

PROPOSTA DE APLICAÇÃO DE UM MATERIAL DIDÁTICO PARA A APRENDIZAGEM DE ESTATÍSTICA POR MEIO DO *SOFTWARE R*, USANDO QUESTÕES DO ENEM

Eduardo César Santos¹
Paulo José Pereira²

Resumo: O ensino de Matemática passa por atualizações uma delas o uso de novas tecnologias. O objetivo desse trabalho foi desenvolver e avaliar um material didático que utilize o *software R*, que se trata de uma aplicação gratuita para ensinar conceitos de Estatística com o uso de questões da principal forma de acesso ao ensino superior ENEM. A pesquisa foi realizada com alunos da turma do terceiro ano de ensino médio de uma escola no município de Remanso, BA. Inicialmente, foi avaliado, via um pré-teste, para identificar o nível em que se encontravam. Na sequência, foram ministradas aulas de resolução de exercícios com o uso do *software R*. Por fim, analisou-se a efetividade do processo de aprendizagem através de um novo teste. A análise dos dados para comparar as notas nos dois momentos foi feita por meio do teste de Wilcoxon ($p < 0,05$), de modo a constatar uma evolução no aprendizado dos estudantes. Também foram coletadas opiniões dos estudantes. A utilização de tecnologias educacionais representa um avanço significativo no processo de ensino-aprendizagem de conteúdos da Estatística.

Palavras-chave: Ensino de Matemática. Ensino de Estatística. Caderno Didático. *Software R*. Tecnologia.

PROPOSAL FOR THE IMPLEMENTATION OF A DIDACTIC MATERIAL FOR LEARNING STATISTICS THROUGH THE R SOFTWARE USING ENEM QUESTIONS

Abstract: The teaching of Mathematics is undergoing updates, one of which is the use of new technologies. The objective of this work was to develop and evaluate a teaching material that uses the R software, which is a free software, to teach Statistics concepts using questions from the ENEM (Brazilian National High School Exam), the main form of access to higher education in Brazil. The research was conducted with students from the third year of high school from a school in the municipality of Remanso, BA. Initially, a pre-test was applied to identify the level at which the students were. Subsequently, classes focused on solving exercises were taught using the R software. Finally, the effectiveness of the learning process was analyzed through a new test. The data analysis to compare the scores at the two moments was done using the Wilcoxon test ($p < 0.05$), in order to verify an improvement in student learning. Student opinions were also collected. The use of educational technologies represents a significant advance in the teaching-learning process of Statistical content.

Keywords: Mathematics Teaching. Statistics Teaching. Didactic Notebook. R Software. Technology.

PROPUESTA DE APLICACIÓN DE UN MATERIAL DIDÁCTICO PARA EL APRENDIZAJE DE ESTADÍSTICA MEDIANTE EL *SOFTWARE R* UTILIZANDO PREGUNTAS DEL ENEM

Resumen: La enseñanza de las Matemáticas se encuentra en constante actualización, una de las cuales es el uso de nuevas tecnologías. El objetivo de este trabajo fue desarrollar y evaluar un material didático que utiliza el *software R*, el cual es un *software* libre, para enseñar conceptos estadísticos mediante preguntas del ENEM (Examen Nacional de Enseñanza Media de Brasil), la principal forma de acceso a

¹ Mestre em Matemática, UNIVASF, Juazeiro, Bahia, Brasil. eduardo.gama1996@hotmail.com. <https://orcid.org/my-orcid?orcid=0009-0006-2785-5502>

² Doutor em Demografia, UNIVASF, Juazeiro, Bahia, Brasil, paulo.pereira@univasf.edu.br, <https://orcid.org/0000-0002-4436-8304>

la educación superior. La investigación se realizó con estudiantes del tercer año de secundaria de una escuela en el municipio de Remanso, BA. Inicialmente, se utilizó una prueba previa para identificar su nivel. Posteriormente, se impartieron clases centradas en la resolución de ejercicios utilizando el software R y, finalmente, se analizó la efectividad del proceso de aprendizaje mediante una nueva prueba. El análisis de los datos para comparar las puntuaciones en ambos momentos se realizó mediante la prueba de Wilcoxon ($p < 0,05$), y se observó una mejora en el aprendizaje de los estudiantes. También se recogieron las opiniones de los estudiantes. El uso de tecnologías educativas representa un avance significativo en el proceso de enseñanza-aprendizaje de los contenidos estadísticos.

Palabras clave: Enseñanza de Matemáticas. Enseñanza de Estadística. Cuaderno Didáctico. Software R. Tecnología.

