

Diretoria de Estatística e Informações (Direi)

**NOTA TÉCNICA:
ÍNDICE DÉFICIT DO SANEAMENTO BÁSICO EM MINAS GERAIS**

Número 2

Belo Horizonte
2021

GOVERNO DO ESTADO DE MINAS GERAIS

Governador

Romeu Zema Neto

Vice-governador

Paulo Eduardo Rocha Brant

SECRETARIA DE ESTADO DE PLANEJAMENTO E GESTÃO

Secretária de Estado de Planejamento e Gestão

Luísa Cardoso Barreto

FUNDAÇÃO JOÃO PINHEIRO

Presidente

Helger Marra Lopes

Vice-presidente

Mônica Moreira Esteves Bernardi

Diretoria de Estatística e Informações (Direi)

Eleonora Cruz Santos (Diretora)

Daniele Oliveira Xavier (Coordenadora-geral)

Coordenação de Habitação e Saneamento

Frederico Poley Martins Ferreira

Equipe técnica

Cláudio Jorge Cançado

Frederico Poley Martins Ferreira

Gabriel do Carmo Lacerda

Plínio de Campos Souza

Revisão

Deysiane Marques Franco Vieira

Normalização

Ana Paula da Silva

SUMÁRIO

Esta nota técnica possui dois objetivos principais: o primeiro é esclarecer e aprofundar a discussão relativa a alguns aspectos conceituais e opções metodológicas na construção do Índice Déficit do Saneamento Básico de Minas Gerais (IDSB). Entre eles, o porquê da escolha de determinados componentes e não de outros e mesmo quanto a definição dos pesos aplicados. O segundo objetivo remete à análise de resultados dos índices obtidos, pelos quais foram aplicadas técnicas de *cluster* e Análise dos Componentes Principais (ACP) para a definição e padronização de categorias quanto aos déficits municipais e mesmo o estudo de sua distribuição espacial.

1 INTRODUÇÃO

Para o aperfeiçoamento metodológico do atual processo de discussão e construção do IDBS de Minas Gerais, propõe-se o desenvolvimento desta segunda nota técnica¹ que aborda, além de um detalhamento e discussão de alguns aspectos conceituais do índice, uma análise específica de alguns dos seus resultados. Para tal, são utilizadas as técnicas de *cluster* e componentes principais, a partir dos municípios e dos TS (bacias hidrográficas), que permitem definir diferentes categorias na classificação dos déficits observados por municípios.

Inicialmente, deve-se ressaltar que na escolha das áreas de atuação (água, esgoto e lixo) e dos indicadores propriamente ditos, procurou-se seguir o princípio fundamental de que esses seriam os “serviços mínimos” que as populações municipais urbanas deveriam ter acesso. Sem dúvida, hoje, no Brasil, estão disponíveis várias pesquisas domiciliares, bancos de dados e registros administrativos sobre saneamento, com uma diversidade muito grande de variáveis, indicadores, índices, inclusive, disponíveis nos mais distintos formatos e temporalidades.

No entanto, para uma abordagem global de “déficit”, de “falta” ou, em última análise, de “necessidades” em saneamento, não faz sentido, pelo menos num primeiro momento, abordar simultaneamente todo esse universo. Por exemplo, serem discutidos aspectos muitos específicos relacionados ao detalhamento das técnicas utilizadas para a prestação, da qualidade, ou das características dos mais diversos serviços prestados, se ainda uma parcela considerável das pessoas, inicialmente, não tem o mínimo de acesso e/ou cobertura garantidos.

Assim, sob esta perspectiva de déficit, ou de problemas prioritários em saneamento, não faz muito sentido, por exemplo, discutir a frequência no abastecimento de água, se ainda há um considerável percentual de domicílios que não possuem acesso ou não estão acessando o abastecimento de água. Da mesma maneira, pode-se pensar em relação ao esgoto coletado e à disposição final dos resíduos sólidos². Parte-se da ideia de que ter o acesso ao mínimo dos serviços mais básicos é pré-condição para quaisquer outros tipos de análise e/ou discussões, sendo que, aqueles que não os possuem, são sempre prioritários na indicação do déficit, em relação aos que já possuem algum acesso, inclusive, mesmo que a “qualidade” do serviço, sob as mais diversas abordagens, ainda não seja a adequada. Nessa lógica, os domicílios que não apresentem o serviço e/ou apresentem os serviços mais precários são os prioritários para as políticas públicas. Isso também se

¹ Para mais detalhamento, vide nota técnica nº 1, disponível em: http://fjp.mg.gov.br/wp-content/uploads/2021/03/14.04_NotaTecnica_IDS_CHS.pdf. Acesso em: 27 ago. 2021.

² No caso da drenagem, conforme discutido na Nota Técnica anterior, o maior problema está relacionado à dificuldade na obtenção de dados padronizados. Porém, assim que for possível obtê-los, a lógica conceitual desenvolvida será a mesma.

reflete nos municípios que apresentam mais domicílios desses tipos, seja em termos absolutos, seja em termos relativos.

Um outro aspecto que deve ser pontuado é que, primeiramente, o índice se refere às populações urbanas. Sem dúvida, as pessoas que residem na área rural no Brasil, de uma forma geral, tendem a ser mais carentes de serviços de saneamento. Exatamente, por se observar esse fato, ao se calcular os déficits “urbanos”, considera-se que, “muito provavelmente”, as suas localidades, com características rurais, se encontram em pior situação. Dessa maneira, o índice serviria, de uma forma ou de outra, para chamar a atenção para o respectivo município de uma maneira global.

De qualquer maneira, o saneamento em áreas rurais precisa ser mais bem avaliado nos desenvolvimentos futuros do índice. A forma de abordagem e as variáveis que devem ser levadas em consideração para as áreas rurais, muitas vezes, são substantivamente distintas das áreas urbanas. Não é trivial a junção dos dois aspectos (urbano e rural) num índice sintético simples, como se pretende com o IDSB.

Especificamente, em relação ao indicador Déficit de Coleta de Esgoto (DCE), o qual leva em consideração o volume de esgotos coletados em relação ao volume de água consumida, um aspecto que pode ser questionado é que esse indicador não considera soluções alternativas e/ou individuais adequadas, como no caso do uso da fossa séptica.

De fato, o ideal seria que fossem descontados no DCE os esgotos destinados às fossas sépticas que tivessem manutenção adequada e, também, destinação final adequada do lodo gerado. Infelizmente, a informação sobre uso de fossas sépticas só é obtida regularmente a partir de pesquisas domiciliares, tais como a Pesquisa Nacional por Amostra de Domicílios Contínua (PnadC) – de periodicidade anual, porém significativa apenas para as capitais, regiões metropolitanas e estados – e nos Censos Demográficos (significativa para os municípios, porém de periodicidade decenal). Mais especificamente com relação ao descarte correto dos resíduos gerados, estas informações são ainda mais difíceis de se obter para todos os municípios.³

Mesmo reconhecendo essas deficiências, um aspecto que, no caso, também deve ser considerado é que, normalmente, os esgotos tratados antes de serem lançados ao meio ambiente, são aqueles normalmente coletados. Assim, um indicador acaba por influenciar o outro (coleta x tratamento). No caso das fossas sépticas, o esgoto não é diretamente coletado e, dessa maneira, muito provavelmente, seus resíduos gerados e esgotados (lodo) também não são tratados.

³ De qualquer maneira, futuramente se pretende incluir, nas análises especialmente mais agregadas, como se dá o uso dessas soluções individuais.

No caso do indicador de Resíduos Sólidos Urbanos (RSU), que mede a proporção de volume destinado inadequadamente (por exemplo, lixões a seu aberto, jogado em rios, córregos e lagos etc.), poderia se questionar o porquê da não inclusão da cobertura da coleta de lixo. Nesse aspecto, considerou-se que a coleta de lixo urbano está praticamente universalizada no estado de Minas Gerais. A Pesquisa Nacional por Amostra de Domicílios Contínua (Pnac), do (Instituto Brasileiro de Geografia e Estatística (IBGE) de 2020, identificou 43.029 domicílios com coleta de lixo inadequada no estado em 2019, em um total de aproximadamente 6,5 milhões de domicílios particulares permanentes urbanos. Nessa dimensão, sem dúvida, os problemas mais agudos se encontram na destinação final do lixo.

Quanto à questão da definição dos diferentes tipos de pesos para os indicadores, essa normalmente é uma discussão (muitas vezes sem solução) que permeia praticamente todos os índices que procuram sintetizar diferentes indicadores e variáveis. Esse também é o caso das discussões no saneamento. Assim, o que seria mais importante: ter acesso aos serviços de água ou aos serviços de esgotamento? Ao tratamento de esgoto ou à coleta de lixo?

