

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

ISABELLA STERZA DE OLIVEIRA BUTZEN

**O USO DO R E RSTUDIO PARA GRÁFICOS ESTATÍSTICOS NA EDUCAÇÃO
BÁSICA**

TOLEDO

2026

ISABELLA STERZA DE OLIVEIRA BUTZEN

**O USO DO R E RSTUDIO PARA GRÁFICOS ESTATÍSTICOS NA EDUCAÇÃO
BÁSICA**

The use of R and RStudio for statistical graphs in Basic Education

Dissertação apresentada como requisito para
obtenção do título de Mestre em do Mestrado
Profissional em Matemática em Rede Nacional
- PROFMAT da Universidade Tecnológica
Federal do Paraná.

Orientador(a): Prof^a. Dr^a. Araceli Ciotti de
Marins

Coorientador(a): Prof^a. Dr^a. Daniela Trentin
Nava

TOLEDO

2026



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



**Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Toledo**



ISABELLA STERZA DE OLIVEIRA BUTZEN

O USO DO R E RSTUDIO PARA GRÁFICOS ESTATÍSTICOS NA EDUCAÇÃO BÁSICA

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Matemática da Universidade Tecnológica Federal do Paraná (UTFPR).
Área de concentração: Matemática Na Educação Básica.

Data de aprovação: 05 de Fevereiro de 2026

Dra. Araceli Ciotti De Marins, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Pablo Chang, Doutorado - Universidade Estadual do Oeste do Paraná (Unioeste)

Dra. Rosangela Aparecida Botinha Assumpcao, Doutorado - Universidade Tecnológica Federal do Paraná

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 04/03/2026.

Dedico este trabalho ao meu marido, meu
porto-seguro.

AGRADECIMENTOS

Este trabalho não poderia ter sido concluído sem o apoio de diversas pessoas, às quais registro meus sinceros agradecimentos. Reconheço que estas linhas não contemplam todas as pessoas que fizeram parte desta etapa tão significativa da minha vida, por isso, desde já peço desculpas àquelas que não estão nomeadas, mas que podem ter certeza de que permanecem presentes em meu pensamento e em minha gratidão.

Agradeço, em primeiro lugar, ao meu marido, que esteve ao meu lado em todos os momentos, inclusive nos mais difíceis. Seu apoio constante, sua paciência e o cuidado cotidiano foram fundamentais para que eu não desistisse e para que eu pudesse sustentar, com tranquilidade, o caminho até a conclusão deste trabalho. Sem esse suporte, este percurso teria sido muito mais árduo e, possivelmente, nem teria sido possível.

Agradeço à minha família, pelo incentivo, pela compreensão e pelo amparo ao longo desta jornada, mesmo quando nem sempre foi possível compartilhar em detalhes as exigências e as inquietações que acompanharam o mestrado.

Agradeço também aos meus amigos, que contribuíram de tantas maneiras para que eu atravessasse este processo com mais leveza, ajudando-me a aliviar a tensão e a lidar com a ansiedade nos momentos de maior pressão. Em especial, agradeço às minhas amigas Alexandra, Aline e Silva, que são referências para mim. Elas me inspiram não só pelo que defendem, mas pelo jeito como vivem. São amigas que acreditam em uma educação pública de verdade e, ao mesmo tempo, cuidam de quem está do lado. Mesmo à distância, estiveram comigo de forma muito concreta, e esse afeto foi essencial para eu chegar até aqui.

Agradeço a todos os professores e colegas do curso, pelas aprendizagens, trocas e contribuições que, direta ou indiretamente, ajudaram a construir este trabalho. De modo especial, agradeço à minha orientadora, Araceli, por conduzir a orientação com sensibilidade e respeito ao meu jeito e ao meu ritmo, tornando possível que este trabalho fosse desenvolvido com seriedade, mas também com o coração leve. Agradeço, ainda, à coorientadora Daniela, pelas contribuições, sugestões e auxílios ao longo do processo.

RESUMO

Este trabalho mapeia e analisa a produção acadêmica sobre o uso da linguagem R no ensino de Estatística na Educação Básica, com foco na formação de professores de Matemática, e propõe um recurso educacional que apoia o uso autônomo do R e do ambiente RStudio na construção e na análise de gráficos estatísticos. Ao situar a pesquisa em um contexto social orientado por dados, marcado pela revolução digital e pela expansão de sistemas baseados em inteligência artificial, o estudo compreende que o raciocínio, o letramento e o pensamento estatístico exigem mais do que cálculo, pois envolvem formular perguntas, organizar e representar informações, interpretar resultados, sustentar argumentos e tomar decisões com base em evidências. O trabalho aplica a metodologia de Estado do Conhecimento, articulada à análise de conteúdo, para levantar, sistematizar e interpretar estudos localizados em bases como o Google Scholar e o Catálogo de Teses e Dissertações da CAPES, a partir de descritores relacionados a Estatística, Matemática, currículo, formação docente e documentos curriculares, incluindo buscas específicas por RStudio e linguagem R. A análise evidencia um descompasso entre a centralidade atribuída à Estatística na Base Nacional Comum Curricular e a forma como a área se materializa na formação inicial e em materiais didáticos, com recorrente predominância de procedimentos e respostas rápidas em detrimento da interpretação crítica e contextualizada. O mapeamento também mostra que propostas formativas com softwares tendem a orientar a replicação de passos, com menor ênfase na explicitação de critérios de escolha de gráficos, na discussão de parâmetros de construção e no desenvolvimento de autonomia docente para adaptar e criar situações investigativas. Como contribuição, o trabalho apresenta um manual digital, organizado de modo progressivo, que orienta instalação e primeiros passos, importação de dados, construção de gráficos de barras, histogramas, setores, linhas e dispersão com funções nativas e com o pacote ggplot2, além de recomendações de replicabilidade e adaptação de scripts a diferentes contextos escolares. Conclui-se que a apropriação autônoma do R e do RStudio fortalece a transparência dos procedimentos, a rastreabilidade das escolhas e a intencionalidade didática, favorecendo um ensino de Estatística mais investigativo, crítico e alinhado às demandas contemporâneas de leitura e interpretação de dados.

Palavras-chave: educação estatística; linguagem r; formação de professores.

ABSTRACT

This study maps and analyzes the academic literature on the use of the R language in the teaching of Statistics in Basic Education, with an emphasis on the education of Mathematics teachers, and proposes an educational resource that supports the autonomous use of R and the RStudio environment for the construction and analysis of statistical graphs. By situating the research in a data-oriented social context shaped by the digital revolution and the expansion of systems based on artificial intelligence, the study argues that statistical reasoning, literacy, and thinking require more than computation, as they involve formulating questions, organizing and representing information, interpreting results, sustaining arguments, and making evidence-based decisions. The study adopts the State of Knowledge methodology, combined with content analysis, to identify, systematize, and interpret studies found in databases such as Google Scholar and the CAPES Catalog of Theses and Dissertations, using descriptors related to Statistics, Mathematics, curriculum, teacher education, and curricular documents, including specific searches for RStudio and the R language. The analysis indicates a mismatch between the centrality assigned to Statistics in Brazil's National Common Curricular Base and the way the field is realized in initial teacher education and in instructional materials, with a recurring emphasis on procedures and quick answers at the expense of critical and contextualized interpretation. The mapping also shows that teacher-education proposals involving software often privilege step-by-step replication, giving less attention to explicit criteria for choosing graphs, to the discussion of construction parameters, and to the development of teacher autonomy to adapt and design investigative learning situations. As a contribution, the study presents a digital manual, progressively organized, that guides installation and initial steps, data import, and the construction of bar charts, histograms, pie charts, line charts, and scatter plots using both base R functions and the ggplot2 package, as well as recommendations on reproducibility and on adapting scripts to different school contexts. It concludes that the autonomous appropriation of R and RStudio strengthens procedural transparency, traceability of choices, and didactic intentionality, fostering a more investigative and critical teaching of Statistics aligned with contemporary demands for reading and interpreting data.

Keywords: statistical education; r programming language; teacher education.

LISTA DE FIGURAS

Figura 1 – Gráfico de barras empilhadas da frequência de presenças por sexo	39
Figura 2 – Gráfico de dispersão com linha de regressão e banda de confiança	51
Figura 3 – Ambiente de desenvolvimento RStudio	55

LISTA DE QUADROS

Quadro 1 – Boas práticas na construção de gráficos de barras	40
Quadro 2 – Boas práticas na construção de histogramas	43
Quadro 3 – Boas práticas na construção de gráficos de setores.....	45
Quadro 4 – Boas práticas na construção de gráficos de linhas	48
Quadro 5 – Boas práticas na construção de gráficos de dispersão	52

SUMÁRIO

1	INTRODUÇÃO	10
2	METODOLOGIA	14
3	ESTADO DA ARTE.....	16
3.1	Revolução Digital, Dataficação e IA na Educação.....	16
3.2	Educação Estatística: entre o raciocínio, literacia e pensamento estatístico	18
3.3	Formação docente e o professor reflexivo	20
3.4	Licenciatura em Matemática e o lugar da estatística no Brasil	22
4	ESTATÍSTICA NA BNCC E NOS LIVROS DIDÁTICOS DE MATEMÁTICA	24
4.1	Estatística na BNCC	24
4.2	Estatística nos Livros Didáticos.....	27
5	CONSTRUÇÃO E ANÁLISE DOS GRÁFICOS À LUZ DA ESTATÍSTICA	34
5.1	Conceitos prévios e critérios gerais	35
5.1.1	Classificação de Variáveis Estatísticas	35
5.1.2	Elementos mínimos e integridade da representação gráfica	36
5.1.3	Escolha do gráfico em função do tipo de variável e do objetivo analítico	37
5.2	Gráfico de Barras	37
5.3	Histograma	40
5.4	Gráfico de Setores	43
5.5	Gráfico de Linhas	45
5.6	Gráfico de Dispersão	49
6	A LINGUAGEM R E O USO DE GRÁFICOS NA FORMAÇÃO DE PROFES- SORES DE MATEMÁTICA	53
6.1	Linguagem R e o RStudio	54
6.2	Linguagem R e Educação Estatística	57
6.2.1	R como ferramenta do pesquisador	57
6.2.2	R como apoio operacional	58
6.2.3	R como meio de investigação e produção	61
6.3	Transversalidade: usos do R em outros tópicos da Matemática	64
6.4	Recurso Educacional: o manual e como utilizá-lo	66
7	CONSIDERAÇÕES FINAIS	68
	REFERÊNCIAS	70

1 INTRODUÇÃO

A sociedade está conectada o tempo todo: celular, internet e redes sociais fazem parte da rotina. Nesse cenário, quase tudo gera dados: compras, buscas, vídeos, trajetos, cliques. Esse volume enorme de informações (o chamado “*big data*”) é tratado por algoritmos e por Inteligência Artificial (IA), como o próprio *ChatGPT* ou as recomendações do *TikTok*, por exemplo. Esses sistemas usam cálculos o tempo inteiro e influenciam decisões de empresas, governos e também da escola. Por isso, os dados e, consequentemente, a estatística ficaram ainda mais presentes no nosso dia a dia.

O uso de estatística foi impulsionado pelo advento da internet comercial, difundido principalmente pelo crescente acesso a computadores pessoais a partir de 1990. A internet começou a ser utilizada ainda no final da década de 1950, principalmente para usos militares, mas seu uso comercial por volta de 1993-1994 (Maynard, 2016) possibilitou inúmeras aplicações, principalmente educacionais. Como uma espécie de corrida de gato e rato, a internet, a linguagem de programação e o uso constante de estatística foram, ao mesmo tempo, impulsionadores e respostas ao crescente “rastros digital” que deixamos a todo instante quando nos conectamos online. Esse movimento fica mais claro quando observamos experiências concretas de gestão educacional orientada por dados no Brasil: no caso do Paraná, por exemplo, a Secretaria de Estado da Educação (SEED-PR) estruturou um painel de gestão baseado em *Business Intelligence* (BI) e, de modo mais amplo, consolidou um ecossistema de plataformas e avaliações com monitoramento contínuo, voltado ao acompanhamento de participação, desempenho e risco de evasão, articulando indicadores de rede, escola e estudante. A partir desse pioneirismo, essa mesma racionalidade de dados se dispersou, reaparecendo em projetos como a política de plataformização como os de São Paulo¹.

Computadores, internet e planilhas eletrônicas passaram a compor, de forma progressiva, a infraestrutura cotidiana do trabalho escolar e as práticas de ensino de Matemática. Nesse cenário, a estatística deixa de ser apenas um conteúdo “a mais” do currículo e passa a operar como linguagem de leitura e interpretação do mundo datafificado: é por meio dela que professores e estudantes podem formular perguntas, coletar e organizar dados, produzir representações e sustentar análises e decisões. Assim, as ferramentas digitais não apenas ampliam o repertório de recursos didáticos disponíveis, como também viabilizam abordagens investigativas com dados reais, aproximando o ensino de estatística das práticas contemporâneas de análise e visualização de dados.

No entanto, apesar desse enorme potencial e de ser reconhecida como área fundamental na educação básica brasileira, a realidade do aprendizado de estatística ainda é um desafio. Silva (2023) aponta que os estudantes apresentam desempenho insatisfatório nessa área. No currículo nacional, seu ensino é obrigatório e integra a Base Nacional Comum Curricular (BNCC;

¹ Para acompanhar as recentes mudanças na política educacional conferir Reis (2025) e Campanha Nacional pelo Direito à Educação (2025) para entender a plataformização e usos de dados na rede educacional de São Paulo e Pickler (2024).

BRASIL, 2018) como eixo de desenvolvimento de competências e habilidades, presente tanto no Ensino Fundamental (anos iniciais e finais) quanto no Ensino Médio. Além disso, a Estatística já figurava nos Parâmetros Curriculares Nacionais (PCN; BRASIL, 1998) do final dos anos 1990, documento que orientou o currículo nacional até a promulgação da BNCC.

Se o uso da estatística já é cada vez mais constante no mundo, o professor de Matemática ocupa um papel de destaque no ensino dessa disciplina dentro de suas aulas. Essa mesma disciplina de estatística é obrigatória nos cursos de formação de Licenciatura em Matemática no Brasil, onde os futuros professores devem compreender elementos constitutivos dessa ciência para também ensinar os estudantes.

Segundo Santos (2020b) o pensamento estatístico é descrito como um modo de pensar através do raciocínio lógico e analítico que permite observar o problema como um todo e compreender o porquê e como as investigações estatísticas são conduzidas. A autora também mobiliza a compreensão de Wild e Pfannkuch(1999) sobre o pensamento estatístico, em que, para os autores, essa “forma de pensar” pode ser dividida em cinco categorias: 1. Reconhecer a necessidade de dados; 2. Transnumeração; 3. Variação; 4. Uso de um conjunto de modelos; e 5. Conhecimento do contexto, conhecimento estatístico e sua síntese. Assim, de maneira geral, para Santos (2020b) se trabalhada essas cinco categorias nas aulas de Matemática, por meio de experimentação, coleta, tratamento e análise dos dados contextualizados, os estudantes conseguem compreender o ciclo investigativo da estatística e relacionar com as diversas utilidades no cotidiano.

Campos (2007) apresenta uma descrição do pensamento estatístico como sendo a relação entre modelos matemáticos a contextos dos problemas. Esse tipo de pensamento também está ligado à outros conceitos de entender a estatística, como o raciocínio estatístico que se concentra na aplicação correta de ferramentas para descrever dados e da literacia/letramento estatístico relacionado a leitura e interpretação dos dados. Assim, o pensamento estatístico vai além das meras aplicações de metodologias e ferramentas, mas busca uma interpretação diagnóstica da situação em que os dados foram gerados, apoiados em análises estatísticas e probabilísticas, não na subjetividade das impressões.

Dessa maneira, Silva (2023) reforça que para o desenvolvimento do pensamento estatístico na Educação Básica é preciso superar o foco no cálculo, principalmente apoiado em ferramentas computacionais, como o Excel, e investir em um ensino de estatística que complexe o nível de leitura e interpretação dos dados dos estudantes. Assim, eles devem entender causas, efeitos e relações através dos dados contextualizados, sugerir explicações para os comportamentos dos dados observados e argumentar a partir de medidas e gráficos.

No conjunto de pesquisas analisadas nesta revisão de literatura, que discutem Educação Estatística na formação de professores de Matemática, emergem desafios recorrentes na formação inicial. Lopes (2008; 2013) destaca, por exemplo, a necessidade de reavaliar a própria disciplina de estatística na formação inicial nos cursos de Licenciatura em Matemática, principalmente ao superar o mero ensino de cálculos e algoritmos para investir em uma disciplina que leve em consideração o uso de tecnologia e a formação pedagógica para o ensino de es-

estatística. Essa mesma consideração foi apontada também por Silva (2011), Rodrigues e Silva (2019) e Schreiber e Porciúncula (2019).

A partir desse panorama, torna-se possível problematizar algumas questões centrais: o que as pesquisas recentes sobre a formação de professores de Matemática têm investigado em relação à estatística? Quais prescrições são dirigidas aos docentes? Que concepções de estatística orientam essas pesquisas? Quais ferramentas tecnológicas são mobilizadas? Como a estatística aparece em documentos norteadores do currículo, como a BNCC? E, por fim, o uso de *softwares* para análise e visualização de dados é suficiente para formar professores, ou se limita a oferecer uma base superficial para aplicação de conteúdos estatísticos?

Neste trabalho, parte-se da hipótese de que uma parcela significativa das pesquisas que investigam a formação de professores para o ensino de estatística, especialmente aquelas que utilizam *softwares* como RStudio, tende a enfatizar a prescrição de métodos, sequências didáticas e passos a passos de aplicação, sem necessariamente promover uma apropriação mais ampla da linguagem e dos conceitos estatísticos pelos docentes. Em outras palavras, o foco recai muitas vezes na replicação de procedimentos, e não na compreensão conceitual que permitiria ao professor planejar e adaptar suas próprias propostas de ensino de forma autônoma.

Nesse contexto, coloca-se o seguinte problema de pesquisa: como professores de Matemática podem apropriar-se da linguagem R e do ambiente RStudio de modo a utilizá-los autonomamente na construção e análise de gráficos estatísticos em suas aulas, em diálogo com as demandas da Educação Estatística e com as orientações da BNCC?

Portanto, o objetivo geral desta pesquisa é mapear e analisar a produção acadêmica que discute o uso da linguagem R no ensino de estatística na Educação Básica, com foco na formação de professores de Matemática, e a partir dessas evidências, propor um recurso educacional que auxilie docentes a utilizar o R/RStudio na construção de gráficos estatísticos de forma autônoma.

Nesse contexto, a análise dessas produções, embora revele contribuições importantes, também evidencia uma tensão, pois, em parte das propostas, o R é mobilizado como recurso para viabilizar atividades e análises em estatística, sem que a apropriação docente da linguagem seja colocada no centro. Assim, permanece a lacuna de formar o professor para compreender escolhas e parâmetros e, a partir disso, produzir e adaptar atividades com autonomia. Além disso, estudos como de Lopes, (2008, 2013), Silva (2011), Rodrigues e Silva (2019) e Schreiber e Porciúncula (2019) sobre a formação inicial em Licenciatura em Matemática existem fragilidades na articulação entre estatística, ensino de Matemática computacional, pensamento computacional e linguagem de programação, o que pode dificultar a integração pedagógica do R na escola.

Foram identificadas lacunas formativas e desafios recorrentes, que orientam a elaboração de um recurso educacional voltado à docentes. Para suprir essa lacuna, o material proposto consiste em um manual digital que se objetiva a sanar dificuldades de professores usando linguagem R no ambiente RStudio. O manual abrange desde a instalação e configuração do *soft-*

ware R e do RStudio até a construção de diferentes tipos de gráficos estatísticos, tanto a funções nativas quanto a funções definidas por bibliotecas específicas. Os conjuntos de dados são oriundos de bases disponíveis e/ou criadas a partir de situações reais. Em cada etapa, foram explicados os parâmetros utilizados, de modo que o material possa ser facilmente utilizado em sala de aula. Com isso, espera-se que este recurso educacional favoreça a autonomia docente, reduzindo a dependência de sequências prontas e incentivando que professores desenvolvam e adaptem seus próprios métodos de ensino de estatística a partir de seus contextos.

Esta dissertação está organizada em seis capítulos, além dos elementos pré e pós-textuais. No Capítulo 2, apresenta-se a metodologia de Estado do Conhecimento. No Capítulo 3, realiza-se o estado da arte. O Capítulo 4 discute o lugar da estatística na Base Nacional Comum Curricular e nos livros didáticos de Matemática. No Capítulo 5, analisam-se, à luz da Educação Estatística, os diferentes tipos de gráficos contemplados no recurso educacional. Por fim, o Capítulo 6 trata da linguagem R e do uso no ensino de estatística e Matemática. O recurso educacional, em formato de manual digital para docentes, é apresentado em anexo.

2 METODOLOGIA

Na presente pesquisa, adota-se a metodologia de Estado do Conhecimento (EC), conforme sistematizada por Kohls-Santos e Morosini (2021). A definição de Estado do Conhecimento pode ser descrita como um tipo de pesquisa bibliográfica que mapeia, registra, categoriza e sintetiza a produção científica (teses, dissertações, artigos e livros) de uma determinada área de pesquisa, em um recorte temporal e espacial para mostrar o estado do tema, apoiando futuras pesquisas e mostrando lacunas e tendências de investigação do campo. Assim, essa metodologia ajuda a compreender o campo de pesquisa, localizar debates e oportunidades de pesquisas inovadoras.

Existem quatro etapas principais da metodologia do EC segundo Kohls-Santos e Morosini (2021). A primeira é a Bibliografia Anotada que consiste na seleção dos textos por descritores, com critérios para inclusão e exclusão dos trabalhos dentro do *corpus* que será analisado, além disso, monta-se uma tabela com dados demográficos do *corpus*, como ano, região, programa e palavras-chave. A segunda etapa é a Bibliografia Sistematiza, consiste em realizar uma leitura flutuante dos resumos para aprofundar o texto, também decide o que entra no *corpus* final, anotando os objetivos, métodos e resultados de cada estudo. A terceira etapa é a Bibliografia Categorizada, onde se reorganiza o *corpus* em categorias temáticas, explicando o viés epistemológico que fundamenta cada categoria. Por fim, a quarta e última etapa é a Bibliografia Propositiva, que vai além da mera revisão bibliográfica, consiste em registrar achados, proposições dos estudos e proposições emergentes do próprio pesquisador a partir da análise realizada. Essa última etapa é uma das propostas deste trabalho, sendo fundamental para realizar um bom Estado do Conhecimento.

EC é aqui articulada à Análise de Conteúdo (AC), conforme proposta por Bardin (2016). Enquanto o EC oferece a estrutura para mapear e organizar a produção científica de uma área, a AC fornece as ferramentas para ir além da descrição, permitindo realizar inferências e interpretações sobre o material coletado. É essa combinação que possibilita, por exemplo, identificar não apenas quais temas são pesquisados, mas como eles são abordados, revelando tendências, lacunas e vieses epistemológicos no campo.

Neste estudo, o EC foi aplicado para levantar e sistematizar textos que investigavam o ensino de Estatística na disciplina de Matemática, consultados nas bases de dados Google Scholar e no Catálogo de Teses e Dissertações da CAPES. Foram utilizadas palavras chave como “estatística”; “matemática”; “currículo”; “formação de professores” e “BNCC” para fazer um levantamento geral dos textos. Parte desses trabalhos foram utilizados para realizar a fundamentação teórica da dissertação. Outra parte, em torno de 25 artigos, textos em anais de eventos, teses e dissertações foram levantadas para discorrer sobre o ensino de estatística dentro da BNCC, na formação de professores e principalmente no uso da linguagem R e do RStudio. Sobre o último critério, foram utilizadas as palavras chave como “RStudio” e “linguagem R” somadas às outras palavras chave.

A finalidade desse levantamento é mapear e analisar como o campo de pesquisa brasileiro que investiga a Educação Estatística, principalmente com foco na formação docente, busca compreender o uso da linguagem R e o RStudio para ensinar estatística. O levantamento resultou em somente 11 produções, mostrando uma baixa adesão de investigações brasileiras. As pesquisas foram organizadas em seções e eixos temáticos, buscando explicitar tanto convergências quanto lacunas identificadas no campo.

