

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO MATEMÁTICA
MESTRADO PROFISSIONAL EM EDUCAÇÃO MATEMÁTICA**

**Victor Ferreira Junqueira
Lupércio França Bessegato**

Uma sequência de atividades de medidas de posição por meio de boxplots e tabelas, na perspectiva da Educação Estatística

Juiz de fora
2026

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO MATEMÁTICA
MESTRADO PROFISSIONAL EM EDUCAÇÃO MATEMÁTICA**

**Victor Ferreira Junqueira
Lupércio França Bessegato**

Uma sequência de atividades de medidas de posição por meio de boxplots e tabelas, na perspectiva da Educação Estatística

Produto Educacional apresentado ao Programa de Pós-Graduação em Educação Matemática da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Educação Matemática. Área de Concentração: Educação Matemática.

Juiz de fora
2026



CC BY-NC 4.0

Este trabalho está licenciado com uma Licença [Creative Commons – Atribuição – NãoComercial 4.0 Internacional](http://creativecommons.org/licenses/by-nc/4.0/).

```
<a rel="license" href="http://creativecommons.org/licenses/by-nc/4.0/"></a><br  
  href="http://creativecommons.org/licenses/by-nc/4.0/">Creative  
  Commons - Atribuição-NãoComercial 4.0 Internacional</a>.
```

LISTA DE GRÁFICOS

Gráfico 1 - Gráfico de pontos da média aritmética da quantidade de sódio dos 20 alimentos matinais.....	17
Gráfico 2 - Gráfico de pontos da média aritmética e mediana da quantidade de sódio dos 20 alimentos matinais.....	19
Gráfico 3 - Comparando a média aritmética e a mediana em distribuições simétricas e assimétricas.....	20
Gráfico 4 - Gráfico de pontos interativos, com a entrada de registro errôneo no valor de 900 mg de sódio, ao invés de 340.....	21
Gráfico 5 - Representação gráfica dos quartis da quantidade de sódio dos 20 alimentos matinais.....	23
Gráfico 6 - Representação gráfica de percentis da pontuação de um aluno em um vestibular.....	24
Gráfico 7 - Representação gráfica dos possíveis outliers da quantidade de sódio dos 20 alimentos matinais.....	26
Gráfico 8 - Gráfico de pontos da quantidade ideal de filhos para uma família classificada por gênero de cada estudante.....	29
Gráfico 9 - Estrutura de um gráfico boxplot.....	39
Gráfico 10 - Renda média x Mesorregiões.....	43
Gráfico 11 - Renda média x Sexo (todos os municípios do estado do RJ).....	47

LISTA DE FIGURAS

- Figura 1 - Tela inicial do aplicativo web, destacando a opção 1 de inserção manual de dados..... 41
- Figura 2 - Opção 2 de inserção dos dados por meio do controle deslizante..... 42

LISTA DE TABELAS

Tabela 1 - Quantidades de sódio e açúcar em 20 alimentos matinais.....	16
Tabela 2 - Quantidades de sódio dos 20 alimentos matinais em lista ordenada de ordem crescente.....	18
Tabela 3 - Quantidade ideal de filhos para uma família classificada por gênero de cada estudante.....	28

LISTA DE ABREVIATURAS E SIGLAS

AED	Análise Exploratória em Dados
<i>CODAP</i>	<i>Common Online Data Analysis Platform</i>
COVID-19	Doença do Coronavírus 2019
IBGE	Instituto Brasileiro de Geografia e Estatística
MCS	Modelo dos Campos Semânticos
TIC	Tecnologia da Informação e Comunicação
<i>VIT</i>	<i>Visual Inference Tools</i>

LISTA DE SÍMBOLOS

n	Observação/Tamanho
X	Variável aleatória quantitativa
$x_{(1)}$	Primeiro valor da observação
$x_{(2)}$	Segundo valor da observação
$x_{(n)}$	N-ésimo valor da observação
\bar{x}	Média aritmética
Σ	Somatório de todos os valores
x_i	Índice de observação
\tilde{x}	Mediana
$x_{\left(\frac{n+1}{2}\right)}$	Ponto de posicionamento
$x_{\left(\frac{n}{2}\right)}$	Ponto de posicionamento
$x_{\left(\frac{n}{2}+1\right)}$	Ponto de posicionamento
%	Porcentagem
q_1	Primeiro quartil
q_2	Segundo quartil
q_3	Terceiro quartil
R	Amplitude
IQ	Distância Interquartil
LS	Limite superior

LI	Limite inferior
s^2	Variância
s	Desvio padrão
m_0	Moda
cv	Coeficiente de variação

SUMÁRIO

1	APRESENTAÇÃO.....	11
2	MEDIDAS RESUMO.....	13
2.1	MÉDIA ARITMÉTICA.....	15
2.2	MEDIANA.....	17
2.3	QUARTIS.....	21
2.4	PERCENTIS.....	24
2.5	AMPLITUDE E A DISTÂNCIA INTERQUARTIL.....	25
2.6	VARIÂNCIA E DESVIO PADRÃO.....	27
2.7	MODA.....	30
2.8	COEFICIENTE DE VARIAÇÃO.....	30
3	TECNOLOGIAS DIGITAIS NO ENSINO DE ESTATÍSTICA.....	32
4	APRESENTAÇÃO DAS TAREFAS.....	38
4.1	TAREFA 1 - O QUE SÃO DADOS EM ESTATÍSTICA E DADOS REPRESENTADOS POR MEIO DE GRÁFICOS BOXPLOTS.....	38
4.2	TAREFA 2 - PROBLEMÁTICA ENVOLVENDO DADOS REPRESENTADOS POR MEIO DE GRÁFICOS BOXPLOTS.....	42
5	SUGESTÕES.....	50
5.1	SUGESTÕES DE LEITURAS.....	50
5.2	SUGESTÕES DE VÍDEOS.....	50
5.3	SUGESTÕES DE CONJUNTO DE DADOS.....	51
5.4	SUGESTÕES DE PORTAIS.....	51
6	ALGUMAS CONSIDERAÇÕES.....	53
	REFERÊNCIAS.....	55
	APÊNDICES.....	59
	APÊNDICE A - TABELAS A SEREM UTILIZADAS NO DESENVOLVIMENTO DAS TAREFAS.....	60

1 APRESENTAÇÃO

Caro(a) professor(a),

Este Produto Educacional é fruto da Dissertação de Mestrado intitulada: “A produção de significados de estudantes do Ensino Médio para as medidas de posição por meio de boxplots e tabelas”, que foi desenvolvida no Programa de Pós-Graduação em Educação Matemática da Universidade Federal de Juiz de Fora e é destinada aos professores de Matemática. As tarefas foram elaboradas e aplicadas para alunos do Ensino Médio, visando a oferecer uma alternativa para o ensino das medidas resumo, com ênfase no conteúdo das medidas de posição por meio de visualização de boxplots, apoiada por tabelas, com o objetivo de fortalecer o processo de compreensão dessas medidas. Desta forma, entendemos que, na pesquisa de campo, além de ter sido reforçado o Raciocínio Estatístico, o conjunto de tarefas também favoreceu a mobilização de aspectos do Letramento e do Pensamento Estatísticos dos alunos, na medida em que eles foram conduzidos a interpretar dados representados em boxplots, com o apoio de tabelas, bem como a discutir e justificar conclusões fundamentadas, por meio de registros escritos. A viabilidade dessa abordagem encontra respaldo em Gómez-Blancarte, Chávez e Aguilar (2021). As tarefas desenvolvidas foram analisadas de acordo com o Modelo dos Campos Semânticos (MCS).

Esse Produto Educacional está estruturado em cinco outros capítulos. O próximo capítulo aborda as medidas resumo utilizadas no Ensino de Estatística na Educação Básica como suportes para basear o planejamento das tarefas, além de ter como propósito a conceituação do conteúdo de medidas resumo para os professores [[capítulo 2](#)]. Dada a disponibilidade crescente de Tecnologias Digitais, no [capítulo 3](#), é discutido seu uso no Ensino de Estatística, descrevendo-se principalmente a relevância das tecnologias e dos *softwares* educacionais no processo de Ensino e Aprendizagem. No [capítulo 4](#), são apresentadas as tarefas sobre medidas de posição, com auxílio de boxplots e tabelas, as quais podem apoiar a didática dos professores e a metodologia de ensino alternativa, proporcionando uma Aprendizagem envolvente. No [capítulo 5](#), apresentamos uma lista de leituras e vídeos que podem complementar a utilização pelos professores deste material em aulas de Matemática e Estatística. Além disso, estão disponíveis algumas

plataformas de dados de estatísticas sociais com livre acesso, bem como uma relação de portais que podem apoiar o processo de Ensino e Aprendizagem de Estatística. No [capítulo 6](#), apresentamos algumas considerações relativas ao uso deste Produto Educacional. Por fim, apresenta-se o [Apêndice A](#), contendo as tabelas que devem ser utilizadas no desenvolvimento das tarefas.

Vale ressaltar que, após a pesquisa de campo, percebemos que a ordem de apresentação e a estrutura das tarefas não foram apropriadas. Por esse motivo, o Produto Educacional foi elaborado com apresentação de tarefas distintas daquelas utilizadas na pesquisa de campo, bem como tabelas de dados e questionário. Desse modo, estes materiais foram adaptados e a proposta apresentada no Produto Educacional será diferente da que foi utilizada na pesquisa de campo.

As tarefas foram planejadas e elaboradas com o propósito de desenvolver a leitura crítica e a tomada de decisão por parte dos estudantes. Ressalta-se que elas podem ser aplicadas na íntegra ou ainda ser adaptadas, conforme o tema em estudo e o público alvo. Nesse caso, deve-se respeitar as especificidades e os interesses dos alunos, assim como a realidade de cada escola.

Portanto, este Produto Educacional tem o objetivo de auxiliar os professores que desejam abordar o tema em sala de aula de forma reflexiva e crítica, explorando tanto o contexto quanto os conceitos estatísticos, de modo a estimular possibilidades de aplicação relacionadas ao assunto.

Desejamos uma ótima leitura e sucesso na utilização deste material nas práticas pedagógicas do Ensino de Matemática e Estatística.

Victor Ferreira Junqueira
Lupércio França Bessegato

2 MEDIDAS RESUMO

Magalhães e Lima (2023) salientam que as medidas resumo, tanto de tendência central quanto de dispersão, são utilizadas para sumarizar os dados disponíveis visando obter uma síntese de informações sobre o comportamento de uma variável¹ a partir de um conjunto de dados qualquer. O objetivo é caracterizar esse conjunto por meio de medidas que resumam a informação, por exemplo, representando principalmente a posição dos dados e sua dispersão ou variabilidade. Alguns autores denominam as medidas de posição por medidas de tendência central ou por medidas de locação. Moore e McCabe (2002) apresentam como medidas resumo de posição as seguintes funções dos dados observados: a média aritmética, ou apenas média, a moda e a mediana. Por outro lado, a amplitude, a variância, o desvio padrão e o intervalo interquartilico são medidas resumo da dispersão dos dados, ou seja, o quanto seus valores estão dispersos ou concentrados. As medidas resumo são consideradas de grande importância na interpretação contextualizada e crítica dos dados. Todas as medidas citadas anteriormente serão definidas e exemplificadas no presente capítulo.

Garfield e Ben-Zvi (2008) destacam que é importante entender o conceito de centro de uma distribuição de dados como um “sinal” em meio a “ruído” (variação). Ele é um componente chave na compreensão do conceito de distribuição, sendo essencial para interpretar gráficos e explorar um conjunto de dados. Os autores recomendam que os alunos desenvolvam, inicialmente, atividades com ideias informais de locação e dispersão, interpretando, descrevendo e representando graficamente o padrão de distribuição dos dados. Mais tarde, eles constroem a ideia de centro de uma maneira mais formal, desenvolvendo diferentes medidas de posição e de dispersão, apresentando como calculá-las, detalhando as informações que elas fornecem e suas aplicações. Além disso, esses autores enfatizam que é impossível analisar as medidas de posição sem considerar também as medidas de dispersão, pois ambas as ideias são necessárias para dar sentido à interpretação dos dados. Cobb e Moore (1997), Fife, James e Peters (2020), Watson, *et al.* (2003), dentre outros, ressaltam a importância da variabilidade na exploração de dados. Eles enfatizam que se não houvesse variabilidade, não haveria análise estatística, pois

¹ É qualquer característica, atributo ou quantidade que pode assumir diferentes valores em um conjunto de dados. Em outras palavras, é aquilo que você observa, mede e registra em uma pesquisa ou experimento, e que varia de um indivíduo para outro ou de uma situação para outra.

essa está baseada na variação dos dados. A variabilidade ajuda a estabelecer uma clara distinção entre a Estatística e outras áreas de conhecimento, como por exemplo a Matemática. Para Shaughnessy (1997), o conceito de variabilidade é complexo e difícil de definir, reforçando a necessidade de ampliar a sua compreensão. Ele é primordial para que os estudantes desenvolvam adequadamente o Pensamento Estatístico. Por sua vez, Burril e Biehler (2011), Shaughnessy (2007), Isoda, Chitmun e Gonzalez (2018), dentre outros, salientam que a ideia de variabilidade vai além do simples cálculo matemático das várias medidas de variação, tais como amplitude, variância, desvio padrão e intervalo interquartil. Segundo esses autores, para entender as perspectivas teóricas de variabilidade e suas aplicações, é preciso superar a preocupação centrada apenas no simples tratamento matemático, a qual faz os estudantes realizarem cálculos mecanicamente, dos quais muitas vezes não entendem seu significado real. Seguindo essa perspectiva, Souza, Cazorla e Monteiro (2024) alertam para a importância da variabilidade nas análises estatísticas, sendo necessário estimular o estudante a questionar e criticar a informação que está sendo gerada. Os autores reforçam que o foco na matematização do conhecimento estatístico dificulta a percepção do fenômeno da variabilidade como componente importante dos dados, o qual está intimamente relacionado com o Letramento Estatístico. Por conta disso, as medidas resumo podem auxiliar bastante na análise estatística de dados, permitindo assim que os cidadãos elaborem análises e interpretações com mais consistência, assimilando as informações do mundo com uma perspectiva mais crítica.

Moore e McCabe (2002) destacam ainda que a descrição de um padrão de distribuição dos dados quase sempre inclui uma medida de posição. As duas medidas de posição mais comuns são a média e a mediana. Para esses autores, a média representa o “valor médio”, enquanto a mediana é o “valor central”. Esses dois conceitos refletem ideias diferentes para a posição de uma distribuição, e as duas medidas se comportam de maneira diferente. Batanero (2000) salienta que é necessário um trabalho específico para que os estudantes possam compreender os distintos significados dos conceitos de média e mediana. Por outro lado, Novaes e Coutinho (2013) reforçam que a análise das medidas de posição e de dispersão complementa as interpretações das representações gráficas e tabulares dos dados, levando a uma análise mais detalhada. Para Agresti, Franklin e Klingenberg (2023), uma medida de posição não é suficiente para descrever adequadamente a

distribuição de uma variável, pois ela não fornece informações sobre a variabilidade dos dados. Por isso, os autores destacam a importância de considerar tanto as medidas de posição quanto as de dispersão, já que ambas são essenciais para uma interpretação completa e significativa dos dados.

