

Uma análise estatística multivariada do desempenho das escolas municipais de Ribeirão Preto^{*, **}

Mozart Neves Ramos^a

Antonio José da Costa Filho^b

João Bosco Paraíso da Silva^c

Ester Fraga Vilas-Bôas Carvalho do Nascimento^d

Resumo

Esse estudo faz uma análise estatística multivariada de componentes principais (ACP) para estabelecer padrões de desempenho escolar para o 5º e o 9º ano do Ensino Fundamental das escolas municipais de Ribeirão Preto. As equações matemáticas obtidas de gráficos bidimensionais de ACP revelam que os percentuais de alunos com aprendizado adequado em língua portuguesa e em matemática são os indicadores mais significativos para explicar tais padrões de desempenho. A partir dessa análise, foi possível estabelecer vetores de desempenho escolar para cada uma das escolas dos dois anos escolares do Ensino Fundamental aqui considerados. Isso permitiu construir um indicador de heterogeneidade para as escolas da rede municipal de Educação de Ribeirão Preto.

Palavras-chave: Desempenho Escolar. Componentes Principais. Heterogeneidade Escolar. Escolas Municipais. Ribeirão Preto.

1 Introdução

Colocar numa mesma equação quantidade e qualidade talvez seja o maior desafio da Educação Pública brasileira. Ao longo das últimas décadas, o país

* Esse trabalho é dedicado à memória do professor Benício de Barros Neto, que muito nos ensinou no uso das técnicas estatísticas de análise de componentes principais.

** Ao Banco Santander, no âmbito do programa Santander Universidades, pelo apoio financeiro para o trabalho da Cátedra Sérgio Henrique Ferreira do IEA/USP – Ribeirão Preto.

^a Universidade de São Paulo, Ribeirão Preto, SP, Brasil.

^b Universidade de São Paulo, Ribeirão Preto, SP, Brasil.

^c Universidade Federal de Pernambuco, Recife, PE, Brasil.

^d Universidade Tiradentes, Aracaju, SE, Brasil.

Recebido em: 07 dez. 2020

Aprovado em: 25 fev. 2021

logrou êxito quanto ao acesso à escola, especialmente no que se refere ao Ensino Fundamental, porém, um esforço ainda precisa ser feito na Pré-escola e no Ensino Médio, se levarmos em conta que 6% das crianças de 4 e 5 anos e 7,5% de jovens de 15 a 17 anos estão fora da escola (OBSERVATÓRIO DO PLANO NACIONAL DE EDUCAÇÃO – OPNE, 2020). Agora esses percentuais são 5,9% das crianças de 4 e 5 anos e 5,5% de jovens de 15 a 17 anos estão fora da escola. Contudo, o grande esforço que o país necessita fazer concentra-se, principalmente, nos baixos índices de aprendizagens escolares, como revertê-los, torná-lo acessível a todos os alunos, pois os níveis de desigualdade educacional são enormes, tanto entre escolas de uma mesma rede de Ensino quanto entre municípios de um mesmo estado.

Para se ter uma ideia do tamanho desse desafio, vamos aqui considerar o percentual de alunos com aprendizado adequado em língua portuguesa e em matemática, em conformidade com a meta 3 estabelecida pelo movimento Todos pela Educação (2007). A meta 3 indica que para o 5º ano do Ensino Fundamental, o aprendizado é considerado adequado, em língua portuguesa e em matemática, quando o aluno obtém um desempenho no exame do Sistema de Avaliação da Educação Básica (Saeb) acima de 200 e de 225 pontos, respectivamente. Para o 9º ano do Ensino Fundamental, por sua vez, o aprendizado é considerado adequado, em língua portuguesa e em matemática, quando o aluno obtém um desempenho, respectivamente, acima de 275 e de 300 pontos no mesmo sistema de avaliação. O Saeb é realizado a cada dois anos, a fim de avaliar estudantes do 5º e do 9º ano do Ensino Fundamental, e do 3º ano do Ensino Médio, nas disciplinas de língua portuguesa e de matemática.

Por exemplo, de acordo com o portal QEdu (2020), o percentual de alunos com aprendizado adequado em matemática, ao concluir o 9º ano do Ensino Fundamental, nas escolas públicas do Brasil, é de apenas 15%; em língua portuguesa, esse percentual é de 34%. Se compararmos esses percentuais nos estados do Ceará e do Maranhão constatamos a nossa preocupação quanto à desigualdade educacional. No Ceará, os percentuais acima, em matemática e em língua portuguesa, correspondem a 20% e a 40%; por outro lado, no Maranhão, tais percentuais simplesmente desabam, ou seja, vão para 5% e para 16%, respectivamente. Mas, essas desigualdades podem também ser vistas, como dissemos, entre escolas de uma mesma rede. Por exemplo, ainda com relação ao 9º ano do Ensino Fundamental na rede municipal de São Luís (MA), os percentuais acima relativos à Escola Aluísio Azevedo são, respectivamente, 31% e 62%, enquanto para a Escola Severiano de Sousa Lima, tais percentuais são 3% e 9%.

