

Relatório da Análise dos Micro Dados do Enem 2021

Sumário

1. Introdução
2. Normalização das notas
3. Performance do participante por perfil social, racial e regional
4. Performance do participante por perfil socioeconômico
5. Considerações finais
6. Conclusão
7. Referências

1. Introdução

Serão analisados como performance dos participantes as notas das provas das quatro áreas do conhecimento, Linguagem e Códigos (LC), Ciências Humanas (CH), Ciências da Natureza (CN) e Matemática (MT).

Não utilizaremos as notas de redação para avaliação da tendência de performance dos participantes, pelo fato que a correção pode conter alguma subjetividade do corretor, ainda que pequena.

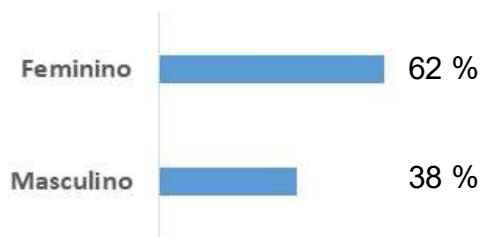
1.1 Condições de contorno da análise dos dados:

- 1) As notas onde o participante faltou à prova ou foi eliminado foram desconsideradas;
- 2) As notas iguais a zero foram desconsideradas da análise uma vez que as provas do Enem têm nota mínima diferente de zero, a nota zero foi considerada uma não conformidade durante o processo de teste;
- 3) Os participantes treineiros foram mantidos na população amostral.

1.2 Participação:

A edição do Enem de 2021 teve um total de inscritos de 3.389.832.

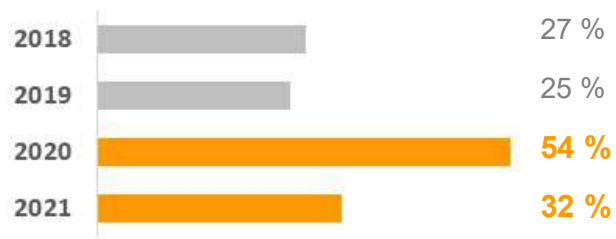
Considerando as condições de contorno:



Comparativo de inscrições com as edições anteriores:



1.3 Ausência média dos dois dias de exame:



Como pode ser observado em seu primeiro ano de pandemia, as abstenções chegaram a mais da metade, em 2021 as abstenções diminuíram em relação à 2020, porém se manteve mais alto que em anos pré-pandemia. Outro dado a ser observado é que em 2021 foi um ano com menor inscrições quando comparado com os três anos anteriores, reflexos da pandemia iniciada em 2020.

2. Normalização das notas

Como as notas de cada prova das quatro áreas do conhecimento têm valores de mínimo e máximos diferentes entre si, a nota de cada prova foi normalizadas pelo método de normalização de MinMaxScaler.

Posteriormente estas notas foram empilhadas, utilizando-se a média das notas para cada classe avaliada, mostrando assim a tendência de performance para os diferentes perfis e características dos participantes.

Método de normalização:

$$X'i = \frac{Xi - Xmin}{Xmax - Xmin}$$

Onde:

Xmin = a menor nota possível da prova

Xmax = a maior nota possível da prova

Xi = a nota do participante

X'i = a nota normalizada do participante

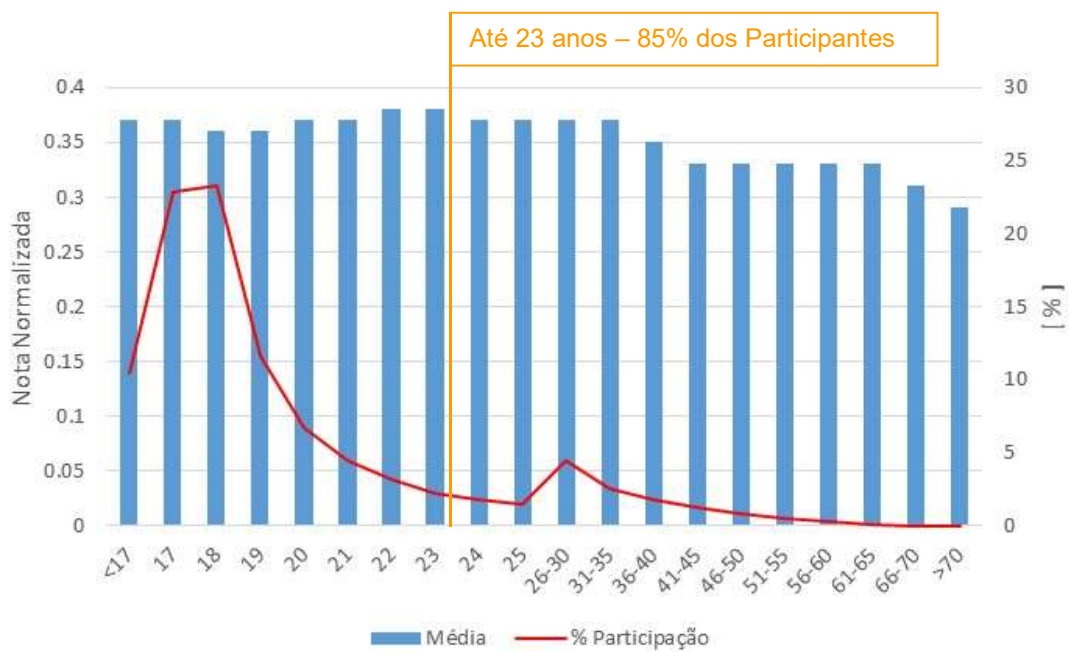
Nota	LC	CH	CN	MT	Normalizada
Min.	295,2	311,6	307,6	310,4	0
Max.	826,1	846,9	867,1	953,1	1

3. Performance do participante por perfil social, racial e regional

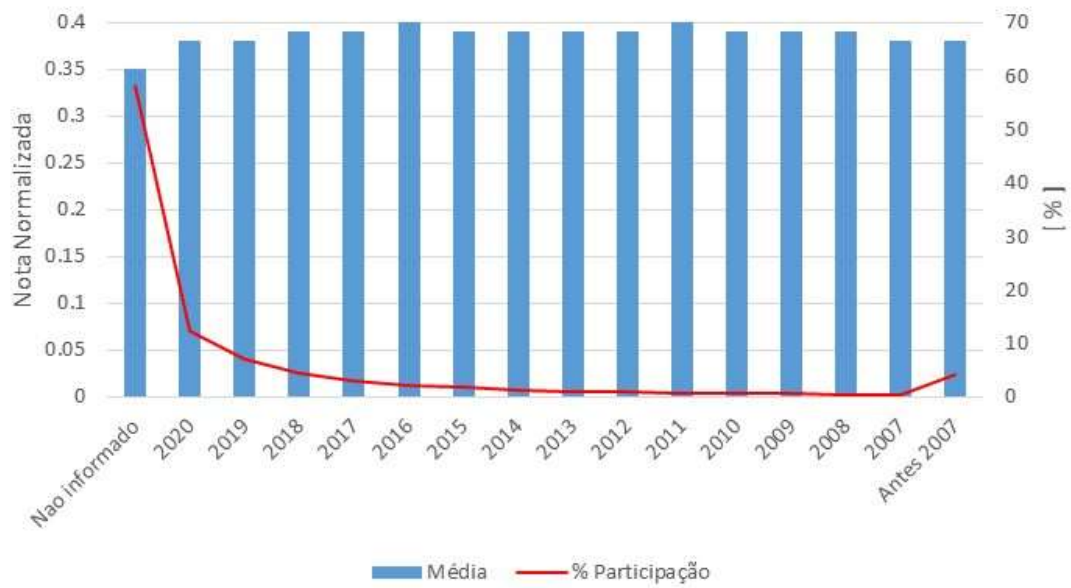
Nota!

