



Universidade de Brasília  
IE- Departamento de Estatística  
Métodos Estatísticos 2

Juliana Magalhães Rosa

**Análise de Dados**  
Amostra dos Resultados do Saeb

Brasília  
2020

## Sumário

<b>1 Introdução</b>	<b>2</b>
<b>2 Variáveis</b>	<b>3</b>
<b>3 Análise Categórica</b>	<b>6</b>
3.1 REGIAO	6
3.2 UF	7
3.3 MUNICIPIO	9
3.4 AREA	11
3.5 DEPENDENCIA_ADM	12
3.6 LOCALIZACAO	13
3.7 SEXO	14
3.8 RACA_COR	15
3.9 MES_NASC	16
3.10 ANO_NASC	18
3.11 COMPUTADOR	20
3.12 ESC_MAE	21
3.13 ESC_PAI	23
3.14 USO_TEMPO_TELAS	24
3.15 AFAZERES_DOM	26
3.16 TRABALHO	27
3.17 PERSPECTIVAS	28
<b>4 Análise Numérica</b>	<b>29</b>
4.1 NOTA_LP	29
4.2 NOTA_MT	33
<b>5 Considerações Finais</b>	<b>37</b>

## 1 Introdução

O Sistema de Avaliação da Educação Básica (Saeb) é um conjunto de avaliações externas em larga escala que permite ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) realizar um diagnóstico da educação básica brasileira e de fatores que podem interferir no desempenho do estudante.

O Saeb tem como população alvo os estudantes do ensino fundamental e médio (e, mais recentemente, também do ensino infantil) de escolas públicas e privadas no Brasil. Porém, a população acessível é composta somente por estudantes daquelas escolas que se dispõem a participar da prova, já que a participação não é obrigatória.

As características levantadas são as habilidades em língua portuguesa e em matemática (e, a partir de 2019, também em ciências da natureza e em ciências humanas para os estudantes do 9º ano), algumas informações sobre as instituições de ensino, e algumas informações pessoais dos estudantes (sexo e raça, por exemplo).

A coleta dos dados é feita a partir de testes e questionários (eletrônicos ou no papel) que são respondidos pelos estudantes e por alguns funcionários das escolas (professores, diretores, etc.).

Neste trabalho será feita uma análise de dados relativos aos resultados dos estudantes do nono ano no Saeb 2017. Será utilizada uma amostra de 2000 estudantes, contendo 20 variáveis.

## 2 Variáveis

ANO: ano de realização do Saeb.

Variável Qualitativa Nominal.

REGIAO: região de localização da escola do estudante (Norte, Nordeste, Sudeste, Sul ou Centro-Oeste).

Variável Qualitativa Nominal.

UF: unidade da federação de localização da escola do estudante.

Variável Qualitativa Nominal.

MUNICIPIO: município de localização da escola do estudante.

Variável Qualitativa Nominal.

AREA: área de localização da escola do estudante (capital ou interior).

Variável Qualitativa Nominal.

LOCALIZACAO: localização da escola do estudante (urbana ou rural).

Variável Qualitativa Nominal.

DEPENDENCIA\_ADM: categoria administrativa da escola do estudante (federal, estadual ou municipal).

Variável Qualitativa Nominal.

NOTA\_LP: proficiência em Língua Portuguesa transformada na escala única do SAEB, com média = 250, desvio = 50 (do SAEB/97).

Variável Quantitativa Contínua.

NOTA\_MT: proficiência do aluno em Matemática transformada na escala única do SAEB, com média = 250, desvio = 50 (do SAEB/97).

Variável Quantitativa Contínua.

SEXO: sexo do estudante (masculino ou feminino).

Variável Qualitativa Nominal.

RACA\_COR: raça/cor do estudante (branca, preta, parda, amarela, indígena ou não quero declarar).

Variável Qualitativa Nominal.

MÊS\_NASC: mês de nascimento do estudante.

Variável Qualitativa Nominal.

ANO\_NASC: ano de nascimento do estudante (2005 ou depois, 2004, 2003, 2002, 2001, 2001, 1999, 1998 ou antes).

Variável Qualitativa Nominal.

COMPUTADOR: tem computador em casa? (não tem; sim, um; sim, dois; sim, três; sim, quatro ou mais).

Variável Qualitativa Ordinal.

ESC\_MAE: escolaridade da mãe (nunca estudou, não completou a 4.<sup>a</sup> série/5.<sup>o</sup> ano do Ensino Fundamental, completou a 4.<sup>a</sup> série/5.<sup>o</sup> ano, mas não completou a 8.<sup>a</sup> série/9.<sup>o</sup> ano do Ensino Fundamental, completou a 8.<sup>a</sup> série/9.<sup>o</sup> ano do Ensino Fundamental, mas não completou o Ensino Médio', completou o Ensino Médio, mas não completou a Faculdade, completou a Faculdade, não sei).

Variável Qualitativa Ordinal.

ESC\_PAI: escolaridade do pai (nunca estudou, não completou a 4.<sup>a</sup> série/5.<sup>o</sup> ano do Ensino Fundamental, completou a 4.<sup>a</sup> série/5.<sup>o</sup> ano, mas não completou a 8.<sup>a</sup> série/9.<sup>o</sup> ano do Ensino Fundamental, completou a 8.<sup>a</sup> série/9.<sup>o</sup> ano do Ensino Fundamental, mas não completou o Ensino Médio', completou o Ensino Médio, mas não completou a Faculdade, completou a Faculdade, não sei).

Variável Qualitativa Ordinal.

USO\_TEMPO\_TELAS: tempo gasto assistindo à TV, navegando na internet ou jogando jogos eletrônicos em dias de aula (menos de 1 hora, entre 1 e 2 horas, mais de 2 horas, até 3 horas, mais de 3 horas, não vejo TV, não navego na internet e não jogo jogos eletrônicos).

Variável Qualitativa Ordinal.

AFAZERES\_DOM: tempo gasto fazendo trabalhos domésticos (ex.: lavando louça, limpando o quintal etc.) em dias de aula (menos de 1 hora, entre 1 e 2 horas, mais de 2 horas, até 3 horas, mais de 3 horas, não faço trabalhos domésticos).

Variável Qualitativa Ordinal.

TRABALHO: atualmente você trabalha fora de casa (recebendo ou não um salário) (sim ou não).

Variável Qualitativa Nominal.

PERSPECTIVAS: o que pretende fazer quando terminar o 9o. ano (8a. série) (somente continuar estudando, somente trabalhar, continuar estudando e trabalhar, ainda não sei).

Variável Qualitativa Nominal.

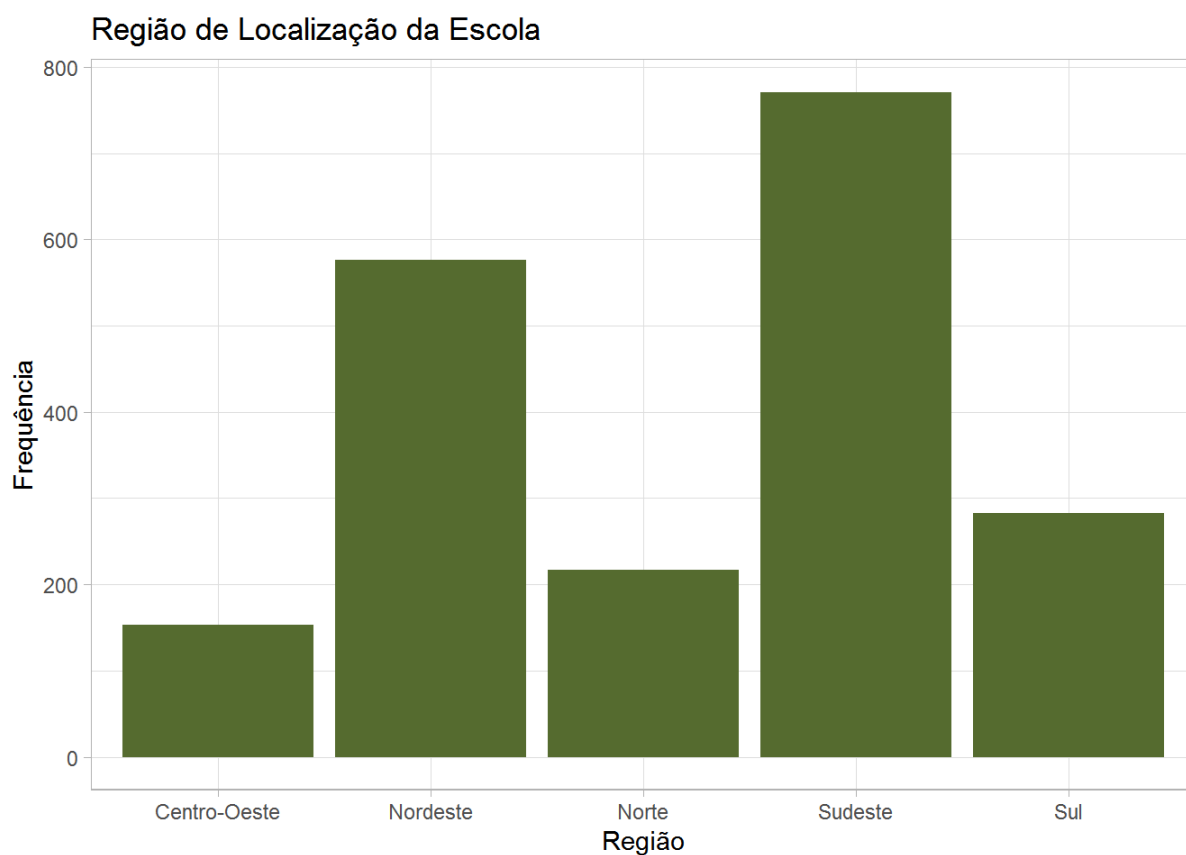
### 3 Análise Categórica

#### 3.1 REGIAO

**Tabela 1-** Distribuição de Escolas entre as Regiões Brasileiras. Saeb, 2017.

Região	Frequência Absoluta	Frequência Relativa (%)
Centro-Oeste	153	7.65
Nordeste	576	28.80
Norte	217	10.85
Sudeste	771	38.55
Sul	283	14.15
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



A região sudeste é a que tem maior quantidade de estudantes do nono ano que participaram do Saeb em 2017, seguida pelo nordeste. De modo geral, as porcentagens mostradas na tabela de frequências condizem com a densidade populacional conhecida das regiões brasileiras.

