



**CENTRO UNIVERSITÁRIO INSTITUTO DE  
EDUCAÇÃO SUPERIOR DE BRASÍLIA**  
**Bacharelado em  
Ciência de Dados e Inteligência Artificial**

**Tássio Lucian de Jesus Sales**

**Análise de dados do Enems 2017, 2018, 2019**

**Brasília  
2021**

# Lista de ilustrações

Figura 1 – Distribuição de frequência por ano dos ENEMs 2017-2019 . . . . .	9
Figura 2 – Distribuição de frequência por Região dos ENEMs 2017-2019 . . . . .	10
Figura 3 – Distribuição de frequência por Sexo dos ENEMs 2017-2019 . . . . .	11
Figura 4 – Distribuição de frequência por Cor e raça dos ENEMs 2017-2019 . . . . .	12
Figura 5 – Distribuição de frequência por Tipo de escola dos ENEMs 2017-2019 . . . . .	13
Figura 6 – Distribuição de frequência por capital dos ENEMs 2017-2019 . . . . .	14
Figura 7 – Distribuição de tabelas por nome da unidade federativa dos ENEMs 2017-2019 . . . . .	15
Figura 8 – Distribuição de frequência por nome da unidade federativa dos ENEMs 2017-2019 . . . . .	16
Figura 9 – Analise dados NU-NOTA-CH (ciências humanas e suas tecnologias) dos ENEMs 2017-2019 . . . . .	17
Figura 10 – Analise dados NU-NOTA-CN (ciências da natureza e suas tecnologias) dos ENEMs 2017-2019 . . . . .	18
Figura 11 – Analise dados NU-NOTA-LC (linguagens, códigos e suas tecnologias) dos ENEMs 2017-2019 . . . . .	19
Figura 12 – Analise dados NU-NOTA-MT (matemática e suas tecnologias)) dos ENEMs 2017-2019 . . . . .	20
Figura 13 – Analise dados da Redação dos ENEMs 2017-2019 . . . . .	21
Figura 14 – Analise dados da Nota Media dos ENEMs 2017-2019 . . . . .	22
Figura 15 – Distribuição de frequência por sexo dos ENEMs 2017-2019 limpo . . . . .	25
Figura 16 – Distribuição de frequência por cor e raca dos ENEMs 2017-2019 limpo . . . . .	26
Figura 17 – Distribuição de frequência por tipo de escola dos ENEMs 2017-2019 limpo . . . . .	27
Figura 18 – Analise dados NU-NOTA-MT (matemática e suas tecnologias)) dos ENEMs 2017-2019 limpo . . . . .	28
Figura 19 – Analise dados NU-NOTA-LC (linguagens, códigos e suas tecnologias) dos ENEMs 2017-2019 limpo . . . . .	29
Figura 20 – Analise dados NU-NOTA-CN (ciências da natureza e suas tecnologias) dos ENEMs 2017-2019 limpo . . . . .	30
Figura 21 – Analise dados NU-NOTA-CH (ciências da natureza e suas tecnologias) dos ENEMs 2017-2019 limpo . . . . .	31
Figura 22 – Analise dados da Redação dos ENEMs 2017-2019 limpo . . . . .	32
Figura 23 – Analise dados da Nota Media dos ENEMs 2017-2019 limpo . . . . .	33

# Sumário

1	INTRODUÇÃO . . . . .	5
2	MOTIVAÇÃO . . . . .	6
3	DESENVOLVIMENTO . . . . .	9
4	MISSINGS . . . . .	24
5	RESULTADOS . . . . .	34
6	CONCLUSÕES . . . . .	37
	 ANEXOS	 39
	ANEXO A – CÓDIGOS DOS PROGRAMAS . . . . .	40
	ANEXO B – CONCLUSÃO . . . . .	42



# 1 INTRODUÇÃO

E louvável a preocupação do Ministério da Educação em obter mais informações sobre o ENEM, visto que a vida de milhões de pessoas tem sido influenciada anualmente por esse exame e 112 instituições federais e estaduais de ensino superior utilizam as notas obtidas pelos candidatos em 26 estados e no Distrito Federal.

O ENEM, transformado em grandioso processo seletivo, deve também ter como objetivo balizar e nortear todo o ensino da Educação Básica. Nesse sentido, MEC e INEP podem e devem promover uma ampla discussão sobre o estabelecimento de um programa curricular unificado para todo o território brasileiro. Essa discussão precisa contemplar opiniões não apenas de mestres e doutores em metodologia e didática, mas também de professores que estão no dia a dia com os alunos. Quais são os conteúdos, habilidades e competências relevantes? Quais assuntos que podem ser postergados para o Ensino Superior? Essas são algumas das importantes questões que precisam ser urgentemente respondidas.

Uma das nobres metas do ENEM é influir e orientar a melhoria do Ensino Médio, servindo também como um eficiente processo de seleção para o acesso à Educação Superior, sendo também um dos propulsores de uma grande mudança na educação nacional, embora alguns ajustes ainda possam ser feitos. Em comparação com a maioria dos vestibulares do país, o ENEM mostra-se superior em diversos aspectos, mas isso não significa que não deva ser aprimorado.

Com o intuito de auxiliar na melhoria do Exame Nacional do Ensino Médio, apresentamos a seguir algumas análises das 4 grandes áreas por ele exploradas (Linguagens e Códigos, Matemática, Ciências da Natureza e Ciências Humanas). Foram realizadas considerações a respeito das matrizes de referências (competências, habilidades e objetos de conhecimento), das qualidades e falhas de algumas questões apresentadas nos últimos exames, de características da prova de redação, entre outras.

## 2 MOTIVAÇÃO

*O que é o Enem?* Enem é a sigla de Exame Nacional do Ensino Médio e foi criado pelo Ministério da Educação e Cultura (MEC) para testar o nível de aprendizado dos alunos que concluíram o ensino médio no Brasil.

Atualmente, os resultados obtidos no Enem ajudam os estudantes a ingressar em universidades públicas ou a ganhar bolsas de estudos em instituições particulares.

Criado em 1998, o Enem é o maior processo seletivo do Brasil, abrangendo aproximadamente mais de 7 milhões de candidatos inscritos para a realização das provas.

### **Finalidade do Enem**

Além de ajudar o governo a traçar um parâmetro de comparação sobre a qualidade do ensino médio brasileiro, o resultado obtido no Enem também pode ser utilizado para que o concursando possa ingressar no ambiente universitário, através de várias possibilidades.

O Enem também é uma “chave” para aqueles que desejam estudar no exterior. Algumas universidades públicas de Portugal, como a Universidade de Coimbra, por exemplo, aceitam a nota do Enem como critério para alguns cursos.

Todas as universidades públicas e institutos federais do Brasil utilizam a nota do Enem como critério para selecionar os alunos que irão ocupar as vagas disponíveis em todos os cursos.

### **Quem pode fazer o Enem?**

Podem participar do Enem os estudantes que estão concluindo o último ano do ensino médio ou aqueles que já o concluíram.

Além disso, é importante referir que aqueles que não tiverem 18 anos completos no dia da prova não poderão utilizá-la como forma de ingresso ou bolsa em instituições.

Os participantes com esse perfil são chamados de treineiros, pois a prova servirá apenas de treino.

No caso de o interessado em fazer a prova estar cumprindo medidas ou penas que o privem de liberdade, ele deve aguardar um edital específico para sua categoria.

A realização do Enem não é obrigatória para que o estudante possa concluir o ensino médio. No entanto, é essencial para aqueles que desejam seguir seus estudos na universidade.

## **Etapas do Enem**

1. Ciências Humanas e suas Tecnologias (História, Geografia, Sociologia e Filosofia).
2. Ciências da Natureza e suas Tecnologias (Química, Física e Biologia).
3. Matemáticas e suas Tecnologias (Matemática).
4. Linguagens, Códigos e suas Tecnologias e Redação (Língua Portuguesa, Literatura, Língua Estrangeira).
5. redação, o candidato deve apresentar as quatro competências exigidas:
  - a) Demonstrar domínio da norma padrão da língua escrita
  - b) Selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista
  - c) Elaborar proposta de intervenção para o problema abordado, respeitando os direitos humanos
  - d) Demonstrar conhecimentos dos mecanismos linguísticos necessários para a construção da argumentação

## **Nota de corte do Enem**

A nota de corte do Enem é a nota mínima necessária para que um candidato possa ser aprovado em um dos programas do governo de acesso ao Ensino Superior

As notas de corte podem variar de acordo com a instituição e/ou com o curso pretendido pelo candidato.

## **Enem e Sisu**

Para auxiliar no processo de seleção do Enem, foi criado o Sistema de Seleção Unificada (Sisu), onde o candidato, a partir da média obtida no Exame Nacional do Ensino Médio, consegue concorrer a uma vaga em um curso de graduação em qualquer universidade pública federal (e algumas estaduais) do país.

A universidade possui total autonomia para decidir como a nota do Enem será utilizada neste processo, seja como requisito para uma primeira fase no processo (tendo o candidato que se submeter a uma outra prova desenvolvida pela universidade à qual se candidata), ou como critério único para ocupação das vagas.

## **Enem e Prouni**

A nota do Enem também é a base para a participação no Programa Universidade para Todos (ProUni), criado pelo governo federal.

O ProUni reserva bolsas de estudo, parciais e integrais, para os estudantes que tenham obtido uma boa média no Enem e que tenham estudado em escolas públicas (ou com bolsa de estudo em escolas particulares) durante todo o ensino médio.

### **A ciência de dados e a inteligência artificial**

A ciência de dados e a inteligência artificial são áreas aplicadas de forma intercambiável nas empresas. É importante saber que, embora os dados possam necessitar de alguns aspectos da IA, eles não refletem tudo. A inteligência artificial está em um processo constante de evolução e ainda há muito o que mudar e explorar.

A ciência de dados é uma tecnologia que vem conquistando indústrias em todo o mundo e contribuiu para a quarta revolução industrial que ouvimos tanto falar.

Isso é uma consequência da explosão maciça de informações disponíveis na internet e a crescente necessidade das organizações de se basear nos dados, para criar produtos melhores ao mercado. Isso porque é a partir deles que os negócios tomam decisões cuidadosas em suas estratégias de atuação.

As várias fases e procedimentos na ciência de dados envolvem coleta, manipulação, visualização e gerenciamento de dados para prever tendências e a ocorrência de eventos futuros. Sendo assim, o profissional precisa ter um conhecimento sobre os algoritmos de aprendizado de máquina. E é aí que entra a inteligência artificial.

As empresas precisam que os cientistas de dados ajudem a tomar as decisões necessárias, baseadas em dados. Eles auxiliam a entender e analisar os seus próprios desempenhos de operação e também sugerem as melhorias necessárias para impulsionar seus resultados. E também ajudam a equipe de desenvolvimento de produtos a aplicar melhorias para os clientes finais, analisando seu comportamento a partir das informações.



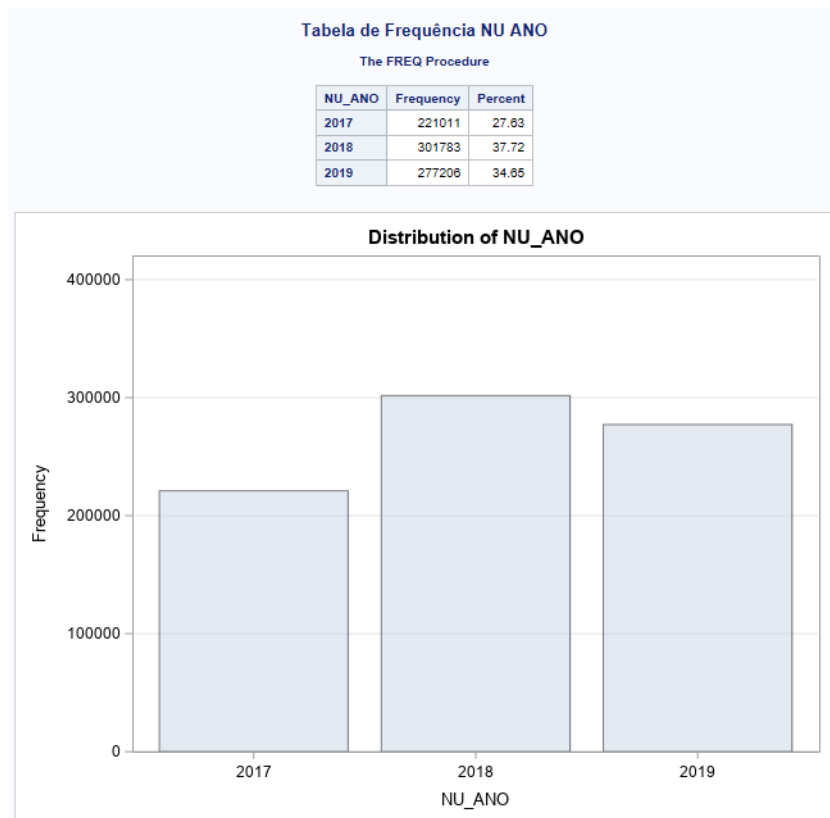
### 3 Desenvolvimento

Aqui onde irei desenvolver uma análise exploratória usando o sas. Os dados usados e uma amostra dos enems de 2017, 2018, 2019. Irei separar nas variáveis qualitativas e quantitativas

**Variáveis qualitativas:** são aquelas cujos valores podem ser separados em diferentes categorias que se distinguem por alguma característica não numérica. Por exemplo: sexo (masculino e feminino), cor dos cabelos (preto, loiro, ruivo, castanho, etc). Podem ser subdivididas em:

1. **Variáveis qualitativas ordinais** – quando existe uma ordem nos seus valores. Por exemplo, a variável “Grau de instrução” pode ter seus valores ordenados (fundamental, médio, superior, etc). O mesmo não ocorre com a variável “cor da pele”.
2. **Variáveis qualitativas nominais** – quando uma ordem não pode ser estabelecida entre seus valores.