A partir desse ponto, consideramos um conjunto de dados com n medidas de uma variável aleatória quantitativa X . Os valores observados dessa amostra podem ser contínuos ou discretos e são denotados por $x_{(1)}, x_{(2)}, \dots, x_{(n)}$. A seguir, definiremos as medidas de posição e de dispersão utilizadas nesse trabalho, e utilizaremos como exemplos os conjuntos de dados apresentados por Agresti, Franklin e Klingenberg (2023) para apresentar as aplicabilidades das medidas resumo.

2.1 MÉDIA ARITMÉTICA

A média aritmética, ou simplesmente média, de um conjunto de observações será denotada por \bar{x} . Sua expressão é dada por:

$$\bar{x} = \frac{x_{(1)} + x_{(2)} + \dots + x_{(n)}}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

Para explicar o conceito de média aritmética, utilizaremos a [Tabela 1](#) de Agresti, Franklin e Klingenberg (2023), apresentando 20 alimentos matinais populares, indicando as quantidades de sódio e açúcar contidas em uma porção. Ambas as variáveis, sódio e açúcar, são quantitativas e contínuas, pois medem quantidades que podem assumir quaisquer valores de números reais positivos, incluindo o valor zero. No entanto, na tabela, os valores foram arredondados para o inteiro mais próximo, em gramas para o açúcar e miligramas para o sódio. Isso faz com que os dados aparentem ser discretos, mas são analisados e interpretados como dados contínuos.

Exemplo 2.1:

Tabela 1 - Quantidades de sódio e açúcar em 20 alimentos matinais.

Alimentos	Sódio (mg)	Açúcar (g)	Tipo ²
Alimento A	0	11	Adulto
Alimento B	340	18	Adulto
Alimento C	70	5	Adulto
Alimento D	140	14	Criança
Alimento E	200	12	Criança
Alimento F	180	1	Criança
Alimento G	210	10	Criança
Alimento H	150	16	Adulto
Alimento I	100	0	Adulto
Alimento J	130	12	Criança
Alimento K	140	14	Criança
Alimento L	180	7	Adulto
Alimento M	190	9	Criança
Alimento N	160	6	Criança
Alimento O	290	3	Criança
Alimento P	50	15	Adulto
Alimento Q	220	4	Adulto
Alimento R	180	4	Adulto
Alimento S	200	3	Adulto
Alimento T	210	11	Criança

Fonte: Síntese do autor a partir do exemplo de Agresti, Franklin e Klingenberg (2023), p.67.

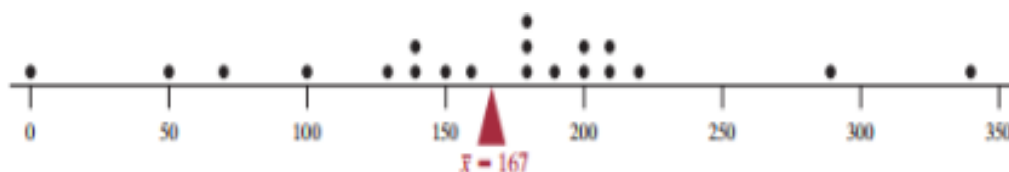
Aplicando a eq. (1), a quantidade do valor médio de sódio dos 20 alimentos consumidos no período da manhã, pelos dados da [Tabela 1](#), é:

$$\bar{x} = \frac{0 + 340 + 70 + \dots + 210}{20} = \frac{3340}{20} = 167$$

Essa média aritmética é o ponto de equilíbrio dos dados. Se distribuíssemos pesos iguais ao longo de uma linha representando as observações dos 20 alimentos matinais, essa linha se equilibraria ao posicionar o ponto de apoio exatamente na média aritmética. O [Gráfico 1](#), a seguir, ilustra essa ideia de forma visual:

² A variável categórica (Tipo) classifica os alimentos como sendo popular para adultos ou crianças.

Gráfico 1 - Gráfico de pontos da média aritmética da quantidade de sódio dos 20 alimentos matinais.



Fonte: Agresti, Franklin e Klingenberg (2023), p.84.

Além disso, a média aritmética pode ser entendida como o “valor da ação”. Por exemplo, ao analisarmos os 20 alimentos, a média de 167 mg pode ser interpretada como a quantidade de sódio presente, aproximadamente, em cada porção de alimento que contém todos os 20 alimentos que têm a mesma quantidade. Dessa forma, compreendemos que a média aritmética é exatamente a soma dos valores observados dividido pela quantidade de observações. Com isso, a média leva em conta os valores numéricos reais de todas as observações analisadas.

2.2 MEDIANA

A mediana, que iremos denotar por \tilde{x} , é o ponto médio de uma distribuição na análise de dados (um valor da variável, tal que metade das observações é inferior a ele, e a outra metade é superior). Conforme Moore e McCabe (2002), para determinar a mediana de uma distribuição, devemos seguir os seguintes passos:

- i) Disponha todas as observações em ordem de tamanho, do menor para o maior (lista ordenada de ordem crescente).
- ii) Se o número n de observações é ímpar, a mediana \tilde{x} é a observação central na lista ordenada de ordem crescente.
- iii) Se o número n de observações é par, a mediana \tilde{x} é a média das duas observações centrais na lista ordenada de ordem crescente.

Em notação matemática, os itens ii) e iii) são apresentados da seguinte forma:

$$\tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & , \text{ se } n \text{ é ímpar;} \\ \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} & , \text{ se } n \text{ é par;} \end{cases} ; \text{ em que: } x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}. \quad (2)$$

Sendo os valores de quantidade de sódio apresentados na [Tabela 1](#), os quais estão apresentados ordenadamente na [Tabela 2](#) (passo i). Para encontrar a mediana da quantidade de sódio dos 20 alimentos matinais, deve-se utilizar o procedimento descrito (passo iii).

Exemplo 2.2:

Tabela 2 - Quantidades de sódio dos 20 alimentos matinais em lista ordenada de ordem crescente.

Sódio (mg)																			
0	50	70	100	130	140	140	150	160	180	180	180	190	200	200	210	210	220	290	340

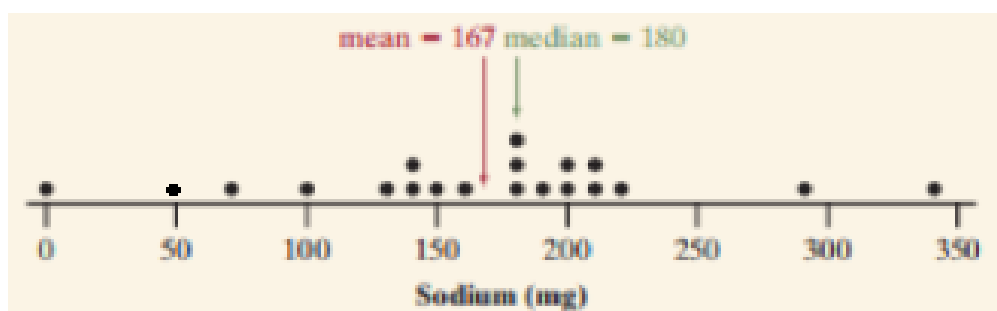
Fonte: Síntese do autor a partir do exemplo de Agresti, Franklin e Klingenberg (2023), p.83.

Na [Tabela 2](#), é possível perceber que há nove observações à esquerda, e nove à direita dos números em negrito. Aplicando a eq. (2), temos:

$$\tilde{x} = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} = \frac{x_{(10)} + x_{(11)}}{2} = \frac{180 + 180}{2} = 180$$

Portanto, existem 10 observações menores ou iguais a 180, e outras 10 observações maiores ou iguais a 180. Assim, a mediana é um valor que divide o conjunto de dados, organizados de forma crescente, em dois subconjuntos com iguais quantidades de elementos. Desta forma, no caso em que a quantidade de observações for par, a mediana não é uma das observações. No [Gráfico 2](#), apresentamos a mediana e a média aritmética:

Gráfico 2 - Gráfico de pontos da média aritmética e mediana da quantidade de sódio dos 20 alimentos matinais.



Fonte: Agresti, Franklin e Klingenberg (2023), p.88

Observe que a média aritmética e a mediana assumem valores diferentes. A mediana representa o centro dos dados, dividindo-os em duas partes iguais, independentemente dos valores numéricos das observações acima ou abaixo dela.

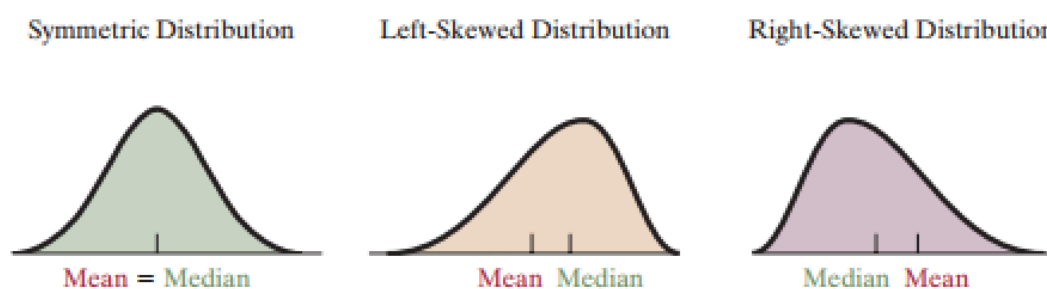
Convém considerar o caso de termos um número ímpar de observações, como 21, em vez das 20 usadas como no exemplo 2.2. Nesse caso, haverá um único valor que ocupa a função central quando os dados são organizados de forma crescente (11ª observação), com dez valores menores à esquerda e dez maiores à direita dele. Dessa forma, quando o conjunto de dados possui um número ímpar de observações, a mediana coincide necessariamente com um dos valores observados, ou seja, a mediana é simplesmente o valor central.

Note-se, ainda, que a mediana, diferentemente da média aritmética, é uma medida robusta³ de valores extremos. No exemplo 2.2, apresentado no [Gráfico 2](#), o último alimento possui uma quantidade de 340 mg de sódio. Caso esse valor tivesse sido registrado erroneamente como 900 mg, a mediana permaneceria inalterada, enquanto a média aritmética seria significativamente afetada. Para Agresti, Franklin e Klingenberg (2023, p.87), “Um resumo numérico das observações é chamado resistente se observações extremas tiverem pouca ou nenhuma influência no seu valor.” Nesse exemplo hipotético, percebe-se que a mediana é resistente ao valor extremo registrado no conjunto de dados, ou, em outras palavras, ela é robusta a valores atípicos (veja no [Gráfico 4](#)).

³ É uma medida pouco sensível a valores extremos. Em outras palavras, resiste a distorções causadas por dados atípicos.

Conforme destacado por Moore e McCabe (2002), a mediana e a média aritmética são as medidas de posição mais comuns de uma distribuição e, dependendo dos seus valores, podemos discutir a simetria da distribuição⁴. Quando a distribuição de um conjunto de dados é simétrica, a média aritmética e a mediana praticamente coincidem. Em distribuições assimétricas, como no caso ilustrado no [Gráfico 2](#), a média aritmética tende a se afastar da mediana na direção da cauda mais longa da distribuição⁵. Esta síntese a respeito das formas de uma distribuição de dados em relação à média aritmética e a mediana pode ser ilustrada no [Gráfico 3](#):

Gráfico 3 - Comparando a média aritmética e a mediana em distribuições simétricas e assimétricas



Fonte: Agresti, Franklin e Klingenberg (2023), p.86

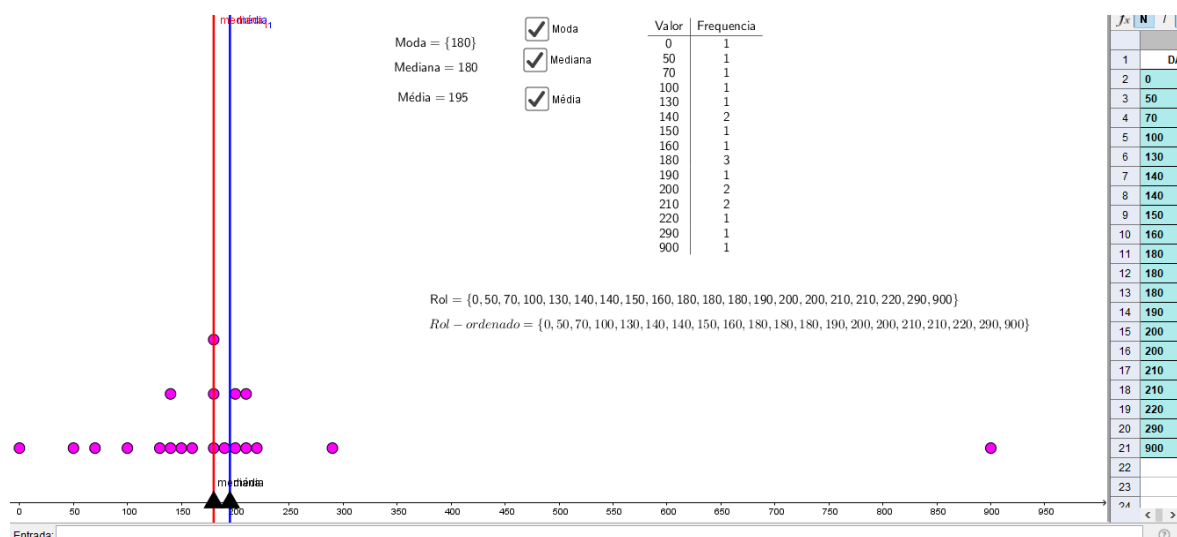
Isso reforça o que afirmamos anteriormente, que a média aritmética é mais sensível à influência de valores atípicos, conhecidos como *outliers*, que são observações excepcionalmente pequenas ou extremamente grandes em relação ao restante dos dados. Para Agresti, Franklin e Klingenberg (2023, p.84), “Um *outlier* é uma observação que fica bem acima ou bem abaixo da massa geral dos dados”. Essa síntese é apresentada no [Gráfico 2](#), com destaque para o alimento que possui 0 mg de sódio, identificado como um potencial *outlier*. Nesse caso, a média aritmética se afasta da mediana na direção da cauda mais longa à esquerda. Além disso, incluímos o exemplo hipotético, em que uma observação foi registrada erroneamente como 900 mg, em vez de 340 mg. No [Gráfico 4](#), observa-se que a

⁴ De acordo com o IBGE, a simetria da distribuição refere-se ao equilíbrio na forma como os dados estão distribuídos em torno das medidas de posição, ou seja, em torno de seus valores centrais.

