

Formação em matemática na UPE: perfilando o curso de licenciatura do campus Garanhuns, Nazaré da Mata e Petrolina mediante a utilização da análise de regressão linear simples

Mathematics training at UPE: outlining the undergraduate course at the Garanhuns, Nazaré da Mata and Petrolina campuses using simple Linear regression analysis

Formación en matemática en la UPE: esquematización del curso de pregrado en los campus de Garanhuns, Nazaré da Mata y Petrolina utilizando análisis de regresión lineal simple

Recebido: 12/07/2022 | Revisado: 26/07/2022 | Aceito: 29/07/2022 | Publicado: 07/08/2022

Letícia Karina Ramos de Lima

ORCID: <https://orcid.org/0000-0001-8014-145X>
Universidade de Pernambuco, Brasil
E-mail: leticia.karina@upe.br

Dâmocles Aurélio Nascimento da Silva Alves

ORCID: <https://orcid.org/0000-0002-7928-1276>
Universidade de Pernambuco, Brasil
E-mail: damocles.aurelio@upe.br

Daniel Nazário Gonçalves

ORCID: <https://orcid.org/0000-0003-2431-8139>
Universidade de Pernambuco, Brasil
E-mail: daniel.nazario@upe.br

Miriam Lecília Farias Ribeiro

ORCID: <https://orcid.org/0000-0002-6439-2563>
Universidade de Pernambuco, Brasil
E-mail: miriam.ribeiro@upe.br

Lauane Raimundo Cordeiro

ORCID: <https://orcid.org/0000-0002-1247-185X>
Universidade de Pernambuco, Brasil
E-mail: lauane.cordeiro@upe.br

Mauricio Costa Goldfarb

ORCID: <https://orcid.org/0000-0002-0909-1514>
Universidade de Pernambuco, Brasil
E-mail: mauricio.goldfarb@upe.br

Manoel Severino Ribeiro Neto

ORCID: <https://orcid.org/0000-0002-3588-215X>
Universidade Católica de Pernambuco, Brasil
E-mail: manoelribeiro20181@gmail.com

Arundo Nunes da Silva Júnior

ORCID: <https://orcid.org/0000-0003-4051-6389>
Secretaria de Educação do Estado de Pernambuco, Brasil
E-mail: arundo.nunes.da.silva.jr@gmail.com

Antônio Lopes Pessoa

ORCID: <https://orcid.org/0000-0002-8057-9408>
Secretaria de Educação do Estado de Pernambuco, Brasil
E-mail: lopespessoaantonio@gmail.com

Danielle Loureiro Roges

ORCID: <https://orcid.org/0000-0001-8502-1510>
Secretaria de Educação do Estado de Pernambuco, Brasil
E-mail: danielle.lou.roges2022@gmail.com

Resumo

O ENADE é uma avaliação obrigatória composta pela média ponderada das notas de Formação Geral - FG e Componentes Específicos - CE que, em concluintes de todas as graduações, mede o desempenho em características como habilidades, conteúdos e competências profissionais acerca de seus cursos. Nesse sentido, neste trabalho, objetivou-se realizar um estudo com o propósito geral de realizar uma modelagem estatística em relação à nota do ENADE do curso de Matemática da Universidade de Pernambuco, identificando os padrões de variação da nota obtida pelo curso e, após isso, através da análise de regressão simples, prever valores das futuras notas desses cursos no

exame. Para isso, foi descrito através de um modelo matemático as relações existentes entre duas variáveis (notas e anos) mediante a utilização da análise de regressão linear simples, foram levados em conta a reta de regressão, o Coeficiente de Correlação de Pearson, o Coeficiente de Determinação e a Análise dos Resíduos. A partir dos resultados observados concluiu-se que as notas do ENADE tendem a decrescer ao longo dos anos, já que há uma relação linear negativa entre as variáveis, isso ocorre mesmo tendo todas as notas de CE e a nota de Garanhuns em FG com uma relação linear positiva.

Palavras-chave: ENADE; Regressão Linear Simples; Ensino.

Abstract

ENADE is a mandatory assessment composed of the weighted average of the General Training grades - FG and Specific Components - CE that, in graduates of all graduations, measures performance in characteristics such as skills, content and professional competences about their courses. In this sense, in this work, the objective was to carry out a study with the general purpose of carrying out a statistical modeling in relation to the ENADE grade of the Mathematics course at the University of Pernambuco, identifying the patterns of variation of the grade obtained by the course and, after that, through simple regression analysis, predict values of future exam scores of these courses. For this, the existing relationships between two variables (grades and years) were described using a simple linear regression analysis, taking into account the regression line, Pearson's Correlation Coefficient, the Determination and Analysis of Waste. From the observed results, it was concluded that the ENADE grades tend to decrease over the years, since there is a negative linear relationship between the variables, this occurs even with all the CE grades and the Garanhuns grade in FG with a positive linear relationship.

Keywords: ENADE; Simple Linear Regression; Teaching.

Resumen

ENADE es una evaluación obligatoria compuesta por el promedio ponderado de las calificaciones de Formación General - FG y Componentes Específicos - CE que, en egresados de todas las carreras, mide el desempeño en características como habilidades, contenidos y competencias profesionales sobre sus cursos. En ese sentido, en este trabajo se tuvo como objetivo realizar un estudio con el propósito general de realizar una modelación estadística en relación a la nota ENADE del curso de Matemáticas de la Universidad de Pernambuco, identificando los patrones de variación de la nota obtenidos por el curso y, luego de eso, a través de análisis de regresión simple, predecir valores de puntajes de exámenes futuros de estos cursos. Para ello, se describieron las relaciones existentes entre dos variables (grados y años) mediante un análisis de regresión lineal simple, teniendo en cuenta la recta de regresión, el Coeficiente de Correlación de Pearson, la Determinación y Análisis de Residuos. De los resultados observados se concluyó que las notas de ENADE tienden a disminuir con los años, ya que existe una relación lineal negativa entre las variables, esto ocurre incluso con todas las notas de CE y la nota de Garanhuns en FG con una relación lineal positiva.

Palabras clave: ENADE; Regresión Lineal Simple; Enseñanza.

