

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS EXATAS
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL
(PROFMAT)

Douglas Barbosa

Estatística no Ensino Fundamental e Médio: Resoluções de Questões Utilizando o
GeoGebra e Planilhas Eletrônicas

Juiz de Fora

2023

Douglas Barbosa

**Estatística no Ensino Fundamental e Médio: Resoluções de Questões Utilizando o
GeoGebra e Planilhas Eletrônicas**

Dissertação apresentada ao Mestrado Profissional em Matemática em Rede Nacional (PROFMAT) da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Matemática. Área de concentração: Ensino de Matemática

Orientador: Prof. Dr. Rogério Casagrande

Juiz de Fora

2023

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF
com os dados fornecidos pelo(a) autor(a)

Barbosa, Douglas.

Estatística no Ensino Fundamental e Médio : Resoluções de Questões Utilizando o GeoGebra e Planilhas Eletrônicas / Douglas Barbosa. – 2023.
140 f. : il.

Orientador: Rogério Casagrande

Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas. Mestrado Profissional em Matemática em Rede Nacional (PROFMAT), 2023.

1. Estatística descritiva. 2. Ensino e Aprendizagem. 3. Ferramentas Tecnológicas. I. Casagrande, Rogério. orient. II. Título.

Douglas Barbosa

Estatística no Ensino Fundamental e Médio: Resoluções de Questões Utilizando o GeoGebra e Planilhas Eletrônicas

Dissertação apresentada ao Mestrado Profissional em Matemática em Rede Nacional (Profmat) da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Matemática. Área de concentração: Ensino de Matemática.

Aprovada em 07 de março de 2023.

BANCA EXAMINADORA

Prof. Dr. Rogério Casagrande - Orientador

Universidade Federal de Juiz de Fora

Prof. Dr. Nelson Dantas Louza Junior

Universidade Federal de Juiz de Fora

Prof. Dra. Gilcélia Regiane de Souza

Universidade Federal de São João del-Rei - Campus Alto Paraopeba

Juiz de Fora, 31/01/2023.



Documento assinado eletronicamente por **Rogério Casagrande, Coordenador(a) em exercício**, em 07/03/2023, às 16:14, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Nelson Dantas Louza Junior, Professor(a)**, em 09/03/2023, às 08:58, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gilcelia Regiane de Souza, Usuário Externo**, em 14/03/2023, às 16:16, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1132864** e o código CRC **CF8AE069**.

Dedico este trabalho a minha mãe Maria das Graças,
aos meus irmãos Denise, Andréa e Anderson e aos meus
sobrinhos. Amo vocês!

AGRADECIMENTOS

Agradeço a minha mãe Maria das Graças, que sempre fez o melhor que pode e lutou para criar seus filhos.

Aos meus irmãos, Denise, Andréa e Anderson, cuja convivência me ensinou a viver em harmonia e em coletividade com a sociedade.

Aos meus sobrinhos, Maria Clara, Amanda, Bernardo e Pedro Henrique, que são a alegria da família.

Aos meus colegas de turma do PROFMAT, pelas dificuldades compartilhadas, momentos de descontração e alegria vivenciados.

Aos meus amigos e aos meus colegas de trabalho, pelo incentivo recebido durante esse período de estudo.

Aos professores do PROFMAT da Universidade Federal de Juiz de Fora, pelo conhecimento proporcionado.

Ao meu orientador, pela dedicação dispensada na elaboração desse trabalho.

Ao coordenador do curso, pelas orientações, suporte e apoio em todas as etapas do mestrado.

E ao Wander, que foi como um pai para mim, que sempre me incentivou e que acreditava na minha capacidade. O senhor não está mais entre nós, mas estará sempre nas minhas lembranças.

RESUMO

Este trabalho propõe a utilização de ferramentas tecnológicas para a resolução de exercícios de estatística descritiva. Como uma proposta de um método de ensino e aprendizagem, por meio dos *softwares* GeoGebra e LibreOffice Calc, foram resolvidos exercícios selecionados, e em alguns casos apresentando a sua representação geométrica e gráfica e a integração com outros conteúdos de forma interdisciplinar. As ferramentas tecnológicas possuem recursos que auxiliam na compreensão dos conceitos de estatística, assim como a visualização dos mesmos, e essa facilidade foi um dos diferenciais para utilizá-las na proposta de resolução dos exercícios. Em cada exercício, inicialmente foi resolvido o que o enunciado solicitava e, após a resolução, foi ampliado a discussão do mesmo, por meio de outros questionamentos relacionados, apresentação de forma gráfica do resultado ou inter-relacionando com outros conteúdos. Utilizando-se de gráficos dispostos na mídia para análise, foi proposto a sua revisão para a análise inicial e posterior correção com as informações mais claras e informativas. Verifica-se que as ferramentas tecnológicas possuem uma variedade de opções que podem facilitar o entendimento dos conceitos de estatística descritiva. A apropriação pelos professores desse método de ensino e aprendizagem contribui para a inclusão dos alunos no processo de análise dos resultados, na apresentação e no questionamento dos dados pesquisados.

Palavras-chave: Estatística descritiva. Ensino e aprendizagem. Ferramentas tecnológicas. GeoGebra. Libreoffice Calc.

ABSTRACT

This work proposes the use of technological tools to solve descriptive statistics exercises. As a proposal for a teaching and learning method, using GeoGebra and LibreOffice Calc softwares, selected exercises were solved, and in some cases presenting their geometric and graphic representation and the integration with other contents in an interdisciplinary way. The technological tools have resources that help in understanding the concepts of statistics, as well as their visualization, and this facility was one of the differentials for using them in the proposed resolution of the exercises. In each exercise, what the statement requested was initially resolved and, after the resolution, the discussion of the same was expanded, by means of other related questions, graphical presentation of the result or interrelation with other contents. Using graphs displayed in the media for analysis, it was proposed to revise it for the initial analysis and subsequent correction with clearer and more informative information. It is verified that the technological tools have a variety of options that can facilitate the understanding of the concepts of descriptive statistics. The appropriation by teachers of this teaching and learning method contributes to the inclusion of students in the process of analysis of results, in the presentation and questioning of researched data.

Keywords: Descriptive Statistics. Teaching and learning. Technological tools. GeoGebra. LibreOffice Calc.

LISTA DE ILUSTRAÇÕES

Figura 1	– Classificação de Variáveis	28
Gráfico 1	– Barras Verticais	36
Gráfico 2	– Barras Horizontais	37
Gráfico 3	– Barras Múltiplas	37
Gráfico 4	– Setores	38
Gráfico 5	– Histograma	39
Gráfico 6	– Linhas	40
Gráfico 7	– Diagrama de Pontos	41
Figura 2	– Construção do Diagrama de Ramos e Folhas	42
Gráfico 8	– Diagrama de Ramo e Folhas	42
Figura 3	– Quartis de uma distribuição	54
Gráfico 9	– Diagrama de Caixa para comparação de notas	58
Figura 4	– Tela Inicial do <i>Software</i> GeoGebra Classic 5	65
Figura 5	– Tela Inicial do <i>Software</i> LibreOffice Calc	66
Figura 6	– Habilitar Planilha no GeoGebra	68
Figura 7	– Inserção de dados na Planilha do GeoGebra	68
Figura 8	– Exibição dos resultados na Janela de Álgebra	69
Gráfico 10	– Exibição gráfica dos resultados	70
Gráfico 11	– Inclusão da reta que identifica a média	71
Figura 9	– Exibição da Janela de Álgebra com os novos resultados	71
Gráfico 12	– Identificação da Mediana no Diagrama de Pontos	72
Figura 10	– Dados digitados na Planilha do LibreOffice Calc	74
Figura 11	– Fórmula da média inserida na planilha	74
Figura 12	– Resultado da média	75
Figura 13	– Classificação decrescente de acordo com a média	75
Gráfico 13	– Gráfico de linhas para os dados do problema	76
Gráfico 14	– Configuração do gráfico de linhas	76
Gráfico 15	– Gráfico de Barras com os dados do problema	78
Gráfico 16	– Análise do Desvio Padrão	79
Figura 14	– Cálculo do Desvio Padrão	79
Figura 15	– Cálculo do Coeficiente de Variação	80
Figura 16	– Inserindo o Controle Deslizante	81
Figura 17	– Controle Deslizante	82
Figura 18	– Dados inseridos na Planilha do GeoGebra	82
Figura 19	– Listas de notas dos alunos	83
Figura 20	– Cálculo da média ponderada do Candidato I	83
Figura 21	– Exibição na Janela de Álgebra das médias calculadas	84

Figura 22 – Exibição dos Resultados	84
Figura 23 – Cálculo final da média	85
Gráfico 17 – Exibição gráfica dos resultados	86
Gráfico 18 – Diagrama resultado final	87
Figura 24 – Comando Tabela de Frequências no GeoGebra	88
Figura 25 – Visualização da Tabela de Frequências	88
Figura 26 – Menu Análise Univariada	89
Gráfico 19 – Histograma da Tabela de Frequências	90
Gráfico 20 – Histograma exibido na Janela de Visualização	90
Figura 27 – Inserção dos dados na planilha	93
Figura 28 – Planilha com as fórmulas inseridas	94
Figura 29 – Planilha com valores calculados	94
Figura 30 – Planilha com valores percentuais	95
Gráfico 21 – Frequência absoluta acumulada	95
Gráfico 22 – Frequência percentual acumulada	96
Figura 31 – Caixa de Confirmação Análise Univariada	97
Figura 32 – Caixa de Diálogo análise de dados	98
Figura 33 – Seleção do Diagrama de Caixa (<i>Box Plot</i>)	98
Figura 34 – Visualização do Diagrama de Caixa calculado	99
Figura 35 – Visualização do Diagrama de Pontos inserido	100
Figura 36 – Identificação da Média da Amostra	100
Figura 37 – Identificação de <i>Outliers</i>	101
Gráfico 23 – Diagrama de Pontos com os gols marcados	103
Gráfico 24 – Diagrama com o Controle deslizante	104
Figura 38 – Exibição dos Quartis	105
Figura 39 – Identificação dos Quartis	106
Figura 40 – Dados digitados na planilha	108
Figura 41 – Cálculo dos resultados	108
Figura 42 – Cálculo da quantidade de jovens que trabalham	109
Gráfico 25 – Gráfico de Pizza	109
Gráfico 26 – Configuração do Gráfico de Pizza	110
Figura 43 – Inflação no Brasil 2009 a 2013	111
Figura 44 – Dados do gráfico digitados na planilha	112
Gráfico 27 – Gráfico da Inflação 2009 a 2013 modificado	113
Figura 45 – Correção do gráfico pela emissora	113
Figura 46 – Divulgação de intenção de voto do Candidato X	114
Figura 47 – Inserção de dados na planilha	115
Gráfico 28 – Gráfico da intenção de votos corrigido	115
Figura 48 – Taxa de desemprego nos EUA em 2011	116

Figura 49 – Digitação dos dados de desemprego nos EUA	117
Gráfico 29 – Taxa de Desemprego com adequação	117
Figura 50 – Comparativo de Reservas Internacionais em 2013	118
Figura 51 – Digitação dos valores das reservas internacionais na planilha	119
Gráfico 30 – Correção do Gráfico de Reservas Internacionais	120
Gráfico 31 – Vendas acumuladas	121
Gráfico 32 – Vendas não acumuladas	121
Figura 52 – Vendas acumuladas de iPhone em 2013	122
Figura 53 – Comparação Vendas acumuladas com a quantidade de vendas por trimestre	122

LISTA DE TABELAS

Tabela 1 – Tempo, em minutos, necessário para transporte de funcionários	31
Tabela 2 – Distribuição de alunos de uma classe, segundo a cor dos cabelos	32
Tabela 3 – Frequência relativa dos alunos de uma classe, segundo a cor dos cabelos	32
Tabela 4 – Frequência percentual dos alunos de uma classe, segundo a cor dos cabelos	33
Tabela 5 – Distribuição dos alunos de uma classe, segundo o número de irmãos ou irmãs	33
Tabela 6 – Altura dos alunos de uma sala de aula	34
Tabela 7 – Idades dos funcionários de uma escola	42
Tabela 8 – Número de gols na 38 ^a rodada	44
Tabela 9 – Número de livros lidos em uma sala de aula	45
Tabela 10 – Número de livros lidos com cálculos auxiliares	45
Tabela 11 – Altura dos alunos de uma sala de aula	46
Tabela 12 – Altura dos alunos de uma sala de aula com cálculos auxiliares	46
Tabela 13 – Número de livros lidos com frequência acumulada	48
Tabela 14 – Distribuição de bolinhas de gudes	49
Tabela 15 – Distribuição de frequência acumulada da altura dos alunos de uma sala de aula	50
Tabela 16 – Distribuição de frequência da altura dos alunos de uma sala de aula . .	52
Tabela 17 – Nota de duas turmas para comparação	57
Tabela 18 – Cálculo da variância	60
Tabela 19 – Cálculo do desvio padrão	61
Tabela 20 – Idades e Pesos de um grupo	62

LISTA DE ABREVIATURAS E SIGLAS

AT	Amplitude Total
BNCC	Base Nacional Comum Curricular
CAS	Cálculo Algébrico Simbólico
IBGE	Instituto Brasileiro de Geografia e Estatística
LDB	Lei de Diretrizes e Bases da Educação
MEC	Ministério da Educação
PM	Ponto Médio

LISTA DE SÍMBOLOS

Σ	Somatório
\vdash	Intervalo fechado a esquerda

SUMÁRIO

1	INTRODUÇÃO	15
2	ANÁLISE DA BNCC COM RELAÇÃO A ESTATÍSTICA . .	17
2.1	A MATEMÁTICA NA BASE NACIONAL COMUM CURRICULAR .	18
2.2	HABILIDADES E COMPETÊNCIAS DE MATEMÁTICA (ESTATÍSTICA) NA BNCC	20
2.2.1	Matemática - 1º Ano	21
2.2.2	Matemática - 2º Ano	22
2.2.3	Matemática - 3º Ano	22
2.2.4	Matemática - 4º Ano	22
2.2.5	Matemática - 5º Ano	23
2.2.6	Matemática - 6º Ano	23
2.2.7	Matemática - 7º Ano	24
2.2.8	Matemática - 8º Ano	24
2.2.9	Matemática - 9º Ano	25
3	CONCEITOS DE ESTATÍSTICA DESCRITIVA NO ENSINO BÁ- SICO	27
3.1	CONCEITOS FUNDAMENTAIS E DEFINIÇÕES	27
3.2	CONSTRUÇÃO DE TABELAS	29
3.2.1	Tabelas de Frequências para variáveis categorizadas	32
3.2.2	Tabelas de Frequências para variáveis numéricas (discretas) . .	33
3.2.3	Tabelas de Frequências para variáveis numéricas (contínuas) . .	34
3.3	REPRESENTAÇÃO GRÁFICA	35
3.3.1	Representação gráfica de variáveis categorizadas	35
<i>3.3.1.1</i>	<i>Gráfico de barras</i>	<i>36</i>
<i>3.3.1.2</i>	<i>Gráfico de setores</i>	<i>37</i>
3.3.2	Representação gráfica de variáveis numéricas	38
<i>3.3.2.1</i>	<i>Histograma</i>	<i>38</i>
<i>3.3.2.2</i>	<i>Gráfico de linhas</i>	<i>39</i>
<i>3.3.2.3</i>	<i>Diagrama de pontos</i>	<i>40</i>
<i>3.3.2.4</i>	<i>Diagrama de ramo e folhas</i>	<i>41</i>
3.4	MEDIDAS DE TENDÊNCIA CENTRAL	43
3.4.1	Média aritmética	43
3.4.2	Média aritmética ponderada	44
3.4.3	Mediana	46
<i>3.4.3.1</i>	<i>Dados não agrupados</i>	<i>47</i>
<i>3.4.3.2</i>	<i>Dados agrupados</i>	<i>48</i>
<i>3.4.3.2.1</i>	<i>Dados agrupados sem intervalo de classe</i>	<i>48</i>

3.4.3.2.2	Dados agrupados com intervalo de classe	49
3.4.4	Moda	51
3.4.4.1	<i>Dados não agrupados</i>	51
3.4.4.2	<i>Dados agrupados</i>	51
3.4.4.2.1	Dados agrupados sem intervalo de classe	52
3.4.4.2.2	Dados agrupados com intervalo de classe	52
3.5	MEDIDAS DE VARIABILIDADE	53
3.5.1	Quartis	53
3.5.1.1	<i>Diagrama de caixa (box plot)</i>	55
3.5.2	Percentis	58
3.5.3	Variância	59
3.5.4	Desvio Padrão	60
3.5.5	Coeficiente de Variação	62
4	ESTATÍSTICA COM O AUXÍLIO DO GEOGEBRA E LIBREOFFICE CALC	63
4.1	GEOGEBRA E SUAS APLICAÇÕES	64
4.2	LIBREOFFICE CALC E SUAS APLICAÇÕES	65
5	RESOLUÇÃO DE QUESTÕES DE ESTATÍSTICA COM SUPORTE DAS FERRAMENTAS ELETRÔNICAS	67
5.1	EXEMPLO 1	67
5.2	EXEMPLO 2	72
5.3	EXEMPLO 3	80
5.4	EXEMPLO 4	87
5.5	EXEMPLO 5	91
5.6	EXEMPLO 6	96
5.7	EXEMPLO 7	101
5.8	EXEMPLO 8	107
6	INTERPRETAÇÃO DE GRÁFICOS	111
6.1	GRÁFICO 1	111
6.2	GRÁFICO 2	113
6.3	GRÁFICO 3	116
6.4	GRÁFICO 4	118
6.5	GRÁFICO 5	120
7	CONCLUSÃO	123
	REFERÊNCIAS	125
	APÊNDICE A – Caderno de Atividades	127

1 INTRODUÇÃO

O ensino de estatística está previsto desde o início da Educação Básica. A Base Nacional Comum Curricular aborda desde os anos iniciais os conceitos e fundamentações para que os alunos desenvolvam habilidades para a coleta, organização, representação e interpretação de dados. Diante da importância desse conteúdo e por verificar que a estatística é utilizada em diferentes áreas e contextos, o objetivo desse trabalho é apresentar uma proposta de resolução de exercícios de estatística envolvendo ferramentas tecnológicas para analisar os dados obtidos e chegar ao resultado esperado, além de ampliar a discussão dos exercícios.

Atualmente as ferramentas tecnológicas assumem um papel muito relevante no processo de ensino e aprendizagem. A utilização de tecnologia para o ensino de estatística está prevista na BNCC, com o uso de calculadoras e planilhas eletrônicas, para a obtenção de medidas estatísticas, comparação de resultados e apresentação de tabelas e gráficos.

Para auxiliar na resolução dos exercícios de estatística da Educação Básica será utilizado *softwares* livres, que estão disponíveis para *download* gratuito na internet. O uso desses *softwares* pode auxiliar não somente ao desenvolvimento das atividades propostas, mas também auxiliar diversas áreas do conhecimento, preparando os alunos para a vida profissional.

Importante salientar que essa proposta apresenta uma visão pessoal de como utilizar essas ferramentas no processo de ensino e aprendizagem, porém as possibilidades de utilização dos *softwares* podem e devem ser ampliadas pelos professores que forem aplicar as ferramentas em sala de aula, contribuindo para o surgimento de outras aplicações, inspirando para trabalhos futuros.

Essa dissertação está dividida em seis capítulos além dessa introdução.

O segundo capítulo faz uma análise da Base Nacional Comum Curricular na área de Matemática, com foco na área de Estatística. Será apresentando as competências e habilidades para o Ensino Fundamental e as competências específicas para o Ensino Médio, sempre voltadas para a área de Estatística.

O terceiro capítulo está destinado ao referencial teórico de estatística que compõe a BNCC. Dividido em cinco seções, será abordado os conceitos fundamentais necessários para a compreensão teórica, a construção de tabelas, a forma de representação gráfica dos dados, as medidas de tendência central e as medidas de variabilidade. Todas as seções apresentarão exemplos resolvidos para a aplicação da teoria.

O quarto capítulo abordará o uso de ferramentas tecnológicas para o ensino de estatística. Será realizado uma breve introdução aos *softwares* GeoGebra e LibreOffice Calc que serão utilizados no decorrer do estudo. O detalhamento das funções dessas ferramentas será realizado no próximo capítulo.

Após esses estudos, no quinto capítulo serão apresentados oito exercícios de estatística que serão resolvidos utilizando as ferramentas tecnológicas. Além da resolução passo a passo dos exercícios, será sugerido uma complementação das atividades para que possa abranger outros entendimentos relacionados utilizando as possibilidades que as ferramentas apresentam.

O sexto capítulo será destinado a interpretação de gráficos. Utilizando exemplos reais, disponíveis em meios de comunicação, os alunos serão motivados a analisarem os gráficos exibidos, verificarem a sua incoerência e os refazerem de maneira que a informação não seja distorcida e que seja apresentada de forma mais clara.

No sétimo capítulo, serão apresentadas as conclusões feitas durante e após a realização do estudo e também perspectivas para trabalhos futuros na mesma área.

E, por fim, no Apêndice A será apresentado um Caderno de Atividades, que representa uma proposta que pode ser adaptada em planos de aulas, com o objetivo de resolução de exercícios de estatística utilizando os *softwares* do presente estudo.

2 ANÁLISE DA BNCC COM RELAÇÃO A ESTATÍSTICA

Neste capítulo, será realizado um breve levantamento sobre as orientações do Ministério da Educação (MEC), no que diz respeito ao ensino-aprendizagem de Estatística no Ensino de Matemática.

Conforme definido na Lei de Diretrizes e Bases da Educação Nacional (LDB, Lei nº 9.394/1996), a Base Nacional Comum Curricular (BNCC) é um documento normativo que estabelece as aprendizagens essenciais que todos os alunos devem desenvolver ao longo da sua educação básica. Em seu artigo 26, tem-se:

Art. 26. Os currículos da educação infantil, do ensino fundamental e do ensino médio devem ter base nacional comum, a ser complementada, em cada sistema de ensino e em cada estabelecimento escolar, por uma parte diversificada, exigida pelas características regionais e locais da sociedade, da cultura, da economia e dos educandos.

Ao definir uma base nacional comum e orientar as políticas educacionais nas escolas de todo o país, a BNCC promove a equidade, garantindo que não haja grande desigualdade na aprendizagem em escolas do setor público e privado, rurais e urbanas, ou entre diferentes regiões do país, permitindo que os alunos de todas as escolas brasileiras tenham acesso aos mesmos conteúdos.

A Educação Básica no Brasil é dividida em Educação Infantil, Ensino Fundamental e Ensino Médio. A BNCC servirá de referência nacional para as escolas públicas e privadas para a elaboração de seus currículos e propostas pedagógicas. É um documento cujas diretrizes devem ser levadas em consideração para elaborar seus currículos cuja responsabilidade é da Secretaria de Educação, e assim, as escolas ficam responsáveis pela elaboração dos projetos pedagógicos.

A educação integral é um concepção da BNCC e, para isso, contempla diferentes dimensões do desenvolvimento humano: Cognitivo; Intelectual; Físico; Social; Emocional e Cultural. Também estabelece que as aprendizagens essenciais devem ser desenvolvidas pelos alunos durante seu percurso na Educação Básica e que sejam sustentadas por dez competências gerais, que são:

- a) Conhecimento;
- b) Raciocínio científico, crítico e criativo;
- c) Repertório Cultural;
- d) Comunicação;
- e) Cultura digital;
- f) Projeto de vida e mundo do trabalho;

- g) Argumentação;
- h) Autoconhecimento e autocuidado;
- i) Empatia e cooperação;
- j) Responsabilidade e cidadania.

Tais competências são obtidas com a mobilização de Conhecimento (conceitos e procedimentos), Habilidades (práticas cognitiva e socioemocionais), Atitudes e Valores que possibilitem resolver questões complexas do dia a dia e do mundo do trabalho.

Além das competências gerais, a BNCC apresenta competências específicas para cada área de conhecimento e estrutura curricular. As competências específicas estabelecem uma aprendizagem mínima que os alunos devem desenvolver nas etapas da educação básica em cada área de conhecimento, contribuindo para que ao final do ciclo da educação básica os alunos alcancem as competências gerais citadas anteriormente.

Será demonstrado a seguir como a Matemática, e no caso o objeto deste estudo, a Estatística, está inserido dentro da BNCC.

2.1 A MATEMÁTICA NA BASE NACIONAL COMUM CURRICULAR

De acordo com a BNCC, a Matemática no Ensino Fundamental não se restringe apenas à quantificação de fenômenos determinísticos (contagem, medição de objetos, grandezas) e das técnicas de cálculo com os números e com as grandezas, pois também estuda a incerteza proveniente de fenômenos de caráter aleatório.

Logo, no Ensino Fundamental a Matemática deve ter compromisso com o desenvolvimento do letramento matemático, ou seja, prezar pelas competências e habilidades de raciocínio, representação, comunicação e argumentação matemática, propiciando o estabelecimento de conjecturas, a formulação e resolução de problemas em uma diversidade de contextos, utilizando conceitos, procedimentos, fatos e ferramentas matemáticas. Segundo a BNCC, o letramento matemático também:

(...) assegura aos alunos reconhecer que os conhecimentos matemáticos são fundamentais para a compreensão e a atuação no mundo e perceber o caráter de jogo intelectual da matemática, como aspecto que favorece o desenvolvimento do raciocínio lógico e crítico, estimula a investigação e pode ser prazeroso (fruição).

Nesse sentido os processos matemáticos como a resolução de problemas, de investigação, de desenvolvimento de projetos e de modelagem são citados no documento como processos potencialmente ricos para o desenvolvimento de competências fundamentais para o letramento matemático.

A BNCC sugere articular diferentes campos que compõe a Matemática em um conjunto de ideias fundamentais (equivalência, ordem, proporcionalidade, interdependência,

representação, variação e aproximação) que são importantes para o desenvolvimento do pensamento matemático. Assim, propõe para o Ensino Fundamental cinco unidades temáticas correlacionadas: números, álgebra, geometria, grandezas e medidas, e probabilidade e estatística.

A incerteza e o tratamento de dados são estudados na unidade temática Probabilidade e Estatística. Abordando conceitos, fatos e procedimentos presentes em muitas situações cotidianas, pretende-se que os alunos desenvolvam habilidades para coleta, organização, representação, interpretação e análise de dados em uma variedade de contextos. Para tanto, o aluno precisa raciocinar e utilizar conceitos, representações e índices estatísticos para descrever, explicar e prever fenômenos.

A BNCC destaca o uso de tecnologias, como calculadoras e planilhas eletrônicas, para comparação de resultados, construção de gráficos e obtenção de medidas estatísticas. Recomenda também a consulta de páginas de institutos de pesquisa, como a do Instituto Brasileiro de Geografia e Estatística (IBGE), que oferece contextos potencialmente ricos para aprendizagem de conceitos e procedimentos, além de servir como compreensão da realidade.

No que se concerne ao estudo de Estatística no Ensino Fundamental - Anos Iniciais, a BNCC inicia com trabalhos que envolvam a coleta e organização de dados de uma pesquisa de interesse dos estudantes. Pretende-se, desta forma, estimular a leitura, a interpretação e a construção de tabelas e gráficos para a compreensão da estatística que resultará em textos escritos que justifiquem a conclusão. No Ensino Fundamental - Anos Finais, espera-se que os alunos planejem e construam relatórios de pesquisas estatísticas descritivas, incluindo medidas de tendência central, construção de tabelas e gráficos variados. Os alunos devem utilizar os conceitos aprendidos como técnicas de amostragem, incluir questões relevantes da população a ser pesquisada, ou definir a utilização de amostra populacional.

No Ensino Médio, a BNCC, em sua área de Matemática e Tecnologia propõe consolidar, ampliar e aprofundar as aprendizagens essenciais que foram desenvolvidas no Ensino Fundamental, interagindo com os conhecimentos já explorados anteriormente. O foco passa a ser a construção de uma visão integrada da Matemática, aplicada à realidade considerando a vivência cotidiana dos alunos.

Conseqüentemente, quando a realidade é a referência, é preciso levar em conta as vivências cotidianas dos estudantes do Ensino Médio – impactados de diferentes maneiras pelos avanços tecnológicos, pelas exigências do mercado de trabalho, pelos projetos de bem viver dos seus povos, pela potencialidade das mídias sociais, entre outros. Nesse contexto, destaca-se ainda a importância do recurso a tecnologias digitais e aplicativos tanto para a investigação matemática como para dar continuidade ao desenvolvimento do

pensamento computacional, iniciado na etapa anterior. (BNCC, 2017)

Segundo a BNCC, os alunos devem apresentar habilidades relativas aos processos de investigação, de construção de modelos e de resolução de problemas, para que atinjam os propósitos nessa área. Os alunos necessitam desenvolver o raciocínio por meio de interação com seus colegas e professores para desenvolver essa habilidade. Utilizando a investigação, explicação e justificação para os problemas resolvidos, com destaque para a argumentação.

Muitas das vezes o aluno precisa das competências de representar e comunicar para a construção de uma argumentação consistente para o seu raciocínio. Verifica-se na Matemática a utilização dessa competência que não é exclusiva, visto que todas as áreas tem seus processos de representação e, na Matemática é importante para a representação e compreensão dos fatos, ideias e conceitos.

De acordo com a BNCC:

Por esse motivo, espera-se que os estudantes conheçam diversos registros de representação e possam mobilizá-los para modelar situações diversas por meio da linguagem específica da matemática – verificando que os recursos dessa linguagem são mais apropriados e seguros na busca de soluções e respostas – e, ao mesmo tempo, promover o desenvolvimento de seu próprio raciocínio.

Após a resolução dos problemas matemáticos, os estudantes se utilizam da competência de comunicar, para apresentar e justificar os seus resultados. Utilizando não somente símbolos matemáticos, os estudantes devem ser capazes de justificar suas conclusões utilizando argumentações orais dos resultados e elaborando relatórios entre outros registros.

Nessa etapa, os estudantes devem ser capazes de mostrar e argumentar seus resultados, apresentando suas justificativas podendo gerar interação com os demais estudantes, desenvolvendo a competência de comunicação com os demais colegas.

A seguir, serão apresentados os Objetos de Conhecimento e Habilidades propostas para o Ensino Fundamental e as Competências Específicas e Habilidades a serem desenvolvidas no Ensino Médio, com relação a Estatística, o objeto desse estudo.

2.2 HABILIDADES E COMPETÊNCIAS DE MATEMÁTICA (ESTATÍSTICA) NA BNCC

Inicialmente é preciso destacar que cada habilidade é identificada por meio de um código alfanumérico cuja composição é a seguinte para o Ensino Fundamental:

EF01MA22

- As duas primeiras letras identificam a etapa educacional
- Os dois próximos números identificam o ano a que se refere a habilidade
- As duas próximas letras identificam o componente curricular
- O último par de números identificam a posição da habilidade na numeração sequencial daquele ano

Assim, o código EF01MA22 refere-se à vigésima segunda habilidade de Matemática para alunos do primeiro ano do Ensino Fundamental.

Para o Ensino Médio as habilidades também são identificadas por um código alfanumérico, cuja composição é a seguinte:

EM13MAT102

- As duas primeiras letras identificam a etapa educacional
- Os dois próximos números identificam que as habilidades descritas podem ser desenvolvidas em qualquer série do Ensino Médio, conforme definição dos currículos
- A próxima sequência de letras identificam a área ou componente curricular
- A sequência final de números identificam a competência específica a qual se relaciona a habilidade (1º número) e a numeração no conjunto de habilidades relativas à cada competência (dois últimos números)

Portanto, o código EM13MAT102 refere-se a segunda habilidade proposta na área de Matemática e suas Tecnologias relacionada à competência 1, que pode ser desenvolvida em qualquer série do Ensino Médio, conforme definições curriculares.

A seguir será apresentado os Objetos de Conhecimento e Habilidades a serem desenvolvidas pelos alunos em cada série do Ensino Fundamental em relação ao tema de estatística.

2.2.1 Matemática - 1º Ano

Objetos de Conhecimento	Habilidades
Leitura de tabelas e de gráficos de colunas simples	(EF01MA21) Ler dados expressos em tabelas e em gráficos de colunas simples.
Coleta e organização de informações Registros pessoais para comunicação de informações coletadas	(EF01MA22) Realizar pesquisa, envolvendo até duas variáveis categóricas de seu interesse e universo de até 30 elementos e organizar dados por meio de representações pessoais.

2.2.2 Matemática - 2º Ano

Objetos de Conhecimento	Habilidades
Coleta, classificação e representação de dados em tabelas simples e de dupla entrada e em gráficos de colunas	(EF02MA22) Comparar informações de pesquisas apresentadas por meio de tabela de dupla entrada e em gráficos de colunas simples ou barras, para melhor compreender aspectos da realidade próxima. (EF02MA23) Realizar pesquisa em um universo de até 30 elementos, escolhendo até três valores categóricos de seu interesse, organizando os dados coletados em listas, tabelas e gráficos de colunas simples.

2.2.3 Matemática - 3º Ano

Objetos de Conhecimento	Habilidades
Leitura, interpretação e representação de dados em tabelas de dupla entrada e gráficos de barras	(EF03MA26) Resolver problemas cujos dados estão apresentados em tabelas de dupla entrada, gráficos de barras ou de colunas. (EF03MA27) Ler, interpretar e comparar dados apresentados em tabelas de dupla entrada, gráficos de barras ou de colunas, envolvendo resultados de pesquisas significativas, utilizando termos como maior e menor frequência, apropriando-se desse tipo de linguagem para compreender aspectos da realidade sociocultural significativos.
Coleta, classificação e representação de dados referentes a variáveis categóricas por meio de tabelas e gráficos	(EF03MA28) Realizar pesquisa envolvendo variáveis categóricas em um universo de até 50 elementos, organizar os dados coletados utilizando listas, tabelas simples ou de dupla entrada e representá-las em gráficos de colunas simples, com e sem uso de tecnologias digitais.

2.2.4 Matemática - 4º Ano

Objetos de Conhecimento	Habilidades
Leitura, interpretação e representação de dados em tabelas de dupla entrada, gráficos de colunas simples e agrupadas, gráfico de barras e colunas e gráficos pictóricos	(EF04MA27) Analisar dados apresentados em tabelas simples ou de dupla entrada e em gráficos de colunas ou pictóricos, com base em informações das diferentes áreas de conhecimento, e produzir texto com a síntese de sua análise.
Diferenciação entre variáveis categóricas e variáveis numéricas Coleta, classificação e representação de dados de pesquisa realizada	(EF04MA28) Realizar pesquisa envolvendo variáveis categóricas e numéricas e organizar dados coletados por meio de tabelas e gráficos de colunas simples ou agrupadas, com e sem uso de tecnologias digitais.

2.2.5 Matemática - 5º Ano

Objetos de Conhecimento	Habilidades
Leitura, coleta, classificação e representação de dados em tabelas de dupla entrada, gráficos de colunas agrupadas, gráficos pictóricos e gráficos de linhas	<p>(EF05MA24) Interpretar dados estatísticos apresentados em textos, tabelas e gráficos (colunas ou linhas), referentes a outras áreas do conhecimento ou a outros contextos, como saúde e trânsito, e produzir textos com o objetivo de sintetizar conclusões.</p> <p>(EF05MA25) Realizar pesquisa envolvendo variáveis categóricas e numéricas, organizar dados coletados por meio de tabelas, gráficos de colunas, pictóricos e de linhas, com e sem uso de tecnologias digitais, e apresentar texto escrito sobre a finalidade da pesquisa e a síntese dos resultados.</p>

2.2.6 Matemática - 6º Ano

Objetos de Conhecimento	Habilidades
Leitura e interpretação de tabelas e gráficos (de colunas ou barras simples ou múltiplas) referentes a variáveis categóricas e variáveis numéricas	<p>(EF06MA31) Identificar as variáveis e suas frequências e os elementos constitutivos (título, eixos, legendas, fontes e datas) em diferentes tipos de gráfico.</p> <p>(EF06MA32) Interpretar e resolver situações que envolvam dados de pesquisas sobre contextos ambientais, sustentabilidade, trânsito, consumo responsável, entre outros, apresentadas pela mídia em tabelas e em diferentes tipos de gráficos e redigir textos escritos com o objetivo de sintetizar conclusões.</p>
Coleta de dados, organização e registro Construção de diferentes tipos de gráficos para representá-los e interpretação das informações	(EF06MA33) Planejar e coletar dados de pesquisa referente a práticas sociais escolhidas pelos alunos e fazer uso de planilhas eletrônicas para registro, representação e interpretação das informações, em tabelas, vários tipos de gráficos e texto.
Diferentes tipos de representação de informações: gráficos e fluxogramas	(EF06MA34) Interpretar e desenvolver fluxogramas simples, identificando as relações entre os objetos representados (por exemplo, posição de cidades considerando as estradas que as unem, hierarquia dos funcionários de uma empresa etc.).

2.2.7 Matemática - 7º Ano

Objetos de Conhecimento	Habilidades
Estatística: média e amplitude de um conjunto de dados	(EF07MA35) Compreender, em contextos significativos, o significado de média estatística como indicador da tendência de uma pesquisa, calcular seu valor e relacioná-lo, intuitivamente, com a amplitude do conjunto de dados.
Pesquisa amostral e pesquisa censitária Planejamento de pesquisa, coleta e organização dos dados, construção de tabelas e gráficos e interpretação das informações	(EF07MA36) Planejar e realizar pesquisa envolvendo tema da realidade social, identificando a necessidade de ser censitária ou de usar amostra, e interpretar os dados para comunicá-los por meio de relatório escrito, tabelas e gráficos, com o apoio de planilhas eletrônicas.
Gráficos de setores: interpretação, pertinência e construção para representar conjunto de dados	(EF07MA37) Interpretar e analisar dados apresentados em gráfico de setores divulgados pela mídia e compreender quando é possível ou conveniente sua utilização.

