

Aplicação da análise multivariada no desempenho e hierarquização das escolas estaduais do núcleo regional **de União da Vitória.**

*Application of multivariate analysis on performance
and hierarchical state schools of the Regional Centre
of União da Vitória.*



Aplicação da análise multivariada no desempenho e hierarquização das escolas estaduais do núcleo regional de União da Vitória.
Application of multivariate analysis on performance and hierarchical state schools of the Regional Centre of União da Vitória.

Jair Mendes Marques¹
Paulo Anselmo Santana Owsiany²

Resumo

Este trabalho busca identificar a importância das variáveis referente ao desempenho e a infraestrutura das escolas públicas estaduais do ensino fundamental pertencentes ao Núcleo Regional de Educação de União da Vitória (NRE-UVA). Para o desenvolvimento desta pesquisa, utilizou-se como instrumento principal a coleta de dados por meio de questionários que foram distribuídos nas escolas divididos em três blocos: escolas, professores e alunos, com o objetivo de obter uma descrição quantitativa de informações. Organizaram-se os dados utilizando parte das variáveis correspondentes formando uma matriz para cada bloco, após utilizou-se para análise dos dados a técnica estatística da Análise Fatorial com o método da rotação Varimax, a fim de obter os indicadores de maior poder de explicação em cada um dos blocos. A partir desta análise utilizou-se sobre os indicadores revelados a análise de agrupamento (Cluster) usando método de Ward, do qual identificou a formação de quatro agrupamentos na distância um vírgula cinco no eixo das ordenadas, tendo destaque positivo em desempenho apenas o primeiro grupo formado pelas escolas que apresentaram as melhores condições de trabalho aos professores e o quarto grupo formado pelas escolas com melhores recursos, demandas, condições de infraestrutura e socioeconômica caracterizado pelos diretores e alunos.

Palavras-chave: Escolas; Desempenho; Análise Fatorial; Análise de Agrupamento.

Abstract

This paper seeks to identify the importance of the variables on the performance and the infrastructure of public schools from elementary schools belonging to the Regional Education Center of the União da Vitória (NRE-UVA). To develop this research the main instrument used to collect data were the questionnaires that were distributed to schools and divided into three blocks: schools, teachers and students, in order to obtain a quantitative description of information. The data was organized by using a portion of the corresponding variables in a matrix for each block, and later the statistical technique of Factor Analysis with Varimax rotation method was used for data analysis, to obtain indicators of greater explanatory power in each block. After this analysis we used on the indicators revealed the cluster analysis (Cluster) using Ward's method, which identified the formation of four clusters of schools in the distance of a point five in the ordinate axis, with positive emphasis on performance, with only the first group of the schools that showed the best working conditions for teachers and the fourth group formed by schools with better resources, demands, infrastructure and socioeconomic conditions characterized by the directors and students.

Keywords: Schools, Performance, Factor Analysis, Cluster Analysis.

¹ Doutor em Ciências Geodésicas pela Universidade Federal do Paraná. Professor da FAE Centro Universitário. E-mail: jair.marques@fae.edu

² Mestre em Métodos Numéricos em Engenharia pela Universidade Federal do Paraná. Agente Educacional do Colégio Estadual Zuleide Samways Portes. E-mail: paulo.owsiany@gmail.com

Introdução

Embora a escola seja responsável por mudanças sociais, políticas e culturais é necessário estar sempre buscando alternativas que correspondam às necessidades reais dos alunos e da comunidade que possa melhorar a qualidade do ensino. Na atualidade, a qualidade do ensino é avaliada pelo Ministério da Educação e Cultura (MEC) por meio do Instituto Nacional de Ensino e Pesquisa (INEP) através do Exame Nacional do Ensino Médio (ENEM), Sistema de Avaliação da Educação Básica (SAEB) e a Prova Brasil, quando adquire o papel central e estratégico do monitoramento do sistema educacional. Mesmo com os esforços do MEC para avaliar a qualidade do ensino e a identificação das causas que as instituições de ensino enfrentam os problemas ainda continuam. Isto ocorre pela necessidade de avaliar variáveis relacionadas diretamente aos educandos, professores e as instituições de ensino.

De acordo com Machniki (2010), o material didático, a sala de aula, as condições de estudo, o trabalho do professor, as condições físicas dos estabelecimentos de ensino, o ambiente escolar e o ambiente social influenciam no rendimento escolar.

Segundo Souza (2001), a qualidade dos serviços educacionais depende, em grande parte, da adequação de suas instalações físicas. O ambiente de uma escola também educa e pode tornar a formação de seus alunos não apenas mais eficiente como também muito agradável. Perceber o aspecto negativo do prédio, do pátio, das salas de aula, dos laboratórios e as questões de limpeza deve ser uma preocupação constante de todos os gestores.

Diante desta situação, este estudo objetiva evidenciar a diferença no desempenho das escolas públicas do ensino fundamental, pertencentes ao Núcleo Regional de Educação de União da Vitória e hierarquizá-las mediante um conjunto de indicadores de cada bloco: escola, professores e alunos contendo variáveis relacionadas aos aspectos de infraestrutura, perfil socioeconômico, localização, cultura e outras com características diferenciadas quanto ao desempenho. Assim será possível conhecer as escolas que se destacam positivamente ou negativamente de acordo com os diretores, professores e os alunos. É também propósito deste trabalho contribuir significativamente na melhoria da qualidade do ensino e das próprias escolas em termos de infraestrutura.

