

## 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

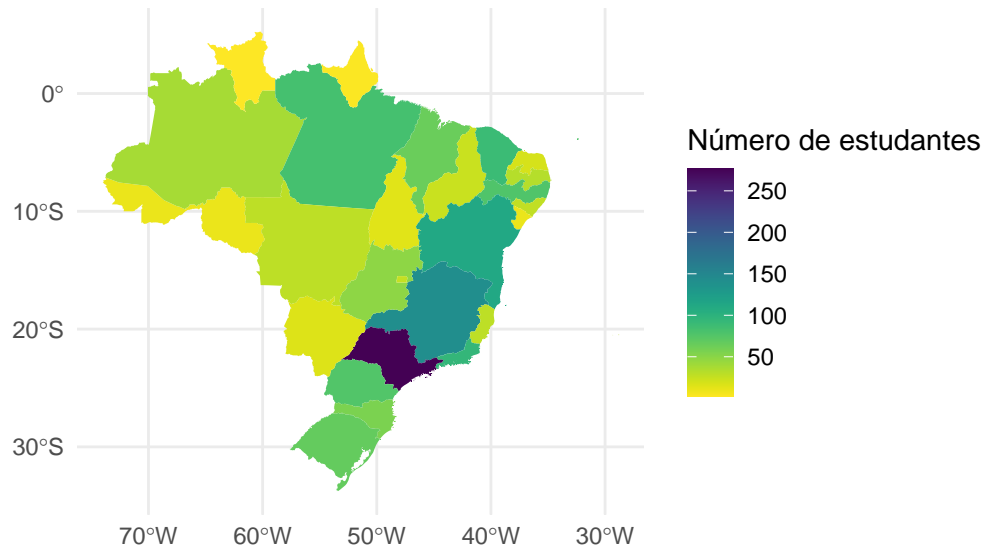


Gráfico 1: Mapa

## Análise de variáveis quantitativas

### Proeficiência do aluno em Língua Portuguesa

```
qtd_classes <- round(1 + 3.3 * log(length(banco)))

dist_freq <- table(cut(banco$NOTA_LP, breaks = qtd_classes))
labels <- as.vector(names(dist_freq))

rownames(dist_freq) <- map_chr(labels, ~gsub("\\((\\d+),(\\d+)\\]", "\\1 |----- \\2", .x))

kable(dist_freq, "latex", caption = "nota dos alunos") %>%
  kable_styling(latex_options = c("hold_position"), full_width = FALSE) %>%
  column_spec(1) %>%
  row_spec(0, bold = TRUE) %>%
  kable_styling(latex_options = c("scale_down"))
```

Unidade Federativa	Frequência absoluta	Frequência relativa
São Paulo	277	0.1846667
Minas Gerais	143	0.0953333
Bahia	113	0.0753333
Rio De Janeiro	95	0.0633333
Ceará	90	0.0600000
Pará	84	0.0560000
Pernambuco	78	0.0520000
Paraná	77	0.0513333
Rio Grande Do Sul	68	0.0453333
Maranhão	65	0.0433333
Santa Catarina	57	0.0380000
Goiás	49	0.0326667
Amazônas	39	0.0260000
Paraíba	33	0.0220000
Espírito Santo	30	0.0200000
Mato Grosso	30	0.0200000
Alagoas	28	0.0186667
Distrito Federal	27	0.0180000
Piauí	25	0.0166667
Rio Grande Do Norte	20	0.0133333
Mato Grosso Do Sul	17	0.0113333
Tocantins	15	0.0100000
Acre	11	0.0073333
Rondônia	11	0.0073333
Sergipe	11	0.0073333
Amapá	4	0.0026667
Roraima	3	0.0020000

**Tabela 2:** qtd de alunos

<b>Var1</b>	<b>Freq</b>
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

**Tabela 3:** nota dos alunos