### 3.2 UF

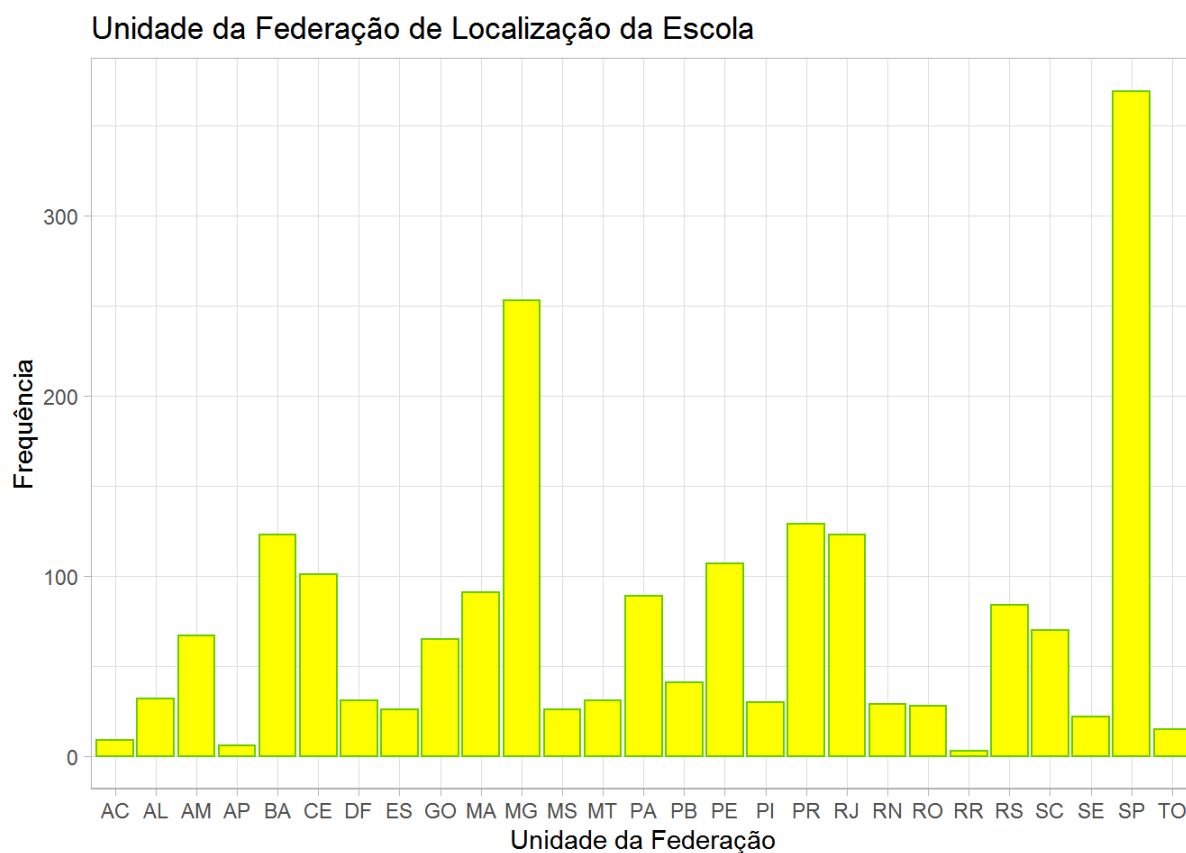
**Tabela 2** – Distribuição de Escolas entre as Unidades Federativas Brasileiras. Saeb, 2017.

Unidade da Federação	Frequência Absoluta	Frequência Relativa (%)
AC	9	0.45
AL	32	1.60
AM	67	3.35
AP	6	0.30
BA	123	6.15
CE	101	5.05
DF	31	1.55
ES	26	1.30
GO	65	3.25
MA	91	4.55
MG	253	12.65
MS	26	1.30
MT	31	1.55
PA	89	4.45
PB	41	2.05
PE	107	5.35
PI	30	1.50
PR	129	6.45
RJ	123	6.15
RN	29	1.45
RO	28	1.40
RR	3	0.15



RS	84	4.20
SC	70	3.50
SE	22	1.10
SP	369	18.45
TO	15	0.75
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



As quantidades de estudantes não variam tanto entre as unidades da federação, com exceção de São Paulo e Minas Gerais. Esses dois estados são, de fato, os mais populosos do Brasil. Essa distinção acentuada fica evidente no gráfico de colunas.

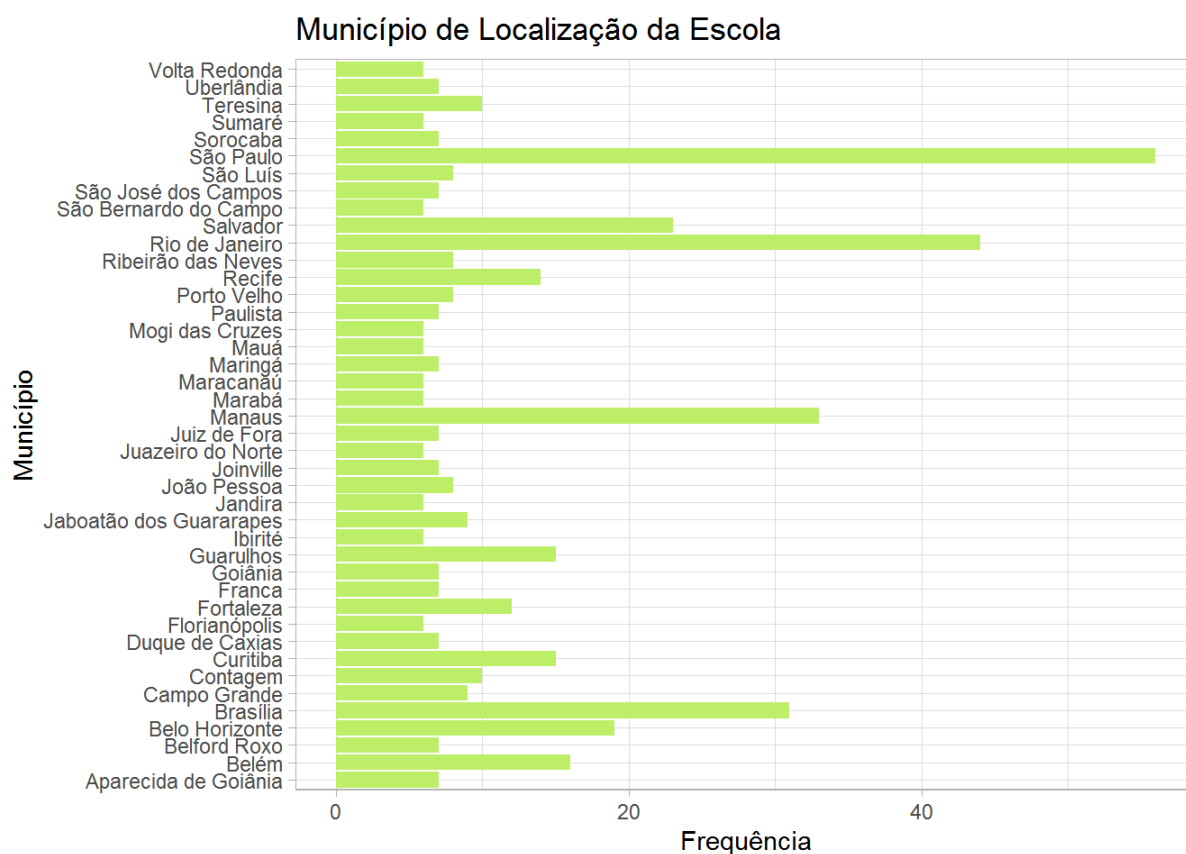
### 3.3 MUNICIPIO

**Tabela 3** – Distribuição de Escolas entre os Municípios Brasileiros. Saeb, 2017.

<b>Município</b>	<b>Frequência Absoluta</b>
Aparecida de Goiânia	7
Belém	16
Belford Roxo	7
Belo Horizonte	19
Brasília	31
Campo Grande	9
Contagem	10
Curitiba	15
Duque de Caxias	7
Florianópolis	6
Fortaleza	12
Franca	7
Goiânia	7
Guarulhos	15
Ibirité	6
Jaboatão dos Guararapes	9
Jandira	6
João Pessoa	8
Joinville	7
Juazeiro do Norte	6
Juiz de Fora	7
Manaus	33
Marabá	6
Maracanaú	6
Maringá	7
Mauá	6
Mogi das Cruzes	6

Paulista	7
Porto Velho	8
Recife	14
Ribeirão das Neves	8
Rio de Janeiro	44
Salvador	23
São Bernardo do Campo	6
São José dos Campos	7
São Luís	8
São Paulo	56
Sorocaba	7
Sumaré	6
Teresina	10
Uberlândia	7
Volta Redonda	6

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



Como são muitos os municípios que participam do Saeb, foram selecionados somente aqueles que possuíam mais de cinco estudantes registrados na prova em 2017, para se obter uma visualização mais informativa nas formas tabular e, principalmente, gráfica. Nessa distribuição de frequências, se destacam os municípios de São Paulo, Rio de Janeiro, Manaus e Brasília, respectivamente, com maiores números de estudantes.

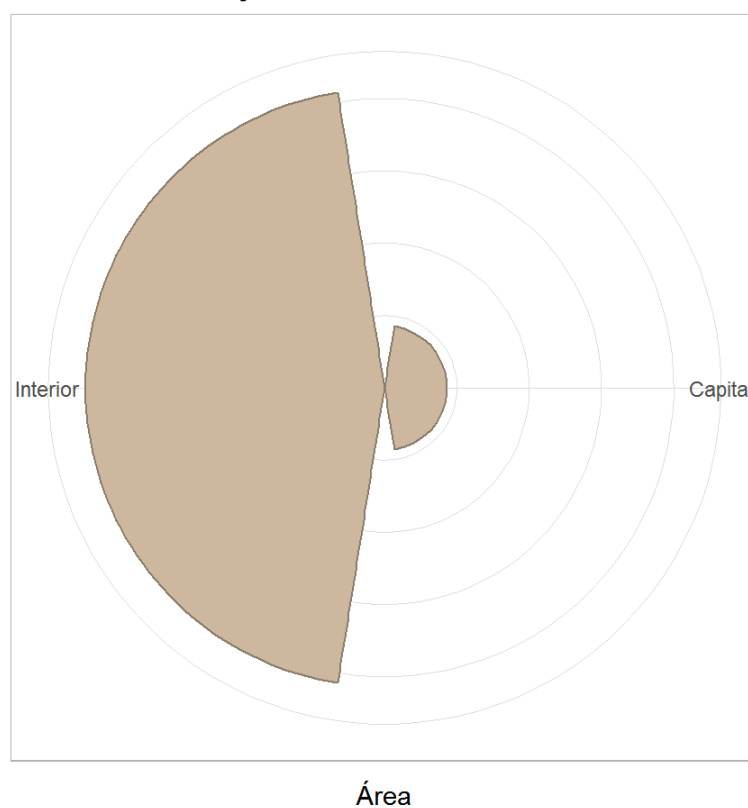
### 3.4 AREA

**Tabela 4 – Distribuição de Escolas entre as capitais e o interior. Saeb, 2017.**

Área	Frequência Absoluta	Frequência Relativa (%)
Capital	345	17.25
Interior	1655	82.75
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>

