

Universidade Federal de Ouro Preto - UFOP Instituto de Ciências Sociais
Aplicadas Graduação em Economia

Breno Henrique Borges De Gusmão

**Eficiência das Escolas Municipais e Estaduais dos
Municípios de Ouro Preto, Mariana e
Conselheiro Lafaiete**

Mariana, Minas Gerais

Dezembro, 2019

Breno Henrique Borges De Gusmão

Eficiência das Escolas Municipais e Estaduais dos Municípios de Ouro Preto, Mariana e Conselheiro Lafaiete

Monografia apresentada ao Curso de Ciências Econômicas do Instituto de Ciências Sociais Aplicadas (ICSA) da Universidade Federal de Ouro Preto (UFOP), como requisito à obtenção de grau de Bacharel em Ciências Econômicas

Orientador: Prof. Dr. Victor Maia Senna Delgado

Mariana, Minas Gerais

Dezembro, 2019

G982e Gusmão, Breno de.
Eficiência das Escolas Municipais e Estaduais dos Municípios de Ouro Preto, Mariana e Conselheiro Lafaiete [manuscrito] / Breno de Gusmão. - 2019.

42f.:

Orientador: Prof. Dr. Victor Delgado.

Monografia (Graduação). Universidade Federal de Ouro Preto. Instituto de Ciências Sociais Aplicadas. Departamento de Ciências Econômicas e Gerenciais.

1. Escolas - Teses. 2. Eficiência organizacional - Teses. 3. Educação básica - Teses. 4. Análise de envoltória de dados - Teses. 5. Censo Escolar - Teses. I. Delgado, Victor. II. Universidade Federal de Ouro Preto. III. Título.

CDU: 373

Catálogo: ficha.sisbin@ufop.edu.br



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIÊNCIAS SOCIAIS E APLICADAS
DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS

**FOLHA DE APROVAÇÃO****Breno Henrique Borges de Gusmão****Eficiência das Escolas Municipais e Estaduais dos Municípios de Ouro Preto, Mariana e Conselheiro Lafaiete**

Membros da banca

Victor Maia Senna Delgado - Dr. - UFOP
Carlos Eduardo da Gama Torres - Dr. - UFOP
Marcelo Aparecido Cabral Nogueira - Dr. - UFOP

Versão final
Aprovado em 19 de Dezembro de 2019

De acordo

Professor (a) Orientador (a) Victor Maia Senna Delgado



Documento assinado eletronicamente por **Victor Maia Senna Delgado, PROFESSOR DE MAGISTERIO SUPERIOR**, em 19/12/2019, às 16:46, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0029832** e o código CRC **2FC1FD31**.

Referência: Caso responda este documento, indicar expressamente o Processo nº 23109.204197/2019-60

SEI nº 0029832

R. Diogo de Vasconcelos, 122, - Bairro Pilar Ouro Preto/MG, CEP 35400-000
Telefone: - www.ufop.br

Agradecimentos

A Universidade Federal de Ouro Preto, e todo seu corpo docente.

A meu orientador Victor Maia Senna Delgado pelo suporte.

A minha mae Izabel, por me ensinar que a vida deve ser encarada com leveza.

A minha mãe-irmã Vivi, sem a qual nada disso teria sido possível.

A minha irmã Carolina, pelas confidências, incentivos, e repreensões.

A meu pai Bernardo, pelo apoio incondicional e pelos valores ensinados.

A minha sobrinha Isabela - pra sempre minha pequena - responsável por mudar minha maneira de enxergar a vida, e por me tornar uma pessoa melhor.

Ao meu sobrinho e afilhado Bernardo, pelas demonstrações puras e simples de amor, carinho, e bom humor.

Ao meu cunhado Leonardo, muito mais que um cunhado.

Aos meus amigos, Ramom, Thomas, Fraldinha, Derick, Tiago Russo, Versieux, Gustavo, José Bruno e Alisson.

RESUMO:

A monografia calculou a eficiência das escolas do 5º ano do Ensino Fundamental nos municípios de Conselheiro Lafaiete, Mariana e Ouro Preto. Foram utilizados quatro modelos diferentes, em todos eles mantendo níveis médios de Proficiência em Português e Matemática por escola (Produtos) e optando, exclusivamente, pela variação dos insumos. O método de Análise de Dados Envoltórios (*Data Envelopment Analysis* – DEA) possibilitou a criação de fronteiras de eficiências, que foram utilizadas como referências comparativas para as demais escolas, *Decision Making Units* (DMUs), presentes na Função de Produção.

Palavras-chave: DEA; DMUs; Escolas; Eficiência; Ensino Básico; Prova Brasil; Censo Escolar

ABSTRACT:

This paper calculated the efficiency of the schools of the 5th year of elementary school in the municipalities of Conselheiro Lafaiete, Mariana and Ouro Preto, Minas Gerais, Brazil. Four different models were used in all of them maintaining the levels of Proficiency by school in Reading and Mathematics (Products) and choosing exclusively four models for the variation of inputs. The Method used was the Data Envelopment Analysis – DEA, it enabled the creation of efficiencies frontiers, which were used as comparative references for the other schools, the Decision Making Units (DMUs), also present in the production function.

Keywords: DEA; DMUs; Efficiency; Elementary School; Prova Brasil; Educational Census

LISTA DE FIGURAS

Gráfico 1 - Fronteira de Produção	21
Gráfico 2 - Função de Produção - Português SAEB	28
Gráfico 3 - Função de Produção Matemática - SAEB.....	29

LISTA DE TABELAS

Tabela 1 - Taxa de Analfabetismo.....	16
Tabela 2 - Taxa de Analfabetismo por Região	16
Tabela 3 - Taxa de Analfabetismo por Idade	17
Tabela 4 - Taxa de Analfabetismo em Relação a Renda.....	18
Tabela 5 - Insumos	26
Tabela 6 – Produtos	26
Tabela 7 - Modelos e Variáveis	27
Tabela 8 - Eficiência Proporcional	30
Tabela 9 - Eficiência Média.....	31
Tabela 10 - Resultado Médio dos Insumos	31
Tabela 11 - Média dos Produtos	32

SUMÁRIO

1 – Introdução	11
2 – Revisão	12
2.1 – Capital Humano.....	12
2.2 – Educação Brasileira.....	14
2.3 – Analfabetismo no Brasil.....	15
2.3.1 - Decomposição por região e Estados:.....	16
2.3.2 - Decomposição por Faixa Etária:.....	17
2.3.3 Decomposição por renda:	17
2.4 - Efeito Escola:	18
3 - MÉTODO.....	20
3.2 Modelo BBC:	23
4 - Variáveis e Suas Fontes.....	25
4.1 – Modelos DEA	27
4.2 Função de Produção	28
5 - Resultados	30
6 - Considerações Finais:.....	33
7 – Referências	35

1 – Introdução

O nível da escolaridade brasileira, embora tenha crescido nos últimos anos, não se encontra em patamares adequados, especialmente se levarmos em conta os países desenvolvidos como medida de comparação. A educação é importante não só como meio de aumentar a produtividade do trabalhador, mas também como um meio de facilitar a convivência do indivíduo com a sociedade. Níveis maiores de educação geram cidadãos mais conscientes e externalidades positivas são criadas a partir disso, por exemplo, queda nos níveis de criminalidade e menores gastos com sistema público de saúde.

Como esperado, devido ao tamanho do território, a educação brasileira apresenta uma infinidade de problemas. Entretanto, nos últimos anos, o país obteve sucesso ao conseguir uma queda vertiginosa nos níveis de analfabetismo e abandono escolar. Muito disso, deve-se a última constituição, criada em 1988, que garante por meio do artigo 205, a educação como direito – tendo o Estado a incumbência de provê-la.

Consciente da importância da educação, esse trabalho se propõe a analisar a eficiência das escolas nas cidades de Conselheiro Lafaiete, Mariana e Ouro Preto. Para uma abordagem qualitativa do problema existe a tese popular de que o problema educacional se restringe à falta de verbas. Ressaltaremos nesse trabalho que, além disso, a questão pode se tratar também de um problema referente a alocação de recursos.

O método adotado pelo trabalho é conhecido como Análise de Dados Envoltórios, do inglês, *Data Envelopment Analysis* (DEA). Por meio deste método é possível calcular a eficiência relativa das escolas. O DEA é um modelo não paramétrico que se utiliza de produtos e insumos agregadamente para traçar o nível de eficiência do produto interessado. No caso do trabalho, a eficiência das escolas.

