

---

arquivos analíticos de  
políticas educativas

Revista acadêmica, avaliada por pares,  
independente, de acesso aberto, e multilíngue



Arizona State University

---

Volume 28 Número 149

19 de outubro de 2020

ISSN 1068-2341

---

## Evasão ou Permanência? Modelos Preditivos para a Gestão do Ensino Superior

*Fernanda Cristina da Silva*

*Thiago Luiz de Oliveira Cabral*



*Andressa Sasaki Vasques Pacheco*

Universidade Federal de Santa Catarina (UFSC)  
Brasil

**Citação:** Silva, F. C., Cabral, T. L. O., & Pacheco, A. S. V. (2020). Evasão ou permanência? Modelos preditivos para a gestão do Ensino Superior. *Arquivos Analíticos de Políticas Educativas*, 28(149). <https://doi.org/10.14507/epaa.28.5387>

**Resumo:** Esta pesquisa tem como objetivo propor modelos estatísticos preditivos para a gestão da evasão em cursos de graduação de uma Instituição de Ensino Superior brasileira. Para isso, realizamos um estudo aplicado em quatro cursos de graduação a distância de uma Universidade pública brasileira. Levantamos os dados de 2.991 alunos a partir dos sistemas institucionais da Universidade e utilizamos o método de Regressão Logística Binária. Ao final, concluímos que, para diferentes cursos, variáveis distintas podem influenciar no fenômeno da evasão, bem como uma mesma variável pode gerar efeitos distintos em diferentes realidades. Além disso, os modelos estatísticos preditivos desenvolvidos permitiram a inferência “evasão” ou “permanência” para os alunos ativos na época da coleta dos dados. Em apuração parcial da acurácia dos modelos, identificamos que nove em cada 10 evasões que ocorreram foram identificadas previamente pelos modelos desenvolvidos.

Isso evidencia o potencial para o uso de modelos preditivos na gestão da evasão estudantil, dando-se base para a revisão de políticas e gestão educacionais por meio da identificação das variáveis que influenciam a evasão e a permanência dos estudantes, bem como por meio da identificação precoce dos alunos em risco de evasão.

**Palavras-chave:** Evasão; Modelo preditivo; Ensino Superior; Gestão Universitária

### **Dropout or permanence? Predictive models for higher education management**

**Abstract:** This research aimed to propose statistical predictive models for the dropout management in undergraduate courses of a Brazilian higher education institution. For this, we conducted an applied study in four undergraduate e-learning courses at a Brazilian public university. We collected the data of 2,991 students from the university's institutional systems and we used the binary logistic regression method. In the end, we conclude that for different courses, different variables can influence the dropout phenomenon, as well as the same variable can generate different effects in different realities. In addition, the statistical predictive models developed allowed the inference “dropout” or “permanence” for active students at the time of data collection. In partial assessment of the accuracy of the models, we identified that 9 out of 10 dropouts that occurred were previously identified by the models developed. This highlights the potential for using predictive models to the student dropout management, providing a basis for reviewing educational policies and management by identifying the variables that influence student dropout and permanence, as well as by early identification of students at risk of dropout.

**Keywords:** Academic dropout; Statistical predictive model; Higher education, University management

### **Deserción o permanencia? Modelos predictivos para la gestión de la enseñanza superior**

**Resumen:** Esta investigación tiene como objetivo proponer modelos estadísticos predictivos para la gestión de la deserción universitaria en los cursos de grado en una institución de educación superior brasileña. Para ello, realizamos un estudio aplicado en cuatro cursos a distancia en una universidad pública del Brasil. Recopilamos datos de 2,991 estudiantes de los sistemas institucionales de la Universidad y utilizamos el método de regresión logística binaria. Al final, concluimos que, para diferentes cursos, distintas variables pueden influir en el fenómeno de la deserción, así como la misma variable puede generar diferentes efectos en diferentes realidades. Además, los modelos estadísticos predictivos desarrollados permitieron hacer inferencias como “deserción” o “permanencia” para los estudiantes activos en el momento de la recopilación de datos. En una evaluación parcial de la precisión de los modelos, identificamos que nueve de cada 10 deserciones que ocurrieron fueron identificadas previamente por los modelos desarrollados. Esto resalta el potencial de usar modelos predictivos para gestionar el abandono escolar, proporcionando una base para revisar las políticas y la gestión educativas mediante la identificación de las variables que influyen en el abandono y la permanencia de los estudiantes, así como mediante la identificación de estudiantes en riesgo de deserción.

**Palabras-clave:** Deserción; Modelos predictivos; Enseñanza Superior; Gestión universitaria

## Introdução

O Ensino Superior no Brasil vem apresentando um crescimento constante no que diz respeito ao número de IES e de alunos matriculados. Entre os anos de 2000 e 2016, o número de Instituições mais que dobrou, o quantitativo de alunos ingressantes cresceu cerca de 92% em um período ainda menor (entre 2003 e 2016), e o número de alunos matriculados aumentou mais de 100% (entre 2003 e 2016; INEP, 2010, 2014, 2015, 2018).

Ao considerar que, desde o ano de 2007, os alunos matriculados em cursos de graduação no Brasil somam ano a ano um quantitativo superior a 5 milhões (superior a 6 milhões a partir de 2010, a 7 milhões a partir de 2012 e a 8 milhões a partir de 2015), em uma primeira análise se estima um volume de conclusões próximo ao número de matrículas. Entretanto, isso não se concretiza na realidade. Conforme o Censo da Educação Superior, apesar do crescimento no número de diplomações entre os anos de 2003 e 2016, o recorde de alunos formados não ultrapassa 1,2 milhão num mesmo ano (INEP, 2010, 2014, 2015, 2018).

Parte dessa diferença pode residir no prolongamento do tempo para formação devido às reprovações e aos trancamentos, por exemplo. Parte pode ser explicada pelo fenômeno da evasão estudantil.

Entre os anos de 2003 e 2008, os cursos de graduação na modalidade presencial apresentaram índices de evasão superiores aos cursos ofertados a distância, chegando a 15,01% em 2003. Entretanto, entre os anos de 2008 e 2013, os índices de evasão da Educação a Distância (EaD) foram superiores (alcançando o marco de 25,23% em 2008), com exceção do ano de 2011, no qual as modalidades apresentaram índices equivalentes. Esses dados evidenciam que a evasão é um desafio comum para as modalidades de ensino presencial e a distância (Silva et al., 2016).

Embora não haja um consenso a respeito do que é a evasão, a partir da literatura e das críticas consultadas, assumimos para esta pesquisa que o aluno evadido é aquele cujo *status* final de matrícula é qualquer outro que não concluinte ou *status* equivalente. Essa definição pode ser utilizada independentemente da modalidade de ensino.

Princiotta e Reyna (2009) elaboraram um guia sobre prevenção da evasão e recuperação estudantil. Esse estudo foi publicado na National Governors Association Center for Best Practices, organização cuja missão é desenvolver e implementar soluções inovadoras para políticas públicas. Para os pesquisadores, por meio de uma política pública de acompanhamento e de combate à evasão, é possível gerar benefícios substanciais para a sociedade, pois diminuir as taxas de evasão permite: expandir oportunidades para os mais jovens, abrindo-se o caminho para o sucesso na faculdade, carreira e vida; e gera comunidades mais fortes, vida cívica melhorada e uma força de trabalho aprimorada. Em longo prazo, incrementar a formação escolar, principalmente a graduação, auxilia os estados no caminho para o crescimento econômico.

Uma vez que um aluno evade, consequências são geradas em diferentes esferas relacionadas ao indivíduo e à instituição. Para o aluno, evadir pode representar a perda da oportunidade da formação superior, bem como o desperdício de tempo e de recursos financeiros destinados a um objetivo não alcançado. Por outro lado, isso pode retratar uma escolha consciente, uma decisão por outras formações ou atividades mais relacionadas com os propósitos de vida do indivíduo.

Já para a instituição, a evasão estudantil representa perda de eficiência, visto que uma estrutura física, tecnológica e de pessoas, inicialmente dimensionada para operar em capacidade máxima, passa a ser subaproveitada.

Diversos autores de expressão nacional e internacional convergem no sentido de apresentarem algumas das principais causas da evasão estudantil (ANDIFES et al., 1997; Biazus, 2004; Cislighi, 2008; Gomes, 1998; Munizagaet al., 2018; Noronhaet al., 2001; Pacheco, 2010;

Paredes, 1994; Pereira, 2003; Tinto, 1989). Quando essa análise das causas da evasão é realizada do ponto de vista da educação a distância, outros fatores podem ser somados, conforme apontam García Aretio (2002) e Pacheco (2010). Esse contexto específico quando se trata da EaD está relacionado às características dessa modalidade de ensino que diferem do ensino presencial especialmente no que se refere ao fato de que durante todo o processo de ensino e aprendizagem, ou pelo menos em grande parte dele, professores e alunos se encontram em local e tempo distintos, pois, muitas vezes, os cursos são ofertados de forma massiva; essa oferta de ensino é baseada numa ação conjunta e sistemática de diferentes atores ao longo do processo; esse modo de educação dispõe de acompanhamento de tutoria; faz uso de meios técnicos como apostilas e ambientes virtuais para a disponibilização do material didático e possibilita uma aprendizagem flexível (Belloni, 1999; García Aretio, 2002; Moore & Kearsley, 2007; Simonson, 2006).

Diante das mais diversas causas para a ocorrência desse fenômeno em cursos de graduação, modelos preditivos têm se apresentado como alternativas para prever a interrupção antecipada dos estudos, possibilitando que gestores universitários façam uso dessas previsões para intervir no percurso estudantil do aluno. De forma geral, os modelos preditivos à evasão são construídos com base em dados históricos que, por meio de diferentes métodos, podem fornecer a probabilidade de evasão de cada estudante matriculado. Esses modelos podem ser construídos para uma turma, um curso, uma IES ou um Sistema de Ensino. Quanto menor a granularidade, maior será a precisão das previsões, uma vez que os dados utilizados correspondem a uma população menos heterogênea.

Diante do exposto, buscamos com a presente pesquisa propor modelos estatísticos preditivos para a gestão da evasão de cursos de graduação de uma Instituição de Ensino Superior brasileira. Os modelos foram desenvolvidos para os cursos de Administração, Administração Pública, Letras Espanhol e Matemática, considerando suas ofertas na modalidade a distância.

Para a consecução da pesquisa, definimos como objetivos específicos a identificação das variáveis significativas para a evasão, bem como o efeito de cada uma no fenômeno da evasão; definição dos modelos preditivos; e suas aplicações à realidade de alunos ativos para apurar sua acurácia. Assim, cumpre destacar que a abordagem empírica do tema reforça a relevância deste estudo, uma vez que buscamos estruturar mecanismos para a gestão da evasão no Ensino Superior e não nos limitamos a discriminar seus determinantes.

O objetivo proposto neste estudo vai ao encontro da resposta a um dos desafios apontados pelo grupo de especialistas em Educação Superior, reunidos pelo Conselho Nacional de Educação (CNE), com a colaboração da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) e da Conferência Nacional de Educação (CONAE). Entre os nove desafios elencados pelos especialistas para o período 2011-2020, destaca-se a necessidade da democratização do acesso, da permanência e do sucesso dos estudantes no Ensino Superior brasileiro, além da pouca compreensão a respeito do fenômeno da evasão estudantil na Educação Superior. Os pesquisadores destacam um desafio a ser superado na prática: a persistência dos estudantes para concluírem seus cursos (Speller et al., 2012).

Alinhado ao trabalho desenvolvido pelo CNE, a Educação Superior é contemplada no Plano Nacional de Educação (PNE), Lei que determina as diretrizes, metas e estratégias para a política educacional do país. O PNE vigente determina vinte metas a serem alcançadas no período entre 2014 e 2024, das quais, uma diz respeito às questões de acesso ao Ensino Superior (BRASIL, 2014).

A referida meta é a Meta 12, que, além de contar com estratégias para otimizar a capacidade instalada da estrutura física e dos recursos humanos (estratégia 12.1) e de estimular mecanismos para ocupar as vagas ociosas em cada período letivo (estratégia 12.17) das IES públicas, buscará, de acordo com a estratégia 12.3, “[...] elevar gradualmente a taxa de conclusão média dos cursos de graduação presenciais nas universidades públicas para 90% (noventa por cento) [...]” (Brasil, 2014).

A preocupação com a evasão na Educação Superior não é exclusividade do Brasil. A atenção a essa problemática é de proporções internacionais, inclusive com projetos em parceria entre diferentes países. Uma iniciativa é o Projeto ALFA-GUIA, coordenado pela Universidad Politécnica de Madrid e ao qual integram 14 universidades da América Latina, localizadas na Argentina, Brasil, Chile, Colômbia, Cuba, Equador, México, Nicarágua, Panamá, Paraguai, Uruguai e Venezuela. O projeto tem como principal objetivo “[...] estabelecer e gerenciar uma rede de IES com o intuito de investigar o abandono estudantil de forma profunda e abrangente, buscando difundir medidas eficazes para incrementar a permanência na educação superior” (PUCRS, [s.d.]). Além de gerar conhecimento, a iniciativa possibilita o compartilhamento de “[...] boas práticas voltadas à melhoria da eficiência e qualidade do sistema educativo, sobretudo nos países Latino Americanos” (PUCRS, [s.d.]).