É importante salientar que o indicador oficial de qualidade do Ensino é o Índice de Desenvolvimento da Educação Básica (Ideb), em conformidade com a meta 7 do Plano Nacional de Educação (PNE) (BRASIL, 2014). O Ideb é calculado pela multiplicação de dois indicadores: (a) a nota média obtida pelos alunos nos exames de língua portuguesa e de matemática da Prova Brasil e (b) o fluxo escolar (taxa de aprovação). Trata-se de um indicador de fácil compreensão e relativamente simples de ser calculado, mas que apresenta algumas limitações, como têm salientado alguns autores (FIGUEIREDO *et al.*, 2018; TRAVITZKI, 2020). É importante destacar que o aprendizado adequado em língua portuguesa e em matemática, considerado na meta 3 do Todos pela Educação, não tem uma relação direta com o aprendizado considerado no cálculo do Ideb.

Outro indicador que é comumente usado é a taxa de distorção idade-série, que mede a proporção de alunos com mais de dois anos de atraso escolar, sendo obtido a partir do Censo Escolar de um dado ano.

Tais indicadores têm sido empregados para analisar a qualidade do Ensino nas escolas e nas redes escolares, como recentemente fizemos para estabelecer alguns padrões de comportamento das escolas, em termos de proficiência escolar, tanto para o 5º ano quanto para o 9º ano do Ensino Fundamental, da rede escolar do município de Ribeirão Preto (RAMOS; COSTA FILHO, 2020).

Nesse trabalho, procuramos ampliar esse olhar, mediante uma análise estatística multivariada, considerando os indicadores supramencionados, ou seja, % de alunos com aprendizado adequado em língua portuguesa e em matemática, o Ideb desdobrado nas suas duas componentes – a nota padronizada e a taxa de aprovação –, e finalmente, a taxa de distorção idade-série para as escolas municipais do Ensino Fundamental de Ribeirão Preto. Essa análise será realizada por meio da técnica de análise de componentes principais, conforme estabelecem Mardia, Kent e Bibby (1979). Essa técnica já tem sido empregada com sucesso (RAMOS *et al.*, 2005) e (SILVA; RAMOS, 2012) em diversos problemas que exigiram uma análise estatística de dados multivariados, permitindo construir gráficos bidimensionais a partir de matrizes multidimensionais.

Nosso objetivo é obter tais gráficos a partir de 2 matrizes: uma de dimensão 27 (escolas) por 5 (indicadores) para o 5º ano do Ensino Fundamental, e outra de dimensão 20 (escolas) por 5 (indicadores) para o 9º ano do Ensino Fundamental da rede municipal de Educação de Ribeirão Preto. Uma das vantagens desses gráficos de componentes principais é que eles permitem, muitas vezes, perceber padrões que não são de simples inspeção na matriz de dados originais. A partir dessas

análises, pretendemos propor um indicador capaz de medir a heterogeneidade escolar para o 5º e para o 9º ano das escolas municipais de Ribeirão Preto, usando, respectivamente 27 e 20 escolas para os dois anos em questão.

2 Matrizes de indicadores e análise de componentes principais

Nessa análise estatística, as escolas municipais de Ribeirão Preto são apresentadas numa matriz de dados $X_{n,p}$ contendo, nas linhas, um número “n” de escolas e, nas colunas, um número $p = 5$ relativas aos indicadores educacionais. Os dados nessa matriz podem ser vistos como os valores das coordenadas de avaliação de um conjunto de “n” escolas representadas num espaço de 5-dimensões.

A análise de componentes principais descreve uma rotação dos eixos originais na busca de novos eixos (componentes principais) onde os pontos possuam uma máxima variância. Matematicamente, essa rotação é realizada pela diagonalização da matriz de covariância $X'X$, onde X' é a transposta da matriz de dados X . Como resultado, os elementos dos autovetores representam os cossenos diretores, ou seja, a contribuição (ou peso) dos eixos originais (nesse trabalho, os cinco indicadores educacionais) para as componentes principais. Os autovalores, por sua vez, representam a variância em cada componente principal. Assim, por exemplo, o primeiro autovalor da matriz de covariância é igual à quantidade de variância estatística explicada pelo primeiro autovetor. Esse autovetor, que define a primeira componente principal (CP1), aponta na direção da variância máxima estatística dos dados investigados. A segunda componente principal (CP2) é perpendicular à CP1 e explica uma quantidade máxima da variância residual dos dados que não foi explicada pela primeira componente. Se as duas primeiras componentes principais explicarem uma quantidade significativa da variância total, um gráfico com a projeção dos pontos nessas CP pode revelar padrões que não são de simples inspeção na matriz de dados originais.

Para o 5º ano do Ensino Fundamental a matriz de dados de dimensão 27×5 é apresentada na Tabela 1, enquanto na Tabela 2 mostramos a matriz dos indicadores e as escolas para o 9º ano do Ensino Fundamental, gerando, assim, uma matriz de dimensão 20×5 . É importante esclarecer que 7 das 27 escolas municipais estudadas para o 5º ano dos anos iniciais não oferecem anos finais do Ensino Fundamental, com base nas informações constantes no portal QEdu (2020). Portanto, 20 escolas foram empregadas para medir o desempenho escolar do 9º do ano do Ensino Fundamental.

As análises de componentes principais foram realizadas com dados centrados na média (i.e., cada elemento em uma coluna nas Tabelas 1 e 2 foi subtraído pela média

da correspondente coluna) por meio de um código próprio escrito na linguagem R (R CORE TEAM, 2019) e confirmadas com o *software* Ein*Sight 3.0 (1991).