2 – Revisão

2.1 – Capital Humano

Um dos assuntos com maior número de estudos na economia é a teoria do crescimento econômico, essa teoria sofreu grandes alterações ao longo dos anos. Até a década de 50, acreditava-se que os fatores determinantes para que um país crescesse eram recursos naturais, capital e trabalho (SOLOW, 1956). Contudo, esse modelo passou do status de vanguarda para obsoleto quando se tornou insuficiente para explicar o que estava acontecendo. Alguns países, sobretudo os mais ricos, apresentavam rendimentos decrescentes em relação ao capital físico, mas continuavam crescendo.

A partir dos trabalhos de Mincer (1958), Schultz (1961) e Becker (1964) surge outro fator de produção capaz de impactar o crescimento econômico dos países, esse fator foi denominado como *Capital Humano*. Esse nome foi dado, pelo caráter indissociável do conhecimento de quem o carrega, uma pessoa nunca deixará de portar aquilo que aprendeu.

A renúncia da remuneração no presente para estudar e se formar e com o objetivo de ganhar mais futuramente foi apresentada como conceito pela primeira vez por Mincer (1958), o aperfeiçoamento profissional adquirido por meio do conhecimento traria ganhos de produtividade a quem o adquiriu Schultz (1964).

Para Becker et al (1999), o investimento em capital humano tem benefícios que se mostram superiores aos aspectos individuais e monetários. Portanto, pode-se notar um ganho coletivo além do ganho individual.

Recentemente, alguns autores, como Pritchett (2001), alegam que o capital humano não tem relevância como causador de crescimento econômico. Esse argumento foi rebatido por uma série de autores, entre eles Nabakashi e Figueiredo (2008). Para estes últimos autores, o diagnóstico feito por Pritchett só é possível devido a definição deficiente de

uma *proxy* para *Capital Humano* e por desconsiderar os aspectos qualitativos da educação, focando quase que de maneira exclusiva no aspecto quantitativo.

Saindo do espectro da produção, Almeida e Pereira (2000) demonstram como o capital humano pode ser responsável por trazer maior igualdade entre as classes da população. O aumento no nível educacional torna-se gerador de bem-estar tanto a nível individual, quanto a nível coletivo. Uma educação boa é algo mais abrangente que um bom nível de instrução, que normalmente torna-se um meio para que se possa atingir maiores níveis de produção. Além disso, pessoas mais educadas, normalmente, são mais responsáveis em relação a saúde e outros aspectos.

Investir em capital humano gera transbordamentos que afetam positivamente outras áreas, como já foi dito. Barros, Henriques e Mendonça (1996) demonstram como esses efeitos são sentidos para aumentar a produtividade, reduzir a expectativa de vida e diminuir o número de filhos por famílias pobres – um dos grandes causadores da desigualdade entre classes – e também impacta na redução da mortalidade infantil.

A alocação de investimentos em capital humano, quando feita antes dos 5 anos de idade, mostra-se ainda mais eficiente. Ao estudar sobre o *Perry Preschool Project*, em 1962, onde um grupo de 162 pessoas foi dividido em dois aleatoriamente,¹ James Heckman (2000) chegou à conclusão que o investimento em educação gera um retorno entre 7 e 10% ao ano. É importante ressaltar a qualidade deste trabalho devido ao caráter de controle-tratamento que foi adotado a perspectiva de realizar um trabalho onde os participantes detinham QI parecidos.

Embora seja amplamente aceita no ambiente acadêmico, a teoria do capital humano não se apresenta como unanimidade. Sen (1997) apresenta sua crítica através de um caminho mais filosófico. Para o economista, a educação muitas vezes é vista apenas como um meio, com intuito de gerar produção, e deveria ser vista como uma finalidade.

Outros autores alegam que o investimento em educação gera maior produtividade, mas não por si só, muito se deve ao caráter disciplinador imposto nas escolas que visa

¹ No *Perry Preschool Project* um dos grupos de pessoas recebeu educação pré-escolar enquanto o outro não.

moldar o ser humano para que ele se torne uma máquina voltada a produção. Teorias como a do *filtro*, *socialização*, *segmentação* e a *credencialistas* também são críticas a teoria do capital humano.

Pode-se concluir que, independente das vertentes contrárias a teoria, a teoria do capital humano apresenta uma grande relevância no meio acadêmico e pode também explicar o crescimento dos países.

2.2 – Educação Brasileira

A educação brasileira tem o seu início atrelado aos padres jesuítas, no século XVII, que além das funções eclesiais se dedicavam ao ensino. Um século depois, os religiosos dessa congregação foram expulsos do país, mas como legado, deixaram a relação entre educação e igreja.

Só a partir de 1934, com a nova constituição, a educação ganhou caráter universal, o intuito por trás disso consistia em reduzir o grau de analfabetismo da população e aumentar o grau de atendimento escolar, que se encontrava-se em patamares extremamente elevados. Embora esse problema venha sendo reduzido ao longo dos anos, quando comparado a países desenvolvidos, o Brasil ainda encontra-se distante de níveis aceitáveis.

A Constituição Federal de 1988 garante, através do artigo 5, que: “Todo homem tem direito à vida, à liberdade e à segurança pessoal.” O mesmo artigo cita a igualdade dos indivíduos diante da lei, por meio do seguinte fragmento: “Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade”. A CF de 88 representa um marco na história do país. A partir dela, direitos imprescindíveis para o bom funcionamento de uma democracia liberal foram garantidos e tornaram-se mais claros quanto a sua importância.

Em Stiglitz, Sen, Fitoussi (2009) foram demonstrados uma série de fatores capazes de influenciar no nível de qualidade de vida das pessoas. Até então, a economia baseava-se

quase que de maneira exclusiva sobre o conceito de bem-estar, muito ligado à teoria clássica e ao arcabouço teórico utilitarista. O primeiro fator apresentado no *Relatório da Comissão Sobre a Mensuração de Desempenho Econômico e Progresso Social* diz respeito sobre a importância da saúde. Seguida por educação, instituições firmes, capazes de garantir segurança física e econômica. Por fim, apresenta-se a importância de reduzir a desigualdade, não só a nível financeiro, mas também a nível social.

2.3 – Analfabetismo no Brasil

O problema referente ao analfabetismo em solo brasileiro, vem sendo atenuado ao longo das últimas décadas. Ao considerar a população de até 15 anos, a queda do nível de analfabetismo em relação ao século passado, apresenta uma trajetória de queda acentuada. 65,3% em 1900 e 13,6% em 2000. Embora proporcionalmente tenha havido queda, quando analisamos os dados em termos absolutos nota-se crescimento no número geral de analfabetos. Grande parte desse crescimento decorre do aumento significativo da população durante o século XX.

Segundo os dados do IBGE (2000), existem, no Brasil, 16 milhões de analfabetos. Entende-se como alfabetizada, para o IBGE, “pessoa capaz de ler e escrever pelo menos um bilhete simples no idioma que conhece”. Ao redor do mundo, academicamente, trabalha-se de maneira usual com o termo analfabeto funcional – aqueles com menos de quatro séries concluídas. Se passarmos a considerar o número de analfabetos funcionais, o número de brasileiros nessa condição passa a ser de 30 milhões.

Tabela 1 - Taxa de Analfabetismo

Ano	População de 15 anos ou mais		
	Total ⁽¹⁾	Analfabeta ⁽¹⁾	Taxa de Analfabetismo
1900	9.728	6.348	65,3
1920	17.564	11.409	65,0
1940	23.648	13.269	56,1
1950	30.188	15.272	50,6
1960	40.233	15.964	39,7
1970	53.633	18.100	33,7
1980	74.600	19.356	25,9
1991	94.891	18.682	19,7
2000	119.533	16.295	13,6

Fonte: IBGE, Censo Demográfico.

Nota: (1) Em milhares

Fonte: Mapa do Analfabetismo no Brasil – INEP

2.3.1 - Decomposição por região e Estados:

O nordeste brasileiro é a região onde o maior número de analfabetos é encontrado. Essa mesma região, quando comparada às demais regiões brasileiras, apresenta o menor nível de desempenho econômico e a menor complexidade. Portanto, é evidente que há uma correlação de nível de educação e de desempenho econômico. Ao passarmos das regiões aos estados, através dos dados fornecidos pelo IBGE, é possível observar quais são os 5 estados com maior número de pessoas sofrendo com alfabetização precária, são eles: Bahia, São Paulo, Minas Gerais, Pernambuco e Ceará.

Tabela 2 - Taxa de Analfabetismo por Região

Unidade Geográfica	Ano		
	1996	1998	2001
Brasil	14,7	13,8	12,4
Norte	12,4	12,6	11,2
Nordeste	28,7	27,5	24,3
Sudeste	8,7	8,1	7,5
Sul	8,9	8,1	7,1
Centro-Oeste	11,6	11,1	10,2

Fonte: IBGE, Pnads de 1996, 1998 e 2001.

