ISSN: 2358-8829

UM RETRATO DA EVASÃO ACADÊMICA EM CURSOS DE LICENCIATURA DA UFCG

Izis Maria Medeiros da Costa [1]
Lucas Silva Ferreira De Carvalho [2]
Kleyton Klaus Guedes de Souza [3]
Vladimir Soares Catão [4]

INTRODUÇÃO

A ampliação do acesso ao ensino superior no Brasil nas últimas décadas (ANDIFES, 2019; INEP, 2024) não foi acompanhada, na mesma proporção, da garantia de que os alunos permaneçam em seus cursos até a conclusão (Schwarz; Dias; Camargo, 2021). Nesse contexto, a evasão, definida pelo abandono do curso, por solicitação do próprio aluno ou por desligamento previsto em regulamento, está entre os principais desafios das Instituições de Ensino Superior (IES). Nas instituições públicas, o problema é mais grave, pois, além dos prejuízos pessoais e familiares, o investimento realizado durante o período em que o estudante esteve matriculado deixa de gerar retorno à sociedade (OCDE, 2025).Nas Licenciaturas, a situação é particularmente preocupante. O Censo da Educação Superior de 2023 registra taxas de desistência acumulada de 58% em Biologia, 66% em Química, 70% em Matemática e 73% em Física (INEP, 2024), resultados semelhantes aos observados na Universidade Federal do Rio Grande do Norte (Carvalho et al., 2019). Esse quadro agrava a carência de professores nos anos finais do Ensino Fundamental e no Ensino Médio, configurando um "apagão" docente (Bof; Caseiro; Mundim, 2023): cada abandono na Licenciatura representa um professor a menos para a educação básica.

A evasão é um fenômeno multidimensional, que envolve fatores internos e externos à universidade (Araújo; Mariano; Oliveira, 2021), e costuma ser precedida pelo

[1] Graduanda do Curso de Licenciatura em Matemática da Universidade Federal de Campina Grande – UFCG – Campus Cuité , izis.maria@estudante.ufcg.edu.br;
[2] Graduando pelo Curso de Licenciatura em Física da Universidade Federal de Campina Grande – UFCG – Campus Cuité, ferreira.carvalho@estudante.ufcg.edu.br;
[3] Mestrando do Programa de Pós-Graduação em Ciências Naturais e Biotecnologia da Universidade Federal de Campina Grande - UFCG, kleyton.klaus@tecnico.ufcg.edu.br;
[4] Doutor pelo Curso de Ciência da Computação, Professor de Informática da Universidade Federal de Campina Grande – UFCG - Campus Cuité, vladimir.soares@professor.ufcg.edu.br.

acúmulo de reprovações e trancamentos (MEC, 2007). Políticas de permanência, como o Programa Nacional de Assistência Estudantil – PNAES (Brasil, 2010) e o Pé-de-Meia Licenciaturas (MEC, 2025), dependem, para ser efetivas, da capacidade de identificar o mais cedo possível os alunos em risco. Embora o histórico acadêmico e as características demográficas estejam entre as variáveis mais utilizadas nos estudos sobre evasão (Tete et al., 2022) e Sales, Balby e Cajueiro (2016) tenham demonstrado o potencial dos registros acadêmicos na própria Universidade Federal de Campina Grande (UFCG), estudos voltados especificamente às Licenciaturas ainda são escassos.

Diante desse contexto, este estudo[5] tem como objetivo identificar os fatores mais associados à evasão nos cursos presenciais de Licenciatura em Biologia, Física, Matemática e Química da UFCG. Especificamente, busca-se: (I) caracterizar o perfil de evadidos e concluintes; (II) descrever em que momento do percurso acadêmico a evasão se concentra; e (III) estimar a associação entre indicadores acadêmicos e demográficos dos três primeiros períodos letivos e a ocorrência de evasão. Trata-se de um estudo observacional, retrospectivo e quantitativo, com 7.745 estudantes com desfecho definido (73,1% evadidos e 26,9% concluintes), ingressantes entre 2010.1 e 2025.2, cujos dados foram obtidos por meio da API do sistema institucional Eureca[6]; os dados são anonimizados, com acesso aberto, sem a necessidade de submissão a Comitê de Ética em Pesquisa com seres humanos. As análises compreenderam estatística descritiva, diagnóstico de multicolinearidade e regressão logística multivariada (James et al., 2023). Os resultados mostram que 56,4% dos evadidos abandonaram o curso até o 3º semestre e que a taxa de sucesso e os créditos do CRA (créditos acumulados que entram no cálculo do Coeficiente de Rendimento Acadêmico) foram fatores protetores independentes ($p < 0,001$); o sexo masculino e o turno noturno foram fatores de risco. Esses indicadores, disponíveis já nos primeiros semestres, podem fundamentar modelos de alerta precoce para alunos em risco e orientar ações de permanência.

