



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 total por nível de escolaridade da mãe com base no tempo de afazeres dos alunos.	14
Figura 2 – Proporção por sexo de tempos em afazeres domésticos por parte dos alunos.	15
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 – Testes de igualdade na variabilidade sobre as relações com o tempo de afazeres domésticos por parte dos alunos.	15
Tabela 2 – Comparações dois a dois entre as ordens das posições sobre os tempos de afazeres domésticos com base na escolaridade das mães 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

σ Letra grega minúscula sigma

μ Letra grega minúscula mu

Sumário

	Introdução	8
I	FUNDAMENTAÇÃO	9
1	METODOLOGIA	10
II	RESULTADOS	12
2	COMPARAÇÕES	13
2.1	Tempo em afazeres domésticos	13
2.2	Notas	17
3	CONCLUSÃO	24
	REFERÊNCIAS	25
	ANEXOS	26
	ANEXO A – AMOSTRA	27

blablabalbalbalbalbalblablablbbbbbblbalbalbalblalbalbalblalbalblaba blablabalbal-
balbalbalblalblabbbbbbblbalbalbalblalbalbalblalbalblaba blablabalbalbalbalbalblalblabbbb-
blbalbalbalblalbalbalblalbalblabal albalblalblalblal

Parte I

FUNDAMENTAÇÃO



1 Metodologia

Este documento tem como objetivo analisar fatores sociais de alunos do 9º ano de 2017 do Brasil, no qual há pressupostos de possibilidade de influenciar no aprendizado básico destes. O banco de dados disponibilizado pelo SAEB de 2017 divulgado pelo [INEP \(2018\)](#), possui variáveis com respeito aos alunos e suas notas em Matemática e Língua Portuguesa, através da Prova Brasil de 2017, do censo de escolas públicas e amostras de escolas particulares, entre outras variáveis com respeito a escola e fatores sociais destes.

A análise parte de uma amostragem aleatória simples de 5271 alunos do 9º deste banco de dados, e relacionar fatores como raça/cor e sexo dos alunos e localizações das escolas e escolaridade da mãe destes, com as variáveis a serem explicadas como a soma destas notas e o tempo de afazeres domésticos realizados por dia com base nos alunos.

O estudo relaciona sobre os alunos a raça/cor, a escolaridade da mãe e o sexo com o tempo de afazeres domésticos, 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 [Kruskal e Wallis \(1952\)](#) para fatores com mais de duas categorias (Teste K), e o teste de [Mann e Whitney \(1947\)](#) para a comparação dois a dois das categorias (Teste W). Estes testes avaliam as distribuições das informações com base na posição, para comparar as distâncias significativas entre as categorias.

O estudo também relaciona a raça/cor, a escolaridade da mãe e as localizações das escolas com a soma das notas em Matemática e Língua Portuguesa com base nos alunos, no intuito de avaliar a influência destas variáveis explicativas sobre o desempenho total 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 [Anderson e Darling \(1954\)](#), [Shapiro e Wilk \(1965\)](#) e [Shapiro e Francia \(1972\)](#) foi possível afirmar que as distribuições das notas são normais e os resultados destes não serão abordados neste documento. Os testes apropriados para avaliar estas relações, são testes paramétricos, no qual aplica o teste ANOVA dado por [Fisher \(1928\)](#) para avaliar os fatores com mais de 2 categorias e o teste T para as comparações dois a dois ([O'CONNOR; ROBERTSON; EDMUND, 1908](#)). Estes testes utilizam das médias (μ) de cada grupo para avaliar as distâncias significativas entre eles, no qual previamente avalia se as variâncias (σ^2) destes são iguais (Teste B), através do teste proposto por [Bartlett \(1954\)](#).



Para avaliar os resultados dos testes, foi proposto o uso da correção de [Bonferroni \(1936\)](#), no qual utiliza-se para os testes com mais de duas categorias. Se houver a evidência de rejeitar igualdade destas, a comparação dois a dois é efetuada e a mesma correção é utilizada. Esta correção é sobre o P-valor, cujo valor é avaliado em uma escala de significância, adotado por este estudo como uma confiança de 95%, no que diz sobre aceitar a hipótese H_0 .

