





## Análise do desempenho acadêmico entre alunos cotistas e não cotistas do IFMG - campus Congonhas e Formiga

### *Analysis of academic performance between quota and non-quota students at IFMG - campus Congonhas and Formiga*

 **Leonardo de Souza Cimino**  
Mestre em Ciência da Computação  
Instituto Federal de Minas Gerais - IFMG.  
Congonhas, Minas Gerais – Brasil.  
[leonardo.cimino@ifmg.edu.br](mailto:leonardo.cimino@ifmg.edu.br)

 **Everthon Valadão dos Santos**  
Mestre em Ciência da Computação  
Instituto Federal de Minas Gerais - IFMG.  
Formiga, Minas Gerais – Brasil.  
[everthon.valadao@ifmg.edu.br](mailto:everthon.valadao@ifmg.edu.br)

 **Lilian Amaral de Carvalho**  
Doutora em Química  
Instituto Federal de Minas Gerais - IFMG.  
Arcos, Minas Gerais – Brasil.  
[lilian.carvalho@ifmg.edu.br](mailto:lilian.carvalho@ifmg.edu.br)

**Resumo:** A implantação das políticas de cotas nas instituições federais de ensino tornou o ambiente acadêmico mais heterogêneo em relação à situação socioeconômica dos alunos. Diante desse cenário tão diverso, o objetivo deste estudo é analisar se o desempenho acadêmico dos alunos mais vulneráveis econômica e socialmente é inferior aos demais estudantes. Para isso, foi analisado o coeficiente de rendimento de alunos de cursos integrados utilizando estatística descritiva e aprendizado de máquina. Na análise descritiva, os alunos foram separados em cotistas e não cotistas e os resultados demonstraram que o desempenho dessas duas classes é similar. Já a análise por aprendizado de máquina encontrou situações socioeconômicas nas quais os alunos tendem a obter um desempenho inferior aos seus pares. As regras obtidas nessa análise podem justificar um acompanhamento pedagógico mais próximo para alunos que se encaixam nelas, evitando a evasão e melhorando o desempenho desses alunos.

**Palavras chave:** cotas; desempenho acadêmico; instituições federais de ensino.

**Abstract:** The implementation of quota policies in federal educational institutions increased the heterogeneity of socioeconomic status of students. The objective of this study is to analyze whether the academic performance of the most economically and socially vulnerable students is inferior to the others. The academic performance of students of integrated courses was analyzed using descriptive statistics and machine learning techniques. In the first one, the students were classified as quota and non-quota students and the results showed that the performance of both is similar. The machine learning analysis found socioeconomic situations where students tend to have a lower performance compared to other students. The obtained rules in this analysis can justify closer pedagogical monitoring for students who fit into them, preventing evasion and improving the performance of these students.

**Keywords:** quota; academic performance; federal educational institutions.

Cite como

(ABNT NBR 6023:2018)

CIMINO, Leonardo de Souza; SANTOS, Everthon Valadão dos; CARVALHO, Lilian Amaral de. Análise do desempenho acadêmico entre alunos cotistas e não cotistas do IFMG - campus Congonhas e Formiga. *Dialogia*, São Paulo, n. 41, p. 1-18, e20206, maio/ago. 2022. Disponível em: <https://doi.org/10.5585/414.2022.20206>.

*American Psychological Association (APA)*

Cimino, L. de S., Santos, E. V. dos., & Carvalho, L. A. de. (2022, jan./abr.) Análise do desempenho acadêmico entre alunos cotistas e não cotistas do IFMG - campus Congonhas e Formiga *Dialogia*, São Paulo, 41, p. 1-18, e20206. <https://doi.org/10.5585/41.2022.20206>.



## 1 Introdução

Houve uma grande expansão da rede pública federal de educação nos últimos anos no Brasil. Grande parte dessa ampliação se deve à criação dos Institutos Federais de Educação, Ciência e Tecnologia (IFs) (CAMPOS, 2018). Além dos IFs, essa rede ainda conta com outras instituições como as Universidades Federais (UFs) e os Centros Federais de Educação Tecnológica (CEFETs). Para democratizar o acesso às instituições federais de ensino, foram promulgados a Lei nº 12.711/2012, o Decreto nº 7.284/2012 e a Portaria Normativa nº 18/2012 do Ministério da Educação implementando um sistema de cotas que reserva 50% das vagas para alunos oriundos de escolas públicas.

Em 2010 já existia uma política de ações afirmativas em quase 70% das universidades federais (TRAGTENBERG, 2010), porém cada instituição definia seus próprios critérios. Somente com a Lei nº 12.711/2012 foi estabelecida uma regulamentação e padronização do sistema de cotas nas instituições federais de ensino do país. A partir da publicação desses instrumentos, as instituições puderam implementar o sistema de forma progressiva até meados de 2016, o prazo limite para o cumprimento integral da referida legislação.

A implementação de um sistema de cotas é um tema que levanta muitas discussões. Há autores a favor e contra esse tipo de política. Souza e Brandalise (2017) afirmam que as cotas são um mecanismo para democratizar o acesso ao ensino superior para indivíduos que foram historicamente excluídos. Zuin e Bastos (2019) afirmam que a Lei de Cotas tem alcançado sua função social e que tem sido um instrumento de inclusão e diversidade cultural. Iensue *et al.* (2009) demonstram que os negros ocupam posições menos favorecidas na sociedade e que isso justifica a implementação de políticas de cotas para garantir que a população negra carente consiga acesso à educação superior.

Em contrapartida, Fry e Maggie (2004) e Azevedo (2004) afirmam que cotas raciais são uma maneira de consolidar o racismo, e não de erradicá-lo, uma vez que o racismo só existe devido à crença em raças. Outro ponto questionado é que a qualidade do ensino superior pode ser comprometida nas universidades devido a uma formação anterior insuficiente de alunos oriundos de escolas públicas (VELLOSO, 2009). Nascimento (2017) afirma que estamos longe de alcançar uma democracia racial e que, diante disso, as cotas exclusivamente sociais são a melhor alternativa, visto que elas cumprem melhor seu objetivo de que pessoas em camadas sociais mais baixas consigam ser incluídas em camadas mais elevadas.

Independentemente dos discursos favoráveis ou contrários à Lei de Cotas, sua implantação fez com que seja encontrado um contraste de estudantes com situações socioeconômicas distintas

dentro das instituições federais de ensino. Em uma mesma sala existem perfis de alunos completamente diferentes, com condições financeiras completamente distintas, diferentes níveis no grau de escolaridade dos pais e outras situações sociais que impactam diretamente a vida do aluno.