No contexto do mestrado profissional de Matemática, foi elaborado de um recurso educacional, materializado em forma de um manual digital voltado aos professores de Matemática da Educação Básica, com foco em professores do Ensino Médio, mas não somente, para a construção de gráficos estatísticos em linguagem R por meio do ambiente RStudio. Esse recurso educacional foi concebido à luz das lacunas apontadas pelo EC, em especial a escassez de materiais que integrem Educação Estatística, formação docente e uso de R como ferramenta pedagógica.

A opção por focar especificamente em gráficos estatísticos decorre de duas justificativas principais. Em primeiro lugar, a BNCC enfatiza no eixo de Probabilidade e Estatística, a leitura, interpretação e produção de tabelas e gráficos ao longo da Educação Básica, o que torna esses objetos centrais para o letramento estatístico. Em segundo lugar, a delimitação do produto aos gráficos, como gráficos de barras, histograma, setores, linhas e dispersão, permite explorar aspectos didático-metodológicos relacionados à escolha de variáveis, escalas, legendas, cores e títulos, tornando a proposta mais exequível para professores que estão iniciando o uso da linguagem R.

O manual foi organizado em seções que apresentam, de forma progressiva, os primeiros passos no uso da linguagem R e no RStudio (instalação, interface, criação e salvamento de *scripts*), e em seguida, a construção e importação de bases de dados em formato .csv, incluindo exemplos gerados com apoio de ferramentas de inteligência artificial para simular contextos escolares. Na sequência, é mostrado a elaboração de diferentes tipos de gráficos tanto por meio de funções nativas do R quanto pelo pacote ggplot2 (Wickham, 2016), com ênfase em gráficos de barras, de setores, de linhas e de dispersão. Por fim, tem-se uma seção de replicabilidade, que discute possibilidades de adaptação dos códigos apresentados a distintas realidades escolares e conjuntos de dados. Ao longo do material, são disponibilizadas imagens (capturas de tela do RStudio, exemplos de gráficos e trechos de código) que buscam tornar o passo a passo mais acessível para docentes com pouca familiaridade prévia com programação.

Do ponto de vista metodológico, o recurso educacional dialoga diretamente com os resultados do EC ao propor um material que, simultaneamente, explora conteúdos de Estatística previstos na BNCC, incorporando a linguagem R como ferramenta de análise e visualização de dados e oferece ao professor exemplos de códigos comentados que podem ser adaptados, replicados e ampliados em diferentes turmas e níveis de ensino. Dessa forma, a pesquisa não apenas descreve o estado do conhecimento, mas também devolve à escola um recurso didático concreto, coerente com as demandas identificadas na literatura e com os objetivos formativos do mestrado profissional.

3 ESTADO DA ARTE

3.1 Revolução Digital, Dataficação e IA na Educação

A sociedade contemporânea vive hoje um cenário marcado por uma verdadeira revolução digital, que vai desde a criação dos primeiros computadores ao desenvolvimento e posterior massificação da internet. Nesse contexto, a organização da vida social se estrutura em redes interconectadas por fluxos constantes de informação, sustentadas por infraestruturas digitais como *datacenters*, plataformas e redes sociais. Não se trata apenas do uso da internet, mas de processos de produção, decisão e convivência mediados por sistemas informacionais que funcionam em tempo quase real. Essa lógica produz um modo de desenvolvimento baseado na inovação permanente, em ciclos contínuos de *feedback* de dados e na codificação do social em bases, métricas e modelos que orientam decisões econômicas, políticas e educacionais.

Ao mesmo tempo, a abundância de dados não garante neutralidade nem objetividade automática. Os dados resultam de escolhas sobre o que coletar, como medir e em que contexto, o que os torna sempre recortes parciais da realidade. A ideia de *Big Data* aprofunda essas tensões: grandes volumes de informação, produzidos de forma contínua e escalável, são frequentemente apresentados como capazes de gerar respostas mais precisas e eficientes, mas continuam dependendo de interpretações, modelos e interesses. No campo educacional, esse cenário torna fundamental uma formação estatística e crítica que possibilite compreender como esses dados são gerados, tratados e usados para orientar políticas, avaliações e práticas pedagógicas.

A dataficação e a digitalização estão reimaginando o futuro da educação. Não se trata somente do uso de tecnologia, mas sim de uma nova infraestrutura sociotécnica baseada em *software*, base de dados, algoritmos, que faz medições, compara dados e governa os processos educativos, desde a sala de aula até grandes políticas públicas. De maneira geral, as promessas desse futuro da educação são a personalização do ensino, *feedback* do aprendizado em tempo real, decisões baseadas em evidências e em eficiência institucional. Porém, os riscos e efeitos colaterais de uma educação voltada em *Big Data* são a intensificação da vigilância, a metrificação da aprendizagem - gerando uma performatividade para atender as métricas e não ao aprendizado -, deslocamento da autoridade sobre a educação para especialistas na técnica e empresas de dados e a redefinição do trabalho docente como a coleta e análise de dados (Williamson, 2017).

A sociedade contemporânea está entrando em uma fase histórica de colonialismo de dados, onde existe uma apropriação sistemática da vida social como um recurso para o capitalismo via captura, processamento e usos de dados. É como um novo modelo de anexação de recursos para sustentar o capitalismo, uma continuidade histórica do próprio colonialismo do século XIX realizado no continente africano e asiático (Couldry; Mejias, 2019).

A ideia de colonialismo digital e de dados também deve ser ressignificada quando observa-se o contexto de países do sul global. Em lugares como o Brasil, a relação do colo-

nialismo é diferente do que ocorre na Europa ou nos EUA. Pesquisas organizadas como a de Silveira, Souza e Cassino (2021) mostram os impactos diretos das *Big Techs*, como Google, Amazon, Facebook, Apple e Microsoft, no Brasil. Um exemplo importante presente no livro é como as universidades federais brasileiras sofrem com a lógica colonial dos dados, adotando plataformas estrangeiras e privadas, alinhando as dinâmicas de mercado e a privatização da educação pública. A obra de Faustino e Lippold (2023) também usa o conceito de colonialismo digital, mas investem em mostrar as relações imperialistas e racistas provocadas pela expansão da tecnologia e no uso de dados brasileiros. Pode-se perceber no Brasil essas consequências do colonialismo digital e de dados. Valeri (2023) discute a presença de viés racial¹. em sistemas de Inteligência Artificial, associado a bases de dados que refletem desigualdades sociais, enquanto Filho (2025) investiga o impacto ambiental da construção de datacenters no Brasil.

Para finalizar essa breve apresentação sobre os impactos dos algoritmos e do *Big Data* na sociedade, destaca-se Zuboff (2021) que traz uma tese central para entender os impactos de uma sociedade que cada vez mais produz dados e vive para interpretá-los. A autora elabora o conceito de capitalismo de vigilância para explicar um novo regime sócio econômico onde as plataformas digitais extraem um *superavit* comportamental dos usuários, convertem isso em previsões e vendem em mercados de futuros comportamentais. É a ideia de controle das emoções e das preferências dos usuários de plataformas de entretenimento, redes sociais e outros aplicativos. Os dados servem para vender publicidade, como é o caso de plataformas como *Instagram* e *TikTok*, ou mesmo intervir em eleições e disputas políticas, como o caso do Brexit em 2016 através da *Cambridge Analytica*².

Diante do exposto, questiona-se qual o impacto dessas ideias na vida social. Barbosa e Alves (2023) ressaltam os problemas trazidos por uma sociedade que vive para gerar dados. A reforma do Novo Ensino Médio e o uso de diversas plataformas criam novos processos regulatórios de vigilância, controle e performatividade que favorecem a privatização da educação, padronizam currículo e desintelectualizam os professores. Fator que aumenta o prejuízo formativo e agrava as desigualdades entre uma educação que não se baseia na performance para atingir os dados necessários. Essa platformização reforça a urgência de que, tanto alunos quanto professores se apropriem criticamente dos dados e pensar estatisticamente para agir com autonomia na tomada de decisões, sejam pessoais ou com impacto social amplo.

Na próxima seção, será discutido sobre o que é o pensamento estatístico e a sua importância para o ensino de Matemática através da Educação Estatística.

¹ Neste trabalho, "viés racial" refere-se à tendência de sistemas de inteligência artificial reproduzirem desigualdades presentes nos dados e em critérios de decisão, o que pode resultar em tratamentos sistematicamente mais rígidos para pessoas negras em comparação a pessoas brancas em aplicações sensíveis, como no apoio a decisões judiciais (Valeri, 2023).

² O Brexit em 2016 foi a votação dos ingleses para decidir se a Inglaterra deveria sair da União Europeia. Investigações mostraram que a empresa de consultoria que utiliza dados, *Cambridge Analytica*, influenciou na decisão. No caso, a empresa extraiu dados do *Facebook* sem consentimento dos usuários para vender uma análise aprofundada dos gostos dos usuários e influenciar em campanhas políticas.

3.2 Educação Estatística: entre o raciocínio, literacia e pensamento estatístico

Dentro do que se conceitua como Educação Estatística estão ideias fundamentais para entender o que é aprender estatística dentro da Matemática. Os conceitos de raciocínio, literacia e pensamento estatístico estão presentes em diversas obras que serão apresentadas nesta seção, além de sua discussão para entender sua importância dentro do ensino de Matemática. Pretende-se compreender como diversas pesquisas mobilizam essas ideias e dialogam com outros conceitos importantes para o ensino.

Uma importante tese sobre a Educação Estatística foi elaborada por Campos (2007) que articula diversos pesquisadores globais para entender os conceitos de literacia estatística, raciocínio estatístico e pensamento estatístico. Para o autor, a literacia estatística é a capacidade de ler, compreender, interpretar, analisar, avaliar e argumentar com base em informações estatísticas, usando corretamente a terminologia da área. Nesse sentido, o termo literacia tem relação com a ideia de letramento, capacidade de ler e interpretar o mundo. Portanto, a literacia estatística é a competência capaz de acompanhar e discutir argumentos que usam dados em diversas mídias e formatos, além das representações visuais como tabelas e gráficos.

Para Gal (2000), entende literacia como a habilidade de interpretar e avaliar criticamente informações estatísticas, como argumentos que circulam nas mídias e discutir opiniões sobre informações que mobilizam dados e estatística. Já para Watson (1997), a definição de literacia é a compreensão do texto e das implicações das informações estatísticas em seu contexto, sugerindo três estágios de desenvolvimento. O primeiro estágio é o terminológico, sendo o entendimento básico da nomenclatura da estatística. O segundo é a linguagem em contexto, onde é preciso compreender a linguagem e conceitos relacionados à estatística em situações sociais diversas. O terceiro é a atitude de questionamento, onde se aplica conceitos mais sofisticados para contradizer alegações e falsidades que não usam fundamentação estatística adequada. Somando a isso, temos a perspectiva de Garfield (1999) que enfatiza o entendimento da linguagem estatística (terminologia, símbolos, termos) e a leitura de tabelas e gráficos como elementos essenciais.

O conceito de raciocínio estatístico é a maneira de como as pessoas raciocinam com ideias estatísticas e dão sentido às informações, interpretando dados, lendo e usando representações, construindo e justificando procedimentos em situações reais através da estatística. Não há um consenso pleno de como desenvolver o raciocínio estatístico e como medir essa progressão. É necessário fazer pesquisa em sala de aula que observe os tipos de raciocínio, os pré-requisitos e os efeitos nos estudantes nas diferentes atividades realizadas (Garfield, 2002).

Ainda no raciocínio estatístico, Campos (2007), mobilizando uma série de referências (Garfield, 1998; Sedlmeier, 1999; delMas, 2002; 2004; Moore, 1998), mostra que treinar fórmulas não produz, por si só, o raciocínio estatístico. Esse tipo de pensamento surge e progride quando o estudante precisar pensar com variabilidade, amostragem e incerteza, explicando o motivo que escolheu métodos e medidas a partir de problemas reais e contextualizados.

Uma última parte é a diferenciação que Campos (2007) realiza do raciocínio matemático e do raciocínio estatístico. Para o autor, seguindo as ideias de Moore (1992); Batanero (2001); Branco (2000) e Kader e Perry (2006), os tipos de raciocínios não são a mesma coisa. A Matemática privilegia a dedução formal e a certeza, já a Estatística trabalha com dados em contexto, variabilidade e incerteza, mudando o tipo de raciocínio exigido. Moore (1992) afirma que a estatística tem substância, conceitos e modos de raciocínios próprios, não é possível tratar a área como apêndice da Matemática. Batanero (2001) dá o exemplo da média: saber calcular não implica entender o que a média representa nem quando usá-la, assim, a introdução do algoritmo antes do conceito prejudica a compreensão pois o raciocínio estatístico exige escolha e justificativa, não só execução. Como lembra Branco (2000), não é possível ensinar Estatística na mesma maneira que a Matemática pois são raciocínios distintos. A chave para distinguir a Estatística da Matemática é proposta por Kader e Perry (2006), em que a variabilidade é o conceito principal de diferenciação, onde as tarefas devem forçar o aluno a compreender e explicar a variabilidade, não só a média, por exemplo.

Quando se coloca o pensamento estatístico no centro, a meta do ensino é mudada. Segundo Campos (2007), o objetivo do aprendizado deixa de ser executar o algoritmo corretamente e passa a ser a construção e sustentação de argumentos com dados, explicitando a variabilidade, inferência e contexto. Parte daí a defesa de ensino e pesquisa com projetos e o ciclo completo de investigação são estruturais no ensino de estatística dentro da Matemática, abordando a coleta, análise, interpretação, discussão e a validação pública das considerações finais. Além disso, é preciso apresentar o processo de pesquisa por completo, antes do cálculo, pode melhorar o entendimento e a retenção dos conceitos pelos estudantes. Assim, o pensamento estatístico fica visível quando o aluno atua em problemas abertos com dados, escreve as justificativas explicando os procedimentos escolhidos e realiza a própria autocrítica da pesquisa, mostrando os limites dos métodos e do próprio entendimento. Portanto, desenvolver o pensamento estatístico é crucial para entender e viver no mundo atual.

Existe uma série de hábitos que os estudantes mostram quando estão colocando o pensamento estatístico em ação. Uma referência pode ser: 1. Planejar dados relevantes ao problema levantado, não somente selecionar qualquer dado ou os que estiveram mais disponíveis; 2. Revisar continuamente as variáveis e as maneiras de examinar os dados; 3. Ver o processo completo de coleta, interpretação e comunicação dos dados com revisão de cada componente; 4. Ser cético quanto a objetividade dos dados, observando fontes e vieses; por fim 5. Relacionar dados e contexto, além de interpretar as conclusões em linguagem não técnica, mostrando ressalvas e condições. A ênfase no contexto é fundamental do pensamento estatístico, já que as conclusões devem ser validadas no contexto do problema. Portanto, o pensamento estatístico significa articular modelos e contexto para explicar e delimitar as conclusões a partir de dados (Campos, 2007).

No trabalho de Campos (2007), o autor reforça que o raciocínio, a literacia e o pensamento estatístico estão interligados, mas não são iguais. A literacia da base da linguagem utilizada pela estatística e a interpretação. O raciocínio cuida do uso adequado do ferramental

conceitual e dos procedimentos. Já o pensamento integra os modelos estatísticos e o contexto para propor uma leitura diagnóstica da situação. Essa delimitação clara é fundamental para não reduzir a literacia a saber gráficos e tabelas, a literacia inclui linguagem, compreensão e julgamento dos dados e através deles. O raciocínio envolve entender e conectar os conceitos da estatística além de explicar os processos que conduzem as investigações. Por fim, o pensamento exige um entendimento da variabilidade, da modelagem e do contexto dos dados.

3.3 Formação docente e o professor reflexivo

A formação de professores de Matemática é um tema muito discutido, induzindo à pesquisas diversas. Um dos principais temas está relacionado à discussão entre o professor de Matemática pesquisador. O texto elaborado por D'Ambrosio e D'Ambrosio (2007) discute o lapso entre a pesquisa e a sala de aula, em que os autores mostram que pesquisas em Matemática, muitas vezes, são compostas por estudos com linguagem estética, forte apego a quadros teóricos e pouca utilidade imediata para a sala de aula, propondo uma aproximação entre a pesquisa e ação docente. As pesquisas acadêmicas, segundo os autores, também costumam tratar o docente como objeto de estudo para que, no fim, criticar seu conhecimento e sua prática docente. Isso alimenta a desconfiança entre a universidade e a escola, distanciando os dois campos de ensino de Matemática, dificultando até mesmo formações continuadas de professores. Uma das alternativas é justamente ter grupos colaborativos que incluam professores como pesquisadores, além de incentivar essa perspectiva na formação inicial, durante o curso de Licenciatura em Matemática, além de dar voz, agência e autonomia para o professor em seu trabalho pedagógico. Além disso, ao realizar esse movimento, também é dar voz para os próprios estudantes, incentivando o aprendizado de Matemática³.

Essas considerações de D'Ambrosio e D'Ambrosio (2007) são fundamentais, mas também devem ser localizadas temporalmente. O trabalho discute problemas ainda presentes, porém, sabe-se que as pesquisas se desenvolveram e hoje essa distância entre a academia e os professores se encurtou. Um dos exemplos máximos dessa tentativa de aproximação é o próprio Mestrado Profissional em Matemática em Rede Nacional (PROFMAT), criado em 2011, como uma política pública fundamental na melhoria do ensino de Matemática no Brasil, atuando também, em partes, como uma formação continuada dos professores.

Ainda nas políticas públicas, Fiorentini (2008), investiga como no Brasil dos anos 1990-2000 essas mesmas políticas impactam a formação inicial e continuada de professores de Matemática. No artigo, o autor situa dentro da globalização, a difusão das Tecnologias da Informação e Comunicação, e de reformas que normalizaram o currículo e avaliação em larga escala no Brasil. Exemplo disso é a Lei de Diretrizes e Base da Educação Nacional número 9.394 de 1996 e os Parâmetros Curriculares Nacionais (PCN) criados entre 1997-1998 e o Plano Naci-

³ No artigo também encontra-se um modelo prático para desenvolver o olhar investigativo do professor sobre a aprendizagem real dos alunos e seu papel reflexivo e investigativo sustentado pelas evidências dentro da sala de aula. Conferir D'Ambrosio e D'Ambrosio (2007).

onal de Educação (2001) que pressionaram por uma titulação de ensino superior para todo o magistério da educação básica até 2007, impulsionando, por exemplo, a expansão da Educação a Distância (EAD) e de licenciaturas aligeiradas.

Cursos emergenciais também foram criados, como as licenciaturas semipresenciais (ou a distância) em tempo reduzido e a Licenciatura Plena Parcelada em Matemática (LPPM, isto é, uma licenciatura organizada em regime parcelado, voltada a habilitar em nível superior professores em exercício na área de Matemática), acabando por reproduzir a racionalidade técnica⁴, pouco discutindo a prática educacional e transformando o ensino de Matemática até então (Fiorentini, 2008).

O outro modelo é o construtivo, sendo situado e colaborativo, onde é arquitetado pensando em uma formação no e com o trabalho real, partindo da realidade da escola que organiza o percurso da formação continuada. O lugar do professor é como protagonista e co-autor (professor-pesquisador), trabalhando em grupos colaborativos. O conteúdo e método dessa formação é a investigação-ação, com estudos em sala de aula, análise da produção dos alunos, projetos de intervenção e portfólios através de uma reflexão sistemática, na ação e sobre a ação pedagógica. A autoavaliação e a coavaliação são incentivadas, os critérios são negociados e as evidências da própria prática dos professores são levadas em consideração. O tempo da formação continuada é contínuo, com ciclos de planejamento focado na ação, observação e reflexão, sempre com acompanhamento na escola. Essa formação promove o desenvolvimento profissional dos professores, com autonomia e incentivo dos saberes práticos com um impacto real na aprendizagem.

Albuquerque e Gontijo (2013) discutem a formação inicial e continuada em um mapeamento de pesquisas na área e a perspectiva formativa dominante entre educadores matemáticos do início dos anos 2010. Os autores mostraram que existem alguns problemas estruturais recorrentes: pouca incorporação de resultados de pesquisas nas licenciaturas, permanecendo uma perspectiva tradicional, a prática de ensino e estágio que estão ao final do curso, são disciplinas burocratizadas e pouco reflexivas, existe um isolamento entre as Instituições de Ensino Superior e as redes de ensino e a desarticulação entre conhecimentos matemáticos e pedagógicos. Mesmo com a expansão dos cursos, persiste uma escassez de professores de Matemática nas redes públicas e baixo desempenho discente, reforçando a necessidade de políticas de valorização da carreira docente e melhoria nas condições de trabalho para fixar e desenvolver profissionais.

Para Giraldo (2018), ainda persiste uma ruptura do que se ensina na universidade dentro do curso de Licenciatura em Matemática e o que o professor realmente precisa saber para ensinar na escola básica. Pois muitas licenciaturas ainda seguem justapondo blocos de conteúdo matemático e didática, mas sem articulação efetiva entre as disciplinas. Em seu artigo, Giraldo (2018) cita autores como Chevallard (1991) e Chervel (1990) para mostrar que a escola não é

⁴ No artigo de Fiorentini (2008), a racionalidade técnica é a lógica que trata a docência como uma mera aplicação de técnicas prescritas por especialistas, sendo que o professor é um mero executor de métodos e materiais padronizados, sem levar em consideração a prática e experiência docente.

apenas uma didatização ou vulgarização da ciência, mas ela própria também produz conhecimento que deve ser levado em consideração pela universidade e estar conectado a formação dos futuros docentes. Assim, a Matemática para o ensino não deve ser um bacharelado mutilado, é um outro tipo de saber que precisa ser também discutido na universidade, em diálogo com a escola.

Em síntese, os trabalhos aqui discutidos convergem para uma ideia central: formar o professor de Matemática como profissional reflexivo e pesquisador da própria prática, superando a racionalidade técnica (Fiorentini, 2008) e da justaposição de conteúdo como prioridade e depois didática (D'Ambrosio; D'Ambrosio, 2007). É necessário construir pontes concretas entre a pesquisa na universidade e o professor na sala de aula dentro da escola. Projetos colaborativos que invistam na autonomia dos futuros professores têm potencial gigantesco (Malheiros, 2014), onde a Matemática para o ensino deve também ser um saber próprio da docência, articulando a universidade com a escola.

Na próxima seção será discutido como o currículo da formação dos futuros professores tematiza o ensino de estatística no Brasil, quais são os maiores desafios e avanços nesse campo.

3.4 Licenciatura em Matemática e o lugar da estatística no Brasil

Ao investigar o ensino de estatística e da probabilidade na educação básica relacionando a formação de professores, Lopes (2008) sustenta que essas áreas são fundamentais no ensino desde os anos iniciais. Aprender estatística e probabilidade sustenta uma cidadania crítica, principalmente pela leitura da informação na mídia, como a interpretação de tabelas, gráficos e índices. A estocástica (jogo de palavra que junta estatística e probabilidade) não deve ser um tópico isolado e algorítmico, sem existir uma problematização e contexto. Sem isso não se desenvolve pensamento estatístico e probabilístico. Além disso, a autora defende uma visão curricular não linear, rompendo com o mito da linearidade: os conceitos, seus usos e aplicações podem emergir de problemas relevantes, não apenas de pré-requisitos rígidos de pesquisa ou em um engessamento de currículo.

O ensino e aprendizagem da estatística também é desenvolvido por Lopes (2008). Para a autora, a resolução de problemas deve ser o eixo: investigar questões reais, planejar coletar de dados, organizar e apresentar dados, interpretar e principalmente argumentar conclusões, evitando realizar somente cálculos. A variabilidade, a incerteza e a aleatoriedade devem aparecer cedo no ensino de estatística, com simulações e experimentos (reais ou digitais com uso de tecnologia, por exemplo). A Educação Estatística deve ser crítica, trabalhando com dados públicos e problemas enfrentados pelo público, relacionados ao meio ambiente, saúde, economia, sustentando à análise e a tomada de decisões com base nas informações.

Diante do exposto, Lopes (2008) sugere recomendações concretas para a formação inicial e continuada: 1. Integrar a estocástica ao longo do curso e não só no fim ou na segunda metade e conectá-la ao Estágio docente e a Didática por meio de projetos investigativos com

dados reais; 2. Usar metodologias ativas investigando na resolução de problemas, simulações, experimentos e portfólios que enfatizem a comunicação de resultados; 3. Explorar a tecnologia para a análise e visualização dos dados, se libertando do cálculo mecânico, discutindo também os limites da inferência e viés na análise dos dados; e 4. Avaliar o que se quer formar nos futuros professores, como leitura de variabilidade, planejamento de estudos, justificativa de decisões e argumentação com dados.