1. Introdução

Em 1995, foi criado o Exame Nacional de Cursos (ENC), denominado Provão, que a partir da avaliação de cursos superiores e seus resultados, tinha como principal objetivo conseguir dados que influenciassem na formulação de ações e melhorias para a educação superior brasileira. O curso de matemática participou do ENC a partir do ano 1998. Em 2004 foi implantado um novo Sistema Nacional de Avaliação da Educação Superior (SINAES), com os mesmos objetivos do Provão, porém com a implantação também do Exame Nacional de Desempenho de Estudantes (ENADE), sendo publicado para os estudantes de matemática somente no ano de 2005.

De acordo com o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (2005), o ENADE é uma avaliação obrigatória que, em concluintes de todas as graduações, mede o desempenho em características como habilidades, conteúdos e competências profissionais acerca de seus cursos e é realizado em períodos que não ultrapassem três anos. Utiliza como instrumentos uma prova sobre o curso e um questionário socioeconômico para obter estatísticas sobre as instituições de ensino superior (IES) e os estudantes. Essas estatísticas são utilizadas para, a partir do nivelamento do desempenho, criar políticas públicas em prol de melhorias para a qualidade e serviços em educação.

A pesquisa em questão tem como propósito geral realizar uma modelagem estatística em relação à nota do ENADE (composto pela média ponderada das notas de Formação Geral - FG e Componentes Específicos - CE) do curso de Matemática

da Universidade de Pernambuco – Campus Garanhuns, Nazaré da Mata e Petrolina, como objetivos específicos pretende identificar os padrões de variação da nota obtida pelo curso e, após isso, através da análise de regressão simples, prever valores das futuras notas desses cursos no exame.

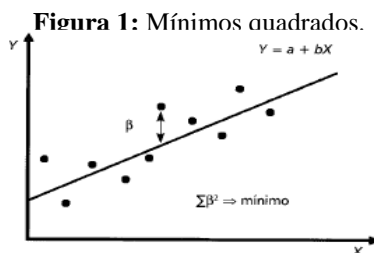
2. Metodologia

Nesse trabalho, a análise de regressão tem por objetivo descrever através de um modelo matemático as relações existentes entre duas ou mais variáveis a partir de n observações dessas variáveis e, desse modo, prever valores futuros de uma variável (dependente), ou seja, com base em uma ou mais variáveis explicar o potencial futuro.

O modelo de regressão linear, segundo Krajewski, Ritzman e Malhotra (2009), é um dos modelos causais mais conhecidos e utilizados, que consiste de uma variável chamada de dependente estar relacionada a uma ou mais variáveis independentes por uma equação linear. Pode-se dizer em uma linguagem técnica que a linha de regressão minimiza os desvios quadrados dos dados reais. Para se obter o cálculo da equação da reta basta aplicar a seguinte equação:

$$y = a + bx$$

Na equação, “y” refere-se a variável dependente e “x” a variável independente. O “a” representa interseção da linha no eixo y e b inclinação da linha. Esta fórmula estabelece a equação que identifica o efeito da variável de previsão (variável independente) sobre a demanda do produto em análise (variável dependente), isso porque se busca prever a demanda de determinado item com base na previsão de outra variável que tenha relação com tal item. Em outras palavras, tem por objetivo encontrar uma equação linear de previsão de modo que a soma dos quadrados dos erros de previsão (beta) seja a mínima possível. Abaixo observa-se um modelo onde uma reta de regressão é formada no sistema cartesiano, note que esta reta é calculada através dos pontos constituídos pelas variáveis dependentes e independentes que estão em análise através do método dos mínimos quadrados (Figura 1):



Fonte: Tubino (2007)

Além disso, as equações para obtenção dos coeficientes a e b que constituem os modelos de regressão linear, conforme descrevem Gaither e Frazier (2006), estão descritas abaixo (Figura 2):

Figura 2
$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$
 regressão
$$a = \frac{\sum x^2 \sum y - \sum x \sum xy}{n\sum x^2 - (\sum x)^2}$$

Fonte: Autores (2019)

De acordo com as figuras acima apresentadas, o cálculo dos coeficientes a e b tem como finalidade minimizar a soma

dos desvios quadrados dos dados reais da linha do gráfico. Para exemplificar o que foi dito, será calculado o modelo de regressão linear para o conjunto de dados analisado durante os resultados deste trabalho que é composto pelas notas do ENADE na cidade de Garanhuns (Quadro 1 e Figura 3):

Quadro 1. Exemplo de aplicação

Garanhuns (X)	Conceito Enade (Y)	X ²	Y ²	XY
2011	2,879	4044121	8,288641	5789,669
2014	2,715	4056196	7,371225	5468,01
2017	2,438	4068289	5,943844	4917,446
Σ=6042	Σ=8,032	Σ=12168606	Σ=21,60371	Σ=16175,125

Fonte: Autores (2019).

Onde:

Figura 3. Coeficientes calculados a partir do quadro 1

$$b = \frac{(3 * 16175,125) - (6042 * 8,032)}{3 * 12168606 - (6042)^2} = \frac{-3,969}{54} = -0,0735$$

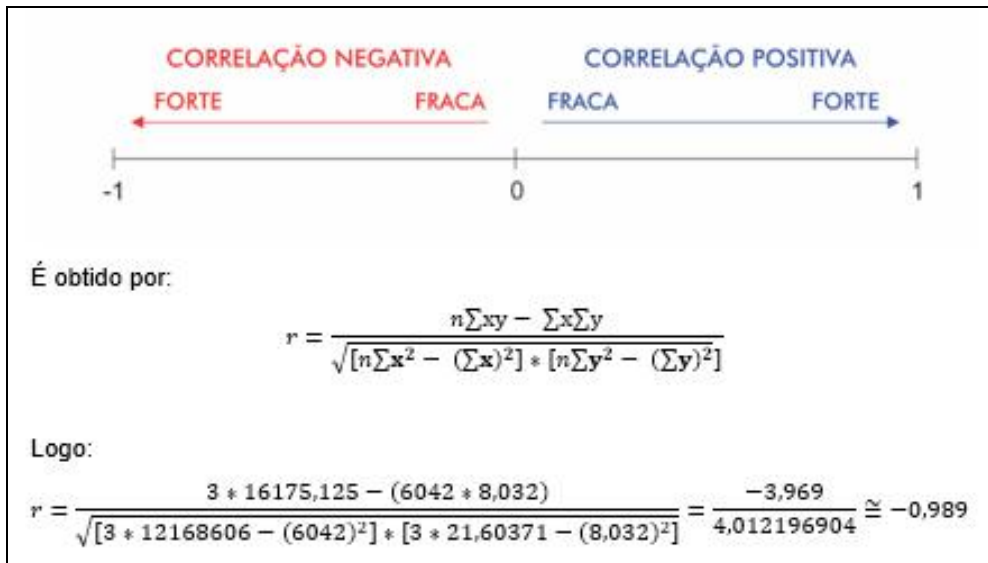
$$a = \frac{(12168606 * 8,032) - (6042 * 16175,125)}{3 * 12168606 - (6042)^2} = \frac{8138,142}{54} \cong 150,7$$

Fonte: Autores (2019).