É necessário estar sempre buscando alternativas que correspondam às necessidades reais dos alunos e da comunidade que possa melhorar a qualidade do ensino.

1 Análise multivariada

A análise multivariada é assim chamada porque os dados incluem medidas simultâneas sobre muitas variáveis, consiste de um conjunto de técnicas estatísticas que possibilitam analisar e comparar dados referentes a diversas variáveis. Tais técnicas vêm sendo ultimamente bastante difundidas e têm ocupado cada vez mais espaço por conferirem maior precisão aos resultados e agilidade na mensuração da satisfação do mesmo com aplicações em diversas áreas do conhecimento, tais como na área econômica, ciências médicas, biologia, física e no meio acadêmico. Na prática, quando se faz uma observação de um indivíduo o mesmo é caracterizado por várias propriedades, justificando-se assim o emprego das técnicas multivariadas.

Segundo Chaves Neto (1997) a análise multivariada, consiste no estudo de soluções para problemas relacionados com:

- a) inferências sobre médias multivariadas;
- b) análise da estrutura de covariância de uma matriz de dados;
- c) técnicas de classificação e agrupamentos.

1.1 Análise fatorial

A análise fatorial é o nome genérico dado a uma classe de métodos estatísticos cujo propósito principal é definir estrutura subjacente em uma matriz de dados, na qual todas as variáveis são simultaneamente consideradas, cada uma relacionada com todas as outras (HAIR *et al*, 2005).

Segundo Johnson e Wichern (1998) a análise fatorial é uma técnica estatística multivariada que trata do relacionamento entre conjuntos de variáveis, em que:

- as variáveis aleatórias são agrupadas de acordo com suas correlações;
- dentro de um grupo as variáveis aleatórias são altamente correlacionadas entre si;
- de um grupo para outro as correlações são baixas;
- admissível que cada grupo represente um fator, o qual é responsável pelas correlações observadas.

O objetivo da análise fatorial é descobrir fatores latentes ou ocultos que geram a estrutura de correlação de um conjunto de variáveis. Esses fatores supõem-se independentes e linearmente correlacionados às variáveis.

1.1.1 Método das Componentes Principais

O método das componentes principais é um dos mais usados para resolver o problema clássico da análise fatorial. Análise de componentes principais está relacionada com a estrutura de covariância por meio de poucas combinações lineares das variáveis originais em estudo. Os objetivos dessa análise são a redução da dimensão original e a facilitação da interpretação das análises realizadas. Em geral a explicação de toda a variabilidade do sistema determinado por p variáveis só pode ser efetuada por p componentes principais. No entanto, uma grande parte dessa variabilidade pode ser explicada por um número r menor de componentes,

$$r \leq p$$

1.1.2 Teste de Esfericidade de Bartlett

De acordo com Marques (2003) o teste de Bartlett testa a hipótese de que as variáveis não sejam correlacionadas na população. A hipótese básica diz que a matriz de correlação da população é uma matriz identidade o que indica que o modelo fatorial é inapropriado. A estatística do teste é dada por:

$$X^2 = - \left[(n-1) - \frac{2p+5}{6} \right] \ln |R|$$

que tem uma distribuição qui-quadrado com:

$$v = \frac{p(p-1)}{2}$$

grau de liberdade, onde:

n = tamanho da amostra;

p = número de variáveis;

$|R|$ = determinante da matriz de correlação.

1.1.3 Medida de Adequação da Amostra (MSA) de Kaiser-Meyer-Olkin (KMO)

Essa medida é representada por um índice (MSA) que avalia a adequação da análise fatorial, sendo calculada por:

$$MSA = \frac{\sum_{j \neq k} \sum r_{jk}^2}{\sum_{j \neq k} \sum r_{jk}^2 + \sum_{j \neq k} \sum q_{jk}^2}$$

Onde: r_{jk}^2 é o quadrado dos elementos da matriz de correlação original (exceto a diagonal); q_{jk}^2 é o quadrado dos elementos fora da diagonal da matriz anti-imagem da correlação.

Valores altos (entre 0,5 e 1) indicam que a análise fatorial é apropriada, enquanto que os valores baixos, abaixo de 0,5 indicam que a análise fatorial pode ser inadequada.

A análise fatorial é o nome genérico dado a uma classe de métodos estatísticos cujo propósito principal é definir estrutura subjacente em uma matriz de dados, na qual todas as variáveis são simultaneamente consideradas,

1.1.4 Escolha do Número de Fatores

A escolha do número de fatores é uma das tarefas mais importantes da Análise Fatorial. Hair Junior *et al.* (2005) discutem que, se o pesquisador opta por um número muito reduzido, não é possível identificar estruturas importantes existentes nos dados; por outro lado, se o número é excessivo, ele pode vir a ter problemas de interpretabilidade dos fatores. Existem na literatura vários critérios que auxiliam na determinação do número de fatores, como por exemplo:

- número de fatores igual ao número de autovalores maiores que 1 (critério de Kaiser);
- número de fatores igual ao número de autovalores maiores que um número V ;
- número de fatores que explicam pelo menos determinado percentual da variância total;
- número de fatores igual a certo valor N .

