

## **Fatores que Impactam no Índice de Desenvolvimento da Educação Básica (IDEB), em Instituições Públicas de Ensino do DF.**

**Ricardo Monteiro, Marcia Longen Zindel, Carlos Henrique Rocha**

**Resumo:** O objetivo do presente trabalho é analisar a influência da etapa de qualificação dos professores, do tempo de serviço dos docentes, do número total de professores por escola e da localização da escola sobre o Índice de Desenvolvimento da Educação Básica (IDEB), levando-se em conta apenas as instituições públicas de ensino do DF. A metodologia utilizada para a análise de dados foi o de regressão linear, multivariada com variáveis explicadoras binárias. O estudo revelou que os principais fatores influenciadores no IDEB são o tempo médio de serviço dos professores e a realização de cursos de especialização, já a quantidade de professores por unidade escolar influencia negativamente na *performance* dos alunos, de tal modo que é preferível que se tenha unidades pequenas ou médias de ensino em oposição a grandes escolas. A presença ou não de professores com Mestrado ou Doutorado não influenciou no IDEB. Já a localização da escola ou sua subordinação administrativa mostra disparidade entre a *performance* das escolas de diversas regiões do DF.

**Palavras chave:** Ensino, Educação de Professores, Instituições Públicas do DF, Desempenho Docente.

## **Factors of Impact the Index of Development of Basic Education (IDEB), in Public Institutions of the DF.**

**Abstract:** The objective of the present study is to analyze the influence of teacher qualification stage, length of service of teachers, total number of teachers per school and school location on the Basic Education Development Index (IDEB), leading to only the public educational institutions of the DF. The methodology used for data analysis was linear regression, multivariate with binary explanatory variables. The study revealed that the main influencing factors in IDEB are the average length of service of teachers and the completion of specialization courses, whereas the number of teachers per school unit negatively influences student performance, so it is preferable to have small or medium schools as opposed to large schools. The presence or absence of masters or doctorate teachers did not influence IDEB. The location of the school or its administrative subordination shows a disparity between the performance of schools in different regions of the DF.

**Keywords:** Teaching, Teacher Education, DF Public Institutions, Teaching Performance.

### **1. Introdução**

Tendo em vista a importância central da educação na evolução de uma sociedade e os esforços realizados pelo governo brasileiro para universalizar e democratizar o ensino público, principalmente após a constituição de 1988 e a publicação da Lei nº 9.394/1996. E, em termos regionais a Lei 5.105/2013 que trata do plano de carreira dos Professores de Escolas Públicas do Distrito Federal, que prevê, dentre outras coisas aumento salarial para os profissionais da educação conforme a

realização de cursos de especialização, mestrado e doutorado. Surge a necessidade de avaliar quais fatores estruturais influenciam na proficiência dos alunos nas disciplinas. Desta forma, o objetivo deste trabalho é analisar a influência da etapa de qualificação dos professores, do tempo de serviço dos docentes, do número total de professores por escola e da localização da escola sobre o Índice de Desenvolvimento da Educação Básica (IDEB), levando-se em conta apenas as instituições públicas de ensino do DF.

## 2. Referencial teórico

### 2.1 IDEB

Segundo a Nota Técnica divulgada pelo INEP (2007), o IDEB é um indicador de qualidade educacional que combina informações de desempenho em exames padronizados obtidos pelos alunos ao final das etapas do ensino básico com informações sobre o rendimento escolar. O IDEB foi criado para sintetizar em um único indicador as informações advindas tanto dos testes quanto do rendimento escolar, sendo que a primeira se origina da aplicação da SAEB e da Prova Brasil e o segundo do Censo Escolar Anual.

Segundo INEP (2007), o IDEB é calculado na forma da Equação 1:

$$\text{IDEB}_{ji} = N_{ji}P_{ji} \quad 0 \leq N_j \leq 10; 0 \leq P_j \leq 1 \text{ e } 0 \leq \text{IDEB}_j \leq 10 \quad (1)$$

Em que:

IDEB<sub>ji</sub> é o valor do IDEB da escola j no ano i;

N<sub>ji</sub> é a média da proficiência em Língua Portuguesa e Matemática, padronizada para um indicador entre zero e dez, dos alunos da unidade j, obtida em determinada edição do exame realizado ao final da etapa de ensino;

P<sub>ji</sub> é o indicador de rendimento baseado na taxa de aprovação da etapa de ensino dos alunos da unidade j.

