

INSTITUTO FEDERAL DE MINAS GERAIS
CAMPUS OURO PRETO
ESPECIALIZAÇÃO EM INTELIGÊNCIA ARTIFICIAL

JOELMA DE FÁTIMA RODRIGUES BATISTA FREITAS

**ANÁLISE PREDITIVA DO DESEMPENHO ACADÊMICO NO ENSINO TÉCNICO
INTEGRADO AO ENSINO MÉDIO: aplicação de estatística, séries temporais e redes
neurais em uma escola profissionalizante da rede federal**

OURO PRETO-MG

Junho/2025



MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE MINAS GERAIS
Campus Ouro Preto
Diretoria de Pesquisa, Inovação e
Pós-Graduação Coordenação do Curso de
Pós-graduação em Inteligência Artificial
Rua Pandiá Calógeras, 898 - Bairro Bauxita - CEP 35400-000 - Ouro Preto - MG
- www.ifmg.edu.br

FOLHA DE APROVAÇÃO

JOELMA DE FÁTIMA RODRIGUES BATISTA FREITAS

ANÁLISE PREDITIVA DO DESEMPENHO ACADÊMICO NO ENSINO TÉCNICO INTEGRADO AO ENSINO MÉDIO: APLICAÇÃO DE ESTATÍSTICA, SÉRIES TEMPORAIS E REDES NEURAIS EM UMA ESCOLA PROFISSIONALIZANTE DA REDE FEDERAL

Trabalho de Conclusão de Curso apresentado ao curso de ESPECIALIZAÇÃO EM INTELIGÊNCIA ARTIFICIAL, ofertado pelo Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais - *Campus* Ouro Preto, como parte dos requisitos para a obtenção do título de ESPECIALISTA EM INTELIGÊNCIA ARTIFICIAL.

Aprovado em 11 de julho de 2025, pela Banca Examinadora:



Documento assinado eletronicamente por Silvia Grasiella Moreira Almeida, Responsável pelas Atividades da Coordenação do Curso de Pós-graduação Lato Sensu em Inteligência Artificial, em 04/08/2025, às 19:17, conforme Decreto nº 10.543, de 13 de novembro de 2020.



Documento assinado eletronicamente por Viviane de Paula Silva, Membro da Banca Examinadora, em 07/08/2025, às 09:59, conforme Decreto nº 10.543, de 13 de novembro de 2020.



Documento assinado eletronicamente por Hudney Alves Faria de Carvalho, Membro da Banca Examinadora, em 07/08/2025, às 10:33, conforme Decreto nº 10.543, de 13 de novembro de 2020.



Documento assinado eletronicamente por Fabiano Gomes da Silva, Membro da Banca Examinadora, em 11/08/2025, às 13:46, conforme Decreto nº 10.543, de 13 de novembro de 2020.



A autenticidade do documento pode ser conferida no site <https://sei.ifmg.edu.br/consultadocs> informando o código verificador 2404346 e o código CRC F4C53D27.

23213.001857/2025-11

2404346v1

F866a Freitas, Joelma de Fátima Rodrigues Batista.
Análise preditiva do desempenho acadêmico no Ensino Técnico integrado ao Ensino Médio [manuscrito] : aplicação de estatística, séries temporais e redes neurais em uma escola profissionalizante da rede federal / Joelma de Fátima Rodrigues Batista Freitas. – 2025.
65 f. : il.

Orientador: Fabiano Gomes da Silva.

Coorientadora: Viviane de Paula Silva.

Trabalho de Conclusão de Curso (especialização) – Instituto Federal de Minas Gerais. *Campus* Ouro Preto, 2025.

1. Rendimento escolar. 2. Séries temporais. 3. Redes Neurais (Computação). I. Silva, Fabiano Gomes da. II. Silva, Viviane de Paula. III. Instituto Federal de Minas Gerais. *Campus* Ouro Preto. III. Título.

CDU: 004.8

Catálogo: Kelly Cristiane Santos Moraes - CRB-6/3217

DEDICATÓRIA

Dedico esse estudo à minha Família, que é o meu alicerce de sempre.

AGRADECIMENTOS

Primeiramente, agradeço a **Deus** pela vida, pelas oportunidades e por me guiar em cada passo dessa jornada.

Aos meus queridos pais, Rosa e Hélio, deixo o meu mais profundo reconhecimento, pois são exemplos de amor, dedicação e luta, sempre me inspirando com seus ensinamentos e valores.

Ao meu esposo, Ronaldo, agradeço pelo apoio incondicional, pela paciência, compreensão e amor que tornaram essa caminhada mais leve.

Aos meus filhos, José Arthur e João Pedro, agradeço pela singela e doce companhia, que enche meus dias de alegria e motivação.

Aos meus irmãos, Elaine, Marco e Elizabeth, deixo meus agradecimentos pela coragem compartilhada, pelo incentivo e por estarem sempre presentes.

Aos meus sobrinhos, que são muitos, mas igualmente especiais, agradeço pelo carinho que aquece o coração e fortalece os laços familiares.

A toda a minha família, que de tantas formas contribuiu para que eu chegasse até aqui, sou grata por tudo o que fizeram por mim.

Um agradecimento especial **aos meus queridos orientadores, o Professor Fabiano, a Técnica Viviane, aos Professores Carlos Severiano e André Klen**, pela paciência, profissionalismo, escuta atenta e pela valiosa experiência compartilhada ao longo desse percurso.

À Profa. Sílvia Grasiella, agradeço pela dedicação e compromisso com a administração do curso e com a formação acadêmica de todos nós.

Ao colega Aislan Rondinelli, agradeço pela ajuda e boa vontade no tratamento dos dados. Ajudou muito!

Aos colegas de curso e de trabalho, pela companhia, pelo apoio mútuo e pelas trocas enriquecedoras ao longo dessa jornada, serei eternamente grata.

Agradeço ao Instituto Federal de Minas Gerais – IFMG – na figura do curso de Inteligência Artificial, pelos ensinamentos que com certeza irão me acompanhar ao longo da minha trajetória pessoal e profissional e pela oportunidade de estudo e trabalho.

A cada pessoa que, direta ou indiretamente, contribuiu para a realização deste trabalho, minha eterna gratidão.

“O Senhor é o meu Pastor, nada me faltará”

SALMO 23:1

RESUMO

O presente trabalho analisa o desempenho acadêmico dos alunos do primeiro ano do Ensino Técnico Integrado ao Ensino Médio em uma escola profissionalizante da Rede Federal de Ensino, com ênfase na alteração do modelo avaliativo que ocorreu entre os períodos de 2014 a 2017 e de 2018 a 2023. Ele compara os resultados das avaliações, que passaram de quatro para três etapas, avaliando o impacto dessa mudança no rendimento dos estudantes nas disciplinas do ciclo básico. Utilizam-se dados quantitativos referentes às notas, frequência escolar e variáveis socioeconômicas, como renda familiar e condições de moradia. A análise estatística demonstra que a redução do número de etapas avaliativas não prejudicou o desempenho dos alunos. Observa-se uma leve melhoria nas médias gerais no período mais recente, indicando que o novo formato pode favorecer um acompanhamento mais eficaz da aprendizagem. A frequência escolar mostra-se um fator significativo para o sucesso acadêmico, destacando a importância do estímulo à presença dos estudantes. Além disso, identificam-se desigualdades relacionadas ao contexto socioeconômico que influenciam o aproveitamento escolar, reforçando a necessidade de políticas educacionais direcionadas à promoção da equidade. Complementarmente, aplicam-se modelos preditivos de inteligência artificial, incluindo redes neurais e Séries Temporais, para se prever o desempenho acadêmico dos alunos. Esses modelos apresentam bom desempenho, possibilitando a antecipação de resultados e a orientação de intervenções pedagógicas personalizadas. A integração dessas ferramentas tecnológicas contribui para a melhoria do processo de gestão educacional e para o desenvolvimento de estratégias mais eficazes de ensino. Conclui-se que a mudança do modelo avaliativo de quatro para três etapas apresenta **resultados positivos para o desempenho escolar, sem comprometer a avaliação do aprendizado. A incorporação de técnicas avançadas de análise de dados amplia o potencial de monitoramento e intervenção no contexto educacional.**

Palavras-chave: Desempenho Escolar. Séries Temporais. Redes Neurais. Inteligência Artificial.

ABSTRACT

This research has analyzed the academic performance of first-year of technical integrated high school students (technical and int at a vocational school in the Federal Network, with an emphasis on the change in the assessment model that occurred between 2014 and 2017 and between 2018 and 2023). The study has compared the results of the assessments, which turned from four to three stages in order to assess the impact of this change on students' performance in basic cycle subjects. Quantitative data are used from regarding grades, school attendance and socioeconomic variables, such as family income and housing conditions. The statistical analysis shows that the reduction in the number of assessment stages did not harm student performance. A slight improvement in overall averages had been observed at the most recent period, indicating that the new format might favor more effective monitoring of learning. School attendance is shown to be a significant factor for academic success, highlighting the importance of encouraging student attendance. In addition, inequalities related to the socioeconomic context that influence academic performance were identified, reinforcing the need for educational policies aimed at promoting equity. In addition, predictive artificial intelligence models, including neural networks and time series, were applied to predict students' academic performance. These models performed some good performance, enabling the anticipation of results and the guidance of personalized pedagogical interventions. The integration of these technological tools has contributed to the improvement of the educational management process and to the development of more effective teaching strategies. It has been concluded that the changing the evaluation model to three stages has presented positive results for academic performance, without compromising the assessment of learning. The incorporation of advanced techniques on data has increased the potential for monitoring and intervention in the educational context.

Keywords: Academic Performance. Time Series. Neural Networks. Artificial Intelligence.

LISTA DE GRÁFICOS

Gráfico 1 - Distribuição de notas por disciplina no período de 2018 a 2023	35
Gráfico 2 - Distribuição Percentual por Residência (2014-2017)	36
Gráfico 3 - Distribuição Percentual por Residência (2018-2023)	37
Gráfico 4 - Distribuição percentual por residência (total).....	37
Gráfico 5 - Distribuição Percentual por Raça/Cor (2014 – 2017).....	39
Gráfico 6 - Distribuição Percentual por Raça/Cor (2018 – 2023)	39
Gráfico 7 - Distribuição percentual por Raça/Cor (Total).....	40
Gráfico 8 - Distribuição das Notas Finais por Raça/Cor (2014-2017)	42
Gráfico 9 - Distribuição por Notas Finais por Raça/Cor (2018-2013).....	42
Gráfico 10 - Média da Renda Familiar por Ano com Variação Percentual no período de 2018 a 2023	43
Gráfico 11 - Distribuição Percentual por Gênero (2014-2017).....	44
Gráfico 12 - Distribuição Percentual por Gênero (2018-2023).....	45
Gráfico 13 - Distribuição Percentual por Gênero (Total).....	45

LISTAS DE TABELAS

Tabela 1	Estatísticas descritivas das notas no período de 2014 a 2017.....	35
Tabela 2	Estatísticas descritivas das notas no período de 2018 a 2023.....	35
Tabela 3	Média e Mediana das Notas por Raça/Cor (2014-2017).....	41
Tabela 4	Média e Mediana das Notas por Raça/Cor (2018-2023).....	41
Tabela 5	Estatísticas Descritivas da Frequência (2014-2017).....	47
Tabela 6	Estatísticas Descritivas da Frequência (2018-2023).....	47
Tabela 7	Taxa de Aprovação 2014-2023.....	51
Tabela 8	Correlação de Pearson (2014 – 2023).....	56
Tabela 9	Correlação de Spearman (2014 a 2017).....	57
Tabela 10	Correlação de Spearman (2018 – 2023).....	58
Tabela 11	Resultados de Acurácia (2014 – 2023).....	59
Tabela 12	Resultados de Acurácia (2018-2023).....	61

LISTA DE SIGLAS E ABREVIATURAS

ADF	Teste de Dickey Fuller
AR	Componentes Autorregressivos
ARIMA	Média Móvel Integrada Autorregressiva
ARMA	Média Móvel Autorregressiva
BNCC	Base Nacional Comum Curricular
ENEM	Exame Nacional de Ensino Médio
I	Integração
IA	Inteligência Artificial
IDEB	Índice de Desenvolvimento da Educação Básica
MA	Médias Móveis
MLP	Perceptron Multicamadas (<i>Multilayer Perceptron</i>)
PNE	Plano Nacional de Educação
PSO	Exame de Partículas
RNA	Redes Neurais Artificiais
SHAP	<i>SHapley Additive exPlanations</i>

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Objetivos	16
1.1.1	Objetivo Geral	16
1.1.2	Objetivos Específicos	16
1.2	Estrutura Organizacional do Estudo	16
2	DESEMPENHO ACADÊMICO NO ENSINO MÉDIO: FATORES E IMPLICAÇÕES	18
3	SÉRIES TEMPORAIS	21
3.1	Componentes das Séries Temporais	21
3.2	Modelos Matemáticos de Séries Temporais	22
3.3	Aplicações em Educação	22
3.4	Desafios na Modelagem de Séries Temporais	23
3.5	Contribuições das Séries Temporais para a Tomada de Decisão	23
3.6	Modelo ARIMA	23
3.6.1	Componente do Modelo ARIMA	24
3.6.1.1	<i>Processo Autorregressivo</i>	24
3.6.1.3	<i>Processo de Médias Móveis (MA)</i>	25
3.6.2	Identificação e Estimação de um Modelo ARIMA	25
3.7	Aplicações do Modelo ARIMA	26
4	REDES NEURAIIS (PERCEPTRON MULTICAMADAS)	27
5	METODOLOGIA	29
5.1	Classificação da Pesquisa Quanto ao seu Objetivo	29
5.2	Classificação da Pesquisa quanto à sua Natureza	29
5.3	Classificação da Pesquisa Quanto ao Objeto de Estudo	30
5.4	Classificação da Pesquisa Quanto à Técnica de Coleta de Dados	30
5.4.1	Pesquisa Bibliográfica e Documental	31
6	RESULTADOS E DISCUSSÕES	33
6.1	Análise Exploratória do Desempenho Acadêmico com Ênfase na Mudança do Número de Etapas Avaliativas (2014-2023)	33
6.2	Análise das Variáveis Sociais e Demográficas	36
6.3	Impacto da Pandemia da Covid-19	47
6.4	Inferência sobre a Mudança de Quatro para Três Etapas	48

6.5	Considerações e Recomendações	48
6.6	Análise de Percentuais	49
6.6.1	Análise da Taxa de Aprovação por Etapa Avaliativa nos Períodos de 2014-2017 e de 2018-2013	49
6.6.2	Resultados do Segundo Código: Análise por Período Letivo e Modelagem Temporal	52
6.6.3	Diferenças entre os Códigos	53
6.6.4	Qualidade e Aplicabilidade dos Resultados	53
6.7	Correlação Pearson e Correlação Spearman para Análise das Correlações entre Desempenho Acadêmico e Variáveis Socio econômicas (2014-2023)	54
6.7.1	Análise das Correlações entre Variáveis (2014-2017 e 2018-2023)	56
6.8	Redes Neurais	59
6.8.1	Análise Comparativa dos Modelos: Redes Neurais, Random Forest e Regressão Logística	59
6.8.1.2	<i>Confiabilidade e Robustez</i>	60
6.8.1.3	<i>Aplicabilidade Prática</i>	60
6.9	Considerações Importantes	61
7	CONSIDERAÇÕES FINAIS	62
	REFERÊNCIAS BIBLIOGRÁFICAS	63

1 INTRODUÇÃO

A educação é uma das áreas mais essenciais e desafiadoras na construção de uma sociedade desenvolvida, justa e equânime. Com os avanços tecnológicos e a necessidade de adaptar o ensino às demandas contemporâneas, métodos tradicionais de avaliação de desempenho acadêmico mostram-se limitados para se captar toda a complexidade do processo de aprendizagem. Por isso, novas abordagens, como o uso de inteligência artificial, ganharam relevância, abrindo caminho para análises mais profundas e preditivas sobre o comportamento e o sucesso escolar dos alunos ao longo do tempo.

Esta pesquisa teve origem a partir do desejo de se contribuir para o entendimento e a análise do desempenho acadêmico por meio de abordagens modernas e integradas, combinando técnicas de inteligência artificial, como redes neurais e Séries Temporais. O histórico da pesquisadora influenciou essa gênese já que possui formação que inclui uma especialização em Inteligência Artificial (IA), iniciada em 2022, além de experiência prévia com modelagem matemática, ferramentas que auxiliaram na condução do estudo. Por atuar na Rede Federal de Ensino, a pesquisadora vivencia diariamente os desafios do Ensino Médio e Técnico-Profissionalizante, convivendo com alunos de realidades diversas. Essa vivência despertou a motivação para se desenvolver um estudo que não apenas explorasse o desempenho acadêmico ao longo do tempo, mas também oferecesse ferramentas para se entenderem as variáveis que o influenciam, utilizando o modelo ARIMA (Média Móvel Integrada Autorregressiva) como uma das principais técnicas de análise e redes neurais (*Perceptron Multicamadas*).

Aplicou-se a Inteligência Artificial como suporte avançado para se identificarem padrões e se preverem comportamentos, permitindo uma compreensão mais profunda das dinâmicas educacionais. Por meio da abordagem integrada, o estudo buscou não apenas descrever o desempenho acadêmico, mas também propor soluções e estratégias educacionais fundamentadas em evidências, contribuindo para a melhoria da qualidade do ensino e para o desenvolvimento de políticas pedagógicas mais eficazes e direcionadas.