2.2.8 Matemática - 8º Ano

Objetos de Conhecimento	Habilidades
Gráficos de barras, colunas, linhas ou setores e seus elementos constitutivos e adequação para determinado conjunto de dados	(EF08MA23) Avaliar a adequação de diferentes tipos de gráficos para representar um conjunto de dados de uma pesquisa.
Organização dos dados de uma variável contínua em classes	(EF08MA24) Classificar as frequências de uma variável contínua de uma pesquisa em classes, de modo que resumam os dados de maneira adequada para a tomada de decisões.
Medidas de tendência central e de dispersão	(EF08MA25) Obter os valores de medidas de tendência central de uma pesquisa estatística (média, moda e mediana) com a compreensão de seus significados e relacioná-los com a dispersão de dados, indicada pela amplitude.
Pesquisas censitária ou amostral Planejamento e execução de pesquisa amostral	(EF08MA26) Selecionar razões, de diferentes naturezas (física, ética ou econômica), que justificam a realização de pesquisas amostrais e não censitárias, e reconhecer que a seleção da amostra pode ser feita de diferentes maneiras (amostra casual simples, sistemática e estratificada).

	(EF08MA27) Planejar e executar pesquisa amostral, selecionando uma técnica de amostragem adequada, e escrever relatório que contenha os gráficos apropriados para representar os conjuntos de dados, destacando aspectos como as medidas de tendência central, a amplitude e as conclusões.
--	---

2.2.9 Matemática - 9º Ano

Objetos de Conhecimento	Habilidades
Análise de gráficos divulgados pela mídia: elementos que podem induzir a erros de leitura ou de interpretação	(EF09MA21) Analisar e identificar, em gráficos divulgados pela mídia, os elementos que podem induzir, às vezes propositadamente, erros de leitura, como escalas inapropriadas, legendas não explicitadas corretamente, omissão de informações importantes (fontes e datas), entre outros.
Leitura, interpretação e representação de dados de pesquisa expressos em tabelas de dupla entrada, gráficos de colunas simples e agrupadas, gráficos de barras e de setores e gráficos pictóricos	(EF09MA22) Escolher e construir o gráfico mais adequado (colunas, setores, linhas), com ou sem uso de planilhas eletrônicas, para apresentar um determinado conjunto de dados, destacando aspectos como as medidas de tendência central.
Planejamento e execução de pesquisa amostral e apresentação de relatório	(EF09MA23) Planejar e executar pesquisa amostral envolvendo tema da realidade social e comunicar os resultados por meio de relatório contendo avaliação de medidas de tendência central e da amplitude, tabelas e gráficos adequados, construídos com o apoio de planilhas eletrônicas.

Para o Ensino Médio, será apresentado na sequência as Competências Específicas e Habilidades referente ao conteúdo de Estatística.

• COMPETÊNCIA ESPECÍFICA 1

Utilizar estratégias, conceitos e procedimentos matemáticos para interpretar situações em diversos contextos, sejam atividades cotidianas, sejam fatos das Ciências da Natureza e Humanas, das questões socioeconômicas ou tecnológicas, divulgados por diferentes meios, de modo a contribuir para uma formação geral.

(EM13MAT102) Analisar tabelas, gráficos e amostras de pesquisas estatísticas apresentadas em relatórios divulgados por diferentes meios de comunicação, identificando, quando for o caso, inadequações que possam induzir a erros de interpretação, como escalas e amostras não apropriadas.

• COMPETÊNCIA ESPECÍFICA 2

Propor ou participar de ações para investigar desafios do mundo contemporâneo e tomar decisões éticas e socialmente responsáveis, com base na análise de problemas

sociais, como os voltados a situações de saúde, sustentabilidade, das implicações da tecnologia no mundo do trabalho, entre outros, mobilizando e articulando conceitos, procedimentos e linguagens próprias da Matemática.

(EM13MAT202) Planejar e executar pesquisa amostral sobre questões relevantes, usando dados coletados diretamente ou em diferentes fontes, e comunicar os resultados por meio de relatório contendo gráficos e interpretação das medidas de tendência central e das medidas de dispersão (amplitude e desvio padrão), utilizando ou não recursos tecnológicos.

- **COMPETÊNCIA ESPECÍFICA 3**

Utilizar estratégias, conceitos, definições e procedimentos matemáticos para interpretar, construir modelos e resolver problemas em diversos contextos, analisando a plausibilidade dos resultados e a adequação das soluções propostas, de modo a construir argumentação consistente.

(EM13MAT316) Resolver e elaborar problemas, em diferentes contextos, que envolvem cálculo e interpretação das medidas de tendência central (média, moda, mediana) e das medidas de dispersão (amplitude, variância e desvio padrão).

- **COMPETÊNCIA ESPECÍFICA 4**

Compreender e utilizar, com flexibilidade e precisão, diferentes registros de representação matemáticos (algébrico, geométrico, estatístico, computacional etc.), na busca de solução e comunicação de resultados de problemas.

(EM13MAT406) Construir e interpretar tabelas e gráficos de frequências com base em dados obtidos em pesquisas por amostras estatísticas, incluindo ou não o uso de *softwares* que inter-relacionem estatística, geometria e álgebra.

(EM13MAT407) Interpretar e comparar conjuntos de dados estatísticos por meio de diferentes diagramas e gráficos (histograma, de caixa (box-plot), de ramos e folhas, entre outros), reconhecendo os mais eficientes para sua análise.

É importante ressaltar que apesar de a Base Nacional Comum Curricular indicar que tal habilidade deve ser desenvolvida em um determinado ano escolar, o conhecimento de Estatística são revisados e ampliados a cada ano. É preciso acompanhar como as habilidades do ano corrente são ligadas as habilidades de anos anteriores, que já foram consolidadas e que servirão de base para aprofundamento do conteúdo nos anos seguintes.

No capítulo a seguir será abordado conceitos teóricos de Estatística Descritiva que compõem o currículo da BNCC.

3 CONCEITOS DE ESTATÍSTICA DESCRITIVA NO ENSINO BÁSICO

A Estatística é um ramo da Matemática usada para coleta, interpretação e análise de dados numéricos de pesquisas sobre a natureza, a sociedade, a economia e o mercado. De acordo com Vieira (2018) a Estatística é “a ciência que fornece os princípios e os métodos para coleta, organização, resumo, análise e interpretação das informações”.

A Estatística teve sua origem quando governos se interessaram em obter informações quantitativas e qualitativas sobre suas riquezas, tributos, populações e moradias.

No Egito antigo, por exemplo, faraós já ordenavam registros estatísticos de suas colheitas. Há mais de dois mil anos, a China já se preocupava com o crescimento populacional por meio de censos. Já no século XIV, o início do Renascimento na Europa também proporcionou novos rumos à ESTATÍSTICA, necessária especialmente para aprimorar a administração de governos (FIOCRUZ, s.d.).

Na atualidade, é visto que a estatística está presente no dia a dia, por meio de índices, tabelas, gráficos e é essencial para o desenvolvimento de todas as ciências.

Segundo Chavante e Prestes (2016) as pesquisas estatísticas são realizadas para diversas finalidades como, por exemplo, verificar a durabilidade de um certo tipo de material ou então a opinião de um certo grupo de pessoas a respeito de um determinado assunto. Seria inviável testar a durabilidade de todo o material, bem como consultar a opinião do grupo inteiro de pessoas, pois demandaria tempo e dinheiro. Imagine se o material a ser testado a durabilidade for uma lâmpada, se a testagem abranger todas as lâmpadas até queimarem, não sobraria lâmpada alguma para ser comercializada.

Observa-se então, que muitas situações necessitam de informações sobre um amplo grupo de elementos (indivíduos, empresas, eleitores, famílias, produtos, clientes, etc.). Devido ao custo, tempo e outros fatores envolvidos, é possível coletar dados somente de uma pequena parte do grupo. (ANDERSON *ET AL.*, 2019).

3.1 CONCEITOS FUNDAMENTAIS E DEFINIÇÕES

Formalmente, as seguintes definições são utilizadas:

População: É o conjunto de todos os elementos de interesse em um determinado estudo.

Amostra: É um subconjunto da população.

As pesquisas estatísticas estão presentes em nosso dia a dia em diversas situações. Quando essas pesquisas envolvem pessoas, geralmente são colhidas informações como idade, massa corporal, estado civil, nível de escolaridade, entre outras. (CHAVANTE

E PRESTES, 2016). Essas informações são exemplos de variáveis estatísticas, assim definidas:

Variável: É uma condição ou característica que descreve uma pessoa, um animal, um lugar, um objeto, uma ideia. A variável assume valores diversos em diferentes unidades.

As variáveis podem ser classificadas em dois tipos:

Variáveis categorizadas (antes chamadas qualitativas): São as variáveis que apresentam uma qualidade ou atributo como resposta.

Variáveis numéricas (antes chamadas quantitativas): São as variáveis que apresentam quantidades como respostas, são obtidas por medição ou contagem.

As variáveis categorizadas (ou qualitativas) são classificadas em dois tipos:

Nominal: Não apresentam uma ordenação entre as possíveis respostas (nacionalidade, religião, afiliação partidária, estado civil, etc.).

Ordinal: Os dados são distribuídos em grupos que tem ordenação natural (escolaridade, satisfação do cliente, etc.).

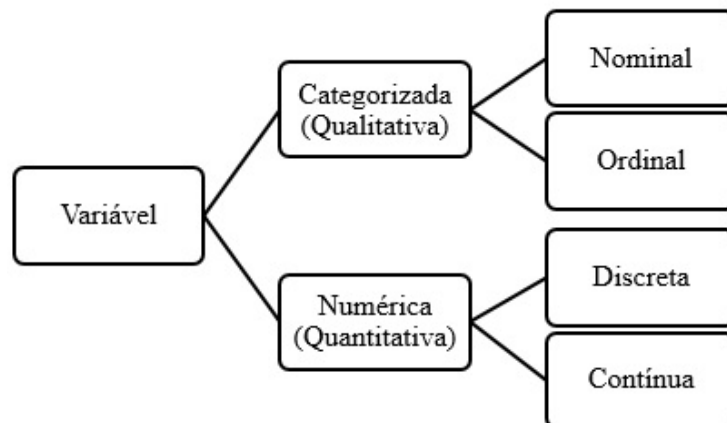
As variáveis numéricas (ou quantitativas) também são classificadas em dois tipos:

Discreta: cujos possíveis valores compõem um conjunto finito ou infinito de números resultantes, geralmente de contagens (quantidade de filhos de um casal, número de cavalos em um haras, quantidade de parafusos em uma cadeira, etc.).

Contínua: cujos possíveis valores formam um intervalo de números reais resultantes, geralmente de mensuração (altura de uma pessoa, tempo de um atleta, quantidade de chuva, etc.).

A Figura 1 apresenta a classificação de uma variável estatística:

Figura 1 - Classificação de Variáveis



Fonte: Vieira (2018).

Nota: Ilustração adaptada.

3.2 CONSTRUÇÃO DE TABELAS

Segundo Larson e Faber (2015), os dados “consistem em informações provenientes de observações, contagens, medições ou respostas”, ou seja, são os valores observados das variáveis. Após a coleta dos dados referente as variáveis, há necessidade de representá-los de forma organizada e prepará-los para análise e interpretação. Essa fase é chamada de apuração. Uma das melhores formas de apresentar os dados organizados é por meio de tabelas (VIEIRA, 2018).

Segundo Vieira (2018), uma tabela deve ter:

1. Título: explica o que a tabela contém.
2. Cabeçalho: especifica o conteúdo de cada coluna.
3. Indicador de linha: especifica o conteúdo de cada linha.
4. Célula: é o espaço na tabela onde os dados são colocados. Esse espaço é criado pelo cruzamento de uma linha com uma coluna.
5. Moldura: configura a tabela. Embora os programas para computador apresentem muitas opções de desenhos de tabelas com grades, as normas brasileiras pedem para evitar traços verticais, principalmente os laterais.

Quando trabalharmos com um grande número de dados, para melhor visualização do comportamento de uma variável, é comum organizá-los por meio de tabelas de frequência ou gráficos. Na sua construção, as variáveis são tratadas de acordo com o tipo (PINHEIRO *ET AL.*, 2009).

A seguir será abordado alguns conceitos importantes para a elaboração e análise das tabelas de frequência.

Dados Brutos: Dados na forma em que foram coletados, sem nenhum tratamento. Geralmente, esse tipo de dado traz pouca ou nenhuma informação ao leitor, sendo necessário organizar os dados, com o intuito de aumentar sua capacidade de informação.

Rol: É a mais simples organização numérica. É a ordenação dos dados em ordem crescente ou decrescente. A simples organização dos dados em um rol aumenta muito a capacidade de informação destes.

Amplitude Total (AT): Corresponde à diferença entre o maior e o menor valor observado em um conjunto de dados.

Classes: Intervalos nos quais os valores da variável analisada são agrupados. A escolha do número de classes (k) é feita pelo pesquisador em função do que se quer apresentar. Se o número de classes for demasiado pequeno, perde-se muita informação. Se o número de classes for grande, há pormenores desnecessários, prejudicial à análise. Não existe um número ideal de classes para um conjunto de dados, embora existam fórmulas para estabelecer quantas classes devem ser construídas.

Alguns critérios utilizados para a determinação do número de classes (k):

- a) Regra de Sturges: $k = 1 + (3,3) \log n$, onde n é o número de observações;
- b) O número de classes será um número inteiro próximo de k , obtido pela fórmula:
 $k = \sqrt{n}$.

Limites de classes: São denominados os extremos de classe. O menor valor é o limite inferior da classe (l_i) e o maior valor é o limite superior da classe (L_i).

Amplitude da Classe (h_i): é a diferença entre os limites superior e inferior da classe.

$$h_i = L_i - l_i.$$

Existem outros procedimentos para determinação da amplitude do intervalo de classe que podem ser encontrados na literatura. Pode-se utilizar a amplitude total dividida pelo número de classes menos 1, ou seja:

$$h_i = \frac{AT}{k - 1}.$$

Ponto Médio da Classe (PM): É dado pela média aritmética ¹ entre os limites superior e inferior da classe.

$$PM = \frac{L_i + l_i}{2}.$$

Frequência absoluta (f_i): Corresponde ao número de observações de uma determinada classe ou um determinado atributo de uma variável categorizada. A soma das frequências absolutas é igual ao número total de observações.

$$\sum f_i = n.$$

O somatório é uma simbologia matemática representada por uma forma ampliada da letra grega Σ (sigma). É utilizada para representar somas sucessivas de elementos de fórmulas e expressões.

Frequência relativa (f_r): Corresponde à proporção do número de observações em uma determinada classe em relação ao total de observações.

$$f_r = \frac{f_i}{\sum f_i} = \frac{f_i}{n}.$$

Esta frequência pode ser expressa em números percentuais. Para isto, basta multiplicar a frequência relativa obtida por 100. A soma das frequências relativas deve ser igual a 1 ou 100%.

¹ Ver definição na Seção 3.4.1.

Frequência Acumulada: Corresponde à soma da frequência daquela classe às frequências de todas as classes anteriores a ela. Assim, tem-se a frequência absoluta acumulada (F_i) e a frequência relativa acumulada (F_r).

Exemplo: Uma empresa decide contratar um serviço de transporte para buscar seus 20 funcionários. Para isso, fez um levantamento do tempo, em minutos, referente a distância da residência de cada um até a empresa. Foram obtidos os seguintes dados: 30, 40, 22, 46, 32, 42, 36, 32, 54, 48, 46, 62, 26, 38, 46, 34, 56, 52, 50 e 56. Calcule a amplitude total e construa uma tabela com as frequências absolutas, relativas e acumuladas.

Inicialmente, coloca-se os dados em ROL. Logo:

22, 26, 30, 32, 32, 34, 36, 38, 40, 42, 46, 46, 46, 48, 50, 52, 54, 56, 56, 62.

A amplitude total será: $AT = 62 - 22 = 40$.

Para definir o número de classes, pode-se utilizar a Regra de Sturges. Assim, têm-se: $k = 1 + (3, 3) \log 20 = 5, 29$. Arredondando o valor encontrado, o número de classes será $k = 6$.

A altura de cada classe, será definida por $h_i = \frac{AT}{k - 1} = \frac{40}{6 - 1} = \frac{40}{5} = 8$.

Para a construção da tabela, deve-se iniciar a primeira classe com o menor valor e somando-se o valor da altura da classe (h_i) obtém-se os limites superiores. Para as demais classes, considera-se o limite inferior o valor calculado do limite superior da classe anterior e o processo é repetido, até encontrar todas as classes. Os demais cálculos, são obtidos utilizando os conceitos de ponto médio, frequências absolutas, relativas e acumulada. O resultado será:

Tabela 1 – Tempo, em minutos, necessário para transporte de funcionários

Tempo (min)	Ponto Médio	Frequência Absoluta (f_i)	Frequência Absoluta Acumulada (F_i)	Frequência Relativa (f_r)	Frequência Relativa Acumulada (F_r)
22 † 30	26	2	2	0,10	0,10
30 † 38	34	5	7	0,25	0,35
38 † 46	42	3	10	0,15	0,50
46 † 54	50	6	16	0,30	0,80
54 † 62	58	4	20	0,20	1,00
Total	-	20	-	1,00	-

Fonte: Elaborada pelo autor (2023).

Observe que a Tabela 1 possui 5 classes, cuja altura de cada uma é 8 minutos. As frequências relativas também podem ser exibidas em valor percentual.

3.2.1 Tabelas de Frequências para variáveis categorizadas

De acordo com Vieira (2018) os “dados categorizados devem ser apresentados em uma tabela de distribuição de frequências, porque esse tipo de tabela pode mostrar quantas unidades foram observadas em cada categoria da variável”.

Em uma tabela de frequências para variáveis categorizadas, observa-se que:

- a) Cada linha corresponde a um possível valor da variável;
- b) Através de um processo de contagem são obtidos os valores que constam na coluna de frequência da tabela. O resultado dessa contagem é a chamada frequência absoluta;
- c) A partir das frequências absolutas podem ser também calculadas frequências relativas, usualmente apresentadas sob a forma de percentuais.

Exemplo: Distribuição de frequência para dados categorizados.

Considere a distribuição de frequência da Tabela 2 que apresenta a quantidade de alunos de uma sala de aula, categorizados pela cor dos cabelos.

Tabela 2 – Distribuição de alunos de uma classe, segundo a cor dos cabelos

Cor de Cabelo	Frequência
Preto	16
Castanho	5
Loiro	8
Ruivo	1
Total	30

Fonte: Elaborada pelo autor (2021).

Nas tabelas de distribuição de frequências é usual fornecer, além das frequências absolutas, as frequências relativas.

Tabela 3 – Frequência relativa dos alunos de uma classe, segundo a cor dos cabelos

Cor de Cabelo	Frequência	Frequência Relativa
Preto	16	0,533
Castanho	5	0,167
Loiro	8	0,267
Ruivo	1	0,033
Total	30	1

Fonte: Elaborada pelo autor (2021).

A frequência relativa da Tabela 3, foi obtida utilizando a fórmula anteriormente definida. Pode-se representar as frequências relativas em porcentagens, conforme a Tabela 4.

Tabela 4 – Frequência percentual dos alunos de uma classe, segundo a cor dos cabelos

Cor de Cabelo	Frequência	Frequência Relativa
Preto	16	53,3%
Castanho	5	16,7%
Loiro	8	26,7%
Ruivo	1	3,3%
Total	30	100,0%

Fonte: Elaborada pelo autor (2021).

3.2.2 Tabelas de Frequências para variáveis numéricas (discretas)

Se os dados são discretos com um pequeno número de valores possíveis (por exemplo: número de filhos), a construção da tabela de frequência segue da mesma maneira que foi visto para variáveis categorizadas.

Desta forma, para organizar a distribuição de frequências:

- Conte quantas vezes cada dado se repete;
- Organize a tabela colocando, na primeira coluna, os dados numéricos em ordem natural (mas sem repetições) e, nas respectivas linhas, as frequências observadas.

Exemplo: Distribuição de frequência para dados discretos.

Considere a Tabela 5 que apresenta o número de irmãos ou irmãs que os alunos de uma sala de aula possuem. É preciso contar quantas vezes cada valor se repete.

Tabela 5 – Distribuição dos alunos de uma classe, segundo o número de irmãos ou irmãs

Número de Irmãos / Irmãs	Frequência	Frequência Relativa
0	7	23,33%
1	11	36,67%
2	8	26,67%
3	3	10,00%
4 ou mais	1	3,33%
Total	30	100,0%

Fonte: Elaborada pelo autor (2021).

3.2.3 Tabelas de Frequências para variáveis numéricas (contínuas)

Se os dados são variáveis numéricas contínuas, para a construção de uma tabela de frequência, deve-se inicialmente dividir o seu intervalo de variação, em subintervalos, chamados de classes, de preferência de mesma amplitude. O mesmo procedimento deve ser adotado para variáveis numéricas discretas em grande número de valores possíveis. Assim, ao adotar esse procedimento, o leitor terá uma visão rápida e global do fenômeno, e sua construção é semelhante ao caso de variáveis discretas.

Deve-se seguir os seguintes passos para a construção da tabela de frequência:

- a) Ache o valor mínimo e o valor máximo do conjunto de dados;
- b) Calcule a amplitude total, que é a diferença entre o valor máximo e o valor mínimo;
- c) Determine o número de classes, pode ser utilizado qualquer um dos critérios apresentados;
- d) Defina a amplitude da classe;
- e) Organize as classes. A primeira deve conter o valor mínimo e a última deve conter o valor máximo. Se o primeiro valor for 0, o limite superior será zero mais o valor da amplitude da classe. Para a próxima classe, o limite inferior será o limite superior da classe anterior, e o limite superior será obtido de maneira análoga a classe anterior. Esse processo será repetido até completar o número de classes.

Usualmente utiliza-se o intervalo de classe misto, ou seja, um dos limites pertence à classe, e o outro, não, representado pelo símbolo \vdash , onde o traço vertical a esquerda significa que o número pertence ao intervalo e o número à direita não pertence ao intervalo.

A Tabela 6 apresenta a distribuição da altura dos alunos de uma sala de aula.

Tabela 6 – Altura dos alunos de uma sala de aula

Altura (cm)	Ponto Médio	Frequência	Porcentagem
140 \vdash 148	144	2	6,67%
148 \vdash 156	152	6	20,00%
156 \vdash 164	160	8	26,67%
164 \vdash 172	168	10	33,33%
172 \vdash 180	176	4	13,33%
Total		30	100,0%

Fonte: Elaborada pelo autor (2021).

3.3 REPRESENTAÇÃO GRÁFICA

Os gráficos constituem uma forma clara e objetiva de apresentar dados estatísticos. Segundo Chavante e Prestes (2016):

O objetivo da representação gráfica é apresentar um conjunto de dados de modo resumido, principalmente para os resultados de uma pesquisa quando se deseja facilitar sua leitura e compreensão a fim de chegar a conclusões a respeito da evolução das variáveis ou de como elas se relacionam. A escolha da representação gráfica mais apropriada à situação depende de vários fatores, mas a simplicidade, clareza e veracidade das informações devem ser consideradas na elaboração do gráfico.

Muito utilizado em jornais, revistas e na televisão por apresentar as informações na forma de ilustrações, os gráficos estatísticos facilitam a compreensão (VIEIRA, 2018). De acordo com Morettin e Bussab (2010, *apud* Chambers *et al.*, 1983), os gráficos são utilizados para diversos fins:

- a) Buscar padrões ou relações;
- b) Confirmar (ou não) certas expectativas que se tinham sobre os dados;
- c) Descobrir novos fenômenos;
- d) Confirmar (ou não) suposições feitas sobre os procedimentos estatísticos usados; e
- e) Apresentar resultados de modo mais rápido e fácil.

Cada tipo de gráfico tem indicação específica e, segundo as normas brasileiras (VIEIRA, 2018):

- a) Todo gráfico deve apresentar título e escala;
- b) O título deve ser colocado acima da ilustração;
- c) As escalas devem crescer da esquerda para a direita e de baixo para cima;
- d) As legendas explicativas devem ser colocadas, de preferência, à direita da figura;
- e) Os gráficos devem ser numerados, na ordem em que são citados no texto.

3.3.1 Representação gráfica de variáveis categorizadas

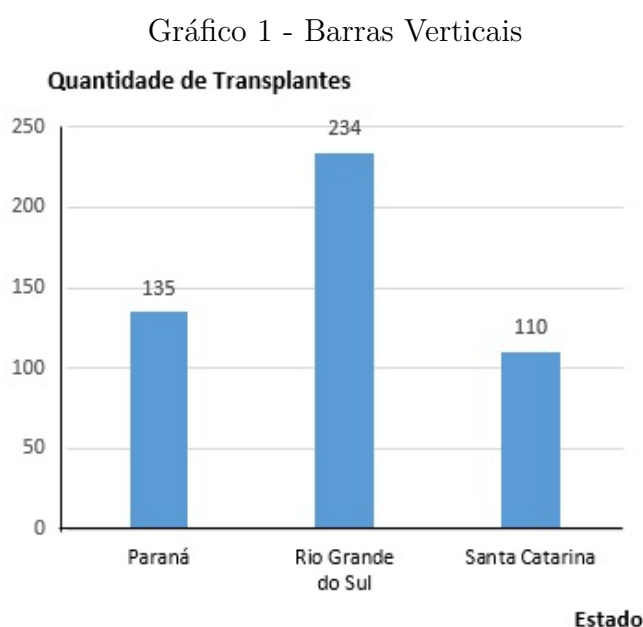
O gráfico de barras e o gráfico de setores são os gráficos mais comuns para representar as variáveis categorizadas.

3.3.1.1 Gráfico de barras

De acordo com Chavante e Prestes (2016) o gráfico de barras, também denominado gráfico de colunas, consiste em retângulos com a mesma largura e altura proporcional aos valores das variáveis, representado em um par de eixos ortogonais. Geralmente, esse tipo de gráfico é utilizado para comparar informações.

O gráfico de barras é um excelente recurso para apresentar dados qualitativos coletados em determinado momento, como as respostas de questionários (VIEIRA, 2018).

Exemplo: O Gráfico 1 apresenta a quantidade de transplantes de rim em pessoas falecidas entre janeiro e junho de 2018, na região Sul do Brasil.



Fonte: Chavante e Prestes (2016, p.204).

Nota: Ilustração adaptada.

O gráfico de barras também pode apresentar as informações coletadas com as barras na posição horizontal. As barras têm larguras iguais e seus comprimentos correspondem à frequência (absoluta ou relativa) dos valores observados. Este tipo de gráfico facilita a leitura, principalmente nos casos em que as categorias têm nomes extensos.

Exemplo: O Gráfico 2 representa o principal motivo da falência, segundo proprietários de 2.000 pequenas empresas criadas no Brasil, que faliram antes de completar dois anos de existência.

Gráfico 2 - Barras Horizontais



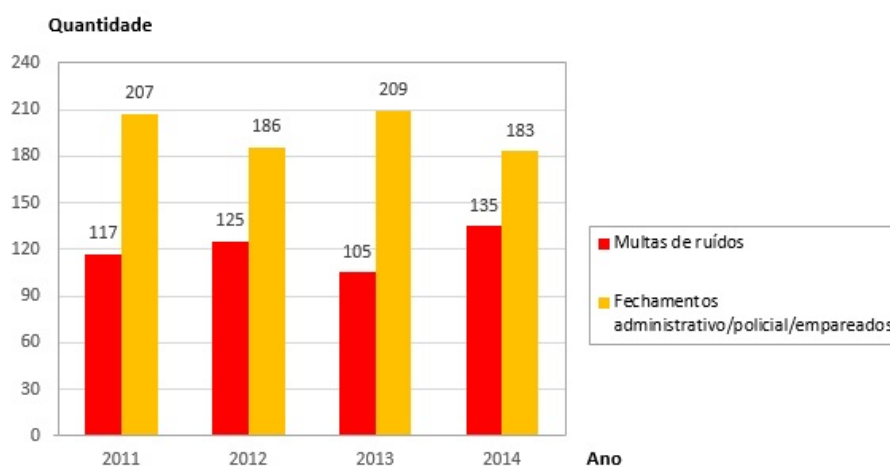
Fonte: Vieira (2018, p.23).

Nota: Ilustração adaptada.

Outra maneira de comparar dados de diferentes categorias no mesmo gráfico, é a utilização de gráficos de barras múltiplas cuja construção segue os mesmos passos do gráfico de barras.

Exemplo: O Gráfico 3 apresenta a fiscalização realizada pela Prefeitura de São Paulo em bares, boates e outros ambientes fechados de modo a combater a poluição sonora nesses ambientes.

Gráfico 3 - Barras Múltiplas



Fonte: Chavante e Prestes (2016, p.205).

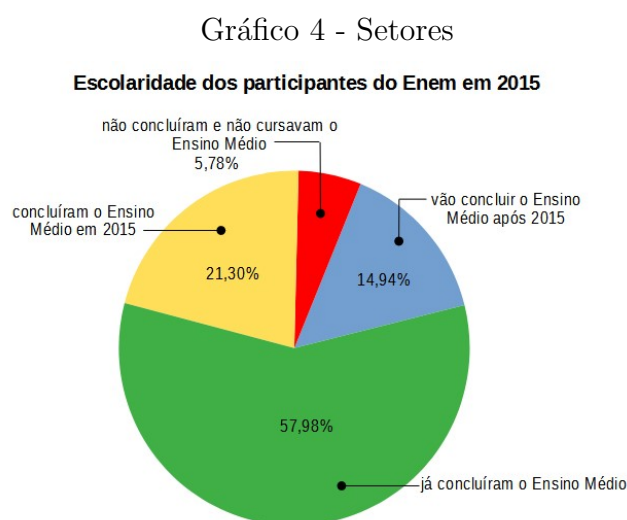
Nota: Ilustração adaptada.

3.3.1.2 Gráfico de setores

Os gráficos de setores são especialmente utilizados para mostrar como se divide o todo. Para Chavante e Prestes (2016), o gráfico de setores (ou popularmente conhecido

como gráfico de pizza) é representado em um círculo por meio de setores circulares. Em sua construção, o círculo é dividido em setores cujas áreas são proporcionais aos valores das variáveis. Cada região dessa representa uma parte do todo e geralmente são indicadas com porcentagem.

Exemplo: O Gráfico 4 apresenta a participação dos alunos no Enem (Exame Nacional do Ensino Médio) em 2015.



Fonte: Balestri (2016, p.120).

Nota: Ilustração adaptada.

O círculo representa o total de alunos (100%) e cada setor representa uma porcentagem desse círculo.

3.3.2 Representação gráfica de variáveis numéricas

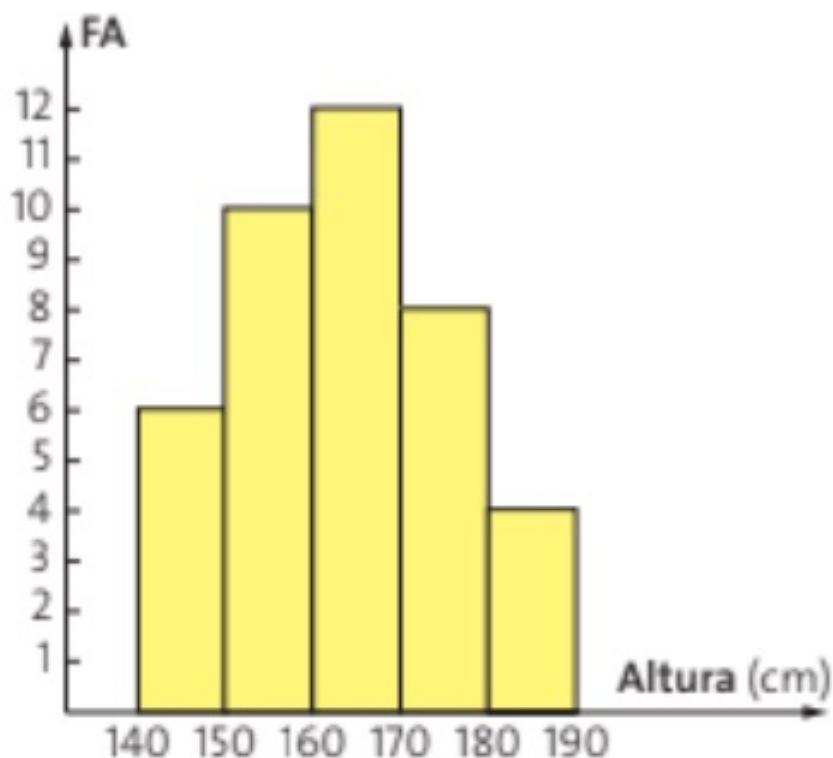
O gráfico mais adequado para representar as variáveis numéricas é o histograma por ser possível condensar os dados. Entretanto, o gráfico de linhas, diagrama de pontos e diagrama de ramo e folhas tem suas particularidades na apresentação dos dados.

3.3.2.1 Histograma

Similarmente ao gráfico de barras, o histograma é representado em um par de eixos ordenados (PINHEIRO *ET AL.*, 2009). Segundo Silveira (2015) o histograma é um gráfico utilizado para representar uma distribuição de frequências em que a base de cada retângulo representa um intervalo de classe e sua altura a frequência da classe correspondente. Se os intervalos estiverem no eixo horizontal e as frequências no eixo vertical, é classificado como um histograma vertical. Se ocorrer o inverso, será denominado histograma horizontal.

Exemplo: O Gráfico 5 representa um histograma com a altura dos alunos de uma classe, agrupadas em intervalos correspondentes a frequência absoluta.

Gráfico 5 - Histograma



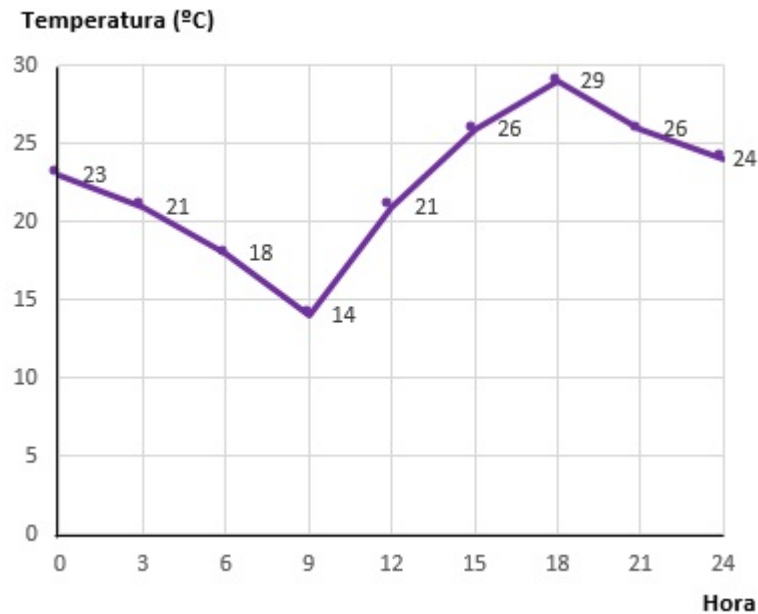
Fonte: Dante (2016, p.43).

3.3.2.2 Gráfico de linhas

Também denominado como gráfico de segmentos, o gráfico de linhas é representado em um par de eixos ortogonais por meio de linhas que indicam uma tendência entre dois pontos consecutivos, apresentando o crescimento, decréscimo ou estabilidade de uma variável ao longo do tempo. Este tipo de gráfico é utilizado para representar a variação de uma grandeza em ordem cronológica, assim, geralmente o eixo horizontal representa o período cronológico e o eixo vertical os valores observados (CHAVANTE E PRESTES, 2016).

Exemplo: O Gráfico 6 representa a temperatura registrada em Dourados (MS) no dia 29/07/2015.

Gráfico 6 - Linhas



Fonte: Chavante e Prestes (2016, p.206).

Nota: Ilustração adaptada.

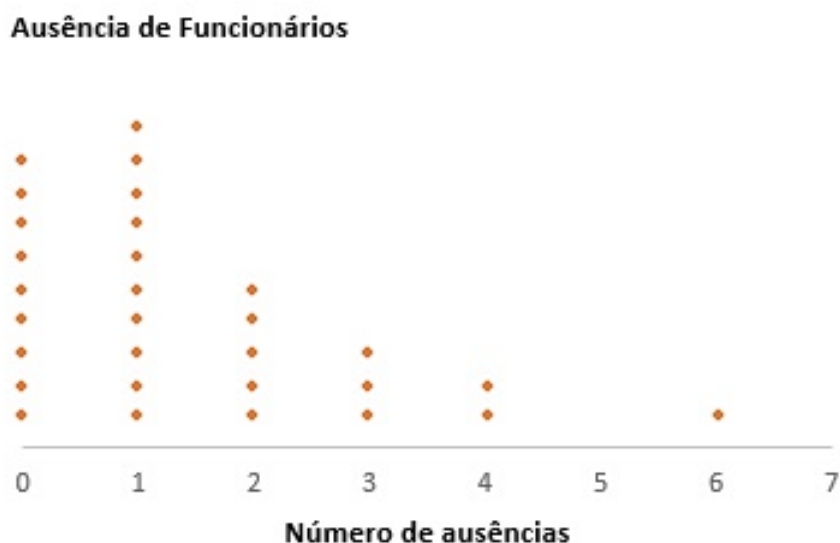
3.3.2.3 Diagrama de pontos

Quando os dados são discretos, as tabelas de distribuição de frequências apresentam valores numéricos em ordem natural, que podem ser representados pelo diagrama de pontos, onde os pontos assumem os valores assumidos pelas variáveis. Os passos para a construção do diagrama de pontos, segundo Vieira (2018) são:

- Desenhe o eixo das abscissas (horizontal) e escreva no eixo os valores assumidos pela variável em estudo, identificando a escala;
- Desenhe o eixo das ordenadas (vertical) e marque uma escala para as frequências. Este eixo pode ser apagado posteriormente;
- Faça um ponto para representar cada dado;
- Escreva a legenda e coloque título na figura.