Dessa forma, observamos que a pesquisa aqui apresentada, mesmo desenvolvida no contexto brasileiro, apresenta possibilidades que extrapolam as fronteiras geográficas, podendo seu método ser replicado e os resultados alcançados inspirarem a superação do desafio da evasão no Ensino Superior em outros países.

Por fim, destacamos que os resultados aqui apresentados conferem um caráter de novidade ao estudo. A partir de pesquisas recentes realizadas (estado da arte e revisão sistemática), tal aspecto pode ser observado no presente trabalho tanto na esfera nacional quanto na internacional, uma vez que traz um novo enfoque à problemática da evasão no Ensino Superior, tanto em relação ao objeto de estudo (modelos preditivos à evasão) quanto ao método, que extrapola as inferências estatísticas e demonstra a acurácia dos modelos na prática.

Na revisão sistemática realizada por Munizaga et al. (2018), em artigos científicos sobre evasão estudantil, eles puderam perceber que parte significativa das pesquisas se concentra na caracterização dos alunos evadidos.

Na pesquisa apresentada neste artigo, ultrapassamos as estatísticas descritivas sobre os alunos evadidos e a análise das variáveis associadas à evasão, utilizando tais variáveis para prever o fenômeno e aplicar os modelos desenvolvidos à realidade dos cursos.

Vale destacar que o uso de modelos preditivos à evasão tem se apresentado na agenda de pesquisadores em todo o mundo. Pesquisas publicadas em periódicos de renome apresentam resultados acerca de modelos preditivos desenvolvidos, fazendo uso de diferentes procedimentos estatísticos (Adroque & García de Fanelli, 2018; Amaral, 2016; Assis, 2017; Continiet al., 2018; Fonseca, 2018; Rovaris Neto, 2002). Alguns modelos foram desenvolvidos para cenários menores, como para diferentes dimensões (curso, IES e área de estudo) (Assis, 2017) e diferentes momentos do aluno no curso (Fonseca, 2018). Outros consideraram não somente informações acadêmicas, mas trouxeram enfoque às questões socioeconômicas dos estudantes (Adroque & García de Fanelli, 2018; Contini et al., 2018). Ainda, identificamos um estudo que verificou a acurácia do modelo em um grupo de controle, alcançando aproximadamente 70% (Amaral, 2016).

Dito isso, ressaltamos que este estudo difere das pesquisas já publicadas por apresentar acurácia superior aos índices já alcançados por outros pesquisadores. Além disso, esta pesquisa pode ser considerada original do ponto de vista do método. Outros estudos já desenvolveram modelos preditivos à evasão, como será apresentado mais adiante. Entretanto, diferente das pesquisas anteriores, nas quais a acurácia foi medida também com dados históricos não utilizados na elaboração do modelo, esta pesquisa traz uma proposta diferente. Para o desenvolvimento dos modelos preditivos, utilizamos a base de dados históricos completa e apuramos a acurácia a partir das informações de alunos ativos nos cursos, simulando, assim, a prática de gestão da evasão por parte da IES, ao longo do tempo, durante o acompanhamento dos seus estudantes em risco de evasão.

## Referencial Teórico

É razoável supor que, num sistema escolar ideal e totalmente eficiente, todos os alunos matriculados, sem exceções, alcançariam a diplomação em prazos normais, conforme destacado por Poignant (1976). No entanto, no cenário real, há diversas intercorrências no processo formativo, como a desistência, o abandono, a troca de curso, entre outros, aspectos que são interpretados por pesquisadores e instituições como evasão estudantil.

Embora pareça um fenômeno de simples definição, já em 1982, Vincent Tinto, renomado estudioso sobre a evasão e a permanência estudantil, ressaltou um desalinhamento no campo de estudo no que se refere à concordância sobre o que poderia, de fato, denominar-se evasão. Nesse sentido, observa-se que a definição desse fenômeno não é um consenso e que ela pode, inclusive, ser modificada em relação aos contextos individual, institucional ou estatal (Tinto, 1989).

Ainda que a temática da evasão estudantil esteja sendo discutida com mais afinco desde os anos 1960 em outros países, no Brasil, os estudos se concentraram nesse fenômeno somente a partir da década de 1990, com o panorama da expansão da Educação Superior no país (Santos Junior & Real, 2017). Não obstante, o entendimento da evasão estudantil no contexto nacional ainda permanece em discussão, bem como não há políticas públicas específicas para esse fim.

Na segunda metade dos anos de 1990, visando a atender um de seus objetivos, a Comissão Especial de Estudos sobre Evasão (ANDIFES et al., 1997, p. 19) entendeu por evasão “[...] a saída definitiva do aluno de seu curso de origem, sem concluí-lo”. O trabalho dessa comissão é utilizado até hoje, e pouco se avançou na discussão de políticas específicas sobre evasão, apesar de essa ainda ser uma realidade que preocupa a todas as Instituições de Ensino Superior.

Assim como na modalidade presencial de ensino, os estudiosos da modalidade a distância também divergem em relação ao conceito da evasão estudantil. Para Favero (2006, p. 50), a evasão pode ser entendida como “[...] o ato da desistência, incluindo os [alunos] que nunca se apresentaram ou se manifestaram de alguma forma para os colegas e mediadores do curso, em qualquer momento”. Nascimento e Esper (2014) optaram por considerar como evadido o aluno que possui ao menos um registro de acesso à sala de aula virtual, porém não concluiu o curso. Há ainda de se ressaltar que a evasão na educação a distância, diferente da educação presencial, pode assumir duas facetas, conforme destacado por García Aretio (2002): “abandono sem começar” e “abandono real”. O primeiro se refere àqueles alunos que não possuem registros de avaliações, ainda que matriculados no curso. Já o último diz respeito aos alunos matriculados que possuem registros de avaliação, porém não alcançaram a diplomação.

Observa-se, assim, que alguns especialistas da modalidade a distância consideram evadidos os alunos que se matricularam e não concluíram o curso, enquanto outros entendem que deve ser considerado como evadido somente aqueles estudantes que marcaram presença no ambiente virtual de aprendizagem ou atividades avaliativas, porém, não concluíram o curso e não formalizaram o trancamento dos estudos.

Com base nos autores supracitados e também embasados no entendimento de Fritsch, Rocha e Vitelli (2005), Gaioso (2005), Castro e Teixeira (2014), Dore, Sales e Castro (2014), Ambiel (2015) e Fritsch (2017) sobre o fenômeno da evasão (Schmitt, 2018), entendemos, e, assim abordamos neste estudo, que o aluno evadido é aquele cujo *status* final de matrícula é qualquer outro que não concluinte ou *status* equivalente. Nesse sentido, a interpretação do fenômeno passa a ser o vínculo do aluno com a instituição, ou seja, a matrícula, e não o aluno em si. Nesse caso, um mesmo aluno pode acumular um histórico acadêmico com diferentes desfechos, independentemente do curso, instituição ou sistema de ensino, tal qual acontece na realidade. Há de se ressaltar ainda que a

definição que assumimos contempla o Ensino Superior de maneira ampla, sem distinção entre as modalidades presencial e a distância.

Há, entretanto, de se ponderar que a questão temporal também acaba por influenciar nas diferentes conceituações de evasão, podendo o fenômeno ser associado aos conceitos de evasão temporária e de evasão definitiva para diferentes autores. Entende-se, no entanto, que a evasão temporária contraria a própria definição de evasão, como anteriormente exposta e, nesse sentido, consideramos neste artigo a evasão como um fenômeno definitivo e irreversível.

De acordo com Ristoff (1995), autor basilar nas definições da Comissão Especial de Estudos sobre Evasão (ANDIFES et al., 1997), é fundamental, ainda, compreender que a evasão estudantil se dá em diferentes dimensões, ou seja, pode ocorrer em um curso, numa instituição ou no sistema educacional como um todo. Nota-se então que há uma ordem de grandeza que se dá do sistema para o curso, visto que ao evadir do sistema, o aluno evade, automaticamente, da instituição e do curso. Contudo, o estudante pode evadir da instituição, por exemplo, e permanecer no sistema educacional.

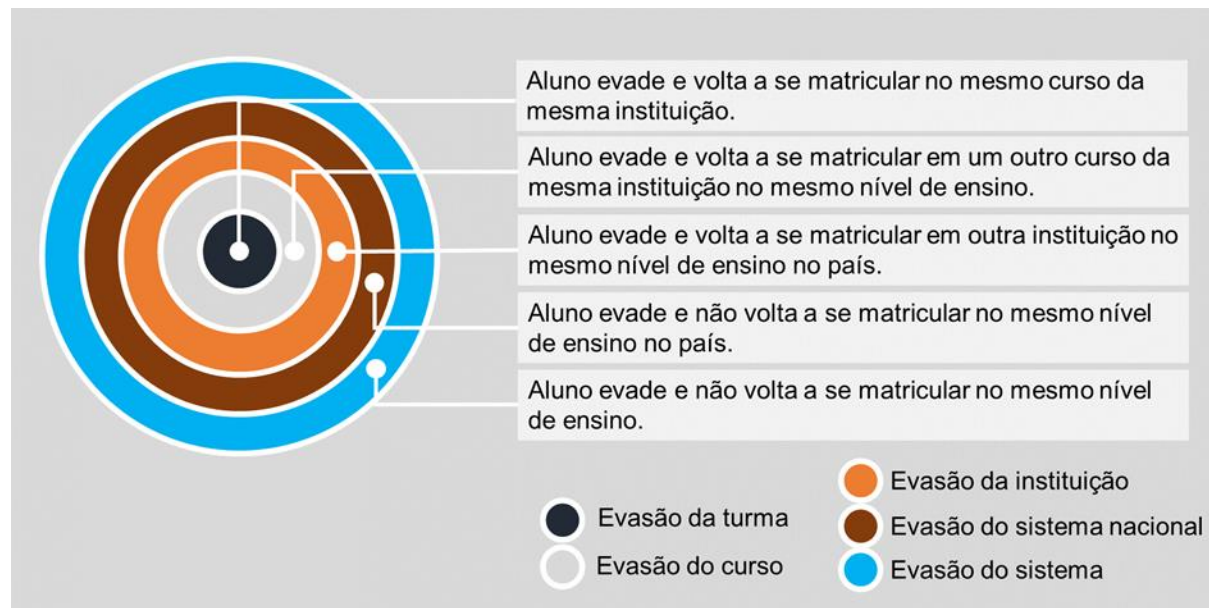
A partir do entendimento da referida Comissão, ponderamos que pode ainda haver dois casos adicionais: a evasão da turma, entendida aqui como o grupo de alunos ingressantes num mesmo curso no mesmo período letivo; e uma dimensão intermediária entre a evasão da instituição e a evasão do sistema, denominada por nós como evasão do sistema nacional.

No que se refere à evasão da turma, Suiter e Pacheco (2019) tratam do caso da reocupação de vagas no Ensino Superior. Nesse caso, cabe considerar a possibilidade de o aluno evadido retornar aos estudos com uma nova matrícula, no mesmo sistema de ensino, na mesma instituição para o mesmo curso no qual evadiu no passado.

Já em relação à evasão do sistema, é importante considerar que um estudante que evade do sistema nacional pode realizar a formação em um outro país, no mesmo nível de ensino e, caso não o faça, aí sim, tem-se a evasão do sistema de ensino como um todo. A figura a seguir apresenta as dimensões da evasão estudantil, segundo nosso entendimento aqui defendido:

### Figura 1

Interpretação das dimensões da evasão



Fonte: Elaborada pelos autores deste artigo

Considerando as diferentes dimensões da evasão, Lobo (2012) comenta sobre as dificuldades em obter dados seguros sobre a evasão da instituição ou da evasão do sistema de ensino frente aos desafios em identificar se a interrupção dos estudos nessas dimensões é temporária ou definitiva. Nesse sentido, percebe-se que há diferentes possibilidades para se analisar a evasão estudantil que acabam por interferir nas formas em que o indicador de evasão é apurado.

Além das preocupações relacionadas à observação e mensuração da evasão, os motivos (ou causas) que levam o aluno a evadir no ensino superior também têm recebido especial atenção por parte dos estudiosos. Tinto (1975), com sua pesquisa, elaborou um modelo teórico para explicar os processos de interação entre o aluno e a instituição de ensino que levam diferentes indivíduos a evadirem de seus cursos, denominado por ele de Modelo Longitudinal de Evasão.

