

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Ciências Econômicas
Programa de Pós-Graduação em Demografia

Estevão de Almeida Vilela

**Trajetórias escolares no Brasil entre 2008 e
2018: uma análise de sequência**

Belo Horizonte

2023

Estevão de Almeida Vilela

Trajetórias escolares no Brasil entre 2008 e 2018: uma análise de sequência

Dissertação apresentada ao Programa de Pós-Graduação em Demografia da Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Demografia.
Orientador: Prof. Bernardo Lanza Queiroz

Belo Horizonte

2023

Ficha catalográfica

V699t Vilela, Estevão de Almeida.
2023 Trajetórias escolares no Brasil entre 2008 e 2018 [manuscrito]:
uma análise de sequência / Estevão de Almeida Vilela. – 2023.
95 f.: il., gráfs., tabs.

Orientador: Bernardo Lanza Queiroz.

Dissertação (mestrado) – Universidade Federal de Minas
Gerais, Centro de Desenvolvimento e Planejamento Regional.
Inclui bibliografia (f. 73-78).

1. Demografia – Teses. 2. Educação – Aspectos econômicos –
Brasil – Teses. 3. Educação básica – Teses. I. Queiroz, Bernardo
Lanza. II. Universidade Federal de Minas Gerais. Centro de
Desenvolvimento e Planejamento Regional. III. Título.

CDD: 304.6

Elaborada por Rosilene Santos CRB6-2527
Biblioteca da FACE/UFMG. RSS – 116/2023



UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE CIÊNCIAS ECONÔMICAS
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL
PROGRAMA DE PÓS-GRADUAÇÃO EM DEMOGRAFIA

ATA DE DEFESA DE DISSERTAÇÃO DE ESTEVÃO DE ALMEIDA VILELA - NÚMERO DE REGISTRO 2021657340.

Às quatorze horas e trinta minutos do dia primeiro do mês de agosto de dois mil e vinte e três, reuniu-se, por videoconferência, a Comissão Examinadora de DISSERTAÇÃO, indicada *ad referendum* pelo Colegiado do Curso em 20/07/2023, para julgar, em exame final, o trabalho final intitulado “Trajetórias escolares no Brasil entre 2008 e 2018: uma análise de sequência”, requisito final para a obtenção do Grau de Mestre em Demografia, área de concentração em Demografia. Abrindo a sessão, o Presidente da Comissão, Prof. Bernardo Lanza Queiroz, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa do candidato. Logo após, a Comissão composta pelos professores Bernardo Lanza Queiroz, Raquel Zanatta Coutinho e José Irineu Rangel Rigotti se reuniu, sem a presença do candidato e do público, para julgamento e expedição do resultado final. A Comissão **APROVOU** o candidato por unanimidade. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 1º de agosto de 2023.

Prof. Bernardo Lanza Queiroz (Orientador) (CEDEPLAR/FACE/UFMG)

Profa. Raquel Zanatta Coutinho (CEDEPLAR/FACE/UFMG)

Prof. José Irineu Rangel Rigotti (CEDEPLAR/FACE/UFMG)

PROFA. PAULA DE MIRANDA RIBEIRO

Coordenadora do Programa de Pós-Graduação em Demografia



Documento assinado eletronicamente por **Raquel Zanatta Coutinho, Professora do Magistério Superior**, em 02/08/2023, às 09:54, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jose Irineu Rangel Rigotti, Professor do Magistério Superior**, em 02/08/2023, às 10:31, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Bernardo Lanza Queiroz, Professor do Magistério Superior**, em 02/08/2023, às 14:30, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Paula de Miranda Ribeiro, Coordenador(a) de curso de pós-graduação**, em 08/08/2023, às 03:23, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2495297** e o código CRC **17FD51B8**.

Referência: Processo nº 23072.245475/2023-80

SEI nº 2495297

*Este trabalho é dedicado às crianças,
que merecem ter condições de se tornar quem quiserem.*

Agradecimentos

Ao Professor Bernardo Lanza Queiroz, pela generosidade em abrir as portas para participar de pesquisas mesmo quando ainda não nos conhecíamos e pela orientação que foi fundamental para meu desenvolvimento e formação.

À banca, Professor José Irineu Rigotti e Professora Raquel Zanatta Coutinho por aceitarem o convite e contribuírem para este trabalho.

À Professora Luciana Luz pelos diálogos sobre educação e demografia.

Aos professores e aos funcionários do CEDEPLAR.

À Paula, que me incentivou a tentar o mestrado em Demografia e esteve presente em todos os momentos difíceis dessa jornada, e os momentos alegres e divertidos também. Principalmente, pelo amor e pelo companheirismo desde que nos conhecessemos.

Ao meu pai, minha mãe e meu irmão, pelo carinho, por estarem sempre presentes e por acreditarem que pode-se sempre ir além.

Aos amigos da demografia que tornaram essa jornada mais alegre e viva.

À CAPES, pelo apoio financeiro que me permitiu dedicar integralmente, durante parte do mestrado.

Resumo

Educação é um direito de todos e dever do Estado e da família, que tem como princípio a igualdade de condições para acesso e permanência na escola. Historicamente, o sistema educacional brasileiro tem conjugado altos índices de repetência e de evasão que evidenciam a baixa eficiência do sistema educacional. As trajetórias escolares dos estudantes brasileiros são marcadas por intercorrências que dificultam a conclusão da educação básica no tempo esperado. O objetivo geral deste trabalho é caracterizar as trajetórias escolares da coorte de estudantes ingressantes no 2º ano do ensino fundamental no Brasil em 2008 e os diferenciais demográficos e socioeconômicos associados. Adota-se uma perspectiva de Curso de Vida Educacional, pela qual se analisa a trajetória educacional como processo de acúmulo de experiências no sistema de ensino. Como metodologia, utiliza-se o arcabouço de análise de sequência e a trajetória escolar é investigada como unidade de análise, considerando-a em sua totalidade. Os resultados mostram que menos da metade dos estudantes da coorte analisada neste trabalho estava na 3ª série do ensino médio no tempo esperado e 27% dos estudantes evadiram em algum ano entre 2008 e 2018. Os padrões de trajetórias identificados mostram que há estudantes que percorreram o sistema de ensino com nenhuma ou pouca intercorrência, enquanto há estudantes em que a trajetória foi marcada por uma ou mais repetências e foi interrompida pela evasão, seja no início, no meio ou no fim da educação básica. Estudantes do sexo masculino, de cor/raça parda ou preta e nível socioeconômico mais baixo estão sobrerrepresentados nos padrões de trajetória escolar associados a maiores graus de repetência e de culminância da evasão, sobretudo a evasão nas etapas iniciais da educação básica. Os padrões de trajetórias escolares identificados podem contribuir para o diagnóstico mais preciso da realidade educacional do sistema de ensino brasileiro.

Palavras-chave: trajetória escolar; curso de vida; análise de sequência.

Abstract

Education is everyone's right and the duty of the State and the family, whose principle is equal conditions for access and permanence in school. Historically, the Brazilian educational system has combined high repetition and dropout rates, which demonstrates the low efficiency of the educational system. The school trajectories of Brazilian students are marked by interurrences that make it difficult to complete basic education in the expected time. The general objective of this work is to characterize the school trajectories of the cohort of students entering the 2nd year of elementary school in Brazil in 2008 and the associated demographic and socioeconomic differentials. An Educational Life Course perspective is adopted, through which the educational trajectory is analyzed as a process of accumulating experiences in the education system. As a methodology, the sequence analysis framework is used and the school trajectory is investigated as a unit of analysis, considering it in its entirety. The results show that less than half of the students in the cohort analyzed in this work were in the 3rd grade of high school in the expected time and 27% of the students dropped out in some year between 2008 and 2018. The patterns of trajectories identified show that there are students who went through the education system with little or no interurrence, while there are students whose trajectory was marked by one or more repetitions and was interrupted by evasion, either at the beginning, in the middle or at the end of basic education. Male, brown, or black students and lower socioeconomic status are overrepresented in the school trajectory patterns associated with higher grades of repetition and dropout rates, especially dropouts in the initial stages of basic education. The patterns of school trajectories identified can contribute to a more accurate diagnosis of the educational reality of the Brazilian education system.

Keywords: school trajectory; life course; sequence analysis.

Lista de ilustrações

Figura 1 – Distribuição de NSE padronizado - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	32
Figura 2 – Matriz de distância entre pares de sequências de trajetórias escolares	40
Figura 3 – Distribuição de NSE padronizado - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	42
Figura 4 – Distribuição relativa dos estados por ano - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	45
Figura 5 – 10 sequências mais comuns - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	46
Figura 6 – Distribuição relativa dos estados por ano e sexo - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	47
Figura 7 – 10 sequências mais comuns por sexo - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	48
Figura 8 – Distribuição relativa dos estados por ano e cor/raça - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	49
Figura 9 – 10 sequências mais comuns por cor/raça - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	50
Figura 10 – Distribuição relativa dos estados por ano e quintil de NSE - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	51
Figura 11 – 10 sequências mais comuns por quintil de NSE - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	52
Figura 12 – Distribuição relativa dos estados por ano - Padrões identificados - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	54

Figura 13 – Tempo médio experimentado por estado - Padrões identificados - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	55
Figura 14 – 10 seqüências mais comuns - Padrões identificados - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018	56
Figura 15 – Árvore de agrupamentos produzido pela função <i>seqtreedisplay</i>	71
Figura 16 – Medidas da qualidade de uma partição - PBC, HG, HC e ASW - 12 agrupamentos	72

Lista de tabelas

Tabela 1 – Coorte inicial, Razões de eliminação de casos inválidos e Coorte final - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental - Brasil	28
Tabela 2 – Distribuição de estudantes por sexo e cor/raça - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental - Brasil - 2008-2018.	30
Tabela 3 – Estados considerados na análise das trajetórias escolares.	36
Tabela 4 – Exemplos de trajetórias escolares hipotéticas	36
Tabela 5 – Exemplos de trajetórias escolares	37
Tabela 6 – Distribuição de estudantes por sexo e cor/raça - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental - Brasil - 2008-2018.	41
Tabela 7 – Estatísticas descritivas - Sexo, Cor/Raça e Nível Socioeconômico por agrupamento - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental - Brasil - 2008-2018.	57
Tabela 8 – Teste Qui-Quadrado de independência de proporções	58
Tabela 9 – Distribuição relativa dos estados por ano - Coorte	74
Tabela 10 – Distribuição relativa dos estados por ano e sexo - Coorte	74
Tabela 11 – Distribuição relativa dos estados por ano e cor/raça - Coorte	75
Tabela 12 – Distribuição relativa dos estados por ano e quintil de NSE - Coorte	76
Tabela 13 – 10 sequências mais comuns - Coorte	77
Tabela 14 – 10 sequências mais comuns por sexo - Coorte	77
Tabela 15 – 10 sequências mais comuns por cor/raça - Coorte	78
Tabela 16 – 10 sequências mais comuns por quintil de NSE - Coorte	79
Tabela 17 – Distribuição relativa dos estados por ano - Padrões identificados - Amostra	80
Tabela 18 – 10 sequências mais comuns - Padrões identificados - Amostra	81
Tabela 19 – Tempo médio por estado - Padrões identificados - Amostra	81

Sumário

1	INTRODUÇÃO	13
2	MARCO TEÓRICO	16
2.1	Trajetórias escolares e o Curso de Vida Educacional	16
2.1.1	Por que os estudantes evadem e qual o papel da repetência?	20
2.2	Desigualdades e o fluxo escolar	22
3	DADOS	25
3.1	Organização da base de dados	25
3.2	Variáveis utilizadas	28
4	METODOLOGIA	33
4.1	Análise de sequências	33
4.2	Sequência de estados e os efeitos <i>timing, duration e sequencing</i>	35
4.3	Análise de sequência e <i>Optimal matching</i>	38
5	RESULTADOS	44
5.1	Padrões de trajetórias escolares	51
5.2	Diferenciais demográficos e socioeconômicos	56
6	DISCUSSÃO E CONCLUSÃO	59
	REFERÊNCIAS	63
	APÊNDICES	69
	APÊNDICE A – SELEÇÃO DO NÚMERO DE AGRUPAMENTOS	70
	ANEXOS	73
	ANEXO A – TABELAS	74
	ANEXO B – SCRIPTS UTILIZADOS	82

1 Introdução

Educação é um direito de todos e dever do Estado e da família, que tem como princípio a igualdade de condições para acesso e permanência na escola (BRASIL, 1988), de modo que um sistema de ensino eficaz deve garantir que os estudantes acessem à escola, permaneçam nela e aprendam (SOARES; ALVES; FONSECA, 2021). No entanto, 2,6% e 9,7% dos estudantes brasileiros de Ensino Fundamental e de Ensino Médio em 2018, respectivamente, não se matricularam em qualquer etapa de ensino no ano de 2019, ou seja, evadiram da escola¹. Estima-se que 17% dos jovens que completaram 16 anos em 2018 não irão concluir a educação básica até os 25 anos de idade (BARROS et al., 2021).

A conclusão da educação básica é de importância significativa tanto para o estudante quanto para a sociedade como um todo. Um estudante que não conclui a educação básica tem escolaridade estimada em 3,6 anos de estudo abaixo dos que concluem a educação básica, o que contribui para uma perda relativa de 37% da remuneração e tem expectativa de vida aos 25 anos estimada 3,3 anos ou 6% inferior à expectativa de vida aos 25 anos de um indivíduo que concluiu a educação básica (BARROS et al., 2021).

No contexto educacional brasileiro, as evidências apontam que a repetência é uma prática que permeia os sistemas de ensino em níveis elevados, de tal modo que poderia haver uma cultura de repetência nas escolas, processo denominado por Ribeiro (1991) como uma "pedagogia da repetência". Ademais, a repetência está intimamente associada à evasão escolar (LEON; MENEZES-FILHO, 2002), (SOUZA et al., 2012), (SHIRASU; ARRAES, 2015), (SOARES et al., 2015).

Sob uma perspectiva de longo prazo, a evasão é a culminância de um processo de desengajamento do estudante (ALEXANDER; ENTWISLE; HORSEY, 1997), no qual a repetência tem papel fundamental em acumular experiências negativas e frustrantes para o estudante (ALEXANDER; ENTWISLE; KABBANI, 2001), contribuir para que os estudantes fiquem em idade mais avançada que seus pares (RODERICK, 1994) e reduzir a participação e a identificação do estudante com a escola, fragilizando as relações (STEARNS et al., 2007).

As trajetórias escolares dos estudantes brasileiros são marcadas por intercorrências que dificultam a conclusão da educação básica no tempo esperado (SOARES; ALVES; FONSECA, 2021), (BOF; OLIVEIRA; BARROS, 2019), (MALAGUTH, 2022), com importantes diferenciais demográficos e socioeconômicos associados. Ainda, há etapas de ensino que apresentam ser mais críticas para determinação de uma trajetória "regular" ou

¹ Taxa de evasão. Indicadores disponíveis em: <<https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/indicadoreseducacionais>>. Acesso em: 31 mai. 2022

"irregular" (MALAGUTH, 2022).

O objetivo geral deste trabalho é caracterizar as trajetórias escolares da coorte de estudantes ingressantes no 2º ano do ensino fundamental no Brasil em 2008 e os diferenciais demográficos e socioeconômicos associados. Para se atingir tal propósito, os seguintes objetivos específicos são perseguidos: i) construir tipologia de estados associados às transições escolares; ii) aplicar métodos de dissimilaridade de trajetórias; iii) identificar padrões de trajetórias escolares; iv) caracterizar as trajetórias escolares segundo as características demográficas – sexo e cor/raça – e socioeconômicas dos estudantes.

Tendo em vista as evidências disponíveis sobre o fluxo escolar dos estudantes no sistema educacional brasileiro e as desigualdades entre grupos sociais observadas, formula-se a hipótese de que as trajetórias escolares associadas à repetência e à evasão, sobretudo aquelas com alto quantitativo de repetência e aquelas com evasão no início da trajetória escolar, são sobrerrepresentadas por grupos socioeconômicos vulnerabilizados.

A estratégia de pesquisa adotada neste trabalho procura avançar no estudo das trajetórias escolares ao propor uma nova abordagem para sua caracterização com: i) a identificação de etapas de ensino que sejam mais críticas para o desenvolvimento dos estudantes em seu percurso educacional, seja por um número acumulado de repetências ou pela culminância da evasão escolar; ii) a diferenciação dos estudantes entre aqueles que evadem nos primeiros anos de escolarização daqueles que persistem por mais anos de ensino.

Adota-se uma perspectiva de Curso de Vida Educacional, pela qual procura-se analisar a trajetória educacional como processo de acúmulo de experiências no sistema de ensino, no qual o resultado é permeado por transições ao longo do percurso escolar (CROSNOE; BENNER, 2016), (PALLAS, 2002). Como metodologia, utiliza-se o arcabouço de análise de sequência que permite investigar um conceito fundamental de Curso de Vida que é a trajetória, considerando-a em sua totalidade (ABBOTT, 1995), (AISENBREY; FASANG, 2010), com o objetivo principal de identificar padrões em trajetórias, a partir de uma comparação entre pares de trajetórias ou em relação a um padrão esperado (HALPIN, 2010), (AISENBREY; FASANG, 2010).

Esta dissertação contém seis capítulos, incluindo esta introdução. No [Capítulo 2](#), discute-se o paradigma de Curso de Vida (ELDER; SHANAHAN, 2007) aplicado ao contexto educacional - Curso de Vida Educacional (CROSNOE; BENNER, 2016) - o qual é definido como o processo de acúmulo de experiência no sistema educacional e que pode ser analisado sob as perspectivas de aprendizado, de desempenho e de resultado. Argumenta-se que a adoção de tal paradigma contribui para o entendimento das trajetórias escolares percorridas pelos estudantes brasileiros na educação básica.

Em seguida, discute-se a relação entre repetência e evasão, apresentando os mecanis-

mos pelos quais essa relação se dá, as evidências empíricas produzidas a nível internacional e nacional e as lacunas não preenchidas por tais evidências no contexto educacional brasileiro. Argumenta-se que ao tratar da relação entre repetência e evasão é necessário colocar tais fenômenos na perspectiva de trajetória.

No [Capítulo 2](#), por fim, apresentam-se os trabalhos que discutem o fluxo escolar no Brasil, discute-se a importância que a repetência tem tido na conformação dos resultados educacionais e as desigualdades demográficas e socioeconômicas que permeiam tais resultados.

O [Capítulo 3](#) apresenta os dados utilizados nessa dissertação, com especial atenção ao tratamento ao qual foram submetidos. A base de dados longitudinal do censo escolar (INEP, 2017) constitui fonte essencial para caracterizar as trajetórias escolares dos estudantes no Brasil. A seleção da coorte de estudantes analisada neste trabalho procurou evidenciar de que modo a repetência e a evasão condicionam trajetórias distintas dado que os estudantes estão na etapa adequada para sua idade.

No [Capítulo 4](#), apresenta-se a metodologia adotada para responder a pergunta de pesquisa deste trabalho. Adota-se a abordagem de análise de sequência, arcabouço metodológico que possibilita investigar trajetórias, considerando-as em sua totalidade (ABBOTT, 1995), (AISENBREY; FASANG, 2010). As trajetórias educacionais são investigadas em relação a três aspectos de sequências de estados: tempo consecutivo experimentado em cada estado (*duration*); quando se dá a ocorrência de cada estado (*timing*) e ordem na qual os distintos estados ocorrem (*sequencing*) (LIAO et al., 2022).

Essencialmente, a análise das trajetórias se dá pela investigação de padrões nas trajetórias consideradas (AISENBREY; FASANG, 2010), (STUDER; RITSCHARD, 2016), e para tal utiliza-se o método *Optimal matching* (ABBOTT; HRYCAK, 1990), adaptado da biologia para identificação de semelhança em cadeias de DNA, o qual gera uma medida da similaridade entre um par de sequências e essa medida pode então ser utilizada para identificar padrões no conjunto de dados.

Os resultados apresentados no [Capítulo 5](#) mostram que menos da metade dos estudantes da coorte analisada neste trabalho estavam na 3ª série do ensino médio no tempo esperado e 27% dos estudantes estavam fora da escola. Os padrões de trajetórias identificados mostram que há estudantes que percorrem o sistema de ensino com nenhuma ou pouca intercorrência, enquanto há estudantes em que a trajetória é marcada por uma ou mais repetências e é interrompida pela evasão, seja no início, no meio ou no fim da educação básica. As desigualdades educacionais relativas às variáveis demográficas - sexo e cor/raça - e socioeconômicas são destacadas.

No [Capítulo 6](#), discutem-se os resultados obtidos, as contribuições deste trabalho, as limitações encontradas e possíveis trabalhos futuros.

2 Marco Teórico

2.1 Trajetórias escolares e o Curso de Vida Educacional

A abordagem de Curso de Vida estabelece um paradigma de pesquisa no qual a vida dos indivíduos se dá pela realização de papéis que são regulados pela idade e pelas instituições sociais (ELDER; SHANAHAN, 2007). Portanto, Curso de Vida se refere no nível individual à trajetória que é percorrida por cada indivíduo dentre um conjunto de possíveis caminhos esperados a partir de normativas sociais, no nível da população, conformando um fenômeno multinível (ELDER, 1998);(CROSNOE; BENNER, 2016).

O paradigma de Curso de Vida pode ser abordado de forma geral ou contingente a uma área específica da vida social. Nesse sentido, no âmbito educacional, o percurso de escolarização formal é caracterizado por um sistema institucional estruturado no qual os indivíduos percorrem em maior ou menor medida caminhos previamente definidos (CROSNOE; BENNER, 2016). Pode se definir, portanto, o Curso de Vida Educacional, o qual é o processo de acúmulo de experiência no sistema educacional, levando em consideração as experiências auferidas antes de sua entrada, durante e após sua conclusão (CROSNOE; BENNER, 2016).

Segundo Crosnoe e Benner (2016), o Curso de Vida Educacional pode ser analisado sob três perspectivas: aprendizado, que diz respeito ao desenvolvimento de habilidades e apreensão de conhecimento; desempenho, que diz respeito a uma avaliação do aprendizado; e resultado, que diz respeito ao progresso no sistema de ensino. Neste trabalho, é de especial interesse o resultado educacional, compreendido como a progressão no sistema de ensino e tal aspecto do Curso de Vida Educacional está diretamente relacionado aos conceitos de caminho, trajetória e transição.

Caminhos se referem às sequências de posições conformadas socialmente e por instituições que ditam, por exemplo, em que idade mínima e máxima e por quanto tempo deve se estar em determinado estado ou que a idade é um importante indicativo para se dizer de transições antecipadas, no tempo ou tardias (ELDER; SHANAHAN, 2007). No contexto educacional, o caminho esperado de um estudante é que ele entre na escolarização formal na idade certa e faça seu percurso no tempo adequado se graduando na idade esperada.

Portanto, ao passo que o caminho conforma social e institucionalmente os percursos esperados, a trajetória é a realização de cada indivíduo sobre esses caminhos possíveis. Intrinsecamente ligadas às trajetórias, são as transições entre estados que dão característica singular à cada trajetória individual (ELDER, 1985), sendo diferenciadas por quando

ocorrem as transições, quanto se experimenta cada estado até a próxima transição e quais transições são experimentadas. Uma trajetória de conclusão da escolarização formal contempla as transições de entrada na escola, progressões sucessivas nas etapas de ensino, graduações nas etapas intermediárias e, por fim, formatura que conclui o processo.

No Brasil, a forma predominante de organização do sistema educacional é a seriada com a progressão para uma determinada etapa de ensino condicionada à aprovação na etapa anterior. A Lei nº 5.692 de 1971 tornou obrigatória a matrícula de estudantes entre 7 e 14 anos de idade no à época chamado "primeiro grau", que contemplava oito anos de estudo. Em 2006, a Lei nº 11.274 ampliou a obrigatoriedade do ensino, instituindo o ensino fundamental de nove anos, com matrícula obrigatória a partir dos 6 anos de idade. Em 2009, a Emenda Constitucional nº 59 expandiu o ensino obrigatório para aqueles entre 4 e 17 anos de idade, contemplando a educação infantil e o ensino médio.

Dessa forma, atualmente, uma trajetória sem intercorrências após a entrada à escola prevê treze progressões ao longo da educação básica, que se inicia aos 4 anos e termina aos 17 anos de idade, culminando em 14 anos de ensino obrigatórios (BRASIL, 1988), (BRASIL, 2006). No entanto, tal "caminho" chamado de "regular" não tem sido observado por todos estudantes na educação brasileira.

O fluxo escolar¹, que se refere ao progresso dos estudantes no sistema de ensino, tem-se mostrado ineficiente: a partir de simulações com dados de taxas de transição entre séries, Klein (2006) identificou pequeno aumento no percentual de jovens concluintes na 3ª série do ensino médio, passando de 30,9% em 1992 para 35,7% em 2003, com um tempo médio de conclusão reduzido de 16 anos para 13,6 anos de estudo (em um ensino de 11 anos, à época).

A melhoria observada no período se deu sobretudo pela redução nas taxas de repetência observadas entre os anos de 1992 e 2003: taxa de repetência média entre 1ª e 4ª série do ensino fundamental passa de 35% para 19,80%; taxa de repetência média entre 5ª e 8ª série do ensino fundamental passa de 30,2% para 18,5%; e a taxa de repetência média entre 1ª e 3ª série do ensino médio passa de 31,8% para 20,62% (KLEIN, 2006). A despeito da melhoria apresentada no período, Klein (2006) aponta que cerca de 50% dos jovens de 15 anos não estavam cursando a série adequada para sua idade em 2003. Portanto, os estudantes avançavam pouco no sistema de ensino, gerando uma ineficiência em termos de oferta de ensino e contribuindo para o maior risco de evasão escolar.

Os indicadores recentes de fluxo escolar apontam melhorias: a taxa de repetência

¹ Fluxo escolar se refere à transição do aluno entre dois anos consecutivos considerando os seguintes cenários possíveis: promoção, repetência, migração para EJA e evasão de escola. Rendimento escolar se refere à situação final do aluno declarada na segunda etapa da coleta do Censo Escolar e consideram os dados de alunos que ao final do ano letivo foram aprovados ou reprovados ou que durante o ano letivo abandonaram a escola. <https://download.inep.gov.br/educacao_basica/censo_escolar/apresentacao/2017/apresentacao_indicadores_de_fluxo_escolar_da_educacao_basica.pdf>

na 1ª série do ensino fundamental (de 8 anos) passou de 60,3% em 1982 para 1,2% em 2017, enquanto a mesma taxa para a 1ª série do ensino médio passou de 35,8% em 1982 para 14,3% em 2017 (KLEIN, 2022). No entanto, as taxas de evasão mostram que para o ensino fundamental houve melhorias enquanto para o ensino médio permanece o desafio: a taxa de evasão na 1ª série do ensino médio passou de 12,2% em 1982 para 13,6% em 2017 (KLEIN, 2022).

Os indicadores de fluxo escolar são componentes essenciais para se avaliar a eficácia do sistema de ensino em garantir a permanência e a progressão dos estudantes. No entanto, tais indicadores são insuficientes para se avaliar como os estudantes percorrem o sistema de ensino.

O estudo de trajetórias escolares em larga escala se tornou possível no Brasil a partir do desenvolvimento de uma base de dados longitudinal, que será abordada detalhadamente no Capítulo 3, construída a partir do Censo Escolar - levantamento estatístico educacional realizado anualmente pelo INEP (INEP, 2017). A partir dessa base de dados é possível realizar o acompanhamento do percurso de cada estudante matriculado a partir de 2007 na educação brasileira, identificando a etapa de ensino que o aluno esteve matriculado, o rendimento e o fluxo escolar.