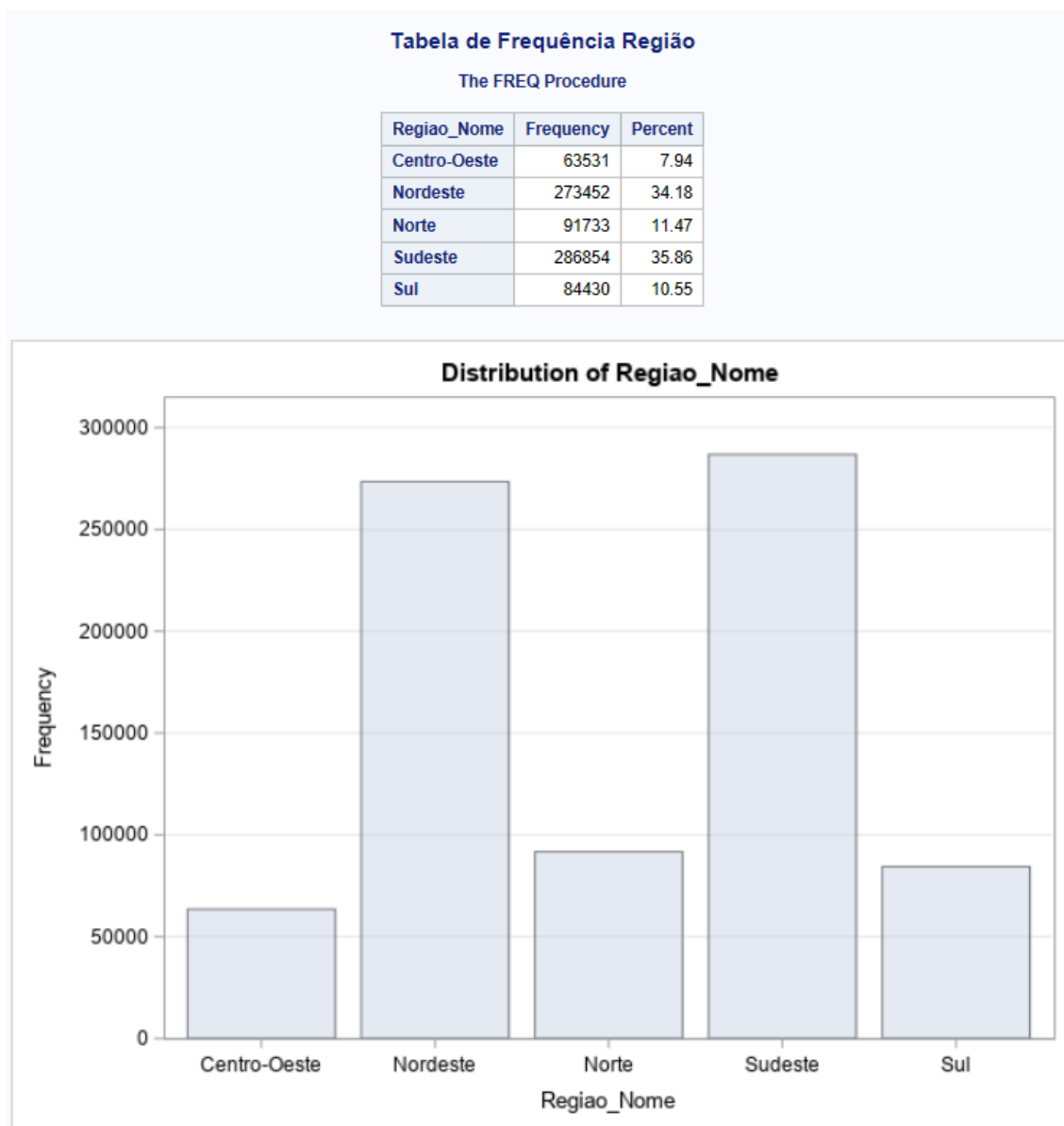
Figura 1 – Distribuição de frequência por ano dos ENEMs 2017-2019



Fonte: SaS

Como podemos verificar a tabela e histograma acima, notamos que de uma amostra de 800 mil inscritos, o ano de 2017, teve 221 mil inscritos, 2018 com 301 mil inscritos e 2019 com 277 mil inscritos. De forma geral o ano de 2018 foi o ano que tivemos mais inscritos. E o ano de 2017 foi o que tivemos menos inscritos. Usei a variável ano com qualitativa, para deixar mais fácil o entendimento dos gráficos já que não existem muito cálculos para serem feitos apenas. Porém vale lembrar que daria muito bem para usar o ano como quantitativa e, na verdade, faz até mais sentido de usarmos.

Figura 2 – Distribuição de frequência por Região dos ENEMs 2017-2019

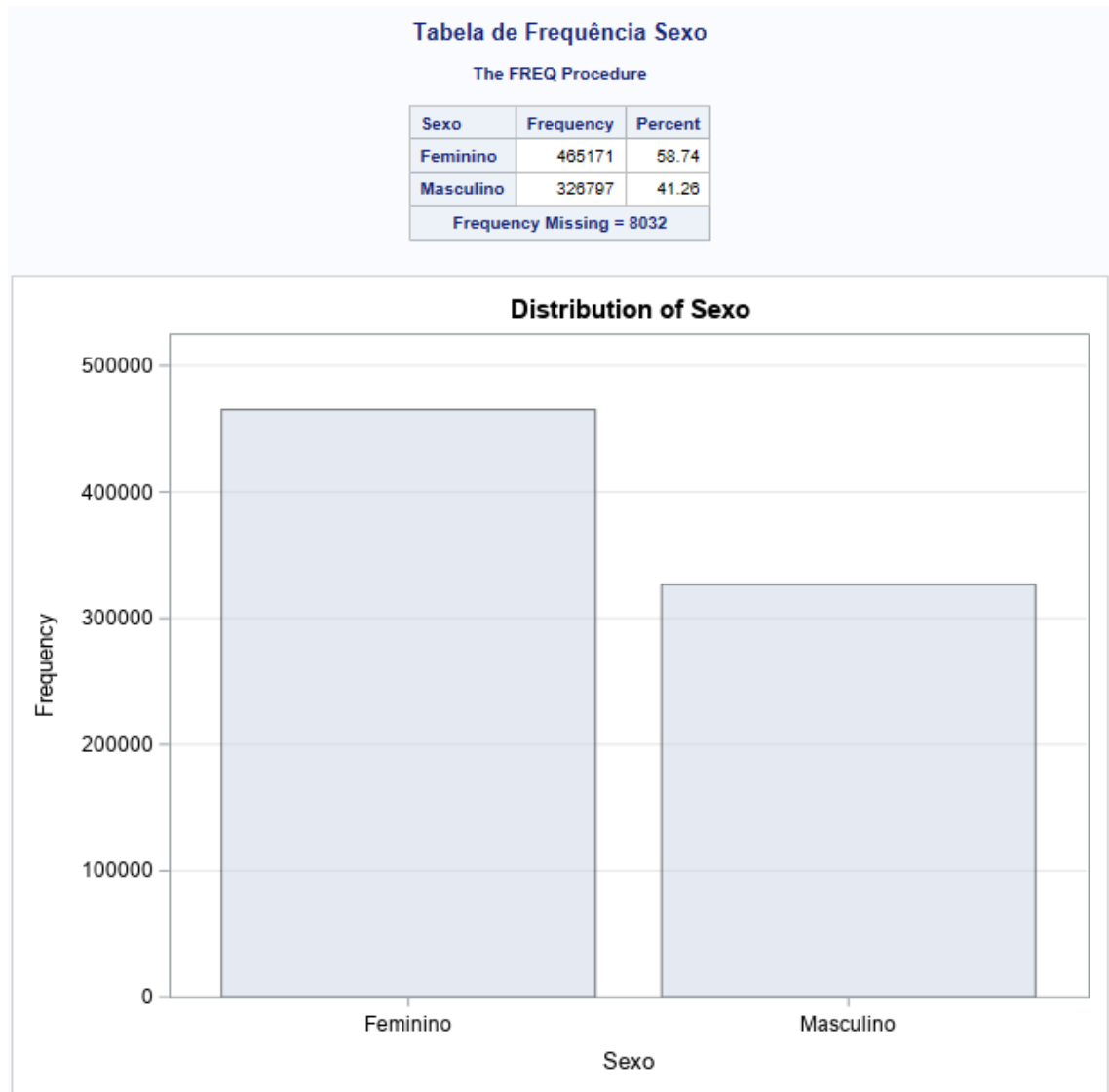


Fonte: SaS

Ao verificar a tabela acima notamos que as regiões centro-oeste, sul e norte são as que tiveram menos participantes creio que por serem as menos populosas, Já as regiões nordeste e sudeste foram as que tiveram mais participantes praticamente o dobro ou o

triplo das outras regiões se formos comparar, mas também acredito que las tem mais inscritos por serem as regiões com maior índice populacional do Brasil

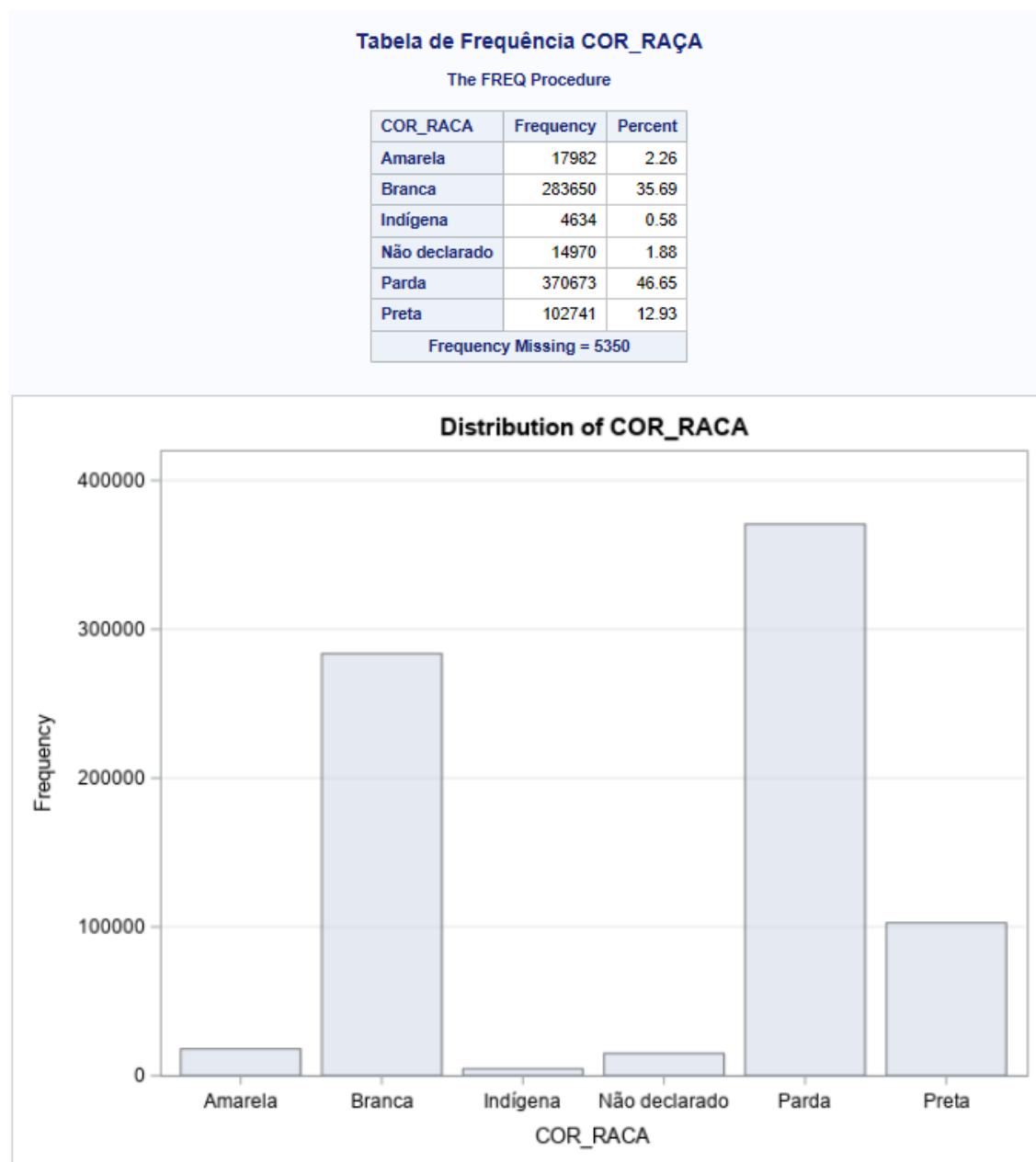
Figura 3 – Distribuição de frequência por Sexo dos ENEMs 2017-2019



Fonte: SaS

Se verificamos o gráfico acima fica notamos que as mulheres são as integrantes com mais participações no somatório de todos os. anos do Enem. E se comparamos com os. homens podemos notar eles têm uns 25% a menos de participação nas provas

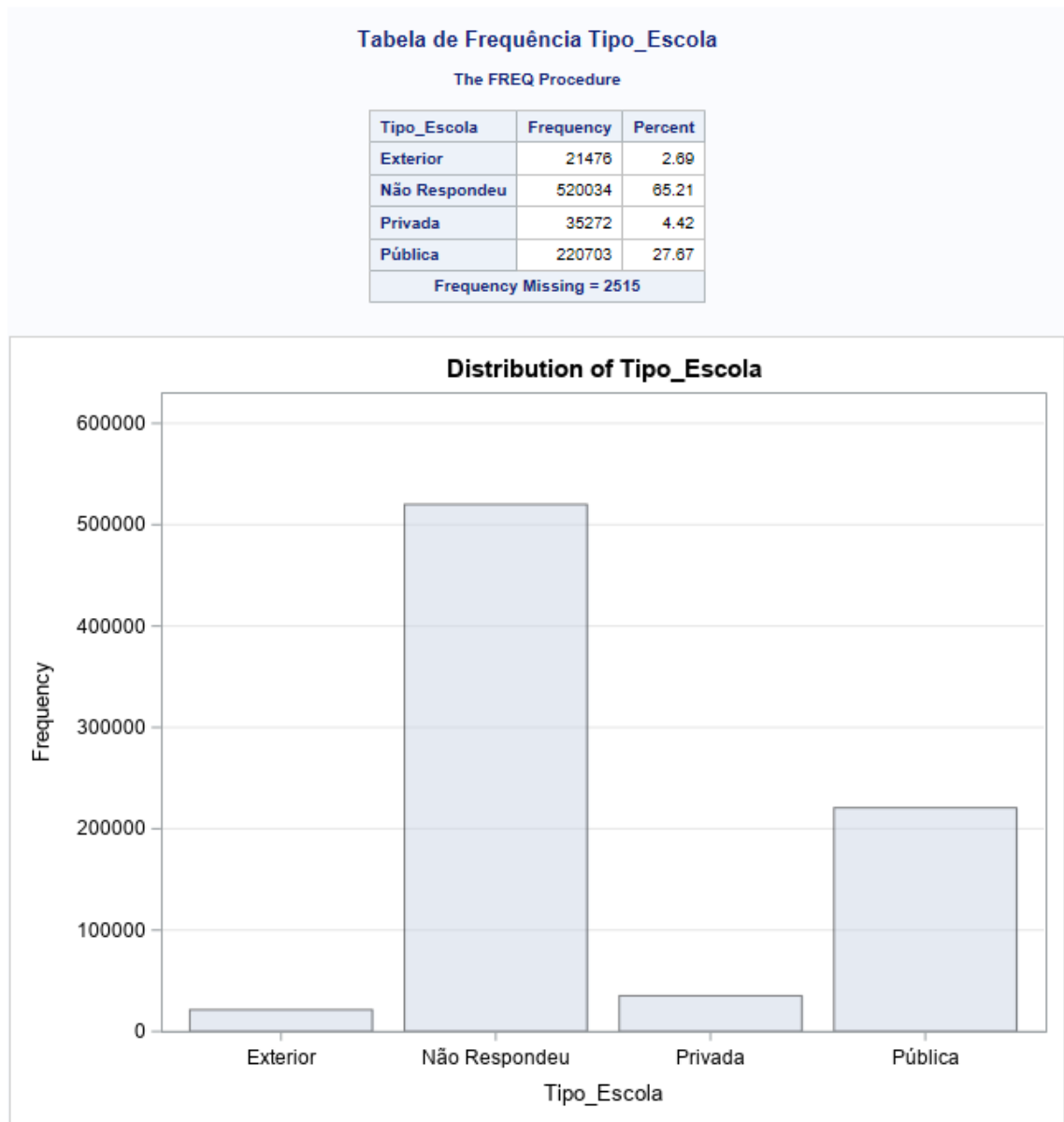
Figura 4 – Distribuição de frequência por Cor e raça dos ENEMs 2017-2019



Fonte: SaS

Aqui como podemos ver a maioria dos alunos que participam da prova do enem são de maioria parda e branca. E uma pequena parte profere não declarar. Nossa minoria é formada por negros, pardos e amarelos. Creio que haja muito desigualdade social ainda no enem.

