



Allan Victor Almeida Faria, Hermes Jasper Winarski, Kassyano Kevyn Andrade  
de Souza

## **Relatório de Pesquisa**

Brasília, DF

Dezembro, 2020



Allan Victor Almeida Faria, Hermes Jasper Winarski, Kassyano Kevyn Andrade  
de Souza

## **Relatório de Pesquisa**

Relatório de Pesquisa do projeto Saeb 2017.

Universidade de Brasília (UnB)  
Instituto de Ciências Exatas (IE)  
Departamento de Estatística (DE)

Brasília, DF  
Dezembro, 2020

# Resumo

blablalbalbalbalbalbalblablalbbbbbbblbalbalbalblalbalbalblalbalblaba blablalbalbalbalbal-  
balblablalbbbbbbblbalbalbalblalbalbalblalbalblaba blablalbalbalbalbalbalblablalbbbbbbbl-  
balbalbalblalbalbalblalbalblabal albalblablalblalbla

**Palavras-chaves:** 1. Saeb 2017. 2. Amostra. 3. Pesquisa.

# Lista de ilustrações

Figura 1 – Proporção por sexo de tempo em afazeres domésticos por parte dos alunos. . . . .	13
Figura 2 – Proporção total por nível de escolaridade da mãe com base no tempo de afazeres dos alunos. . . . .	14
Figura 3 – Distribuições das somas das notas com base na raça/cor dos alunos. . .	18
Figura 4 – Distribuições empíricas das somas das notas com base nas localizações das das escolas dos alunos. . . . .	19
Figura 5 – Distribuições das somas das notas com base nas escolaridades das mães dos alunos. . . . .	20

# Lista de tabelas

Tabela 1 –	Comparações dois a dois entre as ordens das posições sobre os tempos de afazeres domésticos com base na escolaridade das mães dos alunos. .	15
Tabela 2 –	Testes de igualdade na variabilidade sobre as relações com o tempo de afazeres domésticos por parte dos alunos. . . . .	16
Tabela 3 –	Testes para as relações com soma das notas dos alunos. . . . .	21
Tabela 4 –	Comparações dois a dois entre as médias sobre a soma das notas com base na raça/cor dos alunos. . . . .	22
Tabela 5 –	Comparações entre as médias de notas em Matemática e as regiões das escolas dos alunos com base na amostra de tamanho 500. . . . .	23

# Lista de abreviaturas e siglas

INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
SAEB	Sistema de Avaliação da Educação Básica

# Lista de símbolos

$\sigma$  Letra grega minúscula sigma

$\mu$  Letra grega minúscula mu

# Sumário

	<b>Introdução</b>	<b>8</b>
<b>I</b>	<b>FUNDAMENTAÇÃO</b>	<b>9</b>
<b>1</b>	<b>METODOLOGIA</b>	<b>10</b>
<b>II</b>	<b>RESULTADOS</b>	<b>12</b>
<b>2</b>	<b>COMPARAÇÕES</b>	<b>13</b>
<b>2.1</b>	<b>Tempo em afazeres domésticos</b>	<b>13</b>
<b>2.2</b>	<b>Notas</b>	<b>17</b>
<b>3</b>	<b>CONCLUSÃO</b>	<b>24</b>
	<b>ANEXOS</b>	<b>25</b>
	<b>ANEXO A – AMOSTRA</b>	<b>26</b>





## Introdução

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Nisl tincidunt eget nullam non nisi est sit amet. Turpis nunc eget lorem dolor. Posuere urna nec tincidunt praesent semper feugiat nibh. Platea dictumst quisque sagittis purus sit amet volutpat consequat mauris. Velit scelerisque in dictum non consectetur a. Lectus proin nibh nisl condimentum. Nulla pharetra diam sit amet.

Risus in hendrerit gravida rutrum quisque non tellus. Purus gravida quis blandit turpis cursus in hac habitasse. Et malesuada fames ac turpis egestas maecenas pharetra convallis posuere. Id donec ultrices tincidunt arcu non sodales neque. Suspendisse faucibus interdum posuere lorem ipsum dolor sit amet consectetur. Amet purus gravida quis blandit. Sit amet nisl purus in mollis nunc sed. Rutrum quisque non tellus orci. Pellentesque eu tincidunt tortor aliquam nulla facilisi cras fermentum. Lacus suspendisse faucibus interdum posuere lorem ipsum dolor sit amet.