Os valores apresentados são a média das notas normalizadas das provas das quatro áreas do conhecimento, exceto redação.

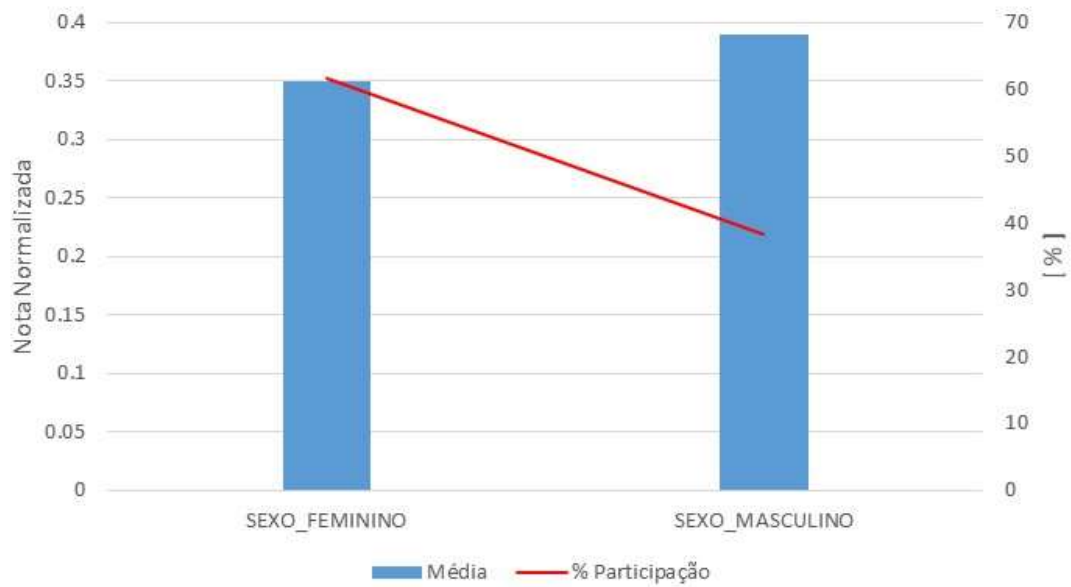
3.1 Por Faixa Etária:



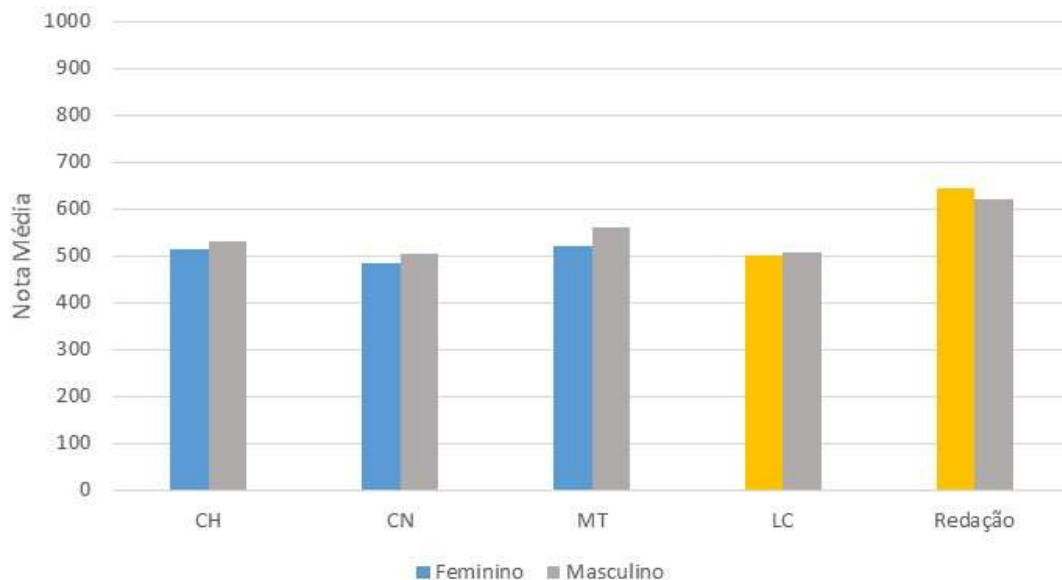
3.2 Por Ano de Conclusão:



3.3 Por Gênero:



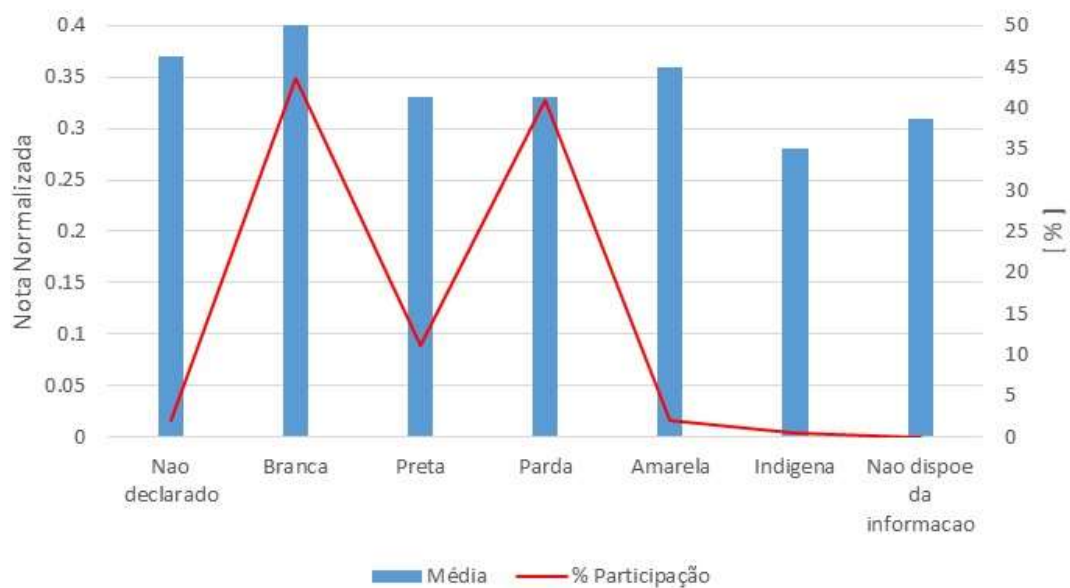
As participantes femininas têm médias ligeiramente menores que os participantes masculinos em três das quatro áreas do conhecimento, já na prova de Linguagem elas praticamente empatam com o público masculino, observando que a única nota máxima foi feminina. Na redação elas tem média superior aos participantes masculinos.



Segundo a 5ª edição da pesquisa Retratos da Leitura no Brasil publicada em 2020 pelo Instituto Pró-Livro, a média das 15 capitais pesquisadas apontam que as mulheres representam 55% dos leitores enquanto que os homens são 45% do público leitor.