A partir do cenário acima, evidenciou-se a questão da pesquisa apresentada, a saber: ***“Como a redução do número de etapas de avaliação de quatro para três, ao longo do período de 2014 a 2023, impactou o desempenho acadêmico dos alunos (4286 – número total vinculado a todas as disciplinas propostas), considerando as variáveis demográficas, socioeconômicas e acadêmicas? ”.***

Objetivando esclarecer a questão acima, registra-se que o sistema de avaliação do IFMG, regulamentado institucionalmente, estabelece a utilização de instrumentos avaliativos diversificados ao longo das etapas, como provas (objetivas, discursivas, orais ou práticas), trabalhos, seminários, portfólios, autoavaliações, entre outros, garantindo que a aferição do aprendizado não se limite a apenas um tipo de atividade (§ 2º do Art. 112 do Regulamento de Ensino). Além disso, o mesmo regulamento define que os cursos integrados devem se organizar em três etapas avaliativas por módulo anual, sendo atribuídos 30 pontos na primeira etapa, 35 na segunda e 35 na terceira. Essa mudança, que substituiu o antigo formato de quatro etapas, buscou adequar o processo avaliativo a uma lógica mais integrada, reduzindo o número de avaliações, mas mantendo a exigência de diversidade nos instrumentos de aferição.

A relevância desse regulamento para esta pesquisa está no fato de que ele marca a transição do sistema de quatro para três etapas avaliativas, justamente o fenômeno analisado nesse estudo. Assim, compreender essa base normativa é fundamental para avaliar se a mudança trouxe impactos positivos ou negativos sobre o desempenho acadêmico dos estudantes.

A proposta visa não só estudar o desempenho acadêmico, mas também entender como fatores de perfil dos discentes – como frequência escolar, situação socioeconômica dos pais, gênero e raça – influenciam nesse desempenho ao longo do tempo, utilizando, para isso, técnicas de inteligência artificial.

A partir da pergunta-problema apresentada, buscou-se examinar a interação entre os perfis dos alunos e seus desempenhos, oferecendo uma análise longitudinal que permitisse identificar padrões e antecipar desafios ao longo de uma década (2014 a 2023).

Quanto à relevância desse estudo, ela se justifica no potencial de inovação que representa ao aplicar técnicas de inteligência artificial em um campo essencial, como a educação, especialmente em um contexto de rede federal de ensino profissionalizante. Busca-se resultados que possam fornecer uma base sólida para a compreensão de fatores que influenciam o sucesso acadêmico, contribuindo para a criação de políticas educacionais mais assertivas e para a implementação de práticas pedagógicas adaptativas que valorizem o desenvolvimento individual dos alunos e promovam a equidade educacional.

Ao oferecer uma abordagem inovadora para análise do desempenho e perfil dos alunos, o estudo não só contribui para a literatura acadêmica, mas também oferece uma ferramenta prática para educadores e gestores, permitindo uma visão mais clara e integrada do processo de ensino e aprendizagem. Dessa forma, busca promover uma maior compreensão sobre como os diferentes fatores interagem e afetam o sucesso escolar, possibilitando que a

inteligência artificial seja aplicada para apoiar a construção de um ambiente educacional mais inclusivo, eficiente e orientado ao desenvolvimento integral dos alunos.

1.1 Objetivos

1.1.1 Objetivo Geral

O objetivo geral da pesquisa é analisar o perfil e o desempenho acadêmico de alunos do primeiro ano do Ensino Técnico Integrado ao Ensino Médio em uma escola profissionalizante da rede federal nas disciplinas de formação geral do ciclo básico (Português, Matemática, Física, Química, Biologia, Geografia e História) ao longo de dez anos (2014 a 2023) utilizando o modelo ARIMA como uma das principais técnicas de análise e redes neurais (*Perceptron Multicamadas*). Essas disciplinas foram selecionadas visto que fazem parte de todos os cursos técnicos e também são importantes quanto ao ENEM – Exame Nacional de Ensino Médio.

1.1.2 Objetivos Específicos

- Analisar a evolução do desempenho dos alunos nas disciplinas do ciclo básico ao longo do período entre 2014 a 2023 (dez anos), identificando padrões e tendências;
- Caracterizar o perfil dos alunos com base em variáveis como frequência escolar, situação socioeconômica dos pais, gênero e raça, e examinar como esses fatores influenciam o desempenho acadêmico;
- Aplicar técnicas de Séries Temporais e redes neurais para identificar padrões de desempenho e possíveis fatores preditores de sucesso escolar ao longo do tempo;
- Analisar o impacto de mudanças socioeconômicas e políticas educacionais sobre o perfil e desempenho dos alunos ao longo dos dez anos.

1.2 Estrutura Organizacional do Estudo

Este estudo está composto de oito capítulos, incluindo esta introdução e as referências bibliográficas utilizadas. A introdução será, portanto, o primeiro capítulo.

O capítulo 2 trata de analisar os fatores e suas implicações no desempenho acadêmico dos alunos do Ensino Técnico Integrado ao Ensino Médio, amostra utilizada nesse estudo. Apontam-se aspectos sociais e financeiros vinculados à rotina dos estudantes bem como a questão da carência de formação continuada de alguns professores, além de metodologias que não se atualizam, o que conduz o aluno à desmotivação e até mesmo à evasão.

O capítulo 3 analisa o conceito, os componentes e os modelos das Séries Temporais e sua aplicação no campo da educação. Também discrimina os desafios de seu uso nesse campo e o aponta como importante para tomadas de decisão. Da mesma forma, trata especificamente do modelo ARIMA (Média Móvel Integrada Autorregressiva): conceito, aplicação nos variados campos do conhecimento. Indica-se sua utilização como recurso metodológico no presente estudo.

O capítulo 4 apresenta o conceito e a aplicação das Redes Neurais no campo da educação e especificamente no estudo apresentado pela autora.

O capítulo 5 tem como objetivo apresentar a metodologia a que se recorreu para desenvolvimento desse estudo. Segue o capítulo 6 com os resultados obtidos e suas discussões e o denominado como “Considerações Finais”. Por fim, apresentam-se as referências bibliográficas.

2 DESEMPENHO ACADÊMICO NO ENSINO TÉCNICO INTEGRADO AO ENSINO MÉDIO: FATORES E IMPLICAÇÕES

Ao ingressarem no primeiro ano do Ensino Técnico Integrado ao Ensino Médio, os alunos enfrentam uma série de desafios que podem impactar significativamente seu desempenho acadêmico. Provenientes de diferentes instituições de Ensino Fundamental, esses jovens deparam-se com a necessidade de adaptação a uma nova estrutura escolar, o que inclui o conhecimento que eles possuem quanto às disciplinas ofertadas e maiores exigências acadêmicas. Além disso, como adolescentes, carregam expectativas de um futuro promissor, mas trazem consigo dificuldades acumuladas ao longo da trajetória escolar anterior.

O desempenho acadêmico é, sem dúvida, um dos indicadores mais importantes para mensurar o sucesso escolar e a qualidade da educação oferecida. Porém, ele não é determinado apenas pelas notas dos alunos; fatores socioeconômicos, culturais, pedagógicos e as características individuais de cada estudante também desempenham papéis cruciais quanto a ele. No contexto do Ensino Médio, essas variáveis se tornam ainda mais evidentes, uma vez que as desigualdades sociais e educacionais se ampliam.

Silva (2023) discrimina vários fatores que podem influenciar direta ou indiretamente no desempenho do aluno no ensino médio. Entre eles, o autor salienta a prática e a formação dos docentes, a estrutura física e organizacional da instituição escolar, além de atributos próprios dos discentes que vão desde sua origem étnica até o local onde vivem/moram, perpassando problemas sociais em que podem estar inseridos. Também evidencia a necessidade de atualização das práticas metodológicas de forma que o aluno se sinta motivado para continuar seus estudos.

Araújo *et al* (2021) registram que a maioria dos fatores que impactam negativa ou positivamente o desempenho escolar de adolescentes se vinculam a questões em que a desigualdade social e econômica é percebida. Os autores expressam que também é preciso especial atenção às metodologias aplicadas para oferta de conteúdos bem como à experiência e a formação adequada dos professores.

A literatura acadêmica, como se pode inferir, aponta que diversos fatores influenciam diretamente o desempenho escolar. A motivação e o interesse dos alunos, por exemplo, são frequentemente apontados como desafios cruciais no Ensino Médio. Segundo Santana e Santos (2025), a falta de motivação e o desinteresse nas escolas públicas, especialmente no Ensino Médio, são problemas significativos que impactam o desempenho dos estudantes. A escassez de recursos didáticos adequados, como aulas práticas em laboratório e

equipamentos multimídia, contribui para a falta de engajamento dos alunos. Esses recursos, segundo os autores, são essenciais para atrair os jovens para o processo de aprendizagem, tornando o conteúdo mais relevante e contextualizado. O estudo reforça a importância de se adotarem estratégias pedagógicas que incentivem a participação ativa dos discentes, mesmo diante da limitação de recursos materiais.

Além dos fatores acima descritos, outro que influencia no desempenho escolar dos alunos em análise, a saber, é a compreensão de leitura, o que tem um papel decisivo quanto a ele. Pires e Mota (2021) destacam que a habilidade de se entenderem textos literários e científicos é um fator preditivo importante do sucesso nas disciplinas de ciências, como física, química e biologia. O estudo revelou que alunos com melhor compreensão de textos, especialmente os científicos, apresentam melhores resultados acadêmicos. Nesse sentido, as estratégias pedagógicas que envolvem a leitura de diferentes tipos de textos e o trabalho com vocabulário são essenciais para se melhorar o desempenho dos alunos. O desenvolvimento de habilidades de leitura crítica, portanto, é uma competência fundamental para o sucesso acadêmico no Ensino Médio.

Thornton e Corso (2022) esclarecem que a não proficiência em leitura impacta diretamente os resultados do ENEM – Exame Nacional de Ensino Médio – e é um dos fatores mais desfavoráveis quanto ao desempenho escolar dos alunos. Aponta, todavia, que esse fator é vinculado a diversas questões proveniente de outros fatores como, por exemplo, o grau acadêmico dos pais, a dificuldade para compra de livros e a ausência de boas bibliotecas nas instituições desde a pré-escola.

Outro aspecto relevante quanto ao tema em epígrafe é a formação dos docentes. Costa, Brito e Waltenberg (2020) afirmam que a formação inadequada dos professores nas disciplinas que lecionam está diretamente relacionada aos baixos resultados acadêmicos no Ensino Médio. Quando os educadores não possuem formação específica para as disciplinas das quais são os regentes, como é o caso em muitas escolas públicas brasileiras, o rendimento discente tende a ser prejudicado, resultando em altas taxas de evasão escolar e distorção idade-série. A qualidade do ensino, portanto, depende não só do domínio dos conteúdos pelos professores, mas também de sua capacidade pedagógica para trabalhar esses conteúdos de forma eficaz e eficiente. A integração do conhecimento disciplinar com a prática pedagógica se torna, assim, essencial para o sucesso acadêmico dos estudantes.

Além de fatores pedagógicos, as características pessoais e socioeconômicas dos alunos também têm um impacto significativo em seu desempenho. Machado *et al.* (2025) identificaram que os fatores como a cor da pele, gênero e nível socioeconômico são

determinantes importantes no rendimento dos estudantes brasileiros. O estudo indicou que alunos de famílias de baixa renda e aqueles que pertencem a grupos raciais minoritários tendem a apresentar desempenho inferior em comparação com seus colegas de classes mais privilegiadas. As escolas privadas, por sua vez, costumam apresentar um desempenho superior, em grande parte devido a essas disparidades socioeconômicas. Fatores como a reprovação escolar e defasagem idade-série também afetam negativamente o desempenho dos alunos, especialmente no Ensino Médio bem como contribuem para a perpetuação das desigualdades no acesso a recursos e oportunidades educacionais.

Considerando o contexto acima apresentado, a análise do desempenho acadêmico no Ensino Médio, especialmente em escolas públicas, deve levar em conta uma gama de fatores interligados. Variáveis como motivação, qualidade docente, compreensão de leitura e características socioeconômicas são determinantes cruciais para entender as disparidades no desempenho escolar. O estudo dessas variáveis é fundamental para a formulação de políticas educacionais mais eficazes e inclusivas capazes de reduzir as desigualdades e promoverem um ambiente de aprendizagem mais equitativo para todos os alunos. Para esse fim, algumas ferramentas são imprescindíveis para essa formulação, entre elas as Séries Temporais e o modelo ARIMA.

3 SÉRIES TEMPORAIS

As Séries Temporais são ferramentas analíticas fundamentais em várias áreas do conhecimento, consistindo em dados coletados sequencialmente ao longo de um período de tempo, com intervalos regulares. Segundo Morettin (1987), essas séries são caracterizadas pela dependência serial, isto é, a relação entre os valores observados em diferentes instantes de tempo. Essa característica única permite a análise e previsão de eventos futuros com base em padrões passados, sendo amplamente utilizada em campos como economia, saúde, educação e meteorologia.

3.1 Componentes das Séries Temporais

As Séries Temporais são compostas por elementos distintos que influenciam diretamente suas dinâmicas. De acordo com Enders (2014), elas podem ser decompostas nos seguintes componentes:

- **Tendência:** representa o comportamento de longo prazo da série, como crescimento ou declínio sistemático ao longo dos anos (MORETTIN, 1987; ENDERS, 2014).
- **Sazonalidade:** reflete padrões que se repetem em intervalos regulares, geralmente associados a fatores como clima ou ciclos econômicos (BOX; JENKINS, 1970).
- **Ciclos:** movimentos ondulatórios que se estendem por períodos mais longos, influenciados por fenômenos estruturais ou econômicos (OLIVEIRA, 2015).
- **Ruído aleatório:** refere-se às variações imprevisíveis nos dados, que não podem ser modeladas diretamente e são tratadas como erro (MORETTIN, 1987; ENDERS, 2014).

Esses componentes desempenham um papel crucial na modelagem e previsão das séries, permitindo que analistas identifiquem padrões subjacentes e extraiam informações significativas dos dados.

3.2 Modelos Matemáticos de Séries Temporais

Os métodos para análise de Séries Temporais evoluíram consideravelmente ao longo das décadas. Os modelos ARIMA, introduzidos por Box e Jenkins (1970), são amplamente reconhecidos pela capacidade de modelar séries estacionárias. Esses modelos combinam componentes autorregressivos (AR), médias móveis (MA) e integração para lidar com séries não estacionárias. Conforme Enders (2014), a estacionaridade é uma característica essencial dos modelos, pois séries com média e variância constantes ao longo do tempo são mais adequadas para modelagem preditiva.

Quando as séries apresentam padrões não lineares, técnicas complementares, como Redes Neurais Artificiais (RNA), têm se mostrado eficazes. Zhang (2001) argumenta que a combinação de métodos lineares e não lineares melhora a precisão preditiva, uma vez que nenhum modelo individual é ideal para se capturar a complexidade dos dados. Oliveira (2015) destaca que sistemas híbridos, que integram ARIMA e RNA, são especialmente úteis para modelar Séries Temporais Educacionais, permitindo previsões mais confiáveis em contextos reais.

3.3 Aplicações em Educação

No campo educacional, as Séries Temporais desempenham um papel estratégico na análise de índices como desempenho acadêmico, evasão escolar e repetência. Bordieri (2021) enfatiza que a previsão desses indicadores auxilia na formulação de políticas educacionais mais eficazes, contribuindo para a melhoria da qualidade do ensino. O Índice de Desenvolvimento da Educação Básica (IDEB), por exemplo, utiliza dados de Séries Temporais para medir e projetar a qualidade da educação no Brasil, permitindo que gestores identifiquem áreas prioritárias para intervenção (OLIVEIRA, 2015).

Outros estudos também exploram como modelos híbridos podem ser aplicados à educação. Oliveira (2015) utilizou uma abordagem que combina algoritmos de enxame de partículas (PSO), modelos autorregressivos e redes neurais para prever dados educacionais. Essa metodologia foi eficaz na análise de Séries Temporais complexas, como taxas de evasão e reprovação, permitindo a elaboração de estratégias específicas para minimizar esses problemas.

3.4 Desafios na Modelagem de Séries Temporais

Apesar de sua utilidade, a modelagem de Séries Temporais apresenta desafios significativos. Conforme Enders (2014), a identificação de padrões em séries não estacionárias ou com alta variabilidade pode ser uma tarefa complexa. Além disso, os métodos tradicionais frequentemente exigem ajustes, como diferenciação, para tornarem as séries adequadas à análise. Zhang (2001) e Oliveira (2015) sugerem que o uso de sistemas híbridos pode superar essas limitações, combinando a força de técnicas lineares e não lineares para se capturarem diferentes características dos dados.

Outro desafio apresentado é a sazonalidade, que exige atenção especial com modelos preditivos. Box e Jenkins (1970) destacam a importância de se considerarem componentes sazonais em séries com padrões cíclicos, como vendas anuais ou taxas de matrículas escolares. Bordieri (2021), por sua vez, observa que, em contextos educacionais, a sazonalidade pode ser influenciada por fatores externos, como calendários acadêmicos e políticas públicas.

3.5 Contribuições das Séries Temporais para a Tomada de Decisão

A análise de Séries Temporais oferece uma base sólida para decisões estratégicas em diversos setores. No caso da educação, esses modelos permitem não apenas monitorar o desempenho dos alunos, mas também antecipar mudanças e planejar intervenções eficazes. Conforme Bordieri (2021), previsões precisas podem orientar investimentos em infraestrutura, currículos e formação de professores, maximizando o impacto de políticas educacionais.