A iaculis at erat pellentesque adipiscing commodo elit. Tellus rutrum tellus pellentesque eu tincidunt tortor aliquam nulla. Eros donec ac odio tempor orci dapibus ultrices. Vitae ultricies leo integer malesuada. Aliquam eleifend mi in nulla posuere sollicitudin aliquam. Vitae purus faucibus ornare suspendisse sed nisi lacus sed.

Consequat interdum varius sit amet mattis. Elementum pulvinar etiam non quam lacus suspendisse faucibus. Morbi non arcu risus quis varius quam quisque id diam. Est ante in nibh mauris cursus mattis molestie a iaculis. Odio aenean sed adipiscing diam donec adipiscing tristique risus nec. Mi tempus imperdiet nulla malesuada pellentesque. Pellentesque habitant morbi tristique senectus et netus et. Luctus venenatis lectus magna fringilla urna porttitor rhoncus. Integer enim neque volutpat ac tincidunt vitae. Viverra ipsum nunc aliquet bibendum.

Parte I

FUNDAMENTAÇÃO



# 1 Metodologia

Este documento visa analisar o aprendizado de alunos do brasileiros no ano de 2017, e como fatores sociais podem influenciar nesse aprendizado. Para a análise, serão utilizados dados do Sistema de Avaliação da Educação Básica (SAEB) de 2017 divulgado pelo ??).

A pesquisa é feita por meio de testes e questionários aplicados bienalmente ao na rede pública e numa amostra da rede privada. A amostra é amostra composta pelo 5º e 9º anos do ensino fundamental, e também pelo 3º ano do ensino médio. Além do desempenho em matemática e língua portuguesa, por meio da prova, são coletadas informações a respeito dos alunos, tais como:

- Dependência administrativa da escola (se é federal, estadual, municipal ou privada);
- Grau de escolaridade dos pais;
- Taxas de reprovação;
- Tempo diário com afazeres domésticos;
- Tempo diário de uso de telas (tais como computadores, celulares e televisões);
- Se o estudante trabalha;
- Quais as suas perspectivas;
- Acesso a computadores;
- Raça/cor.

A análise parte de uma amostra aleatória simples de 5.271 alunos do 9º desse banco de dados.

As variáveis resposta serão a soma das notas em língua portuguesa e matemática e o tempo de afazeres domésticos realizados por dia por cada aluno.

Como variáveis explicativas para a soma das notas, o estudo relaciona a raça/cor, a escolaridade da mãe e as localizações das escolas, e avalia o quanto essas variáveis influenciam no desempenho na Prova Brasil. Para avaliar estas relações, um estudo prévio



foi realizado com amostras de tamanhos 30, 50 e 100 dos 5271 alunos, no qual o com os testes ??), ??) e ??), onde se concluiu que as distribuições das notas são normais. Este resultado será utilizado para atender os testes paramétricos que pressupõem normalidade dos dados. Os testes a serem aplicados serão o teste ANOVA dado por ??) para avaliar os fatores com mais de 2 categorias e o teste t-Student para as comparações dois a dois (??). Estes testes utilizam das médias ( $\mu$ ) de cada grupo para avaliar as distâncias significativas entre eles, no qual previamente avalia se as variâncias ( $\sigma^2$ ) destes são iguais (Teste B), através do teste proposto por ??).

Quanto ao tempo gasto com afazeres domésticos, as variáveis a serem relacionadas são a raça/cor, a escolaridade da mãe e o sexo, para observar se há indícios de diferenças sociais entre a disponibilidade de tempo em casa para outras possíveis tarefas na formação do aprendizado básico. Para estas relações, os testes estatísticos não paramétricos são apropriados, no qual utiliza-se o teste de ??) para fatores com mais de duas categorias (Teste K), e o teste de ??) para a comparação dois a dois das categorias (Teste W). Estes testes avaliam as distribuições das informações com base na posição, para verificar se as distâncias entre as categorias são significativas.

Para avaliar os resultados dos testes, foi proposto o uso da correção de ??) para os testes com mais de duas categorias. Se houver a evidências para rejeitar igualdade destas, a comparação dois a dois é efetuada e a mesma correção é utilizada. Esta correção é sobre o p-valor, que é avaliado em uma escala de significância, adotado por este estudo como uma confiança de 95%, no que diz sobre aceitar a hipótese nula  $H_0$ .

