
Seleção de modelos multiníveis para dados de
avaliação educacional

Fabiano Rodrigues Coelho

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA
UFSCar-USP

FABIANO RODRIGUES COELHO

**SELEÇÃO DE MODELOS MULTINÍVEIS PARA DADOS DE
AVALIAÇÃO EDUCACIONAL**

Dissertação apresentada ao Departamento de Estatística – DEs-UFSCar e ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística.

Orientadora: Profa. Dra. Cibele Maria Russo Noveli

**São Carlos
Outubro de 2017**

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA
UFSCar-USP

FABIANO RODRIGUES COELHO

**SELECTION OF MULTILEVEL MODELS FOR EDUCATIONAL
EVALUATION DATA**

Master dissertation submitted to the Departamento de Estatística – DEs-UFSCar and to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master joint Graduate Program in Statistics.

Advisor: Profa. Dra. Cibele Maria Russo Noveli

**São Carlos
October 2017**



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Fabiano Rodrigues Coelho, realizada em 11/08/2017:

Profa. Dra. Cibele Maria Russo Novelli
USP

Prof. Dr. Caio Lucidius Naberezny Azevedo
UNICAMP

Prof. Dr. Jony Arrais Pinto Junior
UFF

*Este trabalho é dedicado a todos aqueles que amam ciência e
que nunca desiste de seus sonhos, mesmo quando se depara com dificuldades,
não esmorece levanta a cabeça e segue em frente.*

AGRADECIMENTOS

Início os agradecimentos primeiramente a Deus, pois é Nele em que busco minhas forças para suportar os momentos mais difíceis em minha caminhada.

Agradeço também aos meus pais Celino Rodrigues Coelho e Cícera Aparecida da Silva Coelho, minha irmã Danielle da Silva Coelho e minha tia-avó Izabel Maria dos Passos por todo apoio e torcida pelo meu sucesso.

Agradeço ainda, a Eliandra de Mello Bonotto, que é mais do que minha namorada, minha outra metade, meu braço direito, minha maior incentivadora, que está comigo há muitos anos e que torce muito para que eu consiga realizar meus sonhos.

Estendo meus agradecimentos a minha orientadora Prof^a. Dra. Cibele Maria Russo Noveli por tudo, pelas orientações, pela paciência, pelas críticas construtivas, pelos conselhos e também pelo tratamento muito respeitoso em todas as nossas conversas.

Agradeço também os membros das bancas de Qualificação e Defesa deste trabalho: Prof. Dra. Cibele Maria Russo Noveli, Prof. Dr. Jorge Luis Bazán Guzmán, Prof. Dr. Jony Arrais Pinto Junior, Prof. Dr. Caio Lucidius Naberezny Azevedo, pelas reflexões, trocas de ideias e os mais produtivos questionamentos que me proporcionou uma evolução profissional e pessoal muito grande.

Agradeço, em especial, ao Prof. Dr. Jorge Luis Bazán Guzmán, que quando minha orientadora esteve de licença maternidade me deu todo apoio necessário.

Aos docentes: Prof. Dr. Adriano Kamimura Suzuki, Prof. Dr. Marcio Alves Diniz, Prof. Dr. Gustavo Henrique de Araújo Pereira, Prof. Dra. Juliana Cobre, Prof. Dr. Carlos Alberto Ribeiro Diniz e Prof. Dr. Adriano Polpo de Campos, os quais eu tive a honra de ser instruído e aprender muito.

Aos meus amigos da turma de mestrado ingressante em 2015: Alan Henrique de Jesus, Bárbara Beltrame Bettim, Caroline Tenório Mendes de Aquino, Diego Mattozo Bernardes da Silva, Gretta Rossi Ferreira, Juliana Cecília da Silva Teixeira, Karine Zanuto Mendes, Murilo Cantoni, Natália Lombardi de Oliveira, Nicholas Wagner Eugenio, Raul Caram de Assis, Susan Alicia Chumbimune Anyosa e Taís Roberta Ribeiro, pessoas maravilhosas, parceiros de coração com a qual pude dividir os momentos bons e ruins ao longo do mestrado, desejo tudo de bom a eles. Obrigado por tudo.

À servidora Maria Isabel Rinaldo Pessoa de Araújo e aos funcionários do serviço de Pós-Graduação do ICMC-USP, por todo suporte e solicitude quando foi necessário resolver

algum assunto pendente.

Aos amigos que frequentam os laboratórios do ICMC-USP, em especial ao Hélio Azevedo e José Pedro Bello, pelas longas horas de conversas e por todo incentivo e apoio.

Ao CEPID-CeMEAI (Centro de Pesquisa, Inovação e Difusão do Centro de Ciências Matemáticas Aplicadas à Indústria), pela utilização do cluster Euler, sem ele, este trabalho se tornaria praticamente inviável e também aos seus funcionários que sempre que necessário, não deixaram de prestar auxílio, seja pessoalmente ou por e-mail.

A todos com quem tive a oportunidade de conviver no período, muito obrigado por tudo.

*“Não fui eu que lhe ordenei? Seja forte e corajoso!
Não se apavore, nem se desanime, pois o Senhor, o seu Deus,
estará com você por onde você andar”
(Josué 1:9)*

RESUMO

COELHO, F. R. **Seleção de modelos multiníveis para dados de avaliação educacional**. 2017. 173 p. Dissertação (Mestrado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

Quando um conjunto de dados possui uma estrutura hierárquica, uma possível abordagem são os modelos de regressão multiníveis, que se justifica pelo fato de haver uma porção significativa da variabilidade dos dados que pode ser explicada por níveis macro. Neste trabalho, desenvolvemos a seleção de modelos de regressão multinível aplicados a dados educacionais. Esta análise divide-se em duas partes: seleção de variáveis e seleção de modelos. Esta última subdivide-se em dois casos: modelagem clássica e modelagem bayesiana. Buscamos através de critérios como o Lasso, AIC, BIC, WAIC entre outros, encontrar quais são os fatores que influenciam no desempenho em matemática dos alunos do nono ano do ensino fundamental do estado de São Paulo. Também investigamos o funcionamento de cada um dos critérios de seleção de variáveis e de modelos. Foi possível concluir que, sob a abordagem frequentista, o critério de seleção de modelos BIC é o mais eficiente, já na abordagem bayesiana, o critério WAIC apresentou melhores resultados. Utilizando o critério de seleção de variáveis Lasso para abordagem clássica, houve uma diminuição de 34% dos preditores do modelo. Por fim, identificamos que o desempenho em matemática dos estudantes do nono ano do ensino fundamental do estado de São Paulo é influenciado pelas seguintes covariáveis: grau de instrução da mãe, frequência de leitura de livros, tempo gasto com recreação em dia de aula, o fato de gostar de matemática, o desempenho em matemática global da escola, desempenho em língua portuguesa do aluno, dependência administrativa da escola, sexo, grau de instrução do pai, reprovações e distorção idade-série.

Palavras-chave: Modelos multiníveis, Seleção de modelos, Critério de informação e Prova Brasil.

ABSTRACT

COELHO, F. R. **Selection of multilevel models for educational evaluation data.** 2017. 173 p. Dissertação (Mestrado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

When a dataset contains a hierarchical data structure, a possible approach is the multilevel regression modelling, which is justified by the significant amount of the data variability that can be explained by macro level processes. In this work, a selection of multilevel regression models for educational data is developed. This analysis is divided into two parts: variable selection and model selection. The latter is subdivided into two categories: classical and Bayesian modeling. Traditional criteria for model selection such as Lasso, AIC, BIC, and WAIC, among others are used in this study as an attempt to identify the factors influencing ninth grade students' performance in Mathematics of elementary education in the State of São Paulo. Likewise, an investigation was conducted to evaluate the performance of each variable selection criteria and model selection methods applied to fitted models that will be mentioned throughout this work. It was possible to conclude that, under the frequentist approach, BIC is the most efficient, whereas under the Bayesian approach, WAIC presented better results. Using Lasso under the frequentist approach, a decrease of 34% on the number of predictors was observed. Finally, we identified that the performance in Mathematics of students in the ninth year of elementary school in the state of São Paulo is most influenced by the following covariates: mother's educational level, frequency of book reading, time spent with recreation in classroom, the fact of liking Math, school global performance in Mathematics, performance in Portuguese, school administrative dependence, gender, father's educational degree, failures and age-grade distortion.

Keywords: Multilevel models; Model selection; Information criterion and Brazil Exam - Basic Education Assessment.

LISTA DE ILUSTRAÇÕES

Figura 1 – Histograma da variável de interesse, PROFICIENCIA_MT_SAEB.	31
Figura 2 – Gráfico qqplot da variável PROFICIENCIA_MT_SAEB.	31
Figura 3 – Box plot da variável de interesse, PROFICIENCIA_MT_SAEB em cada uma das 15 regiões destacadas anteriormente.	35
Figura 4 – Box plot da variável de interesse, PROFICIENCIA_MT_SAEB pelas covariáveis Q001, Q004, Q019 e Q023.	35
Figura 5 – Box plot da variável de interesse, PROFICIENCIA_MT_SAEB pelas covariáveis Q049, TURNO, DEPENDENCIA_ADM e LOCALIZACAO.	36
Figura 6 – Box plot da variável de interesse, PROFICIENCIA_MT_SAEB pelas covariáveis Q020, Q033, Q044 e Q054.	36
Figura 7 – Esquema representativo de uma estrutura hierárquica de três níveis.	38
Figura 8 – Análise residual do modelo 7 por grupos.	80
Figura 9 – Box plot dos resíduos do modelo 7 por grupos.	81
Figura 10 – Gráfico da função densidade marginal <i>a posteriori</i> dos parâmetros do Modelo 4.	88
Figura 11 – Histogramas da densidade <i>a posteriori</i> simulada via MCMC com dinâmica hamiltoniana dos parâmetros do modelo 4.	88
Figura 12 – Desenvolvimento do processo MCMC durante as 20.000 iterações para cada um dos parâmetros.	89
Figura 13 – Gráfico dos BIC versus λ , para 21 valores de λ tomados entre 0 e 1000.	89

LISTA DE TABELAS

Tabela 1 – Análise Descritiva das Variáveis PROFICIENCIA_MT_SAEB	30
Tabela 2 – Demonstrativo do número de escolas por região amostrada	34
Tabela 3 – Configuração dos modelos estimados em cada réplica.	60
Tabela 4 – Percentual de réplicas onde os critérios de seleção optaram pelo modelo verdadeiro.	61
Tabela 5 – Modelo 1: apenas com os níveis aluno e turma ; Modelo 2: com os níveis aluno, turma e escola.	65
Tabela 6 – Modelo 3: inclusão de variáveis explicativas ao nível de aluno; Modelo 4: inclusão de variáveis explicativas ao nível de escola.	68
Tabela 7 – Teste de Comparação entre os modelos 3 e 4	69
Tabela 8 – Comparação dos modelos ajustados nesta seção utilizando os critérios AIC, BIC, AICc.	70
Tabela 9 – Descrição das covariáveis a serem utilizadas nos ajustes no caso clássico e no bayesiano.	70
Tabela 10 – Estimativa dos parâmetros do modelo (6.7), aqui temos efeitos fixos apenas.	71
Tabela 11 – Procedimento de stepwise utilizado no Modelo 4	72
Tabela 12 – Modelo (6.8) tem intercepto aleatório para turma; Modelo (6.9) tem intercepto aleatório para escola; Modelo (6.10) tem intercepto aleatório para região.	73
Tabela 13 – Modelo (6.11) tem intercepto e inclinações aleatórias para turma; Modelo (6.12) tem intercepto e inclinações aleatórias para escola; Modelo (6.13) tem intercepto e inclinações aleatórias para região.	75
Tabela 14 – Comparação dos modelos ajustados nesta seção utilizando os critérios AIC, BIC, AICc.	76
Tabela 15 – Configuração de ajustes para seleção de modelos.	77
Tabela 16 – Examinando as estimativas dos parâmetros e o valor p dos mesmos, nos modelos 1, 2 e 3 descritos na Tabela 15.	77
Tabela 17 – Examinando as estimativas dos parâmetros e o valor p dos mesmos, nos modelos 4, 5 e 6 descritos na Tabela 15.	78
Tabela 18 – Examinando as estimativas dos parâmetros e o valor p do modelo 7 descritos na Tabela 15.	79
Tabela 19 – Comparação dos modelos clássicos da Tabela 15 mais o modelo completo (6.10), utilizando os critérios AIC, AICc e BIC.	79

Tabela 20 – Estimativas dos efeitos aleatórios do modelo 7	80
Tabela 21 – Estimativas dos parâmetros, intervalos de credibilidade e a estatística \hat{R} no modelo (6.14).	82
Tabela 22 – Examinando as estimativas dos parâmetros, intervalos de credibilidade e a estatística \hat{R} no modelo (6.15).	83
Tabela 23 – Examinando as estimativas dos parâmetros, intervalos de credibilidade e a estatística \hat{R} no modelo (6.16).	84
Tabela 24 – Comparação dos modelos ajustados utilizando os critérios <i>looic</i> , WAIC, DIC,EAIC e EBIC.	84
Tabela 25 – Examinando as estimativas dos parâmetros, intervalos de credibilidade dos modelos 1, 2 e 3 descritos na Tabela 15.	85
Tabela 26 – Examinando as estimativas dos parâmetros, intervalos de credibilidade dos modelos 4, 5 e 6 descritos na Tabela 15.	86
Tabela 27 – Examinando as estimativas dos parâmetros, intervalos de credibilidade do modelo 7 descritos na Tabela 15.	87
Tabela 28 – Comparação dos modelos com interceptos e inclinações aleatórias, utilizando os critérios <i>looic</i> , WAIC, DIC,EAIC e EBIC.	87
Tabela 29 – Resultados referentes ao algoritmo glmmLasso com as covariáveis descritas anteriormente, para $\lambda = 450$, foi selecionado no total 34 variáveis explicativas.	90
Tabela 30 – Descrição do banco de dados geral.	99
Tabela 31 – Codificação das variáveis	126
Tabela 32 – Descrição das variáveis Dicotomizadas da amostra obtida dos dados	148
Tabela 33 – Distribuição amostral das escolas da região de Presidente Prudente segundo os critérios pré-estabelecidos.	159
Tabela 34 – Distribuição amostral das escolas da região de Araçatuba segundo os critérios pré-estabelecidos.	159
Tabela 35 – Distribuição amostral das escolas da região de Jales segundo os critérios pré-estabelecidos.	159
Tabela 36 – Distribuição amostral das escolas da região de São José do Rio Preto segundo os critérios pré-estabelecidos.	160
Tabela 37 – Distribuição amostral das escolas da região de Marília segundo os critérios pré-estabelecidos.	160
Tabela 38 – Distribuição amostral das escolas da região de Bauru segundo os critérios pré-estabelecidos.	160
Tabela 39 – Distribuição amostral das escolas da região de São Carlos segundo os critérios pré-estabelecidos.	161
Tabela 40 – Distribuição amostral das escolas da região de Ribeirão Preto segundo os critérios pré-estabelecidos.	161

Tabela 41 – Distribuição amostral das escolas da região de Registro segundo os critérios pré-estabelecidos.	161
Tabela 42 – Distribuição amostral das escolas da região de Itapetininga segundo os critérios pré-estabelecidos.	162
Tabela 43 – Distribuição amostral das escolas da região de Piracicaba segundo os critérios pré-estabelecidos.	162
Tabela 44 – Distribuição amostral das escolas da região de Campinas segundo os critérios pré-estabelecidos.	162
Tabela 45 – Distribuição amostral das escolas da região de Jundiaí segundo os critérios pré-estabelecidos.	163
Tabela 46 – Distribuição amostral das escolas da região de São José dos Campos segundo os critérios pré-estabelecidos.	163
Tabela 47 – Distribuição amostral das escolas da região de São Paulo segundo os critérios pré-estabelecidos.	163

SUMÁRIO

1	INTRODUÇÃO	25
1.1	Justificativa	26
1.2	Revisão Bibliográfica	27
2	MOTIVAÇÃO	29
2.1	Banco de dados	29
2.2	Análises Descritivas	30
2.3	Amostragem dos dados	32
3	MODELO MULTINÍVEL COM RESPOSTA NORMAL	37
3.1	Modelos de regressão e suposições	37
3.2	Modelo de regressão com dois níveis	38
3.3	Estimação de parâmetros	40
3.4	Testes de Hipóteses	41
3.5	Enfoque bayesiano	42
4	MÉTODOS DE ESCOLHA DE MODELOS	45
4.1	Seleção de Modelos: Caso Clássico	45
4.1.1	<i>AIC (Akaike Information criterion)</i>	46
4.1.2	<i>AICc (AIC Corrigido)</i>	46
4.1.3	<i>BIC (Bayesian Information Criterion)</i>	47
4.2	Seleção de Modelos: Caso Bayesiano	47
4.2.1	<i>DIC (Deviance Information Criterion)</i>	48
4.2.2	<i>WAIC (Widely Applicable Information Criterion)</i>	48
4.2.3	<i>loo (Leave-one-out cross-validation)</i>	50
4.2.4	<i>EAIC (Expected Akaike Information Criterion)</i>	51
4.2.5	<i>EBIC (Expected Bayesian Information Criterion)</i>	52
4.3	Seleção de Variáveis	52
4.3.1	<i>Lasso</i>	52
4.3.2	<i>Modelos lineares com dois níveis e resposta Normal</i>	53
4.3.3	<i>Algoritmo do Gradiente Ascendente</i>	54
4.3.3.1	<i>Algoritmo glmmLasso: caso particular (normal)</i>	54
5	ESTUDO DE SIMULAÇÃO	59

6	APLICAÇÕES A DADOS EDUCACIONAIS	63
6.1	Modelos Multiníveis	63
6.1.1	<i>Modelo 1</i>	64
6.1.2	<i>Modelo 2</i>	64
6.1.3	<i>Modelo 3</i>	65
6.1.4	<i>Modelo 4</i>	66
6.2	Modelagem clássica	70
6.2.1	<i>Modelos com efeitos fixos</i>	71
6.2.2	<i>Modelos com intercepto aleatório</i>	72
6.2.3	<i>Modelos com intercepto e inclinações aleatórios</i>	74
6.3	Seleção de Modelos Clássicos	76
6.4	Modelagem Bayesiana	80
6.4.1	<i>Modelo com efeitos fixos</i>	81
6.4.2	<i>Modelo com interceptos aleatórios</i>	82
6.4.3	<i>Modelo com interceptos e inclinações aleatórios</i>	82
6.5	Seleção de Modelos Bayesianos	83
6.6	Seleção de Variáveis	86
7	DISCUSSÃO E PROPOSTAS FUTURAS	91
	REFERÊNCIAS	95
APÊNDICE A	BANCO DE DADOS	99
APÊNDICE B	AMOSTRAGEM	155
B.1	Regiões amostrais	155
B.2	Distribuições amostrais por região	158
APÊNDICE C	MODELOS CLÁSSICOS	165
C.1	Estimação dos parâmetros	165
C.1.1	<i>Modelos com efeitos fixos</i>	165
C.1.2	<i>Modelo com interceptos aleatórios</i>	165
C.1.3	<i>Modelo com interceptos e inclinações aleatórios</i>	166
C.2	Seleção de Modelos	166
APÊNDICE D	MODELOS BAYESIANOS: STAN	167
D.1	Modelos com efeitos fixos	168
D.2	Modelo com interceptos aleatórios	169
D.3	Modelo com interceptos e inclinações aleatórios	171

INTRODUÇÃO

Quando temos unidades observacionais formando grupos, que por sua vez também podem ser agrupados em unidades maiores, podemos afirmar que esses dados têm estrutura hierárquica. Para estudar fenômenos relativos a esse tipo de dados, uma possibilidade é considerar modelos de regressão multiníveis com dois níveis, que permitem a incorporação de efeitos aleatórios associados a cada um dos níveis. Uma diferença marcante nesses modelos em relação aos modelos de regressão linear múltipla é o fato de que leva em consideração a possível correlação existente entre os dados de um mesmo grupo, como nos diferentes níveis de hierarquia, ver por exemplo [Goldstein \(1986\)](#). Outra suposição comumente considerada na literatura é a de normalidade. Esses modelos têm aplicação em problemas biológicos, médicos, de economia, educacionais, de controle de qualidade, além de diversas outras áreas de aplicação.

Outra característica dos modelos usuais que merece atenção é a suposta independência entre as observações, que pode levar à superestimação dos parâmetros, quando estimados usando o método de máxima verossimilhança restrita. Os modelos multiníveis permitem avaliar a variabilidade dos dados nos diferentes níveis e ajustar um modelo mais adequado à realidade do problema.

Segundo [Castro \(2015\)](#), podemos resumir as vantagens do modelo de regressão multinível, da seguinte maneira:

- apresentam modelos mais flexíveis e estruturados que utilizam um maior nível de informação da amostra e, ainda permitem estabelecer uma equação para cada unidade de agrupamento, o que permite análises individuais para cada grupo;
- possibilitam a formulação e o teste de hipóteses relativas a efeitos entre os níveis a partir do uso da informação de cada agrupamento dos dados;
- possibilitam particionar a variabilidade da variável resposta entre os diferentes níveis, de acordo com a proporção explicada por cada um deles.

Neste capítulo, apresentamos a justificativa para a realização desse trabalho, bem como uma breve revisão bibliográfica a respeito do tema aqui colocado.

1.1 Justificativa

Neste trabalho, apresentamos um estudo de modelos de regressão multiníveis para conjuntos de dados educacionais. Esses modelos são aplicáveis, em geral, para problemas em que os dados são estruturados de forma hierárquica. Realizamos seleção de variáveis e seleção de modelos em modelos multiníveis, como extensão do trabalho de Osio (2013), que podem ser entendidos como modelos onde os componentes aleatórios assumem distribuições de probabilidade. Como aplicação, utilizamos um conjunto de dados de avaliações educacionais do SAEB - Sistema Nacional de Avaliação da Educação Básica, cujo objetivo principal é verificar o rendimento escolar dos alunos na rede pública e privada de ensino brasileira.

A motivação para esse estudo vem da área educacional, em que os dados observados dos alunos são agrupados por turmas, as turmas são agrupadas por escolas e que por sua vez são agrupadas por cidades e regiões.

A educação é uma das principais demandas de nossa atual sociedade. Nos últimos anos, como se tem visto, o Brasil tem alcançado os últimos lugares em avaliações de larga escala, conforme tem sido divulgado pelo Ministério da Educação e organismos internacionais. Em 2012, o desempenho dos estudantes brasileiros em leitura piorou em relação a 2009. De acordo com dados do Pisa (Programa Internacional de Avaliação de Alunos), o país somou 410 pontos em leitura, dois a menos do que sua pontuação na avaliação anterior e 86 pontos abaixo da média dos países da OCDE (Organização e Cooperação e Desenvolvimento Econômico). Matemática foi a única disciplina em que os brasileiros apresentam avanço no desempenho, ainda que pequeno. O Brasil saiu de 386 pontos, em 2009, e foi a 391 pontos em 2012, a média da OCDE é de 494 pontos. A melhora não foi suficiente para que o país avançasse no ranking e o Brasil caiu para 58^o posição em matemática. Apesar da melhora, dois em cada três alunos brasileiros de 15 anos não conseguem interpretar situações que exigem apenas deduções diretas da informação dada, não são capazes de entender percentuais, frações ou gráficos.

Neste trabalho, utilizamos um conjunto de dados do SAEB, do ano de 2011, que consiste em uma avaliação de língua portuguesa e matemática, aplicada aos alunos do nono ano do ensino fundamental, juntamente com informações de diretores, docentes e escolas. Essa prova produz o IDEB, que é o índice de desenvolvimento da educação básica. Realizamos seleção de modelos e seleção de variáveis em problemas com dados correlacionados com um grande volume de dados. Será desenvolvido o processo de estimação, seleção de variáveis e modelos, entre outras análises.

Na próxima seção, fazemos uma breve revisão bibliográfica a respeito de modelos multiníveis, métodos de seleção de variáveis e seleção de modelos e também serão abordados os principais trabalhos com essa classe de modelos aplicados a avaliações educacionais.

1.2 Revisão Bibliográfica

Quando se trata de modelos multiníveis, o que mais se encontra, no que diz respeito à bibliografia, são comparações em situações práticas desses modelos com modelos de regressão linear múltipla. Essas comparações mostram a importância, em alguns casos, de se considerar a correlação existente entre os grupos observados. Com relação à seleção de variáveis e seleção de modelos, pode-se notar que os métodos usuais como AIC, BIC, LASSO e outros, não tem se mostrado eficaz quando se trata de um grande volume de dados, métodos esses que foram propostos a partir da década de 1970, e que foram aperfeiçoados ao longo dos anos.

Entre os trabalhos mais relevantes sobre modelos multiníveis podemos destacar [Osio \(2013\)](#), [Gelman e Hill \(2006\)](#), [Bergamo \(2002\)](#), [Natis \(2013\)](#), com relação à seleção de variáveis e seleção de modelos podemos destacar os trabalhos de [Akaike \(1998\)](#), [Tibshirani \(1996\)](#) e [Weakliem \(1999\)](#).

Muitos estudos foram feitos na área educacional utilizando modelos multiníveis. Eles mostram, por exemplo, que algumas ações realizadas pela família podem aumentar o desempenho em matemática. São elas: incentivar desde a primeira infância, através de jogos e atividades lúdicas, a afinidade pelos conceitos matemáticos; evitar que seus filhos ingressem na escola e permaneçam na escola com distorção idade-série; oferecer para seus filhos com relação à aprendizagem escolar, dentre outras. ([LAROS; MARCIANO; ANDRADE, 2010](#)) aponta ações por parte das escolas para a melhoria dos indicadores educacionais, são elas: aumentar o controle sobre a incidência de faltas em sala de aula; sanar as causas da repetência do ano letivo; estimular o trabalho colaborativo entre os professores; melhoria do clima disciplinar na escola prevenindo roubos, depredações e outros atos violentos.

Já o estudo realizado por [Moreira e Jacinto \(2013\)](#), no Rio Grande do Sul, mostrou que a infraestrutura da residência dos estudantes afeta positivamente no desempenho dos mesmos em matemática, o que não acontece quando se diz respeito à escola, mostrando que a escola pode contribuir muito para melhoria do desempenho em matemática. A mesma conclusão foi obtida por [Andrade e Soares \(2008\)](#) em seu estudo feito a nível nacional em todos os ciclos.

Em seu estudo comparativo entre duas escolas do Distrito Federal, [Santos \(2013\)](#) tomou uma escola com boas notas em matemática na PROVA BRASIL e outra com desempenho abaixo da média e a diferença mais marcante constatada pela autora foi organizacional.

. Utilizando dados do SARESP 97, [Natis \(2000\)](#) conclui que os modelos multiníveis constituem uma nova formulação para dados de efeitos aleatórios que permitem uma descrição e análise mais apropriada das diferentes fontes de variação, a autora deixa como proposta modelos para dados binários, ou ainda dados com distribuição Poisson.

Um texto de extrema importância no que diz respeito a estimação é colocado em [Searle et al. \(1994\)](#). Nele encontramos de maneira bem detalhada a obtenção dos estimadores de máxima verossimilhança restrita e do melhor preditor linear não-viesado dos efeitos aleatórios realizados,

mas não observáveis em modelos multiníveis.

O trabalho mais importante sobre o Lasso (*least absolute shrinkage and selection operator*) é Tibshirani (1996). Neste texto é proposto um novo método para a estimação dos parâmetros em modelos lineares. O Lasso minimiza a soma de quadrados dos resíduos submetidos à soma do valor absoluto dos coeficientes, sendo menos do que uma constante. Devido à natureza deste encolhimento ele tende a produzir alguns coeficientes que são exatamente zero e assim possibilitando a eliminação das variáveis cujo os coeficientes foram estimados como zero. As propriedades favoráveis ao método proposto é verificado neste artigo através de estudos de simulação.

Para obter uma noção intuitiva dos critérios de seleção de modelos, Emiliano (2013), em sua tese de doutorado, utiliza critérios AIC, BIC e outros em muitas áreas, tais como: seleção de modelos normais, séries temporais e modelos de crescimento. Após estudos de simulação ele conclui que quando falamos de modelos normais, o uso do BIC mostrou-se superior para amostras grandes, porém para amostras pequenas o BIC teve desempenho similar ao AIC e AIC corrigido. No caso de modelos de séries temporais, mesmo para amostras de tamanho pequeno, o desempenho dos critérios AIC e AIC corrigido foi similar e inferiores ao BIC, para todos os modelos simulados.

O texto está organizado da seguinte forma: No **Capítulo 2** é apresentada a motivação para esse estudo, que consiste de um banco de dados educacional proveniente de uma avaliação em larga escala aplicada aos alunos da rede pública de ensino brasileira. No **Capítulo 3** apresenta-se aspectos teóricos referentes ao modelo multinível com resposta normal, tais como: formulação teórica, estimação, testes de hipótese e ainda abordamos essa classe de modelos sob o enfoque bayesiano. No **Capítulo 4**, métodos de seleção de modelos e seleção variáveis que foram utilizados neste estudo são discutidos sob ponto de vista teórico. No **Capítulo 5** apresenta-se um breve estudo de simulação para os critérios de seleção de modelos clássicos. No **Capítulo 6** vemos uma aplicação a dados educacionais, utilizando o banco de dados mencionado anteriormente. Por fim, no **Capítulo 7** discute-se os resultados obtidos e coloca-se proposta para trabalhos futuros.

MOTIVAÇÃO

Neste capítulo, apresentamos a motivação para este estudo. Trata-se de um banco de dados educacional advindo de uma prova de larga escala aplicada a estudantes brasileiros. O objetivo principal é selecionar as variáveis que influenciam no desempenho dos alunos em matemática, do nono ano do ensino fundamental do Estado de São Paulo. A seguir destacamos alguns detalhes a respeito dos dados e sobre o processo de amostragem.

2.1 Banco de dados

A educação é uma das demandas mais importantes do país, uma vez que se espera que a educação básica forneça subsídios para que os alunos se tornem cidadãos conscientes de seus deveres e saibam pensar com competência. Desse modo, percebemos um crescimento significativo do número de estudos científicos com o objetivo de identificar os fatores, sejam eles socioeconômicos ou até mesmo estruturais, que interferem na aprendizagem do educando e assim fornecer condições para que os governantes tenham condições de elaborar políticas públicas que contribuam para a melhoria dos indicadores educacionais.

O banco de dados refere-se à Prova Brasil do ano de 2011. Essa prova de larga escala é aplicada de dois em dois anos em todas as escolas do Brasil. Os alunos do quinto e nono do ensino Fundamental são submetidos a esta prova, que verifica a proficiência em língua portuguesa e matemática.

A Prova Brasil faz parte do Sistema de Avaliação da Educação Básica (SAEB), que possui como principal objetivo avaliar a educação básica brasileira e contribuir para a melhoria de sua qualidade e para a universalização do acesso à escola, oferecendo subsídios concretos para a formulação, reformulação e o monitoramento das políticas públicas voltadas para a educação básica. Além disso, procura também oferecer indicadores que possibilitem maior compreensão dos fatores que influenciam o desempenho dos alunos nas áreas e anos avaliados.

Na ocasião da pesquisa, 5.201.730 alunos realizaram a prova em 72.808 escolas no Brasil. Foram entrevistados no total 304.412 professores destas escolas e todos os 72.808 diretores. No entanto, para esse estudo, consideramos apenas os alunos do Estado de São Paulo e do nono ano do Ensino Fundamental, totalizando assim 502.479 alunos que realizaram a Prova Brasil, em 18.139 turmas de 5.018 escolas, em 640 municípios do Estado de São Paulo.

Além de informações referentes ao desempenho dos alunos, possuímos também dados referentes aos questionários sócio-econômico que foram aplicados aos alunos, professores, escolas e diretores. O questionário do aluno é composto por 58 questões, o questionário da escola tem 66 questões, há 212 questões no questionário do diretor e 152 questões no questionário de professor.

Nosso objetivo é identificar fatores que influenciam no desempenho do aluno na disciplina de matemática, utilizando a nota na prova de matemática da PROVA BRASIL. Os dados podem ser encontrados no site do INEP, ver por exemplo [Microdados \(2011\)](#). Na próxima seção, apresentamos a análise descritiva dos dados de modo a entender algumas das particularidades dos mesmos.

2.2 Análises Descritivas

O histograma da variável PROFICIENCIA_MT_SAEB, que representa a proficiência em matemática pelo aluno, é apresentado a seguir e fornece uma ideia de que a distribuição dos dados é simétrica. Apresentamos também a análise descritiva dessa variável, que será a variável resposta do nosso futuro modelo.

A partir da Figura 1 e da Figura 2, não há evidências que nos leve a descartar a suposição de normalidade dos dados. Pela Tabela 1, também temos indício que os dados são normalmente distribuídos em torno da média.

Tabela 1 – Análise Descritiva das Variáveis PROFICIENCIA_MT_SAEB

Nome da variável	Mín.	1º Quartil	Mediana	Média	3º Quartil	Máx.
PROFICIENCIA_MT	109,4	215,2	247,9	247,1	278,9	398,3

Note que a PROFICIENCIA_MT_SAEB, tem um desvio padrão de 45,93, uma curtose de 2,73 (sinalizando uma distribuição platicúrtica), e a assimetria é de -0,01 (indicando uma leve assimetria negativa da distribuição).

No apêndice A, temos a Tabela 30, onde podemos encontrar a organização geral do banco de dados, consta lá, o código original da variável, o novo código da variável, nome da variável, enunciado relativo à variável (caso haja), tipo de variável e o valor p relativo a uma regressão individual, tendo como variável resposta a PROFICIENCIA_MT_SAEB e variável explicativa cada uma das variáveis listadas na referida tabela.

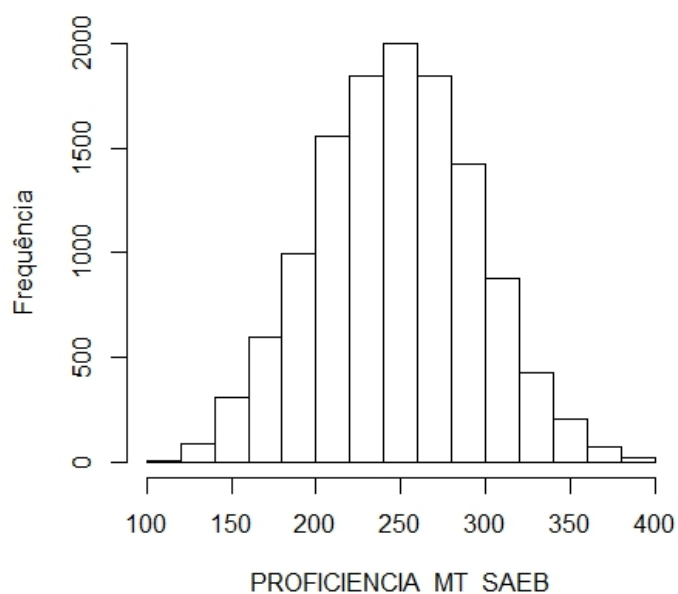


Figura 1 – Histograma da variável de interesse, PROFICIENCIA_MT_SAEB.

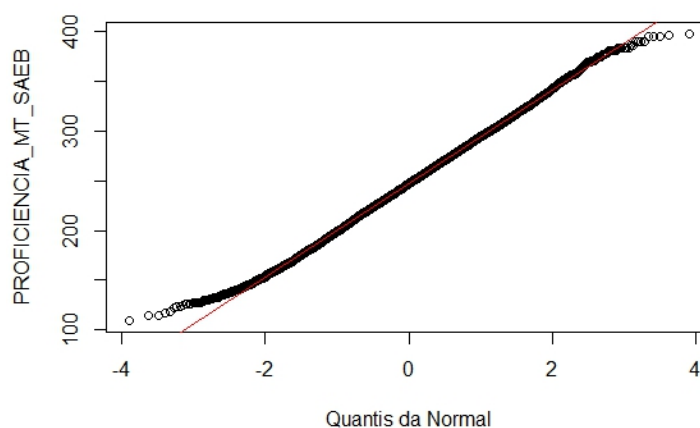


Figura 2 – Gráfico qqplot da variável PROFICIENCIA_MT_SAEB.

Das 498 variáveis explicativas listadas, tomando um nível de significância de 5%, há evidência de que apenas as variáveis D119 (Eventos abertos a comunidade), E027 (Proteção de equipamento de alto custo) e LOCALIZACAO (Localização) não são significativas e ainda foi constatado que em 436 variáveis o valor p é menor do que $2 \cdot 10^{-16}$.

Como existe no banco de dados uma grande quantidade de variáveis significativas, um fator importante a ser considerado é a significância teórica das variáveis em trabalhos acadêmicos, ver, por exemplo, (LAROS; MARCIANO; ANDRADE, 2010; NATIS, 2013; GONÇALVES;

FRANÇA, 2013). A partir desses trabalhos, vemos que as covariáveis mais importantes são: Q001 (Sexo), Q002 (Raça auto-atribuída), Q019 (Escolaridade da mãe ou a mulher responsável), Q023 (Escolaridade do pai ou homem responsável), Q049 (Reprovas), PROFICIENCIA_LP - SAEB (Proficiência em língua portuguesa do SAEB), DEPENDENCIA_ADM (Dependência Administrativa), LOCALIZACAO (Localização) e TURNO (Turno). Essas variáveis serão considerados nos primeiros ajustes deste trabalho.

Esse banco de dados inicial tem 10 variáveis contínuas e 488 variáveis categóricas, dos quais 238 possuem apenas duas opções de respostas, as variáveis restantes têm três ou mais opções de respostas, sendo assim, estas foram dicotomizadas, ficando com duas categorias de respostas.

2.3 Amostragem dos dados

Após a definição do universo da pesquisa como sendo alunos do nono ano do ensino fundamental do Estado de São Paulo, conduzimos um procedimento amostral, após excluir os alunos que estudam no turno noturno, pois esses tem pouca representatividade. Tomamos como unidade amostral as escolas, definimos o tamanho amostral, considerando uma amostragem aleatória simples de proporção para populações finitas. Segundo [Silva \(1998\)](#), considera-se a seguinte relação:

$$n = \frac{NZ^2p(1-p)}{(N-1)e^2 + Z^2p(1-p)}, \quad (2.1)$$

onde:

- n representa o tamanho da amostra que queremos calcular;
- N é o tamanho populacional;
- e é a margem de erro que queremos admitir;
- Z é o desvio do valor médio associado ao nível de confiança desejado;
- p é a proporção que esperamos encontrar.

Para esse estudo consideramos orientações do Banco Mundial, mais detalhes em [Pesquisa \(2011\)](#), com objetivo de manter certa similaridade com outros estudos com mesmos objetivos, baseados nisso consideramos o nível de confiança como sendo 95%, o que significa pela forma da distribuição normal padrão que $Z = 1,96$. A margem de erro que admitimos será de $e = 0,037$ e como não se tem nenhuma informação sobre o valor que esperamos encontrar, usamos $p = 0,5$; o tamanho populacional será $N = 4.900$, ficando assim após a aplicação direta da equação (2.1), com $n = 614$.

Para a realização desse processo foram definidos 4 critérios de precisão: Região, dependência administrativa, porte da escola e a proficiência em matemática de 2009 no IDEB. Primeiramente construímos a variável Região, baseado em classificação colocada pelo IBGE (Instituto Brasileiro de Geografia e Estatística), o Estado de São Paulo está dividido em 15 regiões, são elas: Presidente Prudente, Araçatuba, Jales, São José do Rio Preto, Marília, Bauru, São Carlos, Ribeirão Preto, Registro, Itapetininga, Piracicaba, Campinas, Jundiaí, São José dos Campos, São Paulo.

No Apêndice B, temos a relação das regiões e os municípios que as compõe. Tendo como base critérios empíricos, consideramos sendo uma classe um conjunto de 40 alunos, convencionou-se que quando uma escola tem 2 turmas ou menos ela é uma escola pequena e acima de 5 turmas é uma escola grande, foi construída assim a variável PORTE_ESCOLAR, que será dada da seguinte forma:

- PEQUENO, se a escola tem menos de 80 alunos cursando o nono ano do ensino fundamental;
- MÉDIO, se a escola tem entre 80 e 160 alunos cursando o nono ano do ensino fundamental;
- GRANDE, se a escola tem mais de 160 alunos cursando o nono ano do ensino fundamental.

Baseado na proficiência em matemática da Prova Brasil do ano de 2009, foi construída também a variável QUARTIS onde as notas dos alunos são agrupadas em quatro grupos baseados em seus quartis, essa variável foi construída da forma:

- ABAIXO DO PERCENTIL 25, se a proficiência em matemática média de 2009 foi menor do que 8.785;
- ENTRE O PERCENTIL 25 E 50, se a proficiência em matemática média de 2009 estiver entre 8.785 e 9.807;
- ENTRE O PERCENTIL 50 E 75, se a proficiência em matemática média de 2009 estiver entre 9.807 e 10.910;
- ACIMA DO PERCENTIL 75, se a proficiência em matemática média de 2009 foi maior do que 10.910.

Para alocar as escolas segundo o tamanho amostral considerado, multiplicamos o número de escolas por região pela fração n/N . Na Tabela 2, temos o número de escolas a ser sorteadas por região mediante ao número de escolas por região.

Nas Tabelas 33 até 47, que podem ser encontradas no apêndice B, temos uma visão específica da amostra dos dados em cada região, aqui é importante ressaltar que as informações nas tabelas aqui referidas diferem da Tabela 2, por questões de arredondamento.

Tabela 2 – Demonstrativo do número de escolas por região amostrada

Região	Número de escolas	Quantidade amostrada
Presidente Prudente	119	15
Araçatuba	88	11
Jales	73	9
São José do Rio Preto	137	18
Marília	141	18
Bauru	181	23
São Carlos	104	13
Ribeirão Preto	269	34
Registro	87	11
Itapetininga	147	19
Piracicaba	162	21
Campinas	462	59
Jundiaí	327	42
São José dos Campos	321	41
São Paulo	2182	280

Após serem fixadas, os quatro critérios acima citados e alocados as 614 escolas sorteamos 20 alunos por escola e por fim temos uma amostra com 12.280 alunos, que constituem em 2.204 turmas, 614 escolas, 225 municípios e 15 regiões. Na Figura 3, temos um boxplot da proficiência em matemática em cada região, a partir desse comparativo há evidências de que não há diferença entre a proficiência em matemática nas diferentes regiões.

Na Tabela 32 temos as variáveis dicotomizadas, bem como as categorias resultantes, os códigos resultante (serão sempre 0 ou 1) e a frequência amostral. Na Tabela 31 temos a codificação das variáveis com mais de uma categoria, essa codificação é usada para a dicotomização das variáveis.

Nas Figuras 4 até 6, baseado nos dados amostrados, temos o boxplot das variáveis mencionadas na seção anterior, variáveis essas que foram mais utilizadas em outros estudos educacionais, aos observarmos esses gráficos há indícios de que não há diferença significativa entre grupos para cada uma das variáveis.

No próximo capítulo abordamos os aspectos teóricos relativos aos modelos multiníveis, são eles: formulação matemática, estimação dos parâmetros, testes de hipótese, entre outros.

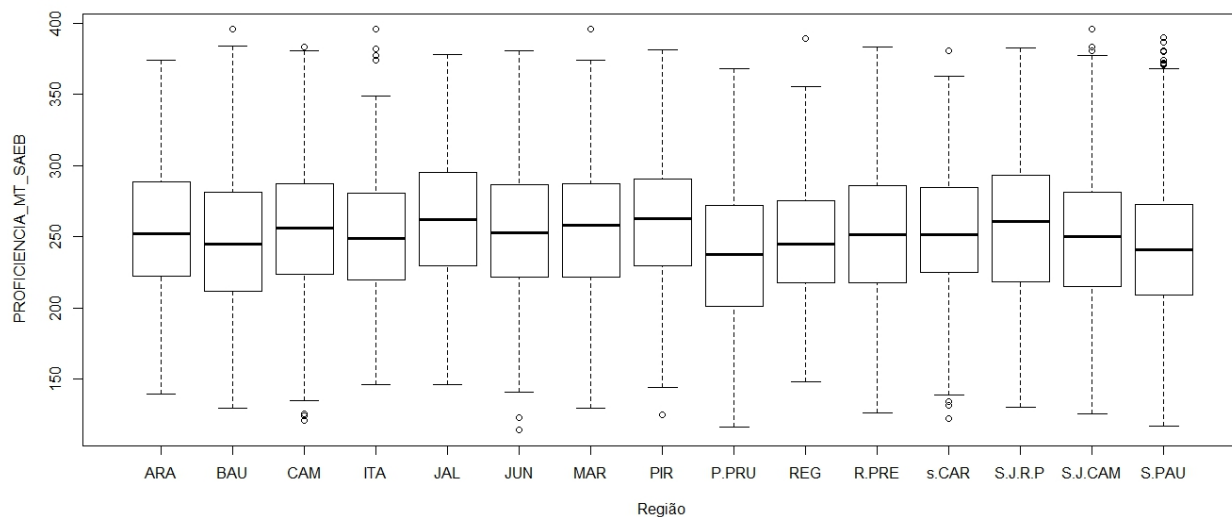


Figura 3 – Box plot da variável de interesse, PROFICIENCIA_MT_SAEB em cada uma das 15 regiões destacadas anteriormente.

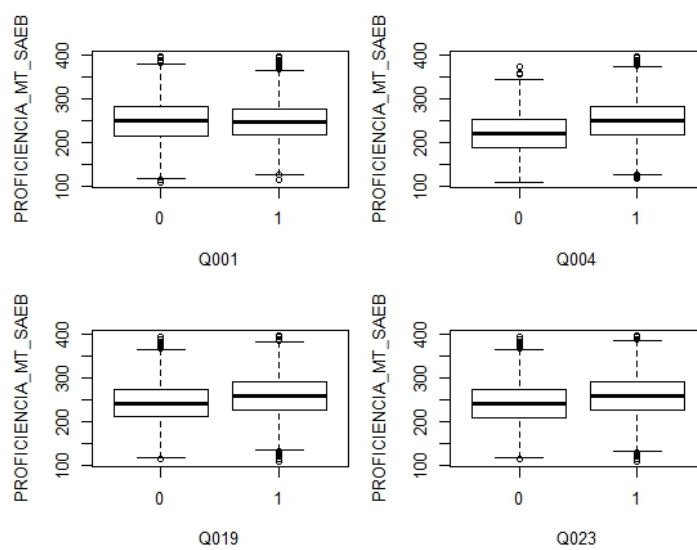


Figura 4 – Box plot da variável de interesse, PROFICIENCIA_MT_SAEB pelas covariáveis Q001, Q004, Q019 e Q023.

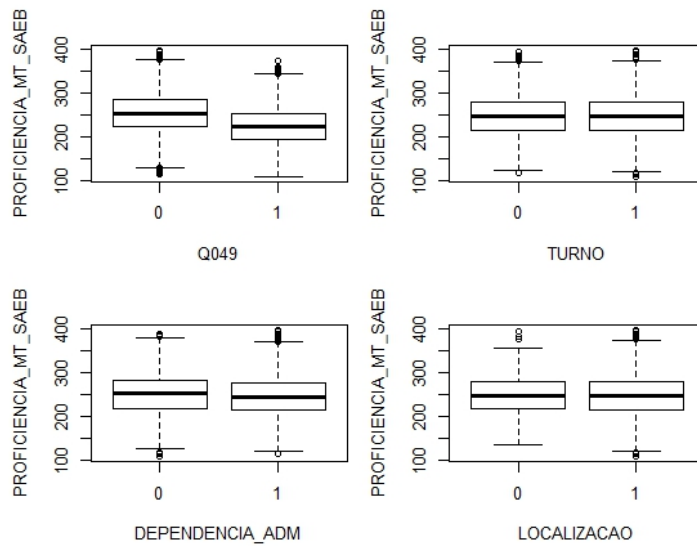


Figura 5 – Box plot da variável de interesse, PROFICIENCIA_MT_SAEB pelas covariáveis Q049, TURNO, DEPENDENCIA_ADM e LOCALIZACAO.

