



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS DE FLORIANÓPOLIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MÉTODOS E GESTÃO EM AVALIAÇÃO

André Felipe Zilio da Silva

**Fatores Associados à Proficiência Abaixo do Básico em Língua Portuguesa no Saeb
2017: Uma Análise de Regressão Hierárquica**

Florianópolis

2020

André Felipe Zilio da Silva

**Fatores Associados à Proficiência Abaixo do Básico em Língua Portuguesa no Saeb
2017: Uma Análise de Regressão Hierárquica**

Dissertação submetida ao Programa de Pós-Graduação
em Métodos e Gestão em Avaliação da Universidade
Federal de Santa Catarina para a obtenção do título de
Mestre em Métodos e Gestão em Avaliação.
Orientador: Prof. Luiz Ricardo Nakamura, Dr.
Coorientador: Prof. Marcelo Menezes Reis, Dr.

Florianópolis

2020

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

DA SILVA, ANDRÉ FELIPE ZILIO
Fatores Associados à Proficiência Abaixo do Básico em
Língua Portuguesa no Saeb 2017: Uma Análise de Regressão
Hierárquica Florianópolis 2020 / ANDRÉ FELIPE ZILIO DA SILVA
; orientador, Luiz Ricardo Nakamura, coorientador,
Marcelo Menezes Reis, 2020.
114 p.

Dissertação (mestrado profissional) - Universidade
Federal de Santa Catarina, Centro Tecnológico, Programa de
Pós-Graduação em , Florianópolis, 2020.

Inclui referências.

1. . 2. Desempenho escolar. 3. Regressão multinível. 4.
Avaliação em larga escala. I. Nakamura, Luiz Ricardo . II.
Reis, Marcelo Menezes . III. Universidade Federal de Santa
Catarina. Programa de Pós-Graduação em . IV. Título.

André Felipe Zilio da Silva

Fatores Associados à Proficiência Abaixo do Básico em Língua Portuguesa no Saeb 2017: Uma Análise de Regressão Hierárquica

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof.^a Andreia Zanella, Dr.^a

Universidade Federal de Santa Catarina - UFSC

Prof. Thiago Gentil Ramires, Dr.

Universidade Tecnológica Federal do Paraná – UTFPR

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Métodos e Gestão em Avaliação.

Prof. Marcelo Menezes Reis, Dr.

Coordenação do Programa de Pós-Graduação e Coorientador

Prof. Luiz Ricardo Nakamura, Dr.

Orientador

Florianópolis, 2020.

Este trabalho é dedicado aos meus avós (in memoriam).

AGRADECIMENTOS

Agradeço à minha família, Ana, Mélo, Eliane, Márcia, José, Bruna e Renan.

Agradeço ao meu orientador, prof. Dr. Luiz Ricardo Nakamura e ao meu coorientador, prof. Dr. Marcelo Reis.

Agradeço aos gestores da Secretaria Municipal da Educação de Curitiba que me permitiram realizar esta conquista, Maria Silvia Bacila, Oséias Santos de Oliveira, Marcia Rover, Elisabeth Dubas, Marcia Peça e Andressa Pereira.

Agradeço aos professores das disciplinas do PPMGA, Dalton, Adriano, Marcelo, Lizandra, Andreia, Silvana, Pedro, Dilvo e Silvia.

Agradeço aos colegas de turma, Adriana, Claudia, Diego, Dirce, Emiliana, Kauê, Leonardo, Leize, Maria Aparecida, Mariane, Nathalia, Paulo, Petterson, Regina, Renan, Romilson, Vivian, Tagma.

Agradeço à minha colega de formação e amiga, Alexandra Russi, por toda contribuição e apoio.

Agradeço a todos colegas de trabalho em especial, Eliane T., Andrea L., Jerusa, Juciele L., Terezinha O., Elisangela M., Sheila G., Juciara, Aline, Claudia H., Claudia M., Claudia A, Rosiane, Clarice M, Samara, Patrícia P., Giovanni, Daniele, Nanci, Maria, Luciana Z. pelo apoio e pela amizade.

Agradeço a todos os membros da banca.

Agradeço a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write. (WELLS, 1903)

RESUMO

Para identificar fatores associados à proficiência abaixo do básico em língua portuguesa de estudantes do 5º ano do ensino fundamental e participantes do Saeb 2017, foram pré-selecionadas 36 variáveis, sendo 28 relacionadas ao nível do estudante e 8 relacionadas ao nível da unidade escolar. Por conta do grande número de participantes na avaliação, optou-se em selecionar uma amostra de estudantes e de escolas para o estudo. A partir de um plano de amostragem, foram selecionados 38.964 estudantes de 1.769 unidades representativas. Um estudo descritivo com as variáveis pré-selecionadas foi realizado a fim de obter as primeiras características dos estudantes e das unidades selecionadas. Como a variável resposta em estudo é de natureza categórica (isto é, estudantes abaixo do básico compõem a categoria de interesse) e as variáveis explicativas são características associadas às escolas e aos estudantes, foi utilizado o modelo de regressão hierárquica logístico. Para selecionar as variáveis significativas para compor o modelo utilizou-se o método iterativo stepwise, por meio do critério de informação bayesiano (BIC). O modelo final selecionado foi composto por 21 variáveis explicativas, sendo que a qualidade do ajuste, baseado em uma análise de resíduos, foi satisfatória, não violando as pressuposições do modelo assumido. Os resultados apontam categorias que possuem mais chance de obter proficiência abaixo do básico em língua portuguesa em relação às demais categorias, como, por exemplo: estudantes do sexo masculino, estudantes com faixa etária abaixo da correta para cursar o 5º ano, estudantes que quase nunca ou nunca leem livros, quadrinhos, estudantes provenientes de escolas da zona rural.

Palavras-chave: Desempenho escolar. Regressão multinível. Avaliação em larga escala

ABSTRACT

In order to identify factors associated with students, from the 5th year of elementary school and Saeb 2017 participants, with proficiency below the basic in Portuguese, 36 variables were pre-selected, 28 of which were related to the level of the student and related to the level of the school unit. Due to the large number of participants in the evaluation, it was decided to select a sample of students and schools for the study. From a sampling plan, 38,964 students and 1,769 representative units were selected. A descriptive study with the pre-selected variables was carried out in order to obtain the first characteristics of the indicators and the selected units. As the response variable under study is categorical in nature (that is, students below the basic make up the category of interest) and as explanatory variables are characteristics associated with schools and students, the logistic hierarchical regression model was used. To select the important variables to compose the model, the interactive method was used, through the Bayesian information criterion (BIC). The selected model is composed of 21 explanatory variables and the quality of the adjustment, based on a residual analysis, was satisfactory, not violating the assumptions of the assumed model. The results point to categories that are more likely to obtain proficiency below the basics in Portuguese in relation to the other categories, such as: male students, students below the correct age group to attend the 5th year, students who almost never or they never read books, comics, students from rural schools.

Keywords: School performance. Multilevel regression. Large-scale evaluation

LISTA DE FIGURAS

Figura 1 - Nuvem de palavras-chave de trabalhos pesquisados para composição do referencial desta dissertação	27
Figura 2 - Fluxograma que descreve o método stepwise utilizando o critério de BIC	42
Figura 3 - Proporção de estudantes em cada uma das categorias de proficiências	44
Figura 4 - Distribuição das proficiências amostradas	45
Figura 5 - Variabilidade das proficiências de Língua Portuguesa – 5º ano por regiões do Brasil.....	46
Figura 6 - Distribuição dos turnos dos estudantes em cada uma das categorias	47
Figura 7 - Distribuição referente ao sexo dos estudantes em cada uma das categorias	48
Figura 8 - Distribuição da idade dos estudantes em cada uma das categorias	50
Figura 9 - Distribuição do estudante ter computador em cada uma das categorias ...	51
Figura 10 - Distribuição do estudante que residem com a mãe em cada uma das categorias	52
Figura 11 - Distribuição do nível de instrução da mãe em cada uma das categorias ..	53
Figura 12 - Distribuição do questionamento se a mãe sabe ler/escrever em cada uma das categorias.....	54
Figura 13 - Nível de instrução do pai em cada uma das categorias	55
Figura 14 - Distribuição do questionamento se o pai sabe ler/escrever em cada uma das categorias.....	56
Figura 15 - Distribuição do questionamento se os pais participam de reunião pedagógica em cada uma das categorias	57
Figura 16 - Distribuição do questionamento se os pais incentivam estudar em cada uma das categorias.....	58
Figura 17 - Distribuição do questionamento se os pais incentivam a realizar o dever em cada uma das categorias.....	59
Figura 18 - Distribuição do questionamento se os pais incentivam a ler em cada uma das categorias.....	60
Figura 19 - Distribuição do questionamento se os pais incentivam a não faltar as aulas em cada uma das categorias.....	61

Figura 20- Distribuição do questionamento se os pais conversam sobre as aulas em cada uma das categorias	62
Figura 21 - Distribuição do questionamento se lê jornais em cada uma das categorias	62
Figura 22 - Distribuição do questionamento se lê livros em cada uma das categorias	64
Figura 23 - Distribuição do questionamento se lê revistas em cada uma das categorias	65
Figura 24 - Distribuição do questionamento se lê quadrinhos em cada uma das categorias	66
Figura 25 - Distribuição do questionamento se o estudante lê notícias na internet em cada uma das categorias	67
Figura 26 - Distribuição do questionamento se o estudante frequenta biblioteca em cada uma das categorias	68
Figura 27 - Distribuição do questionamento se o estudante costuma frequentar espetáculos ou exposições em cada uma das categorias.....	69
Figura 28 - Distribuição do ingresso na escola em cada uma das categorias.....	70
Figura 29 - Distribuição do questionamento se o estudante já reprovou em cada uma das categorias.....	71
Figura 30 - Distribuição do questionamento se o estudante já abandonou em cada uma das categorias.....	72
Figura 31 - Distribuição do questionamento se o estudante faz o dever de caso em cada uma das categorias.....	73
Figura 32 - Distribuição do questionamento se o professor corrige o dever de casa de língua portuguesa em cada uma das categorias	74
Figura 33 - Distribuição do tipo de dependência proveniente do estudante em cada uma das categorias.....	75
Figura 34 - Distribuição da localização da unidade em cada uma das categorias.....	76
Figura 35 - Distribuição dos grupos do NSE em cada uma das categorias.....	77
Figura 36 - Distribuição do item possuir biblioteca em cada uma das categorias.	78
Figura 37 - Distribuição do item possui sala de leitura em cada uma das categorias.	79
Figura 38 -Distribuição do item possuir quadra esportiva em cada uma das categorias.	80

Figura 39 - Distribuição do item possuir laboratório de informática em cada uma das categorias.....	81
Figura 40 - Worm plot do modelo final ajustado	87

LISTA DE QUADROS

Quadro 1 - Caracterização dos constructos dos questionários contextuais Saeb, de estudantes, e quantidades de questões pertinentes.	23
Quadro 2 - Publicações selecionadas para análise.	26
Quadro 3 – Construtos de variáveis a partir da revisão dos autores.....	29
Quadro 4 - Variáveis explicativas – Nível do Estudante utilizadas no estudo.....	32
Quadro 5 – Critério para seleção das unidades (participação da amostragem).....	36
Quadro 6 - Interpretações da Razão de Chances - Variáveis do nível do estudante ..	87
Quadro 7 - Interpretações da Razão de Chances - Variáveis do nível da escola	91
Quadro 8 - Resultados para os aspectos pessoais do estudante.....	101
Quadro 9 - Resultados para os aspectos familiares do estudante	101
Quadro 10 - Resultados para os efeitos das características das escolas e composição da escola	101

LISTA DE TABELAS

Tabela 1 - Distribuição, em percentuais, de estudantes do 5º ano do Ensino Fundamental, por nível de proficiência qualitativo, considerando a componente curricular língua portuguesa nas edições 2013, 2015 e 2017 do Saeb.....	17
Tabela 2 - Seleção da amostra de unidades de acordo com a proporção na população	36
Tabela 3 - Medidas resumo das proficiências dos estudantes amostrados.....	45
Tabela 4 - Frequência Turno do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano	47
Tabela 5 - Sexo do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	48
Tabela 6 - Categorização cor/raça do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	49
Tabela 7 - Categorização idade do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	49
Tabela 8 - Categorização idade do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	50
Tabela 9 - Categorização morar com a mãe x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	51
Tabela 10 - Categorização nível de instrução da mãe x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	53
Tabela 11: Mãe sabe ler? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	54
Tabela 12 - Categorização nível de instrução do pai x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	55
Tabela 13 - Pai sabe ler? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	56
Tabela 14 - Participação dos pais em reuniões pedagógica x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	57
Tabela 15 - Os pais incentivam estudar x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	58

Tabela 16 - Os pais incentivam a realizar o dever x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	59
Tabela 17 - Os pais incentivam a ler x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	60
Tabela 18 - Os pais incentivam a não faltar as aulas x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	60
Tabela 19 - Os pais conversam sobre as aulas x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	61
Tabela 20 - Lê jornais? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	62
Tabela 21 - lê livros? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	63
Tabela 22 - lê revistas? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	64
Tabela 23 - lê quadrinhos? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	65
Tabela 24 - lê notícias na internet? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	66
Tabela 25 - frequenta biblioteca? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	67
Tabela 26 - Costuma ir a espetáculos ou exposições? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	68
Tabela 27 - Ingresso na escola x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	69
Tabela 28 - O estudante já reprovou? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	70
Tabela 29 - O estudante já abandonou? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	71
Tabela 30 - O estudante faz o dever de casa? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	72
Tabela 31 - O professor corrige o dever de casa de língua portuguesa? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	73
Tabela 32 - Tipo de dependência administrativa x Dicotomização das proficiências de Língua Portuguesa – 5º ano.	74

Tabela 33 - Área de localização da unidade x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	75
Tabela 34 - NSE x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	76
Tabela 35 - A unidade possui biblioteca? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	77
Tabela 36 - A unidade possui sala de leitura? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	78
Tabela 37 - A unidade possui quadra esportiva? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	79
Tabela 38 - A unidade possui laboratório de informática? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.....	80
Tabela 39 - Modelo Nulo	82
Tabela 40 - Seleção do melhor modelo através do método Stepwise	82
Tabela 41 - Modelo Selecionado: Estimativas, erro padrão, p-valor e razão de chances	84
Tabela 42 - Resumo dos quantis aleatórios residuais.....	86

LISTA DE ABREVIATURAS E SIGLAS

MPL	Modelo Logístico de Três Parâmetros
ANA	Avaliação Nacional da Alfabetização
Aneb	Avaliação Nacional da Educação Básica
Anresc	Avaliação Nacional do Rendimento Escolar
CCI	Curva Característica do Item
IDEB	Índice de Desenvolvimento da Educação Básica
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
LDB	Lei de Diretrizes e Bases da Educação Nacional
MEC	Ministério da Educação
ML3	Modelo Logístico Unidimensional de Três Parâmetros
MRH	Modelo Regressão Hierárquica
PNE	Plano Nacional de Educação
Saeb	Sistema de Avaliação da Educação Básica
TRI	Teoria de Resposta ao Item

SUMÁRIO

1	INTRODUÇÃO	15
1.1	CONTEXTUALIZAÇÃO DO PROBLEMA.....	15
1.2	OBJETIVOS	18
1.2.1	Objetivo Geral	18
1.2.2	Objetivos Específicos.....	19
1.3	JUSTIFICATIVA	19
1.4	ESTRUTURA DO TRABALHO	20
2	FUNDAMENTAÇÃO TEÓRICA.....	21
2.1	SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA	21
2.1.1	Histórico	21
2.1.2	Saeb - Edição 2017.....	22
2.1.3	Escala de proficiência e a Teoria da Resposta ao Item	24
2.2	EQUIDADE NA EDUCAÇÃO.....	24
2.3	PRINCIPAIS ESTUDOS REALIZADOS SOBRE O SAEB, FATORES ASSOCIADOS À PROFICIÊNCIA NA EDUCAÇÃO BÁSICA E REGRESSÃO HIERÁRQUICA.....	25
3	MATERIAL E MÉTODO	31
3.1	CARACTERIZAÇÃO DOs CONJUNTO DE DADOS	31
3.1.1	Variável resposta ou dependente	31
3.1.2	Variáveis explicativas	32
3.1.3	Seleção da amostra	35
<i>3.1.3.1</i>	<i>Seleção das Escolas.....</i>	<i>35</i>
<i>3.1.3.2</i>	<i>Seleção dos estudantes</i>	<i>37</i>
3.2	MODELAGEM ESTATÍSTICA	37
3.2.1	Modelos de regressão	37
3.2.2	Modelo de regressão logística	38

3.2.3 Modelo de regressão hierárquico de 2 níveis	39
3.2.3.1 <i>Construção do modelo.....</i>	40
3.2.3.2 <i>Critério de seleção de variáveis</i>	41
3.2.3.3 <i>Análise de resíduos.....</i>	43
4 RESULTADOS E DISCUSSÃO.....	44
4.1 ANÁLISE DESCRITIVA	44
4.1.1 Variável resposta ou dependente	44
4.1.2 Variáveis explicativas – nível dos estudantes.....	46
4.1.2.1 <i>V1E - Turno do Estudante</i>	46
4.1.2.2 <i>V2E - Sexo do estudante.....</i>	47
4.1.2.3 <i>V3E - Cor/raça do estudante.....</i>	48
4.1.2.4 <i>V4E - Faixa etária dos estudantes.....</i>	49
4.1.2.5 <i>V5E - Possui computador em casa?</i>	50
4.1.2.6 <i>V6E - Reside com a mãe?.....</i>	51
4.1.2.7 <i>V7E – Nível de instrução da mãe</i>	52
4.1.2.8 <i>V8E – A mãe do estudante sabe ler/escrever?</i>	53
4.1.2.9 <i>V9E – Nível de instrução do pai.....</i>	54
4.1.2.10 <i>V10E – O pai do estudante sabe ler/escrever?.....</i>	55
4.1.2.11. <i>V11E - Os pais participam de reuniões pedagógicas</i>	56
4.1.2.12 <i>V12E - Os pais incentivam estudar</i>	57
4.1.2.13 <i>V13E - Os pais incentivam realizar o dever.....</i>	58
4.1.2.14 <i>V14E - Os pais incentivam a ler.....</i>	59
4.1.2.15 <i>V15E - Os pais incentivam a não faltar às aulas</i>	60
4.1.2.16 <i>V16E - Os pais conversam sobre as aulas?</i>	61
4.1.2.17 <i>V17E - Leitura de jornal.....</i>	62
4.1.2.18 <i>V18E – Leitura de livros.....</i>	63
4.1.2.19 <i>V19E – Leitura de revistas</i>	64

4.1.2.20	<i>V20E – Leitura de Quadrinhos.....</i>	65
4.1.2.21	<i>V21E - Leitura de notícias na internet</i>	66
4.1.2.22	<i>V22E – Utiliza bibliotecas.....</i>	67
4.1.2.23	<i>V23E – O estudante costuma a ir em espetáculos ou exposição.....</i>	68
4.1.2.24	<i>V24E – Entrada na escola.....</i>	69
4.1.2.25	<i>V25E – O estudante já reprovou?</i>	70
4.1.2.26	<i>V26E – O estudante já abandonou?.....</i>	71
4.1.2.27	<i>V27E – O estudante faz o dever de casa.</i>	72
4.1.2.28	<i>V28E – O professor corrige atividades escolares?</i>	73
4.1.3	Variáveis independentes – das unidades escolares	74
4.1.3.1	<i>V1UE – Tipo de dependência administrativa.....</i>	74
4.1.3.2	<i>V2UE - Área de localização da unidade</i>	75
4.1.3.3	<i>V3UE – Taxa de Formação do docente</i>	76
4.1.3.4	<i>V4UE – Nível socioeconômico da unidade (NSE).....</i>	76
4.1.3.5	<i>V5UE – A unidade possui biblioteca.....</i>	77
4.1.3.6	<i>V6UE – A unidade possui sala de leitura.....</i>	78
4.1.3.7	<i>V7UE – A unidade possui quadra esportiva</i>	79
4.1.3.8	<i>V8UE – A unidade possui laboratório de informática.....</i>	80
4.2	MODELO LOGÍSTICO HIERÁRQUICO	81
4.2.1	Ajuste do Modelo Nulo.....	81
4.2.2	Processo de escolha do melhor modelo	82
4.2.3	O modelo selecionado	84
4.2.3.1	<i>Qualidade de ajuste do Modelo selecionado.....</i>	86
4.2.3.2	<i>Interpretação da Razão de Chances.....</i>	87
4.2.4	Contextualização dos resultados	92
4.2.4.1	<i>Aspectos pessoais do estudante</i>	92

4.2.4.2 Aspectos familiares do estudante	96
4.2.4.3 Efeitos das características das escolas e composição da escola	97
5 CONSIDERAÇÕES FINAIS	99
5.1 OBServações gerais.....	99
5.2 Contribuições do trabalho e sugestões para trabalhos futuros.....	102
REFERÊNCIAS.....	104
ANEXO A – Escala de Proficiência	111
APÊNDICE – Script utilizado para análise dos dados.....	114

1 INTRODUÇÃO

O atual sistema educacional brasileiro fundamenta-se, basicamente, em três textos legais: Constituição da República Federativa do Brasil, promulgada em 1988; Lei de Diretrizes e Bases da Educação Nacional (LDB); e Plano Nacional de Educação (PNE). Segundo a Constituição Federal:

A educação constitui um direito social que visa o pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho, além de delimitar competências e atribuições, regulamentando o financiamento e definindo princípios como pluralismo, liberdade e gestão democrática. (BRASIL,1988)

A LDB (Brasil, 1996) tem por finalidade ajustar os princípios enunciados na Constituição Federal para sua aplicação, definindo a educação como atribuição da família e do Estado, inspirada nos princípios de liberdade e solidariedade humana. A partir dela, a Educação Básica passou a ser estruturada em etapas e modalidades de ensino, sendo composta pela pré-escola, ensino fundamental (organizado em nove anos escolares) e o ensino médio (organizado em três anos), como prioritários e gratuitos. Ainda, delimita as competências e responsabilidades de cada ente federado (União, Estados, Distrito Federal e Municípios) na oferta da educação em seus diferentes níveis e modalidades, destacando que eles deverão organizar, em regime de colaboração, seus respectivos sistemas de ensino. Prioritariamente, compete a cada uma das esferas de governo: Municípios – educação infantil e ensino fundamental; Estado – assegurar o ensino fundamental e oferecer prioritariamente o ensino médio; União – organização do sistema de ensino superior e apoio técnico e financeiro aos demais entes federados.

1.1 CONTEXTUALIZAÇÃO DO PROBLEMA

Além de organizar a estrutura de oferta de ensino, é preciso verificar o andamento dos resultados obtidos, se a organização proposta está sendo efetiva, gerando resultados esperados ou mesmo identificando pontos de fragilidade.

Nesse sentido, medidas de verificação da qualidade foram propostas, como, por exemplo, a Meta 7 do PNE (Brasil, 2014) que tem como objetivo alavancar a qualidade da educação básica, melhorando o fluxo de aprovação e aprendizagem. Para assegurar a prerrogativa estabelecida nesta meta, é necessário garantir a apropriação dos saberes pelos

estudantes, principalmente àqueles que não têm as mesmas condições iniciais, devido à singularidade do percurso de desenvolvimento e da vulnerabilidade na trajetória de vida e da aprendizagem. Adicionalmente, é necessário atuar de forma a garantir que haja uma distribuição social da aprendizagem, para promover a emancipação dos estudantes. Algumas estratégias foram traçadas para assegurar esta meta, seja por movimentos órgãos sem fins lucrativos ou mesmo medidas adotadas em documentos oficiais, como preconiza a Meta 7 do PNE, em sua estratégia 7.2.

Assegurar que no quinto ano de vigência deste PNE, pelo menos 70% (setenta por cento) dos (as) alunos (as) do ensino fundamental e do ensino médio tenham alcançado nível suficiente de aprendizado em relação aos direitos e objetivos de aprendizagem e desenvolvimento de seu ano de estudo, e 50% (cinquenta por cento), pelo menos, o nível desejável. (BRASIL,2014)

Como indicador para essa meta, utiliza-se a avaliação de larga escala, bianual, realizada pelo Ministério da Educação (MEC) por intermédio do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). As aplicações ocorrem no final de cada ciclo do ensino fundamental, 5º e 9º anos, e o 3ª e 4ª séries do ensino médio (BRASIL, 2019).

As avaliações que compõem o Sistema de Avaliação da Educação Básica (Saeb) são formadas por conjuntos de questões, tratadas na literatura também como itens (PASQUALI, 2003), construídas com base em matrizes de referência de cada componente curricular (língua portuguesa e matemática). É importante salientar que, conforme INEP (2019), as matrizes do Saeb não englobam o currículo escolar completamente.

Para a interpretação pedagógica dos resultados obtidos por meio da avaliação, o Saeb apresenta uma lista de níveis de desempenhos que compõem a escala de proficiência (veja Anexo A). A escala do 5º ano de língua portuguesa apresenta cinco níveis, enumerados de 0 a 4. Em cada nível é apresentado o intervalo números de escores da avaliação para qual o nível compete, e a descrição das habilidades desenvolvidas pelos estudantes que se enquadram em cada nível (INEP, 2019).

Apesar dessa apresentação, é mais comum a utilização de apenas quatro níveis de proficiência, delimitados por Soares (2018) com base na escala do Saeb. Como modelo de exposição dos resultados adotado pelo portal Qedu¹, os percentuais de alunos, alocados em cada um dos quatro níveis de proficiência qualitativos, são apresentados de forma a evidenciar a distribuição dos alunos nos níveis de proficiência. Os níveis de classificação são Avançado, Adequado (identificado também como Proficiente), Básico e Insuficiente. Alguns autores, como Soares (2009), Oshiro e Scorzafave (2011) e Kulesa (2015) tratam o nível Insuficiente como Abaixo do básico.

O portal Qedu descreve o nível de proficiência Avançado como de aprendizado além das expectativas, cujas recomendações de atividades para os alunos sejam desafiadoras. O nível Proficiente agrupa alunos preparados para continuar os estudos e recomenda atividades de aprofundamento dos conhecimentos. Já o nível básico, formado por alunos que precisam melhorar, sugerindo atividade de reforço. Por fim, o nível insuficiente representa os alunos que demonstraram pouquíssimo aprendizado, sendo indicada a necessidade de recuperação de conteúdo.

A Tabela 1 ilustra os dados das últimas três edições do Saeb, com os percentuais de alunos alocados em cada um dos quatro níveis de proficiência qualitativos da componente curricular de língua portuguesa, considerando escolas municipais e estaduais do país.

Tabela 1 - Distribuição, em percentuais, de estudantes do 5º ano do Ensino Fundamental, por nível de proficiência qualitativo, considerando a componente curricular língua portuguesa nas edições 2013, 2015 e 2017 do Saeb

Nível de proficiência	Saeb 2013	Saeb 2015	Saeb 2017
Avançado (Igual ou maior que 250 pontos)	12,8%	16,9%	21,5%
Proficiente (200 a 249 pontos)	27,0%	33,3%	34,7%
Básico (150 a 199 pontos)	36,0%	35,0%	31,0%
Insuficiente (abaixo de 150 pontos)	24,2%	14,8%	12,8%

Fonte: Dados do INEP/MEC e compilados pelo Qedu

Pode-se observar que, ao longo das aplicações do Saeb, para os estudantes do 5º ano do ensino fundamental, houve diminuição na proporção de estudantes alocados no nível de proficiência insuficiente das provas de língua portuguesa, com variação de 11,4 pontos percentuais (passando de 24,2% dos estudantes no ano de 2013 para 12,8% na última edição). Desde 2006 a Todos pela Educação, organização civil e sem fins lucrativos, monitora a educação básica e traça metas a serem cumpridas até 2021. A alocação dos alunos em níveis de proficiência permite verificar a eficiência, do processo de ensino, com relação à Meta 3¹. Esta estabelece que no mínimo 70% dos estudantes da educação básica, até 2022, tenham o aprendizado adequado a seu ano escolar. Esse desempenho é observado por meio da junção dos percentuais de alunos alocados nos níveis Proficiente e Avançado. De acordo com os resultados da Tabela 1, ao considerar o percentual esperado de 70% da meta 3, observa-se que a educação

¹ Disponível em <https://www.todospelaeducacao.org.br/pag/dados-5-metas#grafico>

básica ainda precisa percorrer um longo caminho de pelo menos 13,8% para atingir a meta dos 70%.

Segundo o portal Qedu, aproximadamente 310 mil (12,8%) estudantes do 5º ano dos anos iniciais que responderam à Prova Saeb 2017 estão no nível insuficiente em língua portuguesa, com um crescimento de taxa de participação entre as edições de 2013, 2015 e 2017, sendo 89%, 90% e 91% respectivamente.

Em 2015, líderes mundiais reuniram-se na sede da ONU e criaram a Agenda 2030, composta por 17 planos de ações das mais diferentes esferas para o desenvolvimento sustentável do mundo (CASTRO FILHO, 2018). O Objetivo 4 desta agenda é a educação de qualidade que, dentre outros aspectos, prevê:

Até 2030, garantir que todas as meninas e meninos completem o ensino fundamental e médio, equitativo e de qualidade, na idade adequada, assegurando a oferta gratuita na rede pública e que conduza a resultados de aprendizagem satisfatórios e relevantes (NAÇÕES, 2016, p 24 e 25.).

Uma possível abordagem, para promoção da Meta 3, é o estudo de fatores que influenciam o desempenho escolar de estudantes, em especial com desempenho insuficiente, de forma a ofertar subsídios que orientem o trabalho escolar de forma equitativa, objetivando a erradicação de alunos alocados no nível de proficiência abaixo do básico. O estudo de fatores associados ao desempenho escolar vem sendo realizado desde a década de sessenta, tendo como estudo pioneiro o Primeiro Relatório de Coleman (Coleman et al., 1966), em que os autores concluem que, dentre várias características, o principal fator de impacto no desempenho escolar seriam as diferenças socioeconômicas entre os estudantes.

1.2 OBJETIVOS

Nas seções abaixo estão descritos o objetivo geral e os objetivos específicos da dissertação.

1.2.1 Objetivo Geral

Identificar fatores associados à proficiência de estudantes no nível de proficiência abaixo do básico, em língua portuguesa do 5º ano, nos diferentes níveis: estudantes e escolas.

1.2.2 Objetivos Específicos

Como desdobramento do objetivo geral, faz-se necessário pontuar os objetivos específicos, apresentados a seguir:

- Investigar variáveis relacionadas aos estudantes e unidades escolares, que possam influenciar e/ou impactar os resultados da avaliação dos alunos, classificados com proficiência abaixo do básico;
- Efetuar estudo descritivo das variáveis explicativas;
- Aplicar métodos estatísticos, por meio de modelagem hierárquica, para estudo dos fatores associados;
- Realizar uma discussão acerca dos resultados oriundos da aplicação do modelo proposto;
- Documentar etapas e cuidados importantes na realização do processo.