Afinal, o aproveitamento de alunos cotistas é inferior ao de alunos não cotistas? Para responder essa pergunta, este artigo apresenta uma análise comparativa do rendimento acadêmico de alunos cotistas e não cotistas dos cursos técnicos integrados no âmbito do Instituto Federal de Minas Gerais (IFMG), nos *campi* Congonhas e Formiga. Também busca-se explorar via técnicas de aprendizado de máquina se há padrões sociais que poderiam interferir no rendimento escolar do aluno.

As seções que compõem esse artigo são: a seção 2 apresenta um levantamento de trabalhos que analisaram o desempenho discente e questões socioeconômicas. O método utilizado na pesquisa é documentado na seção 3 e a seção 4 apresenta os seus resultados. Por fim, as conclusões e sugestões de trabalhos futuros são apresentadas na seção 5.

## 2 Trabalhos relacionados

Ao longo da última década, diversos trabalhos desenvolveram uma análise comparativa entre o desempenho de alunos cotistas e não cotistas. Entretanto, a maioria destes trabalhos, tais como Velloso (2009), Dallabona e Filho (2011), Santos (2012), Maximo, Gandolfi e Lopes (2020), Peixoto *et al.* (2016), Cavalcanti *et al.* (2019), faz uma análise focada em alunos de cursos de graduação. Não foram encontrados muitos trabalhos na literatura analisando e/ou comparando o desempenho de alunos cotistas e não cotistas no Ensino Médio.

Entre as pesquisas realizadas no Ensino Médio, primeiramente destaca-se o trabalho de Lima, Oliveira e Cruz (2020), no qual foi realizado um estudo comparando o coeficiente de rendimento (CR) dos alunos que ingressaram em 2013 e 2014 nos cursos técnicos integrados do Instituto Federal de Minas Gerais (IFMG) - *campus* Ouro Branco e que concluíram os respectivos cursos. Em sua primeira análise, os alunos foram classificados como cotistas e não cotistas, utilizando estatística descritiva e testes de significância para analisar os resultados. A análise descritiva demonstrou que a diferença de desempenho foi inferior a 2% tanto em 2013 quanto em 2014. O teste de significância apontou que as notas não possuem diferença significativa e que as notas dos dois grupos podem ser consideradas equivalentes. Posteriormente, os alunos cotistas e não cotistas foram separados por curso. Os resultados demonstraram que as notas dos dois grupos também foram equivalentes nos três cursos analisados.

Chaves *et al.* (2020) analisaram o desempenho acadêmico, frequência e taxas de permanência de alunos dos cursos técnicos integrados que ingressaram em 2014 no Instituto Federal do Rio Grande do Norte (IFRN) - *campus* Caicó. No referido trabalho, foi observado que os alunos cotistas possuem uma taxa maior de cancelamento de matrículas. Por outro lado, alunos não cotistas possuem uma taxa de conclusão inferior e, portanto, uma taxa de evasão superior. Não foram constatadas diferenças na frequência dos cotistas e não cotistas. O Índice de Rendimento Acadêmico (IRA) dos alunos foi utilizado como parâmetro de comparação. Foram consideradas apenas as disciplinas propedêuticas, as quais são comuns entre todos os quatro cursos integrados no *campus* Caicó do IFRN. A análise separada por curso mostrou que houve algumas disparidades de até 10 pontos de diferença em favor dos não cotistas em determinadas disciplinas, como, por exemplo, “Matemática”. Os resultados globais demonstraram uma discrepância em disciplinas das áreas de Ciências da Natureza e Linguagens em favor dos não cotistas. Considerou-se que os resultados dos alunos cotistas são satisfatórios, apesar de algumas discrepâncias como as apresentadas e, principalmente, nas séries finais.

Pieron (2016) realizou uma análise comparando os índices de evasão, conclusão, reprovação e o rendimento acadêmico de 424 alunos de dez turmas dos cursos técnicos (integrados ao Ensino Médio, concomitante, subsequente) e cursos superiores de tecnologia do Instituto Federal de São Paulo (IFSP) - *campus* Salto. Segundo a referida autora, apesar das diferenças encontradas entre os alunos ingressantes pela reserva de vagas e os de ampla concorrência, em todas as comparações há turmas com melhores desempenhos em ambos os grupos.

Pinto (2015) identificou e analisou a relação entre as características individuais, perfil socioeconômico e o desempenho dos alunos dos cursos de Ensino Médio integrado no Instituto Federal da Bahia (IFBA) - *campus* Barreiras. Segundo essa autora, foi utilizada uma abordagem quantitativa de natureza descritiva e os dados foram analisados através de técnicas estatísticas. Analisando os resultados, Pinto (2015) observou que houve diferença estatisticamente significativa em todos os cursos entre alunos cotistas e não cotistas, porém elas aparecem apenas na primeira série. O nivelamento nas séries seguintes pode indicar um bom acolhimento dos alunos por parte da instituição, corrigindo uma provável formação anterior insuficiente.

### 3 Materiais e método

O objetivo deste trabalho é apresentar uma comparação entre o desempenho acadêmico de alunos cotistas e não cotistas, com o intuito de identificar se os cotistas apresentam desempenho diferente em relação aos não cotistas. Para atingir esse objetivo, foi utilizada uma abordagem

quantitativa usando os dados dos alunos do Instituto Federal de Minas Gerais (IFMG), dos *campi* Congonhas e Formiga. A análise abrangeu o desempenho de 902 alunos dos cursos técnicos integrados, entre os anos de 2017 e 2019, em que 475 (52,6%) ingressaram através de cotas e 427 (47,4%) não são cotistas. Os dados foram extraídos do sistema ERP acadêmico do IFMG, após autorização formal da Diretoria de Planejamento Educacional.

### 3.1 Análise via estatística descritiva

O parâmetro utilizado para medir o desempenho dos alunos foi o Coeficiente de Rendimento (CR). O CR não necessariamente é uma métrica que consegue reproduzir, efetivamente, o conhecimento que o aluno adquiriu. Entretanto, ele ainda é amplamente utilizado como uma métrica objetiva para avaliar o desempenho discente (SILVA; XAVIER; COSTA, 2020). Por isso, diversos trabalhos na literatura utilizam o CR como métrica para avaliar o desempenho acadêmico, dentre os quais destacam-se Silva, Xavier e Costa (2020), Peixoto *et al.* (2016), Maximo, Gandolfi e Lopes (2020), Dallabona e Filho (2011), Chaves *et al.* (2020).