⁵ Uma distribuição com uma "cauda mais longa" indica que os dados não seguem uma distribuição normal (curva de sino), ou seja, simétrica. Em vez disso, a distribuição possui uma assimetria, onde os dados se estendem mais para um dos lados (direita ou esquerda), indicando a presença de possíveis valores atípicos. (veja no [Gráfico 3](#) está síntese)

média aritmética foi influenciada por esse erro, deslocando-se na direção da cauda mais alongada à direita.

Gráfico 4 - Gráfico de pontos interativos, com a entrada de registro errôneo no valor de 900 mg de sódio, ao invés de 340.



Fonte: Elaborado pelo autor (2024).

Esses autores salientam ainda que a identificação de *outliers* e a investigação de suas causas geralmente requerem uma análise mais aprofundada. Isso pode incluir a revisão de possíveis erros na entrada de dados, a análise de eventos inesperados ou a compreensão de ocorrências incomuns. Nesses casos, é fundamental dedicar atenção especial para garantir a precisão dos dados e sua consistência.

Nesses casos, pode-se pensar que a mediana seria uma melhor escolha em comparação à média aritmética, devido à sua resistência a *outliers*. Destaca-se que a escolha entre a mediana e a média aritmética depende do contexto e do objetivo da análise do conjunto de dados.

2.3 QUARTIS

Os quartis são os valores que dividem um conjunto de dados em quatro partes iguais, ou seja, cada uma delas subdivide o conjunto dos dados em quatro novos subconjuntos, de modo que cada um represente 25% dos dados. Denotaremos os quartis por (q_1) , (q_2) e (q_3) . De acordo com Moore e McCabe (2002), para calcular os quartis, devemos seguir os seguintes passos:

- i) Dispomos as observações em ordem crescente e localizamos a mediana global na lista ordenada de observações.
- ii) O primeiro quartil (q_1) é a mediana das observações situadas à esquerda da mediana global.
- iii) O segundo quartil (q_2) é a própria mediana global.
- iv) O terceiro quartil (q_3) é a mediana das observações situadas à direita da mediana global.

Aplicando nos dados do exemplo 2.2 ([Tabela 2](#)), determinaremos os quartis da quantidade de sódio dos 20 alimentos matinais, apresentados na [Tabela 2](#). Seguindo os passos i), ii), iii) e iv), temos:

Exemplo 2.3:

1. O segundo quartil (q_2) é a própria mediana global encontrada no exemplo 2.2:

$$\tilde{x} = q_2 = \frac{x_{(10)} + x_{(11)}}{2} = \frac{180 + 180}{2} = 180$$

2. O primeiro quartil (q_1) é a mediana das dez observações à esquerda da mediana global. Como essa quantidade é par, (q_1) é a média aritmética das 5ª e 6ª observações (em negrito na [Tabela 2](#)):

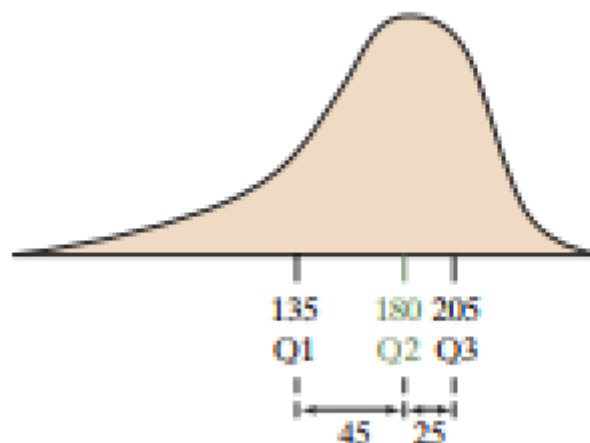
$$q_1 = \frac{x_{(5)} + x_{(6)}}{2} = \frac{130 + 140}{2} = 135$$

3. O terceiro quartil (q_3) é a mediana das dez observações à direita da mediana global, que é a média aritmética das 15ª e 16ª observações (em negrito na [Tabela 2](#)):

$$q_3 = \frac{x_{(15)} + x_{(16)}}{2} = \frac{200 + 210}{2} = 205$$

O [Gráfico 5](#) apresenta uma visualização gráfica desses quartis, considerada a distribuição dos valores da quantidade de sódio dos 20 alimentos matinais.

Gráfico 5 - Representação gráfica dos quartis da quantidade de sódio dos 20 alimentos matinais.



Fonte: Agresti, Franklin e Klingenberg (2023), p.100.

Os quartis informam os intervalos de valores da variável em questão que contém cada uma das quatro partes. Assim, o intervalo entre 0 mg e 135 mg contém os 25% dos menores valores de sódio; o intervalo entre 205 mg e 340 mg contém os 25% dos maiores; as quartas partes restantes estão contidas nos intervalos entre 135 mg e 180 mg e entre 180 mg e 205 mg. Percebe-se assim que o intervalo entre 135 mg e 205 mg contém metade dos dados (porção interna do conjunto de dados). Ao dividir a distribuição em quatro partes, podemos observar que alguns desses intervalos são maiores que os outros e podemos ter uma primeira ideia sobre a forma da distribuição e sobre a dispersão dos dados (veja no [Gráfico 3](#) algumas possíveis formas de distribuição e dispersão dos dados⁶). A diferença de 45 mg entre o primeiro quartil e a mediana é maior do que a diferença de 25 mg entre a mediana e o terceiro quartil. Isso geralmente ocorre quando a distribuição é assimétrica à esquerda. Em outras palavras, a cauda mais longa em relação à mediana refere-se aos menores valores (assimetria à esquerda), como ilustrado no [Gráfico 5](#). Embora cada um desses intervalos possa ser diferente (indicando

⁶ A dispersão dos dados indica o quão espalhados ou concentrados os dados estão em relação às medidas de posição.

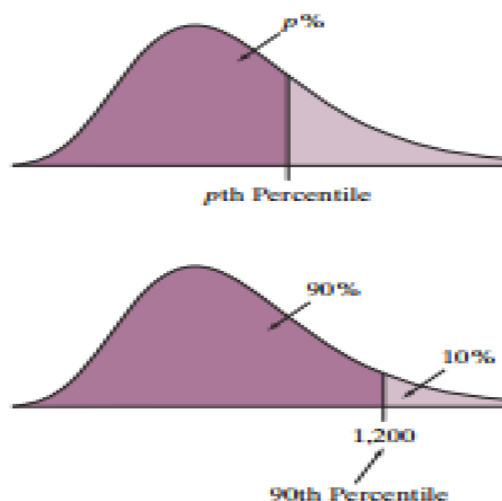
diferentes quantidades de variabilidade na dispersão dentro dos quartis), cada um deles contém a mesma quantidade de observações, ou seja, 25% dos dados.

2.4 PERCENTIS

A mediana e os quartis são casos especiais de um conjunto mais geral de medidas de posição denominado percentis. Para Agresti, Franklin e Klingenberg (2023, p.99), “O p -ésimo percentil é um valor tal que p por cento das observações caia nesse valor ou abaixo dele”. Desta forma, podemos ter outros percentis além da mediana e dos quartis. Para estes autores, as tecnologias digitais podem ser ferramentas aliadas para fazer cálculos de percentis, especialmente para os conjuntos de dados grandes. A seguir, apresentamos um exemplo de Agresti, Franklin e Klingenberg (2023) para explicar o conceito de percentis. Suponha que um aluno seja informado de que sua pontuação foi de 1.200 (de 1.600) em um vestibular. Essa pontuação cai no 90° percentil ($p = 90$). Então, 90% dos que fizeram o exame obtiveram pontuação de 1.200 ou menos. Apenas 10% das pontuações foram superiores às deles. Essa síntese é exemplificada na representação do [Gráfico 6](#) a seguir:

Exemplo 2.4:

Gráfico 6 - Representação gráfica de percentis da pontuação de um aluno em um vestibular.



Fonte: Agresti, Franklin e Klingenberg (2023), p.99.

2.5 AMPLITUDE E A DISTÂNCIA INTERQUARTIL

Para Agresti, Franklin e Klingenberg (2023), a distância interquartil é a diferença entre o terceiro e o primeiro quartis, que é:

$$IQ = q_3 - q_1. \quad (3)$$

Já a amplitude é a diferença entre o maior e o menor valor de um conjunto de dados, que é:

$$R = x_{(n)} - x_{(1)}. \quad (4)$$

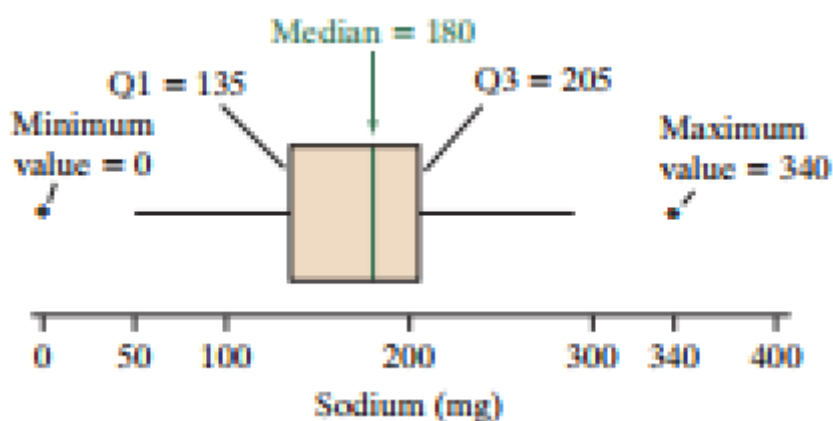
Vale ressaltar que o maior e o menor valor de um conjunto de dados podem influenciar na média aritmética e mediana ([Gráfico 2](#), [Gráfico 3](#) e [Gráfico 4](#)).

Como exemplo, utilizaremos os dados da [Tabela 2](#) para determinar a amplitude (R) e a distância interquartil (IQ). Do exemplo 2.2 e 2.3, sabe-se que os valores do máximo, do mínimo, do primeiro e do terceiro quartil desse conjunto de dados são, respectivamente: $x_{(n)} = 340$, $x_{(1)} = 0$, $q_1 = 135$, $q_3 = 205$.

Assim, aplicando a eq. (4), a amplitude é $R = 340 - 0 = 340$ e aplicando a eq. (3), a distância interquartil é $IQ = 205 - 135 = 70$, o que indica que os 50% dos valores de sódio situados entre o primeiro e o terceiro quartil se estendem por uma faixa de 70 mg. Assim como no caso da amplitude, quanto maior a distância interquartil, maior variabilidade nas observações. No entanto, ao contrário da amplitude, a distância interquartil não é afetada por quaisquer observações abaixo do primeiro quartil ou acima do terceiro quartil; em outras palavras, não é afetada por eventuais valores extremos do conjunto de dados. A amplitude depende exclusivamente dos valores mínimo e máximo, ou seja, dos extremos do conjunto de dados, e varia à medida que qualquer valor extremo se altera. Por exemplo, se o maior valor de sódio fosse 900 mg, em vez de 340 mg, a amplitude mudaria drasticamente, mas a distância interquartil permaneceria inalterada. Por isso, muitas vezes é mais indicado usar o intervalo interquartil em vez da amplitude, para comparar a variabilidade em distribuições altamente distorcidas ou que apresentam discrepâncias significativas.

A distância interquartil pode ser usada para estabelecer regras a fim de identificar valores suspeitos de serem atípicos (*outliers*). Nesse sentido, Agresti, Franklin e Klingenberg (2023, p.101) propõem a seguinte regra: “Uma observação é considerada um potencial *outlier* se está a uma distância de $(1,5)IQ$ abaixo do primeiro quartil ou a uma distância de $(1,5)IQ$ acima do terceiro quartil”. Segundo esses autores, uma observação é identificada como um potencial *outlier*, em vez de ser definida como um *outlier* definitivo, mesmo quando uma distribuição é significativamente assimétrica, pois algumas observações podem estar a mais de $(1,5)IQ$ abaixo do primeiro quartil ou acima do terceiro quartil, sem serem verdadeiramente discrepantes. Isso ocorre porque essas observações podem não estar suficientemente distantes da maior parte dos dados. Por exemplo, em uma distribuição com cauda longa, a maior observação pode não estar tão distante das observações restantes, e não há uma lacuna suficiente para chamá-lo de *outlier*. Essa síntese sobre potenciais *outliers* é exemplificada no [Gráfico 7](#).

Gráfico 7 - Representação gráfica dos possíveis *outliers* da quantidade de sódio dos 20 alimentos matinais.



Fonte: Agresti, Franklin e Klingenberg (2023), p.103.

Desta forma, aplicando a regra descrita acima nos dados ordenados do exemplo 2.2 ([Tabela 2](#) e [Gráfico 7](#)), podem-se identificar potenciais *outliers*. Como já sabemos, $q_1 = 135$ mg, $q_3 = 205$ mg e $IQ = 70$ mg. Assim, os limites inferior e superior são, respectivamente, $LI = 135 - 1,5 \times 70 = 30$ mg e $LS = 205 + 1,5 \times 70 = 310$ mg. Com base nesses limites, a quantidade de sódio de 0 mg para o alimento A está abaixo de 30 mg, e a quantidade de 340 mg para o alimento B está acima de 310 mg. Portanto, ambas as quantidades de alimentos A e B são potenciais *outliers*.

e devem ser estudados no contexto do problema para serem qualificados como valores atípicos ou não.

2.6 VARIÂNCIA E DESVIO PADRÃO

Moore e McCabe (2002) afirmam que os desvios $x_i - \bar{x}$ mostram a dispersão dos valores x_i em torno de sua média \bar{x} , havendo desvios positivos e negativos. A soma desses desvios das observações é sempre zero. Uma maneira de estabelecer uma medida resumo considerando esses desvios é elevá-los ao quadrado, tornando-os todos positivos. Dessa forma, as observações muito distantes da média (de um ou de outro lado) têm grandes desvios quadráticos positivos, e as observações próximas da média, desvios quadráticos mais próximos de zero. A variância poderia ser considerada como a média aritmética desses desvios quadráticos, significando uma quantidade média de dispersão de cada observação. Por razões técnicas, calcula-se uma quantidade média de desvios quadráticos considerando o valor de $n - 1$ no denominador. Dessa maneira, considerando as observações x_1, x_2, \dots, x_n , temos que a variância é:

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1} = \frac{1}{n - 1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (5)$$

Salienta-se que a média dos desvios quadráticos é obtida dividindo o total por $n - 1$, pois apenas $n - 1$ quadrados de desvios podem variar livremente, já que a soma dos desvios é sempre zero, ou seja, pode-se obter o último desvio, desde que conheçamos os outros $n - 1$. De acordo com Moore e McCabe (2002), a quantidade $n - 1$ é denominada de graus de liberdade da variância ou do desvio padrão.