A base de dados longitudinal (INEP, 2017) tem sido utilizada para estudo das trajetórias educacionais no Brasil com diferentes objetivos: Bof, Oliveira e Barros (2019) investigaram a relação entre trajetória e aprendizagem; Soares, Alves e Fonseca (2021) descreveram as trajetórias escolares; e Malaguth (2022) investigou a relação entre migração e fluxo educacional.

As trajetórias escolares dos estudantes brasileiros são marcadas por irregularidades. Analisando a coorte de estudantes nascidos entre 1º de julho de 1999 e 30 de junho de 2000 acompanhados entre os anos de 2007 e 2015, Soares, Alves e Fonseca (2021) identificam que 46,9% dos estudantes não percorreram uma trajetória “regular”, sendo que 16,5% percorreram com “poucas irregularidades” e 30,4% percorreram com “grandes irregularidades”. A partir de outra unidade de análise - estudantes do 5º ano do ensino fundamental que fizeram a Prova Brasil em 2011 - Bof, Oliveira e Barros (2019) identificaram que 31,8% dos estudantes não seguiram uma trajetória regular nos 5 anos acompanhados - de 2011 a 2015. Focando apenas nos estudantes nascidos entre 1º de julho de 2001 e 30 de agosto de 2002 que estavam matriculados no 1º ano do ensino fundamental de 9 anos em Minas Gerais em 2008, Malaguth (2022) identificou que 50,5% da coorte investigada alcançou a última etapa da educação básica no tempo esperado, enquanto 22,4% haviam evadido em algum ano do ensino regular.

Destacam-se também outros resultados importantes: Bof, Oliveira e Barros (2019) identificaram que dentre os estudantes que percorreram uma trajetória regular, 36,9% haviam sido reprovados antes de cursarem o 5º ano do ensino fundamental em 2011, por

outro lado, dentre os estudantes com trajetória irregular, 59,7% haviam sido reprovados antes de 2011; [Malaguth \(2022\)](#) identificou que estudantes que se tornam irregulares se matriculam continuamente até os 15 anos de idade, momento no qual passam a evadir da escola de forma mais acentuada, sobretudo aqueles com maior defasagem idade-série.

Em relação a maior incidência das transições de repetência e de evasão, [Malaguth \(2022\)](#) identificou que a repetência é acentuada no 6º ano do ensino fundamental e no 1º ano do ensino médio tanto para os estudantes "regulares" até então quanto para os "irregulares". Em relação à evasão, essa aumenta a partir do 9º ano do ensino fundamental para estudantes "regulares" até então, enquanto para os estudantes "irregulares" ela aumenta desde o 6º ano do ensino fundamental com uma incidência mais elevada no 1º ano do ensino médio.

As evidências apontam que os estudantes que apresentam histórico de repetência tendem a apresentar resultados educacionais menos satisfatórios com maiores chances de irregularidades em suas trajetórias, seja por repetências recorrentes ou com culminância de evasão escolar.

A trajetória educacional é caracterizada de formas distintas nesses trabalhos. [Soares, Alves e Fonseca \(2021\)](#) definiram trajetórias como “regulares”, “com poucas irregularidades” e “com muitas irregularidades” a partir de cinco critérios que se combinam com pesos diferentes: presença de evasão; número de anos de sucesso escolar; número de anos de atraso escolar; ano escolar atingido em 2015; e presença de repetência. Já [Bof, Oliveira e Barros \(2019\)](#) caracterizaram as trajetórias em "regular", sendo os estudantes que foram continuamente promovidos, e "irregular", sendo todos os demais estudantes que por algum motivo não estavam matriculados no 9º ano do ensino fundamental em 2015, estratégia semelhante é adotada por [Malaguth \(2022\)](#) que definiu trajetória "regular" aquela que consta apenas promoções sucessivas.

As trajetórias escolares dos estudantes brasileiros são marcadas por intercorrências que dificultam a conclusão da educação básica no tempo esperado ([SOARES; ALVES; FONSECA, 2021](#)), ([BOF; OLIVEIRA; BARROS, 2019](#)), ([MALAGUTH, 2022](#)). Ainda, há etapas de ensino que aparentam ser mais críticas para determinação de uma trajetória "regular" ou "irregular" ([MALAGUTH, 2022](#)). Dessa forma, identifica-se que há espaço para que a caracterização das trajetórias escolares proposta por [Soares, Alves e Fonseca \(2021\)](#) incorpore a identificação das etapas "críticas" para escolarização identificadas por [Malaguth \(2022\)](#).

A estratégia de pesquisa adotada neste trabalho procura suprir em parte essa lacuna, avançando no estudo das trajetórias escolares ao propor uma nova abordagem para sua caracterização com: i) a identificação de etapas de ensino que sejam mais críticas para o desenvolvimento dos estudantes em seu percurso educacional, seja por um número acumulado de repetências ou pela culminância da evasão escolar; ii) a diferenciação dos

estudantes entre aqueles que evadem nos primeiros anos de escolarização daqueles que persistem por mais anos de ensino.

2.1.1 Por que os estudantes evadem e qual o papel da repetência?

A evasão escolar é um fenômeno multidimensional, de forma que são diversos os seus determinantes distais e próximos (RUMBERGER; LIM, 2008)². Sob uma perspectiva de longo prazo, a evasão é a culminância de um processo de desengajamento do estudante (ALEXANDER; ENTWISLE; HORSEY, 1997), no qual a repetência tem papel fundamental em acumular experiências negativas e frustrantes para o estudante (ALEXANDER; ENTWISLE; KABBANI, 2001).

A relação entre o desempenho acadêmico, especificamente a reprovação ou a repetência de uma determinada etapa de ensino, e a evasão escolar tem sido explorada e as evidências apontam para uma associação positiva entre reprovação e evasão escolar (RODERICK, 1994); (ALEXANDER; ENTWISLE; HORSEY, 1997); (STEARNS *et al.*, 2007); (GIANO; WILLIAMS; BECNEL, 2022), mas também para uma relação causal entre esses fenômenos (JACOB; LEFGREN, 2009); (ANDREW, 2014).

A reprovação pode contribuir para o aumento do risco de evasão escolar por meio dos seguintes mecanismos: redução na participação e identificação com a escola, redução na autoestima e fragilização da relação com atores escolares – pares, professores e outros agentes da escola (STEARNS *et al.*, 2007). Além disso, por meio de um efeito indireto, a reprovação contribui para que o estudante fique em idade mais avançada que seus pares (RODERICK, 1994). A reprovação em determinadas etapas de ensino pode ter impactos diferenciados no risco de evasão escolar, sendo as que ocorrem nas etapas intermediárias mais significativas que as que ocorrem nas etapas escolares iniciais, em virtude de ocorrerem simultaneamente com outros processos de transformação experimentados pelos jovens, como biológica (puberdade), cognitiva (maior demanda intelectual) e social (complexificação da relação com pares e professores) (GIANO; WILLIAMS; BECNEL, 2022), mas também a reprovação na primeira etapa de ensino reduz a chance de completar o ensino médio (ANDREW, 2014).

A relação entre repetência e evasão é complexa: i) fatores não observáveis podem causar a repetência de forma que sua associação com a evasão se torna mal especificada (JACOB; LEFGREN, 2009); ii) a causalidade pode ocorrer em ambos os sentidos, sendo que a direção da evasão para repetência é menos intuitiva, mas pode ocorrer em virtude

² Rumberger e Lim (2008) apresentam uma detalhada e extensa revisão das causas da evasão escolar que podem ser decompostas em fatores individuais e institucionais. Dentre os fatores individuais, destacam-se i) características demográficas; ii) atitudes, incluídos autopercepção, valores e objetivos; iii) comportamentos, incluídos engajamento, relacionamento com pares e inserção no mercado de trabalho; e, iv) desempenho, incluídos desempenho acadêmico – notas obtidas em provas e testes padronizados, mobilidade escolar e rendimento. Dentre os fatores institucionais, destacam-se i) a família; ii) a escola; e iii) a comunidade.

de o estudante que tem intenção de evadir da escola diminuir suas expectativas e seu desempenho acadêmico incorrendo em reprovação (RODERICK, 1994).

De forma a separar o que é causa e o que é efeito, Jacob e Lefgren (2009) utilizaram uma abordagem causal e identificaram que a reprovação na 8ª série tem impacto direto no risco de o estudante não completar o ensino médio, enquanto a reprovação na 6ª série não apresenta o mesmo risco, em virtude de, possivelmente, os estudantes reprovados nesta série terem mais tempo hábil para recuperar-se academicamente do que os reprovados naquela série.

No contexto educacional brasileiro, diferentes abordagens empíricas apresentam evidências de uma associação positiva entre reprovação e evasão escolar (LEON; MENEZES-FILHO, 2002); (SOUZA et al., 2012); (SHIRASU; ARRAES, 2015); (SOARES et al., 2015). Entre os anos 1984 e 1997, houve redução da taxa de evasão no ensino fundamental, mas o risco para aqueles não-aprovados foi de cinco vezes maior (LEON; MENEZES-FILHO, 2002), resultado similar encontrado por Souza et al. (2012), no entanto, estes encontraram redução da importância da não-aprovação no risco de evasão escolar, analisando a primeira década dos anos 2000. Na educação básica pública do Ceará, a reprovação no ensino médio dobra a chance de evasão escolar (SHIRASU; ARRAES, 2015), enquanto em Minas Gerais cada ano de atraso escolar no ensino médio implica em um risco de 5% de abandono escolar³(SOARES et al., 2015).

A despeito de evidências encontradas no contexto educacional brasileiro sobre a relação entre repetência e evasão escolar, algumas lacunas ainda persistem: unidade de análise espacial restrita à apenas uma Unidade da Federação (SHIRASU; ARRAES, 2015); (SOARES et al., 2015) ou à apenas as seis maiores Regiões Metropolitanas (SOUZA et al., 2012); dependência administrativa da escola restrita apenas à rede pública de ensino (SHIRASU; ARRAES, 2015); (SOARES et al., 2015); variável de interesse definida como abandono escolar ou se está frequentando a escola (SOARES et al., 2015); evasão superestimada em virtude de definição imprecisa do fenômeno (SHIRASU; ARRAES, 2015) ou de desconsideração da migração interescolar (SOUZA et al., 2012); etapas de ensino restritas ao ensino médio (SHIRASU; ARRAES, 2015) ou às etapas de transição de ciclos – 4ª e 8ª séries do ensino fundamental (respectivamente, 5º e 9º ano, no ensino fundamental de 9 anos) e 3ª série do ensino médio (LEON; MENEZES-FILHO, 2002); indefinição de qual etapa de ensino ocorreu a reprovação (SHIRASU; ARRAES, 2015); (SOARES et al., 2015).

A trajetória dos estudantes brasileiros é marcada por alto grau de repetência que se acumulam na vida estudantil e dificultam o progresso para a conclusão das etapas

³ A evasão escolar é identificada na transição entre anos escolares em que se identifica a ausência de matrícula do estudante no ano seguinte. O abandono escolar é identificado dentro do ano escolar em que se identifica o desligamento do estudante com a rede de ensino.

de escolarização da educação básica, inclusive incluindo nesse desfecho trajetórias de rompimento com a escola. Argumenta-se que analisar a trajetória escolar em sua totalidade é uma contribuição relevante para entender como a repetência e a evasão conformam trajetórias escolares distintas.

2.2 Desigualdades e o fluxo escolar

A escolarização é um das principais formas que um indivíduo dispõe para acessar condições de vida melhores a partir de posições ocupacionais e de renda superiores (RIBEIRO, 2011). Os resultados educacionais são determinados por diversos fatores, como recursos familiares, características dos sistemas educacionais e fatores contextuais à localização das escolas (BUCHMANN; HANNUM, 2001), de modo que desigualdades nesses fatores podem contribuir para a manutenção e/ou aumento das desigualdades. Portanto, as políticas educacionais podem contribuir para criar condições para minimizar o impacto das desigualdades de insumos importantes para os resultados educacionais (ALVES; SOARES; XAVIER, 2016).

Ribeiro (2011) procurou investigar como o sistema educacional brasileiro e os recursos familiares contribuíram na superação ou na manutenção das desigualdades de oportunidades e de resultados educacionais. O autor definiu desigualdades de oportunidades educacionais como a chance diferencial de acesso à cada nível do sistema educacional e desigualdades de resultados educacionais como o efeito diferencial de cada transição educacional no total de anos de escolarização obtidos.

Utilizando uma amostra de 8.359 chefes e cônjuges de domicílios entre 25 e 64 anos da pesquisa "Dimensões Sociais das Desigualdades" de 2008, Ribeiro (2011) aplicou modelos de regressão logística para investigar as chances de um indivíduo realizar transições educacionais entre os níveis de ensino no sistema educacional brasileiro: completar as quatro séries iniciais do ensino fundamental; completar as oito séries do fundamental; completar o ensino médio; entrar na universidade; e completar a universidade. Como as transições são condicionais e sequenciais o modelo utilizado é comumente denominado de "modelo de transições educacionais". As variáveis explicativas utilizadas foram: sexo e raça/cor, condição econômica da família, escolarização da mãe e do pai, contexto regional, estrutura familiar, dependência administrativa da escola anterior à cada transição, e coorte de nascimento.

Dentre os resultados encontrados por Ribeiro (2011), destaca-se a concentração das desigualdades em termos de oportunidades educacionais nas transições educacionais iniciais. Condições econômicas mais favoráveis, maior escolaridade do pai e da mãe, ser de cor/raça branca, residir em área urbana e na regiões sul e sudeste contribuíram para maior chance de realizar as transições educacionais com sucesso. Além disso, em termos

de resultados educacionais, [Ribeiro \(2011\)](#) encontrou que as transições de conclusão do ensino fundamental e conclusão do ensino médio foram as que mais contribuíram para as desigualdades observadas de anos de escolarização completos.

No sistema educacional brasileiro, as progressões entre séries e as transições educacionais consideradas por [Ribeiro \(2011\)](#) são condicionadas à promoções nas etapas de ensino intermediárias. No entanto, nesse percurso de escolarização a trajetória escolar do estudante pode ser marcada por uma ou mais repetências e também pode ser interrompida quando há evasão da escola.

As estatísticas oficiais da educação brasileira nos anos 1980 apontavam que um dos principais problemas educacionais era a alta taxa de evasão escolar, notadamente nas primeiras etapas da Educação Básica, que contribuía como principal fator para a baixa escolaridade da população. No entanto, tal como em outros países latino-americanos, no Brasil duas evidências apontavam inconsistências nesses indicadores, de acordo com [Klein e Ribeiro \(1991\)](#): incompatibilidade entre o tamanho das coortes com idade de entrada à escola e o número de novos alunos na 1ª série da Educação Básica; taxas de evasão nas primeiras etapas de ensino sistematicamente elevadas e taxas de repetência nas mesmas etapas sistematicamente baixas.

Portanto, ao analisar esses indicadores educacionais, [Klein e Ribeiro \(1991\)](#) apontaram que uma das causas dessas inconsistências era a definição incorreta de alunos "repetentes". A definição à época no âmbito das estatísticas oficiais considerava que o aluno que durante o ano letivo deixava de frequentar as aulas não era aprovado e nem reprovado e caso se matriculasse na mesma etapa de ensino no ano letivo seguinte era considerado como um "aluno novo". Portanto, a partir dessa definição, os indicadores educacionais sistematicamente subestimavam o número de alunos repetentes e superestimavam os alunos novos ([KLEIN; RIBEIRO, 1991](#)).

A aplicação do modelo PROFLUXO⁴ aos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) por [Klein e Ribeiro \(1991\)](#) e [Ribeiro \(1991\)](#) mostrou que a realidade dos dados educacionais apresentava um panorama diferente: os indicadores oficiais apontavam uma taxa de evasão escolar na primeira etapa de ensino quase 10 vezes maior do que a estimada e a taxa de repetência na mesma série 40% menor do que a estimada; em relação ao acesso, estimou-se um acesso à escola 93% também na primeira etapa de ensino, evidenciando a sua quase universalização.

Dessa forma, trouxe-se à tona evidências de que os problemas educacionais brasileiros não diziam respeito à dificuldade de acesso à escola, mas sim que o problema residia na ineficiência do sistema de ensino: a média de anos de ensino recebida pela população por aluno que se graduava no ensino fundamental era de 19,7 anos, enquanto o tempo

⁴ Para uma explicação do modelo ver [Golgher \(2004\)](#) e [Santos \(2021\)](#).

ideal para cursar esse nível de ensino era de 8 anos (KLEIN; RIBEIRO, 1991). As taxas de repetências eram tal altas que suscitava-se se havia uma cultura de repetências nas escolas, processo que Ribeiro (1991) nomeia como “pedagogia da repetência”.

O sistema de ensino brasileiro tem-se mostrado desigual, com diferentes grupos sociais tendo resultados educacionais díspares, seja no acesso desigual às transições escolares mais avançadas ou seja no acúmulo de anos de escolarização (RIBEIRO, 2011). Um sistema de ensino eficaz deve garantir que os estudantes acessem à escola, permaneçam nela e aprendam (SOARES; ALVES; FONSECA, 2021), sendo a equidade uma dimensão relevante de tal sistema de ensino (ALVES; SOARES; XAVIER, 2016).

A equidade é uma dimensão que informa em que medida as desigualdades podem conformar trajetórias escolares distintas (ALVES; SOARES; XAVIER, 2016). Na investigação das trajetórias escolares dos estudantes no sistema de ensino brasileiro, Soares, Alves e Fonseca (2021) encontraram que: meninos tiveram em média trajetórias mais irregulares que meninas; os estados do Norte e do Nordeste apresentaram maior participação de trajetórias “com muitas irregularidades”; estudantes pretos e pardos tiveram em média trajetórias mais irregulares que estudantes brancos; e estudantes de escolas com maior nível socioeconômico apresentaram em média trajetórias mais regulares.

Os resultados encontrados por Malaguth (2022) em relação às taxas de repetência mostraram uma desigualdade acentuada entre grupos sociais: considerando estudantes regulares, a taxa de repetência no 1º ano do Ensino Médio foi de 21,2% para estudantes do sexo masculino de cor/raça preta, sendo esse valor 60% maior do que estudantes do mesmo sexo e de cor/raça branca e 300% maior em comparação a estudantes brancas.

Em relação à evasão, Malaguth (2022) identificou que as diferenças observadas considerando sexo e cor/raça nas taxas de repetência não se repetiram na mesma intensidade, mesmo considerando separadamente estudantes regulares e não regulares. Apesar dessa pequena diferença nas taxas de evasão, os resultados encontrados por Malaguth (2022) indicaram uma proporção de 10 p.p. há mais de estudantes pardos do sexo masculino evadidos em relação aos estudantes brancos do mesmo sexo e uma proporção de 7 p.p. há mais para estudantes pretos em relação aos brancos também do sexo masculino. Essa diferença de acordo com Malaguth (2022) é atribuída às taxas de repetência, que apresentaram um diferencial substantivo por sexo e raça/cor e ao fato de os estudantes evadidos em quase totalidade experimentarem pelo menos uma repetência antes da evasão.

As trajetórias escolas dos estudantes brasileiros são marcadas por um alto grau de intercorrências, que se distribuem de forma desigual entre os grupos sociais, considerando características como sexo, cor/raça e nível socioeconômico. Considerar tais aspectos no estudo das trajetórias escolares é relevante pois permite a caracterização das diferenças entre grupos sociais e permite a identificação dos estudantes excluídos (SOARES; ALVES; FONSECA, 2021).

3 Dados

3.1 Organização da base de dados

Até o ano de 2006, o INEP coletava informações administrativas do sistema educacional brasileiro como o número de estudantes, de docentes e de turmas agregadas ao nível da escola. A partir de 2007, o Censo Escolar realizado anualmente pelo INEP passa a coletar as informações de matrícula de cada estudante em nível individual, possibilitando a coleta de informações contextualizadas a nível do estudante, como a etapa de ensino, data de nascimento, sexo, cor/raça, entre outras.

A realização periódica do Censo Escolar desde 2007 e a coleta de informações de matrícula a nível do estudante tornou possível a criação de uma base longitudinal a partir dos registros administrativos informados pelas redes de ensino. A construção dessa base necessitou de um esforço para identificação de inconsistências. De forma geral, o processo de tratamento dos censos escolares consistiu em: deduplicação, pela qual se mantém um registro único por estudante; padronização de variáveis, com vistas a harmonizar as informações coletadas entre os censos, realizados anualmente desde 2007; identificação de trajetórias escolares inconsistentes; imputação de informação para os casos inconsistentes com base em indivíduos com trajetórias similares (INEP, 2017). O tratamento de inconsistências de deduplicação e de imputação contribuíram para a correção das estimativas das taxas de transição, notadamente a taxa de evasão (INEP, 2017).

A base de dados longitudinal do censo escolar (INEP, 2017), utilizada nesta dissertação¹, constitui fonte essencial para caracterizar as trajetórias escolares dos estudantes no Brasil haja visto que sua metodologia é a que melhor permite a identificação acurada de eventos relativos ao fluxo escolar.

A base longitudinal do censo escolar (INEP, 2017) tem sido utilizada como fonte de dados por outros trabalhos. Soares, Alves e Fonseca (2021) selecionaram como coorte de estudo os estudantes nascidos entre 1 de julho de 1999 e 30 de junho de 2000 e identificaram uma total de 3.318.649 estudantes com algum registro administrativo de matrícula entre 2007 e 2015. Em 2007, esses estudantes deveriam estar matriculados, no entanto, 8,8% não tinham registro de matrícula nesse ano, apenas nos anos seguintes. Foram três as categorias de exclusão de matrículas adotadas pelos autores: 1) falecidos; 2) estudantes com matrículas em anos escolares incompatíveis com sua idade; 3) “alunos com oito anos, mas com registro de matrícula apenas em 2007”. Ao final, os 3.190.109 estudantes incluídos na análise apresentaram um total de 80.009 trajetórias escolares.

¹ Acordo de Cooperação Técnica INEP-CEDEPLAR n° 3/2021, Processo n° 23036.002230/2020-53.

Já [Malaguth \(2022\)](#) selecionou como coorte de estudo os estudantes nascidos entre 1 de julho de 2001 e 30 de agosto de 2002 que estavam matriculados no 1º ano do ensino fundamental de 9 anos em Minas Gerais em 2008. Do total de estudantes dessa coorte de nascimento, foram excluídos 2,4% que estavam matriculados na educação infantil e 4,6% matriculados em anos escolares que não o 1º ano do ensino fundamental de 9 anos. Após esse primeiro filtro, que resultou em um total de 272.350 estudantes, foram excluídos os estudantes: 1) falecidos – 0,2%; 2) com ano escolar incompatível com a idade – 2,7%, sendo estudantes que no ano calendário $t+1$ se matricularam em um ano escolar menor do que o matriculado em t ou estudantes que se matricularam em um ano escolar superior ao que seria esperado se cumprissem um fluxo regular; 3) por fim, estudantes que possuíam fluxo escolar diferente de “Promovido”, “Repetente”, “Evadido” e “Migração para EJA” – 0,7%. Dos 262.740 estudantes restantes após os filtros realizados, foram excluídos os estudantes de cor/raça “Amarela” (522) e “Indígena” (296).

O tratamento da base de dados longitudinal ([INEP, 2017](#)) realizado por outros trabalhos ([SOARES; ALVES; FONSECA, 2021](#)); ([MALAGUTH, 2022](#)) é relevante para se compreender a variedade de análises possíveis considerando diferentes objetivos de pesquisa. Como se verá a seguir, este trabalho apresenta semelhanças e diferenças em relação ao tratamento dos dados.

A Lei n. 11.274 de 6 de fevereiro de 2006 ampliou o ensino fundamental para nove anos de duração, com a matrícula obrigatória para crianças de seis anos de idade e estabeleceu prazo de implantação pelos sistemas de ensino até 2010 ([BRASIL, 2006](#)). Dessa forma, em 2007 - primeiro ano para qual é possível acompanhar longitudinalmente os estudantes - ainda havia uma diferença no ritmo de implantação do ensino fundamental de 9 anos. Buscando conciliar essa diferença entre as redes de ensino e a disponibilidade da base de dados, que se inicia em 2007 e está disponível até 2018, optou-se por selecionar a coorte de estudantes nascidos entre 1 de julho de 2000 e 30 de junho de 2001 matriculados em 2008.

Essa coorte de estudantes deveria estar matriculada na primeira série do ensino fundamental de 8 anos ou no segundo ano do ensino fundamental de 9 anos em 2008, caso entrassem na escola na idade correta e feito a progressão escolar como esperado. No entanto, 65,77% dos estudantes dessa coorte estavam de fato matriculados em uma dessas duas etapas de ensino. Apesar de não ser a totalidade das matrículas, optou-se pela escolha dessa coorte pois, após 11 anos de escolarização é esperado que esses alunos estivessem concluindo o ensino médio em 2018, último ano de análise deste trabalho, caso seguissem uma trajetória escolar regular, sem repetência e/ou evasão.

Feita a seleção da coorte de nascidos entre 1 de julho de 2000 e 30 de junho de 2001²

² A prática vigente à época considerava os estudantes que nascidos nesse intervalo ([SOARES; ALVES; FONSECA, 2021](#))

matriculados em 2008³, que compõe um total de 3.322.975 estudantes, foram selecionados apenas aqueles que estavam matriculados na 1ª série do ensino fundamental de 8 anos ou no 2º ano do ensino fundamental de 9 anos – 2.185.404 estudantes. Nesse sentido, a seleção da coorte se assemelha ao trabalho de [Malaguth \(2022\)](#) por selecionar uma coorte que se inicia em um determinado ano escolar, e se difere do trabalho de [Soares, Alves e Fonseca \(2021\)](#) os quais selecionaram uma coorte de estudantes independente do ano escolar que estiveram matriculados em 2007 ou mesmo que tivessem registro administrativo em anos posteriores. Foram excluídos 3.072 estudantes falecidos – 0,14%.

A [Tabela 1](#) destaca o quantitativo de estudantes em situações que não contribuem para a análise desse trabalho e, portanto, foram desconsiderados^{4,5}:

Após a remoção dos estudantes da coorte de estudo, tem-se um total de 2.036.185 estudantes. Salienta-se que pode haver sobreposição entre as inconsistências avaliadas, ou seja, um estudante pode ser avaliado em mais de uma situação, no entanto, como foram aplicadas de forma sequencial, o somatório individual das categorias de inconsistência corresponde ao total de estudantes que foram excluídos da análise.

Por fim, se considera-se a combinação entre rendimento escolar – resultado ao final do ano letivo – e fluxo escolar – resultado entre anos calendário seguidos – tem-se situações que não deveriam ocorrer como, por exemplo, um estudante que é reprovado no ano calendário t e é promovido para o próximo ano escolar no ano calendário $t+1$ ou um estudante que é aprovado e é repetente no ano seguinte. No entanto, tais situações de fato ocorrem, como já alertado por [Klein \(2003\)](#): há 34.334 casos de abandono no ano calendário t e promoção no ano $t+1$; há 86.105 casos de reprovação no ano calendário t e promoção no ano $t+1$; como último exemplo, há 163.683 casos de aprovação no ano calendário t e repetência no ano $t+1$. Além disso, há diversos casos de estudantes com informação de rendimento igual a “Sem informação de rendimento”.

O objetivo geral deste trabalho é caracterizar as trajetórias escolares da coorte de estudantes ingressantes no 2º ano do ensino fundamental no Brasil em 2008 e os diferenciais demográficos e socioeconômicos associados. Apesar de se reconhecer que o conhecimento acerca do rendimento dos estudantes – aprovação, reprovação ou abandono – enriquece o entendimento dessa caracterização, pelos fatores apontados de inconsistência dos dados tal investigação não é realizada. Portanto, a variável que indica o rendimento escolar dos

³ A coorte de nascidos entre 1 de julho de 2000 e 30 de junho de 2001 pode ter sua primeira matrícula apenas nos anos seguintes.