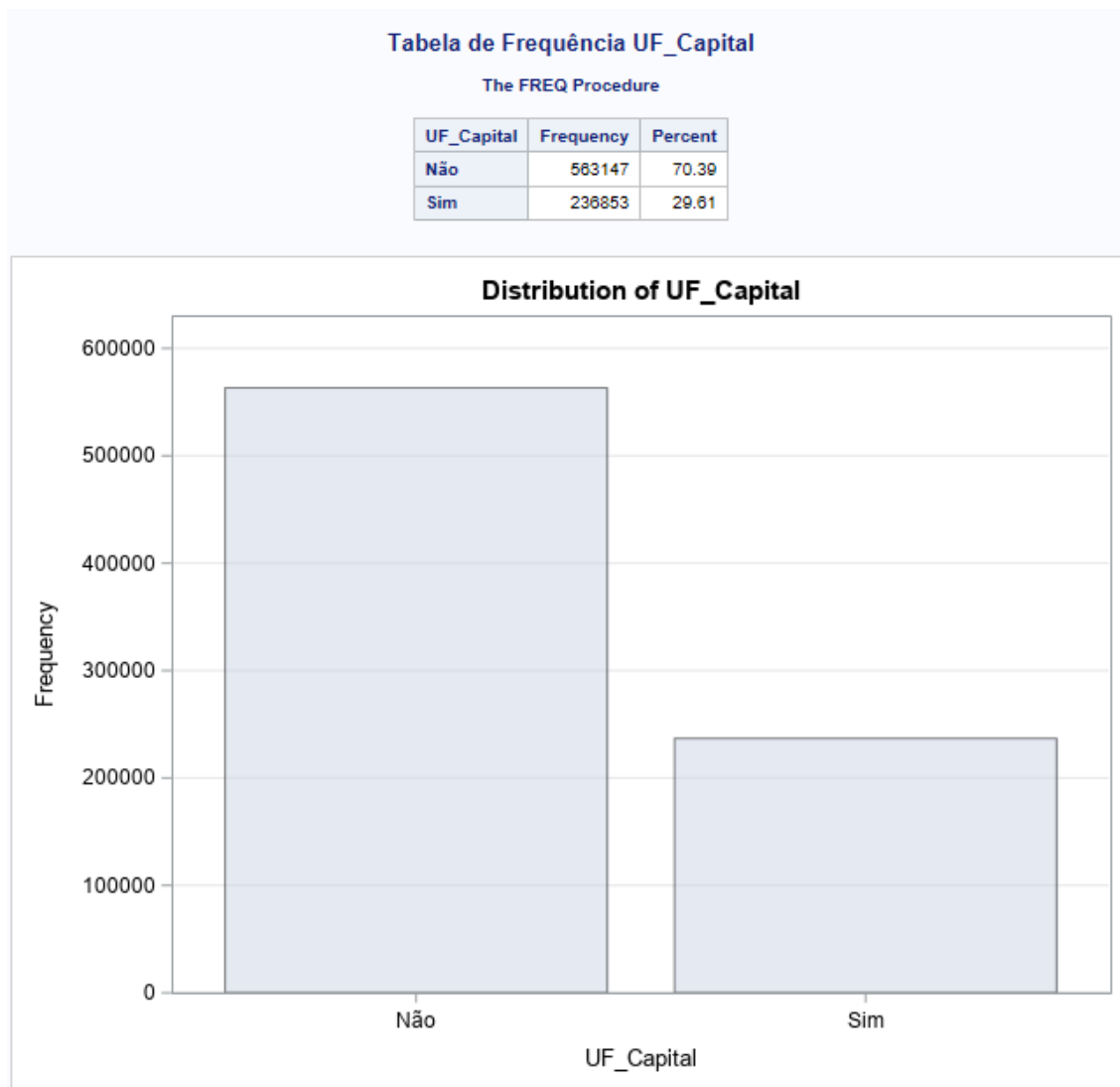
Figura 5 – Distribuição de frequência por Tipo de escola dos ENEMs 2017-2019



Fonte: SaS

Aqui podemos analisar que a maior parte dos participantes optou pela opção de não responde. Já, por outro lado, se analisarmos os alunos que optaram por responder notamos que maior parte dos alunos são de escolas públicas seguido por alunos de escolas privadas. Já a minoria são pessoas que frequentaram escolas do exterior. Bom creio que seria mais viável analisar esses dados se maior parte dos alunos tivessem optado por responder o questionário.

Figura 6 – Distribuição de frequência por capital dos ENEMs 2017-2019



Fonte: SaS

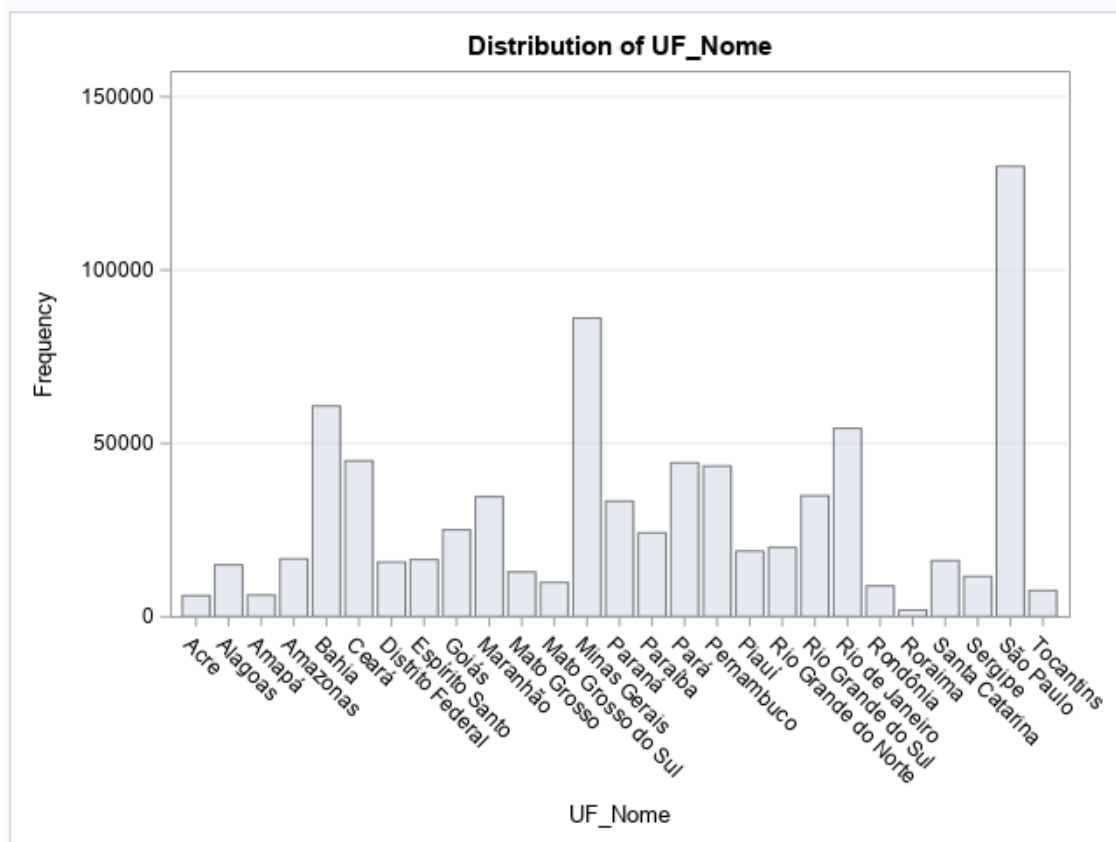
Aqui verificamos onde maior parte das provas foram aplicadas se foram em capitais ou não, E notamos que maior parte das provas não são aplicadas nas capitais brasileiras. O que já era esperado pelo fato de o número de municípios ser bem maior que o número de capitais.

Figura 7 – Distribuição de tabelas por nome da unidade federativa dos ENEMs 2017-2019

Tabela de Frequência UF_Nome		
The FREQ Procedure		
UF_Nome	Frequency	Percent
Acre	6086	0.76
Alagoas	14936	1.87
Amapá	6249	0.78
Amazonas	16685	2.09
Bahia	60781	7.60
Ceará	44967	5.62
Distrito Federal	15750	1.97
Espírito Santo	16450	2.06
Goiás	25032	3.13
Maranhão	34602	4.33
Mato Grosso	12883	1.61
Mato Grosso do Sul	9866	1.23
Minas Gerais	86139	10.77
Paraná	33334	4.17
Paraíba	24203	3.03
Pará	44394	5.55
Pernambuco	43441	5.43
Piauí	18913	2.36
Rio Grande do Norte	19989	2.50
Rio Grande do Sul	34911	4.36
Rio de Janeiro	54339	6.79
Rondônia	8866	1.11
Roraima	1914	0.24
Santa Catarina	16185	2.02
Sergipe	11620	1.45
São Paulo	129926	16.24
Tocantins	7539	0.94

Fonte: SaS

Figura 8 – Distribuição de frequência por nome da unidade federativa dos ENEMs 2017-2019



Fonte: SaS

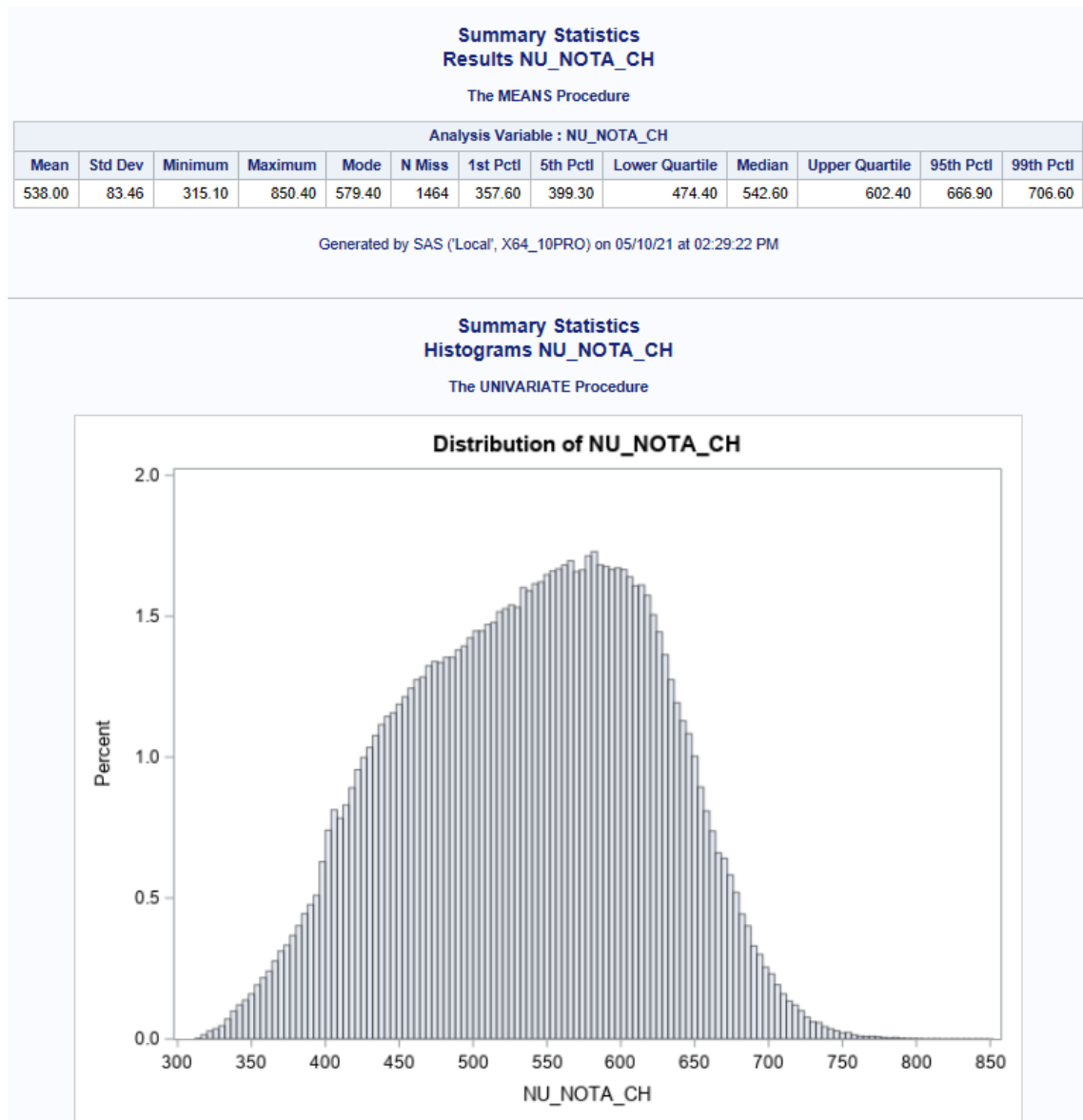
O gráfico acima deixa visível que o estado de São paulo e o estado com mais participantes somente ele tem 16% dos participantes seguidos dos estados de minas gerais com 10% dos integrantes somados chegamos a 26% dos integrantes em somente dois estados. Já a minoria dos participantes está nos estados Roraima, Acre e tocantins todos com menos de 1% e se somarmos eles 1,94% dos integrantes.

**Variáveis quantitativas** – São aquelas cujos valores são expressos em números. As variáveis quantitativas se subdividem, ainda, em:

1. **Variáveis quantitativas discretas** – quando resultam de um conjunto finito (ou enumerável) de valores possíveis. Por exemplo, o número de livros de Matemática da biblioteca da escola e a quantidade de carros presentes, agora, num estacionamento, são variáveis discretas.
2. **Variáveis quantitativas contínuas** – quando resultam de um número infinito de valores possíveis que podem ser associados a pontos em uma escala contínua, de modo que não haja lacunas ou interrupções. Por exemplo, o peso de um pacote de arroz e o comprimento de um parafuso constituem variáveis contínuas.