Percebe-se que a unidade da variância é o quadrado da unidade original da variável. Uma saída para obter uma medida resumo de dispersão na mesma unidade da variável, é calcular a raiz quadrada da variância. Essa medida de dispersão é denominada desvio padrão (s). Essa transformação mantém a mesma ordem de dispersão dos dados, estabelecida pela variância. Em uma notação matemática, temos:

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1}} = \sqrt{\frac{1}{n - 1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (6)$$

Os conceitos de variância e desvio padrão serão exemplificados por meio de conjunto de dados apresentado em Agresti, Franklin e Klingenberg (2023). Imagine o seguinte cenário: Os alunos de uma turma responderam a uma pergunta feita pelo Professor de Matemática no início da aula: “Quantos filhos você acha que são ideais para uma família?” As respostas foram registradas e classificadas de acordo com o gênero (masculino ou feminino) dos estudantes. A [Tabela 3](#) apresenta qual seria o tamanho ideal de família indicado por cada estudante, estratificado por gênero.

Exemplo 2.5:

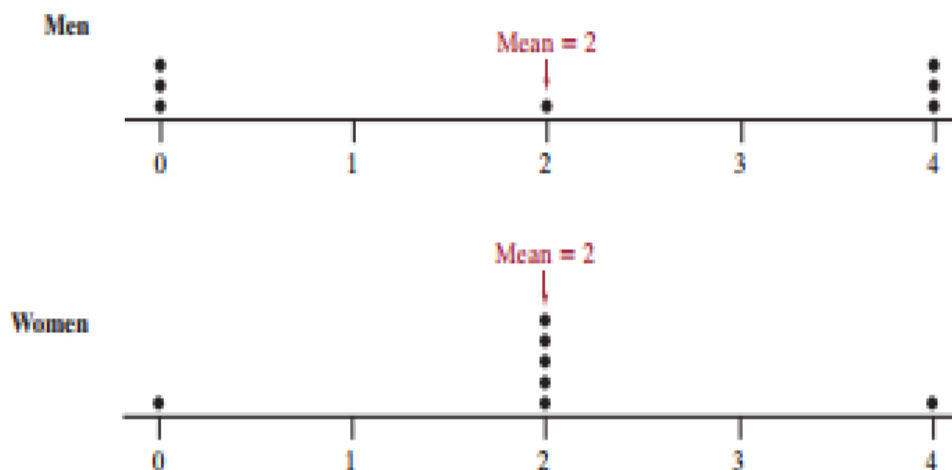
Tabela 3 - Quantidade ideal de filhos para uma família classificada por gênero de cada estudante.

Gênero do estudante	Quantidade ideal de filhos							
Masculino	0	0	0	2	4	4	4	4
Ferminino	0	2	2	2	2	2	2	4

Fonte: Síntese do autor a partir do exemplo de Agresti, Franklin e Klingenberg (2023), p.92.

Baseando-se na [Tabela 3](#), a quantidade média ideal de filhos em uma família é dois em ambos os gêneros. A amplitude das respostas é quatro, que é também a mesma em ambos os estratos. Assim, percebe-se que tanto o gênero masculino quanto o feminino apresentam uma média dois e uma amplitude quatro. O [Gráfico 8](#) apresenta uma síntese visual dos dados por meio de um gráfico de pontos. Percebe-se, entretanto, que, em cada grupo, a dispersão das respostas é bastante distinta.

Gráfico 8 - Gráfico de pontos da quantidade ideal de filhos para uma família classificada por gênero de cada estudante.



Fonte: Agresti, Franklin e Klingenberg (2023), p.92.

O desvio padrão pode perceber essa diferença de dispersão entre os dois grupos. Aplicando a eq. (5), a variância dos dados referentes ao gênero masculino é determinada pela expressão abaixo:

$$s_m^2 = \frac{(0 - 2)^2 + (0 - 2)^2 + \dots + (4 - 2)^2}{7 - 1} = \frac{24}{7 - 1} = 4$$

Aplicando a eq. (6), seu desvio padrão é:

$$s_m = \sqrt{4} = 2$$

Similarmente, aplicando a eq. (5), a variância para o gênero feminino é:

$$s_f^2 = \frac{(0 - 2)^2 + (2 - 2)^2 + \dots + (4 - 2)^2}{7 - 1} = \frac{8}{7 - 1} \approx 1,333\dots$$

E aplicando a eq. (6), seu desvio padrão é:

$$s_f = \sqrt{1,333...} \approx 1,2$$

Verifica-se que, para o gênero masculino, o desvio padrão é maior que para o gênero feminino. Isso mostra que as observações para o gênero masculino tendem a estar mais distantes da média do que as do gênero feminino. Em outras palavras, elas têm uma dispersão maior em torno da média, indicando uma variação maior no tamanho ideal de família. Entretanto, no caso das estudantes do gênero feminino, percebe-se que elas foram mais consistentes em seus pontos de vista sobre o número ideal de filhos para uma família, pois apresentou um valor menor.

2.7 MODA

A moda m_o é definida como “a realização mais frequente do conjunto de valores observados” (Bussab e Morettin, 2017, p. 43). Para estes autores, quando houver duas modas, dizemos que a distribuição é bimodal. Da mesma forma, quando houver três modas, dizemos que a distribuição é trimodal e, quando houver mais de três modas, dizemos que a distribuição é multimodal. O termo amodal é quando não temos uma realização mais frequente do conjunto de dados observados.

Utilizaremos o exemplo 2.2 da [Tabela 2](#) apresentada anteriormente, e determinamos a moda da quantidade de sódio dos 20 alimentos matinais. Podemos verificar que a moda é $m_o = 180$, que é o valor mais frequente do conjunto de dados. A representação gráfica desse conceito pode ser visualizada no [Gráfico 1](#).

2.8 COEFICIENTE DE VARIAÇÃO

O coeficiente de variação (cv) é definido como a razão entre o desvio padrão, s , e a média amostral, \bar{x} , e usualmente expresso em porcentagem para comparar a variabilidade de um conjunto de dados. Em uma notação matemática, temos:

$$cv = \frac{s}{\bar{x}} \cdot 100\% \quad (7)$$

Portanto, utilizando os dados do exemplo 2.5 de variância e desvio padrão e aplicando a eq. (7), o coeficiente de variação para o gênero masculino é:

$$cv = \frac{s_m}{\bar{x}} \cdot 100\% = \frac{2}{2} \cdot 100\% = 100\%$$

Analogamente, aplicando a eq. (7) para o gênero feminino, temos:

$$cv = \frac{s_f}{\bar{x}} \cdot 100\% = \frac{1,333...}{2} \cdot 100\% \approx 66,66\%$$

Esses resultados evidenciam que o grau de variabilidade nas opiniões é menor para o gênero feminino, indicando opiniões mais consistentes em comparação ao gênero masculino.

3 TECNOLOGIAS DIGITAIS NO ENSINO DE ESTATÍSTICA

A crescente presença de dados no cotidiano está fortemente correlacionada com os avanços globais das tecnologias. Conforme Watson (2002), as pessoas estão constantemente expostas a dados por meio dos veículos de comunicação, redes sociais, etc., sendo primordial o desenvolvimento do Letramento, o Raciocínio e o Pensamento Estatísticos, tornando-os críticos e alfabetizados estatisticamente. Vale também ressaltar que Garfield e Ben-Zvi (2008) enfatizam que os dados estão sendo cada vez mais usados para adicionar ou sugerir credibilidade, ou seja, os autores sugerem que as escolas preparem os cidadãos para serem capazes de analisar, interpretar e avaliar criticamente afirmações baseadas em dados. Além disso, eles salientam que a disponibilidade de ferramentas tecnológicas de alta qualidade para a Aprendizagem de Estatística permite alavancar as metodologias da Aprendizagem do Raciocínio Estatístico, permitindo que o campo da Educação Estatística integre e compreenda mais facilmente os progressos relacionados ao Ensino de Estatística (o Letramento, o Raciocínio e o Pensamento Estatísticos). Para Biehler, *et al.* (2013) o desenvolvimento da Estatística sempre esteve interligado com o desenvolvimento da tecnologia. Tukey (1965), Yates (1971) e Chambers (1980) foram pioneiros em prever a natureza das mudanças futuras na elaboração de resultados estatísticos, que surgiriam como consequência dos avanços tecnológicos. Além disso, eles imaginaram a natureza das novas ferramentas que deveriam ser desenvolvidas e necessárias para apoiar a Estatística Aplicada.

Gould, *et al.* (2018) refletiram sobre os desenvolvimentos recentes e futuros no currículo de Estatística e nas formas de Ensino e Aprendizagem da Estatística, destacando-se as tecnologias digitais como um elemento com presença no Ensino da Estatística, sendo também essencial na formação de cidadãos críticos e alfabetizados estatisticamente. A Base Curricular do Chile (*Ministerio de Educación Chile, 2016*) e os autores, Hsu e Chen (2020), enfatizam que vários estudos sugerem que a integração eficaz de tecnologias digitais inovadoras pode melhorar a experiência e o envolvimento dos alunos no processo de Aprendizagem. As Tecnologias da Informação e Comunicação (TIC) constituem-se em um conjunto de recursos tecnológicos que podem ser utilizados no Ensino para potencializar o processo de Aprendizagem. Para Quinõnes, *et al.* (2024), as TIC podem melhorar

significativamente a compreensão de conteúdos e conceitos estatísticos de forma interativa ou dinâmica. No entanto, esses autores realçam que muitas das plataformas disponíveis atualmente, as quais oferecem novas metodologias de comunicação (TIC), podem não ser adequadas no contexto da Educação Básica. Eles apontam, ainda, que existem barreiras linguísticas, ou que a forma de apresentação do conteúdo é confusa, muitas vezes com explicações limitadas ou com a falta de contexto nos problemas abordados, priorizando equivocadamente os cálculos em vez dos conceitos estatísticos. Quinões, *et al.* (2024) salientam ainda que é fundamental enfrentar os desafios de desenvolver uma plataforma centrada no aluno e que ofereça um suporte técnico adequado, garantindo uma experiência de aprendizagem mais eficaz e eficiente. Por outro lado, Samá e Silva (2024) enfatizam que o uso da tecnologia digital pode contribuir no processo do aprendizado estatístico, portanto os professores precisam refletir sobre o que, como e por que ensinar determinados conceitos. Os autores concluem que uma abordagem pedagógica mais abrangente, permeada pela tecnologia digital, é um caminho importante na atualização dos métodos do Ensino de Estatística. Eles pontuam também que a formação de professores, focada em uma pedagogia que integre as tecnologias digitais no Ensino, possibilita abordagens inovadoras, mais alinhadas às demandas atuais dos estudantes, proporcionando a aprendizagem dos conceitos estatísticos de uma maneira mais consistente.

Ainda em relação ao processo de Ensino e Aprendizagem de Estatística, vale ressaltar a Análise Exploratória de Dados (AED), proposta por Tukey (1977), que é uma abordagem Estatística poderosa utilizada para analisar e investigar conjuntos de dados. Seu objetivo é resumir as principais características dos dados, baseando-se fortemente na utilização de métodos visuais. Biehler, *et al.* (2013) destacam que o uso de AED no Ensino de Estatística foi largamente facilitado, devido à simplicidade em trabalhar com dados reais e em construir visualizações aprimoradas de dados, aproveitando-se de recursos computacionais cada vez mais eficazes. Isso contribuiu extensamente para o desenvolvimento do Raciocínio Estatístico, de acordo com o que foi apontado por Garfield e Ben-Zvi (2008). Além disso, Biehler, *et al.* (2013) também enfatizam que a AED foi um passo inicial em direção a um relato mais realista da aplicação da Estatística, com uma abordagem de “aprender com os dados” no decorrer do processo de análise estatística. Esses autores destacam que as novas ferramentas tecnológicas apoiam a verificação

cuidadosa de pressupostos mais complexos de procedimentos tradicionais, podendo assim desenvolver e implementar métodos mais robustos. Eles apontam também que as TIC permitem ao usuário localizar mais facilmente padrões na associação entre as variáveis, na seleção e na validação de modelos funcionais adequados, realçando ainda que esse processo representa um aspecto do Raciocínio Estatístico que é muito mais desafiador do que simplesmente aplicar fórmulas. No contexto do Ensino de Estatística, Wild e Pfannkuch (1999) e Wild, Utts e Horton (2018) [[subseção 2.2.3](#) e [subseção 2.1.1](#)], dentre outros autores, afirmam que “aprender com os dados” é essencial para o desenvolvimento do Pensamento Estatístico. A partir das discussões sobre o uso de tecnologia no Ensino de Estatística, Olive, *et al.* (2010) apontam para duas perspectivas de utilização das TIC. A primeira sugere o desenvolvimento de um processo de Ensino e Aprendizagem de Estatística baseado no apoio contínuo dessas tecnologias, refletindo tanto a teoria quanto a prática dos *softwares* estatísticos. A segunda perspectiva propõe o uso da tecnologia como uma ferramenta eventual para aprimoramento do aprendizado de Estatística. O foco está em explorar também outras possibilidades que a tecnologia oferece, como por exemplo, a visualização interativa ou dinâmica dos resultados estatísticos. Essa abordagem valoriza a compreensão dos conceitos, em vez de se concentrar apenas em cálculos axiomáticos, além de proporcionar aos alunos a oportunidade de explorar e experimentar os dados, tornando o aprendizado mais envolvente e significativo. Por outro lado, Batanero e Borovcnik (2016) ressaltam que os professores podem usar essas novas tecnologias para o Ensino de Estatística, mas recomendam que não se enfatize excessivamente o uso desta ferramenta, em detrimento da compreensão dos conceitos estatísticos. Caso contrário, as metas de aprendizagem podem ser limitadas à mera manipulação de técnicas, sem a devida atenção aos conceitos estatísticos ou à construção do Raciocínio Estatístico subjacente ao método aplicado.

Constata-se também que a tecnologia expandiu profundamente as técnicas gráficas e de visualização, proporcionando aos estudantes novas formas de explorar e analisar os dados, além de incentivar a reflexão crítica sobre ideias estatísticas. Isso permite que os alunos foquem mais na interpretação dos resultados e na compreensão dos conceitos, em vez de se concentrarem apenas no cálculo de fórmulas e nos procedimentos computacionais.

Para Wild, Utts e Horton (2018), os métodos estatísticos apoiados por recursos computacionais são amplamente aplicados em diversas áreas do conhecimento e estão cada vez mais presentes nas práticas de empresas, governos, profissionais de saúde, entre outros, auxiliando nas tomadas de decisões mais assertivas. Além disso, os autores destacam a importância de que todas as pessoas compreendam ideias estatísticas e as suas conclusões. Por exemplo, a ampla disponibilidade de grandes conjuntos de dados públicos e o acesso facilitado a *softwares* de visualização e análise permitem que os dados sejam explorados e compreendidos, de maneira que sejam formuladas e respondidas todas as perguntas importantes no contexto do problema em análise e que permeiam o nosso cotidiano.