Na Equação 1, a média de proficiência padronizada dos estudantes da unidade j é obtida a partir das proficiências médias em Português e Matemática dos estudantes submetidos a determinada edição do exame realizada ao final da etapa educacional considerada. A proficiência é padronizada para estar entre 0 e 10 de modo que isso se reflita no cálculo do IDEB. N<sub>ji</sub> é obtido com base na Equação 2:

$$N_{ji} = \frac{n_{ji}^{lp} + n_{ji}^{mat}}{2} \quad (2)$$

Por sua vez, as proficiências em português e matemática,  $n_{ji}^p$  e  $n_{ji}^m$ , são calculados conforme a Equação 3.

$$n_{ji}^{\alpha} = \frac{S_{ji}^{\alpha} - S_{inf}^{\alpha}}{S_{sup}^{\alpha} - S_{inf}^{\alpha}} \times 10 \quad (3)$$

Onde:

$n_{ji}^{\alpha}$  = proficiência na disciplina  $\alpha$ , obtida pela unidade  $j$ , no ano  $i$ , padronizada para valores entre 0 e 10;

$\alpha$  = disciplina, matemática ou português;

$S_{ji}^{\alpha}$  = proficiência média (em Língua Portuguesa ou matemática), não padronizada, dos alunos da unidade  $j$  obtida no exame do ano  $i$ ;

$S_{inf}^{\alpha}$  = limite inferior da média de proficiência (Língua Portuguesa ou Matemática) do SAEB 1997;

$S_{sup}^{\alpha}$  = limite superior da média de proficiência (Língua Portuguesa ou Matemática) do SAEB 1997.

Para as unidades escolares, ou redes, que obtiverem  $S_{ji}^{\alpha} < S_{inf}^{\alpha}$ , a proficiência média é fixada em  $S_{inf}^{\alpha}$ , já para aquelas que superarem o limite superior, seu desempenho é fixado em  $S_{sup}^{\alpha}$ .

A Tabela 1 apresenta, as médias e os desvios-padrão das proficiências dos alunos da 4ª e da 8ª série do ensino fundamental e da 3ª série do ensino médio no SAEB de 1997.

Tabela 1 - SAEB 1997: Proficiências médias e desvio padrão Brasil

Série	Matemática		Língua Portuguesa	
	Média	Desvio-Padrão	Média	Desvio-Padrão
4ª do EF	190.8	44	186.5	46
8ª do EF	250.0	50	250.0	50
3ª do EM	288.7	59	283.9	56

Fonte: SAEB (1997) – Inep/MEC

A Tabela 2, por sua vez, apresenta os valores dos limites inferiores e superiores utilizados na padronização das proficiências médias em Língua Portuguesa e Matemática dos alunos da 4ª e da 8ª série do ensino fundamental e da 3ª série do ensino médio.

Tabela 2–SAEB 1997 - Limite superior e inferior das proficiências Brasil

Série	Matemática		Língua Portuguesa	
	Sinf	Ssup	Sinf	Ssup
4ª do EF	60	322	49	324
8ª do EF	100	400	100	400
3ª do EM	111	467	117	451

Fonte: SAEB (1997) – Inep/MEC

A partir da média e desvio padrão das proficiências no SAEB 1997 (ano em que a escala do SAEB foi definida), calcularam-se, para cada etapa de ensino, considerando as diferentes disciplinas avaliadas no exame, os limites inferior e superior, de acordo com Equação 4:

$$S_{inf}^{\alpha} = média_{\alpha} - 3\sigma \text{ e } S_{sup}^{\alpha} = média_{\alpha} + 3\sigma \quad (4)$$

Os limites superiores e inferiores apresentados na Tabela 2 que são utilizados para calcular o IDEB das escolas e redes de todo o Brasil desde 1997. Já o indicador de rendimento,  $P_j$ , é obtido conforme a Equação 5, onde a proporção de aprovados em cada uma das séries da etapa considerada,  $p_r$ , é calculada diretamente do Censo Escolar. Se  $p^r$  ( $r = 1, 2, \dots, n$ , em que  $n$  é o número de séries com taxa de aprovação positiva) é a taxa de aprovação da  $r$ -ésima série da etapa educacional considerada, então o tempo médio de duração da série é:

$$T_{ji} = \sum_{r=1}^n \frac{1}{p^r} = \frac{n}{P_{ji}} \quad (5)$$

Na Equação 5,  $P_{ji}$  é a taxa média de aprovação na etapa educacional no ano  $i$ . Note-se que, na ausência de evasão durante a etapa e em equilíbrio estacionário,  $\frac{n}{P_{ji}}$  dá o tempo médio ( $T_{ji}$ ) para conclusão de uma etapa para os estudantes da unidade  $j$ . Se  $P$  é o inverso do tempo médio para conclusão de uma série, então,  $P_{ji} = \frac{1}{T_{ji}}$ . Deste modo têm-se que  $IDEB_{ji} = \frac{N_{ji}}{T_{ji}}$ , ou seja, o indicador fica sendo a pontuação no exame padronizado ajustado pelo tempo médio (em anos) para a conclusão de uma série naquela etapa de ensino.

Como o IDEB é resultado do produto entre o desempenho e do rendimento escolar (ou o inverso do tempo médio de conclusão de uma série) então ele pode ser interpretado da seguinte maneira: para

uma escola A cuja média padronizada da Prova Brasil, 4ª série, é 5,0 e o tempo médio de conclusão de cada série é de 2 anos, o IDEB será igual a 5,0 multiplicado por 0,5, ou seja, 2,5. Já uma escola B com média padronizada da Prova Brasil, 4ª série, igual a 5,0 e tempo médio para conclusão igual a 1 ano, terá IDEB igual a 5,0.

### **3. Metodologia**

#### **3.1. Análise de regressão**

Segundo Gujarati & Porter (2010), a análise de regressão é uma ferramenta estatística que permite o estudo da relação entre uma variável dependente e uma ou mais variáveis explicativas com o objetivo de estimar o valor médio populacional da variável regressada em função dos dados conhecidos das variáveis regressoras.

À análise de regressão interessa o que é conhecido como dependências estatística entre variáveis, diferentemente da dependência funcional ou determinística das ciências naturais. De fato, na primeira se trabalha com variáveis aleatórias ou estocásticas, que em conjunto, jamais irão prever com certeza toda a variação de um fenômeno. Importante ainda é ressaltar que apesar de estudar a dependência de uma variável em relação as outras, a análise de regressão não implica em causalidade necessariamente, por maior que seja a relação estatística entre duas variáveis.

No presente trabalho, o modelo de regressão utilizado foi o de regressão linear, multivariada com variáveis explicadoras binárias. Neste modelo em específico, Gujarati & Porter (2010) se posicionam afirmando que na análise de regressão a variável dependente pode acusar influência de não apenas variáveis em escala de razão, mas também de variáveis qualitativas por natureza ou de escala nominal.

Dentro das variáveis de escala nominal, existe um tipo que é destinado a mostrar se na amostra determinada característica existe ou inexistente, está presente ou não nos indivíduos que a compõe e este determinado tipo é conhecido com variáveis binárias e alguns exemplos deles são: sexo (masculino ou feminino), indicar se a pessoa participou ou não de um curso, se a pessoa votou ou não em um candidato, etc.

Para Wooldridge (2007), as variáveis binárias ou dummy, se originam naturalmente, tendo em vista que os fatores qualitativos frequentemente aparecem de maneira excludente. O autor ressalta a importância de se atribuir corretamente o significado da variável binária, no exemplo citado, ele afirma que se pode caracterizar homem com o valor 1 e mulher com o valor 0, ou vice-versa, mas não

se deve caracterizar gênero com valor 1, pois isso não dá significado a variável, uma vez que não se sabe o gênero da pessoa.

A utilização dos valores binários 0 e 1 é arbitrária, quaisquer outros valores binários poderiam ser utilizados no lugar. No entanto, a utilização deste par leva a modelos de regressão nos quais os parâmetros têm interpretações bastantes naturais.

A Equação 14, demonstra o modelo geral de regressão:

$$y = \beta_0 + \delta_0 x + \beta_1 x_1 + \mu \quad (6)$$

Onde:

$y$  é a variável explicada;

$\beta_0$  é o intercepto;

$\delta_0$  é o coeficiente da variável binária

$x$  é uma variável binária regressora;

$x_1$  é uma variável em escala de razão regressora;

$\beta_1$  é o coeficiente de qualquer variável em escala de razão;

$\mu$  é a média dos erros.