Exemplo: A ausência de 30 funcionários está representada no diagrama do Gráfico 7:

Gráfico 7 - Diagrama de Pontos



Fonte: Vieira (2018, p.30).

Nota: Ilustração adaptada.

3.3.2.4 Diagrama de ramo e folhas

De acordo com Pinheiro (2009), outro tipo de representação gráfica para se mostrar dados referentes a variáveis numéricas é o diagrama de ramo e folhas. Vieira (2018) complementa que esse tipo de representação deve ser utilizada quando não se tem quantidade muito grande de dados (entre 25 e 100). O diagrama é constituído de duas partes: o ramo e as folhas. Os primeiros dígitos dos dados numéricos são os ramos e os últimos dígitos são as folhas. Morettin e Bussab (2010) relata que uma das vantagens do diagrama de ramo e folhas sobre o histograma é a baixa perda de informação sobre os dados em si.

Na construção do diagrama, Vieira (2018) propõe os seguintes passos:

- a) Organize os dígitos correspondentes às dezenas em ordem crescente, em coluna. Esses dígitos formam o ramo;
- b) Faça um traço vertical, à direita desses dígitos;
- c) Escreva os dígitos das unidades, que serão as folhas, à direita do traço, na linha correspondente à sua dezena;
- d) Organize os dígitos das unidades, que são as folhas, em ordem crescente;
- e) Coloque título na figura.

Exemplo: A Tabela 7 representa a idade de 25 funcionários de determinada escola.

Tabela 7 – Idades dos funcionários de uma escola

26	22	32	30	33
37	38	25	22	32
40	47	23	28	60
19	23	19	30	30
18	52	43	31	20

Fonte: Elaborada pelo autor (2021).

Para construir o diagrama, primeiro se escreve o dígito das dezenas em ordem crescente e se desenha um traço vertical, conforme Figura 2:

Figura 2 - Construção do Diagrama de Ramo e Folhas

1	
2	
3	
4	
5	
6	

Fonte: Elaborada pelo autor (2021).

Em seguida, os algarismos que representam as unidades são colocados em frente a cada uma da dezena respectiva, de forma ordenada. Assim, o Gráfico 8, representa o Diagrama de Ramo e Folhas construído.

Gráfico 8 - Diagrama de Ramo e Folhas

1		8	9	9						
2		0	2	2	3	3	5	6	8	
3		0	0	0	1	2	2	3	7	8
4		0	3	7						
5		2								
6		0								

Fonte: Elaborada pelo autor (2021).

O uso do diagrama de ramos e folhas não se restringe a dados discretos ou valores pequenos. Para dados que não são desses tipos, pode-se dividir os números em duas partes, estipular uma escala, etc.

O diagrama de ramo e folhas facilita a visualização da distribuição, uma vez que apresenta os dados em ordem crescente.

3.4 MEDIDAS DE TENDÊNCIA CENTRAL

A apresentação de dados por meio de tabelas de frequência e gráficos, constituem uma importante fonte de informações sobre o comportamento de uma variável mais do que a própria tabela original de dados (MORETTIN E BUSSAB, 2010). Entretanto, conforme Chavante e Prestes (2016), em alguns casos é necessário a apresentação desses dados de forma mais resumida, ou seja, sintetizada, para expressar determinada característica da população em estudo. Isso pode ser feito por meio das medidas de tendência central, que representam os dados de comportamento em torno do centro do qual os dados se distribuem.

Para descrever as medidas de tendência central, será conceituado a média aritmética, a média ponderada, a mediana e a moda.

3.4.1 Média aritmética

A média aritmética (ou simplesmente média) é a medida de tendência central mais conhecida e utilizada para sintetizar as informações de determinada característica em um conjunto de dados (VIEIRA, 2018).

A média de um conjunto de dados é obtida somando-se todos os dados desse conjunto e dividindo esse valor pelo número de dados do conjunto.

Sejam $x_1, x_2, x_3, \dots, x_i, \dots, x_n$ os n valores da variável x . O subscrito i indica a posição do valor na sequência, sendo $i = \{1, 2, \dots, n\}$. Logo, x_1 representa o primeiro valor observado, x_2 representa o segundo valor e assim por diante; x_i é o i -ésimo valor no conjunto de n valores.

A média aritmética de x é representada por \bar{x} (leia-se x-traço ou x-barras) e definida por:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}.$$

A letra grega maiúscula Σ (sigma) é usada para indicar uma soma. A indicação $\sum_{i=1}^n x_i$ significa o somatório dos valores de x_i , para i variando de 1 até n .

Logo:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_i + \dots + x_n}{n}.$$

Exemplo: Na 38ª rodada do Campeonato Brasileiro de Futebol de 2015, foram realizados 10 jogos, cuja quantidade de gols por partida está apresentada na Tabela 8.

Tabela 8 – Número de gols na 38ª rodada

38ª Rodada do Campeonato Brasileiro de 2015										
Partida	1ª	2ª	3ª	4ª	5ª	6ª	7ª	8ª	9ª	10ª
Número de Gols	2	6	1	3	2	3	1	0	1	2

Fonte: Leonardo (2016).

Para calcular a média de gols dessa rodada, deve-se somar o número de gols de cada jogo e dividir o resultado obtido pelo número total de partidas.

$$\bar{x} = \frac{2 + 6 + 1 + 3 + 2 + 3 + 1 + 0 + 1 + 2}{10} = \frac{21}{10} = 2,1.$$

Logo, a média de gols por partida nessa rodada foi de 2,1 gols.

A média de gols serve de parâmetro para, por exemplo, avaliar o desempenho dos times nos diferentes campeonatos que participaram.

Outra observação importante é que a média obtida não coincidiu com nenhum dos números de gols por partida, pois a média não é necessariamente igual a um dos valores da variável.

3.4.2 Média aritmética ponderada

Para o cálculo da média aritmética, todas as observações são somadas individualmente e o resultado dividido pelo número de observações. Porém, há casos em que as observações se repetem e então é utilizado o cálculo da média aritmética ponderada.

Geralmente, essas observações são apresentadas em tabelas de frequências, com dados agrupados ou não. Segundo Vieira (2018): “A média de dados discretos agrupados em uma tabela de distribuição de frequências é dada pela soma dos produtos dos valores da variável pelas respectivas frequências, dividida pela soma das frequências”.

A fórmula da média aritmética ponderada é definida por:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i},$$

sendo que $\sum x_i f_i$ representa a soma de todos os valores da variável multiplicada pela sua respectiva frequência e $\sum f_i$ o número total de observações.

Exemplo: A Tabela 9 mostra o número de livros lidos pelos alunos de uma classe de aula.

Tabela 9 – Número de livros lidos em uma sala de aula

Número de livros lidos	Frequência
0	8
1	10
2	5
3	4
4	2
5	1

Fonte: Elaborada pelo autor (2021).

Para obter a média, deve-se multiplicar o valor de cada número de livros lidos pela sua respectiva frequência e dividir pela soma total de frequências. Ou seja:

$$\bar{x} = \frac{0 \times 8 + 1 \times 10 + 2 \times 5 + 3 \times 4 + 4 \times 2 + 5 \times 1}{8 + 10 + 5 + 4 + 2 + 1} = \frac{45}{30} = 1,5.$$

Logo, a média é de 1,5 livros lidos por aluno.

Pode-se facilitar o entendimento acrescentando uma coluna à tabela original com o cálculo de $x_i f_i$. Assim, utilizando uma linha total, aplica-se a fórmula apresentada, conforme a Tabela 10.

Tabela 10 – Número de livros lidos com cálculos auxiliares

Número de livros lidos (x_i)	Frequência (f_i)	Produto ($x_i f_i$)
0	8	0
1	10	10
2	5	10
3	4	12
4	2	8
5	1	5
Total	30	45

Fonte: Elaborada pelo autor (2021).

Utilizando a fórmula apresentada, tem-se:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{45}{30} = 1,5.$$

Quando são dados contínuos agrupados em classes, e apresentados em tabelas de frequências o procedimento segue o que foi descrito anteriormente, porém, antes de calcular a média deve-se determinar o ponto médio de cada classe. Esse valor será utilizado como parâmetro para o cálculo da média ponderada.

Exemplo: Calcular a média de altura dos alunos de uma sala de aula, conforme a Tabela 11.

Tabela 11 – Altura dos alunos de uma sala de aula

Altura (cm)	Frequência (f_i)
140 † 148	2
148 † 156	6
156 † 164	8
164 † 172	10
172 † 180	4
Total	30

Fonte: Elaborada pelo autor (2021).

Inicialmente deve-se determinar o ponto médio de cada intervalo de classe. Esse valor servirá de referência para os cálculos. Em seguida, uma coluna auxiliar é acrescentada para facilitar os cálculos. Observe a Tabela 12 modificada.

Tabela 12 – Altura dos alunos de uma sala de aula com cálculos auxiliares

Altura (cm)	Ponto Médio (PM)	Frequência (f_i)	Produto ($x_i f_i$)
140 † 148	144	2	288
148 † 156	152	6	912
156 † 164	160	8	1280
164 † 172	168	10	1680
172 † 180	176	4	704
Total		30	4864

Fonte: Elaborada pelo autor (2021).

Utilizando a fórmula, tem-se:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{4864}{30} = 162,1\bar{3}.$$

Portanto, a altura média dos alunos dessa turma é de 162,1 $\bar{3}$ cm.

3.4.3 Mediana

Segundo Silveira (2015) a mediana pode ser definida como o valor que ocupa a posição central de um conjunto de dados ordenados (crescente ou decrescente). Leonardo (2016) complementa que a mediana deve ser utilizada quando os valores da amostra são discrepantes, ou seja, muito maiores ou menores do que os demais. Neste caso é a medida de posição mais adequada para descrever a tendência central dos dados.

Pode ser representada como Me ou Md . Para seu estudo deve-se considerar se os dados estão não agrupados ou agrupados (sem ou com classe).

3.4.3.1 Dados não agrupados

Para dados não agrupados, sendo n o número de termos da distribuição, deve-se considerar que (LEONARDO, 2016):

- a) Quando há um grupo de valores em número ímpar de dados, a mediana é o termo central da distribuição e, nesse caso, pertence ao grupo de valores observado;
- b) Quando há um grupo de valores em número par de dados, a mediana é a média aritmética dos termos centrais e, nesse caso, pode não pertencer ao grupo de valores observado.

Para verificar a posição dos termos centrais, tem-se:

- a) Se n é ímpar, a posição do termo central é dada por $\left(\frac{n+1}{2}\right)$;
- b) Se n é par, as posições dos dois termos centrais são dadas por $\left(\frac{n}{2}\right)$ e $\left(\frac{n}{2} + 1\right)$.

Exemplos:

a) Sete crianças participaram de uma corrida e tiveram os seguintes tempos (em segundos): 180, 227, 197, 205, 195, 215 e 212.

Para calcular a mediana, inicialmente deve-se ordenar os dados, ou seja, escrever o rol dessas observações. Logo, obtêm-se:

$$180 \quad 195 \quad 197 \quad 205 \quad 212 \quad 215 \quad 227.$$

Como o conjunto de elementos possui uma quantidade ímpar de termos, 7 elementos, a mediana será o termo que está na posição central. Assim,

$$Md = x_{\left(\frac{n+1}{2}\right)} = x_{\left(\frac{8}{2}\right)} = x_4.$$

Portanto, o quarto termo ocupa a posição central, logo a mediana é 205.

b) Rafaela vende doces na faculdade e durante 8 dias vendeu a seguinte quantidade: 47, 52, 55, 49, 51, 45, 55 e 46.

O rol correspondente a esses termos é:

$$45 \quad 46 \quad 47 \quad 49 \quad 51 \quad 52 \quad 55 \quad 55.$$

Como o conjunto de dados tem uma quantidade par de termos, a mediana será a média aritmética entre os termos de ordem $\left(\frac{n}{2}\right)$ e $\left(\frac{n}{2} + 1\right)$.

$$Md = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} = \frac{x_4 + x_5}{2}.$$

Os termos centrais ocupam a 4ª e 5ª posição, que correspondem as vendas de 49 e 51, respectivamente. Calculando a média aritmética entre esses dois termos.

$$Md = \frac{49 + 51}{2} = \frac{100}{2} = 50.$$

Logo, a mediana será 50 doces.

3.4.3.2 Dados agrupados

De acordo com Crespo (2002) quando os dados estão agrupados em uma distribuição de frequência, o cálculo da mediana segue de forma similar aos dados não agrupados, implicando, entretanto, a prévia determinação da frequência acumulada.

Deve-se determinar o valor da posição do elemento que divide a distribuição em dois grupos que contenha o mesmo número de elementos. Essa posição será dada pelo quociente do número total de observações ($\sum f_i$) por 2, ou seja:

$$\frac{\sum f_i}{2}.$$

3.4.3.2.1 Dados agrupados sem intervalo de classe

Neste caso basta identificar a frequência acumulada imediatamente superior à posição encontrada. A mediana será aquele valor da variável correspondente a tal frequência acumulada.

Exemplo: Dada a distribuição de frequência do número de livros lidos em uma sala de aula com 30 alunos, conforme Tabela 13, calcule a mediana.

Tabela 13 – Número de livros lidos com frequência acumulada

Número de livros lidos (x_i)	Frequência (f_i)	Frequência Acumulada (F_i)
0	8	8
1	10	18
2	5	23
3	4	27
4	2	29
5	1	30
Total	30	

Fonte: Elaborada pelo autor (2021).

Sendo:

$$\frac{\sum f_i}{2} = \frac{30}{2} = 15.$$

A menor frequência acumulada que supera esse valor é 18, que corresponde ao valor 1 da variável, sendo este o valor mediano. Logo $Md = 1$ livro lido.

Observação: Em caso de existir uma frequência acumulada (F_i) tal que $F_i = \frac{\sum f_i}{2}$, a mediana será dada pela média aritmética entre o valor da variável correspondente a esta frequência acumulada e a seguinte, ou seja,

$$Md = \frac{x_i + x_{i+1}}{2}.$$

Exemplo: Observe a Tabela 14 que representa a quantidade de bolinhas de gudes que um grupo de meninos possui.

Tabela 14 – Distribuição de bolinhas de gudes

Número de bolinhas de gude (x_i)	Frequência (f_i)	Frequência Acumulada (F_i)
10	2	2
11	4	6
12	9	15
14	7	22
18	5	27
22	3	30
Total	30	

Fonte: Elaborada pelo autor (2021).

Para o cálculo da mediana, tem-se:

$$\frac{\sum f_i}{2} = \frac{30}{2} = 15.$$

Como $F_3 = 15$, a mediana será dada pela média de x_3 e x_4 . Logo,

$$Md = \frac{x_3 + x_4}{2} = \frac{12 + 14}{2} = \frac{26}{2} = 13.$$

Assim, a mediana para esse conjunto de dados é 13.

3.4.3.2.2 Dados agrupados com intervalo de classe

Neste caso o problema consiste em determinar o ponto do intervalo em que está compreendido a mediana. Segundo Crespo (2002), os seguintes passos podem ser utilizados:

- Determinar as frequências acumuladas;

- b) Calcular $\frac{\sum f_i}{2}$;
- c) Marcar a classe correspondente à frequência acumulada imediatamente superior à $\frac{\sum f_i}{2}$ – classe mediana – e, em seguida, empregar a fórmula:

$$Md = l^* + \frac{\left[\frac{\sum f_i}{2} - F(\text{ant}) \right] \times h}{f^*}.$$

Na qual:

l^* é o limite inferior da classe mediana;

$F(\text{ant})$ é a frequência acumulada da classe anterior à classe mediana;

f^* é a frequência simples da classe mediana;

h^* é a amplitude do intervalo da classe mediana.

Exemplo: Considere a Tabela 15 de distribuição de frequências acumuladas:

Tabela 15 – Distribuição de frequência acumulada da altura dos alunos de uma sala de aula

i	Altura (cm)	Frequência (f_i)	Frequência Acumulada (F_i)
1	140 – 148	2	2
2	148 – 156	6	8
3	156 – 164	8	16
4	164 – 172	10	26
5	172 – 180	4	30
Total		30	

Fonte: Elaborada pelo autor (2021).

Assim:

$$\frac{\sum f_i}{2} = \frac{30}{2} = 15.$$

A classe mediana é a de ordem 3. Então:

$$l^* = 156; F(\text{ant}) = 8; f^* = 8 \text{ e } h^* = 8.$$

Substituindo esses valores na fórmula, obtém-se o valor da mediana.

$$Md = 156 + \frac{[15 - 8] \times 8}{8} = 163.$$

Logo o valor mediano é 163 cm.

Observação: alguns autores consideram a mediana o ponto médio da classe mediana, ao invés de usar o método de interpolação descrito por Crespo (2002). No exemplo anterior, utilizando esse critério, a mediana encontrada seria igual a 160 cm.

3.4.4 Moda

Segundo Leonardo (2016) a moda são os valores que aparecem com maior frequência em um conjunto de dados. É indicada por *Mo*. Para Vieira (2018) a moda é uma medida de posição muito informativa quando se tem um conjunto de dados muito grande, porém se o conjunto de dados for relativamente pequeno (20 ou 30 observações) a moda, em geral, não tem sentido prático, e nesses casos, a média e a mediana fornecem melhor descrição da tendência central dos dados.

A moda também pode ser utilizada para descrever dados categorizados. Segundo Vieira (2018), nesse caso, a moda é a categoria que ocorre com maior frequência, ou seja, a que concentra maior quantidade de dados.

Sobre um conjunto de dados, Chavante e Prestes (2016) descreve:

Caso um conjunto de dados não apresente moda, ou seja, não contenha valor algum que se repete, dizemos que esse conjunto é amodal. Mas caso haja um conjunto de dados com mais de uma moda, ele é bimodal se tiver duas modas, trimodal se tiver três modas, quadrimodal se tiver quatro modas, e assim sucessivamente.

Um conjunto com três ou mais modas também pode ser chamado de polimodal.

3.4.4.1 Dados não agrupados

Quando os dados estão dispostos um conjunto de valores não agrupados, a moda é facilmente identificada. De acordo com a definição, basta procurar o valor que mais se repete.

Exemplos:

a) As idades de um grupo de trabalho são 18, 18, 19, 19, 19, 21, 21, 23 e 24. Logo, 19 é a moda, pois é a idade que aparece mais vezes nesse grupo.

b) O conceito de notas de uma turma de 20 alunos de graduação em matemática foram:

C, A, E, B, D, A, C, B, D, B, B, C, B, A, D, E, A, C, B, A

Nessa turma, 6 alunos tem o conceito B, 5 alunos, conceito A, 4 alunos, conceito C, 3 alunos, conceito D, e 2 alunos conceito E. Portanto, o conceito B é a moda dessa turma, pois tem a maior frequência.

3.4.4.2 Dados agrupados

Quando os dados estão dispostos em um conjunto de valores agrupados, deve-se utilizar procedimentos distintos para os dados agrupados sem intervalo de classe e para os dados agrupados com intervalo de classe.

3.4.4.2.1 Dados agrupados sem intervalo de classe

Para Crespo (2002) uma vez que os dados estão agrupados, é possível determinar a moda bastando identificar o valor da variável com maior frequência.

Exemplo: Na distribuição de frequência da Tabela 13 (p. 48), observa-se que a frequência que apresenta maior valor é a referente a variável 1, portanto essa é a moda da distribuição.

$$Mo = 1.$$

3.4.4.2.2 Dados agrupados com intervalo de classe

A classe que apresenta a maior frequência é determinada de classe modal. A moda, nesse caso, está compreendida entre os limites da classe modal. Crespo (2002) relata que o método mais simples para obtenção da moda é determinar o ponto médio da classe modal. Sendo l_i o extremo inferior, e L_i o extremo superior, então:

$$Mo = \frac{l_i + L_i}{2}.$$

Exemplo: Calcule a moda para a distribuição dada pela Tabela 16 :

Tabela 16 – Distribuição de frequência da altura dos alunos de uma sala de aula

i	Altura (cm)	Frequência (f_i)
1	140 † 148	2
2	148 † 156	6
3	156 † 164	8
4	164 † 172	10
5	172 † 180	4
Total		30

Fonte: Elaborada pelo autor (2021).

A classe modal é a $i = 4$, com $l_4 = 164$ e $L_4 = 172$. Portanto:

$$Mo = \frac{(164 + 172)}{2} = 168.$$

Logo, $Mo = 168$ cm.

Observação: Há outros métodos mais elaborados para cálculo da moda, como por exemplo o que faz uso da fórmula de Czuber:

$$Mo = l^* + \frac{D_1}{D_1 + D_2} \times h^*.$$

Na qual:

l^* é o limite inferior da classe modal;

h^* é a amplitude da classe modal;

$$D_1 = f^* - f(\text{ant});$$

$$D_2 = f^* - f(\text{post}),$$

sendo:

f^* a frequência simples da classe modal;

$f(\text{ant})$ a frequência simples da classe anterior à classe modal;

$f(\text{post})$ a frequência simples da classe posterior à classe modal.

Exemplo: Para a distribuição de frequência da Tabela 16 (p. 52), tem-se:

$$D_1 = 10 - 8 = 2 \text{ e } D_2 = 10 - 4 = 6.$$

Então:

$$Mo = 164 + \frac{2}{(2 + 6)} \times 8 = 166.$$

Logo, $Mo = 166$ cm.

3.5 MEDIDAS DE VARIABILIDADE

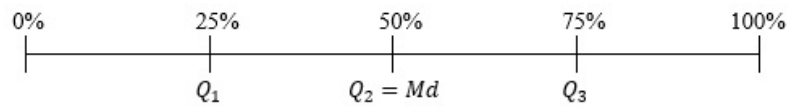
As medidas de tendência central não são suficientes para descrever um conjunto de dados, uma vez que indicam onde ocorre uma tendência de concentração para esse conjunto de dados. Essas medidas não revelam o quanto os dados estão dispersos em relação ao centro. Para descrever um conjunto de dados, segundo Vieira (2018), além das medidas de tendência central, deve ser utilizado medidas que indicam o quanto afastado ou próximo os valores do conjunto de dados estão em relação aos valores centrais.

As medidas estatísticas que descrevem esse comportamento em um conjunto de dados são chamadas de medidas de variabilidade ou de dispersão. Além da amplitude já descrita anteriormente, outras medidas de variabilidade são: quartis, percentis, variância, desvio padrão e coeficiente de variação.

3.5.1 Quartis

A mediana divide um conjunto de dados ordenados em dois subconjuntos com o mesmo número de dados. Os quartis dividem um conjunto de dados ordenados em quatro partes iguais. Observe a Figura 3:

Figura 3 - Quartis de uma distribuição



Fonte: Elaborada pelo autor (2021).

Os quartis são, portanto, três:

- O primeiro quartil (Q_1): 25% dos dados abaixo dele e 75% dos dados acima dele;
- O segundo quartil (Q_2): é a própria mediana, sendo 50% dos dados abaixo e acima dele;
- O terceiro quartil (Q_3): 75% dos dados abaixo dele e 25% dos dados acima dele.

Os métodos utilizados para encontrar os quartis têm pequenas diferenças. Para um conjunto de dados não agrupados, Vieira (2018) propõe os seguintes passos:

- Organize os dados em ordem crescente, isto é, do menor para o maior. Então encontre a mediana (que é também, o segundo quartil) e marque esse valor;
- Encontre o primeiro quartil, da seguinte forma: tome o conjunto de dados à esquerda da mediana; o primeiro quartil é a mediana do novo conjunto de dados;
- Encontre o terceiro quartil, da seguinte forma: tome o conjunto de dados à direita da mediana; o terceiro quartil é a mediana do novo conjunto de dados.

Exemplo: A altura, em centímetros, de 18 alunos de uma classe estão aqui relacionados:

159 159 159 163 163 163 168 168 168
168 168 168 170 170 174 174 178 178

Como os dados já estão ordenados, é necessário calcular inicialmente o valor da mediana. Como o número de dados é par deve-se encontrar os termos centrais da sequência, ou seja o 9º e 10º termo. Como esses termos são iguais, o ponto médio é 168. Portanto a altura mediana será 168 cm.

Para obter o primeiro quartil, separe os dados à esquerda do valor da mediana. Assim:

159 159 159 163 163 163 168 168 168

O conjunto à esquerda contém um número ímpar de dados. Logo o primeiro quartil será o termo central, 163 cm.

Para obter o terceiro quartil, separe os dados à direita do valor da mediana. Assim:

168 168 168 170 170 174 174 178 178

De forma análoga, o terceiro quartil será o termo central desse conjunto de dados, ou seja, 170 cm.

Quando os dados estão agrupados, Crespo (2002) indica que para determinar os quartis, é utilizada a mesma técnica para o cálculo da mediana, substituindo o termo $\frac{\sum f_i}{2}$, por $k \frac{\sum f_i}{4}$, sendo k o número de ordem do quartil.

Portanto, para o primeiro quartil, a fórmula será:

$$Q_1 = l^* + \frac{\left[\frac{\sum f_i}{4} - F(\text{ant}) \right] \times h}{f^*}.$$

E para o terceiro quartil, a seguinte fórmula:

$$Q_3 = l^* + \frac{\left[3 \frac{\sum f_i}{4} - F(\text{ant}) \right] \times h}{f^*}.$$

Exemplo: Utilizando os dados da Tabela 16 (p. 52), o cálculo do primeiro e do terceiro quartil será:

Primeiro Quartil:

$$\frac{\sum f_i}{4} = \frac{30}{4} = 7,5.$$

$$Q_1 = 148 + \frac{[7,5 - 2] \times 8}{6} = 155,33.$$

Terceiro Quartil:

$$3 \frac{\sum f_i}{4} = 3 \frac{30}{4} = 22,5.$$

$$Q_3 = 164 + \frac{[22,5 - 16] \times 8}{10} = 169,2.$$

3.5.1.1 Diagrama de caixa (box plot)

Segundo Pinheiro (2009), o diagrama de caixa (ou *box plot*) é uma representação gráfica, utilizada para sintetizar em uma mesma figura várias informações relativas à distribuição de uma determinada variável numérica. Para Morettin e Bussab (2010):

O *box plot* dá uma ideia da posição, dispersão, assimetria, caudas e dados discrepantes. A posição central é dada pela mediana e a dispersão por d_q . As posições relativas de Q_1 , Q_2 , Q_3 dão uma noção da assimetria da distribuição. Os comprimentos das caudas são dados pelas linhas que vão do retângulo aos valores remotos e pelos valores atípicos.

A dispersão d_q é chamada de distância interquartilica, e representa a distância entre o primeiro e o terceiro quartil, ou seja:

$$d_q = Q_3 - Q_1.$$

Os valores atípicos (também chamados de *outliers*) são as observações destoantes das demais, ou seja, que estão acima ou abaixo dos limites estabelecidos. A partir do retângulo, para cima, segue uma linha até o ponto mais remoto que não exceda o limite superior, que é dado por $LS = q_3 + (1,5)d_q$. De modo análogo, da parte inferior do retângulo, para baixo, segue uma linha até o ponto mais remoto que não seja menor que o limite inferior, dado por $LI = q_1 - (1,5)d_q$.

Vieira (2018) propõe os seguintes passos para a construção de um diagrama de caixa:

- a) Calcular o mínimo e máximo, o primeiro e terceiro quartis e a mediana (segundo quartis);
- b) Criar um segmento de reta em posição vertical, para representar a amplitude dos dados;
- c) Marcar, nesse segmento, o primeiro e o terceiro quartis e a mediana;
- d) Criar uma caixa retangular (box) de maneira que o lado superior e o inferior passem exatamente sobre os pontos que marcam o primeiro e o terceiro quartis;
- e) Fazer um ponto para representar a mediana (obedecendo a escala).

A altura do retângulo do diagrama de caixa é dada pela distância interquartilica, e contém cerca de 50% dos dados centrais, que são chamados de valores adjacentes.

Exemplo: Comparação de dois diagramas de caixa.

Uma professora decide comparar o desempenho de duas turmas para as suas aulas. Para isso selecionou a nota de dez alunos de cada turma em determinada atividade proposta em sala de aula. Ao realizar a correção da atividade, cujo valor máximo era 30, anotou

as notas que estão na Tabela 17. Criar dois diagramas de caixa e realizar a comparação entre eles.

Tabela 17 – Nota de duas turmas para comparação

Turma	
A	B
17	25
22	29
13	15
25	21
18	25
23	19
20	30
18	28
29	27
26	30

Fonte: Elaborada pelo autor (2021).

Calculando as medidas referente as observações do método A e do método B, tem-se:

Turma A:

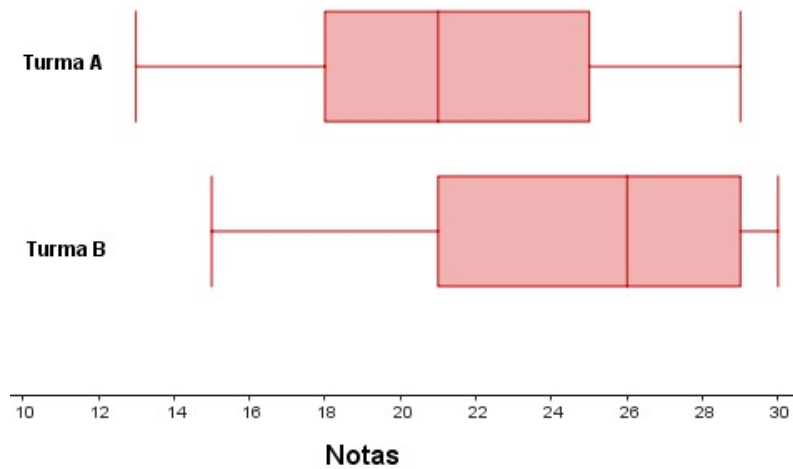
- Mínimo: 13
- Primeiro quartil: 18
- Mediana: 21
- Terceiro quartil: 24,5
- Máximo: 29
- Distância interquartílica: 6,5

Turma B:

- Mínimo: 15
- Primeiro quartil: 22
- Mediana: 26
- Terceiro quartil: 28,75
- Máximo: 30
- Distância interquartílica: 6,75

O Gráfico 9 representa os dois diagramas de caixa construídos para comparação das notas das duas turmas.

Gráfico 9 - Diagrama de Caixa para comparação de notas



Fonte: Elaborada pelo autor (2021).

Observa-se que a mediana da nota da turma B é maior. A variabilidade é praticamente a mesma. A turma B apresenta melhor desempenho do que a turma A. Não há observações discrepantes.

3.5.2 Percentis

São as medidas que dividem as séries em 100 partes iguais. O cálculo de um percentil segue a mesma técnica do cálculo da mediana, substituindo o termo $\frac{\sum f_i}{2}$, por:

$$k \frac{\sum f_i}{100},$$

sendo k o número de ordem do percentil. Assim, a fórmula para o cálculo dos percentis é dada por:

$$P_k = l^* + \frac{\left[k \frac{\sum f_i}{100} - F(\text{ant}) \right] \times h}{f^*}.$$

Exemplo: Considerando a Tabela 16 (p. 52), o oitavo percentil será:

$$k = 8 \quad \Rightarrow \quad 8 \frac{\sum f_i}{100} = 8 \frac{30}{100} = 2,4.$$

Logo:

$$P_8 = 148 + \frac{[2,4 - 2] \times 8}{6} = 148,53.$$

3.5.3 Variância

Quando se deseja calcular a variabilidade dos dados em relação à média, é interessante calcular o desvio de cada observação em relação à média. Observe:

$$d_i = x_i - \bar{x},$$

onde d_i representa o desvio do valor observado (x_i) subtraído da média aritmética de todos os valores (\bar{x}).

Para Bianchini (2018):

O desvio médio absoluto D_m de um conjunto de valores de uma variável estudada mede o grau de dispersão e de concentração dessa variável. Quanto maior o desvio médio absoluto, maior é a dispersão e menor é a concentração. Ou seja, em média, os valores afastam mais da média aritmética. E vice-versa.

Assim, o desvio médio absoluto é dado por:

$$D_m = \frac{\sum_{i=1}^n |d_i|}{n} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}.$$

Considera-se o módulo de cada desvio para se evitar que $\sum_{i=1}^n d_i = 0$.

Se os desvios encontrados forem baixos, os dados apresentam baixa dispersão. Ao contrário se os desvios forem altos, os dados apresentam alta dispersão. A variância é uma medida de variabilidade que é sensível para essas situações. Crespo (2002) define que a variância “baseia-se nos desvios em torno da média aritmética, porém determinando a média aritmética dos quadrados dos desvios”. Denotando a variância por s^2 , alguns autores utilizam "V" ou "Var", obtém-se:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}.$$

Exemplo: A quantidade de vendas realizadas em uma loja durante seis dias foram: 19; 21; 19; 21; 20; 20. Calcule a variância.

O cálculo da média resulta em: $\bar{x} = \frac{(19 + 21 + 19 + 21 + 20 + 20)}{6} = 20$. Os desvios, d_i , calculados são: -1, 1, -1, 1, 0 e 0. A variância será:

$$s^2 = \frac{(-1)^2 + (1)^2 + (-1)^2 + (1)^2 + (0)^2 + (0)^2}{6} = 0,66.$$

A variância 0,66 indica a variabilidade dos dados em termos de desvio da média ao quadrado.

Observação: Quando é utilizado somente uma parte do total dos dados observados, ou seja, uma amostra, ao invés da população total, convém utilizar o denominador $n - 1$ ou invés de n . Portanto, a variância amostral é dada por:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}.$$

Uma fórmula prática para o cálculo da variância pode ser obtida desenvolvendo algebricamente a fórmula indicada. Assim, se obtém:

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n}.$$

Utilizando o exemplo anterior, veja o cálculo da variância, por meio da Tabela 18, utilizando a fórmula prática:

Tabela 18 – Cálculo da variância

x	x^2
19	361
21	441
19	361
21	441
20	400
20	400
$\sum x = 120$	$\sum x^2 = 2404$

Fonte: Elaborada pelo autor (2021).

Aplicando a fórmula prática, tem-se:

$$s^2 = \frac{2404 - \frac{(120)^2}{6}}{6} = 0,66.$$

3.5.4 Desvio Padrão

Conforme Vieira (2018) descreve, é importante analisar que o cálculo da variância envolve o quadrado dos desvios, não sendo possível expressá-la na mesma unidade dos valores da variável. Para obter uma medida de variabilidade com a mesma unidade de medida dos dados, extrai-se a raiz quadrada da variância. Essa medida é o desvio padrão, cuja definição é a raiz quadrada da variância, a qual é denotada por s .

Logo:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}.$$

O desvio padrão facilita a interpretação dos dados pois mede de maneira eficaz a dispersão dos dados.

Exemplo: O número de falta total por mês em uma turma foi computado, mês a mês, durante o determinado semestre. Foram obtidos os seguintes dados:

18 12 13 18 25 28.

Calcular o desvio padrão desse grupo de dados.

Inicialmente deve-se calcular a média desses valores:

$$\bar{x} = \frac{18 + 12 + 13 + 18 + 25 + 28}{6} = \frac{114}{6} = 19.$$

Para obter o desvio padrão, calcula-se primeiro a variância. Assim:

$$s^2 = \frac{(18 - 19)^2 + (12 - 19)^2 + (13 - 19)^2 + (18 - 19)^2 + (25 - 19)^2 + (28 - 19)^2}{6}$$

$$s^2 = \frac{(-1)^2 + (-7)^2 + (-6)^2 + (-1)^2 + (6)^2 + (9)^2}{6}$$

$$s^2 = \frac{1 + 49 + 36 + 1 + 36 + 81}{6} = \frac{204}{6} = 34.$$

Para encontrar o desvio padrão, basta calcular a raiz quadrada da variância. Logo:

$$s = \sqrt{34} = 5,83.$$

Portanto o desvio padrão é de 5,83 faltas.

Pode-se utilizar a fórmula prática para o cálculo do desvio padrão também. Observe a Tabela 19 com os cálculos auxiliares:

Tabela 19 – Cálculo do desvio padrão

x	x^2
18	324
12	144
13	169
18	324
25	625
28	784
$\sum x = 114$	$\sum x^2 = 2370$

Fonte: Elaborada pelo autor (2021).

Aplicando a fórmula prática da variância para o desvio padrão, obtém-se:

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n}}$$

$$s = \sqrt{\frac{2370 - \frac{(114)^2}{6}}{6}}$$

$$= 5,83.$$

3.5.5 Coeficiente de Variação

Segundo Crespo (2002) o fato do desvio padrão estar na mesma unidade de medida dos dados observados, limita a sua utilização quando o objetivo é comparar dois ou mais conjuntos de dados, relativamente a sua variabilidade, quando estão expressos em unidades distintas. Para contornar essa limitação, utiliza-se o coeficiente de variação.