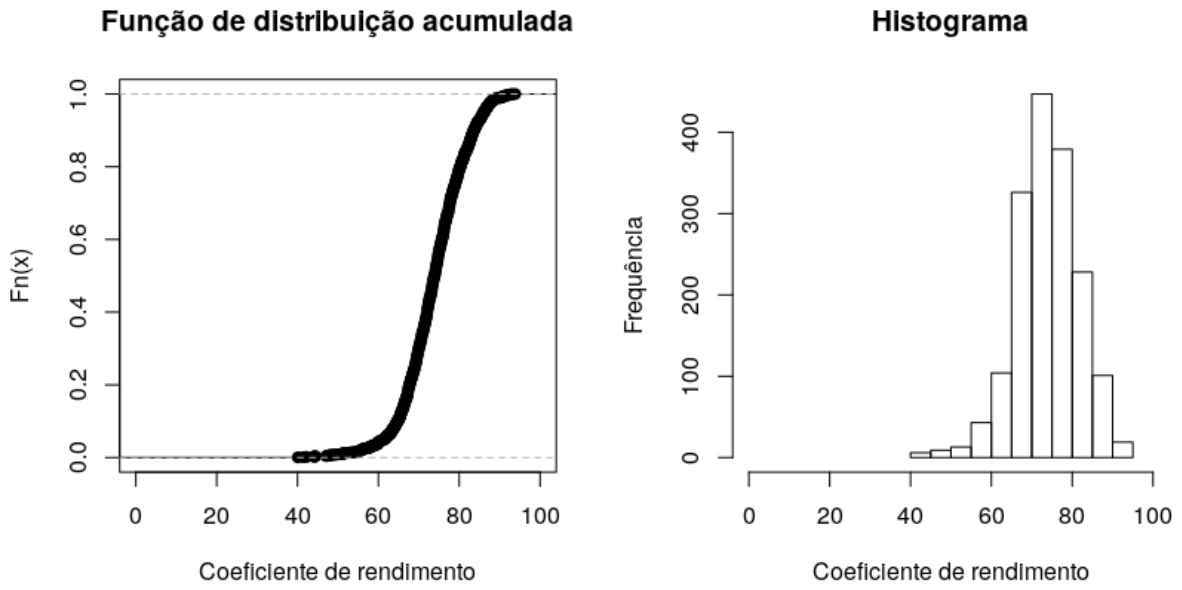
Após a obtenção dos dados, foi feita uma análise exploratória, visando identificar e remover registros com informações incompletas ou incorretas (ex.: campo não preenchido ou campo com valor nulo - *NULL*). A base de dados utilizada possui ao todo 1675 registros, sendo que cada registro indica o CR de um determinado aluno em um determinado período letivo. A Tabela 1 apresenta uma visão geral da distribuição desses registros por curso e *campus*. A Figura 1 apresenta a função de distribuição acumulada e o histograma do coeficiente de rendimento. Já a Figura 2 apresenta o gráfico de caixa (*boxplot*) do CR dos alunos.

**Tabela 1** – Distribuição dos dados

<i>Campus</i>	<i>Curso</i>	<i>Registros</i>
Congonhas	Edificações	307
Congonhas	Mecânica	301
Congonhas	Mineração	312
Formiga	Administração	278
Formiga	Eletrotécnica	211
Formiga	Informática	266

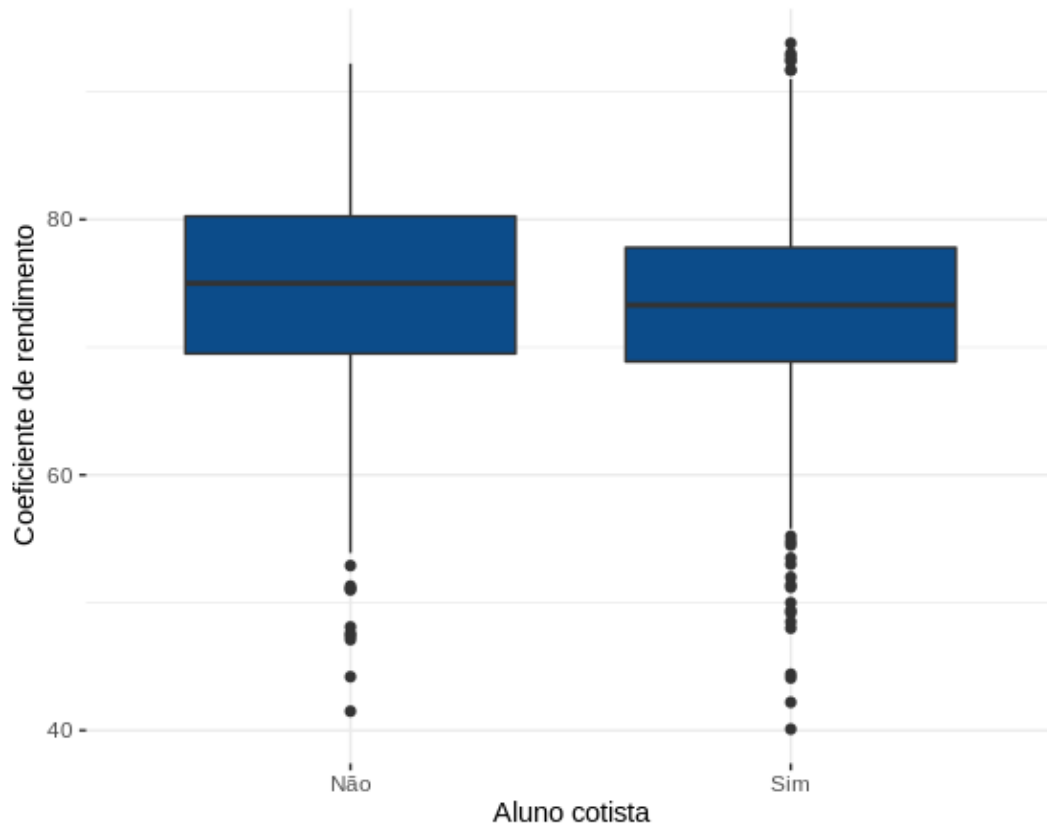
Fonte: Elaborada pelos autores.

**Figura 1** - Função de distribuição acumulada e histograma do CR



Fonte: Elaborada pelos autores.

**Figura 2** – Gráfico de caixa do CR dos alunos



Fonte: Elaborada pelos autores.

### 3.2 Análise via aprendizado de máquina

Adicionalmente à análise estatística, para viabilizar a análise e correlação dos dados de desempenho acadêmico e aspectos socioeconômicos, foi utilizado um ambiente de aprendizado de máquina de código aberto chamado Weka<sup>1</sup>.

Para viabilizar uma análise contextualizada para analisar possíveis diferenças regionais, os registros acadêmicos foram segmentados conforme o *campus* onde o discente estuda, a saber: *campus* Congonhas ou *campus* Formiga.

