



Eixo temático 4: Gestão e Políticas Educacionais no contexto das TDIC

RISCO DE DESCUMPRIMENTO DE META DO PLANO NACIONAL DE EDUCAÇÃO

RISK OF NON-COMPLIANCE WITH THE NATIONAL EDUCATION PLAN GOAL

- **Renata da Paixão e Silva** (ESALQ/USP - prof.renatapsilva@gmail.com)
- **Paulo Fernando do Nascimento Afonso** (PECEGE ESALQ/USP - afonso@conectcor.com.br)

Resumo:

A educação brasileira é uma ação desafiadora, tendo em vista as dimensões do país e a diversidade cultural, social e econômica. Neste contexto, percebeu-se que projetos e planos nacionais de educação precisavam de um estudo muito apurado para serem desenvolvidos de forma bem delineada desde o levantamento de necessidades e definição de metas, até a gestão e acompanhamento. Dessa forma, foi proposto a elaboração de um trabalho que avaliasse os riscos do descumprimento de metas do Plano Nacional de Educação, partindo de um indicador relacionado à jovens concluintes do ensino fundamental, com intuito de verificar se a meta relacionada a tal indicador seria atingida no tempo definido. Além disso, o presente trabalho também pretendeu avaliar outros indicadores, neste caso relacionados aos professores, com a finalidade de buscar um modelo explicativo. O estudo inicialmente utilizou séries temporais, elaborando tendências lineares para apuração dos resultados, em comparação com os dados reais de acompanhamento do plano de educação supracitado. Em seguida, utilizou-se a regressão linear para avaliar possíveis correlações entre dois indicadores associados aos professores. Após verificar-se a correlação com um indicador ligado aos professores, a saber, formação com pós-graduação, o estudo culminou na construção e posterior ajuste de uma equação que expressasse e explicasse a situação apresentada.

Palavras-chave: Ciência de Dados; Educação; Plano Nacional de Educação; Regressão Linear; Séries Temporais.

Abstract:

Brazilian education is a challenging action, given the country's dimensions and cultural, social, and economic diversity. In this context, it was realized that national education projects and plans needed a very detailed study to be developed in a well-outlined way, from needs assessment and goal definition to management and monitoring. Therefore, it was proposed to develop a work that would evaluate the risks of non-compliance with the goals of the National Education Plan, starting from an indicator related to young people who complete elementary school, in order to verify if the goal related to such an indicator would be achieved within the defined time in addition, this work also aimed to evaluate other indicators, in this case related to teachers, in order to seek an explanatory model. The study initially used time series, elaborating linear trends to determine the results, in comparison with the actual data of the education plan. Next, linear regression was used to evaluate possible correlations between two indicators associated with teachers. After verifying the correlation with an indicator related to teachers, namely, postgraduate education, the study culminated in the construction and subsequent adjustment of an equation that expressed and explained the situation presented.

Keywords: Data Science; Education; National Education Plan; Linear Regression; Time Series.

1. Introdução

A educação no Brasil sempre foi um desafio, principalmente nas percepções pedagógica, estrutural, cultural e financeira. E sendo a educação básica um bem público, ela precisa ser estendida a toda a sociedade com qualidade.



Com o objetivo principal de garantir a educação com qualidade a todos, foi iniciado em 2014, o Plano Nacional de Educação (PNE) que é uma lei onde são definidas metas e políticas públicas que precisam ser alcançadas em todas as esferas brasileiras: federal, estadual e municipal, em um período de vigência de 10 anos.

No entanto, se tornou um grande desafio assegurar a conclusão do Ensino Fundamental (EF) na idade recomendada, tendo em vista que apenas 78% dos adolescentes aos 16 anos conseguem concluí-lo (INEP, 2020).

Dentre diversas diretrizes para erradicação do analfabetismo, universalização do atendimento escolar, melhoria da qualidade da educação, existem metas bem definidas no Plano Nacional de Educação, totalizando 20 metas compostas por alguns indicadores.

Desse modo, o presente trabalho propõe-se a abordar a meta 2 do PNE, mais especificamente, o indicador 2B relacionado à população de 16 anos com o ensino fundamental concluído.

A meta 2 objetiva universalizar o ensino fundamental de 9 (nove) anos de duração, para toda a população de 6 (seis) a 14 (quatorze) anos e garantir que pelo menos 95% dos alunos concluam essa etapa na idade recomendada, até o último ano de vigência do PNE (BRASIL, 2014).

E, conforme o relatório do 3º ciclo de monitoramento das metas sugere, a meta 2 não será alcançada, sendo necessário uma melhora dos indicadores, pois as desigualdades regionais e sociais, são expressivas (INEP, 2020).

Partindo desta problemática, pretende-se analisar os dados regionais, com intuito de propor um modelo de melhor ajuste ao objetivo do plano, analisando os possíveis riscos de descumprimento do mesmo.

Utilizando técnicas estatísticas amplamente difundidas, cogita-se identificar nesta pesquisa, a situação atual das metas educacionais, bem como, estabelecer base para uma trajetória explicativa sobre os riscos apresentados, além de estimular novas soluções em favor da educação.