Muito provavelmente, cada especialista terá uma boa quantidade de argumentos para indicar suas preocupações como sendo as mais prioritárias. Nesse sentido, a definição dos pesos para as três áreas de atuação do IDSB procurou sinalizar um equilíbrio na utilização dos indicadores, respectivamente com os seguintes pesos, água: 0,3; esgoto: 0,4; e lixo: 0,3.

Especificamente, no esgotamento sanitário, são incluídos dois indicadores, na medida em que não foi possível identificar em um único indicador simultaneamente as dimensões da coleta e do tratamento de esgotos. Por esse fato, a área de esgoto apresenta mais peso, cada um dos seus componentes apresentam um peso 0,2, implicando num peso global de 0,4.

Por fim, deve-se observar que foram utilizados como parâmetros as metas definidas pelo Plano Nacional de Saneamento Básico (Plansab) para o ano de 2033. Mesmo que para alguns indicadores a meta estabelecida foi de 100% de cobertura, para outros, isso não ocorreu, como é o caso do esgotamento.

O Plansab, além de ser um instrumento de planejamento, pactuado entre os mais diferentes setores do país, procurou identificar a viabilidade da completa universalização ou não dos diferentes aspectos do saneamento estabelecendo prazos para cada serviço.

2 ANÁLISE MULTIVARIADA – CLUSTER E COMPONENTES PRINCIPAIS

2.1 Análise de Cluster

No sentido de melhor definição das faixas de recorte dos dados e compreensão do comportamento do IDSB, foram aplicados métodos de análise multivariada, particularmente a Análise Cluster e a Análise de Componentes Principais. A análise multivariada pode ser definida como:

uma análise estatística que trabalha com medidas – atributos – múltiplas de uma ou mais amostras de indivíduos, tomadas genericamente como um sistema único de medidas, i.e., consideram a interligação geral de variáveis aleatórias simultaneamente. (SIMÕES, 2005, p.16).

Nos objetivos propostos, os “indivíduos” são os municípios mineiros e os atributos são os componentes e subcomponentes que compõem o Índice Déficit de Saneamento Básico.

Primeiramente, foi utilizado o método *cluster*, ou método de agrupamento/classificação, que possui como objetivo “dividir em subconjuntos, os mais semelhantes possíveis, conjuntos de elementos, de maneira que tais elementos pertencentes a um mesmo grupo sejam similares às características (atributos) que forem medidas em cada elemento” (SIMÕES, 2005, p.17). Em outras palavras, o objetivo da *clusterização* é agrupar diferentes elementos em grupos exclusivos, os mais homogêneos internamente (mínima variabilidade intraclasse e máxima variabilidade interclasse), para que, dessa forma, seja possível a geração de “estruturas agregadas significativas”, assim como o desenvolvimento de “tipologias analíticas” (SIMÕES, 2005, p.17-18).

De forma mais específica, aplicou-se o método *Ward*. Esse é um procedimento hierárquico que, a partir da matriz de distância euclidiana⁴, construída com as variáveis selecionadas, consiste na minimização da soma dos erros quadráticos ou da variância *intracluster* das distâncias, em cada estágio de agrupamento dos diferentes indivíduos (EVERITT *et al*, 2011; DONI, 2004). Ou seja, agrupam-se as observações (indivíduos) que apresentam menor erro quadrático ou menor variância até o ponto em que nova observação (indivíduo) incluída(o) no grupo seja capaz de reduzir esse erro ou variância. Portanto, unem-se, em cada etapa, elementos/observações/grupos com menor dissimilaridade entre si (KAUFMAN; ROUSSEUW, 1990).⁵ Em outras palavras, procura-se agrupar os elementos que se apresentam mais próximos entre si, devido às suas similaridades.

⁴ A fórmula da distância euclidiana é: $d(i, j) = [\sum_{k=1}^p w_k^2 (x_{ik} - x_{jk})^2]^{1/2}$

⁵ Formalmente: quer-se minimizar o aumento do total da soma dos erros quadráticos intra-cluster, E, dado por: $E = \sum_{m=1}^g E_m$. Onde: $E_m = \sum_{l=1}^{n_m} \sum_{k=1}^p (x_{ml,k} - \bar{x}_{m,k})^2$. Em que $\bar{x}_{m,k} = (1/n_m) \sum_{l=1}^{n_m} x_{ml,k}$ é a média do m-ésimo cluster da k-ésima variável, sendo $x_{ml,k}$ o score da k-ésima variável ($k = 1, \dots, p$) para o l-ésimo objeto ($l = 1, \dots, n_m$) no m-ésimo cluster ($m = 1, \dots, g$) (EVERITT *et al*, 2011, p. 77-78).

O outro método, a análise de componentes principais (ACP), auxilia na “redução do número de variáveis a serem avaliadas”, assim como na “interpretação das combinações lineares construídas” (MINGOTI, 2020, p.57). Dessa maneira, “a informação contida nas p-variáveis originais é substituída [sumarizada] pela informação contida em k ($k < p$) componentes principais não correlacionados” (MINGOTI, 2020, p. 57). Dito de outra forma, o método ACP⁶ permite a redução do número p de variáveis – ou atributos – explicativos de um conjunto de observações a um pequeno número k de índices, exatamente os componentes principais, que têm como propriedade a não correlação (SIMÕES, 2005, p. 17). Se bem-sucedido, dado que $k < p$, tem-se que um número pequeno de componentes explicaria um número maior de variáveis/atributos

Finalmente, o uso dos dois métodos em conjunto auxilia, por um lado, na identificação de padrões de agrupamentos a partir das características/atributos das observações (Análise *Cluster*); por outro, na identificação das correlações entre as próprias variáveis e o quanto de variância acumulada pode ser explicada pela condensação em cada componente, assim como o peso de cada variável na determinação desse mesmo componente (ACP).

O roteiro de análise consiste em sete etapas:

- a) definição das variáveis para o cálculo dos componentes e subcomponentes do IDSB, além do próprio Índice síntese;
- b) exclusão dos municípios com ausência de alguma informação – variável/componente - que comprometa o cálculo do IDSB;
- c) cálculo do IDSB e dos seus componentes e subcomponentes para os municípios que não foram retirados da amostra;
- d) aplicação das análises multivariadas (*Cluster* e ACP) – utilizando como base os resultados do IDSB, de seus componentes e subcomponentes – para melhor compreensão de seus agrupamentos e variáveis;
- e) cálculo das médias aritméticas e do desvio padrão amostral para cada um dos agrupamentos identificados pela Análise *Cluster*;
- f) definição dos recortes e reagrupamento dos municípios a partir de cinco categorias em relação ao Déficit Saneamento Básico, em ordem crescente de intensidade: I) muito baixo; II) baixo; III) médio; IV) alto; V) muito alto;
- g) sumário de algumas informações descritivas dos grupos, segundo variáveis específicas e sua distribuição espacial nos Territórios do Saneamento (TS).

⁶ Formalmente, dado que as variáveis estão na mesma escala, utilizou-se para a ACP o cálculo de uma matriz de covariância C . Uma descrição detalhada do cálculo se encontra em Manly e Alberto (2016).

Em resumo, a variância dos componentes principais são os autovalores das p-variáveis (λ_p) da matriz de covariância C . Assim, assumindo que os autovalores são ordenados como $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, então λ_i corresponde ao i-ésimo componente principal (Z_i) da combinação linear das variáveis X_1, X_2, \dots, X_p , geradas das p variáveis dos n indivíduos (municípios), onde se tem $Z_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p$. Como $\text{Var}(Z_i) = \lambda_i$ e as constantes $a_{i1} a_{i2} \dots a_{ip}$ são os elementos que correspondem aos autovalores, na mesma escala se tem $a_{i1}^2 + a_{i2}^2 + \dots + a_{ip}^2 = 1$ (MANLY; ALBERTO, 2016).

Uma das propriedades dos autovalores é que eles somam o total dos elementos diagonais da matriz C , ou seja, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p = c_{11} + c_{22} + \dots + c_{pp}$. Portanto, como c_{ii} é a variância dos X_i ; e λ_i é a variância de Z_i , tem-se que a soma das variâncias dos componentes principais é igual a soma das variâncias das variáveis originais. Dessa forma, os componentes principais sumarizam toda variação dos dados originais (MANLY; ALBERTO, 2016).