METODOLOGIA

[5] Vinculado ao projeto de Iniciação Científica (PIVIC) da UFCG intitulado "Aprendizado de Máquina na Predição da Evasão em Cursos de Licenciatura", código PVES1488-2025, vigência 2025-2026.
[6] https://eureca.sti.ufcg.edu.br/

ISSN: 2358-8829

Trata-se de um estudo observacional, retrospectivo, com levantamento documental e abordagem quantitativa. A população compreende os ingressantes nos cursos presenciais de Licenciatura em Ciências Biológicas, Física, Matemática e Química de todos os campi da UFCG entre os semestres 2010.1 e 2025.2. Foram incluídos apenas os alunos com situação final definida (evadidos ou concluintes), excluindo-se os ainda ativos, cujo desfecho não é conhecido. A amostra final foi de 7.745 estudantes: 5.664 (73,1%) evadidos e 2.081 (26,9%) concluintes.

Considerou-se evadido o estudante cujo vínculo com o curso foi encerrado sem a conclusão, por solicitação própria ou por desligamento previsto em regulamento (cancelamento por abandono, reprovação por falta, transferência). A variável dependente foi dicotômica: 1 para evasão e 0 para conclusão. Foram classificados como evadidos os alunos com situação cancelada (5.634) ou trancada (30) e como concluintes os com situação concluída (2.014) ou formada (67), conforme registrado na coleta dos dados; os 30 alunos com matrícula trancada foram contados como evadidos.

Os dados foram obtidos por meio da API do sistema institucional Eureca (https://eureca.sti.ufcg.edu.br/), mantido pela Superintendência de Tecnologia da Informação (STI) da UFCG com informações da Pró-Reitoria de Ensino (PRE/CCA), por meio de um script em Python. Foram coletados apenas dados anonimizados.

As variáveis independentes foram:

- Acadêmicas: taxa de sucesso (razão entre créditos completados e tentados), créditos do CRA (créditos acumulados considerados no cálculo do Coeficiente de Rendimento Acadêmico, denominados “CRA” no resumo), créditos tentados, falhados, suspensos e completados –, calculadas ao final de cada um dos três primeiros períodos letivos cursados (P1, P2 e P3);
- demográficas e institucionais: sexo, idade, turno, forma de ingresso e campus;
- socioeconômicas: tipo de escola no ensino médio, cor/raça e renda.

Variáveis com registros insuficientes foram excluídas da modelagem, como as métricas próprias do sistema SIGAA, implantado apenas em 2024.

A análise foi realizada em três etapas. Na primeira, a estatística descritiva caracterizou o perfil de evadidos e concluintes e a distribuição das evasões ao longo dos semestres; com isso verificou-se a maior concentração nos 3 primeiros períodos de

ISSN: 2358-8829

curso. Na segunda, diagnosticou-se a multicolinearidade pelo coeficiente de correlação de Spearman (Spearman, 1904), medida baseada em postos que não pressupõe distribuição normal (Schober; Boer; Schwarte, 2018), e pelo Fator de Inflação da Variância (VIF) (O'Brien, 2007). Nos pares com correlação elevada ($|\rho| > 0,7$) e nas variáveis com VIF acima de 5, removeu-se a variável redundante, mantendo-se as de maior interpretabilidade para a gestão acadêmica.