2. Material e métodos

O estudo se desenvolve em duas etapas, onde inicialmente, busca-se obter os mesmos resultados do relatório do 3º ciclo de monitoramento do PNE, principalmente como forma de validação do direcionamento do estudo. E na segunda etapa, segue-se comparando o indicador pesquisado, com outros indicadores para avaliar possíveis correlações.

2.1. Indicadores da meta 2

A meta 2 do PNE é composta por dois indicadores, denominados 2A e 2B, sendo o indicador 2A, definido pela taxa de escolarização líquida, que é o percentual da população de 6 a 14 anos que frequentam ou que já concluíram o ensino fundamental. E o indicador 2B é definido pelo percentual da população de 16 anos com pelo menos o ensino fundamental concluído.

O indicador 2B representa a proporção da população de 16 anos de idade que concluiu pelo menos o ensino fundamental em relação a população total de 16 anos de idade, e abrange todos os estados brasileiros.

2.2. Indicadores da meta 16

A meta 16 do PNE também é composta por dois indicadores, denominados 16A e 16B, sendo



o indicador 16A, definido pelo percentual de professores da educação básica com pós-graduação lato sensu ou stricto sensu e o indicador 16B definido pelo percentual de professores da educação básica que realizaram cursos de formação continuada.

2.3. Fontes de informações

As bases de dados da pesquisa são provenientes do Painel de Monitoramento do PNE, que integra o Inep Data, que é o conjunto de painéis do Inep acessível à sociedade, o qual fornece dados abertos com objetivo de possibilitar o acesso de informações por gestores, pesquisadores, sociedade civil, imprensa e outros.

Neste estudo, utiliza-se os valores da tabela: Percentual da população de 16 anos com pelo menos o ensino fundamental concluído, por Brasil, grande região e unidade da Federação 2012-2020, referente ao indicador 2B, supracitado.

Também são usadas as bases de dados do Censo da Educação Básica, para os indicadores 16A e 16B, com as respectivas tabelas: Professores da educação básica com pós-graduação lato sensu ou stricto sensu – por Brasil, grande região e unidade de Federação e Professores da educação básica que realizaram cursos de formação continuada – por Brasil, grande região e unidade de Federação, ambas referente aos anos 2013/2016/2019/2020.

2.4. Métodos

Partindo da base de dados citada anteriormente, que conta com informações de 2012 a 2020, busca-se desenvolver uma análise de série temporal – que é uma forma de observar a variável de estudo, durante um tempo – tendo em vista o registro de fases regulares em anos. Sendo assim, a análise inicial, pesquisa padrões não aleatórios nas séries, averiguando o histórico dos dados nas fases, e com base neles, propõe-se valores futuros.

Portanto, com os dados da base referida, inicia-se uma análise de tendência linear, a fim de validar o método de estudo, principiando com os dados totais do país, em seguida, com os dados das regiões e por fim, com os dados dos estados brasileiros.

Após verificação e validação da série temporal, o estudo segue para uma investigação sobre possíveis correlações entre indicadores relacionados aos jovens concluintes do EF, com indicadores relacionados aos professores.

Nessa etapa, onde busca-se reconhecer variáveis explicativas, emprega-se a regressão linear, preliminarmente, aplicando testes nas amostras para a verificação de normalidade, normalizando as amostras e utilizando testes estatísticos para a verificação da existência do modelo e influência da variável explicadora. Para que, por fim, identifique-se uma equação que melhor se ajuste ao objetivo.

3. Resultados e discussão

Primeiramente, pode-se observar uma tendência crescente de jovens com EF concluído, iniciando com 68,6% em 2012 e finalizando com 82,4% em 2020, aparentemente não há flutuações regulares que configurem sazonalidade. Mas no período entre 2015 e 2018, constata-se uma suave flutuação variando próximo a 1% (Figura 1).

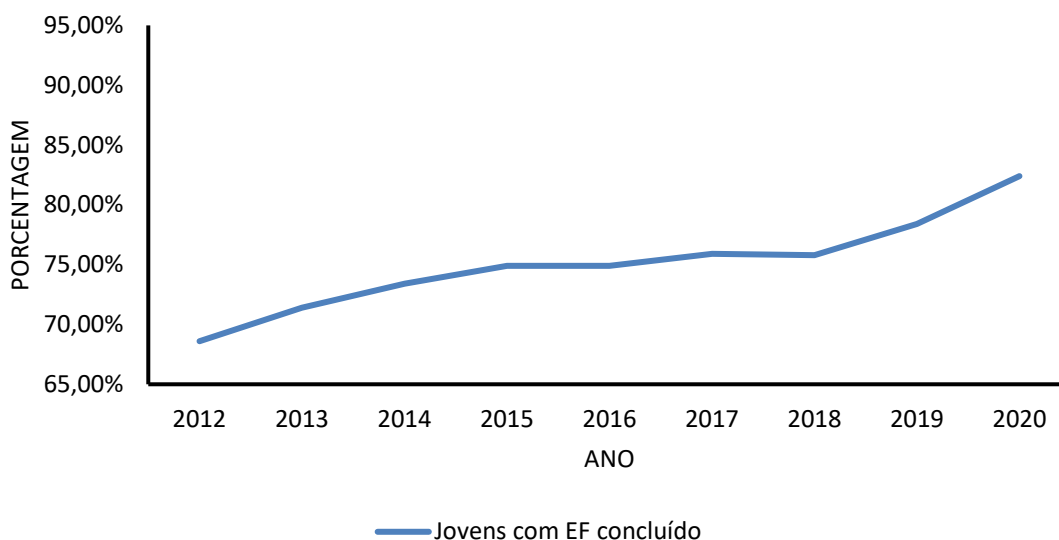


Figura 1. Porcentagem de jovens de 16 anos com EF concluído no Brasil.
Fonte: INEP (2020).

