

# Participação de estudantes de escolas públicas estaduais de Porto Alegre no Enem: uma análise de *cluster*

---

HARLON ROMARIZ RABELO SANTOS<sup>I</sup>

BÁRBARA DONIDA DE BONA<sup>II</sup>

LOURENÇO BRITO FELIN<sup>III</sup>

CÉLIA ELIZABETE CAREGNATO<sup>IV</sup>

<http://dx.doi.org/10.22347/2175-2753v13i41.3565>

## Resumo

Este artigo identifica características de escolas públicas estaduais de ensino médio da cidade de Porto Alegre com o objetivo de entender variações e estratificações entre essas instituições que se relacionam com a taxa de participação no Exame Nacional do Ensino Médio (Enem). Essa investigação busca ampliar a compreensão sobre os fatores que se relacionam à desigualdade educacional brasileira. Foram analisados dados estruturais oferecidos pelo Censo Escolar do Inep. A taxa de participação no Enem 2018 foi elencada como variável dependente. Buscaram-se agrupamentos de escolas que se diferenciam entre si e como ocorre esta variação. Foi possível expor dois *clusters* que mostram que uma maior participação de estudantes no Enem por escola está relacionada a menor distorção idade-série, menor complexidade da gestão e maior índice socioeconômico.

**Palavras-chave:** Escolas públicas. Ensino médio. Enem. Estratificação. Análise de *cluster*.

Submetido em: 17/06/2021

Aprovado em: 19/11/2021

---

<sup>I</sup> Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre (RS), Brasil; <http://orcid.org/0000-0002-5642-0448>; e-mail: harlon.romariz@gmail.com.

<sup>II</sup> Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre (RS), Brasil; <http://orcid.org/0000-0001-6208-7007>; e-mail: barbara.donida@gmail.com.

<sup>III</sup> Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre (RS), Brasil; <http://orcid.org/0000-0001-5764-9693>; e-mail: lofelin@gmail.com.

<sup>IV</sup> Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre (RS), Brasil; <http://orcid.org/0000-0002-9326-590X>; e-mail: celia.caregnato@gmail.com.

# Participation of students from public state school of Porto Alegre in Enem: a cluster analysis

## **Abstract**

This article identifies characteristics of state public high schools in the city of Porto Alegre in order to understand variations and stratifications among these institutions that are related to the participation rate in the National High School Exam (Enem). This investigation seeks to broaden the understanding about the factors of that relate to Brazilian educational inequality. Structural data provided by the Inep's School Census were analyzed. The participation rate in Enem 2018 was listed as a dependent variable. Clusters of schools that differ from each other and how this variation occurs were sought. It was possible to expose two clusters showing that a greater participation of students in Enem per school is related to lower age-grade distortion, lower management complexity and higher socioeconomic index.

**Keywords:** Public schools. High school. Enem. Stratification. Cluster analysis.

# Participación de estudiantes de escuelas públicas estatales de Porto Alegre en Enem: un análisis de *cluster*

## **Resumen**

Este artículo identifica características de las escuelas secundarias públicas estatales de la ciudad de Porto Alegre con el fin de comprender las variaciones y estratificaciones entre estas instituciones que se relacionan con la tasa de participación en el Examen Nacional de Educación Secundaria (Enem). Esta investigación busca ampliar la comprensión de los factores relacionados con la desigualdad educativa brasileña. Se analizaron los datos estructurales proporcionados por el Censo Escolar del Inep. Se definió la tasa de participación en Enem 2018 como variable dependiente. Se buscaron grupos de escuelas que se diferencian entre sí y cómo ocurre esta variación. Se pudo exponer dos *clusters* que muestran que una mayor participación de estudiantes en Enem por escuela se relaciona con menor distorsión edad-grado, menor complejidad en la gestión y mayor índice socioeconómico.

**Palabras clave:** Escuelas públicas. Escuela secundaria. Enem. Estratificación. Análisis de *cluster*.

## Introdução

A rede pública de educação básica no Brasil atingiu a universalização do acesso a vagas no ensino fundamental apenas nos anos 1990. Na década de 2000, ocorreu a ampliação de vagas em outros níveis da educação formal, o ensino médio e a educação superior. É na fase de escolarização média que os jovens precisam desenvolver habilidades e recursos para se situar na fase adulta, progredindo para ocupações no mundo do trabalho, em estudos avançados, ou em ambas as áreas de atividade. A conclusão do ensino médio ou equivalente é um fator significativo para a entrada na maior parte dos postos de trabalho cuja remuneração permite um maior retorno do ponto de vista material. Ademais, os direitos cidadãos só são plenamente acessíveis – isto é, compreensíveis e demandáveis – a partir do domínio de códigos que, na vida escolar, pretende-se que sejam atingidos com a conclusão do ensino médio.

A Emenda Constitucional nº 59 de 2009 (BRASIL, 2009) consolidou legalmente essa expansão ao estabelecer a obrigatoriedade da escolarização para todos os cidadãos entre os 4 e os 17 anos a partir do ano de 2016. Ainda que essa disposição inclua a idade-alvo do ensino médio, 15 a 17 anos, a efetividade da matrícula de jovens nessa etapa é comprometida por condições precárias de permanência e sucesso na escola.

O Plano Nacional de Educação (PNE) 2014-2024 traz como sua Meta 3 “universalizar, até 2016, o atendimento escolar para toda a população de 15 (quinze) a 17 (dezessete) anos e elevar, até o final do período de vigência deste PNE, a taxa líquida de matrículas no ensino médio para 85% (oitenta e cinco por cento)” (BRASIL, 2014, não paginado). Esta meta, cuja parte inicial não foi atingida, defronta-se com as dificuldades de vinculação entre estudantes e sistema escolar, que se refletem em altos índices de evasão e com processos de escolarização precária para aqueles que conseguem se manter na escola. A ampla oferta de vagas no ensino médio ainda convive com importantes limitações, relativas a questões como distorção série-idade e currículos pouco efetivos.

Este artigo analisa aspectos que caracterizam as escolas públicas estaduais de ensino médio de Porto Alegre, com o objetivo de entender a existência de diferenciação entre elas ao verificar fatores de estratificação que podem incidir sobre a participação dos egressos no Exame Nacional do Ensino Médio (Enem). Ele

se situa no âmbito de pesquisa apoiada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)<sup>1</sup>, que empreende esforços para compreender projeções e possibilidades de jovens quanto à transição do ensino médio para a educação superior. Neste caso, queremos entender como ocorrem as variações quanto às condições e dinâmicas escolares a partir de características associadas às instituições, conforme registradas pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). A discussão baseia-se em dados oficiais, examinados de forma quantitativa, com análise de *clusters* relacionando variáveis escolares com a participação dos estudantes no referido exame. Como resultado, constata-se que uma maior participação de estudantes no Enem por escola está relacionada a menor distorção série-idade, menor complexidade da gestão e maior índice socioeconômico.