Na terceira etapa, foi ajustado um modelo de regressão logística (James et al., 2023) para cada um dos três primeiros períodos (P1, P2 e P3), tendo como desfecho a evasão (1) em relação à conclusão (0) e como preditores a taxa de sucesso e os créditos do CRA do próprio período. As variáveis contínuas foram padronizadas, de modo que o OR corresponde ao aumento de um desvio-padrão, e as binárias (sexo e turno) entram com OR por unidade. Como a mesma pessoa pode ter mais de um vínculo (reingresso ou troca de curso), os erros-padrão também foram calculados de forma robusta, agrupados por pessoa. A associação de cada variável com o desfecho foi expressa pela razão de chances, Odds Ratio – OR (James et al., 2023), com intervalo de confiança de 95%: OR menor que 1 indica efeito protetor e OR maior que 1, fator de risco. Adotou-se o nível de significância de 5%.

As análises foram implementadas em linguagem Python (bibliotecas pandas, numpy, scipy, statsmodels e openpyxl), com scripts desenvolvidos com apoio da inteligência artificial generativa Claude Sonnet 4.6, da Anthropic[7]. Todo o código foi revisado e validado pelos pesquisadores, que se responsabilizam pelas análises e pelos resultados apresentados.

REFERENCIAL TEÓRICO

Uma das principais referências para a definição da evasão no Brasil é o relatório da Comissão Especial de Estudos sobre a Evasão nas Universidades Públicas Brasileiras (Brasil, 1996), que a define como a saída definitiva do estudante de seu curso de origem sem concluí-lo e distingue a evasão de curso, da instituição e do sistema. Adota-se aqui a evasão de curso, por ser a que mais diretamente afeta a formação de professores pela instituição.

[7] https://claude.ai

ISSN: 2358-8829

Do ponto de vista teórico, um dos modelos mais influentes é o de Tinto (1975), que compreende a evasão como um processo longitudinal, resultante da interação entre o estudante e a instituição. Os atributos que o estudante traz consigo, como características individuais, contexto familiar e escolarização anterior, influenciam seu comprometimento inicial com a formação, que é reforçado ou enfraquecido ao longo do curso pelo grau de integração acadêmica e social alcançado. A integração acadêmica, expressa sobretudo pelo desempenho nas disciplinas, ocupa papel central: quanto menor a integração, maior a probabilidade de abandono. Essa perspectiva fundamenta o uso da aprovação em disciplinas e dos créditos acumulados como manifestações observáveis do processo que pode levar à evasão.

Os estudos brasileiros reforçam que esse processo se manifesta cedo: a evasão no primeiro ano é consideravelmente maior que nos seguintes, e poucas instituições dispõem de programas sistemáticos de combate ao problema (Silva Filho et al., 2007). Reprovações e trancamentos sucessivos prolongam a permanência e, com frequência, antecedem o abandono (MEC, 2007; Araújo; Mariano; Oliveira, 2021), de modo que as dificuldades acadêmicas dos primeiros semestres podem ser compreendidas também como sinal antecipado do risco de evasão.

Nas Licenciaturas, muitos cursos são noturnos e recebem alunos que precisam conciliar estudo e trabalho, um dos principais motivos apontados para se pensar em abandonar o curso (ANDIFES, 2019), e a baixa atratividade da carreira docente contribui para que parte dos licenciados não siga a profissão (Bof; Caseiro; Mundim, 2023).

Nos estudos sobre evasão, os indicadores utilizados se distribuem em dimensões socioeconômicas, acadêmicas e, em menor número, psicológicas e de saúde (Tete et al., 2022), sendo o histórico acadêmico e as características demográficas os mais empregados, embora nenhum conjunto de variáveis explique o fenômeno em todos os contextos (Silva; Pimentel; Botelho, 2022). Com dados brasileiros, os registros acadêmicos por si só já se mostram informativos: na UFCG, Sales, Balby e Cajueiro (2016) investigaram a evasão exclusivamente a partir do histórico acadêmico. Esses indicadores têm ainda a vantagem de ser registrados rotineiramente, sem custos adicionais de coleta.

