

# **MODELAGEM MATEMÁTICA PARA MINIMIZAÇÃO DAS TAXAS DE EVASÃO DE ALUNOS NA CRIAÇÃO DE NOVOS CURSOS: ABORDAGEM VIA ANÁLISE DE REGRESSÃO**

**MATHEMATICAL MODEL FOR MINIMIZING STUDENT EVASION IN  
THE CREATION OF NEW COURSES: APPROACH VIA RERESSION ANALYSIS**

# MODELAGEM MATEMÁTICA PARA MINIMIZAÇÃO DAS TAXAS DE EVASÃO DE ALUNOS NA CRIAÇÃO DE NOVOS CURSOS: ABORDAGEM VIA ANÁLISE DE REGRESSÃO

## MATHEMATICAL MODEL FOR MINIMIZING STUDENT EVASION IN THE CREATION OF NEW COURSES: APPROACH VIA RERESSION ANALYSIS

Patrick Rogers De Souza<sup>1</sup> • Júlio César Prado Ribeiro<sup>2</sup>  
Kátia Celina da Silva Richetto<sup>3</sup> • Cristie Diego Pimenta<sup>4</sup>

Data de recebimento: 24/10/2023

Data de aceite: 06/11/2023

<sup>1</sup> É especialista em Lean Manufacturing, padronização e processos de melhoria contínua. Term formação em Engenharia da Qualidade e mais de uma década de experiência como consultor e professor no SENAC e SENAI.

E-mail: patrick.souza@sp.senai.br

<sup>2</sup> Possui Técnico em Mecânica pela Escola Senai Félix Guisard de Taubaté, graduação em Engenharia Mecânica pela Universidade de Taubaté (UNITAU) e Pós-graduado em Engenharia da Qualidade pela Faculdade de Tecnologia SENAI Félix Guisard de Taubaté.

E-mail: juliocesar\_p\_r@hotmail.com

<sup>3</sup> Possui graduação em Escola de Engenharia de Lorena (EEL/USP), mestrado em Engenharia Mecânica na Faculdade de Engenharia de Guaratinguetá (UNESP), doutorado em Engenharia de Materiais pela EEL/USP, especialização em Educação a Distância e especialização em Engenharia de Segurança do Trabalho (UNISA). Pós-Doutoranda em Neurociência Aplicada a Educação, cursando Especialização em Engenharia de Segurança do Trabalho. Atualmente é professora assistente doutora da Universidade de Taubaté do Mestrado Profissional em Educação.

E-mail: katia.csrichetto@unitau.br

<sup>4</sup> Pós-doutorado em Estatística Avançada (UNESP). Doutor em Engenharia Mecânica com especialidade em Projeto de Experimentos (UNESP) e Mestre em Engenharia Mecânica (UNESP). Pós-graduado em Engenharia da Qualidade (USP), Projetista de Produto (Bacharelado), Black Belt em Lean Six Sigma e Curso de Extensão em Statistics for Business and Industry no British Columbia Institute of Technology (Vancouver- Canadá).

E-mail: pimentadiego@yahoo.com.br

## RESUMO

A evasão escolar em cursos Técnicos e Profissionalizantes representa um desafio substancial, causando impactos negativos no desenvolvimento educacional e gerando perdas financeiras para instituições. Apesar de sua relevância, há uma lacuna na literatura sobre abordagens preditivas para combater a evasão. Este trabalho aborda a modelagem matemática e a análise de regressão para compreender e antecipar taxas de evasão em novos cursos. Utilizando regressão linear múltipla em dados de 688 turmas de 2017-2018 de uma escola técnica, identificou-se correlações significativas entre variáveis e evasão, com ênfase no período das aulas como fator crítico. O estudo conclui ressaltando a relevância de considerar o horário das aulas em estratégias de retenção e destaca a aplicação de métodos matemáticos na educação. Como resultado, foi gerada uma fórmula, via Minitab, para prever evasão com precisão, oferecendo uma ferramenta valiosa para instituições educacionais.