Parte II

Resultados



2 Comparações

As relações colocadas por este documento, são avaliadas como fatores sociais que podem influenciar no meio escolar sobre a aquisição do conhecimento básico do 9º ano. Dentre as variáveis escolhida por este estudo, as análises sociais são fundamentadas com relação ao aluno, o sexo, a raça/cor, a escolaridade da mãe e a localização da escola. Para avaliar estes fatores, foi explorado o tempo gasto em afazeres domésticos diariamente no intuito de medir a disponibilidade do aluno em casa, e a soma das notas, Matemática e Língua Portuguesa da Prova Brasil, para avaliar se o desempenho médio é afetado por alguma destas relações.

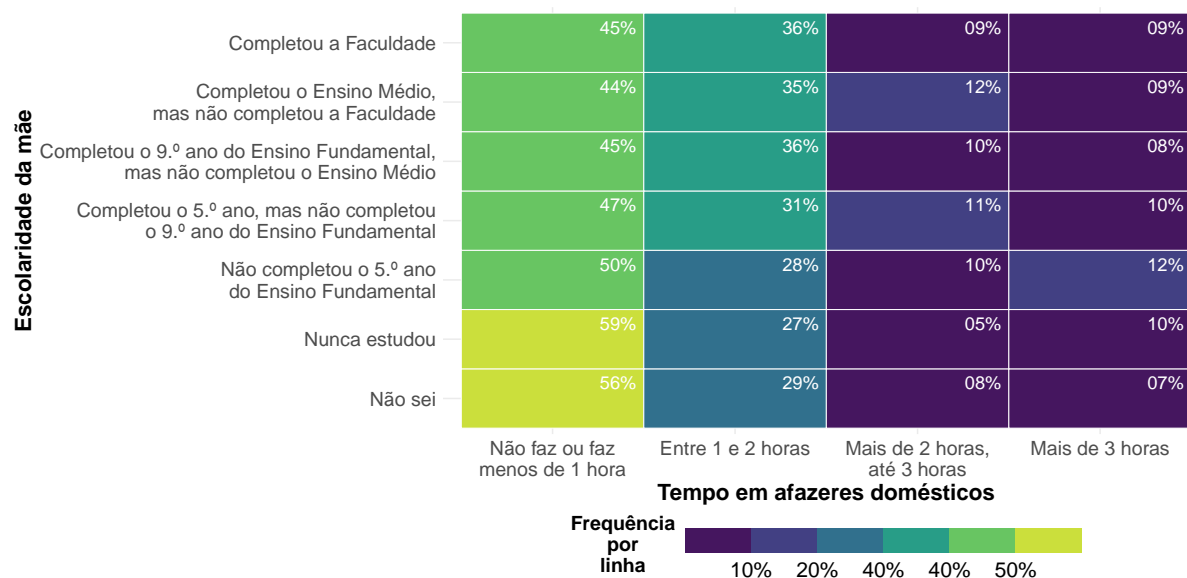
2.1 Tempo em afazeres domésticos

O tempo em afazeres domésticos, priva a disponibilidade dos alunos à tarefas do dia a dia, que por sua vez, grande parte não exerce o processo de conhecimento da educação básica.

As relações sociais propostas para avaliar este interesse, dentre elas, destaca-se a proporção de alunos do sexo Feminino em exercer mais tempo nestes afazeres, pelo menos 60% destes o representam, em cada tempo, de no mínimo uma hora, veja na [Figura 2](#). Ao deparar com a escolaridade da mãe, a proporção de alunos que exercem esses afazeres, para cada nível escolar completo dela, há 70% localizado entre os que não faz ou faz até 2 horas, veja na [Figura 1](#).



Figura 1 – Proporção total por nível de escolaridade da mãe com base no tempo de afazeres dos alunos.

