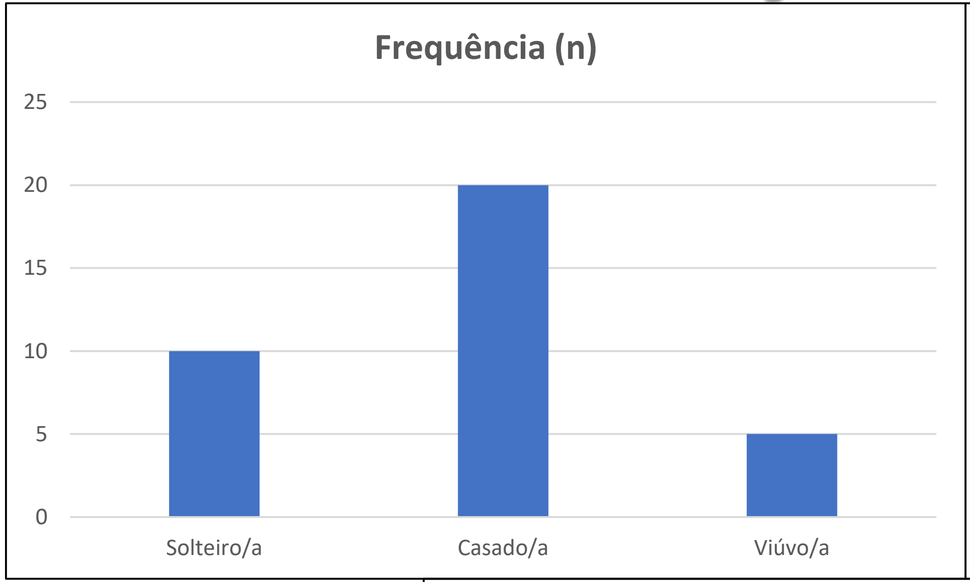
Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

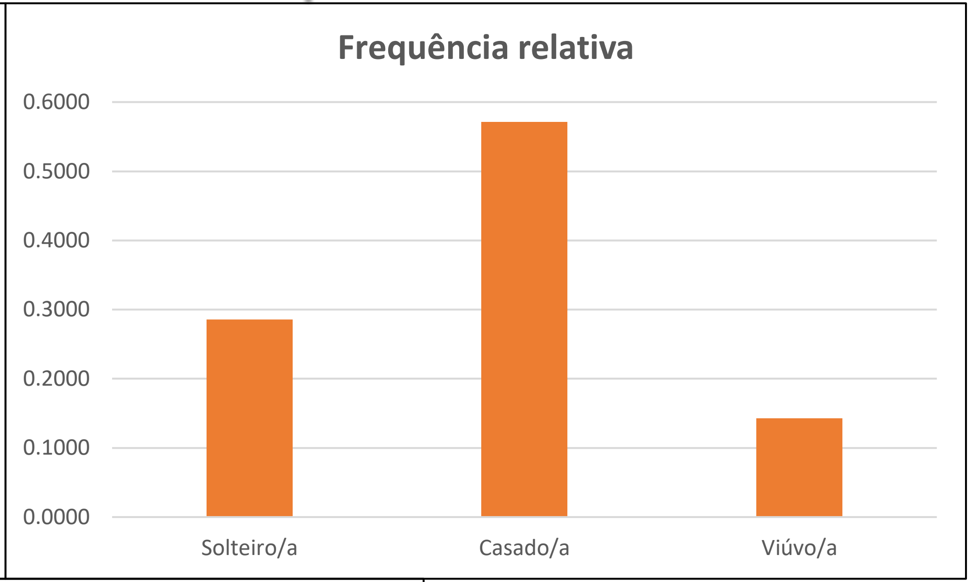
Frequência relativa percentual (FR%): É simplesmente a proporção em que cada valor ocorre multiplicado por 100. Sua fórmula é (f/Σ(f))*100.

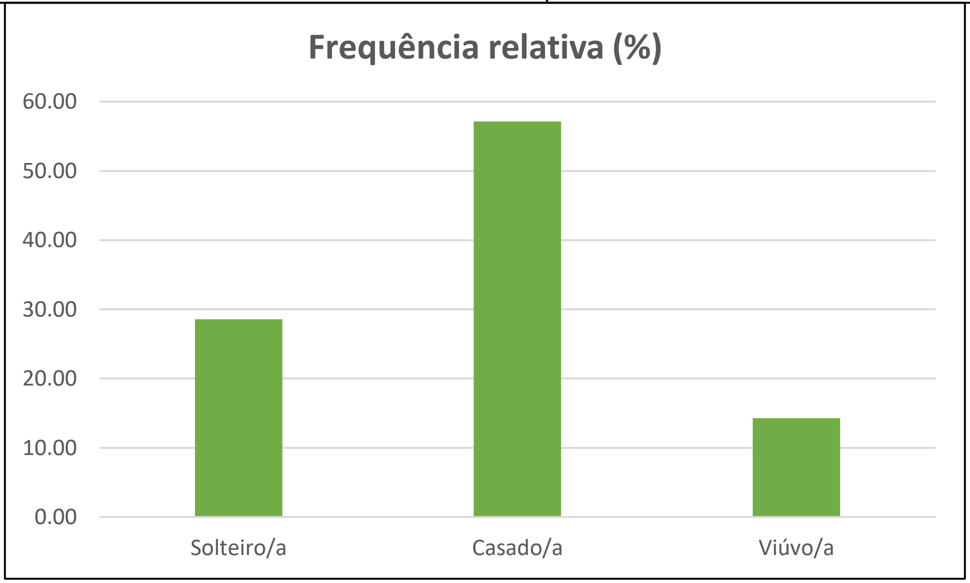
Ou FR*100

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	0.2851*100= **28.51%**
Casado/a	20	20/35= **0.5714**	0.5714*100= **57.14%**
Viúvo/a	5	5/35= **0.1429**	

Distribuição de frequências







Distribuição de frequências

Ex.

Realizei uma amostragem de dados relativo a estado civil de 35 alunos do GECD de Foz do Iguaçu. Agora eu preciso descrever esses dados.

Frequência relativa percentual (FR%): É simplesmente a proporção em que cada valor ocorre multiplicado por 100. Sua fórmula é (f/Σ(f))*100.

Ou FR*100

Estado civil	f	FR	FR%
Solteiro/a	10	10/35= **0.2851**	0.2851*100= **28.51%**
Casado/a	20	20/35= **0.5714**	0.5714*100= **57.14%**
Viúvo/a	5	5/35= **0.1429**	0.1429*100= **14.29%**

Distribuição de frequências

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, $Fai= f1+f2+f3+...+fi$.

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	**Frequência (f) de exercícios**	**FA**
10	10	
20	10	
30	10	
40	10	
50	10	
60	10	

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=
20	10	
30	10	
40	10	
50	10	
60	10	

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	$\boxed{10}$	FA(10m)=
20	10	
30	10	
40	10	
50	10	
60	10	

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=
20	10	
30	10	
40	10	
50	10	
60	10	

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	10	
30	10	
40	10	
50	10	
60	10	

Distribuição de frequências

Ex.
Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	**Frequência (f) de exercícios**	**FA**
10	10	FA(10m)=**10 ex.**
20	10	FA(20m)=
30	10	
40	10	
50	10	
60	10	

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	10	FA(20m)=
30	10	
40	10	
50	10	
60	10	

+

Σ(n.exe10 + n.exe20)

10 + 10

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	+ 10	FA(20m)= **10+10=**
30	10	
40	10	
50	10	
60	10	

$\Sigma(n.\,exe_{10} + n.\,exe_{20})$

$10 + 10$

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	10	FA(20m)= **10+10 = 20ex.**
30	10	
40	10	
50	10	
60	10	

+

$\Sigma(n.exe_{10} + n.exe_{20})$

$10 + 10$

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	10	FA(20m)= **10+10= 20ex.**
30	10	FA(30m)= **10+10+10=**
40	10	
50	10	
60	10	

+

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	+ 10	FA(20m)= **10+10= 20ex.**
30	10	FA(30m)= **10+10+10=**
40	10	
50	10	
60	10	

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	10	FA(20m)= **10+10= 20ex.**
30	10	FA(30m)= **10+10+10= 30ex.**
40	10	
50	10	
60	10	

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	10	FA(20m)= **10+10= 20ex.**
30	10	FA(30m)= **10+10+10= 30ex.**
40	10	
50	10	
60	10	

(...)

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	10	FA(20m)= **10+10= 20ex.**
30	10	FA(30m)= **10+10+10= 30ex.**
40	10	FA(40m) =**10+10+10+10= 40ex.**
50	10	FA(50m)= **10+10+10+10+10= 50ex.**
60	10	FA(60m)= **10+10+10+10+10+10= 60ex.**

(...)

Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o total corrente) de todas as frequências até o ponto presente no conjunto de dados, ou seja, Fai= f1+f2+f3+...+fi.

Tempo	Frequência (f) de exercícios	FA
10	10	FA(10m)=**10 ex.**
20	10	FA(20m)= **10+10= 20ex.**
30	10	FA(30m)= **10+10+10= 30ex.**
40	10	FA(40m) =**10+10+10+10= 40ex.**
50	10	FA(50m)= **10+10+10+10+10= 50ex.**
60	10	FA(60m)= **10+10+10+10+10+10= 60ex.**

(...)

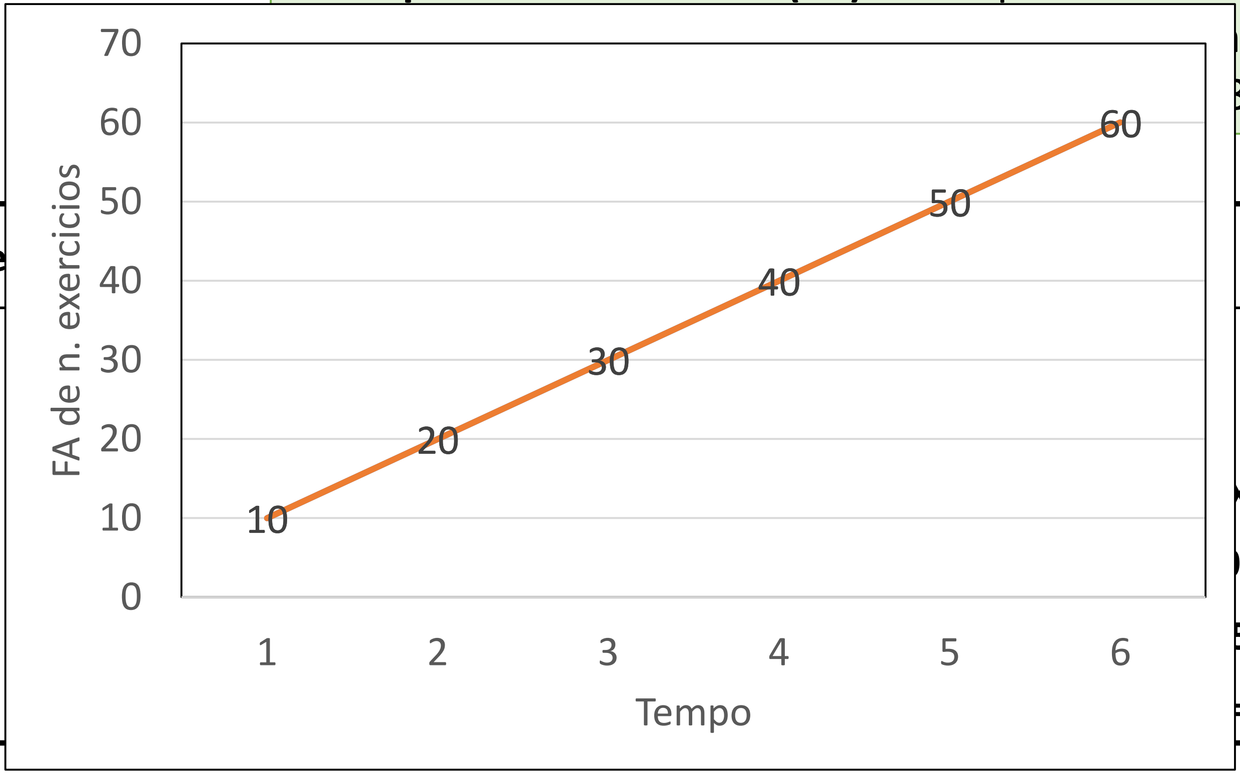
Distribuição de frequências

Ex.

Quantos exercícios de estatística um aluno do GECD consegue responder em um período de 60m?

Frequência acumulada (FA): ela representa a soma (ou o
nto presente
3+...+fi.

Te



x.

0ex.

50ex.

= 60ex.

(...)

Introduzindo a estatística descritiva

- Apenas o início!

- Dados nunca são apresentados de maneira integral, ao invés disso, resumimos nossos dados utilizado a **estatística descritiva**.

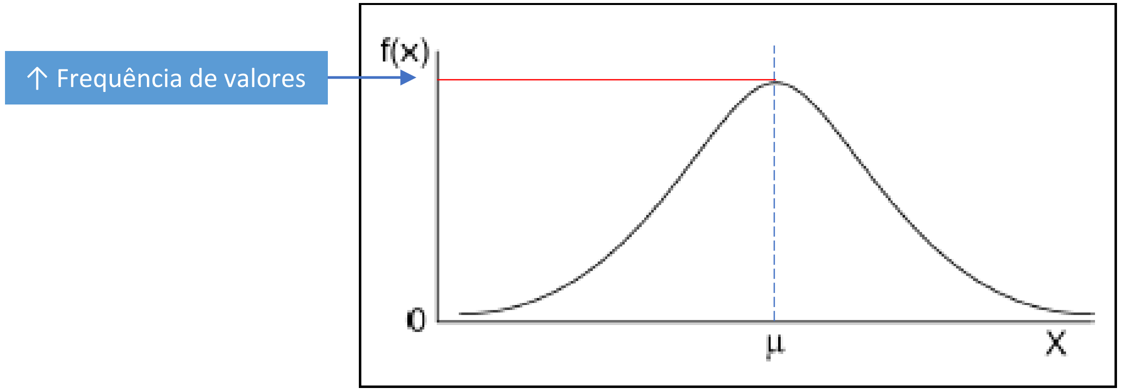
- Para as variáveis qualitativas, usamos as **distribuições de frequência**.

- Para as variáveis quantitativas, podemos descrever os dados pelas **Medidas de tendência central** e pelas **Medidas de variabilidade**.

Medidas de tendência central

Média aritmética

- Os dados de uma população para uma determinada variável ‘x’ normalmente não se distribuem uniformemente, havendo uma certa concentração.



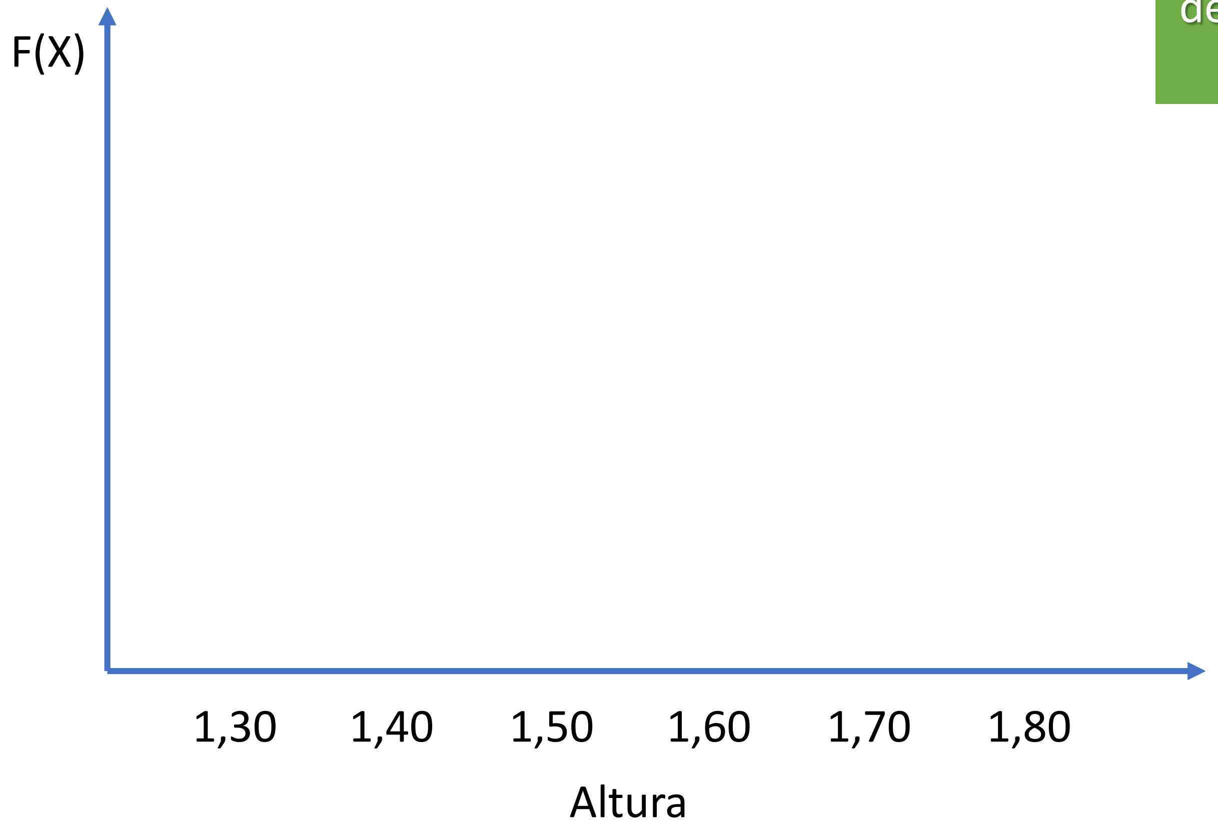
Curva de distribuição normal

- Essa concentração de dados pode ser utilizada para representar os dados (mesmo que esse valor seja uma abstração).