A análise se desenvolve obtendo uma série com tendência, de maneira semelhante ao procedimento de regressão linear, considerando a variável independente como o tempo, que descreve-se por períodos, onde 2012 refere-se ao período 1, seguindo até 2020 como o período 9. Sendo definida uma equação de tendência linear, conforme eq. (1):

$$T = b \times t + a \quad (1)$$

Onde T é o valor da tendência, b é o coeficiente angular da reta, t é o valor do tempo (em períodos) e a é o coeficiente linear da reta.

Os valores de a e b são constantes estimadas respectivamente a partir das funções INTERCEPÇÃO e INCLINAÇÃO da ferramenta MS Excel®, onde relaciona-se a porcentagem de jovens com ensino fundamental concluído, com a quantidade de observações existentes, neste caso 9 observações, referentes aos períodos registrados (2012 a 2020). Vale ressaltar, que b indica a tendência crescente, quando seu valor é positivo e decrescente, quando o valor é negativo.

Considerando os valores calculados, de intercepto (a) e inclinação (b), tem-se a equação de tendência, eq. (2):

$$T = 0,013667 \times t + 0,68244 \quad (2)$$

Propondo-se valores futuros através da tendência, a previsibilidade fica em torno de 86% no ano de 2024, sendo um valor bem abaixo da meta de 95%. Na (Figura 2), observa-se a tendência e projeção até o fim da PNE em 2024.

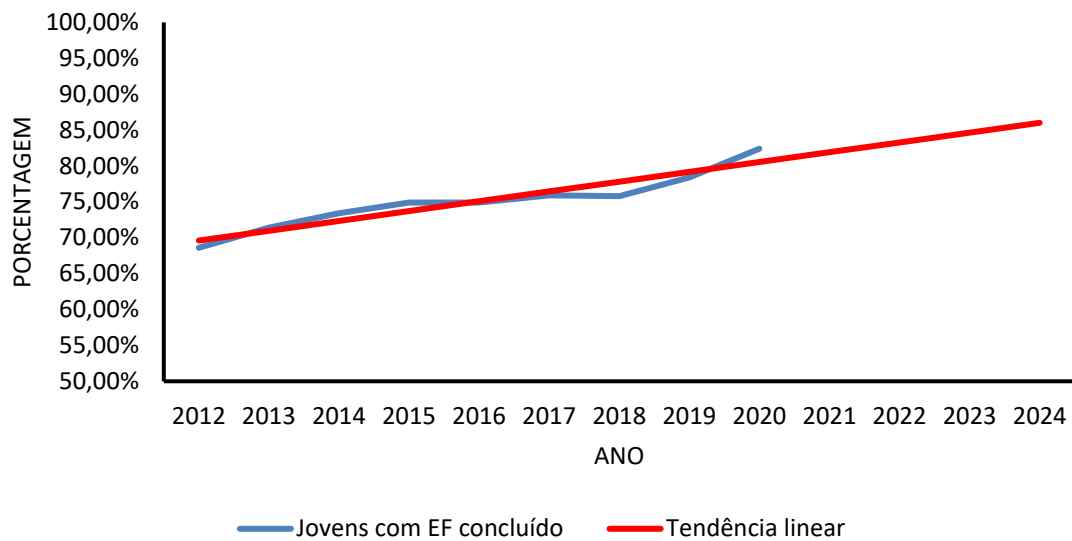


Figura 2. Tendência e projeção de jovens de 16 anos com EF concluído no Brasil.

Fonte: Autoria própria (Resultados originais da pesquisa).

No primeiro momento, a análise fora feita com percentuais de todo o Brasil, e em sequência, as análises seguiram para o nível estadual.

Os primeiros resultados obtidos, mostraram-se similares aos do relatório do 3º ciclo de monitoramento de metas do PNE. Indicando que a meta tende a não ser cumprida no tempo determinado.

Com base nisso, foram analisados em seguida, os estados brasileiros a fim de identificar padrões que possivelmente se associassem aos resultados encontrados anteriormente.

Na região Sul, o estado do Paraná destaca-se pelo crescimento percentual de jovens com o EF concluído, pois iniciou a série com 73,8%, finalizou com 90,2%.

Na região Nordeste, observa-se crescimento também, pois no ano de 2012 as porcentagens dos estados estavam em torno de 43,2% a 69%, e no ano de 2020 entre 58,8% e 91%, sendo o estado do Ceará, quem delineou todo o limite superior das porcentagens na região.

Em relação aos estados da região Norte, Rondônia, Amazonas e Pará, apresentaram tendência crescente, e no Acre, Roraima e Tocantins, as porcentagens de jovens com EF concluído apesar de crescentes, não tiveram um progresso expressivo, e especificamente no estado do Amapá (Figura 3) pode-se observar uma tendência decrescente e muitas oscilações, incluindo duas grandes quedas, uma em 2017 e outra em 2020.