Segundo Batanero (2000),

tradicionalmente, no aprendizado da estatística, tem sido dada grande importância ao cálculo e aos aspectos atuativos, que hoje perdem importância devido às novas tecnologias. Em vez de praticar cálculos e gráficos com lápis e papel, o aluno deve aprender a usar calculadoras gráficas e programas de computador, como planilhas. As novas tecnologias também introduzem novos elementos atuativos e ostensivos, uma vez que a gama de representações disponíveis é muito maior (Batanero, 2000, p.10).

Além disso, Chance, Delmas e Garfield (2004) verificaram que o uso de simulações computacionais em sala de aula atua significativamente no desenvolvimento do Raciocínio Estatístico em tópicos fundamentais, tais como probabilidade, variabilidade, amostragem, distribuições amostrais, etc. Os autores apontam que essas atividades devem ser cuidadosamente planejadas para garantir o envolvimento dos alunos em todo o processo, garantindo o desenvolvimento efetivo do Raciocínio Estatístico. Sharma (2013) aponta também que os avanços no uso de novas tecnologias no Ensino de Estatística ampliam a quantidade de dados disponíveis, coletados por meio da web e das mídias. Ele destaca a importância de iniciar o Ensino com contextos familiares aos alunos, além de utilizar temas de interesse dos estudantes, com o intuito de estimular e consolidar o Letramento Estatístico.

De acordo com Burril e Pfannkuch (2023), a evolução contínua da tecnologia amplia o oferecimento de novos contextos, para que os alunos se envolvam em estudos estatísticos e desenvolvam a compreensão dos fundamentos e conceitos

subjacentes ao problema. Para esses autores, os dados do mundo real são multivariados e podem consistir em uma variedade de formatos como: imagens, texto, sons, datas, símbolos locais e codificados, categorias, medidas, etc. Eles realçam ainda que vários pesquisadores reconhecem o potencial da tecnologia para criar novas formas de visualização dos dados. Seguindo esta perspectiva, Andre e Lavicza (2019) identificaram possibilidades metodológicas com o uso das tecnologias, oferecendo ainda exemplos de como conduzir a integração da visualização de dados na Educação Básica. Engel, *et al.* (2020) também sugerem como as tecnologias podem fornecer ferramentas para visualização de dados que não exijam um conhecimento estatístico profundo para as pessoas explorarem os dados. Esses autores ainda mostram como as visualizações interativas ou dinâmicas de dados podem ser usadas para proporcionar uma compreensão conceitual muito além do que apenas a das representações gráficas usuais. Eles descrevem esse tipo de visualização de dados como uma combinação entre a ciência e a arte, abrindo novos caminhos para a comunicação, por meio de uma abordagem diferente de análise de dados. Sob esta perspectiva, Rubel, Nicol e Chronaki (2021) analisaram o uso de tecnologias nas representações de dados, de maneira a permitir uma popularização da leitura e da interpretação das análises estatísticas. Esses autores usam exemplos relacionados à pandemia (Covid-19), ressaltando dimensões, o que é importante para uma leitura crítica dos dados, que descrevem como:

visualização de dados inter-relacionada com a formulação, o enquadramento e a narração inter-relacionados pode ter ressonância em interpretações privilegiadas, na sustentação de certas hierarquias de poder, na perpetuação de determinados valores em detrimento de outros e na orientação para decisões estabelecidas (Rubel, Nicol e Chronaki, 2021, p. 265).

Burril (2000) realçava que as TIC estavam transformando a forma como se vivia e como os negócios eram efetivados. Alertavam também que ocorreriam mais mudanças nos anos seguintes. No contexto educacional, o autor apontava que as tecnologias impactariam a Educação Estatística, pois poderia ser mudada a natureza do Ensino de Estatística, influenciando significativamente o desenvolvimento do Raciocínio Estatístico dos alunos.

Desta forma, muitas das mudanças na conceituação e na interpretação de resultados estatísticos são hoje possíveis graças às inovações das novas TIC,

alterando tanto as metodologias estatísticas quanto os procedimentos para o Ensino de Estatística. Nesse sentido, os *softwares* estatísticos possuem uma vasta história de utilização no Ensino de Estatística. Wild, Utts e Horton (2018) destacam, dentre outros, os seguintes *softwares* gratuitos para o Ensino de Estatística: *Visual Inference Tools (VIT)*; iNZight; GeoGebra e R. Dentre os *softwares* comerciais, os autores citam: o Tableau; o Minitab; o Fathom e o TinkerPlots. Além disso, eles destacam algumas plataformas gratuitas da web, tais como: *Common Online Data Analysis Platform (CODAP)*; iNZight Lite e TuvaLabs.

Com essa perspectiva, Samá e Silva (2024) salientam a importância de uma abordagem mais abrangente do processo de Ensino e Aprendizagem em Estatística, a qual contemple os aspectos pedagógicos e didáticos, além dos aspectos tecnológicos. Isso implica explorar as estratégias de análise que envolvam o uso de tecnologias digitais e identificar as melhores práticas para o processo e para os desafios enfrentados pelos professores. Além disso, devem-se buscar soluções, em todos os níveis de ensino, que permitam a integração eficaz da tecnologia com o currículo de Estatística. Para os autores, esse procedimento integrado potencializa o uso das TIC no Ensino de Estatística, proporcionando oportunidades para uma abordagem participativa, envolvente e contextualizada, preparando nossos estudantes para um mundo em que a tecnologia seja um auxílio ao pensamento crítico, à resolução de problemas e à tomada de decisões. Destaca-se, assim, a necessidade de a Educação Estatística direcionar seus esforços para o uso de recursos tecnológicos digitais no processo de Ensino e Aprendizagem de Estatística.

4 APRESENTAÇÃO DAS TAREFAS

Neste capítulo, apresentamos o conjunto de tarefas referentes a medidas de posição. Esse tema faz parte do conteúdo de medidas resumo, tendo sido especificamente discutido e trabalhado no ambiente escolar os conceitos e o uso da média aritmética e mediana, estendendo-se também para os conceitos de quartis.

Nossa intenção é contribuir para o Ensino de Estatística de modo que ele seja interessante e estimulante para o aprendizado, a fim de que, assim, esse material possa ser utilizado como aporte às aulas de Matemática que envolvem o conteúdo das medidas de posição.

4.1 TAREFA 1 - O QUE SÃO DADOS EM ESTATÍSTICA E DADOS REPRESENTADOS POR MEIO DE GRÁFICOS BOXPLOTS.

A primeira tarefa é considerada deflagradora de todo o processo, ao introduzir a temática do que são dados em Estatística e sua representação por meio de gráficos boxplots. A proposta consiste em promover uma discussão inicial sobre esses dois conceitos, de modo a possibilitar a continuidade do processo de Ensino e Aprendizagem na tarefa seguinte, com foco no conteúdo das medidas de posição. Nesse sentido, a tarefa tem como objetivo aproximar os alunos do tema em estudo e estabelecer uma base conceitual para o desenvolvimento das demais tarefas.

Com o intuito de estimular a reflexão dos alunos sobre o conceito de dados em Estatística e de sua representação por meio de boxplots, é disponibilizado um texto para reflexão e discussão dos alunos.

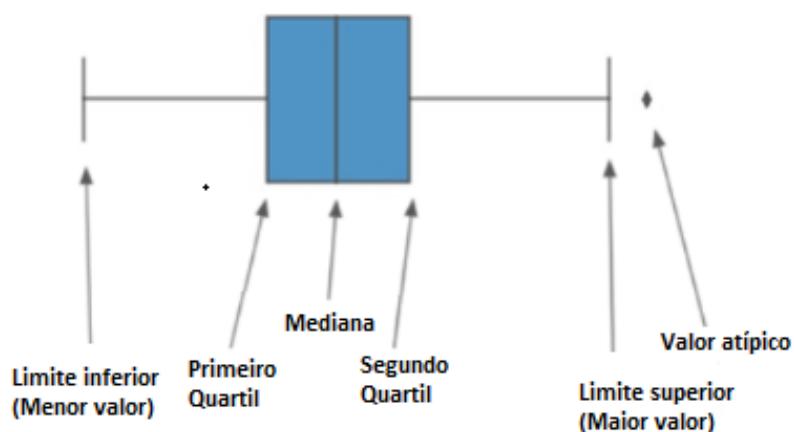
Texto para reflexão e discussão

Na Estatística, os dados representam informações sobre situações do nosso cotidiano, como notas de provas, tempo gasto em atividades, salários ou renda média salarial, consumo de energia, medição de alturas. Para compreender melhor essas informações, não basta apenas olhar os números separadamente; é importante organizá-los e representá-los de maneiras que facilitem a interpretação.

O boxplot é um tipo de gráfico que ajuda a resumir um conjunto de dados e a visualizar sua distribuição. Por meio dele, é possível identificar o menor valor, o

maior valor, a mediana e os quartis, que dividem os dados em quatro partes iguais. Essa representação permite observar como os dados estão distribuídos, se estão mais concentrados em determinada região e se há valores que se afastam dos demais que são chamados de valores atípicos. A seguir, veja no Gráfico 9 a estrutura de um gráfico boxplot:

Gráfico 9 - Estrutura de um gráfico boxplot



Fonte: Elaborado pelo autor (2025).

Além disso, o boxplot facilita a comparação entre diferentes conjuntos de dados. Ao observar dois ou mais boxplots, podemos analisar qual conjunto apresenta maior dispersão, onde os dados estão mais concentrados e se as distribuições são semelhantes ou diferentes. Dessa forma, o uso do boxplot contribui para uma compreensão mais clara dos dados e para uma análise mais crítica das informações.

Logo após o texto apresentado para reflexão e discussão, são propostas perguntas a respeito do texto.

Pergunta 1: Segundo o texto, o que são dados em Estatística? Cite dois exemplos mencionados.

Pergunta 2: Quais informações (conceitos) podem ser identificadas em um boxplot?

Pergunta 3: O que o texto chama de valores atípicos?

Pergunta 4: Qual é a principal função do boxplot na análise de dados?

- Respostas esperadas:

Resposta para a Pergunta 1:

É esperado que os alunos compreendam o que são dados em Estatística citando dois exemplos de acordo com o texto. Entre os exemplos mencionados, estão as notas de provas, o tempo gasto em atividades, salários ou renda média salarial, consumo de energia e medições de altura.

Resposta para a Pergunta 2:

É esperado que os alunos compreendam que em um boxplot, é possível identificar o valor mínimo, o valor máximo, a mediana e os quartis, que dividem os dados em quatro partes iguais, além de observar a distribuição e a concentração dos dados.

Resposta para a Pergunta 3:

É esperado que os alunos compreendam que o texto chama de valores atípicos aqueles valores que se afastam dos demais dados do conjunto analisado.

Resposta para a Pergunta 4:

É esperado que os alunos compreendam que a principal função do boxplot na análise de dados é resumir um conjunto de dados e permitir a visualização de sua distribuição, facilitando a interpretação das informações estatísticas.

Após a discussão das respostas sobre as quatro primeiras perguntas, é efetuada, a seguir, a última pergunta dessa tarefa:

Pergunta 5:

Pedir que os alunos coletem dados, em sala de aula, os quais apresentem alguma variação. Algumas perguntas sugeridas:

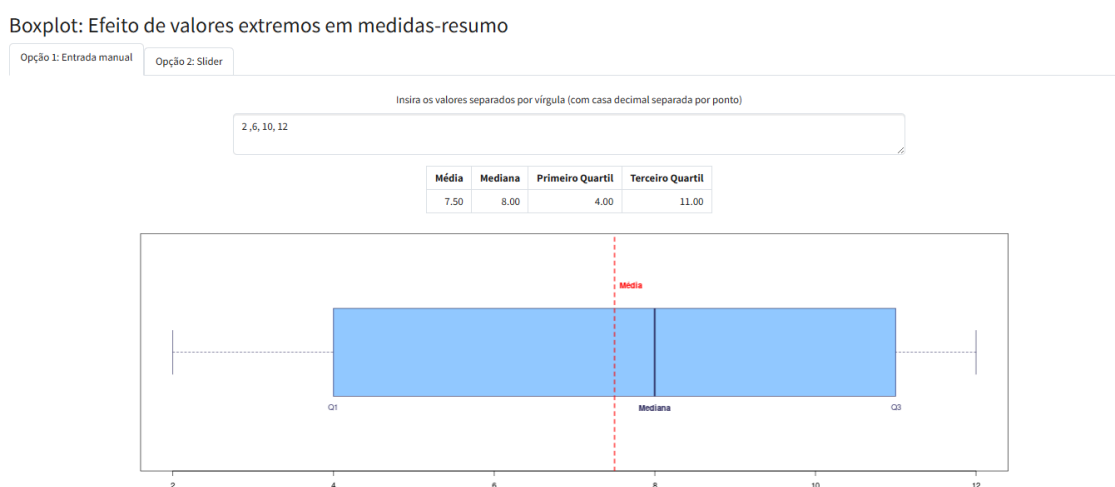
- "Qual foi o seu tempo de uso de tela (*screen time*) no celular ontem, em horas e minutos?"
- "Qual sua altura, em cm?"

- “Qual o tempo, em minutos, de sua casa até a escola?”
- “Quanto tempo, em minutos, você gasta se arrumando de manhã, desde a hora que acorda até sair de casa?”
- "Qual é o tamanho da palma de sua mão (da ponta do dedo à ponta do mindinho), em centímetros?”

Nesse contexto, será então utilizado um aplicativo web para a construção rápida do gráfico boxplot dos dados coletados. Esse aplicativo pode ser acessado de maneira livre e gratuita na url: <https://milenapaz.shinyapps.io/Outliers/>. A seguir, apresenta-se uma breve explicação sobre como acessar o aplicativo e navegar por sua interface ([Figura 1](#) e [Figura 2](#)).

A Figura 1 apresenta a tela inicial do aplicativo web e a opção 1: Entrada manual, com o título Boxplot: Efeito de valores extremos em medidas-resumo.