### **Expansão de oportunidades e reprodução social no ensino médio**

Períodos de expansão da escolarização, como o final do século XX e o início do século XXI no Brasil, indicam a tendência à melhoria da escolarização familiar. Os conhecimentos consolidados pela literatura sociológica da educação mostram que a educação pode servir como recurso para mobilidade social, mas a escola também pode facilitar a reprodução social ao valorizar os capitais cultivados pelas classes dominantes (BOURDIEU, 2010<sup>2</sup>; HAECHELT, 2008). A baixa renda de grande parte das famílias brasileiras exige muitas vezes que os jovens assumam desde cedo a condição de provedores no domicílio, impondo limites à escolarização, seja na trajetória longitudinal, seja na possibilidade de aprofundamento qualitativo da experiência escolar (DAYRELL; JESUS, 2016; CARRANO; MARINHO; OLIVEIRA, 2015). Segmentos de jovens têm possibilidades diferenciadas de fazerem uso da chamada moratória social, ou seja, de contarem com apoio familiar para se dedicarem exclusivamente a sua formação (GROPPO, 2015) para viver a formação escolar com razoáveis chances de sucesso. As tendências gerais convivem com casos fora

---

<sup>1</sup> A pesquisa que originou este artigo intitula-se *Socialização acadêmica em trajetórias estudantis: disposições, reflexividade e reconhecimento na educação superior*, tem o apoio do CNPq e é conduzida pelo Grupo de Pesquisa em Socialização, Estratificação e Trajetórias Juvenis e Educacionais (Geset).

<sup>2</sup> Texto publicado originalmente em 1966 na *Revue Française de Sociologie*.

da curva, como expõem estudos de sucessos improváveis (LAHIRE, 1997; ZAGO, 2000b).

Mesmo com a expansão, há estratificações e seletividades escolares, visíveis em trajetórias de jovens que, mesmo progredindo na escolarização, encontram dificuldades com interrupções, repetências ou desistência temporária (KRAWCZYK; SILVA, 2017), nas quais a organização da escola pode ser decisiva (SOARES; FERNANDES; NÓBREGA; NICOLELLA, 2015). Em geral, os indivíduos progridem no sistema na relação com trajetórias sociais de gerações anteriores e com base no investimento realizado pelas famílias. De acordo com estudos em sociologia da educação de matriz bourdieusiana, a exemplo de Zago (2000a) e Brandão e Carvalho (2011), sabe-se que a escolarização no Brasil ocorre com base em grupos selecionados pelo sistema educacional em interconexão com outros mecanismos de seleção social. Em *survey* realizado recentemente junto a estudantes de ensino médio em 14 escolas públicas na cidade de Porto Alegre, no ano de 2018, foi possível verificar que a maioria dos respondentes tinha pais ou responsáveis com escolarização no mínimo semelhante<sup>3</sup> ao nível no qual os jovens se encontravam, sinalizando a escolaridade de responsáveis como variável significativa na expectativa de vida escolar.

Para além das questões relativas às condições de vida dos indivíduos, cabe conhecer as condições das escolas que proporcionam a formação e que, como instituições públicas, medeiam as oportunidades dos cidadãos na sociedade. A aproximação ao ensino médio revela que essa etapa da educação nacional se encontra concentrada em uma categoria administrativa: a escola pública estadual. De acordo com o Censo Escolar do ano de 2020, as redes estaduais têm uma participação de 84,1% no total de matrículas do ensino médio e concentra 95,9% dos alunos da rede pública (INEP, 2021). Essas redes passaram a ser responsáveis por parcelas mais significativas do estudante de nível médio ao longo da década de 1990, quando concentravam 65,5% das matrículas (SILVA, 2020). O fato de que um percentual majoritário de estudantes está matriculado em escolas públicas estaduais torna pertinente colocar atenção sobre esse segmento de

---

<sup>3</sup> O *survey* foi conduzido como parte da pesquisa referida na nota anterior, contou com uma amostra representativa dos estudantes do terceiro ano de escolas públicas de Porto Alegre e mostrou que um percentual de 67,8% das mães e 67,5% dos pais atingiram o ensino médio de educação.

escolas. No Rio Grande do Sul, no ano de 2020, 75,53% das matrículas no ensino médio ocorreram em escolas públicas estaduais. Na cidade de Porto Alegre, essa fração foi de 68,85% (QEDU, [2021]). Em etapa anterior da pesquisa, verificamos que há uma desigualdade no acesso a vagas na maior universidade da cidade por egressos de escolas de ensino públicas diferenciadas, isto é, que praticam algum tipo de seleção para matrícula (CAREGNATO; SANTOS; FELIN, 2020). Percebemos também que, entre ingressantes oriundos de escolas estaduais, há uma concentração de egressos de escolas específicas.

Compreendemos que o ambiente escolar é fator importante que pode contribuir para aprimorar ou para limitar possibilidades de futuro educacional, profissional e político dos cidadãos. Tomamos o conceito de efeito escola como conjunto de resultados do trabalho escolar que pode ser visualizado no aprendizado do currículo da escola pelos estudantes, a partir de políticas e de práticas institucionais (SOARES; CANDIAN, 2007). Entretanto, há várias teorias e tipos de pesquisa a partir desse conceito (ALVES, 2010; ALVES; SOARES, 2013; BRESSOUX, 2003; DOWNEY; CONDRON, 2016), permitindo explorar diversas variáveis do trabalho escolar. Neste artigo, exploramos a participação dos estudantes no Enem relativamente à complexidade da gestão escolar. É possível ver afinidade entre a variável e o efeito do trabalho escolar, embora ela não se trate de uma relação excludente, já que a condição socioeconômica também é importante com base nos dados em análise.

Destacamos que os estímulos, suportes e desafios que os estudantes experimentam em sua socialização escolar são fundamentais para sua formação e desenvolvimento. Tais elementos, por sua vez, estão relacionados às condições da organização e gestão escolar e à infraestrutura, podendo propiciar maiores ou menores chances de aprofundar processos de ensino-aprendizagem. Para tanto, exploramos elementos que caracterizam diferencialmente as escolas públicas estaduais de ensino médio da cidade de Porto Alegre.

## **Metodologia**

Constituem o universo da pesquisa as escolas estaduais de ensino médio de Porto Alegre com cursos regulares e currículos de três anos de formação. A lista das escolas foi obtida no sítio oficial da Secretaria da Educação do Estado do Rio Grande do Sul (RIO GRANDE DO SUL, 2020). A população estudada é formada por

69 unidades, dentre as quais oito foram omitidas da análise devido à ausência de dados completos. Não foram incluídas escolas estaduais que oferecem exclusivamente educação de jovens e adultos (EJA) e escolas militares. Também foram excluídas as instituições com turmas de quarto ano<sup>4</sup>. Assim, chegou-se a uma amostra de 61 escolas. Esse recorte, limitado a Porto Alegre e não probabilístico, foi escolhido por favorecer maior controle analítico, considerando a proximidade e o conhecimento dos autores em relação ao contexto escolar local. Além disso, esse estudo teve caráter exploratório, descortinando possibilidades para uma agenda de pesquisa mais ampla, extensível inclusive a nível nacional e com uma complexidade modelar maior.