**Palavras-chave:** Desenvolvimento Regional Sustentável; Educação Superior Comunitária; Santa Catarina.

## ABSTRACT

School dropout in Technical and Vocational courses represents a significant challenge, leading to negative impacts on educational development and causing financial losses for institutions. Despite its importance, there exists a gap in literature regarding predictive approaches to address dropout. This study explores mathematical modeling and regression analysis to understand and anticipate dropout rates in new courses. Using multiple linear regression on data from 688 classes from 2017-2018 from a technical school, significant correlations between variables and dropout were identified, with emphasis on the class timings as a critical factor. The study concludes by highlighting the importance of considering class timings in retention strategies and underscores the application of mathematical methods in education. As an outcome, a formula was generated via Minitab to predict dropout accurately, offering a valuable tool for educational institutions.

**Keywords:** School dropout. Regression analysis. Dropout rates. Financial impact. Completion likelihood.

## INTRODUÇÃO

O desafio da evasão escolar, particularmente em cursos de Ensino Técnico e Profissionalizante, tem desencadeado uma série de discussões acadêmicas e práticas em torno de sua natureza e das estratégias para sua mitigação. A evasão não só mina o potencial educacional e desenvolvimento de habilidades dos estudantes, mas também implica em perdas financeiras para as instituições.

Vários estudos têm se concentrado nas causas da evasão. *Morais et al.* (2021), por exemplo, identificaram até 14 fatores possíveis que podem contribuir para a evasão, incluindo a baixa renda e desigualdades sociais. *Anzolin e Krelig* (2013) trouxeram um olhar específico para o contexto do SENAI/SC, enquanto *Queiroga et al.* (2018) lançaram luz sobre a evasão em cursos técnicos a distância.

Em contrapartida, existem limitados estudos que buscam abordar o problema da evasão de maneira preditiva, propondo ferramentas ou modelos que possam prever a evasão antes que ela ocorra. A capacidade de prever a evasão pode levar a estratégias proativas, evitando o fenômeno e garantindo uma melhor retenção de estudantes.

Neste trabalho, o olhar se voltou para a aplicação de modelagem matemática e análise de regressão como ferramentas para entender e, mais crucialmente, prever as taxas de evasão escolar em novos cursos. O foco é utilizar estas ferramentas para auxiliar na criação de cursos com maior potencial de retenção de estudantes.

Os dados analisados neste estudo se concentram em variáveis como duração do curso, custo da hora-aula, mês de início, e período das aulas, que têm sido destacados em estudos anteriores como

fatores críticos na evasão (Lino, 2020). O objetivo geral deste trabalho é oferecer uma solução prática e eficaz para instituições educacionais, permitindo que elas antecipem e abordem proativamente os desafios da evasão escolar.

## REVISÃO DA LITERATURA

A evasão escolar tem sido um ponto focal em diversas pesquisas nos últimos anos, dada a sua importância no contexto educacional e socioeconômico. Moraes *et al.* (2021) exploraram a influência da desigualdade social na evasão escolar, identificando até 14 fatores possíveis que podem contribuir para tal evasão, com destaque para a baixa renda. O estudo enfatizou o potencial do Machine Learning, particularmente os modelos de regressão, como ferramentas valiosas na identificação desses fatores.

Lino (2020) investigaram as principais causas e consequências da evasão escolar, identificando quatro áreas principais de influência: a família, os alunos, a escola e as desigualdades sociais. Este estudo destaca a exclusão social resultante de vários fatores e sugere uma abordagem holística, envolvendo familiares, escolas e políticas públicas.

Anzolin e Krelig (2013) focaram especificamente no SENAI/SC, identificando motivos principais de evasão como mudança de curso e necessidades financeiras, que levam os estudantes a trabalhar, tornando-se relevante para a sua pesquisa no contexto do SENAI.

Sadoyama *et al.* (2020) investigaram a evasão escolar no ensino superior, com foco particular nos desafios encontrados em ambientes presenciais e a distância. Eles identificaram vários fatores, como a escolha inadequada do curso e as dificuldades em equilibrar trabalho e estudos.

Queiroga *et al.* (2018) abordaram o problema da evasão em cursos técnicos a distância, usando mineração de dados para identificar estudantes em risco de evasão com base em suas interações em ambientes virtuais de aprendizagem.

Silva e Sampaio (2022) realizaram uma revisão sistemática, analisando políticas de permanência estudantil no ensino superior, focando principalmente nas avaliações dos impactos dessas políticas na evasão.