Tabela 1 - Escolas da rede municipal de Ribeirão Preto e indicadores educacionais empregados para análise do 5º ano do Ensino Fundamental

Posição	Escola	LP ^a (%)	MAT ^b (%)	APREND ^c (%)	FLUXO ^d (%)	DIS ^e (%)
1	Profa. Elisa Garcia	86	88	76,0	94	12,0
2	Profa. Eponina Rossetto	92	84	73,9	90	15,0
3	Ver. José Delibo	93	82	75,4	98	6,5
4	Prof. Waldemar Roberto	86	79	72,1	91	12,5
5	Prof. Raul Machado	84	79	73,5	93	12,6
6	Profa. Maria Ignez	80	77	71,8	97	5,5
7	Prof. Anísio Teixeira	89	66	70,6	98	10,1
8	Prof. Alfeu Gasparini	82	70	69,5	92	12,6
9	Antônio Palocci	81	70	70,2	94	6,9
10	Profa. Dercy Seixas Ferrari	83	68	70,2	95	7,7
11	Prof. Monte Serrat	78	71	69,3	92	9,7
12	Dom Luís Amaral	77	69	68,1	89	16,7
13	Prof. Eduardo de Souza	82	64	69,8	93	9,9
14	Prof. Paulo Freire	74	65	68,2	94	11,4
15	Profa. Neuza Marzola	74	59	66,5	94	11,9
16	Sebastião Azevedo	71	60	65,1	93	11,3
17	Prof. Salvador Marturano	75	57	65,9	94	13,3
18	Dr. João Gilberto	68	61	65,0	100	7,0
19	Alcina Heck	79	53	62,5	97	7,2
20	Prof. Jarbas Massullo	70	58	65,9	96	11,0
21	Doutor Faustino	64	60	63,3	91	13,5
22	Geralda Spin	72	44	63,8	96	11,2
23	Prof. Domingos Angerami	60	52	62,0	94	8,6
24	Prof. Jaime Monteiro	60	49	60,9	86	18,8
25	Prof. José Rodini	65	47	62,1	95	13,4
26	Prof. Honorato	41	42	55,1	83	22,6
27	Virgílio Salata	48	46	57,7	85	18,9

^aLP: % dos alunos com aprendizado adequado em língua portuguesa no Saeb de 2017,

^bMAT: % dos alunos com aprendizado adequado em matemática no Saeb de 2017, ^cAPREND: Aprendizagem escolar obtida da nota média padronizada dos exames em língua portuguesa e matemática em 2017, ^dFLUXO: Fluxo escolar (taxa de aprovação) em 2017, ^eDIS: Distorção idade-série para 2018

Fonte: Elaborada pelos autores a partir dos dados dos indicadores LP, MAT, APREND e FLUXO foram extraídos da plataforma QÉdu, enquanto DIS foi obtido do Censo Escolar do Inep/MEC

Tabela 2 - Escolas da rede municipal de Ribeirão Preto e indicadores educacionais empregados para análise do 9º ano do Ensino Fundamental

Posição	Escola	LP ^a (%)	MAT ^b (%)	APREND ^c (%)	FLUXO ^d (%)	DIS ^e (%)
1	Vereador José Delibo	76	48	64,7	99	10,9
2	Profa. Maria Ignez	64	48	62,5	94	11,1
3	Profa. Dercy S. Ferrari	60	41	61,7	95	22,6
4	Profa. Elisa Garcia	65	33	61,9	93	23,7
5	Prof. Raul Machado	63	31	60,0	96	13,8
6	Prof. Anísio Teixeira	58	33	59,5	94	19,0
7	Prof. Monte Serrat	60	29	59,8	84	28,3
8	Dom Luís Amaral	57	28	56,1	81	27,5
9	Prof. Eduardo de Souza	55	29	56,5	97	18,6
10	Virgílio Salata	56	20	56,0	82	32,2
11	Prof. Paulo Freire	49	27	54,6	80	28,6
12	Prof. Jarbas Massullo	53	24	58,0	90	22,0
13	Antônio Palocci	49	26	56,3	88	19,1
14	Prof. Waldemar Roberto	50	24	57,0	89	24,1
15	Prof. Alfeu Gasparini	49	23	55,3	86	26,8
16	Prof. José Rodini	45	26	56,3	89	34,3
17	Doutor João Gilberto	48	16	52,4	95	23,5
18	Prof. Jaime Monteiro	31	18	48,5	85	36,5
19	Profa. Geralda Spin	38	16	53,5	93	18,5
20	Prof. Domingos Angerami	36	15	51,3	90	28,2

^aLP: % dos alunos com aprendizado adequado em língua portuguesa no Saeb de 2017, ^bMAT: % dos alunos com aprendizado adequado em matemática no Saeb de 2017, ^cAPREND: Aprendizagem escolar obtida da nota média padronizada dos exames em língua portuguesa e matemática em 2017, ^dFLUXO: Fluxo escolar (taxa de aprovação) em 2017, ^eDIS: Distorção idade-série para 2018
Fonte: Elaborada pelos autores (2020), a partir dos dados dos indicadores LP, MAT, APREND e FLUXO foram extraídos da plataforma QÉdu, enquanto DIS foi obtido do Censo Escolar do Inep/MEC