Tabela 1: Componentes, variáveis e descrição do IDSB

Componente	Subcomponentes	Código Variável no Snis	Descrição da variável	Fórmula do Componente/ Subcomponente	Descrição do componente
Déficit de abastecimento de água	Déficit de Abastecimento de Água (DA)	IN023	Índice de atendimento urbano de água	$DA = \left\{ 1 - \left[\frac{\text{pop urbana atendida com abastecimento}}{\frac{\text{pop urbana}}{\text{Meta Plansab}_{2033}}} \right] \right\} \times 100$	Proporção da população urbana atendida com abastecimento de água em relação ao total da população urbana e quão distante essa proporção se encontra da meta estabelecida pelo Plansab até 2033
Déficit de esgotamento sanitário	Déficit de Coleta de Esgoto (DCE)	IN015	Índice de coleta de esgoto	$DCE = \left\{ 1 - \left[\frac{\text{Volume de esgotos coletados}}{\frac{\text{vol. água consumido} - \text{vol. água trada exportada}}{\text{Meta Plansab}_{2033}}} \right] \right\} \times 100$	Proporção entre o volume total de esgoto coletado na rede geral urbana e a diferença entre o volume de água consumido e o tratado exportado. E o quão distante essa proporção se encontra da meta estabelecida pelo Plansab até 2033
	Déficit de Tratamento de Esgoto (DTE)	IN046	Índice de esgoto tratado referido à água consumida	$DTE = \left\{ 1 - \left[\frac{\text{Vol. esgotos tratados} + \text{vol. esgoto bruto tratado fora}}{\frac{\text{vol. água consumido} - \text{vol. água trada exportada}}{\text{Meta Plansab}_{2033}}} \right] \right\} \times 100$	Proporção entre o volume de esgoto urbano tratado dentro e fora do município em relação ao saldo de água consumida no município. E o quão distante essa proporção se encontra da meta estabelecida pelo Plansab até 2033
Déficit de resíduos sólidos urbanos	Déficit de Resíduos Sólidos Urbanos (DRSU)	UP007	Quantidade anual de resíduos sólidos domiciliares e de resíduos públicos recebido na unidade de processamento	$DRSU = \left\{ 1 - \left[\frac{\text{Qnte total RSU} - \text{Qnte RSU destino inadequado}}{\frac{\text{Quantidade total de RSU}}{\text{Meta Plansab}_{2033}}} \right] \right\} \times 100$	Proporção de volume de resíduos sólidos urbanos destinados inadequadamente (lixões a céu aberto, jogados em rios, córregos e lagos etc.) em relação ao total de resíduos sólidos produzidos na área urbana do município. E quão distante essa proporção se encontra da meta estabelecida pelo Plansab até 2033
		UP003	Tipo de unidade de processamento pelo manejo dos RSU		

Fonte: Elaboração própria.

As variáveis utilizadas na construção de cada componente do IDSB foram retiradas do Sistema Nacional de Informações sobre Saneamento (Snis), relativas ao ano de 2019, e se encontram sumarizadas na tabela acima:

A partir do cálculo de cada componente e subcomponente, calcula-se o IDSB conforme fórmula abaixo (Equação 1). Como dito anteriormente, foram utilizados pesos específicos para cada componente e subcomponente com o objetivo de ponderar a presença de duas variáveis para Esgotamento Sanitário. Dessa maneira, o $IDSB_{município}$ possui a seguinte estrutura de ponderação:

- a) Déficit de Abastecimento de Água (DA) – peso 0,3;
- b) Déficit de Coleta de Esgoto (DCE) – peso 0,2;
- c) Déficit de Tratamento de Esgoto (DTE) – peso 0,2; e, por fim,
- d) Déficit de Resíduos Sólidos Urbanos – peso 0,3.

$$IDSB_{município} = [(DA * 0,3) + (DCE * 0,2) + (DTE * 0,2) + (DRSU * 0,3)] (1)$$

onde:

$IDSB_{município}$ = Índice Déficit do Saneamento Básico

DA = Déficit de Abastecimento de Água;

DCE = Déficit na Coleta de Esgotos;

DTE = Déficit de Tratamento de Esgoto;

DRSU = Déficit de Destinação dos Resíduos Sólidos Urbanos;

Os municípios que possuem ao menos um dos índices ausentes, isto é, não respondidos, foram excluídos da amostra, totalizando 285 municípios faltantes ou 33,4% do total de 853 municípios, e não entraram na análise multivariada. Abaixo, há uma tabela com as informações do total de municípios por Território do Saneamento Básico de Minas Gerais.

É possível constatar disparidade regional de informações, por exemplo, entre o Território São Francisco Alto-Médio (79,9%) e o Território Médio-Baixo (38,1%) de respondentes.

A falta de informações pode ser interpretada como baixa capacidade administrativa em possuir as informações e responder os questionários do Snis.

Tabela 2: Observações por Território do Saneamento

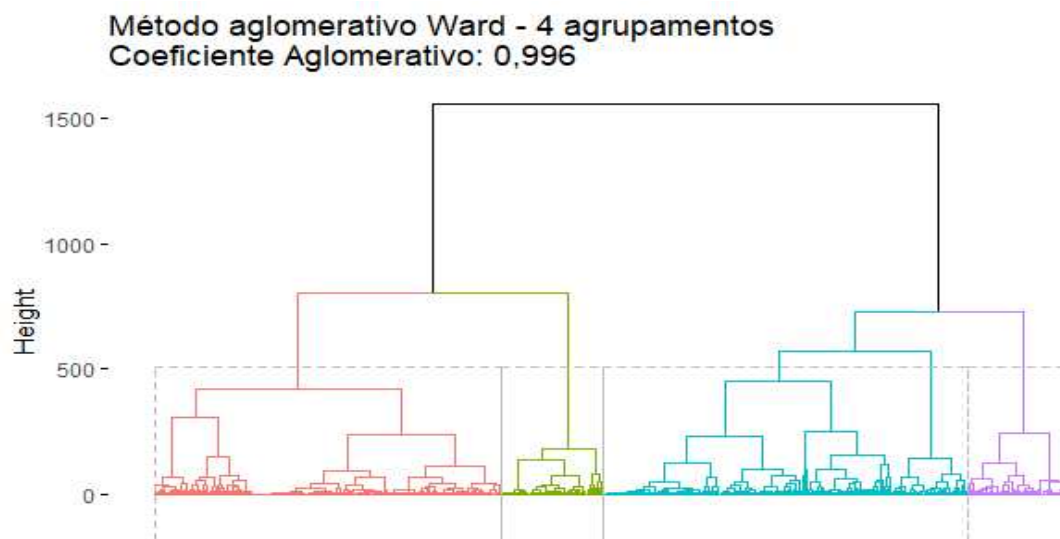
Territórios do Saneamento (TS)	Total de observações por TS	Total de municípios por TS	Observações TS/Total TS
Jequitinhonha	40	67	59,7%
Paraíba do Sul	69	101	68,3%
Rio Doce	142	221	64,3%
Rio Grande	138	182	75,8%
Rio Paranaíba	31	46	67,4%
São Francisco Alto-Médio	111	139	79,9%
São Francisco Médio-Baixo	37	97	38,1%
Total Geral	568	853	66,6%

Fonte: Dados básicos: BRASIL, 2019. Elaboração própria.

Para a análise multivariada, ao final da exclusão, tem-se 568 observações e cinco variáveis ($IDSB_{município}$, DA, DCE, DTE, DRSU). Foram testados diferentes agrupamentos, definindo-se quatro grupos em função da estrutura do dendrograma, dos valores médios intragrupos e intergrupos, da

distribuição das observações no gráfico do ACP e da simplicidade de entendimento das diferenças entre os grupos. O dendrograma, os agrupamentos e o resultado do Coeficiente Aglomerativo (CA)⁷ do método podem ser visualizados abaixo:

Figura 1: Dendrograma dos agrupamentos dos municípios mineiros pelo IDSB



Fonte: Dados básicos: FUNDAÇÃO JOÃO PINHEIRO, 2021. Elaboração própria a partir do IDSB

Do ponto de vista descritivo, a Tabela 3 sumariza as informações das médias intragrupos dos quatro agrupamentos, enquanto na Tabela 4, os desvios padrões amostrais dos *clusters*.