Em outro estudo, Lopes (2013) argumenta e defende o que deveria ser uma disciplina de estatística, diferenciando a estatística da Matemática, já que para a primeira os dados são números com contexto e a variabilidade é central. Assim, para a estatística, é necessário ilustrações reais e juízos críticos, não somente teoria e procedimentos metodológicos. A autora mostra que as diretrizes para a Educação Estatística deve enfatizar a literacia e o pensamento estatístico, usando dados reais e priorizando o entendimento conceitual ao invés de passar fórmulas, adotando uma aprendizagem ativa e uso de tecnologias para análise e visualização de dados.

Para finalizar esta breve discussão sobre a formação de futuros professores de Matemática para ensinar estatística, o artigo de Rodrigues e Silva (2019) é ilustrativo da importância dessa discussão. Os autores investigam a construção e consolidação da disciplina de Estatística nos currículos dos cursos de Licenciatura em Matemática no Brasil através das análises dos Projetos Pedagógicos de Curso (PPCs). Investigando 190 documentos de todo o Brasil através da metodologia de Análise de Conteúdo, os autores apostaram em alguns resultados importantes, indo de encontro com a pesquisa de Silva (2011), a disciplina de estatística existe, estando em 97% das matrizes curriculares analisadas, porém a abordagem predominante é a conceitual/algóritmica, com pouca ênfase no conhecimento pedagógico do conteúdo voltado para os futuros professores mobilizar essas ideias na Educação Básica. Além disso, a estatística aparece nos currículos em quatro categorias: 1. Estatística Descritiva; 2. Estatística Inferencial; 3. Probabilidade/Análise Combinatória; 4. Ensino de Estatística na Educação Básica. Porém, esta última categoria é quase uma raridade nos documentos.

Outra consideração importante é que muitos PPCs dedicam menos de 60 horas para a disciplina de Estatística, sendo considerado baixo por Rodrigues e Silva (2019) considerando as demandas da Educação Básica e para atender os critérios da BNCC. Por fim, as referências bibliográficas apresentam um predomínio de manuais técnico-conceituais com escassez de bibliografia voltada à formação de futuros professores pensando na Educação Estatística. Assim, o licenciado em Matemática sai do curso sabendo realizar contas e testes, mas pouco preparado para ensinar estatística através da investigação com dados, discutir variabilidade e incerteza e comunicar os resultados das pesquisas.

4 ESTATÍSTICA NA BNCC E NOS LIVROS DIDÁTICOS DE MATEMÁTICA

A consolidação da Educação Estatística na Educação Básica brasileira não depende apenas da presença da Estatística nos documentos oficiais, mas também de como esses encaminhamentos curriculares são materializados em materiais de uso cotidiano pelos professores, em especial os livros didáticos de Matemática. Entre o currículo prescrito, representado pela BNCC, e o currículo efetivamente ensinado em sala de aula, os livros didáticos ocupam um lugar central de mediação, servindo, muitas vezes, como principal referência para a seleção de conteúdos, organização de sequências e proposição de atividades. Assim, compreender como a Estatística é tratada nesses dois planos, o normativo e o didático, é fundamental para avaliar em que medida as intenções formativas da BNCC se convertem em oportunidades concretas de letramento estatístico para os estudantes.

4.1 Estatística na BNCC

Nesta seção destaca-se a centralidade que a BNCC teve como documento norteador para normatizar, reorganizar e ampliar de maneira decisiva o lugar na Estocástica (estatística, probabilidade e, em alguns casos, combinatória) no ensino de Matemática na educação básica, ao explicitá-la como unidade temática desde a Educação Infantil até o Ensino Médio. Nesse sentido, a BNCC destaca tanto o conteúdo quanto as competências e habilidades necessárias para aprender estatística. O foco não é listar as habilidades e as competências relacionadas à estatística, mas sim levantar e apresentar investigações sobre a BNCC e o ensino da mesma.

As primeiras análises aqui apresentadas, como as de Giordano, Araújo e Coutinho (2019) e Giordano e Lima (2023), mostram que a BNCC avança em relação aos Parâmetros Curriculares Nacionais (PCN) ao normatizar habilidades ligadas à leitura, produção e interpretação de tabelas e gráficos, a argumentação com dados e ao letramento estatístico em sentido amplo. Estudos posteriores aprofundam esse diagnóstico em diferentes recortes, a saber, Lima *et al.* (2022) examinam a ampliação da Estocástica ao longo de toda a escolaridade; Silva e Castro (2023) e ?? focam no ciclo de alfabetização e os anos iniciais; Pontes *et al.* (2019) discutem o trabalho com probabilidade e estatística por pedagogos; Santos e Bulegon (2023) analisam as potencialidades da BNCC para o desenvolvimento do pensamento crítico e, por fim, Palheta, Porciúncula e Giordano (2025) investigam o diálogo entre a BNCC, currículos estaduais e a metodologia do Projeto de Aprendizagem Estatístico (PAE). Em conjunto, esses trabalhos oferecem um panorama importante e consistente sobre a relação e os impactos da BNCC para a Educação Estatística e das tensões que se colocam para sua efetiva implementação.

O artigo de Giordano, Araújo e Coutinho (2019) é fundamental para entender a recepção e as primeiras impressões da BNCC, sendo esse um documento normativo orientador de 10 competências gerais, destacam-se algumas relacionadas a estatística, como investigar (componente 2); valorizar diversidade (componente 6); argumentar com dados (componente 7) e atuar

com autonomia (componente 10), o que segundo os autores, todos esses componentes são compatíveis com a estatística.

No que concerne à unidade temática, Giordano, Araújo e Coutinho (2019) mostram que a probabilidade e estatística e os objetos e habilidades (como a EF01MA21 e EFO1MA22) estão presentes desde o 1º ano, reforçando ideias como coleta, organização, representação, interpretação de dados e pesquisa com os estudantes. Já no Ensino Médio, observam-se habilidades como a interpretação de histograma, bloxpot, ramos e folhas (EM13MAT407); análise de duas variáveis usando tecnologias da informação (EM13MAT510) e espaços amostrais e hipóteses (EM13MAT511), com ressalvas ao uso de *softwares* e investigação por projetos. Em conclusão, enquanto os PCN sugerem projetos e temas transversais, abordando a estatística de forma orientadora, a BNCC é normativa, assegurando a estatística no currículo, abrindo possibilidades para projetos investigativos em sala de aula.

Giordano e Lima (2023) investigam como a BNCC viabilizou o ensino de tabelas e gráficos na unidade temática da probabilidade e estatística e como isso impactou o Currículo Paulista. Na Educação Infantil a representação de gráficos de maneira básica é prevista, na habilidade EI03ET08. No Ensino Fundamental as habilidades do 1º ao 9º ano também orientam para ler, comparar, planejar pesquisas, escolher e construir gráficos e detectar erros na mídia. Por fim, no Ensino Médio, vê-se análise crítica de tabelas, amostras e gráficos, com usos de softwares e diagramas nas habilidades.

No trabalho de Palheta, Porciúncula e Giordano (2025) o objetivo é analisar em que medida o currículo do Ensino Médio da BNCC e do Documento Curricular do Estado do Pará (DCEPA) se conectam à metodologia de Projeto de Aprendizagem Estatístico (PAE). Como principal achado, tanto a BNCC quanto o DCEPA não explicitam o trabalho com projetos na metodologia do PAE, mas ainda oferecem pistas que podem sustentar práticas pedagógicas com dados reais contextualizados e finalidades cidadãs. No artigo, os autores reconhecem que a BNCC garante a presença do ensino de estatística ao longo de toda escolaridade, persistindo desafios ligados aos materiais didáticos, a formação inicial e continuada. Os dois documentos operam em uma lógica curricular por competências e habilidades, em que o DCEPA não traz novidade substantiva frente à BNCC.

Enquanto isso, A pesquisa de Santos e Bulegon (2023) busca investigar como os conceitos de estatística contribuem para o pensamento crítico na Educação Básica e propor atividades didáticas, tomando a BNCC como referência e a Taxonomia de *Bloom* Revisada como chave analítico-interpretativa. Nesse sentido, o artigo evidencia que há diversas oportunidades de desenvolver o pensamento crítico por meio da estatística, mas que tarefas contextualizadas através do cotidiano do aluno engajam mais e elevam a eficácia da aprendizagem, desde que haja intencionalidade pedagógica do professor.

Já o trabalho de Silva e Castro (2023) é um importante levantamento bibliográfico de pesquisas em Educação Estatística relacionadas ao ciclo de alfabetização (1º e 2º ano do Ensino Fundamental - Anos Iniciais) antes e após a implementação da BNCC. O ensino real da estatística no ciclo de alfabetização é raso e superficial, onde muitas práticas são resumidas em:

professor mostra uma tabela ou gráfico pronto, os alunos apenas observam ou respondem perguntas diretas, sem investigação, coleta de dados nem reflexão crítica, contrariando diretamente a BNCC. O problema mais repetido nos estudos revisados é a falta de formação de professores, onde relatam insegurança com estatística, muitas vezes só repetindo o livro didático, sem formação continuada consistente. Os materiais didáticos são fracos, trabalhando a estatística de forma mecânica, quase sempre somente com interpretação de tabela, sem construção e tomada de decisão.

Em contraponto, a revisão de Silva e Castro (2023) aponta aspectos positivos em relação à BNCC. Esta é mais clara e explícita do que os PCN sobre estatística nos anos iniciais, o que inclui coleta, organização, análise e comunicação de dados já no 1º ano, incentivando práticas investigativas e promovendo leitura crítica e argumentação. Ao mesmo tempo, a BNCC não aborda a formação de professores, aspecto fundamental segundo Silva e Castro (2023) - assim como as outras pesquisas apresentadas nesta dissertação. Essa implementação da BNCC é desigual, criando um abismo entre o documento curricular norteador e a prática. Por fim, algumas habilidades só aparecem tardiamente, como a leitura e interpretação de gráficos de linha somente no 5º ano. Em síntese, para Silva e Castro (2023), a BNCC avançou ao sistematizar a estatística nos anos iniciais, mas os avanços reais na escola são lentos, sendo três problemas estruturantes no Brasil: 1. falta de formação continuada de professores; 2. livros didáticos superficiais; 3. escassez de pesquisas sobre aprendizagem de estatística no ciclo de alfabetização.

Ainda sobre os anos iniciais do Ensino Fundamental a partir da BNCC, Pontes *et al.* (2019) tem como objetivo discutir o ensino de estatística na unidade temática probabilidade e estatística a partir de um estudo bibliográfico. Os autores se propuseram a entender como professores pedagogos estão desenvolvendo o trabalho com estatística após as mudanças trazidas pela BNCC, além de buscar compreender quais mudanças foram essas e os impactos na implementação dessa unidade temática nos anos iniciais. No artigo, os autores elencam uma série de problemas estruturais no ensino de estatística na educação brasileira. Um deles é o histórico de baixa prioridade da estatística na escola, onde mesmo após os PCNs o ensino de estatística é deixado de lado. Isso também ocorre nos livros didáticos, onde o conteúdo é um dos últimos do livro. Isso significa, já que muitos livros didáticos de Matemática também ocupam o espaço de currículo orientando a prática docente, a disciplina de estatística é vista por último e às pressas. A formação inicial em Educação Estatística é precária, muitos pedagogos não cursaram sequer estatística básica na graduação, somando a isso uma baixa carga horária de Matemática. Destacam-se também uma forte dependência de livro didático já que a formação inicial e continuada é ineficiente. O problema é que muitos livros não fazem um trabalho consistente ou contextualizado com estatística.

Além disso, existe também um distanciamento entre a BNCC e a prática em sala de aula, onde, segundo Pontes *et al.* (2019), muitas formações continuadas feitas nas redes básicas costumam focar em ler o documento, sem realizar uma discussão com profundidade, não transformando as habilidades previstas na BNCC em sequências de ensino, projetos e ativida-

des investigativas. Diante do exposto, os autores concluem que, embora a BNCC represente um avanço significativo ao normatizar e explicitar a importância e a presença da probabilidade e estatística desde os anos iniciais e enfatizar o letramento estatístico como condição para atuação crítica na sociedade, existe um forte distanciamento da prescrição do documento oficial e a prática docente.

A partir dos estudos apresentados, os resultados citados apontam a ideia que a BNCC institucionalizou um currículo mais robusto para a Estocástica, mas ainda não garante, por si só, avanços equivalentes nas práticas de sala de aula. Se por um lado os estudos evidenciam a ampliação de habilidades relacionadas à coleta, organização, representação e análise de dados, bem como a abertura para projetos investigados e para o desenvolvimento letramento estatístico e do pensamento crítico (Giordano, Araújo e Coutinho (2019); Giordano e Lima (2023); Santos e Bulegon (2023); Samá e Silva (2020)), por outro lado apontam para um conjunto de desafios estruturais, a saber, fragilidades na formação inicial e continuada em estatística e probabilidade, dependência de livros didáticos, muitas vezes superficiais, e distância entre currículo prescrito e práticas efetivas (Lima *et al.* (2022); Silva e Castro (2023); Pontes *et al.* (2019)).

Além disso, a análise de currículos locais, como o Currículo Paulista e o Documento Curricular do Estado do Pará, indicam que muitas vezes existe apenas uma transposição literal da BNCC, com poucas inovações substantivas e sem explicitação do trabalho com projetos de aprendizagem estatística (Giordano e Lima (2023); Palheta, Porciúncula e Giordano (2025)).

Dessa maneira, a BNCC cria condições institucionais favoráveis à Educação Estatística, mas os estudos ressaltam que a materialização dessas potencialidades depende de políticas formativas mais consistentes, de tempos e recursos pedagógicos adequados e de uma cultura escolar que incorpore práticas investigativas com dados em contextos reais, permitindo a investigação com dados.

4.2 Estatística nos Livros Didáticos

Esta seção reúne pesquisas que analisam como os livros didáticos de Matemática abordam a Estatística, com o objetivo de compreender as oportunidades formativas oferecidas aos estudantes. Partindo do pressuposto de que o livro didático é o principal recurso pedagógico na maioria das salas de aula, as investigações aqui apresentadas avaliam se esses materiais promovem um letramento estatístico crítico ou se limitam o ensino à aplicação de técnicas e fórmulas. O foco comum é, portanto, verificar se as atividades, a organização dos conteúdos e as orientações ao professor favorecem uma abordagem investigativa, alinhada às demandas de uma formação para a cidadania.

Friolani (2007) analisa como os livros didáticos abordam Tratamento da Informação e em que medida essa abordagem favorece a construção do pensamento estocástico. O objetivo é verificar a abordagem do tema em livros do Ensino Fundamental e avaliar se tal abordagem favorece a construção do pensamento estocástico, assumindo o livro didático como apoio pedagógico central na prática docente. O trabalho analisou as coleções 1. *Oficina de Conceitos* de

Spinelli e Souza (2022); 2. *Matemática em Movimento* de Longen (2004) e 3. *Matemática Hoje se Ensina Assim* de Bigode (2002).

Friolani (2007) conclui que a Coleção 1, se aplicada conforme as orientações do material e alinhada aos PCNs, pode levar os alunos ao letramento estatístico cultural, ou seja, compreender termos básicos presentes na mídia e, dependendo da ampliação pelo professor, pode avançar ao letramento funcional, como conversar, ler e escrever informações com termos científicos coerentes. O ponto de diferença é a natureza das tarefas: nessa coleção, aparecem atividades que envolvem pesquisa e resolução de problemas em etapas como definir problemática, coletar dados, representar, interpretar e analisar, fazer deduções e tomar decisões.

Com relação às Coleções 2 e 3, Friolani (2007) faz uma análise conjunta porque identifica perfis equivalentes. O achado central é que a pouca exploração de atividades de Tratamento da Informação possibilita apenas desenvolvimento parcial de conceitos e leva, no máximo, a algumas habilidades de letramento cultural, principalmente porque as tarefas tendem a ser de interpretação de dados já prontos em tabelas e gráficos, com leitura direta, sem exigir etapas de pesquisa, coleta, organização, representação, análise e tomada de decisão. Justamente por isso, não atendem às orientações dos PCNs nem às pesquisas da área.

Nesse sentido, Friolani (2007) aponta que essas duas coleções não apresentam uma organização de conteúdos de Tratamento da Informação que contemple conceitos básicos necessários e, portanto, não favorecem o desenvolvimento do pensamento estocástico.

Por fim, do ponto de vista interpretativo, Friolani (2007) sustenta que, se o livro didático é entendido como principal apoio pedagógico (Lajolo, 1996), então a predominância de tarefas de leitura direta em duas coleções tende a reforçar uma prática de estatística pouco investigativa. Além disso, o texto registra um problema estrutural: mesmo quando há orientação curricular, pode haver insuficiência de preparo docente para ensinar esses conceitos de modo a favorecer pensamento estocástico, em função de formações marcadas por métodos tradicionais e tecnicistas.

O trabalho de Neto (2008) busca investigar se os livros didáticos de Matemática aprovados no Programa Nacional do Livro Didático para o Ensino Médio (PNLEM, 2006) oferecem condições para que estudantes desenvolvam letramento estatístico (isto é, capacidades de interpretar, avaliar criticamente e discutir informações estatísticas) e, sobretudo, em que nível esse letramento é exigido e promovido. A pesquisa parte do problema de verificar se as coleções didáticas do Ensino Médio favorecem (ou não) a construção do letramento estatístico, em vez de reduzirem a Estatística a um ensino tecnicista de fórmulas e procedimentos. Os resultados demonstram que as atividades permitem reconhecer termos e ler informações básicas, mas não consolidam competências mais avançadas de interpretar, avaliar e discutir dados de modo mais autônomo. O autor também conclui que a maioria dos livros não estaria alinhada às recomendações oficiais e, por isso, não desenvolveria suficientemente competências para interpretar, avaliar e discutir dados, e, conseqüentemente, o estudante não estaria bem preparado para tomada de decisão em situações cotidianas se dependesse apenas do que o livro oferece. Já no plano didático-curricular, há um problema recorrente, os conteúdos estatísticos tendem

a ser concentrados em um único volume ou bloco, em vez de serem distribuídos ao longo dos volumes e articulados a diferentes capítulos e contextos.

Como encaminhamento, o autor defende que os materiais deveriam trazer mais atividades que envolvam conteúdos estatísticos e uma distribuição mais integrada, evitando a ilha de estatística no livro. Por fim, ao comparar com análises do Ensino Fundamental, o autor percebe convergência com diagnósticos anteriores, ou seja, a predominância de um ensino tecnicista com procedimentos e fórmulas tende a apoiar apenas o letramento estatístico em nível cultural.

Conti (2013) analisa como os conteúdos de Estatística aparecem em livros didáticos de Matemática do Ensino Médio utilizados na rede estadual do município de Criciúma (SC). Os resultados podem ser classificados em três grandes considerações sobre os livros didáticos. Primeiramente, evidencia-se uma diversidade de materiais e a ausência de padronização entre as coleções, por exemplo, a Estatística é distribuída 33,3% dos conteúdos no 2º e 3º volumes, 50% concentra-se no 3º volume, e 16,6% traz Estatística apenas no 2º volume, o que impacta diretamente quando e como o tema aparece ao longo do Ensino Médio. Em relação a organização curricular do conteúdo de Estatística, a autora mostra que, nas coleções em que a Estatística aparece em dois volumes, há uma lógica de complementaridade, em que o volume 3 dá continuidade ao volume 2 para abranger conteúdos considerados essenciais. Já nas coleções que não possuem uma sequência, o conteúdo se repete, ainda que o volume 3 (o último) seja o mais completo. Em relação as tecnologias, elas aparecem mais como sugestão didática do que como integração estruturante do conteúdo. Assim, as sugestões de tecnologias (*sites*, *softwares* como *Microsoft Excel*, calculadora, mídias e *links*) aparecem apenas na seção de orientações didático-pedagógicas, ou seja, como recomendação de recurso e não necessariamente como elemento articulado ao desenvolvimento conceitual do capítulo. A autora ancora essa discussão na literatura sobre dificuldades docentes de inclusão digital, apontando inclusive resistência e desafios do professor em acompanhar a evolução tecnológica.

Por fim, a autora destaca que há recorrência de dados do cotidiano e da *internet* e de temas atuais, por exemplo de temas relacionados a extinção de fauna e flora, energias renováveis, produção de lixo, desperdício de alimentos, poluição de veículos, além de situações como idade, peso, salário, e também presença de questões do ENEM em ao menos uma coleção. Contudo, a autora observa que, em várias propostas, pede-se pesquisa e construção de gráficos, mas nem sempre se especifica como a tecnologia deve ser usada na atividade, ficando como orientação genérica.

O trabalho de Corrêa (2015) apresenta uma investigação onde o foco é entender ensino e a aprendizagem da amplitude estatística no Ensino Médio, buscando identificar elementos que favoreçam uma melhor abordagem desse conteúdo na escola. O objetivo central da pesquisa é fornecer elementos para qualificar a abordagem da amplitude estatística no Ensino Médio, olhando para o modo como o tema aparece na literatura da área, em livros didáticos e em exames externos. Para examinar livros didáticos, usa como base dados do Guia do Livro Didático do PNLD de 2014. A partir desse levantamento, são selecionadas cinco coleções associadas a uma maior incidência de atividades de estatística. Segundo o texto de Corrêa (2015), as cinco

coleções foram: 1. *Coleção 1* (Editora Saraiva, 8ª edição, 2013); 2. *Novo Olhar Matemática* (Souza, Garcia e Marques, 1ª edição, 2013); 3. *Matemática: Contexto e Aplicações* (Dante, 2ª edição, 2013); 4. *Matemática: Ciências e Aplicações* (Iezzi et al., 7ª edição, 2013); 5. *Matemática: Paiva* (Paiva, 2ª edição, 2013). O texto registra um diagnóstico recorrente: o número de exercícios de Estatística nas primeiras séries do Ensino Médio tende a ser reduzido, o que ajuda a explicar por que certos tópicos ficam para o fim e, muitas vezes, são pouco trabalhados em sala.

O autor reforça que o ensino de Estatística demanda interpretação e análise, e não apenas a execução de procedimentos. O texto destaca, por exemplo, a crítica à ideia de trabalhar atividades sem que a problemática esteja presente, pois isso esvazia o sentido investigativo do conteúdo. Dessa maneira, existe um desalinhamento recorrente entre saber fazer contas e aprender estatística.

Nos livros didáticos, segundo Corrêa (2015) a presença de Estatística pode ser limitada ou concentrada, sobretudo nas primeiras séries do Ensino Médio, o que dificulta a continuidade didática e a consolidação conceitual. Esse ponto é articulado ao diagnóstico de Santos *et al.* (2008) sobre o caráter reduzido de exercícios estatísticos no início do Ensino Médio. Ainda segundo Corrêa (2015), a escolha das coleções com maior incidência de Estatística indica uma preocupação metodológica do estudo: observar livros onde, em tese, haveria mais oportunidades de trabalhar conceitos e medidas, como a amplitude estatística. Isso permite comparar como diferentes coleções tratam o tema e quais tipos de atividade aparecem com mais frequência. Por fim, ao apontar que Probabilidade e Estatística representam 27% das questões no ENEM de 2014, o autor sustenta que o conteúdo não é periférico e que, portanto, o trabalho escolar com Estatística precisa ser consistente e não residual.

A pesquisa de Santos (2020a) tem como objetivo analisar como a Estatística é trabalhada em uma coleção de livros didáticos de Matemática do Ensino Fundamental II utilizada em uma escola pública de Araguaína, em Tocantins, verificando a aderência à BNCC e o potencial de desenvolvimento do letramento estatístico. A autora destaca a centralidade do livro didático como instrumento pedagógico recorrente na escola e, conseqüentemente, como objeto estratégico para verificar se o ensino de Estatística favorece ou limita o desenvolvimento do letramento estatístico. A coleção analisada foi *A Conquista da Matemática* (4. ed., 2018), Volumes do 6º, 7º, 8º e 9º anos de Giovvani Júnior. Os resultados indicam que, na coleção *A Conquista da Matemática* (6º ao 9º ano), a Estatística aparece com distribuição e aprofundamento progressivos, combinando leitura e interpretação de tabelas e gráficos com propostas de coleta de dados e realização de pesquisas, em geral alinhadas às habilidades previstas na BNCC. No 6º ano, o tema surge sobretudo em seções extras (sem unidade própria), com baixa participação no volume (10 páginas; 3,61% e 40 atividades; 5,61%), e o manual do professor apresenta uma definição conceitualmente inadequada ao tratar a Estatística como parte da Matemática; ainda assim, há tarefas com fontes reais e temáticas socialmente relevantes, o que favorece o letramento estatístico, embora haja pouca exploração de tecnologias e a necessidade de fortalecer o trabalho investigativo. No 7º ano, a Estatística ganha espaço e organização, com 26 páginas

(9,38%) e 54 atividades (7,82%), incluindo construção e interpretação de representações, média e amplitude, e pesquisas censitárias e amostrais; a definição do conceito melhora no manual, mas permanece ausente no material do aluno, e a análise crítica de gráficos da mídia ainda aparece pouco, demandando complementações docentes.