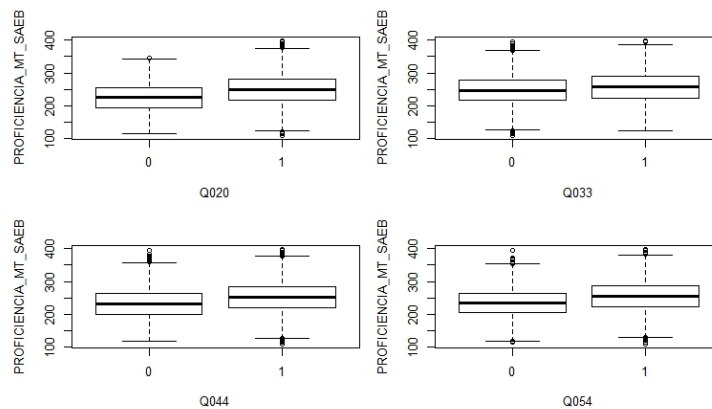


Figura 6 – Box plot da variável de interesse, PROFICIENCIA_MT_SAEB pelas covariáveis Q020, Q033, Q044 e Q054.

MODELO MULTINÍVEL COM RESPOSTA NORMAL

Neste capítulo apresentamos a estrutura geral dos modelos multiníveis, falamos a respeito da estimação de parâmetros dessa classe de modelos e testes de hipóteses. Partimos de uma abordagem bem simples que é um modelo de regressão linear simples para uma mais elaborada, com coeficientes aleatórios e por fim falamos a respeito do enfoque bayesiano.

Quando, em alguns casos, a independência das observações é violada, como o contexto escolar, as observações podem ser separadas em grupos (alunos, turmas, escolas, municípios, etc.), que possuem certas similaridades, que têm uma relevância considerável para o desempenho dos alunos, por exemplo, dois alunos de uma mesma turma podem compartilhar de características similares, por isso é usual incluir essa correlação no modelo. Uma alternativa para resolver esse problema seria o ajuste de um modelo de regressão multinível, tal modelo possibilitaria a correção de problemas relativos à estimação dos parâmetros, testes de hipóteses e intervalos de confiança.

Na Figura 7, apresentamos uma estrutura de três níveis, considerando a ótica de um problema com dados educacionais.

3.1 Modelos de regressão e suposições

No modelo de regressão linear simples, pode-se estudar a relação de uma variável explicativa e uma variável resposta, da seguinte forma:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad (3.1)$$

com $i = 1, 2, \dots, n$, onde:

- y_i representa a resposta do i -ésimo indivíduo;

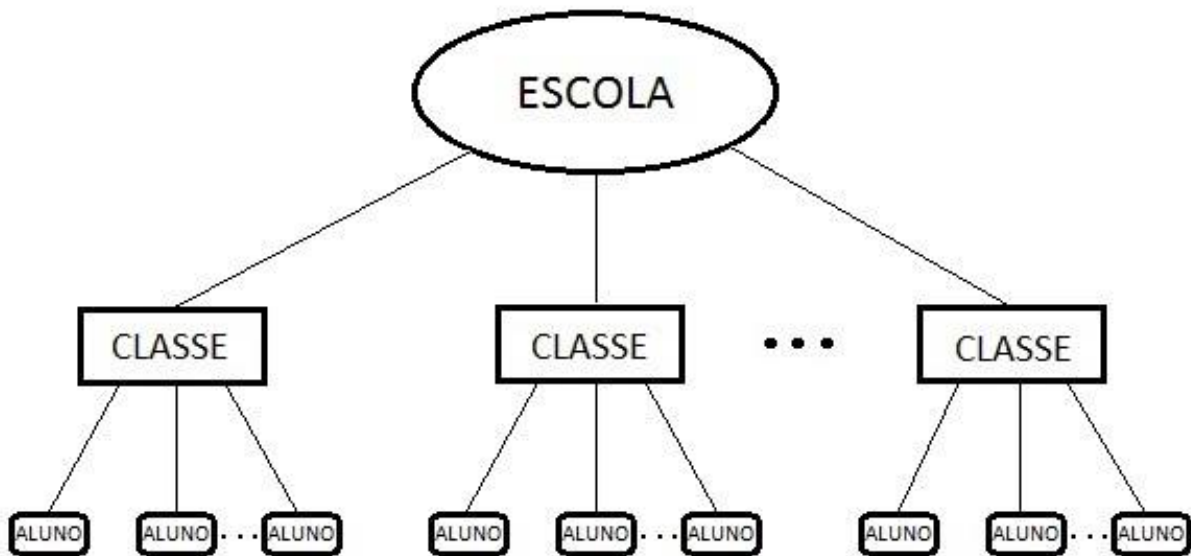


Figura 7 – Esquema representativo de uma estrutura hierárquica de três níveis.

- β_0 é o valor esperado da variável resposta y_i , quando x_i é zero;
- β_1 é a mudança esperada na variável resposta y_i , quando x_i aumenta uma unidade;
- x_i é a variável explicativa do i -ésimo indivíduo;
- $\varepsilon_i \sim N(0, \sigma^2)$ e todos os ε_i são independentes.

Na próxima seção fazemos a construção de modelos multiníveis, para dois níveis e também abordamos sua forma matricial.

3.2 Modelo de regressão com dois níveis

Em um exemplo de modelo de regressão com dois níveis, há n_j unidades do **nível 1**, para cada uma das j unidades do **nível 2**, com $j = 1, 2, \dots, J$. Assim no **nível 1**, o modelo será representado na forma:

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \varepsilon_{ij}, \quad (3.2)$$

com $i = 1, 2, \dots, n_j$ e $j = 1, 2, \dots, J$ em que:

- y_{ij} é a variável resposta do i -ésimo indivíduo do **nível 1**, do j -ésimo grupo;
- x_{ij} é a variável explicativa referente ao i -ésimo indivíduo do **nível 1**, agrupadas para a j -ésimo grupo do **nível 2**;
- β_{0j} é o intercepto para a j -ésimo grupo;

- β_{1j} é a inclinação associada a variável explicativa x_{ij} , da i -ésima unidade do **nível 1** para a j -ésimo grupo;
- ε_{ij} é o erro aleatório associado a i -ésima unidade do **nível 1**, do j -ésimo grupo do **nível 2**, com as suposições: $\varepsilon_{ij} \sim N(0, \sigma^2)$ e todos os ε_{ij} são independentes.

O intercepto β_{0j} é o valor esperado da variável resposta y_{ij} no j -ésimo grupo, quando x_{ij} for igual a zero, assim para o **nível 2** teremos J modelos iguais aos mostrados em (3.2), em que para cada modelo há interceptos aleatórios β_{0j} e inclinações aleatórias β_{1j} diferentes para $j = 1, 2, \dots, J$. Esses coeficientes podem ser modelados na seguinte forma:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + v_{0j}, \\ \beta_{1j} &= \gamma_{10} + v_{1j},\end{aligned}\tag{3.3}$$

onde:

- γ_{00} é o valor esperado dos interceptos nas unidades do **nível 2**;
- γ_{10} é o valor esperado das inclinações na população de **nível 2**;
- v_{0j} é o efeito aleatório no intercepto β_{0j} da j -ésima unidade do **nível 2**;
- v_{1j} é o efeito aleatório na inclinação β_{1j} da j -ésima unidade do **nível 2**.

Assumimos ainda as suposições: $v_{0j} \sim N(0, \tau_{00})$; $v_{1j} \sim N(0, \tau_{11})$; todos os v_{0j} e v_{1j} são independentes entre si e também de cada ε_{ij} correspondente. Logo podemos afirmar que $\beta_{0j} \sim N(\gamma_{00}, \tau_{00})$ e $\beta_{1j} \sim N(\gamma_{10}, \tau_{11})$.

- τ_{00} é a variância populacional dos interceptos;
- τ_{11} é a variância populacional das inclinações;
- τ_{01} é a covariância entre β_{0j} e β_{1j} .

Com o intuito de melhor explicar a variabilidade do **nível 2**, pode ser necessária a inclusão variáveis explicativas no respectivo **nível 2**, considere w_j essa variável, assim podemos reescrever as equações (3.3) da seguinte forma:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}w_j + v_{0j}, \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}w_j + v_{1j},\end{aligned}\tag{3.4}$$

em que:

- γ_{01} é o coeficiente associado a variável explicativa w_j do **nível 2**, referente ao intercepto do **nível 1**;
- γ_{11} é o coeficiente associado a variável explicativa w_j do **nível 2**, referente à inclinação do **nível 1**.

Aqui assumimos as suposições: $v_{0j} \sim N(0, \tau_{00})$; $v_{1j} \sim N(0, \tau_{11})$; todos os v_{0j} e v_{1j} são independentes entre si e também de cada ε_{ij} correspondente; $cov(v_{0j}, v_{1j}) = \tau_{01}$. Observe que τ_{00} , τ_{11} , τ_{01} são componentes de variância. Substituindo as equações (3.4) em (3.2), temos que:

$$y_{ij} = \gamma_{00} + \gamma_{01}w_j + \gamma_{10}x_{ij} + \gamma_{11}w_jx_{ij} + v_{0j} + v_{1j}x_{ij} + \varepsilon_{ij}, \quad (3.5)$$

Assim $y_{ij} = (\gamma_{00} + \gamma_{01}w_j + v_{0j}) + (\gamma_{10} + \gamma_{11}w_j + v_{1j})x_{ij} + \varepsilon_{ij}$. O modelo (3.5) leva em consideração diversos tipos de componentes, sejam eles de **nível 1**, **nível 2**, interações entre os níveis.

Os resultados que apresentamos nesta seção retratam um modelo com apenas uma covariável. De forma natural, eles podem ser estendidos para t covariáveis. Sejam x_1, x_2, \dots, x_r , covariáveis do **nível 1** e w_1, w_2, \dots, w_s , covariáveis de **nível 2**, assim o modelo com essas $t = r + s$ covariáveis será dado, no **nível 1** e **nível 2**, respectivamente, por:

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + \dots + \beta_{rj}x_{rij} + \varepsilon_{ij}, \\ \beta_{qj} &= \gamma_{q0} + \gamma_{q1}w_{1j} + \gamma_{q2}w_{2j} + \dots + \gamma_{qs}w_{sj} + v_{qj}, \end{aligned}$$

onde $i = 1, 2, \dots, n_j$, $j = 1, 2, \dots, J$ e $q = 1, 2, \dots, p$.

Aqui, temos suposições que:

- $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$;
- $\beta_{qj} \sim N(0, \sigma_v^2)$;
- todos os β_{qj} são independentes entre si e com relação aos ε_{ij} .

Na próxima seção apresentamos técnicas de estimação dos parâmetros dos modelos aqui apresentados.

3.3 Estimação de parâmetros

O método de máxima verossimilhança é a maneira mais utilizada para obter as estimativas dos coeficientes nos modelos multiníveis, pois tem a vantagem de produzir estimativas que são eficientes, assintoticamente eficientes e consistentes, quando as condições de regularidade estão satisfeitas, por exemplo, o parâmetro β é identificável, $f(\mathbf{y}|\beta)$ é diferenciável em β , note ainda

que o espaço paramétrico contém um conjunto aberto Ω , do qual o verdadeiro valor do parâmetro está no interior de Ω , a densidade $f(\mathbf{y}|\boldsymbol{\beta})$ é três vezes diferenciável em relação à $\boldsymbol{\beta}$, mais detalhes ver [Casella e Berger \(2002\)](#).

No método de máxima verossimilhança, podem ser usadas duas funções:

- **verossimilhança completa:** coeficientes da regressão e os componentes da variância são incluídos na função de verossimilhança.
- **verossimilhança restrita:** os componentes da variância são incluídos na função de verossimilhança e os coeficientes da regressão são estimados na segunda etapa.

A verossimilhança completa não pode ser maximizada analiticamente, para isso podem-se utilizar alternativas como o algoritmo de escore de Fisher, descrito com detalhes em [Osio \(2013\)](#). Quando trabalhamos com a verossimilhança restrita, os estimadores REML podem ser obtidos, segundo [Patterson e Thompson \(1971\)](#), dividindo-se a função de verossimilhança em duas partes, uma referente aos efeitos fixos e outra aos efeitos aleatórios.

Computacionalmente, a estimação de parâmetros é feita utilizando os pacotes lme4 ou nlme, mais detalhes sobre estes pacotes podem ser vistos em [Bates \(2010\)](#) e [Bliese \(2006\)](#), respectivamente.

3.4 Testes de Hipóteses

O modelo com diversos níveis, deve ser testado em relação a muitos aspectos, um primeiro problema a ser testado é em que nível uma determinada variável deve ser inserida, para obter uma resposta a essa questão deve levar em consideração o contexto em que o banco de dados está inserido e o objetivo da análise.

A hipótese de interesse neste caso é da forma $H_0 : \gamma_k = 0$, onde γ_k é um elemento do vetor de efeitos fixos. A estatística de teste é obtida considerando o estimador de máxima verossimilhança restrita para o erro padrão, dado por:

$$t = \frac{\hat{\gamma}_k}{\sqrt{\hat{m}_k}}, \quad (3.6)$$

onde m_k é a variância de $\hat{\gamma}_k$. A estatística dada em (3.6) segue uma distribuição t-student para dados balanceados e para algumas situações de desbalanceamento, para saber mais a respeito dos graus de liberdade ver, por exemplo [Sullivan, Dukes e Losina \(1999\)](#). Para testar os componentes de variância da matriz D , a hipótese de interesse é da forma $H_0 : \tau_{kl} = 0$. Ressaltamos que pelo fato da hipótese de interesse estar alocada na fronteira do espaço paramétrico, não é possível encontrar um conjunto aberto Ω , que contenha a hipótese $\tau_{kl} = 0$ e ainda esteja contido no

espaço paramétrico. Desta forma não é tarefa fácil realizar teste de hipótese para componentes de variância, Russo, Aoki e Paula (2012) traz uma alternativa viável para solucionar esse problema.

Em Sullivan, Dukes e Losina (1999), há outros testes que também podem ser feitos para efeitos aleatórios. Neste caso a hipótese de interesse será da por $H_0 : v_k = 0$, onde v_k é um elemento do vetor de efeitos aleatórios, segundo Sullivan, Dukes e Losina (1999), tal estatística é obtida levando em consideração a estimativa do efeito aleatório e erro padrão, ela será da forma:

$$t = \frac{\hat{v}_k}{\sqrt{\hat{r}_k}}, \quad (3.7)$$

onde r_k é a variância de \hat{v}_k .

Assim como no caso dos efeitos fixos, a estatística dada em (3.7) segue uma distribuição t-student para dados balanceados e para algumas situações de desbalanceamento.

Na próxima seção apresentamos uma breve introdução de como são tratados os modelos multiníveis bayesianos.

3.5 Enfoque bayesiano

Quando se trata de um modelo de regressão usual, o problema se resume a aplicação do teorema de Bayes para a obtenção de uma distribuição *a posteriori* do parâmetro de interesse, quando assumimos distribuição *a priori* uniforme (proporcional a uma constante), a distribuição *a posteriori* é igual à verossimilhança. Quando se trata de uma regressão multinível, Gelman e Hill (2006) considera um modelo simples sem preditores, com um nível, dado por:

$$\begin{aligned} y_{ij} &\sim N(\beta_{0j}, \sigma_\varepsilon^2), \\ \beta_{0j} &\sim N(\gamma_{00}, \sigma_v^2), \end{aligned} \quad (3.8)$$

para $i = 1, 2, \dots, n_j$ e para $j = 1, \dots, J$. Se pensarmos em estimativas clássicas todos os elementos do segundo nível são considerados equivalentes, isto é, $\beta_{01} = \beta_{02} = \dots = \beta_{0J} = \gamma_{00}$, desta forma o modelo (3.8) se reduz a $y_{ij} \sim N(\gamma_{00}, \sigma_\varepsilon^2 + \sigma_v^2)$, para qualquer observação de y . Assim, para cada um dos j indivíduos do **nível 2**, teremos que $\hat{\beta}_{0j} = \gamma_{00}$.

No modelo multinível, cada elemento do nível é estimado sozinho, de modo que, cada β_{0j} é dado por \bar{y}_{ij} , a média das observações de um dos j indivíduos no **nível 2**. O modelo (3.8) têm coeficientes da regressão ao **nível 2**, dado por $\beta_{01} = \beta_{02} = \dots = \beta_{0J}$ e hiperparâmetros γ_{00} , σ_ε^2 e σ_v^2 , a estimativa multinível é obtida supondo que os hiperparâmetros são conhecidos e daí é feita a estimação de cada β_{0j} , observe que a estimativa estará compreendida entre \bar{y}_{ij} e γ_{00} ,

isto é, se considerarmos os hiperparâmetros conhecidos teremos que:

$$\beta_{0j}|y, \gamma_{00}, \sigma_{\varepsilon}^2, \sigma_v^2 \sim N(\hat{\gamma}_{00}, V_j), \quad (3.9)$$

para $j = 1, \dots, J$. Assim a estimativa e a variância da estimativa são, respectivamente, dadas por:

$$\hat{\beta}_{0j} = \frac{\frac{n_j}{\sigma_{\varepsilon}^2} \bar{y}_{ij} + \frac{1}{\sigma_v^2} \gamma_{00}}{\frac{n_j}{\sigma_{\varepsilon}^2} + \frac{1}{\sigma_v^2}}, V_j = \frac{1}{\frac{n_j}{\sigma_{\varepsilon}^2} + \frac{1}{\sigma_v^2}}, \quad (3.10)$$

onde n_j é o tamanho do grupo j no **nível 2**. Note que a estimativa $\hat{\beta}_{0j}$ pode ser interpretada como a média ponderada entre \bar{y}_{ij} e γ_{00} . Para o grupo j para qual $n_j = \sigma_{\varepsilon}^2/\sigma_v^2$, teremos que $\hat{\beta}_{0j} = 0,5\bar{y}_{ij} + 0,5\gamma_{00}$.

Assim tendo estimado os coeficientes da regressão no **nível 2**, segundo [Gelman e Hill \(2006\)](#), podemos estimar os hiperparâmetros, a variância dos dados σ_{ε}^2 pode ser estimada como:

$$\hat{\sigma}_{\varepsilon}^2 = \frac{1}{n} \sum_{j=1}^J \sum_{i=1}^n (y_{ij} - \beta_{0j})^2. \quad (3.11)$$

A média γ_{00} para o **nível 2**, que pode ser estimado por:

$$\hat{\gamma}_{00} = \frac{1}{J} \sum_{j=1}^J \beta_{0j}, \quad (3.12)$$

com uma variância estimada de σ_v^2/J , como pode-se notar que na equação (3.12), temos que a estimativa $\hat{\gamma}_{00}$ é a média dos interceptos β_{0j} .

A variância no **nível 2**, pode ser estimado por:

$$\hat{\sigma}_v^2 = \frac{1}{J} \sum_{j=1}^J (\beta_{0j} - \hat{\gamma}_{00})^2. \quad (3.13)$$

Para caso com intercepto aleatório, podemos modelar a variável resposta como:

$$\begin{aligned} y_{ij} &\sim N(\beta_{0j} + \sum_{k=1}^m \beta_m x_{mij}, \sigma_{\varepsilon}^2), \\ \beta_{0j} &\sim N(\gamma_{00} + \gamma_{10} w_j, \sigma_v^2), \end{aligned} \quad (3.14)$$

onde, m é o número de variáveis explicativas no **nível 1**, w_j é uma variável explicativa inserida no **nível 2**, $\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon}^2)$ e $v_{0j} \sim N(0, \sigma_v^2)$ são os erros aleatórios respectivamente do **nível 1** e

do **nível 2**. Segundo [Gelman e Hill \(2006\)](#), os β_{0j} têm distribuições *a priori* distintas e γ_{00} , γ_{10} e w_j são estimados a partir dos dados e suas distribuições *a priori* são não informativas.

No processo de estimação dos parâmetros numa perspectiva bayesiana feita ao longo deste trabalho, com exceção dos efeitos aleatórios, no qual assumimos como tendo uma distribuição normal *a priori* com média 0 e variância σ_v^2 , assumimos distribuições *a priori* proporcionais a uma constante para todos os outros hiperparâmetros.

Se considerarmos as inclinações aleatórias podemos reescrever (3.14), como

$$y_{ij} \sim N(\beta_{0j} + \sum_{k=1}^m \beta_{kj} x_{kij}, \sigma_\varepsilon^2),$$

$$\begin{pmatrix} \beta_{0j} \\ \beta_{1j} \\ \dots \\ \beta_{mj} \end{pmatrix} \sim N_{m+1} \left(\begin{pmatrix} \gamma_{00} + \gamma_{01} w_j \\ \gamma_{10} + \gamma_{11} w_j \\ \dots \\ \gamma_{m0} + \gamma_{m1} w_j \end{pmatrix}, \begin{pmatrix} \sigma_{v_{0j}}^2 & \rho_{01} \sigma_{v_{0j}} \sigma_{v_{1j}} & \dots & \rho_{0m} \sigma_{v_{0j}} \sigma_{v_{mj}} \\ \rho_{10} \sigma_{v_{1j}} \sigma_{v_{0j}} & \sigma_{v_{1j}}^2 & \dots & \rho_{1m} \sigma_{v_{1j}} \sigma_{v_{mj}} \\ \vdots & \dots & \ddots & \vdots \\ \rho_{m0} \sigma_{v_{mj}} \sigma_{v_{0j}} & \rho_{m1} \sigma_{v_{mj}} \sigma_{v_{1j}} & \dots & \sigma_{v_{mj}}^2 \end{pmatrix} \right), \quad (3.15)$$

onde ρ_{tu} é a correlação entre β_{tj} e β_{uj} , para $t = 0, 1, \dots, m$ e $u = 0, 1, \dots, m$.

Com relação as distribuições *a priori*, tanto dos parâmetros quanto dos hiperparâmetros, seguem a mesma diretriz dos casos anteriores, conforme é colocado por [Gelman e Hill \(2006\)](#). A estimação dos modelos propostos nesta seção é feita utilizando o pacote RStan, que é uma adaptação da interface Stan, adaptada para ser utilizada na linguagem R. O stan estima modelos bayesianos utilizando um processo Monte Carlo Hamiltoniano (HMC), que é uma adaptação de processo de Monte Carlo cadeia de Markov.

No apêndice D, falamos sobre o processo de HMC e também trazemos os códigos para estimação de modelos de regressão multinível bayesianos. No próximo capítulo falamos a respeito de técnicas para escolhas de modelos, alguns critérios de informação e seleção de variáveis são descritos.

MÉTODOS DE ESCOLHA DE MODELOS

Neste capítulo discutimos alguns métodos de seleção de modelos de regressão multinível, sob o enfoque clássico e também sob o enfoque bayesiano e ainda falamos sobre o critério de seleção de modelos Lasso, para posteriormente serem aplicados a dados educacionais. Falamos a respeito dos critérios de informação que geralmente são os mais usados, para modelos clássicos e também para modelos bayesianos, abordamos as seguintes medidas de informação: AIC, AICc, BIC, DIC, *loic*, WAIC, EAIC e EBIC.

4.1 Seleção de Modelos: Caso Clássico

O critério mais utilizado na atualidade é o AIC (Akaike Information Criterion), proposto por [Akaike \(1998\)](#). Esse método foi construído baseado na distância de Kullback-Leiber, sendo assim ele pode ser enxergado como sendo a distância entre um modelo verdadeiro e um modelo candidato. É aconselhável utilizar AIC apenas para uma amostra de tamanho superior a 40 vezes o número de parâmetros do modelo. Diante de tais dificuldades, [Sugiura \(1978\)](#) propôs uma correção para o critério AIC, o AICc que leva em consideração o tamanho da amostra. É fácil ver que o AIC e o AICc coincidem assintoticamente. Ao contrário desses métodos o BIC (Bayesian Information Criterion), proposto por [Schwarz et al. \(1978\)](#), é um método consistente, ou seja, ele escolhe o modelo correto, desde que ele esteja entre aqueles que estão sendo testados. [Emiliano \(2009\)](#) faz algumas considerações acerca desses métodos e ressalta que:

- tanto o AIC quanto o BIC fundamentam-se no logaritmo da função de verossimilhança, impondo entretanto diferentes penalizações;
- o AIC e o BIC podem ser utilizados não somente para modelos encaixados, como também para modelos não encaixados;

- para $n > 8$, o valor do AIC para um determinado modelo será sempre menor que o valor do BIC, mas os resultados não necessariamente serão menores;
- o AIC e o BIC servem para comparar quaisquer quantidades de modelos, e não somente dois;
- o AIC e o BIC são critérios assintóticos e já existem correções para esses;
- o AIC e o BIC servem para estudar estruturas de covariâncias;
- a seleção de modelos é feita pelo pesquisador e, se somente modelos ruins forem selecionados, o AIC fará a seleção do melhor dentre eles.

A seguir, descrevemos os critérios mencionados.

4.1.1 AIC (Akaike Information criterion)

O critério AIC, proposto por Akaike (1998), é dado por:

$$AIC = -2 \sum_{i=1}^n \ln(\hat{\mathbf{y}}_{ij}, \mathbf{y}_{ij}) + 2m, \quad (4.1)$$

onde \mathbf{y}_{ij} é o i -ésimo valor da variável resposta para o j -ésimo grupo, $\hat{\mathbf{y}}_{ij}$ é o valor predito de \mathbf{y}_{ij} , $\sum_{i=1}^n \ln(\hat{\mathbf{y}}_{ij}, \mathbf{y}_{ij})$ é o logaritmo da função de verossimilhança do modelo e m é o número real de parâmetros. O AIC pode ser enxergado como a distância entre o “modelo verdadeiro” e o “modelo candidato”. Ele surgiu no intuito de escolher o melhor modelo de regressão, quando temos um modelo completo e um modelo reduzido, uma vez que um modelo de ajuste ineficiente pode não trazer uma modelagem estatística adequada, fazendo assim que não seja possível enxergar a verdadeira natureza da variabilidade dos dados. Outro problema que pode ser encontrado é a perda de generalidade do modelo.

O critério AIC é então uma maneira de selecionar o modelo que melhor equilibra esses inconvenientes. Sempre que um melhor modelo é selecionado, outras análises podem ser feitas. É importante destacar que o modelo selecionado é aquele que tem o menor AIC, ressaltando sempre a importância de todos os modelos candidatos estarem bem estimados.

4.1.2 AICc (AIC Corrigido)

Uma das deficiências encontradas ao se utilizar o critério de informação AIC, é que quando se tinha pequenas amostras, há uma situação de *overfit*. Isto se deve ao fato de que o critério de informação de Akaike foi construído levando em consideração propriedades assintóticas,

sendo assim [Sugiura \(1978\)](#) propôs uma correção para este critério de informação, sua expressão pode ser obtida da seguinte forma:

$$AICc = -2 \sum_{i=1}^n \ln(\hat{\mathbf{y}}_{ij}, \mathbf{y}_{ij}) + \left(\frac{2n}{n-p-1} \right) m, \quad (4.2)$$

onde n é o tamanho amostral. Observe que a medida que n cresce os critérios de informação AIC e AICc se aproximam, assim ao comparar modelos escolhemos sempre o modelo com o menor AICc.

4.1.3 BIC (Bayesian Information Criterion)

O Critério BIC, proposto por [Schwarz et al. \(1978\)](#), é dado por:

$$BIC = -2 \sum_{i=1}^n \ln(\hat{\mathbf{y}}_{ij}, \mathbf{y}_{ij}) + m \log(n), \quad (4.3)$$

onde \mathbf{y}_{ij} é o i -ésimo valor da variável resposta para o j -ésimo grupo, $\hat{\mathbf{y}}_{ij}$ é o valor predito de \mathbf{y}_{ij} , m é o número de parâmetros, $\sum_{i=1}^n \ln(\hat{\mathbf{y}}_{ij}, \mathbf{y}_{ij})$ é o logaritmo da função de verossimilhança do modelo e n é o número de observações.

Um das características mais interessantes desse critério é que ele é baseado no cálculo do logaritmo da função de verossimilhança não sendo necessário a especificação de distribuições *a priori*. Um dos critérios de comparação de modelos sob enfoque bayesiano é o fator de Bayes, mais detalhes sobre esse critério, ver por exemplo [Missão \(2010\)](#). Nota-se que em certos casos a seleção de modelos utilizando o fator de Bayes e o BIC são equivalentes, desta forma o BIC é usado em muitos problemas de modelagem bayesiana onde as distribuições *a priori* são difíceis de definir com precisão.

Assim como no caso do AIC, o BIC penaliza modelos com muitas variáveis, sendo que valores menores são preferíveis.

4.2 Seleção de Modelos: Caso Bayesiano

Entre os critérios Bayesianos apresentados, o que é mais utilizado é o DIC (Deviance Information Criterion), proposto por [Spiegelhalter et al. \(2002\)](#). Ele foi construído baseado no número efetivo de parâmetros, essa medida leva em consideração a complexidade do modelo, e talvez essa seja sua maior deficiência, pois se o modelo em questão contiver uma quantidade muito elevada de parâmetros, a medida número efetivo de parâmetros será negativa.

[Spiegelhalter et al. \(2014\)](#) elenca as críticas mais recorrentes a esse critério de informação, são elas:

- o número efetivo de parâmetros é variante a parametrizações, falta de consistência;

- não é baseado em um critério preditivo adequado;
- tem uma justificativa teórica fraca.

Ele também traz algumas iniciativas que têm sido feitas com intuito de melhorar o DIC, como por exemplo, uma medida de complexidade alternativa. Contudo nos últimos anos têm se popularizado o critério de informação WAIC, proposto por [Watanabe \(2010\)](#), ele é visto como uma melhoria do DIC para modelos bayesianos. O WAIC foi construído totalmente numa abordagem bayesiana que se assemelha a um processo de validação cruzada leave-one-out. Ele também é invariante a parametrização, ao contrário do DIC. O WAIC e a validação cruzada leave-one-out, são critérios baseados no logaritmo da função densidade de probabilidade *a posteriori* preditiva. Este trabalho utiliza ainda os critérios EAIC e EBIC, que assim como o DIC, baseiam-se no desvio médio. É importante salientar que esses dois últimos critérios de informação são, o valor esperado dos critérios AIC e BIC.

A seguir descrevemos os critérios de seleção de modelos bayesianos.

4.2.1 DIC (Deviance Information Criterion)

O critério de informação do desvio foi proposto por [Spiegelhalter et al. \(2002\)](#) como uma generalização dos critérios AIC e BIC, ele usa em seu cálculo uma medida que leva em consideração a complexidade do modelo. Uma vantagem é que ele é facilmente calculado em amostras geradas via MCMC, o DIC é definido de maneira análoga ao AIC, da seguinte forma:

$$DIC = p_D + \bar{D}, \quad (4.4)$$

onde $\bar{D} = E_{\boldsymbol{\theta}}(D(\boldsymbol{\theta}))$ é o desvio médio esperado, ou seja, mede o quão bem o modelo se ajusta aos dados. Note que $D(\boldsymbol{\theta}) = -2\log p(\mathbf{y}|\boldsymbol{\theta})$ é o desvio onde \mathbf{y} são os dados, $\boldsymbol{\theta}$ são os parâmetros desconhecidos do modelo e $p(\mathbf{y}|\boldsymbol{\theta})$ é função de verossimilhança, temos também que $p_D = \bar{D} - D(\hat{\boldsymbol{\theta}})$, onde $\hat{\boldsymbol{\theta}}$ é a estimativa pontual da distribuição *a posteriori* dos parâmetros de interesse. Assim como no critério AIC, o modelo com o menor DIC deve ser selecionado. No caso específico de modelos multiníveis, [Gelman e Hill \(2006\)](#) nos fala de certa instabilidade no cálculo do número efetivo de parâmetros sugerindo que, nesta classe de modelos de regressão, o DIC não seja usado como critério definitivo.

4.2.2 WAIC (Widely Applicable Information Criterion)

O critério WAIC proposto por [Watanabe \(2010\)](#) pode ser visto como uma melhoria do critério informação desvio (DIC) para modelos bayesianos, pois o DIC produz estimativas pontuais que em alguns casos podem ser negativos, já o WAIC utiliza distribuições *a posteriori*, o que o torna invariante à parametrizações. Esse critério baseia-se também numa medida conhecida como o número efetivo de parâmetros, que pode ser interpretado segundo [Vehtari, Gelman e](#)

Gabry (2016) como o grau de complexidade do modelo. Lembrando que o critério de informação desvio, proposto por Spiegelhalter *et al.* (2002), é obtido pela expressão $DIC = 2\bar{D} - D(\bar{\theta})$, onde $\bar{D} = E_{\theta}(D(\theta))$ e $D(\theta) = -2\log(p(\mathbf{y}|\theta)) + C$, note que C é uma constante desconhecida.

Para a construção da medida em questão, Vehtari, Gelman e Gabry (2016) define uma medida de precisão preditiva esperada para um novo conjunto de observações, que será dada por:

$$elpd = \sum_{i=1}^{n_j} \sum_{j=1}^J \int p_t(\tilde{y}_{ij}) \log p(\tilde{y}_{ij}|\mathbf{y}) d\tilde{y}_{ij}, \quad (4.5)$$

onde $p_t(\tilde{y}_{ij})$ é função de distribuição preditiva para novas observações de \tilde{y}_{ij} e n_j é o número de observações do grupo j . Lembrando sempre que em inferência bayesiana temos que a distribuição preditiva *a posteriori* será dada por:

$$p(\tilde{y}|\mathbf{y}) = \int p(\tilde{y}_{ij}|\theta) p(\theta|\mathbf{y}) d\theta, \quad (4.6)$$

onde $p(\theta|\mathbf{y})$ é a distribuição *a posteriori* e $p(\theta)$ é a distribuição *a priori*.

Define-se também, conforme pode ser encontrado em Vehtari, Gelman e Gabry (2016), o logaritmo da função densidade preditiva, que pode ser obtida pela seguinte expressão:

$$lpd = \sum_{i=1}^{n_j} \sum_{j=1}^J \log p(y_{ij}|\mathbf{y}) = \sum_{i=1}^{n_j} \sum_{j=1}^J \log p(y_{ij}|\theta) p(\theta|\mathbf{y}) d\theta. \quad (4.7)$$

Na prática, para estimar lpd usaremos métodos de reamostragem, para a obtenção de simulações da distribuição *a posteriori*, usamos $\theta^s, s = 1, 2, \dots, S$, assim teremos que:

$$l\hat{p}d = \sum_{i=1}^{n_j} \sum_{j=1}^J \log\left(\frac{1}{S} \sum_{s=1}^S p(y_{ij}|\theta^s)\right). \quad (4.8)$$

Segundo Watanabe (2010), podemos estimar a medida log de precisão preditiva esperada da seguinte forma:

$$el\hat{p}d_{waic} = l\hat{p}d - \hat{p}_{waic}, \quad (4.9)$$

onde \hat{p}_{waic} é o número de parâmetros efetivos estimados e pode ser calculado usando a variância *a posteriori* da distribuição log da densidade preditiva para cada ponto y_{ij} e então, somando todos os pontos obtemos, segundo Vehtari, Gelman e Gabry (2016), que:

$$\hat{p}_{waic} = \sum_{i=1}^{n_j} \sum_{j=1}^J V_{s=1}^S(\log(y_{ij}|\theta^s)), \quad (4.10)$$

onde $V_{s=1}^S$ representa a variância amostral. Se multiplicarmos a medida obtida na equação (4.9) por -2 , ela pode ser utilizada para comparar os modelos ajustados, isto é, dados dois modelos (modelo A e modelo B), podemos estimar dois valores $-2elpdA_{waic}$ e $-2elpdB_{waic}$, das duas medidas aquela que for menor indicará o melhor modelo.

4.2.3 *loo (Leave-one-out cross-validation)*

Assim como o WAIC, o *loo* (leave-one-out cross-validation) é um método que utiliza o logaritmo da função de verossimilhança avaliada nas simulações da distribuição *a posteriori* dos parâmetros, apresentamos este conceito via amostragem de importância utilizando a distribuição de Pareto generalizada.

A amostragem por importância consiste em forçar a seleção de um maior número de amostras nas partes mais importantes do problema (eventos que mais contribuem para os parâmetros a serem inferidos), isto é, faz um deslocamento da distribuição para áreas de maior interesse. Esta distorção se faz introduzindo uma nova função distribuição e os valores devem ser corrigidos por um fator peso para não alterar os resultados esperados.

Segundo [Vehtari, Gelman e Gabry \(2016\)](#), podemos estimar a medida log de precisão preditiva *loo* da seguinte forma:

$$elpd_{loo} = \sum_{i=1}^{n_j} \sum_{j=1}^J \log p(y_{ij}|y_{ij-1}), \quad (4.11)$$

onde $p(y_{ij}|y_{ij-1}) = \int p(y_{ij}|\boldsymbol{\theta})p(\boldsymbol{\theta}|y_{ij-1})d\boldsymbol{\theta}$ é a densidade preditiva leave-one-out dada a amostra sem o i -ésimo ponto dos dados para o j -ésimo grupo. [Vehtari, Gelman e Gabry \(2016\)](#) propõe uma abordagem utilizando pesos de importância alisados por uma distribuição generalizada de Pareto, ou seja, encaixa-se a essa distribuição na cauda, que consiste nos 20% de maior importância.

Segundo [Vehtari, Gelman e Gabry \(2016\)](#), esse procedimento baseia-se nos seguintes passos:

- ajustar uma distribuição de Pareto para os raios de importância maiores do que 20%, esses raios podem ser obtidos pela expressão:

$$\frac{1}{p(y_{ij}|\boldsymbol{\theta}^s)}; \quad (4.12)$$

- estabilize os raios de importância substituindo os raios maiores do que certo M , pelos valores esperados das estatísticas de ordem da distribuição de Pareto ajustada $G^{-1}\left(\frac{z-0.5}{M}\right)$, onde $z = 1, 2, \dots, M$, onde z é o número de simulações usadas no ajuste da distribuição de Pareto generalizada e G^{-1} é a inversa acumulada da distribuição de Pareto e denote $(w)_{ij}^s$, onde s é o índice de da simulação e ij é o índice associado às observações.

- para garantir a variância finita das estimativas, vamos reescrever o vetor de pesos \mathcal{S} , como $\mathcal{S}^{\frac{3}{4}}\bar{w}$ onde \bar{w} é a média dos pesos suavizados.

Por fim, [Vehtari, Gelman e Gabry \(2016\)](#), ressalta ainda que as etapas anteriormente descritas devem ser realizadas para cada observação e o resultado é um vetor de pesos melhor comportados dos raios de importância bruta descritas acima.

Para analisar a confiabilidade das estimativas é observado o parâmetro de forma da distribuição de Pareto generalizada. Se o parâmetro de forma for menor que 0,5, a variância dos raios de importância bruta é finita e as estimativas convergem rapidamente. Se o parâmetro de forma estiver entre 0,5 e 1, a variância dos raios de importância é infinita, a média existe e assim o teorema do limite central garante a convergência das estimativas. Se o parâmetro de forma for maior que 1, a média dos pesos não existe, sinalizando assim que nesse caso as estimativas *loo* não são confiáveis, daí seria necessário o uso de outras técnicas. Portanto, a estimativa leave-one-out pode ser calculada pela expressão:

$$el\hat{p}d_{loo} = \sum_{i=1}^{n_j} \sum_{j=1}^J \log \left(\frac{\sum_{s=1}^S w_i^s p(y_{ij} | \boldsymbol{\theta}^s)}{\sum_{s=1}^S w_i^s} \right). \quad (4.13)$$

Assim, o critério de informação leave-one-out (*looic*) será dado por:

$$looic = -2el\hat{p}d_{loo}. \quad (4.14)$$

O modelo que deve ser escolhido é aquele que tem o menor *looic*. Outros critérios de informação que podem ser calculados utilizando saídas MCMC, em particular o EAIC e o EBIC.

4.2.4 EAIC (Expected Akaike Information Criterion)

O EAIC, foi proposto por [Brooks et al. \(2002\)](#), como sendo o valor esperado do AIC, sua expressão é obtida da seguinte forma:

$$E\hat{AIC} = \bar{D} + 2m, \quad (4.15)$$

onde, \bar{D} é o desvio médio e m é a quantidade real de parâmetros do modelo em ajuste. Quando estamos comparando dois modelos, consideramos o modelo mais parcimonioso, aquele que possui o menor EAIC.

4.2.5 EBIC (Expected Bayesian Information Criterion)

O EBIC, foi proposto por [Carlim e Louis \(2001\)](#), ele é obtido calculando a esperança matemática da expressão (4.3), sua expressão é dada por:

$$E\hat{BIC} = \bar{D} + m \log(n), \quad (4.16)$$

onde, \bar{D} é o desvio médio, n é o tamanho amostral e m é a quantidade real de parâmetros do modelo em ajuste. Novamente, consideramos o modelo mais preciso, o modelo com o menor EBIC.

4.3 Seleção de Variáveis

Outro método que discutimos é o Lasso, que, ao contrário de outros métodos, seleciona variáveis e não modelos, o que de um ponto de vista prático é muito interessante quando se tem uma grande quantidade de preditores, pois se fossem utilizados critérios de seleção de modelos, seria necessário testar todas as combinações possíveis de modelos, o que seria imensamente trabalhoso quando se tem muitas variáveis explicativas. Utilizar o Lasso significa impor uma penalização a função de verossimilhança do modelo, e obter as estimativas do modelo, maximizando essa função, assim alguns coeficientes do modelo serão estimados como sendo zero.

4.3.1 Lasso

O estimador lasso foi proposto inicialmente por [Tibshirani \(1996\)](#) para **modelos de regressão linear simples**, ele minimiza a soma de quadrados das diferenças entre o resultado observado e o modelo ajustado, desde que a soma em módulo dos coeficientes seja inferior a algum valor s , ou seja, tal método consiste em utilizar a soma

$$S = \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ji})^2$$

para estimar $\hat{\beta}_0$, $\hat{\beta}$, onde $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_p)^T$, sendo que S está sujeito a restrição $\sum_{j=1}^p |\beta_j| \leq g$, onde $g \geq 0$ é o parâmetro de restrição, associado ao ajuste deste modelo. Obtemos para $n = 2$ estimativas de β_1 e β_2 em função das estimativas de mínimos quadrados, assumindo que elas são ambas positivas.

Nesta seção falamos a respeito desse estimador para modelos lineares multiníveis com dois níveis (mistos) com resposta normal, primeiramente fazemos uma breve introdução dessa classe de modelos bem como de sua notação, posteriormente apresentamos o algoritmo do gradiente ascendente, proposto por [Groll e Tutz \(2014\)](#).

4.3.2 Modelos lineares com dois níveis e resposta Normal

Considere um modelo de regressão, onde a i -ésima observação do j -ésimo grupo é obtida através da expressão:

$$y_{ij} = \mathbf{x}_{ij}^\top \boldsymbol{\beta} + \mathbf{w}_{ij}^\top \mathbf{v}_j + \varepsilon_{ij} = \eta_{ij}^{param} + \eta_{ij}^{aleat} + \varepsilon_{ij}, \quad (4.17)$$

onde:

- y_{ij} é a i -ésima observação no j -ésimo grupo, com $i = 1, 2, \dots, n_j$ e $j = 1, 2, \dots, J$;
- $\mathbf{x}_{ij}^\top = (1, x_{i1}, x_{i2}, \dots, x_{ij})$ é o vetor de covariáveis associadas aos efeitos fixos;
- $\mathbf{w}_{ij}^\top = (1, w_{i1}, w_{i2}, \dots, w_{ij})$ é o vetor de covariáveis associadas aos efeitos aleatórios;
- y_{ij} é condicionalmente independente com média $\mu_{ij} = E(y_{ij} | \mathbf{v}_j, x_{ij}, w_{ij})$ e a variância $var(y_{ij} | \mathbf{v}_j) = \Phi v(\mu_{ij})$, onde v é função de variância conhecida e Φ é o parâmetro de escala;
- ε_{ij} é o erro aleatório associado a i -ésima unidade do **nível 1**, do j -ésimo grupo do **nível 2**, com as suposições: $\varepsilon_{ij} \sim N(0, \sigma^2)$ e todos os ε_{ij} são independentes.

Na componente sistemática do modelo temos que $\eta_{ij}^{param} = \mathbf{x}_{ij}^\top \boldsymbol{\beta}$ é dito parâmetro linear, com vetor parâmetro $\boldsymbol{\beta}^\top = (\beta_0, \beta_1, \dots, \beta_p)^\top$ e $\eta_{ij}^{aleat} = \mathbf{w}_{ij}^\top \mathbf{v}_j$ contém um grupo específico de efeitos aleatórios, onde $\mathbf{v}_j \sim N(\mathbf{0}, Q)$.

Alternativamente para o j -ésimo grupo, obtemos que:

$$y_j = X_j \boldsymbol{\beta} + W_j \mathbf{v}_j + \boldsymbol{\varepsilon}_j.$$

Considerando todas as observações, temos que:

$$y = X \boldsymbol{\beta} + W \mathbf{v} + \boldsymbol{\varepsilon},$$

onde $X^\top = [X_1^\top, \dots, X_n^\top]$, $W = \text{diag}(W_1, \dots, W_n)$, $\boldsymbol{\varepsilon}^\top = (\varepsilon_1^\top, \dots, \varepsilon_n^\top)$ e $\mathbf{v}^\top = (v_1^\top, \dots, v_n^\top)$ tem distribuição normal com matriz de covariância $Q_v = \text{diag}(Q, \dots, Q)$.

Considere $f(y_{ij} | x_{ij}, \mathbf{v}_j)$ a função de verossimilhança. A matriz de covariância Q_v dos efeitos aleatórios \mathbf{v}_j , depende do vetor de parâmetros desconhecidos. Segundo [Groll e Tutz \(2014\)](#), em problemas de penalização, a função de probabilidade conjunta é especificada por $\boldsymbol{\gamma}^\top = (\Phi, Q^\top)$ e o vetor $\boldsymbol{\delta}^\top = (\boldsymbol{\beta}^\top, \mathbf{v}^\top)$.

O logaritmo da função de verossimilhança, nesse caso é:

$$l(\boldsymbol{\delta}, \boldsymbol{\gamma}) = \sum_{j=1}^n \log \left(\int f(y_j | \boldsymbol{\delta}, \boldsymbol{\gamma}) p(\mathbf{v}_j, \boldsymbol{\gamma}) d\mathbf{v}_j \right), \quad (4.18)$$

onde $p(v_j, \boldsymbol{\gamma})$ é a densidade dos efeitos aleatórios. A equação (4.18) foi resolvida de forma aproximada, por Breslow e Clayton (1993) utilizando o método de Laplace e a sua solução será:

$$l^{approx}(\boldsymbol{\delta}, \boldsymbol{\gamma}) = \sum_{j=1}^n \log(f(y_j | \boldsymbol{\delta}, \boldsymbol{\gamma})) - \frac{1}{2} \mathbf{v}^\top \mathbf{Q} \mathbf{Q}^{-1} \mathbf{v}, \quad (4.19)$$

onde a penalidade $\mathbf{v}^\top \mathbf{Q} \mathbf{Q}^{-1} \mathbf{v}$ é devido a aproximação do método de Laplace. Incluindo a penalidade $\lambda \sum_{j=1}^p |\beta_j|$ na equação (4.19), teremos que:

$$l_{pen}(\boldsymbol{\beta}, \mathbf{v}, \boldsymbol{\gamma}) = l_{pen}(\boldsymbol{\delta}, \boldsymbol{\gamma}) = l^{approx}(\boldsymbol{\delta}, \boldsymbol{\gamma}) - \lambda \sum_{j=1}^p |\beta_j|. \quad (4.20)$$

Sendo assim, nosso problema de estimação de parâmetros se reduz a maximização da equação (4.20), para um $\hat{\boldsymbol{\gamma}}$ conhecido, analisamos esse problema de maneira mais profunda na subseção seguinte.

