

Análise das oficinas didáticas de estatística no R e RStudio realizadas na Universidade Federal de Pelotas

Ana Rita de A. Mazzini^{1†}, Giselda Maria Pereira¹, Pollyane V. da Silva¹, Isadora M. da Luz Real²,
Sílvia Naiane Jappe³, Luiza Elhert Bierhals³

¹ Universidade Federal de Pelotas; Instituto de Física e Matemática; Departamento de Matemática e Estatística; Pelotas – Rio Grande do Sul, Brasil.

² Universidade Federal do Rio Grande; Instituto de Matemática, Estatística e Física; Rio grande – Rio Grande do Sul, Brasil.

³ Universidade Federal de Pelotas; Faculdade de Agronomia Eliseu Maciel; Pelotas – Rio Grande do Sul, Brasil.

Resumo: A estatística atua em consonância com diferentes áreas do conhecimento e setores da sociedade. Devido a sua importância e alta aplicabilidade torna-se indispensável para alunos de graduação e pós-graduação a aquisição de conhecimento teórico e prático no assunto. Diante disso, foi pensado um projeto de ensino que envolvesse alunos e professores para o desenvolvimento de atividades práticas relacionadas aos conteúdos abordados nas disciplinas de estatística da Universidade Federal de Pelotas (UFPel). O objetivo geral deste trabalho foi relatar a criação, a preparação, os desafios, potencialidades, e então discutir os resultados e avaliações das oficinas didáticas desenvolvidas até o ano de 2023 na UFPel. Os temas das oficinas didáticas foram os seguintes: Estatística descritiva, Inferência Estatística e Estatística Experimental. Os dados utilizados neste trabalho foram obtidos a partir do acervo digital do projeto, memórias dos integrantes da equipe e formulários de avaliação disponibilizados aos estudantes sempre ao final das oficinas. Os dados obtidos nos formulários foram tabulados e submetidos à análise de estatística descritiva com o auxílio do software R e interface RStudio. Notou-se que a criação e preparação das oficinas propiciou a aquisição de conhecimento teórico-prático aos integrantes do projeto. Na análise quantitativa dos dados observou-se uma avaliação satisfatória por parte dos participantes. O formulário de avaliação foi uma ferramenta fundamental para o aprimoramento e reedição de algumas oficinas, bem como, para o preparo de novos temas. Esses resultados mostraram que a contribuição do projeto nos processos de ensino e aprendizagem tem sido efetiva e traz perspectivas positivas para a sua continuidade.

Palavras-chave: Tecnologias Digitais; Software R; Ensino de estatística; Ensino superior, Projetos de Ensino.

Analysis of statistical didactic workshops in R and RStudio held at the Federal University of Pelotas

Abstract: Statistics operates in line with different areas of knowledge and sectors of society. Due to its importance and high applicability, it is essential for undergraduate and postgraduate students to acquire theoretical and practical knowledge on the subject. In view of this, a teaching project was designed that would involve students and teachers to develop practical activities related to the content covered in statistics disciplines at the Federal University of Pelotas (UFPel). The general objective of this work was to report the creation, preparation, challenges, potential, and then discuss the results and evaluations of the didactic workshops developed until 2023 at UFPel. The themes of the didactic workshops were as follows: Descriptive Statistics, Statistical Inference and Experimental Statistics. The data used in this article were obtained from the project's digital collection, team members' memories and evaluation forms made available to students at the end of the workshops. The data obtained from the forms were tabulated and subjected to descriptive statistics analysis with the aid of the R software and RStudio interface. It was noted that the creation and preparation of the workshops provided the project members with the acquisition of theoretical-practical knowledge. In the quantitative analysis of the data, a satisfactory evaluation by the participants was observed. The evaluation form was a fundamental tool for improving and re-editing some workshops, as well as for preparing new themes. These results showed that the project's contribution to the teaching and learning processes has been effective and brings positive perspectives for its continuity.