Nos volumes finais, a autora identifica maior densidade e melhor articulação entre investigação estatística, representações e discussão crítica. No 8º ano, apesar de persistir a fragilidade conceitual retomando a Estatística como parte da Matemática, há aprofundamento em medidas de tendência central e, sobretudo, em propostas de pesquisa, 30 páginas (10,83%) e 44 atividades (6,56%), com tarefas que articulam gráficos, tabelas e infográficos a temas socialmente relevantes e potencial interdisciplinar. No 9º ano, conclui que é o volume com maior contribuição para o pensamento estatístico, por explicitar discussões sobre gráficos na mídia, representação correta, uso de planilhas eletrônicas e postura crítica diante de erros e conclusões indevidas (21 páginas; 7,58% e 34 atividades; 5,18%); embora haja variação de níveis de letramento estatístico, ainda existem menos atividades matematicamente críticas do que seria esperado, reforçando a necessidade de ampliação didática pelo docente para elevar a criticidade ao final da coleção.

Nas conclusões gerais, Santos (2020a) destaca que a coleção apresenta presença mensurável de Estatística em todos os volumes, com variação do percentual de páginas e exercícios por ano e um consolidado da coleção. As seções extras tendem a enfatizar leitura, interpretação, construção e adequação de gráficos e os capítulos tendem a concentrar conceitos, medidas de tendência central e planejamento de pesquisas. Além disso, a coleção é considerada apta a apoiar o professor no ensino de Estatística, pois contempla habilidades da BNCC e apresenta atividades distribuídas por níveis de letramento. No entanto, ainda é necessário corrigir equívocos conceituais sobre o que é Estatística; recomenda-se ampliar e diversificar o trabalho com tecnologias, não apenas construir gráficos, mas explorar mais recursos para tratamento de dados, e recomenda-se intensificar propostas de leitura crítica de mídia e debates orientados, para consolidar cidadania estatística.

O trabalho de Gomes e Duarte (2021) analisa como dois livros didáticos de Matemática (3ª série do Ensino Médio) tratam o conteúdo de Estatística, considerando sua relevância para a formação cidadã e para a leitura crítica de informações no cotidiano. O problema que orienta a pesquisa está relacionado à qualidade e adequação do tratamento de Estatística nos livros didáticos. O livro 1 é *Matemática: Ciência e Aplicações* (Iezzi et al., 2016), capítulo 5 (“Estatística Básica”), já o livro 2 é *Matemática – Contexto & Aplicações* (Dante, 2016). Ao analisarem o capítulo de Estatística do livro de Iezzi et al. (2016), os autores concluem que o material apresenta linguagem acessível, boa explicação conceitual e quantidade adequada de exercícios, com propostas de resolução de problemas, contextualização e interdisciplinaridade, inclusive com informática, abordando noções como população, amostra, variáveis e tabelas de frequência. Os autores destacam a diversidade de representações (setores, barras, linhas, pictogramas e histogramas), a clareza visual dos gráficos e a identificação de fontes das ilustrações; nos

exercícios, apontam incentivo ao uso do Excel, estímulo ao raciocínio e articulação com habilidades da BNCC, como EM13MAT101, EM13MAT104 e EM13MAT202.

Em relação ao livro 2 de Dante (2016), Gomes e Duarte (2021) apontam que o capítulo de Estatística se organiza a partir de estruturas matemáticas simples (como sentenças) para sustentar construções e cálculos de frequências, articulando coleta de dados, procedimentos aritméticos e relações algébricas, com uso de múltiplas representações (tabelas, quadros e diferentes gráficos) e leituras de apoio. A contextualização aparece tanto no cotidiano do aluno (por exemplo, atividades investigativas em sala) quanto como contextualização interna à Matemática, em que conteúdos elementares fundamentam tópicos mais complexos; nos exercícios, além da interdisciplinaridade (como com Física), há incentivo ao uso de recursos digitais (*Canva*, *Minitab*, *R*, *LaTeX*, *Sigma Plot*). Os autores também quantificam os tipos de questões do capítulo — 23 de aplicação, 18 algorítmicas, 10 desafios e 7 resolvidas — indicando a combinação de treino procedimental com problemas contextualizados e itens de maior exigência cognitiva, e observam que esse nível de mapeamento não foi realizado para o livro 1, o que constitui uma lacuna do estudo.

Nas considerações finais, Gomes e Duarte (2021) apontam que ambos os livros tentam aproximar a Estatística do cotidiano para facilitar o processo de ensino-aprendizagem. Entretanto, apontam que contextualizações históricas e sociais são pouco exploradas ao longo do conteúdo, o que poderia enriquecer discussões e ampliar sentidos. A pesquisa também reforça que, apesar de o livro didático ser importante, ele não deve ser tratado como único canal de conhecimento, sugerindo a combinação com outros recursos (apresentações, dinâmicas, dispositivos móveis, jogos etc.).

Em síntese, as pesquisas analisadas convergem ao indicar que a qualidade da formação estocástica proposta pelos livros didáticos depende menos da mera presença de tópicos de Estatística e Probabilidade e mais do lugar curricular atribuído ao tema na coleção, do tipo de tarefa predominante e do discurso que sustenta as técnicas mobilizadas. Quando predominam propostas que envolvem problematização, coleta e organização de dados, uso de múltiplas representações, análise e tomada de decisão, amplia-se a possibilidade de avançar para níveis mais robustos de letramento e pensamento estocástico, sobretudo quando o manual do professor e as orientações didáticas oferecem mediações que favorecem a exploração das atividades. Em contrapartida, quando o Tratamento da Informação aparece de forma residual, concentrado em capítulos finais ou reduzido à leitura direta de tabelas e gráficos prontos, tende a se consolidar um ensino mais restrito, com menor exigência investigativa e pouco espaço para variabilidade, inferência e argumentação.

Além disso, os estudos apontam que, mesmo quando o livro apresenta potencial, a mediação docente permanece decisiva para converter o material em experiência efetiva de aprendizagem: ampliar tarefas, corrigir fragilidades conceituais, incorporar tecnologias e criar situações de leitura crítica de dados frequentemente exigem intervenções que extrapolam o que o texto didático explicita. Assim, esta seção reforça que analisar livros didáticos não é apenas diagnosticar materiais, mas explicitar limites e possibilidades do ensino de Estatística na escola,

evidenciando que a construção do letramento estatístico e do pensamento estocástico resulta da articulação entre livro, currículo e prática pedagógica.

5 CONSTRUÇÃO E ANÁLISE DOS GRÁFICOS À LUZ DA ESTATÍSTICA

Este capítulo discute a construção e a análise de gráficos como formas de representação centrais na Estatística Descritiva, com ênfase no papel dessas visualizações para a leitura, a interpretação e a comunicação de informações a partir de dados. Entende-se que os gráficos não constituem apenas um recurso ilustrativo, mas um modo de organizar e comunicar dados, orientando interpretações e inferências iniciais sobre um fenômeno. No entanto, como ressalta o material do Instituto Brasileiro de Geografia e Estatística (IBGE) (2025), a potência visual dos gráficos é acompanhada pelo risco de induzir leituras equivocadas quando as representações são mal construídas ou carecem de rigor estatístico. Nesse sentido, toma-se como ponto de partida a ideia de que o trabalho com gráficos envolve tanto decisões técnicas de representação quanto um exercício de leitura crítica.

Do ponto de vista didático e comunicacional, os gráficos são ferramentas centrais para explorar, comparar e sintetizar informações, permitindo uma visão rápida do comportamento do fenômeno, além de visualizar tendências. Materiais de divulgação científica e educacional enfatizam esse papel funcional dos gráficos e, ao mesmo tempo, destacam a necessidade de elementos mínimos essenciais para sua identificação e integridade visual, como título, eixos quando aplicável, legenda e fonte dos dados (Instituto Brasileiro de Geografia e Estatística (IBGE), 2025). Esses elementos não são acessórios, pois eles asseguram que a informação seja transparente, rastreável e possa ser comparada com outras fontes, permitindo uma análise bem fundamentada.

Além disso, este capítulo assume como referência um procedimento de leitura crítica que orienta a análise dos gráficos apresentados ao longo das seções. Essa abordagem se baseia em um conjunto de verificações que avaliam a clareza do objetivo do gráfico, a transparência e confiabilidade da fonte, a adequação das escalas e a presença de escolhas que possam distorcer a visualização, como quebras de eixo pouco visíveis ou excesso de elementos visuais (Instituto Brasileiro de Geografia e Estatística (IBGE), 2025). Desse modo, não se busca apenas descrever tipos de gráficos, mas explicitar critérios que sustentem uma leitura estatisticamente consistente.

Serão abordadas as definições e conceitos fundamentais da Estatística Descritiva que fundamentam a construção e a análise de gráficos, bem como algumas considerações sobre o seu ensino e aprendizagem. No entanto, considerando que este trabalho se concentra especificamente em tópicos de representação e análise visual, serão apresentadas apenas as definições e conceitos relevantes para os gráficos que são o foco deste trabalho. Para embasar a exposição teórica e as definições apresentadas, utiliza-se como principais referências as obras de Bussab e Morettin (2010), Lima (2009), Silva, Fernandes e Almeida (2015), o material pedagógico do Instituto Brasileiro de Geografia e Estatística (IBGE) (2025), Wild e Seber (2000), Cleveland (1993), Knaflitz (2015), Cairo (2016), Few (2009), United Nations Economic Commission for Europe (2009a) e United Nations Economic Commission for Europe (2009b).

5.1 Conceitos prévios e critérios gerais

5.1.1 Classificação de Variáveis Estatísticas

A escolha e a interpretação de um gráfico estatístico dependem, em primeiro lugar, da natureza da variável analisada. De forma geral, uma variável pode ser entendida como uma característica observável que assume diferentes valores em um conjunto de elementos. Essa caracterização é decisiva porque orienta tanto o modo de organizar os dados quanto o tipo de representação gráfica mais adequado, assegurando que o gráfico realmente mostre aquilo que se pretende comunicar, sem distorções ou incompatibilidades.

Definição 1. Em uma pesquisa estatística, para cada elemento investigado (isto é, para cada unidade da população ou da amostra) associa-se um resultado que corresponde à realização de uma característica observada ou medida. Essa característica, que pode assumir diferentes valores de um elemento para outro, denomina-se **variável**. Em outras palavras, uma variável é a descrição de um atributo do fenômeno sob estudo, cujos valores constituem os dados coletados na pesquisa.

As variáveis podem ser classificadas dos seguintes modos:

- a) **Qualitativas:** são características de uma população cujos valores assumidos correspondem a categorias (atributos), e não a medidas numéricas. Essas variáveis podem ser:
 - Nominais: quando os valores são expressos por atributos e não há ordenação intrínseca entre as categorias. Ex.: sexo, cor da pele, curso de graduação, nacionalidade etc.
 - Ordinais: quando os valores são expressos por categorias que admitem ordenação, isto é, existe uma hierarquia possível, embora não seja apropriado medir numericamente as diferenças entre categorias. Ex.: escolaridade, cargos em uma empresa, patente militar etc.
- b) **Quantitativas:** são características cujos valores são expressos numericamente, representando quantidades obtidas por contagem ou mensuração. Essas variáveis podem ser:
 - Discretas: quando assumem valores pertencentes a um conjunto enumerável, em geral resultantes de uma contagem. Ex.: número de filhos, quantidade de cursos etc.
 - Contínuas: quando seus possíveis valores pertencem a um intervalo de números reais, em geral resultantes de uma mensuração, podendo assumir qualquer valor dentro de certo intervalo de variação. Ex.: peso, altura, idade medida continuamente, tempo etc.

Essa classificação não é apenas uma formalidade teórica. Compreender se uma variável é qualitativa (nominal/ordinal) ou quantitativa (discreta/contínua) define o que pode ser agrupado em categorias, o que deve ser tratado como uma medida numérica e, principalmente, quais tipos de gráficos são apropriados para cada caso. Portanto, as próximas seções usarão essa tipologia para detalhar as condições de uso, os procedimentos de construção e os limites de interpretação de cada representação gráfica.

5.1.2 Elementos mínimos e integridade da representação gráfica

Independentemente do tipo, um gráfico estatístico deve atender a condições mínimas para que sua leitura seja metodologicamente defensável. Em primeiro lugar, são indispensáveis os elementos de identificação e rastreabilidade: título, eixos quando se tratar de representação no plano cartesiano, legenda quando houver mais de uma categoria ou série, e fonte dos dados. A ausência desses componentes compromete a transparência da comunicação, fragiliza a interpretação e impede que o leitor compreenda o que está sendo representado, em quais unidades, e a partir de qual base de dados.

Em segundo lugar, a integridade do gráfico depende da coerência entre a informação numérica e sua codificação visual. Em gráficos cartesianos, isso envolve atenção às escalas, aos intervalos e à proporcionalidade com que variações numéricas são traduzidas em altura, posição ou comprimento. Mudanças discretas de escala, intervalos inconsistentes e cortes que alterem a relação entre valores e sua representação podem produzir efeitos perceptivos que ampliam ou minimizam diferenças, afetando indevidamente a leitura do fenômeno.

Por fim, recursos visuais devem ser considerados à luz de sua justificativa informacional. Em particular, o uso de efeitos tridimensionais em gráficos de barras, colunas e setores tende a introduzir uma dimensão que não corresponde a uma variável do conjunto de dados, alterando a percepção de tamanho, profundidade e proporção, com potencial de distorção interpretativa. Desse modo, tais recursos devem ser evitados quando não houver uma terceira variável efetivamente representada, preservando-se o princípio de que a forma gráfica deve servir ao dado, e não o contrário.

Além de atender a elementos mínimos de identificação e construção (título, fonte, eixos quando aplicável e legenda), a leitura de um gráfico envolve um percurso interpretativo. Para organizar esse percurso, é possível recorrer a questões norteadoras que conduzem o olhar do leitor para o que está sendo representado, para como a representação foi construída e para quais interpretações são sustentadas pelos dados apresentados. Essas questões não operam como instrumento de validação formal do gráfico, mas como orientação para uma leitura mais consciente e menos imediatista.

Quando for analisar um gráfico, procure observar os seguintes pontos: O título ou objetivo do gráfico é claro? Seu conteúdo faz sentido? A fonte dos dados está explícita no gráfico, na figura ou no texto que a acompanha? As informações foram obtidas de uma fonte confiável? Há rótulos adequados para os eixos, etc., deixando tudo bem identificado? Os eixos começam

em zero ou não? As escalas dos eixos são constantes? Existem “quebras” nos eixos que sejam difíceis de notar? Existem elementos gráficos desnecessários ou que dificultam a visualização da informação?

5.1.3 Escolha do gráfico em função do tipo de variável e do objetivo analítico

A adequação de um gráfico resulta da articulação entre 1. o tipo de variável e 2. a pergunta analítica que se pretende responder. Em termos gerais, pode-se afirmar que diferentes gráficos privilegiam diferentes operações cognitivas e estatísticas: comparar categorias, descrever distribuições, evidenciar composição, acompanhar variações no tempo e explorar associação entre duas grandezas. Assim, a escolha do gráfico não deve ser tratada como decisão meramente estética, mas como decisão metodológica.

Com base nessa articulação, adota-se neste capítulo o seguinte mapeamento orientador, que será detalhado e aplicado nas seções 5.3 a 5.7:

1. Comparação entre categorias (variáveis qualitativas ou quantitativas discretas tratadas como categorias): privilegia-se o gráfico de barras, por favorecer contrastes diretos e leitura comparativa.
2. Composição do todo (partes de um total): privilegia-se o gráfico de setores, desde que haja poucas categorias e percentuais claramente definidos, com explicitação da base.
3. Distribuição de variável quantitativa (especialmente contínua): privilegia-se o histograma, com classes definidas de modo criterioso, por permitir análise de concentração, dispersão e forma da distribuição.
4. Variação temporal (uma variável representando tempo): privilegia-se o gráfico de linhas, por evidenciar movimentos, oscilações e tendências ao longo do período.
5. Associação entre duas variáveis quantitativas: privilegia-se o gráfico de dispersão, por possibilitar identificar padrões, agrupamentos e valores atípicos, preservando o caráter exploratório da análise.

Cumprir destacar, ainda, que a escolha adequada do gráfico não elimina a necessidade de prudência interpretativa. Gráficos apoiam inferências descritivas e exploratórias, mas não estabelecem, por si, explicações causais. Desse modo, sempre que a leitura sugerir interpretações explicativas, torna-se necessário explicitar limites do conjunto de dados, condições de coleta e variáveis não observadas, de modo a preservar coerência com o escopo da Estatística Descritiva e com a metodologia adotada no trabalho.

5.2 Gráfico de Barras

O gráfico de barras é, em geral, a representação mais recorrente quando a pergunta analítica envolve comparar categorias. Nessa modalidade, a série de dados é expressa por retângulos de mesma largura, cuja dimensão variável (altura ou comprimento) é proporcional à medida associada a cada categoria, o que torna a leitura comparativa direta e rápida. Em

alguns autores, reserva-se a denominação gráfico de colunas para a disposição vertical e gráfico de barras para a disposição horizontal; aqui, entretanto, adota-se o termo gráfico de barras para ambos os casos, por se entender que se trata do mesmo procedimento de representação, variando apenas a orientação dos retângulos no plano cartesiano.

Do ponto de vista do tipo de variável, esse gráfico é especialmente apropriado para variáveis qualitativas (nominais e ordinais) e pode também ser utilizado para variáveis quantitativas discretas tratadas como categorias, quando a análise depende de contagens, proporções ou comparações por classe. Sua utilidade decorre do fato de permitir que o leitor compare, de modo imediato, cada categoria com sua frequência (absoluta ou relativa) ou com uma medida-resumo, sem exigir transformações adicionais ou leitura indireta de valores. Em termos estatísticos, portanto, o gráfico de barras responde bem a perguntas do tipo: quais são as categorias mais frequentes?, como se distribui um conjunto de ocorrências entre classes? e qual a diferença entre grupos para uma mesma categoria?.

Na construção, recomenda-se explicitar com clareza a estrutura cartesiana: definir os eixos, nomear as categorias, indicar a unidade de medida e adotar uma escala coerente. Na orientação vertical, usualmente posicionam-se as categorias no eixo horizontal e as frequências (ou percentuais) no eixo vertical; na orientação horizontal, essa disposição pode ser invertida para favorecer a legibilidade, sobretudo quando há muitas categorias ou rótulos longos. Em qualquer caso, as barras devem manter largura uniforme e espaçamento constante, e o gráfico deve apresentar título, rótulos de eixos, legenda e fonte dos dados, assegurando rastreabilidade e interpretação adequada.

A escolha entre a orientação horizontal e a vertical não é meramente estética: depende, principalmente, 1. da quantidade de informações a apresentar e 2. das condições materiais de leitura (tamanho do suporte, espaço disponível e legibilidade dos rótulos). Assim, quando os rótulos são extensos ou a lista de categorias é maior, a disposição horizontal tende a reduzir a sobrecarga visual e a melhorar a leitura; quando se busca evidenciar o contraste entre poucas categorias, a disposição vertical costuma ser suficiente. Em todos os casos, é importante evitar efeitos visuais que dificultem a comparação (por exemplo, excesso de cores sem função informacional ou recursos tridimensionais), pois eles reduzem a precisão perceptiva e deslocam a atenção do leitor para elementos decorativos.

Quando a análise envolve apenas uma variável categórica, utiliza-se o gráfico de barras simples, em que cada barra corresponde a uma categoria e seu comprimento representa a medida escolhida (frequência, porcentagem ou total). Esse formato é o mais indicado para descrever a distribuição de ocorrências entre classes, identificar rapidamente categorias dominantes, localizar concentrações e observar “quebras” de padrão. Nessa leitura, recomenda-se manter a interpretação em nível descritivo: o gráfico evidencia diferenças entre categorias, mas não oferece, por si só, explicação para as causas dessas diferenças.

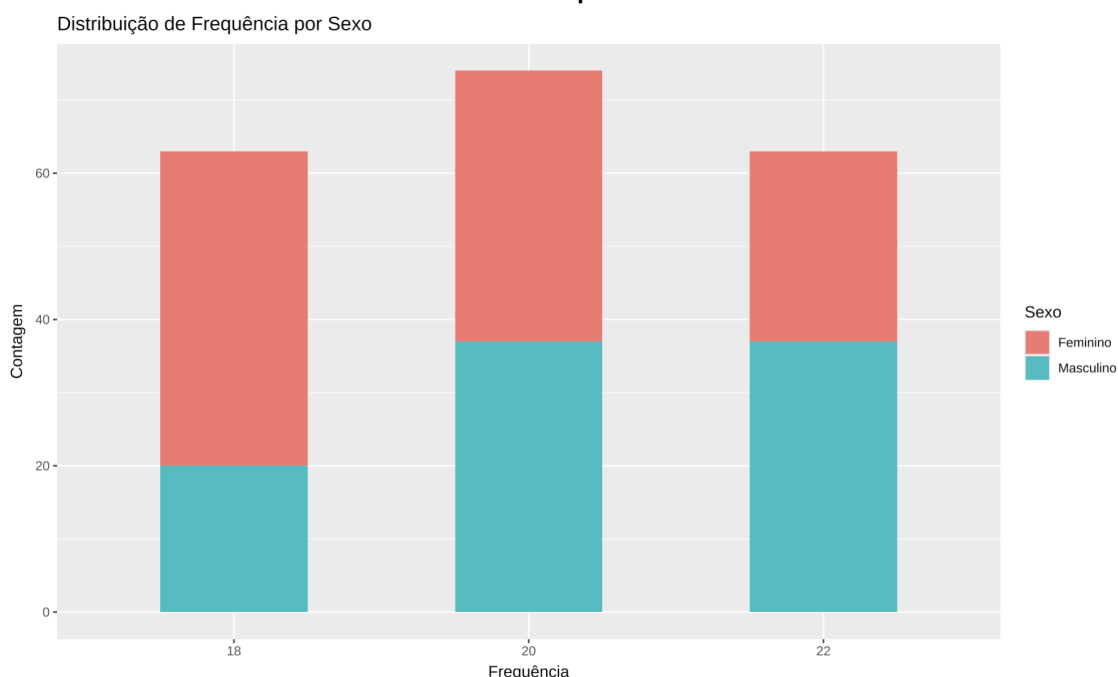
Se, além das categorias, existe uma segunda variável de agrupamento (por exemplo, sexo, turma, turno, rede), o gráfico pode assumir a forma de barras agrupadas lado a lado, permitindo comparar, dentro de cada categoria, como os grupos se distribuem. Nesse caso, a

construção precisa preservar a comparabilidade, ou seja, todas as barras devem compartilhar a mesma escala e a legenda deve ser inequívoca. Também é recomendável limitar o número de grupos para não transformar o gráfico em uma tarefa de busca visual; quando os grupos têm tamanhos muito diferentes, a apresentação em percentuais (em vez de contagens absolutas) tende a ser mais apropriada, desde que o denominador esteja claramente indicado.

Outra variação frequente é o gráfico de barras empilhadas, em que cada barra representa o total de uma categoria e é subdividida em segmentos que expressam componentes desse total. Essa forma é útil quando o objetivo é mostrar simultaneamente volume total e composição, isto é, “quanto há” e “como se reparte”. No entanto, há um limite interpretativo importante: segmentos intermediários não compartilham a mesma linha de base, o que dificulta comparações precisas entre categorias. Por isso, o empilhamento é mais adequado quando a mensagem principal é o total e a composição como panorama geral, e não quando a tarefa central é comparar finamente cada componente entre categorias.

O gráfico a seguir, de autoria própria e integrante do recurso educacional relacionado a esta dissertação, ilustra essa modalidade com a distribuição da frequência de presenças de uma turma de estudantes, segmentada por sexo.

Figura 1 – Distribuição da frequência de presenças de uma turma de estudantes, segmentada por sexo



Fonte: Autoria Própria (2026).