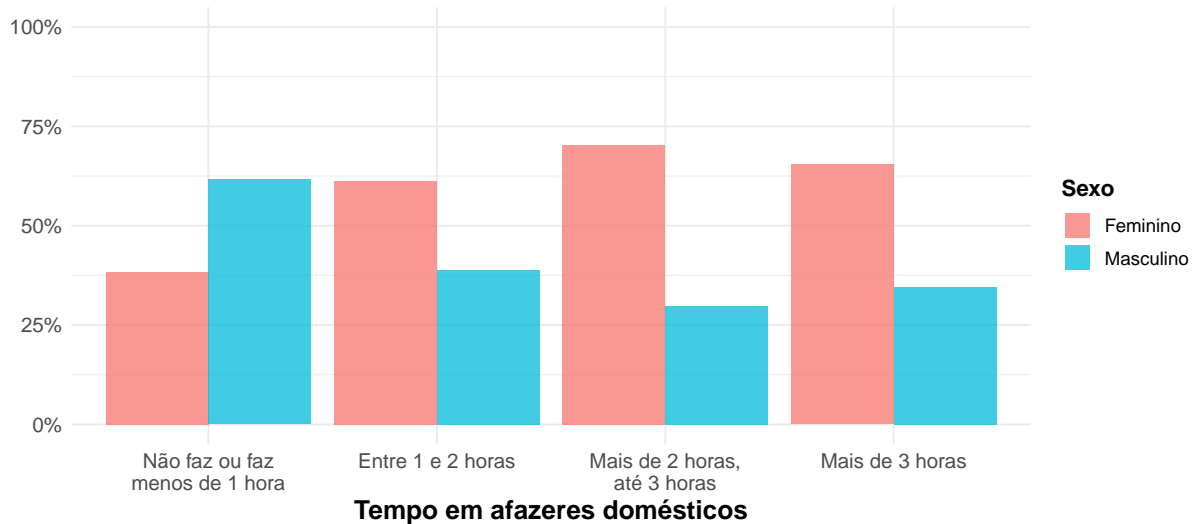


Fonte: Amostra de 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.



Figura 2 – Proporção por sexo de tempos em afazeres domésticos por parte dos alunos.



Fonte: Amostra de 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.

Tabela 1 – 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 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retiradas de uma amostragem aleatórias simples.

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



Tabela 2 – 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 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retiradas de uma amostragem aleatórias simples.

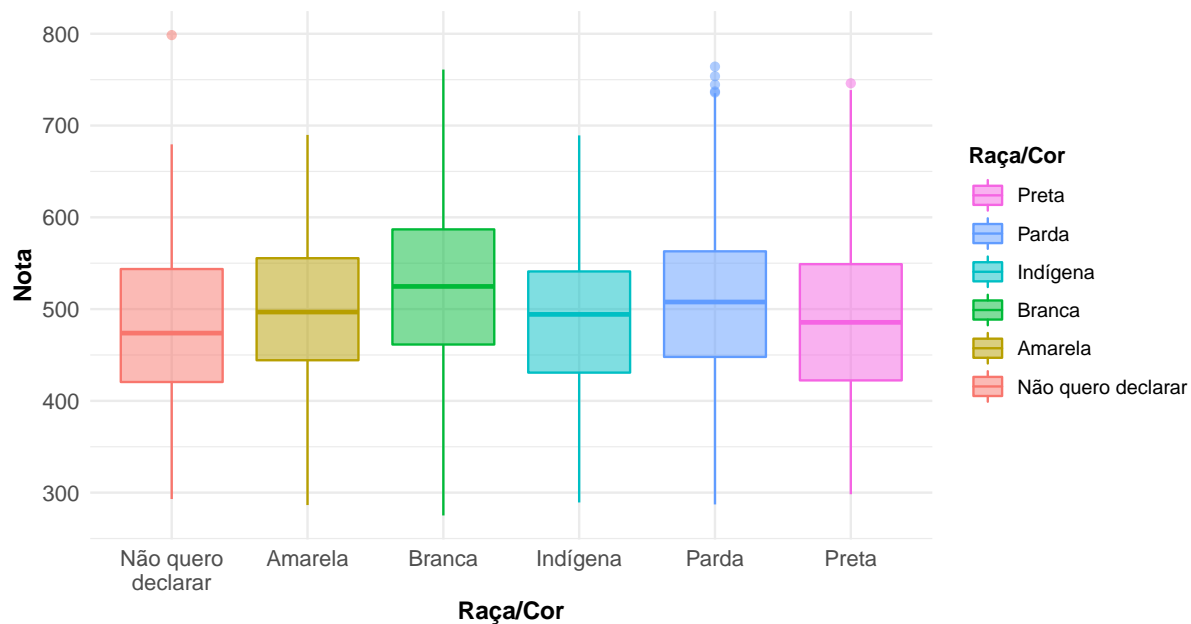
Anotações: Aprox. 0 refere-se à algum número muito pequeno considerando aproximadamente zero.



