

# Usando regressão linear para estimar os pesos de questões de matemática do ENEM na nota final do aluno

Francisco Lima Figueiredo<sup>1</sup>

2/AUG/2025

## Resumo

O presente artigo apresenta os resultados preliminares de um estudo para tentar responder a seguinte pergunta: É possível atribuir um peso, uma nota a cada questão de matemática do ENEM de forma a apresentar os resultados ao aluno em termos equivalentes à de uma avaliação clássica.

Foram feitas análises estatísticas a partir dos microdados do ENEM ...

**Palavras-chave:** ENEM. Matemática. TRI - Teoria de Resposta ao Item. Mínimos Quadrados Ordinários. Avaliação Educacional.

## 1 Introdução

Obviamente o ENEM se tornou hoje fonte de preocupação de alunos e pais em todo país. As notas do ENEM decidem futuros e rumos familiares, uma vez que suas notas determinam o ingresso na rede de ensino superior, pública e particular, bem como definem valores e limites de financiamento para estudantes e avaliação do ensino.

Por mais inegável que existe uma certa aceitação de como o ENEM é feito e sua avaliação, não podemos negar que o seu esquema de notas à olhos leigos assusta, confunde e deixa sem ferramentas os professores, além daquelas obtidas pelos organizadores e pelo empirismo, para argumentar acerca de quanto vale cada questão.

O aluno é, desde criança, imerso em um sistema de notas 'linear' onde cada questão tem um peso e/ou que a quantidade de questões acertadas está diretamente ligada ao seu desempenho.

---

<sup>1</sup>Aluno concluinte do curso de pós-graduação em MATEMÁTICA E ESTATÍSTICA, Universidade ESPG. Currículo Lattes <<http://lattes.cnpq.br/2974257608350963>>. Matrícula 102020777.

Ao fazer o ENEM, os alunos ficam as "cegas" antes, durante e após o certame. E uma simples análise de vídeos no youtube como visto em [Valim \(2019\)](#), [Brasil \(2021\)](#) e [Educação \(2020\)](#) as explicações, embora seja didática e aparentemente simples, não respondem as especificidades da Teoria da Resposta ao Item. Cabe ressaltar que os vídeos citados tem imprecisões quando comparadas ao cerne da TRI, seja por medo ou desconhecimento matemático-estatístico necessário para compreender, como brilhantemente explicado em [Pasquali \(2018\)](#) em relação aos colegas do ramo das ciência sociais:

Além disso, diante do receio da maioria dos cientistas sociais, em particular do psicólogo e do pedagogo, frente ao pensamento matemático, bem como o fraco preparo desses profissionais nas áreas da matemática, este livro procura enveredar o mínimo na sofisticação matemática e estatística que a TRI pode assumir.

Obviamente, não é possível escapar totalmente do pensar matemático nessa área da medida, pois seria utilizar o número, o objeto específico das matemáticas, sem fazer uso dos princípios e métodos dessas. Quem entende e trabalha o número são, necessariamente, as matemáticas. Então não há como eliminá-las no tratamento dos dados empíricos expressos via medida. Contudo para fazer uso inteligente dos princípios e da tecnologia da TRI não é necessário entrar nas altas sofisticações matemáticas e estatísticas que eles permitem. Evidentemente, quem é capaz de seguir por esse caminho tem maiores vantagens na inteligência da problemática psicométrica e do seu ramo mais sofisticado, que é a TRI.

O presente trabalho foca em encontrar pesos para cada questão do ENEM de um determinado ano de forma que o aluno possa, mesmo que por aproximação, verificar o peso de cada questão na sua nota final.

## 2 Referencial Teórico

A teoria envolvida nesse artigo é bem ampla e está melhor detalhada na trabalho, em formato de tese, auxiliar a esse artigo em [Figueiredo \(2021\)](#).

### 2.1 Avaliação - Modelos matemáticos

Segundo [Ferreira \(2018\)](#), A avaliação sempre foi um dos pontos críticos do sistema educacional, pois os métodos utilizados largamente não permitem medir o quão verdadeiro é o conhecimento do indivíduo, de modo que, os testes estão sofrendo revisões e novos instrumentos são criados para resolver e ajudar em uma adequada avaliação, principalmente na área educacional. Os resultados dados pelo número ou pelo percentual de acertos em uma escala própria em geral não diz muito sobre o respondente e prejudica o julgamento sobre um determinado conteúdo. Segundo [Klein \(2013\)](#) costuma-se interpretar que as notas acima de setenta por cento são consideradas adequadas e as abaixo de cinquenta por cento como fracasso. Evidente que, quando as notas são abaixo do esperado, associa-se que, provavelmente a prova estava num nível acima dos alunos ou que os itens estavam em um grau de dificuldade demasiado. Esta percepção não leva em consideração o estado de espírito, a habilidade, a competência do aluno ou a discriminação dos itens.

Para lidar com esse problema, modelos matemáticos foram criados, entre eles a Teoria Clássica dos Testes - TCT e a Teoria de Resposta ao Item - TRI. A teoria TCT expressa escores brutos ou padronizados e tem sido utilizada nos processos de avaliação e

seleção de indivíduos e, as análises estão sempre ligadas à prova como um todo, ficando inviável a comparação entre itens e indivíduos. Já a teoria TRI propõe modelos que analisa características do indivíduo que não podem ser medida diretamente. O método sugere formas de apresentar a probabilidade de um indivíduo dar uma certa resposta, levando em consideração suas habilidades na área de conhecimento. A grande vantagem da TRI sobre a TCT é que essa teoria permite a comparação entre populações, desde que as provas submetidas tenham itens comuns e ainda, a comparação entre os resultados dos indivíduos submetidos à provas diferentes em épocas diferentes. Essa vantagem se dá porque esta teoria tem como elemento central o item e não a prova como um todo. Assim, a introdução da TRI trouxe muitas vantagens sobre o método tradicional, pois permite estimar e comparar os resultados dos alunos, mesmo que respondam itens diferentes, colocando-os em uma mesma escala (Ferreira (2018)).