Figura 1 - Tela inicial do aplicativo web, destacando a opção 1 de inserção manual de dados

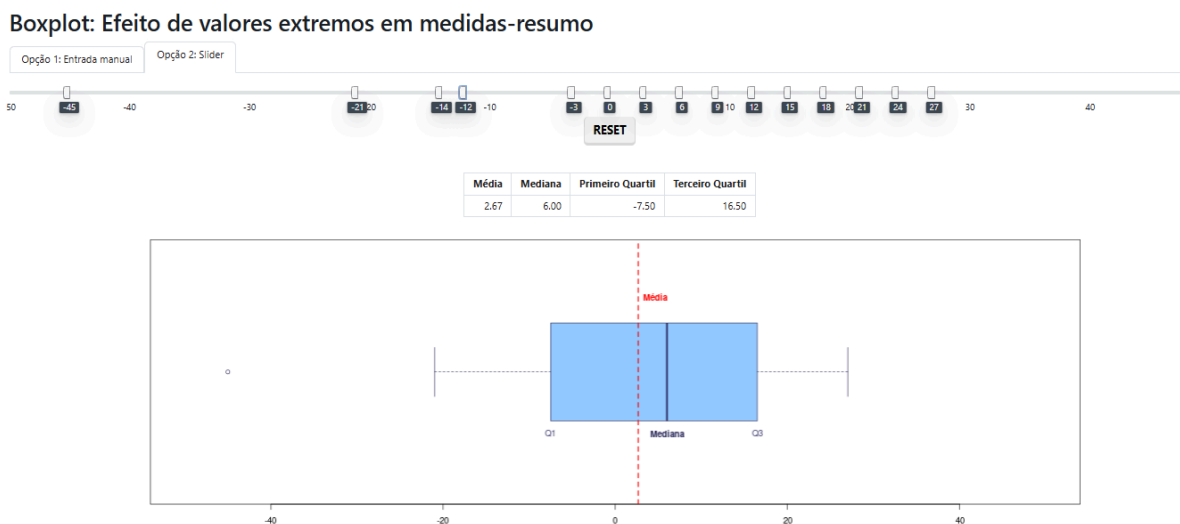


Fonte: Aplicativo desenvolvido pelos autores a partir da pesquisa de campo.

A Figura 1 apresenta as duas opções de entrada diferentes: digitação ou por controle deslizante. O aluno pode inserir os dados pela opção 1, de forma manual, ou pela opção 2, utilizando o controle deslizante. Ao optar pela opção 1, basta inserir os valores, não necessariamente ordenados, reforçando que o separador decimal deve ser o ponto, e os diferentes valores devem ser introduzidos, separando-os por vírgula.

A Figura 2 apresenta a opção 2: Slider, que permite a inserção dos dados por meio do controle deslizante.

Figura 2 - Opção 2 de inserção dos dados por meio do controle deslizante



Fonte: Aplicativo desenvolvido pelos autores a partir da pesquisa de campo.

Nas duas opções do uso do aplicativo web, apresentamos o gráfico com os quartis e a mediana rotulados, facilitando a leitura para aqueles que ainda não estão familiarizados. Além disso, há uma linha em vermelho indicando a média, permitindo a comparação com os demais elementos do gráfico boxplot.

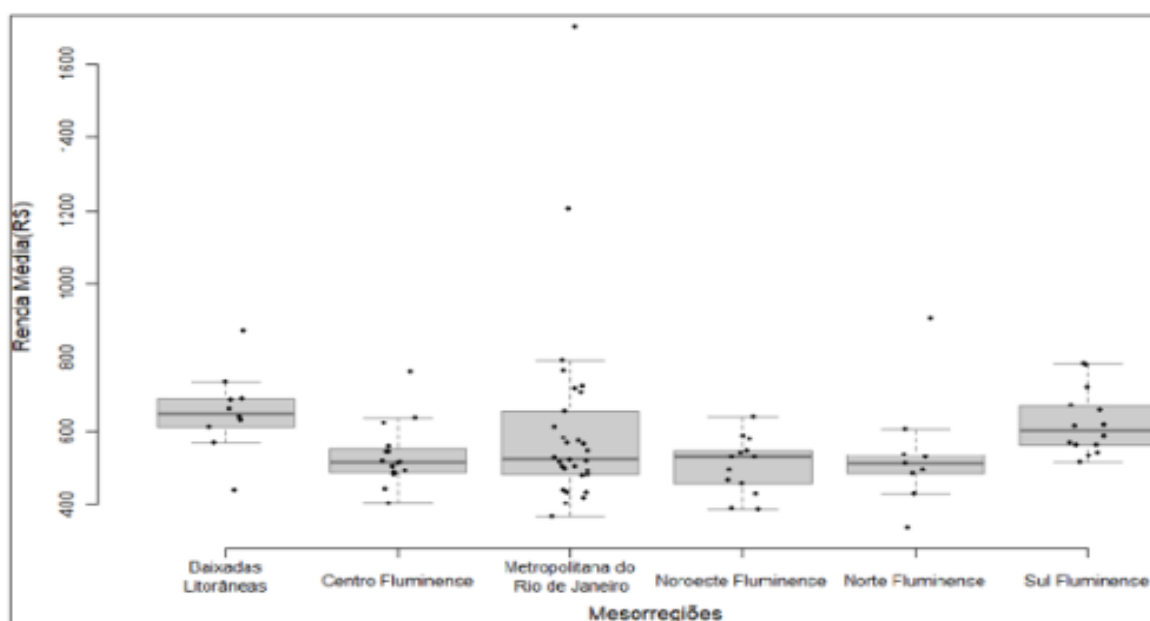
Esse aplicativo web foi desenvolvido a partir da análise das respostas e das falas dos estudantes, após a pesquisa de campo, no âmbito da Dissertação intitulada: “A produção de significados de estudantes do Ensino Médio para as medidas de posição por meio de boxplots e tabelas.”

4.2 TAREFA 2 - PROBLEMÁTICA ENVOLVENDO DADOS REPRESENTADOS POR MEIO DE GRÁFICOS BOXPLOTS.

Nesta tarefa os alunos são convidados a analisar dados representados por boxplots, com o intuito de estimular a leitura crítica de representações gráficas estatísticas, articulando visualização gráfica de medidas de posição, para a construção de significados sobre a variabilidade dos dados.

É apresentada uma situação-problema sobre a qual são discutidas situações utilizando a renda média per capita domiciliar que, conforme o IBGE, é a soma dos rendimentos de todas as fontes dos moradores de um domicílio dividida pelo número de seus residentes, calculada mensalmente. Ela indica o nível socioeconômico e costuma apresentar uma forte assimetria. Os dados tabulados na Tabela 1 ([Apêndice A](#)) apresentam a renda média dos 92 municípios do Rio de Janeiro, além da mesorregião à qual esse município pertence. Mesorregiões são divisões geográficas estabelecidas pelo IBGE, as quais agrupam municípios com características sociais e econômicas similares, utilizadas para planejamento e estudos regionais. Inicialmente, os alunos são convidados a analisar a renda média per capita nos municípios das mesorregiões do estado do Rio de Janeiro por meio dos boxplots, com base nos dados do Censo de 2010 do IBGE. Para isso, devem-se utilizar os gráficos boxplots ([Gráfico 10](#)) e a Tabela 1 ([Apêndice A](#)), permitindo uma comparação da renda média nas várias mesorregiões do estado do Rio de Janeiro. Os objetivos dessa tarefa são: (i) despertar nos alunos a percepção sobre como a forma de apresentação de dados estatísticos pode influenciar a interpretação dos dados e (ii) promover uma discussão sobre como ler e analisar os dados estatísticos apresentados, por meio da observação da relação da renda média de um município para o outro.

Gráfico 10 - Renda média x Mesorregiões



Fonte: Elaborado pelos autores a partir de dados do IBGE (Censo Demográfico 2010).

As perguntas propostas para essa situação-problema são:

Pergunta 1: O que está sendo representado no eixo vertical dos gráficos boxplots?

Pergunta 2: Quais mesorregiões estão sendo comparadas nos gráficos boxplots?

Pergunta 3: Observando os gráficos boxplots e a tabela 1, o que são as bolinhas pretas?

Pergunta 4: Observando a mesorregião da Baixadas Litorâneas e a tabela 1, por que o município de Rio das Ostras está lá em cima (o maior valor de renda média, nessa mesorregião)?

Pergunta 5: Observando a mesorregião da Metropolitana do Rio de Janeiro e a tabela 1, por que os municípios do Rio de Janeiro e de Niterói estão lá em cima (os maiores valores de renda média, nessa mesorregião)?

Pergunta 6: Observando a mesorregião Norte Fluminense e a tabela 1, por que o município de Macaé está lá em cima (o maior valor de renda média, nessa mesorregião), e o município de São Francisco de Itabapoana está lá embaixo (o menor valor de renda média, nessa mesorregião)?

Pergunta 7: Observando a mesorregião Sul Fluminense, o que seria a linha horizontal cinza dentro do retângulo do boxplot? E o que se poderia dizer a respeito das bolinhas acima da linha cinza? E das que estão abaixo?

Pergunta 8: Observando a mesorregião Sul Fluminense e a tabela 1, quais os municípios que compõem a metade com renda média superior?

Pergunta 9: Em qual mesorregião a mediana da renda média é mais alta?

Pergunta 10: Compare as medianas da renda média das mesorregiões Baixadas Litorâneas e Norte Fluminense. O que se pode observar?

Pergunta 11: Por que o retângulo do boxplot da mesorregião Metropolitana do Rio de Janeiro é mais largo do que o da mesorregião Norte Fluminense?

Pergunta 12: É possível afirmar que uma mesorregião é mais homogênea do que outra em termos de renda? Por quê?

- Respostas esperadas:

Resposta para a Pergunta 1:

É esperado que os alunos compreendam que no eixo vertical está representada a renda média, expressa em reais (R\$).

Resposta para a pergunta 2:

É esperado que os alunos compreendam que estão sendo comparadas as seguintes mesorregiões do estado do Rio de Janeiro: Baixadas Litorâneas, Centro Fluminense, Metropolitana do Rio de Janeiro, Noroeste Fluminense, Norte Fluminense e Sul Fluminense.

Resposta para a Pergunta 3:

É esperado que os alunos compreendam que as bolinhas pretas são as representações dos municípios de cada mesorregião em relação à renda média.

Resposta para a Pergunta 4:

É esperado que os alunos compreendam que Rio das Ostras é um município atípico, ou seja, aquele cuja renda média é a mais diferente da dos demais municípios da mesorregião das Baixadas Litorâneas.

Resposta para a Pergunta 5:

É esperado que os alunos compreendam os municípios atípicos da mesorregião Norte Fluminense, com relação à renda média.

Resposta para a Pergunta 6:

É esperado que os alunos compreendam a mediana das rendas médias na mesorregião Sul Fluminense, além de compreender o que representam os municípios com valores acima ou abaixo dessa mediana.

Resposta para a Pergunta 7:

É esperado que os alunos compreendam os dois municípios atípicos da mesorregião Metropolitana do Rio de Janeiro, em relação à renda média.

Resposta para a Pergunta 8:

É esperado que os alunos compreendam quais municípios da mesorregião Sul Fluminense apresentavam renda média superior à mediana das rendas médias.

Resposta para a Pergunta 9:

É esperado que os alunos compreendam que a mesorregião que apresenta a mediana da renda média mais alta é a Baixadas Litorâneas.

Resposta para a Pergunta 10:

É esperado que os alunos compreendam que a mediana da renda média da mesorregião Baixadas Litorâneas é maior do que a da mesorregião Norte Fluminense, indicando que, em termos centrais, a renda média é mais elevada na mesorregião Baixadas Litorâneas.

Resposta para a Pergunta 11:

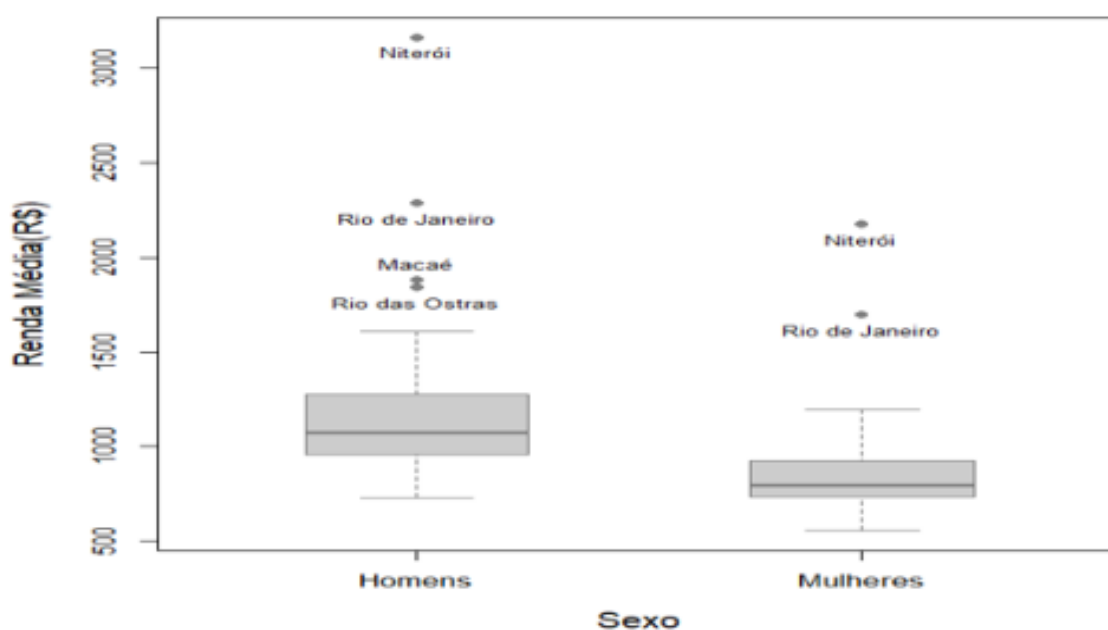
É esperado que os alunos compreendam que o boxplot da mesorregião Metropolitana do Rio de Janeiro é mais largo do que o da mesorregião Norte Fluminense, porque os dados da região Metropolitana apresentam maior dispersão, ou seja, maior variabilidade da renda média entre os municípios dessa mesorregião.

Resposta para a Pergunta 12:

É esperado que os alunos compreendam que sim, porque mesorregiões com boxplots mais estreitos, como o Norte Fluminense, podem ser consideradas mais homogêneas em termos de renda, pois seus dados apresentam menor variabilidade.

No mesmo contexto, os alunos são convidados a analisar a renda domiciliar per capita de todos os municípios do estado do Rio de Janeiro por sexo, com base nos dados do Censo de 2010 do IBGE. Para isso, devem-se utilizar os gráficos boxplots abaixo das rendas médias de todos os municípios do estado do Rio de Janeiro, por sexo, e a Tabela 2 ([Apêndice A](#)), permitindo uma comparação da distribuição da renda per capita por sexo, nesses municípios. O objetivo dessa tarefa foi entender as diferenças das rendas médias e perceber que os homens ganham consistentemente mais do que as mulheres.

Gráfico 11 - Renda média x Sexo (todos os municípios do estado do RJ)



Fonte: Elaborado pelos autores a partir de dados do IBGE (Censo Demográfico 2010).

As perguntas propostas para essa situação-problema são:

Pergunta 1: O que está sendo representado no eixo vertical dos gráficos boxplots?

Pergunta 2: Quais grupos estão sendo comparados no eixo horizontal dos gráficos boxplots?

Pergunta 3: O que os pontos identificados com nomes de cidades indicam?

Pergunta 4: Em qual grupo a mediana da renda média é maior: homens ou mulheres?