Retirado de: Mapa do Analfabetismo no Brasil – INEP ²

² Complexidade Econômica: Mede a sofisticação da estrutura produtiva de um país.

2.3.2 - Decomposição por Faixa Etária:

Se considerarmos a questão do analfabetismo por faixa etária, usando os dados das PNADS de 1996, 1998 e 2001. Nota-se que há queda para todas as faixas de idade, mas para as pessoas mais novas a queda foi ainda maior, e menos sensível para as faixas que englobam pessoas com maior idade.

Tabela 3 - Taxa de Analfabetismo por Idade

Faixa Etária	Ano		
	1996	1998	2001
10 a 14	8,3	6,9	4,2
15 a 19	6,0	4,8	3,2
20 a 29	7,6	6,9	6,0
30 a 44	11,1	10,8	9,5
45 a 59	21,9	20,1	17,6
60 e mais	37,4	35,9	34,0

Fonte: IBGE, Pnads 1995, 1998 e 2001.

Nota: Excluída população rural de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá.

Retirado de: Mapa do Analfabetismo no Brasil – INEP

2.3.3 Decomposição por renda:

O número de analfabetos é bem superior para as pessoas que ganham até um salário mínimo e até mesmo, de um a três salários mínimos. Pode-se notar uma correlação negativa entre renda e nível de analfabetismo. Quanto maior a renda familiar, menor o número de analfabetos. Apenas 1,4% dos analfabetos ganham mais que 10 salários mínimos, enquanto 48,50% dos analfabetos ganham até um salário mínimo, ou entre 1 e 3 salários.

Ao longo dos anos observou-se uma queda significativa do nível de analfabetismo para as famílias mais ricas, que não foi acompanhada para as demais classes. José Ricardo Pires de Almeida (2000) retrata que no Brasil Colônia, era relativamente comum

encontrar pessoas abastadas e analfabetas. Hoje em dia, como demonstram os números, tal fato tornou-se incomum.

Tabela 4 - Taxa de Analfabetismo em Relação a Renda

Unidade Geográfica	Rendimento Domiciliar em Salário Mínimo Corrente ⁽¹⁾⁽²⁾					
	Total	Até 1	Mais de 1 até 3	Mais de 3 até 5	Mais de 5 até 10	Mais de 10
Brasil	12,4	28,8	19,7	9,7	4,7	1,4
Norte	11,2	22,6	15,5	9,9	5,0	2,0
Nordeste	24,3	36,8	29,3	17,2	8,4	1,8
Sudeste	7,5	20,0	13,5	7,5	4,0	1,5
Sul	7,1	19,5	12,4	5,9	3,6	0,8
Centro-Oeste	10,2	23,3	15,3	8,9	5,0	1,4

Fonte: IBGE, Pnad 2001.

Retirado de: Mapa do Analfabetismo no Brasil – INEP

2.4 - Efeito Escola:

O Indicador de efeito Escola é um tópico imprescindível quando se trata de educação, devido a inovação trazida pelo mesmo. Esse indicador, não se atém somente ao desempenho e as variáveis relacionadas diretamente a escola, mas também a aspectos sociais que permeiam a vida social do aluno. É importante levar em conta o seu perfil socioeconômico. Por exemplo, uma família que tem hábito de leitura, em teoria deveria formar um aluno mais hábil, ou pelo menos, com maior número de horas dedicadas aos livros.

A preocupação com o *background* familiar dos estudantes surge pela primeira vez através do trabalho de Coleman (1966), onde a qualidade das escolas é relativizada, e as variáveis que influenciam o desenvolvimento pessoal do aluno seriam determinantes para o desempenho escolar. A partir disso, criou-se a consciência da real importância de variáveis externas ao ambiente da escola.

O indicador permite que se faça uma comparação mais ampla, agrupando os avaliados com perfis semelhantes e dando pesos diferentes aos discrepantes à média. Esse método permite que os erros sejam minimizados, evitando-se de sobrevalorizar ou subvalorizar os dados discrepantes.

3 - MÉTODO

O DEA (*Data Envelopment Analysis*) se baseia nos princípios da teoria da produção para auferir eficiência entre as unidades produtivas. Esse mecanismo utiliza vários conceitos das disciplinas de microeconomia.

O DEA ganhou notoriedade a partir do final da década de 70 com os trabalhos de Rhodes, Cooper e Charnes (1978). Os estudiosos citados pavimentaram um novo caminho através da utilização desta metodologia. O DEA será exposto mais adiante ainda nesse capítulo e visa medir a produtividade e a eficiência de serviços ligados ao setor público.

No que tange aos aspectos quantitativos do modelo, destaca-se o fato dele não ser paramétrico, isto é, não há parâmetros para se submeter aos testes de confiança. O DEA não tem como objetivo achar tendências centrais dos dados, mas sim comparar a produtividade por meio das unidades encontradas que tomam decisões (chamadas pela área de *Decision Making Units*, DMUs), que serão utilizadas como referência.

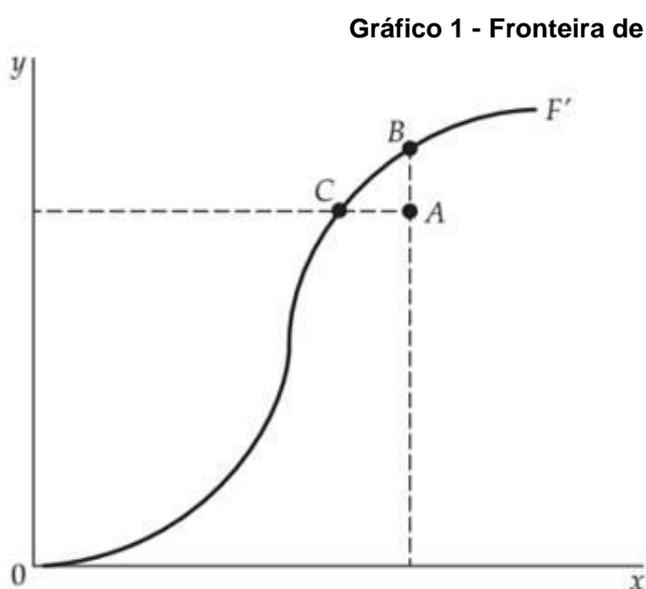
Autores como Ferreira e Gomes (2009) apresentam quatro vantagens na utilização do DEA: 1) Encontrar as DMUs mais eficientes; 2) Avaliar os *benchmarks* - cria-se uma maneira de comparar as referências; 3) Permite a utilização de diferentes insumos para níveis diferentes em uma mesma fronteira; e 4) não exige que as relações entre as variáveis sejam pré-estabelecidas (ou pré-conhecidas).

As curvas de eficiência do modelo podem apresentar rendimentos constantes, crescentes e decrescentes de escala. Além disso, pode-se orientar o modelo por *inputs* (insumos) ou por *outputs* (produtos). O primeiro modo consiste em avaliar a eficiência focando na utilização dos insumos, enfatizando sua redução, já o segundo manteria o foco no manejo dos produtos, visando seu aumento. Embora sejam abordagens diferentes, pode-se chegar a resultados iguais ou semelhantes, dependendo da escala da função de produção, como explicaremos a seguir.

De acordo com Ferreira e Gomes (2009 pp. 23-24), os conceitos importantes para compreensão do DEA são:

- **Eficácia:** Preocupa-se com os fins, visa atingir um objetivo específico. Não tem como preocupação central o processo que culmina com a produção final.
- **Produtividade:** Relaciona-se com uma maneira de pensar mais processual, associa-se com a preocupação em relação a utilização dos insumos. É usualmente representada pela razão entre produto e insumos.
- **Eficiência:** A eficiência diz respeito à relação de Produtividade que evita o desperdícios. Ou seja, a relação de produtividade “mais adequada” para a realização de uma determinada tarefa, ela consiste na comparação entre as razões de produto e insumos observados e desejados.

O gráfico 2 tem o intuito de facilitar a compreensão dos conceitos citados anteriormente. Os pontos B e C são pontos eficientes, o ponto A, por sua vez, é um ponto de ineficiência. Tendo em vista, que com a mesma quantidade de insumo poderia se produzir uma quantidade maior do produto Y. Se traçássemos uma reta partindo do ponto 0, que tangenciasse os pontos C e B, poderíamos dizer que 0C representaria a produtividade média. Até o ponto C se apresentaria de forma crescente, a partir dele se tornaria decrescente.



Fonte: Coelli *et al* (1998)

O gráfico 2 nos permite ainda, ver que o produtor teria duas escolhas se partisse do ponto A. A primeira consistiria em se deslocar em direção ao ponto C, abordagem conhecida como *orientação por insumo*. Já a segunda consistiria em produzir mais com a mesma quantidade de insumos, *orientação por produto*, caminhando de A para B.