Para subsidiar decisões institucionais, é desejável que os modelos sejam interpretáveis. A regressão logística atende a esse requisito, pois expressa a associação

ISSN: 2358-8829

de cada variável com o desfecho pela razão de chances, ajustada pelas demais variáveis (Hosmer; Lemeshow; Sturdivant, 2013). Essa interpretação, contudo, depende da ausência de multicolinearidade elevada, que torna os coeficientes instáveis e infla seus erros-padrão (James et al., 2023; O'Brien, 2007), o que justifica o diagnóstico prévio para identificar os indicadores que contribuem de forma independente para explicar a evasão.

RESULTADOS E DISCUSSÃO

Dos 7.745 estudantes analisados, 5.664 (73,1%) evadiram e 2.081 (26,9%) concluíram o curso: para cada concluinte, quase três abandonaram. As taxas de evasão variam entre as áreas, de 62,1% em Ciências Biológicas a 87,6% em Física (Tabela 1), na mesma ordem das taxas de desistência acumulada registradas nacionalmente, de 58% a 73% (INEP, 2024), embora mais altas, o que pode refletir diferenças de definição, já que aqui só entram os estudantes com desfecho definido. A evasão foi maior no turno noturno (78,9%, contra 67,4% no diurno) e entre os homens (77,7%, contra 69,8% entre as mulheres).

Tabela 1 – Taxa de evasão por área, turno e sexo

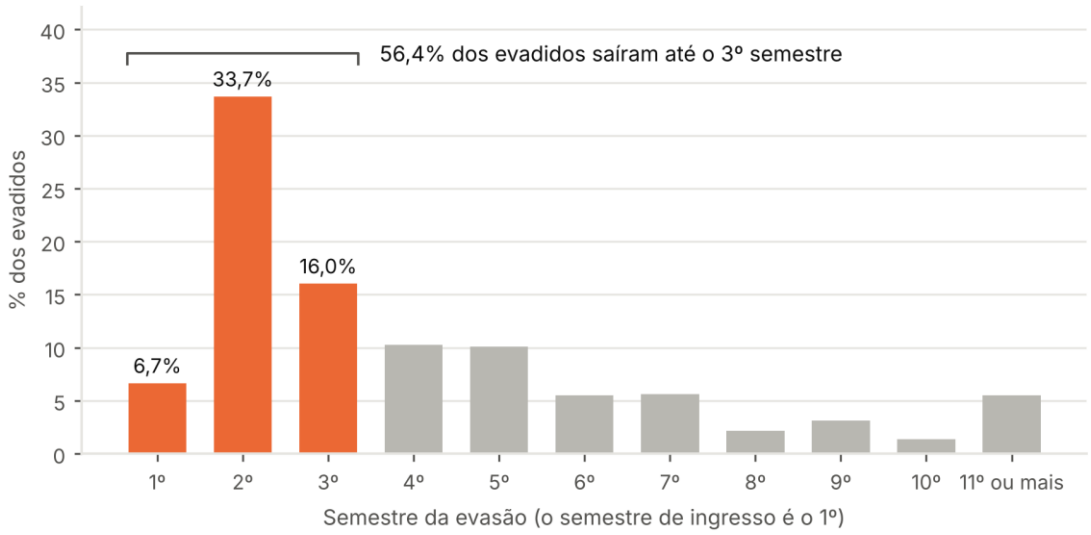
Grupo	Categoria	Alunos	Evadidos	% evadidos
Área	Ciências Biológicas	2.733	1.696	62,1
Área	Física	1.545	1.354	87,6
Área	Matemática	2.089	1.580	75,6
Área	Química	1.378	1.034	75,0
Turno	Diurno	3.895	2.627	67,4
Turno	Noturno	3.850	3.037	78,9
Sexo	Feminino	3.654	2.551	69,8
Sexo	Masculino	3.914	3.041	77,7
Sexo	N	177	72	40,7
Total	**Todos os alunos**	**7.745**	**5.664**	**73,1**

Fonte: elaborada pelos autores, com dados do sistema Eureca (UFCG). **Nota:** Evadido: situação cancelada. N: valor contido no cadastro; possivelmente não do cadastro.

