

MATEUS YUKIO FUKUHARA OIOLI

Análise de clusters aplicada à Educação Básica Brasileira para políticas públicas

Trabalho de Formatura apresentado à Escola
Politécnica da Universidade de São Paulo para
obtenção do diploma de Engenheiro de
Produção.

SÃO PAULO

2022

MATEUS YUKIO FUKUHARA OIOLI

Análise de clusters aplicada à Educação Básica Brasileira para políticas públicas

Trabalho de Formatura apresentado à Escola
Politécnica da Universidade de São Paulo para
obtenção do diploma de Engenheiro de
Produção.

Orientador: Aline Veronese da Silva

SÃO PAULO

2022

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

Oioli, Mateus

Análise de clusters aplicada à Educação Básica Brasileira para políticas públicas/ M. Y. Oioli – São Paulo, 2022

107 p.

Trabalho de Formatura – Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1. Análise de *clusters* 2. Educação 3. K-Means I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Produção II. t.

Dedico este trabalho aos bons amigos que
fiz esses anos, a minha família, e a todos
os professores que me ensinaram a gostar
de aprender

AGRADECIMENTOS

Gostaria de agradecer a todas as pessoas que me apoiaram nessa jornada, se consegui ingressar na Escola Politécnica, e se consegui me graduar nela foi graças ao apoio dessas pessoas.

Agradeço a minha orientadora, Prof. Dr. Aline Veronese pelo suporte e ensinamentos, dados durante a elaboração deste trabalho, sem os quais este não seria possível.

Agradeço aos professores que despertaram meu interesse sobre diferentes matérias em diferentes etapas da minha vida.

Agradeço a todos os amigos que fiz por todas as escolas e cidades que passei, e que ajudaram na minha felicidade e formação como pessoa ao longo desses anos.

Por fim, agradeço aos meus pais e familiares que me apoiaram durante toda a vida, sempre desejando o que era melhor para mim.

“If you are planning for a year,
sow rice; if you are planning for a decade,
plant trees; if you are planning for a
lifetime, educate people.”

(Provérbio chinês)

RESUMO

O principal objetivo desse trabalho é auxiliar a compreensão da complexa realidade da educação no Brasil utilizando técnicas estatísticas sobre diversos dados educacionais do país. A técnica estatística aplicada neste trabalho foi a clusterização por K-means sobre indicadores educacionais dos municípios do país. Os indicadores foram escolhidos de forma a abordar aspectos educacionais de maneira holística, sejam eles intrínsecos ao sistema educacional, como é o caso do número de alunos por turma, ou não, como é o caso da realidade socioeconômica dos estudantes. A análise foi realizada pela formação de *clusters* de municípios em diferentes situações, sendo elas diferentes combinações de indicadores educacionais sobre diferentes etapas do ensino básico nacional. Os resultados obtidos apresentaram a complexidade do espectro educacional sob o qual a educação nacional se encontra, indicando diversas composições distintas de *cluster* para cada situação analisada, expondo o impacto que diferentes fatores podem gerar dependendo da realidade a qual o aluno está inserido. A parte final do trabalho apresentou sugestões breves para cada situação estudada a fim de potencializar os resultados de medidas públicas sobre o aprendizado dos alunos.

Palavras-chave: Análise de *clusters*. Educação. K-means.

ABSTRACT

The main objective of this work is to assist the comprehension of the complex reality of the Brazilian educational system using statistical methods over several educational datasets of the country. The statistical method applied in this paper is clustering by K-means, which was applied over municipal educational indexes. The indexes were chosen to get a holistic perspective of educational aspects, whether they are intrinsic to the educational system, such as the average number of students by class, or not, such as the socioeconomic reality of the students. The analysis was developed by creating clusters of cities across many situations, such as several combinations of clusters across different educational stages. The results demonstrate the complexity of the Brazilian educational system, showing several different clusters composition across different situations, indicating the impact that different aspects can produce depending on the reality of each student. The final part of this paper shows suggestions for every situation observed to leverage future public policies' impact on students learning.

Keywords: Cluster analysis. Education. K-means.

LISTA DE FIGURAS

Figura 1: Evolução das matrículas em creche por rede de ensino no Brasil (2015 a 2020).....	17
Figura 2: Evolução das matrículas em pré-escola por rede de ensino no Brasil (2015 a 2020)	17
Figura 3: Evolução das matrículas nos anos iniciais do ensino fundamental por dependência administrativa no Brasil (2016 a 2020)	19
Figura 4: Evolução das matrículas nos anos finais do ensino fundamental por dependência administrativa no Brasil (2016 a 2020)	20
Figura 5: Evolução das matrículas no ensino médio por dependência administrativa no Brasil (2016 a 2020).....	22
Figura 6: Exemplo de observações agrupadas.....	46
Figura 7: Formação de grupos conforme o Algoritmo de clusterização.	49
Figura 8: Progresso do algoritmo K-means em 6 iterações.....	50
Figura 9: Exemplo de <i>boxplot</i>	53
Figura 10: Resultados do Saeb em língua portuguesa para anos iniciais do ensino fundamental – com <i>outliers</i>	55
Figura 11: Resultados do Saeb em língua portuguesa para anos iniciais do ensino fundamental – sem <i>outliers</i>	55
Figura 12: Número ótimo de <i>clusters</i> (variável <i>number of clusters k</i>).....	57
Figura 13: Alunos dos anos iniciais do ensino fundamental na rede pública.....	58
Figura 14: Proporção de alunos dos anos iniciais do ensino fundamental por rede de ensino.	58
Figura 15: <i>Clusters</i> considerando todos os indicadores para os anos iniciais do ensino fundamental em zonas urbanas.....	59
Figura 16: <i>Clusters</i> considerando indicadores socioeconômicos e estruturais para os anos iniciais do ensino fundamental em zonas urbanas.....	60
Figura 17: <i>Clusters</i> considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para os anos iniciais do ensino fundamental em zonas urbanas	60
Figura 18: <i>Clusters</i> considerando indicadores dos docentes para os anos iniciais do ensino fundamental em zonas urbanas.....	61
Figura 19: <i>Clusters</i> considerando indicadores dos docentes e do desempenho dos alunos para os anos iniciais do ensino fundamental em zonas urbanas.....	62
Figura 20: <i>Clusters</i> considerando todos os indicadores para os anos iniciais do ensino fundamental em zonas rurais	63

Figura 21: <i>Clusters</i> considerando indicadores socioeconômicos e estruturais para os anos iniciais do ensino fundamental em zonas rurais	64
Figura 22: <i>Clusters</i> considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para os anos iniciais do ensino fundamental em zonas rurais	64
Figura 23: <i>Clusters</i> considerando indicadores dos docentes para os anos iniciais do ensino fundamental em zonas rurais	65
Figura 24: <i>Clusters</i> considerando indicadores dos docentes e do desempenho dos alunos para os anos iniciais do ensino fundamental em zonas rurais	66
Figura 25: Alunos dos anos finais do ensino fundamental na rede pública	67
Figura 26: Proporção de alunos dos anos finais do ensino fundamental por rede de ensino ...	67
Figura 27: <i>Clusters</i> considerando todos os indicadores para os anos finais do ensino fundamental em zonas urbanas.....	68
Figura 28: <i>Clusters</i> considerando indicadores socioeconômicos e estruturais para os anos finais do ensino fundamental em zonas urbanas	68
Figura 29: <i>Clusters</i> considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para os anos finais do ensino fundamental em zonas urbanas.....	69
Figura 30: <i>Clusters</i> considerando indicadores dos docentes para os anos finais do ensino fundamental em zonas urbanas.....	70
Figura 31: <i>Clusters</i> considerando indicadores dos docentes e do desempenho dos alunos para os anos finais do ensino fundamental em zonas urbanas	70
Figura 32: <i>Clusters</i> considerando todos os indicadores para os anos finais do ensino fundamental em zonas rurais	71
Figura 33: <i>Clusters</i> considerando indicadores socioeconômicos e estruturais para os anos finais do ensino fundamental em zonas rurais.....	72
Figura 34: <i>Clusters</i> considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para os anos finais do ensino fundamental em zonas rurais	72
Figura 35: <i>Clusters</i> considerando indicadores dos docentes para os anos finais do ensino fundamental em zonas rurais	73
Figura 36: <i>Clusters</i> considerando indicadores dos docentes e do desempenho dos alunos para os anos finais do ensino fundamental em zonas rurais.....	74
Figura 37: Alunos do ensino médio na rede pública	75
Figura 38: Proporção de alunos do ensino médio por rede de ensino	75
Figura 39: <i>Clusters</i> considerando todos os indicadores para o ensino médio em zonas urbanas	

.....	76
Figura 40: <i>Clusters</i> considerando indicadores socioeconômicos e estruturais para o ensino médio em zonas urbanas.....	76
Figura 41: <i>Clusters</i> considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para o ensino médio em zonas urbanas	77
Figura 42: <i>Clusters</i> considerando indicadores dos docentes para o ensino médio em zonas urbanas.....	77
Figura 43: <i>Clusters</i> considerando indicadores dos docentes e do desempenho dos alunos para o ensino médio em zonas urbanas	78
Figura 44: <i>Clusters</i> considerando todos os indicadores para o ensino médio em zonas rurais	79
Figura 45: <i>Clusters</i> considerando indicadores socioeconômicos e estruturais para o ensino médio em zonas rurais	79
Figura 46: <i>Clusters</i> considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para o ensino médio em zonas rurais.....	80
Figura 47: <i>Clusters</i> considerando indicadores dos docentes para o ensino médio em zonas rurais	80
Figura 48: <i>Clusters</i> considerando indicadores dos docentes e do desempenho dos alunos para o ensino médio em zonas rurais.....	81
Figura 49: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para os anos iniciais do ensino fundamental em regiões urbanas	82
Figura 50: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões urbanas	83
Figura 51: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões urbanas	84
Figura 52: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões urbanas.....	85
Figura 53: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para os anos iniciais do ensino fundamental em regiões rurais.....	86

Figura 54: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões rurais.....	87
Figura 55: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões rurais	88
Figura 56: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões rurais	88
Figura 57: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para os anos finais do ensino fundamental em regiões urbanas	90
Figura 58: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões urbanas	90
Figura 59: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões urbanas	91
Figura 60: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões urbanas	92
Figura 61: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para os anos finais do ensino fundamental em regiões rurais.....	93
Figura 62: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões rurais	94
Figura 63: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões rurais	95
Figura 64: Comparação entre indicadores socioeconômicos e estruturais considerando o	

desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões urbanas	96
Figura 65: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para o ensino médio em regiões urbanas	97
Figura 66: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para o ensino médio em regiões urbanas	98
Figura 67: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para o ensino médio em regiões urbanas	99
Figura 68: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para o ensino médio em regiões urbanas	100
Figura 69: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para o ensino médio em regiões rurais	101
Figura 70: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para o ensino médio em regiões rurais	102
Figura 71: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para o ensino médio em regiões rurais.....	103
Figura 72: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para o ensino médio em regiões rurais	103

LISTA DE TABELAS

Quadro 1: Itens abordados pelo questionário Inse.....	26
Quadro 2: Níveis do Daeb para indicadores socioeconômicos	27
Quadro 3: Descrição dos níveis de avaliação de esforço docente	29
Quadro 4: Faixas de regularidade do corpo docente	30
Quadro 5: Descrição dos níveis de desempenho em língua portuguesa nos anos iniciais do ensino fundamental.....	31
Quadro 6: Descrição dos níveis de desempenho em matemática nos anos iniciais do ensino fundamental	33
Quadro 7: Níveis de desempenho em Língua Portuguesa para anos finais do ensino fundamental	36
Quadro 8: Descrição dos níveis de desempenho em Matemática para anos finais do ensino fundamental	37
Quadro 9: Habilidades esperadas para Língua Portuguesa no ensino médio	40
Quadro 10: Habilidades em matemática para alunos do ensino médio	41
Quadro 11: Situações exploradas para definição dos <i>clusters</i>	56
Quadro 12: Cenários para definição dos <i>clusters</i>	56

SUMÁRIO

1	INTRODUÇÃO	15
1.1	A educação básica no brasil.....	15
1.1.1	O ensino infantil no Brasil	15
1.1.2	O ensino fundamental no Brasil	18
1.1.3	O ensino médio no Brasil	20
1.2	Definição do problema	22
1.3	Objetivos.....	23
1.4	estrutura do trabalho	23
2	REFERENCIAL TEÓRICO	25
2.1	Indicadores educacionais analisados	25
2.1.1	Sistema de Avaliação Básica (Saeb).....	25
2.1.2	Indicador socioeconômico	26
2.1.3	Indicadores sobre a condição dos docentes	28
2.1.3.1	Indicador esforço docente.....	29
2.1.3.2	Indicador regularidade do corpo docente	29
2.1.4	Indicador sobre a quantidade de alunos por sala	30
2.1.5	Indicador sobre o aprendizado dos alunos.....	30
2.1.5.1	Níveis do Saeb para os anos iniciais do ensino fundamental	31
2.1.5.1.1	Língua portuguesa	31
2.1.5.1.2	Matemática	32
2.1.5.2	Níveis do Saeb para os anos finais do ensino fundamental	35
2.1.5.2.1	Língua Portuguesa	36
2.1.5.2.2	Matemática	37
2.1.5.3	Níveis do Saeb para o ensino médio.....	40

2.1.5.3.1	Língua Portuguesa	40
2.1.5.3.2	Matemática	41
3	METODOLOGIA.....	45
3.1	Método de clusterização K-means.....	46
3.1.1	Aplicação da técnica K-means computacionalmente	50
3.2	Identificação de <i>outliers</i>	51
3.2.1	Diagrama <i>Boxplot</i>	51
4	RESULTADOS	54
4.1	Análise de <i>Outliers</i>	54
4.2	Definição dos <i>Clusters</i>	55
4.2.1	Resultados obtidos para os anos iniciais do ensino fundamental	57
4.2.1.1	Localização Urbana	58
4.2.1.2	Localização Rural	62
4.2.2	Resultados obtidos para os anos Finais do Ensino Fundamental	66
4.2.2.1	Localização Urbana	67
4.2.2.2	Localização Rural	71
4.2.3	Resultados obtidos para o Ensino Médio.....	74
4.2.3.1	Localização Urbana	75
4.2.3.2	Localização Rural	78
5	ANÁLISE DOS RESULTADOS E DISCUSSÃO.....	82
5.1	Cenário 1 – Anos iniciais do ensino fundamental – Zona Urbana	82
5.1.1	Situação socioeconômico e estrutural com a situação dos docentes	82
5.1.2	Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos.....	82
5.1.3	Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos	83
5.1.4	Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos	84

5.1.5	Todos os indicadores com as outras situações já analisadas	85
5.1.6	Recomendações para o cenário.....	85
5.2	Cenário 2 – Anos iniciais do ensino fundamental – Zona Rural	86
5.2.1	Situação socioeconômico e estrutural com a situação dos docentes	86
5.2.2	Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos.....	87
5.2.3	Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos	87
5.2.4	Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos	88
5.2.5	Todos os indicadores com as outras situações já analisadas	89
5.2.6	Recomendações para o cenário.....	89
5.3	Cenário 3 – Anos finais do ensino fundamental – Zona Urbana.....	89
5.3.1	Situação socioeconômico e estrutural com a situação dos docentes	89
5.3.2	Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos.....	90
5.3.3	Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos	91
5.3.4	Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos	91
5.3.5	Todos os indicadores com as outras situações já analisadas	92
5.3.6	Recomendações para o cenário.....	92
5.4	Cenário 4 – Anos finais do ensino fundamental – Zona Rural.....	93
5.4.1	Situação socioeconômico e estrutural com a situação dos docentes	93
5.4.2	Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos.....	93
5.4.3	Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos	94

5.4.4	Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos	95
5.4.5	Todos os indicadores a partir das comparações já realizadas.....	96
5.4.6	Recomendações para o cenário.....	96
5.5	Cenário 5 – Ensino médio – Zona Urbana.....	97
5.5.1	Situação socioeconômico e estrutural com a situação dos docentes	97
5.5.2	Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos.....	98
5.5.3	Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos	98
5.5.4	Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos	99
5.5.5	Todos os indicadores a partir das comparações já realizadas.....	100
5.5.6	Recomendações para o cenário.....	100
5.6	Cenário 6 – Ensino médio – Zona Rural	101
5.6.1	Situação socioeconômico e estrutural com a situação dos docentes	101
5.6.2	Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos.....	102
5.6.3	Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos	102
5.6.4	Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos	103
5.6.5	Todos os indicadores a partir das comparações já realizadas.....	104
5.6.6	Sugestões para o cenário.....	104
6	Conclusão	105
	REFERÊNCIAS	106

1 INTRODUÇÃO

Este capítulo se inicia apresentando de forma breve as definições legais de cada uma das etapas do ensino básico do país, junto a dados que demonstrem a dimensão e complexidade da educação no país. Em seguida, alguns indicadores são apresentados a fim de mostrar o resultado de anos de políticas públicas malsucedidas, por fim alguns dos problemas educacionais brasileiros são sumarizados, e o objetivo deste trabalho é definido.

1.1 A EDUCAÇÃO BÁSICA NO BRASIL

A qualidade da educação de um país sempre foi determinante para seu desenvolvimento, ela é capaz de gerar empregos mais bem remunerados, cidadãos mais conscientes, produção científica mais avançada, e garantir que esses avanços perdurem durante a mudança de gerações.

O Brasil, apesar de direcionar à educação grande parte do orçamento que é recolhido pelo Estado, nunca conseguiu transformar o nível de orientação da população de forma condizente com os valores que são direcionados, figurando sempre em posições precárias nos principais indicadores internacionais de educação, a situação se mostra ainda mais crítica quando relacionamos esses indicadores ao tamanho das economias dos países.

O sistema educacional brasileiro é dividido em dois macro grupos segundo a Lei de Diretrizes e Bases da Educação Nacional (LDB), sendo eles: Ensino Básico e Ensino Superior. Em 2020 aproximadamente 22% da população (47,3 milhões de habitantes) encontra-se no primeiro grupo (INEP, 2020), enquanto em 2019 4% (8,6 milhões de habitantes) encontra-se no segundo (INEP, 2019); sendo possível perceber que o Ensino Básico, além de se tratar de uma etapa anterior e condicional ao ensino superior, é muito mais representativo em número de alunos da educação no país e será o foco deste trabalho.

A educação básica no Brasil, segundo a LDB, engloba três etapas: Educação Infantil, Ensino Fundamental e Ensino Médio. Cada etapa condiciona o acesso à próxima, seguindo essa ordem. Cada etapa é organizada e gerida por entidades governamentais diferentes, e possui objetivos diferentes para seus estudantes.

1.1.1 O ensino infantil no Brasil

A educação infantil, primeira etapa da educação básica, tem como finalidade o desenvolvimento integral da criança de até cinco anos de idade (MINISTÉRIO DA CASA CIVIL, 1996). Foi considerado dever do Estado apenas a partir de 1988 e passou a integrar a Educação Básica em 1996 com a promulgação da LDB (Lei de Diretrizes e Bases da Educação Nacional). Se tornou obrigatória com a emenda constitucional nº59/2009 para crianças de 4 a 5 anos (MINISTÉRIO DA EDUCAÇÃO, 2015). Segundo a Diretriz, de 0 a 3 anos e 11 meses a criança frequenta creches, e dos 4 anos aos 5 anos e 11 meses ela frequenta as “pré-escolas”.

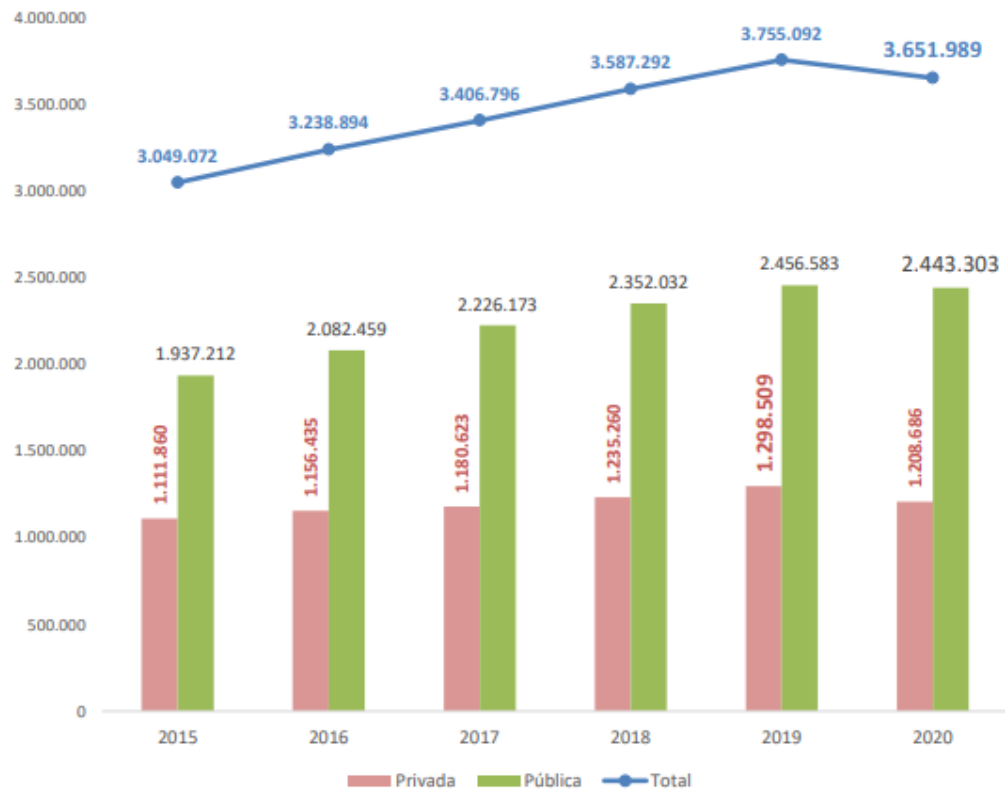
Durante essa etapa, a criança geralmente tem suas primeiras experiências sociais longe de pais e parentes, sendo crucial para descobrir novas formas de convívio social e conhecer melhor tanto a si próprio (do ponto de vista sensorial, emocional e intelectual) quanto do mundo que os cerca (relacionado ao contato com adultos, outras crianças, e a comunidade local). (INEP, 2020). Partindo desses princípios, as Diretrizes Curriculares Nacionais da Educação Infantil (DCNEI) estabelecem dois eixos estruturantes nas práticas pedagógicas: “interações” e “brincadeiras”.

A partir dos eixos delimitados pela DCNEI, a Base Nacional Comum Curricular (BNCC) estrutura cinco campos de experiências, sendo eles: “O eu, o outro, e o nós”, “Corpo, gestos e movimentos”, “Traços, sons, cores e formas”, “Escuta, fala, pensamento e imaginação”, e “Espaços, tempos, quantidades, relações e transformações” (Base Nacional Comum Curricular). Sendo que esses campos de experiência servem como base para os objetivos de aprendizagem e desenvolvimento, e consequentemente para a elaboração das disciplinas que serão lecionadas ao longo da educação infantil (INEP, 2020).

Em 2020, participavam dessa etapa de ensino 8,8 milhões de estudantes distribuídos em 70,9 mil creches e 101 mil pré-escolas. Os estudantes que atendiam às creches correspondiam a apenas 35,6% das crianças na idade ideal, por outro lado, 92,9% das crianças entre 4 e 5 anos atendem às pré-escolas (INEP, 2020).

A participação do poder público, em geral de governos municipais, é consideravelmente mais representativa como é possível observar nas Figura 1 e Figura 2. Vale ressaltar que na rede privada existem convênios com a rede pública, já que 50,9% dos alunos matriculados em creches da rede privada estão em instituições conveniadas com o poder público. No caso das pré-escolas, essa participação é menor: 149 mil dos alunos matriculados estão em instituições conveniadas (INEP, 2020).

Figura 1: Evolução das matrículas em creche por rede de ensino no Brasil (2015 a 2020)



Fonte: INEP, 2020

Figura 2: Evolução das matrículas em pré-escola por rede de ensino no Brasil (2015 a 2020)



Fonte: INEP, 2020

1.1.2 O ensino fundamental no Brasil

O ensino fundamental é a etapa intermediária da educação básica, e a mais longa, já que desde 2006 conta com duração de 9 anos. Se inicia quando a criança tem 6 anos de idade, e, segundo o Ministério da Educação (1996), visa a formação básica do cidadão por meio de:

I - O desenvolvimento da capacidade de aprender, tendo como meios básicos o pleno domínio da leitura, da escrita e do cálculo;

II - A compreensão do ambiente natural e social, do sistema político, da tecnologia, das artes e dos valores em que se fundamenta a sociedade;

III - O desenvolvimento da capacidade de aprendizagem, tendo em vista a aquisição de conhecimentos e habilidades e a formação de atitudes e valores;

IV - O fortalecimento dos vínculos de família, dos laços de solidariedade humana e de tolerância recíproca em que se assenta a vida social.

Essa etapa da educação básica é muitas vezes dividida entre os anos iniciais e os anos finais, sendo que este se caracteriza pelos 4 anos finais e aquele pelos 5 anos iniciais. Essa separação envolve principalmente a profundidade das disciplinas presentes na grade curricular dos alunos e a autonomia que também aumenta com o passar dos anos (MINISTÉRIO DA EDUCAÇÃO, 1996)

Durante os anos iniciais do ensino fundamental, os dois primeiros anos são focados tanto numa continuidade do aprendizado da educação infantil, quanto no processo de alfabetização (MINISTÉRIO DA EDUCAÇÃO, 2015). Nesse período, os alunos possuem aulas com professores generalistas que devem introduzir diferentes áreas do conhecimento, e instigar a curiosidade dos alunos.

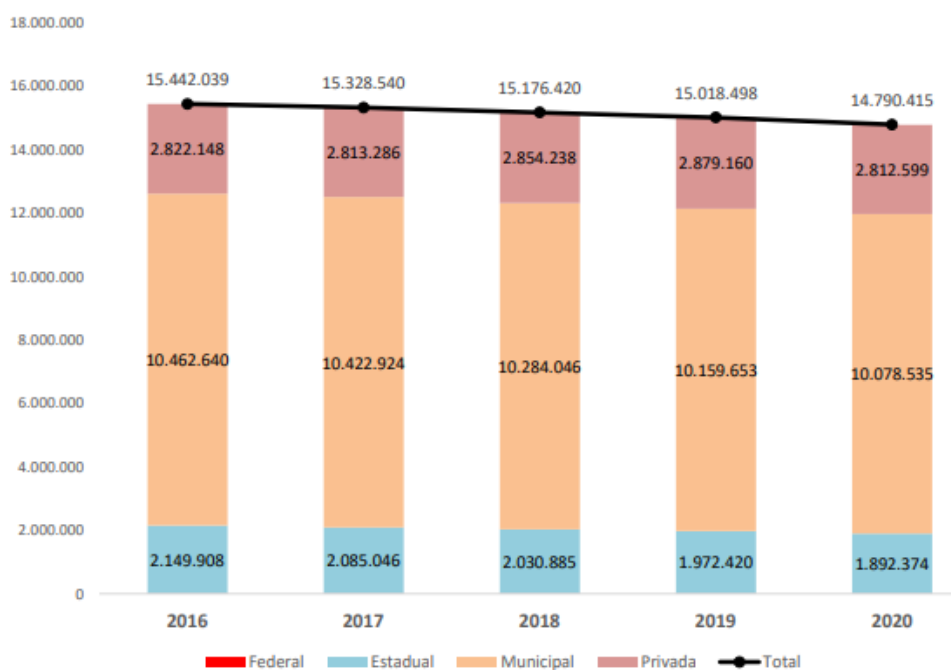
Os anos finais do ensino fundamental são caracterizados pelo aprofundamento dos conhecimentos vistos nos primeiros anos exigindo a compreensão de conceitos mais abstratos, que agora são lecionados por professores específicos para cada disciplina. Além disso, a autonomia dos alunos deve ser estimulada para que eles possam acessar e interagir com diferentes conhecimentos e fontes de informação (MINISTÉRIO DA EDUCAÇÃO, 2015).

As disciplinas que compõem o currículo escolar do ensino fundamental são formados por 5 áreas segundo o BNCC (Base Nacional Comum Curricular): área de Linguagens (da qual pertencem as disciplinas Língua Portuguesa, Arte, Educação Física e, nos anos finais, Língua Inglesa), área de Matemática (que pode ser dividida em Aritmética, Álgebra, Geometria, Estatística e Probabilidade), área de Ciências da Natureza (composta pela disciplina de Ciências), área de Ciências Humanas (que é segmentada nas disciplinas de Geografia e

História), e área de Ensino Religioso, que é facultativa e inclui a disciplina de Ensino Religioso (MINISTÉRIO DA EDUCAÇÃO, 1996; 2015).

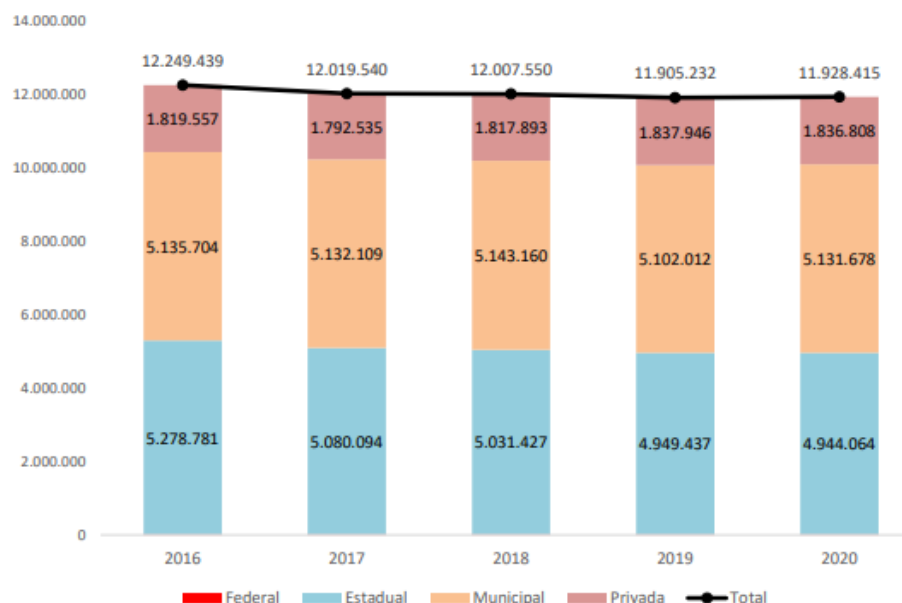
O ensino fundamental possuía o maior número de alunos dentre todas as etapas da educação básica. Em 2020, os anos iniciais possuíam 14,79 milhões de alunos, e os anos finais possuíam 11,9 milhões. Os alunos estavam matriculados em 108 mil escolas nos anos iniciais, e em 61,6 mil escolas nos anos finais. (INEP, 2020). Os números com mais detalhes podem ser vistos na Figura 3 e na Figura 4.

Figura 3: Evolução das matrículas nos anos iniciais do ensino fundamental por dependência administrativa no Brasil (2016 a 2020)



Fonte: INEP, 2020

Figura 4: Evolução das matrículas nos anos finais do ensino fundamental por dependência administrativa no Brasil (2016 a 2020)



Fonte: INEP, 2020

Em 2019, a taxa de atendimento de alunos em escolas do ensino fundamental era bastante elevada, atingindo valores de 99,8% da população de 6 a 10 anos e 99,5% da população de 10 a 14 anos, segundo a Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua – IBGE, 2019).

1.1.3 O ensino médio no Brasil

A última etapa do ensino básico no país é composta por três anos e consiste em aprofundar ainda mais as disciplinas lecionadas nos últimos anos do ensino fundamental. Essa etapa também prepara o aluno de forma básica para o trabalho, cumprimento da cidadania, desenvolvimento de pensamento crítico e autonomia intelectual, e compreensão de fundamentos científico-tecnológicos, de acordo com a Base Nacional Comum Curricular (MINISTÉRIO DA EDUCAÇÃO, 1996).

Recentemente, com a Lei nº 13.415 de 16 de fevereiro de 2017, o ensino médio passou por mudanças significativas em seu arranjo curricular, oferecendo maior flexibilidade aos alunos a partir da segmentação das competências gerais da educação básica em frentes da BNCC e em itinerários formativos (Base Nacional Comum Curricular). Portanto o currículo do ensino médio pode ser segmentado em: linguagens e suas tecnologias, matemáticas e suas

tecnologias, ciências da Natureza e suas tecnologias, ciências humanas e sociais aplicadas, e formação técnica e profissional.

Dentro dessa segmentação, segunda as Diretrizes Curriculares Nacionais do Ensino Médio publicadas em 2018 (DCNEM, MINISTÉRIO DA EDUCAÇÃO, 2018) existem disciplinas e áreas do conhecimento considerados essenciais que compõem a “formação geral básica” e são de caráter obrigatório em qualquer instituição que oferece a última etapa do ensino básico nacional, e áreas do conhecimento que compõe os “itinerários formativos” e possuem alguma flexibilidade no oferecimento pela escola ou na escolha do aluno.

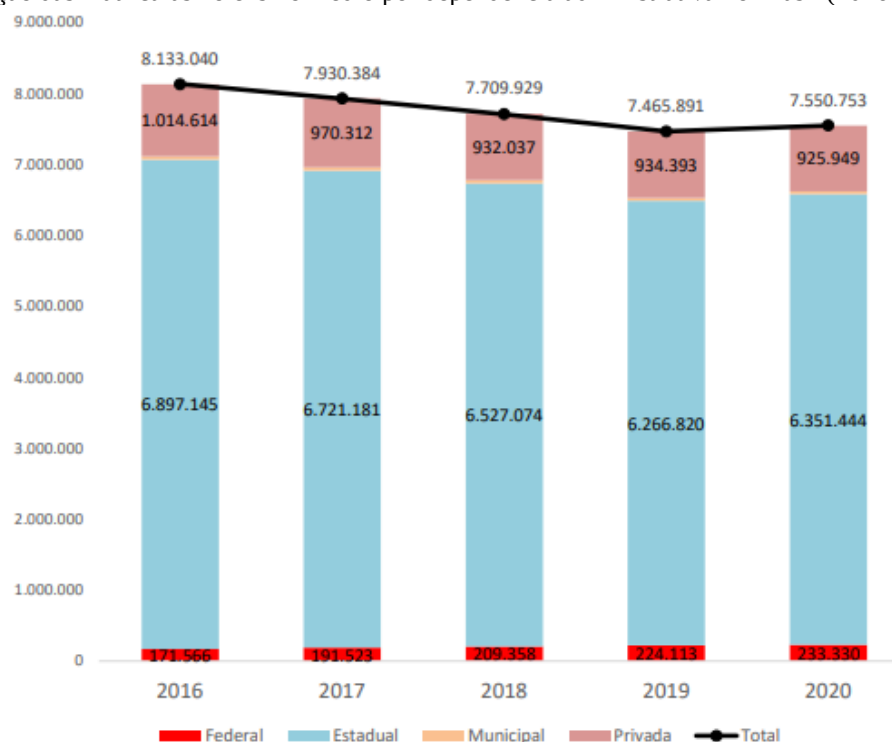
Conforme o DCNEM (MINISTÉRIO DA EDUCAÇÃO, 2018), integram a “formação geral básica” (Base Nacional Comum Curricular): língua portuguesa (ou indígena em suas respectivas comunidades); matemática; conhecimento do mundo físico e natural e da realidade social e política, especialmente do Brasil; educação física com prática facultativa; história do Brasil e do mundo; história e cultura afro-brasileira e indígena especialmente nos estudos de arte de literatura e histórias brasileiras; sociologia e filosofia; língua inglesa, podendo ser oferecidas outras línguas estrangeira em caráter optativo.

Segundo as mesmas Diretrizes (MINISTÉRIO DA EDUCAÇÃO, 2018), os “itinerários formativos” podem ser estruturados seguindo: ênfase em uma área do conhecimento, na formação técnica e de competências profissionais, ou na composição de “itinerários integrados” que consiste na mobilização de competências e habilidades de diferentes áreas (Base Nacional Comum Curricular). A partir dessa definição, formam os “itinerários formativos”: linguagens e suas tecnologias, matemática e suas tecnologias, ciências da natureza e suas tecnologias, ciências humanas e sociais aplicadas, e formação técnica e profissional. Sendo que todos eles consistem em aprofundamento de conhecimentos estruturantes para a aplicação de diferentes conceitos do itinerário em contextos sociais e de trabalho, com exceção do itinerário de formação técnica e profissional que apresenta caráter mais prático.

Comparecem ao ensino médio 89,2% da população entre 15 e 17 anos correspondente a 7,55 milhões de alunos. Essa etapa é majoritariamente gerida pelos governos estaduais (84,1% das matrículas) e possui a menor participação da gestão privada (12,3% das matrículas) (INEP, 2020).

Essa etapa possui algumas particularidades, como: 16,7% dos seus atendentes frequentam o período noturno e há maior participação do governo federal que nas demais etapas (INEP, 2020). Os valores em detalhes podem ser observados na Figura 5.

Figura 5: Evolução das matrículas no ensino médio por dependência administrativa no Brasil (2016 a 2020)



Fonte: INEP, 2020

1.2 DEFINIÇÃO DO PROBLEMA

A complexidade geográfica e demográfica do país sempre foi um empecilho para a definição de políticas públicas eficientes. Diversos experimentos e abordagens diferentes foram utilizadas ao longo dos anos para resolver todos os tipos de desafios que estão sob responsabilidade do poder público, porém sem sucesso em diversos casos. O motivo do fracasso dessas medidas pode ser em parte relacionado à replicação de medidas em cenários extremamente diferentes do retratado no caso de sucesso ou assumido no escopo da pesquisa.

O resultado da falta de aplicações de boas políticas públicas na educação nacional se reflete principalmente em indicadores internacionais de educação. Usando o Programa Internacional de Avaliação de Estudantes (PISA) que é coordenado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE) para avaliar o conhecimento de alunos de até 15 anos em leitura, matemática e ciências como referência, o Brasil se posicionou entre 79 países da seguinte forma (INEP, 2022):

- Leitura: Faixa entre 55° e 59°
- Matemática: Faixa entre 69° e 72°
- Ciências: Faixa entre 64° e 67°

Sendo que a situação se mostra ainda mais grave quando comparamos com os países sul-americanos participantes, em que o Brasil se encontra no último lugar em matemática e ciências, e penúltimo lugar no ranking de leitura.