1.3 JUSTIFICATIVA

Todo processo de avaliação que visa avaliar com seriedade e competência a aprendizagem de estudantes, deve ser construído considerando as desigualdades sociais como uma demanda coletiva que interfere no processo de aprendizado e não apenas com um problema social exclusivo do estudante. Segundo Sposati (2010) “o direito em ter diferenças reconhecidas é o que constituiu a equidade”. No caderno produzido pela secretaria municipal da educação de Curitiba (CURITIBA, 2015) apresenta uma oposição entre igualdade e equidade: “[...] A igual distribuição da educação só reproduz as desigualdades sociais já existentes. Assim, equidade educacional significa disponibilizar a todos e a cada um o necessário para sua emancipação social”.

Para que se compreenda, com plenitude, como contornar essas desigualdades, é importante identificar quais (e em qual magnitude) fatores causam impacto efetivo na aprendizagem.

O princípio da equidade ajuda entender que algumas escolas e alguns estudantes precisam de mais auxílio e ações, a fim de assegurar aos estudantes o direito à aprendizagem. Com o estudo dos fatores associados e com ajuda de políticas públicas pode-se atuar especificamente em ações nos fatores que mais impactam no baixo rendimento dos estudantes e com isso reduzir a chance de influenciar no aprendizado.

1.4 ESTRUTURA DO TRABALHO

Esta dissertação está organizada em cinco capítulos. No Capítulo 1, realizou-se uma breve contextualização da pesquisa, norteando a leitura do documento. Em seguida, descreveu-se o problema de pesquisa, foram descritos os objetivos gerais e específicos, bem como a justificativa para realização do estudo.

No Capítulo 2, por meio da delimitação dos conceitos que permeiam a fundamentação teórica do tema, buscou-se apresentar as principais técnicas e referenciais que foram utilizados como base. Importante destacar uma breve apresentação acerca da avaliação em larga escala – Saeb- e a forma como é construída, avaliada e modelada. Por fim, procura-se delimitar o conceito de equidade na educação, para que o acompanhamento do estudo seja efetivo e pleno.

O Capítulo 3 é formado por toda descrição de métodos escolhidos e utilizados para a condução da pesquisa, bem como a justificativa para a escolha dos mesmos. Também contempla a caracterização e estudo da base de dados.

O Capítulo 4 destina-se à apresentação detalhada de uma análise exploratória das variáveis em estudo, bem como com o emprego da metodologia proposta para verificação dos principais fatores associados à proficiência classificada como insuficiente, em língua portuguesa de estudantes do 5º ano e sua respectiva interpretação em ambos os âmbitos: estatístico e pedagógico.

No Capítulo 5 são relatadas as conclusões geradas a partir dos achados no estudo e apresentadas discussões acerca das conclusões, ao que seguem as referências bibliográficas, os apêndices e os anexos.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentadas três subseções, sendo a primeira dedicada ao Saeb, tratando, brevemente, de sua descrição, histórico e do recorte da edição 2017. Na sequência, a segunda subseção trata da equidade na educação. Finalmente, a terceira subseção aborda uma seleção de artigos e dissertações para embasar o trabalho proposto.

2.1 SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA

Composto por um conjunto de instrumentos avaliativos (testes e questionários), o Saeb é classificado como uma avaliação externa em larga escala, que permite ao INEP a realização de diagnóstico da educação básica brasileira, bem como de fatores que podem impactar, de alguma forma, os desempenhos de aprendizagem dos estudantes (INEP, 2020b).

2.1.1 Histórico

A primeira iniciativa nacional, para se conhecer profundamente o sistema educacional brasileiro, começou a ser desenvolvida no final dos anos 80, resultando na primeira aplicação de uma avaliação de larga escala em 1990. Esta aplicação foi o resultado de uma parceria com o Banco Mundial e abrangeu, de forma amostral, escolas públicas com público alvo de 1ª, 3ª, 5ª e 7ª séries do ensino fundamental (INEP, 2020a). Na época, o ensino fundamental contava com oito anos escolares, e apenas os anos ímpares fizeram parte da avaliação.

A partir do ano de 1993, a avaliação passou a ser aplicada bienalmente, mas ainda mantendo os mesmos anos escolares de aplicação e a participação amostral de escolas públicas. Em 1995, o Saeb passou por uma profunda reestruturação metodológica, adotando a teoria da resposta ao item – TRI - (Lord,1978) como instrumento para mensurar o desempenho do estudante, permitindo a comparabilidade entre os resultados das avaliações ao longo do tempo, além de implantar o levantamento de dados contextuais, com os questionários contextuais. Em 1997 foram desenvolvidas as matrizes de referência do Saeb e delimitadas as escalas de proficiência, únicas para os testes cognitivos. (INEP, 2017).

A expansão do Saeb ocorreu por meio da Portaria MEC nº 931, de 21 de março de 2005 (Brasil, 2005). A reestruturação previa que o Saeb seria composto por duas avaliações: a Avaliação Nacional do Rendimento Escolar (Anresc), comumente conhecida como Prova

Brasil, de aplicação censitária para as 4^a e 8^a séries, contemplando as componentes curriculares Língua portuguesa e Matemática, e a Avaliação Nacional da Educação Básica (Aneb), de forma amostral. A aplicação da Prova Brasil, de forma censitária, proporcionou a divulgação dos resultados dos desempenhos para cada escola e município participante. As notas produzidas pela Prova Brasil, a partir de 2007, passaram a constituir parte do Índice de Desenvolvimento da Educação Básica (IDEB), criado no mesmo ano (INEP, 2007).

Em 2017 a portaria do MEC nº 447/17 (Brasil, 2017) ampliou a população-alvo do Saeb. A Avaliação da Alfabetização (ANA), no período compreendido entre 2013 a 2016, era aplicada, em forma censitária, para a população do 3º ano do Ensino Fundamental de 9 anos. A partir do referido ano, a avaliação passou a ter aplicação obrigatória e censitária, aos estudantes matriculados no 3º ano do ensino fundamental. Em 2018, o decreto nº 9.432, de 29 de junho de 2018, regulamentou a Política Nacional de Avaliação e Exames da Educação Básica (BRASIL, 2018). De acordo com o decreto, o Saeb é um conjunto de instrumentos favoráveis à produção e disseminação de evidências, estatísticas, avaliações e estudos a respeito da qualidade da educação básica (BRASIL, 2017).

2.1.2 Saeb - Edição 2017

Em 2017 a aplicação do Saeb ocorreu entre os dias 23 de outubro e 3 de novembro de 2017, seguindo a portaria nº 447, de 24 de maio de 2017 (Brasil, 2017), que estabeleceu as diretrizes para o planejamento e operacionalização da aplicação, com base nas informações do Censo Escolar 2017, referentes à matrícula, delimitou-se o público-alvo:

- As escolas públicas localizadas em zonas urbanas e rurais que possuíam 10 ou mais estudantes matriculados em cada uma das etapas de 3^a ou 4^a série do Ensino Médio (caso a unidade possua ensino médio integrado com técnico);
- Uma amostra de escola privada localizada em zonas urbanas e rurais que possuíam 10 estudantes matriculados em cada uma das etapas de 5º e 9º ano do Ensino Fundamental e 3º ano do Ensino Médio;
- Por adesão, escolas privadas localizadas em zonas urbanas e rurais que possuíam pelo menos 10 estudantes matriculados na 3ª série do Ensino Médio.

As turmas multisseriadas, de correção de fluxo, de educação especial exclusiva, de educação de jovens e adultos, de ensino médio normal/magistério, bem como as escolas indígenas que não ministravam o ensino em Língua Portuguesa, não compuseram o público alvo do Saeb 2017 (BRASIL, 2019a, p. 19).

Participaram do exame mais de 73 mil escolas e em torno de 5,4 milhões de estudantes de escolas públicas ou privadas. Para divulgação dos resultados, o INEP ajustou alguns de seus critérios. Comparada a aplicações anteriores, a exigência da participação dos estudantes matriculados subiu de 50% para 80% e, além disso, era necessário que pelo menos dez estudantes estivessem presentes no dia da aplicação, conforme dados declarados no Censo Escolar 2017 (Saeb, 2019b).

Os testes de desempenho, aplicados isoladamente sem questionários contextuais, contendo informações do nível socioeconômico e da estrutura das unidades escolares, não fornecem informações sobre a vida escolar do estudante. Assim, os questionários contextuais podem ser utilizados como complemento, oferecendo informações relevantes, que possibilitem um melhor entendimento e compreensão dos resultados obtido pelos estudantes em avaliações de larga escala.

Os questionários contextuais, do Saeb 2017, aplicados aos estudantes de 5º ano, são compostos por 51 questões, distribuídas em cinco grandes constructos, conforme apresentado no Quadro 1.

Quadro 1 - Caracterização dos constructos dos questionários contextuais Saeb, de estudantes, e quantidades de questões pertinentes.

Constructo	Nº de questões	Descrição
Caracterização sócio demográfica	25	Abordam características pessoais
Capital social	6	Abordam comprometimento dos pais nas atividades escolares
Capital cultural	10	Identificam as práticas de leitura e atividades culturais fora da sala de aula
Práticas de Estudo	5	Abordam a realização de deveres de casa e ida à bibliotecas.
Trajetória Escolar	5	Abordam reprovações e frequência a longo da jornada escolar

Fonte: Quadro: Autoria própria (2020), adaptado de Saeb (2019c).

Juntamente com os questionários contextuais dos estudantes, foram aplicados questionários aos professores, com 125 questões, aos diretores, com 111 questões, aos aplicadores da prova, acerca da estrutura da escola, com 74 questões.

2.1.3 Escala de proficiência e a Teoria da Resposta ao Item

Escala de proficiência é uma métrica, na escala dos números reais, obtida por meio da TRI, que representa a medida da proficiência dos estudantes em determinada área de conhecimento (PASQUALI, 2003).

O modelo da TRI, adotado pelo INEP, é o modelo logístico unidimensional de três parâmetros (ML3 ou 3PL), que segundo Andrade, Tavares e Valle (2000), apresenta a seguinte equação

$$P(U_{ij} = 1 | \theta_j, a_i, b_i, c_i) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, i = 1, 2, \dots, I, e j = 1, 2, \dots, n. \quad (1)$$

em que U_{ij} é a variável dicotômica que pode assumir: o valor 1 quando o indivíduo j responde corretamente o item i ; ou o valor 0 caso contrário; θ_j é a habilidade do j -ésimo indivíduo; $P(U_{ij} = 1 | \theta_j, a_i, b_i, c_i)$ é a probabilidade de um indivíduo j , com habilidade θ_j , responder corretamente o item i ; b_i é o parâmetro de dificuldade do item i , medido na mesma escala da habilidade; a_i é o parâmetro da distribuição (ou inclinação) do item i , referente à inclinação da curva característica do item (CCI) no ponto b_i ; c_i é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item i .

O modelo apresentado sustenta-se no fato que estudantes com maior proficiência possuem maior probabilidade de acertar o item.

Uma escala de proficiência construída com base na TRI atribui a cada item do teste uma posição que reflete o seu grau empírico de dificuldade, ou seja, o grau de dificuldade observado de acordo com o comportamento do item quando apresentado a participantes de diversos níveis de proficiência. A proficiência de cada participante do teste, estimada a partir de suas respostas aos itens, é posicionada nessa mesma escala, permitindo a comparação entre estudantes e entre ciclos da avaliação (Saeb, 2019d, p. 45).

Para o Saeb, o INEP adota a escala de proficiência que varia de 0 a 500 pontos, com média de 250 e desvio padrão de 50 (INEP, 2019).

2.2 EQUIDADE NA EDUCAÇÃO

A percepção de discrepâncias no aprendizado, vem impulsionando estudos ao longo dos anos como, por exemplo, os primeiros trabalhos conduzidos por Soares e Andrade (2006)

e Scorzafe e Ferreira (2006). Até a década de 2000, as desigualdades na educação não eram problematizadas, o acesso parecia ser condição necessária e suficiente para se garantir justiça em educação (RIBEIRO, 2012). No estudo conduzido por Lima, Mota e Leal (2019), os autores colocam, além da escola, para se garantir equidade.

A educação para equidade não pode ser apenas tarefa da escola, ela deve envolver ações que refletem no comportamento nos mais variados ambientes (família, escola e empresa). O reconhecimento das diferenças do outro desenvolve hábitos mentais e atitudes de abertura e diálogo, ou seja, equidade na educação básica agrega mudança de mentalidade, transforma o modo de pensar, de sentir, de comportar-se em relação aos outros. (LIMA; MOTA; LEAL, 2019, p. 83)

Autores, como Ronca e Alves (2015), apresentam o uma síntese do plano nacional de educação – edição 2014- articulam com os princípios da equidade, justiça social, inclusão e direitos humanos.

“[...] a equidade reconhece, aprecia os padrões de sociabilidade das várias culturas que são parte da identidade brasileira. Compreende que todos são diversos, que a diversidade é inerente ao conjunto dos alunos, inclusive no que se diz respeito às experiências que trazem para o ambiente escolar e aos modos como aprendem.” (BRASIL, 2017, p.11)

Equidade educacional significa oferecer a todos e a cada um o necessário para sua emancipação social, e, para garantir a equidade, é preciso reconhecer que algumas escolas e alguns estudantes precisam de mais apoio e ações, a fim de assegurar aos estudantes o direito à aprendizagem.

Argumenta-se a abordagem do tema, considerando que, em muitos municípios do Brasil, não existem o atendimento de políticas equânimes com um olhar diferenciado para escolas e para os estudantes que estejam em situações de vulnerabilidade. Existem um rol de políticas instituídas que nem sempre são implementadas, ou são de modos equivocados, sem a devida avaliação de eficiência, efetividade e eficácia.

2.3 PRINCIPAIS ESTUDOS REALIZADOS SOBRE O SAEB, FATORES ASSOCIADOS À PROFICIÊNCIA NA EDUCAÇÃO BÁSICA E REGRESSÃO HIERÁRQUICA

Para a verificação dos estudos realizados nos últimos oito anos referentes ao tema proposto, foi realizado uma revisão da literatura em banco de dados (repositório da Capes) de artigos, teses de dissertações. Para tal, foram utilizadas as expressões ou palavras chaves:

- Sistema Nacional de Avaliação da Educação Básica OR Prova Brasil

- Fatores associados ao desempenho escolar OR Fatores associados ao desempenho do estudante

Mediante o exposto, a partir do resultado da pesquisa de 139 artigos e teses, com as palavras chaves citadas anteriormente, são apresentados aqui alguns resultados. Após a leitura dos resumos, observaram-se sete artigos ou teses que mais se identificaram com o tema proposto. Não foi identificado, na limitação da busca, qualquer artigo que abordasse integralmente da temática proposta neste trabalho. Os artigos selecionados abordam parcialmente o tema proposto e estão disponibilizados no Quadro 2.

Quadro 2 - Publicações selecionadas para análise.

	Artigos Selecionados	Autores	Ano	Tipo
1	Fatores associados ao desempenho escolar em Português: um estudo multinível por regiões	Jacob Arie Laros; João Luiz Marciano; Moura de Andrade	2012	Artigo
2	Fatores associados ao desempenho escolar: Resultado de um modelo multinível de valor adicionado	Paulo A. Meyer M. Nascimento	2012	Artigo
3	Fatores associados ao desempenho escolar: uma análise da proficiência em matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro	Gabrielle A. Palermo; Denise Britz do Nascimento Silva; Maria Salet Ferreira Novellino	2014	Artigo
4	Excelência com Equidade: Fatores Escolares para o sucesso educacional em circunstância desfavorável	Ernesto Martins Faria; Raquel Rangel de Meireles Guimarães	2015	Artigo
5	Análise dos fatores associados ao desempenho em matemática dos estudantes do 5º ano da rede estadual de educação do estado de Goiás no SAEGO de 2013	Marcia Maria de Carvalho Pereira	2016	Dissertação
6	Estudo de fatores associados através de regressão quantílica hierárquica	Pedro Alberto Barbeta; Dalton Francisco de Andrade; Hélio Ribeiro Tavares	2018	Artigo
7	Fatores relacionados ao atraso escolar no estado de Minas Gerais	Josiane Souza de Paula; Ana Maria de Paiva Franco; José Waldemar da Silva	2018	Artigo

Fonte: Autoria própria (2020).

A partir da análise dos sete artigos/dissertação selecionados, inicialmente, construiu-se uma nuvem de palavras com os títulos e palavras chaves, a fim de se verificar os termos de maior incidência para contextualizar as problemáticas e principais assuntos abordados nos

relacionada a diferenças socioeconômicas dos estudantes. Ainda, os autores constataram que recursos culturais na família e atraso escolar foram os fatores que mais afetaram o desempenho a nível estudante e a porcentagem de repetentes na turma para o nível escola.

O trabalho de Nascimento (2012) apresenta um estudo com estudantes do 2º ano do ensino fundamental, da cidade de Salvador-BA. Foram aplicados testes padronizados de leitura em dois momentos do ano de 2005 (início e final de ano) a 4.025 estudantes em 55 escolas e 169 turmas da rede estadual de ensino, com o intuito de verificar, além dos fatores tradicionais, se o conhecimento trazido de anos anteriores influenciava o aproveitamento de suas avaliações. Para isso, o autor utilizou o MRH agregado a três níveis: estudantes, turmas e escolas. O autor constatou que os fatores com maior interferência no aproveitamento do estudante foram, de fato, o conhecimento que ele trazia de anos anteriores, além do capital cultural da família e o tipo de escola.

No trabalho de Palermo, Silva e Novellino (2014) foi utilizada a edição 2007 do Saeb para encontrar fatores associados à proficiência do estudante em matemática baseado em variáveis provenientes dos estudantes, das unidades escolares e das turmas. Para amostragem, os autores utilizam o recorte de estudante e escolas situadas na cidade do Rio de Janeiro – RJ que apresentam pelo menos duas turmas. Os autores justificam utilizar o 5º ano do ensino fundamental por conta da existência de um único professor para todas disciplinas. Foram ajustados três modelos iniciais: efeito turma, efeito escola e efeito escola turma explicando escola. Foram selecionadas 188 variáveis: 17 do perfil do estudante, 70 para o nível turma e 101 para unidades escolares. Para realizar a escolha das variáveis, partiu-se do modelo nulo e foram incluídas variáveis utilizando-se de métodos de seleção. O Quadro 3, sintetiza uma revisão de publicações em que os autores constataram grupos de variáveis que poderiam influenciar no desempenho.

Quadro 3 – Construtos de variáveis a partir da revisão dos autores

Estudante	Aspectos pessoais	Características sócio demográficas
		Habilidade prévias
		Atitudes em relação à escola
		Trajectoria escolar
	Aspectos familiares	Estrutura e arranjos familiares
		Capital socioeconômico
		Capital cultural familiar
Capital social familiar		
Escola	Fatores sociais exógenos	
	Efeitos de composição da Escola	Composição de alunos
		Composição de professores
	Características das escolas	
	Clima Escolar	Clima acadêmico
		Violência ou atos ilícitos na escola
		Percepção dos problemas da escola
	Perfil do diretor	
Políticas ou práticas escolares		

Fonte: adaptação de Palermo, Silva e Novellino (2014).

O trabalho de Faria e Guimarães (2015) apresenta uma análise das 215 escolas selecionadas pelo trabalho realizado pelas fundações Lemann e Itaú BBA (2012). As unidades foram separadas em cinco construtos e utilizaram o modelo da TRI de Samejima (1969). Foram constatados que as instalações com melhor estado de conservação, equipamentos multimídias, bibliotecas, poucos problemas com profissionais da educação, clima escolar bom, professores que corrigem dever de casa, maior incidência de professores com nível superior e diretores bem avaliados pelos professores foram fatores determinantes para o desempenho do estudante.

A dissertação de Pereira (2016), apresenta a análise de fatores associados ao desempenho dos estudantes em matemática na rede estadual de educação de Goiás a partir da avaliação interna Saego de 2013. A autora construiu medidas, a partir da TRI, para o nível socioeconômico da família do estudante e para o clima escolar. A reprovação escolar de anos anteriores foi apontada como o fator mais significativo para o desempenho dos estudantes em matemática.

O trabalho de Barbeta, Andrade e Tavares (2018) utilizou a regressão quantílica hierárquica para o estudo de fatores associados em diferentes quantis da distribuição da proficiência em provas de língua portuguesa e matemática do 8º ano do sistema de avaliação de larga escala do Pará. No estudo, foram considerados 13 fatores dos estudantes e 3 das escolas para língua portuguesa. Os autores, a partir do modelo proposto, relatam a diferença entre o desempenho por sexo nas duas avaliações: meninas foram melhores em língua portuguesa e meninos obtiveram melhor desempenho em matemática. Os que reprovaram obtiveram proficiência bem inferior aos que nunca reprovaram. O modelo proposto apresenta interpretação diferenciada para cinco diferentes faixas quantílicas (10%, 25%, 50%, 75%, 90%) possibilitando melhor precisão nos resultados.

Finalmente, Paula, Franco e Silva (2018) utilizam os MRH para investigar fatores associados ao atraso escolar no estado de Minas Gerais, tendo como variáveis explicativas, a nível do estudante: sexo, cor/raça, mora com pai e mãe, escolaridade da mãe, trabalho doméstico, trabalho fora, tempo de lazer, dever de casa, NSE, nível de capital social do aluno; e para o nível escola: tipo da rede de ensino, zona de localização da unidade, número de matrículas no 5º ano, índice da proficiência média, formação de turmas, programa de redução de abandono e macrorregiões.

Ao analisar os sete trabalhos selecionados, verificou-se a importância da utilização dos MRH para explicar, adequadamente, a proficiência de estudantes a partir de um conjunto de fatores associados. A seleção de tais fatores é tema amplamente discutido em vários trabalhos a partir de avaliações locais e/ou nacionais. O questionário do estudante e as características das unidades são fundamentais para realizar as análises. Neste sentido, Palermo, Silva e Novellino (2014) criaram um protocolo de como selecionar as variáveis para trabalhar com fatores associados a partir da revisão de diversos artigos e dissertações. Posteriormente, Barbeta, Andrade e Tavares (2018) apresentam uma nova metodologia, que geram resultados bem mais aprofundados para identificar fatores associados. Neste trabalho utilizaremos uma terceira abordagem, a categorização: abaixo do básico e acima ou igual ao básico para a proficiência do estudante em língua portuguesa. Devido à grande quantidade de dados será utilizado o critério de informação bayesiano (SCHWARZ, 1978) para a seleção de variáveis que será discutida no Capítulo 4.

3 MATERIAL E MÉTODO

Esta seção apresenta duas subseções: caracterização do conjunto de dados do estudo e modelagem estatística, onde são apresentados os procedimentos e modelos utilizados no tratamento e análise dos dados.

3.1 CARACTERIZAÇÃO DOS CONJUNTO DE DADOS

Os dados utilizados para o estudo são disponibilizados de forma gratuita e aberta a todos no site do INEP² na seção microdados.

Para realização deste trabalho foram utilizados os softwares computacionais R (R CORE TEAM, 2020), RStudio e Microsoft Excel 2016. O Excel foi utilizado para tabulação dos dados e o Rstudio para análise dos dados e para o ajuste dos modelos hierárquicos, por meio do pacote `gamlss.mx` (STASINOPOULOS e RIGBY, 2007).

3.1.1 Variável resposta ou dependente

Para cada área avaliada nos testes do Saeb, existe uma escala com interpretação pedagógica, sendo que a escala de língua portuguesa para os estudantes do 5º ano, varia entre os níveis 1 ao 9 (Saeb, 2019e). Nesta escala, são considerados com proficiência abaixo do básico, os estudantes que se encontram em seu primeiro nível, isto é, estudantes com desempenho entre 125 e 150, conforme Tabela 1. Segundo o INEP (2017), estudantes do Nível 1 são aqueles que somente:

“São capazes de: localizar informações explícitas em textos narrativos curtos, informativos e anúncios. Identificam o tema de um texto. Localizam elementos, como o personagem principal. Estabelecem relação entre partes do texto: personagem e ação; ação e tempo; ação e lugar.” (INEP,2017)

Em contrapartida, estudantes entre os níveis 2 ao 9 foram considerados estudantes com proficiência igual ou acima do básico (desempenho igual ou superior a 150). Maiores informações relacionadas aos diferentes níveis de proficiência podem ser encontradas no Anexo A.

² Disponível em: <http://portal.inep.gov.br/web/guest/microdados>

Assim, a variável resposta em estudo refere-se à proficiência em língua portuguesa obtida no Saeb 2017 por estudantes do 5º ano do ensino fundamental e é considerada aqui como dicotômica. A categoria de interesse (categoria 1), a qual buscamos verificar quais fatores mais contribuem para que um estudante obtenha proficiência abaixo do básico segundo a escala, e a categoria 0 que engloba todos os outros níveis.

3.1.2 Variáveis explicativas

Com a finalidade de identificar os fatores associados à obtenção de proficiência abaixo do básico ou insuficiente de estudantes do 5º ano do ensino fundamental, os quais participaram da edição de 2017 do Saeb, foram utilizados os questionários do estudante e das escolas, no site do INEP³ na seção microdados.

As variáveis explicativas selecionadas como candidatas a compor o modelo estatístico são apresentadas no Quadro 4. Uma descrição mais detalhada acerca de cada uma dessas variáveis é apresentada na Seção 4.1.

Quadro 4 - Variáveis explicativas – Nível do Estudante utilizadas no estudo

Variável			Descrição
Sigla	Nível	Nome da Variável	
V1E	Estudante	Turno	Assume o valor "1" se o estudante é do turno da manhã e "0" se o estudante é do turno da tarde.
V2E	Estudante	Sexo	Assume o valor "1" se o estudante é do sexo masculino e "0" se o estudante é do sexo feminino.
V3E	Estudante	Cor/Raça	Assume o valor "0" se o estudante é pardo, "1" se o estudante não declarou sua raça, "2" se é o estudante é branco, "3" se o estudante é preto, "4" se o estudante é amarelo e "5" se o estudante é indígena.
V4E	Estudante	Faixa etária	Assume o valor "0" se o estudante possui idade correta; "1" se o está abaixo da idade correta e "2" se o estudante está acima da idade correta.

³ Disponível em: <http://portal.inep.gov.br/web/guest/microdados>

V5E	Estudante	Possui Computador?	Assume o valor "1" se o estudante não possui computador e "0" se o estudante possui computador.
V6E	Estudante	Reside com a mãe	Assume o valor "1" se o estudante não reside com a mãe e "0" se o estudante reside com a mãe.
V7E	Estudante	Instrução da mãe	Assume o valor "1" se a mãe nunca estudou ou não completou os anos iniciais do ensino fundamental e "0" caso contrário
V8E	Estudante	Mãe sabe ler e escrever?	Assume o valor "1" se a mãe não sabe ler e "0" caso contrário
V9E	Estudante	Instrução do pai	Assume o valor "1" se o pai nunca estudou ou não completou os anos iniciais do ensino fundamental e "0" caso contrário
V10E	Estudante	Pai sabe ler e escrever?	Assume o valor "1" se o pai não sabe ler e "0" caso contrário
V11E	Estudante	Participação em reunião pedagógica?	Assume o valor "1" se os pais nunca ou quase nunca participam de reuniões pedagógicas e "0" caso contrário.
V12E	Estudante	Pais incentivam ao estudo?	Assume o valor "1" se os pais não incentivam ao estudo e "0" caso contrário.
V13E	Estudante	Pais incentivam a realizar o dever	Assume o valor "1" se os pais não incentivam ao dever e "0" caso contrário.
V14E	Estudante	Pais incentivam a ler?	Assume o valor "1" se os pais não incentivam ler e "0" caso contrário.
V15E	Estudante	Pais incentivam a não faltar às aulas?	Assume o valor "1" se os pais não incentivam a não faltar às aulas e "0" caso contrário.
V16E	Estudante	Pais conversam sobre as aulas?	Assume o valor "1" se os pais não conversam sobre as aulas e "0" caso contrário.
V17E	Estudante	Leitura de Jornal	Assume o valor "1" se o estudante nunca ou quase nunca lê jornal e "0" caso contrário.
V18E	Estudante	Leitura de Livros	Assume o valor "1" se o estudante nunca ou quase nunca lê livros e "0" caso contrário.
V19E	Estudante	Leitura de Revistas	Assume o valor "1" se o estudante nunca ou quase nunca lê revistas e "0" caso contrário.
V20E	Estudante	Leitura de Quadrinhos	Assume o valor "1" se o estudante nunca ou quase nunca lê quadrinhos e "0" caso contrário.

V21E	Estudante	Leitura de Notícias	Assume o valor "1" se o estudante nunca ou quase nunca lê notícias e "0" caso contrário.
V22E	Estudante	Utiliza a biblioteca?	Assume o valor "1" se o estudante nunca ou quase nunca utiliza a biblioteca e "0" caso contrário.
V23E	Estudante	Participa de Eventos Culturais?	Assume o valor "1" se o estudante nunca ou quase nunca participa de eventos culturais e "0" caso contrário.
V24E	Estudante	Entrada na escola	Assume o valor "1" se o estudante entrou na escola a partir do ensino fundamental e "0" caso o estudante entrou na escola a partir da educação infantil.
V25E	Estudante	Reprovou?	Assume o valor "1" se o estudante já reprovou e "0" caso contrário.
V26E	Estudante	Abandonou?	Assume o valor "1" se o estudante já abandonou a escola e "0" caso contrário.
V27E	Estudante	Faz dever de casa?	Assume o valor "1" se o estudante nunca faz ou professor não passa dever de casa e "0" se o estudante faz sempre ou quase sempre o dever de casa.
V28E	Estudante	Professor corrige atividades ?	Assume o valor "1" se o professor não passa dever de casa ou não corrige e "0" sempre ou quase sempre corrige o dever.
V1UE	Escola	Tipo de dependência administrativa	Assume o valor "0" se o tipo de escola é federal ou estadual e "1" se o tipo de escola é municipal.
V2UE	Escola	Localização	Assume o valor "1" se a escola está localizada em zona rural e "0" se a unidade está localizada em zona urbana.
V3UE	Escola	Formação do docente	Proporção de docentes que possuem formação correta para lecionar na unidade escolar.
V4UE	Escola	Nível socioeconômico dos estudantes	Assume o valor "1" se a escola se encontra nos níveis 1 e 2 e "0" se a escola se encontra nos níveis 3 a 5.
V5UE	Escola	Possui biblioteca ?	Assume o valor "1" se a escola não possui biblioteca e "0" caso contrário.
V6UE	Escola	Possui sala de leitura ?	Assume o valor "1" se a escola não possui sala de leitura e "0" caso contrário.