Em suma, as Séries Temporais são ferramentas indispensáveis para a análise e previsão de fenômenos dinâmicos. Seu uso adequado, combinado com modelos híbridos e avanços tecnológicos, possibilita insights valiosos que contribuem para a tomada de decisões mais informadas e eficazes, tanto na educação quanto em outros campos.

3.6 Modelo ARIMA

A análise de Séries Temporais é uma ferramenta essencial em diversas áreas do conhecimento, permitindo compreender padrões históricos e realizar previsões futuras com base em dados sequenciais. Entre os diversos modelos estatísticos utilizados para esse fim, o

modelo ARIMA, introduzido por Box e Jenkins (1976), se consolidou como um dos métodos mais robustos para a modelagem de Séries Temporais.

O modelo citado combina três componentes principais: autorregressão (AR), integração (I) e médias móveis (MA). A especificação de um modelo ARIMA é feita na forma *ARIMA* (p, d, f), onde : p representa a ordem da parte autorregressiva (AR), ou seja, quantas observações passadas influenciam o valor atual; d indica o número de diferenciações aplicadas às séries para torná-las estacionárias; q define a ordem do termo de médias móveis (MA), que capturam a influência dos erros passados no valor presente da série.

A abordagem ARIMA se baseia na suposição de que a Série Temporal pode ser descrita por uma combinação linear desses três componentes. Segundo Morettin e Tolo (2006), essa modelagem é particularmente eficaz para as Séries Temporais que apresentam padrões cíclicos ou tendências, permitindo uma adaptação flexível a diferentes tipos de dados.

3.6.1 Componente do Modelo ARIMA

Apresentam-se, a seguir, informações relativas aos três componentes do modelo ARIMA acima apresentado.

3.6.1.1. *Processo Autorregressivo*

A parte autorregressiva do modelo em epígrafe, representada pelo parâmetro p , indica que o valor presente da série pode ser expresso como uma função linear das observações passadas, acrescida de um termo de erro. Matematicamente, um modelo autorregressivo de ordem p , denotado como $AR(p)$, é expresso por:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

onde $\phi_1, \phi_2, \dots, \phi_p$ são os coeficientes autorregressivos e ε_t é um termo de erro branco, ou seja, um ruído aleatório com média zero e variância constante.

Segundo Brockwell e Davis (2016), os modelos AR são particularmente úteis quando há dependência significativa entre os valores passados da série e os futuros, como ocorre em séries econômicas e financeiras.

3.6.1.2 *Processo de Integração (I)*

A parte integrada do modelo, indicada pelo parâmetro d , refere-se ao número de diferenças necessárias para tornar a série estacionária. Uma Série Temporal é dita estacionária quando sua média e variância são constantes ao longo do tempo e não há dependência sazonal sistemática (HYNDMAN, ATHANASOPOULOS, 2018).

Quando uma série apresenta tendência ou sazonalidade, ela pode ser transformada em uma série estacionária por meio do operador de diferenciação, definido por:

$$Y'_t = Y_t - Y_{t-1}$$

Se a série ainda não for estacionária após uma primeira diferenciação, aplica-se o operador novamente até que a estacionariedade seja alcançada. Modelos ARIMA com $d=0$ são equivalentes a modelos ARMA (*AutoRegressive Moving Average*), indicados para séries que já são originalmente estacionárias (BOX; JENKINS; REINSEL, 2015).

3.6.1.3 *Processo de Médias Móveis (MA)*

O componente de médias móveis, determinada pelo parâmetro q , modela a relação entre o valor presente da série e os erros passados. Um modelo MA de ordem q é representado pela equação:

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

onde $\theta_1, \theta_2, \dots, \theta_q$ são os coeficientes da média móvel e ε_t representa um erro branco. Segundo Wei (2006), esse tipo de modelagem é útil quando os valores passados da série não são diretamente relevantes, mas sim os erros de previsão anteriores.

3.6.2 Identificação e Estimação de um Modelo ARIMA

A modelagem ARIMA segue um processo iterativo baseado na metodologia de Box-Jenkins, que envolve três etapas principais:

- Identificação do modelo: análise da série para determinar os valores adequados de p , d e q por meio de Gráficos de autocorrelação e testes estatísticos como o teste de Dickey-Fuller aumentado (ADF).
- Estimação dos parâmetros: ajuste dos coeficientes ϕ e θ utilizando métodos como Máxima Verossimilhança ou Mínimos Quadrados.
- Validação do modelo: verificação da adequação do modelo ajustado por meio da análise dos resíduos e testes de diagnóstico, como o teste de Ljung-Box.

Hyndman e Athanasopoulos (2018) destacam que um dos desafios na modelagem ARIMA é a escolha correta dos parâmetros, pois um modelo mal especificado pode levar a previsões imprecisas.

3.7 Aplicações do Modelo ARIMA

Estudos demonstram que a aplicação da metodologia Box-Jenkins pode melhorar significativamente a análise das previsões quando comparada a outros modelos, como o de Holt-Winters. Lodeti e Texeira (2023) analisaram a aplicação do ARIMA na previsão de vendas de leite, constatando que esse modelo reduziu os erros de previsão em 50% em relação ao Holt-Winters.

Por sua vez, Camelo *et al.* (2017) demonstraram que a combinação do ARIMA com Redes Neurais Artificiais (RNA) pode melhorar ainda mais a previsão de Séries Temporais, tornando o modelo híbrido uma alternativa promissora para dados complexos.

4 REDES NEURAIIS (PERCEPTRON MULTICAMADAS)

As Redes Neurais Artificiais (RNAs) são modelos computacionais inspirados no funcionamento do cérebro humano, capazes de aprender padrões complexos a partir de dados. Entre suas arquiteturas mais conhecidas, destaca-se o Perceptron Multicamadas (Multilayer Perceptron – MLP), uma rede composta por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. O aprendizado da MLP ocorre via retropropagação do erro (backpropagation), que ajusta iterativamente os pesos sinápticos com base na diferença entre a saída prevista e a real.

No campo educacional, a aplicação do MLP tem crescido significativamente, sobretudo para prever o desempenho acadêmico, identificar estudantes em risco de evasão e orientar intervenções pedagógicas mais eficazes. O estudo conduzido por Rocha *et al.* (2022), por exemplo, utilizou essa técnica para classificar alunos do curso de Licenciatura em Matemática quanto ao risco de desistência. A rede neural foi treinada com dados como horas dedicadas ao estudo semanalmente bem como as para o sono e atividade profissional, tempo dedicado à leitura, frequência às aulas e participação em atividades extracurriculares.

Após o treinamento do modelo, a rede foi capaz de atingir uma taxa de acerto de 81,9672% na identificação de estudantes com potencial de abandono, com taxa de erro de 18,0328%. No total, 50 instâncias foram classificadas corretamente e 11, de forma incorreta. Apesar do número reduzido de alunos no experimento (61), os resultados indicam a eficácia do modelo e sugerem que sua acurácia pode ser ampliada com um conjunto maior de dados. A experiência demonstrou como o MLP pode ser usado como uma ferramenta prática para antecipar comportamentos críticos no percurso formativo dos discentes, auxiliando instituições na tomada de decisão preventiva.

Complementarmente, Nascimento *et al.* (2022) aplicaram técnicas semelhantes para explorar o impacto da gestão do tempo no desempenho acadêmico. Neste caso, as RNAs foram empregadas para classificarem perfis de alunos com base em hábitos e comportamentos relacionados ao uso do tempo, utilizando a técnica SHAP (SHapley Additive Explanations) para interpretação dos modelos. Os resultados demonstraram que o desenvolvimento de um perfil comportamental eficiente favorece diretamente o desempenho acadêmico. A análise dos clusters identificados permitiu reconhecer padrões comuns entre os alunos, reforçando o potencial das redes neurais como instrumentos de apoio à personalização do ensino e ao planejamento de intervenções educacionais.

Conclui-se que a utilização de redes do tipo Perceptron Multicamadas se mostra promissora para a análise de grandes volumes de dados educacionais, oferecendo suporte tanto para a previsão de desempenho quanto para o mapeamento de fatores críticos de sucesso ou de risco acadêmico. Sua capacidade de aprendizado não linear a torna especialmente útil para modelar fenômenos complexos, como a evasão escolar ou a construção de perfis estudantis multifatoriais.

5 METODOLOGIA

Para aprofundar a compreensão dos desafios enfrentados pelos alunos no primeiro ano do Ensino Técnico Integrado ao Ensino Médio, essa pesquisa analisou, ao longo de um período de dez anos (2014-2023), dados referentes a notas, frequência escolar, gênero, raça e arranjo familiar (residência com pais ou outros). Estudos anteriores indicam que essas variáveis desempenham papéis significativos no desempenho acadêmico dos estudantes. A análise longitudinal permitirá observar como essas variáveis interagem e influenciam o desempenho acadêmico ao longo do tempo, fornecendo insights valiosos para a formulação de políticas educacionais mais equitativas e eficazes

A pesquisa apresentada seguirá a classificação apresentada por Oliveira (2011), segundo o qual um estudo pode ser classificado quanto aos objetivos e natureza da pesquisa, quanto à escolha do objeto de pesquisa e quanto à técnica de coleta de dados e de seu tratamento.

Quanto aos seus objetivos, uma pesquisa pode ser classificada como descritiva, exploratória e explicativa e quanto à sua natureza pode ser considerada como qualitativa, quantitativa ou quali quantitativa. Quanto à escolha do objeto de pesquisa, pode-se trabalhar com estudo de caso único ou múltiplos estudos de caso ou com amostragens específicas. Por fim, quanto à coleta de dados, ela pode se valer de busca bibliográfica e/ou documental, entrevistas e questionários, pesquisa de campo e pesquisa-ação. Por fim, os tratamentos oferecidos aos dados podem ser estatísticos ou por análise de conteúdo (OLIVEIRA, 2011).

5.1 Classificação da Pesquisa Quanto ao seu Objetivo

O estudo apresentado pode ser classificado quanto ao seu objetivo como descritivo e explicativo, possuindo também um aspecto exploratório. Todavia, observa-se que os dois primeiros são mais evidentes. Descrevem-se as Séries Temporais e os modelos de análise a que se recorreu para o alcance dos objetivos do estudo com detalhes que não permitem dúvidas quanto à sua importância. Da mesma forma, explicações complementam as descrições o que oferece ao leitor/pesquisador maior conhecimento sobre o assunto tratado.

5.2 Classificação da Pesquisa quanto à sua Natureza

Quanto à natureza, um estudo pode ser quantitativo, qualitativo ou quali quantitativo. Uma pesquisa é qualitativa quando compreende investigação que se

considera específica. Segundo Triviños (1987), a abordagem qualitativa busca um significado para os dados encontrados e tem como base a percepção do fenômeno dentro do seu contexto, captando não apenas a aparência do fenômeno mas também sua essência. Uma abordagem qualitativa pode, muitas vezes, requerer que haja também uma quantificação de amostragem, o que permite que não raras vezes possa haver uma proposta de quantificação de amostragem. Certo é que a abordagem qualitativa estuda mais profundamente o objeto de estudo (OLIVEIRA, 2011). Ainda segundo os autores citados, uma pesquisa é quantitativa quando propicia formulações, gráficos e tabelas, de forma que estatísticas sejam essenciais para sua compreensão.

Dentro do cenário acima, o estudo apresentado é qualiquantitativo pelo aprofundamento na temática e pela apresentação de dados que promovem Gráficos, tabelas e estatísticas variadas.

5.3 Classificação da Pesquisa Quanto ao Objeto de Estudo

A pesquisa apresentada tem como objeto de estudo uma base amostral aleatória, a saber, dados acadêmicos de estudantes do primeiro ano do Ensino Técnico Integrado ao Ensino Médio em uma Instituição Federal de Ensino Superior, ou seja, no sistema conhecido como “Ensino Médio Integrado”. Os dados se referem a um período de 10 anos (2014 a 2023) e que serão submetidos a modelos específicos conhecidos como Séries Temporais. Assim, ela pode ser classificada como pertencente à classificação “Amostragens não Probabilísticas” (MATTAR, 2001; OLIVEIRA, 2011).

5.4 Classificação da Pesquisa Quanto à Técnica de Coleta de Dados

Os dados de um estudo podem ser coletados por meio de pesquisas bibliográficas e documentais, questionários/formulários/entrevistas, observação participante ou não participante, experimentos e pesquisa-ação. As pesquisas bibliográficas e documentais são realizadas quando há busca de referências autorais do campo estudado bem como há a consulta a documentos específicos para se ter acesso a dados específicos. Questionários e outras ferramentas similares são técnicas que implicam a ter contato direto ou indireto com os sujeitos da pesquisa para obtenção de base censitária ou semelhante. As observações participante/não participante dizem respeito a dados colhidos no ambiente de pesquisa quando o pesquisador tem como propósito observar de mais perto a realidade que analisa. Os experimentos dizem

respeito a pesquisas feitas em laboratório ou espaços similares. A pesquisa-ação aponta que o pesquisador faz parte direta ou indireta do campo de amostragem. Isso permita que consiga reunir teoria e prática em sua análise.

A pesquisa em epígrafe teve seus dados coletados a partir de pesquisa bibliográfica e documental, observação participante e pesquisa-ação. Bibliograficamente, buscou-se na literatura do campo referenciais teóricos vinculados aos temas em análise e documentalmente coletaram-se os dados a partir de documentos arquivados com médias de avaliação de alunos do primeiro ano DO Ensino Técnico Integrado ao Ensino Médio na instituição educacional citada em um período de dez anos (2014-2023), mantendo-se a privacidade de todos os envolvidos. As disciplinas incluíam os campos de Português, Matemática, Física, Química, Biologia, Geografia e História. A observação participante se concretiza quando se afirma que a gênese do estudo se deu devido à observação da evolução das médias das avaliações pelo fato de atuar na instituição de ensino que ofertou os dados. Essa observação, por si só, permite que os dados sejam obtidos por pesquisa-ação.

A técnica da pesquisa-ação é estudada com maestria por Thiollent (2011) segundo o qual essa metodologia atende dois propósitos básicos, a saber, a prática e o conhecimento. A prática contribui para a pesquisa na busca de solução para o problema e o conhecimento é gerado a partir dessa solução. O fato de a pesquisadora se conectar ao ambiente do estudo a capacita por trabalhar com o problema buscando uma solução para ele. O conhecimento (ser especializada em Inteligência Artificial e conhecer os modelos de Séries Temporais) permite que ela encontre a possível solução para a questão apresentada.

5.4.1 Pesquisa Bibliográfica e Documental

A coleta de dados bibliográficos abrangeu autores analistas de aspectos relacionados ao desempenho acadêmico, Séries Temporais, Modelo ARIMA e Redes Neurais. Esse referencial serviu como base para interpretar os resultados e validar as metodologias utilizadas. Os dados necessários para a análise foram solicitados ao Departamento de Ensino, responsável pela guarda e sigilo das informações acadêmicas. A solicitação incluiu as notas e o perfil dos alunos ao longo dos 10 anos, entre 2014 a 2023, garantindo o cumprimento das normas de confidencialidade e proteção de dados. As informações foram organizadas em um banco de dados estruturado, assegurando sua integridade e prontidão para análise. Quanto à modelagem e análise dos dados, avaliaram-se o desempenho acadêmico dos alunos, as Séries Temporais, o modelo ARIMA e as Redes Neurais Perceptron.

Desempenho acadêmico é a avaliação do progresso e das conquistas de um aluno no ambiente educacional, refletindo sua capacidade de atingir metas e expectativas estabelecidas. Fatores como como motivação, qualidade de ensino e características individuais influenciam diretamente esse desempenho. Compreender esses aspectos é essencial para se desenvolverem estratégias que promovam um aprendizado mais eficaz e direcionados.

Séries Temporais são as técnicas utilizadas para se analisar a evolução das notas ao longo do tempo, identificando padrões, sazonalidades e tendências que possam auxiliar na compreensão do desempenho acadêmico.

O modelo ARIMA é aplicado como uma ferramenta principal de análise de Séries Temporais, e utilizado para prever comportamentos futuros com base nos padrões históricos identificados, proporcionando insights sobre possíveis trajetórias de desempenho.

As Redes Neurais do tipo Perceptron Multicamadas são empregadas para se modelarem relações complexas entre as variáveis do estudo, permitindo identificar padrões lineares e realizar previsões mais robustas sobre o desempenho dos alunos.

5.5 Classificação da Pesquisa Quanto ao Tratamento Oferecido aos Dados Obtidos

Os dados obtidos foram tratados por meio da Análise de Conteúdo, segundo o teórico Bardin (2010). Essa técnica enriquece a análise das comunicações e objetiva enriquecer a leitura obtida uma vez que extrai conteúdos que, sem ela, poderiam ficar implícitos nos resultados sem serem notados. Essa análise é aplicada tanto nos dados bibliográficos quanto nos documentais, que conduz à explanação do referencial teórico. Da mesma forma, as informações obtidas nos documentos da instituição oferecerem subsídios capazes de permitir a construção de gráficos e tabelas.

Os dados foram analisados com as ferramentas mencionadas, possibilitando a identificação de padrões significativos e fatores que influenciaram o desempenho acadêmico. A interpretação dos resultados foi embasada no referencial teórico, assegurando a robustez das conclusões.

6 RESULTADOS E DISCUSSÕES

Seguem os resultados obtidos e suas respectivas discussões, divididos em setores específicos para melhor análise e reflexão.