Para que haja um avanço no desempenho dos alunos do ensino básico brasileiro, é necessário aplicar políticas públicas eficazes, capazes de resolver os problemas específicos de cada localidade. O agrupamento geográfico que considera as divisões federativas, contudo, nem sempre é capaz de reunir unidades municipais com características semelhantes. A correta identificação desses grupos, para além da divisão política, é um problema para a definição de políticas públicas eficazes.

1.3 OBJETIVOS

O avanço na capacidade de processar e armazenar grandes bases de dados, aliados à aplicação de métodos estatísticos, possibilitaram o surgimento de negócios bem-sucedidos e respostas para problemas mais complexos. Utilizando técnicas semelhantes, é possível replicar parte desse sucesso na resolução de grandes problemas públicos, como é o caso da educação.

Considerando a grande pluralidade de cenários que a educação básica nacional possui, este trabalho tem como principal objetivo: ajudar na identificação de *clusters* de cidades que compartilham características semelhantes (econômicas, sociais e educacionais) a fim de facilitar a replicação de modelos educacionais de sucesso, e a definição de medidas públicas mais condizentes com a sua realidade. Como objetivo secundário desse trabalho, tem-se a recomendação de trabalhar medidas públicas direcionadas a uma característica (condição dos docentes, estrutura das escolas, dentre outras), ou a um conjunto de características.

A definição desses *clusters* ou conglomerados é feita por meio da aplicação de métodos estatísticos e da análise de bases de dados fornecidas pelo governo, e a recomendação de políticas públicas envolverá a convergência entre *clusters* que levam em consideração diferentes indicadores.

1.4 ESTRUTURA DO TRABALHO

Este capítulo apresenta de forma sucinta um panorama da educação nacional, em seguida define os problemas que são enfrentados, e por fim delimita o escopo da solução que será proposta por meio dos objetivos traçados.

No capítulo seguinte, será apresentada a revisão bibliográfica do trabalho a fim de explicar a base teórica sob a qual a solução será construída, junto com a apresentação dos principais indicadores educacionais; em seguida o capítulo 3 que apresentará a metodologia a ser aplicada; na sequência os resultados da aplicação da metodologia serão aplicados no capítulo 4; e as análises e sugestões realizadas a partir do capítulo anterior serão desenvolvidas no capítulo 5; e por fim as conclusões deste trabalho estarão presente no capítulo 6.

2 REFERENCIAL TEÓRICO

2.1 INDICADORES EDUCACIONAIS ANALISADOS

A formação de *clusters* parte da necessidade de identificar semelhanças e diferenças no desempenho educacional das cidades analisadas. Tal desempenho foi analisado por meio de indicadores da área da educação. A coleta de indicadores foi orientada de forma a abordar quatro parâmetros distintos dentro de um sistema educacional, sendo eles: condição socioeconômica dos alunos, condições oferecidas aos docentes, quantidade de alunos por sala de aula, e aprendizado dos alunos ao final de uma etapa da educação.

Os indicadores selecionados buscaram representar todos os parâmetros apresentados dentro de um período semelhante, dessa forma os dados coletados se referem ao ano de 2019, antes da pandemia de Covid-19, em que existe uma maior variedade de dados disponíveis.

No total foram coletados cinco indicadores a partir de fontes do Governo Federal, principalmente do INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), uma autarquia vinculada ao Ministério da Educação, que auxilia na consolidação e divulgação de dados educacionais no Brasil.

Os indicadores coletados apresentaram separações entre localização (total, rural ou urbana) e dependência administrativa (total, municipal, estadual, federal, pública e privada). Neste trabalho, foram utilizados indicadores para as localizações rural e urbana separadamente, considerando apenas a dependência administrativa pública.

2.1.1 Sistema de Avaliação Básica (Saeb)

O Ministério da Educação, junto ao Inep, consolida anualmente o Sistema de Avaliação da Educação Básica (Saeb), que consiste na realização de testes e questionários aplicados em alunos do 5º e 9º ano do ensino fundamental e 3ª e 4ª séries do ensino médio, sendo que a aplicação é realizada de forma censitária para alunos de escolas públicas, e de forma amostral para alunos de escolas privadas. Neste trabalho, apenas dados relativos a alunos de escola pública são analisados (INEP, 2022b).

Os testes aplicados no Saeb têm como objetivo medir o desempenho de alunos em Língua Portuguesa e Matemática. Os testes de Língua Portuguesa mensuram a capacidade do aluno de compreender, analisar e interpretar textos nas mais diferentes situações. Os testes de

Matemática, por sua vez, visam medir a capacidade do aluno em resolver problemas por meio de sua habilidade em estabelecer relações entre o que está exposto no enunciado junto à sua capacidade de intuição, indução, dedução e estimativa de valores (INEP, 2022b). A elaboração do caderno de prova do Saeb utiliza a metodologia denominada Blocos Incompletos Balanceados – BIB (MONTGOMERY, 1984), que permite que muitos itens sejam aplicados aos alunos avaliados, sem que cada aluno precise responder a todos eles. Os testes de desempenho utilizam apenas itens de múltipla escolha como resposta, e são compostos por três partes: texto-base, enunciado e alternativas (INEP, 2022b).

Os questionários aplicados durante o Saeb utilizam como referência a legislação, e temas abordados em questionários aplicados em avaliação internacionais como o Programa Internacional de Avaliação de Estudantes (Pisa), e o *Trends International Results in Mathematics Science Study* (TIMSS). As perguntas presentes no questionário buscam contextualizar o resultado dos testes aplicados de forma a se ter uma visão mais holística sobre o desempenho dos alunos (INEP, 2022b).

Neste trabalho, os resultados do Saeb foram cruzados com os parâmetros de condição socioeconômica dos alunos, a fim de medir o aprendizado que estes obtiveram durante as etapas que são avaliadas nos testes. Os dados utilizados se referem ao ano de 2019, último ano em que os resultados foram divulgados até a data de realização deste trabalho.

2.1.2 Indicador socioeconômico

A condição socioeconômica que o aluno se encontra contribui tanto para a compreensão de suas necessidades, quanto para delimitar políticas públicas que contribuam para tornar uma sociedade mais igualitária (SIRIN, 2005; ALVES; SOARES, 2009; ALVES; SOARES; XAVIER, 2014).

O indicador utilizado neste trabalho é o Indicador de Nível Socioeconômico (Inse), que foi elaborado pela Diretoria de Avaliação da Educação Básica (Daeb), por meio da coleta de respostas de questionários apresentados aos alunos durante a aplicação do Saeb.

O questionário que constitui o Inse envolve a resposta de quinze perguntas sobre a escolaridade dos responsáveis dos alunos, e de características do domicílio em que o estudante reside, as questões podem ser observadas no **Error! Reference source not found..**

Quadro 1: Itens abordados pelo questionário Inse

Código	Item
--------	------

	Qual é a maior escolaridade de sua/seu...
Q008	Mãe (ou mulher responsável por você)?
Q009	Pai (ou homem responsável por você)?
	Dos itens relacionados abaixo, quantos existem na sua casa?
Q019	Geladeira?
Q021	Computador (ou <i>notebook</i>)
Q022	Quartos para dormir
Q023	Televisão
Q024	Banheiro
Q025	Carro
	Na sua casa tem:
Q027	Rede <i>wi-fi</i> ?
Q029	Mesa para estudar (ou escrivaninha)?
Q030	Garagem?
Q031	Forno de micro-ondas?
Q032	Aspirador de pó?
Q033	Máquina de lavar roupa?
Q034	<i>Freezer</i> (independente ou segunda porta da geladeira)?

Fonte: Adaptado de Inep (2022b)

As respostas do questionário são utilizadas no cálculo do Inse e na elaboração de 8 níveis delimitados a partir número de desvio-padrão que uma faixa de alunos se distancia da média. Para o cálculo dos valores do Inse, são considerados apenas municípios que atingem o mínimo de 10 alunos e percentual de participação de 50%. O Fonte: Adaptado de INEP (2022b) apresenta a descrição dos 8 níveis elaborados pelo Daeb/ Inep (INEP, 2022b).

Quadro 2: Níveis do Daeb para indicadores socioeconômicos

Níveis	Descrição
Nível 1	Este é o nível inferior da escala, no qual os estudantes têm dois ou mais desvios-padrão abaixo da média nacional do Inse. Considerando a maioria dos estudantes, o pai/responsável não completou o 5º ano do ensino fundamental e a mãe/responsável tem o 5º ano do ensino fundamental incompleto ou completo. A maioria dos estudantes deste nível possui uma geladeira, um ou dois quartos, uma televisão e um banheiro. Mas não possui muitos dos bens e serviços pesquisados (i.e., computador, carro, wi-fi, mesa para estudar, garagem, micro-ondas, aspirador de pó, máquina de lavar roupa e freezer).
Nível 2	Neste nível, os estudantes estão entre um e dois desvios-padrão abaixo da média nacional do Inse. Considerando a maioria dos estudantes, a mãe/responsável e/ou o pai/responsável tem o 5º ano do ensino fundamental incompleto ou completo. A maioria possui uma geladeira, um ou dois quartos, uma televisão e um banheiro. Mas não possui muitos dos bens e serviços pesquisados, exceto uma parte dos estudantes deste nível passa a ter freezer, máquina de lavar roupa e três ou mais quartos para dormir em sua casa.
Nível 3	Neste nível, os estudantes estão entre meio e um desvio-padrão abaixo da média nacional do Inse. Considerando a maioria dos estudantes, a mãe/responsável e o pai/responsável têm o ensino fundamental incompleto ou completo e/ou ensino médio completo. A maioria possui uma geladeira, um ou dois quartos, uma televisão, um banheiro, wi-fi e máquina de lavar roupas, mas não possui computador, carro, garagem e aspirador de pó. Parte dos estudantes passa a ter também freezer e forno de micro-ondas.
Nível 4	Neste nível, os estudantes estão até meio desvio-padrão abaixo da média nacional do Inse. Considerando a maioria dos estudantes, a mãe/responsável e o pai/responsável têm o ensino fundamental incompleto ou completo e/ou ensino médio completo. A maioria

	possui uma geladeira, um ou dois quartos, um banheiro, wi-fi, máquina de lavar roupas e freezer, mas não possui aspirador de pó. Parte dos estudantes deste nível passa a ter também computador, carro, mesa de estudos, garagem, forno de micro-ondas e uma ou duas televisões.
Nível 5	Neste nível, os estudantes estão até meio desvio-padrão acima da média nacional do Inse. Considerando a maioria dos estudantes, a mãe/responsável tem o ensino médio completo ou ensino superior completo, o pai/responsável tem do ensino fundamental completo até o ensino superior completo. A maioria possui uma geladeira, um ou dois quartos, um banheiro, wi-fi, máquina de lavar roupas, freezer, um carro, garagem, forno de micro-ondas. Parte dos estudantes deste nível passa a ter também dois banheiros.
Nível 6	Neste nível, os estudantes estão de meio a um desvio-padrão acima da média nacional do Inse. Considerando a maioria dos estudantes, a mãe/responsável e/ou o pai/responsável têm o ensino médio completo ou o ensino superior completo. A maioria possui uma geladeira, dois ou três ou mais quartos, um banheiro, wi-fi, máquina de lavar roupas, freezer, um carro, garagem, forno de micro-ondas, mesa para estudos e aspirador de pó. Parte dos estudantes deste nível passa a ter também dois ou mais computadores e três ou mais televisões.
Nível 7	Neste nível, os estudantes estão de um a dois desvios-padrão acima da média nacional do Inse. Considerando a maioria dos estudantes, a mãe/responsável e/ou o pai/responsável têm ensino médio completo ou ensino superior completo. A maioria possui uma geladeira, três ou mais quartos, um banheiro, wi-fi, máquina de lavar roupas, freezer, um carro, garagem, forno de micro-ondas, mesa para estudos e aspirador de pó. Parte dos estudantes deste nível passa a ter também dois ou mais carros, três ou mais banheiros e duas ou mais geladeiras.
Nível 8	Este é o nível superior da escala, no qual os estudantes estão dois desvios-padrão ou mais acima da média nacional do Inse. Considerando a maioria dos estudantes, a mãe/responsável e/ou o pai/responsável têm ensino superior completo. Além de possuírem os bens dos níveis anteriores, a maioria dos estudantes deste nível passa a ter duas ou mais geladeiras, dois ou mais computadores, três ou mais televisões, três ou mais banheiros e dois ou mais carros.

Fonte: Adaptado de INEP (2022b)

O indicador aplicado neste trabalho é uma adaptação a partir dos níveis elaborados no Inse, utilizando o percentual de alunos que estão em um nível acima ou igual ao nível cinco, uma vez que a partir deste nível os alunos se encontram acima da média nacional e os pais ou responsáveis possuem ensino fundamental ou médio completo.

2.1.3 Indicadores sobre a condição dos docentes

Os indicadores selecionados para avaliar a condição dos docentes estão disponíveis no site do Inep, e são o Esforço Docente e a Regularidade do Corpo Docente. Estes indicadores mensuram informações como: a quantidade de alunos para qual o docente leciona, o número de escolas em que o docente atua, o número de etapas de ensino que o docente participa, e o número de anos que o docente está lecionando na mesma escola.

Os dados coletados para a elaboração desses indicadores ocorreram durante o Censo Escolar, e desconsideram turmas que praticam ensino semipresencial ou a distância (EaD) pois os horários praticados por essas turmas não eram coletados pelo Censo Escolar (INEP, 2019).

2.1.3.1 Indicador esforço docente

O indicador de esforço docente está dividido em seis níveis, sendo que níveis mais elevados demonstram maior esforço. Os dados coletados pelo Censo Escolar que delimitam o esforço do docente são: número de escolas em que atua, número de turnos de trabalho, número de alunos atendidos e número de etapas nas quais leciona. Os seis níveis, e as características de cada um podem ser observados no Fonte: Adaptado de INEP (2022b).

Quadro 3: Descrição dos níveis de avaliação de esforço docente	
Nível	Descrição
Nível 1	Docente que, em geral, tem até 25 alunos e atua em um único turno, escola e etapa.
Nível 2	Docente que, em geral, tem entre 25 e 150 alunos e atua em um único turno, escola e etapa.
Nível 3	Docente que, em geral, tem entre 25 e 300 alunos e atua em um ou dois turnos em uma única escola e etapa.
Nível 4	Docente que, em geral, tem entre 50 e 400 alunos e atua em dois turnos, em uma ou duas escolas e em duas etapas.
Nível 5	Docente que, em geral, tem mais de 300 alunos e atua nos três turnos, em duas ou três escolas e em duas etapas ou três etapas.
Nível 6	Docente que, em geral, tem mais de 400 alunos e atua nos três turnos, em duas ou três escolas e em duas etapas ou três etapas.

Fonte: Adaptado de INEP (2022b)

Neste trabalho, o indicador sobre o esforço docente foi adaptado de forma a mensurar o percentual de docentes que atuam em nível de esforço menor ou igual a 3, sendo o nível 3 o valor que precede a participação do docente em mais de uma escola e etapa.

2.1.3.2 Indicador regularidade do corpo docente

A regularidade do corpo docente está relacionada ao tempo de permanência dos professores nas escolas que lecionam ou lecionaram nos últimos cinco anos. A cada período de permanência é atrelada uma pontuação, sendo que quanto mais próximos a cinco mais regular é o vínculo do professor à instituição, e mais próximos a zero menos regular é este vínculo. A regularidade do vínculo depende não só do tempo de permanência, mas também da continuidade de atuação do professor na escola durante esses anos (INEP, 2022b).

De acordo com a sua pontuação, os docentes são alocados dentro de 4 faixas de regularidade no censo escolar. O Fonte: Adaptado de INEP (2022b)

apresenta a classificação dada para cada uma das faixas.

Quadro 4: Faixas de regularidade do corpo docente

Faixa de Regularidade	Descrição
Baixa Regularidade	Indicador de Regularidade do corpo Docente médio igual ou menor que 2
Média-Baixa	Indicador de Regularidade do corpo Docente médio maior que 2 até 3
Média-Alta	Indicador de Regularidade do corpo Docente médio maior que 3 até 4
Alta	Indicador de Regularidade do corpo Docente médio maior que 4 até 5

Fonte: Adaptado de INEP (2022b)

Neste trabalho, os valores fornecidos pelo Daeb (2019) foram adaptados de forma a calcular o valor percentual de docentes que estão nas faixas de regularidade Média-Alta e Alta. Ou seja, valores do Indicador de Regularidade do corpo Docente maior que 3 até 5.

2.1.4 Indicador sobre a quantidade de alunos por sala

O indicador sobre a quantidade de alunos por sala foi o único indicador deste trabalho que não sofreu alterações durante a sua implementação, uma vez que ele se trata uma média simples entre o número de alunos pelo número de turmas na cidade em determinada série ou etapa de ensino. Os dados desse indicador, assim como os de condição dos docentes, foram obtidos durante o Censo Escolar.

Neste trabalho, os valores utilizados foram os de alunos por turma nos anos iniciais do ensino fundamental, alunos por turma nos anos finais do ensino fundamental, e alunos por turma do ensino médio.

2.1.5 Indicador sobre o aprendizado dos alunos

O indicador de aprendizado dos alunos foi obtido a partir do desempenho dos alunos nos testes aplicados no Saeb. De forma semelhante aos indicadores socioeconômico e de condição dos docentes, os testes aplicados pelo Saeb dividem os resultados em níveis de acordo a pontuação. Porém, o Saeb divide os alunos de acordo com as capacidades que estes demonstraram em Língua Portuguesa e Matemática em diferentes etapas de ensino, ou seja, os níveis se ajustam de acordo com o que esperado de um aluno em sua respectiva série ou ano (INEP, 2022b).

Os testes são aplicados para alunos do 5º ano do ensino fundamental (também denominado anos iniciais do ensino fundamental), para alunos do 9º ano do ensino fundamental (também denominados anos finais do ensino fundamental), e para alunos do ensino médio.

2.1.5.1 *Níveis do Saeb para os anos iniciais do ensino fundamental*

2.1.5.1.1 Língua portuguesa

Existem 9 níveis de desempenho em língua portuguesa na medição realizada pelo Saeb para alunos dos anos iniciais do ensino fundamental, sendo que para pontuações inferiores a 125 não é possível inferir quais conhecimentos o aluno obteve. O Fonte: Adaptado de INEP (2022b) traz a descrição dos níveis de desempenho em língua portuguesa nos anos iniciais do ensino fundamental.

Quadro 5: Descrição dos níveis de desempenho em língua portuguesa nos anos iniciais do ensino fundamental

Nível	Descrição das Habilidades Envolvidas
Nível 1 – Desempenho maior ou igual a 125 e menor que 150	Os estudantes provavelmente são capazes de: localizar informações explícitas em textos narrativos curtos, informativos e anúncios. Identificar o tema de um texto. Localizar elementos, como o personagem principal. Estabelecer relação entre partes do texto: personagem e ação; ação e tempo; ação e lugar.
Nível 2 – Desempenho maior ou igual a 150 e menor que 175	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informações explícitas em contos. Identificar o assunto principal e a personagem principal em reportagens e em fábulas. Reconhecer a finalidade de receitas, manuais e regulamentos. Inferir características de personagens em fábulas. Interpretar linguagem verbal e não verbal em tirinhas.
Nível 3 – Desempenho maior ou igual a 175 e menor que 200	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informação explícita em contos e reportagens. Localizar informação explícita em propagandas com ou sem apoio de recursos gráficos. Reconhecer relação de causa e consequência em poemas, contos e tirinhas. Inferir o sentido de palavra, o sentido de expressão ou o assunto em cartas, contos, tirinhas e histórias em quadrinhos, com o apoio de linguagem verbal e não verbal.
Nível 4 – Desempenho maior ou igual a 200 e menor que 225	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: identificar informação explícita em sinopses e receitas culinárias. Identificar assunto principal e personagem em contos e letras de música. Identificar formas de representação de medida de tempo em reportagens. Identificar assuntos comuns a duas reportagens. Identificar o efeito de humor em piadas. Reconhecer sentido de expressão, elementos da narrativa e opinião em reportagens, contos e poemas. Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em fábulas, poemas, contos e tirinhas. Inferir sentido decorrente da utilização de sinais de pontuação e sentido de expressões em poemas, fábulas e contos. Inferir efeito de humor em tirinhas e histórias em quadrinhos.

Nível 5 – Desempenho maior ou igual a 225 e menor que 250	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: identificar assunto e opinião em reportagens e contos. Identificar assunto comum a cartas e poemas. Identificar informação explícita em letras de música e contos. Reconhecer assunto em poemas e tirinhas. Reconhecer sentido de conjunções e de locuções adverbiais em verbetes, lendas e contos. Reconhecer finalidade de reportagens e cartazes. Reconhecer relação de causa e consequência e relação entre pronome e seu referente em tirinhas, contos e reportagens. Inferir elementos da narrativa em fábulas, contos e cartas. Inferir finalidade e efeito de sentido decorrente do uso de pontuação e assunto em fábulas. Inferir informação em poemas, reportagens e cartas. Diferenciar opinião de fato em reportagens. Interpretar efeito de humor e sentido de palavra em piadas e tirinhas.
Nível 6 – Desempenho maior ou igual a 250 e menor que 275	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: identificar opinião e informação explícita em fábulas, contos, crônicas e reportagens. Identificar informação explícita em reportagens com ou sem o auxílio de recursos gráficos. Reconhecer a finalidade de verbetes, fábulas, charges e reportagens. Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em poemas, fábulas e contos. Inferir assunto principal e sentido de expressão em poemas, fábulas, contos, crônicas, reportagens e tirinhas. Inferir informação em contos e reportagens. Inferir efeito de humor e moral em piadas e fábulas.
Nível 7 – Desempenho maior ou igual a 375 e menor que 300	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: identificar assunto principal e informações explícitas em poemas, fábulas e letras de música. Identificar opinião em poemas e crônicas. Reconhecer o gênero textual a partir da comparação entre textos e assunto comum a duas reportagens. Reconhecer elementos da narrativa em fábulas. Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em fábulas, contos e crônicas. Inferir informação e efeito de sentido decorrente do uso de sinais gráficos em reportagens e em letras de música. Interpretar efeito de humor em piadas e contos. Interpretar linguagem verbal e não verbal em histórias em quadrinhos
Nível 8 – Desempenho maior ou igual a 300 e menor que 325	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: identificar assunto principal e opinião em contos e cartas do leitor. Reconhecer sentido de locução adverbial e elementos da narrativa em fábulas e contos. Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em fábulas e reportagens. Reconhecer assunto comum entre textos de gêneros diferentes. Inferir informações e efeito de sentido decorrente do uso de pontuação em fábulas e piadas.
Nível 9 – Desempenho maior ou igual a 325	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: identificar opinião em fábulas e reconhecer sentido de advérbios em cartas do leitor.

Fonte: Adaptado de INEP (2022b)

2.1.5.1.2 Matemática

Os alunos dos anos iniciais do ensino fundamental podem se encaixar em 10 níveis de desempenho em matemática na medição realizada pelo Saeb, sendo que para valores inferiores a 125 não é possível inferir quais conhecimentos o aluno obteve. O **Error! Reference source not found.** mostra a descrição dos níveis de desempenho em matemática nos anos iniciais do ensino fundamental.

Quadro 6: Descrição dos níveis de desempenho em matemática nos anos iniciais do ensino fundamental

Nível	Descrição
Nível 1 – Desempenho maior ou igual a 125 e menor que 150	Os estudantes provavelmente são capazes de: Grandezas e medidas – Determinar a área de figuras desenhadas em malhas quadriculadas por meio de contagem.
Nível 2 – Desempenho maior ou igual a 150 e menor que 175	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Números e operações; álgebra e funções – Resolver problemas do cotidiano envolvendo adição de pequenas quantias de dinheiro. Tratamento de informações – Localizar informações, relativas ao maior ou menor elemento, em tabelas ou gráficos.
Nível 3 – Desempenho maior ou igual a 175 e menor que 200	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Localizar um ponto ou objeto em uma malha quadriculada ou croqui, a partir de duas coordenadas ou duas ou mais referências. Reconhecer, entre um conjunto de polígonos, aquele que possui o maior número de ângulos. Associar figuras geométricas elementares (quadrado, triângulo e círculo) a seus respectivos nomes. Grandezas e medidas – Converter uma quantia, dada na ordem das unidades de real, em seu equivalente em moedas. Determinar o horário final de um evento a partir de seu horário de início e de um intervalo de tempo dado, todos no formato de horas inteiras. Números e operações; álgebra e funções – Associar a fração $\frac{1}{4}$ a uma de suas representações gráficas. Determinar o resultado da subtração de números representados na forma decimal, tendo como contexto o sistema monetário. Tratamento de informações – Reconhecer o maior valor em uma tabela de dupla entrada cujos dados possuem até duas ordens. Reconhecer informações em um gráfico de colunas duplas.
Nível 4 – Desempenho maior ou igual a 200 e menor que 225	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Reconhecer retângulos em meio a outros quadriláteros. Reconhecer a planificação de uma pirâmide entre um conjunto de planificações. Grandezas e medidas – Determinar o total de uma quantia a partir da quantidade de moedas de 25 e/ou 50 centavos que a compõe, ou vice-versa. Determinar a duração de um evento cujos horários inicial e final acontecem em minutos diferentes de uma mesma hora dada. Converter uma hora em minutos. Converter mais de uma semana inteira em dias. Interpretar horas em relógios de ponteiros. Números e operações; álgebra e funções – Determinar o resultado da multiplicação de números naturais por valores do sistema monetário nacional, expressos em números de até duas ordens e efetuar adição posterior. Determinar os termos desconhecidos em uma sequência numérica de múltiplos de cinco. Determinar a adição, com reserva, de até três números naturais com até quatro ordens. Determinar a subtração de números naturais usando a noção de completar. Determinar a multiplicação de um número natural de até três ordens por cinco, com reserva. Determinar a divisão exata por números de um algarismo. Reconhecer o princípio do valor posicional do Sistema de Numeração Decimal. Reconhecer uma fração como representação da relação parte-todo, com o apoio de um conjunto de até cinco figuras. Associar a metade de um total a seu equivalente em porcentagem. Associar um número natural à sua decomposição expressa por extenso. Localizar um número em uma reta numérica graduada onde estão expressos números naturais consecutivos e uma subdivisão equivalente à metade do intervalo entre eles. Tratamento de informações – Reconhecer o maior valor em uma tabela cujos dados possuem até oito ordens. Localizar um dado em tabelas de dupla entrada.
Nível 5 – Desempenho maior ou igual a 225 e menor que 250	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Localizar um ponto entre outros dois fixados, apresentados em uma figura composta por vários outros pontos. Reconhecer a planificação de um cubo entre um conjunto de planificações apresentadas. Grandezas e medidas – Determinar a área de um terreno retangular representado em uma malha quadriculada. Determinar o horário final de um evento a partir do horário de início, dado em horas e minutos, e de um intervalo dado em quantidade

	de minutos superior a uma hora. Converter mais de uma hora inteira em minutos. Converter uma quantia dada em moedas de 5, 25 e 50 centavos e 1 real em cédulas de real. Estimar a altura de um determinado objeto com referência aos dados fornecidos por uma régua graduada em centímetros. Números e operações; álgebra e funções – Determinar o resultado da subtração, com recursos à ordem superior, entre números naturais de até cinco ordens, utilizando as ideias de retirar e comparar. Determinar o resultado da multiplicação de um número inteiro por um número representado na forma decimal, em contexto envolvendo o sistema monetário. Determinar o resultado da divisão de números naturais, com resto, por um número de uma ordem, usando noção de agrupamento. Resolver problemas envolvendo a análise do algoritmo da adição de dois números naturais. Resolver problemas, no sistema monetário nacional, envolvendo adição e subtração de cédulas e moedas. Resolver problemas que envolvam a metade e o triplo de números naturais. Localizar um número em uma reta numérica graduada onde estão expressos o primeiro e o último número representando um intervalo de tempo de dez anos, com dez subdivisões entre eles. Localizar um número racional dado em sua forma decimal em uma reta numérica graduada onde estão expressos diversos números naturais consecutivos, com dez subdivisões entre eles. Reconhecer o valor posicional do algarismo localizado na 4ª ordem de um número natural. Reconhecer uma fração como representação da relação parte-todo, com apoio de um polígono dividido em oito partes ou mais. Associar um número natural às suas ordens e vice-versa.
Nível 6 – Desempenho maior ou igual a 250 e menor que 275	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Reconhecer polígonos presentes em um mosaico composto por diversas formas geométricas. Grandezas e medidas – Determinar a duração de um evento a partir dos horários de início e de término, informados em horas e minutos, sem coincidência nas horas ou nos minutos dos dois horários informados. Converter a duração de um intervalo de tempo, dado em horas e minutos, para minutos. Resolver problemas envolvendo intervalos de tempo em meses, inclusive passando pelo final do ano (outubro a janeiro). Reconhecer que entre quatro ladrilhos apresentados, quanto maior o ladrilho, menor a quantidade necessária para cobrir uma dada região. Reconhecer o m² como unidade de medida de área. Números e operações; álgebra e funções – Determinar o resultado da diferença entre dois números racionais representados na forma decimal. Determinar o resultado da multiplicação de um número natural de uma ordem por outro de até três ordens, em contexto que envolve o conceito de proporcionalidade. Determinar o resultado da divisão exata entre dois números naturais, com divisor de até quatro, e dividendo de até quatro ordens. Determinar 50% de um número natural com até três ordens. Determinar porcentagens simples (25%, 50%). Associar a metade de um total a algum equivalente, apresentado como fração ou porcentagem. Associar números naturais à quantidade de agrupamentos de 1.000. Reconhecer uma fração como representação da relação parte-todo, sem apoio de figuras. Localizar números em uma reta numérica graduada onde estão expressos diversos números naturais não consecutivos e crescentes, com uma subdivisão entre eles. Resolver problemas por meio da realização de subtrações e divisões, para determinar o valor das prestações de uma compra a prazo (sem incidência de juros). Resolver problemas que envolvam soma e subtração de valores monetários. Resolver problemas que envolvam a composição e a decomposição polinomial de números naturais de até cinco ordens. Resolver problemas que utilizam a multiplicação envolvendo a noção de proporcionalidade. Reconhecer a modificação sofrida no valor de um número quando um algarismo é alterado. Reconhecer que um número não se altera ao multiplicá-lo por 1. Tratamento de informações – Interpretar dados em uma tabela simples. Comparar dados representados pelas alturas de colunas presentes em um gráfico.
Nível 7 – Desempenho maior ou igual a 375 e menor que 300	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Interpretar a movimentação de um objeto utilizando referencial diferente do seu. Reconhecer um cubo a partir de uma de suas

	planificações desenhadas em uma malha quadriculada. Grandezas e medidas – Determinar o perímetro de um retângulo desenhado em malha quadriculada, com as medidas de comprimento e largura explicitados. Converter medidas dadas em toneladas para quilogramas. Converter uma quantia, dada na ordem das dezenas de real, em moedas de 50 centavos. Estimar o comprimento de um objeto a partir de outro, dado como unidade padrão de medida. Resolver problemas envolvendo conversão de quilograma para grama. Resolver problemas envolvendo conversão de litro para mililitro. Resolver problemas sobre intervalos de tempo que envolvam adição e subtração e que passem pela meia-noite. Números e operações; álgebra e funções – Determinar 25% de um número múltiplo de quatro. Determinar a quantidade de dezenas presentes em um número de quatro ordens. Resolver problemas que envolvam a divisão exata ou a multiplicação de números naturais. Associar números naturais à quantidade de agrupamentos menos usuais, como 300 dezenas. Tratamento de informações – Interpretar dados em gráficos de setores.
Nível 8 – Desempenho maior ou igual a 300 e menor que 325	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Reconhecer uma linha paralela a outra, dada como referência em um mapa. Reconhecer os lados paralelos de um trapézio expressos em forma de segmentos de retas. Reconhecer objetos com a forma esférica entre uma lista de objetos do cotidiano. Grandezas e medidas – Determinar a área de um retângulo desenhado numa malha quadriculada, após a modificação de uma de suas dimensões. Determinar a razão entre as áreas de duas figuras desenhadas numa malha quadriculada. Determinar a área de uma figura poligonal não convexa desenhada sobre uma malha quadriculada. Estimar a diferença de altura entre dois objetos, a partir da altura de um deles. Converter medidas lineares de comprimento (m/cm). Resolver problemas que envolvem a conversão entre diferentes unidades de medida de massa. Números e operações; álgebra e funções – Resolver problemas que envolvam grandezas diretamente proporcionais requerendo mais de uma operação. Resolver problemas envolvendo divisão de números naturais com resto. Associar a fração $\frac{1}{2}$ à sua representação na forma decimal. Associar 50% à sua representação na forma de fração. Associar um número natural de seis ordens à sua forma polinomial. Tratamento de informações – Interpretar dados em um gráfico de colunas duplas.
Nível 9 – Desempenho maior ou igual a 325 e menor que 350	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Reconhecer a planificação de uma caixa cilíndrica. Grandezas e medidas – Determinar o perímetro de um polígono não convexo desenhado sobre as linhas de uma malha quadriculada. Resolver problemas que envolvam a conversão entre unidades de medida de tempo (minutos em horas, meses em anos). Resolver problemas que envolvam a conversão entre unidades de medida de comprimento (metros em centímetros). Números e operações; álgebra e funções – Determinar o minuendo de uma subtração entre números naturais, de três ordens, a partir do conhecimento do subtraendo e da diferença. Determinar o resultado da multiplicação entre o número oito e um número de quatro ordens com reserva. Reconhecer frações equivalentes. Resolver problemas envolvendo multiplicação com significado de combinatória. Comparar números racionais com quantidades diferentes de casas decimais. Tratamento de informações – Reconhecer o gráfico de linhas correspondente a uma sequência de valores ao longo do tempo (com valores positivos e negativos).
Nível 10 – Desempenho maior ou igual a 350	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Reconhecer entre um conjunto de quadriláteros, aquele que possui lados perpendiculares e com a mesma medida. Grandezas e medidas – Converter uma medida de comprimento, expressando decímetros e centímetros, para milímetros.

Fonte: Adaptado de INEP (2022b)

2.1.5.2 Níveis do Saeb para os anos finais do ensino fundamental

2.1.5.2.1 Língua Portuguesa

Os alunos dos anos finais do ensino fundamental se encaixam em 8 níveis de desempenho em língua portuguesa na medição realizada pelo Saeb, e alunos que tiveram desempenho inferior a 200 se encontram em uma categoria em que não é possível determinar o seu conhecimento. O Fonte: Adaptado de INEP (2022b) mostra a descrição das habilidades envolvidas na avaliação de Língua Portuguesa nos anos finais do ensino fundamental.

Quadro 7: Níveis de desempenho em Língua Portuguesa para anos finais do ensino fundamental

Nível	Descrição das Habilidades Envolvidas
Nível 1 – Desempenho maior ou igual a 200 e menor que 225	Os estudantes provavelmente são capazes de reconhecer expressões características da linguagem (científica, jornalística etc.) e a relação entre expressão e seu referente em reportagens e artigos de opinião. Inferir o efeito de sentido de expressão e opinião em crônicas e reportagens.
Nível 2 – Desempenho maior ou igual a 225 e menor que 250	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informações explícitas em fragmentos de romances e crônicas. Identificar tema e assunto em poemas e charges, relacionando elementos verbais e não verbais. Reconhecer o sentido estabelecido pelo uso de expressões, de pontuação, de conjunções em poemas, charges e fragmentos de romances. Reconhecer relações de causa e consequência e características de personagens em lendas e fábulas. Reconhecer recurso argumentativo em artigos de opinião. Inferir efeito de sentido de repetição de expressões em crônicas.
Nível 3 – Desempenho maior ou igual a 250 e menor que 275	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informações explícitas em crônicas e fábulas. Identificar os elementos da narrativa em letras de música e fábulas. Reconhecer a finalidade de abaixo-assinado e verbetes. Reconhecer relação entre pronomes e seus referentes e relações de causa e consequência em fragmentos de romances, diários, crônicas, reportagens e máximas (provérbios). Interpretar o sentido de conjunções, de advérbios, e as relações entre elementos verbais e não verbais em tirinhas, fragmentos de romances, reportagens e crônicas. Comparar textos de gêneros diferentes que abordem o mesmo tema. Inferir tema e ideia principal em notícias, crônicas e poemas. Inferir o sentido de palavra ou expressão em história em quadrinhos, poemas e fragmentos de romances.
Nível 4 – Desempenho maior ou igual a 275 e menor que 300	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informações explícitas em artigos de opinião e crônicas. Identificar finalidade e elementos da narrativa em fábulas e contos. Reconhecer opiniões distintas sobre o mesmo assunto em reportagens, contos e enquetes. Reconhecer relações de causa e consequência e relações entre pronomes e seus referentes em fragmentos de romances, fábulas, crônicas, artigos de opinião e reportagens. Reconhecer o sentido de expressão e de variantes linguísticas em letras de música, tirinhas, poemas e fragmentos de romances. Inferir tema, tese e ideia principal em contos, letras de música, editoriais, reportagens, crônicas e artigos. Inferir o efeito de sentido de linguagem verbal e não verbal em charges e história em quadrinhos. Inferir informações em fragmentos de romance. Inferir o efeito de sentido da pontuação e da polissemia como recurso para estabelecer humor ou ironia em tirinhas, anedotas e contos.
Nível 5 – Desempenho maior ou igual a 300 e menor que 325	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar a informação principal em reportagens. Identificar ideia principal e finalidade em notícias, reportagens e resenhas. Reconhecer características da linguagem (científica, jornalística etc.) em reportagens. Reconhecer elementos da

			narrativa em crônicas. Reconhecer argumentos e opiniões em notícias, artigos de opinião e fragmentos de romances. Diferenciar abordagem do mesmo tema em textos de gêneros distintos. Inferir informação em contos, crônicas, notícias e charges. Inferir sentido de palavras, da repetição de palavras, de expressões, de linguagem verbal e não verbal e de pontuação em charges, tirinhas, contos, crônicas e fragmentos de romances.
Nível 6	Desempenho maior ou igual a 325 e menor que 350	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Identificar ideia principal e elementos da narrativa em reportagens e crônicas. Identificar argumento em reportagens e crônicas. Reconhecer o efeito de sentido da repetição de expressões e palavras, do uso de pontuação, de variantes linguísticas e de figuras de linguagem em poemas, contos e fragmentos de romances. Reconhecer a relação de causa e consequência em contos. Reconhecer diferentes opiniões entre cartas de leitor que abordam o mesmo tema. Reconhecer a relação de sentido estabelecida por conjunções em crônicas, contos e cordéis. Reconhecer o tema comum entre textos de gêneros distintos. Reconhecer o efeito de sentido decorrente do uso de figuras de linguagem e de recursos gráficos em poemas e fragmentos de romances. Diferenciar fato de opinião em artigos e reportagens. Inferir o efeito de sentido de linguagem verbal e não verbal em tirinhas.
Nível 7	Desempenho maior ou igual a 350 e menor que 375	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informações explícitas, ideia principal e expressão que causa humor em contos, crônicas e artigos de opinião. Identificar variantes linguísticas em letras de música. Reconhecer a finalidade e a relação de sentido estabelecida por conjunções em lendas e crônicas.
Nível 8	Desempenho maior ou igual a 375	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar ideia principal em manuais, reportagens, artigos e teses. Identificar os elementos da narrativa em contos e crônicas. Diferenciar fatos de opiniões e opiniões diferentes em artigos e notícias. Inferir o sentido de palavras em poemas.