1.1.5 Rotação de Fatores

A rotação de fatores é uma técnica para girar os eixos de referência dos fatores em torno da origem, de forma a orientá-los a facilitar a interpretação do conjunto de variáveis em cada fator. A rotação pode ser ortogonal ou oblíqua, caso os eixos se mantiverem ou não em 90 graus entre si durante o giro. O objetivo é facilitar a leitura dos fatores, pois a rotação deixa pesos fatoriais altos em um fator e baixos em outros, definindo mais claramente os grupos de variáveis que fazem parte de um fator estudado. A rotação oblíqua é mais realista, porém mais controversa. O método de rotação mais utilizado é o Varimax o qual simplifica as colunas da matriz de fatores.

1.2 Análise de Cluster

O objetivo da análise é agrupar objetos semelhantes por meio da mínima variância e separar os grupos pela maximização da variância entre eles, onde o agrupamento é feito de tal forma que haja homogeneidade dentro e heterogeneidade entre os agrupamentos. Encontrar nos dados uma estrutura natural de agrupamento é uma importante técnica exploratória. A análise de

Encontrar nos dados uma estrutura natural de agrupamento é uma importante técnica exploratória.

agrupamento por sua vez não considera o número de grupos e é realizada com base na similaridade ou dissimilaridade, ou seja, distâncias (JOHNSON; WICHERN, 1998).

1.2.1 Medidas de similaridade e dissimilaridade

Para agrupar os itens é necessário especificar um coeficiente de similaridade em que quanto maior o valor observado, mais parecidos serão os objetos (como o coeficiente de correlação), ou um coeficiente de dissimilaridade onde quanto maior o valor observado menos parecidos, ou seja, mais dissimilares serão os objetos (como a distância euclidiana).

Neste estudo foi utilizada a distância euclidiana por ser uma das mais conhecidas e adotadas. Para definir-se a distância entre dois objetos no espaço p -dimensional, sejam...

$$\underline{X}'_1 = [X_{11} \ X_{12} \ \dots \ X_{1p}] \quad \text{e} \quad \underline{X}'_2 = [X_{21} \ X_{22} \ \dots \ X_{2p}]$$

observações entre dois objetos (indivíduos), segundo Marques (2003), pode ser calculada da seguinte forma:

$$d(\underline{x}, \underline{y}) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}$$

1.2.2 Método de agrupamento Hierárquico

O método hierárquico pode ser de dois tipos: aglomerativo e divisível, no qual se estabelece uma relação de hierarquia entre o objeto e o conjunto de objetos. Uma vez incorporado a um grupo, o objeto permanece associado a ele até o final do processo de agrupamento. Cabe destacar

ainda, que o método hierárquico apresenta como desvantagens, o fato de que só fundem ou dividem dois clusters de cada vez, e um cluster já formado não se divide. Nesse método, os critérios de agrupamentos mais utilizados são o da associação simples, que é baseado nas menores distâncias entre os objetos, e o da associação completa, baseado nas maiores distâncias entre os objetos e para este estudo destaca-se o método de Ward.

1.2.2.1 Método de Ward

Considerado um procedimento de agrupamento hierárquico baseado em minimizar a “perda de informação” na junção de dois grupos. Este método é geralmente realizado com a perda de informações assumindo ser um aumento no critério do Somatório do Quadrado dos Erros (ESS), que é dado por:

$$ESS = \sum_{j=1}^N (x_j - \bar{x})'(x_j - \bar{x})$$

onde x_j é a medida multivariada associada ao j -ésimo item e \bar{x} é a média de todos os itens.

Os resultados do método de Ward podem ser exibidos como um dendrograma. O eixo vertical apresenta os valores do ESS em que ocorre a fusão. O método de Ward é baseado no conjunto de observações multivariadas que devem ser aproximadamente de forma elíptica (JOHNSON; WICHERN, 1998).

2 Material e métodos

Para testar a viabilidade e a eficiência dos modelos, foram utilizados conjuntos de dados quantitativos extraídos da pesquisa feita nas escolas de ensino fundamental pertencentes ao Núcleo Regional de Educação de União da Vitória realizada através de questionários aplicados, em três segmentos: o primeiro para a escola (diretor responsável por responder), o segundo aos professores e o terceiro aos alunos. O objetivo da coleta de dados é levantar características descritivas sobre as escolas, tais como: educação, estrutura, ambiente, recursos, cultura, condições socioeconômicas da comunidade escolar, localização, entre outras destacando o aproveitamento na melhoria do ensino de acordo com a caracterização das amostras.

Antes da aplicação dos questionários foi obtido junto ao Núcleo de Educação informações das escolas referentes ao início do ano de 2009, onde foi possível verificar o número de alunos e professores de cada instituição podendo avaliar a quantidade de questionários que deveriam ser usados na coleta de dados. Os dados fornecidos pelo Núcleo Regional de Educação relataram a existência de 1.531 professores, 25.020 alunos e 850 turmas do total geral reunindo 42 instituições de ensino distribuídas entre 9 municípios. Para a caracterização das amostras calculou-se o número equivalente a 20% dos alunos e 40% dos professores de cada instituição para que, posteriormente fossem aplicados os questionários.