1 Introdução

O ensino da Matemática é uma preocupação central em diversos centros acadêmicos ao redor do mundo. Muitas discussões incentivam os educadores a refletirem sobre as melhores práticas de ensino, sendo que estas devem promover uma aprendizagem mais significativa para os alunos.

Cardoso *et al.* (2021) uma parte da Matemática, que é bastante contextualizada, é a da Estatística. Seu uso no âmbito educacional pode ser abordado e aplicado não apenas nas Ciências Exatas, mas também de maneira interdisciplinar em outras disciplinas, estabelecendo conexões com o uso de Estatísticas. Pode-se mencionar, como alternativa para o ensino de Estatística, o *software R*, conforme destacado por Ihaka e Gentleman (1996), que o classificam como uma poderosa ferramenta de análise de dados, que permite a visualização e manipulação de grandes volumes de informações. Pensando no desenvolvimento dos alunos do ensino médio voltados ao uso da estatística a integração dessas questões com o *software R*, o emprego desse recurso tem potencial para fornecer um cenário prático para a aplicação de técnicas Estatísticas, facilitando a compreensão dos conteúdos pelos alunos e aumentando a percepção de sua relevância.

Ao longo da vivência em sala de aula, constatou-se que muitos estudantes demonstram grande resistência e dificuldade em compreender conceitos fundamentais da Estatística, tais como média, variância, distribuições de frequência e inferência Estatística relacionados a interpretação de informações contidas em gráficos e tabelas, essa dificuldade foi apontada por Fernandes *et al.* (2004) e, após a apresentação do conteúdo junto aos estudantes, estes apresentaram problemas no aprendizado.

Por outro lado, as questões do Exame Nacional do Ensino Médio (ENEM), que, frequentemente, abordam problemas estatísticos em contextos variados, exigem dos estudantes a aplicação de conceitos em situações reais. Além disso, o referido exame consolidou-se como a maior e principal forma de acesso ao ensino superior no país, apresentando questões

contextualizadas e de relevância para o interesse dos estudantes. A integração dessas questões com o *software* R pode vislumbrar um cenário prático para a aplicação de técnicas estatísticas, facilitando a compreensão dos conteúdos pelos alunos e aumentando a percepção de sua relevância.

Neste cenário, associar uma temática de suma importância em um ensino indisciplinar como a Estatística, que está e sempre esteve presente em todas as provas do ENEM desde sua implementação, associada a um dos principais *softwares* estatísticos de código livre e gratuito, que é o R, surge como possibilidade de melhoria na aprendizagem dos estudantes.

2 Metodologia

A metodologia deste trabalho está subdividida em dois tópicos. O primeiro deles aborda a classificação da pesquisa e explora, de forma objetiva, como este trabalho foi desenvolvido, pormenorizando o processo usado na pesquisa que foi submetida ao comitê de ética.

Vale ressaltar que a presente pesquisa adota uma abordagem quantitativa, que é adequada para a avaliação do impacto de intervenções pedagógicas, pois possibilita a análise de dados numéricos e a verificação de resultados por meio de métodos estatísticos. Segundo Cohen, Manion e Morrison (2011), a pesquisa quantitativa tem como foco principal a mensuração precisa e objetiva de fenômenos, permitindo que as relações entre variáveis sejam analisadas com base em dados numéricos. O estudo é de natureza aplicada e experimental, pois buscou resolver uma questão prática no ensino de Estatística por meio do emprego do *software* R, desenvolvendo um material didático para observar seus efeitos no desempenho dos alunos. Seu objetivo também é descritivo, focado em identificar e descrever fenômenos sem explicar suas causas, tendo em vista que, a pesquisa investiga como o uso de materiais didáticos e ferramentas tecnológicas contribui para o ensino da Estatística no ensino básico, destacando sua eficiência na coleta, tratamento e análise de dados

2.1 Comitê de Ética

Este estudo foi submetido ao Comitê de Ética, sendo aprovado com o número 81613624.0.0000.0282, julgado pelo Comitê de Ética do Hospital Universitário da UNIVASF em 23 de setembro de 2024.

2.2 Local do estudo e população estudada

O estudo foi realizado em uma escola de Ensino Médio localizada na zona urbana do município de Remanso, Bahia. A população-alvo do estudo foram os alunos do 3º ano, turma A. A referida classe era composta por 32 alunos com idades entre 17 e 19 anos.