A evasão mostrou-se fortemente concentrada no início do percurso: 56,4% dos evadidos abandonaram o curso até o 3º semestre (Figura 1), o que concorda com Silva Filho et al. (2007) e indica que os três primeiros semestres são a janela mais crítica para ações de permanência. O 2º semestre concentra, sozinho, 33,7% das evasões. Além disso, cerca de 25% dos evadidos (26,5%, 1.503 de 5.664) não obtiveram aprovação em nenhuma disciplina no 1º período (taxa de sucesso igual a zero), proporção que chega a

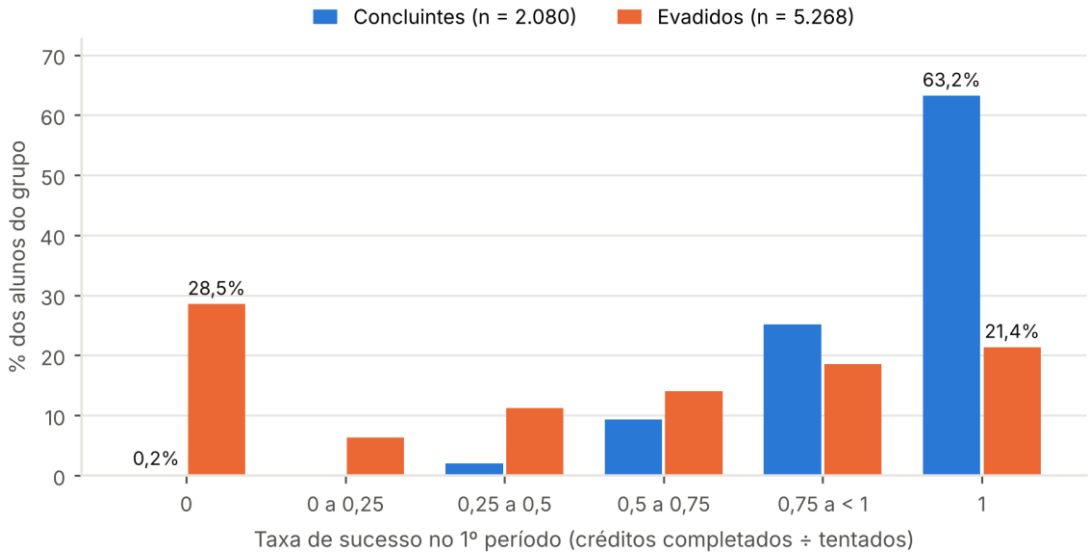
28,5% entre os evadidos com dados do 1º período e contrasta com os 0,2% observados entre os concluintes (Figura 2). Na perspectiva de Tinto (1975), esse quadro revela integração acadêmica muito baixa desde a entrada no curso e evidencia a relação entre retenção e evasão (MEC, 2007; Araújo; Mariano; Oliveira, 2021): reprovações já no primeiro semestre são sinal precoce do risco de evasão.

Figura 1 – Semestre da evasão



Fonte: elaborada pelos autores, com dados do sistema Eureca (UFCG). **Nota:** Percentual dos evadidos com período de evasão válido, por semestre de calendário desde o ingresso (o semestre de ingresso é o 1º).

Figura 2 – Taxa de sucesso no 1º período, por situação final



Fonte: elaborada pelos autores, com dados do sistema Eureca (UFCG). **Nota:** Percentual de cada grupo, entre os alunos com dado no 1º período, em cada faixa da taxa de sucesso; a faixa "0" corresponde a quem não completou nenhum crédito.

ISSN: 2358-8829

No diagnóstico de multicolinearidade, a correlação de Spearman e o VIF indicaram forte sobreposição entre os indicadores de desempenho acadêmico: a taxa de sucesso correlacionou-se de forma negativa e muito forte com os créditos falhados (ρ de −0,93 a −0,96), os créditos do CRA praticamente coincidiram com os créditos tentados (ρ de 0,97 a 0,99) e o VIF foi infinito ou superou 10.000 em cinco dos seis indicadores candidatos, pois as contagens de créditos são combinações lineares umas das outras. Após a remoção das variáveis redundantes, o modelo reteve dois indicadores acadêmicos, independentes entre si (ρ de 0,13 a 0,32; VIF de 1,08 a 1,10; Tabela 2): a taxa de sucesso e os créditos do CRA.