Medidas de tendência central

Curva normal

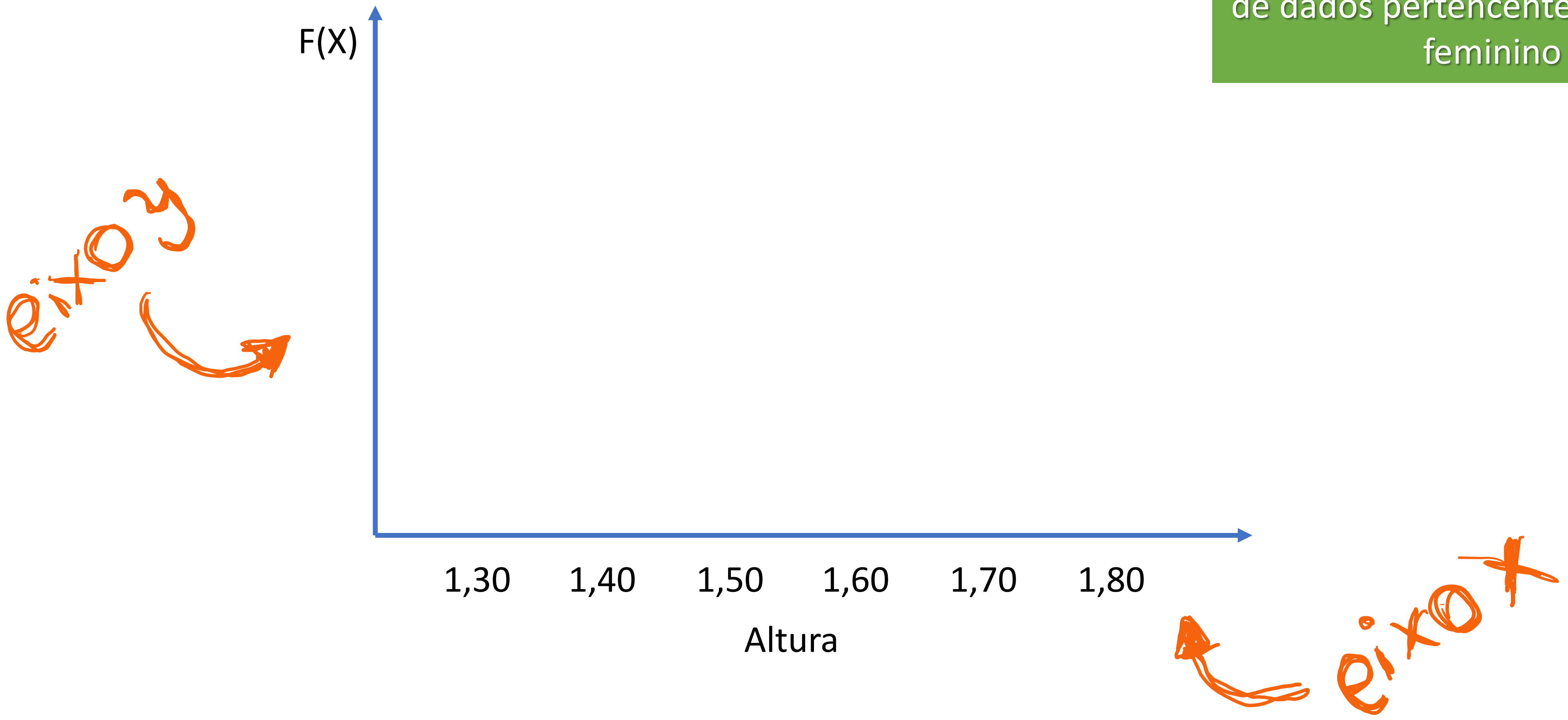
Altura da população de cientistas de dados pertencentes ao gênero feminino



Medidas de tendência central

Curva normal

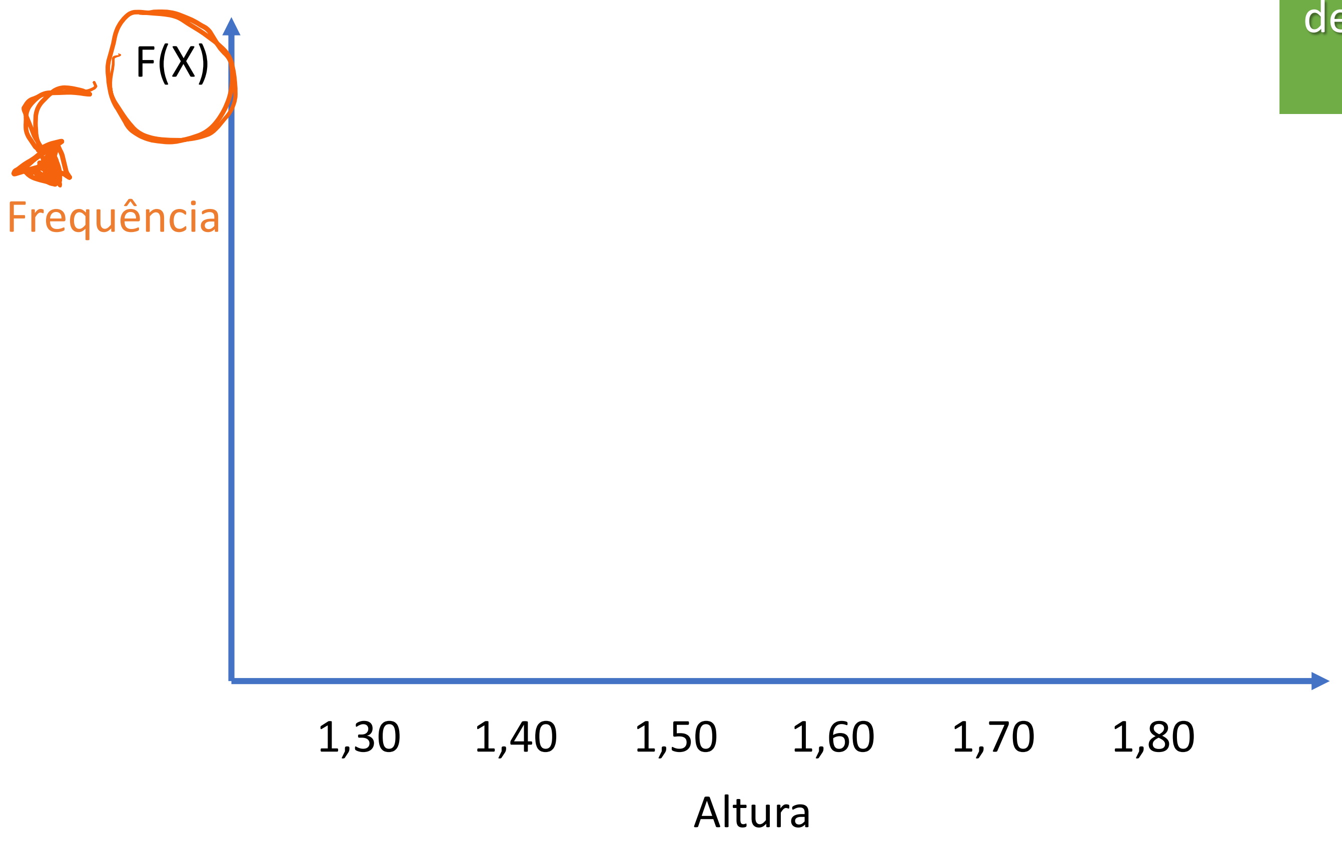
Altura da população de cientistas de dados pertencentes ao gênero feminino



Medidas de tendência central

Curva normal

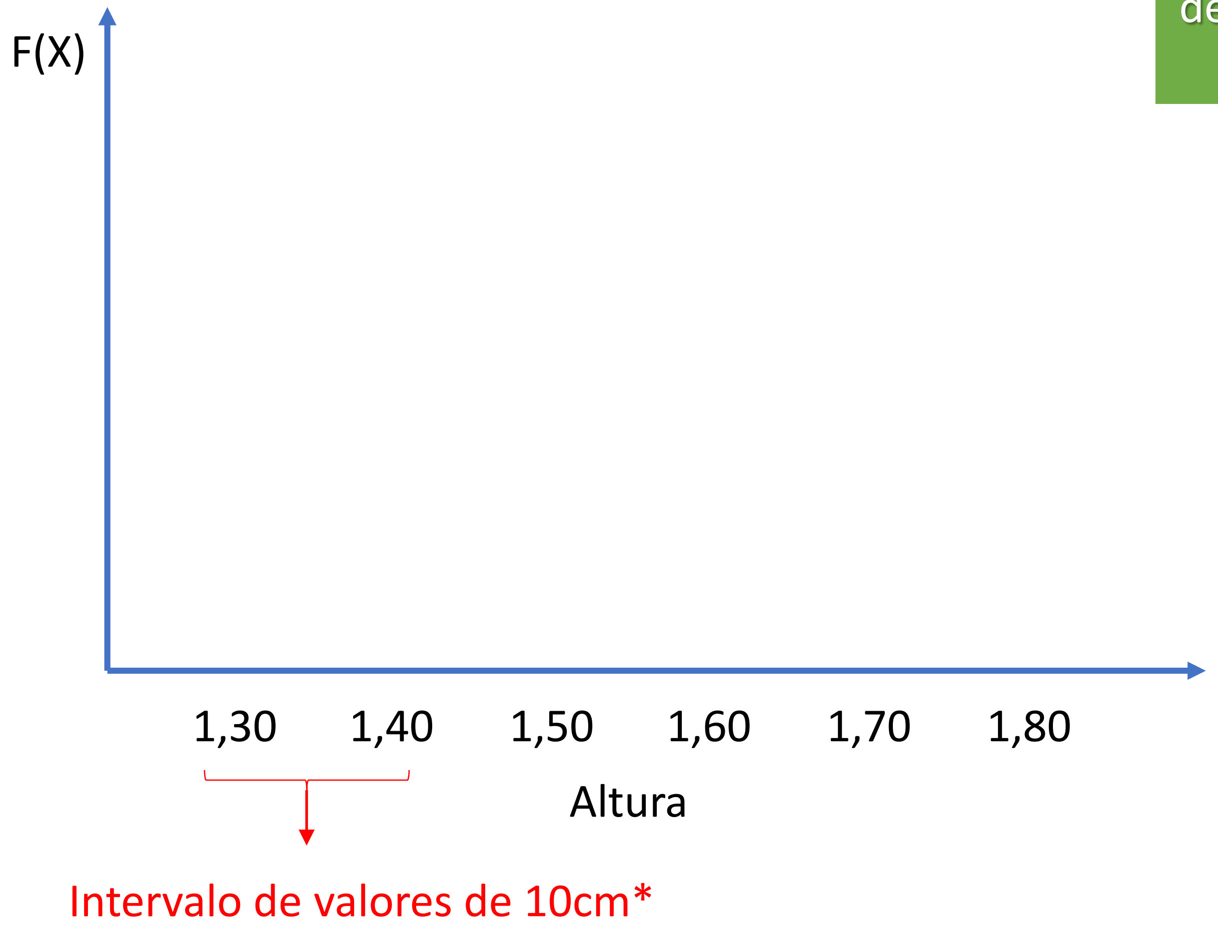
Altura da população de cientistas de dados pertencentes ao gênero feminino



Medidas de tendência central

Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

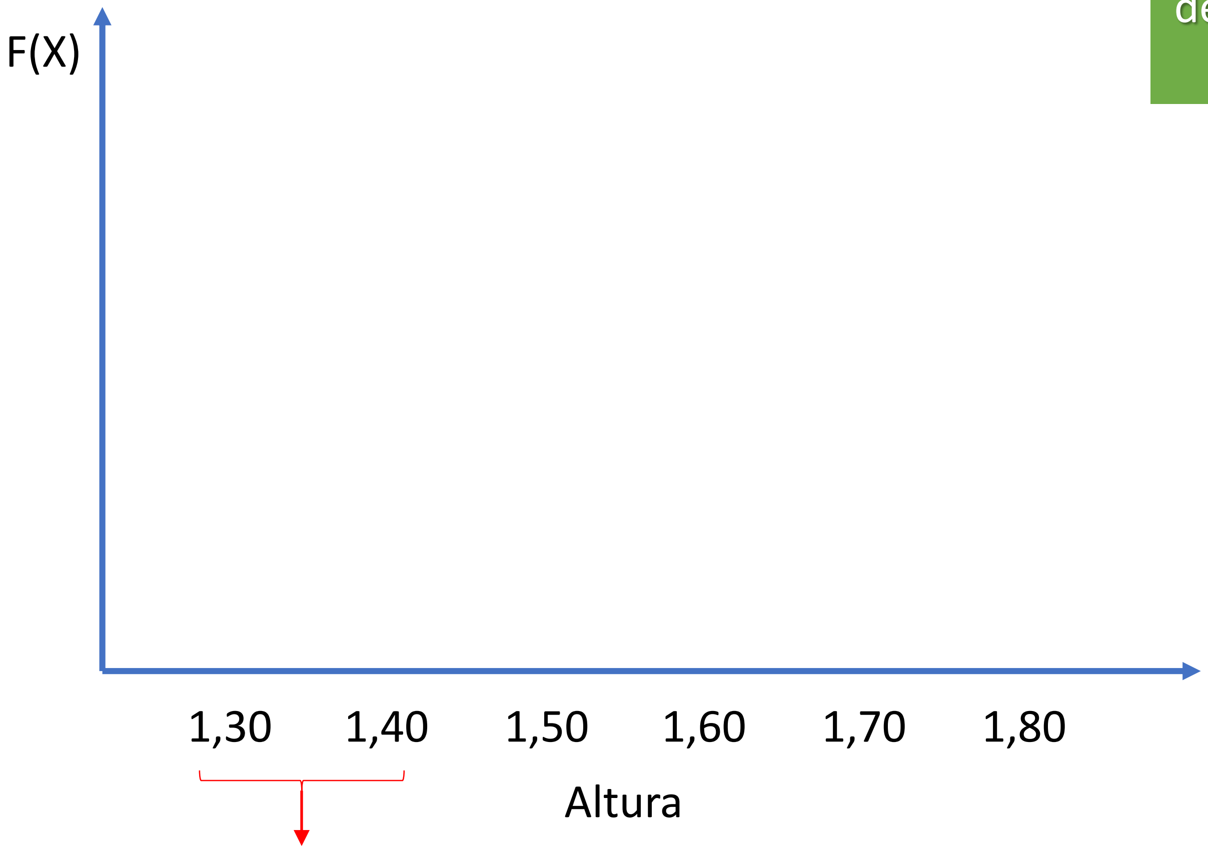


Medidas de tendência central

Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?



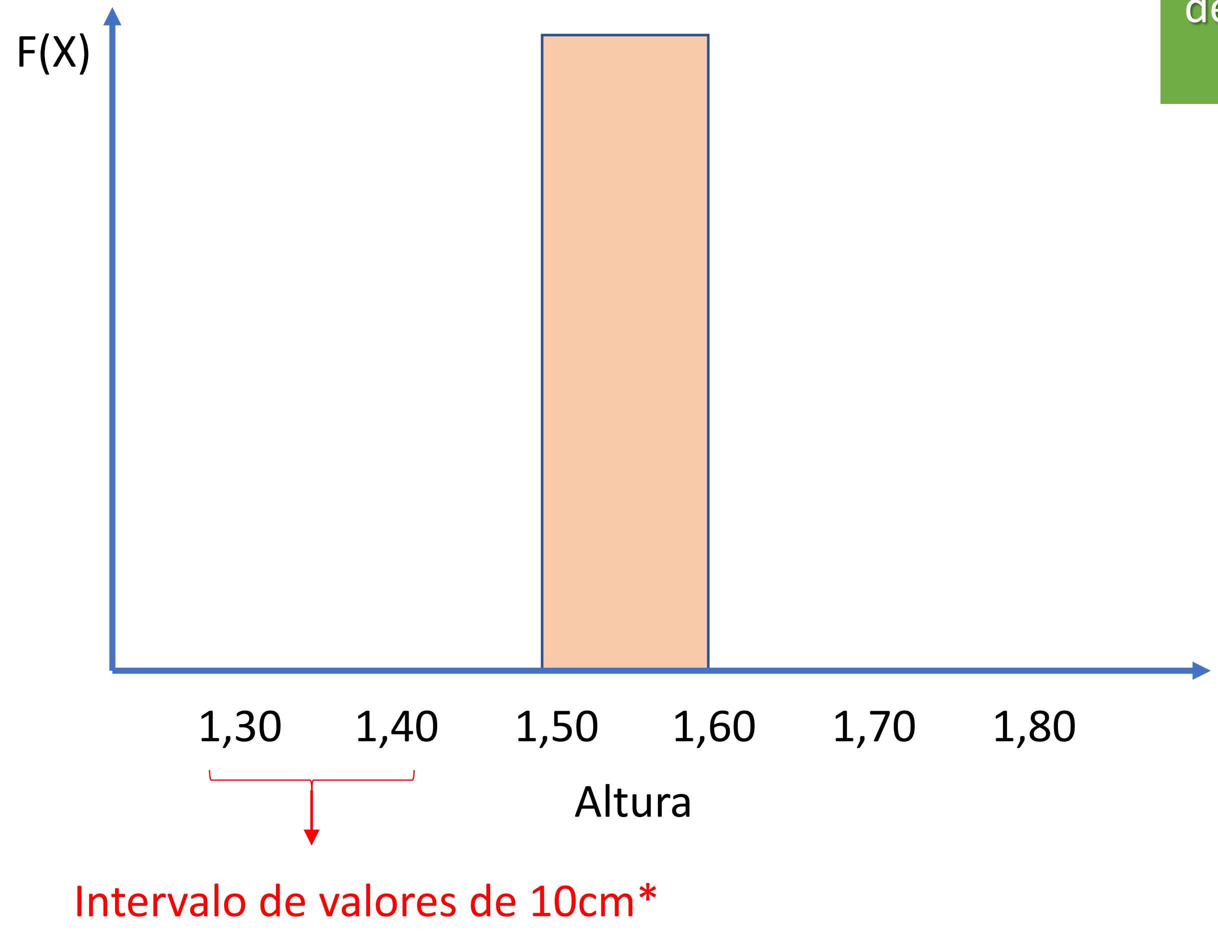
Intervalo de valores de 10cm*

Medidas de tendência central

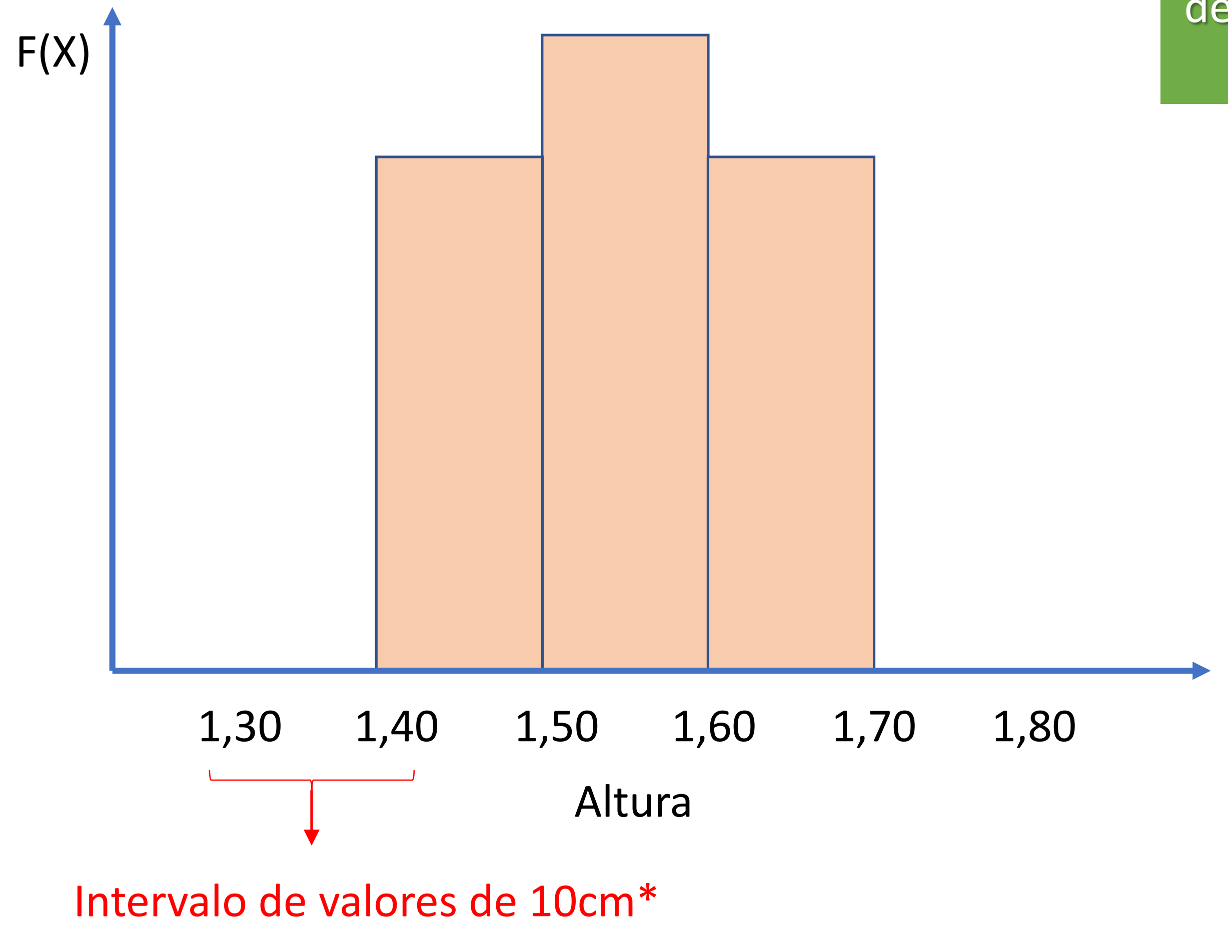
Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?