Como $y = a + bx$, pelas informações obtidas, temos que a equação da reta será $y = 150,7 - 0,0735x$, através dela foi calculado o valor das notas para o ano de 2020, sendo 2020 a variável x e a nota a variável y . Além da reta de regressão foram levados em conta outros fatores determinísticos que influenciam nos parâmetros do modelo, nesse caso, temos o Coeficiente de Correlação de Pearson, o Coeficiente de Determinação e a Análise dos Resíduos, veja como foi obtido cada um:

Coeficiente de Correlação de Pearson (r): o estudo da correlação tem por objetivo medir e avaliar o grau de relação existente entre duas ou mais variáveis na esperança de que qualquer relação encontrada possa ser usada no sentido de fazer estimativas ou predições de uma das variáveis particulares, ou seja, é o número que resume o grau de relacionamento entre a variável dependente e uma ou mais variáveis independentes. Esse grau de correlação varia em escala de -1 a 1 (Figura 4) e é obtido através das disposições dos pontos referentes às variáveis independentes e a variável dependente numa reta ou um plano, para exemplificação do cálculo do coeficiente foram usados os dados do quadro 1 (Figura 4):

Figura 4. Pearson aplicado nos valores do quadro 1.



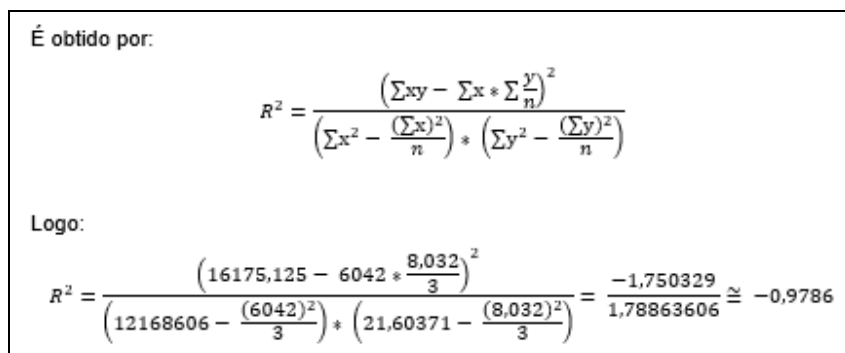
Fonte: Autores (2019).

Como o valor de r resultante foi de aproximadamente -0,99, assumimos uma correlação negativa muito forte, onde a medida que os valores da variável independente aumentam os valores da variável dependente diminuem.

Coefficiente de Determinação (R²): este coeficiente mensura a relação existente entre a variável dependente e as variáveis independentes. Indicando quantos por cento a variação explicada pela regressão representa da variação total (população). Quando:

- R² = 1: Todos os pontos observados se situam exatamente sobre a reta de regressão (ajuste perfeito), ou seja, as variações de y são 100% explicadas pela variação dos x's através da função especificada, não havendo desvios em torno da função estimada.
- R² = 0: Conclui-se que as variações de y são exclusivamente aleatórias e a introdução das variáveis x's no modelo não incorporará informação alguma sobre as variações de y. Para exemplificar o cálculo do R² aplicamos os valores registrados no quadro 1 (Figura 5):

Figura 5. Coeficiente de determinação (R²) do quadro 1.



Fonte: Autores (2019).

O coeficiente de determinação obtido indica que aproximadamente 98% das variáveis dependentes são explicadas pelo modelo de regressão gerado.

Análise de Resíduos: A análise de resíduos desempenha papel fundamental na avaliação do ajuste de um modelo de regressão linear simples, investiga a adequação do modelo quanto às suposições básicas, bem como normalidade, independência dos erros, homocedasticidade, relação linear de X e Y e falta de ajuste. Além dos testes de significância e adequação, a análise de resíduos vem complementar o elenco de procedimentos que devem ser realizados após o ajuste de qualquer modelo (Messeti, 2013). Os resíduos dos valores visualizados no quadro 1 foram descritos abaixo (Quadro 2):

Quadro 2. Análise de resíduos do quadro 1

Garanhuns			
Observação	Y previsto (através da reta de regressão)	Resíduos (diferença entre o y calculado e o y observado)	
1	2,897833	-0,01883	
2	2,677333	0,037667	
3	2,456833	-0,01883	

Fonte: Autores (2019).

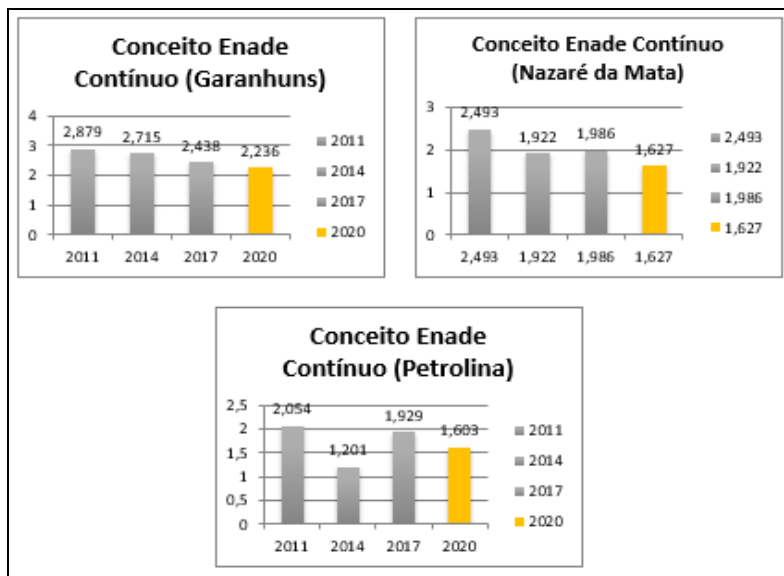
Observe que os valores residuais são pequenos, o que ajuda a entender a qualidade do modelo, além de possuir variância constante, sem outliers ou pontos influentes.