Outra possibilidade oferecida através da análise do gráfico 2 é a de entender com maior facilidade o conceito de eficiência técnica. Que pode ser calculada por meio das formulas seguintes equações:

$$\theta_a^{produto} = \frac{OA}{OB} - \text{através dessa fórmula torna-se possível o cálculo da eficiência}$$

técnica, nesse caso a orientação acontece por meio do produto. Isto é, visa-se um aumento da produção.

$$\theta_a^{insumo} = \frac{OC}{OA} - \text{Nesse caso, a orientação ocorre através dos insumos. Ou seja, prioriza-}$$

se a utilização dos insumos de maneira eficiente.

3.1 Modelo CCR:

O Modelo CCR, formulado por Charles, Cooper e Rhodes (1979), pode ser proposto com a seguinte notação, segundo Ferreira e Gomes (2009):

Temos que o M representa o número total de produtos e o R representa a quantidade total de insumos. O μ e o ϑ representam os pesos, o primeiro se refere aos pesos para os produtos e o segundo referente para os insumos.

$$\begin{aligned} \text{Max } E_0 &= \sum_{j=1}^m \mu_j y_{j0} \\ \text{s.a } &\sum_{i=1}^r \vartheta_i x_{i0} \\ &\sum_{j=1}^m \mu_j y_{jk} - \sum_{i=1}^r \vartheta_i x_{ik} \leq 0, \forall i, j \\ &\mu_j, \vartheta_i \geq 0, \forall i, j \end{aligned}$$

As DMUs, que consistem no intuito do trabalho, podem ser agrupadas em conjuntos, podendo ser representadas através de um conjunto formado por N DMUs $k = \{1, 2, 3, \dots, N\}$.

Pode-se representar de forma alternativa. Por exemplo, partindo do princípio que existe uma DMU₀ (objetivo), os seus produtos seriam representados por y_{m0} , e x_{r0} . O m presente na equação tem o intuito de se referir aos produtos, já o r refere-se aos insumos.

$$\begin{aligned} \text{Max : } E_0 &= \sum_{j=1}^m \mu_j y_{j0} \\ \text{s.a } \sum_{i=1}^r \vartheta_i x_{i0} &= 1 \\ \sum_{j=1}^m \mu_j y_{jk} - \sum_{i=1}^r \vartheta_i x_{ik} &\leq 0, \forall_{i,j} \\ \mu_j, \vartheta_i &\geq 0, \forall_{i,j} \end{aligned}$$

Novamente, μ e ϑ representam os pesos. O intuito é que eles sejam encontrados, para que se possa fazer a maximização da divisão entre produto virtual e insumo virtual. Através de trabalhos envolvendo programação matemática, é possível que cada DMU tenha pesos diferentes. Sem isso, todas as DMUs carregariam os mesmos pesos. Com a possibilidade das *Decision Making Units* terem pesos diversos é possível constatar que quando o valor E_0 da DMU₀ é inferior a 1 denota-se ineficiência. Ou seja, os insumos utilizados ao longo do processo não estão sendo utilizados de maneira correta. Se for exatamente igual a 1, $E_0 = 1$, diz-se que a DMU está utilizando os insumos da melhor maneira encontrada.

3.2 Modelo BBC:

O Modelo BBC, criado após o modelo CCR, permite que DMUs diferentes apresentem rendimentos de escalas diferentes – constantes, crescentes ou decrescentes.

Isso acontece por causa de diversas peculiaridades entre as unidades comparativas, por exemplo, tamanho de estrutura, número de empregados e diversos outros fatores. Matematicamente o modelo pode ser proposto da seguinte forma:

$$\text{Max } \mu, \vartheta = E_{f_0} = \sum_{j=1}^m \mu_j y_{j_0} + \mu_0$$

$$\text{s. a. } \sum_{i=1}^r \vartheta_i x_{i_0} = 1$$

$$\sum_{j=1}^m \mu_j y_{jk} - \sum_{i=1}^r \vartheta_i x_{ik} + \mu_0 \leq 0, \forall k$$

$$\mu_j, \vartheta_i \geq 0(\varepsilon) \forall i, j$$

4 - Variáveis e Suas Fontes

Nessa sessão serão apresentadas as variáveis utilizadas para a realização do trabalho e a origem das mesmas. O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) foi o responsável por fornecer os dados escolhidos para a sua realização. Devido às exigências da monografia, optou-se pelas bases de dados do Microdados do Censo Escolar e a prova do Sistema Nacional de Avaliação da Educação Básica (SAEB), em particular a Prova Brasil, realizada censitariamente para escolas com turmas com mais de 20 alunos nas etapas de ensino avaliadas. Tanto a Prova Brasil quanto o Censo Escolar utilizados nesse trabalho são do ano de 2017. A partir das bases de dados citadas, foram separadas as variáveis em insumos e produtos com o objetivo de auferir a eficiência das escolas nas cidades de Conselheiro Lafaiete, Ouro Preto e Mariana.

Os insumos utilizados no trabalho foram: *Laboratório de Informática, Laboratório de Ciências, Quadra de Esportes, Biblioteca, Sala de Leitura, Auditório, Número de Computadores, Número de Funcionários, Atividades Extracurriculares, Número de Matrículas e Número de Turmas*. Os produtos, por sua vez, são mensurados através da proficiência média nas disciplinas de Português e Matemática da Prova Brasil, que tipicamente variam em uma escala de 0 a 500.

No princípio, o trabalho encontrava-se em um estágio com 59 observações. Ao dividir este número por 5, como sugere a teoria de Nunamaker (1985), chegamos ao resultado de 11,80. Como o resultado final não representa um número inteiro, optou-se por trabalhar com um número inteiro, onze (ver tabela 5 e tabela 6). Sete dessas 11 variáveis retratadas são *dummies* – variáveis binárias. Devido a isso, os 4 modelos utilizados ao longo do trabalho utilizam-se, exclusivamente, de variáveis não *dummies*.

Os produtos foram mantidos em todos os 4 modelos realizados. Esses produtos são a média de resultados da Prova Brasil em Leitura (Língua Portuguesa) e Matemática. O resultado é, portanto, a média da escola nesses dois quesitos. Abaixo, a tabela 5 traz a descrição dos insumos e o nome das variáveis no banco e a tabela 6 apresenta o mesmo para os produtos.

Tabela 5 - Insumos

Insumos	Nome das Variáveis
Laboratório Informática	IN_LABORATORIO_INFORMATICA
Laboratório Ciências	IN_LABORATORIO_CIENCIAS
Quadra de Esportes	IN_QUADRA_ESPORTES
Biblioteca	IN_BIBLIOTECA
Sala de Leitura	IN_SALA_LEITURA
Auditório	IN_AUDITORIO
Número de Computadores	NU_COMPUTADOR
Número de Funcionários	NU_FUNCIONARIOS
Atividades Complementares	TP_ATIVIDADE_COMPLEMENTAR
Número de Matrículas	NU_MATRICULAS
Número de Turmas	QUANTAS_TURMAS

Fonte: Elaboração do Autor

Tabela 6 – Produtos

Produtos	Nome das Variáveis
Proficiência em Português	PROFICIENCIA_LP_SAEB
Proficiência em Matemática	PROFICIENCIA_MAT_SAEB

Fonte: Elaboração do Autor

4.1 – Modelos DEA

Ao longo do trabalho realizaram-se 4 modelos, denominados de: *Modelo Muito Simples*, *Modelo Simples*, *Modelo Normal* e *Modelo Completo*. O primeiro deles contendo 1 insumo, o segundo contendo 2 insumos, o terceiro 3 insumos e o quarto modelo, 4 (quatro). A adoção de mais de um modelo, ao longo do trabalho, deve-se ao fato do aumento de eficiência decorrente do aumento de variáveis. Devido ao foco da monografia, optamos exclusivamente pela orientação por produto (*goal-oriented*) – que prioriza o resultado final. Os aspectos processuais referentes à utilização dos insumos foram deixados de lado.