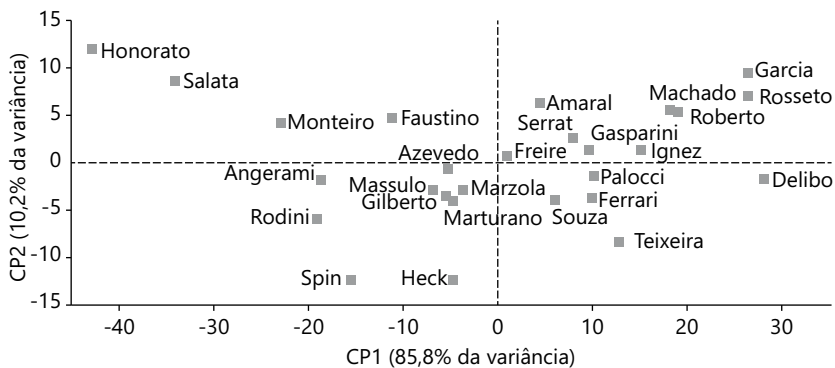
3 Resultados e discussão

A seguir apresentamos os resultados da análise de componentes principais para o 5º e o 9º ano do Ensino Fundamental das escolas municipais de Ribeirão Preto.

3.1 5º ano do Ensino Fundamental

Na Figura 1 apresentamos o gráfico dos escores e as equações das duas primeiras componentes principais (CP1 e CP2) que conjuntamente representam 96,0% da informação original.

Figura 1 - Gráfico dos escores de CP1 x CP2 obtido a partir dos dados centrados na média, para o 5º ano do Ensino Fundamental



$$CP1 = 0,663 LP + 0,673 MAT + 0,287 APREN + 0,102 FLUXO - 0,123 DIS$$

$$CP2 = -0,484 LP + 0,611 MAT + 0,031 APREND - 0,475 FLUXO + 0,407 DIS$$

Fonte: Elaborada pelos autores (2020)

Podemos verificar que a equação de CP1, por si só, representa 85,8% da informação original e é dominada pelas variáveis LP e MAT, que representam os percentuais de alunos com aprendizado adequado em língua portuguesa e em matemática no Saeb de 2017, respectivamente. Claramente, essas duas variáveis possuem maior peso em CP1, em função dos seus coeficientes, ambos positivos e praticamente iguais. Portanto, as escolas que se encontram mais à direita na Figura 1 apresentam melhores proficiências nessas duas disciplinas do que as escolas que se situam mais à esquerda da figura. Por outro lado, CP2, que representa 10,2% da informação, é explicada por quatro variáveis LP, MAT, FLUXO e DIS; onde, MAT tem um peso ligeiramente maior do que as outras três. É interessante verificar que, como esperado, FLUXO e DIS têm sinais opostos entre si, ou seja, quanto maior a taxa de aprovação deve-se esperar menor distorção idade-série.

Nesse ponto, é interessante analisar as três escolas de melhor desempenho escolar, tomando por base os indicadores empregados na Tabela 1. Em consonância com o gráfico dos escores da análise de componentes principais, mostrado na Figura 1, essas 3 escolas do 5º ano do Ensino Fundamental são: Profa. Elisa Garcia, Profa. Eponina Rosseto e Vereador José Delibo. Do ponto de vista das variáveis LP e MAT, todas elas têm, essencialmente, os mesmos resultados, por isso, estão praticamente no mesmo ponto se projetadas no eixo de CP1. Em CP2, as duas primeiras estão com escores positivos, enquanto a Vereador José Delibo situa-se em escore levemente negativo. Isso se deve, principalmente,

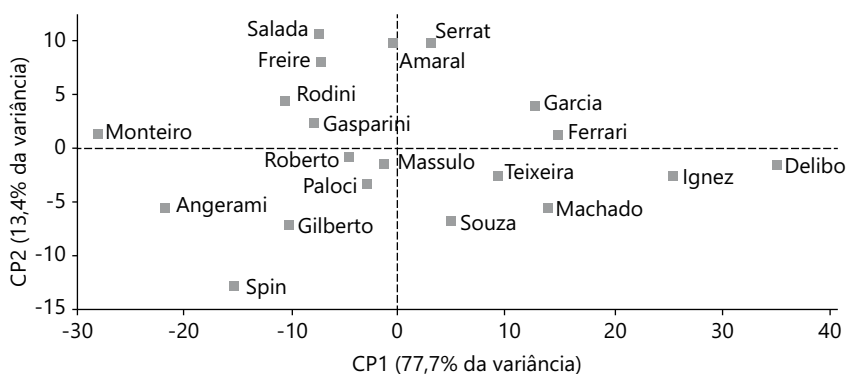
à diferença na distorção idade-série. As duas primeiras têm essa variável com valores bem maiores (12,0% para a Escola Elisa Garcia e 15,0% para a Escola Eponina Rosseto) que aquela de 6,5% para a Escola Vereador José Delibo. As duas primeiras estão no eixo positivo de CP2, porque DIS tem coeficiente positivo nessa equação. Os dados numéricos dessas três escolas estão nas três primeiras posições da Tabela 1.

Outro ponto de interesse é analisar as três escolas com menores níveis de proficiência escolar nessas duas variáveis: Prof. Jaime Monteiro, Prof. Honorato e Virgílio Salata. Elas estão bem à esquerda, ou seja, na parte negativa de CP1, o que significa que devem ter baixa proficiência escolar em LP e MAT, comparativamente às demais escolas. As duas primeiras estão também acima do eixo positivo de CP2, o que deve significar alta distorção idade-série e menor taxa de aprovação relativamente às demais escolas. Isso pode ser confirmado olhando os dados nas posições 24, 26 e 27 da Tabela 1.