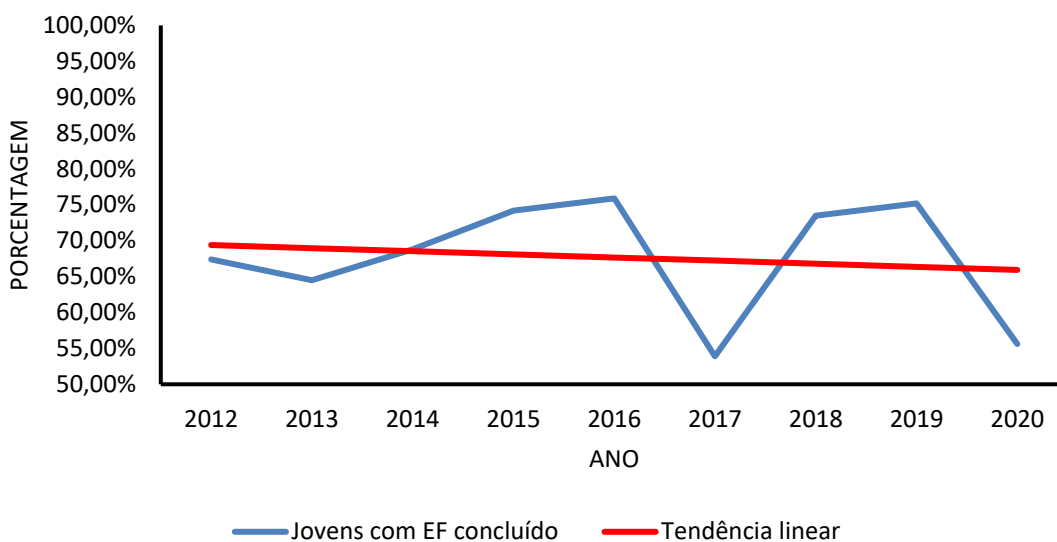


Figura 3. Porcentagem de jovens com EF concluído e tendência (Amapá).
Fonte: Autoria própria (Resultados originais da pesquisa).

Nas regiões Centro-Oeste, Sul e Sudeste, as séries permaneceram acima de 60% em todo o período observado. E especificamente, em São Paulo (Figura 4), a porcentagem de jovens com EF concluído chegou a 94,5% em 2020, tendo começado a série em 2012 com 86,7%, um valor discrepante em relação a maioria dos estados.

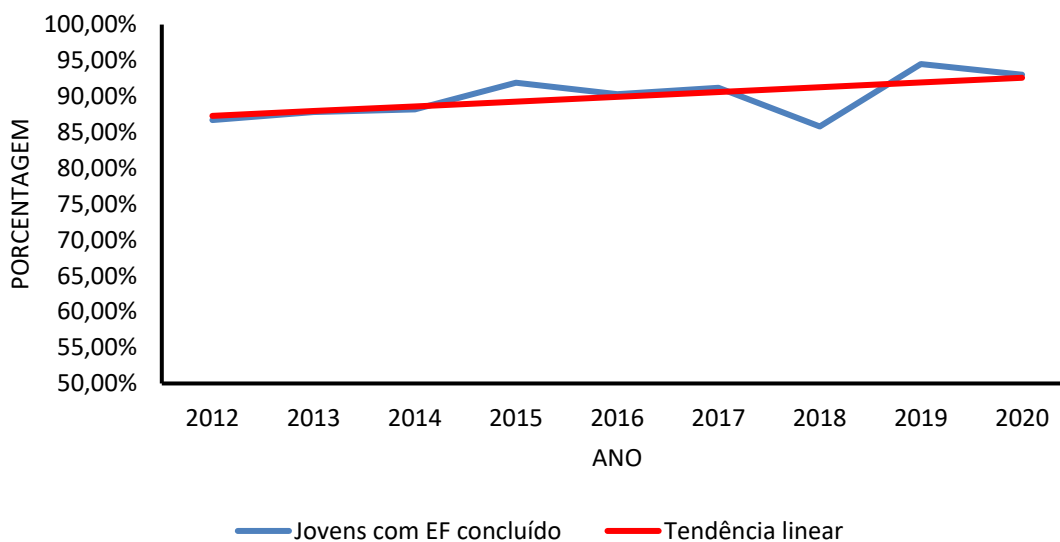


Figura 4. Porcentagem de jovens com EF concluído e tendência (São Paulo).
Fonte: Autoria própria (Resultados originais da pesquisa).

Considerando as visíveis variações observadas nas séries temporais dos estados brasileiros, houve a necessidade de avaliar a relação do indicador 2B em estudo, com outros indicadores classificados no atual PNE. Sendo assim, explora-se a influência de professores da educação básica com pós-graduação e/ou formação continuada nos resultados relacionados ao número de jovens concluintes do EF.

Para isso, parte-se da meta 16: Pós-Graduação de Professores e Formação Continuada dos



Profissionais da Educação Básica, que tem por objetivo formar em nível de pós-graduação 50% dos professores da educação básica, até o último ano de vigência deste PNE, bem como, garantir a todos os profissionais da educação básica, a formação continuada.

