

Exploratory Data Analysis ENEM 2019

1. Introduction

The National High School Exam, also known as ENEM, was created in 1998 aiming to measure the performance of the students after their graduation, becoming the principal way to be accepted in universities in Brazil. Nowadays, there are more than 500 universities that use ENEM as part of their selective process.

This report will present an exploratory data analysis of the ENEM 2019 microdata that is available in the website of the Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), which is the national institute of education and research.

One of the main objectives of educational assessment is to ensure the quality of education, which can be monitored by evaluating the profile of candidates, how this reflects on their grades and whether these factors can be related to the economic and social structure of Brazil.

The project aims to study the correlation between the student's grades and their household income, age, type of school, access to the internet, geographic state and others.

2. General Data of ENEM 2019

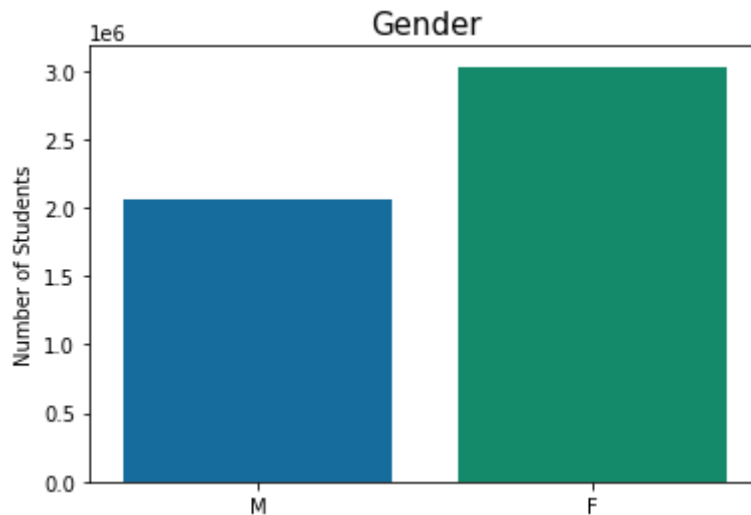
Firstly, we filtered the data to work with those more relevant for this specific analysis, which are:

- geographic state
- gender
- ethnicity
- type of school (public or private)
- access to the internet
- mother's schooling
- household income
- math grades
- essay grades
- human science grades
- natural science grades
- languages, codes and its technologies grade
- presence in math exam
- presence in human science exam

Before moving on to more complex and detailed analyses, it is necessary to know the profile of the exam registrants and have a general idea about the data of the 2019 exam.

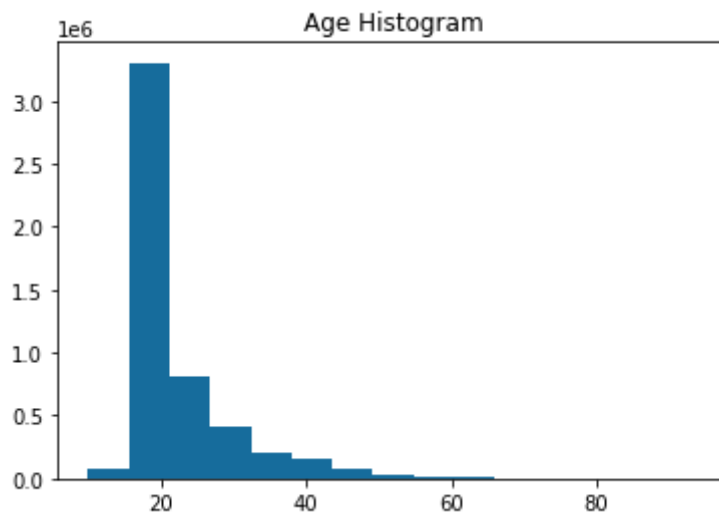
The ENEM 2019 had 5.095.270 subscriptions, but 22.92% of the students did not attend on the first day of exam and 27.17% did not attend on the second day, obtaining the lowest abstention rate in the history of the exam.

When it comes to the profile of the applicants, 40.5% are men and 59.5% are women, as we can see in the graph below:

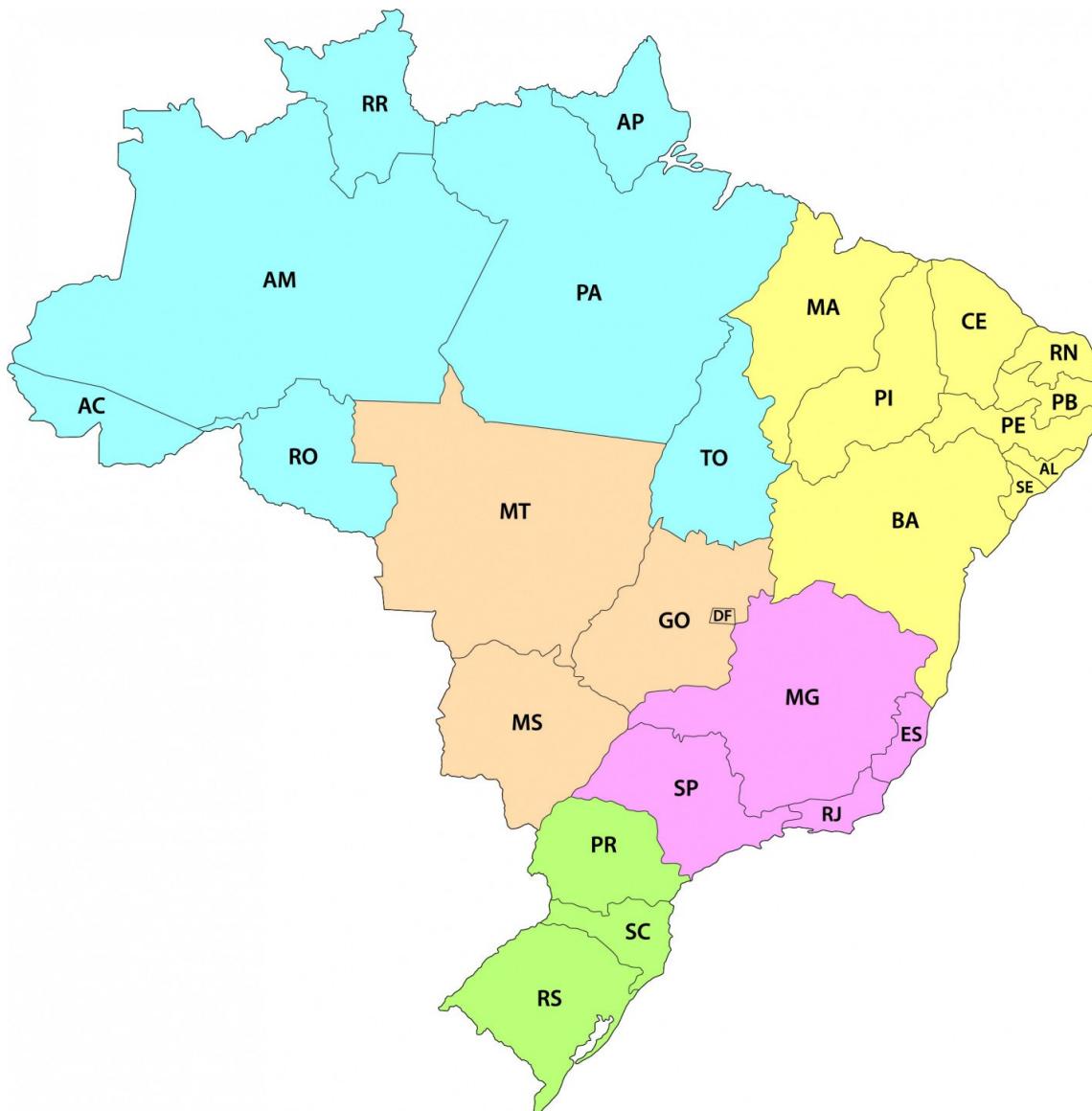
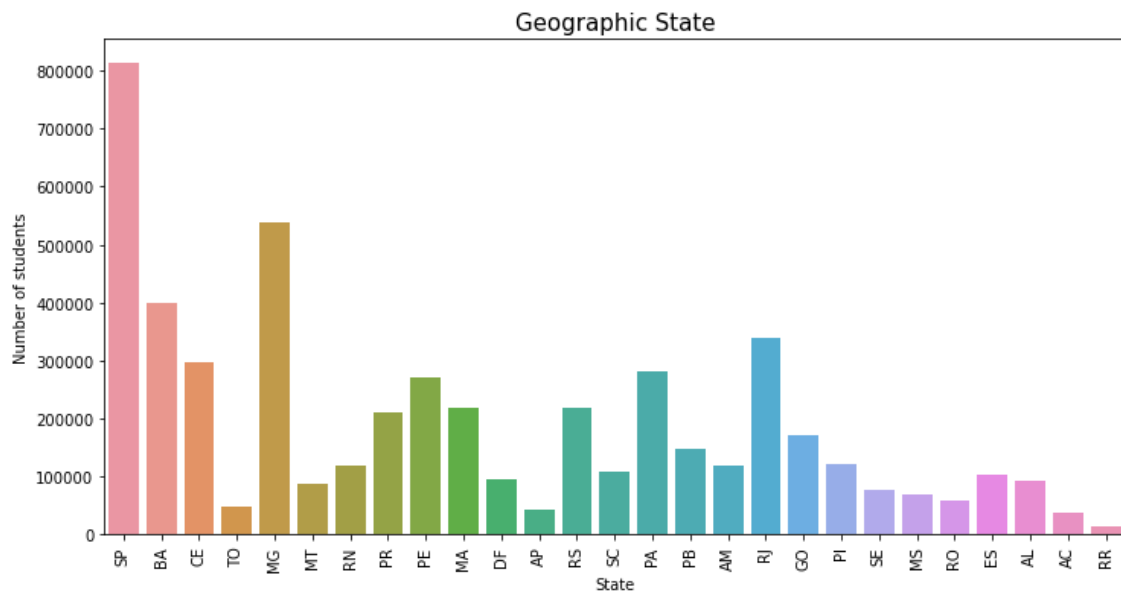


According to INEP (the national institute of education and research), the number of female applicants is the highest since the first edition of the exam in 1998. However, this does not reflect the Brazilian society because according to IBGE (Brazilian Institute of Geography and Statistics) the female population is only 4.7% higher than men. Moreover, if we consider only the population between 15 and 19 years old, men are the majority of the country.

Great part of the candidates are under 20 years old, which is possible to visualize in the histogram below:

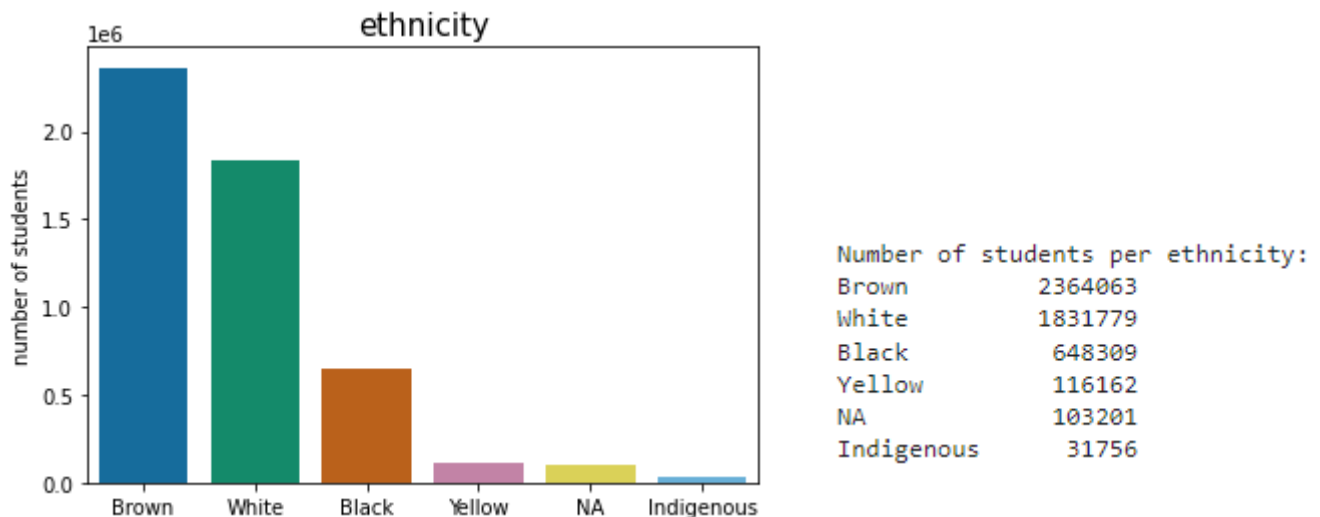


Next, we can see the analysis of the number of candidates per geographic state. São Paulo (SP) is the state with the higher number of applicants, representing 15.97% and it is followed by Minas Gerais (MG) with 10.57%, both in the southeast of Brazil. The following histogram shows the number of candidates in all the states of Brazil:



Nowadays, according to the PNAD (National Household Sample Survey) in 2019 42.7% of the Brazilians declared themselves as white, 46.8% as brown, 9.4% as black and 1.1% as yellow or indigenous. It is possible to see that these values reflect in the candidates of the ENEM 2019.

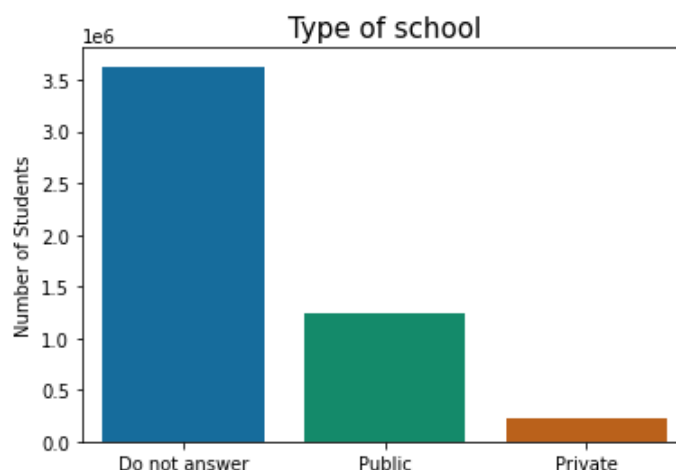
The data analysis about ethnicity of the applicants show that brown people were those with more representativeness, followed by white people. However, indigenous and yellow people were the lower number of applicants, as it is shown in the bar graph below:



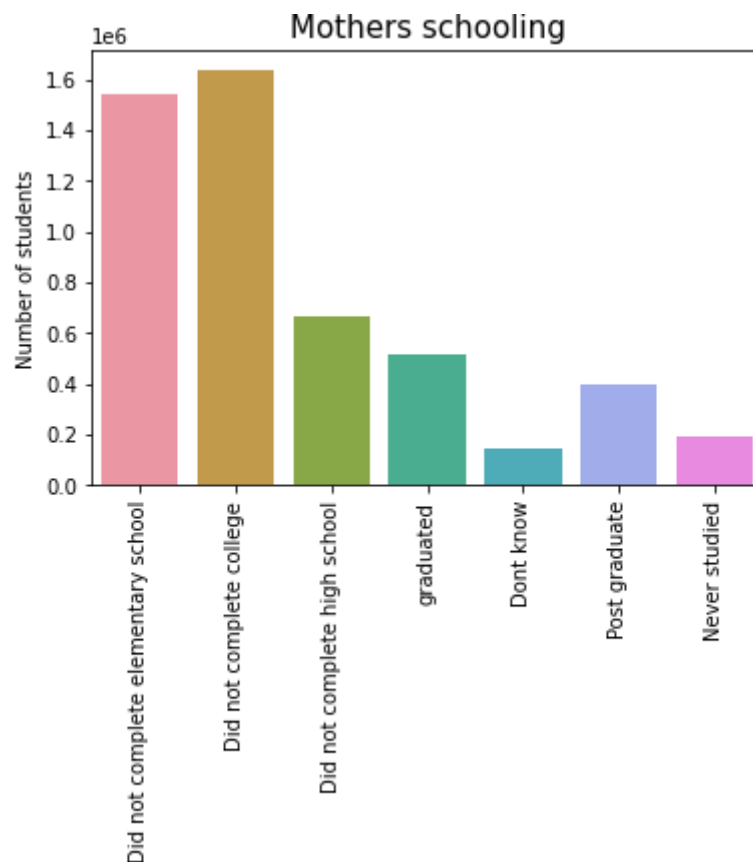
The IBGE (Brazilian Institute of Geography and Statistics) manual defines the meaning attributed to the term brown as people with a mixture of skin colors, whether this mixture is mulatto (descendants of whites and blacks), cabocla (descendants of whites and Amerindians), cafuza (descendants of blacks and indigenous people).

Also according to IBGE (Brazilian Institute of Geography and Statistics) in 2019, almost 40 million Brazilians do not have access to the internet, totaling 21.7% of the population. Regarding ENEM 2019 applicants, 77.62% have access to the internet and 22.38% do not, reflecting Brazilian society.

