

Um estudo sobre a aceitação do uso da linguagem R em disciplinas de Matemática e Estatística

Laura V. Malachias¹, Gilberto R. Liska^{2†}, Josiane Rodrigues²

¹Centro de Ciências Agrárias (CCA), Universidade Federal de São Carlos (UFSCar).

²Departamento de Tecnologia Agroindustrial e Socioeconomia Rural (DTAiSeR), Centro de Ciências Agrárias (CCA), Universidade Federal de São Carlos (UFSCar).

Resumo: O R é uma linguagem de programação de código aberto que permite, além de procedimentos estatísticos, operações matemáticas simples, manipulação de vetores e matrizes, dentre outros diversos procedimentos. Devido a essas características, o R tem sido utilizado como apoio em diversas disciplinas de Matemática e Estatística do Centro de Ciências Agrárias da Universidade Federal de São Carlos, dentre elas as de “Fundamentos e Programação de Computadores”, “Tópicos em Matemática II” e “Noções de Probabilidade e Estatística”, as quais são ofertadas para os cursos de Biotecnologia e Engenharia Agrônômica do Centro. O intuito de utilizar o R nas disciplinas é facilitar a compreensão dos conteúdos abordados, sendo que o objetivo do trabalho foi avaliar a aceitação dos alunos das disciplinas supracitadas sobre o uso da linguagem R durante as aulas. O trabalho utilizou métodos descritivos exploratórios com o auxílio do R para analisar a aceitação dos estudantes na utilização dessa linguagem, por meio da análise das respostas dadas a uma pesquisa de satisfação realizada após a conclusão de cada disciplina. Os índices de aprovação do uso do R nas aulas foram de 100% em Fundamentos e Programação e Computadores, 78% em Tópicos em Matemática II, e 95% em Noções de Probabilidade e Estatística, o que mostra que para essas disciplinas o uso do R concomitantemente aos conteúdos se torna um recurso promissor para auxiliar no processo de ensino-aprendizagem.

Palavras-chave: Ensino de estatística; Ensino de matemática; Linguagem de programação; Estatística aplicada.

Abstract: R language is an open-source programming that allows, in addition to statistical procedures, simple mathematical operations, manipulation of vectors and matrices, among other procedures. Due to these characteristics, R has been used as support to disciplines of Mathematics and Statistics from Centro de Ciências Agrárias of Universidade Federal de São Carlos, such in “Computer Basics and Coding”, “Mathematical Topics II” and “Probability and Statistical Notions”, offered for the courses of Biotechnology and Agronomic Engineering in Center. The intention of using R is that students understand better the subjects taught in classes. In this way, the aim of this work was to evaluate the acceptance of the students of the mentioned disciplines about the use of R programming language during the classes. The research utilized exploratory descriptive methods with the help of R to analyze acceptance of the students about the use of this language, by analyzing the answers of a satisfaction questionnaire that the students answered after the conclusion of each discipline. The approval rates of R usage among classes were 100% in Computer Basics and Coding; 78% in Mathematical Topics II, and 95% in Probability and Statistical Notions, which shows that R utilization, for these disciplines, makes a reliable and promising resource to help the learning and teaching process.

Keywords: Statistical education; Mathematical education; Programming language; Applied statistics.

Introdução

A presença de tecnologias digitais de informação na sociedade atual tem impulsionado mudanças nos processos educacionais das instituições de ensino (VENTURA; GOMES, 2021). Dentro deste contexto, as atividades em sala de aula estão sendo modificadas, a fim de dinamizar o processo de ensino-aprendizagem. Isso vem acontecendo nos mais diversos níveis de ensino, desde a educação básica até o ensino superior. Dessa forma, o aluno da atualidade precisa de um ambiente no qual o professor faça uso de metodologias ativas, utilizando recursos proporcionados pelas tecnologias digitais, com o objetivo de proporcionar motivação e facilitar o aprendizado (SILVA et al., 2018), além de formar profissionais que atuarão de forma dinâmica no mercado de trabalho.

† Autor correspondente: gilbertoliska@ufscar.br.