No modelo estabelecido na Equação 14, somente dois fatores influenciam em  $Y$ . O primeiro deles é o  $x$ , a variável binária, e o outro é o  $x_1$  variável em escala de razão. Como  $x$  assume os valores 0 ou 1, nos momentos em que assume 0, todo o  $Y$  é explicado por  $\beta_0 + \beta_1 x_1 + \mu$ . Já nos momentos em que assume 1,  $Y$  passa a ter a diferença  $\delta_0$  incorporada ao modelo. Dessa forma conclui-se que a utilização dos valores 0 e 1 levam a interpretação de que o coeficiente da variável binária é a diferença atribuída por aquela característica na explicação da variável regressada, dado um mesmo grau de  $x_1$  e erro  $\mu$ .

Para uma expectativa de que o valor esperado de  $\mu$  seja 0,  $\delta_0$  é da do pela Equação 15:

$$\delta_0 = E(Y|x = 1, x_1) - E(Y|x = 0, x_1) \quad (7)$$

Mantendo-se o grau de  $x_1$ , a diferença  $\delta_0$  é devida somente a  $x$ .

### 3.2 Método *stepwise*

Segundo o Portal Action (2018), o método *Stepwise* para seleção de variáveis é muito utilizado em regressão linear. Ele se baseia em um algoritmo que checa a importância das variáveis em termos de uma medida de significância estatística do coeficiente associado à variável para o modelo.

O método é uma modificação da seleção *Forward* em que cada passo é verificado com base na estatística F parcial das variáveis. O método consiste em calcular dois valores que serão utilizados como parâmetros de inserção ou de remoção da variável  $F_{in}$  e  $F_{out}$ , segundo a Equação 16.

$$F = \frac{\left( \frac{SSE_{(j-xr)} - SS_j}{GL_{xr}} \right)}{MSE_j} \quad (16)$$

Onde:

$SSE_{(j-xr)}$  é o Erro SS para o modelo que não contenha  $X_r$ ;

$SSE_j$  é o Erro SS para o modelo que contenha  $X_r$ ;

$MSE_j$  o MS erro para o modelo que contenha  $X_r$ ;

$GL_{xr}$  são os graus de liberdade da variável;

O procedimento segue:

É escolhida, primeiramente, a variável de maior correlação com a variável explicada;

A cada passo *Forward*, depois de incluída uma nova variável é verificado se alguma variável deve ser excluída;

Repete-se os passos até não excluir nem incluir nenhuma variável;

Se  $F_{in} < F_{out}$  é mais fácil adicionar a variável que removê-la;

Se  $F_{in} > F_{out}$  é mais fácil remover a variável que adicioná-la.

### 4. População e Amostra

Para a realização do Estudo, foram consideradas as notas obtidas pelas unidades escolares nos anos de 2005 a 2015, conforme foi possível encontrar na base de dados da SEDF, os professores que lá lecionavam e seus graus de capacitação. A tabela 3 resume os dados da população utilizada no trabalho.

Tabela 3 - Amostra do trabalho, IDEB e Total de Escolas do DF

Ano	População do Trabalho	População Colhida pelo IDEB	Total de Escolas Públicas	% do total de escolas
2005	224	259	612	36,6%
2007	249	288	617	40,3%
2009	279	330	622	44,8%
2011	293	337	645	45,4%
2013	325	330	651	49,9%
2015	322	328	658	48,9%

Fonte: o autor, dados do autor e dos documentos: Série Histórica do Censo Escolar 2004 – 2013 e Série Histórica do Censo Escolar 2006 – 2016.

Para cada uma das escolas e para cada ano estudado, foram encontrados o número total de professores que ministrava na escola no respectivo ano, o tempo médio de serviço dos profissionais, o número total de profissionais com especialização, o número total de profissionais com mestrado, o número total de profissionais com doutorado e a Coordenação Regional de Ensino ao qual a escola estava subordinada.

## 5. Modelo

Para proceder à análise de regressão foi estabelecido o seguinte modelo.

$$IDEB = \beta_0 + \delta_1 E + \delta_2 M + \delta_3 D + \beta_1 T + \beta_2 prof + \sum_{i=1}^{16} \delta_i L_i \quad (8)$$

Onde:

IDEA é a variável explicada e reflete o IDEB calculado pelo INEP para as escolas públicas do DF nos anos de 2005 a 2015.