Fonte: Adaptado de INEP (2022b)

2.1.5.2.2 Matemática

Para os alunos dos anos finais do ensino fundamental, o Saeb delimita 9 níveis de desempenho para matemática. Não é possível inferir quais habilidades foram desenvolvidas pelos Alunos com desempenho inferior a 200 pontos. O Fonte: Adaptado de INEP (2022b) mostra a descrição dos níveis de desempenho em matemática nas séries finais do ensino fundamental.

Quadro 8: Descrição dos níveis de desempenho em Matemática para anos finais do ensino fundamental

Nível	Descrição das Habilidades Envolvidas
Nível 1 Desempenho maior ou igual a 200 e menor que 225	Os estudantes provavelmente são capazes de: Números e operações; álgebra e funções – Reconhecer o maior ou o menor número em uma coleção de números racionais, representados na forma decimal. Tratamento de informações – Interpretar dados apresentados em tabela e gráfico de colunas.
Nível 2 Desempenho maior ou igual a 225 e menor que 250	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Números e operações; álgebra e funções – Reconhecer a fração que corresponde à relação parte-todo entre uma figura e suas partes hachuradas. Associar um número racional que representa uma quantia monetária, escrito por extenso, à sua representação decimal. Determinar uma fração irredutível, equivalente a uma fração dada, a partir da simplificação por três. Tratamento de informações – Interpretar dados apresentados em um gráfico de linhas simples. Associar dados apresentados em gráfico de colunas a uma tabela.

Nível 3	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Reconhecer o ângulo de giro que representa a mudança de direção na movimentação de pessoas/objetos; reconhecer a planificação de um sólido simples, dado por um desenho em perspectiva. Localizar um objeto em representação gráfica do tipo planta baixa, utilizando dois critérios: estar mais longe de um referencial e mais perto de outro. Números e operações; álgebra e funções – Determinar uma fração irredutível, equivalente a uma fração dada, a partir da simplificação por sete; determinar a soma, a diferença, o produto ou o quociente de números inteiros em situações-problema. Localizar o valor que representa um número inteiro positivo associado a um ponto indicado em uma reta numérica. Resolver problemas envolvendo grandezas diretamente proporcionais, representadas por números inteiros. Tratamento de informações – Associar dados apresentados em tabela a gráfico de setores. Analisar dados dispostos em uma tabela simples. Analisar dados apresentados em um gráfico de linhas com mais de uma grandeza representada.
Nível 4	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Localizar um ponto em um plano cartesiano, com o apoio de malha quadriculada, a partir de suas coordenadas. Reconhecer as coordenadas de um ponto dado em um plano cartesiano, com o apoio de malha quadriculada. Interpretar a movimentação de um objeto utilizando referencial diferente do seu. Grandezas e medidas – Converter unidades de medidas de comprimento, de metros para centímetros, na resolução de situação-problema. Reconhecer que a medida do perímetro de um retângulo, em uma malha quadriculada, dobra ou se reduz à metade quando os lados dobram ou são reduzidos à metade. Números e operações; álgebra e funções – Determinar a soma de números racionais em contextos de sistema monetário. Determinar o valor numérico de uma expressão algébrica de 1º grau envolvendo números naturais, em situação-problema. Localizar números inteiros negativos na reta numérica. Localizar números racionais em sua representação decimal. Tratamento de informações – Analisar dados dispostos em uma tabela de dupla entrada.
Nível 5	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Reconhecer que o ângulo não se altera em figuras obtidas por ampliação/redução. Localizar dois ou mais pontos em um sistema de coordenadas. Grandezas e medidas – Determinar o perímetro de uma região retangular, com o apoio de figura, na resolução de uma situação-problema. Determinar o volume mediante contagem de blocos. Números e operações; álgebra e funções – Associar uma fração com denominador dez à sua representação decimal. Associar uma situação-problema à sua linguagem algébrica, por meio de equações do 1º grau ou sistemas lineares. Determinar, em situação-problema, a adição e multiplicação entre números racionais, envolvendo divisão por números inteiros. Determinar a porcentagem envolvendo números inteiros. Resolver problema envolvendo grandezas diretamente proporcionais, representadas por números racionais na forma decimal.
Nível 6	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Reconhecer a medida do ângulo determinado entre dois deslocamentos, descritos por meio de orientações dadas por pontos cardeais. Reconhecer as coordenadas de pontos representados no primeiro quadrante de um plano cartesiano. Reconhecer a relação entre as medidas de raio e diâmetro de uma circunferência, com o apoio de figura. Reconhecer a corda de uma circunferência, as faces opostas de um cubo, a partir de uma de suas planificações. Comparar as medidas dos lados de um triângulo com base nas medidas de seus respectivos ângulos opostos. Resolver problema utilizando o Teorema de Pitágoras no cálculo da medida da hipotenusa, dadas as medidas dos catetos. Grandezas e medidas – Converter unidades de medida de massa, de quilograma para grama, na resolução de situação-problema. Resolver problema fazendo uso de semelhança de triângulos. Números e operações; álgebra e funções – Reconhecer frações equivalentes. Associar um número racional, escrito por extenso, à sua representação decimal, e vice-versa. Estimar o valor da raiz

			quadrada de um número inteiro aproximando-o de um número racional em sua representação decimal. Resolver problema envolvendo grandezas diretamente proporcionais, com constante de proporcionalidade não inteira. Determinar o valor numérico de uma expressão algébrica que contenha parênteses, envolvendo números naturais. Determinar um valor monetário obtido por meio de um desconto ou um acréscimo porcentual. Determinar o valor de uma expressão numérica, com números irracionais, fazendo uso de uma aproximação racional fornecida. Tratamento de informações – Resolver problemas que requerem a comparação de dois gráficos de colunas.
Nível	7	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Reconhecer ângulos agudos, retos ou obtusos de acordo com sua medida em graus. Reconhecer as coordenadas de pontos representados num plano cartesiano localizados em quadrantes que não sejam o primeiro. Determinar a posição final de um objeto, após a realização de rotações em torno de um ponto, de diferentes ângulos, em sentido horário e anti-horário. Resolver problemas envolvendo ângulos, inclusive utilizando a Lei Angular de Tales sobre a soma dos ângulos internos de um triângulo. Resolver problemas envolvendo as propriedades de ângulos internos e externos de triângulos e quadriláteros, com ou sem justaposição ou sobreposição de figuras. Resolver problema utilizando o Teorema de Pitágoras no cálculo da medida de um dos catetos, dadas as medidas da hipotenusa e de um de seus catetos. Grandezas e medidas – Determinar o perímetro de uma região retangular, obtida pela justaposição de dois retângulos, descritos sem o apoio de figuras. Determinar a área de um retângulo em situações-problema. Determinar a área de regiões poligonais desenhadas em malhas quadriculadas. Determinar o volume de um cubo ou de um paralelepípedo retângulo, sem o apoio de figura. Converter unidades de medida de volume, de m³ para litro, em situações-problema. Reconhecer a relação entre as áreas de figuras semelhantes. Números e operações; álgebra e funções – Determinar o quociente entre números racionais, representados na forma decimal ou fracionária, em situações-problema. Determinar a soma de números racionais dados na forma fracionária e com denominadores diferentes. Determinar o valor numérico de uma expressão algébrica de 2º grau, com coeficientes naturais, envolvendo números inteiros. Determinar o valor de uma expressão numérica envolvendo adição, subtração, multiplicação e/ou potenciação entre números inteiros. Determinar o valor de uma expressão numérica com números inteiros positivos e negativos; determinar o valor de uma expressão numérica com números racionais. Comparar números racionais com diferentes números de casas decimais, usando arredondamento. Localizar na reta numérica um número racional, representado na forma de uma fração imprópria. Associar uma fração à sua representação na forma decimal. Associar uma situação-problema à sua linguagem algébrica, por meio de inequações do 1º grau. Associar a representação gráfica de duas retas no plano cartesiano a um sistema de duas equações lineares e vice-versa. Resolver problemas envolvendo equação de 2º grau. Tratamento de informações – Determinar a média aritmética de um conjunto de valores. Estimar quantidades em gráficos de setores. Analisar dados dispostos em uma tabela de três ou mais entradas. Interpretar dados fornecidos em gráficos envolvendo regiões do plano cartesiano. Interpretar gráficos de linhas com duas sequências de valores.
Nível	8	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Resolver problemas utilizando as propriedades das cevianas (altura, mediana e bissetriz) de um triângulo isósceles, com o apoio de figura. Grandezas e medidas – Converter unidades de medida de capacidade, de mililitro para litro, em situações-problema. Reconhecer que a área de um retângulo quadruplica quando seus lados dobram. Determinar a área de figuras simples (triângulo, paralelogramo, trapézio), inclusive utilizando composição/ decomposição. Números e operações; álgebra e funções – Determinar o valor numérico de uma expressão algébrica de 1º grau, com coeficientes racionais, representados na forma decimal. Determinar o valor de uma expressão numérica envolvendo adição, subtração e potenciação entre números racionais, representados na forma decimal. Resolver problemas envolvendo grandezas inversamente proporcionais.

Nível 9 Desempenho maior ou igual a 400	–	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: Espaço e forma – Resolver problemas utilizando a soma das medidas dos ângulos internos de um polígono. Números e operações; álgebra e funções – Reconhecer a expressão algébrica que expressa uma regularidade existente em uma sequência de números ou de figuras geométricas.
--	---	---

Fonte: Adaptado de INEP (2022b)

2.1.5.3 Níveis do Saeb para o ensino médio

2.1.5.3.1 Língua Portuguesa

Os alunos do ensino médio podem se encaixar em 8 níveis de desempenho segundo o Saeb. Para valores inferiores a 225 pontos, não é possível inferir quais habilidades foram desenvolvidas pelo aluno. O Fonte: Adaptado de INEP (2022b) mostra as habilidades esperadas por nível em Língua Portuguesa para os alunos de ensino médio.

Quadro 9: Habilidades esperadas para Língua Portuguesa no ensino médio

Nível	Descrição das Habilidades Envolvidas
Nível 1 – Desempenho maior ou igual a 225 e menor que 250	Os estudantes provavelmente são capazes de: identificar elementos da narrativa em história em quadrinhos; reconhecer a finalidade de recurso gráfico em artigos; reconhecer a relação de causa e consequência em lendas; inferir o sentido de palavra em letras de música e reportagens.
Nível 2 – Desempenho maior ou igual a 250 e menor que 275	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: reconhecer a ideia comum entre textos de gêneros diferentes e a ironia em tirinhas; reconhecer relações de sentido estabelecidas por conjunções ou locuções conjuntivas em letras de música e crônicas; reconhecer o uso de expressões características da linguagem (científica, profissional etc.) e a relação entre pronome e seu referente em artigos e reportagens; inferir o efeito de sentido da linguagem verbal e não verbal em notícias e charges.
Nível 3 – Desempenho maior ou igual a 275 e menor que 300	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informação explícita em artigos de opinião; identificar a finalidade de relatórios científicos; reconhecer relações de sentido marcadas por conjunções, relação de causa e consequência e relação entre o pronome e seu referente em fragmentos de romances; reconhecer o tema de uma crônica; reconhecer variantes linguísticas em artigos; reconhecer o sentido e o efeito de sentido produzido pelo uso de recursos morfosintáticos em contos, artigos e crônicas; reconhecer opiniões divergentes sobre o mesmo tema em diferentes textos; inferir informação, sentido e efeito de sentido produzido por expressão em reportagens e tirinhas.
Nível 4 – Desempenho maior ou igual a 300 e menor que 325	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informações explícitas em infográficos, reportagens, crônicas e artigos; identificar o argumento em contos; identificar a finalidade e a informação principal em notícias; reconhecer a relação entre os pronomes e seus referentes em contos; reconhecer elementos da narrativa em contos; reconhecer variantes linguísticas em contos, notícias e reportagens; reconhecer o efeito de sentido produzido pelo uso de recursos morfosintáticos em poemas; reconhecer ideia comum e opiniões divergentes sobre o mesmo tema na comparação entre diferentes textos; reconhecer ironia e efeito de humor em crônicas e entrevistas; reconhecer a relação de causa e consequência em piadas e fragmentos de romance; comparar poemas que abordem o mesmo tema; diferenciar fato de opinião em contos, artigos e reportagens; diferenciar tese de argumentos em artigos, entrevistas e crônicas; inferir informação, sentido de expressão e efeito de sentido decorrente do uso de recursos morfosintáticos em crônicas; inferir o

	sentido decorrente do uso de recursos gráficos em poemas; inferir o efeito de sentido da linguagem verbal e não verbal e o efeito de humor em tirinhas.
Nível 5 – Desempenho maior ou igual a 325 e menor que 350	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informação explícita em resenhas; identificar a informação principal em reportagens; identificar elementos da narrativa e a relação entre argumento e ideia central em crônicas; reconhecer a finalidade de propagandas; reconhecer variantes linguísticas e o efeito de sentido de recursos gráficos em crônicas e artigos; reconhecer a relação de causa e consequência e relações de sentido marcadas por conjunções em reportagens, artigos e ensaios; reconhecer o tema em poemas; diferenciar fato de opinião em resenhas; inferir o sentido de palavras e expressões em piadas e letras de música; inferir informação em artigos; inferir o sentido de expressão em fragmentos de romances.
Nível 6 – Desempenho maior ou igual a 350 e menor que 375	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: reconhecer efeitos estilísticos em poemas; reconhecer ironia e efeitos de sentido decorrentes da repetição de palavras em sinopses; reconhecer opiniões distintas sobre o mesmo tema, na comparação entre diferentes textos; reconhecer finalidade e traços de humor em reportagens; reconhecer o efeito de sentido do humor em tirinhas; reconhecer o tema em contos e fragmentos de romances; reconhecer relação de sentido marcada por conjunção em crônicas; inferir informação e tema em reportagens, poemas, histórias em quadrinhos e tirinhas; inferir o sentido e o efeito de sentido de palavras ou de expressão em poemas, crônicas e fragmentos de romances.
Nível 7 – Desempenho maior ou igual a 375 e menor que 400	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: identificar a ideia central e o argumento em apresentações de livros, reportagens, editoriais e crônicas; identificar elementos da narrativa em crônicas, contos e fragmentos de romances; identificar ironia e tema em poemas e artigos; reconhecer relações de sentido marcadas por conjunção em artigos, reportagens e fragmentos de romances; reconhecer a relação de causa e consequência em reportagens e fragmentos de romances; reconhecer o efeito de sentido de recursos gráficos em artigos; reconhecer variantes linguísticas em letras de música e piadas; reconhecer a finalidade de reportagens, resenhas e artigos; inferir efeito de humor e ironia em tirinhas e charges.
Nível 8 – Desempenho maior ou igual a 400 e menor que 425	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: reconhecer o efeito de sentido resultante do uso de recursos morfossintáticos em artigos e letras de música.

Fonte: Adaptado de INEP (2022b)

2.1.5.3.2 Matemática

Para os alunos de matemática do ensino médio o Saeb, delimitou 10 níveis de desempenho. Para alunos que obtiveram pontuação inferior a 225, nada se pode concluir a respeito das habilidades que foram desenvolvidas. O Fonte: Adaptado de INEP (2022b).

mostra as habilidades de matemática por nível para alunos do ensino médio.

Quadro 10: Habilidades em matemática para alunos do ensino médio

Nível	Descrição das Habilidades Envolvidas
Nível 1 – Desempenho maior ou igual a 225 e menor que 250	Espaço e forma – Não existem itens-âncora para esse nível. Grandezas e medidas – Não existem itens-âncora para esse nível. Números e operações; álgebra e funções – Não existem itens-âncora para esse nível. Tratamento de informações – O estudante pode ser capaz de associar uma tabela de até duas entradas a informações apresentadas textualmente ou em um gráfico de barras ou de linhas.

Nível 2 Desempenho maior ou igual a 250 e menor que 275	–	Espaço e forma – O estudante pode ser capaz de reconhecer as coordenadas de pontos representados em um plano cartesiano e localizados no primeiro quadrante. Grandezas e medidas – Não existem itens-âncora para esse nível. Números e operações; álgebra e funções – O estudante pode ser capaz de reconhecer os zeros de uma função dada graficamente. Também é bem provável que os alunos determinem: o valor de uma função afim, dada sua lei de formação; um resultado utilizando o conceito de progressão aritmética. Tratamento de informações – O estudante pode ser capaz de associar um gráfico de setores a dados percentuais apresentados textualmente ou em uma tabela.
Nível 3 Desempenho maior ou igual a 275 e menor que 300	–	Espaço e forma – Não existem itens-âncora para esse nível. Grandezas e medidas – Não existem itens-âncora para esse nível. Números e operações; álgebra e funções – O estudante pode ser capaz de reconhecer: o valor máximo de uma função quadrática representada graficamente; em um gráfico, o intervalo no qual a função assume valor máximo. Também pode ser capaz de determinar: por meio de proporcionalidade, o gráfico de setores que representa uma situação com dados fornecidos textualmente; o quarto valor em uma relação de proporcionalidade direta com base em três valores fornecidos em uma situação do cotidiano; um valor reajustado de uma quantia a partir de seu valor inicial e do percentual de reajuste. Além disso, é provável que resolva problemas utilizando operações fundamentais com números naturais. Tratamento de informações – Não existem itens-âncora para esse nível.
Nível 4 Desempenho maior ou igual a 300 e menor que 325	–	Espaço e forma – Não existem itens-âncora para esse nível. Grandezas e medidas – O estudante pode ser capaz de resolver problemas envolvendo área de uma região composta por retângulos, com base em medidas fornecidas em texto e figura. Números e operações; álgebra e funções – O estudante pode ser capaz de reconhecer o gráfico de função considerando valores fornecidos em um texto. Além disso, pode ser capaz de determinar: a lei de formação de uma função linear a partir de dados fornecidos em uma tabela; a solução de um sistema de duas equações lineares; um termo de progressão aritmética, dada sua forma geral; a probabilidade da ocorrência de um evento simples. Também é provável que resolva: problemas utilizando proporcionalidade direta ou inversa, cujos valores devem ser obtidos a partir de operações simples; problemas de contagem usando princípio multiplicativo. Tratamento de informações – Não existem itens-âncora para esse nível.
Nível 5 Desempenho maior ou igual a 325 e menor que 350	–	Espaço e forma – Não existem itens-âncora para esse nível. Grandezas e medidas – O estudante pode ser capaz de determinar medidas de segmentos por meio da semelhança entre dois polígonos. Números e operações; álgebra e funções – O estudante pode ser capaz de determinar: o valor de variável dependente ou independente de uma função exponencial dada; o percentual que representa um valor em relação a outro; o valor de uma expressão algébrica; a solução de um sistema de três equações sendo uma com uma incógnita, outra com duas e a terceira com três incógnitas. Também é provável que seja capaz de resolver problema envolvendo: divisão proporcional do lucro em relação a dois investimentos iniciais diferentes; operações, além das fundamentais, com números naturais; relação linear entre duas variáveis para a determinação de uma delas; probabilidade de união de eventos. Além disso, é provável que os alunos sejam capazes de avaliar o comportamento de uma função, representada graficamente, quanto ao seu crescimento. Tratamento de informações – Não existem itens-âncora para esse nível.
Nível 6 Desempenho maior ou igual a 350 e menor que 375	–	Espaço e forma – O estudante pode ser capaz de reconhecer as coordenadas de pontos representados em um plano cartesiano e localizados em quadrantes que não sejam o primeiro. É provável também que consiga associar um sólido geométrico simples a uma planificação usual dada. Além disso, há uma grande probabilidade de resolver problemas envolvendo o Teorema de Pitágoras, para calcular a medida da hipotenusa de um triângulo pitagórico, a partir de informações apresentadas textualmente e em uma figura. Grandezas e medidas – O estudante pode ser capaz de determinar: a razão de semelhança entre as imagens de um mesmo objeto em escalas diferentes; o volume de um paralelepípedo retângulo, dada sua representação espacial. Números e operações; álgebra e funções – O estudante pode ser capaz de determinar os zeros de uma função quadrática, a partir de sua expressão algébrica. Além disso, é provável que resolva problemas de porcentagem envolvendo números racionais não inteiros. Tratamento de informações – Não existem itens-âncora para esse nível.
Nível 7 Desempenho maior ou igual a	–	Espaço e forma – O estudante pode ser capaz de determinar: a medida de um dos lados de um triângulo retângulo, por meio de razões trigonométricas, sendo fornecidas ou não as fórmulas; com o uso do teorema de Pitágoras, a medida de um dos catetos de um triângulo

375 e menor que 400		retângulo não pitagórico. Grandezas e medidas – O estudante pode ser capaz de determinar a área de um polígono não convexo composto por retângulos e triângulos, a partir de informações fornecidas na figura. Além disso, é provável que consiga resolver problemas: por meio de semelhança de triângulos sem apoio de figura; envolvendo perímetros de triângulos equiláteros que compõem uma figura. Números e operações; álgebra e funções – O estudante pode ser capaz de reconhecer gráfico de função a partir de informações sobre sua variação descritas em um texto; os zeros de uma função quadrática em sua forma fatorada; gráfico de função afim a partir de sua representação algébrica; a equação de uma reta a partir de dois de seus pontos; as raízes de um polinômio apresentado na sua forma fatorada. Além disso, é provável também que os alunos sejam capazes de determinar os pontos de máximo ou de mínimo a partir do gráfico de uma função; o valor de uma expressão algébrica envolvendo módulo; o ponto de interseção de duas retas; a expressão algébrica que relaciona duas variáveis com valores dados em tabela ou gráfico; a maior raiz de um polinômio de 2º grau. Também é provável que os alunos sejam capazes de resolver problemas: para obter valor de variável dependente ou independente de uma função exponencial dada; envolvendo uma equação de 1º grau que requeira manipulação algébrica; envolvendo um sistema linear, dadas duas equações a duas incógnitas; usando permutação; utilizando probabilidade, envolvendo eventos independentes. Tratamento de informações – Não existem itens-âncora para esse nível.
Nível 8 Desempenho maior ou igual a 400 e menor que 425	–	Espaço e forma – O estudante pode ser capaz de reconhecer a proporcionalidade dos elementos lineares de figuras semelhantes. Também é provável que seja capaz de determinar: uma das medidas de uma figura tridimensional, utilizando o Teorema de Pitágoras; a equação de uma circunferência, dados o centro e o raio; a quantidade de faces, vértices e arestas de um poliedro por meio da relação de Euler. É provável também que os alunos sejam capazes de resolver problema envolvendo razões trigonométricas no triângulo retângulo, com apoio de figura. Podem também ser capazes de associar um prisma a uma planificação usual dada. Grandezas e medidas – O estudante pode ser capaz de determinar a área da superfície de uma pirâmide regular; o volume de um paralelepípedo, dadas suas dimensões em unidades diferentes; o volume de cilindros. Números e operações; álgebra e funções – O estudante pode ser capaz de reconhecer: o gráfico de uma função trigonométrica na forma $y=\sin(x)$; um sistema de equações associado a uma matriz. Também é provável que seja capaz de determinar: a expressão algébrica associada a um dos trechos do gráfico de uma função definida por partes; o valor máximo de uma função quadrática a partir de sua expressão algébrica e das expressões que determinam as coordenadas do vértice; a distância entre dois pontos no plano cartesiano. É provável também que os alunos sejam capazes de resolver problema: usando arranjo; envolvendo a resolução de uma equação do 2º grau cujos dados são seus coeficientes. Além disso, existe uma grande probabilidade de que sejam capazes de interpretar o significado dos coeficientes da equação de uma reta, a partir de sua forma reduzida. Tratamento de informações – Não existem itens-âncora para esse nível.
Nível 9 Desempenho maior ou igual a 425 e menor que 450	–	Espaço e forma – O estudante pode ser capaz de reconhecer a equação que representa uma circunferência, entre diversas equações dadas. Também é provável que seja capaz de determinar o centro e o raio de uma circunferência a partir de sua equação geral. É provável também que os alunos sejam capazes de resolver problemas envolvendo relações métricas em um triângulo retângulo que é parte de uma figura plana dada. Grandezas e medidas – O estudante pode ser capaz de determinar o volume de pirâmides regulares. É provável também que os alunos sejam capazes de resolver problema envolvendo: áreas de círculos e polígonos; semelhança de triângulos com apoio de figura na qual os dois triângulos apresentam ângulos opostos pelos vértices; envolvendo cálculo de volume de cilindro. Números e operações; álgebra e funções – O estudante pode ser capaz de reconhecer o gráfico de uma função exponencial do tipo $f(x)=10^x+1$; o gráfico de uma função logarítmica, dada a expressão algébrica da sua função inversa e seu gráfico. Também é provável que seja capaz de determinar a expressão algébrica correspondente a uma função exponencial, com base em dados fornecidos em texto ou gráfico; a inversa de uma função exponencial dada, representativa de uma situação do cotidiano; inclinação ou coeficiente angular de retas a partir de suas equações; um polinômio na forma fatorada, dadas as suas raízes. Tratamento de informações – Não existem itens-âncora para esse nível.

Nível 10	–	Espaço e forma – Não existem itens-âncora para esse nível. Grandezas e medidas – Não existem itens-âncora para esse nível. Números e operações; álgebra e funções – O estudante pode ser capaz de determinar a solução de um sistema de três equações lineares, a três incógnitas, apresentado na forma matricial escalonada. Tratamento de informações
Desempenho maior ou igual a 450		– Não existem itens-âncora para esse nível.

Fonte: Adaptado de INEP (2022b).

Neste trabalho, o indicador utilizado para analisar o aprendizado do aluno foi o valor percentual de alunos que estiveram em nível igual ou superior a 5 em língua portuguesa, uma vez que nos anos iniciais do ensino fundamental este aluno consegue diferenciar opinião e informação dentro de uma reportagem

3 METODOLOGIA

Nessa seção são apresentadas as metodologias utilizadas para a análise conduzida.

Após apresentar os principais conceitos e categorias sobre os quais o sistema educacional brasileiro é estruturado, a próxima etapa do estudo consiste em introduzir o referencial teórico que permitiria agrupar um conjunto de instituições de ensino ou unidades comparáveis, como municípios, com características semelhantes.

O desenvolvimento deste trabalho envolve a aplicação de métodos de clusterização para identificar as semelhanças entre escolas que apresentam características semelhantes dado que, segundo James *et al.*, 2013, o processo de *Clusterizar* ou agrupar envolve um conjunto de técnicas que permite formar subgrupos, ou *clusters*, a partir de uma base de dados. Formando grupos que possuem membros semelhantes entre si, mas distintos quando comparados a membros de outros grupos. Sendo assim, é necessário inicialmente definir as características que tornam duas observações semelhantes ou diferentes entre si.

De acordo com James *et al.* (2013), a formação de *clusters* pode ocorrer, em geral, de duas maneiras: (i) agrupamento das observações a partir das características de seus integrantes que permitem identificar subgrupos dentre as observações; ou (ii) agrupamento das características dos indivíduos dos grupos, tomando como base as observações, para então descobrir subgrupos dentre as características. Neste estudo, assim como proposto por James *et al.* (2013), será aplicada a clusterização de observações a partir de suas características.

Dentro do escopo das técnicas de aprendizado estatístico ou aprendizado de máquina (*statistical learning* ou *machine learning*), os métodos de clusterização são aqueles que correspondem ao aprendizado não supervisionado: são problemas em que temos dados “não rotulados” (PROVOST; FAWCETT, 2013). Ou seja, nesse tipo de problema, para cada observação $i = 1, \dots, n$, temos um vetor de medidas x_i , mas não temos uma resposta associada y_i . São casos nos quais não é possível aplicar um modelo de regressão linear (JAMES *et al.*, 2013). Para Provost e Fawcett (2013), nesse grupo de técnicas estão os algoritmos que buscam identificar padrões subjacente entre os dados, muitas vezes não conhecidos *a priori*.

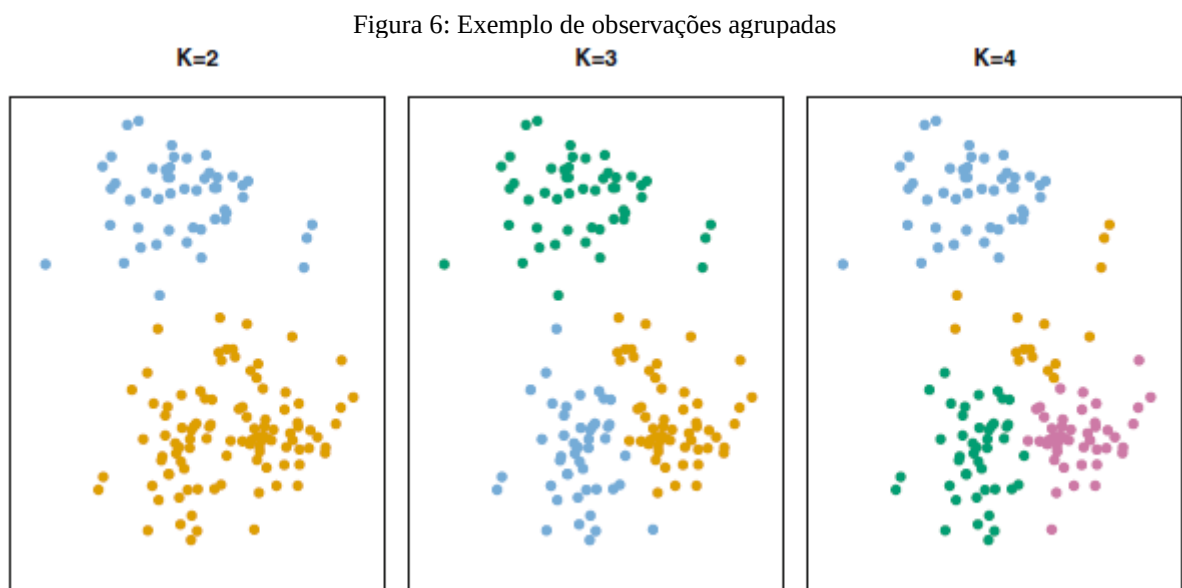
Há diversas técnicas para clusterizar grupos a partir de suas características. Nesse trabalho, exploramos duas das principais: “K-Means” e “Clusterização por hierarquia”. Ao final da análise, após aplicar ambas as técnicas, teremos dois conjuntos de *clusters* distintos (elementos distintos). A partir, tenta-se identificar uma convergência dentre eles e entender

quais as melhores sugestões para o desenvolvimento dos alunos das instituições que integram cada *cluster*.

3.1 MÉTODO DE CLUSTERIZAÇÃO K-MEANS

A metodologia de clusterização “K-means” envolve dividir um banco de dados em K clusters distintos e sem sobreposição. Para isso, o primeiro passo consiste em definir o número de *clusters* desejado, e então o algoritmo do método vai designar cada observação para um dos clusters. (JAMES *et al.*, 2013). De forma simplificada, Provost e Fawcett (2013) descrevem esse método como um algoritmo que minimiza a distância euclidiana entre pontos em um plano.

A **Error! Reference source not found.** ilustra como cada observação, representada por um ponto, é designada a um *cluster*, que é representado por uma cor distinta das demais, de forma que o número de *clusters* se dá pelo valor do número K .



Fonte: Adaptado de James *et al.* (2013), p. 387

James *et al.* (2013) descrevem a notação a partir da qual o método é definido: $C_1, \dots, C_k$ são conjuntos que contêm os índices das observações dentro de cada *cluster*. Esses conjuntos devem satisfazer duas propriedades:

1. $C_1 \cup C_2 \cup \dots \cup C_K = \{1, \dots, n\}$: Cada n observação deve pertencer a pelo menos um dos K *clusters*.
2. $C_1 \cap C_{k'} = \emptyset$ para todo $k \neq k'$: Os *clusters* não se sobrepõem, ou seja, nenhuma observação pertence a mais de um *cluster*.

Dando continuidade à notação fornecida pelos autores, teríamos que a i -ésima observação pertence ao k -ésimo *cluster*, $i \in C_k$. Uma boa clusterização feita por meio do método K-means define a menor variação possível entre as observações pertencentes a um *cluster*. Para James *et al.* (2013), a variação dentro de um *cluster* C_k é uma medida pela função $W(k)$, que mensura a variação entre as observações de um mesmo *cluster*. A menor variação entre as observações do *cluster* é dada pela **Error! Reference source not found.**, que minimiza a função $W(k)$.

Equação 1

$$\min_{C_1, \dots, C_k} \left\{ \sum_{k=1}^K W(C_k) \right\}$$

A resolução da equação 1 passa pela definição da “Variação dentro de um *cluster*” $W(C_k)$, sendo que neste projeto será utilizado um dos métodos mais comuns para a resolução desse problema, envolvendo o *quadrado da distância Euclidiana*, que é definida pela equação 2.

Equação 2

$$W(C_k) = \frac{1}{|C_k|} \sum_{i, i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2$$

Em que $|C_k|$, representa o número de observações dentro do k -ésimo *cluster*.

Unindo a equação 1 e a equação 2, obtemos o problema que define a clusterização K-means pela equação 3:

Equação 3

$$\text{minimize}_{C_1, \dots, C_k} \left\{ \sum_{k=1}^K \frac{1}{|C_k|} \sum_{i, i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2 \right\}$$

O processo de resolução da função objetivo representada pela fórmula XXXX, passa pela elaboração de um algoritmo que divide as observações em K-clusters de forma a minimizar

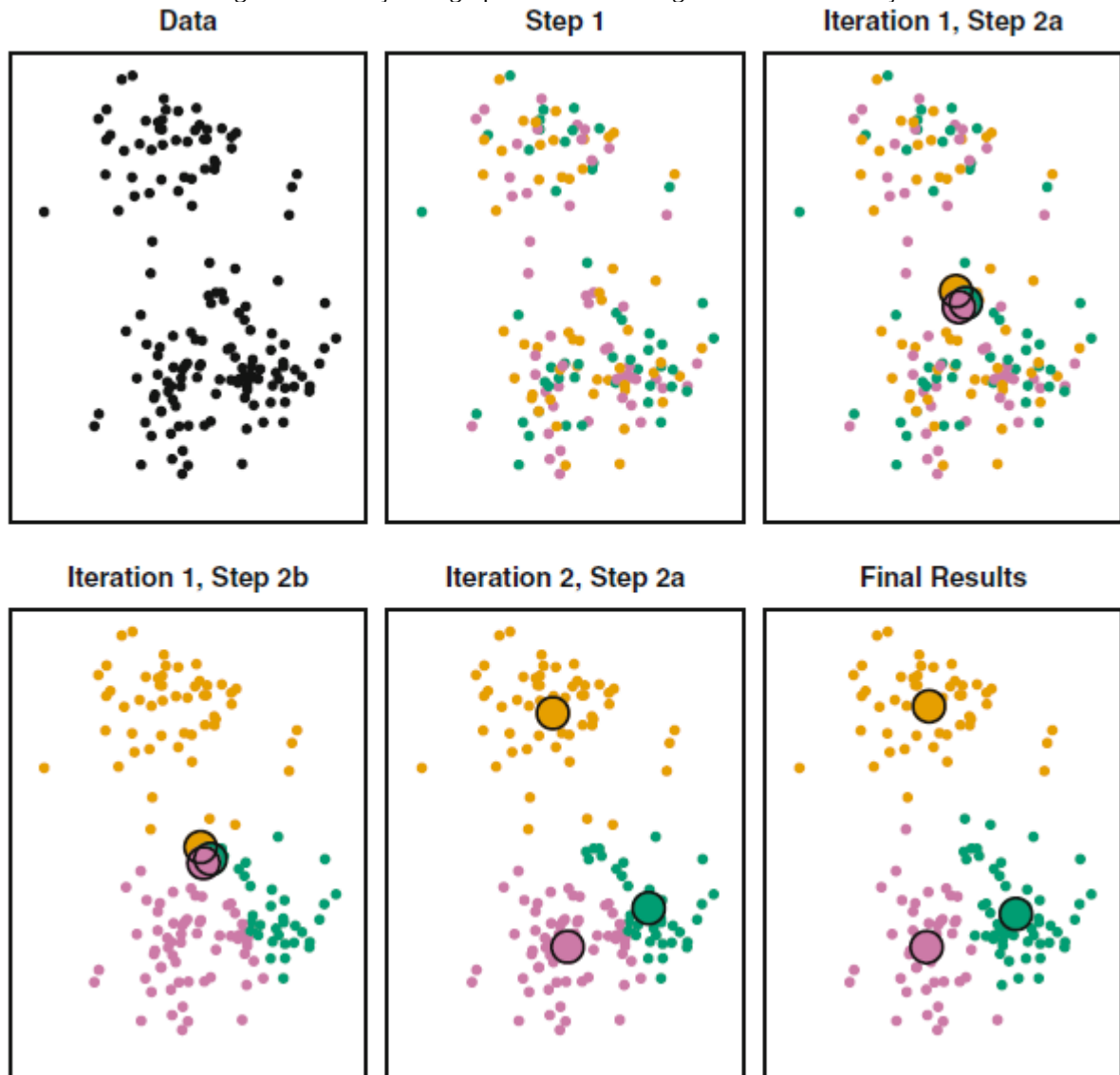
o resultado da função objetivo. O algoritmo 1 sintetizado por James *et al.* (2013) busca alcançar um valor ótimo local, seguindo dois passos:

Algoritmo 1:

1. Designar para cada observação um valor aleatório entre 1 e K, formando um *cluster* inicial para cada observação
2. Iterar até que as observações deixem de ser designadas a diferentes *clusters*, seguindo as etapas:
 - a. Para cada cluster K, calcula-se o *centróide* do *cluster*. O centroide do K-ésimo cluster é o vetor da média da característica p para as observações dentro do K-ésimo *cluster*
 - b. Designar cada observação para o *cluster* cujo centroide está mais próximo, utilizando a definição Euclidiana de distância.