Tabela 3: Médias intragrupos dos componentes e subcomponentes do IDSB

Grupo	Déficit Abastecimento (DA)	Déficit Coleta de Esgoto (DCE)	Déficit Tratamento de Esgoto (DTE)	Déficit Resíduos Sólidos Urbanos (DRSU)	Índice Déficit do Saneamento Básico (IDSB)
1	4,9	14,3	11,3	6,7	8,6
2	6,0	21,8	90,6	0,6	24,4
3	7,8	25,2	22,5	96,7	40,9
4	8,2	27,3	93,5	87,0	52,7

Fonte: Dados básicos: BRASIL, 2019; FUNDAÇÃO JOÃO PINHEIRO, 2021. Elaboração própria.

A partir da Tabela 3, verifica-se os agrupamentos dos municípios no sentido das melhores condições (i.e., menores valores médios para seus componentes, conseqüentemente, menor média

⁷ Quanto mais próximo de 1 está o Coeficiente Aglomerativo, mais a estrutura apresenta agrupamentos bem definidos (KAUFMAN; ROUSSEUW, 1990).

do IDSB) para piores condições (i.e., maiores valores médios para suas componentes, conseqüentemente, maior média IDSB). É possível notar, pelas médias, dois grandes grupos de carências, um relacionado ao tratamento de esgoto (DTE) e outro ao manejo de resíduos sólidos (DRSU). Ademais, constata-se valores mais próximo da universalidade no serviço de abastecimento público de água (DA).

De forma mais específica, nota-se, por um lado, que, enquanto o grupo 4 apresenta uma precariedade generalizada no tratamento de esgoto e no manejo dos resíduos sólidos, o grupo 3 apresenta uma precariedade concentrada na gestão dos resíduos sólidos.

Por outro lado, com médias IDSB menores, o grupo 1 tem como principal carência média a coleta do esgoto, enquanto, no grupo 2, a característica marcante é a carência de tratamento de esgoto.

Tabela 4: Desvio padrão intragrupos dos componentes e subcomponentes do IDSB

Grupo	Déficit Abastecimento (DA)	Déficit Coleta de Esgoto (DCE)	Déficit Tratamento de Esgoto (DTE)	Déficit Resíduos Sólidos Urbanos (DRSU)	Índice Déficit de Saneamento Básico (IDSB)
1	5,4	11,3	9,3	17,2	6,9
2	7,5	23,5	18,5	2,9	5,7
3	8,1	17,7	17,8	6,9	8,0
4	14,1	29,3	13,1	22,9	10,1

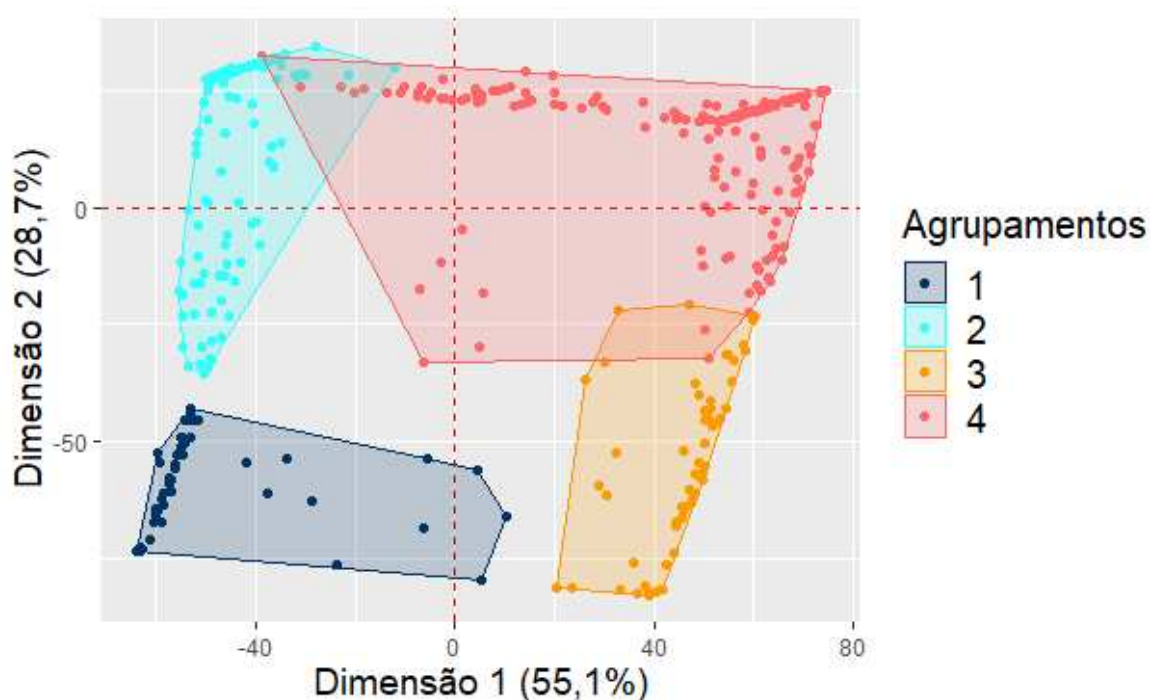
Fonte: Dados básicos: BRASIL, 2019; FUNDAÇÃO JOÃO PINHEIRO, 2021. Elaboração própria.

A Tabela 4 apresenta os desvios padrões amostrais intragrupos dos componentes do IDSB. Saliencia-se a maior dispersão dos dados de DCE, conferindo mais heterogeneidade a esse subcomponente nos diferentes grupos. Essa diversidade é, inclusive, reforçada pelas médias na Tabela

2.2 Análise dos Componentes Principais (ACP)

Do ponto de vista da ACP, os mesmos agrupamentos da Análise *Cluster* estão dispostos no Gráfico 1 abaixo. Nele também constam dois dos componentes principais que possuem maior variância explicada por cada um deles, cuja soma acumulada é 83,8%. Os pesos do quanto cada variável está sendo explicada por cada componente está na Tabela 5.

Gráfico 1: Análise de Componentes Principais dos municípios pelo IDSB



Fonte: Dados básicos: BRASIL, 2019; FUNDAÇÃO JOÃO PINHEIRO, 2021. Elaboração própria.

Tabela 5: Pesos das variáveis nos Componentes Principais

Variável	Dimensão 1	Dimensão 2	Dimensão 3	Dimensão 4	Dimensão 5
Déficit Abastecimento (DA)	0,02	0,01	0,05	0,96	0,27
Déficit Coleta de Esgoto (DCE)	0,14	0,03	0,97	-0,10	0,18
Déficit Tratamento de Esgoto (DTE)	0,07	0,98	-0,08	-0,05	0,18
Déficit Resíduos Sólidos Urbanos (DRSU)	0,93	-0,14	-0,19	-0,08	0,27
Índice Déficit de Saneamento Básico (IDSB)	0,33	0,16	0,13	0,23	-0,89

Fonte: Dados básicos: BRASIL, 2019; FUNDAÇÃO JOÃO PINHEIRO, 2021. Elaboração própria.

Em primeiro lugar, é possível perceber uma boa distribuição nos quadrantes de cada um dos grupos – a exceção é o agrupamento 4, com um conjunto de observações em “transição” para o grupo 2.

Em segundo lugar, a partir dos pesos de explicação de cada variável dentro de cada componente principal, se averigua, em linhas gerais, que a Dimensão 1, dada sua elevada *correlação positiva* com DRSU, é fortemente determinada pela *ausência* ou pela *gestão inadequada* dos RSU. Assim, quando a direita (positivo) no eixo X, maiores são os DRSU. Enquanto, pelo contrário, valores

negativos nas abscissas, menores os DRSU. Ademais, há um peso relativo do indicador síntese (IDSB) no mesmo sentido, isto é, quanto maior IDSB, mais a direita; quanto menor IDSB, menor déficit, mais à esquerda.

Já em relação a Dimensão 2, em linhas gerais, dada a elevada *correlação positiva* com DTE, quanto maior a *ausência* de tratamento de esgoto adequado, maiores os valores de DTE, portanto, observa-se os pontos positivos no eixo Y. Pelo contrário, a presença de tratamento adequado do esgoto se vincula a valores negativos no eixo das ordenadas.

Novamente, ressalta-se que os vetores são ponderados pela condensação das variáveis, por isso as relações não são lineares nos eixos. Contudo, se pode notar as médias da análise *cluster*, que os agrupamentos seguem o padrão dos quadrantes, isto é, o grupo 4 (quadrante superior direito) é aquele com piores indicadores de DTE e DRSU; o grupo 3 (quadrante inferior direito) com piores (maiores) resultados no DRSU, mas melhores (menores) DTE; o grupo 2 (quadrante superior esquerdo) com maiores DTE e menores no DRSU; e, por fim, o grupo 1 (quadrante inferior esquerdo) com indicadores melhores (menores) em ambos os componentes.