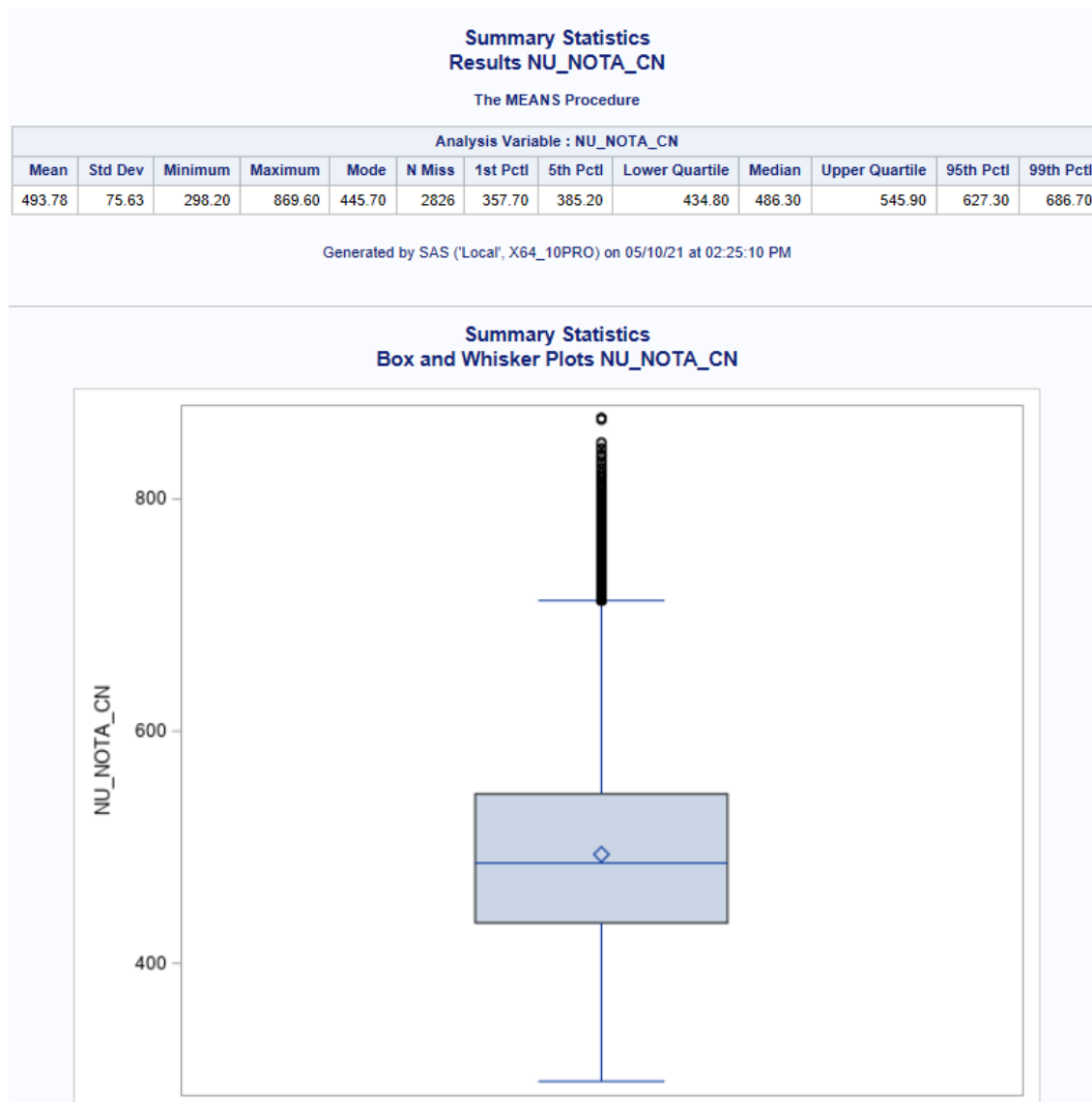
Figura 9 – Análise dados NU-NOTA-CH (ciências humanas e suas tecnologias) dos ENEMs 2017-2019



Fonte: SaS

Analisando a tabela vemos que a menor nota encontrada é 315,10 a nota máxima é 850,40. A quantidade de missings presentes é de 1464 o que é pequeno para uma amostra de 800 mil. A média e mediana estão bem próximas em valores. Nossa média e moda tem uma pequena diferença também o suficiente para que nosso histograma tenha uma pequena assimetria para direita.

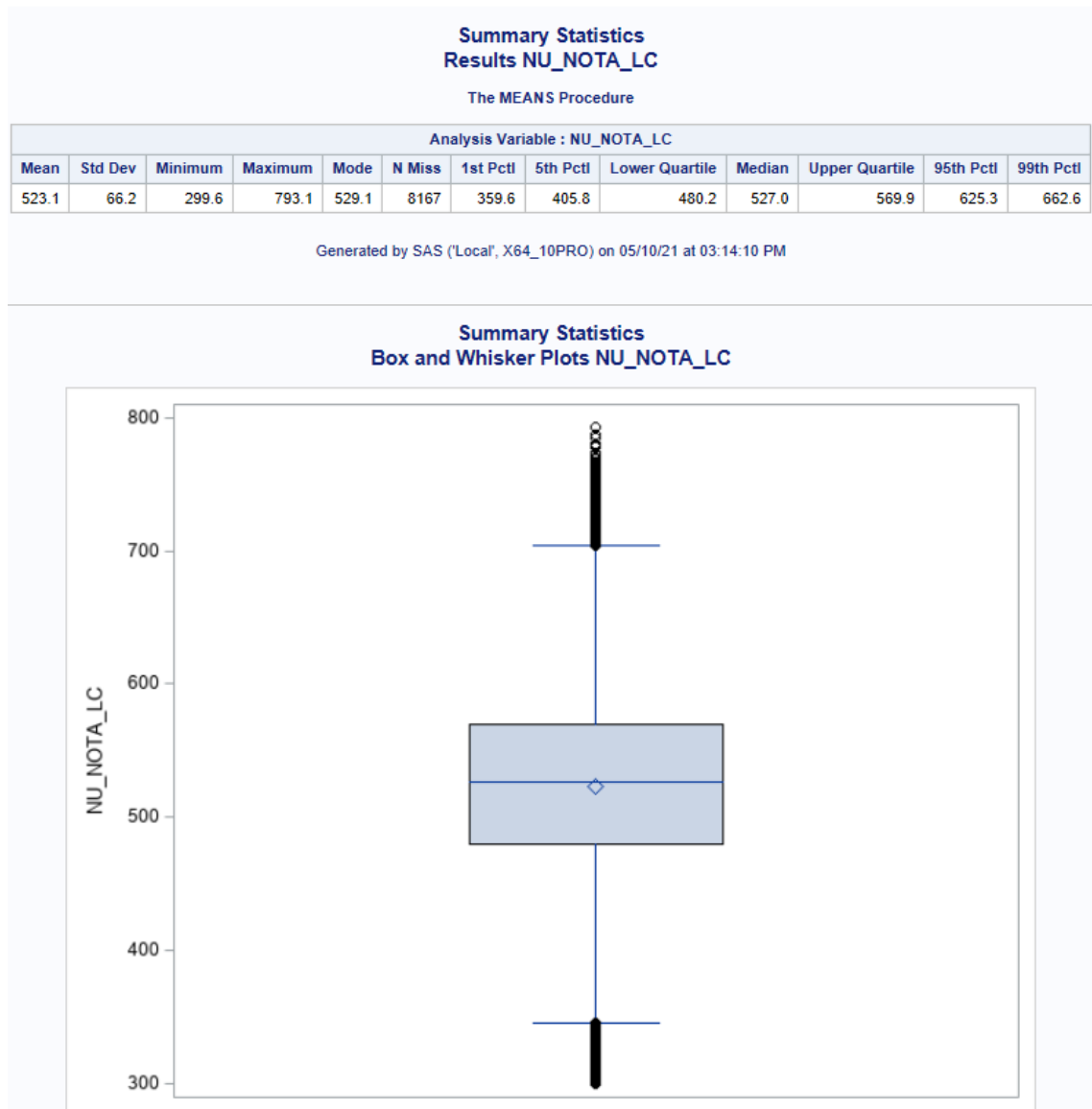
Figura 10 – Analise dados NU-NOTA-CN (ciências da natureza e suas tecnologias) dos ENEMs 2017-2019



Fonte: SaS

Aqui encontramos 2876 missings. A nota média dos alunos é de 493,78 o que é bem próximo da nossa mediana que é de 486,30. A maior nota encontrada foi 869,60 e a menor nota é 298,20. Temos muitos outliers. O que significa que temos muitos valores discrepantes acima do terceiro quartil.

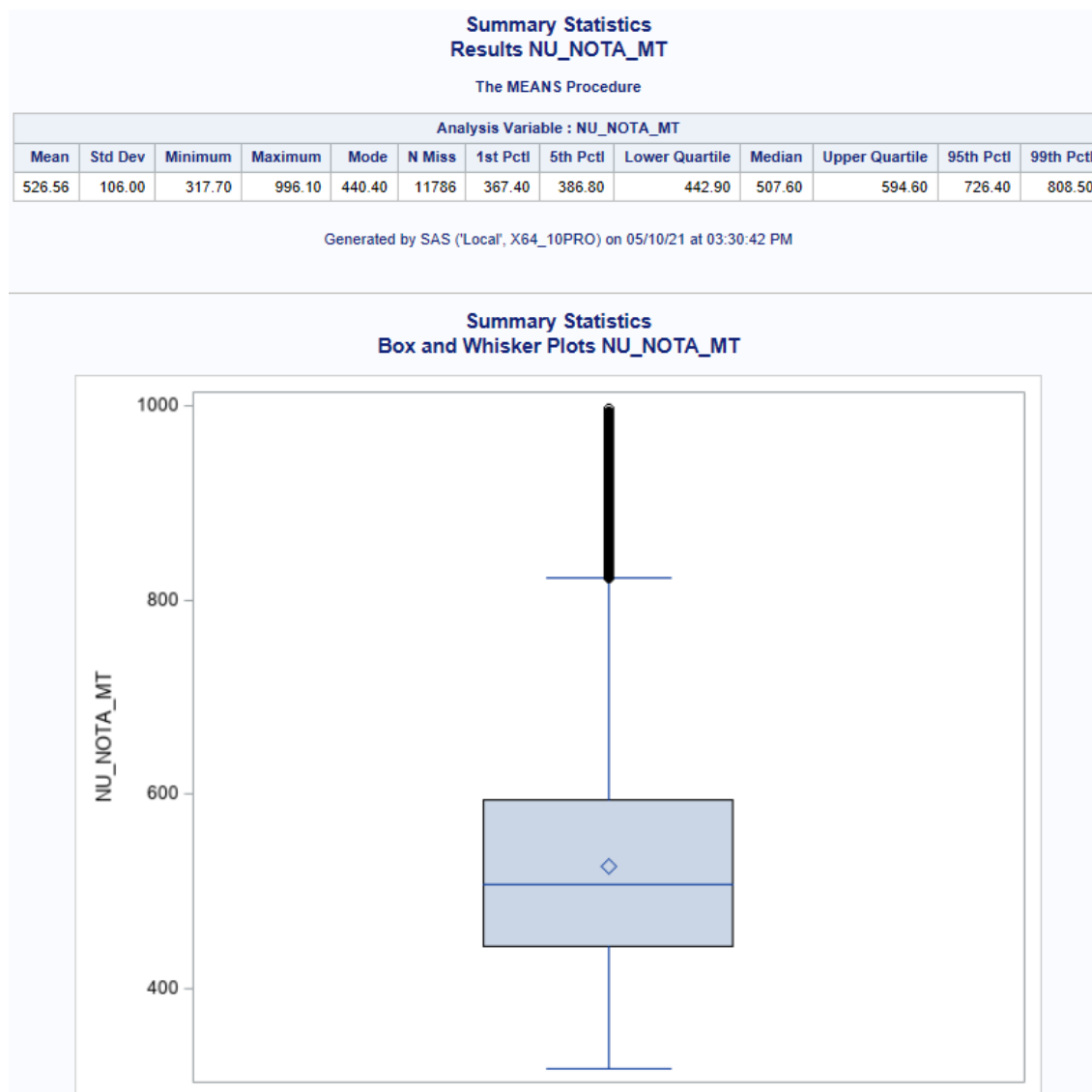
Figura 11 – Análise dados NU-NOTA-LC (linguagens, códigos e suas tecnologias) dos ENEMs 2017-2019



Fonte: SaS

Aqui como podemos observar temos muito outlaiers. Muito abaixo do primeiro quartil e muito cima do terceiro. Também podemos observar que nossas média e media estão bem próximas. Média com 523,1 e mediana como 527,0. O número de missings na prova de linguagens e condigo cresceu baste se comparado as matérias anteriores chegando a 8167, o que já e quase 1% dos dados porem e insignificantes ainda. Nossa nota mínima e 296,6 e nossa nota máxima e 793,1.

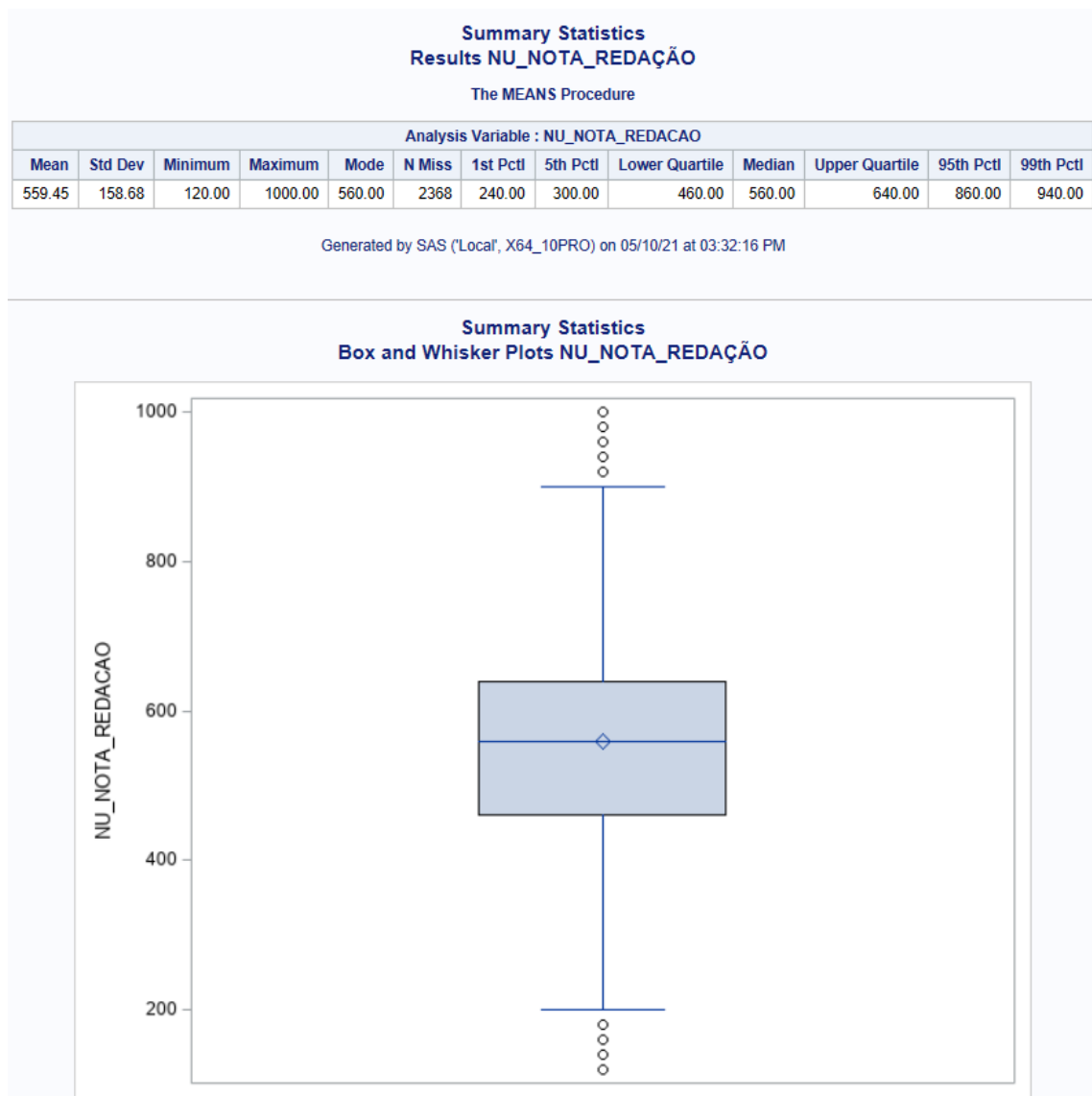
Figura 12 – Análise dados NU-NOTA-MT (matemática e suas tecnologias)) dos ENEMs 2017-2019



Fonte: SaS

No gráfico acima observamos que nossa média e media já estão um pouco afastados. E temos muitos acima do terceiro quartil. Nossa moda e 440,40 bem distante da nossa média e mediana. Nosso número de missings novamente de um salto e chegamos a 11 786. A nota máxima também deu salto chegou a 996,10 o que e bem próximo da nota total da prova. Nosso valor mínimo manteve o padrão das provas anteriores que fica entre 290 a 330. Nesse caso nossa nota mínima foi de 317,70.