Keywords: Digital technologies; R Software; Teaching statistics; University education, Teaching Project.

[†] Autor Correspondente: anarita.mazzini@gmail.com

Manuscrito recebido em: 06/06/2024

Manuscrito revisado em: 30/09/2024

Manuscrito aceito em: 30/09/2024

Introdução

A Estatística atua em consonância com diferentes áreas do conhecimento e setores da sociedade. Assim, o letramento e o desenvolvimento do pensamento estatístico são fundamentais para a compreensão do mundo que nos cerca (CAZORLA; SANTANA, 2019).

O processar informações do cotidiano de forma crítica e reflexiva. E no caso do acréscimo “Estatística letramento, também entendido como literacia, segundo Silva e Carmo (2023) é a forma de compreender, assimilar e”, em Literacia Estatística, caminhamos nos mesmos percursos de Gal (2002) em que uma criança, um jovem ou um adulto se constitui letrado estatisticamente quando consegue realizar interpretações e análises críticas sobre informações estatísticas apresentadas em uma ótica social globalizada.

Devido a importância da estatística e sua alta aplicabilidade torna-se indispensável para alunos de graduação e pós-graduação a aquisição de conhecimento teórico e prático no assunto. Na universidade, as disciplinas de estatística estão presentes nos currículos de cursos de diferentes áreas do conhecimento, tais como Ciências Exatas e da Terra, Ciências Biológicas, Engenharias, Ciências da Saúde, Ciências Agrárias, Ciências Sociais Aplicadas e Ciências Humanas.

O uso crescente de *softwares* educacionais para apoiar atividades de aprendizagem demanda desenvolvimento de novas competências aos docentes, como domínio do uso do computador, compreensão do potencial pedagógico inerente à utilização do *software* educacional, planejamento e execução de atividades de aprendizagem, estabelecimento de orientações adequadas e atuação como mediador durante a utilização dos *softwares* pelos estudantes (RAABE et al., 2007).

Em vista disso, buscar e inserir novas alternativas de ensinar e aprender estatística é essencial e, pensando nesse movimento, temos as Tecnologias Digitais (TD) enquanto recursos de exploração e experimentação que fomentam a autonomia do estudante, a curiosidade e uma forma dinâmica de produzir conhecimentos (BORBA; VILLARREAL, 2005).

Acredita-se que as tecnologias digitais, principalmente, as educacionais existentes, podem promover o acesso, a apropriação e o desenvolvimento das competências esperadas pelo ensino de Estatística, de modo que os alunos possam se valer desses recursos para auxiliar no tratamento de dados e na leitura de mundo ao utilizarem conceitos estatísticos embasados em considerações extraídas de ambientes computacionais (DAMIN et al., 2018).

O *Software R* foi desenvolvido por Robert Gentleman e Ross Ihaka em 1991, no departamento de Estatística na Universidade de Auckland em Nova Zelândia, sendo seu objetivo produzir um *software* para as aulas de laboratório baseado na linguagem S, do software comercial S-Plus (VENABLES; SMITH, 2016). É um programa que utiliza a abordagem construtiva, ou seja, as análises estatísticas não são realizadas com um simples clicar de ícone. É necessário construir, por meio de linhas de comando, os procedimentos da análise, sendo fundamental o conhecimento teórico e construtivo de cada metodologia empregada para que a análise possa ser realizada (CARDOSO, 2019).

Ademais, no ambiente R, é possível utilizar uma ampla gama de pacotes disponíveis, desenvolvidos por pesquisadores e disponibilizados em uma biblioteca. Os pacotes permitem a realização de análises específicas que facilitam a utilização dos recursos pelos usuários, ainda que estes não tenham conhecimento de programação. A oferta de pacotes é ampla, contínua e envolve diversas aplicações. Além disso, é possível solucionar dúvidas utilizando as comunidades e fóruns de discussão on-line.