4.3.3 Algoritmo do Gradiente Ascendente

Para cada $\boldsymbol{\delta}$ e toda direção $\mathbf{h} \in \mathbb{R}^{p+nq}$ podemos definir:

$$l'_{pen}(\boldsymbol{\delta}, \mathbf{h}, \boldsymbol{\gamma}) = \lim_{t \rightarrow 0} \frac{1}{t} (l_{pen}(\boldsymbol{\delta} + t\mathbf{h}, \boldsymbol{\gamma}) - l_{pen}(\boldsymbol{\delta}, \boldsymbol{\gamma})). \quad (4.21)$$

O algoritmo do gradiente ascendente usa aproximação em série de Taylor localmente a partir de uma estimativa de $\hat{\boldsymbol{\delta}}$ na direção do gradiente da aproximação de Taylor de segunda ordem, dada por:

$$l_{pen}(\hat{\boldsymbol{\delta}} + t s^{pen}(\hat{\boldsymbol{\delta}}, \boldsymbol{\gamma}), \boldsymbol{\gamma}) \approx l_{pen}(\hat{\boldsymbol{\delta}}, \boldsymbol{\gamma}) + t l'_{pen}(\hat{\boldsymbol{\delta}}, s^{pen}(\hat{\boldsymbol{\delta}}, \boldsymbol{\gamma}), \boldsymbol{\gamma}) + 0,5 t^2 l''_{pen}(\hat{\boldsymbol{\delta}}, s^{pen}(\hat{\boldsymbol{\delta}}, \boldsymbol{\gamma})), \quad (4.22)$$

onde $t > 0$ e onde $s^{pen}(\cdot, \cdot)$ e $l''_{pen}(\cdot, \cdot)$ são definidos no passo 2 do algoritmo. Uma questão central nesse processo é encontrar o t ótimo para a penalização.

4.3.3.1 Algoritmo glmLasso: caso particular (normal)

Inicialização

Computar os valores iniciais $\hat{\boldsymbol{\beta}}^{(0)}$, $\hat{\mathbf{v}}^{(0)}$, $\hat{\boldsymbol{\gamma}}^{(0)}$ e $\hat{y}^{(0)} = X \hat{\boldsymbol{\beta}}^{(0)} + W \hat{\mathbf{v}}^{(0)}$.

Iteração

Para $l = 1, 2, \dots$ até a convergência utilize os seguintes procedimentos:

Calcule o gradiente da função log verossimilhança para um dado $\hat{\boldsymbol{\gamma}}^{(l-1)}$, para tanto considere

$$s(\boldsymbol{\delta}) = \frac{\partial l^{approx}(\boldsymbol{\delta})}{\partial \boldsymbol{\delta}}$$

e a partir daí teremos que $s_0^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}) = s_0(\hat{\boldsymbol{\delta}}^{(l-1)})$, $s_i^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}) = s_i(\hat{\boldsymbol{\delta}}^{(l-1)})$, $i = p+1, \dots, p+ns$. Além disso, para $i = 1, \dots, p$ temos que $s_i^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}) = s_i(\hat{\boldsymbol{\delta}}^{(l-1)}) - \lambda \text{sign}(\hat{\boldsymbol{\beta}}_i^{(l-1)})$, se $\hat{\boldsymbol{\beta}}_i^{(l-1)} \neq 0$; $s_i^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}) = s_i(\hat{\boldsymbol{\delta}}^{(l-1)}) - \lambda \text{sign}(\hat{\boldsymbol{\delta}}_i^{(l-1)})$, se $\hat{\boldsymbol{\beta}}_i^{(l-1)} = 0$ e $|s_i(\hat{\boldsymbol{\delta}}_i^{(l-1)})| > \lambda$; $s_i^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}) = 0$, caso contrário.

$$\text{Onde } \text{sign}(x) = \begin{cases} 1, & \text{se } x > 0; \\ 0, & \text{se } x = 0; \\ -1, & \text{se } x < 0. \end{cases}$$

Calculando a segunda derivada direcional

Seja $A := [X, W]$ e $K = \text{diag}(0, \dots, 0, Q^{-1}, \dots, Q^{-1})$ uma matriz bloco-diagonal com uma diagonal de $p+1$ zeros que correspondentes aos efeitos fixos e, em seguida, a matriz Q^{-1} n vezes. Então a matriz de informação de Fisher é dada em sua forma fechada por $F^{pen}(\boldsymbol{\delta}) = A^\top R(\boldsymbol{\delta})A + K$ com $R(\boldsymbol{\delta}) = D(\boldsymbol{\delta})\Sigma^{-1}(\boldsymbol{\delta})D(\boldsymbol{\delta})^\top$ e $D(\boldsymbol{\delta}) = \mathbb{I}$, $\Sigma(\boldsymbol{\delta}) = \text{cov}(y|\boldsymbol{\delta})$, onde \mathbb{I} denota a matriz identidade.

A segunda derivada direcional é dada para cada $\boldsymbol{\delta}$ e cada vetor direcional $h \in \mathbb{R}^{p+1+ns}$ por $l''_{pen}(\boldsymbol{\delta}; h) = -h^\top F^{pen}(\boldsymbol{\delta})h$.

Aproximação de Taylor ótima

Maximizando a aproximação de Taylor (4.22) com relação a t , usando $l'_{pen}(\boldsymbol{\delta}; s^{pen}(\boldsymbol{\delta}, \boldsymbol{\gamma}), \boldsymbol{\gamma}) = \|s^{pen}(\boldsymbol{\delta})\|_2$,

$$t^{l-1}_{otimo} = -\frac{\|s^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)})\|_2}{l''_{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}, s^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}))}$$

e

$$t^{(l-1)}_{edge} = \min_i \left\{ -\frac{\hat{\boldsymbol{\delta}}_i^{(l-1)}}{s_i^{pen}(\hat{\boldsymbol{\delta}}_i^{(l-1)})} : \text{sign}(\boldsymbol{\delta}_i^{(l-1)}) = -\text{sign}[s_i^{pen}(\boldsymbol{\delta}_i^{(l-1)})] \neq 0. \right.$$

com norma $\|\cdot\|_2$ denotando a norma L_2 (norma euclidiana).

Atualização

$$\hat{\boldsymbol{\delta}}^{(l)} = \begin{cases} \hat{\boldsymbol{\delta}}^{(l-1)} + t_{edge}^{(l-1)} s^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}), & \text{se } t_{otimo}^{(l-1)} \geq t_{edge}^{(l-1)}; \\ \hat{\boldsymbol{\delta}}_{FS}^{(l-1)}, & \text{se } t_{otimo}^{(l-1)} < t_{edge}^{(l-1)} \text{ e } \text{sign}(\hat{\boldsymbol{\delta}}_{FS}^{(l)}) = \text{sign}(\hat{\boldsymbol{\delta}}^{(l-1)}); \\ \hat{\boldsymbol{\delta}}^{(l-1)} + t_{edge}^{(l-1)} s^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}), & \text{caso contrário.} \end{cases}$$

onde $\hat{\boldsymbol{\delta}}_{FS}^{(l)}$ denota o escore de Fisher estimado.

Estimativas \hat{Q}^l são obtidos como estimativas aproximadas do tipo EM ou por métodos alternativos que produzam a atualização Q^l . Se for necessário, todo o vetor de $\hat{\boldsymbol{\gamma}}^{(l)}$ é completado por uma estimativa do parâmetro de dispersão.

Reestimação

Numa etapa final, utilizando apenas as variáveis correspondentes aos parâmetros cujos $\hat{\boldsymbol{\beta}}$ que são não nulos, uma nova estimativa dos parâmetros é feita, para isso utilizamos um escore de Fisher, conforme descrito no capítulo 3.

Para uma iteração arbitrária definimos $J = \{j : \beta_j \neq 0, j = 0, 1, \dots, p\}$ o conjunto dos índices das covariáveis ativas, correspondendo para as $m = \text{card}(J) \leq p + 1$ coeficientes não-nulos. Considerando, $\tilde{\boldsymbol{\delta}}^\top = (\beta_{J_1}, \dots, \beta_{J_m}, \mathbf{v}^T)$, e seja

$$\tilde{s}^{pen}(\boldsymbol{\delta}) = \left\{ s_{J_1}^{pen}(\boldsymbol{\delta}), \dots, s_{J_m}^{pen}(\boldsymbol{\delta}), s_{p+1}^{pen}(\boldsymbol{\delta}), \dots, s_{p+ns}^{pen}(\boldsymbol{\delta}) \right\}^\top$$

o gradiente em domínio restrito e $\tilde{F}^{pen}(\boldsymbol{\delta}) = A_j^\top R(\boldsymbol{\delta}) A_j + K_j$, com $A_j := [X_j, W]$ enquanto que X_j contém apenas aquelas colunas de X correspondente a J ,

$$K_j = \text{diag}(0, \dots, 0, Q^{-1}, \dots, Q^{-1})$$

é uma matriz de penalização em bloco com uma diagonal de m zeros correspondente aos efeitos fixos não nulos e n vezes a matriz Q^{-1} .

Em algum passo do escore de Fisher as estimativas assumem a forma

$$\hat{\boldsymbol{\delta}}^{(l)} = \hat{\boldsymbol{\delta}}^{(l-1)} + (\tilde{F}^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}))^{-1} \tilde{s}^{pen}(\hat{\boldsymbol{\delta}}^{(l-1)}). \quad (4.23)$$

O estimador obtido na equação (4.23), pode ser expresso por um vetor da forma $\hat{\boldsymbol{\delta}}_{FS}^{(l)}$ aumentando $\hat{\boldsymbol{\delta}}^{(l)}$ com zeros para todas as covariáveis não-ativas, para que a aproximação de Taylor obtida no passo seguinte se mantenha dentro das características colocadas anteriormente, isto é, $\hat{\boldsymbol{\delta}}_{FS}^{(l)}$ é aceito somente quando $\text{sign}(\hat{\boldsymbol{\delta}}_{FS}^{(l)}) = \text{sign}(\hat{\boldsymbol{\delta}}_{FS}^{(l-1)})$.

Observamos que as sugestões de escolha de modelos ou variáveis são baseadas em critérios objetivos. Porém, muitas vezes a experiência do pesquisador da área pode se sobrepor aos métodos aqui utilizados. No próximo capítulo falaremos a respeito de aplicações a dados educacionais, tanto no caso clássico como no caso bayesiano.

Para o caso clássico usamos os critérios de seleção de modelos AIC, BIC e AICc e o critério de seleção de variáveis Lasso. Para o caso bayesiano usamos os critérios WAIC, *looic*, DIC, EAIC e EBIC.

ESTUDO DE SIMULAÇÃO

O objetivo deste capítulo é desenvolver um breve estudo de simulação para analisar a eficiência dos critérios de seleção clássicos AIC, AICc e BIC. Dessa forma, consideramos alguns cenários para gerar observações vindas de um modelo normal e assim comparar os critérios de seleção citados anteriormente. Primeiramente, geramos três covariáveis (x_1, x_2 e x_3) dicotômicas, que foram fixadas para todos os cenários, e daí fixamos valores teóricos para os efeitos fixos de nosso modelo $\gamma_0 = 100$, $\gamma_1 = 3$, $\gamma_2 = 3$ e $\gamma_3 = -15$. Feito isso, geramos uma variável de grupo, em que a quantidade de grupos varia de acordo com o tamanho da amostra, com exceção da amostra de tamanho $n = 50$, onde são considerado grupos de tamanho 10.

Para a variância dos erros e da matriz de variâncias e covariâncias, fixamos como valores teóricos $\tau_{00} = 11$, $\tau_{01} = 5$, $\tau_{02} = 1$, $\tau_{03} = 1$, $\tau_{11} = 5$, $\tau_{12} = 1$, $\tau_{13} = 1$, $\tau_{22} = 6$, $\tau_{23} = 1$, $\tau_{33} = 4$ e $\sigma^2 = 33$. Foram geradas 1000 réplicas, de acordo com o seguinte procedimento:

- geramos os efeitos aleatórios b_0, b_1, b_2 e b_3 , respectivamente, para o intercepto;
- obtemos os coeficientes, utilizamos as equações $\beta_{0j} = \gamma_0 + b_{0j}$, $\beta_{1j} = \gamma_1 + b_{1j}$, $\beta_2 = \gamma_2 + b_2$ e $\beta_3 = \gamma_3 + b_3$;
- geramos o vetor de erros, com a distribuição normal com média zero e variância 33;
- geramos a variável resposta y , utilizando uma equação da forma $y_{ij} = \beta_{0j} + \beta_{1j}x_{1ij} + \beta_2x_{2ij} + \beta_3x_{3ij} + \varepsilon_{ij}$, dados $(x_1, x_2$ e $x_3)$, que foram fixados inicialmente.

Note que a equação utilizada no último item varia de acordo com o cenário em questão, devido à variação no número de covariáveis. Os tamanhos amostrais utilizados são 50, 100, 1000, 5000, 10000. Os cenários utilizados foram os seguintes:

- **Cenário 1:** utilizamos apenas a covariável x_{1ij} ;

- **Cenário 2:** utilizamos apenas a covariável x_{2ij} ;
- **Cenário 3:** utilizamos apenas a covariável x_{3ij} ;
- **Cenário 4:** utilizamos as covariáveis x_{1ij} e x_{2ij} ;
- **Cenário 5:** utilizamos as covariáveis x_{1ij} e x_{3ij} ;
- **Cenário 6:** utilizamos as covariáveis x_{2ij} e x_{3ij} ;
- **Cenário 7:** utilizamos as covariáveis x_{1ij} , x_{2ij} e x_{3ij} .

Após a geração da variável resposta, estimam-se sete modelos de acordo com a configuração descrita na Tabela 3 e calculam-se o valor das estatísticas de comparação de modelo, daí verifica-se quantas vezes dentre as réplicas foi escolhido o modelo teórico. Observe que o modelo teórico é aquele sob o qual os dados foram gerados, por exemplo, no cenário 1, o modelo 1 é o modelo teórico.

O percentual de vezes onde as estatísticas de comparação de modelos nos levaram a optar pelo modelo teórico encontram-se disponíveis na Tabela 4.

Tabela 3 – Configuração dos modelos estimados em cada réplica.

	x_{1ij}	x_{2ij}	x_{3ij}
Modelo 1	x		
Modelo 2		x	
Modelo 3			x
Modelo 4	x	x	
Modelo 5	x		x
Modelo 6		x	x
Modelo 7	x	x	x

Ao observar a Tabela 4, podemos notar que quando aumentamos o tamanho amostral há um aumento na eficiência dos três critérios de informação, ao compararmos os três critérios pode-se afirmar que o critério BIC é o mais eficiente dos três para todos tamanhos amostrais, quando temos um número menor de variáveis explicativas. Outro ponto importante a se destacar é que para grandes tamanhos amostrais o AIC e o AICc assemelham-se no percentual de escolha do modelo correto. Como o contexto prático em que estamos trabalhando é de uma amostra grande e uma grande quantidade de covariáveis, é razoável esperar que o BIC seja o critério mais eficiente. No próximo capítulo trabalharemos com as estatísticas de seleção de modelos utilizadas no capítulo 4 numa situação de dados reais.

	$n = 50$			$n = 100$			$n = 1000$			$n = 5000$			$n = 10000$		
	AIC	AICc	BIC	AIC	AICc	BIC	AIC	AICc	BIC	AIC	AICc	BIC	AIC	AICc	BIC
Cenário 1	59,3%	67,8%	68,9%	73,9%	79,2%	83,4%	93,6%	94,1%	100%	97,1%	97,1%	100%	98,7%	98,7%	100%
Cenário 2	57,2%	66,0%	66,9%	71,4%	76,4%	80,2%	91,0%	91,4%	100%	93,5%	93,5%	100%	98,0%	98,0%	100%
Cenário 3	78,0%	94,3%	98,8%	85,9%	94,6%	99,8%	91,5%	91,6%	100%	93,5%	93,5%	100%	97,5%	97,5%	100%
Cenário 4	24,4%	5,9%	0,8%	46,8%	34,9%	3,2%	97,7%	97,7%	98,8%	99,2%	99,2%	100%	99,3%	99,3%	100%
Cenário 5	46,9%	23,2%	10,0%	67,5%	58,5%	18,1%	98,4%	98,6%	99,5%	99,2%	99,2%	100%	99,3%	99,3%	100%
Cenário 6	44,3%	23,1%	9,5%	59,6%	54,1%	15,9%	98,3%	98,4%	99,2%	98,9%	98,9%	100%	99,2%	99,2%	100%
Cenário 7	14,7%	0,5%	0,4%	32,2%	14,6%	0,8%	100%	100%	93,5%	100%	100%	100%	100%	100%	100%

Tabela 4 – Percentual de réplicas onde os critérios de seleção optaram pelo modelo verdadeiro.

APLICAÇÕES A DADOS EDUCACIONAIS

Neste capítulo, analisamos uma aplicação de modelos multiníveis com resposta normal para dados educacionais. Temos aqui o objetivo principal de analisar e identificar os fatores que influenciam no desempenho educacional em matemática, dos alunos do 9º ano do ensino fundamental do Estado de São Paulo. Para isso foi usado o banco de dados da Prova Brasil de 2011, disponível em [Microdados \(2011\)](#), que em sua forma amostral consiste em 12.280 alunos, em 614 escolas do Estado de São Paulo. Foram analisados além das notas dos alunos, seus questionários socioeconômicos e também relativos a entes escolares. Por se tratar de dados com uma estrutura hierárquica propomos a utilização de uma abordagem multinível. Começamos com um modelo de dois níveis (aluno e turma) e depois inserimos o nível escola.

6.1 Modelos Multiníveis

Quando temos dados estruturados de forma que haja uma correlação entre seus grupos, faz-se necessário a construção de modelos multiníveis com dois ou mais níveis. Segundo [Laros, Marciano e Andrade \(2010\)](#), faz parte do processo básico para a construção do modelo hierárquico começar com um modelo simplificado, ou seja, sem covariáveis, a partir daí faz-se a inclusão de variáveis explicativas até chegar a um modelo mais completo. O modelo nulo serve como base para analisar um modelo mais completo e a variância da variável resposta do modelo é representada pela soma das variações entre os níveis. O modelo nulo (sem variáveis explicativas) é útil para estimar a variância explicada versus variância não explicada em comparação a outros modelos condicionais estimados *a posteriori*, o que facilita o cálculo do coeficiente de correlação intra-classe (ICC). Note que quando o ICC é próximo de zero, significa que as unidades nível macro são homogêneas entre si, caracterizando pouca ou nenhuma influência no desempenho do aluno, se o ICC for zero, não há necessidade do uso de regressão multinível.

Como foi dito no Capítulo 2, a amostra dos dados é constituída de 12.280 alunos, que constituem 2.187 turmas, 614 escolas, 236 municípios e 15 regiões. Vimos também a análise

descritiva da variável PROFICIENCIA_MT_SAEB. Podemos afirmar que não há evidências que nos leve a rejeitar suposição de normalidade dos dados. Nosso objetivo aqui é apontar os fatores que influenciam no desempenho do aluno na disciplina de matemática, utilizando a prova de matemática da PROVA BRASIL. Por se tratar de uma estrutura hierárquica propomos uma abordagem multinível.

Para ajustar os modelos sob abordagem clássica, usaremos o software R, com os pacotes lme4, nlme, entre outros, mais detalhes em [Gelman e Hill \(2006\)](#) e [Bates et al. \(2015\)](#). Os ajustes são obtidos via REML (estimativa de máxima verossimilhança restrita), conceito esse que foi detalhado no Capítulo 3.

Para escolha das covariáveis utilizadas nessa seção, conforme dito na Seção 2.2, consideramos critérios de significância teórica, ou seja, ao analisar trabalhos publicados na área recentemente, foi feito um levantamento dos preditores mais utilizados, dentre esses trabalhos podemos destacar [Goldstein \(1986\)](#), [Castro \(2015\)](#), [Osio \(2013\)](#), [Natis \(2013\)](#), [Tibshirani \(1996\)](#) e [Weakliem \(1999\)](#).

6.1.1 Modelo 1

O **Modelo 1** é um modelo sem variáveis explicativas com os níveis aluno e turma. Esse modelo no **nível 1** será dado por:

$$y_{ij} = \beta_{0j} + \varepsilon_{ij}, \quad (6.1)$$

onde $\beta_{0j} = \gamma_{00} + v_{0j}$, com $i = 1, \dots, n_j$ e $j = 1, \dots, 2.187$. Note que, y_{ij} é a nota de matemática do i -ésimo aluno da j -ésima turma, $E(\varepsilon_{ij}) = 0$, $cov(\varepsilon) = \sigma^2$, γ_{00} é o efeito fixo e v_{0j} é o efeito aleatório de turma, assumimos aqui que $v_{0j} \sim N(\mathbf{0}, \tau_{00})$. Assumimos ainda, independência entre o efeito fixo e os efeitos aleatórios, bem como entre os efeitos aleatórios.

6.1.2 Modelo 2

O **Modelo 2** é um modelo sem variáveis explicativas com os níveis aluno, turma e escola. As equações desse modelo no **nível 1**, **nível 2** e **nível 3**, respectivamente, são dadas por:

$$\begin{aligned} y_{ijk} &= \beta_{0jk} + \varepsilon_{ijk}, \\ \beta_{0jk} &= \gamma_{00k} + v_{0jk}, \\ \gamma_{00k} &= \alpha_{000} + u_{00k}. \end{aligned} \quad (6.2)$$

onde $i = 1, \dots, n_j$, $j = 1, \dots, n_k$ e $k = 1, \dots, 614$. As suposições são:

- $\varepsilon_{ijk} \sim N(0, \sigma_{\varepsilon}^2)$;
- $v_{0jk} \sim N(0, v_0)$;

- $u_{00k} \sim N(0, \tau_{00})$.

Note que, α_{000} é o efeito fixo do modelo, u_{00k} é o efeito aleatório da escola e v_{0jk} é o efeito aleatório de turma. Observe também que u_{00k} e v_{0jk} são independentes para diferentes valores de k e também são independentes entre si para diferentes valores de j e k .

Na Tabela 5, vemos as estimativas do modelo 1 (à esquerda) e do modelo 2 (à direita).

Tabela 5 – Modelo 1: apenas com os níveis aluno e turma ; Modelo 2: com os níveis aluno, turma e escola.

	Modelo 1		Modelo 2	
Variáveis Explicativas				
Efeito fixo	Estimativa	valor p	Estimativa	valor p
Intercepto	246,79	< 0,001	247,08	< 0,001
Efeito aleatório				
Nível 1 (Aluno)				
Variância (Resíduo)	1886,60	-	1884,20	-
Nível 2 (Turma)				
Variância (intercepto)	326,60	-	121,80	-
Nível 3 (Escola)				
Variância (intercepto)	-	-	204,60	-
ICC	0,14	-	-	-

No modelo 1 (ver Tabela 5 à esquerda), temos que o ICC é de 0,147, ou seja, aproximadamente 14,75% da variabilidade do desempenho escolar em matemática dos alunos estão associados à variabilidade das turmas, o intercepto 246,79 é a média das proficiências em matemática de cada grupo. Observe que, o coeficiente de correlação intra-classe (ICC), é dado por:

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma^2}, \quad (6.3)$$

onde σ_u^2 é a variância do efeito aleatório associado as turmas e σ^2 é a variância dos resíduos.

Ao inserirmos o nível escola (ver Tabela 5 à direita) vemos que aproximadamente 9,25% da variabilidade do desempenho escolar em matemática dos alunos está associado à variabilidade das escolas e aproximadamente 5,50% está associado à turma. É importante ressaltar que com a inclusão do nível escola o intercepto do modelo passa ser 247,08. Destacamos também que aproximadamente 62,68% da variabilidade acumulada nas turmas é explicada pela variabilidade entre as escolas, indicando uma importância relativamente alta da escola no processo de ensino e aprendizagem de suas turmas.

6.1.3 Modelo 3

O **Modelo 3** tem a mesma estrutura do **Modelo 2**, acrescido das variáveis explicativas gênero, raça auto atribuída, escolaridade da mãe e escolaridade do pai ao nível de aluno. As

equações desse modelo no **nível 1**, **nível 2** e **nível 3**, respectivamente, são:

$$\begin{aligned} y_{ijk} &= \beta_{0jk} + \beta_{1jk}Q001_{ijk} + \sum_{q=2}^6 \beta_{qjk}Q002_{ijk}^{(q)} + \beta_{7jk}Q019_{ijk} + \beta_{8jk}Q023_{ijk} + \varepsilon_{ijk}, \\ \beta_{qjk} &= \gamma_{q0k} + v_{qjk}, q = 0, \dots, 8, \\ \gamma_{q0k} &= \alpha_{q00} + u_{q0k}. \end{aligned} \quad (6.4)$$

onde $i = 1, \dots, n_j$, $j = 1, \dots, n_k$ e $k = 1, \dots, 614$. Temos ainda que:

$$Q002_{ijk}^{(2)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça amarela;} \\ 0, & \text{caso contrário.} \end{cases}$$

$$Q002_{ijk}^{(3)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça branca;} \\ 0, & \text{caso contrário.} \end{cases}$$

$$Q002_{ijk}^{(4)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça indígena;} \\ 0, & \text{caso contrário.} \end{cases}$$

$$Q002_{ijk}^{(5)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça parda;} \\ 0, & \text{caso contrário.} \end{cases}$$

$$Q002_{ijk}^{(6)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça preta;} \\ 0, & \text{caso contrário.} \end{cases}$$

As suposições são:

- $\varepsilon_{ijk} \sim N(0, \sigma_{\varepsilon}^2)$;
- $v_{qjk} \sim N(0, v_{qjk})$;
- $u_{q0k} \sim N(0, \tau_{q0k})$.

Note que, α_{q00} é o efeito fixo do modelo, u_{q0k} é o efeito aleatório da escola e v_{qjk} é o efeito aleatório de turma. Observe também que u_{q0k} e v_{qjk} são independentes para diferentes valores de k e também são independentes entre si para diferentes valores de j e k .

6.1.4 Modelo 4

O **Modelo 4** tem a mesma estrutura do **Modelo 3**, acrescido das variáveis explicativas dependência administrativa, localização, turno e proficiência em língua portuguesa ao nível de aluno e dependência administrativa e localização ao nível de escola. As equações desse modelo

no **nível 1**, **nível 2** e **nível 3**, respectivamente, por:

$$\begin{aligned}
 y_{ijk} &= \beta_{0jk} + \beta_{1jk}Q001_{ijk} + \sum_{q=2}^6 \beta_{qjk}Q002_{ijk}^{(q)} + \beta_{7jk}Q019_{ijk} + \beta_{8jk}Q023_{ijk} + \\
 &\beta_{9jk}DEPENDENCIA_ADM_{ijk} + \beta_{10jk}LOCALIZACAO_{ijk} + \beta_{11jk}TURNO_{ijk} + \\
 &\beta_{12jk}PROFICIENCIA_LP_SAEB_{ijk} + \varepsilon_{ijk}, \\
 \beta_{qjk} &= \gamma_{q0k} + v_{qjk}, q = 0, \dots, 12, \\
 \gamma_{q0k} &= \alpha_{q000} + \alpha_{q001}DEPENDENCIA_ADM_k + \alpha_{q002}LOCALIZACAO_k + u_{q0k}.
 \end{aligned} \tag{6.5}$$

onde $i = 1, \dots, n_j$, $j = 1, \dots, n_k$ e $k = 1, \dots, 614$. Temos ainda que:

$$Q002_{ijk}^{(2)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça amarela;} \\ 0, & \text{caso contrário.} \end{cases}$$

$$Q002_{ijk}^{(3)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça branca;} \\ 0, & \text{caso contrário.} \end{cases}$$

$$Q002_{ijk}^{(4)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça indígena;} \\ 0, & \text{caso contrário.} \end{cases}$$

$$Q002_{ijk}^{(5)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça parda;} \\ 0, & \text{caso contrário.} \end{cases}$$

$$Q002_{ijk}^{(6)} = \begin{cases} 1, & \text{se o estudante autodeclarou ser da raça preta;} \\ 0, & \text{caso contrário.} \end{cases}$$

As suposições são:

- $\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$;
- $v_{qjk} \sim N(0, v_{qjk})$;
- $u_{q0k} \sim N(0, \tau_{q0k})$.

Note que, α_{q00} é o efeito fixo do modelo, u_{q0k} é o efeito aleatório da escola e v_{qjk} é o efeito aleatório de turma. Observe também que u_{q0k} e v_{qjk} são independentes para diferentes valores de k e também são independentes entre si para diferentes valores de j e k . Na Tabela 6 à esquerda, temos o ajuste do modelo 3 e à direita o ajuste do modelo 4.

No modelo 3 (Tabela 6 à esquerda) inserimos as variáveis explicativas gênero, raça autodeclarada, escolaridade do pai e escolaridade da mãe, consideramos ainda três níveis: alunos, turma e escolas, temos 8,01% da variabilidade do desempenho escolar em matemática dos alunos é explicada pelas escolas, ou seja, ao inserirmos variáveis explicativas diminuí a influência das escolas na nota dos alunos e 0,36% da variabilidade do desempenho escolar em matemática é explicado pelas turmas. Temos ainda que 31,40% da variabilidade da turma é explicada pela escola. Pode-se concluir então que quando inserimos preditores no nível aluno, há uma

Tabela 6 – Modelo 3: inclusão de variáveis explicativas ao nível de aluno; Modelo 4: inclusão de variáveis explicativas ao nível de escola.

Variáveis Explicativas	Modelo 3		Modelo 4	
	Estimativa	valor p	Estimativa	valor p
Efeito fixo				
Intercepto	221,03	< 0,001	95,58	< 0,001
Sexo: Feminino	-2,41	5,38.10 ⁻³	-14,01	< 0,001
Sexo: Masculino	-	-	-	-
Raça: Amarelo	23,02	< 0,001	9,47	0,00
Raça: Branco	25,90	< 0,001	8,18	0,00
Raça: Indígena	11,35	0,01	1,22	0,75
Raça: Pardo	19,43	< 0,001	6,44	0,10
Raça: Preto	10,50	1,22.10 ⁻⁴	1,68	0,09
Raça: Não Sei	-	-	-	-
Escolaridade(Mãe): Ingressou no Ensino Superior ou mais	10,32	< 0,001	3,47	0,00
Escolaridade(Mãe): Terminou o Ensino Médio ou Menos	-	-	-	-
Escolaridade(Pai): Ingressou no Ensino Superior ou mais	10,00	< 0,001	3,38	< 0,001
Escolaridade(Pai): Terminou o Ensino Médio ou Menos	-	-	-	-
DEPENDENCIA_ADM: Estadual	-	-	-4,24	0,00
DEPENDENCIA_ADM: Municipal	-	-	-	-
LOCALIZACAO: Urbana	-	-	-5,66	0,06
LOCALIZACAO: Rural	-	-	-	-
TURNO: Matutino	-	-	1,10	0,86
TURNO: Vespertino	-	-	-	-
PROFICIENCIA_LP_SAEB	-	-	0,64	< 0,001
Efeito aleatório				
Nível 1 (Aluno)				
Variância (Resíduo)	1.825,83	-	1.131,00	-
Nível 2 (Turma)				
Variância (intercepto)	75,89	-	3,68	-
Nível 3 (Escola)				
Variância (intercepto)	165,77	-	< 0,001	-
Variância (DEPENDENCIA_ADM)	-	-	41,99	-
Variância (LOCALIZACAO)	-	-	83,59	-

diminuição da influência da escola na variabilidade da proficiência em matemática, no entanto, há uma diminuição da variabilidade total que era 2.210,60 e agora passa a ser 2.067,74.

Aparentemente, o modelo 3 sugere que a raça autodeclarada interfere na proficiência em matemática, sendo a categoria “preto” a que leva a menores notas. Outro fator importante a ser destacado é que a escolaridade do pai ou da mãe tem uma influência menor na variável resposta do que raça auto declarada, como pode ser visto nas estimativas da Tabela 6 à esquerda.

No modelo 4 (Tabela 6 á direita), incluímos as variáveis explicativas proficiência em língua portuguesa, dependência administrativa, localização e turno no nível de aluno e incluímos também os preditores dependência administrativa e localização ao nível de escola, obviamente que com a inclusão das variáveis explicativas ao nível de escola, há um aumento da variabilidade da proficiência em matemática explicada pelas escolas, que passa de 8,01% para aproximadamente 9,96%, por outro lado 0,29% da variabilidade das notas em matemática é explicada pelas turmas, temos ainda que 97,15% da variabilidade da turma é explicada pela escola. Assim como no modelo anterior, há indícios de que a raça autodeclarada interfere na proficiência em matemática, note que os alunos que se autodeclararam pretos, pardos e indígenas parecem ter notas menores do que os alunos que autodeclararam amarelo ou branco, mas os valores p para esses casos são grandes, assim esta hipótese é prontamente descartada.

Depois que houve uma inserção de quatro variáveis explicativas no nível aluno, houve uma diminuição da variabilidade das notas explicadas pelas escolas, e no último modelo inserimos mais quatro variáveis no nível aluno e duas variáveis explicativas ao nível de escola, e por fim temos que aproximadamente 89,74% da variabilidade da variável resposta é explicada pelos alunos, aproximadamente 0,29% é explicada pelas turmas e aproximadamente 9,96% é explicada pelas escolas, em contraste com outros estudos onde a escola tem uma importância maior na variabilidade do desempenho.

As estimativas dos modelos 3 e 4 nos dão evidências de que a escolaridade da mãe e a escolaridade do pai tem uma importância considerável no desempenho do estudante. Observe que tanto no caso do pai quanto no caso da mãe, o fato deles terem uma escolaridade mais elevada contribui positivamente em seu desempenho em matemática. Na Tabela 7, temos um teste de comparações múltiplas entre os modelos 3 e 4, como valor p é pequeno, há evidência, considerando um nível de significância de 5%, que o modelo 4 é melhor.

Tabela 7 – Teste de Comparação entre os modelos 3 e 4

	Df	AIC	BIC	log Veros.	desvio	χ	g.l.	valor p
Modelo 3	12	108951,16	109038,20	-54463,58	108927,16			
Modelo 4	21	103428,26	103580,59	-51693,13	103386,26	5540,90	9	0

Na Tabela 7, temos um teste de razão de verossimilhança, onde a estatística teste será dada por:

$$\chi_{(m_4-m_3)} = -2\ln(M_3) + 2\ln(M_4), \quad (6.6)$$

em que M_3 é a função de verossimilhança do modelo 3, M_4 é a função de verossimilhança do modelo 4, m_4 é o número de parâmetros do modelo 4 e m_3 é o número de parâmetros do modelo 3.

Por fim, vamos comparar todos os modelos ajustados nesta seção usando os critérios AIC, BIC e AICc, temos que nos três casos o modelo 4 é o modelo mais adequado aos dados, mais detalhes a respeito podem ser vistos na Tabela 8.

Tabela 8 – Comparação dos modelos ajustados nesta seção utilizando os critérios AIC, BIC, AICc.

	AIC	BIC	AICc
Modelo1	109649,7	109671,4	109649,7
Modelo2	109474,8	109503,8	109474,8
Modelo3	108928,1	109015,1	108928,1
Modelo4	103411,0	103563,3	103411,1

Assim, devido a porção considerável da variabilidade da variável resposta que é explicada pela variabilidade entre as turmas e entre as escolas, não há evidências que nos façam descartar a utilização de modelos multiníveis.

Na seção seguinte apresentamos novos ajustes, outra configuração de variáveis é utilizada, primeiramente trabalhamos com o enfoque clássico, posteriormente o enfoque bayesiano.

6.2 Modelagem clássica

Considerando novamente como variável resposta a proficiência em matemática do SAEB (PROFICIENCIA_MT_SAEB), vamos ajustar modelos usando os preditores listados na Tabela 9.

Tabela 9 – Descrição das covariáveis a serem utilizadas nos ajustes no caso clássico e no bayesiano.

Covariável	Código na base de Dados	Código no ajuste
Dependência administrativa	DEPENDENCIA_ADM	x_1
Localização	LOCALIZACAO	x_2
Turno	TURNO	x_3
Sexo do aluno	Q001	x_4
Idade	Q004	x_5
Escolaridade da mãe	Q019	x_6
Escolaridade do pai	Q023	x_7
Reprovas	Q049	x_8
Proficiência em língua portuguesa do SAEB	PROFICIENCIA_LP_SAEB	X_9

Com exceção da variável explicativa proficiência em língua portuguesa do SAEB, todos os outros preditores são variáveis dicotômicas, sua definição pode ser encontrada na Tabela 32, no apêndice A.

Nesta seção ajustamos modelos de regressão multiníveis com dois níveis: aluno e escola; aluno e turma e também aluno e região. Primeiramente buscamos qual a metodologia de ajuste deve ser escolhida (efeitos fixos, intercepto aleatórios e intercepto e inclinações aleatórias),

utilizando os critérios de informação (AIC, AICc e BIC) e partir daí verificamos quais das variáveis explicativas estariam no modelo final.

6.2.1 Modelos com efeitos fixos

A equação de um modelo de regressão linear com efeitos fixos, considerando todos os preditores e a variável resposta acima, será da seguinte forma:

$$y_i = \beta_0 + \sum_{m=1}^9 \beta_m x_{mi} + \varepsilon_i, \quad (6.7)$$

com $i = 1, 2, \dots, 12.280$, onde:

- y_i representa a proficiência em matemática do i -ésimo indivíduo;
- $\varepsilon_i \sim N(0, \sigma^2)$ e todos os ε_i são independentes.

Na Tabela 10, temos as estimativas do modelo (6.7).

Tabela 10 – Estimativa dos parâmetros do modelo (6.7), aqui temos efeitos fixos apenas.

Parâmetro	Estimativa	Erro Padrão	valor t	valor p
$\hat{\beta}_0$ (Intercepto)	103,90	3,14	33,00	< 0,001
$\hat{\beta}_1$ (Dep. Admin.)	-4,19	0,77	-5,40	< 0,001
$\hat{\beta}_2$ (Localização)	-5,74	2,17	-2,63	0,008
$\hat{\beta}_3$ (Turno)	0,94	0,77	1,21	0,224
$\hat{\beta}_4$ (Q001)	-14,77	0,68	-21,46	< 0,001
$\hat{\beta}_5$ (Q004)	2,04	1,33	1,53	0,125
$\hat{\beta}_6$ (Q019)	2,93	0,81	3,61	< 0,001
$\hat{\beta}_7$ (Q023)	3,26	0,82	3,94	< 0,001
$\hat{\beta}_8$ (Q049)	-9,84	1,01	-9,72	< 0,001
$\hat{\beta}_9$ (Nota Port.)	0,64	0,00	84,67	< 0,001

Ao analisar o teste t de significância individual, ao nível de significância de 5%, há evidências, a partir da consulta do valor p na Tabela 10, de que as variáveis explicativas TURNO e distorção idade-série (Q004) são não significativas.

Ao analisarmos as estimativas dos parâmetros do modelo (6.7), destaca-se que o aluno ter sofrido reprovos em sua vida escolar influencia negativamente seu desempenho em matemática na Prova Brasil, fixado os outros preditores, o mesmo pode-se dizer o fato do aluno declarar ser do sexo feminino. Por outro lado, as estimativas contidas na Tabela 10, mostram que os alunos cujos pais tem escolaridade mais elevada tendem a ter um desempenho mais elevado em matemática, fixado outras covariáveis.

Contudo ao realizarmos o procedimento de stepwise, eliminamos apenas a variável TURNO, indicando a importância do aluno não ter defasagem idade série para o modelo. Mais detalhes pode ser visto na Tabela 11.

Tabela 11 – Procedimento de stepwise utilizado no Modelo 4

Step	g.l.	Desvio	g.l. Resid.	Resid. Dev	AIC	
1			10424	12177770,07	73708,01	
2	- TURNO	1	1723,95	10425	12179494,02	73707,49

6.2.2 Modelos com intercepto aleatório

Nesta subseção ajustamos mais três modelos, será colocado no modelo (6.7), efeitos aleatórios relacionados à turma, escola e região. O modelo linear com intercepto aleatório para turma será dado no nível 1, por:

$$y_{ij} = \beta_{0j} + \sum_{m=1}^9 \beta_m x_{mij} + \varepsilon_{ij}, \quad (6.8)$$

com $i = 1, 2, \dots, n_j$ e $j = 1, 2, \dots, 2187$, onde:

- y_{ij} representa a proficiência em matemática do i -ésimo aluno para a j -ésima turma;
- $\varepsilon_{ij} \sim N(0, \sigma^2)$ e todos os ε_{ij} são independentes.

Temos também que $\beta_{0j} = \gamma_{00} + v_{0j}$, onde γ_{00} é a proficiência em matemática esperada de cada turma, v_{0j} é o efeito aleatório da j -ésima turma no intercepto do modelo (6.8) e $v_{0j} \sim N(0, \tau_{00})$.

O modelo linear com intercepto aleatório para escola será dado no nível 1, por:

$$y_{ij} = \beta_{0j} + \sum_{m=1}^9 \beta_m x_{mij} + \varepsilon_{ij}, \quad (6.9)$$

com $i = 1, 2, \dots, n_j$ e $j = 1, 2, \dots, 614$, onde:

- y_{ij} representa a proficiência em matemática do i -ésimo aluno da j -ésima escola;
- $\varepsilon_{ij} \sim N(0, \sigma^2)$ e todos os ε_{ij} são independentes.

Temos também que $\beta_{0j} = \gamma_{00} + v_{0j}$, onde γ_{00} é a proficiência em matemática esperada de cada escola, v_{0j} é o efeito aleatório da j -ésima escola no intercepto do modelo (6.9) e $v_{0j} \sim N(0, \tau_{00})$.

O modelo linear com intercepto aleatório para região será dado no nível 1, por:

$$y_{ij} = \beta_{0j} + \sum_{m=1}^9 \beta_m x_{mij} + \varepsilon_{ij}, \quad (6.10)$$

com $i = 1, 2, \dots, n_j$ e $j = 1, 2, \dots, 15$, onde:

- y_{ij} representa a proficiência em matemática do i -ésimo aluno da j -ésima região;
- $\varepsilon_{ij} \sim N(0, \sigma^2)$ e todos os ε_{ij} são independentes.

Temos também que $\beta_{0j} = \gamma_{00} + v_{0j}$, onde γ_{00} é a proficiência em matemática esperada de cada região, v_{0j} é o efeito aleatório da j -ésima região no intercepto do modelo (6.9) e $v_{0j} \sim N(0, \tau_{00})$. Na Tabela 12, temos as estimativas para o modelo (6.8), modelo (6.9) e modelo (6.10).

Tabela 12 – Modelo (6.8) tem intercepto aleatório para turma; Modelo (6.9) tem intercepto aleatório para escola; Modelo (6.10) tem intercepto aleatório para região.

	Modelo (6.8)		Modelo (6.9)		Modelo (6.10)	
Efeito fixo						
Parâmetro	Estimativa	Valor p	Estimativa	valor p	Estimativa	valor p
$\widehat{\gamma}_{00}$	105,80	< 0,001	107,20	< 0,001	106,40	< 0,001
$\widehat{\beta}_1$ (Dep. Admin.)	-4,09	< 0,001	-4,25	< 0,001	-4,61	< 0,001
$\widehat{\beta}_2$ (Localização)	-6,03	0,02	-5,85	0,04	-2,69	0,22
$\widehat{\beta}_3$ (Turno)	0,74	0,38	1,04	0,26	-0,24	0,75
$\widehat{\beta}_4$ (Q001)	-14,79	< 0,001	-14,62	< 0,001	-14,64	< 0,001
$\widehat{\beta}_5$ (Q004)	2,35	0,07	2,25	0,08	2,35	0,07
$\widehat{\beta}_6$ (Q019)	2,81	< 0,001	3,02	< 0,001	3,54	< 0,001
$\widehat{\beta}_7$ (Q023)	3,35	< 0,001	3,38	< 0,001	3,46	< 0,001
$\widehat{\beta}_8$ (Q049)	-9,78	< 0,001	-10,03	< 0,001	-9,98	< 0,001
$\widehat{\beta}_9$ (Nota Port.)	0,63	< 0,001	0,63	< 0,001	0,63	< 0,001
Efeito aleatório						
Nível 1 (Aluno)						
Var. (Resíduo)	1116,19	-	1117,53	-	1152,53	-
Nível 2 (Turma)						
Var. (intercepto)	52,43	-				
Nível 2 (Escola)						
Var. (intercepto)			51,63	-		
Nível 2 (Região)						
Var. (intercepto)					8,234	-

Observando a Tabela 12, podemos notar que há evidências, a um nível de significância de 5%, que as variáveis explicativas turno e distorção idade série são não significativas nos três

modelos disponíveis. Quando tomamos interceptos aleatórios para região, a variável explicativa localização não é significativa, diferentemente dos outros modelos.

Assim, a interpretação das estimativas do modelo (6.8), modelo (6.9) e do modelo (6.10) são parecidas com a interpretação das estimativas do modelo (6.7).

6.2.3 Modelos com intercepto e inclinações aleatórios

O modelo linear com intercepto e inclinações aleatórios para turma será dado por:

$$y_{ij} = (\beta_{0j} + v_{0j}) + \sum_{m=1}^9 (\beta_{mj} + v_{mj})x_{mij} + \varepsilon_{ij}, \quad (6.11)$$

com $i = 1, 2, \dots, n_j$ e $j = 1, 2, \dots, 2187$, onde:

- y_{ij} representa a proficiência em matemática do i -ésimo aluno para a j -ésima turma;
- $\varepsilon_{ij} \sim N(0, \sigma^2)$ e todos os ε_{ij} são independentes;
- v_{0j} é o efeito aleatório no intercepto β_{0j} da j -ésima turma;
- v_{mj} é o efeito aleatório na inclinação β_{mj} da j -ésima turma, onde $m = 1, \dots, 9$.

As suposições aqui são que: $v_{mj} \sim N(0, \tau_{mm})$, onde os coeficientes aleatórios v_{mj} são independentes entre si e também de cada ε_{ij} , onde $m = 0, \dots, 9$.

O modelo linear com intercepto e inclinações aleatórios para escola será dado por:

$$y_{ij} = (\beta_{0j} + v_{0j}) + \sum_{m=1}^9 (\beta_{mj} + v_{mj})x_{mij} + \varepsilon_{ij}, \quad (6.12)$$

com $i = 1, 2, \dots, n_j$ e $j = 1, 2, \dots, 614$, onde:

- y_{ij} representa a proficiência em matemática do i -ésimo aluno para a j -ésima escola;
- $\varepsilon_{ij} \sim N(0, \sigma^2)$ e todos os ε_{ij} são independentes;
- v_{0j} é o efeito aleatório no intercepto β_{0j} da j -ésima escola;
- v_{mj} é o efeito aleatório na inclinação β_{mj} da j -ésima escola, onde $m = 1, \dots, 9$.