2.3 Procedimentos de recrutamento e coleta de dados

Os alunos foram informados sobre a pesquisa durante uma aula e convidados a participar voluntariamente. Foi solicitado o consentimento informado dos alunos e, no caso de menores de idade, também o consentimento dos pais ou responsáveis.

2.4 Coleta de dados

Os dados foram coletados em três momentos distintos:

2.4.1.1 Pré-teste

O teste era composto por sete exercícios objetivos do ENEM e três subjetivos, elaborados pelo autor: todos os dez foram retirados do caderno elaborado, com nota máxima de 10 pontos.

2.4.1.2 Intervenção

Implementação de um curso intensivo de Estatística, com duração de um mês, sendo oito horas-aula, utilizando o *software* R com questões do ENEM como ferramenta principal. O curso foi desenvolvido ao longo de quatro semanas seguidas, abrangendo desde da instalação do *software* até a resolução de questões complexas com auxílio da ferramenta.

2.4.1.3 Pós-teste

Aplicação de um teste de Estatística similar ao pré-teste para avaliar o ganho de conhecimento dos alunos após a intervenção, feito com dez questões divididas em sete exercícios objetivos do ENEM e três subjetivos criados pelo autor, com nota máxima de 10 pontos.

2.5 Análise dos dados

Os dados coletados no pré e no pós-teste foram analisados para comparar o desempenho dos alunos antes e depois da intervenção. Pelo fato de os dados obtidos não serem paramétricos, foi analisada a conformidade dos resultados via teste de Wilcoxon para amostras pareadas de

um mesmo grupo, para determinar se houve um aumento significativo no desempenho dos alunos considerando um $p < 0,05$.

2.5.1 Software R

A criação da linguagem R, orientada a objetos, ocorreu em 1996 por Ross Ihaka e Robert Gentleman que aliada a um ambiente integrado permite a manipulação de dados, realização de cálculos e geração de Gráficos. É importante ressaltar que o R não é um programa estatístico, mas que devido a suas rotinas permite a manipulação, avaliação e interpretação de procedimentos estatísticos aplicado a dados (De Souza *et al.* 2014).

Assim, o *software* foi usado no momento da intervenção junto aos estudantes sendo um instrumento de ensino e de onde se coletará as informações do momento após a intervenção. O R é uma linguagem de programação e ambiente de *software* desenvolvido, especificamente, para computação estatística e de gráficos. Ele é amplamente utilizado em diversos campos, como ciência de dados, pesquisa acadêmica, bioinformática, economia, finanças e muitos outros, devido à sua flexibilidade e capacidade de lidar com análise estatística complexa.

3 Resultados e Discussão

Na etapa inicial do estudo, o “pré-teste” avaliou o nível de conhecimentos dos alunos. Na data de realização do pré-teste, havia 29 alunos presentes, de modo que os resultados de medidas de tendência central e a dispersão das notas da avaliação podem ser vistos na Tabela 1, disponível abaixo.

Tabela 1 - Medidas de tendência central e dispersão do pré-teste

Medida	Resultado
Media	2,1
Moda	1 e 3
Mediana	2
Variância	2
Desvio Padrão	1,45
Coefficiente de Variação	69,05%

Fonte: Elaborado pelo autor.

A seguir, dispomos a Tabela 2, com as medidas obtidas no pós-teste, sendo realizada com auxílio do programa.

Tabela 2 – Medidas de tendência central e dispersão do pós-teste

Medida	Resultado
Media	4,15
Moda	5
Mediana	4,5
Variância	2,65
Desvio Padrão	1,61
Coefficiente de Variação	39,51%

Fonte: Elaborado pelo autor.

A análise da Tabela 2 evidencia uma evolução significativa do desempenho da turma após o ensino do *software* R, destacando-se, inicialmente, um aumento aproximado de 97% na média geral da sala. Tal crescimento expressivo indica uma melhoria consistente no rendimento dos estudantes como um todo. Observa-se que, no período anterior à intervenção, a moda apresentava valor superior à média e esta, por sua vez, era inferior à mediana, revelando maior assimetria e variação entre as notas. Após a utilização do R, a elevação da moda contribuiu para compensar a presença de valores mais baixos, resultando no aumento da média, além de uma maior aproximação entre média e mediana, o que sugere uma distribuição mais equilibrada dos dados. Outro aspecto relevante refere-se ao aumento tanto da nota mínima quanto da máxima da turma, indicando que houve progresso não apenas entre os estudantes com melhor desempenho, mas também entre aqueles que apresentavam maiores dificuldades. Ademais, a redução do coeficiente de variação no período posterior demonstra que a turma se tornou mais homogênea, evidenciando que os ganhos de aprendizagem foram mais uniformes.