**Área de Localização da Escola**



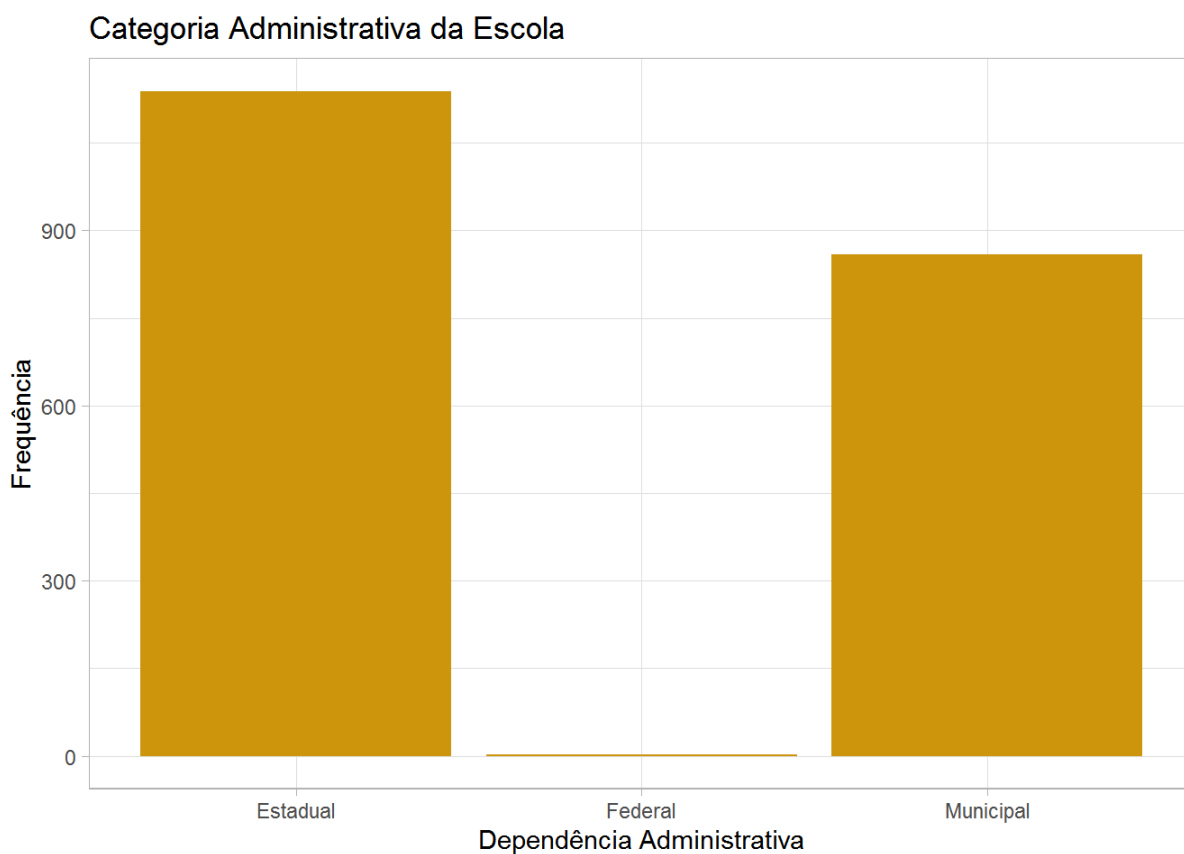
Olhando tanto para a tabela como para o gráfico, fica clara a discrepância entre a quantidade de estudantes nas capitais e no interior. Essa diferença se dá ao fato de existir apenas uma capital por uf, sendo todos os outros municípios classificados como interior.

### 3.5 DEPENDENCIA\_ADM

**Tabela 5** – Distribuição de Frequências entre Escolas Estaduais, Federais e Municipais. Saeb, 2017.

<b>Dependência Administrativa</b>	<b>Frequência Absoluta</b>	<b>Frequência Relativa (%)</b>
Estadual	1139	56.95
Federal	2	0.10
Municipal	859	42.95
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



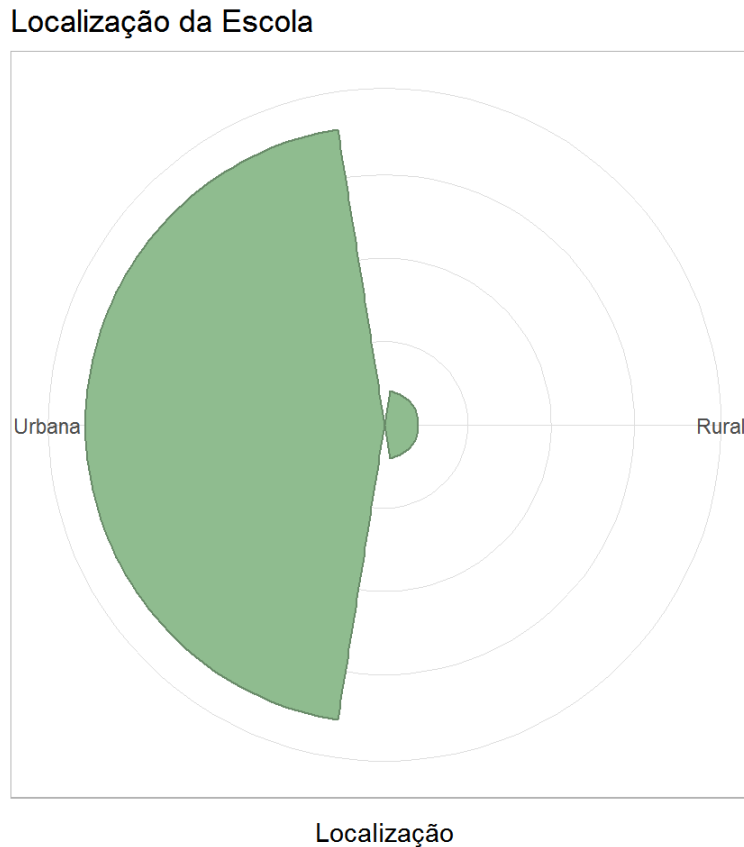
Os números de escolas com dependência administrativa municipal e estadual são similares. O que chama a atenção é o número quase insignificante de escolas com dependência federal. São apenas duas, sendo uma de Brasília e a outra do Rio de Janeiro.

### 3.6 LOCALIZACAO

**Tabela 6** – Distribuição de Escolas entre áreas rurais e urbanas. Saeb, 2017.

Localização	Frequência Absoluta	Frequência Relativa (%)
Rural	205	10.25
Urbana	1795	89.75
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



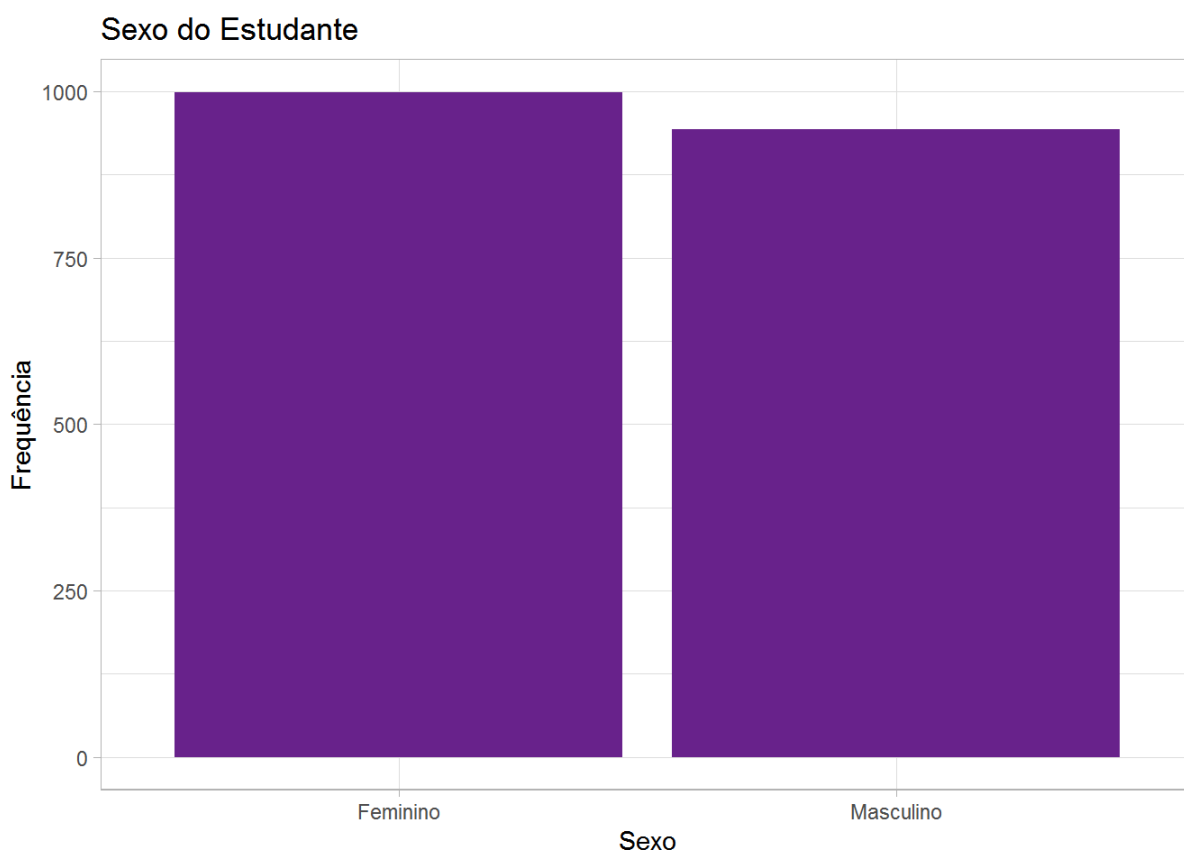
Existe um número quase nove vezes maior de estudantes na área urbana quando comparada à rural. Assim como ocorre para hospitais, faculdades, etcetera, a concentração de escolas é muito maior nas cidades do que nas áreas rurais.

### 3.7 SEXO

**Tabela 7 – Distribuição de Estudantes quanto ao Sexo. Saeb, 2017.**

Sexo	Frequência Absoluta	Frequência Relativa (%)
Feminino	999	49.95
Masculino	943	47.15
NA	58	2.90
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



A proporção de estudantes do sexo feminino é bem próxima da proporção de estudantes do sexo masculino, sendo apenas levemente mais alta. Essa pequena diferença é esperada, já que no Brasil existe um número um pouco maior de mulheres do que de homens.

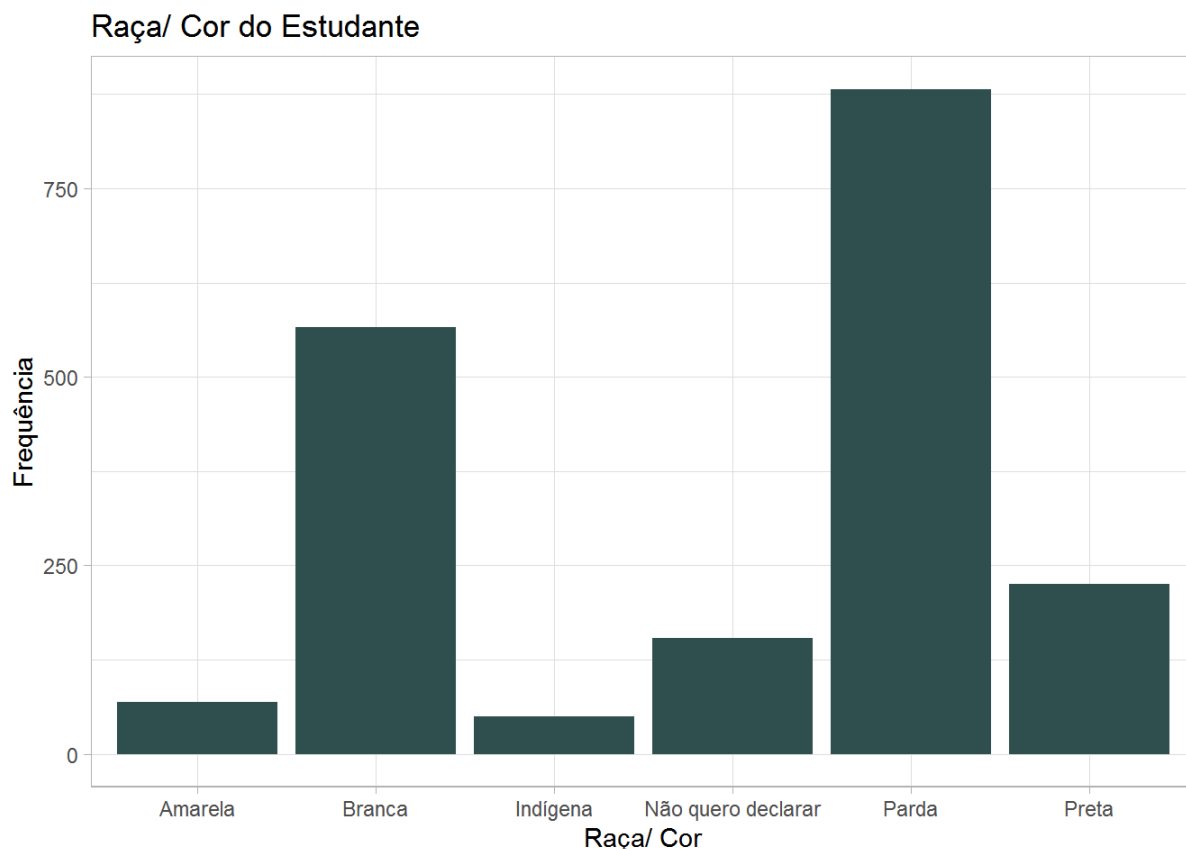
### 3.8 RACA\_COR

**Tabela 8** – Distribuição de Estudantes quanto à Raça/ Cor. Saeb, 2017.

<b>Raça/ Cor</b>	<b>Frequência Absoluta</b>	<b>Frequência Relativa (%)</b>
Amarela	69	3.45
Branca	566	28.30
Indígena	50	2.50
Não quero declarar	154	7.70
Parda	882	44.10
Preta	225	11.25
NA	54	2.70
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>





A maioria dos estudantes (44,1%) se declarou como parda, mas também há uma boa proporção de estudantes (28,3%) que se considera branca. Vale comentar que mais de 7% dos estudantes não quis declarar sua raça/cor.