3. Resultados e Discussão

O Exame Nacional de Desempenho de Estudantes (Enade), como já se sabe, é composto pela média ponderada das notas de Formação Geral (FG) e Componentes Específicos (CE). As notas observadas foram organizadas por cidade (Garanhuns, Nazaré da Mata e Petrolina), Categorias (Enade, Formação Geral ou Componentes Específicos) e os anos em que foram obtidas as notas, buscando, através da regressão linear simples, prever as notas que serão obtidas pelos cursos de Licenciatura em Matemática da Universidade de Pernambuco no ano de 2020 (ainda não divulgadas).

3.1 ENADE: Os gráficos a seguir mostram as notas do Enade obtidas pelos alunos nos anos de 2011, 2014 e 2017, além das notas esperadas para o ano de 2020:

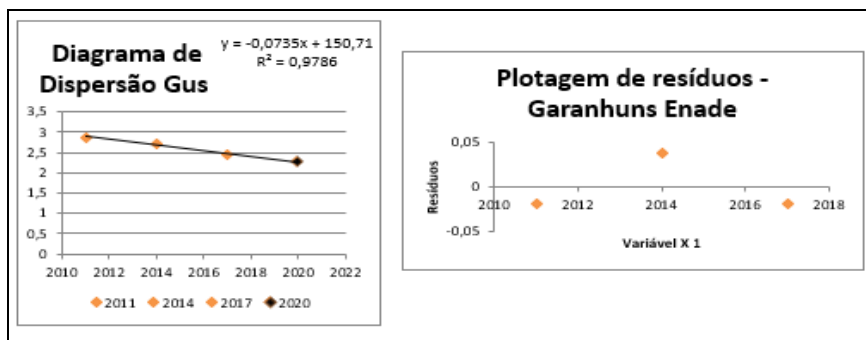
Conjunto 1 de Gráficos. Notas do Enade nos anos de 2011, 2014 e 2017 e notas previstas para ano de 2020 de acordo com o modelo de regressão nas cidades de Garanhuns, Nazaré da Mata e Petrolina .



Fonte: Autores (2019).

As informações descritas neste conjunto 1 de gráficos foram utilizadas para realização dos cálculos dos modelos de regressão linear para cada cidade nas notas gerais do ENADE. No conjunto 2 de gráficos estão descritas as informações da cidade de Garanhuns:

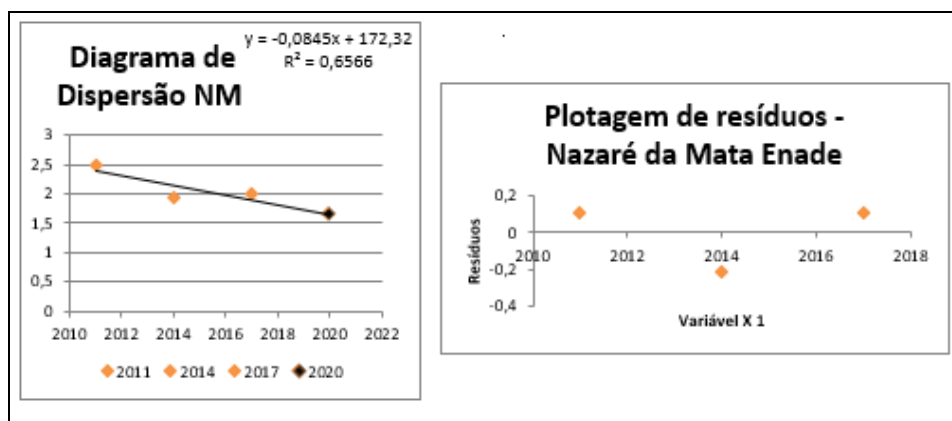
Conjunto 2 de Gráficos. Diagrama de dispersão e plotagem de resíduos das notas do Enade na cidade de Garanhuns.



Fonte: Autores (2019).

Notamos que na cidade de Garanhuns em todas as variáveis enquanto X cresce Y decresce, logo, há uma relação linear negativa onde o Coeficiente de Correlação de Pearson (r) é igual a $-0,98923$ (Conjunto 2 de Gráficos), caracterizando uma correlação muito forte. A reta de regressão se dá pela equação $y = -0,0735x + 150,71$, onde o Coeficiente de Determinação (R^2) é igual a $0,9786$, ou seja, a reta explica quase que 100% os valores observados. Os resíduos entre as notas observadas e as previstas são quase iguais, variando entre $0,05$ e $-0,05$. Os dados acima sugerem que a relação entre as variáveis analisadas é forte, como consequência espera-se que a nota do Enade em Garanhuns em 2020 seja ainda menor que nos anos anteriores. No conjunto 3 de gráficos estão descritas as informações da cidade de Nazaré da Mata:

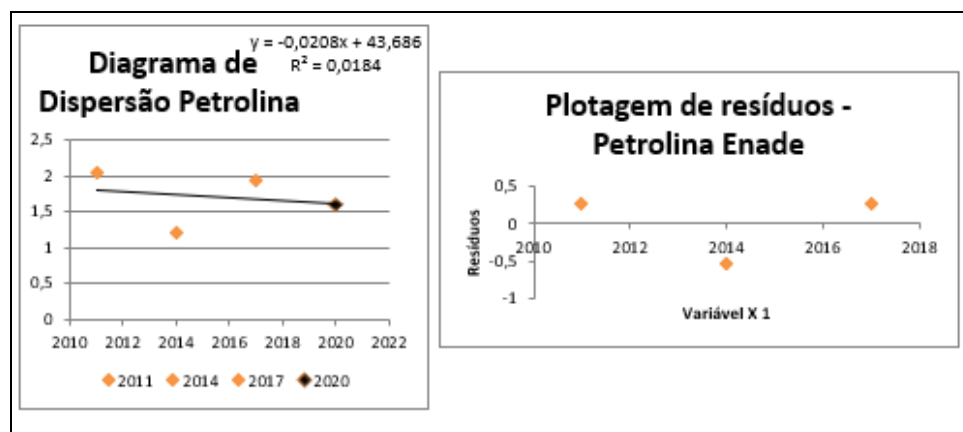
Conjunto 3 de Gráficos. Diagrama de dispersão e plotagem de resíduos das notas do Enade na cidade de Nazaré



Fonte: Autores (2019).