2.2 Notas

blablalbalbalbalbalbalblablalbbbbbbblbalbalbalblalbalbalblalbalblaba blablalbalbal-
balbalbalblablalbbbbbbblbalbalbalblalbalbalblalbalblaba blablalbalbalbalbalbalblablalbbbb-
blbalbalbalblalbalbalblalbalblabal albalblablalbla

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

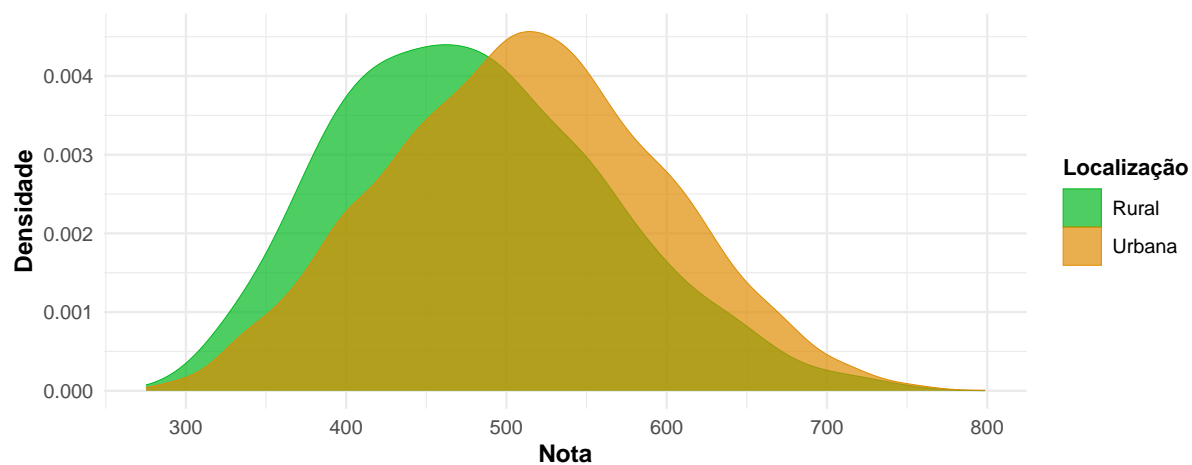


Fonte: Amostra de 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.