⁴ Em relação à EJA, foi adotado como critério válido de entrada na EJA Ensino fundamental o corte etário de 15 anos completos e para a EJA Ensino Médio o corte etário de 18 anos completos, de acordo com o Parecer CNE/CEB nº 6/2010.

⁵ O fluxo dos alunos não é avaliado para as seguintes etapas (observar valores da variável de etapas auxiliar TP_ETAPA_INDICADOR - ver Anexo 1): Creche (1101), Pré-escola (1102), 4ª série E.M. (1304), Não seriado E.M. (1305), E.P. Concomitante (2001), E.P. Subsequente (2002) e também as etapas da EJA (3200, 3300) para as quais só se avalia a migração para o ensino regular.

Tabela 1 – Coorte inicial, Razões de eliminação de casos inválidos e Coorte final - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental - Brasil

Coorte inicial	2.185.404
Estudantes com matrícula no 1º ano do ensino fundamental de 9 anos, Pré escola ou Creche.	23.389
Estudantes com matrícula na EJA Ensino Fundamental antes de 2016 ou com matrícula na EJA Ensino Médio antes de 2018.	18.775
Estudantes com matrícula em ano escolar superior ao previsto para a idade caso estivessem em fluxo regular, entre 2009 e 2015.	57.691
Estudantes com matrículas nos anos escolares 4ª série do ensino médio, ensino médio não seriado, ensino profissional concomitante, ensino profissional subsequente, por não ter fluxo escolar avaliado.	12.268
Estudantes com matrícula na 2ª série ou 3ª série do ensino médio em 2015 ou na 3ª série do ensino médio em 2016.	996
Estudantes com matrícula no ano calendário t+1 em ano escolar inferior ao matriculado no ano calendário t. Para os anos em que o aluno esteve fora da escola ou esteve matriculado na EJA ensino fundamental ou ensino médio, considerou-se o último ano escolar do ensino regular cursado.	29.531
Estudantes que possuem fluxo escolar "9 - Não se aplica", mas o ano escolar não é EJA, sendo que os outros anos escolares para os quais esse fluxo é válido já foram desconsiderados.	2.595
Estudantes em ano escolar na EJA que possuem fluxo escolar diferente de "6 – Migração para regular" e "9 - Não se aplica" os únicos fluxos válidos.	902
Coorte final	2.036.185

alunos em cada ano calendário não é utilizada.

Ademais, não se excluem os casos de inconsistência entre rendimento escolar e fluxo escolar encontrados, pois considera-se que a informação de fluxo escolar está correta, haja visto que o objetivo de tratamento de inconsistências realizado pelo INEP (2017) teve o objetivo de construir taxas de fluxo acuradas e proporcionar o acompanhamento das trajetórias escolares dos estudantes.

3.2 Variáveis utilizadas

As variáveis sociodemográficas utilizadas para contextualizar as trajetórias escolares dos estudantes são: *sexo*, *cor/raça* e *nível socioeconômico*. A variável *sexo* assume os valores "Feminino" e "Masculino", enquanto a variável *cor/raça* assume os valores "Amarela", "Branca", "Indígena", "Parda", "Preta", "Não declarada". A Tabela 2 apresenta a composição da coorte sob análise por *sexo* e *cor/raça*.

A variável de *cor/raça* foi introduzida no Censo Escolar em 2005 a partir de uma recomendação do INEP para os estabelecimentos de ensino por meio da Portaria nº 156, de 2004, com objetivo de orientar as escolas a inserir esse campo nas fichas de matrícula. No Censo Escolar, a variável *cor/raça* dos estudantes é autodeclarada para estudantes com 16 anos ou mais completos e declarada pelos pais ou responsáveis para estudantes abaixo dessa idade - portanto, heterodeclarada - e também é heteropreenchida, ou seja, o preenchimento no sistema que coleta tal informação é realizado por um terceiro (SENKEVICS; MACHADO; OLIVEIRA, 2016).

Adicionalmente aos campos possíveis de preenchimento no quesito "cor/raça" no Censo Demográfico, foi adicionada a possibilidade "não declarada" no Censo Escolar (SENKEVICS; MACHADO; OLIVEIRA, 2016). A inclusão desse campo se justifica para aqueles que não tem interesse em declarar sua *cor/raça* (INEP, 2015); o índice de "não declaração" de "cor/raça" por estudantes no Censo Escolar em 2009 foi de 47% que se reduziu para 32,7% em 2014, apresentando importante variação dentre os grupos etários, chegando a 25% de não declaração entre os estudantes de 6 a 10 anos (SENKEVICS; MACHADO; OLIVEIRA, 2016).

Os índices relativamente altos de "não declaração" mostram a certa fragilidade no uso da variável "cor/raça". No entanto, procurando investigar o por quê desses altos índices, a partir de procedimento de inspeção in loco nas equipes gestoras do Censo Escolar em alguns municípios brasileiros em 2014, a equipe do INEP constatou que a opção de "não declaração" era utilizada por quem preenchia o sistema em virtude da indisponibilidade da informação, pois não estava presente na ficha de matrícula dos estudantes (INEP, 2015). Dessa forma, vê-se que parcela da "não declaração" advém da ausência da informação e não pelo interesse em não declarar uma "cor/raça".

O índice de "não declarados" na coorte sob análise neste trabalho é de 28,5% (Tabela 2), menor que o índice global do Censo Escolar de 2014. Em virtude do exposto, decide-se por utilizar tal variável por se considerar que o aspecto racial é um importante componente das desigualdades educacionais no Brasil e deve ser levado em consideração, mesmo que sua mensuração ainda careça de melhorias.

A variável de *nível socioeconômico* utilizada neste trabalho corresponde ao proposto por Soares e Alves (2023), que procuram apresentar uma medida de nível socioeconômico (NSE) para a quase totalidade das escolas de educação no Brasil.⁶

O indicador oficial de nível socioeconômico (INSE) (INEP, 2023) abrange menos da metade das escolas brasileiras, estando indisponível principalmente para escolas pequenas, localizadas em áreas rurais, indígenas e quilombolas, e das regiões Norte e Nordeste; também não possuem INSE as escolas privadas que não ofertam ensino médio (SOARES;

⁶ Agradeço à Professora Maria Teresa Gonzaga Alves (FaE/UFMG) por disponibilizar as estimativas de NSE calculadas.

Tabela 2 – Distribuição de estudantes por sexo e cor/raça - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental - Brasil - 2008-2018.

	Total	N	%
		2.035.451	100
Sexo	Masculino	1.039.078	51,05
	Feminino	996.373	48,95
	Não declarada	580.048	28,5
Cor/Raça	Branca	595.874	29,27
	Preta	79.720	3,92
	Parda	763.105	37,49
	Amarela	7.666	0,38
	Indígena	9.038	0,44

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

ALVES, 2023).

O INSE pode ser calculado para cada estudante individualmente e também para uma escola, geralmente tomando-se a média do NSE de seus estudantes. O NSE para os estudantes é construído a partir de suas respostas em questionários contextuais sendo as duas principais fontes o Sistema Nacional de Avaliação da Educação Básica (Saeb) e o Exame Nacional de Ensino Médio (ENEM). As respostas dos estudantes sobre questões como educação dos pais, posse de bens, renda familiar direta e indireta, e contratação de serviço doméstico são utilizadas para se estimar o NSE dos estudantes (SOARES; ALVES, 2023).

O INSE (INEP, 2023), tal qual outros indicadores de nível socioeconômico construídos para estudantes e para escolas, são calculados a partir de questionários contextuais. Como nem todas as escolas participam do Saeb ou seus estudantes realizam o ENEM, não há INSE calculado para elas.

A proposta de Soares e Alves (2023) consistiu, portanto, em utilizar fontes primárias e secundárias para expandir o cálculo do nível socioeconômico para a quase totalidade das escolas. Os procedimentos adotados pelos autores consistiram em: calcular o nível socioeconômico dos estudantes a partir dos questionários contextuais do Saeb e do ENEM a partir das edições realizadas entre 2007 e 2019 e calcular o nível socioeconômico da escola a partir da média do indicador dos estudantes - chamado de nível socioeconômico primário; para as escolas que não tiveram NSE primário calculado, o indicador foi calculado a partir do NSE de alunos que estiveram matriculados em escola com NSE primário calculado, haja visto que a mobilidade interescolar é significativa e o NSE dos estudantes e das escolas é relativamente estável - chamado de NSE secundário. Utilizando-se o NSE primário, o NSE secundário e quatro características das escolas - nível de renda do município em que

a escola se localiza, localização (rural e urbana), dependência administrativa (pública e privada) e o percentual de estudantes com Bolsa Família na escola - foi calculado o NSE para todas as escolas utilizando a metodologia de Teoria de Resposta ao Item.

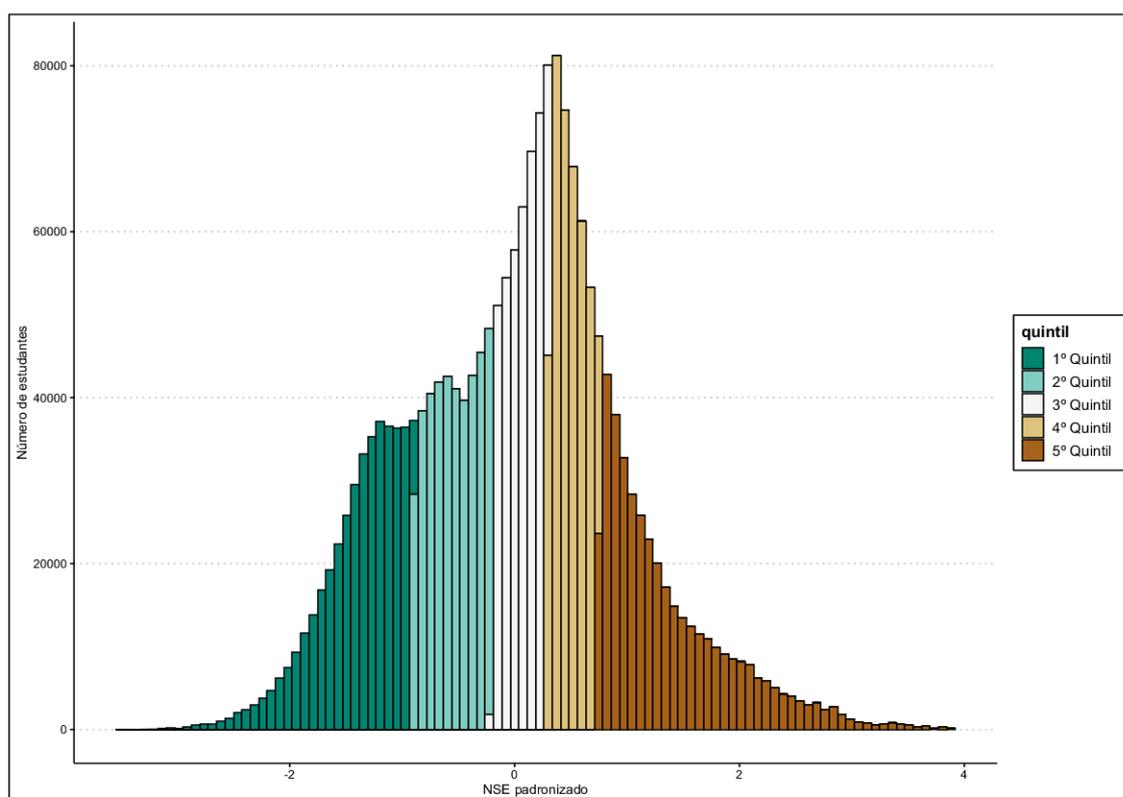
A metodologia empregada por Soares e Alves (2023) para cálculo do NSE resultou em um indicador disponibilizado para 185.616 escolas de educação básica no Brasil. A metodologia foi validada a partir da comparação com INSE (INEP, 2023) e com dados de resultados educacionais, apresentando resultados satisfatórios.

Do total de escolas nas quais os estudantes da coorte neste trabalho estudaram, 3.109 não possuem o NSE calculado por Soares e Alves (2023). Para calcular o NSE dos estudantes, foi tomada a média do NSE das escolas pelas quais o estudante passou ao longo de sua trajetória escolar, desconsiderando-se os anos em que esteve fora da escola. Dado que não é a totalidade das escolas que possuem o indicador de NSE, nem todos estudantes possuem NSE da escola correspondente ao ano em que esteve matriculado nessa escola.

Do total de estudantes da coorte considerada, 99,76% passaram por escolas que tiveram NSE calculado em todos os anos que estiveram na escola; 734 estudantes não possuem nenhuma escola com NSE calculado por Soares e Alves (2023), de forma que esses estudantes foram desconsiderados da base o que resultou em uma coorte de 2.035.451 estudantes considerados para análise. Na Figura 1, que apresenta a distribuição do NSE padronizado da coorte sob análise, vê-se que a distribuição apresenta cauda à direita mais longa, o que reflete estudantes em escolas com NSE relativamente mais alto - chegando à 4 desvios-padrão da média - mas em menor número (SOARES; ALVES, 2023).

Este capítulo evidenciou o processo de tratamento dos dados utilizados neste trabalho. A seleção de estudantes que estejam matriculados no 2º ano do ensino fundamental em 2008 nascidos entre 1 de julho de 2000 e 30 de junho de 2001 implica que as trajetórias são distintas exclusivamente por movimentos de fluxo escolar distintos e não por entrada tardia no sistema de ensino ou retenção em etapas anteriores. Ou seja, dado que os estudantes estão na etapa adequada para sua idade, de que modo a repetência e a evasão condicionam trajetórias distintas?

Figura 1 – Distribuição de NSE padronizado - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018



Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

4 Metodologia

Neste capítulo apresenta-se a metodologia adotada neste trabalho. Na [seção 4.1](#) descreve-se o arcabouço metodológico. Na [seção 4.2](#) descreve-se a construção da tipologia de estados associados às transições escolares. E na [seção 4.3](#) descrevem-se o método adotado para identificar dissimilaridades nas trajetórias, os procedimentos para identificar padrões de trajetórias escolares e as abordagens utilizadas para caracterizar as trajetórias escolares segundo as características demográficas – sexo e cor/raça – e socioeconômicas dos estudantes.

4.1 Análise de sequências

A análise de sequências é um arcabouço metodológico que permite investigar um conceito fundamental de Curso de Vida que é a trajetória, considerando-a em sua totalidade ([ABBOTT, 1995](#)), ([AISENBREY; FASANG, 2010](#)). Seu objetivo principal é identificar padrões em trajetórias, a partir de uma comparação entre pares de trajetórias ou em relação a um padrão esperado ([HALPIN, 2010](#)), ([AISENBREY; FASANG, 2010](#)).

Como arcabouço metodológico, segundo [Aisenbrey e Fasang \(2010\)](#), a análise de sequências possibilita investigar o conjunto de dados de interesse sem nenhum pressuposto sobre o processo gerador de dados subjacente e permite identificar cursos de vida (educacionais) que se diferenciam do padrão esperado.

O elemento fundamental para a análise de sequência é uma sequência, a qual é definida como uma lista ordenada de elementos, na qual está subjacente uma dimensão temporal. Os elementos que compõem a sequência são eventos, os quais estão contidos em uma “cesta” de possíveis eventos – chamado de universo. Exemplificando, os eventos podem ser “casamento”, “ter um filho” e “separação” dos quais se derivam estados, como, respectivamente, “estar casado”, “ser mãe ou pai” e “estar divorciado”, os quais passam a compor uma possível sequência. A representação de uma trajetória por uma sequência de eventos pode ser transformada em uma sequência de estados, e vice-versa ([BILLARI, 2001](#)).

[Abbott \(1995\)](#) apresenta quatro propriedades de sequência que são úteis para construção de um marco analítico:

- a) eventos em uma sequência podem se repetir ou não;
- b) sequências podem ter dependência entre seus estados;
- c) pode haver dependência entre as sequências de indivíduos;

d) sequências podem ser estudadas como variáveis independentes ou dependentes.

Cada trajetória em análise pode ser representada como uma “palavra” ou uma sequência de caracteres que simbolizam os possíveis estados ou eventos considerados. Seguindo o exemplo de [Billari \(2001\)](#), suponha três estados - solteiro (S), divorciado (D) e casado (C) – e que o tempo que um indivíduo permaneceu em cada um deles é de 2 meses, 2 meses e 8 meses, respectivamente. Dada à lógica do exemplo, pode-se representar tal configuração assim:

SSCCCCCCCCDD

De forma análoga, caso se dispusesse do início dos eventos “casamento” e “divórcio” poderia-se reconstruir a trajetória no período anual e estabelecer o tempo que o indivíduo permaneceu em cada estado, supondo que iniciasse no estado solteiro.

A abordagem de análise de sequência tem sido aplicada em diversas áreas do conhecimento: Psicologia, Economia, Arqueologia, Linguística e Ciências Políticas, inclusive na Sociologia e, particularmente, nos estudos de Curso de Vida. No que tange aos estudos demográficos, [Billari \(2001\)](#) explora o fenômeno de transição para a vida adulta em jovens italianos de 20 a 35 anos utilizando seis variáveis que descrevem os eventos representativos dessa transição: ter completado a educação; ter entrado no mercado de trabalho; ter deixado a casa dos pais; ter tido relação sexual; ter entrado em uma união; ter tido um filho. A combinação das sequências permitiu a identificação de quais variáveis são mais representativas na determinação de transição para a vida adulta.

O arcabouço de análise de sequência tem sido também aplicado em estudos qualitativos. [Pereira \(2016\)](#) procura estudar de que forma eventos podem alterar percursos de vida que estariam “pré-determinados”, quando condicionados a condições similares de partida, notadamente o nível socioeconômico. Por meio de entrevistas realizadas com jovens de 18 a 24 anos e aplicando o método de trajetórias narrativas, encontra-se como principal resultado que determinados eventos na vida dos jovens podem restringir ou ampliar as possibilidades futuras em termos de estudo e de mercado de trabalho.

A abordagem de análise de sequências tem utilidade quando se dispõe de dados de registro administrativo, dados em painel ou quando se tem dados retrospectivos de eventos que permitem a criação de uma sequência de estados de interesse ([BILLARI, 2001](#)).

Nesse sentido, a base de dados longitudinal do censo escolar ([INEP, 2017](#)) constitui uma fonte para o estudo das trajetórias escolares e aplicação da abordagem de análise de sequências. Considera-se que a aplicação dessa metodologia contribui para caracterizar as trajetórias escolares da coorte de estudantes ingressantes no 2º ano do ensino fundamental no Brasil em 2008 e os diferenciais demográficos e socioeconômicos associados.

4.2 Sequência de estados e os efeitos *timing*, *duration* e *sequencing*

Análise de sequência tem como objetivo principal identificar padrões nas trajetórias consideradas. Para atingir tal objetivo, podem ser utilizadas duas abordagens: algébrica, pela qual se reduz cada sequência a uma forma mais simples; e métrica, a partir da qual se desenvolve uma medida de distância entre um par de sequências (ABBOTT, 1995). Esta última envolve uma decisão a respeito de qual medida de distância se deve utilizar tendo como critério algum fator que seja relevante para as ciências sociais, quando aplicado nesse campo (LIAO et al., 2022).

As sequências de estados apresentam informações que auxiliam a identificação do critério relevante para a escolha da métrica adequada (STUDER; RITSCHARD, 2016). Essas informações dizem respeito, por exemplo, a ocorrência de determinado evento que configura em um estado específico, quanto tempo se permaneceu em um estado e a ordem em que os eventos ocorreram configurando diferentes sequências de estados.

São cinco os aspectos de sequências de estados destacados por Studer e Ritschard (2016): os distintos estados experimentados por cada indivíduo em sua trajetória; tempo total experimentado em cada estado durante toda a trajetória, que pode ser utilizado como referência para tempo de exposição, aspecto que, no entanto, não informa como o estado foi experimentado ao longo da trajetória; idade ou data em que ocorre o evento de interesse ou o estado de interesse é experimentado; tempo consecutivo experimentado em um mesmo estado; ordem ou sequência em que os distintos estados são experimentados.

Esses aspectos são interrelacionados e se complementam, de modo que podem ser resumidos em três aspectos principais segundo Liao et al. (2022): tempo consecutivo experimentado em cada estado (*duration*); quando se dá a ocorrência de cada estado (*timing*) e ordem na qual os distintos estados ocorrem (*sequencing*).

No ensino seriado do sistema educacional brasileiro, os estudantes percorrem as etapas de escolarização do ensino regular de forma anualizada. Em cada ano-calendário, portanto, um estudante está matriculado em uma etapa de ensino caso esteja na escola. Portanto, a definição dos estados adotados neste trabalho decorre da etapa de ensino que cada estudante está matriculado ou, se não o está, o estudante é considerado como "fora da escola".

No processo de construção da base longitudinal a partir dos Censos Escolares (INEP, 2017) as etapas de ensino foram simplificadas e agregadas: as etapas do ensino fundamental de 8 anos foram compatibilizadas com as etapas do ensino fundamental de 9 anos; as categorias normal/magistério e curso técnico integrado foram agrupadas nas etapas do ensino médio. Neste trabalho, os estudantes matriculados na Educação de Jovens e Adultos foram considerados como "fora da escola", em virtude da evasão do sistema de ensino regular, e os estudantes matriculados em etapas de ensino que não tem fluxo escolar

avaliado foram desconsiderados, tal como descrito no [Capítulo 3](#).

Os estudantes que evadiram ou migraram para a Educação de Jovens e Adultos foram considerados permanentemente evadidos e para todos os anos após o evento "evasão" ou "migração para EJA" foram considerados "fora da escola". Essa é uma importante limitação que implica em uma superestimação do número de estudantes fora da escola ao não considerar possíveis re-entradas no sistema de ensino.

À cada estado considerado, portanto, foi atribuída uma codificação que identifica cada estado por uma letra do alfabeto, a qual é livre e não interfere na análise dos resultados. A [Tabela 3](#) apresenta os estados e a codificação utilizada.

Tabela 3 – Estados considerados na análise das trajetórias escolares.

A - 2º ano E.F.	G - 8º ano E.F.
B - 3º ano E.F.	H - 9º ano E.F.
C - 4º ano E.F.	I - 1ª série E.M.
D - 5º ano E.F.	J - 2ª série E.M.
E - 6º ano E.F.	K - 3ª série E.M.
F - 7º ano E.F.	L - Fora da escola

Nota: E.F. – Ensino Fundamental, E.M. – Ensino Médio.

As sequências que caracterizam as trajetórias escolares podem ser representadas por um conjunto de símbolos ordenados. A [Tabela 4](#) apresenta exemplos hipotéticos de trajetórias a partir da codificação apresentada na [Tabela 3](#), que serão úteis para explorar os aspectos da sequência de estados destacados por [Studer e Ritschard \(2016\)](#), seguindo o exemplo de [Liao et al. \(2022\)](#).

Tabela 4 – Exemplos de trajetórias escolares hipotéticas

1 - LAAAAAAAAA	Estudante fora da escola com repetência contínua
2 - AAAAAAAAAAL	Estudante com repetência contínua seguida de evasão
3 - ALLLLLLLLL	Estudante com evasão no 2º Ano E.F.
4 - LALALALAL	Estudante com evasão e retorno à escola
5 - ALALALALALA	Estudante com evasão e retorno à escola seguidos

Nota: A – 2º ano Ensino Fundamental, L – Fora da escola.

As sequências 1 e 2 apresentam os efeitos *timing* e *duration* similares, pois estão na maior parte do tempo na mesma etapa de ensino - 2º ano do ensino fundamental, no entanto, elas se diferem quanto à ordem dos distintos estados experimentados - *sequencing*. Na sequência 2, ocorre uma evasão escolar enquanto na sequência 1 ocorre um retorno à escola. As sequências 2 e 3 apresentam similaridade quanto ao efeito *sequencing* haja visto que ambas apresentam evasão escolar; essas duas sequências diferem em relação aos efeitos *timing* e *duration* pois na maior parte do tempo não estão no mesmo estado. Por fim, as sequências 4 e 5 apresentam uma alternância entre evasão e retorno à escola, de modo que tem o efeito *sequencing* similar e, da mesma forma, o efeito *duration*, pois permanecem

em cada estado apenas 1 ano; já em relação ao efeito *timing* as sequências diferem, pois não estão no mesmo estado em nenhum ano.

As trajetórias apresentadas na [Tabela 4](#) são consideradas hipotéticas neste trabalho em virtude das decisões metodológicas deste trabalho. A estratégia de seleção da coorte adotada bem como a desconsideração de casos inconsistentes descrita no [Capítulo 3](#) implica em um total de 12.118 trajetórias escolares distintas, quantidade menor que encontrado por [Soares, Alves e Fonseca \(2021\)](#) – 80.009.

A modelagem dos dados utilizada neste trabalho exclui diversos casos de inconsistência que, em conjunto com a escolha de evasão em definitivo e consideração de migração para EJA como evasão do ensino regular, configura as trajetórias a partir de estados que são ordenados. Ou seja, em todas as trajetórias consideradas o 2º ano do Ensino Fundamental precede o 3º ano, que precede o 4º ano e, assim, sucessivamente até a 3ª série do Ensino Médio. Além disso, o estado "Fora da escola" não é seguido por nenhum outro estado além dele mesmo. A [Tabela 5](#) apresenta exemplos de trajetórias escolares possíveis neste trabalho.

Tabela 5 – Exemplos de trajetórias escolares

1 - ABCDEFGHIJK	Estudante continuamente promovido
2 - AABCDEFGHIJ	Estudante com uma repetência no 2º ano E.F.
3 - ABCDEFGHIJJ	Estudante com uma repetência na 2ª série E.M.
4 - AABBCCDDEEF	Estudante com duas repetências em cada etapa de ensino até o 5º ano E.F.
5 - ABBCCDDEEFF	Estudante com duas repetências em cada etapa de ensino até o 6º ano E.F.
6 - AAAAAAAAAAL	Estudante continuamente repetente seguido de evasão
7 - ALLLLLLLLLLL	Estudante evadido no 2º ano E.F.

A sequência 1 apresenta o caminho esperado de um estudante no sistema educacional brasileiro: continuamente promovido chegando à 3ª série do ensino médio após 12 anos de escolarização sem nenhum intercorrência - como repetência ou evasão. As sequências 2 e 3 se diferem da sequência 1 pois apresentam uma repetência - no 2º ano do ensino fundamental e na 2ª série do ensino médio, respectivamente. Entre essas sequências - 1, 2 e 3 - observa-se que o efeito *duration* é similar, diferenciando-se apenas ao tempo de permanência nas etapas de ensino onde houve repetência, mas o efeito *timing* é diferente: as sequências 1 e 3 experimentam o mesmo estado na maior parte do tempo, enquanto a sequência 2 não. Em relação ao efeito *sequencing* não há diferenciação entre as sequências, haja visto que experimentam os distintos estados na mesma ordem.

As sequências 4 e 5 apresentam situações em que há duas repetências no 3º, 4º e 5º ano do Ensino Fundamental, sendo que na sequência 4 há uma repetência também no 2º ano e na sequência 5 há uma repetência no 6º ano. Essas duas sequências apresentam o efeito *duration* similar, pois permanecem em geral o mesmo tempo consecutivo em cada um dos estados; já em relação ao efeito *timing* a similaridade diminui pois em alguns anos os estudantes cursam etapas de ensino distintas. Por fim, em relação ao efeito *sequencing*

não há diferenciação, pois experimentam os distintos estados na mesma ordem.

Por último, as sequências 6 e 7 não se diferenciam em relação ao efeito *sequencing*. Em relação aos efeitos *timing* e *duration* há diferenças marcantes: na sequência 6 o estudante permanece durante 10 anos no 2º ano do ensino fundamental e 1 ano fora da escola, enquanto na sequência 7 o estudante permanece apenas 1 ano no 2º ano do ensino fundamental e 10 anos fora da escola.