Em síntese, os agrupamentos da Análise *Clusters* se distribuíram bem segundo os quadrantes da ACP, indicando, primeiro, a correlação entre variáveis, realçando as mais relevantes que condensam a variação – ou seja, os problemas no manejo de resíduos sólidos e no tratamento de esgoto. E, segundo, agrupando os municípios pelas características observadas nas variáveis, permitindo uma hierarquização da intensidade da condição deficitária dos diferentes elementos que compõe o saneamento básico.

Com base nos resultados acima, considerou-se as médias do IDSB dos *clusters* como um bom recorte dos agrupamentos e das características entre os municípios mineiros. Dessa forma, utilizando as médias dos quatro agrupamentos, chegou-se aos cinco grupos de análise em relação às condições de déficit do saneamento básico no conjunto dos municípios mineiros.

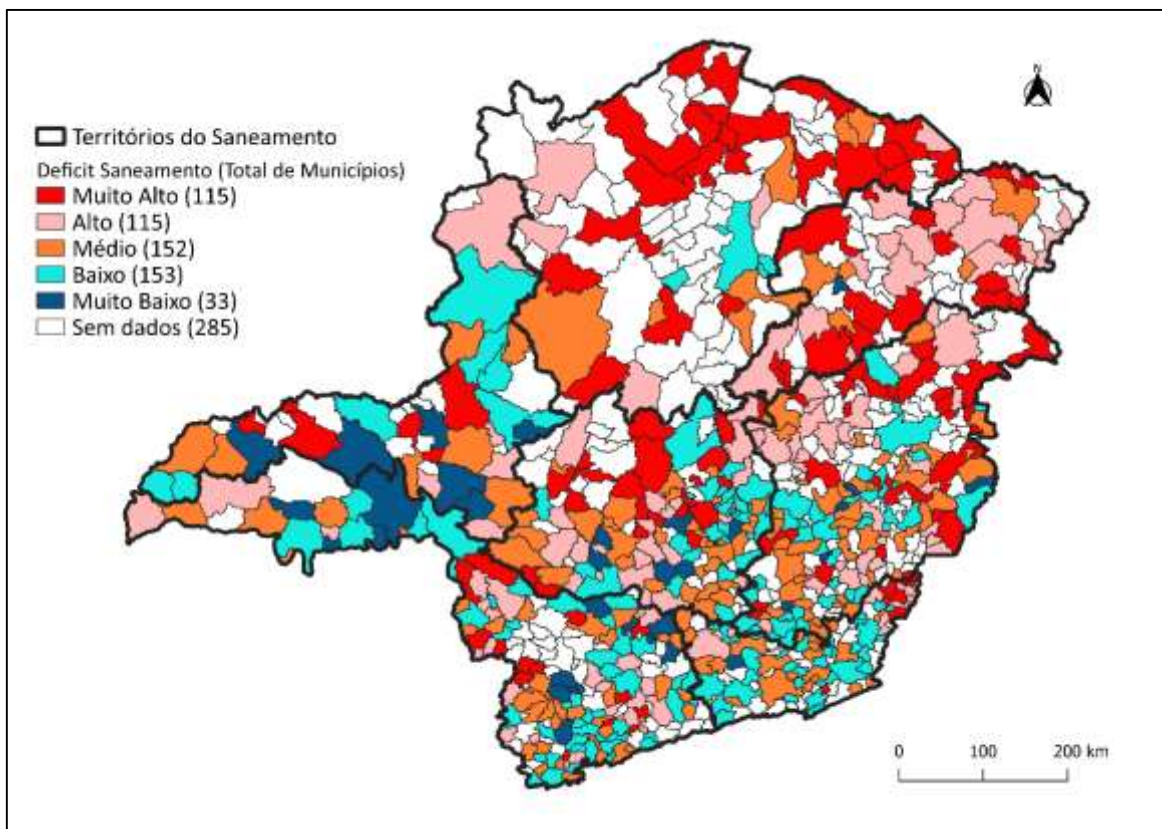
Tabela 6: Categorias do IDSB, valores de corte e quantidade de municípios por grupo

Categoria de IDSB	Valor do IDSB	Membros
Muito Alto	Maior que 52,7	115
Alto	Menor ou igual a 52,7 e maior que 40,9	115
Médio	Menor ou igual a 40,9 e maior que 24,4	152
Baixo	Menor ou igual a 24,4 e maior que 8,6	153
Muito Baixo	Menor ou igual que 8,6	33

Fonte: Elaboração própria.

Do ponto de vista espacial, a disposição se encontra no Mapa 1. Além dos municípios por categoria do IDSB, tem-se também os TS e os municípios sem informação. No sentido de ter uma visão mais acurada, as Tabelas 7 a 9 apresentam, respectivamente, o quantitativo de cada categoria por TS, as proporções da distribuição de cada categoria por TS e a composição categórica em cada TS.

Mapa 1: Municípios por categoria IDSB



Fonte: Dados básicos: BRASIL, 2019. Elaboração própria.

Tabela 7: Municípios por categoria do IDSB e Território do Saneamento

Território do Saneamento	Total de municípios por categoria do IDSB				
	Muito Alto	Alto	Médio	Baixo	Muito Baixo
Jequitinhonha	16	18	5	0	1
Paraíba do Sul	6	8	21	33	1
Rio Doce	29	31	42	35	5
Rio Grande	15	26	37	47	13
Rio Paranaíba	6	5	7	7	6
São Francisco Alto-Médio	21	22	34	27	7
São Francisco Médio-Baixo	22	5	6	4	0
Total	115	115	152	153	33

Fonte: Dados básicos: BRASIL, 2019. Elaboração própria.

Tabela 8: Distribuição das categorias IDSB por território

Território do Saneamento	Categoria do IDSB				
	Muito Alto	Alto	Médio	Baixo	Muito Baixo
Jequitinhonha	13,9%	15,7%	3,3%	0,0%	3,0%
Paraíba do Sul	5,2%	7,0%	13,8%	21,6%	3,0%
Rio Doce	25,2%	27,0%	27,6%	22,9%	15,2%
Rio Grande	13,0%	22,6%	24,3%	30,7%	39,4%
Rio Paranaíba	5,2%	4,3%	4,6%	4,6%	18,2%
São Francisco Alto-Médio	18,3%	19,1%	22,4%	17,6%	21,2%
São Francisco Médio-Baixo	19,1%	4,3%	3,9%	2,6%	0,0%
Total	100,0%	100,0%	100,0%	100,0%	100,0%

Fonte: Dados básicos: BRASIL, 2019. Elaboração própria.

Tabela 9: Distribuição nos Territórios do Saneamento das diferentes categorias do IDSB

Território do Saneamento	Categoria do IDSB					Total
	Muito Alto	Alto	Médio	Baixo	Muito Baixo	
Jequitinhonha	40,0%	45,0%	12,5%	0,0%	2,5%	100,0%
Paraíba do Sul	8,7%	11,6%	30,4%	47,8%	1,4%	100,0%
Rio Doce	20,4%	21,8%	29,6%	24,6%	3,5%	100,0%
Rio Grande	10,9%	18,8%	26,8%	34,1%	9,4%	100,0%
Rio Paranaíba	19,4%	16,1%	22,6%	22,6%	19,4%	100,0%
São Francisco Alto Médio	18,9%	19,8%	30,6%	24,3%	6,3%	100,0%
São Francisco Médio-Baixo	59,5%	13,5%	16,2%	10,8%	0,0%	100,0%
Total de municípios por categoria	20,2%	20,2%	26,8%	26,9%	5,8%	100,0%

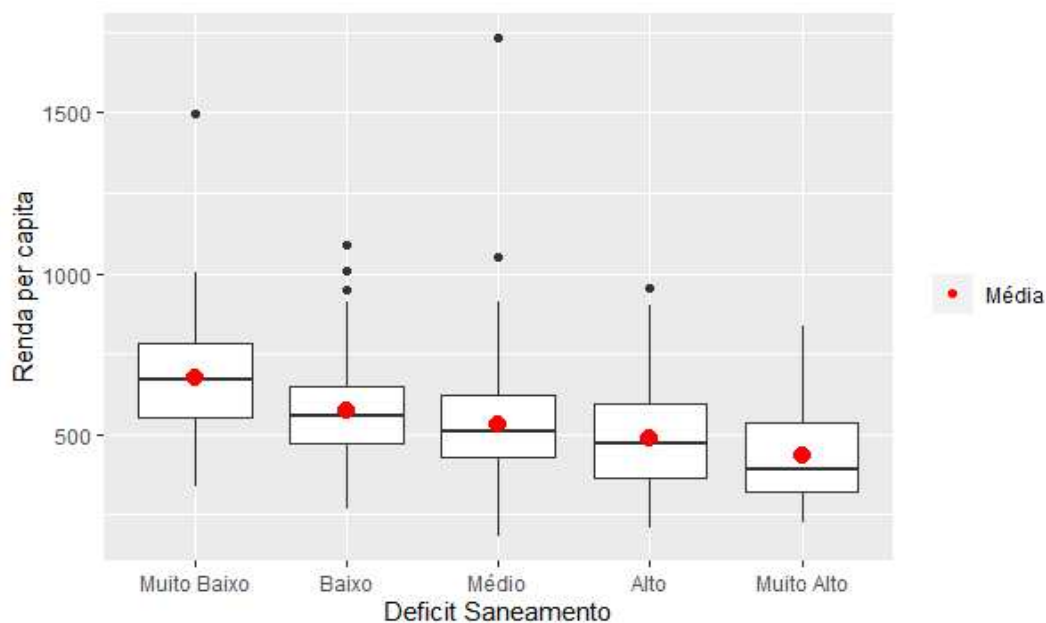
Fonte: Dados básicos: BRASIL, 2019. Elaboração própria.