Parte II

Resultados

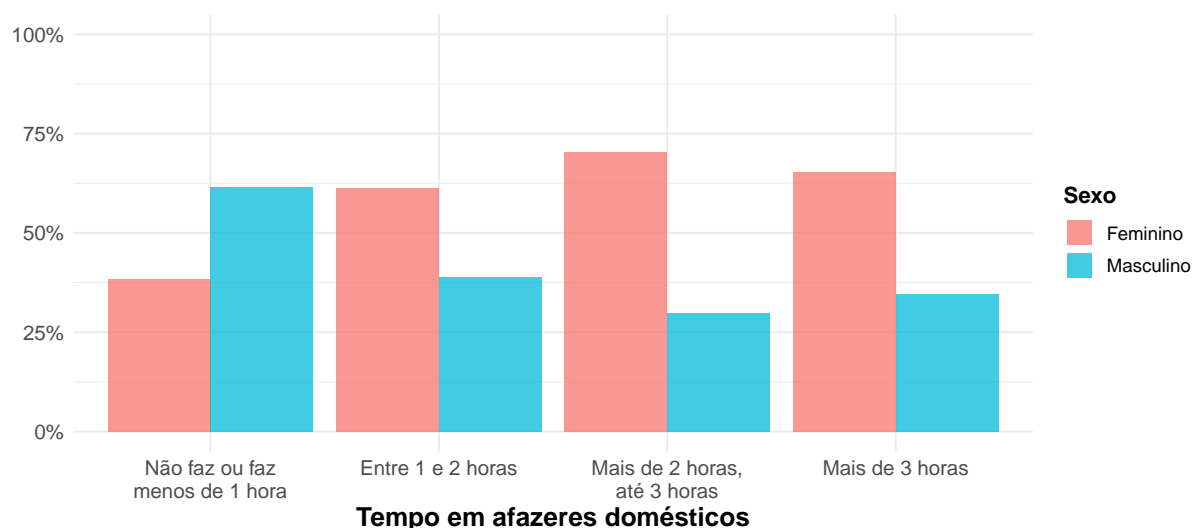
## 2 Comparações

### 2.1 Tempo em afazeres domésticos

Dado que o tempo gasto com afazeres domésticos pode privar um aluno de atividades educacionais, foram feitas análises de diferenças entre esses tempos.

A Figura 2.1 mostra como se distribuem esses tempos em cada sexo.

Figura 1 – Proporção por sexo de tempo em afazeres domésticos por parte dos alunos.



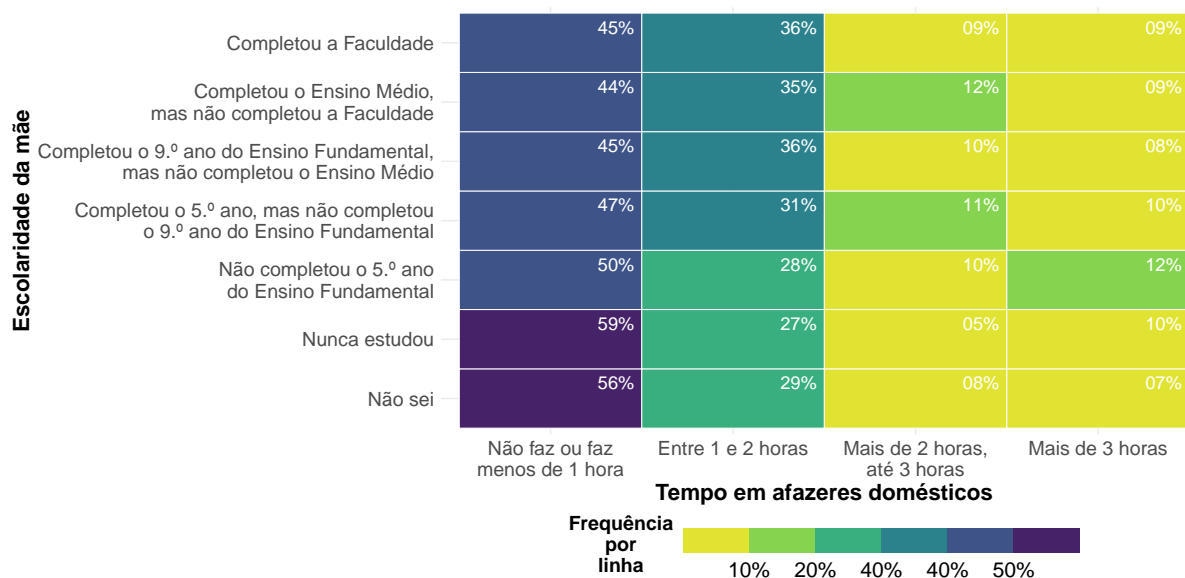
Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.

