

CONTRIBUIÇÕES DA MINERAÇÃO DE DADOS PARA ANÁLISE DE DESEMPENHO SOCIAL

CONTRIBUTIONS OF DATA MINING FOR SOCIAL PERFORMANCE ANALYSIS

DEGENHART, Larissa¹
VOGT, Mara²
KREUZBERG, Fernanda³
RODRIGUES JUNIOR, Moacir Manoel⁴

RESUMO: O objetivo da pesquisa consiste em identificar as contribuições da análise fatorial, como ferramenta de mineração de dados, para a análise do desempenho social das empresas do setor de consumo cíclico. Para alcançar o objetivo, realizou-se uma pesquisa descritiva, documental e quantitativa. A população constituiu-se de empresas do setor de consumo cíclico devido à sua representatividade na economia nacional, sendo assim a amostra composta por 43 empresas. Para analisar o desempenho social, consideraram-se os indicadores oriundos da Demonstração do Valor Adicionado (DVA). Concluiu-se que para o período de 2009, os fatores selecionados foram Gastos com Pessoal e Gastos com Tributos. Para o período de 2010 e 2011, os fatores foram as variáveis Quociente de Valor Adicionado, Taxa de Variação do Valor Adicionado Bruto, Taxa de Variação Total do Valor Adicionado Bruto.

Palavras-chaves: Mineração de Dados. Análise Fatorial. Desempenho Social. Empresas Brasileiras. Consumo Cíclico.

ABSTRACT: *The objective of the research is to identify the contributions of the factorial analysis, as a data mining tool, for the analysis of social performance of the companies of the cyclical consumption sector. To achieve the purpose, we performed a descriptive, archival and quantitative research. As population, we delimited company in the consumer cyclic sector due representativeness in economy, therefore, the final sample is represented by 43 companies. To analyze the social performance, we used the indicators derived from the Value-Added Statement (VAS). We concluded that for 2009 the selected factors were Employees/Workers Spending and Tax Spending. For 2010 and 2011, the factors were Quotient of Value Added, Variation in Gross Added Value and Variation in Total Gross Value Added.*

¹ Doutoranda em Ciências Contábeis e Administração pela Universidade Regional de Blumenau – FURB. E-mail: lari_ipo@hotmail.com

² Doutoranda em Ciências Contábeis e Administração pela Universidade Regional de Blumenau – FURB. E-mail: maravogtcco@gmail.com

³ Doutoranda em Ciências Contábeis pela Universidade Federal de Santa Catarina – UFSC. E-mail: fernandakreuzberg@gmail.com

⁴ Doutor em Métodos Numéricos em Engenharia pela Universidade Federal do Paraná – UFPR. E-mail: moacir_ro@hotmail.com



Keywords: *Data Mining. Factor analysis. Social Performance. Brazilian Companies. Cyclic consumption sector.*

1 INTRODUÇÃO

A Contabilidade é de grande relevância no âmbito econômico, não somente como uma ferramenta necessária para o recolhimento de impostos, mas sim apresentando-se como uma ferramenta relevante para os usuários das informações contábeis, auxiliando no processo de tomada de decisões das organizações.

As informações das demonstrações financeiras não evidenciam de forma aprofundada o desempenho social das empresas. Geralmente, as informações referentes às relações das empresas com a sociedade, a geração de riqueza das empresas, bem como, a distribuição aos funcionários, governo e investidores, figuram em segundo plano. Esse cenário indica a contabilidade como uma importante ferramenta para a análise das relações sociais, não objetivando somente a lucratividade, mas também a evidenciação da sua relação com a sociedade e o meio ambiente (CUNHA, 2002; SANTOS, 2007).

Para Santos e Silva (2009), a Demonstração do Valor Adicionado (DVA) é considerada o demonstrativo mais adequado para as empresas evidenciarem o seu compromisso com o crescimento social, como ferramenta de análise de desempenho (SANTOS, 2007), visto que o desempenho social pode ser definido como o compromisso voluntário assumido por uma empresa, desenvolvido na esfera social, além de configurar um compromisso responsável com os grupos de interesse com quem a empresa interage (CUADRADO-BALLESTEROS et al., 2016). Contudo, o desempenho social das empresas, requer manutenção constante, ajustes, melhorias, assim como, realinhamentos (YUEN; THAI; WONG, 2016), diante das necessidades dos indicadores que abrangem a DVA.