O modelo linear com intercepto e inclinações aleatórios para escola será dado por:

$$y_{ij} = (\beta_{0j} + v_{0j}) + \sum_{m=1}^9 (\beta_{mj} + v_{mj})x_{mij} + \varepsilon_{ij}, \quad (6.13)$$

com $i = 1, 2, \dots, n_j$ e $j = 1, 2, \dots, 15$, onde:

- y_{ij} representa a proficiência em matemática do i -ésimo aluno para a j -ésima região;
- $\varepsilon_{ij} \sim N(0, \sigma^2)$ e todos os ε_{ij} são independentes;
- v_{0j} é o efeito aleatório no intercepto β_{0j} da j -ésima região;
- v_{mj} é o efeito aleatório na inclinação β_{mj} da j -ésima região, onde $m = 1, \dots, 9$.

As suposições para os modelos (6.12) e (6.13) são as mesmas do modelo (6.11). Na Tabela 13, temos as estimativas para o modelo (6.11), modelo (6.12) e modelo (6.13).

Tabela 13 – Modelo (6.11) tem intercepto e inclinações aleatórias para turma; Modelo (6.12) tem intercepto e inclinações aleatórias para escola; Modelo (6.13) tem intercepto e inclinações aleatórias para região.

	Modelo (6.11)		Modelo (6.12)		Modelo (6.13)	
Efeito fixo						
Parâmetro	Estimativa	Valor p	Estimativa	valor p	Estimativa	valor p
$\widehat{\beta}_{0j}$ (Intercepto)	103,90	0,00	106,58	0,00	104,13	0,00
$\widehat{\beta}_{1j}$ (Dep. Admin.)	-4,19	0,00	-3,87	0,00	-8,42	0,00
$\widehat{\beta}_{2j}$ (Localização)	-5,74	0,00	-4,75	0,25	-1,56	0,63
$\widehat{\beta}_{3j}$ (Turno)	0,94	0,22	0,61	0,53	1,78	0,08
$\widehat{\beta}_{4j}$ (Q001)	-14,77	0,00	-14,53	0,00	-14,07	0,00
$\widehat{\beta}_{5j}$ (Q004)	2,04	0,12	2,70	0,05	2,93	0,05
$\widehat{\beta}_{6j}$ (Q019)	2,93	0,00	2,92	0,00	4,46	0,00
$\widehat{\beta}_{7j}$ (Q023)	3,26	0,00	3,32	0,00	2,75	0,00
$\widehat{\beta}_{8j}$ (Q049)	-9,84	0,00	-9,79	0,00	-10,82	0,00
$\widehat{\beta}_{9j}$ (Nota Port.)	0,64	0,00	0,62	0,00	0,64	0,00
Efeito aleatório						
Nível 1 (Aluno)						
Var. (Resíduo)	134,17	-	32,69	-	33,81	-
Nível 2	Turma		Escola		Região	
Var. (\widehat{v}_{0j})	$6,65 \cdot 10^{-11}$	-	11,75	-	7,40	-
var. (\widehat{v}_{1j})	$5,54 \cdot 10^{-14}$	-	8,88	-	5,12	-
var. (\widehat{v}_{2j})	$3,90 \cdot 10^{-7}$	-	12,04	-	7,04	-
var. (\widehat{v}_{3j})	$1,33 \cdot 10^{-14}$	-	5,98	-	2,06	-
var. (\widehat{v}_{4j})	$1,02 \cdot 10^{-13}$	-	4,65	-	2,37	-
var. (\widehat{v}_{5j})	$2,11 \cdot 10^{-11}$	-	9,03	-	2,04	-
var. (\widehat{v}_{6j})	$1,10 \cdot 10^{-13}$	-	6,11	-	1,91	-
var. (\widehat{v}_{7j})	$6,71 \cdot 10^{-14}$	-	4,80	-	1,75	-
var. (\widehat{v}_{8j})	$1,37 \cdot 10^{-13}$	-	8,07	-	2,89	-
var. (\widehat{v}_{9j})	$3,26 \cdot 10^{-9}$	-	0,06	-	0,02	-

Observando a Tabela 13, considerando um nível de 5% de significância, há evidências que:

- as covariáveis turno e série são não significativas no modelo (6.11);
- as covariáveis localização e turno são não significativas no modelo (6.12);
- as covariáveis localização e turno são não significativas no modelo (6.13).

Podemos ainda notar nos três modelos da Tabela 13, que o fato do aluno declarar ser do sexo feminino e já ter obtido reprovação escolar são os fatores que mais influenciam negativamente, fixado as outras covariáveis. Por outro lado, é possível destacar que o fato dos pais dos estudantes terem pelo menos ingressado no ensino superior, exerce uma influência positiva na proficiência em matemática.

Na próxima seção realizamos seleção de modelos, e decidimos qual o melhor modelo clássico deve ser utilizado a partir das covariáveis testadas.

6.3 Seleção de Modelos Clássicos

Primeiramente, comparamos os modelos ajustados na seção anterior e daí definimos qual será a metodologia de ajuste a ser utilizada, para isso utilizamos três critérios de informação, são eles: AIC, AICc e BIC. Na Tabela 14, temos os valores dos critérios de informação AIC, BIC e AICc, para cada um dos sete modelos, observe que nos três critérios o modelo (6.10) tem o menor valor, assim este deve ser o modelo que deve ser escolhido.

Tabela 14 – Comparação dos modelos ajustados nesta seção utilizando os critérios AIC, BIC, AICc.

	AIC	BIC	AICc
Modelo (6.7)	103320,4	103400,2	103320,4
Modelo (6.8)	103269,9	103356,9	103269,9
Modelo (6.9)	103207,2	103294,3	103207,3
Modelo (6.10)	103195,4	103282,4	103195,4
Modelo (6.11)	103425,1	103903,7	103425,9
Modelo (6.12)	103293,1	103771,7	103294,0
Modelo (6.13)	103260,0	103738,6	103260,8

A metodologia de ajuste clássico escolhida é um modelo de regressão multinível com dois níveis com intercepto aleatório para região. Ao analisarmos as estimativas do modelo (6.10), conforme já destacamos anteriormente, podemos notar que as variáveis explicativas turno, localização e idade são não significativas, isto pode ser visto, há uma nível de significância de 5%, observando os valores p associado a cada um dos coeficientes. Assim ajustamos sete novos modelos com todas as variações possíveis de covariáveis, levando em consideração o descarte ou não de uma, duas ou das três covariáveis não significativas para o modelo (6.10) e posteriormente

calculamos o AIC, BIC e AICc para os três modelos. Na Tabela 15, temos a configuração desses ajustes.

Tabela 15 – Configuração de ajustes para seleção de modelos.

Covariáveis	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	Mod. 6	Mod. 7
PROFIC._LP_SAEB	x	x	x	x	x	x	x
DEPENDEN._ADM	x	x	x	x	x	x	x
LOCALIZAÇÃO		x	x			x	
TURN0	x		x		x		
Q001 (Sexo)	x	x	x	x	x	x	x
Q004 (Idade)	x	x		x			
Q019 (Escol. Mãe)	x	x	x	x	x	x	x
Q023 (Escol. Pai)	x	x	x	x	x	x	x
Q049 (Reprovas)	x	x	x	x	x	x	x

Nas Tabelas 16, 17 e 18 temos as estimativas dos modelos descritos na Tabela 15, reparem que dos sete modelos que estão apresentados nessas tabelas, apenas no modelo 7 temos todas as covariáveis significativas, para isso basta observar o valor p associado a cada coeficiente estimado, a um nível de significância de 5%.

Tabela 16 – Examinando as estimativas dos parâmetros e o valor p dos mesmos, nos modelos 1, 2 e 3 descritos na Tabela 15.

	Modelo 1		Modelo 2		Modelo 3	
Efeito fixo						
Parâmetro	Estimativa	Valor p	Estimativa	valor p	Estimativa	valor p
$\hat{\gamma}_{00}$	103,87	0,00	106,50	0,00	108,40	0,00
$\hat{\beta}_1$ (Dep. Admin.)	-4,66	0,00	-4,60	0,00	-4,57	0,00
$\hat{\beta}_2$ (Localização)	-	-	-2,65	0,23	-2,68	0,22
$\hat{\beta}_3$ (Turno)	0,19	0,80	-	-	0,26	0,73
$\hat{\beta}_4$ (Q001)	-14,63	0,00	-14,63	0,00	-14,61	0,00
$\hat{\beta}_5$ (Q004)	2,34	0,07	2,35	0,07	-	-
$\hat{\beta}_6$ (Q019)	3,51	0,00	3,54	0,00	3,57	0,00
$\hat{\beta}_7$ (Q023)	3,43	0,00	3,46	0,00	3,49	0,00
$\hat{\beta}_8$ (Q049)	-9,98	0,00	-9,98	0,00	-10,93	0,00
$\hat{\beta}_9$ (Nota Port.)	0,63	0,00	0,63	0,00	0,63	0,00
Efeito aleatório						
Nível 1 (Aluno)						
Var. (Resíduo)	1152,57	-	1152,42	-	1152,78	-
Nível 2 (Região)						
Var. (intercepto)	8,37	-	8,28	-	8,14	-

Na Tabela 19, temos o valor dos três critérios de seleção (AIC, AICc e BIC) para o modelo completo (6.10) e para os sete modelos reduzidos. Ao analisarmos o desempenho do três critérios que estão sendo utilizados, constatamos que o AIC e o AICc tem comportamentos parecidos,

Tabela 17 – Examinando as estimativas dos parâmetros e o valor p dos mesmos, nos modelos 4, 5 e 6 descritos na Tabela 15.

	Modelo 4		Modelo 5		Modelo 6	
Efeito fixo						
Parâmetro	Estimativa	Valor p	Estimativa	valor p	Estimativa	valor p
$\widehat{\gamma}_{00}$	104,01	0,00	105,92	0,00	108,56	0,00
$\widehat{\beta}_1$ (Dep. Admin.)	-4,66	0,00	-4,63	0,00	-4,57	0,00
$\widehat{\beta}_2$ (Localização)	-	-	-	-	-2,64	0,23
$\widehat{\beta}_3$ (Turno)	-	-	0,21	0,78	-	-
$\widehat{\beta}_4$ (Q001)	-14,63	0,00	-14,61	0,00	-14,61	0,00
$\widehat{\beta}_5$ (Q004)	2,35	0,07	-	-	-	-
$\widehat{\beta}_6$ (Q019)	3,51	0,00	3,55	0,00	3,57	0,00
$\widehat{\beta}_7$ (Q023)	3,44	0,00	3,46	0,00	3,49	0,00
$\widehat{\beta}_8$ (Q049)	-9,98	0,00	-10,93	0,00	-10,94	0,00
$\widehat{\beta}_9$ (Nota Port.)	0,63	0,00	0,63	0,00	0,63	0,00
Efeito aleatório						
Nível 1 (Aluno)						
Var. (Resíduo)	1152,46	-	1152,81	-	1152,67	-
Nível 2 (Região)						
Var. (intercepto)	8,41	-	8,27	-	8,19	-

isto se deve ao fato de termos 12.280 observações, como já foi destacado anteriormente o AIC têm problemas assintóticos, ou seja, quando temos poucas observações ele deve ser substituído por outros critérios.

Observando a Tabela 19, vemos que os critérios AIC e o AICc optam pelo modelo 2, que exclui apenas a covariável turno, contudo as covariáveis localização e Q004(mede a distorção idade-série) são não significativas para a variável resposta. O critério BIC seleciona o modelo 7, que exclui as covariáveis turno, localização e Q004(mede a distorção idade-série). Como o modelo 7 não tem covariáveis não significativas, fato esse que pode ser notado via teste t de significância individual com 5% de significância, ele será escolhido como melhor modelo sob enfoque clássico, lembrando sempre que esse modelo não é necessariamente o modelo verdadeiro e sim o melhor dentre aqueles aqui ajustados.

Por fim sugerimos o uso do critério BIC para modelos multiníveis com dados normais, pois além de selecionar modelo apenas com variáveis significativas, seleciona também modelo com menor número de covariáveis, facilitando a interpretação das estimativas. Fato esse que [Emiliano \(2009\)](#), verificou ao concluir em seu trabalho que para dados simulados com distribuição normal, o critério BIC é o mais eficiente.

Na Figura 8 temos o gráfico dos resíduos por região, na Figura 9 temos o box plot dos resíduos por região, na Tabela 20 temos a estimativa dos efeitos aleatórios associados a cada uma das regiões que foram estimados no modelo 7.

Tabela 18 – Examinando as estimativas dos parâmetros e o valor p do modelo 7 descritos na Tabela 15.

	Modelo 7	
Efeito fixo		
Parâmetro	Estimativa	Valor p
$\widehat{\gamma}_0$	106,08	0,00
$\widehat{\beta}_1$ (Dep. Admin.)	-4,62	0,00
$\widehat{\beta}_2$ (Localização)	-	-
$\widehat{\beta}_3$ (Turno)	-	-
$\widehat{\beta}_4$ (Q001)	-14,61	0,00
$\widehat{\beta}_5$ (Q004)	-	-
$\widehat{\beta}_6$ (Q019)	3,55	0,00
$\widehat{\beta}_7$ (Q023)	3,47	0,00
$\widehat{\beta}_8$ (Q049)	-10,93	0,00
$\widehat{\beta}_9$ (Nota Port.)	0,63	0,00
Efeito aleatório		
Nível 1 (Aluno)		
Var. (Resíduo)	1152,70	-
Nível 2 (Região)		
Var. (intercepto)	8,31	-

Tabela 19 – Comparação dos modelos clássicos da Tabela 15 mais o modelo completo (6.10), utilizando os critérios AIC, AICc e BIC.

	AIC	BIC	AICc
Modelo (6.10)	103195,4	103282,4	103195,4
Modelo 1	103198,3	103278,1	103198,3
Modelo 2	103194,8	103274,6	103194,9
Modelo 3	103198,9	103278,7	103199,0
Modelo 4	103197,7	103270,2	103197,7
Modelo 5	103201,8	103274,4	103201,9
Modelo 6	103198,4	103270,9	103198,4
Modelo 7	103201,2	103266,5	103201,3

Ao analisar as estimativas do modelo 7 (18), podemos concluir que os alunos do sexo feminino tem desempenho inferior em matemática fixado as outras variáveis, o mesmo podemos afirmar com relação aos alunos que foram reprovados pelo menos uma vez em sua vida escolar. Por outro lado vemos que alunos cuja mãe ou mulher responsável pelo menos ingressou no ensino superior tem um desempenho superior em matemática, fixado os outros preditores, similarmente, os alunos cujo pai ou homem responsável por ele, ao menos ingressou no ensino superior têm um desempenho superior aos demais.

Analisando a Tabela 20, temos que as regiões que têm pior desempenho em matemática na Prova Brasil são as regiões de São Carlos e São Paulo, por outro lado as regiões que têm melhor desempenho em matemática na Prova Brasil são as regiões de Jales e Piracicaba, fixando todas as covariáveis com sendo zero. Quando observamos as Figuras 8 até 9, podemos validar as

Tabela 20 – Estimativas dos efeitos aleatórios do modelo 7

Região	Estimativa
Araçatuba	-0,20
Bauru	-0,09
Campinas	1,75
Itapetininga	0,17
Jales	2,51
Jundiaí	1,73
Marília	-0,27
Piracicaba	2,47
Presidente Prudente	-0,86
Registro	-1,49
Ribeirão Preto	1,62
São Carlos	-3,14
São José do Rio Preto	1,41
São José dos Campos	1,15
São Paulo	-6,78

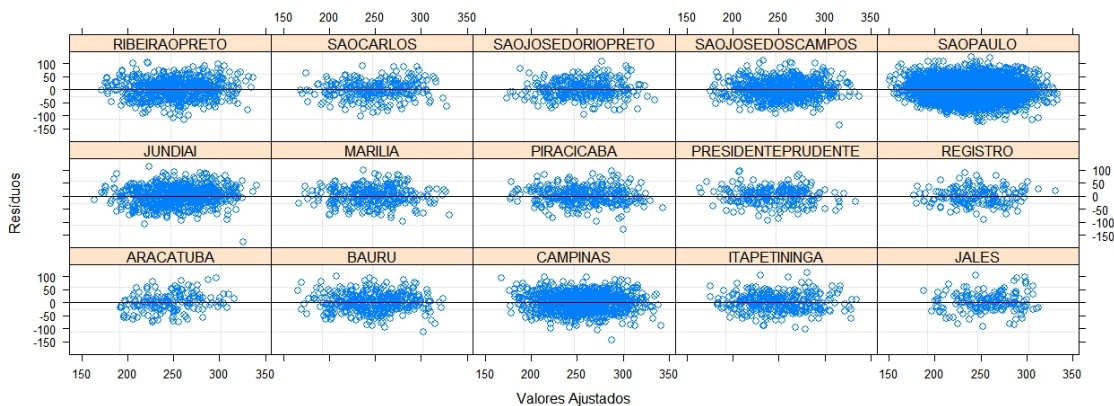


Figura 8 – Análise residual do modelo 7 por grupos.

suposições do modelo, isto é, os erros têm variância constante e têm distribuição normal com média zero.

Note que, embora na Figura 3 não haja indícios de diferenças entre a proficiência em matemática das 15 regiões, a inclusão do efeito aleatório para região se justifica, conforme nós pudemos notar ao longo dessa seção. Na próxima seção temos a análise bayesiana, com efeitos aleatórios para escola.

6.4 Modelagem Bayesiana

Nesta seção ajustamos modelos com abordagem bayesiana. Primeiramente estimamos os parâmetros de três modelos e comparamos três metodologias distintas, são elas: modelo com efeitos fixos, modelo com interceptos aleatórios e modelo com interceptos e inclinações

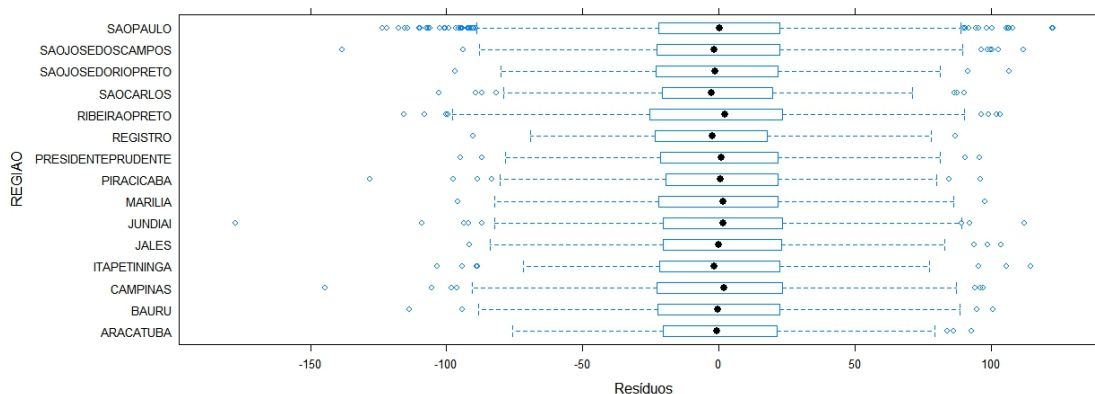


Figura 9 – Box plot dos resíduos do modelo 7 por grupos.

aleatórias. Aqui usamos efeitos aleatórios de escola, para ajustar esses modelos utilizamos a interface Stan, adaptada ao software R, para mais detalhes a respeito, ver por exemplo, [Sorensen e Vasishth \(2015\)](#). O Stan utiliza como método de estimação o processo MCMC com dinâmica hamiltoniana, mais detalhes a respeito pode ser visto em [Neal et al. \(2011\)](#), em cada cadeia ele trabalha com um período de aquecimento, ou seja, na primeira metade da cadeia há ajustes à distribuição dos dados, o período de aquecimento pode ser ajustado pelo usuário.

Para os estudos realizados nesta seção foram excluídos os missings dos dados, ficando assim com 10311 alunos, o que corresponde a aproximadamente 83,96% dos dados da amostra. No Apêndice D, temos os códigos no stan que foram usados para a obtenção das estimativas desta seção, ressaltamos que foram utilizadas distribuições *a priori* não informativas para os efeitos fixos. As covariáveis aqui utilizadas foram escolhidas utilizando critérios empíricos, conforme visto no capítulo 2.

6.4.1 Modelo com efeitos fixos

Considere um modelo com a seguinte estrutura:

$$y_i \sim N(\mu, \sigma_\epsilon), \quad (6.14)$$

onde $i = 1, \dots, 10311$. Temos ainda que $\mu = \beta_0 + \sum_{m=1}^9 \beta_m x_{mi}$. Esse é o modelo de efeitos fixos, suas estimativas poderão ser encontradas na Tabela 21.

Na Tabela 21, temos a distribuição *a posteriori* conjunta dos parâmetros $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9$ e σ_ϵ do modelo (6.14), observe que o intervalo com 95% de credibilidade para β_3 e β_5 contém o zero, sinalizando que as covariáveis turno e idade, nesse caso é não significativa para a variável resposta desse modelo, note que a estatística \hat{R} é 1 em todos os casos sinalizando que as cadeias de distribuições que contêm 2.000 vetores de parâmetros estimados

Tabela 21 – Estimativas dos parâmetros, intervalos de credibilidade e a estatística \widehat{R} no modelo (6.14).

Parâmetro	média	2,5%	97,5%	\widehat{R}
$\widehat{\beta}_0$ (Intercepto)	103,92	97,99	109,74	1
$\widehat{\beta}_1$ (Dep. Admin.)	-4,22	-5,59	-2,85	1
$\widehat{\beta}_2$ (Localização)	-5,76	-9,84	-1,74	1
$\widehat{\beta}_3$ (Turno)	0,95	-0,46	2,37	1
$\widehat{\beta}_4$ (Q001)	-14,75	-16,12	-13,45	1
$\widehat{\beta}_5$ (Q004)	2,10	-0,50	4,78	1
$\widehat{\beta}_6$ (Q019)	2,92	1,38	4,48	1
$\widehat{\beta}_7$ (Q023)	3,32	1,74	4,74	1
$\widehat{\beta}_8$ (Q049)	-9,83	-11,83	-7,79	1
$\widehat{\beta}_9$ (Nota Port.)	0,65	0,63	0,65	1
$\widehat{\sigma}_\varepsilon$	34,19	33,77	34,63	1

convergem.

6.4.2 Modelo com interceptos aleatórios

Vamos considerar agora um modelo com a seguinte estrutura:

$$\begin{aligned} y_{ij} &\sim N(\mu, \sigma_\varepsilon + \sigma_v), \\ \varepsilon_{ij} &\sim N(0, \sigma_\varepsilon), \end{aligned} \quad (6.15)$$

onde $i = 1, \dots, 10311$ e $j = 1, \dots, 614$. Temos ainda que $\mu = \beta_{0j} + \sum_{m=1}^9 \beta_m x_{mij}$. Temos ainda que $\beta_{0j} \sim N(0, \sigma_v)$. Na Tabela 22, vemos a distribuição conjunta *a posteriori* dos parâmetros β_{0j} , β_1 , β_2 , β_3 , β_4 , β_5 , β_6 , β_7 , β_8 , β_9 , σ_ε e σ_v do modelo, foram estimados 10.000 vetores de parâmetros em 1 cadeia onde os 5.000 primeiros são um período de aquecimento, observe que a estatística \widehat{R} é 1, indicando que as cadeias estão convergindo.

Observe na Tabela 22 que o intervalo com 95% de credibilidade para a estimativa *a posteriori* marginal de β_3 e β_5 , contém o 0 indicando novamente que as covariáveis turno e idade são não significativa, note também que há uma pequena diminuição $\widehat{\sigma}_\varepsilon$, isto ocorre devido a existência do componente de variância $\widehat{\sigma}_v$.

6.4.3 Modelo com interceptos e inclinações aleatórios

O modelo com intercepto e inclinações aleatórias para escola será dado pela equação:

$$y_{ij} = (\beta_{0j} + v_{0j}) + \sum_{m=1}^9 (\beta_{mj} + v_{mj}) x_{mij} + \varepsilon_{ij}, \quad (6.16)$$

Tabela 22 – Examinando as estimativas dos parâmetros, intervalos de credibilidade e a estatística \hat{R} no modelo (6.15).

Parâmetro	média	2,5%	97,5%	\hat{R}
$\hat{\beta}_0$ (Intercepto)	107,16	100,29	114,40	1
$\hat{\beta}_1$ (Dep. Admin.)	-4,26	-6,23	-2,26	1
$\hat{\beta}_2$ (Localização)	-5,83	-11,51	-0,37	1
$\hat{\beta}_3$ (Turno)	1,05	-0,82	2,92	1
$\hat{\beta}_4$ (Q001)	-14,62	-15,91	-13,32	1
$\hat{\beta}_5$ (Q004)	2,25	-0,28	4,83	1
$\hat{\beta}_6$ (Q019)	3,02	1,44	4,63	1
$\hat{\beta}_7$ (Q023)	3,40	1,77	5,01	1
$\hat{\beta}_8$ (Q049)	-10,04	-11,94	-8,07	1
$\hat{\beta}_9$ (Nota Port.)	0,63	0,62	0,65	1
$\hat{\sigma}_\varepsilon$	32,44	32,97	33,92	1
$\hat{\sigma}_v$	7,20	6,25	8,17	1

onde $i = 1, \dots, 10311$ e $j = 1, \dots, 614$. Na Tabela 23, temos a distribuição de probabilidade *a posteriori* conjunta dos parâmetros do modelo (6.16), inclusive dos componentes de variância associados a cada um dos efeitos aleatórios.

No modelo (6.16), o intervalo com 95% de credibilidade para β_2 , β_3 e β_5 contém o 0, sendo assim, há indícios de que a localização da escola não tem interferência na proficiência em matemática dos alunos avaliados, assim turno em que o aluno estuda bem como sua idade.

6.5 Seleção de Modelos Bayesianos

Agora vamos selecionar qual tipo de modelo será utilizado, para isso usamos os critérios de informação *looic*, WAIC, DIC, EAIC e EBIC, vistos no capítulo 4. Na Tabela 24, temos o valor desses critérios de informação para cada um dos modelos bayesianos ajustados até aqui.

Analisando a Tabela 24, podemos ver que em qualquer um dos cinco critérios o modelo (6.16) tem o menor valor. Assim o melhor modelo seria um modelo com intercepto e inclinações aleatórias com efeito aleatório para escola.

Na Tabela 23, podemos ver as estimativas do modelo (6.16). Observe que há três covariáveis (Localização, Turno e Q004) cujo intervalo de credibilidade contém o zero, o que nos fornece evidências de que elas podem não ser significativas para modelar a variabilidade de PROFICIENCIA_MT_SAEB. Para analisar essa questão com mais profundidade ajustamos mais sete modelos, cuja configuração pode ser vista na Tabela 15.

Na Tabela 25, temos as estimativas dos modelos 1, 2 e 3, da mesma forma na Tabela 26, temos as estimativas dos modelos 4, 5 e 6, e na Tabela 27 temos as estimativas do modelo 7. Observe que, em cada um dos modelos, foi calculado a média *a posteriori* conjunta dos

Tabela 23 – Examinando as estimativas dos parâmetros, intervalos de credibilidade e a estatística \hat{R} no modelo (6.16).

Parâmetro	média	2,5%	97,5%	\hat{R}
$\hat{\beta}_0$ (Intercepto)	107,04	99,85	114,34	1
$\hat{\beta}_1$ (Dep. Admin.)	-4,16	-6,20	-2,10	1
$\hat{\beta}_2$ (Localização)	-5,48	-11,10	0,16	1
$\hat{\beta}_3$ (Turno)	0,86	-0,99	2,71	1
$\hat{\beta}_4$ (Q001)	-14,59	-15,94	-13,23	1
$\hat{\beta}_5$ (Q004)	2,39	-0,24	5,01	1
$\hat{\beta}_6$ (Q019)	2,95	1,34	4,55	1
$\hat{\beta}_7$ (Q023)	3,36	1,72	4,96	1
$\hat{\beta}_8$ (Q049)	-9,95	-11,99	-7,93	1
$\hat{\beta}_9$ (Nota Port.)	0,63	0,61	0,64	1
$\hat{\sigma}_\varepsilon$	32,24	32,77	33,71	1
$\hat{\sigma}_{v0}$	2,17	0,08	5,83	1
$\hat{\sigma}_{v1}$	2,74	0,13	6,75	1
$\hat{\sigma}_{v2}$	2,41	0,12	6,20	1
$\hat{\sigma}_{v3}$	2,04	0,09	5,38	1
$\hat{\sigma}_{v4}$	1,87	0,07	4,88	1
$\hat{\sigma}_{v5}$	4,45	0,41	7,87	1
$\hat{\sigma}_{v6}$	3,29	0,19	6,82	1
$\hat{\sigma}_{v7}$	1,66	0,07	4,45	1
$\hat{\sigma}_{v8}$	3,75	0,25	7,75	1
$\hat{\sigma}_{v9}$	0,02	0,00	0,04	1

Tabela 24 – Comparação dos modelos ajustados utilizando os critérios *loaic*, WAIC, DIC, EAIC e EBIC.

	EAIC	EBIC	WAIC	<i>loaic</i>	DIC
modelo (6.14)	103330,8	103410,6	103319,3	103319,3	103319,10
modelo (6.15)	102871,9	102958,9	103131,3	103131,8	-4017,71
modelo (6.16)	102765,6	102917,9	103112,1	103113,9	-431541

parâmetros em questão. Na Tabela 28 comparamos os sete modelos reduzido ajustados e também com o modelo completo (Modelo (6.16)), para isso utilizamos os critérios de informação EAIC, EBIC, WAIC, Looic e DIC.

Observando a Tabela 28, temos que os critérios EBIC, WAIC, *loaic* indicam que o Modelo 4 (sem as covariáveis localização e turno) é o melhor. Já o critério EAIC, nos fala que o Modelo 1 (sem a covariável localização) é o melhor. O Critério DIC nos leva a optar pelo Modelo 6 (sem as covariáveis turno e idade). Note que os valores do critério de informação desvio (DIC), diferem dos demais em todos os oito modelos disponíveis na Tabela 28, isto se deve ao fato de que ao se construir o DIC usa-se uma medida que quantifica a complexidade do modelo, que segundo Spiegelhalter *et al.* (2014) e Gelman e Hill (2006), para modelos com muitos parâmetros, que é o caso do modelo multinível com dois níveis, essa medida se distancia

Tabela 25 – Examinando as estimativas dos parâmetros, intervalos de credibilidade dos modelos 1, 2 e 3 descritos na Tabela 15.

Parâmetro	Modelo 1 média (I.C.)	Modelo 2 média (I.C.)	Modelo 3 média (I.C.)
$\hat{\beta}_0$ (Intercepto)	101,77 (96,77;106,73)	107,60 (100,47;114,73)	108,54 (101,54;115,52)
$\hat{\beta}_1$ (Dep. Admin.)	-4,20 (-6,26;-2,16)	-4,12 (-6,16;-2,11)	-4,02 (-6,07;-1,95)
$\hat{\beta}_2$ (Localização)	-	-5,49 (-11,05;0,01)	-5,07 (-10,77;0,56)
$\hat{\beta}_3$ (Turno)	0,86 (-1,05;2,65)	-	0,91 (-0,97;2,75)
$\hat{\beta}_4$ (Q001)	-14,58 (-15,94;-13,24)	-14,60 (-15,97;-13,23)	-14,60 (-15,99;-13,21)
$\hat{\beta}_5$ (Q004)	2,40 (-0,22;5,00)	2,42 (-0,22;5,10)	-
$\hat{\beta}_6$ (Q019)	2,93 (1,37;4,57)	2,96 (1,34;4,56)	3,06 (1,43;4,68)
$\hat{\beta}_7$ (Q023)	3,31 (1,71;4,93)	3,36 (1,71;4,99)	3,36 (1,74;5,03)
$\hat{\beta}_8$ (Q049)	-9,93 (-11,95;-7,92)	-9,95 (-11,93;-7,93)	-10,96 (-12,66;-9,25)
$\hat{\beta}_9$ (Nota Port.)	0,63 (0,61;0,64)	0,63 (0,61;0,65)	0,63 (0,62;0,65)
$\hat{\sigma}_\varepsilon$	32,23 (32,76;32,72)	32,24 (32,76;33,72)	33,26 (32,79;33,74)
$\hat{\sigma}_{v0}$	2,42 (0,11;6,35)	2,33 (0,10;6,22)	2,59 (0,12;6,61)
$\hat{\sigma}_{v1}$	3,24 (0,17;7,40)	3,16 (0,15;7,25)	3,21 (0,17;7,58)
$\hat{\sigma}_{v2}$	-	2,44 (0,09;6,42)	2,80 (0,14;6,87)
$\hat{\sigma}_{v3}$	2,07 (0,07;5,65)	-	2,15 (0,09;5,70)
$\hat{\sigma}_{v4}$	1,95 (0,10;4,94)	1,88 (0,08;4,87)	1,96 (0,09;4,95)
$\hat{\sigma}_{v5}$	4,20 (0,33;7,72)	4,62 (0,54;8,09)	-
$\hat{\sigma}_{v6}$	3,16 (0,19;6,59)	3,30 (0,22;6,77)	3,28 (0,22;6,75)
$\hat{\sigma}_{v7}$	1,67 (0,07;4,35)	1,71 (0,07;4,60)	1,71 (0,07;4,52)
$\hat{\sigma}_{v8}$	3,73 (0,20;7,81)	3,67 (0,20;7,65)	4,04 (0,26;7,99)
$\hat{\sigma}_{v9}$	0,03 (0,01;0,04)	0,02 (0,00;0,04)	0,03 (0,01;0,04)

do verdadeiro número de parâmetros, ocasionando assim discrepâncias como por exemplo, o número efetivo de parâmetros negativo.

Analisamos com mais detalhes o Modelo 4. Na Figura 10 temos o gráfico da densidade *a posteriori* dos parâmetros, bem como da variância dos erros e dos componentes de variância desse modelo, na Figura 11, temos os histogramas da densidade *a posteriori* simulada via MCMC com dinâmica hamiltoniana dos parâmetros do modelo, na Figura 12 vemos o desenvolvimento da cadeia de 20.000 iterações, observe que não há padrões que possam indicar falta de convergências.

Observando a Figura 10 e a Figura 11, os parâmetros aparentemente tem distribuições simétricas, assim como a variância dos erros, já os componentes de variância notamos que eles tem distribuição assimétrica à direita.

Observando as estimativas do modelo bayesiano 4 na Tabela 26, podemos notar que há indícios que o aluno declaram ser do sexo feminino e ter sido reprovado influencia de maneira negativa no desempenho do aluno em matemática, fixando as outras covariáveis como sendo zero. Por outro lado, os alunos cujos pais são mais escolarizados tendem a ter um desempenho mais elevado em matemática, destaca-se ainda que apesar do modelo 4 for o escolhido por alguns

Tabela 26 – Examinando as estimativas dos parâmetros, intervalos de credibilidade dos modelos 4, 5 e 6 descritos na Tabela 15.

Parâmetro	Modelo 4 média (I.C.)	Modelo 5 média (I.C.)	Modelo 6 média (I.C.)
$\hat{\beta}_0$ (Intercepto)	102,33 (97,62;106,93)	103,69 (99,43;108,05)	109,03 (102,33;115,73)
$\hat{\beta}_1$ (Dep. Admin.)	-4,16 (-6,22;-2,12)	-4,07 (-6,13;-2,04)	-3,96 (-6,00;-1,90)
$\hat{\beta}_2$ (Localização)	-	-	-4,96 (-10,56;0,68)
$\hat{\beta}_3$ (Turno)	-	0,83 (-1,01;2,69)	-
$\hat{\beta}_4$ (Q001)	-14,59 (-15,92;-13,23)	-14,59 (-15,97;-13,24)	-14,60 (-15,94;-13,23)
$\hat{\beta}_5$ (Q004)	2,41 (-0,21;4,98)	-	-
$\hat{\beta}_6$ (Q019)	2,94 (1,34;4,51)	3,04 (1,44;4,68)	3,05 (1,40;4,71)
$\hat{\beta}_7$ (Q023)	3,32 (1,67;5,00)	3,31 (1,74;4,90)	3,38 (1,72;5,03)
$\hat{\beta}_8$ (Q049)	-9,94 (-11,95;-7,97)	-10,94 (-12,67;-9,28)	-10,96 (-12,69;-9,27)
$\hat{\beta}_9$ (Nota Port.)	0,63 (0,61;0,64)	0,63 (0,62;0,65)	0,63 (0,62;0,65)
$\hat{\sigma}_\varepsilon$	32,25 (32,76;32,73)	33,26 (32,79;33,76)	33,26 (32,79;33,74)
$\hat{\sigma}_{v0}$	2,35 (0,10;6,36)	2,70 (0,12;7,39)	2,68 (0,12;6,96)
$\hat{\sigma}_{v1}$	3,50 (0,17;7,69)	3,66 (0,17;8,05)	3,63 (0,16;8,09)
$\hat{\sigma}_{v2}$	-	-	2,84 (0,12;7,16)
$\hat{\sigma}_{v3}$	-	2,19 (0,10;5,59)	-
$\hat{\sigma}_{v4}$	1,90 (0,08;4,84)	1,93 (0,08;4,89)	1,92 (0,07;5,03)
$\hat{\sigma}_{v5}$	4,09 (0,23;7,59)	-	-
$\hat{\sigma}_{v6}$	3,10 (0,19;6,59)	3,17 (0,18;6,65)	3,24 (0,20;6,71)
$\hat{\sigma}_{v7}$	1,69 (0,07;4,47)	1,69 (0,06;4,61)	1,75 (0,07;4,68)
$\hat{\sigma}_{v8}$	3,72 (0,21;7,69)	3,85 (0,21;7,93)	4,09 (0,25;8,06)
$\hat{\sigma}_{v9}$	0,03 (0,01;0,04)	0,03 (0,02;0,04)	0,03 (0,01;0,04)

dos critérios de informação, o intervalo com 95% de credibilidade para o parâmetro associado a covariável que identifica a distorção idade série contém o zero.

6.6 Seleção de Variáveis

Nesta seção realizamos seleção de variáveis, ressaltamos que ao contrário de seleção de modelos, onde é feito a comparação do modelo completo com o modelo reduzido (sem algumas das covariáveis utilizadas anteriormente), utilizamos o modelo completo e daí extraímos os preditores que não são significativos, ou seja, dentre todas as variáveis explicativas disponíveis, devemos encontrar um subconjunto de variáveis importantes para o modelo.

Consideramos modelos de regressão multinível com dois níveis, com efeito aleatório de escola, cuja variável resposta é a proficiência em matemática, que conforme foi visto anteriormente, há indícios de que é aderente a uma distribuição normal. O objetivo aqui é utilizar o algoritmo glmLasso que descrevemos no capítulo 4 e realizar seleção de variáveis num contexto educacional. Para tanto consideramos todas as variáveis que são provenientes do questionário dos alunos e também as variáveis dependência administrativa, localização, turno, nota de

Tabela 27 – Examinando as estimativas dos parâmetros, intervalos de credibilidade do modelo 7 descritos na Tabela 15.

Parâmetro	Modelo 7 média (I.C.)
$\hat{\beta}_0$ (Intercepto)	104,24 (100,16;108,36)
$\hat{\beta}_1$ (Dep. Admin.)	-3,99 (-6,07;-1,89)
$\hat{\beta}_2$ (Localização)	-
$\hat{\beta}_3$ (Turno)	-
$\hat{\beta}_4$ (Q001)	-14,59 (-15,94;-13,25)
$\hat{\beta}_5$ (Q004)	-
$\hat{\beta}_6$ (Q019)	3,04 (1,42;4,70)
$\hat{\beta}_7$ (Q023)	3,32 (1,74;4,92)
$\hat{\beta}_8$ (Q049)	-10,97 (-12,68;-9,25)
$\hat{\beta}_9$ (Nota Port.)	0,63 (0,61;0,65)
$\hat{\sigma}_\varepsilon$	33,26 (32,78;33,74)
$\hat{\sigma}_{v0}$	2,81 (0,13;6,82)
$\hat{\sigma}_{v1}$	4,40 (0,40;8,62)
$\hat{\sigma}_{v2}$	-
$\hat{\sigma}_{v3}$	-
$\hat{\sigma}_{v4}$	1,92 (0,08;4,93)
$\hat{\sigma}_{v5}$	-
$\hat{\sigma}_{v6}$	3,17 (0,20;6,62)
$\hat{\sigma}_{v7}$	1,72 (0,07;4,52)
$\hat{\sigma}_{v8}$	4,16 (0,30;8,05)
$\hat{\sigma}_{v9}$	0,03 (0,02;0,04)

Tabela 28 – Comparação dos modelos com interceptos e inclinações aleatórias, utilizando os critérios *loaic*, WAIC, DIC, EAIC e EBIC.

	EAIC	EBIC	WAIC	<i>loaic</i>	DIC
Modelo (6.16)	102765,6	102917,9	103112,1	103113,9	-431541
Modelo 1	102760	102897,8	103110,9	103112,7	-600863,2
Modelo 2	102763,9	102901,7	103110,9	103112,7	-613839,8
Modelo 3	102903	103040,9	103242,9	103244,6	-559262,6
Modelo 4	102761	102884,3	103108,6	103110,4	-38513777
Modelo 5	102899,6	103022,9	103240,5	103242,2	-38676171
Modelo 6	102901,6	103024,9	103241,9	103243,6	-38700918
Modelo 7	102894,9	103003,7	103237,7	103239,4	-633529,9

matemática da escola, nota de português da escola, taxa de participação da escola, número de alunos presentes por escola, número de alunos matriculados por escola e proficiência em língua portuguesa. Ressaltamos ainda que as variáveis contínuas usadas nesse procedimento foram centralizadas. Mais detalhes a respeito das covariáveis que são utilizadas pode ser encontrado no apêndice A.

O primeiro passo é encontrar o λ ótimo, para tanto realizamos simulações computacionais

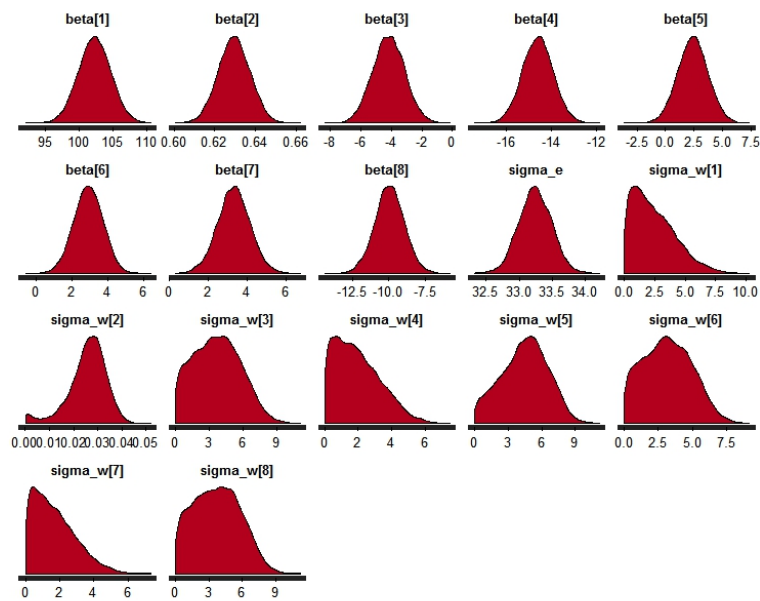


Figura 10 – Gráfico da função densidade marginal *a posteriori* dos parâmetros do Modelo 4.

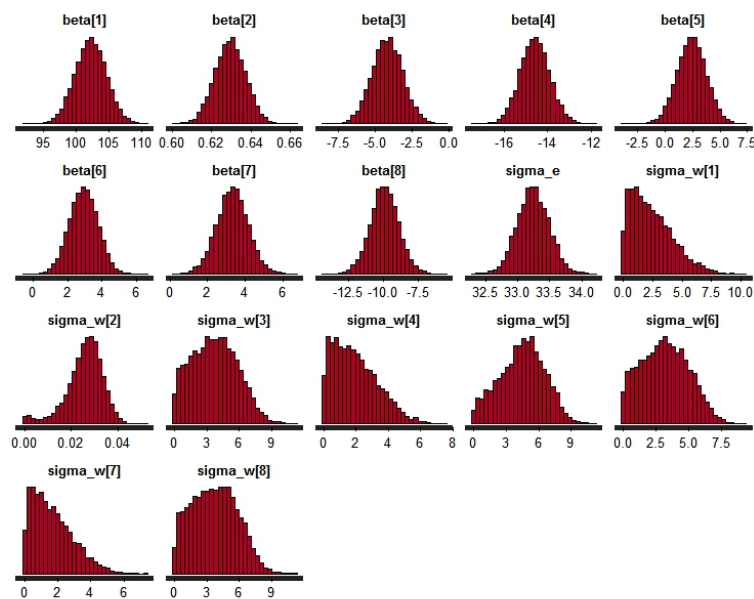


Figura 11 – Histogramas da densidade *a posteriori* simulada via MCMC com dinâmica hamiltoniana dos parâmetros do modelo 4.

comparando os modelos estimados e seus respectivos BIC , foi gerado um conjunto de 21 valores para λ variando de 0 até 1000. Constatamos que para $\lambda = 450$ foi estimado o modelo com o menor $BIC(-3403,94)$ e na Figura 13 temos o gráfico dos valores do BIC para os λ citados anteriormente.

Na Tabela 29 temos o modelo estimado via escore de Fisher, apenas com as covariáveis selecionadas. Observe que 34 preditores foram selecionados, vemos ainda como fatores que influenciam positivamente a proficiência em matemática são: mãe alfabetizada, leitura de livros

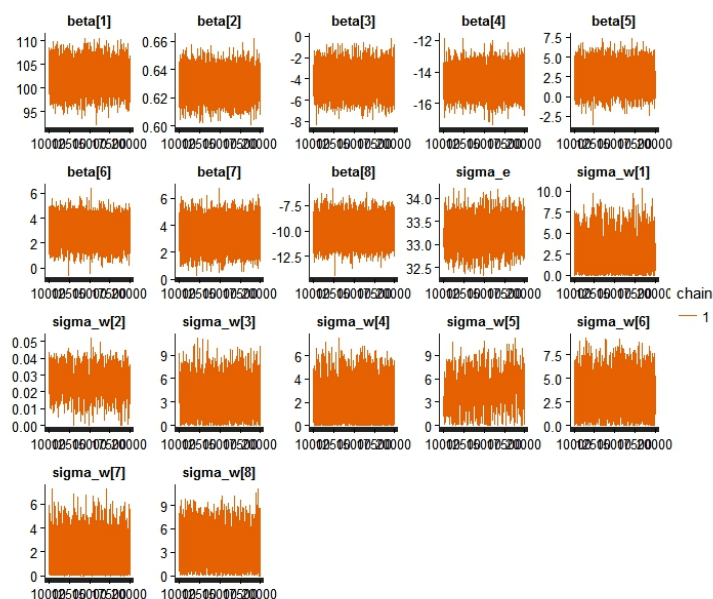


Figura 12 – Desenvolvimento do processo MCMC durante as 20.000 iterações para cada um dos parâmetros.