A **Error! Reference source not found.** ilustra a formação de cada *cluster* a partir do Algoritmo 2 apresentado, em que os passo 1 representa a designação aleatória de cada observação a um *cluster*, e o passo 2 mostra o cálculo do centroide de cada *cluster* e a designação das observações a partir das iterações que são realizadas.

Figura 7: Formação de grupos conforme o Algoritmo de clusterização.



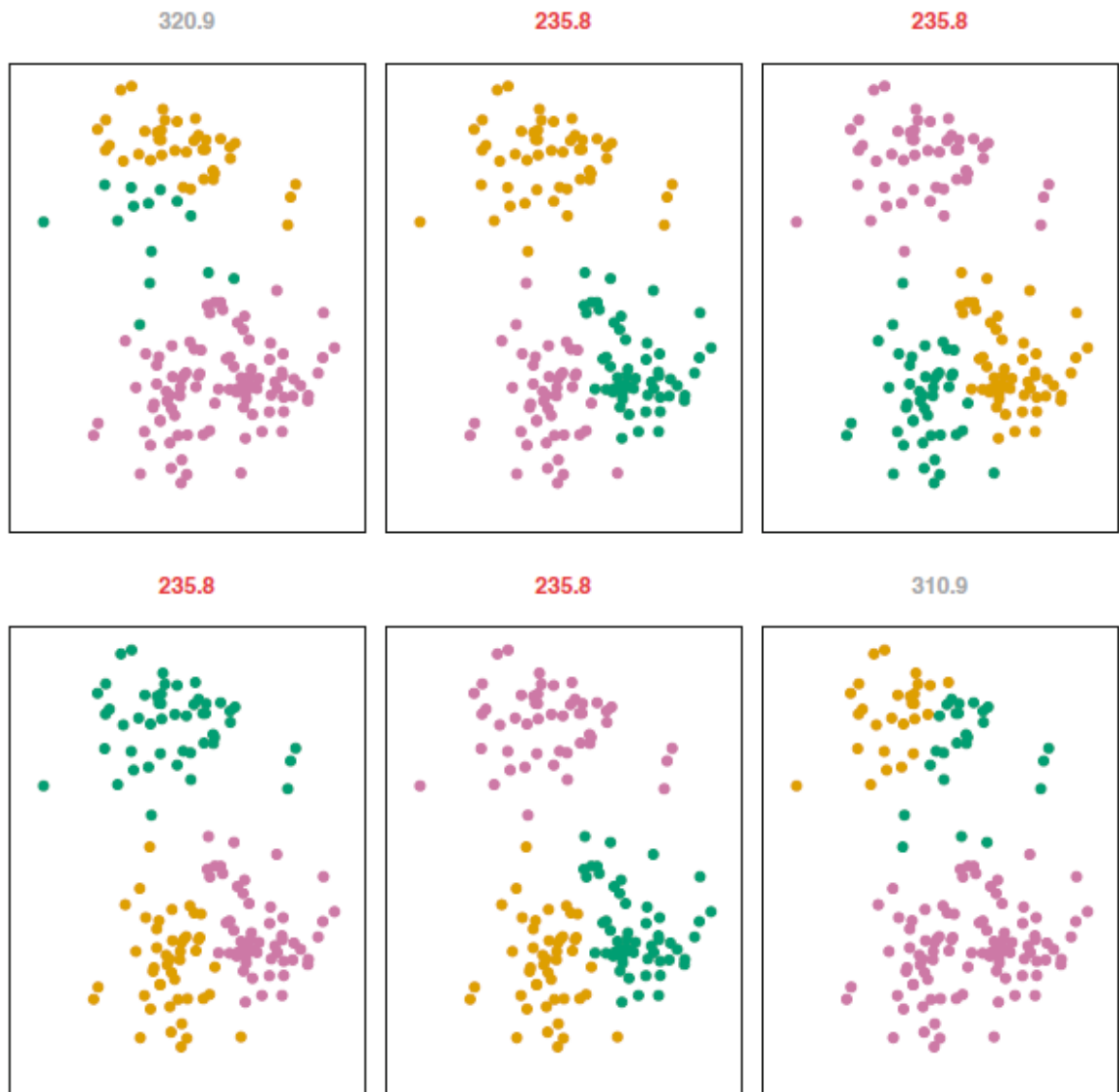
Fonte: Adaptado de James *et al.* (2013, p. 389)

O algoritmo 1 garante a redução do valor da função objetivo a cada passo. Outra característica do algoritmo 1 é que ele encontra um mínimo local e não global, como consequência, o resultado depende da designação aleatória realizada no primeiro passo do algoritmo. Sendo importante realizar o algoritmo diversas vezes com configurações iniciais distintas até que o valor da função objetivo seja o menor dentre as execuções realizadas, alcançando assim a melhor solução possível.

A variação dos valores da função objetivo de acordo com a designação aleatória de K é a apresentada na figura 8, em que quatro observações apresentam valores da função objetivo

idênticos e duas delas apresentam valores superiores, estas representam resultados menos precisos que as demais.

Figura 8: Progresso do algoritmo K-means em 6 iterações



Fonte: Adaptado de James *et al.* (2013, p. 390)

3.1.1 Aplicação da técnica K-means computacionalmente

A técnica K-means pode ser aplicada segundo alguns algoritmos diferentes, sendo os três algoritmos mais difundidos: o algoritmo de Forgy/ Lloyd, MacQueen, e Hartigan & Wong. Para executar esses algoritmos primeiro é necessário definir o número de *clusters* a serem

criados, e para sua aplicação é recomendada a padronização dos dados caso os componentes não estejam na mesma escala. (MORISSETTE *et al.*, 2013).

Morissette *et al.* (2013) definem o algoritmo de MacQueen como mais eficiente que o de Forgy/ Lloyd, pois ele atualiza a posição dos centróides de forma mais frequente. Em comparação com o algoritmo de Hartigan & Wong, o algoritmo de MacQueen foi mais eficiente computacionalmente durante a aplicação dos dados.

Neste trabalho, o algoritmo selecionado para a aplicação da técnica K-means foi o de MacQueen pelos motivos listados anteriormente, e sua sequência está explícita no Algoritmo 2, conforme definido por Morissette *et al.* (2013) como:

Algoritmo 2:

1. Escolher o número de *clusters*
2. Escolher a unidade métrica a ser utilizada
3. Utilizar o método de escolha dos centroides iniciais
4. Designar os Centroides iniciais
5. Enquanto os centroides deixarem de ser atualizados ou até atingirem um valor de iterações máximo
 - a. Para o *i*-ésimo caso:
 - i. Designar o caso *i* para o *cluster* mais próximo de acordo com a unidade métrica definida
 - ii. Atualizar os centroides para os *clusters* afetados
 - b. Atualizar todos os centroides

3.2 IDENTIFICAÇÃO DE *OUTLIERS*

As metodologias de análise de *cluster* utilizadas nesse trabalho são determinísticas, isto é, não fazem inferência sobre a existência de erros aleatórios nos dados utilizados, fazendo uso de todo o conjunto amostral para a determinação do resultado mensurado. Sendo assim, foi necessário excluir possíveis pontos espúrios das bases de dados analisadas. Se não removidos, esses pontos, também denominados *outliers*, poderiam definir um *cluster* individualmente, prejudicando a análise.

3.2.1 Diagrama *Boxplot*

O diagrama de caixa, ou *boxplot* em inglês, é uma forma de sintetizar diversas informações presentes em uma base de dados, sendo as principais: mediana, primeiro quartil, terceiro quartil, limite superior, e limite inferior (MANN, PREM S. *et al.*, 2010).

Costa Neto *et al.* (2002) definem a mediana para um conjunto de n valores ordenados, sendo n ímpar como igual ao valor de ordem $(n + 1)/2$ desse conjunto. Se n for par, a mediana poderia ser definida como qualquer valor situado entre o de ordem $n/2$ e o de ordem $(n/2) + 1$.

Segundo Devore *et al.* (2010), o valor que delimita o primeiro quartil é a mediana entre o menor valor observado e a mediana de todos os dados, enquanto o valor que delimita o terceiro quartil é a mediana entre o maior valor observado e a mediana de todos os dados. A diferença entre o valor do terceiro quartil e o valor do primeiro quartil é chamada de amplitude interquartil, conforme a equação 3

Equação 4

$$\textit{Amplitude interquartil} = \textit{Terceiro quartil} - \textit{Primeiro quartil}$$

Em um *boxplot*, o limite inferior da caixa é definido como a diferença entre o valor do primeiro quartil e uma vez e meia o valor da amplitude interquartil, conforme equação 5. O valor do limite superior é igual ao valor do terceiro quartil mais uma vez e meia o valor da amplitude interquartil, conforme equação 6 (DEVORE *et al.*, 2010)

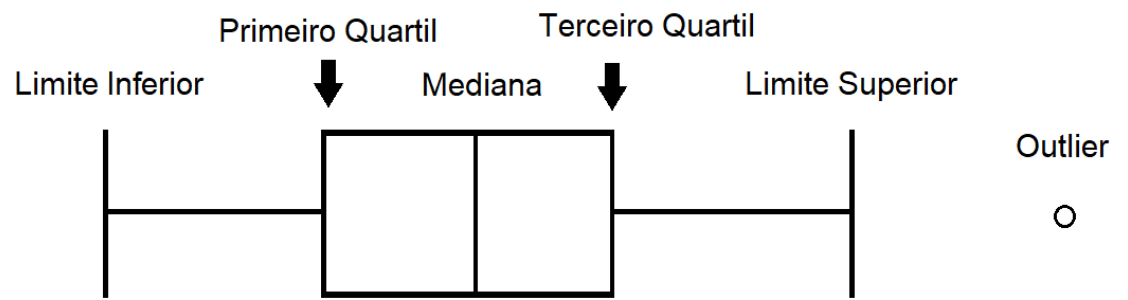
Equação 5

$$\textit{Limite inferior} = \textit{Primeiro Quartil} - 1,5x(\textit{Amplitude interquartil})$$

Equação 6

$$\textit{Limite superior} = \textit{Terceiro Quartil} + 1,5x(\textit{Amplitude interquartil})$$

Como descrito em Devore *et al.*, 2010 Os valores que são menores que o limite inferior ou maiores que o limite superior são considerados outliers. A Figura 9 ilustra os valores apresentados por um diagrama de caixa.

Figura 9: Exemplo de *boxplot*

Fonte: Elaboração Própria

4 RESULTADOS

Nessa seção, são apresentados os resultados obtidos a partir da aplicação do algoritmo de MacQueen para a técnica K-means nos indicadores estudados. O primeiro passo envolveu a remoção de *outliers*, seguido por uma análise considerando diferentes combinações de indicadores em cenários distintos.

4.1 ANÁLISE DE OUTLIERS

Dentro da base de dados analisada, alguns dados tiveram que ser removidos por representarem prováveis erros de coleta, ausência de dados fornecidos pelas fontes, erros no registro dos dados coletados, ou outros problemas que iriam distorcer consideravelmente os resultados podendo gerar novos *clusters* provenientes de observações errôneas.

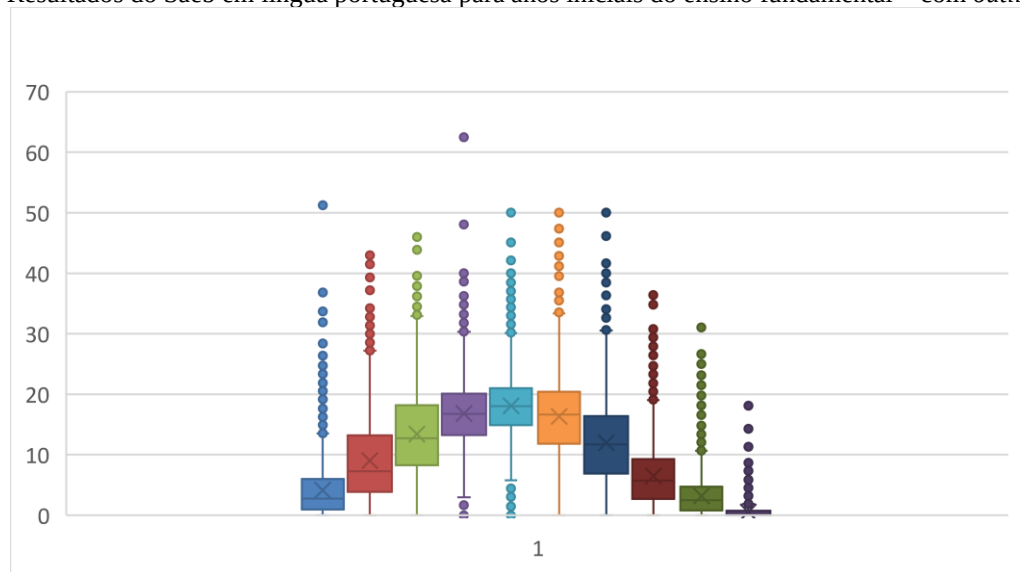
Esses dados são denominados *outliers*, segundo James *et. al* (2013), e foram removidos em duas etapas consecutivas: remoção de valores que apresentavam valores iguais a zero apesar de não ser possível essa condição (como a média dos alunos no Saeb), e análise de outros *outliers* a partir de um diagrama de caixa (*boxplot*) elaborado após a remoção dos *outliers* iniciais. Sendo assim, nesse trabalho os *outliers* são caracterizados por indicadores resultantes a valores acima ou abaixo da Amplitude Interquartil da amostra, conforme as equações 5 e 6.

O primeiro passo foi realizado por meio a verificação de valores que não poderiam ser nulos, mas apresentavam valores iguais a zero, sendo eles: média de todos os alunos no Saeb, número de alunos por turma, percentual de todos os alunos em todos os níveis do Inse, e percentual de todos os docentes em todos os níveis de esforço docente ou regularidade docente. Seguido pela remoção desses valores.

O segundo passo envolveu a aplicação do diagrama de caixa nos dados remanescentes, sendo que valores inferiores ao limite inferior, e superiores ao limite superior foram removidos já que esses valores também são considerados *outliers*.

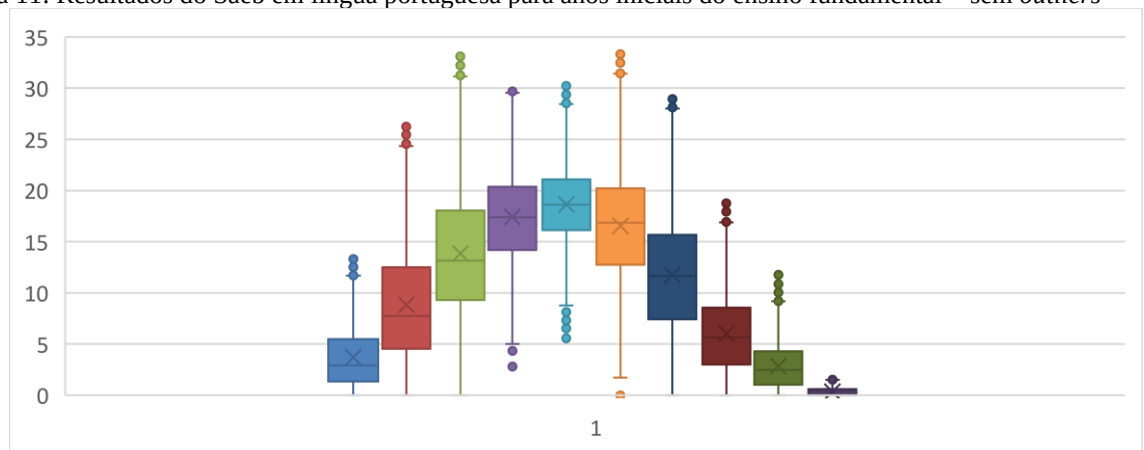
Alguns exemplos de *outliers* que foram removidos podem ser observados na **Error! Reference source not found.** e na **Error! Reference source not found.**, que demonstram a remoção de *outliers* que surgiram no percentual de alunos por nível do Saeb em língua portuguesa para os anos iniciais.

Figura 10: Resultados do Saeb em língua portuguesa para anos iniciais do ensino fundamental – com *outliers*



Fonte: INEP (2022b)

Figura 11: Resultados do Saeb em língua portuguesa para anos iniciais do ensino fundamental – sem *outliers*



Fonte: INEP (2022b)

4.2 DEFINIÇÃO DOS CLUSTERS

A definição dos *clusters* em cada etapa da educação básica seguiu três etapas consecutivas: (i) identificação e remoção de *outliers*, (ii) definição do número ideal de *clusters* e (iii) alocação das cidades em diferentes *clusters*.

A primeira etapa foi realizada a partir da análise dos diagramas *boxplot* para todos os indicadores, em que todas as cidades cujos valores foram considerados outliers foram removidas da análise.

As duas etapas seguintes foram executadas a partir do software “R” (R Project), sendo que a segunda etapa utilizou a função $W(k)$ para definir o número ideal de *clusters*, enquanto a terceira aplicou a técnica K-means segundo o algoritmo de MacQueen para alocar as cidades em cada *cluster*.

Os indicadores selecionados neste trabalho (percentual de alunos em nível cinco ou superior no Inse, percentual de docentes em nível de esforço menor ou igual a três, percentual de docentes com Indicador de Regularidade do Corpo Docente maior que 3, quantidade de alunos por sala, e percentual de alunos com nível Saeb igual ou superior a 5 em língua portuguesa) foram alocados em 5 situações diferentes apresentadas no Fonte: elaborado pelo autor

Quadro 11: Situações exploradas para definição dos *clusters*

Situação	Indicadores
Todos os Indicadores	• Percentual de alunos em nível cinco ou superior no Inse • Percentual de docentes em nível de esforço menor ou igual a três • Percentual de docentes com Indicador de Regularidade do Corpo Docente maior que 3 • Quantidade de alunos por sala • Percentual de alunos com nível Saeb igual ou superior a 5 em Língua Portuguesa
Indicadores socioeconômico e de estrutura das escolas	• Percentual de alunos em nível cinco ou superior no Inse • Quantidade de alunos por sala
Indicadores sobre os docentes	• Percentual de docentes em nível de esforço menor ou igual a três • Percentual de docentes com Indicador de Regularidade do Corpo Docente maior que 3
Indicadores socioeconômico, de estrutura das escolas e desempenho dos alunos	• Percentual de alunos em nível cinco ou superior no Inse • Quantidade de alunos por sala • Percentual de alunos com nível Saeb igual ou superior a 5 em Língua Portuguesa
Indicadores sobre os docentes e desempenho dos alunos	• Percentual de docentes em nível de esforço menor ou igual a três • Percentual de docentes com Indicador de Regularidade do Corpo Docente maior que 3 • Percentual de alunos com nível Saeb igual ou superior a 5 em Língua Portuguesa

Fonte: elaborado pelo autor

Todas as situações apresentadas foram aplicadas em seis cenários distintos considerando apenas instituições que pertencem ao ensino público para a formação dos *clusters*, esses cenários podem ser observados no Fonte: Elaborado pelo autor

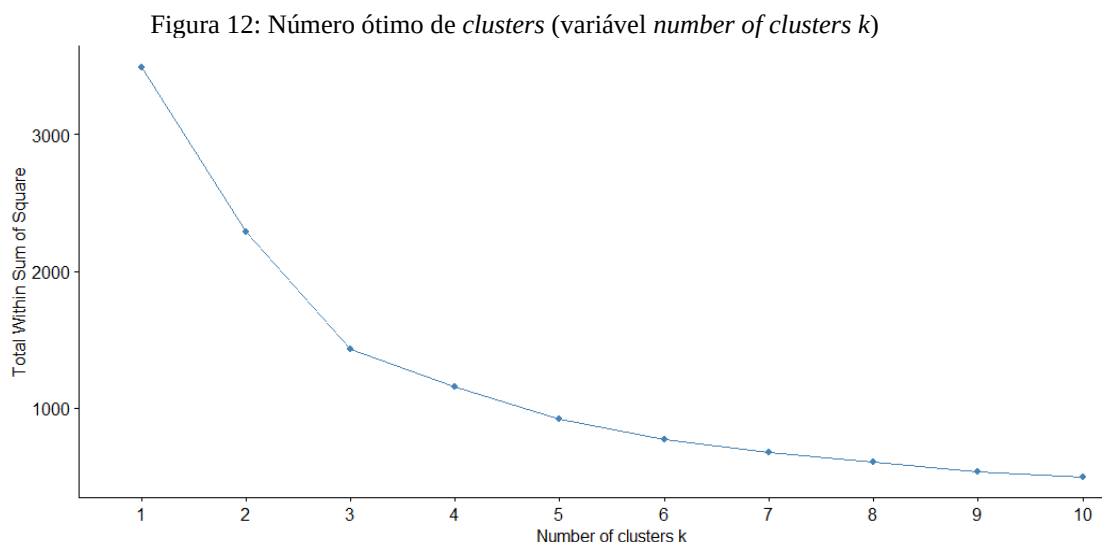
Quadro 12: Cenários para definição dos *clusters*

Cenário	Etapas de ensino	Localização
Cenário 1	Anos iniciais do ensino fundamental	Zona urbana
Cenário 2	Anos iniciais do ensino fundamental	Zona rural

Cenário 3	Anos finais do ensino fundamental	Zona urbana
Cenário 4	Anos finais do ensino fundamental	Zona rural
Cenário 5	Ensino médio	Zona urbana
Cenário 6	Ensino médio	Zona rural

Fonte: Elaborado pelo autor

O processo de definição do número ideal de *clusters* teve como base uma análise dos gráficos gerado pela função $W(k)$ em todos os indicadores descritos, sendo possível observar que o número ideal de *clusters* sugeridos variava entre 4 e 5 para qualquer situação de indicadores apresentada, ou seja, não há ganho na redução da soma do quadrado dos erros quando o número de *clusters* é maior que 5. A **Error! Reference source not found.** mostra um exemplo de um desses gráficos, onde é possível observar o resultado da aplicação da função $W(k)$ graficamente gerada pelo R, e como a partir de 5 *clusters* o valor da somatória do das distâncias dentro de um *cluster* tende a se estabilizar.



Fonte: Elaborado pelo autor

4.2.1 Resultados obtidos para os anos iniciais do ensino fundamental

Os anos iniciais do ensino fundamental podem ser definidos como os anos entre a 1º ano do ensino fundamental e o 5º ano do ensino fundamental em que mais de 15 milhões de alunos estão matriculados em mais de 109 mil escolas (INEP, 2019).

Aproximadamente 80% dos alunos nessa etapa estão matriculados em escolas públicas (**Error! Reference source not found.**) sendo a maioria administrada pela esfera municipal, e

aproximadamente 82% desses alunos estudam em escolas da zona urbana (**Error! Reference source not found.**).

Figura 13: Alunos dos anos iniciais do ensino fundamental na rede pública

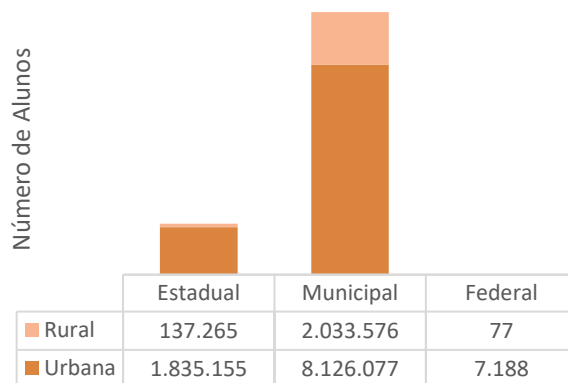
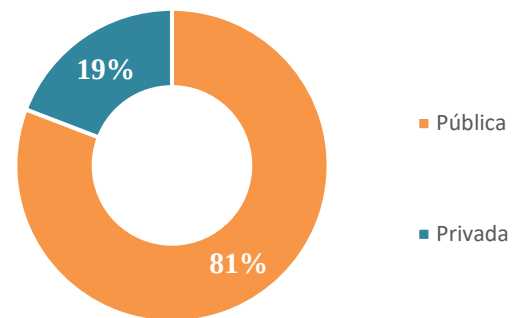


Figura 14: Proporção de alunos dos anos iniciais do ensino fundamental por rede de ensino



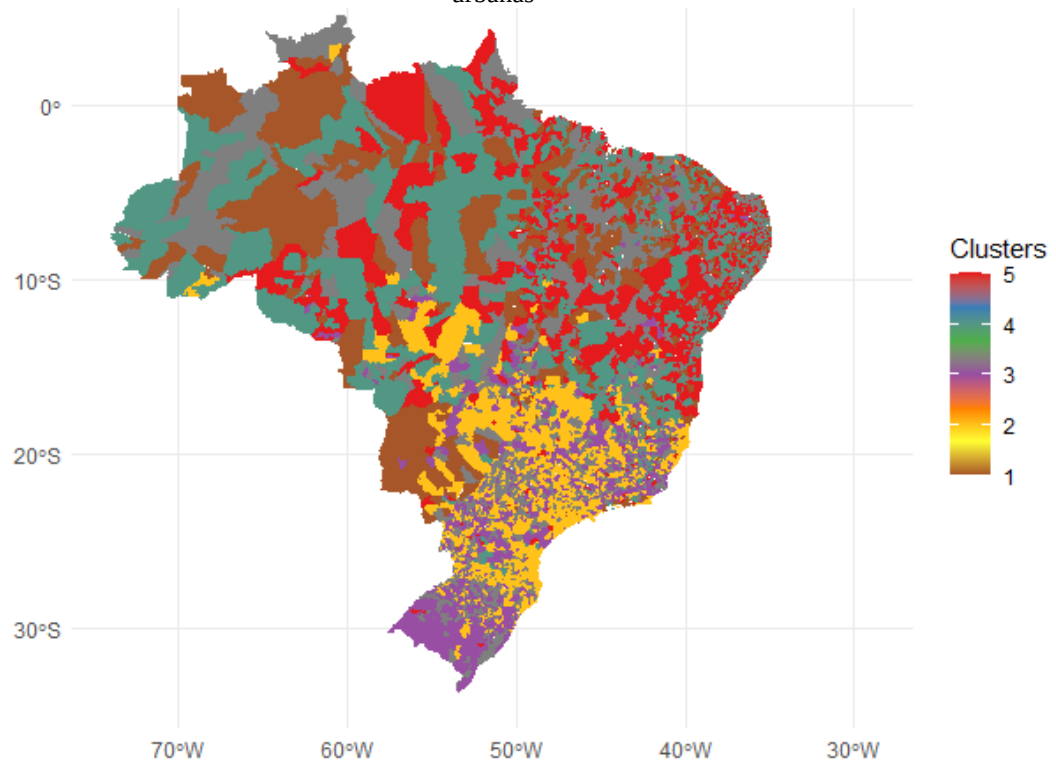
Fonte: INEP (2019)

4.2.1.1 Localização Urbana

O primeiro cenário a ser analisado envolve municípios que oferecem os anos iniciais do ensino fundamental em regiões urbanas, sendo que, após a remoção de *outliers*, 3772 municípios fizeram parte dessa análise.

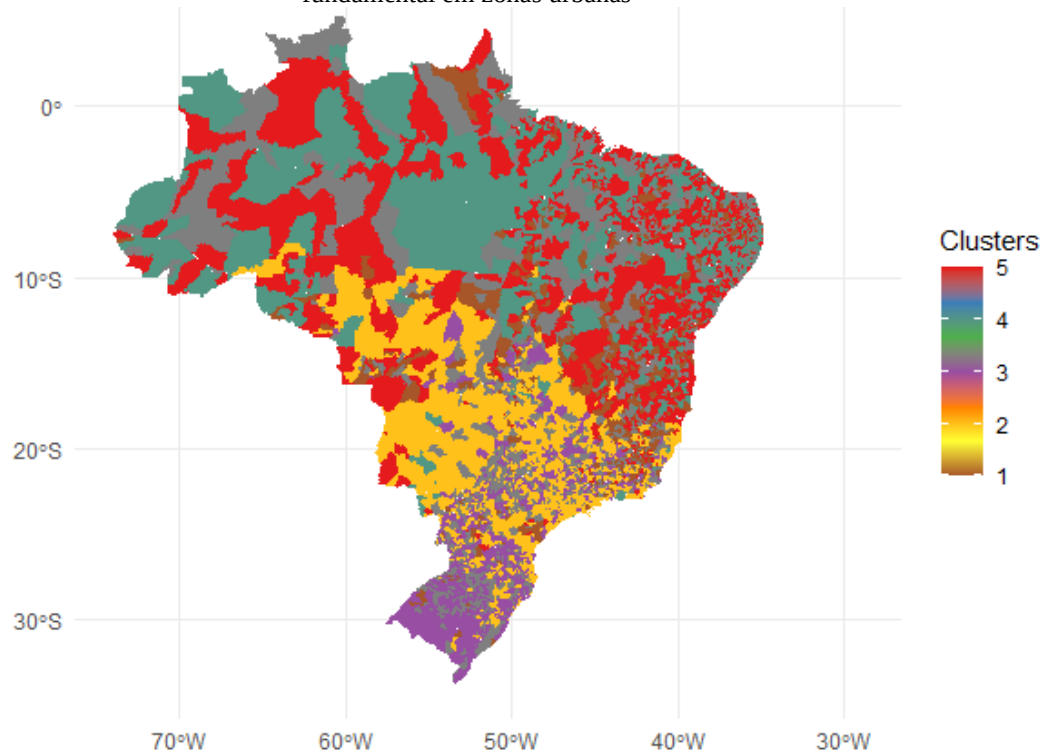
No total 5 *clusters* foram formados e a alocação de municípios em cada *cluster* seguiu as situações descritas no Quadro 11. A Figura 15 apresenta a aplicação de todos os indicadores, a Figura 16 e a Figura 17 mostram o impacto socioeconômico, estrutural (número de alunos por turma) e de desempenho dos alunos, enquanto as Figuras 18 e 19 apresentam a aplicação de indicadores sobre a condição dos docentes e de desempenho dos alunos.

Figura 15: *Clusters* considerando todos os indicadores para os anos iniciais do ensino fundamental em zonas urbanas



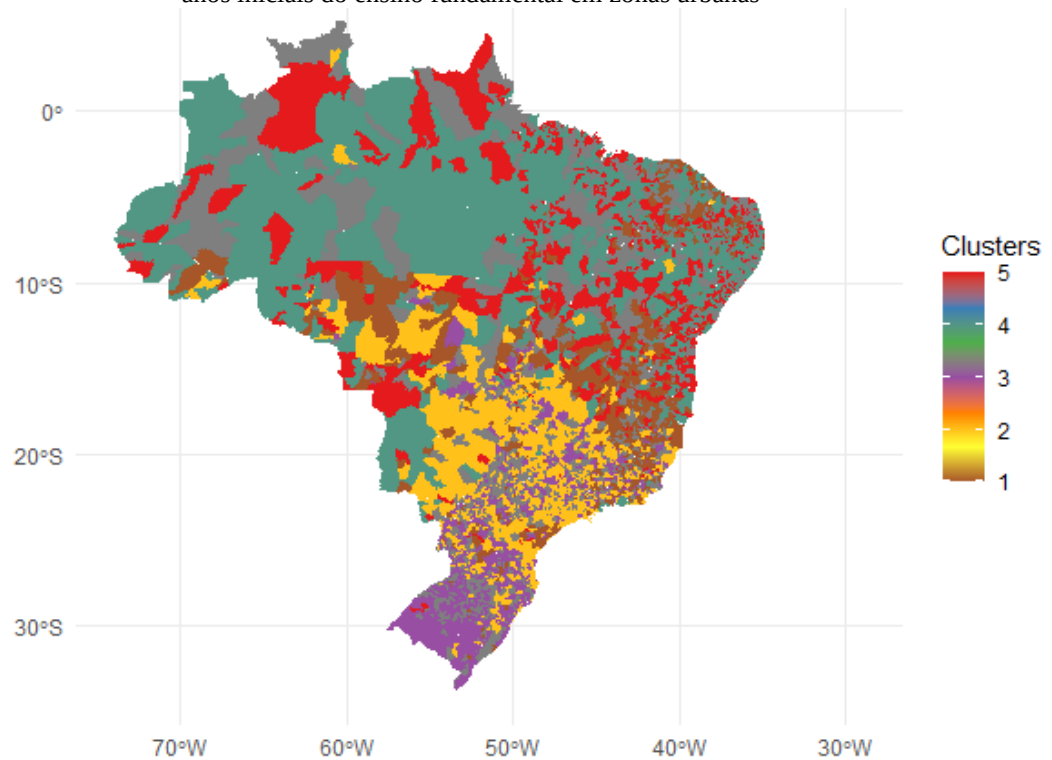
Fonte: Elaboração Própria

Figura 16: *Clusters* considerando indicadores socioeconômicos e estruturais para os anos iniciais do ensino fundamental em zonas urbanas



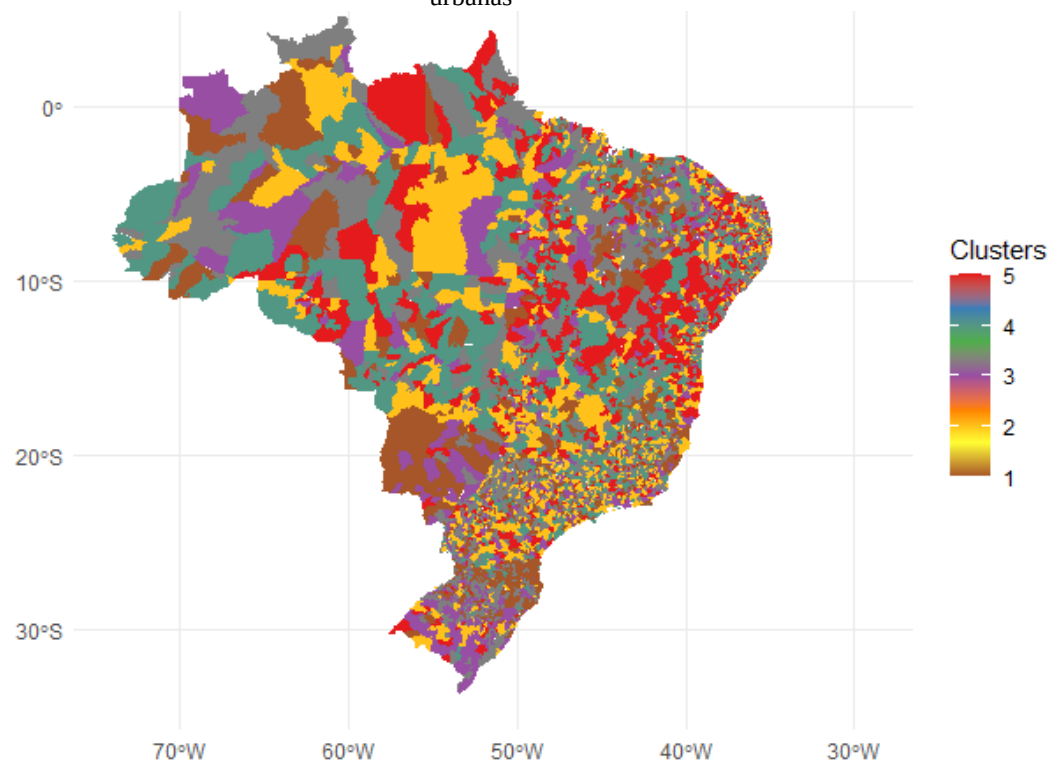
Fonte: Elaboração Própria

Figura 17: *Clusters* considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para os anos iniciais do ensino fundamental em zonas urbanas



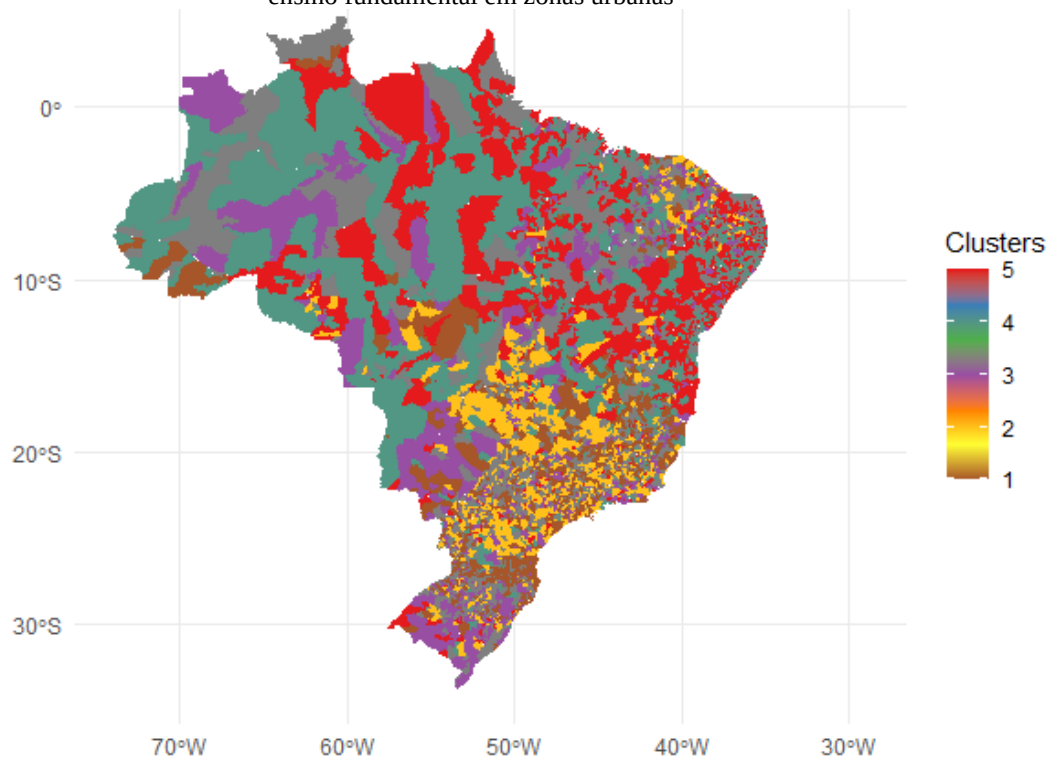
Fonte: Elaboração Própria

Figura 18: *Clusters* considerando indicadores dos docentes para os anos iniciais do ensino fundamental em zonas urbanas



Fonte: Elaboração Própria

Figura 19: *Clusters* considerando indicadores dos docentes e do desempenho dos alunos para os anos iniciais do ensino fundamental em zonas urbanas



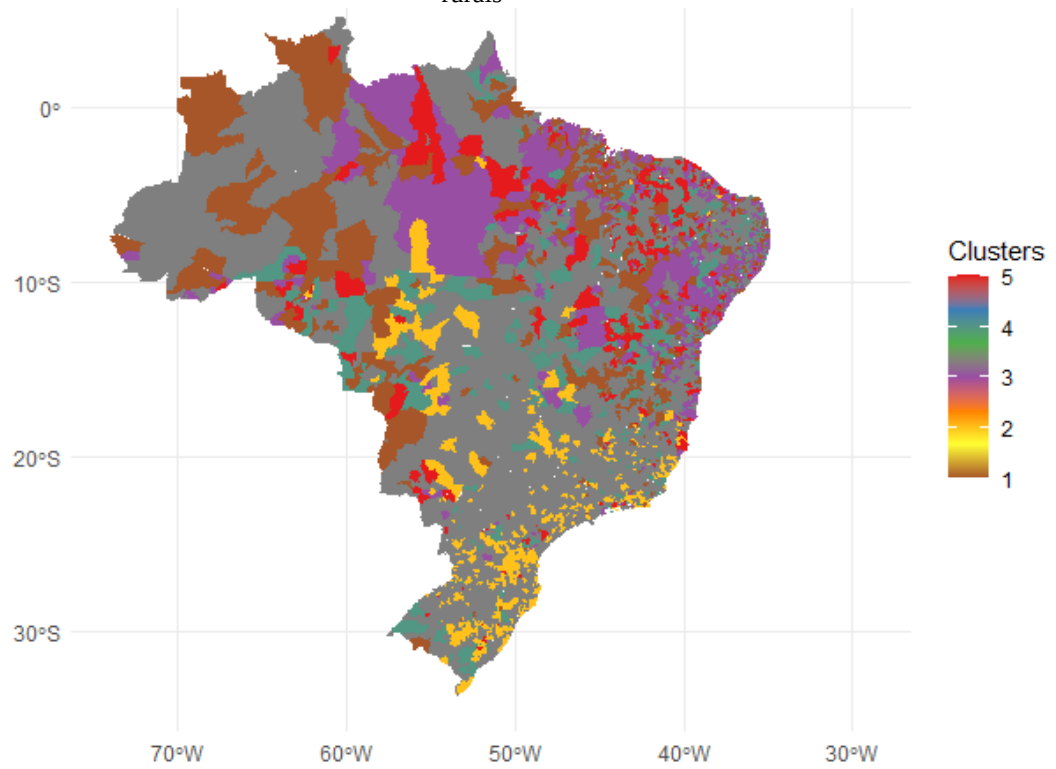
Fonte: Elaboração Própria

4.2.1.2 Localização Rural

O segundo cenário apresentado pelo Quadro 12 envolve municípios que oferecem os anos iniciais do ensino fundamental para regiões rurais. Após a remoção de *outliers*, 1966 municípios foram analisados.