V7UE	Escola	Possui quadra ?	Assume o valor "1" se a escola não possui quadra esportiva e 0" caso contrário.
V8UE	Escola	Possui laboratório de Informática?	Assume o valor "1" se a escola não possui laboratório de informática e 0" caso contrário.
V9UE	Escola	Taxa de Abandono	Proporção de estudantes que abandonaram a unidade escolar no ano de 2017.
V10UE	Escola	Proficiência Média da unidade em Língua Portuguesa - 5º ano	Score médio dos estudantes participantes do Saeb 2017 em Língua Portuguesa.

Fonte: Saeb (2017).

3.1.3 Seleção da amostra

Segundo Bolfarine e Bussab (2005), o objetivo da amostragem é realizar constatação para população utilizando-se os resultados de uma amostra. Para realizar uma amostragem, é necessário planejamento para não infringir a seleção aleatória e representativa (KISH,1967).

Para o estudo corrente, foram tomados todos os cuidados com planejamento e a coleta dos dados. Os próximos parágrafos descrevem todo o processo de seleção da amostragem.

O Saeb 2017, contou com a participação de 2.173.332 estudantes do 5º ano do ensino fundamental e mais de 48 mil unidades escolares (Saeb, 2019a). Por conta do tamanho do banco de dados, para atingir o objetivo proposto por esta dissertação, optou-se em criar um plano de amostragem que representasse a população de estudantes e escolas participantes da avaliação. Para isso, foram seguidos alguns passos de redução do banco de dados e também critérios de seleção. Observou-se todos procedimentos para manter aleatoriedade e representatividade de estudantes e unidades escolares. Observando estudos realizados por institutos de pesquisa, como PNAD, definiu-se que o número inicial de estudantes amostrado seria de 50 mil. O número é referência para os demais procedimentos.

3.1.3.1 Seleção das Escolas

Conforme organizado no Quadro 5, inicialmente, foram previstas 49.164 escolas para participarem do Saeb 2017. Deste número, 48.730 escolas realmente efetivaram sua participação e, posteriormente, somente 41.520 unidades atenderam os critérios de divulgação de resultado (INEP, 2017).

Posteriormente, foram removidas as unidades que não apresentaram os seguintes requisitos: i) questionários da escola validados e respondidos, ii) taxa de formação do docente

dos anos iniciais do ensino fundamental e iii) índice de nível sócio econômico calculado. Com o novo filtro, o número final de 33.132 unidades foi obtido.

Quadro 5 – Critério para seleção das unidades (participação da amostragem)

Critério para seleção das unidades - participação da amostragem		Quantidade
1º	Unidades previstas para participar do Saeb 2017 - 5º ano	49.164
2º	Unidades que participaram do Saeb 2017 - 5º ano	48.730
3º	Unidades com divulgação do resultado	41.520
4º	Unidades que apresentaram todos critérios para participar do estudo	33.132

Fonte: Autoria própria (2020).

Com o número de unidades definido, verificou-se a proporção de unidades provenientes de zona urbana e de zona rural. Das 33.132 unidades, 27.458 são provenientes de zona urbana e 5.674 de zona rural, obtendo-se uma proporção de, aproximadamente, 83% e 17%, respectivamente. Para facilitar o estudo adotou-se a proporção 80% - 20%.

Conforme mencionado na Subseção 3.1.3, no primeiro momento definiu-se 50 mil como o número de estudantes da amostra. Assim, o número médio de estudantes que realizaram a avaliação por unidade é 28,02. Realizando a divisão entre o número médio de estudantes e a definição de 50 mil amostrados, obtêm-se 1.785 unidades a serem amostradas.

Finalmente, utilizando a proporção de 80% de unidades urbanas e 20% de unidades rurais supracitada, uma amostra aleatória estratificada com um total de 1.428 unidades urbanas e 357 unidades rurais foi obtida, como mostra a Tabela 2.

Tabela 2 - Seleção da amostra de unidades de acordo com a proporção na população

Tipo da Unidade	População de Unidades	%	Amostra
Urbana - Região Sul	4.945	18,0%	257
Urbana - Região Sudeste	10.727	39,1%	558
Urbana - Região Centro Oeste	2.347	8,5%	122
Urbana - Região Nordeste	6.859	25,0%	357
Urbana - Região Norte	2.58	9,4%	134
Total da Zona Urbana	27.458	100%	1.428
Rural - Região Sul	437	7,7%	27
Rural - Região Sudeste	775	13,7%	49
Rural - Região Centro Oeste	210	3,7%	13
Rural - Região Nordeste	3.291	58,0%	207
Rural - Região Norte	961	16,9%	61
Total da Zona Rural	5.674	100%	357

Fonte: Autoria própria (2020).

3.1.3.2 Seleção dos estudantes

Após realizar o sorteio aleatório das unidades para composição da amostra, foi realizada uma amostragem dos estudantes dentro das unidades sorteadas com o intuito de coletar as 50 mil observações inicialmente previstas. Por conta da utilização dos métodos de modelagem, é necessário que todas as observações (estudantes e unidades escolares) tenham respondido na totalidade as perguntas que vão ser utilizadas para as variáveis explicativas. Com isso foram excluídos 21.036 estudantes e 15 unidades escolares que não atenderam esse pré-requisito. No final de todo procedimento, a base de dados contou com 38.964 estudantes e 1.769 unidades escolares para o estudo.

3.2 MODELAGEM ESTATÍSTICA

Segundo Magalhães e Lima (2015), em diversas situações há interesse em investigar a existência de relação entre variáveis que possibilite a quantificação do efeito que uma variável provoca noutra. Uma das formas de estabelecer esse tipo de relação é por meio de modelos de regressão, sendo o mais simples o modelo de regressão linear.

3.2.1 Modelos de regressão

Elian (1988) coloca a regressão linear como um método global, baseado na utilização de uma equação capaz de explicar a relação entre variáveis de interesse. Modelos de regressão linear podem ser classificados como simples ou múltiplos, de acordo com o número de variáveis explicativas consideradas, sendo classificado como simples os modelos de regressão linear com apenas uma variável explicativa e múltiplo quando mais de uma variável explicativa é introduzida no estudo. Neste tipo de modelagem, o relacionamento entre a variável dependente e cada uma das variáveis explicativas é exclusivamente linear.

O modelo de regressão linear múltipla pode ser escrito como

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i$$

em que y_i é o i -ésimo valor da variável resposta (variável de interesse no estudo), $\beta_0, \beta_1, \dots, \beta_k$, são parâmetros do modelo que devem ser estimados (coeficientes de regressão), os quais estão relacionados às variáveis explicativas, x_1, x_2, \dots, x_n são os i -ésimos valores de cada uma das variáveis explicativas e ε_i são os erros ditos independentes e com distribuição $N(0, \sigma^2)$.

A estimação dos coeficientes β , em geral, é realizada por meio do método de mínimos quadrados ordinários, que consiste em minimizar a soma dos quadrados dos resíduos obtida por meio da diferença entre os valores observados de y e valores esperados $E(Y|X = x)$. (CHARNET et al., 2008).

O coeficiente β_j , $j=1, \dots, k$, de uma variável explicativa representa a variação (aumento ou redução) na média da variável resposta, ou seja, mantendo constante os valores das demais variáveis explicativas, isto é, considerando as mesmas condições para todas as outras variáveis, e aumentando uma unidade em x_j .

A utilização dos modelos de regressão linear clássicos para a explicação de uma variável resposta com base em variáveis explicativas demanda que as seguintes pressuposições sejam atendidas:

- 1) As variáveis explicativas devem ser controladas, sem estarem sujeitas a variações aleatórias;
- 2) para cada valor de uma variável explicativa, os erros devem estar em torno da média, ou seja, $E(\varepsilon_1) = 0$.
- 3) os erros têm a mesma variabilidade em todos os níveis da variável explicativa, ou seja, $V(\varepsilon_i) = \sigma^2$;
- 4) os erros são não correlacionados, ou seja, $Cov(\varepsilon_i, \varepsilon_j) = 0$, para todo $i \neq j$

Nos casos em que a pressuposição de normalidade não é atendida, diferentes metodologias ou abordagens podem ser utilizadas. Exclusivamente para respostas dicotômicas, pode-se, por exemplo, utilizar o modelo de regressão logística (HOSMER, LEMESHOW e STURDIVANT, 2013), conforme apresentado na sequência.

3.2.2 Modelo de regressão logística

Modelos de regressão logística estão inseridos na classe dos modelos lineares generalizados (MLG), propostos por Nelder e Wedderburn (1972), como uma extensão do modelo linear clássico, possibilitando a utilização de outras distribuições (pertencentes à família exponencial) para explicar a variável resposta em estudo, com base em uma função de ligação, que é responsável por descrever a relação matemática entre a componente sistemática e o valor esperado da componente aleatória. (CORDEIRO e DEMÉTRIO, 2011). Assim, os MLGs

podem descrever variáveis que assumem diferentes características, tais como dados de contagem, simétricos e assimétricos, dicotômicos e categóricos.

Dentre os modelos que são capazes de lidar com variáveis resposta dicotômicas, isto é, cuja resposta é zero ou um (evento de interesse), destaca-se o modelo de regressão logística. Neste modelo, a exata probabilidade de a resposta retornar o evento de interesse - em nosso estudo, o estudante apresentar proficiência em língua portuguesa insuficiente - pode ser descrita como

$$P(Y = 1) = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)} \quad (1)$$

em que, Y é a variável dependente dicotômica, $\beta_0, \beta_1, \dots, \beta_k$ são os parâmetros do modelo e X_k são as variáveis independentes. Outro conceito importante e de relevância prática segundo Hosmer, Lemeshow e Sturdivant (2013), é a chamada razão de chances (OR), obtida simplesmente por meio de $OR = \exp(\beta_j)$, $j = 1, \dots, k$. Esta medida indica o quão maior (ou menor) é a chance de uma resposta apresentar um evento de interesse (neste estudo, a proficiência abaixo do básico) em relação ao evento complementar, se uma determinada característica (categoria de uma variável explicativa) for observada. Por exemplo, suponha uma variável explicativa com duas categorias (A e B) em que A é dita a categoria de referência. Se $OR=2$, implica que a chance de um estudante apresentar proficiência abaixo do básico dado que ele possui a característica B é duas vezes maior do que aqueles com característica A.

3.2.3 Modelo de regressão hierárquico de 2 níveis

Apesar dos modelos de regressão usuais serem bastante úteis para o entendimento do comportamento de uma determinada variável resposta de interesse, em certas ocasiões, determinadas características ou comportamentos observados, como, por exemplo, a existência de diferentes níveis nas variáveis explicativas, podem prejudicar as estimativas, bem como a análise de diagnóstico (checagem das pressuposições) do modelo ajustado. Para essas situações, pode-se utilizar como alternativa os modelos de regressão hierárquicos ou multiníveis, como é apresentado na sequência.

Um exemplo clássico para aplicação de um modelo de regressão multinível ou hierárquico são dados relacionados à educação, em que uma clara hierarquia pode ser observada.

É preciso notar que os dados obtidos são estruturados hierarquicamente, isto é, alunos, turmas, escolas, constituem uma sequência natural de agrupamentos aninhados, de tal

forma, que as variáveis representativas das características nos diversos níveis podem interagir com outras variáveis dentro do mesmo nível hierárquico e, também, com variáveis de outro nível. (SOARES e MENDONÇA, 2003, p. 442)

Modelos hierárquicos possibilitam que essas estruturas presentes nos dados sejam consideradas na análise de regressão, permitindo verificar os efeitos que as características de cada nível da variável de interesse exercem sobre a variável resposta.

Os resultados dessas pesquisas indicam que existe sempre uma influência de agrupamento, presente no sistema escolar, no modelo de impacto das variáveis sobre o rendimento escolar, reforçando a ideia de que os alunos não estão distribuídos aleatoriamente pelas escolas. (SOARES e MENDONÇA, 2003, p. 423)

Segundo Soares (2001) as vantagens dos modelos hierárquicos são:

- 1) É possível obter melhor estimativas para os parâmetros por unidades específicas;
- 2) É possível detectar os efeitos entre os níveis;
- 3) É possível calcular a variância de cada componente;

3.2.3.1 Construção do modelo

Hox (2010) propõe passos para construção de modelos hierárquicos. No primeiro passo, apresenta-se o modelo nulo, sendo o modelo mais simples dentre todos apresentados, no qual não são consideradas variáveis explicativas, e é diretamente relacionado ao componente da variância. Verifica-se quanto da variância da variável resposta pode ser explicada pelas diferenças de cada nível (FERRÃO, 2003).

Considerando um modelo clássico de forma hierárquica de dois níveis, tem-se

- 1) para o nível 1: $y_{ij} = \beta_{0j} + e_{ij}$
- 2) para o nível 2: $\beta_{0j} = \gamma_{00} + u_{0j}$

em que y_{ij} é o desempenho do estudante i da escola j ; β_{0j} é o desempenho esperado (desempenho médio) da escola j ; γ_{00} é o desempenho esperado (desempenho médio) das escolas; e_{ij} é o termo aleatório do estudante i da escola j ; e u_{0j} é o termo aleatório da escola j .

Ou ainda, podemos reescrever o modelo como

$$y_{ij} = \gamma_{00} + u_{0j} + e_{ij}.$$

Quando consideramos variáveis independentes para explicar a variável resposta em estudo, escrevemos o modelo hierárquico da seguinte maneira

$$y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + \dots + e_{ij}$$

em que

$$\text{Para o nível 1 : } y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \dots + e_{ij}$$

$$\text{Para o nível 2: } \beta_{0j} = \gamma_{00} + \gamma_{01}W_j + \dots + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + \dots + u_{1j}$$

Para o trabalho proposto, a variável resposta apresenta estrutura categórica dicotômica (dois fatores). As proficiências dos estudantes serão categorizadas em: 1, para estudantes que apresentaram proficiência abaixo do básico e 0, para estudantes que apresentaram proficiência igual ou acima do básico, com isso, como já descrito anteriormente, o modelo de regressão logística é o mais apropriado. Como serão consideradas características referentes aos estudantes e às unidades escolares o modelo hierárquico é o mais indicado.

Proposto por Goldstein (1995), o modelo de regressão logística hierárquico (MLH) apresenta a variável resposta categórica (0 ou 1), em que o sucesso (objetivo do estudo, neste trabalho a proficiência abaixo do básico) é representado por 1 e fracasso por 0. Analogamente à equação (1), no MLH, a exata probabilidade de a resposta retornar o evento de interesse pode ser descrita como

$$P(Y = 1) = \frac{\exp(\beta_{0j} + \beta_1 X_1 + \dots + \beta_k X_k + u_{0j})}{1 + \exp(\beta_{0j} + \beta_1 X_1 + \dots + \beta_k X_k + u_{0j})}$$

3.2.3.2 Critério de seleção de variáveis

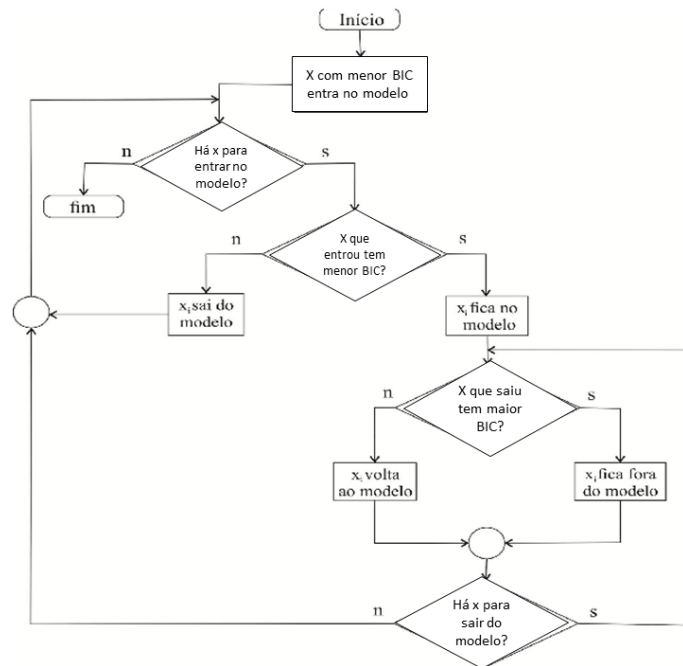
Muitas variáveis explicativas pré-selecionadas para compor o modelo, podem, na prática, não influenciar (não serem significativas) na presença das demais. Recomenda-se, então, a realização de processos interativos de seleção de variáveis independentes.

Existem diversos métodos, dentre eles, alguns dos mais conhecidos algoritmos são: *All Regressions* (todas regressões possíveis); *best subsets* (melhor subconjunto); *forward* (inclusão passo à frente); *backward* (eliminação passo atrás) e *stepwise* (seleção passo a passo) (BERK, 1978).

No trabalho proposto, devido a quantidade de variáveis, escolheu-se o algoritmo de seleção passo a passo (*stepwise*). O método foi apresentado pela primeira vez por Efroymson (1960) e é uma mistura de dois outros procedimentos (*forward* e *backward*).

A variável que é selecionada para o modelo em um passo pode sair nos próximos passos, assim como a variável que sai em um passo poderá retornar em um passo seguinte. O trabalho de Alves, Lotufo e Lopes (2013) apresenta um fluxograma que ilustra o processo *stepwise*. Vale ressaltar que os autores utilizam como critério de seleção das variáveis o coeficiente de correlação linear. A Figura 2 apresenta uma adaptação do fluxograma proposto pelos autores.

Figura 2 - Fluxograma que descreve o método *stepwise* utilizando o critério de BIC



Fonte: Adaptação de Alves, Lotufo e Lopes (2013).

Outros diferentes critérios são apresentados na literatura como o critério de informação de Akaike (AIC) (AKAIKE, 1974) e o critério de informação bayesiano (BIC) (SCHWARZ, 1978).

Neste trabalho, devido à estrutura dos dados, que apresenta um grande número de observações e de variáveis, optou-se pelo BIC que, devido a sua formulação, que em geral, seleciona menos variáveis para o modelo final do que o AIC. O critério descrito por Schwarz (1978) é definido por $BIC = -2 \cdot \ln \hat{L} + k \cdot \ln(n)$, em que L é a função de máxima verossimilhança

do modelo, k é a quantidade de parâmetros que compõem o modelo e n é o número de observações. O modelo que apresenta o menor valor de BIC é o melhor e por consequência o que será escolhido como melhor ajustado.

Após a seleção do modelo final por meio do BIC, podemos, ainda, realizar testes de hipóteses e verificar se, de fato, as variáveis selecionadas são significativas para explicar a variável resposta. Entretanto, vários autores, como Lee et al. (2016), ponderam que deve-se ter bastante cautela em remover variáveis após sua seleção por meio de critérios de qualidade de ajuste.

3.2.3.3 *Análise de resíduos*

Após a seleção do melhor modelo, devemos verificar a qualidade de seu ajuste, bem como com o cumprimento de todas as pressuposições necessárias para sua aplicação. Neste trabalho, utilizamos os resíduos quantílicos aleatorizados normalizados (DUNN e SMYTH, 1996). O interessante deste tipo de resíduo, conforme descrito por Stasinopoulos et al. (2017), é que, independentemente da distribuição assumida para a variável resposta em estudo, se o modelo é adequado para descrever o relacionamento dos dados, então os resíduos seguem, necessariamente, uma distribuição normal padrão. Ainda, para sua visualização, Stasinopoulos et al. (2017) recomendam a utilização dos gráficos de minhoca (VAN BUUREN e FREDRIKS, 2001), que são uma espécie de gráfico quantil-quantil sem tendência. Neste gráfico, espera-se que 95% dos pontos (resíduos) estejam compreendidos dentro de bandas de 95% de confiança.

4 RESULTADOS E DISCUSSÃO

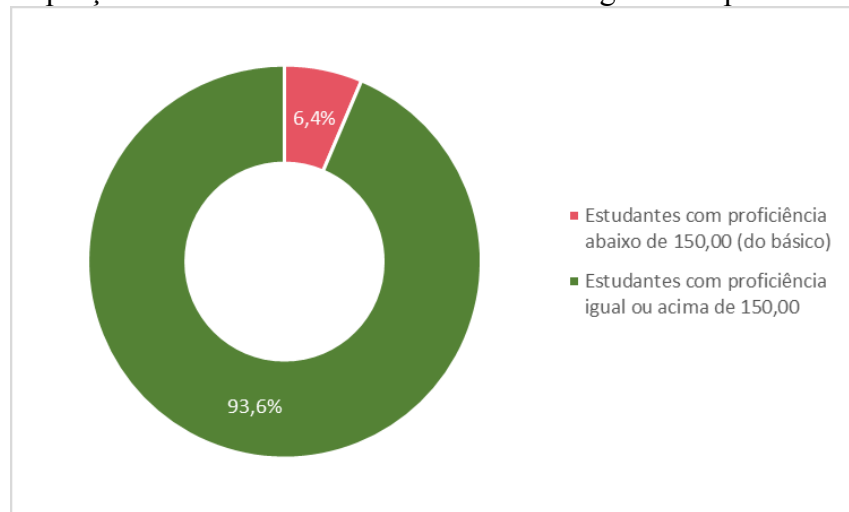
4.1 ANÁLISE DESCRITIVA

A primeira etapa de um trabalho quantitativo é a verificação da natureza dos dados, assim como a realização de uma análise descritiva e exploratória de acordo com as características das variáveis a serem trabalhadas.

4.1.1 Variável resposta ou dependente

Como citado na Seção 3.1.3, 38.964 estudantes foram selecionados para compor o estudo. Deste total, 36.470 (93,6%) observações apresentaram proficiência igual ou superior a 150,00 pontos em língua portuguesa no Saeb 2017 e 2.494 (6,4%) estudantes com proficiências abaixo de 150,00 em língua portuguesa (Figura 3).

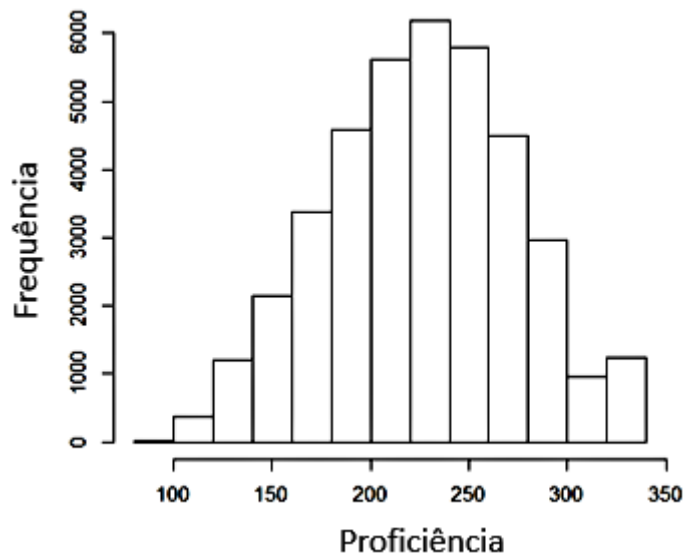
Figura 3 - Proporção de estudantes em cada uma das categorias de proficiências



Fonte: Autoria própria (2020).

A Figura 4 e a Tabela 3 apresentam a distribuição e as medidas resumo das proficiências em língua portuguesa, respectivamente. Percebe-se uma maior concentração em torno da média (225,77) com desvio padrão de 47,77, a mediana é um pouco maior que a média. Utilizando o pacote e1071 do software estatístico R, calculou-se as estimativas de assimetria e curtose para as proficiências dos estudantes amostrados. Para assimetria o valor estimado é -0,063 e para curtose o valor estimado é de -0,486. Com as duas estimativas pode-se colocar que a distribuição dos dados de proficiência possui uma curva mais achatada do que a distribuição normal de probabilidade e conseqüentemente classificada como platicúrtica e quanto a assimetria, identifica-se uma assimetria quase nula.

Figura 4 - Distribuição das proficiências amostradas



Fonte: Autoria própria (2020).

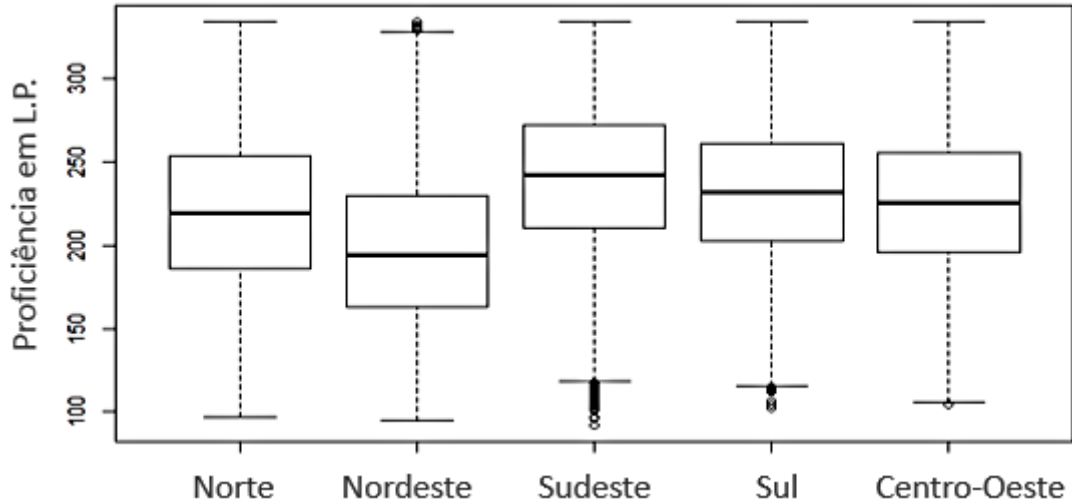
Tabela 3 - Medidas resumo das proficiências dos estudantes amostrados.

Mínima	Máxima	Média	Desvio Padrão	Mediana	1º quartil	3º quartil
91,15	334,23	225,77	47,77	227,19	191,97	259,81

Fonte: Autoria própria (2020).

Um outro gráfico importante para o estudo é o gráfico de caixa, que apresenta a variabilidade dos dados além de indicar possíveis valores discrepantes. A Figura 5 mostra a variabilidade das proficiências comparando estudantes das cinco regiões do país. Percebe-se que a região nordeste apresenta mais da metade dos estudantes amostrados abaixo da proficiência 200,00 ao passo que as regiões Sudeste e Sul apresentam, em média, as melhores distribuições de proficiências comparadas com as demais regiões do país. As regiões: sudeste, sul e centro-oeste apresentam valores discrepantes na parte inferior da distribuição e a região nordeste apresenta valor discrepante na parte superior.

Figura 5 - Variabilidade das proficiências de Língua Portuguesa – 5º ano por regiões do Brasil



Fonte: Autoria própria (2020).

4.1.2 Variáveis explicativas – nível dos estudantes

Nas próximas subseções são realizadas análises descritivas para cada uma das variáveis pré-escolhidas para o nível dos estudantes.

4.1.2.1 VIE - Turno do Estudante

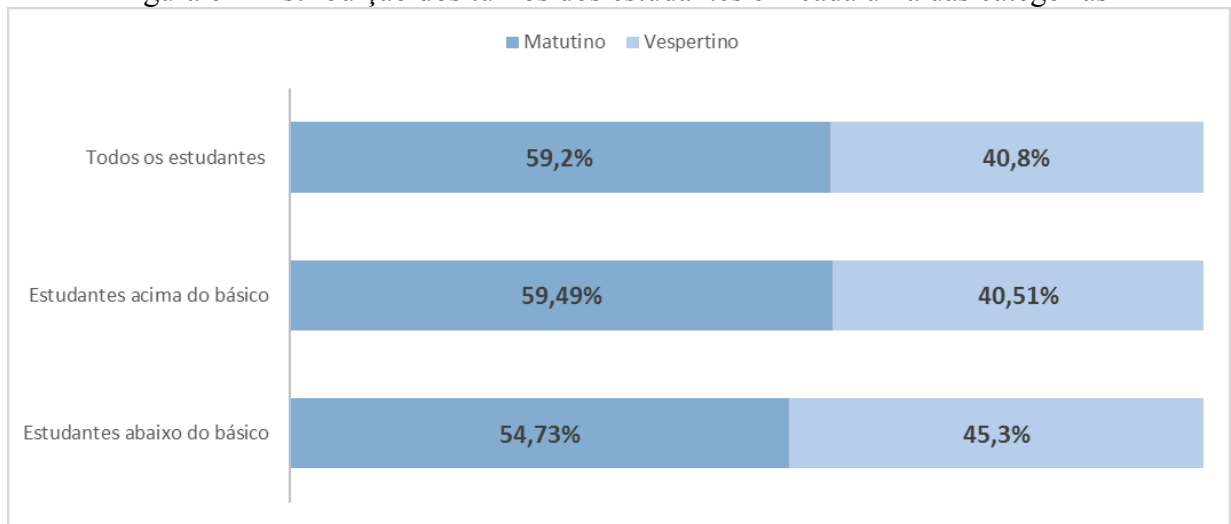
A primeira variável explicativa a ser analisada faz referência ao turno em que se encontravam matriculados os estudantes selecionados (Tabela 4 e Figura 6). Observa-se que a maioria dos estudantes eram provenientes do turno matutino (59,2%). Ao analisar a proporção de estudantes em cada uma das duas categorias, nota-se que a proporção de estudantes abaixo do básico nos turnos é semelhante à do total de estudantes amostrados. Estudantes com proficiência acima do básico possuem uma maior proporção de matrículas no turno vespertino. Em diversos estudos, como é o caso de Almeida (2013), o autor chega à conclusão que estudantes do período matutino apresentam melhor rendimento escolar do que os do vespertino. A proporção de estudantes no vespertino aumenta no grupo de estudantes com proficiência abaixo do básico se comparado com os demais grupos.

Tabela 4 - Frequência Turno do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano

Turno	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Matutino	1.365	21.695	23.060
Vespertino	1.129	14.775	15.904
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 6 - Distribuição dos turnos dos estudantes em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.2 V2E - Sexo do estudante

Os autores Martinelli e Osti (2014) em um estudo sobre desempenho escolar em estudantes de 5º ano do ensino fundamental apontam em seus resultados que as meninas matriculadas nesta faixa etária apresentam um desempenho melhor do que os meninos em língua portuguesa.