### 2.1.1 Teoria Clássica dos Testes - TCT

Segundo Klein (2013) a Teoria Clássica dos Testes - TCT, orientou por muito tempo o desenvolvimento dos testes educacionais, psicológicos, avaliativos e etc. A TCT baseia-se no escore bruto de cada indivíduo, ou seja, o resultado é obtido apenas comparando a quantidade de itens respondidos corretamente por cada indivíduo, logo a habilidade é estimada pelo número de acertos, sendo X escore observável obtido por um estudante em um teste.

A confiabilidade depende dos alunos testados, do tamanho e dos itens do teste. Itens com coeficiente ponto bisserial, correlação entre acerto e número de acertos na prova, baixo ou negativo acrescentam pouco à confiabilidade.

A Teoria Clássica dos Testes, apesar de amplamente utilizada nos diversos tipos de avaliação, têm limitações, dentre as quais foram listadas por Klein (2013)

- As estatísticas que descrevem os itens de teste dependem do grupo de estudantes que fazem o teste;
- Os escores de teste que descrevem o desempenho dos alunos dependem dos itens apresentados aos alunos;
- A Teoria Clássica dos Testes só pode ser utilizada em situações nas quais todos os alunos fazem o mesmo teste (ou formas “paralelas” de teste);
- A Teoria Clássica dos Testes não fornece um modelo do desempenho de um aluno em um item.
- A maioria das aplicações da Teoria Clássica dos Testes assume incorretamente que os erros de medida têm a mesma variabilidade para todos os alunos.

### 2.1.2 Teoria de Resposta ao Item - TRI

Segundo Klein (2013) a Teoria Clássica dos Testes - TCT, orientou por muito tempo o desenvolvimento dos testes educacionais, psicológicos, avaliativos e etc. A TCT baseia-se no escore bruto de cada indivíduo, ou seja, o resultado é obtido apenas comparando a quantidade de itens respondidos corretamente por cada indivíduo, logo a habilidade é estimada pelo número de acertos, sendo X escore observável obtido por um estudante em um teste.

A confiabilidade depende dos alunos testados, do tamanho e dos itens do teste. Itens com coeficiente ponto bisserial, correlação entre acerto e número de acertos na prova, baixo ou negativo acrescentam pouco à confiabilidade.

O TRI trouxe diversos avanços, como mencionam [Pasquali \(2018\)](#) referenciando-se à [Rogers \(1991\)](#):

- a) O *cálculo do nível de aptidão do sujeito* é independente da lista de itens utilizados: demonstra-se que a habilidade do sujeito é independente do teste (not test-dependent). Na Psicometria Clássica, o escore do sujeito dependia e variava segundo a facilidade do teste aplicado, ou dos erros produzidos. Assim, tais escores tinham base comparativa e, mesmo após ajustes, os escores ainda não permitiam comparações, sobretudo porque os testes produziam diferenças nas variâncias nos erros de medida. No caso da TRI, não importa as questões ou conjunto de questões que você utilize, obviamente que estejam medindo o mesmo traço latente, eles irão produzir o mesmo nível de aptidão do sujeito, dentro dos sempre presentes erros de medida em qualquer ramo da ciência;
- b) O *cálculo ou estimativa dos parâmetros da função de probabilidade dos itens* (dificuldade e discriminação) é realizada independente da amostra de sujeitos utilizada: diz-se que são independentes dos sujeitos ou não dependente de grupo (not group-dependent). Na teoria clássica, os parâmetros dependiam demais dos indivíduos amostrados terem maior ou menor aptidão;
- c) A TRI permite *comparabilidade dos itens com a aptidão do sujeito*. em outras palavras que se avalia a proficiência de um sujeito utilizando-se questões com dificuldade tal que se situam ao redor do requerido tamanho da aptidão do sujeito, sendo, assim, possível utilizar desde questões mais fáceis para sujeitos com habilidades inferiores e questões com dificuldade mais elevada para sujeitos com maior proficiência, e ainda sim produzindo scores comparáveis em ambos os casos. Na psicometria clássica, era aplicado o mesmo teste sempre, hermeticamente fechado, para todos os sujeitos, de tal sorte que, se o teste fosse fácil, seriam bem avaliados os sujeitos de aptidão menor e mal avaliados os sujeitos de aptidão superior e, caso o teste fosse difícil, faria o contrário.
- d) A TRI é um modelo que não precisa fazer suposições que aparentam serem improváveis, tais como os erros de medida serem iguais para todos os testandos;
- e) A TRI não necessita trabalhar com testes estritamente paralelos como exige a psicometria clássica.

O TRI por si só merece uma tese a parte, mas basicamente a teoria se baseia em algumas premissas:

- a) A intenção do modelo é mensurar a habilidade, traço latente ou aptidão que ele possua, seja numa matéria escolar ou avaliação psíquica. A essa habilidade vamos atribuir um score  $\theta_j$ <sup>1</sup>, mensurável numa escala de média 0 e desvio padrão 1. No

---

<sup>1</sup> Leia-se habilidade do indivíduo  $i$

ENEM, por exemplo, eles ajustam para, a grosso modo, uma nota com média 500 e desvio padrão 100 usando algo parecido com  $N_j = 100\theta_j + 500$  <sup>2</sup>

- b) Ao efetuar uma prova, para cada questão é mensurado ou atribuído uma função de probabilidade para que, com o indivíduo de habilidade  $\theta_j$  acerte a questão  $U_i$  que é definido, para o modelo de 3 parâmetros, como:

$$P(U_{ji} = 1|\theta_j) = c_i + (1 - c_i) \cdot \frac{1}{1 + e^{-Da_i(\theta_i - b_i)}}$$

Onde  $P(U_{ji} = 1|\theta_j)$  é a probabilidade do indivíduo  $j$  acertar a questão  $i$ , dado sua habilidade ou traço latente  $\theta_j$ ,  $c_i$  é o **parâmetro de acerto ao acaso** do item  $i$  que representa a probabilidade de indivíduos com baixa ou nenhuma habilidade  $\theta_j$  responderem corretamente o item  $i$  (muitas vezes referido como a probabilidade de acerto casual) e  $b_i$  é o **parâmetro de dificuldade** (ou de posição) do item  $i$ , medido na mesma escala da habilidade. Em outras palavras, esse parâmetro mede a habilidade  $\theta_j$  necessária ao indivíduo para que tenha 50% de chance de acertar a questão; e  $a_i$  é o **parâmetro de discriminação** (ou de inclinação) do item  $i$ , com valor proporcional à inclinação da Curva Característica do Item — CCI no ponto  $b_i$ . Segundo [Gomes \(2014\)](#) para entendermos melhor esse parâmetro, devemos ter clareza a respeito do que venha ser a discriminação de um item. É plausível pensarmos que numa prova feita com vários respondentes com habilidades distintas, alguns itens serão considerados fáceis pelos indivíduos com uma proficiência alta, mas podem ser considerados difíceis pelos indivíduos com uma baixa proficiência. A discriminação é exatamente essa característica do item capaz de diferenciar indivíduos com habilidades distintas.

