

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
Bacharelado em Ciências da Computação

Axell Brendow Batista Moreira
Cecilia Capurucho Bouchardet
Maria Aparecida Brenda de Souza
Maycon Bruno de Jesus
Victor Colen Costa

**IDENTIFICAR COMO A NOTA NAS PROVAS OBJETIVAS DO ENEM FOI
IMPACTADA PELA DESIGUALDADE SOCIOECONÔMICA**

1 INTRODUÇÃO

O ENEM(Exame Nacional do Ensino Médio) é uma prova de âmbito nacional que testa os conhecimentos adquiridos desde o ensino básico até o fundamental de estudantes de todo o país, com o intuito de disponibilizar vagas em universidades públicas e privadas do Brasil.

O Brasil é o 84º país no ranking de IDH mundial. Esse ranking considera os aspectos econômicos do desenvolvimento de um país. Um dos fatores que influenciam essa posição baixa do Brasil nessa ranking é a desigualdade social. Dados do último Estudo IBGE estudo feito pelo Instituto Brasileiro de Geografia e Estatística (IBGE), mostram que o rendimento mensal dos 1% mais ricos do país é quase 34 vezes maior do que o rendimento da metade mais pobre da população. Ainda, o estudo mostrou que a renda dos 5% mais pobres caiu em 3%, enquanto a renda dos 1% mais ricos aumentou em 8%.

Essa desigualdade se reflete também no âmbito escolar brasileiro. Segundo a Revista Educação Anuário Brasileiro da Educação Básica 2020, a aprendizagem adequada em matemática no 9º do ensino fundamental de alunos ricos foi de 54,7%, já para os estudantes pobres, apenas 8,8% chegaram no nível tido como condizente. A desigualdade é ainda maior quando se analisa a qualidade adequada da mesma disciplina nos alunos do 3º ano do ensino médio: enquanto os ricos adquiriram 45,7% do nível esperado, somente 3,2% dos estudantes pobres obtiveram o mesmo patamar de aprendizagem. Sobre laboratórios de ciências, nos anos iniciais do ensino fundamental, somente 9,7% das escolas possuem tal área. Nos anos finais há um pequeno aumento, chegando a 23,7% e somando as duas etapas, 12,5% das escolas com ensino fundamental possuem laboratórios de ciências. Por fim, já no ensino médio, 48% têm infraestrutura laboratorial.

Ao analisar o banco de dados do Enem(Exame Nacional do Ensino Médio), pode-se ver atributos que contenha informações do participante. E com elas retiraria-se dados que comprovaria o problema. As provas objetivas do Enem são divididas em dois dias, onde o primeiro é a prova de linguagens e humanas e a segunda e matemática e natureza. Elas abrangem todo o ensino médio. O problema começa que por ser uma prova grande e que precisa de uma alta capacidade de interpretação de textos, conteúdos e raciocínio lógico. Assim, quem não teve um bom ensino médio, provavelmente não fará uma boa prova. A parte socioeconômica afeta os estudantes sem condições de uma escola particular, já que a maioria das escolas públicas têm o ensino de má qualidade. Além disso, a maioria dos alunos que estuda nas escolas públicas não são brancos. E essas pessoas não tem condições muitas vezes de pagar livros e cursinhos preparatórios.

1.1 Motivação

Analisando os últimos estudos do IBGE e do Anuário Brasileiro da Educação Básica é visível que existe uma desigualdade social muito grande no Brasil e uma grande desigualdade na educação desde o ensino fundamental até o médio, que é a porta de entrada dos alunos para o ENEM. Sendo assim, a probabilidade dessa desigualdade se refletir em notas inferiores para o mais pobre e de raças menos favorecidas socialmente é alto.

1.2 Objetivos

Analisar as provas objetivas do primeiro e segundo dia do ENEM e comprovar o quanto a desigualdade socioeconômica influencia nas notas dessas provas.

1.3 Resultados Esperados

Espera-se encontrar uma desigualdade nas notas das provas objetivas do ENEM, favorecendo as pessoas com maiores rendas e de raças mais privilegiadas socialmente.