Atuando como docentes na área de estatística em um período superior a dez anos foi constatado que os estudantes apresentam dificuldade em compreender a importância da estatística

nas suas vidas o que leva a um desinteresse em estudar seus conteúdos. De acordo com Porciúncula e Samá, 2015, acredita-se que o trabalho por meio de projetos de ensino seja uma estratégia eficiente para construção do aprender a aprender do estudante.

Diante disso, em 2019, criou-se o projeto de ensino “Explorando R como ferramenta auxiliar no ensino da estatística”. O projeto foi desenvolvido com o intuito de motivar os discentes para a construção do conhecimento estatístico visando melhorar a assimilação dos conceitos e práticas estatísticas, proporcionando aos alunos uma experiência mais interativa e aplicada usando ferramentas computacionais e tecnologia digital. A partir do exposto, organizou-se uma oficina didática piloto para verificar o interesse e adesão dos estudantes a esta iniciativa, como de responder a questões tais: Como o uso do *software* R pode facilitar o ensino da estatística? Quais são as principais barreiras enfrentadas pelos estudantes no aprendizado de estatística com e sem o apoio de ferramentas computacionais e tecnologias digitais?

A oficina piloto foi muito bem aceita pelos estudantes participantes, que demonstraram interesse em realizar novas oficinas e sugeriram temas de interesse acadêmico. A partir desta demanda a equipe do projeto desenvolveu e aplicou novas oficinas com diferentes conteúdos de estatística. Essas experiências culminaram na criação do projeto de ensino intitulado “Ferramentas computacionais no ensino da estatística”, com intuito de utilizar outras tecnologias (além do R e RStudio) para o ensino da estatística.

O objetivo geral deste artigo foi relatar a criação, a preparação, os desafios, potencialidades, então discutir os resultados e avaliações das oficinas didáticas desenvolvidas pelos projetos de ensino até o ano de 2023 na UFPel.

Materiais e métodos

Foram realizadas oficinas didáticas, que abrangeram os seguintes temas: Estatística Descritiva, Inferência Estatística e Estatística Experimental. O grupo de estudos contou com a participação de quatro professoras da área de Estatística do Departamento de Matemática e Estatística (DME) do Instituto de Física e Matemática (IFM) da UFPel, alunos de diferentes cursos de graduação da UFPel e uma pós-graduanda da Universidade Federal do Rio Grande (FURG).

A organização das oficinas foi realizada em encontros semanais, de forma conjunta onde cada integrante ficou responsável pela elaboração de uma parte do material. A primeira oficina foi realizada no ano de 2019 de forma presencial. As oficinas no período de 2020 a 2022 foram realizadas de forma remota, com o auxílio de ferramentas como o *Google Meet* e *WebConf* da UFPel. Em 2023, as oficinas retornaram para o formato presencial. Todas as oficinas presenciais ocorreram no Laboratório de Novas Tecnologias, do Instituto de Física e Matemática da UFPel, no *campus* Capão do Leão. As oficinas remotas foram realizadas no turno da noite enquanto que as presenciais no intervalo de almoço. A avaliação das oficinas foi registrada ao término das atividades a partir de formulários impressos ou *Google Forms*, de acordo com a modalidade da oficina.

Os dados deste trabalho foram extraídos dos formulários de avaliação e se referem a sete oficinas didáticas realizadas para alunos de graduação e pós-graduação da UFPel, no período de 2019 a 2023. Os dados foram tabulados e submetidos a análise de estatística descritiva com o auxílio do *software* R (R Core Team, 2023 - versão 4.3.1) e interface RStudio (RStudio, 2020). Os gráficos foram gerados com a utilização do pacote *ggplot2* (WICKHAN, 2016).