3.2 9º ano do Ensino Fundamental

Na Figura 2, apresentamos o gráfico dos escores e as equações das duas primeiras componentes principais (CP1 e CP2) que, conjuntamente, representam 91,1% da informação original.

Figura 2 - Gráfico dos escores de CP1 x CP2 obtido a partir dos dados centrados na média, para o 9º ano do Ensino Fundamental



$$CP1 = 0,663 LP + 0,578 MAT + 0,250 APREN + 0,201 FLUXO - 0,352 DIS$$

$$CP2 = -0,364 LP + 0,167 MAT + 0,078 APREND - 0,646 FLUXO + 0,645 DIS$$

Fonte: Elaborada pelos autores (2020)

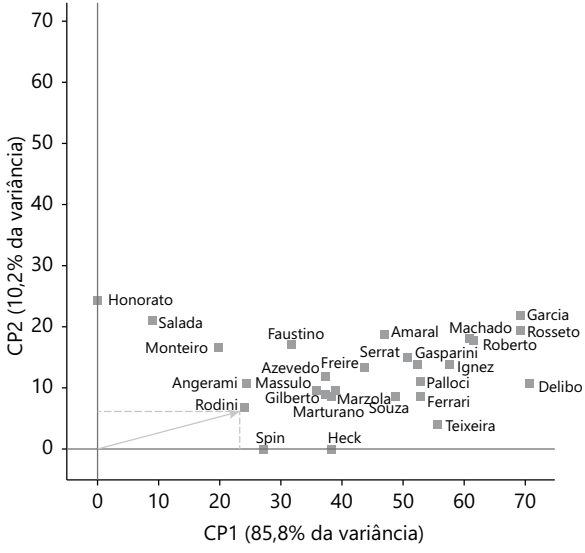
Podemos verificar que a equação da CP1, por si só, representa 77,7% da informação original e é dominada pelas variáveis LP e MAT em conformidade com os coeficientes encontrados, e, como já sabemos, elas representam os percentuais de alunos com aprendizado adequado em língua portuguesa e em matemática, respectivamente, no Saeb de 2017. É interessante verificar que tal comportamento segue aquele que foi também encontrado para o 5º ano do Ensino Fundamental, ou seja, essas duas variáveis têm altos coeficientes. Portanto, as escolas que se encontram mais à direita na Figura 2 apresentam melhores proficiências nessas duas disciplinas do que as escolas que se situam mais à esquerda. Isso pode ser visto olhando os valores desses indicadores nas posições 1 e 2 da Tabela 2 para as escolas Vereador José Delibo e Profa. Maria Ignez (que estão em escores positivos em CP1) e nas posições 18 e 20 da Tabela 2 para as escolas Prof. Jaime Monteiro e Prof. Domingos Angerami (que estão em escores negativos em CP1). Por exemplo, enquanto os percentuais de alunos com aprendizado adequado em língua portuguesa e em matemática são 76% e 48%, respectivamente, na escola Vereador José Delibo, na escola Prof. Jaime Monteiro esses percentuais são 31% e 18%, respectivamente.

Por outro lado, CP2 explica 13,4% da informação original e é, essencialmente, dominada pelo fluxo escolar e pela distorção idade-série e, como esperado, esses indicadores têm sinais opostos. O fluxo tem coeficiente negativo de -0.646, enquanto a distorção tem valor positivo e igual a +0.645. Além disso, são, em módulo, praticamente iguais, o que significa que o peso dessas variáveis para explicar CP2 é essencialmente o mesmo. Assim, quanto mais positivo a escola estiver nos escores em CP2 maior deve ser a distorção idade-série e menor o fluxo escolar, tais como nas escolas Virgílio Salata, Prof. Mont Serrat e Dom Luís Amaral (ver entradas 10, 7 e 8 na Tabela 1). Por outro lado, as escolas situadas na parte negativa de CP2 têm relativamente maiores fluxos escolares e menores taxas de distorção idade-série, como as escolas Profa. Geralda Spin, Prof. Eduardo de Souza e Doutor João Gilberto (ver entradas 19, 9 e 17 na Tabela 1).

4 Análise da heterogeneidade das escolas

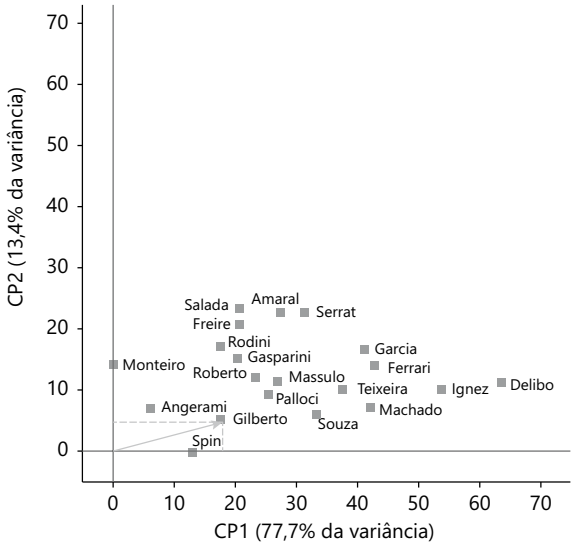
Uma forma de ordenar as escolas, levando em conta os indicadores aqui considerados, é combinar os escores nas duas primeiras componentes principais, de modo que a escola de pior desempenho numa determinada componente passe a ser o zero da escala correspondente, como se vê nas Figuras 3 e 4, para o 5º e para o 9º ano do Ensino Fundamental, respectivamente.