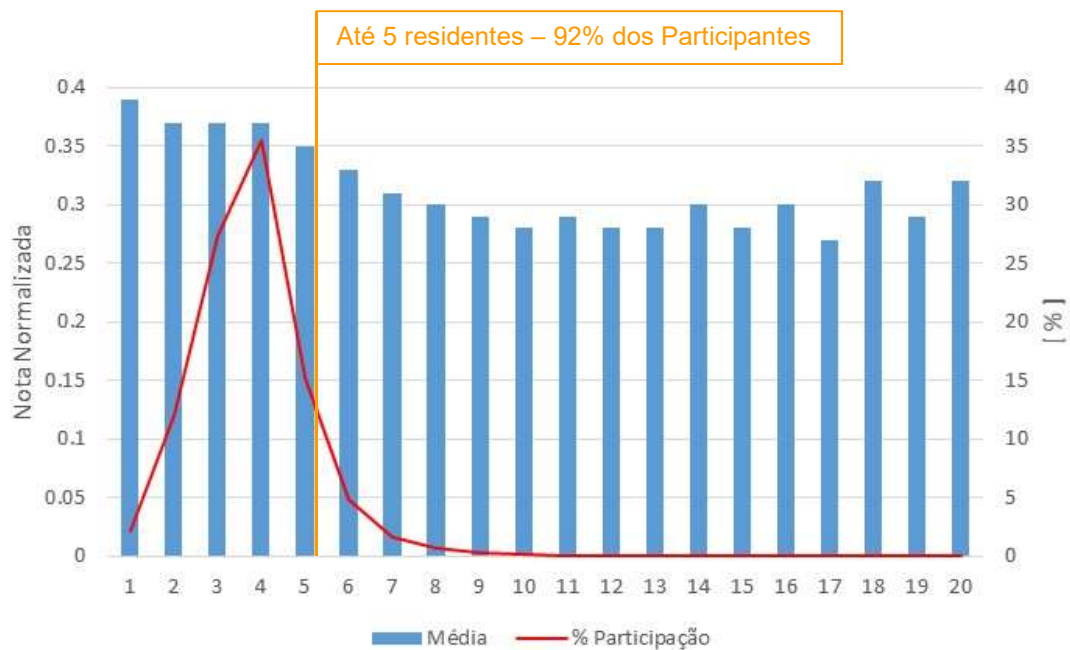
Talvez isto explique a melhor performance das mulheres em redação. Elas provavelmente leem mais que os homens desde a escola média e levam este hábito para a vida adulta.

3.4 Por Raça:



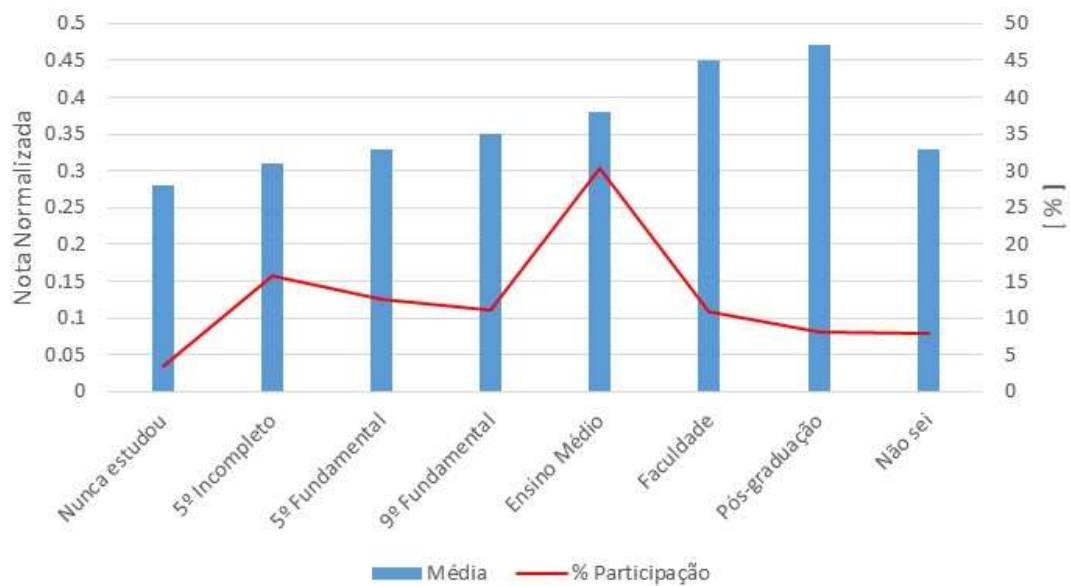
A raça branca 43,6% e a parda 41%, representam juntas aproximadamente 85% dos participantes.