O coeficiente de variação (*C.V.*), segundo Martins (2006) é uma medida de dispersão relativa, calculada pelo quociente entre o desvio padrão (*s*) pela média (\bar{x}). O coeficiente de variação é adimensional, ou seja, não tem unidade de medida, uma vez que a média e o desvio padrão estão na mesma unidade, e se anulam. Portanto, pode ser expresso em percentual. Assim:

$$C.V. = \frac{s}{\bar{x}} \times 100.$$

Exemplo: Observe os resultados das idades e dos pesos de um mesmo grupo de indivíduos, de acordo com a Tabela 20:

Tabela 20 – Idades e Pesos de um grupo

	\bar{x}	<i>S</i>
Idade	41,6 anos	0,82 anos
Peso	82 kg	2,2 kg

Fonte: Elaborada pelo autor (2021).

Calculando o coeficiente de variação para cada variável, obtém-se:

Para a idade:

$$C.V._I = \frac{0,82}{41,6} \times 100 = 1,97\%.$$

Para o peso:

$$C.V._P = \frac{2,2}{82} \times 100 = 2,68\%.$$

Logo, nesse grupo de indivíduos, os pesos apresentam maior grau de dispersão que as idades.

Na sequência deste trabalho, o próximo capítulo traz uma análise sobre a utilização de recursos tecnológicos para o ensino de estatística na Educação Básica.

4 ESTATÍSTICA COM O AUXÍLIO DO GEOGEBRA E LIBREOFFICE CALC

Os recursos tecnológicos tem contribuído, nas últimas décadas, para a construção do conhecimento. Essas novas tecnologias de comunicação, segundo Kenski (2003) estão cada vez mais presentes no dia a dia. A integração desses novos meios de comunicação, já estão presentes em nosso comportamento e sem sentir, adaptamos a maneira de agir, de pensar e de nos comunicar. Para Muran (2007) essas tecnologias caminham para a convergência, a integração e a multifuncionalidade.

Na educação, se observa que essas mudanças tecnológicas funcionam como importante ferramenta de ensino. Segundo Muran (2007):

A digitalização permite registrar, editar, combinar, manipular toda e qualquer informação, por qualquer meio, em qualquer lugar, a qualquer tempo, traz a multiplicação de possibilidades de escolha, de interação. A mobilidade e a virtualização nos libertam de espaços e tempos rígidos, previsíveis, determinados. Na educação, o presencial *se virtualiza* e a distância *se presencializa*.

De acordo com Muran (2007) essas tecnologias são um suporte e transformaram-se em importantes instrumentos para a mudança na educação. Algumas etapas são necessárias para a inclusão dessas tecnologias no cotidiano. Inicialmente é preciso definir quais são as tecnologias mais adequadas para cada instituição. Em seguida deve-se realizar a aquisição e, nesta etapa, é importante definir o quanto gastar e que modelo adotar, se será utilizado *software* livre ou proprietário. E por último o domínio técnico-pedagógico, orientando o uso didático da ferramenta no processo de ensino e aprendizagem.

Um fator importante a considerar é a infraestrutura tecnológica necessária da escola para investimento. Kenski (2003) relata que a escola precisa estar conectada ao ambiente tecnológico. Para isso deve possuir infraestrutura adequada como computadores em número suficiente, *modems* e formas diversificadas e velozes de conexão à internet. Para a autora essa questão é motivo de preocupação, uma vez que muitas escolas no país não possuem condições mínimas de infraestrutura.

O ensino privado dispõe de recursos próprios e, em geral, pode garantir com maior facilidade sua integração no mundo das redes eletrônicas de comunicação e informação. O desafio é garantir essa mesma possibilidade às instituições públicas de ensino. A democratização do acesso ao conhecimento e ao uso das novas tecnologias passa pela necessidade de que as escolas públicas tenham condições de oferecer com qualidade essas atividades e possibilidades tecnológicas a seus alunos. (KENSKI, 2003)

Logo é importante que o incentivo ao uso de tecnologias aplicadas ao ambiente escolar seja fortalecido. A utilização de projetos pedagógicos com o uso dessas tecnologias

demonstra-se eficiente, porém requerem dedicação e tempo para sua elaboração e aplicação. No ensino de estatística é importante alinhar que o foco do ensino esteja na compreensão da estatística ao invés da tecnologia utilizada.

Para Chance *et al.* (2007) as ferramentas tecnológicas para o ensino de estatística podem ser divididas em diversas categorias: pacotes de *software* estatísticos, *software* educacional, planilhas, *applets*/aplicativos autônomos, calculadoras gráficas, materiais multimídias e repositório de dados. Segundo Ben-Zvi (2000) e Biehler (1997) *apud* Chance *et al.* (2007) alguns recursos disponíveis nessas categorias se sobrepõem e atendem ao mesmo objetivo, porém, nenhuma ferramenta contempla todos os recursos educacionais possíveis da tecnologia.

Para o presente estudo optou-se por utilizar os *softwares* GeoGebra e Libre Office Calc, ambos possuem licença livre e podem compor um ambiente apropriado para o ensino de estatística. Apesar de não serem *softwares* destinados para essa finalidade, apresentam variado pacote de recursos tecnológicos que podem auxiliar os alunos na compreensão de conceitos estatísticos e desenvolvimento de tarefas relacionadas ao tema. Na próxima seção, será abordado os *softwares* citados com seus principais recursos.

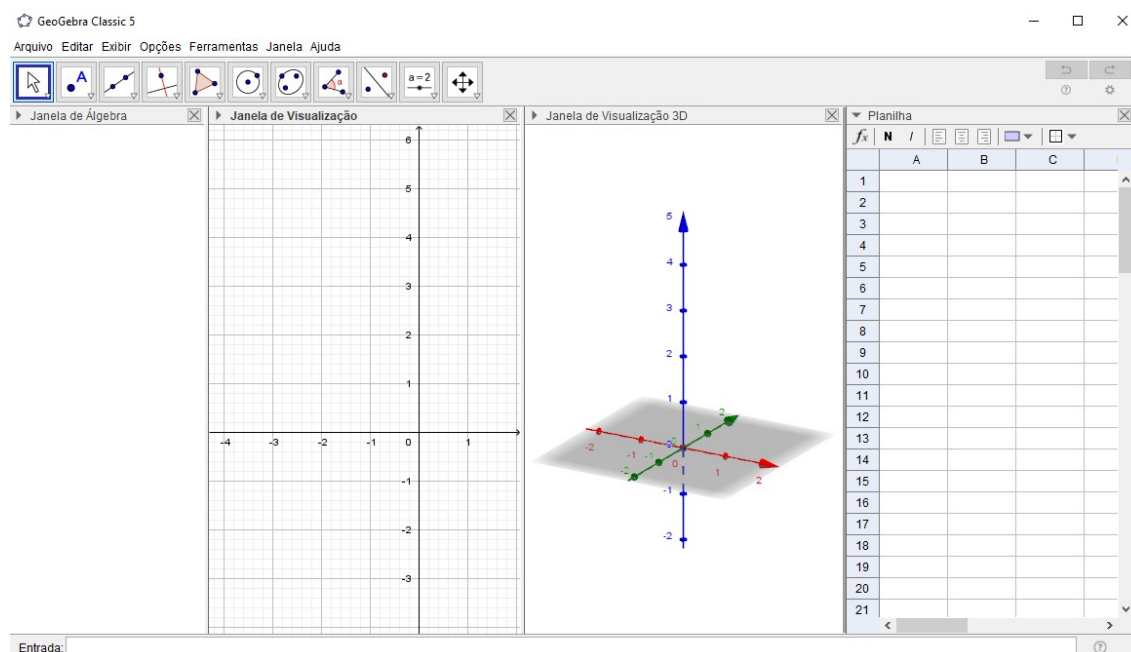
4.1 GEOGEBRA E SUAS APLICAÇÕES

O GeoGebra é um *software* dinâmico de matemática para todos os níveis de educação que reúne geometria, álgebra, planilhas, gráficos, estatísticas e cálculos em uma única plataforma (GEOGEBRA, 2022). O uso do GeoGebra proporciona a interatividade no auxílio do processo de aprendizagem. O *software* possibilita a busca de novas formas na solução de atividades propostas em sala de aula. É um *software* gratuito disponível em diversas plataformas (Microsoft Windows, Linux, Macintosh, etc.) e pode ser utilizado em todos os níveis de ensino e “vai de encontro de novas estratégias de ensino e aprendizagem, permitindo explorar, conjecturar, investigar tais conteúdos na construção do conhecimento matemático.”(SOUZA, R.; CALEJON, L. 2019).

O GeoGebra permite que na resolução de uma atividade, um mesmo objeto possibilite diferentes representações que podem ser modificadas dinamicamente para análise.

A Figura 4, apresenta a área de trabalho do GeoGebra com algumas das janelas de visualização disponíveis que podem ser exibidas conforme a necessidade de utilização.

A barra de menus apresenta um variado conjunto de opções que proporcionam ao usuário a manipulação dos arquivos (abrir arquivo, salvar, iniciar um novo arquivo, etc.), a exibição das janelas disponíveis para execução das tarefas, as ferramentas disponíveis, entre outras. A Janela de Álgebra exibe os registros algébricos das construções realizadas pelo usuário. A Janela de Visualização permite a exibição de gráficos e figuras geométricas

Figura 4 - Tela Inicial do *Software* GeoGebra Classic 5

Fonte: *Software* GeoGebra (2022).

bidimensionais, enquanto a Janela de Visualização 3D permite a exibição de gráficos e figuras geométricas tridimensionais. A Planilha permite a construção de tabelas e lista de pares ordenados de pontos que podem ser visualizados na Janela de Visualização e na Janela de Álgebra. Abaixo, no campo de Entrada, o usuário pode inserir comandos e equações a serem executados. Além disso, o *software* possui uma Janela CAS (Cálculo Algébrico Simbólico) que permite, por exemplo, a resolução de equações de forma simbólica. Para ter acesso a todas essas janelas o usuário pode configurar sua exibição por meio do menu Exibir.

4.2 LIBREOFFICE CALC E SUAS APLICAÇÕES

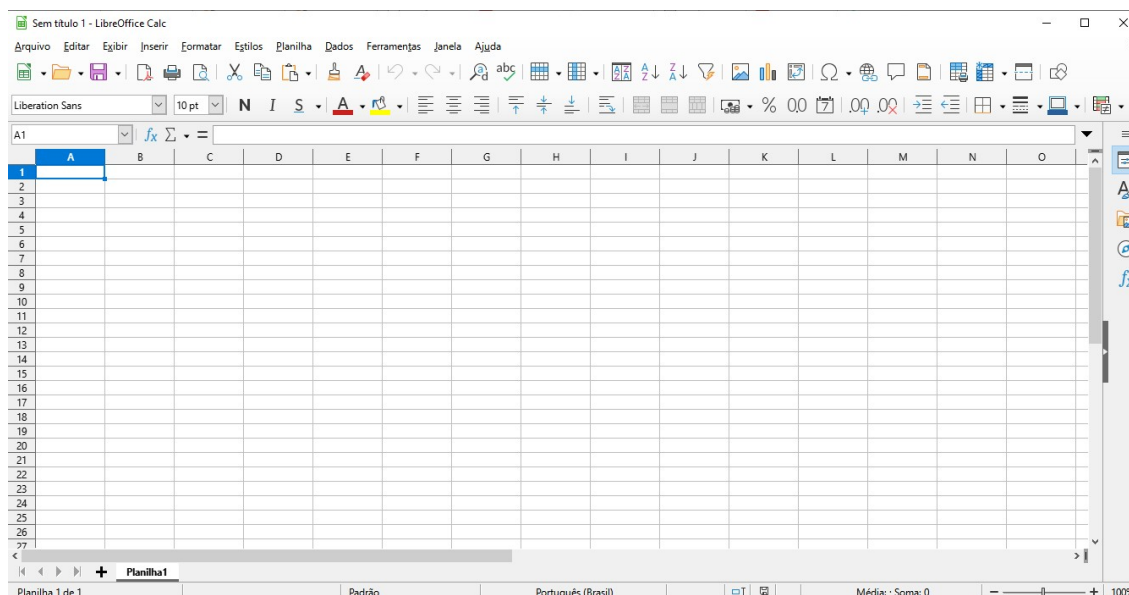
O LibreOffice é um pacote de programas para uso profissional ou pessoal. O serviço traz opções para criar e editar textos, tabelas, apresentações, desenhos, fórmulas matemáticas e até organizar um banco de dados (TECHTUDO, 2015).

O pacote conta com editor de texto (Writer), planilhas (Calc), criador de apresentações (Impress), editor de desenho e fluxogramas (Draw), banco de dados (Base) e editor de equações (Math). Os programas são compatíveis com os arquivos do Microsoft Office e vêm com conversor de PDF, entre outros recursos.

O LibreOffice Calc é um *software* gratuito que faz parte do LibreOffice e possibilita a criação, edição e apresentação de planilhas eletrônicas. A Figura 5 apresenta a tela de

início do Calc.

Figura 5 - Tela Inicial do *Software* LibreOffice Calc



Fonte: *Software* LibreOffice Calc (2022).

Quando o LibreOffice Calc é iniciado aparece uma janela de documento com o nome de Sem Título 1, que será substituído no momento que o arquivo for salvo. Abaixo é apresentado a aba Planilha 1, e o usuário pode incluir ou excluir abas a seu critério. Cada planilha contém colunas, dispostas lado a lado, e linhas que estão de cima a baixo e, que traçadas compreendem as células, cuja identificação é composta da letra de sua coluna com o número de sua linha. Na barra de endereços é mostrado a célula (ou conjunto de células) ativas. Na Figura 5, a célula ativa é a A1.

As células podem conter dados numéricos ou textual digitados pelo usuário. Também podem conter resultado de fórmulas ou função pré-definida. São nas células que o usuário entra com as informações com as quais irá realizar as análises ou classificações.

O *software* compreende ainda a barra de menus e ferramentas que auxiliam na manipulação dos arquivos, formatação da planilha, visualização dos dados, entre outras funções.

No capítulo a seguir, utilizando as ferramentas tecnológicas aqui apresentadas, será resolvido uma lista de exercícios. Para isso, será demonstrado de maneira detalhada, como acessar as funções necessárias, tanto do GeoGebra como do LibreOffice Calc, para a execução do passo a passo de cada uma.

5 RESOLUÇÃO DE QUESTÕES DE ESTATÍSTICA COM SUPORTE DAS FERRAMENTAS ELETRÔNICAS

Este capítulo tem por objetivo selecionar questões que envolvam o conteúdo de Estatística e propor uma resolução utilizando os recursos computacionais, os *softwares* GeoGebra e LibreOffice Calc, para auxiliar a compreensão do aluno na interpretação das questões.

Pretende-se no decorrer deste capítulo mostrar, com o auxílio das ferramentas eletrônicas, diversos conceitos relacionados a estatística, para a compreensão e análise crítica dos dados apresentados. A obtenção do resultado da questão será aprofundada com análise de outros parâmetros estatísticos e servirá de auxílio para que os professores possam estender a aprendizagem dos conceitos de estatística por meio destas ferramentas.

5.1 EXEMPLO 1

O número de faltas de cada aluno de uma turma com 25 alunos do Ensino Fundamental, durante um ano, foram levantadas pela direção da escola e ordenadas da seguinte forma: 5, 7, 9, 6, 8, 7, 5, 6, 10, 10, 9, 8, 5, 10, 9, 6, 7, 8, 10, 6, 11, 5, 7, 6 e 10 faltas. Calcule a média, moda, mediana desses dados informados.

Objetivos da tarefa:

- a) Compreensão dos conceitos de Média, Moda e Mediana.
- b) Mostrar as características geométricas das estatísticas calculadas.
- c) Compreender a estrutura de um Diagrama de Pontos e como relacioná-lo às medidas de tendência central.

Competências e Habilidades na BNCC

- EF07MA35, EF08MA25 e EF09MA22.

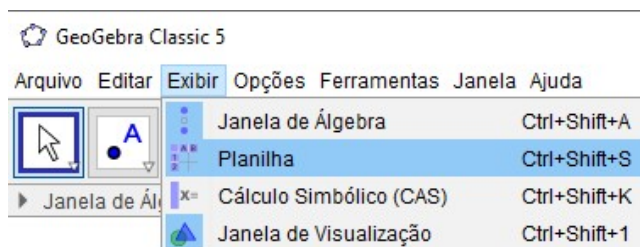
Para a solução proposta a seguir, os discentes irão aplicar a teoria aprendida utilizando uma ferramenta tecnológica. O professor, além de apresentar a solução, irá ampliar as possibilidades de utilização do *software* com questões complementares a tarefa.

Resolução proposta:

Será utilizado o *software* GeoGebra para a apresentação da resolução.

1º passo: Inicialmente os alunos irão habilitar no GeoGebra a planilha para lançar os dados coletados. Para isso, clicar em *Menu Exibir » Planilha*, conforme Figura 6.

Figura 6 - Habilitar Planilha no GeoGebra



Fonte: Elaborada pelo autor (2022).

2º passo: Logo em seguida, começando da célula A1, deslocando-se para baixo, até a célula A25, irá preencher com os dados do problema. Observe a Figura 7. Nesse momento, pode-se explicar que a ordem de preenchimento não precisa seguir um padrão para esse exercício, pois somente uma variável (faltas) está sendo tratada.

Figura 7 - Inserção de dados na Planilha do GeoGebra

The image shows the 'Planilha' (Spreadsheet) window in GeoGebra. The spreadsheet has columns A, B, C, and D. Column A contains the following values from row 1 to row 19: 5, 7, 9, 6, 8, 7, 5, 6, 10, 10, 9, 8, 5, 10, 9, 6, 7, 8, 10.

	A	B	C	D
1	5			
2	7			
3	9			
4	6			
5	8			
6	7			
7	5			
8	6			
9	10			
10	10			
11	9			
12	8			
13	5			
14	10			
15	9			
16	6			
17	7			
18	8			
19	10			

Fonte: Elaborada pelo autor (2022).

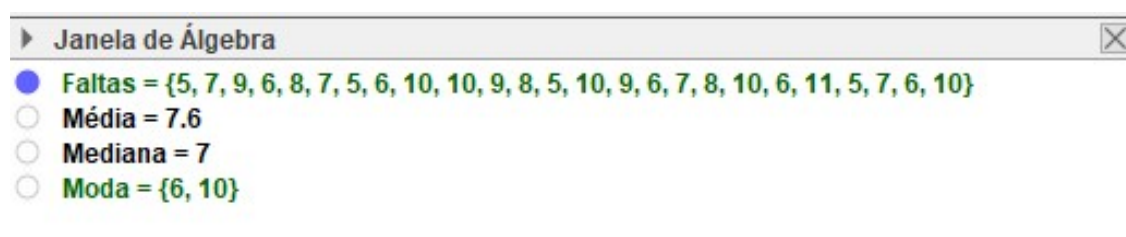
3º passo: Deve-se agora criar uma lista com esses dados, ou seja, torná-los um conjunto de dados. Para isso os alunos devem selecionar todos os elementos, clicar com o botão direito do mouse, selecionar *Criar Lista*. A lista criada será exibida na *Janela de Álgebra*, que pode ser renomeada para *Faltas* da seguinte maneira: clicar com o botão direito do mouse sobre a lista e na opção *Renomear*, digite o nome sugerido na caixa de diálogo que irá aparecer.

4º passo: Para calcular as medidas de tendência central, será usado os comandos listados abaixo no campo de *Entrada*. O resultado será exibido na *Janela de Álgebra*. Pode-se renomear os resultados clicando com o botão direito do mouse em cima dos valores visualizados na *Janela de Álgebra*, escolhendo a opção *Renomear*.

- Para encontrar o valor da média entre os dados, digite *Média(Faltas)*, e renomeie o resultado para ***Média***.
- Para localizar o valor central entre os dados, digite *Mediana(Faltas)* e renomeie o resultado para ***Mediana***.
- Para encontrar o valor com maior frequência entre os dados, digite *Moda(Faltas)* e renomeie o resultado para ***Moda***.

Os resultados obtidos estão listados na Figura 8:

Figura 8 - Exibição dos resultados na Janela de Álgebra



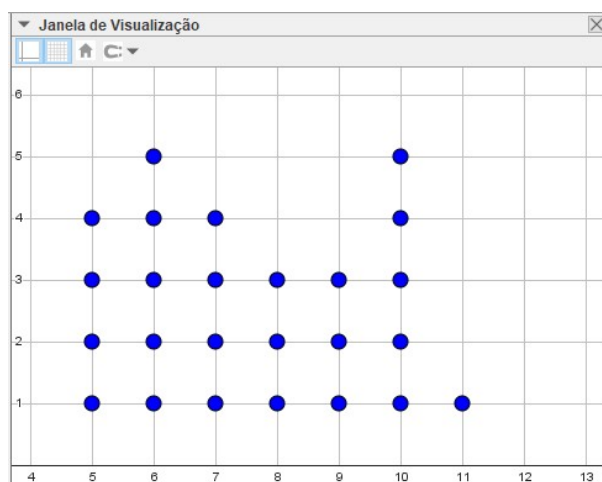
Fonte: Elaborada pelo autor (2022).

Assim, os valores da *Média*, *Mediana* e *Moda* são, respectivamente, 7,6 ; 7 e {6, 10}.

Os alunos deverão perceber que os valores são facilmente obtidos utilizando o GeoGebra. A partir deste momento, como complemento ao exercício, será realizada uma análise do comportamento desses dados, por meio de um diagrama de pontos.

5º passo: No campo de *Entrada* digite *DiagramaDePontos(Faltas)* e renomeie para ***Gráfico***. Na *Janela de Visualização* será mostrado o diagrama da ***Gráfico 10***.

Gráfico 10 - Exibição gráfica dos resultados



Fonte: Elaborada pelo autor (2022).

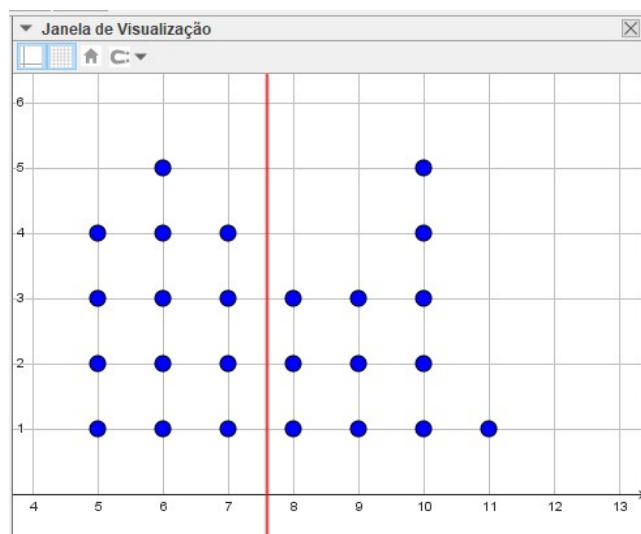
Deve-se mostrar aos alunos que o diagrama exibe a falta de todos os alunos. O *Eixo X* representa a quantidade de faltas dos alunos e o *Eixo Y* a quantidade de alunos com a respectiva falta. Os alunos devem ser motivados a realizarem configurações de cor, formato, tamanho, etc. para se familiarizarem com as possibilidades de exibição do *software*.

Agora será analisado o comportamento das medidas de tendência central calculadas algebricamente.

Para compreender como analisar a moda, os alunos devem observar qual “coluna de pontos” apresenta o maior valor. Eles devem chegar à conclusão de que 6 e 10 são as faltas com maiores frequências, 5 cada uma. Algumas observações auxiliares podem ser vistas como a menor quantidade de faltas é 5 e a maior quantidade de faltas é 11, sendo esta observada somente uma vez. A amplitude dos dados é 6, que representa o maior valor observado subtraído do menor valor observado.

Para identificar a média, será utilizado uma reta paralela ao Eixo Y, com o valor calculado de 7,6. Para isso no campo *Entrada* será digitado: $x=7,6$. Os alunos também podem utilizar a sintaxe no campo de *Entrada*: $x=Média(Faltas)$. O diagrama irá apresentar uma reta que representa a média entre as idades, conforme Gráfico 11.

Gráfico 11 - Inclusão da reta que identifica a média

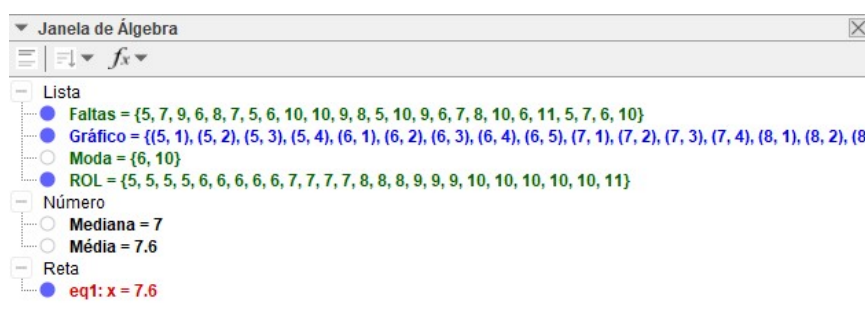


Fonte: Elaborada pelo autor (2022).

Deve-se explicar aos alunos que pode ocorrer que a média seja um número diferente de todos os elementos apresentados e, que neste caso, a média não tem existência concreta.

Para exibir graficamente a mediana, deve-se inicialmente criar uma lista auxiliar com os dados ordenados. No campo de *Entrada* digite *Ordenar(Faltas)* e renomeie para **ROL**. O resultado será exibido na Janela de Álgebra, conforme a Figura 9.

Figura 9 - Exibição da Janela de Álgebra com os novos resultados



Fonte: Elaborada pelo autor (2022).

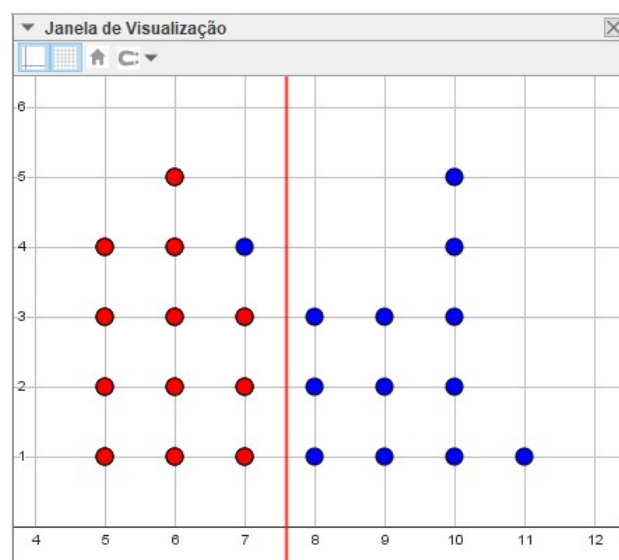
A lista ordenada (ROL) irá evidenciar as faltas em ordem crescente. A mediana irá destacar os elementos que estão abaixo de 50% do tamanho da amostra e para isto será criada outra lista, com a metade dos valores das faltas.

No campo de *Entrada* digite *ParteDaLista(ROL, 1, Comprimento(ROL)/2)* e renomeie para *Med*. Esse comando irá criar uma lista a partir de ROL com metade dos elementos de ROL. Para finalizar, será criado outro diagrama de pontos, mas desta vez

com a lista *Med*. Basta digitar *DiagramaDePontos(Med)*, e modificar as configurações de cor para destacar com relação ao diagrama de pontos das faltas.

O Gráfico 12 mostra o novo diagrama com a mediana na Janela de Visualização .

Gráfico 12 - Identificação da Mediana no Diagrama de Pontos



Fonte: Elaborada pelo autor (2022).

Por meio desses comandos os alunos devem observar que o diagrama ficou dividido em duas partes, e que a parte inferior compreende 50% dos elementos, e como calculado, a mediana é 7.

Como complementação, poderá ser solicitado o cálculo e exibição dos quartis. Os alunos também podem alterar o tamanho da amostra e seus respectivos valores para verificar o comportamento do gráfico e das estatísticas calculadas.

5.2 EXEMPLO 2

QUESTÃO 170 / ENEM 2012 (Caderno Amarelo): A tabela a seguir mostra a evolução da receita bruta anual dos três últimos anos de cinco microempresas (ME) que se encontram à venda.

ME	2009 (em milhares de reais)	2010 (em milhares de reais)	2011 (em milhares de reais)
Alfinetes V	200	220	240
Balas W	200	230	200
Chocolates X	250	210	215
Pizzaria Y	230	230	230
Tecelagem Z	160	210	245

Um investidor deseja comprar duas das empresas listadas na tabela. Para tal ele calcula a média da receita bruta anual dos últimos três anos (de 2009 a 2011) e escolhe as duas empresas de maior média anual.

As empresas que este investidor escolhe comprar são:

- (A) Balas W e Pizzaria Y.
- (B) Chocolates X e Tecelagem Z.
- (C) Pizzaria Y e Alfinetes V.
- (D) Pizzaria Y e Chocolates X.
- (E) Tecelagem Z e Alfinetes V.

Objetivos da tarefa:

- a) Compreensão dos conceitos de Média.
- b) Analisar e interpretar conjunto de média de grupos distintos.
- c) Compreender a estrutura de um Gráfico de linhas e de barras e relacioná-los às medidas de tendência central e à variabilidade dos dados.

Competências e Habilidades na BNCC

- EF02MA22, EF09MA22 e EM13MAT316.

Na solução proposta a seguir, além de resolver o exercício os discentes irão analisar outras formas de apresentar o resultado por meio de gráficos que permitem a interdisciplinaridade. Além disso, será proposto cálculos complementares de outras estatísticas.

Resolução proposta:

Será utilizado o *software* LibreOffice Calc para a apresentação da resolução.

1º passo: Os alunos irão digitar o quadro do exercício na área da planilha eletrônica. Espera-se que cheguem no resultado da Figura 10. As configurações de formatação ficam a critério dos alunos.

Figura 10 - Dados digitados na Planilha do LibreOffice Calc

	A	B	C	D
1	ME	2009 (em milhares de reais)	2010 (em milhares de reais)	2011 (em milhares de reais)
2	Alfinetes V	200	220	240
3	Balas W	200	230	200
4	Chocolates X	250	210	215
5	Pizzaria Y	230	230	230
6	Tecelagem Z	160	210	245
7				

Fonte: Elaborada pelo autor (2022).

2º passo: Será solicitado aos alunos que insiram uma nova coluna, com o título de Média. Em cada célula dessa coluna, os alunos devem inserir a fórmula $=MÉDIA()$ e dentro do parênteses colocar o intervalo das células que contenham os dados de cada empresa, ou seja, para a Empresa Alfinetes V, será $=MÉDIA(B2:D2)$. Repetir o processo para as demais empresas. Na Figura 11 o resultado esperado é exibido.

Figura 11 - Fórmula da média inserida na planilha

	A	B	C	D	E
1	ME	2009 (em milhares de reais)	2010 (em milhares de reais)	2011 (em milhares de reais)	Média
2	Alfinetes V	200	220	240	=MÉDIA(B2:D2)
3	Balas W	200	230	200	=MÉDIA(B3:D3)
4	Chocolates X	250	210	215	=MÉDIA(B4:D4)
5	Pizzaria Y	230	230	230	=MÉDIA(B5:D5)
6	Tecelagem Z	160	210	245	=MÉDIA(B6:D6)
7					

Fonte: Elaborada pelo autor (2022).

Após o preenchimento de todos os campos da planilha, o cálculo executado está exibido na Figura 12:

Figura 12 - Resultado da média

	A	B	C	D	E
1	ME	2009 (em milhares de reais)	2010 (em milhares de reais)	2011 (em milhares de reais)	Média
2	Alfinetes V	200	220	240	220
3	Balas W	200	230	200	210
4	Chocolates X	250	210	215	225
5	Pizzaria Y	230	230	230	230
6	Tecelagem Z	160	210	245	205

Fonte: Elaborada pelo autor (2022).

3º passo: Após esse cálculo os alunos podem facilmente, por meio de observação, identificar quais são as duas empresas que possuem a maior média no período analisado, mas é interessante apresentar aos alunos o conceito de ordenar no *software*, uma vez que será de grande auxílio quando forem lidar com uma amostra, que possua grande quantidade de dados. Para isso, devem selecionar a tabela inteira, ou seja, o intervalo (A2:E6), clicar no menu *Dados » Ordenar...* Na caixa de diálogo selecionar na *Chave de ordenação 1* a coluna *Média*, depois clicar em *Decrescente* e em seguida *OK*. A planilha apresentará os dados com ordem decrescente de Média, conforme a Figura 13:

Figura 13 - Classificação decrescente de acordo com a média

	A	B	C	D	E
1	ME	2009 (em milhares de reais)	2010 (em milhares de reais)	2011 (em milhares de reais)	Média
2	Pizzaria Y	230	230	230	230
3	Chocolates X	250	210	215	225
4	Alfinetes V	200	220	240	220
5	Balas W	200	230	200	210
6	Tecelagem Z	160	210	245	205

Fonte: Elaborada pelo autor (2022).

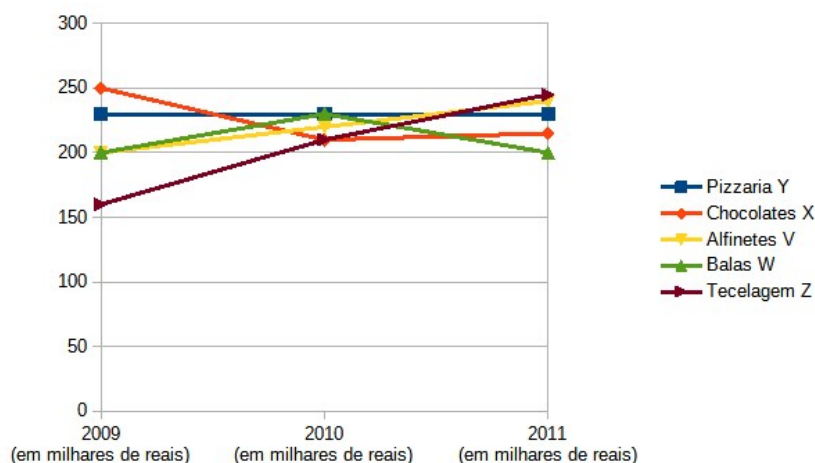
Assim, a resposta do exercício é a **alternativa D**, *Pizzaria Y* e *Chocolates X*.

O exercício foi resolvido no passo anterior, todavia como complemento da atividade será interessante mostrar aos alunos como apresentar esses dados de forma gráfica e realizar algumas análises complementares.

4º passo: Como os dados do exercício se trata de uma série histórica, o gráfico ideal para análise é o gráfico de linhas. Para isso, selecione os dados, exceto a coluna média, clique no menu *Inserir » Gráfico*. Na caixa de diálogo em Tipo de Gráfico escolha

a opção de gráfico *Linha*. Em intervalo de dados escolha *Série de Dados em Linha*. O Gráfico 13 será apresentado no *software*.

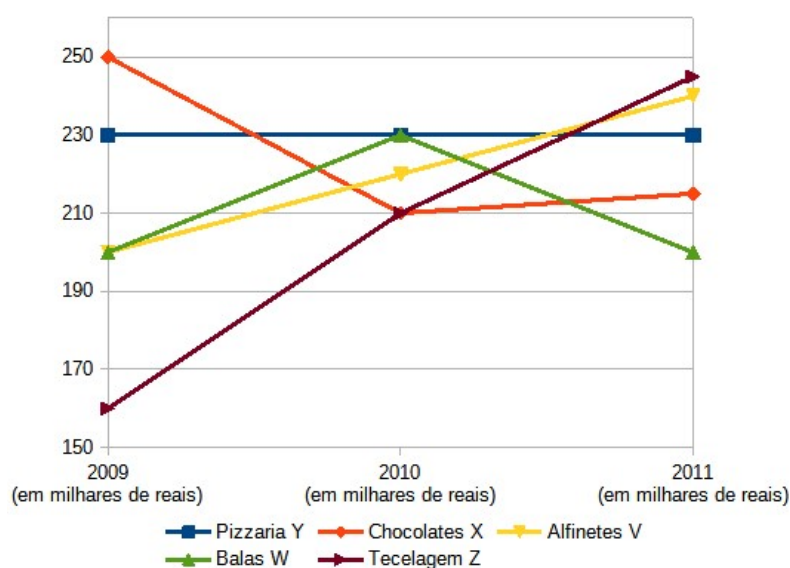
Gráfico 13 - Gráfico de linhas para os dados do problema



Fonte: Elaborada pelo autor (2022).

5º passo: Solicite aos alunos que melhorem a visualização do gráfico, alterando a posição das legendas, intervalo dos eixos, entre outros. Nesse momento é importante mostrar aos alunos como a visualização de um gráfico facilita na interpretação dos dados. Uma sugestão de configuração seria a do Gráfico 14.

Gráfico 14 - Configuração do gráfico de linhas



Fonte: Elaborada pelo autor (2022).

Para obtê-la os alunos devem executar os seguintes comandos:

- Clicar com o botão direito do mouse em cima do Eixo Y e escolher a opção Formatar eixo...
- Na aba Escala alterar os valores para:
 - Mínimo: 150
 - Máximo: 260
 - Intervalo principal: 20
- Clicar no botão OK
- Para alterar a posição da legenda, clicar com o botão direito do mouse em cima da legenda e escolher a opção Formatar legenda...
- Na aba Posição, escolher a opção Embaixo.
- Clicar no botão OK.