A metodologia empregada resume-se a dois procedimentos. Primeiramente, foi realizada uma regressão múltipla para analisar os coeficientes e a influência das variáveis selecionadas sobre a taxa de participação no Enem 2018. Em seguida, foi efetuada uma análise de *cluster* que pudesse agrupar as escolas conforme semelhanças em valores de variáveis influentes e significativas observadas no modelo ajustado final da regressão.

A análise busca verificar a influência de fatores escolares, em termos de efeito escola, sobre a taxa de participação do Enem. Tendo em vista a importância de condições socioeconômicas dos estudantes e suas famílias, foram adotadas variáveis de controle relativas a tais aspectos. As variáveis utilizadas são apresentadas no Quadro 1 e comentadas em seguida.

Quadro 1 – Apresentação das variáveis

Variável	Posição na relação	Tipo
Participação no Enem	Dependente	Contínua
Nível de complexidade da gestão	Independente	Ordinal
Adequação da formação	Independente	Contínua
Esforço docente	Independente	Contínua
Equipamentos	Independente	Contínua
Espaços	Independente	Discreta
Distorção idade-série	Independente	Contínua
Indicador de nível socioeconômico	Controle	Contínua
Proporção de mulheres	Controle	Contínua
Proporção de pretos e pardos	Controle	Contínua

Fonte: Os autores (2021).

<sup>4</sup> A taxa de participação no Enem não pode ser calculada de forma precisa em relação a escolas com turmas de quarto ano por não haver correspondência entre os microdados de matrícula e do Enem.

*Participação no Enem:* as porcentagens de participação no Enem por escola foram obtidas através da divisão entre a quantidade de estudantes concluintes que participaram da prova do Enem e a de estudantes do terceiro ano do ensino médio. Foram utilizados os microdados do Enem 2018 e os dados de matrículas disponíveis no Censo Escolar 2018 (INEP, 2019c)<sup>5</sup>.

*Nível de complexidade da gestão:* segundo nota técnica do Inep (2014c), a complexidade da gestão escolar é dividida em seis níveis. As unidades são agrupadas segundo o porte (quantidade de matrículas), número de etapas, complexidade das etapas (consideradas mais complexas quanto mais elevada a idade do alunado) e número de turnos de funcionamento.

*Adequação da formação:* o Inep separa os professores em cinco níveis de adequação da formação. O primeiro grupo conta com "docentes com formação superior de licenciatura na mesma disciplina que lecionam, ou bacharelado na mesma disciplina com curso de complementação pedagógica concluído" (INEP, 2014a, p. 5). A variável refere-se ao percentual desse grupo. Os dados são do ensino médio, baseados no Censo Escolar 2018 (INEP, 2019a).

*Esforço docente:* segundo nota técnica do Inep (2014b, p. [6]), o nível mais alto de esforço docente, formado por "docente que tem mais de 400 alunos e atua nos três turnos, em duas ou três escolas e em duas etapas ou três etapas". A variável refere-se ao percentual desse nível. Os dados são do ensino médio, baseados no Censo Escolar 2018 (INEP, 2019b).

*Equipamentos:* a fim de padronizar os dados sem associar pesos a intervalos, o índice de equipamentos é formado pela soma de dois escores Z. A padronização é aplicada às razões: a) quantidade de computadores disponíveis para utilização dos alunos e o número total de estudantes da escola; e b) quantidade de equipamentos multimídia presentes na escola e a quantidade de salas utilizadas como sala de aula. Os dados são do Censo Escolar 2018. O cálculo dos escores situa os valores resultantes das razões a certa quantidade de desvios padrões em relação à média. O escore Z de um caso da primeira razão ( $Z_a$ ) é dado através da subtração da média de  $a$  ( $\mu$ ) do valor do caso ( $x$ ), dividido pelo desvio-padrão de  $a$

---

<sup>5</sup> Originalmente do Inep, a variável deixou de ser disponibilizada em 2015.

( $\sigma$ ). O escore  $Z$  de casos da segunda razão ( $Z_b$ ) é obtido pelo mesmo tipo de cálculo.

*Espaços*: com valores de 0 a 5, a variável conta a quantidade de espaços presentes na escola dentre o seguinte conjunto: sala de leitura, biblioteca, quadra de esporte coberta ou descoberta, laboratório de informática e laboratório de ciências<sup>6</sup>. Os dados provêm do Censo Escolar de 2018.

*Distorção idade-série*: porcentagem de alunos com idade de dois ou mais anos superior à recomendada para a série em questão. Os dados são das turmas de terceiro ano, baseados no Censo Escolar 2018 (INEP, 2019d).

*Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse)*: índice produzido pelo Inep nos anos de 2011, 2013 e 2015. Foram utilizados os dados de 2015 (INEP, 2016), baseados nas provas do Sistema de Avaliação da Educação Básica (Saeb) (2015 e 2018) e do Enem (2015 e 2018). Segundo a nota técnica veiculada pelo Inep (2015), o cálculo do indicador é feito através da análise dos itens das provas referentes aos bens que a família do estudante possui em casa e a formação dos pais.

*Proporção de mulheres*: porcentagem de estudantes do sexo feminino obtida através dos dados de matrículas disponíveis no Censo Escolar 2018.

*Proporção de pretos ou pardos*: porcentagem de estudantes autodeclarados pretos ou pardos obtida por meio dos dados de matrícula disponíveis no Censo Escolar 2018.

Outras variáveis foram cogitadas para a exploração da taxa de participação no Enem. As taxas de rendimento (reprovação, abandono e aprovação) foram consideradas a fim de expressar as trajetórias dos estudantes de ensino médio. Contudo, como os dados utilizados para construir tais taxas são endógenos à distorção idade-série, optou-se por empregar apenas por esta última variável. Matrículas totais da escola e matrículas totais do ensino médio poderiam evidenciar o tamanho da escola e aspectos da gestão escolar. Foram desconsiderados por sua endogeneidade em relação à classificação de complexidade da gestão já

---

<sup>6</sup> Para a construção do índice de espaços foram exploradas outras variáveis de infraestrutura, a saber: presença de *internet* e *internet* banda larga, bem como de pátio coberto ou descoberto, água tratada, energia elétrica e saneamento básico. Todos se mostraram constantemente presentes ou ausentes nas escolas de Porto Alegre, não apresentando, portanto, interesse estatístico em sua análise.

disponibilizada pelo Inep. Ademais, o baixo número de casos limita a quantidade de variáveis que podem ser inseridas no modelo.