Com o intuito de mensurar estes aspectos, devido ao número pequeno de escolas entre muitas variáveis, foram selecionadas apenas um número reduzido de variáveis através de consultas as equipes pedagógicas, diretores, professores e alunos para cada um dos blocos:

- Variáveis selecionadas do questionário escola:
 - E1 - Adaptação para deficientes;
 - E2 - Quadra de esportes;
 - E3 - Condição biblioteca;
 - E4 - Disponibilidade computadores aos alunos;
 - E5 - Demanda de pedagogos;
 - E6 - Premiação na Olimpíada de Matemática;
 - E7 - Recursos do Governo;
 - E8 - Grau de risco da escola devido à localização;
 - E9 - Nível clientela escolar;
 - E10 - Serviços gerais;
 - E11 - Condição dos banheiros;
 - E12 - Condição das salas de aula.

O critério usado na identificação de determinadas variáveis foi o uso de respostas binárias (0 ou 1) por se tratar de uma opção para cada resposta (apenas o diretor), considerando o valor 1 (um) para resposta mais adequada, melhor condição ou satisfatório e o valor 0 (zero) para resposta considerada inadequada, pior condição ou insatisfatório.

- Variáveis selecionadas do questionário professor:
 - P1 - Qualidade do material didático do aluno;
 - P2 - Condição biblioteca;

- P3 - Desinteresse do professor devido ao salário;
 - P4 - Desinteresse dos alunos pela condição social, política e econômica;
 - P5 - Nível de aprendizagem dos alunos;
 - P6 - Condições de trabalho do professor;
 - P7 - Grau de risco da escola devido à localização;
 - P8 - Limpeza geral;
 - P9 - Condições dos banheiros;
 - P10 - Condições das salas de aulas;
 - P11 - Condição quadro-negro;
 - P12 - Inexistência de agressão física entre alunos.
- Variáveis selecionadas do questionário aluno:
- A1 - Escolaridade do pai médio ou superior;
 - A2 - Renda familiar até um salário mínimo;
 - A3 - Condições da biblioteca;
 - A4 - Satisfação com a direção da escola;
 - A5 - Serviços gerais;
 - A6 - Condições dos banheiros;
 - A7 - Condições das salas de aula;
 - A8 - Famílias beneficiadas pelo programa Bolsa Família;
 - A9 - Frequência dos responsáveis na escola;
 - A10 - Estudam no mínimo trinta minutos por dia;
 - A11 - Condição quadra de esportes;
 - A12 - Condição do pátio escolar.

O critério usado para os professores e alunos na identificação das variáveis foi o uso de respostas percentuais, ou seja, quanto maior for o percentual (mais próximo de 1) melhor é a condição ou satisfação dos mesmos para determinada variável.

O procedimento estatístico foi aplicar sobre as matrizes o teste da Medida de Adequacidade da Amostra (MSA) e o teste de esfericidade de Bartlett. Após a aplicação de ambos os testes foi realizada em cada matriz a análise fatorial, sendo a extração dos fatores feita pelo método das componentes principais através da rotação Varimax, tendo como critério o número de fatores explicando no mínimo 75% da variância total a fim de obter uma melhor interpretabilidade dos fatores em relação à correlação entre as variáveis, usando o programa computacional Matlab 6.0 (MATSUMOTO, 2002).

Baseado no resultado da análise fatorial sobre a matriz dos questionários das escolas realizou-se uma interpretação sobre os escores fatoriais finais ponderados (obtidos através a média ponderada entre os escores fatoriais e a

proporção de explicação dos respectivos fatores) das 42 instituições convertidos na escala entre 0 e 1 para identificar o nível de rendimento percentual das escolas. Em seguida realizou-se a interpretação dos escores fatoriais finais nas análises dos professores e também dos alunos e posteriormente procedeu-se a análise de agrupamento (Cluster), sobre os escores fatoriais finais resultantes das análises fatoriais dos três blocos usando o *software* Statistica 7.0.

3 Resultados e discussão

Resultados da análise das escolas:

Teste de Esfericidade de Bartlett e Medida de Adequacidade da Amostra de Kaiser_Meyer Olkin (KMO):

(1) Teste de Esfericidade de Bartlett;

Q2 = 98.242, com p valor = 0 ao nível de 5% indicando que existe uma correlação significativa entre as variáveis.

(2) Medida de Adequacidade da Amostra de Kaiser_Meyer Olkin (KMO).

MSA = 0.5145, como MSA > 0.5 indica que a análise fatorial é adequada.

Na sequência realizou-se a análise fatorial pelo método das componentes principais, da qual revelou sete fatores cujo percentual de variância explicada pelo primeiro fator é 20,5%, o segundo 16,17%, o terceiro 12,25%, o quarto 10,67%, o quinto 8,58%, o sexto 6,83% e o sétimo fator 6,25% de explicação. Dessa forma os sete fatores explicam 81,08% da variabilidade total.