2 METODOLOGIA

2.0.1 *Descrição da base e dicionário*

Foi extraída uma base (Disponível em: https://download.inep.gov.br/microdados/microdados_enem_2019.zip) do site do INEP(Instituto Nacional de EStudos e Pesquisas Educacionais Anísio Teixeira) com cerca de 5 milhões de microdados da prova do ENEM do ano de 2019. Esta base contém as seguintes informações sobre um participante da prova.

- Dados de participante
- Dados da escola
- Dados dos pedidos de atendimento especializado
- Dados dos pedidos de específico
- Dados dos pedidos de recursos especializados e específicos para realização das provas
- Dados do local de aplicação da prova
- Dados da prova objetiva
- Dados da redação
- Dados do questionário socioeconômico

Cada tópico acima contém diversos atributos numéricos e nominais, totalizando 105 atributos.

Foi feita uma limpeza da base considerando apenas os atributos relevantes para a investigação socio-econômica relacionada as notas das provas. Sendo assim no final restaram os seguintes atributos que foram analisados.

- TP_COR_RACA: Cor/raça
- TP_ESCOLA: Tipo de escola do Ensino Médio
- NU_NOTA_CN: Nota da prova de Ciências da Natureza
- NU_NOTA_CH: Nota da prova de Ciências Humanas
- NU_NOTA_LC: Nota da prova de Linguagens e Códigos
- NU_NOTA_MT: Nota da prova de Matemática
- Q006: Qual é a renda mensal de sua família? (Some a sua renda com a dos seus familiares.)

Também foi feito um filtro através dos atributos TP_ST_CONCLUSAO (Situação de conclusão do Ensino Médio) e IN_TREINEIRO(Indica-se o inscrito fez a prova com intuito de apenas treinar seus conhecimentos) onde pegamos apenas participantes que já tinham concluído o ensino médio e não eram treineiros.

2.0.2 *Materiais*

1. Site do INEP.
2. Excel.
3. Python.

2.0.3 *Metodos*

Foi utilizado o Excel para entender o dicionário de dados e conseguir separar os atributos que fossem pertinentes à análise. Como a base era muito grande decidimos usar o Python para filtrar e analisar os dados. Então primeiramente foi feito um filtro com apenas os participantes que tinham concluído o ensino médio e não eram treineiros, que gerou uma nova base com 2.900.000 instâncias. A partir dessas instâncias foi feita a análise exploratória dos dados extraindo seus fatos e oportunidades.

Para a comprovação da hipótese inicial, foi calculado a média geral das notas de todas as provas objetivas agrupadas por Cor/Raça (TP_COR_RACA) e por Renda (Q006) e depois a média de cada Cor/Raça e cada Renda. Então foi plotado gráficos com as informações obtidas dessas análises estatísticas.

3 FUNDAMENTOS

3.0.1 Análise dos gráficos

A linha azul dos gráficos representa a média de cada raça. Já a linha laranja, a média global.

Ao analisar os gráficos que estão relacionados a renda que são as Figura 1, Figura 2, Figura 3 e Figura 4 e os resultados das provas percebe-se que quanto maior a renda dos participantes mais altas são suas médias. Sendo que em algumas disciplinas, participantes que tinham uma determinada renda ficaram mais acima da média do que em outras.

O gráfico da Figura 5 referente as notas na prova de matemática, observando-o percebe-se que a porcentagem de participantes que alcançaram a média na mesma e baixa, sendo as raças Branca e não declarada as mais altas, a Amarela um pouco acima da média e as raças Preta e Parda abaixo da média, já a raça Indígena e a menor de todas as notas apresentadas.

A Figura 6 representa o gráfico das notas na prova de Ciências da Natureza, tendo como foco a Cor/Raça. Quando analisado percebe-se que a cor Preta e Parda, assim como no outro gráfico obtém baixa na nota, a Indígena possui a menor nota e as raças Branca e não definida alcançaram os maiores resultados.