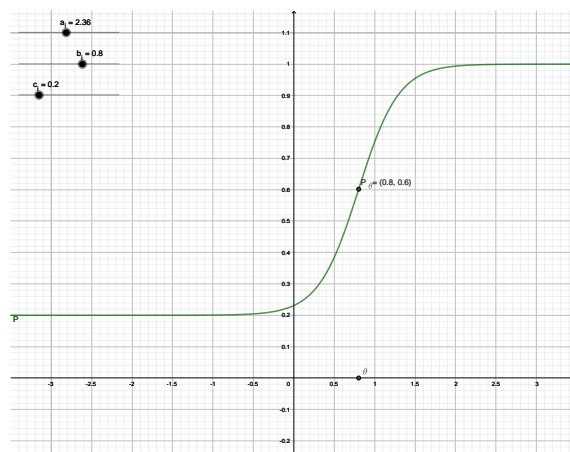


Figura 1 – CCI - Curva Característica do Item, disponível para manipulação em <https://www.geogebra.org/classic/nq8dk5x9>.

Fonte: Autor

Essa curva de probabilidade é caracterizada por uma curva em formato de S, basicamente a medida que aumenta a habilidade  $\theta$  do aluno, maior a probabilidade de acerto da questão  $Q_i$ , como demonstra a figura 1

<sup>2</sup> Leia-se: A nota do indivíduo  $j$  é 100x a habilidade  $\theta$  do indivíduo somado a 500 que é a média esperada

## 2.2 Mínimos Quadrados Ordinários

De acordo com [Wikipedia \(2021\)](#), o Método dos Mínimos Quadrados (MMQ), ou Mínimos Quadrados Ordinários (MQO) ou OLS (do inglês Ordinary Least Squares) é uma técnica de otimização matemática que procura encontrar o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados (tais diferenças são chamadas resíduos).

Para o caso geral de 2 dimensões ( $\mathbb{R}^2$ ), temos um conjunto finito de  $n$  dados  $\mathbb{D} = \{(x_i, y_i) | 1 \leq i \leq n\}$ , e uma função parametrizável  $f(x, \alpha, \beta, \gamma, \dots)$ , onde  $\alpha, \beta, \gamma, \dots$  são parâmetros que se desejam achar com o objetivo de minimizar o somatório dos quadrados dos resíduos  $r_i = y_i - f(x_i, \alpha, \beta, \gamma, \dots)$  ou seja :

$$\text{minimizar } S_r = \sum_{i=1}^n [y_i - f(x_i, \alpha, \beta, \gamma, \dots)]^2$$

Essa minimização pode ser obtida, derivando-se em relação aos parâmetros e igualando a zero, ou seja

$$\frac{\partial S_r}{\partial \alpha} = 0, \frac{\partial S_r}{\partial \beta} = 0, \frac{\partial S_r}{\partial \gamma} = 0, \dots$$

Na vida prática, usamos algoritmos numéricos, muito populares em pacotes de estatística como Python e R, para o cálculo aproximando desses parâmetros, como usado nesse artigo.

## 3 Metodologia

Foi utilizado o Python como ferramenta de análise de dados e técnicas de data science para avaliar os microdados do ENEM em busca de insights e testes de hipótese sobre os dados em busca de informações úteis sobre o assunto.

### 3.1 Catalogação e classificação das questões do ENEM

Para iniciar os estudos e coleta e gravação das informações, Anteriormente ao processamento acima descrito, foi criado um banco de dados Access <sup>3</sup> contendo todas as questões dos ENEM, contendo, para cada questão:

1. **Imagem da questão:** cada questão possui uma imagem fidedigna, as vezes ajustada, da questão no caderno de questões azul de cada prova.
2. **Habilidade:** habilidade avaliada na questão, conforme parâmetros retirado dos microdados e comparado com o documento **Matriz de referência ENEM**, conforme [Inep \(2020\)](#).
3. **Gabarito da questão:** tirada dos microdados do ENEM, equivale a alternativa considerada correta para fins de constatação estatística

---

<sup>3</sup> Microsoft Access, integrante do pacote Office

4. **Conhecimentos mapeados para resolução** - lista padronizada de conhecimentos, mapeados a partir do conhecimento necessários para resolver de forma simples a questão, baseados no conhecimento do autor.
5. **Mapa de Respostas** - lista com o número da questão, conforme a prova. Já reservada futuramente caso a questão mude a ordem das alternativas.
6. **Dificuldade** - Classificação de dificuldade que o aluno teria para resolução da questão, baseado na opinião do autor baseada em sua experiência e na classificação de ao menos 2 professores que corrigiram as questões em vídeos do Youtube.

Além das variáveis acima mapeados, foram criadas campos nesse arquivo de banco de dados para recepcionar as variáveis geradas nesse e em outros estudos, sendo que para esse artigo nosso escopo será o campo **Peso Nota** ou registrado em [Figueiredo \(2021\)](#) denominado **Peso do Item na Nota Final de Matemática**.