Em Nazaré da Mata temos que, em média, X cresce e Y decresce, r é igual a $-0,81034$, indicando uma correlação forte, a equação da reta de regressão é $y = -0,0845x + 172,32$ e $R^2 = 0,6566$, sendo então aproximadamente 65% o quanto o modelo de regressão consegue explicar os valores observados. Quanto aos resíduos, não variam muito, não são maiores que 0,2 ou menores que 0,3, por isso, a nota predita para 2020 pode ser levada em consideração, ou seja, os dados sugerem que no ano de 2020 a nota será novamente a menor no conjunto de dados. No conjunto 4 de gráficos estão descritas as informações da cidade de Petrolina:

Conjunto 4 de Gráficos. Diagrama de dispersão e plotagem de resíduos das notas do Enade na cidade de Petrolina.

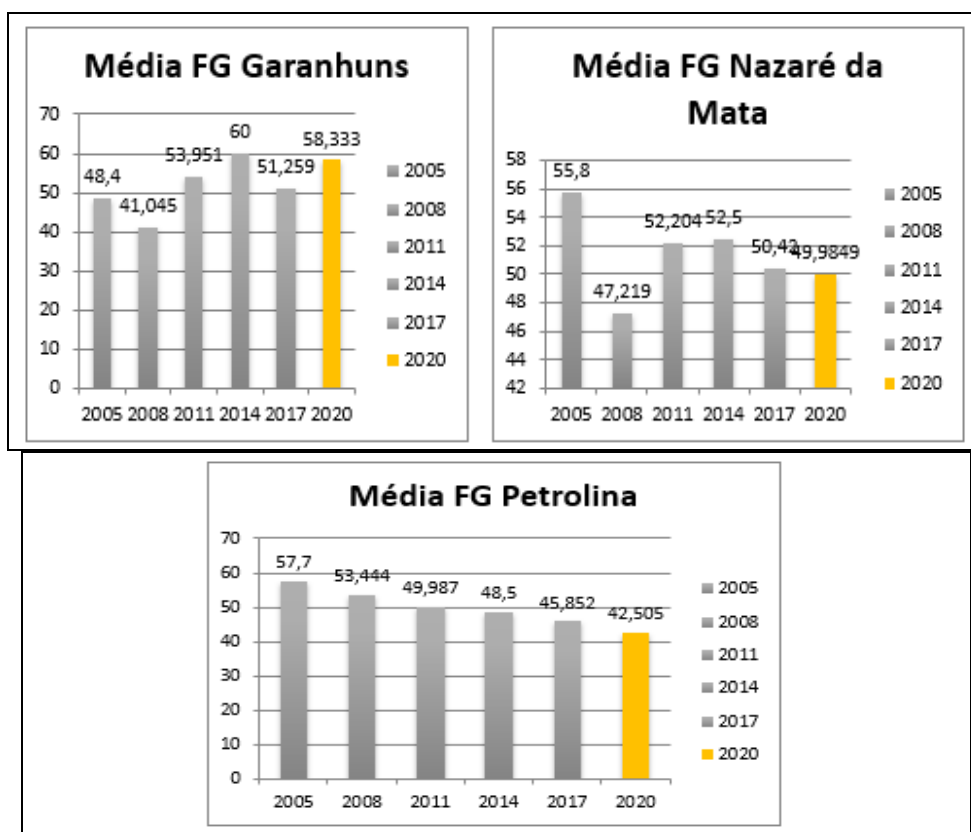


Fonte: Autores (2019).

Na cidade de Petrolina, em média, enquanto X cresce Y diminui, ao contrário dos outros campi, temos que a correlação entre as variáveis que foram analisadas é quase que desprezível, isso ocorre porque r é igual a $-0,13568$. A Reta de Regressão é dada pela equação $y = -0,0208x + 43,686$ e o Coeficiente de Determinação é 0,0184, também baixíssimo, significando que a reta não chega a explicar nem 1% dos valores observados, essa situação nos mostra que praticamente não há uma relação entre as variáveis e, portanto, na nota predita para o ano de 2020 existe grandes chances que ocorra um resíduo maior da nota observada para a nota calculada.

3.2 Formação Geral (FG): Das notas de Formação Geral foram pegos os anos de 2005, 2008, 2011, 2015 e 2017, nos gráficos constam as notas observadas nesses anos mais a nota predita para o ano de 2020:

Conjunto 5 de Gráficos: Notas de Formação Geral nos anos de 2005, 2008, 2011, 2015 e 2017 e notas previstas para ano de 2020 de acordo com o modelo de regressão nas cidades de Garanhuns, Nazaré da Mata e Petrolina .



Fonte: Autores (2019).

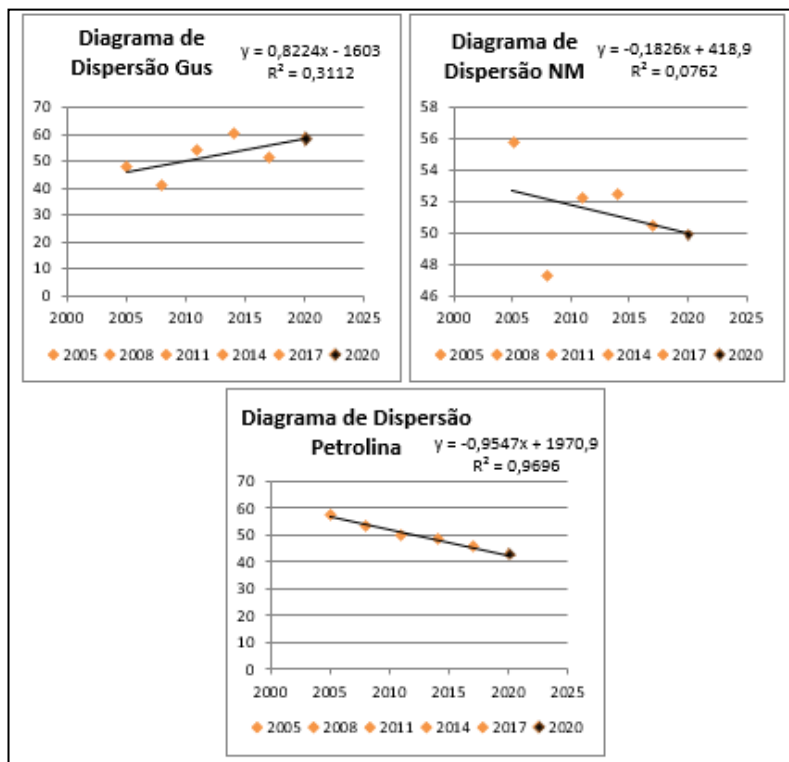
As informações descritas neste conjunto 5 de gráficos foram utilizadas para realização dos cálculos dos modelos de regressão linear para cada cidade nas notas de Formação Geral, os dados que estão descritos no conjunto 6 de gráficos abaixo estão também resumidos no quadro 3:

Quadro 3. Dados da regressão para as cidades nas notas de Formação Geral

	Garanhuns:	Nazaré da Mata	Petrolina
r	0,557856	-0,27604	-0,9847
Reta	$y = 0,8224x - 1603$	$y = -0,1826x + 418,9$	$y = -0,9547x + 1970,9$
R ²	0,3112	0,0762	0,9696

Fonte: Autores (2019)

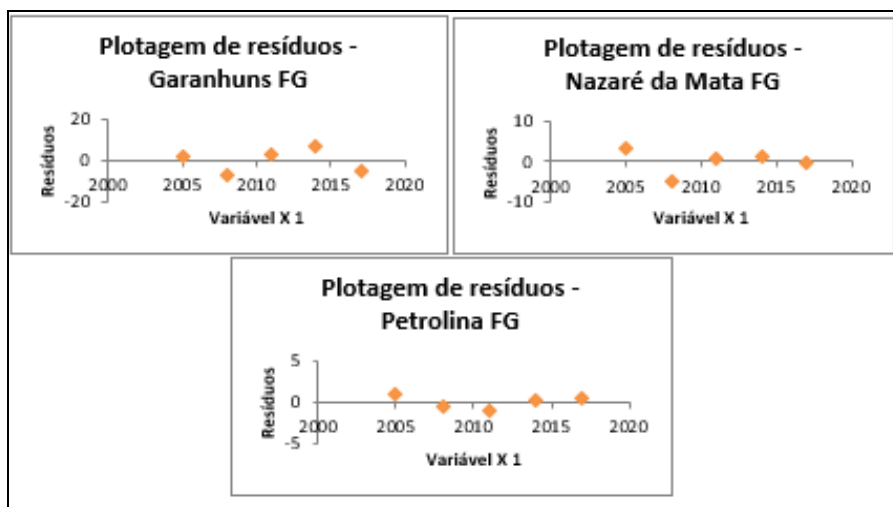
Conjunto 6 de Gráficos. Diagramas de dispersão de todas as cidades nas notas de Formação Geral:



Fonte: Autores (2019).

Observe as informações, nas notas de Formação Geral (FG) percebemos que onde há correlação mais forte é em Petrolina, classificada em muito forte, seguida de Garanhuns, com uma correlação moderada e Nazaré da Mata, que possui correlação desprezível, isso reflete nos Coeficientes de Determinação, onde a Reta de regressão, em Nazaré da Mata, explica menos de 1% dos dados observados, logo após vem Garanhuns, com 31% e Petrolina com 96%. Quanto a análise de resíduos (conjunto 7 de gráficos):

Conjunto 7 de Gráficos. Gráficos de plotagem de resíduos de todas as cidades nas notas de FG



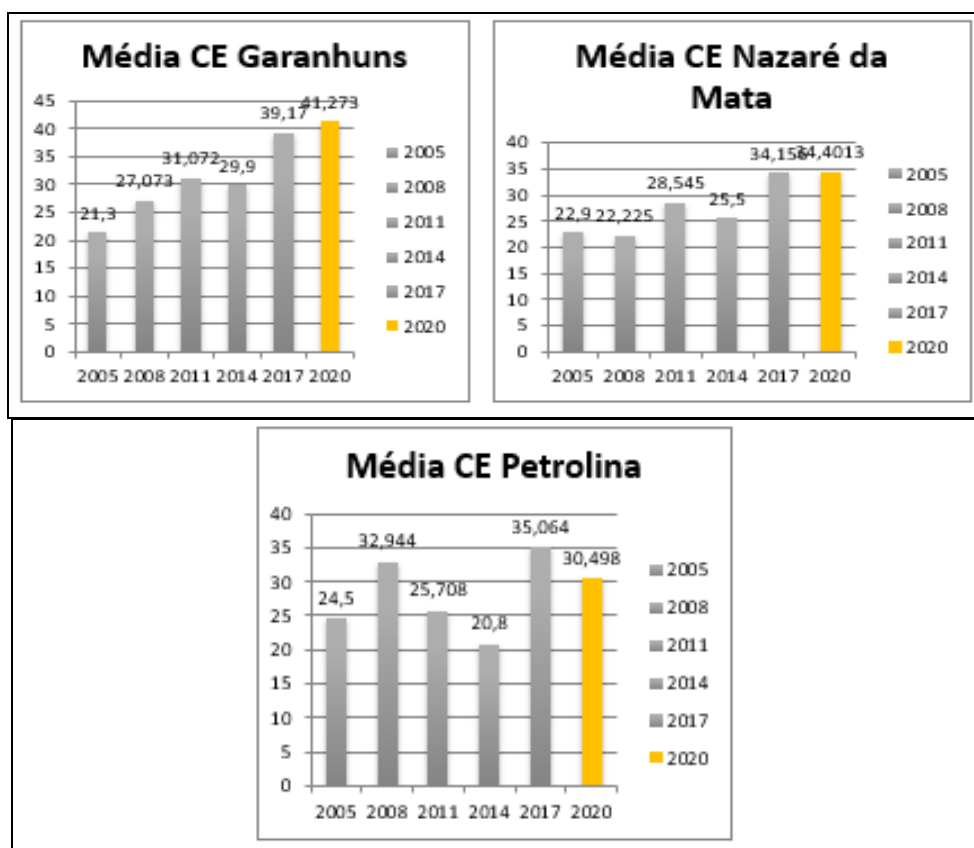
Fonte: Autores (2019).

Note que os resíduos, em Garanhuns, variam consideravelmente entre si. Nazaré da Mata, por sua vez, embora tenha uma correlação desprezível entre as variáveis, não possui resíduos muito distintos, já Petrolina naturalmente possui os menores resíduos, com a maioria chegando a quase zero.

Com base nas observações percebe-se que a análise da relação linear entre as duas variáveis (ano e nota FG) nas três cidades nos sugere que Petrolina tem uma relação direta de causa e efeito entre as variáveis, portanto, é esperado que sua menor nota em Formação Geral seja no ano de 2020, em Nazaré não existe uma relação linear considerável, sendo a nota predita mais incerta, já Garanhuns, caso o valor obtido for próximo do que foi predito, alcançará sua segunda maior nota se comparada aos anos anteriores.