Do ponto de vista inferencial, o teste de Wilcoxon confirmou os resultados descritivos. O p-valor obtido (0,0002264), significativamente inferior ao nível de significância adotado de 0,05, permite rejeitar a hipótese nula, indicando que a diferença observada entre os grupos “antes” e “depois” não pode ser atribuída ao acaso, mas sim a uma mudança real no desempenho dos estudantes. Dessa forma, conclui-se que a mediana do grupo “antes” é estatisticamente menor do que a mediana do grupo “depois”, validando a hipótese alternativa e comprovando uma evolução significativa da aprendizagem em Estatística.

Essa evolução pode ser caracterizada como multidimensional, envolvendo aspectos procedimentais, conceituais e interpretativos. No âmbito procedimental, os estudantes passaram

a demonstrar maior domínio na execução de procedimentos estatísticos e no uso adequado do *software* R. Em termos conceituais, a melhoria nas medidas de tendência central e a maior consistência das notas indicam uma compreensão mais sólida dos conceitos estatísticos, indo além da simples reprodução de comandos. Já no aspecto cognitivo e interpretativo, os avanços podem ser associados ao fato de que os testes aplicados eram compostos por questões do ENEM, que exigem leitura atenta, interpretação de enunciados e tomada de decisões fundamentadas, competências essenciais para a aprendizagem em Estatística.

Ressalta-se que os testes aplicados iam além da execução mecânica de comandos, exigindo dos alunos a articulação entre conhecimento matemático e interpretação contextualizada dos problemas. Essa prática adicional de resolução de exercícios pode ter contribuído significativamente para o aumento das notas, especialmente considerando que o 3º ano do ensino médio dispõe atualmente de apenas duas aulas semanais de Matemática. Embora parte da turma tenha demonstrado certo desinteresse, possivelmente em função da ausência de recompensas avaliativas formais, os resultados obtidos indicam uma evolução clara e consistente do desempenho da turma como um todo. Assim, com base nas evidências estatísticas e pedagógicas apresentadas, conclui-se que a utilização do *software* R teve impacto positivo e significativo na aprendizagem em Estatística, promovendo avanços quantitativos e qualitativos no desempenho dos estudantes.

Além disso, os alunos foram interpelados sobre a contribuição do aplicativo para o aprendizado de estatística, os dados estão presentes na tabela 3 abaixo.

Tabela 3 – Resultado dos questionários

Pergunta	Concordo (%)	Discordo (%)	Indiferente (%)
Facilidade de aprendizado	41	31	28
Ajuda na Compreensão de Estatística	65,52	24,14	10,34
Atividades Interativas e Aprendizado Envolvente	65,52	17,24	17,24
Motivação para aprender Estatística	27,59	44,82	27,59
Facilidade na Resolução de Problemas	62,07	13,79	24,14
Design e Navegação	37,93	51,73	10,34
Clareza na Explicação e Exemplos	41,38	34,48	24,14
Relevância do Conteúdo	79,31	13,79	6,90
Satisfação Geral com o <i>Software</i>	62,07	10,34	27,59

Pergunta	Concordo (%)	Discordo (%)	Indiferente (%)
Recomendação para outros alunos	68,97	17,24	13,79

Fonte: Elaborado pelo autor.

Foi possível observar que, o *software* foi bem avaliado em relação à compreensão dos conceitos de Estatística, atividades interativas e relevância do conteúdo. Mais de 60% dos alunos demonstraram-se satisfeitos e recomendariam o uso para outros. Por outro lado, houve críticas sobre dificuldade em relação ao uso e a interface do programa, observando, ainda, que o uso de tal programa não foi um ponto tão motivador dos estudantes. A discordância sobre o *design* e a navegação pode ter impactado a percepção de facilidade no uso do *software*. Estudantes que acharam o *design* confuso, provavelmente, enfrentaram dificuldades adicionais no aprendizado, sugerindo que, embora as atividades interativas tenham sido vistas de forma positiva, elas podem não ter sido suficientes para influenciar a motivação geral dos alunos.

A falta de motivação pode ser causada por outros fatores, como a dificuldade de usar a aplicação ou pela pouca experiência em uso de computadores, podendo ser mencionado ainda a falta de recompensas. A taxa de recomendação, ligeiramente maior que a de satisfação, sugere que, mesmo entre os que não estavam totalmente satisfeitos, muitos ainda viam valor no *software* como uma boa oportunidade de melhora nos conhecimentos relacionados a Estatística, dando-o como sugestão para outros estudantes. Esse resultado pode indicar que os alunos consideram o *software* como uma ferramenta útil, apesar de suas limitações, possivelmente por reconhecerem o potencial de aprendizado que ele oferece e as facilidades que ele proporciona.