Resultados e discussão

Inicialmente as oficinas abordaram o emprego de metodologias de estatística descritiva as quais apresentaram atividades sobre representação gráfica, medidas de resumo e análise exploratória de dados. Posteriormente, foram desenvolvidas oficinas contemplando conteúdos de inferência estatística e estatística Experimental. Para a criação e o desenvolvimento das oficinas foi realizada uma ampla pesquisa bibliográfica e, a partir de artigos selecionados, bancos de dados foram simulados nas áreas de ciências agrárias, ciências exatas, ciências biológicas e ciências humanas. Os diferentes bancos de dados possibilitaram exemplificar a utilização da estatística para diversos cursos de graduação da universidade. Silva (2018) afirma que a estatística é uma ciência multidisciplinar, que permite a análise estatística de dados em todas as áreas científicas.

Para a prática das atividades propostas a equipe do projeto dava início às oficinas apresentando uma breve revisão teórica dos conteúdos abordados. No decorrer das atividades, a interpretação das saídas computacionais era enfatizada relacionando os resultados obtidos com suas implicações práticas na área de estudo. Além disso, foram desenvolvidos materiais de apoio para cada oficina, os quais apresentam a fundamentação teórica e o passo a passo da prática computacional. Estes materiais foram disponibilizados via e-mail ou na plataforma *e-projeto* da UFPel, de tal forma que os participantes das oficinas pudessem acessá-los livremente. Com relação ao formato presencial ou remoto das oficinas, o grupo de trabalho realizou ajustes na execução destas para adequação à modalidade a ser utilizada.

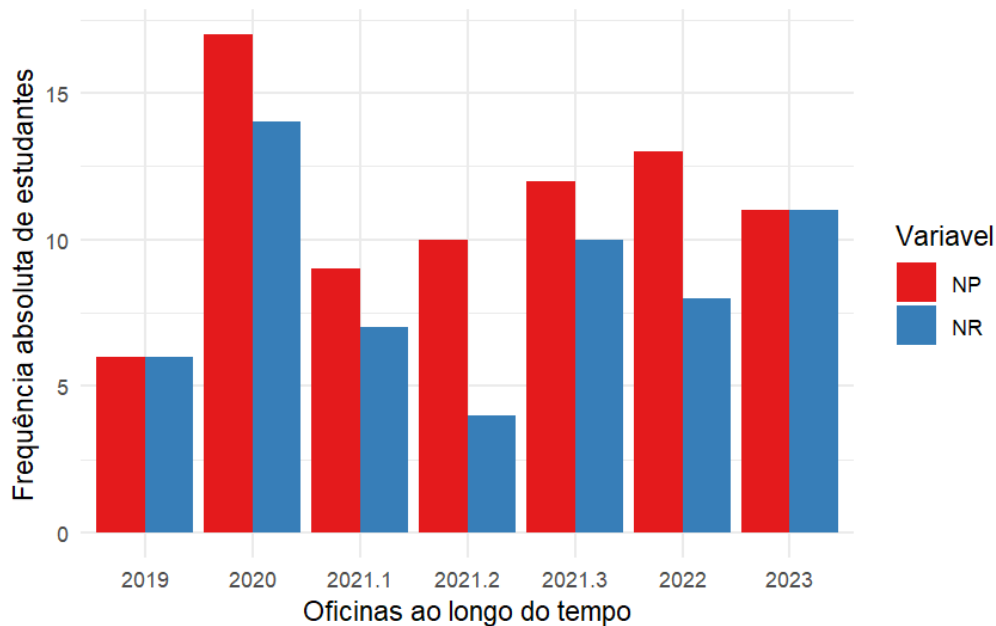
O processo de preparo e adequação das oficinas possibilitou aos alunos integrantes do projeto aquisição de experiências, troca de saberes, desenvolvimento de novas habilidades e competências relacionadas à utilização de ferramentas computacionais para análises estatísticas e para o ensino remoto, construção de conhecimentos para aplicação dos conteúdos abordados em trabalhos acadêmicos, bem como, experiência docente supervisionada.

Um recurso importante para continuidade das ações dos projetos foi o desenvolvimento de instrumento de avaliação (questionário) das oficinas. O gráfico apresentado na Figura 1 corresponde ao número de participantes (NP) nas sete oficinas e número de respondentes (NR) do formulário de avaliação.