Pergunta 5: Compare as medianas da renda média dos dois grupos. Qual diferença é possível observar?

Pergunta 6: O tamanho do retângulo do boxplot é o mesmo para homens e mulheres? O que essa diferença indica?

Pergunta 7: Existem valores atípicos nos gráficos boxplots? Em qual grupo aparecem mais?

Pergunta 8: É possível afirmar que há maior desigualdade de renda em um dos grupos? Justifique com base nos gráficos boxplots.

Pergunta 9: Apenas com esses gráficos boxplots, podemos concluir que todos os homens ganham mais do que todas as mulheres? Por quê?

- Respostas esperadas:

Resposta para a Pergunta 1:

É esperado que os alunos compreendam que no eixo vertical está representada a renda média, expressa em reais (R\$).

Resposta para a Pergunta 2:

É esperado que os alunos compreendam que no eixo horizontal estão sendo comparados dois grupos, definidos pelo sexo: homens e mulheres.

Resposta para a Pergunta 3:

É esperado que os alunos compreendam que os pontos identificados com nomes de cidades indicam valores atípicos, ou seja, cidades cuja renda média se distancia significativamente da maioria dos dados do grupo ao qual pertencem.

Resposta para a Pergunta 4:

É esperado que os alunos compreendam que a mediana da renda média é maior no grupo dos homens.

Resposta para a Pergunta 5:

É esperado que os alunos compreendam que a mediana da renda média dos homens é superior à das mulheres, indicando que, de modo geral, a renda média masculina tende a ser maior do que a feminina.

Resposta para a Pergunta 6:

É esperado que os alunos compreendam que não, pois o retângulo do boxplot dos homens é maior do que o das mulheres. Essa diferença indica que a variação da renda média entre os homens é maior do que entre as mulheres.

Resposta para a Pergunta 7:

É esperado que os alunos compreendam que sim, pois existem valores atípicos no gráfico. Eles aparecem principalmente no grupo dos homens, representados por cidades com renda média bem acima da maioria dos dados.

Resposta para a Pergunta 8:

É esperado que os alunos compreendam que sim, pois é possível afirmar que há maior desigualdade de renda entre os homens, pois o boxplot desse grupo apresenta maior dispersão e mais valores atípicos.

Resposta para a Pergunta 9:

É esperado que os alunos compreendam que não, pois o boxplot mostra tendências gerais, como mediana, quartis e a dispersão, mas não tem como concluir com precisão sobre os indivíduos ou cidades específicas, ou seja, isso indica que podem existir mulheres com renda média maior do que a de alguns homens. Os gráficos de boxplots podem evidenciar diferenças entre os grupos, mas não sustenta a afirmação de que todos os homens ganham mais do que todas as mulheres.

5 SUGESTÕES

Neste capítulo, apresentamos uma lista de leituras e vídeos que podem complementar o uso deste material em aulas de Matemática e Estatística. Além disso, disponibilizamos uma lista de conjuntos de dados que podem ser utilizados na Aprendizagem das medidas de posição, bem como uma relação de portais que podem apoiar o processo de Ensino e Aprendizagem de Estatística.

5.1 SUGESTÕES DE LEITURAS

ATLASSIAN/AUS. **O que é um boxplot**, [s.d]. Um guia sobre boxplot. Disponível em: <https://www.atlassian.com/data/charts/box-plot-complete-guide>. Acesso em: 18 de nov. de 2025.

LABDGE/UFF. **Boxplot**, 2016. Um guia sobre boxplot. Disponível em: <https://labdge.uff.br/boxplot/>. Acesso em: 18 de nov. de 2025.

LANDIM, F. *et al.* **Medidas de posição e dispersão**. Versão 2.0.alpha, 20 nov. 2025. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada (IMPA-OS), 2025. Disponível em: <https://livroaberto.uniriotec.br/ensino-medio/>. Acesso em: 25 de nov. de 2025.

MAGALHÃES, M. N. **Atividades para o Ensino de Probabilidade e Estatística na Educação Básica**. 1º Edição. Rio de Janeiro: SBM, 2019. Disponível em: <https://sbm.org.br/wp-content/uploads/2021/10/magalhaes-simposio.pdf>. Acesso em: 18 de nov. de 2025.

5.2 SUGESTÕES DE VÍDEOS

ALTMAYER, S. **Boxplot - como interpretar e fazer**. Publicado pelo canal Medicina SA - simplificando aprendizado!. Youtube. 3 de maio de 2021. Disponível em: <https://www.youtube.com/watch?v=Yjg7turJApc>. Acesso em: 15 de maio de 2025.

MURAKAMI, M. **Gráfico de boxplot** ● **Diagrama de caixa**. Publicado pelo canal Prof. Murakami - Matemática Rapidola. Youtube. 24 de maio de 2019. Disponível em: <https://www.youtube.com/watch?v=ebkzrOMjGXU> . Acesso em: 15 de maio de 2025.

RAPOSO, L. **Resumo - Boxplot**. Publicado pelo canal R, Estatística e Aprendizado de Máquina. Youtube. 22 de abril de 2021. Disponível em: <https://www.youtube.com/watch?v=2VZGijVHldo&t=294s>. Acesso em: 25 de nov. de 2025.

STATPLACE. **Boxplot: Como interpretar?**. Publicado pelo canal Statplace. Youtube. 12 de jul. de 2022. Disponível em: <https://www.youtube.com/watch?v=qU2IANG4hYQ&t=226s>. Acesso em: 15 de maio de 2025.

5.3 SUGESTÕES DE CONJUNTO DE DADOS

IBGE/SIDRA. **Banco de Tabelas Estatística**, [s.d]. Plataforma online do Instituto Brasileiro de Geografia e Estatística (IBGE) que disponibiliza os dados de suas pesquisas para acesso público. Disponível em: <https://sidra.ibge.gov.br/home/pnadct/brasil>. Acesso em: 14 de maio de 2025.

MPO. **IPEADATA - Instituto de Pesquisa Aplicada**, [s.d]. Plataforma que reúne séries de dados macroeconômicos, financeiros, sociais e regionais do Brasil. Disponível em: <https://www.ipeadata.gov.br/Default.aspx>. Acesso em: 18 de nov. de 2025.

MS. **Covid-19 no Brasil**, 2020. Painel oficial do Governo Federal para acompanhamento de dados da situação da Covid-19 no Brasil, mantido pelo Ministério da Saúde. Disponível em: https://infoms.saude.gov.br/extensions/covid-19_html/covid-19_html.html. Acesso em: 18 de nov. de 2025.

ONU. **UNdata - Banco de Dados das Nações Unidas**, [s.d]. Portal de dados estatísticos da ONU que fornece acesso a uma ampla gama de informações de países e áreas do mundo. Disponível em: <https://data.un.org/>. Acesso em: 18 de nov. de 2025.

SESP. **ISP Dados Abertos - Instituto de Segurança Pública**, 1999. Página de dados abertos do Instituto de Segurança Pública do Rio de Janeiro. Disponível em: <https://www.ispdados.rj.gov.br/>. Acesso em: 18 de nov. de 2025.

5.4 SUGESTÕES DE PORTAIS

CONRE-3. **Educação Estatística**, [s.d]. Portal que oferece atividades, vídeos e recursos pedagógicos para professores que ensinam Estatística. Disponível em: https://www.conre3.org.br/portal/educacao_estatistica/. Acesso em: 26 de nov. de 2025.

GAPMINDER. **Visualizações de dados e informações globais**, 2005. Organização educacional independente sem fins lucrativos que combate concepções errôneas em todo o mundo. Disponível em: <https://www.gapminder.org/>. Acesso em: 7 de maio de 2025.

IBGE. **IBGEeduca**, 2018. Portal do IBGE voltado para a educação com conteúdos atualizados e lúdicos sobre o Brasil. Disponível em: <https://educa.ibge.gov.br/>. Acesso em: 7 de maio de 2025.

IME/UNICAMP. **Matemática Multimídia**, 2007. Recursos educacionais multimídia para a matemática do ensino médio. Disponível em: <https://m3.ime.unicamp.br/>. Acesso em: 18 de nov. de 2025.

IME/USP. **AtivEstat - Atividades de Estatística**, [s.d]. Portal educacional que oferece atividades e recursos pedagógicos para professores que ensinam Estatística. Disponível em: <https://www.ime.usp.br/ativestat/>. Acesso em: 7 de maio de 2025.

INE/PT. **ALEA – Ação Local de Estatística Aplicada**, 1999. Projeto de parceria que visa a elaboração e a disponibilização de instrumentos de apoio ao ensino e à aprendizagem da Estatística. Disponível em: <https://www.alea.pt/index.php>. Acesso em: 18 de nov. de 2025.

MEC. Disponível em: <https://portaldoprofessor.mec.gov.br/index.html>. Acesso em: 7 de maio de 2025.

MEC. **Portal do Professor**, 2008. Portal para o professor acessar sugestões de planos de aula, baixar mídias de apoio, ter notícias sobre educação e iniciativas do

UNIRIO/IMPA. **Livro Aberto de Matemática**, 2016. Projeto colaborativo para criar e distribuir materiais didáticos de Matemática. Disponível em: <https://livroaberto.uniriotec.br/>. Acesso em: 25 de nov. de 2025.

6 ALGUMAS CONSIDERAÇÕES

Cabe ressaltar que, durante a aplicação do Produto Educacional na pesquisa de campo, algumas sugestões e percepções foram destacadas pelos alunos, entre elas, a importância do uso de Tecnologias Digitais no Ensino de Estatística. Da mesma forma, o pesquisador também reconheceu o potencial das Tecnologias Digitais como recurso capaz de potencializar o processo de Ensino e Aprendizagem de Estatística, conforme discutido no [capítulo 3](#).

Dessa forma, foi desenvolvido um recurso digital em conjunto com o projeto de iniciação científica “Ensino de Estatística e Probabilidade na Educação Básica por Projetos” (Freitas, 2025), com acesso livre e gratuito em: <https://milenapaz.shinyapps.io/Outliers/>. Esse aplicativo web tem como objetivo explorar a conceituação de medidas resumo por meio de boxplot, proporcionando uma maior interatividade e dinamismo no processo de Ensino e Aprendizagem. Entendemos que tal ferramenta possa favorecer tanto o trabalho docente, quanto a construção do conhecimento pelos alunos, tornando o estudo das medidas de posição envolvente.

Além disso, conforme apresentado e discutido no [capítulo 3](#), foi indicada a existência de algumas plataformas gratuitas da web, tais como: *Common Online Data Analysis Platform (CODAP)*; *iNZight Lite* e *Tuvalabs*. Essas ferramentas se destacam na exploração de gráficos. O uso dessas plataformas possibilita práticas de Análise Exploratória de Dados, permitindo a visualização e comparação de informações por meio de diversos gráficos exibidos simultaneamente, assim, complementando o recurso digital que apresentamos anteriormente para Análise Exploratória de Dados com gráficos simultâneos. O *CODAP* está disponível em: <https://codap.concord.org/>; o *iNZight Lite*, em: <https://lite.docker.stat.auckland.ac.nz/>; e o *Tuvalabs*, em: <https://tuvalabs.com/>.

Por fim, desejamos que este conjunto de tarefas estimule uma nova forma de abordagem em relação ao processo de Ensino e Aprendizagem de Estatística e que sirva como mais uma peça como uma metodologia alternativa para o Ensino de Estatística, promovendo, assim, um ensino transversal e interdisciplinar. Esperamos, ainda, que inspire professores e pesquisadores a desenvolverem metodologias e reflexões integradas ao contexto educacional, capazes de impactar o aprendizado dos estudantes e contribuir com uma formação de pessoas mais críticas e

questionadoras. Salienta-se que é possível utilizar sequência de atividades similares, baseadas em outros conjuntos de dados, de maneira a trabalhar com temas que sejam de especial interesse dos alunos.

REFERÊNCIAS

- AGRESTI, A.; FRANKLIN, C.; KLINGENBERG, B. **Statistics**:The art and science of learning from data. 5° eds. UK:Pearson, 2023.
- ANDRE, M.; LAVICZA, Z. Technology changing statistics education: Defining possibilities, opportunities and obligations. **Electronic Journal of Mathematics & Technology**, v. 13, n. 3, 2019.
- BATANERO, C. Significado y comprensión de las medidas de posición central. **Uno: Revista de Didáctica de las Matemáticas**, v. 25, p. 41–58, 2000.
- BATANERO, C.; BOROVCNIK, M. **Statistics and probability in high school**. Springer, 2016.
- BIEHLER, R. *et al.* Technology for enhancing statistical reasoning at the school level. *In*: BISHOP A. J. *et al.* Ed(s). **Third international handbook of mathematics education, v27**. New York, EUA: Springer, 2013. p. 643-689.
- BURRIL, G. Statistics education and the role of technology. **Mathematical Sciences Education Board**, National Research Council, Washington, D.C. 2000.
- BURRILL, G.; BIEHLER, R. Fundamental statistical ideas in the school curriculum and in training teachers. *In*: BATANERO, C.; BURRIL, G.; READING, C. Ed(s). **Teaching statistics in school mathematics-challenges for teaching and teacher education: A Joint ICMI/IASE study: The 18th ICMI study**, Dordrecht, NLD: Springer, 2011. p. 57-69, 2011.
- BURRILL, G.; PFANNKUCH, M. Emerging trends in statistics education. **ZDM–Mathematics Education**, v. 56, n. 1, p. 19-29, 2023.
- CHAMBERS, J. M. Statistical computing: History and trends. **The American Statistician**, v. 34, n. 4, p. 238-243, 1980.
- CHANCE, B. L.; DELMAS, R.; GARFIELD, J. Reasoning about sampling distributions. *In*: BEN-ZVI, D.; GARFIELD, J. Ed(s). **The challenge of developing statistical literacy, reasoning and thinking**. Dordrecht, NLD: Springer, 2004. p. 295-323.
- CHILE. Ministerio de Educación. **Bases curriculares 7° a 2° medio**. Santiago: Ministerio de Educación, 2016. Disponível em: <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/654> . Acesso em 12 de out. de 2024.
- COBB, G. W.; MOORE, D. S. Mathematics, statistics, and teaching. **The American Mathematical Monthly**, v. 104, n. 9, p. 801-823, 1997.
- ENGEL, J. *et al.* Visualizing multivariate data: Graphs that tell stories. *In*: New skills in the changing world of statistics education. **Proceedings of the Roundtable Conference of the International Association for Statistical Education (IASE)**. 2020.

FIFE, J. H.; JAMES, K.; PETERS, S. A learning progression for variability. **ETS research report series**, v. 2020, n. 1, p. 1-22, dez. 2020.

FREITAS, M. P. **Ensino de Estatística e Probabilidade na Educação Básica por Projetos**. Juiz de Fora: ProPP/UFJF, 2025.

GARFIELD, J.; BEN-ZVI, D. **Developing students' statistical reasoning: Connecting research and teaching practice**. New York: Springer, 2008.