Referente a Figura 7 o grupo pertencente as raças Preta, Parda, Amarela e Indígena não atingiram a média na prova de Ciências Humanas, ficando assim abaixo dela. Já as Raças Brancas e que não foram definidas alcançaram resultados acima da média estabelecida.

Ao observar a Figura 8 relacionada as provas de linguagens pode se ver que as notas dos participantes de raça parda e negra ficaram quase completamente iguais ao comparar com a Figura 6 da prova de matemática.

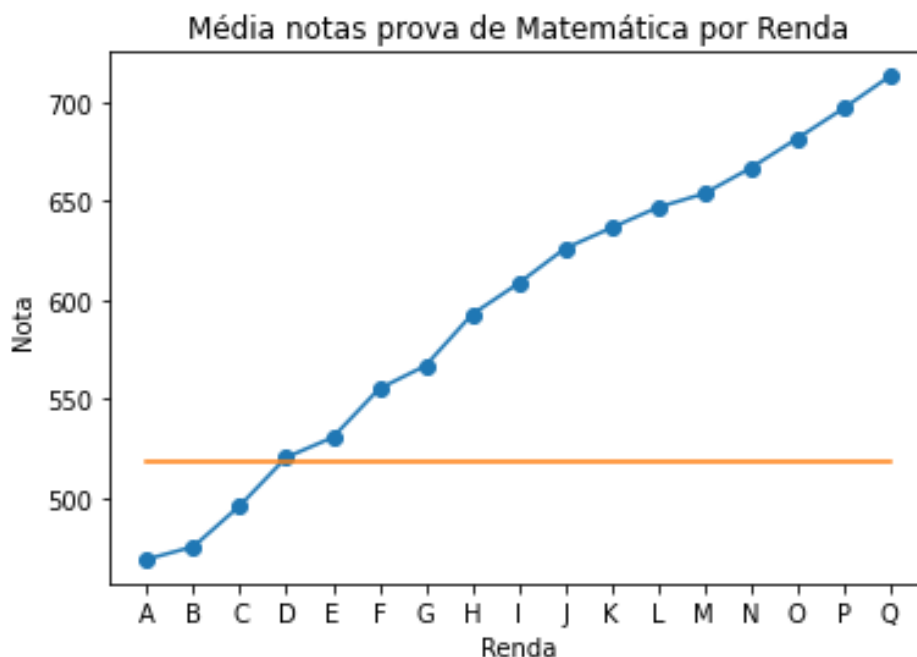


Figura 1 – Nota de Matemática em relação a renda do participante

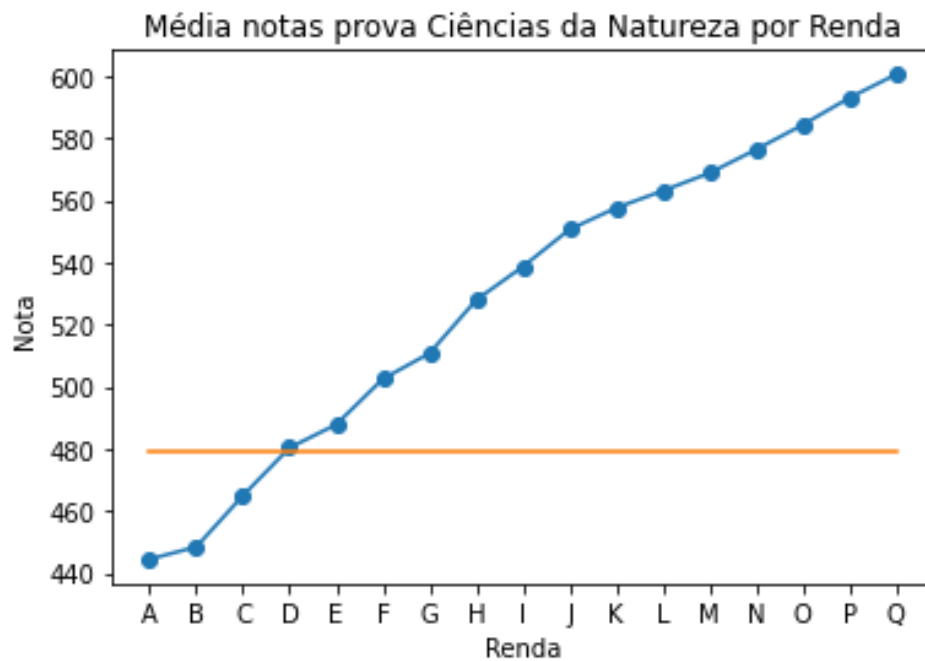


Figura 2 – Nota de Natureza em relação a renda do participante

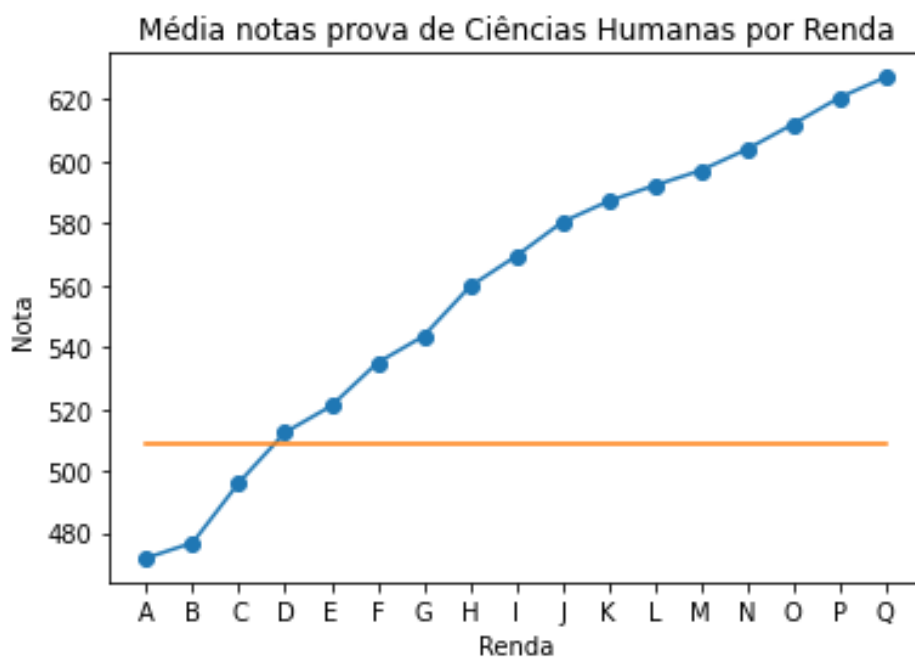