$\beta_0$  é a média do IDEB para as Escolas de Brazlândia, grupo de referência;

$\delta_1$  é a diferença atribuída ao modelo em função do percentual de especialistas no corpo docente

$\delta_2$  é a diferença atribuída ao modelo em função da presença ou não de mestres no corpo docente.

$\delta_3$  é a diferença atribuída ao modelo em função da presença ou não de doutores no corpo docente.

$\delta_{4-16}$  são a diferença atribuída ao modelo em função da localidade  $L_4 \rightarrow 16$ ;

$L_i$  representa cada uma das 13 Coordenações Regionais de Ensino, (com exceção de Brazlândia que é o grupo de referência) às quais as escolas estão vinculadas e assume o valor 1 se a escola pertence àquela regional e 0 caso não pertença.

“E” é uma variável binária que assume os valores 1 para um corpo docente composto por mais de 50% de especialistas e 0 para um corpo docente composto por menos de 50% de especialistas.

“M” é uma variável binária que assume os valores 1 para um corpo docente composto por um ou mais mestres e 0 para um corpo docente sem nenhum mestre.

“D” é uma variável binária que assume os valores 1 para um corpo docente composto por um ou mais doutores e 0 para um corpo docente sem nenhum doutor.

$\beta_1$  é a influência da variável Tempo médio de Serviço.

T é o tempo médio de serviço dos profissionais de uma escola.

$\beta_2$  é a influência no modelo do número total de docentes da escola.

Prof é o número total de docentes da escola.

Os resultados obtidos referente à regressão estão resumidos na Tabela 4, a primeira parte da tabela trata das estatísticas do modelo de regressão, enquanto que a segunda parte mostra os coeficientes padronizados obtidos.

Tabela 4 - Resumo dos Resultados Obtidos pela Regressão (2005 – 2015, Distrito Federal)

<b>Estatística</b>	<b>Regressão</b>
R <sup>2</sup>	0,51
Erro Padrão	0,49
Durbin-Watson	1,83
SQR	404,245
<b>Coeficientes Padronizados</b>	
Nº de Professores	-0,063
Tempo Médio	0,468
Especialistas	0,346
Mestres	-
Doutores	-
Brazlândia	3,856
Plano Piloto	+0,156
Santa Maria	+0,064
Paranoá	+0,051
S. Sebastião	+0,041
Ceilândia	-
N.Bandeirante	-

R. das Emas	-
Samambaia	-
Guará	-0,053
Taguatinga	-0,069
Gama	-0,072
Sobradinho	-0,087
Planaltina	-0,111

Fonte: o autor.

## 6. Análise dos Resultados.

Os resultados obtidos mostram que o modelo de regressão apresentou um  $R^2$  de 0,51 significando que o modelo proposto consegue explicar 51% da variação do IDEB das escolas públicas do DF. O ideal é que sejam incorporadas mais variáveis com a finalidade de se aumentar o grau de explicação do modelo. Provavelmente as novas variáveis a serem incorporadas deverão traduzir características relativas aos discentes e suas famílias.

O erro padrão foi de 0,49 e a estatística Durbin-Watson mostrou que não houve autocorrelação positiva nos resíduos da regressão gerada, validando o modelo.

No que se refere aos coeficientes padronizados, houve uma influência negativa no IDEB na variável número total de professores por escola, revelando que quanto maior o número de professores lotados na escola, menor tende a ser o seu IDEB. Já com relação às influências positivas foram, destacam-se o tempo médio de serviço dos profissionais da escola e a presença de um corpo docente composto por mais de 50% de especialistas. No caso do primeiro, revela que as escolas cujos professores exercem suas funções a mais tempo tendem a ter um resultado no IDEB melhor do que as escolas cujos profissionais tem menos experiência.

Já com relação à especialização, a presença de um corpo docente composto majoritariamente por professores com esse nível de capacitação influenciou positivamente no IDEB das escolas, o que não foi verificado para os níveis de capacitação de mestrado ou doutorado.