No que diz respeito ao ensino da Matemática e da Estatística no ensino superior, sabe-se que se trata de áreas de grande importância, visto que são capazes, dentre outras coisas, de desenvolver o senso crítico dos alunos. Entretanto, apesar de sua importância, verifica-se que os alunos apresentam algumas dificuldades no seu aprendizado. Assim sendo, existe a necessidade de investigar quais ferramentas facilitam o processo educativo, bem como permitem aos alunos aprendizagens mais significativas (ROSA; RODRIGUES; SILVA, 2011).

No contexto do ensino dessas disciplinas a nível de graduação com o auxílio de tecnologias digitais, o R é uma linguagem de programação de código aberto que pode ser implementada em sala de aula. A linguagem surgiu na década de 1990 e originalmente foi desenvolvida por Ross Ihaka e Robert Gentleman, pesquisadores do Departamento de Estatística da Universidade de Auckland, na Nova Zelândia (PASSOS, 2021). A linguagem permite, além de procedimentos estatísticos, operações matemáticas simples, manipulação de vetores e matrizes, dentre outros diversos procedimentos (CRAWLEY, 2012). O software é muito utilizado nas mais diversas áreas da ciência, e tem como uma das maiores qualidades a grande quantidade de pacotes disponíveis e que possibilitam manipular dados de forma eficaz, assim como armazená-los (DALL'AGNOL; BETZEK; BAZZI, 2018).

O uso do software em sala de aula possui uma série de vantagens: é uma fonte de informação, auxilia no processo de construção do conhecimento, amplia a autonomia de raciocínio e de reflexão dos alunos, bem como os ampara no processo de criação de soluções na resolução de problemas (GLADCHEFF; SILVA; ZUFFI, 2001). Segundo Barros e Pacheco (2013), o uso do software em aula amplia os conceitos aprendidos, uma vez que faz com que eles sejam compreendidos de forma interativa e intuitiva.

Diante do exposto, o presente trabalho teve como objetivo avaliar a aceitação dos alunos ao uso do software R nas disciplinas de “Tópicos em Matemática II (TMII)”, “Fundamentos e Programação de Computadores (FPC)” e “Noções de Probabilidade e Estatística (NPE)”, as quais são ministradas para os cursos de Biotecnologia e Engenharia Agrônoma do Centro de Ciências Agrárias (CCA) da Universidade Federal de São Carlos (UFSCar).

Em TMII, os temas abordados nas aulas com o uso do R foram: derivadas e anti derivadas; integrais indefinidas e definidas; funções de duas ou mais variáveis, e equações diferenciais de 1ª ordem e suas aplicações. Já em FPC, o R foi utilizado na construção de algoritmos utilizando estruturas condicionais e de repetições, e na elaboração de matrizes, tabelas e gráficos. Por fim, em NPE, o R foi utilizado no estudo da amostragem, onde foram trabalhadas técnicas para a obtenção de amostras representativas da população, na análise exploratória de dados, em probabilidade e em inferência estatística.

Com o trabalho, pretende-se compreender a percepção dos alunos sobre o uso da linguagem aberta R em disciplinas de Matemática e Estatística, o que, de modo geral, auxiliará os professores da área que já utilizam do recurso em suas aulas a melhorarem suas práticas didáticas, assim também como estimulará docentes que não a utilizam a se sentirem encorajados a adotar o recurso em suas práticas em sala de aula.

Materiais e Métodos

O software R e o RStudio

Desde o seu desenvolvimento, a linguagem de programação R tem sido mantida por um vasto grupo de colaboradores voluntários que disponibilizam toda a documentação do avanço e aprimoramento da linguagem em uma plataforma na internet. Contudo, a interface do R não é muito amigável para iniciantes, de modo que um grande progresso na utilização da linguagem foi a criação da IDE (*Integrated Development Environment*) RStudio, a qual nada mais é do que um ambiente integrado com o R, que possui inúmeras linguagens de marcação de texto, facilitando assim o aprendizado dos alunos, uma vez que permite uma melhor experiência entre hardware e software.

Assim sendo, o R em cada uma das disciplinas de TMII, FPC e NPE, foi utilizado em conjunto com a IDE RStudio (RSTUDIO TEAM, 2020), a fim de facilitar a interação dos alunos com a linguagem.