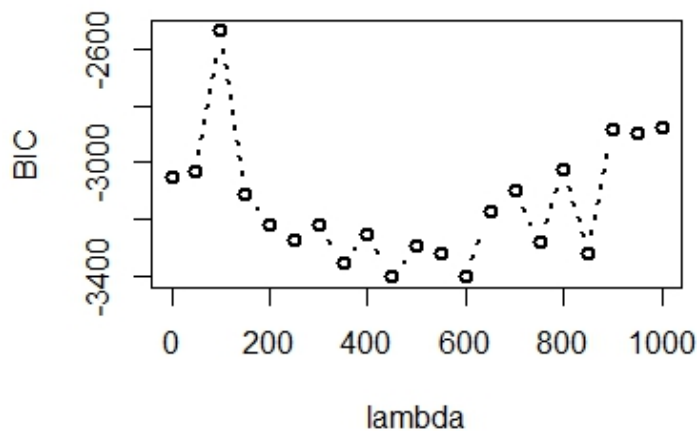


Figura 13 – Gráfico dos BIC versus λ , para 21 valores de λ tomados entre 0 e 1000.

com alta frequência por parte do aluno, tempo gasto com recreação em dia de aula, aluno gostar de matemática, o desempenho em matemática global da escola e o desempenho em língua portuguesa do aluno. Por outro lado podemos destacar como fatores que influenciam negativamente a proficiência em matemática são: o aluno declarar ser do sexo feminino, o aluno já ter sido reprovado e gostar de língua portuguesa.

Tabela 29 – Resultados referentes ao algoritmo glmLasso com as covariáveis descritas anteriormente, para $\lambda = 450$, foi selecionado no total 34 variáveis explicativas.

Variáveis Explicativas	Estimativa	Valor p
Efeito fixo		
Intercepto	-0,582	< 0,001
Q001 (1: Sexo Feminino)	-0,257	< 0,001
Q002: Amarelo	0,154	0,067
Q002: Branco	0,149	< 0,001
Q002: Indígena	0,019	0,856
Q002: Pardo	0,134	0,038
Q002: Preto	0,072	0,601
Q002: Não Sei	-	-
Q004 (1: aluno está na série ideal)	0,052	0,294
Q012 (1: tem carro em casa)	-0,000	0,973
Q013 (1: tem computador em casa)	0,054	0,080
Q019 (1: mãe com ensino superior ou mais)	0,036	0,192
Q020 (1: mãe alfabetizada)	0,143	0,022
Q022 (1: aluno mora com o pai)	-0,002	0,923
Q023 (1: pai com ensino superior ou mais)	0,042	0,141
Q024 (1: pai alfabetizado)	0,044	0,328
Q026 (1: alta frequência dos pais nas reuniões escolares)	0,029	0,255
Q030 (1: há incentivo dos pais para que os alunos não faltem)	0,154	0,152
Q033 (1: leitura de livros com alta frequência)	0,057	0,053
Q037 (1: alta frequência de leitura de revistas comportamentais)	-0,065	0,013
Q038 (1: leitura de sites de internet com alta frequência)	0,001	0,955
Q040 (1: ir ao cinema em alta frequência)	-0,070	0,021
Q041 (1: ver apresentações teatrais em alta frequência)	-0,058	0,354
Q042 (1: ver shows musicais ou de dança em alta frequência)	-0,041	0,343
Q043 (1: participar de festas comunitárias em alta frequência)	-0,049	0,105
Q044 (1: alunos gasta muito tempo com recreação em dia de aula)	0,072	0,012
Q046 (1: aluno trabalha fora de casa)	0,012	0,696
Q047 (1: ingressou na escola antes da 1ª série)	-0,029	0,391
Q049 (1: aluno já foi reprovado)	-0,151	< 0,001
Q050 (1: aluno já abandonou a escola)	0,010	0,882
Q051 (1: aluno gosta de estudar Língua Portuguesa)	-0,180	< 0,001
Q052 (1: aluno faz o dever de casa de Língua Portuguesa com alta frequência)	0,021	0,716
Q054 (1: aluno gosta de estudar Matemática)	0,308	< 0,001
Q055 (1: aluno faz o dever de casa de Matemática com alta frequência)	0,084	0,094
Q056 (1: professor corrige o dever de casa de Matemática com alta frequência)	0,034	0,610
MEDIA_MT	0,321	< 0,001
MEDIA_LP	-0,191	< 0,001
TAXA_PARTICIPACAO	-0,004	0,723
PROFICIENCIA_LP_SAEB	0,579	< 0,001
DEPENDENCIA_ADM (1: Estadual)	-0,030	0,272
Efeito aleatório		
Desvio Padrão(Intercepto da Escola)	0,041	

DISCUSSÃO E PROPOSTAS FUTURAS

O presente trabalho tem como objetivo principal construir modelos de regressão multinível tendo como variável resposta o desempenho escolar dos estudantes do nono ano do ensino fundamental do Estado de São Paulo na PROVA BRASIL 2011 em matemática. O uso de modelos multiníveis com dois níveis se justificam pelo fato de que há uma porção significativa da variabilidade dos dados que é explicada por escolas. Nosso intuito foi identificar quais variáveis explicativas afetam o desempenho dos estudantes em matemática. É importante destacar que ao apontarmos essas variáveis, damos um grande passo para compreender o que influencia no bom ou no mau desempenho do estudante e a partir daí, entender o funcionamento do sistema educacional paulista.

Nossa análise se dividiu em seleção de modelos e seleção de variáveis. Quando falamos em seleção de variáveis o método utilizado foi o Lasso, desenvolvido inicialmente para modelos de regressão linear simples e depois aperfeiçoado para ser utilizado em modelos lineares generalizados mistos (multiníveis com dois níveis), considerando resposta normal e função de ligação identidade. Constatamos uma diminuição de aproximadamente 45% dos preditores, contudo deve-se ressaltar que após a execução do algoritmo do gradiente ascendente, são obtidas as estimativas dos parâmetros utilizando um procedimento de escore de Fisher, e nesse caso verificamos que vários dos preditores selecionados via Lasso não são significativos se utilizarmos um teste t com 5% de significância. Assim podemos afirmar que os fatores que influenciam positivamente a proficiência em matemática dos alunos da rede pública paulista são: grau de instrução da mãe, frequência de leitura de livros por parte dos alunos, tempo gasto com recreação em dia de aula, o aluno declarou gostar de matemática, o desempenho em matemática global da escola e o desempenho em língua portuguesa do aluno. Por outro lado, podemos destacar como fatores que influenciam negativamente a proficiência em matemática: o aluno declara ser do sexo feminino, o aluno já ter sido reprovado e gostar de língua portuguesa.

Com relação à seleção de modelos, dividimos nossa análise entre enfoque clássico e

enfoque bayesiano. Na modelagem clássica utilizamos para selecionar modelos três critérios de informação: AIC, AICc e BIC. No fim os três critérios utilizados foram favoráveis à seleção do modelo de regressão multinível com intercepto aleatório para região, modelo esse em que as covariáveis turno, localização e idade (mede a distorção idade e série) são não significativas. Baseado em todas as configurações existentes relativas à extração ou não das covariáveis citadas acima, ajustamos mais sete modelos e aplicamos novamente os critérios de seleção e daí notamos que o BIC optou pelo modelo onde é feita a exclusão das três covariáveis, por outro lado o AIC e AICc selecionou o modelo onde apenas a variável explicativa turno é excluída. Como o modelo selecionado pelo AIC e AICc contem variáveis explicativas não significativas se utilizarmos um teste t com 5% de significância, o que não ocorre com o modelo selecionado pelo BIC, que possui todas os preditores significativos, foi realizado ainda um breve estudo de simulação para o caso frequentista, onde foi constatado que para grande amostras e muitos preditores o critério BIC é mais eficiente que o AIC e o AICc. Sendo assim, sugerimos a utilização do BIC como critério de seleção de modelos, onde a abordagem for de modelos de regressão multiníveis com resposta normal.

Na modelagem bayesiana utilizamos as mesmas covariáveis da parte clássica. Os critérios de informação que empregamos para comparação desses modelos foram: WAIC, *looic*, DIC, EAIC e EBIC. Primeiramente ajustamos três modelos bayesianos: um com efeitos fixos apenas, outro com intercepto aleatório e por último um com intercepto e inclinações aleatórias. Ressaltamos que os efeitos aleatórios são de escola.

Depois de ajustados esses três primeiros modelos constatamos que os cinco critérios são favoráveis à seleção do modelo com intercepto e inclinações aleatórias para escola. Analisando os intervalos de credibilidade desses modelos, concluímos que os intervalos de credibilidade para localização, turno e distorção idade-série contêm o zero, evidenciando assim que esses preditores podem ou não fazer parte do modelo. Logo se faz necessário ajustar mais sete modelos levando em consideração todas as configurações possíveis com essas três covariáveis. Feito isso e calculado os critérios de informação, concluímos que em três dos cinco critérios (WAIC, *Looic* e EBIC), o modelo sem os preditores localização e turno é o melhor, contudo o intervalo de credibilidade para o coeficiente da covariável distorção idade-série contêm o zero.

De uma forma geral, sugerimos a utilização do critério de informação WAIC, para seleção de modelos de regressão multinível, pois ao contrário das outras medidas de informação, é invariante a parametrização, insensível ao tamanho amostral e também computacionalmente eficiente, além do mais sua construção é feita totalmente numa abordagem bayesiana, pois se baseia na função $\log a posteriori$ preditiva.

Para trabalhos futuros podem-se testar outros preditores, sejam eles já existentes na amostra dos dados, ou até mesmo outro construído via técnicas de análise multivariada e a partir daí é possível realizar também:

-
- a extensão de suposições de distribuição de probabilidade para componentes aleatórios;
 - análise de diagnóstico;
 - estudos de simulação mais aprofundados.

Ressaltamos que a base de dados em questão nos fornece uma grande quantidade de análises possíveis a serem realizadas além daquelas que aqui apresentamos, inclusive outras questões práticas do contexto educacional em análise podem ser respondidas, uma das mais importantes que consideramos é encontrar quais são os preditores que são significativos para a covariável reprovação, bem como a analisar as covariáveis provenientes do questionário de professor, diretor e da escola são significativas para o desempenho do aluno em matemática. Devido à alta correlação entre a proficiência em matemática e proficiência em língua portuguesa é possível o ajuste de modelos de regressão multivariados.

No contexto prático abordado neste trabalho há uma imensa quantidade de preditores (mais de 500), por isso é esperado que métodos de seleção de modelos sejam mais eficazes neste caso, pois para estimar todos os modelos possíveis e calcular os seus respectivos critérios de informação seria muito trabalhoso e ainda demandaria muito tempo. Futuramente outros métodos de seleção de modelos podem ser testados, dentre os mais importantes podemos destacar: Regressão Ridge, Adaptive Lasso e Lasso Bayesiano.

Como já foi dito anteriormente, a educação é uma das demandas mais desafiadoras em nossa sociedade atual, por isso, pesquisas com dados educacionais são de vital importância. Tais pesquisas fornecem subsídios para diagnósticos e reformulações de políticas educacionais, fornecem ainda ferramentas para que o docente em sala de aula possa reformular sua prática pedagógica e assim contribuir mais incisivamente para a evolução do aluno no processo de ensino e aprendizagem.

REFERÊNCIAS

- AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: **Selected Papers of Hirotugu Akaike**. [S.l.]: Springer, 1998. p. 199–213. Citado nas páginas [27](#), [45](#) e [46](#).
- ANDRADE, R. J. de; SOARES, J. F. O efeito da escola básica brasileira. **Estudos em avaliação educacional**, v. 19, n. 41, p. 379–406, 2008. Citado na página [27](#).
- BATES, D.; MÄCHLER, M.; BOLKER, B.; WALKER, S. Fitting linear mixed-effects models using lme4. **Journal of Statistical Software**, v. 67, n. 1, p. 1–48, 2015. Citado nas páginas [64](#) e [165](#).
- BATES, D. M. **lme4: Mixed-effects modeling with R**. [S.l.]: Springer New York, 2010. Citado na página [41](#).
- BERGAMO, G. C. **Aplicação de modelos multiníveis na análise de dados de medidas repetidas no tempo**. Dissertação (Mestrado) — Universidade de São Paulo, 2002. Citado na página [27](#).
- BLIESE, P. **Multilevel Modeling in R (2.2)—A Brief Introduction to R, the multilevel package and the nlme package**. [S.l.]: October, 2006. Citado na página [41](#).
- BOLKER, B.; TEAM, R. bbmle: Tools for general maximum likelihood estimation. **R package version 0.9**, v. 5, 2010. Citado na página [165](#).
- BRESLOW, N. E.; CLAYTON, D. G. Approximate inference in generalized linear mixed models. **Journal of the American statistical Association**, Taylor & Francis Group, v. 88, n. 421, p. 9–25, 1993. Citado na página [54](#).
- BROOKS, S.; SMITH, J.; VEHTARI, A.; PLUMMER, M.; STONE, M.; ROBERT, C.; TITTERINGTON, D.; NELDER, J.; ATKINSON, A.; DAWID, A.; LAWSON, A.; CLARK, A.; BERNARDO, J.; SAHU, S.; RICHARDSON, S.; GREEN, P.; BURNHAM, K.; DEIORIO, M.; ROBERT, C.; DRAPER, D.; GELFAND, A.; TREVISANI, M.; HODGES, J.; LEE, Y.; De Luna, X.; MENG, X. Discussion on the paper by Spiegelhalter, Best, Carlin and Van der Linde. **Journal of the Royal Statistical Society. Series B: Statistical Methodology**, Wiley-Blackwell, v. 64, n. 4, p. 616–639, 2002. ISSN 1369-7412. Citado na página [51](#).
- CARLIM, B.; LOUIS, T. Bayes and empirical bayes methods for data analysis essays on item response theory. **Chapman and Hall**, New York, 2001. Citado na página [52](#).
- CASELLA, G.; BERGER, R. L. **Statistical inference**. [S.l.]: Duxbury Pacific Grove, CA, 2002. v. 2. Citado na página [41](#).
- CASTRO, C. S. **Aplicação de Modelos Multinível para o Estudo da Sinistralidade no Retalho Alimentar**. Dissertação (Mestrado) — Faculdade de Ciências da Universidade do Porto, 2015. Citado nas páginas [25](#) e [64](#).

- EMILIANO, P. C. **Fundamentos e aplicações dos critérios de Informação: akaike e bayesiano**. Dissertação (Mestrado), 2009. Citado nas páginas 45 e 78.
- EMILIANO, P. C. **Critérios de Informação: Como eles se comportam em diferentes modelos**. Tese (Doutorado) — Universidade Federal de Lavras, 2013. Citado na página 28.
- GELMAN, A.; HILL, J. **Data analysis using regression and multilevel/hierarchical models**. [S.l.]: Cambridge university press, 2006. Citado nas páginas 27, 42, 43, 44, 48, 64 e 84.
- GOLDSTEIN, H. Multilevel mixed linear model analysis using iterative generalized least squares. **Biometrika**, v. 73, n. 1, p. 43–56, 1986. Citado nas páginas 25 e 64.
- GONÇALVES, F. d. O.; FRANÇA, M. T. A. Qualidade educacional nos municípios nordestinos: Evidências a partir do prova brasil 2005. **Planejamento e Políticas Públicas**, n. 40, 2013. Citado nas páginas 31 e 32.
- GROLL, A.; TUTZ, G. Variable selection for generalized linear mixed models by l1-penalized estimation. **Statistics and Computing**, Springer, v. 24, n. 2, p. 137–154, 2014. Citado nas páginas 52 e 53.
- KUZNETSOVA, A.; BROCKHOFF, P. B.; CHRISTENSEN, R. H. B. Package ‘lmertest’. **R package version**, v. 2, 2015. Citado na página 165.
- LAROS, J. A.; MARCIANO, J. L. P.; ANDRADE, J. M. de. Fatores que afetam o desempenho na prova de matemática do saeb: Um estudo multinível. **Avaliação Psicológica: Interamerican Journal of Psychological Assessment**, Instituto Brasileiro de Avaliação Psicológica (IBAP), v. 9, n. 2, p. 173–186, 2010. Citado nas páginas 27, 31, 32 e 63.
- MICRODADOS. 2011. <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Accessed: 2016-03-15. Citado nas páginas 30 e 63.
- MISSÃO, É. C. M. **Uma revisão do fator de Bayes com aplicação à modelos com misturas**. Dissertação (Mestrado), 2010. Citado na página 47.
- MOREIRA, K. d. S. G.; JACINTO, P. de A. Determinantes da proficiência em matemática no rio grande do sul: Uma análise a partir de modelos hierárquicos. **PPGE/PUCRS**, 2013. Citado na página 27.
- NATIS, L. **Modelos lineares hierárquicos**. Dissertação (Mestrado) — Instituto de Matemática e Estatística, 2000. Citado na página 27.
- NATIS, L. Modelos hierárquicos lineares. **Estudos em avaliação educacional**, n. 23, p. 3–29, 2013. Citado nas páginas 27, 31, 32 e 64.
- NEAL, R. M. *et al.* Mcmc using hamiltonian dynamics. **Handbook of Markov Chain Monte Carlo**, v. 2, p. 113–162, 2011. Citado nas páginas 81, 167 e 168.
- OSIO, M. M. G. **Análise de modelos de regressão multiníveis simétricos**. Dissertação (Mestrado) — Instituto de Ciências Matemáticas e de Computação, 2013. Citado nas páginas 26, 27, 41 e 64.
- PATTERSON, H. D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, Biometrika Trust, v. 58, n. 3, p. 545–554, 1971. Citado na página 41.

PESQUISA. 2011. <<https://www.educacao.mg.gov.br/images/documentos/Pesquisa%20sobre%20a%20Qualidade%20do%20Ensino%20nas%20Escolas%20Estaduais%20de%20MG.pdf>>. Accessed: 2016-08-07. Citado na página 32.

PINHEIRO, J.; BATES, D.; DEBROY, S.; SARKAR, D.; R Core Team. **nlme: Linear and Nonlinear Mixed Effects Models**. [S.l.], 2017. R package version 3.1-131. Disponível em: <<https://CRAN.R-project.org/package=nlme>>. Citado na página 165.

RUSSO, C. M.; AOKI, R.; PAULA, G. A. Assessment of variance components in nonlinear mixed-effects elliptical models. **Test**, Springer, p. 1–27, 2012. Citado na página 42.

SANTOS, L. S. G. dos. Fatores associados ao sucesso escolar em matemática: um estudo comparativo entre duas escolas públicas do DF a partir dos dados da prova brasil. 2013. Citado na página 27.

SCHWARZ, G. *et al.* Estimating the dimension of a model. **The annals of statistics**, Institute of Mathematical Statistics, v. 6, n. 2, p. 461–464, 1978. Citado nas páginas 45 e 47.

SEARLE, S. R. *et al.* Biometrics unit technical reports: Number bu-1256-m: On mixed models, reml and blup. 1994. Citado na página 27.

SILVA, N. N. d. **Análise probabilística: um curso introdutório**. 18. ed. [S.l.]: Edusp, 1998. Citado na página 32.

SORENSEN, T.; VASISHTH, S. Bayesian linear mixed models using stan: a tutorial for psychologists, linguists, and cognitive scientists. **arXiv preprint arXiv:1506.06201**, 2015. Citado na página 81.

SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. V. D. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 64, n. 4, p. 583–639, 2002. Citado nas páginas 47, 48 e 49.

SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. The deviance information criterion: 12 years on. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 76, n. 3, p. 485–493, 2014. Citado nas páginas 47 e 84.

SUGIURA, N. Further analysts of the data by akaike's information criterion and the finite corrections: Further analysts of the data by akaike's. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 7, n. 1, p. 13–26, 1978. Citado nas páginas 45 e 47.

SULLIVAN, L. M.; DUKES, K. A.; LOSINA, E. Tutorial in biostatistics. an introduction to hierarchical linear modelling. **Stat Med**, v. 18, n. 7, p. 855–888, 1999. Citado nas páginas 41 e 42.

TIBSHIRANI, R. Regression shrinkage and selection via the lasso. **Journal of the Royal Statistical Society. Series B (Methodological)**, te, p. 267–288, 1996. Citado nas páginas 27, 28, 52 e 64.

VEHTARI, A.; GELMAN, A.; GABRY, J. Practical bayesian model evaluation using leave-one-out cross-validation and waic. **Statistics and Computing**, Springer, p. 1–20, 2016. Citado nas páginas 49, 50 e 51.

WATANABE, S. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. **Journal of Machine Learning Research**, v. 11, n. Dec, p. 3571–3594, 2010. Citado nas páginas 48 e 49.

WEAKLIEM, D. L. A critique of the bayesian information criterion for model selection. **Sociological Methods & Research**, Sage Publications, v. 27, n. 3, p. 359–397, 1999. Citado nas páginas 27 e 64.

BANCO DE DADOS

Neste apêndice apresentamos a sumarização dos dados em sua forma original, como pode ser visto na Tabela 30. Nesta tabela temos o código original das covariáveis, o código em que cada covariável foi realocada, nome da variável, enunciado referente, tipo e o valor-p associado ao teste t de significância individual. Após o procedimento de amostragem descrito no capítulo 2, algumas covariáveis foram dicotomizadas, na Tabela 31, vemos como foi codificado as variáveis explicativas com mais de uma categoria e a partir daí foram dicotomizadas. Na Tabela 32 destacamos as variáveis que foram dicotomizadas, sua categorias, o código que essas categorias assumiram e sua frequência amostral.

Tabela 30 – Descrição do banco de dados geral.

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
MEDIA_LP	MEDIA_LP	Nota média dos participantes em Língua Portuguesa na escola		Contínua	$< 2.10^{-16}$
MEDIA_MT	MEDIA_MT	Nota média dos participantes em matemática na escola		Contínua	$< 2.10^{-16}$
TAXA_PARTICIPACAO	TAXA_PARTICIPACAO	Taxa de participação dos estudantes da escola		Contínua	$< 2.10^{-16}$
NU_PRESENTES	NU_PRESENTES	Número de presentes		Contínua	$< 2.10^{-16}$
NU_MATRICULADOS_CENSO	NU_MATRICULADOS_CENSO	Número de matriculados na escola		Contínua	$< 2.10^{-16}$
PROFICIENCIA_LP_SAEB	PROFICIENCIA_LP_SAEB	Proficiência em língua portuguesa no SAEB		Contínua	$< 2.10^{-16}$
ID_DEPENDENCIA_ADM	DEPENDENCIA_ADM	Dependência administrativa da escola		Catégorica nominal	$< 2.10^{-16}$
ID_LOCALIZACAO	LOCALIZACAO	Localização da escola		Catégorica nominal	0,205