Apesar da publicação dos resultados do Enem por escola ter sido descontinuada pelo Inep em 2017, ainda seria possível calcular a nota média no Enem por escola a partir dos microdados do Exame. Contudo, poucas escolas (11 entre as 61 escolas) tiveram mais de 50% e no mínimo 10 estudantes do terceiro ano do ensino médio fazendo a prova. Isso reduziria muito a amostra, dificultando a interpretação por falta de representatividade dos dados.

Após a seleção das variáveis, na primeira etapa foi executada uma regressão linear múltipla a fim de explorar a relação entre as variáveis explicativas da participação no Enem. Para a seleção do modelo final, empregou-se o método *stepwise* e se verificou a presença de multicolinearidade pelo fator de inflação da variância (aceitando variáveis quando  $< 4$ ). Foram mantidas na equação final as variáveis significativas (nível de significância  $\leq 0,05$ ). A análise foi realizada com o *software R*, versão 4.0.3, no ambiente RStudio.

Através da regressão, é possível avaliar o quanto a variação da dependente é influenciada pelas independentes. Também é observada a importância explicativa de uma variável para o modelo e a natureza das relações entre as independentes e a dependente (HAIR; BLACK; BABIN; ANDERSON; TATHAM, 2009). Conforme orientado pela literatura (HAIR; BLACK; BABIN; ANDERSON; TATHAM, 2009; AGRETI; FINLAY, 2012), foram considerados pressupostos da regressão múltipla: presença de relação linear entre a dependente e as independentes; ausência de multicolinearidade e endogeneidade crítica; ausência de sobreajuste; e presença de homoscedasticidade.

Na segunda etapa, foi realizada uma análise de agrupamentos hierárquica, técnica multivariada que aglomera casos de acordo com os valores que apresentam para as variáveis (HAIR; BLACK; BABIN; ANDERSON; TATHAM, 2009). As escolas são classificadas de forma a aproximar as mais semelhantes e afastar as diferentes, criando grande homogeneidade intragrupo e grande heterogeneidade extragrupo. A técnica foi escolhida por conta do seu uso na criação de taxonomias por meio da classificação de objetos, possibilitando avaliar as diferenças entre os grupos de escolas.

Hair, Black, Babin, Anderson e Tatham (2009) apresentam três críticas relacionadas à análise de agrupamentos. Primeiro, apontam a impossibilidade de realizar inferências de uma amostra para a população. Muitas soluções diferentes podem ser alcançadas devido a mudanças nos procedimentos do método, como redução ou aumento de casos e de variáveis. Segundo, problematizam a criação de agrupamentos independentemente da existência real de qualquer estrutura nos dados. O algoritmo sempre mostrará agrupamentos, mesmo que não faça sentido no mundo real. Com suporte conceitual e validação dos agrupamentos, é possível encontrar evidência real de agrupamentos em potencial. Terceiro, alertam para a impossibilidade de generalização, por ser uma técnica completamente dependente das variáveis. A eliminação ou adição de variáveis legítimas ou ilegítimas causa grandes diferenças no resultado, mais do que em outras técnicas multivariadas. Esses limites, entre outros fatores, levaram-nos a combinar a técnica da regressão com a análise de agrupamentos hierárquica, permitindo um resultado mais ampliado e analítico.

A análise foi executada através da linguagem de programação Python, com a biblioteca *scikit-learn* para *machine learning*, no ambiente de desenvolvimento *Jupyter Notebook*. Para os agrupamentos, foram empregadas somente as variáveis avaliadas como significativas na regressão linear multivariada. Para evitar atribuição não intencional de peso às variáveis, os dados foram padronizados através do escore Z antes da mensuração da distância entre os casos.

A fim de medir a dissimilaridade entre os casos, foi utilizada a distância euclidiana, ou seja, a distância em linha reta entre os casos. Em um plano cartesiano, posiciona-se os casos através das suas medidas nos eixos X e Y. Um triângulo-retângulo é traçado no plano, de forma que a linha reta entre os pontos seja a hipotenusa. Abaixo, expressamos a fórmula da Distância Euclidiana (DE) apresentada por Hair, Black, Babin, Anderson e Tatham (2009, p. 442) para dois casos com dois atributos, calculados pelo teorema de Pitágoras.

$$DE = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$$

Generalizando para todas as variáveis (i) e cada par de casos (p e q), resulta a seguinte fórmula:

$$DE = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

Com a matriz de similaridade calculada, decide-se como serão particionados os agrupamentos. Optou-se pelo método de Ward, uma forma hierárquica de classificação que gera um gráfico de dendrograma. Neste método, “a similaridade entre dois agrupamentos não é uma única medida de similaridade, mas a soma dos quadrados dentro dos agrupamentos feita sobre todas as variáveis” (HAIR; BLACK; BABIN; ANDERSON; TATHAM, 2009, p. 452). A cada iteração através dos casos, o método combina aqueles que apresentam menor aumento na variância intragrupo.

Para a decisão do número de grupos, foi levada em conta a distância entre eles, a qual pode ser avaliada através das linhas verticais no dendrograma. O algoritmo aglomerativo inicia com cada caso como um *cluster* e, a cada iteração, há a união entre os grupos segundo o método de Ward. As repetições cessam quando restar apenas um agrupamento. A medida da linha vertical no dendrograma mostra as semelhanças na aglomeração, de modo que, quanto maior a distância, maior a diferença entre os membros do grupo que estão abaixo da linha.

## Resultados e discussão

A análise descritiva dos dados evidenciou a baixa presença de equipamentos multimídia (vídeo-projetor e computadores) disponíveis para uso dos estudantes. Independentemente do número de estudantes, a maioria das escolas possui entre zero e 30 computadores. Das quatro escolas com mais de 80 computadores, três são voltadas para a oferta de educação profissional técnica. Quanto aos equipamentos multimídia, contam-se entre zero e 4 por escola, independentemente da quantidade de salas utilizadas como salas de aula. Os computadores não necessariamente têm acesso à *internet* e conexão de qualidade. O mesmo ocorre com os equipamentos multimídia nas salas.

Os dados de cor/raça apontam que duas a cada dez escolas da capital apresentam estudantes que não declararam cor/raça como a categoria mais numerosa. Segundo o Censo Demográfico 2010 do Instituto Brasileiro de Geografia e

Estatística (IBGE), 20,24% da população de Porto Alegre é formada por pretos ou pardos (IBGE, 2010). Há nove escolas dentro da amostra com porcentagem de estudantes pretos ou pardos superior a 35%, evidenciando uma concentração de jovens pardos e pretos proporcionalmente maior entre os autodeclarados entre os habitantes da cidade. Essa configuração deve ser levada em consideração ao interpretar os resultados da regressão, ligados à situação das escolas da cidade. Outras capitais com diferentes proporções de grupos de cor/raça podem ter outros resultados quanto à relevância desta variável para a explicação da participação no Enem.