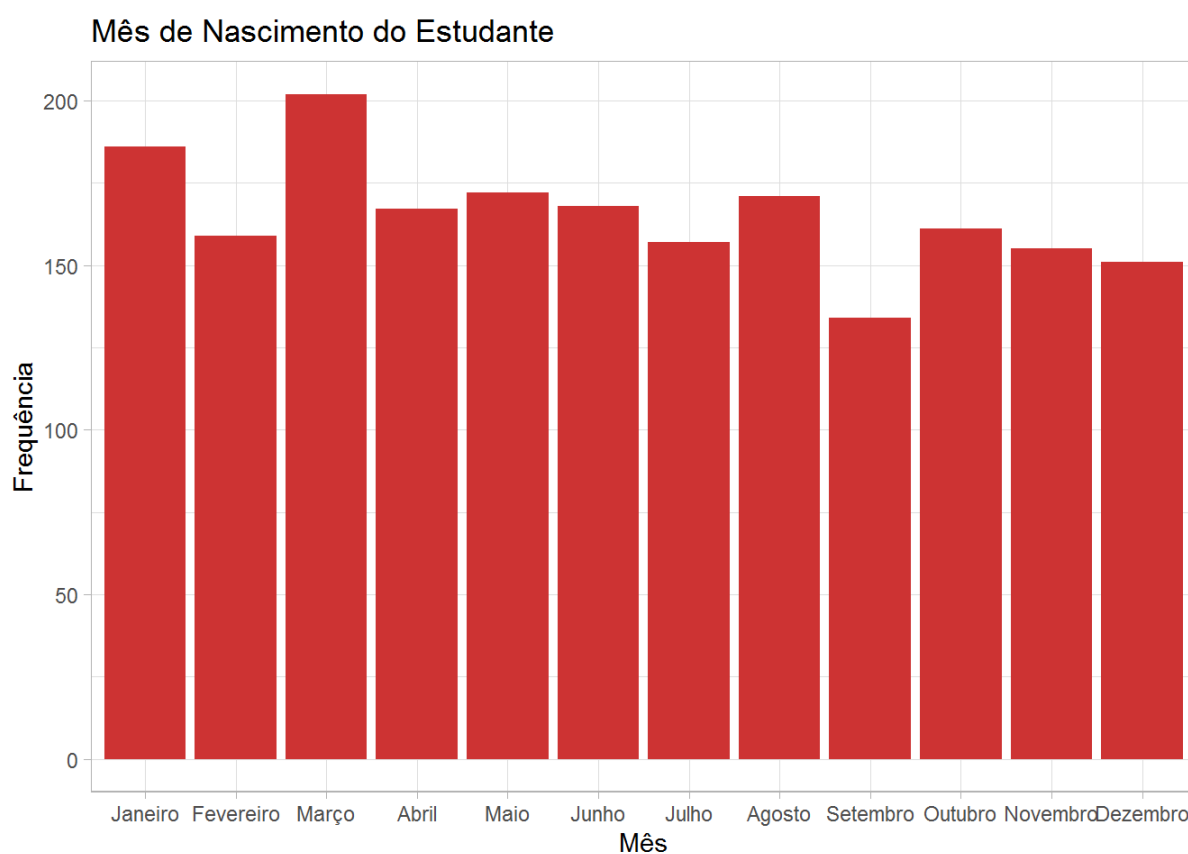
### 3.9 MES\_NASC

**Tabela 9** – Distribuição de Estudantes quanto ao seu Mês de Nascimento. Saeb, 2017.

Mês de Nascimento	Frequência Absoluta	Frequência Relativa (%)
Janeiro	186	9.30
Fevereiro	159	7.95
Março	202	10.10
Abril	167	8.35
Maio	172	8.60
Junho	168	8.40

Julho	157	7.85
Agosto	171	8.55
Setembro	134	6.70
Outubro	161	8.05
Novembro	155	7.75
Dezembro	151	7.55
NA	17	0.85
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



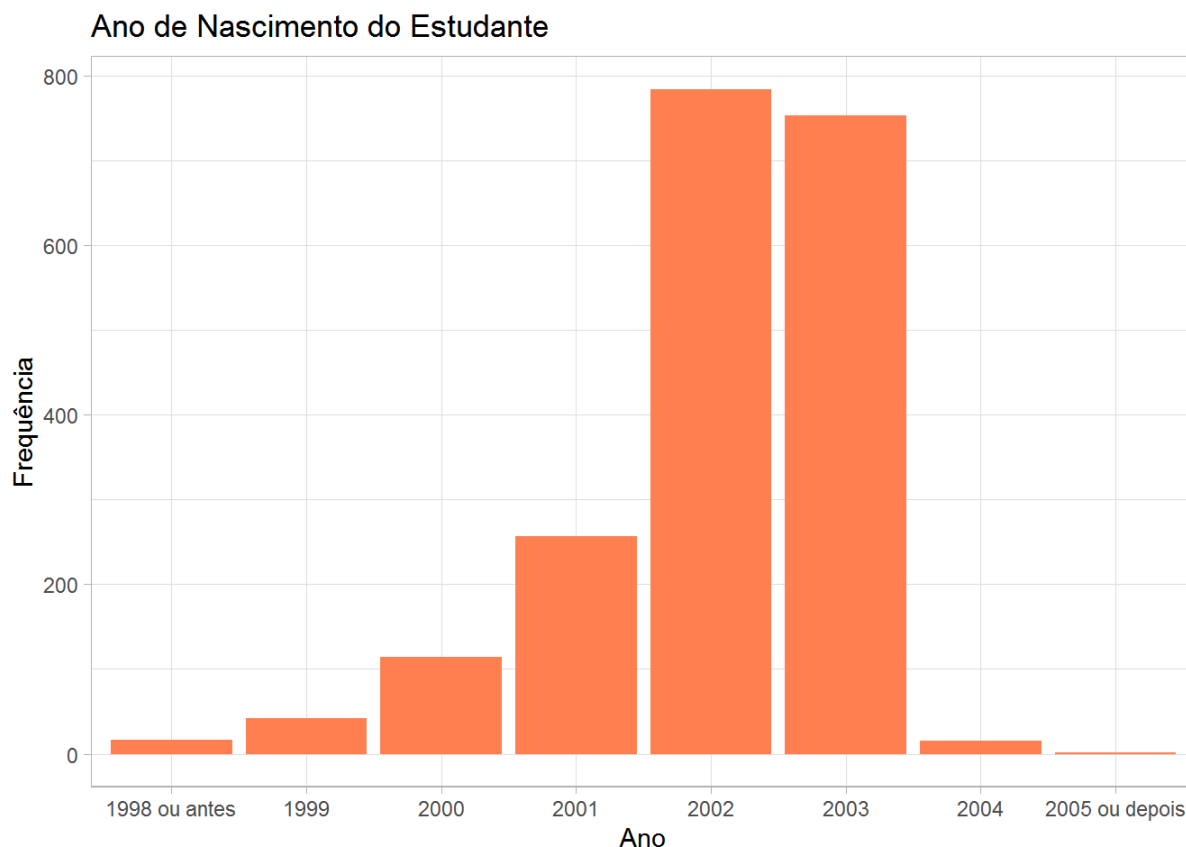
A distribuição dos meses de nascimento dos estudantes é relativamente homogênea. Seria curioso se fosse heterogênea, porque a princípio não há motivo para um mês ter mais ou menos nascimentos que o outro.

### 3.10 ANO\_NASC

**Tabela 10** – Distribuição de Estudantes quanto ao seu Ano de Nascimento. Saeb, 2017.

<b>Ano de Nascimento</b>	<b>Frequência Absoluta</b>	<b>Frequência Relativa (%)</b>	<b>Frequência Acumulada Abaixo de (%)</b>
1998 ou antes	16	0.80	0.80
1999	42	2.10	2.90
2000	114	5.70	8.60
2001	257	12.85	21.45
2002	784	39.20	60.65
2003	753	37.65	98.30
2004	15	0.75	99.05
2005 ou depois	2	0.10	99.15
NA	17	0.85	100.00
Total	2000	100.00	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



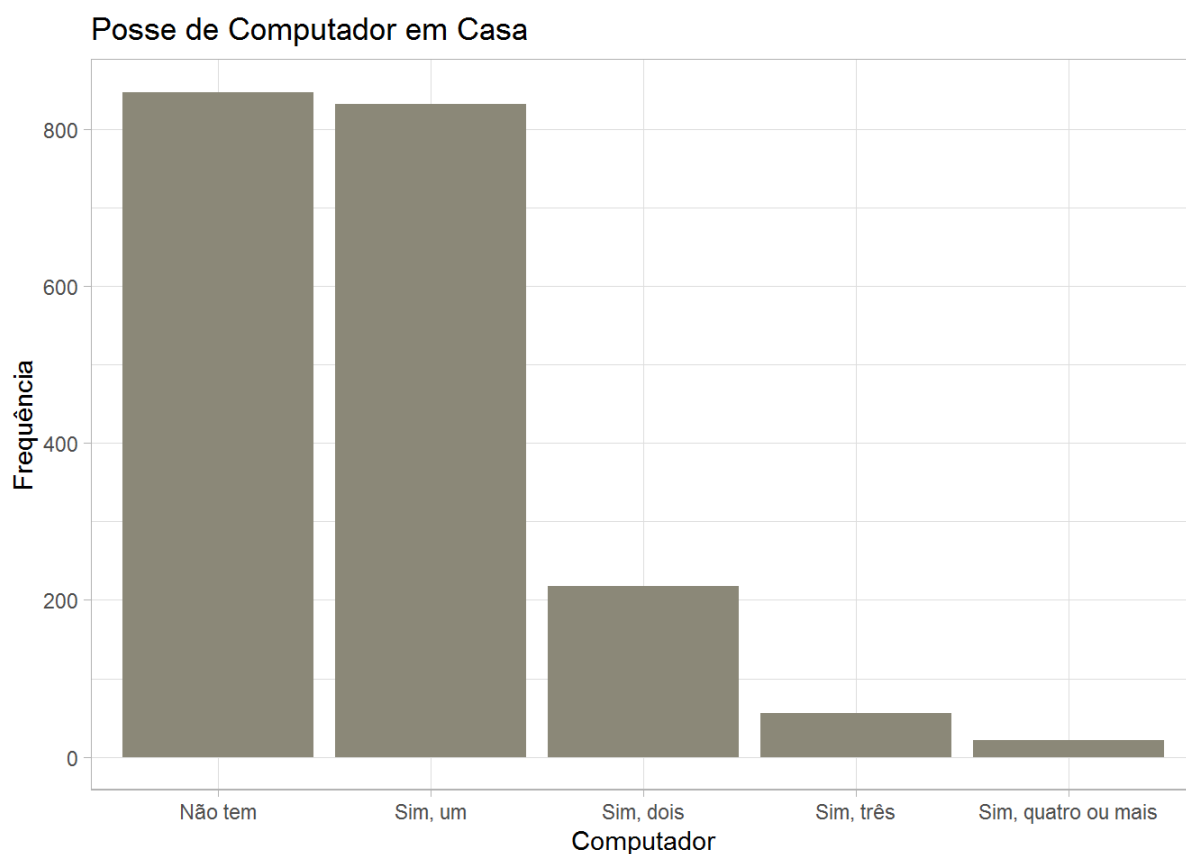
O mais comum é que os estudantes tenham nascido em 2002 ou 2003. Essa informação faz sentido, já que 14 anos é a idade padrão para um aluno do nono ano e a amostra é relativa a 2017. É raro existirem estudantes mais novos do que isso na oitava série, então já era esperado que as frequências a partir de 2004 fossem baixas. O surpreendentemente é que existem vários estudantes de até 17 anos (que nasceram em 2000 ou 2001) e que estavam nessa série quando a prova foi realizada. O gráfico indica, então, que é mais frequente haverem estudantes que reprovaram do que estudantes que pularam séries.