O indicador 16A, relaciona-se ao número de professores da educação básica com pós-graduação lato sensu ou stricto sensu, e o indicador 16B ao número de professores da educação básica que realizaram cursos de formação continuada.

Segue-se a análise, empregando a técnica de regressão linear múltipla, com o uso de duas variáveis explicativas no modelo.

Portanto, algebricamente temos eq. (3)

$$Y_i = a + b_1 \cdot X_{1i} + b_2 \cdot X_{2i} + u_i \quad (3)$$

Onde Y representa a variável dependente, X_1 e X_2 variáveis explicativas, i representa cada uma das observações da amostra, u é o termo de erro ou resíduo e a e b são respectivamente o intercepto e a inclinação como exposto anteriormente.

Dessa forma, com o intuito de relacionar uma variável dependente com uma ou mais variáveis, se estabelece o indicador 2B, número da população de 16 anos de idade que concluiu pelo menos o ensino fundamental, como uma variável dependente quantitativa, e os indicadores 16A e 16B como variáveis explicativas.

Pode-se então modelar o problema como uma relação entre a ocorrência de jovens concluintes do EF (*jovens*) em função do número de professores com pós-graduação (*pos*), e professores com formação continuada (*cont*), conforme eq. (4)

$$jovens_i = a + b_1 \cdot pos_i + b_2 \cdot cont_i + u_i \quad (4)$$

Assim, o valor esperado da variável dependente será dado por eq. (5):

$$jovêns_i = \alpha + \beta_1 \cdot pos_i + \beta_2 \cdot cont_i \quad (5)$$

Em que α e β , são as estimativas dos parâmetros a e b , respectivamente.

Considera-se neste estudo o conceito de *ceteris paribus*, ou seja, mantém-se constante todos outros fatores e condições que possam influenciar o modelo, tendo em vista que o parâmetro de cada variável será analisado de forma isolada. Bem como, duas condições relacionadas aos resíduos: a somatória dos resíduos igual a zero e a somatória dos resíduos ao quadrado a mínima possível.

Parte-se para a elaboração da (Tabela 1) com dados referentes ao ano de 2020, último ano registrado da pesquisa, totalizando 27 observações referentes aos estados brasileiros.

Tabela 1. Jovens com EF concluído x Professores com pós-graduação e formação continuada (continua)

Estados	Jovens com EF concluído (Y_i)	Professores com pós-graduação (X_1)	Professores com formação continuada (X_2)
Rondônia	24.958	9.647	7.225
Acre	13.134	3.307	5.365
Amazonas	59.241	13.336	14.372
Roraima	6.008	3.056	4.092
Pará	133.842	26.631	31.667



Tabela 1. Jovens com EF concluído x Professores com pós-graduação e formação continuada (conclusão)

Estados	Jovens com EF concluído (Y_i)	Professores com pós-graduação (X_1)	Professores com formação continuada (X_2)
Amapá	8.485	4.144	4.707
Tocantins	17.322	6.367	8.659
Maranhão	109.203	32.080	45.583
Piauí	50.661	18.853	24.920
Ceará	151.355	38.634	46.152
Rio Grande do Norte	58.958	15.883	13.827
Paraíba	43.534	21.111	21.609
Pernambuco	123.102	34.343	28.349
Alagoas	47.045	13.417	14.124
Sergipe	23.767	10.114	9.655
Bahia	153.404	67.720	62.958
Minas Gerais	258.808	94.580	77.382
Espírito Santo	42.728	35.837	32.608
Rio de Janeiro	191.158	45.766	57.156
São Paulo	616.113	175.965	103.351
Paraná	142.340	106.144	103.861
Santa Catarina	89.547	51.834	56.391
Rio Grande do Sul	104.054	59.604	52.023
Mato Grosso do Sul	34.105	17.444	16.243
Mato Grosso	43.303	20.935	14.042
Goiás	80.510	30.144	21.053
Distrito Federal	54.228	14.517	16.747
Total	2.680.913	971.413	894.121

Fonte: INEP (2020).

No entanto, para o desenvolvimento do estudo, necessita-se garantir a distribuição normal como um pressuposto para o modelo estatístico de regressão linear e para os testes estatísticos da análise de variância.

Para isso, a partir de um pacote de recursos para o MS Excel® denominado Real Statistics Resource Pack¹, aplica-se o teste de normalidade de Anderson-Darlin para verificação se a amostra supracitada apresenta-se oriunda de uma distribuição normal.

Para a verificação, utiliza-se os dados das amostras ordenados, e calcula-se o valor z, considerando o valor da observação menos a média amostral, dividido pelo desvio padrão amostral. Aplica-se a função DIST.NORMP.N aos valores z-score para gerar uma função de distribuição cumulativa normal $f(z)$, em seguida avalia-se toda a distribuição, verificando as hipóteses do teste de Anderson-Darlin, que mede se os dados seguem uma distribuição específica ou não, sendo assim, as hipóteses são H_0 : os dados seguem a distribuição específica e H_1 : os dados não seguem uma distribuição específica.