3.5 Por Quantidade de Residentes:

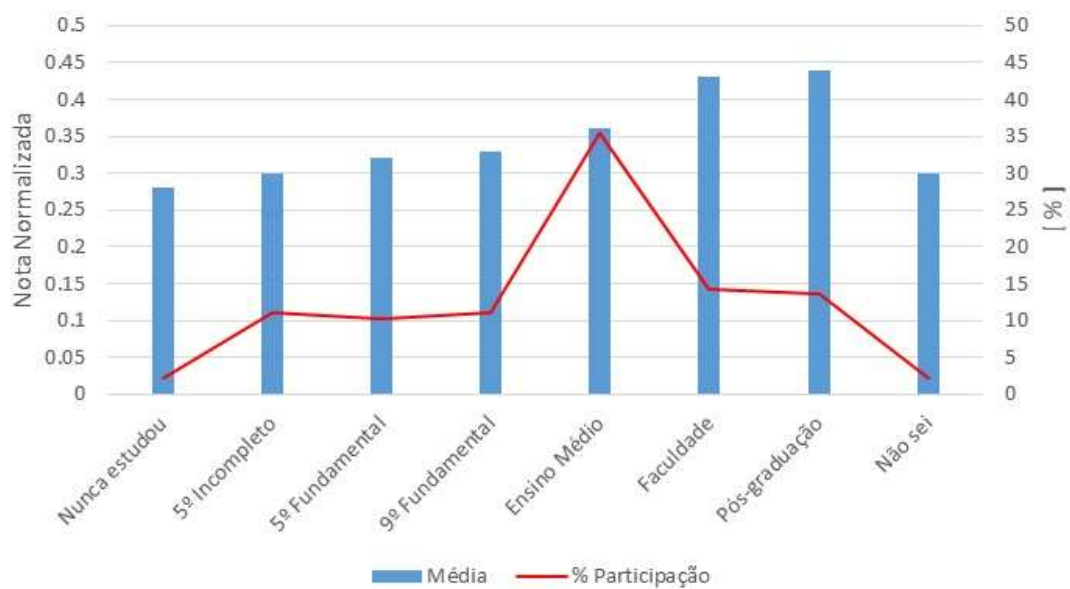


3.6 Por Nível de Instrução do Responsável:

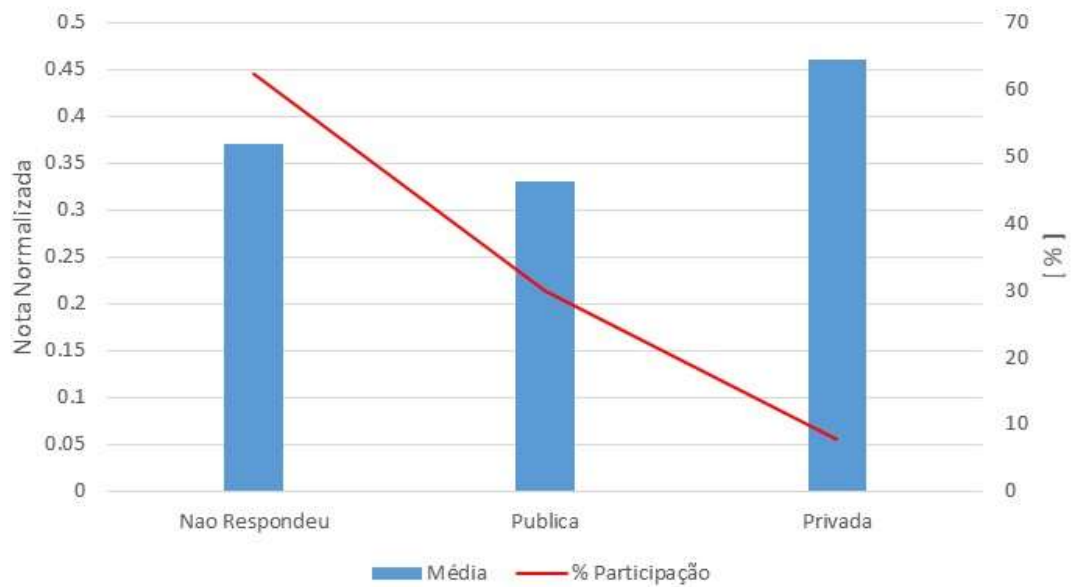
Responsável homem:



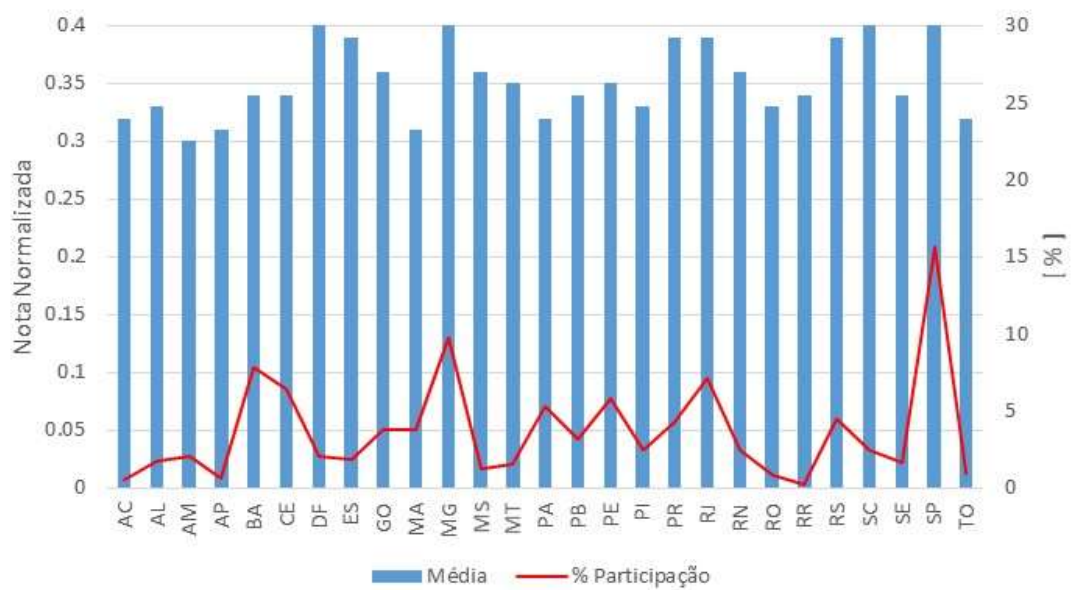
Responsável Mulher:

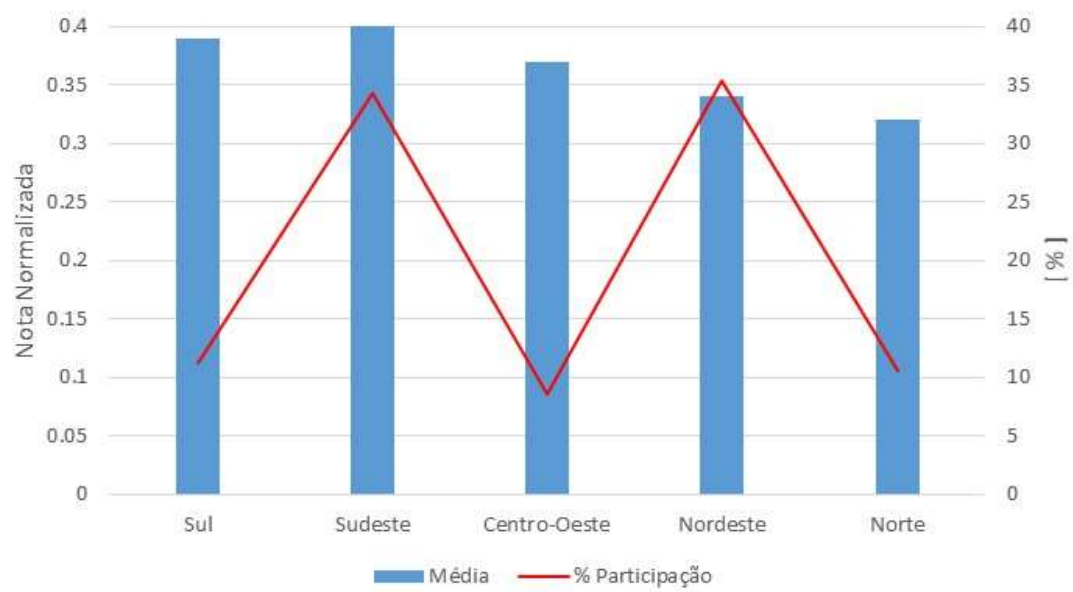


3.7 Por Instituição de Ensino:



3.8 Por Região de Realização da Prova:



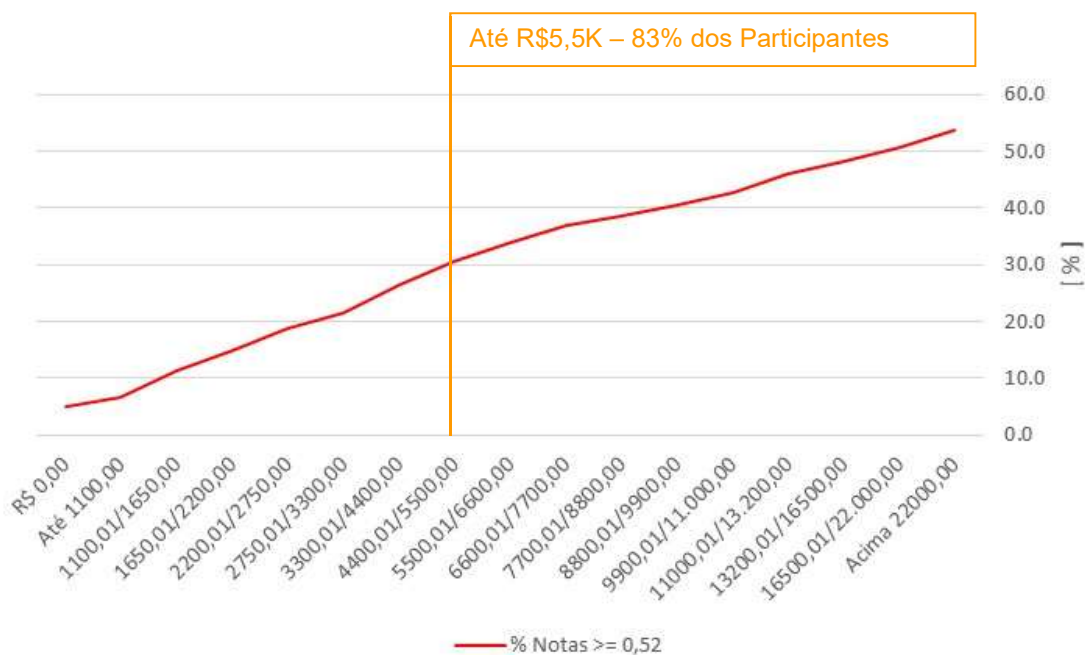
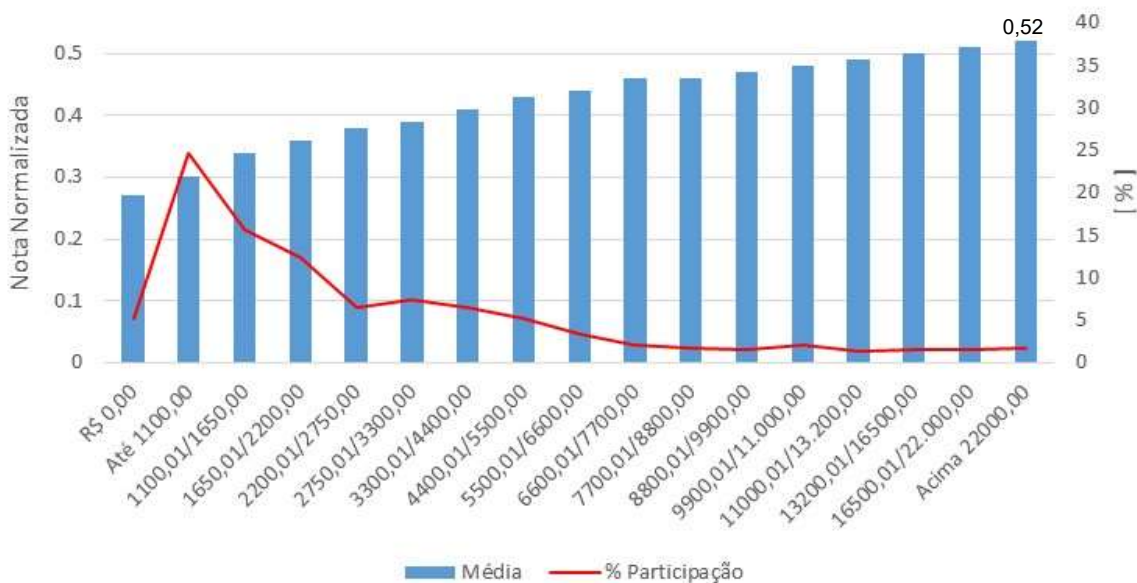


4. Performance do participante por perfil socioeconômico

Nota!

Os valores apresentados são a média das notas normalizadas das provas das quatro áreas do conhecimento, exceto redação.

4.1 Por Faixa de Renda:



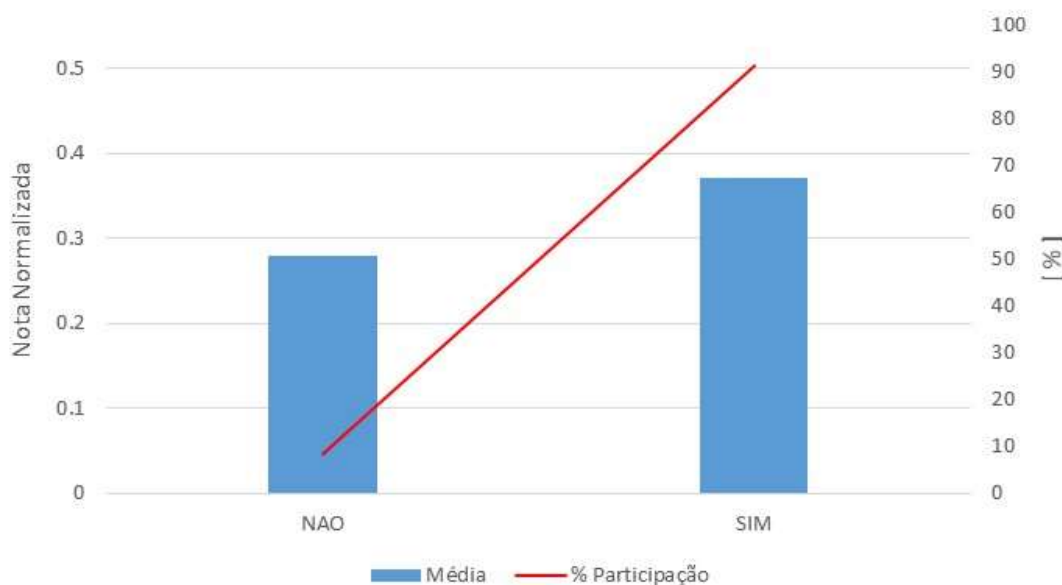
A estratificação por faixa de renda se mostrou mais sensível à média das notas, apresentou um valor mínimo de média de 0,27 e um valor máximo de 0,52, sendo diretamente proporcional ao valor da faixa de renda.

Nenhuma outra análise apresentou médias acima de 0,5 sendo que a maioria das análises apresentaram médias abaixo de 0,4.

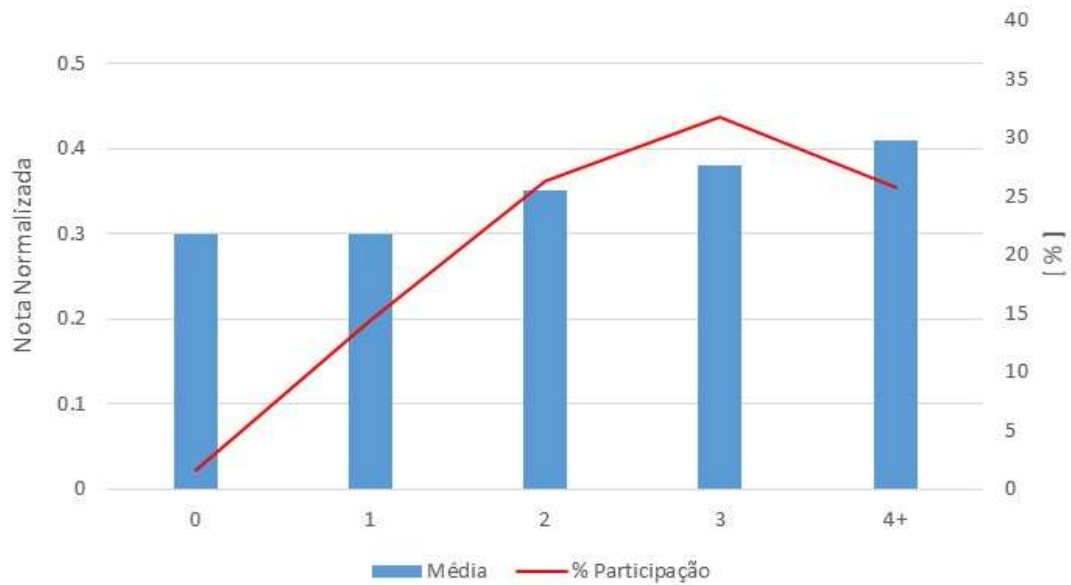
Se fizermos um corte a partir da maior média que foi 0,52 veremos que a população de participantes de R\$0,00 a R\$5,5K apenas 14% de todas as notas desta população são iguais ou maiores que 0,52.

Enquanto que na população de maior renda 54% das notas são iguais ou maiores que 0,52.