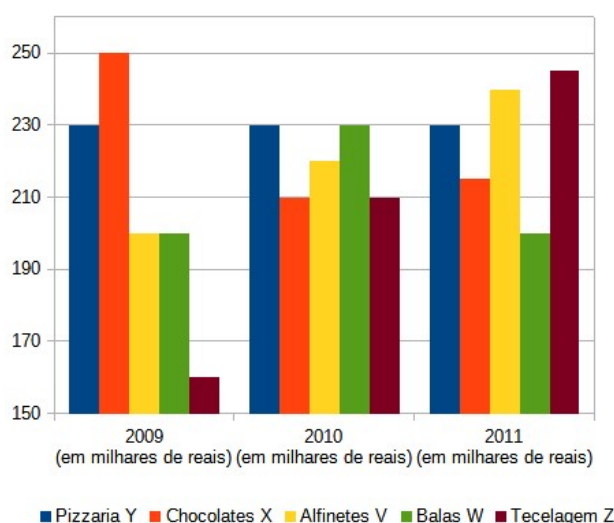
6º passo: Para a análise do gráfico os alunos poderão avaliar outras situações particulares de cada empresa, e não somente a média. Verifica-se que:

- a *Pizzaria Y* apresenta a melhor regularidade, uma vez que seu gráfico é uma constante;
- a *Tecelagem Z* é a empresa que está apresentando o melhor resultado, pois seu gráfico apresenta crescimento no período analisado;
- a empresa *Alfinetes V* também apresenta um crescimento no período analisado;
- a *Chocolates X* apresentou uma queda brusca em 2010 e se reergueu um pouco em 2011;
- as *Balas W* tiveram um crescimento em 2010 e logo em seguida caiu novamente.

Essa análise gráfica pode ser utilizada como uma ferramenta para interdisciplinaridade, pois os alunos podem avaliar o cenário econômico do país, estratégias de marketing e mercado, entre outros fatores e não somente a análise da média da Receita.

Pode-se sugerir que os alunos realizem a construção de outros gráficos, por exemplo Gráfico de barras vertical, conforme o Gráfico 15. Por meio desse gráfico, os alunos podem comparar o desempenho das empresas em cada ano.

Gráfico 15 - Gráfico de Barras com os dados do problema



Fonte: Elaborada pelo autor (2022).

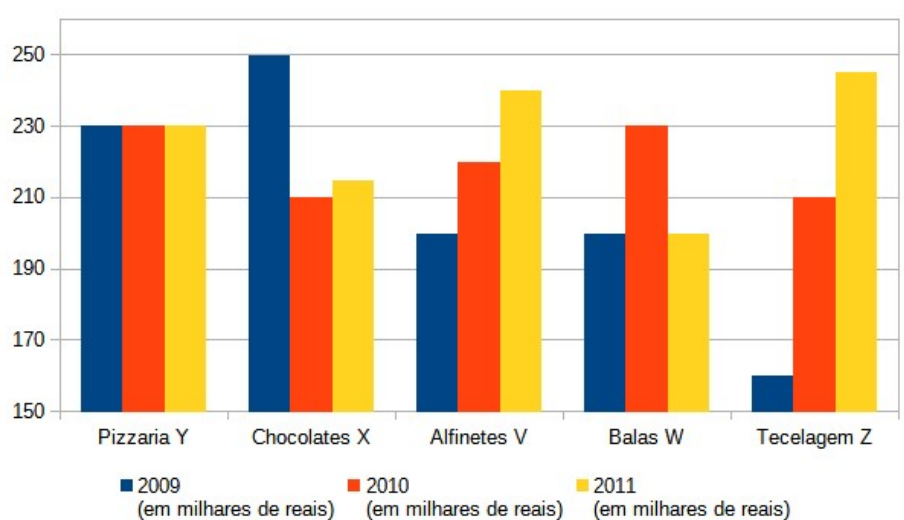
Para obter o Gráfico 15 os alunos devem realizar os seguintes comandos:

- Selecionar os dados da planilha;
- Clicar no menu Inserir » Gráfico;
- Na caixa de diálogo devem executar as seguintes alterações:
 - Em tipo de gráfico: escolher a opção Coluna;
 - Em Intervalo de dados: selecionar a opção Série de dados em linhas;
 - Em Série de dados: manter as configurações sugeridas;
 - Em elementos gráficos: escolher a opção Embaixo para Exibir legendas;
 - Clicar em Finalizar;
- Com o gráfico pronto alterar as configurações do Eixo Y, conforme o passo anterior.

7º passo: A análise gráfica do passo anterior nos leva a inserir o conceito de Desvio Padrão aos alunos. Conforme visto na Seção 3.5.4 o desvio padrão facilita a interpretação da dispersão dos dados. Logo, será modificada a configuração do gráfico para analisar o comportamento de cada empresa anualmente. Para isso deve-se seguir os mesmos passos anteriores, somente alterando em Intervalo de dados a opção *Série de dados em colunas*. Espera-se chegar ao resultado do Gráfico 16.

8º passo: Será analisado o comportamento de cada empresa de acordo com a altura de suas colunas. A *Pizzaria Y* não teve alteração, logo mostra pouca variabilidade. As empresas *Chocolate X*, *Alfinetes V* e *Balas W*, tiveram pouca variabilidade no período analisado. Entretanto a empresa *Tecelagem Z* apresenta a maior variabilidade entre as

Gráfico 16 - Análise do Desvio Padrão



Fonte: Elaborada pelo autor (2022).

empresas. Para confirmar essa interpretação, pode-se sugerir aos alunos o cálculo do desvio padrão de cada empresa. Utilizando o *software* LibreOffice Calc, na planilha existente, será inserida uma coluna após a Média com o nome de Desvio Padrão. Nas células correspondente de cada empresa, será digitado o comando $=DESVPAD()$. Entre os parênteses será digitado o intervalo referente aos dados das empresas. A Figura 14 exibe o resultado encontrado.

Figura 14 - Cálculo do Desvio Padrão

	A	B	C	D	E	F
1	ME	2009 (em milhares de reais)	2010 (em milhares de reais)	2011 (em milhares de reais)	Média	Desvio Padrão
2	Pizzaria Y	230	230	230	230	0,00
3	Chocolates X	250	210	215	225	21,79
4	Alfinetes V	200	220	240	220	20,00
5	Balas W	200	230	200	210	17,32
6	Tecelagem Z	160	210	245	205	42,72

Fonte: Elaborada pelo autor (2022).

Observe que a análise gráfica corresponde aos valores encontrados do Desvio Padrão. Aplicando os conceitos vistos anteriormente, quanto menor o desvio padrão, menor a variabilidade entre os dados.

Outra maneira dos discentes verificarem a variabilidade dos dados seria utilizar o Coeficiente de Variação. Para isso, na planilha acrescentar após a coluna Desvio Padrão a coluna Coeficiente de Variação. Nas células correspondentes de cada empresa deverá

selecionar o valor referente ao Desvio Padrão e dividir pelo valor da Média. Assim, para a empresa *Pizzaria Y*, deverá digitar $=F2/E2$. Essa fórmula será copiada para as demais empresas. Será obtido o resultado da Figura 15. Peça para os alunos configurarem as células com o formato de Porcentagem.

Figura 15 - Cálculo do Coeficiente de Variação

	A	B	C	D	E	F	G
1	ME	2009 (em milhares de reais)	2010 (em milhares de reais)	2011 (em milhares de reais)	Média	Desvio Padrão	Coeficiente de Variação
2	Pizzaria Y	230	230	230	230	0,00	0,00%
3	Chocolates X	250	210	215	225	21,79	9,69%
4	Alfinetes V	200	220	240	220	20,00	9,09%
5	Balas W	200	230	200	210	17,32	8,25%
6	Tecelagem Z	160	210	245	205	42,72	20,84%

Fonte: Elaborada pelo autor (2022).

Observe que, quanto menor o coeficiente de variação, maior a uniformidade entre os dados. Esse conceito é aplicado quando o objetivo é comparar dois ou mais conjunto de dados, com tamanho de amostra distintos ou dados que estão em unidades de medida diferentes.

Os alunos devem ser motivados a pesquisarem dados reais de empresas e realizarem os mesmos procedimentos ou então alterar os dados na planilha, acrescentar outros anos para verificar o comportamento dos dados analisados.

5.3 EXEMPLO 3

QUESTÃO 150 / ENEM 2014 (Caderno Amarelo): Ao final de uma competição de ciências em uma escola, restaram apenas três candidatos. De acordo com as regras, o vencedor será o candidato que obtiver a maior média ponderada entre as notas das provas finais nas disciplinas química e física, considerando, respectivamente, os pesos 4 e 6 para elas. As notas são sempre números inteiros. Por questões médicas, o candidato II ainda não fez a prova final de química. No dia em que sua avaliação for aplicada, as notas dos outros dois alunos, em ambas as disciplinas, já terão sido divulgadas.

O quadro apresenta as notas obtidas pelos finalistas nas provas finais.

Candidato	Química	Física
I	20	23
II	X	25
III	21	18

A menor nota que o candidato II deverá obter na prova final de química para vencer a competição é:

- (A) 18.
- (B) 19.
- (C) 22.
- (D) 25.
- (E) 26.

Objetivos da tarefa:

- a) Compreensão dos conceitos de Média Ponderada.
- b) Retomar o conceito de solução conjunto das inequações.
- c) Utilizar o Controle Deslizante do GeoGebra para simulações de resultados.

Competências e Habilidades na BNCC

- EF02MA22 e EF03MA26.

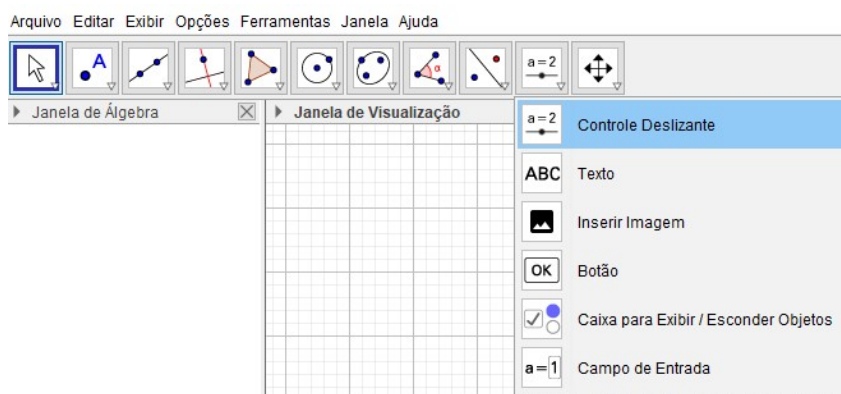
A resolução apresentada a seguir, levará em consideração uma série de valores possíveis para encontrar a solução. Nesse caso, os alunos devem observar o valor mínimo a ser considerado para que o candidato II vença a competição. Conceitos complementares de inequação devem ser apresentados aos alunos.

Resolução proposta:

Será utilizado o *software* GeoGebra para a apresentação da resolução.

1º passo: Inicialmente os alunos irão habilitar a Planilha do GeoGebra para inserir os dados do exercício, clicar em *Menu Exibir » Planilha*. Antes de lançar os valores na planilha, será criado um controle deslizante que representará a nota X do candidato II. Para isso, os alunos devem ir na Barra de Ferramentas e clicar no botão correspondente conforme Figura 16.

Figura 16 - Inserindo o Controle Deslizante



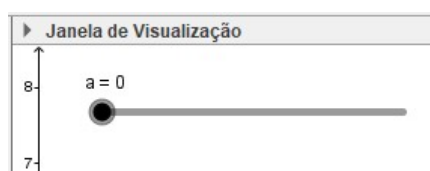
Fonte: Elaborada pelo autor (2022).

Após clicar na Janela de Visualização, o *software* irá apresentar a caixa de diálogo com as configurações do Controle Deslizante. Realizar as seguintes modificações:

- Será um controle do tipo: Número;
- O controle terá o nome de a ;
- Em intervalo: min: 0 e max: 30 (o exercício não informa o valor máximo da avaliação e neste caso será considerado que a nota máxima possível é 30);
- O incremento será de 1, pois as notas só podem ser valores inteiros.

Após confirmar as modificações, deixar o valor inicial como 0, uma vez que o aluno não fez a avaliação. Na Janela de Visualização, será exibido o Controle Deslizante, conforme a Figura 17.

Figura 17 - Controle Deslizante



Fonte: Elaborada pelo autor (2022).

A seguir, os alunos devem habilitar a planilha no GeoGebra, e ao preencher a célula que representa a nota X , será colocado o nome do controle deslizante a , e ficará com o formato da Figura 18.

Figura 18 - Dados inseridos na Planilha do GeoGebra

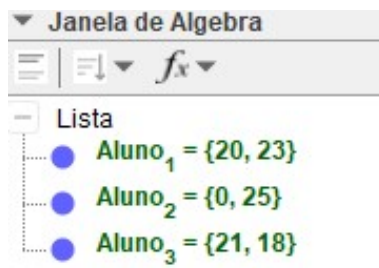
Planilha			
	A	B	C
1	Candidato	Química	Física
2	I	20	23
3	II	0	25
4	III	21	18

Fonte: Elaborada pelo autor (2022).

2º passo: Logo em seguida, será criada 3 listas: $Aluno_1$, $Aluno_2$ e $Aluno_3$ que representam as notas dos candidatos I, II e III, respectivamente. Para isso, selecione as células $B2:C2$, que representam as notas do candidato I, clique com o botão direito do

mouse e escolha: *Criar » Lista*. Renomeie a lista para *Aluno_1*. Repita o procedimento para os demais candidatos. O resultado será visualizado na Janela de Álgebra exibido na Figura 19:

Figura 19 - Listas de notas dos alunos

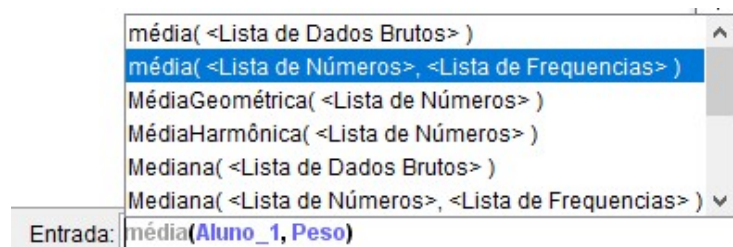


Fonte: Elaborada pelo autor (2022).

3º passo: Deve-se agora criar uma lista que represente os pesos das disciplinas. Para isso, na Caixa de Entrada digitar: $Peso=\{4,6\}$, nesta ordem, uma vez que representam os pesos das disciplinas de Química e Física, respectivamente.

4º passo: Neste momento é oportuno revisar os conceitos de média ponderada com os alunos. Após, utilizando a Entrada de Comandos será calculado a média de cada candidato. Quando o objetivo é calcular a média ponderada, deve-se utilizar a seguinte sintaxe: $média(<Lista\ de\ Números>, <Lista\ de\ Freqüências>)$ onde *Lista de Números* representa a lista com a nota do candidato e *Lista de Freqüências* representa o peso da nota das disciplinas. Para o candidato I, a entrada será $média(Aluno_1, Peso)$, conforme a Figura 20:

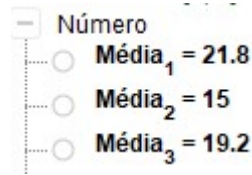
Figura 20 - Cálculo da média ponderada do Candidato I



Fonte: Elaborada pelo autor (2022).

Renomeie o resultado obtido para *Média_1* e repita o processo para cada candidato. O resultado será exibido na Janela de Álgebra. Observe a Figura 21.

Figura 21 - Exibição na Janela de Álgebra das médias calculadas



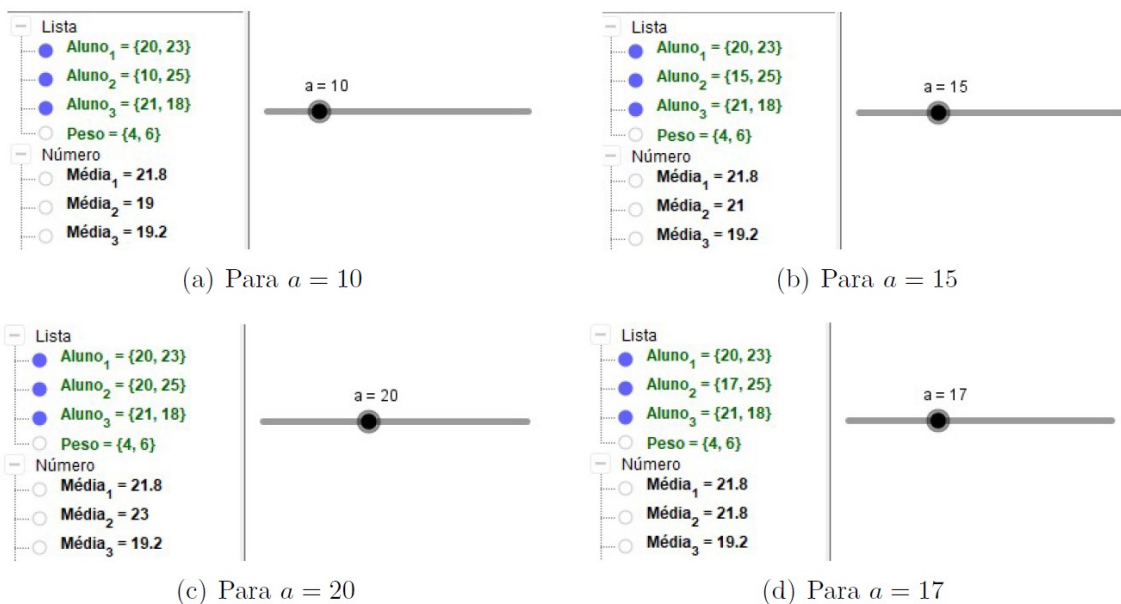
Fonte: Elaborada pelo autor (2022).

Portanto o candidato II tem uma média inicial de 15, e para vencer a competição, deve possuir uma nota que eleve sua média para acima da média do candidato I.

5^o passo: Os alunos irão utilizar o controle deslizante para simular algumas notas para o candidato II, e devem notar que ao aumentar o valor do controle deslizante, a média do candidato II irá aumentar. Nessa parte é importante destacar que toda vez que se aumenta o valor de determinado dado de uma amostra, a média também aumenta. O mesmo ocorrerá ao se diminuir o valor de determinado dado da amostra, ou seja, a média também irá diminuir.

Os alunos poderão simular notas a partir do 0, até encontrar o resultado. Para fins de apresentação, será utilizado os valores 10, 15, 20 e 17 para avaliar o comportamento da média. A Figura 22 mostra os resultados encontrados.

Figura 22 - Exibição dos Resultados



Fonte: Elaborada pelo autor (2022).

Com nota 10 na prova, a média do candidato II subiu para 19, porém é insuficiente para vencer a competição. Observe a Figura 22(a).

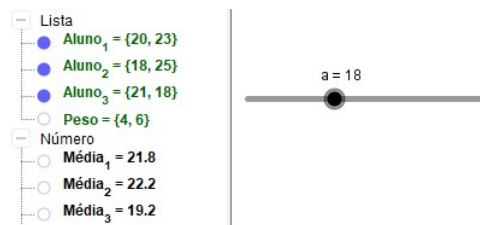
Com nota 15 na prova, o candidato II ultrapassa o candidato III, conforme Figura 22(b), porém ainda não obteve nota suficiente para vencer a competição.

Com nota 20 na prova, verifica-se que o candidato II ganha a competição Figura 22(c). Mas como o exercício solicita a nota mínima para vencer, deve-se verificar se essa nota é a mínima necessária.

Com nota 17, o candidato II empata a competição. Observe a Figura 22(d).

Logo com a nota 18, o candidato II é o vencedor, lembrando que o incremento foi de 1, uma vez que o exercício determinou as notas com valores inteiros. Pode-se observar o resultado na Figura 23.

Figura 23 - Cálculo final da média



Fonte: Elaborada pelo autor (2022).

Logo para qualquer valor acima de 18, o candidato II será o campeão da competição. Ou seja:

$$S = \{x \in \mathbb{R} | x \geq 18\}$$

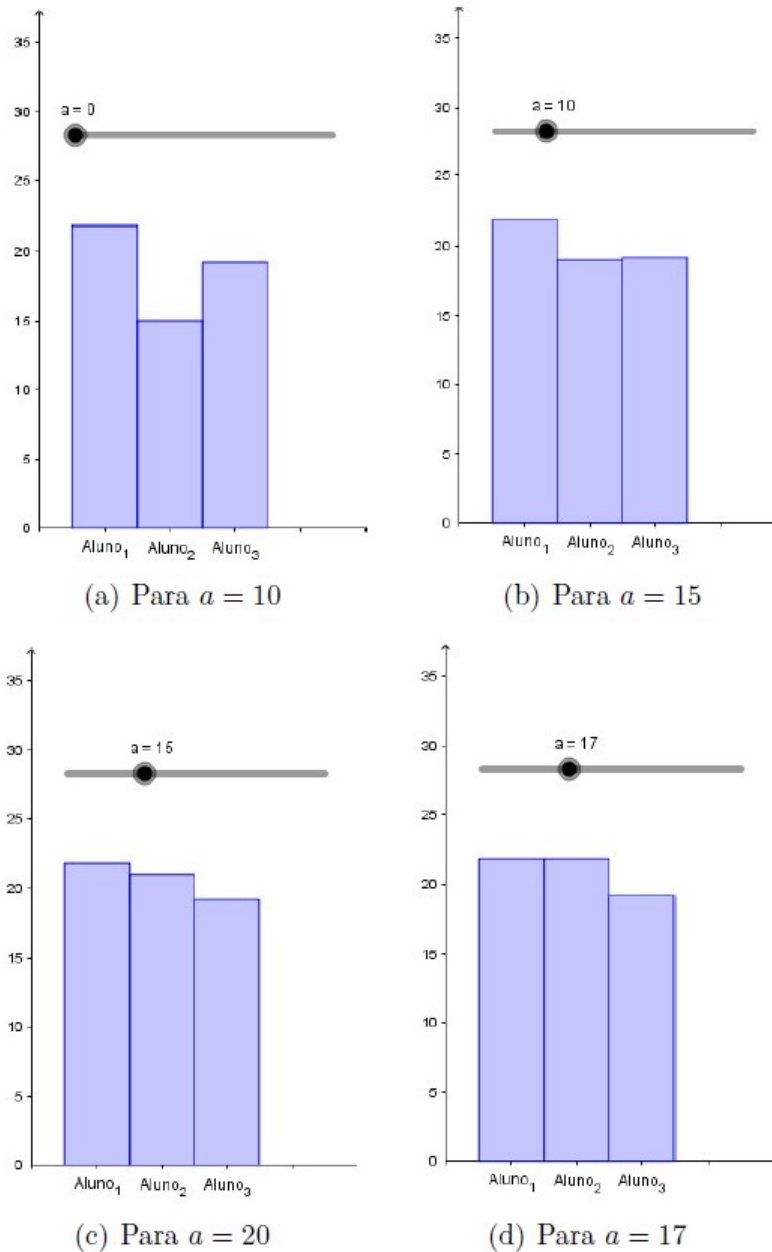
Com os testes devidos, os alunos conseguem chegar ao resultado que representa o valor mínimo, ou seja, 18. Logo a **alternativa A** é a correta.

É importante frisar com os alunos que a solução proposta não representa um método de tentativa e erro. Por meio do controle deslizante é possível identificar um conjunto solução que representa todos os valores que o candidato II precisa para vencer a competição.

Os alunos podem observar graficamente os testes obtidos no passo anterior e verificar o aumento gradual da média. Para isso, será necessário exibir um diagrama de barras. Na Caixa de Entrada digite o seguinte comando: `DiagramaDeBarras({"Aluno_1", "Aluno_2", "Aluno_3"}, {Média_1, Média_2, Média_3})`.

Alterando o valor do controle deslizante, para os mesmos valores sugeridos no passo anterior, o gráfico apresentará as seguintes configurações, para $a=0$ o diagrama visto no Gráfico 17(a), para $a=10$ o diagrama do Gráfico 17(b), para $a=15$ o diagrama do Gráfico 17(c) e para $a=17$ o diagrama do Gráfico 17(d):

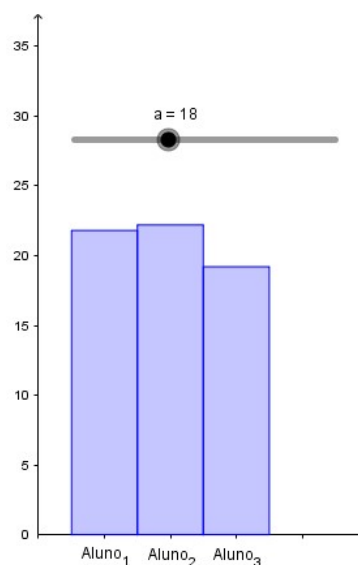
Gráfico 17 - Exibição gráfica dos resultados



Fonte: Elaborada pelo autor (2022).

Os alunos devem observar que a medida que alteram o valor do controle deslizante, a coluna central, que representa a nota média do candidato II, vai alterando o valor. O Gráfico 18 exibe o resultado para $a=18$, cuja coluna então ultrapassa a dos demais candidatos.

Gráfico 18 - Diagrama resultado final



Fonte: Elaborada pelo autor (2022).

Os alunos devem ser motivados a propor outras notas para os alunos afim de verificar o comportamento da média e efetuar novos testes.

5.4 EXEMPLO 4

SARESP/2007 (Adaptada): Após corrigir as provas de 30 alunos da mesma classe de 8^a série, a professora de Matemática anotou, em ordem crescente, as notas a eles atribuídas.

1,0	2,0	2,5	3,0	3,0	4,0	4,0	4,0	4,0	5,0
5,0	5,0	5,5	5,5	6,0	6,0	6,0	6,0	6,0	6,5
6,5	7,0	7,5	7,5	7,5	8,0	8,0	8,5	9,0	9,0

Considerando essas notas, obtenha a moda, a mediana e a média aritmética.

Objetivos da tarefa:

- Compreensão dos conceitos de Média, Moda e Mediana para dados agrupados.
- Compreender a estrutura de uma tabela de frequência com os dados agrupados e como relacioná-la às medidas de tendência central.
- Analisar os dados por meio de um histograma.

Competências e Habilidades na BNCC

- EF03MA28 e EM13MAT406.

Para a resolução proposta a seguir será utilizado o conceito de tabelas de frequências com dados agrupados. A utilização da tabela de frequências facilita a apresentação de uma amostra com grande número de dados.

Resolução proposta:

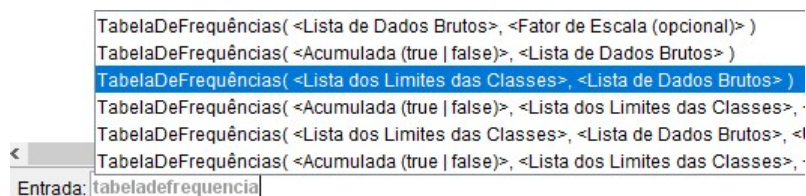
Será utilizado o *software* GeoGebra para a apresentação da resolução.

1º passo: Habilitar a planilha no GeoGebra: *Menu » Planilha*, e inserir as notas dos alunos. Após a inserção, os alunos irão selecionar os dados e com o botão direito do mouse clicar em *Criar » Lista*. Renomear para *Notas*.

2º passo: Será orientado aos alunos que as notas sejam agrupadas em 4 classes, sejam: $[0; 2,5)$, $[2,5; 5)$, $[5; 7,5)$ e $[7,5; 10]$. Logo, deve-se criar uma lista com os extremos de cada classe. Na caixa de entrada digitar: $Classes = \{0, 2.5, 5, 7.5, 10\}$. Essa lista compreende o primeiro limite inferior e os limites superiores de cada classe e representam a altura de 2,5.

3º passo: Para a inclusão da tabela de frequências, será utilizado o comando *TabelaDeFrequências(<Lista dos Limites das Classes>, <Lista de Dados Brutos>)*, onde a Lista dos Limites das Classes é a lista Classes e a Lista de Dados Brutos a lista Notas. Esse comando será digitado na Caixa de Entrada, conforme a Figura 24.

Figura 24 - Comando Tabela de Frequências no GeoGebra



Fonte: Elaborada pelo autor (2022).

Na Janela de Visualização será exibida a tabela criada, conforme a Figura 25.

Figura 25 - Visualização da Tabela de Frequências

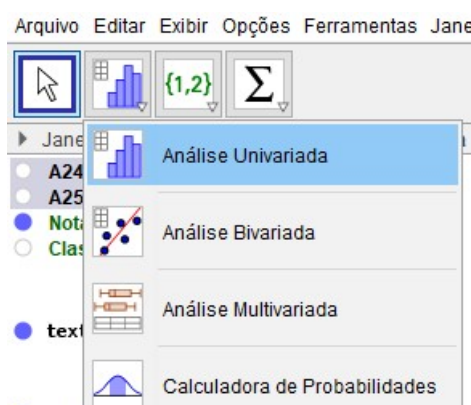
Janela de Visualização	
Intervalo	Contagem
0 – 2.5	2
2.5 – 5	7
5 – 7.5	13
7.5 – 10	8

Fonte: Elaborada pelo autor (2022).

Essa tabela é dinâmica e deve ser explicado aos alunos que qualquer alteração efetuada nos dados da Planilha, a tabela será automaticamente atualizada. A coluna Contagem representa a frequência dos dados observados em cada classe.

4^o passo: É possível criar a mesma tabela de frequências utilizando um histograma. Para isso, selecione os dados na planilha. Clique no ícone *Análise Univariada* e na Caixa de Diálogo que abrir, clique em *Analisar*. Observe a Figura 26.

Figura 26 - Menu Análise Univariada



Fonte: Elaborada pelo autor (2022).

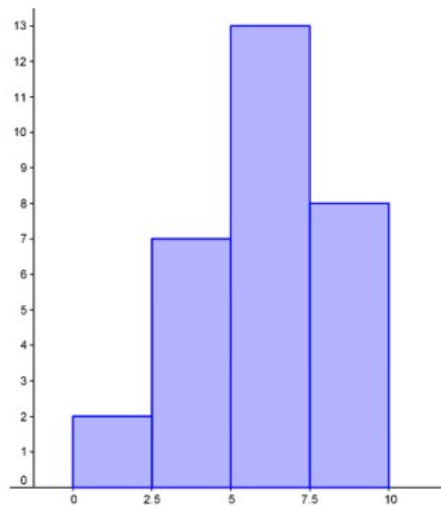
Será exibido um histograma. Oriente aos alunos a efetuarem as seguintes configurações:

- Na aba histograma:
 - Clique em Definir Classes Manualmente e escreva “0” e “2,5” em “Início” e “Largura”, respectivamente.
 - Mantenha em Regra de Classe a opção $\leq x <$
 - Habilite “Tabela de Frequências”.

- Na aba Visualização:
 - Desmarque Ajuste Automático.
 - Nas dimensões insira: -2 para x Mín, 12 para x Máx, 2,5 para *Passo em x*, 13,5 para y Max e 1 para *Passo em y*.

Será obtido o histograma do Gráfico 19:

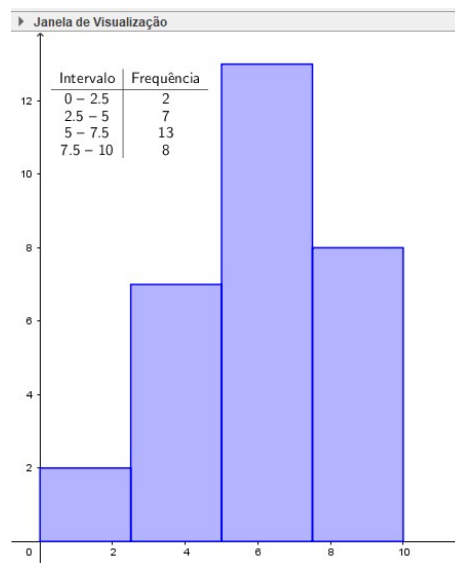
Gráfico 19 - Histograma da Tabela de Frequências



Fonte: Elaborada pelo autor (2022).

5º passo: Clique com o botão direito do mouse em cima do histograma e escolha a opção *Copiar para a Janela de Visualização*. O histograma irá aparecer na Janela de Visualização, juntamente com a tabela de frequências, conforme o Gráfico 20.

Gráfico 20 - Histograma exibido na Janela de Visualização



Fonte: Elaborada pelo autor (2022).

Mostre aos alunos que qualquer modificação das notas da Planilha, o histograma é automaticamente atualizado.

Para encontrar a média, moda e mediana será utilizado somente o conceito da frequência de dados agrupados.

6º passo: Para o cálculo da moda os alunos devem notar que a 3ª classe é a que possui mais elementos, logo é a classe modal. Porém, como pela tabela não é possível determinar qual é esse elemento, pode ser utilizado qualquer uma das fórmulas para a sua identificação. Utilizando o ponto médio será obtido:

$$PM = \frac{L_S + L_I}{2} = \frac{7,5 + 5}{2} = \frac{12,5}{2} = 6,25$$

Logo a moda é 6,25.

7º passo: Para o cálculo da média, será utilizado o conceito de média ponderada. Inicialmente, deve-se criar uma lista com as frequências de cada classe. Na caixa de entrada, inserir o seguinte comando: *Frequência(Classes, Notas)*. Será obtido uma lista que contém as frequências da tabela. Esse método de identificação é mais eficiente do que somente criar uma lista com a frequência observada, uma vez que as alterações na planilha serão atualizadas automaticamente. Renomeie essa lista para *Frequencia*.

Em seguida, na Caixa de Entrada, digite: *média(Classes, Frequencia)*. Será obtido o valor de 6. Mostre aos alunos que a referência para cada classe é o ponto médio da classe.

8º passo: Para o cálculo da mediana, digite na Caixa de Entrada: *Mediana(Classes, Frequencia)*. Será exibido o valor de 6,15.

Como complementação ao exercício, pode solicitar aos alunos que o refaçam, porém, utilizando uma tabela de frequências simples. Neste caso, verificar se houve mudança nos valores calculados e em caso afirmativo, questionar o motivo de tal resultado.

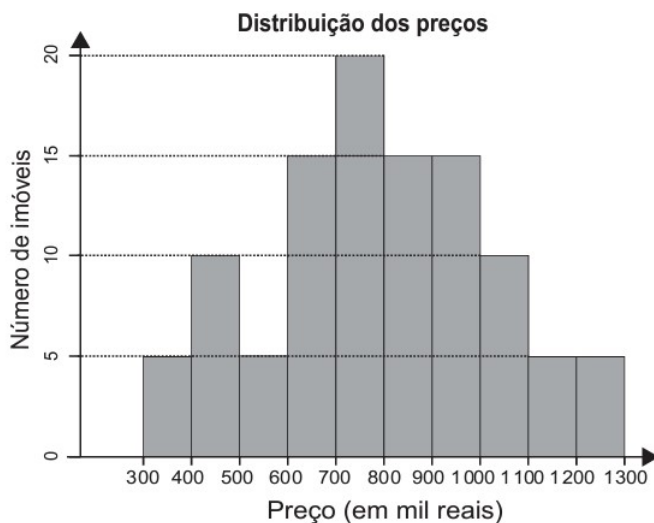
5.5 EXEMPLO 5

QUESTÃO 167 / ENEM 2021 (Caderno Amarelo): Um casal está planejando comprar um apartamento de dois quartos num bairro de uma cidade e consultou a página de uma corretora de imóveis, encontrando 105 apartamentos de dois quartos à venda no bairro desejado. Eles usaram um aplicativo da corretora para gerar a distribuição dos preços do conjunto de imóveis selecionados.

O gráfico ilustra a distribuição de frequências dos preços de venda dos apartamentos dessa lista (em mil reais), no qual as faixas de preço são dadas por]300, 400],]400, 500],]500, 600],]600, 700],]700, 800],]800, 900],]900, 1000],]1000, 1100],]1100, 1200] e]1200, 1300].

A mesma corretora anuncia que cerca de 50% dos apartamentos de dois quartos nesse bairro, publicados em sua página, têm preço de venda inferior a 550 mil reais. No

entanto, o casal achou que essa última informação não era compatível com o gráfico obtido.



Com base no gráfico obtido, o menor preço, p (em mil reais), para o qual pelo menos 50% dos apartamentos apresenta preço inferior a p é

- (A) 600.
- (B) 700.
- (C) 800.
- (D) 900.
- (E) 1000.

Objetivos da tarefa:

- a) Compreensão dos conceitos de Frequência Relativa e Frequência Acumulada.
- b) Analisar e interpretar uma tabela de frequências.
- c) Compreender a estrutura de um Gráfico de Colunas e relacioná-los às medidas de tendência central.

Competências e Habilidades na BNCC

- EF03MA27, EF04MA27 e EM13MAT406.

Para a resolução proposta a seguir, será utilizado uma tabela de frequência acumulada. Os alunos deverão coletar os dados por meio de observação do gráfico apresentado, e efetuar os passos para a obtenção do resultado e apresentação do gráfico sugerido.

Resolução proposta:

Será utilizado o *software* LibreOffice Calc para a apresentação da resolução.

1º passo: Os alunos irão digitar os valores gráfico na planilha do LibreOffice Calc. Espera-se que cheguem ao resultado mostrado na Figura 27:

Figura 27 - Inserção dos dados na planilha

	A	B
1	Intervalo	Frequência
2] 300, 400]	5
3] 400, 500]	10
4] 500, 600]	5
5] 600, 700]	15
6] 700, 800]	20
7] 800, 900]	15
8] 900, 1000]	15
9] 1000, 1100]	10
10] 1100, 1200]	5
11] 1200, 1300]	5
12		

Fonte: Elaborada pelo autor (2022).