É notável o expressivo interesse dos estudantes em participarem das oficinas, estes desempenharam um papel fundamental no andamento do projeto e na proposição de novas oficinas. Vale ressaltar que pelo fato das oficinas ofertadas de 2020 a 2022 terem sido realizadas de forma remota, não houve restrições no número de participantes. No entanto, na oficina de 2023, que marcou o primeiro encontro presencial pós-pandemia, testemunhamos um interesse maciço dos estudantes, o que levou a equipe limitar o número de vagas para acomodar todos os interessados no laboratório de informática.

As perguntas que fizeram parte do questionário correspondem a variáveis do tipo qualitativas, as quais foram cuidadosamente formuladas com o objetivo de obter informações que possibilitaram direcionar melhorias contínuas nas ações do projeto, sempre com foco no atendimento dos interesses do público-alvo. Sendo assim, estão apresentadas na Tabela 1 as questões objetivas com as respostas dos estudantes e as respectivas porcentagens.

Figure 1: Number of participants (NP) and number of respondents (NR) in the workshop evaluation questionnaire.



Source: from the authors (2024).

Table 1: Questions included in the workshop evaluation forms and their respective responses.

Perguntas	Respostas (%)
A oficina atendeu suas expectativas?	
Sim	93,33
Parcialmente	6,67
Não	-
Você diria que seu aproveitamento na oficina foi bom?	
Sim	88,33
Parcialmente	11,67
Não	-
Quanto ao horário de realização desta oficina, você achou adequado?	
Sim	96,67
Parcialmente	3,33
Não	-
Os tópicos abordados poderão ser utilizados na sua vida acadêmica?	
Sim	96,67
Parcialmente	3,33
Não	-
Você tem interesse em participar de outras oficinas de análises em R?	
Sim	98,33
Não	1,67

Source: from the authors (2024).

Ao analisar os dados da Tabela 1, observou-se que as oficinas atenderam as expectativas de mais de 93% dos respondentes. No que se refere a opinião sobre a adequação do horário de realização das oficinas e se os tópicos abordados contribuem com a vida acadêmica, evidenciou-se que 96,67% destes afirmaram que sim. Sobre o aproveitamento das oficinas realizadas, 88,33% responderam que seu aproveitamento foi bom, enquanto 11,67% responderam que foi parcialmente bom. Além disso, 98,33% declararam ter interesse em participar de novas oficinas. Essas respostas mostram que as oficinas atingiram seu propósito e que a continuidade destas ações tem grande relevância para formação dos discentes.

A partir das respostas do questionário foi possível identificar os pontos que poderiam ser melhorados, as necessidades e expectativas dos participantes para então propor novas atividades. Para isso, no mesmo formulário foi colocado o item “Sugestões” e deixou-se a resposta de forma dissertativa para que o estudante participante tivesse a oportunidade de colocar seu parecer em relação à oficina e também fazer sugestões para as próximas. Algumas das sugestões mais relevantes foram: ofertar mais oficinas, aumentar a carga horária das mesmas, período de realização das oficinas (início de semestre), oficinas presenciais, horários extras de atendimento (monitorias) relacionados a prática dos conteúdos abordados.

Para as oficinas realizadas a partir de 2020, a única modificação feita no formulário foi a adição da pergunta “Você tem interesse em participar de outras oficinas de análise em R?”, neste caso, se a resposta fosse sim, perguntou-se: “Qual o assunto?”. Nesta questão quase que a totalidade dos respondentes demonstraram interesse em participar de outras oficinas e sugeriram os seguintes temas: inferência estatística, estatística experimental, construção de funções no R, gráficos interativos, análise de dados categóricos, análise multivariada, uso de outros pacotes do R, sendo que uma das sugestões foi o pacote ggplot2. Tendo em vista as sugestões recebidas naquele momento, a equipe já ofertou oficinas com os seguintes temas: inferência estatística, estatística experimental e estatística descritiva com o uso do pacote ggplot2.