Tabela 2 – Multicolinearidade entre a taxa de sucesso e os créditos do CRA

Período	n	Spearman (taxa × créditos do CRA)	VIF da taxa de sucesso	VIF dos créditos do CRA
P1	7.348	0,13	1,08	1,08
P2	5.483	0,27	1,09	1,09
P3	4.460	0,32	1,10	1,10

Fonte: elaborada pelos autores, com dados do sistema Eureca (UFCG).

Os dois indicadores foram significativos em todos os períodos ($p < 0,001$), e tiveram efeito protetor (Tabela 3). Para a taxa de sucesso, o OR por desvio-padrão foi de 0,145 em P1, 0,135 em P2 e 0,130 em P3 (IC 95% entre 0,113 e 0,165): a cada desvio-padrão a mais na taxa de sucesso (0,39 em P1, 0,27 em P2 e 0,22 em P3), a chance de evasão cai entre 85,5% e 87,0%, mantido o outro indicador. Para os créditos do CRA, cujo desvio-padrão é de 27 a 29 créditos, o OR ficou entre 0,733 e 0,816, uma redução de 18,4% a 26,7% na chance de evasão a cada desvio-padrão. Como o desvio-padrão da taxa de sucesso diminui de P1 a P3, os OR por desvio-padrão são próximos entre os períodos, mas o OR por unidade, que compara quem não é aprovado em nenhum crédito com quem é aprovado em todos, passa de 0,0069 em P1 a 0,0001 em P3.

ISSN: 2358-8829

Tabela 3 – Odds ratios por período: taxa de sucesso e créditos do CRA

Período	n	Variável	OR (IC 95%)	P
P1	7.348	Taxa de sucesso	0,145 (0,128–0,165)	< 0,001
P1	7.348	Créditos do CRA	0,816 (0,775–0,859)	< 0,001
P2	5.483	Taxa de sucesso	0,135 (0,118–0,154)	< 0,001
P2	5.483	Créditos do CRA	0,792 (0,742–0,845)	< 0,001
P3	4.460	Taxa de sucesso	0,130 (0,113–0,149)	< 0,001
P3	4.460	Créditos do CRA	0,733 (0,677–0,794)	< 0,001

Fonte: elaborada pelos autores, com dados do sistema Eureca (UFCG). **Nota:** Regressão logística por período, com a taxa de sucesso e os créditos do CRA do próprio período (OR por desvio-padrão, IC: Intervalo de Confiança).

A significância simultânea dos dois indicadores sugere que captam dimensões distintas: a taxa de sucesso reflete a capacidade de concluir as disciplinas em que o estudante se matricula, e os créditos do CRA, o volume de créditos cursados, que reflete o ritmo de progressão no curso. Esses resultados concordam com a literatura, que aponta o histórico acadêmico como um dos grupos de variáveis de maior poder explicativo (Tete et al., 2022; Silva; Pimentel; Botelho, 2022), e com Sales, Balby e Cajueiro (2016).

Entre as variáveis demográficas e institucionais, o sexo masculino e o turno noturno foram fatores de risco. A razão de chances do sexo masculino foi de 1,51 (IC 95%: 1,36–1,67; $p < 0,001$); a do turno noturno foi de 1,64 a 1,94, todas com $p < 0,001$. Assim, os homens têm chance de evasão 33% a 38% maior que as mulheres, e os estudantes do noturno, 64% a 94% maior que os do diurno. Na análise bivariada, o tipo de escola do ensino médio (V de Cramér = 0,003; $p = 0,79$; sem informação para 1.253 alunos), a cor/raça ($V = 0,033$; $p = 0,19$) e o estado civil ($V = 0,030$; $p = 0,28$) não apresentaram associação significativa com a evasão, e a renda estava registrada para apenas 13,7% dos alunos (1.062 de 7.745). Isso não permite afirmar que sejam irrelevantes, mas indica que, com os dados do sistema institucional, são os indicadores acadêmicos que concentram a maior parte da informação sobre o risco de evasão.