2º passo: Será solicitado aos alunos para inserirem 3 colunas à direita da coluna Frequência com os seguintes títulos: Frequência Acumulada, Frequência Relativa e Frequência Relativa Acumulada.

3º passo: Utilizando os conceitos de Frequência Acumulada, os alunos devem preencher a respectiva coluna com os valores acumulados, ou seja, a partir da 2ª linha somar os valores da coluna Frequência com o valor anterior da Frequência Acumulada.

4º passo: Para preencher a coluna de Frequência Relativa, inicialmente deve-se criar uma soma na coluna Frequência. Na coluna de Frequência Relativa inserir a fórmula para divisão da Frequência Absoluta de cada classe pela Soma. Oriente aos alunos para utilizar o caractere \$ na identificação da célula que contém a soma, pois assim, a célula estará "travada" e será possível copiar essa fórmula para as demais células desta coluna. Assim, na célula D2, terá a fórmula $=B2/BS12$ que poderá ser copiada as demais, permanecendo a Soma (célula B12) fixa.

5º passo: Para a coluna de Frequência Relativa Acumulada, realizar os mesmos procedimentos da Frequência Acumulada, porém substituindo a coluna de Frequência pela coluna de Frequência Relativa.

A Figura 28 exibe as fórmulas inseridas para cada coluna:

Figura 28 - Planilha com as fórmulas inseridas

	A	B	C	D	E
1	Intervalo	Frequência	Frequência Acumulada	Frequência Relativa	Frequência Relativa Acumulada
2] 300, 400]	5	=B2	=B2/\$B\$12	=D2
3] 400, 500]	10	=C2+B3	=B3/\$B\$12	=E2+D3
4] 500, 600]	5	=C3+B4	=B4/\$B\$12	=E3+D4
5] 600, 700]	15	=C4+B5	=B5/\$B\$12	=E4+D5
6] 700, 800]	20	=C5+B6	=B6/\$B\$12	=E5+D6
7] 800, 900]	15	=C6+B7	=B7/\$B\$12	=E6+D7
8] 900, 1000]	15	=C7+B8	=B8/\$B\$12	=E7+D8
9] 1000, 1100]	10	=C8+B9	=B9/\$B\$12	=E8+D9
10] 1100, 1200]	5	=C9+B10	=B10/\$B\$12	=E9+D10
11] 1200, 1300]	5	=C10+B11	=B11/\$B\$12	=E10+D11
12	SOMA	=SOMA(B2:B11)			

Fonte: Elaborada pelo autor (2022).

O resultado obtido está exibido na Figura 29:

Figura 29 - Planilha com valores calculados

	A	B	C	D	E
1	Intervalo	Frequência	Frequência Acumulada	Frequência Relativa	Frequência Relativa Acumulada
2] 300, 400]	5	5	0,05	0,05
3] 400, 500]	10	15	0,10	0,14
4] 500, 600]	5	20	0,05	0,19
5] 600, 700]	15	35	0,14	0,33
6] 700, 800]	20	55	0,19	0,52
7] 800, 900]	15	70	0,14	0,67
8] 900, 1000]	15	85	0,14	0,81
9] 1000, 1100]	10	95	0,10	0,90
10] 1100, 1200]	5	100	0,05	0,95
11] 1200, 1300]	5	105	0,05	1
12	SOMA	105			

Fonte: Elaborada pelo autor (2022).

As frequências relativas podem ser configuradas para a apresentação de forma percentual. Para isto, basta selecionar os dados e clicar no menu *Formatar* » *Formato Numérico* » *Porcentagem*, ou apertar as teclas de atalho *Ctrl + Shift + 5*. Observe a Figura 30:

Figura 30 - Planilha com valores percentuais

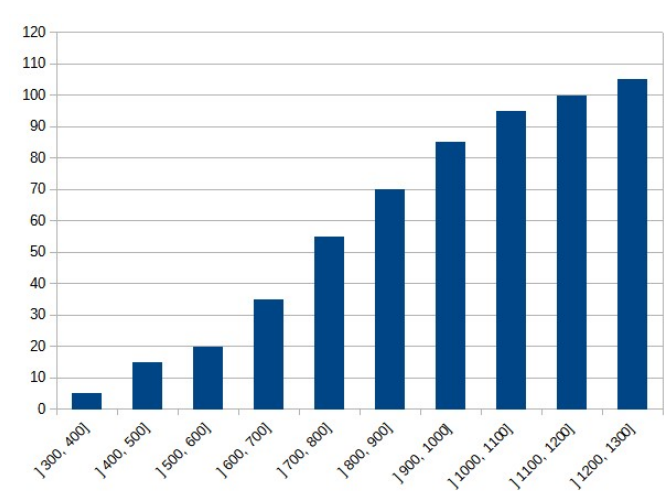
	A	B	C	D	E
1	Intervalo	Frequência	Frequência Acumulada	Frequência Relativa	Frequência Relativa Acumulada
2] 300, 400]	5	5	4,76%	4,76%
3] 400, 500]	10	15	9,52%	14,29%
4] 500, 600]	5	20	4,76%	19,05%
5] 600, 700]	15	35	14,29%	33,33%
6] 700, 800]	20	55	19,05%	52,38%
7] 800, 900]	15	70	14,29%	66,67%
8] 900, 1000]	15	85	14,29%	80,95%
9] 1000, 1100]	10	95	9,52%	90,48%
10] 1100, 1200]	5	100	4,76%	95,24%
11] 1200, 1300]	5	105	4,76%	100,00%
12	SOMA	105			

Fonte: Elaborada pelo autor (2022).

6º passo: Com a apresentação da tabela, pode-se mostrar aos alunos vários conceitos relacionados a medidas separatrizes (quartil, percentil, etc). Para a resolução do exercício, basta os alunos observarem que a frequência relativa acumulada apresenta o percentual de dados referente a soma de frequência daquela classe com as frequências das classes anteriores. Observa-se que a classe que representa que 50% dos dados é a classe]700, 800], cuja qual está a resposta do exercício, ou seja, **alternativa C**.

7º passo: Para a exibição dos dados em um gráfico com valores acumulados, solicitar aos alunos que selecionem a coluna Intervalo e a coluna Frequência Acumulada. Após, ir no Menu *Inserir » Gráfico*. Selecionar o gráfico de colunas e clicar em Finalizar. Será apresentado o Gráfico 21:

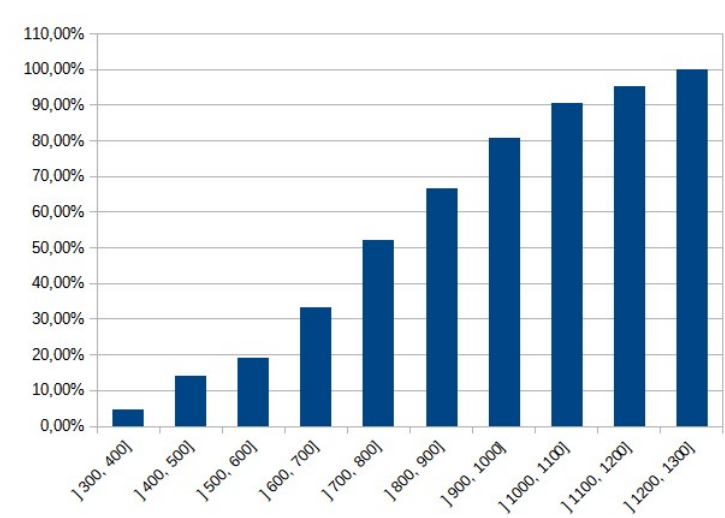
Gráfico 21 - Frequência absoluta acumulada



Fonte: Elaborada pelo autor (2022).

Neste momento deve-se mostrar aos alunos a diferença apresentada entre os gráficos. É possível também que o gráfico seja apresentado com valores percentuais. Para isso basta selecionar as colunas Intervalo e Frequência Relativa Acumulada e repetir o procedimento anterior para a inserção do gráfico. Observe o Gráfico 22:

Gráfico 22 - Frequência percentual acumulada



Fonte: Elaborada pelo autor (2022).

Mostre aos alunos que o gráfico apresenta o mesmo formato, mas o eixo Y possui escala diferente. Enquanto no primeiro gráfico é necessário o aluno entender o conceito de mediana para identificar a resposta do exercício, no segundo gráfico essa identificação é imediata, pois basta identificar a coluna que apresenta o percentual de 50%.

5.6 EXEMPLO 6

QUESTÃO 169 / ENEM 2021 (Caderno Amarelo): O quadro apresenta o número de terremotos de magnitude maior ou igual a 7, na escala Richter, ocorridos em nosso planeta nos anos de 2000 a 2011.

Ano	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Terremotos	15	16	13	15	16	11	11	18	12	17	24	20

Disponível em: <https://earthquake.usgs.gov/earthquakes/browse/m7-world.php>. Acesso em: 13 ago. 2012 (adaptado).

Um pesquisador acredita que a mediana representa bem o número anual típico de terremotos em um período. Segundo esse pesquisador, o número anual típico de terremotos de magnitude maior ou igual a 7 é

- (A) 11.
- (B) 15.

- (C) 15,5.
- (D) 15,7.
- (E) 17,5.

Objetivos da tarefa:

- a) Compreensão dos conceitos de Mediana.
- b) Analisar e interpretar um Diagrama de Caixa.

Competências e Habilidades na BNCC

- EF02MA22, EF03MA27, EF05MA24 e EM13MAT407

Na resolução proposta a seguir, será exposto a utilização de uma ferramenta do *software* GeoGebra. Será apresentado o Diagrama de Caixa para a interpretação das medidas de tendência central, entre elas a mediana, e alguns conceitos complementares.

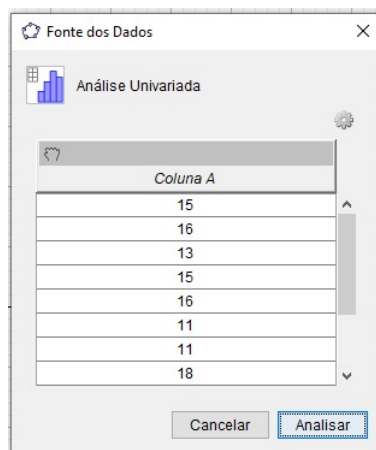
Resolução proposta:

Será utilizado o *software* GeoGebra para a apresentação da resolução e aprofundamento sobre as medidas de tendência central.

1º passo: Inicialmente deve-se habilitar a Planilha do GeoGebra e inserir os valores das magnitudes. Não há necessidade de incluir os anos.

2º passo: Para o cálculo imediato da Mediana, será utilizado a Análise Univariada, conforme apresentado na Figura 26. O objetivo é apontar a facilidade do *software* para os cálculos e acrescentar análises aos resultados encontrados. Para isso deve-se selecionar os dados, clicar no ícone correspondente e confirmar na Caixa de Diálogo, como aparece na Figura 31.

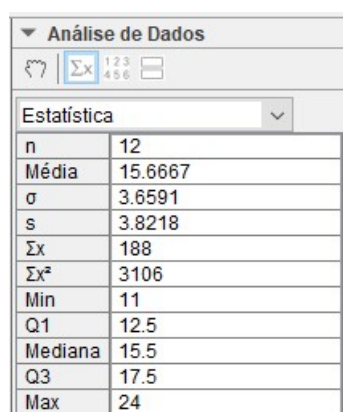
Figura 31 - Caixa de Confirmação Análise Univariada



Fonte: Elaborada pelo autor (2022).

3º passo: Em seguida na janela Análise de Dados, clicar em Exibir Estatística. Será exibido uma caixa com as principais estatísticas para os dados apresentados. Observe a Figura 32:

Figura 32 - Caixa de Diálogo análise de dados



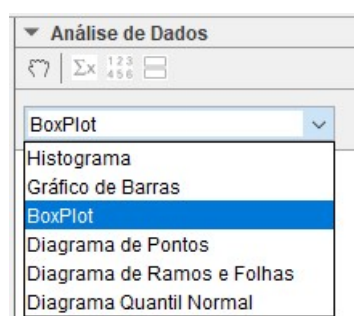
Análise de Dados	
Estadística	
n	12
Média	15.6667
σ	3.6591
s	3.8218
Σx	188
Σx^2	3106
Min	11
Q1	12.5
Mediana	15.5
Q3	17.5
Max	24

Fonte: Elaborada pelo autor (2022).

A apresentação dos cálculos é realizada de forma direta e deve ser utilizada para que os alunos possam conferir os dados e analisar o comportamento destes. Os alunos irão verificar o valor da mediana, 15,5; bem como outras medidas de tendência central e de variabilidade. Portanto, a resposta correta é a **alternativa C**.

4º passo: Será construído o Diagrama de Caixa e demonstrado sua importância para a análise de Dados. Na janela Análise de Dados, clicar no seletor e escolher *Box Plot* (Diagrama de Caixa), conforme a Figura 33.

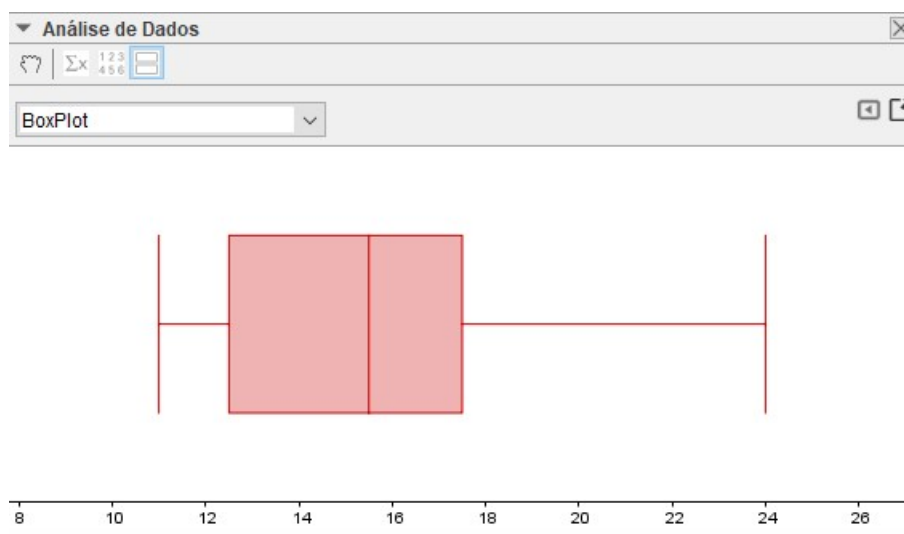
Figura 33 - Seleção do Diagrama de Caixa (*Box Plot*)



Fonte: Elaborada pelo autor (2022).

Após a confirmação será apresentado, logo abaixo o Diagrama de Caixa. A Figura 34 exibe o resultado.

Figura 34 - Visualização do Diagrama de Caixa calculado



Fonte: Elaborada pelo autor (2022).

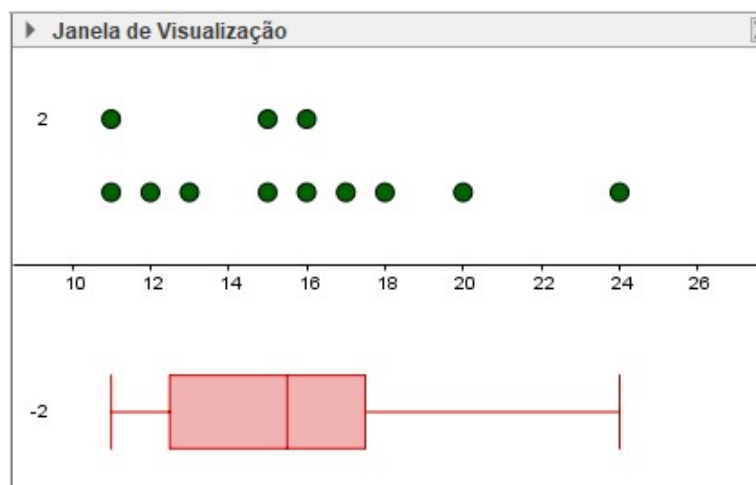
Deve-se explicar aos alunos o que significa cada parte do diagrama. O traço inicial e o final, que estão ligados pela linha horizontal são os extremos da amostra, ou seja, o menor e o maior valor, respectivamente.

No retângulo que forma a caixa, a primeira linha vertical representa o valor do primeiro quartil, a segunda linha representa o valor da mediana e a terceira linha o terceiro quartil. Nesse momento os alunos são motivados a conferirem o diagrama com a estatística exibida anteriormente, observando esses valores com o eixo enumerado logo abaixo do diagrama.

Após a identificação dos valores, deve-se mostrar aos alunos o que essa denominação implica na amostra apresentada. Essa divisão representa que cada intervalo possui a mesma quantidade de elementos, ou seja, como a amostra tem 12 elementos, cada intervalo possui 3 valores. Isto possibilita entender onde os dados estão mais concentrados (mais valores e, portanto, as caixas ou as linhas serão menores) ou mais dispersos (menos valores e, portanto, as caixas ou linhas serão maiores).

Caso haja dificuldade para os alunos identificarem essa distribuição, pode-se inserir um gráfico de pontos para exibir os dados da amostra. Inicialmente na Janela Análise de Dados, clicar em cima do diagrama e *Copiar para a Janela de Visualização*. Logo após selecionar os dados na janela Planilha, clicar com o botão direito e escolher *Criar » Lista*. Renomear a lista para Magnitude. Na caixa de diálogo escreva *DiagramaDePontos(Magnitude)*. Esse procedimento já foi realizado em exercícios anteriores. O diagrama está apresentado na Figura 35:

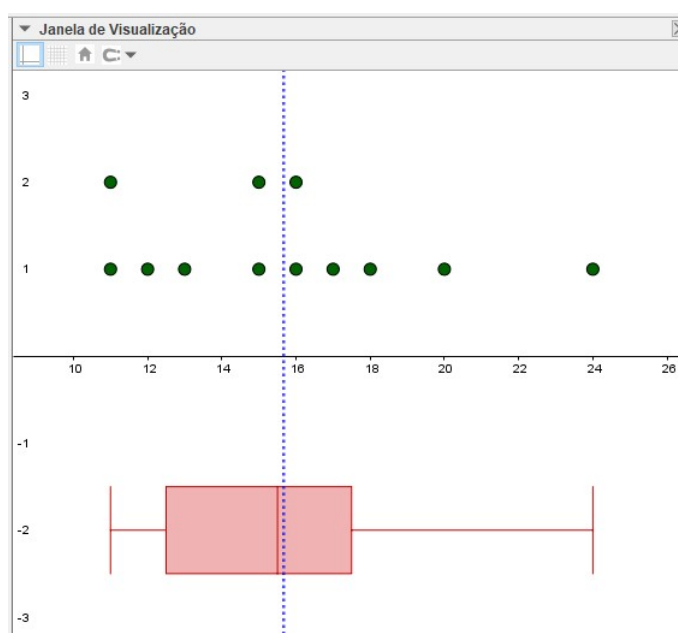
Figura 35 - Visualização do Diagrama de Pontos inserido



Fonte: Elaborada pelo autor (2022).

Os alunos poderão identificar na Figura 35 a distribuição dos dados, verificando que cada quartil possui a mesma quantidade de elementos. Também podem incluir uma reta com o valor da média dos dados. Deverá para isso, observando a tabela de estatística, onde a média é 15,6667, digitar na Caixa de Entrada $x = 15,6667$. Será obtido o resultado exibido na Figura 36:

Figura 36 - Identificação da Média da Amostra



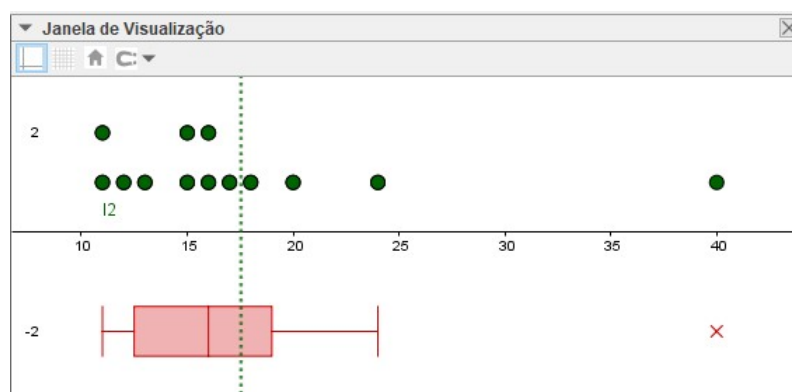
Fonte: Elaborada pelo autor (2022).

Os alunos devem notar que a média é superior ao valor da mediana.

5º passo: Outro conceito importante para os alunos são os *outliers*, ou seja, os valores discrepantes da amostra. Será solicitado aos alunos para acrescentar um valor, por exemplo 40, na planilha de dados e observar o que acontece com o diagrama.

Realizando as alterações sugeridas, os cálculos das estatísticas modificam, juntamente com apresentação dos diagramas. Observe a Figura 37:

Figura 37 - Identificação de *Outliers*



Fonte: Elaborada pelo autor (2022).

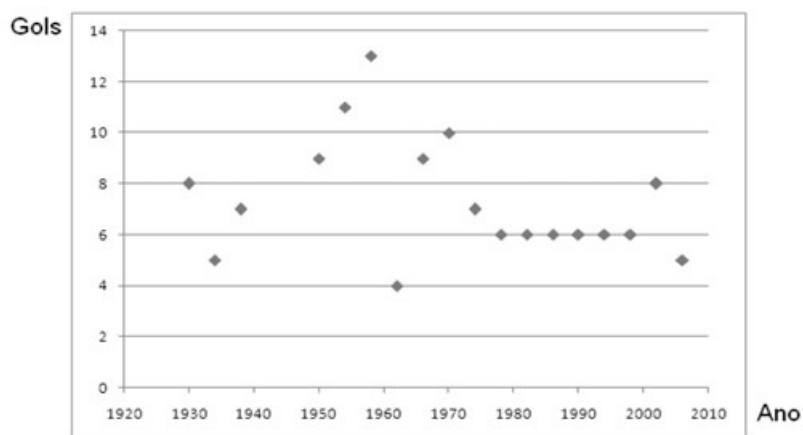
Observe que o novo valor inserido, 40, não está dentro do diagrama de caixa. Deve-se explicar aos alunos esse motivo. O diagrama de caixa identifica os valores discrepantes de uma amostra por meio de um “X” que se localiza além dos limites, inferior e/ou superior, e representam valores que tem uma grande variabilidade em relação a média. Logo esses valores, quando identificados, devem ser tratados para que não interfiram no resultado da análise. Quando uma amostra apresenta *outliers*, a estatística de posição mais indicada para a representação dos dados é a mediana, ao invés da média, pois essa não sofre grandes alterações devido a esses valores discrepantes, como se pode observar com o distanciamento entre a reta que representa a média e a mediana do Diagrama de Caixa da Figura 37.

Finalize o exercício solicitando aos alunos, que realizem novas modificações nos valores da planilha, para identificar o comportamento do diagrama. Pode sugerir também a visualização dos dados por meio de um gráfico de linhas.

5.7 EXEMPLO 7

QUESTÃO 167 / ENEM 2010 (Caderno Amarelo): O gráfico apresenta a quantidade de gols marcados pelos artilheiros das Copas do Mundo desde a Copa de 1930 até a de 2006.

Quantidades de Gols dos Artilheiros das Copas do Mundo



Disponível em: <http://www.suapesquisa.com>. Acesso em: 23 abr. 2010 (adaptado).

A partir dos dados apresentados, qual a mediana das quantidades de gols marcados pelos artilheiros das Copas do Mundo?

- (A) 6 gols
- (B) 6,5 gols
- (C) 7 gols
- (D) 7,3 gols
- (E) 8,5 gols

Objetivos da tarefa:

- a) Compreensão dos conceitos dos Quartis.
- b) Apresentar as características geométricas das estatísticas calculadas.
- c) Compreender a estrutura de um Diagrama de Pontos e como relacioná-lo às medidas de tendência central.

Competências e Habilidades na BNCC

- EF01MA21, EF05MA24, EF08MA23 e EF09MA22.

Para a resolução proposta a seguir, além do cálculo da mediana, os alunos serão motivados a calcular os demais quartis da amostra e representar graficamente cada um deles.

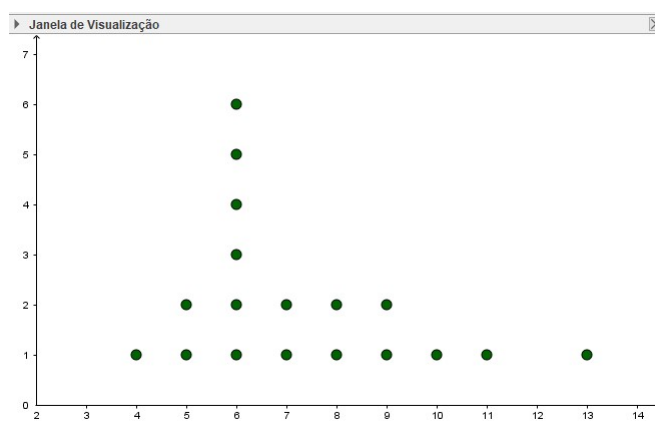
Resolução proposta:

Será utilizado o *software* GeoGebra para a apresentação da resolução e aprofundamento sobre os quartis.

1º passo: Criar uma lista com os valores observados no gráfico. Não será utilizado desta vez a planilha do GeoGebra. Na caixa de Entrada digitar: $Gols=\{8, 5, 7, 9, 11, 13, 4, 9, 10, 7, 6, 6, 6, 6, 6, 8, 5\}$, onde *Gols* será o nome da lista criada.

2º passo: Em seguida criar um Diagrama de Pontos com a lista *Gols*. Na caixa de Entrada digite: $DiagramaDePontos(Gols)$. Renomeie o diagrama para *Gols_1*. Será apresentado o Diagrama do Gráfico 23.

Gráfico 23 - Diagrama de Pontos com os gols marcados



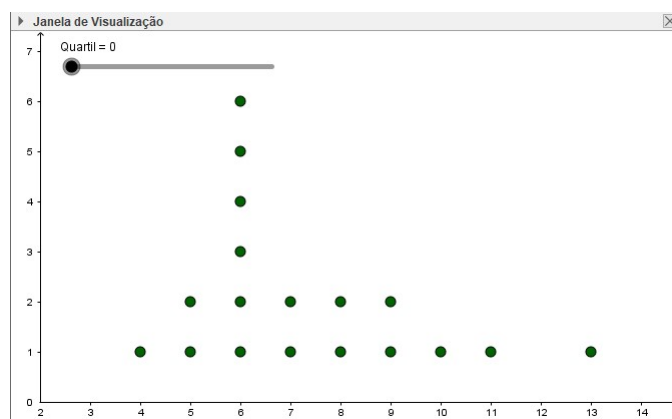
Fonte: Elaborada pelo autor (2022).

3º passo: Para o estudo dos quartis será utilizado um controle deslizante. Na barra de ferramentas clique em Controle Deslizante, e na caixa de propriedades insira:

- Opção: Número
- Nome: Quartil
- Intervalo:
 - Mínimo: 0
 - Máximo: 4
 - Incremento: 1

Os valores do Controle Deslizante representarão o respectivo quartil, quando selecionado. Após confirmar as propriedades será exibido na Janela de Visualização o Controle Deslizante, conforme o Gráfico 24. Altere seu valor para 0.

Gráfico 24 - Diagrama com o Controle deslizante



Fonte: Elaborada pelo autor (2022).

O controle deslizante irá variar de 0 a 4, e identificará os elementos que estão em cada quartil. É importante mostrar aos alunos os conceitos de quartil.

4º passo: Serão criados quatro diagramas de pontos representando, cada um, o quartil correspondente. Para isso, será utilizado o comando *ParteDaLista*.

O comando *ParteDaLista* seleciona parte de uma lista de acordo com um comprimento estabelecido. Sua sintaxe é: *ParteDaLista(<Lista>, <Posição Inicial>, <Posição Final>)*. Assim, *<Lista>* representa o nome da nossa lista, no caso *Gols_1*, *<Posição Inicial>* representa a posição do elemento que será iniciada a nova lista, ou seja, o primeiro e, *<Posição Final>* a posição do elemento final da lista. Será utilizado o comando *Comprimento*, para identificar o tamanho da lista. Para o primeiro quartil, deve-se dividir o comprimento da lista *Gols_1* por 4, para o segundo quartil, dividir o comprimento da lista por 2, para o terceiro quartil, dividir o comprimento da lista por 4 e multiplicar por 3, e o quarto quartil representa todos os elementos da lista.

A sintaxe para criação será:

$$Q1 = \text{ParteDaLista}(\text{Gols_1}, 1, \text{Comprimento}(\text{Gols_1})/4)$$

$$Q2 = \text{ParteDaLista}(\text{Gols_1}, 1, \text{Comprimento}(\text{Gols_1})/2)$$

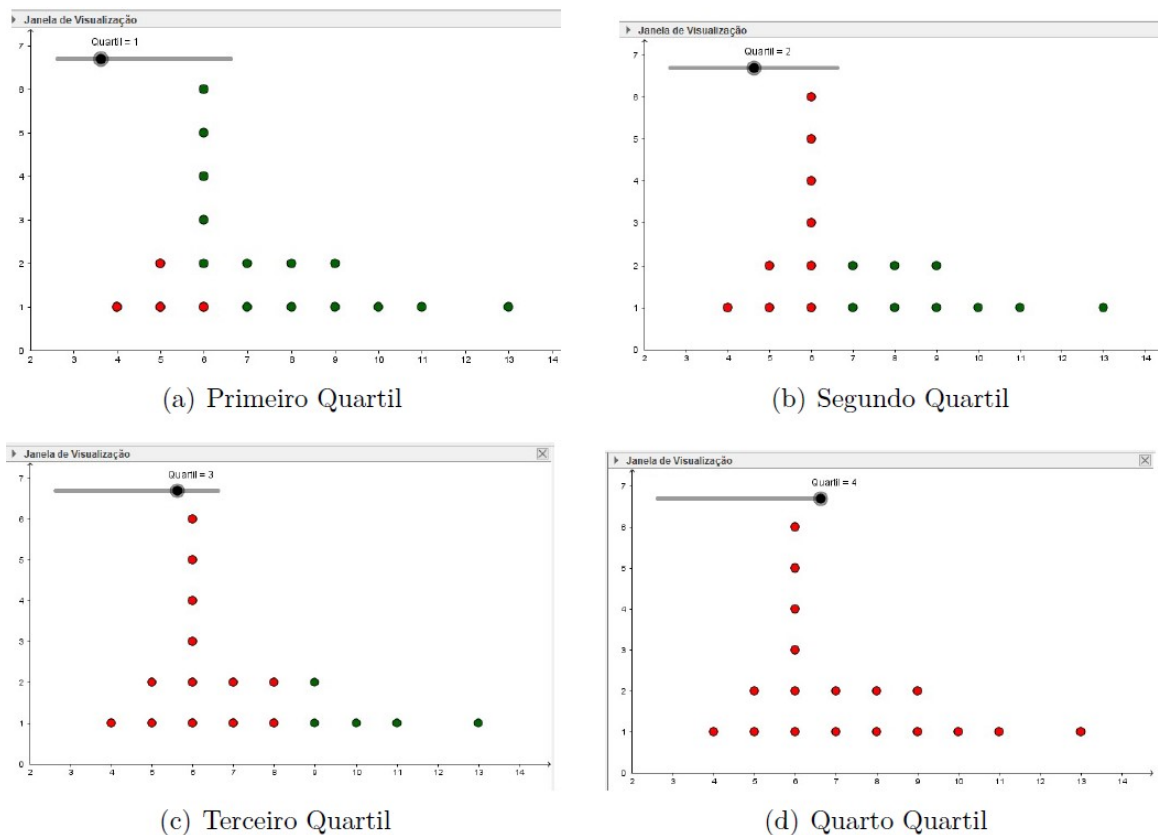
$$Q3 = \text{ParteDaLista}(\text{Gols_1}, 1, 3 * \text{Comprimento}(\text{Gols_1})/4)$$

$$Q4 = \text{ParteDaLista}(\text{Gols_1}, 1, \text{Comprimento}(\text{Gols_1}))$$

Na caixa propriedades dos diagramas criados, alterar a cor padrão para Vermelho e na aba Avançado, em *Condição para Exibir Objeto(s)* digite *Quartil = 1*, para o primeiro quartil, *Quartil = 2* para segundo quartil e assim sucessivamente.

5º passo: Ao movimentar o controle deslizante, o diagrama de pontos mostrará quais são os elementos que compõem cada quartil. Observe a Figura 38.

Figura 38 - Exibição dos Quartis



Fonte: Elaborada pelo autor (2022).

Observando a Figura 38(b) que representa que 50% dos dados estão abaixo do valor 6 e, como a amostra possui quantidade par, a mediana será o ponto médio entre 6 e 7, logo:

$$Me = \frac{6 + 7}{2} = 6,5$$

Assim, a mediana é 6,5 e a **alternativa B** é a correta.

Esse método mostra aos alunos os elementos que compõe cada quartil, porém não identifica, com precisão qual é o valor do quartil. Para facilitar essa exibição, será criado uma janela de cálculo, que mostra o valor de cada quartil.

6º passo: Na caixa de entrada, digite os seguintes comandos:

$Q_1 = \text{Quartil1}(\text{Gols})$

$Q_2 = \text{Mediana}(\text{Gols})$

$Q_3 = \text{Quartil3}(\text{Gols})$

$Q_4 = \text{Máximo}(\text{Gols})$

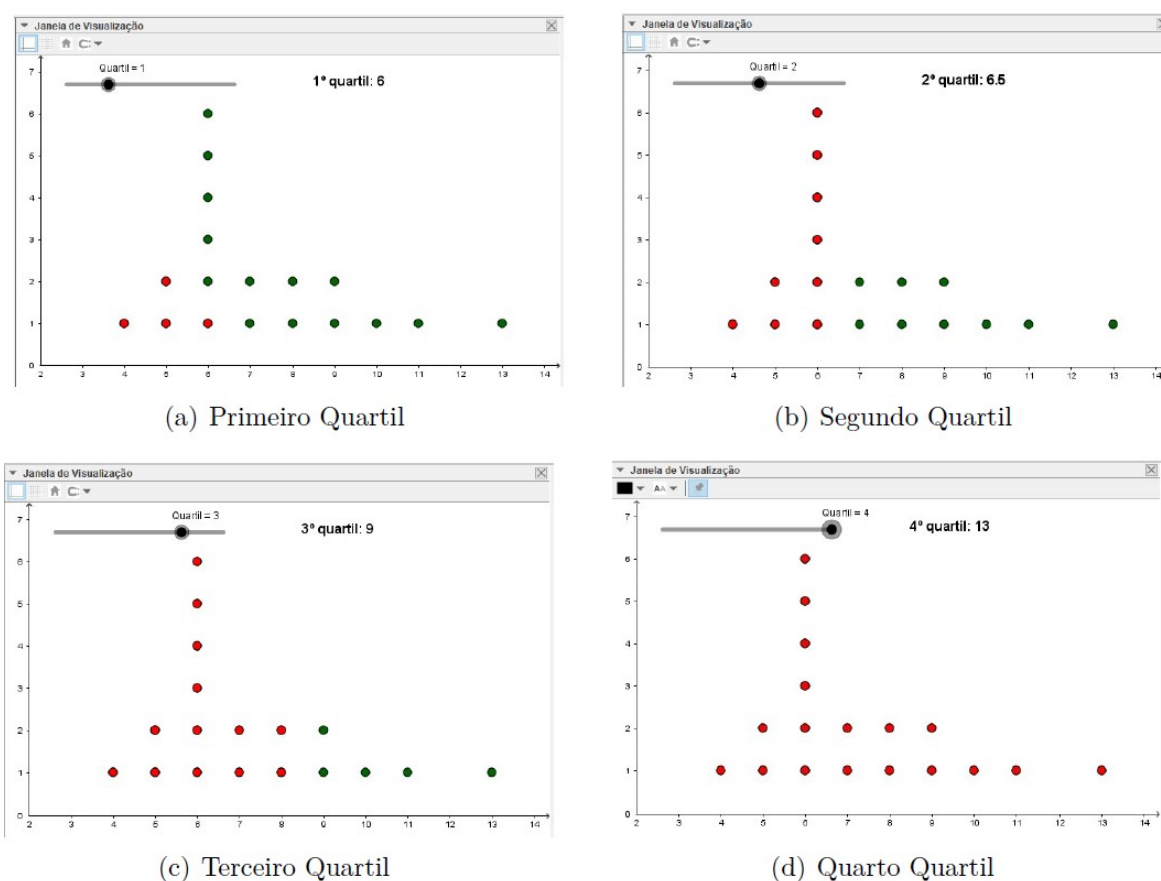
Essas serão as variáveis que irão indicar o valor de cada quartil. Neste momento mostre aos alunos que o 2º quartil é a mediana dos dados, pois identifica que 50% dos

elementos estão abaixo dela, bem como o 4º quartil é o último valor do rol, já que representa que 100% dos elementos estão abaixo.

7º passo: Na barra de ferramentas selecione a opção *Texto* e no campo *Editar* escreva 1º quartil: e selecione na opção objetos Q_1 . Repita o mesmo procedimento para os demais quartis correspondentes e sobreponha os textos.

8º passo: Nas propriedades de cada Texto, na guia Avançado, repita o procedimento do 4º passo, ou seja, para cada texto a condição para exibir será quando o controle deslizante estiver no quartil correspondente. Assim, o resultado será exibido conforme a Figura 39.

