

A coleta de dados

A Prova Brasil é uma avaliação diagnóstica realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep/MEC) com o intuito de avaliar os alunos de ensino básico do quinto ano e do nono ano do Ensino Fundamental. É também um exame para avaliar o quão bom é o sistema educacional no Brasil. A prova é realizada em escolas públicas e particulares. Podemos declarar que a população alvo são esses próprios alunos de quinto e nono ano. Entretanto não temos o mesmo grupo como população acessível, visto que apenas escolas mais de 20 alunos matriculados nesses anos são avaliadas, perdendo assim uma parte da sua população alvo. As principais características levantadas incluem o nome, idade, nota das provas (Português e Matemática) como também escolaridade do pai e da mãe, sexo, entre outras variáveis. A maneira que eles coletam esses dados é justamente através da prova e através de questionário feito pelos alunos.

Amostra do aluno 232014905

Segue a seguir uma tabela de todas variáveis encontradas na amostra fornecida para o aluno com matrícula 232014905 junto com descrições de cada uma e o tipo delas, além da escala.

Variável	Tipo.de.variável	Escala	Descrição.da.variável
Ano	Quantitativa discreta.	Nominal.	Ano de realização da SAEB.
Região	Qualitativa nominal.	Nominal.	Região onde é localizada a escola.
Unidade Federativa	Qualitativa nominal.	Nominal.	UF da localização da escola.
Município	Qualitativa nominal.	Nominal.	Município onde é localizada a escola.
Área	Qualitativa nominal.	Nominal.	Área de localização do estudante.
Dependência administrativa	Qualitativa nominal.	Nominal.	Âmbito de subordinação da escola.
Localização	Qualitativa nominal.	Nominal.	Localização da escola
Nota em LP	Quantitativa contínua.	Intervalar.	Nota na prova de língua portuguesa.
Nota em MAT	Quantitativa contínua.	Intervalar.	Nota na prova de matemática.
Sexo	Qualitativa nominal.	Nominal.	Sexo do estudante.
Raça	Qualitativa nominal.	Nominal.	Raça do estudante.
Mês de Nasc.	Qualitativa nominal.	Nominal.	Mês de nascimento.
Ano de Nasc.	Quantitativa discreta.	Nominal.	Ano de nascimento.
Computador	Qualitativa ordinal.	Razão.	Se o aluno tem computador em casa.
Esc. Mãe	Qualitativa ordinal.	Ordinal.	Escolaridade da mãe.
Esc. Pai	Qualitativa ordinal.	Ordinal.	Escolaridade do pai.
Tempo de tela	Qualitativa ordinal.	Ordinal.	Quanto tempo o estudante gasta em telas eletrônicas.
Afazer dom.	Qualitativa ordinal.	Ordinal.	Tempo gasto fazendo afazeres domésticos.
Trabalho	Qualitativa nominal.	Nominal.	Se o aluno trabalha ou não.
Perspectivas	Qualitativa nominal.	Nominal.	Perspectivas do futuro.