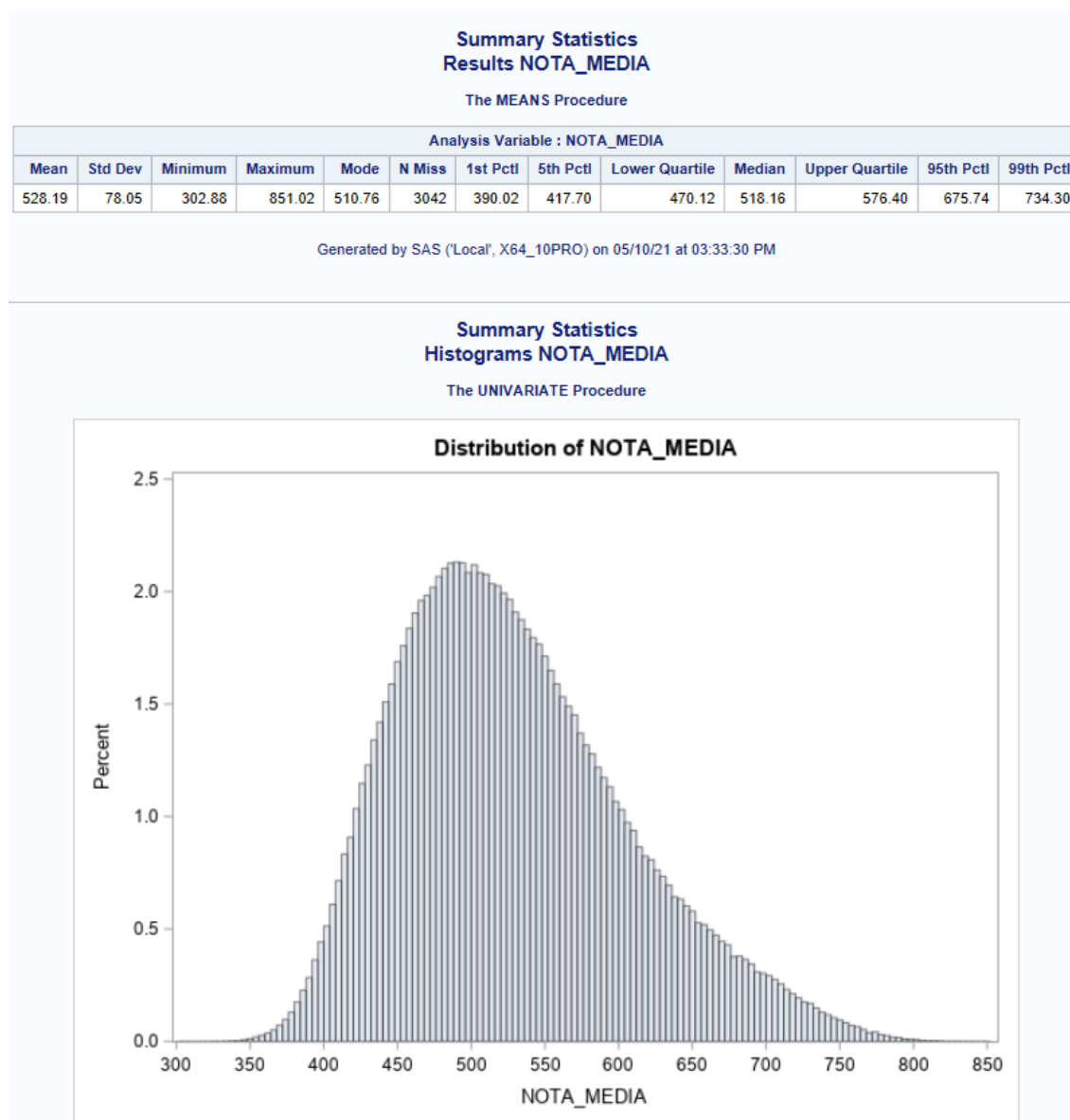
Figura 13 – Análise dados da Redação dos ENEMs 2017-2019



Fonte: SaS

. Aqui tivemos que a nota mais alta foi 1000. Assim concluímos que a única matéria que obteve nota máxima que e 1000 foi a de redação. Porem também tivemos que a menor nota em todas as matérias foram em redação que foi 120, nossa média, moda, e mediana são basicamente idênticas. Que nos dá, um gráfico assimétrico. Nosso missings e de 2368 e nosso outliers são poucos.

Figura 14 – Analise dados da Nota Media dos ENEMs 2017-2019



Fonte: SaS

Aqui quase obtemos um grafico assimetrico. porem nosso grafico tem um leve assimetria a direita. Obtivemso uma maxima bem impressionante que e de 851,02. Nossa media geral foi de 528 e nossa moda de 510. o que nos mostra que maior concentração de notas esta exatamente entre nossa moda e nossa media.

M

## 4 Missings

Um passo importante quando se trata de missing values é entender os motivos pelos quais esses dados estão ausentes, para entender por que isso acontece.

A missing value é um valor ou uma observação, que não foi calculado, devido ao impossível de existir tal valor ou variável. Ao inserirmos um missing value simplesmente estamos ignorando a observação, que não realizamos o cálculo desse valor. No entanto, ao ignorarmos a sua existência, o beneficiamos em uma possível análise futura. As análises podem ser realizadas posteriormente, através de tabelas e gráficos, com os valores redimensionados como todos os outros. No entanto, não existem as mesmas possibilidades de análise no caso de um missing value.

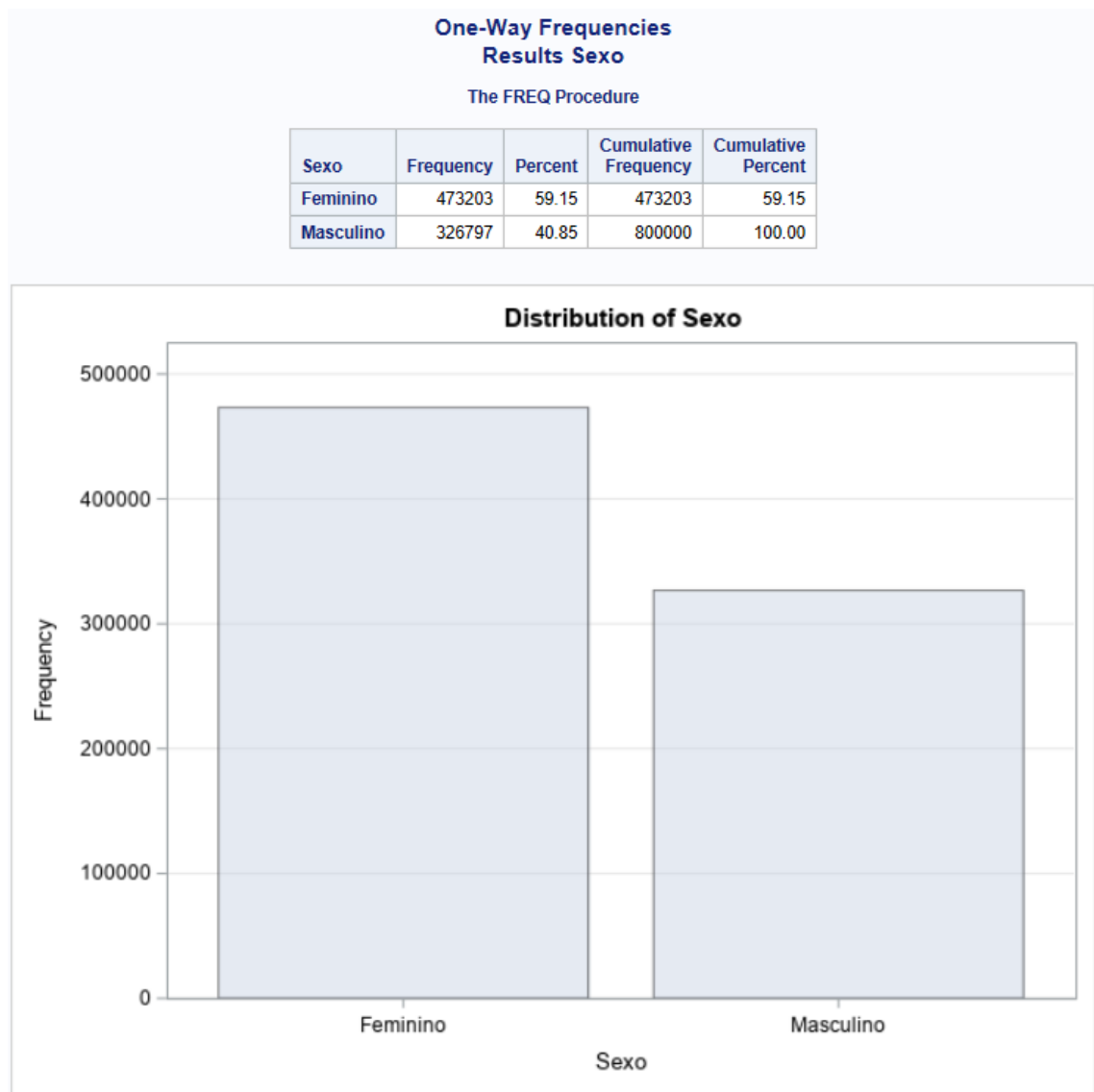
### Métodos

1. Excluir observações que possuam qualquer tipo de missing: esse método é chamado de listwise deletion, e envolve a exclusão de toda linha que tenha pelo menos um valor missing.
2. Excluir observações que possuam missing na variável que você quer utilizar para a análise: esse método é chamado de pairwise deletion, e baseia-se na exclusão daquela linha durante ou para a execução de uma análise. Assim, essa linha não é excluída do dataset necessariamente.
3. Substituir o valor missing por outros valores: isso envolve a substituição do valor missing de uma célula por um valor relevante para aquela observação, como a média, mediana, imputação por regressão etc.



## Analises dos variáveis que tinham valores missings

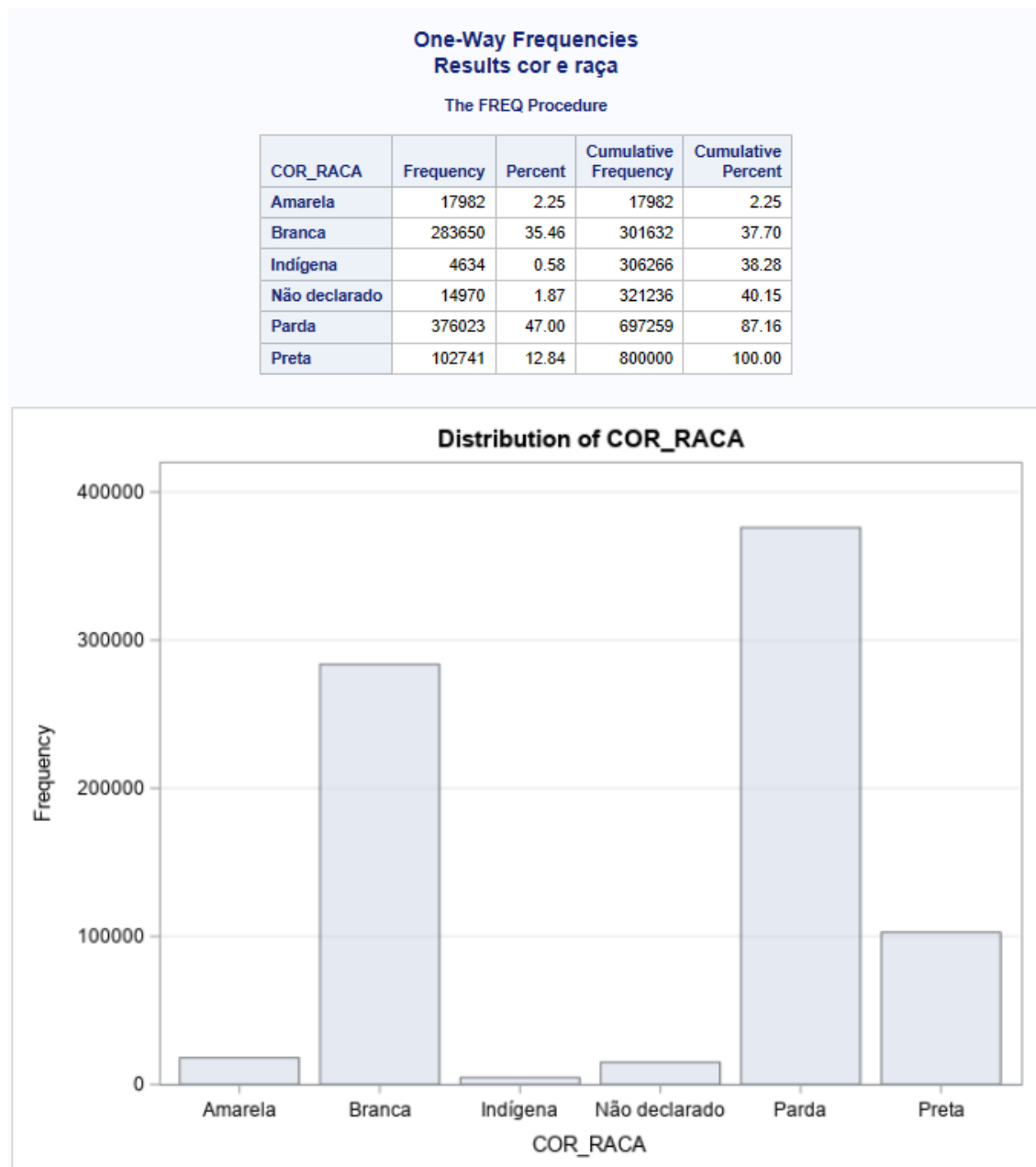
Figura 15 – Distribuição de frequência por sexo dos ENEMs 2017-2019 limpo



Fonte: SaS

Aqui eu realizei a realocação dos valores missings para o valor dentro da variável. Assim vemos que a quantidade de mulheres se comparado a tabela de sexo anterior subiu. Porém a análise continua a mesma com as mulheres ainda sendo maioria nas provas do enem.