GÓMEZ-BLANCARTE, A. L.; CHÁVEZ, R. R.; AGUILAR, R. D. C. A survey of the teaching of statistical literacy, reasoning and thinking: teachers classroom practice in Mexican high school education. **Statistics Education Research Journal**, v. 20, n. 2, p. 13-13, 2021.

GOULD, R. *et al.* Revolutions in Teaching and Learning Statistics: A Collection of Reflections. *In*: BEN-ZVI, D.; MAKAR, K.; GARFIELD, J. Ed(s). **International handbook of research in statistics education**. Cham: Springer, 2018. p. 457–472.

HSU, J. L.; CHEN, Y. Integrating big data in introductory statistics education-challenges for instructors and students. *In*: HUANG, T. *et al.* (Eds.). **Innovative Technologies and Learning: Third International Conference, ICITL 2020, Porto, Portugal, November 23–25, 2020, Proceedings**. Porto, Portugal: Springer International Publishing, 2020. p. 263-267.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo Demográfico 2010: Resultados**. Rio de Janeiro: IBGE, 2012. Disponível em: <https://www.ibge.gov.br/>. Acesso em 1 abr. de 2025.

ISODA, M.; CHITMUN, S.; GONZALEZ, O. Japanese and thai senior high school mathematics teachers' knowledge of variability. **Statistics Education Research Journal**, v. 17, n. 2, p. 196-215, nov. 2018.

MAGALHÃES, M. N. ; LIMA, A. C. P. **Noções de probabilidade e estatística**. 7ª edição. São Paulo: Edusp, 2023.

MOORE, D. S.; MCCABE, G. P. **Introdução à prática da estatística**. Tradução: FARIAS, A. A. 3º ed. Rio de Janeiro: LTC, 2002.

MORETTIN, P. A.; BUSSAB, W. O. **Estatística básica**. 9º edição. São Paulo:Saraiva, 2017.

NOVAES, D. V.; COUTINHO, C. Q. S. **Estatística para educação profissional e tecnológica**. 2º ed. São Paulo: Atlas, 2013.

OLIVE, J. *et al.* Mathematical knowledge and practices resulting from access to digital technologies. *In*: HOYLES C.; LAGRANGE J. B. Ed(s). **Mathematics education and technology-rethinking the terrain: the 17th ICMI study**, p. 133-177, 2010.

QUIÑONES, D. *et al.* Innovating statistics education: The design of a novel app using design thinking. **Applied Sciences**, v. 14, n. 18, p. 8515, 2024.

RUBEL, L. H.; NICOL, C.; CHRONAKI, A. A critical mathematics perspective on reading data visualizations: Reimagining through reformatting, reframing, and renarrating. **Educational Studies in Mathematics**, v. 108, n. 1, p. 249-268, 2021.

SHARMA, S. Developing statistical literacy with Year 9 students: A collaborative research project. **Research in Mathematics Education**, v. 15, n. 2, p. 203-204, 2013.

SHAUGHNESSY, J. M. Missed opportunities in research on the teaching and learning of data and chance. *In*: BIDDULPH, F.; CARR, K. (Eds.). **People in mathematics education: Proceedings Of The 20th Annual Conference of The Mathematics Education Research Group of Australasia**. Rotorua, NZ: Merga, 1997, v.1, p. 6-22.

SHAUGHNESSY, J. M. Research on statistics' reasoning and learning. *In*: Frank K. Lester, Jr. (Eds.). **Second handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics**. Charlotte, NC: Information Age Publishing, 2007, p. 957-1009.

SILVA, A. M. **O Modelo dos Campos Semânticos: Um modelo epistemológico em Educação Matemática**. Rio de Janeiro: Editora Ciência Moderna Ltda, 2022.

SILVA, B.; SAMÁ, S. Tecnologia digital no ensino de estatística: Perspectivas para uma abordagem pedagógica abrangente a partir da ICOTS. **Revista Sergipana de Matemática e Educação Matemática**, v. 9, n. 1, p. 129–149-129–149, 2024.

SOUSA, E. P.; CAZORLA, I. M.; MONTEIRO, C. E. F. Abordagens conceituais sobre variabilidade estatística em pesquisas brasileiras: Uma revisão sistemática de literatura. **Ensino em Re-Vista**, v.31, p. 1-24, 2024.

TUKEY, J. W. Exploratory data analysis. **Reading/Addison-Wesley**, 1977.

TUKEY, J. W. The technical tools of statistics. **The American Statistician**, v. 19, n. 2, p. 23-28, 1965.

WATSON, J. M. Doing research in statistics education: More than just data. *In*: **Proceedings of the Sixth International Conference on Teaching Statistics: Developing a Statistically Literate Society**. Voorburg, NLD: International Statistical Institute. 2002.

WATSON, J. M. *et al.* The measurement of school students' understanding of statistical variation. **International Journal of Mathematical Education in Science and Technology**, v. 34, n. 1, p. 1-29, 2003.

WILD, C. J.; PFANNKUCH, M. Statistical thinking in empirical enquiry. **International statistical review**, v. 67, n. 3, p. 223-248, 1999.

WILD, C. J.; UTTS, J. M.; HORTON, N. J. What is statistics? *In*: BEN-ZVI, D.; BAKKER, A.; MAKAR, K. Ed(s). **International handbook of research in statistics education**. Cham, CHE: Springer, 2018. p. 4-36.

YATES, F. The use of computers for statistical analysis: A review of aims and achievements. **Bulletin-International Statistical Institute**, v. 44, n. 1, p. 39-53, 1971.

APÊNDICES

APÊNDICE A - TABELAS A SEREM UTILIZADAS NO DESENVOLVIMENTO DAS TAREFAS

Tabela 1: Rendimento domiciliar per capita nos municípios das mesorregiões do estado do Rio de Janeiro

Baixadas Litorâneas

Município	Renda Média	1º Quartil	Mediana	3º quartil
Araruama	609	203	368	666
Armação dos Búzios	734	272	466	765
Arraial do Cabo	638	254	425	740
Cabo Frio	688	233	402	750
Casimiro de Abreu	660	255	440	765
Iguaba Grande	686	253	426	767
Rio das Ostras	871	290	510	1000
São Pedro da Aldeia	630	224	400	690
Saquarema	566	220	350	594
Silva Jardim	439	170	293	510

Centro Fluminense

Município	Renda Média	1º Quartil	Mediana	3º quartil
Areal	514	208	357	600
Bom Jardim	559	242	400	625
Comendador Levy Gasparian	493	230	364	564
Cantagalo	543	200	350	600
Carmo	516	204	342	550
Cordeiro	636	255	443	703

(continuada)

Município	Renda Média	1º Quartil	Mediana	3º quartil
Duas Barras	483	204	345	510
Macuco	503	195	350	550
Nova Friburgo	762	314	510	800
Parafba do Sul	545	210	360	583
Santa Maria Madalena	491	192	340	540
São Sebastião do Alto	403	166	255	510
Sapucaia	517	200	352	590
Sumidouro	442	167	300	510
Trajano de Moraes	481	170	315	510
Três Rios	622	242	400	667

Metropolitana do Rio de Janeiro

Município	Renda Média	1º Quartil	Mediana	3º quartil
Belford Roxo	418	200	327	510
Cachoeiras de Macacu	520	204	340	570
Duque de Caxias	498	204	300	600
Engenheiro Paulo de Frontin	505	228	345	528
Guapimirim	505	200	340	533
Itaboraí	482	200	340	553
Itaguaí	546	230	380	634
Japeri	368	170	278	483
Magé	478	200	340	550
Mangaratiba	704	255	475	800
Maricá	792	283	510	933
Mendes	583	250	400	638
Mesquita	566	237	400	667
Miguel Pereira	718	250	407	746
Nilópolis	652	267	496	770
Niterói	1700	400	900	2153
Nova Iguaçu	493	200	340	565
Paracambi	517	233	383	612
Paty do Alferes	432	170	299	503
Petrópolis	766	264	475	770
Queimados	432	186	325	510

(continuada)

Município	Renda Média	1º Quartil	Mediana	3º quartil
Rio Bonito	563	224	375	605
Rio de Janeiro	1204	300	563	1250
São Gonçalo	572	255	425	674
São João de Meriti	515	240	388	623
São José do Vale do Rio Preto	437	184	318	510
Seropédica	528	204	358	600
Tanguá	404	175	300	510
Teresópolis	723	255	416	702
Vassouras	609	204	375	630

Noroeste Fluminense

Município	Renda Média	1º Quartil	Mediana	3º quartil
Aperibé	467	233	340	510
Bom Jesus do Itabapoana	580	200	357	618
Cambuci	459	178	326	510
Italva	496	208	350	570
Itaocara	529	210	365	566
Itaperuna	640	255	423	680
Laje do Muriaé	429	170	278	510
Miracema	530	177	337	545
Natividade	588	204	370	600
Porciúncula	540	170	340	592
Santo Antônio de Pádua	547	204	340	553
São José de Ubá	389	150	255	489
Varre-Sai	386	170	262	473

Norte Fluminense

Município	Renda Média	1º Quartil	Mediana	3º quartil
Carapebus	537	241	400	650
Campos dos Goytacazes	604	189	340	625
Cardoso Moreira	429	170	290	510

(continuada)

Município	Renda Média	1º Quartil	Mediana	3º quartil
Conceição de Macabu	530	204	355	600
Macaé	907	290	510	1000
Quissamã	512	193	328	543
São Francisco de Itabapoana	338	128	251	407
São Fidélis	496	204	340	524
São João da Barra	485	188	310	510

Sul Fluminense

Município	Renda Média	1º Quartil	Mediana	3º quartil
Angra dos Reis	672	255	452	755
Barra do Piraí	614	250	407	680
Barra Mansa	615	255	432	700
Itaiala	719	250	425	720
Paraty	655	244	465	706
Pinheiral	561	225	383	650
Piraí	567	233	390	643
Porto Real	534	242	390	625
Quatis	516	213	370	605
Resende	784	270	500	850
Rio Claro	541	216	383	625
Rio das Flores	561	190	306	510
Valença	587	216	378	630
Volta Redonda	781	300	510	900

Tabela 2: Rendimento em reais médio e mediano por sexo nos municípios do estado do Rio de Janeiro

Baixasdas Litorâneas

Município	Média	Média	Mediana	Mediana
	Masculina	Feminina	Masculina	Feminina
Araruama	1229	934	700	510
Armação dos Búzios	1441	1063	900	650
Arraial do Cabo	1331	899	800	510
Cabo Frio	1417	1033	800	600
Casimiro de Abreu	1310	936	800	550
Iguaba Grande	1367	1020	800	510
Rio das Ostras	1843	1188	1000	700
São Pedro da Aldeia	1351	917	800	510
Saquarema	1145	870	700	510
Silva Jardim	921	672	580	510

Centro Fluminense

Município	Média	Média	Mediana	Mediana
	Masculina	Feminina	Masculina	Feminina
Areal	1020	778	700	510
Bom Jardim	994	823	600	510
Comendador Levy Gasparian	941	670	600	510
Cantagalo	1049	800	600	510

(continuada)

Município	Média Masculina	Média Feminina	Mediana Masculina	Mediana Feminina
Carmo	978	759	550	510
Cordéiro	1208	891	800	510
Duas Barras	866	681	510	510
Macuco	1023	767	600	510
Nova Friburgo	1343	996	800	600
Paraíba do Sul	1035	822	600	510
Santa Maria Madalena	893	750	510	510
São Sebastião do Alto	734	697	510	510
Sapucaia	1002	784	650	510
Sunadouro	785	634	510	510
Trajano de Moraes	849	776	510	510
Três Rios	1181	918	700	510

Metropolitana do Rio de Janeiro

Município	Média Masculina	Média Feminina	Mediana Masculina	Mediana Feminina
Belford Roxo	924	673	700	510
Cachoeiras de Macacu	1047	753	650	510
Duque de Caxias	1079	774	800	510
Engenheiro Paulo de Frontin	942	733	510	510
Guapimirim	1079	787	700	510
Itaboraí	1025	721	700	510
Itaguaí	1193	798	900	510
Japeri	835	606	600	510
Magé	1050	735	710	510
Mangaratiba	1449	982	900	600
Maricá	1563	1101	1000	600
Menzes	1064	833	600	510
Mesquita	1199	853	800	530
Miguel Pereira	1342	1031	700	510
Nilópolis	1319	952	973	600
Niterói	3166	2176	1500	1020
Nova Iguaçu	1072	791	750	510

(continuada)

Município	Média Masculina	Média Feminina	Mediana Masculina	Mediana Feminina
Paracambi	1020	747	700	510
Paty do Alferes	864	682	510	510
Petrópolis	1456	1060	800	600
Queimados	962	700	700	510
Rio Bonito	1122	826	650	510
Rio de Janeiro	2288	1698	1000	800
São Gonçalo	1162	822	800	510
São João de Meriti	1073	767	800	510
São José do Vale do Rio Preto	871	625	600	510
Seropédica	1111	803	800	510
Tanguá	829	630	600	510
Teresópolis	1394	1046	765	510
Vassouras	1146	928	600	510

Noroeste Fluminense

Município	Média Masculina	Média Feminina	Mediana Masculina	Mediana Feminina
Aperibé	865	671	600	510
Bom Jesus do Itabapoana	1116	850	510	510
Cambuci	851	700	510	510
Italva	954	688	600	510
Itaocara	928	763	510	510
Itaperuna	1199	876	700	510
Laje do Muriaé	788	721	510	510
Miracema	1026	790	510	510
Natividade	1052	857	510	510
Porciúncula	1030	835	510	510
Santo Antônio de Pádua	1037	796	510	510
São José de Ubá	728	629	510	510
Varre-Sai	771	651	510	510

Norte Fluminense

Município	Média Masculina	Média Feminina	Mediana Masculina	Mediana Feminina
Carapebus	1105	775	800	510
Campos dos Goytacazes	1275	964	700	510
Cardoso Moreira	877	676	510	510
Conceição de Macabu	1063	801	700	510
Macaré	1884	1195	1000	700
Quissamã	1074	775	650	510
São Francisco de Itabapoana	742	556	510	510
São Fidélis	924	758	510	510
São João da Barra	990	721	600	510

Sul Fluminense

Município	Média Masculina	Média Feminina	Mediana Masculina	Mediana Feminina
Angra dos Reis	1424	981	1000	600
Barra do Piraí	1227	875	800	510
Barra Mansa	1272	854	900	510
Itaíaiá	1470	931	800	510
Paraty	1325	971	800	600
Pinheiral	1199	758	800	510
Piraí	1145	822	700	510
Porto Real	1121	746	800	510
Quatis	1053	753	720	510
Resende	1607	1062	1000	600
Rio Claro	1096	759	750	510
Rio das Flores	1217	710	510	510
Valença	1096	883	600	510
Volta Redonda	1595	1024	1000	600