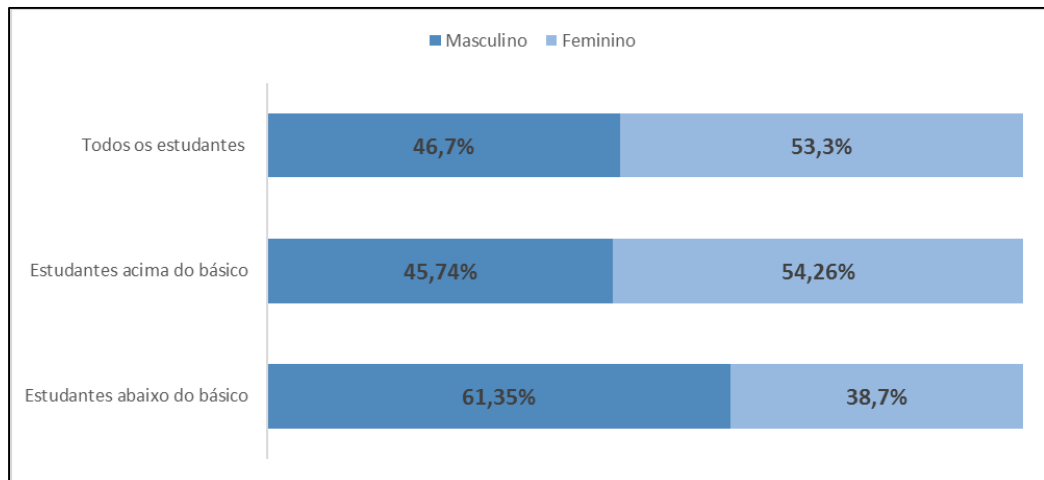
Observa-se que 46,7% dos estudantes selecionados são do sexo masculino e 53,3% são do sexo feminino (Tabela 5 e Figura 7). Ao analisar a proporção de estudantes em cada uma das duas categorias, nota-se uma predominância do sexo masculino (61,3%) na categoria de estudantes com proficiência abaixo do básico e a predominância de estudantes do sexo feminino na categoria acima do básico, resultado que corrobora com as conclusões de Martinelli e Osti (2014).

Tabela 5 - Sexo do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Sexo	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Masculino	1.530	16.681	18.211
Feminino	964	19.789	20.753
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020)

Figura 7 - Distribuição referente ao sexo dos estudantes em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.3 V3E - Cor/raça do estudante

A terceira variável a ser analisada pertence ao grupo de variáveis autodeclaradas pelos estudantes que realizaram a avaliação e amostrados no estudo. O item pertence ao construto de caracterização sociodemográfica do questionário do estudante. No questionário original, o INEP apresentou todas as categorias possíveis de cor/raça dentre elas a opção da não declaração.

Estudantes com cor parda, são a maioria, representando 48,7% do total de estudantes amostrados. Confrontando com as análises entre as categorias de estudantes abaixo e acima do básico, nota-se que a proporção de estudantes brancos decai para 20,33% no grupo com estudantes com proficiência abaixo do básico e cresce para 29,24% na categoria de estudantes acima do básico. A raça preta que apresenta participação 7,3% no grupo de estudante com proficiência igual ou acima do básico e 18,2% no grupo de estudantes com proficiência abaixo do básico.

Tabela 6 - Categorização cor/raça do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Cor/Raça	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Branca.	507	10.663	11170
Preta.	455	2.648	3103
Parda.	1.121	17.713	18834
Amarelo.	78	898	976
Indígena.	78	954	1032
Não declarada	255	3.594	3849
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

4.1.2.4 V4E - Faixa etária dos estudantes

A variável faixa etária dos estudantes pertence ao grupo das variáveis autodeclaradas pelos estudantes que realizaram a avaliação e amostrados no estudo. O item pertence ao construto de caracterização sociodemográfica do questionário do estudante. No questionário original, o INEP apresentou as seguintes categorias: 8 ou menos; 9 anos; 10 anos; 11 anos; 12 anos; 13 anos; 14 anos; 15 anos ou mais. Estudante sem distorção de idade-série e que ingressou em idade correta o ensino fundamental, deverá completar 10 anos no ano em que cursará o 5º ano do ensino fundamental. Por esse motivo, resolveu-se categorizar em três grupos: estudantes entre 8 a 9 anos (abaixo da idade recomendada), estudantes com 10 anos (idade recomendada) e estudantes com 11 anos ou mais (acima da idade recomendada).

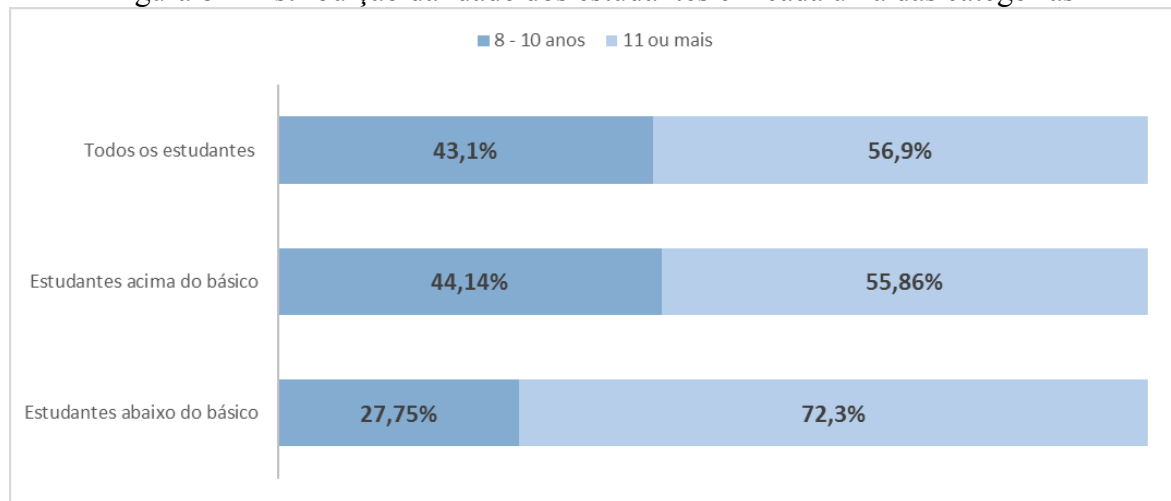
Mais da metade (56,9%) dos estudantes amostrados estão acima da faixa etária recomendada para cursar o 5º ano dos anos iniciais. O índice é bem maior para estudantes com proficiência abaixo do básico (72,3%). Somente 26,6% dos estudantes que se encontra no grupo abaixo do básico possuem idade correta.

Tabela 7 - Categorização idade do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Faixa etária	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
8 ou 9 anos	29	584	613
10 anos	663	15.514	16.177
11 anos ou mais	1.802	20.372	22.174
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 8 - Distribuição da idade dos estudantes em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.5 V5E - Possui computador em casa?

Outro questionamento realizado aos estudantes que realizaram o Saeb/2017, é a quantidade de computadores que possui em casa. O item pertence ao construto de caracterização sociodemográfica do questionário do estudante. No questionário original, o INEP apresentou as seguintes categorias: não tenho; sim, 1; sim, 2; sim, 3; sim, quatro ou mais. Para o estudo a variável foi categorizada em: não tenho ou tenho.

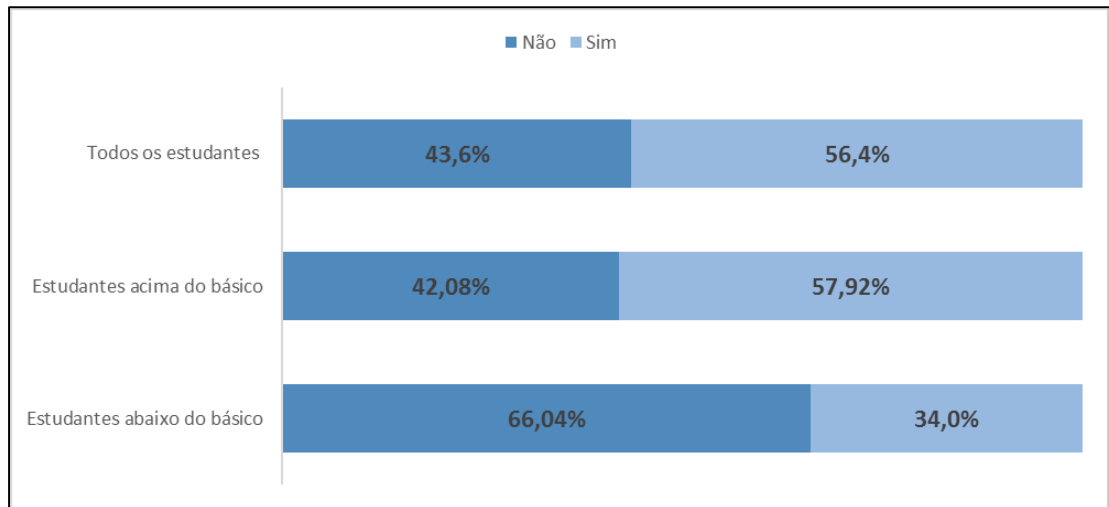
Aproximadamente 22 mil dos quase 39 mil estudantes amostrados possuem pelo menos um computador em casa, correspondendo a uma taxa de 56,4% do total. Ao contrário, 66,04% dos estudantes com proficiência abaixo do básico não possuem computador. A proporção leva a crer que possuir o equipamento está diretamente relacionado ao desempenho em língua portuguesa na avaliação estudada.

Tabela 8 - Categorização idade do estudante x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Possui computador?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	1.647	15.348	16.995
Sim	847	21.122	21.969
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 9 - Distribuição do estudante ter computador em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.6 V6E - Reside com a mãe?

A variável morar com sua mãe pertence ao grupo das variáveis autodeclaradas pelos estudantes que realizaram a avaliação e amostrados no estudo. O item pertence ao construto de caracterização sociodemográfica do questionário do estudante. No questionário original, o INEP apresentou as seguintes categorias: não; sim; não, mas moro com outra mulher responsável por mim. Para o estudo a variável foi categorizada em: sim ou não.

A literatura especializada em pesquisas educacionais tem obtido resultados que associam o bom desempenho escolar do aluno com o fato deste morar com sua mãe (ANDRADE, BRANDÃO e MARTINS 2017).

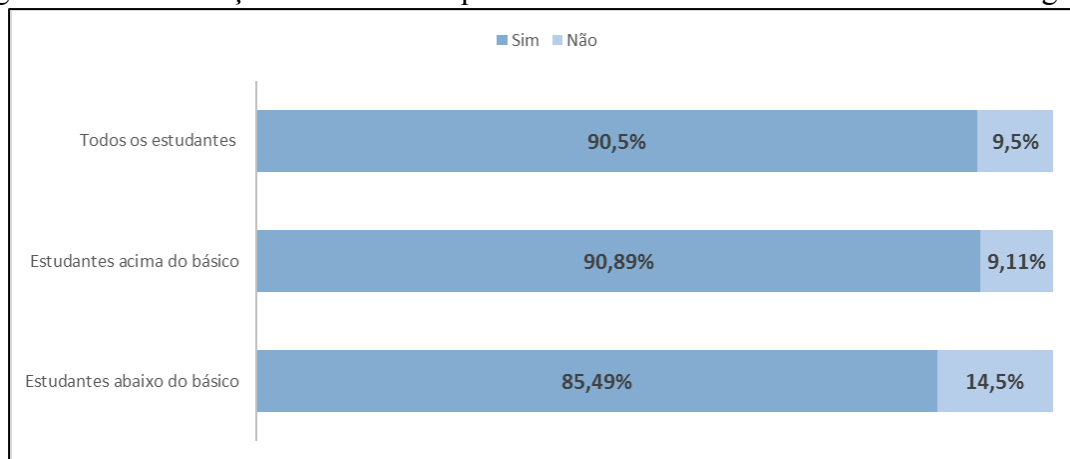
Observa-se que 90,5% do total de estudantes que compõem a amostragem moram com a mãe. Nota-se que na categoria de estudantes com proficiência abaixo do básico, apresenta uma ligeira queda para 85,5%.

Tabela 9 - Categorização morar com a mãe x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Mora com a mãe?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	362	3.324	3.686
Sim	2.132	33.146	35.278
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 10 - Distribuição do estudante que residem com a mãe em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.7 V7E – Nível de instrução da mãe

Outro questionamento realizado aos estudantes que realizaram o Saeb/2017, até que série sua mãe, ou a mulher responsável por você estudou. O item pertence ao construto de caracterização sociodemográfica do questionário do estudante. No questionário original, o INEP apresentou as seguintes categorias: nunca estudou; não completou o ensino fundamental- anos finais; completou os anos iniciais, porém não completou os anos finais; ensino fundamental completo, porém sem completar o ensino médio; completou o ensino médio, porém não completou a faculdade; completou a faculdade; não sei. Para o estudo, a variável foi categorizada em: não sei/nunca estudou/ não completou o ensino básico e completou o ensino básico.

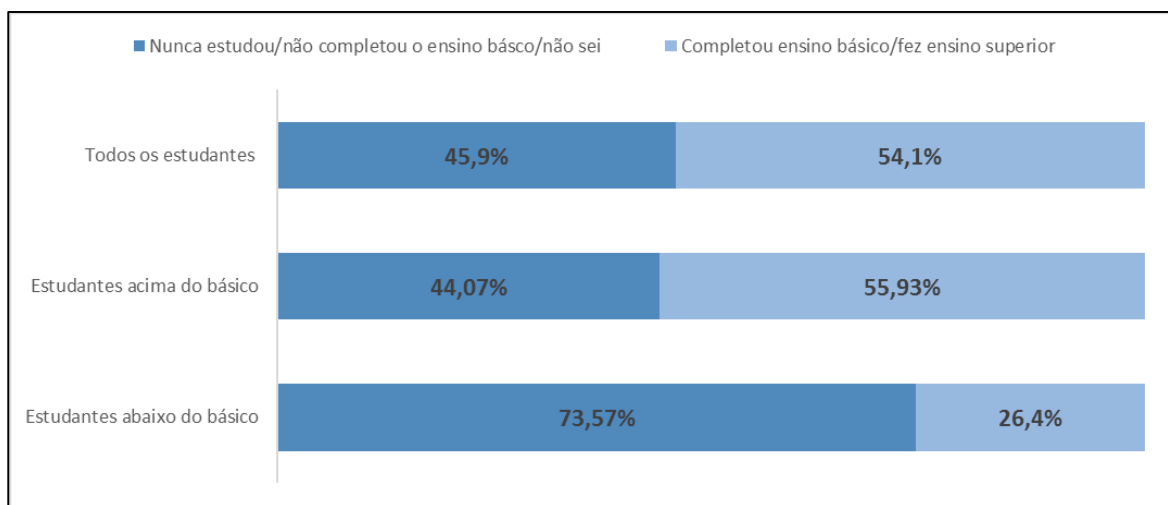
Excluindo a categoria não sei, verifica-se que 54,1% das mães dos estudantes amostrados concluíram o ensino básico e/ou possuem nível superior. Dos estudantes abaixo do básico somente 26,4%.

Tabela 10 - Categorização nível de instrução da mãe x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Nível de instrução da mãe	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Nunca estudou/não completou o ensino básico/não sei	2.142	24.956	35.278
Completou ensino básico/fez ensino superior	352	11.514	3.686
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 11 - Distribuição do nível de instrução da mãe em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.8 V8E – A mãe do estudante sabe ler/escrever?

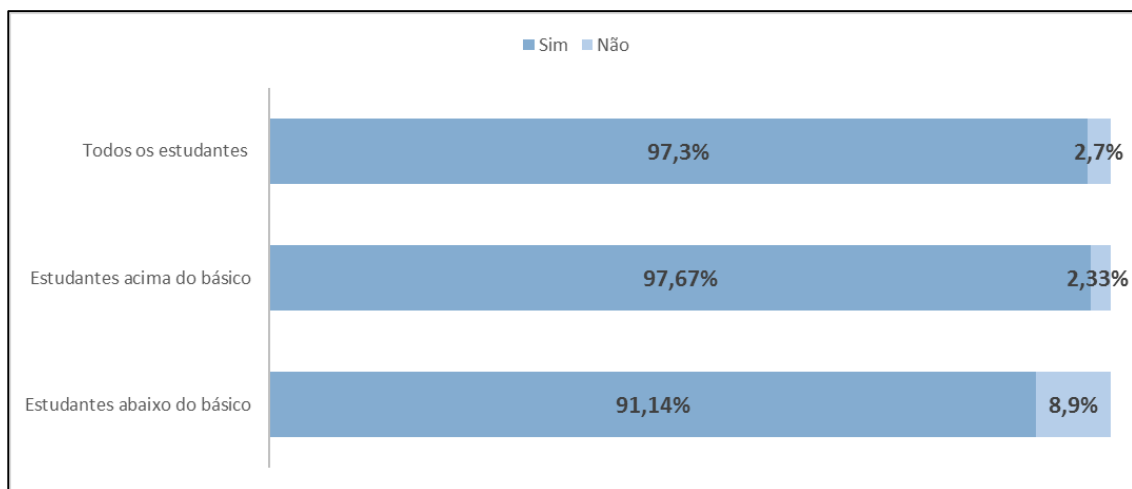
Outro questionamento realizado sobre o perfil da mãe, perguntava se a mãe do estudante sabe ler/escrever. 97,3% do total de estudantes amostrados afirmaram que a sua mãe sabe ler e escrever. Ao categorizar os grupos, percebe-se que a proporção do índice de mães que sabem ler e escrever decaiu no grupo abaixo do básico. 91,1% contra 97,7% do segundo grupo.

Tabela 11: Mãe sabe ler? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Mãe sabe ler/escrever?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	221	848	1.069
Sim	2.273	35.622	37.895
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 12 - Distribuição do questionamento se a mãe sabe ler/escrever em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.9 V9E – Nível de instrução do pai

Agora, perguntou-se aos estudantes até que série seu pai, ou o homem responsável pelo estudante estudou. O item pertence ao construto de caracterização sociodemográfica do questionário do estudante. No questionário original, o INEP apresentou as seguintes categorias: nunca estudou; não completou o ensino fundamental- anos finais; completou os anos iniciais, porém não completou os anos finais; ensino fundamental completo, porém sem completar o ensino médio; completou o ensino médio, porém não completou a faculdade; completou a faculdade; não sei. Para o estudo, a variável foi categorizada em: não sei/nunca estudou/ não completou o ensino básico e completou o ensino básico.

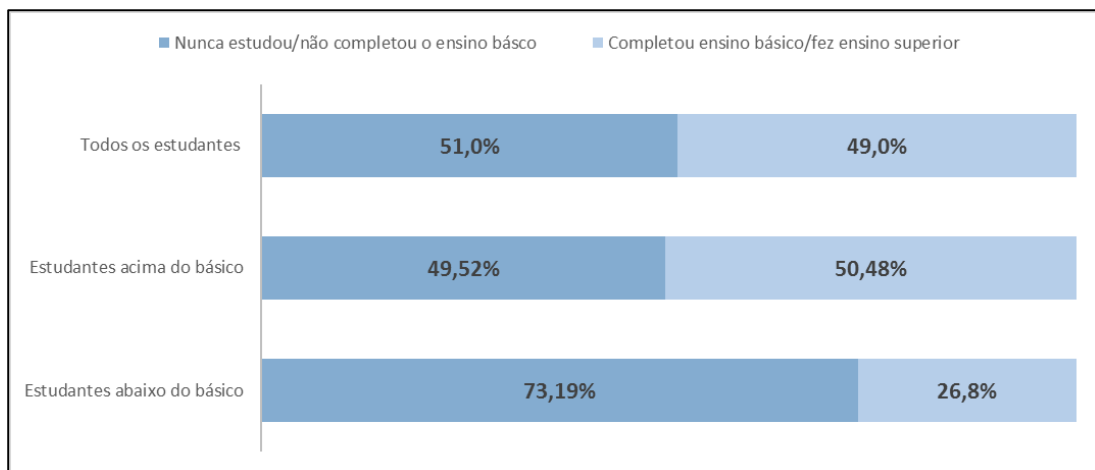
Nota-se que 56,0% dos estudantes não souberam informar. Para melhor análise, considerou-se somente os estudantes que informaram o nível de escolarização de seu pai. Ao analisar a distribuição de respostas por categorias, nota-se 51,0% tem ensino fundamental não completo ou nunca estudaram. Para estudantes que possuem proficiência abaixo do básico, a taxa é muito maior (73,19%).

Tabela 12 - Categorização nível de instrução do pai x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Nível de instrução do pai	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Nunca estudou/não completou o ensino básico	811	8.016	8.827
Completou ensino básico/fez ensino superior	297	8.173	8.470
Total	1.108	16.189	17.297

Fonte: Autoria própria (2020).

Figura 13 - Nível de instrução do pai em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.10 V10E – O pai do estudante sabe ler/escrever?

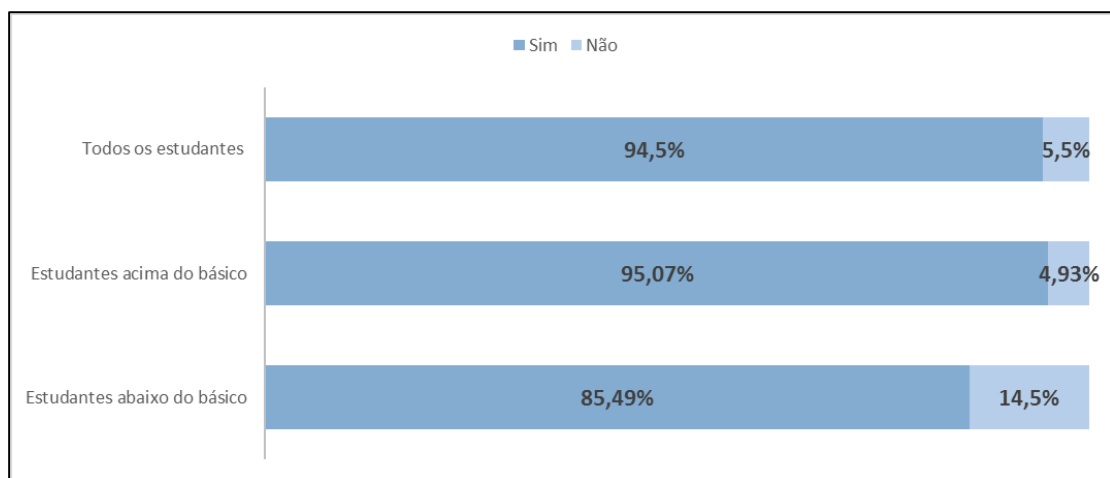
Outro questionamento realizado sobre o perfil do pai, perguntava se o pai do estudante sabe ler/escrever. 94,5% do total de estudantes amostrados afirmaram que a seu pai sabe ler e escrever. Ao categorizar os grupos, percebe-se que a proporção do índice de pais que sabem ler e escrever decaiu no grupo abaixo do básico. 85,49% contra 95,07% do segundo grupo.

Tabela 13 - Pai sabe ler? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Pai sabe ler/escrever?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	362	1.797	2.159
Sim	2.132	34.673	36.805
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 14 - Distribuição do questionamento se o pai sabe ler/escrever em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.11. VIIIE - Os pais participam de reuniões pedagógicas

Outro questionamento realizado aos estudantes que realizaram o Saeb/2017, refere-se à participação dos pais ou responsáveis em reuniões pedagógicas. O item pertence ao construto do capital social no questionário do estudante. No questionário original, o INEP apresentou as seguintes categorias: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para a análise cauterizou-se em duas categorias: nunca ou quase nunca; sempre, quase sempre ou de vez e quando.

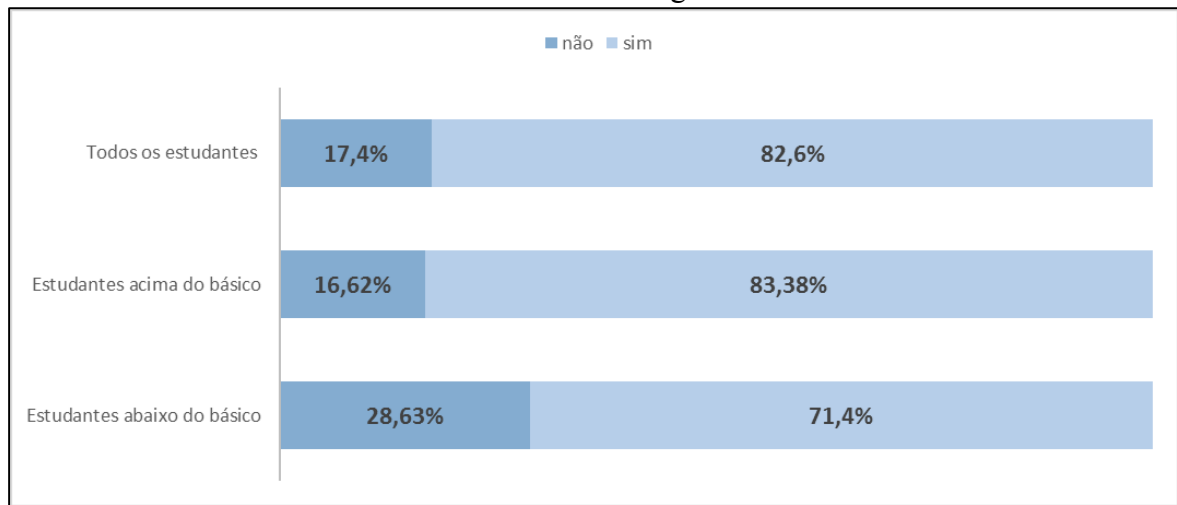
Nota-se que 82,6% dos estudantes apontam que seus responsáveis, participam pelo menos de vez em quando, sempre ou quase sempre. Ao verificar a distribuição por grupos, nota-se que no grupo de estudantes com proficiência abaixo do básico a taxa é menor (71,4%) em contrapartida, no grupo de proficiência maior o índice é um pouco mais elevado (83,4%).

Tabela 14 - Participação dos pais em reuniões pedagógica x Dicotomização das proficiências de Língua Portuguesa – 5º ano

Participação dos pais em reuniões de pais	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	714	6.060	6.774
Sim	1.780	30.410	32.190
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 15 - Distribuição do questionamento se os pais participam de reunião pedagógica em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.12 V12E - Os pais incentivam estudar

Agora, perguntou-se aos estudantes se os pais incentivam a estudar. O item pertence ao construto do capital social no questionário do estudante. No questionário original, o INEP apresentou as seguintes categorias: sim ou não.

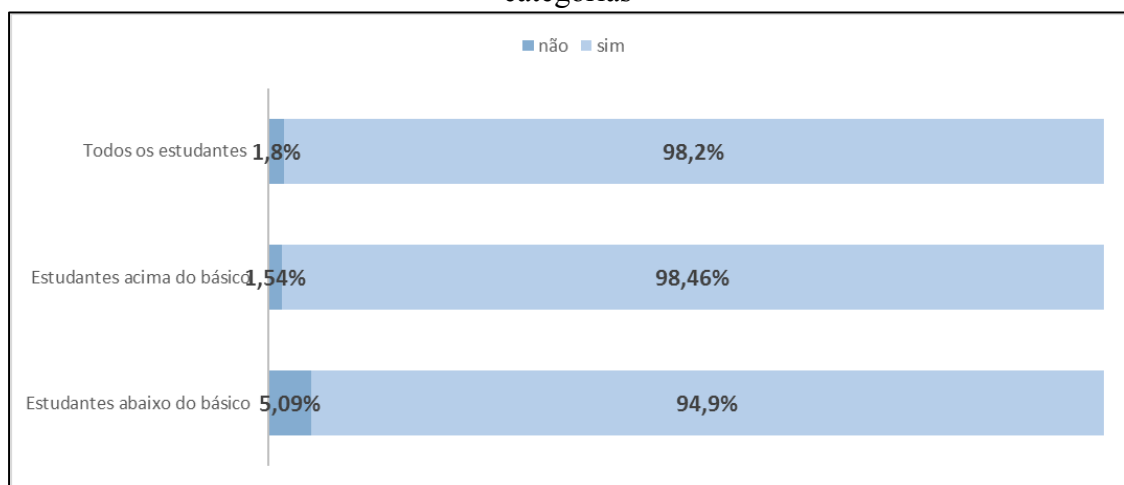
Somente 690 (1,8%) do total de estudantes amostrados, afirmaram que os pais não incentivam a estudar. Esta mesma característica para estudantes com proficiência abaixo do básico é ligeiramente maior (5,09%).

Tabela 15 - Os pais incentivam estudar x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Os pais incentivam a estudar?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	127	563	690
Sim	2.367	35.907	38.274
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 16 - Distribuição do questionamento se os pais incentivam estudar em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.13 V13E - Os pais incentivam realizar o dever

Outro questionamento realizado aos estudantes que realizaram o Saeb/2017, refere-se ao incentivo dos pais a realizar dever de casa e/ou trabalho de escola. O item pertence ao construto do capital social no questionário do estudante. No questionário, o INEP apresentou as seguintes categorias: sim ou não.

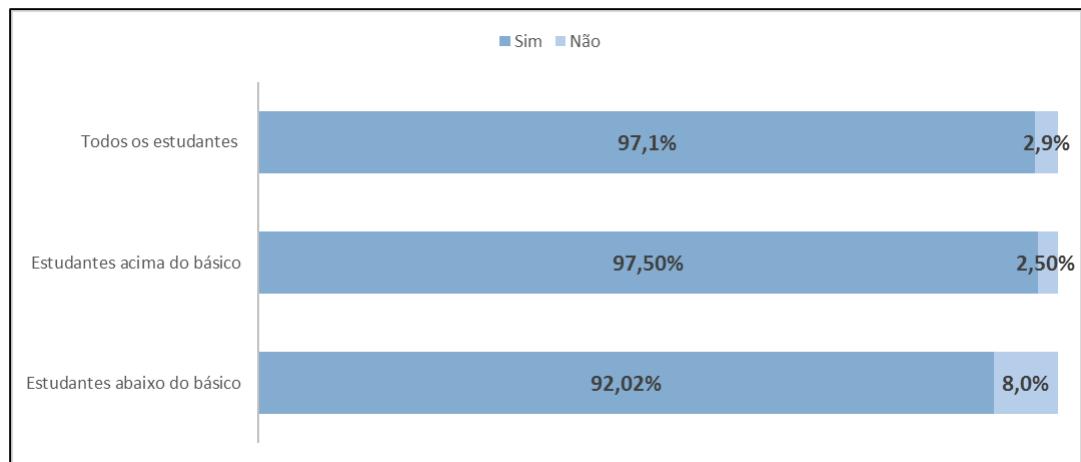
Do total, 97,1% dos estudantes amostrados afirmam que os pais incentivam a fazer a lição de casa. Para os estudantes do grupo com proficiência abaixo do básico a proporção de pais que incentivam decaiu para 92,02%.

Tabela 16 - Os pais incentivam a realizar o dever x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Os pais incentivam a fazer a lição de casa?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	199	912	1.111
Sim	2.295	35.558	37.853
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 17 - Distribuição do questionamento se os pais incentivam a realizar o dever em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.14 V14E - Os pais incentivam a ler

Para o próximo item do grupo capital social do questionário do estudante, pergunta-se se os pais incentivam a ler, com duas alternativas: sim ou não.