Conforme elencado no Pronunciamento Técnico CPC 09 (2008), a DVA proporciona um conhecimento acerca das informações econômicas e sociais. Ressalta-se ainda que na DVA apresenta-se a destinação de recursos de natureza social efetuada pelas empresas, diferenciando-se da Demonstração de Resultado do Exercício (DRE). Além disso, uma das importantes decisões de gestão concentra-se no nível de compromisso que as empresas possuem com as práticas sociais, estas que determinam o nível de desempenho social corporativo (CUADRADO-BALLESTEROS et al., 2016).



Nesse sentido, esse demonstrativo visa apresentar atividades de Responsabilidade Social Corporativa (RSC), possibilitando que as organizações mantenham e melhorem sua relação com a sociedade e com as partes interessadas (ALLEN; PELOZA, 2015), pois a RSC trata da obrigação que a empresa possui para com o bem-estar destes. Também, é importante que exceda as expectativas éticas, legais, comerciais e públicas (KAHREH et al., 2014).

Diante disso, é válido e imprescindível que a administração repense seus sistemas de medição de desempenho, implemente estratégias eficientes, para aumentar seu desempenho social (MADRUGA, 2014). Empresas que divulgarem informações sobre o desempenho social, de acordo com Angelia e Suryaningsih (2015), serão bem vistas pela sociedade. Para Kadlubek (2015), essa ideia é um grande desafio que as empresas precisam enfrentar, mas é uma maneira de aumentar o valor da organização e fortalecer sua vantagem competitiva.

A DVA é reconhecida pela sua elevada compreensibilidade entre os diversos usuários (MEEK; GRAY, 1998; STAINBANK, 2009). Verifica-se, portanto ainda uma grande incipiência de estudos nessa área, representando um vasto campo para pesquisas, principalmente voltado para análise dos indicadores (BORBA, 2005).

De acordo com Dočekalová e Kocmanová (2016), o conjunto de indicadores de desempenho social permite uma análise detalhada, bem como, a determinação do impacto em várias áreas de desempenho corporativo. Contudo, as questões relacionadas ao desempenho social, mais precisamente à análise dos indicadores, às relações entre as partes interessadas, as ligações para com o desempenho e a ética das empresas, possibilitam o suporte para a prática da RSC (DUCASSY; MONTANDRAU, 2015).

Nesse cenário e frente a crescente importância da avaliação de desempenho para as empresas, insere-se a teoria de mineração dos dados, que diante de um conjunto de dados consegue otimizar uma análise, no intuito de gerar o melhor aproveitamento das informações, sem ocasionar perdas (GOEBEL; GRUENWALD, 1999; DIAS, 2002; CĂRBUREANU, 2012).

O foco principal da mineração de dados, de acordo com Amani e Fadlalla (2017) é a alavancagem de dados a fim de obter benefícios financeiros ou benefícios não financeiros. Verifica-se assim, a aplicação da mineração de dados em diversas áreas, incluindo a contabilidade. Ishibashi et al. (2016) salientam que as técnicas de mineração de dados desempenham um papel crucial na análise, bem como, na modelagem orientada por dados.



Nesse sentido, insere-se a seguinte questão problema: Quais as contribuições da análise fatorial, como ferramenta de mineração de dados, para a análise do desempenho social das empresas do setor de consumo cíclico? Para responder à problemática proposta, o presente estudo apresenta por objetivo identificar as contribuições da análise fatorial, como ferramenta de mineração de dados, para a análise do desempenho social das empresas do setor de consumo cíclico.