Tabela 7 - Modelos e Variáveis

Modelo	Nome das Variáveis
Modelo Muito Simples	NU_MATRICULAS
Modelo Simples	NU_MATRICULAS, NU_FUNCIONARIOS
Modelo Normal	NU_MATRICULAS, NU_FUNCIONARIOS, QUANTAS TURMAS
Modelo Completo	NU_MATRICULAS, NU_FUNCIONARIOS, QUANTAS_TURMAS, NU_COMPUTADORES

Elaboração do Autor

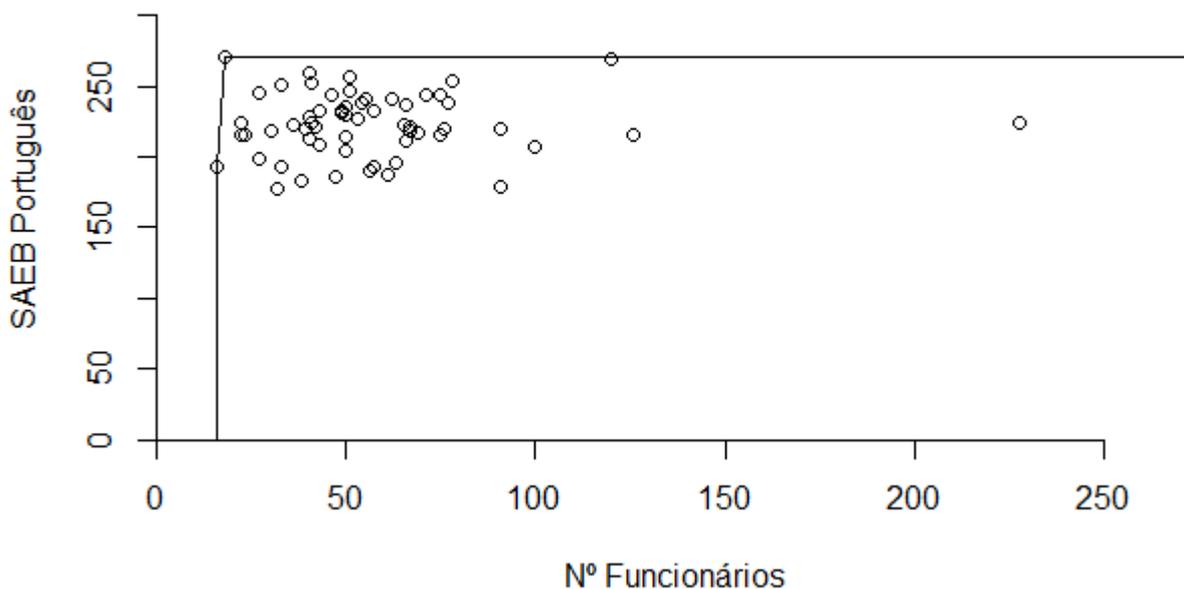
O primeiro modelo recebeu a alcunha de *Modelo Muito Simples*. Os seus produtos, assim como para os demais, são *Proficiência em Português e Proficiência em Matemática*. Diferenciando-se, portanto, quanto aos insumos utilizados para o cálculo das eficiências. O primeiro modelo, como ilustrado na tabela acima, utiliza-se exclusivamente do insumo *número de matriculas*. No Segundo, *Modelo Simples*, as variáveis utilizadas foram *número de matriculas e número de funcionários*. Para o terceiro, *modelo normal*, os insumos utilizados no *Modelo Simples* foram repetidos e adicionou-se a variável responsável por medir o *número de turmas*. O *modelo completo*, o último a ser rodado, utilizou-se das três variáveis do *Modelo Normal*, completado por o insumo *número de computadores*.

4.2 Função de Produção

Ao considerar os insumos necessários para a produção de vários produtos, o modelo de análise envoltória (DEA), caracteriza-se pela facilidade em criar funções de produções, que por sua vez, retratam a eficiência dos pontos também conhecidos como *Decision-Making-Units* (DMUs). Como já dito anteriormente nesse mesmo trabalho, os pontos sobre a fronteira de produção são os mais bem sucedidos.

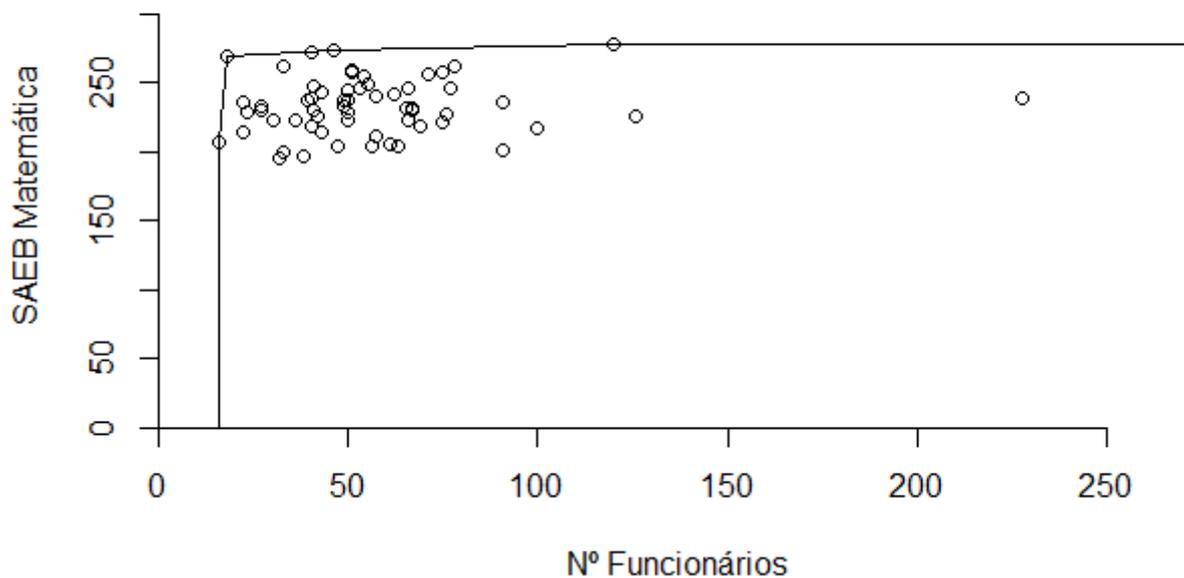
A fim de exemplificar o conteúdo previamente apresentando por vias teóricas, podemos inserir graficamente as DMUs e uma fronteira parcial do modelo mais simples. Os gráficos a seguir representam as notas médias das escolas, usando exclusivamente número de funcionários como insumos. Trata-se um gráfico de um modelo DEA-BCC, ou seja, rendimentos variáveis de escala, para um insumo e um produto.

Gráfico 2 - Função de Produção - Português SAEB



Fonte: Elaboração Pelo Autor

Gráfico 3 - Função de Produção Matemática - SAEB



Fonte: Elaboração do Autor.

Sabemos da possibilidade de expor cada função de produção e a sua gama de combinações diante dos insumos. Entretanto, optamos por demonstrar os gráficos com intuito único de facilitar a compreensão e visualização do conteúdo, respeitando o foco e escopo inicial da monografia.

5 - Resultados

Nessa sessão, pretende-se discutir os resultados observados através da utilização do DEA. Deve-se ressaltar que o modelo possui estrutura não paramétrica e embora possa ser utilizado de diversas formas, optamos por uma abordagem *output-oriented*, priorizando o pragmatismo dos resultados, ou seja, o modelo prioriza a melhor dos resultados na Prova Brasil.

Os resultados da tabela 8 representam o número de escolas eficientes levados em conta a totalidade de colégios nessas cidades. Por exemplo, em *Efic1*, o número de escolas eficientes são 4, quando dividimos pelo total de centros escolares disponíveis em Conselheiro Lafaiete (28) chegamos ao resultado de *Efic1*, que é o percentual de eficiência daquele município. Os demais resultados da tabela seguiram exatamente os mesmos critérios para serem calculados.

A utilização de quatro modelos diferentes nos permite a obtenção de quatro resultados de eficiência diversos por município. Em todos eles, Conselheiro Lafaiete demonstra maior eficiência. Outro fator facilmente observável por meio da imagem é o aumento de eficiência, ao passo que se aumenta o número de insumos.

Tabela 8 - Eficiência Proporcional

Cidades	<i>Efic1</i>	<i>Efic2</i>	<i>Efic3</i>	<i>Efic4</i>
Conselheiro Lafaiete	14,3%	17,9%	25,0%	42,9%
Mariana	0,0%	8,3%	8,3%	16,7%
Ouro Preto	10,5%	21,1%	21,1%	26,3%

Fonte: Elaboração do Autor

A tabela 9 apresenta a eficiência média das escolas em Conselheiro Lafaiete, Mariana e Ouro Preto. Essa tabela diferencia-se da anterior por levar em conta todas as escolas disponíveis para as cidades analisadas, enquanto a primeira foca exclusivamente nas eficientes. Interpreta-se como tal, apenas resultados exatamente iguais 1. Quanto maior o distanciamento desse referencial, maior o grau de ineficiência, ou seja, índices muito maiores do que 1 representam maior grau de ineficiência.

