

**Universidade de São Paulo
Instituto de Matemática e Estatística**

Centro de Estatística Aplicada

Relatório de Análise Estatística

RAE-CEA-25P14

RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:

“Epidemiologia do baixo Letramento em Saúde: contribuições da educação escolar”

**Cristine Airi Kimura
Gabriel Medeiros Cupini
Vicenzo Carlim de Sousa
Anatoli Iambartsev
Eduardo Janotti Cavalcante**

São Paulo, junho de 2025

CENTRO DE ESTATÍSTICA APLICADA - CEA – USP

TÍTULO: Relatório de Análise Estatística sobre o Projeto: “Epidemiologia do baixo Letramento em Saúde: contribuições da educação escolar”

PESQUISADORA: Luci Ana Santos da Cunha

ORIENTADORAS: Profa. Dra. Raquel Ajub Moyses e Profa. Dra. Tatiana Natasha Toporcov

INSTITUIÇÃO: Faculdade de Saúde Pública da Universidade de São Paulo (FSP-USP)

FINALIDADE DO PROJETO: Doutorado

RESPONSÁVEIS PELA ANÁLISE: Cristine Airi Kimura

Gabriel Medeiros Cupini

Vicenzo Carlim de Sousa

Anatoli Iambartsev

Eduardo Janotti Cavalcante

REFERÊNCIA DESTE TRABALHO: KIMURA, C.A.; CUPINI, G.M.; DE SOUSA, V.C.; IAMBARTSEV, A.; CAVALCANTE, E.J. **Relatório de análise estatística sobre o projeto: “Epidemiologia do baixo Letramento em Saúde: contribuições da educação escolar”**. São Paulo, IME-USP, 2025 (RAE-CEA-25P14).

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

- AGRESTI, A. (2013). **Categorical Data Analysis**. 3ª edição. Wiley, New York.
- AKAIKE, H. (1974). A new look at statistical model identification. **IEEE Transactions on Automatic Control**, AU-19:716-722.
- BÉLAND, S.; PICHETTE, F.; JOLANI, S. (2016). Impact on Cronbach's α of simple treatment methods for missing data. **Quant. Methods Psychology**, 12, 57–73.
- BRASIL (2013). Ministério da Educação. Secretaria de Educação Básica. Secretaria de Educação Continuada, Alfabetização, Diversidade e Inclusão. Conselho Nacional da Educação. **Diretrizes Curriculares Nacionais Gerais da Educação Básica**. Brasília, DF: MEC, SEB, DICEI.
- BUSSAB, W.O.; MORETTIN, P.A. (2024). **Estatística Básica**. 10.ed. São Paulo: Saraiva.
- CORTINA, J.M. (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. **J. Appl. Psychol**, 78, 98–104.
- COX, D.R. (1970). **The Analysis of Binary Data**. 1ª edição, Methuen, Londres.
- CRONBACH, L.J. (1951). Coefficient alpha and the internal structure of tests. **Psychometrika**, 16, 297–334.
- DA CUNHA, L.A.S. (2023). **Letramento em saúde: métrica para apoiar a educação escolar a ser tornar base para uma vida saudável. Projeto de pesquisa do programa de doutorado em Epidemiologia**. Faculdade de Saúde Pública USP.
- DE SOUZA, A.C.; ALEXANDRE, N.M.C.; GUIRARDELLO, E.B. (2017). Propriedades psicométricas na avaliação de instrumentos: Avaliação da confiabilidade e da validade. **Epidemiol. Serv. Saúde**, 26, 649–659.

FREIRE, P. (2007). **Pedagogia da autonomia: saberes necessários à prática educativa**. São Paulo: Paz e Terra.

IZBICKI, R.; DOS SANTOS, T.M. (2020). **Aprendizado de máquina: uma abordagem estatística**. 1ª edição, livro eletrônico.

KESZEI, A.P.; NOVAK, M.; STREINER, D.L. (2010). Introduction to health measurement scales. **J. Psychosom. Res.** 68, 319–323.

KRUSKAL, W.H.; WALLIS, W.A. (1952). Use of ranks in one-criterion variance analysis. **Journal of the American Statistical Association**, 47, 583–621.

KUSSUNOKI, E.Y.; AMORIM, M.R.; GIAMPAOLI, V. (2024). **Relatório de análise estatística sobre o projeto: “Letramento em saúde: métrica para apoiar a educação escolar a se tornar base para uma vida saudável”**. São Paulo, IME-USP, 2024 (RAE – CEA – 24P05).

MATTA, G.C., et al. eds. **Os impactos sociais da Covid-19 no Brasil: populações vulnerabilizadas e respostas à pandemia**. Rio de Janeiro: Observatório Covid 19; Fiocruz, 2021. Disponível em: <https://books.scielo.org/id/r3hc2>. Acesso em: 30 abr. 2025.

MIALHE, F.L.; SAMPAIO, H.A.C.; MORAES, K.L.; BRASIL, V.V.; REBUSTINI, F. (2022). Propriedades psicométricas da versão brasileira do European Health Literacy Questionnaire. **Promoção da Saúde Internacional**, 37, 1-10, Oxford.

PELIKAN, J.M.; GANAHL, K.; VAN DEN BROUCKE, S.; SØRENSEN, K. (2019). Measuring Health Literacy in Europe: Introducing the European Health Literacy Survey Questionnaire (HLS-EU-Q). **International Handbook of Health Literacy**. Bristol: Policy Press PP - Bristol. p. 115–38.

PEDRO, A.R.; RAPOSO, B.; LUÍS, L.; AMARAL, O.; ESCOVAL, A.; SIMÕES DIAS, S. (2023). Portuguese version of the HLS-EU-Q6 and HLS-EU-Q16 Questionnaire: Psychometric Properties. **Int. J. Environ. Res. Public Health**.

- REYNOLDS, D. (2009). Gaussian Mixture Models. In: Li, S.Z., Jain, A. (eds) **Encyclopedia of Biometrics**. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-73003-5_196.
- RIGBY, R.A.; STASINOPOULOS, D.M. (2005). Generalized additive models for location, scale and shape (with discussion). **Applied Statistics**, 54, 507-554.
- RIGBY, R.A.; STASINOPOULOS, D.M. (2006). Using the Box–Cox t distribution in GAMLSS to model skewness and kurtosis. **Stat Model**, 6:209–229.
- SCHEAFFER, R.L.; MENDENHAL III, W.; OTT, R.L.; GEROW, K.G. (2011). **Elementary Survey Sampling**. Cengage Learning.
- SHAPIRO, S.S.; WILK, M.B. (1965). An analysis of variance test for normality (complete samples). **Biometrika**. 52 (3–4): 591–611.
- SØRENSEN, K.; VAN DEN BROUCKE, S.; FULLAM, J.; DOYLE, G.; PELIKAN, J.; SLONSKA, Z. (2012). Alfabetização em saúde e saúde pública: uma revisão sistemática e integrada de definições e modelos. **BMC Saúde Pública**, 12, 1-13.
- SPEARMAN, C. (1904). The Proof and Measurement of Association between Two Things. **Am. J. Psychol.** 15, 72–101.
- STEVANIM, L.F. (2020). **Nós: os vulneráveis**. RADIS. Revista da Fundação Oswaldo Cruz, n. 212.
- TANIGUCHI, M.; HIRUKAWA, J. (2012). Generalized information criterion. **Journal of Time Series Analysis**, 33: 287-297. <https://doi.org/10.1111/j.1467-9892.2011.00759.x>.
- TERWEE, C.B.; BOT, S.D.M.; DE BOER, M.R.; VAN DER WINDT, D.A.W.M.; KNOL, D.L.; DEKKER, J.; BOUTER, L.M.; DE VET, H.C.W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. **J. Clin. Epidemiol**, 60, 34–42.
- VAN BUUREN, S.; FREDRIKS, M. (2001). Worm plot: a simple diagnostic device for modelling growth reference curves. **Stat Med**. 20(8):1259-77.

WILLIAMS, D.A. (1984). Residuals in generalized linear models. **Proceedings of the 12th. International Biometrics Conference**, Tokyo, pp. 59-68.

WORLD HEALTH ORGANIZATION (WHO) (1998). **Health Promotion Glossary**. Geneva: WHO.

WORLD HEALTH ORGANIZATION (WHO) (2016). **9a Global Conference on Health Promotion. Policy brief 4: Health Literacy**. Geneva: WHO.

PROGRAMAS COMPUTACIONAIS UTILIZADOS:

Microsoft Word para Microsoft 365 MSO (versão 2503)

Microsoft Excel para Microsoft 365 MSO (versão 2503)

R Studio para Windows (versão 2024.04.2, Build 764)

TÉCNICAS ESTATÍSTICAS UTILIZADAS

Análise Descritiva Unidimensional (03:010)

Análise Descritiva Multidimensional (03:020)

Testes de Hipóteses Não Paramétricas (05:070)

Outros: GAMLSS (07:990)

Outros: Árvores de regressão (07:990)

Regressão Logística (07:090)

ÁREA DE APLICAÇÃO

Outros: Saúde Pública (14:990)

Resumo

O letramento em saúde (LS) diz respeito à apropriação e ao uso social de conceitos sobre saúde no cotidiano e é essencial para que os cidadãos possam entender, avaliar, aplicar e promover informações relacionadas à saúde. Indivíduos com baixo nível de LS têm dificuldade em usufruir dos serviços de saúde, compreender as informações fornecidas pelos profissionais de saúde e podem ter suas condições de vida prejudicadas. Uma vez que o letramento em saúde deve ser desenvolvido ainda na vida escolar, é de grande interesse identificar os fatores que impactam o nível de LS dos profissionais da educação, em especial, a fim de planejar intervenções de saúde pública e buscar capacitação dos profissionais em Educação para a Saúde (ES).

O questionário europeu de literacia em saúde (HLS-Q16) é um recurso bastante utilizado para medir o letramento em saúde. O objetivo deste estudo é descrever o nível de LS dos gestores e professores da educação básica da rede municipal de ensino da cidade de São Paulo, além de analisar os fatores sociodemográficos e profissionais associados às suas condições de LS. A amostra consiste em 1015 profissionais da educação e os dados foram coletados por meio da aplicação da versão brasileira do HLS-Q16 e de questionários sociodemográfico e de autoeficácia do profissional.

Além da análise descritiva dos perfis de letramento em saúde, foram construídos um modelo de regressão linear generalizado com distribuição Box-Cox t , uma árvore de regressão e um modelo de mistura Gaussiana complementado por um modelo de regressão logística. Foi observado que alguns fatores sociodemográficos, como o tipo de educação básica; profissionais, como o cargo e a satisfação no trabalho; e pedagógicos, como a percepção sobre o nível de ES dos alunos, a auto apropriação de conceitos de ES e a formação sobre informações de saúde, apresentam efeitos significativos em relação ao letramento.

Concluiu-se que os profissionais com satisfação frequente em relação ao trabalho ou que receberam formação em ES apresentam maior letramento, bem como aqueles que trabalham em escolas onde há incentivo à formação e discussão sobre saúde. Esses resultados podem influenciar a proposta de políticas educacionais,

aumentar os investimentos em educação, promover um debate sobre o estreitamento das relações entre escola e saúde, enfatizando o papel da escola na promoção da saúde, e discutir os determinantes sociais da condição de saúde dos profissionais do sistema educacional público.

Sumário

1. Introdução.....	12
2. Objetivos.....	13
3. Descrição do estudo	14
3.1. Cálculo do tamanho amostral.....	14
3.2. Amostra e coleta dos dados	15
4. Descrição das variáveis	19
4.1. Variáveis geradas pelo questionário HLS-Q16 e variáveis respostas.....	19
4.2. Variáveis explicativas geradas pelo questionário sociodemográfico.....	21
4.3. Variáveis explicativas geradas pelo questionário de autoeficácia em Educação para a Saúde	22
4.3.1. Variáveis relacionadas à formação docente.....	23
4.3.2. Variáveis relacionadas à transposição didática	24
5. Análise descritiva	26
5.1. Análise descritiva unidimensional.....	26
5.1.1. Respostas do questionário HLS-Q16.....	26
5.1.2. Variável de interesse Letramento.....	27
5.1.3. Variável de interesse Escore.....	28
5.1.4. Dados sociodemográficos	28
5.1.5. Perfil pedagógico-profissional	29
5.2. Análise descritiva multidimensional	31
5.2.1. Variável de interesse Letramento.....	31
5.2.2. Variável de interesse Escore.....	35
5.3. Validação do questionário e consistência interna	37

6. Análise inferencial.....	38
6.1. Modelo GAMLSS	39
6.2. Árvore de regressão.....	42
6.3. Modelo de mistura Gaussiana.....	43
6.3.1. Modelo de regressão logística	45
7. Conclusão.....	47
APÊNDICE A.....	48
APÊNDICE B.....	75
APÊNDICE C.....	94
ANEXO	98

1. Introdução

O Brasil é um país de grande extensão territorial e com muita desigualdade social. Segundo estudo lançado mundialmente pelo *World Inequality Lab*, o Brasil permanece como um dos países com maior desigualdade social e de renda do mundo. Nos anos de 2020 e 2021, essas desigualdades foram acentuadas pela pandemia do novo coronavírus 2019 (Covid-19) e pelo necessário isolamento social (Stevanim, 2020; Matta et al., 2021).

Durante este período, os epidemiologistas ganharam maior relevância, pois foram responsáveis por darem explicações, além de realizarem análises estatísticas, orientações e recomendações. Em contrapartida, foi também uma época de bombardeio de informações falsas e infodemia, crescimento do movimento antivacina e de falta de crítica por parte da população para entender as informações veiculadas. Esse cenário evidenciou a importância do letramento em saúde e a necessidade de formação adequada em relação aos saberes que ele abrange.

Alfabetização em saúde ou letramento em saúde é um campo do conhecimento que está associado à capacidade e às habilidades das pessoas para acessar, compreender, avaliar e aplicar informações em saúde (Mialhe et al., 2022; WHO, 1998; WHO, 2016). O domínio destas competências possibilita às pessoas fazerem julgamentos e tomarem decisões cotidianas sobre cuidados de saúde, prevenção de doenças e promoção da saúde para manter ou melhorar a qualidade de vida (Sørensen et al., 2012).

Os indivíduos com baixo LS apresentam dificuldades para buscar e usufruir ao máximo dos serviços de saúde. Quando têm acesso à saúde, eles demonstram menor probabilidade de compreensão da informação, tanto escrita quanto oral, fornecida pelos profissionais de saúde. Logo, um letramento inadequado em saúde está associado a piores condições gerais de saúde e a uma baixa utilização de serviços de prevenção de doenças, impactando a qualidade de vida dos indivíduos.

É notório que o processo de aprendizagem dos estudantes, sejam crianças ou adolescentes, não depende apenas da metodologia de ensino promovida, mas também

da própria condição de saúde e qualidade de vida dos alunos. Nesse contexto, a escola revela-se como um espaço de LS e promoção de saúde e, portanto, faz-se mister uma discussão a respeito da mobilização em torno do direito à saúde, do fortalecimento da participação da comunidade escolar nos espaços de controle social e de gestão participativa do Sistema Único de Saúde (SUS) no sistema educacional.

Em vista disso, é de interesse saber o nível de LS dos responsáveis pela educação, como professores e gestores das escolas. Entre os instrumentos atualmente existentes para medir o LS, destaca-se a versão curta do *European Health Literacy Survey Questionnaire* de 16 questões (HLS-Q16) (Pelikan et al., 2019), originalmente um questionário de 47 questões, desenvolvido por um consórcio europeu de instituições de pesquisa. Ele é baseado na definição de letramento em saúde proposta pela Organização Mundial de Saúde. O HLS é um instrumento internacionalmente reconhecido por sua robustez metodológica e grande aplicabilidade clínica, validado em mais de 15 idiomas, apresentando excelentes resultados nas análises de validação e confiabilidade.

2. Objetivos

O objetivo principal do estudo é descrever e analisar o grau de letramento em saúde de gestores e professores da rede pública municipal de ensino de São Paulo, avaliando os possíveis fatores associados ao nível de letramento em saúde dos profissionais. Para atingir esse objetivo, o estudo visa adquirir um perfil geral dos respondentes, investigando se existe associação entre o letramento e variáveis sociodemográficas e da vida profissional dos professores e gestores, além de indicar a direção e o significado dessas associações. Ademais, vale destacar que o perfil profissional será caracterizado pela formação e pela capacidade de transposição didática dos docentes e educadores.

Com base nisso, os resultados desse estudo podem contribuir para o planejamento de políticas públicas na área da educação, como a construção e revisão de currículos e programas escolares, a partir das respostas obtidas, além de justificar

investimentos na formação inicial e continuada de gestores e docentes e enfatizar o protagonismo da escola na discussão a respeito dos determinantes sociais da saúde e da doença no espaço que envolve a comunidade escolar (Brasil, 2013; Freire, 2007; Da Cunha, 2023).

Dessarte, o desenvolvimento de habilidades e competências que possibilitam a aprendizagem é um direito garantido a todas as pessoas. Os educadores têm a responsabilidade de orientar e promover esse processo; a escola representa o ambiente ideal para tornar esse acesso ao conhecimento mais amplo e acessível a todos, assegurando que a educação seja realmente inclusiva e democrática (Brasil, 2013; Freire, 2007; Da Cunha, 2023).

3. Descrição do estudo

O estudo é do tipo transversal analítico com indivíduos que ocupam os cargos de gestor ou professor da educação básica na rede municipal de ensino da cidade de São Paulo, selecionados aleatoriamente. O período de coleta de dados, provenientes dos questionários respondidos, foi entre março de 2023 e setembro de 2024. O planejamento amostral bem como os questionários coletados serão explicitados nas seções subsequentes.

3.1. Cálculo do tamanho amostral

O delineamento amostral foi proposto pela Prof^a. Dra. Daiane Leite da Roza, do Departamento de Epidemiologia da Faculdade de Saúde Pública da Universidade de São Paulo (FSP-USP), e encontra-se no Anexo 1.

A cidade de São Paulo conta com 1.507 unidades educativas municipais, divididas em 13 Diretorias Regionais de Saúde (DRE) espalhadas por toda a cidade de São Paulo, compreendendo os cinco distritos regionais (Norte, Sul, Leste, Oeste e Central). As unidades escolares são divididas quanto à sua categoria, sendo: Educação Infantil (CEI, EMEI e CEMEI) e Ensino Fundamental e Médio (EMEF e EMEFM). As

unidades escolares foram agrupadas em 26 estratos de acordo com a diretoria regional a que essa unidade pertence, bem como o seu tipo.

Em suma, o tamanho amostral foi obtido considerando um plano amostral estratificado (Scheaffer et al., 2011), um coeficiente de confiança de 95% e uma precisão absoluta de 3% para a proporção de professores e gestores. Estimou-se um tamanho amostral de 1.056 entrevistas, sendo o número de entrevistas a ser realizada em cada estrato proporcional ao respectivo tamanho populacional.

3.2. Amostra e coleta dos dados

Para a composição da amostra, foram excluídos os indivíduos que estavam em gozo de licenças de qualquer natureza. Aqueles que não concordaram com a participação na pesquisa foram substituídos. Vale ressaltar que todos os profissionais podiam participar da pesquisa, mas, para composição da amostra, foram apenas selecionados os que haviam sido sorteados.

Foi selecionada uma amostra aleatória de 1.015 profissionais da educação (professores e gestores), com idades maiores ou iguais a 18 anos, recolhida em 31 escolas.

A coleta de dados foi realizada por meio da aplicação *on-line* e presencial de três questionários em conjunto:

- O *European Health Literacy Survey* (HLS-Q16) adaptado para o português (Anexo 2);
- Questionário sociodemográfico (Anexo 3);
- Questionário de autoeficácia do profissional: autodeclarado e elaborado pela pesquisadora, contém perguntas referentes à formação docente e à transposição didática (Anexo 4).

Os questionários totalizam 57 questões, distribuídas em: 16 questões do HLS-Q16, 11 questões do questionário sociodemográfico e 30 questões do questionário de autoeficácia do profissional. Eles foram submetidos ao Comitê de Ética em Pesquisa da

Faculdade de Saúde Pública da Universidade de São Paulo e foram aprovados. Além disso, os participantes tiveram de assinar um Termo de Consentimento Livre e Esclarecido durante o seu preenchimento.

As 16 questões do questionário HLS-Q16 são em escala Likert de 4 pontos (Likert, 1932). Cada questão admite cinco respostas: muito fácil, fácil, difícil, muito difícil e não sei. Este instrumento de avaliação identifica duas dimensões do letramento em saúde explicadas a seguir.

A primeira dimensão refere-se às quatro competências que um indivíduo deve ter para utilizar a informação de maneira a promover e manter uma boa saúde, a saber:

- Acesso, ou seja, ter capacidade para procurar, encontrar e obter informação relacionada à saúde;
- Compreensão, ou seja, ter capacidade para entender a informação procurada;
- Avaliação, ou seja, ter capacidade de selecionar e avaliar a informação procurada;
- Utilização, ou seja, ter capacidade para comunicar e utilizar a informação encontrada.

A segunda dimensão identifica que a informação sobre saúde está dividida em três domínios, a saber:

- Cuidados da saúde (CS), representado pelas questões de 1 a 7;
- Prevenção da doença (PD), representado pelas questões de 8 a 12;
- Promoção da saúde (PS), representado pelas questões de 13 a 16.

O Quadro 1 abaixo apresenta a matriz das dimensões de letramento em saúde baseada no modelo conceitual do HLS-Q16, utilizada na construção do questionário.

Quadro 1: Matriz das dimensões de LS. Adaptado de Sørensen et al., 2012.

	Acessar	Compreender	Avaliar	Aplicar
Cuidados da saúde	1) Capacidade de acesso à informação relacionada com problemas médicos ou clínicos.	2) Capacidade de compreensão da informação médica e do seu significado.	3) Capacidade de interpretação e da avaliação de informações médicas.	4) Capacidade de tomar decisões informadas sobre questões médicas.
Prevenção da doença	5) Capacidade de acesso à informação sobre fatores de risco.	6) Capacidade de compreensão dos fatores de risco e do seu significado.	7) Capacidade de interpretação e de avaliação de informações relacionadas com os fatores de risco.	8) Capacidade de julgar a relevância das informações sobre fatores de risco.
Promoção da saúde	9) Capacidade de atualização sobre questões de saúde.	10) Capacidade de compreensão da informação relacionada com a saúde e do seu significado.	11) Capacidade de interpretação e de avaliação de informações sobre questões relacionadas com a saúde.	12) Capacidade de tomar uma opinião consciente sobre questões de saúde.

Em cada domínio do questionário, há questões que abordam as quatro competências da primeira dimensão. Portanto, este instrumento de avaliação pode ser visto como uma matriz de competências com quatro grupos de questões, representados, por domínio, no Quadro 2.

Quadro 2: Matriz de competências no questionário HLS-Q16.

	Acessar	Compreender	Avaliar	Aplicar
Cuidados da saúde	Questões 1 e 2	Questões 3 e 4	Questão 5	Questões 6 e 7
Prevenção da doença	Questão 8	Questões 9 e 10	Questão 11	Questão 12
Promoção da saúde	Questão 13	Questões 14 e 15	Questão 16	

A partir do questionário HLS-Q16, cria-se um escore que representa o nível de letramento de um indivíduo ou de um grupo de indivíduos (amostra) em relação a um dos domínios ou em relação ao LS global. Para cada resposta possível, é atribuído um peso: Muito difícil (peso 1), Difícil (peso 2), Fácil (peso 3) e Muito fácil (peso 4). As respostas “Não sei” não são levadas em consideração no cálculo do escore. Para avaliar corretamente o nível de LS dos respondentes, eles precisam ter respondido no mínimo 14 das 16 questões. Por esse motivo, sete respondentes foram substituídos por outros, pois eles apresentavam mais do que duas respostas Não sei ou Em branco. Para os indivíduos que responderam exatamente 14 ou 15 questões, o cálculo do escore, que será descrito abaixo, é análogo; entretanto, a quantidade de questões que será considerada corresponde ao total de questões respondidas, e não ao total de 16 questões.

Para obter o escore de um indivíduo, calcula-se a pontuação parcial somando os pesos aplicados às respostas dadas por ele. Então, calcula-se a média dividindo a pontuação parcial pela quantidade de questões do domínio de interesse ou pela quantidade de questões do questionário.

A fim de garantir uma comparação efetiva entre os domínios, os escores são padronizados em uma escala métrica de 0 a 50, por meio da fórmula abaixo:

$$\text{Escore} = \frac{\text{média} - 1}{3} \times 50$$

$$\text{Média} = \frac{\text{pontuação parcial}}{\text{quantidade de questões}}$$

Por exemplo, um indivíduo que tenha respondido todas as questões com Muito fácil, receberá nota 4 para todas elas. Sua pontuação parcial será $16 \times 4 = 64$; por sua vez, sua média será $\frac{64}{16} = 4$. Logo, seu escore global será $\frac{(4-1)}{3} \times 50 = \frac{3}{3} \times 50 = 50$. Se um indivíduo respondeu todas as questões com Fácil, receberá nota 3 para todas elas. Sua pontuação parcial será $16 \times 3 = 48$; por sua vez, sua média será 3 e seu escore global será 33,33.

4. Descrição das variáveis

4.1. Variáveis geradas pelo questionário HLS-Q16 e variáveis respostas

As variáveis abaixo estão atreladas ao questionário HLS-Q16 (descrito no Anexo 2) e apresentam as categorias: Muito difícil, Difícil, Fácil, Muito fácil e Não sei.

- **Q1:** Com que facilidade você consegue encontrar informações sobre os tratamentos de doenças que preocupam você?
- **Q2:** Com que facilidade você consegue descobrir onde conseguir ajuda profissional quando está doente? (p.ex., um médico, farmacêutico, psicólogo)
- **Q3:** Com que facilidade você consegue entender o que o seu médico diz a você?
- **Q4:** Com que facilidade você consegue entender as instruções do seu médico ou farmacêutico sobre como tomar um medicamento (remédio) que foi receitado para você?
- **Q5:** Com que facilidade você consegue avaliar quando você precisa de uma segunda opinião de outro médico?
- **Q6:** Com que facilidade você consegue usar as informações que seu médico passa a você para tomar decisões sobre a sua doença?
- **Q7:** Com que facilidade você consegue seguir as instruções do seu médico ou farmacêutico?
- **Q8:** Com que facilidade você consegue encontrar informações sobre como lidar com problemas de saúde mental, como o estresse ou depressão?

- **Q9:** Com que facilidade você consegue entender os avisos sobre comportamentos que prejudicam a saúde, tais como fumar, praticar pouca atividade física e o consumo excessivo de álcool?
- **Q10:** Com que facilidade você consegue entender por que você precisa fazer exames periódicos de saúde? (p.ex., exame de mamas, teste de açúcar no sangue, pressão arterial)
- **Q11:** Com que facilidade você consegue avaliar se as informações sobre os riscos à saúde disponíveis nos meios de comunicação são confiáveis? (p.ex., TV, *internet* ou outros meios de comunicação)
- **Q12:** Com que facilidade você consegue decidir como você pode se proteger de uma doença com base nas informações dos meios de comunicação? (p.ex., jornais, *internet* ou outros meios de comunicação)
- **Q13:** Com que facilidade você consegue encontrar informações sobre as atividades que são boas para o seu bem-estar mental? (p.ex., meditação, exercício, caminhada, pilates etc.)
- **Q14:** Com que facilidade você consegue entender os conselhos de saúde que os membros da sua família ou seus amigos dão a você?
- **Q15:** Com que facilidade você consegue entender as informações disponíveis nos meios de comunicação sobre como ficar mais saudável? (p.ex., *internet*, jornais, revistas)
- **Q16:** Com que facilidade você consegue avaliar que comportamentos do seu dia a dia afetam a sua saúde? (p.ex., consumo de bebidas e hábitos alimentares, exercícios etc.)