When it comes to the type of school, this was one of the lowest answer rates, approximately 71.23% prefer not to answer, 24.47% of the students are from public schools and only 4.29% are from private schools. The following bar graph shows the number of students per type of school:

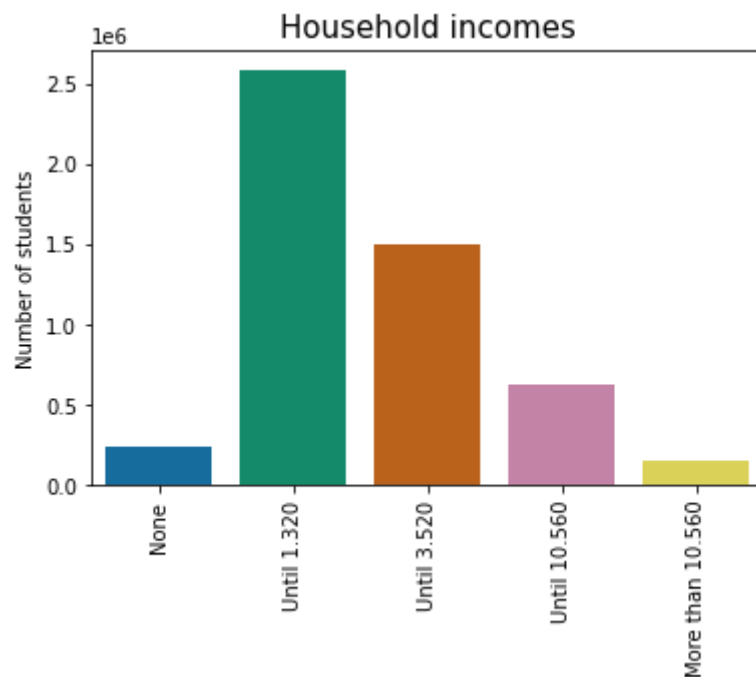


Furthermore, it is possible to see that 32.06% of the students' mothers do not complete college and 30.24% do not complete elementary school. Only 10.08% graduated and this number is even lower when it comes to postgraduate, with only 7.84%. The graph below show this correlation:

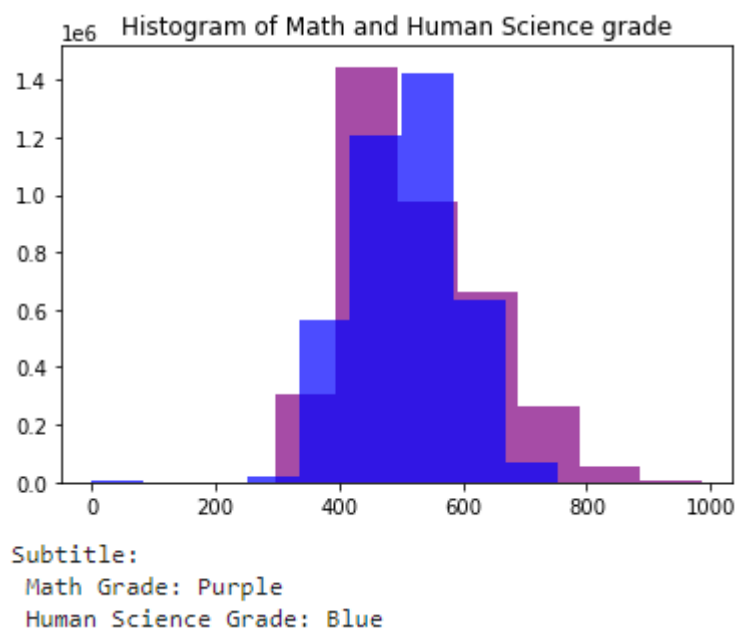


Number of students according to their mother's schooling:	
Did not complete college	1633692
Did not complete elementary school	1541024
Did not complete high school	667415
graduated	513711
Post graduate	399470
Never studied	194625
Dont know	145333

According to IBGE (Brazilian Institute of Geography and Statistics), the per capita household income in Brazil was R\$1439 in 2019. This reflects consistency with the data from the ENEM 2019 candidates, as the vast majority of students have a family income of up to R\$1,320.00 per month and only a minority responded who have no income or who have income greater than R\$10,560.00. This relationship can be seen in the following chart household income *versus* number of students:



Lastly, comparing the performance of the students, we can notice that the grades are worse in the math exam. Because the exam has a higher level of complexity, we will use the math grades in the future analysis.



3. Correlation of Variables

In this topic, the correlation between two variables will be addressed. This way it is possible to carry out analyzes and reach more assertive conclusions about whether or not certain conditions influence the performance of candidates.

Despite several factors, in this report only three aspects will be addressed: ethnicity, household income and mother's schooling. These were considered the most relevant among the topics studied.

a. Ethnicity

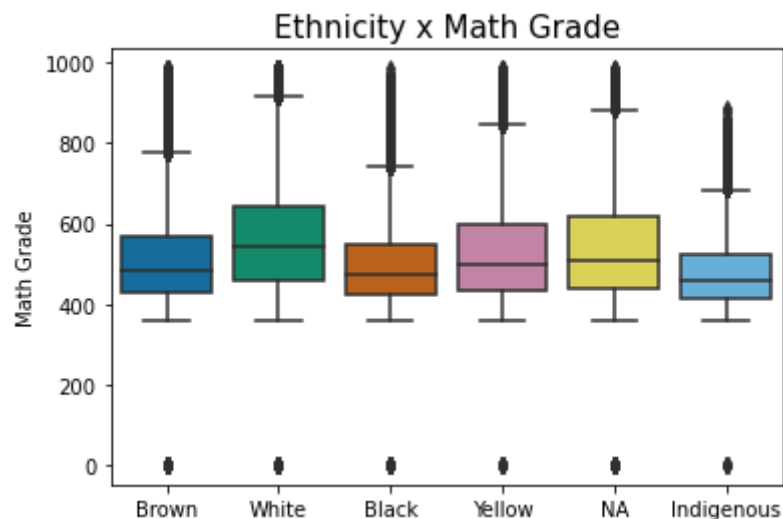
Firstly, it will be addressed if the ethnicity of the candidates could affect their grades. However, in order to understand the approaches that will be used, we need a historical contextualization of Brazil.