Questionário de Satisfação

Os dados da pesquisa foram obtidos por meio de questionário de satisfação aplicado aos alunos após a finalização de cada uma das disciplinas de TMII, FPC e NPE. O questionário continha dez perguntas

sobre a qualidade das aulas, materiais e sobre o software utilizado. Considerando que completar o questionário de satisfação não era uma etapa obrigatória para a finalização da disciplina, muitos discentes deixaram de respondê-lo, de modo que a taxa de resposta ao questionário foi de no máximo 55,9% por turma. Além disso, vale ressaltar que as aulas, bem como os questionários, aconteceram durante o período da pandemia de COVID-19, nos anos de 2020 e 2021, quando do Ensino Não Presencial Emergencial (ENPE) da UFSCar, em que as aulas da Instituição, devido à pandemia, ocorreram por meio virtuais.

A pergunta chave do questionário para o presente estudo foi: “*Em uma escala de 0 a 10, que nota você atribui ao uso do R nas aulas? Atribua 0 para muito insatisfeito e 10 para muito satisfeito*”. De modo a facilitar as análises, foram atribuídas categorias às respostas, da seguinte forma: 0 - plenamente insatisfeito; 1 a 5 - moderadamente insatisfeito; 6 a 9 - moderadamente satisfeito e 10 - plenamente satisfeito.

Como os dados são quantitativos discretos, optou-se, após a síntese das respostas, à criação de gráficos de barras para melhor visualização dos dados. Além disso, foram calculadas medidas descritivas relativas ao conjunto, bem como os limites do intervalo de confiança de 95% para a média. Para realizar a análise dos resultados foi utilizada a versão 4.1.2 do R (R CORE TEAM, 2021) e pacotes estatísticos como *asbio* (AHO, 2022) e *fBasics* (WUERTZ et al., 2022).

Resultados e Discussão

Dos alunos que responderam à pesquisa de satisfação na disciplina de TMII, 22% mostraram-se plenamente satisfeitos quanto ao uso do R nas aulas, enquanto 56% mostraram-se moderadamente satisfeitos. Já na disciplina de FPC, o percentual de alunos plenamente satisfeitos foi de 43%, enquanto a porcentagem de alunos moderadamente satisfeitos foi de 57%. Por fim, na disciplina de NPE, o percentual de alunos plenamente satisfeitos foi de 32%, enquanto os alunos moderadamente satisfeitos foram 63%. Nota-se, portanto, a partir dos resultados encontrados, que a maioria dos alunos aprova o uso da linguagem em conjunto com as disciplinas.

Tal aspecto reflete que o aluno do ensino superior, cada vez mais, é a favor do uso de metodologias diferenciadas em sala de aula, as quais vão além dos artefatos convencionais utilizados nas formações de nível superior (BLASS; GOERGEN; IRALA, 2020). É válido ressaltar que as metodologias tradicionais são de extrema importância no processo de ensino-aprendizagem. Entretanto, a utilização de metodologias ativas nas aulas, tal qual a utilização de softwares, por exemplo, auxilia o aluno e o professor no processo de construção do conhecimento, amplia a autonomia do raciocínio dos alunos, e auxilia na criação de soluções, dentre outros diversos benefícios (GLADCHEFF; SILVA; ZUFFI, 2001).

Observando as medidas descritivas dos dados obtidos com a pesquisa (Tabela 2), verifica-se que a disciplina com menor aprovação dos alunos no que diz respeito ao uso do R nas aulas foi a de TMII. O mesmo pode ser notado através do gráfico de barras que apresenta as frequências percentuais, por disciplina, das notas atribuídas pelos alunos à pergunta (Figura 1). A nota média atribuída pelos discentes em TMII foi de 7,22, com desvio padrão de 2,74. Além disso, o limite inferior do intervalo de confiança de 95% para a média foi de 6,19, sendo o limite superior igual a 8,25. Partindo do princípio de que os respondentes constituem uma amostra dos alunos da referida disciplina e, sob a suposição de normalidade, com 95% de confiança, pode-se afirmar que a nota média está contida no intervalo entre 6,19 e 8,25.