Medidas de tendência central

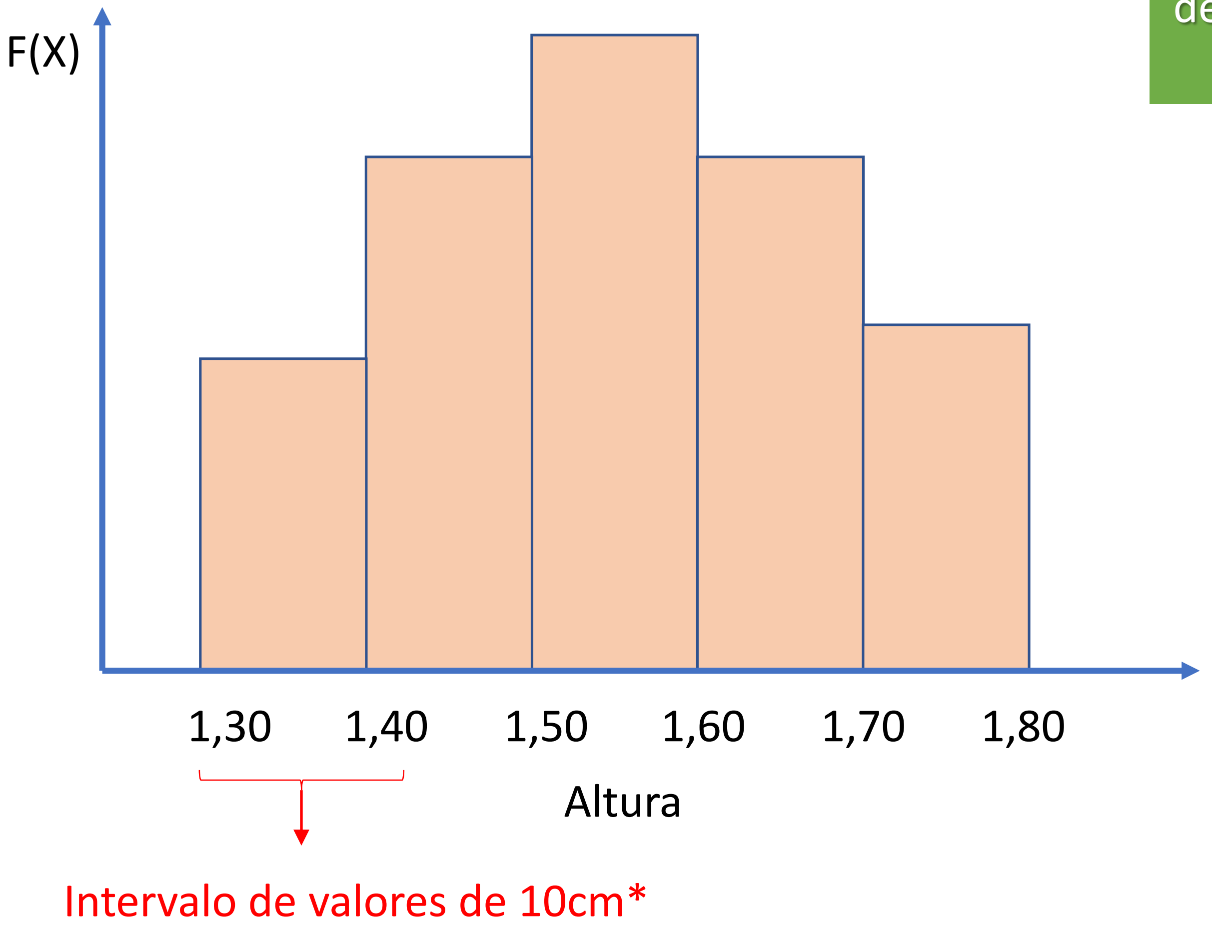


Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?

Medidas de tendência central

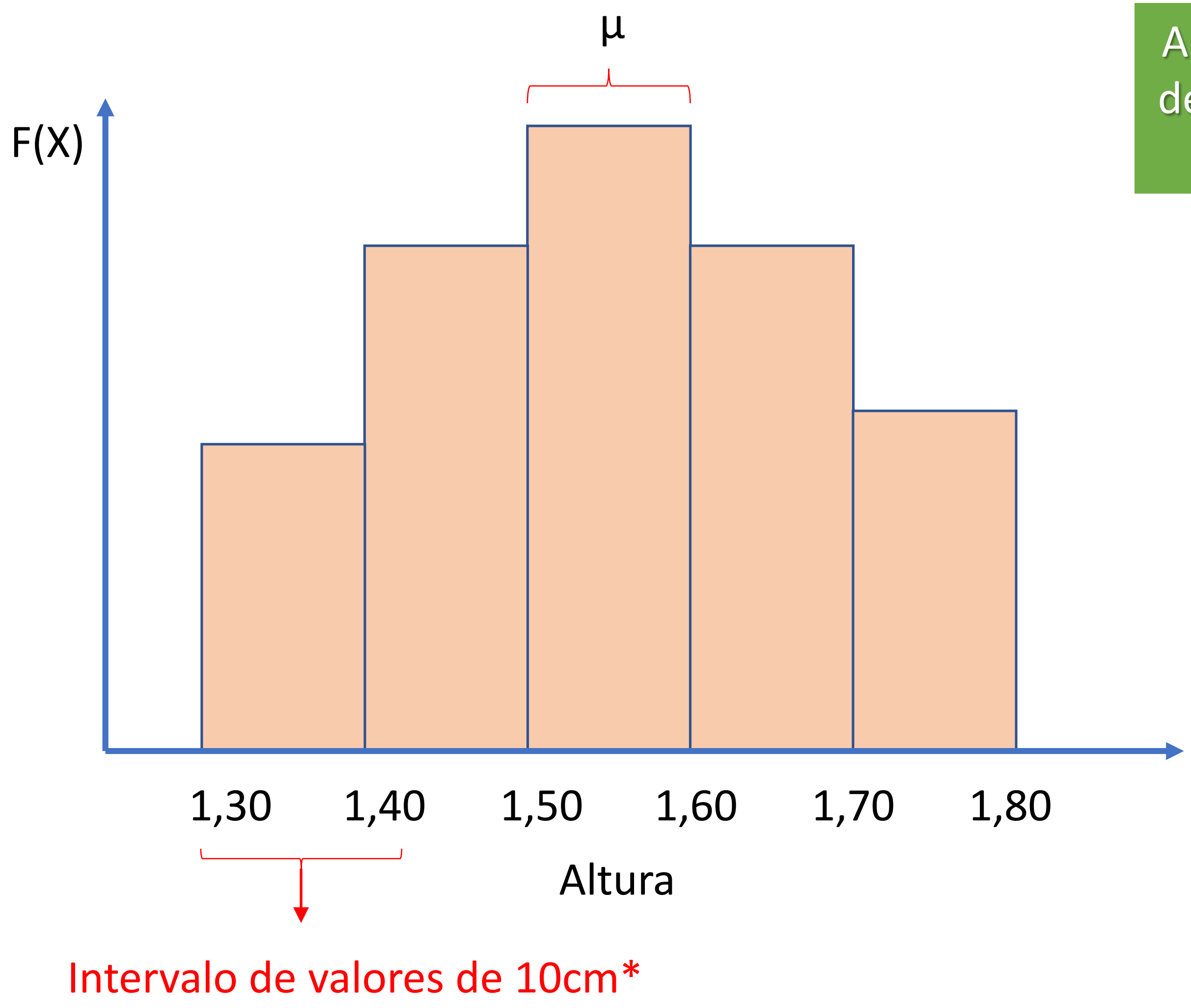


Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?

Medidas de tendência central

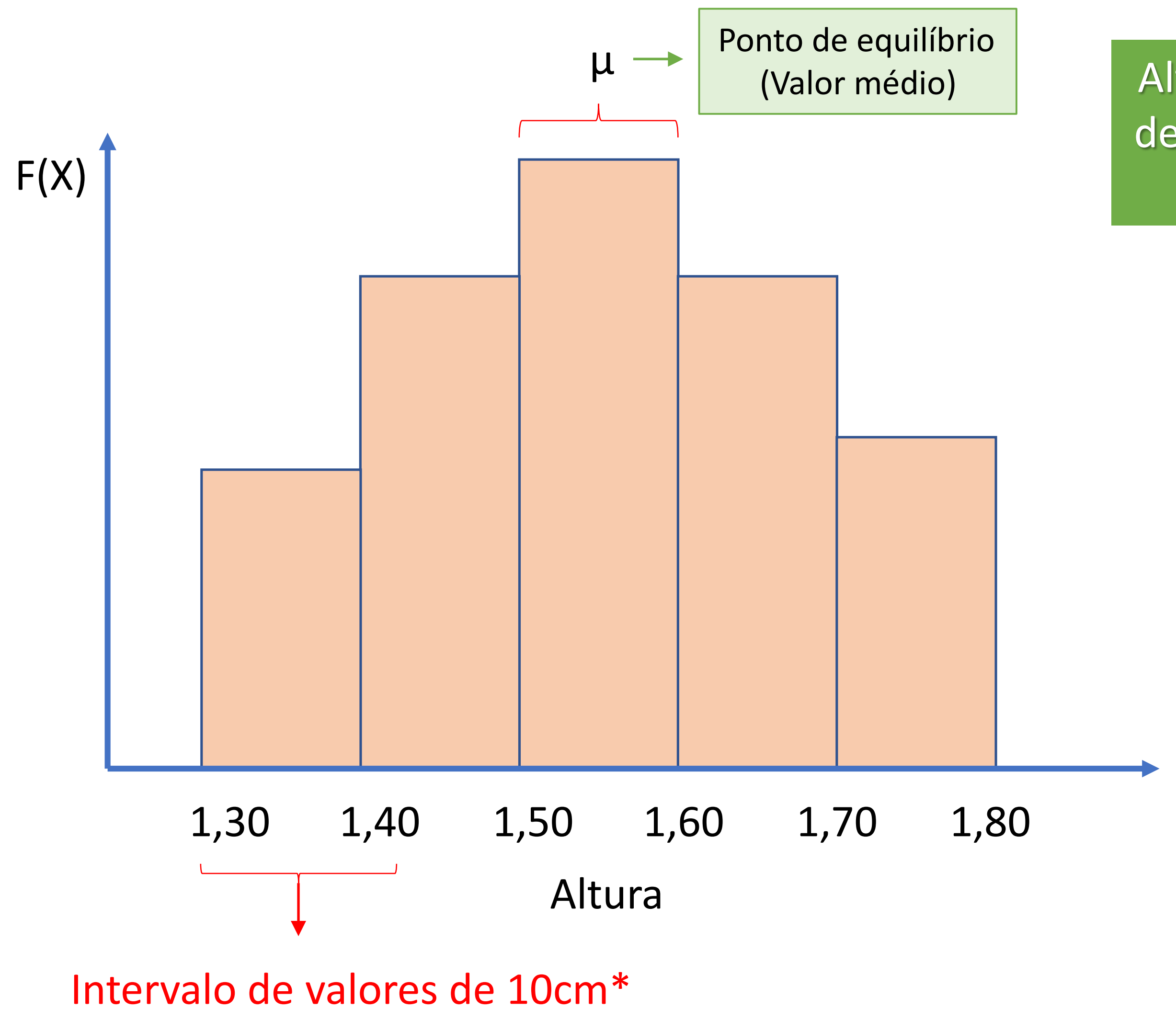


Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?

Medidas de tendência central

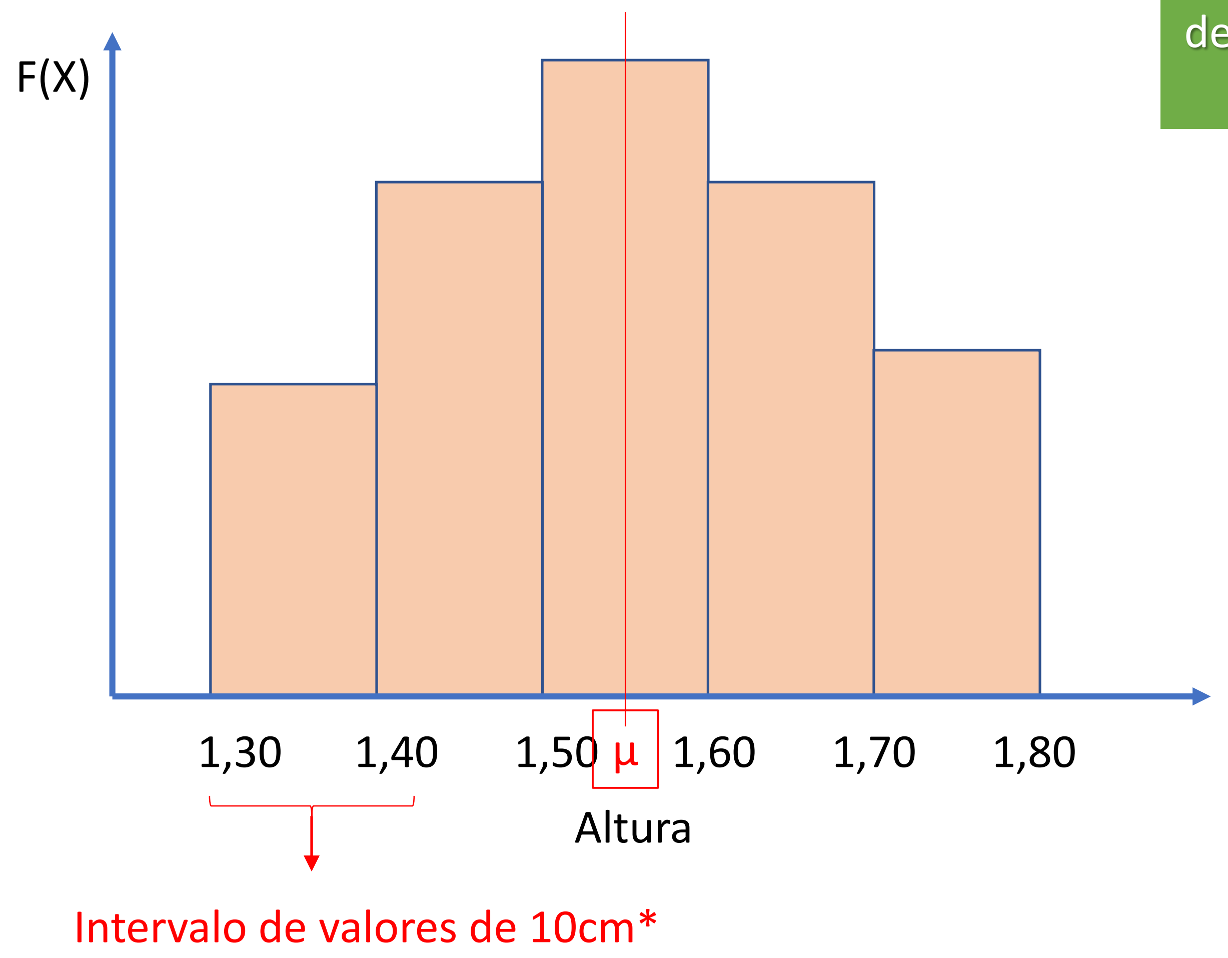


Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?

Medidas de tendência central

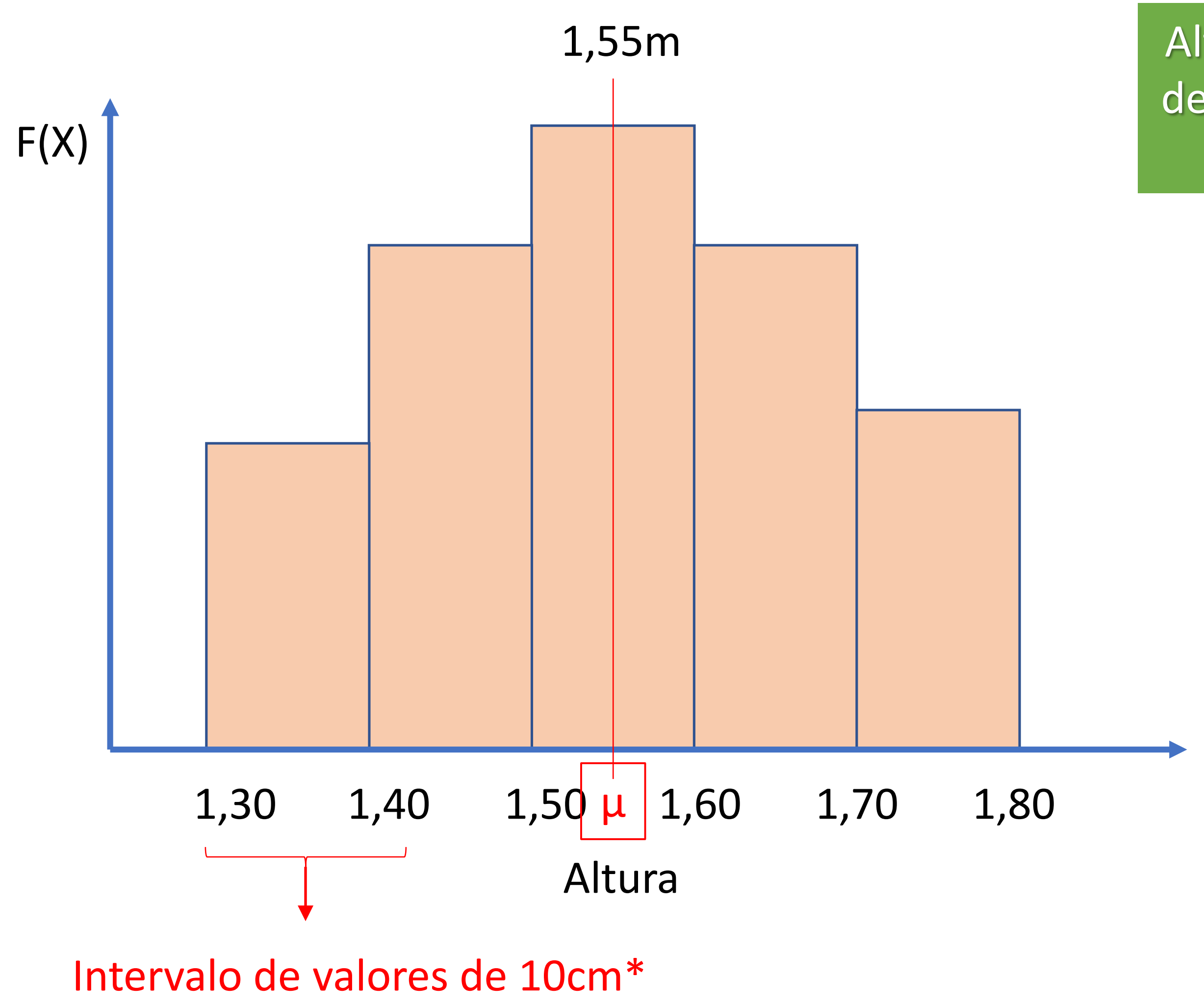


Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?