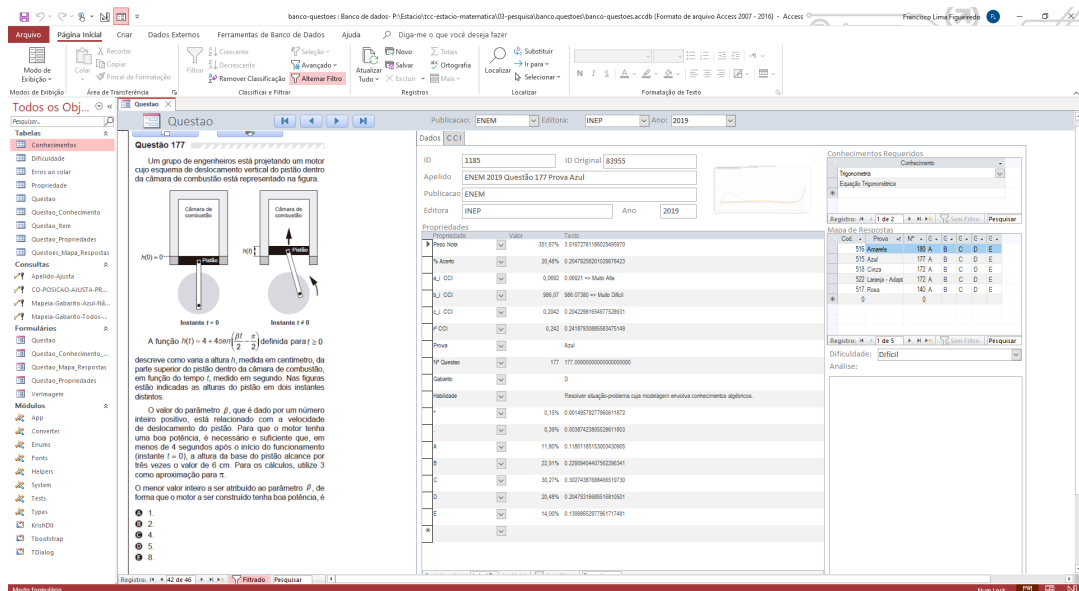


Figura 2 – Banco de Questões desenvolvido pelo Autor para catalogação das variáveis de cada questão, disponível no link <https://github.com/ChicoFigueiredo/Estacio-TCC-Estacio-Matematica/tree/main/03-pesquisa/banco.questoes> para download

Todos os resultados desse banco estão disponíveis de maneira organizada em [Figueiredo \(2021\)](#) no Anexo C, disponível em <https://github.com/ChicoFigueiredo/Estacio-TCC-Estacio-Matematica/tree/main/03-pesquisa/banco.questoes> conforme exemplo na figura 3:

## c.110 ENEM 2018 Questão 155 Prova Azul

## Questão publicada

## QUESTÃO 155

O remo de assento deslizante é um esporte que faz uso de um barco e dois remos do mesmo tamanho.

A figura mostra uma das posições de uma técnica chamada afastamento.



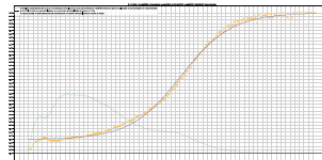
Disponível em: [www.remobrasil.com](http://www.remobrasil.com). Acesso em: 6 dez. 2017 (adaptado).

Nessa posição, os dois remos se encontram no ponto A e suas outras extremidades estão indicadas pelos pontos B e C. Esses três pontos formam um triângulo ABC cujo ângulo  $\widehat{BAC}$  tem medida de  $170^\circ$ .

O tipo de triângulo com vértices nos pontos A, B e C, no momento em que o remador está nessa posição, é

- Ⓐ retângulo escaleno.
- Ⓑ acutângulo escaleno.
- Ⓒ acutângulo isósceles.
- Ⓓ obtusângulo escaleno.
- Ⓔ obtusângulo isósceles.

## Curva CCI



## Parâmetros estatísticos da questão

- ID Questão INEP: 1211
- Habilidade: Identificar características de figuras planas ou espaciais.
- Conhecimentos mapeados para resolução:
  - Geometria
  - Ângulos
- Gabarito: E
- Percentual de Acerto: 21,64% Difícil
- Percentual de Erro: 78,36%
- Padrão de Resposta
  - Dupla Resposta: 0,14%
  - Em Branco: 0,21%
  - Letra A: 25,51%
  - Letra B: 15,74%
  - Letra C: 17,28%
  - Letra D: 19,48%
  - Letra E: 21,64%
- Peso do Item na Nota Final de Matemática: 23,19
- Curva Característica do Item TRI Modelada:
  - Discriminação  $a_i$ : 0,0182 Muito Alta
  - Dificuldade  $b_i$ : 701,93 Muito Difícil
  - % Acerto ao acaso (chute)  $c_i$ : 9,97% pouco
  - Coeficiente de determinação  $r^2$  da curva: 0,99704
- Classificação do Autor: Fácil

Figura 3 – Questão detalhada em Figueiredo (2021) elaborada pelo autor, página 202

Para geração rápida de informações, também foi produzida a partir do banco de dados também uma planilha Excel, com informações mais importantes, disponível em [https://github.com/ChicoFigueiredo/Estacio-TCC-Estacio-Matematica/blob/main/03-pesquisa/banco.questoes/Mapa\\_Questoes.xlsm?raw=true](https://github.com/ChicoFigueiredo/Estacio-TCC-Estacio-Matematica/blob/main/03-pesquisa/banco.questoes/Mapa_Questoes.xlsm?raw=true)

### 3.2 Processamento e análise dos microdados ENEM

Como cita Morita (2021), a etapa mais demorada e trabalhosa do projeto é a obtenção e tratamento de dados. Normalmente, quando estudamos Data Science, as bases de dados usadas estão prontas para análise e são de fácil acesso. Na prática, é o oposto! Dados têm diferentes fontes e formatos. Podemos analisar tabelas, imagens, áudios, textos vindos de redes sociais, sites, bancos de dados, pesquisas, documentos digitalizados, etc. Além disso, eles estão longe de prontos para serem analisados: precisam antes serem tratados e integrados.

Os microdados do Enem são o menor nível de desagregação de dados recolhidos por meio do exame. Eles atendem a demanda por informações específicas ao disponibilizar as provas, os gabaritos, as informações sobre os itens, as notas e o questionário respondido pelos inscritos no Enem (Inep (2021)).

Foram realizados o download dos arquivos de microdados dos ENEM entre os anos



de 2016 a 2021 que consistem, basicamente, de arquivos CSV <sup>4</sup> gigantescos, que variam na casa de 3 Gb <sup>5</sup> de tamanho cada arquivo.

Para fins desse estudo, foi feito um tratamento dos dados para normalização e padronização dos dados, pois algumas variáveis se alteraram de um ano para outro, tais como posição das questões conforme a cor do caderno, a presença ou não, ou posição, de algumas variáveis sociais, entre outros tratamentos. Isso está detalhado nos notebooks python (extensão ipynb) presente no endereço <<https://github.com/ChicoFigueiredo/Estacio-TCC-Estacio-Matematica/tree/main/03-pesquisa>> ou detalhado em Figueiredo (2021).