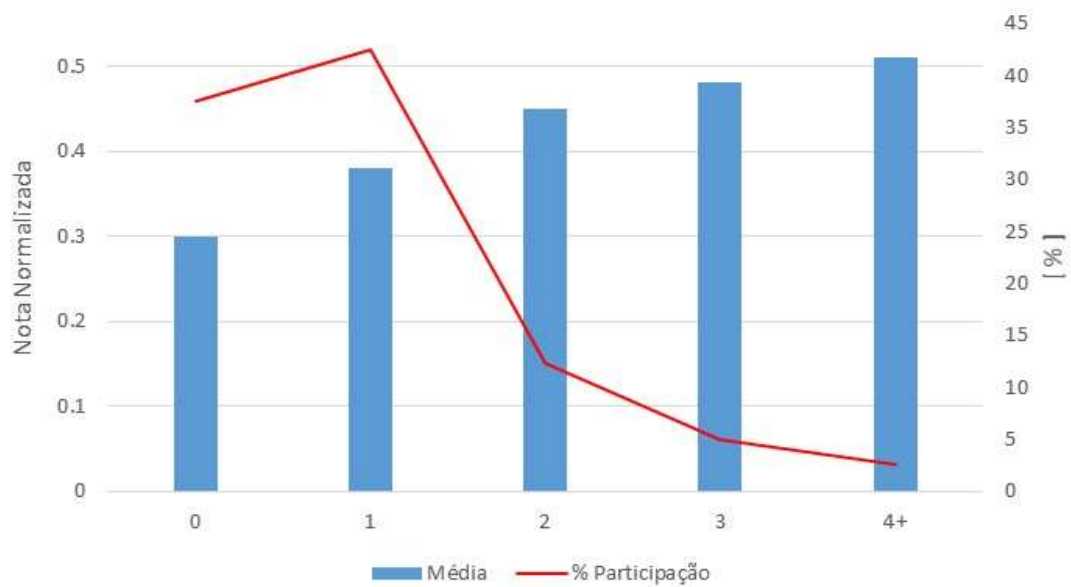
4.2 Por Acesso à Internet:



4.3 Por Quantidade de Celular na Residência:



4.4 Por Quantidade de Computador na Residência:



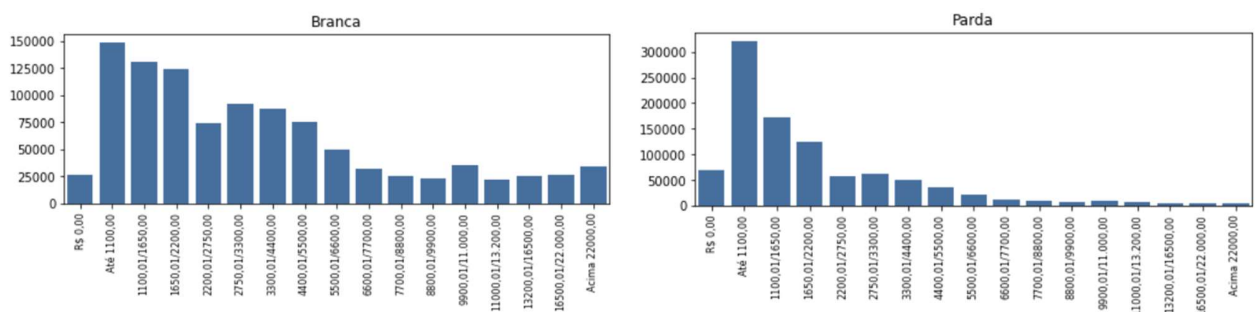
5. Considerações Finais

_ A média normalizada global com base nas provas das quatro áreas do conhecimento, Linguagem e Códigos (LC), Ciências Humanas (CH), Ciências da Natureza (CN) e Matemática (MT), se mostrou mais sensível ao estrato da faixa de renda do que aos demais estratos aqui analisados.

Ela apresentou uma variação diretamente proporcional à faixa de renda, sendo a média de 0,27 para a menor faixa de renda e uma média de 0,52 para a maior faixa de renda.

_ Situações de moradia como por exemplo quantidade de residentes até 4 independentemente da renda apresentam as melhores médias, quando comparados com situações de um número maior de residentes, mas ainda inferiores às médias da faixa de renda de R\$3,3k/4,4k independentemente ao número de residentes. Isto porque sabemos que quanto maior a renda no Brasil menos numerosas são as residências familiares.

_ Se olharmos para as duas raças que mais se declararam, brancos e pardos totalizando 85% dos participantes a nota média dos brancos foi de 0,41 enquanto que a de pardos foi de 0,33. Analisando a distribuição de renda entre as duas maiores participações de raças temos:

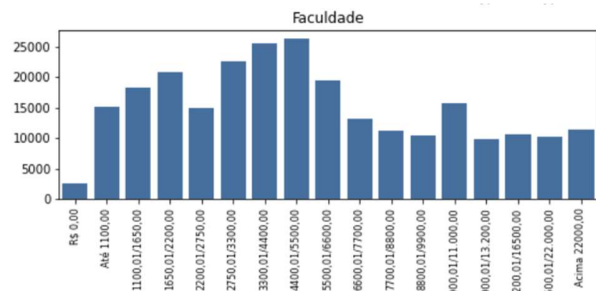
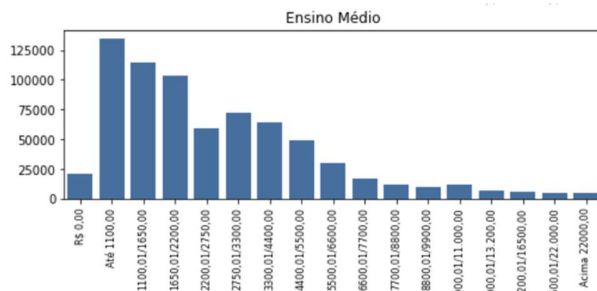


Mais que o dobro de pardos está na faixa de R\$1,1K quando comparados aos brancos.

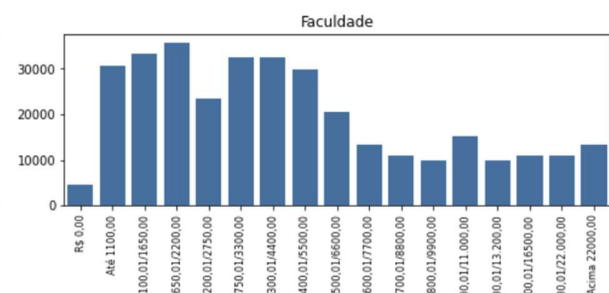
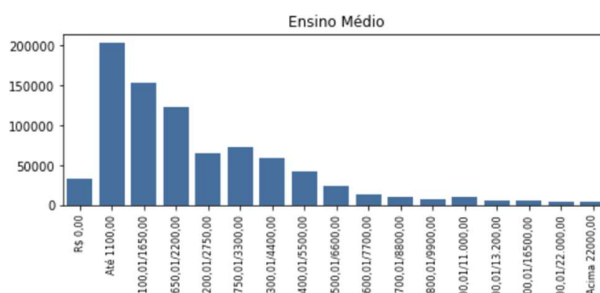
_ Participantes com responsáveis com formação superior também apresentam as melhores médias por volta de 0,45 quando comparados aos responsáveis sem ensino superior, ficando próxima ao valor da mediana da análise da variação da nota pela faixa de renda.