Silva *et al.* (2019) se concentraram em ações de retenção em cursos de Ciências Contábeis, identificando várias estratégias e práticas que podem ajudar a minimizar a evasão. Colpani (2018) explorou a aplicação de técnicas de Mineração de Dados, utilizando indicadores educacionais do INEP,

para compreender e abordar a evasão escolar.

O trabalho de Oliveira *et al.* (2022) propôs a criação de um banco de dados dimensional para analisar características relacionadas à evasão escolar, usando dados da Plataforma Nilo Peçanha. Prim e Fávero (2013) focaram em cursos tecnológicos em Blumenau/SC, identificando a evasão principalmente entre estudantes do gênero masculino, entre 18 e 24 anos, e no início de seus cursos.

Oliveira e Carmo (2021) realizaram uma revisão integrativa no contexto do PROEJA, buscando entender os desafios da evasão e destacando a necessidade de preparação adequada dos professores para o ensino de jovens e adultos.

Hoyos Osorio e Daza Santacoloma (2023) destacaram a importância de identificar precocemente estudantes em risco de evasão no ensino superior e propuseram um sistema de alerta precoce para essa finalidade.

Por último, Pimenta *et al.* (2016) compararam dois métodos de otimização para modelar processos em arames de aço, demonstrando a importância da aplicação de metodologias científicas rigorosas em diferentes campos, incluindo a educação.

## MÉTODO - COLETA DE DADOS

Os dados foram coletados a partir de diferentes instituições de Ensino Técnico e Profissionalizante. Ao todo, conseguimos os registros de 688 cursos em diferentes instituições. Foram reunidos registros históricos dos anos de 2017 e 2018 relacionados à inscrição, progresso e conclusão ou evasão dos estudantes nos respectivos cursos.

## VARIÁVEIS

Foram consideradas as seguintes variáveis para a análise:

Carga Horária do Curso: dispensada a variável devido ao registro obtido de todos os cursos com a mesma quantidade de horas, ou seja, 160 horas no total.

Duração do Curso: em dias, variável muito impactante devido a carga de horas diária refletindo em cursos com duração desde 38 dias até 336 dias.

Dias da Semana: avaliar o impacto da quantidade de dias de aula na semana. Para o cálculo foram utilizados os valores conforme tabela 1:

**Tabela 1 | Transformação de variáveis discretas dias da semana em contínuas**

Variável dias da semana	Quantidade de dias
aos Sábados	1
às Segunda-feiras	2
de 2ª e 3ª feira	3
de 2ª e 4ª feira	3
de 2ª e 5ª feira	3
de 2ª e 6ª feira	3
de 2ª, 3ª e 4ª feira	4
de 2ª, 3ª e 5ª feira	4
de 2ª, 3ª e 6ª feira	4
de 2ª, 3ª, 4ª e 5ª feira	5
de 2ª, 3ª, 4ª e 6ª feira	5
de 2ª, 3ª, 4ª, 5ª e 6ª feira	6
de 2ª, 3ª, 5ª e 6ª feira	5
de 2ª, 4ª e 5ª feira	4
de 2ª, 4ª e 6ª feira	4
de 2ª, 5ª e 6ª feira	4
de 3ª e 4ª feira	3
de 3ª e 5ª feira	3
de 3ª, 4ª e 5ª feira	4
de 3ª, 4ª e 6ª feira	4
de 3ª, 4ª, 5ª e 6ª feira	5
de 3ª, 5ª e 6ª feira	4
de 4ª e 5ª feira	3
de 4ª e 6ª feira	3
de 4ª, 5ª e 6ª feira	4
de 5ª e 6ª feira	3
de 5ª, 6ª e Sábado	7

Fonte: Os Autores (2023)

Custo da Hora-Aula: Valor financeiro correspondente a uma hora de aula.