Figura 3 – Nota de Humanas em relação a renda do participante

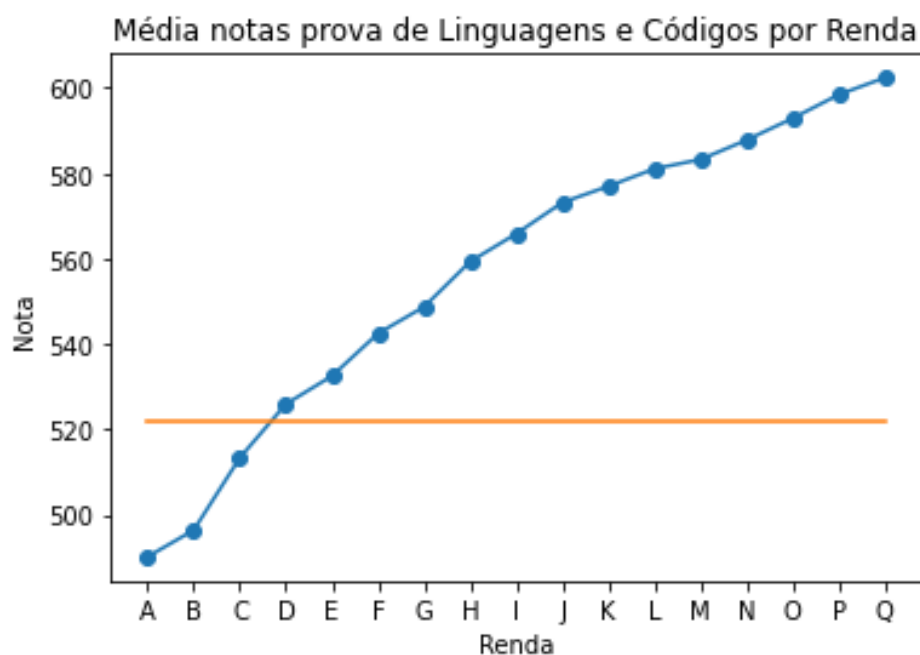


Figura 4 – Nota de Linguagens em relação a renda do participante

Categoria	Descrição
A	Nenhuma renda.
B	Até R\$ 998,00.
C	De R\$ 998,01 até R\$ 1.497,00.
D	De R\$ 1.497,01 até R\$ 1.996,00.
E	De R\$ 1.996,01 até R\$ 2.495,00.
F	De R\$ 2.495,01 até R\$ 2.994,00.
G	De R\$ 2.994,01 até R\$ 3.992,00.
H	De R\$ 3.992,01 até R\$ 4.990,00.
I	De R\$ 4.990,01 até R\$ 5.988,00.
J	De R\$ 5.988,01 até R\$ 6.986,00.
K	De R\$ 6.986,01 até R\$ 7.984,00.
L	De R\$ 7.984,01 até R\$ 8.982,00.
M	De R\$ 8.982,01 até R\$ 9.980,00.
N	De R\$ 9.980,01 até R\$ 11.976,00.
O	De R\$ 11.976,01 até R\$ 14.970,00.
P	De R\$ 14.970,01 até R\$ 19.960,00.
Q	Mais de R\$ 19.960,00.