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q032	Q032	Jornais	Você lê: Jornais	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q033	Q033	Livros em geral	Você lê: Livros em geral	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q034	Q034	Livros de literatura infanto-juvenil	Você lê: Livros de literatura infanto-juvenil	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q035	Q035	Revistas em geral	Você lê: Revistas em geral quadrinhos	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q036	Q036	Revistas em quadrinhos	Você lê: Revistas em quadrinhos	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q037	Q037	Revistas sobre comportamento, celebridades, esporte e TV	Você lê: Revistas sobre comportamento, celebridades, esporte e TV	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q038	Q038	Internet	Você lê: Sites da internet	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q039	Q039	Biblioteca	Você Costuma: Frequentar bibliotecas	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q040	Q040	Cinema	Você Costuma: Ir ao cinema.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q041	Q041	Apresentações Teatrais	Você Costuma: Ver apresentações teatrais	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q042	Q042	Apresentações Teatrais	Você Costuma: Ver apresentações musicais ou de dança	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q043	Q043	Festas abertas à comunidade	Você Costuma: Participar de festas abertas à comunidade	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q044	Q044	Lazer	Em dia de aula, quanto tempo você gasta assistindo à TV, navegando na internet ou jogando jogos eletrônicos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q045	Q045	Trabalhos domésticos	Em dias de aula, quanto tempo você gasta fazendo trabalhos domésticos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q046	Q046	Trabalho fora de casa	Você trabalha fora de casa?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q047	Q047	Entrada na escola	Quando você entrou na escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q048	Q048	Tipo de escola desde a quinta série	Desde a quinta série em que tipo de escola você estudou?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q049	Q049	Reprovas	Você já foi reprovado?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q050	Q050	Abandono Escolar	Você já abandonou a escola durante o período de aulas e ficou fora da escola o resto do ano?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q051	Q051	Língua Portuguesa	Você gosta de estudar língua portuguesa?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q052	Q052	Dever de casa de L.P.	Você faz o dever de casa de língua portuguesa?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q053	Q053	Correção do dever de casa de L.P.	O professor corrige o dever de casa de língua portuguesa?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q054	Q054	Matemática	Você gosta de estudar matemática?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q055	Q055	Dever de casa de matemática	Você faz o dever de casa de matemática?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q056	Q056	Correção do dever de casa de matemática	O professor corrige o dever de casa de matemática?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q057	Q057	Sala de Leitura	Você utiliza a biblioteca ou sala de leitura da sua escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q058	Q058	Pretensão ao terminar 9º ano	Quando você terminar o 9º ano (8ª série), você pretende:	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q001	E001	Telhado	Avalie o estado de conservação dos itens e equipamentos do prédio - telhado	categórica nominal	$7,08.10^{-14}$
TX_RESP_Q002	E002	Paredes	Avalie o estado de conservação dos itens e equipamentos do prédio-paredes	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q003	E003	Piso	Avalie o estado de conservação dos itens e equipamentos do prédio-piso	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q004	E004	Entrada do prédio	Avalie o estado de conservação dos itens e equipamentos do prédio-entrada do prédio	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q005	E005	Pátio	Avalie o estado de conservação dos itens e equipamentos do prédio-pátio	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q006	E006	Corredores	Avalie o estado de conservação dos itens e equipamentos do prédio-corredores	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q007	E007	Salas de aula	Avalie o estado de conservação dos itens e equipamentos do prédio-salas de aula	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q008	E008	Portas	Avalie o estado de conservação dos itens e equipamentos do prédio-portas	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q009	E009	Janelas	Avalie o estado de conservação dos itens e equipamentos do prédio-janelas	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q010	E010	Banheiros	Avalie o estado de conservação dos itens e equipamentos do prédio-banheiros	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q011	E011	Cozinha	Avalie o estado de conservação dos itens e equipamentos do prédio-cozinha	categórica nominal	0,000519
TX_RESP_Q012	E012	Instalações hidráulicas	Avalie o estado de conservação dos itens e equipamentos do prédio - instalações hidráulicas	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q013	E013	Instalações elétricas	Avalie o estado de conservação dos itens e equipamentos do prédio- instalações elétricas	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q014	E014	Iluminação de sala de aula	Avalie os seguintes aspectos em relação às salas de aula: (considere sim quando mais de 50% das salas de aula apresentarem o aspecto questionado) - são iluminadas?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q015	E015	Ventilação de sala de aula	Avalie os seguintes aspectos em relação às salas de aula: (considere sim quando mais de 50% das salas de aula apresentarem o aspecto questionado) - são arejadas?	categórica nominal	$7,73.10^{-11}$
TX_RESP_Q016	E016	Ventilação de sala de aula	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há muros, grades ou cercas em condições de garantir a segurança dos alunos? (caso existam buracos ou aberturas que permitam o acesso de estranhos, responder não)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q017	E017	Fluxo de alunos	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há controle de entrada e saída de alunos?	categórica nominal	$5,36.10^{-12}$
TX_RESP_Q018	E018	Fluxo de pessoas estranhas	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há controle de entrada de pessoas estranhas na escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q019	E019	Portões de acesso à parte externa da escola	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: os portões que dão acesso à parte externa permanecem trancados durante o horário de funcionamento da escola?	categórica nominal	0,000703
TX_RESP_Q020	E020	Segurança no período diurno	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum tipo de vigilância para o período diurno?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q021	E021	Segurança no período noturno	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum tipo de vigilância para o período noturno?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q022	E022	Segurança nos fins de semana	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum tipo de vigilância para os finais de semana e feriados?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q023	E023	Policimento anti - furtos	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum esquema de policimento para inibição de furtos, roubos e outras formas de violência?	categórica nominal	0,0418
TX_RESP_Q024	E024	Policimento anti - drogas dentro da escola	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum esquema de policimento para inibição de tráfico de tóxicos/ drogas dentro da escola?	categórica nominal	$4,37.10^{-6}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q025	E025	Policiamento anti - drogas nas imediações escola	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum esquema de policiamento para inibição de tráfico de tóxicos/ drogas nas imediações da escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q026	E026	Proteção contra incêndio	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: a escola possui algum sistema de proteção contra incêndio (alarme de fumaça e temperatura, extintores contra incêndio dentro do prazo de validade, mangueira, etc.)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q027	E027	Proteção para equipamentos de alto custo	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: as salas onde são guardados os equipamentos mais caros (computadores, projetores, televisão, vídeo etc.) Possuem dispositivos para serem trancados (cadeados, grades, travas, trancas, etc.)?	categórica nominal	0,561
TX_RESP_Q028	E028	Depredação	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: a escola apresenta sinais de depredação (vidros, portas e janelas quebradas, lâmpadas estouradas etc.)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q029	E029	Iluminação externa	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há uma boa iluminação do lado de fora da escola?	categórica	$< 2.10^{-16}$
TX_RESP_Q030	E030	Segurança nas imediações	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: a escola adota alguma medida de segurança para proteger os alunos nas suas imediações?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q031	E031	Computadores escolares	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Computadores para uso dos alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q032	E032	Internet da escola	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Acesso à internet para uso dos alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q033	E033	Computadores para professores	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Computadores para uso dos professores.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q034	E034	Internet para professores	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Acesso à internet para uso dos professores.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q035	E035	Computadores administrativos	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Computadores exclusivamente para o uso administrativo.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q036	E036	Mídias educativas	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Fitas de vídeo ou DVD (educativas).	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q037	E037	Mídias de lazer	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Fitas de vídeo ou DVD (lazer).	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q038	E038	Máquina copiadora	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Máquina copiadora.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q039	E039	Impressora	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Impressora.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q040	E040	Retroprojeter	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Retroprojeter.	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q041	E041	Projektor de slides	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Projektor de slides.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q042	E042	Videocassete ou DVD escolar	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Videocassete ou DVD.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q043	E043	Televisão	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Televisão.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q044	E044	Mimeógrafo	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Mimeógrafo.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q045	E045	Câmera fotográfica	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Câmera fotográfica.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q046	E046	Antena parabólica	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Antena parabólica.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q047	E047	Linha telefônica	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Linha telefônica.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q048	E048	Fax	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Aparelho de fax.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q049	E049	Aparelho de som	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Aparelho de som.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q050	E050	Biblioteca	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Biblioteca.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q051	E051	Quadra de esportes	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Quadra de esportes.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q052	E052	Laboratório	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Laboratório.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q053	E053	Auditório	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Auditório.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q054	E054	Sala de música	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Sala para atividades de música.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q055	E055	Sala de artes plásticas	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Sala para atividades de artes plásticas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q056	E056	Sala de leitura	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Sala de leitura.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q057	E057	Acervo da sala de leitura	Em relação à biblioteca ou sala de leitura: possui acervo diversificado que desperte o interesse dos alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q058	E058	Brinquedoteca da sala de leitura	Em relação à biblioteca ou sala de leitura: possui brinquedoteca.	categórica nominal	$3,2.10^{-8}$
TX_RESP_Q059	E059	Sala de estudos coletivos	Em relação à biblioteca ou sala de leitura: possui espaço para estudos coletivos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q060	E060	Empréstimo de livros	Em relação à biblioteca ou sala de leitura: os livros podem ser manuseados e emprestados.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q061	E061	Comunidade na sala de leitura	Em relação à biblioteca ou sala de leitura: a comunidade pode utilizar o espaço e os livros.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q062	E062	Climatização da sala de leitura	Em relação à biblioteca ou sala de leitura: o espaço é arejado e bem iluminado.	categórica nominal	0,00363
TX_RESP_Q063	E063	Responsável pela sala de leitura	Em relação à biblioteca ou sala de leitura: existe uma pessoa responsável pelo atendimento na biblioteca ou na sala de leitura?	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q064	E064	Retirada de livros por alunos	Os usuários da biblioteca (ou sala de leitura) levam livros para casa? Os alunos.	categórica	$< 2.10^{-16}$
TX_RESP_Q065	E065	Retirada de livros por professores	Os usuários da biblioteca (ou sala de leitura) levam livros para casa? Os professores.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q066	E066	Retirada de livros pela comunidade	Os usuários da biblioteca (ou sala de leitura) levam livros para casa? Os membros da comunidade.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q001	D001	Sexo do diretor	Sexo	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q002	D002	Idade do diretor	Idade	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q003	D003	Raça do diretor	Como você se considera?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q004	D004	Escolaridade do diretor	Qual o seu nível de escolaridade (até a graduação).	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q005	D005	Tempo de escolaridade	Há quantos anos você obteve o nível de escolaridade assinalado anteriormente?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q006	D006	Tipo de instituição onde obteve seu título profissional	Em que tipo de instituição você fez o curso superior? Se você estudou em mais de uma instituição, assinale aquela em que obteve seu título profissional.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q007	D007	Natureza da instituição onde obteve seu título profissional	Qual era a natureza dessa instituição?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q008	D008	Forma de realização do curso superior	De que forma você realizou o curso superior?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q009	D009	Modalidade da pós - graduação	Indique a modalidade de cursos de pós-graduação de mais alta titulação que você possui.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q010	D010	Temática da pós - graduação	Indique qual a área temática do curso de pós-graduação de mais alta titulação que você possui.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q011	D011	Formação Continuada	Você participou de alguma atividade de formação continuada (atualização, treinamento, capacitação, etc.) nos últimos dois anos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q012	D012	Carga horária da atividade mais relevante	Qual a carga horária da atividade mais relevante da qual você participou?	categórica nominal	$1,71.10^{-11}$
TX_RESP_Q013	D013	Aplicação da formação continuada	Você utiliza os conhecimentos adquiridos nas atividades de formação continuada de que você participou?	categórica nominal	$1,08.10^{-15}$
TX_RESP_Q014	D014	Salário bruto do diretor	Qual é aproximadamente o seu salário bruto como diretor(a), com adicionais, se houver? (preencha os espaços com o valor e, abaixo de cada algarismo, marque o campo correspondente.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q015	D015	Renda familiar bruta do diretor	Qual é aproximadamente sua renda familiar bruta? (preencha os espaços com o valor e, abaixo de cada algarismo, marque o campo correspondente.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q016	D016	Atividade extraescolar	Além da direção desta escola, você exerce outra atividade que contribui para sua renda pessoal?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q017	D017	Tempo de trabalho na educação	Há quantos anos você trabalha em educação?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q018	D018	Tempo de como direção	Há quantos anos você exerce funções de direção?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q019	D019	Tempo como direção desta escola	Há quantos anos você é diretor(a) desta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q020	D020	Carga horária nesta escola	Qual é a sua carga horária de trabalho nesta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q021	D021	Forma que assumiu a direção	Você assumiu a direção desta escola por:	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q022	D022	Promoção de formação continuada	Você promoveu alguma atividade de formação continuada (atualização, treinamento, capacitação, etc.) Nesta escola?	categórica nominal	$4,55.10^{-10}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q023	D023	Proporção de professores que participaram das atividades de formação	Qual foi à proporção de docentes da sua escola que participou das atividades de formação continuada promovidas por você nos últimos dois anos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q024	D024	Números de reunião de conselho de escola	Conselho de escola é um colegiado constituído por representantes da escola e da comunidade que tem como objetivo acompanhar as atividades escolares. Neste ano, quantas vezes o conselho desta escola se reuniu?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q025	D025	Professores no conselho de escola	O conselho de escola é composto por professores. (marque sim ou não em cada linha)	categórica nominal	$3,25.10^{-9}$
TX_RESP_Q026	D026	Alunos no conselho de escola	O conselho de escola é composto por alunos. (marque sim ou não em cada linha)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q027	D027	Funcionários no conselho de escola	O conselho de escola é composto por funcionários. (marque sim ou não em cada linha)	categórica nominal	$7,72.10^{-7}$
TX_RESP_Q028	D028	Pais no conselho de escola	O conselho de escola é composto por pais. (marque sim ou não em cada linha)	categórica nominal	$2,04.10^{-7}$
TX_RESP_Q029	D029	Conselho de classe	Conselho de classe é um órgão formado por todos os professores que lecionam em cada turma/série. Neste ano, quantas vezes se reuniram os conselhos de classe desta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q030	D030	Projeto Pedagógico	Quanto ao projeto pedagógico desta escola neste ano (marque apenas uma opção)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q031	D031	Critério Admissão Escolar	Qual é o critério para a admissão de alunos nesta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q032	D032	Oferta de vagas escolar	Neste ano letivo, como foi a situação da oferta de vagas nesta escola? (marque apenas uma alternativa.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q033	D033	Critério de composição de turmas	Qual o critério utilizado para formação das turmas nesta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q034	D034	Atribuição de turmas ensino fundamental I	Neste ano, qual foi o critério mais importante para a atribuição das turmas de 1. ^a a 4. ^a séries do ensino fundamental aos professores?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q035	D035	Professores estáveis	Qual é o percentual de professores com vínculo estável nesta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q036	D036	Taxas de abandono	Nesta escola, há algum programa de redução das taxas de abandono?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q037	D037	Taxas de reprova	Nesta escola, há algum programa de redução das taxas de reprovação?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q038	D038	Política anti - faltas dos alunos feita pelos professores	Para evitar que os alunos faltem às aulas, os professores falam com os alunos.	categórica nominal	$5,66.10^{-5}$
TX_RESP_Q039	D039	Comunicação com os pais por causa das faltas dos alunos	Para evitar que os alunos faltem às aulas, os pais/responsáveis são avisados por comunicação escrita.	categórica nominal	$4,32.10^{-6}$
TX_RESP_Q040	D040	Conversa com os pais por causa das faltas dos alunos em reunião	Para evitar que os alunos faltem às aulas, os pais/responsáveis são chamados à escola para conversar sobre o assunto em reunião de pais.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q041	D041	Conversa com os pais por causa das faltas dos alunos individualmente	Para evitar que os alunos faltem às aulas, os pais/responsáveis são chamados à escola para conversar sobre o assunto individualmente.	categórica nominal	$3,33.10^{-10}$
TX_RESP_Q042	D042	Visita aos alunos	Para evitar que os alunos faltem às aulas, a escola envia alguém à casa do aluno.	categórica nominal	$2,86.10^{-13}$
TX_RESP_Q043	D043	Reforço escolar	Esta escola desenvolve, regularmente, algum programa de apoio ou reforço de aprendizagem para os alunos (monitoria, aula de reforço etc.)	categórica nominal	$8,32.10^{-14}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q044	D044	Cultura afro - brasileira	Você tem conhecimento do conteúdo da lei 11.645 de 2008 que determina a obrigatoriedade do estudo da temática "história e cultura afro-brasileira e indígena" nos estabelecimentos de ensino do país?	categórica nominal	$1,87.10^{-9}$
TX_RESP_Q045	D045	Atividades de cultura afro - brasileira	Neste ano, foram desenvolvidas atividades para atender o determinado pela lei 11.645 de 2008 nesta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q046	D046	Recursos financeiros para merenda escolar	Em relação à merenda escolar, como você considera recursos financeiros.	categórica nominal	$3,34.10^{-16}$
TX_RESP_Q047	D047	Quantidade de merenda escolar	Em relação à merenda escolar, como você considera quantidade de alimentos.	categórica nominal	$1,56.10^{-14}$
TX_RESP_Q048	D048	Qualidade da merenda escolar	Em relação à merenda escolar, como você considera qualidade dos alimentos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q049	D049	Variedade da merenda escolar	Em relação à merenda escolar, como você considera variedade do cardápio.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q050	D050	Conservação da merenda escolar	Em relação à merenda escolar, como você considera armazenamento e conservação.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q051	D051	Cozinha escolar	Em relação à merenda escolar, como você considera espaço físico para cozinhar.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q052	D052	Higiene da merenda escolar	Em relação à merenda escolar, como você considera higiene.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q053	D053	Merendeira	Em relação à merenda escolar, como você considera disponibilidade de pessoal.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q054	D054	Qualificação: Merendeira	Em relação à merenda escolar, como você considera qualificação do pessoal.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q055	D055	Recursos financeiros	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola insuficiência de recursos financeiros?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q056	D056	Falta de professores	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola inexistência de professores para algumas disciplinas ou séries?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q057	D057	Falta de pessoal administrativo	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola carência de pessoal administrativo?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q058	D058	Falta de pessoal pedagógico	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola carência de pessoal de apoio pedagógico (coordenador, supervisor, orientador educacional)?	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q059	D059	Falta de recursos pedagógicos	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola falta de recursos pedagógicos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q060	D060	Interrupção das atividades escolares	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola interrupção das atividades escolares?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q061	D061	Índice de faltas por professores	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola alto índice de faltas por parte de professores?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q062	D062	Índice de faltas pelos alunos	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola alto índice de faltas por parte de alunos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q063	D063	Rotatividade de corpo docente	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola rotatividade do corpo docente?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q064	D064	Indisciplina dos alunos	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola problemas disciplinares causados por alunos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q065	D065	Interferências externas na gestão escolar	Considere as condições existentes para o exercício do cargo de diretor nesta escola. Há interferências externas em sua gestão?	categórica nominal	$1,56.10^{-14}$
TX_RESP_Q066	D066	Apoio na gestão escolar	Considere as condições existentes para o exercício do cargo de diretor nesta escola. Há apoio de instâncias superiores?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q067	D067	Troca de informações entre professores	Considere as condições existentes para o exercício do cargo de diretor nesta escola. Há troca de informações com diretores de outras escolas?	categórica nominal	$1,35.10^{-12}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q068	D068	Apoio da comunidade na gestão escolar	Considere as condições existentes para o exercício do cargo de diretor nesta escola. Há apoio da comunidade à sua gestão?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q069	D069	Pavimentação asfáltica próximo à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola, as calçadas possuem pavimento regular, plano, sem buracos ou degraus?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q070	D070	Largura da calçada próximo à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola, as calçadas possuem largura mínima de 1,20 m?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q071	D071	Obstáculos na calçada próximo à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola, as calçadas garantem uma faixa de circulação livre de obstáculos com largura mínima de 80 cm? (obstáculos: poste, árvore, floreira, orelhão, lixeira, placa de sinalização etc.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q072	D072	Semáforos próximo à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: há semáforos para pedestres?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q073	D073	Semáforos sonoros próximo à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: o semáforo é sonoro?	categórica nominal	3.10^{-12}
TX_RESP_Q074	D074	Faixa de pedestres próximo à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: há faixa de travessia (faixa de pedestres ou faixa de segurança)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q075	D075	Rampas de Travessia próximo à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: as calçadas são rebaixadas com rampas dos dois lados da rua no trecho da travessia?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q076	D076	Vagas em estacionamento para deficientes próximo à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: há vagas sinalizadas para veículos para pessoas com deficiência?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q077	D077	Acesso à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: o portão de acesso de pedestres à escola é separado da entrada de carros?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q078	D078	Acesso à escola	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: o portão de acesso tem largura mínima de 80 cm?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q079	D079	Entrada principal	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: os corredores e as calçadas tem largura mínima de 1,20 m?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q080	D080	Faixa de circulação interna	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: há faixa de circulação livre de obstáculos com largura mínima de 80 cm? (obstáculos: extintor de incêndio, vasos, lixeiras, móveis, bebedouros etc.)	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q081	D081	Piso interno	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: o piso é plano, sem degraus e sem desníveis superiores a 1,5 cm?	categórica nominal	$8,71.10^{-16}$
TX_RESP_Q082	D082	Rampas interna	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: há rampas?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q083	D083	Rampas interna	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: as rampas têm largura mínima de 80 cm?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q084	D084	Rampas internas com corrimão	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: há rampas com corrimão dos dois lados?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q085	D085	Escadas internas com corrimão	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: as escadas têm corrimãos dos dois lados?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q086	D086	Escadas internas com corrimão	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: elevadores ou plataformas elevatórias?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q087	D087	Soleiras das portas	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas salas de aula, na administração, na biblioteca e outros ambientes: nas portas, há soleiras niveladas com o piso do corredor, pequenas rampas ou degraus de no máximo 1,5 cm?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q088	D088	largura das portas	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas salas de aula, na administração, na biblioteca e outros ambientes: as portas de entrada possuem largura mínima de 80 cm?	categórica nominal	$4,35.10^{-8}$
TX_RESP_Q089	D089	Maçanetas das portas	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas salas de aula, na administração, na biblioteca e outros ambientes: as maçanetas são de alavanca?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q090	D090	Sinalização em braile	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas salas de aula, na administração, na biblioteca e outros ambientes: há algum tipo de sinalização em braile nos ambientes?	categórica nominal	$3,62.10^{-9}$
TX_RESP_Q091	D091	Banheiro para deficientes	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há pelo menos um sanitário adequado a pessoas com deficiência?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q092	D092	Banheiro para pessoas deficientes para cada sexo	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há pelo menos um sanitário adequado a pessoas com deficiência para cada sexo?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q093	D093	Largura das portas dos banheiros	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: as portas possuem largura mínima de 80 cm?	categórica nominal	$2,77.10^{-7}$
TX_RESP_Q094	D094	Direção de abertura das portas dos banheiros	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há banheiros ou mictórios com portas que abrem para fora?	categórica nominal	$3,92.10^{-9}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q095	D095	Maçaneta das portas dos banheiros	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: as maçanetas são de alavanca?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q096	D096	Cadeira de rodas nos banheiros	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há espaço para cadeiras de rodas e para transferência de pelo menos 80 cm?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q097	D097	Barras de apoio nos banheiros	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há barras de apoio fixadas nas paredes de fundo e na lateral do vaso sanitário?	categórica nominal	$4,63.10^{-14}$
TX_RESP_Q098	D098	Torneira da pia dos banheiros	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há pia com torneira com acionamento por alavanca, pressão ou cruzeta (em cruz)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q099	D099	Pia dos banheiros	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há pia com altura entre 78 e 80 cm e com vão inferior livre com altura de 73 cm?	categórica nominal	$2,12.10^{-8}$
TX_RESP_Q100	D100	Atendimento Educacional	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: sala de recursos multifuncionais para atendimento educacional especializado (aee)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q101	D101	Materiais didáticos	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: materiais didáticos e paradidáticos acessíveis? (braille, caracteres ampliados, libras, texturas, contrastes etc).	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q102	D102	Recursos ópticos	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: recursos ópticos? (ex: lupa, telulupa, telescópio, lentes especiais, etc.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q103	D103	Recursos não ópticos	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: recursos não ópticos? (ex: celofane amarelo, livro falado, material impresso com tipos maiores, tábua de apoio para leitura, lápis e canetas especiais, guia de leitura etc.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q104	D104	Soroban	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: soroban?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q105	D105	Reglete e punção	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: reglete e punção?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q106	D106	Calculadora sonora	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: calculadora sonora?	categórica nominal	$5,03.10^{-12}$
TX_RESP_Q107	D107	Calculadora para pessoas com baixa visão	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: calculadora para pessoas de baixa visão?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q108	D108	Máquina perkins	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: máquina perkins? (máquina de escrever em braille)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q109	D109	Software leitor de tela	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: computador com software leitor de tela?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q110	D110	Software de comunicação alternativa	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: software de comunicação alternativa aumentativa (caa)?	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q111	D111	Mesas e cadeiras para cadeirantes	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: mesas e cadeiras acessíveis para pessoas em cadeiras de rodas?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q112	D112	Quadro negro que permita o alcance de cadeirantes	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: quadro negro (ou branco) sem degraus com altura que permita seu alcance por pessoa em cadeira de rodas?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q113	D113	Bebedouros para deficientes	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: bebedouros que permitem a aproximação frontal de pessoas em cadeiras de rodas?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q114	D114	Atividades esportivas	Que atividades extracurriculares são desenvolvidas regularmente com os alunos nesta escola? Esportivas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q115	D115	Atividades artísticas	Que atividades extracurriculares são desenvolvidas regularmente com os alunos nesta escola? Artísticas (música, teatro, trabalhos artesanais).	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q116	D116	Atividades comunitárias	Os espaços desta escola são utilizados para atividades comunitárias?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q117	D117	Atividades comunitárias com recursos da escola	Neste ano, ocorreram nesta escola eventos da comunidade usando instalações, equipamentos ou recursos da escola?	categórica nominal	$9,62.10^{-6}$
TX_RESP_Q118	D118	Atividades comunitárias realizados por terceiros	Neste ano, ocorreram nesta escola eventos de terceiros realizados na escola e abertos para a comunidade (shows, teatro, palestras)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q119	D119	Eventos á comunidade externa	Neste ano, ocorreram nesta escola eventos da escola e destinados à comunidade externa (cursos, práticas esportivas, palestras)?	categórica nominal	0,078
TX_RESP_Q120	D120	Campanhas de solidariedade	Neste ano, ocorreram nesta escola campanhas de solidariedade promovidas pela escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q121	D121	Campanhas de solidariedade da comunidade	Neste ano, ocorreram nesta escola campanhas de solidariedade propostas pela comunidade, envolvendo a escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q122	D122	Hortas, pomar, jardins	Neste ano, ocorreram nesta escola comunidade colaborando na manutenção de hortas, pomar, jardins?	categórica nominal	$3,02.10^{-6}$
TX_RESP_Q123	D123	Mutirão de limpeza	Neste ano, ocorreram nesta escola comunidade participando em mutirão para limpeza da escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q124	D124	Mutirão de manutenção	Neste ano, ocorreram nesta escola comunidade participando em mutirão para manutenção da estrutura física da escola?	categórica nominal	$9,66.10^{-9}$
TX_RESP_Q125	D125	Financiamento Federal	Esta escola recebe apoio financeiro de programa de financiamento do governo federal?	categórica nominal	$2,26.10^{-12}$
TX_RESP_Q126	D126	Financiamento estadual	Esta escola recebe apoio financeiro de programa de financiamento do governo estadual?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q127	D127	Financiamento municipal	Esta escola recebe apoio financeiro de programa de financiamento do governo municipal?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q128	D128	Livros didáticos	Quem escolheu os livros didáticos utilizados nesta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q129	D129	Época que os livros didáticos chegaram	Neste ano, ocorreram as seguintes situações: os livros chegaram a tempo hábil para o início das aulas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q130	D130	Faltaram livros didáticos	Neste ano, ocorreram as seguintes situações: faltaram livros para os alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q131	D131	Sobraram livros didáticos	Neste ano, ocorreram as seguintes situações: sobraram livros.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q132	D132	Recebimento dos livros didáticos	Neste ano, ocorreram as seguintes situações: os livros escolhidos foram os recebidos.	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q133	D133	Atentado á vida de professores(estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: atentado à vida de professores ou funcionários dentro da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q134	D134	Atentado á vida de alunos(estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: atentado à vida de alunos dentro da escola. agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q135	D135	Furto aos professores ou funcionários (estranho a escola)	Neste ano, aconteceram os seguintes fatos nesta escola: furto a professores ou funcionários dentro da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q136	D136	Furto aos alunos (estranho a escola)	Neste ano, aconteceram os seguintes fatos nesta escola: furto a alunos dentro da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q137	D137	Roubo(com uso de violência)aos professores	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) a professores ou funcionários dentro da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q138	D138	Roubo(com uso de violência)aos alunos	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) a alunos dentro da escola. Agente causador externo (estranho à escola)	categórica nominal	$6,39.10^{-8}$
TX_RESP_Q139	D139	Furtos de equipamentos	Neste ano, aconteceram os seguintes fatos nesta escola: furto de equipamentos e materiais didáticos ou pedagógicos da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q140	D140	Roubos de equipamentos	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) de equipamentos e materiais didáticos ou pedagógicos da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q141	D141	Quebra intencional de equipamentos (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: quebra intencional de equipamento. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q142	D142	Pichações externas (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: pichação de muros ou paredes das dependências externas da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q143	D143	Depredações externas (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: depredação das dependências externas da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q144	D144	Sujeira nas dependências externas (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: sujeira nas dependências externas da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q145	D145	Sujeira nas dependências internas (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: sujeira nas dependências internas da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q146	D146	Pichações nas dependências internas (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: pichação de muros ou paredes das dependências internas da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q147	D147	Depredação nas dependências internas (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: depredação das dependências internas da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q148	D148	Depredação de banheiros (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: depredação de banheiros. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q149	D149	Bebidas alcoólicas (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de bebidas alcoólicas nas dependências da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q150	D150	Consumo de drogas (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de drogas nas dependências da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q151	D151	Consumo de drogas nas proximidades da escola (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de drogas nas proximidades da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q152	D152	Tráfico de drogas nas dependências da escola (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: tráfico de drogas nas dependências da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q153	D153	Tráfico de drogas nas proximidades da escola (estranho à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: tráfico de drogas nas proximidades da escola. Agente causador externo (estranho à escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q154	D154	Atentado à vida de professores dentro da escola(interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: atentado à vida de professores ou funcionários dentro da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q155	D155	Atentado à vida de alunos dentro da escola (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: atentado à vida de alunos dentro da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q156	D156	Furto aos professores dentro da escola (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: furto a professores ou funcionários dentro da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q157	D157	Furto aos alunos dentro da escola (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: furto a alunos dentro da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q158	D158	Roubo aos professores (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) a professores Agente causador interno	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q159	D159	Roubo aos alunos (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) a alunos dentro da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q160	D160	Furto de equipamentos (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: furto de equipamentos e materiais didáticos ou pedagógicos da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q161	D161	Roubo de equipamentos (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) de equipamentos e materiais didáticos ou pedagógicos da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q162	D162	Quebra intencional de equipamentos (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: quebra intencional de equipamento. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q163	D163	Pichação de muros externos (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: pichação de muros ou paredes das dependências externas da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q164	D164	Depredação de áreas externas (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: depredação das dependências externas da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q165	D165	Sujeira nas áreas externas (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: sujeira nas dependências externas da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q166	D166	Sujeira nas áreas internas (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: sujeira nas dependências internas da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q167	D167	Pichação nas áreas internas (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: pichação de muros ou paredes das dependências internas da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q168	D168	Depredação nas dependências internas (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: depredação das dependências internas da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q169	D169	Depredação de banheiros nas dependências internas (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: depredação de banheiros. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q170	D170	Consumo de bebidas alcoólicas nas dependências internas (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de bebidas alcoólicas nas dependências da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q171	D171	Consumo de drogas nas dependências internas (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de drogas nas dependências da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q172	D172	Consumo de drogas nas proximidades da escola (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de drogas nas proximidades da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q173	D173	Tráficos de drogas nas dependências da escola (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: tráfico de drogas nas dependências da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q174	D174	Tráfico de drogas nas proximidades da escola (interno à escola)	Neste ano, aconteceram os seguintes fatos nesta escola: tráfico de drogas nas proximidades da escola. Agente causador interno (da própria escola)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q175	D175	Porte de arma de fogo	Neste ano, os seguintes eventos fizeram ou não parte do cotidiano desta escola: membros da comunidade escolar portando arma de fogo.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q176	D176	Porte de arma branca	Neste ano, os seguintes eventos fizeram ou não parte do cotidiano desta escola: membros da comunidade escolar portando arma branca (faca, canivete, estilete, etc.).	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q177	D177	Ação de gangues na área externa a escola	Neste ano, os seguintes eventos fizeram ou não parte do cotidiano desta escola: ação de gangues nas dependências externas da escola.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q178	D178	Ação de gangues na área interna a escola	Neste ano, os seguintes eventos fizeram ou não parte do cotidiano desta escola: ação de gangues nas dependências internas da escola.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q179	D179	Agressão verbal aos professores por alunos	Neste ano, houve agressão verbal a professores. Quem foi o agressor? Aluno	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q180	D180	Agressão física aos professores por alunos	Neste ano, houve agressão física a professores. Quem foi o agressor? Aluno	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q181	D181	Agressão verbal aos alunos por alunos	Neste ano, houve agressão verbal a alunos. Quem foi o agressor? Aluno	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q182	D182	Agressão física aos alunos por alunos	Neste ano, houve agressão física a alunos. Quem foi o agressor? Aluno	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q183	D183	Agressão verbal aos funcionários por alunos	Neste ano, houve agressão verbal a funcionários. Quem foi o agressor? Aluno	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q184	D184	Agressão física aos funcionários por alunos	Neste ano, houve agressão física a funcionários. Quem foi o agressor? Aluno	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q185	D185	Agressão verbal aos professores por professores	Neste ano, houve agressão verbal a professores. Quem foi o agressor? Professor	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q186	D186	Agressão física aos professores por professores	Neste ano, houve agressão física a professores. Quem foi o agressor? Professor	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q187	D187	Agressão verbal aos alunos por professores	Neste ano, houve agressão verbal a alunos. Quem foi o agressor? Professor	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q188	D188	Agressão física aos alunos por professores	Neste ano, houve agressão física a alunos. Quem foi o agressor? Professor	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q189	D189	Agressão verbal aos funcionários por professores	Neste ano, houve agressão verbal a funcionários. Quem foi o agressor? Professor	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q190	D190	Agressão física aos funcionários por professores	Neste ano, houve agressão física a funcionários. Quem foi o agressor? Professor	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q191	D191	Agressão verbal aos professores por funcionários	Neste ano, houve agressão verbal a professores. Quem foi o agressor? Funcionário	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q192	D192	Agressão física aos professores por funcionários	Neste ano, houve agressão física a professores. Quem foi o agressor? Funcionário	categórica nominal	$6,02.10^{-14}$
TX_RESP_Q193	D193	Agressão verbal aos alunos por funcionários	Neste ano, houve agressão verbal a alunos. Quem foi o agressor? Funcionário	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q194	D194	Agressão física aos alunos por funcionários	Neste ano, houve agressão física a alunos. Quem foi o agressor? Funcionário	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q195	D195	Agressão verbal aos funcionários por funcionários	Neste ano, houve agressão verbal a funcionários. Quem foi o agressor? Funcionário	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q196	D196	Agressão física aos funcionários por funcionários	Neste ano, houve agressão física a funcionários. Quem foi o agressor? Funcionário	categórica nominal	$7,53.10^{-16}$
TX_RESP_Q197	D197	Projetos sobre violência	Nesta escola há projetos nas seguintes temáticas violência.	categórica nominal	$6,01.10^{-16}$
TX_RESP_Q198	D198	Projetos sobre drogas	Nesta escola há projetos nas seguintes temáticas uso abusivo de drogas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q199	D199	Projetos sobre racismo	Nesta escola há projetos nas seguintes temáticas racismo.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q200	D200	Projetos sobre machismo	Nesta escola há projetos nas seguintes temáticas machismo e homofobia.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q201	D201	Projetos sobre bullying	Nesta escola há projetos nas seguintes temáticas bullying.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q202	D202	Projetos sobre sexualidade	Nesta escola há projetos nas seguintes temáticas sexualidade e gravidez na adolescência.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q203	D203	Projetos sobre desigualdades sociais	Nesta escola há projetos nas seguintes temáticas desigualdades sociais.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q204	D204	Projetos sobre conflitos religiosos	Nesta escola há projetos nas seguintes temáticas conflitos religiosos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q205	D205	Projetos sobre meio ambiente	Nesta escola há projetos nas seguintes temáticas meio ambiente.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q206	D206	Músicas religiosas ou orações	Nesta escola há costume de se fazer oração ou cantar músicas religiosas (na entrada, antes do lanche, etc).	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q207	D207	Símbolos religioso	Nesta escola há objetos, imagens, frases ou símbolos de teor religioso.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q208	D208	Ensino religioso	Nesta escola há aula de ensino religioso.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q209	D209	Alternativas ao ensino religioso	Nesta escola há atividades alternativas para os estudantes que não queiram participar das aulas de ensino religioso.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q210	D210	Obrigatoriedade do ensino religioso	Nesta escola, o ensino religioso (se não houver ensino religioso deixe as próximas questões em branco): é de presença obrigatória.	categórica nominal	0,0079
TX_RESP_Q211	D211	Caráter do ensino religioso	Nesta escola, o ensino religioso (se não houver ensino religioso deixe as próximas questões em branco): é de caráter confessional (segundo uma religião específica).	categórica nominal	$1,32.10^{-7}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q212	D212	Diversidade do ensino religioso	Nesta escola, o ensino religioso (se não houver ensino religioso deixe as próximas questões em branco): contempla a diversidade religiosa.	categórica nominal	$3,82.10^{-8}$
TX_RESP_Q001	P001	Sexo do professor	Sexo	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q002	P002	Idade do professor	Idade	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q003	P003	Raça do professor	Como você se considera?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q004	P004	Escolaridade do professor	Qual o seu nível de escolaridade (Até a Graduação). Caso Marque A, B ou C, passe para questão 11	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q005	P005	Tempo de escolaridade do professor	Há Quantos anos você obteve o nível de escolaridade assinalado anteriormente?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q006	P006	Tipo de instituição onde o professor obteve seu título profissional	Em que tipo de instituição você fez o curso superior? Se você estudou em mais de uma instituição, assinale aquela em que obteve o seu título profissional.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q007	P007	Natureza da instituição onde o professor obteve seu título profissional	Qual era a natureza dessa instituição?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q008	P008	Forma de realização do curso superior pelo professor	De que forma você realizou o curso superior?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q009	P009	Modalidade da pós - graduação do professor	Indique a modalidade de cursos de pós-graduação de mais alta titulação que você possui.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q010	P010	Temática da pós - graduação do professor	Indique qual a área temática do curso de pós-graduação de mais alta titulação que você possui.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q011	P011	Formação continuada do professor	Você participou de alguma atividade de formação continuada (atualização, treinamento, capacitação nos últimos dois anos?etc.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q012	P012	Carga horária da atividade mais relevante do professor	Qual a carga horária da atividade mais relevante da qual você participou?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q013	P013	Aplicação da formação continuada pelo professor	Você utiliza os conhecimentos adquiridos nas atividades de formação continuada para a melhoria de sua prática em sala de aula?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q014	P014	Salário bruto do professor na escola	Nesta escola, qual, aproximadamente, o seu salário bruto? (com adicionais, se houver) (preencha os espaços com o valor e marque, na folha de respostas, os campos correspondentes a cada algarismo.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q015	P015	Atividade de renda complementar do professor	Além da atividade como docente nesta escola, você exerce outra atividade que contribui para sua renda pessoal?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q016	P016	Salário bruto do professor	Qual é, aproximadamente, seu salário bruto (com adicionais, se houver) como professor(a)? (soma de tudo o que você ganha como professor(a)) (preencha os espaços com o valor e marque, na folha de respostas, os campos correspondentes a cada algarismo.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q017	P017	Tempo de trabalho na educação do professor	Há quantos anos você leciona?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q018	P018	Tempo como professor desta escola	Há quantos anos você trabalha nesta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q019	P019	Tempo como professor desta série	Há quantos anos você ministra aulas para alunos das séries desta turma?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q020	P020	Carga horária semanal do professor nesta escola	Nesta escola, qual a sua carga horária semanal? (considere a carga horária contratual: horas-aula mais horas para atividades, se houver.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q021	P021	Número de escolas que o professor trabalha	Em quantas escolas você trabalha? (caso marque a alternativa a, passe para a questão 23)	categórica nominal	$4,48.10^{-16}$
TX_RESP_Q022	P022	Horas - aulas semanais pelo professor	Ao todo, quantas horas-aula você ministra por semana? (não considere aulas particulares)	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q023	P023	Situação trabalhista do professor nesta escola	Qual é a sua situação trabalhista nesta escola? (marque apenas uma opção)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q024	P024	Biblioteca (professor)	Você costuma: frequentar bibliotecas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q025	P025	Cinema (professor)	Você costuma: ir ao cinema.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q026	P026	Museu (professor)	Você costuma: ir ao museu.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q027	P027	Apresentações teatrais (professor)	Você costuma: ver apresentações teatrais.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q028	P028	Apresentações musicais (professor)	Você costuma: ver apresentações musicais ou de dança	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q029	P029	Jornais (professor)	Em seu tempo livre, você lê: jornais.	categórica nominal	$1,69.10^{-11}$
TX_RESP_Q030	P030	Revistas de informação em geral (professor)	Em seu tempo livre, você lê: revistas de informação geral.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q031	P031	Revistas de humor (professor)	Em seu tempo livre, você lê: revistas de humor/quadrinhos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q032	P032	Revistas de divulgação científica (professor)	Em seu tempo livre, você lê: revistas de divulgação científica/cultural.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q033	P033	Revistas de comportamento (professor)	Em seu tempo livre, você lê: revistas sobre comportamento, celebridades, esporte e tv.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q034	P034	Livros (professor)	Em seu tempo livre, você lê: livros.	categórica nominal	$2,82.10^{-16}$
TX_RESP_Q035	P035	Sites da internet (professor)	Em seu tempo livre, você lê: sites da internet.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q036	P036	Carga horária extra-classe do professor	Dentro da sua carga horária, quantas horas semanais são dedicadas à atividades extraclasse (formação e estudo, planejamento, produção de recursos didáticos etc)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q037	P037	Utilização de jornais em salas de aulas	Indique se você utiliza ou não nesta escola: jornais e revistas informativas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q038	P038	Utilização de livros de literatura em salas de aulas	Indique se você utiliza ou não nesta escola: livros de literatura em geral.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q039	P039	Utilização de projetor de slides em salas de aulas	Indique se você utiliza ou não nesta escola: projetor de slides.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q040	P040	Utilização de retroprojetor em salas de aulas	Indique se você utiliza ou não nesta escola: retroprojetor.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q041	P041	Utilização de máquinas copiadoras em salas de aula	Indique se você utiliza ou não nesta escola: máquina copiadora.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q042	P042	Projeto pedagógico	Como foi desenvolvido o projeto pedagógico desta escola neste ano? (marque apenas uma opção)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q043	P043	Reuniões do conselho de classe	Conselho de classe é um órgão formado por todos os professores que lecionam em cada turma/série. Neste ano, quantas vezes se reuniram os conselhos de classe desta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q044	P044	História e cultura afro-brasileira	Você tem conhecimento do conteúdo da lei número 11.645 de 2008 que determina a obrigatoriedade do estudo da temática história e cultura afro-brasileira e indígena nos estabelecimentos de ensino do país?	categórica nominal	$7,73.10^{-14}$
TX_RESP_Q045	P045	Atividades de história e cultura afro-brasileira	Neste ano, foram desenvolvidas atividades para atender o determinado pela lei n.º 11.645 de 2008 nesta escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q046	P046	Infra-estrutura física ou pedagógica	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem na escola devido à carência de infraestrutura física e/ou pedagógica.	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q047	P047	Insegurança física da escola	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): são decorrentes do ambiente de insegurança física da escola.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q048	P048	Desenvolvimento das capacidades intelectuais	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem na escola por oferecer poucas oportunidades de desenvolvimento das capacidades intelectuais do aluno.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q049	P049	Dificuldades dos alunos relacionada aos conteúdos curriculares	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos avaliada(s): estão relacionados aos conteúdos curriculares, que são inadequados às necessidades dos alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q050	P050	Dificuldades dos alunos relacionada ao não cumprimento do conteúdo curricular.	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): estão relacionados ao não cumprimento do conteúdo curricular.	categórica nominal	$3,36.10^{-14}$
TX_RESP_Q051	P051	Dificuldades dos alunos relacionada à sobrecarga de trabalho dos professores.	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): relacionam-se à sobrecarga de trabalho do(as) professores(as), dificultando o planejamento e o preparo das aulas.	categórica nominal	$5,68.10^{-14}$
TX_RESP_Q052	P052	Dificuldades dos alunos relacionada ao baixo salário dos professores	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem devido ao baixo salário dos insatisfação e desestímulo para a atividade docente.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q053	P053	Dificuldades dos alunos relacionada são devido ao meio que vive	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): são decorrentes do meio em que o aluno vive.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q054	P054	Dificuldades dos alunos relacionada são devido ao meio cultural dos pais do aluno	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): são decorrentes do nível cultural dos pais dos alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q055	P055	Dificuldades dos alunos relacionada à falta de acompanhamento familiar	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): estão relacionadas à falta de assistência e acompanhamento da família nos deveres de casa e pesquisas dos alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q056	P056	Dificuldades dos alunos relacionada à falta de aptidão do aluno	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem devido à falta de aptidão e habilidades do aluno.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q057	P057	Dificuldades dos alunos relacionada à baixa autoestima do aluno	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): estão vinculados à baixa autoestima dos alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q058	P058	Dificuldades dos alunos relacionada à falta de esforço do aluno	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem devido ao desinteresse e falta de esforço do aluno.	categórica nominal	$1,39.10^{-8}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q059	P059	Dificuldades dos alunos relacionada à indisciplina do aluno	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): são decorrentes da indisciplina dos alunos em sala de aula.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q060	P060	Motivação do diretor ao trabalho do professor	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): ao(a) diretor(a) me anima e me motiva para o trabalho.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q061	P061	Confiança no diretor	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): tenho plena confiança no(a) diretor(a) como profissional.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q062	P062	Comprometimento do professor	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) consegue que os professores se comprometam com a escola.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q063	P063	Estimulo às atividades inovadoras	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) estimula as atividades inovadoras.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q064	P064	Atenção do diretor à aprendizagem dos alunos	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) dá atenção especial a aspectos relacionados com a aprendizagem dos alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q065	P065	Atenção do diretor às normas administrativas	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) dá atenção especial a aspectos relacionados com as normas administrativas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q066	P066	Atenção do diretor às manutenções da escola	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) dá atenção especial a aspectos relacionados com a manutenção da escola.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q067	P067	Respeito do diretor ao professor	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): sinto-me respeitado(a) pelo(a) diretor(a).	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q068	P068	Respeito ao diretor pelo professor	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): respeito o(a) diretor(a).	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q069	P069	Participação das decisões relacionadas à escola pelo professor	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): participo das decisões relacionadas com o meu trabalho.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q070	P070	Professores leva em consideração minhas ideias.	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): a equipe de professores leva em consideração minhas ideias.	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q071	P071	Eu levo em consideração as ideias de outros colegas.	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): eu levo em consideração as ideias de outros colegas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q072	P072	Interação entre os professores	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o ensino que a escola oferece aos alunos é muito influenciado pela troca de ideias entre os professores.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q073	P073	Interação entre os conteúdos das disciplinas	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): os professores desta escola procuram coordenar o conteúdo das disciplinas entre as diferentes séries.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q074	P074	Colaboração entre os membros da comunidade escolar	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o diretor, os professores e os demais membros da equipe da escola colaboram para fazer esta escola funcionar bem.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q075	P075	Carência de recursos financeiros	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: insuficiência de recursos financeiros?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q076	P076	Falta de professores	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: inexistência de professores para algumas disciplinas ou séries?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q077	P077	Falta de pessoal administrativo	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: carência de pessoal administrativo?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q078	P078	Falta de apoio pedagógico	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: carência de pessoal de apoio pedagógico (coordenador, supervisor, orientador educacional)?	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q079	P079	Falta de recurso pedagógico	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: falta de recursos pedagógicos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q080	P080	Índice de falta dos professores	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: alto índice de faltas por parte dos professores?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q081	P081	Índice de falta dos alunos	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: alto índice de faltas por parte dos alunos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q082	P082	Indisciplinas dos alunos	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: problemas disciplinares causados pelos alunos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q083	P083	Agressão verbal a professores feita por alunos	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno): agressão verbal a professores.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q084	P084	Agressão física aos professores feita por alunos	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno): agressão física a professores.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q085	P085	Agressão verbal aos alunos feitas por alunos	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno): agressão verbal a alunos.	categórica nominal	$1,51.10^{-12}$
TX_RESP_Q086	P086	Agressão física aos alunos feita por alunos	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno): agressão física a alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q087	P087	Agressão verbal aos funcionários feita por alunos	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno): agressão verbal a funcionários.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q088	P088	Agressão física aos funcionários feitos por alunos	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno): agressão física a funcionários.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q089	P089	Agressão verbal aos professores feito por professores	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor): agressão verbal a professores.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q090	P090	Agressão física aos professores feito por professores	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor): agressão física a professores.	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q091	P091	Agressão verbal aos alunos por professor	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor): agressão verbal a alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q092	P092	Agressão física aos alunos feita por professor	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor): agressão física a alunos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q093	P093	Agressão verbal aos funcionários feita por professor	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor): agressão verbal a funcionários.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q094	P094	Agressão física aos funcionários feita por professor	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor): agressão física a funcionários.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q095	P095	Agressão verbal aos professores feita por funcionários	Agressão verbal a professores. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q096	P096	Agressão física aos professores feita por funcionários	Agressão física a professores. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q097	P097	Agressão verbal aos alunos feita por funcionários	Agressão verbal a alunos. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q098	P098	Agressão física aos alunos feita por funcionários	Agressão física a alunos. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q099	P099	Agressão verbal aos funcionários feita por funcionários	Agressão verbal a funcionários. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q100	P100	Agressão física aos funcionários feita por funcionários	Agressão física a funcionários. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q101	P101	Atentado à vida	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi vítima de atentado à vida?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q102	P102	Ameaças à vida feita por alunos	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi ameaçado por algum aluno?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q103	P103	Agressão verbal por algum aluno	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi agredido verbalmente por algum aluno?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q104	P104	Agressão física por algum aluno	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi agredido fisicamente por algum aluno?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q105	P105	Furto aos professores	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi vítima de furto?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q106	P106	Roubo aos professores	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi vítima de roubo (com uso de violência)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q107	P107	Bebida alcoólicas	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: alunos frequentaram as suas aulas sob efeito de bebida alcoólica?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q108	P108	Drogas ilícitas	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: alunos frequentaram as suas aulas sob efeito de drogas ilícitas?	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q109	P109	Porte de arma branca	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: alunos frequentaram as suas aulas portando arma branca (facas, canivetes, etc.)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q110	P110	Porte de arma de fogo	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: alunos frequentaram as suas aulas portando arma de fogo?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q111	P111	Prova Brasil de 2009	A sua escola participou da Prova Brasil de 2009?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q112	P112	Conhecimento dos resultados do SAEB da sua escola	Você conhece os resultados do sistema de avaliação da educação básica, SAEB: da sua escola?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q113	P113	Conhecimento dos resultados do SAEB do seu município	Você conhece os resultados do sistema de avaliação da educação básica, SAEB: do seu município?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q114	P114	Conhecimento dos resultados do SAEB do seu estado	Você conhece os resultados do sistema de avaliação da educação básica, SAEB: do seu estado?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q115	P115	Utilização de computadores pelos professores em sala de aula	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): computadores.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q116	P116	Utilização de internet pelos professores em sala de aula	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): internet.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q117	P117	Utilização de fitas de vídeo ou DVD pelos professores em sala de aula	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): fitas de vídeo ou DVD.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q118	P118	Utilização de jornais e revistas pelos professores em sala de aula	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): jornais e revistas informativas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q119	P119	Utilização de revistas em quadrinhos pelos professores em sala de aula	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): revistas em quadrinhos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q120	P120	Utilização de livros didáticos pelos professores em sala de aula	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): livros didáticos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q121	P121	Quantidade de conteúdos curriculares cumpridas pelo professor	Quanto dos conteúdos previstos você conseguiu desenvolver com os alunos desta turma, neste ano?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q122	P122	Perspectiva de conclusão do Ensino Fundamental I	Quantos dos alunos desta turma você acha que: concluirão os anos iniciais do ensino fundamental (4. ^a série/5. ^o ano)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q123	P123	Perspectiva de conclusão do Ensino Fundamental II	Quantos dos alunos desta turma você acha que: concluirão o ensino fundamental (8. ^a série/9. ^o ano)?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q124	P124	Perspectiva de conclusão do Ensino Médio	Quantos dos alunos desta turma você acha que: concluirão o ensino médio?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q125	P125	Perspectiva de ingresso nas universidades	Quantos dos alunos desta turma você acha que: entrarão para a universidade?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q126	P126	Porte de livros didáticos pelos alunos	Os alunos desta turma têm livros didáticos?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q127	P127	Recebimento de livros didáticos pelos alunos	Os alunos desta turma receberam o livro didático no início do ano letivo?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q128	P128	Escolha e recebimento de livros didáticos	O livro didático escolhido foi o recebido?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q129	P129	Qualidade do livros didáticos desta turma	Como você considera o(s) livro(s) didático(s) utilizado(s) por você na(s) disciplina(s) que ministra nesta turma?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q130	P130	Guia de livros didáticos do MEC	Na escolha do livro didático utilizado nesta turma em que você leciona, foi consultado o "guia de livros didáticos" da SEB/MEC?	categórica nominal	$4,31.10^{-15}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q131	P131	Forma de escolha do livro didático	Para a disciplina que você ministra, como foi escolhido o livro didático para utilização nesta turma? (marque apenas uma opção.)	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q132	P132	Forma de utilização da sala de leitura	Como você utiliza a biblioteca ou sala de leitura da escola em apoio às suas aulas nesta turma?	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q133	P133	O professor de Língua Portuguesa pede que faça cópias do quadro negro ou de livros	(Professor de língua portuguesa) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: copiar textos do livro didático ou do quadro negro ou lousa.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q134	P134	O professor de Língua Portuguesa promove discussões a partir de textos de jornais e revistas	(Professor de língua portuguesa) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: promover discussões a partir de textos de jornais e revistas.	categórica nominal	$3,08.10^{-12}$
TX_RESP_Q135	P135	Exercícios gramaticais relacionados com textos de jornais e revistas	(Professor de língua portuguesa) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: fazer exercícios sobre gramática relacionados com textos de jornais ou revistas.	categórica nominal	$2,88.10^{-13}$
TX_RESP_Q136	P136	Leitura, discussão e escrita de textos relacionados com projetos temáticos	(Professor de língua portuguesa) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: ler, discutir com colegas e escrever textos relacionados com o desenvolvimento de projeto temático.	categórica nominal	$9,83.10^{-8}$
TX_RESP_Q137	P137	Leitura de contos, crônicas, poesias ou romances	(Professor de língua portuguesa) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: ler contos, crônicas, poesias ou romances.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q138	P138	Conversa sobre contos, crônicas, poesias ou romances	(Professor de língua portuguesa) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: conversar sobre contos, crônicas, poesias ou romances.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q139	P139	Uso de contos, crônicas e poesias ou romances para o estudo de gramática	(Professor de língua portuguesa) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: utilizar contos, crônicas, poesias ou romances para exercitar aspectos da gramática.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q140	P140	Fixação de conceitos gramaticais e linguísticos	(Professor de língua portuguesa) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: fixar os nomes de conceitos gramaticais e linguísticos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q141	P141	Promoção de debates levando em consideração a diferença de opiniões	(Professor de língua portuguesa) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: discutir um texto, explorando as diferenças entre fatos e opiniões.	categórica nominal	$5,86.10^{-10}$
TX_RESP_Q142	P142	Resolução de exercícios matemáticos	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: fazer exercícios para fixar procedimentos e regras.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q143	P143	Raciocínios complexos	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: lidar com situações problemas que exigem raciocínios diferentes e mais complexos que a maioria dos exemplos usuais.	categórica nominal	$< 2.10^{-16}$

Código original	Código na base de dados	Nome da variável	Enunciado	Tipo de variável	valor p
TX_RESP_Q144	P144	Análise de soluções de problemas	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: falar sobre suas soluções, discutindo os caminhos usados para encontrá-las.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q145	P145	Fixação de regras matemáticas	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: gravar as regras que permitem obter as respostas certas dos cálculos e problemas.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q146	P146	Contextualização dos temas de jornais e revistas em matemática	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: lidar com temas que aparecem em jornais e/ou revistas, discutindo a relação dos temas com a matemática.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q147	P147	Interpretação de resultados obtidos em resolução de problemas	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q148	P148	Temas matemáticos de interesse dos alunos	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: lidar com situações que lhes sejam familiares e que apresentem temas de interesse dos alunos.	categórica nominal	$7,71.10^{-10}$
TX_RESP_Q149	P149	Exploração de diferentes métodos de resolução de problemas	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: experimentar diferentes modos de resolver um problema ou de efetuar um cálculo.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q150	P150	Precisão de execução de cálculos	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: aprimorar a precisão e a velocidade de execução de cálculos.	categórica nominal	$< 2.10^{-16}$
TX_RESP_Q151	P151	Táticas de resolução de problemas	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: experimentar diferentes ações (coletar informações, recortar, analisar, explorar, discutir, manipular, etc.) Para resolver problemas.	categórica nominal	$3,39.10^{-10}$
TX_RESP_Q152	P152	Análise crítica de resultados matemáticos	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: incentivar e estimular o aluno a analisar criticamente se os resultados obtidos na resolução de um problema são plausíveis.	categórica nominal	$< 2.10^{-16}$
PESO	PESO	PESO		Contínua	$< 2.10^{-16}$
ID_TURNO	TURNO	Turno da turma		Categórica nominal	$< 2.10^{-16}$

Tabela 31 – Codificação das variáveis

Questão	Questão	Alternativa	Código
Q001	Sexo	A	0
		B	1
Q004	Em que ano você nasceu?	A	8
		B	7
		C	6
		D	5
		E	4
		F	3
		G	2
		H	1
Q005	Na sua casa tem televisão em cores?	D	1
		A	2
		B	3
		C	4
Q006	Na sua casa tem rádio?	D	1
		A	2
		B	3
		C	4
Q007	Na sua casa tem videocassete ou DVD?	B	1
		A	2
Q008	Na sua casa tem geladeira?	C	1
		A	2
		B	3
Q009	Na sua casa tem freezer junto à geladeira?	C	1
		B	2
		A	3
Q010	Na sua casa tem freezer separado da geladeira?	C	1
		B	2
		A	3
Q011	Na sua casa tem máquina de lavar roupas?	B	1
		A	2
Q012	Na sua casa tem carro?	D	1
		A	2
		B	3
		C	4
Q013	Na sua casa tem computador?	C	1
		B	2
		A	3
Q014	Dentro da sua casa tem banheiro?	E	1
		A	2
		B	3
		C	4
		D	5
Q015	Na sua casa trabalha alguma empregada doméstica?	D	1
		A	2
		B	3
		C	4
Q016	Na sua casa tem quartos para dormir?	E	1
		A	2
		B	3
		C	4
		D	5
Q017	Quantas pessoas moram com você?	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
Q018	Você mora com sua mãe?	B	1
		C	2
		A	3
Q019	Até que série sua mãe ou a mulher responsável por você estudou?	G	1
		A	2
		B	3
		C	4
		D	5
		E	6
F	7		
Q020	Sua mãe ou mulher responsável por você sabe ler e escrever?	C	1
		B	2
		A	3

Questão	Questão	Alternativa	Código
Q021	Você vê sua mãe ou mulher responsável por você lendo?	B A	1 2
Q022	Você mora com seu pai?	B C A	1 2 3
Q023	Até que série seu pai ou o homem responsável por você estudou?	G A B C D E F	1 2 3 4 5 6 7
Q024	Seu pai ou homem responsável por você sabe ler e escrever?	C B A	1 2 3
Q025	Você vê o seu pai ou homem responsável por você lendo?	B A	1 2
Q026	Com que frequência seus pais ou responsáveis vão á reunião de pais?	C B A	1 2 3
Q027	Seus pais ou responsáveis incentivam você a estudar?	B A	1 2
Q028	Seus pais ou responsáveis incentivam você a fazer o dever de casa e os trabalhos da escola?	B A	1 2
Q029	Seus pais ou responsáveis incentivam você a ler?	B A	1 2
Q030	Seus pais ou responsáveis incentivam você a ir à escola e não faltar às aulas?	B A	1 2
Q031	Seus pais ou responsáveis conversam com você sobre o que acontece na escola?	B A	1 2
Q032	Você lê: Jornais	C B A	1 2 3
Q033	Você lê: Livros em geral	C B A	1 2 3
Q034	Você lê: Livros de literatura infanto-juvenil.	A B C	3 2 1
Q035	Você lê: Revistas em geral.	A B C	3 2 1
Q036	Você lê: Revistas em quadrinhos.	A B C	3 2 1
Q037	Você lê: Revistas sobre comportamento, celebridades, esporte e TV.	A B C	3 2 1
Q038	Você lê: Sites da internet.	A B C	3 2 1
Q039	Você Costuma: Frequentar bibliotecas.	A B C	3 2 1
Q040	Você Costuma: Ir ao cinema.	A B C	3 2 1
Q041	Você Costuma: Ver apresentações teatrais.	A B C	3 2 1
Q042	Você Costuma: Ver apresentações musicais ou de dança.	A B C	3 2 1
Q043	Você Costuma: Participar de festas abertas à comunidade.	A B C	3 2 1
Q044	Em dia de aula, quanto tempo você gasta assistindo à TV, navegando na internet ou jogando jogos eletrônicos?	A B C D	1 2 3 4
Q045	Em dias de aula, quanto tempo você gasta fazendo trabalhos domésticos?	A B C D E	2 3 4 5 1

Questão	Questão	Alternativa	Código
Q046	Você trabalha fora de casa?	A	1
		B	0
Q047	Quando você entrou na escola?	A	1
		B	2
		C	3
		D	4
Q048	Desde a quinta série em que tipo de escola você estudou?	A	1
		B	3
		C	2
Q049	Você já foi reprovado?	A	1
		B	2
		C	3
Q050	Você já abandonou a escola durante o período de aulas e ficou fora da escola o resto do ano?	A	1
		B	2
		C	3
Q051	Você gosta de estudar língua portuguesa?	A	1
		B	0
Q052	Você faz o dever de casa de língua portuguesa?	A	3
		B	2
		C	1
Q053	O professor corrige o dever de casa de língua portuguesa?	A	3
		B	2
		C	1
Q054	Você gosta de estudar matemática?	A	1
		B	0
Q055	Você faz o dever de casa de matemática?	A	3
		B	2
		C	1
Q056	O professor corrige o dever de casa de matemática?	A	3
		B	2
		C	1
Q057	Você utiliza a biblioteca ou sala de leitura da sua escola?	A	4
		B	3
		C	2
		D	1
Q058	Quando você terminar o 9º ano(8ª série), você pretende:	A	4
		B	3
		C	2
		D	1
D001	Sexo do diretor	A	0
		B	1
D002	Idade	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
D004	Qual o seu nível de escolaridade (até a graduação).	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
		I	9
		J	10
D005	Há quantos anos você obteve o nível de escolaridade assinalado anteriormente?	A	1
		B	2
		C	3
		D	4
		E	5
D006	Em que tipo de instituição você fez o curso superior? Se você estudou em mais de uma instituição, assinale aquela em que obteve seu título profissional.	A	5
		B	4
		C	3
		D	2
		E	1
D007	Qual era a natureza dessa instituição?	A	3
		B	2
		C	4
		D	1
D008	De que forma você realizou o curso superior?	A	4
		B	3
		C	2
		D	1

Questão	Questão	Alternativa	Código
D009	Indique a modalidade de cursos de pós-graduação de mais alta titulação que você possui.	A B C D E	2 3 4 5 1
D010	Indique qual a área temática do curso de pós-graduação de mais alta titulação que você possui.	A B C D E	5 4 3 2 1
D011	Você participou de alguma atividade de formação continuada (atualização, treinamento, capacitação nos últimos dois anos?etc.)	A B	1 0
D012	Qual a carga horária da atividade mais relevante da qual você participou?	A B C D	1 2 3 4
D013	Você utiliza os conhecimentos adquiridos nas atividades de formação continuada de que você participou?	A B C D	4 3 2 1
D014	Qual é aproximadamente o seu salário bruto como diretor(a), com adicionais, se houver? (preencha os espaços com o valor e, abaixo de cada algarismo, marque o campo	A B C D E F G H I J K	1 2 3 4 5 6 7 8 9 10 11
D015	Qual é aproximadamente sua renda familiar bruta? (preencha os espaços com o valor e, abaixo de cada algarismo, marque o campo correspondente.)	A B C D E F G H I J K	1 2 3 4 5 6 7 8 9 10 11
D016	Além da direção desta escola, você exerce outra atividade que contribui para sua renda pessoal?	A B C	3 2 1
D017	Há quantos anos você trabalha em educação?	A B C D E F G H	1 2 3 4 5 6 7 8
D018	Há quantos anos você exerce funções de direção?	A B C D E F G H	1 2 3 4 5 6 7 8
D019	Há quantos anos você é diretor(a) desta escola?	A B C D E F G H	1 2 3 4 5 6 7 8
D020	Qual é a sua carga horária de trabalho nesta escola?	A B C D	1 2 3 4
D022	Você promoveu alguma atividade de formação continuada (atualização, treinamento, capacitação etc.) Nesta escola?	A B	1 0

Questão	Questão	Alternativa	Código
D023	Qual foi a proporção de docentes da sua escola que participou das atividades de formação continuada promovidas por você nos últimos dois anos?	A	2
		B	3
		C	4
		D	5
		E	1
D024	Conselho de escola é um colegiado constituído por representantes da escola e da comunidade que tem como objetivo acompanhar as atividades escolares. Neste ano, quantas vezes o conselho desta escola se reuniu?	A	3
		B	4
		C	5
		D	2
		E	1
D025	O conselho de escola é composto por professores. (marque sim ou não em cada linha)	A	1
		B	0
D026	O conselho de escola é composto por alunos. (marque sim ou não em cada linha)	A	1
		B	0
D027	O conselho de escola é composto por funcionários. (marque sim ou não em cada linha)	A	1
		B	0
D028	O conselho de escola é composto por pais. (marque sim ou não em cada linha)	A	1
		B	0
D029	Conselho de classe é um órgão formado por todos os professores que lecionam em cada turma/série. Neste ano, quantas vezes se reuniram os conselhos de classe desta escola?	A	3
		B	4
		C	5
		D	2
		E	1
D030	Quanto ao projeto pedagógico desta escola neste ano (marque apenas uma opção)	A	6
		B	7
		C	8
		D	9
		E	5
		F	4
		G	3
		H	1
		I	2
D031	Qual é o critério para a admissão de alunos nesta escola?	A	6
		B	5
		C	4
		D	3
		E	2
		F	1
D032	Neste ano letivo, como foi a situação da oferta de vagas nesta escola? (marque apenas uma alternativa.)	A	1
		B	2
		C	3
		D	4
D033	Qual o critério utilizado para formação das turmas nesta escola?	A	5
		B	4
		C	3
		D	2
		E	1
D034	Neste ano, qual foi o critério mais importante para a atribuição das turmas de 1. ^a a 4. ^a séries do ensino fundamental aos professores?	A	6
		B	7
		C	8
		D	9
		E	5
		F	4
		G	3
		H	2
		I	1
D035	Qual é o percentual de professores com vínculo estável nesta escola?	A	1
		B	2
		C	3
		D	4
		E	5
D036	Nesta escola, há algum programa de redução das taxas de abandono?	A	4
		B	3
		C	2
		D	1
D037	Nesta escola, há algum programa de redução das taxas de reprovação?	A	4
		B	3
		C	2
		D	1
D038	Para evitar que os alunos faltem às aulas, os professores falam com os alunos.	A	1
		B	0
D039	Para evitar que os alunos faltem às aulas, os pais/responsáveis são avisados por comunicação escrita.	A	1
		B	0
D040	Para evitar que os alunos faltem às aulas, os pais/responsáveis são chamados à escola para conversar sobre o assunto em reunião de pais.	A	1
		B	0

Questão	Questão	Alternativa	Código
D041	Para evitar que os alunos faltem às aulas, os pais/responsáveis são chamados à escola para conversar sobre o assunto individualmente.	A	1
		B	0
D042	Para evitar que os alunos faltem às aulas, a escola envia alguém à casa do aluno.	A	1
		B	0
D043	Esta escola desenvolve, regularmente, algum programa de apoio ou reforço de aprendizagem para os alunos (monitoria, aula de reforço etc.)?	A	1
		B	0
D044	Você tem conhecimento do conteúdo da lei 11.645 de 2008 que determina a obrigatoriedade do estudo da temática "história e cultura afro-brasileira e indígena" nos estabelecimentos de ensino do país?	A	1
		B	0
D045	Neste ano, foram desenvolvidas atividades para atender o determinado pela lei 11.645 de 2008 nesta escola?	A	3
		B	2
		C	1
D046	Em relação à merenda escolar, como você considera recursos financeiros.	A	1
		B	2
		C	3
		D	4
		E	5
D047	Em relação à merenda escolar, como você considera quantidade de alimentos.	A	1
		B	2
		C	3
		D	4
		E	5
D048	Em relação à merenda escolar, como você considera qualidade dos alimentos.	A	1
		B	2
		C	3
		D	4
		E	5
D049	Em relação à merenda escolar, como você considera variedade do cardápio.	A	1
		B	2
		C	3
		D	4
		E	5
D050	Em relação à merenda escolar, como você considera armazenamento e conservação.	A	1
		B	2
		C	3
		D	4
		E	5
D051	Em relação à merenda escolar, como você considera espaço físico para cozinhar.	A	1
		B	2
		C	3
		D	4
		E	5
D052	Em relação à merenda escolar, como você considera higiene.	A	1
		B	2
		C	3
		D	4
		E	5
D053	Em relação à merenda escolar, como você considera disponibilidade de pessoal.	A	1
		B	2
		C	3
		D	4
		E	5
D054	Em relação à merenda escolar, como você considera qualificação do pessoal.	A	1
		B	2
		C	3
		D	4
		E	5
D055	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola insuficiência de recursos financeiros?	A	3
		B	2
		C	1
D056	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola inexistência de professores para algumas disciplinas ou séries?	A	3
		B	2
		C	1
D057	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola carência de pessoal administrativo?	A	3
		B	2
		C	1

Questão	Questão	Alternativa	Código
D058	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola.(marque apenas uma opção em cada linha.) Ocorreu na escola carência de pessoal de apoio pedagógico (coordenador, supervisor, orientador educacional)?	A B C	3 2 1
D059	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola.(marque apenas uma opção em cada linha.) Ocorreu na escola falta de recursos pedagógicos?	A B C	3 2 1
D060	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola.(marque apenas uma opção em cada linha.) Ocorreu na escola interrupção das atividades escolares?	A B C	3 2 1
D061	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola.(marque apenas uma opção em cada linha.) Ocorreu na escola alto índice de faltas por parte de professores?	A B C	3 2 1
D062	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola.(marque apenas uma opção em cada linha.) Ocorreu na escola alto índice de faltas por parte de alunos?	A B C	3 2 1
D063	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola.(marque apenas uma opção em cada linha.) Ocorreu na escola rotatividade do corpo docente?	A B C	3 2 1
D064	As perguntas de 55 a 64 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola.(marque apenas uma opção em cada linha.) Ocorreu na escola problemas disciplinares causados por alunos?	A B C	3 2 1
D065	Considere as condições existentes para o exercício do cargo de diretor nesta escola. Há interferências externas em sua gestão?	A B	1 0
D066	Considere as condições existentes para o exercício do cargo de diretor nesta escola. Há apoio de instâncias superiores?	A B	1 0
D067	Considere as condições existentes para o exercício do cargo de diretor nesta escola. Há troca de informações com diretores de outras escolas?	A B	1 0
D068	Considere as condições existentes para o exercício do cargo de diretor nesta escola. Há apoio da comunidade à sua gestão?	A B	1 0
D069	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola, as calçadas possuem pavimento regular, plano, sem buracos ou degraus?	A B	1 0
D070	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola, as calçadas possuem largura mínima de 1,20 m?	A B	1 0
D071	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola, as calçadas garantem uma faixa de circulação livre de obstáculos com largura mínima de 80 cm? (obstáculos: poste, árvore, floreira, orelhão, lixeira, placa de sinalização etc.)	A B	1 0
D072	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: há semáforos para pedestres?	A B	1 0
D073	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: o semáforo é sonoro?	A B	1 0
D074	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: há faixa de travessia (faixa de pedestres ou faixa de segurança)?	A B	1 0
D075	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: as calçadas são rebaixadas com rampas dos dois lados da rua no trecho da travessia?	A B	1 0
D076	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: há vagas sinalizadas para veículos para pessoas com deficiência?	A B	1 0
D077	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: o portão de acesso de pedestres à escola é separado da entrada de carros?	A B	1 0
D078	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas imediações ou na frente da escola: o portão de acesso tem largura mínima de 80 cm?	A B	1 0