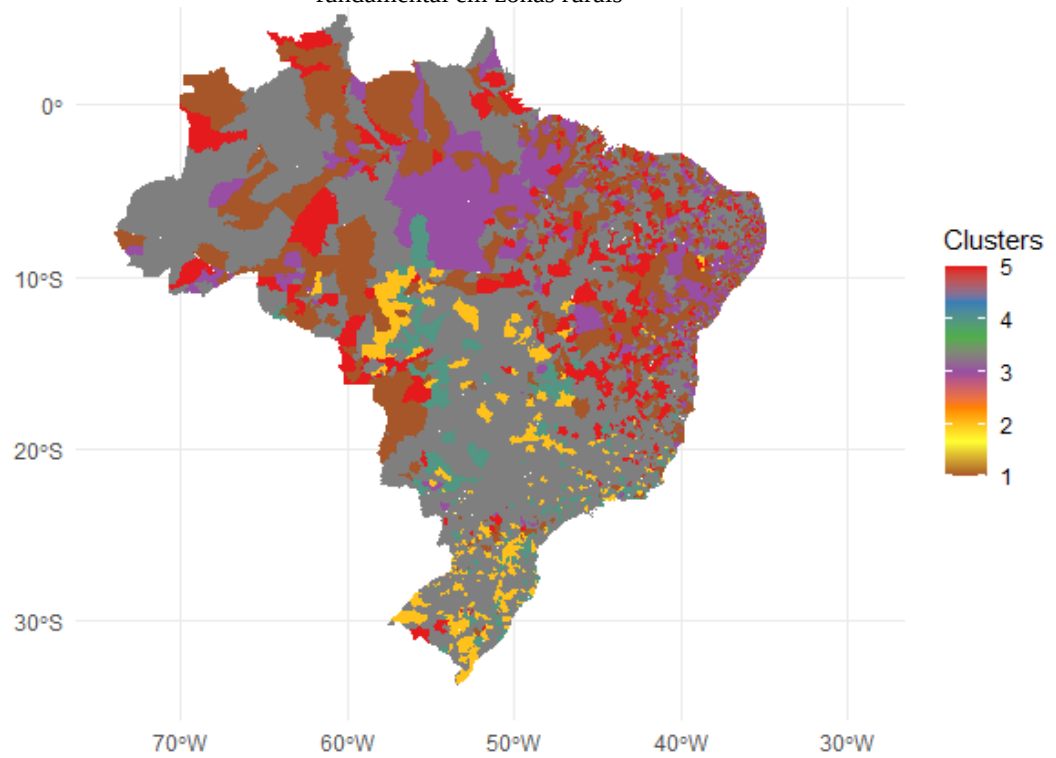
Assim como o primeiro cenário, 5 *clusters* foram analisados em diferentes combinações de indicadores. A Figura 20 apresenta a alocação de municípios em cada *cluster* quando todos os indicadores são levados em consideração, as Figuras 21 e 22 demonstram essa alocação quando indicadores socioeconômicos, estruturais e de desempenho dos alunos são levados em consideração, e as Figuras 23 e 24 mostram a distribuição de municípios em cada *cluster* quando indicadores sobre a condição dos docentes e de desempenho dos alunos são considerados.

Figura 20: *Clusters* considerando todos os indicadores para os anos iniciais do ensino fundamental em zonas rurais



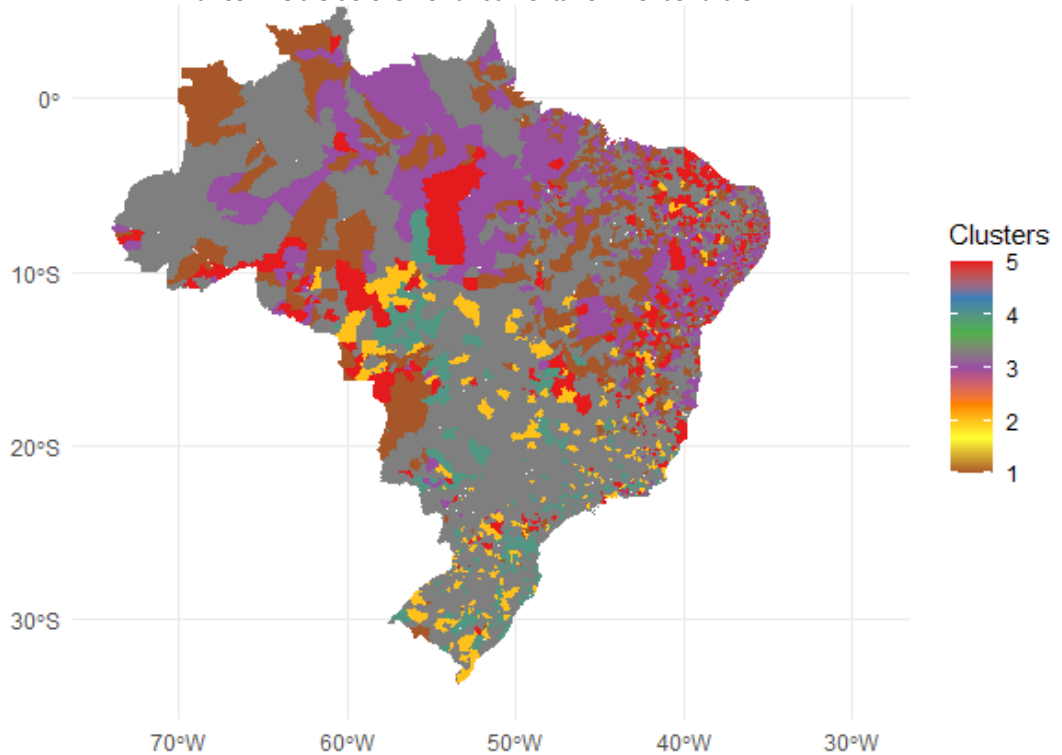
Fonte: Elaboração Própria

Figura 21: *Clusters* considerando indicadores socioeconômicos e estruturais para os anos iniciais do ensino fundamental em zonas rurais



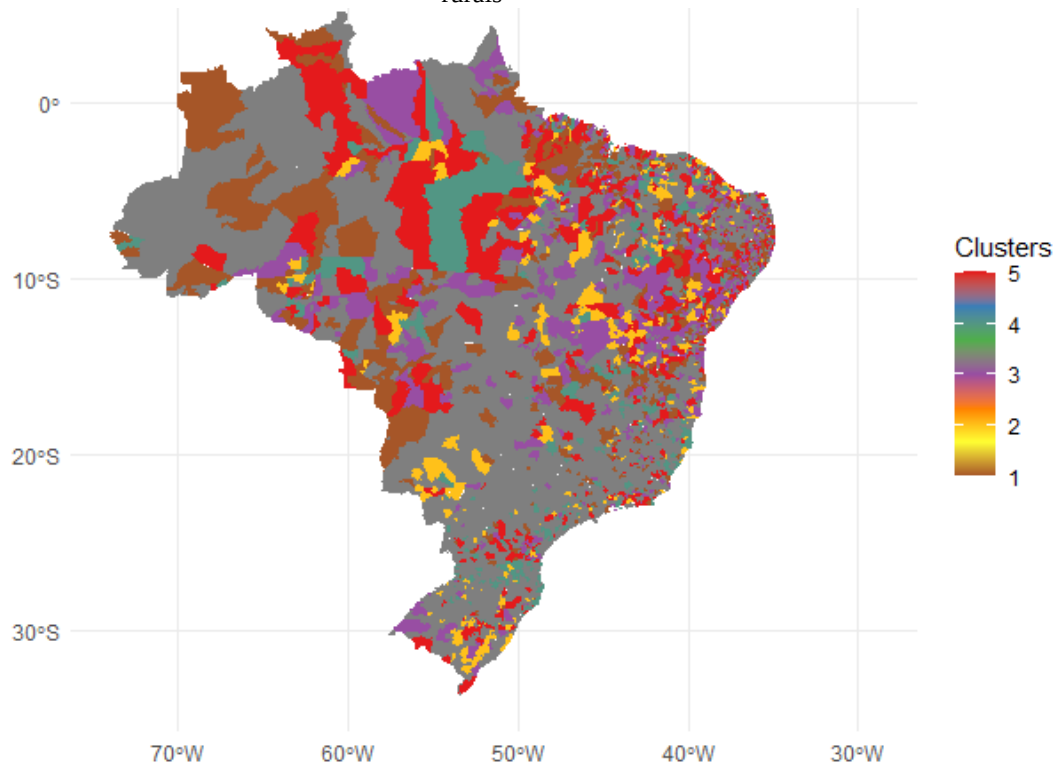
Fonte: Elaboração Própria

Figura 22: *Clusters* considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para os anos iniciais do ensino fundamental em zonas rurais



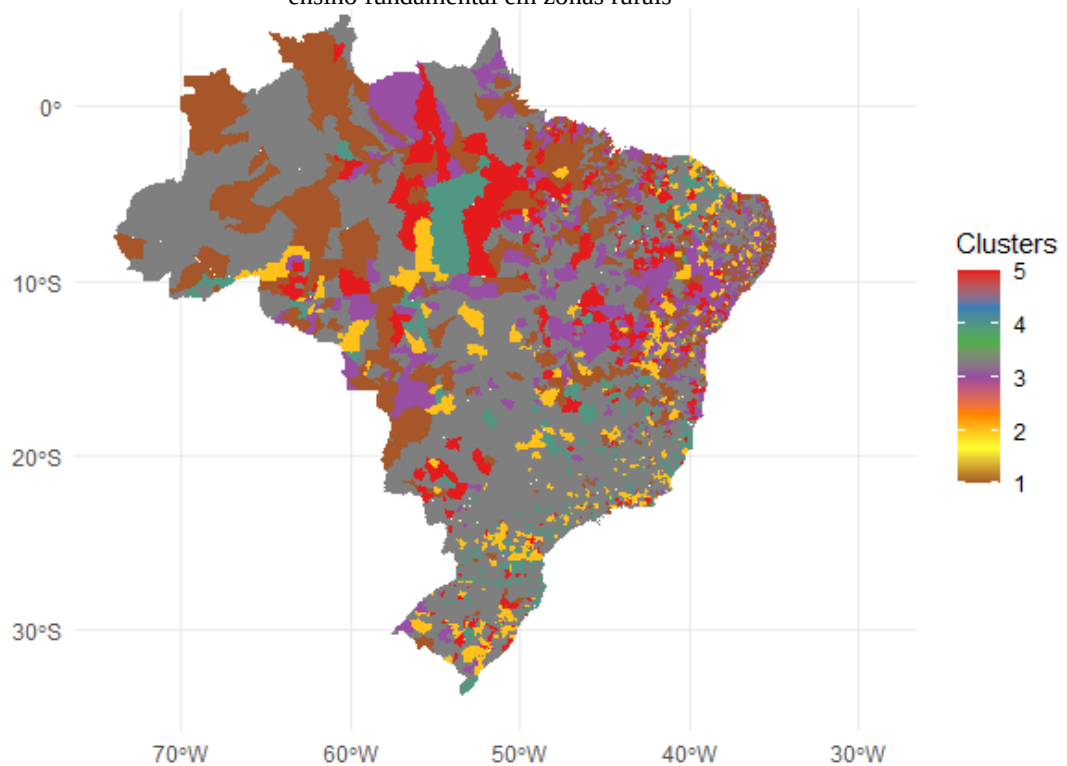
Fonte: Elaboração Própria

Figura 23: *Clusters* considerando indicadores dos docentes para os anos iniciais do ensino fundamental em zonas rurais



Fonte: Elaboração Própria

Figura 24: Clusters considerando indicadores dos docentes e do desempenho dos alunos para os anos iniciais do ensino fundamental em zonas rurais



Fonte: Elaboração Própria

4.2.2 Resultados obtidos para os anos Finais do Ensino Fundamental

Correspondente as turmas que vão do 6º ano do ensino fundamental ao 9º ano do ensino fundamental, os anos finais do ensino médio possuem aproximadamente 12 milhões de alunos matriculados em mais de 61 mil escolas.

A gestão das escolas nas quais esses alunos estão matriculados estão divididos predominantemente entre esfera Estadual e Municipal na rede pública de ensino. Sendo que aproximadamente 87% desses alunos atendem as aulas nas regiões urbanas dos municípios.

Figura 25: Alunos dos anos finais do ensino fundamental na rede pública

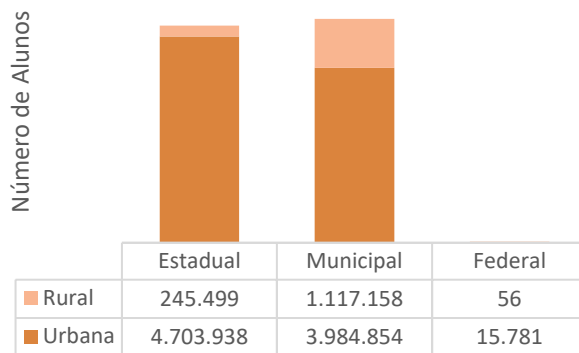
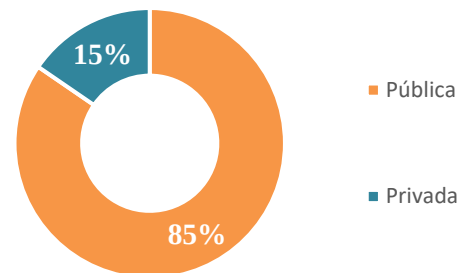


Figura 26: Proporção de alunos dos anos finais do ensino fundamental por rede de ensino



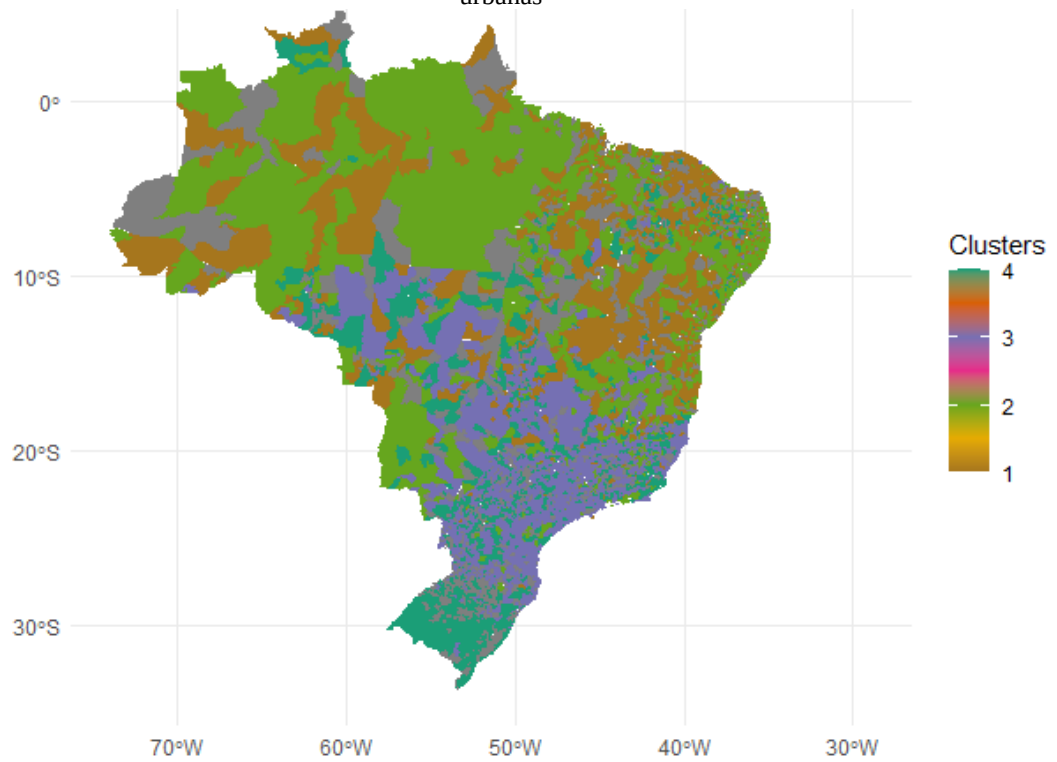
Fonte: INEP (2019)

4.2.2.1 Localização Urbana

Neste estudo, 4093 municípios que oferecem os anos finais do ensino fundamental em áreas urbanas foram analisados após a remoção de *outliers*. Esses municípios foram distribuídos em 4 *clusters* considerando as situações apresentadas no Quadro 11.

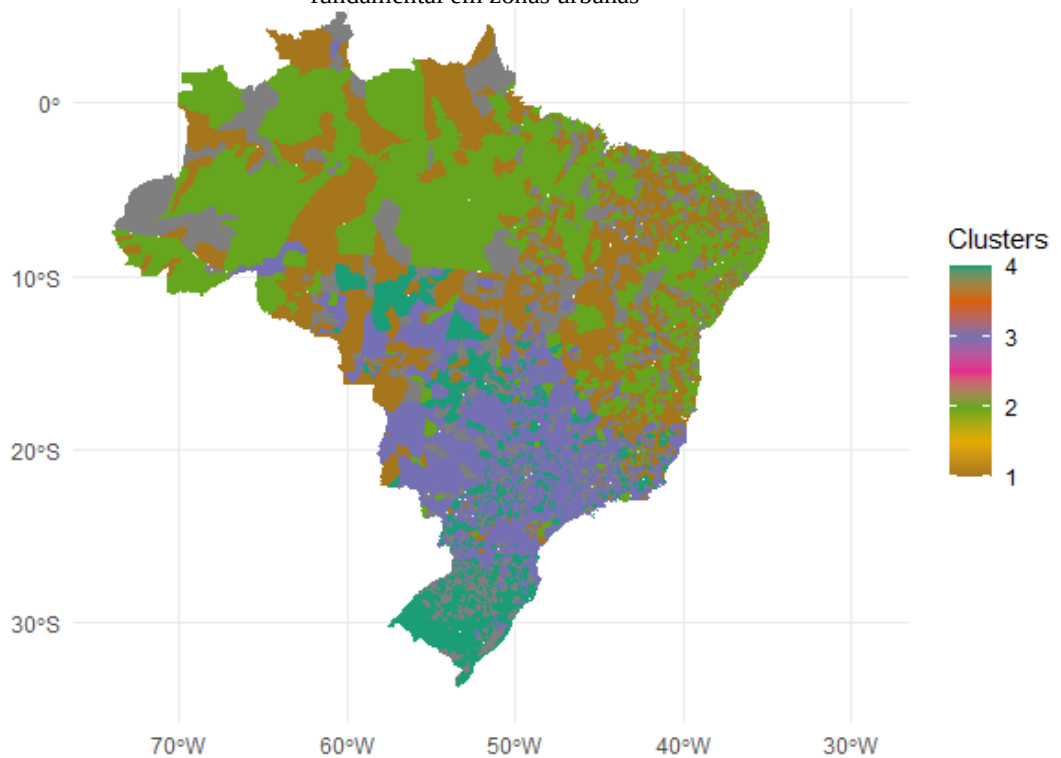
A alocação de municípios na situação em que todos os indicadores foram levados em consideração pode ser observada na Figura 27, a composição de *clusters* quando indicadores socioeconômicos, de número de alunos por turma e de desempenho dos alunos são considerados pode ser observada nas Figuras 28 e 29; e as Figuras 30 e 31 apresentam a distribuição dos municípios em *clusters* quando indicadores de condição dos docentes e de desempenho dos alunos são analisados.

Figura 27: *Clusters* considerando todos os indicadores para os anos finais do ensino fundamental em zonas urbanas



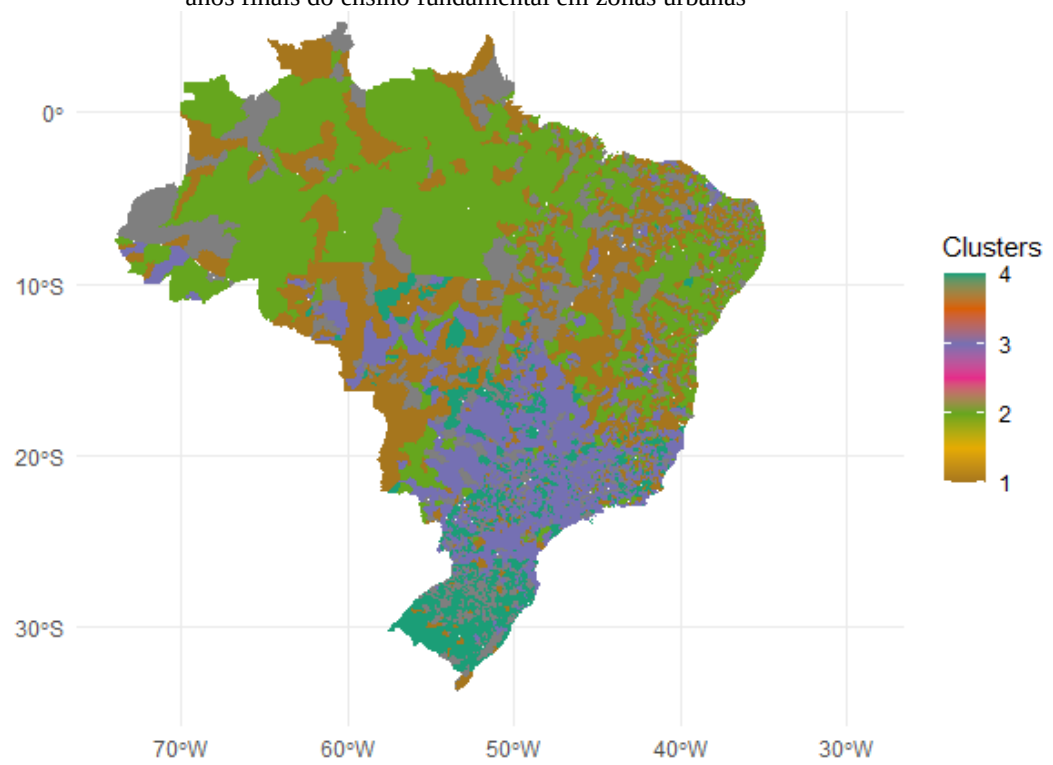
Fonte: Elaboração Própria

Figura 28: *Clusters* considerando indicadores socioeconômicos e estruturais para os anos finais do ensino fundamental em zonas urbanas



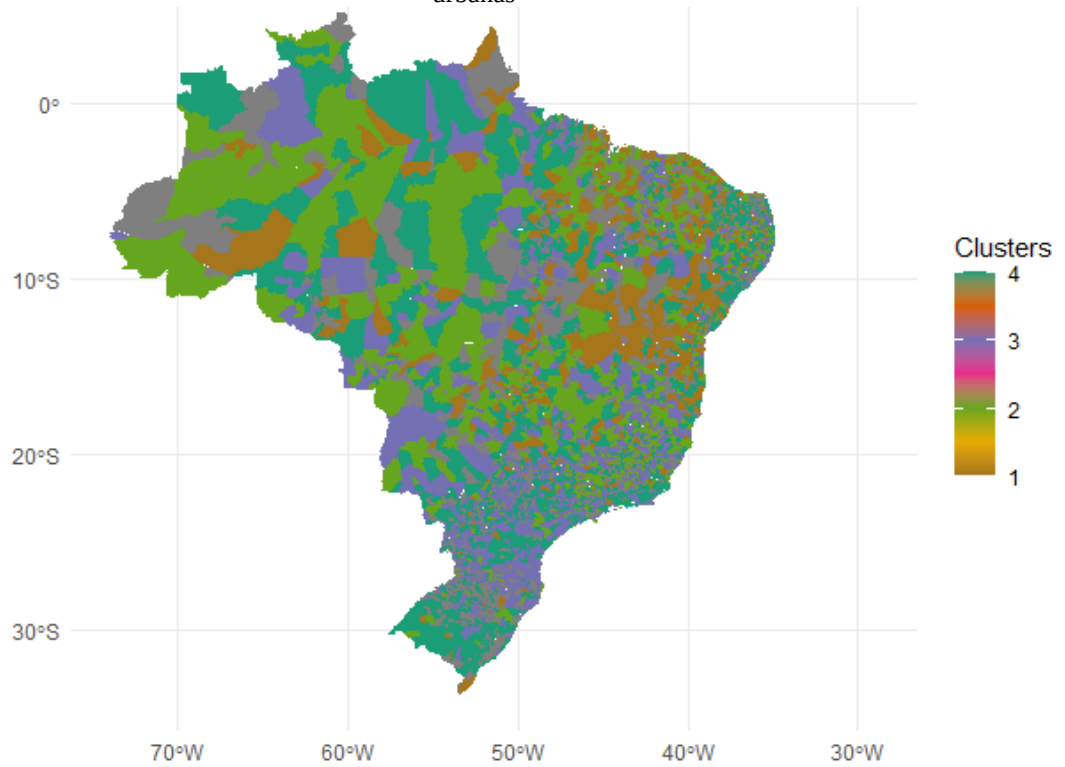
Fonte: Elaboração Própria

Figura 29: *Clusters* considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para os anos finais do ensino fundamental em zonas urbanas



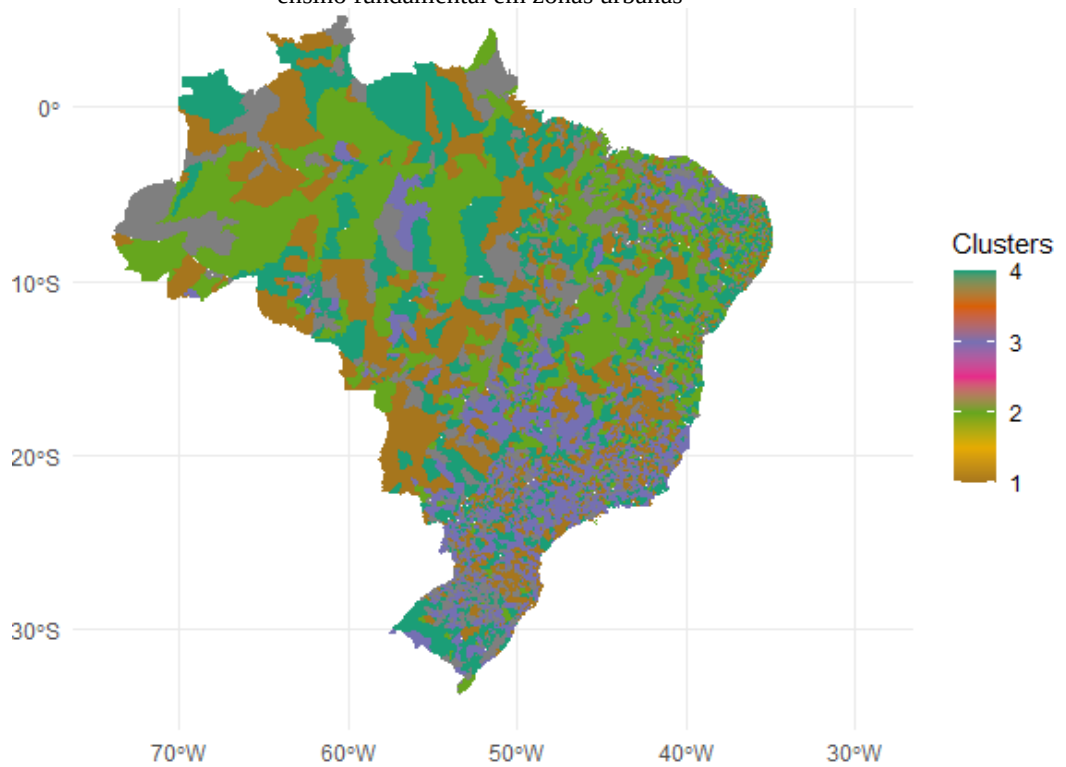
Fonte: Elaboração Própria

Figura 30: *Clusters* considerando indicadores dos docentes para os anos finais do ensino fundamental em zonas urbanas



Fonte: Elaboração Própria

Figura 31: *Clusters* considerando indicadores dos docentes e do desempenho dos alunos para os anos finais do ensino fundamental em zonas urbanas



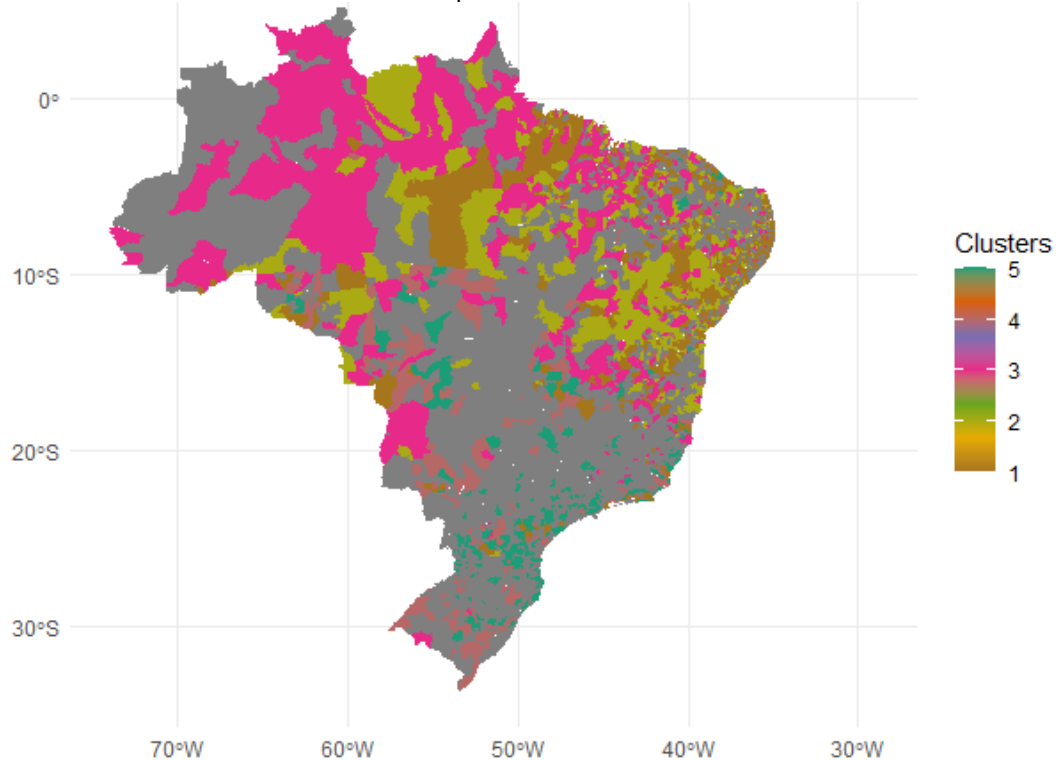
Fonte: Elaboração Própria

4.2.2.2 Localização Rural

O quarto cenário apresentado pelo Quadro 12 envolveu a alocação de 1748 municípios que oferecem os anos finais do ensino fundamental para áreas rurais em 5 *clusters* após a remoção de *outliers*.

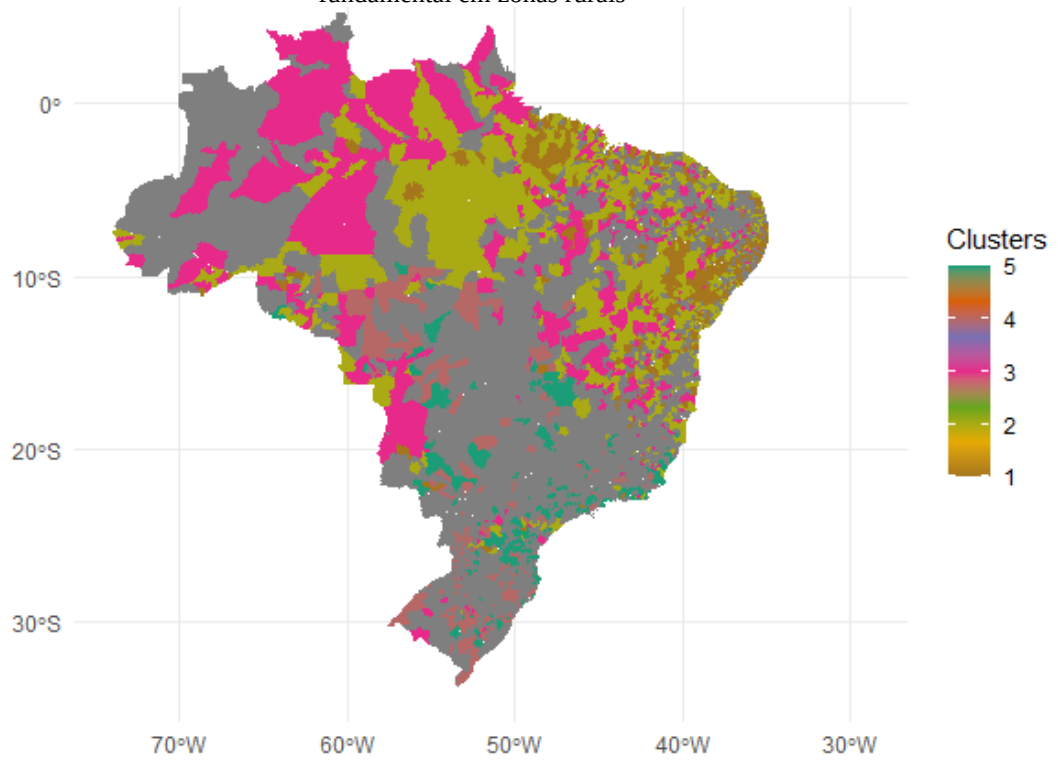
A situação em que todos os indicadores são analisados pode ser observada na Figura 32, e as situações que levam em consideração indicadores socioeconômicos, estruturais e de desempenho dos alunos podem ser observadas nas Figuras 33 e 34; enquanto as Figuras 35 e 36 apresentam a alocação de municípios nos 5 *clusters* quando indicadores de condição dos docentes e desempenho dos alunos considerados.