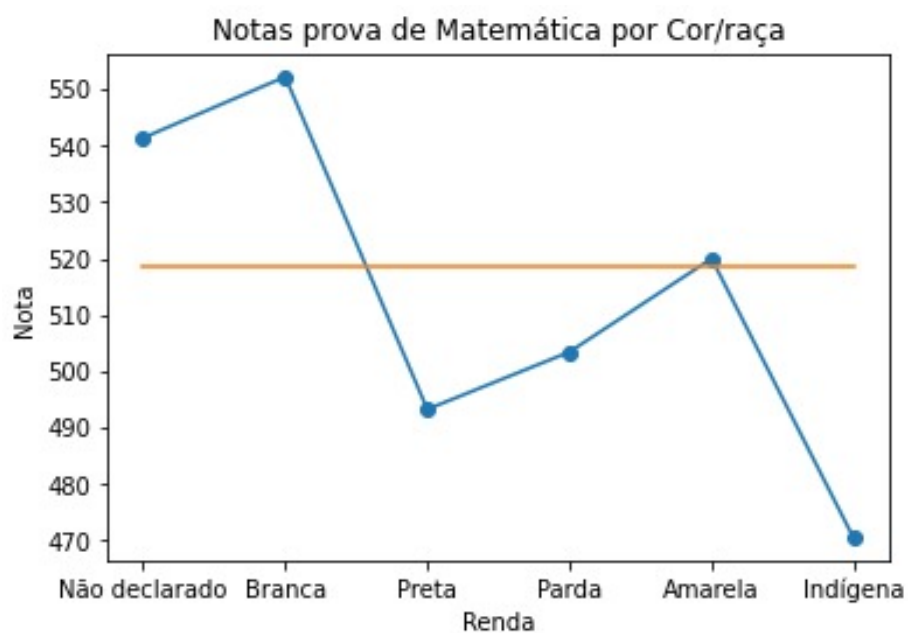


Figura 5 – Nota de Matemática em relação a raça do participante

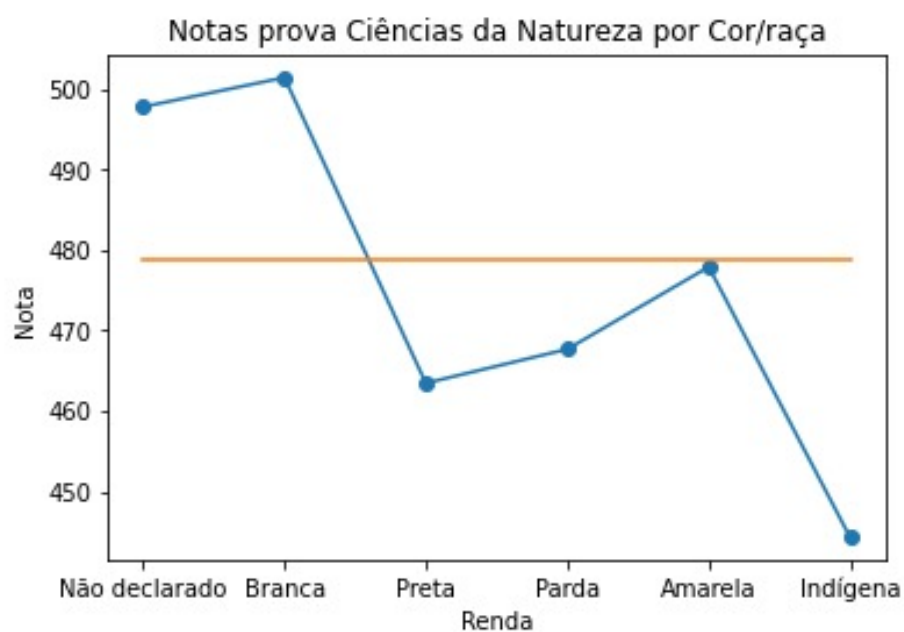


Figura 6 – Nota de Natureza em relação a raça do participante

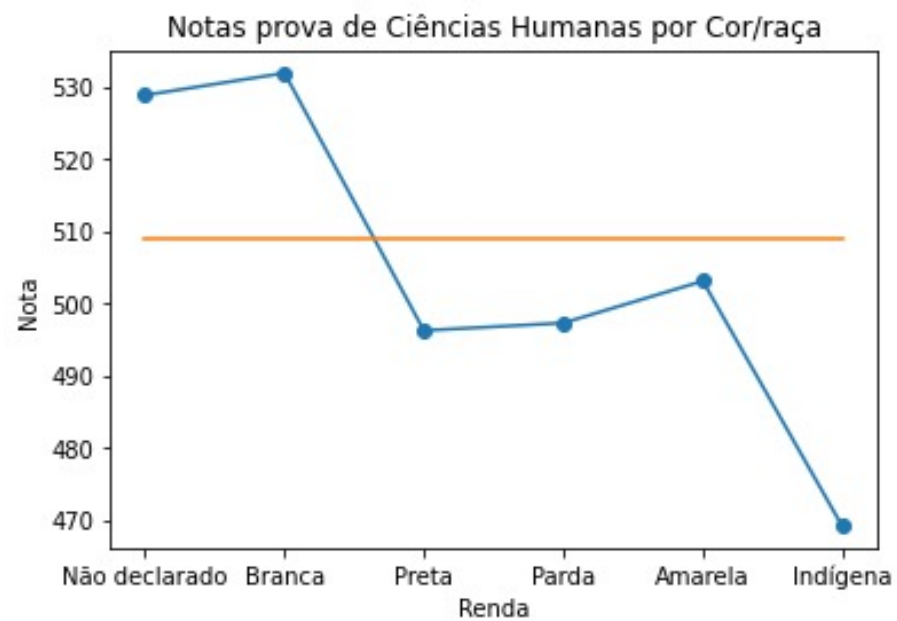


Figura 7 – Nota de Humanas em relação a raça do participante

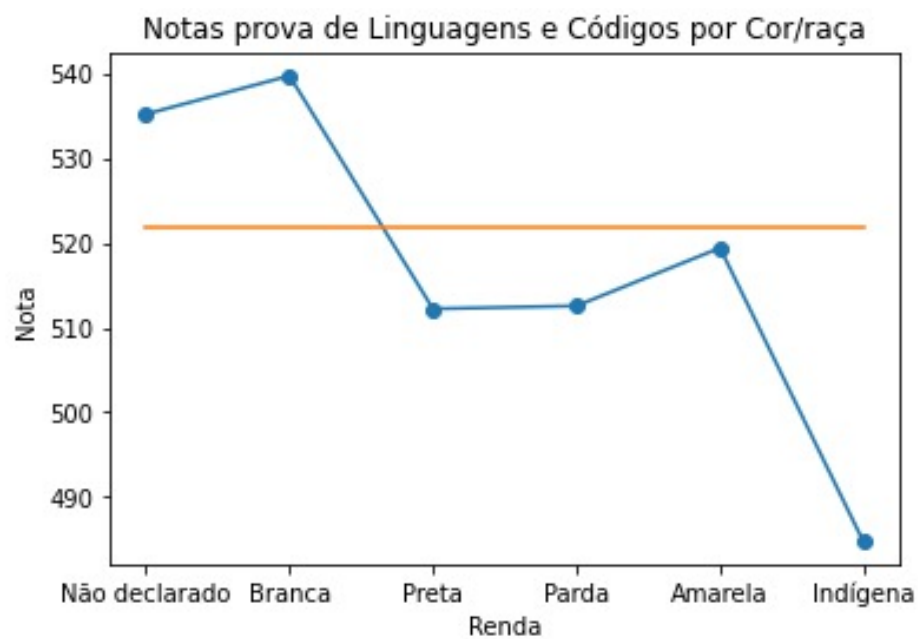


Figura 8 – Nota de Linguagens em relação a raça do participante

4 Cronograma

Cronograma

Data	Entrega
10 de outubro	primeira reunião
11 de outubro	base do código
16 de outubro	produção final do artigo
17 de outubro	entrega final

5 Resultados

Pode-se ver que as notas por rendas são bastante parecidas ao comparar as provas objetivas. Mas ao olharmos a média global, representada pela linha laranja, pode-se ver que a prova de natureza obteve pior resultado em relação às outras notas.