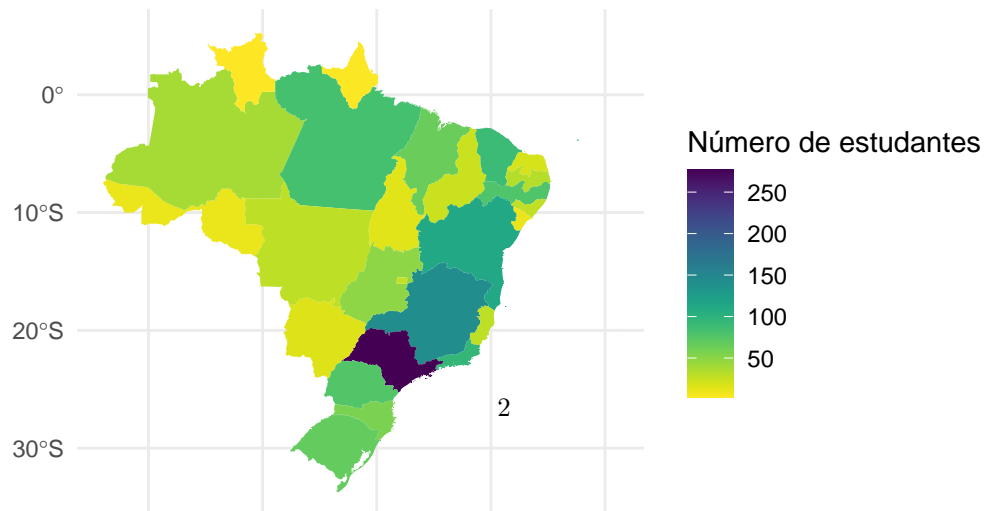
Tabela 1: Explicação das Variáveis

Análise das variáveis categóricas

As variáveis selecionadas foram: sexo(1) e raça(2) do estudante, região(3) onde é localizada a escola, se tem ou não computadores(4) e

Unidade Federativa	Frequência absoluta	Frequência relativa
São Paulo	277	0.185
Minas Gerais	143	0.095
Bahia	113	0.075
Rio De Janeiro	95	0.063
Ceará	90	0.06
Pará	84	0.056
Pernambuco	78	0.052
Paraná	77	0.051
Rio Grande Do Sul	68	0.045
Maranhão	65	0.043
Santa Catarina	57	0.038
Goiás	49	0.033
Amazônas	39	0.026
Paraíba	33	0.022
Espírito Santo	30	0.02
Mato Grosso	30	0.02
Alagoas	28	0.019
Distrito Federal	27	0.018
Piauí	25	0.017
Rio Grande Do Norte	20	0.013
Mato Grosso Do Sul	17	0.011
Tocantins	15	0.01
Acre	11	0.007
Rondônia	11	0.007
Sergipe	11	0.007
Amapá	4	0.003
Roraima	3	0.002
NA	1500	1

Tabela 2: qtd de alunos



Análise de variáveis quantitativas

Proeficiência do aluno em Língua Portuguesa

Tabela 3: Distribuição da nota dos alunos

Intervalo de Classe	Frequência
135 [—) 157	27
157 [—) 179	96
179 [—) 201	111
201 [—) 222	162
222 [—) 244	238
244 [—) 266	238
266 [—) 287	238
287 [—) 309	200
309 [—) 331	114
331 [—) 352	59
352 [—) 374	17
Total	1500

Tabela 4: Medidas resumo

Média	Mediana	Erro padrão	Coef. de Variação	Coef. de Assimetria de Pearson	Curtose
253.529	255.934	48.805	0.193	-2.817	-0.571