### 3.11 COMPUTADOR

**Tabela 11** – Distribuição de Estudantes quanto à Posse de Computador em Casa.  
Saeb, 2017.

<b>Computador</b>	<b>Frequência Absoluta</b>	<b>Frequência Relativa (%)</b>	<b>Frequência Acumulada Abaixo de (%)</b>
Não tem	848	42.40	42.40
Sim, um	833	41.65	84.05
Sim, dois	218	10.90	94.95
Sim, três	56	2.80	97.75
Sim, quatro ou mais	21	1.05	98.80
NA	24	1.20	100.00
Total	2000	100.00	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



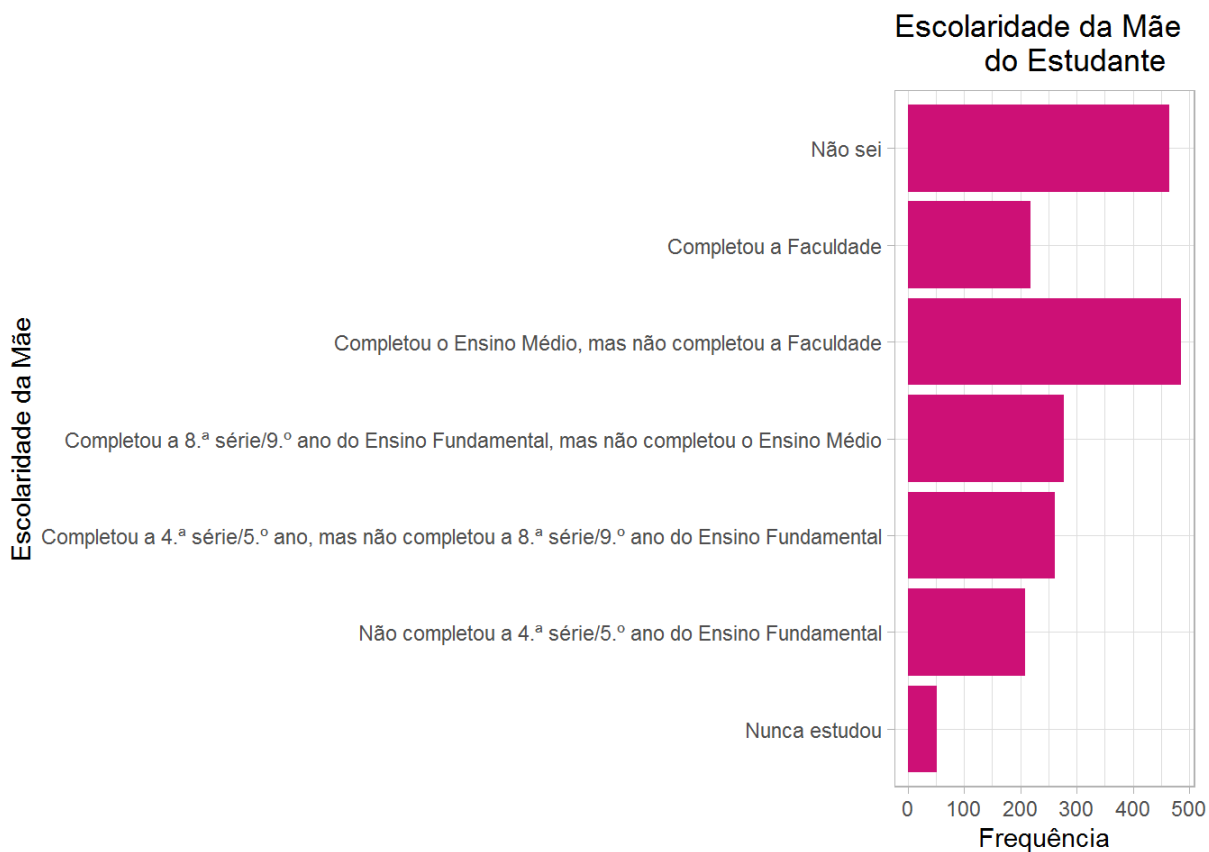
Quase metade dos estudantes (42,40%) não possui computador. O segundo maior grupo é o daqueles que possuem apenas um computador (41,65%). Apesar disso, quase 16% dos estudantes possui dois ou mais computadores. Essa disparidade é um reflexo da desigualdade de renda no Brasil.

### 3.12 ESC\_MAE

**Tabela 12** – Distribuição de Frequências quanto à Escolaridade da Mãe do Estudante. Saeb, 2017.

<b>Escolaridade da Mãe</b>	<b>Frequência Absoluta</b>	<b>Frequência Relativa (%)</b>
Nunca estudou	51	2.55
Não completou a 4. <sup>a</sup> série/ 5. <sup>o</sup> ano do Ensino Fundamental	209	10.45
Completou a 4. <sup>a</sup> série/ 5. <sup>o</sup> ano do Ensino Fundamental, mas não completou a 8. <sup>a</sup> série/ 9. <sup>o</sup> ano do Ensino Fundamental	262	13.10
Completou a 8. <sup>a</sup> série/ 9. <sup>o</sup> ano do Ensino Fundamental, mas não completou o Ensino Médio	277	13.85
Completou o Ensino Médio, mas não completou a faculdade	485	24.25
Completou a Faculdade	218	10.90
Não sei	465	23.25
NA	33	1.65
<b>Total</b>	<b>2000</b>	<b>100.00</b>

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



Para a variável escolaridade, o ideal seria que as escolaridades mais altas tivessem frequências maiores, mas isso não ocorre. Apenas 10,95% das mães completou a faculdade. Outro ponto que chama a atenção é que a segunda categoria mais comum é “Não Sei”, o que sugere que muitos estudantes têm pouco ou nenhum contato com suas mães.

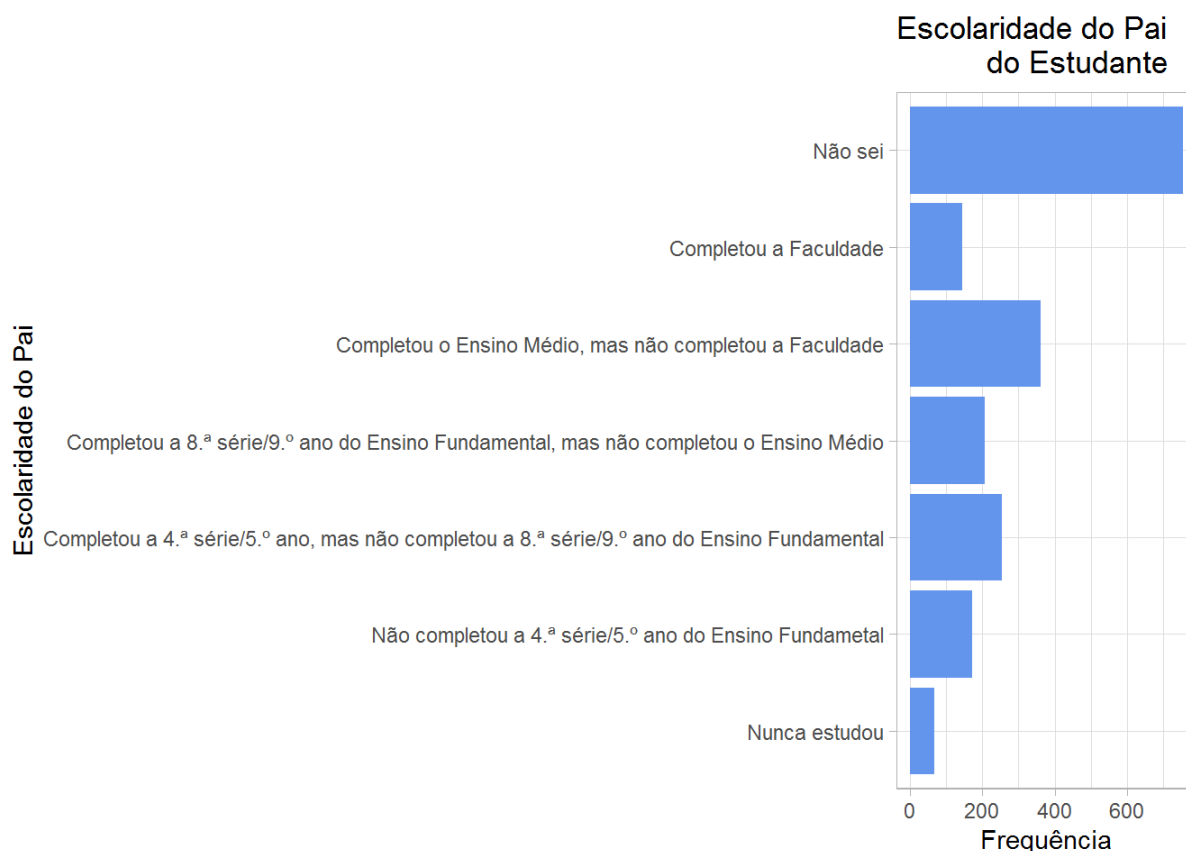
### 3.13 ESC\_PAi

**Tabela 13** – Distribuição de Frequências quanto à Escolaridade do Pai do Estudante. Saeb, 2017.

<b>Escolaridade do Pai</b>	<b>Frequência Absoluta</b>	<b>Frequência Relativa (%)</b>
Nunca estudou	67	3.35
Não completou a 4. <sup>a</sup> série/ 5. <sup>o</sup> ano do Ensino Fundamental	173	8.65
Completou a 4. <sup>a</sup> série/ 5. <sup>o</sup> ano do Ensino Fundamental, mas não completou a 8. <sup>a</sup> série/ 9. <sup>o</sup> ano do Ensino Fundamental	253	12.65
Completou a 8. <sup>a</sup> série/ 9. <sup>o</sup> ano do Ensino Fundamental, mas não completou o Ensino Médio	207	10.35
Completou o Ensino Médio, mas não completou a faculdade	360	18.00
Completou a Faculdade	145	7.25
Não sei	753	37.65
NA	42	2.10
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>





Entre os pais dos estudantes, também é pouco comum a finalização da faculdade. A proporção de estudantes que não sabe a escolaridade do pai é ainda maior (37,65%). Há um número maior de pais que completou a quarta série do que a oitava.