Por meio da obtenção dos escores a partir do questionário HLS-Q16, definem-se duas variáveis respostas de interesse:

- **Escore:** índice associado ao grau de letramento em saúde do profissional (subdividido em **global**, envolvendo todas as questões, e por **domínio**).
- **Letramento:** Inadequado, Problemático, Suficiente, Excelente (subdividido em **global**, envolvendo todas as questões, e por **domínio**).

Segundo Pedro et al. (2023), a classificação do letramento segundo o valor do escore é feita da seguinte forma: Inadequado (0, 25]; Problemático (25, 33]; Suficiente (33; 42]; e Excelente (42, 50].

4.2. Variáveis explicativas geradas pelo questionário sociodemográfico

A partir do questionário descrito no Anexo 3, para se investigar a associação entre o letramento e o perfil sociodemográfico consideram-se as seguintes variáveis que caracterizam a amostra:

- **Idade** (anos);
- **Faixa etária:** 18-29 anos, 30-35 anos, 36-41 anos, 42-47 anos, 48-53 anos, 54-59 anos, 60-65 anos, 66-75 anos;
- **Gênero:** Feminino, Masculino, Não binário;
- **Cor/Raça:** Amarela, Branca, Indígena, Parda, Preta;
- **Cargo:** Assistente de diretor de escola, Coordenador pedagógico, Diretor de escola, Professor de educação infantil, Professor de educação infantil e ensino fundamental I, Professor de ensino fundamental II e médio;
 - **Disciplina:** Se for Professor de ensino fundamental II e médio, qual disciplina leciona? Artes, Ciências, Educação Física, Geografia, História, Inglês, Matemática, Português;
- **Função:** Gestor, Professor;
- **Tempo de trabalho:** Menos de 3 anos, 3 a 4 anos, 5 a 10 anos, 11 a 15 anos, 16 a 20 anos, 21 a 25 anos, Mais de 25 anos;
- **Escolaridade:** Ensino médio (Magistério), Graduação, Pós-graduação (*Lato Sensu* / Especialização), Mestrado, Doutorado, Pós-doutorado;
- **Tipo de educação básica:** tipo de escola frequentada na educação básica, Pública, Privada, Pública e privada;
- **Tipo de graduação:** tipo de escola frequentada na graduação, Pública, Privada, Pública e privada;

- **Renda:** 2 a 3 salários mínimos, 4 a 5 salários mínimos, 6 a 7 salários mínimos, 8 a 9 salários mínimos, 10 salários mínimos ou mais.
- **Escola:** A (EMEF Ten. Aviador Frederico Gustavo dos Santos), B (EMEI Ministro Bilac Pinto), C (EMEI Rosa e Carolina Agazzi), D (EMEF João Ramos – Pernambuco – Abolicionista), E (EMEI Luís Gama), F (CEI CEU Jaguaré), G (CEU EMEF Perus), H (EMEF Desembargador Amorim Lima), I (EMEI Professor Benedicto Castrucci), J (EMEI Assis Chateaubriand), K (EMEI Catulo da Paixão Cearense), L (EMEF General de Gaulle), M (EMEF Jornalista Millôr Fernandes), N (CEU CEI Parelheiros), O (EMEF Professora Marina Melander Coutinho), P (EMEF Prof^a. Liliane Verzini Silva), Q (EMEI Cruz e Sousa), R (CEI Silvia Covas), S (CEI Vereador José de Moura), T (CEI Prof^a. Maria Augusta de Paula), U (EMEF Vladimir Herzog), V (EMEF Benedito Calixto), W (EMEI Prof. Alberto Mesquita de Camargo), X (EMEF Espaço de Bitita), Y (EMEF Guilherme de Almeida), Z (EMEI Alfredo da Rocha Viana Filho – Pixinguinha), AA (EMEF Coelho Neto), AB (EMEI Alfredo Volpi), AC (CEU CEI Parque São Carlos), AD (CEU EMEF Vila Curuçá), AE (EMEF Padre José de Anchieta);
 - **DRE:** Diretoria Regional de Educação, Butantã, Campo Limpo, Capela do Socorro, Freguesia/Brasilândia, Guaianases, Ipiranga, Itaquera, Jaçanã/Tremembé, Penha, Pirituba/Jaraguá, Santo Amaro, São Mateus, São Miguel;
 - **Zona:** Norte, Sul, Leste, Oeste.

4.3. Variáveis explicativas geradas pelo questionário de autoeficácia em Educação para a Saúde

Para investigar a associação entre o letramento e o perfil pedagógico-profissional (formação docente e transposição didática) a partir do questionário do Anexo 4, consideram-se as variáveis listadas a seguir:

4.3.1. Variáveis relacionadas à formação docente

- **Q5.1:** Você está satisfeito(a) com sua contribuição na escola? Sim, frequentemente; Sim, mas estou sentindo que minha satisfação tem diminuído; Não;
- **Q5.2:** Se pudesse, escolheria novamente ser profissional da educação? Sim, Não;
 - **Q5.2a:** Se você respondeu não, relacione o motivo. Salários baixos e sobrecarga de trabalho, Indisciplina e dificuldade com alunos e famílias, Más condições de trabalho, Desgaste emocional e físico, Desvalorização e falta de reconhecimento;
- **Q6.1:** Na sua formação escolar, você recebeu formação sobre Educação para a Saúde? Sim, Não;
 - **Q6.1a:** Se você respondeu sim, relacione o(s) tema(s) trabalhado(s). Alimentação saudável, Doenças e patologias, Educação Física e Ciências Biológicas, Saúde e educação, Outros;
- **Q6.2:** Você já participou de formação continuada sobre Educação para a Saúde? Sim, Não;
 - **Q6.2a:** Se você respondeu sim, quais foram a(s) instituição(ões) que ofereceu(ram) o curso? Sindicato/Entidades sindicais, Programas de saúde na escola, Secretaria Municipal de Educação, Instituições de ensino superior, Formações religiosas ou comunitárias, Outros;
- **Q9.2:** Os educadores recebem formação sobre notícias/informações de saúde veiculadas nas redes sociais? Sim, Não;
 - **Q9.2a:** Se você respondeu sim, relacione ações. Ações da UBS e parcerias de saúde, Discussões e reflexões coletivas, Informações sobre alimentação saudável e propaganda, Informações sobre dengue e prevenção de doenças específicas, Informações sobre campanhas de vacinação e saúde pública, Informações gerais e comunicados oficiais, Formações sobre *fake news* e combate à desinformação, Formações e momentos pedagógicos, Outros;

- **Q9.3:** Os educadores recebem formação sobre como encontrar um equilíbrio entre cumprir o papel profissional e seus requisitos e prestar atenção nos estudantes e responder às suas necessidades de saúde? Sim, Não;
 - **Q9.3a:** Se você respondeu sim, relacione ações. Formação de professores, Reuniões pedagógicas, Observação e identificação de questões de saúde dos alunos, Parcerias externas, Práticas de cuidado e promoção de saúde.

4.3.2. Variáveis relacionadas à transposição didática

- **Q6.3:** No Projeto Político Pedagógico (PPP) da escola em que trabalha constam temas sobre Educação para a Saúde? Sim, Não, Não conheço o PPP da escola em que trabalho;
- **Q6.4:** A escola em que você trabalha participa do Programa Saúde na Escola ou Programa Primeira Infância? Sim; Sim, Programa Saúde na Escola; Sim, Programa Primeira Infância; Não participa;
- **Q7.1:** No período de planejamento são discutidas na escola em que trabalha questões sobre Educação para a Saúde, para constar nos planos? Sim, Não;
- **Q7.2:** No seu plano de ensino ou plano de trabalho, constam temas voltados à Educação para a Saúde? Sim, Não;
 - **Q7.2a:** Se você respondeu sim, relacione o(s) tema(s) que consta(m). Ações intersetoriais e parcerias, Alimentação saudável, Higiene pessoal e saúde bucal, Saúde mental e emocional, Vacinação e prevenção de doenças;
- **Q8.1:** Você considera que seus alunos (se professor) ou os alunos da escola (outros cargos) possuem uma boa Educação para a Saúde? Sim, Não;
- **Q8.2:** Você realiza atividades que envolvam a Educação para a Saúde? Sim, Não;
 - **Q8.2a:** Se você respondeu sim, relacione qual(is) atividade(s). Alimentação saudável, Atividades físicas, Autocuidado e bem-estar,

Conversas sobre saúde, Convivência e outros temas, Educação ambiental, Higiene pessoal e bucal, Prevenção de doenças;

- **Q8.3:** Você considera que se apropriou dos conteúdos específicos e pedagógicos necessários para realizar ações de Educação para a Saúde no seu trabalho, ou seja, se sentiu seguro(a) para trabalhar com os conteúdos sobre Educação para a Saúde? Sim, Não;
 - **Q8.3a:** Se você respondeu não, relacione o(s) motivo(s). Comentários variados, Conhecimento superficial/Temas abordados superficialmente, Desconhecimento completo do tema, Falta de apoio e recursos, Falta de formação específica, Formação básica não abrangente, Necessidade de aperfeiçoamento/Interesse em mais formação;
- **Q8.4:** Seu aluno (se professor) ou o aluno da escola (outros cargos) é capaz de se comunicar com um adulto ou profissional da saúde sobre comportamentos de saúde? (Por ex., sobre escovação de dentes, alimentação saudável). Sim, Não;
- **Q8.5:** O que falta para você trabalhar melhor (se já trabalha) ou iniciar o trabalho com Educação para a Saúde? Currículos e programas indicando conteúdos a serem trabalhados; Formação (cursos) sobre conteúdos, metodologias, atividades; Materiais didáticos (livros, apostilas etc.) para educadores e estudantes; Não falta nada, não trabalho com Ed. para a Saúde porque não disponho de tempo ou não considero necessário; Não falta nada e com os recursos de que disponho já trabalho com Educação para a Saúde;
- **Q9.1:** Os familiares/responsáveis pelos estudantes têm a oportunidade de conversar com os profissionais da escola sobre Educação para a Saúde? Sim, Não;
 - **Q9.1a:** Se você respondeu sim, relacione os momentos. Abertura da escola para comunicação com famílias, Atendimento individualizado e orientação direta, Dia da Família e eventos na escola, Formação e palestras sobre saúde, Reuniões com famílias e responsáveis, Outros;
- **Q9.4:** No geral, a infraestrutura do seu local de trabalho é: Adequada, Inadequada;

- **Q9.4a:** Caso considere necessário, justifique. Espaço físico insuficiente ou inadequado, Falta de acessibilidade, Higiene ou limpeza inadequadas, Necessidade de melhoria na infraestrutura geral, Problemas de acústica, Reformas ou necessidade de ajustes, Salas com superlotação, Outros;
- **Q10.1:** Indique os temas relacionados à Educação para a Saúde que você normalmente trabalha, tendo como referência o Currículo da Cidade. Alimentação saudável, Atividades físicas, Cuidado com o meio ambiente, Educação socioemocional, Higiene, Prevenção de doenças, Saúde bucal, Saúde mental, Vacinação, Outros.

5. Análise descritiva

Nesta seção, apresentamos a análise descritiva dos dados, que nos permite ter uma visão inicial dos resultados do estudo (Bussab e Morettin, 2024).

5.1. Análise descritiva unidimensional

5.1.1. Respostas do questionário HLS-Q16

A Tabela A.1 apresenta a distribuição das respostas das questões do questionário HLS-Q16. É importante destacar que nenhum participante foi excluído do cálculo do escore, pois não houve excessos de respostas Não sei ou Em branco. As porcentagens de respostas Não sei/Em branco variaram de 0% a 0,3%. As questões 9 e 10 apresentaram as maiores porcentagens de respostas Muito fácil (37,7% e 36,5%, respectivamente); ambas fazem parte do domínio de prevenção da doença. Por sua vez, os itens 8 e 2 tiveram as maiores porcentagens de respostas Muito difícil (8,4% e 4,7%, respectivamente). O item 8 faz parte do domínio de prevenção da doença, enquanto o item 2 integra o domínio de cuidados da saúde.

Observa-se na Figura B.1 que a maioria das respostas são Fácil e Muito fácil para todas as questões. Principalmente as questões 3 e 4 apresentam maior porcentagem de respostas Fácil e Muito fácil, sugerindo que as pessoas não têm

dificuldade em entender o que o médico diz e entender as instruções do médico ou farmacêutico sobre como tomar um medicamento que foi receitado.

Na Figura B.2, nota-se que a maioria das respostas são Fácil e Muito fácil para todas as questões. Principalmente as questões 9 e 10 apresentam maior porcentagem de respostas Fácil e Muito fácil, sugerindo que as pessoas não têm dificuldade em entender os avisos sobre comportamentos que prejudicam a saúde e entender por que é preciso fazer exames periódicos de saúde. Em contrapartida, a questão 8 apresenta maior proporção de respostas Difícil e Muito difícil, o que indica que as pessoas têm dificuldade em encontrar informações sobre como lidar com problemas de saúde mental.

Na Figura B.3, pode-se observar que a maioria das respostas são Fácil e Muito fácil para todas as questões, que possuem proporções de respostas semelhantes.

5.1.2. Variável de interesse Letramento

De acordo com a Tabela A.2, considerando os pontos de corte sugeridos para a classificação dos níveis de LS, a maioria dos participantes apresenta letramento global Suficiente (40,3%), seguido do nível Problemático (36%), Excelente (13,7%) e, em menor número, Inadequado (10%).

A partir da Figura B.4, pode-se observar que cerca de 46% dos participantes apresentam letramento global Inadequado ou Problemático. No que se refere aos cuidados da saúde, 57% dos participantes apresentam letramento Suficiente ou Excelente. No que diz respeito à prevenção da doença, 60% revelam ter letramento Suficiente ou Excelente. É no domínio promoção da saúde que se destacam os melhores resultados, visto que 75% apresentam letramento Suficiente ou Excelente.

Na Tabela A.3, apresentam-se a média, a mediana, o desvio padrão e o coeficiente de variação do escore global e por domínio. A média do escore global foi igual a 33,85, com desvio padrão igual a 7,25. Para o domínio cuidados da saúde, a média foi igual a 33,28, com desvio padrão igual a 7,70; para o domínio prevenção da doença, a média foi igual a 33,46, com desvio padrão igual a 8,27; e, para o domínio promoção da saúde, a média foi a maior, igual a 35,32, com desvio padrão igual a 8,48. Os escores medianos foram iguais a 33,33 para todos os domínios.

5.1.3. Variável de interesse Escore

Nas Figuras B.5 e B.6, estão indicados o histograma da distribuição do escore global com a curva de densidade e o respectivo *box plot*. Observa-se que os dados são assimétricos à esquerda, devido à presença de alguns pontos atípicos (escores muito baixos). Os valores da média e da mediana estão bem próximos.

O teste de normalidade de Shapiro-Wilk (Shapiro e Wilk, 1965) foi aplicado e rejeitou-se a hipótese de normalidade dos escores, com valor- $p < 0,001$. Portanto, os escores não são normalmente distribuídos.

A Figura B.7 revela que não existe uma tendência do escore global em relação à idade dos participantes. O valor do coeficiente de correlação linear de Pearson indica que não existe correlação entre as variáveis escore e idade.

5.1.4. Dados sociodemográficos

Os dados sociodemográficos estão resumidos nas Tabelas A.4 e A.5. Destaca-se que dos 1.015 entrevistados:

- a maior porcentagem da amostra apresenta entre 42 e 47 anos (27,4%), seguida por 48-53 anos (19,7%) e 54-59 anos (17,9%), enquanto a faixa etária de 66 a 75 anos apresenta a menor porcentagem (1,8%);
- 85,7% dos entrevistados são do gênero Feminino;
- mais da metade (52,7%) é da cor/raça Branca, seguida da Parda (28,6%) e Preta (16%). Indígenas e amarelos têm menor representação;
- 8,6% são gestores e 91,4% são professores;
- 3,3% são coordenadores e apenas 2,3% são diretores;
- a maioria dos professores são do Ensino Fundamental I (51,2%), seguidos dos professores do Fundamental II e Médio (23,3%);
- 25,2% dos indivíduos apresentam tempo de trabalho entre 5 e 10 anos, seguido por Menos de 3 anos (21,8%) e 11 a 15 anos (18,1%), enquanto apenas 4,7% trabalham entre 3 a 4 anos na rede municipal de ensino;

- 60% possuem pós-graduação (*Lato Sensu*/especialização) e apenas 4,8%, mestrado ou doutorado;
- 81,9% frequentaram escola pública na educação básica e 78,9% frequentaram escola privada na educação superior;
- 45,6% da amostra apresenta renda mensal de 4 a 5 salários mínimos e 22,2%, renda mensal de 2 a 3 salários mínimos, enquanto apenas 4,1% têm renda mensal de 10 salários mínimos ou mais;
- as três escolas com as maiores porcentagens de profissionais entrevistados são: CEU EMEF Perus (5,4%), CEU EMEF Vila Curuçá (5%) e EMEF Jornalista Millôr Fernandes (4,7%);
- as diretorias regionais de educação com o maior número de entrevistados são Campo Limpo (140; 13,8%) e São Miguel (132; 12,9%);
- 42,1% das escolas estão localizadas na Zona Leste e 29,9% das escolas estão localizadas na Zona Sul.

5.1.5. Perfil pedagógico-profissional

Nesta seção, são apresentados os percentuais de respondentes em cada categoria de resposta para as perguntas relacionadas à formação docente e à transposição didática.

Formação docente

A partir das respostas relacionadas à formação docente, pode-se observar da Tabela A.6 que:

- 64,1% estão satisfeitos com seu trabalho na escola e 74,8% escolheriam ser profissional da educação novamente, se pudessem;
- 85,8% não recebeu formação sobre Educação para a Saúde na graduação e 87,2% não participou de formação continuada sobre este assunto;
- 67,3% dos professores e gestores não receberam formação para lidar com notícias ou informações veiculadas nas redes sociais;

- 69,5% dos entrevistados não receberam formação para encontrar um equilíbrio entre cumprir o seu papel profissional e orientar os estudantes com relação aos hábitos de saúde.

Pode-se afirmar que a maioria dos profissionais entrevistados considera que não possui formação em Educação para a Saúde.

Transposição didática

Os resultados sobre transposição didática estão indicados na Tabela A.7 e pode-se mencionar que:

- 49% dos entrevistados trabalham em escolas que incluem temas sobre Educação para a Saúde no Projeto Político Pedagógico;
- 20,3% dos entrevistados trabalham em escolas que participam do Programa Saúde na Escola e 21,5% trabalham em escolas que participam do Programa Primeira Infância;
- 49,2% trabalham em escolas que discutem questões sobre Educação para a Saúde e 51,8% trabalham em escolas que possuem temas voltados sobre Educação para a Saúde nos planos de ensino;
- 50,7% não consideram que os alunos da escola possuem uma boa Educação para a Saúde;
- 56,7% realizam atividades que envolvem Educação para a Saúde;
- 48,4% consideram que se apropriaram dos conteúdos necessários para exercer a docência em Educação para a Saúde;
- 59,3% pensam que o aluno da escola é capaz de se comunicar com um adulto ou profissional da saúde sobre comportamentos de saúde;
- 49,4% disseram que os familiares e responsáveis têm a oportunidade de conversar com professores e equipe gestora da escola sobre Educação para a Saúde;
- 79,8% consideram a infraestrutura do local de trabalho adequada;

Assim, apesar da formação formal em Educação para a Saúde não estar muito presente, cerca de metade dos entrevistados realizam a transposição didática desse tema.

5.2. Análise descritiva multidimensional

5.2.1. Variável de interesse Letramento

Na Figura B.8, as classificações positivas (Suficiente/Excelente) e negativas (Problemático/Inadequado) estão bem equilibradas. As proporções de Suficiente/Excelente são um pouco maiores que as de Problemático/Inadequado, com exceção da faixa etária 66-75, em que 67% das classificações são Problemático/Inadequado. É importante observar que essa faixa etária conta com a menor quantidade de respostas (apenas 18). Fora 66-75, a faixa etária 30-35 é a que apresenta maior proporção de classificações Excelente (20%), mas também é a que apresenta uma das maiores proporções de Inadequado (13%). Além disso, as faixas etárias 36-41 e 42-47 apresentam as maiores porcentagens de classificações Suficiente/Excelente (56%).

Não binário apresenta apenas 3 observações, então tirar conclusões a partir do gráfico da Figura B.9 seria equivocado. Entre os gêneros Feminino e Masculino, a diferença entre as proporções de Suficiente/Excelente e Problemático/Inadequado é apenas de 1%. O gênero Feminino apresenta a maior proporção de classificação Excelente (14%), mas também apresenta maior proporção de classificação Inadequado (10%).

A Figura B.10 mostra que temos duas classes, Amarela e Indígena, que não apresentam a classificação Inadequado, devido ao fato de haver poucas observações em cada uma dessas categorias (14 e 5, respectivamente); seria, então, equivocado dizer que pessoas amarelas/indígenas possuem melhor classificação. Fora essas duas categorias, observamos que Branca tem maior proporção de respostas Suficiente/Excelente (56%), mas Preta possui maior proporção de pessoas classificadas como Excelente (15%).

Dentre os gestores, Coordenador pedagógico é a categoria que obteve maior porcentagem de Suficiente/Excelente (76%), enquanto Diretor de escola apresentou a menor porcentagem de Suficiente/Excelente (43%). Dentre os professores, Professor de educação infantil é o que apresenta menor proporção de Suficiente/Excelente (47%), menor classificação de Excelente (10%) e maior classificação de Inadequado (15%).

Em contrapartida, Professor de ensino fundamental II e médio teve menor classificação de Inadequado (8%).

Em geral, é possível observar na Figura B.12 que todas as categorias estão equilibradas e apresentam proporções semelhantes. Menos de 3 anos e 21 a 25 anos apresentam as maiores proporções de classificações positivas (Suficiente/Excelente). Mais de 25 anos e 16 a 20 anos apresentam maior proporção de Excelente. Além disso, 3 a 4 anos e 5 a 10 anos são as categorias com as maiores porcentagens de Problemático.

A Figura B.13 mostra que Doutorado e Pós-doutorado não possuem classificação Inadequado por terem poucas observações (7 e 5, respectivamente). Ensino médio (Magistério) possui grande proporção de Problemático/Inadequado, mas também apresenta poucas observações (apenas 24). Entre as escolaridades restantes, todos possuem proporção de classificações positivas e negativas semelhantes, com Pós-graduação (*Lato Sensu*/Especialização) apresentando a maior proporção de Excelente (15%) e Inadequado (11%).

A Figura B.14 revela que indivíduos que estudaram tanto em escola pública quanto privada na educação básica apresentam a maior proporção de classificações Suficiente/Excelente (60%). A categoria Privada apresenta a maior porcentagem de Excelente (24%), enquanto a categoria Pública detém a maior porcentagem de Inadequado (11%).

De acordo com a Figura B.15, a distribuição dos níveis de letramento global é bastante similar em todas as categorias da variável tipo de graduação. Indivíduos que frequentaram apenas faculdade Privada apresentam a maior proporção de Inadequado (10%). Por sua vez, indivíduos que cursaram apenas faculdade Pública têm a maior proporção de classificações Suficiente/Excelente (55%).

Segundo a Figura B.16, 10 salários mínimos ou mais tem a maior proporção de Suficiente/Excelente, mas vale observar que a categoria apresenta 42 observações. Levando em conta as categorias restantes, 8 a 9 salários mínimos tem a maior proporção de classificações Suficiente/Excelente (64%), e junto com 4 a 5 salários mínimos possuem a maior proporção de classificação Excelente (14%). Por sua vez, 6

a 7 salários mínimos apresenta a menor proporção de classificações Suficiente/Excelente (48%).

Segundo a Figura B.17, a proporção de classificações positivas é consideravelmente menor nas respostas Não (42%) e Sim, mas tem diminuído (43%) comparadas a Sim, frequentemente (60%).

A Figura B.18 indica que pessoas que responderam que seriam professores novamente (Sim) têm maior proporção de classificações Suficiente/Excelente (56%) quando comparadas àquelas que não seriam (48%). Em contraste, os que responderam Não têm maior proporção de Excelente (16%), mas também de Inadequado (16%).

Na Figura B.19, observa-se que aqueles que receberam formação inicial em Educação em Saúde apresentam maior proporção de classificações positivas. A proporção de Excelente para Sim e Não são semelhantes, mas aqueles que tiveram formação têm menor proporção de Inadequado (5%) do que aqueles que nunca tiveram (11%).

Segundo a Figura B.20, igualmente à formação inicial, aqueles que receberam formação continuada em Educação em Saúde apresentam maior proporção de classificações positivas. A resposta Sim apresenta a maior proporção de Excelente (16%) e a menor proporção de Inadequado (5%).

A Figura B.21 indica que aqueles que receberam formação sobre notícias de saúde veiculadas nas redes sociais apresentam maior proporção de classificações Suficiente/Excelente. A resposta Sim apresenta a maior proporção de Excelente (17%) e a menor proporção de Inadequado (8%).

A Figura B.22 mostra que aqueles que conseguem equilibrar profissão e saúde apresentam maior proporção de classificações positivas. A resposta Sim apresenta a maior proporção de Excelente (18%) e a menor proporção de Inadequado (7%).

Na Figura B.23, observa-se que para aqueles que no PPP da escola constam temas sobre saúde a proporção de classificações Excelente (15%) é a maior e a proporção de Inadequado (8%) é a menor.

Pela Figura B.24, observa-se que para os indivíduos cuja escola tem algum programa de saúde (Programa Primeira Infância ou Programa Saúde na Escola) a

proporção de classificações Suficiente/Excelente em comparação às escolas sem programa de saúde é a maior.

A partir da Figura B.25, nota-se que, para os indivíduos cuja escola discute questões sobre Educação para a Saúde, a proporção de classificações Suficiente/Excelente é maior (60%). Além disso, nas escolas onde não se discute questões sobre ES, a proporção de Inadequado é maior (13%).

Para os indivíduos cujo plano de ensino ou trabalho contém temas de Educação para a Saúde, a proporção de classificações Suficiente/Excelente é maior (57%), como mostra a Figura B.26. Por outro lado, aqueles que não incluem ES em seu plano de ensino/trabalho têm a maior proporção de letramento global Inadequado (13%).

Pela Figura B.27, é possível notar que a proporção de classificações de letramento global Suficiente/Excelente é maior (65%) entre os profissionais que consideram que seus alunos têm boa ES em comparação com os profissionais que não consideram (47%), os quais inclusive apresentam a maior proporção de Inadequado (12%).

Com base na Figura B.28, observa-se que profissionais que realizam atividades que envolvem Educação para a Saúde têm a maior proporção de classificações Suficiente/Excelente (57%), enquanto os profissionais que não realizam este tipo de atividade apresentam a maior proporção de Inadequado (11%).

A proporção de indivíduos com classificação de letramento global Suficiente/Excelente é maior entre aqueles que consideram que se apropriaram de conteúdos de ES (61%) em relação aos que não consideram (47%), dentre os quais a proporção de Inadequado é maior (13%), como pode ser visto na Figura B.29.

A proporção de indivíduos com classificação de letramento global Suficiente/Excelente é maior entre aqueles que consideram que seus alunos conseguem se comunicar sobre temas de saúde (57%) em relação aos que não consideram (44%), dentre os quais a proporção de Inadequado é maior (14%), como pode ser visto na Figura B.30.

A partir da Figura B.31, é possível notar que profissionais que trabalham em uma escola onde os responsáveis podem conversar sobre Educação para a Saúde têm a maior proporção de classificações Suficiente/Excelente (62%), enquanto os

profissionais que trabalham em escolas onde não há este incentivo apresentam a maior proporção de Inadequado (19%).