Figura 39 - Identificação dos Quartis

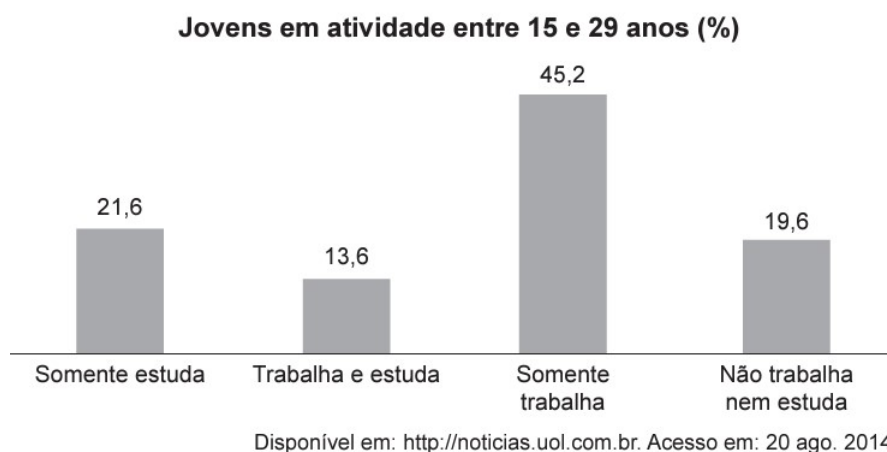


Fonte: Elaborada pelo autor (2022).

Espera-se que os alunos compreendam os conceitos relacionados aos quartis. Como complemento, pode-se solicitar aos alunos que repitam os passos anteriores aplicando o conceito de percentis. Sugere-se, neste caso, que seja incluído uma amostra com mais elementos.

5.8 EXEMPLO 8

QUESTÃO 155 / ENEM 2020 (Caderno Amarelo): A Pesquisa Nacional por Amostra de Domicílios (Pnad) é uma pesquisa feita anualmente pelo IBGE, exceto nos anos em que há Censo. Em um ano, foram entrevistados 363 mil jovens para fazer um levantamento sobre suas atividades profissionais e/ou acadêmicas. Os resultados da pesquisa estão indicados no gráfico.



De acordo com as informações dadas, o número de jovens entrevistados que trabalha é

- (A) 114 708.
- (B) 164 076.
- (C) 213 444.
- (D) 284 592.
- (E) 291 582.

Objetivos da tarefa:

- a) Compreensão dos conceitos de frequência absoluta.
- b) Compreender a estrutura de um Gráfico de Setores e relacioná-lo às frequências relativas e absolutas.

Competências e Habilidades na BNCC

- EF07MA37, EF08MA23 e EF09MA22.

A proposta de resolução apresentada a seguir utilizará uma tabela com os dados categóricos exibidos em percentual e cálculo da frequência absoluta dos mesmos. Em seguida os dados serão apresentados por meio de um gráfico de setores.

Resolução proposta:

Será utilizado o *software* LibreOffice Calc para a apresentação da resolução.

1º passo: Os alunos irão digitar na planilha eletrônica os dados obtidos do gráfico. Espera-se que cheguem ao resultado da Figura 40:

Figura 40 - Dados digitados na planilha

	A	B
1	Atividades	%
2	Somente Estuda	21,6
3	Trabalha e Estuda	13,6
4	Somente Trabalha	45,2
5	Não trabalha nem Estuda	19,6

Fonte: Elaborada pelo autor (2022).

2º passo: Em uma célula qualquer, por exemplo B8, os alunos irão inserir o total de jovens entrevistados, 363000.

3º passo: Será solicitado aos alunos que insiram uma nova coluna, com o título de Frequência Absoluta. Na célula C2 dessa coluna, os alunos devem inserir a fórmula $=B2/100*\$C\8 . O sinal de \$ incluído na fórmula serve para travar seu conteúdo e ao copiar a fórmula para outras células, não alterar o valor. Deve-se explicar aos alunos que esse é o motivo de se colocar o sinal de \$ na fórmula. Para as células restantes, ou seja de C3 a C5, deve-se arrastar o conteúdo de C2. Para isso, posiciona-se o mouse em cima do canto inferior direito, o ponteiro irá assumir a forma de uma cruz, e arrasta-se para as células pretendidas. O *software* calculará os resultados, que estão exibidos na Figura 41.

Figura 41 - Cálculo dos resultados

	A	B	C
1	Atividades	%	Freq. Absoluta
2	Somente Estuda	21,6	78408
3	Trabalha e Estuda	13,6	49368
4	Somente Trabalha	45,2	164076
5	Não trabalha nem Estuda	19,6	71148
6			
7			
8	Jovens Entrevistados	363000	
9			

Fonte: Elaborada pelo autor (2022).

4º passo: Para encontrar o número de jovens entrevistados que trabalham, deve-se somar os valores referentes as informações contidas em C3 e C4, 49368 e 164076,

respectivamente. Assim, na célula B9, digitar $=C3+C4$, obtendo a resposta do exercício, conforme Figura 42. Logo, a quantidade de jovens entrevistados que trabalham é de 213.444, e a resposta correta é a **alternativa C**.

Figura 42 - Cálculo da quantidade de jovens que trabalham

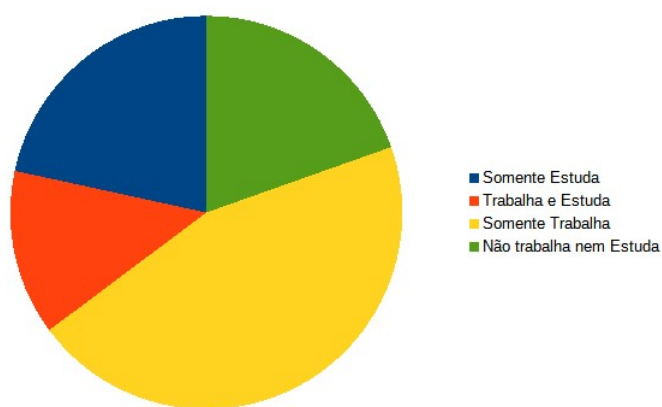
	A	B	C
1	Atividades	%	Freq. Absoluta
2	Somente Estuda	21,6	78408
3	Trabalha e Estuda	13,6	49368
4	Somente Trabalha	45,2	164076
5	Não trabalha nem Estuda	19,6	71148
6			
7			
8	Jovens Entrevistados	363000	
9	Jovens que Trabalham	213444	

Fonte: Elaborada pelo autor (2022).

5º passo: A partir deste momento os alunos irão utilizar gráficos para a interpretação de resultados. Como as informações estão se referindo a variáveis categorizadas é importante mostrar que o gráfico mais indicado é o de setores, ou seja, gráfico de pizza.

6º passo: Para a construção do gráfico de setores (ou gráfico de pizza) os alunos não devem utilizar o percentual dado pelo exercício e sim o valor absoluto. Deve-se orientar aos alunos que o *software* calcula o percentual, sem ter a necessidade de cálculos manuais para a sua obtenção. Os alunos devem selecionar a coluna Atividades, e apertando a tecla *CTRL* selecionar a coluna Frequência Absoluta. Em seguida clicar no menu *Inserir » Gráfico*. Escolher o tipo *PIZZA* e clicar em Finalizar. Será obtido o Gráfico 25.

Gráfico 25 - Gráfico de Pizza

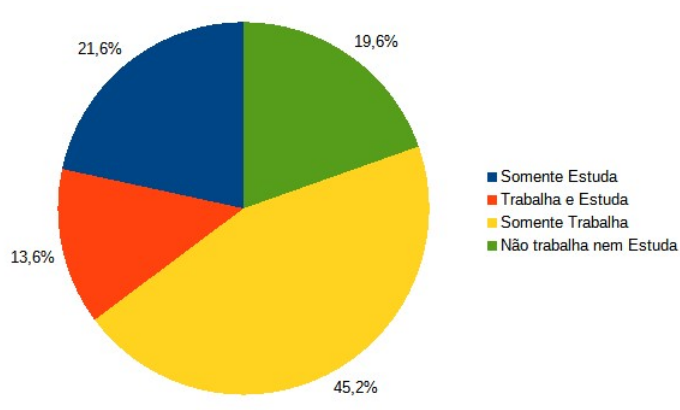


Fonte: Elaborada pelo autor (2022).

Configurações adicionais de formatos podem ser realizados de acordo com os estudantes.

7º passo: Para inserir os rótulos com os percentuais, clicar em cima do gráfico e no menu suspenso escolher *Formatar rótulo de dados...* Em atributos de texto, marcar *Valor como porcentagem*. Em *Formato de Porcentagem* desmarcar a opção *Formato de origem* e em opções aumentar para uma casa decimal. Confirme as alterações clicando em OK e será exibido o Gráfico 26.

Gráfico 26 - Configuração do Gráfico de Pizza



Fonte: Elaborada pelo autor (2022).

Os alunos devem perceber que mesmo utilizando somente os valores absolutos, o *software* calcula automaticamente o percentual de cada categoria. Outra observação importante é com relação a visualização do gráfico de setores (pizza) o qual é o mais indicado para indicar percentual de variáveis categorizadas.

6 INTERPRETAÇÃO DE GRÁFICOS

Em estatística uma das melhores maneiras de mostrar os resultados é através de gráficos. Bastante utilizado em meios de comunicação e periódicos, os gráficos permitem uma análise sintética de determinada pesquisa, seja para mostrar o resultado de um experimento ou dados de uma população. Entretanto, os gráficos que podem ser facilmente utilizados também servem para enganar.

O objetivo desse capítulo será mostrar alguns gráficos que foram exibidos na mídia e que apresentam algum erro de interpretação e, solicitar aos alunos para refazê-los de forma a corrigir o gráfico original. Abaixo será listado algumas formas mais comuns em que as visualizações podem ser enganosas e todas elas foram retiradas de algum meio de comunicação.

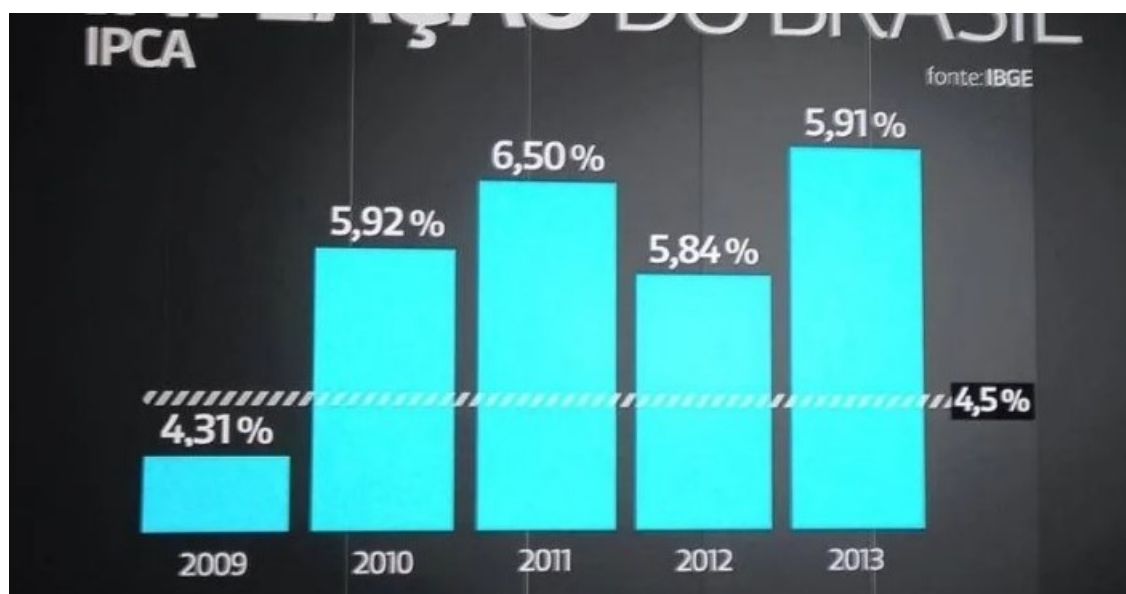
Competências e Habilidades na BNCC

- EF06MA31, EF09MA21, EM13MAT102

6.1 GRÁFICO 1

Observe a Figura 43 que apresenta um gráfico exibido por uma rede de comunicação. Inicialmente os alunos devem ser motivados a identificar e discutir os erros apresentados no gráfico.

Figura 43 - Inflação no Brasil 2009 a 2013



Fonte: Disponível em: <https://gizmodo.uol.com.br/mentir-visualizacao-dados/> (Acesso em 03.jul.2022)

Após a identificação das divergências apresentadas, os alunos irão refazer o gráfico, corrigindo a sua apresentação. Para isso irão utilizar o *software* LibreOffice Calc.

Os alunos devem digitar na coluna A da planilha os anos do gráfico acima e na coluna B os dados relacionados ao percentual. Observe o resultado obtido na Figura 44:

Figura 44 - Dados do gráfico digitados na planilha

	A	B
1	2009	4,31%
2	2010	5,92%
3	2011	6,50%
4	2012	5,84%
5	2013	5,91%

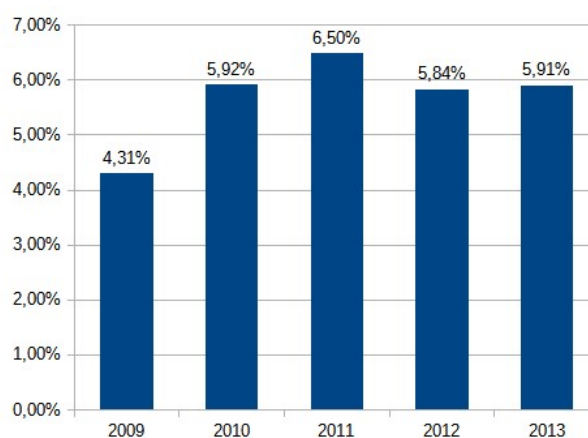
Fonte: Elaborada pelo autor (2022).

Para incluir o gráfico deve-se selecionar os dados da coluna B e clicar em *Inserir » Gráfico*, e seguir as seguintes orientações:

1. Em **Tipo de gráfico**, selecionar Coluna e clicar em Próximo;
2. Em **Intervalo de dados**, selecionar Série de dados em colunas e deixar desmarcado as demais caixas de seleção;
3. Em **Série de dados**, clicar no ícone em frente a Categorias e selecionar o intervalo A1 : A5, que representará os rótulos do eixo X;
4. Em **Elementos do gráfico**, desmarcar a opção Exibir legenda;
5. Clicar em **Finalizar**.
6. Em cima do gráfico, clicar com o botão direito do mouse e escolher a opção Inserir rótulos de dados.

O resultado será exibido conforme o Gráfico 27:

Gráfico 27 - Gráfico da Inflação 2009 a 2013
modificado



Fonte: Elaborada pelo autor (2022).

A formatação do gráfico fica a critério do aluno e para finalizar a atividade deve ser solicitado a comparação entre o gráfico encontrado e o gráfico exibido.

Nesse caso específico, dias após a divulgação do gráfico de forma irregular, o meio de comunicação efetuou a correção do gráfico e exibiu o resultado de forma mais clara. Observe a Figura 45:

Figura 45 - Correção do gráfico pela emissora



Fonte: Disponível em: <https://gizmodo.uol.com.br/mentir-visualizacao-dados/> (Acesso em 03.jul.2022)

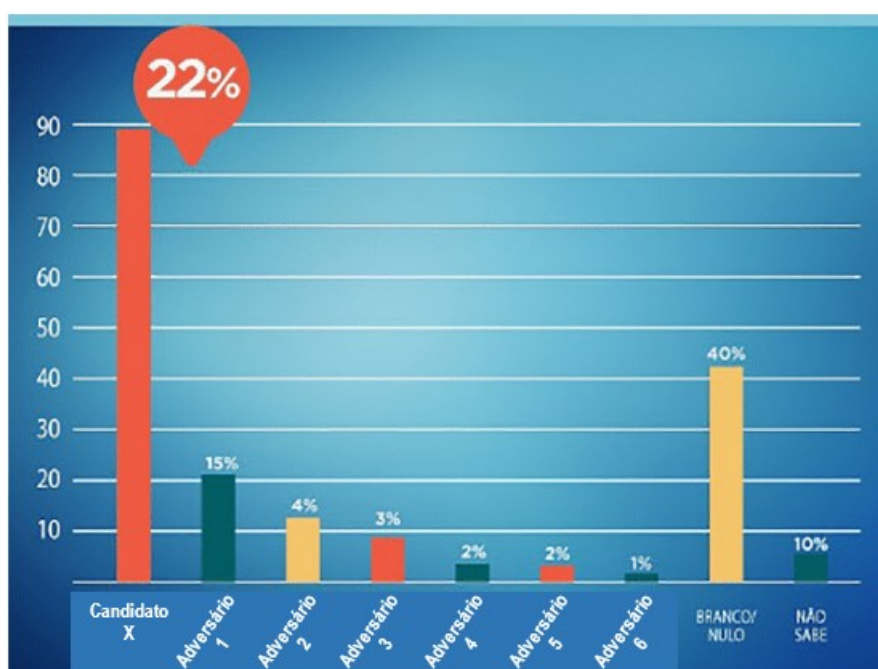
6.2 GRÁFICO 2

Em campanhas eleitorais é comum a divulgação da intenção de votos dos candidatos. Nesse exemplo, para fins somente didáticos, optou-se pela não divulgação dos nomes dos

candidatos da pesquisa.

Quando a equipe da campanha do candidato X divulgou a pesquisa da eleição, conforme a Figura 46, verificou-se inúmeros erros. A porcentagem do candidato, apesar de ser 22%, no gráfico está representada próxima ao valor de 90%. Além disso, o valor que representa o número de votos Brancos/Nulos é aproximadamente o dobro da intenção de voto do candidato X é representada pela metade da intenção do voto do candidato X . Observa-se também que a porcentagem dos demais adversários não estão corretamente relacionadas com o eixo y .

Figura 46 - Divulgação de intenção de voto do Candidato X



Fonte: Disponível em:

<https://noticias.uol.com.br/politica/eleicoes/2018/noticias/2018/05/30/psdb-sp-divulga-grafico-desproporcional-de-doria-e-tira-do-ar-apos-criticas.htm> (Acesso em 03.jul.2022)

Para a correção do gráfico, os alunos devem lançar os valores extraídos da Figura 46, na planilha do LibreOffice Calc, sendo a coluna A o nome do Candidato e a coluna B , o percentual apresentado. Deve-se obter o resultado da Figura 47:

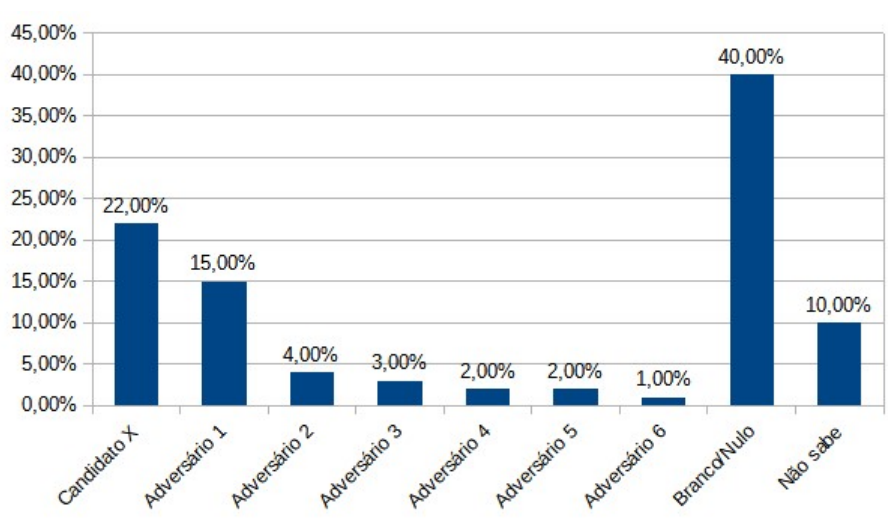
Figura 47 - Inserção de dados na planilha

	A	B
1	Candidato X	22,00%
2	Adversário 1	15,00%
3	Adversário 2	4,00%
4	Adversário 3	3,00%
5	Adversário 4	2,00%
6	Adversário 5	2,00%
7	Adversário 6	1,00%
8	Branco/Nulo	40,00%
9	Não sabe	10,00%

Fonte: Elaborada pelo autor (2022).

Para a inserção do gráfico, os alunos devem selecionar os dados da planilha e clicar no menu *Inserir » Gráfico*. Solicite para retirar a legenda do gráfico e será obtido o Gráfico 28:

Gráfico 28 - Gráfico da intenção de votos corrigido



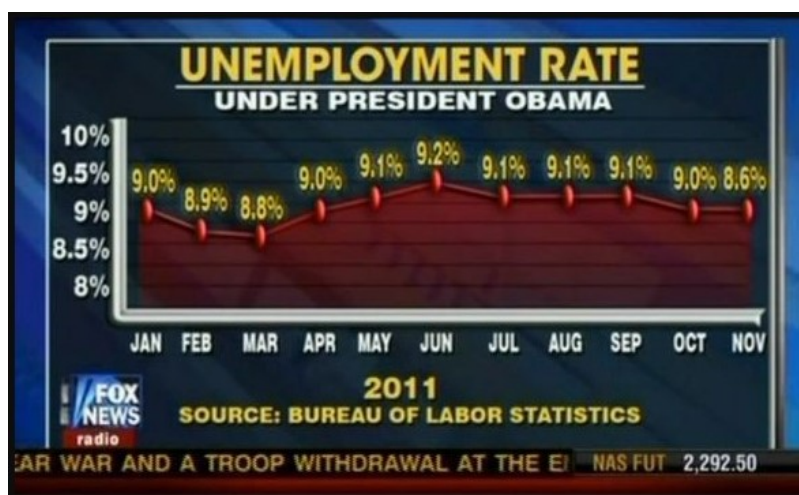
Fonte: Elaborada pelo autor (2022).

Assim, a comparação entre os gráficos torna mais visível a discrepância entre a informação divulgada e os dados reais. A partir dessa comparação deve ser motivada a discussão entre os alunos sobre o critério de apresentação do gráfico pela equipe do Candidato X.

6.3 GRÁFICO 3

A Figura 48 exibe o gráfico que representa a evolução da taxa de desemprego durante o ano de 2011 do governo do presidente americano.

Figura 48 - Taxa de desemprego nos EUA em 2011



Fonte: Disponível em: <https://medium.com/@alxcrv/fake-news-em-gr%C3%A1ficos-d68eff4462e4> (Acesso em 03.jul.2022).

Observe que há discrepância na escala apresentada. O último valor apresentado, 8,6% está no mesmo nível do valor anterior 9,0% e bem superior ao valor de março que é de 8,8%.

Solicite aos alunos para lançar os dados na planilha do LibreOffice Calc. Na coluna A devem colocar o nome do mês e na coluna B o percentual correspondente, conforme a Figura 49.

Para inserir o gráfico, deve selecionar os dados, clicar em *Inserir » Gráfico* e efetuar as seguintes orientações:

1. Em **Tipo de gráfico**, selecionar Linha e a opção Pontos e linhas. Depois clicar em Próximo;
2. Em **Intervalo de dados**, selecionar Série de dados em colunas e deixar marcado Primeira coluna como rótulo;
3. Em **Série de dados**, não realizar alterações;
4. Em **Elementos do gráfico**, desmarcar a opção Exibir legenda;
5. Clicar em **Finalizar**.

Figura 49 - Digitação dos dados de desemprego nos EUA

	A	B
1	JAN	9,00%
2	FEB	8,90%
3	MAR	8,80%
4	APR	9,00%
5	MAY	9,10%
6	JUN	9,20%
7	JUL	9,10%
8	AUG	9,10%
9	SEP	9,10%
10	OCT	9,00%
11	NOV	8,60%

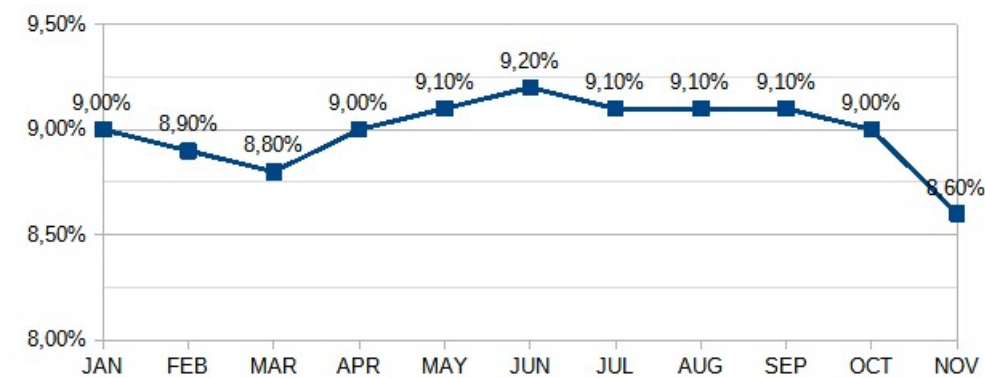
Fonte: Elaborada pelo autor (2022).

Algumas configurações adicionais podem ser realizadas para melhorar a visualização do gráfico.

Para incluir o rótulo dos dados, clicar com o botão direito em cima do gráfico e escolher Inserir rótulos dos dados.

Para mudar a escala do eixo Y, clicar com o botão direito do mouse e escolher a opção *Formatar eixo..* Na aba Escala, desmarcar a opção Automático de Mínimo e colocar o valor de 8%. Em Intervalo Principal, desmarcar a opção Automático e colocar o valor de 0,5%. Será obtido o Gráfico 29:

Gráfico 29 - Taxa de Desemprego com adequação



Fonte: Elaborada pelo autor (2022).

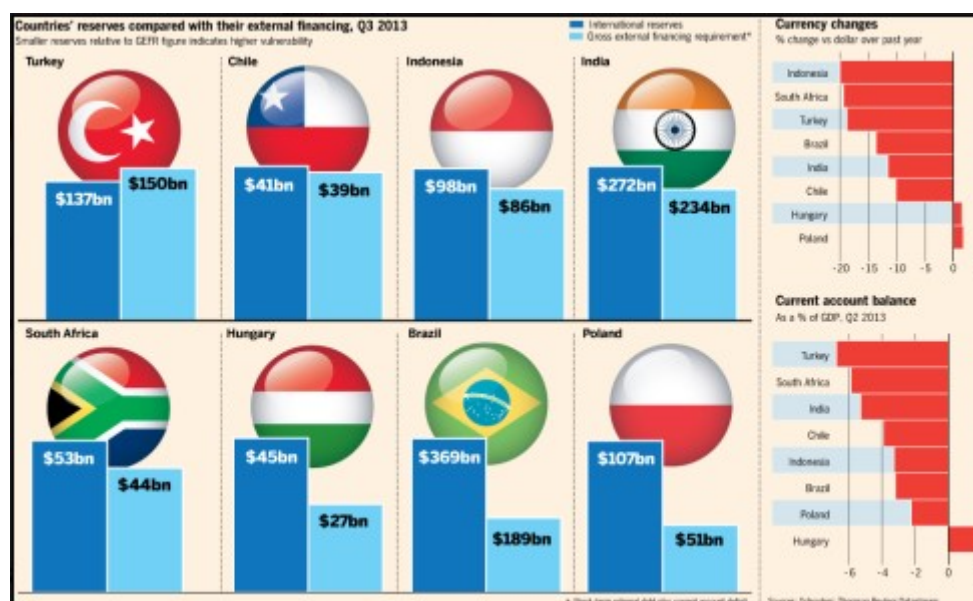
Logo, os alunos devem perceber que o gráfico mostra uma tendência de queda em

relação ao gráfico mostrado anteriormente.

6.4 GRÁFICO 4

A Figura 50 mostra a comparação entre a relação das reservas internacionais e a necessidade de financiamento de diversos países emergentes. Discuta com os alunos as discrepâncias apresentadas com relação as escalas utilizadas.

Figura 50 - Comparativo de Reservas Internacionais em 2013



Fonte: Disponível em: <http://blog.brasilacademico.com/2014/04/graficos-que-mentem-manipulando.html> (Acesso em: 03.jul.2022)

Observe que, visualmente, mesmo o Brasil possuindo US\$ 369 bi, o tamanho da barra representativa do valor está no mesmo tamanho dos demais países, apesar de não serem valores semelhantes.

Para correção do gráfico os alunos devem primeiramente digitar os valores de cada país referentes as Reservas Internacionais e Necessidade de Financiamento Externo na planilha do Libreoffice Calc, nas colunas *B* e *C*, respectivamente, conforme a Figura 51:

Figura 51 - Digitação dos valores das reservas internacionais na planilha

	A	B	C
1	País	Reservas Internacionais	Necessidade de Financiamento Externo
2	Turquia	137	150
3	Chile	41	39
4	Indonésia	98	86
5	Índia	272	234
6	África do Sul	53	44
7	Hungria	45	27
8	Brasil	369	189
9	Polônia	107	51

Fonte: Elaborada pelo autor (2022).

Para a inclusão do gráfico, os alunos devem selecionar o intervalo $A1 : C9$ e clicar em *Inserir » Gráfico*. Na caixa de diálogo devem fazer as seguintes configurações:

1. Em **Tipo de gráfico**: Escolher o gráfico de coluna;
2. Em **Intervalo de dados**: Escolher Série de dados em colunas e deixar as duas caixas de seleção marcadas;
3. Em **Série de Dados**: Não efetuar alterações;
4. Em **Elementos do Gráfico**: Exibir a legenda Embaixo;
5. Clicar em **Finalizar**.

Para melhorar a visualização do gráfico, peça aos alunos para efetuarem as seguintes alterações:

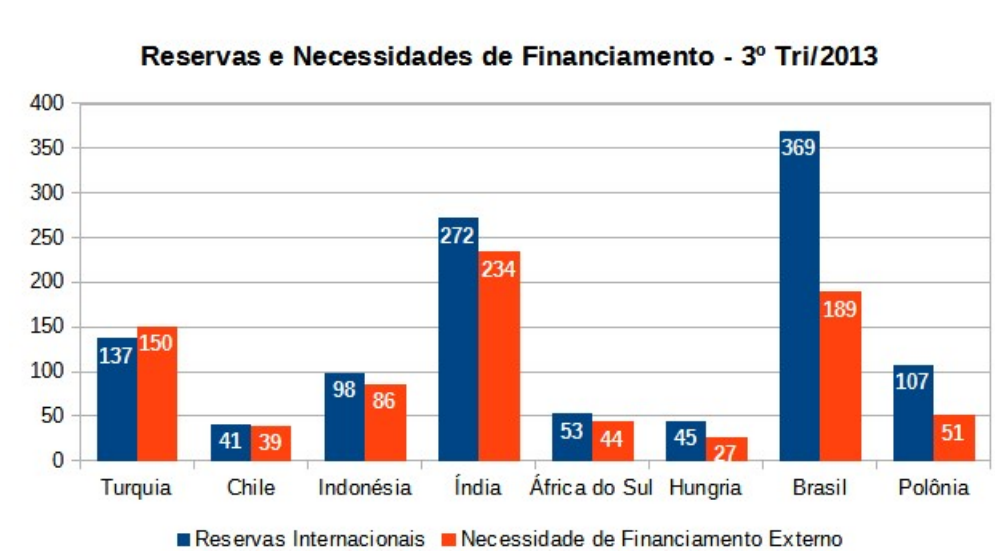
1. Clicar com o botão direito do mouse em cima da barra do gráfico e escolher a opção *Inserir rótulo de dados*;
2. Clicar novamente com o botão direito do mouse em cima da barra e escolher a opção *Formatar rótulos de dados*;
3. Na guia Rótulo de dados, da caixa de diálogo, escolher em *Opções de Atributos* a *Colocação Dentro*. Isto fará que os rótulos fiquem dispostos internamente na barra vertical;
4. Na guia Fonte, escolher o *Estilo Negrito* e na guia *Efeitos da Fonte* a cor *Branca*.

5. Para confirmar as alterações deve-se clicar em *OK*.

O procedimento acima deve ser realizado para cada uma das barras.

Por último, deve-se inserir o título do gráfico. Para isto, clicar em cima do gráfico com o botão direito do mouse e escolher a opção: *Inserir títulos..* e digitar: Reservas e Necessidades de Financiamento - 3º Tri/2013. O resultado aparece no Gráfico 30.

Gráfico 30 - Correção do Gráfico de Reservas Internacionais



Fonte: Elaborada pelo autor (2022).

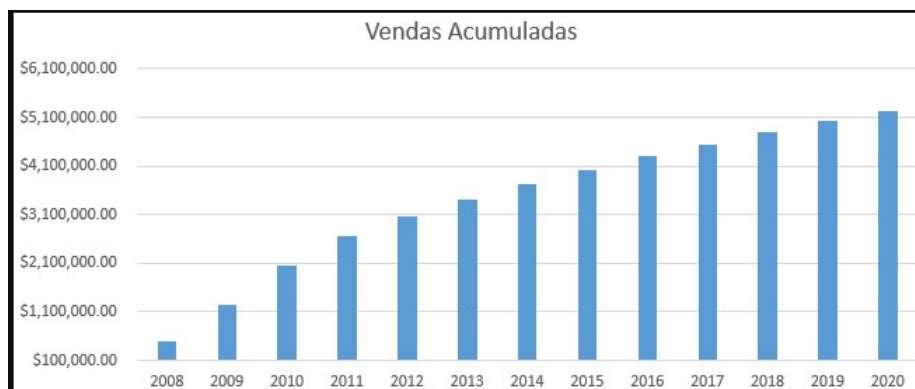
Pode-se trabalhar com os alunos a configuração do gráfico acima. Solicite para que mudem as cores das colunas, escolham a opção 3D, a cor de fundo, entre outras configurações.

E por fim, discuta com os alunos qual a principal modificação apresentada entre o gráfico exibido, e o que foi realizado em sala de aula.

6.5 GRÁFICO 5

Muitas das vezes a utilização de gráficos cumulativos visa a mostrar o bom desempenho dos dados, sejam eles: número de usuários, receita, *downloads* ou outras medidas importantes. Geralmente esses gráficos tendem a mostrar crescimento, o que indicará bom rendimento. Observe o Gráfico 31 com as vendas acumuladas de uma empresa ao longo dos anos:

Gráfico 31 - Vendas acumuladas

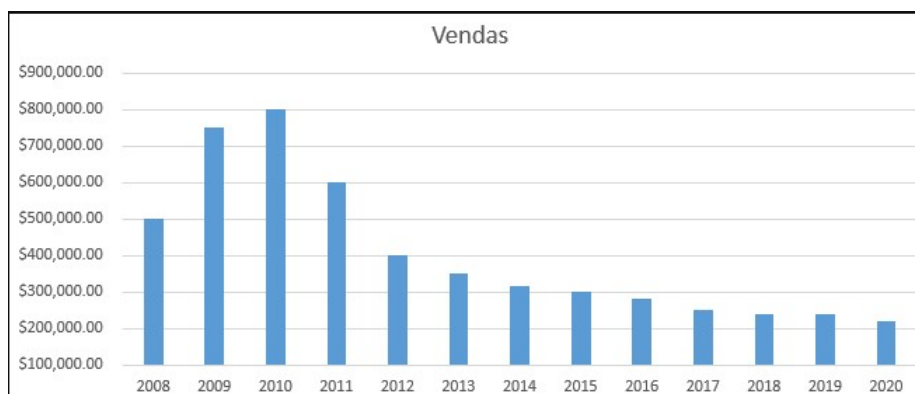


Fonte: Disponível em:

<https://medium.com/big-data-blog/a-estat%C3%ADstica-pode-te-enganar-como-fugir-das-armadilhas-de-dados-5f417823afe8> (Acesso em: 03.jul.2022)

É muito comum interpretar esse gráfico como um bom rendimento da empresa em questão. Todavia, observe o Gráfico 32 que exhibe as vendas não acumuladas ao longo dos anos:

Gráfico 32 - Vendas não acumuladas



Fonte: Disponível em:

<https://medium.com/big-data-blog/a-estat%C3%ADstica-pode-te-enganar-como-fugir-das-armadilhas-de-dados-5f417823afe8> (Acesso em: 03.jul.2022)

A interpretação do segundo gráfico mostra que as vendas da empresa vêm diminuindo ao passar dos anos, o que pode representar uma queda representativa da receita. Assim, o primeiro gráfico pode levar a uma interpretação equivocada da situação atual.