3.3 Componentes Específicos (CE): Por fim, das notas de Componentes específicos também foram pegos os anos de 2005, 2008, 2011, 2015 e 2017, nos gráficos constam as notas observadas nesses anos mais a nota predita para o ano de 2020:

Conjunto 8 de Gráficos. Notas de Componentes Específicos nos anos de 2005, 2008, 2011, 2015 e 2017 e notas previstas para ano de 2020 de acordo com o modelo de regressão nas Cidades de Garanhuns, Nazaré da Mata e Petrolina.



Fonte: Autores (2019).

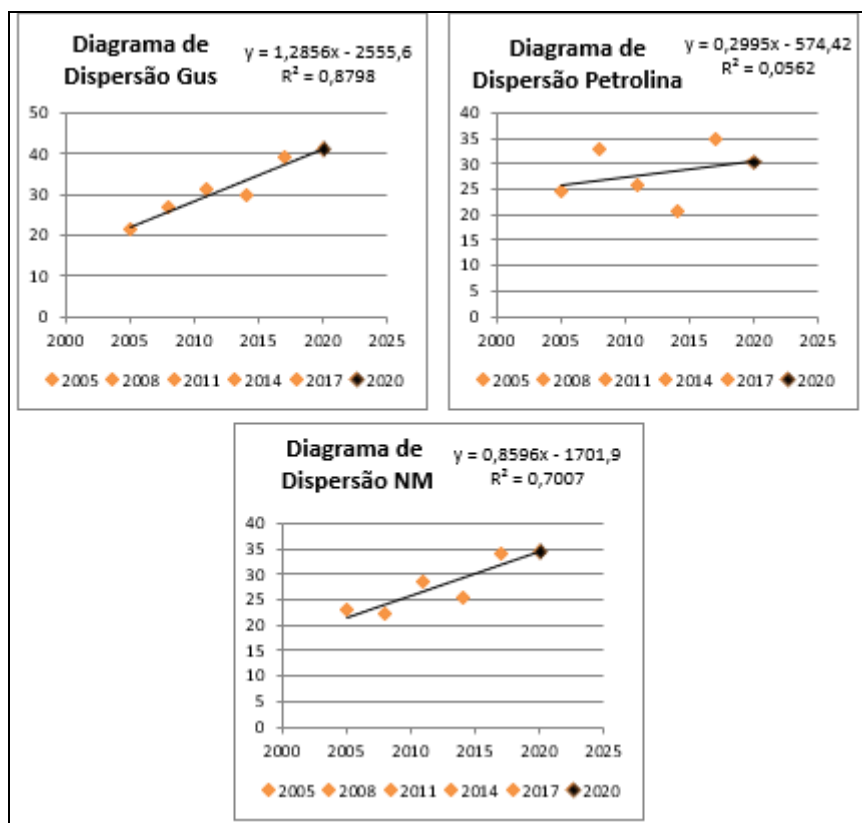
As informações descritas neste conjunto 8 de gráficos foram utilizadas para realização dos cálculos dos modelos de regressão linear para cada cidade nas notas de Componentes Específicos, os dados que estão descritos no conjunto 9 de gráficos abaixo estão também resumidos no Quadro 4:

Quadro 4. Dados da regressão para as cidades nas notas de Componentes Específicos

	Garanhuns:	Nazaré da Mata	Petrolina
r	0,937972	0,837101	0,237167
Reta	$y = 1,2856x - 2555,6$	$y = 0,8596x - 1701,9$	$y = 0,2995x - 574,42$
R ²	0,8798	0,7007	0,0562

Fonte: Autores (2019).

Conjunto 9 de Gráficos. Diagramas de dispersão de todas as cidades nas notas de Componentes Específicos:

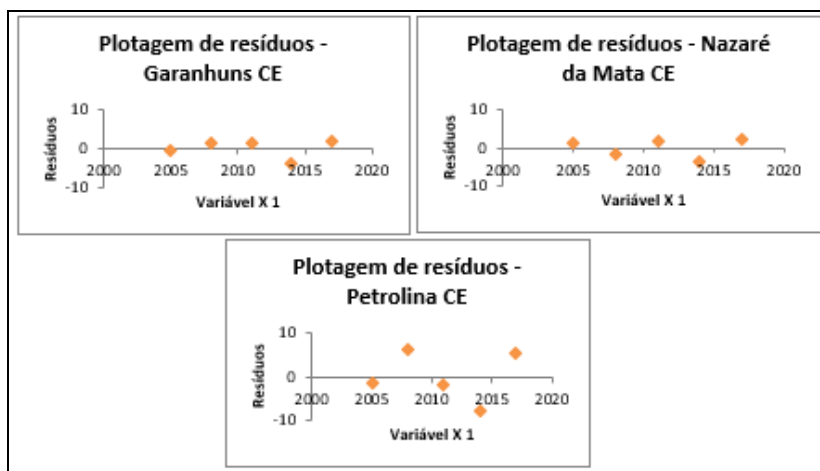


Fonte: Autores (2019).

Na cidade de Garanhuns, ao analisarmos as notas, percebemos que, em média, se X cresce Y cresce, com exceção de uma ocasião entre os anos de 2011 e 2014, isso nos deixa com uma correlação linear positiva onde o Coeficiente de correlação de Pearson tem valor próximo de um, o que acarreta numa correlação forte, uma situação parecida acontece com Nazaré da Mata, que obtém uma correlação forte, já Petrolina apresenta correlação desprezível. Quanto ao Coeficiente de Determinação (R²), pela sua análise nota-se que a reta de regressão explica menos de 1% dos valores observados na cidade de Petrolina, em Nazaré da Mata esse valor é de aproximadamente 70% e em Garanhuns 87%.

Ainda pela análise do conjunto de gráficos 9 vimos que as notas preditas para o ano de 2020 em Garanhuns e Nazaré possuem grande relação de causa e efeito, com as maiores já alcançadas por ambas, o que definitivamente não ocorre em Petrolina, que, além de seus valores terem relação linear quase nula, apenas alcançaria sua terceira maior nota se ocorresse o que foi predito. Observe agora os resíduos (conjunto 10 de gráficos):

Conjunto 10 de Gráficos. Gráficos de plotagem de resíduos de todas as cidades nas notas de CE



Fonte: Autores (2019).