Portanto, vê-se que as inconsistências identificadas e desconsideradas da análise de dados apresentadas no [Capítulo 3](#) e a definição de evasão permanente evitam diferenciações entre trajetórias em relação ao efeito *sequencing*, haja visto que a sequência de estados quando o estudante se encontra na escola é sempre ordenada e o estado "fora da escola" não é seguido por nenhum outro estado além dele mesmo. Dessa forma, a ordem dos estados entre duas sequências pode se diferir em consequência de quais e de quando os estados são experimentados (*timing*) e quanto tempo se permanece consecutivamente em cada estado (*duration*).

A análise de dados neste trabalho foi conduzida no *software* R ([R Core Team, 2021](#)) e o pacote *TraMineR* ([GABADINHO et al., 2011](#)) foi utilizado para análise das trajetórias escolares. A representação das trajetórias foi realizada a partir do formato STS (*'STates-sequence'*), no qual a trajetória de cada estudante é representado em uma linha do conjunto de dados e as colunas representam o estado em que se encontra entre 2008 e 2018, anualmente.

4.3 Análise de sequência e *Optimal matching*

Para se caracterizar as trajetórias educacionais da coorte de estudantes ingressantes no 2º ano do ensino fundamental no Brasil em 2008 e investigar os diferenciais demográficos e socioeconômicos associados adota-se o método *Optimal matching* proposto por [Abbott e Hrycak \(1990\)](#) e [Abbott \(1995\)](#), o qual permite a mensuração de semelhança entre sequências e a identificação de padrões em um conjunto de sequências.

Adaptado da biologia para identificação de semelhança em cadeias de DNA, o método *Optimal matching* ([ABBOTT; HRYCAK, 1990](#)) gera uma medida da similaridade entre um par de sequências e essa medida pode então ser utilizada para identificar padrões no conjunto de dados. A medida é gerada a partir da menor quantidade de operações – inserção, deleção e substituição – necessárias para se transformar uma sequência – representada por uma cadeia de caracteres – em outra.

De forma geral, o método pode ser aplicado em três etapas ([AISENBREY; FASANG, 2010](#)): especificação do espaço de estados e dos custos de transformação; aplicação do algoritmo de *Optimal matching* para produzir pares de distâncias entre as sequências;

agrupamento das sequências.

O espaço de estados apresentado na [Tabela 3](#) é constituído pelas etapas de ensino e o estado "Fora da escola". As operações de transformação que permitem se obter uma sequência de caracteres a partir de outra são inserção, deleção e substituição. No campo da biologia essas operações estão associadas a um processo de mutação entre cadeias de DNA; no entanto, nas ciências sociais não há de se falar em uma transformação de uma trajetória individual a partir de uma mutação ([STUDER; RITSCHARD, 2016](#)).

No campo das ciências sociais, o significado dessas operações, portanto, é outro: as operações de substituição estão interessadas se o mesmo estado ocorre ao mesmo tempo em duas sequências, enfatizando o aspecto de *timing*; já os custos de inserção/deleção estão interessados se os mesmos estados ocorrem nas duas sequências, independente de quando, enfatizando a ocorrência de estados e, portanto, o aspecto de *duration* ([AISENBREY; FASANG, 2010](#)).

Os aspectos de *timing* e *duration* são os dois aspectos a serem evidenciados nas trajetórias escolares consideradas neste trabalho e, portanto, o *Optimal matching* é um método adequado para se calcular distâncias entre as sequências representadas por essas trajetórias ([STUDER; RITSCHARD, 2016](#)). Particularmente, adota-se a variação do método conhecida como OM 1&2 na qual as operações de inserção e deleção tem custo unitário enquanto a operação de substituição tem custo igual a dois.

Supondo as sequências 4 e 5 apresentadas na [Tabela 5](#) a quantidade mínima de operações necessárias para se transformar uma sequência na outra são duas: deleção do primeiro estado 'A' na primeira sequência e inserção de um estado 'F' ao final da mesma sequência; de outro modo, poderia-se inserir um estado 'A' no início da segunda sequência seguido de uma deleção do estado 'F' ao final. A soma dos custos dessas operações resulta em dois - uma inserção e uma deleção:

$$AABBCCDDEEF$$

$$ABBCCDDEEFF$$

Supondo as sequências 6 e 7 apresentadas na [Tabela 5](#) a quantidade mínima de operações necessárias para se transformar uma sequência na outra são nove: substituição do estado 'A' por 'L'; de outro modo, poderia-se se inserir o estado 'A' nove vezes e se deletar o estado 'L' também nove vezes na segunda sequência para obter a primeira. A soma dos custos das operações necessárias resulta em dezoito - seja por meio das operações de substituição ou das operações de inserção/deleção:

$$AAAAAAAAAAL$$

$$ALLLLLLLLLLL$$

A partir dos exemplos de trajetórias escolares da Tabela 5 apresenta-se na Figura 2 uma matriz de distância entre todos os pares de sequência. A matriz resultante é simétrica com diagonal nula, pois a distância entre sequências independe de qual se aplica as operações e a distância de uma sequências para ela mesma é zero.

Figura 2 – Matriz de distância entre pares de sequências de trajetórias escolares

	ABCDEFGHIJK	AABCDEFGHIJ	ABCDEFGHIJ	AABBCCDDEEF	ABBCCDDEEFF	AAAAAAAAAAL	ALLLLLLLLL
ABCDEFGHIJK	0	2	2	10	10	20	20
AABCDEFGHIJ	2	0	2	8	10	18	20
ABCDEFGHIJ	2	2	0	10	10	20	20
AABBCCDDEEF	10	8	10	0	2	18	20
ABBCCDDEEFF	10	10	10	2	0	20	20
AAAAAAAAAAL	20	18	20	18	20	0	18
ALLLLLLLLL	20	20	20	20	20	18	0

A última etapa da aplicação do método de análise de sequências constitui o agrupamento de sequências a partir de uma medida de dissimilaridade entre elas (AISENBREY; FASANG, 2010). Essa medida de dissimilaridade é utilizada para agrupar sequências em grupos que apresentam alta homogeneidade interna e alta heterogeneidade entre si (STUDER, 2013).

Um total de 2.035.451 estudantes compõem a coorte selecionada neste trabalho. A definição de uma distância entre a trajetória escolar, representada por uma sequência de caracteres, entre cada estudante resulta em uma matriz quadrada com $2.035.451^2$ elementos. Para armazenar tal matriz, seriam necessários aproximadamente 33 Tb de memória computacional.

Adicionalmente a essa inviabilidade operacional, a função *seqdist* do pacote *TraMineR* (GABADINHO et al., 2011) armazena a matriz de distância entre as sequências a partir de uma indexação de 32bit, o que limita o número máximo de sequências a ser utilizada nessa função em 46.341, de forma a não exceder o número de elementos possíveis com a arquitetura de 32 bit ($2^{32} - 1$).¹

Portanto, para a análise de agrupamentos foi selecionado a maior amostra possível - 46.341 estudantes - aleatoriamente, a qual representa 2,27% do total de estudantes e apresenta 2.499 sequências de trajetórias distintas, o que representa 20,6% do total de sequências de trajetórias da coorte analisada.

A Tabela 6 apresenta a composição da amostra da coorte sob análise por *sexo* e *cor/raça*, a qual se aproxima com pequenos desvios da distribuição da coorte apresentada na Tabela 2. A Figura 3 apresenta a distribuição do NSE padronizado da amostra da coorte sob análise.

¹ <https://r-forge.r-project.org/tracker/index.php?func=detail&aid=6512&group_id=743&atid=2975> (Acesso em 21 de junho de 2023).

Tabela 6 – Distribuição de estudantes por sexo e cor/raça - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental - Brasil - 2008-2018.

Total		N	%
		46.341	100
Sexo	Masculino	23.605	50,94
	Feminino	22.736	49,06
	Não declarada	13.336	28,78
Cor/Raça	Branca	13.535	29,21
	Preta	1.799	3,88
	Parda	17.288	37,31
	Amarela	179	0,39
	Indígena	204	0,44

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Existem diversos algoritmos para agrupar as sequências a partir da matriz de dissimilaridade, de modo que foi escolhido o algoritmo de *agrupamento aglomerativo hierárquico*, pois é recomendado quando não se sabe a priori o número de agrupamentos esperado (STUDER, 2013). Esse algoritmo opera construindo os agrupamentos a partir das observações individuais, partindo, portanto, do total de observações até se chegar a um agrupamento que inclui todas as observações, construindo uma árvore de agrupamento; o processo de agrupamento permite selecionar o número de agrupamentos de interesse e tal processo está descrito no Apêndice A.

Para operacionalizar tal algoritmo foi utilizada a função *agnes* do pacote *cluster* (MAECHLER et al., 2022), com um custo de tempo de 3,74 dias a partir de uma matriz de dissimilaridade que ocupa 17,2Gb de memória computacional.

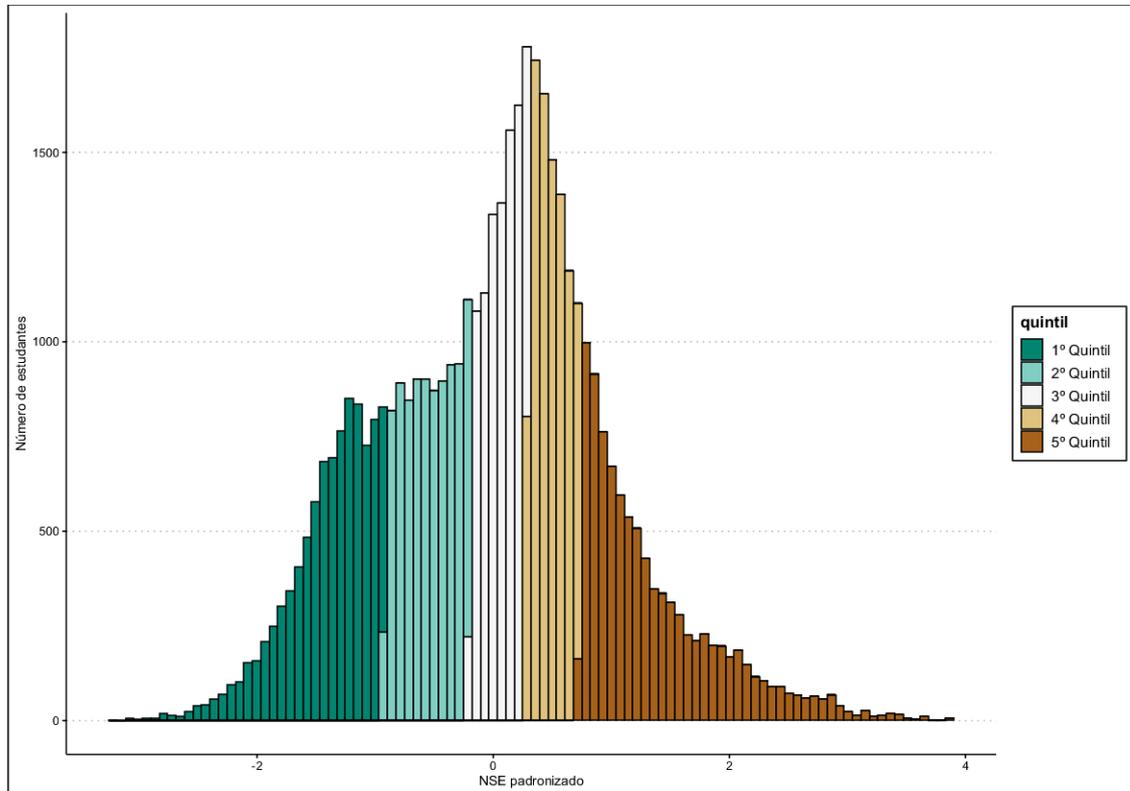
A partir dos padrões identificados na análise de agrupamentos, foram utilizadas estatísticas descritivas a fim de computar a distribuição de estudantes por sexo, cor/raça e NSE por padrão identificado. Com o objetivo de testar se há uma associação entre os pares de variáveis categóricas - sexo e padrão, cor/raça e padrão e quintil de NSE e padrão - foi aplicado um teste Qui-Quadrado de independência entre proporções, sendo

H_0 : hipótese nula – não há associação entre as variáveis

H_a : hipótese alternativa – há associação entre as variáveis

Para aplicar o teste estatístico, utiliza-se uma tabela de contingência na qual nas colunas se apresentam os valores de uma variável categórica e nas linhas os valores de outra variável categórica. Os valores das células dessa tabela representam o número de

Figura 3 – Distribuição de NSE padronizado - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018



Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

casos observados na amostra e esses valores são comparados com o valor de casos esperados caso as variáveis fossem independentes.

A estatística do teste Qui-Quadrado é calculada pela seguinte fórmula:

$$\chi^{2*} = \sum_{i=1}^{lc} \frac{(O_i - E_i)^2}{E_i}$$

l : número de linhas

c : número de colunas

O_i : número de casos observados

E_i : número de casos esperados

A estatística do teste é comparada com uma distribuição Qui-Quadrado com graus de liberdade igual a $(l - 1) * (c - 1)$ e utiliza-se a estratégia do p-valor para se verificar se

a hipótese nula pode ser rejeitada. Realizou-se o teste Qui-Quadrado a partir da função *chi.test* do pacote *stats* (R Core Team, 2021).

A análise de dados neste trabalho foi operacionalizada a partir do software R (R Core Team, 2021), utilizando o pacote que viabiliza análise de sequências *TraMineR* (GABADINHO et al., 2011), os pacotes de manipulação de dados *dplyr* (WICKHAM et al., 2023), *tidyr* (WICKHAM; VAUGHAN; GIRLICH, 2023), *lubridate* (GROLEMUND; WICKHAM, 2011) e *stringr* (WICKHAM, 2022), o pacote de visualização de dados *ggplot2* (WICKHAM, 2016), os pacotes para clusterização de sequências *cluster* (MAECHLER et al., 2022) e *WeightedCluster* (STUDER, 2013) e o pacote para visualização de sequências *ggseqplot* (RAAB, 2022). Os scripts utilizados constam no Apêndice B.

5 Resultados

Este capítulo apresenta os resultados das análises das trajetórias escolares da coorte de estudantes nascidos entre 1 de julho de 2000 e 30 de junho de 2001 matriculados no 2º ano do ensino fundamental em 2008. Inicia-se com a descrição das trajetórias da coorte, destacando-se a distribuição dos estudantes por estado em cada ano entre 2008 e 2018 e as 10 trajetórias mais comuns dessa coorte. Em seguida, descrevem-se as trajetórias da coorte considerando as características demográficas - sexo e cor/raça - e socioeconômicas.

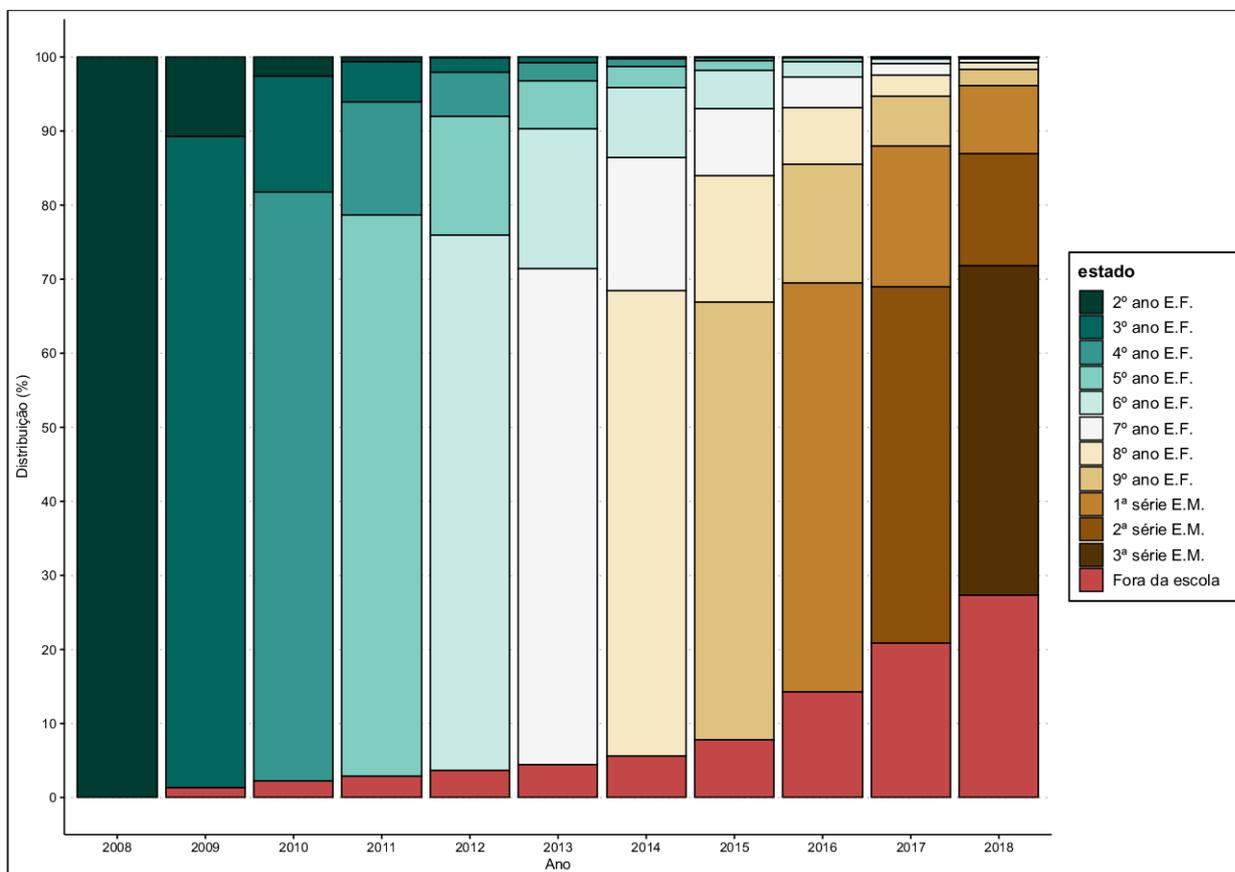
Na [seção 5.1](#), descrevem-se os padrões de trajetórias escolares identificados destacando a distribuição dos estudantes por estado em cada ano entre 2008 e 2018, o tempo médio em cada estado, que consiste no número médio de vezes que cada estado é observado em uma sequência ([GABADINHO et al., 2011](#)), e as 10 trajetórias mais comuns. Na [seção 5.2](#), os padrões de trajetórias escolares são descritos considerando-se as características demográficas - sexo e cor/raça - e socioeconômicas dos estudantes e são apresentados os resultados do teste Qui-Quadrado de independência entre proporções.

Em 2008, em virtude do critério de seleção da coorte em análise, todos os estudantes estão matriculados no 2º ano do ensino fundamental, já em 2018, após 11 anos, é esperado que os estudantes estivessem matriculados na 3ª série do ensino médio, no entanto, menos da metade dos estudantes - 44,5% - estavam matriculados nessa etapa de ensino ([Figura 4](#)), resultado semelhante foi identificado por [Malaguth \(2022\)](#), que identificou 50,5% dos estudantes na 3ª série do ensino médio, considerando apenas o estado de Minas Gerais. A evasão e a repetência contribuíram para que os estudantes não estivessem matriculados na última etapa de ensino da educação básica no tempo esperado.

O percentual de estudantes fora da escola é crescente pela metodologia adotada que considera a evasão escolar como permanente e a migração para EJA também como evasão permanente. Ressalta-se que essa é uma importante limitação deste trabalho, que não considera todas as entradas e as saídas dos estudantes na escola no período entre 2008 e 2018 e, portanto, superestima o número de estudantes fora da escola. Considera-se que tal limitação deve ser tratada em próximos trabalhos dentro da linha de pesquisa de caracterização das trajetórias escolares.

Destaca-se o aumento entre os anos de 2015 e 2016 de estudantes que estavam fora da escola passando de 7,8% para 14,2% e, nos anos seguintes, essa proporção continua aumentando, passando para 20,9% em 2017 e chegando a 27,3% em 2018 ([Figura 4](#)), valor próximo ao identificado por [Malaguth \(2022\)](#) que identificou 22,4% dos estudantes fora da escola. Tal resultado mostra que o fator idade pode ser um componente importante para a saída da escola, combinada com a inadequação da idade para a etapa que o estudante está

Figura 4 – Distribuição relativa dos estados por ano - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018



Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

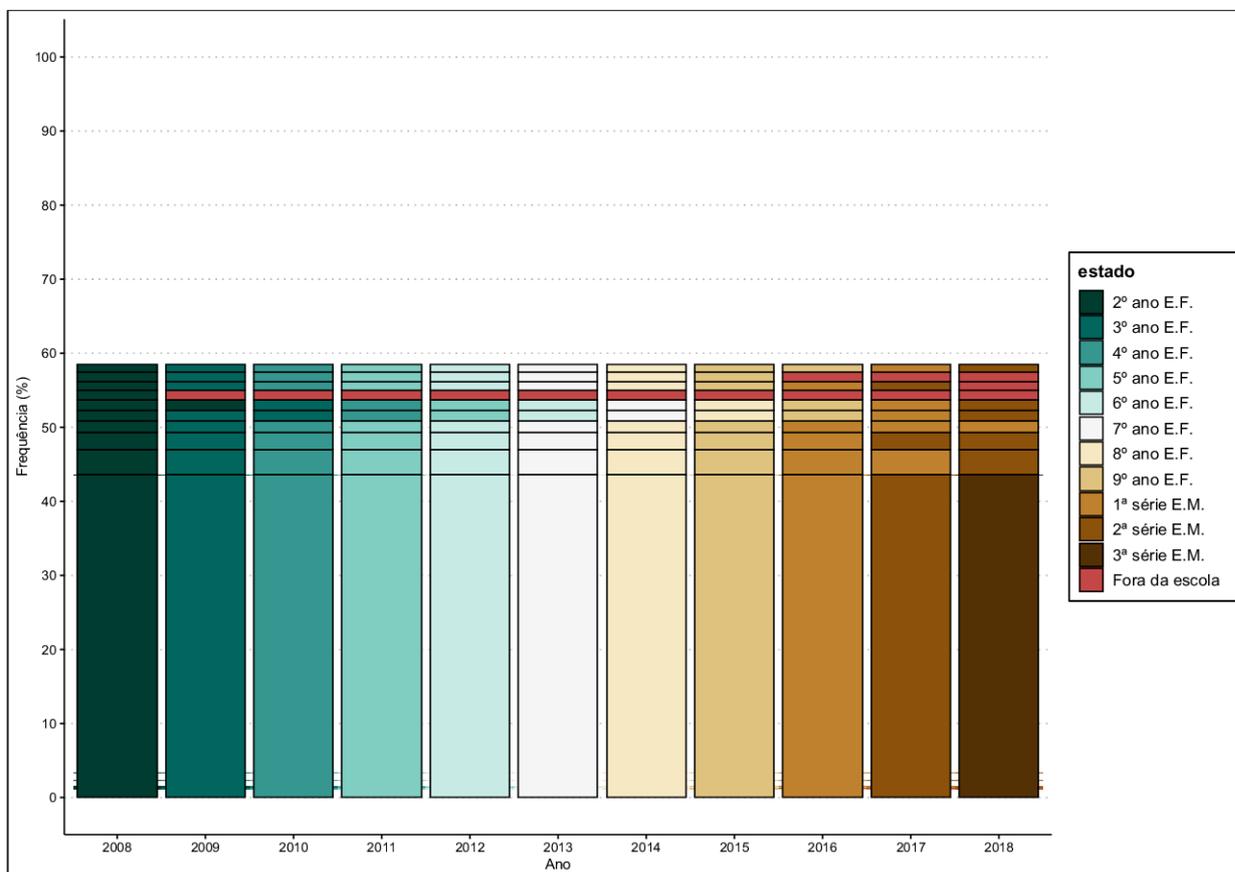
matriculado (RODERICK, 1994); (MALAGUTH, 2022).

Em relação à repetência, observa-se que se inicia no 2º ano do ensino fundamental, com 10,7% dos estudantes repetindo essa etapa de ensino em 2009, percentual que representa a taxa de repetência no biênio 2008/2009 para essa etapa de ensino¹ (Figura 4). Nos anos seguintes, observa-se uma participação crescente dos estudantes que estão matriculados em "atraso escolar", ou seja, em etapas anteriores ao que era esperado, caso seguissem uma trajetória regular, atingindo um pico de 33,1% em 2015 (Figura 4).

Considerando a sequência de estados que são experimentados pelos estudantes, ou

¹ A taxa oficial calculada para o INEP para essa etapa é de 16%. Enquanto neste trabalho a taxa considera apenas estudantes que ingressaram no 2º ano do ensino fundamental na idade esperada, a taxa calculada pelo INEP considera todos os estudantes matriculados no 2º ano do ensino fundamental incluindo, portanto, estudantes que não entraram no sistema de ensino na idade adequada ou também que já são repetentes. Taxa de repetência. Indicadores disponíveis em: <<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/indicadoreseducacionais>>. Acesso em: 31 mai. 2022.

Figura 5 – 10 sequências mais comuns - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018

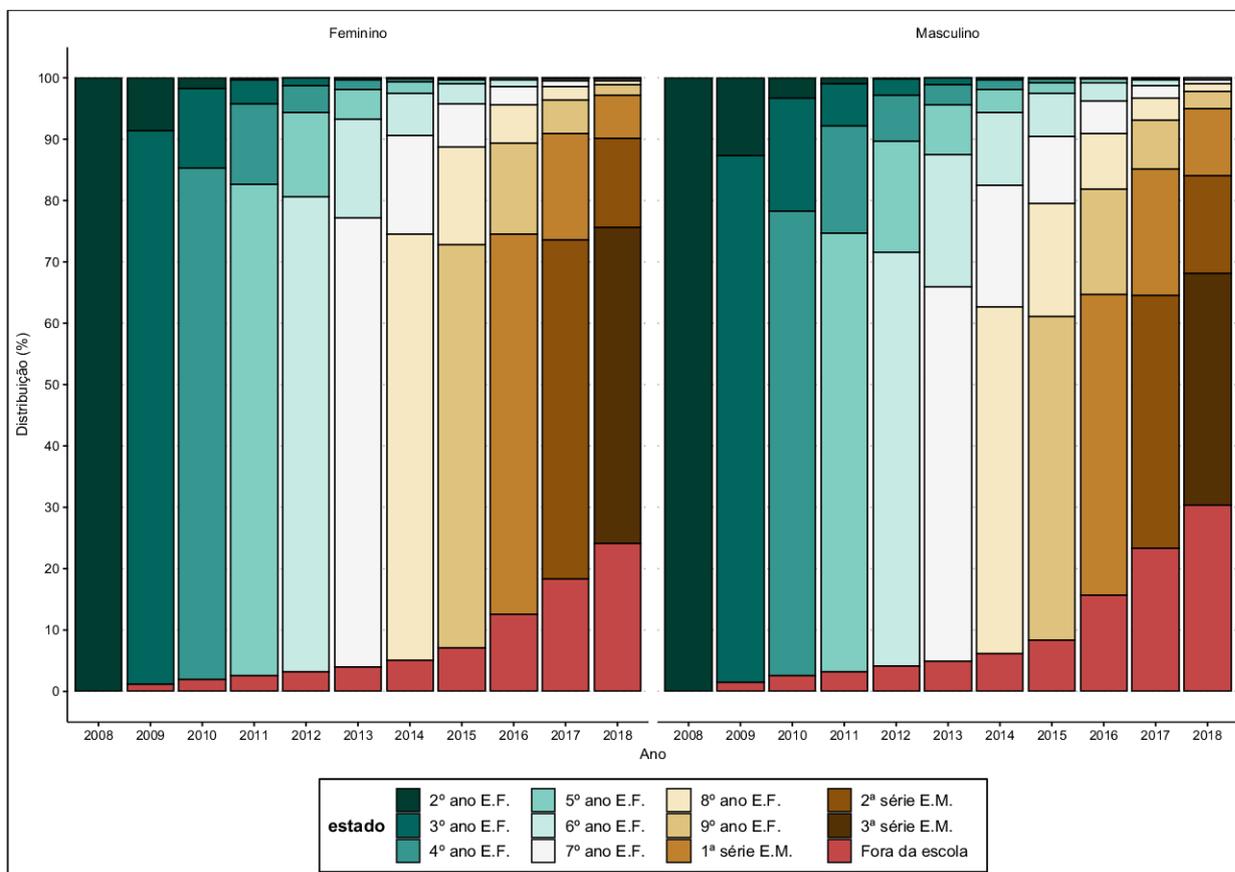


Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

seja, como percorreram o sistema de ensino, a Figura 5 apresenta as 10 sequências mais comuns e o quanto cada uma representa do total de estudantes. Como já observado, 44,5% dos estudantes estavam matriculados na 3ª série do ensino médio (Figura 4), sendo que 43,57% dos estudantes atingiram essa etapa de ensino em 2018 percorrendo uma trajetória regular desde 2008². A segunda e a terceira trajetórias mais comuns representam em conjunto 5,73% dos estudantes e consistem-se de estudantes que foram reprovados na 1ª série ou na 2ª série do ensino médio. A quarta trajetória mais comum é aquela dos estudantes que cursaram a 1ª série do ensino médio três vezes, representando 1,58% dos estudantes. As 10 sequências mais comuns contemplam 58,51% das trajetórias da coorte, sendo que, exceto a trajetória regular, representam 14,94% do total de trajetórias, de modo que há uma heterogeneidade substancial nas trajetórias percorridas pelos estudantes dessa coorte.