A partir do Mapa 1 e das Tabelas 7 a 9, acima, depreende-se as diversidades e desigualdades regionais do IDSB por TS. Como o indicador utiliza todos os municípios mineiros com dados completos, o que se está calculando é a posição relativa de cada um deles em relação a todos os demais.

No intuito de compreender melhor a especificação inter e intragrupos, foram realizados cruzamentos de distribuição segundo variáveis de renda per capita, IDHM e taxa de urbanização. Os *boxplots*⁸ abaixo especificam as características socioeconômicas e sua vinculação com o próprio IDSB. Dessa maneira, dos Gráficos 2, 3 e 4, é possível verificar que os maiores déficits de saneamento básico se relacionam a, geralmente, aos municípios com menores renda per capita, IDHM e taxa de urbanização. Importante acrescentar que, regionalmente, o maior número de municípios considerados deficitários, pelo IDSB, localizava-se próximos a municípios cujas informações eram incompletas e que, portanto, não puderam compor o índice.

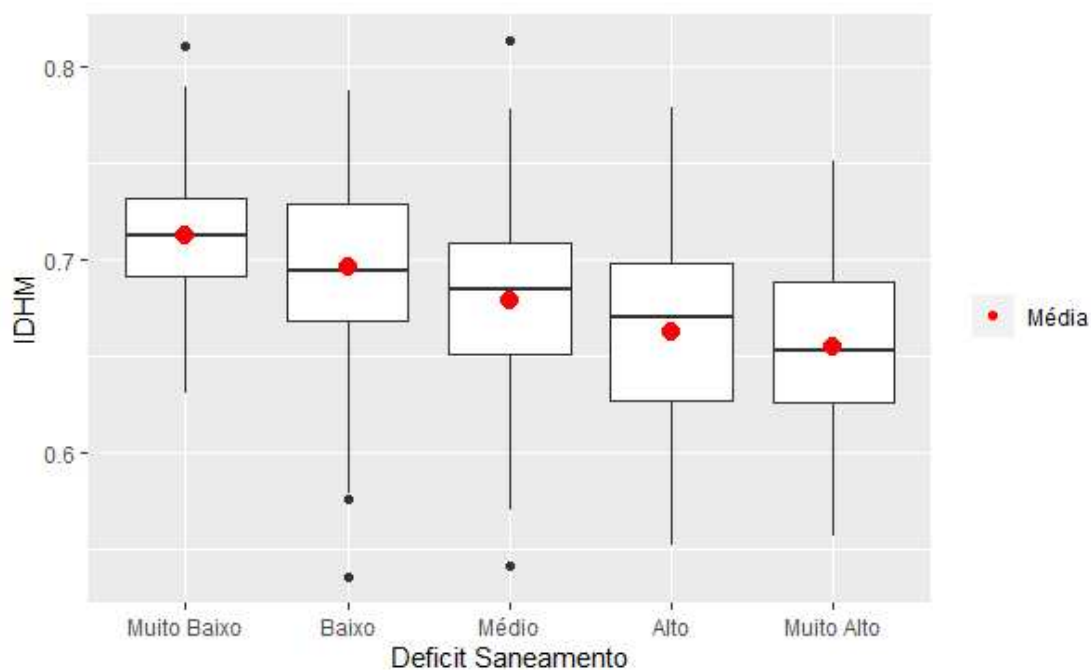
⁸ As caixas (*box*) são lidas da seguinte forma: a linha inferior representa o primeiro quartil; a linha no meio da caixa representa a mediana; a linha superior da caixa representa o terceiro quartil. Os pontos seriam os *outliers*. Por fim, foi adicionado um ponto vermelho às as médias dos grupos.

Gráfico 2: *Boxplot* da renda per capita por categoria IDSB



Fonte: Dados básicos: IBGE, 2010.; FJP, 2021. Elaboração própria.

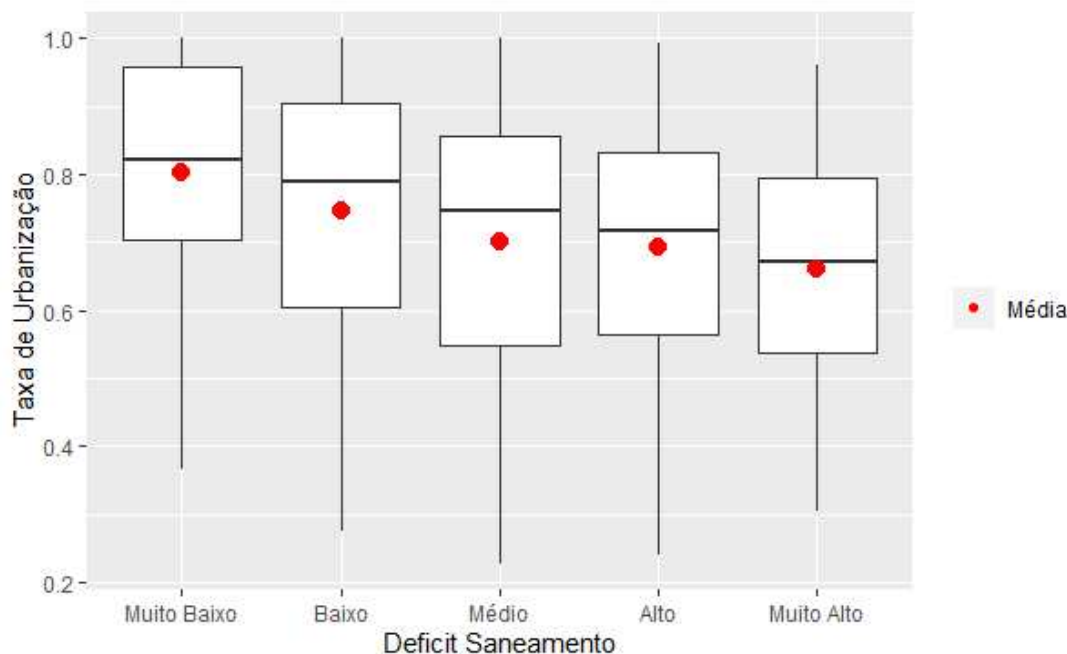
Gráfico 3: *Boxplot* do Índice de Desenvolvimento Humano Municipal (IDHM) por categoria IDSB



Fonte: Dados básicos: PROGRAMA..., 2020; FJP, 2021. Elaboração própria.

Fonte:

Gráfico 4: Boxplot da Taxa de Urbanização por categoria do IDSB



Fonte: Dados básicos: BRASIL, 2019; FUNDAÇÃO JOÃO PINHEIRO, 2021. Elaboração própria.