Tabela 9 - Eficiência Média

Cidades	Eficiência Média			
	<i>Efic1</i>	<i>Efic2</i>	<i>Efic3</i>	<i>Efic4</i>
Conselheiro Lafaiete	1,133	1,113	1,108	1,186
Mariana	1,217	1,206	1,202	1,398
Ouro Preto	1,175	1,165	1,158	1,232

Fonte: Elaboração do Autor

Independente de analisarmos separadamente eficiência em proporção de número de escolas e da eficiência média das escolas, Conselheiro Lafaiete ocupa o primeiro lugar para todos os modelos, seguido respectivamente por Ouro Preto e Mariana. As escolas de Mariana fiam sempre na última posição em eficiência.

O nível de eficiência do produto no DEA é resultante da melhor aquisição de resultados diante de determinado nível de utilização dos insumos, isso demonstra que Conselheiro Lafaiete utiliza de forma mais adequada (segundo as especificações da fronteira) os recursos a sua disposição. Em *Efic4*, por exemplo, adicionamos o insumo *número de computadores*, que impactou numa variação positiva de 0,1962 para o município de Mariana, demonstrando que essa variável não desempenha grande importância para determinar o nível de proficiência em matemática e português dos alunos.

Na tabela 10 representamos as variáveis não *dummies* e seus valores médios. Mariana apresenta níveis superiores em *quantidades de turmas* e muitos mais elevados para *número de funcionários*, o que possivelmente explicaria a baixa eficiência em relação aos municípios comparados. Em negrito estão os valores mais altos para cada uma das variáveis de insumos.

Tabela 10 - Resultado Médio dos Insumos

Média dos Insumos	Cons. Lafaiete	Mariana	Ouro Preto
Número de Computadores	17,43	19,5	19,95
Número de Funcionários	55,75	67,5	51,26
Número de Matrículas	21,76	20,0	18,88
Quantidade de Turmas	2,04	2,33	1,68

Fonte: Elaboração do Autor

Na tabela 11, apresentamos os valores médios dos produtos para cada município.

Municípios	Produtos	
	Média Português	Média Matemática
Conselheiro Lafaiete	230.9	240.9
Mariana	211.8	223.8
Ouro Preto	218.0	226.5

Fonte: Elaboração do Autor

6 - Considerações Finais:

A realização do trabalho tornou possível a observação da variação comportamental causada pelos insumos nos produtos, em particular, a proficiência em Português e Matemática. Priorizando as turmas de 5º ano do ensino fundamental, nas cidades de Conselheiro Lafaiete, Mariana e Ouro Preto.

A fim de mensurar a eficiência das escolas, escolhemos o modelo de Análise Envoltória dos Dados (DEA), orientando-se em relação ao produto, sempre apresentado em média por escola durante a realização dessa monografia.

O trabalho não pretende assumir um intuito prescritivo nem definitivo sobre o funcionamento das escolas analisadas. Sabe-se da limitação de análises baseadas em resultados médios, e da exclusão de variáveis socioeconômicas, cujo papel é sabidamente imprescindível.

O papel da educação transcende o escopo quantitativo e também a importância do desempenho em exames escolares realizados pontualmente. Trata-se, portanto, de um meio pelo qual o cidadão insere-se na sociedade e obtém dignidade. Os ganhos educacionais extrapolam a esfera individual e impactam a sociedade de maneira coletiva, como ressaltado por Becker (1999). Em Sen (1997), a *Educação* é vista como uma ferramenta, pela qual o ser humano pode viver de maneira mais prazerosa e produtiva, contrapondo-se a ideia de que sua principal função seria aumentar a produtividade do trabalhador.

O investimento em *Capital Humano* pode representar a oportunidade de aproximação do Brasil em relação aos países desenvolvidos. Essa diferença é resultado de um longo processo histórico, de investimentos insuficientes e de pouca eficiência alocativa. As cidades de Conselheiro Lafaiete, Mariana e Ouro Preto foram selecionadas devido ao relacionamento pessoal do graduando com as mesmas, motivando o processo de entendimento por trás da real situação escolar vivenciada.

Embora apresente número médio inferior de insumos, o município de Ouro Preto demonstra-se menos eficiente que o município de Conselheiro Lafaiete, valorizando ainda mais a relação do último com as variáveis insumos. Intuitivamente, um número inferior de insumos facilitaria seu manejo, resultando em maior nível de produto. Fato que não se observa no trabalho.

Os resultados decepcionantes referentes à cidade de Mariana devem-se ao excesso de variáveis com pouca contribuição, como *número de funcionários* e também, a má utilização dos recursos disponíveis. O que pode ser exemplificado, através da inserção da variável *número de computadores*, cujo impacto sobre o produto foi bastante insignificante. Além dos computadores, qualquer outra variável, se má utilizada surtiria pouco efeito para aumentar os níveis de eficiência.

Para expor ações e medidas capazes de melhorar a alocação de verbas destinadas a educação, seria necessário a realização de outro trabalho, orientado prioritariamente pelos insumos (*input-oriented*). Abordagem com maior cuidado em relação aos aspectos processuais, não representando a prioridade delineada pela monografia.

7 – Referências

ALMEIDA, José Ricardo Pires de. **Instrução pública no Brasil (1500-1889): história e legislação**. São Paulo: EDUC, 2000.

BANKER, R. D.; CHARNES, H.; COOPER W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. **Management Science**, v. 30, n. 9, p. 1078-1092, 1984.

BARROS, R. P. de, MENDONÇA, R. **Salário e educação no Brasil**. Rio de Janeiro: IPEA, 1996.

BECKER, G. S. Human Capital: **A Theoretical and Empirical Analysis with Special Reference to Education**. 3 ed. Chicago: The University of Chicago Press, 1964.

BECKER, Gary S.; GLAESER, Edward L.; MURPHY, Kevin M. Population and economic growth. **American Economic Review**, v. 89, n. 2, p. 145-149, 1999.

CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, v. 2, 1978.

COELI, T. J. et al. **An Introduction to Efficiency and Productivity Analysis**. 2 ed. Nova York: Editora Springer Science & Business Media, Inc. 2005.

COLEMAN, James S. et al. York. 1966. **Equality of educational opportunity**, v. 2, 1966.

FERREIRA, C. M. C.; GOMES, A. P. **Introdução à Análise Envoltória de Dados: Teoria, Modelos e Aplicações**. 1 ed. Viçosa-MG: Editora UFV, 2009.

HECKMAN, James J. Causal parameters and policy analysis in economics: A twentieth century retrospective. **The Quarterly Journal of Economics**, v. 115, n. 1, p. 45-97, 2000.

MINCER, Jacob. Investment in human capital and personal income distribution. **Journal of Political Economy**, v. LXVI, n. 4, p. 281-302, 1958.

NAKABASHI, L.; FIGUEIREDO, L. de. Mensurando os impactos diretos e indiretos do capital humano sobre o crescimento. **Economia Aplicada**, v. 12, n. 1, p. 151-71, jan/mar 2008.

NUNAMAKER, Thomas R. Using data envelopment analysis to measure the efficiency of non-profit organizations: A critical evaluation. **Managerial and decision Economics**, v. 6, n. 1, p. 50-58, 1985.

PRITCHETT, L. Where has all the education gone? **The World Bank Economic Review**, v. 15, n. 3, p. 367-91, 2001.

SCHULTZ, T. W. **O Capital Humano: Investimentos em Educação e Pesquisa**. 1 ed. Nova York: The Free Press, 1971.

SCHULTZ, T. W. Investment in Human Capital. **The American Economic Review**, vol. 51, no. 1, 1961, pp. 1–17.

SCHULTZ, T. W. **O Valor Econômico da Educação**. Rio de Janeiro: Zahar Editores, 1964

SEN, Amartya. **Choice, welfare and measurement**. Harvard University Press, 1997.

SOLOW, R. M. A Contribution to the Theory of Economic Growth. **The Quarterly Journal of Economics**, Fevereiro, 1956.

STIGLITZ, Joseph; SEN, Amartya K.; FITOUSSI, Jean-Paul. **The measurement of economic performance and social progress revisited: reflections and overview**. 2009.