² Foi identificado que alguns estudantes progrediram mais de um etapa em um biênio, portanto, o percentual de estudantes que estão na 3ª série do ensino médio (44,5%) é superior ao percentual de estudantes que percorreram todas as etapas de ensino nos 11 anos considerados (43,57%)

Figura 6 – Distribuição relativa dos estados por ano e sexo - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018



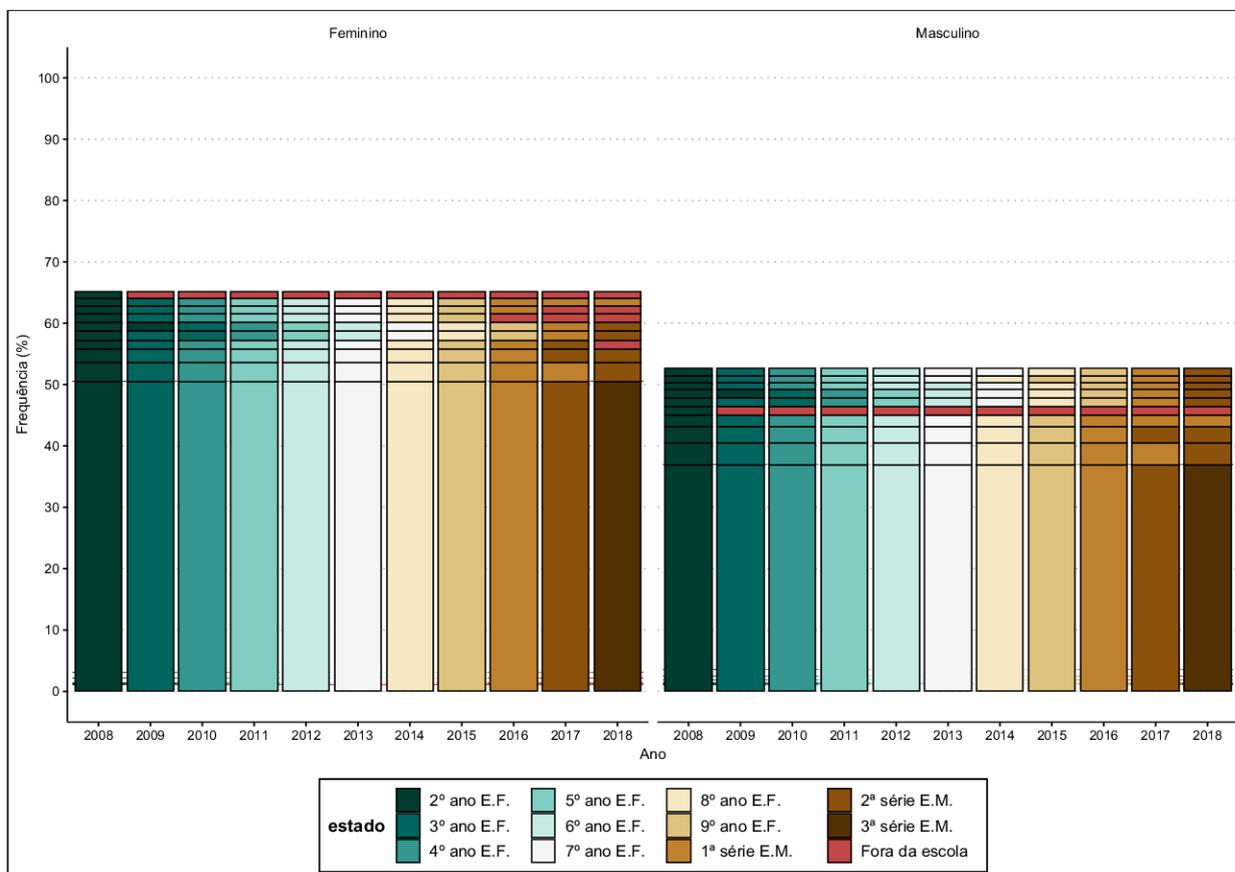
Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Todos os estudantes da coorte em análise neste trabalho estão matriculados na mesma etapa de ensino em 2008. Como já visto, apenas uma parcela desses estudantes chega à 3ª série do ensino médio em 2018, no tempo esperado (Figura 4). A distribuição dos estudantes que chegaram a essa etapa de ensino, bem como dos estudantes que evadiram da escola, mostrou-se desigual entre os grupos sociais, seja por sexo, por cor/raça ou por nível socioeconômico.

Em relação à variável sexo, 51,5% das estudantes do sexo feminino chegaram à 3ª série do ensino médio em 2018 e, nesse mesmo ano, 24,1% estavam fora da escola (Figura 6). Por outro lado, 37,7% dos estudantes do sexo masculino chegaram à 3ª série do ensino médio em 2018 - 13,8% p.p. a menos - e 30,4% estavam fora da escola no mesmo ano (Figura 6).

As 10 trajetórias mais comuns percorridas pelas estudantes do sexo feminino

Figura 7 – 10 seqüências mais comuns por sexo - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018



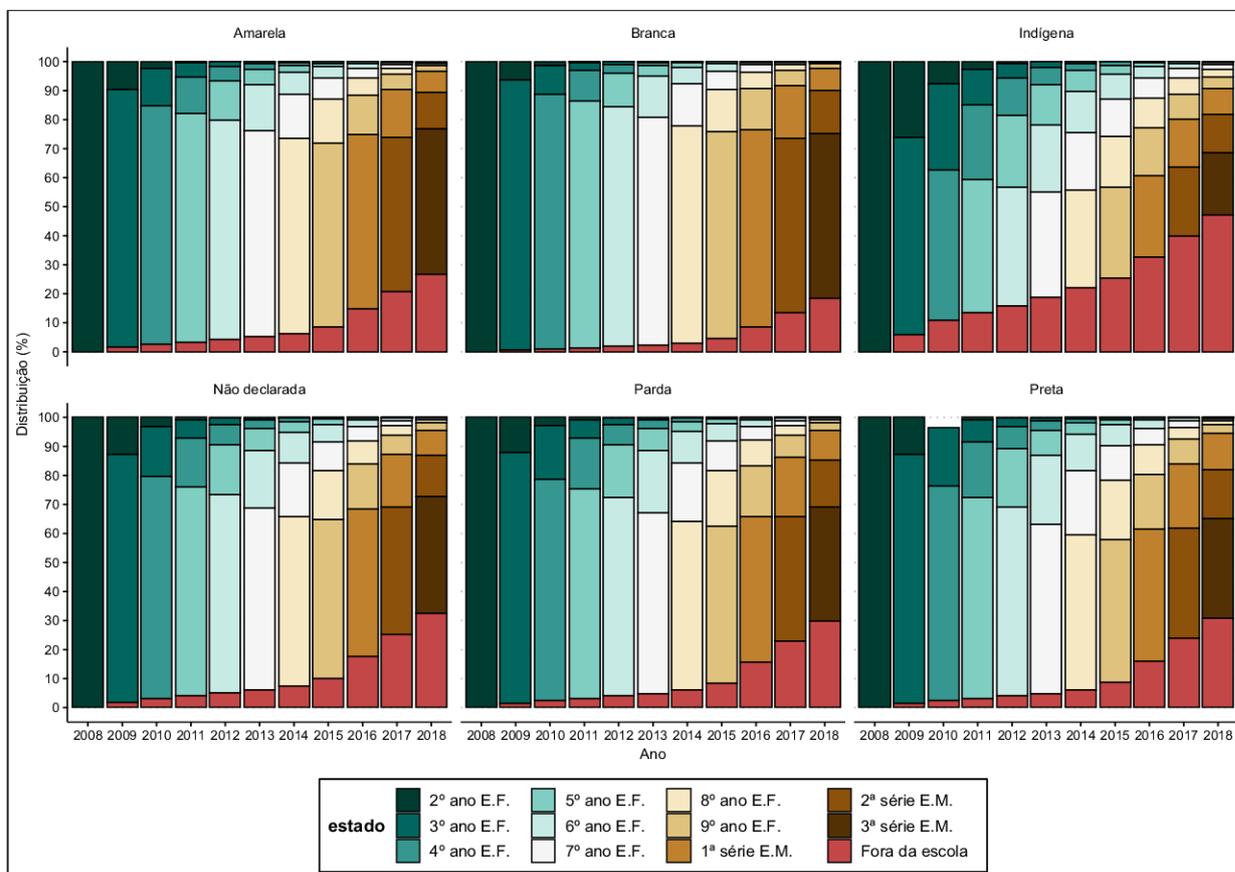
Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

representaram 65,2% do total de trajetórias, sendo que a trajetória regular representou 50,2% (Figura 7). Comparativamente, as 10 trajetórias mais presentes entre os estudantes do sexo masculino representaram 52,62% do total de trajetórias, sendo que a trajetória regular representou 36,91% (Figura 7).

Em relação à variável cor/raça, observa-se os seguintes percentuais de estudantes que chegaram à 3ª série do ensino médio em 2018: 39,4% de cor/raça parda; 34,3% dos estudantes de cor/raça preta; 56,6% dos estudantes de cor/raça branca; 50,2% dos estudantes de cor/raça amarela; 21,3% dos estudantes de cor/raça indígena; e 40,4% dos estudantes de cor/raça não declarada (Figura 8). Observa-se uma disparidade racial substantiva nos resultados educacionais dos estudantes analisados.

Considerando os estudantes que estão fora da escola em 2018, tem-se os seguintes resultados: 29,8% de cor/raça parda; 30,8% dos estudantes de cor/raça preta; 18,5% dos

Figura 8 – Distribuição relativa dos estados por ano e cor/raça - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018



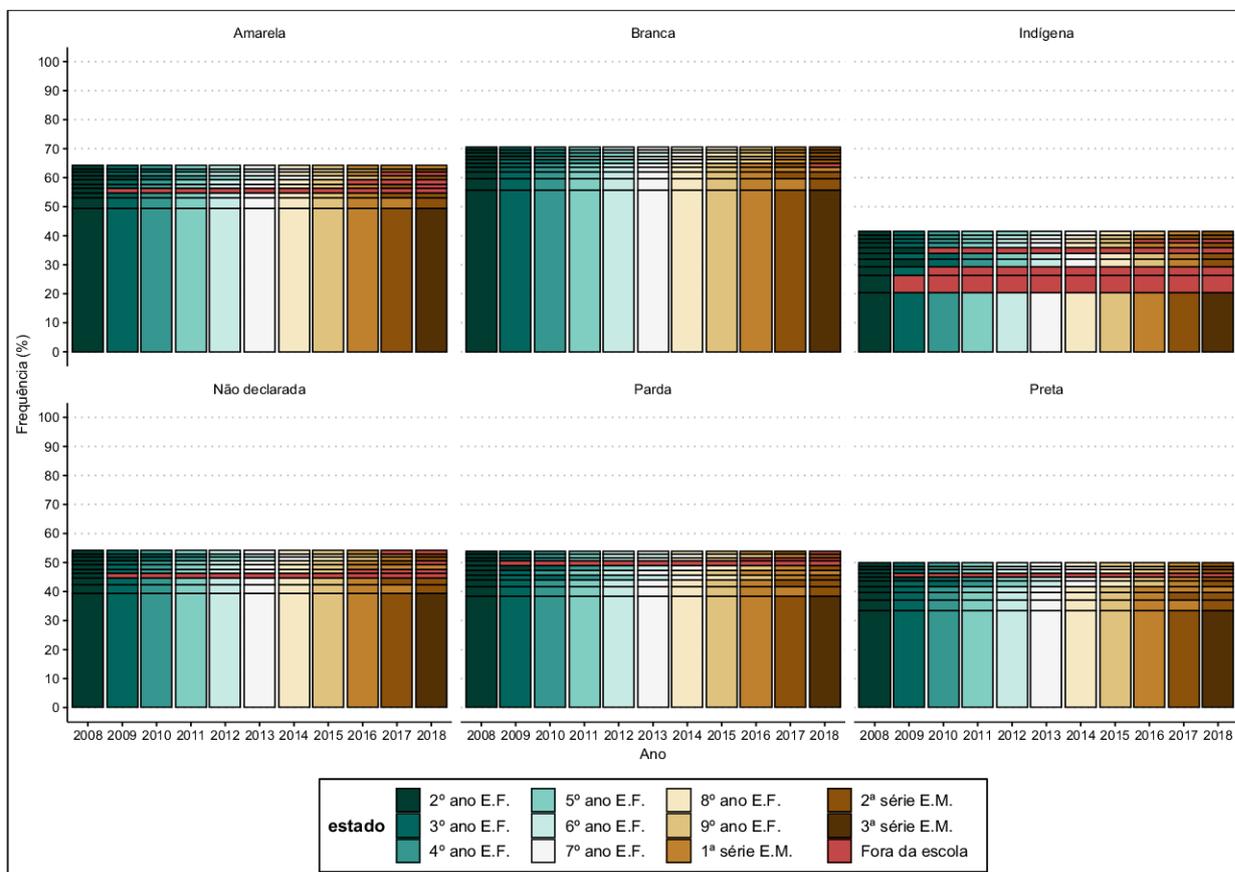
Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

estudantes de cor/raça branca; 26,5% dos estudantes de cor/raça amarela; 47,1% dos estudantes de cor/raça indígena; e 32,3% dos estudantes de cor/raça não declarada (Figura 8).

Destaca-se que essa diferença observada nos resultados educacionais se inicia já no 2º ano do ensino fundamental, visto que ao considerar a repetência no biênio 2008/2009, a porcentagem de estudantes de cor/raça branca (6,4%) é próxima da metade em relação aos estudantes de cor/raça parda (12,3%), de cor/raça preta (12,7%) e de cor/raça não declarada (12,7%) (Figura 8).

Observa-se que um percentual maior de estudantes de cor/raça branca percorreu as etapas de ensino de forma regular - 55,83% - comparativamente aos estudantes de cor/raça parda - 38,45%, preta - 33,45%, amarela - 49,52%, indígena - 20,47% e não declarada - 39,4% (Figura 9).

Figura 9 – 10 sequências mais comuns por cor/raça - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018

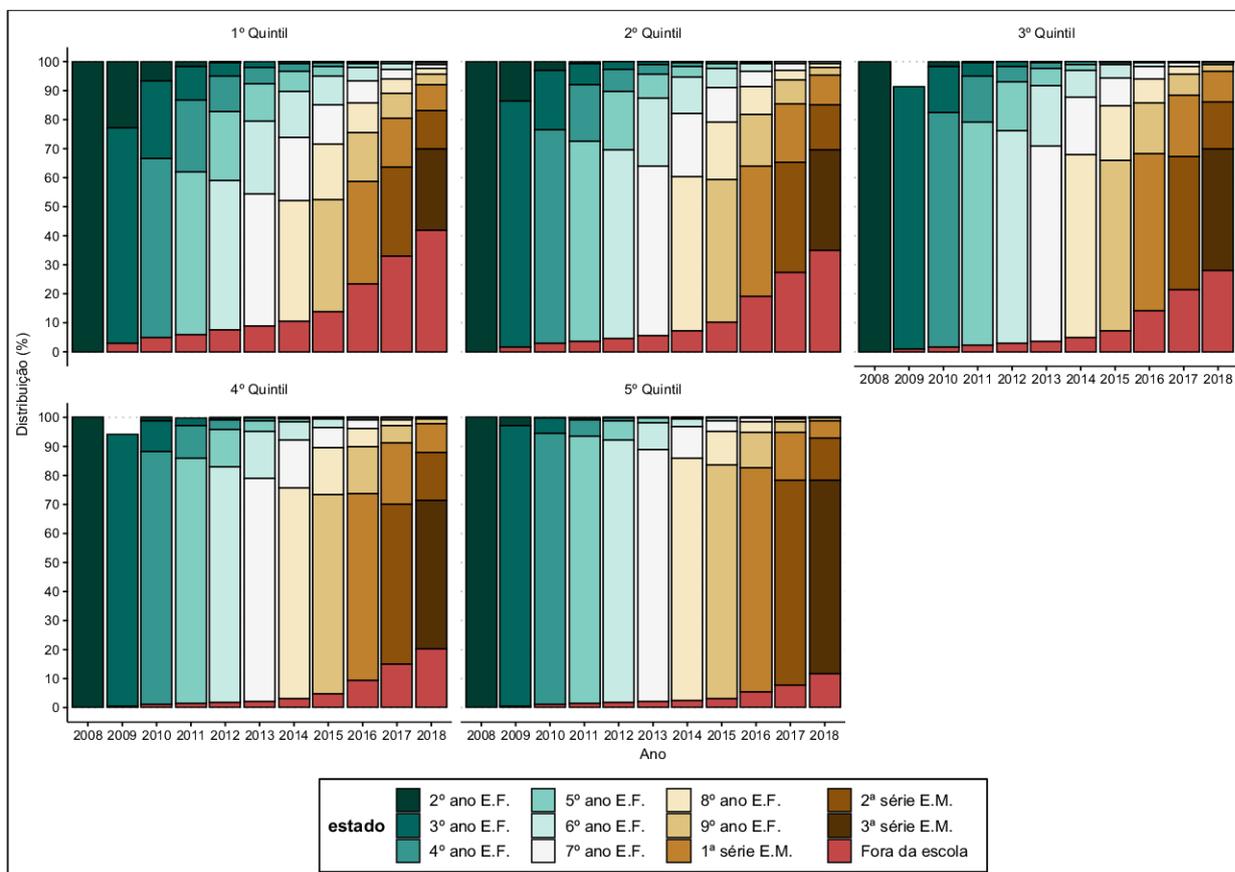


Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Em relação ao NSE, observa-se um gradiente nos resultados educacionais que melhoram com o maior NSE dos estudantes (Figura 10). Os seguintes percentuais de estudantes chegaram à 3ª série do ensino médio em 2018: 28,0% do 1º quintil; 34,6% do 2º quintil; 42,0% do 3º quintil; 51,2% do 4º quintil; e 66,8% do 5º quintil. Considerando os estudantes que estão fora da escola em 2018, tem-se os seguintes resultados: 41,8% do 1º quintil; 35,0% do 2º quintil; 28,0% do 3º quintil; 20,3% do 4º quintil; e 11,5% do 5º quintil. Portanto, quase 4 vezes mais estudantes do 1º quintil de NSE se encontraram fora da escola em 2018, comparativamente aos estudantes do 5º quintil, e aproximadamente 3 vezes menos estudantes se encontraram na etapa esperada em 2018. Além disso, o percentual de estudantes do 2º ano do ensino fundamental reprovados no biênio 2008/2009 do 1º quintil de NSE foi 8 vezes maior do que o percentual entre os estudantes do 5º quintil (Figura 10).

Por fim, as 10 sequências de trajetórias escolares mais comuns entre os estudantes do 1º quintil de NSE representaram 42,05% do total de trajetórias nesse grupo, enquanto

Figura 10 – Distribuição relativa dos estados por ano e quintil de NSE - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018



Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

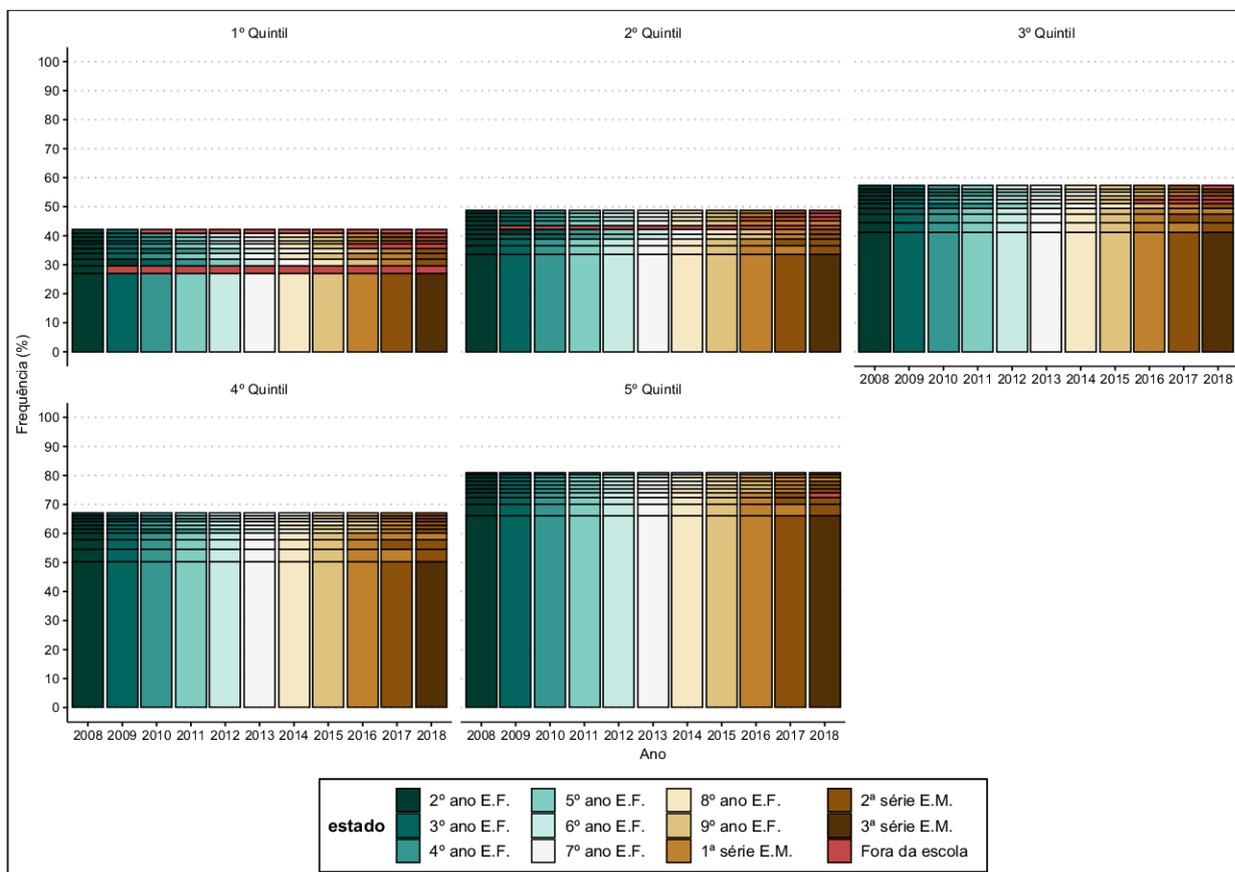
as 10 mais comuns entre os estudantes do 5º quintil representaram 81,08% (Figura 11). Os estudantes de grupos socioeconômicos mais baixos, portanto, percorreram trajetórias mais diversificadas que os estudantes dos grupos mais altos.

Na seção 5.1 a seguir exploram-se os resultados da identificação de padrões de trajetórias escolares que procuram resumir a complexidade de trajetórias encontrada na coorte analisada.

5.1 Padrões de trajetórias escolares

Nesta seção apresentam-se os padrões de trajetórias escolares identificados a partir da análise de agrupamentos realizada para uma amostra de estudantes da coorte de estudo. O processo de seleção do número de agrupamentos está descrito no Apêndice A.

Figura 11 – 10 sequências mais comuns por quintil de NSE - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018



Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Os padrões identificados foram nomeados com o objetivo de se identificar um comportamento que seja predominante em cada grupo de trajetórias e os nomes dados foram: "regular"; "pouca irregularidade"; "evasão - ensino médio"; "evasão - anos finais"; "evasão - anos iniciais"; e "evasão precoce". A seguir descrevem-se os padrões identificados.

O padrão "regular" é caracterizado pela quase totalidade (99,99%) dos estudantes desse grupo estarem matriculados na 3ª série do Ensino Médio em 2018, percorrendo a trajetória escolar esperada no tempo adequado, sem repetências ou evasão (Figura 12). Esse grupo representa 43,94% do total da amostra, sendo que Malaguth (2022) identificou 50,5% de trajetórias "regulares", considerando o estado de Minas Gerais, Bof, Oliveira e Barros (2019) identificaram 67% de trajetórias "regulares" considerando a trajetória do 5º ano ao 9º ano do ensino fundamental e Soares, Alves e Fonseca (2021) identificaram 53,1% trajetórias "regulares" considerando o ensino fundamental como um todo.

Os estudantes desse grupo experimentaram cada uma das etapas de ensino desde o 2º ano do Ensino Fundamental até à 3ª série do Ensino Médio em aproximadamente 1 ano (Figura 13). São, portanto, estudantes que chegam na última etapa da Educação Básica no tempo certo.

Os estudantes do padrão definido como "pouca irregularidade" estavam matriculados majoritariamente na 1ª série e na 2ª série do Ensino Médio em 2018, 29,8% e 54,3%, respectivamente (Figura 12). Portanto, 84,1% dos estudantes nesse padrão estão em até 2 etapas de ensino "em atraso", considerando que em 2018 deveriam estar na 3ª série do Ensino Médio, caso tivessem uma trajetória regular. Ademais, 5,47% dos estudantes estavam fora da escola em 2018 nesse grupo de estudantes.

O padrão "pouca irregularidade" contém 27,91% do total de estudantes da amostra e as 10 sequências mais observadas representaram 54,72% do total de trajetórias desse grupo (Figura 14), de modo que se observa uma heterogeneidade maior do que o grupo "regular". A sequência mais comum nesse grupo é dos estudantes que repetiram a 1ª série do Ensino Médio uma vez, representando 12% dos estudantes desse grupo, seguida da trajetória dos estudantes que repetiram a 2ª série do Ensino Médio, representando 8,60% dos estudantes.