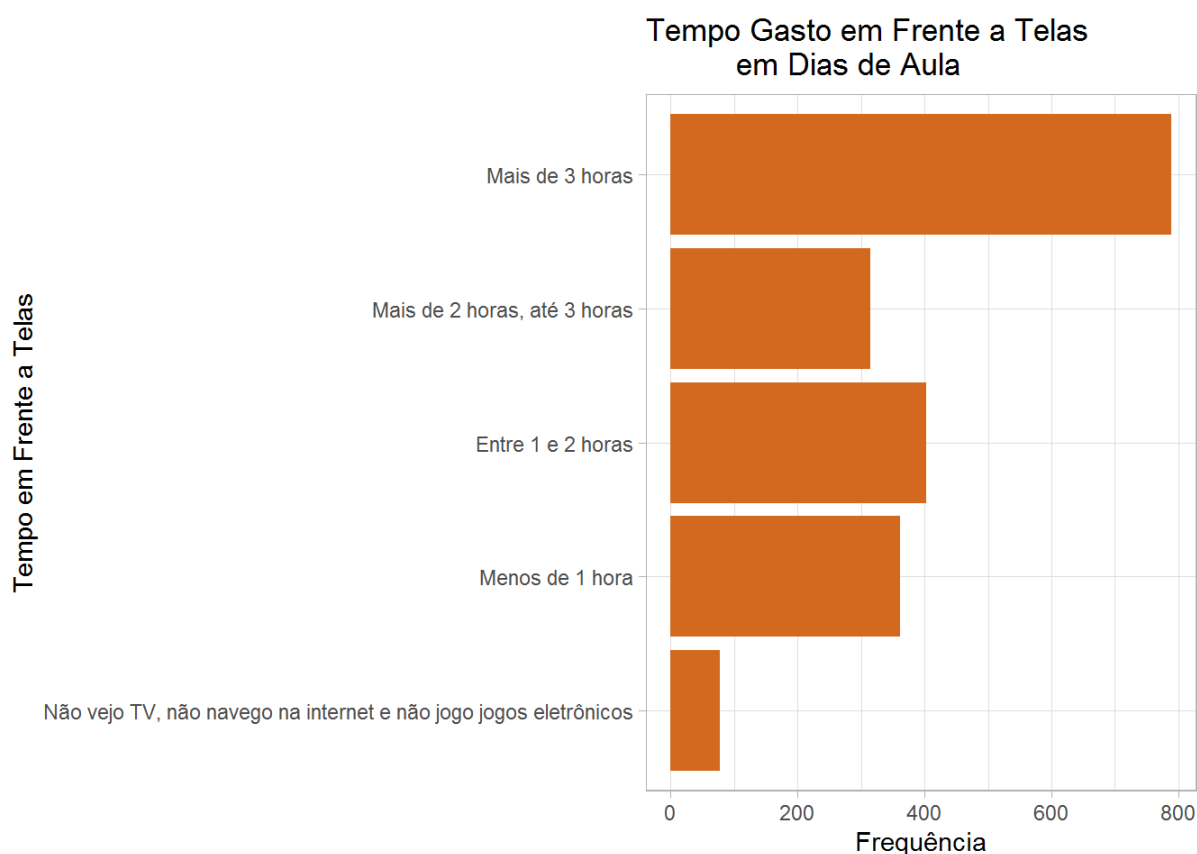
### 3.14 USO\_TEMPO\_TELAS

**Tabela 14** – Distribuição de Estudantes quanto ao Tempo Gasto em Frente a Telas. Saeb, 2017.

Tempo em Frente a Telas	Frequência Absoluta	Frequência Relativa (%)
Não vejo TV, não navego na internet e não jogo jogos eletrônicos	77	3.85
Menos de 1 hora	362	18.10

Entre 1 e 2 horas	403	20.15
Mais de 2 horas, até 3 horas	315	15.75
Mais de 3 horas	789	39.45
NA	54	2.70
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



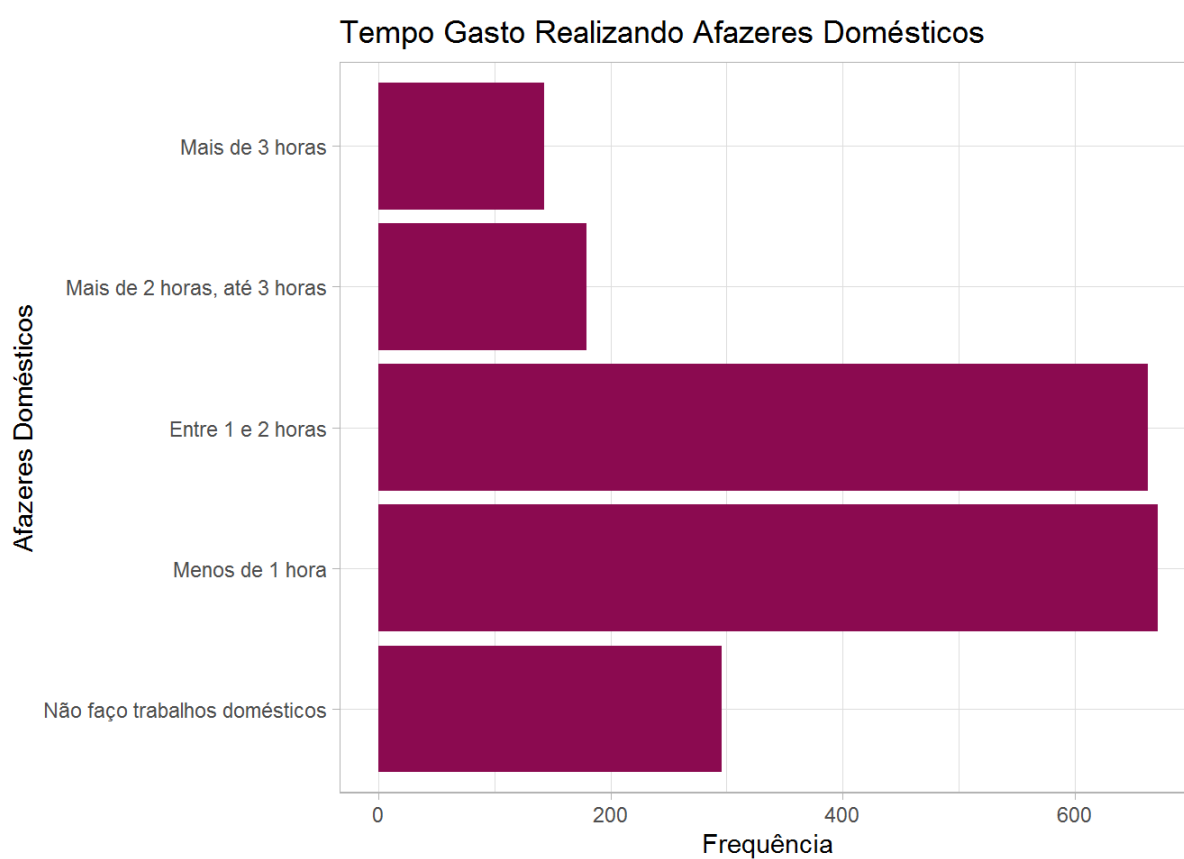
São poucos os estudantes que não tem nenhum contato com eletrônicos. A maioria passa bastante tempo em frente a telas nos dias de aula. Quase 40% gasta mais de 3 horas vendo TV, usando a internet ou jogando.

### 3.15 AFAZERES\_DOM

**Tabela 15** – Distribuição de Estudantes quanto ao Tempo Gasto Realizando Afazeres Domésticos. Saeb, 2017.

<b>Tempo Gasto em Afazeres Domésticos</b>	<b>Frequência Absoluta</b>	<b>Frequência Relativa (%)</b>
Não faço trabalhos domésticos	296	14.80
Menos de 1 hora	671	33.55
Entre 1 e 2 horas	663	33.15
Mais de 2 horas, até 3 horas	179	8.95
Mais de 3 horas	143	7.15
NA	48	2.40
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



Mais de 80% dos estudantes gastam até 2 horas com afazeres domésticos. São poucos os que gastam mais tempo com isso

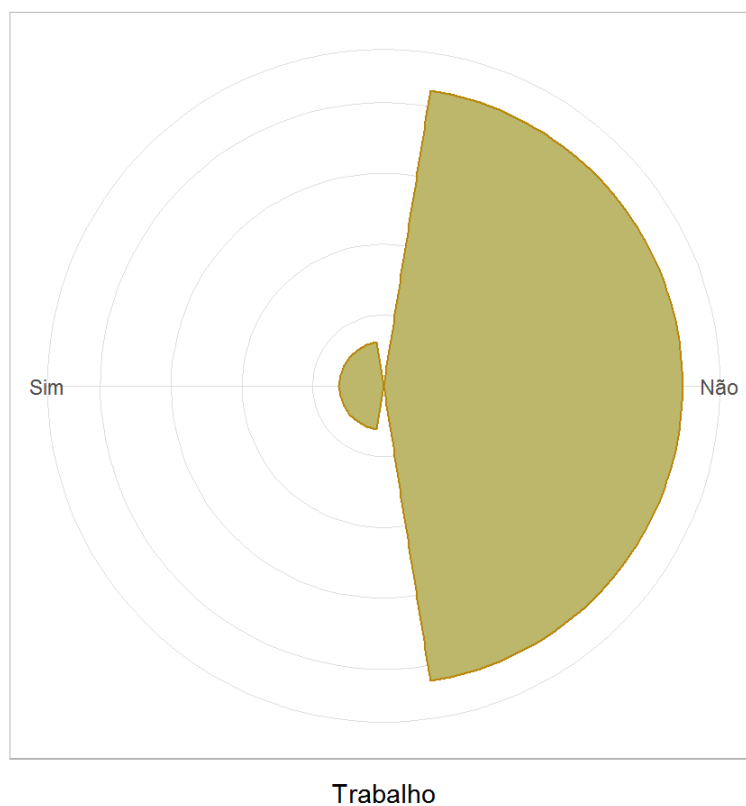
### 3.16 TRABALHO

**Tabela 16** – Distribuição de Frequência para Estudantes que Trabalham ou Não. Saeb, 2017.

Trabalho	Frequência Absoluta	Frequência Relativa (%)
Não	1689	84.45
Sim	251	12.55
NA	60	3.00
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>