Por fim, o gráfico de barras proporcionais é uma variação do empilhado em que todas as barras são normalizadas para o mesmo total, deslocando o foco para a composição relativa. Trata-se de um recurso adequado quando a pergunta é "como a participação de cada componente muda entre categorias?", independentemente do tamanho absoluto dos grupos. Ainda

assim, é necessário explicitar que a leitura é proporcional e, quando os totais absolutos forem relevantes para a interpretação, apresentá-los como complemento (em anotação, tabela ou outra visualização), evitando que o leitor confunda mudanças de proporção com mudanças de quantidade.

Em síntese, a construção e a análise de gráficos de barras exigem cuidado com a integridade visual e com a coerência entre o tipo de gráfico e a pergunta estatística. O Chapter 1 evidencia boas práticas na construção do gráfico de barras.

Quadro 1 – Boas práticas na construção de gráficos de barras

Boa prática	Descrição
Eixo de valores iniciando em zero	Para evitar distorções na percepção das proporções, o eixo que representa os valores deve iniciar em zero.
Ordenação das barras	Ordenar as barras de forma crescente ou decrescente facilita a identificação dos maiores e menores valores e aprimora a comparação entre categorias.
Rótulos claros e concisos	Eixos, categorias e unidades de medida devem estar explicitados com rótulos legíveis, de modo que o leitor compreenda imediatamente o que está sendo representado.
Uso de cores	As cores devem ser utilizadas com intencionalidade (por exemplo, para destacar informações relevantes ou agrupar categorias), e não apenas por motivos estéticos.

Fonte: Autoria Própria (2026).

Em tempo, o gráfico de barras descreve diferenças e tendências de frequências (ou proporções), mas não explica suas causas, uma vez que é um ferramental descritivo apenas. Mesmo quando uma variação parece expressiva, a explicação do fenômeno não está contida no gráfico, o que demanda o uso da Inferência Estatística.

5.3 Histograma

O histograma é a representação mais adequada quando a pergunta analítica envolve compreender a distribuição de uma variável quantitativa: onde os valores se concentram, como se espalham ao longo de uma escala numérica e qual é a distribuição de probabilidades dos dados. É comum que, à primeira vista, ele seja confundido com um gráfico de barras, mas a diferença é conceitual e afeta diretamente a leitura, no histograma, cada barra representa um intervalo de valores (uma classe) e, por isso, as barras aparecem contíguas, sem espaçamento, sugerindo continuidade. Em termos estatísticos, a finalidade do histograma não é comparar categorias isoladas, mas visualizar como a frequência (ou a proporção) se distribui ao longo de classes numéricas, permitindo reconhecer padrões como pico(s), caudas, assimetrias, lacunas e, eventualmente, múltiplas concentrações.

Do ponto de vista do tipo de variável, o histograma se aplica a variáveis quantitativas, sobretudo quando faz sentido tratá-las como contínuas e agrupá-las em classes. Seu uso, entre-

tanto, exige uma decisão que não é meramente técnica, é preciso definir como o eixo numérico será dividido em intervalos e o que o eixo vertical representará. Em outras palavras, o desenho que o leitor vê não depende apenas dos dados, mas também da forma como esses dados foram organizados em classes. Por isso, ao apresentar histogramas, é importante explicitar a unidade da variável, tornar claros os eixos e assegurar que a construção adotada seja compatível com o objetivo analítico, evitando escolhas que comprimam o corpo principal dos dados ou que criem intervalos arbitrários, difíceis de interpretar. Em termos estatísticos, portanto, o histograma responde bem a perguntas do tipo: em quais faixas de valores os dados se concentram?, a distribuição é mais concentrada ou mais dispersa?, há assimetria (cauda para a direita ou para a esquerda)?, existem lacunas ou mais de um pico (indicando possíveis subgrupos)? e há valores extremos que merecem atenção?.

Na construção, algumas regras simples sustentam integridade e legibilidade. Em primeiro lugar, as classes devem ser definidas de modo que cada observação pertença a apenas um intervalo, sem sobreposições; em segundo, salvo justificativa explícita, é recomendável manter larguras consistentes entre os intervalos, porque classes de tamanhos diferentes alteram o formato percebido da distribuição e podem induzir interpretações equivocadas. Também é boa prática trabalhar com limites “redondos” e coerentes com a escala do fenômeno, pois isso facilita a comunicação: intervalos que fazem sentido para o contexto (por exemplo, notas de 0,5 em 0,5; idades ano a ano) costumam gerar uma leitura mais transparente. Além disso, histogramas exigem atenção aos valores extremos, *outliers*¹ podem ampliar a faixa total e comprimir visualmente a maior parte das observações. Nesses casos, o tratamento deve ser metodologicamente explícito, seja discutindo os extremos como parte do resultado, seja complementando o histograma com anotações e resumos que preservem a leitura do conjunto sem mascarar informação relevante.

Quando a leitura é bem conduzida, o histograma permite interpretar sobretudo forma e estrutura. Ele ajuda a localizar onde estão as maiores concentrações, a perceber se os dados estão mais aglomerados ou mais espalhados e a reconhecer simetria ou assimetria (inclusive a direção da cauda). Também pode indicar a presença de mais de um pico, o que, em muitos contextos, sugere a mistura de subgrupos na mesma amostra e pode orientar investigações posteriores. Ao mesmo tempo, convém delimitar o que o histograma não oferece com precisão: ele é excelente para descrever a distribuição de uma única variável, mas pode não ser a melhor escolha quando se exige comparação fina entre grupos, especialmente se a intenção for identificar diferenças sutis no formato. Nessas situações, o histograma tende a sustentar descrições mais gerais e pode precisar ser combinado com outras representações, dependendo do objetivo da análise.

Uma maneira frequente de construir histogramas é definir o particionamento pelo número de classes, isto é, decidir quantos intervalos serão usados para cobrir toda a faixa de

¹ *Outlier* é uma observação atípica (valor extremo) que se afasta de maneira incomum do conjunto de dados e pode influenciar de forma relevante medidas-resumo e modelos, como correlação e regressão.

valores. Essa escolha regula o equilíbrio entre síntese e detalhe, um número pequeno de classes simplifica a leitura e tende a destacar o formato geral, mas pode ocultar características importantes; um número grande revela variações locais e possíveis lacunas, porém pode produzir um contorno irregular e dificultar a percepção do padrão essencial. Por isso, a escolha do número de classes deve ser guiada pelo propósito: quando se busca uma visão global, tende-se a trabalhar com menos classes; quando se quer explorar nuances, pode-se aumentar esse número, desde que a leitura permaneça inteligível. Como prática exploratória, é recomendável testar mais de uma configuração e comparar as formas que emergem, evitando sustentar interpretações conclusivas a partir de uma única escolha.

Outra possibilidade é definir as classes pela largura dos intervalos, isto é, pelo tamanho de cada classe no eixo numérico. Essa opção costuma ser especialmente útil quando a unidade de medida é substantiva para a interpretação, pois torna o recorte mais alinhado ao contexto: intervalos de 1 em 1, de 2 em 2 ou de 0,5 em 0,5, por exemplo, podem ser mais comunicáveis do que uma divisão abstrata em tantas classes. O ponto crítico é reconhecer que a largura do intervalo não é um detalhe, ao ampliá-la o histograma tende a suavizar o formato, destacando o padrão global; ao reduzi-la, tende a revelar variações locais, mas também pode amplificar ruídos. Em termos práticos, a melhor largura é aquela que permite enxergar o padrão do conjunto sem perder coerência com a escala do fenômeno e sem criar uma leitura visual excessivamente fragmentada.

Por fim, uma decisão que altera o modo de leitura é escolher entre frequência absoluta e densidade no eixo vertical. O histograma em contagem é mais intuitivo e comunica quantos casos caem em cada classe, sendo apropriado quando o tamanho da amostra é parte relevante da mensagem. Já a densidade é preferível quando se pretende comparar distribuições de tamanhos diferentes ou quando se deseja interpretar o histograma em termos de concentração relativa ao longo do eixo, de modo que a área total representada seja padronizada. Nessa modalidade, a comparação do formato tende a ser mais robusta, mas exige cuidado: o eixo vertical precisa indicar claramente que se trata de densidade, para evitar que o leitor interprete valores como contagens. Quando os intervalos têm larguras iguais, contagem e proporção podem produzir gráficos visualmente semelhantes; quando a leitura envolve padronização por largura, a distinção se torna conceitualmente central e deve ser explicitada no gráfico e no texto.

Em síntese, a construção e a análise de histogramas exigem coerência entre o objetivo estatístico e as decisões de agrupamento em classes. Boas práticas incluem definir intervalos consistentes, explicitar eixos e unidades, testar diferentes configurações (número de classes e largura dos intervalos) para avaliar a estabilidade do padrão percebido, tratar valores extremos com critério e escolher a escala vertical (contagem, proporção ou densidade) de acordo com a mensagem analítica.

Do mesmo modo, cumpre reforçar um limite interpretativo: o histograma é excelente para descrever a forma geral da distribuição de uma variável quantitativa, mas pode ocultar estruturas relevantes ou induzir leituras apressadas quando se exige comparação detalhada entre grupos;

Quadro 2 – Boas práticas na construção de histogramas

Boa prática	Descrição
Classes com largura consistente	Manter intervalos de mesma largura evita distorções no formato percebido da distribuição e sustenta comparações visuais coerentes entre classes.
Definição criteriosa de número de classes ou largura das classes	Poucas classes podem ocultar padrões; classes demais podem gerar contorno irregular. Recomenda-se testar configurações e escolher aquela que melhor equilibra síntese e detalhe para o objetivo analítico.
Eixos e unidades explícitos	O eixo horizontal deve indicar a variável quantitativa e sua unidade; o eixo vertical deve deixar claro se representa contagem, proporção ou densidade, evitando ambiguidades interpretativas.
Tratamento transparente de valores extremos	Outliers podem comprimir o corpo principal dos dados e reduzir a resolução do gráfico. Quando necessário, deve-se justificar a estratégia adotada e, se pertinente, complementar com anotações ou resumos.
Escolha adequada entre contagem e densidade	Contagem é útil quando o tamanho amostral é parte da mensagem; densidade é indicada para comparar distribuições com tamanhos diferentes e para leitura em termos de concentração relativa (área total igual a 1).

Fonte: Autoria Própria (2026).

nessas situações, ele funciona melhor como etapa exploratória ou como parte de um conjunto de representações complementares.

5.4 Gráfico de Setores

O gráfico de setores (frequentemente chamado de gráfico de pizza) é uma representação voltada a comunicar relações parte–todo, isto é, mostrar como um total se distribui entre categorias. Sua leitura se apoia na participação relativa de cada “fatia” no conjunto, o que o torna útil quando a intenção principal é evidenciar composição e destacar parcelas claramente maiores ou menores dentro de um total bem definido. Em termos estatísticos, portanto, o gráfico de setores responde bem a perguntas do tipo: como o total se distribui entre categorias?, qual categoria representa a maior parcela do todo? e quais partes são predominantes e quais são residuais na composição?.

Do ponto de vista do tipo de variável, o gráfico de setores se aplica a uma variável categórica (as partes) associada a uma medida que componha um todo (contagem, frequência, proporção ou percentual). Para que faça sentido, é essencial que o todo esteja claramente estabelecido e que as categorias sejam mutuamente exclusivas (cada caso pertence a uma única categoria) e, ao mesmo tempo, exaustivas em relação ao total representado. Além disso, o gráfico funciona melhor com poucas categorias, pois o aumento do número de fatias eleva o esforço de leitura, reduz a distinção visual e enfraquece a mensagem de composição.

Na construção, recomenda-se que o gráfico se explique com o mínimo de fricção. Isso implica explicitar qual é o total e qual medida está sendo representada (percentual, proporção ou frequência relativa), manter rótulos claros e, sempre que possível, privilegiar rotulagem direta (categoria e valor junto à fatia) em vez de depender exclusivamente de legenda distante, que exige um “vai e volta” visual e torna a leitura mais lenta. A organização das fatias também não é neutra, ou seja, ordenar as categorias do maior para o menor tende a reduzir ambiguidade e orientar a atenção do leitor para a estrutura principal da composição. Do mesmo modo, recursos decorativos que criam profundidade, inclinação, sombras ou perspectiva devem ser evitados, pois acrescentam aparência sem acrescentar informação e podem introduzir distorções perceptivas no tamanho aparente das parcelas.

Em termos de escolhas de forma, a questão central não é estética, mas funcional: o gráfico de setores comunica melhor quando preserva simplicidade, hierarquia visual e correspondência imediata entre categoria e parcela. Quando o desenho força o leitor a adivinhar diferenças pequenas entre fatias ou a depender de números para responder perguntas básicas, isso indica que a tarefa analítica talvez esteja melhor atendida por alternativas baseadas em comprimento em escala comum (como barras), que favorecem comparações mais precisas e rápidas.

Uma primeira variação comum diz respeito à apresentação em porcentagens, que costuma ser a opção mais coerente com a lógica parte–todo, pois torna explícita a soma em torno do total (100%) e facilita a leitura direta de participação relativa. Nesse formato, a interpretação esperada é identificar categorias dominantes, reconhecer a presença de parcelas residuais e compreender o desenho geral da composição. Ainda assim, o gráfico não favorece comparações finas entre parcelas próximas, por isso, não é adequado sustentar conclusões do tipo “A é um pouco maior que B” quando as fatias têm tamanhos semelhantes, pois a leitura por ângulo e área tende a ser pouco precisa.

Uma segunda variação é o uso de valores absolutos (por exemplo, contagens) como rótulos. Nesse caso, é importante lembrar que, mesmo quando números absolutos são exibidos, o gráfico continua comunicando proporção, ou seja, a fatia codifica quanto do todo aquela categoria ocupa. Assim, valores absolutos fazem sentido quando ajudam a contextualizar o tamanho do conjunto (por exemplo, evidenciando volume total), mas não transformam o gráfico em uma boa ferramenta para comparar magnitudes com precisão. Se a pergunta principal for “quanto maior é A do que B?”, especialmente quando as diferenças são pequenas, o gráfico de setores tende a ser uma escolha frágil, e a apresentação em barras costuma atender melhor à tarefa comparativa.

Uma terceira variação recorrente é o uso de múltiplos gráficos de setores para comparar composição entre grupos, anos ou contextos. Esse arranjo amplia o custo cognitivo porque o leitor passa a comparar várias “pizzas” sem referência linear comum, o que dificulta perceber mudanças e hierarquias de modo confiável. Em situações de comparação temporal (o que aumentou, o que diminuiu e quanto), esse formato frequentemente empurra o leitor a abandonar a leitura pela forma e recorrer apenas aos números, o que enfraquece o papel do gráfico como

síntese visual. Nesses casos, tende a ser mais consistente adotar representações que facilitem a comparação em escala comum (como barras proporcionais a 100%) ou outras soluções voltadas à leitura de mudança.

Em síntese, o gráfico de setores é mais defensável quando a mensagem é realmente composicional, com poucas categorias e diferenças relativamente claras entre parcelas.

Quadro 3 – Boas práticas na construção de gráficos de setores

Boa prática	Descrição
Total bem definido e categorias exclusivas	O gráfico deve representar partes de um mesmo todo; as categorias precisam ser mutuamente exclusivas e corresponder integralmente ao total analisado.
Poucas categorias e foco na mensagem	Evitar muitas fatias; quando necessário, agrupar parcelas muito pequenas para preservar legibilidade e tornar a composição interpretável.
Rótulos diretos e eixos informativos claros	Preferir rótulos junto às fatias (categoria e valor) a depender apenas de legenda; explicitar se os valores são percentuais, proporções ou frequências relativas.
Ordenação consistente das fatias	Ordenar do maior para o menor reduz ambiguidade e facilita identificar rapidamente categorias dominantes e residuais.
Evitar 3D, perspectiva e efeitos decorativos	Recursos tridimensionais e sombras podem distorcer a percepção de áreas e ângulos, deslocando o foco do dado para o efeito visual.
Preferir alternativas quando a tarefa for comparação precisa	Quando a pergunta exigir ranking claro, comparação fina ou comparação entre grupos e tempos, considerar barras (inclusive 100%) ou outras formas com escala comum.

Fonte: Autoria Própria (2026).

Do mesmo modo, cumpre reforçar que o gráfico de setores descreve composição e participação relativa, mas não sustenta interpretações finas quando as diferenças são pequenas e tampouco é adequado para explicar causas: ele mostra como se reparte, não por que se reparte assim.

5.5 Gráfico de Linhas

O gráfico de linhas é uma representação cartesiana indicada quando a pergunta analítica envolve variação ao longo de uma dimensão ordenada, em especial o tempo. Ao conectar valores sucessivos por segmentos, ele favorece a percepção de trajetória e torna visíveis padrões como tendência, oscilações, mudanças de patamar e momentos de ruptura. Em termos estatísticos, portanto, o gráfico de linhas responde bem a perguntas do tipo: como o fenômeno evolui ao longo do tempo?, há tendência de crescimento, queda ou estabilidade?, existem ciclos ou sazonalidades?, em que momentos ocorrem mudanças mais abruptas? e há períodos que destoam do comportamento esperado?.

Do ponto de vista do tipo de variável, o gráfico de linhas pressupõe uma variável quantitativa no eixo vertical e uma dimensão ordenada no eixo horizontal. Embora o tempo seja o caso mais comum (dias, meses, anos, bimestres, etapas), o eixo horizontal pode representar outras sequências em que a ordem e o espaçamento tenham significado interpretativo. Esse pré-requisito é decisivo: o gráfico de linhas não é adequado para categorias sem ordenação intrínseca, porque a conexão entre pontos sugere continuidade e pode induzir o leitor a perceber um movimento onde ele não existe. Também é importante explicitar a unidade e a periodicidade das observações (por exemplo, semanal, mensal), pois a leitura de oscilações e mudanças depende diretamente do intervalo de coleta e do espaçamento no eixo do tempo.

Na construção, recomenda-se explicitar com clareza a estrutura do gráfico para que o leitor compreenda a série de relance. Isso envolve um título informativo (o que está sendo medido, em que contexto e em qual período), rótulos completos dos eixos (com unidade de medida), e indicação de fonte dos dados, garantindo rastreabilidade. A escala do eixo vertical deve preservar a resolução do padrão sem criar efeitos visuais enganosos: ampliar demais pode exagerar oscilações pequenas, comprimir demais pode esconder mudanças relevantes. Já o eixo do tempo precisa ter marcações suficientes para orientar a leitura, evitando tanto a falta de referência quanto o excesso de ruído. Elementos auxiliares, como linhas de grade, podem ser úteis quando efetivamente ajudam a comparar níveis ao longo do tempo, mas devem ser discretos e subordinados aos dados.

Em termos de decisões de forma, o gráfico de linhas exige cuidado com a geometria do desenho: o formato do gráfico (mais alongado ou mais achatado) altera a percepção de inclinações e, conseqüentemente, a sensação de aceleração, desaceleração ou estabilidade. Por isso, a escolha do tamanho e da proporção do gráfico não é meramente estética, pois ela influencia diretamente o que o leitor percebe como mudança e intensidade de variação. Além disso, quando há comparação entre séries (por exemplo, turmas, grupos, escolas), é necessário controlar a complexidade para que a leitura não se transforme em uma tarefa de busca, nesse caso, a organização e a consistência visual são parte do rigor interpretativo.

No gráfico de linhas simples (uma série), a finalidade é evidenciar o comportamento global da variável ao longo do período, privilegiando a leitura de padrões em vez de valores isolados. A interpretação costuma começar pela tendência (direção geral), passar pelas oscilações de curto prazo e, por fim, identificar irregularidades locais, como picos, quedas abruptas e pontos que destoam do padrão. Um alerta interpretativo importante é evitar explicar visualmente flutuações pequenas quando a variabilidade do processo e a escala do gráfico sugerem incerteza, ou seja, nem toda oscilação é significativa, e a leitura deve considerar magnitude, contexto e estabilidade do comportamento ao longo do tempo. Outro cuidado é não inferir causalidade a partir da forma da curva, pois a linha evidencia variação, mas não fornece, por si só, explicações.

No gráfico de linhas com linhas de referência (por exemplo, média, máximo e mínimo), acrescenta-se um componente de apoio que serve como ancoragem interpretativa. Essas referências funcionam como marcos: a média ajuda a identificar períodos acima e abaixo de um

nível típico, o máximo e o mínimo ajudam a localizar extremos e avaliar amplitude, e outras linhas podem indicar metas, limites ou patamares relevantes para a análise. Para que esse recurso realmente ajude, ele precisa ser aplicado com parcimônia e com hierarquia visual bem definida, isto é, as linhas de referência devem ser discretas, claramente rotuladas (preferencialmente no próprio plano do gráfico) e metodologicamente coerentes com o recorte adotado (por exemplo, média do período inteiro ou de um intervalo específico). O alerta interpretativo aqui é duplo: primeiro, referências mal rotuladas ou competindo visualmente com a série tendem a confundir, segundo, é comum o leitor transformar a referência em regra normativa (acima é bom, abaixo é ruim) sem considerar o contexto do dado, o que exige que o texto acompanhe o gráfico com uma orientação de leitura consistente.

Quando se trabalha com mais de uma série no mesmo gráfico (comparação de grupos), o objetivo passa a incluir comparação de trajetórias: não apenas como evolui, mas como evolui em relação a outras séries. Essa variação pode ser muito informativa, mas também é a que mais rapidamente se torna ilegível quando há excesso de linhas. Por isso, a condição de uso é controlar quantidade de séries e garantir comparabilidade: escalas comuns, rótulos claros e organização visual consistente. Quando a comparação envolve muitas séries, a boa prática é buscar estratégias que preservem a leitura (por exemplo, organizar comparações de forma que o leitor não precise alternar continuamente entre legenda e linhas), evitando sobrecarga visual. O limite interpretativo específico é que, com muitas linhas, o gráfico tende a perder a função de síntese e o leitor passa a depender de leitura numérica pontual, o que enfraquece a vantagem do gráfico de linhas como ferramenta para reconhecer padrões.

Em síntese, o gráfico de linhas é particularmente útil quando se deseja comunicar trajetória e mudança em uma dimensão ordenada, destacando tendências, oscilações e pontos de inflexão. O Quadro 4 apresenta boas práticas na construção desse tipo de gráfico.

Como fechamento, cumpre reforçar um limite interpretativo: o gráfico de linhas descreve mudanças ao longo do tempo e evidencia padrões de comportamento, mas não explica suas causas. Além disso, a percepção de mudança relevante é sensível à escala e ao formato do gráfico, portanto, a interpretação deve considerar magnitude, contexto e consistência do comportamento ao longo do período, evitando tanto a superinterpretação de flutuações pequenas quanto conclusões causais baseadas apenas na forma da curva.

Quadro 4 – Boas práticas na construção de gráficos de linhas

Boa prática	Descrição
Eixo do tempo explícito e coerente	Informar período, unidade e frequência das observações, garantindo que a conexão entre pontos represente uma ordenação interpretável.
Escala do eixo vertical transparente	Definir limites e marcações que preservem a leitura do padrão sem exagerar oscilações pequenas nem esconder variações relevantes.
Título, rótulos e fonte dos dados	Incluir título informativo, eixos com unidades e fonte dos dados, assegurando interpretação adequada e rastreabilidade.
Design sóbrio e legível	Usar grades e marcas de referência com parcimônia, mantendo foco na série e reduzindo ruído visual.
Comparação entre séries com controle de complexidade	Evitar excesso de linhas; quando houver comparação, manter escala comum e rótulos claros para reduzir esforço de leitura.
Linhas de referência com função analítica clara	Incluir média, máximo e mínimo (ou outros patamares) apenas quando ajudam a interpretar; rotular explicitamente e manter hierarquia visual subordinada à série principal.

Fonte: Autoria Própria (2026).

5.6 Gráfico de Dispersão

O gráfico de dispersão é uma representação cartesiana indicada quando a finalidade estatística é examinar a relação entre duas variáveis quantitativas, preservando a variabilidade observada caso a caso. Nele, cada observação aparece como um ponto no plano, definido por dois eixos numéricos, o que permite enxergar tendências, concentrações, lacunas e exceções que dificilmente seriam percebidas apenas por valores-resumo. Em termos estatísticos, portanto, o gráfico de dispersão responde bem a perguntas do tipo: à medida que X aumenta, Y tende a aumentar, diminuir ou não apresentar padrão?, há associação aparente entre as variáveis?, o padrão parece aproximadamente linear ou curvilíneo?, existem agrupamentos (subgrupos) de observações? e há valores atípicos que merecem investigação?.