The abolition of slavery officially took place in 1888 with the Áurea law, but no social or economic policy was applied in order to insert those people in the labor market with dignified conditions, allowing the social marginalization of black people.

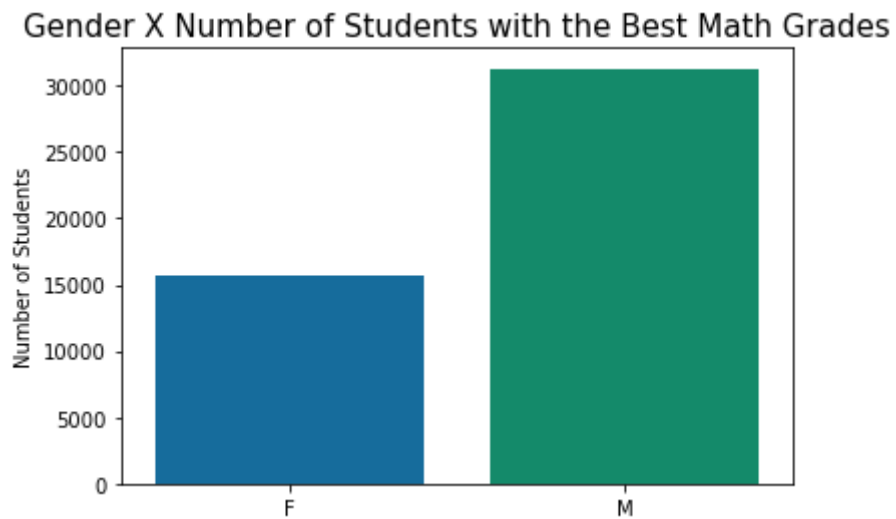
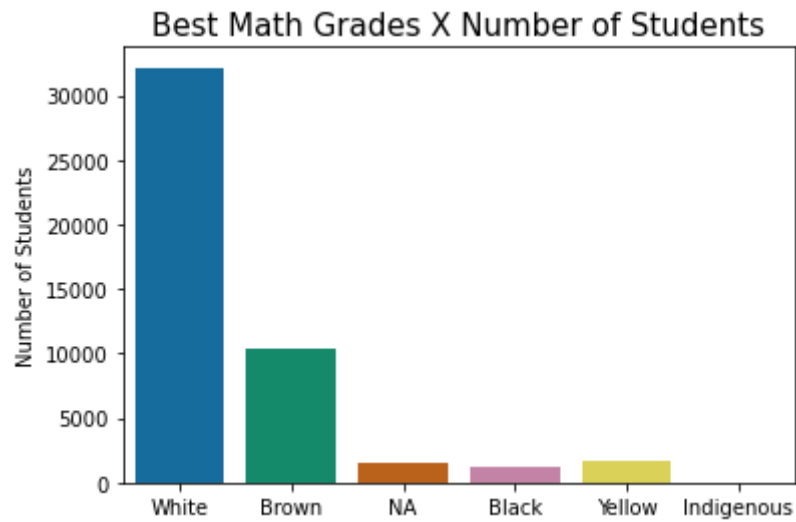
For this reason, a lot of unemployed ex-slaves went to live in collective housing in the center of the city, also known as cortiços, because they were unable to legally buy or rent a house. Next, with the demolition of the cortiços, this population settled in devalued areas such as hills, caves and swamps, giving rise to the favelas.

Slavery began in Brazil in the mid-1530s, that is, it took place for about 358 years and was abolished only 133 years ago. For this reason, it is possible to see that the abolition of slavery is recent and there may still have many vestiges of that time until today.

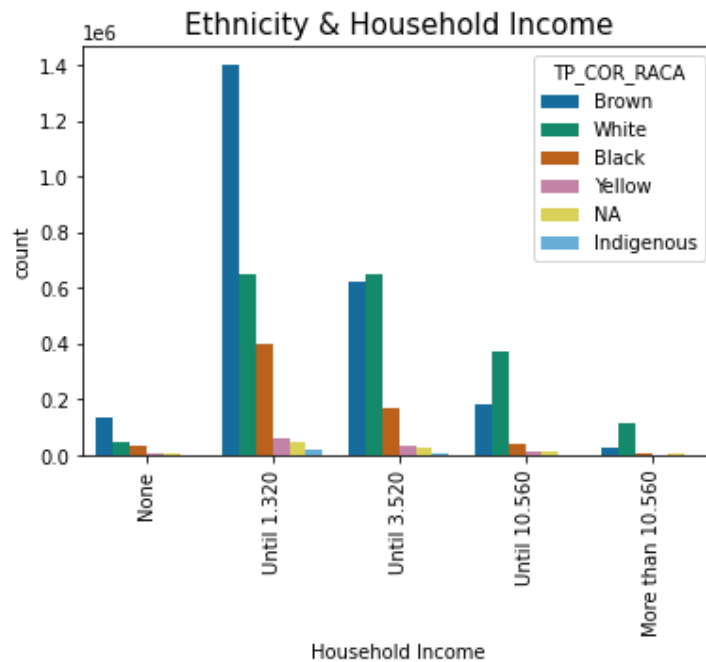
These vestiges can be found comparing the math grades. Black and indigenous people have a considerable difference from other ethnicities, having their max and medians smaller. The following graph shows this correlation:



Filtering the best math grades, that is, those above 800 points (corresponding to 80% of the test) the difference shown above is even bigger. It is possible to see that the majority of candidates with the best grades are white men, as it is shown in the next bar graph.