Estudantes que Trabalham ou Não



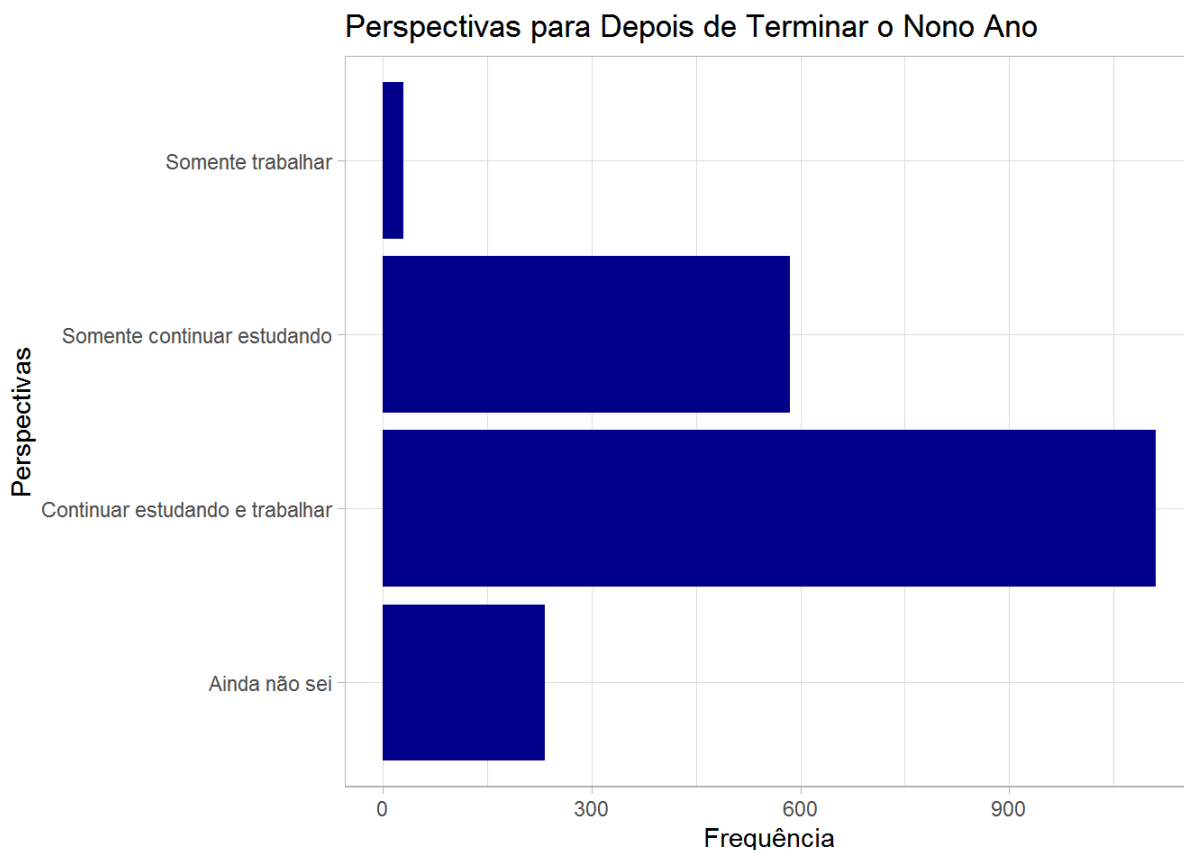
Apenas 251 entre os 2000 estudantes trabalham. Ou seja, é incomum que alunos do nono ano trabalhem. Talvez essa porcentagem aumente para os estudantes do Ensino Médio.

### 3.17 PERSPECTIVAS

**Tabela 17** – Distribuição de Estudantes quanto às suas perspectivas para depois de Finalizarem o Ensino Fundamental. Saeb, 2017.

Perspectivas	Frequência Absoluta	Frequência Relativa
Somente Trabalhar	29	1.45
Ainda não sei	233	11.65
Somente continuar estudando	585	29.25
Continuar estudando e trabalhar	1111	55.55
NA	42	2.10
Total	2000	100.00

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



Mais da metade dos estudantes pretende trabalhar além de estudar no Ensino Médio. 84,8% dos alunos planeja seguir estudando.

## 4 Análise Numérica

### 4.1 NOTA\_LP

```

12 |
13 | 045789
14 | 0112224444555677778999
15 | 0111233445555566677777788899
16 | 00000000011112222233333344455556667888888999
17 | 00000000011112223333333333334444444445555555566666677777888889999
18 | 000000011111222222333333333344445555566666666677777888888888999
19 | 00000001111112222222222333333333334444444444455555556666666+12
20 | 00000011111222222222223333333344444444444455555556666666667777+22
21 | 0001111112222222333333333344444444444455555555555666666677777+23
22 | 00000000111111112222222222222223333333344444444444555555556+43
23 | 000000001111111111122222222222333333333333444444444444444+62
24 | 0000000001111111111112222222222333333333333344444444455555555+54
25 | 000000000000011111111111111222222222222223333333333333333334+97
26 | 0000000000000000011111111111112222222222233333333333333333333+72
27 | 000000000000011111111111111122222222222223333333333333333334444+64
28 | 0000000000000000011111111112222222222222222223333333333333334444+65
29 | 0000000000000000011111111122222222222222333333344444444445555555+45
30 | 00000000000001111111222222222222333333333333333444444444444555555+30
31 | 0000000000011111111222222222222333344444444555555566666667777777+2
32 | 00000000111122233333334444555555566677888888888999
33 | 000000011122223334666777888
34 | 0011122233333445666677789
35 | 111111222233344555

```

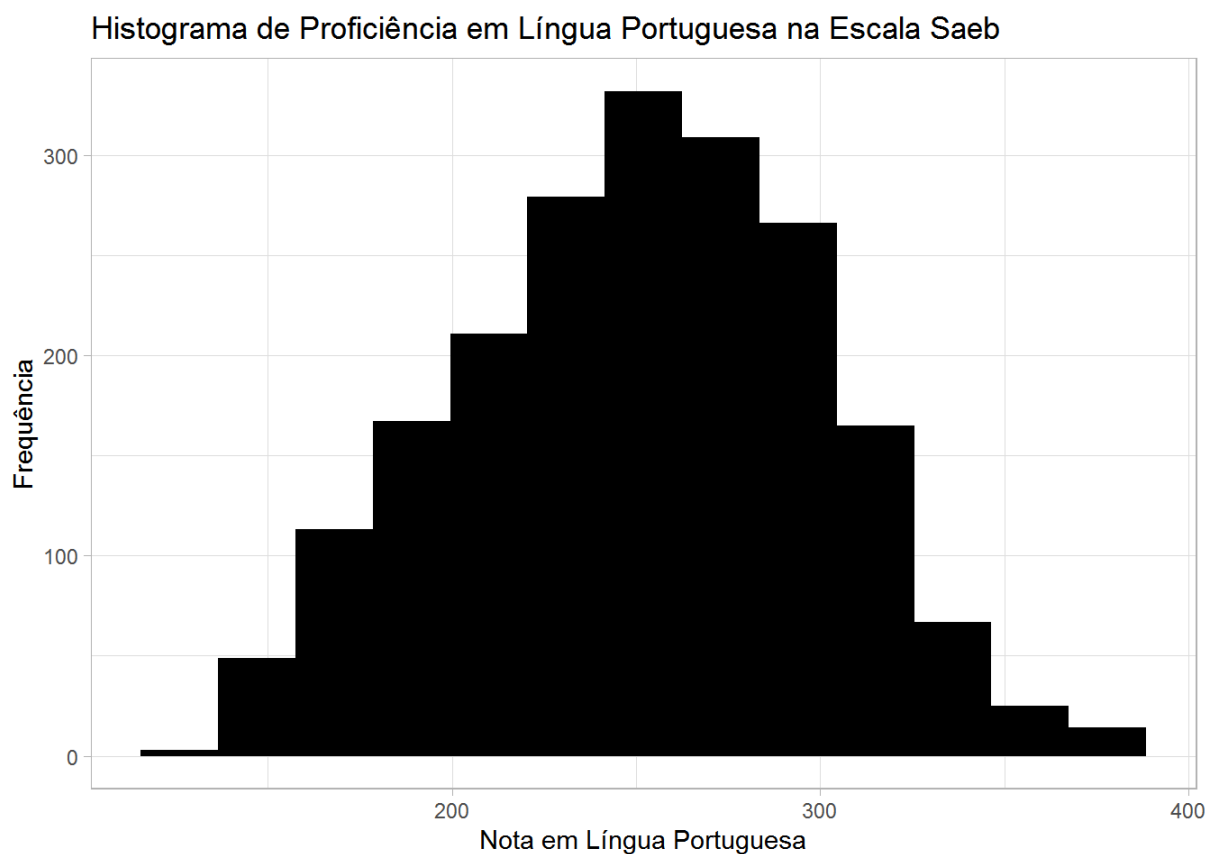
36 | 8899999

37 | 00000

**Tabela 18-** Distribuição de Frequências das Notas em Língua Portuguesa. Saeb, 2017.

<b>Intervalos de Classes das Notas de Língua Portuguesa</b>	<b>Frequência Absoluta</b>
(129, 150]	29
(150, 171]	85
(171, 192]	148
(192, 213]	193
(213, 234]	260
(234, 255]	320
(255, 276]	322
(276, 297]	278
(297, 318]	214
(318, 339]	91
(339, 360]	46
(360, 381]	14

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>

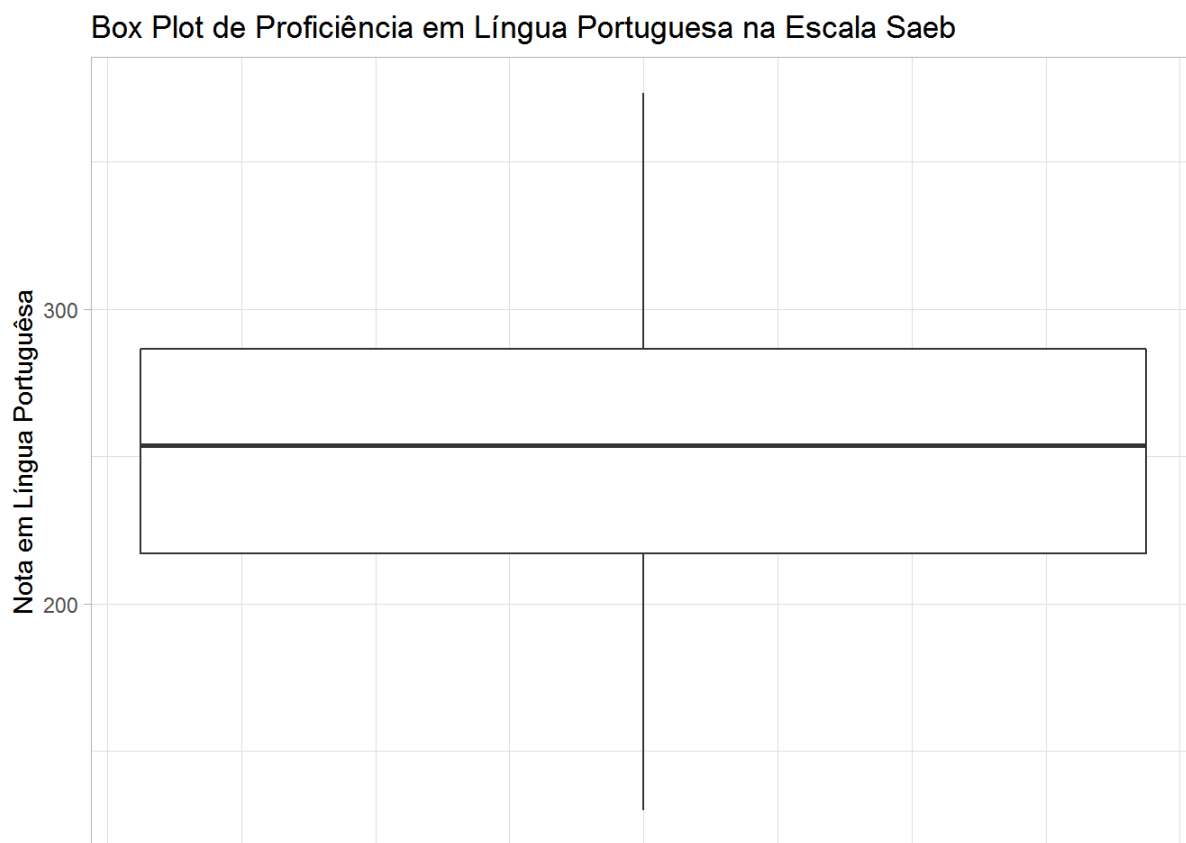


**Tabela 19-** Medidas Descritivas para as Notas em Língua Portuguesa. Saeb, 2017.

Medidas Descritivas	Valores
Média	250.998
Primeiro Quartil	216.931
Mediana/ Segundo Quartil	253.736
Terceiro Quartil	286.454
Quarto Quartil	373.400
Amplitude	243.405
Intervalo Interquartil	69.522
Variância	2330.598
Desvio Padrão	48.276
Coeficiente de Variação	0.192
Coeficiente de Assimetria de Pearson	-0.170
Coeficiente Quartil de Assimetria	-0.059
Coeficiente de Achatamento	0.271

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>





Observando o ramo e folhas das notas em língua portuguesa, percebe-se que a maior densidade de notas está entre os valores de 190 e 320. A tabela de frequências por intervalos de classes confirma isso, basta notar que mais de 1587 estudantes, entre os 2000, tiveram uma pontuação entre esses dois números. O histograma indica que as notas em português giram em torno de 250 e que há uma possível assimetria à esquerda, significando que, de modo geral, os estudantes tiram notas melhores em português. Por causa dessa assimetria, a distribuição não pode ser considerada normal, apesar de ter maior densidade no centro do que nas extremidades.