Mês de Início: Mês no qual os cursos começaram visando entender a influência de períodos específicos do ano na evasão. Para o cálculo foram utilizados os valores conforme tabela 2:

**Tabela 2 | Transformação de variáveis discretas mês de início em contínuas**

Mês de Início	Nº Referência
Janeiro	1
Fevereiro	2
Março	3
Abril	4
Mai	5
Junho	6
Julho	7
Agosto	8
Setembro	9
Outubro	10
Novembro	11

Fonte: Os Autores (2023)

Período das Aulas: Se as aulas eram ministradas durante a Manhã, Tarde, Noite, Integral ou Vespertino. Para o cálculo foram utilizados os valores conforme tabela 3:4

**Tabela 3 | Transformação de variáveis discretas período em contínuas**

Mês de Início	Nº Referência
Janeiro	1
Fevereiro	2
Março	3
Abril	4
Maio	5

Fonte: Os Autores (2023)

Quantidade de alunos matriculados: Avaliar o impacto sobre a evasão com turmas reduzidas de 10 até salas em grande volume até 32 alunos.

### ANÁLISE DE REGRESSÃO

Utilizou-se a técnica de análise de regressão para entender o relacionamento entre as variáveis citadas e a evasão escolar. A regressão permitiu prever as taxas de evasão com base nessas variáveis. As análises foram realizadas utilizando o software Minitab, devido às suas robustas ferramentas estatísticas e capacidade de lidar com grandes conjuntos de dados.

### VALIDADE E CONFIABILIDADE

Para garantir a validade e confiabilidade dos resultados, os dados foram submetidos a testes preliminares, assegurando que não houvesse multicolinearidade ou autocorrelação que poderiam distorcer os resultados da regressão. A precisão dos modelos foi testada usando métodos de validação cruzada.

### CONSIDERAÇÕES ÉTICAS

Todos os dados coletados foram anonimizados para garantir a privacidade dos estudantes e das instituições envolvidas. O estudo foi conduzido de acordo com as diretrizes éticas para pesquisas envolvendo seres humanos.

### RESULTADOS E DISCUSSÃO - APRESENTAÇÃO DOS RESULTADOS

Foram analisados dados de 688 cursos distribuídos em diversas instituições de ensino. A seguir, apresentamos os principais resultados obtidos:

## Figura 1 | Resultado do Modelo

### Sumário do Modelo

S	R2	R2(aj)	R2(pred)
1,90401	77,01%	76,81%	76,27%

Fonte: Os Autores (2023)

Neste resultado o R2 ajustado de 76,81% significa que aproximadamente 76,81% da variação na variável dependente é explicada pelo nosso modelo de regressão, levando em consideração o número de preditores que temos.

Em termos práticos, isso é um valor relativamente alto, indicando que nosso modelo capta bem a variação dos dados. No entanto, ainda existem cerca de 23,6% da variação que não são explicados por ele, e esses podem ser atribuídos a outras variáveis não incluídas aqui, erro aleatório, entre outros fatores.

**Duração do Curso:** Os cursos avaliados têm durações que variam de 38 a 336 dias. Observou-se variações nas taxas de evasão entre cursos de diferentes durações. Cursos mais curtos tendem a ter taxas de evasão distintas de cursos mais longos.

**Dias da Semana:** Foi analisado o impacto da quantidade de dias de aula por semana na evasão. Cursos que funcionam em diferentes frequências semanais apresentaram taxas de evasão variáveis.

**Custo da Hora-Aula:** O custo da hora-aula mostrou-se uma variável relevante.

Cursos com diferentes faixas de custo hora-aula tiveram taxas de evasão distintas.

**Mês de Início:** A análise procurou entender se o mês de início de um curso influencia as taxas de evasão dos estudantes. No entanto, com um valor-p de 0,610 os resultados indicam que o mês de início não é uma variável estatisticamente significativa no modelo. Isso sugere que a taxa de evasão, dentro deste conjunto de dados e nas condições analisadas, não é fortemente influenciada pelo mês em que o curso começa. Vale ressaltar que esta conclusão não nega a possibilidade de existir um efeito, mas dentro do contexto do modelo e dos dados, não foi possível estabelecer uma relação significativa.

**Período das Aulas:** A análise buscou entender se o período em que as aulas são ministradas (manhã, tarde, noite, integral ou vespertino) influencia as taxas de evasão dos estudantes. Com um valor-p de 0,05, os resultados indicam que o período é uma variável estatisticamente significativa no modelo. Isso sugere que a taxa de evasão pode ser influenciada pelo período do curso. Por exemplo, cursos ministrados à noite podem ter taxas de evasão diferentes daqueles oferecidos durante a manhã

ou tarde. Esta descoberta destaca a importância de considerar o horário das aulas ao analisar e planejar estratégias de retenção de estudantes.