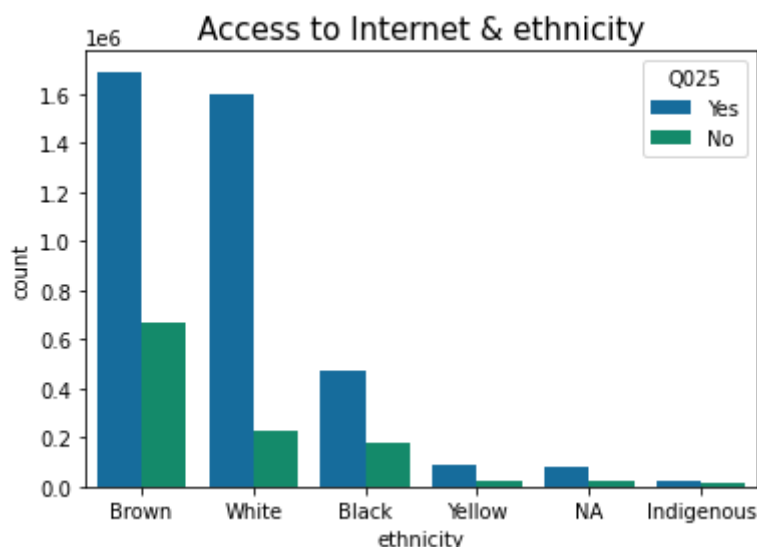


The reasons for this difference to take place could be innumerable, but one of them could be the household income. This is because with more disposable income it is possible to invest in specialized preparation materials, such as preparatory schools.



Looking at the graph we can see that the household income until R\$1320 is dominated by brown people. When it comes to income up to R\$3520, the majority of candidates is white. This difference is amplified when the household income is up to R\$10560.

The access to the internet can also influence the students' performance, because they can have access to different types of sources and knowledge to get rid of doubts more easily. The correlation between ethnicity and access to the internet is shown in the graph below:

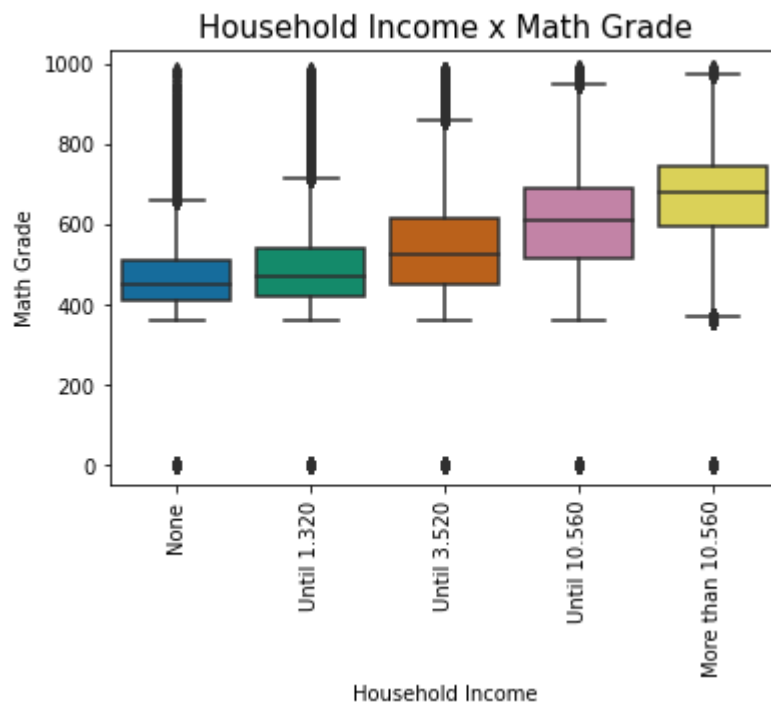


The graph shows that brown people are those who have the lowest internet accessibility. However, note that the relationship between the number of white and black people and those who do not have access, both have very close values, that is, this corresponds to only 10.91% of white people enrolled in ENEM 2019 and 30.84 % of black people. Showing a considerable difference in terms of accessibility.

b. Household Income

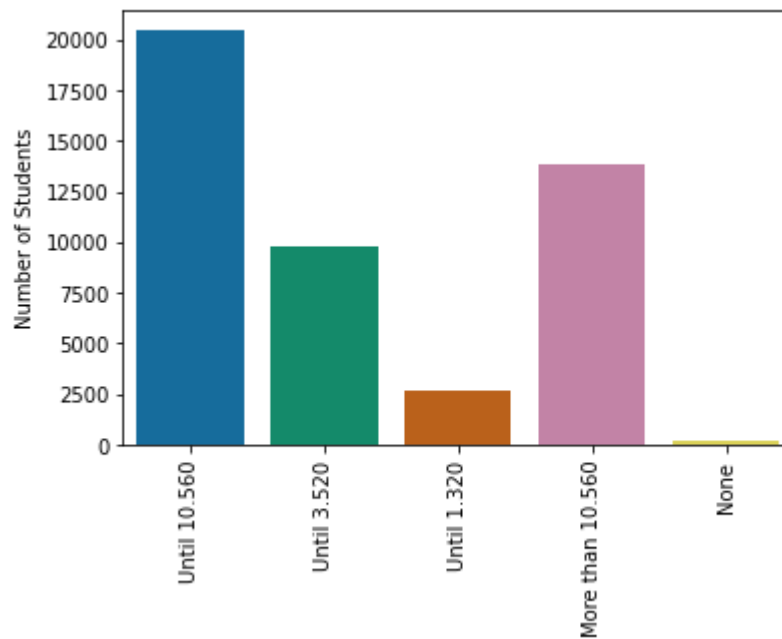
Household income can also be an aspect that directly influences the students' performance. This happens because when the student has a higher family income, he has more time available to concentrate on his studies, not needing to work. Young people who work and study simultaneously need to learn to deal with mental and physical fatigue in order to perform well on the exam.