É fácil observar que estudantes do sexo feminino tendem a gastar mais tempo com atividades domésticas. A única categoria em que o sexo feminino teve menos representatividade foi justamente a categoria onde menos tempo diário é gasto com atividades domésticas. Para todas as outras categorias, onde mais de uma hora por dia é gasta, o sexo feminino representa pelo menos 60% de todos os indivíduos, o que indica que estudantes de sexo masculino têm mais disponibilidade de tempo para estudo.



Figura 2 – Proporção total por nível de escolaridade da mãe com base no tempo de afazeres dos alunos.



Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.

Ao comparar com a escolaridade da mãe, a proporção de alunos que exercem esses afazeres, para cada nível escolar, há 70% localizado entre os que não fazem ou fazem até 2 horas.



Tabela 1 – Comparações dois a dois entre as ordens das posições sobre os tempos de afazeres domésticos com base na escolaridade das mães dos alunos.

Comparações	P-valor	Evidência (RA 95%)
Não sabe = Nunca estudou	Aprox. 0	Desiguais
Não sabe = Incompleto 5.º ano do EF	1.0000	Iguais
Não sabe = Completou 5.º ano do EF	0.0084	Desiguais
Não sabe = Completou 9.º ano do EF	0.1927	Iguais
Não sabe = Completou EM	Aprox. 0	Desiguais
Não sabe = Completou Faculdade	Aprox. 0	Desiguais
Nunca estudou = Incompleto 5.º ano do EF	0.0038	Desiguais
Nunca estudou = Completou 5.º ano do EF	Aprox. 0	Desiguais
Nunca estudou = Completou 9.º ano do EF	Aprox. 0	Desiguais
Nunca estudou = Completou EM	Aprox. 0	Desiguais
Nunca estudou = Completou Faculdade	Aprox. 0	Desiguais
Incompleto 5.º ano do EF = Completou 5.º ano do EF	0.0002	Desiguais
Incompleto 5.º ano do EF = Completou 9.º ano do EF	0.0048	Desiguais
Incompleto 5.º ano do EF = Completou EM	Aprox. 0	Desiguais
Incompleto 5.º ano do EF = Completou Faculdade	Aprox. 0	Desiguais
Completo 5.º ano do EF = Completou 9.º ano do EF	1	Iguais
Completo 5.º ano do EF = Completou EM	Aprox. 0	Desiguais
Completo 5.º ano do EF = Completou Faculdade	Aprox. 0	Desiguais
Completo 9.º ano do EF = Completou EM	Aprox. 0	Desiguais
Completo 9.º ano do EF = Completou Faculdade	Aprox. 0	Desiguais
Completou EM = Completou Faculdade	1.0000	Iguais

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples.

Anotações: Aprox. 0 refere-se à algum número muito pequeno considerando aproximadamente zero.





A Tabela 2 mostra resultados dos testes de igualdade nas variabilidades dos tempos gastos com atividades domésticas entre raças, escolaridade da mãe e sexos.

Tabela 2 – Testes de igualdade na variabilidade sobre as relações com o tempo de afazeres domésticos por parte dos alunos.

Teste	$H_0$	P-valor	Decisão de $H_0$ (95%)
K	$\mu_{Raça/Cor}$ iguais	0.369	Aceita
K	$\mu_{Esc(mãe)}$ iguais	Aprox. 0	Rejeita
W	$\mu_M = \mu_F$ iguais	Aprox. 0	Rejeita

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples.

Anotações: Os subíndices M e F referem-se, respectivamente, aos sexos Masculino e Feminino dos alunos. Esc (mãe) diz respeito à escolaridade da mãe destes. O Aprox. 0 refere-se a um número muito pequeno considerado por este estudo aproximadamente zero.