O mesmo pode ser observado na Figura 52, onde o CEO da Apple, Tim Cook, fez a apresentação do gráfico das vendas dos iPhones em 2013:

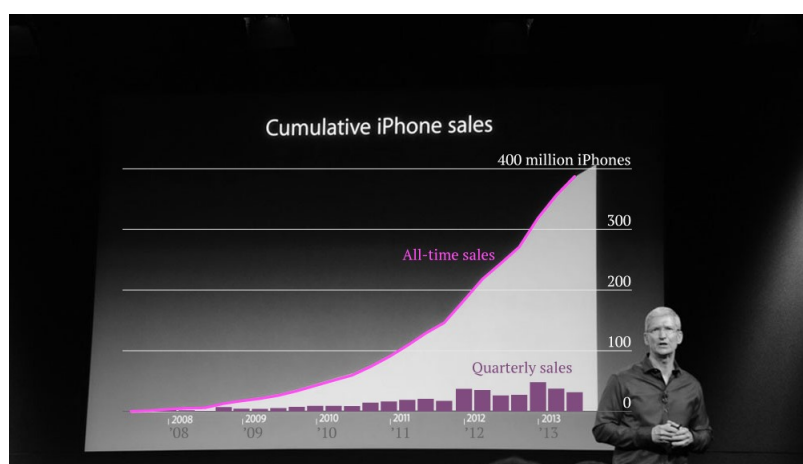
Figura 52 - Vendas acumuladas de iPhone em 2013



Fonte: Disponível em:
<http://blog.brasilacademico.com/2014/04/graficos-que-mentem-manipulando.html> (Acesso em: 03.jul.2022)

Porém, quando é realizada a comparação com a quantidade de vendas por trimestre, obtém-se o gráfico exibido na Figura 53:

Figura 53 - Comparação Vendas acumuladas com a quantidade de vendas por trimestre



Fonte: Disponível em:
<http://blog.brasilacademico.com/2014/04/graficos-que-mentem-manipulando.html> (Acesso em: 03.jul.2022)

Logo a intenção de mostrar vendas acumuladas não representa que o número de vendas está em estabilidade com tendência de queda.

Nesse exemplo, como não se tem os dados listados para a elaboração do gráfico, os alunos devem entender a necessidade da interpretação. Sugere-se que os alunos obtenham dados de outras empresas, ou fictícios, para que possam verificar as comparações.

7 CONCLUSÃO

O objetivo deste trabalho é apresentar uma proposta de ensino e aprendizagem, utilizando ferramentas tecnológicas para a resolução de questões relacionadas à Estatística do ensino básico. A pesquisa foi motivada para que o ensino de estatística fosse desenvolvido dando enfoque nos procedimentos voltados à aprendizagem dos alunos. Uma das vantagens de utilizar essas ferramentas é a possibilidade de empregar diferentes abordagens para a resolução do problema, expandir a análise dos exercícios, proporcionar uma melhoria na interpretação gráfica dos resultados e se familiarizar com o uso da tecnologia.

Para que fosse possível alcançar esse objetivo, inicialmente foi realizado uma pesquisa da Base Nacional Comum Curricular com foco na área de Matemática e Tecnologia voltadas para o ensino de Estatística. Posteriormente foi necessária a introdução dos principais conceitos de estatística descritiva, listando cada uma das medidas de tendência central, medidas de variabilidade e também a construção de tabelas e gráficos. Com o conceito previamente formalizado foi necessária uma análise da utilização de ferramentas computacionais para o ensino de estatística. Uma breve explanação sobre os *softwares* GeoGebra e LibreOffice Calc foi realizada para a sua aplicação. Em seguida para cada um dos conteúdos listados, foram apresentados exemplos de sua aplicação utilizando o GeoGebra ou o LibreOffice Calc, não se limitando somente a resolução do exercício, mas expandindo a exploração da atividade com complementação das tarefas e ampliação da aprendizagem utilizando as possibilidades dos *softwares*.

A utilização das ferramentas tecnológicas não deve ser vista como uma forma única de se compreender a estatística, mas como uma maneira de consolidar a aprendizagem desses conteúdos, motivo pelo qual a revisão bibliográfica é importante para o desenvolvimento dos conceitos estatísticos.

Para que se chegue a um bom resultado, a integração dessas ferramentas em sala de aula necessita da preparação dos professores para que consigam desenvolver atividades relacionadas as ferramentas tecnológicas. É necessário que os professores desenvolvam materiais, que possibilitem a adaptação para que tenha eficácia na aplicação das atividades. É importante uma preparação, para que as atividades sejam executadas de acordo com o nível da turma, favoreçam a integração com o cotidiano do aluno, valorize a interação entre aluno e professor, e que o tempo proposto para a sua realização seja adequado. Deve-se observar se as atividades propostas atendem as competências e habilidades relacionadas ao conteúdo ministrado para cada turma.

A utilização do GeoGebra e LibreOffice Calc deve ser verificada não somente como uma ferramenta para solução dos problemas apresentados, mas como uma importante ferramenta de ensino-aprendizagem que podem contribuir para a construção do conhecimento com professores e alunos. A utilização dessas ferramentas tecnológicas torna a aula

mais dinâmica e facilita a visualização gráfica do comportamento dos dados observados. Além disso, apesar das características que os *softwares* apresentam para a abordagem dos conceitos estatísticos, deve ser considerado que é importante a maneira como se esboça, realiza e avalia as atividades propostas, de acordo com os objetivos traçados inicialmente e a utilização dos ambientes computacionais disponíveis, evitando somente a aplicação de fórmulas para o ensino do conteúdo.

Durante a execução deste trabalho, procurou-se mostrar os vários recursos que os *softwares* apresentam, sejam eles estáticos (como gráficos e tabelas) ou dinâmicos (por meio de exploração e manipulação), permitindo a interação dos alunos com o ambiente. É claro que são inúmeros as possibilidades de utilização das ferramentas, cabendo ao professor escolher a forma mais apropriada de adequá-las aos objetivos da aula, fazendo com que os alunos possam chegar às suas próprias conclusões.

A escolha dessas duas ferramentas tecnológicas foi motivada por serem *softwares* livres, pois viabilizam a utilização nas escolas públicas, sem custos adicionais. Porém, existem variados tipos de *softwares* capazes de promover o desenvolvimento básico dos conceitos estatísticos. O uso de ferramentas tecnológicas pode contribuir expressivamente para o ensino de estatística em sala de aula e deve ser estimulado para que os alunos se tornem capacitados para as próximas etapas de ensino e se tornarem bem-sucedidos profissionalmente. Os professores podem ser motivados a utilizar a inovação tecnológica no processo de ensino, acrescentando a tecnologia às aulas expositivas.

A facilidade com que essas ferramentas proporcionam a elaboração de tabelas e gráficos, possibilita aos alunos a capacidade de interpretar estatisticamente o significado dos dados. Nesse momento o professor deve instigar o aluno a avaliar a compatibilidade da construção que foi realizada com os resultados que foram alcançados anteriormente. Assim, o professor será capaz de analisar a aprendizagem dos alunos sobre a interpretação do pensamento estatístico.

Em relação à continuidade desse trabalho vê-se a importância de avançar a exploração das possibilidades das ferramentas no ensino de estatística. A criação de *applets* e outras atividades que possam ser disponibilizadas online, podem ampliar a diversidade de conteúdos dinâmicos para o ensino de estatística.

Por fim, espera-se que este trabalho contribua para o desenvolvimento do processo de ensino e aprendizagem no que concerne à estatística no ensino básico. Espera-se que os exemplos aqui propostos possam ser implementados pelos professores, para que consigam produzir materiais complementares que sejam capazes de motivar o desenvolvimento dos alunos, desencadeando a assimilação do conhecimento e que desperte o interesse dos alunos pela estatística e seu papel no cotidiano.

REFERÊNCIAS

- ANDERSON, David R. *et al.* **Estatística aplicada a administração e economia.** São Paulo: Cengage, 2019.
- BALESTRI, Rodrigo. **Matemática: interação e tecnologia, volume 3. 2ª ed.** São Paulo: Leya, 2016.
- BIANCHINI, Edwaldo. **Matemática 9º ano - Bianchini. 9ª ed.** São Paulo: Moderna, 2018
- BRASIL. Ministério da Educação. **Base Nacional Comum Curricular.** Brasília, 2018. Disponível em: http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_-versaofinal_site.pdf. Acesso em: 21 ago. 2021
- CHANCE, B. *et al.* The Role of Technology in Improving Student Learning of Statistics. **Technology Innovations in Statistics Education Journal**, v. 1, n. n. 1, p. 1-26, 2007.
- CHAVANTE, Eduardo. PRESTES, Diego. **Quadrante Matemática, 3º ano: ensino médio. 1ª ed.** São Paulo: Edições SM, 2016.
- CRESPO, Antônio Arnot. **Estatística Fácil. 17ª ed.** São Paulo: Saraiva, 2002.
- DANTE, Luiz Roberto. **Matemática: Contexto & Aplicações: ensino médio. 3ª ed.** São Paulo: Ática, 2016.
- FIOCRUZ: O que é Estatística. **Fiocruz**, s.d. Disponível em: <http://www.juventudect.fiocruz.br/estatistica>. Acesso em: 12 de mar. de 2023.
- GEOGEBRA, 2022. Página inicial. Disponível em: <https://www.geogebra.org/about>. Acesso em: 03 de jul. de 2022.
- GRÁFICOS que mentem: Manipulando gráficos e dados. **Brasil Acadêmico**, 2014. Disponível em: <http://blog.brasilacademico.com/2014/04/graficos-que-mentem-manipulando.html>. Acesso em: 03 de jul. de 2022.
- KENSKI, Vani Moreira. **Tecnologias e ensino presencial e a distância. 4ª ed.** Campinas, SP: Papirus, 2003.
- LARSON, Ron. FARBER, Betsy. **Estatística Aplicada. 6ª ed.** São Paulo: Pearson Education do Brasil, 2015.
- LEONARDO, Fábio Martins de. **Conexões com a matemática.** São Paulo: Moderna, 2016.
- LIBREOFFICE: O que é, como instalar e usar no seu PC. **Techtudo**, 2015. Disponível em: <https://www.techtudo.com.br/noticias/2015/12/libreoffice-o-que-e-como-instalar-e-usar-no-seu-pc.ghtml>. Acesso em: 03 de jul. de 2022.
- LOPES, Alex. Fake News em Gráficos. **Medium.com**, 2018. Disponível em: <https://medium.com/@alxcrv/fake-news-em-gr%C3%A1ficos-d68eff4462e4>. Acesso em: 03 de jul. de 2022.

MARTINS, Gilberto de Andrade. **Estatística Geral e Aplicada. 3ª ed.** São Paulo: Atlas, 2006.

MORETTIN, Pedro A. BUSSAB, Wilton O. **Estatística Básica. 6ª ed.** São Paulo: Saraiva, 2010.

MURAN, José Manuel. **A educação que desejamos: novos desafios e como chegar lá. 2ª ed.** Campinas, SP: Papirus, 2007.

PARIKH, Ravi. Como são feitos os gráficos enganosos – e como não ser enganado por eles. **Gizmodo**, 2014. Disponível em: <https://gizmodo.uol.com.br/mentir-visualizacao-dados/>. Acesso em: 03 de jul. de 2022.

PINHEIRO, João Ismael D. *et al.* **Estatística Básica: a arte de trabalhar com dados.** Rio de Janeiro: Elsevier, 2009.

SILVEIRA, Ênio. **Matemática: compreensão e prática. 3ª ed.** São Paulo: Moderna, 2015.

SOUZA, R.; CALEJON, L. Uso da tecnologia da informação e comunicação em uma sequência didática incluindo software GeoGebra no Ensino da Estatística Descritiva. **Revista de Ensino de Ciências e Matemática**, v. 10, n. 4, p. 227-244, 18 jul. 2019.

TEIXEIRA, Lucas Borges. Após críticas, PSDB-SP apaga gráfico que supervalorizava Doria em pesquisa. **UOL Notícias**, São Paulo, 30 de mai. de 2018. Disponível em: <https://noticias.uol.com.br/politica/eleicoes/2018/noticias/2018/05/30/psdb-sp-divulga-grafico-desproporcional-de-doria-e-tira-do-ar-apos-criticas.htm>. Acesso em: 03 de jul. de 2022.

TOLEDO JR., Tiago. CUNHA, Luna. A Estatística Pode Te Enganar: Como Fugir das Armadilhas de Dados. **Medium.com**, 2020. Disponível em: <https://medium.com/big-data-blog/a-estat%C3%ADstica-pode-te-enganar-como-fugir-das-armadilhas-de-dados-5f417823afe8>. Acesso em: 03 de jul. de 2022.

VIEIRA, Sônia. **Estatística Básica. 2ª ed.** São Paulo: Cengage, 2018.

APÊNDICE A – Caderno de Atividades

**Caderno
de
Atividades**

Exercícios de Estatística Utilizando Ferramentas Tecnológicas para
Alunos do Ensino Fundamental e Médio.

Orientações Gerais

O presente Apêndice contém um Caderno de Atividades, voltado aos professores como sugestão de aplicação de exercícios de estatística, utilizando ferramentas tecnológicas para alunos do Ensino Fundamental e Médio. As atividades abrangem grande parte do conteúdo de Estatística proposto na BNCC, e têm o objetivo de trabalhar a familiaridade dos alunos com o uso de tecnologias, tornando a aprendizagem mais atraente, para a obtenção de resultados e análises críticas dos dados observados, adequando-se ao cotidiano dos alunos.

Será necessário, para a resolução dessas atividades, que os alunos tenham acesso a computadores. Assim, sugere-se que já estejam instalados nos respectivos computadores os *softwares* GeoGebra e LibreOffice Calc. Não é necessário estarem conectados à internet. Se o número de computadores não for suficiente para todos os alunos, deve-se dividir a turma em grupos, ou estipular revezamento para que todos possam participar do desenvolvimento das atividades.

Aos professores que não tiverem conhecimento da utilização dos *softwares*, é recomendável que assistam tutoriais disponíveis na internet para a familiarização na sua utilização. Pode-se também, utilizar como apoio o desenvolvimento dos exemplos desta dissertação, para a realização passo-a-passo das atividades propostas.

Inicialmente para a realização das atividades 1, 2 e 3, será necessário que os alunos preencham um formulário, cujas respostas serão utilizadas para o desenvolvimento das mesmas. O formulário deverá conter: a Idade e Altura dos alunos; e se praticam atividades extraclasses (Pratica esportes, Curso de Inglês, Aula de Música). Para a coleta dos dados pode ser utilizado um formulário impresso, uma lista a ser preenchida manualmente ou até mesmo um formulário para preenchimento on-line.

Um formulário simples para coleta de dados, pode ter o seguinte formato:

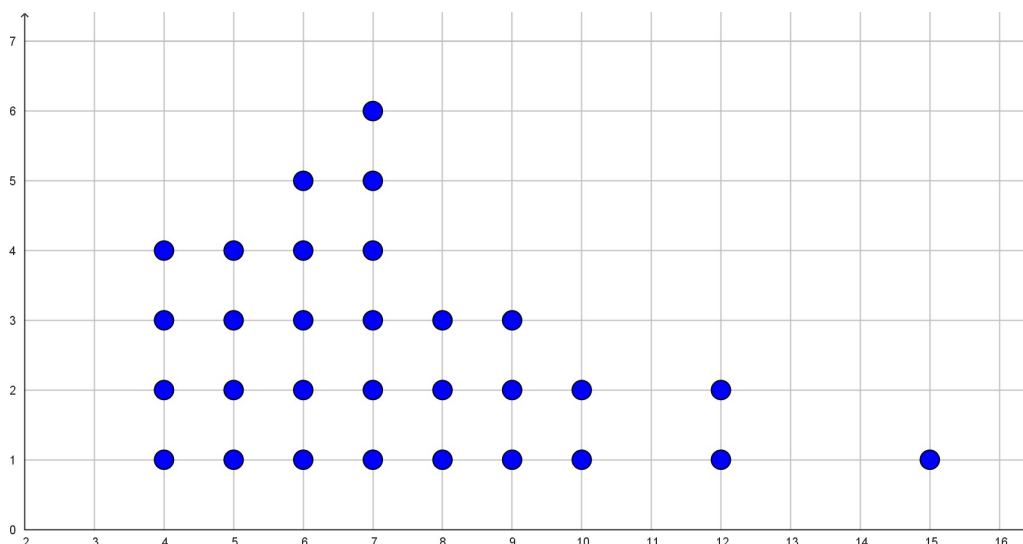
Dados para atividades

1. Idade: _____
2. Altura: _____ (em cm)
3. Pratica Esportes? () Sim () Não
4. Faz curso de inglês? () Sim () Não
5. Faz aula de música? () Sim () Não

Nota: Os professores podem adaptar as variáveis sugeridas para a solução das atividades, sugerindo outras que sejam mais apropriadas ao dia a dia dos alunos em sala de aula.

Atividade 1 - Medidas de Tendência Central

O diagrama de pontos abaixo representa a quantidade de livros lidos pelos alunos de uma turma durante o ano escolar.



Responda aos seguintes questionamentos:

1. Classifique o tipo de variável apresentada no diagrama.
2. Quantos alunos essa turma possui?
3. Qual é a menor quantidade de livros lidos? E a maior quantidade de livros lidos?
4. Identifique a moda dos dados. Como chegou a este resultado?
5. Qual é a mediana dos dados apresentados? Identifique-a no gráfico.
6. A partir do diagrama de pontos, estime a média aritmética desse conjunto de dados. Justifique sua resposta.
7. Calcule a média aritmética do conjunto de dados e compare com o valor estimado.
8. Exclua dos cálculos a quantidade de livros lidos igual a 15 e calcule novamente a mediana e a média. Houve alteração dos valores?
9. Em ambos os casos, qual é a melhor medida representativa desse conjunto de dados?

Nesta primeira etapa, os alunos devem analisar o diagrama de pontos e identificar as medidas de tendência central. Em seguida, será utilizada uma ferramenta tecnológica para aprofundar esses questionamentos.

► **Utilizando o *software* GeoGebra.**

Esta atividade tem por objetivo realizar a coleta das idades dos alunos e, com o auxílio do GeoGebra, calcular a média, a moda e a mediana desses dados, além de analisar graficamente o comportamento dessas medidas.

Será necessário, que os alunos tenham preenchido o formulário “Dados para atividades” solicitado anteriormente. Será utilizado os dados referente a IDADES.

Antes de montar uma tabela com as idades coletadas do formulário, é importante que os alunos classifiquem a variável e sua unidade de medida. Neste caso, a variável IDADES é numérica discreta e medida em anos.

O professor deve distribuir a tabela com os dados coletados das idades para os alunos (ou grupos), e orientá-los para que os acrescentem esses dados na Planilha do GeoGebra. Explique como realizar esse procedimento.

Para realizar o cálculo das medidas de tendência central (Média, Moda e Mediana), utilize o auxílio do *software*. Inicialmente os alunos devem criar uma lista na planilha do GeoGebra com as idades informadas. Assim, por meio da lista criada, utilize os comandos do GeoGebra para encontrar os resultados das medidas. As sintaxes para o cálculo das medidas são:

- Média(<Lista de Dados Brutos>)
- Mediana(<Lista de Dados Brutos>)
- Moda(<Lista de Números>)

Após o cálculo das medidas de tendência central, os alunos irão analisar graficamente a mediana. Acrescente um Diagrama de Pontos com as idades que estão na lista. A sintaxe para sua construção é: DiagramaDePontos(<Lista de Dados Brutos>).

Com o diagrama de pontos exibido na Janela de Visualização, identifique a moda, ou seja, a coluna que apresenta mais pontos.

Para analisar a mediana, utilize o Comando ParteDaLista, identificando a metade dos valores das idades. Reforce com os alunos o conceito de mediana, uma vez que é necessário que os dados estejam ordenados. É possível alterar as configurações de cores dessa lista para analisar graficamente a mediana.

Para verificar a média, insira uma reta que identifique o valor médio das idades. O valor encontrado, possui existência concreta?

E por fim, solicite aos alunos que comparem os dados calculados pelo *software* com a análise gráfica. Seria possível chegar as mesmas conclusões, sem realizar o cálculo algébrico?

Atividade 2 - Tabela de Frequências com Dados Agrupados

A tabela a seguir apresenta o peso dos 30 alunos de uma turma.

Peso (kg)	Frequência Absoluta
54 – 58	4
58 – 62	5
62 – 66	8
66 – 70	6
70 – 74	4
74 – 78	2
78 – 82	1

Responda aos seguintes questionamentos:

1. Classifique o tipo de variável (Peso) apresentada na tabela.
2. Quantas classes tem essa tabela de dados? Qual é a altura de cada classe?
3. Identifique a classe modal, ou seja, aquela que possui mais elementos.
4. É possível identificar a moda desses dados? Como pode-se calcular esse valor?
5. Construa um histograma com base nessa tabela. O eixo x representará a classe dos pesos dos alunos e o eixo y a quantidade de elementos de cada classe.
6. Analisando o histograma, é possível identificar a moda e a mediana desses dados? Justifique.

Nesta primeira etapa, os alunos devem analisar a tabela de frequência com dados agrupados e construir um histograma com seus valores. Em seguida, deverão utilizar uma ferramenta tecnológica para aprofundar esses questionamentos.

► Utilizando o *software* GeoGebra.

Nesta atividade os alunos aprenderão como montar uma tabela de frequência com dados agrupados a partir de uma coleta de informações simples. Será utilizado os dados referente à ALTURA do formulário “Dados para atividades”.

Identifique a variável ALTURA e sua unidade de medida. No caso é uma variável numérica contínua, medida em centímetros (cm).

Entregue a tabela com as alturas dos alunos coletadas no formulário e solicite a

digitação dos valores das alturas na planilha do GeoGebra. Oriente para que utilizem a mesma unidade de medida.

Inicialmente será inserido uma tabela de frequência com dados simples. Para isso, crie uma lista com os dados inseridos, e utilize a sintaxe `TabelaDeFrequências(<Lista de Dados Brutos>)`. O resultado será exibido na Janela de Visualização, e algumas observações sobre medidas estatísticas já podem ser realizadas, como a moda e a mediana.

Para a inserção da Tabela de Frequências com dados agrupados, será necessário utilizar a Análise Univariada e incluir um Histograma. Para a configuração do Histograma, verifique qual o extremo inferior e defina esse valor em Início e para a altura das classes, defina a Largura como 5cm. Clique em Exibir Tabela de Frequências, e observe o histograma.

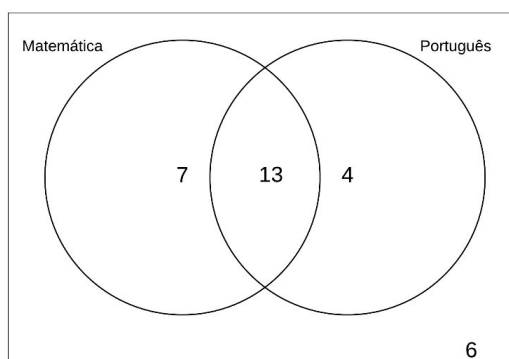
Quantas classes foram criadas? E se a altura de cada classe fosse 10cm, quantas classes seriam?

Adicione o Histograma na Janela de Visualização, e observe a Tabela de Frequência de Dados Acumulados. Analisando o Histograma, identifique a Moda desses dados.

Utilizando os comandos na Caixa de Entrada do GeoGebra, calcule a Média e a Mediana da altura dos alunos. Observe se a análise corresponde com os valores calculados.

Atividade 3 - Diagrama de Venn e Gráfico de Setores

O diagrama de Venn abaixo, representa as disciplinas preferidas, entre Matemática e Português, em uma turma de 30 alunos.



Responda aos seguintes questionamentos:

1. Quantos alunos gostam de matemática?
2. Quantos alunos gostam de português?
3. Quantos alunos não gostam de nenhuma das duas disciplinas?
4. Qual o percentual de alunos que gostam somente de matemática?
5. Qual o percentual de alunos que gostam somente de português?
6. Qual o percentual de alunos que gostam das duas disciplinas?
7. Construa uma tabela com os dados obtidos do Diagrama de Venn, com suas frequências absolutas e relativas.

Nesta primeira etapa, os alunos irão analisar um Diagrama de Venn e calcular a frequência relativa de cada variável. Além disso, devem construir uma tabela com os dados observados. Em seguida, utilizando uma ferramenta tecnológica, irão aprofundar esses questionamentos.

► Utilizando o *software* LibreOffice Calc.

Esta atividade tem o objetivo de criar um gráfico de setores, utilizando inicialmente informações contidas em um Diagrama de Venn. Será utilizado os dados referente às atividades praticadas constantes no formulário “Dados para atividades”.

Com as respostas obtidas no formulário e, com o auxílio do LibreOffice Calc, monte

uma tabela com as informações sobre as atividades extraclasses dos alunos. Para isso será necessário inserir três colunas (Esportes, Inglês, Música), e em cada linha identificar a resposta dos alunos individualmente. Sugere-se escrever “Sim” e “Não” para identificar se a atividade é realizada ou não.

Peça para os alunos criarem uma tabela auxiliar, em que seja possível identificar a quantidade de alunos, que realizam as seguintes atividades:

- Somente Pratica esportes
- Somente faz Curso de Inglês
- Somente faz Aula de música
- Pratica esportes e faz Curso de Inglês
- Pratica esportes e faz Aula de Música
- Faz Curso de Inglês e Aula de Música
- Pratica Esportes, faz Curso de Inglês e Aula de Música
- Não realiza nenhuma atividade

Os alunos podem obter essas informações utilizando o comando =CONT.SE(Intervalo; Critérios), para contar as atividades individuais. Em Intervalo deve-se colocar as células, que compreendem a variável analisada e em Critérios se o aluno realiza a atividade.

Para contar as atividades com múltiplos critérios, por exemplo, se Pratica Esportes e Faz Curso de Inglês utiliza-se o comando =CONT.SES(Intervalo 1;Critérios 1; Intervalo 2;Critérios 2;...), em que cada intervalo e critério, identifica a variável requisitada.

Construa um Diagrama de Venn com os dados observados, e crie uma tabela auxiliar para identificar cada atividade desenvolvida.

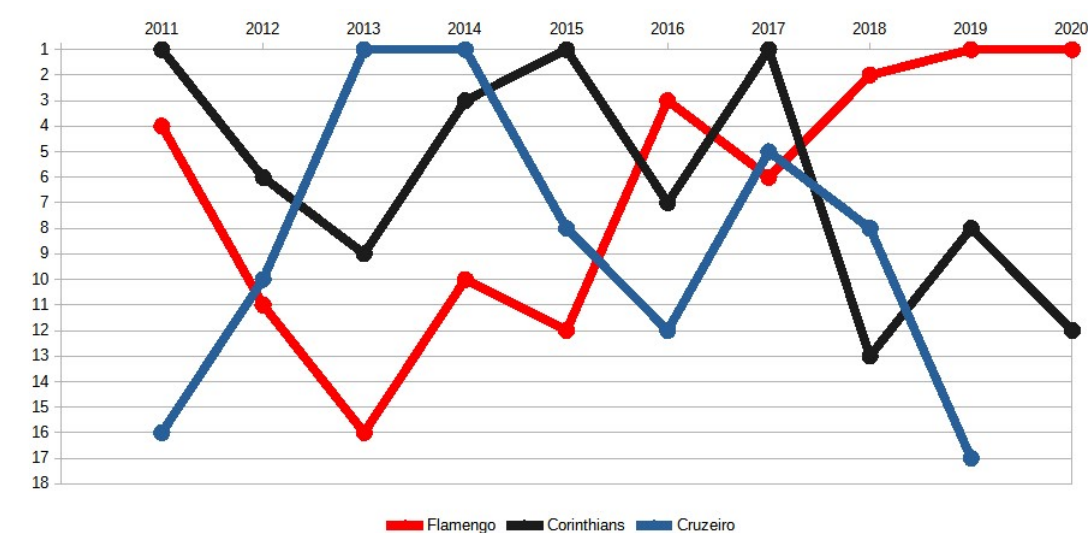
Calcule, no LibreOffice Calc, o percentual de cada uma das atividades listadas, considerando a tabela auxiliar.

Insira um Gráfico de Setores com os dados da Tabela auxiliar, utilizando os valores encontrados no Diagrama de Venn. Realize as configurações necessárias para visualização.

Para finalizar, no Gráfico de Setores, insira os rótulos de dados e exiba em formato percentual.

Atividade 4 - Interdisciplinaridade

O gráfico abaixo, representa o desempenho no Campeonato Brasileiro da Série A, de 2011 a 2020, de três clubes: Flamengo, Corinthians e Cruzeiro. O Cruzeiro não participou da série A em 2020.



Responda aos seguintes questionamentos:

1. Qual foi o ano da pior posição de cada um dos times?
2. Quais anos cada time foi o campeão?
3. Identifique em quais anos o Flamengo foi melhor que o Corinthians. Faça a mesma comparação com o Cruzeiro.
4. Identifique em quais anos o Corinthians foi melhor que o Flamengo. Faça a mesma comparação com o Cruzeiro.
5. Identifique em quais anos o Cruzeiro foi melhor que o Corinthians. Faça a mesma comparação com o Flamengo.
6. Entre 2011 e 2014, qual time obteve o melhor desempenho? Explique comparando com os demais times.
7. Entre 2017 e 2020, qual time obteve o melhor desempenho? Explique comparando os demais times.
8. A partir de 2013, qual time teve a melhor variação de sua posição? Como chegou a essa conclusão?

Na etapa inicial, será requisitado aos alunos que procedam à análise de um gráfico de linhas, com a finalidade de observar o desempenho dos times ao longo de um determinado período e efetuar uma comparação do desempenho de cada equipe. Posteriormente, será

empregada uma ferramenta tecnológica para aprofundar tais indagações.

► Utilizando o *software* LibreOffice Calc.

O objetivo desta atividade é analisar como as medidas estatísticas podem ser utilizadas para servir de apoio à tomada de decisões.

Divida a turma em cinco equipes, sendo que cada uma deverá pesquisar o faturamento anual de determinada empresa (local ou nacional) para servir de análise do problema. Oriente para pesquisar em meios de comunicação (internet, jornais, etc.).

Após a coleta desses dados, reúna as informações de todas as equipes, e crie uma tabela identificando a empresa e o faturamento por ano.

Distribua essa tabela aos alunos, e peça para inserir esses valores em uma planilha do LibreOffice Calc. Os valores devem estar na mesma unidade de medida.

Calcule a média de cada empresa nos últimos cinco anos, utilizando a fórmula de média do LibreOffice Calc, e ordene as empresas da maior para a menor média calculada. Identifique as duas empresas com maior média.

Logo após a obtenção da média, impõe-se aos estudantes a avaliação do crescimento de faturamento das empresas que detêm a melhor média nos últimos anos. Com o intuito de realizar tal tarefa, utilizar-se-á um gráfico para fins de análise.

Insira um gráfico de linhas com os valores anuais de cada empresa. Observe que não se deve incluir a média calculada no gráfico. Faça uma análise da evolução de cada empresa. Quais tiveram um crescimento positivo durante esse período? Alguma empresa apresentou estabilidade ou não teve crescimento durante esse período?

Agora para a análise de variabilidade das empresas, insira um gráfico de colunas com os mesmos dados. Configure o gráfico para que os dados de cada empresa sejam exibidos lado a lado. Observando o gráfico de colunas, identifique quais empresas possuem maior variabilidade, ou seja, menor diferença entre a altura das colunas comparadas anualmente.

Para confirmar a variabilidade anual, insira na tabela uma coluna de Desvio Padrão e utilizando o comando =DESVPAD(), calcule o valor para cada empresa.

As empresas que tiveram a menor variabilidade, são as que apresentaram a maior média?

Atividade 5 - Analisando os dados discrepantes de uma amostra

Uma fábrica de móveis apresenta a seguinte produção diária de mesas durante um período de 20 dias:

5 - 7 - 6 - 5 - 8 - 9 - 3 - 5 - 7 - 6 - 6 - 8 - 3 - 9 - 13 - 5 - 6 - 8 - 7 - 6

Responda aos seguintes questionamentos:

1. Organize os dados apresentados em um ROL.
2. Identifique o menor e o maior elemento dessa amostra e calcule a amplitude.
3. Calcule a mediana dessa amostra.
4. Calcule o primeiro e o terceiro quartil dessa amostra.
5. Calcule a média dessa amostra.
6. Construa um Diagrama de Caixa, identificando os quartis e os valores máximos e mínimos.
7. Existe algum valor discrepante nessa amostra? Justifique.
8. Elimine esse valor discrepante e calcule novamente os quartis. O que pode se observar quando se elimina um valor discrepante da amostra?

Na etapa inicial, caberá aos alunos a elaboração de um diagrama de caixa, o qual permitirá a observação da distribuição dos elementos da amostra. Através de análise, deverão identificar a presença de um valor discrepante na amostra. Posteriormente, utilizando uma ferramenta tecnológica, deverão aprofundar esses questionamentos.

► Utilizando o *software* GeoGebra.

O objetivo desta atividade será verificar, por meio do Diagrama de Caixa, dados discrepantes de uma amostra e discutir como esses valores interferem nas medidas de tendência central.

Para realizar essa tarefa, será necessário que os alunos pesquisem o número de vitórias de determinado time no Campeonato Brasileiro nos últimos 10 anos. Pode-se sugerir um único time para toda turma, ou dividir a turma em grupos, e cada grupo escolhe um time diferente.

Após a coleta desses dados, monte uma tabela com os resultados encontrados, e insira na planilha do GeoGebra.

Utilizando os comandos do *software*, calcule a média de vitórias desse time nos últimos dez anos, e identifique quais anos o time ficou acima da média. Oriente aos alunos para inserirem um diagrama de pontos com os dados coletados, e por meio de uma reta paralela ao eixo x, traçar a média. Esse procedimento irá facilitar a verificação dos anos, em que o número de vitórias foi superior à média calculada, ao invés de fazer uma análise visual.

Logo em seguida, selecione a coluna que contém o número de vitórias, e habilite a Análise Univariada do *software*. Clique em Exibir Estatística e, assim, identifique a mediana, o valor mínimo, o valor máximo e calcule a amplitude dos valores.

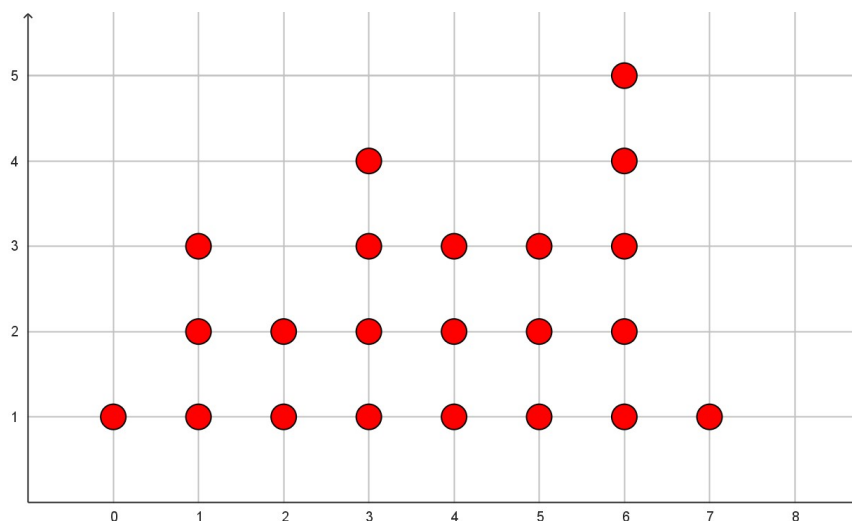
Para analisar o comportamento dos dados, construa um Diagrama de Caixa. Para isto, basta selecionar em Análise de Dados a opção BoxPlot.

Verifique se existem valores discrepantes (*outliers*) dentre os dados coletados. Caso existam, elimine esses valores da amostra. Verifique as modificações que foram executadas na análise estatística, e se os anos analisados anteriormente permanecem os mesmos.

Caso não existam valores discrepantes, acrescente um valor que destoe dos demais, e realize a mesma análise. Qual a conclusão a que se chega?

Atividade 6 - Quartis

O diagrama de pontos abaixo apresenta o número de gols marcados em finais de copas do mundo de 1930 a 2022.



Responda aos seguintes questionamentos:

1. Identifique o menor e o maior números de gols marcados em finais de copas do mundo.
2. Marque no diagrama a mediana dos dados apresentados. Qual é seu valor?
3. Marque no diagrama o primeiro e o terceiro quartil. Quais são esses valores?
4. Qual é a quantidade de gols mais frequente em finais de copas do mundo? Qual é este valor?
5. Usando a fórmula de média ponderada, calcule a média e trace uma reta vertical no diagrama para identificar.

Na fase inicial, os discentes devem efetuar a análise de um diagrama de pontos e identificar as medidas de posição, mais precisamente os quartis. Ademais, devem proceder ao cálculo da média, obtendo-se os valores a partir do referido diagrama. Posteriormente, por meio de uma ferramenta tecnológica, é necessário aprofundar tais indagações.

► Utilizando o *software* GeoGebra.

Nesta tarefa, os alunos terão como objetivo apresentar graficamente os quartis de uma amostra e identificar algumas medidas de tendência central.

Inicialmente os alunos devem pesquisar o número de gols marcados pelo Brasil em

cada edição da Copa do Mundo, e montar uma tabela com esses dados.

Insira o número de gols na planilha do GeoGebra, e crie uma lista com esses valores. Os alunos também podem criar uma lista inserindo os valores diretamente na Caixa de Entrada. É interessante mostrar as duas opções aos alunos.

Por meio da lista criada anteriormente, insira um Diagrama de Pontos. Na Janela de Visualização, identifique qual é o número de gols mais frequente do Brasil em Copas do Mundo, ou seja, o valor modal.

Para analisar graficamente os quartis, crie listas auxiliares utilizando o comando `ParteDaLista` para representar cada quartil desse conjunto de dados, ou seja, será necessário criar diagramas de pontos, um para cada quartil. Modifique a configuração de cores, para visualizar os quartis.

Acrescente um Controle Deslizante com valores de 0 a 4, e configure para que cada diagrama de ponto seja exibido quando modificar o Controle deslizante.

Analisando o diagrama de pontos, verifique qual é a mediana dos números de gols feitos pelo Brasil em Copas do Mundo.

Digite na Caixa de Entrada os comandos necessários para o cálculo da média, e insira uma reta que identifique a média. Verifique quantas vezes o Brasil fez mais gols do que a média geral.