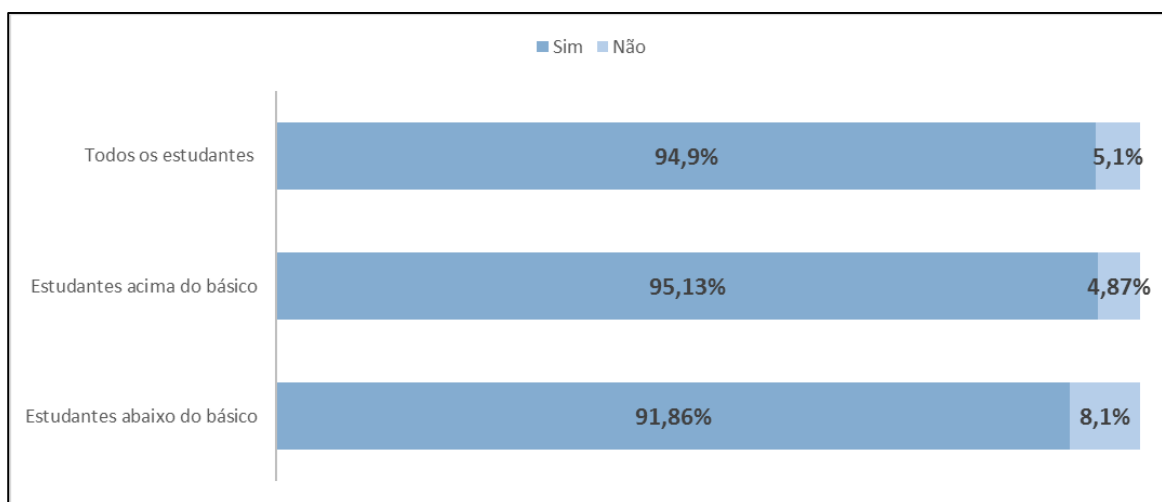
Analisando todos os estudantes amostrados, verifica-se que 94,9% dos estudantes tem o incentivo dos pais para leitura. A proporção é um pouco menor se considerar somente do grupo abaixo do básico (91,86%).

Tabela 17 - Os pais incentivam a ler x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Pais incentivam a ler?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	203	1.775	1.978
Sim	2.291	34.695	36.986
Total	2.494	36.470	38.964

Fonte: Aatoria própria (2020).

Figura 18 - Distribuição do questionamento se os pais incentivam a ler em cada uma das categorias



Fonte: Aatoria própria (2020).

4.1.2.15 V15E - Os pais incentivam a não faltar às aulas

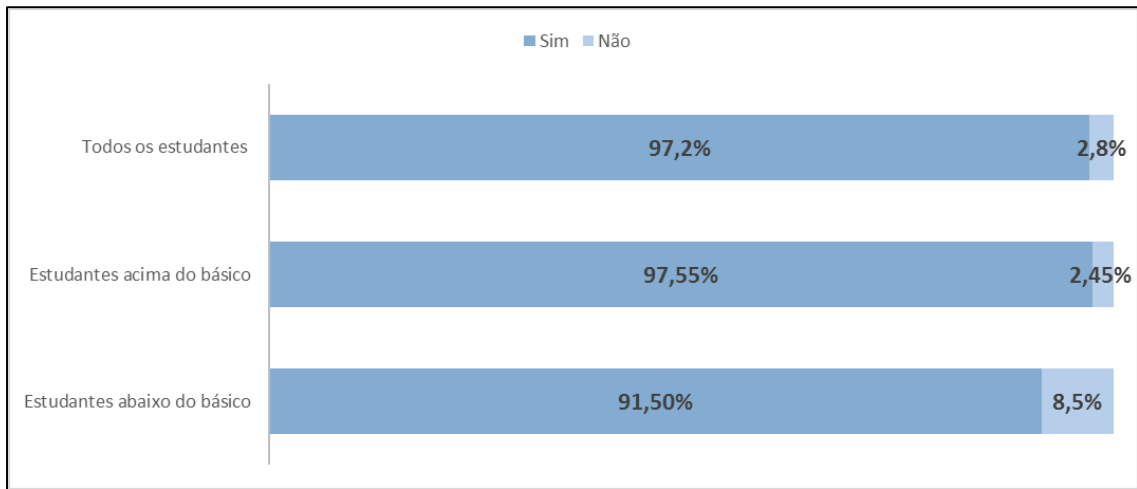
Continuando no conjunto de itens do grupo social, o questionamento é em relação ao incentivo a não faltar às aulas. Observa-se que 97% entre todos estudantes apontaram que os pais incentivam a não faltar às aulas. Verificando e comparando os dois grupos sob estudo, os índices são: 91,5% no grupo 1 e 97,5% no grupo 2.

Tabela 18 - Os pais incentivam a não faltar as aulas x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Os pais incentivam a não faltar a aula?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	212	895	1.107
Sim	2.282	35.575	37.857
Total	2.494	36.470	38.964

Fonte: Aatoria própria (2020).

Figura 19 - Distribuição do questionamento se os pais incentivam a não faltar as aulas em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.16 V16E - Os pais conversam sobre as aulas?

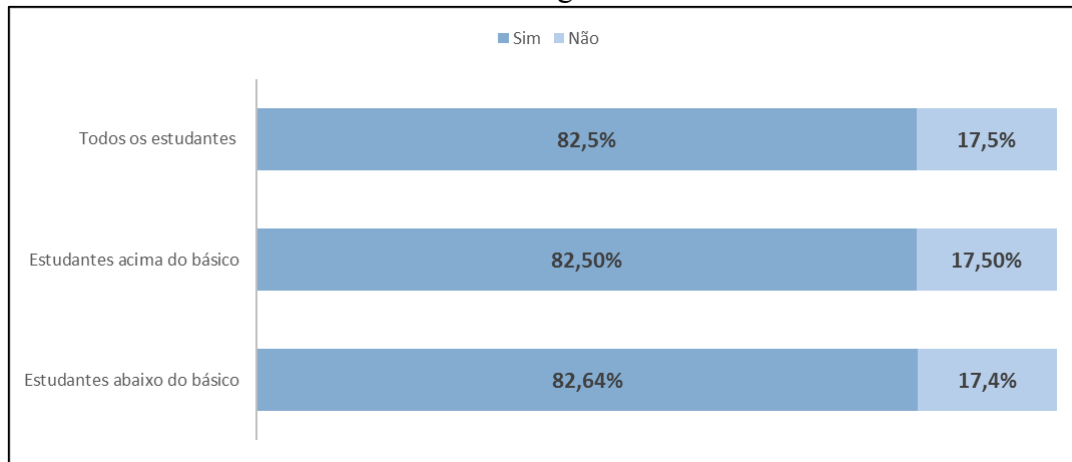
O último questionamento do construto capital social, refere-se se os pais conversam sobre as aulas. É apresentado com duas alternativas: sim ou não. Verificando e comparando os dois diferentes grupos, verifica-se que os índices são muito próximos: 82,6% no grupo dos estudantes com proficiência abaixo do básico e 82,5% no grupo contrário.

Tabela 19 - Os pais conversam sobre as aulas x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Os pais conversam como foi a aula?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	433	6.383	6.816
Sim	2.061	30.087	32.148
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 20- Distribuição do questionamento se os pais conversam sobre as aulas em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.17 V17E - Leitura de jornal

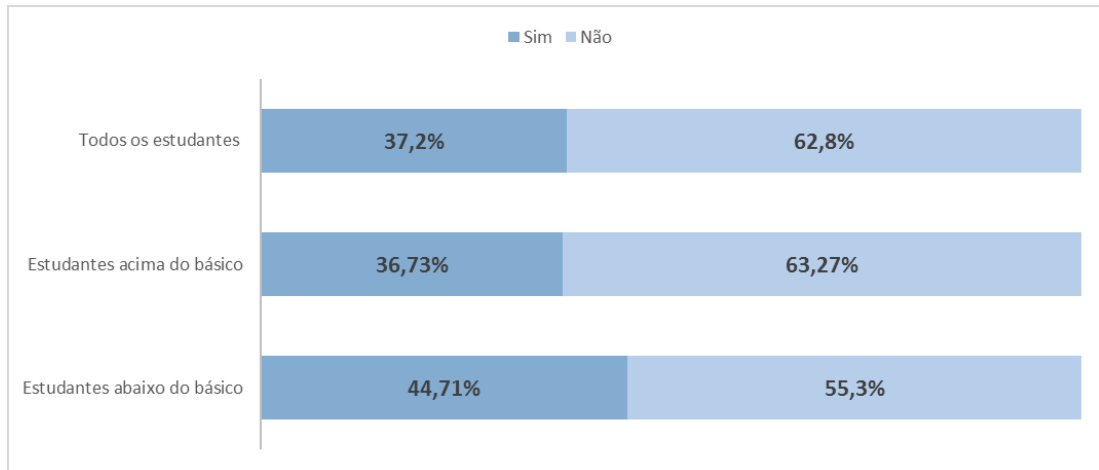
O item, com que frequência você lê jornais, pertence ao construto capital cultural. No questionário do estudante foram relacionadas três categorias para o item: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para o estudo foi considerada a categorização dos itens em: nunca ou quase nunca (não); sempre, quase sempre ou de vez em quando (sim). A maioria dos estudantes (62,8%), afirmam que não leem jornal impresso. No grupo dos estudantes com a proficiência abaixo do básico, ao contrário do esperado, o índice de estudantes que não leem jornal impresso é menor (55,3%).

Tabela 20 - Lê jornais? x Dicotomização das proficiências de Língua Portuguesa – 5º ano

Lê jornais?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	1.379	23.073	24.452
Sim	1.115	13.397	14.512
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 21 - Distribuição do questionamento se lê jornais em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.18 V18E – Leitura de livros

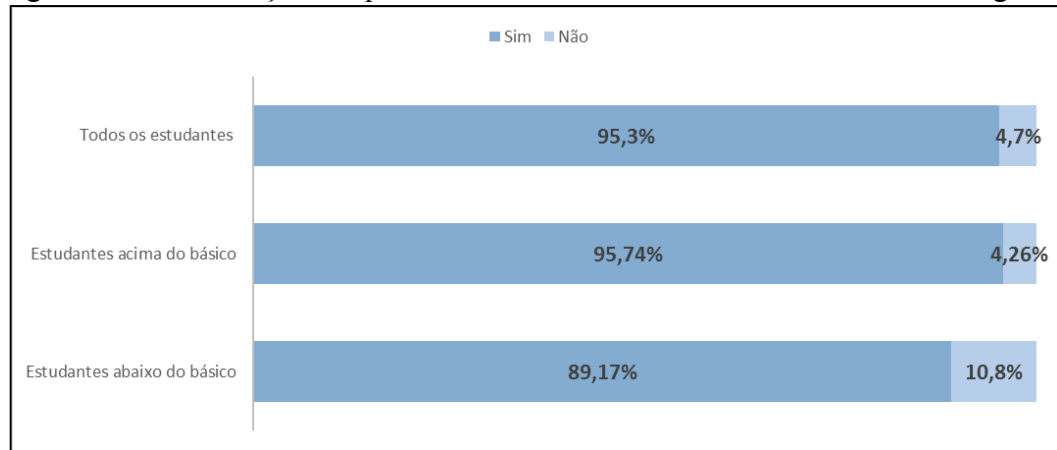
O item com que frequência você lê livros, pertence ao construto capital cultural. No questionário do estudante foram relacionadas três categorias para o item: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para o estudo foi considerada a categorização dos itens em: nunca ou quase nunca (não); sempre, quase sempre ou de vez em quando (sim). Dos estudantes amostrados, 95,3% leem livro com frequências sempre, quase sempre ou de vez em quando. Para o grupo abaixo do básico o índice é inferior (89,17%).

Tabela 21 - lê livros? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Lê livros?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	270	1.554	1.824
Sim	2.224	34.916	37.140
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 22 - Distribuição do questionamento se lê livros em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.19 V19E – Leitura de revistas

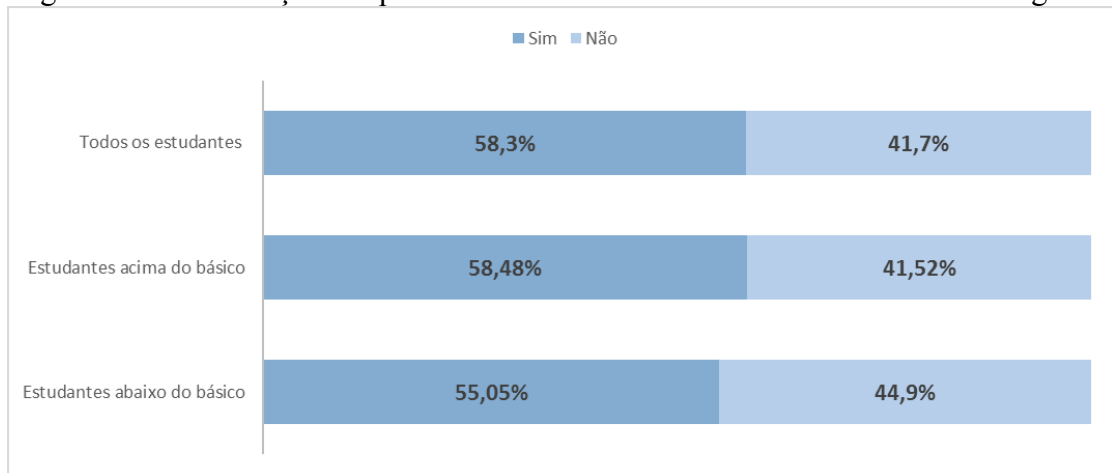
O item com que frequência você lê revistas, pertence ao construto capital cultural. No questionário do estudante foram relacionadas três categorias para o item: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para o estudo foi considerada a categorização dos itens em: nunca ou quase nunca (não); sempre, quase sempre ou de vez em quando (sim). Dos 38.964 estudantes amostrados, 22.701 (58,3%) leem revistas com uma certa frequência. Para o grupo de estudantes com proficiência abaixo do básico essa porcentagem é um pouco menor (55,05%).

Tabela 22 - lê revistas? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Lê revistas?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	1.121	15.142	16.263
Sim	1.373	21.328	22.701
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 23 - Distribuição do questionamento se lê revistas em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.20 V20E – Leitura de Quadrinhos

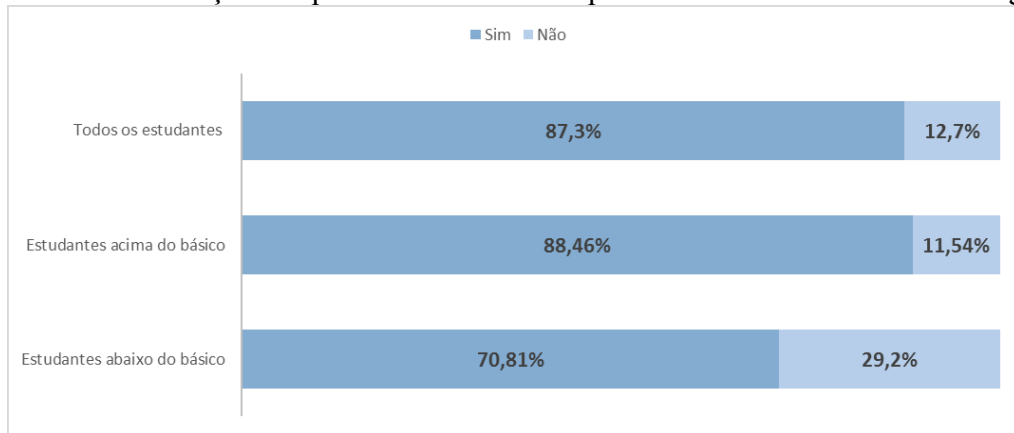
Outro questionamento refere-se à frequência que os estudantes leem quadrinhos. O item pertence ao construto capital cultural. No questionário do estudante foram relacionadas três categorias para o item: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para o estudo foi considerada a categorização dos itens em: nunca ou quase nunca (não); sempre, quase sempre ou de vez em quando (sim). Um pouco mais de 34 mil estudantes, dos 38 mil amostrados (87,3%) leem quadrinhos com frequência sempre, quase sempre ou de vez em quando. Para estudantes do grupo com proficiência abaixo do básico, essa porcentagem cai para 70,81%. Para estudantes do grupo acima do básico 88,46% leem quadrinhos.

Tabela 23 - lê quadrinhos? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Lê quadrinhos?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	728	4.207	4.935
Sim	1.766	32.263	34.029
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 24 - Distribuição do questionamento se lê quadrinhos em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.21 V21E - Leitura de notícias na internet

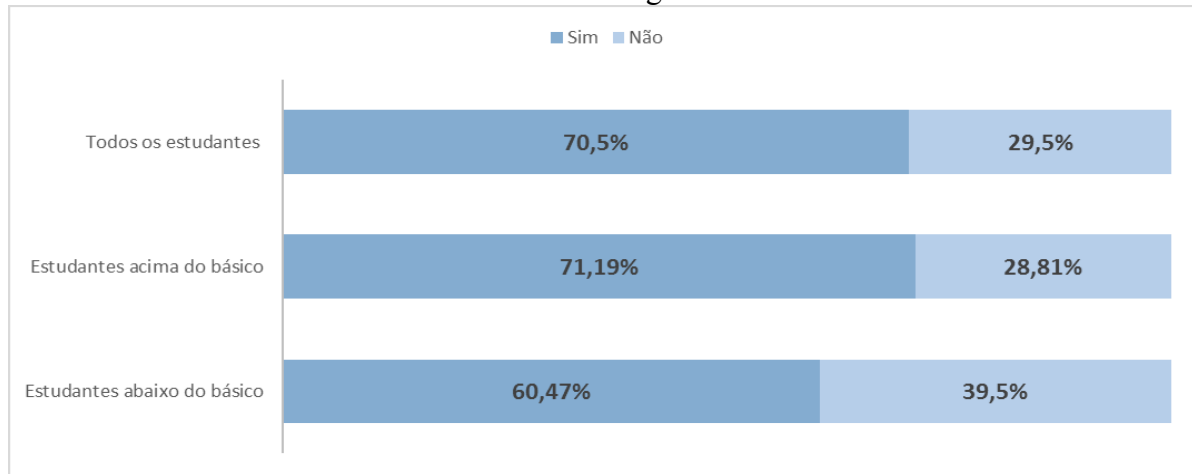
Seguindo com os itens do construto capital cultural, pergunta-se a frequência que os estudantes leem notícias na internet. No questionário do estudante foram relacionadas três categorias para o item: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para o estudo foi considerada a categorização dos itens em: nunca ou quase nunca (não); sempre, quase sempre ou de vez em quando (sim). 70,5% dos estudantes apontaram leem notícias na internet. No grupo dos estudantes abaixo do básico a porcentagem é menor (60,47%).

Tabela 24 - lê notícias na internet? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Lê notícias na internet?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	986	10.508	11.494
Sim	1.508	25.962	27.470
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 25 - Distribuição do questionamento se o estudante lê notícias na internet em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.22 V22E – Utiliza bibliotecas

No construto capital cultura, perguntou-se aos estudantes qual é a frequência que vai à biblioteca. No questionário do estudante foram relacionadas três categorias para o item: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para o estudo foi considerada a categorização dos itens em: nunca ou quase nunca (não); sempre, quase sempre ou de vez em quando (sim).

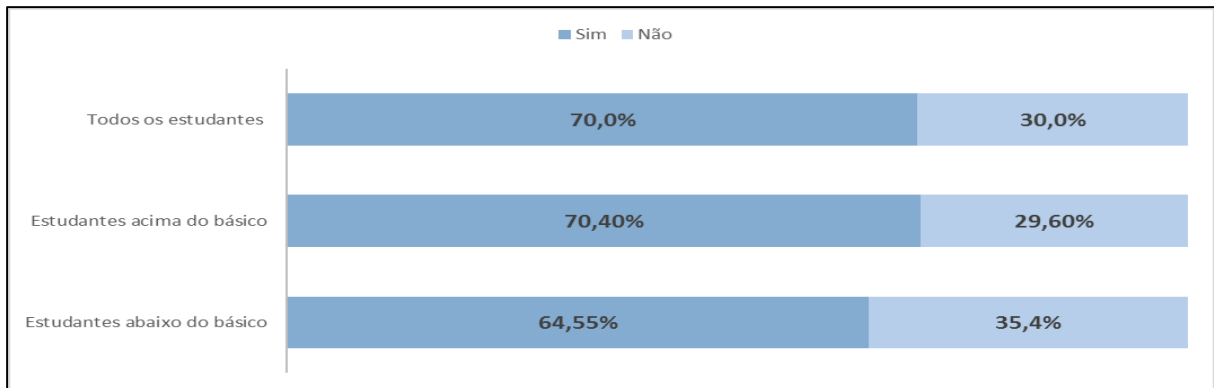
De todos estudantes que compõem a amostra, 70% dos estudantes frequentam biblioteca com uma certa frequência. O grupo dos estudantes acima do básico possui uma taxa parecida (70,4%). Entretanto, o grupo dos estudantes com proficiência abaixo do básico o índice é um pouco abaixo do geral (64,55%).

Tabela 25 - frequenta biblioteca? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Frequenta biblioteca?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	884	10.795	11.679
Sim	1.610	25.675	27.285
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 26 - Distribuição do questionamento se o estudante frequenta biblioteca em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.23 V23E – O estudante costuma a ir em espetáculos ou exposições.

Um segundo item do construto capital cultural, questiona os estudantes sobre a frequência costuma ir em espetáculo ou exposição (teatro, museu, dança ou música). No questionário do estudante foram relacionadas três categorias para o item: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para o estudo foi considerada a categorização dos itens em: nunca ou quase nunca (não); sempre, quase sempre ou de vez em quando (sim).

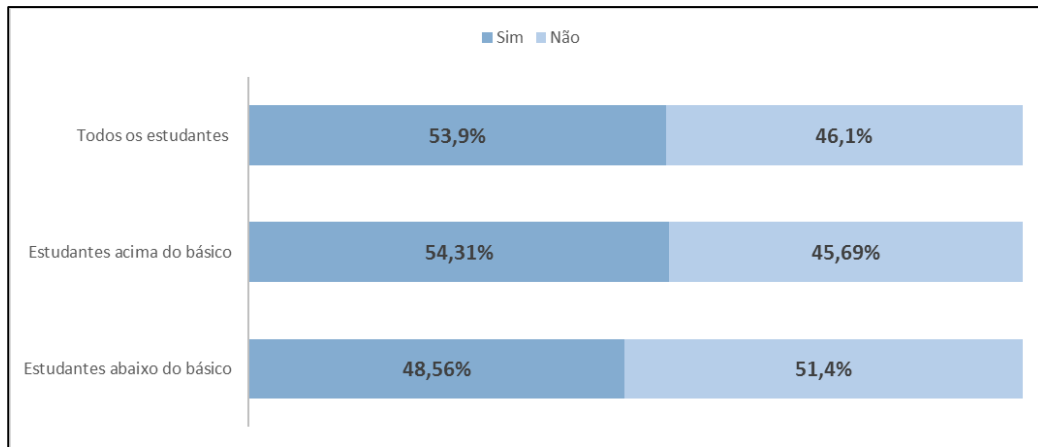
Ao analisar as respostas dos estudantes amostrados, verifica-se que 21.018 estudantes (53,9%) costumam frequentar espetáculos ou exposições. Verificando os grupos, estudantes acima do básico possuem uma porcentagem parecida (54,31%) ao contrário do grupo dos estudantes abaixo do básico, que possuem 48,56%.

Tabela 26 - Costuma ir a espetáculos ou exposições? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Costuma ir a espetáculos ou exposições?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	1.283	16.663	17.946
Sim	1.211	19.807	21.018
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 27 - Distribuição do questionamento se o estudante costuma frequentar espetáculos ou exposições em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.24 V24E – Entrada na escola

O item 24E questiona quando o estudante entrou na escola pertence ao construto trajetória escolar. No questionário do estudante foram relacionadas quatro categorias para o item: na creche (0-3 anos); na pré-escola (4-5 anos); no primeiro ano (6 ou 7 anos); depois do primeiro ano. Para o estudo foi considerada a categorização dos itens em: na educação infantil (0-5 anos); no ensino fundamental (a partir dos 6 anos).

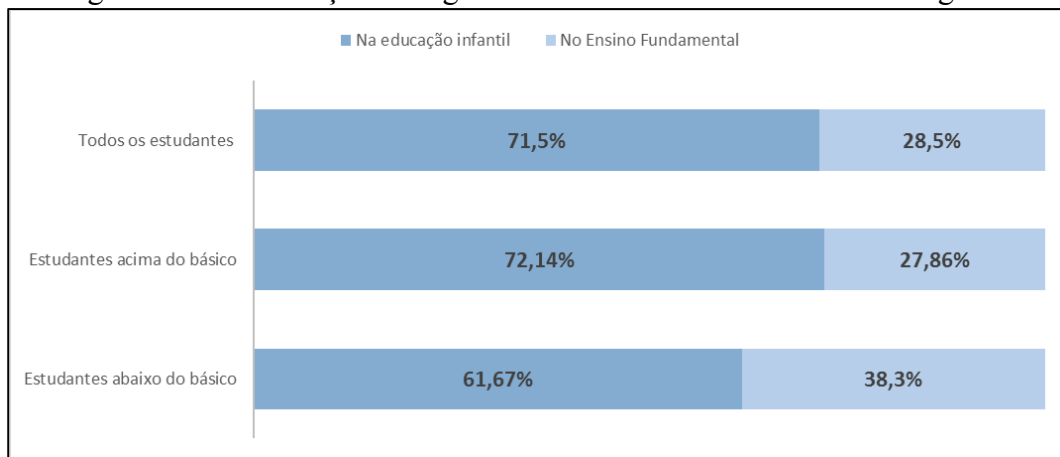
De acordo com os dados, nota-se que 71,5% dos estudantes ingressaram na escola a partir da educação infantil. Estudantes com proficiência abaixo do básico apresentam um índice ligeiramente inferior (61,67%) a porcentagem geral das estudantes.

Tabela 27 - Ingresso na escola x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Ingresso na escola	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Na educação infantil	1.538	26.309	27.847
No Ensino Fundamental	956	10.161	11.117
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 28 - Distribuição do ingresso na escola em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.25 V25E – O estudante já reprovou?

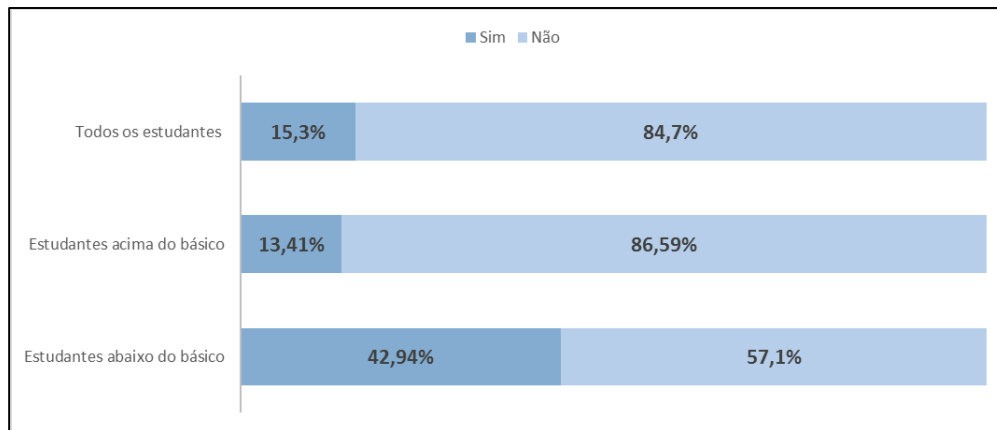
O item V25E questiona se o estudante já reprovou e pertence ao construto trajetória escolar. No questionário do estudante foram relacionadas três categorias para o item: não; sim, uma vez; sim, duas ou mais vezes. Para o estudo o item foi categorizado em: sim; não. Verificando todos os estudantes amostrados, somente 15,3% dos estudantes reprovaram pelo menos uma vez. Para estudantes do grupo abaixo do básico o índice é muito elevado (42,94%). Os estudantes do grupo acima do básico possuem uma taxa aproximada com a de todos os estudantes amostrados (13,41%).

Tabela 28 - O estudante já reprovou? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Já reprovou?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	1.423	31.581	33.004
Sim	1.071	4.889	5.960
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 29 - Distribuição do questionamento se o estudante já reprovou em cada uma das categorias



Fonte: Autoria própria (2020).

4.1.2.26 V26E – O estudante já abandonou?

O item V26E questiona se o estudante já abandonou a escola durante o período de aulas e ficou fora da escola durante o resto do ano, pertence ao construto trajetória escolar. No questionário do estudante foram relacionadas três categorias para o item: não; sim, uma vez; sim, duas ou mais vezes. Para o estudo, o item foi categorizado em: sim; não.

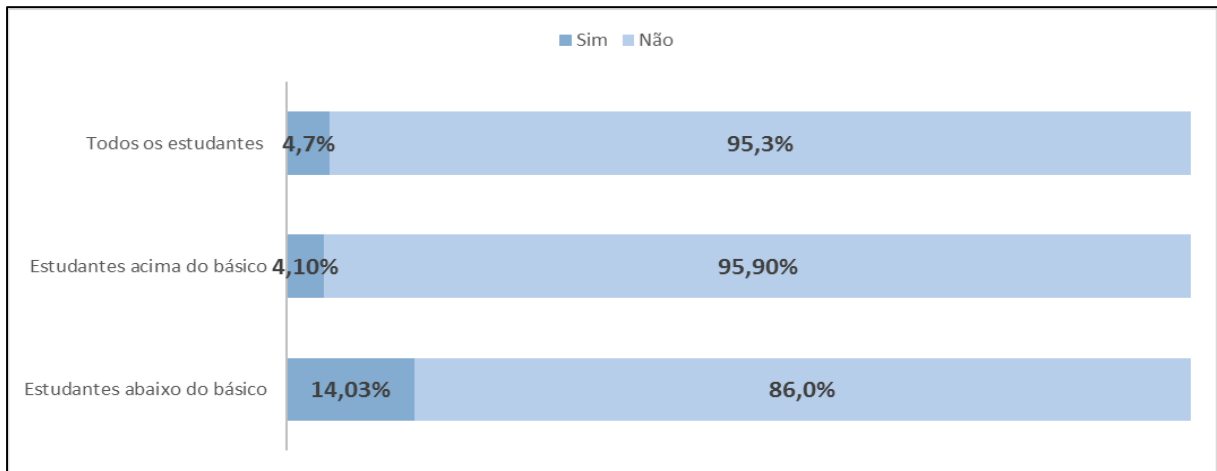
Do total de estudantes amostrados, 1.846 (4,7%) já abandonaram pelo menos uma vez a escola. Para estudantes do grupo abaixo do básico o índice é muito elevado (14,03%). Os estudantes do grupo acima do básico possuem uma taxa aproximada com a de todos estudantes amostrados (4,1%).