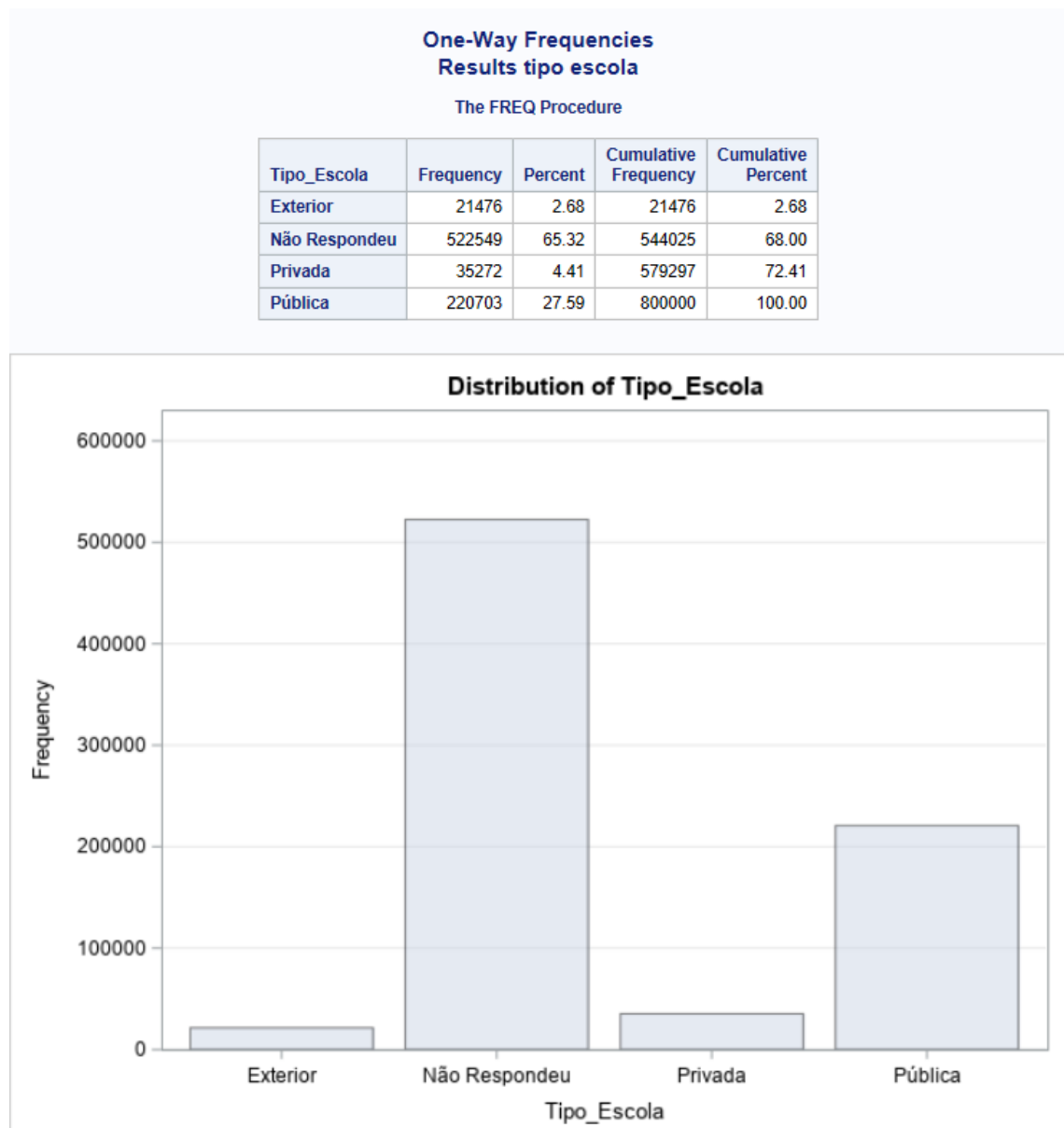
Figura 16 – Distribuição de frequência por cor e raça dos ENEMs 2017-2019 limpo



Fonte: SaS

Aqui eu realizei a realocação dos valores missings para o valor na variável mais constantes. Bom se comparamos com a análise anterior vemos que o número de pardos cresceu por ela ser o valor mais constante na variável de cor e raça. Bom após analisar creio que teria sido melhor alocar o missings para os valores não declarado.

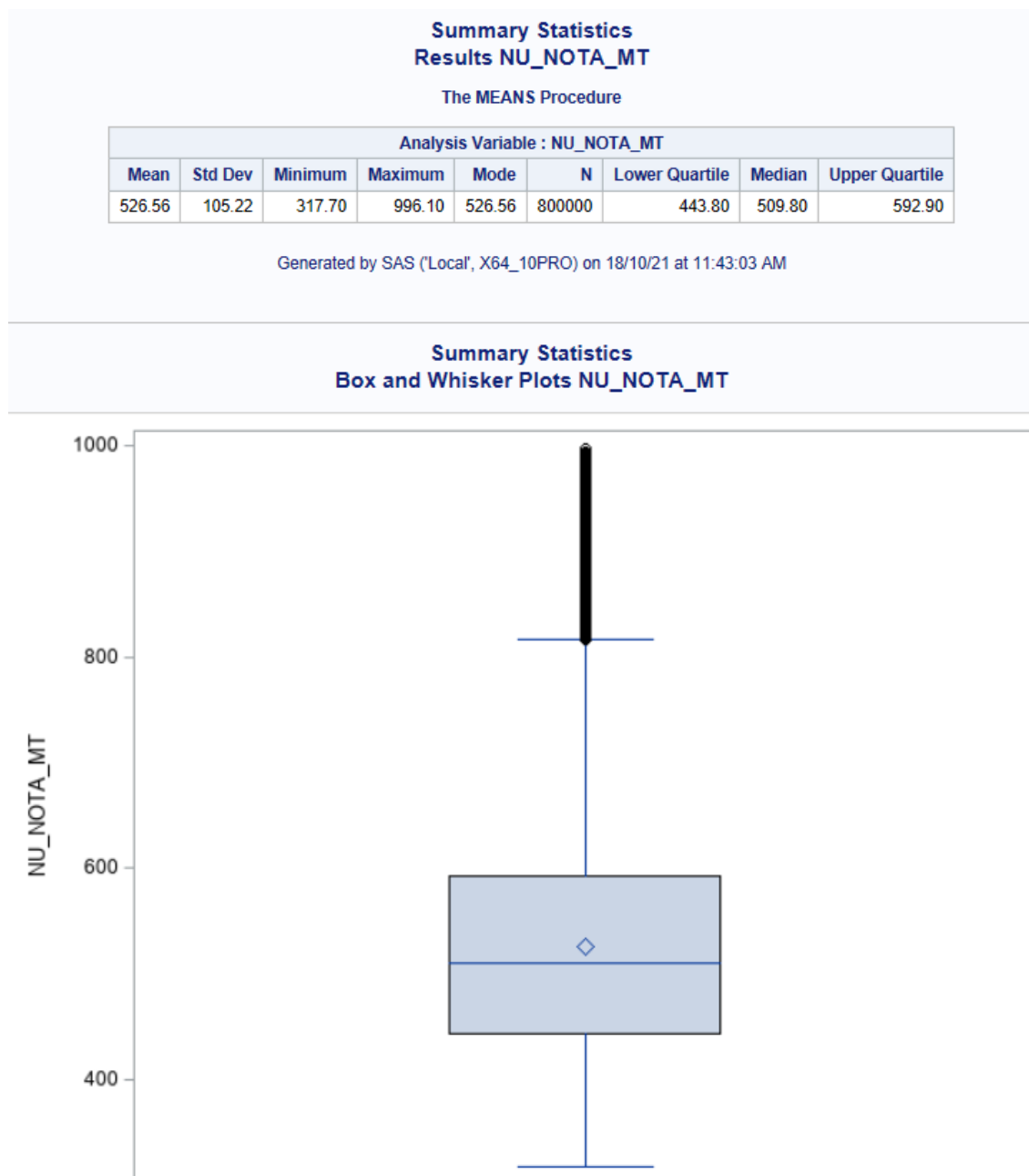
Figura 17 – Distribuição de frequência por tipo de escola dos ENEMs 2017-2019 limpo



Fonte: SaS

Aqui também realizei a realocação dos valores missings para o valor na variável mais constantes. O mesmo aconteceu a varivel mais constantes nesse caso são os alunos que preferiram não responder. Sendo assim o resultado continua basicamente o mesmo da análise anterior.

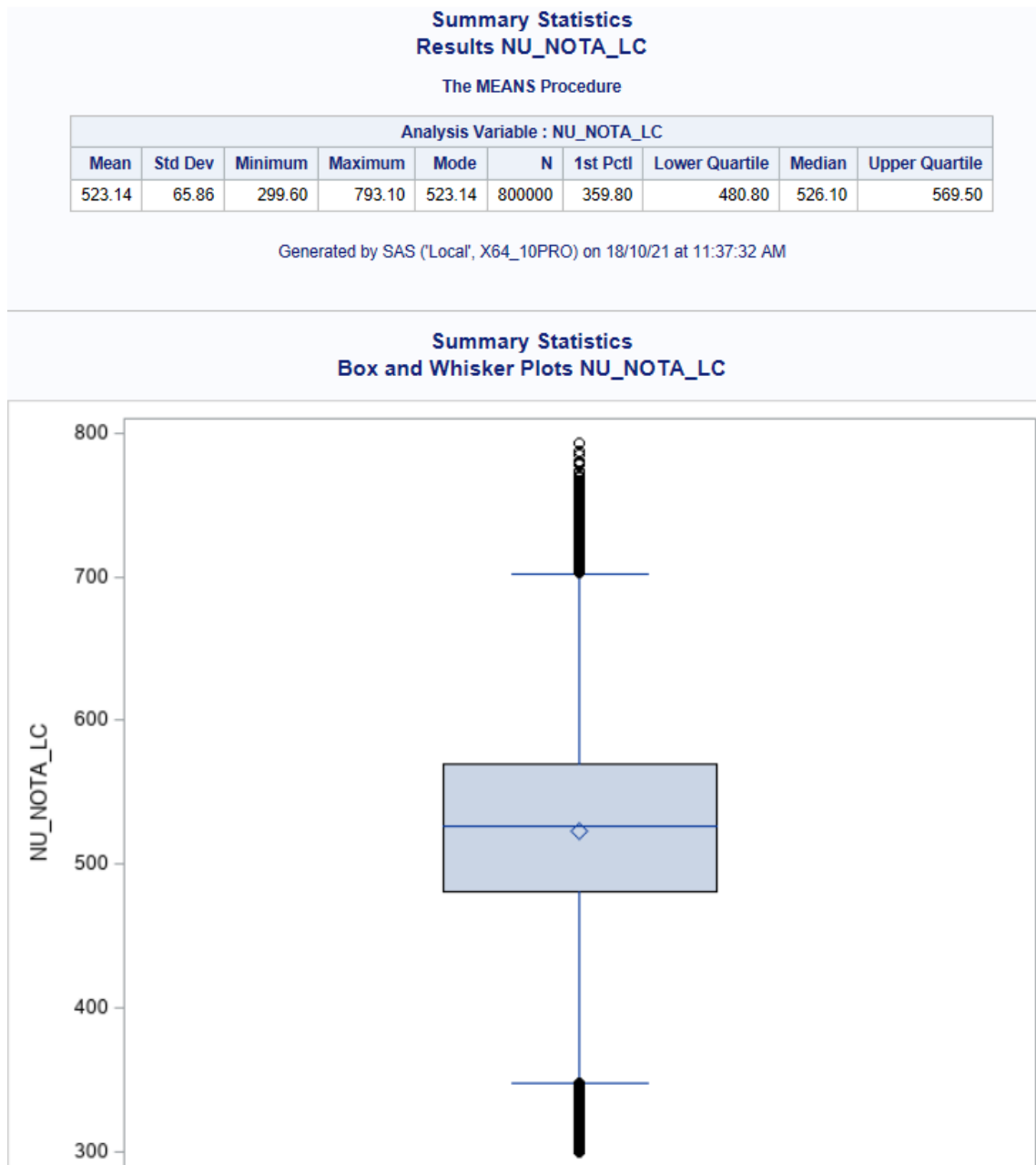
Figura 18 – Analise dados NU-NOTA-MT (matemática e suas tecnologias)) dos ENEMs 2017-2019 limpo



Fonte: SaS

Na variável NU NOTA MT eu dei preferência por substituir os missings pela nota média. Assim a análise manteve o padrão da anterior, com sua média em 526,56 a moda passou a ser a mesma que a da média, e a média manteve a mesma. e ainda contendo muitos a outliers.