Vale destacar que para viabilizar a identificação de regras de associação por algoritmos de aprendizado de máquina, no lugar dos valores numéricos (ex.: coeficiente de rendimento) é necessário definir algumas classes de valores. Inicialmente foi utilizada a seguinte classe de valores para o coeficiente de rendimento (doravante, CR): A, B, C, D, E e F, com diferença fixa de 10 pontos entre classes, i.e., A ( $CR \geq 90$ ), B ( $90 < CR \leq 80$ ), C ( $80 < CR \leq 70$ ), D ( $70 < CR \leq 60$ ), E ( $60 < CR \leq 50$ ) e F ( $CR < 50$ ). Porém, com a utilização dessas seis classes, não foram encontradas regras de associação significativas pelos algoritmos de aprendizado de máquina. Então, foi analisado o intervalo de valores do CR dos 902 alunos conforme sua distribuição estatística (vide histograma na Figura 1 e *boxplot* na Figura 2) de CR dos discentes em ambos os *campi*. Dessa observação, definiu-se para o CR as classes de valores, a saber: “ótimo” ( $CR \geq 80$ ), “bom” ( $80 < CR \leq 65$ ), “regular” ( $65 < CR \leq 60$ ) e “insuficiente” ( $CR < 60$ ).

Com os valores numéricos devidamente segmentados, os registros acadêmicos foram “minerados” com o software Weka em busca de regras de associação utilizando o algoritmo Apriori, no qual o suporte mínimo é reduzido iterativamente até encontrar a quantidade de regras necessárias com o nível mínimo de confiança desejado (AGRAWAL; SRIKANT, 1994). Os registros acadêmicos também foram “minerados” com o classificador JRip, um algoritmo de aprendizado de regras proposicionais proposto por Cohen (1995). Adicionalmente, foi utilizado o classificador Logit Boost, que realiza a classificação usando um esquema de regressão como aprendizado base, por meio de regressão logística aditiva (FRIEDMAN; HASTIE; TIBSHIRANI, 1998). Por fim, os dados foram analisados com um classificador Naive Bayes, que utiliza classes de estimadores, cujos valores de precisão são escolhidos com base na análise dos dados de treinamento (JOHN; LANGLEY, 1995).

## 4 Resultados e análise

A Tabela 2 apresenta uma comparação da média do CR por curso. A mediana e o coeficiente de variação (CV) também foram calculados para auxiliar na interpretação dos dados.



Nessa primeira análise, os alunos foram classificados entre cotistas e não cotistas, independentemente do tipo da cota de ingresso (vide Quadro 1).

**Tabela 2** – Análise do CR por curso

<i>Campus</i>	Curso	Cotistas			Não cotistas		
		Média	Mediana	CV	Média	Mediana	CV
Congonhas	Edificações	76.6	75.8	7,7%	77.8	78.2	8,3%
Congonhas	Mecânica	70.3	71.2	10,2%	70.0	70.0	11,1%
Congonhas	Mineração	72.3	72.2	8,3%	71.5	71.0	9,6%
Formiga	Administração	76.7	76.8	9,3%	79.0	79.4	7,9%
Formiga	Eletrotécnica	71.3	71.0	12,1%	72.8	72.9	11,0%
Formiga	Informática	71.5	72.1	11,9%	74.7	75.0	10,5%

Fonte: Elaborada pelos autores.

Inicialmente foi observado que, em todos os cenários, a mediana está muito próxima à média. Isso demonstra que não existem dados discrepantes os quais poderiam interferir na média geral para cima ou para baixo. Tal informação é expressa pelo coeficiente de variação (CV), que se manteve em índices baixos e, portanto, ratificando que existe uma certa homogeneidade nos dados, sem variações “bruscas” na distribuição dos dados.

Ao observar a média do CR de alunos cotistas e não cotistas, percebeu-se que, em geral, a média do CR desses dois grupos é muito próxima. Nos cursos técnicos integrados, os alunos cotistas apresentaram um CR médio superior em dois dos seis cursos avaliados. Entretanto, na maioria dos casos, a diferença das médias entre os dois grupos é inferior a 2,5%, seja em favor dos cotistas ou não cotistas. A maior diferença registrada em tais cursos foi no curso Técnico em Informática, mas com uma diferença de apenas 4,4% no CR médio de alunos cotistas e não cotistas, sendo que a maior média coube a não cotistas. Portanto, ao analisar tais dados, não se pode afirmar que o desempenho de alunos cotistas difere superior ou inferiormente àquele de estudantes não cotistas. Conforme os dados foram explorados, percebeu-se que a diferença do CR médio entre os alunos é muito pequena na maioria dos casos. Assim, como o desempenho acadêmico desses dois grupos é muito similar, não é possível afirmar que um grupo possui desempenho superior ou inferior ao outro.

Para se ter uma visão mais geral do desempenho discente, foi feita uma análise agrupando os dados por *campus*, com resultados apresentados na Tabela 3. De forma semelhante à análise



anterior, a mediana se manteve próxima à média e o coeficiente de variação (CV) foi baixo, indicando maior homogeneidade nos dados, ou seja, não existe grande quantidade de outliers interferindo na média em relação à mediana. Ao agrupar os dados, percebeu-se que o desempenho (CR médio) dos alunos não cotistas foi ligeiramente superior ao dos não cotistas: no *campus* Congonhas, a diferença foi de 0,5%; no campus Formiga, a diferença foi de 3,1%. Entretanto, reforça-se que, como os valores são muito próximos, não foi possível identificar qualquer indício com significância estatística de que o desempenho de um grupo é superior ao outro.

**Tabela 3** – Análise do CR por modalidade e campus

<i>Campus</i>	Cotistas			Não cotistas		
	Média	Mediana	CV	Média	Mediana	CV
Congonhas	73.0	73.0	9,4%	73.3	73.4	10,7%
Formiga	73.4	73.8	11,5%	75.6	75.7	10,3%

Fonte: Elaborada pelos autores.

Nos cenários explorados anteriormente, todos os alunos cotistas foram unificados em um único grupo. Porém, existem diversos tipos de cotas pelas quais um aluno pode ingressar, conforme detalhado no Quadro 1.

**Quadro 1** – As siglas das cotas e seu significado

Sigla	Significado
AC	Ampla concorrência (não cotista)
L1	Renda familiar bruta <i>per capita</i> igual ou inferior a 1,5 (um e meio) salário mínimo, estudantes que tenham cursado integralmente o Ensino Fundamental ou Médio, conforme o caso, em escolas públicas.
L2	Alunos autodeclarados negros (pretos ou pardos) ou indígenas, com renda familiar bruta <i>per capita</i> inferior a 1,5 (um e meio) salário mínimo e que tenham cursado integralmente o Ensino Fundamental ou Médio, conforme o caso, em escolas públicas.
L5	Alunos com renda familiar bruta <i>per capita</i> igual ou superior a 1,5 (um e meio) salário mínimo que cursaram integralmente o Ensino Fundamental ou Médio, conforme o caso, em escolas públicas.
L6	Alunos autodeclarados pretos, pardos ou indígenas com renda familiar bruta <i>per capita</i> superior a 1,5 (um e meio) salário mínimo que, independentemente da renda, cursaram integralmente o Ensino Fundamental ou Médio, conforme o caso, em escolas públicas.