QUADRO 1 variáveis que compõem cada fator, de acordo com os pesos para cada um dos fatores - questionário escola

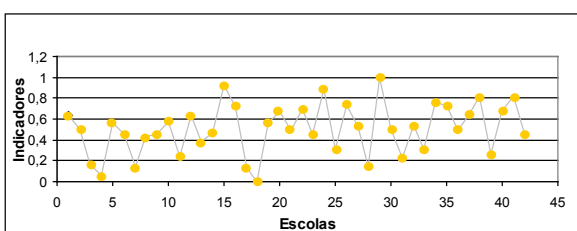
Fator	Variáveis de cada fator
F1	E5: Demanda de pedagogos E7: Recursos do governo
F2	E12: Condição das salas de aula E3: Condição biblioteca E11: Condição dos banheiros
F3	E1: Adaptação para deficientes E10: Serviços gerais
F4	E2: Quadra de esportes E4: Disponibilidade computadores na biblioteca aos alunos
F5	E9: Nível clientela escolar
F6	E6: Premiação na Olimpíada de Matemática
F7	E8: Grau de risco da escola devido à localização

Fonte: O autor

No gráfico 1 verifica-se que 14,29% das escolas não atingiram o indicador 0,2, isso significa que 6 instituições desempenham menos de 20% dos requisitos relacionados ao nível de infraestrutura e investimento público pelo fato dos fatores F1 e F2 juntos apresentarem um poder explicativo de quase 37%, ou seja, são os fatores que mais influenciam no nível de indicação. Deve-se levar em conta, também, os demais fatores que correspondem às necessidades de adaptações do espaço físico, recursos para contratação de funcionários e investimentos tecnológicos e na segurança em torno da escola. Os resultados apresentados para esse grupo de escolas estão condizentes com a realidade devido à falta de investimento no ambiente escolar, pela necessidade de aumento na demanda funcional, alocação de recursos para reformas e/ou construção de novas escolas e nas condições socioculturais da comunidade. No intervalo de 0,2 até 0,4 apresenta um percentual de 14,29% das escolas com desempenhos que variam de 20% a 40% de suas necessidades. Os resultados, embora um pouco melhor do que o grupo anterior, apresentam as mesmas deficiências significativas nas novas variáveis (fatores). Nos indicadores maiores que 0,4 e menores ou iguais a 0,6, são representadas por 35,71% que atingiram um desempenho de até 60% de aproveitamento, na outra escala cujos valores são maiores que 0,6 estão localizadas 15 escolas que representam cerca de 35,71% do total e são consideradas a partir desta escala as que apontam ter as melhores condições de infraestrutura, geralmente são as escolas que estão inseridas mais próximas aos centros urbanos ou estão localizadas nos municípios com melhores condições econômicas da região, podendo estar caracterizada numa comunidade com melhores condições sociais e culturais possibilitando maior favorecimento dos líderes políticos. De acordo com a localização dessas 15 escolas, 10 estão localizadas nos dois principais municípios: São Mateus do Sul e União da Vitória (sede do Núcleo Regional), ambos com mais de 40 mil habitantes.

No geral, 64,28% das escolas não conseguiram obter mais de 60% dos itens necessários para uma melhor colocação em relação às variáveis da análise.

GRÁFICO 1 - indicador de desempenho por escola



FONTE: O autor

Para os resultados da análise da amostra dos professores tem-se:

Teste de Esfericidade de Bartlett e Medida de Adequacidade da Amostra de Kaiser_Meyer Olkin (KMO):

(1) Teste de Esfericidade de Bartlett;

$Q^2 = 128,4091$, com p valor = 0 ao nível de 5% indicando que existe uma correlação significativa entre as variáveis.

(2) Medida de Adequacidade da Amostra de Kaiser_Meyer Olkin (KMO).

$MSA = 0,5217$, como $MSA > 0,5$ indica que a análise fatorial é adequada para a amostra utilizada.

Resultaram na análise fatorial seis fatores, usando como critério que a explicação dos fatores seja no mínimo de 75%. Sendo assim, o primeiro fator explica 24,43% da variabilidade total e os demais fatores apresentam uma proporção de 16,5%, 12,5%, 11,5%, 7,67% e 6,42% respectivamente e os seis fatores juntos explicam 79% de variabilidade, este valor é a soma da proporção de todos os fatores.

QUADRO 2 variáveis que compõem cada fator, de acordo com os pesos para cada um dos fatores - questionário professor.

FATOR	VARIÁVEIS DE CADA FATOR
F1	P2: Condição biblioteca P6: Condições de trabalho do professor P11: Condição quadro-negro
F2	P1: Qualidade do material didático do aluno P5: Nível aprendizagem dos alunos
F3	P4: Desinteresse dos alunos pela condição social, política e econômica P3: Desinteresse Professor devido ao salário P7: Grau de risco da escola devido à localização
F4	P8: Limpeza geral
F5	P12: Inexistência de agressão física entre alunos
F6	P9: Condição dos banheiros P10: Condições das salas de aula

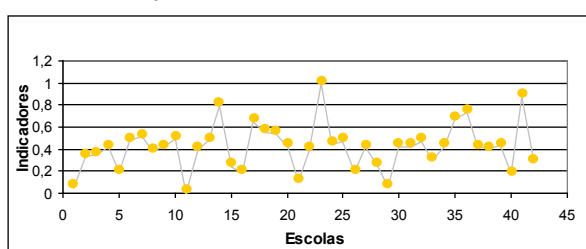
FONTE: O autor

O gráfico 2, mostra que 8 escolas tiveram indicadores menores ou iguais a 0,2 e isso revela que 19,40% das escolas resumem seu rendimento e condições de infraestrutura abaixo de 20%. O baixo rendimento pode estar caracterizado pelas condições de trabalho do professor: pode ser pela dificuldade de acesso à escola, violência em torno e principalmente por falta de estrutura no trabalho apresentando um nível explicativo de 24,43% no fator F1, sendo o fator mais influencia-