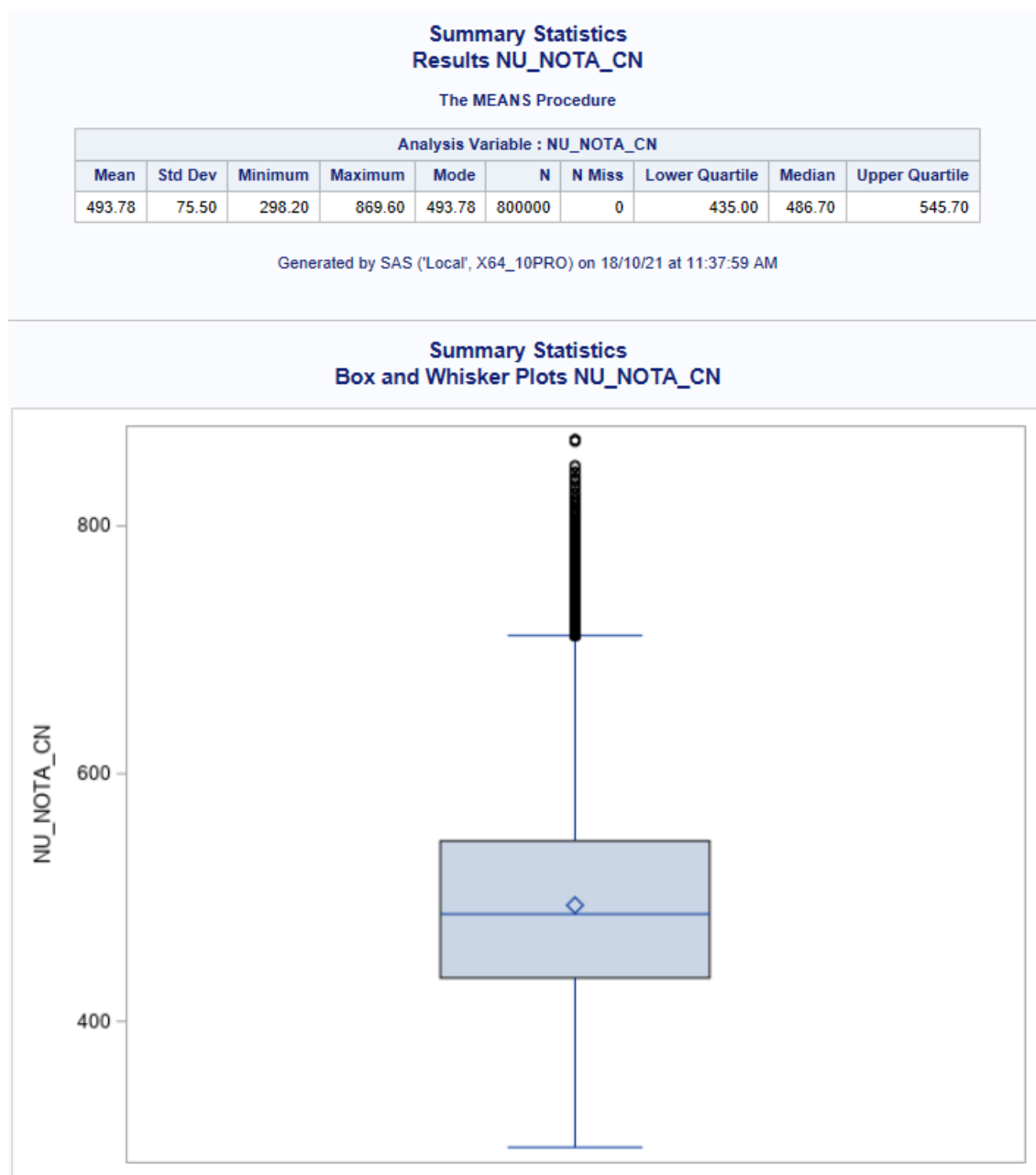
Figura 19 – Análise dados NU-NOTA-LC (linguagens, códigos e suas tecnologias) dos ENEMs 2017-2019 limpo



Fonte: SaS

Na variável NU NOTA LC eu dei preferencia por substituir os missings pela nota média. E analisei que a nossa moda mudou se comparada a análise anterior nossa média teve uma leve diminuição. Porem os demais valores se matem.

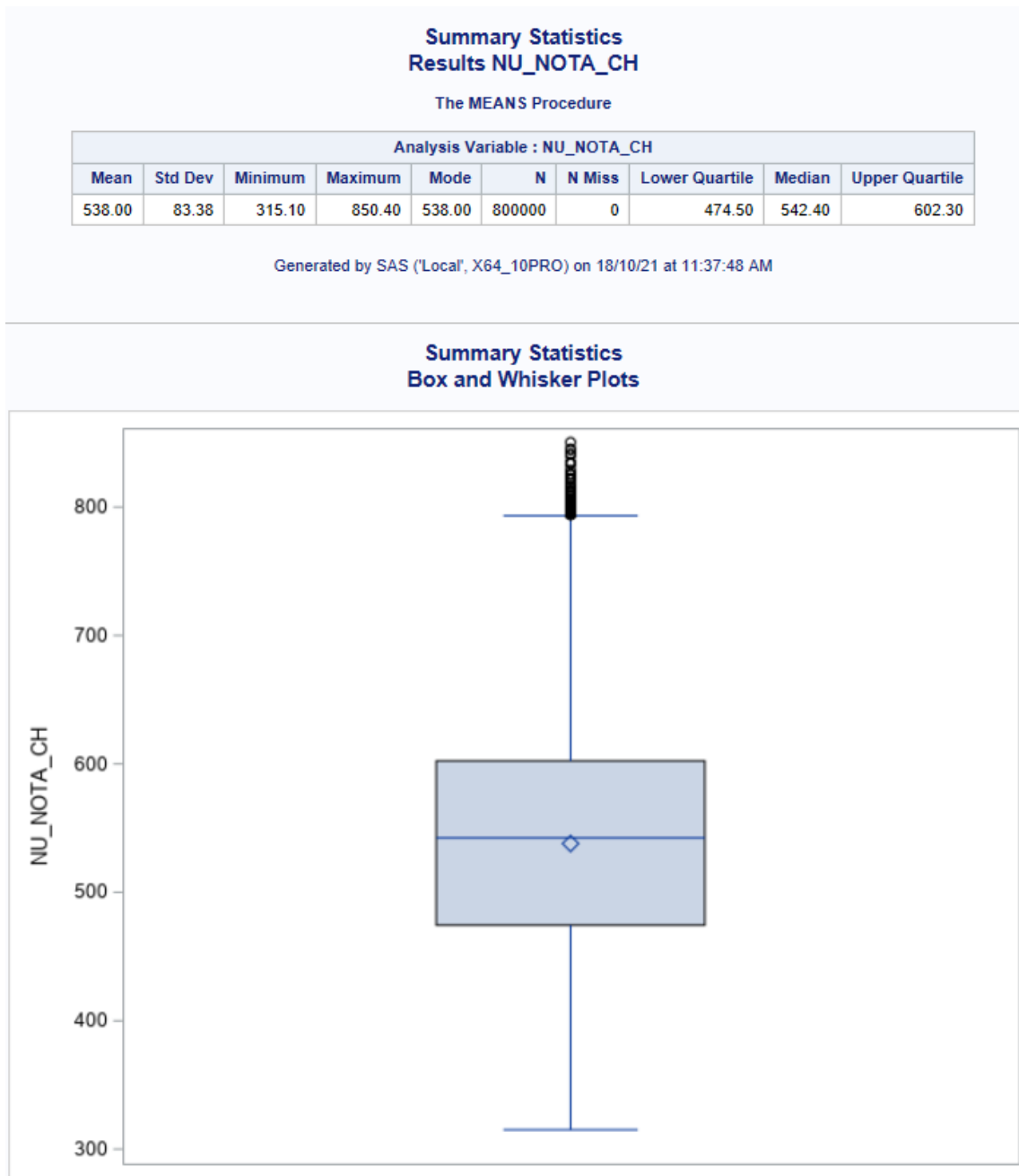
Figura 20 – Analise dados NU-NOTA-CN (ciências da natureza e suas tecnologias) dos ENEMs 2017-2019 limpo



Fonte: SaS

Após alterar os missings pela média. Nossa moda foi altera para 493,78 nossa média se manteve e nossa mediana aumentou um pouco e foi para 486,70.

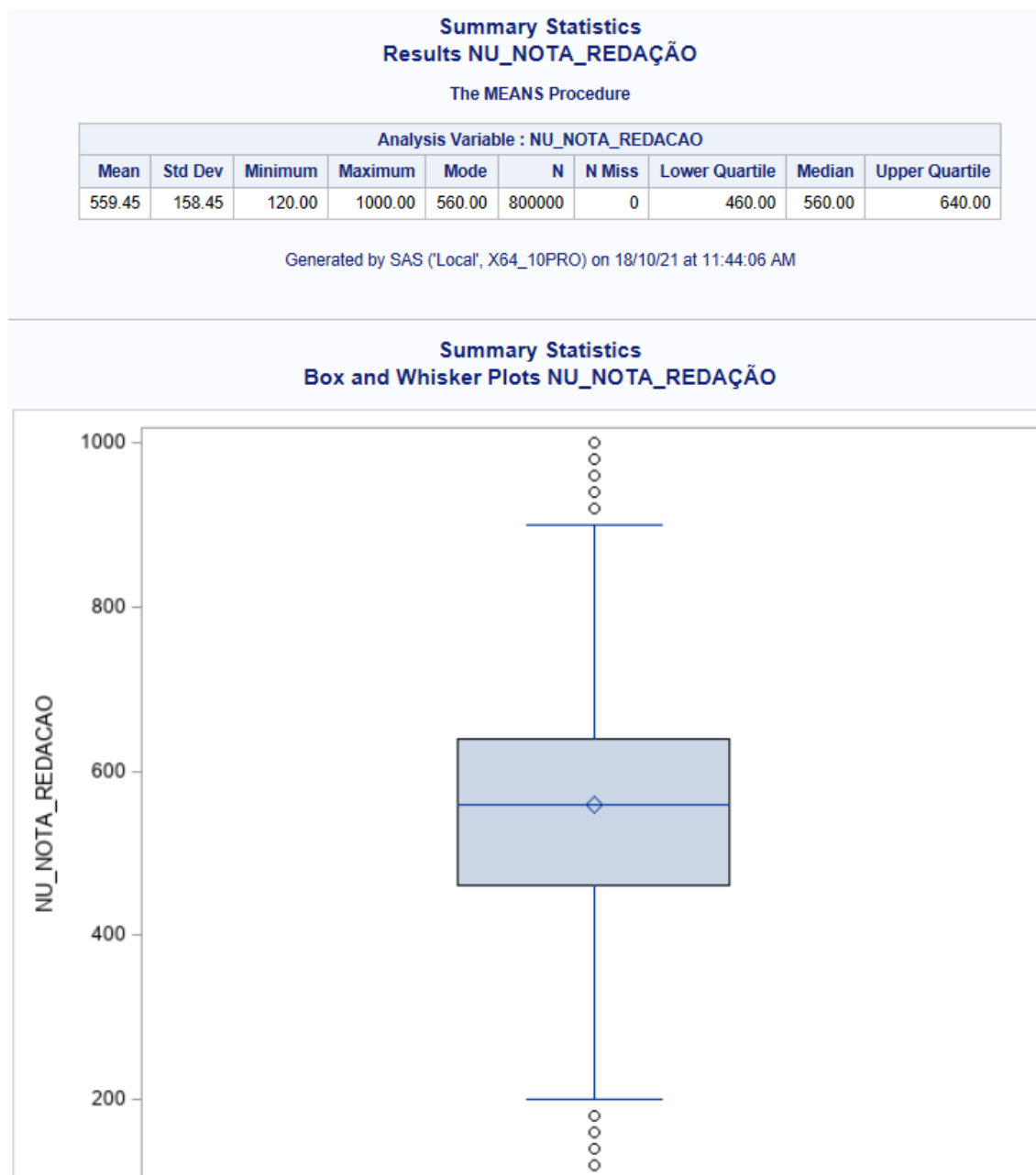
Figura 21 – Analise dados NU-NOTA-CH (ciências da natureza e suas tecnologias) dos ENEMs 2017-2019 limpo



Fonte: SaS

Após alterar os missings pela média. Nossa moda foi alterada para 538 nossa média se manteve e nossa mediana se diminuiu e foi para 542,40.

Figura 22 – Analise dados da Redação dos ENEMs 2017-2019 limpo

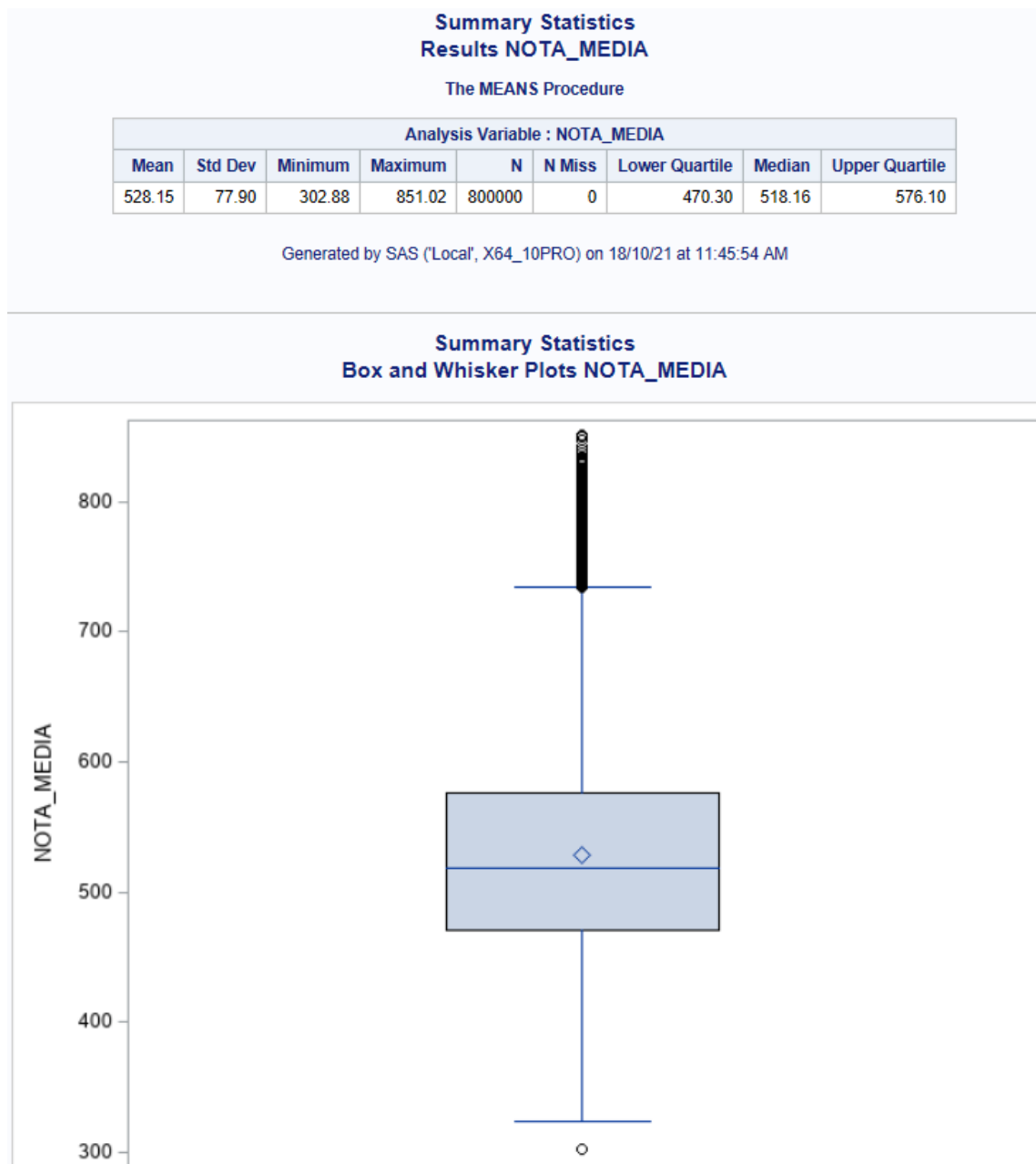


Fonte: SaS

No caso da notas de redação modifiquei os missings pela mediana. porem não obtivemos muita diferença das análises anteriores. a única coisa que modificou foi a moda todos os valores restantes se matem igual a análise anterior.