Utiliza-se a função ADPROB, para encontrar o Valor-P, que neste caso (Tabela 2), apresenta-se menor que 0,05, em todas as amostras. Dessa forma, não se aceita a hipótese nula, desconsiderando a distribuição determinada.

¹ Real Statistics Resource Pack versão Rel 1.8.6: é um suplemento do Excel criado por Dr. Charles Zaiontz, que amplia os recursos das ferramentas de análise de dados. Disponível em <www.real-statistics.com>. Acesso em: 28 jun. 2022.



Tabela 2. Resumo dos resultados dos dados brutos e valores de Anderson-Darlin, para cada amostra

Resultados	Jovens com EF concluído	Professores com pós-graduação	Professores com formação continuada
Média	99293,074	35978,259	33115,593
Desvio Padrão	120444,422	38424,250	28554,638
Anderson-Darlin	2,580408	2,059134	1,340781
Valor-P	0,000001651	0,000030953	0,001786711

Fonte: Autoria própria (Resultados originais da pesquisa).

Por fim, utiliza-se a Transformação de Box Cox, através da função BOXCOX para gerar um conjunto de dados transformados, e em seguida, normaliza-se e verifica-se novamente por Anderson-Darlin se as amostras seguem a distribuição normal.

Desta vez, pode-se observar que o Valor-P aumentou para acima de 79% em todas as amostras (Tabela 3).

Tabela 3. Resumo dos resultados dos dados transformados, normalizados e valores de Anderson-Darlin, para cada amostra

Dados das amostras	Jovens com EF concluído	Professores com pós-graduação	Professores com formação continuada
Lambda p/ Box Cox	0,070457	0,062876	0,074236
Média	16,711	14,021	14,958
Desvio Padrão	2,277	1,924	1,950
Anderson-Darlin	0,195440861	0,124812928	0,23292483
Valor-P	0,890841601	0,986400825	0,798431057

Fonte: Autoria própria (Resultados originais da pesquisa).

Dessa forma, conclui-se que o novo conjunto de dados apresenta uma distribuição normal. E então, segue-se para a análise de regressão linear múltipla com os novos valores encontrados conforme (Tabela 4).

Tabela 4. Dados normalizados de jovens com EF concluído x Professores com pós-graduação e formação continuada (continua)

Estados	Jovens com EF concluído (Y_i)	Professores com pós-graduação (X_1)	Professores com formação continuada (X_2)
Rondônia	14,77304	12,41208	12,58202
Acre	13,49201	10,56869	12,01266
Amazonas	16,59203	12,99453	13,94667
Roraima	12,00769	10,43762	11,50537
Pará	18,41163	14,27894	15,60267
Amapá	12,65277	10,94691	11,76633
Tocantins	14,03719	11,68186	12,93454
Maranhão	17,94759	14,63430	16,39957
Piauí	16,25453	13,63049	15,09010
Ceará	18,69534	14,99335	16,42709
Rio Grande do Norte	16,58164	13,31387	13,86810
Paraíba	15,93101	13,84131	14,78943
Pernambuco	18,22004	14,76547	15,36476
Alagoas	16,09609	13,00553	13,91126
Sergipe	14,67342	12,49637	13,14883
Bahia	18,72651	16,10317	17,12431



Tabela 4. Dados normalizados de jovens com EF concluído x Professores com pós-graduação e formação continuada (conclusão)

Estados	Jovens com EF concluído (Y _i)	Professores com pós-graduação (X ₁)	Professores com formação continuada (X ₂)
Minas Gerais	19,96222	16,78259	17,59644
Espírito Santo	15,89138	14,84770	15,66593
Rio de Janeiro	19,24082	15,32423	16,90550
São Paulo	22,11456	18,08379	18,27105
Paraná	18,55335	17,02052	18,28265
Santa Catarina	17,50133	15,56965	16,87513
Rio Grande do Sul	17,83840	15,84729	16,69405
Mato Grosso do Sul	15,41736	13,48660	14,19689
Mato Grosso	15,91972	13,82566	13,89943
Goiás	17,26466	14,51501	14,73480
Distrito Federal	16,40085	13,14912	14,25972
Total	451,1972	378,5567	403,8553

Fonte: Autoria própria (Resultados originais da pesquisa).

A partir dos dados acima, calcula-se os valores estatísticos, por meio da ferramenta MS Excel®, seguindo o caminho: Dados > Análise de Dados > Regressão, onde apresenta-se os resultados na (Tabela 5).

Tabela 5. Estatísticas de regressão linear múltipla da variável jovens com EF concluído e das variáveis explicativas, professores com pós-graduação e formação continuada

Estatísticas de regressão	Resultados
R múltiplo	0,935319863
R-quadrado	0,874823246
R-quadrado ajustado	0,86439185
Erro padrão	0,83839031
Observações totais	27

Fonte: Autoria própria (Resultados originais da pesquisa).

Inicialmente observa-se o coeficiente de ajuste da regressão R-quadrado (R²), e pode-se dizer que 87,48% da variabilidade dos jovens que concluíram o EF é descrito pelas variáveis explicativas estudadas.

Segundo Fávero et al. (2009) em um modelo de regressão múltipla o coeficiente de ajuste R² mostra quanto do comportamento da variável Y é decorrente da variação conjunta das variáveis X do modelo. No entanto, este coeficiente somente captura uma relação entre as variáveis, sendo preciso identificar se X é estatisticamente significativa para explicar a variabilidade de Y.