do no rendimento. Os demais fatores contribuem 54,57% de explicação da variabilidade total, juntos podem influenciar no baixo desempenho pela falta de investimentos na qualidade da educação e nos demais setores da sociedade. Outras 9 escolas com uma representação de 21,43% têm um indicador de aproveitamento acima de 0,2 e abaixo de 0,4, representando um desempenho de 20% a 40% de suas características necessárias, de acordo com a análise matemática. No intervalo entre 0,4 e 0,6 estão representadas boa parte das escolas totalizando 19, estimando 45,24% que apresentam aproveitamento intermediário entre 40% e 60%. Estas escolas apresentam situações regulares de trabalho para o professor assim como na infraestrutura do prédio escolar e nos materiais didáticos para uso dos alunos. Encontram-se entre os indicadores 0,6 e 1 um total de 6 escolas representando 14,3% das instituições, das quais, são as que obtiveram os melhores resultados favoráveis aos professores em termos de condições de trabalho, satisfação com o salário, organização do ambiente escolar e na segurança em torno da escola. Estes resultados estão baseados na maior representatividade de explicação relacionados aos fatores F1, F2, F3 e F4 (24,43%, 16,5%, 12,5% e 11,5%).

No geral, 85,7% das escolas não conseguiram obter mais de 60% dos itens necessários para uma melhor colocação em relação às variáveis da análise. Este resultado, na realidade pode estar relacionado à insatisfação do professor devido ao salário e nas condições de trabalho que as escolas oferecem diante de uma sociedade cada vez mais dependente do sistema educacional.

GRÁFICO 2 - indicador de desempenho dos professores por escola



FONTE: O autor

Para os resultados da análise da amostra dos alunos têm-se:

Teste de Esfericidade de Bartlett e Medida de Adequacidade da Amostra de Kaiser_Meyer Olkin (KMO):

(1) Teste de Esfericidade de Bartlett;

$Q2 = 211,6622$ com p valor = 0 (para $gl = v = 66$), portanto, o teste é altamente significativo, a matriz de correlação é adequada para a análise fatorial.

(2) Medida de Adequacidade da Amostra de Kaiser - Meyer - Olkin (KMO).

$MSA = 0,6715$, verifica-se que a medida está acima do valor mínimo aceitável 0,5.

Os resultados obtidos do Teste de Esfericidade de Bartlett e Medida de Adequacidade da Amostra indicam que a análise fatorial é adequada.

Resultou na análise 5 fatores, usando como critério o número de fatores que juntos expliquem pelo menos 75% da variabilidade. Sendo assim o primeiro fator explica um percentual de 33,33% da variabilidade total e os demais apresentam uma explicação de 18,92%, 10,08%, 7,92% e 7,67% respectivamente e os cinco fatores juntos explicam 77,91% de variabilidade, este valor é a soma da proporção de todos os fatores.

QUADRO 3 - variáveis que compõem cada fator, de acordo com os pesos para cada um dos fatores - questionário aluno

FATOR	VARIÁVEIS QUE COMPÕE CADA FATOR
F1	A3: Condições da biblioteca A4: Satisfação com a direção da escola A5: Serviços gerais A6: Condição dos banheiros A7: Condições das salas de aula
F2	A2: Renda familiar até um salário mínimo A8: Famílias beneficiadas pelo Bolsa Família
F3	A10: Estudam no mínimo 30 min por dia
F4	A1: Escolaridade do pai médio ou superior A9: Frequência dos responsáveis na escola
F5	A11: Condição quadra de esportes A12: Condição do pátio da escola

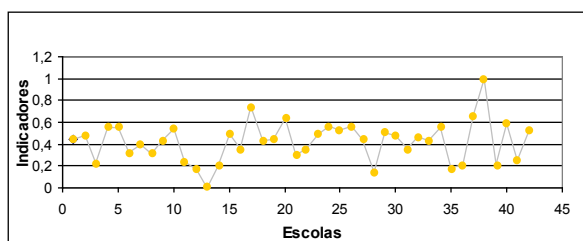
FONTE: O autor

De acordo com o gráfico 3, mostrou-se na análise que 5 escolas possuem como indicadores valores menores ou iguais a 0,2, isso evidencia que dos itens que fazem parte dos questionários dos alunos, 11,90% das escolas resumem suas condições de infraestrutura e as condições sociais e econômicas das famílias dos alunos um rendimento inferior a 20%. Estes resultados baseiam-se principalmente no fator F1, do qual tem o maior percentual de explicação (33,33%) e no fator F2 (18,92%) que corresponde à situação econômica das famílias dos alunos. De antemão, pode-se afirmar que o investimento nesses fatores pelos governantes traria melhores resultados, com exceção da variável A8 (famílias beneficiadas pelo Bolsa Família), devido ao aumento na renda familiar.