## 2.2 Notas

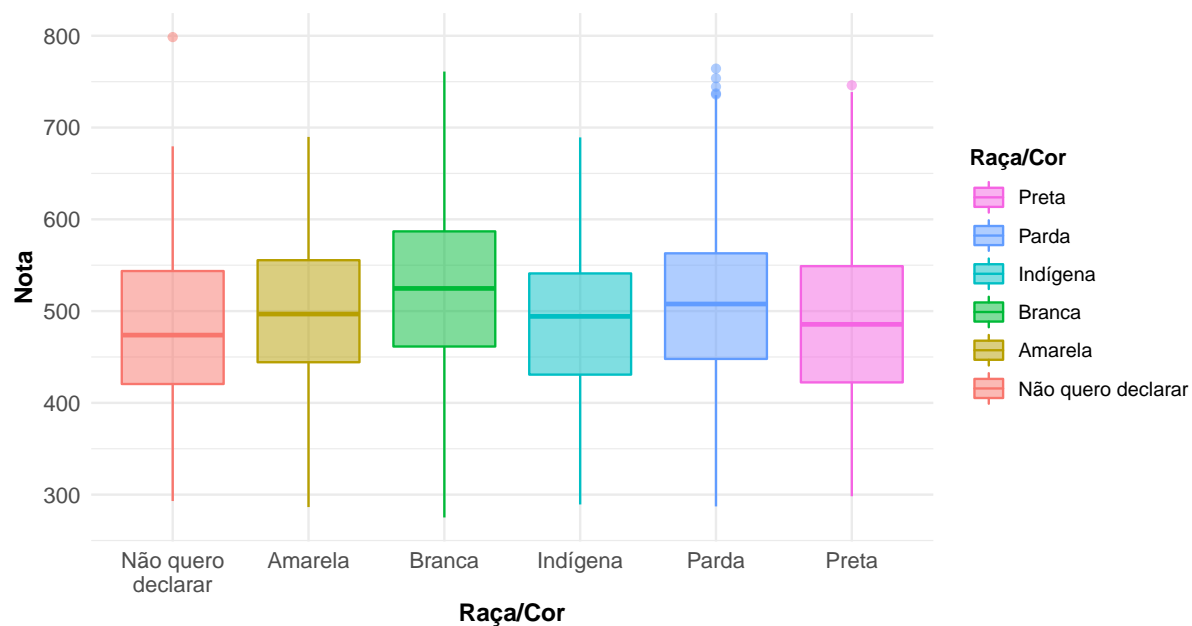
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Nisl tincidunt eget nullam non nisi est sit amet. Turpis nunc eget lorem dolor. Posuere urna nec tincidunt praesent semper feugiat nibh. Platea dictumst quisque sagittis purus sit amet volutpat consequat mauris. Velit scelerisque in dictum non consectetur a. Lectus proin nibh nisl condimentum. Nulla pharetra diam sit amet.

Risus in hendrerit gravida rutrum quisque non tellus. Purus gravida quis blandit turpis cursus in hac habitasse. Et malesuada fames ac turpis egestas maecenas pharetra convallis posuere. Id donec ultrices tincidunt arcu non sodales neque. Suspendisse faucibus interdum posuere lorem ipsum dolor sit amet consectetur. Amet purus gravida quis blandit. Sit amet nisl purus in mollis nunc sed. Rutrum quisque non tellus orci. Pellentesque eu tincidunt tortor aliquam nulla facilisi cras fermentum. Lacus suspendisse faucibus interdum posuere lorem ipsum dolor sit amet.

A iaculis at erat pellentesque adipiscing commodo elit. Tellus rutrum tellus pellentesque eu tincidunt tortor aliquam nulla. Eros donec ac odio tempor orci dapibus ultrices. Vitae ultricies leo integer malesuada. Aliquam eleifend mi in nulla posuere sollicitudin aliquam. Vitae purus faucibus ornare suspendisse sed nisi lacus sed.

Consequat interdum varius sit amet mattis. Elementum pulvinar etiam non quam lacus suspendisse faucibus. Morbi non arcu risus quis varius quam quisque id diam. Est ante in nibh mauris cursus mattis molestie a iaculis. Odio aenean sed adipiscing diam donec adipiscing tristique risus nec. Mi tempus imperdiet nulla malesuada pellentesque. Pellentesque habitant morbi tristique senectus et netus et. Luctus venenatis lectus magna fringilla urna porttitor rhoncus. Integer enim neque volutpat ac tincidunt vitae. Viverra ipsum nunc aliquet bibendum.