Quando comparamos a distribuição de renda entre ensino médio e nível superior temos:

Responsável Homem

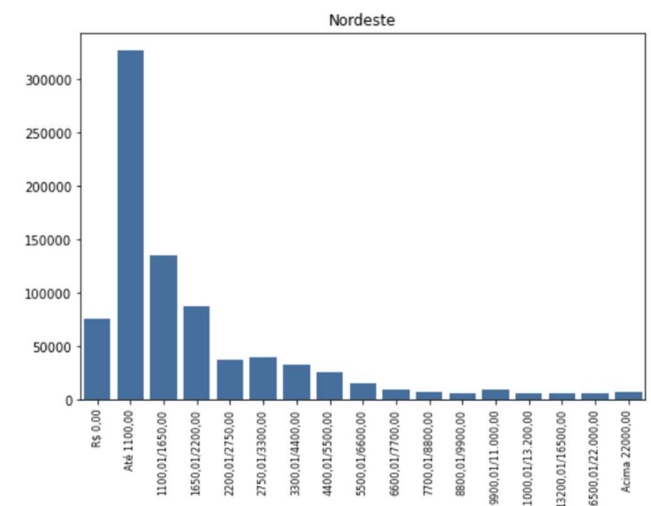
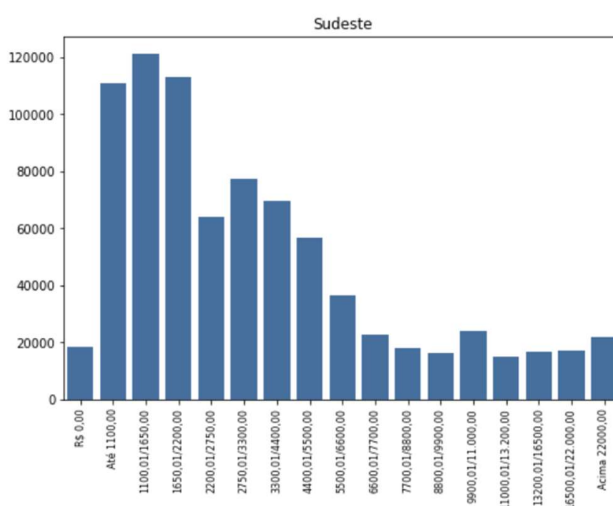


Responsável Mulher



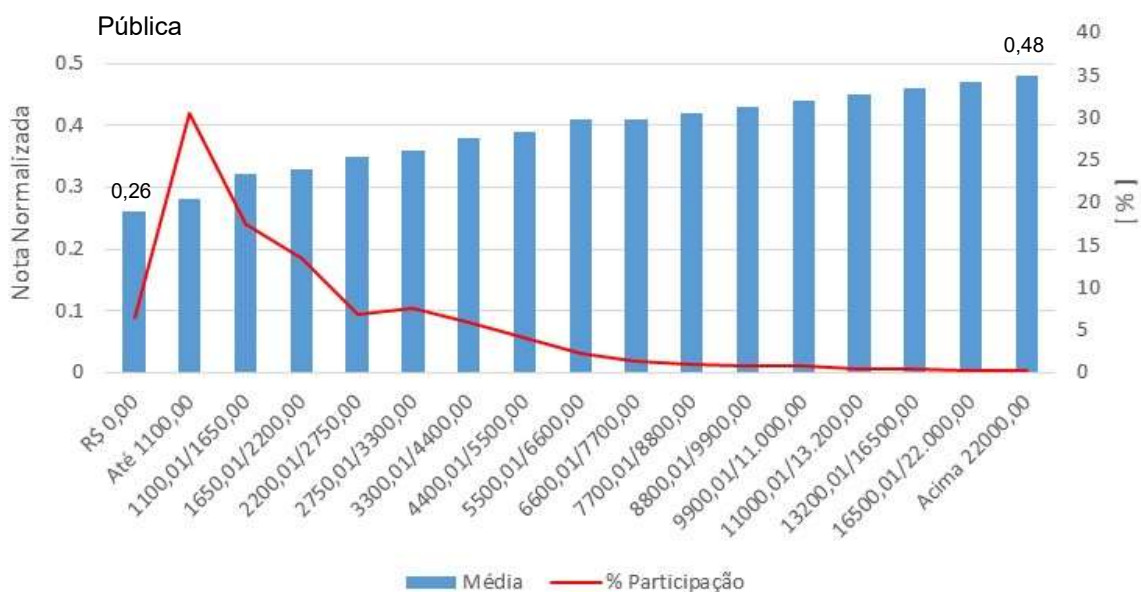
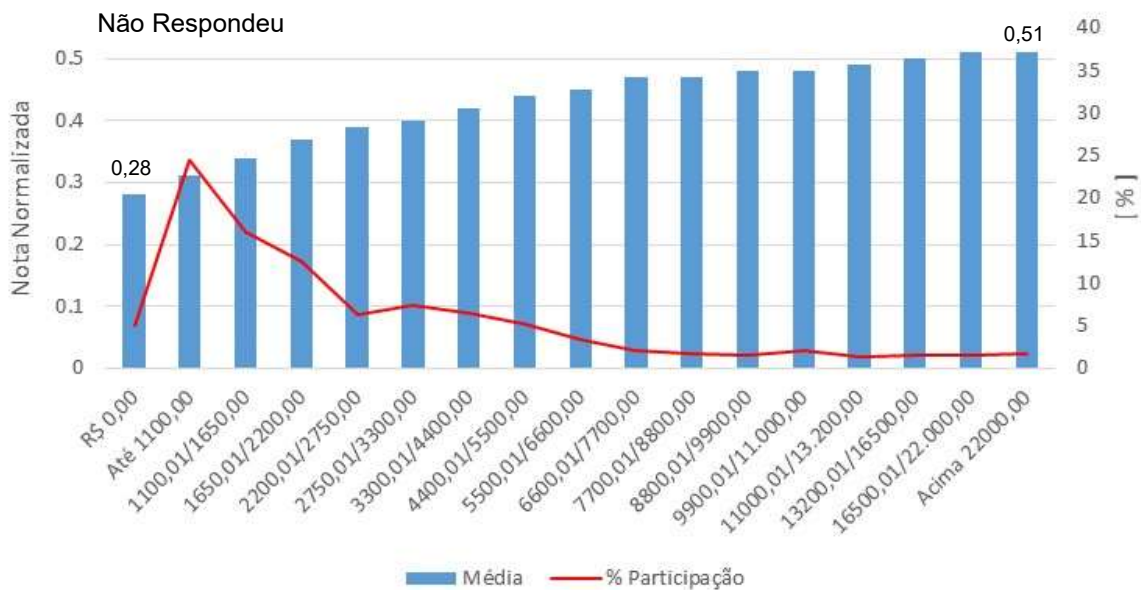
_ Analisando as duas regiões com maior número de participantes o Sudeste com 34% e o Nordeste com 35%, observamos respectivamente 0,4 e 0,34 de valores médios.