Inicialmente, o modelo desenvolvido por Tinto (1975) considera que cada indivíduo que ingressa na vida universitária traz consigo um conjunto de atributos (sexo, raça, etc.), experiências pré-universitárias (desempenhos escolares, realizações acadêmicas, etc.) e antecedentes familiares que impactam de maneira direta e indireta em seu desempenho em um curso de graduação, bem como nas expectativas e compromissos que o aluno traz para o ambiente universitário (Tinto, 1975). Para o autor, a integração do indivíduo nos sistemas acadêmico e social da instituição é aquilo que mais possibilita a continuidade do vínculo entre o aluno e a IES. Dessa forma, quanto maior o grau de integração existente, maior tende a ser o compromisso do estudante com a instituição e, consequentemente, com a conclusão do curso.

Além de Vincent Tinto, outros estudiosos também apontaram os principais motivos que levam os estudantes a evadir. Desses motivos, pode-se destacar a falta de clareza pela escolha do curso e seu consequente desinteresse; matrícula no curso como segunda opção; pressão familiar para início do curso; imaturidade do estudante; pouco envolvimento com o curso; dificuldades de conciliar os estudos com a rotina do trabalho e da família; realização de mais de um curso simultaneamente; problemas de saúde; mudança de cidade; dificuldade de adaptação à vida universitária; baixo desempenho acadêmico; baixa qualidade das IES; infraestrutura deficitária da IES; falta de assessoramento aos alunos calouros; alto custo do curso; rigidez, pouca motivação e despreparo dos docentes; metodologias educacionais inadequadas; falta de prestígio da profissão; falta de integração entre IES e empresas; currículos desatualizados (ANDIFES et al., 1997; Biazus, 2004; Cislighi, 2008; Gomes, 1998; Munizaga et al., 2018; Noronha et al., 2001; Paredes, 1994; Pereira, 2003; Tinto, 1989)

Ademais, Pacheco (2010) destaca fatores determinantes relacionados à evasão e à permanência estudantil na EaD que se somam aos já apresentados, como a falta de identificação com a modalidade de ensino; a rotatividade de tutores ao longo do curso; um acolhimento institucional que faça o aluno se sentir parte da IES, mesmo com a distância física existente; um processo comunicativo claro, pronto e cordial pelos meios disponíveis; as tecnologias oferecidas; e o relacionamento com os profissionais da tutoria.

Frente aos desafios da evasão estudantil no Ensino Superior, tanto no ensino presencial quanto na EaD, é crescente o número de pesquisas no campo da gestão universitária que têm buscado sugerir ou testar estratégias que favoreçam a permanência dos estudantes. Uma vez sabendo-se da diversidade de motivos que podem levar um aluno a evadir, estima-se também uma diversidade de estratégias para minimizar os impactos desse fenômeno. Assim, diversos pesquisadores têm concentrado seus estudos na gestão da evasão no Ensino Superior para possibilitar uma melhor compreensão acerca do fenômeno e de sua possível mitigação (Aguado Franco, 2017; Arendt, 2013; Jones & Braxton, 2009; Kori et al., 2016; Martín et al., 2011; Sánchez-Hernández et al., 2017).

Para além de descrever o fenômeno da evasão, apurar estatísticas descritivas a respeito do fenômeno e testar estratégias que visam à permanência estudantil em cursos de graduação, observa-



se também o desenvolvimento de pesquisas relacionadas à previsão dos estudantes em risco de evadir. Os modelos desenvolvidos contribuem sobremaneira à gestão universitária na medida em que apontam, de forma antecipada e personalizada, a probabilidade de evadir de um dado aluno, conhecidos como modelos preditivos (Adroque & García de Fanelli, 2018; Amaral, 2016; Assis, 2017; Carvajal et al., 2018; Contini et al., 2018; Fonseca, 2018; Pérez et al., 2018; Rovaris Neto, 2002).

Destarte, os modelos preditivos apresentam oportunidades para a gestão universitária e para a elaboração de políticas públicas, que residem especialmente na possibilidade de prever, com certo grau de confiança, a conclusão ou evasão do aluno e no conhecimento de quais aspectos influenciam mais na permanência estudantil.

Entretanto, para a elaboração de tais modelos, não basta a disponibilização dos dados, mas a obtenção de informações consistentes e atualizadas. Inicialmente, para a construção do próprio modelo, o qual depende de dados históricos para seu desenvolvimento, e, nesse caso, dados inconsistentes podem levar a modelos não confiáveis. Em um segundo momento, tal importância é ressaltada pela necessidade de se aplicar ao modelo as informações dos estudantes em curso. Caso as informações não sejam consistentes e atualizadas, mesmo com base num modelo confiável, as predições de evasão podem se dar de maneira equivocada, levando a análises e a tomadas de decisão não coerentes.

Dessa forma, consideramos não somente necessário garantir a validade na utilização de modelos preditivos de evasão aplicados à gestão universitária, como, talvez mais importante do que isso, assegurar a qualidade das informações utilizadas para o desenvolvimento de tais modelos, visto que os dados são a matéria-prima para sua construção.

## Procedimentos Metodológicos

A seção a seguir apresenta as principais características dos cursos e dos alunos que compõem o universo desta análise. Na sequência, descrevemos a população estudada, os procedimentos para a coleta dos dados, bem como o procedimento metodológico utilizado para o desenvolver os modelos estatísticos preditivos à evasão.

### População Estudada

A população estudada pertence a uma universidade pública federal brasileira, cuja secretaria institucional manifestou interesse em participar da pesquisa e em disponibilizar os dados necessários.

Por ser uma universidade referência e precursora no que diz respeito à educação a distância, elegemos os cursos de graduação EaD para o desenvolvimento da pesquisa.

A ideia inicial era trabalhar com todos os cursos de graduação a distância da universidade. Após realizarmos contato com todos os coordenadores dos referidos cursos, obtivemos retorno positivo das coordenações dos cursos de Administração, Administração Pública, Filosofia, Letras Espanhol e Matemática. Devido ao baixo número de alunos no curso de Filosofia e em função de este não estar ativo no período da pesquisa, optou-se por não desenvolver o modelo preditivo para esse curso.

Assim, a população estudada considerou os 2.991 alunos dos cursos de graduação a distância em Administração, Administração Pública, Letras Espanhol e Matemática, os quais foram categorizados com “matrícula ativa” ou “matrícula não ativa”. A matrícula foi considerada ativa no caso em que o aluno ainda não possuía um *status* final no curso, ou seja, nos casos em que o aluno ainda estava em processo de formação, seja cursando um semestre letivo ou mesmo com a matrícula trancada, já que esse *status* é temporário, podendo o aluno retornar ao curso para concluí-lo. Já a matrícula não ativa foi considerada no caso em que o aluno já possuía um *status* final no curso, não

podendo este ser alterado (*status* abandono, concluinte, desistente, falecido, formado, transferido e troca de curso). Essa distinção se fez necessária para que o modelo preditivo fosse desenvolvido com base nas matrículas não ativas e a acurácia fosse apurada aplicando-se o modelo preditivo aos alunos com matrículas ativas. Os quantitativos são apresentados na tabela a seguir:

**Tabela 1**

*Matrículas ativas e não ativas nos cursos estudados*

Curso	Matrículas ativas	Matrículas não ativas	Total
Administração	314	1.132	1.446
Administração Pública	82	540	622
Letras Espanhol	74	371	445
Matemática	19	459	478
Total	489	2.502	2.991

*Nota:* elaborada pelos autores deste artigo a partir dos dados dos sistemas institucionais.

A população estudada possuía idade média de 36 anos, sendo que o aluno mais novo tinha 20 anos e o mais velho 70 anos completos no período da pesquisa. Foi formada por 51% de indivíduos do sexo feminino, aproximadamente 82% proveniente de escola pública, 83% se autodeclararam brancos e mais de 54% possuíam renda de até três salários mínimos. No que diz respeito à experiência com a modalidade de ensino, 69% declararam nunca ter realizado um curso a distância antes de se matricular no curso de graduação.

Cumprе ressaltar que, mesmo havendo um perfil predominante na população estudada, o modelo desenvolvido pode ser aplicado a qualquer aluno matriculado em um dos cursos contemplados na pesquisa, já que a garantia de uma significância de 5% e um grau de confiança de 95% é dada pelo método e não pelas características dos indivíduos, conforme demonstramos na seção “Regressão Logística Binária”.

### Coleta de Dados

Para a realização da pesquisa, solicitamos à Universidade acesso aos dados coletados no vestibular (processo seletivo para ingresso), dados do Sistema Acadêmico e do Ambiente Virtual de Aprendizagem (AVA).

Por intermédio da Secretaria de Educação a Distância da Universidade, os setores responsáveis pela disponibilização dos dados foram acionados e as informações foram cedidas em arquivos nos formatos “csv” e “txt”. Em função da não disponibilidade das informações do AVA para os períodos anteriores a 2012, essas informações não foram utilizadas.

No levantamento dos dados, efetuado em setembro de 2016, consideramos as matrículas realizadas a partir do ano de 2008. Esse corte temporal se fez necessário, uma vez que nos anos anteriores a oferta de cursos de graduação a distância era incipiente na instituição e os dados daquele período poderiam comprometer a confiabilidade do modelo estatístico preditivo. Destaca-se, ainda, que a identidade dos alunos foi preservada pela instituição, não sendo disponibilizada na coleta de dados da pesquisa, já que a informação era irrelevante para a consecução deste estudo.

A seleção das informações para o desenvolvimento da pesquisa se deu a partir dos dados disponíveis nos sistemas acadêmicos, do referencial teórico consultado (considerando as causas da evasão estudantil apontadas por estudos já publicados) e da *expertise* do grupo de pesquisa constituída a partir das experiências e de outras investigações. A lista das variáveis selecionadas e suas categorias

podem ser visualizadas no Apêndice A. Do sistema acadêmico extraímos dez informações, a saber: situação da matrícula; sexo; idade; cor; estado civil; cidade de residência; cidade do polo de apoio presencial; categoria de ingresso; pontos no vestibular; e Índice de Aproveitamento Acumulado (IAA). A partir das informações de cidade de residência e cidade do polo, foi criada mais uma informação: se o aluno residia ou não na cidade do polo de apoio presencial. Assim, ao final, totalizaram dez informações do sistema acadêmico da instituição para serem utilizadas como variáveis independentes e uma como variável dependente (situação da matrícula).

Já a partir do questionário aplicado no momento da inscrição no vestibular, foram consideradas dez informações socioeconômicas. A seguir são apresentadas as questões contempladas no questionário socioeconômico, mantendo-se a enumeração original disponibilizada pela Instituição:

- Questão 1) Renda total mensal familiar (valor de referência – um salário mínimo).
- Questão 2) Número de pessoas que residem com você (incluindo você).
- Questão 3) Qual o tempo aproximado que você leva para chegar até o polo de ensino a distância de sua cidade.
- Questão 4) Concluiu a maior parte do ensino médio em (escola pública ou particular).
- Questão 5) Em relação ao ensino superior (experiência).
- Questão 12) Já realizou algum curso a distância.
- Questão 14) Com que frequência você utiliza o computador.
- Questão 15) Onde você tem acesso a computador conectado à internet.
- Questão 16) Como é sua conexão com a internet
- Questão 18) Como é o seu conhecimento sobre a utilização do computador e da internet.

Os dados disponibilizados pelos setores eram consistentes, não sendo apresentados dados corrompidos ou incoerentes. As categorias das variáveis estavam padronizadas e foi disponibilizado um dicionário de dados que apresentava a definição de cada variável. Cabe destacar que, para a variável “estado civil”, houve casos em que a categoria assinalada foi “Não informado”. Ou seja, nessas situações, a IES não dispunha daquela informação sobre os referidos alunos. Entretanto, tal constatação não impediu a aplicação do procedimento de regressão logística. Nas demais variáveis, mesmo possuindo a categoria “Não informado”, para nenhum dos alunos ela foi selecionada.

A respeito das informações socioeconômicas coletadas, estas foram cadastradas pelos próprios estudantes no ato da inscrição para o vestibular. Já as informações pessoais obtidas por meio do sistema acadêmico foram cadastradas pelas secretarias dos cursos no momento da matrícula, com possibilidade de atualização posterior por parte dos alunos para aquelas informações que podem mudar ao longo do tempo, como endereço, por exemplo. As informações de desempenho acadêmico e de *status* de matrícula foram atualizadas ao final de cada semestre letivo.

Por fim, registramos que, sobre os dados levantados, foram desenvolvidas tabelas de contingência para identificar as necessidades de agrupamento de categorias das variáveis qualitativas. Esse agrupamento se faz necessário quando existem poucas observações em alguma categoria, o que pode comprometer a análise. Para Hair et al. (2009), cinco é um número razoável de observações para cada categoria de uma variável independente. Diante disso, foi necessário realizar o agrupamento de categorias qualitativas em nove variáveis.