Como limitações, trata-se de um estudo observacional, de modo que as razões de chances expressam associações, e não relações de causa e efeito; além disso, foram analisados apenas estudantes com situação final definida, e a falta de dados socioeconômicos limitou a análise a indicadores acadêmicos e demográficos. O tamanho da amostra diminui ao longo dos períodos (7.348 alunos em P1, 5.483 em P2 e 4.460 em P3), porque quem evade cedo não tem os indicadores dos períodos seguintes, de modo que cada modelo é condicional a ter chegado ao período correspondente. Os valores de p não foram corrigidos para comparações múltiplas, e a inclusão dos 30

alunos com matrícula trancada entre os evadidos pode superestimar a evasão, já que alguns podem retornar ao curso.

CONSIDERAÇÕES FINAIS

Este estudo mostrou que a evasão nas Licenciaturas em Biologia, Física, Matemática e Química da UFCG é elevada e concentrada nos três primeiros semestres, e que a taxa de sucesso e os créditos do CRA são os indicadores mais fortemente associados a ela, ao lado do sexo masculino e do turno noturno. Os resultados indicam que o desempenho nos primeiros períodos concentra a maior parte da informação sobre o risco de evasão. Como esses indicadores já são registrados rotineiramente pela instituição, podem subsidiar coordenações de curso e pró-reitorias no acompanhamento dos estudantes e na adoção de ações de permanência.

Palavras-chave: Evasão no ensino superior; Estatística descritiva; Indicadores acadêmicos; Dados educacionais abertos.

REFERÊNCIAS

ANDIFES. Associação Nacional dos Dirigentes das Instituições Federais de Ensino Superior. **V Pesquisa Nacional de Perfil Socioeconômico e Cultural dos(as) Graduandos(as) das IFES – 2018**. Brasília: ANDIFES, 2019. Disponível em: https://www.andifes.org.br/wp-content/uploads/2019/05/V-Pesquisa-Nacional-de-Perfil-Socioeconomico-e-Cultural-dos-as-Graduandos-as-das-IFES-2018.pdf. Acesso em: 12 jun. 2025.

ARAÚJO, A. C. P. L. de; MARIANO, F. Z.; OLIVEIRA, C. S. de. Determinantes acadêmicos da retenção no Ensino Superior. **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 29, n. 113, p. 1045–1066, dez. 2021. DOI: https://doi.org/10.1590/S0104-403620210002902255.

BOF, A. M.; CASEIRO, L. Z.; MUNDIM, F. C. Carência de professores na educação básica. **Cadernos de Estudos e Pesquisas em Políticas Educacionais**, Brasília, v. 9, 2023. DOI: https://doi.org/10.24109/9786558011026.ceppe.v9.5967. Disponível em: https://cadernosdeestudos.inep.gov.br/ojs3/index.php/cadernos/article/view/5967. Acesso em: 16 jun. 2025.

BRASIL. Ministério da Educação. Comissão Especial de Estudos sobre a Evasão nas Universidades Públicas Brasileiras. **Diplomação, retenção e evasão nos cursos de graduação em instituições de ensino superior públicas**. Brasília: ANDIFES/ABRUEM/SESu/MEC, 1996.

BRASIL. Decreto nº 7.234, de 19 de julho de 2010. Dispõe sobre o Programa Nacional de Assistência Estudantil – PNAES. **Diário Oficial da União**, Brasília, DF, 20 jul. 2010. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/decreto/d7234.htm. Acesso em: 13 jun. 2025.

CARVALHO, D. da R.; NUNEZ, I. B.; MELO, M. M. P. de; SANTOS, R. B. dos; OLIVEIRA, R. M. A. de. Diagnóstico da evasão nas licenciaturas da UFRN: o caso das licenciaturas em física, matemática e química da UFRN. **Brazilian Journal of Development**, Curitiba, v. 5, n. 6, p. 6795–6809, 2019. DOI: https://doi.org/10.34117/bjdv5n6-172.

FIELD, Andy. **Discovering Statistics Using IBM SPSS Statistics**. 6. ed. London: Sage Publications Ltd, 2024.

HOSMER, D. W.; LEMESHOW, S.; STURDIVANT, R. X. **Applied Logistic Regression**. 3. ed. Hoboken: Wiley, 2013. DOI: https://doi.org/10.1002/9781118548387.

INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Resumo Técnico do Censo da Educação Superior 2023**. Brasília, DF: INEP, 2024. Disponível em: https://download.inep.gov.br/educacao_superior/censo_superior/documentos/2023/resumo_tecnico_do_censo_da_educacao_superior_2023.pdf. Acesso em: 16 jun. 2025.

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R.; TAYLOR, J. **An Introduction to Statistical Learning: with Applications in Python**. Cham: Springer, 2023. DOI: https://doi.org/10.1007/978-3-031-38747-0.

MEC. Ministério da Educação. **Reestruturação e Expansão das Universidades Federais: Diretrizes Gerais**. Brasília, 2007. Disponível em: http://portal.mec.gov.br/sesu/arquivos/pdf/diretrizesreuni.pdf. Acesso em: 12 jun. 2025.

MEC. Ministério da Educação. **Pé-de-Meia Licenciaturas**. Brasília, 2025. Disponível em: https://www.gov.br/mec/pt-br/mais-professores/pe-de-meia-licenciaturas. Acesso em: 16 jun. 2025.

O'BRIEN, R. M. A caution regarding rules of thumb for variance inflation factors. **Quality & Quantity**, v. 41, n. 5, p. 673–690, 2007. DOI: https://doi.org/10.1007/s11135-006-9018-6.

OCDE. Organização para a Cooperação e Desenvolvimento Econômico. **Education at a Glance 2025: Brazil**. Paris: OECD Publishing, set. 2025. Disponível em: https://www.oecd.org/en/publications/education-at-a-glance-2025_1a3543e2-en/brazil_d42263a0-en.html.

SALES, A.; BALBY, L.; CAJUEIRO, A. Exploiting academic records for predicting student drop out: a case study in Brazilian higher education. **Journal of Information and Data Management**, v. 7, n. 2, p. 166–180, ago. 2016.

SCHOBER, P.; BOER, C.; SCHWARTE, L. A. Correlation coefficients: appropriate use and interpretation. **Anesthesia & Analgesia**, v. 126, n. 5, p. 1763–1768, 2018. DOI: https://doi.org/10.1213/ANE.0000000000002864.

SCHWARZ, J. C.; DIAS, M. S. de L.; CAMARGO, D. de. Dificuldades encontradas por estudantes no ensino superior e práticas institucionais adotadas para superá-las. **Quaestio – Revista de Estudos em Educação**, Sorocaba, v. 23, n. 3, p. 741–761, nov. 2021. DOI: https://doi.org/10.22483/2177-5796.2021v23n3p741-761.

SILVA, B. J. da; PIMENTEL, E. P.; BOTELHO, W. T. Predição de desempenho de estudantes: uma revisão sistemática de literatura. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO (SBIE), 33., 2022, Manaus. **Anais** [...]. Porto Alegre: Sociedade Brasileira de Computação, 2022. p. 1040–1052. DOI: https://doi.org/10.5753/sbie.2022.225093.

SILVA FILHO, R. L. L.; MOTEJUNAS, P. R.; HIPÓLITO, O.; LOBO, M. B. C. M. A evasão no ensino superior brasileiro. **Cadernos de Pesquisa**, São Paulo, v. 37, n. 132, p. 641–659, set./dez. 2007. DOI: https://doi.org/10.1590/S0100-15742007000300007.

SPEARMAN, C. The proof and measurement of association between two things. **The American Journal of Psychology**, v. 15, n. 1, p. 72–101, 1904. DOI: https://doi.org/10.2307/1412159.

TETE, M. F.; SOUSA, M. de M.; SANTANA, T. S. de; FELLIPE, S. Aplicação de métodos preditivos em evasão no ensino superior: uma revisão sistemática da literatura. **Education Policy Analysis Archives**, v. 30, n. 149, out. 2022. DOI: https://doi.org/10.14507/epaa.30.6845.

TINTO, V. Dropout from higher education: a theoretical synthesis of recent research. **Review of Educational Research**, v. 45, n. 1, p. 89–125, 1975. DOI: https://doi.org/10.3102/00346543045001089.