Pela Figura B.32, nota-se que a distribuição dos níveis de letramento global é bastante similar em todas as categorias da variável Q9.4. Indivíduos que consideram a infraestrutura da escola Adequada apresentam a maior proporção de Suficiente/Excelente (55%). Por sua vez, profissionais que consideram a infraestrutura da escola Inadequada têm a maior proporção de letramento Inadequado (12%).

5.2.2. Variável de interesse Escore

Para cada variável explicativa, foi realizado um teste não paramétrico com o objetivo de verificar a igualdade das medianas do escore global nas categorias da variável, já que a suposição de normalidade do escore não foi atendida. Para os testes de hipóteses, foi considerado um nível de significância de 5%.

Para variáveis explicativas com duas categorias, aplicou-se o teste de Mann-Whitney (ou teste de Wilcoxon) (Bussab e Morettin, 2024). Já para variáveis com três ou mais categorias, utilizou-se o teste de Kruskal-Wallis (Kruskal e Wallis, 1952). A hipótese nula em ambos os testes é de que as medianas dos escores dos níveis da variável são iguais. Caso a hipótese nula seja rejeitada, isto é, se o teste for estatisticamente significativo, há evidências de que os escores medianos são diferentes entre as categorias. Os valores-p obtidos estão apresentados na Tabela A.8.

Com base nos testes realizados, observou-se, com valor-p < 0,05, diferença entre os escores medianos de pelo menos duas categorias das seguintes variáveis sociodemográficas: cargo, função, tipo de educação básica e renda. Para uma abordagem mais conservadora, considerando valor-p < 0,01, apenas cargo e renda foram significantes.

Para as variáveis provenientes do questionário de autoeficácia do professor, ao nível de significância de 5%, não foi observada diferença dos escores medianos entre pelo menos duas categorias das variáveis Q6.4, Q8.5, Q9.4 e Q10.1. Vale ressaltar que, ao nível de significância de 1%, as variáveis em que foi observada diferença dos escores medianos foram: Q5.1, Q9.2, Q9.3, Q7.1, Q8.1, Q8.3, Q8.4 e Q9.1.

A Tabela A.8 também fornece algumas medidas descritivas do escore por todas as variáveis explicativas. Por exemplo, para o cargo, os coordenadores pedagógicos apresentam o maior escore mediano, enquanto os diretores detêm o menor escore mediano e o menor desvio padrão do escore. O menor escore dentre todos (igual a zero) está presente na categoria Professor de educação infantil e ensino fundamental I. O nível Professor de educação infantil apresenta a maior variabilidade de escore.

Para a variável tipo de educação básica, nota-se que os participantes cuja formação aconteceu apenas em escola Pública apresentam o menor escore médio. O escore igual a zero está contido nessa categoria.

Em relação à renda, nota-se que os participantes cuja renda pessoal líquida é de 10 salários mínimos ou mais apresentam o maior escore mediano, o maior escore médio e o maior desvio padrão dos escores. A categoria 6 a 7 salários mínimos apresenta o menor escore mediano e contém o escore igual a zero.

Com respeito à variável Q5.1, percebe-se que tanto a mediana quanto a média do escore é maior para os professores satisfeitos com sua contribuição na escola. Para a variável Q5.2, o escore mediano é maior dentre os participantes que escolheriam novamente ser um profissional da educação, enquanto os escores são mais dispersos entre aqueles que não escolheriam.

Para as variáveis Q6.1, Q6.2, Q9.2 e Q9.3, percebe-se que a média do escore é mais alta dentre os profissionais que receberam algum tipo de formação sobre Educação para a Saúde.

Em relação à variável Q6.3, nota-se que o escore médio é o maior dentre os profissionais que trabalham em uma escola cujo Projeto Político Pedagógico contém temas sobre ES.

Para a variável Q7.1, percebe-se que os professores que discutem questões sobre Educação para a Saúde no planejamento de aula apresentam o maior escore mediano. Por sua vez, acerca da variável Q7.2, nota-se que profissionais cujo plano de ensino ou trabalho contém temas voltados à ES apresentam o maior escore médio.

Ademais, profissionais que consideram que os alunos da escola possuem uma boa ES (variável Q8.1) apresentam o maior escore mediano. Em relação à variável

Q8.2, os profissionais que realizam atividades que envolvam a ES detêm o maior escore médio.

Para a variável Q8.3, observa-se que profissionais que consideram ter se apropriado dos conteúdos necessários para realizar ações de ES no trabalho têm o maior escore mediano. Já para a variável Q8.4, os profissionais que acreditam que os alunos são capazes de se comunicar com um adulto ou profissional de saúde sobre comportamentos de saúde apresentam o maior escore mediano.

Com respeito à variável Q9.1, percebe-se que, nas escolas em que os responsáveis pelos estudantes têm a oportunidade de conversar com os profissionais sobre saúde, estes profissionais da educação apresentam maiores escores médio e mediano.

5.3. Validação do questionário e consistência interna

Confiabilidade é a habilidade do instrumento de reproduzir um resultado consistentemente. Ela está relacionada, principalmente, com estabilidade, consistência interna e equivalência de medidas; ela não é uma propriedade fixa de um instrumento e pode ser influenciada por alguns fatores (por exemplo, a função do instrumento, a população, o contexto etc.) (De Souza et al., 2017; Keszei et al., 2010).

A consistência interna é uma propriedade importante para questionários que pretendem medir um único construto utilizando itens múltiplos. “Uma baixa consistência interna pode significar que os itens medem atributos distintos ou que as respostas aos assuntos são inconsistentes”. (Keszei et al., 2010)

O alfa de Cronbach (Cronbach, 1951) é frequentemente utilizado para medir a consistência interna (De Souza et al., 2017; Béland et al., 2016), embora seja importante levar em conta que os valores de alfa são influenciados pelo número de itens (De Souza et al., 2017; Cortina, 1993; Terwee et al., 2007). É uma medida desenvolvida por Lee Cronbach, em 1951.

Um limite inferior bastante aceito para o alfa de Cronbach é 0,7 (Cortina, 1993), logo, uma escala de questionário possui uma boa consistência interna se o alfa de Cronbach é maior do que 0,7. Se $\alpha = 1$, então os itens estão perfeitamente

correlacionados uns aos outros; se, pelo contrário, $\alpha = 0$, então os itens não estão relacionados entre si.

A fiabilidade do instrumento, baseada no coeficiente de alfa de Cronbach, na escala de 16 itens, foi de 0,92. Assim, no presente estudo, foi verificado que o HLS-16 tem consistência interna. Os valores do HLS-Q16 para o letramento global e para os domínios, analisados individualmente, são apresentados na Tabela A.9. Analisando a tabela, verifica-se que os coeficientes de alfa de Cronbach oscilam entre 0,78-0,92, portanto, eles são satisfatórios. Os coeficientes de alfa de Cronbach para o HLS-Q16 foram: 0,84 para o domínio de cuidados da saúde, 0,78 para o domínio de prevenção da doença e 0,81 para o domínio de promoção da saúde.

Como se pode observar na Tabela A.10, todas as 16 questões são significativamente correlacionadas (valores- $p < 0,001$), com coeficientes de correlação de Spearman (Spearman, 1904) variando entre 0,284 e 0,641. Nota-se que todas as correlações são positivas. Observa-se também que as correlações entre questões consecutivas do mesmo domínio são mais altas. Por exemplo, as questões Q1 e Q2, pertencentes ao domínio de cuidados da saúde, apresentam correlação moderada (0,605), analogamente para as questões Q3 e Q4 (0,613); as questões Q9 e Q10, que integram o domínio de prevenção da doença, também apresentam correlação moderada (0,628); as questões Q14 e Q15, pertencentes ao domínio de promoção da saúde, apresentam correlação moderada (0,565).

6. Análise inferencial

Nesta seção, apresentam-se análises inferenciais dos dados por meio de três modelos: modelo de regressão generalizado, árvore de decisão e modelo de mistura Gaussiana complementado pelo modelo logístico. O objetivo desses modelos estatísticos é explicar a relação entre a variável resposta escore global e as variáveis explicativas.

6.1. Modelo GAMLSS

O modelo generalizado aditivo para localização, escala e forma (GAMLSS – *Generalized Additive Models for Location, Scale and Shape*) está implementado no pacote computacional “gamlss” do R (Rigby e Stasinopoulos, 2005). A distribuição da variável resposta, no GAMLSS, pode ser escolhida a partir de uma família de distribuições muito ampla, que abrange aquelas com grande assimetria e curtose, discretas ou contínuas. No ajuste do modelo, foi considerada a distribuição Box-Cox t (BCT).

A distribuição Box-Cox t é utilizada em situações nas quais os dados apresentam uma distribuição assimétrica, com caudas pesadas ou contêm *outliers*, sem exigir transformação prévia dos dados, o que facilita a interpretação direta dos parâmetros e melhora a modelagem dos dados. Estas características estão de acordo com o comportamento dos dados observado no histograma com curva de densidade da distribuição do escore global (Figura B.5). A distribuição BCT é um caso especial da ampla classe das distribuições simétricas de Box-Cox (BCS), proposta por Rigby e Stasinopoulos (2006). De forma geral, ela possui quatro parâmetros: μ (parâmetro de localização), que é a mediana da distribuição; σ (parâmetro de escala) é a dispersão relativa; ν (parâmetro de forma) é a assimetria da distribuição; e τ (parâmetro de forma) está relacionado à curtose da distribuição. Alguns detalhes desta distribuição podem ser encontrados no Apêndice C.1.

Para o ajuste, foram desconsiderados todos os dados ausentes. Logo, o número de observações ajustadas foi igual a 945 profissionais da educação. É importante ressaltar que a distribuição BCT não pode assumir valor nulo, pois o zero não pertence ao domínio da distribuição, isto é, todos os valores possíveis da variável aleatória. Portanto, para a observação de número 77, que apresentou escore igual a zero, o valor do escore foi alterado para 0,001.

No modelo inicial, todos os parâmetros foram modelados, considerando-se todas as variáveis explicativas geradas pelos questionários sociodemográfico e de autoeficácia do profissional. A seleção das variáveis foi realizada por meio do procedimento *stepwise* baseado no critério GAIC (*Generalized Akaike Information Criterion*) (Taniguchi e Hirukawa, 2012), para cada parâmetro da distribuição BCT. A

função “stepGAIC” do GAMLSS seleciona automaticamente o melhor modelo a fim de otimizar a medida GAIC de qualidade do ajuste.

Após um processo exaustivo de seleção de variáveis para todos os parâmetros da distribuição BCT, obteve-se um modelo cujas estimativas para o parâmetro de locação, que diz respeito aos efeitos das covariáveis no escore global mediano, encontram-se na Tabela A.11.

A interpretação das estimativas dos parâmetros do modelo permite analisar o efeito de cada variável selecionada no escore global. Os parâmetros de escala e forma foram modelados no sentido de controlar a distribuição e suas interpretações serão omitidas, pois elas não integram o escopo da análise. Suas estimativas podem ser encontradas nas Tabelas A.12 a A.14.

A seguir, interpreta-se o modelo para o parâmetro μ , que denota a mediana do escore global. No ajuste do modelo, foi utilizada a ligação identidade, logo, as estimativas podem ser interpretadas diretamente. Pela Tabela A.11, pode-se analisar cada um dos efeitos significantes da seguinte maneira, supondo indivíduos com as mesmas características à exceção da variável considerada, isto é, fixadas as demais covariáveis:

- Cargo: o escore global mediano dos Professores de educação infantil é menor em relação aos Assistentes de diretor de escola, sendo a diferença de 3,85 pontos;
- Renda: em relação aos indivíduos que recebem 10 ou mais salários mínimos, profissionais com renda líquida mensal de 6 a 7 salários mínimos têm escore global mediano menor, com uma diferença de 2,58 pontos;
- Q8.1: o escore global mediano dos profissionais que consideram que os alunos de sua escola têm uma boa Educação para a Saúde é acrescido de 1,35 pontos em relação ao escore global mediano dos profissionais que não o consideram;
- Q8.3: os profissionais que consideram terem se apropriado dos conteúdos necessários para realizar ações de ES no trabalho apresentam escore global mediano superior aos profissionais que não o consideram, sendo a diferença de 1,14 pontos;

- Q9.1: em comparação com os profissionais que consideram que os responsáveis pelos estudantes não têm a oportunidade de conversar com os profissionais da escola sobre ES, o escore global mediano dos profissionais que não sabem dizer se os responsáveis têm essa oportunidade é superior em 1,47 pontos, enquanto o escore global mediano dos profissionais que responderam que os responsáveis têm essa oportunidade é maior em 1,97 pontos;
- Q9.2: os educadores que recebem formação sobre notícias de saúde veiculadas nas redes sociais apresentam escore global mediano maior (diferença de 1,22 pontos) em comparação aos que não a recebem.

Além disso, é possível realizar uma interpretação para o valor do intercepto: um indivíduo que é assistente de diretor, tem doutorado, fez a educação básica em escola privada, tem renda de 10 salários mínimos ou mais, não está satisfeito com sua contribuição na escola, não considera que os alunos da escola possuem uma boa Educação para a Saúde, não considera que se apropriou dos conteúdos pedagógicos necessários para realizar ações de ES no trabalho, afirma que os responsáveis pelos estudantes não têm a oportunidade de conversar com os profissionais da escola sobre ES e que não recebeu formação sobre informações de saúde veiculadas nas redes sociais apresenta escore global mediano de 43,4.

Para se obter qualquer estimativa do escore global mediano, é preciso combinar o valor do intercepto com os valores dos efeitos estimados para cada variável explicativa desejada. Por exemplo, se o interesse é estimar o escore global mediano de um professor da educação infantil (-3,849), que fez alguma especialização (-0,404), cuja formação básica foi realizada em escola pública (-8,043), com renda de 4 a 5 salários mínimos (-1,201), cuja satisfação com o trabalho tem diminuído (-1,386), que considera que os seus alunos têm boa ES (1,346), que considera que se apropriou dos conteúdos de ES (1,143), que trabalha em uma escola onde os responsáveis podem conversar sobre ES (1,969) e que recebeu formação sobre notícias de ES (1,221), o seu escore global mediano estimado será a soma do intercepto com os efeitos estimados. Assim, o escore global mediano deste indivíduo é:

$$\text{Escore global mediano} = 43,390 - 3,849 - 0,404 - 8,043 - 1,201 - 1,386 + 1,346 + 1,143 + 1,969 + 1,221 \cong 34,2$$

Após o ajuste, é necessária a realização de um diagnóstico e análise dos resíduos, que correspondem aos erros estimados do modelo, para avaliar se há afastamentos importantes das suposições feitas pelo modelo. Na Figura B.33A, o gráfico dos valores ajustados pelos resíduos quantílicos normalizados é utilizado para verificar a suposição de homocedasticidade, isto é, de que a variância dos erros do modelo é constante. Como os pontos apresentam uma distribuição retangular, uniforme, sem qualquer padrão, verifica-se a homocedasticidade dos erros.

Na Figura B.33B, é apresentado o gráfico dos resíduos por seus índices. Esse gráfico é utilizado para verificar se existe alguma tendência temporal. Como a distribuição dos resíduos não apresenta qualquer padrão, pode-se afirmar que não existe correlação temporal dos dados.

Para a Figura B.33C, sob a distribuição suposta para os dados, no caso a distribuição BCT, os resíduos devem ser normais. A partir da figura, nota-se que a normalidade é suave. Na Figura B.33D, é apresentado o *QQ-plot* dos resíduos (Bussab e Morettin, 2024); se os pontos estiverem próximos da reta vermelha, há indicativo de bom ajuste do modelo. Pelo gráfico, nota-se que os pontos seguem a linha vermelha, com uma ligeira oscilação nas caudas. Portanto, o modelo está bem ajustado.

Na Figura B.34, é apresentado o gráfico *worm plot* (Van Buuren e Fredriks, 2001). O ideal é que os pontos estejam em cima da linha vermelha e não ultrapassem as meias-luas. Assim, é possível concluir que o modelo apresentou bom ajuste. O modelo também está bem ajustado em relação aos valores discrepantes.

6.2. Árvore de regressão

Com a finalidade de avaliar o impacto das variáveis explicativas em relação ao escore global, construiu-se uma árvore de decisão de regressão (Izbicki e Dos Santos, 2020). As árvores de regressão consistem em uma metodologia não paramétrica, ou seja, que não assume uma distribuição para os dados, e que leva a resultados extremamente interpretáveis. Embora elas sejam muito utilizadas com a finalidade de previsão, no presente estudo, essa técnica de aprendizado de máquina foi aplicada a fim de se obter interpretações para os efeitos das variáveis explicativas no escore.

Uma árvore é construída por partições recursivas entre as covariáveis. Cada particionamento recebe o nome de nó e cada resultado recebe o nome de folha. A partir do resultado, são criadas ramificações que permitem delinear como as características do profissional da educação impactam no seu nível de letramento em saúde, expresso por seu valor de escore. Assim, é possível observar se alguma das variáveis explicativas tem um efeito significativo. Vale destacar que a árvore de regressão se expande quando há uma divisão estatisticamente significativa.

A partir da árvore mostrada na Figura B.35, o nó inicial revela que o escore global médio é de 33,84. Nota-se que, ao separar os profissionais que não consideram que os alunos da escola possuem uma boa Educação para a Saúde dos que o consideram, o primeiro grupo apresenta escore médio de 32,75, enquanto o segundo grupo apresenta escore global médio de 34,99, ou seja, a média do escore é ligeiramente superior neste grupo.

Dentre os profissionais que não consideram que os alunos têm uma boa ES, aqueles cuja satisfação com a contribuição na escola tem diminuído ou não existe apresentam escore global médio inferior (31,51) em comparação aos que estão frequentemente satisfeitos (33,85).

Ainda sob esse recorte, para os profissionais com tempo de trabalho entre 3 a 10 anos, o escore global médio é de 31,25, enquanto para os profissionais com tempo de trabalho menor do que 3 anos ou superior a 10 anos, é 35,07.

Retomando o grupo de profissionais que consideram que os alunos têm uma boa ES, aqueles que não recebem formação sobre notícias e informações de saúde veiculadas nas redes sociais apresentam escore global médio inferior (33,96) em comparação aos que recebem (36,45).

Portanto, a partir da análise da árvore de regressão, observa-se que as variáveis Q8.1, Q5.1, Q9.2 e tempo de trabalho são bastante relevantes na predição do escore global.

6.3. Modelo de mistura Gaussiana

Uma outra abordagem que pode ser interessante é investigar a distribuição dos escores nos três domínios, ao invés de considerar apenas o escore global. Para tanto,

pode-se utilizar o modelo de mistura Gaussiana (*Gaussian Mixture Model*, ou apenas GMM) (Reynolds, 2009), um modelo probabilístico paramétrico cuja principal suposição é de que todos os dados são gerados de uma mistura finita de distribuições Gaussianas com parâmetros desconhecidos. Esse modelo é utilizado como um método de clusterização que incorpora informações sobre a estrutura de covariância dos dados, bem como sobre as médias das distribuições Gaussianas.

Por sua vez, a clusterização (ou análise de agrupamentos) é um algoritmo não-supervisionado de aprendizado de máquina que organiza e classifica diferentes observações em grupos (clusters) baseado na similaridade entre os dados. É uma técnica que visa encontrar agrupamentos naturais nos dados, com o intuito de identificar quais categorias podem existir e o que as determina.

Foi ajustado um modelo GMM com dois clusters, cada um contendo sua própria estrutura de matriz de covariância. A Figura B.36 exibe uma visualização tridimensional dos escores por domínio de cada indivíduo. A partir dela, pode-se concluir que a distribuição dos escores em relação aos três domínios é bastante similar, pois se nota que a nuvem de pontos forma, no espaço tridimensional, uma espécie de cilindro. Logo, os escores são bem distribuídos. Existe apenas um único ponto discrepante, o escore de valor zero, que se desvia do padrão dos dados.

Com base na Figura B.36 e na Tabela A.15, nota-se que a técnica de clusterização separou os pontos em dois clusters: o cluster 0, que representa os profissionais com escore máximo no domínio PS (todos os escores neste domínio são iguais a 50) e alto escore no domínio PD (todos os escores neste domínio são maiores ou iguais a 30); e o cluster 1, que inclui os profissionais com escore menor do que 50 no domínio PS. Uma possível interpretação é de que existem dois grupos segmentados: um deles contém os profissionais de educação com um excelente escore de promoção da saúde e escores médios altos para os outros domínios, o que indica que eles possuem uma excelente capacidade de compreender, interpretar, se atualizar e tomar uma opinião consciente sobre questões de saúde; por sua vez, o outro grupo inclui os profissionais com escore baixo para algum dos três domínios, sugerindo que tais profissionais da educação necessitam de aprimoramento nas competências de um ou mais domínios.

Com base nesses resultados, uma análise que pode ser realizada é a regressão logística com o cluster como variável resposta e as mesmas variáveis explicativas utilizadas nos outros modelos.

6.3.1. Modelo de regressão logística

A regressão logística (Cox, 1970) é um dos modelos estatísticos mais utilizados na análise de dados com resposta binária, ou seja, resposta que admite apenas dois resultados. Considera-se o cluster obtido no ajuste do modelo GMM como a variável resposta. Então, é de interesse estimar a probabilidade (denotada por π) de se pertencer ao cluster 1, dados os valores das variáveis explicativas.

Inicialmente, foi ajustado um modelo logístico binomial com todas as variáveis explicativas. A fim de encontrar um modelo reduzido que inclua apenas as covariáveis mais importantes para explicar a probabilidade de ser do cluster 1, foi aplicado o procedimento *stepwise*, que se baseia em um algoritmo misto de inclusão e eliminação de variáveis explicativas segundo a importância delas, de acordo com o critério de Akaike (AIC) (Akaike, 1974). O submodelo selecionado contém as seguintes covariáveis: Tipo de graduação, Q5.1, Q5.2, Q6.2, Q6.4, Q8.1 e Q8.3.

As estimativas resultantes são apresentadas na Tabela A.16. Algumas interpretações podem ser feitas com base nesses valores:

- A razão de chances entre um indivíduo que respondeu Sim à questão 5.2 e um indivíduo que respondeu Não à questão 5.2 de pertencer ao cluster 1, fixadas as demais variáveis, é estimada por $e^{0,73} \cong 2,075$, ou seja, um indivíduo que respondeu Sim à questão 5.2 tem, aproximadamente, 2,1 vezes a chance de um indivíduo que respondeu Não à questão 5.2 de pertencer ao cluster 1;
- Um profissional da educação que respondeu Não sei à questão 6.4 tem, aproximadamente, 2,1 ($e^{0,751}$) vezes a chance de um profissional que respondeu Não à questão 6.4 de pertencer ao cluster 1, fixadas as demais variáveis;
- A chance de um indivíduo que respondeu Sim à questão 8.1 de pertencer ao cluster 1 é, aproximadamente, 43% ($1 - e^{-0,560}$) menor do que um indivíduo que respondeu Não à questão 8.1, fixadas as demais variáveis;

- A chance de um indivíduo cuja graduação se deu tanto em instituição pública quanto privada de pertencer ao cluster 1 é, aproximadamente, 220% ($e^{1,164} - 1$) maior do que um indivíduo cuja graduação se deu apenas em instituição privada, fixadas as demais variáveis.

Para verificar se o modelo logístico está bem ajustado aos dados, faz-se mister realizar uma análise e diagnóstico dos resíduos. A partir da Figura B.37, que corresponde ao envelope normal de probabilidades para os resíduos do modelo (Williams, 1984), pode-se notar que não há indícios de afastamentos da suposição de distribuição Binomial para a resposta, uma vez que os pontos estão dentro do envelope. Além disso, os pontos estão bem ajustados à reta de referência.

Na Figura B.38, o gráfico da distância de Cook aproximada destaca alguns pontos, correspondentes às unidades observacionais 279, 851 e 896, como possivelmente influentes. No entanto, levando-se em consideração que há muitas unidades observacionais, não há grandes problemas. Portanto, o modelo logístico binomial apresenta bom ajuste aos dados.

A Figura B.39 apresenta a curva característica de operação do receptor (curva ROC) (Agresti, 2013), um procedimento gráfico de discriminação dos dados binários que varia conforme variam as taxas de verdadeiros positivos e falsos positivos. O valor da área sob a curva ROC ajustada (AUC) é uma medida de desempenho do modelo. Quanto mais próxima de 1 (100%), melhor é a capacidade do modelo de distinguir os dois clusters. Foi obtida uma área sob a curva de 66,5%. Portanto, a classificação segundo a curva ROC é considerada aceitável.

7. Conclusão

O estudo revelou que o letramento em saúde está relacionado a características sociodemográficas, formação do docente e transposição didática. Por meio da análise descritiva, constatou-se que indivíduos com renda de 10 salários mínimos ou mais e satisfeitos com sua contribuição na escola demonstram maior grau de letramento, podendo sugerir que a satisfação profissional desempenha um papel crucial nesse aspecto.

Em contraste, os resultados do ajuste do modelo linear generalizado mostraram que professores da educação infantil apresentam grau de letramento, representado pelo escore global mediano, inferior aos assistentes de diretor. Além disso, os profissionais que receberam formação sobre notícias e informações de saúde têm um grau de letramento em saúde superior aos que não tiveram formação; de semelhante modo, os indivíduos que consideram ter se apropriado dos conceitos de ES apresentam um escore global mediano maior em comparação aos que não consideram. Analogamente, aqueles profissionais que consideram que seus alunos possuem uma boa ES apresentam melhor nível de LS em relação aos que não consideram.

Pela árvore de decisão, conclui-se também que os profissionais com mais do que 10 anos de tempo de trabalho na rede pública ou com menos do que 3 anos apresentam, em média, maior nível de letramento em saúde em comparação aos profissionais com tempo de trabalho entre 3 a 10 anos.

Além disso, o modelo de mistura Gaussiana separou os indivíduos em dois clusters: aqueles com escore máximo para o domínio de prevenção da saúde e aqueles com escore menor do que 50 para o domínio de prevenção da saúde. A partir da regressão logística, concluiu-se que profissionais que consideram que seus alunos possuem uma boa ES têm maior chance de pertencer ao grupo dos indivíduos com altos escores para os três domínios em relação aos que não consideram.

APÊNDICE A

Tabelas

Tabela A.1 Respostas às questões do HLS-Q16 (frequência absoluta; porcentagem).