As escolas pesquisadas se concentram nos níveis IV e V do Inse<sup>7</sup>. De modo geral, os estudantes das escolas estaduais de Porto Alegre sinalizaram ter acesso a bens elementares, complementares e suplementares, além de responsáveis com ao menos ensino médio completo. Em 2015, ano de elaboração do Inse, a renda mensal domiciliar *per capita* do Rio Grande do Sul era de R\$1.434, ou seja, quase dois salários-mínimos da época (IBGE, 2016), valor condizente com os níveis apresentados pelos estudantes da capital.

Tabela 1 – Sumário da Participação no Enem

Medida	Valor
Mínimo	11,63
Primeiro quartil	28,57
Mediana	35,42
Média	37,62
Terceiro quartil	47,62
Máximo	75,54

Fonte: Os autores (2021).

A variável participação no Enem apresentou distribuição aproximadamente normal<sup>8</sup> e indica uma média baixa de participação no Enem, 37,62%. A média está

<sup>7</sup> No Nível IV, os estudantes, de modo geral, indicaram que havia em sua casa bens elementares como dois ou três quartos para dormir, um banheiro, uma geladeira, três ou mais telefones celulares, e um ou dois televisores; bens complementares como máquina de lavar roupas, micro-ondas, computador (com ou sem *internet*), um telefone fixo e um carro; e bens suplementares como *freezer*. Além disso, a renda familiar mensal está entre 1,5 e 3 salários-mínimos e os responsáveis completaram o ensino médio ou a faculdade. No Nível V, os estudantes, de modo geral, indicaram que há em suas casas dois ou mais banheiros, três quartos para dormir, quatro ou mais telefones celulares, dois ou três televisores; máquina de lavar roupas, um ou dois computadores (com ou sem *internet*), um telefone fixo, um carro, além de serviço de TV por assinatura; *freezer* e um aspirador de pó. Além disso, a renda familiar mensal está entre 2,5 e 7 salários-mínimos e os responsáveis completaram o ensino médio ou curso superior (INEP, 2015).

<sup>8</sup> Verificada por teste Shapiro-Wilk ( $p = 0,09$ ) e coeficiente de assimetria ( $z$ -assimetria = 0,50).

relacionada ao índice das escolas, não à totalidade dos estudantes de Porto Alegre. Isso significa que há baixa participação na maioria das escolas, informação também evidenciada pelo dado de que, dentre as 61 escolas, apenas 11 têm mais de 50% dos alunos do terceiro ano participando do exame.

Feita a análise preliminar dos dados e suas ponderações, é apresentado o modelo de regressão com todas as variáveis de estudo. A regressão do modelo completo inicial, com todas as variáveis, gerou os valores apresentados na Tabela 2.

Tabela 2 – Regressão múltipla, modelo completo inicial

	<b>B</b>	<b>valor-p</b>
Intercepto	-95,88391	0,14
Complexidade da gestão	-2,81948	0,11
Adequação da formação	0,28409	0,07
Esforço docente	-0,09658	0,77
Equipamentos	0,75852	0,62
Espaços	1,25224	0,22
Distorção idade-série	-0,48486	0,00
Inse	2,87424	0,01
% Mulheres	-0,14343	0,68
% Pretos/Pardos	-0,11641	0,48
$R^2 = 0,49$ ; $R^2$ ajustado = 0,40		

Fonte: Os autores (2021).

Com aplicação do método *stepwise*, restaram as variáveis complexidade da gestão, distorção idade-série e Inse, que alcançaram valor-p inferior ao nível de significância estabelecido de 0,05. Ou seja, é rejeitada a hipótese nula de que essas variáveis independentes têm efeito nulo na variância da dependente. Pode-se afirmar que é possível que complexidade da gestão, Inse e distorção idade-série tenham efeito real na taxa de participação no Enem, conforme Tabela 3.

Tabela 3 – Coeficientes B e Betas do modelo final

	<b>B</b>	<b>Betas (padronizados)</b>
Intercepto	-77,6329	-
Complexidade da gestão	-3,6460	-0,2270163
Distorção idade-série	-0,5316	-0,4995903
Inse	2,9063	0,3178183
$R^2 = 0,43$ ; $R^2$ ajustado = 0,40		

Fonte: Os autores (2021).

Ressalta-se que a normalidade dos resíduos foi verificada através do teste de Shapiro-Wilk ( $p = 0,72$ ) e visualização do gráfico quantil-quantil.

Apesar do  $R^2$  do primeiro modelo ser maior, com a ponderação do tamanho da amostra e da quantidade de variáveis presentes no modelo, o  $R^2$  ajustado do modelo final permaneceu igual. O novo modelo ajustado é mais parcimonioso e, com apenas três variáveis, explicou 40% da variação da dependente.

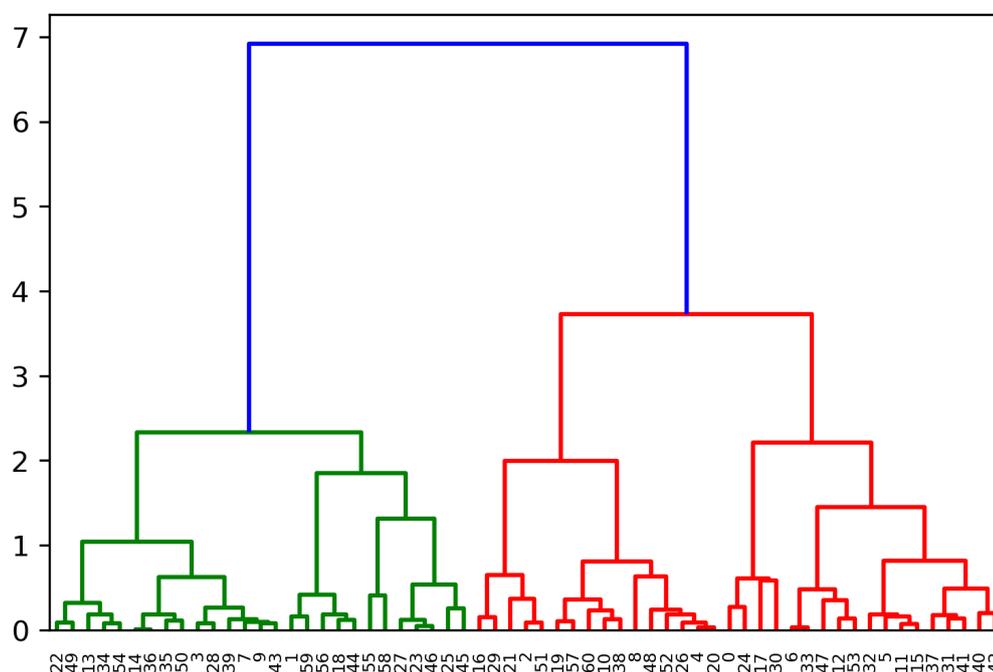
A variável explicativa com maior influência é a distorção idade-série, seguida do Inse, e a menos influente é a complexidade da gestão, vistos os coeficientes padronizados. Conforme os coeficientes da regressão, a participação no Enem decresce em média 3,64 unidades para cada aumento na complexidade da gestão. Também há uma relação inversa com a distorção idade-série: a cada aumento de unidade dessa variável, a participação tende a diminuir em média 0,53 unidade. Por outro lado, para cada aumento da unidade de Inse, aumenta-se em média 2,91 a participação no Enem.