6.1 Análise Exploratória do Desempenho Acadêmico com Ênfase na Mudança do Número de Etapas Avaliativas (2014-2023)

A análise do desempenho acadêmico dos alunos do primeiro ano do Ensino Técnico Integrado ao Ensino Médio de uma escola profissionalizante da Rede Federal, ao longo do período de 2014 a 2023, teve como principal objetivo compreender os impactos decorrentes da mudança no número de etapas avaliativas – de quatro etapas (2014–2017) para três etapas (2018–2023). No entanto, para garantir uma análise abrangente e contextualizada, também foram consideradas variáveis socioeconômicas e demográficas, tais como gênero, raça/cor, com quem o aluno reside, renda familiar e frequência escolar. Esse conjunto de variáveis permitiu uma compreensão mais ampla dos fatores que influenciam o desempenho acadêmico, oferecendo subsídios relevantes para a formulação de políticas educacionais mais eficazes.

Além da análise exploratória, o estudo se aprofundou por meio da aplicação de modelos de Séries Temporais e redes neurais artificiais, com o objetivo de identificar padrões históricos, prever tendências futuras e avaliar, com maior rigor, os impactos da mudança no número de etapas avaliativas sobre o desempenho dos estudantes. Essa abordagem permitiu, ainda, realizar simulações e inferências estatísticas robustas, capazes de fornecer suporte às decisões institucionais no âmbito pedagógico e administrativo.

Quanto à contextualização e estrutura da avaliação, expõe-se que, entre os anos de 2014 e 2017, o modelo de avaliação utilizado na instituição era composto por quatro etapas (denominadas 1_ET, 2_ET, 3_ET e 4_ET). A partir de 2018, a estrutura foi modificada para contemplar apenas três etapas (1_ET, 2_ET e 3_ET). Essa mudança visou, entre outros aspectos, otimizar o tempo pedagógico, reduzir a sobrecarga dos estudantes e tornar o processo avaliativo mais eficiente.

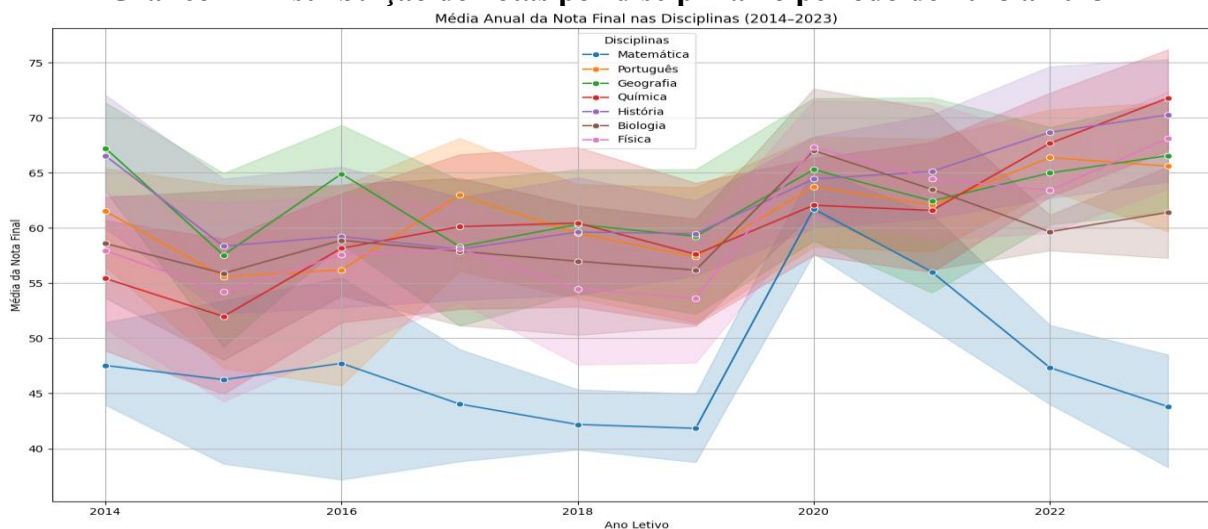
Do ponto de vista pedagógico, a mudança para três etapas se justificou por uma reestruturação curricular orientada pela necessidade de aprofundamento dos conteúdos e práticas interdisciplinares. A instituição buscou alinhar essa decisão às diretrizes da Base Nacional Comum Curricular (BNCC) e às metas do Plano Nacional de Educação (PNE), que recomendam estratégias avaliativas que privilegiem a aprendizagem contínua. **Os dados**

analisados mostram que, mesmo com a mudança de quatro para três etapas avaliativas, o desempenho médio dos estudantes manteve-se estável, confirmando a eficácia da proposta prevista no regulamento institucional, que buscava simplificar o processo avaliativo sem reduzir a diversidade dos instrumentos de aferição

Quanto à análise por disciplina e desempenho acadêmico, contemplaram-se as disciplinas Matemática, Português, Biologia, Física, Química, Geografia e História, que apresentaram comportamentos distintos ao longo do tempo. No período de 2014 a 2017, o desempenho médio dos estudantes oscilou com mais intensidade, indicando uma possível sobrecarga pela quantidade de avaliações. Por outro lado, de 2018 a 2023, observou-se uma tendência de maior estabilidade nas médias, especialmente em disciplinas como Português e História, que apresentaram menor variabilidade nas notas.

Na disciplina de Matemática identificou-se uma média geral inferior às das demais disciplinas em ambos os períodos citados, com destaque para a maior incidência de notas abaixo de 60% entre 2014 e 2017. A mudança para três etapas, a partir de 2018, não resultou em uma melhoria significativa das médias nessa disciplina, mas reduziu a dispersão dos dados, indicando maior homogeneidade nos resultados. A análise estatística mostra que o desvio padrão das notas em Matemática reduziu de 2,3 (no modelo com quatro etapas) para 1,7 (no modelo com três etapas), o que sugere que, mesmo sem aumento substancial nas médias, houve um nivelamento dos resultados entre os alunos.

Quanto à Biologia e à Química, houve uma leve melhoria nas médias no período de três etapas, especialmente após a retomada presencial pós-pandemia, o que pode indicar que a carga de avaliações mais enxuta favoreceu uma melhor assimilação dos conteúdos. Em Geografia e História, observou-se um aumento gradativo nas médias e menor percentual de reprovação, sobretudo nos anos de 2022 e 2023. Os resultados são apresentados no Gráfico 1 que, por sua vez, permitiu a construção das Tabelas 1 e 2.

Gráfico 1 Distribuição de notas por disciplina no período de 2018 a 2023

Fonte: Elaborado pela autora (2025)

Tabela 1 Estatísticas descritivas das notas no período de 2014 a 2017

Estatística	1_ET	2_ET	3_ET	4_ET	S_NOTA	N_FINAL
Count	12826	12718	12364	12061	12836	12836
Mean	11,6	11	16,3	16,3	55,4	56
Std	4,8	5,3	8,3	9,1	25,4	25,3
Min	0	0	0	0	0	0
25%	8,8	7,7	11,8	10,6	40,3	41,1
50%	12,4	12	18,1	18,5	63	63,3
75%	15,1	15	22,5	23,1	73,2	73,5
Max	20	20	30	30	100	100

Fonte: Elaborado pela autora (2025)

Tabela 2 Estatísticas descritivas das notas no período de 2018 a 2023

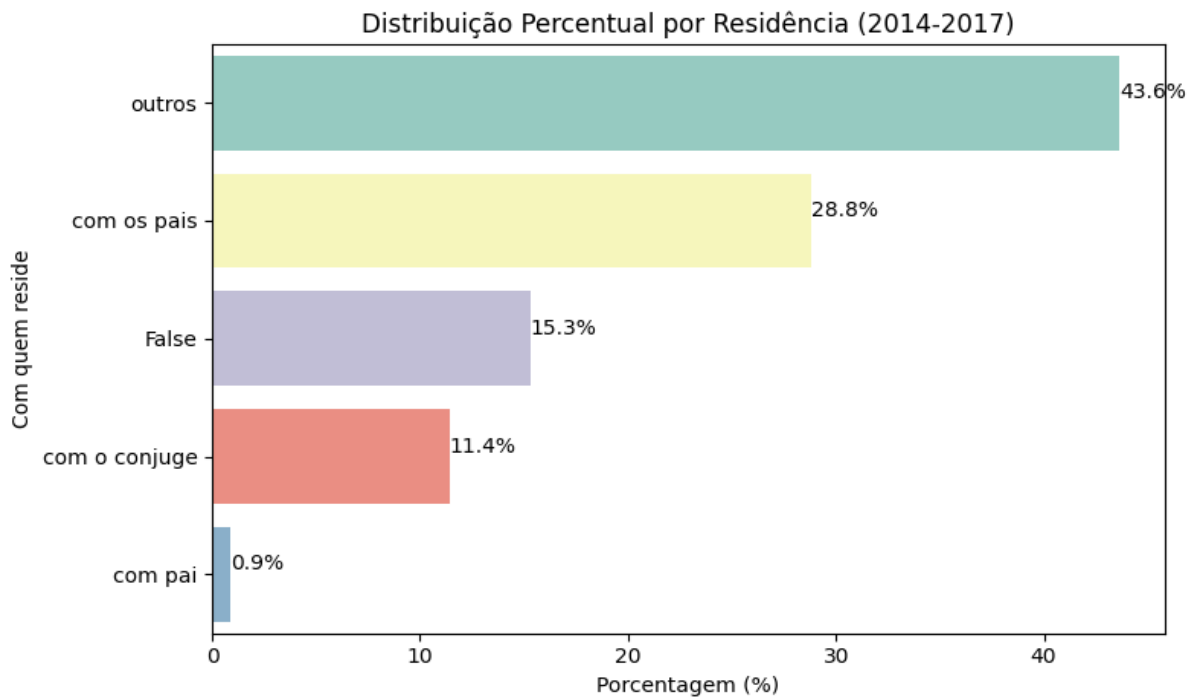
Estatística	1_ET	2_ET	3_ET	S_NOTA	N_FINAL
Count	18053	18053	18053	17674	17674
Mean	17,8	19,9	19,0	59,1	60,0
Std	8,6	11,3	12,0	27,9	27,4
Min	0	0	0	0	0
25%	13,2	12	7,5	43	48,6
50%	19,5	22,9	22,7	66,5	66,5
75%	24,2	28,8	28,3	79,1	79,1
Max	30	35	35	100	100

Fonte: Elaborado pela autora (2025)

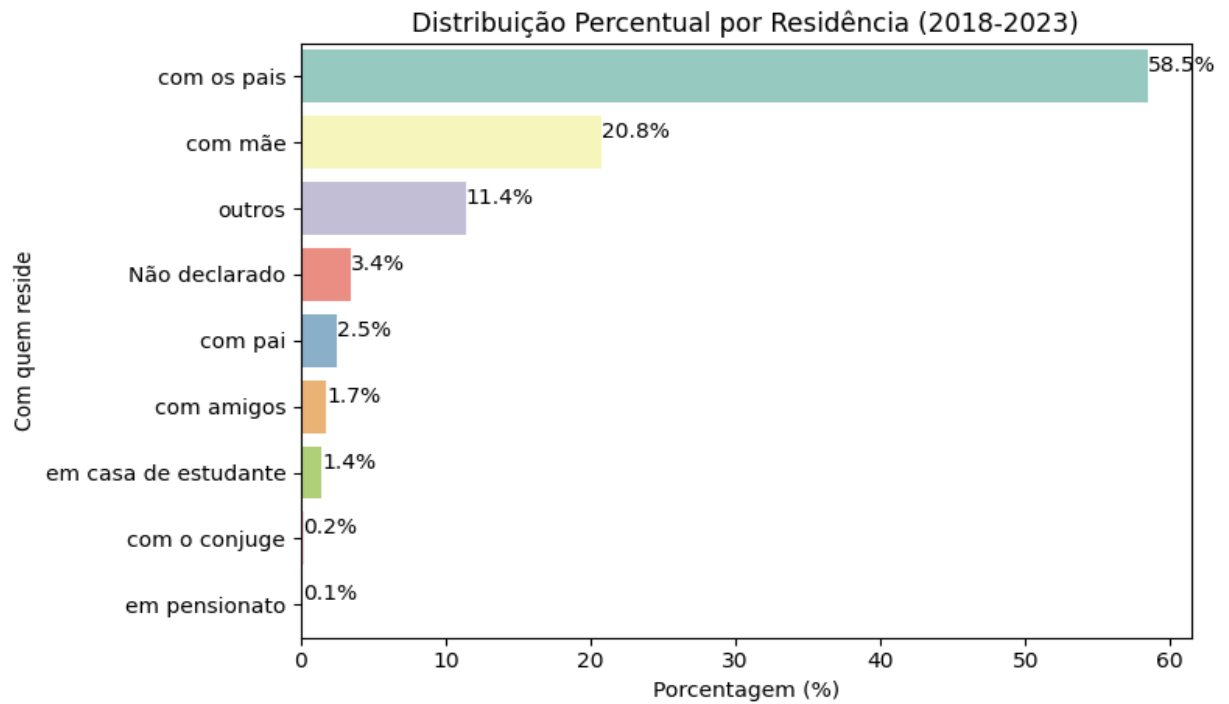
6.2 Análise das Variáveis Sociais e Demográficas

As variáveis categóricas analisadas – gênero, raça/cor de pele, residência, renda familiar e frequência escolar – também apresentaram correlação com o desempenho acadêmico. Os resultados demonstram, por exemplo, que estudantes que residem com ambos os pais têm, em média, desempenho superior aos demais. Isso sugere que o suporte familiar pode ser um fator protetivo importante no processo de aprendizagem. Esses resultados podem ser percebidos nos Gráficos 2 a 4.

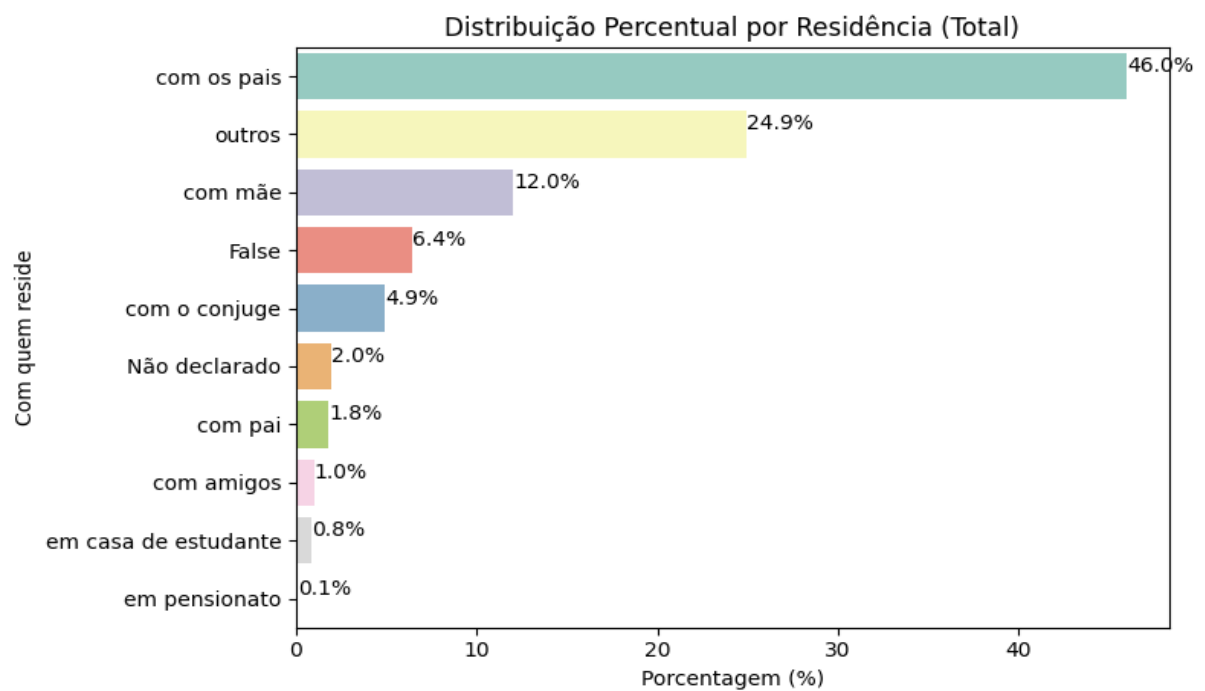
Gráfico 2 Distribuição Percentual por Residência (2014-2017)



Fonte: Elaborado pela autora (2025)

Gráfico 3 - Distribuição Percentual por Residência (2018-2023)

Fonte: elaborado pela autora (2015)

Gráfico 4 - Distribuição percentual por residência (total)

Fonte: elaborado pela autora (2025)