Fonte: Elaborado pelos autores.

Cada cota agrupa um perfil de alunos diferente e, por isso, classificá-los todos em um único grupo, como foi feito, poderia ocultar informações relevantes. Portanto, foram feitas análises

complementares para avaliar o desempenho discente por tipo de cota de ingresso. A Tabela 4 apresenta o CR médio por tipo de cota de ingresso.

**Tabela 4** – Análise do CR por tipo de cota e campus

<i>Campus</i>	AC	L1	L2	L5	L6
Congonhas	73.3	74.3	71.7	73.7	72.9
Formiga	75.6	73.6	73.4	73.8	72.8

Fonte: Elaborada pelos autores.

Em ambos os *campi*, o CR médio de todos os tipos de cota se mantiveram próximos. No *campus* Congonhas, o grupo de alunos com melhor desempenho foi o de cotistas L1. Já no *campus* Formiga, os alunos não cotistas tiveram desempenho médio superior aos demais. Contudo, ressalta-se que a diferença percentual foi bem pequena.

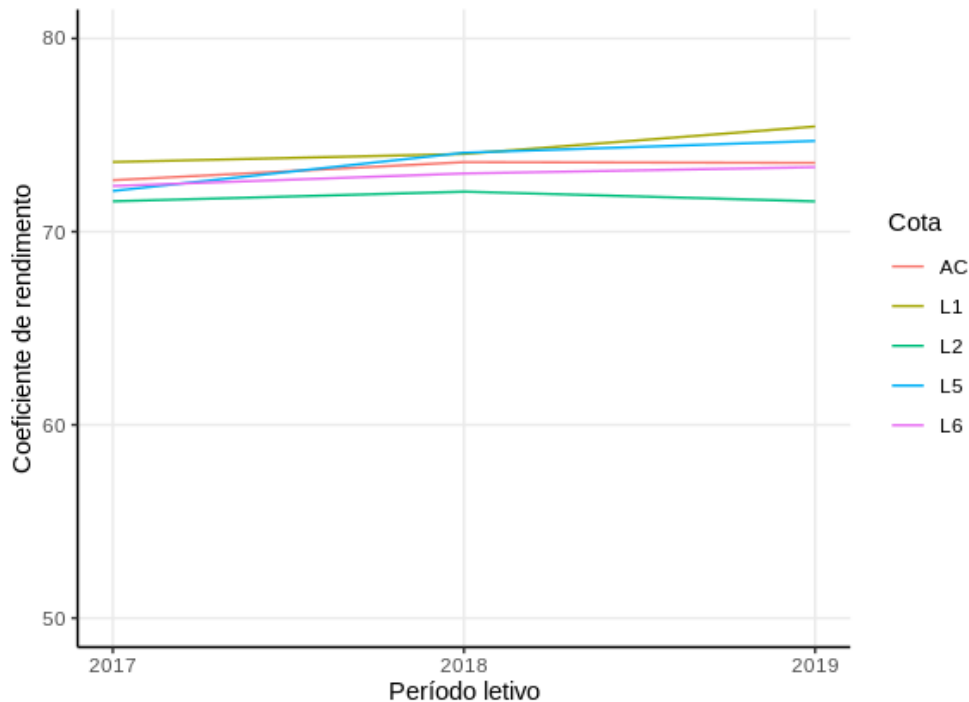
Para se obter uma visão cronológica da evolução de desempenho dos discentes, agruparam-se os dados por período letivo. As Figuras 3 e 4 apresentam os resultados dos cursos técnicos integrados ao Ensino Médio nos *campi* Congonhas e Formiga, respectivamente.

A Figura 4 demonstra que, entre os alunos dos cursos integrados do *campus* Formiga, os alunos não cotistas obtiveram um CR médio ligeiramente superior ao longo dos anos de 2017, 2018 e 2019. Entre os cotistas, o CR médio se manteve próximo, sendo que, nos períodos letivos de 2019, houve uma queda mais significativa entre os alunos de cota L6, fenômeno também observado no CR médio dos alunos não cotistas (AC).

#### 4.1 Resultados obtidos via aprendizagem de máquina

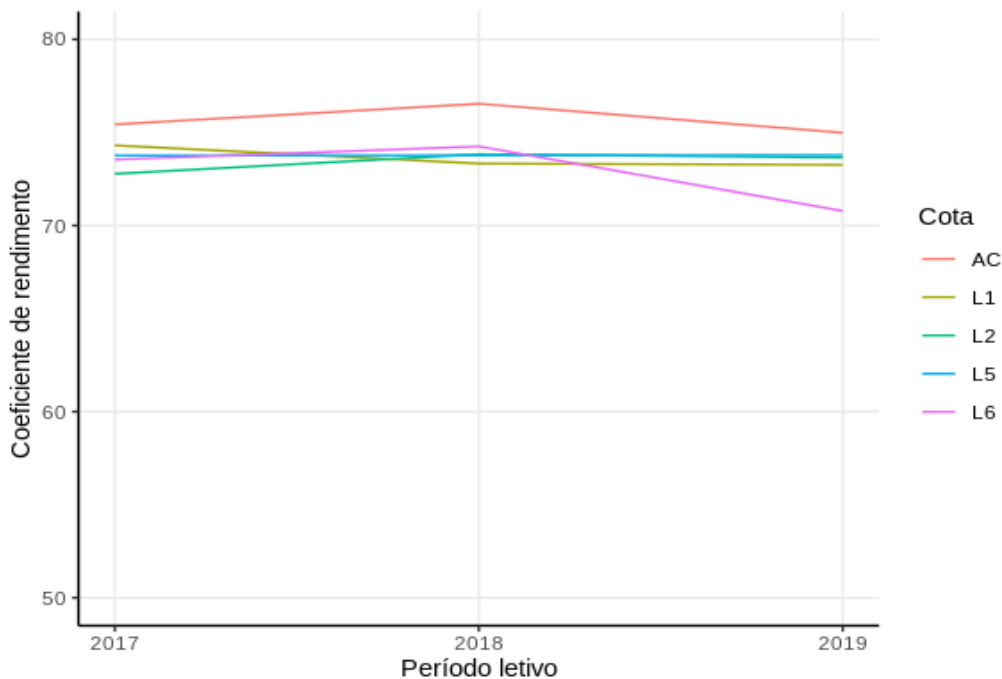
A seguir, são apresentados os resultados apontados pelos algoritmos de aprendizagem de máquina, quando se analisaram os dados de registro acadêmico de alunos dos cursos técnicos integrados ao Ensino Médio no *campus* Congonhas e no *campus* Formiga. Vale ressaltar que adiante são apresentadas as regras de associação encontradas de forma automatizada por algoritmos de aprendizado de máquina e, portanto, desprovidos de qualquer juízo de valor por parte dos autores.