Com relação à FPC, entretanto, observa-se que esta disciplina foi a que obteve maior aprovação dos alunos com relação ao uso do R nas aulas (Tabela 2 e Figura 1). A nota média para a pergunta foi equivalente a 9,22, com desvio padrão de 0,90. Por sua vez, o limite inferior do intervalo de 95% de confiança para a média foi de 8,88, sendo o limite superior igual a 9,56 pontos.

Tabela 1. Tabela de frequências das notas dos respondentes da pergunta: “Em uma escala de 0 a 10, que nota você atribui ao uso do R nas aulas? Atribua 0 para muito insatisfeito e 10 para muito satisfeito”. TMII - Tópicos em Matemática II; FPC - Fundamentos e Programação de Computadores; NPE - Noções de Probabilidade e Estatística.

Nota	TMII				FPC				NPE			
	Frequência		Frequência Acumulada		Frequência		Frequência Acumulada		Frequência		Frequência Acumulada	
	f_i	fr_i	F_i	Fr_i	f_i	fr_i	F_i	Fr_i	f_i	fr_i	F_i	Fr_i
0	1	0,037	1	0,037	0	0,000	0	0,000	1	0,053	1	0,053
1	0	0,000	1	0,037	0	0,000	0	0,000	0	0,000	1	0,053
2	2	0,074	3	0,111	0	0,000	0	0,000	0	0,000	1	0,053
3	0	0,000	3	0,111	0	0,000	0	0,000	0	0,000	1	0,053
4	1	0,037	4	0,148	0	0,000	0	0,000	0	0,000	1	0,053
5	2	0,074	6	0,222	0	0,000	0	0,000	0	0,000	1	0,053
6	2	0,074	8	0,296	0	0,000	0	0,000	2	0,105	3	0,158
7	4	0,148	12	0,444	1	0,043	1	0,043	2	0,105	5	0,263
8	4	0,148	16	0,593	4	0,174	5	0,217	5	0,263	10	0,526
9	5	0,185	21	0,778	8	0,348	13	0,565	3	0,158	13	0,684
10	6	0,222	27	1,000	10	0,435	23	1,000	6	0,316	19	1,000
Total	27	1			23	1			19	1		

Frequência absoluta (f_i); Frequência relativa (fr_i); Frequência acumulada absoluta (F_i); Frequência acumulada relativa (Fr_i).

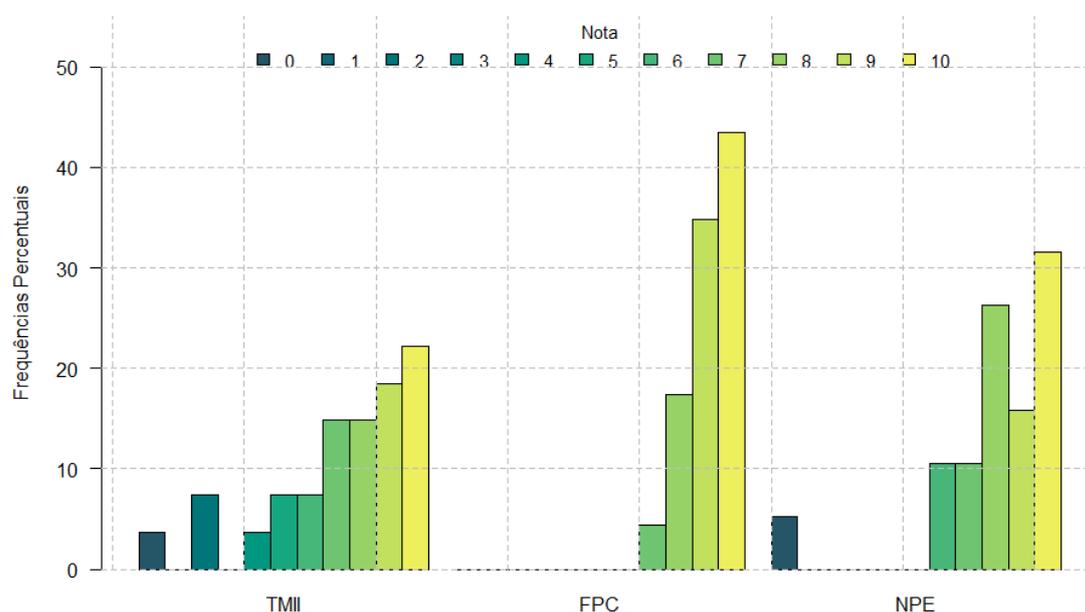
Fonte: Dos autores.