Entre os indicadores de 0,2 a 0,4 do intervalo estão localizadas 11 escolas representando 26,19% do total, cujas características de desempenho não ultrapassam dos 40% conforme os indicadores do

gráfico 3. Ao verificar os indicadores maiores que 0,4 e menores ou iguais a 0,6, encontra-se dentro desta faixa de intervalo o maior número de escolas pesquisadas, atribuindo 52,38% representando mais da metade do total das instituições com desempenhos entre 40% e 60%. Neste intervalo estão as instituições com nível de infraestrutura regular e que podem estar situadas em uma comunidade menos precária nas condições social e econômica. Entre os indicadores 0,6 e 1 encontram-se situadas nesta faixa apenas 4 escolas, representando 9% do total das instituições pesquisadas. Desta forma, segundo os dados da análise fatorial, resume-se que as estruturas físicas do ambiente dessas escolas apresentam melhores condições em comparação com as demais pela influência da proporção acumulada dos fatores F1 e F5 (41%). Os demais fatores podem contribuir significativamente nesses resultados. Conforme um estudo realizado por Barbosa (2009) o fator renda familiar e o fator tempo de estudo são uns dos que mais influenciam o resultado do desempenho das escolas e ainda de um modo mais positivo se tiver uma clientela mais participativa a seu favor, tornando evidente que as avaliações das escolas devem ser baseadas num leque bastante amplo de variáveis e que juntas trarão ganhos eficientes para a educação.

GRÁFICO 3 - Indicador de desempenho dos alunos por escola



FONTE: O autor

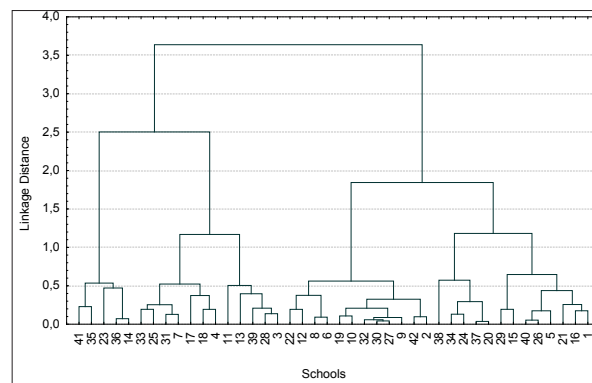
3.1 Análise de agrupamentos

Para o uso da análise de agrupamentos, primeiramente foi usado a Análise Fatorial para obter os indicadores (escores) que foram convertidos para o intervalo entre 0 e 1 a fim de facilitar a interpretação dos resultados das escolas, professores e alunos. Após, calcularam-se então as distâncias euclidianas entre os escores resultando em um total de 861 distâncias.

O resultado dessas ligações pode ser disposto em um diagrama em árvore ou dendrograma como é mais conhecido, do qual possui uma escala para observar os níveis de distância. O critério de

similaridade é representado pela altura resultante de cada combinação distinta entre as variáveis, ou seja, quanto mais baixa for essa junção, maior será o grau de similaridade entre as situações estudadas (São Paulo, 1992). Desta forma, com intuito de não perder estruturas importantes e não correr o risco de ter problemas de interpretabilidade, optou-se para este estudo apenas o intervalo de estabilidade compreendido por volta do valor 1,5 do eixo das distâncias da ligação com uma representação de 4 agrupamentos, conforme mostra o gráfico 4.

GRÁFICO 4 - Dendrograma referente aos agrupamentos das escolas - Método de Ward



FONTE: O autor

Embora exista certa dificuldade em estabelecer uma medida que caracterize a magnitude das condições das escolas no geral e das condições socioeconômicas da região, é possível, tomando-se por base os indicadores e os clusters obtidos, fazer alguma diferenciação entre os grupos formados, quanto à intensidade do desenvolvimento.

Nesse sentido, pode-se dizer que o primeiro grupo composto pelas escolas 41, 35, 23, 36 e 14 apresenta as melhores condições de trabalho ao professor e bons resultados em investimento na demanda funcional, nas adaptações físicas às necessidades especiais e nos recursos materiais para uso dos alunos. Por outro lado é o grupo que apresenta a pior condição de infraestrutura física; revela também um índice menos significativo na infraestrutura familiar dos alunos relacionado à renda e a necessidade de ajuda de algum programa social. Resume-se que embora este grupo apresente boas condições de trabalho e uma melhor alocação de recursos na área pedagógica, sugere-se investir mais em um ambiente que satisfaça a necessidade do aluno como, por exemplo, a melhoria nas salas de aula, banheiros e bibliotecas. O segundo agrupamento está representado pelas instituições 33, 25, 31, 7, 17, 18, 4, 11, 13, 39, 28 e 3 que possui indicadores baixos em pelo menos uma das três análises, considerando-se o grupo que apresenta uma significativa variação entre as esco-

las, correspondente aos recursos de investimento, tanto na escola quanto na sociedade. A falta desses recursos, no conjunto das variáveis utilizadas na pesquisa, reflete baixos níveis de desenvolvimento no aspecto social e educacional. O terceiro agrupamento é formado pelas escolas 22, 12, 8, 6, 19, 10, 32, 30, 27, 9, 42 e 2 que revelou indicadores médios nas análises dos três questionários não tendo variação expressiva entre os indicadores. Pode-se considerar que este grupo apresenta condições regulares em infraestrutura e nos demais fatores que envolvem as questões socioeconômicas. No quarto grupo estão as escolas 34, 24, 37, 20, 29, 15, 40, 26, 5, 21, 16, 1 e 38 que apresentaram as melhores condições de infraestrutura no geral, juntamente revelou uma comunidade escolar mais participativa no desenvolvimento educacional de seus educandos, além de possuir mais recursos para as necessidades dos alunos, porém este grupo caracterizou aos professores as piores condições de trabalho, podendo estar relacionado ao salário e principalmente na segurança. Embora essas escolas supram as necessidades físicas e pedagógicas é recomendável rever e analisar as necessidades do educador com a intenção de melhorar a educação.