**Figura 3 – CR por cota de ingresso dos cursos integrados - *campus* Congonhas**  
**CR médio dos cursos integrados - campus Congonhas**



Fonte: Elaborada pelos autores.

**Figura 4 – CR por cota de ingresso dos cursos integrados - *campus* Formiga**  
**CR médio dos cursos integrados - campus Formiga**



Fonte: Elaborada pelos autores.

#### 4.1.1 Regras de inferência na lógica proposicional

Antes de apresentar os resultados, é importante ressaltar a necessidade de atenção ao interpretar regras de associação, de maneira que se devem observar as regras de inferência válidas na lógica proposicional. Por exemplo, a afirmação “se P então Q” significa que “se P é verdadeiro, então Q será verdadeiro também”. Porém, deve-se atentar que o simples fato de Q ser verdadeiro não implica que P seria necessariamente verdadeiro também. Ou seja, a regra de associação deve ser lida e interpretada na ordem apresentada ( $P \rightarrow Q$ ), e não ao contrário ( $Q \rightarrow P$ ).

Entretanto, destaca-se que, na lógica proposicional, pode-se utilizar o argumento dedutivo *modus tollens* como regra de inferência válida. O *modus tollens* consiste no nome formal para a prova indireta via negação do conseqüente (ROSEN, 2009). Nele, há uma negação do conseqüente na segunda premissa, ou seja, afirma-se que o conseqüente é falso. Disso, conclui-se que o antecedente também é falso. Portanto “se P então Q” e “Q é falso”, logo, “P é falso” consiste em uma interpretação logicamente válida e correta.

#### 4.1.2 Cursos técnicos integrados no campus Congonhas

As informações de desempenho acadêmico e dados socioeconômicos de 920 registros do curso técnico integrado do *campus* Congonhas foram computacionalmente analisados, utilizando duas abordagens de aprendizado de máquina. Através do classificador JRip, foi realizado o aprendizado de regras proposicionais, das quais realçam-se as seguintes:

- “Discentes do sexo ‘masculino’ E com situação da mãe ‘ausente’ E situação do pai ‘ausente’ APRESENTARAM desempenho ‘insuficiente’ ( $CR < 60$ )”.

É importante observar que tal regra proposicional foi apenas observada em três instâncias e sem nenhum contraexemplo. Entretanto, aplicando-se a regra de inferência *modus tollens* pode-se igualmente afirmar que, nos cursos técnicos integrados do *campus* Congonhas, “nenhum discente do sexo ‘masculino’ que tenha apresentado desempenho acima de ‘insuficiente’ ( $CR \geq 60$ ) possui ambos os pais ‘ausentes’”. Portanto, a partir de tal constatação é recomendável que o setor pedagógico e diretoria de ensino efetuem um acompanhamento maior dos discentes que se encontrem em tal situação parental (estudantes do sexo masculino com ambos os pais ausentes).

- “Discentes do curso ‘Téc.Int.Mineração’ E do sexo ‘masculino’ E mora com ‘pai’ APRESENTARAM desempenho ‘regular’ ( $65 < CR \leq 60$ )”.

Enfatiza-se que tal regra proposicional foi apenas observada em cinco instâncias e sem nenhum contraexemplo. Mesmo assim, apesar de essa regra proposicional encontrada pelo JRip não ser uma regra de ampla aplicação, tal constatação teria utilidade, considerando-se que o setor pedagógico e a coordenação do curso Técnico Integrado em Mineração poderiam direcionar um acompanhamento personalizado aos discentes do sexo ‘masculino’ e que morem com o ‘pai’.

Como contraponto, considere a regra proposicional:

- “Discentes do sexo ‘feminino’ E grau de instrução da mãe é ‘ensino superior completo’ APRESENTARAM desempenho ‘ótimo’ (CR  $\geq$  80).

Salienta-se que tal regra proposicional foi observada em dez instâncias e teve um contraexemplo. Ainda assim, poderia indicar uma tendência de que discentes do sexo ‘masculino’, cujas mães tenham grau de instrução diferente de ‘ensino superior completo’, estejam apresentando desempenho acadêmico inferior a ‘ótimo’ (CR  $<$  80).

#### 4.1.3 Cursos técnicos integrados no campus Formiga

As informações de desempenho acadêmico e dados socioeconômicos de 755 registros do curso técnico integrado do *campus* Formiga foram computacionalmente analisados através do classificador JRip. Dentre as proposições encontradas, ressaltam-se as seguintes regras:

- “Discentes do curso ‘Téc.Int.Informática’ E reside com ‘outros’ E cotista ‘L1’ APRESENTARAM desempenho ‘insuficiente’ (CR  $\leq$  60)”.

Vale ressaltar que tal regra proposicional foi apenas observada em três instâncias e sem nenhum contraexemplo. Tal constatação poderia ser levada em conta pelo setor pedagógico e a coordenação do curso Técnico Integrado em Informática, visando proporcionar um acompanhamento cuidadoso aos discentes que não residam com os pais e que sejam cotistas ‘L1’.

- “Discentes do curso ‘Téc.Int.Administração’ E área de procedência ‘rural’ E não cotista (‘AC’) APRESENTARAM desempenho ‘ótimo’ (CR  $\geq$  80)”.

Destaca-se que tal regra proposicional foi observada apenas em cinco instâncias e sem nenhum contraexemplo. Entretanto, aplicando-se a regra de inferência modus tollens, pode-se

igualmente afirmar que, nos cursos técnicos integrados do campus Formiga, SE o discente apresentou desempenho inferior a 'ótimo' ( $CR < 80$ ) ENTÃO o discente atende ao menos uma das seguintes condições: pertence a um curso diferente de 'Administração' ('Eletrotécnica' ou 'Informática') OU sua área de procedência é diferente de 'rural' ('urbana') OU é um aluno cotista. Conforme as regras da lógica proposicional, vale destacar que não é válido inferir a partir disso que um aluno cotista necessariamente apresente desempenho inferior a 'ótimo' ( $CR \leq 80$ ), conforme devidamente explicado na Seção 4.1.1.