Apêndices

Apêndice A – Rotinas para replicação do Trabalho

Códigos do programa R para o desenvolvimento do trabalho. Na primeira rotina, organizamos os dados retirados do site do INEP, de acordo com ano e município.

```

1 getwd()
2 setwd("C:\\Users\\Usuario\\Desktop\\R - programin\\Dados")
3
4 escolas <- read.csv2("ESCOLAS.csv", sep = "|")
5
6 head(escolas)
7 dim(escolas)
8
9 # So Mariana e Ouro Preto
10
11 escolas_mariana <- escolas[escolas$CO_MUNICIPIO == 3140001, ]
12 escolas_mariana$NO_ENTIDADE
13
14 # So escolas ativas:
15
16 escolas_mariana <- escolas_mariana[escolas_mariana$TP_SITUACAO_FUNCIONAMENTO == 1,]
17
18 # Apenas escolas pùblicas e Urbanas:
19
20 escolas_mariana <- escolas_mariana[escolas_mariana$TP_DEPENDENCIA != 4 & escolas_mariana$TP_
    LOCALIZACAO == 1,]
21
22 escolas_mariana$NO_ENTIDADE
23
24 # Mesmos procedimentos para Ouro Preto
25
26 escolas_op <- escolas[escolas$CO_MUNICIPIO == 3146107, ]
27 escolas_op <- escolas_op[escolas_op$TP_SITUACAO_FUNCIONAMENTO == 1,]
28 escolas_op <- escolas_op[escolas_op$TP_DEPENDENCIA != 4 & escolas_op$TP_LOCALIZACAO == 1,]
29
30 escolas_op$NO_ENTIDADE
31
32 # Escolas de Conselheiro Lafaiete código
33
34 # Mesmos procedimentos para Ouro Preto
35
36 escolas_cl <- escolas[escolas$CO_MUNICIPIO == 3118304, ]
37 escolas_cl <- escolas_cl[escolas_cl$TP_SITUACAO_FUNCIONAMENTO == 1,]
38 escolas_cl <- escolas_cl[escolas_cl$TP_DEPENDENCIA != 4 & escolas_cl$TP_LOCALIZACAO == 1,]
39
40 escolas_cl$NO_ENTIDADE
41
42 # Vamos ler o banco de Turmas Primeiro:
43
44 turmas <- read.csv2("TURMAS.csv", sep = "|")
45
46 turmas_mariana <- turmas[turmas$CO_MUNICIPIO == 3140001,]

```

```

47 turmas_op <- turmas[turmas$CO_MUNICIPIO == 3146107,]
48 turmas_cl <- turmas[turmas$CO_MUNICIPIO == 3118304,]
49
50 # Merge Mariana
51
52 df_mariana <- merge(escolas_mariana, turmas_mariana, by = "CO_ENTIDADE", all.y = TRUE)
53 df_mariana <- df_mariana[is.na(df_mariana$NU_ANO_CENSO.x) == FALSE,]
54 table(df_mariana$TP_ETAPA_ENSINO) # Para ver se existem ainda turmas com 8 anos no Ensino
Fundamental
55 df_mariana <- df_mariana[which(df_mariana$TP_ETAPA_ENSINO == 18),]
56
57 # AGGREGATE Mariana
58 ## DF <- Data Frame
59
60 agg_m1 <- aggregate(df_mariana, by = list(df_mariana$CO_ENTIDADE, df_mariana$NO_ENTIDADE), FUN
= "mean")
61 agg_m1$NO_ENTIDADE <- agg_m1$Group.2
62 df_mariana$QUANTAS_TURMAS <- 1
63 agg_m2 <- aggregate(df_mariana[,c(1,173,254)], by = list(df_mariana$CO_ENTIDADE), FUN = "sum")
64 names(agg_m2)[3] <- "NU_MATRICULAS_SOMA"
65
66 df_mariana <- cbind(agg_m1, agg_m2[,3:4])
67 df_mariana <- df_mariana[,3:257] # So para tirar a primeira coluna. Banco de Mariana esta
pronto.
68
69 #Merge Ouro Preto
70 df_op <- merge(escolas_op, turmas_op, by = "CO_ENTIDADE", all.y = TRUE)
71 df_op <- df_op[is.na(df_op$NU_ANO_CENSO.x) == FALSE,]
72 table(df_op$TP_ETAPA_ENSINO) # Olhar se ainda existem escolas aderindo o modelo de 8 anos (Cod
4).
73 #Estamos interessados no codigo 18
74 df_op <- df_op[which(df_op$TP_ETAPA_ENSINO == 18),]
75
76 #Aggregate Ouro Preto:
77 agg_m1 <- aggregate(df_op, by = list(df_op$CO_ENTIDADE, df_op$NO_ENTIDADE), FUN = "mean")
78 agg_m1$NO_ENTIDADE <- agg_m1$Group.2
79 df_op$QUANTAS_TURMAS <- 1
80 agg_m2 <- aggregate(df_op[,c(1,173,254)], by = list(df_op$CO_ENTIDADE), FUN = "sum")
81 names(agg_m2)[3] <- "NU_MATRICULAS_SOMA"
82 df_op <- cbind(agg_m1, agg_m2[,3:4])
83
84 df_op <- df_op[,3:257]
85
86 #Merge Conselheiro Lafaiete
87
88 df_cl <- merge(escolas_cl, turmas_cl, by = "CO_ENTIDADE", all.y = TRUE)
89 df_cl <- df_cl[is.na(df_cl$NU_ANO_CENSO.x) == FALSE,]
90 table(df_cl$TP_ETAPA_ENSINO)
91
92 df_cl <- merge(escolas_cl, turmas_cl, by = "CO_ENTIDADE", all.y = TRUE)
93 df_cl <- df_cl[is.na(df_cl$NU_ANO_CENSO.x) == FALSE,]
94 table(df_cl$TP_ETAPA_ENSINO)
95 df_cl <- df_cl[which(df_cl$TP_ETAPA_ENSINO == 18),]
96 #Aggregate Conselheiro Lafaiete
97 agg_m1 <- aggregate(df_cl, by = list(df_cl$CO_ENTIDADE, df_cl$NO_ENTIDADE), FUN = "mean")
98 agg_m1$NO_ENTIDADE <- agg_m1$Group.2
99 df_cl$QUANTAS_TURMAS <- 1
100 agg_m2 <- aggregate(df_cl[,c(1,173,254)], by = list(df_cl$CO_ENTIDADE), FUN = "sum")
101 names(agg_m2)[3] <- "NU_MATRICULAS_SOMA"
102 df_cl <- cbind(agg_m1, agg_m2[,3:4])
103
104 df_cl <- df_cl[,3:257]
105
106 # Estamos com os dados todos prontos:
107 # vamos juntar os bancos e na ordem alfabetica do municipio
108
109 dados_prontos <- rbind(df_cl, df_mariana)
110 dados_prontos <- rbind(dados_prontos, df_op)
111 dim(dados_prontos)
112 write.csv(dados_prontos, file = "DADOS_ESCOLAS_QUASE_PRONTOS.csv")

```

A rotina seguinte consiste em obter os valores médios de proficiência em Português e Matemática, referentes a prova Brasil.

```
1 getwd()
2 setwd("C:\\Users\\Usuario\\Desktop\\R - programin'\\Dados\\Prova Brasil 2017\\DADOS")
3
4 prova <- read.csv(file = 'NSEpronto_4s_17.csv')
5
6 dim(prova)
7 head(prova)
8
9 prova_mariana <- prova[prova$ID_MUNICIPIO == 3140001,]
10 prova_op <- prova[prova$ID_MUNICIPIO == 3146107,]
11 prova_cl <- prova[prova$ID_MUNICIPIO == 3118304,]
12
13 agg_mar <- aggregate(prova_mariana, by = list(prova_mariana$ID_ESCOLA), FUN = "mean", na.rm =
  TRUE)
14 agg_op <- aggregate(prova_op, by = list(prova_op$ID_ESCOLA), FUN = "mean", na.rm = TRUE)
15 agg_cl <- aggregate(prova_cl, by = list(prova_cl$ID_ESCOLA), FUN = "mean", na.rm = TRUE)
16
17 agg_mar <- agg_mar[,3:39]
18 agg_op <- agg_op[,3:39]
19 agg_cl <- agg_cl[,3:39]
20
21 dados_juntos <- rbind(agg_cl, agg_mar)
22 dados_juntos <- rbind(dados_juntos, agg_op)
23 dim(dados_juntos)
24
25 write.csv(dados_juntos, file = "PROVA_BRASIL_PRONTO.csv")
```

Na rotina a seguir, procuramos juntar as bases de dados previamente apresentadas, a fim de aumentar o número de insumos disponíveis para o modelo.

```
1 getwd()
2 setwd("C:\\Users\\Usuario\\Desktop\\R - programin'\\Dados")
3
4 PB <- read.csv("PROVA_BRASIL_PRONTO.csv")
5 ESCOLAS <- read.csv("DADOS_ESCOLAS_QUASE_FRONTOS.csv")
6
7 names(PB)[7] <- "CO_ENTIDADE"
8
9 dados <- merge(ESCOLAS, PB, by = "CO_ENTIDADE")
10
11 # Vamos reduzir so para as variaveis que vamos utilizar:
12
13 dados[,c(1,4,6)]
14
15 dados <- dados[,c(1,4:6,13:14,25:50)]
```