Tanto os resíduos de Garanhuns quanto os de Nazaré não possuem grande variância entre eles, em ambas as cidades o menor valor não passa de -4 e o maior valor não passa de 3, tendo, além disso, vários resíduos com valores quase iguais. Petrolina, por sua vez, possui uma variação bem maior, com resíduos entre -8 e 7 e bastante distintos entre si.

4. Considerações Finais

Com a realização deste trabalho, percebeu-se que os dados obtidos a partir da aplicação do método proposto nos revelaram o valor das notas preditas para os anos de 2020, além disso, através do Coeficiente de Correlação Linear foi possível quantificar as correlações como fortes ou fracas, através do Coeficiente de Determinação explicou-se a proporção da variabilidade dos valores preditos para y que é explicada pelo modelo de regressão linear simples em cada caso, e, por fim, através da Análise de Resíduos confirmaram-se as suposições básicas de cada modelo.

Ao final do estudo constatou-se, quanto às notas do Enade, que as cidades que obtiveram boa relação entre as variáveis estudadas têm como supostas notas para o ano de 2020 um valor inferior a todos os anteriores, ou seja, a nota tende a decrescer a cada ano, o mesmo acontece com a nota de Formação Geral em Petrolina. A nota de Formação Geral em Garanhuns, que tem correlação moderada, tende a crescer e pressupõe chegar a sua segunda maior nota no ano de 2020.

Quanto às notas em Componentes Específicos, as duas cidades que possuem boa correlação alcançam suas maiores notas no ano de 2020, ou seja, a nota de CE tende a crescer ao longo dos anos. Por fim, das nove análises feitas, as três que obtiveram correlação fraca ou insignificante (Enade de Petrolina, FG de Nazaré da Mata e CE de Petrolina) têm, naturalmente, uma variação grande em suas notas, sem um padrão específico, consequentemente, as notas preditas para o ano de 2020 têm a mesma característica.

Para trabalhos futuros, sugere-se utilizar a regressão linear simples para uma análise do ENADE em algum município específico do estado em outros cursos de licenciatura ou Bacharelado da Universidade de Pernambuco.

Referências

- Barros, J. L. D. C., Campos, M. Z. D., Teixeira, D. D. C., & Cabral, B. G. D. A. T. (2020). Reflexões sobre o nível de conhecimentos específicos dos estudantes de licenciatura em Educação Física no Enade 2014. *Revista Brasileira de Estudos Pedagógicos*, 101, 99-119.
- Beltrão, K. I., & Mandarino, M. C. F. (2014). Evidências do ENADE-mudanças no perfil do matemático graduado. *Ensaio: Avaliação e Políticas Públicas em Educação*, 22, 733-753.

Da Silva Dias, J., de Magalhães Porto, C., & Nunes, A. K. F. (2016). Formação geral e conhecimento específico na prova Enade. *Encontro Internacional de Formação de Professores e Fórum Permanente de Inovação Educacional*, 9(9).

De Almeida, D. A., Almeida, S. P. N. de C. e., & Amorim, M. M. T. (2021). Profile of graduates in mathematics degree: an analysis from enade's data (2005-2017). In *SciELO Preprints*. <https://doi.org/10.1590/SciELOPreprints.2561>.

De Melo, A. M. (2009). Análise de Correlação e Regressão Linear Simples: Contabilometria Aplicada em Indicadores Econômico-Financeiros de 2009 das Empresas de Capital Aberto do Seguimento de Construção Civil Integrantes do Novo Mercado. *Congresso UFSC de Controladoria e Finanças e Iniciação Científica em Contabilidade*, Brasil.

De Oliveira Filho, M. L. (2002). A utilização da regressão linear como ferramenta estratégica para a projeção dos custos produção. In *Anais do Congresso Brasileiro de Custos-ABC*.

Do Amaral, L. S., dos Santos, A. L. P., de Figueiredo, M. P. S., de Almeida Ferreira, D. S., Silva, J. E., dos Santos, H. C. T., & Moreira, G. R. (2020). Interiorização do Covid-19: Uma análise da evolução dos casos/10 mil habitantes em municípios da Microrregião de Garanhuns no Estado de Pernambuco, através de modelos de Regressão não linear. *Research, Society and Development*, 9(9), e293996582-e293996582.

Morettin, P. A., & Bussab, W. O. (2017). *Estatística básica*. Saraiva Educação SA.

Gaither, N., & Frazier, G. (2001). *Administração da produção e operações*. Pioneira Thomson Learning.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2020). *Conceito Enade*. Retrieved from <http://portal.inep.gov.br/conceito-enade>.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (2020). *Exame Nacional de Desempenho dos Estudantes (Enade)*. Retrieved from <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enade>.

Krajewski, L. J.; Ritzman, L. P.; Malhotra, M. (2009). *Administração da produção e operações*. 8. ed. São Paulo: Pearson Prentice Hall.

Lima, B. A. T., & dos Santos, M. B. H. (2021). Percepção dos discentes sobre a formação na universidade através do questionário do estudante do ENADE. *Research, Society and Development*, 10(3), e19510313150-e19510313150.

Malhotra, N. K., Rocha, I., Laudisio, M. C., Altheman, É., Borges, F. M., & Taylor, R. B. (2005). *Introdução à pesquisa de marketing*.

Medeiros, F. S. B., & Bianchi, R. C. (2009). A aplicação do método regressão linear simples na demanda de produtos sazonais: um estudo de caso. *Disciplinarum Scientia| Sociais Aplicadas*, 5(1), 35-53.

Mesetti, Ana Verginia Libos. (2013). Análise de correlação e regressão. Curso de especialização "Lato Sensu" em estatística. *Universidade Estadual de Londrina, Londrina*.

Ramos, P. N. F., da Silveira, O. R., & de Souza Maia, J. C. (2022). Determinação da análise de regressão linear simples para explicar a influência dos atributos físicos do solo na produção do algodão. *Research, Society and Development*, 11(8), e28411830591-e28411830591.

Smith, M. M. *ENADE Comentado: Componente: Formação Geral: edições 2006, 2007, 2008*. EDIPUCRS.

Triola, M. F. (2008). Introdução à estatística. In *Introdução à estatística* (pp. xxvi-310).

Tubino, D. F. (2007). *Planejamento e controle da produção: teoria e prática*. Atlas.