Quanto ao tipo de variável e aos pré-requisitos, o gráfico de dispersão é apropriado quando se têm duas variáveis numéricas medidas para os mesmos casos, isto é, pares ordenados (X, Y) para cada unidade de observação (por exemplo, estudantes, turmas, escolas, municípios). Para que a leitura seja consistente, é necessário que o recorte seja claro (mesma população, período e contexto de coleta) e que o leitor identifique imediatamente o que está no eixo X e o que está no eixo Y. Também é decisivo explicitar unidades de medida e critérios de inclusão dos dados, evitando misturar grupos incomparáveis sem sinalização, pois isso pode produzir padrões visuais artificiais e interpretações equivocadas.

Na construção, a exigência central é garantir que o gráfico ofereça apoios suficientes para que o padrão possa ser lido com segurança. Isso envolve um título informativo (capaz de situar o que está sendo comparado e em qual contexto), rótulos completos dos eixos (com unidades quando houver) e indicação de fonte dos dados. A escala dos eixos deve ser escolhida de modo a representar as relações de forma transparente, evitando parametrizações que ampliem artificialmente pequenas variações ou comprimam diferenças relevantes. Além disso, quando há muitas observações, pode ocorrer sobreposição de pontos, nesse caso, a apresentação precisa preservar a inteligibilidade do conjunto, porque a utilidade do gráfico depende de permitir ao leitor perceber densidades, agrupamentos e lacunas, e não apenas um amontoado de marcas no plano. Elementos auxiliares, como linhas de grade, devem ser discretos e usados apenas quando efetivamente facilitam a leitura, mantendo os dados como foco principal e evitando ornamentos que introduzam ruído visual.

Em termos de decisões de forma, o gráfico de dispersão exige atenção especial ao equilíbrio entre simplicidade e explicitação. Por um lado, trata-se de um gráfico que pode parecer mais técnico para quem não está habituado, justamente porque não há categorias explícitas, o sentido está na posição relativa dos pontos. Por outro, essa mesma característica é o que o torna potente para a análise, desde que o leitor tenha clareza sobre as variáveis, escalas e recorte. Assim, escolhas de design devem ser guiadas pelo objetivo de tornar o padrão visível com o menor esforço possível, quando houver subgrupos relevantes, por exemplo, é possível distingui-los, mas essa distinção precisa ser feita de forma econômica, para não elevar a carga cognitiva e transformar o gráfico em um quebra-cabeça.

No gráfico de dispersão simples, o objetivo é favorecer a leitura do padrão global sem impor interpretações indevidas. A interpretação esperada começa pela direção do relacionamento (associação positiva, negativa ou ausência de direção), avança para a forma (linearidade, curvatura, mudanças de comportamento em faixas diferentes de X) e inclui a avaliação da força visual do padrão, observando o quão dispersos os pontos estão em torno de uma tendência implícita. Também é parte da leitura identificar possíveis agrupamentos (clusters) e valores atípicos (outliers), que podem sinalizar erro, caso especial ou um fenômeno real que merece explicação. Um alerta específico é não forçar uma mensagem que o gráfico não sustenta: uma nuvem de pontos pode sugerir associação, mas pode também refletir mistura de subgrupos, curvatura não linear ou grande variabilidade, e essas características precisam ser reconhecidas antes de qualquer conclusão mais forte.

Quando se adiciona uma linha de regressão ao gráfico de dispersão, a finalidade passa a incluir uma síntese visual da tendência média do relacionamento entre as variáveis. Isso pode ajudar o leitor a perceber mais rapidamente o sentido e a inclinação do padrão, especialmente quando a nuvem de pontos é grande ou quando a tendência não se destaca à primeira vista. A interpretação, contudo, deve ocorrer em duas camadas: a linha como resumo e os pontos como evidência da variabilidade e das exceções. O alerta interpretativo central é que a linha não substitui os dados, ela pode parecer convincente mesmo quando há dispersão elevada, quando existem subgrupos com comportamentos diferentes ou quando poucos pontos extremos influenciam fortemente o ajuste. Por isso, a leitura correta não é “seguir a linha”, mas comparar a linha com a estrutura da nuvem de pontos, reconhecendo limites do resumo diante da heterogeneidade observada.

No caso do gráfico de dispersão com linha de regressão e faixa ao redor da linha (banda), ilustrado na Figura 2 abaixo, acrescenta-se um componente de apoio voltado a comunicar variabilidade e incerteza em torno do padrão sintetizado. A faixa pode ajudar a evitar uma falsa sensação de precisão, tornando visualmente explícito que a linha é uma aproximação e que há dispersão ao redor da tendência. Para que esse recurso seja válido, porém, ele precisa ser introduzido com parcimônia e com rotulagem clara, de modo que o leitor entenda o que a faixa representa no contexto do gráfico. O alerta interpretativo específico é duplo: primeiro, a banda não deve encobrir a nuvem de pontos nem virar protagonista; segundo, ela não pode ser tratada como prova do relacionamento, mas como um recurso de leitura que deve ser interpretado junto com a distribuição real das observações.

Em síntese, o gráfico de dispersão é particularmente útil quando a pergunta analítica envolve associação, tendência e heterogeneidade entre duas medidas numéricas, pois preserva as observações e permite identificar padrões e exceções com transparência. O Quadro 5 reúne boas práticas para sua construção.

Figura 2 – Gráfico de dispersão com linha de regressão ajustada e banda de confiança ao redor da tendência



Fonte: Autoria Própria (2026).

Quadro 5 – Boas práticas na construção de gráficos de dispersão

Boa prática	Descrição
Eixos claros, com unidades e recorte explícito	Indicar o que está em X e em Y, com unidades quando houver, e situar período, local e população ou amostra analisada.
Escalas coerentes e leitura transparente	Definir limites e escalonamento de modo a evitar compressões ou ampliações artificiais que distorçam a percepção do padrão.
Design sóbrio e foco nos dados	Manter elementos auxiliares discretos (por exemplo, grades leves) e evitar ornamentos que competem com a nuvem de pontos.
Tratar sobreposição e densidade de pontos	Quando houver muitos pontos, ajustar a apresentação para tornar visíveis concentrações e lacunas, preservando a interpretação do conjunto.
Linha de regressão como apoio, não como substituta	Usar a linha para sintetizar tendência quando fizer sentido; interpretar sempre junto com dispersão, clusters e outliers.
Faixa ao redor da linha com significado explícito	Quando utilizada, a banda deve apoiar a leitura de variabilidade e incerteza sem encobrir os dados; seu sentido deve estar claro no gráfico ou em nota.

Fonte: Autoria Própria (2026).

Como fechamento, cumpre reforçar um limite interpretativo fundamental: o gráfico de dispersão evidencia padrões de co-variação, mas não estabelece, por si só, relações de causa e efeito. Mesmo quando há tendência clara (com ou sem linha), é necessário cuidado para não transformar associação visual em explicação causal. Além disso, correlações e ajustes podem ser sensíveis a valores extremos, subgrupos e não linearidades. Assim, o gráfico cumpre sua função de revelar relações, padrões e exceções com clareza, desde que a interpretação permaneça ancorada na estrutura dos pontos e no contexto de produção dos dados, e não apenas em um traço de síntese.

6 A LINGUAGEM R E O USO DE GRÁFICOS NA FORMAÇÃO DE PROFESSORES DE MATEMÁTICA

A presença crescente de dados no cotidiano torna insuficiente um ensino de Estatística restrito à execução de cálculos e à aplicação mecânica de fórmulas. Na Educação Básica, espera-se que os estudantes consigam ler e interpretar informações estatísticas que circulem em diferentes mídias, avaliar argumentos sustentados por gráficos e tabelas e tomar decisões a partir de evidências. Essa perspectiva se aproxima do debate sobre literacia e letramento estatístico tal como vem sendo sistematizado na área, inclusive pela distinção entre “aprender a fazer” e “aprender a interpretar criticamente”, com ênfase na leitura de representações e na avaliação de afirmações baseadas em dados.

No entanto, a literatura e a experiência docente indicam que, com frequência, o ensino de Estatística permanece centrado no cálculo e em exercícios de resposta única, com pouco espaço para investigação, produção de dados, discussão de variabilidade e argumentação. Esse formato tende a deslocar o foco do “o que os dados dizem” para “qual fórmula usar”, empobrecendo a leitura de gráficos e reduzindo a Estatística a um apêndice da Matemática escolar. Em contrapartida, documentos curriculares brasileiros, como a BNCC, ao situar Probabilidade e Estatística como unidade temática ao longo da escolaridade, reforçam a necessidade de práticas que articulem coleta, organização, representação e interpretação de dados; no mesmo sentido, os PCNs, ainda que em outro desenho curricular, já indicavam a centralidade do Tratamento da Informação como componente formativo para compreender e comunicar informações no mundo social.

É nesse ponto que a tecnologia deixa de ser acessório e passa a operar como condição didática para deslocar o ensino do cálculo para a interpretação e a tomada de decisão. Ao automatizar operações repetitivas e permitir explorar rapidamente diferentes representações, escalas e formas de organização dos dados, ferramentas computacionais podem liberar tempo cognitivo e didático para o que é central na Educação Estatística: formular perguntas, testar escolhas de representação, analisar padrões, discutir limites e sustentar conclusões com base em evidências. Nessa direção, Lemes e Santos (2017) defendem o RStudio como um micromundo de Ciência de Dados, no qual estudantes podem explorar dados reais e produzir artefatos públicos (como gráficos, explicações e resultados de investigações), fortalecendo aprendizagens por construção, experimentação e comunicação. Assim, justificar o uso do R não se reduz a afirmar sua potência técnica, mas a defendê-lo como meio para organizar situações de aprendizagem mais investigativas, nas quais a construção e a análise de gráficos se tornem práticas de leitura, argumentação e comunicação.

Diante desse quadro, a discussão que segue não toma o R apenas como ferramenta instrumental, mas como ambiente de trabalho que organiza condições de produção, exploração e comunicação de informações baseadas em dados. Assim, antes de discutir usos pedagógicos, é necessário explicitar o que é o R e o RStudio, quais são seus elementos constitutivos e quais

implicações didáticas decorrem do modo como o ambiente estrutura a interação com dados e visualizações.

6.1 Linguagem R e o RStudio

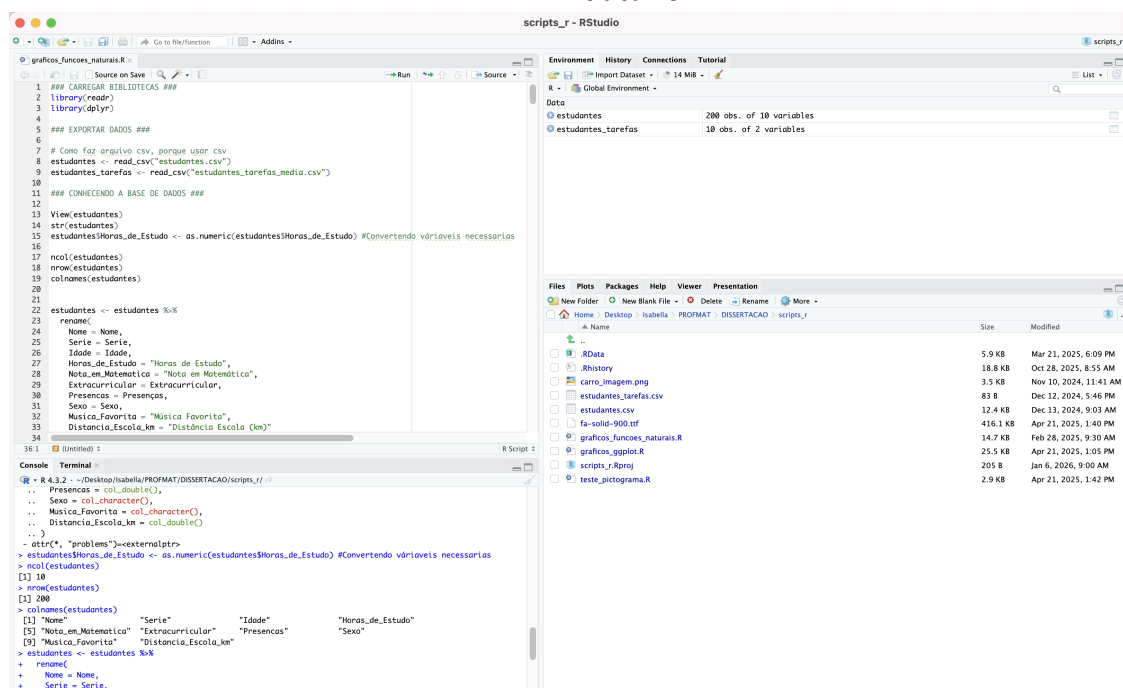
A linguagem R, em si, e seu uso junto com o programa RStudio já vêm sendo alvo de investigação. Diferentemente de softwares comuns para análise de dados, como o Microsoft Excel, que é pago, o software R se encaixa em um cenário de ferramentas computacionais interativas e gratuitas. Ele se destaca como uma linguagem livre, potente e com uma grande comunidade ativa, podendo ser usado também no contexto escolar (Portela; Oliveira; Viola, 2023). O R foi projetado especificamente para a computação estatística e a criação de gráficos, o que o torna particularmente adequado para os objetivos de um ensino de Estatística focado na investigação.

No acesso ao R, um caminho recorrente na literatura é o uso do RStudio, entendido não como a linguagem em si, mas como um Ambiente de Desenvolvimento Integrado (*Integrated Development Environment* – IDE) que organiza, em uma única interface, os recursos necessários para escrever, executar, registrar e revisar análises realizadas em R. Nessa distinção, o R corresponde à linguagem (e ao sistema que interpreta os comandos), enquanto o RStudio opera como a “sala de trabalho” na qual o professor desenvolve suas atividades de forma mais estruturada, com maior controle sobre procedimentos, resultados e versões do que foi feito. Em pesquisas que discutem seu uso educacional, esse ambiente é frequentemente indicado por ser gratuito e por favorecer a articulação entre múltiplas representações, como tabelas e gráficos, contribuindo para a exploração de medidas e para a leitura e discussão de resultados em atividades de Estatística.

Do ponto de vista da interface, o RStudio distribui o trabalho em quatro áreas principais, o que auxilia o usuário a acompanhar, com clareza o que foi escrito, o que foi executado e o que foi produzido ao longo da análise. Como se observa na Figura 3, o editor de scripts (painel superior esquerdo) é o espaço de escrita e salvamento do código, permitindo registrar procedimentos e reutilizá-los em outros contextos; o console (inferior esquerdo) apresenta a execução dos comandos e seus retornos imediatos; o painel de ambiente (*Environment*, superior direito) lista os objetos criados na sessão, como tabelas, vetores, conjuntos de dados e resultados intermediários; e a área que integra arquivos, gráficos e pacotes (inferior direito) organiza a navegação por pastas, a visualização das representações geradas e o gerenciamento de bibliotecas necessárias para diferentes tarefas. Essa configuração favorece a rastreabilidade do trabalho, pois permite ao professor acompanhar, em uma única tela, o código em elaboração, a saída produzida, os dados disponíveis e os gráficos construídos, reforçando a organização e a possibilidade de retomar e justificar cada etapa do processo.

Para iniciar o uso, tanto o R quanto o RStudio são obtidos gratuitamente em seus respectivos *websites* oficiais, e o processo de instalação é direto. É importante ressaltar que o foco deste trabalho não é o passo a passo operacional, mas sim a justificativa pedagógica para a

Figura 3 – Interface do ambiente de desenvolvimento RStudio e seus principais painéis de trabalho



Fonte: Autoria Própria (2026).

adoção da ferramenta. Orientações detalhadas sobre a instalação e o primeiro uso podem ser encontradas no recurso educacional vinculado a essa dissertação.

No ecossistema do R, um elemento fundamental para compreender sua adoção em contextos educacionais é a noção de pacotes (*packages*), também chamados, em uso corrente, de bibliotecas. Pacotes são conjuntos de funções, rotinas e recursos desenvolvidos pela comunidade para ampliar as capacidades da linguagem, permitindo que o usuário utilize ferramentas especializadas sem precisar “programar do zero” cada procedimento. É por meio desses pacotes que o R se consolida como um ambiente flexível e expansível: o professor pode instalar e carregar bibliotecas voltadas à análise estatística, à visualização de dados ou à construção de interfaces, incorporando ao seu trabalho funcionalidades já testadas e documentadas. Assim, quando este capítulo menciona ferramentas como R Commander, Swirl e Shiny, está se referindo a pacotes do R, isto é, extensões que criam diferentes formas de interação com a linguagem e podem reduzir, em graus variados, a barreira inicial de uso.

Uma das principais barreiras para a adoção de linguagens de programação na escola é a curva de aprendizado inicial e, nesse sentido, determinados pacotes operam como pontes didáticas para o início do trabalho com o R. Ciente disso, a comunidade R desenvolveu diversas "camadas de entrada" que funcionam como pontes, permitindo que professores e alunos utilizem o poder do R sem a necessidade de programar desde o início. Uma dessas ferramentas é o R Commander (Rcmdr), um pacote que adiciona uma interface gráfica ao R, baseada em menus e botões. Como destaca Almeida (2014), o R Commander acaba por "mascarar" a linguagem de

programação, facilitando o ambiente para trabalhar com a Estatística e tornando a abordagem mais atrativa e manipulável para iniciantes. Em contrapartida, ao privilegiar a lógica de menus, essa camada pode não mobilizar, de forma aprofundada, a aprendizagem da linguagem em si, funcionando mais como uma porta de entrada para análise e representação de dados do que como um caminho para o domínio de programação.

Outra abordagem para facilitar a entrada no universo R é a aprendizagem interativa dentro do próprio ambiente. O pacote Swirl, por exemplo, transforma o console do R em um ambiente de tutoria, com exercícios e perguntas que guiam o usuário durante o uso do R. No trabalho de Cardoso, Sturion e Sturion (2021), o Swirl é utilizado para que os estudantes aprendam conceitos de Estatística de forma guiada, recebendo feedback imediato sobre seus acertos e erros. Essa metodologia, segundo os autores, favorece tomadas de decisão e reflexão, promovendo uma aprendizagem mais ativa e interativa.

Para um nível ainda maior de interatividade e para a criação de produtos educacionais completos, destaca-se o pacote Shiny. Trata-se de uma biblioteca que permite criar aplicativos web e painéis interativos diretamente a partir do código R, sem exigir programação em outras linguagens como HTML ou JavaScript. Nos trabalhos de Cruz (2024) e Araújo (2020), o Shiny é a ferramenta central para desenvolver, respectivamente, jogos digitais para o ensino de frações e páginas web que apoiam a investigação estatística. Nesses casos, o aluno não interage com o código R, mas sim com a aplicação final (o jogo ou o site), que utiliza o poder de cálculo e visualização do R nos bastidores.

Essas alternativas de acesso (R Commander, Swirl e Shiny) não são excludentes, mas sim complementares, e organizam percursos pedagógicos possíveis. Elas permitem que o professor escolha o nível de contato com a programação mais adequado ao seu público e aos seus objetivos, desde o uso de interfaces prontas até a manipulação direta do código. O recurso educacional desenvolvido neste trabalho assume essa lógica de percursos graduais para justificar o desenho de suas atividades.

Quando a instalação local representa um obstáculo (por infraestrutura, permissões em computadores da escola ou limitações de laboratório), também aparecem alternativas de uso em ambiente on-line. Um exemplo citado na literatura é o Google Colab, caracterizado como um ambiente de notebooks interativo para escrever e executar códigos em contexto on-line, o que pode reduzir barreiras técnicas e facilitar acesso e compartilhamento de materiais. Outras soluções equivalentes, como o Posit Cloud (anteriormente RStudio Cloud), oferecem o ambiente RStudio completo diretamente no navegador. A lógica por trás dessas plataformas é a mesma: remover a barreira técnica da instalação e da manutenção de software, facilitando a reprodutibilidade e o acesso ao material didático.

Estabelecidas as características do ambiente e suas possibilidades de acesso e uso, o passo seguinte consiste em situar como a literatura em Educação Matemática e Educação Estatística tem mobilizado o R em práticas formativas e de sala de aula.

6.2 Linguagem R e Educação Estatística

Bossi e Schimiguel (2023) realizam um estudo do tipo “estado da arte” com o objetivo de mapear como tecnologias digitais vêm sendo mobilizadas em dissertações e teses sobre o ensino de Estatística, identificando tendências e lacunas na produção acadêmica recente. Metodologicamente, os autores consultam a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), delimitam um recorte temporal de dez anos e, após critérios de seleção e leitura analítica, constituem um corpus final de 21 trabalhos, a partir do qual elaboram categorias de tecnologias investigadas.

Nos resultados, Bossi e Schimiguel (2023) sinalizam que, mesmo com o recorte de uma década, a produção ainda é relativamente reduzida em algumas frentes, indicando um campo amplo a explorar e tecnologias pouco investigadas. No caso específico da categoria linguagem R, os autores identificam três pesquisas que tratam explicitamente do uso de R e RStudio, reportando efeitos positivos de recepção e impacto em diferentes níveis, embora também apareçam desafios ligados à formação e, sobretudo, à consolidação da interpretação crítica dos dados. Nesse sentido, o levantamento reforça tanto uma tendência de crescimento do interesse por ferramentas como o R quanto a necessidade de aprofundar a discussão sobre condições de implementação, incluindo a formação e o suporte ao professor, para que a tecnologia deixe de ser um recurso periférico e se torne um meio efetivo de trabalho com dados e representações.

O presente trabalho, portanto, dialoga com essa lacuna ao propor um recurso educacional voltado à formação do professor, não sendo um plano de aula prescritivo, mas um manual técnico-pedagógico que explicita parâmetros, escolhas e procedimentos de construção de gráficos no RStudio, visando fortalecer a autonomia docente para adaptar o uso do R às suas demandas curriculares e às características de suas turmas.

Diante desse panorama, a análise da literatura permite identificar diferentes modos de integração da linguagem R no ensino de Estatística, isto é, diferentes funções que a ferramenta assume nas propostas analisadas, o que orienta a organização das três subseções a seguir. Em um primeiro conjunto de trabalhos, o R aparece principalmente como ferramenta do pesquisador, utilizado para organizar, tratar e apresentar dados, sem necessariamente estruturar a atividade do estudante. Em um segundo grupo, o R é mobilizado como apoio operacional ao ensino, reduzindo barreiras iniciais e liberando tempo didático para interpretação e discussão de tabelas e gráficos. Por fim, há estudos em que o R atua como meio de investigação e produção, sustentando atividades com dados e visualizações e, em alguns casos, a construção de artefatos como aplicações e materiais interativos.

6.2.1 R como ferramenta do pesquisador

Neste nível, o R é utilizado principalmente pelo pesquisador ou professor para a análise de dados coletados em suas investigações, sem que o estudante interaja diretamente com o

código ou com a ferramenta. O uso do R aqui atesta a robustez e a credibilidade da análise estatística subjacente ao estudo.

Partindo da premissa de que a literacia estatística é uma competência cidadã fundamental (Gal, 2002), Souza (2012) buscou avaliar o impacto de uma oficina didático-pedagógica sobre Estatística e Probabilidade em 24 estudantes de licenciatura em Matemática. A intervenção, estruturada em pré-teste, oficina e pós-teste, foi fundamentada em metodologias ativas e na resolução de problemas, visando desenvolver o raciocínio estatístico dos participantes em temas como medidas de tendência central, variabilidade e incerteza.

Para avaliar a eficácia da proposta metodológica, o software R, por meio do pacote R Commander (Rcmdr), foi mobilizado pelo pesquisador para realizar a análise quantitativa dos dados dos questionários (pré e pós-testes), produzindo os gráficos e as estatísticas que fundamentaram a validação dos resultados da oficina. O papel do R, portanto, foi o de uma ferramenta analítica para a avaliação da proposta de ensino, operando nos bastidores da pesquisa e não como um instrumento pedagógico direto utilizado pelos estudantes durante a intervenção. Essa distinção é crucial para compreender o modo de integração do R neste estudo, que o posiciona como um recurso para a validação científica de práticas educacionais.

6.2.2 R como apoio operacional

Este nível abrange os estudos que utilizam interfaces gráficas ou ambientes guiados para reduzir a barreira da programação, permitindo que o aluno se concentre na análise e interpretação dos resultados.