Nessa última rotina, separamos as variáveis consideradas importantes para rodar os quatro modelos responsáveis por balizar o trabalho.

```
1 library(Benchmarking)
2 library(tidyverse)
3 getwd()
4 setwd("C:\\Users\\[usuário]\\Desktop\\R - programa\\Dados")
5
6 PB <- read.csv("PROVA_BRASIL_PRONTO.csv")
7 ESCOLAS <- read.csv("DADOS_ESCOLAS_QUASE_PRONTOS.csv")
8
9 names(PB)[7] <- "CO_ENTIDADE"
10
11 dados <- merge(ESCOLAS, PB, by = "CO_ENTIDADE")
12
13 # Vamos reduzir ao para as variáveis que vamos utilizar:
14
15 dados <- dados[,c
16   (1,4,6,13:15,27:36,38:50,51,62:75,80:87,101:106,111:114,119:129,139,143:145,148:149,160,16
17   )
18   ]
19 head(dados)
20
21 # selecionando variáveis
22
23 dados2 <- select(dados, IN_LABORATORIO_INFORMATICA, IN_LABORATORIO_Ciencias,
24   IN_QUADRA_ESPORTES, IN_BIBLIOTECA, IN_SALA_LEITURA, IN_AUDITORIO,
25   NU_COMPUTADOR, NU_FUNCIONARIOS, TP_ATIVIDADE_COMPLEMENTAR, NU_MATRICULAS,
26   QUANTAS_TURMAS)
27
28 dados3 <- select(dados, PROFICIENCIA_LP_SABE, PROFICIENCIA_MT_SABE)
29
30 # Variável de Infraestrutura
31
32 infra <- apply(INSUMOS[,c(1:6)], MARGIN = 1, FUN = sum)
33 dados2$infra <- infra
34
35 INSUMOS <- as.matrix(dados2)
36 PRODUTOS <- as.matrix(dados3)
37
38 table(infra)
39
40 # como observar esses os pontos na fronteira?
41 # Fazendo alguns Gráficos:
42
43 # dea
44
45 dea.plot(x = INSUMOS[,8], y = PRODUTOS[,1], xlab = "N Funcionarios", ylab = "SAEB Portugues")
46
47 # Para Matemática:
48
49 dea.plot(x = INSUMOS[,8], y = PRODUTOS[,2], xlab = "N Funcionarios", ylab = "SAEB Matemática")
50
51 text(x = INSUMOS[,8], y = PRODUTOS[,2], labels = dados$NO_ENTIDADE, cex = 0.5)
52
53 # Para algumas escolas apenas
54
55 dea.plot(x = INSUMOS[,8:head], y = PRODUTOS[,2], xlab = "N Funcionarios", ylab = "SAEB
56   Matemática")
```

```

52 text(x = INSUMOS[c(22, 36, 42, 43),8], y = PRODUTOS[c(22, 36, 42, 43),2] + 10, labels = dados$
NO_ENTIDADE[c(22, 36, 42, 43)], cex = 0.5)
53 text(x = INSUMOS[c(35),8] + 5, y = PRODUTOS[c(35),2] - 6, labels = dados$NO_ENTIDADE[c(35)],
cex = 0.5)
54 text(x = INSUMOS[c(24),8] + 6, y = PRODUTOS[c(24),2] + 10, labels = dados$NO_ENTIDADE[c(24)],
cex = 0.5)
55
56 dea.plot(x = INSUMOS[,12], y = PRODUTOS[,2], xlab = "Variavel Infra", ylab = "SAEB Matematica"
)
57
58 # valores de eficiencia (?)
59 # Mais simples de todos:
60
61 efic1 <- dea(X = INSUMOS[,10], Y = PRODUTOS, RTS = "VRS", ORIENTATION = "out")
62 summary(efic1)
63
64 # Para incluir no banco de escolas
65
66 banco_efic <- cbind(dados[,c(1,2,4)], efic1$eff)
67 names(banco_efic)[3] <- "CO_MUNICIPIO"
68 banco_efic$NO_MUNICIPIO[banco_efic$CO_MUNICIPIO == 3140001] <- "Mariana"
69 banco_efic$NO_MUNICIPIO[banco_efic$CO_MUNICIPIO == 3146107] <- "Ouro Preto"
70 banco_efic$NO_MUNICIPIO[banco_efic$CO_MUNICIPIO == 3118304] <- "Cons. Lafaiete"
71 names(banco_efic)[4] <- "efic1"
72
73 table(banco_efic$NO_MUNICIPIO, banco_efic$efic1)
74 table(banco_efic$NO_MUNICIPIO)
75
76 #Simples
77 #Codigo para rodar a eficiencia
78
79 efic2 <- dea(X = INSUMOS[,c(8,10)], Y = PRODUTOS, RTS = "VRS", ORIENTATION = "out")
80 summary(efic2)
81
82 banco_efic$efic2 <- efic2$eff
83
84 table(banco_efic$NO_MUNICIPIO, banco_efic$efic2)
85
86 efic3 <- dea(X = INSUMOS[,c(8)], Y = PRODUTOS, RTS = "CRS", ORIENTATION = "out")
87 summary(efic3)
88 banco_efic$efic3 <- efic3$eff
89
90 efic4 <- dea(X = INSUMOS[,c(7, 8, 10, 11)], Y = PRODUTOS, RTS = "CRS", ORIENTATION = "out")
91 summary(efic4)
92 banco_efic$efic4 <- efic4$eff
93
94 cor(INSUMOS[,8], PRODUTOS[,2])
95
96 cor(INSUMOS[,7], PRODUTOS[,2])
97 cor(INSUMOS[,10], PRODUTOS[,2])
98
99 plot(INSUMOS[,10], PRODUTOS[,2], xlim = c(1, 50), ylim = c(150,300), xlab = "Matriculas", ylab
= "Nota Matematica", pch = "")
100 dea.plot(x = INSUMOS[,10], y = PRODUTOS[,2], add = TRUE)
101 points(INSUMOS[,10], PRODUTOS[,2], pch = 19)
102 abline(lm(PRODUTOS[,2]~INSUMOS[,10]), col = "red")
103
104 efic <- dea(X = INSUMOS, Y = PRODUTOS, RTS = "CRS", ORIENTATION = "out")
105 summary(efic)
106

```

```

107 # Eficiencia um pouco mais completo
108
109 efic3 <- dea(X = INSUMOS[,c(8,10,11)], Y = PRODUTOS, RTS = "VRS", ORIENTATION = "out")
110 summary(efic2)
111
112 banco_efic$efic3 <- efic3$eff
113 table(banco_efic$NO_MUNICIPIO, banco_efic$efic3)
114 7/28
115 dea.plot(efic3)
116 cor(INSUMOS[,8], PRODUTOS[,2])
117
118
119 banco_efic %>%
120 group_by(NO_MUNICIPIO) %>%
121 summarise(escolas = n(),
122 eficiente_mod1 = sum(efic1 == 1) / escolas,
123 eficiente_mod2 = sum(efic2 == 1) / escolas,
124 eficiente_mod3 = sum(efic3 == 1) / escolas,
125 eficiente_mod4 = sum(efic4 == 1) / escolas
126 )
127 summary(efic1)
128
129 media_efic_mar = mean(efic1$eff[banco_efic$NO_MUNICIPIO == "Mariana"])
130 media_efic_mar
131
132
133 media_efic_op = mean(efic1$eff[banco_efic$NO_MUNICIPIO == "Ouro Preto"])
134 media_efic_op
135
136 media_efic4_mar = mean(efic3$eff[banco_efic$NO_MUNICIPIO == "Mariana"])
137 media_efic4_mar
138 media_efic4_cl = mean(efic4$eff[banco_efic$NO_MUNICIPIO == "Cons. Lafaiete"])
139 media_efic4_cl
140 media_efic4_op = mean(efic4$eff[banco_efic$NO_MUNICIPIO == "Ouro Preto"])
141 media_efic4_op
142 table(efic_1)
143 #media produtos e insumos
144 summary(INSUMOS)
145 summary(PRODUTOS)
146
147 mar <- which(banco_efic$NO_MUNICIPIO == "Mariana")
148 op <- which(banco_efic$NO_MUNICIPIO == "Ouro Preto")
149 cl <- which(banco_efic$NO_MUNICIPIO == "Cons. Lafaiete")
150 #numero medio de insumos
151 summary(INSUMOS[cl, ])
152 summary(INSUMOS[op, ])
153
154 summary(PRODUTOS[cl,])
155
156 summary(INSUMOS[mar, ])

```