Figura 3 - Redefinição por translação de eixos para ordenamento das escolas relativas ao estudo do 5º ano do Ensino Fundamental



Fonte: Elaborada pelos autores (2020)

Figura 4 - Redefinição por translação de eixos para ordenamento das escolas relativas ao estudo do 9º ano do Ensino Fundamental



Fonte: Elaborada pelos autores (2020)

Na Figura 3, a escola de pior desempenho em CP1 é a Prof. Honorato, enquanto em CP2, Geralda Spin é a pior escola. Já na Figura 4, a escola de pior desempenho em CP1 é a Prof. Jaime Monteiro, enquanto em CP2, a escola Profa. Geralda Spin aparece como a pior. Geometricamente, isso é uma simples translação do sistema de eixos, que não altera o ordenamento, mas tem a vantagem de mostrar que as escolas que estiverem mais afastadas da origem serão as de melhor desempenho em ambas as componentes. Usando os valores assim obtidos como catetos de um triângulo retângulo, a extensão da hipotenusa desse triângulo, calculada pelo Teorema de Pitágoras, pode ser usada para ordenar as escolas em termos dos seus respectivos desempenhos escolares a partir dos indicadores aqui empregados. Esse procedimento está ilustrado (seta vermelha) nas Figuras 3 e 4 para as escolas Prof. José Rodini e Doutor João Gilberto, cujos vetores da origem à localização dessas escolas no gráfico bidimensional leva aos valores de 24,5 e 18,5, respectivamente. Estamos designando a norma (ou seja, o comprimento) desse vetor por “N”. Como se pode ainda notar, os intervalos nos eixos x e y são iguais em cada gráfico, isto é, ambos os eixos variam de -5 a 75 na Figura 3 e de -5 a 70 na Figura 4. Essa equivalência nas escalas de ambos os eixos é realizada para permitir uma visualização correta e imediata da verdadeira distância de cada escola até a origem, bem como da informação relativa explicada por cada componente. Os resultados dos valores de N para todas as escolas, de ambas as etapas escolares, são mostrados na Tabela 3, indo da escola de menor desempenho para aquela de maior desempenho.

Tabela 3 - Ordenamento das escolas municipais de Ribeirão Preto para o 5º e para o 9º ano do Ensino Fundamental

Escolas do 5º ano do EF	Valor de Nª	Escolas do 9º ano do EF	Valor de Nª
Virgílio Salata	22,8	Prof. Domingos Angerami	9,5
Prof. Honorato	24,4	Geralda Spin	12,7
Prof. José Rodini	24,5	Prof. Jaime Monteiro	14,2
Prof. Jaime Monteiro	25,8	Doutor João Gilberto	18,5
Prof. Domingos Angerami	26,4	Prof. José Rodini	24,5
Geralda Spin	27,1	Prof. Alfeu Gasparini	25,2
Dr. Faustino	35,9	Prof. Waldemar Roberto	26,3
Prof. Jarbas Massullo	37,1	Antônio Palocci	27,0
Alcina Heck	38,0	Prof. Jarbas Massullo	28,9
Doutor João Gilberto	38,2	Prof. Paulo Freire	29,2
Prof. Salvador Marturano	38,9	Virgílio Salata	31,1
Sebastião Azevedo	39,1	Prof. Eduardo de Souza	33,5

Continua

Continuação

Escolas do 5º ano do EF	Valor de N^a	Escolas do 9º ano do EF	Valor de N^a
Profa. Neuza Marzola	40,2	Dom Luís Amaral	35,5
Prof. Paulo Freire	45,5	Prof. Monte Serrat	38,6
Prof. Eduardo de Souza	49,5	Prof. Anísio Teixeira	38,7
Dom Luís Amaral	50,6	Prof. Raul Machado	42,4
Prof. Monte Serrat	52,8	Profa. Elisa Garcia	44,0
Profa. Dercy Seixas Ferrari	53,4	Profa. Dercy S. Ferrari	45,0
Antônio Palocci	53,9	Profa. Maria Ignez	54,4
Prof. Alfeu Gasparini	54,1	Vereador José Delibo	64,0
Prof. Anísio Teixeira	55,6		
Profa. Maria Ignez	59,3		
Prof. Raul Machado	63,5		
Prof. Waldemar Roberto	64,2		
Vereador José Delibo	71,5		
Profa. Eponina Rossetto	71,8		
Profa. Elisa Garcia	72,6		

^{na}: magnitude do vetor da origem à localização da escola nos gráficos bidimensionais das Figuras 3 e 4

Fonte: Elaborada pelos autores (2020)

De acordo com os valores de N, para o 5º ano do Ensino Fundamental, as 3 melhores escolas, como esperado, são Profa. Elisa Garcia, Profa. Eponina Rossetto e Vereador José Delibo, enquanto para o 9º ano, as 3 melhores são Vereador José Delibo, Profa. Maria Ignez e Profa. Dercy Seixas Ferrari. Entretanto, para o 9º ano, há uma distância significativa entre elas nos seus respectivos valores de N, ou seja, 64,0, 54,4 e 45,0. Uma diferença de, praticamente, 10 pontos entre cada uma delas, mostrando, assim, uma grande heterogeneidade já entre as 3 melhores escolas do 9º ano, tomando por base os indicadores considerados nesse estudo. Para o 5º ano do Ensino Fundamental, por sua vez, isso não ocorre, como podemos ver na Tabela 3, cujos valores de N para as 3 melhores escolas são 72,6, da Profa. Elisa Garcia; 71,8 da escola Profa. Eponina Rossetto e 71,5, da Vereador José Delibo.