Em resumo, primeiro foi montado um programa em um notebook python para executar o tratamento de dados e normalização de algumas tabelas como demonstra o ambiente presente na figura 4

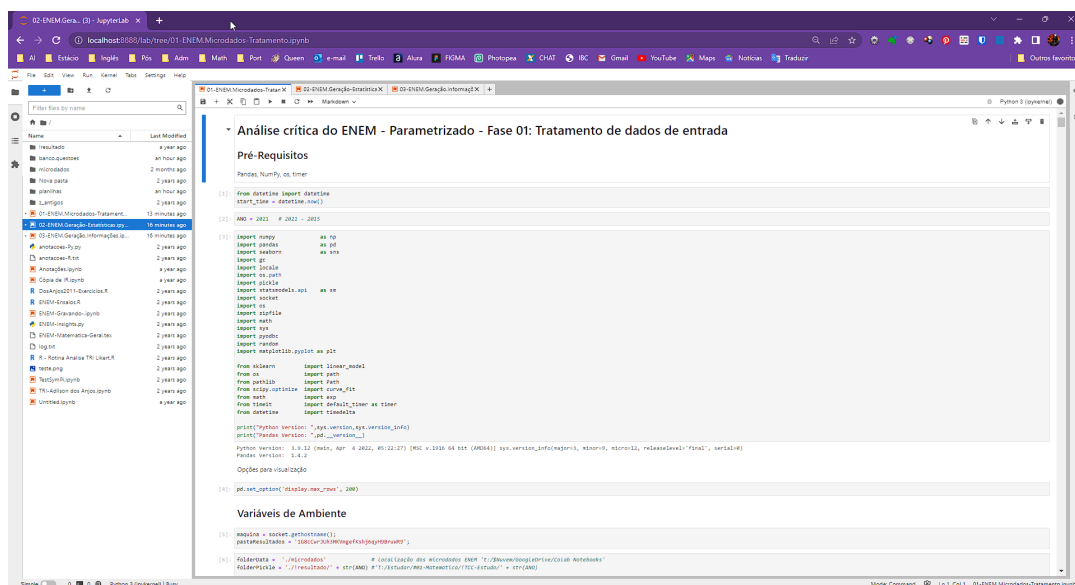


Figura 4 – Ambiente de desenvolvimento Jupyter Lab, contendo o programa de limpeza de dados encontrado em <<https://github.com/ChicoFigueiredo/Estacio-TCC-Estacio-Matematica/blob/main/03-pesquisa/01-ENEM.Microdados-Tratamento.ipynb>> Fonte: O Autor.

Em seguida, foi montado um segundo programa para gerar as informações necessárias para determinar os pesos de cada questão, em cada ano, conforme a figura 5

<sup>4</sup> CSV = Comma-separated values, valores separados por vírgula, formato de arquivo de texto de formato regulamentado pelo RFC 4180, que faz uma ordenação de bytes ou um formato de terminador de linha, separando valores com vírgulas (podendo ser separado por ponto-e-virgula). Ele comumente é usado em softwares de escritório, tais como o Microsoft Excel e o LibreOffice Calc, ou de troca de dados estatístico (R ou Python).

<sup>5</sup> 3 Gb =  $3 \cdot 1024 \cdot 1024 \cdot 1024 \approx 3,75$  bilhões de caracteres

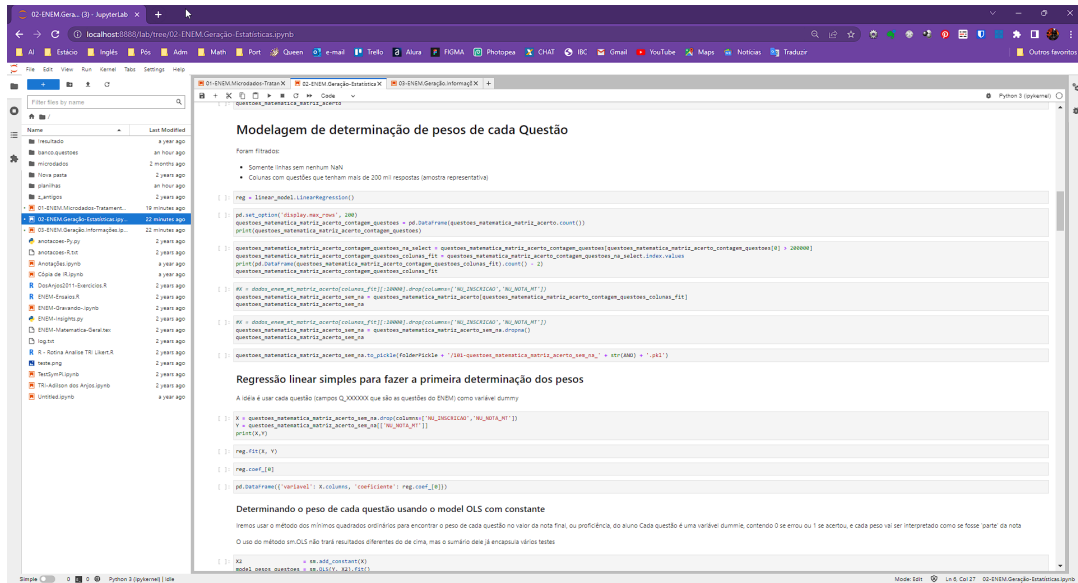


Figura 5 – Ambiente de desenvolvimento Jupyter Lab, contendo o programa de regressão linear encontrado em <https://github.com/ChicoFigueiredo/Estacio-TCC-Estacio-Matematica/blob/main/03-pesquisa/02-ENEM.Gera%CC%83%80-Estat%CC%83%84sticas.ipynb>. Fonte: O Autor.

O mesmo programa gravou os dados no banco de dados access relatado acima.