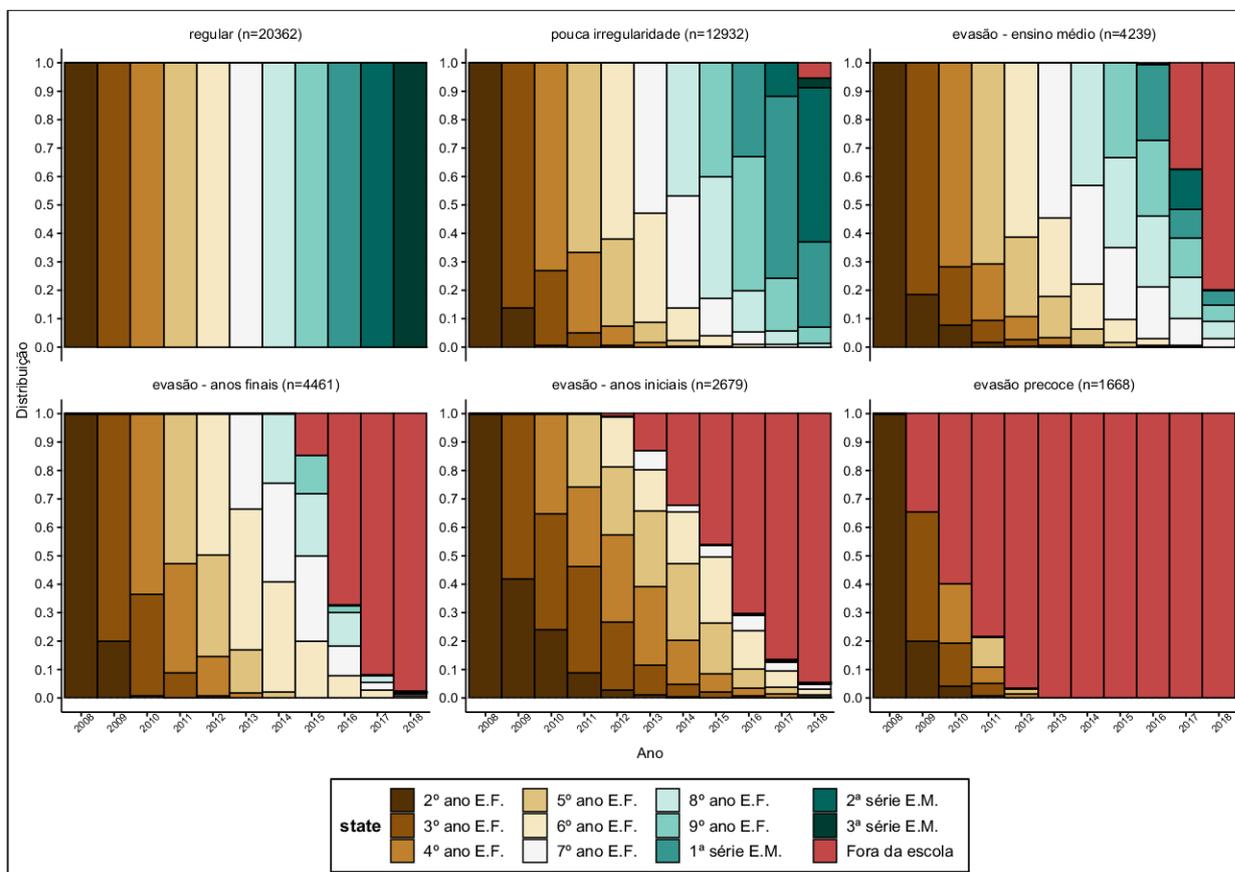
No padrão "pouca irregularidade", os estudantes experimentaram mais do que 1 ano em todas as etapas desde o 2º ano do Ensino Fundamental até a 1ª série do Ensino Médio (Figura 13). Destacam-se 1,18 anos experimentados no 3º ano do Ensino Fundamental, 1,16 no 6º ano do Ensino Fundamental e 1,27 na 1ª série do Ensino Médio. Dos 11 anos considerados, 10,25 foram vivenciados até a 1ª série do Ensino Médio.

O padrão "evasão - ensino médio" representa 9,15% do total da amostra e nesse grupo a evasão se iniciou em 2017, com 37,5% dos estudantes fora da escola nesse ano, aumentando para 79,92% no ano seguinte (Figura 12). Em 2016, caso os estudantes seguissem uma trajetória regular eles deveriam estar matriculados na 1ª série do Ensino Médio, no entanto, apenas 26,63% dos estudantes estavam nessa etapa nesse ano.

As duas trajetórias mais comuns nesse grupo - representando 13,78% e 12,43% do total de trajetórias - são de estudantes que evadiram em 2018 e 2017, respectivamente, após uma trajetória regular. As 10 trajetórias mais comuns nesse grupo representaram 44,56% do total de trajetórias (Figura 14). Em relação ao tempo médio vivido em cada um dos estados considerados, nesse grupo 1,28 anos foram experimentados no 2º ano do Ensino Fundamental e 1,45 no 7º ano do Ensino Fundamental. Dos 11 anos considerados na análise, 1,18 anos foram vivenciados fora da escola (Figura 13).

Os estudantes do padrão "evasão - anos finais" representam 9,63% do total da amostra e a evasão nesse grupo iniciou em 2015, com 14,6% dos estudantes fora da escola nesse ano, aumentando para 67,3% no ano seguinte e chegando à 97,7% em 2018

Figura 12 – Distribuição relativa dos estados por ano - Padrões identificados - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018

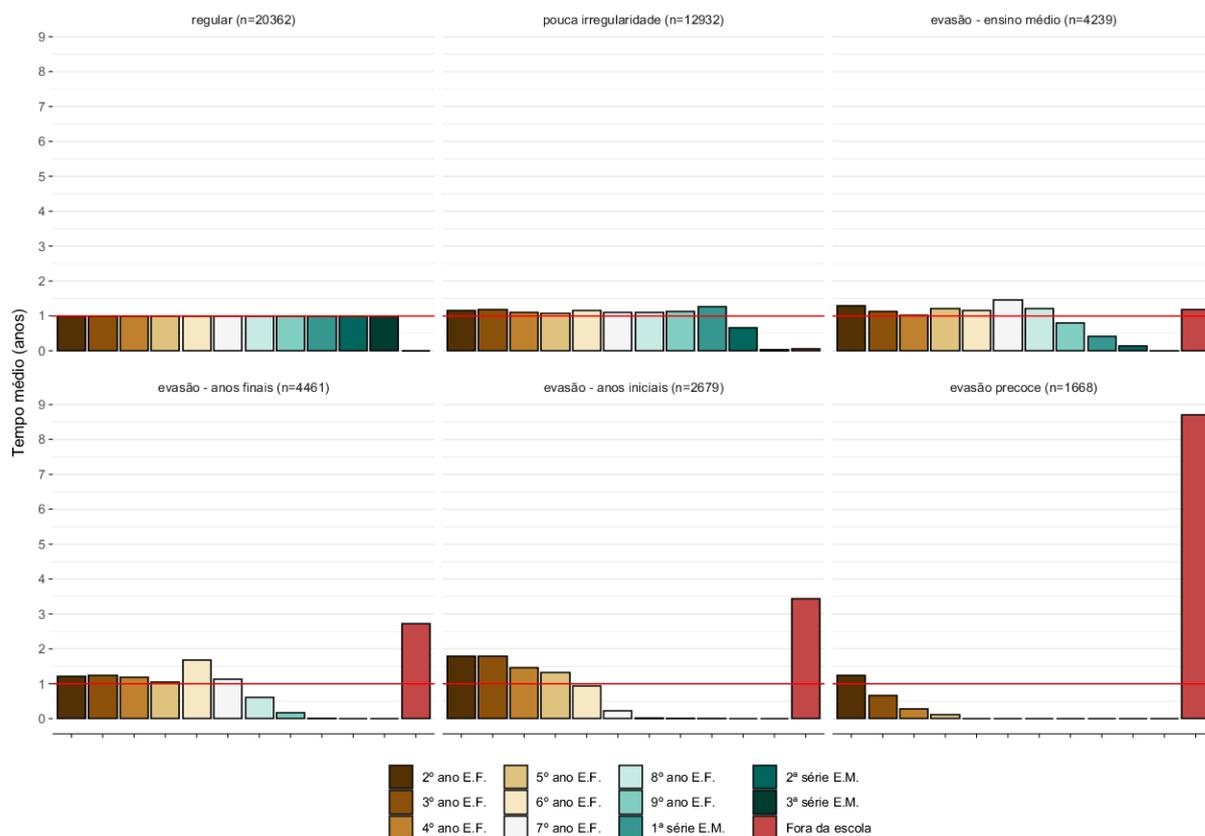


Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

(Figura 12). Em 2015, caso os estudantes desse grupo seguissem uma trajetória regular deveriam estar matriculados no 9º ano do Ensino Fundamental, no entanto, 13,6% dos estudantes estavam nessa etapa nesse ano e 20% dos estudantes estavam matriculados no 6º ano do ensino fundamental. As 10 trajetórias mais comuns no grupo "evasão - anos finais" representam 38,89% do total de trajetórias e as duas trajetórias mais comuns - representando 12,91% e 5,9% do total de trajetórias - são de estudantes que evadiram em 2016 e 2015, respectivamente, após uma trajetória regular (Figura 14). Dos 11 anos considerados entre 2008 e 2018, os estudantes desse grupo vivenciaram 1,68 anos no 6º ano do Ensino Fundamental e 2,71 anos fora da escola (Figura 13).

O padrão "evasão - anos iniciais" representa 5,78% do total da amostra. Nesse grupo, 41,96% dos estudantes repetiram o 2º ano do Ensino Fundamental em 2009, 6,79% dos estudantes estavam no 7º ano do Ensino Fundamental em 2013, ano em que se observa 13,03% estudantes fora da escola. O número de estudantes fora da escola passou para 32,06%

Figura 13 – Tempo médio experimentado por estado - Padrões identificados - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018

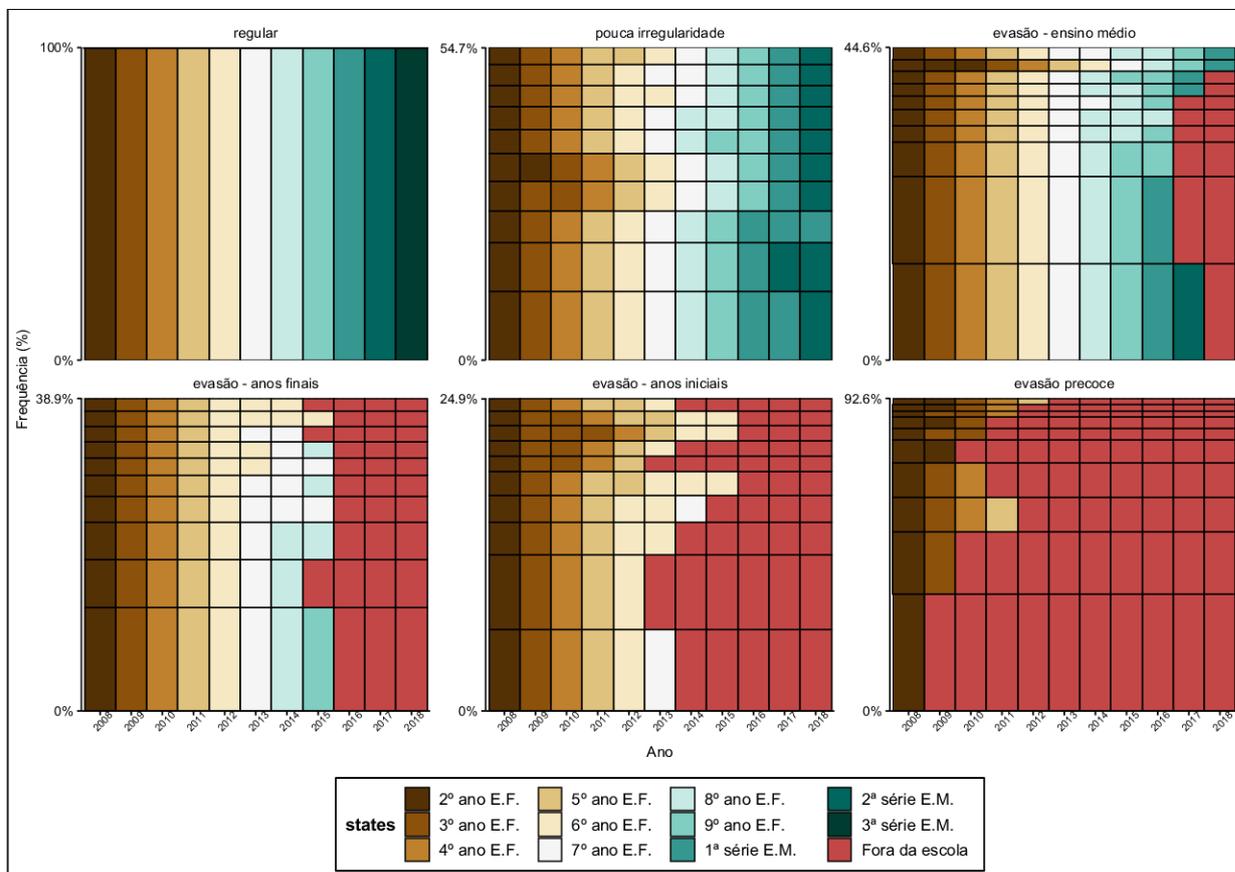


Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

em 2014 e chegou à 94,44% em 2018 (Figura 12). As 10 trajetórias mais comuns nesse grupo representam 24,93% do total de trajetórias (Figura 14) e os estudantes passaram 1,79 anos no 2º ano, 1,78 no 3º ano, 1,47 no 4º ano e 1,31 no 5º ano do Ensino Fundamental, somando 6,36 anos nessas quatro etapas de ensino, que deveriam ser cursadas em 4 anos (Figura 13). Além disso, 3,43 anos foram experimentados fora da escola.

Por fim, o padrão "evasão precoce" representa 3,6% do total da amostra e todos os estudantes desse grupo estão fora da escola a partir do ano de 2013. Em 2009, 45,5% dos estudantes progrediram de etapa e estavam no 3º ano do Ensino Fundamental. As 10 sequências mais comuns desse grupo representam 92,57% do total de sequências do grupo e a sequência mais comum (34,65%) é dos estudantes que evadem logo após o 2º ano do Ensino Fundamental (Figura 14). Dos 11 anos considerados, os estudantes experimentaram 1,25 anos no 2º ano do Ensino Fundamental e o tempo médio no estado "fora da escola" foi de 8,69 anos (Figura 13).

Figura 14 – 10 seqüências mais comuns - Padrões identificados - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental em 2008 - Brasil - 2008-2018



Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

5.2 Diferenciais demográficos e socioeconômicos

Nesta seção são apresentadas as características demográficas - sexo e cor/raça - e socioeconômicas dos estudantes tendo em vista os seis padrões de trajetórias que foram identificados. Primeiramente, faz-se uma descrição da Tabela 7 que apresenta o total e a proporção de estudantes em cada um dos padrões pelas variáveis consideradas, tendo em vista a distribuição da amostra apresentada na Tabela 6. Em seguida, são apresentados os resultados do teste Qui-Quadrado de independência entre proporções na Tabela 8.

O percentual de estudantes do sexo masculino no padrão "regular" é de 43,29%, sendo 7,65 pontos percentuais (p.p.) abaixo do presente na amostra. Em todos os demais padrões identificados os estudantes do sexo masculino são sobrerrepresentados em relação às estudantes do sexo feminino. Em relação aos padrões "evasão - anos iniciais", "evasão - anos finais" e "evasão - ensino médio" há uma gradação em que se diminui a sobrerrepresentação dos estudantes do sexo masculino, passando de 62,97% para 57,16% e chegando à 53,08%.

Os estudantes de cor/raça parda representam 37,31% do total de estudantes da amostra e dentre os estudantes do padrão "regular" representam 33,50%. Em todos os demais padrões identificados esses estudantes são sobrerrepresentados, sendo que no padrão "evasão - anos iniciais" representam 44,08% do total de estudantes, aproximadamente 7 p.p. a mais do que na amostra.

Tabela 7 – Estatísticas descritivas - Sexo, Cor/Raça e Nível Socioeconômico por agrupamento - Amostra - Coorte de estudantes nascidos entre 01/07/2000 e 30/06/2001 matriculados no 2º Ano do ensino fundamental - Brasil - 2008-2018.

Variável	"regular"		"pouca irregularidade"		"evasão - ensino médio"		"evasão - anos finais"		"evasão - anos iniciais"		"evasão precoce"		Total		
	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	
Sexo	Masculino	8.815	43,29	7.340	56,76	2.250	53,08	2.550	57,16	1.687	62,97	963	57,73	23.605	50,94%
	Feminino	11.547	56,71	5.592	43,24	1.989	46,92	1.911	42,84	992	37,03	705	42,27	22.736	49,06%
Cor/Raça	Parda	6.821	33,50	5.133	39,69	1.677	39,56	1.818	40,75	1.181	44,08	658	39,45	17.288	37,31%
	Branca	7.560	37,13	3.447	26,65	1.013	23,90	920	20,62	371	13,85	224	13,43	13.535	29,21%
	Não declarada	5.259	25,83	3.627	28,05	1.353	31,92	1.475	33,06	949	35,42	673	40,35	13.336	28,78%
	Preta	589	2,89	612	4,73	155	3,66	208	4,66	153	5,71	82	4,92	1.799	3,88%
	Indígena	47	0,23	66	0,51	22	0,52	24	0,54	19	0,71	26	1,56	204	0,44%
	Amarela	86	0,42	47	0,36	19	0,45	16	0,36	6	0,22	5	0,30	179	0,39%
Quintil de NSE	1º Quintil	2.498	12,27	2.573	19,90	1.173	27,67	1.249	28,00	1.132	42,25	700	41,97	9.325	20,12%
	2º Quintil	3.042	14,94	2.809	21,72	1.011	23,85	1.181	26,47	692	25,83	398	23,86	9.133	19,71%
	3º Quintil	3.925	19,28	2.799	21,64	868	20,48	970	21,74	475	17,73	253	15,17	9.290	20,05%
	4º Quintil	4.715	23,16	2.621	20,27	713	16,82	718	16,10	266	9,93	163	9,77	9.196	19,84%
	5º Quintil	6.182	30,36	2.130	16,47	474	11,18	343	7,69	114	4,26	154	9,23	9.397	20,28%
Total	20.362	43,94%	12.932	27,91%	4.239	9,15%	4.461	9,63%	2.679	5,78%	1.668	3,60%	46.341	100,00%	

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Os estudantes de cor/raça branca estão sobrerrepresentados no padrão "regular", com aproximadamente 8 p.p. a mais do que presentes na amostra. Tal qual foi identificado para os estudantes do sexo masculino, há uma gradação na participação de estudantes de cor/raça branca em relação aos padrões "evasão - anos iniciais", "evasão - anos finais" e "evasão - ensino médio" que passa de 13,85% para 20,62% e chega a 23,90%.

Os estudantes de cor/raça preta representam 3,88% do total da amostra. Em relação ao padrão "regular", esses estudantes representam 2,89% do total e representam 5,71% do padrão "evasão - anos iniciais", 4,66% do padrão "evasão - anos finais" e 3,66% do padrão "evasão - ensino médio".

É relevante destacar que os estudantes de cor/raça amarela e indígena representam 0,83% dos estudantes da amostra e a pequena quantidade de estudantes em cada um dos grupos identificados dificulta uma interpretação adequada. Os estudantes de cor/raça "não declarada" representam 28,78% dos estudantes da amostra e a distribuição dos estudantes entre os padrões de trajetória se assemelha aos estudantes de cor/raça parda.

Em relação ao Nível Socioeconômico (NSE) dos estudantes, os estudantes do quintil mais baixo (1º Quintil) representam 42,25% do total de estudantes do padrão "evasão - anos iniciais" e 12,27% do padrão "regular". Já os estudantes do quintil mais alto (5º Quintil) representam 4,26% e 30,36% nesses mesmos padrões, respectivamente. Dentre os padrões "evasão - anos iniciais", "evasão - anos finais" e "evasão - ensino médio" há um gradiente

entre os estudantes por quintil de NSE, sendo sobrerepresentados os dos quintis mais baixos e subrepresentados os dos quintis mais altos. Resultado semelhante é encontrado por Soares, Alves e Fonseca (2021), que identificaram maior percentual de trajetórias "regulares" com aumento do NSE.

O teste Qui-Quadrado de independência de proporções apresentou resultados que permitem rejeitar a hipótese nula de que não há associação entre os três pares de variáveis considerados: sexo e padrão; cor/raça e padrão; e quintil de NSE e padrão. Portanto, pode-se concluir com alto grau de confiança que há uma associação entre as variáveis e a distribuição dos estudantes pelas características demográficas e socioeconômicas consideradas não é um evento aleatório.

Tabela 8 – Teste Qui-Quadrado de independência de proporções

Variáveis	Estatística	Graus de liberdade	P-valor
Sexo e padrão de trajetória	914,67	5	< 0.000
Cor/raça e padrão de trajetória	1.547,1	25	< 0.000
Quintil de NSE e padrão de trajetória	4.910,9	20	< 0.000

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Os resultados descritos neste capítulo apontam que mais da metade da coorte dos estudantes nascidos entre 1 de julho de 2000 e 30 de junho de 2001 matriculados no 2º ano do ensino fundamental em 2008 repetiram uma ou mais etapas de ensino ou evadiram da escola até o ano de 2018. Foram identificadas trajetórias escolares que se caracterizam por nenhuma ou poucas irregularidades contemplando 43,94% e 27,91% dos estudantes, respectivamente. Também foram identificadas trajetórias que se caracterizaram pela evasão escolar, seja ela precoce, representando 3,6% dos estudantes, nos anos iniciais do ensino fundamental - 5,78%, nos anos finais do ensino fundamental - 9,63%, e no ensino médio - 9,15%.

6 Discussão e Conclusão

Historicamente, o sistema educacional brasileiro tem apresentado problemas relativos ao fluxo escolar, conjugando índices altos de repetência e de evasão que evidenciam a baixa eficiência do sistema educacional (RIBEIRO, 1991); (KLEIN, 2006); e (KLEIN, 2022), sendo que as transições escolares e o acesso aos níveis de ensino mais elevados se distribuem de forma desigual entre os grupos sociais (RIBEIRO, 2011). As trajetórias escolares dos estudantes brasileiros são marcadas por irregularidades que dificultam a conclusão da educação básica no tempo esperado com importantes diferenciais demográficos e socioeconômicos associados (SOARES; ALVES; FONSECA, 2021), (BOF; OLIVEIRA; BARROS, 2019), (MALAGUTH, 2022).

A estratégia de pesquisa adotada neste trabalho procura avançar no estudo das trajetórias escolares propondo-se uma nova abordagem para sua caracterização. Adota-se a perspectiva de Curso de Vida Educacional (CROSNOE; BENNER, 2016) e de Curso de Vida (ELDER; SHANAHAN, 2007), levando-se em conta que o caminho esperado no sistema educacional é aquele que o estudante percorre de forma regular. Adota-se o arcabouço metodológico de análise de sequências e toma-se a trajetória dos estudantes como elemento de análise em sua totalidade (ABBOTT, 1995), (AISENBREY; FASANG, 2010).

Os resultados encontrados por este trabalho mostram que 44,5% dos estudantes da coorte analisada estavam na 3ª série do ensino médio após 11 anos de escolarização, considerando que eles estavam no 2º ano do ensino fundamental em 2008. Resultado semelhante foi encontrado por Malaguth (2022), que identificou que 50,5% da coorte analisada estava na 3ª série do ensino médio, considerando os estudantes do estado de Minas Gerais. Essas evidências apontam que a repetência e a evasão condicionaram as trajetórias escolares e contribuíram para menos da metade da coorte analisada estar na etapa de ensino adequada, pois a escolha da coorte permitiu descartar fatores explicativos como entrada tardia na escola e/ou repetências anteriores.

No ano de 2018, 27% dos estudantes da coorte analisada neste trabalho estavam fora da escola. Considerando apenas o estado de Minas Gerais, Malaguth (2022) identificou que 22,4% dos estudantes estavam fora da escola, sendo que também adotou o critério de evasão permanente, tornando os resultados comparáveis. Destaca-se o aumento relevante de estudantes fora da escola na transição entre 14 e 15 anos de idade: em 2015, quando os estudantes da coorte acompanhada estavam com 14 anos, 7,8% estava fora da escola e em 2016 esse número passa para 14,2%, aumentando nos anos seguintes para 20,9% em 2017 e 27,3% em 2018. Tal efeito mostra que a evasão é uma combinação de atraso escolar, em

virtude de uma ou mais repetências, e idade, sendo que completar 15 anos pode ser um marco importante na decisão do estudante de sair do ensino regular (RODERICK, 1994), (MALAGUTH, 2022), tendo em vista que nessa idade torna-se permitido a entrada na EJA.

Procura-se analisar as trajetórias escolares a partir do arcabouço metodológico de análise de sequências e, particularmente, aplica-se o método *Optimal matching*, com o objetivo de se quantificar as similaridades entre as sequências (ABBOTT; HRYCAK, 1990), (ABBOTT, 1995). A partir de uma amostra com 46.301 estudantes, calculam-se as distâncias entre todos os pares de sequências e aplica-se o método de agrupamento aglomerativo hierárquico. Os padrões de trajetórias identificados representam estudantes que apresentaram trajetórias regulares (43,94%), com poucas irregularidades (27,91%), que evadiram no ensino médio (9,15%), que evadiram nos anos finais do ensino fundamental (9,63%), que evadiram nos anos iniciais do ensino fundamental (5,78%) e que evadiram precocemente (3,60%).

As trajetórias analisadas por Soares, Alves e Fonseca (2021), acompanhadas entre 2007 e 2015, são classificadas em "regular- 53,1%, com "poucas irregularidades- 16,5%, e com "grandes irregularidades" - 30,4%. A adoção de critérios de definição de coorte de análise distintos dificulta uma comparação dos resultados. Porém, pode se especular que as trajetórias classificadas como "regular" por Soares, Alves e Fonseca (2021) se assemelham ao padrão "regular" deste trabalho, pois caracterizam estudantes que tem uma ou nenhuma intercorrência na trajetória escolar; o grupo com "poucas irregularidades" se assemelha ao padrão "poucas irregularidades" deste trabalho, por caracterizar estudantes que tenham acumulado entre 1 e 3 repetências e estudantes, em minoria, que tenham evadido; o último grupo de estudantes com "grandes irregularidades" corresponde aos padrões "evasão anos finais", "evasão anos iniciais" e "evasão precoce" identificados neste trabalho; e o padrão "evasão ensino médio" identificado neste trabalho não teria uma correspondência direta, pois a coorte analisada por Soares, Alves e Fonseca (2021) é acompanhada até o ano de 2015, ano que deveria estar matriculada na última etapa do ensino fundamental.

As desigualdades de oportunidades e de resultados educacionais entre grupos sociais identificadas por Ribeiro (2011) se mostram presentes nas trajetórias da coorte analisada. Estudantes do sexo masculino, de cor/raça parda ou preta e nível socioeconômico mais baixo estão sobrerrepresentados nos padrões de trajetória escolar associados a maiores graus de repetência e de culminância da evasão, sobretudo a evasão nas etapas iniciais da educação básica.

Considerando os padrões "evasão - anos iniciais", "evasão - anos finais" e "evasão - ensino médio", que refletem diferentes etapas onde ocorre a evasão escolar, observou-se uma gradação negativa na participação de estudantes do sexo masculino, de estudantes de cor/raça parda e preta e de estudantes do 1º quintil de NSE. Ou seja, à medida que a evasão

ocorre mais cedo no percurso escolar, mais estão presentes os estudantes do sexo masculino, de cor/raça parda e preta e de nível socioeconômico mais baixo. Resultado semelhante é encontrado por [Soares, Alves e Fonseca \(2021\)](#), que identificaram "um gradiente ascendente na proporção de trajetórias regulares (tipo 1) ao se passar do NSE mais baixo (32,5%) para o NSE mais alto (73,5%). Entre as escolas de NSE mais baixo, a maioria dos estudantes encontra-se no tipo 3, trajetórias com muitas irregularidades."