## Regressão Logística Binária

Aos dados disponibilizados e organizados, aplicamos o processo de Regressão Logística Binária<sup>1</sup> para a identificação dos aspectos que impactam na evasão dos cursos. A regressão é um modelo estatístico-matemático que relaciona uma variável dependente  $Y$  (ou variável resposta) a uma variável independente  $X$  (ou variável explicativa). Nesse sentido, pode-se estabelecer uma dependência entre as duas variáveis a partir de uma relação linear, por exemplo, dada por  $y = \alpha + \beta x$ , que pode ser obtida a partir de uma regressão linear, processo que permite obter valores aproximados para  $\alpha$  e  $\beta$  a partir de observações de  $x$  e  $y$  (Barbetta, 2014).

De acordo com Fávero (2015), diferente do método de regressão linear, em que a variável dependente é quantitativa e linear, a regressão logística não é linear e é utilizada quando o fenômeno se apresenta de forma qualitativa, representado por uma ou mais variáveis. Nesse caso, é preciso transformar a variável qualitativa em uma variável quantitativa, e, para isso, utiliza-se a variável binária no formato “0 – 1”.

Assim, Fávero (2015, p. 104) define o principal objetivo da regressão logística binária como sendo

[...] estudar a probabilidade de ocorrência de um evento definido por  $Y$  que se apresenta na forma qualitativa dicotômica ( $Y=1$  para descrever a ocorrência do evento de interesse e  $Y=0$  para descrever a ocorrência do não evento), com base no comportamento de variáveis explicativas.

Nesse caso, ao final, obtém-se um modelo matemático que permite calcular a probabilidade de ocorrência de um fenômeno a partir das variáveis independentes, que pode ser expresso por:

$$p_i = \frac{1}{1 + e^{-(\alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})}}$$

onde  $p_i$  é a probabilidade do evento  $Y$  ocorrer;  $\alpha$  representa uma constante;  $B_j$  ( $j=1,2,\dots,k$ ) são os parâmetros estimados de cada variável independente (também chamados de coeficientes logísticos);  $x_j$  ( $j=1,2,\dots,k$ ) são as variáveis independentes e  $i$  representa uma observação da amostra (Hair et al., 2009; Fávero, 2015). Para o presente estudo, a Regressão Logística Binária foi realizada no *software* R.

Inicialmente, definimos as hipóteses nula e alternativa:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{existe pelo menos um } \beta_j \neq 0$$

A confirmação da hipótese nula indica que nenhuma das variáveis independentes influencia na evasão dos estudantes dos cursos estudados. Já a comprovação da hipótese alternativa confirma que uma ou mais informações contidas nos sistemas institucionais influenciam significativamente esse fenômeno.

Para a verificação da significância de cada parâmetro do modelo de regressão logística binária, utilizamos a estatística  $Z$  de Wald. Nesse caso, após calcular  $Z$  para  $\alpha$  e para  $\beta_j$ , recorreremos à

---

<sup>1</sup> Para o desenvolvimento dos modelos preditivos, optamos pelo modelo logit (ao invés do probit), já que esse permite o cálculo da razão de chances e possibilita, assim, medir o efeito das variáveis no fenômeno, conforme objetivo específico definido. Além disso, para as amostras utilizadas neste estudo, o desempenho dos modelos logit e probit é semelhante, conforme apontam Freitas et al. (2013).

tabela de distribuição normal padrão para verificar se as variáveis independentes eram estatisticamente significantes para o modelo. Para uma significância de 5% e um grau de confiança de 95%, definidas para esse estudo, excluimos do modelo os parâmetros que apresentaram  $Z$  entre -1,96 e 1,96, ou  $p - value > 0,05$ .

Para a exclusão das variáveis, utilizamos, inicialmente, o método *stepwise*, e, posteriormente, o Modelo Linear Generalizado do *software* R, garantindo, assim, a permanência apenas das variáveis relevantes para a determinação da evasão (Fávero, 2015; Hocking, 1976; Hosmer Jr. & Lemeshow, 1989).

Com base nos coeficientes  $\alpha$  e  $\beta_j$  obtidos para cada variável identificada como significativa, elaboramos os modelos matemáticos preditivos para a evasão dos cursos de graduação estudados. Além da análise dos dados a partir da significância das variáveis, analisamos também a razão de chances (*odds ratios*).

Por definição, a chance é expressa por  $\frac{p_i}{1-p_i}$ , onde  $p_i$  é a probabilidade de ocorrência do fenômeno associada à observação  $i$ . Assim, caso a probabilidade de um aluno evadir do curso seja de 20%, a sua chance de evasão é de  $\frac{1}{4}$  (um para quatro), já que  $\frac{20\%}{1-20\%} = \frac{20\%}{80\%} = \frac{1}{4}$ .

A partir da chance, é possível calcular o quanto a probabilidade de ocorrer o fenômeno de interesse varia quando se altera uma das variáveis independentes (mantendo as demais constantes).

Esse processo pode ser realizado a partir da razão  $\frac{Chance_1}{Chance_2}$ , definida como “razão de chances”, que mede o efeito de uma variável sobre o fenômeno de interesse quando se mantêm as demais variáveis constantes.

O resultado desse processo mostra o quanto aumenta ou diminui a chance de o evento acontecer, conforme se altera a variável em análise. Assim, neste estudo, foi possível medir o quanto a chance de o aluno evadir aumenta ou diminui com a alteração de apenas uma variável, mantendo as demais constantes.

Por fim, aplicamos o modelo estatístico desenvolvido aos dados observados dos alunos com matrícula ativa nos cursos estudados, o que possibilitou prever a probabilidade de os referidos estudantes evadirem.

## Resultados e Discussão

Nesta seção serão apresentados os resultados da pesquisa na seguinte estrutura:

- a) Variáveis significativas e o efeito de cada uma no fenômeno da evasão (análise da razão de chances).
- b) Modelos preditivos à evasão.
- c) Aplicação e acurácia dos modelos preditivos desenvolvidos.

### As Variáveis Significativas e o Efeito de Cada uma no Fenômeno da Evasão

Para a identificação dos principais aspectos que impactam na evasão dos cursos estudados, inicialmente definimos as seguintes hipóteses:

$H_0$ : não existe relação entre a evasão dos alunos e as informações dos sistemas institucionais.

$H_1$ : há relação entre a evasão dos alunos e as informações dos sistemas institucionais.

Essas hipóteses foram consideradas para cada um dos cursos estudados, já que desenvolvemos um modelo preditivo para cada curso, com o propósito de não homogeneizar as informações em uma análise ampla e generalizada.

Antes de iniciar o processo de regressão logística, definimos a categoria referência para cada variável qualitativa. Para isso, utilizamos como critério a categoria com maior frequência apurada. Essa definição não altera os resultados, mas se fez necessária para o procedimento no *software*. No caso em que a maior frequência foi conferida à categoria “Não informado”, consideramos como categoria referência a que apresentou a segunda maior frequência. A tabela a seguir apresenta a categoria referência de cada variável qualitativa.

**Tabela 2**

Variáveis qualitativas e categorias referência

Variável	Administração	Administração Pública	Letras Espanhol	Matemática
Sexo	M (masculino)	M (masculino)	F (feminino)	M (masculino)
Cor	Branca	Branca	Branca	Branca
Estado civil	Casado/União	Casado/União	Solteiro	Casado/União
	Estável	Estável		estável
UF residência	Região Sul	Região Sul	*	*
UF polo	Região Sul	*	*	*
Reside cidade polo	Não	Sim	Sim	Sim
Categoria de ingresso	Classificação geral	Classificação geral	Classificação geral	*
Renda familiar	De 1 a 3 salários	De 4 a 8 salários	De 1 a 3 salários	De 1 a 3 salários
Tamanho da família	De 3 a 4 pessoas	De 3 a 4 pessoas	De 3 a 4 pessoas	De 3 a 4 pessoas
Tempo de deslocamento até o polo	Até 30 minutos	Até 30 minutos	Até 30 minutos	Até 30 minutos
Onde estudou a maior parte do Ensino Médio	Escola pública	Escola pública	Escola pública	Escola pública
Experiência no Ensino Superior	Nunca ingressou	Já concluiu	Já concluiu	Nunca ingressou
Experiência na EaD	Não	Não	Não	Não
Frequência de uso do computador	Todos os dias	Todos os dias	Todos os dias	Todos os dias
Local de acesso à internet	Casa	Casa	Casa	Casa
Tipo de conexão à internet	ADSL	ADSL	ADSL	ADSL
Nível de conhecimento para o uso do computador e da internet	Médio	Médio	Médio	Médio

*Nota:* Os campos indicados com \* sinalizam que a variável não foi considerada para o respectivo curso, visto que todos os alunos possuíam a mesma informação relativa àquela variável. Tabela elaborada pelos autores deste artigo.

Definidas as categorias referências para as variáveis qualitativas, procedemos com a aplicação da Regressão Logística Binária individualmente para cada curso.

Com base nas informações dos alunos categorizados com matrícula “não ativa”, utilizamos o comando de Ajuste de Modelos > Modelo Linear Generalizado (GLM) do *software* estatístico R para realizar o processo de ajuste dos dados com base na regressão logística binária.

Após o ajuste dos dados, o *software* retornou as estimativas para os coeficientes do modelo de regressão logística. Aos dados ajustados, aplicamos o método *stepwise* com a direção *forward/backward* para que somente aquelas variáveis identificadas como importantes para o fenômeno da evasão permanecessem.

Nesse processo, o *software* retornou as seguintes variáveis:

- Administração: IAA; Experiência no Ensino Superior (Questão 5); Estado Civil; Categoria de ingresso; pontos Vestibular; Reside na cidade polo; idade; e cor.
- Administração Pública: IAA; Estado Civil; Categoria de ingresso; Nível de conhecimento para o uso do computador e da internet (Questão 18); UF Residência; e Onde estudou a maior parte do Ensino Médio (Questão 4).
- Letras Espanhol: IAA; Tamanho da família (Questão 2); idade; e Estado Civil.
- Matemática: IAA; Nível de conhecimento para o uso do computador e da internet (Questão 18); Estado Civil; Local de acesso à internet (Questão 15); Tamanho da família (Questão 2); e Experiência na EAD (Questão 12).

Conforme aponta a metodologia proposta, a obtenção do modelo deveria garantir que permanecessem somente variáveis significativas, considerando um nível de significância de 5%.

Dessa forma, processamos o Modelo Linear Generalizado novamente, considerando, nesse caso, somente as variáveis que o método *stepwise* resultou como importantes para cada curso.

Uma a uma, excluímos as variáveis com valores de  $p$  superiores ao desejado, iniciando a exclusão daquelas com maior valor de  $p$ . A partir de cada exclusão, apuramos novamente o modelo generalizado até que permanecessem somente as variáveis dentro do nível de significância definido.

Até a obtenção das variáveis finais para cada curso, o processo nos levou a excluir outras variáveis, as quais foram descartadas na seguinte ordem:

- Administração: Categoria de ingresso; Estado civil; Experiência no Ensino Superior (Questão 5); Idade (que deixou de ser significativa com a exclusão da variável “Experiência no Ensino Superior (Questão 5)”); Cor.
- Pública: UF residência; Categoria de ingresso; Onde estudou a maior parte do Ensino Médio (Questão 4) (que deixou de ser significativa a partir da exclusão da variável “Categoria de ingresso”); Nível de conhecimento para o uso do computador e da internet (Questão 18) (pois a categoria “pouco ou nenhum” apresentava significância superior a 5%); e Estado civil (que apresentou a significância desejada em apenas uma de suas categorias).
- Letras Espanhol: Estado Civil; e Tamanho da família (Questão 2).
- Matemática: Tamanho da família (Questão 2); Nível de conhecimento para o uso do computador e da internet (Questão 18); Estado civil; Local de acesso à internet (Questão 15); Experiência na EAD (Questão 12) (que deixou de ser significante para o modelo após a exclusão da variável “Local de acesso à internet (Questão 15)”).

Assim, restaram como significativas para a evasão em cada curso as seguintes variáveis:

- Administração: IAA; pontos Vestibular; e Reside na cidade polo.
- Administração Pública: IAA.
- Letras Espanhol: IAA; e idade.
- Matemática: IAA.

Com o término desse procedimento obtivemos o *odds ratios* para cada uma das variáveis significativas para a evasão nos cursos analisados. O *odds ratios*, ou razão de chances, permite medir o efeito das variáveis no fenômeno, mantendo-se as demais variáveis constantes. As razões de chances obtidas podem ser observadas na tabela a seguir.