A raça branca é a que tem as melhores notas, comprovando a hipótese de que as outras raças, que estão em maior vulnerabilidade social apresentam melhores notas nas provas objetivas, demonstrando a desigualdade racial existente no Brasil. E a pior nota na média global foi da prova de Natureza.

6 Conclusões

O problema esta exatamente no atributo socioeconômico do participante, já que muitos não obtiveram um acesso ao um ensino de qualidade, por não conseguir um acesso as escolas privadas ou ao um cursinho preparatório, e muitas vezes aos livros. O fato do problema estar associado a classe que o participante da prova pertence não poderia ser um fato que impactaria a sua nota,

Foi proposto que dentro das provas objetivas a de matemática seria a mais afetada pelo seu grau de dificuldade. Mas ao concluir, pode-se perceber que não é a somente a prova de matematica que foi afetada, mas sim, todas as provas objetivas tiveram quase a mesmo impacto diante da condição socioeconômica. E que tanto por raça ou renda a prova de Natureza teve o pior resultado na média global. Também foi visto que a raça indígena e a mais afetada. E que o problema existe e afeta grande parte da população e , assim, e preciso agir diante desse problema.

REFERÊNCIAS

INEP. Bancos de dados fornecido no site do Inep. Out. 2021.
<https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enem>

REVISTA EDUCACAO. Educação brasileira avançou, mas desigualdade de aprendizagem ainda é latente. Agos. 2020.
<https://revistaeducacao.com.br/2020/08/03/educacao-brasileira-desigualdade/>

ESTUDO IBGE. Aumenta desigualdade social no país. Out. 2019.
www.anfip.org.br/geral/aumenta-desigualdade-social-no-pais-revela-pesquisa-do-ibge/