Os dados do modelo ajustado indicam que variáveis relacionadas com condições socioeconômicas estão positivamente relacionadas com a participação no Enem, mostrando a influência das condições sociais e familiares nessa possibilidade de acesso à educação superior. A complexidade da gestão indica um fator escolar que, entre outras variáveis com potencial explicativo, pode dificultar a promoção dessa oportunidade junto a seus alunos. Maior complexidade na gestão está diretamente relacionada com a quantidade de níveis de formação, de turnos de atividades e do número de alunos matriculados.

Disso decorre a interpretação de que escolas que se dedicam unicamente ao nível médio têm gestão mais simplificada e, portanto, podem atender com maior acuidade às demandas e expectativas desse nível de formação. Com os resultados da regressão e a aplicação da técnica de análise de agrupamentos, foi gerado o dendrograma apresentado na Figura 1.

O gráfico exibe a possibilidade de dois grupos de escolas, o primeiro representado pela cor verde e o segundo em vermelho. A distância horizontal indica a distância relativa entre os casos. Como esperado, devido às particularidades do método de Ward, os agrupamentos têm quantidade semelhante de casos (HAIR; BLACK; BABIN; ANDERSON; TATHAM, 2009). A fim de compreender melhor as características das divisões, calculou-se a média de cada variável segundo cada agrupamento, obtendo os resultados expostos na Tabela 4.

Figura 1 – Dendrograma das escolas públicas estaduais de Ensino Médio do município de Porto Alegre, 2019, conforme a participação no Enem, distorção idade-série, complexidade da gestão e Inse



Fonte: Os autores (2021).

Como indicado pela regressão, escolas com distorção idade-série mais alta, maior complexidade da gestão e Inse mais baixo têm participação no Enem mais baixa. A análise mostra agrupadas as escolas com tais características no Grupo 2 (com 27 escolas). No Grupo 1 (com 34 escolas), estão aquelas com maior participação no Enem e indicadores mais favoráveis.

Tabela 4 – Resumo da média de cada variável segundo agrupamentos

	<b>Grupo 1 (n = 34) Cor vermelha</b>	<b>Grupo 2 (n = 27) Cor verde</b>	<b>Amostra (n = 61)</b>
Participação no Enem	46,54	26,4	37,62
Complexidade da gestão	4,44	4,63	4,52
Distorção idade-série	31,66	50,15	39,84
Inse	53,69	52,82	53,31

Fonte: Os autores (2021).

### Considerações finais

Vimos que 70% dos estudantes de terceiro ano das escolas amostradas encontram-se em distorção idade-série de dois anos ou mais. Apenas 11 das 61 escolas da amostra (15,5%) evidenciaram ter mais do que 50% dos alunos do terceiro ano participando do Enem. Esses dados compõem um panorama da

situação de estudantes matriculados em escolas públicas estaduais na cidade de Porto Alegre.

Constatou-se uma relação positiva entre participação no Enem e menor distorção série-idade, gestão menos complexa e Inse mais alto. Fatores como adequação da formação dos professores, esforço docente, equipamentos disponíveis na escola, espaços nela presentes, porcentagem de mulheres e pessoas negras não foram constatados como significativos para explicar a variação na participação no Enem.

A identificação de dois agrupamentos de escolas, conforme apresentado, não pode ser tomada como um diagnóstico definitivo sobre a situação das escolas que compõem cada um deles. Entretanto, é possível entender que, para a análise da participação dos estudantes na prova do Enem, há tendências que precisam ser observadas. Esses padrões subsidiam a problematização de trajetórias escolares no ensino médio ao indicar maior probabilidade e projeto de continuidade de estudos por parte dos estudantes do ensino médio público estadual que frequentam escolas do Grupo 1, que se aproximam por características de menor complexidade da gestão, menor distorção idade-série e maior Inse. É importante destacar que as escolas do grupo com maior participação no Enem também apresentaram média de notas (543,4) maior que o grupo com menor participação (519,8)<sup>9</sup>.

A heterogeneidade na caracterização escolar demonstra que desigualdades nas perspectivas de ingresso na educação superior, para além de estarem vinculadas diretamente a condições socioeconômicas e culturais, são mediadas por condições institucionais dos ambientes escolares. O modelo ajustado alcançou uma capacidade de explicação da participação no Enem de 40% a partir de três variáveis. A partir desse resultado, é possível visualizar possibilidades de avanço na construção metodológica para exploração da realidade. Se uma parte da explicação reside nas trajetórias dos indivíduos e nas singularidades dos meios familiares, a continuidade do desenvolvimento da metodologia aqui apresentada com novas pesquisas, amostras e modelos pode responder às questões relativas à incidência do meio institucional, através da ação escolar. Na continuidade da pesquisa, cabe também nos aproximarmos destas escolas para desenvolver com a

---

<sup>9</sup> Considerando a média aritmética das quatro provas do Enem 2018.

comunidade escolar estudos qualitativos que permitam entender dinâmicas internas e recursos acionados no cotidiano do trabalho escolar que caracterizam essa mediação institucional, que pode agravar ou remediar as desigualdades de origem social entre estudantes.

A discussão aqui apresentada busca expor uma agenda de pesquisa para reconhecer heterogeneidade entre escolas públicas no Brasil. Com sua aplicação, é possível esmiuçar fenômenos educacionais aliando análise de fatores escolares e extraescolares, avançando a partir de uma perspectiva teórica de causalidade multidirecional, verificada no debate em sociologia da educação. Essa agenda, ao robustecer o reconhecimento da realidade escolar, favorece a identificação de características e práticas que impactam as chances de acesso à educação superior por egressos do ensino médio público. Ao trabalhar com amostras maiores, será possível reinserir as variáveis formação de professores e condições de trabalho, bem como incluir variáveis relativas ao currículo em modelos ajustados posteriores. O conhecimento resultante dessa agenda de pesquisa poderá fomentar agendas de ação pedagógica e políticas públicas de qualificação das práticas de governança e formação de escolas, grupos de escolas e redes de ensino.

## Referências

AGRESTI, A.; FINLAY, B. *Métodos estatísticos para as ciências sociais*. 4. ed. Porto Alegre: Penso, 2012.