**Tabela 3**

Odds ratios (razão de chances) das variáveis significativas para a evasão em cada curso

Curso	Variável	Pr ( $>  z $ )	Odds ratios
Administração	IAA	$<2e-16$	1.173675e-01
	pontosVestibular	0.000155	1.866797e+00
	Reside.na.cidade.polo.[T.sim]	0.011268	2.316121e+00
Administração Pública	IAA	$<2e-16$	0.2506229
Letras Espanhol	IAA	$3.12e-08$	8.364696e-02
	Idade	0.0411	1.068559e+00
Matemática	IAA	$1.80e-07$	0.197359

*Nota:* elaborada pelos autores deste artigo a partir das saídas do *Software R*.

De acordo com os dados apresentados na tabela anterior, podemos notar que no curso de graduação em **Administração**, a cada 1,00 ponto a mais no IAA (considerando notas de 0 a 10), a chance de o aluno evadir é multiplicada por 0,12, ou seja, reduz em 88%. A respeito da pontuação do aluno no vestibular, observamos que a cada 1,00 ponto a mais no vestibular (considerando notas de 0 a 10), a chance de o aluno evadir é multiplicada por 1,87, ou seja, aumenta em 87%. Por fim, no que se refere à residência do estudante, para alunos que residem na cidade do polo, a chance de evadir é multiplicada por 2,31, em outras palavras, aumenta 131%.

Para o curso de graduação em **Administração Pública**, identificamos que a cada 1,00 ponto a mais no IAA (considerando notas de 0 a 10), a chance de o aluno evadir é multiplicada por 0,25, ou seja, reduz em 75%.

Já no curso de graduação **Letras Espanhol**, a cada 1,00 ponto a mais no IAA (considerando notas de 0 a 10), a chance de o aluno evadir é multiplicada por 0,083, ou seja, reduz em 91,7%. No que se refere à idade do estudante, observamos que a cada 1 ano a mais de idade, a chance de o aluno evadir é multiplicada por 1,068, ou seja, aumenta em 6,8%.

Por fim, no curso de graduação em **Matemática**, a cada 1,00 ponto a mais no IAA (considerando notas de 0 a 10), a chance de o aluno evadir é multiplicada por 0,1973, ou seja, reduz em 80,27%.



A partir da análise da razão de chances da variável IAA, a qual é significativa para a evasão em todos os cursos estudados, resgatamos a teoria consultada, a qual aponta diferentes aspectos que relacionam o desempenho do aluno à evasão estudantil.

Os resultados que alcançamos mostram que qualquer aumento no IAA do aluno pode reduzir significativamente a chance de evasão, o que sugere a possibilidade de se promoverem iniciativas que possibilitem ao estudante o alcance de um melhor desempenho nas disciplinas. Por outro lado, destacamos que um baixo desempenho pode resultar, também, de aspectos não relacionados de maneira direta às questões acadêmicas ou de aprendizagem, como fatores socioeconômicos referentes à situação financeira, à vida familiar e às questões laborais, por exemplo.

No que tange à variável “Pontos vestibular”, esta torna-se significativa apenas para o curso de Administração, pois percebe-se que um melhor desempenho impacta em uma maior chance de evasão, um resultado, inicialmente, curioso para nós. Entretanto, ao analisá-lo a partir do referencial teórico, situação que pode motivar esse fenômeno, ressalta-se a aprovação em outros processos seletivos, ocasião que pode levar o aluno a evadir, uma vez que este pode escolher realizar outro curso superior, no qual tenha sido aprovado e esteja mais alinhado aos seus objetivos acadêmicos e profissionais.

Em relação à residência do aluno, também significativa apenas para o curso de Administração, observa-se que aqueles que não residem na cidade do polo possuem uma chance inferior de evadir do curso em comparação àqueles que moram na mesma cidade em que o polo de apoio presencial está localizado.

No referencial teórico consultado, a distância que o aluno precisa percorrer entre sua casa e a instituição de ensino não é mencionada como um dos fatores que pode influenciar na evasão. Todavia, vale destacar que alunos que residem em outra cidade, que não são as mesmas em que o polo está localizado, mesmo que tenham em seu município de residência a oferta de ensino superior, podem optar por permanecer no curso devido à aprendizagem flexível proporcionada pela EaD (modalidade dos cursos estudados), que não exige presença física de maneira tão frequente como ocorre na educação presencial. Além disso, deve-se considerar também a possibilidade de não existir oferta de cursos de graduação na cidade do aluno ou nas proximidades, o que limitaria as alternativas do estudante para cursar o ensino superior.

Finalmente, no que se refere à idade, variável significativa apenas para o curso de Letras Espanhol, vale destacar que quanto mais tempo de vida o aluno tem, possivelmente, maior o tempo desse aluno fora da vida escolar, o que pode impactar na adaptação do estudante à vida acadêmica. Ainda, salienta-se que pessoas com maior idade podem ter mais compromissos familiares e laborais quando comparadas às pessoas mais jovens, aspectos que, conforme indicado na teoria consultada, podem concorrer com a dedicação do aluno ao curso e levá-lo a evadir. Entretanto, há de se ponderar que, apesar de estudos apontarem que a imaturidade do aluno pode causar a evasão em cursos de graduação, o resultado obtido para o curso de Letras Espanhol demonstra que alunos mais jovens possuem menores probabilidades de evadir.

### **Modelos Preditivos à Evasão**

Ao término do processo para a obtenção das variáveis significativas para a evasão nos cursos estudados, além do *odds ratios*, obtivemos também os coeficientes de cada variável (por curso), os quais apresentamos a seguir.

**Tabela 4**

Coeficientes das variáveis significativas para a evasão nos cursos estudados

Curso	Variável	Pr (>  z )	Estimative
Administração	Intercept	2.69e-16	10.7175
	IAA	<2e-16	-2.1424
	pontosVestibular	0.000155	0.6242
	Reside.na.cidade.polo.[T.sim]	0.011268	0.8399
Administração Pública	Intercept	<2e-16	9.0576
	IAA	<2e-16	-1.3838
Letras Espanhol	Intercept	4.12e-07	15.84605
	IAA	3.12e-08	-2.48115
	Idade	0.0411	0.06631
Matemática	Intercept	1.42e-07	11.0728
	IAA	1.80e-07	-1.6227

*Nota:* Elaborada pelos autores deste artigo.

Tais coeficientes são conhecidos como coeficientes logísticos, os quais compõem do modelo preditivo de evasão para cada curso. O valor apresentado como “intercept” se refere à constante do modelo.

Diante dos coeficientes logísticos que obtivemos, concluímos que a hipótese nula ( $H_0$ ) deveria ser rejeitada, pois encontramos coeficientes  $\beta_j \neq 0$ . Ou seja, **identificamos que existe para cada curso estudado um modelo logístico binário a partir do qual podemos afirmar, com 95% de confiança, a probabilidade de um aluno evadir.**

Considerando o modelo geral da regressão logística (apresentado nos procedimentos metodológicos da pesquisa), construímos o modelo preditivo à evasão para cada um dos cursos. Os quatro modelos construídos são apresentados nas equações que seguem.

### Equação 1

*Modelo preditivo para a evasão no curso de graduação em Administração*

$$p_i = \frac{1}{1 + e^{-(10,7175 - 2,1424IAA_i + 0,6242Vestib_i + 0,8399Res_i)}}$$

*Nota:* elaborada pelos autores deste artigo.

### Equação 2

*Modelo preditivo para a evasão no curso de graduação em Administração Pública*

$$p_i = \frac{1}{1 + e^{9,0576 - 1,3838IAA_i}}$$

*Nota:* elaborada pelos autores deste artigo.

### Equação 3

*Modelo preditivo para a evasão no curso de graduação em Letras Espanhol*

$$p_i = \frac{1}{1 + e^{15,84605 - 2,48115IAA_i + 0,6631idade_i}}$$

Nota: elaborada pelos autores deste artigo.

#### Equação 4

Modelo preditivo para a evasão no curso de graduação em Matemática

$$p_i = \frac{1}{1 + e^{11,0728 - 1,62271AA_i}}$$

Nota: elaborada pelos autores deste artigo.

onde:

$p_i$  = probabilidade de o aluno  $i$  evadir;

$IAA_i$  = IAA do aluno  $i$ ;

$Vestib_i$  = nota (de 0 a 10) do aluno  $i$  no vestibular;

$Res_i$  = informação do aluno  $i$  sobre residir na cidade do polo de apoio presencial.

$idade_i$  = idade do aluno  $i$  em anos completos.

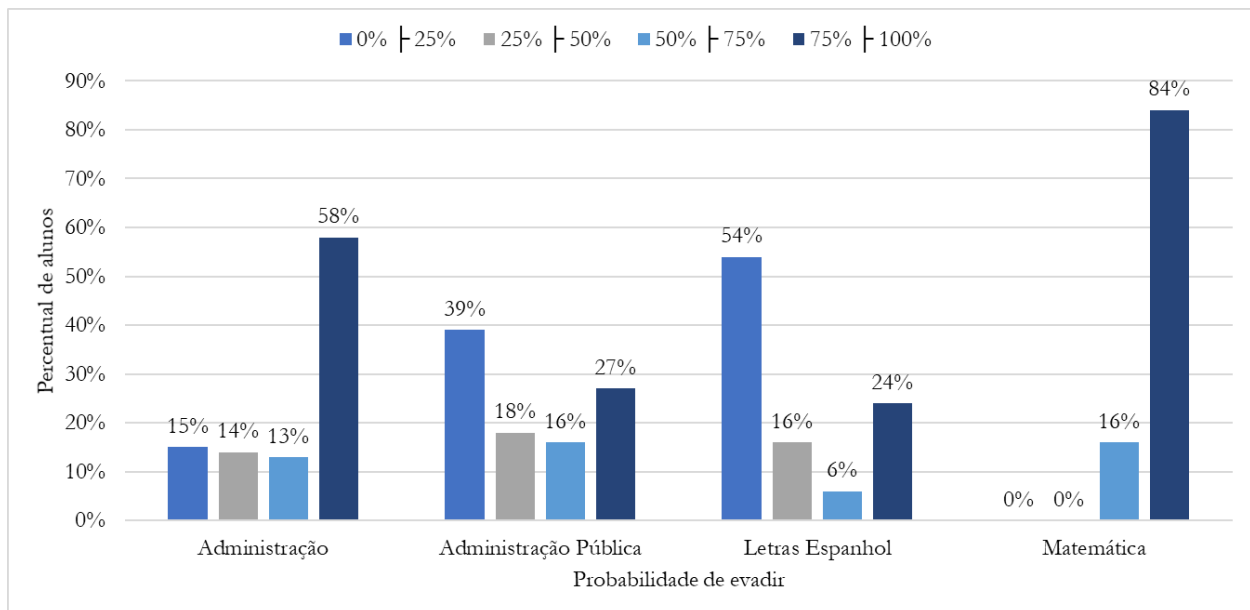
Com as equações definidas, partimos para a aplicação delas à realidade dos cursos no momento da pesquisa.

#### Aplicação e Acurácia dos Modelos Preditivos à Evasão

Para a aplicação dos modelos preditivos, utilizamos os dados dos alunos com matrícula ativa no momento da pesquisa para prever seus *status* finais nos respectivos cursos. Após calcular a probabilidade de evasão de cada aluno, distribuímos as frequências em faixas de probabilidade de evasão, conforme apresentado no gráfico a seguir.

#### Gráfico 1

Distribuição dos alunos com matrícula ativa de acordo com suas probabilidades de evadir



Com a aplicação do modelo, verificamos que:

- a) 58% dos alunos ativos no curso de graduação em Administração tinham probabilidade de evadir igual ou superior a 75%.
- b) 43% dos alunos ativos no curso de graduação em Administração Pública tinham probabilidade de evadir igual ou superior a 50%, enquanto 57% dos alunos no curso tinham probabilidade inferior a 50%.
- c) 30% dos alunos ativos no curso de graduação em Letras Espanhol tinham probabilidade de evadir igual ou superior a 50%, enquanto 70,3% dos alunos no curso tinham probabilidade inferior a 50%.
- d) 84% dos alunos ativos no curso de graduação em Matemática tinham probabilidade de evadir igual ou superior a 75%, enquanto nenhum aluno resultou em probabilidade inferior a 50%.

Com base na probabilidade individual de evasão dos alunos, inferimos para cada um “evasão” ou “permanência”. Consideramos como “permanência” todo aluno que apresentou probabilidade de evasão inferior a 50%. Já para alunos com probabilidade de evasão igual ou superior a 50%, inferimos como *status* final “evasão”.

Assim, da lista de 489 alunos ativos cujos dados foram aplicados ao modelo preditivo do respectivo curso, previmos que 299 evadiriam, enquanto 190 persistiriam e alcançariam a diplomação.

Em junho de 2019, com a finalidade de apurar a acurácia do modelo preditivo desenvolvido, foram analisadas as situações atualizadas dos 489 alunos ativos no momento do levantamento dos dados. Desses, 157 ainda estavam em curso, não apresentando um *status* final de matrícula. Assim, apurou-se a acurácia parcialmente, considerando os 332 alunos que já chegaram a um *status* final.