Figura 23 – Analise dados da Nota Media dos ENEMs 2017-2019 limpo



Fonte: SaS

No caso da nota media modifiquei os missings pela mediana. porem não obtivemos muita diferenca das analises anteriores. a única coisa que modificou foi a moda todos os valores restantes se matem igual a analise anterior.

## 5 Resultados

Para chegar a essa conclusão, o jornal analisou informações de 3,87 milhões de candidatos do Enem 2011 que responderam ao questionário socioeconômico no ato da inscrição e que fizeram a prova de redação naquele ano. Esses dados são os mais recentes disponíveis em relação ao exame que se tornou a principal porta de entrada para o ensino superior no Brasil. Neste fim de semana, acontece a próxima edição do exame, que tem 7,1 milhões de inscritos.

Ao comparar renda familiar e desempenho na redação, prova que tem o maior peso no exame, percebeu-se um aumento contínuo da nota junto com a situação financeira e a escolaridade dos pais. Enquanto a nota média entre aqueles com renda de até um salário mínimo foi de 460 pontos, o grupo com renda acima de 15 salários chegou a 642 pontos. Diferença de 40%.

Na comparação entre as unidades da federação, essa disparidade é mais ampla no Piauí, onde a diferença entre a menor e a maior médias é de 50%. Santa Catarina e Amapá são os que apresentam menor discrepância: 27%.

- O Enem reproduz brutalmente as nossas desigualdades, e outros estudos que consideraram outras variáveis sociais chegaram às mesmas conclusões. O pobre não é burro, mas ele participa de um concurso com jovens que têm acesso a experiências educacionais muito mais ricas. Nesse sentido, a sociedade não se dá conta de que vivemos uma situação de cartas marcadas, que reproduz nosso padrão socioeconômico. A solução para isso não é fazer uma avaliação mais leniente com quem vem de família com baixa renda, mas melhorar a escola, pôr a questão do aprendizado no centro da atenção - diz o professor Francisco Soares, do grupo de avaliação e medidas socioeducativas da Universidade Federal de Minas Gerais (UFMG).

Quando as notas são avaliadas segundo a rede de ensino, as diferenças persistem. Em 2011, cerca de 1,4 milhão de alunos que fizeram a redação do Enem estavam no ensino médio. A nota média entre os candidatos de escolas estaduais (78% desse universo) foi de 486 pontos. A rede municipal alcançou 498. Já a média entre os colégios privados chegou a 612, pouco abaixo do ensino federal, com 623. A porcentagem de alunos de escolas federais no Enem, porém, está em 1,8%.

O baixo desempenho nas redes estadual e municipal é explicado também pela renda das famílias. Cerca de 80% dos estudantes das escolas estaduais e municipais que fizeram o Enem 2011 afirmaram ter renda de até dois salários mínimos. Na rede federal, esse percentual cai para 55%, e na privada é de apenas 30%.

Também há muita discrepância quando se comparam notas entre alunos de baixa e

alta renda dentro da mesma rede de ensino. Nas escolas municipais, a média entre alunos com renda de até um salário é de 433 pontos, enquanto entre os de renda de mais de 15 salários é de 553 (diferença de 28%). Na rede estadual, as notas vão de 443 a 562 (27%). Na federal, de 550 a 689 (25%). E na particular, de 539 a 652 (21%).

#### Desistência maior na rede estadual

O coordenador de projetos da Fundação Lemann, Ernesto Martins Faria, explica que jovens de famílias com poucos recursos vivem em condições desfavoráveis que afetam o aprendizado, como, por exemplo, espaço inadequado em casa para se dedicar aos estudos, baixo acesso a livros e até mesmo um vocabulário pouco diversificado utilizado pelos pais. Para Faria, a relação entre desempenho na redação e renda familiar, contudo, deve ser vista com cautela quando se trata de alunos da rede federal.

- O patamar das notas dos alunos de baixa renda é bem mais baixo nas redes públicas estadual e municipal. Esses alunos têm um background extra escolar mais desfavorável. Alunos da rede particular provavelmente têm pais mais engajados, e o gasto com educação privada, apesar da baixa renda familiar, ilustra isso. Já entre os alunos da rede federal, alguns devem ter passado por processos seletivos que são feitos em certas escolas. Para alunos que passam por processos seletivos a renda não é uma boa ilustração do background ou das oportunidades educacionais - afirma Faria.

Com poucos recursos e enfrentando situações por vezes desfavoráveis, boa parte dos alunos da rede pública desiste no meio do concurso. Pelos dados analisados pelo jornal, quanto menor a renda familiar, menor é a probabilidade de os alunos participarem da redação, aplicada no segundo dia de provas. A desistência entre os alunos na rede municipal chegou a 24%, seguida da estadual, com 19,7%. Nas escolas federais, a desistência foi de apenas 6%, patamar muito próximo da rede privada (5%).

- Esses dados revelam algo que merece uma maior atenção do poder público. O Enem gera um incentivo à participação dos alunos, porque eles querem o ensino superior. Quem faz a redação está envolvido com essa perspectiva. A desistência maior entre alunos da rede pública indica, a meu ver, uma falta de perspectiva dos alunos. Eles pensam que não poderão ser aprovados ou, caso sejam, pensam em como poderão se manter financeiramente no ensino superior. Isso tudo tem a ver com as políticas que podem ser criadas para permitir que esses jovens se dediquem aos estudos ou possam se manter durante a faculdade - observa Faria.

#### Falta de professores prejudica estudante

Aluna do Colégio Estadual João Alfredo, Rayane Florêncio, de 17 anos, vai fazer o Enem este ano. A moradora do bairro de Jacaré, na Zona Norte, ficou sem professor de química durante meses, e está sem professor de geografia devido à greve de profissionais da categoria nas redes estadual e municipal do Rio. Para correr atrás, entrou num cursinho

pré-vestibular comunitário.

- Gostaria de estar num colégio particular para não ter esses problemas, mas não teria como pagar - conta a aluna, filha de um caminhoneiro e uma dona de casa, cujo sonho é estudar Letras na UFF. - Tenho um pouco de medo. Sei que a prova vai cobrar coisas que não aprendi.

Até agosto, Rayane trabalhava numa pizzaria à noite, para ter seu próprio dinheiro, mas isso atrapalhava demais sua preparação.

- Tinha a escola pela manhã e, depois, o cursinho das 13h às 18h. Saía correndo para o trabalho, onde ficava até meia-noite. Era cansativo. Abri mão do trabalho para focar no Enem - desabafa.

No estudo feito pelo GLOBO, também foram comparadas as médias por estados. Como a participação no Enem é voluntária, os dados servem apenas para ilustrar o desempenho dos alunos que fizeram as provas, e não para explicar disparidades socioeconômicas nos estados como um todo. No Piauí, onde a discrepância entre as notas de alunos com baixa e alta renda chega a 50%, os estudantes de famílias que vivem com até um salário mínimo tiveram média da redação de 450 pontos, enquanto os com renda acima de 15 salários alcançaram 676. Em Mato Grosso do Sul, a disparidade foi de 46%. As menores diferenças foram no Amapá e em Santa Catarina (ambos com 27%), seguido de São Paulo (33%). O impacto da situação socioeconômica no rendimento dos alunos foi analisado pelo doutor em Educação pela Universidade de São Paulo (USP) Rodrigo Travitzki. Na pesquisa, feita para defesa da sua tese este ano, ele comparou a média das escolas no Enem e concluiu que mais de 80% das variações são explicadas por fatores que não podem ser controlados pelas escolas, como renda e escolaridade familiar.

- Esse dado revela que a educação de um país não pode ser muito melhor que o país. As escolas sozinhas não resolvem. Precisamos melhorar as escolas, mas precisamos também reduzir nossas desigualdades. Minha tese procurou discutir esse tema, porque não adianta focar no ranking das melhores escolas do Enem. Acaba virando marketing das escolas, quando sabemos que elas, sozinhas, pouco podem fazer. Presidente do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), órgão do MEC responsável pela aplicação do Enem, Luiz Claudio Costa reconhece que o exame, por si só, não vai melhorar os rumos da educação. Ele sabe das discrepâncias entre as notas de alunos de baixa e alta renda familiar.

- Educação é maratona. É preciso transformar toda uma realidade. Nossa história é de exclusão, e o Brasil vem mudando isso, colocando jovens nas escolas. O Enem, assim como as cotas, é uma ferramenta no processo. Antigamente, dois ou três vestibulares influenciavam muito no ensino, e só dialogavam com escolas particulares. O Enem promove diálogo com a escola pública. Mas não é uma mudança rápida.

## 6 Conclusões

Marco avaliativo final de um processo de escolarização ainda notadamente discriminatório, as notas do Enem expressam o funil excludente que marca a educação brasileira.

apesar de nos últimos anos ser observada queda do abandono no Ensino Médio tanto entre estudantes brancos quanto negros, a distância nos últimos anos pouco se alterou (permanece estável entre 2,5 e 3 pontos percentuais). Além disso, o último dado, de 2018, mostra que a queda segue entre os estudantes brancos, ao contrário do índice entre os estudantes negros (subiu de 7,7% em 2017 para 7,8% em 2018), com possibilidade de agravamento diante da pandemia do Coronavírus, conforme conteúdo da seção Em Debate, que também abordou o assunto.

Além da evasão, diversos outros índices reafirmam essa desigualdade estrutural, expressão do racismo presente na sociedade brasileira desde sua constituição até a atualidade. Se observados os dados de 2018, bem como sua trajetória histórica, em todas as áreas de conhecimento, o desempenho dos estudantes brancos é maior do que dos estudantes negros, chegando a 65,2 pontos de diferença na Redação.

Se observamos as diferenças entre estudantes mulheres e homens também é possível refletir sobre desigualdades de gênero e o sexismo na educação. Apesar das edições do Enem, contarem, desde 2011, com taxa de participação repetidamente maior entre as estudantes do sexo feminino, é possível também verificar a persistência de diferença de desempenho entre elas e eles.

Se observamos as diferenças entre estudantes mulheres e homens também é possível refletir sobre desigualdades de gênero e o sexismo na educação. Apesar das edições do Enem, contarem, desde 2011, com taxa de participação repetidamente maior entre as estudantes do sexo feminino, é possível também verificar a persistência de diferença de desempenho entre elas e eles.

Segundo levantamento do jornal Folha de S.Paulo, realizado com base no perfil de participantes do Enem 2018, 34% dos estudantes da rede pública que prestaram a prova naquele ano não tinham acesso à internet. Entre os alunos da rede privada, a parcela desprovida deste acesso era de apenas 3,7%. O estudo usou os dados mais recentes disponíveis e constatou, ainda, que a desigualdade aumenta quando se observam recortes de renda e raça.

Recente publicação do Jornal Nexo demonstrou como as desigualdades raciais e de gênero foram acentuadas pela pandemia. Ao analisar os índices de infecção e de mortalidade entre pessoas negras e brancas, através de dados disponibilizados pelo Ministério da Saúde,

constatou-se que enquanto as pessoas negras representam cerca de 23,1% das pessoas internadas com Síndrome Respiratória Aguda Grave, elas somam mais de 30% dos óbitos por Covid-19. Além disso, dados da PNAD Covid-19, demonstram como a crise econômica afetou de maneira desproporcional pessoas negras e mulheres, sobretudo no que refere à diminuição da carga horária de trabalho e da renda. O impacto desproporcional da pandemia na vida de pessoas negras e mulheres tem consequências diretas na garantia de condições socioeconômicas mínimas para que essas juventudes obtenham resultados satisfatórios no próximo Exame Nacional do Ensino Médio (Enem).

A despeito do acesso à educação apresentar índices que se aproximam da universalidade, o direito à aprendizagem ainda esbarra nas desigualdades que estruturam a nossa sociedade. Dessa forma, a diferença na média dos resultados de estudantes negros e brancos, homens e mulheres, pessoas cis e trans, revela a necessidade de aprimorar as políticas públicas educacionais de modo a promover a equidade na Educação Básica.

Enquanto política pública educacional, o Enem cumpre o papel de cada vez mais fomentar e possibilitar a inserção de pessoas negras, de baixa renda e de mulheres no Ensino Superior. A despeito das críticas, necessárias a contínua melhoria e ampliação dessa política pública, o Exame Nacional do Ensino Médio possibilitou, por exemplo, o crescimento expressivo do percentual de pessoas negras no ensino superior: de 5,5% em 2005 para 12,8% em 2015 (IB

# Anexos

# ANEXO A – Códigos dos Programas

*#A seguir estão descritos todos os códigos de programas desenvolvidos*

*#O programa a seguir foi utilizado para realizar a limpeza dos dados missings*

*#Importando base de dados*

```
import pandas as pd
import numpy as np
from sklearn.impute import SimpleImputer
```

*#Realizando leitura da base de dados*

```
df = pd.read_excel("Tassio_'_ENEM_DADOS_Anomalias.xlsx")
```

*#visualizando a tabela*

```
display(df)
```

*#Verificando os dados nulos*

```
df.isnull().sum().sort_values(ascending = False)
```

*#Removendo missings das variaveis quantitativas*

*#Variaveis quantitativas*

```
df['NU_NOTA_MT'] = quantitativas.fit_transform(df[['NU_NOTA_MT']])
df['NU_NOTA_LC'] = quantitativas.fit_transform(df[['NU_NOTA_LC']])
df['NU_NOTA_CN'] = quantitativas.fit_transform(df[['NU_NOTA_CN']])
df['NU_NOTA_CH'] = quantitativas.fit_transform(df[['NU_NOTA_CH']])
df['NU_NOTA_REDACAO'] = quantitativas_2.fit_transform(df[['NU_NOTA_REDACAO']])
df['NOTA_MEDIA'] = quantitativas_2.fit_transform(df[['NOTA_MEDIA']])
df['NU_IDADE'] = quantitativas.fit_transform(df[['NU_IDADE']])
```

*#verificando missings novamente*

```
df.isnull().sum().sort_values(ascending = False)
```

*#Removendo missings das variaveis qualitativas*

*#Variaveis qualitativas*

```
df['Sexo'] = qualitativas.fit_transform(df[['Sexo']])
df['COR_RACA'] = qualitativas.fit_transform(df[['COR_RACA']])
df['Tipo_Escola'] = qualitativas.fit_transform(df[['Tipo_Escola']])
```



```
#verificando missings novamente
```

```
df.isnull().sum().sort_values(ascending = False)
```

```
#Salvando dataset em formato xlsx(excel)
```

```
df.to_excel('Tassio_ENEM_DADOS_testes.xlsx', encoding = "utf-8", index=False)
```

## ANEXO B – Conclusão

Agora sim, Eu estou pronto para lidar com campos missings / nulos quando receber uma base incompleta. Veja que a forma de lidar, preenchendo com algum valor ou simplesmente excluindo a linha, varia de acordo com cada situação. Muitas vezes excluir uma linha ou outra sequer vai atrapalhar no todo e pode ser mais rápido e fácil. Em outros casos, para não perder muita informação, você pode optar por preencher com a média, mediana ou até mesmo com um valor fixo. Se estiver trabalhando com variáveis categóricas, criar uma nova categoria para o missing é uma opção também.