ALVES, M. T. G. Dimensões do efeito das escolas: explorando as interações entre famílias e estabelecimentos de ensino. *Estudos em Avaliação Educacional*, São Paulo, v. 21, n. 46, p. 271-296, 2010.

ALVES, M. T. G.; SOARES, J. F. Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional. *Educação e Pesquisa*, São Paulo, v. 39, n. 1, p. 177-194, jan./mar. 2013.

BRESSOUX, P. As pesquisas sobre o efeito-escola e o efeito-professor. *Educação em Revista*, Belo Horizonte, n. 38, p. 17-88, 2003.

BOURDIEU, P. A escola conservadora: as desigualdades frente à escola e à cultura. In: NOGUEIRA, M. A.; CATANI, A. (org.). *Escritos de educação*. Petrópolis, RJ: Vozes, 2010. p. 45-72.

BRANDÃO, Z.; CARVALHO, C. P. de. Processos de produção das elites escolares. *Educação e Sociedade*, Campinas, SP, v. 32, n. 115, p. 507-522, abr./jun. 2011. DOI: <https://doi.org/10.1590/S0101-73302011000200015>. Disponível em: <https://www.scielo.br/j/es/a/YvgKjXCt67n6zycPJLwHQhp/?format=pdf&lang=pt>. Acesso em: 1 ago. 2020.

BRASIL. Emenda Constitucional nº 59, de 11 de novembro de 2009. Acrescenta § 3º ao art. 76 do Ato das Disposições Constitucionais Transitórias para reduzir, [...] e dá nova redação ao § 4º do art. 211 e ao § 3º do art. 212 e ao caput do art. 214, com inserção neste dispositivo de inciso VI. *Diário Oficial da União*, Brasília, DF, 12 nov. 2009. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/constituicao/emendas/emc/emc59.htm](http://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc59.htm). Acesso em: 21 jul. 2020.

BRASIL. Lei nº 13.005, de 24 de junho de 2014. Aprova o Plano Nacional de Educação - PNE e dá outras providências. *Diário Oficial da União*, Brasília, DF, 26 jun. 2014. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2014/lei/l13005.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l13005.htm). Acesso em: 21 jul. 2020.

CAREGNATO, Célia Elizabete; SANTOS, Harlon Romariz Rabelo; FELIN, Lourenço Brito. Origem escolar e acesso à educação superior: análise da ocupação de vagas de ações afirmativas na UFRGS. *Educação em Revista*, Belo Horizonte, v. 36, n. e231759, p. 1-21. 2020. DOI: 10.1590/0102-4698231759.

CARRANO P. C. R.; MARINHO, A. C.; OLIVEIRA, V. N. M. de. Trajetórias truncadas, trabalho e futuro: jovens fora de série na escola pública de ensino médio. *Educação e Pesquisa*, São Paulo, v. 41, n. esp., p. 1439-1454, 2015. DOI: <https://doi.org/10.1590/S1517-9702201508143413>. Disponível em: <https://www.scielo.br/j/ep/a/P9xX7fLyt8MSgLVmcGK4ft/?format=pdf&lang=pt>. Acesso em: 12 jul. 2020.

DAYRELL, J. T.; JESUS, R. E. de. Juventude, ensino médio e os processos de exclusão escolar. *Educação e Sociedade*, Campinas, SP, v. 37, n. 135, p. 407-423, abr./jun. 2016. DOI: <https://doi.org/10.1590/ES0101-73302016151533>. Disponível em: <https://www.scielo.br/j/es/a/vDyjXnzDWz5VsFKFzVytMmp/?format=pdf&lang=pt>. Acesso em: 10 jul. 2020.

DOWNEY, D. B.; CONDRON, D. J. Fifty years since the Coleman Report: rethinking the relationship between schools and inequality. *Sociology of Education*, New York, v. 89, n. 3, p. 207-220, 2016. DOI: <https://doi.org/10.1177/0038040716651676>. Disponível em: <https://journals.sagepub.com/doi/pdf/10.1177/0038040716651676>. Acesso em: 27 jun. 2020.

GROPPO, L. A. Teorias críticas da juventude: geração, moratória social e subculturas juvenis. *Em Tese*, Florianópolis, v. 12, n. 1, jan./jul. 2015. DOI: <https://doi.org/10.5007/1806-5023.2015v12n1p4>. Disponível em: <https://periodicos.ufsc.br/index.php/emtese/article/view/1806-5023.2015v12n1p4/29763>. Acesso em: 1 jun. 2020.

HAECHT, A. van. *Sociologia da educação: a escola posta à prova*. 3. ed. Porto Alegre: Artmed, 2008.

HAIR, J. F.; BLACK, W.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. *Análise multivariada de dados*. 6. ed. Porto Alegre: Bookman, 2009.

IBGE. Tabela 3175: população residente, por cor ou raça, segundo a situação do domicílio, o sexo e a idade. In: IBGE. *Censo demográfico*. Brasília, DF: IBGE, 2010. Disponível em: <https://sidra.ibge.gov.br/Tabela/3175>. Acesso em: 9 jul. 2020.

IBGE. *IBGE divulga renda domiciliar per capita 2015*. Brasília, DF: IBGE, 2016. Disponível em: [ftp://ftp.ibge.gov.br/Trabalho\\_e\\_Rendimento/Pesquisa\\_Nacional\\_por\\_Amostra\\_de\\_Domicilios\\_continua/Renda\\_domiciliar\\_per\\_capita/Renda\\_domiciliar\\_per\\_capita\\_2015\\_20160420.pdf](ftp://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilios_continua/Renda_domiciliar_per_capita/Renda_domiciliar_per_capita_2015_20160420.pdf). Acesso em: 09 jul. 2020.

INEP. Diretoria de Estatísticas Educacionais. *Adequação da formação docente: escolas 2018*. Brasília, DF: Inep, 2019a. Disponível em: [http://download.inep.gov.br/informacoes\\_estatisticas/indicadores\\_educacionais/2018/AFD\\_2018\\_ESCOLAS.zip](http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2018/AFD_2018_ESCOLAS.zip). Acesso em: 9 jul. 2021.

INEP. Diretoria de Estatísticas Educacionais. *Censo da educação básica 2020: resumo técnico*. Brasília, DF: Inep, 2021. Disponível em: [https://download.inep.gov.br/publicacoes/institucionais/estatisticas\\_e\\_indicadores/resumo\\_tecnico\\_censo\\_escolar\\_2020.pdf](https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/resumo_tecnico_censo_escolar_2020.pdf). Acesso em: 6 maio 2021.

INEP. Diretoria de Estatísticas Educacionais. *Indicador de esforço docente: escolas 2018*. Brasília, DF: Inep, 2019b. Disponível em: [http://download.inep.gov.br/informacoes\\_estatisticas/indicadores\\_educacionais/2018/IED\\_2018\\_ESCOLAS.zip](http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2018/IED_2018_ESCOLAS.zip). Acesso em: 9 jul. 2021.