Ao considerar o total de predições de “evasão” e de “permanência”, 272 das 332 se concretizaram conforme previsto. Ou seja, a aplicação do modelo preditivo possibilitou identificar antecipadamente e de maneira precisa o *status* final de 81,93% dos alunos.

Já no que diz respeito aos alunos evadidos, das 146 evasões que ocorreram até a apuração parcial, 133 puderam ser previstas pelos modelos propostos, ou seja, 9 a cada 10.

Esse resultado demonstra a validade do modelo estatístico desenvolvido e, conseqüentemente, indica que o uso de tais modelos pode ser uma alternativa para que dirigentes universitários façam uma gestão mais efetiva no que diz respeito ao fenômeno da evasão.

### **Limitações da Pesquisa**

Os modelos preditivos à evasão apresentados nesta pesquisa são válidos para a realidade dos cursos contemplados no estudo. No entanto, o método aqui empregado pode ser amplamente utilizado em qualquer IES, conforme destacamos na introdução do artigo. Salientamos ainda que a não participação de todos os cursos inicialmente convidados não comprometeu a realização da pesquisa, visto que as informações de cada curso são independentes e os procedimentos adotados foram aplicados individualmente.

Para a realização do estudo, foram consideradas apenas as matrículas realizadas a partir do ano de 2008. Esse corte temporal se fez necessário, uma vez que, nos anos anteriores, a oferta de cursos de graduação a distância era incipiente na instituição, e os dados daquele período poderiam comprometer a confiabilidade dos modelos preditivos.

Por fim, destacamos que a pesquisa se restringiu à extração de dados ao sistema acadêmico e às informações do vestibular, devido à falta de informações disponíveis no Ambiente Virtual de Aprendizagem (AVA) para o corte temporal realizado, conforme explicitado anteriormente.

### Considerações Finais

A presente pesquisa buscou propor modelos estatísticos preditivos para a gestão da evasão de quatro cursos de graduação a distância de uma Instituição de Ensino Superior brasileira.

A partir da literatura consultada, este estudo trouxe inicialmente duas contribuições teóricas. A primeira diz respeito ao conceito de evasão assumido nesta pesquisa, o qual considera que o aluno evadido é aquele cujo *status* final de matrícula é qualquer outro que não concluinte ou *status* equivalente. Esse conceito foi construído com base na combinação do que propõem autores de renome em conjunto com uma interpretação que faz sentido na prática. Assim, o referido conceito de evasão permitiu considerar que um mesmo aluno pode acumular um histórico acadêmico com diferentes desfechos, independentemente da modalidade de ensino e da dimensão em que ocorre, tal qual acontece na realidade.

Ainda, a partir da literatura sobre a temática evasão, as reflexões aqui apresentadas levaram à necessidade de considerar novas dimensões da evasão estudantil. Além da evasão do curso, da IES ou do sistema educacional, conforme aponta a teoria disponível, somamos duas novas dimensões: a evasão da turma e a evasão do sistema nacional de ensino. A primeira, abarca os casos em que o aluno evade e volta a se matricular no mesmo curso da mesma IES. Por sua vez, a última se refere ao aluno que evade e não volta a se matricular no mesmo nível de ensino no país, mas sim no exterior.

Posteriormente, identificamos os principais aspectos que impactam na evasão dos cursos estudados. Para isso, utilizamos o método estatístico denominado Regressão Logística Binária para identificar, com um nível de significância de 5%, quais variáveis influenciam na evasão dos estudantes. Para tal procedimento, utilizou-se o *software* estatístico R.

O processo foi realizado separadamente para cada curso, no qual identificamos que o Índice de Aproveitamento Acumulado (IAA) (todos os cursos), a pontuação do aluno no vestibular (Administração); o fato de o aluno residir ou não na cidade em que o polo de apoio presencial está localizado (Administração); e a idade dos estudantes (Letras Espanhol) são os aspectos que impactam na evasão dos cursos estudados.

Observamos que, em todos os cursos, a variável IAA figurou como importante para explicar a evasão dos alunos, corroborando a teoria consultada no que diz respeito à necessidade de se monitorar o desempenho dos estudantes, bem como de realizar ações preventivas e corretivas no processo de aprendizagem, de modo que os índices de insucesso nas disciplinas sejam reduzidos. Além disso, o resultado da pesquisa permitiu verificar que uma mesma variável pode produzir efeitos diferentes no fenômeno estudado.

A partir da identificação das variáveis que influenciam na evasão dos cursos pesquisados, construímos os modelos estatísticos preditivos para a evasão. No processo de Regressão Logística Binária realizado no *software* R, foram obtidas as estimativas dos coeficientes logísticos associados a cada variável, que foram utilizados para a obtenção dos modelos apresentados nos resultados do estudo.

Por fim, aplicamos nos cursos pesquisados os modelos estatísticos preditivos de evasão propostos. Essa etapa possibilitou inferir que, dos 489 alunos ativos cujos dados foram aplicados ao modelo preditivo do respectivo curso, 299 evadiriam, enquanto 190 persistiriam e alcançariam a diplomação. Com a apuração parcial da acurácia dos modelos desenvolvidos em junho de 2019, identificamos que, dos 489 alunos, 332 que já haviam chegado a um *status* final.

No que se refere ao total de predições de “evasão” ou “permanência”, 272 se concretizaram conforme previsto. Já em relação ao fenômeno da evasão, das 146 evasões que ocorreram até a apuração parcial, 133 puderam ser previstas pelos modelos desenvolvidos.

Os resultados desta pesquisa permitiram concluir que a utilização de modelos preditivos pode fornecer aos gestores universitários e às políticas públicas educacionais o direcionamento necessário para tomar conhecimento dos principais aspectos que impactam na evasão dos cursos de graduação, assim como apontar, com alto nível de confiança, os alunos com maiores riscos de evasão.

Uma política pública visa a assegurar um direito do cidadão. A educação é um direito universal e a evasão é conflitante com esse direito. Para combater a evasão, é necessário conhecê-la, a fim de direcionar nossos recursos de forma mais eficiente. Esses recursos são finitos, principalmente se pensarmos em relação ao tempo e ao capital disponível, e a decisão de onde direcioná-los pode levar a uma infinidade de possibilidades.

Nesse sentido, entendemos que os modelos preditivos de evasão são mecanismos que podem dar base a políticas públicas. Como os referidos modelos são desenvolvidos a partir de dados, consideramos essencial que eles sejam corretamente coletados, tratados, armazenados, atualizados e distribuídos para que possibilitem que ações voltadas às políticas em comento sejam implementadas. A resposta de quais são os dados essenciais para a elaboração mais ajustada de modelos preditivos decorrerá de investigações futuras. De posse de dados de qualidade, gestores universitários podem aplicar o método aqui demonstrado e, quiçá, dispor de sistema informatizado com interface de gestão para visualizarem os alunos em risco de evasão em tempo real e garantirem ações voltadas para a permanência do aluno antes que ele evada.

A partir de todas as reflexões postas, entendemos que, com uma política pública clara e com informações disponibilizadas, cada instituição poderá realizar a gestão da evasão e permanência de seus estudantes de acordo com sua realidade, que é única. Assim, estudos relacionados à evasão estudantil certamente não encontrarão uma solução geral para esse problema, mas somarão esforços na busca de práticas educacionais e de gestão universitária que possam contribuir para a redução dos índices de evasão.

Como agenda para pesquisas futuras, destacamos a importância de se desenvolver políticas públicas e institucionais com a utilização de ferramentas de gestão que apresentem aos dirigentes probabilidades de evasão dos estudantes, contribuindo para se ter uma visão instantânea da situação dos alunos e viabilizando, a partir dessas informações, a implementação de medidas preventivas à evasão estudantil.

Por fim, salientamos a oportunidade de se investigar quais as variáveis são indispensáveis para o desenvolvimento de modelos preditivos à evasão, uma vez que este e outros estudos têm se limitado ao que aponta a teoria sobre as causas da evasão e ao que se tem de dados disponíveis. Entretanto, entendemos que assumir tal procedimento pode diminuir a confiabilidade dos modelos desenvolvidos.

## Agradecimentos

Agradecemos ao Professor Dr. Pedro Alberto Barbetta pelos esclarecimentos e pelo apoio no uso do *software* R e no uso do método de Regressão Logística, bem como na validação dos resultados estatísticos alcançados. Ainda agradecemos à Secretaria de Estado de Educação de

Santa Catarina que, por meio do Programa UNIEDU, concedeu auxílio financeiro para a realização do estudo por meio de bolsa universitária de pós-graduação.

## Referências

- Adroque, C., & García de Fanelli, A. M. (2018). Gaps in persistence under open-access and tuition-free public higher education policies. *Education Policy Analysis Archives*, 26(126).  
<https://doi.org/10.14507/epaa.26.3497>
- Aguado Franco, J. C. (2017). ¿Pueden los MOOC favorecer el aprendizaje y hacer disminuir las tasas de abandono universitario? *RIED. Revista Iberoamericana de Educación a Distancia*, 20(1), 125.  
<https://doi.org/10.5944/ried.20.1.16684>
- Amaral, M. G. do. (2016). *Mineração de dados aplicada à classificação do risco de evasão de discentes ingressantes em Instituições Federais de Ensino Superior* [Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação - Universidade Federal de Pernambuco].  
[https://repositorio.ufpe.br/bitstream/123456789/19502/1/projeto\\_v26016.pdf](https://repositorio.ufpe.br/bitstream/123456789/19502/1/projeto_v26016.pdf)
- ANDIFES, ABRUEM, & SESU/MEC. (1997). *Diplomação, retenção e evasão nos cursos de graduação em Instituições de Ensino Superior públicas*. MEC - Secretaria de Educação Superior.
- Arendt, J. N. (2013). The effect of public financial aid on dropout from and completion of university education: Evidence from a student grant reform. *Empirical Economics*, 44(3), 1545–1562. <https://doi.org/10.1007/s00181-012-0638-5>
- Assis, L. R. S. de. (2017). *Perfil de evasão no ensino superior brasileiro: Uma abordagem de mineração de dados* [Dissertação de Mestrado, Mestrado Profissional em Computação Aplicada - Universidade de Brasília]. <http://repositorio.unb.br/handle/10482/32139>
- Barbetta, P. A. (2014). *Estatística aplicada às ciências sociais* (9ª ed). Editora da UFSC.
- Belloni, M. L. (1999). *Educação a distância*. Autores Associados.
- Biazus, C. A. (2004). *Sistema de fatores que influenciam o aluno a evadir-se dos cursos de graduação na UFSM e na UFSC: Um estudo no curso de ciências contábeis* [Tese de Doutorado]. Programa de Pós-Graduação em Engenharia de Produção - Universidade Federal de Santa Catarina.
- Brasil. (2014). Lei 13.005 *Plano Nacional de Educação - PNE - Aprova o Plano Nacional de Educação—PNE e dá outras providências* Brasília, Distrito Federal, 2014. Disponível em <http://pne.mec.gov.br/18-planos-subnacionais-de-educacao/543-plano-nacional-de-educacao-lei-n-13-005-2014>.
- Carvajal, C. M., González, J. A., & Sarzoza, S. J. (2018). Variables sociodemográficas y académicas explicativas de la deserción de estudiantes en la facultad de ciencias naturales de la Universidad de Playa Ancha (Chile). *Formación Universitaria*, 11(2), 3–12.  
<https://doi.org/10.4067/S0718-50062018000200003>
- Cislaghi, R. (2008). *Um modelo de Sistema de gestão do conhecimento em um framework para a promoção da permanência discente no ensino de graduação* [Tese de Doutorado]. Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento - Universidade Federal de Santa Catarina.
- Contini, D., Cugnata, F., & Scagni, A. (2018). Social selection in higher education. Enrolment, dropout and timely degree attainment in Italy. *Higher Education*, 75(5), 785–808.  
<https://doi.org/10.1007/s10734-017-0170-9>
- Fávero, L. P. (2015). *Análise de dados: Modelos de Regressão com Excel®, Stata® e SPSS®*. Elsevier.
- Favero, R. V. M. (2006). *Dialogar ou evadir: Eis a questão!: Um estudo sobre a permanência e a evasão na Educação a Distância* [Dissertação de Mestrado, Programa de Pós-Graduação em Educação - Universidade Federal do Rio Grande do Sul].  
<https://www.lume.ufrgs.br/handle/10183/14846>