Para isso, propõe-se a estatística F, onde se verifica o modelo que está sendo estimado examinando se pelo menos um parâmetro β é considerado estatisticamente diferente de 0, para haver influência dele na variável dependente.

Pela (Tabela 6) observa-se o F significação, pois no caso do MS Excel® não apresenta-se o valor de F crítico para se comparar com o valor do F calculado.

Sendo assim, se F significação for menor que 0,05 (5%), logo existe diferença significativa, e considerando as hipóteses H₀: $\beta_1 = \beta_2 = 0$ e H₁: existe pelo menos um $\beta \neq 0$, não se aceita a hipótese nula.



Tabela 6. Estatística F de Snedecor

Análise da variância	Grau de liberdade	Soma de quadrados	Quadrado médio	F calculado	F de significação
Regressão	2	117,896353	58,948	83,86	1,4801E-11
Resíduo	24	16,8695595	0,7029		
Total	26	134,765912			

Fonte: Autoria própria (Resultados originais da pesquisa).

No entanto, Belfiore (2015) explica que como o teste F avalia a significância conjunta das variáveis explicativas, se faz necessário a consideração da significância estatística de cada parâmetro, para apurar a inclusão da variável X no modelo.

Para isso propõe-se a estatística t, fazendo a análise do Valor-P, se é menor do que 0,05 para cada parâmetro. E considerando as hipóteses da estatística t como $H_0: \alpha = 0$ e $H_1: \alpha \neq 0$ para o parâmetro α , bem como $H_0: \beta = 0$ e $H_1: \beta \neq 0$, para cada parâmetro β , observa-se (Tabela 7) que existe diferença significativa somente para o parâmetro β_1 , relacionado à variável explicativa X_1 , o que sugere que grande parte da variabilidade está sendo explicada pelo parâmetro professores com pós-graduação.

Tabela 7. Estatística t de Student

Parâmetros	Coefficientes	Erro padrão	Estatística t	Valor-P
Interseção	0,873074474	1,27757476	0,6834	0,5009
X_1 = Professores com Pós-Graduação	0,797361073	0,37054073	2,1519	0,0417
X_2 = Professores com Formação Continuada	0,311442789	0,36553225	0,852	0,4026

Fonte: Autoria própria (Resultados originais da pesquisa).

Sobre o parâmetro α não significante, Belfiore (2015) diz que impor que α seja zero, favorece a geração de vieses com consequências na elaboração de previsões.

No caso do parâmetro β_2 , referente à variável explicativa X_2 , segue-se eliminando-o do modelo. Com isso, inicia-se nova análise de regressão, agora, simples, pois conta somente com os valores dos parâmetros α e β_1 . Como apresentado em (Tabela 8) não se observa muita diferença em relação aos valores da regressão linear múltipla.

Tabela 8. Estatísticas de regressão linear simples da variável jovens com EF concluído e da variável explicativa, professores com pós-graduação

Estatísticas de regressão	Resultados
R múltiplo	0,933293589
R-quadrado	0,871036923
R-quadrado ajustado	0,8658784
Erro padrão	0,833782386
Observações totais	27

Fonte: Autoria própria (Resultados originais da pesquisa).

Em (Tabela 9) com o valor de F significação, menor que 0,05, entende-se que o parâmetro β é diferente de zero. Confirmando que há influência desse parâmetro na variável dependente.

Tabela 9. Estatística F de Snedecor

Análise da variância	Grau de liberdade	Soma de quadrados	Quadrado médio	F calculado	F de significação
Regressão	1	117,3860857	117,38609	168,8539	1,2796E-12
Resíduo	25	17,37982669	0,6951931		
Total	26	134,7659124			

Fonte: Autoria própria (Resultados originais da pesquisa).

Novamente, calcula-se a estatística t, para verificar os valores estimados dos coeficientes dos parâmetros interseção e da variável explicativa relacionada aos professores com pós-graduação, conforme (Tabela 10), afim de serem utilizados na elaboração de uma nova equação para o modelo.

Tabela 10. Estatística t de Student

Parâmetros	Coefficientes	Erro padrão	Estatística t	Valor-P
Interseção	1,224433408	1,202543654	1,0182029	0,3183346
X = Professores com Pós-Graduação	1,104557186	0,085002679	12,994381	1,2796E-12

Fonte: Autoria própria (Resultados originais da pesquisa).

Abaixo, têm-se a eq. (6) do modelo de regressão linear com os coeficientes estimados.

$$jovêns_i = 1,224433408 + 1,104557186 \cdot pos_i \quad (6)$$

E também, pode-se observar o gráfico de ajuste do modelo (Figura 5), onde apresenta-se os jovens de 16 anos com ensino fundamental concluído e seu valor esperado, em função dos professores com pós-graduação.