Domínio	Questões	Muito fácil	Fácil	Difícil	Muito difícil	Não sei/Em branco
CS	Q1	208; 20,5%	627; 61,8%	148; 14,5%	30; 3,0%	2; 0,2%
CS	Q2	189; 18,6%	583; 57,5%	194; 19,1%	48; 4,7%	1; 0,1%
CS	Q3	180; 17,7%	726; 71,6%	104; 10,2%	4; 0,4%	1; 0,1%
CS	Q4	288; 28,4%	677; 66,7%	46; 4,5%	4; 0,4%	0; 0,0%
CS	Q5	144; 14,2%	527; 51,9%	302; 29,8%	39; 3,8%	3; 0,3%
CS	Q6	155; 15,2%	624; 61,5%	222; 21,9%	13; 1,3%	1; 0,1%
CS	Q7	224; 22,0%	685; 67,5%	88; 8,7%	15; 1,5%	3; 0,3%
PD	Q8	144; 14,2%	454; 44,7%	329; 32,4%	85; 8,4%	3; 0,3%
PD	Q9	383; 37,7%	566; 55,8%	60; 5,9%	6; 0,6%	0; 0,0%
PD	Q10	370; 36,5%	581; 57,2%	59; 5,8%	5; 0,5%	0; 0,0%
PD	Q11	166; 16,4%	582; 57,3%	243; 23,9%	23; 2,3%	1; 0,1%
PD	Q12	156; 15,3%	624; 61,5%	219; 21,6%	16; 1,6%	0; 0,0%
PS	Q13	272; 26,8%	594; 58,5%	126; 12,4%	23; 2,3%	0; 0,0%
PS	Q14	219; 21,6%	654; 64,4%	125; 12,3%	15; 1,5%	2; 0,2%
PS	Q15	266; 26,2%	657; 64,7%	88; 8,7%	4; 0,4%	0; 0,0%
PS	Q16	287; 28,3%	605; 59,6%	109; 10,7%	14; 1,4%	0; 0,0%

Tabela A.2 Distribuição de frequências da variável Letramento global

Classificação	Letramento global
Inadequado	100 (10,0%)
Problemático	367 (36,0%)
Suficiente	409 (40,3%)
Excelente	139 (13,7%)

Tabela A.3 Medidas-resumo do Escore global e por domínio (n = 1015)

	Média	Desvio padrão	Mínimo	1º quartil	Mediana	3º quartil	Máximo	Coefficiente de variação
Global	33,85	7,25	0	29,17	33,33	37,50	50	21,4%
Cuidados da saúde	33,28	7,70	0	28,57	33,33	35,71	50	23,2%
Prevenção da doença	33,46	8,27	0	29,58	33,33	40,00	50	24,7%
Promoção da saúde	35,32	8,48	0	29,17	33,33	41,67	50	24,0%

Tabela A.4 Distribuição de frequências das variáveis sociodemográficas

Variável	Categoria	Contagem (n)	Porcentagem (%)
Faixa etária	18-29 anos	26	2,6%
	30-35 anos	83	8,2%
	36-41 anos	176	17,3%
	42-47 anos	278	27,4%
	48-53 anos	200	19,7%
	54-59 anos	182	17,9%
	60-65 anos	52	5,1%
	66-75 anos	18	1,8%
Gênero	Feminino	870	85,7%
	Masculino	139	13,7%
	Não binário	3	0,3%
	Prefiro não responder	3	0,3%
Cor/Raça	Amarela	14	1,4%
	Branca	535	52,7%
	Indígena	5	0,5%
	Parda	290	28,6%
	Preta	163	16%
	Não informado	8	0,8%

Tabela A.4 (Continuação) Distribuição de frequências das variáveis sociodemográficas

Variável	Categoria	Contagem (n)	Porcentagem (%)
Cargo	Assistente de diretor	30	3,0%
	Coordenador	34	3,3%
	Diretor	23	2,3%
	Professor Ed. Infantil	171	16,8%
	Professor Fund. I	520	51,3%
	Professor Fund. II e Médio	237	23,3%
Disciplina (apenas para Professor Fund. II e Médio)	Artes	31	3,1%
	Ciências	41	4,0%
	Educação Física	23	2,3%
	Geografia	25	2,5%
	História	29	2,9%
	Inglês	18	1,8%
	Matemática	30	3,0%
	Português	40	3,9%
	Não se aplica	778	76,5%
Função	Gestor	87	8,6%
	Professor	928	91,4%
Tempo de trabalho	Menos de 3 anos	221	21,8%
	3 a 4 anos	48	4,7%
	5 a 10 anos	256	25,2%
	11 a 15 anos	184	18,1%
	16 a 20 anos	134	13,2%
	21 a 25 anos	96	9,5%
	Mais de 25 anos	75	7,4%
	Não informado	1	0,1%

Tabela A.4 (Continuação) Distribuição de frequências das variáveis sociodemográficas

Variável	Categoria	Contagem (n)	Porcentagem (%)
Escolaridade	Ensino médio	24	2,4%
	Graduação	328	32,3%
	Pós-graduação	609	60,0%
	Mestrado	42	4,1%
	Doutorado	7	0,7%
	Pós-doutorado	5	0,5%
Tipo de educação básica	Privada	76	7,5%
	Pública	831	81,9%
	Pública e privada	108	10,6%
Tipo de graduação	Privada	801	78,9%
	Pública	151	14,9%
	Pública e privada	63	6,2%
Renda	2 a 3 sal. mínimos	225	22,2%
	4 a 5 sal. mínimos	463	45,6%
	6 a 7 sal. mínimos	184	18,1%
	8 a 9 sal. mínimos	99	9,8%
	10 sal. mínimos ou mais	42	4,1%
	Não informado	2	0,2%

Tabela A.5 Organização dos dados da variável Escola em função da quantidade de gestores e professores por DRE e Zona (n; porcentagem)

Escola	Gestores	Professores	Total	DRE	Zona
A	4; 0,4%	31; 3,1%	35; 3,5%	Freguesia/Brasilândia 66; 6,6%	Norte 133; 13,2%
B	2; 0,2%	14; 1,4%	16; 1,6%		
C	3; 0,3%	12; 1,2%	15; 1,5%		
D	3; 0,3%	34; 3,5%	37; 3,8%	Jaçanã/Tremembé 67; 6,6%	
E	2; 0,2%	28; 2,7%	30; 2,9%		
F	3; 0,3%	39; 3,8%	42; 4,1%	Pirituba/Jaraguá 97; 9,5%	Oeste 151; 14,8%
G	3; 0,3%	52; 5,1%	55; 5,4%		
H	5; 0,5%	25; 2,4%	30; 2,9%	Butantã 54; 5,3%	
I	3; 0,3%	21; 2,1%	24; 2,4%		
J	3; 0,3%	27; 2,7%	30; 3,0%	Campo Limpo 140; 13,8%	
K	3; 0,3%	30; 3,0%	33; 3,3%		
L	5; 0,5%	23; 2,3%	28; 2,8%		
M	4; 0,4%	45; 4,3%	49; 4,7%		
N	3; 0,3%	41; 3,9%	44; 4,2%	Capela do Socorro 78; 7,7%	Sul 302; 29,9%
O	4; 0,4%	30; 3,1%	34; 3,5%		
P	1; 0,1%	31; 3,1%	32; 3,2%	Santo Amaro 62; 6,2%	
Q	3; 0,3%	27; 2,7%	30; 3,0%		
R	0; 0,0%	10; 1,0%	10; 1,0%	Ipiranga 22; 2,2%	
S	0; 0,0%	12; 1,2%	12; 1,2%		

Tabela A.5 (Continuação) Organização dos dados da variável Escola em função da quantidade de gestores e professores por DRE e Zona

Escola	Gestores	Professores	Total	DRE	Zona
T	3; 0,3%	27; 2,6%	30; 2,9%	Guaianases	Leste 429; 42,1%
U	5; 0,5%	31; 3,1%	36; 3,6%	66; 6,5%	
V	0; 0,0%	39; 3,8%	39; 3,8%	Itaquera	
W	2; 0,2%	26; 2,6%	28; 2,8%	67; 6,6%	
X	5; 0,5%	19; 1,9%	24; 2,4%	Penha 82; 8,0%	
Y	0; 0,0%	12; 1,2%	12; 1,2%		
Z	0; 0,0%	46; 4,4%	46; 4,4%		
AA	5; 0,5%	41; 4,0%	46; 4,5%	São Mateus	
AB	3; 0,3%	33; 3,3%	36; 3,6%	82; 8,1%	
AC	1; 0,1%	41; 3,9%	42; 4,0%	São Miguel 132; 12,9%	
AD	5; 0,5%	46; 4,5%	51; 5,0%		
AE	4; 0,4%	35; 3,5%	39; 3,9%		
Total	87; 8,6%	928; 91,4%	1015; 100%	1015; 100%	1015; 100%

Tabela A.6 Distribuição de frequências das variáveis relacionadas à formação docente

Questão	Resposta	Contagem (n)	Porcentagem (%)
Q5.1	Sim, frequentemente	651	64,1%
	Sim, mas tem diminuído	333	32,8%
	Não	26	2,6%
	Não informado	5	0,5%
Q5.2	Sim	759	74,8%
	Não	256	25,2%
Q5.2a	Desgaste emocional e físico	29	2,9%
	Desvalorização e falta de reconhecimento	70	6,9%
	Indisciplina e dificuldade com alunos e famílias	33	3,2%
	Más condições de trabalho	99	9,7%
	Salários baixos e sobrecarga de trabalho	19	1,9%
	Não informado	765	75,4%
Q6.1	Sim	142	14,0%
	Não	871	85,8%
	Não informado	2	0,2%
Q6.1a	Alimentação saudável	4	0,4%
	Doenças e patologias	6	0,6%
	Ed. Física e Biologia	42	4,1%
	Saúde e educação	32	3,2%
	Outros	21	2,1%
	Não informado/Não sei/Não lembro	910	89,6%
Q6.2	Sim	129	12,7%
	Não	885	87,2%
	Não informado	1	0,1%

Tabela A.6 (Continuação) Distribuição de frequências das variáveis relacionadas à formação docente

Questão	Resposta	Contagem (n)	Porcentagem (%)
Q6.2a	Formações religiosas ou comunitárias	3	0,3%
	Instituições de ensino superior	4	0,4%
	Programas de saúde na escola	36	3,5%
	Secretaria Municipal de Educação	36	3,5%
	Sindicato/Entidades sindicais	9	0,9%
	Outros	12	1,2%
	Não informado/Não sei	915	90,2%
Q9.2	Sim	326	32,1%
	Não	683	67,3%
	Não informado	6	0,6%
Q9.2a	Ações da UBS e parcerias de saúde	45	4,4%
	Discussões e reflexões coletivas	27	2,7%
	Formações e momentos pedagógicos	58	5,7%
	Formações sobre <i>fake news</i> e combate à desinformação	42	4,1%
	Informações gerais e comunicados oficiais	43	4,2%
	Informações sobre alimentação saudável	9	0,9%
	Informações sobre campanhas de vacinação e saúde pública	26	2,6%
	Informações sobre dengue e prevenção de doenças específicas	21	2,1%
	Outros	33	3,3%
	Não informado/Não sei	711	70,0%

Tabela A.6 (Continuação) Distribuição de frequências das variáveis relacionadas à formação docente

Questão	Resposta	Contagem (n)	Porcentagem (%)
Q9.3	Sim	298	29,3%
	Não	705	69,5%
	Não informado	12	1,2%
Q9.3a	Formação de professores (Horário Coletivo de Estudos – JEIF e PEA: Projeto Especial de Ação)	34	3,3%
	Observação e identificação de questões de saúde dos alunos	17	1,7%
	Parcerias externas	9	0,9%
	Práticas de cuidado e promoção de saúde	9	0,9%
	Reuniões pedagógicas	20	2,0%
	Não informado	926	91,2%

Tabela A.7 Distribuição de frequências das variáveis relacionadas à transposição didática

Questão	Resposta	Contagem (n)	Porcentagem (%)
Q6.3	Sim	497	49,0%
	Não	370	36,5%
	Não conheço o PPP da escola	142	14,0%
	Não informado	6	0,5%
Q6.4	Sim, Programa Saúde na Escola	206	20,3%
	Sim, Programa Primeira Infância	218	21,5%
	Não	424	41,8%
	Não sei	167	16,4%

Tabela A.7 (Continuação) Distribuição de frequências das variáveis relacionadas à transposição didática

Questão	Resposta	Contagem (n)	Porcentagem (%)
Q7.1	Sim	499	49,2%
	Não	509	50,1%
	Não informado	7	0,7%
Q7.2	Sim	526	51,8%
	Não	480	47,3%
	Não informado	9	0,9%
Q7.2a	Ações intersetoriais e parcerias	15	1,5%
	Alimentação saudável	129	12,7%
	Higiene pessoal e saúde bucal	94	9,3%
	Saúde mental e emocional	24	2,4%
	Vacinação e prevenção de doenças	39	3,8%
	Não informado	714	70,3%
Q8.1	Sim	223	22,0%
	Não	515	50,7%
	Não sei responder	277	27,3%
Q8.2	Sim	576	56,7%
	Não	428	42,2%
	Não informado	11	1,1%
Q8.2a	Alimentação saudável	74	7,3%
	Atividades físicas	25	2,5%
	Autocuidado e bem-estar	24	2,3%
	Conversas sobre saúde	21	2,1%
	Convivência e outros temas	6	0,6%
	Educação ambiental	19	1,8%
	Higiene pessoal e bucal	59	5,8%
	Prevenção de doenças	17	1,7%
	Não informado	770	75,9%

Tabela A.7 (Continuação) Distribuição de frequências das variáveis relacionadas à transposição didática

Questão	Resposta	Contagem (n)	Porcentagem (%)
Q8.3	Sim	491	48,4%
	Não	511	50,3%
	Não informado	13	1,3%
Q8.3a	Comentários variados	6	0,6%
	Conhecimento superficial	32	3,1%
	Desconhecimento completo do tema	17	1,6%
	Falta de apoio e recursos	9	0,9%
	Falta de formação específica	81	8,0%
	Formação básica não abrangente	10	1,0%
	Necessidade de aperfeiçoamento	27	2,7%
	Não informado	833	82,1%
Q8.4	Sim	601	59,3%
	Não	174	17,1%
	Não sei responder	240	23,6%
Q8.5	Currículos e programas	286	28,2%
	Formação sobre conteúdos e atividades	277	27,3%
	Materiais didáticos	216	21,2%
	Não falta nada e já trabalho com Educação para a Saúde	230	22,7%
	Não informado	6	0,6%
Q9.1	Sim	501	49,4%
	Não	182	17,9%
	Não sei responder	332	32,7%

Tabela A.7 (Continuação) Distribuição de frequências das variáveis relacionadas à transposição didática

Questão	Resposta	Contagem (n)	Porcentagem (%)
Q9.1a	Abertura da escola para comunicação com famílias	29	2,8%
	Atendimento individualizado e orientação direta	17	1,7%
	Dia da Família e eventos na escola	22	2,1%
	Formação e palestras sobre saúde	13	1,3%
	Reuniões com famílias e responsáveis	83	8,2%
	Outros	30	3,0%
	Não informado	821	80,9%
Q9.4	Adequada	810	79,8%
	Inadequada	203	20,0%
	Não informado	2	0,2%
Q9.4a	Espaço físico insuficiente ou inadequado	13	1,3%
	Falta de acessibilidade	6	0,6%
	Higiene ou limpeza inadequadas	3	0,3%
	Melhoria na infraestrutura geral	16	1,5%
	Problemas de acústica	8	0,8%
	Reformas e necessidade de ajustes	11	1,1%
	Salas com superlotação	8	0,8%
	Outros	5	0,5%
	Não informado	945	93,1%

Tabela A.7 (Continuação) Distribuição de frequências das variáveis relacionadas à transposição didática

Questão	Resposta	Contagem (n)	Porcentagem (%)
Q10.1	Alimentação saudável	81	7,9%
	Atividades físicas	54	5,2%
	Cuidado com o meio ambiente	21	2,1%
	Educação socioemocional	17	1,7%
	Higiene	76	7,5%
	Prevenção de doenças	32	3,2%
	Saúde bucal	31	3,1%
	Saúde mental	44	4,3%
	Vacinação	25	2,5%
	Não informado	634	62,5%

Tabela A.8 Medidas-resumo do Escore por variável explicativa, com os valores-p dos testes de igualdade de medianas. Os valores-p em destaque indicam que o teste de Kruskal-Wallis ou Mann-Whitney (Wilcoxon) foi significativo, ao nível de significância de 5%.

Variável	Categoria	n	Média	Desvio padrão	Mínimo	Mediana	Máximo	Valor-p
Faixa etária	18-29 anos	26	33,85	7,48	21,88	33,33	50,00	0,706
	30-35 anos	83	34,09	8,27	15,63	33,33	50,00	
	36-41 anos	176	33,91	6,80	14,58	33,33	50,00	
	42-47 anos	278	34,31	7,20	14,58	33,33	50,00	
	48-53 anos	200	33,81	7,30	0,00	33,33	50,00	
	54-59 anos	182	33,63	6,93	13,54	33,33	50,00	
	60-65 anos	52	32,09	7,62	13,54	32,81	50,00	
	66-75 anos	18	32,64	9,10	22,92	29,17	50,00	
Gênero	Feminino	870	33,89	7,37	0,00	33,33	50,00	0,762
	Masculino	139	33,58	6,65	15,63	33,33	50,00	
	Não binário	3	36,06	5,31	31,11	35,42	41,67	

Tabela A.8 (Continuação) Medidas-resumo do Escore por variável explicativa, com os valores-p dos testes de igualdade de medianas

Variável	Categoria	n	Média	Desvio padrão	Mínimo	Mediana	Máximo	Valor-p
Cor/Raça	Amarela	14	36,68	8,40	26,04	34,38	50,00	0,194
	Branca	535	34,23	7,22	13,54	33,33	50,00	
	Indígena	5	35,21	7,88	29,17	32,29	47,92	
	Parda	290	33,09	6,87	14,58	32,81	50,00	
	Preta	163	33,58	7,78	0,00	33,33	50,00	
Cargo	Assistente de diretor	30	36,01	7,05	22,92	33,33	50,00	0,002
	Coordenador	34	37,29	7,18	22,92	34,90	50,00	
	Diretor	23	32,16	6,75	13,54	31,25	43,75	
	Professor Ed. Infantil	171	32,38	7,47	13,54	32,29	50,00	
	Professor Fund. I	520	33,89	7,11	0,00	33,33	50,00	
	Professor Fund. II e Médio	237	34,22	7,28	15,63	33,33	50,00	
Função	Gestor	87	35,49	7,25	13,54	33,33	50,00	0,026
	Professor	928	33,69	7,24	0,00	33,33	50,00	
Tempo de trabalho	Menos de 3 anos	221	34,08	6,71	15,63	33,33	50,00	0,627
	3 a 4 anos	48	34,20	7,81	18,75	33,33	50,00	
	5 a 10 anos	256	33,29	7,13	13,54	32,29	50,00	
	11 a 15 anos	184	33,53	7,11	13,54	33,33	50,00	
	16 a 20 anos	134	34,81	7,95	14,58	33,33	50,00	
	21 a 25 anos	96	33,48	7,56	0,00	33,33	50,00	
	Mais de 25 anos	75	34,35	7,57	17,71	33,33	50,00	

Tabela A.8 (Continuação) Medidas-resumo do Escore por variável explicativa, com os valores-p dos testes de igualdade de medianas

Variável	Categoria	n	Média	Desvio padrão	Mínimo	Mediana	Máximo	Valor-p
Escolaridade	Ensino médio	24	30,71	5,97	18,75	31,77	48,96	0,125
	Graduação	328	33,49	6,90	13,54	33,33	50,00	
	Pós-graduação	609	34,14	7,46	0,00	33,33	50,00	
	Mestrado	42	33,21	7,07	13,54	33,33	48,96	
	Doutorado	7	37,50	8,69	27,08	34,38	50,00	
	Pós-doutorado	5	36,88	6,97	31,25	35,42	48,96	
Tipo de educação básica	Privada	76	35,70	7,69	18,75	33,33	50,00	0,026
	Pública	831	33,52	7,25	0,00	33,33	50,00	
	Pública e privada	108	35,09	6,68	20,83	33,33	50,00	
Tipo de graduação	Privada	801	33,81	7,34	0,00	33,33	50,00	0,838
	Pública	151	34,16	6,80	17,71	33,33	50,00	
	Pública e privada	63	33,58	7,32	14,58	33,33	50,00	
Renda	2 a 3 sal. mínimos	225	33,26	7,31	14,58	33,33	50,00	0,005
	4 a 5 sal. mínimos	463	33,97	7,34	13,54	33,33	50,00	
	6 a 7 sal. mínimos	184	33,00	7,15	0,00	32,29	50,00	
	8 a 9 sal. mínimos	99	35,11	6,59	21,88	33,33	50,00	
	10 sal. mínimos ou mais	42	36,28	7,37	13,54	34,38	50,00	

Tabela A.8 (Continuação) Medidas-resumo do Escore por variável explicativa, com os valores-p dos testes de igualdade de medianas

Variável	Categoria	n	Média	Desvio padrão	Mínimo	Mediana	Máximo	Valor-p
Q5.1	Sim, frequentemente	651	34,61	7,22	13,54	33,33	50,00	< 0,001
	Sim, mas tem diminuído	333	32,40	7,19	0,00	32,29	50,00	
	Não	26	32,61	6,40	18,75	32,29	43,75	
Q5.2	Sim	759	34,11	6,90	13,54	33,33	50,00	0,029
	Não	256	33,07	8,18	0,00	32,29	50,00	
Q6.1	Sim	142	35,09	6,64	15,63	33,33	50,00	0,011
	Não	871	33,66	7,34	0,00	33,33	50,00	
Q6.2	Sim	129	35,32	7,02	20,83	33,33	50,00	0,014
	Não	885	33,64	7,27	0,00	33,33	50,00	
Q9.2	Sim	326	35,22	7,26	14,58	33,33	50,00	< 0,001
	Não	683	33,14	7,15	0,00	32,29	50,00	
Q9.3	Sim	298	35,15	7,54	0,00	33,33	50,00	< 0,001
	Não	705	33,28	7,10	13,54	32,29	50,00	
Q6.3	Sim	497	34,43	7,38	0,00	33,33	50,00	0,029
	Não	370	33,32	7,18	13,54	32,29	50,00	
	Não conheço o PPP da escola	142	33,23	6,98	13,54	33,33	50,00	
Q6.4	Sim, Programa Saúde na Escola	206	34,59	6,96	16,67	33,33	50,00	0,055
	Sim, Programa Primeira Infância	218	34,43	7,15	14,58	33,33	50,00	
	Não	424	33,64	7,57	0,00	33,33	50,00	
Q7.1	Sim	499	34,69	7,04	13,54	33,33	50,00	< 0,001
	Não	509	33,06	7,40	0,00	32,29	50,00	
Q7.2	Sim	526	34,50	7,13	14,58	33,33	50,00	0,014
	Não	480	33,15	7,29	0,00	33,33	50,00	

Tabela A.8 (Continuação) Medidas-resumo do Escore por variável explicativa, com os valores-p dos testes de igualdade de medianas

Variável	Categoria	n	Média	Desvio padrão	Mínimo	Mediana	Máximo	Valor-p
Q8.1	Sim	223	35,52	7,41	13,54	33,33	50,00	< 0,001
	Não	515	32,93	7,14	0,00	32,29	50,00	
Q8.2	Sim	576	34,28	7,35	0,00	33,33	50,00	0,024
	Não	428	33,28	7,14	13,54	33,33	50,00	
Q8.3	Sim	491	34,89	7,30	0,00	33,33	50,00	< 0,001
	Não	511	32,86	7,12	13,54	32,29	50,00	
Q8.4	Sim	601	34,28	7,30	0,00	33,33	50,00	0,007
	Não	174	32,60	7,46	13,54	32,29	50,00	
Q8.5	Currículos e programas	286	33,12	7,58	13,54	32,29	50,00	0,078
	Formação sobre conteúdos e atividades	277	33,85	7,06	14,58	33,33	50,00	
	Materiais didáticos	216	34,10	7,20	0,00	33,33	50,00	
	Não falta nada e já trabalho com ES	230	34,54	7,10	16,67	33,33	50,00	
Q9.1	Sim	501	34,94	7,12	0,00	33,33	50,00	< 0,001
	Não	182	31,82	7,76	13,54	31,25	50,00	
Q9.4	Adequada	810	33,98	7,19	0,00	33,33	50,00	0,169
	Inadequada	203	33,33	7,53	13,54	32,29	50,00	
Q10.1	Alimentação saudável	81	34,44	7,71	17,71	33,33	50,00	0,514
	Atividades físicas	54	33,68	8,89	0,00	33,33	50,00	
	Cuidado meio ambiente	21	37,10	7,35	25,00	36,46	50,00	
	Educação socioemocional	17	33,33	6,47	23,96	32,29	48,96	
	Higiene	76	34,58	6,87	22,92	33,33	50,00	
	Prevenção de doenças	32	33,33	7,28	15,63	31,77	50,00	
	Saúde bucal	31	33,17	8,04	20,83	32,29	50,00	
	Saúde mental	44	34,40	6,05	19,79	33,85	48,96	
	Vacinação	25	33,08	6,84	23,96	31,25	50,00	

Tabela A.9 Valores do Alfa de Cronbach

Domínio	Alfa de Cronbach
Letramento global	0,92
Cuidados da saúde	0,84
Prevenção da doença	0,78
Promoção da saúde	0,81

Tabela A.10 Correlação de Spearman entre as questões do HLS-Q16

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16
Q1	1															
Q2	0,605 *	1														
Q3	0,414 *	0,460 *	1													
Q4	0,383 *	0,400 *	0,613 *	1												
Q5	0,352 *	0,429 *	0,458 *	0,343 *	1											
Q6	0,380 *	0,420 *	0,513 *	0,447 *	0,564 *	1										
Q7	0,374 *	0,435 *	0,459 *	0,522 *	0,390 *	0,547 *	1									
Q8	0,445 *	0,536 *	0,446 *	0,384 *	0,484 *	0,489 *	0,438 *	1								
Q9	0,329 *	0,320 *	0,417 *	0,440 *	0,289 *	0,365 *	0,410 *	0,371 *	1							
Q10	0,323 *	0,354 *	0,416 *	0,473 *	0,341 *	0,376 *	0,435 *	0,341 *	0,628 *	1						
Q11	0,336 *	0,335 *	0,400 *	0,355 *	0,386 *	0,457 *	0,365 *	0,394 *	0,397 *	0,396 *	1					
Q12	0,337 *	0,330 *	0,406 *	0,393 *	0,392 *	0,450 *	0,339 *	0,411 *	0,376 *	0,408 *	0,641 *	1				
Q13	0,412 *	0,404 *	0,387 *	0,409 *	0,366 *	0,383 *	0,382 *	0,500 *	0,542 *	0,492 *	0,434 *	0,460 *	1			
Q14	0,331 *	0,339 *	0,398 *	0,367 *	0,398 *	0,421 *	0,423 *	0,419 *	0,462 *	0,452 *	0,366 *	0,387 *	0,523 *	1		
Q15	0,417 *	0,367 *	0,448 *	0,446 *	0,340 *	0,402 *	0,468 *	0,451 *	0,561 *	0,551 *	0,464 *	0,499 *	0,613 *	0,565 *	1	
Q16	0,284 *	0,341 *	0,345 *	0,327 *	0,343 *	0,422 *	0,387 *	0,386 *	0,516 *	0,552 *	0,404 *	0,388 *	0,538 *	0,502 *	0,534 *	1

Legenda: *valor-p < 0,001.

Tabela A.11 Estimativas dos parâmetros associados a μ no modelo GAMLSS com distribuição Box-Cox t ajustado aos dados

Efeito	Estimativa	Erro padrão	Valor-t	Valor-p
Intercepto	43,390	6,663	6,512	< 0,001 **
Cargo Coordenador	0,519	1,700	0,306	0,760
Cargo Diretor	-2,462	1,702	-1,446	0,148
Cargo Prof. Ed. Infantil	-3,849	1,397	-2,756	0,006 **
Cargo Prof. Ed. Infantil e Fund I	-2,360	1,332	-1,771	0,077
Cargo Prof Fund. II e Médio	-1,778	1,364	-1,304	0,193
Escolaridade Ensino Médio	-3,869	3,034	-1,275	0,202
Escolaridade Graduação	-1,533	2,864	-0,535	0,593
Escolaridade Mestrado	-2,890	2,882	-1,003	0,316
Escolaridade Pós-doutorado	0,912	3,856	0,237	0,813
Escolaridade Pós-graduação	-0,404	2,823	-0,143	0,886
Tipo de educação básica Pública	-8,043	4,717	-1,705	0,088
Tipo de educação básica Pública e privada	-7,023	4,706	-1,493	0,136
Renda 2 a 3 sal. mín.	-2,052	1,100	-1,866	0,062
Renda 4 a 5 sal. mín.	-1,201	1,011	-1,188	0,235
Renda 6 a 7 sal. mín.	-2,580	1,062	-2,430	0,015 *
Renda 8 a 9 sal. mín.	-1,359	1,109	-1,225	0,221

Legenda: *O valor-p é significante para um nível de significância de 5%; **O valor-p é significante para um nível de significância de 1%.