Tabela 2. Medidas descritivas das notas dos respondentes da pergunta: “Em uma escala de 0 a 10, que nota você atribui ao uso do R nas aulas? Atribua 0 para muito insatisfeito e 10 para muito satisfeito”. TMII - Tópicos em Matemática II; FPC - Fundamentos e Programação de Computadores; NPE - Noções de Probabilidade e Estatística.

Disciplina	Média	Desvio-padrão	Intervalo de Confiança (95%)	
			Limite inferior	Limite superior
TMII	7,22	2,74	6,19	8,25
FPC	9,22	0,90	8,88	9,56
NPE	8,05	2,36	7,16	8,95

Fonte: Dos autores.

Gráfico 1. Gráfico de barras das frequências percentuais dos respondentes da pergunta: “Em uma escala de 0 a 10, que nota você atribui ao uso do R nas aulas? Atribua 0 para muito insatisfeito e 10 para muito satisfeito”. TMII - Tópicos em Matemática II; FPC - Fundamentos e Programação de Computadores; NPE - Noções de Probabilidade e Estatística.



Fonte: Dos autores.

Por fim, a disciplina de NPE ficou em situação intermediária (Tabela 2 e Figura 1), sendo a nota média atribuída pelos alunos à pergunta equivalente a 8,05, com desvio padrão de 2,36 pontos. O limite inferior do intervalo de confiança de 95% para a média foi de 7,16 pontos, e o limite superior foi de 8,95.

Quanto aos alunos que se mostraram insatisfeitos quanto ao uso da linguagem R nas aulas, 22% mostraram-se parcialmente ou totalmente insatisfeitos na disciplina de TMII, sendo este percentual nulo para FPC, e de apenas 5% para NPE. A maior insatisfação sobre o uso do R em TMII pode ser justificada pelo fato dela tratar do Cálculo Diferencial e Integral. Apesar de ser de extrema importância, há diversas pesquisas sobre o ensino de Cálculo que apontam o baixo desempenho dos alunos na disciplina (ALMEIDA, 2016) e, portanto, é possível supor que as notas de insatisfação sejam devidas, dentre outros fatores, à dificuldade em compreender conceitos inerentes ao Cálculo. É válido ressaltar, entretanto, que o uso de recursos tecnológicos pode ser uma importante ferramenta no ensino da disciplina.

De modo geral, é possível afirmar que as estratégias didáticas adotadas com o uso do software R em cada uma das disciplinas de TMII, FPC e NPE revelaram-se eficientes para impulsionar uma visão favorável dos alunos quanto às aulas. Por sua vez, a criação de um ambiente favorável em sala de aula acaba se tornando um facilitador do processo de ensino-aprendizagem. De forma similar, outras pesquisas realizadas para analisar o uso de softwares em atividades de ensino, dentre eles o R, mostraram que a utilização do recurso auxilia no processo de construção de conhecimento e tornam as aulas mais dinâmicas e participativas (CARDOSO, 2019).

Segundo Giorgi et al. (2022), a linguagem R está entre as mais usadas no mundo, ocupando a 7ª linguagem mais popular de acordo com o índice PYPL (CARBONNELLE, 2023). Além disso, em dezembro de 2021, o R foi classificado como a 8ª linguagem mais útil em termos de número de empregos que a exigem (EASTWOOD, 2023). Frente a essas informações, e considerando a aprovação dos alunos quanto ao uso do R nas aulas, bem como a importância do uso de tecnologias no processo de ensino-aprendizagem, isso representa um estímulo para que docentes continuem com o uso da linguagem R em sala de aula, de forma particular aqui nas disciplinas das áreas de Matemática e Estatística.

Conclusão

Com base na pesquisa de satisfação realizada, verificou-se que o percentual de aprovação dos alunos quanto ao uso da linguagem R nas aulas das disciplinas de TMII, FPC e NPE foi de 78%, 100% e 95%, respectivamente, o que mostra que para essas disciplinas o uso do R concomitantemente aos conteúdos é um recurso viável e promissor para auxiliar no processo de ensino-aprendizagem.