- Fonseca, J. P. S. da. (2018). *Modelo preditivo de evasão no ensino superior* [Dissertação de Mestrado, Mestrado Profissional em Administração - Universidade de Fortaleza].  
<https://uol.unifor.br/oul/ObraBdtdSiteTrazer.do?method=trazer&ns=true&obraCodigo=104172#>
- Freitas, L. dos R., Martins Filho, S., Ribeiro Júnior, J. I., & Fonseca e Silva, F. (2013). Comparação das funções de ligação logit e probit em regressão binária considerando diferentes tamanhos amostrais. *Enciclopédia Biosfera*, 9.  
<https://www.conhecer.org.br/enciclop/2013b/CIENCIAS%20EXATAS/comparacao.pdf>
- García Aretio, L. (2002). *La educación a distancia: De la teoría a la práctica*. Ariel.
- Gomes, A. A. (1998). *Evasão e evadidos: O discurso dos ex-alunos sobre evasão escolar nos cursos de licenciatura* [Tese de Doutorado, Programa de Pós-graduação em Educação - Universidade Estadual Paulista].  
[https://repositorio.unesp.br/bitstream/handle/11449/102247/gomes\\_aa\\_dr\\_mar.pdf?sequence=1&isAllowed=y](https://repositorio.unesp.br/bitstream/handle/11449/102247/gomes_aa_dr_mar.pdf?sequence=1&isAllowed=y)
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Análise multivariada de dados* (6ª ed). Bookman.
- Hocking, R. R. (1976). The analysis and selection of variables in linear regression. *Biometrics*, 32(1), 1–49. <https://doi.org/10.2307/2529336>
- Hosmer Jr., D. W., & Lemeshow, S. (1989). *Applied logistic regression*. Wiley.
- INEP. (2010). *Resumo técnico—Censo da Educação Superior de 2009* (p. 37). Ministério da Educação.  
[http://download.inep.gov.br/download/superior/censo/2009/resumo\\_tecnico\\_2009.pdf](http://download.inep.gov.br/download/superior/censo/2009/resumo_tecnico_2009.pdf)
- INEP. (2014). *Resumo técnico—Censo da Educação Superior 2012* (p. 138). Ministério da Educação.  
[http://download.inep.gov.br/download/superior/censo/2012/resumo\\_tecnico\\_censo\\_educacao\\_superior\\_2012.pdf](http://download.inep.gov.br/download/superior/censo/2012/resumo_tecnico_censo_educacao_superior_2012.pdf)
- INEP. (2015). *Resumo técnico—Censo da Educação Superior 2013* (p. 82). Ministério da Educação.  
[http://download.inep.gov.br/download/superior/censo/2013/resumo\\_tecnico\\_censo\\_educacao\\_superior\\_2013.pdf](http://download.inep.gov.br/download/superior/censo/2013/resumo_tecnico_censo_educacao_superior_2013.pdf)
- INEP. (2018). *Resumo técnico—Censo da Educação Superior 2016* (p. 101). Ministério da Educação.  
[http://download.inep.gov.br/educacao\\_superior/censo\\_superior/resumo\\_tecnico/resumo\\_tecnico\\_censo\\_da\\_educacao\\_superior\\_2016.pdf](http://download.inep.gov.br/educacao_superior/censo_superior/resumo_tecnico/resumo_tecnico_censo_da_educacao_superior_2016.pdf)
- Jones, W. A., & Braxton, J. M. (2009). Cataloging and comparing institutional efforts to increase student retention rates. *Journal of College Student Retention: Research, Theory & Practice*, 11(1), 123–139. <https://doi.org/10.2190/CS.11.1.g>
- Kori, K., Pedaste, M., Altin, H., Tonisson, E., & Palts, T. (2016). Factors that influence students' motivation to start and to continue studying information technology in Estonia. *IEEE Transactions on Education*, 59(4), 255–262. <https://doi.org/10.1109/TE.2016.2528889>
- Lobo, M. B. de C. M. (2012). Panorama da evasão no ensino superior brasileiro: Aspectos gerais das causas e soluções. In *Evasão no ensino superior brasileiro* (p. 9–58). Associação Brasileira de Mantenedoras de Ensino Superior.  
<https://abmes.org.br/arquivos/publicacoes/Cadernos25.pdf>
- Martín, F. D. F., Tirado, J. L. A., Ortega, S. L., & Díaz, V. A. H. (2011). Prevención del fracaso académico universitario mediante tutoría entre iguales. *Revista Latinoamericana de Psicología*, 43(1), 59–71.
- Moore, M. G., & Kearsley, G. (2007). *Educação a distância: Uma visão integrada*. Thomson Learning.
- Munizaga, F. R., Cifuentes, M. B., & Beltrán, A. (2018). Retención y abandono estudiantil en la educación superior universitaria en América Latina y el Caribe: Una revisión sistemática. *Education Policy Analysis Archives*, 26(61). <https://doi.org/10.14507/epaa.26.3348>



- Nascimento, T. P., & Esper, A. (2014). Evasão em cursos de educação continuada a distância: Um estudo na Escola Nacional de Administração Pública. *Revista do Serviço Público*, 60, 159–173. <https://doi.org/10.21874/rsp.v60i2.19>
- Noronha, A. B., Carvalho, B. M., & Santos, F. F. F. dos. (2001). *Perfil dos alunos evadidos da Faculdade de Economia, Administração e Contabilidade campus Ribeirão Preto e avaliação do tempo de titulação dos alunos atualmente matriculados*. <http://nupps.usp.br/downloads/docs/dt0101.pdf>
- Pacheco, A. S. V. (2010). *Evasão e permanência dos estudantes de um curso de Administração do sistema Universidade Aberta do Brasil: Uma teoria fundamentada em fatos e na gestão do conhecimento* [Tese de Doutorado, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento - Universidade Federal de Santa Catarina]. <http://btd.egc.ufsc.br/wp-content/uploads/2010/06/Andressa-Pacheco.pdf>
- Paredes, A. S. (1994). *A evasão do terceiro grau em Curitiba*. Núcleo de pesquisas sobre ensino superior – NUPES. Universidade de São Paulo.
- Pereira, F. C. B. (2003). *Determinantes da evasão de alunos e os custos ocultos para as instituições de ensino superior: Uma aplicação na Universidade do Extremo Sul Catarinense* [Tese de Doutorado, Programa de Pós-graduação em Engenharia de Produção - Universidade Federal de Santa Catarina]. <http://repositorio.ufsc.br/xmlui/handle/123456789/86403>
- Pérez, A. M., Escobar, C. R., Toledo, M. R., Gutierrez, L. B., & Reyes, G. M. (2018). Modelo de predicción de la deserción estudiantil de primer año en la Universidad Bernardo O'Higgins. *Educación e Pesquisa*, 44(0). <https://doi.org/10.1590/s1678-4634201844172094>
- Poignant, R. (1976). *Curso de planejamento da educação*. (uls.28149). Catálogo do Sistema de Bibliotecas da UFSC. <http://search.ebscohost.com/login.aspx?direct=true&db=cat07205a&AN=uls.28149&site=eds-live>
- Princiotta, D., & Reyna, R. (2009). *Achieving graduation for all: A governor's guide to dropout -prevention and recovery* (p. 48). NGA Center for Best Practices. <https://eric.ed.gov/?id=ED507071>
- PUCRS, P. U. C. do R. G. do S. ([s.d.]). *PROJETO ALFA GUIA*. PROJETO ALFA GUIA. Recuperado 2 de julho de 2019, de <http://www.pucrs.br/humanidades/pesquisa/projetos/projeto-alfa-guia/>
- Rovaris Neto, E. (2002). *E-BAYES - Sistema especialista para a análise da evasão discente de cursos de graduação no ensino superior* [Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação - Universidade Federal de Santa Catarina]. <https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/84446/PGCC0204-D.pdf?sequence=1&isAllowed=y>
- Sánchez-Hernández, G., Barboza-Palomino, M., & Castilla-Cabello, H. (2017). Análisis de la deserción y los factores asociados a la permanencia estudiantil en una universidad peruana. *Actualidades Pedagógicas*, 69, 169–191. <https://doi.org/10.19052/ap.4075>
- Santos Junior, J. da S., & Real, G. C. M. (2017). A evasão na educação superior: O estado da arte das pesquisas no Brasil a partir de 1990. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 22(2), 385–402. <https://doi.org/10.1590/s1414-40772017000200007>
- Schmitt, J. (2018). *Construção de uma escala de propensão à evasão estudiantil em cursos de graduação* [Tese de Doutorado]. Universidade Federal de Santa Catarina.
- Silva, F. C. D., Cabral, T. L. de O., & Pacheco, A. S. V. (2016). Evasão em cursos de graduação: Uma análise a partir do censo da educação superior brasileira. In *XVI Colóquio Internacional de Gestão Universitária* (p. 16). [https://repositorio.ufsc.br/bitstream/handle/123456789/171098/OK%20-%20101\\_00387%20OK.pdf?sequence=1&isAllowed=y](https://repositorio.ufsc.br/bitstream/handle/123456789/171098/OK%20-%20101_00387%20OK.pdf?sequence=1&isAllowed=y)
- Simonson, M. (2006). *Educación abierta y a distancia*. UOC.

- Speller, P., Robl, F., & Meneghel, M. M. (Orgs.). (2012). *Desafios e perspectivas da educação superior brasileira para a próxima década*. UNESCO, CNE, MEC.
- Suiter, D., & Pacheco, A. S. V. (2019). Gestão da reocupação de vagas no ensino superior público. *Revista Gestão Universitária*, 10(2018). <http://www.gestaouniversitaria.com.br/artigos-cientificos/gestao-da-reocupacao-de-vagas-no-ensino-superior-publico>
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89-125 <https://doi.org/10.3102/00346543045001089>
- Tinto, V. (1989). Definir la desercion: Una cuestion de perspectiva. *Revista de la Educación Superior*, Número 71, 1-9

## Apêndice A

### Variáveis Seleccionadas para a Elaboração dos Modelos Preditivos e suas Categorias

**Tabela A1**

Variáveis seleccionadas para a elaboração dos modelos preditivos e suas categorias

Variável	Categorias
Sexo	Feminino; Masculino.
Cor	Amarela; Branca; Indígena; Parda; Preta; Não informada.
Estado civil	Casado; Divorciado; Separado Judicialmente; Solteiro; União Estável; Viúvo; Não Informado; Outros.
UF residência	Região Centro-oeste; Região Nordeste; Região Norte; Região Sudeste; Região Sul.
UF polo	Região Centro-oeste; Região Nordeste; Região Norte; Região Sudeste; Região Sul.
Reside cidade polo	Sim; Não.
Categoria de ingresso	Classificação geral; Programa de Ações Afirmativas.
Renda familiar	não informado; de 1 a 3 salários; de 4 a 8 salários; de 9 a 12 salários; de 13 a 20 salários; acima de 20 salários.
Tamanho da família	não informado; de 1 a 2 pessoas; de 3 a 4 pessoas; de 5 a 6 pessoas; de 7 a 8 pessoas; 9 ou mais pessoas.
Tempo de deslocamento até o polo	não informado; até 30 minutos; entre 30 minutos e 1 hora; entre 1 e 2 horas; acima de 2 horas.
Onde estudou a maior parte do Ensino Médio	não informado; escola pública; escola particular.
Experiência no Ensino Superior	não informado; nunca ingressou; já concluiu; ingressou, mas não concluiu; está cursando.
Experiência na EAD	não informado; sim; não.
Frequência de uso do computador	não informado; não utiliza; todos os dias; raramente; uma vez por semana.
Local de acesso à internet	não informado; casa; trabalho; escola; amigo; parente; outros; não tenho acesso.
Tipo de conexão à internet	não informado; não tenho; linha discada; ADSL; a cabo; outra.
Nível de conhecimento para o uso do computador e da internet	não informado; nenhum; pouco; médio; aprofundado.

*Nota:* elaborada pelos autores deste artigo a partir de dados dos sistemas institucionais.

## **Sobre o Autores**

### **Fernanda Cristina da Silva**

Universidade Federal de Santa Catarina - UFSC

fernanda.silva.f.cs@posgrad.ufsc.br

ORCID: <https://orcid.org/0000-0001-9059-0689>

Doutoranda e mestra em Administração pela Universidade Federal de Santa Catarina - UFSC, na linha de pesquisa Gestão Universitária. Licenciada em Matemática pela UFSC. Membro do Instituto de Pesquisas e Estudos em Administração Universitária – INPEAU – UFSC. Pesquisa temas relacionados à evasão, permanência, egressos, educação a distância e gestão universitária.

### **Thiago Luiz de Oliveira Cabral**

Universidade Federal de Santa Catarina - UFSC

thiago.cabral@ufsc.br

ORCID: <https://orcid.org/0000-0003-2999-3512>

Doutorando em Administração pela Universidade Federal de Santa Catarina - UFSC. Bacharel e Mestre em Administração pela Universidade Federal de Santa Catarina. É Servidor técnico-administrativo da UFSC. Membro do Instituto de Pesquisas e Estudos em Administração Universitária – INPEAU – UFSC. Pesquisa temas relacionados à evasão, permanência, egressos, internacionalização e gestão universitária.