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

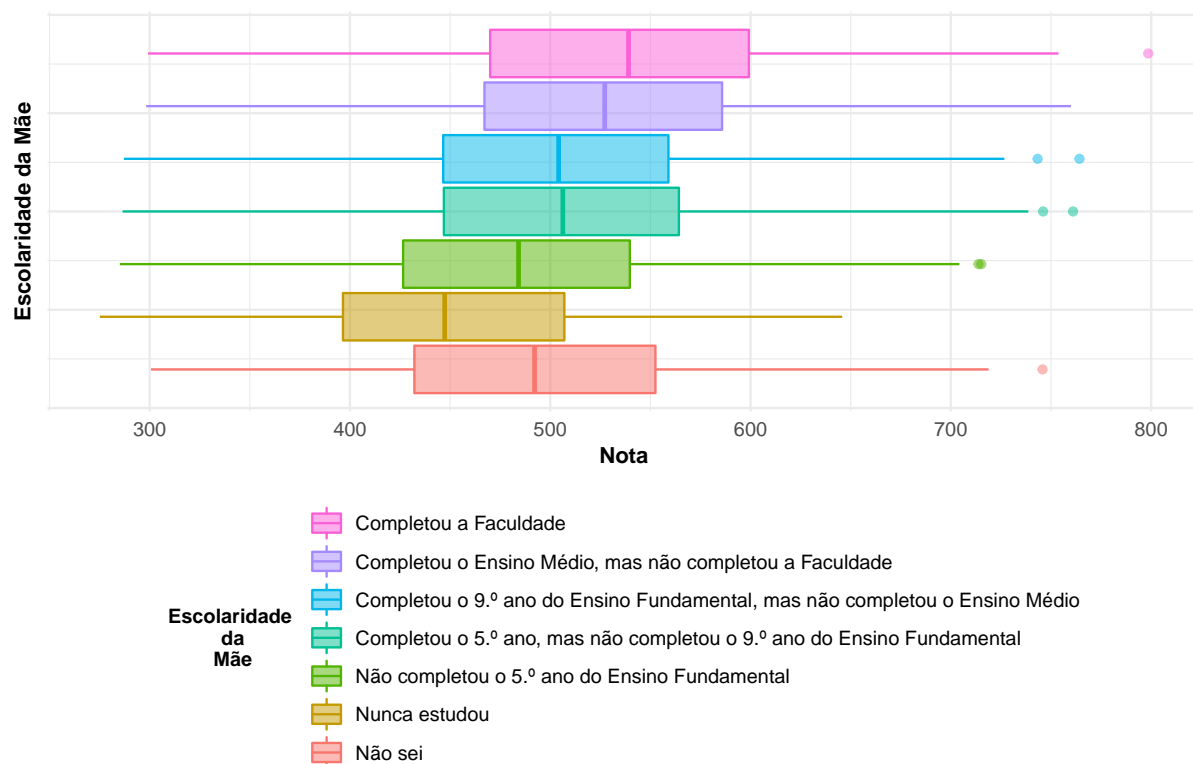


Fonte: Amostra de 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.



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



Fonte: Amostra de 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.



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 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retiradas de uma amostragem aleatórias 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 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retiradas de uma amostragem aleatórias 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 5271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retiradas de uma amostragem aleatórias simples.

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

3 Conclusão

blablabalbalbalbalbalblablablbbbbbblbalbalbalblalbalbalblalbalblaba blablabalbal-
balbalbalblablablbbbbbblbalbalbalblalbalbalblalbalblaba blablabalbalbalbalbalblablablbbbbb-
blbalbalbalblalbalbalblalbalblabal albalblablablalba



Referências

ANDERSON, T. W.; DARLING, D. A. A test of goodness of fit. *Journal of the American statistical association*, Taylor & Francis, v. 49, n. 268, p. 765–769, 1954. Citado na página 10.

BARTLETT, M. S. A note on the multiplying factors for various χ^2 approximations. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 296–298, 1954. Citado na página 10.

BONFERRONI, C. Teoria statistica delle classi e calcolo delle probabilit . *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, v. 8, p. 3–62, 1936. Citado na p gina 11.

FISHER, R. A. The general sampling distribution of the multiple correlation coefficient. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, The Royal Society London, v. 121, n. 788, p. 654–673, 1928. Citado na p gina 10.

INEP. *SAEB 2017 MICRODADOS DA ANEB E DA ANRESC (PROVA BRASIL)*: Saeb (aneb/prova brasil 2017). [S.l.], 2018. 33 p. Dispon vel em: <<http://portal.inep.gov.br/web/guest/microdados>>. Citado na p gina 10.

KRUSKAL, W. H.; WALLIS, W. A. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, Taylor & Francis Group, v. 47, n. 260, p. 583–621, 1952. Citado na p gina 10.

MANN, H. B.; WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, JSTOR, p. 50–60, 1947. Citado na p gina 10.

O’CONNOR, J. J.; ROBERTSON, E. F.; EDMUND, F. Student’s t-test. *MacTutor History of Mathematics archive*, 1908. Citado na p gina 10.

SHAPIRO, S. S.; FRANCIA, R. An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 67, n. 337, p. 215–216, 1972. Citado na p gina 10.

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). *Biometrika*, JSTOR, v. 52, n. 3/4, p. 591–611, 1965. Citado na p gina 10.

Anexos



ANEXO A – Amostra

O Banco de dados dos alunos participantes do Saeb de 2017 foi disponibilizado como amostras de 2000 alunos em formato Excel referentes à cada matrícula para os cursandos de Métodos Estatísticos 2 da Universidade de Brasília (Unb), no qual este documento analisa o arquivo: **amostra_190127180.csv**.