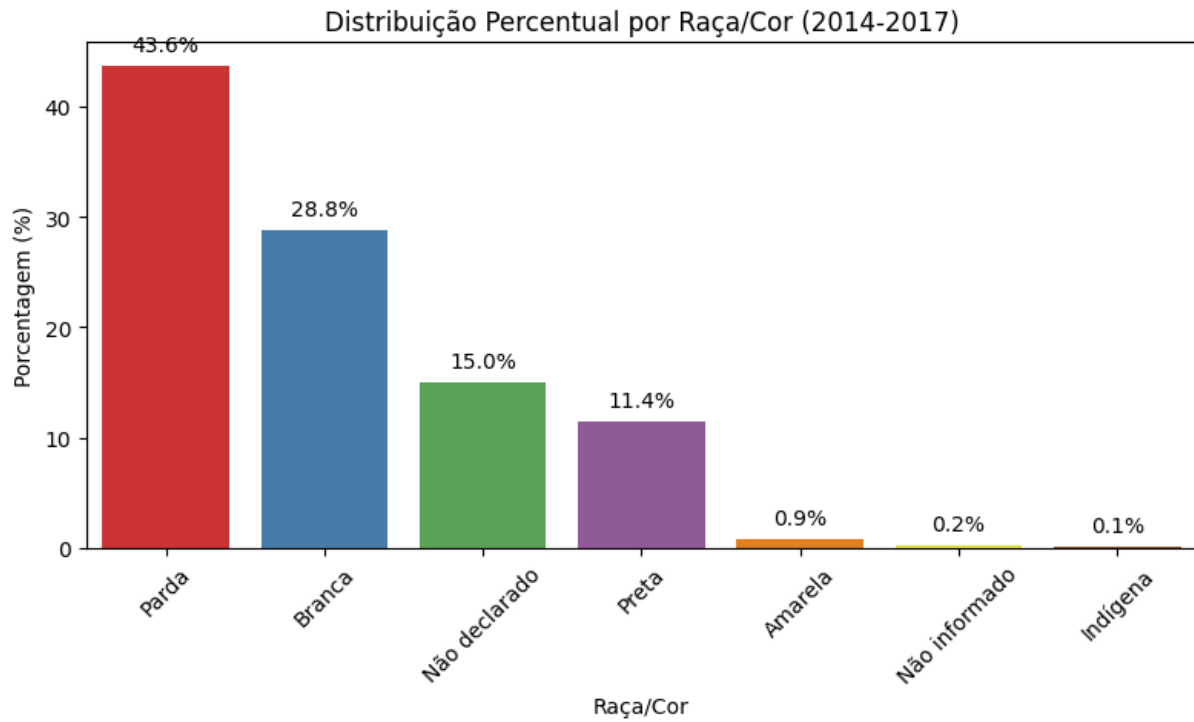
Observou-se que, no período de 2014 a 2017, os alunos que moravam com os pais estavam na faixa de 29%, mas que no período de 2018 a 2023, essa porcentagem praticamente

dobrou. Quando se analisa o índice em sua totalidade, observa-se a supremacia dessa variável. Salienta-se também a variável “com a mãe” que aparece no segundo período analisado em segundo lugar (20,8%) e, quando se analisa a totalidade entre os dois períodos, aparece em terceiro (12%). A literatura do campo expressa a importância da família para o desenvolvimento acadêmico dos filhos. Alves *et al* (2013) registram que essa influência é analisada desde o período após a Segunda Guerra Mundial e, desde então, se observa a influência positiva da família no desempenho discente. Os autores apresentam uma abrangente bibliografia que aponta que os resultados acadêmicos dos filhos são sempre melhores quando há uma interação entre pais e filhos. Rosa, Magalhães e Silveira (2024) salientam que o desempenho discente não será totalmente positivo se não houve a atenção da rede familiar. Mesmo que haja bom programa curricular, ferramentas didático-pedagógicas apropriadas e boa formação do corpo docente, o papel da família é essencial para o desempenho acadêmico dos filhos.

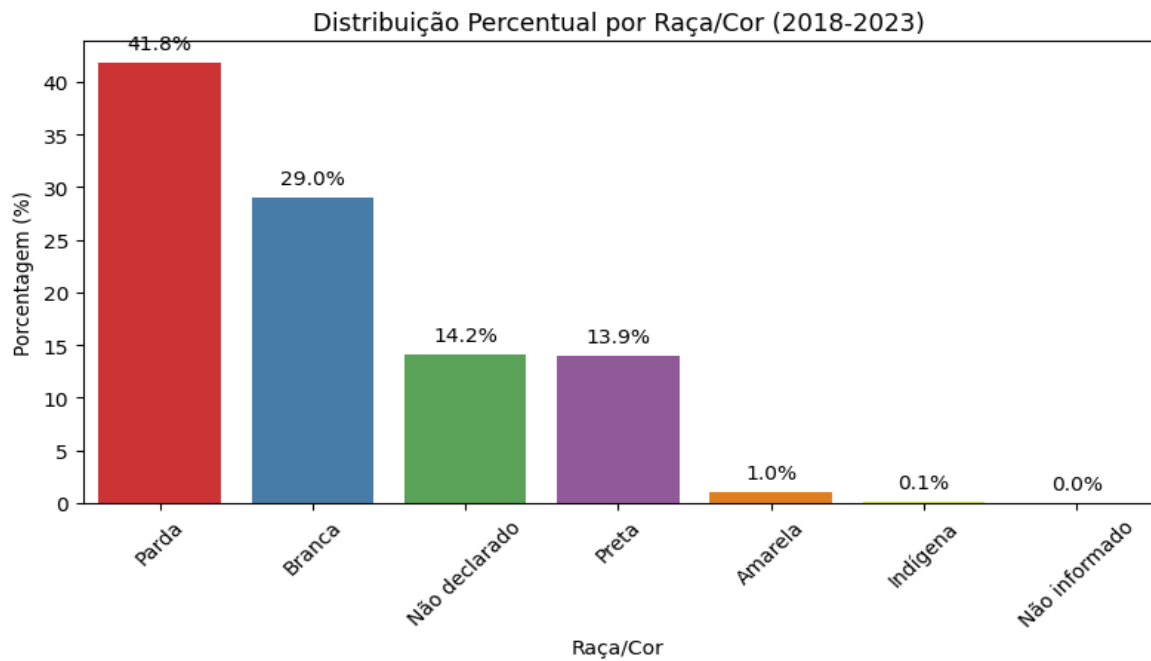
Quanto à variável raça/cor de pele, ela revelou disparidades importantes: estudantes que se autodeclararam brancos apresentaram, em média, notas superiores aos alunos autodeclarados pretos e pardos. Essa diferença se manteve nos dois períodos analisados, indicando a persistência de desigualdades estruturais que afetam o processo educacional. De 2018 em diante, notou-se um leve estreitamento dessa diferença, especialmente em disciplinas de ciências humanas, o que pode ser resultado de estratégias pedagógicas mais inclusivas e alinhadas à diversidade racial. Os resultados podem ser confirmados nos Gráficos 5 a 7.

A classificação étnico-racial dos estudantes foi obtida por meio de **autodeclaração**, seguindo a metodologia utilizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), na qual o indivíduo pode se identificar como **branco, preto, pardo, amarelo** ou **indígena** (IBGE, 2023). O termo *pardo* refere-se a pessoas com ascendência mista — como mulatos, caboclos, cafuzos ou mestiços — conforme definido oficialmente pelo IBGE (IBGE, 2023).

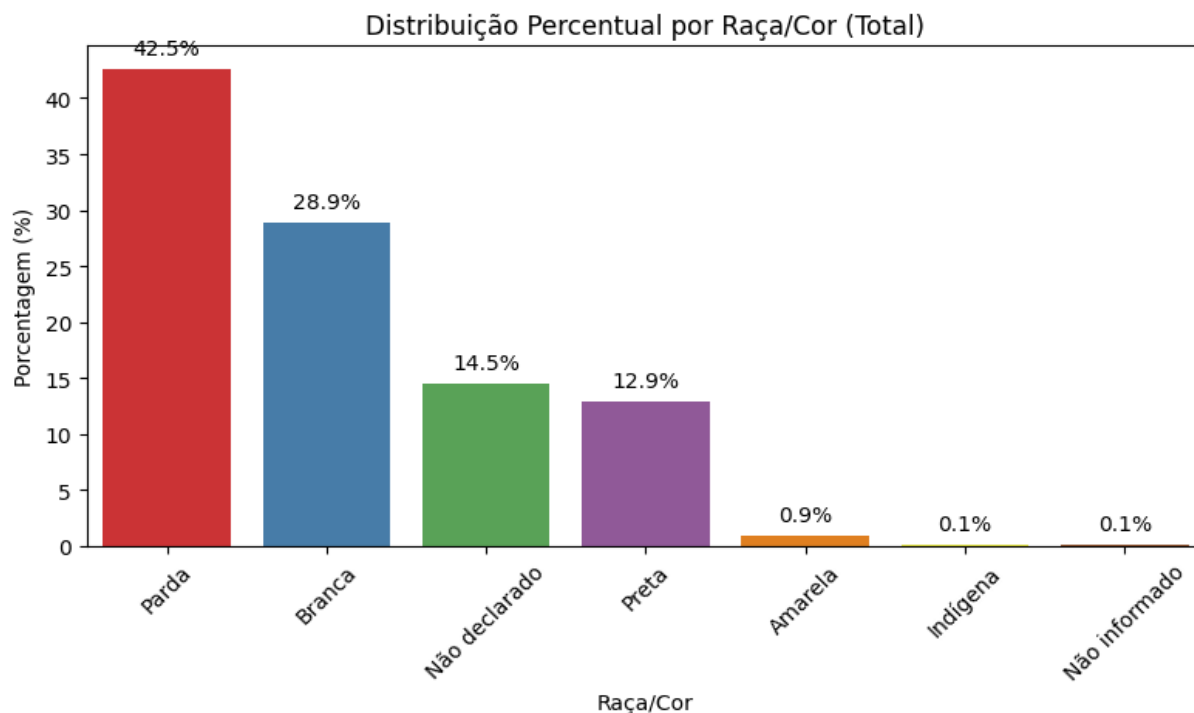
Adicionalmente, em consonância com o **Estatuto da Igualdade Racial** (Lei nº 12.288, de 20 de julho de 2010), adota-se a expressão **população negra** para designar o conjunto de pessoas que se autodeclararam **pretas e pardas** (BRASIL, 2010). Essa abordagem segue a prática das estatísticas oficiais brasileiras e possibilita comparações consistentes com pesquisas populacionais e educacionais conduzidas em nível nacional.

Gráfico 5 Distribuição Percentual por Raça/Cor (2014 – 2017)

Fonte: Elaborado pela autora (2025)

Gráfico 6 - Distribuição Percentual por Raça/Cor (2018 – 2023)

Fonte: elaborado pela autora (2025)

Gráfico 7 - Distribuição percentual por Raça/Cor (Total)

Fonte: elaborado pela autora (2025)

Os resultados encontrados estão em consonância com estudos encontrados na literatura do campo em análise. Silveira, Rosa e Tenório (2013) destacam, por exemplo, que os alunos negros apresentam menor proficiência em Matemática, seguidos pelos brancos, amarelos, pardos e indígenas. Já em Português, o mesmo ocorre, havendo apenas uma mudança de posição entre brancos e amarelos. Os autores sugerem que o desempenho escolar inferior do negro pode ser consequência do preconceito ainda existente contra os alunos negros no interior das instituições de ensino, mesmo em cenário inclusivo.

Carneiro, Shirasu e Irfê (2023) registram que, no Ensino Médio, observa-se uma diferença significativa não só no desempenho acadêmico dos alunos pretos, mas também fracasso escolar, evasão, abandono e reprovação. Segundo os autores, esse desempenho negativo se reflete quando os alunos negros chegam ao curso superior, os alunos brancos obtendo médias melhores do que os pretos.

As tabelas 3 e 4 trazem os dados quanto à média e à mediana das notas por raça/cor nos dois períodos estudados. Observe-se que a maior média pertence aos indígenas, seguida da dos brancos e dos não declarados. Os pretos apresentam a menor das médias. De qualquer forma, excetuando-se a média indígena, todas se encontram abaixo de 6,0. Quanto às medianas, no período de 2014 a 2017, novamente as dos indígenas são as maiores, seguidas pelas dos

brancos e pelas dos não declarados. Os discentes pretos apresentam as menores medianas. No segundo período analisado, as maiores médias se encontram no grupo dos não informados, seguidas pelas dos alunos brancos e pardos. A mesma situação do primeiro período de análise se repete para os alunos pretos. Os Gráficos 8 e 9 registram os dados das tabelas citadas.

Tabela 3 Média e Mediana das Notas por Raça/Cor (2014-2017)

Cor/raça	Média	Mediana
Amarela	52,7	60,5
Indígena	83,1	83,6
Não declarado	56,5	62,5
Não informado	54,4	63,9
Parda	53,9	62,6
Preta	52,3	60,1

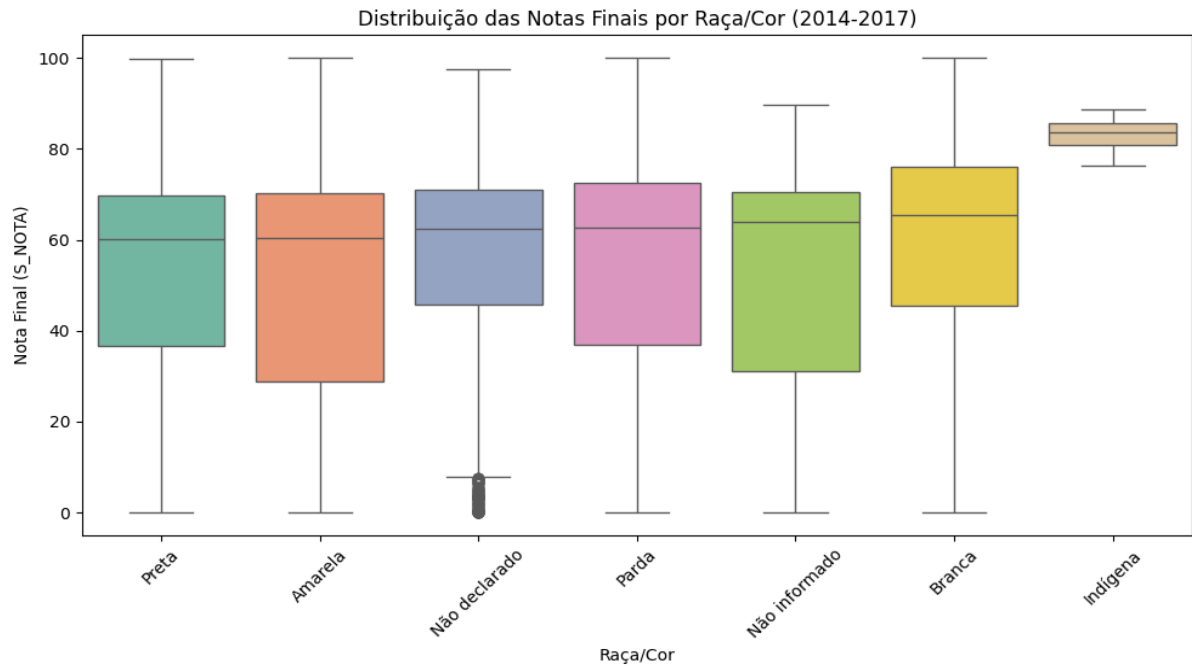
Fonte: elaborado pela autora (2025)

Tabela 4 Média e Mediana das Notas por Raça/Cor (2018-2023)

Cor/raça	Média	Mediana
Amarela	53,2	63,6
Branca	63,4	70,0
Indígena	56,7	62,0
Não declarado	55,6	64,2
Não informado	70,2	67,7
Parda	58,6	66,0
Preta	55,4	64,0

Fonte: elaborado pela autora (2025)

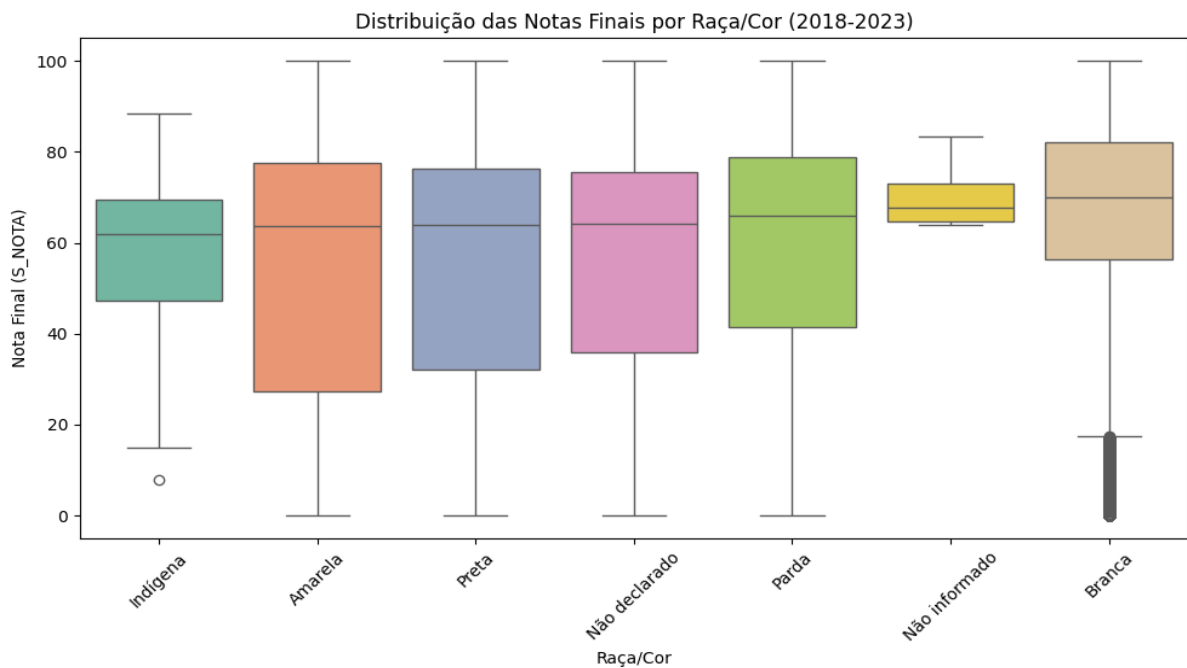
Gráfico 8 - Distribuição das Notas Finais por Raça/Cor (2014-2017)



□ O boxplot mostra a mediana (linha central), os quartis (caixa) e outliers (pontos fora da caixa).

Fonte: elaborado pela autora (2025)

Gráfico 9 Distribuição por Notas Finais por Raça/Cor (2018-2013).