Tabela 29 - O estudante já abandonou? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Já abandonou a escola?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	2.144	34.974	37.118
Sim	350	1.496	1.846
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 30 - Distribuição do questionamento se o estudante já abandonou em cada uma das categorias.



Fonte: Autoria própria (2020).

4.1.2.27 V27E – O estudante faz o dever de casa.

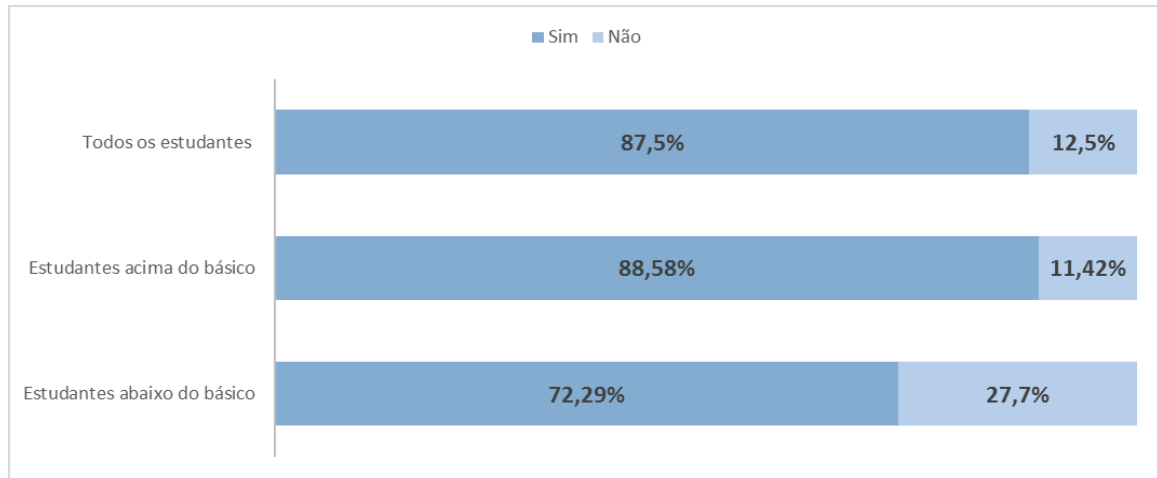
O item V27E questiona se o estudante faz o dever de casa de língua portuguesa, esse item pertence ao construto práticas de estudo. No questionário do estudante foram relacionadas três categorias para o item: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para o estudo foi considerada a categorização dos itens em: nunca ou quase nunca (não); sempre, quase sempre ou de vez em quando (sim). Do total de estudantes participantes da amostra, 87,5% declaram que fazem o dever de casa de língua portuguesa. No grupo dos estudantes que possuem proficiência abaixo do básico, o índice é de 72,2%.

Tabela 30 - O estudante faz o dever de casa? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Faz o dever de casa de L.P.?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Sim	1.803	32.304	34.107
Não	691	4.166	4.857
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 31 - Distribuição do questionamento se o estudante faz o dever de caso em cada uma das categorias.



Fonte: Autoria própria (2020).

4.1.2.28 V28E – O professor corrige atividades escolares?

O item V28E questiona se corrige o dever de casa de língua portuguesa, esse item pertence ao construto práticas de estudo. No questionário do estudante foram relacionadas três categorias para o item: sempre ou quase sempre; de vez em quando; nunca ou quase nunca. Para o estudo foi considerada a categorização dos itens em: nunca ou quase nunca (não); sempre, quase sempre ou de vez em quando (sim).

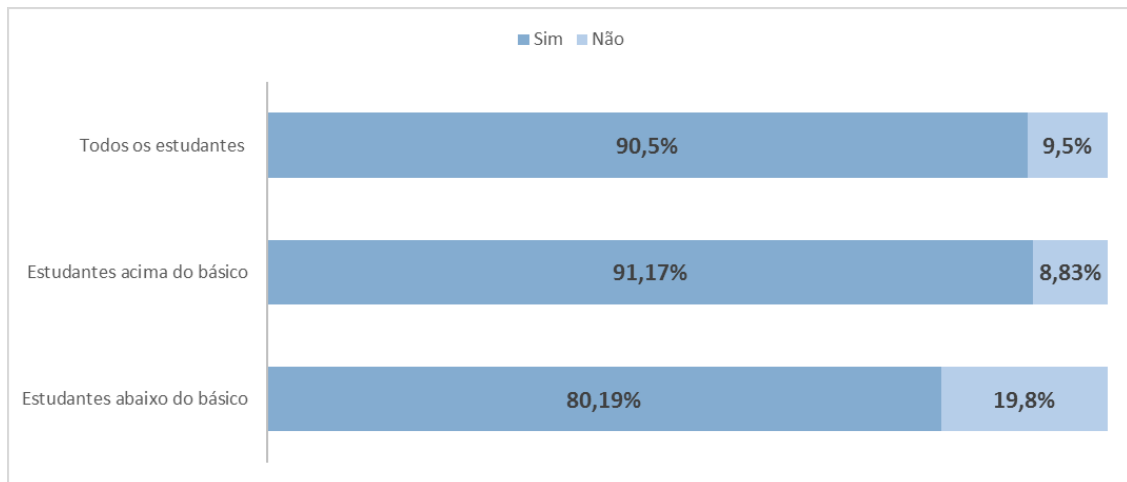
Do total de estudantes participantes da amostra, 90,5% declaram que o professor corrige o dever de casa de língua portuguesa. No grupo dos estudantes que possuem proficiência abaixo do básico, o índice é de 80,2%.

Tabela 31 - O professor corrige o dever de casa de língua portuguesa? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

O professor corrige o dever de casa de L.P.?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	494	3.222	3.716
Sim	2.000	33.248	35.248
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 32 - Distribuição do questionamento se o professor corrige o dever de casa de língua portuguesa em cada uma das categorias.



Fonte: Autoria própria (2020).

4.1.3 Variáveis independentes – das unidades escolares

Nas próximas subseções são realizadas análises descritivas para cada uma das variáveis pré-escolhidas para o nível escola.

4.1.3.1 VIUE – Tipo de dependência administrativa

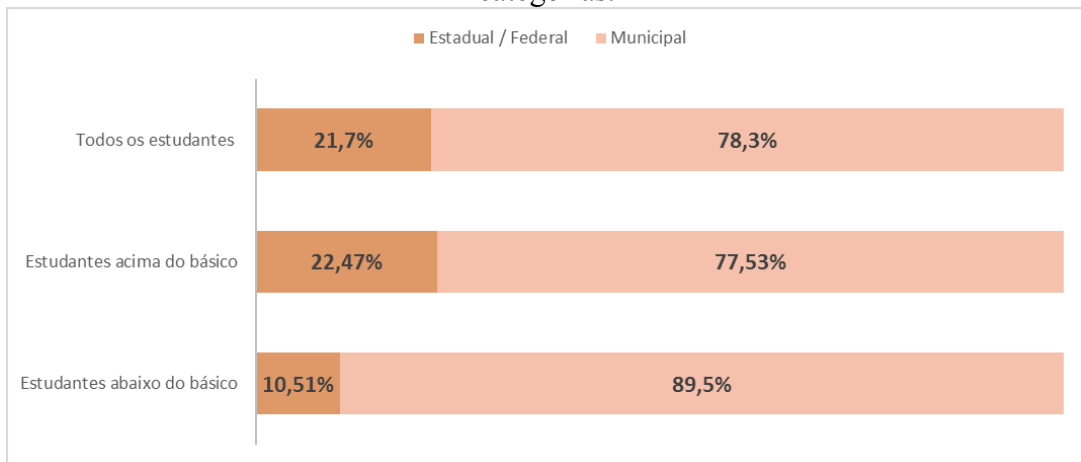
O item VIUE apresenta qual o tipo de dependência administrativa que unidade em estudo se encaixa. Os estudantes do 5º ano que realizaram a avaliação são provenientes de três tipos de dependência administrativa: municipal, estadual ou federal. Para o estudo, escolas provenientes de redes estaduais e federais foram categorizadas no mesmo grupo. Observa-se que 78,3% do total de estudantes selecionados pela amostra são provenientes da rede municipal de ensino. No grupo dos estudantes com proficiência abaixo do básico existe um aumento dessa proporção passando para 89,5%.

Tabela 32 - Tipo de dependência administrativa x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Tipo de dependência	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Estadual / Federal	262	8.194	8.456
Municipal	2.232	28.276	30.508
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 33 - Distribuição do tipo de dependência proveniente do estudante em cada uma das categorias.



Fonte: Autoria própria (2020).

4.1.3.2 V2UE - Área de localização da unidade

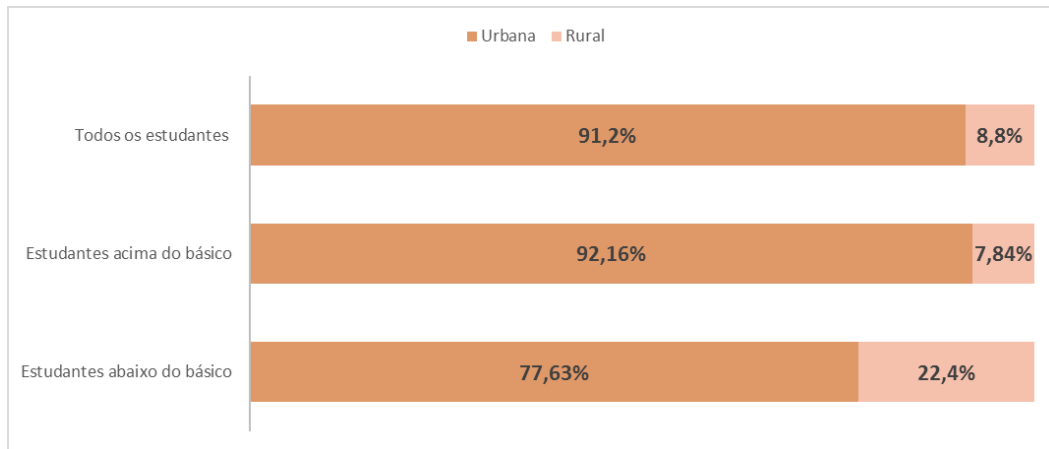
O item V2UE apresenta qual o tipo de localização a unidade encontra-se situada. São duas categorias: urbana ou rural. Do total de estudantes amostrados, somente 3.419 (8,8%) estudantes são provenientes de escola urbana. No grupo dos estudantes com proficiência abaixo do básico, a porcentagem é muito maior, chegando a 22,4%.

Tabela 33 - Área de localização da unidade x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Localização?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Urbana	1.936	33.609	35.545
Rural	558	2.861	3.419
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 34 - Distribuição da localização da unidade em cada uma das categorias.



Fonte: Autoria própria (2020).

4.1.3.3 V3UE – Taxa de Formação do docente

O item V3UE apresenta a taxa de formação correta de docente para atuar nos anos iniciais do ensino fundamental nas unidades amostradas. A variável é numérica. Variando de 0% a 100%. São dados coletados pelo censo escolar 2017 (INEP, 2017).

4.1.3.4 V4UE – Nível socioeconômico da unidade (NSE)

O item V4UE apresenta o nível socioeconômico das unidades escolares (NSE). O NSE é um indicador calculado pelo INEP por meio de questionário dos estudantes do Saeb 2015 e do Censo Escolar também daquele ano. O INEP divide as escolas em cinco grupos. Grupo 1 são as escolas com nível sócio econômico mais baixo e o grupo 5 escolas com nível sócio econômico mais alto. Para o estudo os grupos foram agrupados: categoria 1 (grupos 1 e 2) e categoria 0 (grupos 3, 4 e 5).

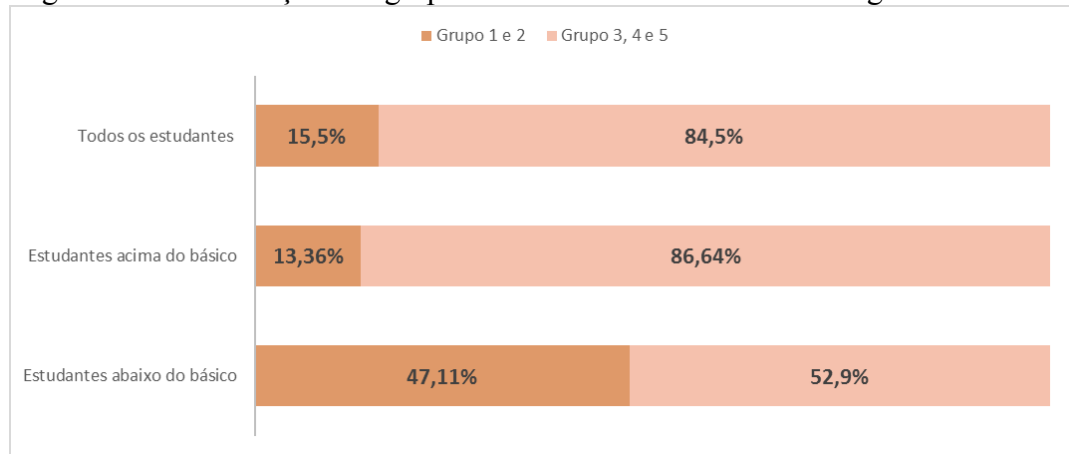
Dos 38.964 estudantes amostrados, 6.047 (15,55%) são provenientes de escolas com nível socioeconômico muito baixo (grupos 1 e 2). Analisando o grupo de estudantes com proficiência abaixo do básico, a proporção de estudantes provenientes de unidades no grupo socioeconômico muito baixo é de 47,11%.

Tabela 34 - NSE x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

NSE	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Grupo 1 e 2	1.175	4.872	6.047
Grupo 3, 4 e 5	1.319	31.598	32.917
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 35 - Distribuição dos grupos do NSE em cada uma das categorias.



Fonte: Autoria própria (2020).

4.1.3.5 V5UE – A unidade possui biblioteca

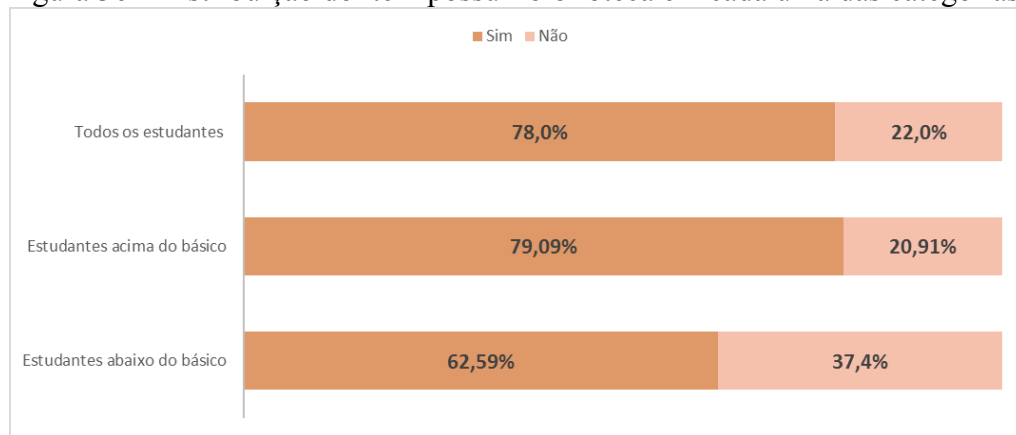
O item V5UE pertence ao questionário escola. É coletado a informação sobre a existência e as condições de uso da biblioteca na unidade. São apresentadas quatro categorias: bom; regular; ruim; inexistente. Para o estudo a variável foi categorizada em: possui (sim) ou não possui (não). Nota-se que 22,0% dos estudantes selecionados para o estudo, são de escolas que não possuem biblioteca. Ao analisar o grupo de estudantes com proficiência abaixo do básico, a taxa de estudantes provenientes de escolas sem biblioteca aumenta para 37,4%.

Tabela 35 - A unidade possui biblioteca? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Biblioteca?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	933	7.626	8.559
Sim	1.561	28.844	30.405
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 36 - Distribuição do item possuir biblioteca em cada uma das categorias.



Fonte: Autoria própria (2020).

4.1.3.6 V6UE – A unidade possui sala de leitura

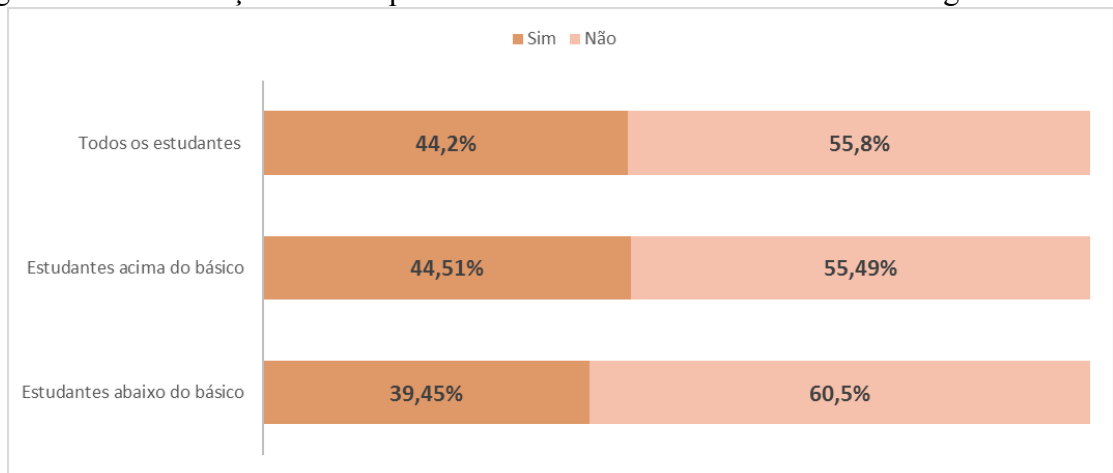
O item V6UE pertence ao questionário escola. É coletado a informação sobre a existência e as condições de uso da sala de leitura na unidade. São apresentadas quatro categorias: bom; regular; ruim; inexistente. Para o estudo a variável foi categorizada em: possuiu (sim) ou não possuiu (não). Observa-se que 55,8% dos estudantes selecionados para o estudo, são de escolas que não possuem sala de leitura. Ao analisar o grupo de estudantes com proficiência abaixo do básico, a taxa de estudantes provenientes de escolas sem sala de leitura aumenta para 60,5%.

Tabela 36 - A unidade possui sala de leitura? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Sala de Leitura?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	1.510	20.239	21.749
Sim	984	16.231	17.215
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 37 - Distribuição do item possui sala de leitura em cada uma das categorias.



Fonte: Autoria própria (2020).

4.1.3.7 V7UE – A unidade possui quadra esportiva

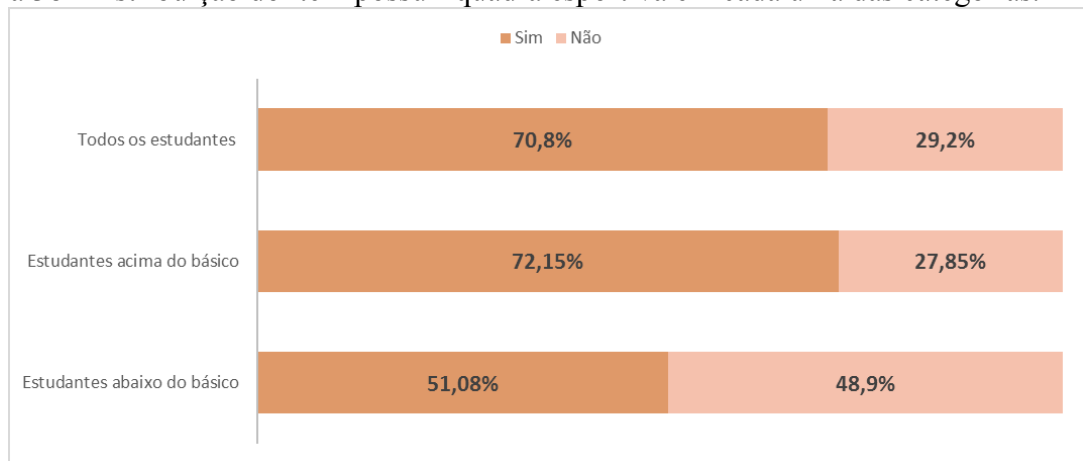
O item V7UE pertence ao questionário escola. É coletado a informação sobre a existência e as condições de uso da quadra esportiva na unidade. São apresentadas quatro categorias: bom; regular; ruim; inexistente. Para o estudo a variável foi categorizada em: possuiu (sim) ou não possuiu (não). Nota-se que 29,2% dos estudantes selecionados para o estudo, são de escolas que não possuem quadra esportiva. Ao analisar o grupo de estudantes com proficiência abaixo do básico, a taxa de estudantes provenientes de escolas sem quadra esportiva aumenta para 48,9%.

Tabela 37 - A unidade possui quadra esportiva? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Quadra esportiva?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Não	1.220	10.156	11.376
Sim	1.274	26.314	27.588
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 38 -Distribuição do item possuir quadra esportiva em cada uma das categorias.



Fonte: Autoria própria (2020).

4.1.3.8 V8UE – A unidade possui laboratório de informática

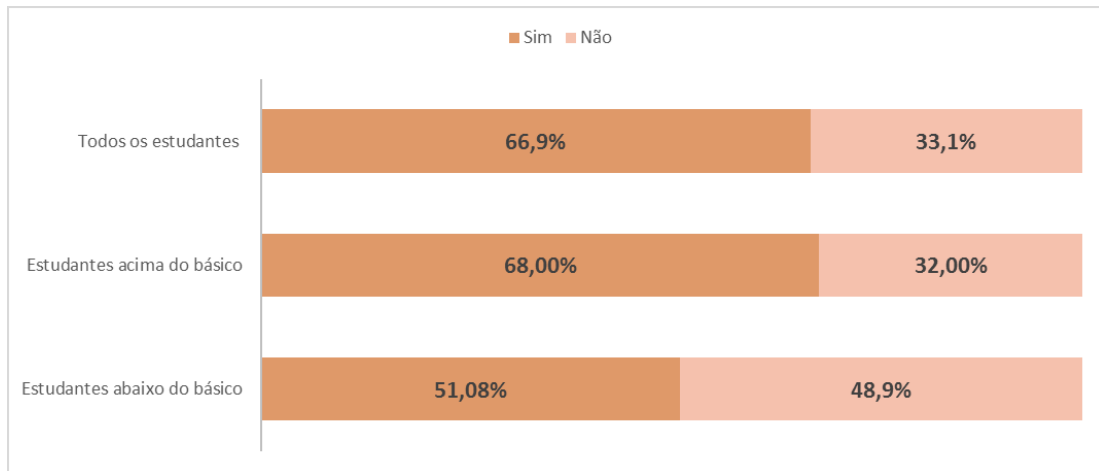
O item V8UE pertence ao questionário escola. É coletado a informação sobre a existência e as condições de uso do laboratório de informática na unidade. São apresentadas quatro categorias: bom; regular; ruim; inexistente. Para o estudo a variável foi categorizada em: possuiu (sim) ou não possui (não). Observa-se que 33,1% dos estudantes selecionados para o estudo, são de escolas que não possuem laboratório de informática. Ao analisar o grupo de estudantes com proficiência abaixo do básico, a taxa de estudantes provenientes de escolas sem laboratório de informática aumenta para 48,9%.

Tabela 38 - A unidade possui laboratório de informática? x Dicotomização das proficiências de Língua Portuguesa – 5º ano.

Laboratório de Informática?	Estudantes com proficiência abaixo do básico	Estudantes com proficiência igual ou acima do básico	Total
Sim	1.274	24.801	26.075
Não	1.220	11.669	12.889
Total	2.494	36.470	38.964

Fonte: Autoria própria (2020).

Figura 39 - Distribuição do item possuir laboratório de informática em cada uma das categorias.



Fonte: A autoria própria (2020).

4.2 MODELO LOGÍSTICO HIERÁRQUICO

Retomando os capítulos anteriores, para encontrar os fatores associados a estudantes com proficiência abaixo do básico em língua portuguesa no 5º ano foram pré-escolhidas 28 variáveis do nível estudante e oito variáveis do nível unidade escolar. Todas variáveis passaram por uma análise descritiva e foram validadas para o passo seguinte: a modelagem logística hierárquica.

Para a execução do modelo foi utilizado o pacote `gamlss.mx` (STASINOPOULOS e RIGBY, 2007) do software estatístico R. O primeiro passo a ser realizado para a modelagem é o ajuste do modelo nulo.

4.2.1 Ajuste do Modelo Nulo

O modelo Nulo (M_0) apresenta a influência de cada nível no modelo. São apresentados o percentual da variância explicada, pelas diferenças entre escolas e pelas diferenças do estudante de uma mesma escola. A Tabela 39 apresenta os resultados do M_0 para a probabilidade de proficiência abaixo do básico em língua portuguesa para estudantes do 5º ano. Para realizar a interpretação dos parâmetros é necessário utilizar a razão de chances (odds ratio) apresentada na Seção 4.2.3.2.

Tabela 39 - Modelo Nulo

Efeito Fixo	Estimativa	Erro Padrão
Intercepto	-2,8468	0,0217
Efeito Aleatório		Variância
Variância do Intercepto	1,291738	
Variância residual	0,8098	

Fonte: Autoria própria (2020).

Interpretando os resultados do modelo M0 verifica-se que um determinado estudante do 5º ano que realizou a prova do Saeb, têm 17,24 vezes mais chance de obter proficiência igual ou acima do básico na prova de língua portuguesa do que obter proficiência abaixo do básico.

4.2.2 Processo de escolha do melhor modelo

Para seleccionar o Modelo de regressão adequado, diversas técnicas estatísticas se fazem necessárias. Dentre elas, podem-se citar os métodos de seleção de variáveis. Optou-se pelo método de seleção Stepwise (passo a passo). A seguir são apresentados na Tabela 40 os resultados obtidos, após aplicação do método de seleção escolhido. Pode-se observar que o modelo M21 obteve o melhor (menor) BIC dentre todos ajustados e, conseqüentemente, é o selecionado para análise no trabalho.

Tabela 40 - Seleção do melhor modelo através do método Stepwise

	Modelo	BIC
M0	$VR \sim 1 + re(random = \sim 1 U.E.)$	26.333,20
M1	$VR \sim (re(random = \sim 1 U.E.)) + V4UE$	23.891,92
M2	$VR \sim (re(random = \sim 1 U.E.)) + V4UE + V25E$	22.750,74
M3	$VR \sim (re(random = \sim 1 U.E.)) + V4UE + V25E + V3UE$	22.376,85
M4	$VR \sim (re(random = \sim 1 U.E.)) + V4UE + V25E + V3UE + V20E$	22.154,10
M5	$VR \sim (re(random = \sim 1 U.E.)) + V4UE + V25E + V3UE + V20E + V17E$	21.980,81
M6	$VR \sim (re(random = \sim 1 U.E.)) + V4UE + V25E + V3UE + V20E + V17E + V27E$	21.807,47
M7	$VR \sim (re(random = \sim 1 U.E.)) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE$	21.663,96

M8	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E$	21.555,78
M9	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E$	21.463,69
M10	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE$	21.384,03
M11	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E$	21.307,84
M12	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E$	21.233,62
M13	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E + V18E$	21.161,07
M14	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E + V18E + V15E$	21.103,79
M15	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E + V18E + V15E + V2E$	21.059,82
M16	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E + V18E + V15E + V2E + V22E$	21.021,37
M17	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E + V18E + V15E + V2E + V22E + V1UE$	21.003,02
M18	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E + V18E + V15E + V2E + V22E + V1UE + V8E$	20.988,90
M19	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E + V18E + V15E + V2E + V22E + V1UE + V8E + V8UE$	20.979,30

M20	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E + V18E + V15E + V2E + V22E + V1UE + V8E + V8UE + V4E$	20.970,06
M21	$VR \sim (\text{re}(\text{random} = \sim 1 \mid \text{U.E.})) + V4UE + V25E + V3UE + V20E + V17E + V27E + V2UE + V5E + V3E + V7UE + V26E + V7E + V18E + V15E + V2E + V22E + V1UE + V8E + V8UE + V4E + V11E$	20.966,85

Fonte: Autoria própria (2020).

4.2.3 O modelo selecionado

Após o processo de seleção, obtêm-se o melhor modelo utilizando o critério do BIC. O modelo M21 é composto por 21 variáveis explicativas, sendo que, seis são do nível escola e 15 do nível do estudante. A Tabela 41 apresenta todas as estimativas, p-valor e a razão de chances para cada covariável inserida no modelo. Observe que coeficientes (estimativas) positivos indicam que se for observada uma determinada característica no estudante, em um determinado fator (característica) em estudo, existe uma maior probabilidade deste estudante ter proficiência abaixo do básico. Analogamente, coeficientes negativos implicam características que diminuem tal probabilidade. Pelo teste estatístico de significância das estimativas dos coeficientes, algumas variáveis não se mostram significativa, no entanto, como citado na Seção 3.2.3.2, devido à grande quantidade de observações e de variáveis, testes estatísticos são afetados, e por consequência optou-se por não excluir nenhuma variável com significância fora do recomendado.

Tabela 41 - Modelo Selecionado: Estimativas, erro padrão, p-valor e razão de chances

Efeito Fixo	Estimativa	Erro Padrão	p-valor	Razão de Chances
Intercepto	-3,621	0,121	< 2e-16 ***	0,03
V2E - Sexo (Masculino)	0,444	0,047	3,02e-08 ***	1,56
V3E - Cor/ Raça (não declarada x parda)	0,128	0,079	0,104289	1,14
V3E - Cor/ Raça (branca x parda)	0,037	0,060	0,542301	1,04
V3E - Cor/ Raça (preta x parda)	0,780	0,067	< 2e-16 ***	2,18

V3E - Cor/ Raça (amarela x parda)	0,224	0,136	0,099672 .	1,25
V3E - Cor/ Raça (indígena x parda)	0,195	0,133	0,142865	1,22
V4E - Faixa etária (abaixo da idade correta x idade correta)	0,280	0,215	0,191982	1,32
V4E - Faixa etária (acima da idade correta x idade correta)	0,152	0,054	0,004696 **	1,16
V5E - Possui computador em casa (não)	0,203	0,051	0,030666 *	1,23
V7E - Escolaridade da mãe (nunca estudou x não completou os anos iniciais)	0,305	0,048	0,030666 *	1,36
V8E - Mãe sabe ler? (não)	0,312	0,093	0,030666 *	1,37
V11E - Os pais participam de reunião pedagógica (não)	0,073	0,054	0,172095	1,08
V15E - Incentivo a não faltar as aulas (não)	0,611	0,091	1,81e-11 ***	1,84
V17E - Lê jornal (nunca/quase nunca)	-0,340	0,048	9,80e-13 ***	0,71
V18E - Lê Livros (nunca/quase nunca)	0,641	0,084	2,86e-14 ***	1,90
V20E - Leitura de quadrinhos (nunca/quase nunca)	0,608	0,057	< 2e-16 ***	1,84
V22E - Frequenta biblioteca (nunca/quase nunca)	-0,391	0,052	8,48e-14 ***	0,68
V25E - Reprovou (sim)	1,109	0,052	< 2e-16 ***	3,03
V26E - Abandonou (Sim)	0,549	0,075	3,48e-13 ***	1,73
V27E - Faz o dever de casa (não/professor não passa)	0,310	0,056	3,02e-08 ***	1,36
V1UE - Dependência Administrativa (Federal/Estadual)	0,164	0,076	0,030666 *	1,18
V2UE - Localização (zona rural)	0,355	0,064	3,02e-08 ***	1,43
V3UE - Formação do docente	-0,007	0,001	0,030666 *	0,99

V4UE -NSE (nível 1 e 2)	0,988	0,061	0,030666 *	2,69
V7UE - A escola possui quadra (não)	0,237	0,051	0,030666 *	1,27
V8UE - Escola possui Laboratório de Informatica (Não)	0,127	0,050	0,000806 ***	1,14
Efeito Aleatório	Variância	Erro Padrão	Razão de Chances	
Intercepto	0,775	0,084	2,17	

Código Significativo : 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fonte: Autoria própria (2020).