#### 4.1.4 Cursos Técnicos Integrados em ambos os campi (Congonhas e Formiga)

As informações de desempenho acadêmico e dados socioeconômicos de 1675 registros dos cursos técnicos integrados de ambos os campi Congonhas e Formiga foram computacionalmente analisados, utilizando-se de duas abordagens de aprendizado de máquina. Inicialmente, foi utilizado o algoritmo Apriori (AGRAWAL; SRIKANT, 1994) para "minerar" por regras de associação, tendo como métrica o nível de confiança mínimo de 90%. Das regras de associação por ele identificadas, destaca-se a seguinte:

- É possível afirmar com 97% de confiança que "SE o discente apresenta desempenho 'bom' ( $80 < CR \leq 65$ ) ENTÃO a situação da mãe é 'presente'".

Vale destacar que o suporte para tal afirmação é de 1110 instâncias (96,62%), uma vez que apenas no caso de 38 dos 1148 discentes com desempenho caracterizado como 'bom' a situação da mãe era 'ausente' ou 'falecida'. Ainda, considerando-se a regra de inferência *modus tollens* (conforme Seção 4.1.1), pode-se afirmar que "SE a situação da mãe é 'ausente', ENTÃO o discente apresenta desempenho diferente de 'bom', portanto, poderia tanto apresentar um desempenho 'regular' ou 'insuficiente' quanto apresentar um desempenho 'ótimo'".

Ainda, um classificador Logit Boost (FRIEDMAN; HASTIE; TIBSHIRANI, 1998) foi capaz de classificar corretamente 68,97% dos dados do registro acadêmico, apresentando a seguinte observação:

- "Discentes que residem em 'pensionato' apresentaram desempenho abaixo de 'ótimo' ( $CR < 80$ )", sem nenhum contra exemplo.

Por fim, utilizando-se o classificador Naive Bayes (JOHN; LANGLEY, 1995), foi possível realizar as seguintes constatações nos dados do registro acadêmico:

1. 75,5% dos discentes com desempenho 'insuficiente' ( $CR < 60$ ) são do sexo 'masculino';
2. 63,9% dos discentes com desempenho 'ótimo' ( $CR \geq 80$ ) são do sexo 'feminino';

## 5 Conclusões e trabalhos futuros

A diferença entre os coeficientes de rendimento (CR) médios foi inferior a 2 pontos percentuais em quase todos os cenários analisados, sendo que alunos cotistas tiveram um desempenho ligeiramente melhor em alguns cenários de comparação. No *campus* Congonhas, ao avaliar os alunos por tipo de cota de ingresso, pode-se verificar que os alunos com melhor desempenho (CR médio) foram alunos que ingressaram por cotas. O desempenho dos alunos não cotistas do *campus* Formiga foi ligeiramente superior ao dos cotistas. Também não foi identificado um desempenho inferior dos alunos de acordo com o tipo de cota, ou seja, não é possível afirmar, por exemplo, que o desempenho de alunos da cota L5 é inferior (ou superior) aos demais.

Assim, através da análise das estatísticas descritivas dos registros acadêmicos, observa-se que não existe diferença estatística significativa em relação ao desempenho de alunos cotistas e não cotistas. Portanto, a avaliação do desempenho acadêmico dos alunos dos cursos técnicos integrados ao Ensino Médio do IFMG, nos *campi* Congonhas e Formiga, demonstrou que os alunos cotistas e não cotistas apresentaram desempenho semelhante de 2017 a 2019, intervalo de tempo dos dados analisados. Também, não foi detectado que alunos de determinados tipos de cota possuam desempenho inferior (ou superior) aos demais. Dessa forma, os resultados são compatíveis com os estudos realizados por Lima, Oliveira e Cruz (2020) no *campus* Ouro Branco do IFMG, e, também, compatíveis com os resultados dos trabalhos de Chaves *et al.* (2020) no IFRN, de Pironi (2016) no IFSP e de Pinto (2015) no IFBA.

Complementarmente, os resultados obtidos através do modelo de aprendizagem de máquina demonstraram regras de inferência (lógica proposicional) através das quais pode-se recomendar ao setor pedagógico e/ou diretoria de ensino do respectivo *campus* um acompanhamento mais específico para discentes que se encontrem em determinada situação socioeconômica (de vulnerabilidade) para a qual os algoritmos de aprendizagem de máquina identificaram determinada tendência de desempenho acadêmico insuficiente em relação aos seus pares. Por exemplo, alguns dos discentes com a mãe ausente ou com ambos os pais ausentes, alguns dos discentes que morem com o pai ou que residam com outros, para alguns dos discentes com os perfis identificados é recomendável um acompanhamento didático-pedagógico e/ou de assistência social, caso ainda não seja prática recorrente nas comunidades escolares referidas.

Porém, vale esclarecer que tais regras de inferência se aplicam apenas ao conjunto de dados analisado e, nele, se manifestaram em alguns dos discentes (não todos com determinado “perfil”). Portanto, tais regras de inferência não devem ser extrapoladas como se fossem “universais”. Discentes que se encaixem nos perfis identificados não estão destinados a apresentar pior (ou



melhor) desempenho acadêmico, assim como para algumas das regras de inferência encontradas (“perfis”) há contraexemplos dentro do conjunto de dados. Considerando-se o exposto, alguns dos resultados deste trabalho indicam que a utilização da ferramenta Weka para analisar os registros acadêmicos discentes (através de classificadores de aprendizado de máquina) pode ser um método valioso para auxiliar os setores de apoio ao discente e pedagógico na triagem e identificação de perfis discentes para os quais provavelmente uma maior atenção e acompanhamento didático-pedagógico seriam demandados.

### Referências

AGRAWAL, R.; SRIKANT, R. *Fast algorithms for mining association rules in large databases*. In: 20th International Conference on Very Large Data Bases. [S.l.]: Morgan Kaufmann, Los Altos, CA, 1994. p. 478–499.

AZEVEDO, C. M. M. de. *Cota racial e estado: abolição do racismo ou direitos de raça?* Cadernos de pesquisa, SciELO Brasil, v. 34, n. 121, p. 213–239, 2004. Disponível em: <https://doi.org/10.1590/S0100-15742004000100010>. Acesso em 10 mar. 2022.

CAMPOS, M. M. S. *A expansão da rede federal de educação profissional, técnica e tecnológica no período 2003-2014: uma análise a partir da abordagem das capacitações*. Revista Economia e Desenvolvimento, Santa Maria, v. 28, n. 2, p. 497–516, 2018. Disponível em: <https://doi.org/10.5902/1414650926166>. Acesso em 10 mar. 2022.