□ O boxplot mostra a mediana (linha central), os quartis (caixa) e outliers (pontos fora da caixa).

Fonte: elaborado pela autora (2025)

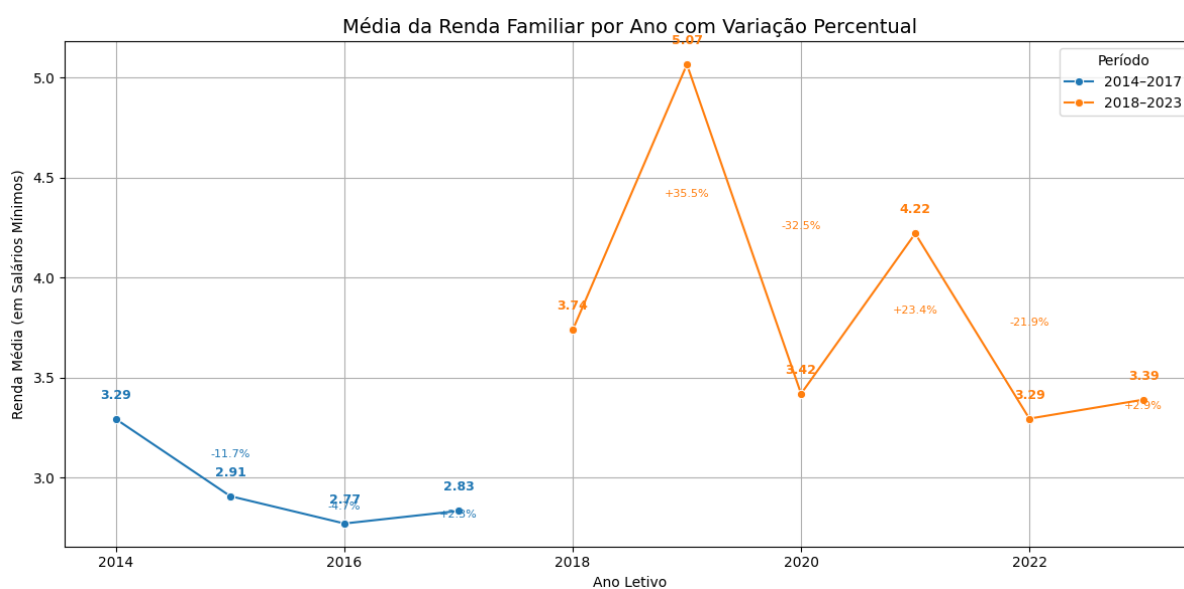
Quanto à variável renda familiar, ela demonstrou impacto no desempenho escolar. Alunos com renda familiar acima de dois salários mínimos apresentaram médias mais altas do que aqueles com renda inferior. Esse fator também interage com os demais indicadores, reforçando a importância de políticas de equidade. Vale destacar que estudantes com renda mais baixa foram mais afetados durante o período de ensino remoto.

Basseto, Ursino e Guedes (2023) apontam que o elemento socioeconômico é essencial para se avaliar o desempenho escolar dos estudantes. Isso se explica porque os recursos financeiros dos pais é que controlam as despesas educacionais dos filhos. Quanto menor é a renda familiar, menor será a capacidade da família de atender a todas as necessidades de seus filhos. Esse fator é importante porque o nível socioeconômico dos pais tende a refletir também o seu nível de escolaridade.

Peres, Campos e Moraes (2025), ao analisarem as médias alcançadas pelos candidatos matriculados no ENEM 2022, registram que a média das notas de Matemática dos candidatos cuja renda familiar era menor ou igual a um salário mínimo foi muito inferior à média das notas daqueles com renda familiar superior a nove salários mínimos, um índice de diferença em torno de 26%.

Os resultados encontrados para a variável média de renda familiar se encontram no Gráfico 10.

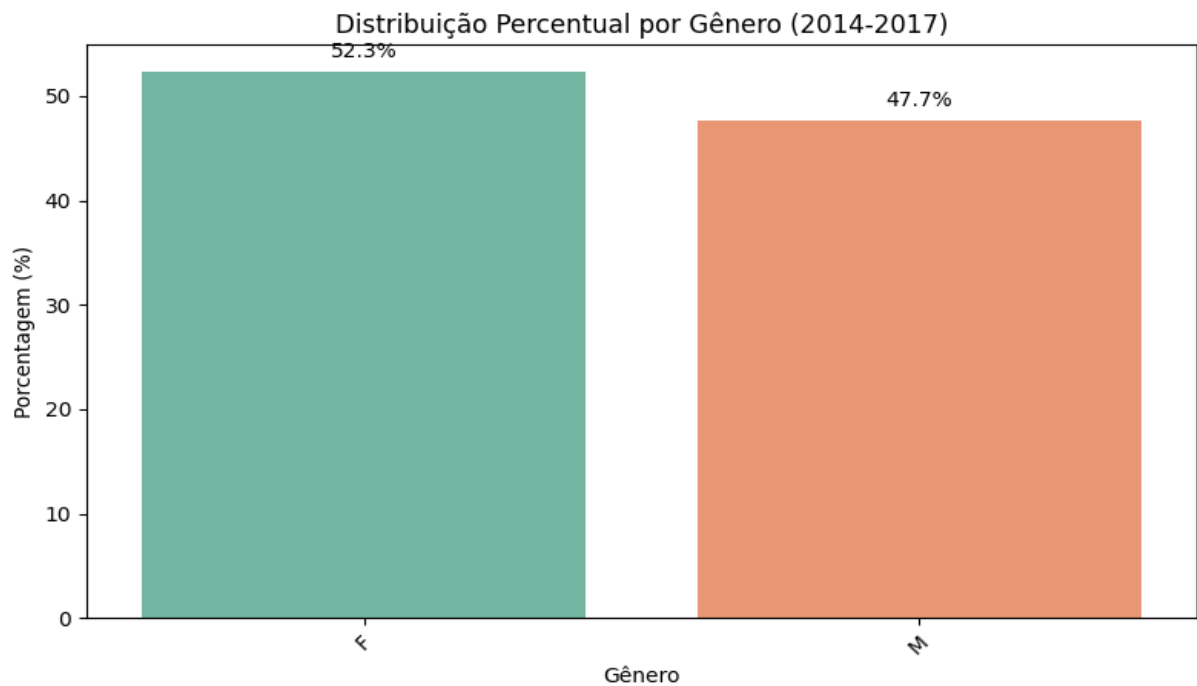
Gráfico 10 - Média da Renda Familiar por Ano com Variação Percentual no período de 2018 a 2023



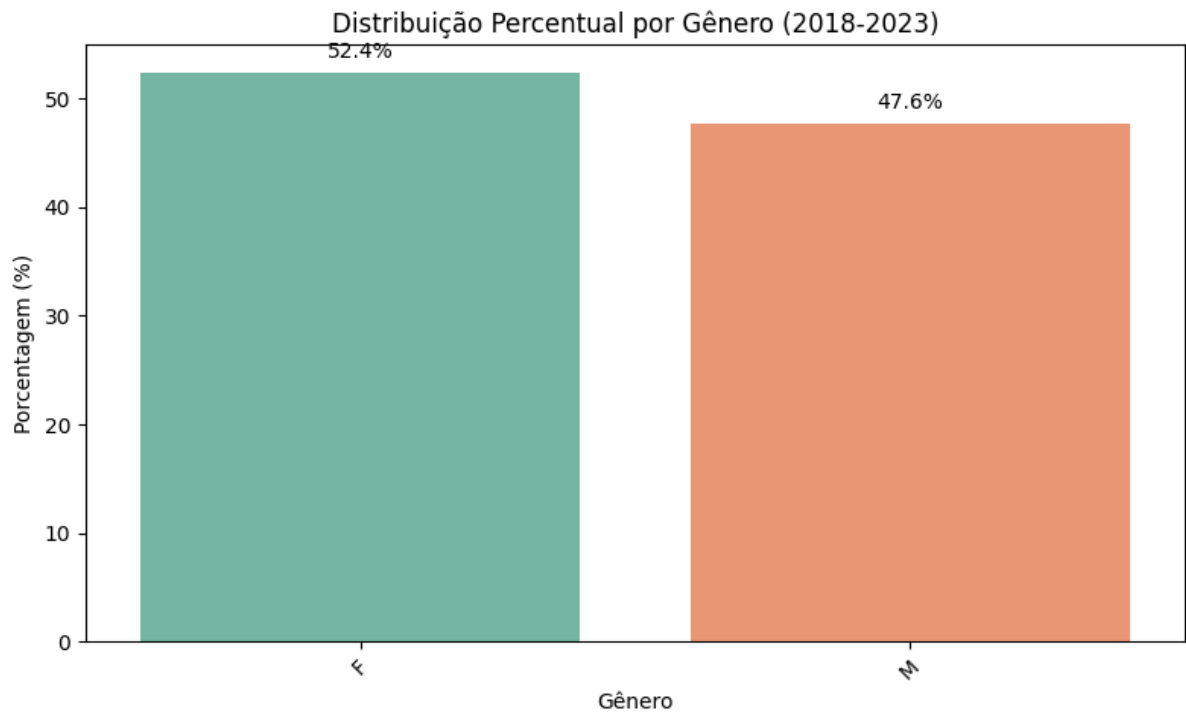
Fonte: Elaborado pela autora (2025)

Em relação à variável gênero, não se observou diferença estatisticamente significativa no desempenho médio geral, mas algumas disciplinas apresentaram variações: meninas tiveram desempenho levemente superior em Português e Biologia, enquanto os meninos obtiveram médias ligeiramente melhores em Física e Matemática. Os dados apontam que, no modelo de três etapas, essas diferenças foram suavizadas, indicando uma possível maior equidade nos critérios de avaliação. Os resultados podem ser avaliados nos Gráficos 11 a 13.

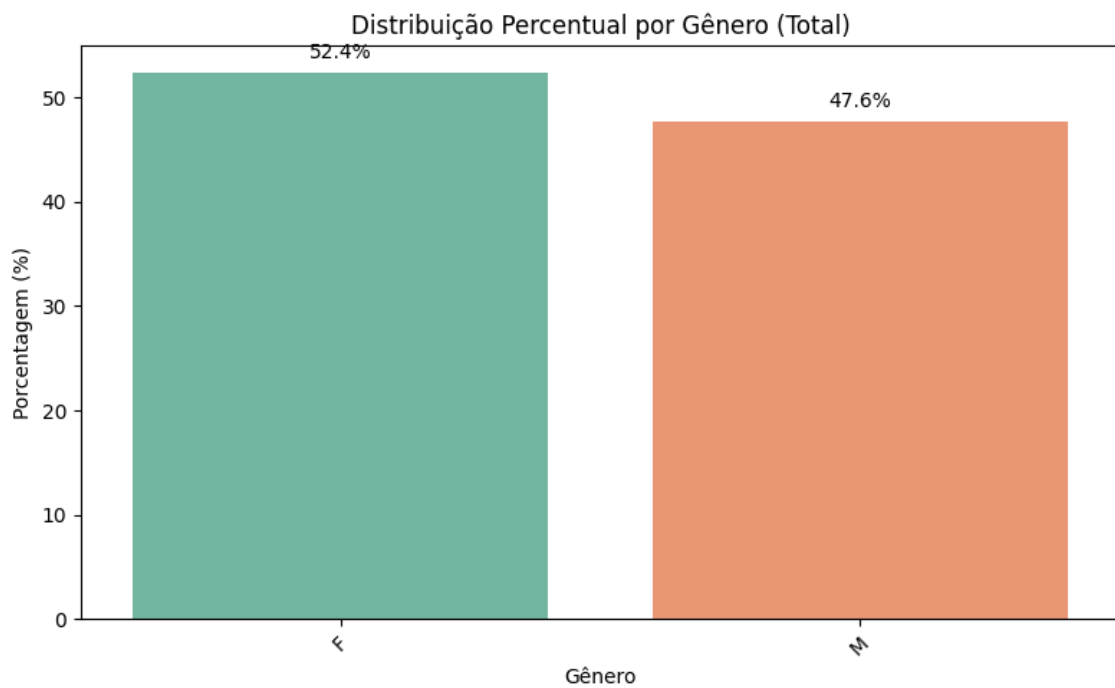
Gráfico 11 Distribuição Percentual por Gênero (2014-2017)



Fonte: Elaborado pela autora (2025)

Gráfico 12 Distribuição Percentual por Gênero (2018-2023)

Fonte: Elaborado pela autora (2025)

Gráfico 13 - Distribuição Percentual por Gênero (Total)

Fonte: Elaborado pela autora (2025)

Segundo Sanches e Guerra (2021), após exaustivo estudo estatístico, concluíram que não existe diferença estatística quanto ao desempenho escolar entre alunos e alunas. O que pode ocorrer é que a variável gênero se acople a outras variáveis e isso conduza a algumas influências nas médias obtidas por ambos.

Quanto à variável “frequência escolar”, houve uma forte correlação entre a assiduidade e o desempenho: alunos com frequência superior a 90% obtiveram médias significativamente maiores do que aqueles com frequência inferior a esse patamar. A mudança de quatro para três etapas parece ter contribuído para uma leve melhoria na frequência média dos alunos, especialmente a partir de 2019. Em 2023, por exemplo, 64% dos alunos apresentaram frequência superior a 90%, contra 52% em 2016.

Cicuto e Torres (2020) apontam que há uma correlação exata entre frequência e rendimento escolar. Alunos com alta frequência normalmente são aprovados com excelentes médias e, no caso contrário, as médias são inferiores e maiores são as reprovações. Os autores salientam a necessidade de os professores tornarem suas aulas mais atraentes de forma que os alunos se sintam incentivados a participarem mais ativamente das aulas. Também registram que a falta de planejamento das equipes pedagógicas pode levar à baixa frequência dos discentes e até mesmo para a evasão escolar. Em sua pesquisa, alunos se queixaram de que há acúmulos de trabalhos em outras disciplinas e que antes das avaliações de uma disciplina, professores pedem atividades que lhes desviam o foco. Em geral, a queixa é de sobrecarga de conteúdo e de número excessivo de aulas, sobrando pouco tempo para estudar.

As tabelas 5 e 6 abaixo registram os resultados encontrados nesse estudo para os dois períodos analisados. A tabela 5 avalia a frequência entre 2014 a 2017 enquanto que a 6 avalia a mesma variável entre 2018 e 2023.

Tabela 5 Estatísticas Descritivas da Frequência (2014-2017)

Estatística	F_ETP1	F_ETP2	F_ETP3	F_ETP4	F_TOTAL
Count	13090	13090	13090	13090	13090
Mean	1,5	0	2,2	2,2	5,8
Std	4,4	0	5,6	5,8	13,5
Min	0	0	0	0	0
25%	0	0	0	0	0
50%	0	0	0	0	0
75%	1	0	2	2	6
Max	60	0	0	62	156

Fonte: elaborado pela autora (2025)

Tabela 6 Estatísticas Descritivas da Frequência (2018-2023)

Estatística	F_ETP1	F_ETP2	F_ETP3	F_TOTAL
Count	18053	18053	8053	18053
Mean	3,3	3,7	0	6,9
Std	7,6	8,6	0	14,6
Min	0	0	0	0
25%	0	0	0	0
50%	0	0	0	2
75%	3	4	0	8
Max	82	88	0	150

Fonte: Elaborado pela autora (2025)

6.3 Impacto da Pandemia da Covid-19

O período da pandemia (2020 e 2021) impactou consideravelmente o desempenho dos alunos. Durante o ensino remoto, observou-se uma queda acentuada nas médias em todas as disciplinas, com recuperação gradual a partir de 2022. Este fator atua como uma variável interveniente relevante na análise do desempenho geral no segundo intervalo (2018–2023) e deve ser considerado nas inferências sobre o efeito da mudança nas etapas avaliativas.

A evasão escolar e a redução na frequência atingiram picos nesse período. Embora o modelo de três etapas tenha sido mantido, a implementação remota das avaliações gerou distorções nos dados, principalmente nas disciplinas exatas, que exigem maior

acompanhamento presencial. Os efeitos da pandemia foram mais severos entre os alunos com baixa renda, baixa frequência e residindo com apenas um responsável.

6.4 Inferência sobre a Mudança de Quatro para Três Etapas

A principal pergunta da pesquisa refere-se à efetividade da mudança no número de etapas avaliativas. A partir da análise estatística e descritiva dos dados, observou-se que:

- A mudança para três etapas não causou deterioração no desempenho geral dos estudantes; ao contrário, em algumas disciplinas houve leve melhoria ou maior estabilidade nos resultados;
- A homogeneização das notas, com menor desvio padrão a partir de 2018, sugere que o novo formato avaliativo contribuiu para reduzir as discrepâncias de desempenho entre os alunos;
- A frequência escolar melhorou após a mudança, o que pode indicar que um calendário menos fragmentado foi mais favorável ao engajamento dos estudantes;
- A redução no número de provas pode ter favorecido uma maior assimilação dos conteúdos e menor estresse para os discentes;
- Disciplinas com maior complexidade, como Matemática, continuam exigindo estratégias pedagógicas específicas, independentemente do número de etapas.

Os dados analisados mostram que, mesmo com a mudança de quatro para três etapas avaliativas, o desempenho médio dos estudantes manteve-se estável, confirmando a eficácia da proposta prevista no regulamento institucional, que buscava simplificar o processo avaliativo sem reduzir a diversidade dos instrumentos de aferição. Contudo, é importante reconhecer que a mudança por si só não é suficiente para garantir melhorias significativas na aprendizagem. É necessário que ela seja acompanhada de ações pedagógicas articuladas, formação continuada dos professores, adequação do currículo e suporte psicopedagógico contínuo.