Figura 32: *Clusters* considerando todos os indicadores para os anos finais do ensino fundamental em zonas rurais



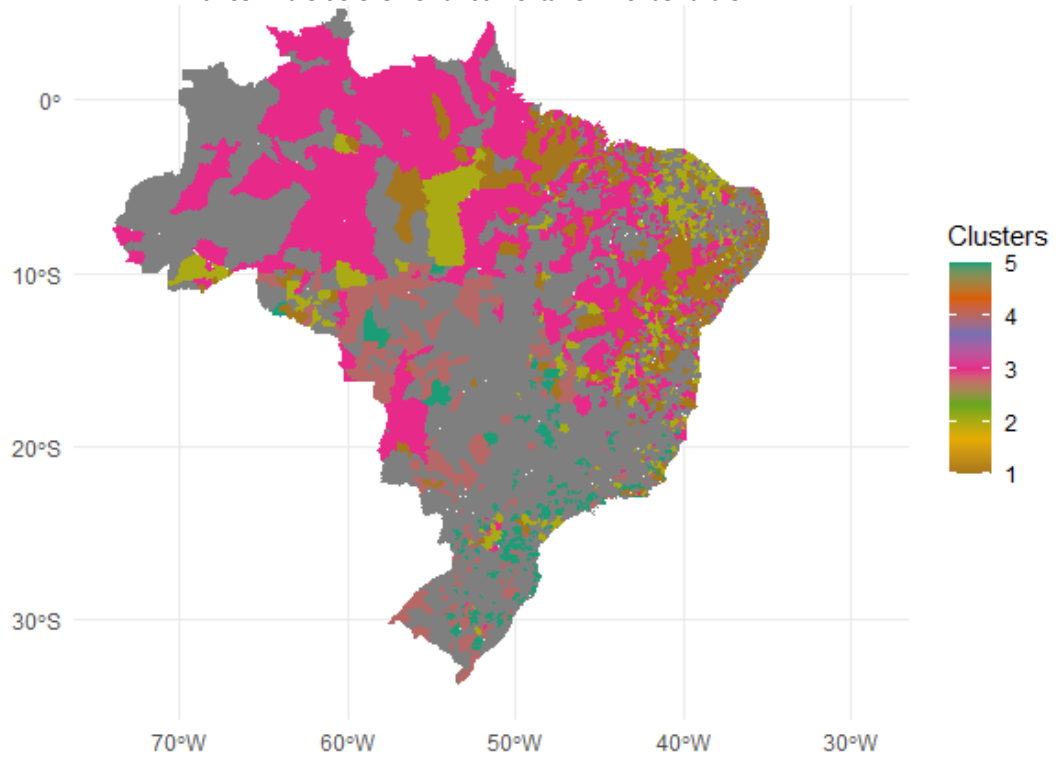
Fonte: Elaboração Própria

Figura 33: *Clusters* considerando indicadores socioeconômicos e estruturais para os anos finais do ensino fundamental em zonas rurais



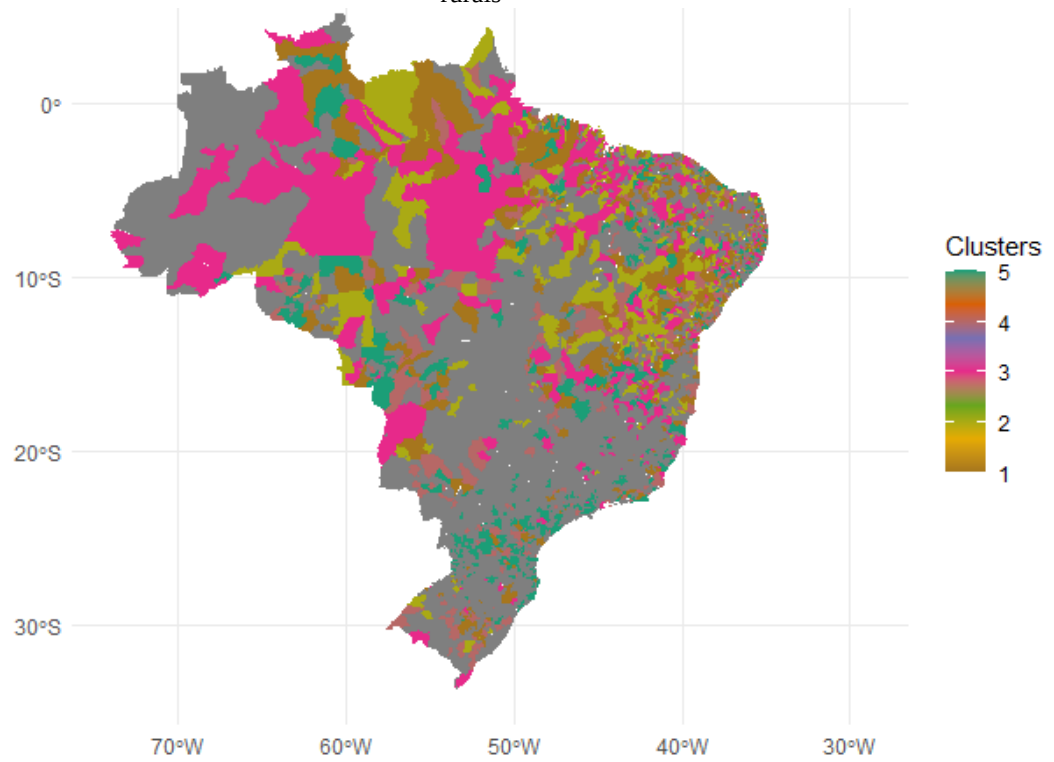
Fonte: Elaboração Própria

Figura 34: *Clusters* considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para os anos finais do ensino fundamental em zonas rurais



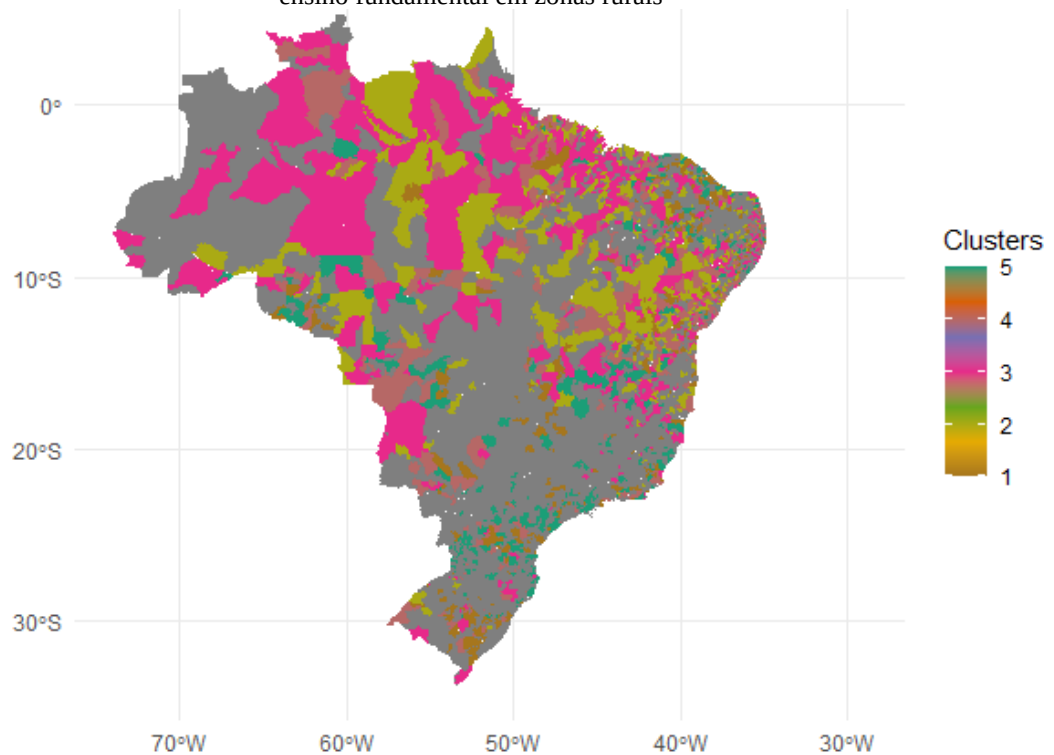
Fonte: Elaboração Própria

Figura 35: *Clusters* considerando indicadores dos docentes para os anos finais do ensino fundamental em zonas rurais



Fonte: Elaboração Própria

Figura 36: *Clusters* considerando indicadores dos docentes e do desempenho dos alunos para os anos finais do ensino fundamental em zonas rurais



Fonte: Elaboração Própria

4.2.3 Resultados obtidos para o Ensino Médio

O ensino médio corresponde aos anos finais da educação básica, com aproximadamente 7,5 milhões de alunos matriculados em 28860 escolas (INEP, 2019). Neste estudo só foram levados em consideração alunos entre a 1ª série e a 3ª série do ensino médio, ou seja, turmas agregadas não foram levadas em consideração.

Dos alunos matriculados na rede pública de ensino, 95% atendem a aulas em regiões urbanas, sendo que nesta etapa de ensino a gestão das escolas é predominantemente realizada na esfera Estadual.

Figura 37: Alunos do ensino médio na rede pública

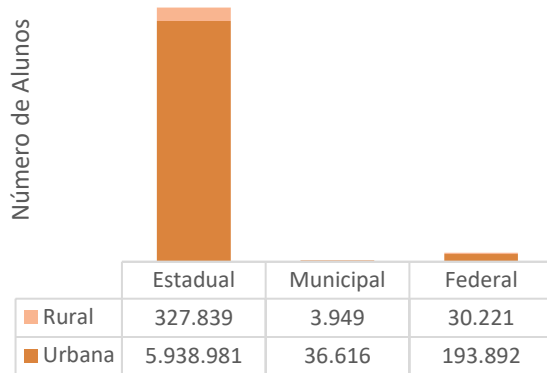
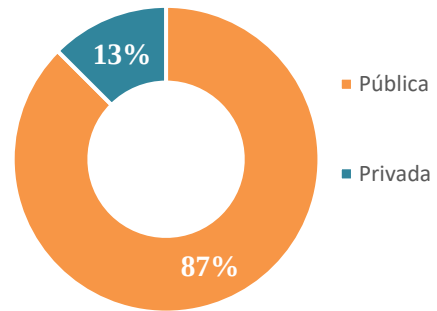


Figura 38: Proporção de alunos do ensino médio por rede de ensino



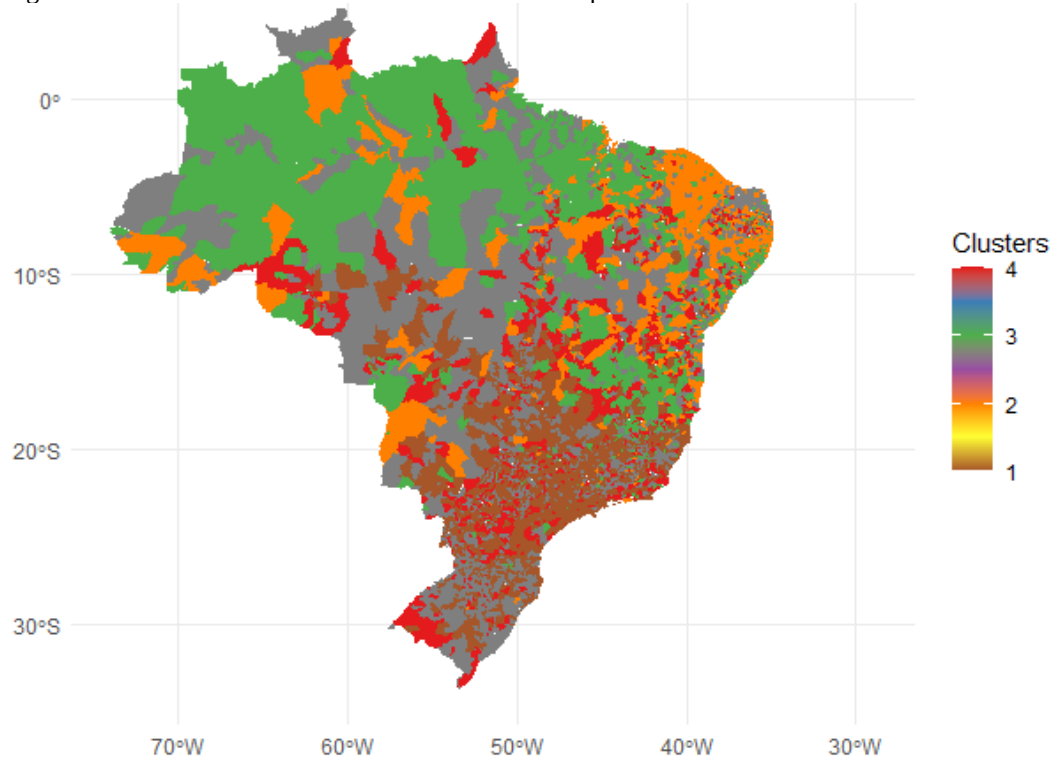
Fonte: INEP (2019)

4.2.3.1 Localização Urbana

Após a remoção de *outliers*, 3486 municípios que oferecem o ensino médio em regiões urbanas foram analisados. Esses municípios foram distribuídos entre 4 *clusters* após levar em consideração as situações presentes no Quadro 11.

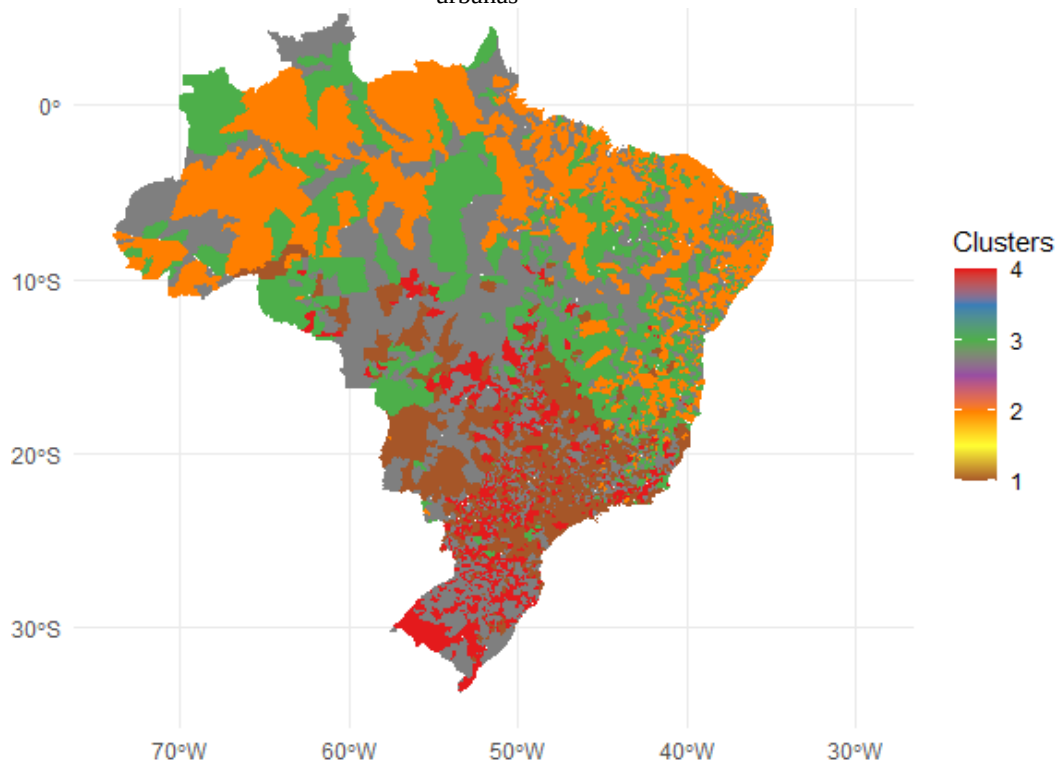
A distribuição de municípios por *cluster* quando todos os indicadores são considerados pode ser observada na Figura 39, as situações em que indicadores socioeconômicos, de alunos por turma e de desempenho dos alunos são apresentadas nas Figuras 40 e 41; a composição de *clusters* quando indicadores de condição dos docentes e de desempenho dos alunos são levados em consideração podem ser observados nas Figuras 42 e 42.

Figura 39: *Clusters* considerando todos os indicadores para o ensino médio em zonas urbanas



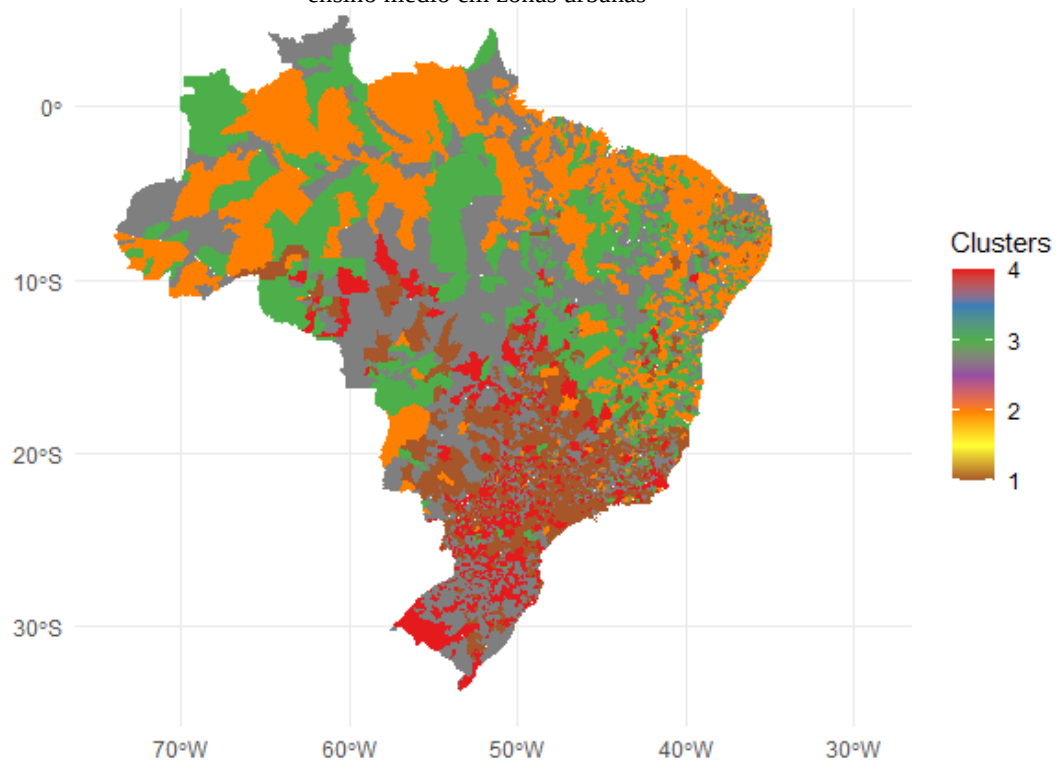
Fonte: Elaboração Própria

Figura 40: *Clusters* considerando indicadores socioeconômicos e estruturais para o ensino médio em zonas urbanas



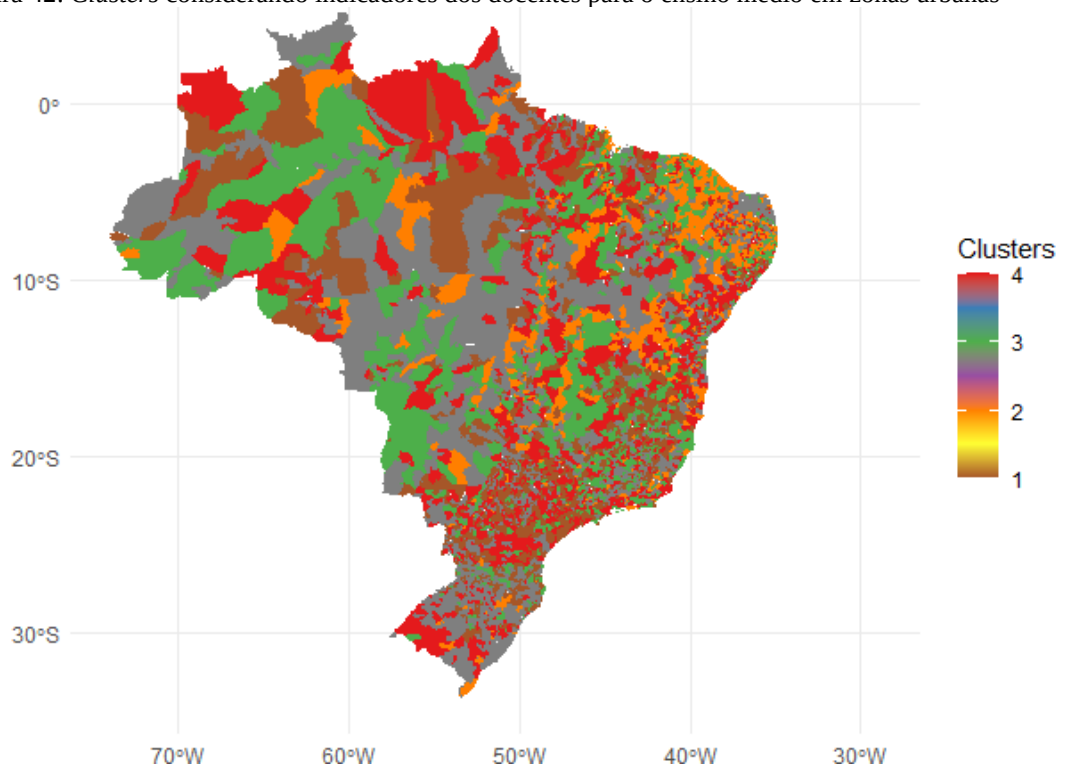
Fonte: Elaboração Própria

Figura 41: *Clusters* considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para o ensino médio em zonas urbanas



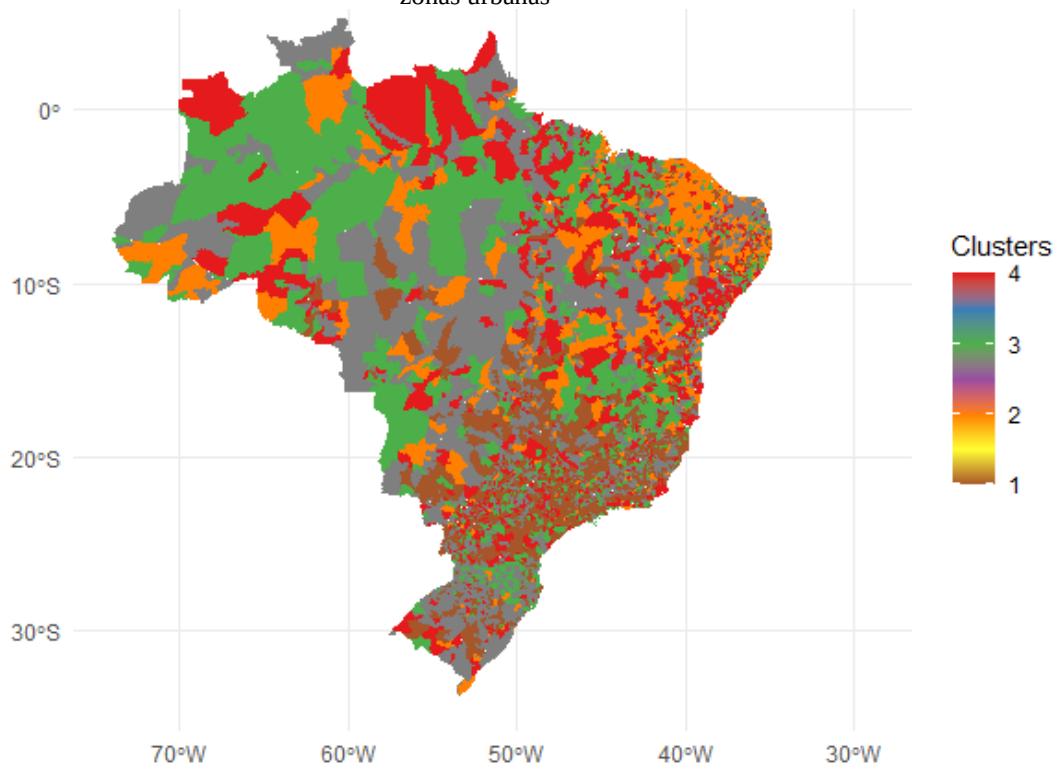
Fonte: Elaboração Própria

Figura 42: *Clusters* considerando indicadores dos docentes para o ensino médio em zonas urbanas



Fonte: Elaboração Própria

Figura 43: *Clusters* considerando indicadores dos docentes e do desempenho dos alunos para o ensino médio em zonas urbanas



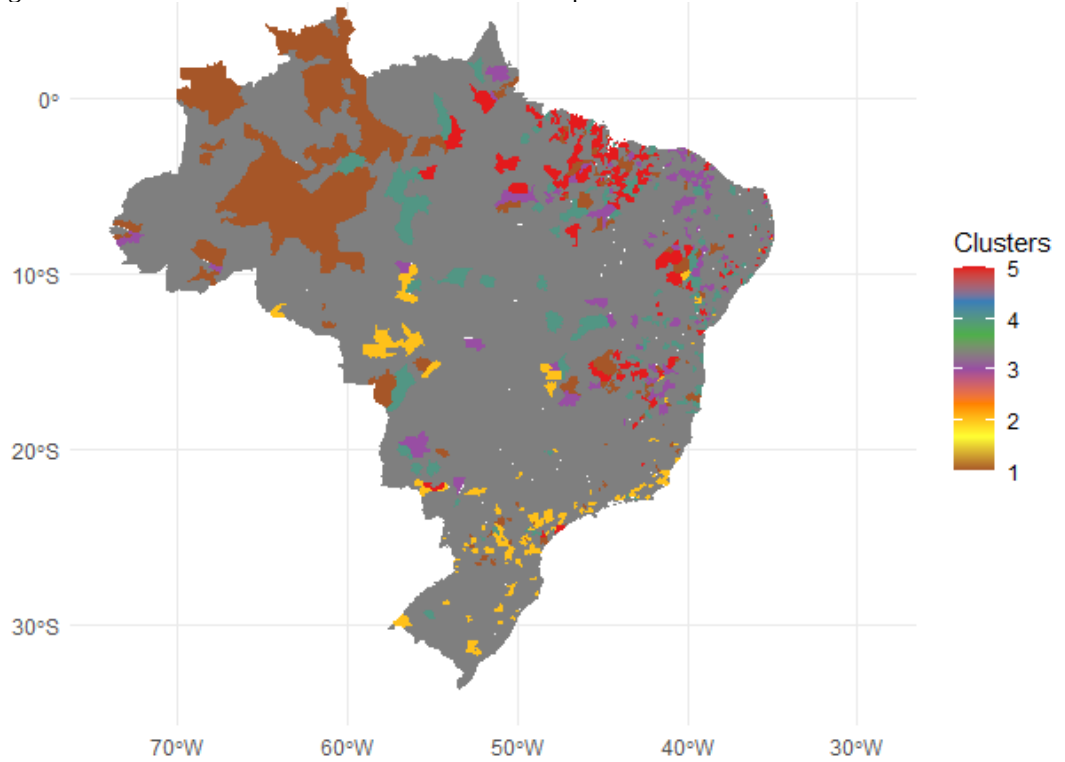
Fonte: Elaboração Própria

4.2.3.2 Localização Rural

Apenas 587 municípios que oferecem o ensino médio em regiões rurais foram analisados após a remoção de *outliers*. Esses municípios foram distribuídos entre 5 *clusters* após a aplicação da função $W(k)$.

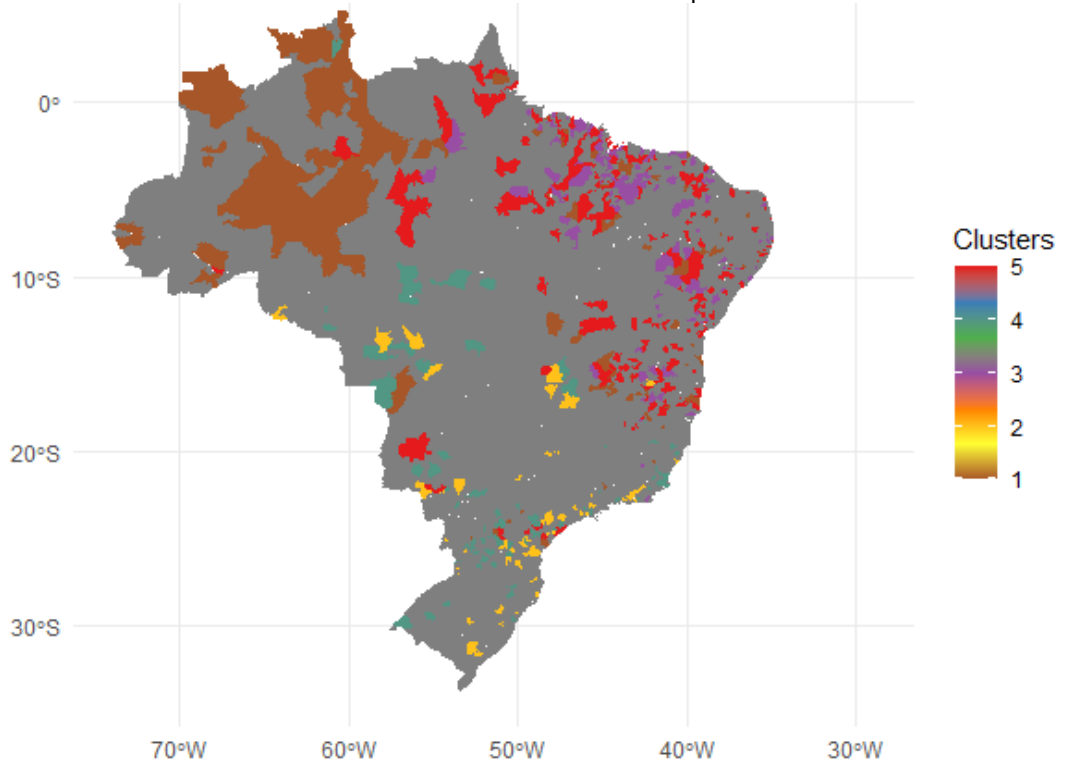
A Figura 44 mostra a distribuição dos municípios por *cluster* quando todos os indicadores são considerados, as Figuras 45 e 46 apresentam a composição dos *clusters* quando os indicadores socioeconômicos, estruturais e de desempenho dos alunos são levados em consideração; e as Figuras 47 e 48 apresentam as situações em que indicadores de condição dos docentes e desempenho dos alunos são estudados.

Figura 44: *Clusters* considerando todos os indicadores para o ensino médio em zonas rurais



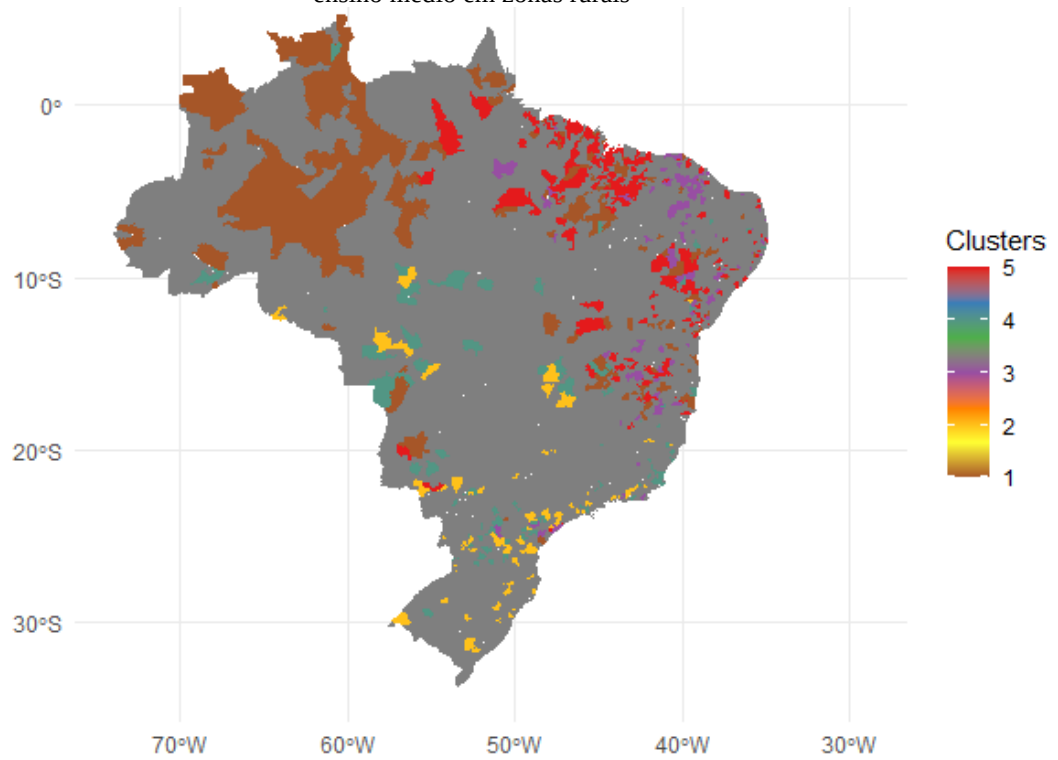
Fonte: Elaboração Própria

Figura 45: *Clusters* considerando indicadores socioeconômicos e estruturais para o ensino médio em zonas rurais



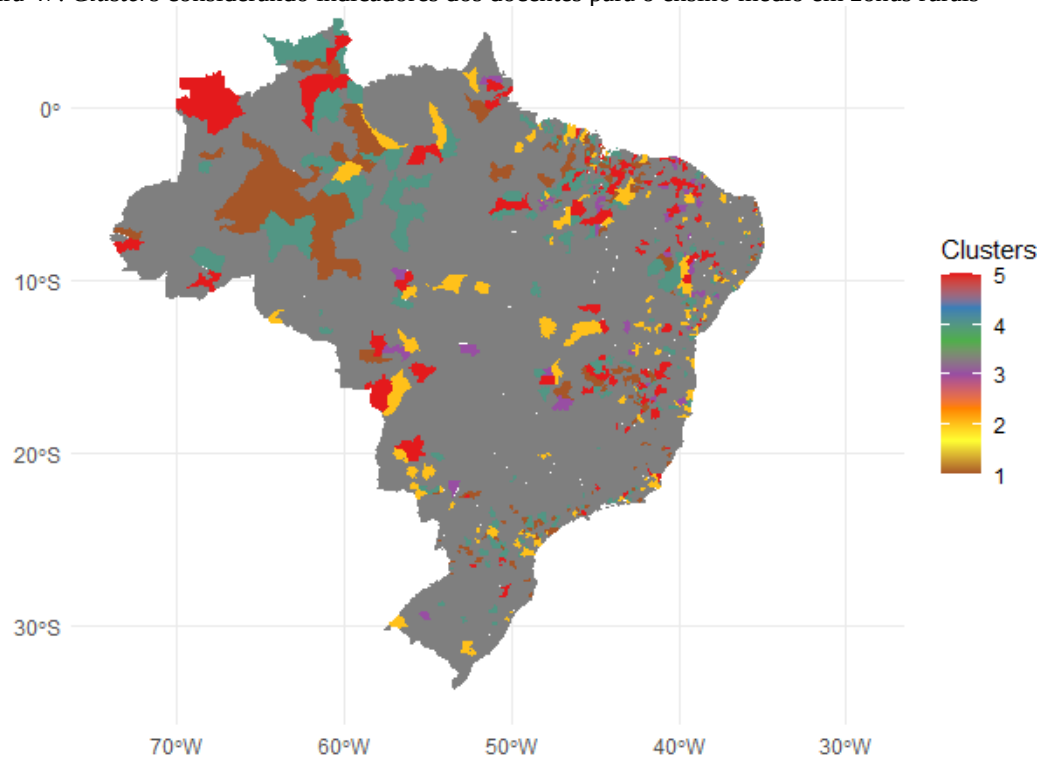
Fonte: Elaboração Própria

Figura 46: *Clusters* considerando indicadores socioeconômicos, estruturais e do desempenho dos alunos para o ensino médio em zonas rurais



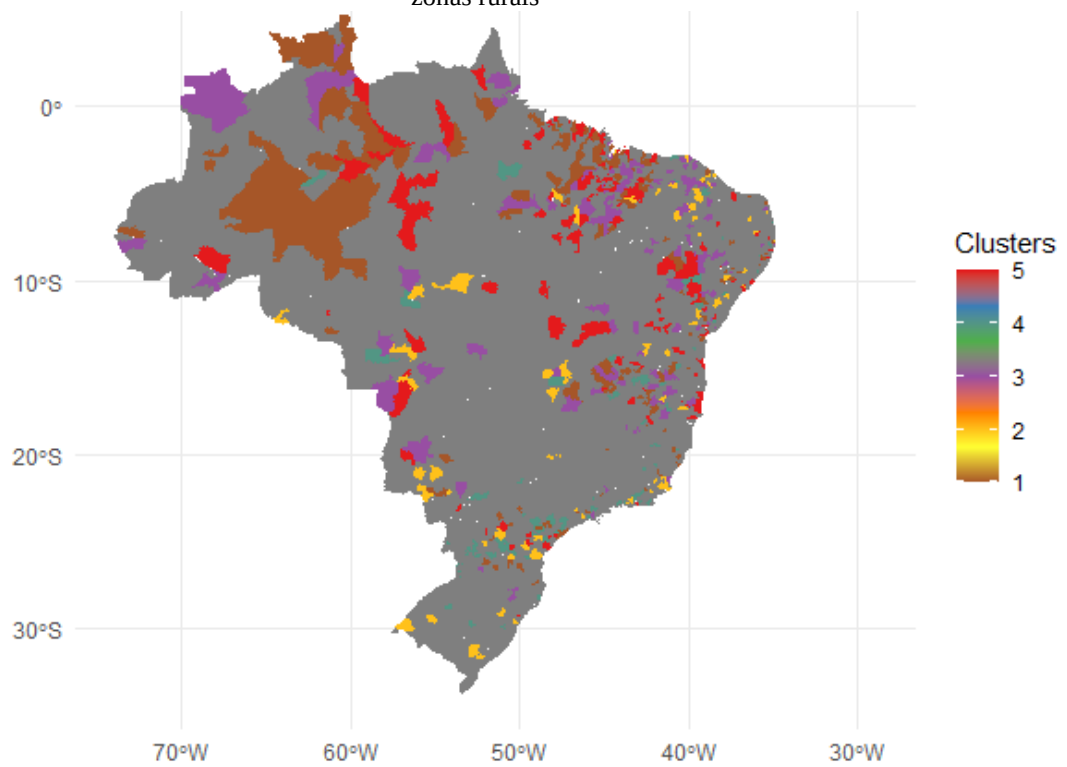
Fonte: Elaboração Própria

Figura 47: *Clusters* considerando indicadores dos docentes para o ensino médio em zonas rurais



Fonte: Elaboração Própria

Figura 48: *Clusters* considerando indicadores dos docentes e do desempenho dos alunos para o ensino médio em zonas rurais



Fonte: Elaboração Própria

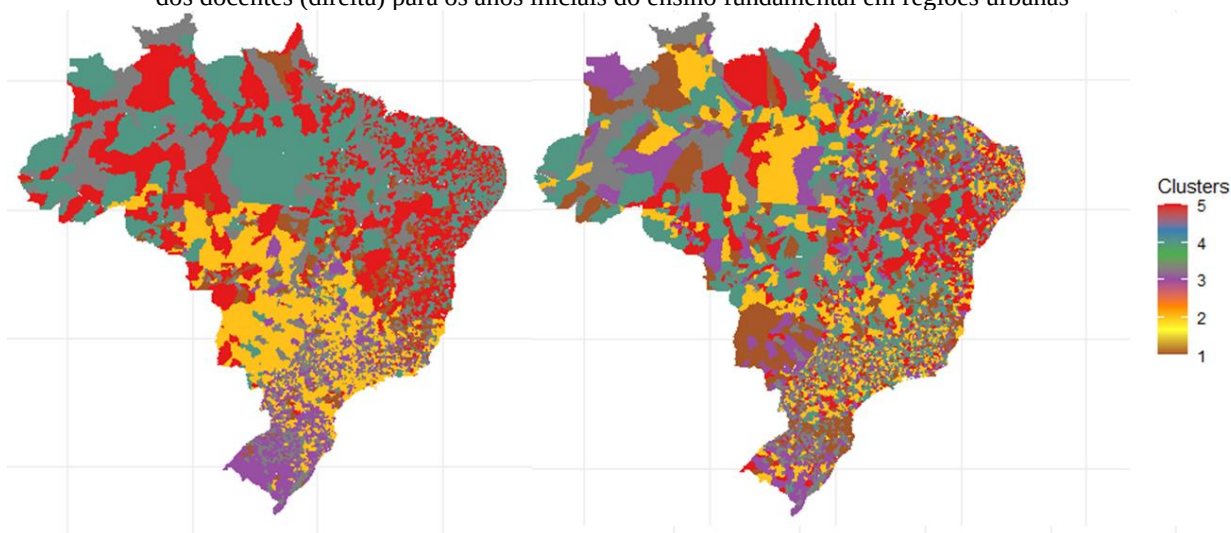
5 ANÁLISE DOS RESULTADOS E DISCUSSÃO

Neste capítulo, os *clusters* resultantes de cada cenário apresentado no Quadro 12 serão analisados em todas as situações apresentadas no Quadro 11, para que seja possível concluir semelhanças entre as situações dentro de um mesmo cenário, e assim sugerir medidas públicas condizentes com o objetivo deste trabalho no próximo capítulo.