Para medir o grau de heterogeneidade das escolas municipais de Ribeirão Preto, para o 5º e para o 9º ano do Ensino Fundamental, vamos fazer uso do coeficiente de variação (CV), que é uma medida de dispersão normalizada pela média aritmética, representado por:

$$CV = (\sigma/\mu) \times 100 \quad (1)$$

onde, μ representa a média e σ o desvio padrão. Quanto menor for o seu valor, mais homogêneos serão os dados. Os resultados de μ , σ e CV são mostrados na Tabela 4 para o 5º e para o 9º ano do Ensino Fundamental.

Tabela 4 - Parâmetros estatísticos para cada etapa escolar

Parâmetro	5º ano do EF	9º ano do EF
Média (μ)	45,81	32,15
Desvio padrão (σ)	15,50	13,78
Coefficiente de variação (σ/μ) x100	33,83	42,87

Fonte: Elaborada pelos autores (2020)

Os dados de CV na Tabela 4 indicam que as normas para os 2 anos estudados possuem alta dispersão em relação à média (> 33%), portanto, indica um caráter de heterogeneidade educacional relativamente alto, especialmente para o 9º ano do Ensino Fundamental.

5 Conclusão

Com base nos gráficos bidimensionais de componentes principais (CP) e no coeficiente de variação (CV), foi possível estabelecer padrões de desempenho escolar e de heterogeneidade nas escolas municipais de Ribeirão Preto para o 5º e para o 9º ano do Ensino Fundamental. Olhando o conjunto de escolas estudadas, uma estratégia que poderia ser adotada pela Secretaria Municipal de Educação, para o 5º ano do Ensino Fundamental, seria o de promover uma política de colaboração entre escolas (por pares) de melhor desempenho com aquelas de pior desempenho, na perspectiva não só de reduzir o grau de heterogeneidade em termos de proficiência escolar, mas também fortalecer o desejável regime de colaboração na rede escolar. Isso, no nosso entendimento, poderia ser feito ao longo dos 5 anos do Ensino Fundamental, e não apenas no 5º ano. Para isso, a Secretaria de Educação poderia lançar um programa que estimulasse as escolas a trabalharem em pares.

Para o 9º ano do Ensino Fundamental cabe, particularmente, um esforço de melhorar a formação docente em matemática em todas as escolas, cujos resultados puxam para baixo a proficiência escolar. Isso pode ser verificado tanto para as escolas de melhor desempenho quanto para aquelas de pior desempenho. Por exemplo, a melhor escolar para essa etapa é a Vereador José Delibo, cujo percentual de alunos

com aprendizado adequado em língua portuguesa é 76%, enquanto esse percentual em matemática é de apenas 48%, uma diferença de 28 pontos percentuais. Uma das escolas de pior desempenho é a Prof. Domingos Angerami, cujos respectivos percentuais em LP e MAT são 36% e 15%, respectivamente, com uma diferença de 21 pontos percentuais. Prover uma capacitação docente em matemática ao longo dos 4 anos (do 6º ao 9º ano) seria um caminho altamente desejável para melhorar o desempenho da rede escolar dos anos finais de Ribeirão Preto.

Espera-se, portanto, que esse trabalho possa contribuir para aperfeiçoar as políticas públicas de Educação do município de Ribeirão Preto, especialmente quanto à proficiência escolar e ao regime de colaboração. O indicador de heterogeneidade, aqui proposto, pode ser uma bússola a ser usada, a cada dois anos, em conformidade com os exames do Saeb, no contexto dessa política local visando a uma melhor oferta educacional com equidade. Notadamente, outros indicadores poderiam ter sido aqui empregados, mas entendemos que o desafio maior da Educação brasileira, e não só de Ribeirão Preto, concentra-se, como falamos no início desse trabalho, na melhoria da aprendizagem escolar e na redução da desigualdade educacional. Assim sendo, acreditamos que os indicadores aqui empregados são suficientes para dar uma visão desse desafio, e seu enfrentamento passa, necessariamente, pela formação e pela valorização do professor, por uma gestão escolar profissional e por escolas bem estruturadas.

Por fim, cabe ressaltar que a análise estatística multivariada aqui discutida pode ser estendida a qualquer outro município de nosso país, e que as matrizes contendo os fatores em análise podem ser ajustadas para incluir outros aspectos ou indicadores julgados pertinentes pelo gestor.

A multivariate statistical analysis of the performance of the Ribeirão Preto municipal schools

Abstract

This study performs a multivariate statistical analysis of principal components (PCA) to establish school performance standards for the 5th and 9th grades of the municipal schools in Ribeirão Preto. The mathematical equations obtained from two-dimensional PCA graphs reveal that the percentages of students with adequate learning in Portuguese language and mathematics are the most significant indicators to explain such performance patterns. From this analysis, it was possible to establish school performance vectors for each of the schools in the two elementary school years considered here. Finally, it was possible to establish a quantitative indicator of the heterogeneity for each of the school levels investigated.