6.5 Considerações e Recomendações

A análise integrada dos dados permite concluir que a mudança de quatro para três etapas avaliativas não teve impacto negativo no desempenho acadêmico dos alunos do primeiro ano do Ensino Técnico Integrado ao Ensino Médio e, em algumas dimensões, trouxe efeitos positivos. A simplificação do processo avaliativo parece ter favorecido a aprendizagem, principalmente entre os alunos com maior regularidade na frequência. Contudo, é necessário

observar que o desempenho está profundamente associado a variáveis socioeconômicas e estruturais. Políticas públicas que visem melhorar as condições de vida, a equidade racial e o apoio pedagógico contínuo são fundamentais para garantir que mudanças estruturais como essa tenham efeitos mais consistentes e duradouros.

Recomenda-se:

1. A manutenção do modelo com três etapas, aliado a estratégias pedagógicas focadas no suporte aos alunos em situação de vulnerabilidade;
2. O acompanhamento sistemático do desempenho e da frequência por meio de sistemas informatizados e relatórios periódicos;
3. O fortalecimento do apoio psicopedagógico e da atuação da família no processo de ensino-aprendizagem;
4. A realização de avaliações qualitativas que complementem os dados quantitativos, permitindo uma visão mais ampla da realidade educacional;
5. A ampliação de programas de recuperação paralela e tutorias, com foco em disciplinas com maior índice de reprovação;
6. A utilização de tecnologias educacionais e análise de dados preditivos para identificar padrões de evasão e dificuldades de aprendizagem com antecedência.

6.6 Análise de Percentuais

Seguem a análise de percentuais encontrados no estudo.

6.6.1 Análise da Taxa de Aprovação por Etapa Avaliativa nos Períodos de 2014-2017 e de 2018-2013

A presente seção apresenta a análise comparativa das taxas de aprovação dos alunos do primeiro ano do Ensino Técnico Integrado ao Ensino Médio em uma escola profissionalizante da Rede Federal, considerando os dois períodos avaliados neste estudo: de 2014 a 2017, quando o processo avaliativo contemplava quatro etapas, e de 2018 a 2023, com três etapas. O objetivo dessa análise foi verificar o impacto da alteração no número de etapas avaliativas sobre o desempenho acadêmico dos estudantes.

No período de 2014 a 2017, os dados indicam que a taxa de aprovação variou entre as quatro etapas da seguinte forma: 86,50% na primeira etapa, 81,38% na segunda, 84,48% na terceira e 80,83% na quarta etapa. A primeira etapa destacou-se como a fase com maior índice

de aprovação, alcançando 86,50%. Este resultado sugere que os alunos apresentaram um desempenho mais expressivo no início do ano letivo, possivelmente em decorrência da maior motivação ou de um planejamento curricular que favoreceu um melhor aproveitamento inicial. A taxa média geral de aprovação nesse período foi de 83,30%, indicando um desempenho global satisfatório ao longo das quatro etapas.

Por outro lado, no período de 2018 a 2023, com a redução para três etapas, observou-se a seguinte distribuição das taxas de aprovação: 86,26% na primeira etapa, 81,22% na segunda e 76,15% na terceira etapa. Assim como no período anterior, a primeira etapa apresentou a melhor taxa de aprovação, embora ligeiramente inferior à observada anteriormente (86,26%). Nota-se uma tendência de queda progressiva da taxa de aprovação nas etapas subsequentes, especialmente na terceira, que registrou o índice mais baixo entre todas as etapas dos dois períodos avaliados. A taxa média geral de aprovação no intervalo 2018–2023 foi de 81,21%, inferior à observada no período anterior. Os resultados se apresentam na Tabela 7, a seguir.

Tabela 7 Taxa de Aprovação 2014-2023

Período 2014-2017	
Etapa	Taxa de aprovação (%)
1ª etapa	86,5
2ª etapa	81,38
3ª etapa	84,48
4ª etapa	80,83
Média geral	83,3
Melhor etapa	1ª etapa
Período 2018-2023	
1ª etapa	86,26
2ª etapa	81,22
3ª etapa	76,15
Média geral	81,21
Melhor etapa	1ª etapa
Resultados	
Melhor taxa de aprovação individual	2014-2017, 1ª etapa (86,50%)
Taxa global (4 etapas – 2014-2017)	91,38%
Taxa global (3 etapas – 2018-2023)	89,81%
Qual apresentou melhor resultado?	Estratégia de 4 etapas apresentou melhor resultado geral

Fonte: Elaborado pela autora (2025)

A comparação dos dois períodos analisados evidencia que a melhor taxa de aprovação global foi registrada no período de 2014 a 2017, especialmente na primeira etapa (86,50%). A manutenção de um maior número de etapas avaliativas pode ter proporcionado uma distribuição mais equilibrada das avaliações ao longo do ano, facilitando a recuperação e o acompanhamento do desempenho dos alunos. Em contrapartida, a diminuição para três etapas a partir de 2018 pode ter concentrado a carga avaliativa, dificultando a obtenção de melhores resultados em etapas posteriores.

Os resultados sugerem que o modelo de avaliação em quatro etapas favorece o desempenho acadêmico dos estudantes, apresentando uma taxa média de aprovação superior e um comportamento mais estável nas etapas avaliativas. Recomenda-se, portanto, uma reflexão sobre a organização do processo avaliativo, considerando os impactos no aprendizado e na

aprovação dos alunos, bem como a possibilidade de se adotarem estratégias que possam amenizar os efeitos observados com a redução do número de etapas.

Em suma, a análise realizada demonstra que a estrutura de avaliação adotada entre 2014 e 2017 foi ligeiramente mais eficaz em promover o sucesso escolar dos alunos do primeiro ano do Ensino Técnico Integrado ao Ensino Médio , enquanto o modelo vigente entre 2018 e 2023 pode demandar ajustes para otimizar o rendimento acadêmico. Esses achados são relevantes para a formulação de políticas educacionais e para a melhoria contínua dos processos pedagógicos na instituição estudada.

6.6.2 Resultados do Segundo Código: Análise por Período Letivo e Modelagem Temporal

O segundo código realiza a análise da taxa de aprovação a partir do agrupamento da variável “*SITUACAO*” por “*PERIODOLETIVO*”, calculando a taxa percentual de alunos aprovados em cada ano letivo dentro dos dois intervalos estudados. O procedimento consistiu em agrupar os dados por ano letivo, contabilizar o número de alunos aprovados e dividir pelo total de alunos daquele período, obtendo-se, assim, uma taxa anual.

Os resultados obtidos evidenciam a evolução da taxa de aprovação ao longo dos anos para cada período, possibilitando a identificação de tendências e variações que não são captadas em análises estáticas. Essa abordagem permite visualizar o comportamento temporal do desempenho escolar, considerando os anos individualmente, o que é fundamental para identificar efeitos pontuais, tais como mudanças pedagógicas, administrativas ou impactos externos, como a pandemia da COVID-19.

Além da análise descritiva, foi empregada a modelagem de Séries Temporais utilizando o modelo ARIMA. Esta técnica estatística é amplamente reconhecida para a análise de dados temporais, permitindo compreender a dinâmica das taxas de aprovação ao longo do tempo e projetar tendências futuras com base nos padrões históricos.

Os Gráficos gerados ilustram claramente a evolução da taxa de aprovação observada e a taxa ajustada pelo modelo ARIMA, facilitando a visualização da aderência do modelo aos dados reais.

6.6.3 Diferenças entre os Códigos

A principal diferença entre os dois códigos reside no foco e no nível de detalhamento da análise realizada. O primeiro código calcula a taxa de aprovação por etapa avaliativa (1ª, 2ª, 3ª e 4ª etapas para o período de 2014 a 2017; 1ª, 2ª e 3ª etapas para 2018 a 2023), levando em consideração somente as notas das etapas, independentemente da situação final do aluno. Essa análise destaca quais etapas apresentaram maior sucesso na aprovação dos alunos, porém não considera a aprovação ou reprovação global do estudante no ano letivo.

Por sua vez, o segundo código calcula a taxa de aprovação global por ano letivo, utilizando a variável “*SITUACAO*”, que indica se o aluno foi aprovado ou não no final do período. Além disso, utiliza uma abordagem de Séries Temporais para captar tendências, variações e comportamentos ao longo do tempo, não se limitando a médias agregadas.

Essas diferenças metodológicas explicam a disparidade nos resultados. O primeiro código é mais granular e foca nas etapas individuais, enquanto o segundo oferece uma visão mais ampla e contextualizada da aprovação anual.

6.6.4 Qualidade e Aplicabilidade dos Resultados

Ambos os métodos apresentam alta qualidade e oferecem *insights* complementares relevantes para a análise do desempenho acadêmico.

O primeiro código é especialmente útil para gestores educacionais e docentes que desejam identificar quais etapas da avaliação apresentam maior ou menor taxa de aprovação, permitindo intervenções específicas nas etapas com índices mais elevados de reprovação. Trata-se de uma análise imediata e focada nas avaliações parciais.

Já o segundo código oferece uma visão macro, considerando o resultado final do aluno em cada ano letivo. A utilização do modelo ARIMA eleva o nível da análise, pois possibilita inferências sobre o comportamento temporal das taxas de aprovação, auxiliando na detecção de tendências, sazonalidades e efeitos externos ao longo dos anos. Essa abordagem é valiosa para o planejamento estratégico, avaliação de políticas educacionais e tomada de decisão baseada em evidências quantitativas.

A aplicabilidade prática do segundo método é mais ampla em estudos que busquem compreender a evolução da qualidade do ensino ao longo do tempo e para a construção de modelos preditivos que possam antecipar o desempenho futuro dos alunos.

Em síntese, o segundo código proporciona uma análise robusta, completa e temporalmente contextualizada da aprovação escolar, enquanto o primeiro permite um olhar detalhado sobre o desempenho nas etapas individuais de avaliação. Recomenda-se a utilização conjunta das duas metodologias para obtenção de uma compreensão abrangente e aprofundada do desempenho acadêmico e da eficácia das avaliações aplicadas.

6.7 Correlação Pearson e Correlação Spearman para Análise das Correlações entre Desempenho Acadêmico e Variáveis Socio econômicas (2014-2023)

Para compreender a relação entre o desempenho dos alunos nas diferentes etapas avaliativas ao longo dos períodos estudados, realizou-se uma análise de correlação utilizando os coeficientes de Pearson e Spearman. Esses métodos estatísticos permitem identificar o grau de associação entre as notas obtidas nas etapas, proporcionando uma melhor compreensão da consistência do desempenho dos estudantes durante o ano letivo. Além disso, incluíram-se na análise algumas variáveis socioeconômicas, como renda, sexo, situação de residência e situação final no curso, a fim de investigar sua possível influência no desempenho acadêmico.

O coeficiente de correlação de Pearson é uma medida que quantifica a relação linear entre duas variáveis. Valores próximos de +1 indicam uma associação positiva forte, o que sugere que aumentos nas notas de uma etapa acompanham aumentos nas notas das demais. No período de 2014 a 2017, quando o processo avaliativo era composto por quatro etapas, os coeficientes de Pearson entre as etapas variaram entre 0,637 e 0,779, indicando uma associação linear significativa. Já no período de 2018 a 2023, com apenas três etapas avaliativas, os valores de Pearson ficaram entre 0,711 e 0,815, demonstrando uma correlação ainda mais forte entre os desempenhos das etapas.

Além disso, calculou-se o coeficiente de correlação de Spearman, que avalia a associação monotônica entre variáveis, ou seja, a tendência de manutenção da ordem dos valores, mesmo quando a relação não é estritamente linear. Para o período de 2014 a 2017, os coeficientes de Spearman variaram entre 0,572 e 0,700, enquanto que para o período de 2018 a 2023, os valores oscilaram entre 0,651 e 0,742. Esses resultados reforçam a existência de uma relação estável na classificação relativa dos alunos, demonstrando que estudantes com bom desempenho em uma etapa tendem a manter um desempenho similar nas etapas subsequentes.

Os resultados acima evidenciam que o desempenho dos alunos ao longo do ano letivo é consistente, independentemente do número de etapas avaliativas. A estabilidade observada sugere que as avaliações sequenciais refletem adequadamente o progresso dos

estudantes, sendo possível inclusive prever o desempenho futuro com base nas notas obtidas nas primeiras etapas. Essa constatação tem implicações importantes para a gestão educacional. Com base nos padrões identificados, é possível implementar ações pedagógicas preventivas e personalizadas, voltadas principalmente aos alunos que apresentam baixo rendimento nas etapas iniciais. O acompanhamento contínuo do desempenho, aliado à análise das correlações entre as etapas, pode contribuir significativamente para a melhoria dos resultados acadêmicos e para o aprimoramento das estratégias avaliativas adotadas pela instituição.

Além das correlações entre as etapas, também foram analisadas as relações entre o desempenho acadêmico e algumas variáveis socioeconômicas e demográficas:

- A variável SITUAÇÃO, que indica se o aluno foi aprovado ou reprovado, apresentou forte correlação com as etapas, especialmente com as etapas finais de cada ano. No período de 2014 a 2017, o coeficiente de correlação de Pearson entre a quarta etapa e a situação final foi de aproximadamente 0,73, evidenciando a relação entre desempenho e aprovação. No período de 2018 a 2023, essa correlação também permaneceu forte, com a terceira etapa apresentando coeficiente próximo de 0,68.
- A renda familiar apresentou correlação fraca com o desempenho nas etapas, tanto para Pearson (valores entre 0,07 e 0,09) quanto para Spearman (entre 0,21 e 0,23). Isso sugere que, embora as condições econômicas possam influenciar a vida dos estudantes, esse impacto não se refletiu de forma marcante no rendimento acadêmico dentro da instituição analisada.
- A variável sexo demonstrou uma correlação baixa com as notas, variando de 0,11 a 0,17, o que indica diferenças pouco expressivas entre os gêneros quanto ao desempenho escolar.
- A variável VOCERESIDE, que indica com quem o aluno reside (pais, responsáveis, outros ou sozinho), apresentou correlação muito baixa ou próxima de zero no período de 2014 a 2017 e levemente negativa no período de 2018 a 2023 (com destaque para a 2ª etapa, com -0,09). Essa variação pode estar associada a mudanças sociais e pedagógicas vivenciadas durante e após o período pandêmico, refletindo possíveis dificuldades de organização e apoio familiar no contexto de ensino remoto e pós-pandemia.

6.7.1 Análise das Correlações entre Variáveis (2014-2017 e 2018-2023)

Seguem-se as análises das correlações entre variáveis nos dois períodos estudados bem como suas interpretações. Primeiramente, trabalha-se com a correlação de Pearson (2014-2017) sob o título de Tabela 8.

Tabela 8 Correlação de Pearson (2014 – 2023)

Variáveis	1_ET	2_ET	3_ET	4_ET	RendaFam	SEXO	SITUACAO	VOCERESIDE
1_ET	1	0,687	0,613	0,56	0,078	0,107	0,55	0,034
2_ET	0,687	1	0,675	0,632	0,088	0,135	0,601	0,03
3_ET	0,613	0,675	1	0,709	0,075	0,172	0,66	0,01
4_ET	0,56	0,632	0,709	1	0,076	0,172	0,732	-0,002
RendaFam	0,078	0,088	0,075	0,076	1	-0,043	0,066	0,015
SEXO	0,107	0,135	0,172	0,172	-0,043	1	0,13	0,024
SITUACAO	0,55	0,601	0	0,732	0,066	0,13	1	0,023
VOCERESIDE	0,034	0,03	0,010	-0,002	0,015	0,024	0,023	1

Fonte: Elaborado pela autora (2025)

Interpretação:

- As correlações entre as etapas de avaliação (1_ET, 2_ET, 3_ET, 4_ET) são moderadas a fortes, entre 0,56 e 0,71, indicando que as notas nas diferentes etapas estão positivamente relacionadas;
- A variável "SITUACAO" (provavelmente situação do aluno, como aprovado ou não) tem correlação relativamente forte com as etapas, especialmente com a 4_ET (0,732);
- A renda familiar (RendaFam) tem correlação fraca com as notas (por volta de 0,07 a 0,08);
- A variável SEXO apresenta correlação positiva fraca com as notas, variando de 0,107 a 0,172;

- VOCERESIDE (com quem o aluno reside) tem correlações muito baixas, quase nulas.

Segue a Correlação de Spearman para o período de 2014 a 2017, sob o título de Tabela 9.

Tabela 9 Correlação de Spearman (2014 a 2017)

Variáveis	1_ET	2_ET	3_ET	4_ET	RendaFam	SEXO	SITUACAO	VOCERESIDE
1_ET	1	0,65	0,578	0,502	0,226	0,102	0,519	0,038
2_ET	0,65	1	0,639	0,572	0,226	0,131	0,567	0,037
3_ET	0,578	0,639	1	0,62	0,227	0,166	0,612	0,02
4_ET	0,502	0,572	0,62	1	0,212	0,164	0,655	0,001
RendaFam	0,226	0,226	0,227	0,212	1	-0,054	0,25	-0,037
SEXO	0,102	0,131	0,166	0,164	-0,054	1	0,13	0,024
SITUACAO	0,519	0,567	0,612	0,655	0,25	0,13	1	0,023
VOCERESIDE	0,038	0,037	0,02	0,001	-0,037	0,024	0,023	1

Fonte: Elaborado pela autora (2025)

Interpretação:

- As correlações de Spearman indicam uma relação monotônica entre as variáveis, e geralmente são um pouco menores que as de Pearson;
- Destaca-se que a correlação de renda familiar é maior (em torno de 0,22 a 0,25), indicando que a renda tem alguma influência monotônica mais perceptível nas notas;
- As demais interpretações são similares às de Pearson, confirmando que a renda e situação do aluno influenciam, mas com intensidade moderada, enquanto residência praticamente não influencia.