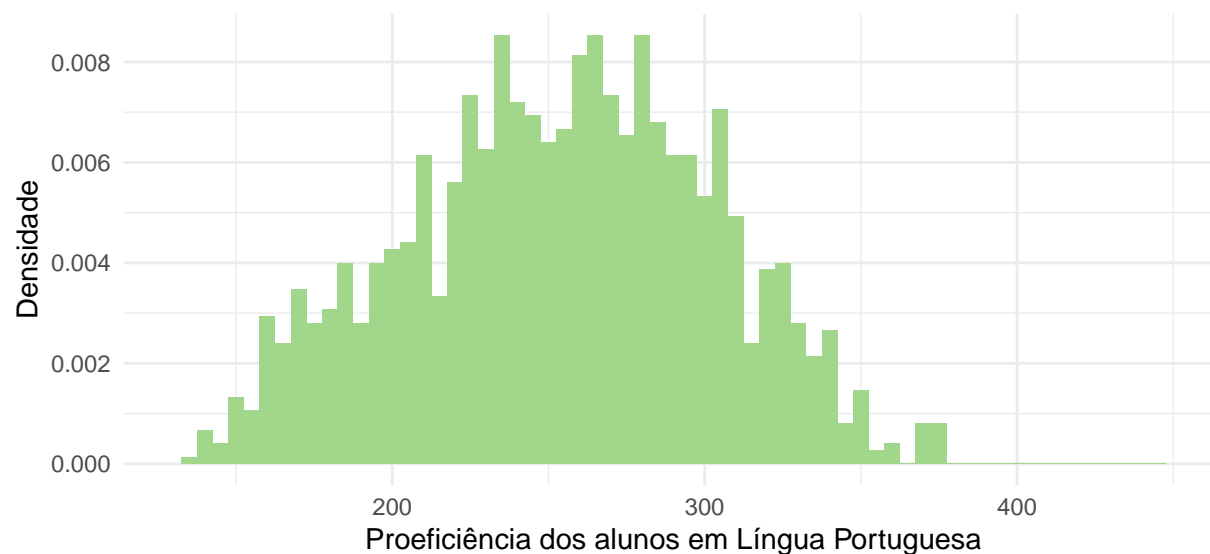


Gráfico 2: Histograma das notas de LP

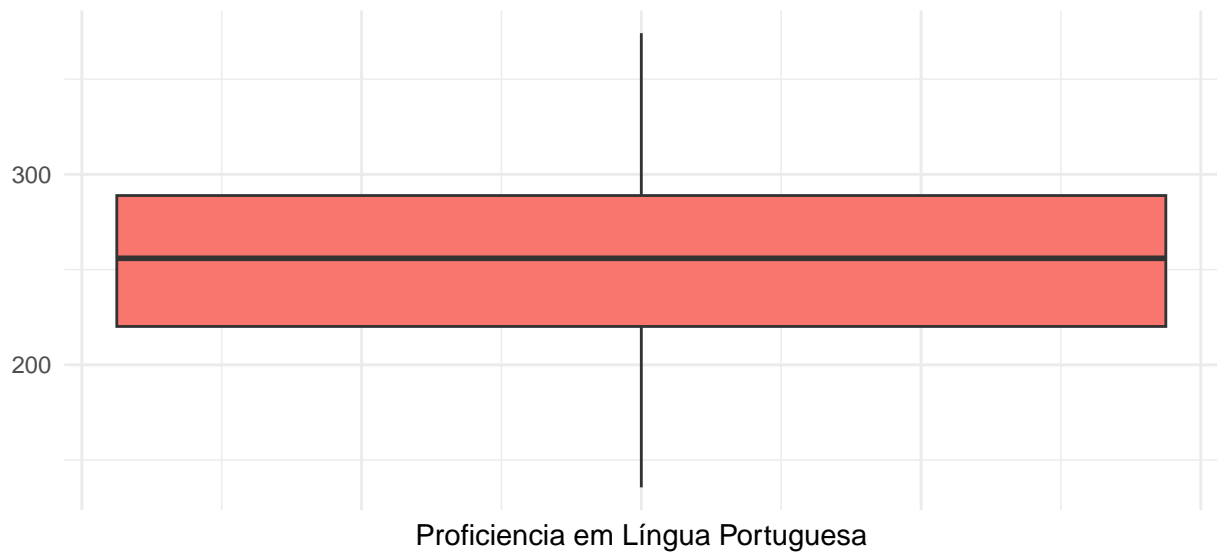


Gráfico 3: Box-plot das notas de LP

Proeficiência do aluno em Matemática

Tabela 5: Distribuição da nota dos alunos

Intervalo de Classe	Frequência
133 [—) 160	37
160 [—) 187	108
187 [—) 214	201
214 [—) 240	299
240 [—) 267	290
267 [—) 294	255
294 [—) 321	193
321 [—) 347	83
347 [—) 374	28
374 [—) 401	4
401 [—) 428	2
Total	1500

Tabela 6: Medidas resumo

Média	Mediana	Erro padrão	Coef. de Variação	Coef. de Assimetria de Pearson	Curtose
250.859	249.162	48.681	0.194	3.786	-0.374