Figura 3 – Distribuições das somas das notas com base na raça/cor dos alunos.



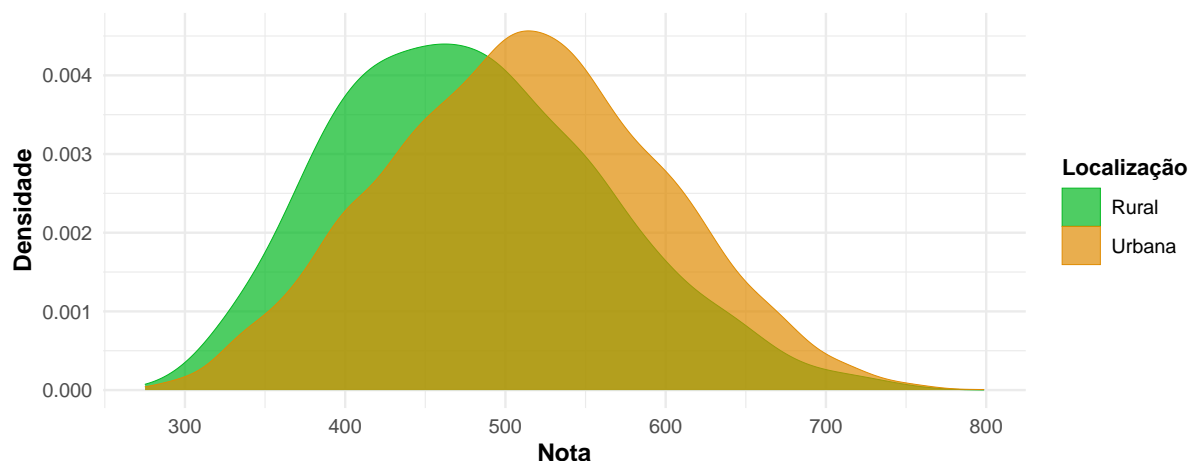
Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.



A Figura 2.2 mostra as distribuições das somas das notas por localização.

Figura 4 – Distribuições empíricas das somas das notas com base nas localizações das das escolas dos alunos.



Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.

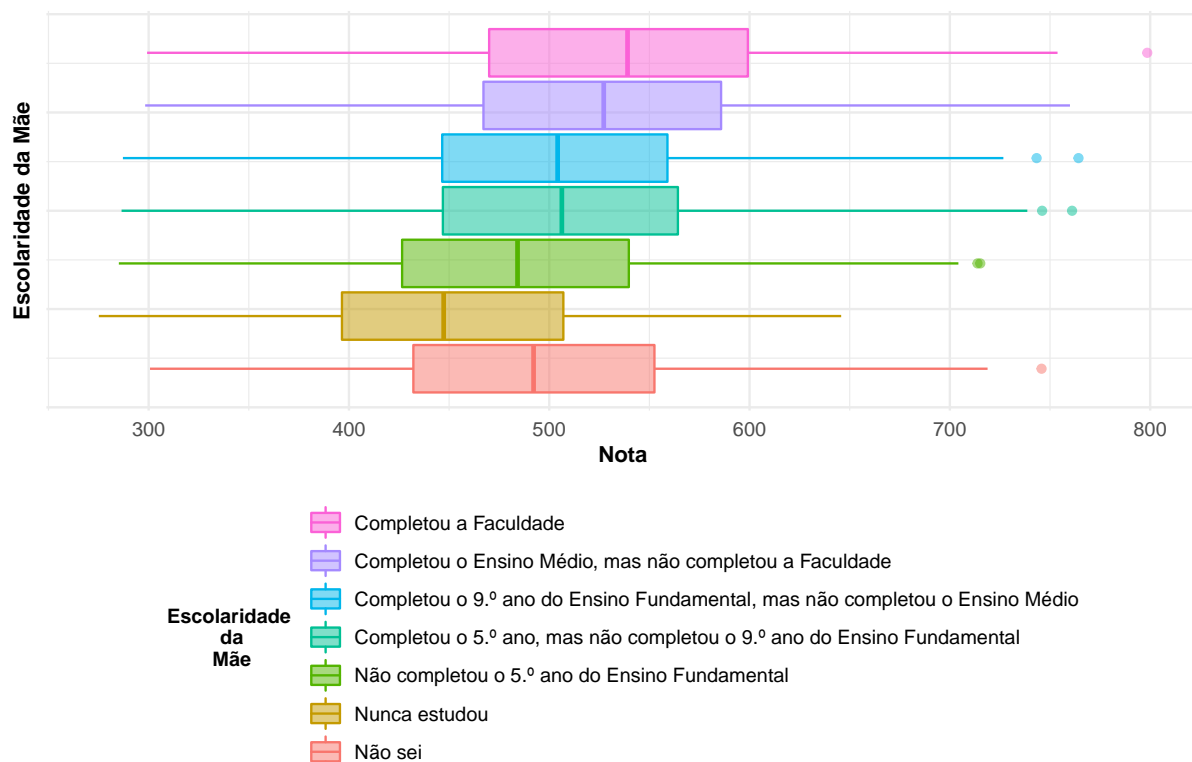
É possível observar um grau de assimetria um pouco maior nas notas para a região rural em relação à região urbana. A zona rural teve sua distribuição um pouco mais inclinada para a esquerda, enquanto a zona urbana foi mais centralizada. É possível observar que a moda das notas foi inferior na região rural em relação à região urbana.