5.1 CENÁRIO 1 – ANOS INICIAIS DO ENSINO FUNDAMENTAL – ZONA URBANA

5.1.1 Situação socioeconômico e estrutural com a situação dos docentes

Figura 49: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para os anos iniciais do ensino fundamental em regiões urbanas

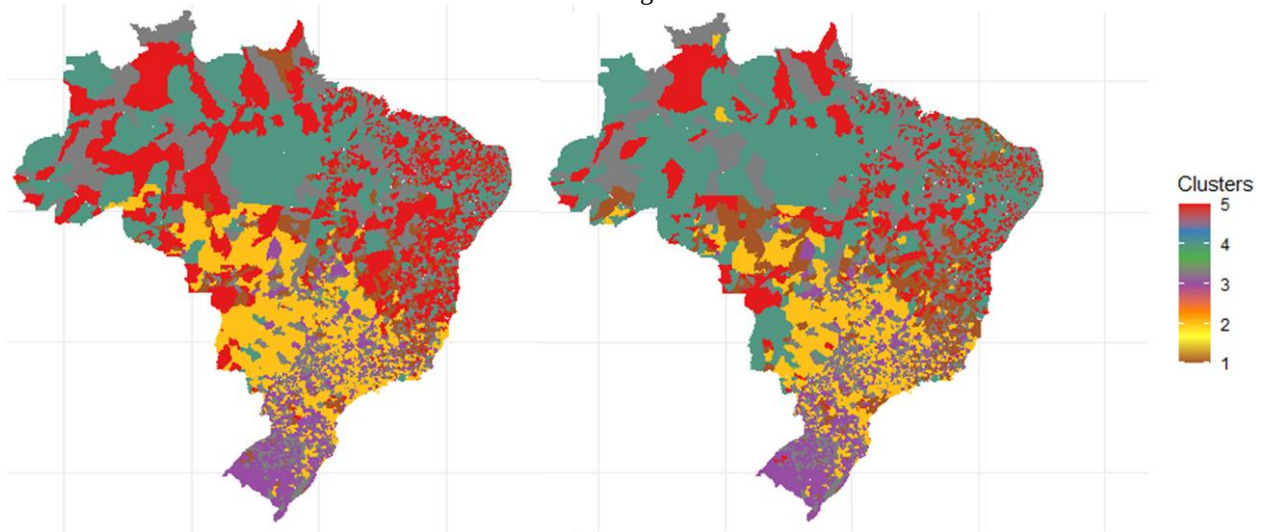


Fonte: Elaboração própria

Ao analisar a Figura 49, é possível perceber que poucos municípios compartilham o mesmo *cluster* nas duas combinações de indicadores realizadas. Mostrando que, na maioria dos casos, as condições socioeconômicas dos alunos e de infraestrutura das escolas não se relacionam diretamente com as condições dos docentes nessas escolas.

5.1.2 Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos

Figura 50: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões urbanas

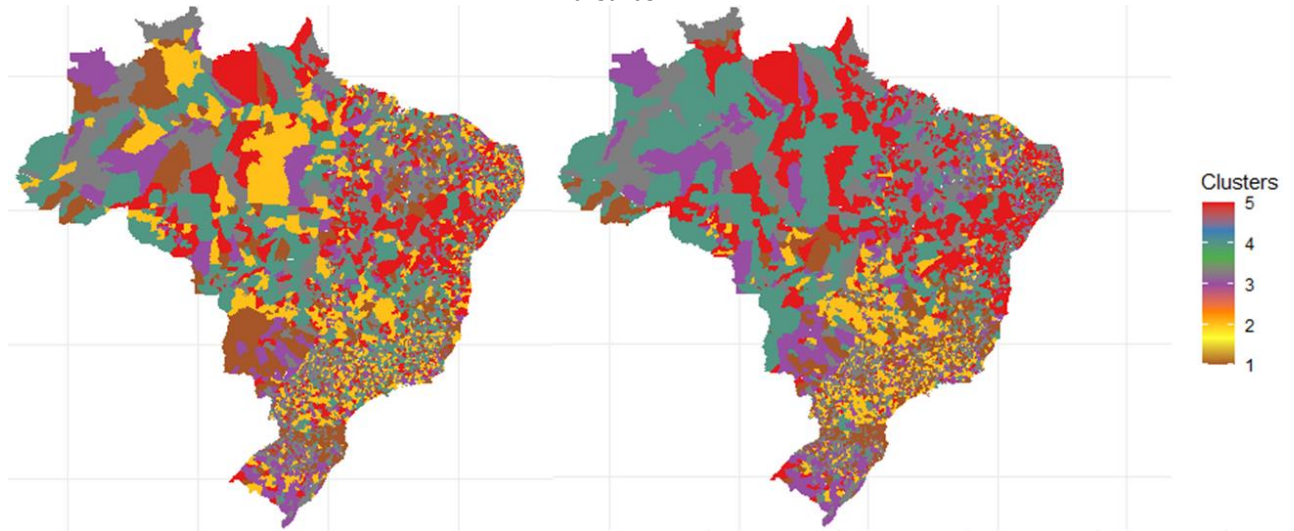


Fonte: Elaboração própria

As situações expostas na Figura 50 mostram que o desempenho dos alunos provavelmente está atrelado a sua condição socioeconômica e o número de alunos em sua turma. Porém, algumas exceções se mostram mais presentes nas regiões centro-oeste e norte.

5.1.3 Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos

Figura 51: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões urbanas



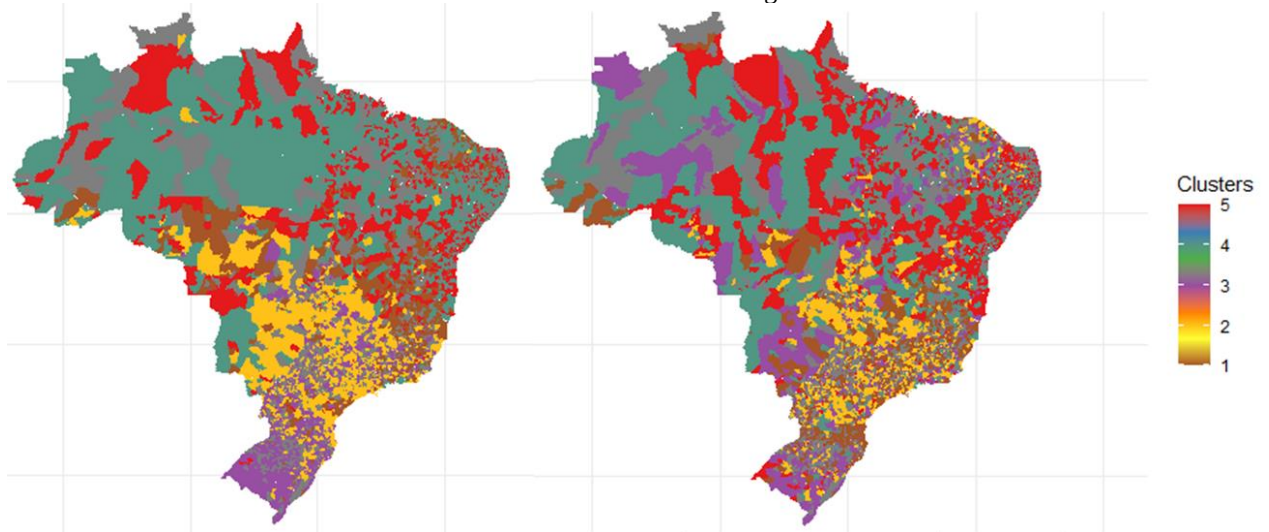
Fonte: Elaboração própria

A Figura 51 expõe que, principalmente nas regiões nordeste, sudeste e sul, existe relação entre as condições as quais os docentes estão submetidos e o nível de aprendizado dos alunos. As regiões centro-oeste e norte mais uma vez mostram uma maior heterogeneidade entre os *clusters* alocados a cada município.

Uma exceção dentro da região nordeste aparenta ser o estado do Ceará que mostra indicadores de resultado mais semelhantes aos da região sudeste que dos estados que está mais próximo.

5.1.4 Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos

Figura 52: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões urbanas



Fonte: Elaboração própria

A Figura 52 se assemelha bastante à Figura 49, representando que o desempenho dos alunos está atrelado ou a condições socioeconômicas dos alunos e de estrutura das escolas ou a condição dos docentes, e não relacionado a um conjunto desses fatores.

5.1.5 Todos os indicadores com as outras situações já analisadas

A Figura 15 reforça a ideia de que não existe uma relação clara entre condições socioeconômicas dos alunos, estrutura das escolas, condição dos docentes e desempenho dos alunos em conjunto.

5.1.6 Recomendações para o cenário

Após analisar as comparações feitas dentro do cenário 1, é possível perceber que condições socioeconômicas e de número de alunos por turma devem ser tratados de forma separada ao de condição dos docentes para aumentar o desempenho dos alunos.

Nas regiões nordeste, sudeste, e sul, parece existir uma relação entre desempenho dos alunos com sua condição socioeconômica e o número de alunos por turma, assim como existe uma relação entre o desempenho dos alunos e a condição dos docentes.

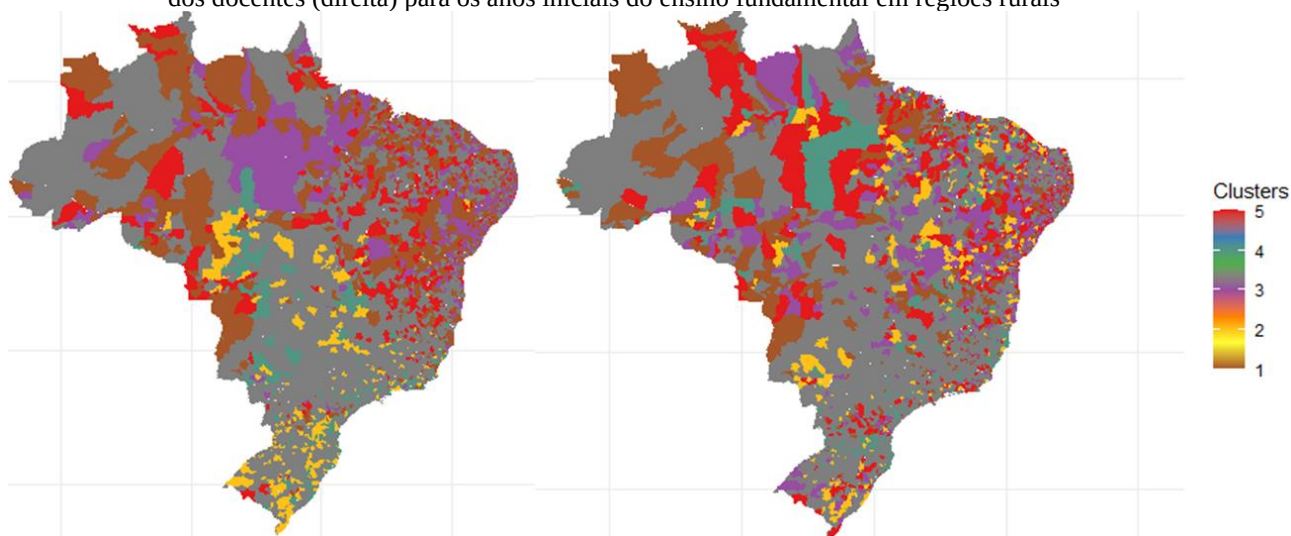
No caso dos estados do centro-oeste e do norte, nenhuma das combinações parece afetar diretamente o desempenho dos alunos, sendo necessário aprofundar a análise de quais

fatores, ou combinação de fatores, podem gerar maior impacto no desempenho dos alunos nessas regiões.

5.2 CENÁRIO 2 – ANOS INICIAIS DO ENSINO FUNDAMENTAL – ZONA RURAL

5.2.1 Situação socioeconômico e estrutural com a situação dos docentes

Figura 53: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para os anos iniciais do ensino fundamental em regiões rurais

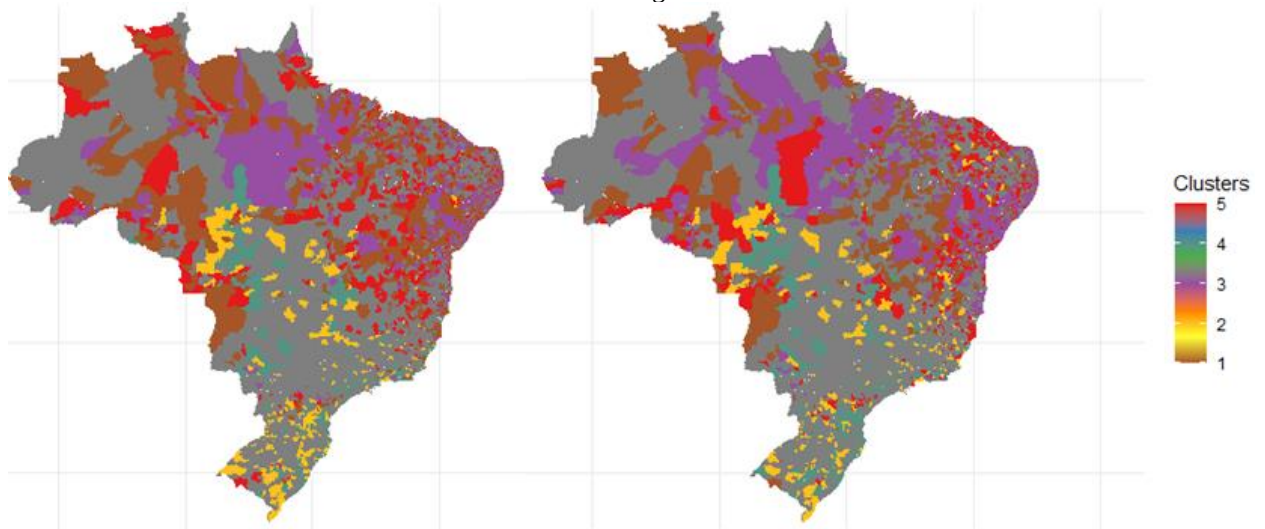


Fonte: Elaboração própria

As situações expostas pela Figura 53 mostram que, assim como nas localidades urbanas, nos anos iniciais do ensino fundamental não existe uma relação clara entre condição socioeconômica dos alunos e número de alunos por turma com a condição dos docentes em um mesmo município.

5.2.2 Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos

Figura 54: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões rurais

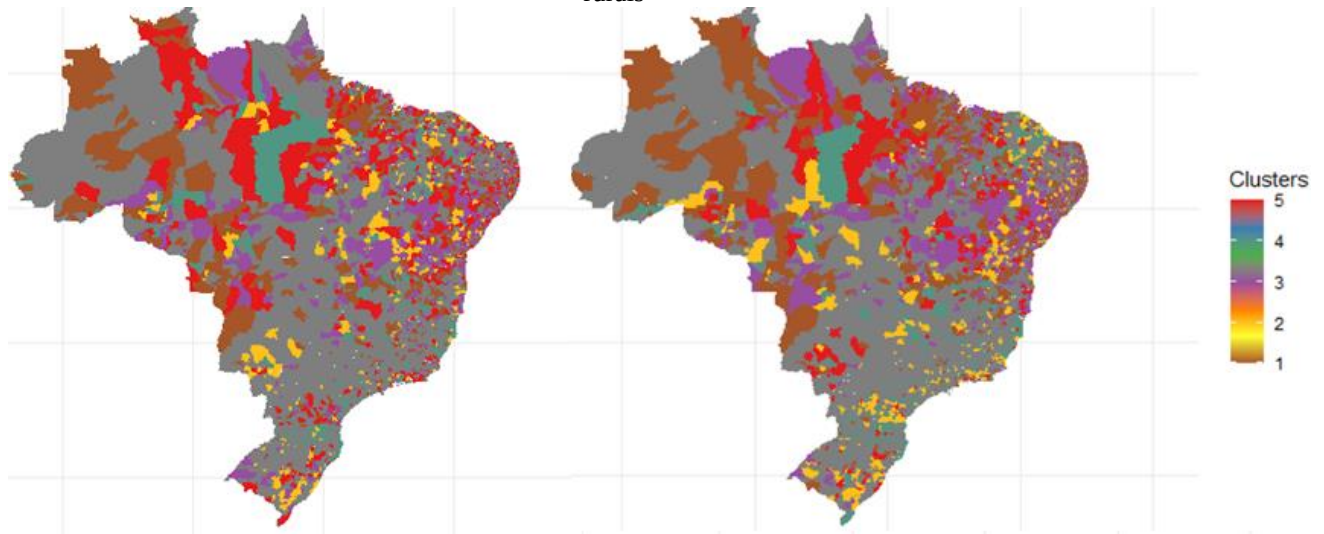


Fonte: Elaboração própria

Analisando a Figura 54 é possível observar que o desempenho dos alunos e sua condição socioeconômica, junto ao número de alunos por turma, se mostra bastante relevante nas regiões centro-oeste, sudeste, sul e em partes da região nordeste. Porém, a região norte e parte da região nordeste não mostra uma relação significativa entre os fatores citados.

5.2.3 Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos

Figura 55: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões rurais

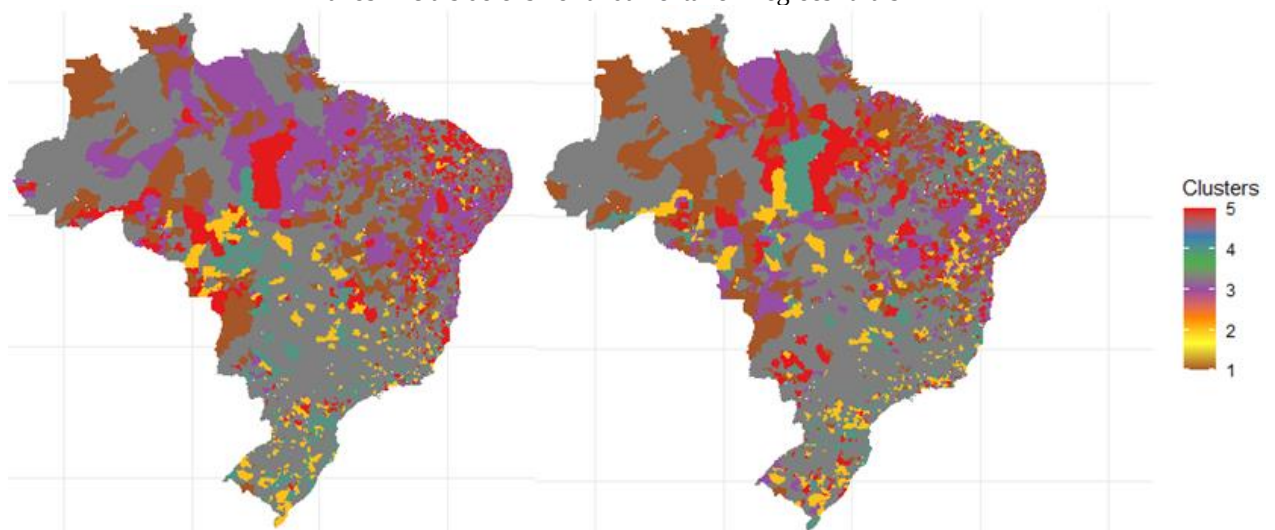


Fonte: Elaboração própria

A Figura 55 mostra que o desempenho dos alunos se relaciona de forma mais significativa a condição dos docentes na região norte. Enquanto nas demais regiões do país essa relação aparece de forma mais pontual, como é o caso do estado do Ceará, Rio Grande do Sul, norte do estado de Minas Gerais, e oeste do Estado da Bahia.

5.2.4 Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos

Figura 56: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos iniciais do ensino fundamental em regiões rurais



Fonte: Elaboração própria

Ao analisar a Figura 56 é possível perceber que não existe uma relação clara entre as duas combinações de indicadores, o que corrobora com a ideia de que o desempenho dos alunos não pode ser observado a partir de uma visão unificada sobre os fatores que compõe esses indicadores.

5.2.5 Todos os indicadores com as outras situações já analisadas

Comparando a situação em que todos os indicadores estão sendo analisados com as demais combinações entre indicadores, torna-se visível a heterogeneidade presente nesse cenário, e a independência entre os fatores quando estes não estão relacionados ao desempenho dos alunos.

5.2.6 Recomendações para o cenário

Observando as comparações realizadas no cenário 2, é possível visualizar algumas relações entre os indicadores estudados e o desempenho dos alunos. Mais uma vez, indicadores de condição socioeconômica dos alunos e de número de alunos por turma devem ser analisados de forma distinta ao de condição dos docentes.

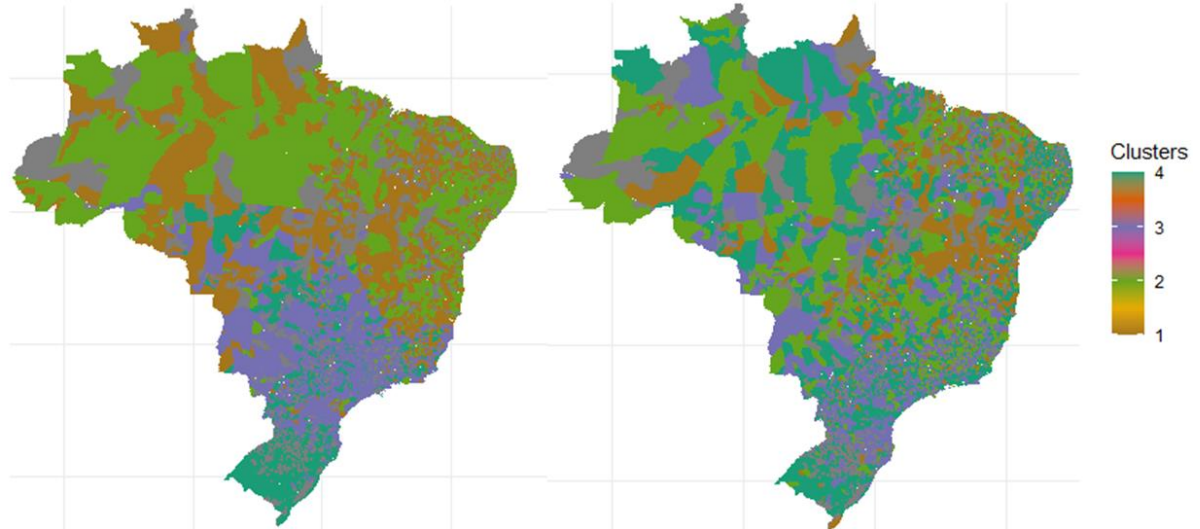
Aparentemente, nas regiões centro-oeste, sudeste, sul, e no estado do Pará parece existir uma relação entre indicadores socioeconômicos e estruturais com o desempenho dos alunos, enquanto na região norte a condição dos docentes aparenta se relacionar com o desempenho dos alunos.

A região nordeste não apresenta uma relação clara com quaisquer indicadores apresentados, sendo necessário aprofundar a análise de quais fatores gerariam maior impacto na região. Ademais, uma análise mais aprofundada sobre o impacto dos docentes sobre o desempenho dos alunos nas regiões centro-oeste, sudeste e sul, seria bastante produtivo. De forma semelhante, um outro estudo sobre o impacto socioeconômico e de número de alunos na região norte também seria recomendável.

5.3 CENÁRIO 3 – ANOS FINAIS DO ENSINO FUNDAMENTAL – ZONA URBANA

5.3.1 Situação socioeconômico e estrutural com a situação dos docentes

Figura 57: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para os anos finais do ensino fundamental em regiões urbanas

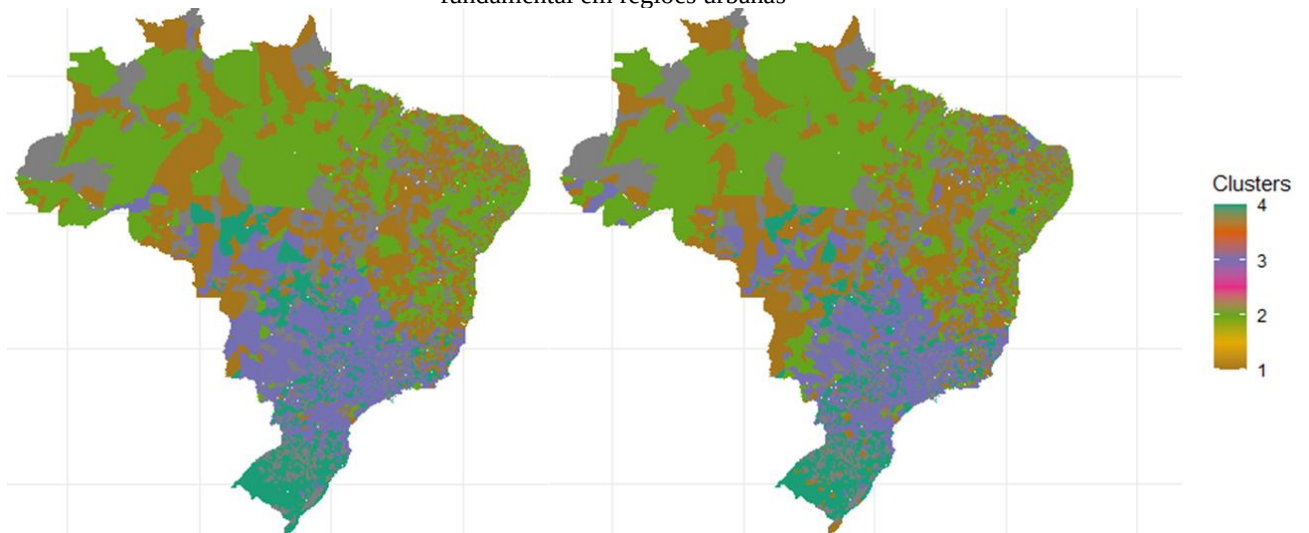


Fonte: Elaboração própria

A Figura 57 apresenta poucas semelhanças entre os indicadores apresentados, mostrando que existe pouca relação entre as condições socioeconômicas dos alunos e de número de alunos por turma com a condição dos docentes dentro de um município.

5.3.2 Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos

Figura 58: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões urbanas

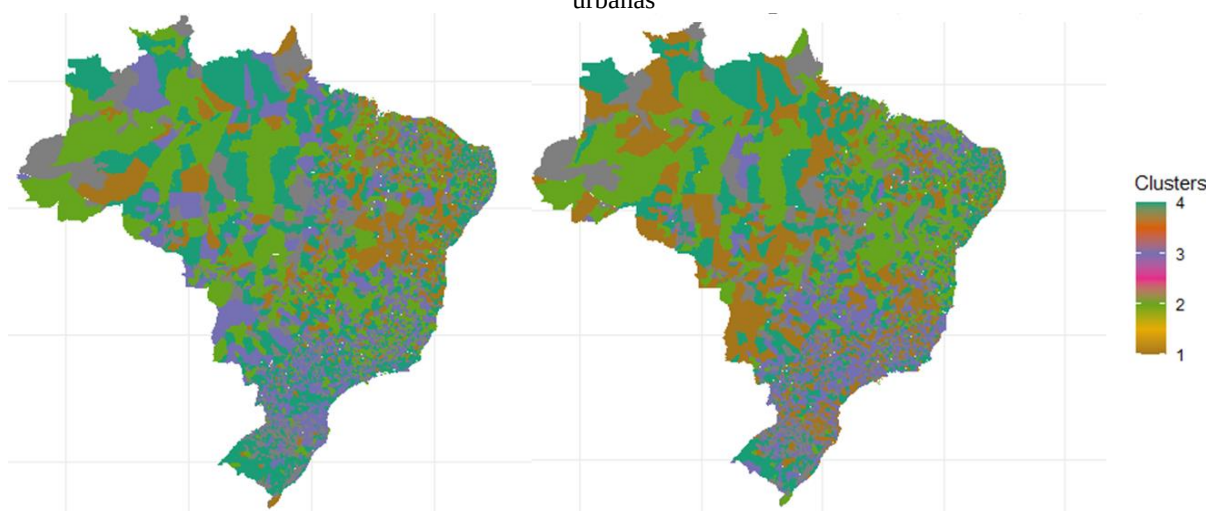


Fonte: Elaboração própria

Ao analisar a Figura 58 é possível perceber uma forte relação entre o desempenho dos alunos com a sua condição socioeconômica e o número de alunos presentes na sua turma. Porém, alguns estados se destacam pela diferença entre os *clusters* os quais seus municípios estão designados nas duas condições, sendo eles: o Mato Grosso do Sul, que se assemelha bastante a região sudeste; e partes do estado do Acre, que se assemelha a região sudeste; e partes do Amazonas que se assemelha mais a partes predominantes nos estados do Tocantins e do Maranhão.

5.3.3 Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos

Figura 59: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões urbanas

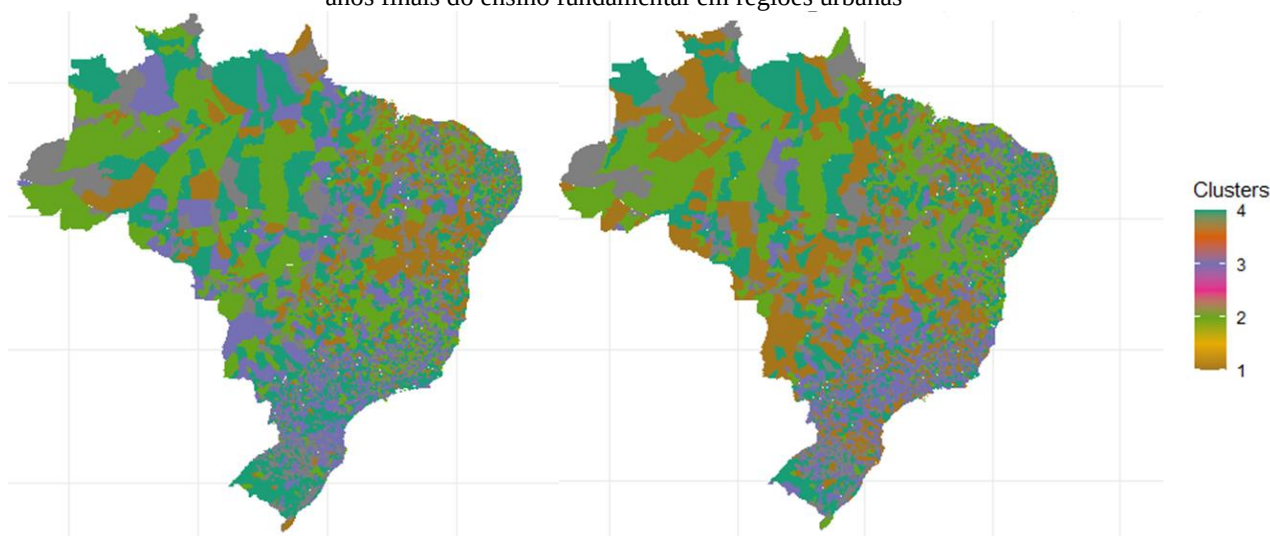


Fonte: Elaboração própria

Os indicadores de condição dos docentes e de condição dos docentes levando em consideração o desempenho dos alunos não apresenta grande semelhanças quando se analisa a Figura 59. Algumas exceções podem ser observadas principalmente na região Nordeste e Norte, e em parte da região Sudeste.

5.3.4 Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos

Figura 60: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões urbanas



Fonte: Elaboração própria

A Figura 60 apresenta poucas semelhanças entre os *clusters* presentes nas duas combinações de indicadores, mostrando mais uma vez que não é possível estabelecer uma boa relação que consiga unificar o desempenho dos alunos a sua condição socioeconômica, o número de alunos de sua turma, e a condição dos docentes sob uma perspectiva únicas.

5.3.5 Todos os indicadores com as outras situações já analisadas

Ao comparar a situação em que todos os indicadores são levados em consideração com os demais, é possível perceber que não existe uma relação clara entre os indicadores de condição socioeconômica, do número de alunos por turma, de condições dos docentes, e do desempenho dos alunos. Sendo necessário analisar cada situação a partir de uma perspectiva distinta.

5.3.6 Recomendações para o cenário

É possível perceber que nesse cenário em quase todos os estados existe uma forte relação entre o desempenho dos alunos com sua condição socioeconômica, e o número de alunos em sua turma.

A relação entre desempenho dos alunos e a condição dos docentes, é bem menos delimitada, sendo necessário compreender quais fatores podem ser mais impactantes para o desempenho do aluno. Algumas regiões se mostraram exceções e seria interessante

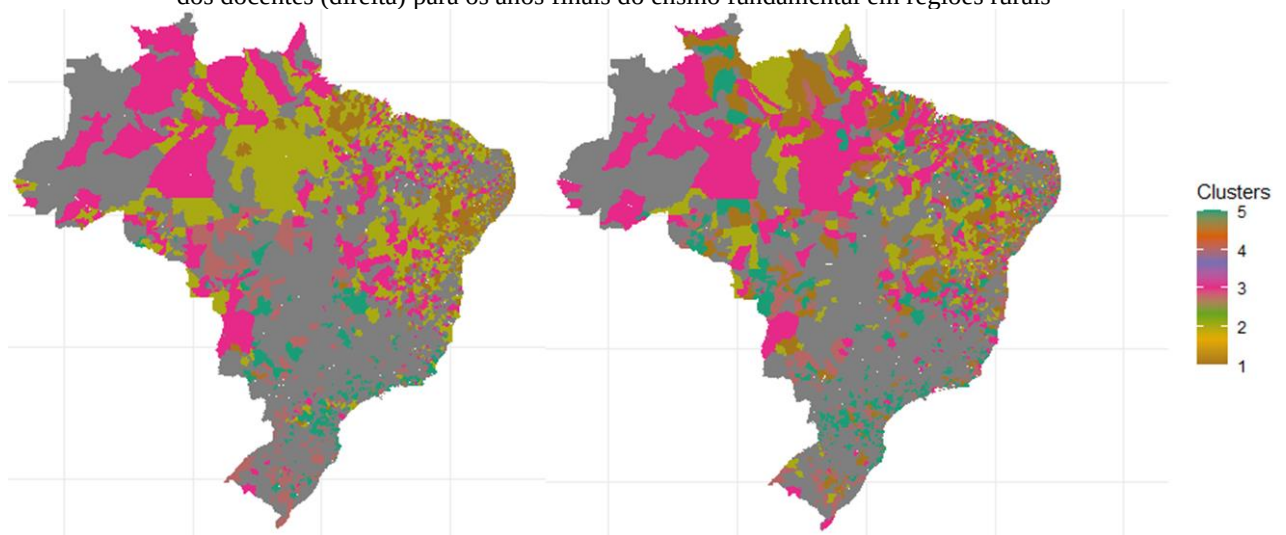
compreender o porquê dessa relação ser mais forte nesses locais, para então compreender qual, ou quais, fatores também deveriam ser analisados.