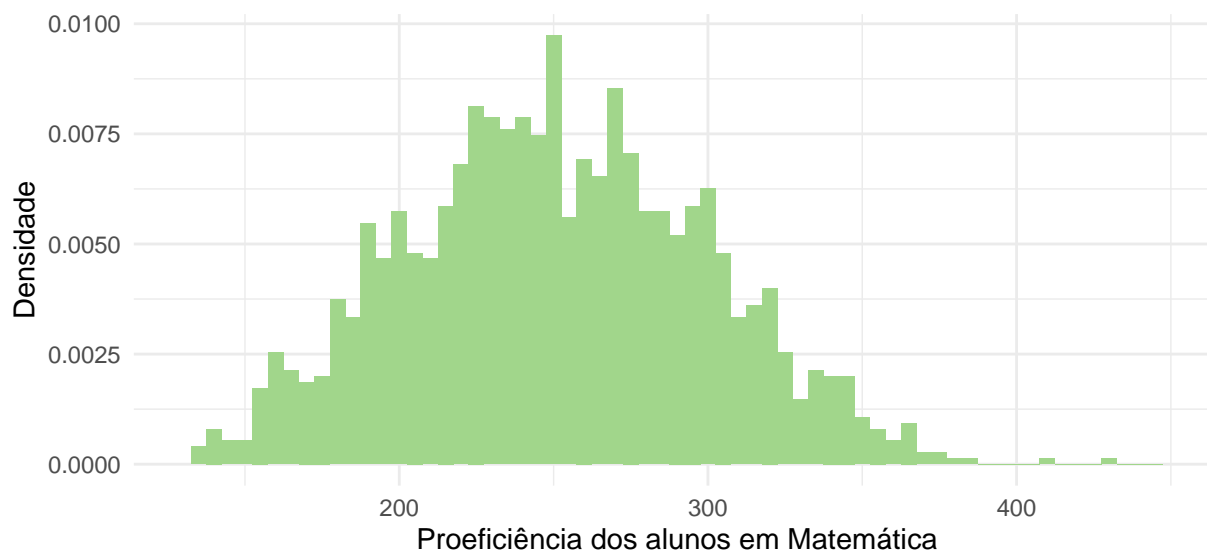


Gráfico 4: Histograma das notas de Mat.

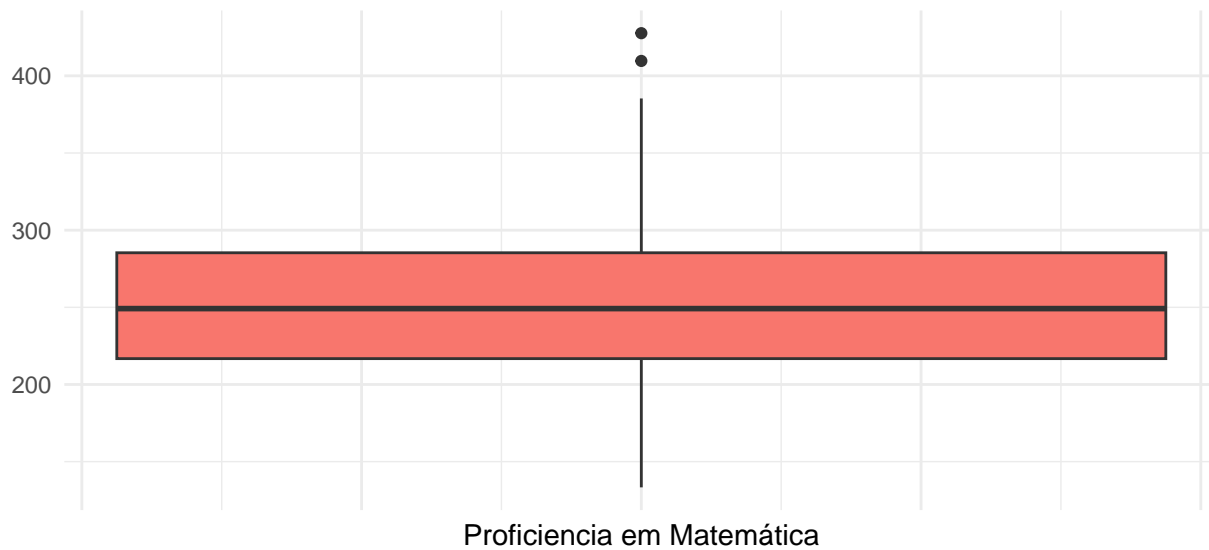


Gráfico 5: Box-plot das notas de Mat.