Keywords: School Performance. Principal Components. School Heterogeneity. Municipal Schools. Ribeirão Preto.

Un análisis estadístico multivariado del desempeño en las escuelas municipales de Ribeirão Preto

Resumen

Este estudio realiza un análisis estadístico multivariado de componentes principales (ACP) para establecer estándares de desempeño escolar para los grados 5° y 9° de la escuela primaria en las escuelas municipales de Ribeirão Preto. Las ecuaciones matemáticas obtenidas de los gráficos bidimensionales de ACP revelan que los porcentajes de estudiantes con un aprendizaje adecuado en portugués y en matemáticas son los indicadores más significativos para explicar dichos patrones de desempeño. A partir de este análisis, fue posible establecer vectores de desempeño escolar para cada una de las escuelas de los dos años de Educación Primaria aquí considerados. Esto permitió la construcción de un indicador de heterogeneidad para las escuelas del sistema municipal de Educación de Ribeirão Preto.

Palabras clave: Desempeño Escolar. Componentes Principales. Heterogeneidad Escolar. Escuelas Municipales. Ribeirão Preto.

Referências

BRASIL. Lei nº 13.005, de 24 de junho de 2014. Aprova o Plano Nacional de Educação – PNE e dá outras providências. *Diário Oficial da União*, Brasília, DF, 26 jun. 2014.

EIN*SIGHT 3.0. Seattle: Infometrix; 1991.

FIGUEIREDO, D. *et al.* Os cavalos também caem: tratado das inconsistências do IDEB. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 26, n. 100, p. 552-572, jul./set. 2018. <https://doi.org/10.1590/s0104-403620180026001178>

MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. *Multivariate analysis*. London: Academic Press, 1979.

OBSERVATÓRIO DO PLANO NACIONAL DE EDUCAÇÃO – OPNE. O Plano Nacional de Educação. Disponível em: <https://www.opne.org.br/>. Acesso em: 15 fev. 2021.

QEdU. Use dados. *Transforme a educação*. 2020. Disponível em: <https://www.qedu.org.br/aprendizado>. Acesso em: 15 fev. 2021.

RAMOS, M. N. *et al.* Aplicação de análise de componentes principais para verificação de atribuições de sinais nos espectros de RMN 1H: o caso dos 3-Aril(1,2,4)-Oxadiazol-5-Carboidrazida. *Química Nova*, São Paulo, v. 28, n. 3, p. 492-496, maio/jun. 2005. <https://doi.org/10.1590/S0100-40422005000300022>.

RAMOS, M. N.; COSTA FILHO, A. Uma análise do aprendizado escolar nas escolas municipais de Ribeirão Preto. *Nova Escola*, São Paulo, maio 2020. Disponível em: <https://novaescola.org.br/conteudo/19149/padrões-de-aprendizagem-para-pensar-sobre-desigualdade>. Acesso em: 17 mar 2021.

R CORE TEAM. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2019. Disponível em: <https://www.R-project.org/>. Acesso em: 15 fev. 2021.

SILVA, A. G.; RAMOS, M. N. A multivariate statistical analysis of the QSAR of 2-(substituted phenyl)indan-1,3-diones with hypolipidemic activity. *Journal of the Brazilian Chemical Society*, São Paulo, v. 23, n. 9, p. 1747-1755, Sept. 2012. <https://doi.org/10.1590/S0103-50532012005000040>


TODOS PELA EDUCAÇÃO. *Em 10 anos, aprendizado adequado no ensino médio segue estagnado apesar dos avanços do 5º ano do fundamental*. 2007. Disponível em: <https://todospelaeducacao.org.br/noticias/meta-3-em-10-anos-aprendizado-adequado-ensino-medio-segue-estagnado-avancos-5-ano-fundamental/>. Acesso em: 15 fev. 2021.

TRAVITZKI, R. Qual é o grau de incerteza do Ideb e por que isso importa? *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 28, n. 107, p. 500-520, abr./jun. 2020. <https://doi.org/10.1590/s0104-40362020002801770>




Informações sobre os autores


Mozart Neves Ramos: Doutor em Química pela Universidade Estadual de Campinas. Professor Emérito da Universidade Federal de Pernambuco.

 <http://orcid.org/0000-0002-6062-3897>


Antonio José da Costa Filho: Doutor em Física pela Universidade de São Paulo. Professor Titular da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo. Pesquisador Nível 1B do Conselho Nacional de Desenvolvimento Científico e Tecnológico. Contato: ajcosta@usp.br

 <http://orcid.org/0000-0001-6730-8737>

João Bosco Paraiso da Silva: Doutor em Química pela Universidade Federal de Pernambuco. Professor Associado IV do Departamento de Química Fundamental e do Instituto de Educação para Inovação da Universidade Federal de Pernambuco. Docente da Universidade Federal de Pernambuco. Contato: paraiso@ufpe.br

 <https://orcid.org/0000-0002-8937-3369>

Ester Fraga Vilas-Bôas Carvalho do Nascimento: Doutora em Educação pela Pontifícia Universidade Católica de São Paulo. Professora PPG II Nível 3 da Universidade Tiradentes. Integra o Programa de Pós-Graduação em Educação. Pesquisadora em Educação Nível 2 do Conselho Nacional de Desenvolvimento Científico e Tecnológico. Contato: esterfraga@gmail.com

 <https://orcid.org/0000-0002-4050-767X>