A área rural apresentou uma nota média de 479, que também é inferior à nota média da área urbana, que foi de 512.



A Figura 5 mostra a distribuição das notas em relação à escolaridade das mães.

Figura 5 – Distribuições das somas das notas com base nas escolaridades das mães dos alunos.



Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.

É possível observar um crescimento das notas no geral à medida que o grau de escolaridade das mães é maior, de modo que filhos cujas mães têm grau de escolaridade alto têm tendência a terem melhor desempenho em provas. Esse comportamento é observado mesmo entre os outliers, onde a maior nota registrada vem da parte de um aluno cuja mãe completou a faculdade.



Tabela 3 – Testes para as relações com soma das notas dos alunos.

Teste	$H_0$	P-valor	Decisão de $H_0$ (95%)
B	$\sigma_R^2 = \sigma_U^2$	0.503	Aceita
B	$\sigma_{Raça/Cor}^2$ iguais	0.265	Aceita
B	$\sigma_{Esc(mãe)}^2$ iguais	0.132	Aceita
T	$\mu_R = \mu_U$ iguais	Aprox. 0	Rejeita
T	$\mu_M = \mu_D$	0.905	Aceita
ANOVA	$\mu_{Raça/Cor}$ iguais	Aprox. 0	Rejeita
ANOVA	$\mu_{Esc(mãe)}$ iguais	Aprox. 0	Rejeita

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples.

Anotações: Os subíndices, com base nos alunos,  $R$  e  $U$  refere-se as localizações das escolas rurais e urbanas, M e F sobre os sexos Masculino e Feminino respectivamente e  $Esc(mãe)$  diz respeito a escolaridade da mãe. O Aprox. 0 refere-se à algum número muito pequeno considerado por este estudo aproximadamente zero.



Tabela 4 – Comparações dois a dois entre as médias sobre a soma das notas com base na raça/cor dos alunos.

Comparações	P-valor	Evidência (RA 95%)
Amarela = Não quero declarar	0.4113	Iguais
Amarela = Branca	0.0005	Desiguais
Amarela = Indígena	1.0000	Iguais
Amarela = Parda	1.0000	Iguais
Amarela = Preta	1.0000	Iguais
Branca = Não quero declarar	Aprox. 0	Desiguais
Branca = Indígena	0.0010	Desiguais
Branca = Parda	Aprox. 0	Desiguais
Branca = Preta	Aprox. 0	Desiguais
Indígena = Não quero declarar	1.0000	Iguais
Indígena = Parda	0.7758	Iguais
Indígena = Preta	1.0000	Iguais
Parda = Não quero declarar	Aprox. 0	Desiguais
Parda = Preta	Aprox. 0	Desiguais

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples.

Anotações: Aprox. 0 refere-se à algum número muito pequeno considerando aproximadamente zero.



Tabela 5 – Comparações entre as médias de notas em Matemática e as regiões das escolas dos alunos com base na amostra de tamanho 500.