Tabela A.11 (Continuação) Estimativas dos parâmetros associados a μ no modelo GAMLSS com distribuição Box-Cox t ajustado aos dados

Efeito	Estimativa	Erro padrão	Valor-t	Valor-p
Q5.1 Sim, frequentemente	0,254	1,311	0,194	0,846
Q5.1 Sim, mas tem diminuído	-1,386	1,307	-1,060	0,289
Q8.1 Não sei	0,985	0,513	1,921	0,055
Q8.1 Sim	1,346	0,547	2,463	0,013 *
Q8.3 Sim	1,143	0,436	2,620	0,009 **
Q9.1 Não sei	1,472	0,609	2,417	0,016 *
Q9.1 Sim	1,969	0,567	3,474	< 0,001 **
Q9.2 Sim	1,221	0,493	2,475	0,013 *

Legenda: *O valor-p é significativo para um nível de significância de 5%; **O valor-p é significativo para um nível de significância de 1%.

Tabela A.12 Estimativas dos parâmetros associados a $\log(\sigma)$ no modelo GAMLSS com distribuição Box-Cox t ajustado aos dados

Efeito	Estimativa	Erro padrão	Valor-t	Valor-p
Intercepto	-0,303	0,616	-0,491	0,623
Gênero Masculino	-0,136	0,076	-1,785	0,075
Gênero Não binário	-1,202	0,658	-1,827	0,068
Tipo de educação básica Pública	-1,392	0,620	-2,244	0,025 *
Tipo de educação básica Pública e privada	-1,585	0,632	-2,508	0,012 *
Q5.1 Sim, frequentemente	0,410	0,202	2,033	0,042 *
Q5.1 Sim, mas tem diminuído	0,197	0,201	0,980	0,327
Q5.2 Sim	-0,274	0,065	-4,223	< 0,001 **
Q6.1 Sim	-0,188	0,081	-2,329	0,020 *
Q7.1 Sim	-0,140	0,055	-2,553	0,011 *
Q8.3 Sim	0,037	0,058	0,637	0,524

Legenda: *O valor-p é significativo para um nível de significância de 5%; **O valor-p é significativo para um nível de significância de 1%.

Tabela A.13 Estimativas dos parâmetros associados a ν no modelo GAMLSS com distribuição Box-Cox t ajustado aos dados

Efeito	Estimativa	Erro padrão	Valor-t	Valor-p
Intercepto	-5,823	1,981	-2,927	0,003 **
Gênero Masculino	0,681	0,474	1,435	0,152
Gênero Não binário	-11,890	15,829	-0,751	0,453
Tipo de educação básica Pública	4,071	0,621	6,552	< 0,001 **
Tipo de educação básica Pública e privada	3,740	0,808	4,626	< 0,001 **
Renda 2 a 3 sal. mín.	3,828	2,169	1,765	0,078
Renda 4 a 5 sal. mín.	3,912	2,130	1,845	0,065
Renda 6 a 7 sal. mín.	3,295	2,191	1,504	0,133
Renda 8 a 9 sal. mín.	2,255	2,004	1,125	0,261
Q5.1 Sim, frequentemente	-1,734	0,948	-1,829	0,068
Q5.1 Sim, mas tem diminuído	-1,263	0,957	-1,320	0,187
Q6.4 Não sei	0,364	0,414	0,878	0,380
Q6.4 Sim, Programa Primeira Infância	0,745	0,390	1,910	0,056
Q6.4 Sim, Programa Saúde na Escola	-0,676	0,497	-1,360	0,174
Q7.2 Sim	-0,619	0,338	-1,831	0,067
Q9.2 Sim	0,617	0,400	1,540	0,124
Q9.3 Sim	-0,917	0,397	-2,312	0,021 *

Legenda: *O valor-p é significativo para um nível de significância de 5%; **O valor-p é significativo para um nível de significância de 1%.

Tabela A.14 Estimativas dos parâmetros associados a $\log(\tau)$ no modelo GAMLSS com distribuição Box-Cox t ajustado aos dados

Efeito	Estimativa	Erro padrão	Valor-t	Valor-p
Intercepto	0,937	3,775	0,248	0,804
Idade	0,239	0,102	2,352	0,019 *
Tempo de trabalho 16 a 20 anos	0,852	1,775	0,480	0,631
Tempo de trabalho 21 a 25 anos	-5,207	1,813	-2,873	0,004 **
Tempo de trabalho 3 a 4 anos	0,126	1,219	0,103	0,918
Tempo de trabalho 5 a 10 anos	0,999	1,325	0,754	0,451
Tempo de trabalho Mais de 25 anos	-0,012	3,740	-0,003	0,997
Tempo de trabalho Menos de 3 anos	3,603	3,787	0,951	0,341
Tipo de graduação Pública	-1,300	1,028	-1,264	0,207
Tipo de graduação Pública e privada	-3,295	1,665	-1,979	0,048 *
Q6.1 Sim	-0,696	0,896	-0,776	0,438
Q6.2 Sim	0,902	1,057	0,853	0,394
Q6.3 Não conheço o PPP	0,030	4,807	0,006	0,995
Q6.3 Sim	-3,550	2,649	-1,340	0,181
Q8.3 Sim	-1,960	1,286	-1,524	0,128
Q9.3 Sim	-3,477	1,397	-2,489	0,013 *
Q9.4 Inadequada	2,232	1,560	1,430	0,153

Legenda: *O valor-p é significativo para um nível de significância de 5%; **O valor-p é significativo para um nível de significância de 1%.

Tabela A.15 Medidas-resumo do Escore de cada domínio pelo cluster obtido no ajuste do modelo GMM

Domínio	Cluster	Média	Desvio padrão	Mínimo	Mediana	Máximo
CS	0	42,89	7,82	23,81	45,25	50,00
CS	1	31,81	6,54	0,00	33,33	50,00
PD	0	45,16	5,70	30,00	46,67	50,00
PD	1	31,67	7,05	0,00	33,33	50,00
PS	0	50,00	0,00	50,00	50,00	50,00
PS	1	33,07	6,69	0,00	33,33	45,83

Tabela A.16 Estimativas dos parâmetros associados a $\log \left[\frac{\pi}{1-\pi} \right]$ no modelo logístico binomial ajustado aos dados

Efeito	Estimativa	Erro padrão	Valor-t	Valor-p
Intercepto	16,579	482,147	0,034	0,973
Tipo de graduação Pública	-0,032	0,271	-0,118	0,906
Tipo de graduação Pública e privada	1,164	0,611	1,905	0,057
Q5.1 Sim, frequentemente	-15,172	482,146	-0,031	0,975
Q5.1 Sim, mas tem diminuído	-14,614	482,146	-0,030	0,976
Q5.2 Sim	0,730	0,237	3,076	0,002 **
Q6.2 Sim	-0,436	0,266	-1,641	0,101
Q6.4 Não sei	0,752	0,341	2,205	0,027 *
Q6.4 Sim, Programa Primeira Infância	0,130	0,249	0,522	0,601
Q6.4 Sim, Programa Saúde na Escola	0,409	0,274	1,492	0,136
Q8.1 Não sei responder	-0,436	0,240	-1,816	0,069
Q8.1 Sim	-0,560	0,255	-2,194	0,028 *
Q8.3 Sim	-0,392	0,216	-1,810	0,070

Legenda: *O valor-p é significativo para um nível de significância de 5%; **O valor-p é significativo para um nível de significância de 1%.

APÊNDICE B

Figuras

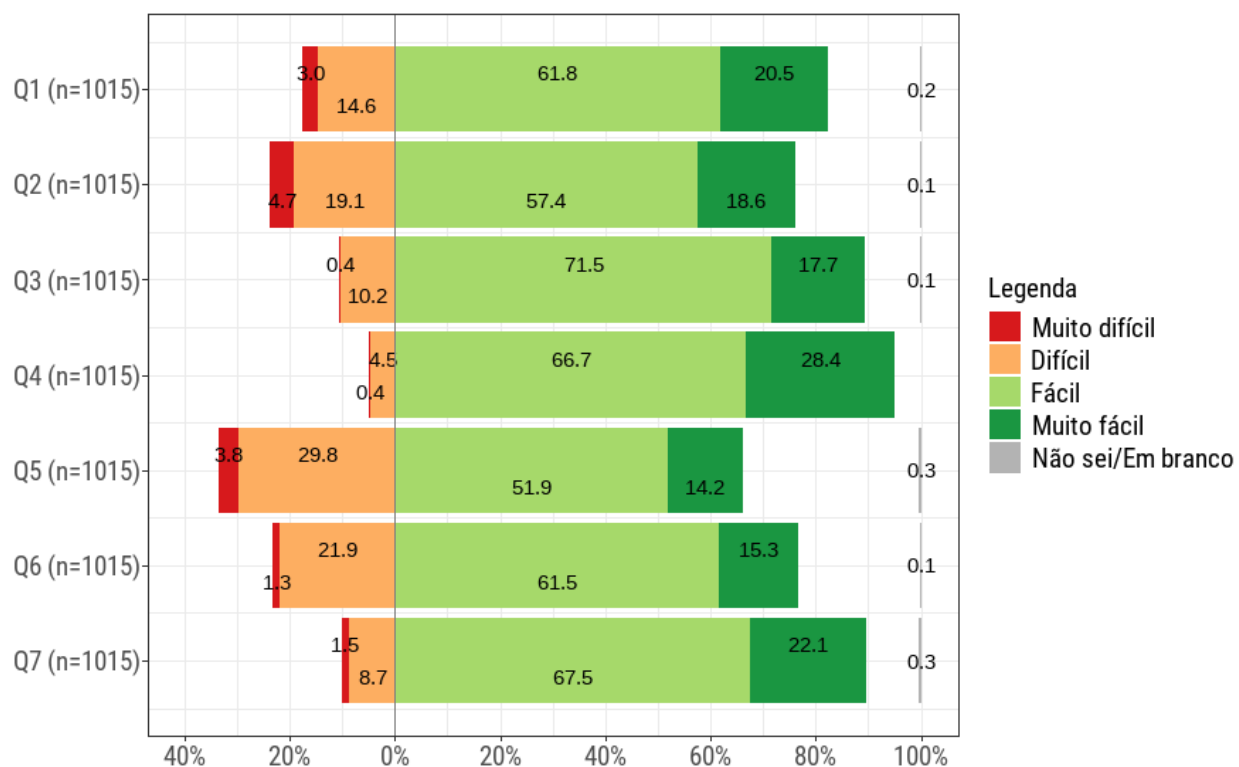


Figura B.1 Distribuição de frequências das questões do HLS-Q16 para o domínio Cuidados da saúde

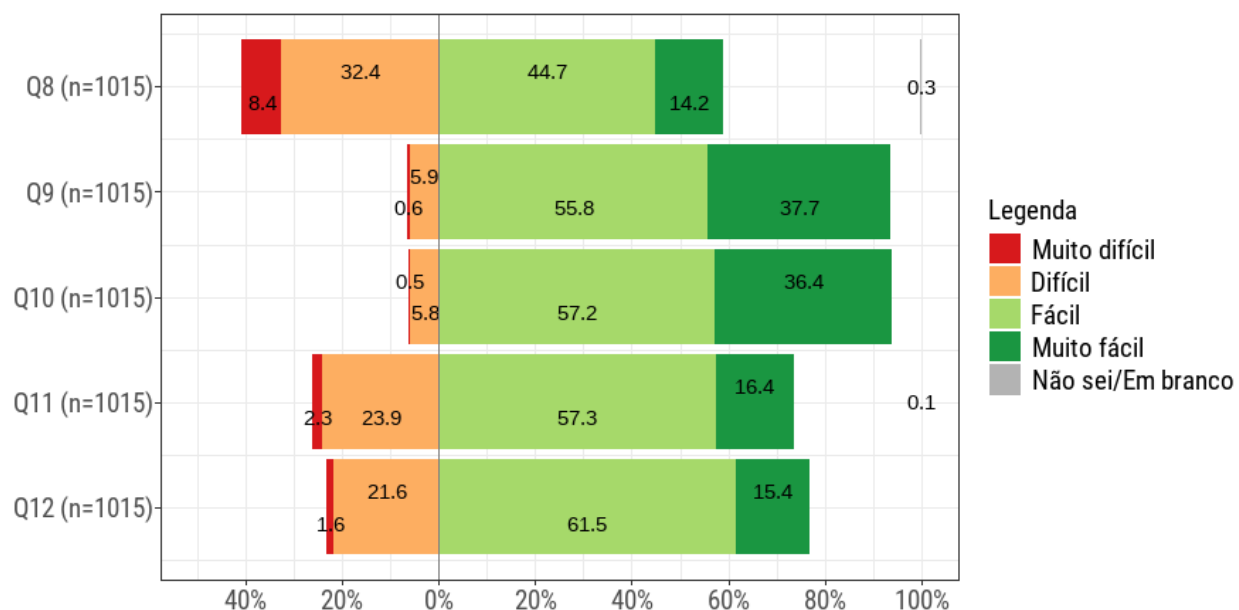


Figura B.2 Distribuição de frequências das questões do HLS-Q16 para o domínio Prevenção da doença

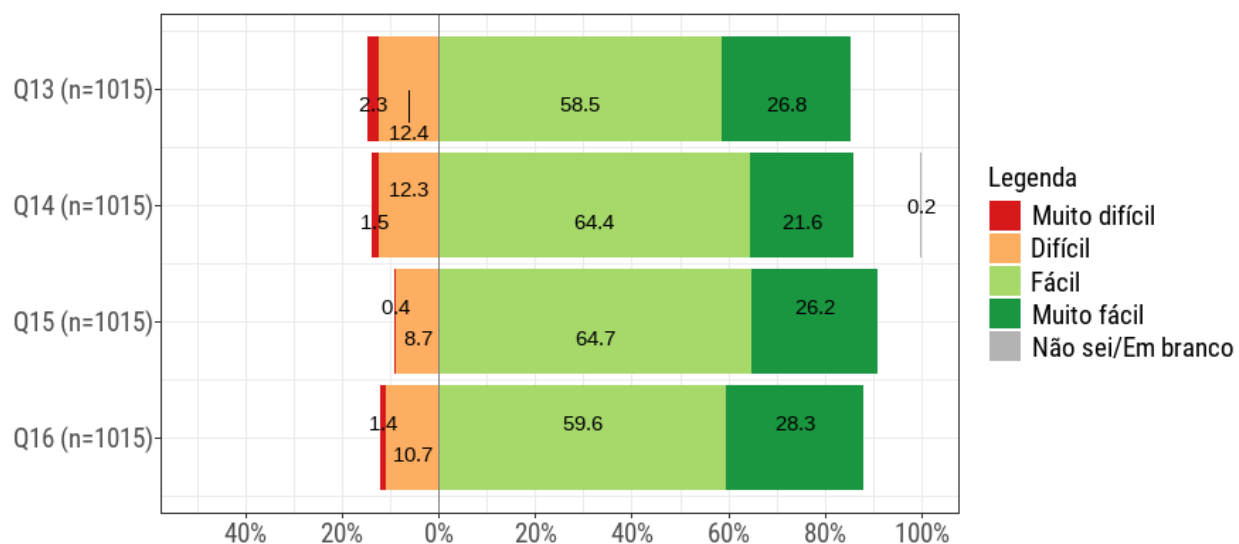


Figura B.3 Distribuição de frequências das questões do HLS-Q16 para o domínio Promoção da saúde

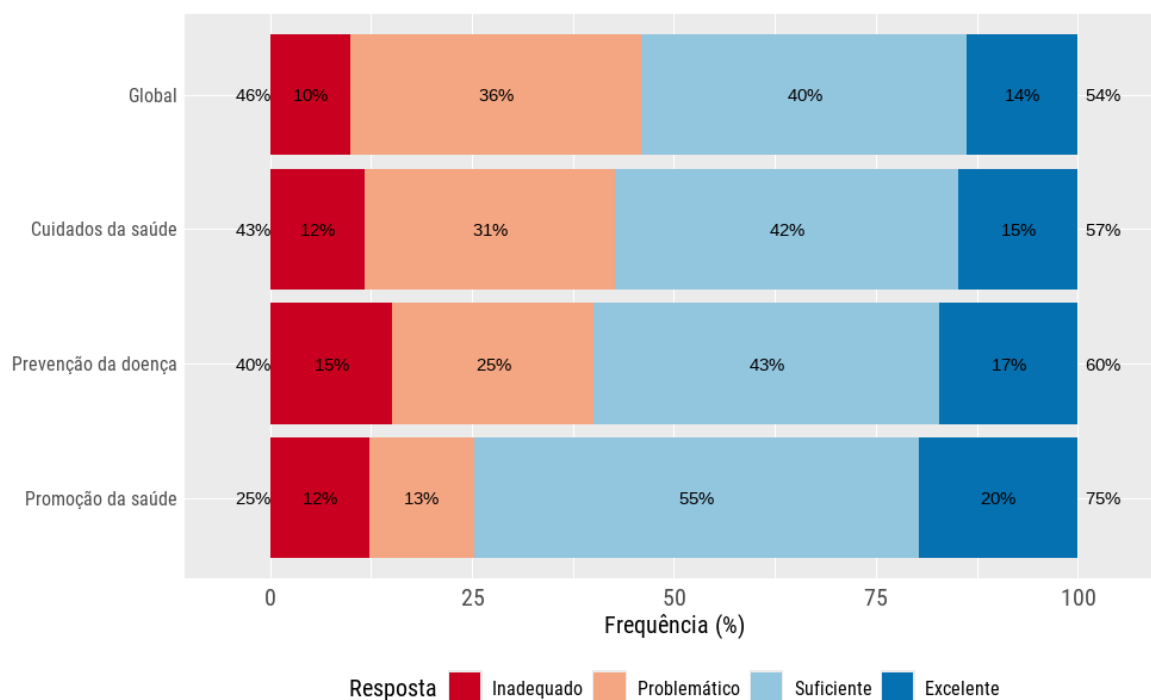


Figura B.4 Distribuição de frequências do Letramento global e por domínio. As porcentagens na extremidade do lado esquerdo referem-se à soma das porcentagens de classificações negativas (Inadequado e Problemático); as porcentagens na extremidade do lado direito referem-se à soma das porcentagens de classificações positivas (Suficiente e Excelente).