A equipe segue realizando oficinas didáticas e está planejando novas ações que abordem os demais temas sugeridos, bem como, a utilização de novas ferramentas computacionais. Assim que possível, preferencialmente com auxílio de bolsistas, a equipe do projeto pretende ampliar as ações de atendimento extra (monitorias) referente aos conteúdos das oficinas, com o objetivo de atender uma das pertinentes sugestões encaminhadas pelos alunos. Outra sugestão relevante, foi o aumento do tempo de duração das oficinas que será levada em consideração pela equipe do projeto de acordo com a disponibilidade dos laboratórios de multimídias da UFPel.

Observou-se que uma das motivações dos participantes da oficina presencial de 2023 foi aprender a utilizar uma ferramenta computacional para análise e interpretação dos dados gerados em seus TCC's (Trabalho de Conclusão Curso). Este fato evidencia a importância da estatística que conforme Morettin (1981) é responsável pelo planejamento de experimentos, interpretação dos dados obtidos através de pesquisas de campo e apresentação de resultados de maneira a facilitar a tomada de decisões por parte do pesquisador/gestor. Além disso, Bosoni e Bruzzone (2018) destacam diversas vantagens do uso do *software* R, como a replicabilidade, eficiência e controle nos processos analíticos. Essas características permitem que os alunos aprimorem seus trabalhos acadêmicos, proporcionando uma abordagem mais precisa e profissional no tratamento dos dados.

Alguns desafios foram enfrentados pela equipe do projeto especialmente no início das atividades como: atrair os estudantes para as oficinas, realizar a implementação das atividades, definir o número ideal de estudantes para atuar no projeto, dentre outras. Com a chegada da pandemia da COVID-19, houve a necessidade de adequar as oficinas para o formato remoto, sendo necessário repensar a forma de treinar os estudantes para realizar as tarefas do projeto. Além disso, observou-se dificuldade de ter estudantes dispostos a participar do projeto, principalmente no

período de pandemia. No ano de 2020, o projeto foi contemplado com uma bolsa de ensino da própria universidade o que favoreceu a execução das atividades e atraiu a atenção dos discentes para o projeto. Nos anos subsequentes contamos apenas com alunos voluntários para o andamento do projeto, o número reduzido de alunos dificulta o atendimento de algumas demandas. Este fato evidencia a importância das bolsas de estudo aos projetos de ensino.

Por fim, pode-se dizer que os resultados das oficinas mostraram que a utilização de ferramentas computacionais atrai o interesse dos estudantes, contribuindo com a construção do conhecimento de Estatística e favorecendo o desenvolvimento de competências que poderão ser utilizadas em suas vidas acadêmicas e profissionais. Deve-se ressaltar, contudo, que conforme Riedner e Pischetola (2021), o uso de tecnologias digitais pode ser considerado como uma inovação a qual atua como meio e não como fim em si mesmo.

Considerações Finais

A criação e preparação das oficinas propiciou a aquisição de conhecimento teórico-prático para os integrantes do projeto. As sete oficinas de estatística realizadas utilizando o *software* R e interface RStudio tiveram uma avaliação satisfatória por parte dos participantes. O formulário/questionário de avaliação foi uma ferramenta fundamental para o aprimoramento e a reedição das oficinas, bem como para o preparo de novos temas. Os desdobramentos dessas oficinas foram significativos. Além de fortalecer a compreensão dos conceitos estatísticos, o projeto despertou nos participantes o interesse pela exploração de ferramentas computacionais, como o R. Isso foi especialmente relevante para os alunos que buscavam melhorar suas habilidades de análise de dados em projetos de pesquisa, como os Trabalhos de Conclusão de Curso (TCC). As oficinas estimularam a integração entre discentes e docentes, formando um grupo de estudos que contribui para a evolução do processo ensino-aprendizagem. Este grupo, por sua vez, está criando uma rede de apoio entre seus membros, fomentando discussões sobre temas de estatística e novas metodologias de ensino utilizando ferramentas computacionais. Com base nos resultados obtidos até o momento, as perspectivas para a continuidade do projeto são extremamente positivas. O engajamento dos participantes e o impacto gerado pelas oficinas sugerem que o uso do R não só facilita o aprendizado estatístico, mas também abre portas para novos horizontes na formação acadêmica e profissional dos alunos, preparando-os para os desafios do mercado de trabalho e da pesquisa científica. Por fim, acreditamos que as oficinas didáticas utilizando o *Software* R e a interface RStudio estimulam a utilização destas importantes ferramentas computacionais de livre acesso no ambiente acadêmico da UFPel.