Comparações	P-valor	Evidência (RA 95%)
Não sabe = Nunca estudou	1.0000	Iguais
Não sabe = Incompleto 5.º ano do EF	0.0078	Desiguais
Não sabe = Completou 5.º ano do EF	0.0005	Desiguais
Não sabe = Completou 9.º ano do EF	0.0001	Desiguais
Não sabe = Completou EM	Aprox. 0	Desiguais
Não sabe = Completou Faculdade	0.0011	Desiguais
Nunca estudou = Incompleto 5.º ano do EF	1.0000	Iguais
Nunca estudou = Completou 5.º ano do EF	0.5598	Iguais
Nunca estudou = Completou 9.º ano do EF	0.4165	Iguais
Nunca estudou = Completou EM	0.1114	Iguais
Nunca estudou = Completou Faculdade	0.4707	Iguais
Incompleto 5.º ano do EF = Completou 5.º ano do EF	1.0000	Iguais
Incompleto 5.º ano do EF = Completou 9.º ano do EF	1.0000	Iguais
Incompleto 5.º ano do EF = Completou EM	1.0000	Iguais
Incompleto 5.º ano do EF = Completou Faculdade	1.0000	Iguais
Completo 5.º ano do EF = Completou 9.º ano do EF	1.0000	Iguais
Completo 5.º ano do EF = Completou EM	1.0000	Iguais
Completo 5.º ano do EF = Completou Faculdade	1.0000	Iguais
Completo 9.º ano do EF = Completou EM	1.0000	Iguais
Completo 9.º ano do EF = Completou Faculdade	1.0000	Iguais
Completou EM = Completou Faculdade	1.0000	Iguais

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples.

Anotações: Aprox. 0 refere-se à algum número muito pequeno considerando aproximadamente zero.





## 3 Conclusão

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Nisl tincidunt eget nullam non nisi est sit amet. Turpis nunc eget lorem dolor. Posuere urna nec tincidunt praesent semper feugiat nibh. Platea dictumst quisque sagittis purus sit amet volutpat consequat mauris. Velit scelerisque in dictum non consectetur a. Lectus proin nibh nisl condimentum. Nulla pharetra diam sit amet.

Risus in hendrerit gravida rutrum quisque non tellus. Purus gravida quis blandit turpis cursus in hac habitasse. Et malesuada fames ac turpis egestas maecenas pharetra convallis posuere. Id donec ultrices tincidunt arcu non sodales neque. Suspendisse faucibus interdum posuere lorem ipsum dolor sit amet consectetur. Amet purus gravida quis blandit. Sit amet nisl purus in mollis nunc sed. Rutrum quisque non tellus orci. Pellentesque eu tincidunt tortor aliquam nulla facilisi cras fermentum. Lacus suspendisse faucibus interdum posuere lorem ipsum dolor sit amet.

A iaculis at erat pellentesque adipiscing commodo elit. Tellus rutrum tellus pellentesque eu tincidunt tortor aliquam nulla. Eros donec ac odio tempor orci dapibus ultrices. Vitae ultricies leo integer malesuada. Aliquam eleifend mi in nulla posuere sollicitudin aliquam. Vitae purus faucibus ornare suspendisse sed nisi lacus sed.

Consequat interdum varius sit amet mattis. Elementum pulvinar etiam non quam lacus suspendisse faucibus. Morbi non arcu risus quis varius quam quisque id diam. Est ante in nibh mauris cursus mattis molestie a iaculis. Odio aenean sed adipiscing diam donec adipiscing tristique risus nec. Mi tempus imperdiet nulla malesuada pellentesque. Pellentesque habitant morbi tristique senectus et netus et. Luctus venenatis lectus magna fringilla urna porttitor rhoncus. Integer enim neque volutpat ac tincidunt vitae. Viverra ipsum nunc aliquet bibendum.

# Anexos



## ANEXO A – Amostra

Os bancos de dados dos alunos participantes do SAEB de 2017 foram disponibilizados como amostras de tamanho 2.000 em formato CSV. Cada banco de dados era uma amostra aleatória diferente do banco de dados original do SAEB, para cada aluno da Turma A de Métodos Estatísticos 2 do semestre 1/2020 da Universidade de Brasília (UnB).

Os dados utilizados desse documento foram a junção dos três bancos de dados dos integrantes, com os seguintes nomes: **amostra\_190015853.csv**, **amostra\_190029498.csv** e **amostra\_190127180.csv**