### 3.3 Análise estatística e obtenção de parâmetros das questões

Para a modelagem do problema, foi suposto que um aluno  $i$  conseguiu a nota ponderada  $N_i$  com o gabarito de questões marcadas  $(x_{i,1}, x_{i,2}, x_{i,3}, \dots, x_{i,45})$  onde  $x_{i,j}$  são variáveis dummies que possuem o valor  $x_{i,j} = 1$  caso o aluno acerte ou a questão seja anulada e  $x_{i,j} = 0$  caso ele erre ou deixe em branco.

O objetivo é encontrar os pesos  $\rho_1, \rho_2, \rho_3, \dots, \rho_{45}$  que permitam assertivamente ter a nota  $N_i$  a partir do gabarito do aluno, com o menor erro  $\varepsilon_i$  possível para o aluno.

$$N_i = \rho_1 \cdot x_{i,1} + \rho_2 \cdot x_{i,2} + \rho_3 \cdot x_{i,3} + \dots + \rho_{45} \cdot x_{i,45} + \varepsilon_i$$

ou resumidamente

$$N_i = \sum_{j=1}^{45} \rho_j \cdot x_{i,j} + \varepsilon_i$$

Em termos práticos, a otimização é:

$$\text{minimizar } \varepsilon_i = N_i - \sum_{j=1}^{45} \rho_j \cdot x_{i,j}$$

Usando o programa montado no arquivo <https://github.com/ChicoFigueiredo/Estacio-TCC-Estacio-Matematica/blob/main/03-pesquisa/02-ENEM.Gera%CC%83%80-Estat%CC%83%84sticas>.

ipynb> foi possível, a partir dos dados processados, conseguir facilmente os pesos  $\rho_j$  de cada questão usando o método de mínimos quadrados ordinários.

Para demonstrar o poder de aproximação desse modelo, o output abaixo é o sumário da regressão gerado pelo framework statsmodels.

OLS Regression Results						
=====						
Dep. Variable:	NU_NOTA_MT	R-squared:	0.971			
Model:	OLS	Adj. R-squared:	0.971			
Method:	Least Squares	F-statistic:	1.493e+06			
Date:	Sat, 08 Jul 2023	Prob (F-statistic):	0.00			
Time:	21:32:08	Log-Likelihood:	-8.4452e+06			
No. Observations:	1938906	AIC:	1.689e+07			
Df Residuals:	1938861	BIC:	1.689e+07			
Df Model:	44					
Covariance Type:	nonrobust					
=====						
coef	std err	t	P> t	[0.025	0.975]	
-----						
const	332.7341	0.035	9427.572	0.000	332.665	332.803
Q_111459	2.8457	0.038	74.448	0.000	2.771	2.921
Q_111476	16.1274	0.030	535.996	0.000	16.068	16.186
Q_111521	27.5088	0.028	976.559	0.000	27.454	27.564
Q_111716	1.2208	0.036	34.369	0.000	1.151	1.290
Q_117624	12.3872	0.036	345.981	0.000	12.317	12.457
Q_117674	-7.929e-15	3.38e-17	-234.558	0.000	-7.99e-15	-7.86e-15
Q_117675	6.5048	0.032	203.341	0.000	6.442	6.567
Q_117692	15.9844	0.033	484.028	0.000	15.920	16.049
Q_117731	0.3077	0.035	8.738	0.000	0.239	0.377
Q_117831	3.4616	0.033	104.379	0.000	3.397	3.527
Q_117832	6.1131	0.036	170.013	0.000	6.043	6.184
Q_117852	12.8248	0.030	422.071	0.000	12.765	12.884
Q_117902	21.6895	0.030	723.647	0.000	21.631	21.748
Q_117923	6.6577	0.035	192.616	0.000	6.590	6.725
Q_126019	11.1770	0.034	331.536	0.000	11.111	11.243
Q_15209	15.4162	0.031	501.540	0.000	15.356	15.476
Q_15454	11.7075	0.034	343.359	0.000	11.641	11.774
Q_24895	12.9875	0.031	413.256	0.000	12.926	13.049
Q_27471	6.3829	0.030	210.799	0.000	6.324	6.442
Q_30420	9.3008	0.033	285.676	0.000	9.237	9.365
Q_30751	17.9715	0.030	603.191	0.000	17.913	18.030
Q_30781	20.0875	0.030	679.904	0.000	20.030	20.145
Q_30836	16.6620	0.031	538.522	0.000	16.601	16.723
Q_37576	15.2257	0.030	513.126	0.000	15.168	15.284
Q_39923	17.3462	0.038	460.508	0.000	17.272	17.420
Q_59786	4.2191	0.034	123.573	0.000	4.152	4.286
Q_59925	3.8115	0.029	133.686	0.000	3.756	3.867
Q_66618	7.1930	0.033	217.297	0.000	7.128	7.258
Q_67554	23.9873	0.030	807.730	0.000	23.929	24.046

Q_67990	6.8781	0.031	219.870	0.000	6.817	6.939
Q_81869	8.7419	0.033	267.797	0.000	8.678	8.806
Q_82880	7.0816	0.031	229.864	0.000	7.021	7.142
Q_83741	16.6998	0.029	578.695	0.000	16.643	16.756
Q_84258	31.5590	0.031	1007.481	0.000	31.498	31.620
Q_85219	8.4527	0.030	279.341	0.000	8.393	8.512
Q_86433	37.2710	0.033	1135.362	0.000	37.207	37.335
Q_86767	15.7305	0.031	501.159	0.000	15.669	15.792
Q_88357	2.3361	0.036	65.187	0.000	2.266	2.406
Q_88461	11.4151	0.038	303.603	0.000	11.341	11.489
Q_88518	21.5323	0.031	698.227	0.000	21.472	21.593
Q_89238	3.2530	0.030	107.585	0.000	3.194	3.312
Q_95820	37.8954	0.030	1261.341	0.000	37.837	37.954
Q_96222	3.9435	0.034	114.926	0.000	3.876	4.011
Q_96226	12.8109	0.033	391.496	0.000	12.747	12.875
Q_97598	29.6158	0.031	962.402	0.000	29.556	29.676

Omnibus:	655529.554	Durbin-Watson:	2.000
Prob(Omnibus):	0.000	Jarque-Bera (JB):	30045210.953
Skew:	-0.894	Prob(JB):	0.00
Kurtosis:	22.202	Cond. No.	2.28e+15

Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.  
[2] The smallest eigenvalue is 2.33e-24. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

O modelo obteve um  $r^2$  de 97,1%, ou seja, 97,1% do modelo pode ser explicado exclusivamente pelo gabarito de respostas do aluno.<sup>6</sup>

## 4 Considerações Finais

### 4.1 Tabulação dos resultados

A partir dos processamentos, ano a ano, dos microdados do ENEM, foi possível tabular as informações na seguinte tabela posição da questão no caderno X Ano de aplicação. Quanto mais vermelho a célula, maior foi o peso dessa questão na nota final do aluno.