INEP. Diretoria de Estatísticas Educacionais. *Indicador de Nível Socioeconômico: escolas 2015*. Brasília, DF: Inep, 2016. Disponível em:

[http://download.inep.gov.br/informacoes\\_estatisticas/indicadores\\_educacionais/2015/nivel\\_socioeconomico/Indicador\\_Inse\\_2015.zip](http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2015/nivel_socioeconomico/Indicador_Inse_2015.zip). Acesso em: 9 jul. 2021.

INEP. Diretoria de Estatísticas Educacionais. *Microdados do censo da educação básica 2018*. Brasília, DF: Inep, 2019c. Disponível em: [http://download.inep.gov.br/microdados/microdados\\_educacao\\_basica\\_2018.zip](http://download.inep.gov.br/microdados/microdados_educacao_basica_2018.zip). Acesso em: 9 jul. 2021.

INEP. Diretoria de Estatísticas Educacionais. *Nota técnica*. Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse). Brasília, DF: Inep, 2015. Disponível em: [https://download.inep.gov.br/informacoes\\_estatisticas/indicadores\\_educacionais/2015/nota\\_tecnica/nota\\_tecnica\\_inep\\_inse\\_2015.pdf](https://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2015/nota_tecnica/nota_tecnica_inep_inse_2015.pdf). Acesso em: 9 jul. 2020.

INEP. Diretoria de Estatísticas Educacionais. *Nota técnica nº 020/2014*. Indicador de adequação da formação do docente da educação básica. Brasília, DF: Inep, 2014a. Disponível em: [https://download.inep.gov.br/informacoes\\_estatisticas/indicadores\\_educacionais/2014/docente\\_formacao\\_legal/nota\\_tecnica\\_indicador\\_docente\\_formacao\\_legal.pdf](https://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2014/docente_formacao_legal/nota_tecnica_indicador_docente_formacao_legal.pdf). Acesso em: 9 jul. 2020.

INEP. Diretoria de Estatísticas Educacionais. *Nota técnica nº 039/2014*. Indicador de esforço docente. Brasília, DF: Inep, 2014b. Disponível em: [http://download.inep.gov.br/informacoes\\_estatisticas/indicadores\\_educacionais/2014/docente\\_esforco/nota\\_tecnica\\_indicador\\_docente\\_esforco.pdf](http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2014/docente_esforco/nota_tecnica_indicador_docente_esforco.pdf). Acesso em: 9 jul. 2020.

INEP. Diretoria de Estatísticas Educacionais. *Nota técnica nº 040/2014*. Indicador para mensurar a complexidade da gestão nas escolas a partir dos dados do censo escolar da educação básica. Brasília, DF: Inep, 2014c. Disponível em: [https://download.inep.gov.br/informacoes\\_estatisticas/indicadores\\_educacionais/2014/escola\\_complexidade\\_gestao/nota\\_tecnica\\_indicador\\_escola\\_complexidade\\_gestao.pdf](https://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2014/escola_complexidade_gestao/nota_tecnica_indicador_escola_complexidade_gestao.pdf). Acesso em: 9 jul. 2020.

INEP. Diretoria de Estatísticas Educacionais. *Taxa de distorção idade-série: escolas 2018*. Brasília, DF: Inep, 2019d. Disponível em: [http://download.inep.gov.br/informacoes\\_estatisticas/indicadores\\_educacionais/2018/TDI\\_2018\\_ESCOLAS.zip](http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2018/TDI_2018_ESCOLAS.zip). Acesso em: 9 jul. 2021.

KRAWCZYK, N.; SILVA, C. J. O. Desigualdades educacionais no ensino médio brasileiro: uma análise de perfil socioeconômico de jovens que realizaram o exame nacional do ensino médio. *Sensos-e*, [Porto] v. 4, n. 1, p. 12-23, 2017. DOI: <https://doi.org/10.34630/sensos-e.v4i1.2253>. Disponível em: <https://parc.ipp.pt/index.php/sensos/article/view/2253/640>. Acesso em: 27 jun. 2020.

LAHIRE, B. *Sucesso escolar nos meios populares: as razões do improvável*. São Paulo: Ática, 1997.

QEDU. *Matrículas e infraestrutura*. Brasília, DF, [2021]. Disponível em: <https://qedu.org.br/brasil/censo-escolar?year=2020&dependence=0&localization=0&item=>. Acesso em: 6 maio 2021.

- RIO GRANDE DO SUL. Secretaria de Educação. CRE 01. In: Rio Grande do Sul. Secretaria de Educação. *Coordenadorias regionais de educação*. Porto Alegre, 2020. Disponível em: <https://educacao.rs.gov.br/divisao-porto-alegre>. Acesso em: 18 out. 2020.
- SILVA, M. R. da. Ampliação da obrigatoriedade escolar no Brasil: o que aconteceu com o ensino médio?. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 28, n. 107, p. 274-291, abr./jun. 2020. DOI: <https://doi.org/10.1590/S0104-40362019002701953>. Disponível em: <https://www.scielo.br/j/ensaio/a/mqfqFrxbKWWpcjhwrGNqsgn/?format=pdf&lang=pt>. Acesso em: 6 maio 2021.
- SOARES, J. F.; CANDIAN, J. F. O efeito da escola básica brasileira: as evidências do Pisa e do Saeb. *Revista Contemporânea de Educação*, Rio de Janeiro, v. 2, n. 4, p. 163-182, 2007. DOI: <https://doi.org/10.20500/rce.v2i4.1522>. Disponível em: <https://revistas.ufrj.br/index.php/rce/article/view/1522/1371>. Acesso em: 3 ago. 2021.
- SOARES, T. M.; FERNANDES, N. da S.; NÓBREGA, M. C.; NICOLELLA, A. C. Fatores associados ao abandono escolar no ensino médio público de Minas Gerais. *Educação e Pesquisa*, São Paulo, v. 41, n. 3, p. 757-772, jul./set. 2015. DOI: <http://dx.doi.org/10.1590/S1517-9702201507138589>. Disponível em: <https://www.scielo.br/j/ep/a/XhMWFmKSzSrKCsDPhbsYs5P/?format=pdf&lang=pt>. Acesso em: 9 jul. 2020.
- ZAGO, N. Processos de escolarização nos meios populares: as contradições da obrigatoriedade escolar. In: NOGUEIRA, M. A.; ROMANELLI, G.; ZAGO, N. (org.). *Família e escola: trajetórias de escolarização em camadas médias e populares*. Petrópolis, RJ: Vozes, 2000a. p. 19-43.
- ZAGO, N. Quando os dados contrariam as previsões estatísticas: os casos de êxito escolar nas camadas socialmente desfavorecidas. *Paidéia*, Ribeirão Preto, SP, v. 10, n. 18, p. 70-80, jan./jul. 2000b.