Os valores da média (251) e da mediana (253,7) ratificam o centro da distribuição observado no histograma. Também é confirmada a assimetria à esquerda pelo coeficiente de assimetria de Pearson e pelo coeficiente quartil de assimetria, que são ambos negativos. O desvio padrão de 48,3 aponta certa variabilidade nos dados. A partir do coeficiente de achatamento, pode-se classificar a distribuição como platicúrtica. No box plot percebe-se uma distância maior entre o primeiro quartil e a mediana, servindo como mais uma evidência de assimetria.

## 4.2 NOTA\_MT

```

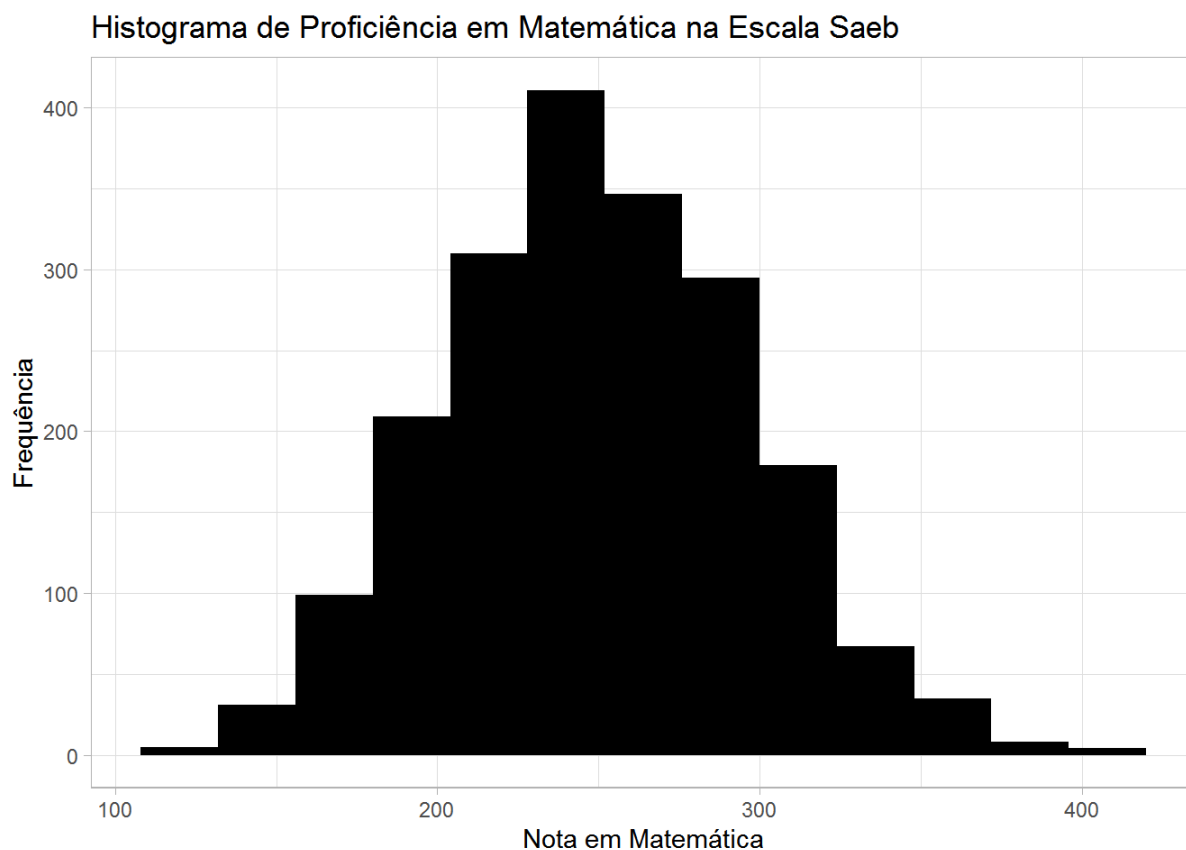
12 | 8801145777
14 | 0001113356772223344455556666677789999
16 | 00001112223344444466666667777888999900001111111222333334445556666677+1
18 | 00000000011111122222223333344455555555555566666777777778899999999+90
20 | 00000011111112222222222333333333333334444444444455555555556666+144
22 | 0000000000001111111111111111111222222222222222333333333333333444+245
24 | 000000000000000000111111111112222222222222222233333333333333333+268
26 | 000000000111111111112222222222222222222223333333333333333344444+193
28 | 0000000000000000000011111112222222222222333333333334444444444+168
30 | 00000011111111122222222223333333344444455555666666777777888888+75
32 | 00000000011222222333333444455555666677777889999900011122233333344+1
34 | 00123446667779999001122355556678899
36 | 01256678901114577
38 | 3333
40 | 2230

```

**Tabela 20-** Distribuição de Frequências das Notas em Matemática. Saeb, 2017.

Intervalos de Classes das Notas de Matemática	Frequência Absoluta
(127, 151]	22
(151, 175]	84
(175, 199]	184
(199, 223]	282
(223, 247]	403
(247, 271]	375
(271, 295]	316
(295, 319]	191
(319, 343]	88
(343, 367]	36
(367, 391]	13
(391, 415]	6

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>

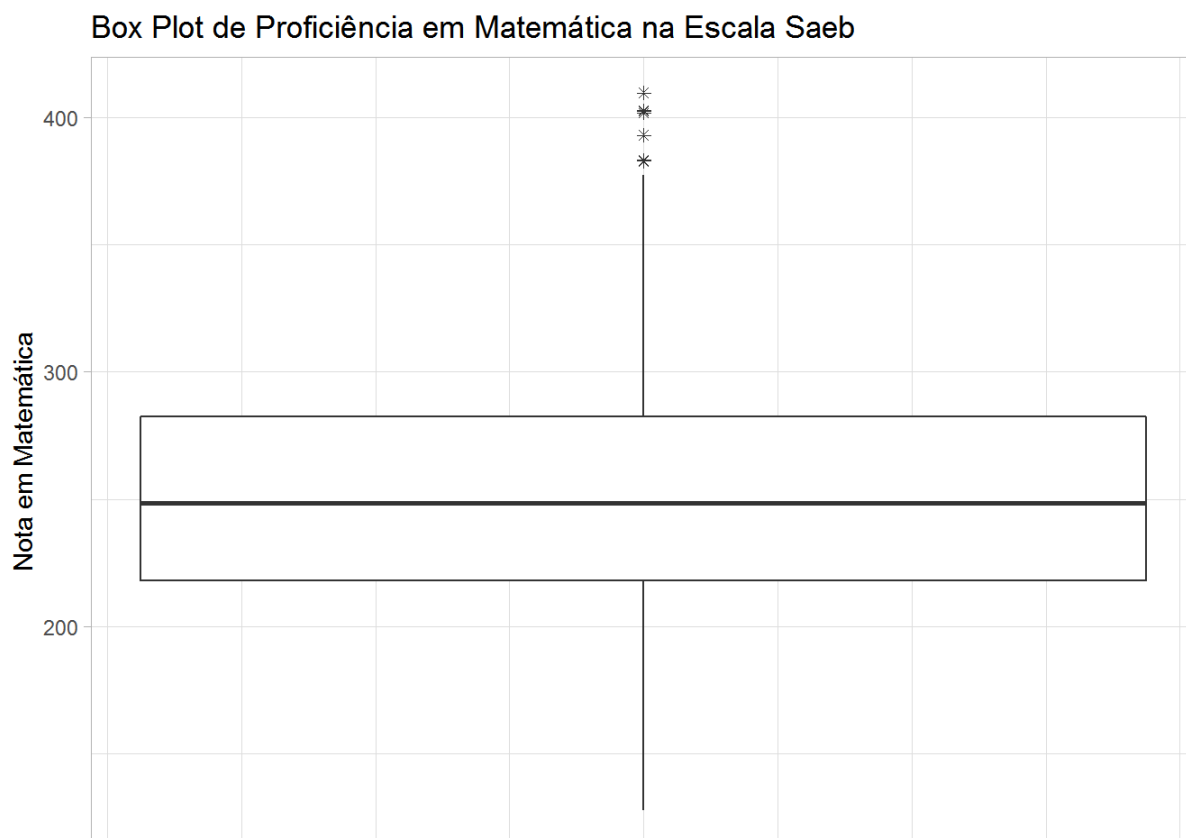


**Tabela 21-** Medidas Descritivas para as Notas em Matemática. Saeb, 2017.

Medidas Descritivas	Valores
Média	249.799
Primeiro Quartil	217.900
Mediana/ Segundo Quartil	248.528
Terceiro Quartil	282.406
Quarto Quartil	409.655
Amplitude	281.810
Intervalo Interquartil	64.506
Variância	2210.856
Desvio Padrão	47.02
Coeficiente de Variação	0.188

Coeficiente de Assimetria de Pearson	0.081
Coeficiente Quartil de Assimetria	0.050
Coeficiente de Achatamento	0.266

Fonte: <http://portal.inep.gov.br/educacao-basica/saeb>



O ramo e folhas das notas em matemática indica que as notas se concentram principalmente entre 160 e 320. E de fato, pela tabela é possível ver que as classes mais frequentes são as de 175 a 199 até 295 a 319. O histograma mostra evidências de uma distribuição assimétrica à direita, centrada um pouco abaixo de 250. A assimetria revela que, de forma geral, os estudantes tiram notas piores em matemática. Assim como as notas em língua portuguesa, as de matemática não formam uma distribuição normal, porque existe essa assimetria. Talvez com uma amostra maior, essas variáveis possam ser aproximadas por uma normal.

A média de 249,8, a mediana de 217,9, e os coeficientes de assimetria positivos validam as evidências observadas no histograma. O coeficiente de variação (0,188) quando comparado com o da distribuição das notas em língua

portuguesa (0,192) acusa uma variabilidade levemente menor. Pode-se concluir também que a distribuição é quase mesocúrtica, mas ainda platicúrtica, olhando para a medida de curtose (0,266). Novamente a assimetria se revela no box plot, tanto pela distância da mediana até o terceiro quartil, como pela distância do terceiro quartil até o valor máximo. Além disso, podemos identificar valores extremos pelos asteriscos no gráfico, os quais representam aqueles estudantes com proficiência excepcional em matemática.

## 5 Considerações Finais

O Sistema de Avaliação da Educação Básica coleta dados dos estudantes para poder analisar, tirar conclusões e, se possível, usar esses resultados para criar e implementar medidas para a melhoria do ensino brasileiro.

Além das variáveis numéricas referentes ao desempenho dos estudantes na prova, são, também, de extrema importância as informações geográficas das localizações das escolas, para se identificar onde é necessária a maior atenção no momento de planejar e aplicar políticas.

Também é importante investigar possíveis variáveis de explicação para o desempenho dos estudantes. Para isso são coletadas informações pessoais do aluno, incluindo condições de estudo (ter ou não um computador, gastar tempo com trabalho ou afazeres domésticos, etc.).

Em suma, este relatório foi feito com o objetivo de expor os resultados de uma exploração inicial das variáveis. Para se obter informações mais substanciais, é necessária uma análise aprofundada, incluindo estudo de associação e uso de modelos estatísticos.