CAVALCANTI, I. T. d. N. et al. *Desempenho acadêmico e o sistema de cotas no ensino superior: evidência empírica com dados da Universidade Federal da Bahia*. Avaliação: Revista da Avaliação da Educação Superior (Campinas), SciELO Brasil, v. 24, n. 1, p. 305–327, 2019. Disponível em: <https://doi.org/10.1590/S1414-407720190001000016>. Acesso em 10 mar. 2022.

CHAVES, K. G. O. et al. *Permanência e êxito escolar: análise do desempenho acadêmico de estudantes ingressantes no ensino médio integrado do IFRN/Caicó através das cotas étnico-raciais no ano de 2014*. Revista Educação e Políticas Em Debate, v. 9, n. 3, p. 672–691, 2020. Disponível em: <https://doi.org/10.14393/REPOD-v9n3a2020-57880>. Acesso em 10 mar. 2022.

COHEN, W. W. *Fast effective rule induction*. In: Twelfth International Conference on Machine Learning. [S.l.]: Morgan Kaufmann, 1995. p. 115–123. Disponível em: <https://doi.org/10.1016/B978-1-55860-377-6.50023-2>. Acesso em 10 mar. 2022.

DALLABONA, C. A.; FILHO, M. F. de O. S. *Desempenho acadêmico de estudantes oriundos de escolas públicas: cursos de graduação do campus Curitiba da UTFPR*. In: Anais do Congresso Brasileiro de Educação em Engenharia. Blumenau, SC, Brasil. [S.l.: s.n.], 2011. v. 39.

FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. *Additive Logistic Regression: a Statistical View of Boosting*. Stanford University, 1998. Disponível em: <https://doi.org/10.1214/aos/1016218223>. Acesso em 10 mar. 2022.

- FRY, P.; MAGGIE, Y. *Cotas raciais: construindo um país dividido*. *Econômica*, v. 6, n. 1, p. 153–161, 2004. Disponível em: <https://doi.org/10.22409/economica.6i1.p202>. Acesso em 10 mar. 2022.
- IENSUE, G. et al. *Política de cotas raciais em universidades brasileiras: entre a legitimidade e a eficácia*. UNIVERSIDADE ESTADUAL DE PONTA GROSSA, 2009.
- JOHN, G. H.; LANGLEY, P. *Estimating continuous distributions in bayesian classifiers*. In: Eleventh Conference on Uncertainty in Artificial Intelligence. San Mateo: Morgan Kaufmann, 1995. p. 338–345.
- LIMA, C. N. de; OLIVEIRA, A. R. de; CRUZ, T. L. B. da. *Análise estatística do desempenho de alunos cotistas versus não cotistas: Um estudo sobre o rendimento escolar de estudantes de curso técnico integrado*. *Revista Brasileira da Educação Profissional e Tecnológica*, v. 1, n. 18, p. 7900, 2020. Disponível em: <https://doi.org/10.15628/rbept.2020.7900>. Acesso em 10 mar. 2022.
- MAXIMO, R. de O.; GANDOLFI, P. E.; LOPES, J. E. F. *Cotas universitárias: Estudo do desempenho acadêmico na graduação após a implementação da política de cotas na Universidade Federal de Uberlândia*. *Revista Educação e Políticas em Debate*, v. 9, n. 3, p. 636–654, 2020. Disponível em: <https://doi.org/10.14393/REPOD-v9n3a2020-57854>. Acesso em 10 mar. 2022.
- NASCIMENTO, C. J. P. do. *Políticas públicas: cotas raciais à luz da constituição brasileira*. In: CONGRESSO INTERINSTITUCIONAL UNISC/URCA. [S.l.: s.n.], 2017.
- PEIXOTO, A. de L. A. et al. *Cotas e desempenho acadêmico na UFBA: um estudo a partir dos coeficientes de rendimento*. Avaliação: Revista da Avaliação da Educação Superior (Campinas), SciELO Brasil, v. 21, n. 2, p. 569–592, 2016. Disponível em: <https://doi.org/10.1590/S1414-40772016000200013>. Acesso em 10 mar. 2022.
- PIERONI, A. R. *Um estudo sobre o desempenho de alunos cotistas e não cotistas no Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – Campus Salto*. Dissertação (Mestrado) — Universidade Federal de Alfenas, campus Varginha, setembro 2016.
- PINTO, J. de C. A. *Desempenho acadêmico e fatores socioeconômicos: uma análise no ensino médio integrado do IFBA/Campus Barreiras*. Dissertação (Mestrado) — Faculdade de Educação da Universidade Federal da Bahia, julho 2015.
- ROSEN, K. *Matemática discreta e suas aplicações*. São Paulo: McGraw-Hill, 2009.
- SANTOS, C. T. *Ações afirmativas no ensino superior: análise do perfil socioeconômico e da experiência universitária de bolsistas do PROUNI na PUC-Rio*. *Revista Brasileira de Estudos Pedagógicos*, SciELO Brasil, v. 93, n. 235, p. 770–790, 2012. Disponível em: <https://doi.org/10.24109/2176-6681.rbep.93i235.415>. Acesso em 10 mar. 2022.
- SILVA, B. C. M.; XAVIER, W. S.; COSTA, T. de M. T. da. *Sistema de cotas e desempenho: uma comparação entre estudantes cotistas e não cotistas*. *Administração Pública e Gestão Social*, 2020. Disponível em: <https://doi.org/10.21118/apgs.v12i3.6125>. Acesso em 10 mar. 2022.
- SOUZA, A. C. de; BRANDALISE, M. Â. T. *Política de cotas e democratização do ensino superior: a visão dos implementadores*. *Revista Internacional de Educação Superior*, v. 3, n. 3, p. 515–538, 2017. Disponível em: <https://doi.org/10.22348/riesup.v3i3.7763>. Acesso em 10 mar. 2022.

TRAGTENBERG, M. H. R. *Programa de ações afirmativas da Universidade Federal de Santa Catarina: resultados preliminares*. Revista Gestão Universitária na América Latina-GUAL, Universidade Federal de Santa Catarina, v. 3, n. 1, p. 1–16, 2010. Disponível em: <https://doi.org/10.5007/1983-4535.2010v3n1p144>. Acesso em 10 mar. 2022.

VELLOSO, J. *Cotistas e não-cotistas: rendimento de alunos da Universidade de Brasília*. Cadernos de pesquisa, SciELO Brasil, v. 39, n. 137, p. 621–644, 2009. Disponível em: <http://doi.org/10.1590/S0100-15742009000200014>. Acesso em 10 mar. 2022.

ZUIN, A. L. A.; BASTOS, E. *A justiça social por meio das cotas na Universidade Federal de Rondônia*. Educação & Formação, v. 4, n. 12set/dez, p. 104–123, 2019. Disponível em: <https://doi.org/10.25053/redufor.v4i12.945>. Acesso em 10 mar. 2022.