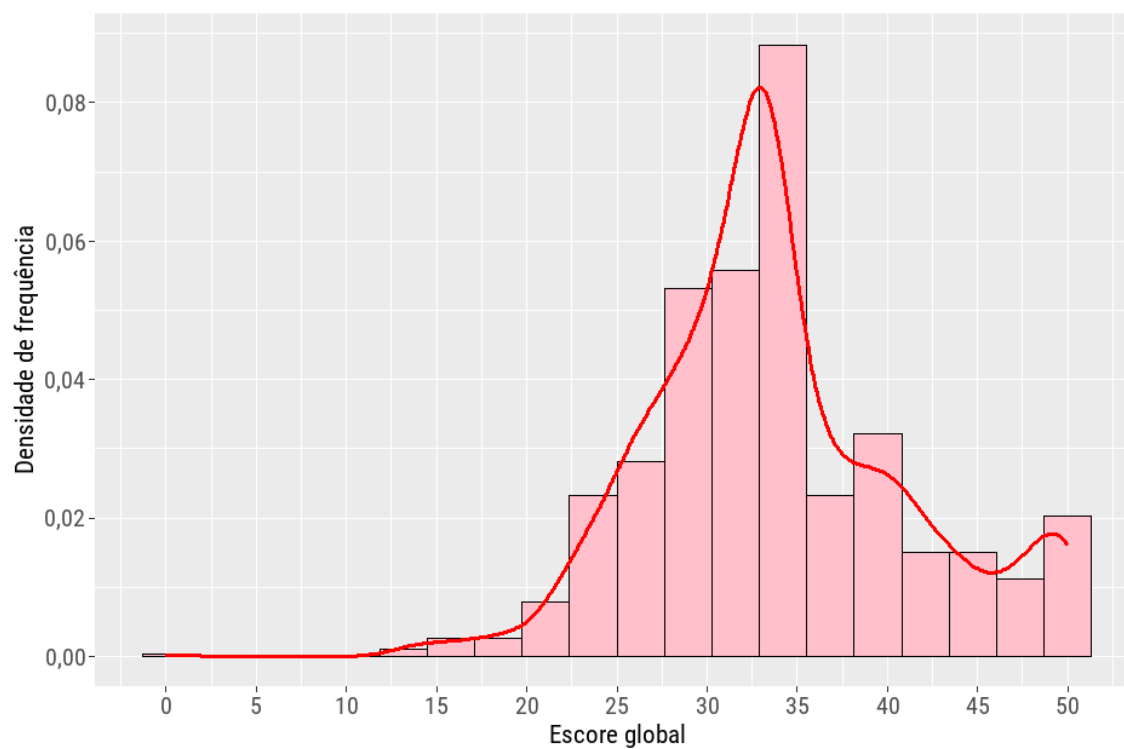


Figura B.5 Histograma do Escore com densidade suavizada

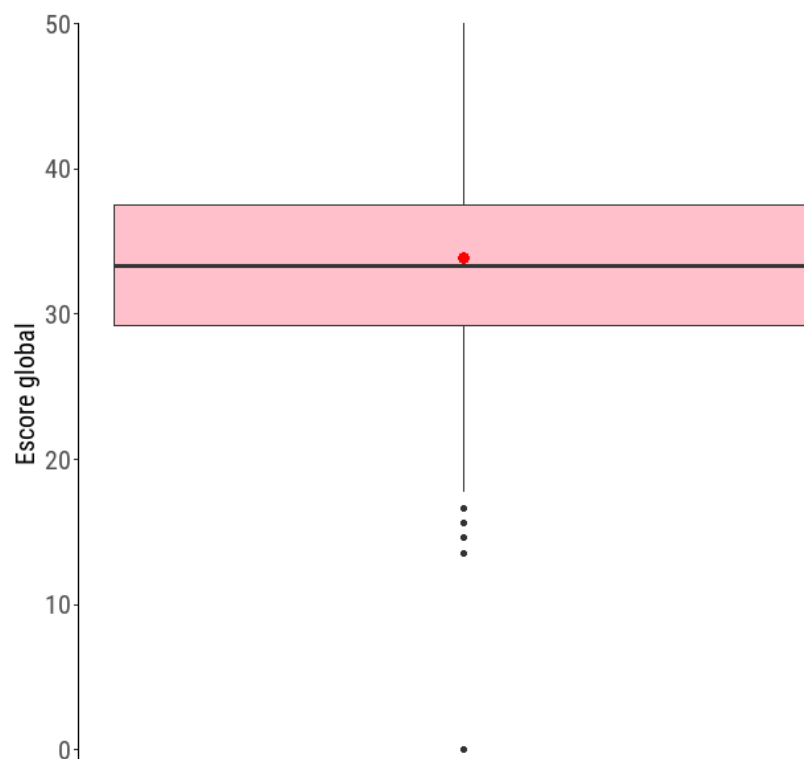


Figura B.6 *Box plot* do Escore global

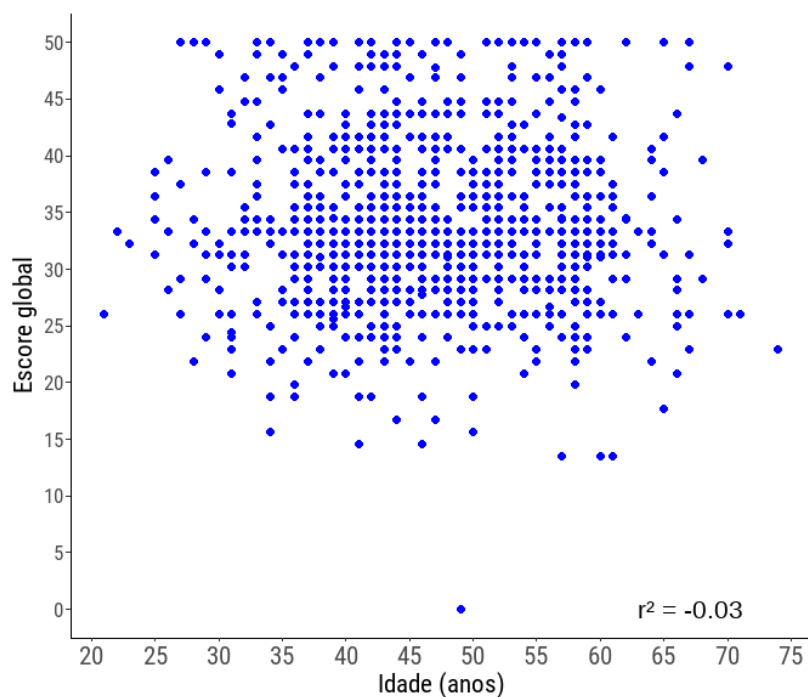


Figura B.7 Gráfico de dispersão entre Idade e Escore global

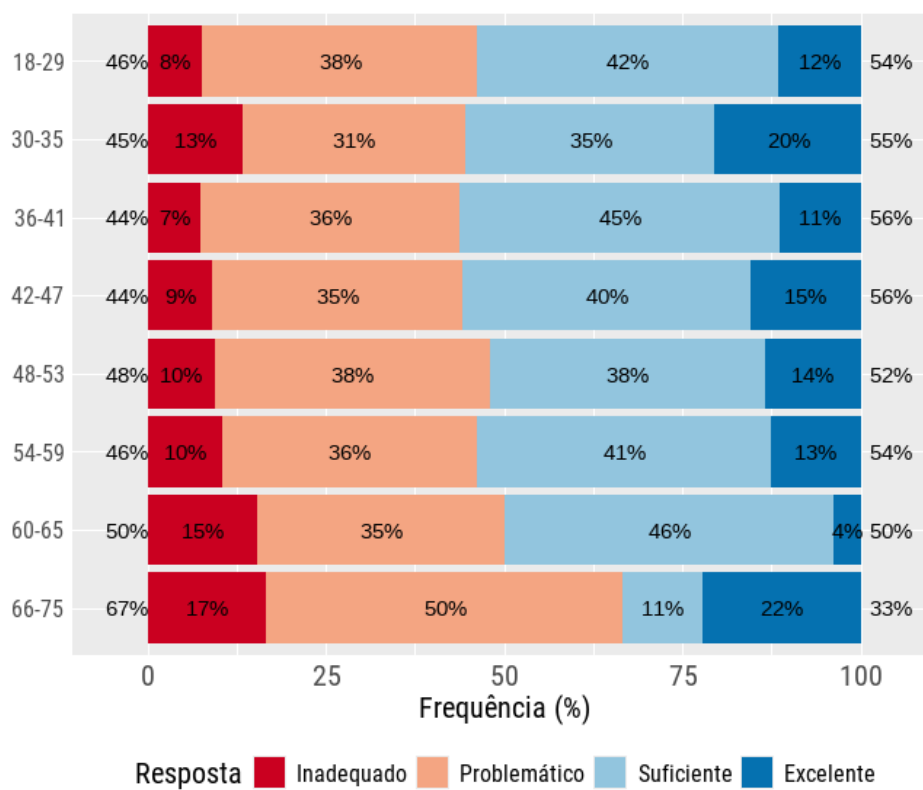


Figura B.8 Distribuição do Letramento global por Faixa etária

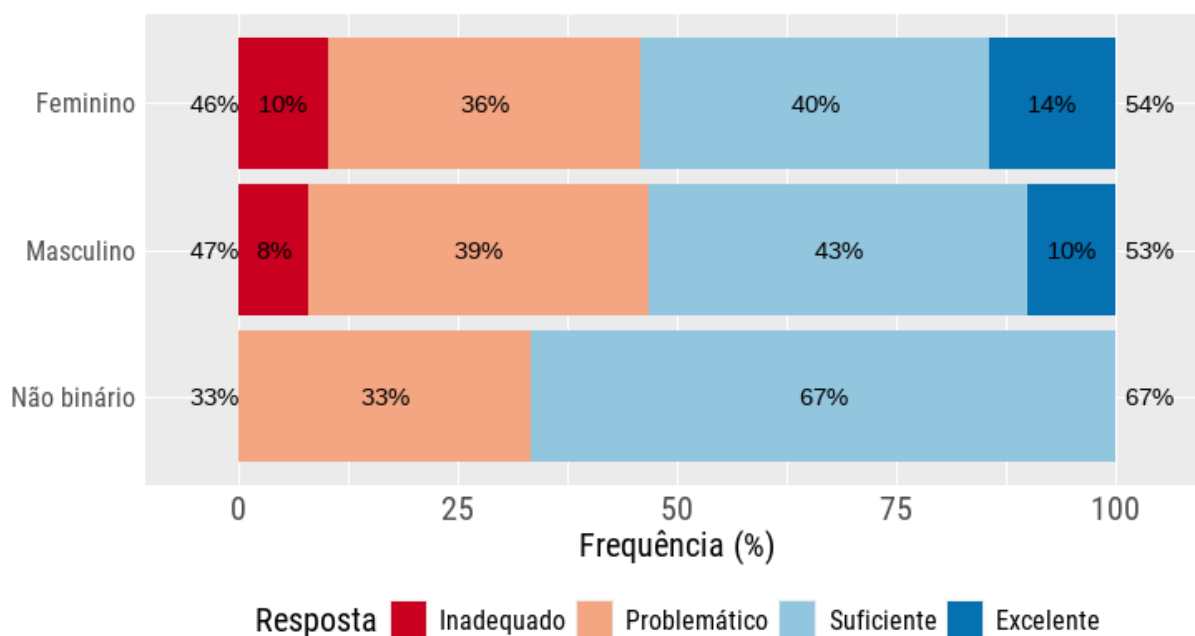


Figura B.9 Distribuição do Letramento global por Gênero

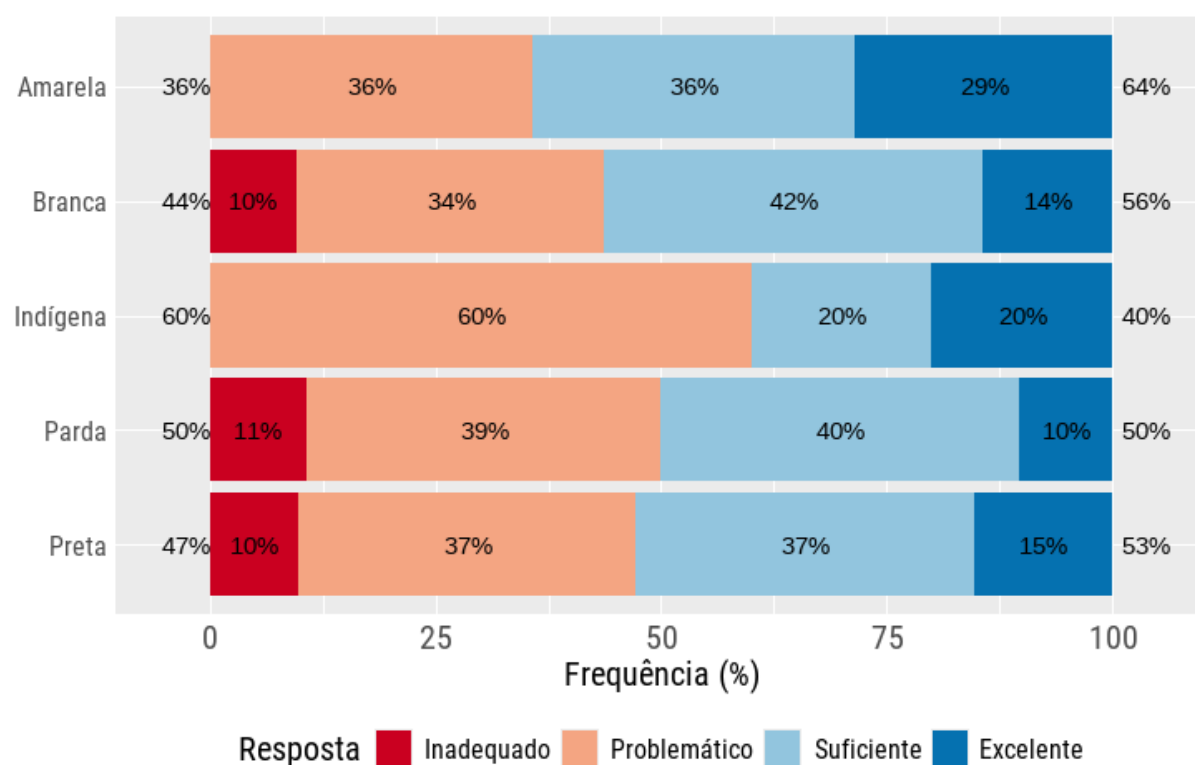


Figura B.10 Distribuição do Letramento global por Cor/Raça

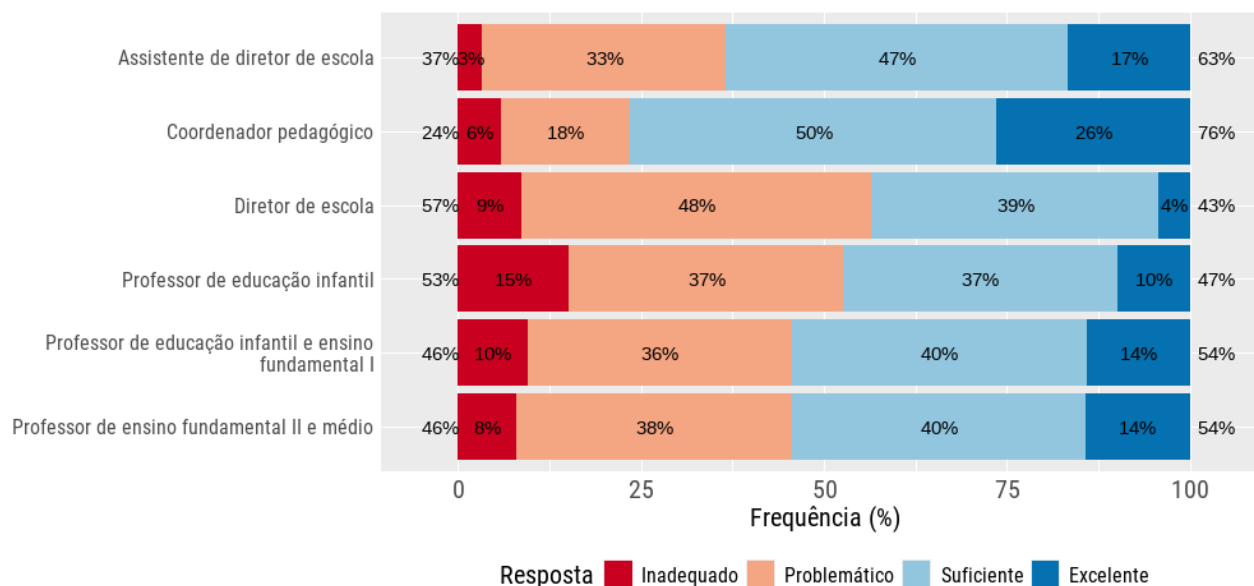


Figura B.11 Distribuição do Letramento global por Cargo

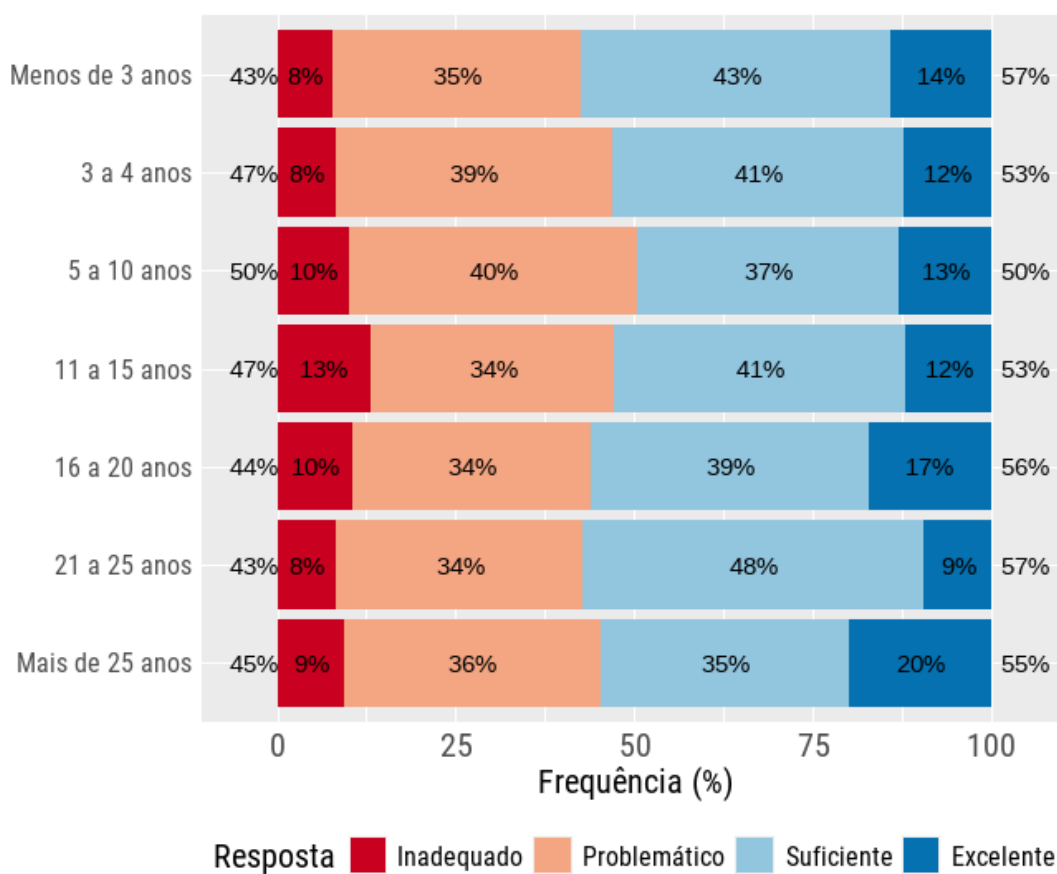


Figura B.12 Distribuição do Letramento global por Tempo de trabalho

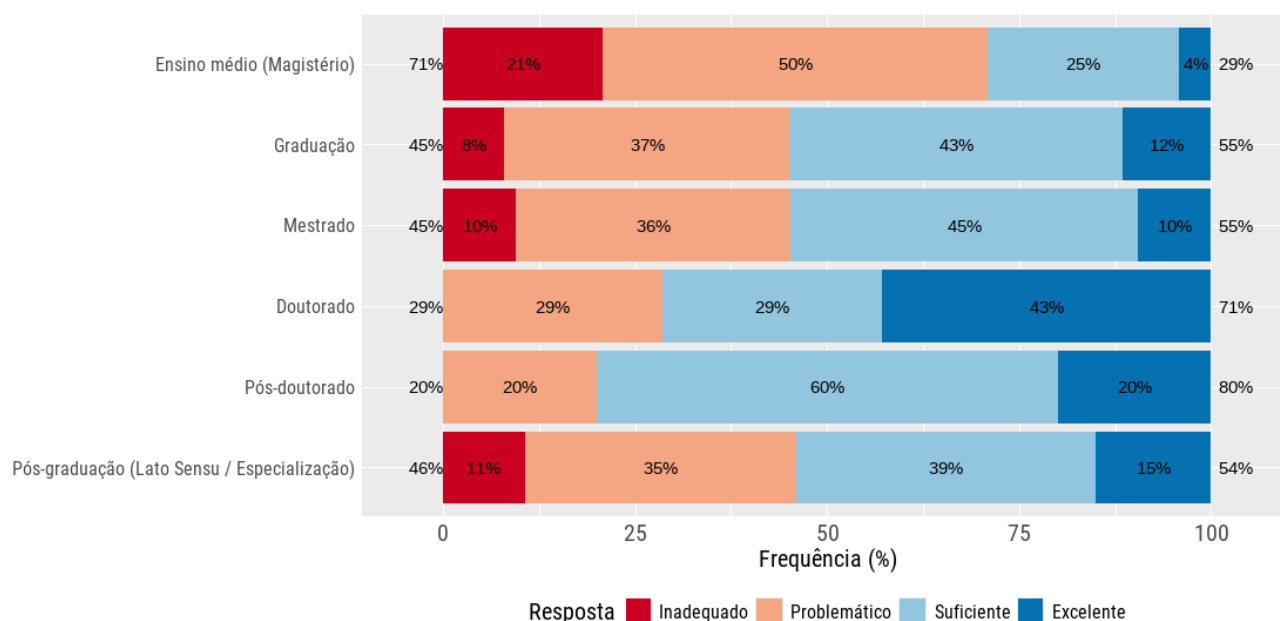


Figura B.13 Distribuição do Letramento global por Escolaridade

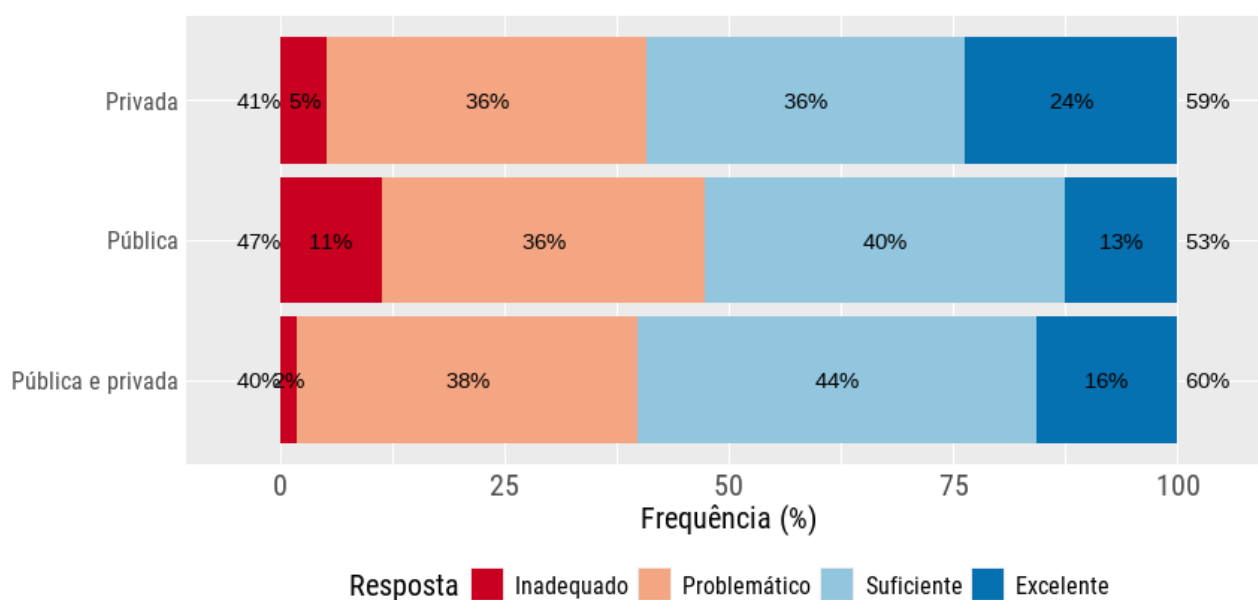


Figura B.14 Distribuição do Letramento global por Tipo de educação básica

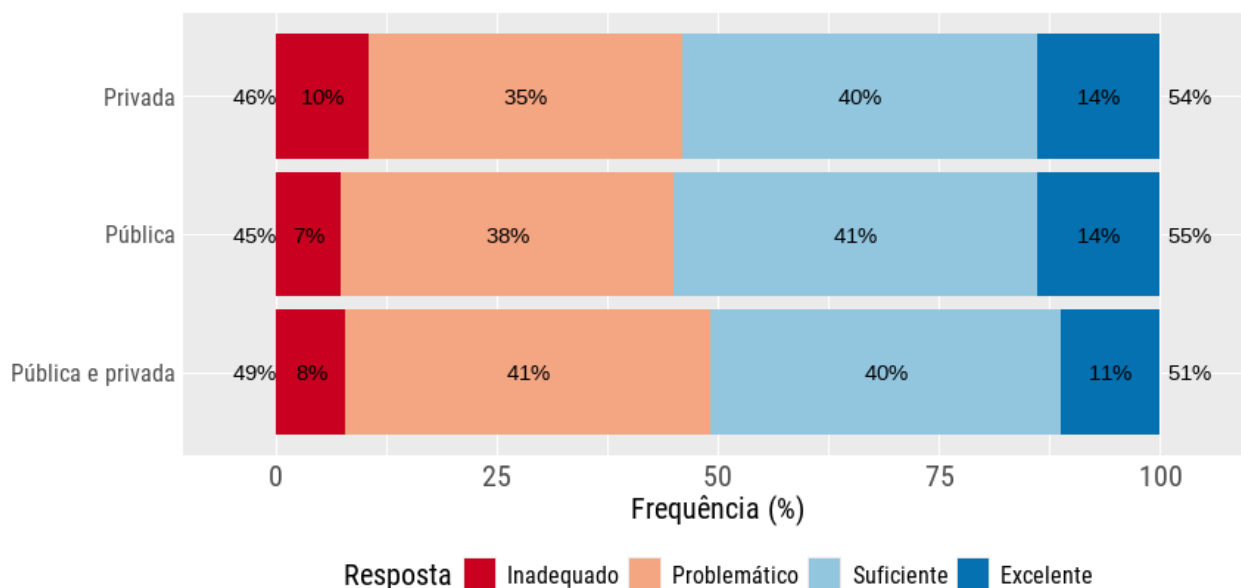


Figura B.15 Distribuição do Letramento global por Tipo de graduação

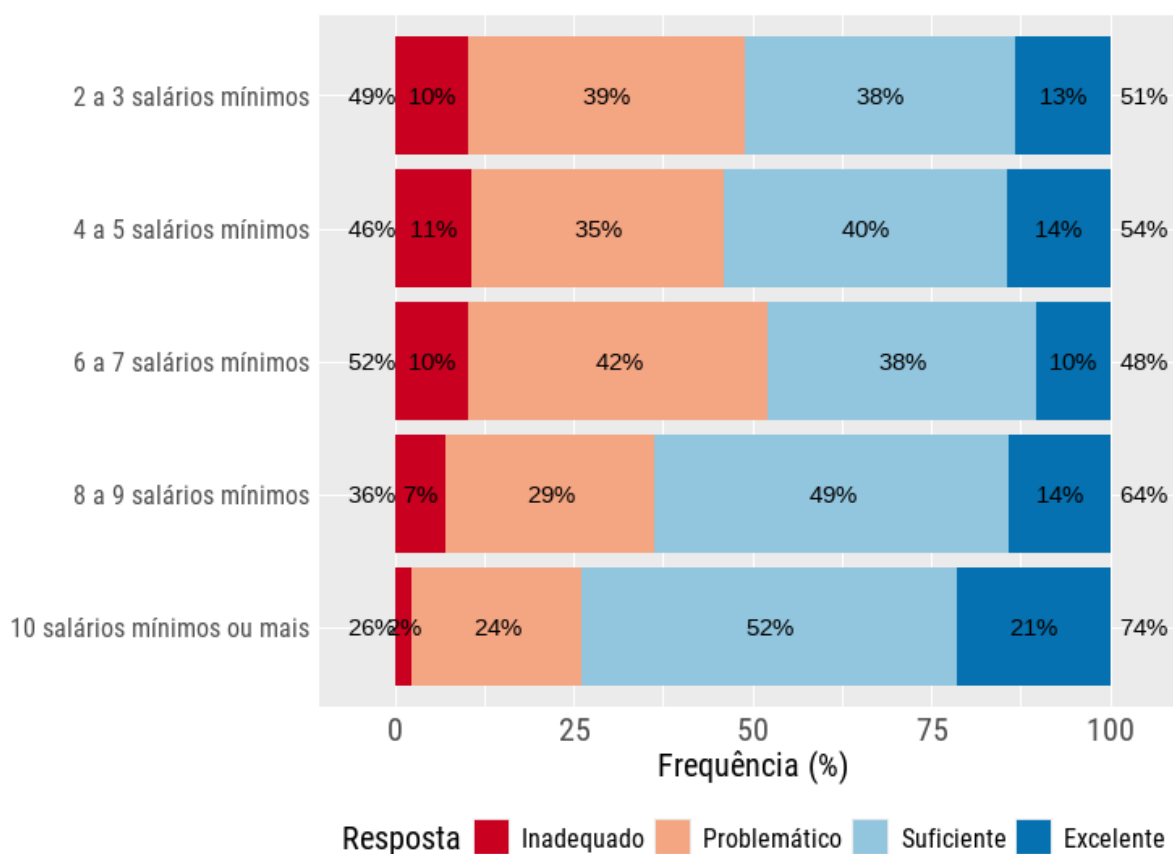


Figura B.16 Distribuição do Letramento global por Renda

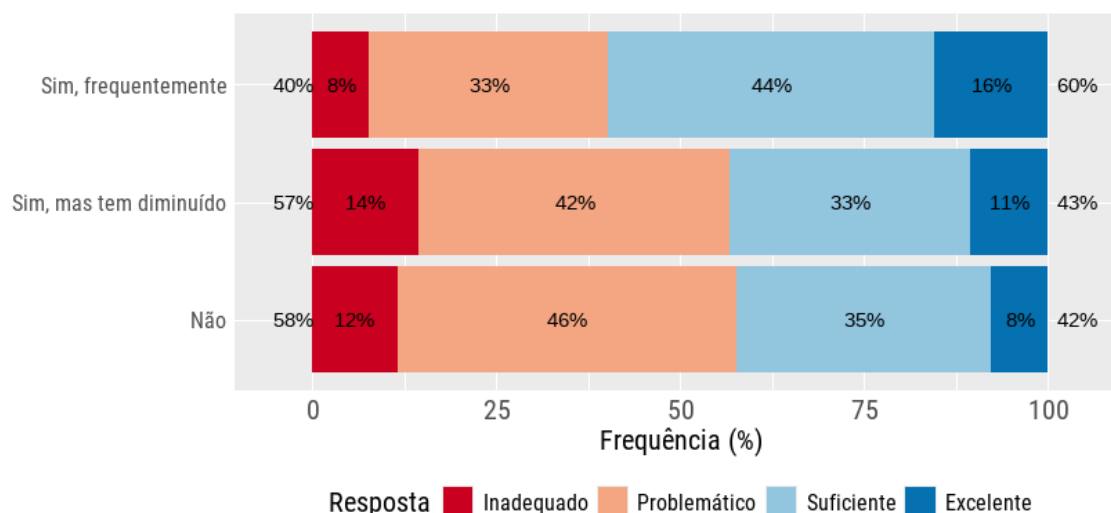


Figura B.17 Distribuição do Letramento global por Q5.1

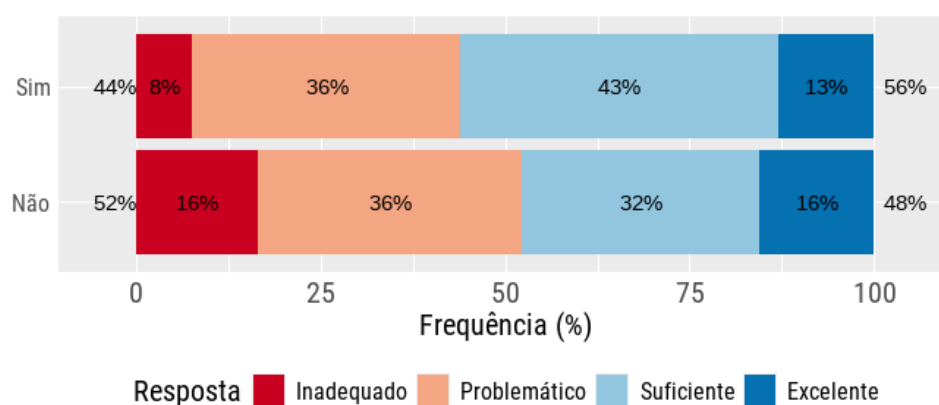


Figura B.18 Distribuição do Letramento global por Q5.2

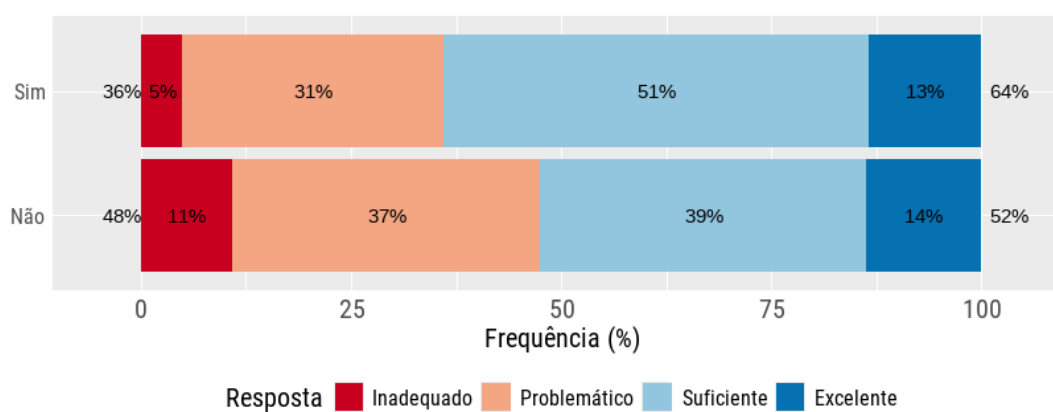


Figura B.19 Distribuição do Letramento global por Q6.1

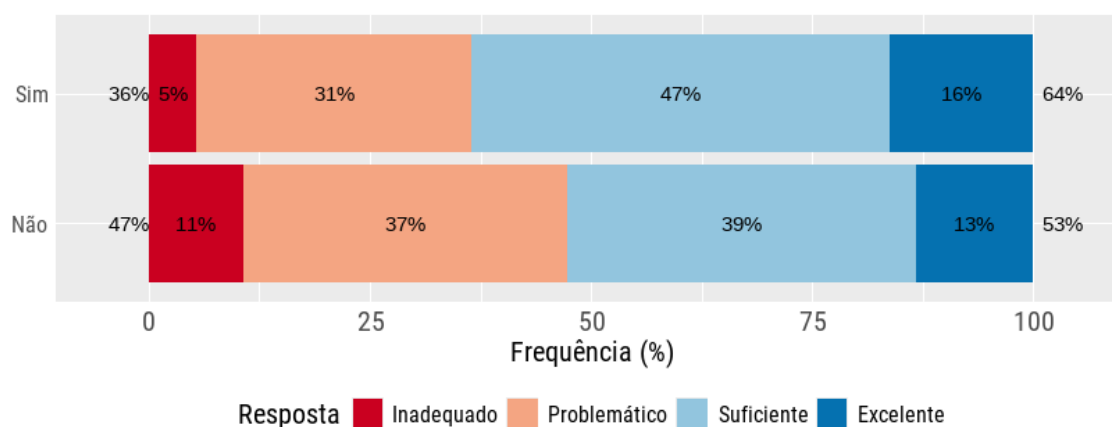


Figura B.20 Distribuição do Letramento global por Q6.2

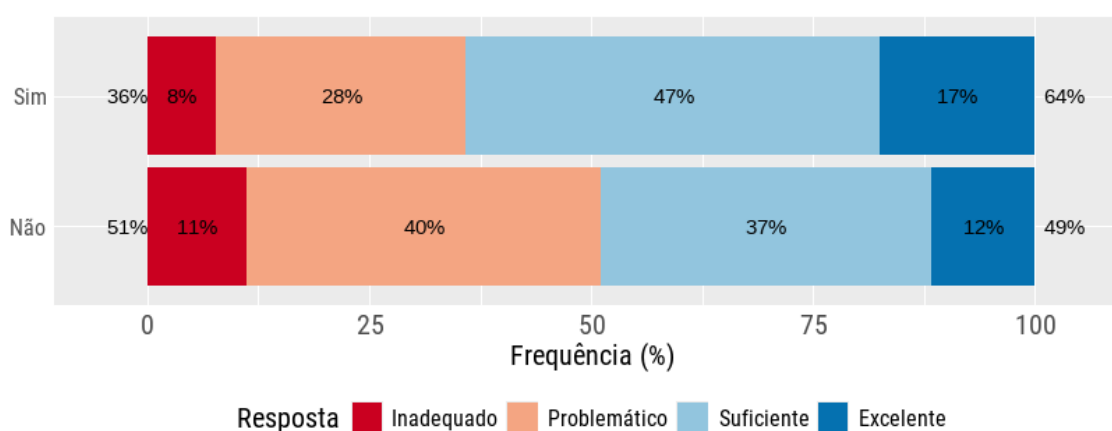