4 Conclusão

Este trabalho procurou através de um número significativo de variáveis de cada bloco de questionário, analisar e avaliar o desempenho e infraestrutura das escolas levando em consideração várias situações que são comuns no dia a dia destas instituições.

Ao analisar as características das 42 escolas públicas do ensino fundamental do NRE de União da Vitória, observou-se, através dos diagnósticos da realidade das escolas que, apesar da maioria apresentar estruturas físicas necessárias, faltam recursos financeiros para o investimento de melhorias e com isso poder alcançar bons resultados.

Um dos principais desafios dos gestores responsáveis por estas instituições e dos professores consiste em administrar e a trabalhar os fatores que não dependem somente das ações das próprias escolas, como o baixo nível cultural da comunidade escolar, situação socioeconômica das famílias, a indiferença e a falta de interesse dos pais na educação dos filhos entre outros.

Utilizado a ferramenta estatística de análise fatorial e análise de agrupamentos sobre as três matrizes de dados foi possível, através dos escores

fatoriais finais, analisar e classificar as escolas do Núcleo Regional de União da Vitória em função dos segmentos selecionados e aplicados a esta pesquisa. O método proposto demonstrou que, a eficiência no sistema de educação varia sistematicamente, quanto às características social, econômica e cultural, das localizações que esses estabelecimentos se encontram nos centros urbanos, periferias e zonas rurais, quanto a gerências educacionais. Em tais casos, precisa-se considerar e/ou colocar mais condições que visem no suprimento destas defasagens, compreender a situação socioeconômica para tentar sanar as restrições desfavoráveis e, assim, implicar positivamente nos estabelecimentos.

Com os resultados obtidos conclui-se que 11,9% das instituições atingiram resultados favoráveis aos professores em termos de condições de trabalho, satisfação com o salário, nível de aprendizagem, organização do ambiente escolar e na segurança em torno da escola, mas revelou serem as escolas com menores indicadores nas condições de infraestrutura física e socioeconômica da clientela escolar; 28,57% apresentaram resultados insatisfatórios no tocante aos fatores revelados; 28,57% atingiram resultados menos insatisfatórios em infraestrutura, condições de trabalho e questões socioeconômicas; 30,95% apresentaram melhores indicadores de desempenhos em infraestrutura, recursos, demandas e nas condições sociais e econômicas na caracterização dos diretores e alunos, porém, obtiveram os menores indicadores nas condições de trabalho

Administrar e trabalhar os fatores que não dependem somente das ações das próprias escolas, como o baixo nível cultural da comunidade escolar, situação socioeconômica das famílias, a indiferença e a falta de interesse dos pais na educação dos filhos entre outros.

Os escores e agrupamentos fornecidos pela análise fatorial e análise de Cluster respectivamente, podem ser um dos instrumentos valiosos para os gestores das escolas e líderes políticos.

Os escores e agrupamentos fornecidos pela análise fatorial e análise de Cluster respectivamente, podem ser um dos instrumentos valiosos para os gestores das escolas e líderes políticos a discutirem e direcionarem seus projetos de forma mais eficaz na eficiência de resultados.

dos professores. Destas análises ficou evidenciado um aproveitamento desequilibrado entre as escolas, assim sugere-se que políticas compensatórias não devam concentrar-se, basicamente, em somente um dos grupos revelados. Além disso, algumas escolas podem se espelhar nos trabalhos pedagógicos e administrativos de outras para melhorar seu desempenho.

- Recebido em: 01/12/2010
- Aprovado em: 20/05/2011

Referências

- BARBOSA, S. G. Influência dos fatores sociais econômicos no desempenho de escolas públicas. **Revista Diálogos & Saberes**, Mandaguari, v. 5, n. 1, p. 93-109, 2009.
- CHAVES NETO, A. **Análise multivariada aplicada à pesquisa**. Curitiba: Universidade Federal do Paraná, 1997. Notas de aula.
- HAIR JR., J. F. et al. **Análise multivariada de dados**. Porto Alegre: Bookman, 2005.
- JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 4th. ed. Upper Saddle River, N.J.: Prentice Hall, 1998.
- MACHNICKI, A. C. **Estudo e aplicação da análise multivariada na validação das estruturas físicas e pedagógicas de escolas públicas do Estado do Paraná na Região Sul**. 2010. 143 f. Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Setores de Tecnologia e de Ciências Exatas, UFPR, Curitiba, 2010.
- MARQUES, J. M. **Análise multivariada**. Curitiba, 2003. Notas de Aula
- MATSUMOTO, E. Y. **Matlab 6.5: fundamentos de programação**. 2. ed. São Paulo: Érica, 2002.
- SÃO PAULO. Fundação Sistema Estadual de Análise de Dados. **Definição e mensuração da pobreza na região metropolitana de São Paulo: uma abordagem multissetorial**. São Paulo, 1992.
- SOUZA, Paulo N. Pereira de. **LDB e educação superior**. São Paulo: Thomson Learning, 2001.