Este trabalho adota um arcabouço metodológico que permite lidar com dados longitudinais categóricos, como são as trajetórias escolares. Os resultados encontrados na identificação dos padrões de trajetórias apontam para um detalhamento da categoria de trajetórias classificadas como com "grandes irregularidades" por [Soares, Alves e Fonseca \(2021\)](#). Os padrões identificados analisados neste trabalho refletem a análise do número de agrupamentos descrita [Apêndice A](#), de modo que pode-se analisar um número maior de padrões. Por fim, apresenta-se uma tipologia de trajetórias que podem ser utilizadas de forma sintética como variável independente por outros trabalhos.

Destaca-se que os resultados deste trabalho apresentam limitações. A quantidade de trajetórias distintas observadas por [Soares, Alves e Fonseca \(2021\)](#) – 80.019 – é cerca de 7,5 vezes maior que as 12.118 trajetórias que foram analisadas nesse trabalho, isso significa que parte substantiva da complexidade de trajetórias que os estudantes brasileiros experimentaram no percurso educacional não foi analisada. Ademais, apenas uma parcela - 2.499 - das trajetórias foi considerada para definição dos padrões de trajetórias.

Além disso, decidiu-se por agregar a migração dos estudantes para a Educação de Jovens e Adultos juntamente à evasão por considerar que ambas dizem respeito à evasão do ensino regular. Tal decisão implica que ambos os movimentos por parte dos estudantes são considerados iguais, o que, no entanto, pode não ser o caso. Por fim, a consideração da evasão como permanente implica em limitações no sentido de não se analisar as entradas e as saídas dos estudantes o que implica em sistematicamente superestimar o total de estudantes que estão fora da escola.

Esta pesquisa procura caracterizar as trajetórias escolares dos estudantes matriculados no 2º ano do ensino fundamental em 2008 até o ano de 2018. O acompanhamento dessa coorte pode ser estendido considerando tanto os estudantes que permaneceram matriculados na educação básica por mais anos e aqueles estudantes que ingressaram no ensino superior. Dessa forma, pode-se estender a caracterização das trajetórias escolares tendo em vista um período maior e a inserção dos estudantes no ensino superior.

Ademais, outros fatores relevantes para determinação dos resultados educacionais ([BUCHMANN; HANNUM, 2001](#)) podem ser incluídos na análise, como fatores regionais ([RIBEIRO, 2011](#)), ([SOARES; ALVES; FONSECA, 2021](#)), a rede de ensino - estadual, municipal, federal ou privada - ([RIBEIRO, 2011](#)), ([SOARES; ALVES; FONSECA, 2021](#)) e se houve movimento migratório por parte dos estudantes ([MALAGUTH, 2022](#)).

A trajetória dos estudantes brasileiros é marcada por alto grau de repetência que se acumulam na vida estudantil e dificultam o progresso para a conclusão das etapas de escolarização da educação básica, inclusive incluindo nesse desfecho rompimento com a escola. A multiplicidade de trajetórias escolares no sistema de ensino brasileiro implica em uma complexidade da qual se é difícil extrair informações, de modo que são necessárias categorizações. Trabalhos recentes categorizaram as trajetórias de forma dicotômica - regular e irregular (BOF; OLIVEIRA; BARROS, 2019), (MALAGUTH, 2022) - ou em graus de irregularidade, totalizando ao máximo 3 categorias (SOARES; ALVES; FONSECA, 2021).

Neste trabalho aplica-se uma abordagem teórica e metodológica que permiti abordar o problema da complexidade das trajetórias construindo categorias ou padrões que sejam mais informativos e, além disso, que pode ser aplicada com outros objetivos de pesquisa. Conclui-se que os padrões de trajetórias escolares identificados neste estudo podem contribuir para o diagnóstico mais preciso da realidade educacional do sistema de ensino brasileiro.

Referências

ABBOTT, A. Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology*, v. 21, p. 93–113, 1995. Citado 7 vezes nas páginas 14, 15, 33, 35, 38, 59 e 60.

ABBOTT, A.; HRYCAK, A. Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers. *American Journal of Sociology*, v. 96, n. 1, p. 144–185, jul. 1990. ISSN 0002-9602, 1537-5390. Disponível em: <<https://www.journals.uchicago.edu/doi/10.1086/229495>>. Citado 3 vezes nas páginas 15, 38 e 60.

AISENBREY, S.; FASANG, A. E. New Life for Old Ideas: The "Second Wave" of Sequence Analysis Bringing the "Course" Back Into the Life Course. *Sociological Methods & Research*, v. 38, n. 3, p. 420–462, fev. 2010. ISSN 0049-1241, 1552-8294. Disponível em: <<http://journals.sagepub.com/doi/10.1177/0049124109357532>>. Citado 7 vezes nas páginas 14, 15, 33, 38, 39, 40 e 59.

ALEXANDER, K. L.; ENTWISLE, D. R.; HORSEY, C. S. From First Grade Forward: Early Foundations of High School Dropout. *Sociology of Education*, v. 70, n. 2, p. 87–107, 1997. ISSN 00380407, 19398573. Publisher: [Sage Publications, Inc., American Sociological Association]. Disponível em: <<http://www.jstor.org/stable/2673158>>. Citado 2 vezes nas páginas 13 e 20.

ALEXANDER, K. L.; ENTWISLE, D. R.; KABBANI, N. S. The Dropout Process in Life Course Perspective: Early Risk Factors at Home and School. *Teachers College Record*, v. 103, n. 5, p. 760–822, out. 2001. Citado 2 vezes nas páginas 13 e 20.

ALVES, M. T. G.; SOARES, J. F.; XAVIER, F. P. Desigualdades educacionais no ensino fundamental de 2005 a 2013: hiato entre grupos sociais. *Revista Brasileira de Sociologia - RBS*, v. 4, n. 7, p. 49–82, jul. 2016. ISSN 2318-0544, 2317-8507. Disponível em: <<https://rbs.sbsociologia.com.br/index.php/rbs/article/view/181>>. Citado 2 vezes nas páginas 22 e 24.

ANDREW, M. The Scarring Effects of Primary-Grade Retention? A Study of Cumulative Advantage in the Educational Career. *Social Forces*, v. 93, n. 2, p. 653–685, 2014. ISSN 00377732, 15347605. Publisher: Oxford University Press. Disponível em: <<http://www.jstor.org/stable/43287841>>. Citado na página 20.

BARROS, R. P. d. et al. *Consequências da violação do direito à educação*. 1. ed. Rio de Janeiro: Autografia, 2021. ISBN 978-65-5943-345-2. Citado na página 13.

BILLARI, F. C. Sequence Analysis in Demographic Research. *Canadian Studies in Population*, v. 28, n. 2, p. 439, dez. 2001. ISSN 1927-629X, 0380-1489. Disponível em: <<https://journals.library.ualberta.ca/csp/index.php/csp/article/view/15880>>. Citado 2 vezes nas páginas 33 e 34.

BOF, A. M.; OLIVEIRA, A. S. d.; BARROS, G. T. d. F. Trajetória escolar, aprendizagem e desigualdades no ensino fundamental no Brasil. *Cadernos de Estudos e Pesquisas em Políticas Educacionais*, v. 1, p. 55–78, ago. 2019. Disponível em:

- <<http://cadernosdeestudos.inep.gov.br/ojs3/index.php/cadernos/article/view/1003>>. Citado 6 vezes nas páginas 13, 18, 19, 52, 59 e 62.
- BRASIL. *Constituição da República Federativa do Brasil*. 1988. Disponível em: <https://www2.senado.leg.br/bdsf/bitstream/handle/id/518231/CF88_Livro_EC91_2016.pdf>. Citado 2 vezes nas páginas 13 e 17.
- BRASIL. *Lei n. 11.274 de 6 de fevereiro de 2006*. 2006. Publisher-place: Brasília. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/111274>. Citado 2 vezes nas páginas 17 e 26.
- BUCHMANN, C.; HANNUM, E. Education and Stratification in Developing Countries: A Review of Theories and Research. *Annual Review of Sociology*, v. 27, n. 1, p. 77–102, ago. 2001. ISSN 0360-0572, 1545-2115. Disponível em: <<https://www.annualreviews.org/doi/10.1146/annurev.soc.27.1.77>>. Citado 2 vezes nas páginas 22 e 61.
- CROSNOE, R.; BENNER, A. D. Educational Pathways. In: SHANAHAN, M. J.; MORTIMER, J. T.; JOHSON, M. K. (Ed.). *Handbook of the Life Course*. Cham: Springer International Publishing, 2016, (Handbooks of Sociology and Social Research, II). p. 179–200. ISBN 978-3-319-20879-4. Disponível em: <<http://link.springer.com/10.1007/978-3-319-20880-0>>. Citado 3 vezes nas páginas 14, 16 e 59.
- ELDER, G. H. The Life Course as Developmental Theory. *Child Development*, v. 69, n. 1, p. 1–12, fev. 1998. Disponível em: <<https://www.jstor.org/stable/1132065>>. Citado na página 16.
- ELDER, G. H.; SHANAHAN, M. J. The Life Course and Human Development. In: DAMON, W.; LERNER, R. M. (Ed.). *Handbook of Child Psychology*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2007. p. chpsy0112. ISBN 978-0-470-14765-8. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/9780470147658.chpsy0112>>. Citado 3 vezes nas páginas 14, 16 e 59.
- ELDER, G. H. J. Perspectives on the Life Course. In: ELDER, G. H. J. (Ed.). *Life course dynamics: trajectories and transitions, 1968-1980*. Ithaca, NY: Cornell University Press, 1985. Citado na página 16.
- GABADINHO, A. et al. Analyzing and Visualizing State Sequences in *R* with **TraMineR**. *Journal of Statistical Software*, v. 40, n. 4, 2011. ISSN 1548-7660. Disponível em: <<http://www.jstatsoft.org/v40/i04/>>. Citado 4 vezes nas páginas 38, 40, 43 e 44.
- GIANO, Z.; WILLIAMS, A. L.; BECNEL, J. N. Grade Retention and School Dropout: Comparing Specific Grade Levels Across Childhood and Early Adolescence. *The Journal of Early Adolescence*, v. 42, n. 1, p. 33–57, jan. 2022. ISSN 0272-4316, 1552-5449. Disponível em: <<http://journals.sagepub.com/doi/10.1177/02724316211010332>>. Citado na página 20.
- GOLGHER, A. B. Modelo PROFLUXO e indicadores derivados. In: *Introdução à Demografia da Educação*. Campinas, SP: ABEP, 2004. p. 207. Organizadores: Rios-Neto, Eduardo Luiz G; Riani, Juliana de Lucena Ruas. Citado na página 23.

- GROLEMUND, G.; WICKHAM, H. Dates and times made easy with lubridate. *Journal of Statistical Software*, v. 40, n. 3, p. 1–25, 2011. Disponível em: <<https://www.jstatsoft.org/v40/i03/>>. Citado na página 43.
- HALPIN, B. Optimal Matching Analysis and Life-Course Data: The Importance of Duration. *Sociological Methods & Research*, v. 38, n. 3, p. 365–388, fev. 2010. ISSN 0049-1241, 1552-8294. Disponível em: <<http://journals.sagepub.com/doi/10.1177/00491241110363590>>. Citado 2 vezes nas páginas 14 e 33.
- INEP. *O item cor-raça no censo escolar da educação básica*. Brasília, DF, 2015. Disponível em: <https://download.inep.gov.br/educacao_basica/educacenso/documentos/2015/cor_raca.pdf>. Citado na página 29.
- INEP. *Estimativas de fluxo escolar a partir do acompanhamento longitudinal dos registros de aluno do Censo Escolar do período 2007-2016*. Brasília, 2017. Citado 34 vezes nas páginas 15, 18, 25, 26, 28, 30, 32, 34, 35, 41, 42, 45, 46, 47, 48, 49, 50, 51, 52, 54, 55, 56, 57, 58, 71, 72, 74, 75, 76, 77, 78, 79, 80 e 81.
- INEP. *Indicador de Nível Socioeconômico do Saeb 2021: nota técnica*. Brasília, DF, 2023. 28 p. Disponível em: <<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/indicadores-educacionais/nivel-socioeconomico>>. Citado 3 vezes nas páginas 29, 30 e 31.
- JACOB, B. A.; LEFGREN, L. The Effect of Grade Retention on High School Completion. *American Economic Journal: Applied Economics*, v. 1, n. 3, p. 33–58, 2009. ISSN 19457782, 19457790. Publisher: American Economic Association. Disponível em: <<http://www.jstor.org/stable/25760170>>. Citado 2 vezes nas páginas 20 e 21.
- KLEIN, R. Produção e utilização de indicadores educacionais: metodologia de cálculo de indicadores do fluxo escolar da educação básica. *Revista Brasileira de Estudos Pedagógicos*, v. 84, n. 206-07-08, p. 107–157, 2003. ISSN 2176-6681, 0034-7183. Disponível em: <<http://rbep.inep.gov.br/ojs3/index.php/rbep/article/view/1378>>. Citado na página 27.
- KLEIN, R. Como está a educação no Brasil? O que fazer? *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 14, n. 51, p. 139–171, jun. 2006. ISSN 0104-4036. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-40362006000200002&lng=pt&tlng=pt>. Citado 2 vezes nas páginas 17 e 59.
- KLEIN, R. Revisitando a pedagogia da repetência. *Em Aberto*, v. 35, n. 113, p. p. 243–252, maio 2022. ISSN 2176-6673, 0104-1037. Disponível em: <<http://emaberto.inep.gov.br/ojs3/index.php/emaberto/article/view/5082>>. Citado 2 vezes nas páginas 18 e 59.
- KLEIN, R.; RIBEIRO, S. C. O censo educacional e o modelo de fluxo: o problema da repetência. *Revista Brasileira de Estatística*, v. 52, n. 197-198, p. 5–45, 1991. ISSN 0034-7175. Citado 2 vezes nas páginas 23 e 24.
- LEON, F. L. L. d.; MENEZES-FILHO, N. A. Reprovação, avanço e evasão escolar no Brasil. *pesquisa e planejamento econômico*, v. 32, n. 3, p. 36, dez. 2002. Citado 2 vezes nas páginas 13 e 21.

- LIAO, T. F. et al. Sequence analysis: Its past, present, and future. *Social Science Research*, v. 107, p. 102772, set. 2022. ISSN 0049089X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0049089X22000783>>. Citado 3 vezes nas páginas 15, 35 e 36.
- MAECHLER, M. et al. *cluster: Cluster Analysis Basics and Extensions*. [S.l.], 2022. R package version 2.1.4 — For new features, see the 'Changelog' file (in the package source). Disponível em: <<https://CRAN.R-project.org/package=cluster>>. Citado 2 vezes nas páginas 41 e 43.
- MALAGUTH, T. Z. *Migrações e fluxo escolar da coorte de estudantes de 2008 a 2019, em Minas Gerais*. Tese (Dissertação de mestrado) — Universidade Federal de Minas Gerais, Belo Horizonte, 2022. Disponível em: <<http://hdl.handle.net/1843/47288>>. Citado 14 vezes nas páginas 13, 14, 18, 19, 24, 26, 27, 44, 45, 52, 59, 60, 61 e 62.
- PALLAS, A. M. Educational transitions, trajectories, and pathways. In: MORTIMER, J. T.; SHANAHAN, M. J. (Ed.). *Handbook of the Life Course*. [S.l.]: Kluwer Academic Publishers, 2002, (Handbooks of Sociology and Social Research). p. 749. ISBN 0-306-47498-0. Citado na página 14.
- PEREIRA, R. B. Variações nos percursos de vida de jovens de mesma origem social: sequências, eventos e pontos de virada. *Revista Brasileira de Sociologia - RBS*, v. 4, n. 7, p. 83–110, jul. 2016. ISSN 2318-0544, 2317-8507. Disponível em: <<https://rbs.sbsociologia.com.br/index.php/rbs/article/view/182>>. Citado na página 34.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>. Citado 3 vezes nas páginas 38, 43 e 82.
- RAAB, M. *ggseqplot: Render Sequence Plots using 'ggplot2'*. [S.l.], 2022. Disponível em: <<https://maraab23.github.io/ggseqplot/>>. Citado na página 43.
- RIBEIRO, C. A. C. Desigualdade de oportunidades e resultados educacionais no Brasil. *DADOS – Revista de Ciências Sociais*, v. 54, n. 1, p. 41–87, 2011. ISSN 0011-5258. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0011-52582011000100002&lng=pt&tlng=pt>. Citado 6 vezes nas páginas 22, 23, 24, 59, 60 e 61.
- RIBEIRO, S. C. A pedagogia da repetência. *Estudos Avançados*, v. 5, n. 12, p. 07–21, ago. 1991. ISSN 0103-4014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40141991000200002&lng=pt&tlng=pt>. Citado 4 vezes nas páginas 13, 23, 24 e 59.
- RODERICK, M. Grade Retention and School Dropout: Investigating the Association. *American Educational Research Journal*, v. 31, n. 4, p. 729–759, 1994. ISSN 00028312, 19351011. Publisher: [American Educational Research Association, Sage Publications, Inc.]. Disponível em: <<http://www.jstor.org/stable/1163393>>. Citado 5 vezes nas páginas 13, 20, 21, 45 e 60.
- RUMBERGER, R. W.; LIM, S. A. Why students drop out of school: a review of 25 years of research. *California Dropout Research Project Report #15 October 2008*, p. 130, jan. 2008. Citado na página 20.

SANTOS, F. V. L. d. Pequeno manual para aplicação do modelo Profluxo. *Revista Brasileira de Educação*, v. 26, p. e260059, 2021. ISSN 1809-449X, 1413-2478. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-24782021000100241&tlng=pt>. Citado na página 23.

SENKEVICS, A. S.; MACHADO, T. d. S.; OLIVEIRA, A. S. A cor ou raça nas estatísticas educacionais : uma análise dos instrumentos de pesquisa do Inep. *Brasília : Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira*, v. 41, p. 48, 2016. ISSN 1414-0640. Citado na página 29.

SHIRASU, M. R.; ARRAES, R. d. A. e. Determinantes da evasão e repetência escolar no ensino médio do Ceará. *Revista Econômica do Nordeste*, v. 46, n. 4, p. 117–136, out. 2015. Citado 2 vezes nas páginas 13 e 21.

SOARES, J. F.; ALVES, M. T. G. UMA MEDIDA DO NÍVEL SOCIOECONÔMICO DAS ESCOLAS BRASILEIRAS UTILIZANDO INDICADORES PRIMÁRIOS E SECUNDÁRIOS. SSRN, 2023. Disponível em: <<https://ssrn.com/abstract=4325674>>. Citado 3 vezes nas páginas 29, 30 e 31.

SOARES, J. F.; ALVES, M. T. G.; FONSECA, J. A. Trajetórias educacionais como evidência da qualidade da educação básica brasileira. *Revista Brasileira de Estudos de População*, v. 38, p. 1–21, set. 2021. ISSN 1980-5519, 0102-3098. Disponível em: <<https://rebep.org.br/revista/article/view/1833>>. Citado 14 vezes nas páginas 13, 18, 19, 24, 25, 26, 27, 37, 52, 58, 59, 60, 61 e 62.

SOARES, T. M. et al. Fatores associados ao abandono escolar no ensino médio público de Minas Gerais. *Educação e Pesquisa*, v. 41, n. 3, p. 757–772, set. 2015. ISSN 1517-9702. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-97022015000300757&lng=pt&tlng=pt>. Citado 2 vezes nas páginas 13 e 21.

SOUZA, A. P. d. et al. FATORES ASSOCIADOS AO FLUXO ESCOLAR NO INGRESSO E AO LONGO DO ENSINO MÉDIO NO BRASIL. *pesquisa e planejamento econômico*, v. 42, n. 1, p. 36, 2012. Citado 2 vezes nas páginas 13 e 21.

STEARNS, E. et al. Staying Back and Dropping out: The Relationship between Grade Retention and School Dropout. *Sociology of Education*, v. 80, n. 3, p. 210–240, 2007. ISSN 00380407, 19398573. Publisher: [Sage Publications, Inc., American Sociological Association]. Disponível em: <<http://www.jstor.org/stable/20452707>>. Citado 2 vezes nas páginas 13 e 20.

STUDER, M. WeightedCluster Library Manual: A practical guide to creating typologies of trajectories in the social sciences with R. *LIVES Working Papers*, n. 24, 2013. Citado 4 vezes nas páginas 40, 41, 43 e 70.

STUDER, M.; RITSCHARD, G. What Matters in Differences Between Life Trajectories: A Comparative Review of Sequence Dissimilarity Measures. *Journal of the Royal Statistical Society Series A: Statistics in Society*, v. 179, n. 2, p. 481–511, fev. 2016. ISSN 0964-1998, 1467-985X. Disponível em: <<https://academic.oup.com/jrssa/article/179/2/481/7068212>>. Citado 4 vezes nas páginas 15, 35, 36 e 39.

WICKHAM, H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. Disponível em: <<https://ggplot2.tidyverse.org>>. Citado na página 43.

WICKHAM, H. *stringr: Simple, Consistent Wrappers for Common String Operations*. [S.l.], 2022. <https://stringr.tidyverse.org>, <https://github.com/tidyverse/stringr>. Citado na página 43.

WICKHAM, H. et al. *dplyr: A Grammar of Data Manipulation*. [S.l.], 2023. <https://dplyr.tidyverse.org>, <https://github.com/tidyverse/dplyr>. Citado na página 43.

WICKHAM, H.; VAUGHAN, D.; GIRLICH, M. *tidyr: Tidy Messy Data*. [S.l.], 2023. <https://tidyr.tidyverse.org>, <https://github.com/tidyverse/tidyr>. Citado na página 43.

Apêndices

APÊNDICE A – Seleção do número de agrupamentos

A seleção do número de agrupamentos foi realizado utilizando inspeção visualização das partições realizadas pelo algoritmo agrupamento aglomerativo hierárquico e também por uma série de métricas que procuram medir a qualidade de um determinado particionamento.

A [Figura 15](#) apresenta os particionamentos realizados no conjunto de dados até o limite de 12 agrupamentos. Essa figura é construída a partir da função `seqtreedisplay` do pacote `WeightedCluster` (STUDER, 2013) e utiliza em sua construção a árvore de agrupamentos, construída pela função `as.seqtree` também do pacote `WeightedCluster`.

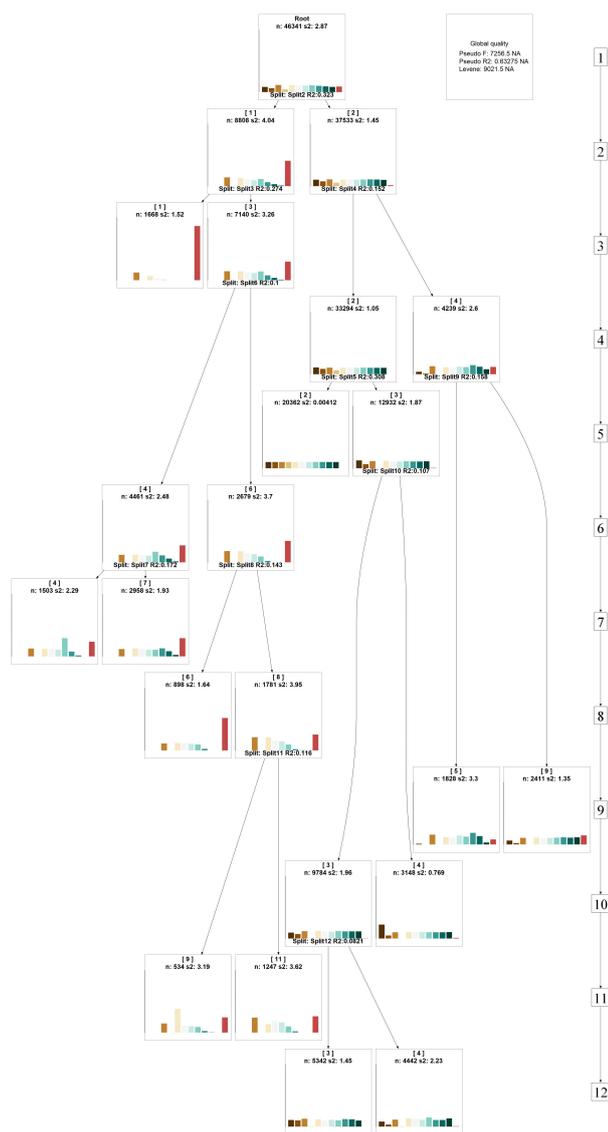
A [Figura 15](#) apresenta o tempo médio em cada um dos 12 estados considerados e em cada um dos agrupamentos resultantes das partições. O primeiro particionamento é considerado o mais importante (STUDER, 2013). Vê-se, portanto, que o primeiro particionamento separa aqueles estudantes que experimentaram tempo significativo fora da escola dos demais estudantes. O próximo particionamento mais importante separa os estudantes que experimentaram tempo significativo fora da escola entre aqueles que evadiram logo no início do período considerado daqueles que vivenciaram mais etapas de ensino. Os particionamentos se seguem até que sejam identificados os 12 agrupamentos definidos a priori.

As métricas de qualidade de uma partição utilizadas foram PBC (*Point Biserial Correlation*), HG (*Hubert's Gamma*), HC *Hubert's* e ASW *Average Silhouette Width*. Utiliza-se aqui as recomendações propostas por Studer (2013): deve-se procurar escolher agrupamentos que maximizem as métricas PBC, HG e ASW e minimizem a métrica HC.

Vê-se pela [Figura 16](#) que entre 2 e 4 agrupamentos tem-se uma maximização das métricas PBC e ASW, e também da métrica HG, apesar de que esta também é maximizada para valores maiores do número de agrupamentos. A métrica HC permanece em valores baixos e os pontos de máximo e mínimo não se diferem substancialmente. Além disso, as métricas não variam muito no intervalo entre 5 e 10 agrupamentos.

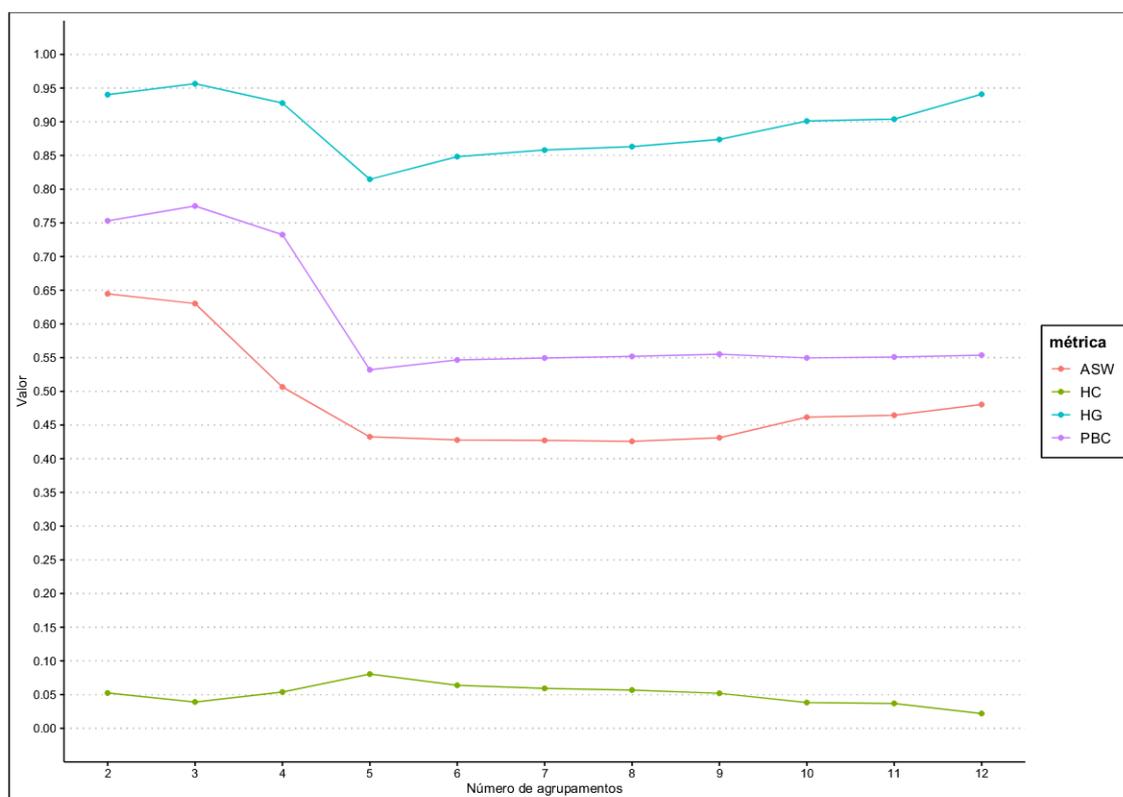
Portanto, a decisão do número de agrupamentos a ser analisado envolveu a inspeção visual dos tempos médios em cada agrupamento na [Figura 15](#), as métricas apresentadas na [Figura 16](#) e a baixa variação apresentada, e também o número de observações em cada um dos agrupamentos. Escolheu-se seis agrupamentos para análise.