4.2.3.1 Qualidade de ajuste do Modelo selecionado

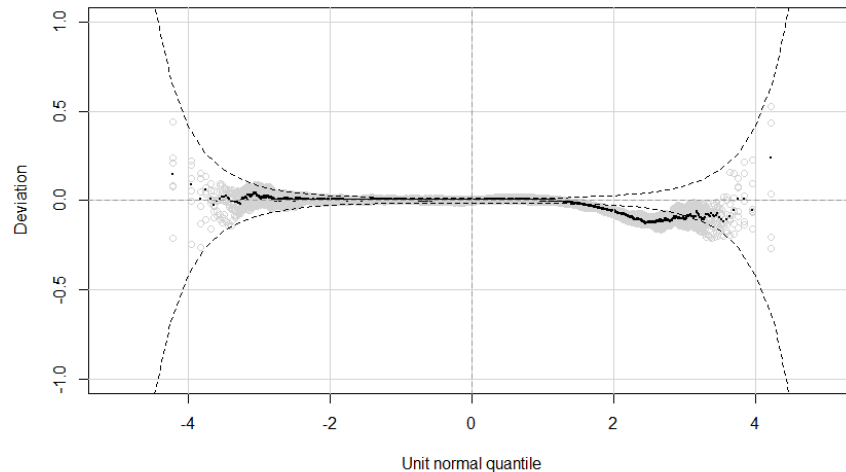
Para avaliar a qualidade do ajuste dos resíduos tem-se a Tabela 41. Observa-se, pelos resultados, que os resíduos apresentam momentos próximos aos de uma distribuição normal padrão, isto é, média próxima a zero, variância próxima a 1 e coeficiente de assimetria e grau de curtose próximos a zero. Para a verificação visual dos resíduos, apresenta-se na Figura 40 o gráfico wormplot (de minhoca), que indica uma pequena irregularidade nos resíduos na parte superior da distribuição. Entretanto, em geral, o modelo proposto é razoável para explicar o comportamento dos dados, uma vez que 95% dos resíduos estão dentro das bandas de confiança apresentadas.

Tabela 42 - Resumo dos quantis aleatórios residuais

Medida	Valor
Média	-0,0013
Variância	0,978
Coeficiente de assimetria	-0,0438
Coeficiente de curtose	-0,0788
Coeficiente de correlação de Filliben	0,9998

Fonte: Autoria própria (2020).

Figura 40 - Worm plot do modelo final ajustado



Fonte: A autoria própria (2020).

4.2.3.2 Interpretação da Razão de Chances

A partir do modelo final ajustado (Tabela 40) pode-se obter as razões de chances de cada uma das variáveis selecionadas pelo modelo, fixando todas as demais, isto é, considerando a mesma situação para todas as outras características. Em todas as variáveis categóricas, foi fixada uma categoria como base (também chamada de basal). É a partir dessa categoria que serão realizadas as comparações. Vale ressaltar, uma vez mais que, por se tratar de uma amostra com muitas observações, devemos ser extremamente cautelosos no que tange à utilização de testes estatísticos para validação das estimativas, especialmente após o processo de seleção do modelo (LEE et al., 2016). Os Quadros 6 e 7 apresentam, respectivamente, as interpretações das razões de chances das variáveis do nível do estudante e do nível das escolas.

Quadro 6 - Interpretações da Razão de Chances - Variáveis do nível do estudante

Variável Selecionada pelo modelo – Nível do Estudante	Interpretação da Razão de Chances
V2E - Sexo do estudante	A partir do modelo proposto, conclui-se que estudante de 5º ano, do sexo masculino, tem 1,56 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao do sexo feminino.

<p>V3E - Cor/ Raça do estudante</p>	<p>Para a análise optou-se em escolher a categoria com maior frequência de estudantes como a ser comparada (referência). No caso, estudantes autodeclarados pardos representam 48,7% do total dos estudantes amostrados.</p> <p>O estudante de 5º ano cuja cor/raça não foi declarada foi mantido no conjunto de dados apenas para que essas observações pudessem ser utilizadas, uma vez que uma característica não observada em um indivíduo o exclui da análise, e não será aqui analisado.</p> <p>Estudantes autodeclarados da cor/raça preta possuem 2,18 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que se autodeclara da cor parda. A mesma comparação é realizada com as raças/cor amarela (1,25), branca (1,04) e indígena (1,22).</p>
<p>V4E - Faixa etária do estudante:</p>	<p>Para a análise optou-se em escolher a categoria com maior frequência de estudante como a ser comparada. No caso, estudantes com faixa etária correta (sem distorção, ou seja, 10 anos de idade) representam a maioria do total dos estudantes amostrados. O estudante de 5º ano que possui faixa etária abaixo da idade correta (8 e 9 anos), tem 1,32 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante sem distorção (10 anos) e estudantes que possuem faixa etária acima da idade correta (11 ou mais anos) possuem 1,16 mais chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante sem distorção etária (10 anos).</p>

V5E - Possui computador em casa:	A partir dos dados obtidos pelo modelo proposto constata-se que um estudante de 5º ano que não possui computador, tem 1,23 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que possui o equipamento.
V7E - Escolaridade da mãe	O estudante de 5º ano, que possui mãe que nunca estudou ou não concluiu o ensino básico, tem 1,36 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que possuem mãe com ensino básico completo e/ou ensino superior.
V8E - Mãe sabe ler	O estudante de 5º ano cuja mãe não sabe ler tem 1,37 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que possui mãe que sabe ler. Uma pesquisa brasileira comprova os dados da análise.
V11E - Os pais participam de reunião pedagógica	O modelo apresenta o seguinte resultado para a participação dos pais em reuniões pedagógicas. O estudante de 5º ano cujos pais não participam de reunião pedagógica, tem 1,08 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que os pais participam das reuniões pedagógicas. Ressalta-se que o valor se encontra próximo de 1 e, por consequência, não existe uma diferença significativa entre a participação dos pais em reunião pedagógica para os grupos estudados. Entretanto, como já mencionado, do ponto de vista estatístico devemos ter bastante cautela em analisar a significância de um coeficiente com amostras muito grandes e após a etapa de seleção do modelo.

V15E - Incentivo a não faltar às aulas	Para esta variável, o modelo proposto aponta que um estudante de 5º ano que não tem incentivo dos pais para não faltar às aulas, tem 1,84 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante cujos pais incentivam sua participação.
V17E – Leitura de jornal	O estudante de 5o ano que nunca ou quase nunca lê jornal, possui 29% menos chance de obter proficiência abaixo do básico em relação ao estudante que lê jornal impresso.
V18E – Leitura de Livros	O estudante de 5º ano que nunca ou quase nunca lê livros tem 1,90 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação aos estudantes que leem livros com grande frequência.
V20E - Leitura de quadrinhos	A variável leitura de quadrinhos encontra-se dentro do construto capital cultural. De acordo com o modelo proposto, estudante de 5º ano, que não lê revista em quadrinho, tem 1,84 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que lê revista em quadrinhos.
V22E – Frequenta de ida a biblioteca	O estudante de 5º ano que nunca ou quase nunca frequenta a biblioteca tem 32% menos chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que frequenta a biblioteca.
V25E – Já reprovou?	A variável reprovação encontra-se dentro do construto trajetória escolar. De acordo com o modelo proposto, o estudante de 5º ano, que já reprovou, tem 3,03 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que nunca reprovou.

V26E - Abandonou	Para o abandono escolar, o modelo proposto aponta que o estudante de 5º ano, que já abandonou a escola em um determinado momento, tem 1,73 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que nunca abandonou a escola.
V27E - Faz o dever de casa ?	O estudante de 5º ano que nunca ou quase nunca faz o dever de casa ou que o professor não passa dever tem 1,36 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante faz o dever de casa.

Fonte: Autoria própria (2020).

Quadro 7 - Interpretações da Razão de Chances - Variáveis do nível da escola

Variável Selecionada pelo modelo – Nível do Estudante	Interpretação da Razão de Chances
V1UE - Dependência Administrativa	O modelo escolhido sugere que um estudante de 5º ano proveniente de uma escola municipal tem 1,18 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante proveniente de escola estadual ou federal.
V2UE - Localização	De acordo com o modelo proposto, o estudante de 5º ano, proveniente de escola situada em zona rural, tem 1,43 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que estuda em escola situada em zona urbana.
V3UE - Formação do docente	De acordo com os dados do modelo proposto, observa-se que a cada 1% a mais observado na proporção de docentes que possuem formação correta para lecionar na unidade escolar em estudo, a chance de um estudante obter proficiência abaixo do básico diminui em 1%

V4UE -NSE	O estudante de 5º ano proveniente de escola que possui nível socioeconômico baixo (nível 1 e 2), possui 2,69 mais chances de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante que estuda em escolas com nível 3, 4 ou 5.
V7UE - A escola possui quadra	O estudante de 5º ano, proveniente de escolas que não possuem quadra esportiva, tem 1,27 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante proveniente de escola com quadra esportiva
V8UE - Escola possui Laboratório de Informática	O estudante de 5º ano, proveniente de escola que não possui laboratório de informática, tem 1,14 vezes maior chance de obter proficiência abaixo do básico em língua portuguesa em relação ao estudante proveniente de escola que possui laboratório de informática.

Fonte: Autoria própria (2020).

4.2.4 Contextualização dos resultados

Após as devidas interpretações da razão de chances para todas as variáveis selecionadas pelo modelo, faz-se necessário a contextualização dos resultados com outros autores e publicações relevantes. Como visto na Subseção 2.3, Palermo, Silva e Novellino (2014) criaram um protocolo de como selecionar as variáveis para trabalhar com fatores associados a partir da revisão de diversos artigos e dissertações. O Quadro 3 apresentado reproduz resumidamente o protocolo com os grupos de variáveis que mais se encaixam para o estudo. Os autores dividem em dois grandes construtos o nível do estudante: aspectos pessoais e aspectos familiares. Já para o nível escola, são outros seis grandes construtos: fatores sociais exógenos, características das escolas, efeitos de composição da escola, clima escolar, perfil do diretor e políticas ou práticas escolares.

4.2.4.1 Aspectos pessoais do estudante

Este grande construto apresenta outros pequenos construtos como características sócio demográficas, habilidades prévias, atitudes em relação à escola e trajetória escolar. A partir dos

resultados constatou-se as principais características (ou chances) dos estudantes que apresentaram proficiência abaixo do básico em língua portuguesa no Saeb 2017. Os tópicos abaixo resumem os achados pelo modelo para os aspectos pessoais do estudante.

- Do sexo masculino (em relação ao sexo feminino).
- Com faixa etária abaixo da correta ou acima da correta para cursar o 5º ano do ensino fundamental (em relação aos que estão sem distorção).
- Que nunca ou quase nunca leem livros (em relação aos que leem com maior frequência).
- Que nunca ou quase nunca leem quadrinhos (em relação aos que leem com maior frequência).
- Que leem jornal impresso (em relação aos que não leem jornal impresso com menor frequência)
- Que nunca ou quase nunca frequenta biblioteca (em relação aos que frequentam com muita frequência à biblioteca).
- Que já reprovou pelo menos uma vez alguma etapa dos anos iniciais do ensino fundamental (em relação aos que nunca reprovaram).
- Que já abandonou pelo menos uma vez alguma etapa dos anos iniciais do ensino fundamental (em relação aos que nunca abandonaram).

Para a variável sexo, estudos apontam que estudantes do sexo masculino possuem melhor rendimento em matemática e estudantes do sexo feminino possuem melhor rendimento em língua portuguesa, como menciona Pereira (2016) em seu trabalho. Ainda, os autores Osti e Martinelli (2014) apresentam em seu trabalho autores que comprovam os achados no presente trabalho.

Uma revisão de estudos conduzida por Liederman, Kantrowitz e Flannery (2005), em que eram investigadas as dificuldades de leitura, também revelou haver um predomínio de meninos em relação às meninas com dificuldades. Assim, pode-se afirmar que são os meninos, de maneira geral, a encabeçarem as estatísticas por apresentarem maiores índices de reprovação, evasão e dificuldades. (OSTI; MARTINELLI, 2014, p.56)

Verificando a variável cor/raça nota-se quase a maioria dos estudantes (48,7%) amostrados se auto declaram com a cor parda, motivo pelo qual foi a utilizada como categoria basal para comparações com as demais raças. Silveira (2012) apresenta uma hipótese para as diferenças entre raças.

O capital cultural surgiu como uma “hipótese” imprescindível para auxiliar na compreensão das desigualdades de desempenho escolar de crianças oriundas dos diversos grupos sociais. (SILVEIRA, 2012, p. 64)

Já Alves e Soares (2002), realizaram cruzamento de variáveis em seu trabalho e identificaram situações socioeconômicas e culturais, conforme trecho abaixo.

[...]os alunos brancos, em comparação com os pardos e negros, vêm de famílias com condições socioeconômicas melhores, têm pais com níveis de instrução mais elevados, têm menos problema de defasagem idade-série e contam com recursos culturais que favorecem o aprendizado. Estas desigualdades se refletem no contexto escolar frequentado por esses alunos e no desempenho escolar: os alunos negros estão em escolas com pior infraestrutura e onde há alta incidência de alunos com problema de distorção idade-série. Vivem, portanto, um contexto menos estimulante e menos propício para as realizações escolares. [...] (ALVES e SOARES, 2002, p.12)

A distorção idade-série é um fator muito preocupante e significativo (de acordo com o modelo proposto) para estudantes que realizaram a avaliação estudada. Os resultados apresentados pelo modelo corroboram com outros estudos, como, por exemplo, Girandi e Orzechowski (2016). Segundo Soares e Satyro (2008):

[...] “é importante ressaltar que, quanto maior a distorção, pior o desempenho escolar. [...] Há uma íntima relação entre a taxa de distorção idade-série e a taxa de repetência, sendo que esta é causa daquela. Embora a relação não seja linear, ela é monótona crescente – quanto maior for a repetência nos anos anteriores ao ano em questão, maior será a defasagem.” (apud GIRANDI e ORZECHOWSKI, 2008, p.10).

No campo das leituras (livros, revistas, jornais impressos), espera-se, a princípio, que estudantes que leem obtenham desempenho superior em língua portuguesa aos que não leem. Segundo um trecho das diretrizes curriculares da educação básica do estado do Paraná

[...] Mesmo vivendo numa época denominada “era da informação”, a qual possibilita acesso rápido à leitura de uma gama imensurável de informações, convivemos com o índice crescente de analfabetismo funcional, e os resultados das avaliações educacionais revelam baixo desempenho do aluno em relação a compreensão de textos [...] (SEED, 2014 p.48)

Porém, identifica-se uma situação curiosa. De acordo com os dados dos estudantes amostrados, proporcionalmente, analisando os dados antes do modelo, somente 36,7% dos estudantes com proficiência acima do básico realizava leitura de jornais com grande frequência, ao passo que para o grupo dos estudantes com proficiência abaixo do básico este valor era um pouco superior (44,71%). Após a modelagem identificou-se que quem não lê jornal com frequência tem 29% menos chance de obter proficiência abaixo do básico do que os que leem. Revisitando a literatura, nenhuma constatação similar foi identificada. Uma possível justificativa para tal achado poderia ser que, possivelmente, há um entendimento equivocado

do estudante em relação à essa pergunta em específico, uma vez que a leitura de jornal impresso nos dias atuais é muito menos frequente do que no ano de criação do questionário utilizado, em 2005. Assim, uma possível readaptação ou atualização desta pergunta poderia ser considerada.

Nos demais campos de leitura, os resultados apareceram como os verificados na literatura existente. Para leitura em quadrinhos, os autores Perrelli e Stryer (2012), relatam que tal leitura ajuda os estudantes a melhorar regras da língua portuguesa e ajuda no reconhecimento de expressões. Ademais,

Alunos que leem gibis têm melhor desempenho escolar do que aqueles que usam apenas o livro didático – entre os estudantes da rede pública, a HQ aumenta significativamente o desempenho do aluno: entre os que acompanham quadrinhos, o percentual das notas nas provas aplicadas foi de 17,1%, contra 9,9% dos que não leem. (CARVALHO, 2006, p. 38)

Na frequência a ida à biblioteca, a princípio, esperava-se que o modelo apontasse que estudantes que frequentam a biblioteca com maior frequência possuíssem uma menor probabilidade ou chance de obter proficiência abaixo do básico em comparação aos estudantes que não frequentam a biblioteca. Contudo, o modelo obtido na Tabela 40 sugere exatamente o oposto. Uma possível justificativa prática que pode ser levantada para ilustrar tal resultado, seria que a ida a biblioteca não está, necessariamente, atrelada a leitura de livros, revistas, jornais ou estudo. Para a faixa etária utilizada para o estudo, são diversas formas de atividades que podem ser exercidas como atividades lúdicas, contos ou até mesmo momentos sem nenhuma atividade.

Segundo dados do censo escolar, em 2018, das 180 mil escolas brasileiras, aproximadamente 98 mil não tem bibliotecas escolar ou sala de leitura (INEP, 2018).

Do ponto de vista estatístico, uma possível justificativa seria algum tipo de correlação parcial entre variáveis (como uma espécie de confundimento latente) que também foram incluídas no modelo final. Este tipo de comportamento não é incomum na literatura referente à modelagem estatística, como pode ser visto, por exemplo, em Nakamura et al. (2019).

No construto trajetória escolar encontram-se três variáveis estudadas neste trabalho: reprovação, abandono e realização de deveres escolares. Diversos estudos referentes a fatores associados à proficiência do estudante apontam a reprovação do estudante em sua trajetória, como um dos fatores associados à proficiência baixa.

[...]. No que se refere aos aspectos pessoais dos alunos que afetam seu desempenho, uma primeira dimensão relevante foi a sua trajetória escolar. Dessa forma, se o aluno

já possuía um histórico de repetência, com mais de uma reprovação ou muitos abandonos, por exemplo, sua proficiência, como de fato era esperado, em termos de resultados a serem preditos, é menor, da mesma forma que observam os estudos de Albernaz et al.(2002), Franco et al. (2007), Machado et al. (2008) e Ribeiro e Koslinski (2010). (PALERMO; SILVA e NOVELLINO, 2014, pg. 368)

No trabalho de Barbeta, Andrade e Tavares (2018) é mencionado que a reprovação é um fator bastante estudado em avaliação de larga escola e possui um efeito negativo para aqueles estudantes que possuem reprovação em relação aos que não possuem.

Soares (2011), em seu trabalho apresenta várias relações entre o dever de casa e as avaliações de aprendizagens, como, no trecho

[...] assim, pode-se afirmar que as lições de casa vêm ganhando destaque nos últimos anos, ao mesmo tempo em que se fortalecem as políticas de incentivo ao envolvimento dos pais na vida estudantil de seus filhos como prática eficaz para a obtenção de um melhor desempenho escolar. (SOARES,2011, pg. 34)

4.2.4.2 Aspectos familiares do estudante

Neste grande construto apresentam-se outros pequenos construtos como a estrutura e arranjos familiares, o capital socioeconômico, o capital cultural familiar e o capital social familiar. A partir dos resultados constatou-se as principais características (ou chances) dos estudantes que apresentaram proficiência abaixo do básico em língua portuguesa no Saeb 2017. Os tópicos abaixo resumem os achados pelo modelo para os aspectos familiares do estudante:

- Mãe não sabe ler (em relação a mãe do estudante saber ler).
- Mães que nunca estudaram ou não concluíram o ensino fundamental (em relação as que concluíram o ensino fundamental, básico ou superior)
- Os pais participarem frequentemente de reuniões pedagógica (em relação aos que participam com pouco frequência de reuniões pedagógicas).
- Pais que não incentivam a frequentar as aulas (em relação a pais que incentivam a frequentar as aulas).
- Que não possui computadores em casa (em relação aos que possuem computador em casa).

Vivemos a era da tecnologia onde praticamente tudo está interligado por meio da internet. O aprendizado vem se adaptando ao longo dos anos para uma nova era. Antes tudo estático, agora ganham movimentos, formas e diversas maneiras. O computador se tornou uma das principais formas de aprendizado, desenvolvendo o raciocínio lógico. Os dados mostram

que estudantes que não possuem computador em casa, possuem maior chance de obter proficiência abaixo do básico em relação aos que possuem.

Em um artigo para o portal brasil escola é mencionado um parecer pedagógico para o aprendizado através do computador.

[...]a criança dessa faixa etária encontra-se na fase de latência na teoria freudiana, esta é a idade do domínio versus inferioridade, que vai dos seis aos doze anos. A principal realização deste estágio de aprendizagem das habilidades tanto na escola quanto fora dela. (Junior, Portal Brasil Escola)

Ao analisar a escolaridade da mãe e se a mãe sabe ler, percebe-se que 73,5% dos estudantes amostrados abaixo do básico possuem mães que nunca estudaram ou não concluíram o ensino básico. Segundo Rios-Neto, Cezar e Riani (2002) “Analisando a progressão por série, por meio de um modelo de regressão logística hierárquico com dados da Pesquisa Nacional por Amostra de Domicílios (Pnad), verificaram que a escolaridade da mãe e a educação média do professor têm um efeito positivo na probabilidade de progressão escolar” (apud PAULA, FRANCO e SILVA, 2018, pg.891).

O incentivo dos pais a não faltar a escola, aparece como um dos importantes itens do grupo de aspectos familiares. A maioria dos estudantes amostrados (97,2%) apresentam incentivo dos pais a não faltar na escola. Segundo o modelo, estudantes que não possuem incentivo dos pais tem 1,84 vezes mais chance de obter proficiência abaixo do básico. A revista Brasil Escola (2010), menciona um importante fala de Ferreira e Barrera (2010) reforçando a constatação obtida no modelo proposto.

Segundo Ferreira e Barrera (2010) quando não há valorização dos familiares pelo processo estudantil, os estudantes tendem a não valorizar também o espaço escolar. O mesmo autor acrescenta que a ausência da família é motivo de desânimo e não interesse por parte dos alunos. Com isso, pode-se dizer que a ausência familiar pode ser um dos fatores que causam vários prejuízos no contexto escolar dos estudantes. (BRASIL ESCOLA, 2010)

4.2.4.3 Efeitos das características das escolas e composição da escola

Somente dois dos seis grandes construtos do efeito escola foram contemplados com as variáveis no estudo: características das escolas e composição da escola. Onde também se destaca a composição de alunos e composição de professores inseridos dentro do construto composição da escola. A partir dos resultados constatou-se as principais características (ou chances) dos estudantes que apresentaram proficiência abaixo do básico em língua portuguesa

no Saeb 2017. Os tópicos abaixo resumem os achados pelo modelo para os aspectos familiares do estudante:

- Estudantes provenientes de escolas municipais (em relação aos provenientes de escolas estaduais ou federais).
- Estudantes provenientes da zona rural (em relação a estudantes provenientes da zona urbana).
- Estudantes provenientes de escolas que possuem nível socioeconômico baixo (em relação aos provenientes de escolas com nível socioeconômico intermediário ou alto).
- Estudantes provenientes de escolas que possuem quadra esportiva (em relação aos estudantes provenientes de escolas que não possuem quadra esportiva).
- Estudantes provenientes de escolas que possuem laboratório de informática (em relação aos estudantes provenientes de escolas que não possuem laboratório de informática)

Conforme o censo escolar 2018, a tecnologia não está acessível aos estudantes em cerca da metade das escolas de ensino fundamental. Apenas 46,8% das escolas dispõem de laboratório de informática; 65,6% têm acesso a internet.

Considerando uma maior diversidade de redes municipais em comparação com as redes estaduais é difícil chegar algum tipo de conclusão sem realizar um estudo prévio da oferta, da demanda e principalmente da infraestrutura das redes estudadas.

Estudo realizado pelo INEP (2004), aponta que o desempenho de estudantes de área urbana é 29,2 pontos maior do que a dos estudantes de área rural na prova de leitura do Saeb. Assim, pode-se afirmar que, mesmo após mais de uma década, esta diferença ainda é significativa.

5 CONSIDERAÇÕES FINAIS

O presente trabalho apresentou a proposição de elencar fatores associados à proficiência abaixo do básico em língua portuguesa em estudantes do 5º ano do ensino fundamental que realizaram a avaliação Saeb 2017.

5.1 OBSERVAÇÕES GERAIS

O foco em língua portuguesa deve-se ao fato de que, ao analisar o questionário do estudante, percebe-se muitos itens focados para captar a habilidade de leitura do estudante como leitura de quadrinhos, leitura de livros, leitura de revistas, leitura de jornais entre outros itens.

Ao tratar do tema fatores associados ao desempenho, sejam eles em avaliações nacionais em larga escala (como no caso dessa dissertação) ou em contextos específicos (como em avaliações locais), é importante ressaltar que para ter uma boa qualidade de análise e de resultados, faz-se necessário planejar instrumentos que capturem coerentemente as características dos estudantes. Os questionários devem estar bem alinhados com a capacidade de entendimentos dos estudantes (principalmente a faixa etária estudada) e também atualizado com a realidade. O questionário apresentado para os estudantes na avaliação Saeb 2017, apresenta alguns itens que, possivelmente, não fazem mais parte da realidade da maioria dos brasileiros, como questionamentos sobre leitura de jornal impresso ou itens relacionados a posse de bens (como ter vídeo cassete/DVD ou rádio). Como sugestão, estes itens, particularmente, poderiam ser revistos para melhor adaptação à nova realidade.

Neste trabalho apresentou-se como motivação o conceito de equidade. Espera-se que ao analisar os fatores associados para o grupo de estudantes com proficiência abaixo do básico, unidades escolares, redes educacionais apliquem na prática ações equânimes a fim de reduzir ou minimizar as diferenças obtidas nas mais diversas variáveis estudadas. Como por exemplo, pode-se citar o trabalho que a Rede Municipal de Educação de Curitiba vem desempenhando. A partir do Programa Transformando Realidades – equidade na educação são analisadas variáveis sociais, pedagógicas e financeiras das unidades que compõe a rede, do entorno das unidades e dos estudantes e, posteriormente, as escolas mais vulneráveis ganham uma ajuda extra financeira e pedagógica. São princípios equânimes sendo aplicados na prática.

Ao trabalhar com banco de dados muito grandes (como é o caso dos microdados do Saeb 2017), é comum aparecer várias dificuldades técnicas e muito dados faltantes. Para a obtenção da amostra, optou-se excluir da base os estudantes que não responderam, pelo menos uma, dos itens a serem estudados como possíveis candidatos para explicar a proficiência abaixo

do básico, com isso, vários estudantes foram excluídos da base antes de realizar o sorteio para amostragem. Uma das ressalvas ao se realizar a exclusão de observações em trabalhos cujo foco é o desempenho escolar, é o fato de que o estudante que não conseguiu responder perguntas (parcial ou integralmente) poderia ser objeto de estudo pelo fato de não conseguir compreender, em sua totalidade, o questionário. Como um dos objetivos da dissertação era trabalhar com modelagem estatística, mais especificamente com modelos estatísticos de predição, um dos pressupostos é ter todas variáveis respondidas integralmente. Além da exclusão das observações (no caso, o estudante), existem diversas técnicas estatísticas e de banco de dados para imputação que podem suprimir os dados faltantes, como, por exemplo, interpolação, substituição de dados faltantes pela média, mediana ou até mesmo por regressão linear. No trabalho, entretanto, optou-se por excluir as observações onde alguma pergunta não estivesse respondida uma vez que a amostra resultante utilizada possuía um grande número de respondentes (38.964 estudantes).

Outro cuidado bastante importante foi a categorização das variáveis explicativas. Para este tipo de procedimento, é necessário a execução de uma análise descritiva das variáveis. Os possíveis problemas que podem aparecer no resultado final do modelo, podem, em diversas vezes, ser verificados em uma simples análise descritiva dos dados. Neste trabalho, uma extensa análise foi realizada, onde optou-se por confrontar, cada uma das variáveis explicativas com as categorias da variável resposta e, assim, vários resultados após a execução do modelo puderam ser ilustrados com esta análise prévia.

Para modelagem estatística, optou-se por utilizar o modelo de regressão logística hierárquica com foco em dois níveis: estudantes e escolas. O modelo selecionado é composto por 21 das 36 variáveis explicativas pré-selecionadas antes do procedimento. Como pôde ser visto por meio da análise de diagnóstico, apesar de apenas pouco mais da metade das variáveis serem utilizadas (58,33%), o modelo ajustado foi razoável para explicar o comportamento dos dados. Destaca-se aqui a importância do procedimento utilizado para o processo de sua seleção, escolhido devido ao grande número de observações e ao elevado número de variáveis. Mais importante, a maioria dos resultados encontrados corroboram com a literatura especializada e alguns aspectos detectados pelo modelo demandam uma atenção especial, principalmente, os resultados que são, de determinada maneira, contra intuitivos, como, por exemplo, o que aponta uma maior chance de o estudante de 5º ano obter proficiência abaixo do básico se possuir uma maior frequência de leitura de jornais. Esses resultados apontam, possivelmente, a necessidade

da readequação de alguns itens do questionário aplicado. Os quadros 8, 9 e 10 resumem os achados mais importantes neste trabalho por ordem de relevância.

Quadro 8 - Resultados para os aspectos pessoais do estudante

OR	Categoria Referência	Categoria comparada
3,03	Reprovou pelo menos uma vez	Nunca reprovou
1,90	Nunca ou quase nunca lê livros	Sempre ou quase sempre lê livro livros
1,84	Nunca ou quase nunca lê quadrinhos	Sempre ou quase sempre lê quadrinhos
1,73	Já abandonou pelo menos uma vez alguma etapa dos anos iniciais	Nunca abandonou alguma etapa dos anos iniciais
1,56	Sexo Masculino	Sexo Feminino
1,32	Abaixo da idade correta (5º ano)	sem distorção de idade (5º ano)
1,16	acima da idade correta (5º ano)	sem distorção de idade (5º ano)
29% a menos de chance	Nunca ou quase nunca frequenta biblioteca escolar	Sempre ou quase sempre frequenta a biblioteca escolar
32% a menos de chance	Que nunca ou quase nunca lê jornal impresso	Que leem jornal impresso

Fonte: Aatoria própria (2020).