The previous statement is supported by the graph below, which shows that the medians of students' math grades increase according to their household income.



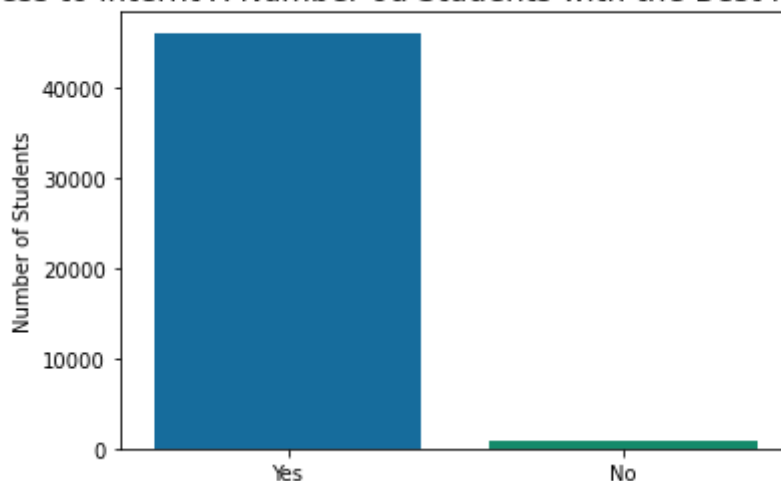
The number of candidates with grades above 800 points in the math test was mostly students with a household income of up to R\$10,560.00 or higher. It is worth remembering that both categories together represent only 15.22% of the total ENEM 2019 candidates and have a representation of 72.95% in the best grades. The following graph shows this correlation:

Household Income X Number of Students with the Best Math Grades



Continuing to analyze the math grades, it is possible to observe that 98.24% of the students with the best results have access to the internet. The IBGE (Brazilian Institute of Geography and Statistics) highlights that the financial issue is the main factor presented as the reason for people not having access to the internet. This proves the fact that it is an extremely relevant factor in the good performance of the student.

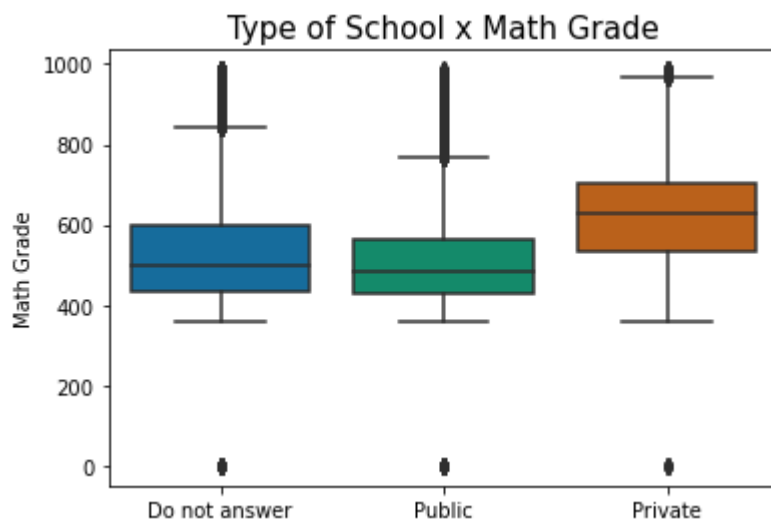
Access to Internet X Number of Students with the Best Math Grades



The availability of income to invest in private education and ENEM preparatory courses also influence the students' performance. Unfortunately, nowadays many public schools do not have the same quality of education as private schools and this may have worsened during the first two years of Jair Bolsonaro's government. Investments by the

Ministry of Education were the lowest since 2010, which may directly reflect on the performance of candidates coming from public schools.

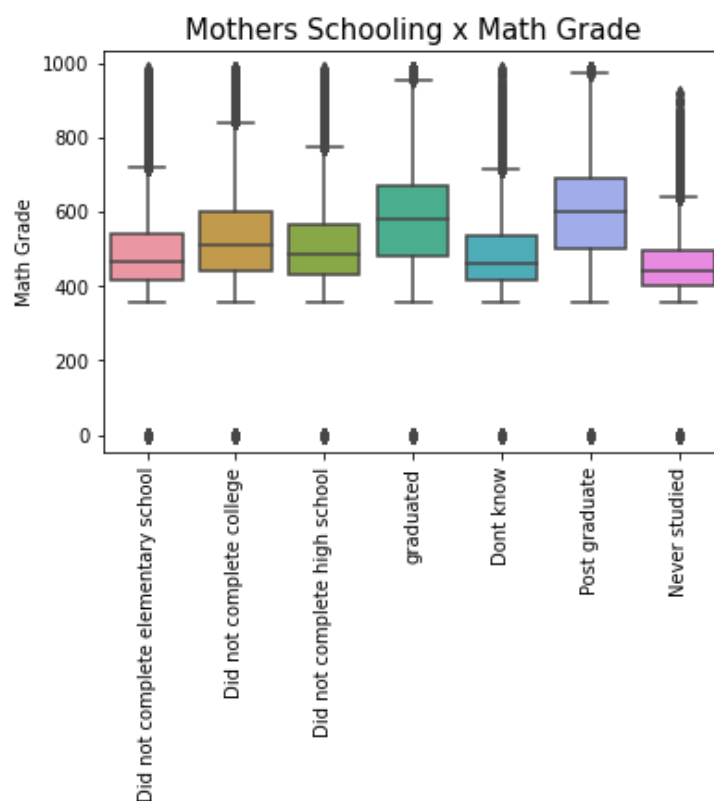
In the graph below, it is possible to observe that the median math scores were considerably higher for students from private schools:



c. Mother's Schooling

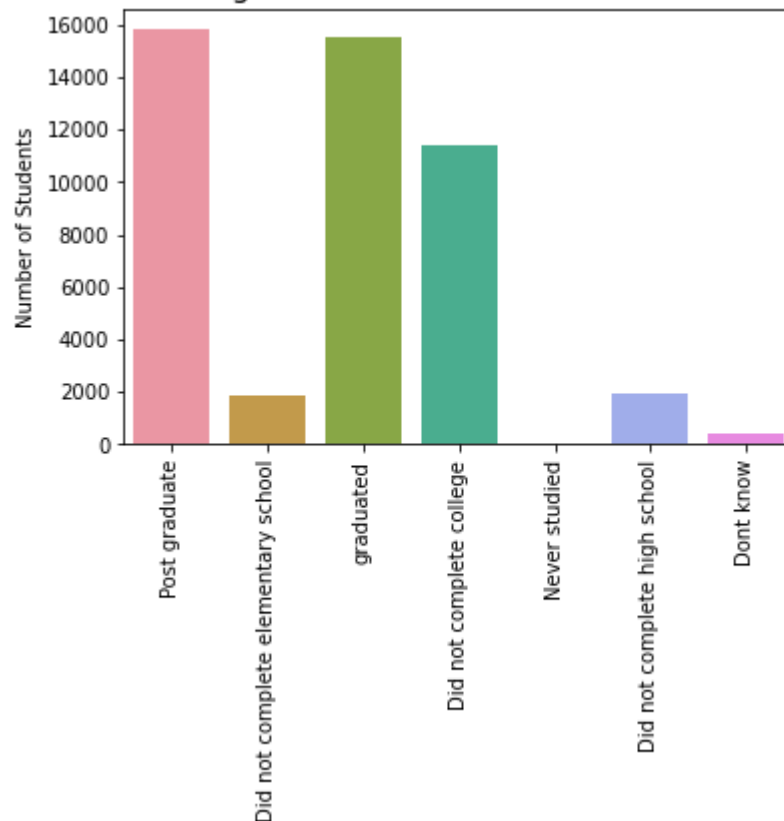
A study published in 2019 by Ipea (Institute for Applied Economic Research) points out that the lower the mother's education, the lower the literacy level.

This study is reflected in the performance of students in ENEM 2019 through math grades, as it is possible to observe that students who have undergraduate and graduate mothers performed better in the test. The following graph shows this relationship:



Students who performed more than 800 points on the math test make this relationship even more evident in the following graph:

Mothers Schooling X Number of Students with Best Math Grades



4. Conclusion:

Through the data analyzes presented above, it is clear to see that household income, color/race and mother's education are related to each other and affect the students' performances. This was proven through statistical calculations and graphics based on ENEM 2019 microdata.

The Quota Law for federal educational institutions (Law 12.711/2012) was created in 2012 and will expire in 2022, if not renewed by the Congress. Therefore, it is essential that social, racial and disability quotas are renewed, as they help to alleviate the disparity of these inequalities. In addition, this serves as a historical reparation for the Brazilian people and guarantees representation to minorities.

Every Brazilian citizen has the right to have access to quality education, regardless of their origin, color, gender or sexual orientation and they must be treated with equality.

5. References:

MEC (Ministério da Educação). Available in: [http://portal.mec.gov.br/enem-sp-2094708791#:~:text=Criado%20em%201998%2C%20o%20Exame,ao%20fim%20da%20escolaridade%20b%C3%A1sica.&text=O%20Enem%20%C3%A9%20utilizado%20como,Universidade%20para%20Todos%20\(ProUni\)>](http://portal.mec.gov.br/enem-sp-2094708791#:~:text=Criado%20em%201998%2C%20o%20Exame,ao%20fim%20da%20escolaridade%20b%C3%A1sica.&text=O%20Enem%20%C3%A9%20utilizado%20como,Universidade%20para%20Todos%20(ProUni)>) Accessed in: 05 dec. 2021

MEC (Ministério da Educação). Available in: <http://portal.mec.gov.br/component/tags/tag/enem-2019> Accessed in: 03 dez. 2021

IPEA (Instituto de Pesquisa Econômica Aplicada). Available in: https://www.ipea.gov.br/desafios/index.php?option=com_content&view=article&id=1111:catid=28&Itemid=23 Accessed in: 03 dez. 2021

IPEA (Instituto de Pesquisa Econômica Aplicada). Available in: https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=35191 Accessed in: 03 dez. 2021

UNICAP (Universidade Católica De Pernambuco). Available in: <http://www.unicap.br/coloiodehistoria/wp-content/uploads/2013/11/4Col-p.309.pdf> Accessed in: 04 dez. 2021

Folha de São Paulo. Available in: <https://www1.folha.uol.com.br/educacao/2021/02/sobbolsonaro-gasto-do-mec-com-investimentos-e-o-menor-desde-2015.shtml> Accessed in: 04 dez. 2021

IBGE (Instituto Brasileiro de Geografia e Estatística) < <https://www.ibge.gov.br/> > Accessed in: 03 dez. 2021

Globo
<<https://g1.globo.com/educacao/noticia/2021/02/13/lei-de-cotas-ainda-nao-cumpriuseu-objetivo-e-precisa-ser-renovada-em-2022-diz-reitor-da-zumbi-dos-palmares.ghtml>>
Accessed in: 07 dez. 2021