Medidas de tendência central

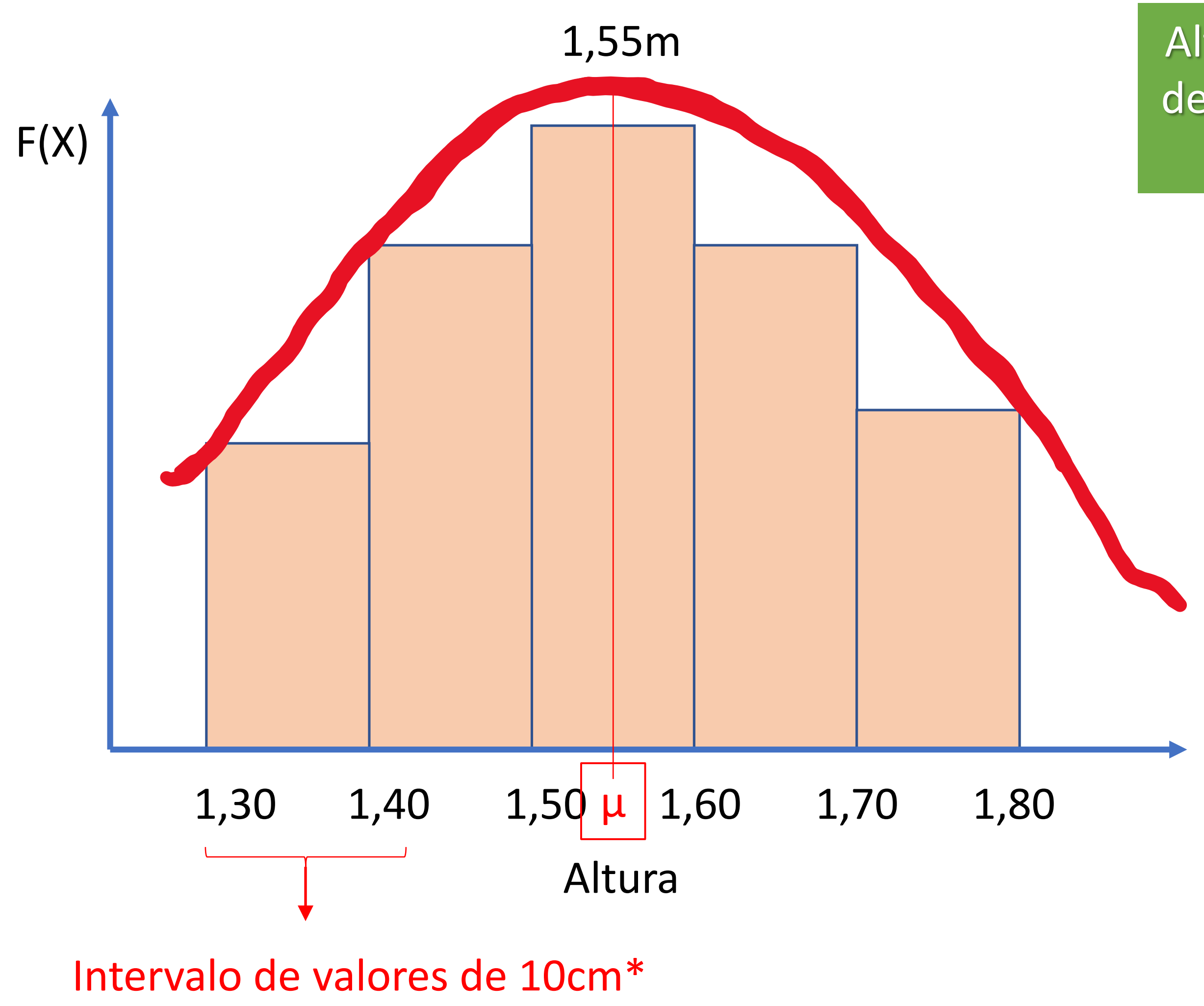


Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?

Medidas de tendência central

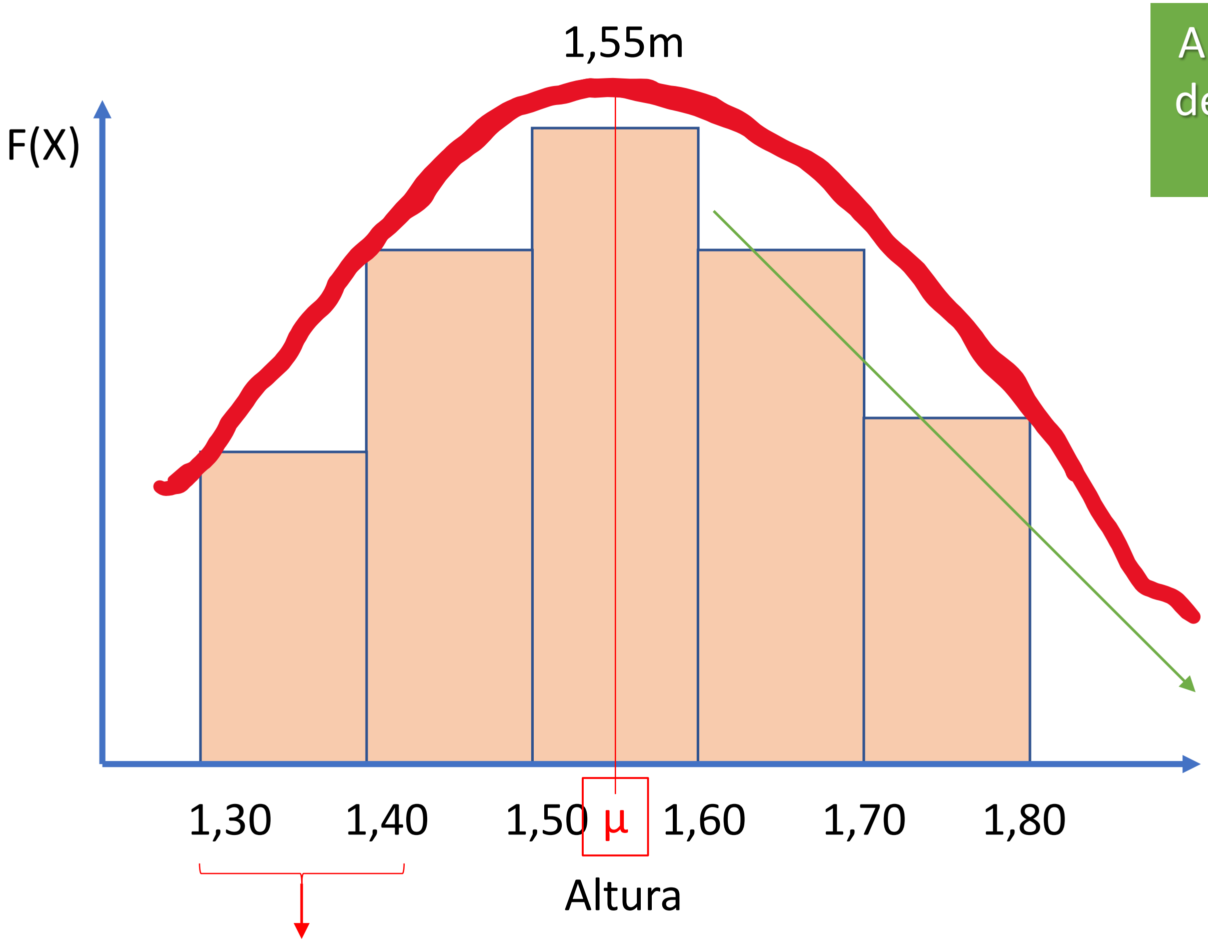


Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?

Medidas de tendência central

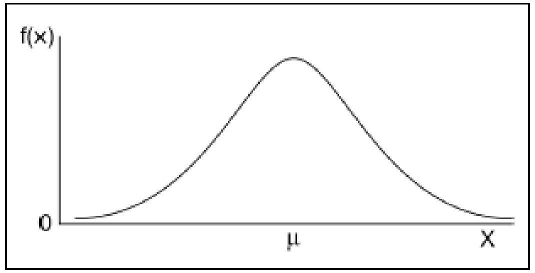


Intervalo de valores de 10cm*

Curva normal

Altura da população de cientistas de dados pertencentes ao gênero feminino

Qual intervalo de valores eu espero que seja o mais frequente em minha população?



Medidas de tendência central

Média aritmética

$$\bar{x} = \frac{\sum x}{n}$$

$\sum x$: Soma de todos os valores de x.

n: número de unidades amostrais, ou número de amostras x

Medidas de tendência central

Média aritmética

$$\bar{x} = \frac{\sum x}{n}$$

$\sum x$: Soma de todos os valores de x.
n: número de unidades amostrais, ou número de amostras x

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a média de altura dessas mulheres?

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **média** de altura dessas mulheres?

$$\bar{x} = \frac{\sum x}{n} = \frac{1.46+1.67+1.92+1.79+1.43+1.60+2.16+1.66+2.14+1.84+1.37+1.7+1.82+1.58+1.74}{15}$$

$\bar{x}$= 1.73m

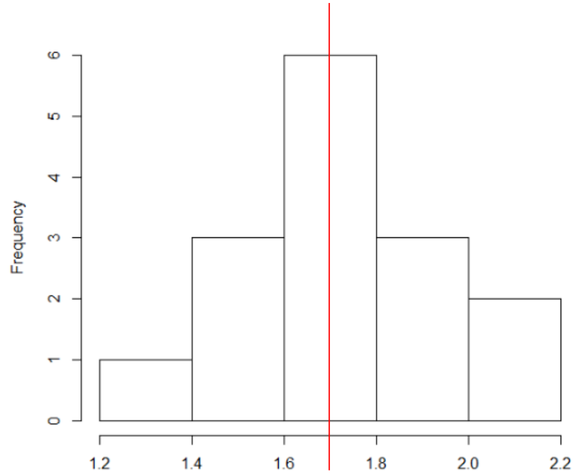
Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **média** de altura dessas mulheres?

$$\bar{x} = \frac{\sum x}{n} = \frac{1.46+1.67+1.92+1.79+1.43+1.60+2.16+1.66+2.14+1.84+1.37+1.7+1.82+1.58+1.74}{15}$$

$\bar{x}$= 1.73m

$\bar{x}$ =~1.7



Medidas de tendência central

Média aritmética

- Medida de tendência central mais utilizada por ser fácil de calcular e possuir uma interpretação familiar.
- Propriedades que a torna útil no momento de comparar populações ou amostras.
- Ela representa o valor "provável" de uma variável, por isso, é muitas vezes chamada de valor esperado, ou esperança, quando calculada para a população: E(X).
- Usamos o símbolo 'μ' (letra m do alfabeto grego; leia-se mü) quando nos referimos a média populacional e $\bar{x}$ para a média amostral.

Medidas de tendência central

Existem outros tipos de Médias também!

- Média ponderada.
- Média geométrica.
- Média para dados agrupados por intervalos de classe.
- Etc.

Vou enviar material sobre essas médias para que vocês estudem em casa (:

Medidas de tendência central

Mediana

- A mediana (md) é o valor de x, em uma série ordenada de dados, que divide a série em dois subgrupos de igual tamanho.
- Em outras palavras, é um valor tal que tenha igual quantidade de valores menores e maiores que ele.
- !! Um valor importante é que ela não é afetada pelos valores extremos da série como a média !! (explico melhor depois).
- $Md = \frac{n+1}{2}$

Seguindo com o nosso exemplo de antes!

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

O primeiro passo para calcular a mediana é organizar o nosso rol de dados em ordem crescente:

<u>Rol original:</u>

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

Qual é o menor valor?

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

O primeiro passo para calcular a mediana é organizar o nosso rol de dados em ordem crescente:

Rol original:

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, **1.37**, 1.70, 1.82, 1.58, 1.74}

Qual é o menor valor?

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

O primeiro passo para calcular a mediana é organizar o nosso rol de dados em ordem crescente:

Rol original:

1

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, **1.37**, 1.70, 1.82, 1.58, 1.74}

Qual é o menor valor?

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

O primeiro passo para calcular a mediana é organizar o nosso rol de dados em ordem crescente:

Rol original:

1

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, **1.37**, 1.70, 1.82, 1.58, 1.74}

Qual é o segundo menor valor?

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

O primeiro passo para calcular a mediana é organizar o nosso rol de dados em ordem crescente:

<u>Rol original:</u>

2 1

x={1.46, 1.67, 1.92, 1.79, **1.43**, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

Qual é o segundo menor valor?

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

O primeiro passo para calcular a mediana é organizar o nosso rol de dados em ordem crescente:

Rol original:

x={1.46, 1.67, 1.92, 1.79, $\overset{2}{1.43}$, 1.60, 2.16, 1.66, 2.14, 1.84, $\overset{1}{1.37}$, 1.70, 1.82, 1.58, 1.74}

Qual é o terceiro menor valor?

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

O primeiro passo para calcular a mediana é organizar o nosso rol de dados em ordem crescente:

Rol original:

2 1

x={**1.46**, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

3

Qual é o terceiro menor valor?

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

O primeiro passo para calcular a mediana é organizar o nosso rol de dados em ordem crescente:

Rol original:

2 1

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

3

Qual é o terceiro menor valor?

Etc Etc.

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

O primeiro passo para calcular a mediana é organizar o nosso rol de dados em ordem crescente:

Rol original:

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14, 2.16}

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

Rol original:

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

11 5 1 14 6 8 2 12 15 4 13 10 3 9 7

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14, 2.16}

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

Rol original:

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

11	5	1	14	6	8	2	12	15	4	13	10	3	9	7

Qual é o valor no rol que divide a minha série no meio?
Ou seja, qual é a minha mediana?

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14, 2.16}

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

Rol original:

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

11	5	1	14	6	8	2	12	15	4	13	10	3	9	7

$$Md = \frac{n+1}{2}$$

Qual é o valor no rol que divide a minha série no meio? Ou seja, qual é a minha mediana?

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14, 2.16}

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

<u>Rol original:</u>

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

11 5 1 14 6 8 2 12 15 4 13 10 3 9 7

$$Md = \frac{n+1}{2} = \frac{15+1}{2} = 8$$

Qual é o valor no rol que divide a minha série no meio? Ou seja, qual é a minha mediana?

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14, 2.16}

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

Rol original:

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

11 5 1 14 6 8 2 12 15 4 13 10 3 9 7

$$Md = \frac{n+1}{2} = \frac{15+1}{2} = 8$$

Qual é o valor no rol que divide a minha série no meio? Ou seja, qual é a minha mediana?

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, $\boxed{1.70,}$ 1.74, 1.79, 1.82, 1.84, 1.92, 2.14, 2.16}

1 2 3 4 5 6 7 $\boxed{8}$ 9 10 11 12 13 14 15

Md

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

Rol original:

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

11 5 1 14 6 8 2 12 15 4 13 10 3 9 7

$$Md = \frac{n+1}{2} = \frac{15+1}{2} = 8$$

Qual é o valor no rol que divide a minha série no meio? Ou seja, qual é a minha mediana?

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, $\boxed{1.70,}$ 1.74, 1.79, 1.82, 1.84, 1.92, 2.14, 2.16}

1 2 3 4 5 6 7 $\boxed{8}$ 9 10 11 12 13 14 15

Md

Medidas de tendência central

Exercício: Amostramos a altura de 15 mulheres do GECD, sendo o rol de valores o seguinte: x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.7, 1.82, 1.58, 1.74}. Qual é a **mediana** de altura dessas mulheres?

Rol original:

x={1.46, 1.67, 1.92, 1.79, 1.43, 1.60, 2.16, 1.66, 2.14, 1.84, 1.37, 1.70, 1.82, 1.58, 1.74}

11 5 1 14 6 8 2 12 15 4 13 10 3 9 7

$$Md = \frac{n+1}{2} = \frac{15+1}{2} = 8$$

Qual é o valor no rol que divide a minha série no meio? Ou seja, qual é a minha mediana?

Rol ordenado:

7 valores para trás | 7 valores para frente

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, $\boxed{1.70,}$ 1.74, 1.79, 1.82, 1.84, 1.92, 2.14, 2.16}

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Md

Medidas de tendência central

Mas e se o meu ROL de dados fosse um número par, e não impar?

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1	2	3	4	5	6	7	8	9	10	11	12	13	14

Medidas de tendência central

Mas e se o meu ROL de dados fosse um número par, e não impar?

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1	2	3	4	5	6	7	8	9	10	11	12	13	$\boxed{14}$

n=14

Medidas de tendência central

Mas e se o meu ROL de dados fosse um número par, e não impar?

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1	2	3	4	5	6	7	8	9	10	11	12	13	14

n=14

$$Md = \frac{n+1}{2} =$$

Medidas de tendência central

Mas e se o meu ROL de dados fosse um número par, e não impar?

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1	2	3	4	5	6	7	8	9	10	11	12	13	14

n=14

14

$$Md = \frac{n + 1}{2} =$$

Medidas de tendência central

Mas e se o meu ROL de dados fosse um número par, e não impar?

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1	2	3	4	5	6	7	8	9	10	11	12	13	14

n=14

$$Md = \frac{n+1}{2} = \frac{14+1}{2} = 7,5$$

Medidas de tendência central

Mas e se o meu ROL de dados fosse um número par, e não impar?

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 2 3 4 5 6 7 8 9 10 11 12 13 14

7,5

n=14

$$Md = \frac{n+1}{2} = \frac{14+1}{2} = 7{,}5$$

Medidas de tendência central

Mas e se o meu ROL de dados fosse um número par, e não impar?

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1	2	3	4	5	6	7	8	9	10	11	12	13	14

7,5

n=14

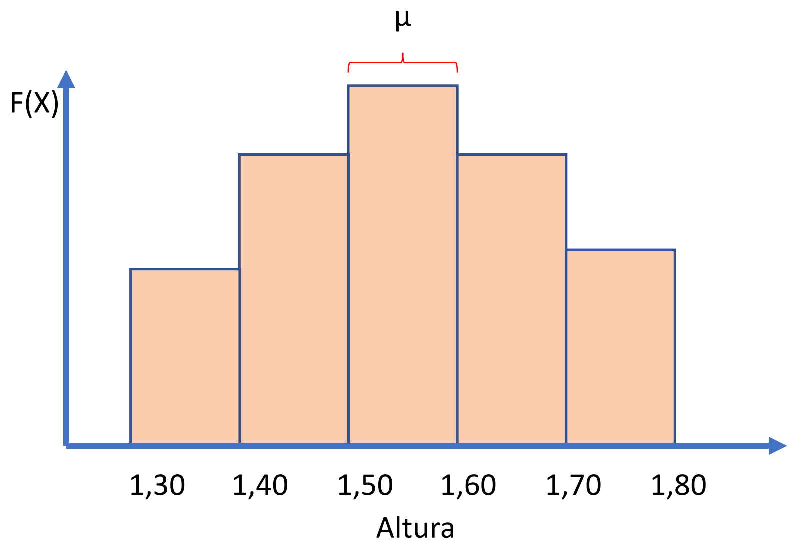
$$Md = \frac{n+1}{2} = \frac{14+1}{2} = 7{,}5$$

Média dos valores em 7 e 8!
Ou seja, (1.67+1.70)/2= **1.69m**

Medidas de tendência central

Mediana

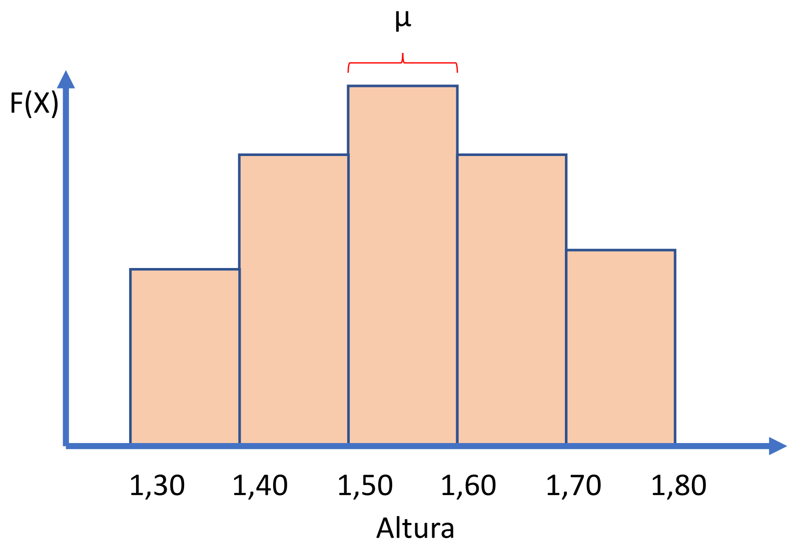
A mediana é uma medida de tendência central útil para descrever dados que se distribuem de maneira assimétrica; ou seja, quando os dados não se distribuem normalmente.

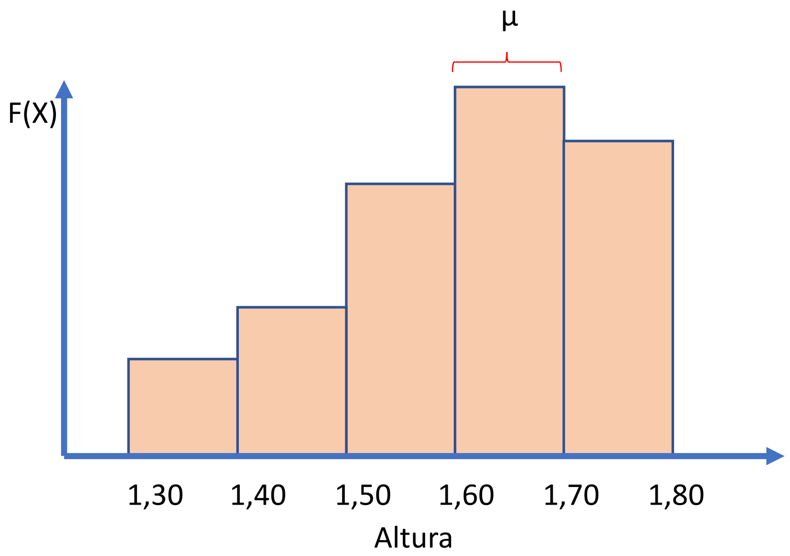


Medidas de tendência central

Mediana

A mediana é uma medida de tendência central útil para descrever dados que se distribuem de maneira assimétrica; ou seja, quando os dados não se distribuem normalmente.