Em alguns estados, algumas de suas regiões se mostraram exceções em todas as análises como é o caso de parte do estado Acre, do Amazonas, e do Mato Grosso do Sul. Mostrando que estudos sobre essas regiões poderiam colaborar para o compreender o impacto de outros indicadores.

5.4 CENÁRIO 4 – ANOS FINAIS DO ENSINO FUNDAMENTAL – ZONA RURAL

5.4.1 Situação socioeconômico e estrutural com a situação dos docentes

Figura 61: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para os anos finais do ensino fundamental em regiões rurais



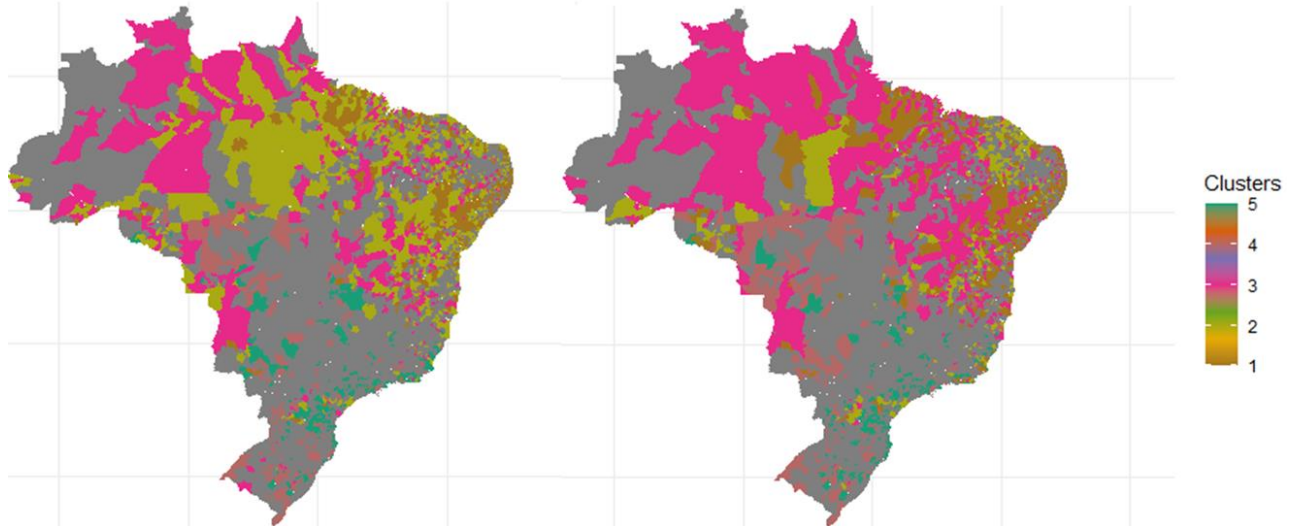
Fonte: Elaboração própria

A Figura 61 mostra diversas semelhanças nas regiões centro-oeste, sudeste e sul, podendo indicar que fatores socioeconômicos e estruturais poderiam ser trabalhados em conjunto a fatores de condição dos docentes nessas regiões.

Nas regiões norte e nordeste existe uma maior diferença na alocação de *clusters*, mostrando que esses fatores não estão fortemente relacionados.

5.4.2 Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos

Figura 62: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões rurais

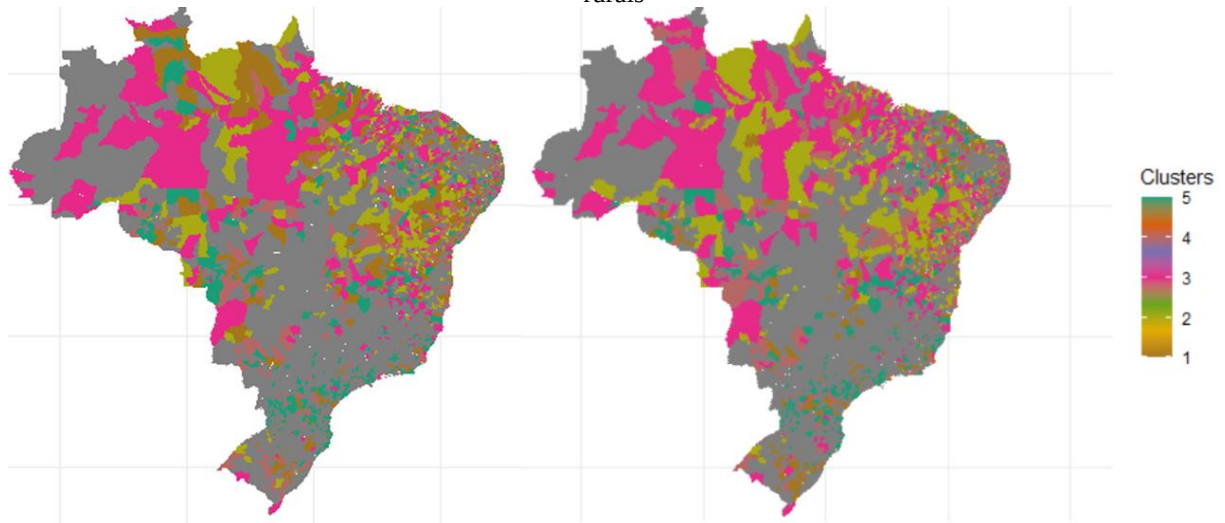


Fonte: Elaboração própria

As regiões centro-oestes, sudeste, e sul apresentam uma relação bem delimitada entre condições socioeconômicas dos alunos e de número de alunos por turma com o desempenho desses alunos. Enquanto as regiões nordeste e norte, e parte da região norte do estado de Minas Gerais, mostram que essa relação não está bem delimitada nesses locais.

5.4.3 Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos

Figura 63: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões rurais



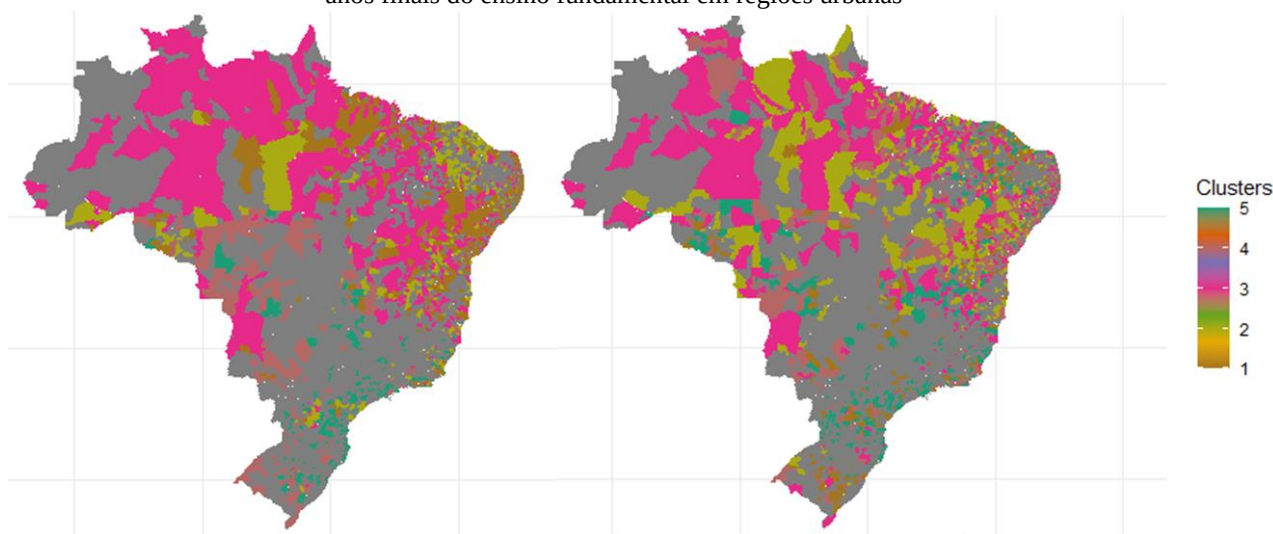
Fonte: Elaboração própria

Os conjuntos de *clusters* apresentados na Figura 63 mostram que existe uma relação entre o desempenho dos alunos com a condição dos docentes nas regiões centro-oeste, nordeste, sudeste, sul, e em partes da região norte.

Em alguns estados da região norte como Amapá e Roraima a relação entre desempenho dos alunos e condição dos docentes não está bem delimitada.

5.4.4 Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos

Figura 64: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para os anos finais do ensino fundamental em regiões urbanas



Fonte: Elaboração própria

Ao analisar a Figura 64 é possível perceber que as regiões centro-oeste, sudeste e sul apresentam diversos municípios com *clusters* coincidentes. Enquanto as regiões nordeste e norte apresentam maiores diferenças entre os *clusters* apresentados.

5.4.5 Todos os indicadores a partir das comparações já realizadas

A Figura 32 se assemelha as demais comparações realizadas em que as regiões norte e nordeste apresentam *clusters* distintos para o mesmo município, enquanto os municípios das regiões centro-oeste, sudeste, e sul tendem a continuar no mesmo *cluster* ao longo de todas as análises.

5.4.6 Recomendações para o cenário

No cenário 4 é possível perceber que para as regiões centro-oeste, sudeste e sul existe uma clara relação entre os indicadores, permitindo que medidas públicas possam ser tomadas levando em consideração todos esses fatores em conjunto a fim de aumentar o desempenho dos alunos.

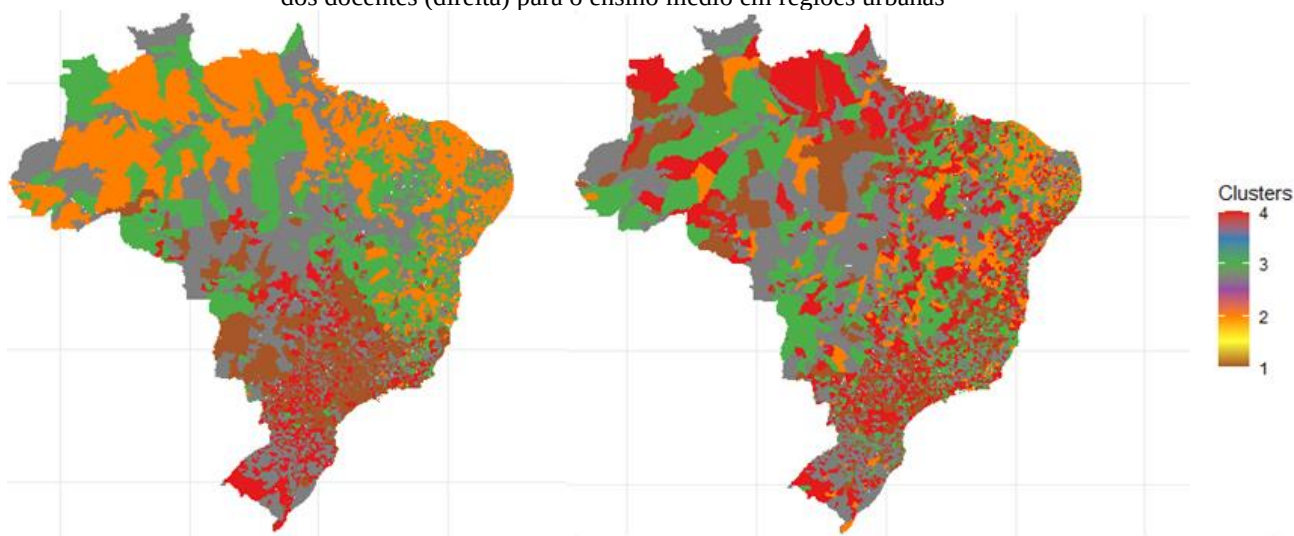
As regiões nordeste e norte apresentam relações mais fortes apenas entre desempenho dos alunos e condição dos docentes, sendo necessário estudar outros fatores que sejam mais

impactantes no desempenho dos alunos considerando aspectos socioeconômicos e estruturais das escolas.

5.5 CENÁRIO 5 – ENSINO MÉDIO – ZONA URBANA

5.5.1 Situação socioeconômico e estrutural com a situação dos docentes

Figura 65: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para o ensino médio em regiões urbanas

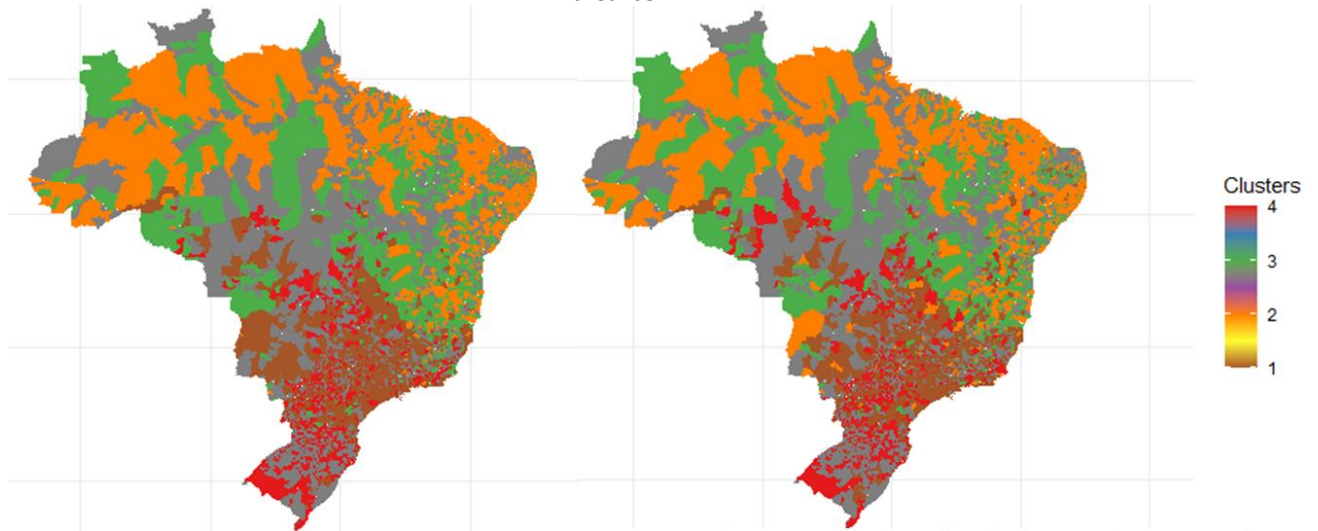


Fonte: Elaboração própria

Poucas relações entre as duas situações expostas podem ser obtidas a partir de uma análise da Figura 65. Dessa forma é possível perceber que a condição dos docentes, no geral, não se relaciona com a condição socioeconômica dos alunos e com quantidade de alunos por sala para uma mesma escola.

5.5.2 Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos

Figura 66: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para o ensino médio em regiões urbanas

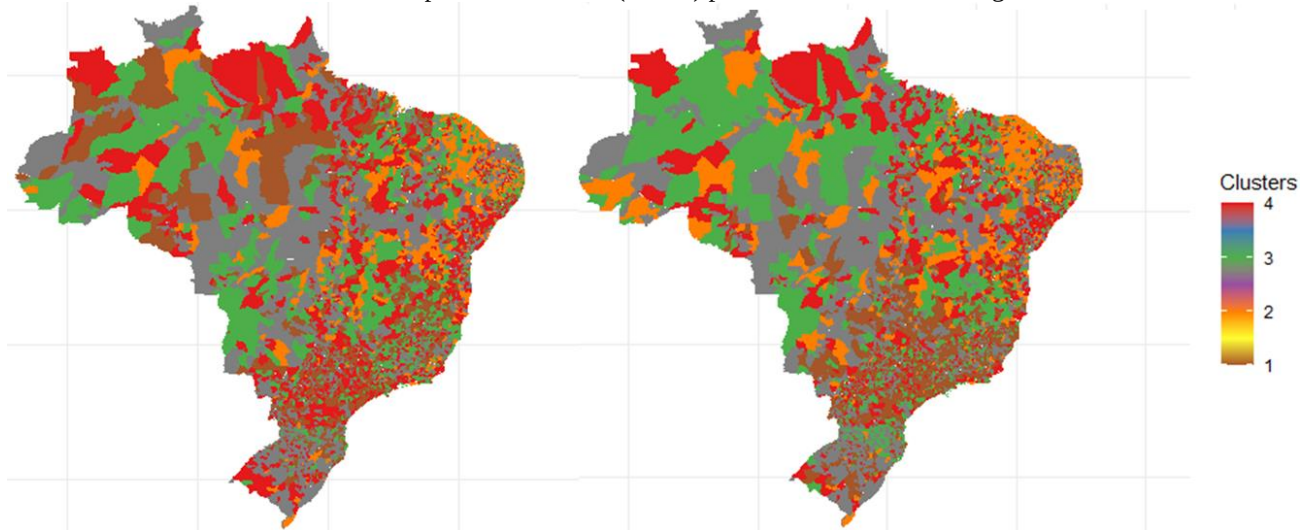


Fonte: Elaboração própria

A Figura 66 apresenta diversas semelhanças entre os dois conjuntos de *clusters* apresentados, mostrando que provavelmente existe uma forte relação entre o desempenho dos alunos com sua condição socioeconômica e o número de alunos em sua turma.

5.5.3 Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos

Figura 67: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para o ensino médio em regiões urbanas



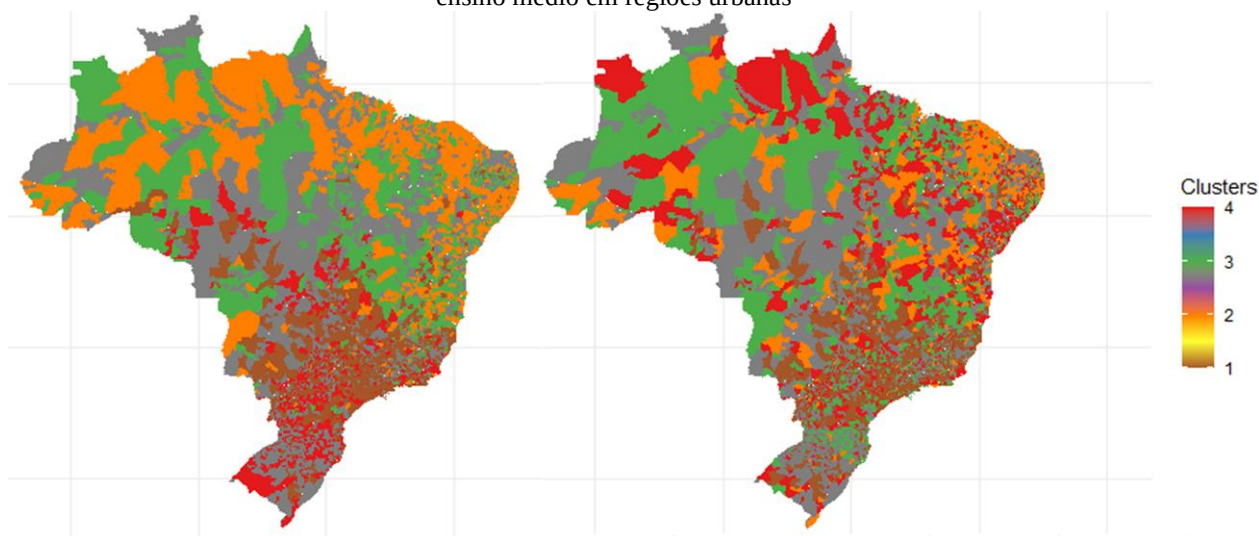
Fonte: Elaboração própria

Ao analisar a Figura 67 percebe-se que existe uma relação entre indicadores de desempenho dos alunos e condição dos docentes principalmente nas regiões centro-oeste, nordeste e sudeste.

As regiões norte e sul apresentam uma menor consistência entre os *clusters* designados a cada município nas duas condições apresentadas.

5.5.4 Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos

Figura 68: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para o ensino médio em regiões urbanas



Fonte: Elaboração própria

Ao comparar os casos apresentados pelas Figura 68, é possível perceber que os *clusters* 1 e 3 são semelhantes nos dois casos, porém os demais *clusters* sugeridos apresentam diversas diferenças em todas as regiões do país, especialmente na região Nordeste.

5.5.5 Todos os indicadores a partir das comparações já realizadas

A Figura 39 expõe mais uma vez a diferença entre as condições analisadas, mostrando que não existe uma relação clara entre indicadores socioeconômicos, estruturais, de condição dos docentes e de desempenho dos alunos que possa unificar todos esses fatores sob uma perspectiva única.

5.5.6 Recomendações para o cenário

Analisando o quinto cenário presente no Quadro 12 é possível perceber mais uma vez que indicadores socioeconômicos e de número de alunos por turma não devem ser analisados com indicadores de condição dos docentes de forma unificada para todas as regiões ou *clusters*. Porém, um estudo sobre os *clusters* 1 e 3 os quais vários municípios permaneceram sob diferentes análises nos mesmos *clusters* seria bastante promissor.

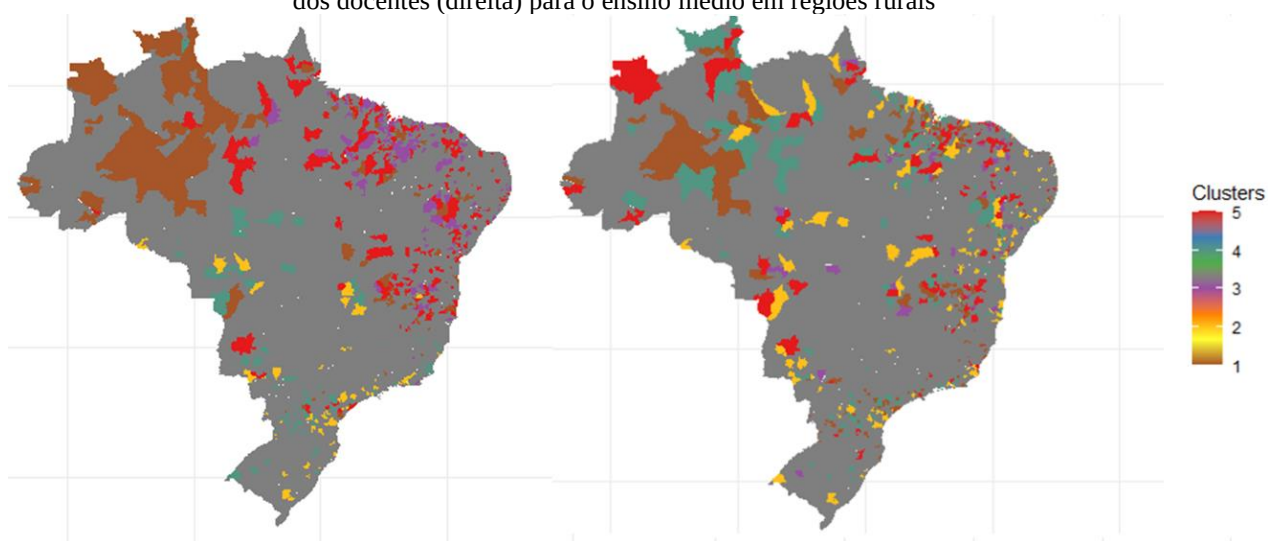
Outra conclusão que pode ser tomada é uma relação entre desempenho dos alunos com sua condição socioeconômica e o número de alunos em sua turma. Enquanto, os indicadores de

condição dos docentes apresentaram uma relação menos delimitada com o desempenho dos alunos, o que torna recomendável estudos mais aprofundados sobre outro fator, ou fatores, que seriam mais relevantes.

5.6 CENÁRIO 6 – ENSINO MÉDIO – ZONA RURAL

5.6.1 Situação socioeconômico e estrutural com a situação dos docentes

Figura 69: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores de condição dos docentes (direita) para o ensino médio em regiões rurais

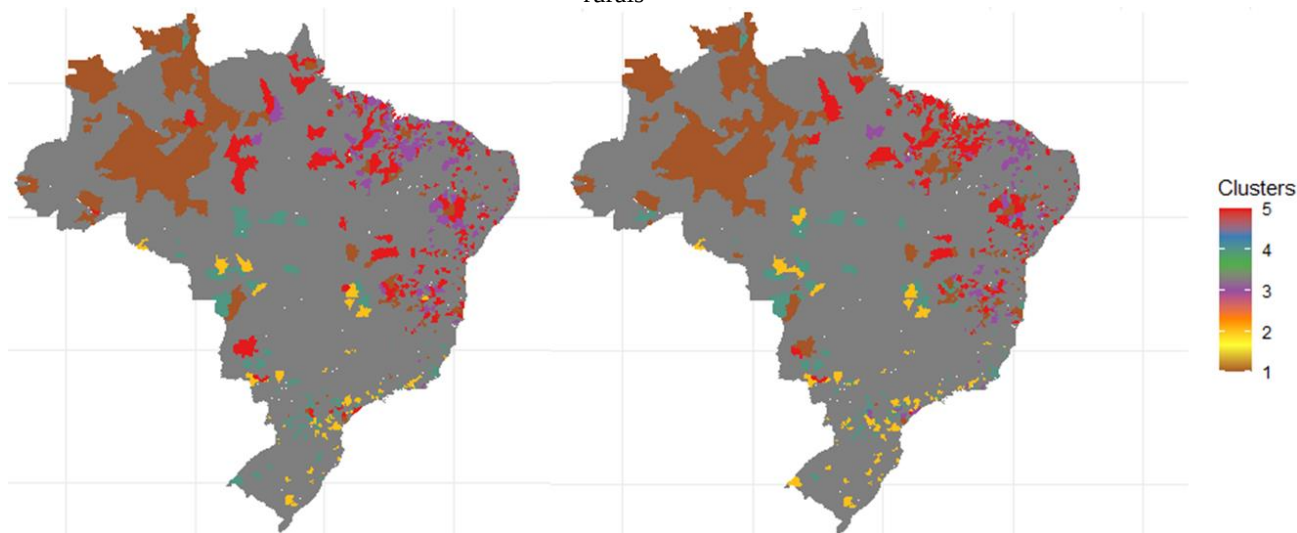


Fonte: Elaboração própria

Analisando os casos apresentados pela Figura 69, foi possível perceber poucas semelhanças entre as duas situações expostas, mostrando que indicadores socioeconômicos dos alunos e estruturais possuem pouca relação com a condição de docentes considerando uma mesma escola.

5.6.2 Situação socioeconômico e estrutural com a socioeconômico e estrutural considerando o desempenho dos alunos

Figura 70: Comparação entre indicadores socioeconômicos e estruturais (esquerda) com indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (direita) para o ensino médio em regiões rurais

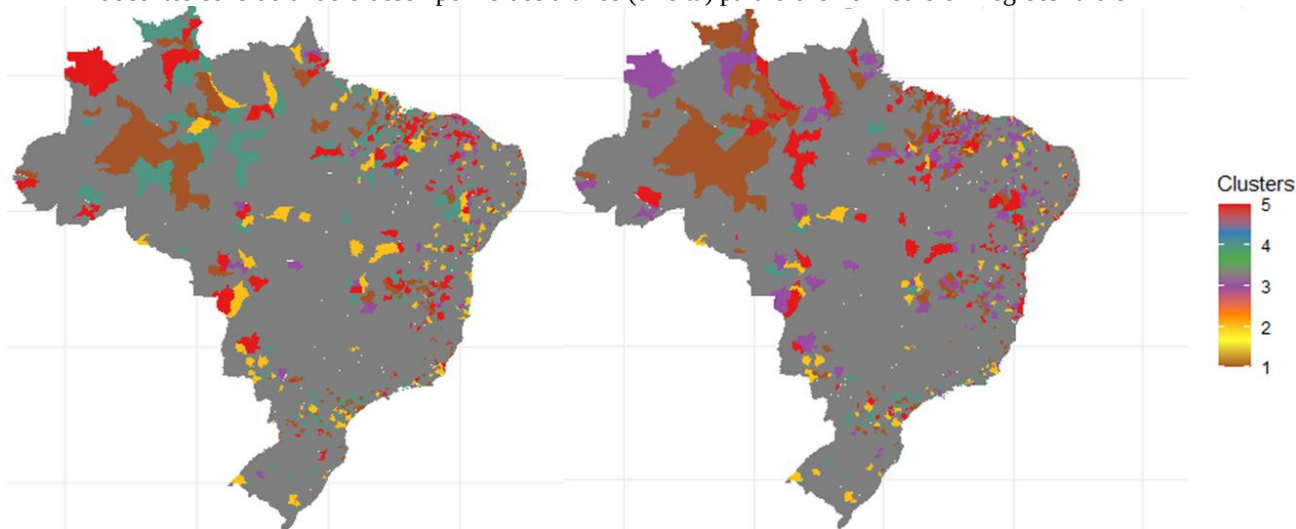


Fonte: Elaboração própria

Ao comparar os dois casos apresentados pelas Figura 70 diversas semelhanças podem ser encontradas entre os *clusters* designados, com uma diferença maior na alocação de *clusters* dentro do estado do Maranhão, e em outras partes do semiárido nordestino.

5.6.3 Situação dos docentes com a situação dos docentes considerando o desempenho dos alunos

Figura 71: Comparação entre indicadores de condição dos docentes (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para o ensino médio em regiões rurais

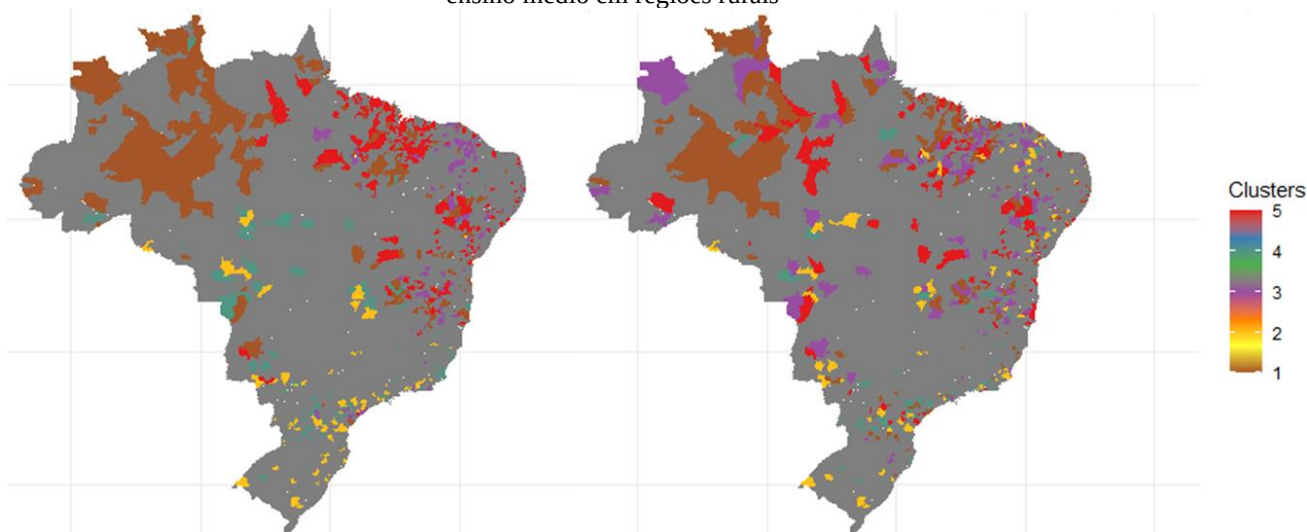


Fonte: Elaboração própria

Considerando os dois casos apresentados pela Figura 71, é possível perceber várias sobreposições nas regiões sudeste e sul, porém nas demais regiões poucas semelhanças podem ser identificadas.

5.6.4 Situação socioeconômico e estrutural considerando o desempenho dos alunos com a situação dos docentes considerando o desempenho dos alunos

Figura 72: Comparação entre indicadores socioeconômicos e estruturais considerando o desempenho dos alunos (esquerda) com indicadores de condição dos docentes considerando o desempenho dos alunos (direita) para o ensino médio em regiões rurais



Fonte: Elaboração própria

Poucas semelhanças podem ser encontradas ao analisar a Figura 72 mostrando, assim como a Figura 69, que não existe uma relação bem delimitada entre indicadores socioeconômicos e estruturais com indicadores de condição dos docentes.

As poucas coincidências entre os *clusters* apresentados não estão concentradas em regiões inteiras, mas estão dispersas e pontualmente presentes em algumas regiões do país como: estado do Pará, semiárido nordestino, parte do estado do Maranhão e do Piauí, norte do estado de Minas Gerais e sul do estado da Bahia, e sul do estado do Paraná e norte do Estado de Santa Catarina.

5.6.5 Todos os indicadores a partir das comparações já realizadas

Ao comparar as demais situações analisadas com Figura 44, nota-se que existe pouca relação entre os indicadores de condição docente com os indicadores de condição socioeconômica dos alunos e número de alunos por turma.

5.6.6 Sugestões para o cenário

Ao analisar o cenário 6 é possível concluir que não existe uma relação clara entre desempenho dos alunos, condição dos docentes, número de alunos por turma e condição socioeconômica dos alunos.

Porém, existe uma relação entre desempenho dos alunos com condição socioeconômica dos alunos e número de alunos por turma em quase todas as regiões. A relação entre condição dos docentes e desempenho dos alunos se mostra mais acentuada nas regiões sudeste e sul.

Nas regiões centro-oeste, nordeste e norte seria interessante realizar outros estudos que permitam aumentar a compreensão sobre qual outro fator, ou fatores, podem estar relacionados ao desempenho dos alunos e condição dos docentes que não foram analisados.

6 CONCLUSÃO

Este trabalho apresenta uma análise geral da situação do ensino básico no Brasil, que corresponde as turmas que vão desde o 1º ano do ensino fundamental até a 3ª série do ensino médio. Sendo que, em 2019, aproximadamente 34 milhões de brasileiros frequentavam escolas que ofereciam essa etapa de ensino.

Dentro do ensino básico no Brasil, este trabalho deu ênfase na análise de alunos que frequentavam instituições públicas, uma vez existe uma diferença significativa entre o desempenho dos alunos provenientes de instituições públicas e privadas segundo o Pisa. Além disso, entender melhor a situação da educação pública auxilia na compreensão da realidade de mais de 83% dos alunos do país.

O escopo deste trabalho se restringiu aos alunos da educação básica uma vez que para alunos do ensino infantil existem poucos dados disponíveis, e um ensino superior de qualidade está condicionado a um ensino básico de qualidade.

Após definir qual setor da educação atuar, foi necessário definir técnicas que possibilitassem a compreensão da conjuntura das escolas do país, para essa finalidade foram utilizadas adaptações de indicadores presentes em sites governamentais, junto a métodos estatísticos e de programação que tornassem possível analisar a grande quantidade de dados disponível.

Por fim, a partir de uma análise dos mapas gerados a partir dos métodos estatísticos foi possível perceber quais regiões, ou grupos de municípios, poderiam definir medidas públicas levando em consideração fatores socioeconômicos, de estrutura das escolas e de condição dos docentes em conjunto, ou se cada um desses fatores deveria ser tratado de forma separada. Outra opção levantada por este trabalho é que outros fatores externos aos analisados poderiam ser relevantes, e essa conclusão também foi alcançada algumas vezes.

Neste trabalho, foi possível perceber a heterogeneidade das condições educacionais do Brasil e que, resultados semelhantes podem ser observados entre regiões que possuem grandes diferenças socioeconômicas ou estão distantes por milhares de quilômetros. Mostrando a necessidade de uma visão mais holística de agentes municipais, estaduais e federais, para que a educação pública nacional passe a apresentar resultados mais concretos no futuro.

REFERÊNCIAS

- ALVES, M. T. G.; SOARES, J. F. Medidas de nível socioeconômico em pesquisas sociais: uma aplicação aos dados de uma pesquisa educacional. *Opinião Pública*, Campinas, v. 15, n. 1, p. 1-30, jun. 2009.
- ALVES, M. T.; SOARES, J. F.; XAVIER, F. P.; Índice socioeconômico das escolas de educação básica brasileiras. *Ensaio: Avaliação de Políticas Públicas Educacionais*, Rio de Janeiro, v. 22, n. 84, p. 671-704, jul/set. 2014.
- COSTA NETO, Pedro Luiz de Oliveira (2002). *Estatística*
- DEVORE, Jay L. (2010) *Probability & Statistics for Engineering and the Sciences* 8th ed
- IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Pesquisa Nacional por Amostra de Domicílios Contínua. Brasília, DF, 2019.
- INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Censo do Ensino Superior, Brasília, 2019.
- INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Censo Escolar, Brasília, 2020.
- INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Portal INEP, Brasília, DF, 2022. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/pisa>
- INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Portal INEP: SAEB – Sistema de Avaliação da Educação Básica, Brasília, DF, 2022b. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb>
- MANN, PREM S. (2010). *Introductory Statistics* 7^a ed. [S.l.]: Wiley.
- MINISTÉRIO DA CASA CIVIL. Lei nº 9.394 – Diretrizes e Bases da educação nacional. Brasília, DF, 20 de dezembro de 1996.
- MINISTÉRIO DA EDUCAÇÃO. Diretrizes Curriculares Nacionais para o Ensino Médio (DCENEM): Resolução nº 3, de 21 de novembro de 2018. Brasília, DF, 2018. Disponível em: <http://portal.mec.gov.br/docman/novembro-2018-pdf/102481-rceb003-18/file#:~:text=2%C2%BA%20As%20Diretrizes%20Curriculares%20Nacionais,Federal%20e%20dos%20Munic%C3%ADpios%20na>
- MINISTÉRIO DA EDUCAÇÃO. Base Nacional Comum Curricular. Brasília, DF, 2015. Disponível em: http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_versaofinal_site.pdf

MONTGOMERY, D. C. Design and analysis of experiments. 2nd ed. New York: John Wiley & Sons, 1984.

MORISSETTE, Laurence & Chartier, Sylvain. (2013). The k-means clustering technique: General considerations and implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology*. 9. 15–24. 10.20982/tqmp.09.1.p015.

PROVOST, Foster; FAWCETT, Tom. **Data Science for Business: What you need to know about data mining and data-analytic thinking**. " O'Reilly Media, Inc.", 2013.

R PROJETO. Disponível em: <https://www.r-project.org/>

SIRIN, S. R; Socioeconomic status and academic achievement: a meta-analytic review of research. *Review of Educational Research*, [S.l.], v. 75, n. 3, p. 417-453, 2005.