Figura B.21 Distribuição do Letramento global por Q9.2

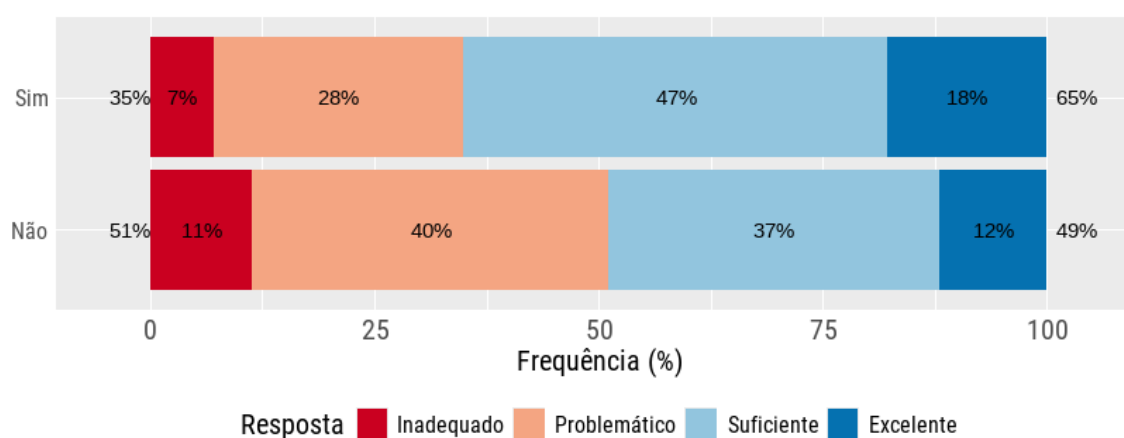


Figura B.22 Distribuição do Letramento global por Q9.3

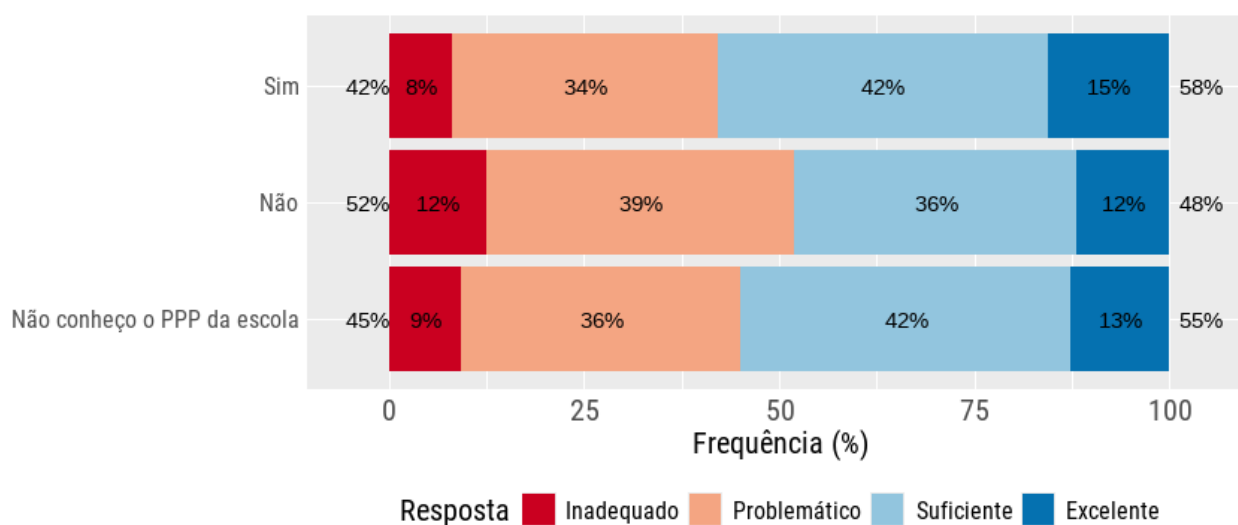


Figura B.23 Distribuição do Letramento global por Q6.3

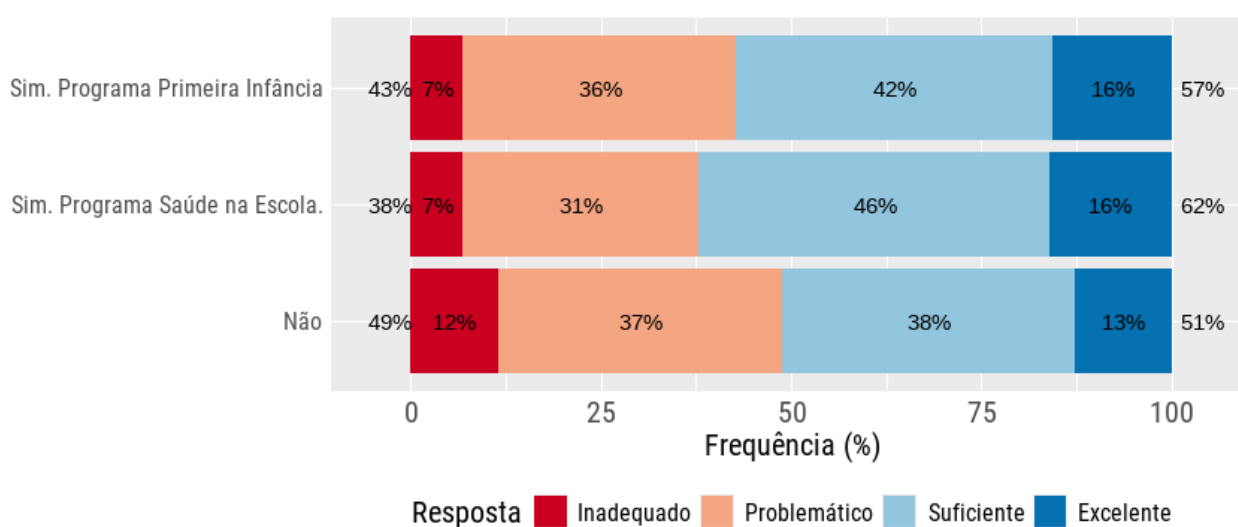


Figura B.24 Distribuição do Letramento global por Q6.4

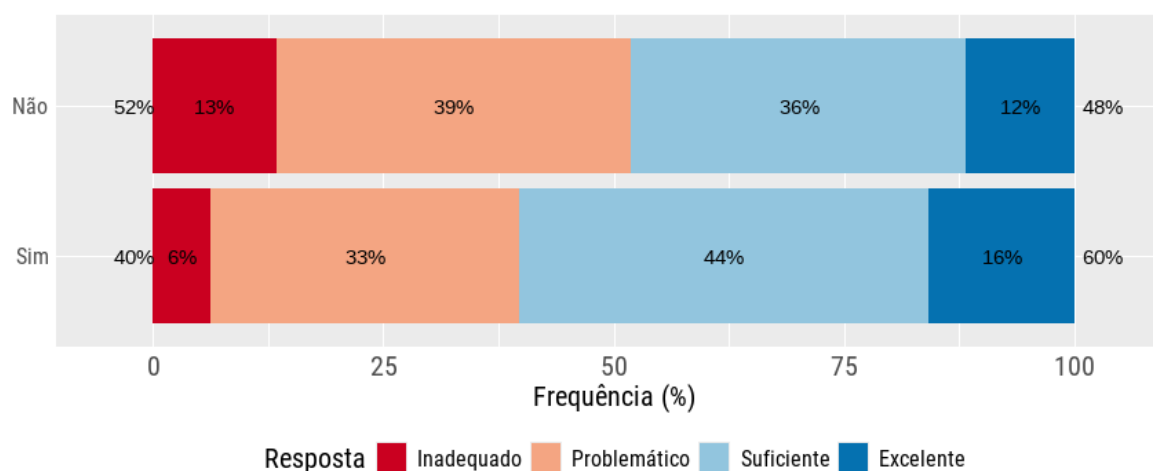


Figura B.25 Distribuição do Letramento global por Q7.1

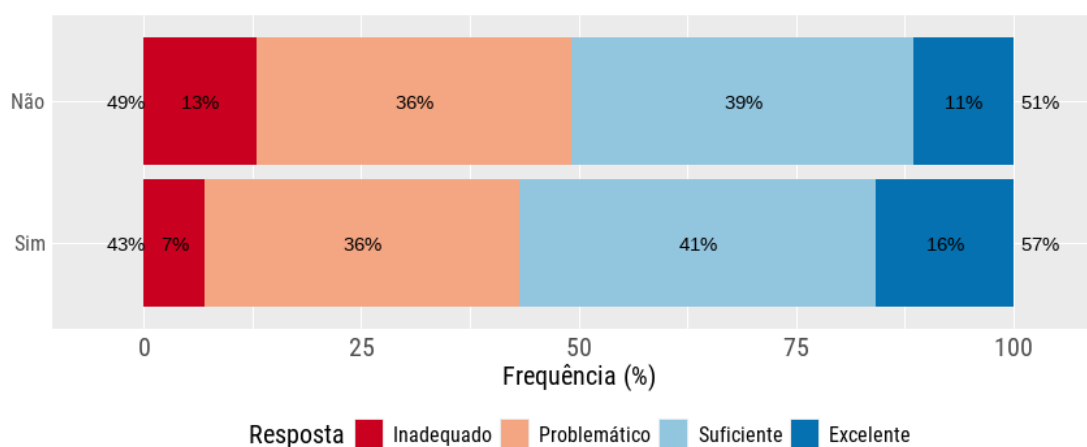


Figura B.26 Distribuição do Letramento global por Q7.2

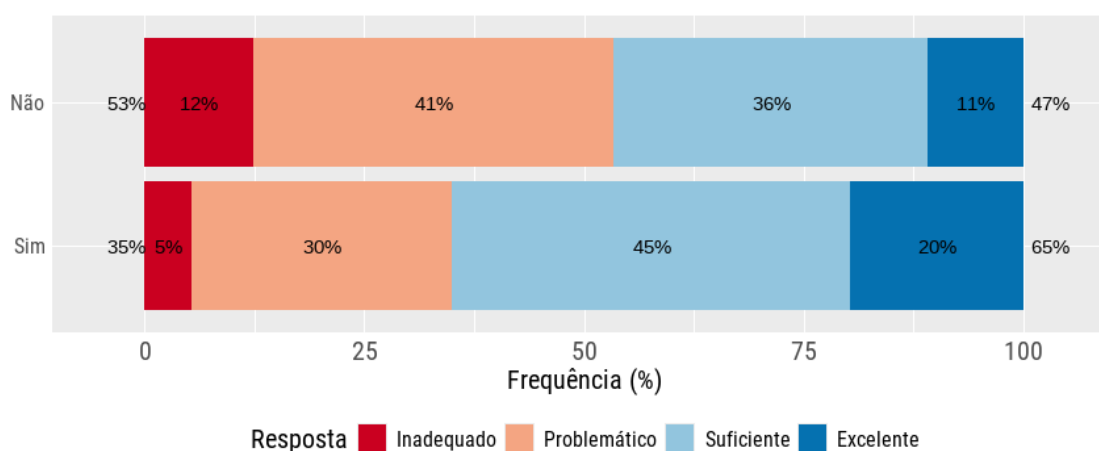


Figura B.27 Distribuição do Letramento global por Q8.1

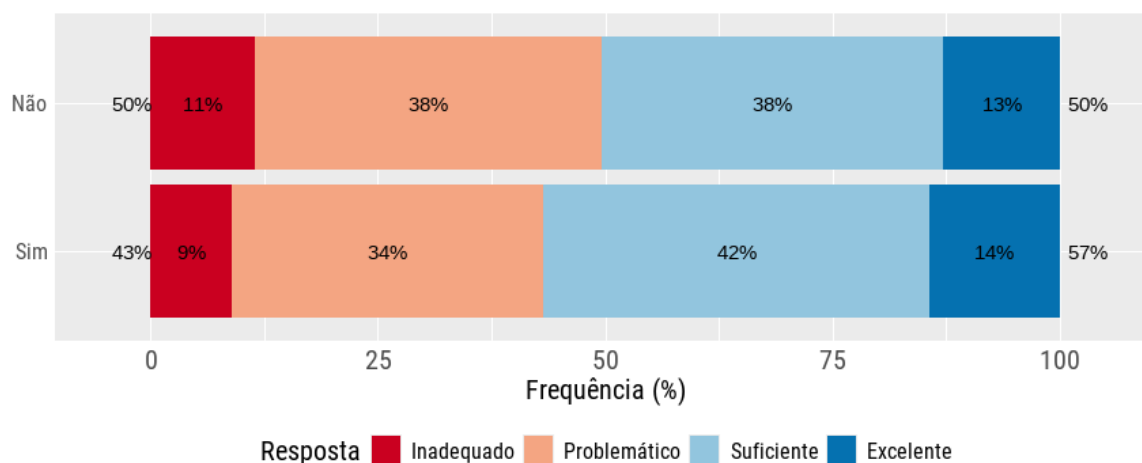


Figura B.28 Distribuição do Letramento global por Q8.2

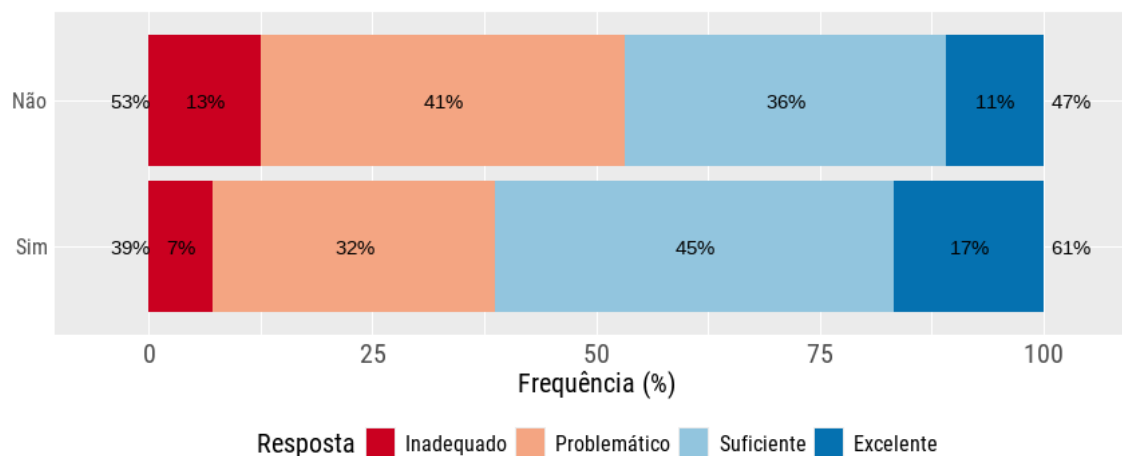


Figura B.29 Distribuição do Letramento global por Q8.3

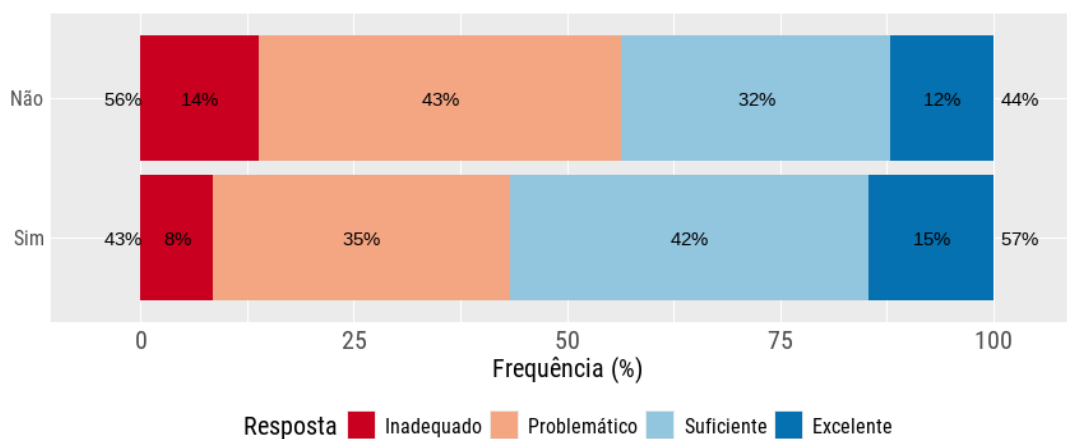


Figura B.30 Distribuição do Letramento global por Q8.4

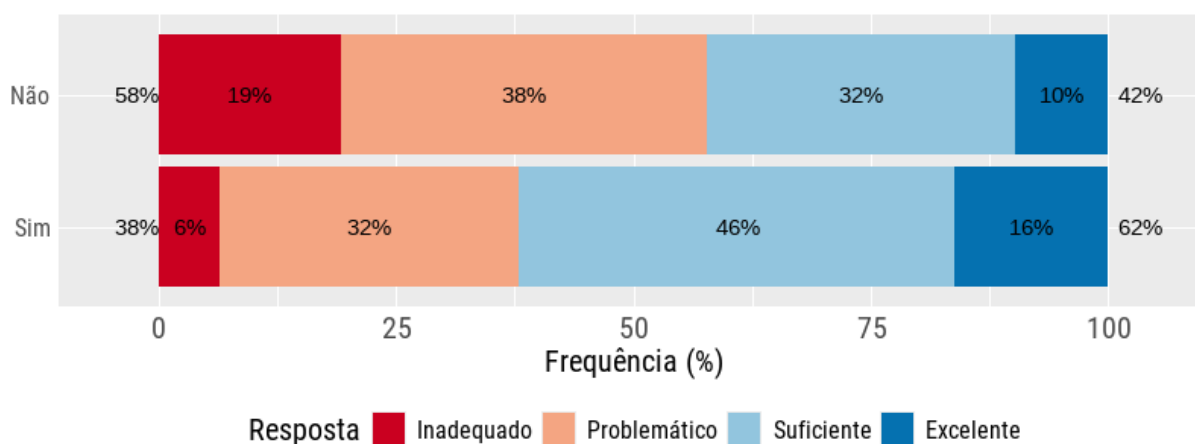


Figura B.31 Distribuição do Letramento global por Q9.1

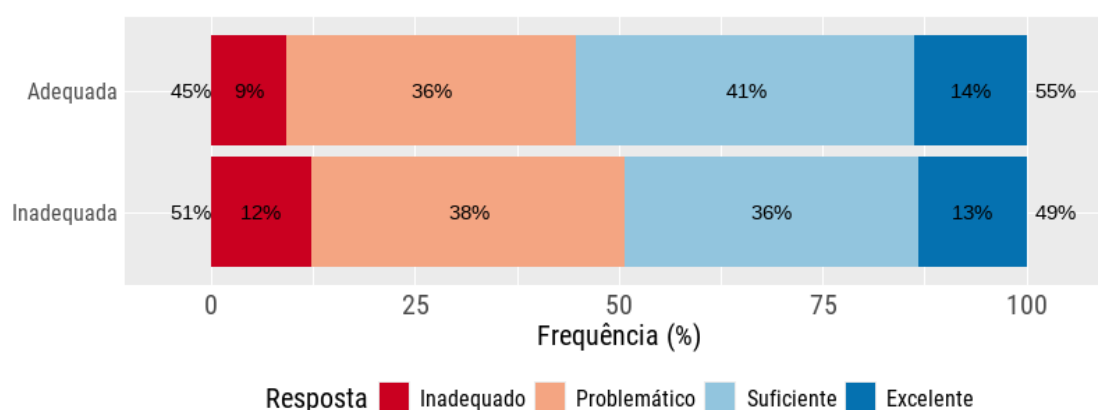


Figura B.32 Distribuição do Letramento global por Q9.4

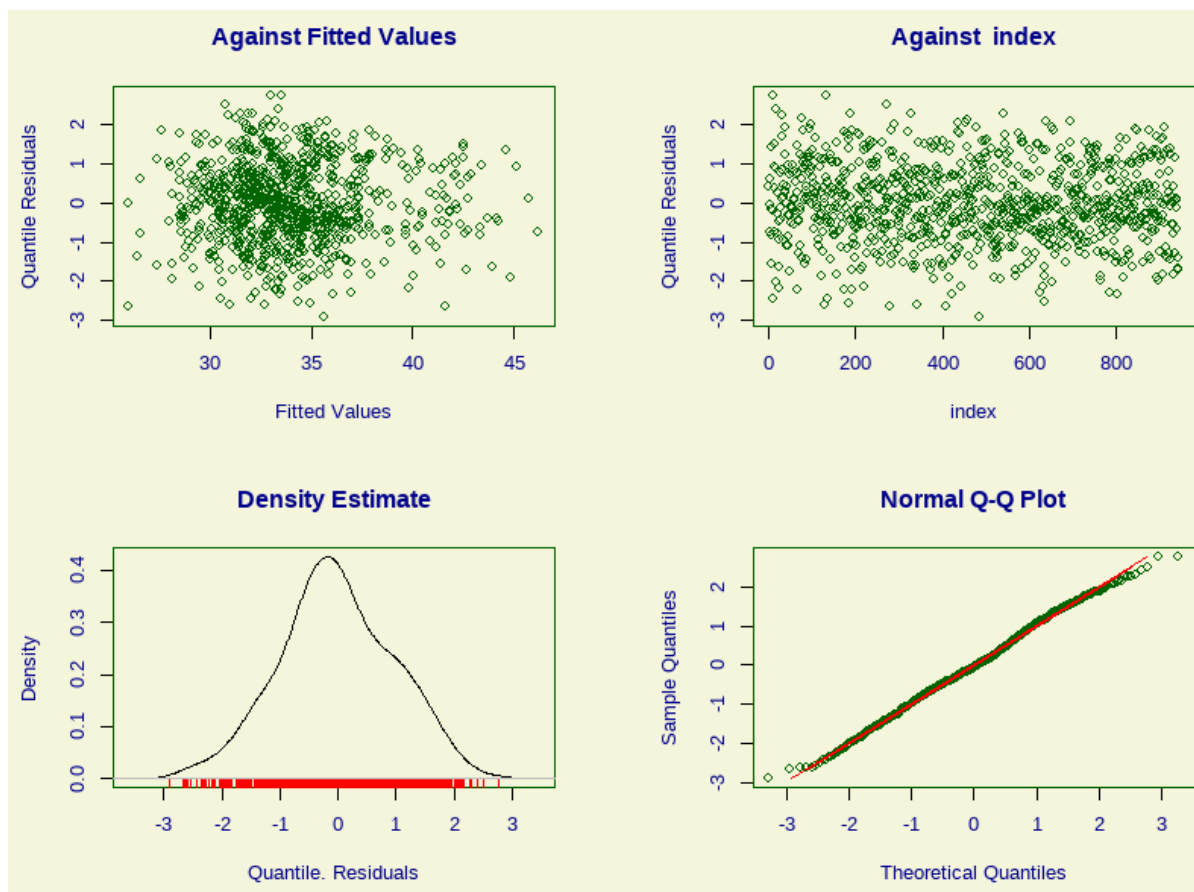


Figura B.33 Gráficos dos resíduos quantílicos normalizados. No canto superior esquerdo (Figura B.23A), é apresentado o gráfico dos valores ajustados pelos resíduos quantílicos; no canto superior direito (Figura B.23B), é apresentado o gráfico dos índices dos resíduos pelos resíduos quantílicos; no canto inferior esquerdo (Figura B.23C), é apresentada a densidade estimada dos resíduos quantílicos; no canto inferior direito (Figura B.23D), é apresentado o gráfico *QQ-Plot* dos resíduos quantílicos pelos resíduos normais teóricos.