Figura 15 – Árvore de agrupamentos produzido pela função *seqtreedisplay*



Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Figura 16 – Medidas da qualidade de uma partição - PBC, HG, HC e ASW - 12 agrupamentos



Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Anexos

ANEXO A – Tabelas

Tabela 9 – Distribuição relativa dos estados por ano - Coorte

Estado	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
2º ano E.F.	100,0%	10,7%	2,6%	0,7%	0,2%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
3º ano E.F.	0,0%	88,0%	15,7%	5,3%	1,9%	0,7%	0,3%	0,1%	0,0%	0,0%	0,0%
4º ano E.F.	0,0%	0,0%	79,5%	15,4%	5,9%	2,4%	1,0%	0,4%	0,1%	0,0%	0,0%
5º ano E.F.	0,0%	0,0%	0,0%	75,7%	16,1%	6,5%	2,8%	1,3%	0,5%	0,2%	0,1%
6º ano E.F.	0,0%	0,0%	0,0%	0,0%	72,3%	18,9%	9,4%	5,1%	2,0%	0,7%	0,2%
7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	66,9%	18,0%	9,1%	4,2%	1,5%	0,5%
8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	62,8%	17,1%	7,6%	2,9%	0,9%
9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	59,1%	16,1%	6,7%	2,2%
1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	55,2%	19,1%	9,1%
2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	48,1%	15,2%
3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	44,5%
Fora da escola	0,0%	1,3%	2,2%	2,9%	3,7%	4,5%	5,6%	7,8%	14,2%	20,9%	27,3%

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Tabela 10 – Distribuição relativa dos estados por ano e sexo - Coorte

Sexo	Estado	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Feminino	2º ano E.F.	100,0%	8,7%	1,8%	0,4%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Feminino	3º ano E.F.	0,0%	90,2%	12,9%	3,8%	1,2%	0,4%	0,1%	0,1%	0,0%	0,0%	0,0%
Feminino	4º ano E.F.	0,0%	0,0%	83,4%	13,1%	4,3%	1,6%	0,6%	0,2%	0,1%	0,0%	0,0%
Feminino	5º ano E.F.	0,0%	0,0%	0,0%	80,2%	13,9%	4,8%	1,9%	0,7%	0,3%	0,1%	0,0%
Feminino	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	77,3%	16,0%	6,8%	3,2%	1,1%	0,4%	0,1%
Feminino	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	73,2%	16,2%	7,1%	2,9%	1,0%	0,3%
Feminino	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	69,4%	15,8%	6,2%	2,1%	0,6%
Feminino	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	65,7%	14,9%	5,5%	1,7%
Feminino	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	61,8%	17,4%	7,2%
Feminino	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	55,2%	14,4%
Feminino	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	51,5%
Feminino	Fora da escola	0,0%	1,1%	2,0%	2,5%	3,2%	3,9%	5,0%	7,1%	12,6%	18,3%	24,1%
Masculino	2º ano E.F.	100,0%	12,7%	3,3%	0,9%	0,3%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
Masculino	3º ano E.F.	0,0%	85,9%	18,3%	6,8%	2,6%	1,0%	0,4%	0,2%	0,0%	0,0%	0,0%
Masculino	4º ano E.F.	0,0%	0,0%	75,8%	17,5%	7,4%	3,3%	1,5%	0,7%	0,2%	0,1%	0,0%
Masculino	5º ano E.F.	0,0%	0,0%	0,0%	71,4%	18,2%	8,1%	3,8%	1,8%	0,6%	0,2%	0,1%
Masculino	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	67,4%	21,7%	11,8%	6,9%	2,8%	1,0%	0,3%
Masculino	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	60,9%	19,8%	10,9%	5,3%	2,0%	0,7%
Masculino	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	56,5%	18,4%	9,0%	3,6%	1,2%
Masculino	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	52,7%	17,2%	7,9%	2,7%
Masculino	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	48,9%	20,7%	10,9%
Masculino	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	41,2%	15,9%
Masculino	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	37,7%
Masculino	Fora da escola	0,0%	1,4%	2,5%	3,3%	4,1%	5,0%	6,2%	8,4%	15,7%	23,3%	30,4%

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Tabela 11 – Distribuição relativa dos estados por ano e cor/raça - Coorte

Cor/Raça	Estado	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Não declarada	2º ano E.F.	100,0%	12,7%	3,3%	0,9%	0,2%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
Não declarada	3º ano E.F.	0,0%	85,6%	17,1%	6,2%	2,3%	0,9%	0,3%	0,1%	0,0%	0,0%	0,0%
Não declarada	4º ano E.F.	0,0%	0,0%	76,5%	16,9%	7,0%	3,0%	1,3%	0,6%	0,2%	0,1%	0,0%
Não declarada	5º ano E.F.	0,0%	0,0%	0,0%	72,1%	17,1%	7,5%	3,5%	1,7%	0,6%	0,2%	0,1%
Não declarada	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	68,4%	19,7%	10,7%	6,1%	2,5%	0,9%	0,3%
Não declarada	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	62,8%	18,3%	10,0%	4,8%	1,8%	0,6%
Não declarada	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	58,5%	16,8%	8,0%	3,2%	1,0%
Não declarada	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	54,8%	15,5%	6,7%	2,4%
Não declarada	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	50,7%	18,0%	8,5%
Não declarada	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	43,9%	14,3%
Não declarada	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	40,4%
Não declarada	Fora da escola	0,0%	1,8%	3,1%	3,9%	4,9%	5,9%	7,3%	9,9%	17,6%	25,2%	32,3%
Branca	2º ano E.F.	100,0%	6,4%	1,3%	0,3%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Branca	3º ano E.F.	0,0%	93,0%	10,1%	2,7%	0,8%	0,3%	0,1%	0,0%	0,0%	0,0%	0,0%
Branca	4º ano E.F.	0,0%	0,0%	87,6%	10,5%	3,1%	1,1%	0,4%	0,2%	0,1%	0,0%	0,0%
Branca	5º ano E.F.	0,0%	0,0%	0,0%	85,1%	11,6%	3,7%	1,4%	0,5%	0,2%	0,1%	0,0%
Branca	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	82,5%	14,2%	5,7%	2,7%	1,0%	0,3%	0,1%
Branca	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	78,4%	14,6%	6,1%	2,5%	0,8%	0,2%
Branca	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	74,9%	14,6%	5,6%	1,9%	0,5%
Branca	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	71,4%	14,4%	5,2%	1,5%
Branca	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	67,7%	18,0%	7,7%
Branca	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	60,4%	14,7%
Branca	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	56,6%
Branca	Fora da escola	0,0%	0,6%	1,1%	1,4%	1,8%	2,2%	3,0%	4,5%	8,7%	13,3%	18,5%
Preta	2º ano E.F.	100,0%	12,7%	3,4%	1,0%	0,3%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
Preta	3º ano E.F.	0,0%	85,9%	20,3%	7,5%	2,8%	1,1%	0,5%	0,2%	0,1%	0,0%	0,0%
Preta	4º ano E.F.	0,0%	0,0%	73,9%	19,0%	7,8%	3,3%	1,5%	0,7%	0,2%	0,1%	0,0%
Preta	5º ano E.F.	0,0%	0,0%	0,0%	69,4%	20,1%	8,6%	3,7%	1,7%	0,6%	0,2%	0,1%
Preta	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	65,1%	23,8%	12,6%	7,2%	3,0%	1,0%	0,3%
Preta	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	58,2%	22,0%	11,9%	5,7%	2,2%	0,7%
Preta	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	53,5%	20,5%	10,1%	3,9%	1,4%
Preta	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	49,2%	18,6%	8,7%	3,0%
Preta	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	45,6%	22,1%	12,5%
Preta	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	37,9%	16,9%
Preta	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	34,3%
Preta	Fora da escola	0,0%	1,4%	2,4%	3,1%	3,9%	4,9%	6,1%	8,6%	16,1%	24,0%	30,8%
Parda	2º ano E.F.	100,0%	12,3%	3,0%	0,8%	0,2%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
Parda	3º ano E.F.	0,0%	86,3%	18,3%	6,4%	2,3%	0,9%	0,3%	0,1%	0,0%	0,0%	0,0%
Parda	4º ano E.F.	0,0%	0,0%	76,3%	17,6%	6,9%	2,9%	1,2%	0,5%	0,2%	0,1%	0,0%
Parda	5º ano E.F.	0,0%	0,0%	0,0%	72,0%	18,3%	7,6%	3,3%	1,5%	0,5%	0,2%	0,1%
Parda	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	68,3%	21,4%	10,8%	5,9%	2,4%	0,8%	0,2%
Parda	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	62,3%	20,2%	10,4%	4,8%	1,8%	0,6%
Parda	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	58,0%	19,1%	8,8%	3,3%	1,1%
Parda	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	54,1%	17,6%	7,6%	2,6%
Parda	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	50,2%	20,4%	10,2%
Parda	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	42,9%	16,0%
Parda	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	39,4%
Parda	Fora da escola	0,0%	1,4%	2,4%	3,2%	4,0%	4,8%	6,1%	8,5%	15,5%	22,9%	29,8%
Amarela	2º ano E.F.	100,0%	9,8%	2,4%	0,5%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Amarela	3º ano E.F.	0,0%	88,6%	12,9%	4,7%	1,7%	0,7%	0,3%	0,1%	0,0%	0,0%	0,0%
Amarela	4º ano E.F.	0,0%	0,0%	82,1%	12,6%	5,0%	2,1%	1,1%	0,5%	0,2%	0,1%	0,0%
Amarela	5º ano E.F.	0,0%	0,0%	0,0%	78,9%	13,5%	5,3%	2,2%	1,3%	0,5%	0,2%	0,1%
Amarela	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	75,5%	15,7%	7,7%	3,9%	1,7%	0,7%	0,2%
Amarela	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	71,0%	15,2%	7,1%	3,2%	1,3%	0,5%
Amarela	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	67,1%	15,1%	6,1%	2,0%	0,6%
Amarela	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	63,6%	13,3%	5,2%	1,8%
Amarela	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	60,2%	16,6%	7,3%
Amarela	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	53,4%	12,7%
Amarela	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	50,2%
Amarela	Fora da escola	0,0%	1,6%	2,6%	3,3%	4,3%	5,2%	6,4%	8,4%	14,8%	20,6%	26,5%
Indígena	2º ano E.F.	100,0%	26,0%	7,8%	2,6%	0,7%	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%
Indígena	3º ano E.F.	0,0%	68,0%	29,5%	12,2%	4,9%	1,7%	0,6%	0,2%	0,1%	0,0%	0,0%
Indígena	4º ano E.F.	0,0%	0,0%	52,0%	25,8%	13,0%	6,1%	2,5%	1,1%	0,3%	0,1%	0,1%
Indígena	5º ano E.F.	0,0%	0,0%	0,0%	46,0%	24,8%	13,7%	7,0%	3,2%	1,4%	0,6%	0,3%
Indígena	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	40,8%	23,2%	14,4%	8,6%	3,9%	1,6%	0,6%
Indígena	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	36,3%	19,8%	12,6%	7,0%	3,5%	1,7%
Indígena	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	33,7%	17,6%	10,2%	5,4%	2,6%
Indígena	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	31,2%	16,3%	8,6%	4,2%
Indígena	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	28,1%	16,5%	8,7%
Indígena	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	23,7%	13,4%

Tabela 12 – Distribuição relativa dos estados por ano e quintil de NSE - Coorte

Quintil de NSE	Estado	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
1º Quintil	2º ano E.F.	100,0%	23,0%	6,6%	1,8%	0,5%	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%
1º Quintil	3º ano E.F.	0,0%	74,2%	26,7%	11,5%	4,4%	1,8%	0,7%	0,3%	0,1%	0,0%	0,0%
1º Quintil	4º ano E.F.	0,0%	0,0%	61,8%	24,7%	12,3%	5,6%	2,7%	1,3%	0,5%	0,2%	0,1%
1º Quintil	5º ano E.F.	0,0%	0,0%	0,0%	56,0%	23,8%	13,0%	6,7%	3,5%	1,5%	0,6%	0,2%
1º Quintil	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	51,6%	25,1%	16,0%	9,8%	4,7%	2,0%	0,7%
1º Quintil	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	45,4%	21,6%	13,6%	7,3%	3,4%	1,4%
1º Quintil	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	41,6%	19,2%	10,5%	4,9%	2,0%
1º Quintil	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	38,6%	16,6%	8,6%	3,7%
1º Quintil	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	35,4%	16,8%	8,8%
1º Quintil	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	30,5%	13,3%
1º Quintil	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	28,0%
1º Quintil	Fora da escola	0,0%	2,8%	4,8%	6,0%	7,4%	8,8%	10,6%	13,7%	23,4%	33,0%	41,8%
2º Quintil	2º ano E.F.	100,0%	13,5%	3,1%	0,8%	0,2%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
2º Quintil	3º ano E.F.	0,0%	84,9%	20,3%	7,0%	2,6%	1,0%	0,4%	0,1%	0,0%	0,0%	0,0%
2º Quintil	4º ano E.F.	0,0%	0,0%	73,8%	19,7%	7,6%	3,2%	1,3%	0,5%	0,1%	0,0%	0,0%
2º Quintil	5º ano E.F.	0,0%	0,0%	0,0%	68,9%	20,1%	8,5%	3,7%	1,6%	0,5%	0,1%	0,0%
2º Quintil	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	64,9%	23,5%	12,5%	6,8%	2,6%	0,7%	0,2%
2º Quintil	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	58,1%	21,6%	11,8%	5,5%	2,0%	0,6%
2º Quintil	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	53,2%	19,8%	9,5%	3,6%	1,1%
2º Quintil	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	49,1%	17,7%	8,1%	2,9%
2º Quintil	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	45,0%	20,0%	10,2%
2º Quintil	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	38,0%	15,5%
2º Quintil	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	34,6%
2º Quintil	Fora da escola	0,0%	1,6%	2,8%	3,6%	4,6%	5,7%	7,3%	10,2%	19,0%	27,4%	35,0%
3º Quintil	2º ano E.F.	100,0%	8,6%	1,8%	0,5%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
3º Quintil	3º ano E.F.	0,0%	90,5%	15,6%	4,7%	1,5%	0,5%	0,2%	0,1%	0,0%	0,0%	0,0%
3º Quintil	4º ano E.F.	0,0%	0,0%	80,9%	15,7%	5,3%	2,0%	0,7%	0,3%	0,1%	0,0%	0,0%
3º Quintil	5º ano E.F.	0,0%	0,0%	0,0%	76,9%	16,9%	5,9%	2,2%	0,8%	0,2%	0,1%	0,0%
3º Quintil	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	73,2%	20,5%	9,2%	4,6%	1,5%	0,4%	0,1%
3º Quintil	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	67,4%	19,8%	9,4%	4,0%	1,2%	0,3%
3º Quintil	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	63,1%	18,9%	8,3%	2,8%	0,8%
3º Quintil	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	58,7%	17,6%	7,3%	2,2%
3º Quintil	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	54,3%	21,1%	10,6%
3º Quintil	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	45,9%	16,1%
3º Quintil	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	42,0%
3º Quintil	Fora da escola	0,0%	0,9%	1,7%	2,2%	2,9%	3,6%	4,8%	7,3%	14,0%	21,3%	28,0%
4º Quintil	2º ano E.F.	100,0%	5,8%	1,1%	0,3%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
4º Quintil	3º ano E.F.	0,0%	93,7%	10,6%	2,6%	0,8%	0,2%	0,1%	0,0%	0,0%	0,0%	0,0%
4º Quintil	4º ano E.F.	0,0%	0,0%	87,4%	11,1%	3,2%	1,0%	0,3%	0,1%	0,0%	0,0%	0,0%
4º Quintil	5º ano E.F.	0,0%	0,0%	0,0%	84,6%	12,9%	3,7%	1,2%	0,4%	0,1%	0,0%	0,0%
4º Quintil	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	81,4%	16,1%	6,3%	3,0%	0,9%	0,2%	0,0%
4º Quintil	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	76,7%	16,4%	6,9%	2,7%	0,7%	0,2%
4º Quintil	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	72,7%	16,3%	6,3%	2,0%	0,5%
4º Quintil	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	68,6%	16,3%	5,9%	1,6%
4º Quintil	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	64,2%	20,9%	9,8%
4º Quintil	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	55,4%	16,4%
4º Quintil	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	51,2%
4º Quintil	Fora da escola	0,0%	0,6%	1,0%	1,3%	1,7%	2,2%	2,9%	4,7%	9,4%	14,8%	20,3%
5º Quintil	2º ano E.F.	100,0%	2,9%	0,4%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
5º Quintil	3º ano E.F.	0,0%	96,5%	5,1%	0,8%	0,2%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
5º Quintil	4º ano E.F.	0,0%	0,0%	93,6%	5,7%	1,1%	0,3%	0,1%	0,0%	0,0%	0,0%	0,0%
5º Quintil	5º ano E.F.	0,0%	0,0%	0,0%	92,1%	6,6%	1,5%	0,4%	0,1%	0,0%	0,0%	0,0%
5º Quintil	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	90,4%	9,3%	2,8%	1,1%	0,3%	0,1%	0,0%
5º Quintil	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	86,9%	10,7%	3,6%	1,2%	0,3%	0,1%
5º Quintil	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	83,6%	11,7%	3,6%	1,0%	0,2%
5º Quintil	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	80,4%	12,2%	3,7%	0,8%
5º Quintil	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	77,2%	16,6%	6,1%
5º Quintil	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	70,5%	14,5%
5º Quintil	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	66,8%
5º Quintil	Fora da escola	0,0%	0,6%	1,0%	1,3%	1,7%	2,0%	2,4%	3,1%	5,4%	7,8%	11,5%

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Tabela 13 – 10 sequências mais comuns - Coorte

id	Frequência	%	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
1	886918	43,57	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.	3ª série E.M.
2	68514	3,37	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	1ª série E.M.	2ª série E.M.
3	48042	2,36	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.	2ª série E.M.
4	32136	1,58	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	1ª série E.M.	1ª série E.M.
5	29249	1,44	2º ano E.F.	3º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.
6	28000	1,38	2º ano E.F.	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.
7	26367	1,30	2º ano E.F.	Fora da escola									
8	24544	1,21	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.	Fora da escola
9	24452	1,20	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	Fora da escola	Fora da escola	Fora da escola
10	22707	1,12	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Tabela 14 – 10 sequências mais comuns por sexo - Coorte

Sexo	id	Frequência	%	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Feminino	1	503408	50,52	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.	3ª série E.M.
Feminino	2	305600	3,07	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	1ª série E.M.	2ª série E.M.
Feminino	3	21927	2,20	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.	2ª série E.M.
Feminino	4	14475	1,45	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.	Fora da escola
Feminino	5	14342	1,44	2º ano E.F.	3º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.
Feminino	6	14327	1,44	2º ano E.F.	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.
Feminino	7	14106	1,42	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	Fora da escola	Fora da escola	Fora da escola
Feminino	8	12766	1,28	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	Fora da escola	Fora da escola
Feminino	9	12425	1,25	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	1ª série E.M.	1ª série E.M.
Feminino	10	11318	1,14	2º ano E.F.	Fora da escola									
Masculino	1	383510	36,91	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.	3ª série E.M.
Masculino	2	37954	3,65	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	1ª série E.M.	2ª série E.M.
Masculino	3	26115	2,51	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.	2ª série E.M.
Masculino	4	19711	1,90	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	1ª série E.M.	1ª série E.M.
Masculino	5	15049	1,45	2º ano E.F.	Fora da escola									
Masculino	6	14907	1,43	2º ano E.F.	3º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.
Masculino	7	13673	1,32	2º ano E.F.	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.
Masculino	8	12116	1,17	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.
Masculino	9	11918	1,15	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.
Masculino	10	11764	1,13	2º ano E.F.	3º ano E.F.	4º ano E.F.	5º ano E.F.	6º ano E.F.	7º ano E.F.	8º ano E.F.	9º ano E.F.	9º ano E.F.	1ª série E.M.	2ª série E.M.

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

Tabela 17 – Distribuição relativa dos estados por ano - Padrões identificados - Amostra

Padrão	Estado	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
evasão precoce	2º ano E.F.	100,0%	19,8%	4,3%	0,6%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	3º ano E.F.	0,0%	45,5%	15,1%	4,6%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	4º ano E.F.	0,0%	0,0%	20,8%	5,5%	1,5%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	5º ano E.F.	0,0%	0,0%	0,0%	10,8%	1,9%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão precoce	Fora da escola	0,0%	34,7%	59,8%	78,5%	96,6%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
regular	2º ano E.F.	100,0%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
regular	3º ano E.F.	0,0%	99,9%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
regular	4º ano E.F.	0,0%	0,0%	99,9%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
regular	5º ano E.F.	0,0%	0,0%	0,0%	99,9%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
regular	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	99,9%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
regular	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	99,9%	0,0%	0,0%	0,0%	0,0%	0,0%
regular	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	99,9%	0,0%	0,0%	0,0%	0,0%
regular	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	100,0%	0,1%	0,0%	0,0%
regular	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	99,9%	0,0%	0,0%
regular	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	100,0%	0,0%
regular	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	100,0%
regular	Fora da escola	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
pouca irregularidade	2º ano E.F.	100,0%	13,9%	0,7%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
pouca irregularidade	3º ano E.F.	0,0%	86,1%	26,3%	4,9%	0,8%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
pouca irregularidade	4º ano E.F.	0,0%	0,0%	73,0%	28,3%	6,6%	1,5%	0,2%	0,0%	0,0%	0,0%	0,0%
pouca irregularidade	5º ano E.F.	0,0%	0,0%	0,0%	66,7%	30,8%	7,3%	2,0%	0,3%	0,0%	0,0%	0,0%
pouca irregularidade	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	61,8%	38,1%	11,5%	3,8%	0,8%	0,1%	0,0%
pouca irregularidade	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	53,0%	39,5%	12,9%	4,4%	0,8%	0,1%
pouca irregularidade	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	46,8%	42,9%	14,6%	4,8%	1,2%
pouca irregularidade	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	40,1%	47,1%	18,6%	5,8%
pouca irregularidade	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	33,1%	64,0%	29,9%
pouca irregularidade	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	11,7%	54,3%
pouca irregularidade	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	3,2%
pouca irregularidade	Fora da escola	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	5,5%
evasão - anos finais	2º ano E.F.	100,0%	20,1%	0,9%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão - anos finais	3º ano E.F.	0,0%	79,9%	35,5%	8,8%	0,6%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão - anos finais	4º ano E.F.	0,0%	0,0%	63,6%	38,5%	14,1%	1,6%	0,1%	0,0%	0,0%	0,0%	0,0%
evasão - anos finais	5º ano E.F.	0,0%	0,0%	0,0%	52,7%	35,7%	15,4%	2,1%	0,1%	0,0%	0,0%	0,0%
evasão - anos finais	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	49,6%	49,5%	38,5%	20,0%	7,8%	2,8%	0,6%
evasão - anos finais	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	33,5%	34,7%	29,8%	10,4%	2,7%	0,9%
evasão - anos finais	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	24,6%	21,9%	11,8%	2,3%	0,4%
evasão - anos finais	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	13,6%	2,4%	0,1%	0,2%
evasão - anos finais	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,3%	0,2%	0,2%
evasão - anos finais	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão - anos finais	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão - anos finais	Fora da escola	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	14,6%	67,3%	91,9%	97,7%
evasão - ensino médio	2º ano E.F.	100,0%	18,4%	7,7%	1,8%	0,3%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão - ensino médio	3º ano E.F.	0,0%	81,6%	20,6%	7,7%	2,4%	0,5%	0,1%	0,0%	0,0%	0,0%	0,0%
evasão - ensino médio	4º ano E.F.	0,0%	0,0%	71,7%	19,7%	8,1%	2,7%	0,7%	0,1%	0,0%	0,0%	0,0%
evasão - ensino médio	5º ano E.F.	0,0%	0,0%	0,0%	70,9%	27,9%	14,4%	5,7%	1,5%	0,5%	0,1%	0,0%
evasão - ensino médio	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	61,4%	27,7%	15,6%	8,1%	2,5%	0,4%	0,1%
evasão - ensino médio	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	54,6%	34,9%	25,3%	18,2%	9,4%	2,9%
evasão - ensino médio	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	43,1%	31,7%	25,0%	14,6%	5,9%
evasão - ensino médio	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	33,4%	26,6%	14,0%	5,9%
evasão - ensino médio	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	26,6%	10,0%	5,1%
evasão - ensino médio	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	14,0%	0,0%
evasão - ensino médio	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão - ensino médio	Fora da escola	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,5%	37,5%	79,9%
evasão - anos iniciais	2º ano E.F.	100,0%	42,0%	23,9%	8,8%	2,8%	1,0%	0,4%	0,1%	0,1%	0,0%	0,0%
evasão - anos iniciais	3º ano E.F.	0,0%	58,0%	40,9%	37,4%	24,1%	10,7%	4,5%	1,8%	0,8%	0,2%	0,0%
evasão - anos iniciais	4º ano E.F.	0,0%	0,0%	35,2%	27,9%	30,5%	27,4%	15,5%	6,5%	2,6%	1,2%	0,3%
evasão - anos iniciais	5º ano E.F.	0,0%	0,0%	0,0%	25,9%	23,8%	26,6%	27,0%	18,0%	6,8%	2,5%	0,7%
evasão - anos iniciais	6º ano E.F.	0,0%	0,0%	0,0%	0,0%	17,8%	14,6%	18,0%	23,1%	13,3%	5,7%	2,2%
evasão - anos iniciais	7º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	6,8%	2,5%	4,0%	5,4%	3,1%	1,5%
evasão - anos iniciais	8º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	0,6%	0,5%	0,5%
evasão - anos iniciais	9º ano E.F.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,3%	0,1%	0,1%	0,2%
evasão - anos iniciais	1ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,2%	0,1%	0,1%
evasão - anos iniciais	2ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão - anos iniciais	3ª série E.M.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
evasão - anos iniciais	Fora da escola	0,0%	0,0%	0,0%	0,0%	1,0%	13,0%	32,1%	45,9%	70,2%	86,6%	94,4%

Fonte: Base Longitudinal do Censo Escolar 2007-2018 (INEP, 2017)

ANEXO B – Scripts utilizados

Os códigos em desenvolvidos em linguagem R (R Core Team, 2021) estão disponíveis em repositório público no GitHub na seguinte página da Web: <https://github.com/estevaovilela/dissertacao_mestrado_demografia>