Questão	Questão	Alternativa	Código
D079	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: os corredores e as calçadas tem largura mínima de 1,20 cm?	A	1
		B	0
D080	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: há faixa de circulação livre de obstáculos com largura mínima de 80 cm? (obstáculos: extintor de incêndio, vasos, lixeiras, móveis, bebedouros etc.)	A	1
		B	0
D081	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: o piso é plano, sem degraus e sem desníveis superiores a 1,5 cm?	A	1
		B	0
D082	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: há rampas?	A	1
		B	0
D083	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: as rampas têm largura mínima de 80 cm?	A	1
		B	0
D084	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: há rampas com corrimão dos dois lados?	A	3
		B	2
		C	1
D085	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: as escadas têm corrimãos dos dois lados?	A	3
		B	2
		C	1
D086	As perguntas de 69 a 113 referem-se às condições de acessibilidade. No interior da escola, na entrada principal e nas circulações internas: há elevadores ou plataformas elevatórias?	A	3
		B	2
		C	1
D087	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas salas de aula, na administração, na biblioteca e outros ambientes: nas portas, há soleiras niveladas com o piso do corredor, pequenas rampas ou degraus de no máximo 1,5 cm?	A	1
		B	0
D088	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas salas de aula, na administração, na biblioteca e outros ambientes: as portas de entrada possuem largura mínima de 80 cm?	A	1
		B	0
D089	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas salas de aula, na administração, na biblioteca e outros ambientes: as maçanetas são de alavanca?	A	1
		B	0
D090	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Nas salas de aula, na administração, na biblioteca e outros ambientes: há algum tipo de sinalização em braille nos ambientes?	A	1
		B	0
D091	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há pelo menos um sanitário adequado a pessoas com deficiência?	A	1
		B	0
D092	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há pelo menos um sanitário adequado a pessoas com deficiência para cada sexo?	A	1
		B	0
D093	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: as portas possuem largura mínima de 80 cm?	A	1
		B	0
D094	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há banheiros ou mictórios com portas que abrem para fora?	A	1
		B	0
D095	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: as maçanetas são de alavanca?	A	1
		B	0
D096	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há espaço para cadeira de rodas e para transferência de pelo menos 80 cm?	A	1
		B	0
D097	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há barras de apoio fixadas nas paredes de fundo e na lateral do vaso sanitário?	A	1
		B	0
D098	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há pia com torneira com acionamento por alavanca, pressão ou cruzeta (em cruz)?	A	1
		B	0
D099	As perguntas de 69 a 113 referem-se às condições de acessibilidade. Os banheiros: há pia com altura entre 78 e 80 cm e com vão inferior livre com altura de 73 cm?	A	1
		B	0
D100	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: sala de recursos multifuncionais para atendimento educacional especializado (aee)?	A	1
		B	0
D101	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: materiais didáticos e paradidáticos acessíveis? (braille, caracteres ampliados, libras, texturas, contrastes etc).	A	1
		B	0
D102	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: recursos ópticos? (ex: lupa, telêlupa, telescópio, lentes especiais etc.)	A	1
		B	0
D103	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: recursos não ópticos? (ex: celofane amarelo, livro falado, material impresso com tipos maiores, tábua de apoio para leitura, lápis e canetas especiais, guia de leitura etc.)	A	1
		B	0

Questão	Questão	Alternativa	Código
D104	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: soroban?	A B	1 0
D105	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: reglete e punção?	A B	1 0
D106	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: calculadora sonora?	A B	1 0
D107	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: calculadora para pessoas de baixa visão?	A B	1 0
D108	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: máquina perkins? (máquina de escrever em braile)	A B	1 0
D109	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: computador com software leitor de tela?	A B	1 0
D110	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: software de comunicação alternativa aumentativa (caa)?	A B	1 0
D111	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: mesas e cadeiras acessíveis para pessoas em cadeiras de rodas?	A B	1 0
D112	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: quadro negro (ou branco) sem degraus com altura que permita seu alcance por pessoa em cadeira de rodas?	A B	1 0
D113	As perguntas de 69 a 113 referem-se às condições de acessibilidade. A escola possui: bebedouros que permitem a aproximação frontal de pessoas em cadeiras de rodas?	A B	1 0
D114	Que atividades extracurriculares são desenvolvidas regularmente com os alunos nesta escola? Esportivas.	A B	1 0
D115	Que atividades extracurriculares são desenvolvidas regularmente com os alunos nesta escola? Artísticas (música, teatro, trabalhos artesanais).	A B	1 0
D116	Os espaços desta escola são utilizados para atividades comunitárias?	A B C D	2 3 4 1
D117	Neste ano, ocorreram nesta escola eventos da comunidade usando instalações, equipamentos ou recursos da escola?	A B	1 0
D118	Neste ano, ocorreram nesta escola eventos de terceiros realizados na escola e abertos para a comunidade (shows, teatro, palestras)?	A B	1 0
D119	Neste ano, ocorreram nesta escola eventos da escola e destinados à comunidade externa (cursos, práticas esportivas, palestras)?	A B	1 0
D120	Neste ano, ocorreram nesta escola campanhas de solidariedade promovidas pela escola?	A B	1 0
D121	Neste ano, ocorreram nesta escola campanhas de solidariedade propostas pela comunidade, envolvendo a escola?	A B	1 0
D122	Neste ano, ocorreram nesta escola comunidade colaborando na manutenção de hortas, pomar, jardins?	A B	1 0
D123	Neste ano, ocorreram nesta escola comunidade participando em mutirão para limpeza da escola?	A B	1 0
D124	Neste ano, ocorreram nesta escola comunidade participando em mutirão para manutenção da estrutura física da escola?	A B	1 0
D125	Esta escola recebe apoio financeiro de programa de financiamento do governo federal?	A B C	3 2 1
D126	Esta escola recebe apoio financeiro de programa de financiamento do governo estadual?	A B C	3 2 1
D127	Esta escola recebe apoio financeiro de programa de financiamento do governo municipal?	A B C	3 2 1
D128	Quem escolheu os livros didáticos utilizados nesta escola?	A B C D E F	6 5 4 3 2 1
D129	Neste ano, ocorreram as seguintes situações: os livros chegaram a tempo hábil para o início das aulas.	A B C	3 2 1
D130	Neste ano, ocorreram as seguintes situações: faltaram livros para os alunos.	A B C	3 2 1
D131	Neste ano, ocorreram as seguintes situações: sobraram livros.	A B C	3 2 1
D132	Neste ano, ocorreram as seguintes situações: os livros escolhidos foram os recebidos.	A B C	3 2 1

Questão	Questão	Alternativa	Código
D133	Neste ano, aconteceram os seguintes fatos nesta escola: atentado à vida de professores ou funcionários dentro da escola. Agente causador externo (estranho a escola)	A B	1 0
D134	Neste ano, aconteceram os seguintes fatos nesta escola: atentado à vida de alunos dentro da escola. Agente causador externo (estranho a escola)	A B	1 0
D135	Neste ano, aconteceram os seguintes fatos nesta escola: furto a professores ou funcionários dentro da escola. Agente causador externo (estranho a escola)	A B	1 0
D136	Neste ano, aconteceram os seguintes fatos nesta escola: furto a alunos dentro da escola. Agente causador externo (estranho a escola)	A B	1 0
D137	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) a professores ou funcionários dentro da escola. Agente causador externo (estranho a escola)	A B	1 0
D138	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) a alunos dentro da escola. Agente causador externo (estranho a escola)	A B	1 0
D139	Neste ano, aconteceram os seguintes fatos nesta escola: furto de equipamentos e materiais didáticos ou pedagógicos da escola. Agente causador externo (estranho a escola)	A B	1 0
D140	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) de equipamentos e materiais didáticos ou pedagógicos da escola. Agente causador externo (estranho a escola)	A B	1 0
D141	Neste ano, aconteceram os seguintes fatos nesta escola: quebra intencional de equipamento. Agente causador externo (estranho a escola)	A B	1 0
D142	Neste ano, aconteceram os seguintes fatos nesta escola: pichação de muros ou paredes das dependências externas da escola. Agente causador externo (estranho a escola)	A B	1 0
D143	Neste ano, aconteceram os seguintes fatos nesta escola: depredação das dependências externas da escola. Agente causador externo (estranho a escola)	A B	1 0
D144	Neste ano, aconteceram os seguintes fatos nesta escola: sujeira nas dependências externas da escola. Agente causador externo (estranho a escola)	A B	1 0
D145	Neste ano, aconteceram os seguintes fatos nesta escola: sujeira nas dependências internas da escola. Agente causador externo (estranho a escola)	A B	1 0
D146	Neste ano, aconteceram os seguintes fatos nesta escola: pichação de muros ou paredes das dependências internas da escola. Agente causador externo (estranho a escola)	A B	1 0
D147	Neste ano, aconteceram os seguintes fatos nesta escola: depredação das dependências internas da escola. Agente causador externo (estranho a escola)	A B	1 0
D148	Neste ano, aconteceram os seguintes fatos nesta escola: depredação de banheiros. Agente causador externo (estranho a escola)	A B	1 0
D149	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de bebidas alcoólicas nas dependências da escola. Agente causador externo (estranho a escola)	A B	1 0
D150	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de drogas nas dependências da escola. Agente causador externo (estranho a escola)	A B	1 0
D151	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de drogas nas proximidades da escola. Agente causador externo (estranho a escola)	A B	1 0
D152	Neste ano, aconteceram os seguintes fatos nesta escola: tráfico de drogas nas dependências da escola. Agente causador externo (estranho a escola)	A B	1 0
D153	Neste ano, aconteceram os seguintes fatos nesta escola: tráfico de drogas nas proximidades da escola. Agente causador externo (estranho a escola)	A B	1 0
D154	Neste ano, aconteceram os seguintes fatos nesta escola: atentado à vida de professores ou funcionários dentro da escola. Agente causador interno (da própria a escola)	A B	1 0
D155	Neste ano, aconteceram os seguintes fatos nesta escola: atentado à vida de alunos dentro da escola. Agente causador interno (da própria a escola)	A B	1 0
D156	Neste ano, aconteceram os seguintes fatos nesta escola: furto a professores ou funcionários dentro da escola. Agente causador interno (da própria a escola)	A B	1 0
D157	Neste ano, aconteceram os seguintes fatos nesta escola: furto a alunos dentro da escola. Agente causador interno (da própria a escola)	A B	1 0
D158	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) a professores ou funcionários dentro da escola. Agente causador interno (da própria a escola)	A B	1 0
D159	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) a alunos dentro da escola. Agente causador interno (da própria a escola)	A B	1 0
D160	Neste ano, aconteceram os seguintes fatos nesta escola: furto de equipamentos e materiais didáticos ou pedagógicos da escola. Agente causador interno (da própria a escola)	A B	1 0
D161	Neste ano, aconteceram os seguintes fatos nesta escola: roubo (com uso de violência) de equipamentos e materiais didáticos ou pedagógicos da escola. Agente causador interno (da própria a escola)	A B	1 0
D162	Neste ano, aconteceram os seguintes fatos nesta escola: quebra intencional de equipamento. Agente causador interno (da própria a escola)	A B	1 0

Questão	Questão	Alternativa	Código
D163	Neste ano, aconteceram os seguintes fatos nesta escola: pichação de muros ou paredes das dependências externas da escola. Agente causador interno (da própria a escola)	A B	1 0
D164	Neste ano, aconteceram os seguintes fatos nesta escola: depredação das dependências externas da escola. Agente causador interno (da própria a escola)	A B	1 0
D165	Neste ano, aconteceram os seguintes fatos nesta escola: sujeira nas dependências externas da escola. Agente causador interno (da própria a escola)	A B	1 0
D166	Neste ano, aconteceram os seguintes fatos nesta escola: sujeira nas dependências internas da escola. Agente causador interno (da própria a escola)	A B	1 0
D167	Neste ano, aconteceram os seguintes fatos nesta escola: pichação de muros ou paredes das dependências internas da escola. Agente causador interno (da própria a escola)	A B	1 0
D168	Neste ano, aconteceram os seguintes fatos nesta escola: depredação das dependências internas da escola. Agente causador interno (da própria a escola)	A B	1 0
D169	Neste ano, aconteceram os seguintes fatos nesta escola: depredação de banheiros. Agente causador interno (da própria a escola)	A B	1 0
D170	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de bebidas alcoólicas nas dependências da escola. Agente causador interno (da própria a escola)	A B	1 0
D171	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de drogas nas dependências da escola. Agente causador interno (da própria a escola)	A B	1 0
D172	Neste ano, aconteceram os seguintes fatos nesta escola: consumo de drogas nas proximidades da escola. Agente causador interno (da própria a escola)	A B	1 0
D173	Neste ano, aconteceram os seguintes fatos nesta escola: tráfico de drogas nas dependências da escola. Agente causador interno (da própria a escola)	A B	1 0
D174	Neste ano, aconteceram os seguintes fatos nesta escola: tráfico de drogas nas proximidades da escola. Agente causador interno (da própria a escola)	A B	1 0
D175	Neste ano, os seguintes eventos fizeram ou não parte do cotidiano desta escola: membros da comunidade escolar portando arma de fogo.	A B	1 0
D176	Neste ano, os seguintes eventos fizeram ou não parte do cotidiano desta escola: membros da comunidade escolar portando arma branca (faca, canivete, estilete etc.).	A B	1 0
D177	Neste ano, os seguintes eventos fizeram ou não parte do cotidiano desta escola: ação de gangues nas dependências externas da escola.	A B	1 0
D178	Neste ano, os seguintes eventos fizeram ou não parte do cotidiano desta escola: ação de gangues nas dependências internas da escola.	A B	1 0
D179	Neste ano, houve agressão verbal a professores. Quem foi o agressor? Aluno	A B	1 0
D180	Neste ano, houve agressão física a professores. Quem foi o agressor? Aluno	A B	1 0
D181	Neste ano, houve agressão verbal a alunos. Quem foi o agressor? Aluno	A B	1 0
D182	Neste ano, houve agressão física a alunos. Quem foi o agressor? Aluno	A B	1 0
D183	Neste ano, houve agressão verbal a funcionários. Quem foi o agressor? Aluno	A B	1 0
D184	Neste ano, houve agressão física a funcionários. Quem foi o agressor? Aluno	A B	1 0
D185	Neste ano, houve agressão verbal a professores. Quem foi o agressor? Professor	A B	1 0
D186	Neste ano, houve agressão física a professores. Quem foi o agressor? Professor	A B	1 0
D187	Neste ano, houve agressão verbal a alunos. Quem foi o agressor? Professor	A B	1 0
D188	Neste ano, houve agressão física a alunos. Quem foi o agressor? Professor	A B	1 0
D189	Neste ano, houve agressão verbal a funcionários. Quem foi o agressor? Professor	A B	1 0
D190	Neste ano, houve agressão física a funcionários. Quem foi o agressor? Professor	A B	1 0
D191	Neste ano, houve agressão verbal a professores. Quem foi o agressor? Funcionário	A B	1 0
D192	Neste ano, houve agressão física a professores. Quem foi o agressor? Funcionário	A B	1 0
D193	Neste ano, houve agressão verbal a alunos. Quem foi o agressor? Funcionário	A B	1 0
D194	Neste ano, houve agressão física a alunos. Quem foi o agressor? Funcionário	A B	1 0
D195	Neste ano, houve agressão verbal a funcionários. Quem foi o agressor? Funcionário	A B	1 0
D196	Neste ano, houve agressão física a funcionários. Quem foi o agressor? Funcionário	A B	1 0

Questão	Questão	Alternativa	Código
D197	Nesta escola há projetos nas seguintes temáticas violência.	A	1
		B	0
D198	Nesta escola há projetos nas seguintes temáticas uso abusivo de drogas.	A	1
		B	0
D199	Nesta escola há projetos nas seguintes temáticas racismo.	A	1
		B	0
D200	Nesta escola há projetos nas seguintes temáticas machismo e homofobia.	A	1
		B	0
D201	Nesta escola há projetos nas seguintes temáticas bullying.	A	1
		B	0
D202	Nesta escola há projetos nas seguintes temáticas sexualidade e gravidez na adolescência.	A	1
		B	0
D203	Nesta escola há projetos nas seguintes temáticas desigualdades sociais.	A	1
		B	0
D204	Nesta escola há projetos nas seguintes temáticas conflitos religiosos.	A	1
		B	0
D205	Nesta escola há projetos nas seguintes temáticas meio ambiente.	A	1
		B	0
D206	Nesta escola há costume de se fazer oração ou cantar músicas religiosas (na entrada, antes do lanche etc).	A	1
		B	0
D207	Nesta escola há objetos, imagens, frases ou símbolos de teor religioso.	A	1
		B	0
D208	Nesta escola há aula de ensino religioso.	A	1
		B	0
D209	Nesta escola há atividades alternativas para os estudantes que não queiram participar das aulas de ensino religioso.	A	1
		B	0
D210	Nesta escola, o ensino religioso (se não houver ensino religioso deixe as próximas questões em branco): é de presença obrigatória.	A	1
		B	0
D211	Nesta escola, o ensino religioso (se não houver ensino religioso deixe as próximas questões em branco): é de caráter confessional (segundo uma religião específica).	A	1
		B	0
D212	Nesta escola, o ensino religioso (se não houver ensino religioso deixe as próximas questões em branco): contempla a diversidade religiosa.	A	1
		B	0
E001	Avalie o estado de conservação dos itens e equipamentos do prédio - telhado	A	4
		B	3
		C	2
		D	1
E002	Avalie o estado de conservação dos itens e equipamentos do prédio - paredes	A	4
		B	3
		C	2
		D	1
E003	Avalie o estado de conservação dos itens e equipamentos do prédio - piso	A	4
		B	3
		C	2
		D	1
E004	Avalie o estado de conservação dos itens e equipamentos do prédio - entrada do prédio	A	4
		B	3
		C	2
		D	1
E005	Avalie o estado de conservação dos itens e equipamentos do prédio - pátio	A	4
		B	3
		C	2
		D	1
E006	Avalie o estado de conservação dos itens e equipamentos do prédio - corredores	A	4
		B	3
		C	2
		D	1
E007	Avalie o estado de conservação dos itens e equipamentos do prédio - salas de aula	A	4
		B	3
		C	2
		D	1
E008	Avalie o estado de conservação dos itens e equipamentos do prédio - portas	A	4
		B	3
		C	2
		D	1
E009	Avalie o estado de conservação dos itens e equipamentos do prédio - janelas	A	4
		B	3
		C	2
		D	1
E010	Avalie o estado de conservação dos itens e equipamentos do prédio - banheiros	A	4
		B	3
		C	2
		D	1

Questão	Questão	Alternativa	Código
E011	Avalie o estado de conservação dos itens e equipamentos do prédio - cozinha	A	4
		B	3
		C	2
		D	1
E012	Avalie o estado de conservação dos itens e equipamentos do prédio - instalações hidráulicas	A	4
		B	3
		C	2
		D	1
E013	Avalie o estado de conservação dos itens e equipamentos do prédio - instalações elétricas	A	4
		B	3
		C	2
		D	1
E014	Avalie os seguintes aspectos em relação às salas de aula: (considere sim quando mais de 50% das salas de aula apresentarem o aspecto questionado) - são iluminadas?	A	1
E015	Avalie os seguintes aspectos em relação às salas de aula: (considere sim quando mais de 50% das salas de aula apresentarem o aspecto questionado) - são arejadas?	A	1
		B	0
E016	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há muros, grades ou cercas em condições de garantir a segurança dos alunos? (caso existam buracos ou aberturas que permitam o acesso de estranhos, responder não)	A	1
		B	0
E017	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há controle de entrada e saída de alunos?	A	1
		B	0
E018	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há controle de entrada de pessoas estranhas na escola?	A	1
		B	0
E019	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: os portões que dão acesso à parte externa permanecem trancados durante o horário de funcionamento da escola?	A	1
		B	0
E020	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum tipo de vigilância para o período diurno?	A	1
		B	0
E021	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum tipo de vigilância para o período noturno?	A	1
		B	0
E022	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum tipo de vigilância para os finais de semana e feriados?	A	1
		B	0
E023	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum esquema de policiamento para inibição de furtos, roubos e outras formas de violência?	A	1
		B	0
E024	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum esquema de policiamento para inibição de tráfico de tóxicos/ drogas dentro da escola?	A	1
		B	0
E025	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há algum esquema de policiamento para inibição de tráfico de tóxicos/ drogas nas imediações da escola?	A	1
		B	0
E026	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: a escola possui algum sistema de proteção contra incêndio (alarme de fumaça e temperatura, extintores contra incêndio dentro do prazo de validade, mangueira, etc.)?	A	1
		B	0
E027	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: as salas onde são guardados os equipamentos mais caros (computadores, projetores, televisão, vídeo etc.) Possuem dispositivos para serem trancados(cadeados, grades, travas, trancas etc.)?	A	1
		B	0
E028	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: a escola apresenta sinais de depredação (vidros, portas e janelas quebradas, lâmpadas estouradas etc.)?	A	1
		B	0
E029	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: há uma boa iluminação do lado de fora da escola?	A	1
		B	0
E030	Avalie os seguintes aspectos em relação à segurança da escola e dos alunos: a escola adota alguma medida de segurança para proteger os alunos nas suas imediações?	A	1
		B	0
E031	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Computadores para uso dos alunos.	A	4
		B	3
		C	2
		D	1
E032	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Acesso à internet para uso dos alunos.	A	4
		B	3
		C	2
		D	1
E033	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Computadores para uso dos professores.	A	4
		B	3
		C	2
		D	1
E034	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Acesso à internet para uso dos professores.	A	4
		B	3
		C	2
		D	1
E035	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Computadores exclusivamente para o uso administrativo.	A	4
		B	3
		C	2
		D	1

Questão	Questão	Alternativa	Código
E036	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Fitas de vídeo ou dvd (educativas).	A B C D	4 3 2 1
E037	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Fitas de vídeo ou dvd (lazer).	A B C D	4 3 2 1
E038	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Máquina copiadora.	A B C D	4 3 2 1
E039	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Impressora.	A B C D	4 3 2 1
E040	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Retroprojektor.	A B C D	4 3 2 1
E041	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Projetor de slides.	A B C D	4 3 2 1
E042	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Videocassete ou dvd.	A B C D	4 3 2 1
E043	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Televisão.	A B C D	4 3 2 1
E044	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Mimeógrafo.	A B C D	4 3 2 1
E045	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Câmera fotográfica.	A B C D	4 3 2 1
E046	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Antena parabólica.	A B C D	4 3 2 1
E047	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Linha telefônica.	A B C D	4 3 2 1
E048	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Aparelho de fax.	A B C D	4 3 2 1
E049	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Aparelho de som.	A B C D	4 3 2 1
E050	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Biblioteca.	A B C D	4 3 2 1
E051	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Quadra de esportes.	A B C D	4 3 2 1
E052	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Laboratório.	A B C D	4 3 2 1
E053	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Auditório.	A B C D	4 3 2 1
E054	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Sala para atividades de música.	A B C D	4 3 2 1

Questão	Questão	Alternativa	Código
E055	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Sala para atividades de artes plásticas.	A	4
		B	3
		C	2
		D	1
E056	Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Sala de leitura.	A	4
		B	3
		C	2
		D	1
E057	Em relação à biblioteca ou sala de leitura: possui acervo diversificado que desperte o interesse dos alunos.	A	1
		B	0
E058	Em relação à biblioteca ou sala de leitura: possui brinquedoteca.	A	1
		B	0
E059	Em relação à biblioteca ou sala de leitura: possui espaço para estudos coletivos.	A	1
		B	0
E060	Em relação à biblioteca ou sala de leitura: os livros podem ser manuseados e emprestados.	A	1
		B	0
E061	Em relação à biblioteca ou sala de leitura: a comunidade pode utilizar o espaço e os livros.	A	1
		B	0
E062	Em relação à biblioteca ou sala de leitura: o espaço é arejado e bem iluminado.	A	1
		B	0
E063	Em relação à biblioteca ou sala de leitura: existe uma pessoa responsável pelo atendimento na biblioteca ou na sala de leitura?	A	1
		B	0
E064	Os usuários da biblioteca (ou sala de leitura) levam livros para casa? Os alunos.	A	3
		B	2
		C	1
E065	Os usuários da biblioteca (ou sala de leitura) levam livros para casa? Os professores.	A	3
		B	2
		C	1
E066	Os usuários da biblioteca (ou sala de leitura) levam livros para casa? Os membros da comunidade.	A	3
		B	2
		C	1
P001	Sexo do Professor	A	0
		B	1
P002	Idade do Professor	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
P004	Qual o seu nível de escolaridade (Até a Graduação). Caso Marque A, B ou C, passe para questão 11	A	1
		B	2
		C	3
		D	6
		E	8
		F	7
		G	5
		H	4
P005	Há quantos anos você obteve o nível de escolaridade assinalado anteriormente?	A	1
		B	2
		C	3
		D	4
		E	5
P006	Em que tipo de instituição você fez o curso superior? Se você estudou em mais de uma instituição, assinale aquela em que obteve o seu título profissional.	A	5
		B	4
		C	3
		D	2
		E	1
P007	Qual era a natureza dessa instituição?	A	3
		B	2
		C	4
		D	1
P008	De que forma você realizou o curso superior?	A	4
		B	3
		C	2
		D	1
P009	Indique a modalidade de cursos de pós-graduação de mais alta titulação que você possui	A	2
		B	3
		C	4
		D	5
		E	1
P010	Indique qual a área temática do curso de pós-graduação de mais alta titulação que você possui.	A	5
		B	4
		C	3
		D	2
		E	1

Questão	Questão	Alternativa	Código
P011	Você participou de alguma atividade de formação continuada (atualização, treinamento, capacitação, etc) nos últimos dois anos?	A	1
		B	0
P012	Qual a carga horária da atividade considerada mais relevante dentre as quais você participou?	A	1
		B	2
		C	3
		D	4
P013	Você utiliza os conhecimentos adquiridos nas atividades de formação continuada para a melhoria de sua prática em sala de aula?	A	4
		B	3
		C	2
		D	1
P014	Nesta escola, qual, aproximadamente, o seu salário bruto? (com adicionais, se houver) (preencha os espaços com o valor e marque, na folha de respostas, os campos correspondentes a cada algarismo.)	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
		I	9
		J	10
		K	11
P015	Além da atividade como docente nesta escola, você exerce outra atividade que contribui para sua renda pessoal?	A	3
		B	2
		C	1
P016	Qual é, aproximadamente, seu salário bruto (com adicionais, se houver) como professor(a)? (soma de tudo o que você ganha como professor(a)) (preencha os espaços com o valor e marque, na folha de respostas, os campos correspondentes a cada algarismo.)	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
		I	9
		J	10
		K	11
P017	Há quantos anos você leciona?	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
P018	Há quantos anos você trabalha nesta escola?	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
P019	Há quantos anos você ministra aulas para alunos das série desta turma?	A	1
		B	2
		C	3
		D	4
		E	5
P020	Nesta escola, qual a sua carga horária semanal? (considere a carga horária contratual: horas-aula mais horas para atividades, se houver.)	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
		I	9
P021	Em quantas escolas você trabalha? (caso marque a alternativa a, passe para a questão 23)	A	1
		B	2
		C	3
		D	4

Questão	Questão	Alternativa	Código
P022	Ao todo, quantas horas-aula você ministra por semana? (não considere aulas particulares) atividades, se houver.)	A	1
		B	2
		C	3
		D	4
		E	5
		F	6
		G	7
		H	8
		I	9
P023	Qual é a sua situação trabalhista nesta escola? (marque apenas uma opção)	A	5
		B	4
		C	3
		D	2
		E	1
P024	Você costuma: frequentar bibliotecas.	A	4
		B	3
		C	2
		D	1
P025	Você costuma: ir ao cinema.	A	4
		B	3
		C	2
		D	1
P026	Você costuma: ir ao museu.	A	4
		B	3
		C	2
		D	1
P027	Você costuma: ver apresentações teatrais.	A	4
		B	3
		C	2
		D	1
P028	Você costuma: ver apresentações musicais ou de dança	A	4
		B	3
		C	2
		D	1
P029	Em seu tempo livre, você lê: jornais.	A	3
		B	2
		C	1
P030	Em seu tempo livre, você lê: revistas de informação geral.	A	3
		B	2
		C	1
P031	Em seu tempo livre, você lê: revistas de humor /quadrinhos.	A	3
		B	2
		C	1
P032	Em seu tempo livre, você lê: revistas de divulgação científica/cultural.	A	3
		B	2
		C	1
P033	Em seu tempo livre, você lê: revistas sobre comportamento, celebridades, esporte e tv .	A	3
		B	2
		C	1
P034	Em seu tempo livre, você lê: livros.	A	3
		B	2
		C	1
P035	Em seu tempo livre, você lê: sites da internet.	A	3
		B	2
		C	1
P036	Dentro da sua carga horária, quantas horas semanais são dedicadas à atividades extra-classe (formação e estudo, planejamento, produção de recursos didáticos etc)?	A	1
		B	2
		C	3
P037	Indique se você utiliza ou não nesta escola: jornais e revistas informativas.	A	3
		B	2
		C	1
P038	Indique se você utiliza ou não nesta escola: livros de literatura em geral.	A	3
		B	2
		C	1
P039	Indique se você utiliza ou não nesta escola: projetor de slides.	A	3
		B	2
		C	1
P040	Indique se você utiliza ou não nesta escola: retroprojetor.	A	3
		B	2
		C	1
P041	Indique se você utiliza ou não nesta escola: máquina copiadora.	A	3
		B	2
		C	1

Questão	Questão	Alternativa	Código
P043	Conselho de classe é um órgão formado por todos os professores que lecionam em cada turma/série. Neste ano, quantas vezes se reuniram os conselhos de classe desta escola?	A B C D E	3 4 5 2 1
P044	Você tem conhecimento do conteúdo da lei número 11.645 de 2008 que determina a obrigatoriedade do estudo da temática História e cultura afro-brasileira e indígena nos estabelecimentos de ensino do país?	A B	1 0
P045	Neste ano, foram desenvolvidas atividades para atender o determinado pela lei n.º 11.645 de 2008 nesta escola?	A B C	3 2 1
P046	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem na escola devido à carência de infraestrutura física e/ou pedagógica.	A B	1 0
P047	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): são decorrentes do ambiente de insegurança física da escola.	A B	1 0
P048	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem na escola por oferecer poucas oportunidades de desenvolvimento das capacidades intelectuais do aluno	A B	1 0
P049	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): estão relacionados aos conteúdos curriculares, que são inadequados às necessidades dos alunos.	A B	1 0
P050	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): estão relacionados ao não-cumprimento do conteúdo curricular.	A B	1 0
P051	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): relacionam-se à sobrecarga de trabalho do(as) professores(as), dificultando o planejamento e o preparo das aulas.	A B	1 0
P052	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem devido ao baixo salário dos professores, que gera insatisfação e desestímulo para a atividade docente.	A B	1 0
P053	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): são decorrentes do meio em que o aluno vive.	A B	1 0
P054	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): são decorrentes do nível cultural dos pais dos alunos.	A B	1 0
P055	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): estão relacionadas à falta de assistência e acompanhamento da família nos deveres de casa e pesquisas dos alunos.	A B	1 0
P056	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem devido à falta de aptidão e habilidades do aluno.	A B	1 0
P057	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): estão vinculados à baixa autoestima dos alunos.	A B	1 0
P058	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem devido ao desinteresse e falta de esforço do aluno.	A B	1 0
P059	Assinale sua posição em relação às afirmações abaixo, que se referem aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): são decorrentes da indisciplina dos alunos em sala de aula.	A B	1 0
P060	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) me anima e me motiva para o trabalho.	A B C D E	5 4 3 2 1
P061	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): tenho plena confiança no(a) diretor(a) como profissional.	A B C D E	5 4 3 2 1
P062	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) consegue que os professores se comprometam com a escola.	A B C D E	5 4 3 2 1
P063	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) estimula as atividades inovadoras.	A B C D E	5 4 3 2 1

Questão	Questão	Alternativa	Código
P064	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) dá atenção especial a aspectos relacionados com a aprendizagem dos alunos.	A	5
		B	4
		C	3
		D	2
		E	1
P065	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) dá atenção especial a aspectos relacionados com as normas administrativas.	A	5
		B	4
		C	3
		D	2
		E	1
P066	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o(a) diretor(a) dá atenção especial a aspectos relacionados com a manutenção da escola.	A	5
		B	4
		C	3
		D	2
		E	1
P067	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): sinto-me respeitado(a) pelo(a) diretor(a).	A	5
		B	4
		C	3
		D	2
		E	1
P068	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): respeito o(a) diretor(a).	A	5
		B	4
		C	3
		D	2
		E	1
P069	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): participo das decisões relacionadas com o meu trabalho.	A	5
		B	4
		C	3
		D	2
		E	1
P070	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): a equipe de professores leva em consideração minhas ideias.	A	5
		B	4
		C	3
		D	2
		E	1
P071	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): eu levo em consideração as ideias de outros colegas.	A	5
		B	4
		C	3
		D	2
		E	1
P072	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o ensino que a escola oferece aos alunos é muito influenciado pela troca de ideias entre os professores.	A	5
		B	4
		C	3
		D	2
		E	1
P073	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): os professores desta escola procuram coordenar o conteúdo das disciplinas entre as diferentes séries.	A	5
		B	4
		C	3
		D	2
		E	1
P074	Os itens de 60 a 74 apresentam algumas afirmações. Indique seu grau de concordância/discordância com cada uma delas. (marque apenas uma opção em cada linha.): o diretor, os professores e os demais membros da equipe da escola colaboram para fazer esta escola funcionar bem.	A	5
		B	4
		C	3
		D	2
		E	1
P075	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: insuficiência de recursos financeiros?	A	3
		B	2
		C	1
P076	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: inexistência de professores para algumas disciplinas ou séries?	A	3
		B	2
		C	1
P077	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: carência de pessoal administrativo?	A	3
		B	2
		C	1
P078	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: carência de pessoal de apoio pedagógico (coordenador, supervisor, orientador educacional)?	A	3
		B	2
		C	1
P079	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: falta de recursos pedagógicos?	A	3
		B	2
		C	1

Questão	Questão	Alternativa	Código
P080	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: alto índice de faltas por parte dos professores?	A B C	3 2 1
P081	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: alto índice de faltas por parte dos alunos?	A B C	3 2 1
P082	As perguntas de 75 a 82 apresentam alguns problemas que podem ocorrer nas escolas. Responda se cada um deles ocorreu ou não neste ano. Caso tenha ocorrido, assinale se foi ou não um problema grave, dificultando o funcionamento da escola. (marque apenas uma opção em cada linha.) Ocorreu na escola: problemas disciplinares causados pelos alunos?	A B C	3 2 1
P083	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno) : agressão verbal a professores.	A B	1 0
P084	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno) : agressão física a professores.	A B	1 0
P085	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno) : agressão verbal a alunos.	A B	1 0
P086	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno) : agressão física a alunos.	A B	1 0
P087	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno) : agressão verbal a funcionários.	A B	1 0
P088	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: aluno) : agressão física a funcionários.	A B	1 0
P089	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor):agressão verbal a professores.	A B	1 0
P090	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor):agressão física a professores.	A B	1 0
P091	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor):agressão verbal a alunos.	A B	1 0
P092	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor):agressão física a alunos.	A B	1 0
P093	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor): agressão verbal a funcionários.	A B	1 0
P094	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: professor): agressão física a funcionários.	A B	1 0
P095	Agressão verbal a professores. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	A B	1 0
P096	Agressão física a professores. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	A B	1 0
P097	Agressão verbal a alunos. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	A B	1 0
P098	Agressão física a alunos. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	A B	1 0
P099	Agressão verbal a funcionários. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	A B	1 0
P100	Agressão física a funcionários. Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano nesta escola. (agressor: funcionário):	A B	1 0
P101	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi vítima de atentado à vida?	A B	1 0
P102	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi ameaçado por algum aluno?	A B	1 0
P103	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi agredido verbalmente por algum aluno?	A B	1 0
P104	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi agredido fisicamente por algum aluno?	A B	1 0
P105	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi vítima de furto?	A B	1 0
P106	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: você foi vítima de roubo (com uso de violência)?	A B	1 0
P107	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: alunos frequentaram as suas aulas sob efeito de bebida alcoólica?	A B	1 0
P108	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: alunos frequentaram as suas aulas sob efeito de drogas ilícitas?	A B	1 0
P109	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: alunos frequentaram as suas aulas portando arma branca (facas, canivetes etc.)?	A B	1 0
P110	Sobre os fatos listados abaixo, diga se eles aconteceram ou não este ano, nesta escola: alunos frequentaram as suas aulas portando arma de fogo?	A B	1 0
P111	A sua escola participou da Prova Brasil de 2009?	A B	1 0
P112	Você conhece os resultados do sistema de avaliação da educação básica (Saeb): da sua escola?	A B	1 0
P113	Você conhece os resultados do sistema de avaliação da educação básica (Saeb): do seu município?	A B	1 0

Questão	Questão	Alternativa	Código
P114	Você conhece os resultados do sistema de avaliação da educação básica (Saeb): do seu estado?	A	1
		B	0
P115	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): computadores.	A	3
		B	2
		C	1
P116	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): internet.	A	3
		B	2
		C	1
P117	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): fitas de vídeo ou dvd.	A	3
		B	2
		C	1
P118	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): jornais e revistas informativas.	A	3
		B	2
		C	1
P119	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): revistas em quadrinhos.	A	3
		B	2
		C	1
P120	Indique se você utiliza ou não nesta turma: (marque apenas uma opção em cada linha.): livros didáticos.	A	3
		B	2
		C	1
P121	Quanto dos conteúdos previstos você conseguiu desenvolver com os alunos desta turma, neste ano?	A	1
		B	2
		C	3
		D	4
P122	Quanto dos alunos desta turma você acha que: concluirão os anos iniciais do ensino fundamental?	A	5
		B	4
		C	3
		D	2
		E	1
P123	Quanto dos alunos desta turma você acha que: concluirão o ensino fundamental (8. ^a série/9. ^o ano)?	A	5
		B	4
		C	3
		D	2
		E	1
P124	Quanto dos alunos desta turma você acha que: concluirão o ensino médio?	A	5
		B	4
		C	3
		D	2
		E	1
P125	Quanto dos alunos desta turma você acha que: entrarão para a universidade?	A	5
		B	4
		C	3
		D	2
		E	1
P126	Os alunos desta turma têm livros didáticos?	A	5
		B	4
		C	3
		D	2
		E	1
P127	Os alunos desta turma receberam o livro didático no início do ano letivo?	A	1
		B	0
P128	O livro didático escolhido foi o recebido?	A	1
		B	0
P129	Como você considera o(s) livro(s) didático(s) utilizado(s) por você na(s) disciplina(s) que ministra nesta turma?	A	5
		B	4
		C	3
		D	2
		E	1
P130	Na escolha do livro didático utilizado nesta turma em que você leciona, foi consultado o guia de livros didáticos da SEB/MEC?	A	1
		B	0
P131	Para a disciplina que você ministra, como foi escolhido o livro didático para utilização nesta turma? (marque apenas uma opção.)	A	7
		B	6
		C	5
		D	4
		E	3
		F	2
		G	1
P132	Como você utiliza a biblioteca ou sala de leitura da escola em apoio às suas aulas nesta turma?	A	1
		B	2
		C	3
		D	4
		E	5

Questão	Questão	Alternativa	Código
P142	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: fazer exercícios para fixar procedimentos e regras.	A B C D E	5 4 3 2 1
P143	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: lidar com situações problemas que exigem raciocínios diferentes e mais complexos que a maioria dos exemplos usuais.	A B C D E	5 4 3 2 1
P144	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: falar sobre suas soluções, discutindo os caminhos usados para encontrá-las.	A B C D E	5 4 3 2 1
P145	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: gravar as regras que permitem obter as respostas certas dos cálculos e problemas.	A B C D E	5 4 3 2 1
P146	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: lidar com temas que aparecem em jornais e/ou revistas, discutindo a relação dos temas com a matemática.	A B C D E	5 4 3 2 1
P147	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema.	A B C D E	5 4 3 2 1
P148	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: lidar com situações que lhes sejam familiares e que apresentem temas de interesse dos alunos.	A B C D E	5 4 3 2 1
P149	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: experimentar diferentes modos de resolver um problema ou de efetuar um cálculo.	A B C D E	5 4 3 2 1
P150	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: aprimorar a precisão e a velocidade de execução de cálculos.	A B C D E	5 4 3 2 1
P151	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: experimentar diferentes ações (coletar informações, recortar, analisar explorar, discutir, manipular etc.) Para resolver problemas.	A B C D E	5 4 3 2 1
P152	(Professor de matemática) indique a frequência com a qual você desenvolve as seguintes práticas pedagógicas com seus alunos nesta turma: incentivar e estimular o aluno a analisar criticamente se os resultados obtidos na resolução de um problema são plausíveis.	A B C D E	5 4 3 2 1

Tabela 32 – Descrição das variáveis Dicotomizadas da amostra obtida dos dados

Variáveis Dicotomizadas	Categorias	Código	Frequência amostral
Q001	Masculino	0	6105
	Feminino	1	5940
Q004	1996 ou antes	0	1274
	1997 ou depois	1	10857
Q005	Sim	1	11903
	Não	0	161
Q006	Sim	1	10897
	Não	0	1186
Q007	Sim	1	11199
	Não	0	854
Q008	Sim	1	12043
	Não	0	73
Q009	Sim	1	9316
	Não	0	2709
Q010	Sim	1	1951
	Não	0	10099
Q012	Sim	1	7443
	Não	0	4676
Q013	Sim	1	8940
	Não	0	3185
Q014	Sim	1	12059
	Não	0	77
Q015	Sim	1	1056
	Não	0	11063
Q016	Sim	1	11955
	Não	0	71
Q017	Mais de três pessoas	1	9932
	Duas pessoas ou menos	0	2177
Q018	Sim	1	11024
	Não	0	1070
Q019	Ingressou no Ensino Superior ou mais	1	3698
	Terminou o Ensino Médio ou Menos	0	8202
Q020	Sim	1	11395
	Não	0	489
Q022	Sim	1	8010
	Não	0	3997
Q023	Ingressou no Ensino Superior para mais	1	3004
	Concluiu o Ensino Médio ou menos	0	7756
Q024	Sim	1	9843
	Não	0	918
Q026	Alta	1	8263
	Baixa	0	3795
Q032	Alta	1	1008
	Baixa	0	10726
Q033	Alta	1	2559
	Baixa	0	9175
Q034	Alta	1	1808
	Baixa	0	9799
Q035	Alta	1	2900
	Baixa	0	8756
Q036	Alta	1	3069
	Baixa	0	8578
Q037	Alta	1	3744
	Baixa	0	7929
Q038	Alta	1	6462
	Baixa	0	5172
Q039	Alta	1	978
	Baixa	0	10725
Q040	Alta	1	2481
	Baixa	0	9284
Q041	Alta	1	504
	Baixa	0	11129
Q042	Alta	1	1163
	Baixa	0	10506
Q043	Alta	1	2443
	Baixa	0	9304
Q044	Muito Tempo	1	9040
	Pouco Tempo	0	2999
Q045	Muito Tempo	1	4524
	Pouco Tempo	0	7507
Q047	Antes da 1ªSerie	1	1866
	Na 1ªSerie ou depois	0	10144

Variáveis Dicotomizadas	Categorias	Código	Frequência amostral
Q048	Somente em Particular	1	148
	Estudou em escola pública em algum momento	0	11872
Q049	Sim	1	2591
	Não	0	9452
Q050	Sim	1	
	Não	0	
Q052	Alta	1	11282
	Baixa	0	737
Q053	Alta	1	11615
	Baixa	0	415
Q055	Alta	1	10860
	Baixa	0	1182
Q056	Alta	1	11517
	Baixa	0	522
Q057	Sim	1	6104
	Não	0	5951
Q058	Sim	1	3251
	Não	0	8817
E001	Tem	1	605
	Não Tem	0	1
E002	Bom	1	593
	Ruim	0	15
E003	Bom	1	542
	Ruim	0	65
E004	Bom	1	576
	Ruim	0	33
E005	Bom	1	569
	Ruim	0	35
E006	Bom	1	589
	Ruim	0	16
E007	Bom	1	581
	Ruim	0	23
E008	Bom	1	543
	Ruim	0	66
E009	Bom	1	575
	Ruim	0	32
E010	Bom	1	521
	Ruim	0	84
E011	Bom	1	559
	Ruim	0	42
E012	Bom	1	524
	Ruim	0	84
E013	Bom	1	500
	Ruim	0	109
E031	Bom	1	532
	Ruim	0	72
E032	Bom	1	460
	Ruim	0	146
E033	Bom	1	549
	Ruim	0	59
E034	Bom	1	512
	Ruim	0	94
E035	Bom	1	589
	Ruim	0	17
E036	Bom	1	588
	Ruim	0	19
E037	Bom	1	535
	Ruim	0	69
E038	Bom	1	393
	Ruim	0	214
E039	Bom	1	574
	Ruim	0	33
E040	Bom	1	520
	Ruim	0	84
E041	Bom	1	341
	Ruim	0	258
E042	Bom	1	576
	Ruim	0	32
E043	Bom	1	589
	Ruim	0	20
E044	Bom	1	344
	Ruim	0	264
E045	Bom	1	577
	Ruim	0	30

Variáveis Dicotomizadas	Categorias	Código	Frequência amostral
E046	Bom	1	249
	Ruim	0	343
E047	Bom	1	589
	Ruim	0	17
E048	Bom	1	491
	Ruim	0	116
E049	Bom	1	586
	Ruim	0	23
E050	Bom	1	448
	Ruim	0	155
E051	Bom	1	513
	Ruim	0	94
E052	Bom	1	213
	Ruim	0	390
E053	Bom	1	119
	Ruim	0	484
E054	Bom	1	44
	Ruim	0	562
E055	Bom	1	76
	Ruim	0	530
E056	Bom	1	343
	Ruim	0	262
E063	Bom	1	396
	Ruim	0	190
E064	Bom	1	554
	Ruim	0	32
E065	Bom	1	550
	Ruim	0	34
E066	Bom	1	219
	Ruim	0	362
D001	Feminino	1	506
	Masculino	0	96
D002	40 anos ou mais	1	497
	Menos de 40 anos	0	82
D004	Com Ensino Superior	1	596
	Sem Ensino Superior	0	6
D005	15 anos ou mais	1	416
	Menos de 15 anos	0	183
D006	Pública	1	54
	Privada	0	545
D007	Universidade	1	339
	Não era universidade	0	262
D008	Presencial	1	584
	Não - Presencial	0	20
D009	Mestrado ou Doutorado	1	31
	Não tem Mestrado ou Doutorado	0	571
D010	Educação	1	464
	Outras Áreas que não a educação	0	127
D012	21 horas ou mais	1	517
	Menos de 20 horas	0	24
D013	Sim	1	542
	Não	0	0
D014	2 salários mínimos ou mais	1	208
	Menos de 2 salários mínimos	0	228
D015	2 salários mínimos ou mais	1	91
	Menos de 2 salários mínimos	0	365
D016	Sim	1	178
	Não	0	422
D017	10 anos ou mais	1	590
	Menos de 10 anos	0	12
D018	10 anos ou mais	1	236
	Menos de 10 anos	0	363
D019	10 anos ou mais	1	94
	Menos de 10 anos	0	506
D020	Mais de 40 horas	1	136
	40 horas ou menos	0	464
D023	Mais de 50%	1	265
	50% ou menos	0	122
D024	Uma vez ou mais	1	592
	Nenhuma vez	0	7
D029	Uma vez ou mais	1	601
	Nenhuma vez	0	3
D030	Elaborado pelo diretor	1	170
	Elaborado por outros	0	426