Segue a Correlação de Spearman (2018 a 2023), sob o título de Tabela 10.

Tabela 10 Correlação de Spearman (2018 – 2023)

Variáveis	1_ET	2_ET	3_ET	RendaFam	SEXO	SITUACAO	VOCERESIDE
1_ET	1	0,651	0,546	0,008	0,111	0,492	-0,079
2_ET	0,651	1	0,666	0,02	0,121	0,591	-0,092
3_ET	0,546	0,666	1	0,014	0,13	0,679	-0,087
RendaFam	0,008	0,02	0,014	1	-0,037	0,021	0,003
SEXO	0,111	0,121	0,13	-0,037	1	0,077	-0,025
SITUACAO	0,492	0,591	0,679	0,021	0,077	1	-0,11
VOCERESIDE	-0,079	-0,092	-0,087	0,003	-0,025	-0,11	1

Fonte: Elaborado pela autora (2025)

Interpretação:

A correlação monotônica mostra padrão semelhante à Pearson, porém um pouco mais forte para renda familiar, que apresenta correlação positiva moderada (em torno de 0,12 a 0,15) com as notas.

- A variável situação mantém boa correlação positiva com as etapas;
- VOCERESIDE continua apresentando correlações negativas baixas ou nulas.

Em resumo tem-se:

- Relação entre etapas (1_ET, 2_ET, 3_ET, 4_ET): consistente e forte em ambos os períodos e métodos de correlação, indicando coerência entre as notas ao longo das avaliações.
- SITUACAO (situação do aluno): mostra forte correlação positiva com as notas, o que é esperado (provavelmente indica aprovação/reprovação);
- Renda familiar: apresenta correlação muito fraca em Pearson, mas um pouco mais forte em Spearman, sugerindo que a relação não é linear, mas existe alguma monotonicidade entre renda e notas;
- Sexo: correlações fracas, positivas, pouco impacto;
- VOCERESIDE: correlações muito baixas, indicando que o local de residência do aluno tem pouca influência no desempenho medido pelas notas.

Em suma, os dados analisados indicam que o principal fator associado à aprovação dos alunos é o próprio desempenho nas etapas avaliativas. A consistência observada entre as etapas valida a estrutura atual do modelo avaliativo da escola, demonstrando que tanto o modelo

anterior (com quatro etapas) quanto o atual (com três etapas) permitem monitorar com eficácia o progresso dos estudantes ao longo do ano. A análise de correlação, portanto, reforça o papel das etapas como instrumentos diagnósticos valiosos, que podem e devem ser utilizados para orientar intervenções pedagógicas precoces e eficazes.

6.8 Redes Neurais

Seguem-se análises dos dados quanto às redes neurais.

6.8.1 Análise Comparativa dos Modelos: Redes Neurais, Random Forest e Regressão Logística

A modelagem do desempenho acadêmico dos alunos requer o uso de técnicas de aprendizado de máquina que possam lidar com a complexidade e o desequilíbrio dos dados, bem como fornecer previsões precisas e confiáveis para auxiliar na tomada de decisões educacionais. Este estudo avaliou três métodos distintos — Redes Neurais, Random Forest e Regressão Logística — aplicados aos dois períodos avaliados (2014–2017 e 2018–2023), com o objetivo de comparar seu desempenho preditivo, confiabilidade e aplicabilidade prática.

6.8.1.1 Resultados de Acurácia

Os resultados para a acurácia se encontram nas Tabelas 11 e 12.

Tabela 11 Resultados de Acurácia (2014 – 2023)

Modelo	Período	Acurácia
Redes Neurais	2014–2017	85,85%
Random Forest	2014–2017	85,85%
Regressão Logística	2014–2017	85,94%

Fonte: Elaborado pela autora (2025)

Tabela 12 Resultados de Acurácia (2018 – 2023)

Modelo	Período	Acurácia
Redes Neurais	2018–2023	82,05%
Random Forest	2018–2023	82,05%
Regressão Logística	2018–2023	82,09%

Fonte: Elaborado pela autora (2025)

Observa-se que a acurácia dos três modelos é bastante próxima, a Regressão Logística apresentando um leve desempenho superior em ambos os períodos. No entanto, essa diferença é pequena e pode estar dentro da margem de erro estatístico.

6.8.1.2 *Confiabilidade e Robustez*

A partir dos dados obtidos, infere-se que:

- **Redes Neurais:** apresentam boa capacidade de capturar padrões não lineares e relações complexas nos dados, o que pode aumentar a confiabilidade em cenários com múltiplas variáveis interagindo. Contudo, são modelos mais complexos e difíceis de interpretar, o que pode dificultar a comunicação dos resultados para gestores educacionais.
- **Random Forest:** é um modelo robusto que combina múltiplas árvores de decisão para reduzir o *overfitting* e melhorar a generalização. Possui boa capacidade para lidar com dados desbalanceados e variáveis categóricas, além de fornecer indicadores de importância das variáveis, o que auxilia na interpretação.
- **Regressão Logística:** apesar de ser um modelo linear e relativamente simples, a regressão logística é altamente interpretável, facilitando a compreensão dos impactos das variáveis no desempenho acadêmico. Entretanto, tem limitações em modelar relações não lineares complexas e, conforme os resultados, apresenta dificuldades em classificar corretamente classes minoritárias.

6.8.1.3 *Aplicabilidade Prática*

- **Redes Neurais:** indicadas quando a base de dados é grande e complexa, e há interesse em maximizar a capacidade preditiva, mesmo que isso implique menor

interpretabilidade: ideal para projetos de longo prazo com equipe técnica capacitada.

- Random Forest: equilibra boa acurácia com interpretabilidade razoável, sendo apropriado para aplicações práticas que exigem explicações das decisões do modelo, como em instituições educacionais que buscam justificar intervenções pedagógicas.
- Regressão Logística: adequada para contextos onde a transparência e a simplicidade são prioritárias, permitindo que gestores entendam e validem facilmente o modelo, apesar das limitações no desempenho para classes minoritárias.

6.9 Considerações Importantes

Embora a Regressão Logística tenha apresentado a melhor acurácia nominal, essa vantagem é marginal e deve ser considerada à luz das demais métricas e das necessidades específicas da aplicação. O Random Forest surge como uma alternativa sólida, conciliando desempenho e interpretabilidade, e pode ser preferido em cenários onde a compreensão das variáveis é tão importante quanto a acurácia. As Redes Neurais, por sua vez, podem ser recomendadas para contextos em que a maximização da capacidade preditiva é essencial e há infraestrutura para lidar com sua complexidade.

A partir do contexto acima, a escolha do modelo mais confiável e aplicável deve levar em conta não apenas a acurácia, mas também o perfil dos dados, a necessidade de interpretabilidade e os recursos disponíveis para manutenção e atualização do sistema preditivo.

7 CONSIDERAÇÕES FINAIS

O presente estudo teve como objetivo analisar o desempenho acadêmico dos alunos do primeiro ano do Ensino Técnico Integrado ao Ensino Médio em uma escola profissionalizante da Rede Federal, com ênfase na mudança do modelo avaliativo, que passou de quatro etapas no período de 2014 a 2017 para três etapas entre 2018 e 2023. Para tanto, foram avaliadas as notas das disciplinas do ciclo básico, considerando variáveis socioeconômicas e demográficas, além da frequência escolar, a fim de compreender os efeitos dessa alteração no processo de avaliação e seu impacto sobre o rendimento dos estudantes.

Os resultados evidenciam que a redução no número de etapas avaliativas não comprometeu a qualidade da avaliação. Ao contrário, houve uma leve melhora no desempenho médio dos alunos no período mais recente. Tal constatação sugere que a reorganização do calendário avaliativo pode ter contribuído para um acompanhamento mais consistente e menos fragmentado da aprendizagem, beneficiando a concentração dos esforços dos alunos e dos docentes. Além disso, a análise estatística demonstrou que a frequência escolar continua sendo um fator determinante para o sucesso acadêmico, reforçando a importância do controle e estímulo da assiduidade.

Ademais, o estudo apontou a relevância de fatores socioeconômicos, como renda familiar e condições de moradia, na variação do desempenho escolar, evidenciando desigualdades que impactam diretamente o aproveitamento dos estudantes. Tais resultados corroboram a literatura acadêmica que destaca a influência do contexto socioeconômico no rendimento escolar, ressaltando a necessidade de políticas educacionais que promovam a equidade.

Por fim, a aplicação de modelos preditivos baseados em inteligência artificial, como redes neurais e Séries Temporais, demonstrou-se eficaz para a previsão do desempenho acadêmico, possibilitando uma análise prospectiva capaz de orientar intervenções pedagógicas mais precisas e personalizadas.

Dessa forma, conclui-se que a mudança do modelo avaliativo para três etapas foi positiva para o desempenho dos alunos, sem prejuízo na avaliação do aprendizado, e que a incorporação de técnicas avançadas de análise de dados pode ampliar significativamente o suporte à gestão educacional, o que responde a questão-problema que foi a gênese desse estudo e alcança o objetivo geral proposto na sua introdução. Recomenda-se que futuras pesquisas aprofundem o estudo do impacto da pandemia e explorem estratégias de mitigação das desigualdades socioeconômicas, fortalecendo o processo de ensino-aprendizagem e

contribuindo para a melhoria contínua da qualidade da educação profissionalizante na rede pública federal.

REFERÊNCIAS BIBLIOGRÁFICAS

ALVES, Maria T. G. *et al.* Fatores Familiares e Desempenho Escolar: Uma Abordagem Multidimensional. **DADOS – Revista de Ciências Sociais**, Rio de Janeiro, vol. 56, no 3, 2013, pp. 571 a 603.

ARAÚJO, Juliana M. de *et al.* Fatores escolares como determinantes do desempenho dos alunos da educação básica. *Linhas Críticas*, vol. 27, e37190, Enero-Diciembre/2021.

BARDIN, Lawrence de. **Análise de Conteúdo**. Lisboa: Edições 70, 2010.

BASSETO, Camila F.; URSINO, Driely T.; GUEDES, Álvaro M. Escolaridade materna e renda familiar : Impactos sobre o nível de proficiência em matemática medidos a partir do SARESP. **Revista Ibero-Americana de Estudos de Educação**. V. 18, 2023.

BORDIERI, E. D. Prevendo o Desempenho no ENADE: uma aplicação de algoritmos de aprendizagem de máquina. **Diversità – Revista Multidisciplinar do Centro Universitário Cidade Verde**, v. 7, n. 1, 2021.

BOX, G. E. P.; JENKINS, G. M. **Time Series Analysis: Forecasting and Control**. San Francisco: Holden-Day, 1970.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis: Forecasting and Control**. 5th ed. Hoboken, NJ: Wiley, 2015.

BRASIL. IBGE. **Cor ou raça**. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 2023. Disponível em: <https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18319-cor-ou-raca.html>. Acesso em: 13 ago. 2025.

BRASIL. PRESIDÊNCIA DA REPÚBLICA. **Lei nº 12.288, de 20 de julho de 2010**. *Institui o Estatuto da Igualdade Racial*. Diário Oficial da União, Brasília, DF, 21 jul. 2010.

Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/lei/112288.htm. Acesso em: 13 ago. 2025.

BROCKWELL, P. J.; DAVIS, R. A. **Introduction to Time Series and Forecasting**. 3rd ed. New York: Springer, 2016.

CAMELO, H. N. *et al.* Métodos de previsão de Séries Temporais e modelagem híbrida ambos aplicados em médias mensais de velocidade do vento para regiões do Nordeste do Brasil. **Revista Brasileira de Meteorologia**, v. 32, n. 4, p. 565-574, 2017.

CARNEIRO, Diego; SHIRASU, Maitê; IRFFI, Guilherme. Identificando a discriminação racial pelo diferencial de desempenho dos estudantes do Ensino Médio. **Revista de Economia Política**, vol. 43, nº 2, abril-junho/2023.

CICUTO, Camila A. T.; TORRES, Bayardo B. Influência da frequência e participação no desempenho em um ambiente de aprendizagem centrado no aluno. **Educação • Quím. Nova**, v. 43, n. 2, Fevereiro 2020.

COSTA, Roberta; BRITTO, Ariana; WALTENBERG, Fábio. Efeitos da formação docente sobre resultados escolares do Ensino Médio. **Revista Estudos Econômicos** (São Paulo), v. 50, n. 3, 2020, pp. 369-409.

ENDERS, W. **Applied Econometric Time Series: an Introduction**. 4. ed. Hoboken: Wiley, 2014.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: Principles and Practice**. 2nd ed. Melbourne: OTexts, 2018.

LODETTI, L.; TEIXEIRA, L. L. Aplicação da metodologia Box & Jenkins na previsão das vendas de leite de um laticínio em Guaira-PR. **Revista Eletrônica Científica Inovação e Tecnologia**, 2023.

MACHADO, Flávia F. da . *et al.* Fatores explicativos do desempenho escolar entre estudantes brasileiros. **Revista Observatorio de la Economía Latinoamericana**. Curitiba, v.21, n.8, p. 9172-9204. 2023.

MATTAR, F. N. **Pesquisa de marketing**. 3.ed. São Paulo: Atlas, 2001.

MORETTIN, P. A. **Análise de Séries Temporais**. São Paulo: Blucher, 1987.

MORETTIN, C. E.; TOLOI, C. M. C. **Análise de Séries Temporais**. 2. ed. São Paulo: Edgard Blücher, 2006

NASCIMENTO, Pricylla S. C. do *et al.* **Análise dos impactos da gestão do tempo no desempenho acadêmico através da mineração de dados educacionais**. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO (SBIE), 33., 2022, Manaus. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021. p. 1909–1918. Disponível em: <<https://sol.sbc.org.br/index.php/sbie/article/view/18106/17940>>. Acesso em: 26 mai. 2025.

OLIVEIRA, G. H. F. M. **Sistema Híbrido para Previsão de Séries Temporais Educacionais**. Universidade de Pernambuco, 2015. Disponível em: <<https://repositorio.ufpe.br/handle/123456789/18353>>. Acesso em: 29 dez. 2024.

OLIVEIRA, Maxwell F. de **Metodologia Científica: um manual para a realização de pesquisas em administração**. Catalão (GO): UFG, 2011.

PERES, Rodrigo T.; CAMPOS, Letícia B.; MORAES, Carolina P. de. Considerações acerca dos impactos da renda familiar e da escolaridade materna no desempenho em Matemática no ENEM 2022. **REPOD-Revista Educação e Políticas em Debate**. V.14, n.2, maio/agosto 2025.

PIRES, Marlene A. B.; MOTA, Marcia M. P. E. da. Compreensão de leitura no Ensino Médio e desempenho acadêmico em diferentes matérias escolares. **Estudos e Pesquisas em Psicologia**, Rio de Janeiro, v. 21, n. 2, p. 572-589, maio/ago. 2021.

ROCHA, Alyne de C. S. *et al.* Rede neural artificial: identificação de acadêmicos em curso de Licenciatura em Matemática com possibilidade de desistência. **Cadernos Cajuína**, Teresina, v. 7, n. 2, p. 1–15, jul./dez. 2022.

ROSA, Tatiana da ; MAGALHÃES, Cledilene R.; SILVEIRA, Luiz M. de O. B.. Envolvimento família-escola e suas implicações no desempenho escolar na educação básica. **Psicologia Escolar e Educacional**, v. 28, 2024.

SANCHES, Eder A.; GUERRA, Oscar U. Estudo Relacional entre Gênero e Rendimento Escolar de Alunos do Ensino Fundamental e Médio de Brasília-Brasil. **Revista Ibero-Americana de Estudos em Educação**, vol. 16, núm. 1, pp. 37-62, 2021.

SANTANA, Luciene C.; SANTOS, Luzia C. de M. **Análise da falta de interesse e a motivação dos alunos do primeiro ano do Ensino Médio**. Anais do IV Colóquio Internacional Educação e Contemporaneidade. Universidade Federal de Sergipe. 2010. Disponível em: <https://ri.ufs.br/bitstream/riufs/10349/27/26.pdf>. Acesso em: 22 fev. 2025.

SILVA, João Gomes da. **Desempenho escolar do ensino médio no Brasil: o papel da formação docente e da infraestrutura da escola no contexto regional e estadual**. Tese (Doutorado em Demografia) - Centro de Ciências Exatas e da Terra, Universidade Federal do Rio Grande do Norte, Natal, 2023.

SILVEIRA, Andrea C.; ROSA, Dora L.; TENÓRIO, Robson M. Iniquidade de resultados no desempenho escolar de crianças negras e brancas: um estudo em escolas integrantes do Projeto Geres em Salvador – Bahia. **P O I É S I S – Revista Do Programa De Pós-Graduação Em Educação – Mestrado – Universidade do Sul de Santa Catarina**. Vol. 7, n. 14, Jun/Dez. 2013

THIOLLENT, Michel. **Metodologia da Pesquisa-ação**. 18 ed. São Paulo: Cortez, 2011.

THORNTON, Alessandra F.; CORSO, Luciana V. Comprensión lectora y rendimiento en la Escuela Secundaria. **Revista e-Curriculum**, vol .20, n. 2, São Paulo, abr./jun 2022.

TRIVIÑOS, Augusto M. S. **Introdução à Pesquisa em Ciências Sociais: a Pesquisa Qualitativa em Educação**. São Paulo: Atlas, 1987.

ZHANG, G. P. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, v. 50, p. 159-175, 2001.