### **Andressa Sasaki Vasques Pacheco**

Universidade Federal de Santa Catarina - UFSC

andressa.pacheco@ufsc.br

ORCID: <https://orcid.org/0000-0002-7403-5148>

Professora do Departamento de Ciências da Administração da Universidade Federal de Santa Catarina - UFSC. Doutora em Engenharia e Gestão do Conhecimento, mestra e bacharel em Administração pela Universidade Federal de Santa Catarina. Professora do Programa de Pós-graduação em Administração e do Programa de Pós-graduação em Administração Universitária. Coordenadora do curso de especialização em Gestão Pública. Experiência na área de Administração, com ênfase em Gestão Universitária. Educação a distância e Administração de Materiais, atuando principalmente nos seguintes temas: evasão, egressos, avaliação, empreendedorismo, estoques e compras. Participante dos grupos de pesquisa INPEAU - Instituto de Pesquisas e Estudos em Administração Universitária e NEOGAP - Núcleo Estudos e Observação de Gestão, Aprendizagem e Pessoas.

# archivos analíticos de políticas educativas

Volume 28 Número 149

19 de outubro 2020

ISSN 1068-2341



Los/as lectores/as pueden copiar, mostrar, distribuir, y adaptar este artículo. Cualquier otro uso debe ser aprobado en conjunto por el autor/es, o AAPE/EPAA. La sección en español para Sud América de AAPE/EPAA es publicada por el *Mary Lou Fulton Teachers College, Arizona State University* y la *Universidad de San Andrés* de Argentina. Los artículos que aparecen en AAPE son indexados en CIRC (Clasificación Integrada de Revistas Científicas, España) DIALNET (España), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), PubMed, QUALIS A1 (Brazil), Redalyc, SCImago Journal Rank, SCOPUS, SOCOLAR (China). Creative Commons se encuentran en <https://creativecommons.org/licenses/by-sa/4.0/>. Cualquier otro uso debe ser aprobado en conjunto por el autor/es, o AAPE/EPAA. La sección en español para Sud América de AAPE/EPAA es publicada por el *Mary Lou Fulton Teachers College, Arizona State University* y la *Universidad de San Andrés* de Argentina. Los artículos que aparecen en AAPE son indexados en CIRC (Clasificación Integrada de Revistas Científicas, España) DIALNET (España), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), PubMed, QUALIS A1 (Brazil), Redalyc, SCImago Journal Rank, SCOPUS, SOCOLAR (China).

Por errores y sugerencias contacte a [Fischman@asu.edu](mailto:Fischman@asu.edu)

Síguenos en EPAA's Facebook comunidad at <https://www.facebook.com/EPAAAPE> y en Twitter feed @epaa\_aape.

arquivos analíticos de políticas educativas  
conselho editorial

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Editoras Coordenadoras: **Marcia Pletsch, Sandra Regina Sales** (Universidade Federal Rural do Rio de Janeiro)

Editores Associadas: **Andréa Barbosa Gouveia** (Universidade Federal do Paraná), **Kaizo Iwakami Beltrao**, (EBAPE/FGVI), **Sheizi Calheira de Freitas** (Federal University of Bahia), **Maria Margarida Machado**, (Federal University of Goiás / Universidade Federal de Goiás), **Gilberto José Miranda**, (Universidade Federal de Uberlândia)

**Almerindo Afonso**

Universidade do Minho  
Portugal

**Alexandre Fernandez Vaz**

Universidade Federal de Santa  
Catarina, Brasil

**José Augusto Pacheco**

Universidade do Minho, Portugal

**Rosanna Maria Barros Sá**

Universidade do Algarve  
Portugal

**Regina Célia Linhares Hostins**

Universidade do Vale do Itajaí,  
Brasil

**Jane Paiva**

Universidade do Estado do Rio de  
Janeiro, Brasil

**Maria Helena Bonilla**

Universidade Federal da Bahia  
Brasil

**Alfredo Macedo Gomes**

Universidade Federal de Pernambuco  
Brasil

**Paulo Alberto Santos Vieira**

Universidade do Estado de Mato  
Grosso, Brasil

**Rosa Maria Bueno Fischer**

Universidade Federal do Rio Grande  
do Sul, Brasil

**Jefferson Mainardes**

Universidade Estadual de Ponta  
Grossa, Brasil

**Fabiany de Cássia Tavares Silva**

Universidade Federal do Mato  
Grosso do Sul, Brasil

**Alice Casimiro Lopes**

Universidade do Estado do Rio de  
Janeiro, Brasil

**Jader Janer Moreira Lopes**

Universidade Federal Fluminense e  
Universidade Federal de Juiz de Fora,  
Brasil

**António Teodoro**

Universidade Lusófona  
Portugal

**Suzana Feldens Schwertner**

Centro Universitário Univates  
Brasil

**Debora Nunes**

Universidade Federal do Rio Grande  
do Norte, Brasil

**Lílian do Valle**

Universidade do Estado do Rio de  
Janeiro, Brasil

**Geovana Mendonça Lunardi**

**Mendes** Universidade do Estado de  
Santa Catarina

**Alda Junqueira Marin**

Pontifícia Universidade Católica de  
São Paulo, Brasil

**Alfredo Veiga-Neto**

Universidade Federal do Rio Grande  
do Sul, Brasil

**Flávia Miller Naethe Motta**

Universidade Federal Rural do Rio de  
Janeiro, Brasil

**Dalila Andrade Oliveira**

Universidade Federal de Minas  
Gerais, Brasil

## archivos analíticos de políticas educativas consejo editorial

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Coordinador (Español/Latinoamérica): **Ignacio Barrenechea, Axel Rivas** (Universidad de San Andrés)

Editor Coordinador (Español/Norteamérica): **Armando Alcántara Santuario** (Universidad Nacional Autónoma de México)

Editor Coordinador (Español / España): **Antonio Luzon** (Universidad de Granada)

Editores Asociados: **Felicitas Acosta** (Universidad Nacional de General Sarmiento), **Jason Beech** (Universidad de San Andrés), **Angelica Buendia**, (Metropolitan Autonomous University), **Alejandra Falabella** (Universidad Alberto Hurtado, Chile), **Veronica Gottau** (Universidad Torcuato Di Tella), **Carolina Guzmán-Valenzuela** (Universidad de Chile), **Cesar Lorenzo Rodriguez Uribe** (Universidad Marista de Guadalajara)

**María Teresa Martín Palomo** (University of Almería), **María Fernández Mellizo-Soto** (Universidad Complutense de Madrid), **Tiburcio Moreno** (Autonomous Metropolitan University-Cuajimalpa Unit), **José Luis Ramírez**, (Universidad de Sonora), **Maria Veronica Santelices** (Pontificia Universidad Católica de Chile)

**Claudio Almonacid**

Universidad Metropolitana de Ciencias de la Educación, Chile

**Miguel Ángel Arias Ortega**

Universidad Autónoma de la Ciudad de México

**Xavier Besalú Costa**

Universitat de Girona, España

**Xavier Bonal Sarro** Universidad Autónoma de Barcelona, España

**Antonio Bolívar Boitia**

Universidad de Granada, España

**José Joaquín Brunner** Universidad Diego Portales, Chile

**Damián Canales Sánchez**

Instituto Nacional para la Evaluación de la Educación, México

**Gabriela de la Cruz Flores**

Universidad Nacional Autónoma de México

**Marco Antonio Delgado Fuentes**

Universidad Iberoamericana, México

**Inés Dussel**, DIE-CINVESTAV, México

**Pedro Flores Crespo** Universidad Iberoamericana, México

**Ana María García de Fanelli**

Centro de Estudios de Estado y Sociedad (CEDES) CONICET, Argentina

**Juan Carlos González Faraco**

Universidad de Huelva, España

**María Clemente Linuesa**

Universidad de Salamanca, España

**Jaume Martínez Bonafé**

Universitat de València, España

**Alejandro Márquez Jiménez**

Instituto de Investigaciones sobre la Universidad y la Educación, UNAM, México

**María Guadalupe Olivier Tellez**,

Universidad Pedagógica Nacional, México

**Miguel Pereyra** Universidad de

Granada, España

**Mónica Pini** Universidad Nacional de San Martín, Argentina

**Omar Orlando Pulido Chaves**

Instituto para la Investigación Educativa y el Desarrollo Pedagógico (IDEP)

**José Ignacio Rivas Flores**

Universidad de Málaga, España

**Miriam Rodríguez Vargas**

Universidad Autónoma de Tamaulipas, México

**José Gregorio Rodríguez**

Universidad Nacional de Colombia, Colombia

**Mario Rueda Beltrán** Instituto de Investigaciones sobre la Universidad y la Educación, UNAM, México

**José Luis San Fabián Maroto**

Universidad de Oviedo, España

**Jurjo Torres Santomé**, Universidad de la Coruña, España

**Yengny Marisol Silva Laya**

Universidad Iberoamericana, México

**Ernesto Treviño Ronzón**

Universidad Veracruzana, México

**Ernesto Treviño Villarreal**

Universidad Diego Portales Santiago, Chile

**Antoni Verger Planells**

Universidad Autónoma de Barcelona, España

**Catalina Wainerman**

Universidad de San Andrés, Argentina

**Juan Carlos Yáñez Velazco**

Universidad de Colima, México

education policy analysis archives  
editorial board

Lead Editor: **Audrey Amrein-Beardsley** (Arizona State University)

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Associate Editors: **Melanie Bertrand, David Carlson, Lauren Harris, Danah Henriksen, Eugene Judson, Mirka Koro-Ljungberg, Daniel Liou, Scott Marley, Molly Ott, Iveta Silova** (Arizona State University)

**Madelaine Adelman** Arizona State University

**Cristina Alfaro**  
San Diego State University

**Gary Anderson**  
New York University

**Michael W. Apple**  
University of Wisconsin, Madison

**Jeff Bale** University of Toronto, Canada

**Aaron Benavot** SUNY Albany

**David C. Berliner**  
Arizona State University

**Henry Braun** Boston College

**Casey Cobb**  
University of Connecticut

**Arnold Danzig**  
San Jose State University

**Linda Darling-Hammond**  
Stanford University

**Elizabeth H. DeBray**  
University of Georgia

**David E. DeMatthews**  
University of Texas at Austin

**Chad d'Entremont** Rennie Center for Education Research & Policy

**John Diamond**  
University of Wisconsin, Madison

**Matthew Di Carlo**  
Albert Shanker Institute

**Sherman Dorn**  
Arizona State University

**Michael J. Dumas**  
University of California, Berkeley

**Kathy Escamilla**  
University of Colorado, Boulder

**Yariv Feniger** Ben-Gurion University of the Negev

**Melissa Lynn Freeman**  
Adams State College

**Rachael Gabriel**  
University of Connecticut

**Amy Garrett Dikkers** University of North Carolina, Wilmington

**Gene V Glass**  
Arizona State University

**Ronald Glass** University of California, Santa Cruz

**Jacob P. K. Gross**  
University of Louisville

**Eric M. Haas** WestEd

**Julian Vasquez Heilig** California State University, Sacramento

**Kimberly Kappler Hewitt**  
University of North Carolina

Greensboro

**Aimee Howley** Ohio University

**Steve Klees** University of Maryland

**Jaekyung Lee** SUNY Buffalo

**Jessica Nina Lester**  
Indiana University

**Amanda E. Lewis** University of Illinois, Chicago

**Chad R. Lochmiller** Indiana University

**Christopher Lubienski** Indiana University

**Sarah Lubienski** Indiana University

**William J. Mathis**  
University of Colorado, Boulder

**Michele S. Moses**  
University of Colorado, Boulder

**Julianne Moss**  
Deakin University, Australia

**Sharon Nichols**  
University of Texas, San Antonio

**Eric Parsons**  
University of Missouri-Columbia

**Amanda U. Potterton**  
University of Kentucky

**Susan L. Robertson**  
Bristol University

**Gloria M. Rodriguez**  
University of California, Davis

**R. Anthony Rolle**  
University of Houston

**A. G. Rud**  
Washington State University

**Patricia Sánchez** University of Texas, San Antonio

**Janelle Scott** University of California, Berkeley

**Jack Schneider** University of Massachusetts Lowell

**Noah Sobe** Loyola University

**Nelly P. Stromquist**  
University of Maryland

**Benjamin Superfine**  
University of Illinois, Chicago

**Adai Tefera**  
Virginia Commonwealth University

**A. Chris Torres**  
Michigan State University

**Tina Trujillo**  
University of California, Berkeley

**Federico R. Waitoller**  
University of Illinois, Chicago

**Larisa Warhol**  
University of Connecticut

**John Weathers** University of Colorado, Colorado Springs

**Kevin Welner**  
University of Colorado, Boulder

**Terrence G. Wiley**  
Center for Applied Linguistics

**John Willinsky**  
Stanford University

**Jennifer R. Wolgemuth**  
University of South Florida

**Kyo Yamashiro**  
Claremont Graduate University

**Miri Yemini**  
Tel Aviv University, Israel