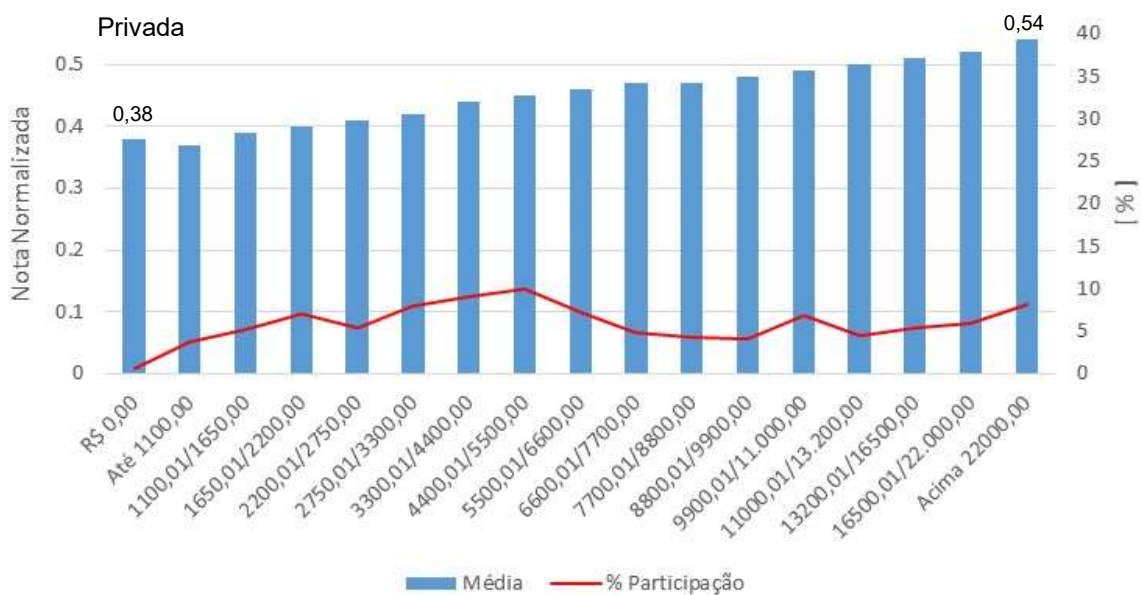
Quando comparamos a distribuição de renda entre Sudeste e Nordeste temos:



Podemos notar uma forte concentração de renda do Nordeste na faixa de R\$1,1K e depois uma queda abrupta nas faixas seguintes, enquanto que a concentração nas faixas iniciais da região Sudeste é aproximadamente a metade.

_ Analisando as instituições de ensino com base na renda observamos a mesma característica da variação da média diretamente proporcional a renda.





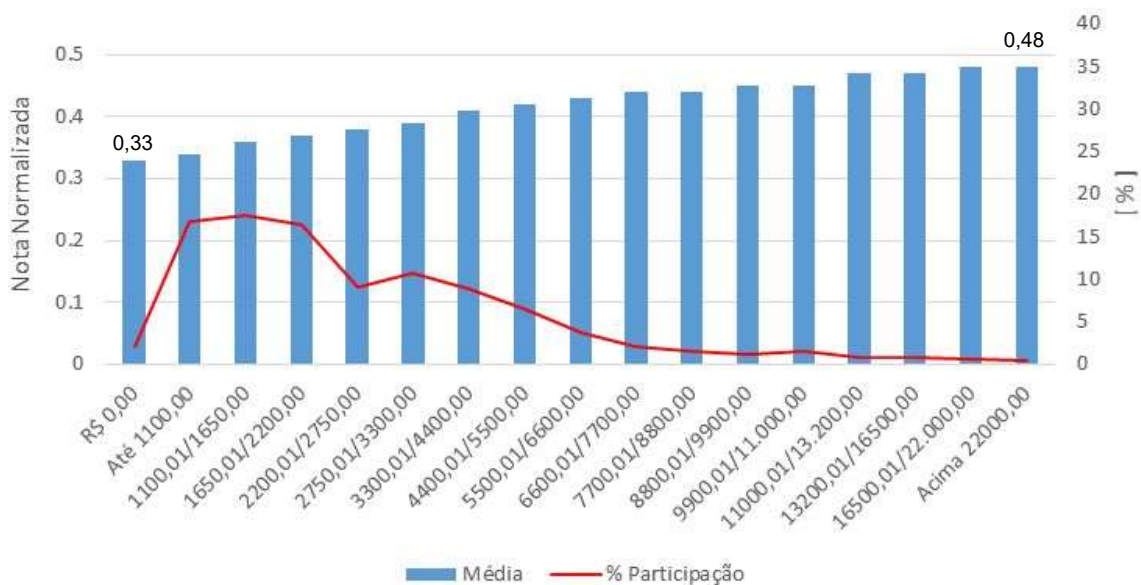
Porém podemos ver a influência do tipo de instituição em relação a média das notas, enquanto a pública vai da média de 0,26 a 0,48 em função da faixa de renda, a privada tem índices maiores indo de 0,38 a 0,54.

Podemos dizer também que para a instituição “Não Respondeu” encontramos a média menor mais próxima a média da pública e a média das duas últimas faixas de renda mais próximas a privada.

_ Com relação ao acesso à internet e o uso de dispositivos digitais (Computador e Telefone Celular).

O acesso à internet entre os participantes é de 91%, é um valor bem expressivo. Entre os que tem acesso à internet 46% possui ao menos um computador em casa.

Analisando este grupo específico com base na renda observamos a mesma característica da variação da média diretamente proporcional a renda.



Porém não ter computador e nem acesso a internet pegando as três primeiras faixas de renda a nota média não ultrapassa 0,29.

Se olharmos a média dos participantes que declaram ter quatro ou mais computadores veremos um valor de 0,51 muito próximo às médias das rendas mais altas, quando analisamos as notas por faixa de renda. Porém o percentual de participantes que indicaram ter mais de dois computadores é 20%. Isto porque a quantidade de computadores está muito próxima ao padrão de renda.

Já a quantidade de celulares de quatro ou mais apresenta uma nota média de 0,41, ligeiramente menor. Porém o percentual de participantes que indicaram ter mais de dois celulares é 84%. Isto porque o celular apresenta opções no mercado com valores menores aos computadores e também por ser um dispositivo muito mais pessoal do que um computador que pode ser compartilhado com mais facilidade entre os moradores da residência. Lembrando que 92% dos participantes declararam ter em média 3 moradores na residência.

6. Conclusão

Sabemos que a faixa de renda apenas não é o fator principal e sim o acesso a que a renda permite, como cultura, viagem, intercâmbio e infraestrutura adequada. Histórico de pais ou responsáveis com formação acadêmica superior também se mostrou importante para o desempenho do participante seja pelo fomento ao conhecimento ou pela melhor renda familiar. Este arca bolso ao qual a renda está intimamente ligada permite que o indivíduo que esteja nesta situação tenha acesso a um conhecimento geral mais completo e robusto.

A internet sozinha também não é responsável por promover uma melhora no conhecimento do participante é necessário todo um ecossistema que fomente tal conhecimento.

7. Referências

- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS
ANÍSIO TEIXEIRA. Microdados do Enem 2018. Brasília: Inep, 2018.
Disponível em:
<<http://portal.inep.gov.br/web/guest/microdados>>. Acesso em: 30 abr. 2017.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS
ANÍSIO TEIXEIRA. Microdados do Enem 2019. Brasília: Inep, 2019.
Disponível em:
<<http://portal.inep.gov.br/web/guest/microdados>>. Acesso em: 20 jun. 2020.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS
ANÍSIO TEIXEIRA. Microdados do Enem 2020. Brasília: Inep, 2020.
Disponível em:
< <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>>. Acesso em: 20 jun. 2021.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS
ANÍSIO TEIXEIRA. Microdados do Enem 2021. Brasília: Inep, 2022.
Disponível em:
<<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>>. Acesso em: 25 mai. 2022.
- 5ª Edição da Pesquisa Retratos da Leitura no Brasil – Divulgada pelo Instituto Pró-Livro – 09/2020.