<sup>6</sup> O R-quadrado ou  $r^2$  é uma medida estatística de quão próximos os dados estão da linha de regressão ajustada. Ele também é conhecido como o coeficiente de determinação ou o coeficiente de determinação múltipla para a regressão múltipla.

Peso Posição	Ano						
	2015	2016	2017	2018	2019	2020	2021
1º	13,166	3,043	8,294	8,447	21,892	8,992	15,416
2º	29,700	6,008	6,044	7,386	7,544	44,785	11,177
3º	33,073	40,325	18,824	3,226	20,429	1,319	0,000
4º	21,835	13,628	15,949	3,922	16,579	45,867	37,895
5º	35,936	12,353	2,158	18,144	47,879	8,432	12,811
6º	5,463	5,281	13,472	1,686	2,832	3,614	15,984
7º	26,874	26,624	8,341	18,195	6,810	24,796	0,308
8º	22,175	4,822	0,027	6,227	48,616	8,510	37,271
9º	4,569	41,029	12,907	48,846	29,592	10,917	17,346
10º	26,993	12,146	9,454	15,056	-0,090	3,399	3,811
11º	15,373	17,491	21,709	24,544	2,451	1,789	2,846
12º	21,953	9,482	10,236	32,774	4,432	2,826	31,559
13º	18,960	18,632	13,127	9,130	8,734	23,929	17,972
14º	10,983	11,445	38,317	10,705	39,853	3,364	29,616
15º	13,299	11,892	41,592	11,554	18,268	9,473	11,707
16º	13,875	10,926	8,344	0,413	16,610	7,098	8,453
17º	7,677	1,195	6,827	-0,100	2,004	6,583	8,742
18º	20,512	34,613	9,414	10,632	23,011	30,198	12,825
19º	5,625	32,626	67,069	6,106	1,521	5,173	3,253
20º	7,892	39,029	28,743	23,192	8,138	14,953	6,505
21º	17,490	10,796	1,568	0,592	0,218	21,457	6,878
22º	20,221	34,393	31,678	14,633	2,300	0,000	3,943
23º	14,566	22,051	1,013	5,810	2,159	1,671	4,219
24º	17,407	2,620	17,176	62,431	1,870	8,002	2,336
25º	15,627	5,344	7,894	13,099	23,825	4,897	21,689
26º	14,156	9,154	16,200	4,255	8,770	3,719	1,221
27º	11,040	10,872	1,011	20,378	4,192	8,913	21,532
28º	13,662	9,798	13,901	0,000	0,627	10,836	3,462
29º	13,670	11,958	35,402	23,834	3,321	39,695	27,509
30º	29,947	10,529	17,347	2,537	17,788	3,770	6,113
31º	16,657	3,738	2,016	9,239	12,696	5,069	7,193
32º	24,517	2,256	2,545	9,124	2,856	20,676	7,082
33º	10,975	3,887	11,167	26,619	1,190	8,785	12,387
34º	19,089	3,273	10,611	3,219	18,018	8,466	9,301
35º	16,712	13,321	23,008	9,112	8,058	10,960	6,383
36º	12,266	3,686	7,254	28,308	19,655	11,005	11,415
37º	17,012	6,037	18,437	39,189	3,720	8,294	15,730
38º	19,711	13,588	4,274	2,194	1,515	1,458	20,087
39º	15,453	7,630	17,664	6,177	9,610	12,058	16,700
40º	4,747	0,314	4,142	1,144	0,769	18,925	16,127
41º	39,360	10,717	3,675	15,906	2,750	22,872	23,987
42º	31,062	33,233	14,913	18,579	3,517	38,126	15,226
43º	22,140	12,464	34,467	2,799	9,324	21,642	16,662
44º	11,437	36,812	5,329	16,134	27,071	12,503	6,658
45º	35,467	3,056	11,768	13,717	35,780	16,972	12,987

Para melhor compreensão, a 19ª questão de 2017 registrou um peso de 67,069

pontos, a maior da série histórica. Isso quer dizer que o aluno que eventualmente acertou essa questão, ganhou 67 pontos na nota final. Nesse ano a média da matéria Matemática e suas Tecnologias registrou uma média de 518,5, o que significa que com essa única questão, entre as 45, permitiu o aluno ganhar 13% da nota da prova.

Isso está de acordo com as diretrizes da TRI, em que o aluno com habilidade  $\theta$  tende a ter muito mais chance de acertar questões que necessitam de pouca habilidade.

Como curiosidade, a 19ª questão de 2017 mapeada no banco de dados é essa da figura 6

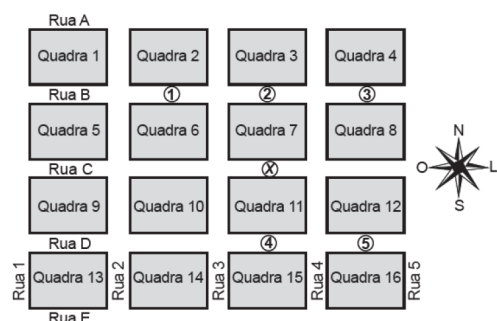
#### c.109 ENEM 2017 Questão 19 Prova Azul

Questão publicada

Curva CCI

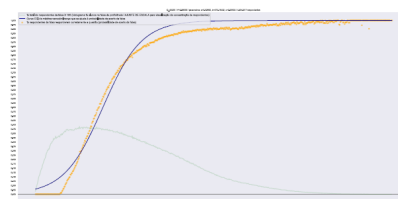
##### QUESTÃO 154

Um menino acaba de se mudar para um novo bairro e deseja ir à padaria. Pediu ajuda a um amigo que lhe forneceu um mapa com pontos numerados, que representam cinco locais de interesse, entre os quais está a padaria. Além disso, o amigo passou as seguintes instruções: a partir do ponto em que você se encontra, representado pela letra X, ande para oeste, vire à direita na primeira rua que encontrar, siga em frente e vire à esquerda na próxima rua. A padaria estará logo a seguir.