Agradecimentos

A PRAE/UFPel pela concessão de bolsa a estudante do projeto.

Referências

BORBA, M. C.; VILLARREAL, M. E. **Humans-with-media and the reorganization of mathematical thin-king**: Information and communication technologies, modeling, visualization and experimentation. Springer Science & Business Media, 2005.

BOSONI, G. B.; BRUZZONE, F. R. Uso de RStudio para Estadística Univariada en Ciencias Sociales. **no. June**, 2018.

CARDOSO, M. G. **Ensino de estatística: o estudo de conceitos potencializado pelo software RStudio**. Dissertação (Mestrado em Ensino de Matemática) – Universidade Tecnológica Federal do Paraná, Londrina (PR), 2019.

CAZORLA, I. M.; SANTANA, E. R. S. Estatística para a leitura de mundo. In: CONFERÊNCIA INTERAMERICANA DE EDUCAÇÃO MATEMÁTICA 15 – CIAEM, 2019.

DAMIN, W.; SANTOS JUNIOR, G.; COELHO NETO, J.; ROBIM, B. N. P. A. S.; PEREIRA, R. S. G. As Tecnologias Digitais Educacionais e o Ensino de Estatística e Probabilidade. **Revista de Ensino, Educação e Ciências Humanas**, [S. l.], v. 20, n. 1, p. 53–57, 2019.

GAL, I. Adults' statistical literacy: Meanings, Componentes, Responsibilities. **International Statistical Review**, 70(1), p. 1-25, 2002.

MORETTIN, P. A. **Introdução à estatística para ciências exatas**. São Paulo: Atual, 1981.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2023. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

RAABE, A. L. A.; ALVES, A. G. Oficinas de utilização de Software Educacional: Um relato de experiência. In: **Anais do Workshop de Informática na Escola**. 2007. p. 264-271.

RIEDNER, D. D. T.; PISCHETOLA, M. A inovação das práticas pedagógicas com uso de tecnologias digitais no ensino superior: um estudo no âmbito da formação inicial de professores. **ETD Educação Temática Digital**, v. 23, n. 1, p. 64-81, 2021.

RSTUDIO, Team. RStudio: integrated development for R. RStudio Team, PBC, Boston, MA, 2020. URL <http://www.rstudio.com/>.

SILVA, M.P.M.da; SAMÁ,S.; Educação Estatística: ações e estratégias pedagógicas no Ensino Básico e Superior. In: SAMÁ, S.; SILVA, M. P. M. da. (Org). **PROJETOS DE APRENDIZAGEM: uma proposta pedagógica para a sala de aula de Estatística**. Curitiba: Editora CRV, 2015. p.133-141.

SILVA, P. V.; CARMO, J. V. G. Educação Estatística e software RStudio: perspectivas na educação matemática. In: **Anais do IV SITEM - Simpósio Internacional de Tecnologias em Educação Matemática**. UNESP/IGCE, 2023.

WICKHAM, H.; CHANG, W.; WICKHAM, H. Package ‘ggplot2’. Create elegant data visualisations using the grammar of graphics. Version, v. 2, n. 1, p. 1-189, 2016.