Variáveis Dicotomizadas	Categorias	Código	Frequência amostral
D031	Diversos critérios são utilizados	1	432
	Não existe critério de seleção	0	168
D032	Tinha vagas	1	342
	Não tinha vagas	0	262
D033	Homogeneidade e heterogeneidade	1	491
	Não houve critério	0	104
D034	Diversos critérios	1	541
	Não houve critério	0	18
D035	Mais de 50%	1	415
	50% ou menos	0	184
D036	Sim	1	490
	Não	0	112
D037	Sim	1	561
	Não	0	43
D045	Sim	1	557
	Não	0	31
D046	Ruim	1	493
	Bom	0	91
D047	Ruim	1	584
	Bom	0	14
D048	Ruim	1	581
	Bom	0	20
D049	Ruim	1	555
	Bom	0	44
D050	Ruim	1	566
	Bom	0	32
D051	Ruim	1	534
	Bom	0	64
D052	Ruim	1	587
	Bom	0	10
D053	Ruim	1	538
	Bom	0	63
D054	Ruim	1	556
	Bom	0	44
D055	Sim	1	221
	Não	0	381
D056	Sim	1	476
	Não	0	128
D057	Sim	1	409
	Não	0	193
D058	Sim	1	151
	Não	0	452
D059	Sim	1	113
	Não	0	491
D060	Sim	1	42
	Não	0	562
D061	Sim	1	468
	Não	0	135
D062	Sim	1	357
	Não	0	246
D063	Sim	1	288
	Não	0	315
D064	Sim	1	508
	Não	0	95
D084	Sim	1	143
	Não	0	455
D085	Sim	1	291
	Não	0	309
D086	Sim	1	92
	Não	0	509
D116	Sim	1	508
	Não	0	85
D125	Sim	1	521
	Não	0	76
D126	Sim	1	455
	Não	0	143
D127	Sim	1	153
	Não	0	444
D128	Professores	1	487
	Outros	0	105
D129	Sim	1	499
	Não	0	97
D130	Sim	1	260
	Não	0	333

Variáveis Dicotomizadas	Categorias	Código	Frequência amostral
D131	Sim	1	188
	Não	0	404
D132	Sim	1	530
	Não	0	62
P001	Feminino	1	1193
	Masculino	0	481
P002	40 anos ou mais	1	951
	Menos de 40 anos	0	718
P004	Com ensino superior	1	1662
	Sem ensino superior	0	7
P005	15 anos ou mais	1	785
	Menos de 15 anos	0	876
P006	Pública	1	192
	Privada	0	1475
P007	Universidade	1	951
	Não era universidade	0	704
P008	Presencial	1	1640
	Não - Presencial	0	20
P009	Mestrado ou Doutorado	1	63
	Não tem mestrado ou Doutorado	0	1581
P010	Educação	1	760
	Outras áreas que não a educação	0	64
P012	21 horas ou mais	1	1062
	Menos de 20 horas	0	163
P013	Sim	1	1214
	Não	0	14
P014	2 salários mínimos ou mais	1	1239
	Menos de 2 salários mínimos	0	180
P015	Sim	1	652
	Não	0	949
P016	2 salários mínimos ou mais	1	513
	Menos de 2 salários mínimos	0	246
P017	10 anos ou mais	1	1151
	Menos de 10 anos	0	517
P018	10 anos ou mais	1	361
	Menos de 10 anos	0	1305
P019	6 anos ou mais	1	427
	Menos de 6 anos	0	1233
P020	Até 20 horas	1	1468
	Mais de 20 horas	0	192
P021	Uma escola	1	610
	Mais de uma escola	0	1062
P022	Até 20 horas	1	1425
	Mais de 20 horas	0	87
P024	Sim	1	1577
	Não	0	93
P025	Sim	1	1606
	Não	0	64
P026	Sim	1	1521
	Não	0	150
P027	Sim	1	1548
	Não	0	121
P028	Sim	1	1473
	Não	0	0
P029	Alta	1	1014
	Baixa	0	656
P030	Alta	1	1053
	Baixa	0	618
P031	Alta	1	224
	Baixa	0	1446
P032	Alta	1	620
	Baixa	0	1049
P033	Alta	1	269
	Baixa	0	1402
P034	Alta	1	1024
	Baixa	0	644
P035	Alta	1	1380
	Baixa	0	290
P036	Mais de um terço da carga horária	1	512
	Um terço ou menos da carga horária	0	1153
P037	Sim	1	1278
	Não	0	388
P038	Sim	1	778
	Não	0	878

Variáveis Dicotomizadas	Categorias	Código	Frequência amostral
P039	Sim	1	572
	Não	0	1088
P040	Sim	1	537
	Não	0	1123
P041	Sim	1	537
	Não	0	1123
P043	Uma vez ou mais	1	1636
	Não se reuniu	0	32
P045	Sim	1	1167
	Não	0	189
P060	Concordo	1	1187
	Não Concordo	0	491
P061	Concordo	1	1406
	Não Concordo	0	276
P062	Concordo	1	1290
	Não Concordo	0	392
P063	Concordo	1	1287
	Não Concordo	0	392
P064	Concordo	1	1334
	Não Concordo	0	345
P065	Concordo	1	1550
	Não Concordo	0	131
P066	Concordo	1	1553
	Não Concordo	0	128
P067	Concordo	1	1542
	Não Concordo	0	140
P068	Concordo	1	1644
	Não Concordo	0	35
P069	Concordo	1	1534
	Não Concordo	0	145
P070	Concordo	1	1419
	Não Concordo	0	263
P071	Concordo	1	1585
	Não Concordo	0	97
P072	Concordo	1	1350
	Não Concordo	0	332
P073	Concordo	1	1389
	Não Concordo	0	291
P074	Concordo	1	1539
	Não Concordo	0	140
P075	Sim	1	797
	Não	0	876
P076	Sim	1	1178
	Não	0	499
P077	Sim	1	823
	Não	0	857
P078	Sim	1	364
	Não	0	1314
P079	Sim	1	607
	Não	0	1076
P080	Sim	1	952
	Não	0	728
P081	Sim	1	1014
	Não	0	670
P082	Sim	1	1488
	Não	0	193
P115	Sim	1	858
	Não	0	752
P116	Sim	1	862
	Não	0	749
P117	Sim	1	813
	Não	0	800
P118	Sim	1	1157
	Não	0	457
P119	Sim	1	331
	Não	0	1282
P120	Sim	1	1544
	Não	0	79
P121	Mais de 60%	1	1384
	Até 60%	0	230
P122	Mais da Metade	1	743
	Metade ou menos	0	224
P123	Mais da Metade	1	1547
	Metade ou menos	0	71

Variáveis Dicotomizadas	Categorias	Código	Frequência amostral
P124	Mais da Metade	1	1187
	Metade ou menos	0	245
P125	Mais da Metade	1	390
	Metade ou menos	0	1165
P126	Sim	1	1561
	Não	0	61
P129	Não é Ruim	1	1584
	Ruim	0	38
P131	Professor escolheu	1	1230
	Outros membros da equipe escolar escolheram	0	370
P132	Encaminhar os alunos que atrapalham	1	1478
	Atividades pedagógicas	0	24
P142	Alta	1	1562
	Baixa	0	32
P143	Alta	1	1562
	Baixa	0	37
P144	Alta	1	1587
	Baixa	0	14
P145	Alta	1	1478
	Baixa	0	121
P146	Alta	1	1169
	Baixa	0	430
P147	Alta	1	1555
	Baixa	0	46
P148	Alta	1	1430
	Baixa	0	171
P149	Alta	1	321
	Baixa	0	15
P150	Alta	1	1381
	Baixa	0	217
P151	Alta	1	1255
	Baixa	0	346
P152	Alta	1	1546
	Baixa	0	53

AMOSTRAGEM

Primeiramente, temos as regiões administrativas do Estado de São Paulo e os municípios que as compõe, posteriormente temos a distribuição amostral dos dados em cada uma das regiões, baseado nos critérios amostrais colocados no capítulo 2.

B.1 Regiões amostrais

Presidente Prudente: ADAMANTINA, ALFREDO MARCONDES, ALVARES MACHADO, ANHUMAS, CAIABU, CAIUA, DRACENA, EMILIANOPOLIS, ESTRELA DO NORTE, EUCLIDES DA CUNHA PAULISTA, FLORA RICA, FLORIDA PAULISTA, INDIANA, INUBIA PAULISTA, IRAPURU, JOAO RAMALHO, JUNQUEIROPOLIS, LUCELIA, MARABA PAULISTA, MARIAPOLIS, MARTINOPOLIS, MIRANTE DO PARANAPANEMA, MONTE CASTELO, NARANDIBA, NOVA GUATAPORANGA, OSVALDO CRUZ, OURO VERDE, PACAEMBU, PANORAMA, PARAPUA, PAULICEIA, PIQUEROBI, PIRAPOZINHO, PRACINHA, PRESIDENTE BERNARDES, PRESIDENTE EPITACIO, PRESIDENTE PRUDENTE, PRESIDENTE VENCESLAU, RANCHARIA, REGENTE FEIJO, RIBEIRAO DOS INDIOS, RINOPOLIS, ROSANA, SAGRES, SALMOURAO, SANDOVALINA, SANTA MERCEDES, SANTO ANASTACIO, SANTO EXPEDITO, SAO JOAO DO PAU D'ALHO, TACIBA, TARABAI, TEODORO SAMPAIO e TUPI PAULISTA.

Araçatuba: ALTO ALEGRE, ANDRADINA, ARAÇATUBA, AVANHANDAVA, BARBOSA, BENTO DE ABREU, BILAC, BIRIGUI, BRAÚNA, BREJO ALEGRE, BURITAMA, CASTILHO, CLEMENTINA, COROADOS, GABRIEL MONTEIRO, GLICERIO, GUARAÇAI, GUARARAPES, ILHA SOLTEIRA, ITAPURA, LAVINIA, LOURDES, LUIZIANIA, MIRANDOPOLIS, MURUTINGA DO SUL, NOVA INDEPENDENCIA, PENÁPOLIS, PEREIRA BARRETO, PENÁPOLIS, PIACATU, RUBIACEA, SANTO ANTONIO DO ARACANGUA, SANTOPOLIS DO AGUAPEI, SUD MENNUCCI, SUZANAPOLIS, TURIUBA e VALPARAISO.

Jales: ALVARES FLORENCE, AMERICO DE CAMPOS, APARECIDA D'OESTE, ASPASIA, AURIFLAMA, CARDOSO, COSMORAMA, DIRCE REIS, DOLCINOPOLIS, ESTRELA D'OESTE, FERNANDOPOLIS, FLOREAL, GASTAO VIDIGAL, GENERAL SALGADO, GUARANI D'OESTE, GUZOLANDIA, INDIAPORA, JALES, MACEDONIA, MAGDA, MARINOPOLIS, MERIDIANO, MESOPOLIS, MIRA ESTRELA, NOVA CANAA PAULISTA, NOVA CASTILHO, NOVA LUZITANIA, OUROESTE, PALMEIRA D'OESTE, PARANAPUA, PARISI, PEDRANOPOLIS, PONTALINDA, PONTES GESTAL, POPULINA, RIOLANDIA, RUBINEIA, SANTA ALBERTINA, SANTA CLARA D'OESTE, SANTA FÉ DO SUL, SANTA RITA D'OESTE, SANTA SALETE, SANTANA DA PONTE PENSA, SÃO FRANCISCO, SÃO JOAO DAS DUAS PONTES, SÃO JOAO DE IRACEMA, TRÊS FRONTEIRAS, TURMALINA, URÂNIA, VALENTIM GENTIL, VITORIA BRASIL e VOTUPORANGA.

São José do Rio Preto: ADOLFO, ALTAIR, BADCY BASSITT, ARIRANHA, BALSAMO, CAJOBI, CATANDUVA, CATIGUA, CEDRAL, ELISIARIO, EMBAUBA, GUAPIACU, GUARACI, IBIRA, ICEM, IPIGUA, IRAPUA, ITAJOBI, JACI, JOSÉ BONIFACIO, MACAUBAL, MARAPOAMA, MENDONÇA, MIRASSOL, MIRASSOLANDIA, MONCOES, MONTE APRAZIVEL, NEVES PAULISTA, NHANDEARA, NIPOA, NOVA ALIANÇA, NOVA GRANADA, NOVAIS, NOVO HORIZONTE, OLÍMPIA, ONDA VERDE, ORINDIÚVA, PALESTINA, PALMARES PAULISTA, PARAÍSO, PAULO DE FARIA, PINDORAMA, PLANALTO, POLONI, POTIRENDABA, SALES, SANTA ADÉLIA, SÃO JOSE DO RIO PRETO, SEBASTIANÓPOLIS DO SUL, SEVERINIA, TABAPUA, TANABI, UBARANA, UCHOA, UNIÃO PAULISTA, URUPÊS e ZACARIAS.

Marília: ALVARO DE CARVALHO, ALVINLANDIA, ARCO-IRIS, ASSIS, BASTOS, BERNARDINO DE CAMPOS, BORA, CAMPOS NOVOS PAULISTA, CANDIDO MOTA, CANITAR, CHAVANTES, CRUZÁLIA, ECHAPORÃ, ESPIRITO SANTO DO TURVO, FARTURA, FERNÃO, FLORINIA, GÁLIA, GARÇA, HERCULÂNDIA, IACRI, IBIRAREMA, IEPE, IPAUSSU, LUPÉRCIO, LUTÉCIA, MANDURI, MARACAÍ, MARÍLIA, NANTES, OCAUCU, ÓLEO, ORIENTE, OSCAR BRESSANE, OURINHOS, PEDRINHAS PAULISTA, PALMITAL, PARAGUAÇU PAULISTA, PIRAJU, PLATINA, POMPÉIA, QUATÁ, QUEIROZ, QUINTANA, RIBEIRÃO DO SUL, SALTO GRANDE, SANTA CRUZ DO RIO PARDO, SÃO PEDRO DO TURVO, SARUTAIÁ, TAGUAI, TARUMÃ, TEJUPA, TIMBURI, TUPÃ e VERA CRUZ.

Bauru: AGUAS DE SANTA BARBARA, AGUDOS, ANHEMBI, ARANDU, AREALVA, AREIÓPOLIS, AVAÍ, AVARÉ, BALBINOS, BARIRI, BARRA BONITA, BAURU, BOCAINA, BOFETE, BORACÉIA, BOREBI, BOTUCATU, CABRÁLIA PAULISTA, CAFELÂNDIA, CERQUEIRA CÉSAR, CONCHAS, DOIS CORREGOS, DUARTINA, GETULINA, GUAICARA, GUAIMBE, GUARANTA, IACANGA, IARAS, IGARACU DO TIETE, ITAI, ITAJU, ITAPUI, ITATINGA, JAU, JULIO MESQUITA, LENCOIS PAULISTA, LINS, LUCIA-

NOPOLIS, MACATUBA, MINEIROS DO TIETE, PARANAPANEMA, PARDINHO, PAULISTANIA, PEDERNEIRAS, PIRAJUI, PIRATININGA, PONGAI, PRATANIA, PRESIDENTE ALVES, PROMISSAO, REGINOPOLIS, SABINO, SAO MANUEL, UBIRAJARA e URU.

São Carlos: AMERICO BRASILIENSE, ANALANDIA, ARARAQUARA, BOA ESPERANCA DO SUL, BORBOREMA, DESCALVADO, DOBRADA, DOURADO, GAVIAO PEIXOTO, IBATE, IBITINGA, ITAPOLIS, MATAO, MOTUCA, NOVA EUROPA, RIBEIRAO BONITO, RINCAO, SANTA LUCIA, SAO CARLOS, TABATINGA e TRABIJU.

Ribeirão Preto: ALTINOPOLIS, ARAMINA, BARRETOS, BARRINHA, BATATAIS, BEBEDOURO, BRODOWSKI, BURITIZAL, CAJURU, CANDIDO RODRIGUES, CASSIA DOS COQUEIROS, COLINA, COLOMBIA, CRAVINHOS, CRISTAIS PAULISTA, DUMONT, FERNANDO PRESTES, FRANCA, GUAIRA, GUARA, GUARIBA, GUATAPARA, IGARAPAVA, IPUA, ITIRAPUA, ITUVERAVA, JABORANDI, JABOTICABAL, JARDINOPOLIS, JERQUARA, LUIS ANTONIO, MIGUELOPOLIS, MONTE ALTO, MONTE AZUL PAULISTA, MORRO AGUDO, NUPORANGA, ORLANDIA, PATROCINIO PAULISTA, PEDREGULHO, PIRANGI, PITANGUEIRAS, PONTAL, PRADOPOLIS, RESTINGA, RIBEIRAO CORRENTE, RIBEIRAO PRETO, RIFAINA, SALES OLIVEIRA, SANTA CRUZ DA ESPERANCA, SANTA ERNESTINA, SANTA RITA DO PASSA QUATRO, SANTA ROSA DE VITERBO, SANTO ANTONIO DA ALEGRIA, SÃO JOAQUIM DA BARRA, SÃO JOSE DA BELA VISTA, SÃO SIMÃO, SERRA AZUL, SERRANA, SERTÃOZINHO, TAIACU, TAIUVA, TAQUARAL, TAQUARITINGA, TERRA ROXA, VIRADOURO e VISTA ALEGRE DO ALTO.

Registro: BARRA DO TURVO, CAJATI, CANANÉIA, ELDORADO, IGUAPE, ILHA COMPRIDA, ITANHAEM, ITARIRI, JACUPIRANGA, JUQUIA, MIRACATU, MONGAGUA, PARIQUERA-AÇU, PEDRO DE TOLEDO, PERUÍBE, REGISTRO e SETE BARRAS.

Itapetininga: ALAMBARI, ANGATUBA, APIAI, BARAO DE ANTONINA, BARRA DO CHAPEU, BOITUVA, BOM SUCESSO DE ITARARE, BURI, CAMPINA DO MONTE ALEGRE, CAPAO BONITO, CERQUILHO, CESARIO LANGE, CORONEL MACEDO, GUAPIARA, GUAREI, IPORANGA, ITABERA, ITAOCA, ITAPETININGA, ITAPEVA, ITAPIRAPUA PAULISTA, ITAPORANGA, ITARARE, LARANJAL PAULISTA, NOVA CAMPINA, PEREIRAS, PORANGABA, QUADRA, RIBEIRA, RIBEIRAO BRANCO, RIBEIRAO GRANDE, RIVERSUL, TAQUARITUBA, TAQUARIVAI, TATUÍ e TORRE DE PEDRA.

Piracicaba: AGUAS DE SAO PEDRO, ARARAS, BROTAS, CAPIVARI, CHARQUEADA, CONCHAL, CORDEIROPOLIS, CORUMBATAI, IPEUNA, IRACEMAPOLIS, ITIRAPINA, JUMIRIM, LEME, LIMEIRA, MOMBUCA, PIRACICABA, RAFARD, RIO CLARO, RIO DAS PEDRAS, SALTINHO, SANTA CRUZ DA CONCEICAO, SANTA GERTRUDES, SANTA MARIA DA SERRA, SAO PEDRO, TIETE e TORRINHA.

Campinas: AGUAI, AGUAS DA PRATA, AGUAS DE LINDOIA, AMERICANA, AM-

PARO, ARTUR NOGUEIRA, CACONDE, CAMPINAS, CASA BRANCA, COSMOPOLIS, DIVINOLANDIA, ELIAS FAUSTO, ENGENHEIRO COELHO, ESPIRITO SANTO DO PINHAL, ESTIVA GERBI, HOLAMBRA, HORTOLANDIA, INDAIATUBA, ITAPIRA, ITOBI, JAGUARIUNA, LINDOIA, MOCOCA, MOGI GUACU, MOGI MIRIM, MONTE ALEGRE DO SUL, MONTE MOR, NOVA ODESSA, PAULINIA, PEDRA BELA, PEDREIRA, PINHALZINHO, PIRASSUNUNGA, PORTO FERREIRA, SANTA BARBARA D'OESTE, SANTA CRUZ DAS PALMEIRAS, SANTO ANTONIO DE POSSE, SANTO ANTONIO DO JARDIM, SAO JOAO DA BOA VISTA, SAO JOSE DO RIO PARDO, SAO SEBASTIAO DA GRAMA, SERRA NEGRA, SOCORRO, SUMARE, TAMBAU, TAPIRATIBA, VALINHOS, VARGEM GRANDE DO SUL e VINHEDO.

Jundiaí: ALUMINIO, ARACARIGUAMA, ARACOIABA DA SERRA, ATIBAIA, BOM JESUS DOS PERDOES, BRAGANCA PAULISTA, CABREUVA, CAMPO LIMPO PAULISTA, CAPELA DO ALTO, IBIUNA, IPERO, ITATIBA, ITU, ITUPEVA, JARINU, JOANOPOLIS, JUNDIAI, LOUVEIRA, MAIRINQUE, MORUNGABA, NAZARE PAULISTA, PIEDADE, PILAR DO SUL, PIRACAIA, PORTO FELIZ, SALTO, SALTO DE PIRAPORA, SAO MIGUEL ARCANJO, SAO ROQUE, SARAPUI, SOROCABA, TAPIRAI, TUIUTI, VARGEM, VARZEA PAULISTA e VOTORANTIM.

São Paulo: ARUJA, BARUERI, BERTIOGA, BIRITIBA-MIRIM, CAIEIRAS, CAJAMAR, CARAPICUIBA, COTIA, CUBATAO, DIADEMA, EMBU, EMBU-GUACU, FERRAZ DE VASCONCELOS, FRANCISCO MORATO, FRANCO DA ROCHA, GUARAREMA, GUARUJA, GUARULHOS, ITAPECERICA DA SERRA, ITAPEVI, ITAQUAQUECETUBA, JANDIRA, JUQUITIBA, MAIRIPORA, MAUA, MOGI DAS CRUZES, OSASCO, PIRAPORA DO BOM JESUS, POA, PRAIA GRANDE, RIBEIRAO PIRES, RIO GRANDE DA SERRA, SALESOPOLIS, SANTA ISABEL, SANTANA DE PARNAIBA, SANTO ANDRE, SANTOS, SAO BERNARDO DO CAMPO, SAO CAETANO DO SUL, SAO LOURENCO DA SERRA, SAO PAULO, SAO VICENTE, SUZANO, TABOAO DA SERRA e VARGEM GRANDE PAULISTA.

São José dos Campos: APARECIDA, ARAPEI, AREIAS, BANANAL, CACAPAVA, CACHOEIRA PAULISTA, CAMPOS DO JORDAO, CANAS, CARAGUATATUBA, CRUZEIRO, CUNHA, GUARATINGUETA, IGARATA, ILHABELA, JACAREI, JAMBEIRO, LAGOINHA, LAVRINHAS, LORENA, MONTEIRO LOBATO, NATIVIDADE DA SERRA, PARAIBUNA, PINDAMONHANGABA, PIQUETE, POTIM, QUELUZ, REDENCAO DA SERRA, ROSEIRA, SANTA BRANCA, SANTO ANTONIO DO PINHAL, SAO BENTO DO SAPUCAI, SAO JOSE DO BARREIRO, SAO JOSE DOS CAMPOS, SAO LUIS DO PARAITINGA, SAO SEBASTIAO, SILVEIRAS, TAUBATE, TREMEMBE e UBATUBA.

B.2 Distribuições amostrais por região

Tabela 33 – Distribuição amostral das escolas da região de Presidente Prudente segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	1	1	0
	ENTRE O PERCENTIL 25 E 50	2	1	0
	ENTRE O PERCENTIL 50 E 75	1	2	1
	ACIMA DO PERCENTIL 75	1	1	1
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	0	0	0
	ENTRE O PERCENTIL 50 E 75	0	1	0
	ACIMA DO PERCENTIL 75	0	0	0

Tabela 34 – Distribuição amostral das escolas da região de Araçatuba segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	0	1	0
	ENTRE O PERCENTIL 25 E 50	1	1	1
	ENTRE O PERCENTIL 50 E 75	1	2	0
	ACIMA DO PERCENTIL 75	1	2	0
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	0	0	0
	ENTRE O PERCENTIL 50 E 75	0	0	0
	ACIMA DO PERCENTIL 75	0	0	0

Tabela 35 – Distribuição amostral das escolas da região de Jales segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	1	0	0
	ENTRE O PERCENTIL 50 E 75	2	1	0
	ACIMA DO PERCENTIL 75	3	2	0
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	0	0	0
	ENTRE O PERCENTIL 50 E 75	0	0	0
	ACIMA DO PERCENTIL 75	0	0	0

Tabela 36 – Distribuição amostral das escolas da região de São José do Rio Preto segundo os critérios pré-estabelecidos.

Dependência	Quartis	Pequena	Média	Grande
Administrativa				
Estadual	ABAIXO DO PERCENTIL 25	1	0	0
	ENTRE O PERCENTIL 25 E 50	1	1	0
	ENTRE O PERCENTIL 50 E 75	2	2	1
	ACIMA DO PERCENTIL 75	1	3	1
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	0	1	0
	ENTRE O PERCENTIL 50 E 75	0	0	0
	ACIMA DO PERCENTIL 75	2	2	0

Tabela 37 – Distribuição amostral das escolas da região de Marília segundo os critérios pré-estabelecidos.

Dependência	Quartis	Pequena	Média	Grande
Administrativa				
Estadual	ABAIXO DO PERCENTIL 25	1	1	0
	ENTRE O PERCENTIL 25 E 50	2	2	0
	ENTRE O PERCENTIL 50 E 75	2	3	0
	ACIMA DO PERCENTIL 75	2	3	1
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	0	0	0
	ENTRE O PERCENTIL 50 E 75	0	0	0
	ACIMA DO PERCENTIL 75	0	1	0

Tabela 38 – Distribuição amostral das escolas da região de Bauru segundo os critérios pré-estabelecidos.

Dependência	Quartis	Pequena	Média	Grande
Administrativa				
Estadual	ABAIXO DO PERCENTIL 25	1	3	0
	ENTRE O PERCENTIL 25 E 50	2	2	1
	ENTRE O PERCENTIL 50 E 75	1	4	1
	ACIMA DO PERCENTIL 75	1	3	1
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	0	0	0
	ENTRE O PERCENTIL 50 E 75	0	0	0
	ACIMA DO PERCENTIL 75	1	1	0

Tabela 39 – Distribuição amostral das escolas da região de São Carlos segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	1	1	0
	ENTRE O PERCENTIL 25 E 50	1	1	0
	ENTRE O PERCENTIL 50 E 75	1	2	0
	ACIMA DO PERCENTIL 75	1	2	1
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	1	0	0
	ENTRE O PERCENTIL 50 E 75	1	0	0
	ACIMA DO PERCENTIL 75	1	1	0

Tabela 40 – Distribuição amostral das escolas da região de Ribeirão Preto segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	1	1	0
	ENTRE O PERCENTIL 25 E 50	1	3	1
	ENTRE O PERCENTIL 50 E 75	1	4	2
	ACIMA DO PERCENTIL 75	1	5	2
Municipal	ABAIXO DO PERCENTIL 25	1	1	0
	ENTRE O PERCENTIL 25 E 50	1	1	0
	ENTRE O PERCENTIL 50 E 75	2	2	1
	ACIMA DO PERCENTIL 75	3	2	1

Tabela 41 – Distribuição amostral das escolas da região de Registro segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	2	1	0
	ENTRE O PERCENTIL 25 E 50	1	1	0
	ENTRE O PERCENTIL 50 E 75	2	2	0
	ACIMA DO PERCENTIL 75	1	0	0
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	0	0	0
	ENTRE O PERCENTIL 50 E 75	0	0	0
	ACIMA DO PERCENTIL 75	0	0	0

Tabela 42 – Distribuição amostral das escolas da região de Itapetininga segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	2	1	0
	ENTRE O PERCENTIL 25 E 50	2	1	0
	ENTRE O PERCENTIL 50 E 75	1	2	0
	ACIMA DO PERCENTIL 75	1	1	0
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	1	1	0
	ENTRE O PERCENTIL 50 E 75	1	1	0
	ACIMA DO PERCENTIL 75	2	1	0

Tabela 43 – Distribuição amostral das escolas da região de Piracicaba segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	0	1	0
	ENTRE O PERCENTIL 25 E 50	1	2	0
	ENTRE O PERCENTIL 50 E 75	1	2	1
	ACIMA DO PERCENTIL 75	2	6	1
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	0	0	0
	ENTRE O PERCENTIL 50 E 75	1	0	0
	ACIMA DO PERCENTIL 75	1	1	0

Tabela 44 – Distribuição amostral das escolas da região de Campinas segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	1	2	1
	ENTRE O PERCENTIL 25 E 50	3	5	2
	ENTRE O PERCENTIL 50 E 75	4	8	3
	ACIMA DO PERCENTIL 75	5	10	3
Municipal	ABAIXO DO PERCENTIL 25	0	0	0
	ENTRE O PERCENTIL 25 E 50	1	0	0
	ENTRE O PERCENTIL 50 E 75	1	1	0
	ACIMA DO PERCENTIL 75	4	3	1

Tabela 45 – Distribuição amostral das escolas da região de Jundiaí segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	1	2	1
	ENTRE O PERCENTIL 25 E 50	1	4	2
	ENTRE O PERCENTIL 50 E 75	1	6	3
	ACIMA DO PERCENTIL 75	2	5	3
Municipal	ABAIXO DO PERCENTIL 25	1	0	0
	ENTRE O PERCENTIL 25 E 50	1	1	0
	ENTRE O PERCENTIL 50 E 75	1	1	0
	ACIMA DO PERCENTIL 75	3	3	0

Tabela 46 – Distribuição amostral das escolas da região de São José dos Campos segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	2	2	0
	ENTRE O PERCENTIL 25 E 50	3	2	1
	ENTRE O PERCENTIL 50 E 75	3	5	1
	ACIMA DO PERCENTIL 75	2	3	1
Municipal	ABAIXO DO PERCENTIL 25	2	0	0
	ENTRE O PERCENTIL 25 E 50	1	2	0
	ENTRE O PERCENTIL 50 E 75	2	3	0
	ACIMA DO PERCENTIL 75	2	4	1

Tabela 47 – Distribuição amostral das escolas da região de São Paulo segundo os critérios pré-estabelecidos.

Dependência Administrativa	Quartis	Pequena	Média	Grande
Estadual	ABAIXO DO PERCENTIL 25	13	33	36
	ENTRE O PERCENTIL 25 E 50	8	22	29
	ENTRE O PERCENTIL 50 E 75	6	16	14
	ACIMA DO PERCENTIL 75	3	10	9
Municipal	ABAIXO DO PERCENTIL 25	6	20	4
	ENTRE O PERCENTIL 25 E 50	4	14	4
	ENTRE O PERCENTIL 50 E 75	4	11	1
	ACIMA DO PERCENTIL 75	3	7	2

MODELOS CLÁSSICOS

Neste apêndice apresentamos os procedimentos computacionais utilizados para a estimação de modelos clássicos, para tanto foi utilizados os pacotes `lme4`, `nlme`, `lmerTest` e `bbml`. Mais detalhes com relação aos pacotes citados podem ser encontrados em (BATES *et al.*, 2015; PINHEIRO *et al.*, 2017; KUZNETSOVA; BROCKHOFF; CHRISTENSEN, 2015; BOLKER; TEAM, 2010)

C.1 Estimação dos parâmetros

Nesta seção temos as rotinas computacionais para estimação de parâmetros nos casos: efeitos fixos, intercepto aleatório e intercepto e inclinações aleatórias.

C.1.1 Modelos com efeitos fixos

O modelo (6.7), foi estimado numa abordagem clássica, vejamos o código a seguir.

```
Modelo1 <- lm(PROFICIENCIA_MT_SAEB ~ as.factor(Q001) + as.factor(Q004) +
as.factor(Q019) + as.factor(Q023) + as.factor(Q049) +
as.factor(DEPENDENCIA_ADM) + as.factor(TURNO) + as.factor(LOCALIZACAO) +
PROFICIENCIA_LP_SAEB, data = dados)
summary(Modelo1)
```

C.1.2 Modelo com interceptos aleatórios

Os códigos para estimação do modelo (6.10), encontram - se a seguir.

```
require(lme4)
require(lmerTest)
```

```
Modelo2 <- lmer(PROFICIENCIA_MT_SAEB ~ as.factor(Q001) + as.factor(Q004) +
as.factor(Q019) + as.factor(Q023) + as.factor(Q049) +
as.factor(DEPENDENCIA_ADM) + as.factor(TURNO) + as.factor(LOCALIZACAO) +
PROFICIENCIA_LP_SAEB + (1|REGIAO), data = dados)
summary(Modelo2)
```

C.1.3 Modelo com interceptos e inclinações aleatórios

Abaixo temos os códigos, para estimação do modelo (6.13).

```
require(nlme)

Modelo3 <- lme(PROFICIENCIA_MT_SAEB ~ as.factor(Q001) + as.factor(Q004) +
as.factor(Q019) + as.factor(Q023) + as.factor(Q049) +
as.factor(DEPENDENCIA_ADM) + as.factor(TURNO) + as.factor(LOCALIZACAO) +
PROFICIENCIA_LP_SAEB, random = ~ as.factor(Q001) + as.factor(Q004) +
as.factor(Q019) + as.factor(Q023) + as.factor(Q049) +
as.factor(DEPENDENCIA_ADM) + as.factor(TURNO) + as.factor(LOCALIZACAO) +
PROFICIENCIA_LP_SAEB|REGIAO, control=lmeControl(opt="optim",maxIter=
50000, msMaxIter=200, tolerance=1e-4, niter=50, msTol=1e-5,
nlmStepMax=500, msVerbose=TRUE, returnObject=TRUE), data=dados)
summary(Modelo3)
```

C.2 Seleção de Modelos

Como já foi dito nos capítulos 4 e 5 a seleção de modelos no caso clássico foi feita usando os critérios AIC, AICc e BIC, a seguir podemos ver como foi computacionalmente realizado esse procedimento.

```
require(bbmle)

AICctab(Modelo1, Modelo2, Modelo3, base = T, weights = T)
AICtab(Modelo1, Modelo2, Modelo3, base = T, weights = T)
BICtab(Modelo1, Modelo2, Modelo3, base = T, weights = T)
```

MODELOS BAYESIANOS: STAN

Quando se trabalha com dados em que há medidas repetidas e com grande número de observações, uma possível abordagem seria o ajuste de modelos mistos. Essa classe de modelos tem se popularizado muito nos últimos anos, principalmente em ciências cognitivas e educacionais, por se tratar de vertentes científicas que possuem dados com grande quantidade de observações. Houve um amplo desenvolvimento computacional para dar suporte a esses profissionais. Com esse pensamento surgiu a interface Stan, que posteriormente foi adaptada para ser utilizada na linguagem R, para ajustes bayesianos. O Stan possibilita que o pesquisador leve em consideração seus conhecimentos prévios sobre o tema para a modelagem do problema e utiliza como método de estimação o Monte Carlo Hamiltoniano.

Segundo Neal *et al.* (2011), Monte Carlo Hamiltoniano (HMC) é um método de Monte Carlo (MCMC) de cadeia de Markov que utiliza a derivada da função de densidade dos dados que estão sendo amostrados para gerar transições eficientes que abrangem a distribuição *a posteriori* geradas dos dados originais.

O primeiro passo é escrever a função de densidade da distribuição que se pretende amostrar como função de energia potencial e cinética, essa função será da forma:

$$H(\rho, \theta) = -\log p(\rho, \theta), \quad (\text{D.1})$$

onde ρ é uma variável auxiliar de impulso, daí simula-se cadeias de Markov e em seguida faz-se um atualização via algoritmo de Metropolis. Em Mecânica Estatística, a função H é conhecida como distribuição canônica da distribuição que se pretende amostrar.

Cada iteração do algoritmo HMC tem duas etapas. Na primeira etapa, novos valores para as variáveis de ρ são aleatoriamente extraídos de sua distribuição, este passo, obviamente, deixa a distribuição canônica comum invariante. Na segunda etapa, uma atualização Metropolis é realizada, usando a dinâmica hamiltoniana para propor um novo estado. Começando com

o estado atual, (ρ, θ) . Se o estado proposto não for aceito (ou seja, ele é rejeitado), o estado seguinte é o mesmo que o estado atual.

Segundo Neal *et al.* (2011), se olharmos para HMC como amostragem da distribuição conjunta de ρ e θ , o passo Metropolis deixa a função densidade de probabilidade para (ρ, θ) inalterada ou quase inalterada. O movimento para pontos (ρ, θ) com uma densidade de probabilidade diferente é realizado somente pelo primeiro passo em uma iteração de HMC, em que ρ é substituído por um novo valor.

Neste apêndice encontram-se os códigos em Stan que foram usados para construção dos modelos bayesianos do capítulo 5.

D.1 Modelos com efeitos fixos

O modelo (6.14), foi estimado numa abordagem bayesiana, vejamos o código no stan

```
data{
int <lower=1> N; //Número de observações
real PROFICIENCIA_LP_SAEB[N]; // preditor
real PROFICIENCIA_MT_SAEB[N]; // variável resposta
int <lower=0, upper = 1> DEPENDENCIA_ADM [N]; // preditor
int <lower=0, upper = 1> LOCALIZACAO [N]; // preditor
int <lower=0, upper = 1> TURNO [N]; // preditor
int <lower=0, upper = 1> Q001 [N]; // preditor
int <lower=0, upper = 1> Q004 [N]; // preditor
int <lower=0, upper = 1> Q019 [N]; // preditor
int <lower=0, upper = 1> Q023 [N]; // preditor
int <lower=0, upper = 1> Q049 [N]; // preditor
}
parameters {
vector[10] beta; // intercepto e inclinações
real<lower=0> sigma_e; // erro padrão
}
model{
real mu;
for (i in 1:N) {
mu= beta[1] + beta[2]*PROFICIENCIA_LP_SAEB[i] +
beta[3]*DEPENDENCIA_ADM[i] +
beta[4]*LOCALIZACAO[i] +
beta[5]*TURNO[i] + beta[6]*Q001[i] +
beta[7]*Q004[i] + beta[8]*Q019[i] +
```

```

beta[9]*Q023[i] + beta[10]*Q049[i];    //verossimilhança
PROFICIENCIA_MT_SAEB[i] ~ normal (mu, sigma_e);
}
}
generated quantities{
real dev;
vector[N] log_lik;
dev <- 0;
for(n in 1:N){
dev <- dev + (-2)*normal_log(PROFICIENCIA_MT_SAEB[n],
beta[1] + beta[2]*PROFICIENCIA_LP_SAEB[n] +
beta[3]*DEPENDENCIA_ADM[n] + beta[4]*LOCALIZACAO[n] +
beta[5]*TURNO[n] + beta[6]*Q001[n] + beta[7]*Q004[n] +
beta[8]*Q019[n] + beta[9]*Q023[n] + beta[10]*Q049[n],
sigma_e);    //Calculo do deviance
log_lik[n] <- normal_log(PROFICIENCIA_MT_SAEB[n],
beta[1] + beta[2]*PROFICIENCIA_LP_SAEB[n] +
beta[3]*DEPENDENCIA_ADM[n] + beta[4]*LOCALIZACAO[n] +
beta[5]*TURNO[n] + beta[6]*Q001[n] + beta[7]*Q004[n] +
beta[8]*Q019[n] + beta[9]*Q023[n] + beta[10]*Q049[n],
sigma_e);    //Calculo da log-verossimilhança
}
}

```

D.2 Modelo com interceptos aleatórios

Os códigos do Stan para estimação bayesiana do modelo (6.15), encontram - se a seguir.

```

data{
int<lower=1> N;    // número de observações
real PROFICIENCIA_LP_SAEB[N];    //preditor
real PROFICIENCIA_MT_SAEB[N];    // variável resposta
int<lower=1> J;    //número de escolas
int<lower=1, upper= J> ESCOLA[N];    //identificação de escolas
int <lower=0, upper = 1> DEPENDENCIA_ADM [N]; //preditor
int <lower=0, upper = 1> LOCALIZACAO [N]; //preditor
int <lower=0, upper = 1> TURNO [N]; //preditor
int <lower=0, upper = 1> Q001 [N]; //preditor
int <lower=0, upper = 1> Q004 [N]; //preditor

```

```

int <lower=0, upper = 1> Q019 [N];      //preditor
int <lower=0, upper = 1> Q023 [N];      //preditor
int <lower=0, upper = 1> Q049 [N];      //preditor
}
parameters {
vector[10] beta;      //intercepto e inclinações fixas
vector[J] w;          //interceptos de escola
real<lower=0> sigma_e; //erro padrão
real<lower=0> sigma_w; //componentes de variância de escola
}
model{
real mu;
w ~ normal(0,sigma_w);          //priori do efeito aleatório de escola
for (i in 1:N){
mu= beta[1] + w[ESCOLA[i]] +
beta[2]*PROFICIENCIA_LP_SAEB[i] + beta[3]*DEPENDENCIA_ADM[i] +
beta[4]*LOCALIZACAO[i] + beta[5]*TURNO[i] + beta[6]*Q001[i] +
beta[7]*Q004[i] + beta[8]*Q019[i] +
beta[9]*Q023[i] + beta[10]*Q049[i];
PROFICIENCIA_MT_SAEB[i] ~ normal (mu, sigma_e); //maximaverossimilhança
}
}
generated quantities{
real dev;          //Deviance
vector[N] log_lik; //Log - Verossimilhança
dev <- 0;
for(n in 1:N){
dev <- dev + (-2)*normal_log(PROFICIENCIA_MT_SAEB[n],
beta[1] + w[ESCOLA[n]] + beta[2]*PROFICIENCIA_LP_SAEB[n] +
beta[3]*DEPENDENCIA_ADM[n] + beta[4]*LOCALIZACAO[n] +
beta[5]*TURNO[n] + beta[6]*Q001[n] + beta[7]*Q004[n] +
beta[8]*Q019[n] + beta[9]*Q023[n] + beta[10]*Q049[n],
sigma_e);
log_lik[n] <- normal_log(PROFICIENCIA_MT_SAEB[n],
beta[1] + w[ESCOLA[n]] + beta[2]*PROFICIENCIA_LP_SAEB[n] +
beta[3]*DEPENDENCIA_ADM[n] + beta[4]*LOCALIZACAO[n] +
beta[5]*TURNO[n] + beta[6]*Q001[n] + beta[7]*Q004[n] +
beta[8]*Q019[n] + beta[9]*Q023[n] + beta[10]*Q049[n],
sigma_e);
}
}

```

```
}
}
```

D.3 Modelo com interceptos e inclinações aleatórios

Abaixo temos os códigos, da interface Stan, para estimação do modelo (6.16).

```
data{
  int<lower=1> N;           //número de observações
  real PROFICIENCIA_LP_SAEB[N];      //preditor
  real PROFICIENCIA_MT_SAEB[N];      //variável resposta
  int<lower=1> J;           //número de escolas
  int<lower=1, upper= J> ESCOLA[N];   //preditor
  int <lower=0, upper = 1> DEPENDENCIA_ADM [N]; //preditor
  int <lower=0, upper = 1> LOCALIZACAO [N]; //preditor
  int <lower=0, upper = 1> TURNO [N]; //preditor
  int <lower=0, upper = 1> Q001 [N]; //preditor
  int <lower=0, upper = 1> Q004 [N]; //preditor
  int <lower=0, upper = 1> Q019 [N]; //preditor
  int <lower=0, upper = 1> Q023 [N]; //preditor
  int <lower=0, upper = 1> Q049 [N]; //preditor
}
parameters {
  vector[10] beta; //interceptos e inclinações fixas
  real<lower=0> sigma_e; //erro padrão
  vector<lower=0>[10] sigma_w; // componentes de variância de escola
  cholesky_factor_corr[10] L_w;
  matrix [10,J] z_w;
}
transformed parameters{
  matrix[10,J] w;
  w <- diag_pre_multiply(sigma_w,L_w)*z_w; //efeito aleatório de escola
}
model{
  real mu;
  L_w ~ lkj_corr_cholesky(10.0); //priors
  to_vector(z_w) ~ normal(0,1);
  for (i in 1:N){
    mu= beta[1] + w[1,ESCOLA[i]] +
```

```

(beta[2] + w[2,ESCOLA[i]])*PROFICIENCIA_LP_SAEB[i] +
(beta[3] + w[3,ESCOLA[i]])*DEPENDENCIA_ADM[i] +
(beta[4] + w[4,ESCOLA[i]])*LOCALIZACAO[i] +
(beta[5] + w[5,ESCOLA[i]])*TURNO[i] +
(beta[6] + w[6,ESCOLA[i]])*Q001[i] +
(beta[7] + w[7,ESCOLA[i]])*Q004[i] +
(beta[8] + w[8,ESCOLA[i]])*Q019[i] +
(beta[9] + w[9,ESCOLA[i]])*Q023[i] +
(beta[10] + w[10,ESCOLA[i]])*Q049[i];
PROFICIENCIA_MT_SAEB[i] ~ normal(mu, sigma_e);    maximaverossimilhança
}
}
generated quantities{
real dev;          //Deviance
vector[N] log_lik; //log-verossimilhança
dev <- 0;
for(n in 1:N){
dev <- dev + (-2)*normal_log(PROFICIENCIA_MT_SAEB[n],
beta[1] + w[1,ESCOLA[n]] +
(beta[2] + w[2,ESCOLA[n]])*PROFICIENCIA_LP_SAEB[n] +
(beta[3] + w[3,ESCOLA[n]])*DEPENDENCIA_ADM[n] +
(beta[4] + w[4,ESCOLA[n]])*LOCALIZACAO[n] +
(beta[5] + w[5,ESCOLA[n]])*TURNO[n] +
(beta[6] + w[6,ESCOLA[n]])*Q001[n] +
(beta[7] + w[7,ESCOLA[n]])*Q004[n] +
(beta[8] + w[8,ESCOLA[n]])*Q019[n] +
(beta[9] + w[9,ESCOLA[n]])*Q023[n] +
(beta[10] + w[10,ESCOLA[n]])*Q049[n], sigma_e);
log_lik[n] <- normal_log(PROFICIENCIA_MT_SAEB[n],
beta[1] + w[1,ESCOLA[n]] +
(beta[2] + w[2,ESCOLA[n]])*PROFICIENCIA_LP_SAEB[n] +
(beta[3] + w[3,ESCOLA[n]])*DEPENDENCIA_ADM[n] +
(beta[4] + w[4,ESCOLA[n]])*LOCALIZACAO[n] +
(beta[5] + w[5,ESCOLA[n]])*TURNO[n] +
(beta[6] + w[6,ESCOLA[n]])*Q001[n] +
(beta[7] + w[7,ESCOLA[n]])*Q004[n] +
(beta[8] + w[8,ESCOLA[n]])*Q019[n] +
(beta[9] + w[9,ESCOLA[n]])*Q023[n] +
(beta[10] + w[10,ESCOLA[n]])*Q049[n], sigma_e);
}
}

```

}
}