Ressalta-se, ainda, que, embora a maioria dos estudantes tenha avaliado o uso do R como indiferente ou difícil, os resultados do teste, em comparação ao pré-teste, evidenciaram um aumento considerável nas notas. Tal resultado permite inferir que, mesmo quando se avalia seu uso como difícil, ele contribuiu de forma significativa para oportunizar o aprendizado. Constata-se, também, que essa dificuldade não os fez discordarem da sua funcionalidade e de uma recomendação para outros estudantes.

4 Conclusão

A vivência em sala de aula demonstrou que a dificuldade do aprendizado de Estatística por parte dos alunos se faz presente em grande parte destes. Tal ponto foi confirmado por um pré-teste com a turma utilizada, em que a nota média atingida pela turma foi consideravelmente inferior à média do sistema de ensino. Dessa forma, mostrando a necessidade de uma intervenção para a melhoria do aprendizado destes estudantes.

Ao serem realizados estudos diante dessa temática, surgiu como possível solução o uso de tecnologias para que fossem auxiliares no aprendizado da disciplina, sendo escolhido o *software* R, que já é amplamente difundido no ensino de estatística e é gratuito. O material utilizado buscou contemplar exercícios contextualizados que possuem uma natureza que faz o aluno refletir situações reais de aplicação dos conteúdos. Para tanto, tomou-se como referência questões aplicadas no ENEM, possibilitando o uso do caderno de forma isolada, sem a necessidade do uso do *software* R. Os resultados indicam que a utilização de recursos tecnológicos, como o que se empregou, pode potencializar o aprendizado de disciplinas complexas, como a Estatística, ao proporcionar *feedback* imediato e permitir uma prática constante com questões contextualizadas.

A possibilidade de simular condições semelhantes às de uma prova real, como o ENEM, também foi um fator que contribuiu para a maior preparação dos estudantes, indicando a conclusão dos objetivos do trabalho de produzir e avaliar a efetividade do material. Além disso, a análise dos dados coletados revelou que o uso desse recurso tecnológico ajudou a aumentar a motivação dos alunos para estudar Estatística, mostrando que o aprendizado ativo, aliado a ferramentas tecnológicas, tem grande potencial para melhorar a qualidade do ensino. A aplicação de questões reais do ENEM proporcionou uma conexão direta com o conteúdo abordado na avaliação nacional, o que pode ter aumentado a relevância do estudo para os alunos. É importante pontuar que, para a aplicação desses projetos, a escola esteja equipada com computadores que tenham acesso à internet. Tais ferramentas auxiliam tanto o ensino do professor das mais diversas disciplinas quanto a aprendizagem do 48 estudante. Conclui-se que a utilização de tecnologias educacionais, como o *software* analisado, representa um avanço significativo no processo de ensino-aprendizagem de Estatística, especialmente, em contextos de avaliação como o ENEM. De fato, a combinação de questões aplicadas e ferramentas digitais pode ser uma estratégia eficaz para aprimorar o aprendizado e a preparação dos estudantes, contribuindo para o sucesso acadêmico.

Referências

CARDOSO, M. G.; STURION, L.; STURION, LHC. **O ensino de Estatística por meio das potencialidades do Software RStudio**. São Paulo: Científica do Brasil, 2021.

COHEN, L.; MANION, L.; MORRISON, K. **Research Methods in Education**. 7th ed. London: Routledge, 2011.

FERNANDES, J. A.; SOUSA, M. V.; RIBEIRO, S. A. O ensino de Estatística no ensino básico e secundário: Um estudo exploratório. *In*: ENCONTRO DE PROBABILIDADES E

ESTATÍSTICA NA ESCOLA, 1., 2004, Braga. **Atas...** Braga: Centro de Investigação em Educação da Universidade do Minho, 2004. p. 165–193.

IHAKA, Ross; GENTLEMAN, Robert. R: a language for data analysis and graphics. **Journal of computational and graphical statistics**, [S. l.], v. 5, n. 3, p. 299–314, 1996.

PACHECO, José Adson Demétrio; BARROS, Janaina Viana. O uso de softwares educativos no ensino de Matemática. **Revista Diálogos**, [S. l.], v. 8, p. 5–13, 2013.