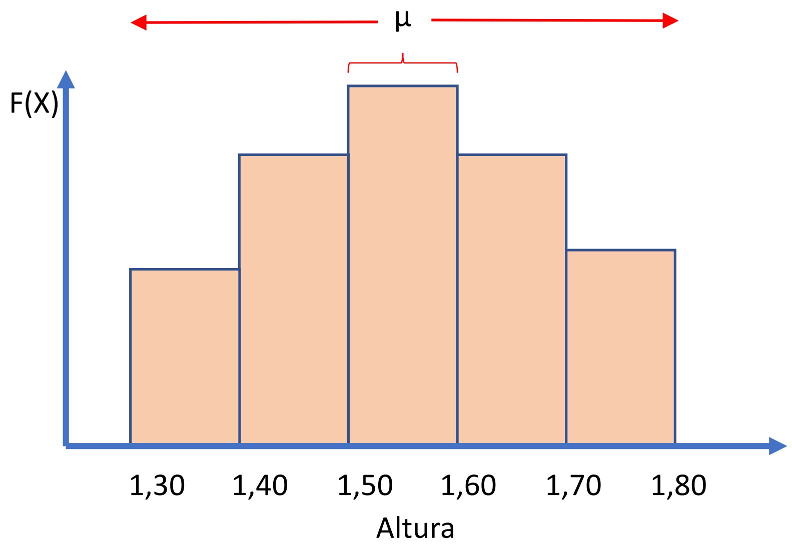


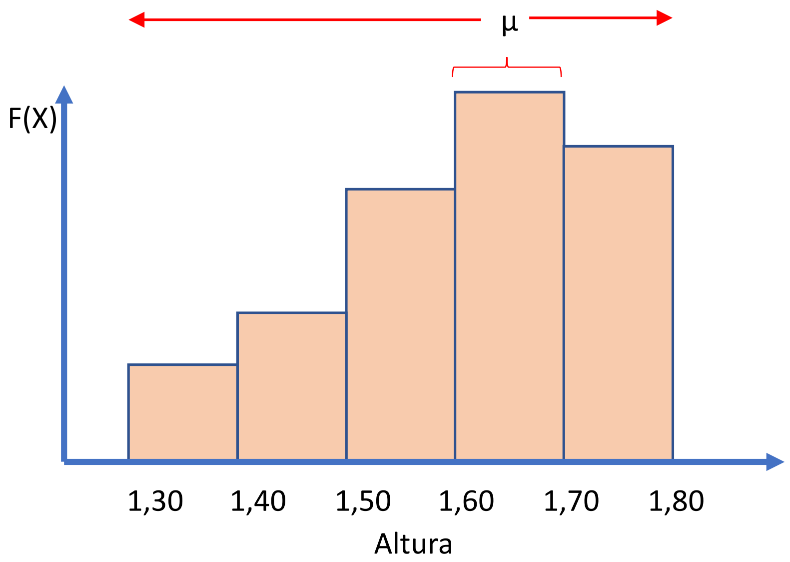


Medidas de tendência central

Mediana

A mediana é uma medida de tendência central útil para descrever dados que se distribuem de maneira <u>assimétrica</u>; ou seja, quando os dados não se distribuem normalmente.

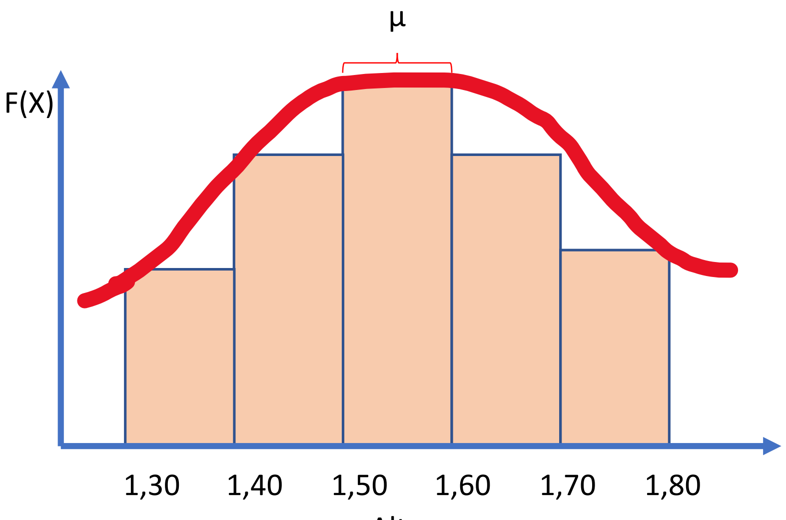


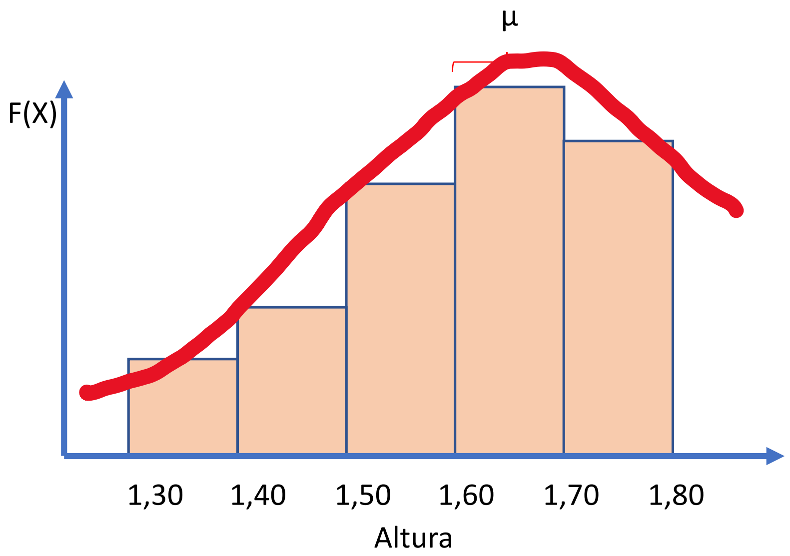


Medidas de tendência central

Mediana

A mediana é uma medida de tendência central útil para descrever dados que se distribuem de maneira assimétrica; ou seja, quando os dados não se distribuem normalmente.

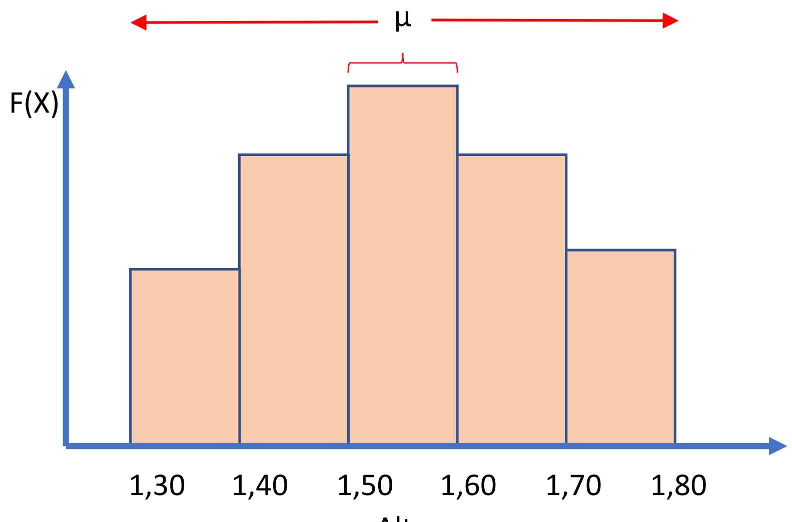


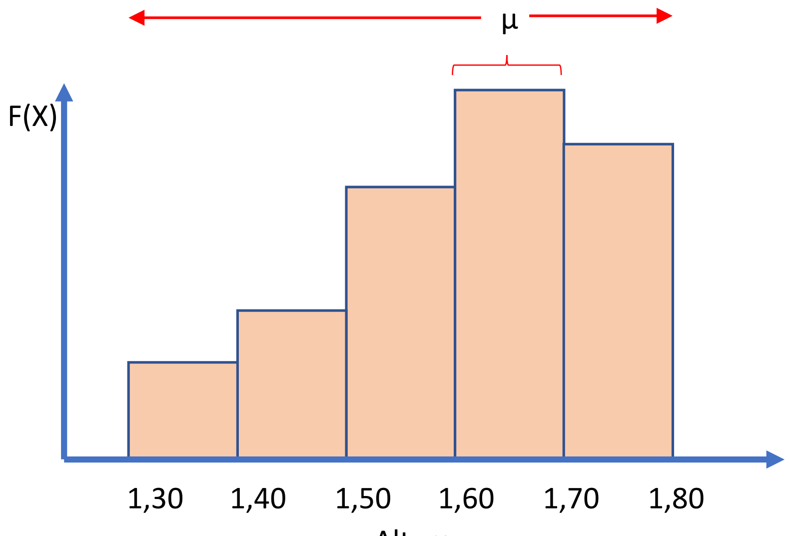


Medidas de tendência central

Mediana

A mediana é uma medida de tendência central útil para descrever dados que se distribuem de maneira assimétrica; ou seja, quando os dados não se distribuem normalmente.

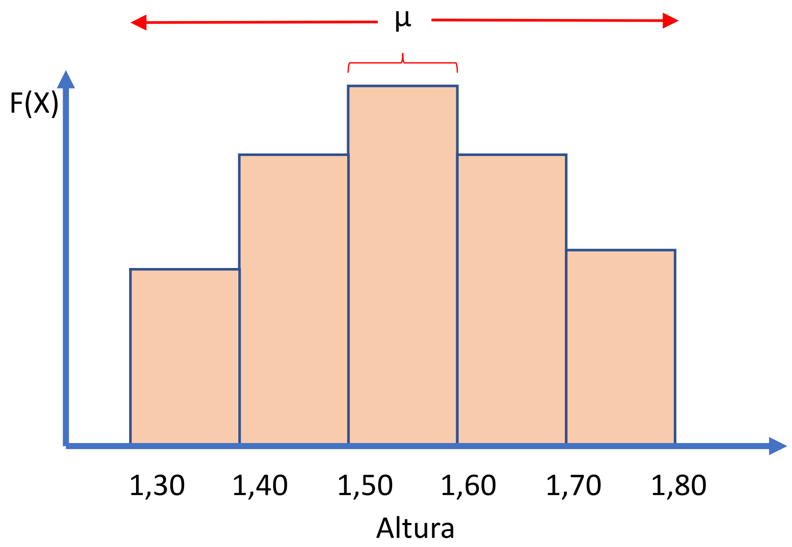


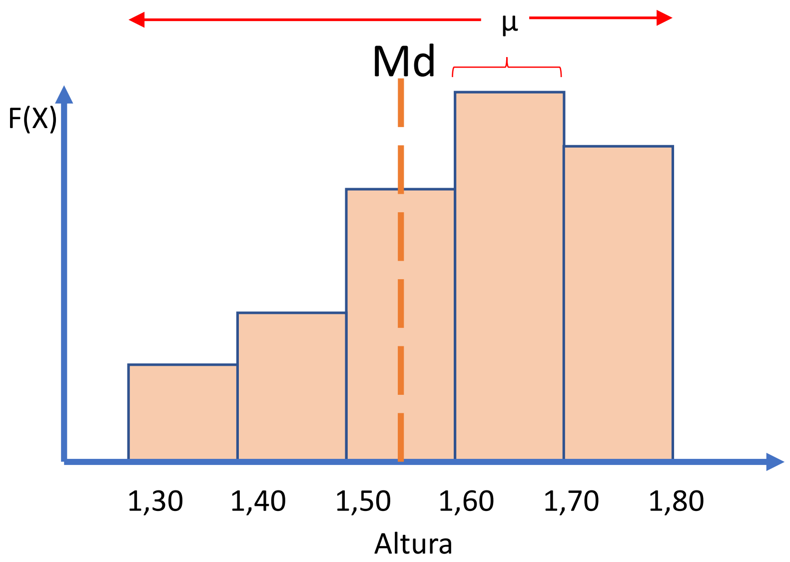


Medidas de tendência central

Mediana

A mediana é uma medida de tendência central útil para descrever dados que se distribuem de maneira assimétrica; ou seja, quando os dados não se distribuem normalmente.



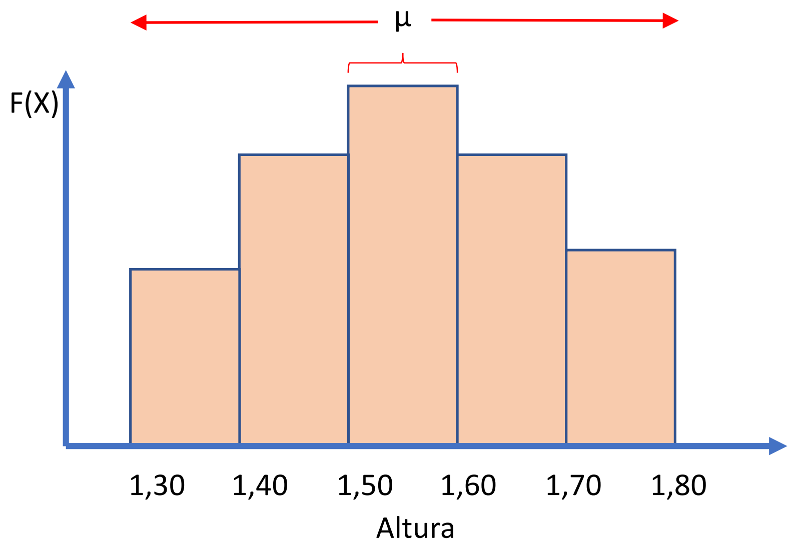


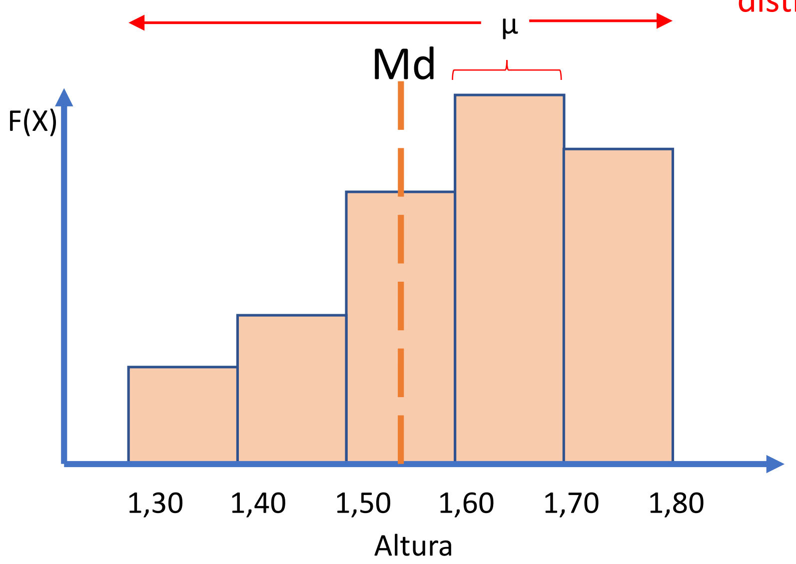
Medidas de tendência central

Mediana

A mediana é uma medida de tendência central útil para descrever dados que se distribuem de maneira <u>assimétrica</u>; ou seja, quando os dados não se distribuem normalmente.

Usamos a mediana pq ela não se altera com os valores dos dados, sempre estará no meio da distribuição.



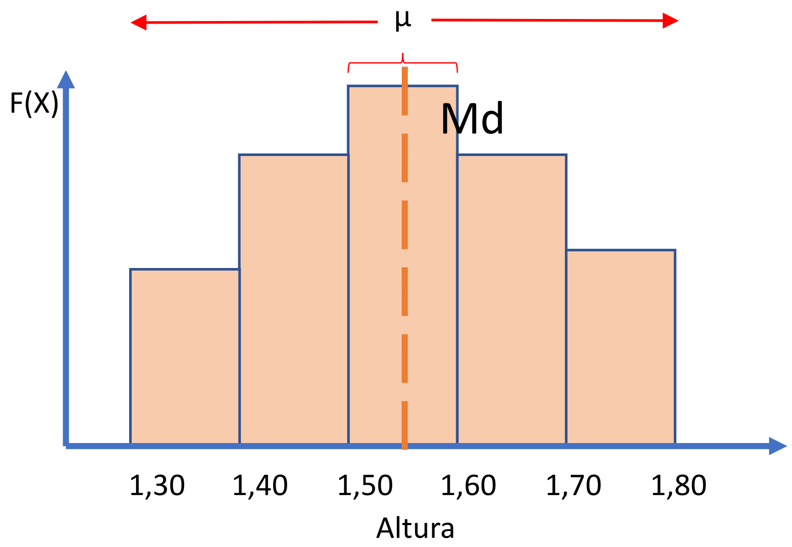


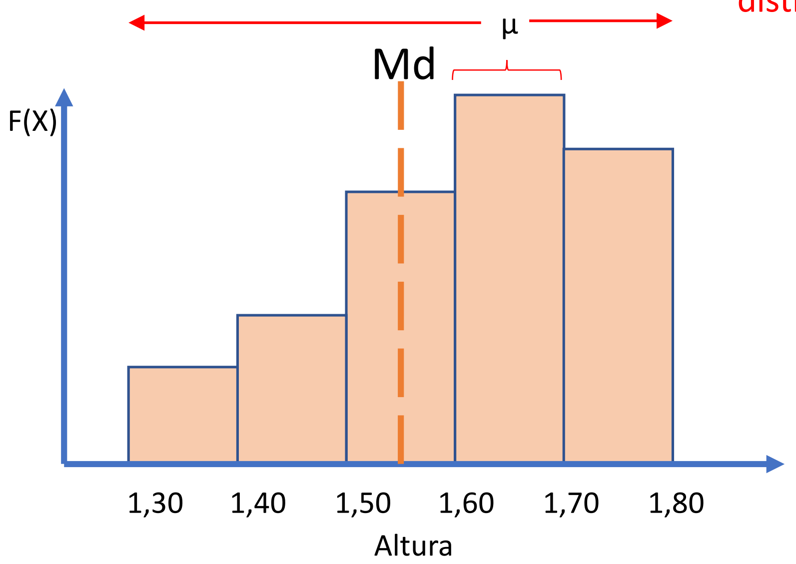
Medidas de tendência central

Mediana

A mediana é uma medida de tendência central útil para descrever dados que se distribuem de maneira <u>assimétrica</u>; ou seja, quando os dados não se distribuem normalmente.

Usamos a mediana pq ela não se altera com os valores dos dados, sempre estará no meio da distribuição.