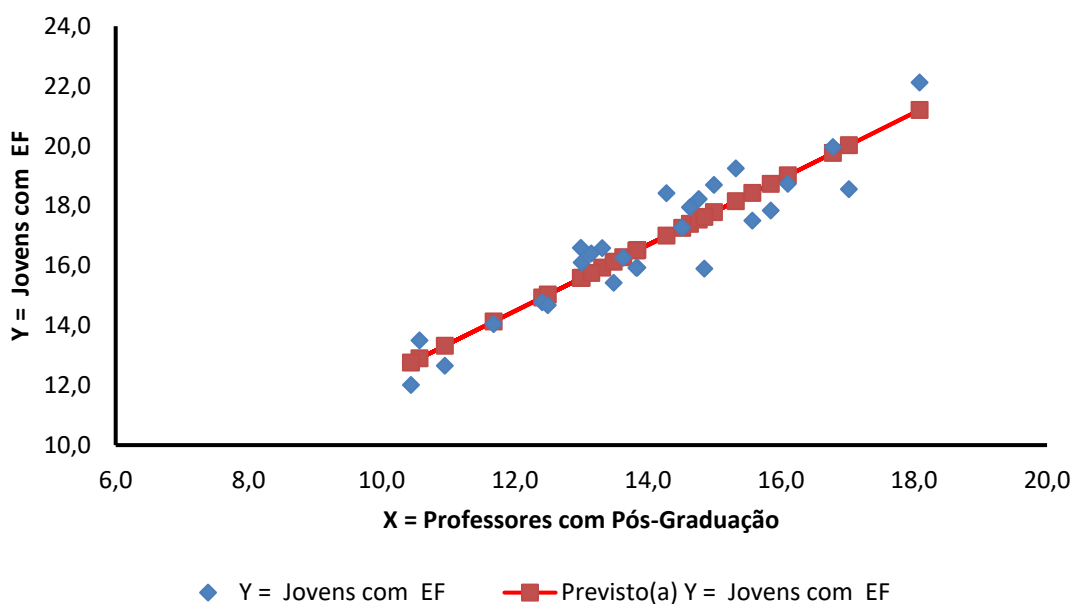


Figura 5. Jovens concluintes do EF em função dos Professores com Pós-Graduação (modelo).

Fonte: Autoria própria (Resultados originais da pesquisa).

Nota: Valores de X e Y normalizados.



A partir da equação elaborada, propõe-se a seguinte pergunta para análise: quantos professores com pós-graduação seriam necessários para que a meta de 95% dos jovens com 16 anos concluintes do EF, seja alcançada segundo o modelo proposto?

Com base nos dados do Inep Data, podemos considerar um total de 3.254.775 jovens, e dentre esses, 95% que equivale a 3.092.036, substitui-se no modelo.

Em sequência, chega-se ao resultado de 2.799.343 professores com pós-graduação, valor que ultrapassa o total geral de professores 2.235.935.

Dessa forma, percebe-se que é um indicador que influencia, mas que se torna necessário identificar outros indicadores relacionados, ou aprofundar-se nos indicadores de professores, para explicar o problema de jovens que não concluem a educação básica na idade especificada. Sabe-se que outras variáveis explicativas podem ser inseridas no modelo a fim de gerar melhor previsibilidade.

4. Conclusão

Após estas observações, conforme os resultados avaliados, conclui-se que realmente ocorre relação entre os indicadores associados aos jovens concluintes e o indicador associado à formação dos professores, especificamente, pós-graduação. Sendo grande parte da variabilidade explicada por esta variável. O indicador formação continuada dos professores, não apresenta influência significativa na variável dependente.

Nos dados do PNE, ao tratar-se a pós-graduação dos professores, não se apresenta dados com a distinção entre mestrado e doutorado, ou áreas de interesse, mas percebe-se que uma formação com maior rigor científico torna-se significativa neste estudo.

Propõe-se outras análises mais aprofundadas adentrando nos indicadores de professores, para isso sugere-se a modelagem linear hierárquica, também chamada de multinível, como mecanismo de investigação, pois tem-se observado que esta análise ofereceu uma boa compreensão de fenômenos no campo de estudo da Educação e das Ciências Sociais.

No entanto, sabe-se também que existem muitos outros fatores explicativos que possivelmente relacionam-se ao indicador 2B, principalmente os indicadores sociais, e econômicos que não foram analisados por hora. Contudo, pode-se reforçar a grande importância do desenvolvimento de planos educacionais com estratégias cada vez mais pautadas em dados a fim de se obter melhoramento neste quesito.

5. Referências

BELFIORE, P. **Estatística aplicada a administração, contabilidade e economia com Excel e SPSS**. 1ed. Rio de Janeiro, RJ: Elsevier, 2015.

BRASIL. **Lei n. 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação - PNE**.

Brasília, DF: Presidência da República, 2014. Disponível em:

<https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l13005.htm >. Acesso em: 07 nov. 2021.

FÁVERO, L.P.; BELFIORE, P.; SILVA, F.L.; CHAN, B.L. 2009. **Análise de dados: modelagem multivariada para tomada de decisões**. Rio de Janeiro, RJ: Elsevier, 2009.



INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Painel de Indicadores do Plano Nacional de Educação**. Brasília, 2020. Disponível em: <<https://inepdata.inep.gov.br/analytics/saw.dll?dashboard>>. Acesso em: 18 out. 2021.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Relatório do 3º ciclo de monitoramento das metas do Plano Nacional de Educação – 2020: sumário executivo**. Brasília, 2020.