Quadro 9 - Resultados para os aspectos familiares do estudante

OR	Categoria Referência	Categoria comparada
1,84	Pais que não incentivam a frequentar as aulas	Pais que sempre ou quase sempre incentivam a frequentar as aulas
1,37	Mãe não sabe ler/escrever	Mãe sabe ler/escrever
1,36	Mãe nunca estudou ou não concluiu o ensino fundamental	Mãe conduiu o ensino fundamental, médio ou superior
1,23	Não possui computador em casa	Possui computador em casa
1,08	Pais participam frequentemente de reuniões pedagógicas	Pais participam com pouca frequência de reuniões pedagógicas

Fonte: Aatoria própria (2020).

Quadro 10 - Resultados para os efeitos das características das escolas e composição da escola

OR	Categoria Referência	Categoria comparada
2,69	Estudante oriundo de escolas com nível socioeconômico baixo	Estudante oriundo de escolas com nível socioeconômico intermediário ou alto
1,43	Estudante oriundo da zona rural	Estudante oriundo da zona urbana
1,27	Estudante oriundo de escola sem quadra esportiva	Estudante oriundo de escola com quadra esportiva
1,18	Estudante oriundo de escola municipal	Estudante oriundo de escola estadual ou federal
1,14	Estudante oriundo de escola que não possui laboratório de informática	Estudante oriundo de escola que possui laboratório de informática

Fonte: Aatoria própria (2020).

A partir dos achados, várias discussões pedagógicas e sociais podem ser inseridas para ilustrar os achados do modelo. Apresento um breve relato sobre as 3 variáveis que foram destaques em seus grupos: reprovação, incentivo dos pais a não faltar as aulas e nível socioeconômico das escolas.

Muitos estudiosos e profissionais da educação acreditam que a reprovação escolar não contribui para aprendizagem, inclusive apontam que a reprovação apresenta um efeito negativo. Diversas rede de ensino ofertam ensino ciclado, no qual o estudante é caminha sem retenção por alguns períodos e no final de cada ciclo é verificado o cumulativo dos anos. Uma abordagem que pode ser uma alternativa para possíveis desistências e desmotivação dos estudantes.

Com uma concorrência com jogos on-line e redes sociais, o incentivo dos pais a não faltar as aulas cada vez se torna muito importante para crianças. As crianças que declaram que não possuem incentivo dos pais para frequentar a escola apresentam 84% mais chance de possuir proficiência abaixo do básico em língua portuguesa. É importante destacar a ausência dos pais no papel de cobrar resultados dos filhos. Na maioria dos casos, são estudantes que são desassistidos pela família ou existe a falta feedback entre pais e filhos.

O nível socioeconômico escolar baixo é reflexo do baixo poder aquisitivo dos estudantes que estão matriculados na mesma. Para o cálculo do indicador são coletadas informações dos bens que as famílias dos estudantes possuem em suas residências. Escolas com essa realidade, geralmente, estão situados em locais de grande vulnerabilidade social e com grande índice de violência e pobreza em seu entorno.

5.2 CONTRIBUIÇÕES DO TRABALHO E SUGESTÕES PARA TRABALHOS FUTUROS

Na semana em que se conclui a presente dissertação, o INEP/MEC divulgou o resultado de mais uma edição do Saeb e o IDEB (edição 2019). Novos trabalhos atrelados a fatores associados ao desempenho do estudante estarão por vir, porém espera-se que gestores das redes de ensino ou de unidades coloquem em prática ações equânimes que possam resgatar estudantes que não possuem desempenho adequado para a etapa.

Como sugestões para trabalhos futuros:

- 1). Pode-se pensar em um terceiro nível (turma do estudante) ou então considerar três diferentes categorias para a variável resposta: 0, para nenhuma proficiência abaixo do básico; 1, para pelo menos uma proficiência abaixo do básico (em língua portuguesa ou em matemática) e 2, para duas proficiências abaixo do básico (língua portuguesa e matemática). Neste último caso, um Modelo de regressão multinomial seria apropriado.

- 2). A criação de índices, por meio da análise fatorial ou pela TRI, agrupando variáveis de construtos parecidos, poderia ser realizada afim de reduzir o número de variáveis explicativas no modelo.
- 3). Adicionar um 3º nível – Turma do estudante.
- 4). Aplicação do modelo nas redes municipais, estaduais ou federais de ensino para verificar resultados específicos.

O desafio que já era grande, torna-se muito maior: 2020, o ano em que o mundo parou, as escolas fecharam, o ensino tornou-se distante (online para uma grande parte ou inexistente para outra grande parte). A pandemia da COVID-19 e suas consequências, na educação ainda estão incalculáveis. No presente momento, o ensino sem retorno presencial para grande parte dos estudantes do ensino básico, refletirá drasticamente nos resultados das próximas avaliações de aprendizado. O momento é de muita apreensão, mas também de cautela. Trabalhos futuros provavelmente abordarão reflexões sobre o impacto da pandemia no aprendizado e por consequência novos fatores associados a proficiência do estudante.

REFERÊNCIAS

- AKAIKE, Hirotugu. A new look at the statistical model identification. **IEEE transactions on automatic control**, v. 19, n. 6, p. 716-723, 1974.
- ALBERNAZ, Ângela; FERREIRA, Francisco HG; FRANCO, Creso. **Qualidade e equidade na educação fundamental brasileira**. Texto para discussão, 2002.
- ALMEIDA, Marcos Felipe Camarinha (2013). Relação entre Ritmo Circadiano, Turno e Rendimento Escolar de Alunos do Ensino Fundamental. *Revista de Neurociência*, 21 (2), 171-172.
- ALVES, Maria Teresa Gonzaga; SOARES, José Francisco. Raça e desempenho escolar: as evidências do Sistema de Avaliação da Educação Básica-SAEB. **XXVI Encontro Anual da ANPOCS. Caxambu**, 2002.
- ALVES, Marleide F.; LOTUFO, Anna Diva P.; LOPES, Mara Lúcia M. Seleção de variáveis stepwise aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 1, n. 1, 2013.
- ANDRADE, D. F., TAVARES, H. R., VALLE, R. C. 2000. Teoria de Resposta ao Item: conceitos e aplicações. ABE — Associação Brasileira de Estatística, 4º SINAPE, 2000.
- ANDRADE, Doherty, BRANDÃO, Bernadete Maria Suaki, MARTINS, Ana Beatriz Tozzo. Características comuns entre alunos com baixo desempenho em matemática na prova brasil de 2011: um estudo logístico. **Revista Paranaense de Educação Matemática**, v. 6, n. 10, 2017.
- BARBETTA, Pedro Alberto; DE ANDRADE, Dalton Francisco; TAVARES, Heliton R. Estudo de fatores associados através de regressão quantílica hierárquica. **Estudos em Avaliação Educacional**, v. 29, n. 71, p. 320-349, 2018.
- BERK, Kenneth N. Comparing subset regression procedures. **Technometrics**, v. 20, n. 1, p. 1-6, 1978.
- BLAND, J. M.; ALTMAN, D. G. Cronbach's alpha. *British Medical Journal*. p. 314-572. 1997.
- BOLFARINE, Heleno; DE OLIVEIRA BUSSAB, Wilton. **Elementos de amostragem**. Editora Blucher, 2005.
- BRASIL. [Constituição (1988)]. Constituição da República Federativa do Brasil de 1988. Brasília, DF: Presidência da República, [2016]. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 11 jan. 2020.
- BRASIL. Decreto nº 9.432, de 29 de junho de 2018. Regulamenta a Política Nacional de Avaliação e Exames da Educação Básica. Disponível em: http://download.inep.gov.br/educacao_basica/saeb/2018/legislacao/decreto_n_9432_2906201

8_regulamenta_politica_nacional_de_avaliacao_e_exames_da_educacao_basica.pdf. Acesso em: 20 ago. 2020.

BRASIL. Lei de Diretrizes e Bases da Educação Nacional. Lei número 9394, 20 de dezembro de 1996.

BRASIL. Lei n.º 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação – PNE e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/113005.htm. Acesso em: 22 jul 2020.

BRASIL. Ministério da Educação. Portaria nº 482, de 7 de junho de 2013. Dispõe sobre o Sistema de Avaliação da Educação Básica – Saeb. Diário Oficial da União, Brasília, DF, 10 jun. 2013. Seção 1, p. 17. Disponível em: http://www.in.gov.br/materia//asset_publisher/Kujrw0TZC2Mb/content/id/30677986/do1-2013-06-10-portaria-n-482-de-7-de-junho-de-2013-30677978. Acesso em: 23 maio 2020.

BRASIL. Portaria nº 931 de 21 de março de 2005. Institui o Sistema de Avaliação da Educação Básica-Saeb, que será composto por dois processos de avaliação: a Avaliação Nacional da Educação Básica-Aneb, e a Avaliação Nacional do Rendimento Escolar-ANRESC. Diário Oficial da União, 2005.

BRASIL. Portaria nº 447 de 24 de maio de 2017. Estabelece diretrizes para o planejamento e a operacionalização do Sistema de Avaliação da Educação Básica (Saeb) no ano de 2017. Diário Oficial da União, 2017a.

BRASIL. Base Nacional Comum Curricular: Educação é a base. Terceira versão. Ministério da Educação: Brasil, 2017b.

BUUREN, Stef van; FREDRIKS, Miranda. Worm plot: a simple diagnostic device for modelling growth reference curves. **Statistics in medicine**, v. 20, n. 8, p. 1259-1277, 2001.

CARVALHO, D. A educação está no gibi. São Paulo: Papyrus, 2006 p. 38.

CASTRO FILHO, C. M. Agenda 2030 para o desenvolvimento sustentável: uma leitura de política pública na clave da biblioteca escolar. **Revista Digital de Biblioteconomia & Ciência da Informação**, n. 3, v. 16, p. 355-372, 2018. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8650931>. Acesso em: 29 maio 2019.

COLEMAN, James S. et al. York. 1966. Equality of educational opportunity, v. 2, 1966.

CORDEIRO, Gauss Moutinho; DEMÉTRIO, Clarice GB. Modelos lineares generalizados e extensões. **Piracicaba: USP**, 2008.

DE PAULA, Josiane Souza; DE PAIVA FRANCO, Ana Maria; DA SILVA, José Waldemar. Fatores relacionados ao atraso escolar no estado de Minas Gerais. **Estudos em Avaliação Educacional**, v. 29, n. 72, p. 886-917, 2018.

DUNN, Peter K.; SMYTH, Gordon K. Randomized quantile residuals. **Journal of Computational and Graphical Statistics**, v. 5, n. 3, p. 236-244, 1996.

ELIAN, S.N. Análise de regressão. São Paulo: USP / Dep. Estatística, 1988. p.232.

EFROYMSON, M. A. Multiple regression analysis. **Mathematical methods for digital computers**, p. 191-203, 1960.

ERIKSON, Erik. Infância e Sociedade [1963]. 1971.

FARIA, Ernesto Martins; DE MEIRELES GUIMARÃES, Raquel Rangel. Excelência com equidade: fatores escolares para o sucesso educacional em circunstâncias desfavoráveis. **Estudos em Avaliação Educacional**, v. 26, n. 61, p. 192-215, 2015.

FERRÃO, Maria Eugénia. **Introdução aos modelos de regressão multinível em educação**. Komed, 2003.

FIGUEIRA, Cleonis Viater. Modelos de regressão logística. 2006.

FRANCO, Creso et al. Qualidade e equidade em educação: reconsiderando o significado de "fatores intra-escolares". **Ensaio: avaliação e políticas públicas em educação**, v. 15, n. 55, p. 277-298, 2007.

FREITAS, Luiz Carlos de; et al. Avaliação educacional: caminhando pela contramão. 5. ed. Petrópolis: Vozes, 2013.

FUNDAÇÃO LEMANN; ITAÚ BBA. Excelência com equidade: as lições de escolas que oferecem um ensino de qualidade aos alunos com baixo nível socioeconômico. São Paulo: Fundação Lemann, Itaú BBA, 2012.

GABRIELLE, A. et al. Fatores associados ao desempenho escolar: uma análise da proficiência em matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro. *Revista Brasileira de Estudos de População*, v. 31, n. 2, p. 367-394, 2014.

GALVÃO, T. F.; PEREIRA, M. G. Revisões sistemáticas da literatura: passos para sua elaboração. Brasília, 2014. Disponível em: <http://www.scielo.br/pdf/ress/v23n1/2237-9622-ress-23-01-00183.pdf>. Acesso em: 07 maio 2018.

GIOLO, S.R. Introdução a Análise de Dados Categóricos com Aplicações: 1.ed. Blucher - Projeto Fisher ABE, 2017

GOLDSTEIN, H. (1995). **Multilevel Statistical Models** (2nd. Ed.). London: Arnold.

HOSMER JR, David W.; LEMESHOW, Stanley; STURDIVANT, Rodney X. **Applied logistic regression**. John Wiley & Sons, 2013.

HOX, Joop J.; MOERBEEK, Mirjam; VAN DE SCHOOT, Rens. **Multilevel analysis: Techniques and applications**. Routledge, 2010.

INEP. Descrição dos Níveis da Escala de Desempenho de Língua Portuguesa – Saeb: 5º e 9º. Ano do Ensino Fundamental. 2011a. Disponível em: http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/escala/2011/escala_desempenho_portugues_fundamental.pdf. Acesso em: 22 jul. 2020.

INEP. Descrição dos Níveis da Escala de Desempenho de Língua Portuguesa – Saeb: 5º e 9º. Ano do Ensino Fundamental. 2011b p. 1. Disponível em: http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/escala/2011/escala_desempenho_portugues_fundamental.pdf. Acesso em: 22 jul. 2020.

INEP. Resultados. 2017. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb/resultados>. Acesso em: 23 jul. de 2020.

INEP. Matrizes de Escalas. 2019. Disponível em: <http://portal.inep.gov.br/web/guest/educacao-basica/saeb/matrizes-e-escalas>. Acesso em: 23 jul. de 2020.

INEP. Histórico Saeb. 2020a Disponível em: <http://inep.gov.br/educacao-basica/saeb/historico> Acesso em: 23 jul. 2020.

INEP. Saeb. 2020b. Disponível em: <http://portal.inep.gov.br/web/guest/educacao-basica/saeb>. Acesso em: 23 jul. de 2020.

KISH, L. Survey sampling. 2ªed. New York: John Wiley & Sons, Inc., 1967.

KULESSA, Erika. Práticas de escrita nas aulas de Sociologia: implicações para o processo de apropriação da linguagem sociológica. Em Debate, n. 14, p. 82-100, 2015.

LAROS, Jacob Arie; MARCIANO, João Luiz; ANDRADE, Josemberg Moura de. Fatores associados ao desempenho escolar em Português: um estudo multinível por regiões. **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 20, n. 77, p. 623-646, 2012.

LIMA, Letícia Alencar; DE SOUSA MOTA, Vanêssa; LEAL, Willany Palhares Palhares. Políticas públicas e gestão escolar: equidade na educação básica. **Humanidades & Inovação**, v. 6, n. 7, p. 78-87, 2019a.

LIMA, Letícia Alencar; DE SOUSA MOTA, Vanêssa; LEAL, Willany Palhares Palhares. Políticas públicas e gestão escolar: equidade na educação básica. **Humanidades & Inovação**, v. 6, n. 7, p. 78-87, 2019b p. 83.

LIEDERMAN, Jacqueline; KANTROWITZ, Lore; FLANNERY, Kathleen. Male vulnerability to reading disability is not likely to be a myth: A call for new data. **Journal of learning disabilities**, v. 38, n. 2, p. 109-129, 2005.

LORD, Frederic M. **Applications of item response theory to practical testing problems**. Routledge, 1980.

MAGALHÃES, M. N.; LIMA, A. C. P. Noções de probabilidade e estatística. 7ª edição 3ª reimpressão revista. **São Paulo**, 2015.

OSTI, Andréia; MARTINELLI, Selma de Cássia. Desempenho escolar: análise comparativa em função do sexo e percepção dos estudantes. **Educação e pesquisa**, v. 40, n. 1, p. 49-59, 2014.

MACHADO, A. F.; MORO, S.; MARTINS, L.; RIOS, J. Qualidade do ensino em matemática: determinantes do desempenho de alunos em escolas públicas estaduais mineiras. **Revista Economia**, v. 9, n. 1, jan.-abr. 2008.

MUNDO, Transformando Nosso. a Agenda 2030 para o Desenvolvimento Sustentável. Recuperado em, v. 15, 2016.

MUNDO, Transformando Nosso. a Agenda 2030 para o Desenvolvimento Sustentável. Recuperado em, v. 15, 2016, p. 24 e 25.

NAKAMURA, LUIZ R.; CERQUEIRA, PEDRO H. R.; RAMIRES, THIAGO G.; PESCIM, RODRIGO R.; RIGBY, R. A.; STASINOPOULOS, DIMITRIOS M. A new continuous distribution on the unit interval applied to modelling the points ratio of football teams. **JOURNAL OF APPLIED STATISTICS**, v. 46, p. 416-431, 2019.

NASCIMENTO, P. A. M. M. Fatores associados ao desempenho escolar: resultados de um modelo multinível de valor adicionado. **Estudos em Avaliação Educacional**, v. 23, n. 51, p. 16-39, jan./abr. 2012

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models-JR Statist. Soc. A, 135: 370-384. Nelder370135J. R. **Statist. Soc A**, v. 1972, 1972.

ONU. Organização das Nações Unidas no Brasil. Documentos Temáticos: Objetivos de Desenvolvimento Sustentável 1 · 2 · 3 · 5 · 9 · 14. Brasília: **ONUBR**, 2017. Disponível em: <http://www.br.undp.org/content/dam/brazil/docs/publicacoes/documentos-tematicos-ods-07-2017.pdf>>. Acesso em: 12 jun. 2018.

OSHIRO, Cláudia Hiromi; SCORZAFAVE, Luiz Guilherme. Efeito do pagamento de bônus aos professores sobre a proficiência escolar no Estado de São Paulo. **Anais do Encontro Nacional de Economia**, v. 39, 2011.

OSTI, Andréia; MARTINELLI, Selma de Cássia. Leitura: Desempenho escolar: análise comparativa em função do sexo e percepção dos estudantes. *Educ. Pesqui.*, São Paulo, v. 40, n. 1, p. 49-59, jan./mar. 2014

PALERMO, Gabrielle A.; SILVA, Denise Britz do Nascimento; NOVELLINO, Maria Salet Ferreira. Fatores associados ao desempenho escolar: uma análise da proficiência em matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro. **Revista Brasileira de Estudos de População**, v. 31, n. 2, p. 367-394, 2014.

PEREIRA, Márcia Maria de Carvalho et al. Análise dos fatores associados ao desempenho em matemática dos estudantes do 5º ano da rede estadual de educação do estado de Goiás no Saego de 2013. 2016.

PERRELLI, Marcia Regina; STRYER, Fábio Augusto. *Leitura: A contribuição das histórias em quadrinhos para a formação do leitor*. Paraná, 2012

PINO, Francisco Alberto. Modelos de decisão binários: uma revisão. *Rev. de Economia Agrícola*, São Paulo, v. 54, n. 1, p. 43-57, 2007.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

RIBEIRO, Vanda Mendes. **Justiça na escola e regulação institucional em redes de ensino do estado de São Paulo**. 2012. Tese de Doutorado. Universidade de São Paulo.

RIBEIRO, L. C. de Q.; KOSLINSKI, M. C. Fronteiras urbanas da democratização das oportunidades educacionais: o caso do Rio de Janeiro. In: RIBEIRO, L. C. de Q.; KOSLINSKI, M. C.; ALVES, F.; LASMAR, C. (Orgs.). **Desigualdades urbanas, desigualdade escolares**. Rio de Janeiro: Letra Capital, Observatório das Metrópoles, Ippur/UFRJ, 2010, p. 121-153.

RODRIGUES, Giovane. **Jovens que leem livros durante as férias têm melhor rendimento escolar**. Universidade Metodista de São Paulo, São Paulo, 06 de dez. de 2019. Disponível em: < <http://www.metodista.br/rroonline/jovens-que-leem-livros-durante-as-ferias-tem-melhor-rendimento-escolar>>. Acesso em: 29 de Ago. de 2020.

RONCA, Antônio Carlos Caruso; ALVES, Luiz Roberto. *O Plano Nacional de Educação e o Sistema Nacional de Educação: educar para a equidade*. **São Paulo: Fundação Santillana**, 2015.

SAEB, INEP Relatório. 2017. In: SAEB, INEP. Público-alvo. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2019a p. 19.

SAEB, INEP Relatório. 2017. In: SAEB, INEP. Etapas da divulgação. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2019b p. 20.

SAEB, INEP Relatório. 2017. In: SAEB, INEP. Questionários: Alunos. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2019c p. 41.

SAEB, INEP Relatório. 2017. In: SAEB, INEP. Escala de proficiência/desempenho. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2019d p. 45.

SAEB, INEP Relatório. 2017. In: SAEB, INEP. Conceitos utilizados no relatório. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2019e p. 21.

SAMEJIMA, Fumiko. Estimation of latent ability using a response pattern of graded scores. **Psychometrika monograph supplement**, 1969.

SCHWARZ, Gideon et al. Estimating the dimension of a model. **The annals of statistics**, v. 6, n. 2, p. 461-464, 1978.

SEED. DCE - Diretrizes Curriculares Estaduais de Língua Portuguesa do Estado do Paraná. Curitiba: **SEED**, 2008.

SILVEIRA, Andréia Cardoso. Raça e desempenho escolar: uma análise comparativa do desempenho de crianças negras e brancas em escolas integrantes do Projeto GERES em Salvador-BA. 2012.

SOARES, Jose Francisco. Índice de desenvolvimento da educação de São Paulo–Idesp. **São Paulo Perspec**, v. 23, n. 1, p. 29-41, 2009.

SOARES, José Francisco. Qual desempenho é adequado nos testes da Prova Brasil? Nova Escola, 04 de set. de 2018. Disponível em: <<https://novaescola.org.br/conteudo/12526/qual-desempenho-e-adequado-nos-testes-da-prova-brasil>>. Acesso em: 22 de jul. de 2020.

SOARES, Jose Francisco; ANDRADE, Renato Júdice de. Nível socioeconômico, qualidade e equidade das escolas de Belo Horizonte. **Ensaio: avaliação e políticas públicas em educação**, v. 14, n. 50, p. 107-125, 2006.

SOARES, Tufi Machado; MENDONÇA, Márcia Cristina Meneghin. Construção de um modelo de regressão hierárquico para os dados do SIMAVE-2000. **Pesquisa Operacional**, v. 23, n. 3, p. 421-441, 2003a pg. 423.

SOARES, Tufi Machado; MENDONÇA, Márcia Cristina Meneghin. Construção de um modelo de regressão hierárquico para os dados do SIMAVE-2000. **Pesquisa Operacional**, v. 23, n. 3, p. 421-441, 2003b pg. 442.

SOARES, Enilvia Rocha Morato. O dever de casa no contexto da avaliação de aprendizagens. 2011. 232 f. Dissertação (Mestrado em Educação) - **Universidade de Brasília**, Brasília, 2011.

SOARES, Sergei; SÁTYRO, Natália. O impacto de infra-estrutura escolar na taxa de distorção idade-série das escolas brasileiras de ensino fundamental: 1998 a 2005. 2008.

SPOSATI, A. Equidade. In: OLIVEIRA, D.A.; DUARTE, A.M.C.; VIEIRA, L.M.F. DICIONÁRIO: trabalho, profissão e condição docente. Belo Horizonte: **UFMG/Faculdade de Educação**, 2010.

STASINOPOULOS, D. Mikis et al. Generalized additive models for location scale and shape (GAMLSS) in R. **Journal of Statistical Software**, v. 23, n. 7, p. 1-46, 2007

ANEXO A – Escala de Proficiência



ESCALA DE PROFICIÊNCIA DE LÍNGUA PORTUGUESA 5º ANO DO ENSINO FUNDAMENTAL

Nível	Descrição do Nível
Nível 0 Desempenho menor que 125	<p>A Prova Brasil não utilizou itens que avaliam as habilidades deste nível.</p> <p>Os estudantes localizados abaixo do nível 125 requerem atenção especial, pois não demonstram habilidades muito elementares.</p>
Nível 1 Desempenho maior ou igual a 125 e menor que 150	<p>Os estudantes provavelmente são capazes de:</p> <ul style="list-style-type: none"> Localizar informações explícitas em textos narrativos curtos, informativos e anúncios. Identificar o tema de um texto. Localizar elementos como o personagem principal. Estabelecer relação entre partes do texto: personagem e ação; ação e tempo; ação e lugar.
Nível 2 Desempenho maior ou igual a 150 e menor que 175	<p>Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de:</p> <ul style="list-style-type: none"> Localizar informações explícitas em contos. Identificar o assunto principal e a personagem principal em reportagem e em fábulas. Reconhecer a finalidade de receitas, manuais e regulamentos. Inferir características de personagens em fábulas. Interpretar linguagem verbal e não-verbal em tirinhas.
Nível 3 Desempenho maior ou igual a 175 e menor que 200	<p>Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de:</p> <ul style="list-style-type: none"> Localizar informação explícita em contos e reportagens. Localizar informação explícita em propagandas com ou sem apoio de recursos gráficos. Reconhecer relação de causa e consequência em poemas, contos e tirinhas. Inferir o sentido de palavra, o sentido de expressão ou o assunto em cartas, contos, tirinhas e histórias em quadrinhos com o apoio de linguagem verbal e não verbal.
Nível 4 Desempenho maior ou igual a 200 e menor que 225	<p>Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de:</p> <ul style="list-style-type: none"> Identificar informação explícita em sinopses e receitas culinárias. Identificar assunto principal e personagem em contos e letras de música. Identificar formas de representação de medida de tempo em reportagens. Identificar assuntos comuns a duas reportagens. Identificar o efeito de humor em piadas. Reconhecer sentido de expressão, elementos da narrativa e opinião em reportagens, contos e poemas. Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em fábulas, poemas, contos e tirinhas. Inferir sentido decorrente da utilização de sinais de pontuação e sentido de expressões em poemas, fábulas e contos. Inferir efeito de humor em tirinhas e histórias em quadrinhos.



Nível	Descrição do Nível
<p>Nível 5 Desempenho maior ou igual a 225 e menor que 250</p>	<p>Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de:</p> <ul style="list-style-type: none"> • Identificar assunto e opinião em reportagens e contos. • Identificar assunto comum a cartas e poemas. • Identificar informação explícita em letras de música e contos. • Reconhecer assunto em poemas e tirinhas. • Reconhecer sentido de conjunções e de locuções adverbiais em verbetes, lendas e contos. • Reconhecer finalidade de reportagens e cartazes. • Reconhecer relação de causa e consequência e relação entre pronome e seu referente em tirinhas, contos e reportagens. • Inferir elementos da narrativa em fábulas, contos e cartas. • Inferir finalidade e efeito de sentido decorrente do uso de pontuação e assunto em fábulas. • Inferir informação em poemas, reportagens e cartas. • Diferenciar opinião de fato em reportagens. • Interpretar efeito de humor e sentido de palavra em piadas e tirinhas.
<p>Nível 6 Desempenho maior ou igual a 250 e menor que 275</p>	<p>Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de:</p> <ul style="list-style-type: none"> • Identificar opinião e informação explícita em fábulas, contos, crônicas e reportagens. • Identificar informação explícita em reportagens com ou sem o auxílio de recursos gráficos. • Reconhecer a finalidade de verbetes, fábulas, charges e reportagens. • Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em poemas, fábulas e contos. • Inferir assunto principal e sentido de expressão em poemas, fábulas, contos, crônicas, reportagens e tirinhas. • Inferir informação em contos e reportagens. • Inferir efeito de humor e moral em piadas e fábulas.
<p>Nível 7 Desempenho maior ou igual a 275 e menor que 300</p>	<p>Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de:</p> <ul style="list-style-type: none"> • Identificar assunto principal e informações explícitas em poemas, fábulas e letras de música. • Identificar opinião em poemas e crônicas. • Reconhecer o gênero textual a partir da comparação entre textos e assunto comum a duas reportagens. • Reconhecer elementos da narrativa em fábulas. • Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em fábulas, contos e crônicas. • Inferir informação e efeito de sentido decorrente do uso de sinais gráficos em reportagens e em letras de música. • Interpretar efeito de humor em piadas e contos. • Interpretar linguagem verbal e não verbal em histórias em quadrinhos.



Nível	Descrição do Nível
<p>Nível 8 Desempenho maior ou igual a 300 e menor que 325</p>	<p>Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de:</p> <ul style="list-style-type: none"> ▪ Identificar assunto principal e opinião em contos e cartas do leitor. ▪ Reconhecer sentido de locução adverbial e elementos da narrativa em fábulas e contos. ▪ Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em fábulas e reportagens. ▪ Reconhecer assunto comum entre textos de gêneros diferentes. ▪ Inferir informações e efeito de sentido decorrente do uso de pontuação em fábulas e piadas.
<p>Nível 9 Desempenho maior ou igual a 325</p>	<p>Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de:</p> <ul style="list-style-type: none"> ▪ Identificar opinião em fábulas e reconhecer sentido de advérbios em cartas do leitor.

APÊNDICE – Script utilizado para análise dos dados

```
library(gamlss.mx)
```

Para o modelo nulo

```
m0 <- gamlss(VR ~ 1 + re(random=~1|U.E.), family=BI, data=BaseR)
```

Para seleção do melhor modelo via stepwise utilizado BIC.

```
m1 <- stepGAIC(m0, scope=list(lower=~1 + re(random=~1|U.E.),
    upper=~re(random=~1|U.E.)+V1E+V2E+
    V3E + V4E + V5E + V6E + V7E + V8E +
    V9E + V10E + V11E + V12E + V13E + V14E +
    V15E + V16E + V17E + V18E + V19E + V20E +
    V21E + V22E + V23E + V24E + V25E + V26E +
    V27E + V28E + V1UE + V2UE + V3UE + V4UE +
    V5UE + V6UE + V7UE + V8UE),
    direction = 'both', k=log(length(BaseR$V1E)))
```

Informações e resumo do modelo nulo

```
m0
```

```
summary(m0)
```

```
getSmo(m0)
```

Informações e resumo do modelo com as covariáveis

```
plot(m1)
```

```
wp(m1)
```

```
rqres.plot(m1, xlim = c(-5,5), ylim = c(-1,1), plot.type = "all", pch=20, cex=.5,
col='black')
```