Por último, quanto à localização das escolas, verificou-se uma influência positiva para as escolas subordinadas à regional do Plano Piloto, Santa Maria, Paranoá e São Sebastião em relação às escolas

da Coordenação Regional de Ensino de Brasília. Escolas das Regionais de Ensino de Ceilândia, Núcleo Bandeirante, Recanto das Emas e Samambaia, não apresentaram influência quanto a localização e escolas do Guará, Taguatinga, Gama, Sobradinho e Planaltina, obtém, em média, IDEB inferiores. Não há infelizmente dados para se dizer como a localização das escolas influencia no IDEB obtido pelos alunos, para isso deve-se proceder um novo estudo, reunindo-se as características das escolas de determinada regional que podem interferir no IDEB dos alunos e analisar essas características cruzando com as escolas de outras regionais.

## **7. Conclusão.**

O estudo revelou que os principais fatores influenciadores no IDEB dos alunos, no que diz respeito à infraestrutura fornecida, são o tempo médio de serviço dos professores e se estes realizaram cursos de Especialização para dar aula. A quantidade de professores por unidade escolar influencia negativamente na performance dos alunos, de tal modo que é preferível que se tenha unidades pequenas ou médias de ensino em oposição a grandes escolas. A presença ou não de professores com Mestrado ou Doutorado não influenciou no IDEB. Já a localização da escola ou sua subordinação administrativa mostra disparidade entre a performance das escolas de diversas regiões do DF, devendo-se comparar a infraestrutura fornecida pela Regional do Plano Piloto, regional com maior média no IDEB, com as das demais escolas para se descobrir quais fatores estão presentes ou ausentes que possam levar os alunos da Regional do Plano a serem mais bem sucedidos que os demais nas provas de proficiência.

## **REFERÊNCIAS**

**ATKINSON**, Anthony A., et al. Contabilidade Gerencial. São Paulo, Atlas, 2015.

**GUJARATI**, Damodar N. e **PORTER**, Dawn C. **Econometria**. Cidade do México, McGRAW-HILL, 2010.

**INSTITUTO NACIONAL DE ESTUDO E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA**. Documentos e Legislação. Site do INEP. Disponível em:  
<[http://download.inep.gov.br/educacao\\_basica/portaI\\_ideb/o\\_que\\_sao\\_as\\_metas/Artigo\\_projecoes.pdf](http://download.inep.gov.br/educacao_basica/portaI_ideb/o_que_sao_as_metas/Artigo_projecoes.pdf)> Acesso em: 05 mai. 2018.

**INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA**. [2007]. SAEB. Site do INEP. [Online] INEP, 2007. Disponível em: <<http://portal.inep.gov.br/educacao-basica/saeb>> Acesso em: 14 de maio de 2018.

**INSTITUTO NACIONAL DE PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA.** 2007. Censo Escolar. Site do INEP, [2007]. Disponível em: <<http://inep.gov.br/censo-escolar>.> Acesso em: 18 de Maio de 2018.

**DISTRITO FEDERAL** (Estado). Lei distrital nº 5.105 de maio de 2013. Diário Oficial do Distrito Federal, Poder Executivo, Brasília, DF, 05 de maio de 2013.

**MORETTIN**, Pedro A. e **BUSSAB**, Wilton de O. Estatística Básica. São Paulo, SP, Saraiva, 2004.

**PORTAL ACTION.** 2018. Seleção Stepwise. Site do Portal Action. Portal Action, 2018. Disponível em: <<http://www.portalaction.com.br/analise-de-regressao/4251-selecao-stepwise>.> Acesso em: 13 de 06 de 2018.

**PRODANOV**, Cleber Cristiano e **FREITAS**, Ernani Cesar de. Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico. [ed.] Daiane Thomé Scariot e Grazielle Borguetto Souza. 2ª ed., Novo Hamburgo, RS, Feevale, 2013.

**SECRETARIA DE ESTADO DE EDUCAÇÃO DO DISTRITO FEDERAL.** SAEB e Prova Brasil. Site da Secretaria de Estado de Educação do DF. SEDF, 2018. Disponível em: <<http://www.educacao.df.gov.br/saeb-e-prova-brasil/>.> Acesso em: 19 de maio de 2018.

**SECRETARIA DE ESTADO DE EDUCAÇÃO DO DISTRITO FEDERAL.** Série Histórica do Censo Escolar 2004 - 2013. Brasília, DF, [2014]

**SECRETARIA DE ESTADO DE EDUCAÇÃO DO DISTRITO FEDERAL.** Série Histórica do Censo Escolar 2007 - 2016. Brasília, DF, SEDF, 2017.

**WOOLDRIDGE**, Jeffrey M. Introdução à Econometria. São Paulo, SP, Cengage Learning, 2007