Quantidade de Alunos Matriculados: A quantidade de alunos matriculados em cada sala também mostrou impacto nas taxas de evasão. Salas com diferentes quantidades de alunos apresentaram diferentes taxas de evasão.

## DISCUSSÃO

Os resultados da análise destacam a complexa interação de variáveis que influenciam as taxas de evasão em cursos de ensino técnico e profissionalizante.

**Figura 2 | Análise do Impacto das Variáveis**

### Análise de Variância

Fonte	GL	SQ (Aj.)	QM (Aj.)	Valor F	Valor-P
Regressão	6	8271,3	1378,55	380,26	0,000
Período	1	13,1	13,06	3,60	0,058
Início	1	0,9	0,95	0,26	0,610
Total de Dias	1	37,9	37,88	10,45	0,001
QTD Dias Aula	1	63,1	63,14	17,42	0,000
Nº de Matrículas	1	7192,2	7192,25	1983,93	0,000
Valor Hora	1	14,0	14,03	3,87	0,050
Erro	681	2468,8	3,63		
Falta de ajuste	582	2272,8	3,91	1,97	0,000
Erro puro	99	196,0	1,98	*	†
Total	687	10740,1			

Fonte: Os Autores (2023)

O período das aulas surgiu como uma variável estatisticamente significativa no estudo, com um valor-p de 0,05. Isso sugere que o horário em que as aulas são ministradas tem um impacto palpável sobre a taxa de evasão dos estudantes. Por exemplo, aulas ministradas à noite podem enfrentar desafios diferentes quando comparadas às da manhã ou tarde. Isso pode ser devido a diversos fatores, como compromissos de trabalho dos alunos, disponibilidade de transporte ou mesmo preferências pessoais. Esta descoberta ressalta a importância de considerar o período das aulas ao planejar novos cursos ou ao reavaliar cursos existentes.

Por outro lado, o mês de início não mostrou uma relação estatisticamente significativa com as taxas de evasão. Com um valor-p de 0,602, é possível que o mês em que um curso começa não tenha uma influência forte ou clara sobre a evasão. Isso pode indicar que fatores mais intrínsecos ao curso, como o

conteúdo ou a metodologia de ensino, ou fatores externos, como as condições econômicas, desempenhem um papel mais crucial do que o mês de início do curso.

Outras variáveis, como duração do curso, dias da semana em que o curso é ministrado, custo da hora-aula, e quantidade de alunos matriculados, também mostraram variações nas taxas de evasão. Estas variações enfatizam a necessidade de uma abordagem holística ao considerar a formulação de cursos e estratégias de retenção de estudantes.

Ao comparar os resultados com pesquisas anteriores, fica evidente que a evasão escolar é influenciada por uma combinação de fatores tanto internos quanto externos à instituição. Enquanto alguns estudos anteriores se concentraram em aspectos individuais ou socioeconômicos dos estudantes, este estudo adiciona uma nova dimensão ao debate, realçando a importância de aspectos estruturais dos cursos.

## CONCLUSÃO

A pesquisa em foco conseguiu identificar, com base em uma modelagem matemática robusta, o impacto significativo de várias variáveis na evasão de alunos em cursos de Ensino Técnico e Profissionalizante. Com uma probabilidade de influência de 76,4%, fica evidente que as variáveis em análise têm um papel crucial na determinação do número de alunos que concluirão seus cursos. A equação de regressão resultante:

### Figura 3 | Equação Matemática para Geração da Probabilidade

#### Equação de Regressão

$$\begin{aligned} \text{N}^\circ \text{ de Concluintes} = & -2,221 + 0,1828 \text{ Período} + 0,0116 \text{ Início} + 0,00736 \text{ Total de Dias} \\ & + 0,2690 \text{ QTD Dias Aula} + 0,8612 \text{ N}^\circ \text{ de Matrículas} + 0,0986 \text{ Valor Hora} \end{aligned}$$

Fonte: Os Autores (2023)

Esta fórmula, proporciona às instituições educacionais uma ferramenta valiosa para prever, com um grau considerável de precisão, o número de concluintes com base nas variáveis definidas. Além da aplicação prática da fórmula proposta, o estudo reforça a importância de considerar fatores como a duração do curso, o período das aulas, o mês de início, entre outros, ao formular e ofertar cursos. Ao integrar esses insights em seus processos de planejamento e execução, as instituições podem adotar estratégias mais eficazes para maximizar a retenção de estudantes, reduzindo assim as taxas de evasão.