A padaria está representada pelo ponto numerado com

- A** 1.
- B** 2.
- C** 3.
- D** 4.
- E** 5.



##### Parâmetros estatísticos da questão

- ID Questão INEP: **1256**
- Habilidade: **Interpretar a localização e a movimentação de pessoas/objetos no espaço tridimensional e sua representação no espaço bidimensional.**
  - Conhecimentos mapeados para resolução:
    - Visão Espacial
    - Coordenadas
- Gabarito: **A**
- Percentual de Acerto: **59,22%** Fácil
- Percentual de Erro: **40,78%**
- Padrão de Resposta
  - Dupla Resposta: **0,15%**
  - Em Branco: **0,14%**
  - Letra A: **59,22%**
  - Letra B: **11,44%**
  - Letra C: **9,61%**
  - Letra D: **10,62%**
  - Letra E: **8,82%**
- Peso do Item na Nota Final de Matemática: **67,07**
- Curva Característica do Item TRI Modelada:
  - Discriminação  $a_i$ : **0,0286** Muito Alta
  - Dificuldade  $b_i$ : **472,13** Mediano
  - % Acerto ao acaso (chute)  $c_i$ : **0,00%** muito pouco
  - Coeficiente de Correlação  $r^2$ : **0,98535**
- Classificação do Autor: **Muito Fácil**

Figura 6 –

É possível notar que 59,2% dos alunos acertaram a questão.

## 5 Conclusões

Esses dados permitem mapear as questões mais fáceis e mostrar aos alunos que o treino e a disciplina de estudos podem facilitar o entendimento, as vezes não trivial, que acertar as questões consideradas fáceis permitem uma melhor nota.

Embora o TRI tenha um esquema não linear, é possível a partir dos dados aplicar a TCT nas provas do ENEM e obter o eventual score que ele teria caso fizesse aquela prova.

Está feito o convite para ao leitor se debruçar sobre os dados e debater eventuais novas conclusões.

## Referências

BRASIL, T. *Entenda como é calculada a nota do Enem*. 2021. 070. Disponível em: <<https://www.youtube.com/watch?v=65sNn25R1rM>>. Citado na página 2.

EDUCAÇÃO, M. da. *Enem: como calcular a nota?* 2020. 070. Disponível em: <[https://www.youtube.com/watch?v=7\\_VQzASukU8](https://www.youtube.com/watch?v=7_VQzASukU8)>. Citado na página 2.

FERREIRA, E. A. *Teoria de resposta ao item – TRI : análise de algumas questões do ENEM – habilidades 24 a 30*. 87 p. mathesis, 2018. 086. Disponível em: <[https://sca.profmatt-sbm.org.br/sca\\_v2/get\\_tcc3.php?id=160330321](https://sca.profmatt-sbm.org.br/sca_v2/get_tcc3.php?id=160330321)>. Citado 2 vezes nas páginas 2 e 3.

FIGUEIREDO, F. L. *Defesa do artigo 'Ensaio sobre o desempenho dos alunos na resolução de questões de matemática do ENEM à luz da Teoria da Resposta ao Item e parâmetros estatísticos obtidos a partir dos microdados do ENEM de 2016 à 2019.'* mathesis, 2021. Disponível em: <<https://gitlab.com/fran.fig/tcc-matematica-estacio/-/raw/main/02-Defesa/zz.pdf/002-Defesa-TCC-Francisco.Lima.Figueiredo-202004137859.pdf>>. Citado 4 vezes nas páginas 2, 7, 8 e 9.

GOMES, L. da S. *A teoria de resposta ao item na avaliação em larga escala: um estudo sobre o Exame Nacional de Acesso do Mestrado Profissional em Matemática em Rede Nacional –PROFMAT 2012*. mathesis, 2014. 080. Citado na página 5.

INEP, I. N. de Estudos e P. E. A. T. . (Ed.). *Matriz de referência ENEM*. 2020. Disponível em: <[https://download.inep.gov.br/download/enem/matriz\\_referencia.pdf](https://download.inep.gov.br/download/enem/matriz_referencia.pdf)>. Citado na página 6.

INEP, I. N. de Estudos e P. E. A. T. . *Os microdados do Enem*. 2021. Disponível em: <<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>>. Citado na página 8.

KLEIN, R. Alguns aspectos da teoria de resposta ao item relativos à estimação das proficiências. *Ensaio: Avaliação e Políticas Públicas em Educação*, FapUNIFESP (SciELO), v. 21, n. 78, p. 35–56, jul 2013. 060. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0104-40362013000100003&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-40362013000100003&lng=en&nrm=iso)>. Citado 2 vezes nas páginas 2 e 3.

MORITA, F. *As 5 fases de um projeto de Data Science*. 2021. Disponível em: <<https://awari.com.br/fases-projeto-data-science/>>. Citado na página 8.

PASQUALI, L. *TRI – Teoria de Resposta ao Item: teoria, procedimentos e aplicações*. [S.l.: s.n.], 2018. 277 p. 070. Citado 2 vezes nas páginas 2 e 4.

PASQUALI, L.; PRIMI, R. Fundamentos da teoria da resposta ao item: Tri. *Avaliação Psicológica*, scieloapsic, v. 2, p. 99 – 110, 12 2003. ISSN 1677-0471. Disponível em: <[http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1677-04712003000200002&nrm=iso](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712003000200002&nrm=iso)>. Nenhuma citação no texto.

ROGERS, R. K. H. H. S. H. J. *Fundamentals of item response theory*. [S.l.: s.n.], 1991. Citado na página 4.

VALIM, Q. com P. P. *A nota do ENEM e o TRI - Entenda!* 2019. 070. Disponível em: <<https://www.youtube.com/watch?v=QZFG5Q8sezW>>. Citado na página 2.

WIKIPEDIA. *Método dos mínimos quadrados*. 2021. Disponível em: <[https://pt.wikipedia.org/wiki/M%C3%A9todo\\_dos\\_m%C3%ACnimos\\_quadrados](https://pt.wikipedia.org/wiki/M%C3%A9todo_dos_m%C3%ACnimos_quadrados)>. Citado na página 6.