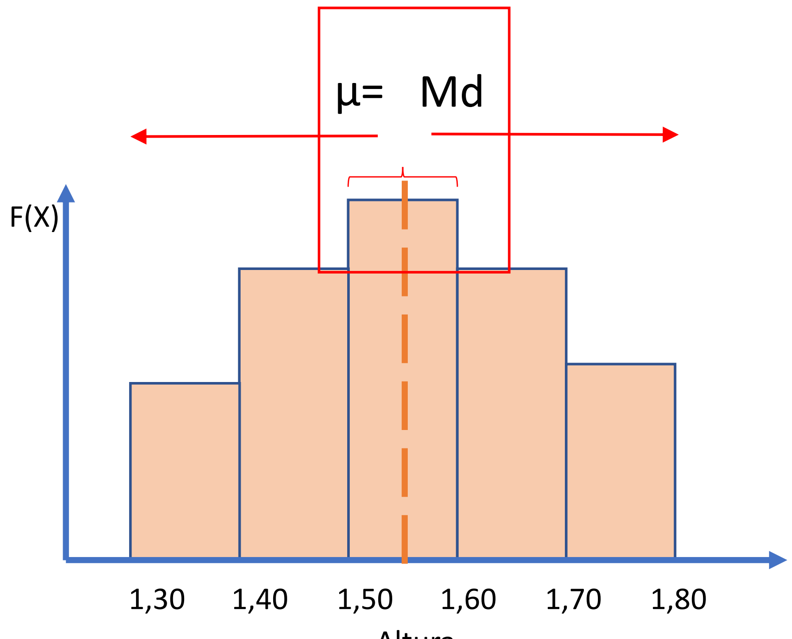




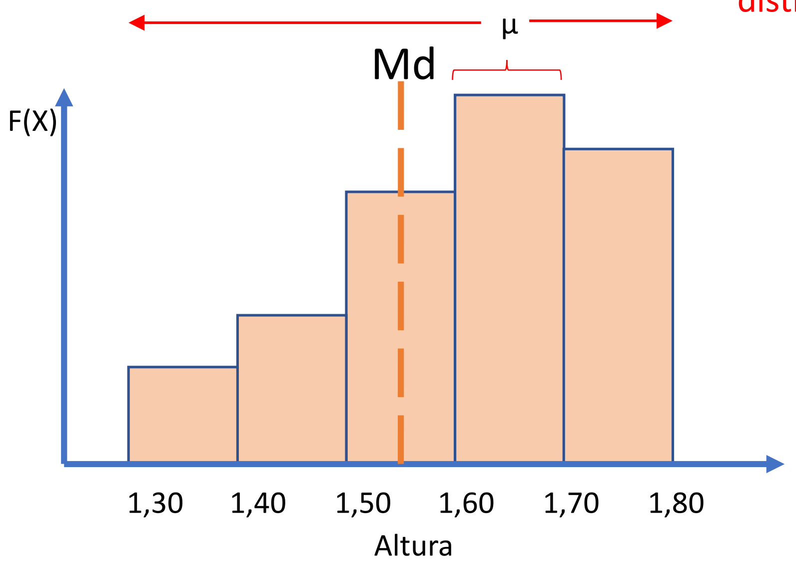
Medidas de tendência central

Mediana

A mediana é uma medida de tendência central útil para descrever dados que se distribuem de maneira assimétrica; ou seja, quando os dados não se distribuem normalmente.



Usamos a mediana pq ela não se altera com os valores dos dados, sempre estará no meio da distribuição.



Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1 1

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1 1 1

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1 1 1 1

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1 1 1 1 ....

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1 1 1 1 1 1 1 1 1 1 1

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1 1 1 1 1 1 1 1 1 1 1

- Não temos um valor modal, todos tem a mesma frequência.

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.43, 1.43, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1 1 1 1 1 1 1 1 1 1 1

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.43, 1.43, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1 1 1 1 1 1 1 1 1 1 1

1 + 1 + 1 = 3

Moda = 1.43

Medidas de tendência central

Moda

- A moda (Mo) é o valor mais frequente de uma série de valores.

Rol ordenado:

x={1.37, 1.43, 1.43, 1.43, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1 1 1 1 1 1 1 1 1 1 1 1 1 1

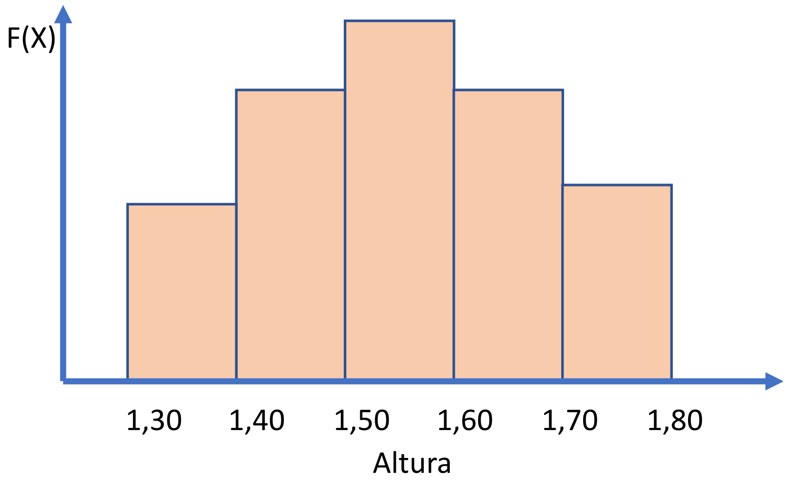
1 + 1 + 1 = 3
Moda = 1.43

Agora temos um valor modal igual a 1.43m

Medidas de tendência central

Moda

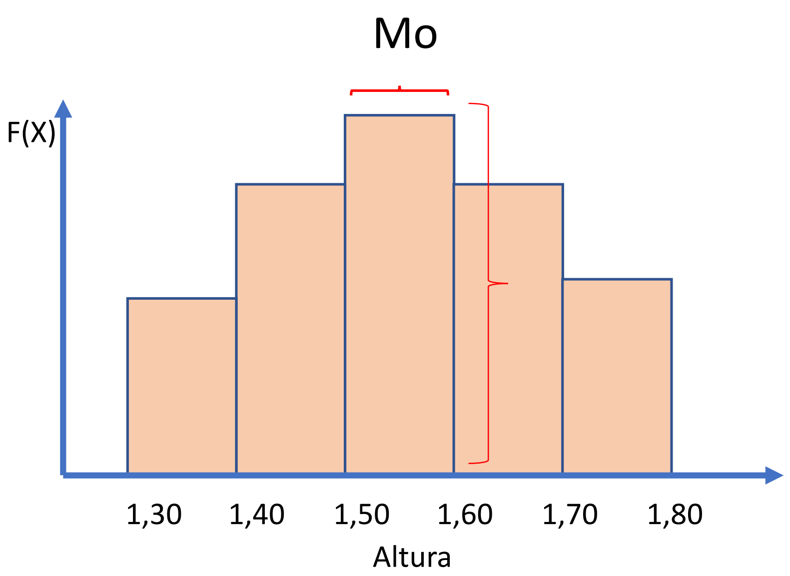
A moda (Mo) pode ser usada quando não temos distribuição normal dos dados e/ou quando queremos enfatizar os valores que foram mais frequentes.



Medidas de tendência central

Moda

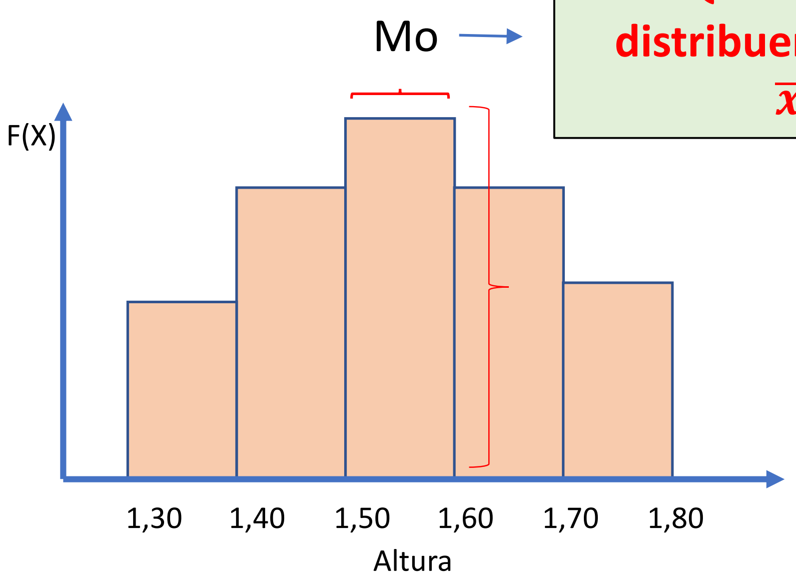
A moda (Mo) pode ser usada quando não temos distribuição normal dos dados e/ou quando queremos enfatizar os valores que foram mais frequentes.



Medidas de tendência central

Moda

A moda (Mo) pode ser usada quando não temos distribuição normal dos dados e/ou quando queremos enfatizar os valores que foram mais frequentes.

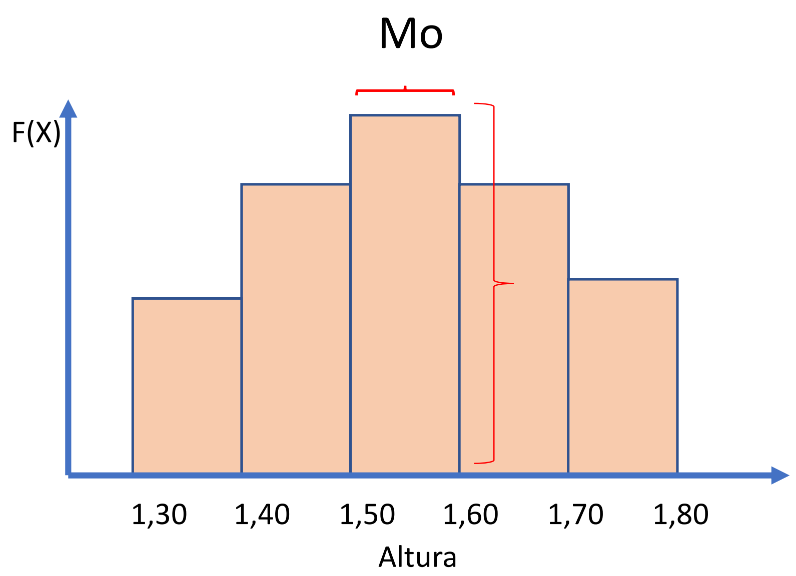


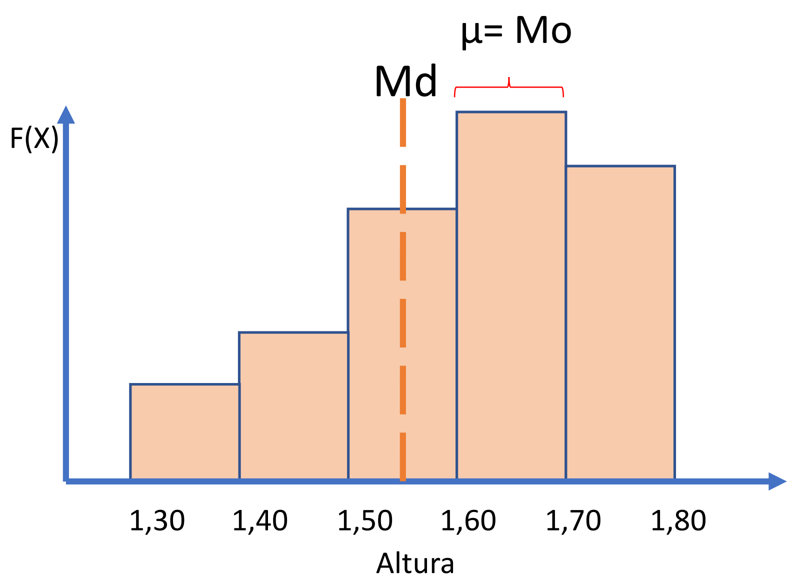
Quando os dados se distribuem normalmente a $\bar{x}$=Md=Mo

Medidas de tendência central

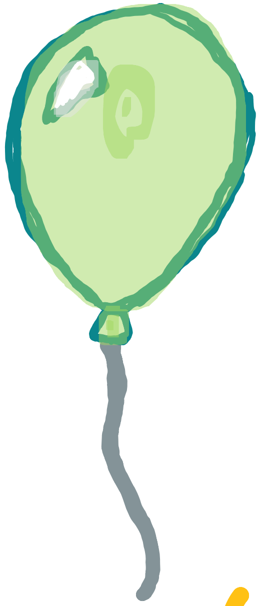
Moda

A moda (Mo) pode ser usada quando não temos distribuição normal dos dados e/ou quando queremos enfatizar os valores que foram mais frequentes.





Um intervalo
para um café,
por favor!

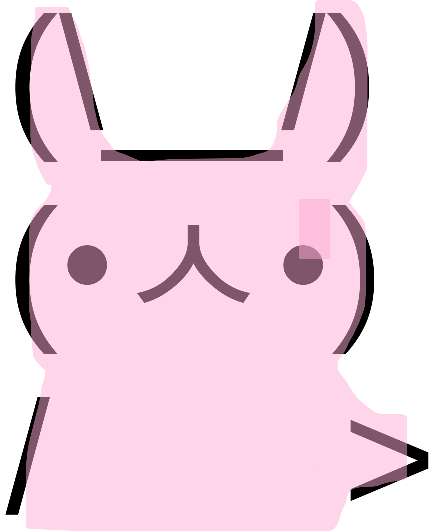


Voltamos em breve!

Medidas de dispersão ou de variabilidade

As medidas de tendência central são insuficientes para representar adequadamente conjuntos de dados, pois nada revelam sobre sua variabilidade.

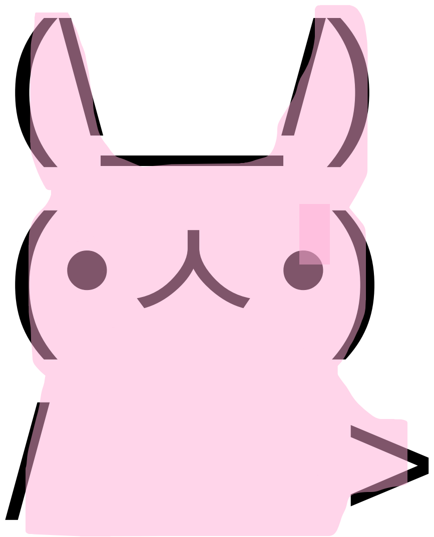
Ok! A média de altura da minha população é 1.70m, mas qual é a variabilidade de alturas da minha população? Existem pessoas muito pequenas ou muito grandes na minha população?



Medidas de dispersão ou de variabilidade

As medidas de tendência central são insuficientes para representar adequadamente conjuntos de dados, pois nada revelam sobre sua variabilidade.

Para responder essas perguntas vamos ter que usar outros tipos de medidas! A amplitude de variação, a variância e o desvio-padrão, o coeficiente de variação e por fim os quartis. **Vamos lá**!



Medidas de dispersão ou de variabilidade

Amplitude de variação (a)

- A medida mais simples de dispersão
- O calculo é muito simples: o último valor **-** o primeiro valor.

Medidas de dispersão ou de variabilidade

Amplitude de variação (a)

- A medida mais simples de dispersão
- O calculo é muito simples: o último valor **-** o primeiro valor.

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Medidas de dispersão ou de variabilidade

Amplitude de variação (a)

- A medida mais simples de dispersão
- O calculo é muito simples: o último valor **-** o primeiro valor.

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Medidas de dispersão ou de variabilidade

Amplitude de variação (a)

- A medida mais simples de dispersão
- O calculo é muito simples: o último valor **-** o primeiro valor.

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Medidas de dispersão ou de variabilidade

Amplitude de variação (a)

- A medida mais simples de dispersão
- O calculo é muito simples: o último valor **-** o primeiro valor.

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

a= 2.14 – 1.37 = 0.77cm

Medidas de dispersão ou de variabilidade

Amplitude de variação (a)

- A medida mais simples de dispersão
- O calculo é muito simples: o último valor **-** o primeiro valor.

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

a= 2.14 – 1.37 = 0.77cm

Entre o valor inicial e final eu tenho uma variação de 0.77 cm.

Medidas de dispersão ou de variabilidade

Variância

- A variância e o desvio padrão visam representar quanto os valores de ‘x’ se desviam da média.
- A variância é a base do cálculo do desvio-padrão, por isso vamos vê-la primeiro.
- O símbolo da variância populacional é “” e da amostral é “S^2”.
- A variância é a **soma dos quadrados** dividida pelo número de observações do conjunto. A variância é representada por S^2, sendo calculada pela fórmula:

$$\frac{\sum(xi\ - média)^2}{n-1}$$

Medidas de dispersão ou de variabilidade

Variância

- A variância e o desvio padrão visam representar quanto os valores de ‘x’ se desviam da média.
- A variância é a base do cálculo do desvio-padrão, por isso vamos vê-la primeiro.
- O símbolo da variância populacional é “” e da amostral é “S^2”.
- A variância é a **soma dos quadrados** dividida pelo número de observações do conjunto. A variância é representada por S^2, sendo calculada pela fórmula:

$$\frac{\sum(xi - média)^2}{n-1}$$

xi= O valor de x para cada unidade experimental i

Medidas de dispersão ou de variabilidade

Variância

- A variância e o desvio padrão visam representar quanto os valores de ‘x’ se desviam da média.
- A variância é a base do cálculo do desvio-padrão, por isso vamos vê-la primeiro.
- O símbolo da variância populacional é “” e da amostral é “S^2”.
- A variância é a **soma dos quadrados** dividida pelo número de observações do conjunto. A variância é representada por S^2, sendo calculada pela fórmula:

$$\frac{\sum(xi - média)^2}{n-1}$$!!

xi= O valor de x para cada unidade experimental i

Medidas de dispersão ou de variabilidade

Variância

- A variância e o desvio padrão visam representar quanto os valores de ‘x’ se desviam da média.
- A variância é a base do cálculo do desvio-padrão, por isso vamos vê-la primeiro.
- O símbolo da variância populacional é “” e da amostral é “S^2”.
- A variância é a **soma dos quadrados** dividida pelo número de observações do conjunto. A variância é representada por S^2, sendo calculada pela fórmula:

$$S^2 = \frac{\sum(xi - média)^2}{n-1} \rightarrow !!$$

MQ!!

xi= O valor de x para cada unidade experimental i

Medidas de dispersão ou de variabilidade

Variância

- A variância e o desvio padrão visam representar quanto os valores de ‘x’ se desviam da média.
- A variância é a base do cálculo do desvio-padrão, por isso vamos vê-la primeiro.
- O símbolo da variância populacional é “” e da amostral é “S^2”.
- A variância é a **soma dos quadrados** dividida pelo número de observações do conjunto. A variância é representada por S^2, sendo calculada pela fórmula:

xi= O valor de x para cada unidade experimental i

$$S^2 = \frac{\sum(xi - média)^2}{n-1}$$

!! MQ!!

GL

Medidas de dispersão ou de variabilidade

Variância

- A variância e o desvio padrão visam representar quanto os valores de ‘x’ se desviam da média.
- A variância é a base do cálculo do desvio-padrão, por isso vamos vê-la primeiro.
- O símbolo da variância populacional é “” e da amostral é “S^2”.
- A variância é a **soma dos quadrados** dividida pelo número de observações do conjunto. A variância é representada por S^2, sendo calculada pela fórmula:

xi= O valor de x para cada unidade experimental i

$$S^2 = \frac{\sum(xi - média)^2}{n - \cancel{1}}$$

!!

MQ!!

GL

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi\ - média)^2}{n-1}$$

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - média)^2}{n-1}$$

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1º passo é calcular a média da amostra!

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - média)^2}{n-1}$$

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1º passo é calcular a média da amostra!

$\bar{x}$ =1.69

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - 1.69)^2}{n-1}$$

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

1º passo é calcular a média da amostra!