2.3 Os componentes do IDSB por categoria de déficit

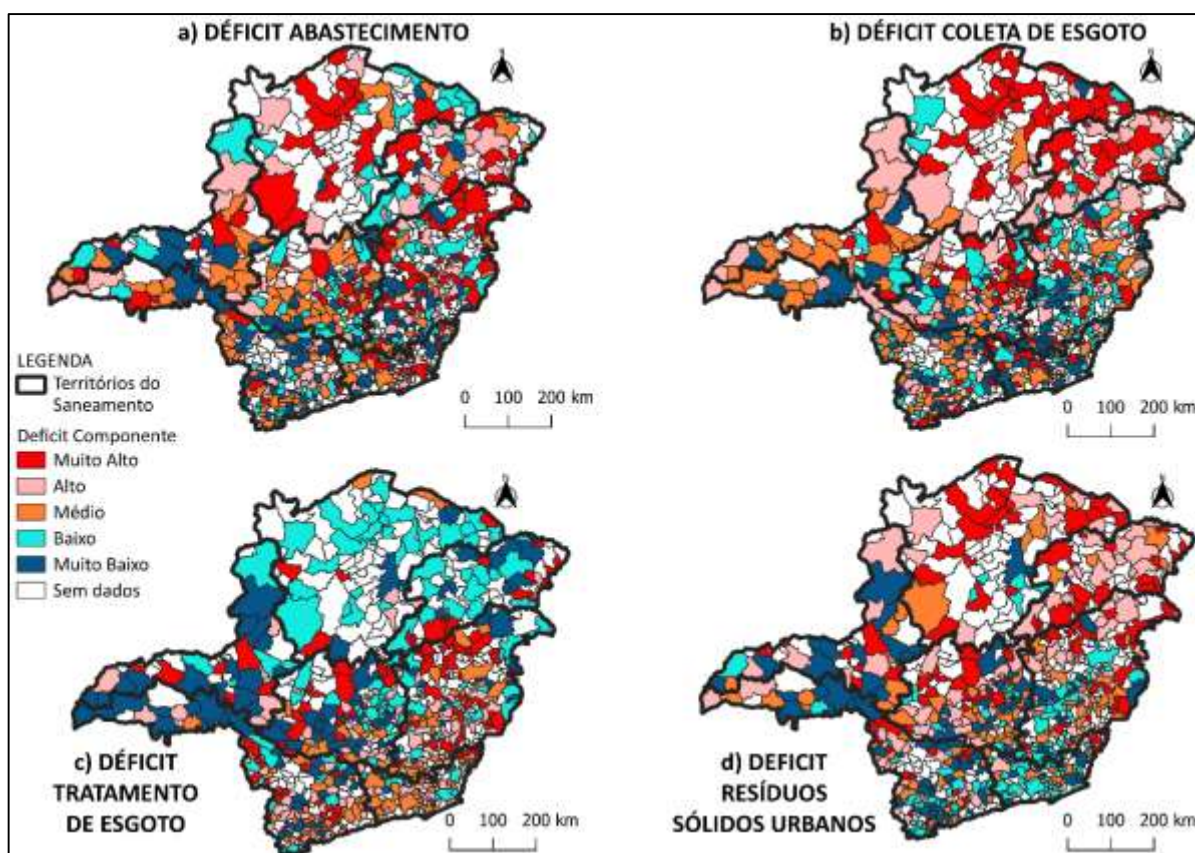
Já a especificação de intensidade do déficit de saneamento básico por componente está sintetizada na Figura 2. Diferentemente do indicador sintético, os recortes considerados foram a distribuição dos respectivos indicadores por quintis. A Tabela 10 apresenta o valor do quintil por cada componente e subcomponente do IDSB. Pondera-se que a classificação por categoria, caso o quintil seja do mesmo valor, deu-se segundo os menores valores do IDSB.

Com exceção do componente de tratamento de esgoto, nos mapas da Figura 2, a disparidade regional das condições de provimento dos serviços de abastecimento, coleta e manejo dos RSU está bem representada, com os valores alto e muito alto nas regiões Norte e Leste de Minas Gerais.

Já em relação com DTE, a disposição espacial menos evidente, com melhor colocação dos municípios das regiões Leste e Norte nas categorias baixo ou muito baixo, se justifica pelos elevados valores do DTE, conforme a Tabela 10. Exemplificando, nota-se por essa tabela que, se determinado município tiver um terço de sua população sem tratamento adequado de esgoto, ainda assim estaria, relativamente ao demais municípios mineiros, na categoria muito baixo de déficit de tratamento.

Na Tabela 10, nota-se, em primeiro lugar, como o serviço de abastecimento é generalizado nos municípios mineiros. Em segundo lugar, como o serviço de coleta de esgoto está em situação de universalização muito heterogênea no âmbito estadual. Por fim, como a gestão de resíduos sólidos também é problema generalizado, porém mais bem gerido nos municípios e regiões mais desenvolvidas do estado.

Mapa 2: Categorias de déficit por Componente



Fonte: Dados básicos; BRASIL, 2019; FUNDAÇÃO JOÃO PINHEIRO, 2021.

Elaboração própria.

Tabela 10: Recortes por categorias dos componentes do IDSB por quintis dos respectivos indicadores – 2019

Território do Saneamento	Categoria				
	Muito Alto	Alto	Médio	Baixo	Muito Baixo
DA	Menor igual a 100 e maior que 12,1	Menor igual a 12,1 e maior que 6,4	Menor igual 6,4 e maior que 1,6	Menor igual a 1,6 e IDSB maior que 47,5	Igual a 0 e IDSB menor igual 47,5
DCE	Menor igual a 100 e maior que 41,5	Menor igual 41,5 e maior que 24,8	Menor igual 24,8 e maior que 10,8	Menor igual 10,8 e IDSB maior que 33,0	Igual a 0 e IDSB menor igual 33,0
DTE	Igual a 100 e IDSB maior que 48,5	Igual a 100 e IDSB entre 48,5 e 25,3	Menor igual 100 e maior que 90,2	Menor igual 90,2 e maior que 37,0	Menor igual 37,0
DRSU	Igual a 100 e IDSB maior que 50,9	Menor igual a 100 e maior que 84,9	Menor igual 84,9 e IDSB maior que 31,2	Menor igual 84,9 e IDSB entre 31,2 e 20,6	Igual a 0 e IDSB menor que 20,6

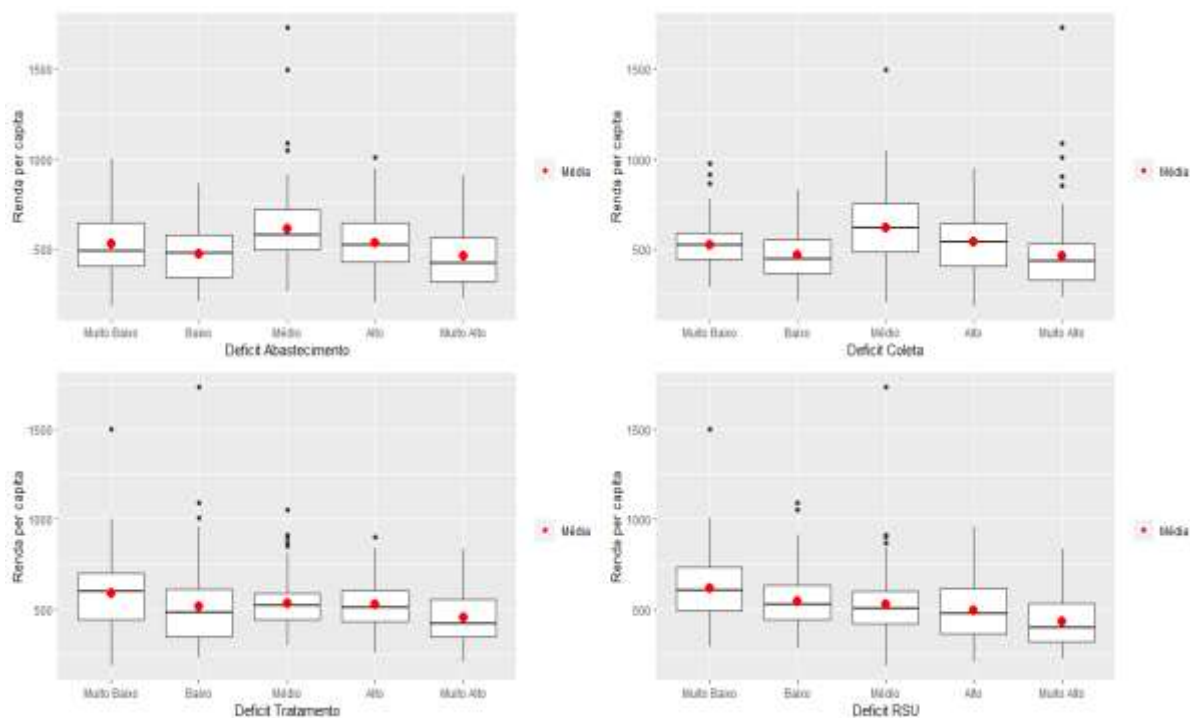
Fonte: Dados básicos: BRASIL, 2019.

Elaboração própria.