Almeida (2014) elabora uma sequência didática para estudantes do 9º ano do Ensino Fundamental, articulando o ensino de Estatística à Modelagem Matemática. No estudo, o autor destaca que a proposta se apoia no R Commander (Rcmdr) como interface gráfica para operar procedimentos estatísticos. O papel do R, nesse caso, é o de reduzir barreiras operacionais e dinamizar a produção de tabelas, gráficos e medidas, favorecendo que os estudantes concentrem energia nas conclusões e interpretações, ainda que não se trate de uma aprendizagem da linguagem de programação em si.

Nos resultados, o autor afirma que a dinâmica investigativa estimulou maior protagonismo discente e reposicionou o professor como orientador do processo, além de indicar que o uso do R Commander contribuiu para a compreensão estatística: em comparações entre momentos com e sem o software, o trabalho reporta mais dificuldades de significados e representações quando o recurso computacional não foi utilizado, e maior concentração nas conclusões e interpretações quando os estudantes puderam operar as representações com o programa.

Esse estudo se enquadra principalmente como exemplo de R (via R Commander) enquanto interface de apoio ao ensino, reduzindo dificuldades iniciais de uso, simplificando procedimentos técnicos e liberando tempo didático para interpretação. Embora haja menções operacionais ao carregamento do pacote (por exemplo, iniciar o Rcmdr), o foco não está na aprendizagem da linguagem R, mas no uso de uma camada de entrada que facilita a operacionalização

de procedimentos estatísticos em um contexto investigativo. Assim, a contribuição central de Almeida (2014) para o presente trabalho está em sustentar a pertinência de interfaces que diminuam barreiras técnicas e ampliam a possibilidade de trabalhar com representações gráficas e leitura de resultados, sem tornar a programação o primeiro obstáculo.

Cardoso, Sturion e Sturion (2021) investigam o ensino de Estatística mediado pelo RStudio, buscando analisar tanto possíveis contribuições do software para a aprendizagem quanto o grau de aceitação pelos estudantes. A experiência pedagógica que sustenta o estudo envolve 14 estudantes do 8º ano do Ensino Fundamental, em Londrina (PR). Os autores defendem que o uso pedagógico da tecnologia pode favorecer maior foco na interpretação e na compreensão conceitual, reduzindo o peso de procedimentos mecânicos (Chance *et al.*, 2007), ainda que esse potencial dependa de formação docente e de letramento tecnológico do professor.

Na intervenção, os autores descrevem o RStudio como ambiente gratuito e propício ao trabalho com múltiplas representações (tabelas e gráficos) e com medidas como média, mediana e moda, e apresentam o Swirl como um pacote que transforma o console em um espaço interativo, com exercícios e feedback imediato. Metodologicamente, a pesquisa se organiza em três momentos: aplicação de questionário diagnóstico via Google Forms, apresentação orientada do RStudio e execução de uma atividade no Swirl com questões sobre medidas de tendência central, seguida de uma pergunta aberta sobre percepções dos estudantes. Os resultados indicam boa aceitação e maior engajamento com a dinâmica proposta, mas também registram obstáculos práticos, como o idioma do software, a inexistência de versão para celular e a necessidade de digitar comandos para iniciar o curso. No desempenho, os autores relatam acertos imediatos em alguns itens e dificuldades conceituais mais evidentes em mediana (confusões com média), além de erros em média ponderada associados ao tratamento dos pesos, destacando que o feedback do Swirl favoreceu reflexão e tomada de decisão sobre as respostas.

Como síntese, Cardoso, Sturion e Sturion (2021) concluem que o uso de RStudio e Swirl tornou as aulas mais dinâmicas e interativas, com indícios de contribuição para a aprendizagem estatística, ainda que as evidências apresentadas se apoiem sobretudo em registros de tentativas, observações e respostas qualitativas. Este estudo exemplifica o R como camada de entrada que facilita a condução de atividades e pode apoiar a interpretação de resultados, ao mesmo tempo em que reforça a necessidade de investimento em formação de professores para integração consistente de tecnologias no ensino de Matemática.

Matos (2023), em dissertação desenvolvida no PROFMAT, discute a implementação da linguagem R no ensino de Estatística a partir de dados de Saúde Pública, tomando como referência a base do Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel), vinculada ao Ministério da Saúde. O autor sustenta que conteúdos de Estatística Descritiva no Ensino Médio, como medidas de tendência central e dispersão e diferentes representações gráficas, podem ganhar mais significado quando trabalhados com dados autênticos e contextualizados, favorecendo articulações interdisciplinares com áreas como Biologia, Educação Física e Geografia. A escolha por uma base extensa, com dezenas de milhares de registros, é mobilizada para evidenciar que determinados tratamentos e visualizações

se tornam pouco viáveis apenas com procedimentos manuais, reforçando a pertinência do uso de software já na Educação Básica.

Do ponto de vista metodológico, trata-se de um trabalho propositivo, que sistematiza conceitos de Estatística Descritiva e apresenta orientações operacionais para uso do R em atividades com base real, com o objetivo de apoiar o cálculo de medidas e a construção de gráficos, discutindo a adequação de representações aos diferentes tipos de variáveis. Na fundamentação, Matos (2023) defende a integração da tecnologia ao currículo com mediação docente, mobilizando Bonilla (1995) para reforçar que o software não atua de forma autônoma, e articula sua justificativa às competências previstas na BNCC (2017), com ênfase em problemas socialmente relevantes, como os de Saúde Pública.

Nesse sentido, o estudo se insere, nesta seção, como exemplo de R como apoio operacional ao ensino, ao oferecer um conjunto de procedimentos, rotinas e exemplos que facilitam a implementação da ferramenta pelo professor e dão viabilidade ao trabalho com dados e representações em sala de aula. A dissertação detalha etapas de importação e tratamento de dados, instalação de pacotes e geração de representações, além de indicar o Posit Cloud (anteriormente RStudio Cloud) como alternativa para executar R em nuvem e reduzir obstáculos ligados à instalação local, permitindo o uso via navegador em laboratório e, em alguns casos, em dispositivos móveis. Ao explorar o Vigitel, o texto exemplifica rotinas de leitura de arquivo em formato csv, organização de variáveis, produção de tabelas de frequência e construção de gráficos, mencionando pacotes usados para manipulação e importação de dados. Como síntese, o trabalho contribui principalmente como proposta de implementação e material de orientação para o professor, ainda que não apresente uma etapa robusta de avaliação empírica dos efeitos sobre a aprendizagem dos estudantes.

Mascarenhas (2021) tem como objetivo elaborar um material didático auxiliar para o ensino de Estatística com apoio do software R e do RStudio, partindo de dificuldades recorrentes na Educação Estatística. O trabalho aponta que, na prática escolar, o tema muitas vezes aparece de modo incipiente ou restrito a leituras pontuais de gráficos, o que se relaciona à insegurança docente e ao pouco contato com Estatística na formação inicial, favorecendo aulas muito guiadas pelo livro e com menor espaço para análise e discussão de resultados. Nesse contexto, a pesquisa propõe um material de apoio aos livros de Matemática dos anos finais do Ensino Fundamental, oferecendo resoluções de atividades de Estatística com uso do R como alternativa para tornar as aulas mais práticas e apoiar a compreensão conceitual.

O principal resultado é o recurso educacional, um caderno intitulado “Estatística para os anos finais do Ensino Fundamental – Compreensão e Prática”, composto por 50 questões resolvidas com apoio do software R, organizado por ano escolar e conteúdos. O material abrange, por exemplo, tabelas e gráficos no 6º ano, gráficos e medidas de tendência central no 7º ano, distribuições de frequências e diferentes representações no 8º ano e aprofundamentos no 9º ano, incluindo menções a histogramas. O caderno explicita os comandos utilizados em cada questão e inclui uma seção introdutória sobre instalação e uso do RStudio como ambiente de trabalho.

Como síntese, o estudo se destaca por oferecer um material operativo voltado ao professor, com orientações e comandos suficientes para resolver atividades estatísticas e produzir representações, sem pretensão de ensinar programação em profundidade. Ao mesmo tempo, o próprio trabalho reconhece limites, como a opção por evitar comandos mais complexos em algumas construções gráficas e a ausência de validação em turma com evidências empíricas de aprendizagem, apontando como encaminhamentos futuros a aplicação do material em sala e a realização de oficinas de RStudio para docentes. Assim, Mascarenhas (2021) contribui principalmente como proposta de implementação e suporte didático, organizada para facilitar a adoção do R no Ensino Fundamental II.

Santos e Guimarães (2025) investigam a eficácia de integrar a linguagem de programação R ao ensino de Estatística Descritiva no Ensino Médio Técnico, com a intenção de favorecer aprendizagem significativa e desenvolvimento de habilidades analíticas. Para isso, os autores propõem uma intervenção baseada em vídeos curtos publicados no YouTube e articulados ao uso do R no Google Colab, entendido como ambiente on-line de notebooks interativos para escrever e executar códigos.

A metodologia foi aplicada em uma instituição pública municipal da Região Metropolitana de São Paulo, com participação de 319 estudantes do 3º ano de cursos técnicos integrados (Análises Clínicas, Farmácia e Química). A intervenção consistiu na produção de vídeos explicativos, focados no uso do R no Google Colab e em conteúdos introdutórios, com ênfase em média aritmética, além de uma sequência de tópicos como vetores e operações aritméticas. Para avaliar a experiência, os autores aplicaram três questionários com finalidades distintas, incluindo mapeamento de conhecimentos prévios, interesse pelos vídeos e percepção sobre contribuições para a aprendizagem, utilizando esses instrumentos para apreciar a efetividade do formato adotado.

Nos resultados, Santos e Guimarães (2025) observam que, embora a Estatística seja relativamente familiar aos estudantes, a programação em R é pouco conhecida, e a adesão aos vídeos foi limitada, com aproximadamente 30% relatando ter assistido aos conteúdos. Os autores interpretam que essa baixa participação pode estar associada ao lugar do YouTube como espaço de entretenimento e ao desinteresse pelo tema, ainda que, entre os que acessaram os vídeos, predominem avaliações de contribuição para a compreensão, mesmo com dúvidas remanescentes. Como síntese, concluem que a integração do R tem potencial para tornar a Estatística Descritiva mais prática e interessante, mas não opera de modo isolado, exigindo mediação pedagógica e atividades mais estruturadas para engajar os estudantes e transformar curiosidade em participação efetiva.

6.2.3 R como meio de investigação e produção

Neste nível, o R é utilizado de forma mais direta, permitindo que os estudantes manipulem dados reais e participem de todas as etapas de um ciclo investigativo.

Ribeiro (2022), em pesquisa desenvolvida no PROFMAT, investiga o R como ferramenta didático-pedagógica para o ensino de Estatística Descritiva no Ensino Médio. A autora parte de uma dificuldade comum no ensino de estatística, em que quando trabalhada estatística “de forma real”, com bases mais amplas e dados coletados, torna-se difícil apenas com lápis e papel, pois a organização, o tratamento e a representação das informações exigem um volume de procedimentos que tende a deslocar a atenção do estudante para tarefas mecânicas. O autor propõe o uso do R para apoiar análises exploratórias, construção de gráficos e discussões baseadas em dados, com o objetivo de analisar suas contribuições para o desenvolvimento de competências associadas à literacia, ao raciocínio e ao pensamento estatístico, articuladas a uma perspectiva de formação crítica e investigativa.

Metodologicamente, Ribeiro (2022) estrutura uma sequência didática inspirada em Zabala (1998), organizada em três etapas. A primeira compreende aulas voltadas a conceitos e aplicações de Estatística Descritiva por meio de discussões e situações-problema, em contexto de reorganização do ensino durante a pandemia de COVID-19. A segunda etapa introduz o software R e o RStudio, apresentando funções e procedimentos básicos, com a intenção de promover um primeiro contato com a linguagem e estimular elementos de pensamento computacional. Na terceira etapa, os estudantes assumem um papel investigativo: escolhem temas, elaboram questionários, coletam e organizam dados e produzem gráficos no R, culminando na socialização dos resultados, com uso de estratégias como a aula invertida (Valente, 2018) para articular estudo prévio e atividades práticas.

Nos resultados, Ribeiro (2022) aponta indícios de que o uso do R viabilizou a realização de tarefas investigativas com dados em escala mais ampla, favorecendo a passagem de um trabalho centrado em cálculos e confecção manual de gráficos para práticas de organização, representação, interpretação e comunicação de resultados. A autora registra, por um lado, dificuldades iniciais associadas sobretudo à entrada na “cultura da programação”, e não necessariamente ao conteúdo matemático; por outro, relata alta aceitação da proposta, com respostas finais indicando posicionamento favorável dos estudantes à integração de tecnologia e avaliação positiva das contribuições do R para aprender Estatística. Assim, o estudo se enquadra, neste capítulo, como exemplo de R como meio de investigação, em que a ferramenta funciona como mediação para sustentar atividades com dados reais e visualizações, sem se configurar como um curso de programação, mas como suporte para desenvolver competências estatísticas no Ensino Médio.

Araújo (2020), em sua pesquisa realizada pelo PROFMAT, propõe uma sequência de atividades para o ensino de Estatística no Ensino Médio com apoio de tecnologias digitais, articulando a BNCC (2017) ao uso de páginas web construídas em R por meio do pacote Shiny. A autora combina proposta didática, recurso educacional e relato de experiência, com o objetivo de tornar o trabalho com Estatística mais investigativo e menos centrado em fórmulas, incorporando etapas de coleta, organização, análise e comunicação de resultados. Nesse desenho, as páginas web funcionam como suporte para cálculos e visualizações, como medidas resumo

e gráficos, e o material inclui anexos e um apêndice com códigos em R Shiny, o que reforça a reprodutibilidade do artefato.

Metodologicamente, não se trata de um experimento controlado. O texto se organiza no desenvolvimento do recurso educacional e no relato de uma aplicação realizada em 2019 com uma turma de 1ª série de uma escola do norte do Paraná. Na experiência descrita, os estudantes trabalharam com o tema do Índice de Massa Corporal (IMC), definiram o público e o procedimento de coleta e realizaram medições, o que permitiu discutir aspectos como erros de medida e impacto de pequenas variações nos resultados. No tratamento e na apresentação dos dados, surgiram dificuldades operacionais e conceituais, como uso de calculadora, notação decimal e escolhas inadequadas de gráficos para certos tipos de variáveis, situações que foram exploradas didaticamente pela professora para discutir adequação entre variáveis, representações e intenção comunicativa.

Como síntese, o trabalho contribui por explicitar um ciclo de investigação estatística (tema, coleta, tratamento, apresentação e relato) e por oferecer um artefato digital que pode apoiar a condução dessas etapas, sustentando atividades investigativas por meio de aplicações interativas, mais do que como ferramenta operacional direta na mão do estudante. No que se refere ao papel do R, a linguagem aparece principalmente como infraestrutura de produção do material, isto é, a autora utiliza R e Shiny para construir as páginas web e as aplicações, enquanto os estudantes interagem com as representações e resultados, sem programação em R. Assim, Araújo (2020) se destaca por evidenciar como artefatos em Shiny podem apoiar práticas investigativas e ampliar possibilidades de visualização e análise, ao mesmo tempo em que desloca a mediação tecnológica para o professor, responsável por planejar e conduzir a atividade em sala.

Silva (2023), em dissertação do PROFMAT, desenvolve um trabalho robusto sobre o uso da linguagem R na Educação Básica voltado ao ensino e à aprendizagem de Estatística, partindo de uma dificuldade recorrente na área. Muitos estudantes conseguem executar cálculos quando os dados já estão organizados, mas apresentam limitações na leitura, interpretação e atribuição de sentido a tabelas e gráficos, o que é associado a níveis baixos de letramento estatístico e dificulta avançar para formas mais elaboradas de raciocínio e pensamento estatístico. Como resposta, o autor propõe reduzir o peso do cálculo como fim em si mesmo e ampliar a centralidade da interpretação e da argumentação baseada em dados, utilizando a tecnologia para tornar os procedimentos mais transparentes e abrir espaço para análise.

O recurso educacional consiste em um site acessível pela internet, implementado em R com Shiny, voltado a apoiar professores e estudantes por meio de funcionalidades como importação de dados (de arquivos, links e planilhas), geração de medidas e construção de gráficos. Além de apresentar o recurso, Silva (2023) explicita o processo de criação, personalização e publicação do site, incluindo orientações sobre instalação do R e do RStudio, uso de pacotes e disponibilização em servidor, com a intenção de favorecer reuso e adaptação por outros docentes. Na fundamentação, o trabalho mobiliza referências sobre literacia, raciocínio e pensamento

estatístico, e discute a presença de Estatística nos PCN e na BNCC (2017), reforçando a importância de competências estatísticas ao longo da Educação Básica.

Os resultados de Silva (2023) reportam melhora no desempenho após o uso do site e destacam que a ferramenta contribui ao evidenciar o percurso do dado até a informação, embora o autor enfatize que o recurso não substitui a mediação docente. Assim, o trabalho se destaca por apresentar o R como infraestrutura de um artefato interativo que sustenta práticas de leitura e análise estatística, permitindo que estudantes utilizem a ferramenta sem programar, ao mesmo tempo em que oferece caminhos para que professores repliquem e personalizem o aplicativo.

A síntese dos estudos analisados permite afirmar que a linguagem R e o ambiente RStudio vêm sendo mobilizados, na Educação Estatística, segundo ao menos três configurações recorrentes. Em alguns trabalhos, o R aparece sobretudo como ferramenta, empregada para organizar, tratar e representar dados no âmbito da pesquisa, sem constituir o núcleo das atividades de sala de aula, como se observa em Souza (2012). Em outro conjunto, a tecnologia assume função de apoio ao ensino, ao oferecer interfaces, tutoriais e percursos guiados que facilitam o contato inicial e deslocam parte do esforço técnico para a ferramenta, favorecendo tempo e atenção para interpretação de tabelas e gráficos (Almeida, 2014; Cardoso; Sturion; Sturion, 2021). Por fim, há propostas em que o R sustenta práticas investigativas com dados e, em alguns casos, a produção de artefatos digitais, seja como suporte à exploração de bases extensas, seja como infraestrutura para aplicações interativas (Ribeiro, 2022; Silva, 2023; Araújo, 2020).

Essas diferentes formas de integração ajudam a compreender por que, no contexto desta dissertação, faz sentido deslocar o foco para a formação do professor e para a explicitação do processo de construção de gráficos no RStudio, isto é, para as decisões que estão por trás das representações e que condicionam a leitura e a interpretação dos resultados. Com base nesse panorama, a seção seguinte apresenta um conjunto breve de trabalhos que mobilizam o R em outros tópicos da Matemática, como nota de transversalidade, sem alterar o eixo central do capítulo, que permanece ancorado no ensino de Estatística.

6.3 Transversalidade: usos do R em outros tópicos da Matemática

Portela, Oliveira e Viola (2023) investigam o uso do software R como apoio ao ensino e à aprendizagem de Análise Combinatória, partindo de uma dificuldade didática recorrente. Estudantes tendem a memorizar fórmulas de arranjos, permutações e combinações sem compreender critérios de escolha, como o papel da ordem, a existência de repetição e a relação entre número de objetos e de posições, o que também desafia a mediação docente quando a aula se restringe a listas de exercícios e cálculo mecânico. Para enfrentar esse problema, os autores propõem uma ferramenta baseada em uma árvore de decisão, implementada em R por meio da função *acombi()*, que conduz o estudante, por perguntas simples, à identificação do

tipo de situação combinatória e às informações necessárias para estruturar a resolução, sem exigir que a programação seja o foco da aprendizagem.

Os autores em um estudo piloto com estudantes do Ensino Médio, afirmam que a ferramenta é comparada a materiais concretos e à árvore impressa, e os resultados indicam boa aceitação, embora o desempenho em acertos permaneça limitado, sinalizando que a tecnologia pode favorecer engajamento, explicitação de critérios e diagnóstico de dificuldades, mas não substitui a construção conceitual. Para esta seção, o estudo é relevante por evidenciar a transversalidade do R na Matemática e por reforçar um argumento central do capítulo. Tecnologias digitais, quando bem mediadas, podem deslocar o trabalho do uso automático de fórmulas para a explicitação de decisões, critérios e interpretações, o que dialoga com a ideia de tecnologia como meio de investigação e não apenas como ferramenta de cálculo.

Cruz (2024) parte de um problema didático associado ao ensino de frações no Ensino Fundamental. A autora busca compreender, a partir de uma revisão bibliográfica, em que medida jogos podem contribuir para o engajamento e para a aprendizagem, especialmente quando o conteúdo é trabalhado de forma pouco contextualizada ou reduzida a procedimentos repetitivos. Nesse sentido, o trabalho enfatiza a importância de atividades mais dinâmicas e alinhadas às habilidades previstas na BNCC (2017), defendendo que recursos lúdicos podem favorecer participação, atenção e construção de significados no estudo de frações.

O R entra no estudo como meio de viabilização do recurso educacional, pois a autora implementa dois jogos por meio do pacote Shiny, que permite construir aplicações interativas a partir de código em R e disponibilizá-las como interface web. A pesquisa exemplifica um uso em que a linguagem opera predominantemente como infraestrutura técnica de produção e organização do material didático, sem se tornar o foco de aprendizagem do estudante.

Julio e Ferreira (2021) descrevem uma experiência formativa no ensino superior, na disciplina optativa “Construindo Fractais”, tomando como problema didático a baixa presença da geometria fractal na formação de professores, apesar da expansão de pesquisas na área. O objetivo central é apresentar a dinâmica da disciplina e, em especial, discutir a primeira aula, em que a construção de um “cartão fractal” é usada como disparador para conectar diferentes tópicos matemáticos, como sequências, recorrências, algoritmos e transformações, além de favorecer discussões didático-formativas para futura prática docente.

Nesse contexto, o R entra como recurso de exploração e experimentação, compondo o conjunto de ferramentas computacionais mobilizadas ao lado de materiais manipuláveis e de recursos como planilhas. A linguagem é utilizada para gerar imagens digitais e investigar padrões por meio de procedimentos iterativos, articulando iteração, algoritmos, recorrências e matrizes, o que reconfigura o modo de abordar o objeto matemático e amplia possibilidades de visualização e análise. O trabalho evidencia um uso do R que vai além do suporte operacional, pois a tecnologia atua como meio produtivo de investigação, contribuindo para a produção de significados em sala e para deslocar a aprendizagem de um repertório de fórmulas para a explicitação de processos, justificativas e relações matemáticas.

Em síntese, ao aparecer tanto em práticas voltadas à Estatística quanto em outros objetos matemáticos, o R se consolida como recurso formativo que pode sustentar diferentes modalidades de trabalho matemático, especialmente quando orientado por problemas, critérios e interpretação. Com base nessas evidências, apresenta-se, na próxima seção, o recurso educacional desenvolvido nesta pesquisa, concebido para apoiar a apropriação docente do R e do RStudio em atividades de leitura, análise e comunicação de dados.

6.4 Recurso Educacional: o manual e como utilizá-lo

Com base nas discussões desenvolvidas neste capítulo, a tecnologia é compreendida como uma condição didática para deslocar o ensino de Estatística de abordagens centradas em cálculo e procedimentos para práticas que privilegiem leitura, interpretação, argumentação e tomada de decisão a partir de evidências. Ao mesmo tempo, a literatura revisada aponta que a integração de ferramentas como o R depende de condições de implementação que envolvem, de modo decisivo, a formação e o suporte ao professor. É nesse enquadramento que se situa o recurso educacional desta pesquisa, concebido como uma resposta prática às necessidades identificadas, ao oferecer um percurso de apropriação do R e do RStudio orientado à construção de visualizações que sustentem investigações e discussões com dados no contexto escolar.

O recurso educacional vinculado a esta dissertação é o manual digital, elaborado para professores da Educação Básica interessados em utilizar a linguagem R como recurso para trabalhar Estatística e visualização de dados. O objetivo do material não é apresentar um receituário fechado de comandos, nem prescrever uma sequência única de aula, mas favorecer a autonomia docente no uso do RStudio para construir e ajustar gráficos, compreendendo parâmetros, escolhas e implicações de cada configuração. Assim, o professor é convidado a tomar o R como meio de trabalho e de mediação didática, adaptando scripts e exemplos às demandas curriculares e às características de suas turmas.