$\bar{x}$ =1.69

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - \mathbf{1.69})^2}{n-1}$$

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

2º já vamos incluir o valor de n-1 na Formula também (:

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - \mathbf{1.69})^2}{n-1}$$

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

2º já vamos incluir o valor de n-1 na Formula também (:

n – 1= 14 – 1= 13

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - \mathbf{1.69})^2}{\mathbf{13}}$$

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

2º já vamos incluir o valor de n-1 na Formula também (:

n – 1= 14 – 1= 13

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - \mathbf{1.69})^2}{\mathbf{13}}$$

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - \mathbf{1.69})^2}{\mathbf{13}}$$

Rol ordenado:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Medidas de altura (x)	xi – média	(xi – média)²

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - \mathbf{1.69})^2}{\mathbf{13}}$$

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Medidas de altura (x)	xi – média	(xi – média)²

Medidas de dispersão ou de variabilidade

$$S^2 = \frac{\sum(xi - \mathbf{1.69})^2}{\mathbf{13}}$$

<u>Rol ordenado</u>:

x={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Medidas de altura (x)	xi – média	(xi – média)²
1.37		
1.43		
1.46		
1.58		
1.60		
1.66		
1.67		
1.70		
1.74		
1.79		
1.82		

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	xi – média	(xi – média)²
1.37		
1.43		
1.46		
1.58		
1.60		
1.66		
1.67		
1.70		
1.74		
1.79		
1.82		
1.84		
1.92		
2.14		
∑=		

$$S^2 = \frac{\sum(xi - \mathbf{1.69})^2}{\mathbf{13}}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37		
1.43		
1.46		
1.58		
1.60		
1.66		
1.67		
1.70		
1.74		
1.79		
1.82		
1.84		
1.92		
2.14		
∑=		

$$S^2 = \frac{\sum(xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)=	
1.43		
1.46		
1.58		
1.60		
1.66		
1.67		
1.70		
1.74		
1.79		
1.82		
1.84		
1.92		
2.14		
∑=		

$$S^2 = \frac{\sum(xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)=	
1.43	(1.43-1.69)=	
1.46	(1.46-1.69)=	
1.58	(1.58-1.69)=	
1.60	(1.60-1.69)=	
1.66	(1.66-1.69)=	
1.67	(1.67-1.69)=	
1.70	(1.70-1.69)=	
1.74	(1.74-1.69)=	
1.79	(1.79-1.69)=	
1.82	(1.82-1.69)=	
1.84	(1.84-1.69)=	
1.92	(1.92-1.69)=	
2.14	(2.14-1.69)=	
∑=		

$$S^2 = \frac{\sum(xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	
1.43	(1.43-1.69)= -0.26	
1.46	(1.46-1.69)= -0.23	
1.58	(1.58-1.69)= -0.11	
1.60	(1.60-1.69)= -0.09	
1.66	(1.66-1.69)= -0.03	
1.67	(1.67-1.69)= -0.02	
1.70	(1.70-1.69)= 0.01	
1.74	(1.74-1.69)= 0.05	
1.79	(1.79-1.69)= 0.10	
1.82	(1.82-1.69)= 0.13	
1.84	(1.84-1.69)= 0.15	
1.92	(1.92-1.69)= 0.23	
2.14	(2.14-1.69)= 0.45	
∑=		

$$S^2 = \frac{\sum(xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	
1.43	(1.43-1.69)= -0.26	
1.46	(1.46-1.69)= -0.23	
1.58	(1.58-1.69)= -0.11	
1.60	(1.60-1.69)= -0.09	
1.66	(1.66-1.69)= -0.03	
1.67	(1.67-1.69)= -0.02	
1.70	(1.70-1.69)= 0.01	
1.74	(1.74-1.69)= 0.05	
1.79	(1.79-1.69)= 0.10	
1.82	(1.82-1.69)= 0.13	
1.84	(1.84-1.69)= 0.15	
1.92	(1.92-1.69)= 0.23	
2.14	(2.14-1.69)= 0.45	
∑=		

$$S^2 = \frac{\sum (xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	$(xi – média)^2$
1.37	(1.37-1.69)= - 0.32	$(-0.32)^2$=
1.43	(1.43-1.69)= -0.26	$(-0.26)^2$=
1.46	(1.46-1.69)= -0.23	$(-0.23)^2$=
1.58	(1.58-1.69)= -0.11	$(-0.11)^2$=
1.60	(1.60-1.69)= -0.09	$(-0.09)^2$=
1.66	(1.66-1.69)= -0.03	$(-0.03)^2$=
1.67	(1.67-1.69)= -0.02	$(-0.02)^2$=
1.70	(1.70-1.69)= 0.01	$(0.01)^2$=
1.74	(1.74-1.69)= 0.05	$(0.05)^2$=
1.79	(1.79-1.69)= 0.10	$(0.10)^2$=
1.82	(1.82-1.69)= 0.13	$(0.13)^2$=
1.84	(1.84-1.69)= 0.15	$(0.15)^2$=
1.92	(1.92-1.69)= 0.23	$(0.23)^2$=
2.14	(2.14-1.69)= 0.45	$(0.45)^2$=
∑=		

$$S^2 = \frac{\sum(xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	$(-0.32)^2$= 0.10
1.43	(1.43-1.69)= -0.26	$(-0.26)^2$= 0.07
1.46	(1.46-1.69)= -0.23	$(-0.23)^2$= 0.05
1.58	(1.58-1.69)= -0.11	$(-0.11)^2$= 0.01
1.60	(1.60-1.69)= -0.09	$(-0.09)^2$= 0.01
1.66	(1.66-1.69)= -0.03	$(-0.03)^2$= 0.00
1.67	(1.67-1.69)= -0.02	$(-0.02)^2$= 0.00
1.70	(1.70-1.69)= 0.01	$(0.01)^2$= 0.00
1.74	(1.74-1.69)= 0.05	$(0.05)^2$= 0.00
1.79	(1.79-1.69)= 0.10	$(0.10)^2$= 0.01
1.82	(1.82-1.69)= 0.13	$(0.13)^2$= 0.02
1.84	(1.84-1.69)= 0.15	$(0.15)^2$= 0.02
1.92	(1.92-1.69)= 0.23	$(0.23)^2$= 0.05
2.14	(2.14-1.69)= 0.45	$(0.45)^2$= 0.20
∑=		

$$S^2 = \frac{\sum (xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	$(-0.32)^2$= 0.10
1.43	(1.43-1.69)= -0.26	$(-0.26)^2$= 0.07
1.46	(1.46-1.69)= -0.23	$(-0.23)^2$= 0.05
1.58	(1.58-1.69)= -0.11	$(-0.11)^2$= 0.01
1.60	(1.60-1.69)= -0.09	$(-0.09)^2$= 0.01
1.66	(1.66-1.69)= -0.03	$(-0.03)^2$= 0.00
1.67	(1.67-1.69)= -0.02	$(-0.02)^2$= 0.00
1.70	(1.70-1.69)= 0.01	$(0.01)^2$= 0.00
1.74	(1.74-1.69)= 0.05	$(0.05)^2$= 0.00
1.79	(1.79-1.69)= 0.10	$(0.10)^2$= 0.01
1.82	(1.82-1.69)= 0.13	$(0.13)^2$= 0.02
1.84	(1.84-1.69)= 0.15	$(0.15)^2$= 0.02
1.92	(1.92-1.69)= 0.23	$(0.23)^2$= 0.05
2.14	(2.14-1.69)= 0.45	$(0.45)^2$= 0.20
∑=		

$$S^2 = \frac{\sum (xi - \mathbf{1.69})^2}{\mathbf{13}}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	$(-0.32)^2$= 0.10
1.43	(1.43-1.69)= -0.26	$(-0.26)^2$= 0.07
1.46	(1.46-1.69)= -0.23	$(-0.23)^2$= 0.05
1.58	(1.58-1.69)= -0.11	$(-0.11)^2$= 0.01
1.60	(1.60-1.69)= -0.09	$(-0.09)^2$= 0.01
1.66	(1.66-1.69)= -0.03	$(-0.03)^2$= 0.00
1.67	(1.67-1.69)= -0.02	$(-0.02)^2$= 0.00
1.70	(1.70-1.69)= 0.01	$(0.01)^2$= 0.00
1.74	(1.74-1.69)= 0.05	$(0.05)^2$= 0.00
1.79	(1.79-1.69)= 0.10	$(0.10)^2$= 0.01
1.82	(1.82-1.69)= 0.13	$(0.13)^2$= 0.02
1.84	(1.84-1.69)= 0.15	$(0.15)^2$= 0.02
1.92	(1.92-1.69)= 0.23	$(0.23)^2$= 0.05
2.14	(2.14-1.69)= 0.45	$(0.45)^2$= 0.20
$\sum$=		

+ + + + + +

$$S^2 = \frac{\sum (xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	$(-0.32)^2$= 0.10
1.43	(1.43-1.69)= -0.26	$(-0.26)^2$= 0.07
1.46	(1.46-1.69)= -0.23	$(-0.23)^2$= 0.05
1.58	(1.58-1.69)= -0.11	$(-0.11)^2$= 0.01
1.60	(1.60-1.69)= -0.09	$(-0.09)^2$= 0.01
1.66	(1.66-1.69)= -0.03	$(-0.03)^2$= 0.00
1.67	(1.67-1.69)= -0.02	$(-0.02)^2$= 0.00
1.70	(1.70-1.69)= 0.01	$(0.01)^2$= 0.00
1.74	(1.74-1.69)= 0.05	$(0.05)^2$= 0.00
1.79	(1.79-1.69)= 0.10	$(0.10)^2$= 0.01
1.82	(1.82-1.69)= 0.13	$(0.13)^2$= 0.02
1.84	(1.84-1.69)= 0.15	$(0.15)^2$= 0.02
1.92	(1.92-1.69)= 0.23	$(0.23)^2$= 0.05
2.14	(2.14-1.69)= 0.45	$(0.45)^2$= 0.20
∑=		0.54

$$S^2 = \frac{\sum (xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	$(-0.32)^2$= 0.10
1.43	(1.43-1.69)= -0.26	$(-0.26)^2$= 0.07
1.46	(1.46-1.69)= -0.23	$(-0.23)^2$= 0.05
1.58	(1.58-1.69)= -0.11	$(-0.11)^2$= 0.01
1.60	(1.60-1.69)= -0.09	$(-0.09)^2$= 0.01
1.66	(1.66-1.69)= -0.03	$(-0.03)^2$= 0.00
1.67	(1.67-1.69)= -0.02	$(-0.02)^2$= 0.00
1.70	(1.70-1.69)= 0.01	$(0.01)^2$= 0.00
1.74	(1.74-1.69)= 0.05	$(0.05)^2$= 0.00
1.79	(1.79-1.69)= 0.10	$(0.10)^2$= 0.01
1.82	(1.82-1.69)= 0.13	$(0.13)^2$= 0.02
1.84	(1.84-1.69)= 0.15	$(0.15)^2$= 0.02
1.92	(1.92-1.69)= 0.23	$(0.23)^2$= 0.05
2.14	(2.14-1.69)= 0.45	$(0.45)^2$= 0.20
∑=		0.54

$$S^2 = \frac{\sum(xi - 1.69)^2}{13}$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	$(-0.32)^2$= 0.10
1.43	(1.43-1.69)= -0.26	$(-0.26)^2$= 0.07
1.46	(1.46-1.69)= -0.23	$(-0.23)^2$= 0.05
1.58	(1.58-1.69)= -0.11	$(-0.11)^2$= 0.01
1.60	(1.60-1.69)= -0.09	$(-0.09)^2$= 0.01
1.66	(1.66-1.69)= -0.03	$(-0.03)^2$= 0.00
1.67	(1.67-1.69)= -0.02	$(-0.02)^2$= 0.00
1.70	(1.70-1.69)= 0.01	$(0.01)^2$= 0.00
1.74	(1.74-1.69)= 0.05	$(0.05)^2$= 0.00
1.79	(1.79-1.69)= 0.10	$(0.10)^2$= 0.01
1.82	(1.82-1.69)= 0.13	$(0.13)^2$= 0.02
1.84	(1.84-1.69)= 0.15	$(0.15)^2$= 0.02
1.92	(1.92-1.69)= 0.23	$(0.23)^2$= 0.05
2.14	(2.14-1.69)= 0.45	$(0.45)^2$= 0.20
∑=		0.54

$$S^2 = \frac{\sum (xi - 1.69)^2}{13}$$

$$S^2 = \frac{0.54}{13} = 0.04$$

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	$(-0.32)^2$= 0.10
1.43	(1.43-1.69)= -0.26	$(-0.26)^2$= 0.07
1.46	(1.46-1.69)= -0.23	$(-0.23)^2$= 0.05
1.58	(1.58-1.69)= -0.11	$(-0.11)^2$= 0.01
1.60	(1.60-1.69)= -0.09	$(-0.09)^2$= 0.01
1.66	(1.66-1.69)= -0.03	$(-0.03)^2$= 0.00
1.67	(1.67-1.69)= -0.02	$(-0.02)^2$= 0.00
1.70	(1.70-1.69)= 0.01	$(0.01)^2$= 0.00
1.74	(1.74-1.69)= 0.05	$(0.05)^2$= 0.00
1.79	(1.79-1.69)= 0.10	$(0.10)^2$= 0.01
1.82	(1.82-1.69)= 0.13	$(0.13)^2$= 0.02
1.84	(1.84-1.69)= 0.15	$(0.15)^2$= 0.02
1.92	(1.92-1.69)= 0.23	$(0.23)^2$= 0.05
2.14	(2.14-1.69)= 0.45	$(0.45)^2$= 0.20
∑=		0.54

$$S^2 = \frac{\sum (xi - 1.69)^2}{13}$$

$$S^2 = \frac{0.54}{13} = 0.04$$

Quanto maior a variância de uma série, maior a dispersão dos valores que a compõe. Assim, se uma amostra tem variância igual a 0.04 e outra, da mesma variável, igual a 0.10, nesta segundo os dados variam mais em relação a média que a primeira.

Medidas de dispersão ou de variabilidade

Medidas de altura (x)	(xi – média)	(xi – média)²
1.37	(1.37-1.69)= - 0.32	$(-0.32)^2$= 0.10
1.43	(1.43-1.69)= -0.26	$(-0.26)^2$= 0.07
1.46	(1.46-1.69)= -0.23	$(-0.23)^2$= 0.05
1.58	(1.58-1.69)= -0.11	$(-0.11)^2$= 0.01
1.60	(1.60-1.69)= -0.09	$(-0.09)^2$= 0.01
1.66	(1.66-1.69)= -0.03	$(-0.03)^2$= 0.00
1.67	(1.67-1.69)= -0.02	$(-0.02)^2$= 0.00
1.70	(1.70-1.69)= 0.01	$(0.01)^2$= 0.00
1.74	(1.74-1.69)= 0.05	$(0.05)^2$= 0.00
1.79		.01
1.82		.02
1.84	(1.84-1.69)= 0.15	$(0.15)^2$= 0.02
1.92	(1.92-1.69)= 0.23	$(0.23)^2$= 0.05
2.14	(2.14-1.69)= 0.45	$(0.45)^2$= 0.20
∑=		0.54

Quando não houver variabilidade, a variância é igual a zero.

$$S^2 = \frac{\sum(xi - 1.69)^2}{13}$$

$$S^2 = \frac{0.54}{13} = 0.04$$

Quanto maior a variância de uma série, maior a dispersão dos valores que a compõe. Assim, se uma amostra tem variância igual a 0.04 e outra, da mesma variável, igual a 0.10, nesta segundo os dados variam mais em relação a média que a primeira.

Medidas de dispersão ou de variabilidade

Desvio-padrão

- Uma dificuldade com a variância, como medida de dispersão, é o fato de não poder ser apresentada com a mesma unidade com que a variável foi medida.

$$S^2 = \frac{0.54}{13} = 0.04$$

Medidas de dispersão ou de variabilidade

Desvio-padrão

- Uma dificuldade com a variância, como medida de dispersão, é o fato de não poder ser apresentada com a mesma unidade com que a variável foi medida.

$$S^2 = \frac{0.54}{13} = 0.04$$

Medidas de dispersão ou de variabilidade

Desvio-padrão

- Uma dificuldade com a variância, como medida de dispersão, é o fato de não poder ser apresentada com a mesma unidade com que a variável foi medida.

$$S^2 = \frac{0.54}{13} = 0.04$$

- A solução é extrair a raiz quadrada positiva da variância, já que, com isso, se volta à unidade original da variável. Essa medida é chamada de Desvio-Padrão (S).

Medidas de dispersão ou de variabilidade

Desvio-padrão

- Uma dificuldade com a variância, como medida de dispersão, é o fato de não poder ser apresentada com a mesma unidade com que a variável foi medida.

$$S^2 = \frac{0.54}{13} = 0.04$$

- A solução é extrair a raiz quadrada positiva da variância, já que, com isso, se volta à unidade original da variável. Essa medida é chamada de Desvio-Padrão (S).

$$S = \sqrt{S^2}$$

Medidas de dispersão ou de variabilidade

Desvio-padrão

- Uma dificuldade com a variância, como medida de dispersão, é o fato de não poder ser apresentada com a mesma unidade com que a variável foi medida.

$$S^2 = \frac{0.54}{13} = 0.04$$

- A solução é extrair a raiz quadrada positiva da variância, já que, com isso, se volta à unidade original da variável. Essa medida é chamada de **Desvio-Padrão (S)**.

$$S = \sqrt{S^2} = \sqrt{0.04}$$

Medidas de dispersão ou de variabilidade

Desvio-padrão

- Uma dificuldade com a variância, como medida de dispersão, é o fato de não poder ser apresentada com a mesma unidade com que a variável foi medida.

$$S^2 = \frac{0.54}{13} = 0.04$$

- A solução é extrair a raiz quadrada positiva da variância, já que, com isso, se volta à unidade original da variável. Essa medida é chamada de **Desvio-Padrão (S)**.

$$S = \sqrt{S^2} = \sqrt{0.04} = 0.21$$

Medidas de dispersão ou de variabilidade

Desvio-padrão

- Uma dificuldade com a variância, como medida de dispersão, é o fato de não poder ser apresentada com a mesma unidade com que a variável foi medida.

$$S^2 = \frac{0.54}{13} = 0.04$$

Resultado:
Os valores se desviam da média de tamanho (1.69m) em ±0.21m.

- A solução é extrair a raiz quadrada positiva da variância, já que, com isso, se volta à unidade original da variável. Essa medida é chamada de Desvio-Padrão (S).

$$S = \sqrt{S^2} = \sqrt{0.04} = 0.21$$

Medidas de dispersão ou de variabilidade

Desvio-padrão

- Uma dificuldade com a variância, como medida de dispersão, é o fato de não poder ser apresentada com a mesma unidade com que a variável foi medida.

Quanto maior o desvio-padrão de uma série, maior a dispersão dos valores que a compõe. Assim, se uma amostra tem desvio de dados igual a 0.21 e outra, da mesma variável, igual a 0.40, nesta segundo os dados se desviam mais em relação a média que a primeira.

<u>Resultado</u>:
Os valores se desviam da média de tamanho (1.69m) em ±0.21m.

- A soluçã ... tiva da variância, já que, com isso, se volta à un dade original da variável. Essa medida é chamada

Quando não houver variabilidade, a variância é igual a zero.

$$\mathbf{S}= \sqrt{S^2} = \sqrt{0.04} = 0.21$$

Medidas de dispersão ou de variabilidade

Coeficiente de variação

- E quando temos duas variáveis com medidas diferentes? Por exemplo, queremos saber o que varia mais, peso (kg) ou altura (m)?

Rol ordenado:

Peso (kg)={40, 42, 45, 46, 50, 53, 55, 56, 58, 59, 60, 61, 63, 65}

Altura (m)={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

- Neste caso, é possível utilizar o coeficiente de variação percentual (CV%), que é uma medida de dispersão independente da unidade de mensuração da variável.

$$CV\% = \frac{S}{\bar{x}} * 100,$$

S= variância
$\bar{x}$= média

Medidas de dispersão ou de variabilidade

Rol ordenado:

Altura (m)={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Peso (kg)={49, 49, 50, 51 , 50, 53, 55, 56, 56, 57, 58, 60, 60, 61}

Medidas de dispersão ou de variabilidade

<u>Rol ordenado</u>:

Altura (m)={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Peso (kg)={49, 49, 50, 51 , 50, 53, 55, 56, 56, 57, 58, 60, 60, 61}

$Altura$

$$CV\% = \frac{S}{\bar{x}} * 100$$

$Peso$

$$CV\% = \frac{S}{\bar{x}} * 100$$

Medidas de dispersão ou de variabilidade

<u>Rol ordenado</u>:

Altura (m)={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Peso (kg)={49, 49, 50, 51 , 50, 53, 55, 56, 56, 57, 58, 60, 60, 61}

$Altura$

$$CV\% = \frac{S}{\bar{x}} * 100 \longrightarrow$$

$Altura$

$$CV\% = \frac{0.04}{1.69} * 100 = 12.16\%$$

$Peso$

$$CV\% = \frac{S}{\bar{x}} * 100$$

Medidas de dispersão ou de variabilidade

<u>Rol ordenado</u>:

Altura (m)={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Peso (kg)={49, 49, 50, 51 , 50, 53, 55, 56, 56, 57, 58, 60, 60, 61}

Altura

$$CV\% = \frac{S}{\bar{x}} * 100 \longrightarrow CV\% = \frac{0.04}{1.69} * 100 = 12.16\%$$

Peso

$$CV\% = \frac{S}{\bar{x}} * 100 \longrightarrow CV\% = \frac{18.55}{54.64} * 100 = 33.96\%$$

Medidas de dispersão ou de variabilidade

Rol ordenado:

Altura (m)={1.37, 1.43, 1.46, 1.58, 1.60, 1.66, 1.67, 1.70, 1.74, 1.79, 1.82, 1.84, 1.92, 2.14}

Peso (kg)={49, 49, 50, 51 , 50, 53, 55, 56, 56, 57, 58, 60, 60, 61}

$Altura$

$$CV\% = \frac{S}{\bar{x}} * 100 \longrightarrow$$

$Altura$

$$CV\% = \frac{0.04}{1.69} * 100 = 12.16\%$$

$Peso$

$$CV\% = \frac{S}{\bar{x}} * 100 \longrightarrow$$

$Peso$

$$CV\% = \frac{18.55}{54.64} * 100 = 33.96\%$$

Pode-se, agora, verificar que o Peso (kg) é uma característica mais variável que a Altura (m).