Os *boxplots* a seguir, com os cruzamentos de renda per capita municipal e do IDHM com as categorias de déficit por componente do IDSB, demonstram, por um lado, como os municípios mais ricos e desenvolvidos se classificam na categoria média em relação ao DA e ao DCE. Como tais resultados não assumem valores elevados, ou seja, não são condição generalizada de precariedade nos municípios em análise, seus valores não impactam tanto o IDSB. Principalmente no caso do abastecimento de água, a situação não está tão precária na maioria dos municípios do estado.

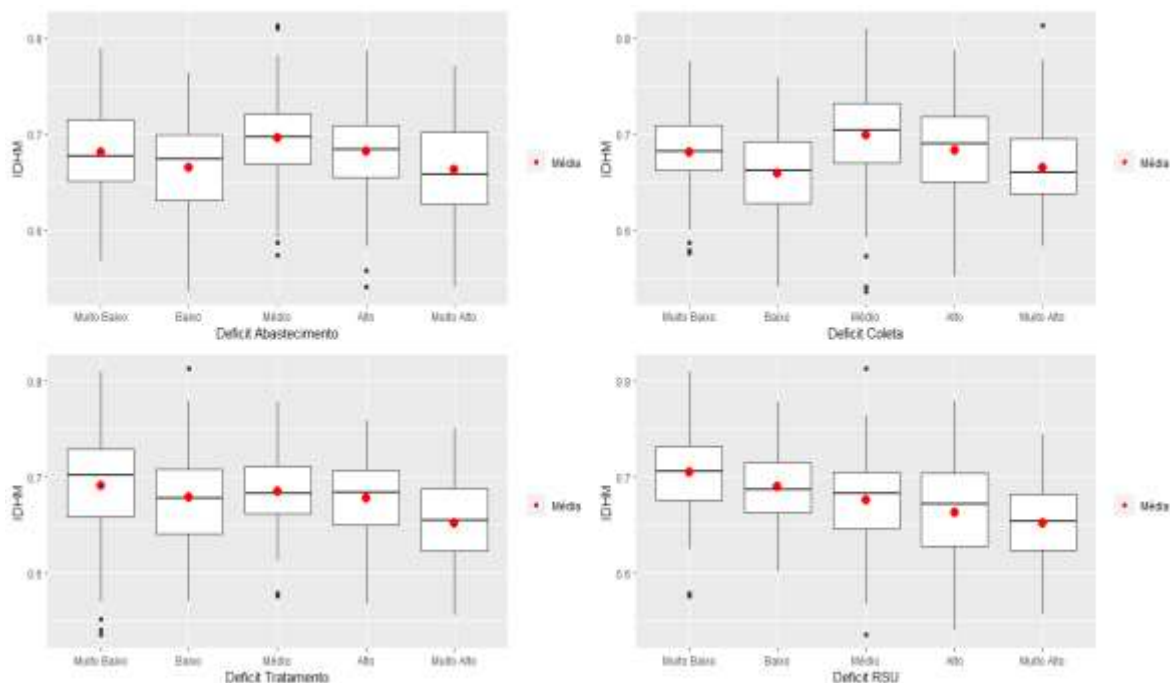
Por outro lado, a situação referente aos componentes DTE e DRSU se reflete como na ACP: municípios mais ricos e desenvolvidos estão em posição bem melhor que os demais do estado. Isso aponta que há uma situação de correlação inversa entre a presença desses serviços e os indicadores de renda per capita e de desenvolvimento humano. Nesse sentido, reforçam-se como a ausência de tratamento de esgoto e de manejo de resíduos sólidos se configuram como os principais responsáveis pela elevação do IDSB, em síntese.

Gráfico 5: *Boxplots* da renda per capita municipal por categoria de déficit, segundo componente do IDSB



Fonte: Dados básicos: IBGE, 2010; FJP, 2021. Elaboração própria.

Gráfico 6: *Boxplots* IDHM por categoria do déficit, segundo componentes do IDSB



Fonte: PROGRAMA..., 2020; FUNDAÇÃO JOÃO PINHEIRO, 2021. Elaboração própria.

3 CONSIDERAÇÕES FINAIS

Esta nota técnica teve como objetivo, além de retomar algumas discussões teórico-conceituais sobre a construção do IDBS de Minas Gerais, apresentar uma análise dos dados obtidos e submetê-los a testes quantitativos mais refinados.

Dessa maneira, a partir da definição de *clusters* e da ACP, foi possível a delimitação de “cortes” de categorias de déficit (muito alto, alto, médio, baixo e muito baixo). Isso permitiu uma melhor comparabilidade, inclusive no tempo, da evolução dos indicadores, sejam eles voltados para o estado, sejam para os TS (bacias hidrográficas) e municípios, o que facilitou visão mais “qualitativa” nas distintas categorias definidas.

A aplicação da metodologia proposta apresentou resultados bastante promissores em relação aos objetivos iniciais, quando do projeto de criação do índice, tanto do ponto de vista espacial quanto, mais especificamente, da compreensão e interpretação de cada componente abordado.

Nessas análises chama a atenção o peso dos componentes do esgotamento, principalmente a falta de tratamento de esgoto e a disposição dos RSU, que, realmente, assumem grande peso na explicação da variabilidade do índice de déficit.

Quanto à distribuição espacial do IDSB, os resultados também foram coerentes com aspectos relacionados à renda, ao Índice de Desenvolvimento Humano (IDH) e à taxa de urbanização. Normalmente, os maiores déficits estão concentrados nas bacias/regiões mais ao norte e nordeste do estado. Por sua vez, os menores déficits estão distribuídos em municípios próximos à Região Metropolitana de Belo Horizonte, no Triângulo e Sul de Minas.

Finalmente, salienta-se que as análises sobre o desenvolvimento desse índice continuam. Um dos pontos que se pretende aprimorar, em futuro próximo, é a análise sobre a qualidade dos dados utilizados, principalmente no que se refere à elevada frequência de informações faltantes.

REFERÊNCIAS

- BRASIL. Ministério do Desenvolvimento Regional. **Diagnóstico dos serviços de água e esgotos**: 2019. Brasília, DF: SNIS, 2019. Disponível em: <http://www.snis.gov.br/diagnostico-anual-agua-e-esgotos/diagnostico-dos-servicos-de-agua-e-esgotos-2019>. Acesso em: 9 abr. 2021.
- DONI, Marcelo Viana. **Análise de cluster**: métodos hierárquicos e de particionamento. Trabalho de Graduação (Bacharelado em Sistemas de Informação) – Faculdade de Computação e Informática, Universidade Presbiteriana Mackenzie, São Paulo, 2004.
- EVERITT, Brian S. et al. **Cluster analysis**. 5th ed. New Jersey: Wiley, 2011.
- FUNDAÇÃO JOÃO PINHEIRO. Diretoria de Estatística e Informações. Índice deficit do saneamento básico em Minas Gerais. Belo Horizonte: FJP, 2021. Nota técnica. Disponível em: http://fjp.mg.gov.br/wp-content/uploads/2021/03/14.04_NotaTecnica_IDS_CHS.pdf. Acesso em: 30 ago. 2021.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **[Dados PNADC: visita 1: 2019]**. Rio de Janeiro: IBGE, 2020. Disponível em: https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html?caminho=Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilio_s_continua/Anual/Microdados/Visita. Acesso em: 26 mar. 2020.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo 2010**: glossário. Rio de Janeiro: IBGE, 2010. Disponível em: <https://censo2010.ibge.gov.br/materiais/guia-do-censo/glossario.html>. Acesso em: 21 julho, 2021.
- KAUFMAN, Leonard; ROUSSEEUW, Peter J. **Finding groups in data**: an introduction to cluster analysis. New York: J. Wiley, 1990.
- MANLY, Bryan F.J; ALBERTO, Jorge A. Navarro. **Multivariate statistical methods**: a primer. Boca Raton, FL: CRC Press, 2016.
- MINGOTI, Sueli Aparecida. **Análise de dados através de métodos estatística multivariada**: uma abordagem aplicada. Belo Horizonte: editora UFMG, 2020.
- PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO; INSTITUTO DE PESQUISA ECONÔMICA APLICADA; FUNDAÇÃO JOÃO PINHEIRO. **Atlas do Desenvolvimento Humano do Brasil**. Brasília, DF: Pnud Brasil, Ipea, FJP, 2020. Disponível em: <http://www.atlasbrasil.org.br/consulta>. Acesso em: 10 jun. 2021.
- SIMÕES, Rodrigo Ferreira. **Métodos de análise regional e urbana**: diagnóstico aplicado ao planejamento. Belo Horizonte: Cedeplar/UFMG, 2005.