Agradecimentos

À Pró-Reitoria de Extensão da Universidade Federal de São Carlos (ProEx) pela bolsa de extensão concedida à aluna Laura Vieira Malachias.

Referências

AHO, K. *asbio*: A Collection of Statistical Tools for Biologists. R package version 1.8-3. 2022. Disponível em: <https://CRAN.R-project.org/package=asbio>.

ALMEIDA, W.Q. *Dificuldades dos alunos no aprendizado de Cálculo Diferencial e Integral I*: uma reflexão. Monografia (Especialização em Matemática para professores) – Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

BARROS, J.V.; PACHECO, J.A.D. O uso de softwares educativos no ensino de matemática. *Revista de Estudos Culturais e da Contemporaneidade*, n.8, p.5-13, 2013.

BLASS, L.; GOERGEN, G.; IRALA, V.B. Análise das percepções e desempenhos discentes com a utilização do software R em uma disciplina de graduação de natureza multicurso. *Revista Prática Docente*, v.5, n.3, p.1568-1592, 2020.

CARBONNELLE, P. *PYPL - Popularity of Programming Language Index*. Disponível em: <https://pypl.github.io/PYPL.html>. Acesso em: 19 jan. 2023.

CARDOSO, M.G. *Ensino de Estatística: o estudo de conceitos potencializado pelo software Rstudio*. 2019. Dissertação (Mestrado) - Curso de Matemática, Universidade Tecnológica Federal do Paraná, Londrina, 2019.

CRAWLEY, M.J. *The R book*. 2.ed. Wiley, 2012. 1080p.

DALL'AGNOL, R.W.; BETZEK, N.M.; BAZZI, C.L. Aplicação do software R para análise geoestatística, interpolação de dados por krigagem ordinária e geração de mapas temáticos de produtividade agrícola. *Revista de Computação aplicada ao Agronegócio*, v.1, n.1, p.24-37, 2018.

EASTWOOD, B. *The 10 Most Popular Programming Languages to Learn in 2021*. Disponível em: <https://www.northeastern.edu/graduate/blog/most-popular-programming-languages/>. Acesso em: 19 jan. 2023.

GIORGI, F.M.; CERAIOLO, C.; MERCATELLI, D. The R Language: An Engine for Bioinformatics and Data Science. *Life*, v.12, n.5, p.648, 2022.

GLADCHEFF, A.P.; SILVA, D.M.; ZUFFI, E.M. Um instrumento para avaliação da qualidade de softwares educacionais de matemática para o ensino fundamental. In: *Congresso da Sociedade Brasileira de Computação*, 2001. Fortaleza. Anais [...] Fortaleza: VII Workshop de Informática na Escola, 2001. p.1-12.

PASSOS, I.C. Análise de dados qualitativos com o R: uma introdução ao pacote RQDA. *Inter-Legere*, v.4, n.30, p.1-19, 2021.

R CORE TEAM (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <https://www.R-project.org>.

R STUDIO TEAM (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. Disponível em: <http://www.rstudio.com/>.

ROSA, O.; RODRIGUES, C.; SILVA, P. Aspectos Motivacionais na Disciplina de Cálculo Diferencial e Integral. *Revista Eletrônica TECCEN*, v.4, n.2, p.49-62, 2011.

SILVA, J.B.; ANDRADE, M.H.; OLIVEIRA, R.R.; SALES, G.L.; ALVES, F.R.V. Tecnologias digitais e metodologias ativas na escola: o contributo do Kahoot para gamificar a sala de aula. *Revista Thema*, v.15, n.2, p.780-791, 2018.

VENTURA, J.P.C.; GOMES, C.R. Softwares no ensino de matemática: um olhar sobre a BNCC. *Boletim Cearense de Educação e História da Matemática*, v.8, n.23, p.846-860, 2021.

WUERTZ, D.; TOBIAS, S.; YOHAN, C.; MARTIN, M. *fBasics: Rmetrics - Markets and Basic Statistics*. Rpackage version 4021.93. 2022. Disponível em: <https://cran.r-project.org/web/packages/fBasics/index.html>.