Medidas de dispersão ou de variabilidade

Quartis

- Quando dividimos a nossa série de dados para uma variável 'x' em quatro grupos, cada um contendo 25% dos dados.

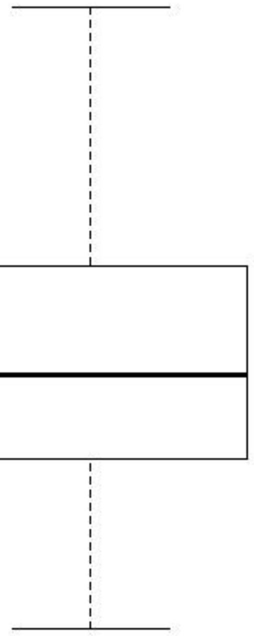


Diagrama de caixa
(boxplot)

Medidas de dispersão ou de variabilidade

Quartis

- Quando dividimos a nossa série de dados para uma variável ‘x’ em quatro grupos, cada um contendo 25% dos dados.

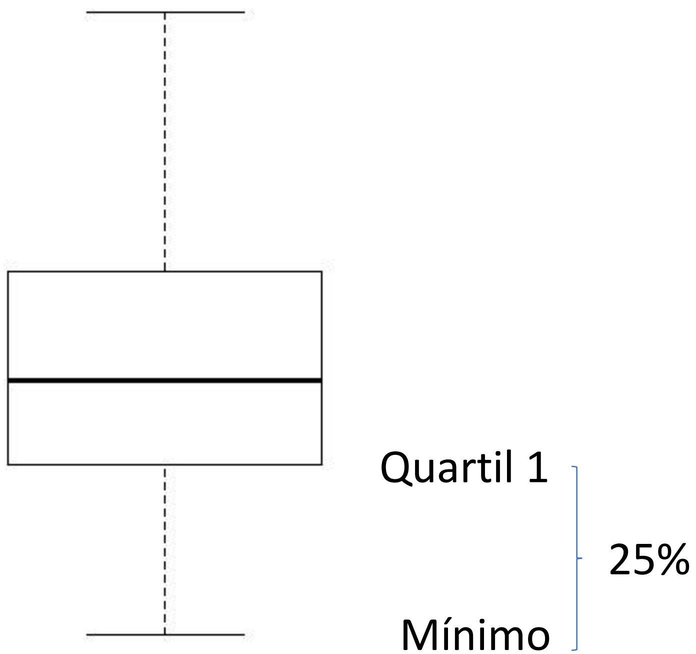


Diagrama de caixa
(boxplot)

Medidas de dispersão ou de variabilidade

Quartis

- Quando dividimos a nossa série de dados para uma variável ‘x’ em quatro grupos, cada um contendo 25% dos dados.

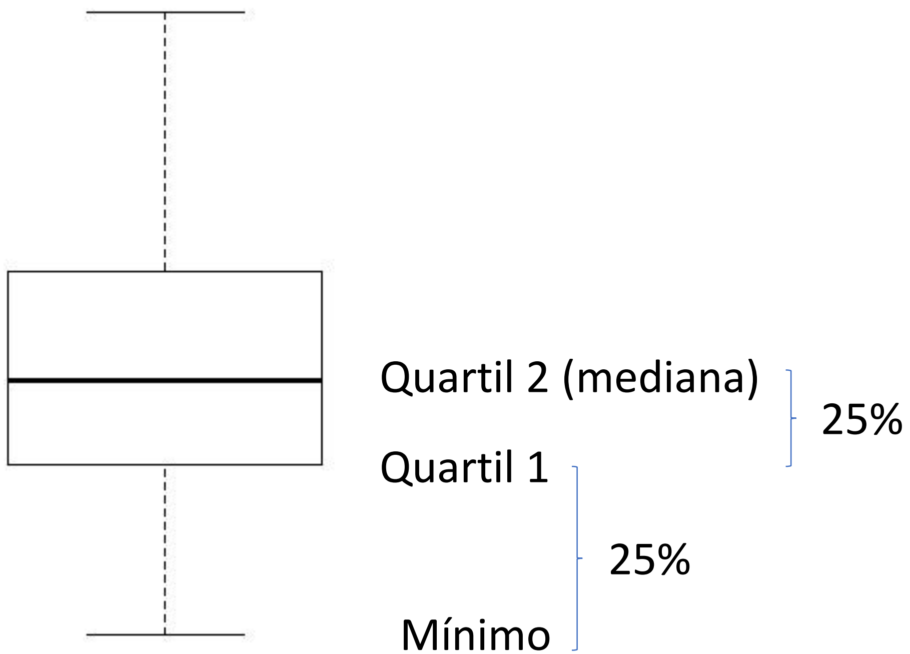


Diagrama de caixa
(boxplot)

Medidas de dispersão ou de variabilidade

Quartis

- Quando dividimos a nossa série de dados para uma variável ‘x’ em quatro grupos, cada um contendo 25% dos dados.

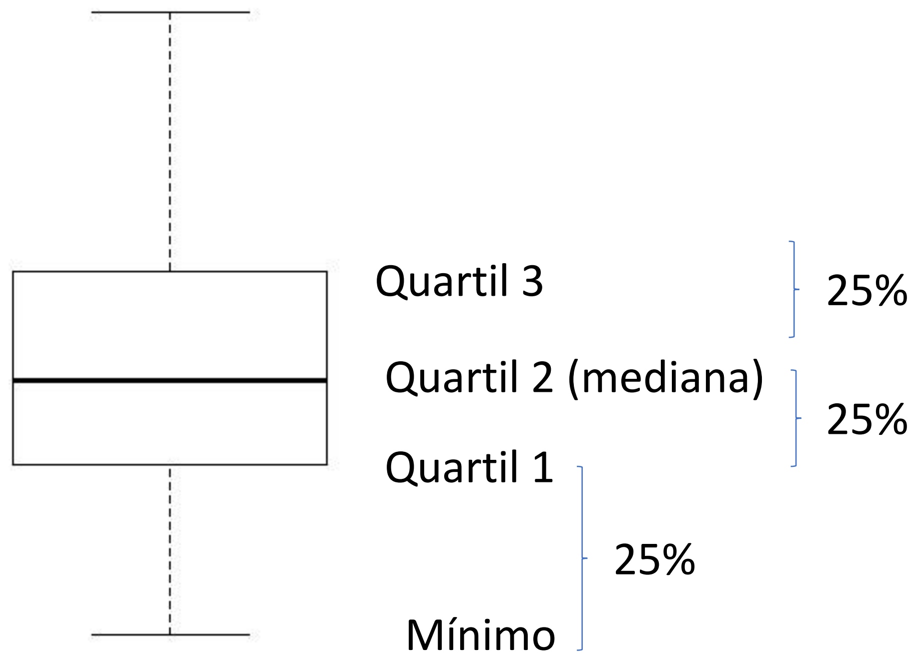


Diagrama de caixa (boxplot)

Medidas de dispersão ou de variabilidade

Quartis

- Quando dividimos a nossa série de dados para uma variável ‘x’ em quatro grupos, cada um contendo 25% dos dados.

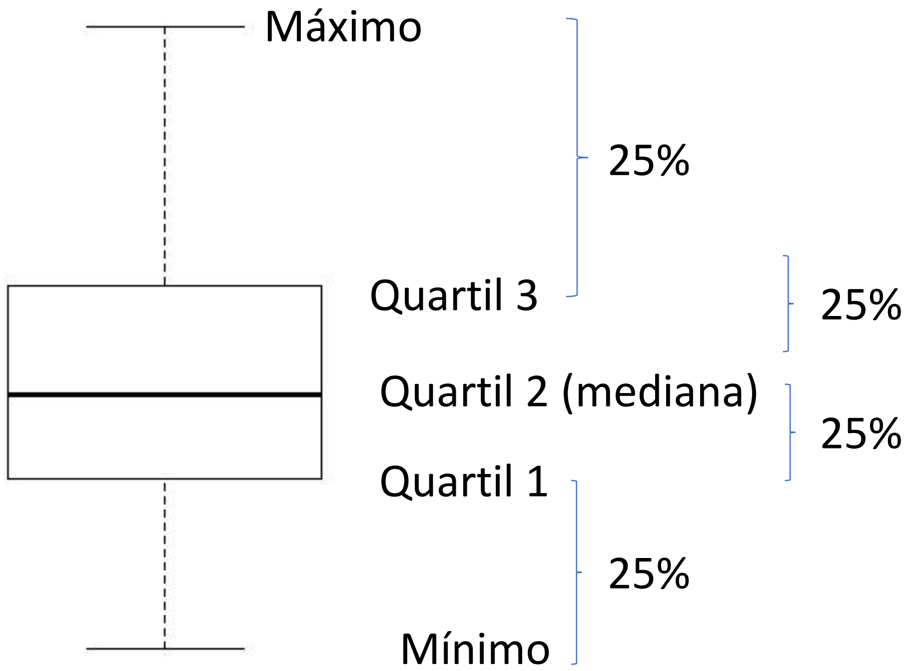


Diagrama de caixa
(boxplot)

Medidas de dispersão ou de variabilidade

Quartis

- Quando dividimos a nossa série de dados para uma variável ‘x’ em quatro grupos, cada um contendo 25% dos dados.

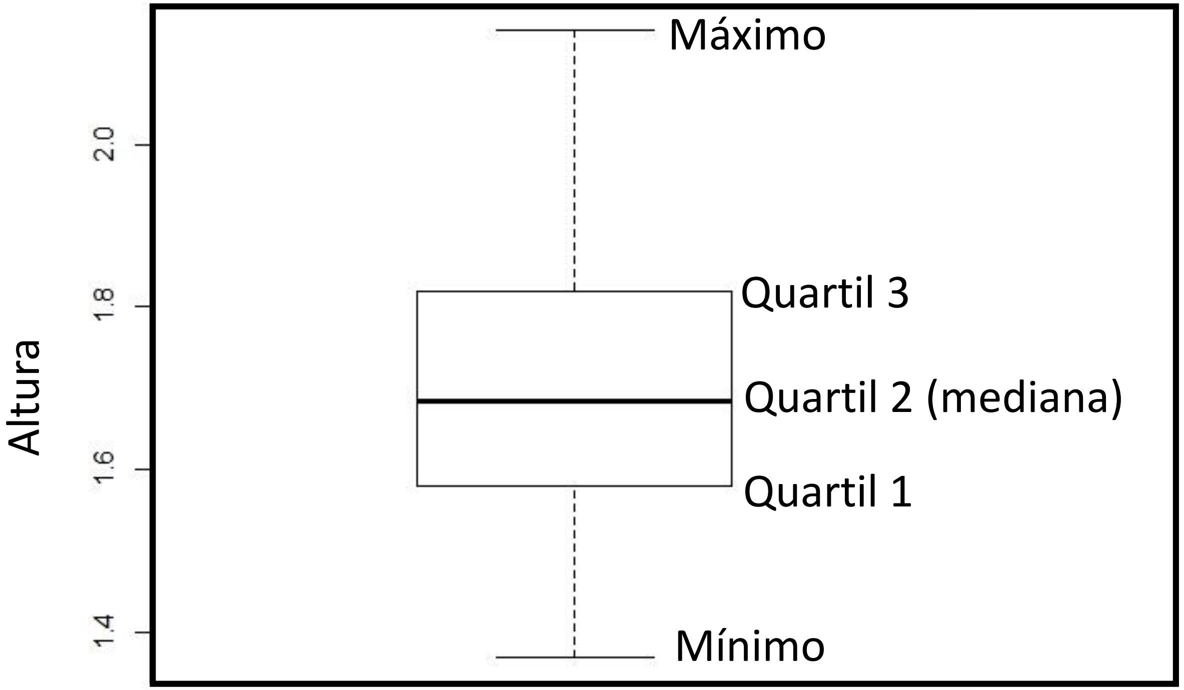


Diagrama de caixa
(boxplot)

Medidas de dispersão ou de variabilidade

Quartis

- Quando dividimos a nossa série de dados para uma variável ‘x’ em quatro grupos, cada um contendo 25% dos dados.

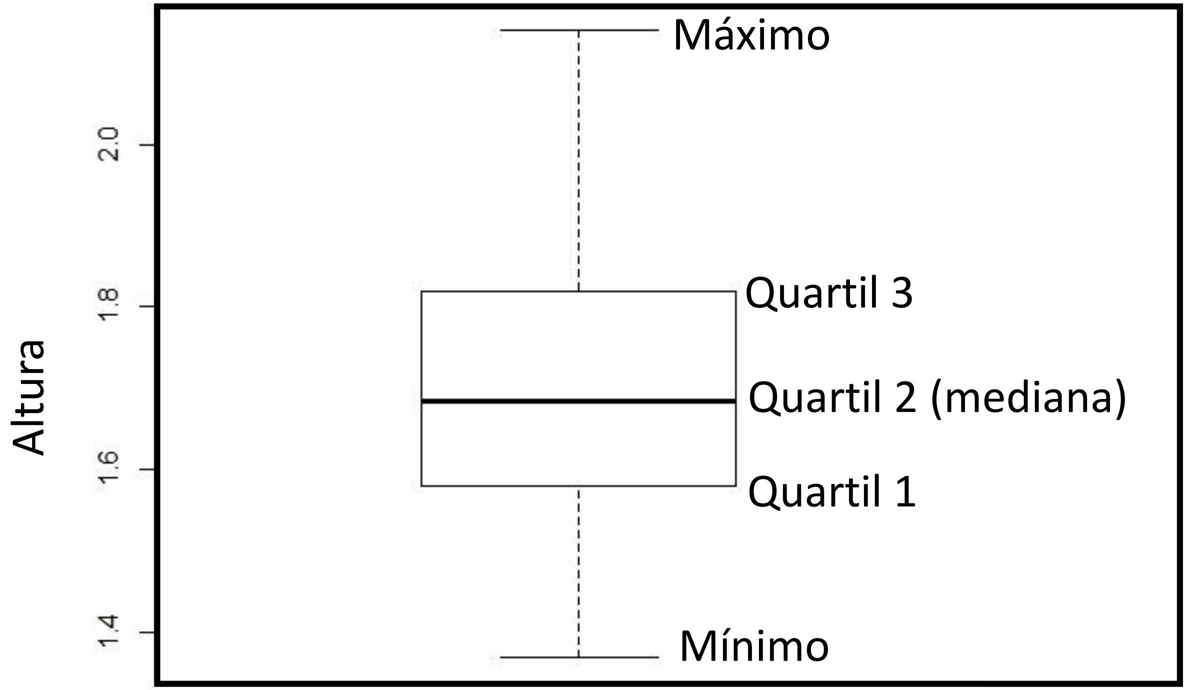
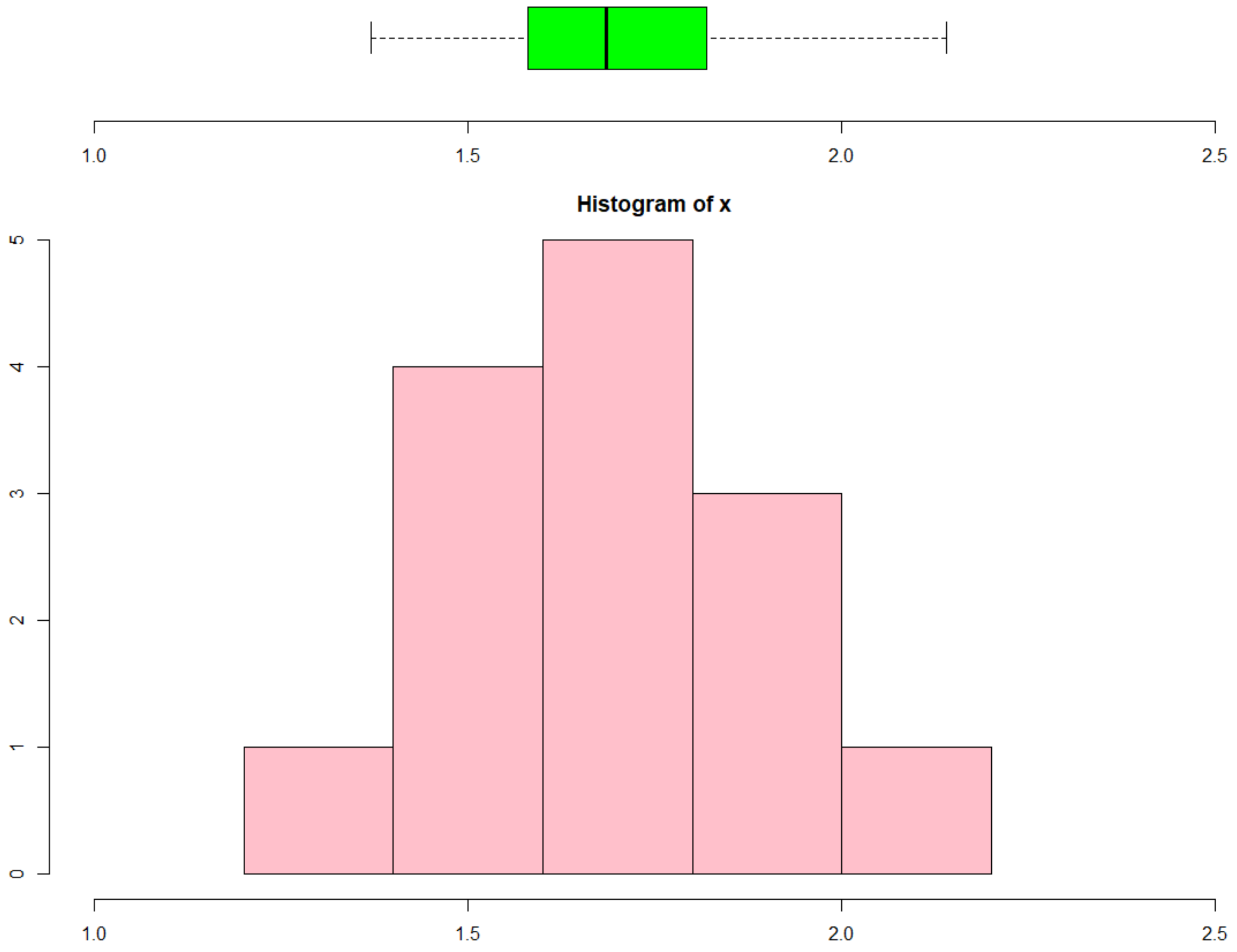


Diagrama de caixa (boxplot)

Quando se descreve um conjunto de dados de distribuição assimétrica, a distância entre quartis representa melhor a variação do que a amplitude ou o desvio padrão, porque não é afetada pelos valores extremos.

Medidas de dispersão ou de variabilidade



Atividade

- Você foi incumbido de caracterizar os cientistas de dados que estão participando do Grupo de Estudos em Ciência de Dados (GECD) de Foz do Iguaçu.
- Para tal, selecione dois tipos de variáveis quantitativas discretas, duas contínuas, duas qualitativas nominais e duas ordinais para a descrição.
- Selecione aleatoriamente 15 colegas de turma para participar de seu estudo.
- Você precisará tabelar os dados e depois descrever cada um deles. Você poderá usar tabelas com análises estatísticas descritivas e gráficos para tal.