Em última análise, o sucesso desta pesquisa reafirma a necessidade e o valor da aplicação de métodos matemáticos e estatísticos na área educacional. Com abordagens como a presente, é possível não só entender os desafios da educação, mas também prever e, mais importante, moldar o futuro da aprendizagem.

## REFERÊNCIAS

- ANZOLIN, R.M., KRELING, W.L. **Análise das causas de evasão escolar nos cursos de aprendizagem industrial de uma unidade de educação profissional do SENAI/SC no ano de 2012**. Revista e-TECH: Tecnologias para Competitividade Industrial-ISSN-1983-1838, 73-90, 2013
- COLPANI, R. **Mineração de Dados Educacionais: um estudo da evasão no ensino médio com base nos indicadores do Censo Escolar**. Informática na educação: teoria & prática, v. 21, n. 3, 2018.
- HOYOS OSORIO, J.K., DAZA SANTACOLOMA, G. **Modelo preditivo para identificar estudantes universitários com alto risco de evasão**. Revista electrónica de investigación educativa, v. 25, 2023.
- LINO, E.R.O. **A problemática da evasão escolar: uma revisão bibliográfica integrativa**. 2020.
- MORAIS, F.L., MELO, A., MOUTINHO, M. FAGUNDES, R. **Modelos de regressão aplicados na previsão da evasão escolar do ensino básico: uma revisão sistemática da literatura**. Anais do XXXII Simpósio Brasileiro de Informática na Educação, p. 168-178, 2021.
- OLIVEIRA, I., ANDRADE, F., MEDEIROS, F., CASIMIRO, C., & FARIAS, F. **Construção de um data warehouse para uma análise multidimensional da evasão escolar: estudo de caso da Plataforma Nilo Peçanha**. Revista Principia, 2022.
- OLIVEIRA, P.L., CARMO, N.C. **A temática evasão escolar no contexto do PROEJA: uma revisão integrativa**. Revista Ponto de Vista, v. 10, n. 1, p. 01-21, 2021.
- PIMENTA, C.D., SILVA, M.B., COSTA, A.F.B., SALOMON, V.A.P. **Otimização e escolha de modelos probabilísticos no processo de tratamento térmico em arames de aço temperados e revenidos**. Revista Eletrônica Produção & Engenharia, v. 8, n. 1, p. 640-652, 2016.
- PRIM, A.L., FÁVERO, J.D. **Motivos da evasão escolar nos cursos de ensino superior de uma faculdade na cidade de Blumenau**. Revista E-Tech: Tecnologias para Competitividade Industrial-ISSN-1983-1838, p. 53-72, 2013.
- QUEIROGA, E.M., LOPES, J.L.B., ARAÚJO, R.M., CECHINEL, C. **Modelo de predição da evasão de estudantes em cursos técnicos a distância a partir da contagem de interações**. Revista Thema, v. 15, n. 2, p. 425-438, 2018.
- SADUYAMA, A., REZENDE, K., SANTOS, M., SADUYAMA, G. **Evasão escolar no ensino superior: um estudo de revisão sistemática**. Revista de Psicologia, Educação e Cultura, v. 24, n. 3, p. 92-103, 2020.
- SILVA, I.J.A., MIRANDA, G.J., LEAL, E.A., PEREIRA, J.M. **Estratégias das coordenações dos cursos de ciências contábeis para combater a evasão**. Revista Universo Contábil, v. 14, n. 2, p. 61-81, 2019.
- SILVA, P.T.F., SAMPAIO, L.M.B. **Políticas de permanência estudantil na educação superior: reflexões de uma revisão da literatura para o contexto brasileiro**. Revista de Administração Pública, 56, 603-631, 2022



**UNITAU**  
Universidade de Taubaté