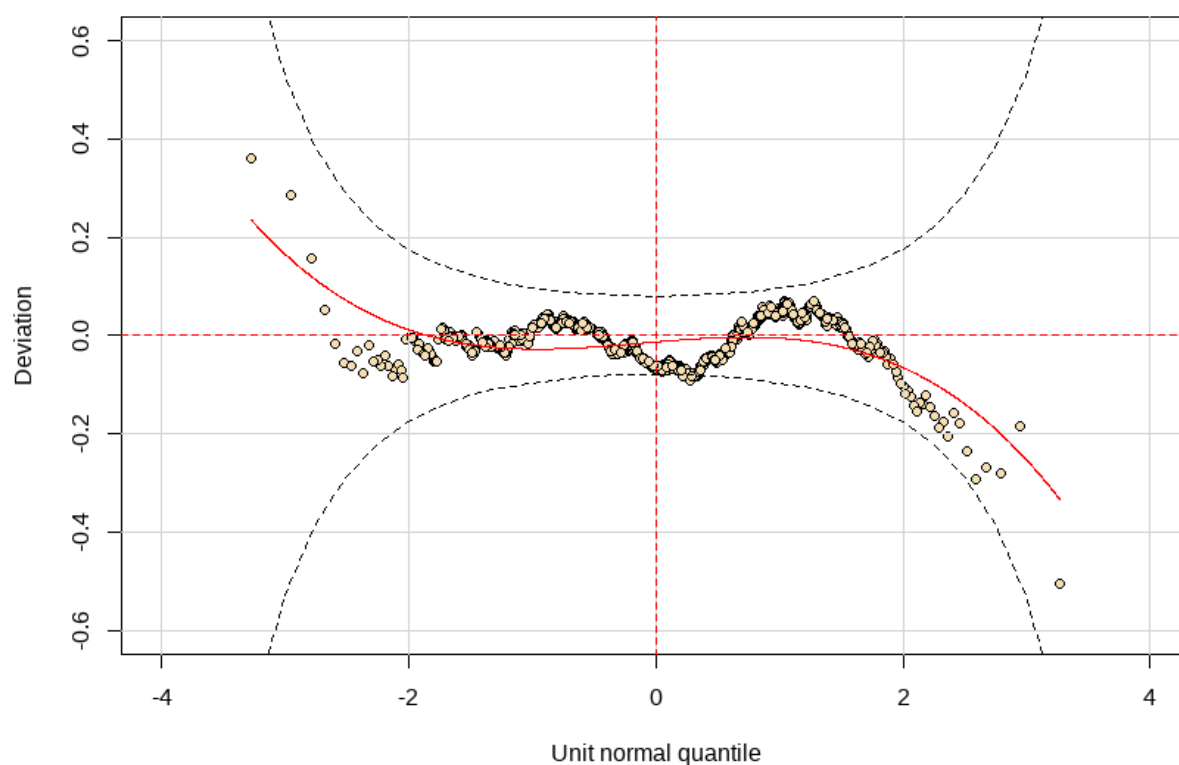


Figura B.34 Worm plot do ajuste do modelo Box-Cox t

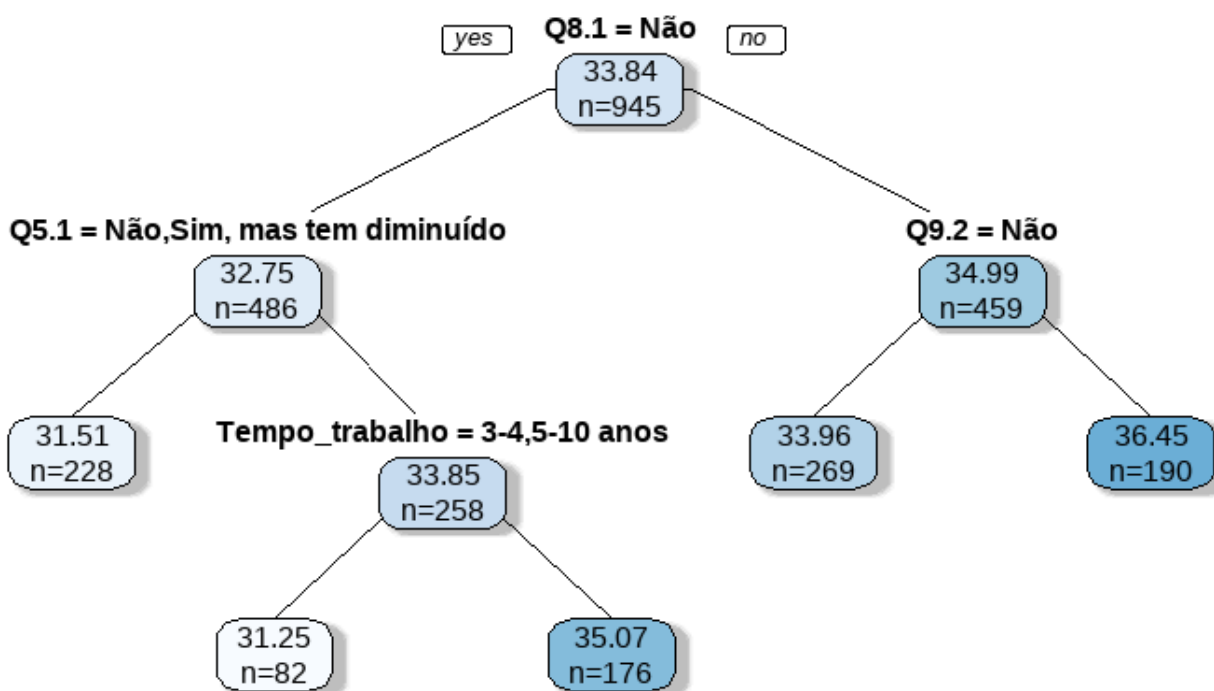


Figura B.35 Árvore de regressão para o Escore global

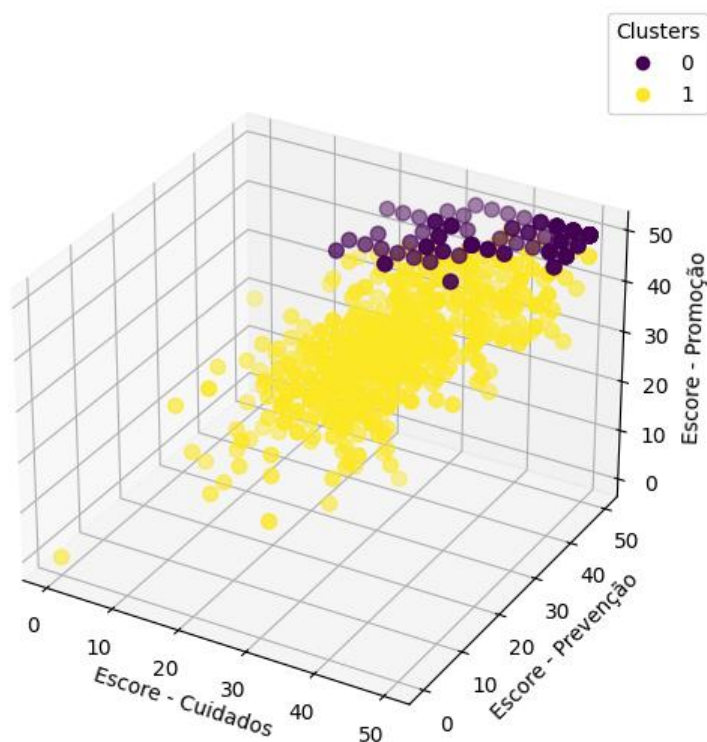


Figura B.36 Representação tridimensional dos clusters do Escore para os três domínios do questionário HLS-Q16

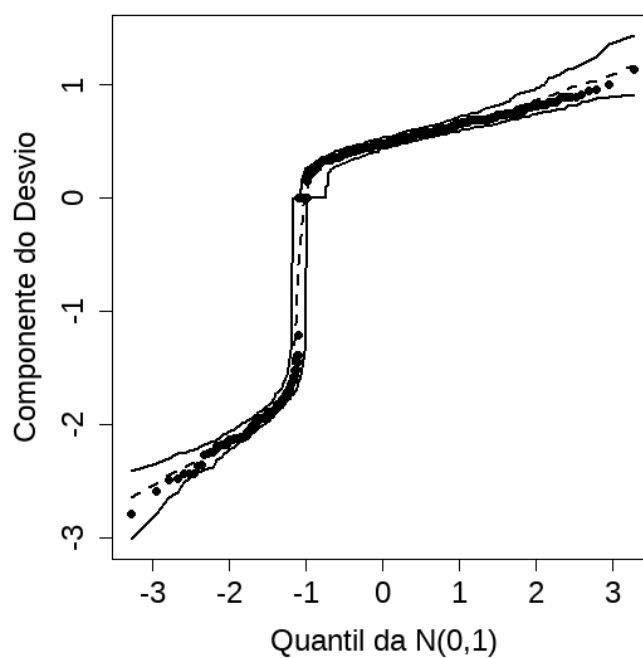


Figura B.37 Gráfico normal de probabilidades para o resíduo componente do desvio referente ao modelo logístico ajustado aos dados

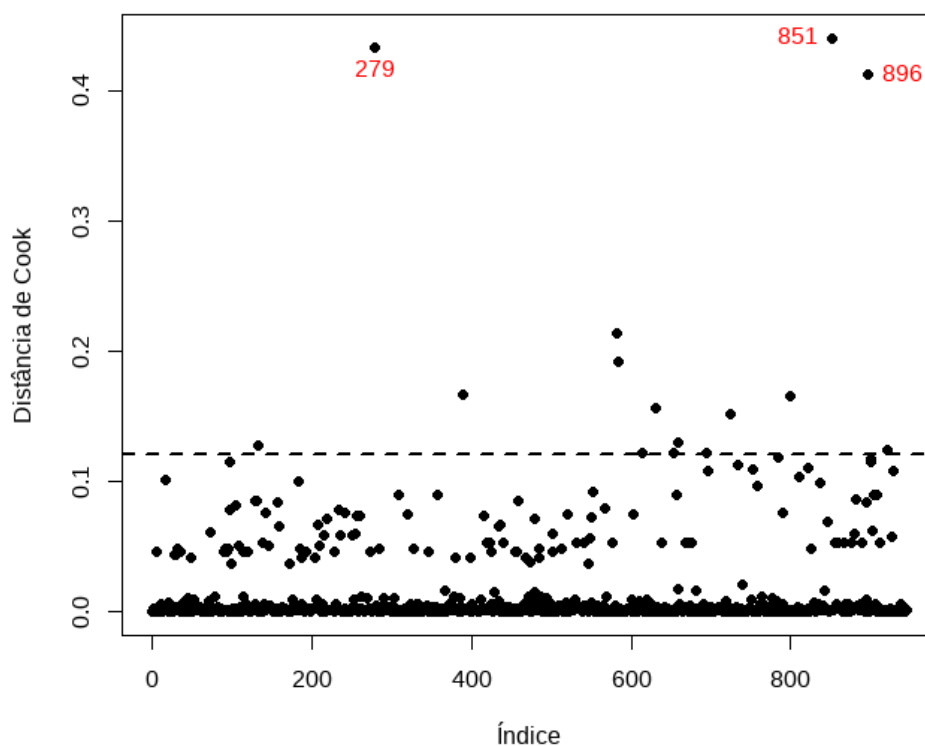


Figura B.38 Distância de Cook referente ao modelo logístico ajustado aos dados

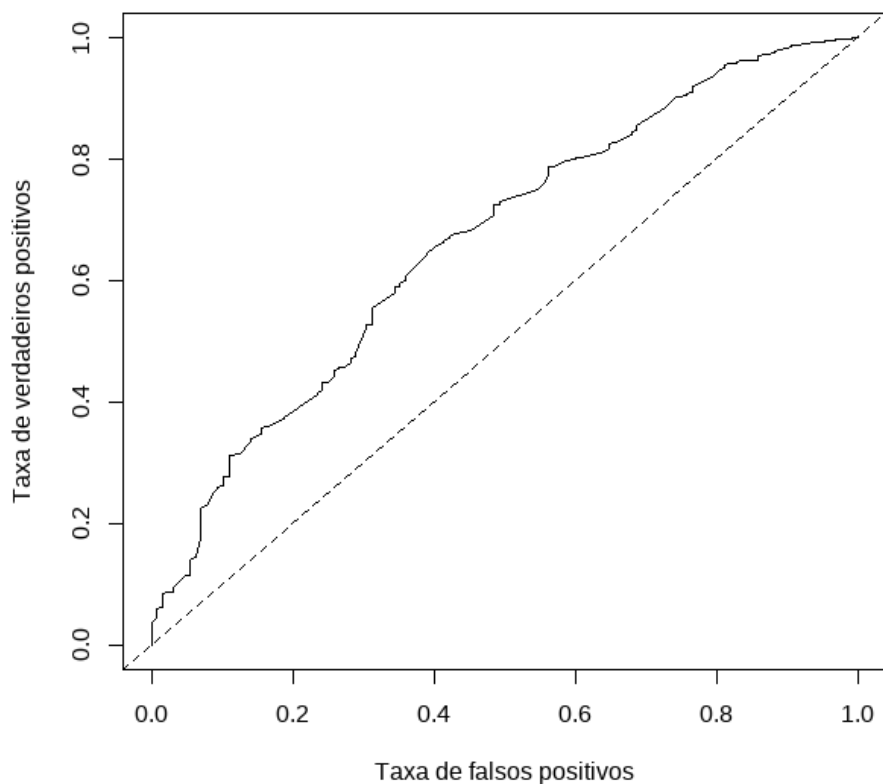


Figura B.39 Curva ROC do modelo logístico ajustado aos dados

APÊNDICE C

Metodologia

Apêndice C.1 Distribuição Box-Cox t

Seja Y uma variável aleatória contínua e positiva. A distribuição Box-Cox t (BCT) (Rigby e Stasinopoulos, 2006) é definida assumindo que:

$$Z = Z(Y; \mu, \sigma, \lambda) = \begin{cases} \frac{1}{\sigma\lambda} \left[\left(\frac{Y}{\mu} \right)^\lambda - 1 \right], & \text{se } \lambda \neq 0, \\ \frac{1}{\sigma}, & \text{se } \lambda = 0, \end{cases}$$

em que $\mu > 0, \sigma > 0, -\infty < \lambda < \infty$, tem uma distribuição t-Student padrão com τ graus de liberdade ($\tau > 0$) truncada em $A(\sigma, \lambda)^c = \mathbb{R} \setminus A(\sigma, \lambda)$, com $A(\sigma, \lambda) = \{z \in \mathbb{R} \mid \lambda z > -1/\sigma\}$ se $\lambda \neq 0$ e $A(\sigma, \lambda) = \mathbb{R}$, se $\lambda = 0$. Então, diz-se que Y tem distribuição BCT com parâmetros μ, σ, λ e τ e se denota $Y \sim BCT(\mu, \sigma, \lambda, \tau)$. Uma representação de $Y \sim BCT(\mu, \sigma, \lambda, \tau)$ é:

$$Y = \begin{cases} \mu(\sigma\lambda Z + 1)^{1/\lambda}, & \text{se } \lambda \neq 0, \\ \mu e^{\sigma Z}, & \text{se } \lambda = 0. \end{cases}$$

A função densidade de probabilidade de Y é dada por:

$$f_Y(y) = \frac{y^{\lambda-1}}{\mu^\lambda \sigma} \frac{f_T(z)}{F_T\left(\frac{1}{\sigma|\lambda|}\right)},$$

em que $z = Z(y; \mu, \sigma, \lambda)$ com $f_T(\cdot)$ e $F_T(\cdot)$ representando, respectivamente, a função densidade de probabilidade e a função de distribuição acumulada de uma distribuição t-Student padrão com τ graus de liberdade.

Uma característica bem interessante da distribuição Box-Cox t é a interpretabilidade dos seus parâmetros. Por exemplo, pode-se verificar que, para $\lambda = 0$, a mediana de Y é o próprio parâmetro μ . Rigby e Stasinopoulos (2006) mostram que μ é proporcional a todos os quantis da distribuição $BCT(\mu, \sigma, \lambda, \tau)$ quando os demais parâmetros são fixos, enquanto σ pode ser interpretado como a dispersão relativa. Além disso, λ atua como um parâmetro de assimetria, dado seu caráter de transformação, e τ controla o peso da cauda à direita da distribuição.

Apêndice C.2 Árvores de regressão

Uma árvore de regressão é uma representação gráfica de ramificações da teoria de decisão, em que a partir de um modelo base é gerada uma predição dos valores da variável resposta e se utiliza de uma técnica recursiva que particiona os dados e cria um modelo.

A utilização da árvore para prever uma nova observação é feita do seguinte modo: começando pelo topo, verificamos se a condição descrita no topo (primeiro nó) é satisfeita. Caso seja, seguimos à esquerda. Caso contrário, seguimos à direita. Assim prosseguimos até atingir uma folha.

Formalmente, uma árvore cria uma partição do espaço das covariáveis em regiões distintas e disjuntas: R_1, \dots, R_j . A predição para a resposta Y de uma observação com covariáveis x que estão em R_k é dada por

$$g(x) = \frac{1}{|\{i: x_i \in R_k\}|} \sum_{i: x_i \in R_k} y_i$$

Então, para prever o valor da resposta de x , observamos a região a qual a observação x pertence e, então, calculamos a média dos valores da variável resposta das amostras do conjunto de treinamento pertencentes àquela mesma região.

A criação da estrutura de uma árvore de regressão é feita através de duas grandes etapas: (i) a criação de uma árvore completa e complexa e (ii) a poda dessa árvore, com a finalidade de evitar o sobreajuste.

No passo (i), avalia-se o quão razoável uma dada árvore T é através de seu erro quadrático médio, em que y_R é o valor predito para a resposta de uma observação pertencente à região R . Utiliza-se uma heurística para encontrar uma árvore com erro quadrático médio baixo que consiste na criação de divisões binárias recursivas. Inicialmente, o algoritmo particiona o espaço de covariáveis em duas regiões distintas. Para escolher essa partição, busca-se, dentre todas as covariáveis x_i e cortes t_1 , aquela combinação que leva a uma partição (R_1, R_2) com predições de menor erro quadrático:

$$\sum_{i: x_i \in R_1}^n (y_i - \widehat{y}_{R_1})^2 + \sum_{i: x_i \in R_2}^n (y_i - \widehat{y}_{R_2})^2,$$

em que y_{R_k} é a predição fornecida para a região R_k . Assim, define-se

$$R_1 = \{\mathbf{x}: x_i < t_1\} \text{ e } R_2 = \{\mathbf{x}: x_i \geq t_1\}$$

em que x_i é a variável escolhida e t_1 é o corte definido.

Uma vez estabelecidas tais regiões, o nó inicial da árvore é então fixado. No próximo passo busca-se particionar R_1 ou R_2 em regiões menores. Para escolher a nova divisão, a mesma estratégia é utilizada: busca-se, dentre todas as covariáveis x_i e cortes t_2 , aquela combinação que leva a uma partição com menor erro quadrático. Note que agora também é necessário escolher qual região deve ser particionada: R_1 ou R_2 . Assuma que R_1 foi a região escolhida, juntamente com a covariável x_j e o corte t_2 . Chamemos a partição de R_1 de $R_{1,1}$, $R_{1,2}$. Assim,

$$R_{1,1} = \{\mathbf{x}: x_i < t_1, x_j < t_2\}, R_{1,2} = \{\mathbf{x}: x_i < t_1, x_j \geq t_2\} \text{ e } R_2 = \{\mathbf{x}: x_i \geq t_1\}.$$

O procedimento continua recursivamente, até que cheguemos a uma árvore com poucas observações em cada uma das folhas.

A árvore criada utilizando-se esse processo produz bons resultados para o conjunto de treinamento, mas é muito provável que ocorra o sobreajuste (*overfitting*). Isso gera uma performance preditiva ruim em novas observações. Assim, prossegue-se para o passo (ii), que é chamado de poda. Nessa etapa do processo cada nó é retirado, um por vez, e observa-se como o erro de predição varia no conjunto de validação. Com base nisso, decide-se quais nós permanecerão na árvore.

ANEXO

Anexo 1: Delineamento amostral

Estrato	Número de Unidades de Ensino	Quantidade de Professores	Quantidade de Gestores	Tamanho do estrato	Tamanho amostral
1	48	1411	174	1585	24
2	34	1777	136	1913	30
BUTANTÃ	82	3188	310	3498	54
3	120	4529	419	4948	77
4	70	3996	280	4276	66
CAMPO LIMPO	190	8525	699	9224	143
5	73	2594	260	2854	44
6	38	2035	152	2187	34
CAPELA DO SOCORRO	111	4629	412	5041	78
7	54	1819	197	2016	31
8	37	2092	148	2240	35
FREGUESIA DO Ó / BRASILÂNDIA	91	3911	345	4256	66
9	66	2119	235	2354	36
10	36	2147	144	2291	36
GUAINASES	102	4266	379	4645	72
11	75	2223	276	2499	39
12	36	1844	144	1988	31
IPIRANGA	111	4067	420	4487	70
13	71	2285	258	2543	39
14	30	1655	120	1775	28
ITAQUERA	101	3940	378	4318	67
15	63	2175	227	2402	37
16	38	2131	152	2283	36
JAÇANÃ / TREMEMBÉ	101	4306	379	4685	73
17	84	2685	300	2985	47
18	38	2216	152	2368	36
PENHA	122	4901	452	5353	83
19	79	2460	287	2747	42
20	60	3277	240	3517	55
PIRITUBA	139	5737	527	6264	97
21	51	1726	186	1912	30
22	36	1927	144	2071	32
SANTO AMARO	87	3653	330	3983	62
23	82	2655	294	2949	46
24	52	2790	208	2998	46
SÃO MATEUS	134	5445	502	5947	92
25	84	2818	295	3113	48
26	52	3079	208	3287	51
SÃO MIGUEL PAULISTA	136	5897	503	6400	99
TOTAL	1507	62465	5636	68101	1056

Anexo 2: Healthy Literacy Survey na versão em português (HLS-Q16)

Em uma escala que vai de “muito fácil” a “muito difícil”, com que facilidade você consegue:						
		Muito fácil	Fácil	Difícil	Muito difícil	Não sei (espontâneo)*
1encontrar informações sobre os tratamentos de doenças que preocupam você?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	...descobrir onde conseguir ajuda profissional quando está doente? (p.ex., um médico, farmacêutico, psicólogo)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	...entender o que o seu médico diz a você?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	... entender as instruções do seu médico ou farmacêutico sobre como tomar um medicamento (remédio) que foi receitado para você?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	...avaliar quando você precisa de uma segunda opinião de outro médico?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	...usar as informações que seu médico passa a você para tomar decisões sobre a sua doença?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7	... seguir as instruções do seu médico ou farmacêutico?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8encontrar informações sobre como lidar com problemas de saúde mental, como o estresse ou depressão?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9	...entender os avisos sobre comportamentos que prejudicam a saúde, tais como fumar, praticar pouca atividade física e o consumo excessivo de álcool?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10entender por que você precisa fazer exames periódicos de saúde? (p.ex., exame de mamas, teste de açúcar no sangue, pressão arterial)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11	... avaliar se as informações sobre os riscos à saúde disponíveis nos meios de comunicação são confiáveis? (p.ex., TV, internet ou outros meios de comunicação).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12	...decidir como você pode se proteger de uma doença com base nas informações dos meios de comunicação? (p.ex., jornais, internet ou outros meios de comunicação)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13	... encontrar informações sobre as atividades que são boas para o seu bem-estar mental? (p.ex., meditação, exercício, caminhada, pilates etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14	... entender os conselhos de saúde que os membros da sua família ou seus amigos dão a você?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15	... entender as informações disponíveis nos meios de comunicação sobre como ficar mais saudável? (p.ex., internet, jornais, revistas)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16	... avaliar que comportamentos do seu dia-a-dia afetam a sua saúde? (p.ex., consumo de bebidas e hábitos alimentares, exercícios etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

*Essa questão é utilizada somente quando o instrumento é aplicado no formato de entrevista.

Anexo 3: Questionário sociodemográfico

I. Questionário sociodemográfico
1. Características do entrevistado
1.1. Nome:
1.2. Idade (assinalar faixa etária)
1.3. Identidade de gênero. Selecione a opção que melhor descreve você: () Feminino () Masculino () Não binário
1.4. Você se declara uma pessoa: () Amarela () Branca () Indígena () Parda () Preta
2. Trabalho
2.1. Cargo/ função: () Assistente de diretor de escola () Coordenador pedagógico () Diretor de escola () Professor de educação infantil () Professor de educação infantil e ensino fundamental I () Professor de ensino fundamental II e médio. 2.1.a. Disciplina: _____
2.2. Tempo de trabalho na rede pública municipal de São Paulo: () Menos de 3 anos () De 3 a 4 anos () De 5 a 10 anos () 11 a 15 anos () 16 a 20 anos () 21 a 25 anos () Mais de 25 anos
2.2.a. Atualmente, em qual unidade educativa trabalha:
3. Escolaridade (último nível completo)
3.1. Nível de escolaridade () Ensino médio (Magistério) () Graduação () Pós-graduação (Lato Sensu / Especialização) () Mestrado () Doutorado () Pós-doutorado
3.2. Tipo de escola frequentada na educação básica: () Pública () Privada () Pública e privada
3.3. Tipo de escola frequentada na graduação: () Pública () Privada () Pública e privada
4. Renda
4.1. Renda pessoal (líquida) () 2 a 3 salários mínimos () 4 a 5 salários mínimos () 6 a 7 salários mínimos () 8 a 9 salários mínimos () 10 salários mínimos ou mais

Anexo 4: Questionário de autoeficácia do professor

II. Autoeficácia do professor _ Educação para a Saúde
<p>Caso necessite, para responder as perguntas, considere os seguintes conceitos, definidos pela Organização Mundial de Saúde (2021):</p> <p>- Letramento em Saúde: O LS pode ser definido como o conhecimento e as competências que determinam a motivação e a capacidade dos indivíduos em terem acesso, compreenderem e usarem informação de modo a promover e a manter a saúde, visando à melhor qualidade de vida.</p> <p>- Educação para a saúde: qualquer combinação de experiências de aprendizagem projetadas para ajudar indivíduos e comunidades a melhorar sua saúde, aumentando o conhecimento, influenciando a motivação e melhorando o letramento em saúde.</p>
5.Satisfação com o seu trabalho
<p>5.1.Você está satisfeito(a) com sua contribuição na escola?</p> <p>() Sim, frequentemente.</p> <p>() Sim, mas estou sentindo que minha satisfação tem diminuído.</p> <p>() Não.</p>
<p>5.2.Se pudesse, escolheria novamente ser professor(a)?</p> <p>() Sim. () Não.</p> <p>Se você respondeu não, relacione o motivo:</p>
6. Currículo
<p>6.1.No seu curso de graduação (inicial) você recebeu formação sobre Educação para a Saúde (Letramento em Saúde)?</p> <p>() Sim. () Não.</p> <p>Se você respondeu sim, relacione:</p> <p>A. O(s) nome(s) da(s) disciplina(s):</p> <p>B. O(s) tema(s) trabalhado (s):</p>
<p>6.2.Você já participou de formação continuada sobre Educação para a Saúde?</p> <p>() Sim. () Não.</p> <p>Se você respondeu sim, relacione:</p> <p>A. A (s) instituição(s) que ofereceu o curso:</p> <p>B. O (s) tema(s) trabalhado (s):</p>
<p>6.3.No Projeto Político Pedagógico (PPP) da escola em que trabalha consta temas sobre Educação para a Saúde ?</p> <p>() Sim. () Não. () Não conheço o PPP da escola em que trabalho.</p>
<p>6.4.A unidade educativa em que você trabalha participa do Programa Saúde na Escola (Ensino fundamental) ou Programa Primeira Infância (Educação Infantil)?</p> <p>() Sim. () Não. () Não sei responder.</p>

Anexo 4 (Continuação): Questionário de autoeficácia do professor

7.Planejamento e planos
7.1.No período de planejamento são discutidas na escola em que trabalha, questões sobre Educação para a Saúde, para constar nos planos?
() Sim. () Não.
7.2.No seu plano de ensino constam temas voltados à Educação para a Saúde? () Sim. () Não. Se você respondeu sim, relacione: O(s) tema(s) que consta (m):
8. Construção de conhecimentos (na aula)
8.1.Você considera que seus alunos possuem uma boa Educação para a Saúde ? () Sim. () Não. () Não sei responder.
8.2.Você realiza atividades que envolvam a Educação para a Saúde? () Sim. () Não. Se você respondeu sim, relacione tipo(s) de atividade(s):
8.3.Você considera que se apropriou dos conteúdos específicos e pedagógicos necessários para exercer a docência em Educação para a Saúde, ou seja, se sentiu seguro (a) ao ministrar os conteúdos sobre Educação para a Saúde nas suas aulas? () Sim. () Não. Se você respondeu não, relacione o(s) motivo(s):
8.4.Seu aluno é capaz de comunicar-se com um adulto, cuidador ou profissional de saúde sobre comportamentos de saúde? (Por exemplo, sobre escovação de dentes, atividade física, alimentação saudável) () Sim. () Não. () Não sei responder.
8.5.O que falta para você trabalhar melhor (se já trabalha) ou iniciar o trabalho com Educação para a Saúde, nas suas aulas? (Pode assinalar mais de uma alternativa) () Formação (cursos) sobre conteúdos e metodologias para professores e gestores. () Materiais didáticos (livros, apostilas etc.) para professores e alunos. () Currículos e programas indicando conteúdos a serem trabalhados. () Não falta nada, não trabalho com Educação para a Saúde porque não disponho de tempo ou não considero necessário. () Não falta nada e com os recursos de que disponho já trabalho com Educação para a Saúde.
9. Na unidade educativa em que trabalha
9.1.Os familiares / responsáveis têm a oportunidade de conversar com professores e equipe gestora da escola sobre Educação para a Saúde? () Sim. () Não. () Não sei responder. Se você respondeu sim, relacione os momentos:
9.2.Os professores e equipe gestora recebem formação sobre notícias/informações de saúde veiculadas nas redes sociais (distinguir notícias verdadeiras das falsas)? () Sim. () Não. Se você respondeu sim, relacione ações:

Anexo 4 (Continuação): Questionário de autoeficácia do professor

9.3.Os professores recebem formação sobre como encontrar um equilíbrio entre cumprir o papel profissional como professor e seus requisitos (ensinar uma classe
numerosa) e prestar atenção aos indivíduos e responder às suas necessidades de saúde? () Sim. () Não. Se você respondeu sim, relacione ações:
9.4.No geral, a infraestrutura de seu local de trabalho, é: () Adequada. () Inadequada. Caso considere necessário, justifique:
III. Educação para a Saúde no Currículo da Cidade
10.Indique os temas relacionados a Educação para Saúde que você normalmente trabalha, tendo como referência o Currículo da Cidade. (Se necessário, use o verso da folha).

Fonte: elaborado pela autora.