Em termos de estrutura, o manual organiza-se como um percurso que combina fundamentos operacionais e encaminhamentos didático-técnicos. Inicialmente, apresenta orientações para instalação e configuração do ambiente, contemplando os passos necessários para obter e iniciar o R e o RStudio. Em seguida, discute estratégias para a criação de conjuntos de dados que possam ser utilizados em atividades, incluindo possibilidades de dados fictícios, dados reais e alternativas de produção e organização em diferentes contextos. Na parte central do material, são desenvolvidos tutoriais de construção dos principais tipos de gráficos trabalhados no produto, sendo gráficos de barras, histogramas, gráfico de setores, de linhas e dispersão, com atenção à lógica de escolha da representação e às possibilidades de personalização. Nessa seção, o manual também enfatiza a comparação entre a sintaxe nativa do R e a sintaxe baseada no pacote ggplot2, de modo a explicitar vantagens, limites e critérios de uso conforme o objetivo da visualização e o nível de familiaridade do professor com o ambiente.

Para reduzir a barreira inicial e favorecer a adoção da ferramenta, o manual sugere um percurso de entrada em três movimentos. O primeiro consiste em instalar o R e o RStudio e

reconhecer a interface do ambiente, compreendendo onde o código é escrito e salvo, onde os comandos são executados e onde os resultados são visualizados.

O segundo movimento é selecionar ou produzir um conjunto de dados para trabalho. Para isso, o material disponibiliza exemplos e propostas de bases de dados para coleta, permitindo que o professor parta de um ponto concreto sem precisar construir tudo do zero. Além disso, o professor tem autonomia para personalizar os temas escolhidos, aproximando os dados da realidade e dos interesses da própria turma. Nesse processo, a inteligência artificial pode ser uma aliada relevante: o material oferece prompts de comandos prontos para uso, um recurso especialmente valioso dado que grande parte dos professores ainda não possui familiaridade com essas ferramentas. Com apoio desses prompts, é possível gerar contextos, adaptar exemplos e criar situações-problema personalizadas de forma acessível.

Por fim, o terceiro movimento é construir o primeiro gráfico, executando e adaptando scripts, de modo a tornar explícitas as escolhas realizadas (títulos, eixos, escalas, categorias, intervalos) e sua relação com a intenção comunicativa e interpretativa da atividade.

O recurso educacional está disponibilizado em repositório público no GitHub, o que favorece o acesso, a atualização e a reutilização dos materiais, garantindo que os arquivos possam ser consultados e utilizados por outros docentes em diferentes contextos. O endereço do repositório é <https://github.com/isasterza/manual-graficos-rstudio>.¹ Com isso, encerra-se a apresentação do recurso e, no capítulo seguinte, aprofunda-se a discussão sobre os gráficos trabalhados, explicitando suas finalidades estatísticas, critérios de construção e elementos de interpretação, de modo a sustentar o uso do R como recurso de leitura, análise e comunicação de informações baseadas em dados.

¹ <https://github.com/isasterza/manual-graficos-rstudio>.

7 CONSIDERAÇÕES FINAIS

O presente trabalho buscou mapear e analisar a produção acadêmica sobre o uso da linguagem R no ensino de Estatística na Educação Básica, com foco na formação de professores de Matemática, e, a partir dessa análise, propor um recurso educacional que auxiliasse docentes a utilizar o R e o ambiente RStudio de forma autônoma na construção e na análise de gráficos estatísticos.

Em um cenário social cada vez mais orientado por dados, impulsionado pela Revolução Digital e pela Inteligência Artificial, a Estatística assume importância crescente como linguagem de leitura e interpretação do mundo. Nessa perspectiva, desenvolver raciocínio, letramento e pensamento estatístico ultrapassa o mero domínio do cálculo, envolvendo capacidades de formular perguntas, organizar e representar informações, interpretar resultados, sustentar argumentos e tomar decisões com base em evidências. No entanto, a análise discutida ao longo do trabalho evidencia um descompasso entre essa demanda formativa e as condições de preparação docente no Brasil, embora a BNCC estabeleça a Estatística como componente obrigatório, a formação inicial e os materiais didáticos frequentemente privilegiam procedimentos e respostas rápidas, com menor ênfase na interpretação crítica, na contextualização e na explicação das escolhas envolvidas na produção e na representação dos dados.

Os resultados do mapeamento mostram que uma parte expressiva das pesquisas e propostas formativas relacionadas ao uso de softwares no ensino de Estatística tende a enfatizar a prescrição de métodos e a replicação de procedimentos. Ainda que tais abordagens tenham relevância, observou-se que, em muitos casos, elas não priorizam suficientemente a apropriação conceitual e a autonomia do professor para adaptar, criar e justificar suas próprias propostas de ensino. Em termos didáticos, isso se expressa em usos da tecnologia que operam como execução guiada de passos, sem tornar explícitos os critérios de escolha do tipo de gráfico, os parâmetros de construção e os elementos fundamentais para interpretação e argumentação.

Esta dissertação sustenta que a apropriação autônoma é não apenas possível, mas necessária para qualificar o ensino. Isso implica compreender a tecnologia não como um recurso que automatiza a aula, mas como um meio de explicitar escolhas e fortalecer práticas investigativas: ao organizar análises em scripts, o professor torna visível o caminho entre dado, tratamento e comunicação de resultados, ampliando a rastreabilidade dos procedimentos, a replicabilidade das produções e a intencionalidade didática das decisões de visualização. Nesse contexto, o R é mobilizado como linguagem de programação voltada à análise, e o RStudio como Ambiente de Desenvolvimento Integrado (Integrated Development Environment – IDE), favorecendo a organização do trabalho e a transparência do processo analítico.

Como contribuição prática, foi proposto um manual digital voltado a professores da Educação Básica, estruturado de forma progressiva e orientado ao desenvolvimento de autonomia: instalação e primeiros passos no R e no RStudio; construção e importação de bases de dados; produção e análise de visualizações por funções nativas e pelo pacote ggplot2; e uma seção de replicabilidade, com orientações para adaptação de códigos a diferentes contextos escolares.

Ao explicitar parâmetros, escolhas e implicações na construção e no ajuste dos gráficos, o manual busca deslocar o uso do software da mera repetição para uma prática em que o professor faça escolhas informadas e coerentes com a pergunta analítica, o tipo de variável e os objetivos de aprendizagem, fortalecendo a visualização como componente do raciocínio estatístico e não como etapa final de ilustração de resultados.

Ao enfatizar a autonomia, a justificativa das escolhas e a adaptação contextualizada, esta pesquisa defende que empoderar o professor no uso de ferramentas como R e RStudio é condição para que o ensino de Estatística se torne mais investigativo, crítico e relevante para os estudantes, especialmente em uma sociedade marcada pela circulação intensiva de dados e por disputas de interpretação.

REFERÊNCIAS

- ALBUQUERQUE, L. C. d.; GONTIJO, C. H. A complexidade da formação do professor de matemática e suas implicações para a prática docente. **Espaço Pedagógico**, Passo Fundo v. 20, n. 1, p. 76–87, 2013. Disponível em: https://repositorio.unb.br/bitstream/10482/14022/1/ARTIGO_ComplexidadeForma%C3%A7%C3%A3oProfessor.pdf.
- ALMEIDA, L. H. P. d. **Proposta de ensino de Estatística em uma turma de nono ano do Ensino Fundamental com uso do programa R-Commander**. 2014. Dissertação (Dissertação (Mestrado Profissional em Ensino de Matemática)) — Universidade Federal do Rio Grande do Sul (UFRGS) Porto Alegre, RS 2014.
- ARAÚJO, C. D. d. **Estatística no Ensino Médio**: uma proposta de atividades com o uso de tecnologias. 2020. Dissertação (Dissertação (Mestrado Profissional em Matemática)) — Universidade Tecnológica Federal do Paraná (UTFPR) Cornélio Procopio, PR 2020.
- BARBOSA, R. P.; ALVES, N. Reforma do ensino médio e a plataformização da educação. **Revista E-Curriculum**, Pontifical Catholic University of São Paulo (PUC-SP) v. 21,, p. 1–21, set. 2023. Disponível em: <https://revistas.pucsp.br/index.php/curriculum/article/view/61619>.
- BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2016. 3. reimp. da 1. edição de 2016. Título original: L'analyse de contenu. ISBN 978-85-62938-04-7.
- BOSSI, K.; SCHIMIGUEL, J. Utilização das tecnologias digitais no ensino da estatística: uma visão baseada em análise de dissertações e teses. **Revista Sociedade Científica**, v. 6, n. 1, p. 3358–3379, 2023.
- BRASIL. **Parâmetros Curriculares Nacionais**: Matemática. Brasília, 1998. Disponível em: <http://200.144.244.96/cda/PARAMETROS-CURRICULARES/PCN-5a-a-8a/matematica.pdf>.
- Brasil. Ministério da Educação (MEC). **Base Nacional Comum Curricular**: educação é a base. Brasília, DF, 2018. Disponível em: http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518-versaofinal_site.pdf.
- BUSSAB, W. d. O.; MORETTIN, P. A. **Estatística Básica**. 6. ed. São Paulo: Editora Saraiva, 2010. ISBN 978-85-02-08177-2.
- CAIRO, A. **The Truthful Art: Data, Charts, and Maps for Communication**. 1. ed. Berkeley, CA: New Riders, 2016. (Voices That Matter). ISBN 978-0-13-344049-2.
- Campanha Nacional pelo Direito à Educação. **SEDUC-SP gasta quase R\$ 500 milhões em plataformas sem melhorar resultados educacionais, apontam pesquisadores**, . 2025. Notícia. Disponível em: <https://campanha.org.br/noticias/2025/07/03/seduc-sp-gasta-quase-r-500-milhoes-em-plataformas-sem-melhorar-resultados-educacionais-apontam-pe>
- CAMPOS, C. R. **A Educação Estatística**: uma investigação acerca dos aspectos relevantes à didática da estatística em cursos de graduação. 2007. Tese (Tese (Doutorado em Educação Matemática)) — Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista Rio Claro, SP 2007.
- CARDOSO, M. d. G.; STURION, L.; STURION, L. H. C. O ensino de estatística por meio das potencialidades do software rstudio. *In*: **Educação matemática em pesquisa**: perspectivas e tendências. volume 2. [S.l.: s.n.] 2021. p. 60–73.
- CLEVELAND, W. S. **Visualizing Data**. Summit, NJ: Hobart Press, 1993.

- CONTI, T. **A estatística nos livros didáticos do ensino médio: uma abordagem histórico-cultural**. 2013. Tese (Monografia (Especialização em Educação Matemática)) — Universidade do Extremo Sul Catarinense (UNESC) Criciúma, SC 2013.
- CORRÊA, M. V. B. **Ensino de estatística: concepções e análises**. 2015. Tese (Trabalho de Conclusão de Curso (Licenciatura em Matemática)) — Universidade Federal do Rio Grande do Sul (UFRGS), Instituto de Matemática e Estatística, Departamento de Matemática Pura e Aplicada Porto Alegre, RS 2015.
- COULDRY, N.; MEJIAS, U. A. **The Costs of Connection: How data is colonizing human life and appropriating it for capitalism**. Stanford: Stanford University Press, 2019.
- CRUZ, I. S. **Utilização dos jogos como recursos didáticos no ensino de frações matemáticas**. 2024. Dissertação (Dissertação (Mestrado Profissional em Matemática em Rede Nacional – PROFMAT)) — Universidade Federal de São João del-Rei (UFSJ), Campus Alto Paraopeba Ouro Branco, MG 2024.
- D'AMBROSIO, B. S.; D'AMBROSIO, U. Formação de professores de matemática: professor-pesquisador. **Atos de Pesquisa em Educação**, v. 1, n. 1, p. 75–85, 2007. Disponível em: <https://ojsrevista.furb.br/ojs/index.php/atosdepesquisa/article/view/65>.
- FAUSTINO, D.; LIPPOLD, W. **Colonialismo digital: por uma crítica hacker-fanoniana**. 1. ed. São Paulo: Boitempo, 2023.
- FEW, S. **Now You See It: Simple Visualization Techniques for Quantitative Analysis**. Oakland, CA: Analytics Press, 2009. ISBN 9780970601988.
- FILHO, L. A. **Data centers no Nordeste, minérios na África, lucros no Vale do Silício**, . 2025. Disponível em: <https://diplomatie.org.br/data-centers-no-nordeste-minerios-na-africa-lucros-no-vale-do-silicio/>.
- FIORENTINI, D. A pesquisa e as práticas de formação de professores de matemática em face das políticas públicas no Brasil. **Bolema: Boletim de Educação Matemática**, Rio Claro, SP v. 21, n. 29, p. 43–70, 2008.
- FRIOLANI, L. C. **O pensamento estocástico nos livros didáticos do ensino fundamental**. 2007. Dissertação (Dissertação (Mestrado Profissional em Ensino de Matemática)) — Pontifícia Universidade Católica de São Paulo (PUC-SP) São Paulo, SP 2007.
- GIORDANO, C. C.; ARAÚJO, J. R. A.; COUTINHO, C. d. Q. e. S. Educação estatística e a base nacional comum curricular: o incentivo aos projetos. **REVEMAT: Revista Eletrônica de Educação Matemática**, Florianópolis, SC v. 14,, p. 1–20, 2019. Edição Especial Educação Estatística.
- GIORDANO, C. C.; LIMA, R. F. O ensino de representações estatísticas no Brasil, frente às demandas pós-bncc. **Revista Paideia**, v. 15, n. 27, p. 4–21, abr. 2023. UNIMES Virtual.
- GIRALDO, V. Formação de professores de matemática: para uma abordagem problematizada. **Ciência e Cultura**, São Paulo v. 70, n. 1, p. 37–42, mar. 2018. Edição jan./mar.
- GOMES, R. S. d. S.; DUARTE, G. d. O. Estatística para o terceiro ano do ensino médio: análise de livros didáticos de matemática. In: **Teorias da educação matemática na prática de sala de aula: anais do 11. seminário de cognição e educação matemática (scem)**. Belém, PA: Universidade do Estado do Pará (UEPA) 2021. p. 84–96.
- Instituto Brasileiro de Geografia e Estatística (IBGE). **Principais tipos de gráficos para a educação básica: recursos para o trabalho em sala de aula**. Rio de Janeiro, 2025. Disponível

em: <https://educa.ibge.gov.br/professores/educa-recursos/20773-tipos-de-graficos-no-ensino.html>.

JULIO, R. S.; FERREIRA, J. C. Construindo seu fractal: uma experiência no ensino superior. *In: Educação matemática em pesquisa: perspectivas e tendências*. volume 2. [s.n.] 2021. p. 438–453. Disponível em: <https://downloads.editoracientifica.org/articles/201202526.pdf>.

KNAFLIC, C. N. **Storytelling with Data: A Data Visualization Guide for Business Professionals**. Hoboken, NJ: John Wiley & Sons, 2015.

KOHL-SANTOS, P.; MOROSINI, M. C. O revisitar da metodologia do estado do conhecimento para além de uma revisão bibliográfica. **Revista Panorâmica Online**, v. 33,, 2021. Disponível em: <https://periodicoscientificos.ufmt.br/revistapanoramica/index.php/revistapanoramica/article/view/1318>.

LEMES, I. L.; SANTOS, R. P. Aplicações de rstudio na aprendizagem matemática: elaboração de artefatos e ciência de dados como matética. **Educação Matemática em Revista**, [S.l.] v. 2, n. 18, p. 87–97, 2017.

LIMA, L. d. **Estatística Aplicada: Tecnologia em hotelaria**. Fortaleza: UAB/IFCE, 2009. ISBN 978-85-475-0017-7.

LIMA, S. d. O. *et al.* Ensino de estatística, probabilidade e combinatória na educação básica: os novos desafios da bncc. **Revista Baiana de Educação Matemática**, Salvador v. 3, n. 1, p. 1–20, 2022.

LOPES, C. E. Educação estatística no curso de licenciatura em matemática. **Bolema**, v. 27, n. 47, p. 901–915, dez. 2013.

LOPES, C. E. O ensino da estatística e da probabilidade na educação básica e a formação dos professores. **Cadernos CEDES**, v. 28, n. 74, p. 57–73, 2008.

MALHEIROS, A. P. d. S. Contribuições de paulo freire para uma compreensão do trabalho com a modelagem na formação inicial de professores de matemática. **Boletim GEPEM**, , n. 64, p. 46–56, 2014. Disponível em: <https://periodicos.ufrj.br/index.php/gepem/article/view/12>.

MASCARENHAS, L. R. **Proposta de um material didático auxiliar para o ensino de Estatística por meio do software R**. 2021. Dissertação (Dissertação (Mestrado Profissional em Matemática em Rede Nacional – PROFMAT)) — Universidade Federal do Vale do São Francisco (UNIVASF), Campus Juazeiro Juazeiro, BA 2021.

MATOS, M. A. d. S. A. **Implementação da linguagem R no ensino de Estatística: utilizando dados de saúde pública**. 2023. Dissertação (Dissertação (Mestrado Profissional em Matemática em Rede Nacional)) — Universidade Federal do Maranhão (UFMA) São Luís, MA 2023.

MAYNARD, D. C. S. Passado eletrônico: notas sobre história digital. **Acervo**, Rio de Janeiro v. 29, n. 2, p. 103–116, dez. 2016. Edição semestral jun./dez.

NETO, F. d. S. **Análise do letramento estatístico nos livros didáticos do Ensino Médio**. 2008. Dissertação (Dissertação (Mestrado Profissional em Ensino de Matemática)) — Pontifícia Universidade Católica de São Paulo (PUC-SP) São Paulo, SP 2008.

PALHETA, H. B. B.; PORCIÚNCULA, M. M. d. S.; GIORDANO, C. C. A educação estatística na bncc e dcepa no ensino médio: Possibilidades de ensino envolvendo projeto de aprendizagem estatístico (pae). **Revista de Investigação e Divulgação em Educação Matemática**, Juiz de Fora, MG v. 9, n. 1, p. 1–18, 2025.

PICKLER, A. P. **Sociedade olha para o professor com pena e não com admiração profissional**, . 2024. Notícia. Disponível em: <https://www.gazetadopovo.com.br/parana/claudia-costin-roni-miranda-debatem-desafios-educacao/>.

PONTES, M. M. d. *et al.* A temática “probabilidade e estatística” nos anos iniciais do ensino fundamental a partir da promulgação da bncc: percepções pedagógicas. **Educitec – Revista de Estudos e Pesquisas sobre Ensino Tecnológico**, Manaus, AM v. 5, n. 12, p. 221–237, dez. 2019. Disponível em: <https://sistemascmc.ifam.edu.br/educitec/index.php/educitec/article/view/957>.

PORTELA, A. C. T.; OLIVEIRA, H. H. G. d. S.; VIOLA, D. N. Utilização do software r para auxiliar no processo de ensino-aprendizagem da análise combinatória. **Ensino em Re-Vista**, Uberlândia v. 30,, p. 1–25, 2023.

REIS, D. M. Educação plataformizada: novos contornos da política educacional do estado de são paulo. **Revista da FAEEBA – Educação e Contemporaneidade**, Salvador v. 34, n. 78, p. 191–209, 2025. Disponível em: <https://www.revistas.uneb.br/faeeba/article/view/22785>.

RIBEIRO, M. A. **Contribuição ao estudo do Software R como ferramenta didático-pedagógica para o desenvolvimento de Estatística Descritiva no Ensino Médio**. 2022. Dissertação (Dissertação (Mestrado Profissional em Ensino de Ciências e Matemática)) — Universidade Federal de Sergipe (UFS) São Cristóvão, SE 2022.

RODRIGUES, M. U.; SILVA, L. D. d. Disciplina de estatística na matriz curricular dos cursos de licenciatura em matemática no brasil. **REVEMAT**, v. 14,, p. 1–21, 2019. Edição Especial Educação Estatística.

SAMÁ, S.; SILVA, R. C. a. S. d. Probabilidade e estatística nos anos iniciais do ensino fundamental a partir da bncc. **Zetetiké**, Campinas, SP v. 28,, p. e020011, abr. 2020. ISSN 2176-1744. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/zetetike/article/view/8656990>.

SANTOS, A. L. F. d.; GUIMARÃES, P. H. S. Aplicação da linguagem de programação r no ensino de estatística descritiva no ensino médio técnico. **Revista Científica – RECIMA21**, v. 6, n. 7, p. e676628, 2025.

SANTOS, G. F. d. **A estatística em uma coleção de livros didáticos de Matemática do Ensino Fundamental II**: um olhar investigativo acerca do desenvolvimento do letramento estatístico. 2020. Tese (Trabalho de Conclusão de Curso (Licenciatura em Matemática)) — Universidade Federal do Tocantins (UFT) Araguaína, TO 2020.

SANTOS, L. R. O. d. **Formação reflexiva do professor de matemática**: uma proposta de desenvolvimento do pensamento estatístico. 2020. Tese (Tese (Doutorado em Educação)) — Universidade Federal do Piauí, Centro de Ciências da Educação Teresina, PI 2020.

SANTOS, P. A. d.; BULEGON, A. M. Contribuições dos conceitos de estatística para o desenvolvimento do pensamento crítico: análise da bncc na perspectiva da taxonomia de bloom. **Revista Poíesis**, Tubarão, SC v. 17, n. Especial, p. 139–158, 2023.

SCHREIBER, K. P.; PORCIÚNCULA, M. Mapeamento de pesquisas sobre educação estatística na biblioteca digital brasileira de teses e dissertações: um olhar para a formação de professores de matemática. **REVEMAT**, v. 14,, p. 1–17, 2019. Edição Especial Educação Estatística.

SILVA, D. G. d.; CASTRO, J. B. d. Educação estatística no ciclo de alfabetização: uma revisão sistemática sobre pesquisas realizadas antes e após a implementação da bncc. **Revista Espaço do Currículo**, João Pessoa, PB v. 16, n. 1, p. 1–16, 2023. ISSN 2177-2886. Edição jan./abr.

SILVA, J. B. A. d. **A linguagem R na educação básica**: o desenvolvimento de um site para auxiliar no ensino e aprendizagem de estatística. 2023. Dissertação (Dissertação (Mestrado Profissional em Matemática)) — Universidade Federal do Estado do Rio de Janeiro (UNIRIO) Rio de Janeiro, RJ 2023.

SILVA, J. L. d. C.; FERNANDES, M. W.; ALMEIDA, R. L. F. d. **Estatística e Probabilidade**. 3. ed. Fortaleza: EdUECE, 2015. ISBN 978-85-7826-439-0.

SILVA, M. A. d. A presença da estatística e da probabilidade no currículo prescrito de cursos de licenciatura em matemática: uma análise do possível descompasso entre as orientações curriculares para a educação básica e a formação inicial do professor de matemática.

BOLEMA, v. 24, n. 40,, 2011.

SILVEIRA, S. A. d.; SOUZA, J.; CASSINO, J. F. (Ed.). **Colonialismo de dados e modulação algorítmica**. São Paulo: Autonomia, 2021. (orgs.).

SOUZA, F. d. S. **Atividades para o ensino de probabilidade e estatística**: avaliando proposta metodológica de ensino através do software r, . 2012. Trabalho apresentado na III Jornada de Educação Matemática, Universidade Federal Fluminense, Instituto do Noroeste Fluminense de Educação Superior.

United Nations Economic Commission for Europe . **Making Data Meaningful. Part 1: A Guide to Writing Stories about Numbers**. New York and Geneva: , 2009. Disponível em: <https://ec.europa.eu/eurostat/documents/64157/4374310/32-UNECE-making-data-meaningful-Part1-EN.pdf/82cf87d2-4bc9-431c-be96-2257cf77c810>.

United Nations Economic Commission for Europe . **Making Data Meaningful. Part 2: A Guide to Presenting Statistics**. Geneva: , 2009. Disponível em: <https://ec.europa.eu/eurostat/documents/64157/4374310/33-UNECE-making-data-meaningful-Part2-EN.pdf/d5b954e2-b110-469b-a2b5-78aa7c12ab62>.

VALERI, J. **Inteligência artificial utiliza base de dados que refletem preconceitos e desigualdades**. São Paulo, 2023. Disponível em: <https://jornal.usp.br/atualidades/inteligencia-artificial-utiliza-base-de-dados-que-refletem-preconceitos-e-desigualdades/>.

WICKHAM, H. **ggplot2: Elegant Graphics for Data Analysis**. New York: Springer, 2016.

WILD, C. J.; SEBER, G. A. F. **The Language of Uncertainty: A First Course in Data**. New York, NY: John Wiley & Sons, 2000. ISBN 9780471329367.

WILLIAMSON, B. **Big Data in Education**: The digital future of learning, policy and practice. London: SAGE Publications, 2017.

ZUBOFF, S. **A era do capitalismo de vigilância**: a luta por um futuro humano na nova fronteira do poder. 1. ed. Rio de Janeiro: Intrínseca, 2021. Tradução para o português.