Justifica-se como objeto de análise o desempenho social pelo fato de atualmente muitas empresas incorporarem a RSC nas suas atividades, com vistas a obter vantagens competitivas, devido ao cenário turbulento e incerto ao qual estamos inseridos (MADORRAN; GARCIA, 2016). Dessa forma, as empresas buscam pela maximização do desempenho social, satisfazendo assim os acionistas e demais interessados, como a sociedade (RODRIGUEZ-FERNANDEZ, 2016).

Vale destacar também o crescimento e sofisticação de *software*, impulsionado por interesses de investigação e urgente utilização prática, social e econômica (GOEBEL; GRUENWALD, 1999). A mineração de dados busca evidenciar informações que possam demonstrar alguma tendência ou algum padrão, de modo a melhorar o processo de tomada de decisão e gerar um reflexo nos resultados das empresas. Esta técnica é de suma importância para a cultura de gestão e comportamento empresarial, o que permite trabalhar da melhor forma na consecução dos objetivos organizacionais (WITTEN; FRANK, 2005; BLANCO; ORTEGA, VANTI, JOHANN, 2013).

De forma complementar, Amani e Fadlalla (2017) destacam que a mineração de dados é uma das ferramentas atuais mais importantes para análise de negócios e essa importância é reconhecida pelos principais órgãos profissionais de contabilidade. Além disso, cabe frisar que a mineração de dados auxilia as organizações no que diz respeito à concentração de informações e conhecimentos relevantes disponíveis nos bancos de dados existentes.

Devido à vasta aplicabilidade da mineração de dados em diversas áreas do conhecimento, busca-se neste estudo apresentar a aplicação dessa ferramenta para a contabilidade, no intuito de analisar o desempenho social das empresas, contribuindo para a literatura nesta área. Para tanto, vale destacar que a utilização da Análise Fatorial se consubstancia no fato desta medir tanto a variância explicada por fatores quanto por variáveis.



O presente estudo está estruturado em outras cinco seções. A próxima seção, do referencial teórico, versará sobre o Desempenho Social e a Análise Fatorial como ferramenta de mineração de dados (*data mining*). Na sequência, apresenta-se a metodologia adotada nesta pesquisa. Na quarta seção apresenta-se a análise dos resultados e por fim, a última seção evidencia as considerações finais acerca do trabalho, destacando os principais resultados alcançados.

2 REFERENCIAL TEÓRICO

O alicerce teórico da pesquisa pauta-se na avaliação de desempenho das empresas, especificamente o desempenho social. Além disso, descreve-se acerca da importância da análise fatorial como ferramenta de mineração de dados (*data mining*).

2.1 Desempenho Social

Ao analisar a avaliação de desempenho, faz-se necessário avaliar primeiramente o instrumento de mensuração. Além do foco econômico as empresas apresentam ainda objetivos sociais, os quais podem ser utilizados de maneira integrada (GUERREIRO, 1989). Enquanto o aspecto econômico busca a maximização dos resultados, o aspecto social pauta-se nas relações das empresas com a sociedade. Buscando aprofundamento do aspecto social, ressalta-se a importância do Balanço Social, que conforme Tinoco (2001) consiste em um instrumento de gestão e informação cuja finalidade é a evidenciação de informações que refletem o desempenho de maneira transparente.

Portanto, recentemente grande importância é atribuída a uma das vertentes do Balanço Social que é a DVA que, segundo Santos (2007), é uma ferramenta de análise a ser utilizada por todos os usuários interessados nas relações entre empresa e sociedade. O autor complementa que a inserção da DVA, no conjunto de demonstrações poderá auxiliar na comparação das grandezas econômicas. Conforme Santos (1999), a DVA possibilita contribuir com diversas informações úteis acerca da relação entre empresa e sociedade.

O cerne da discussão do valor adicionado na literatura contábil internacional alcançou grande popularidade da década de 70, principalmente, nos países europeus (RIAHIBELKAOUI, 1999). Segundo Van Staden (2000), o conceito de valor adicionado foi inicialmente introduzido por Suojanen, com uma pesquisa publicada em 1954. Conforme



Suojanen (1954, p. 396), a “DVA seria um relatório complementar de análise do valor agregado na produção e sua origem ou da distribuição entre os participantes da organização”.

As informações referentes ao valor adicionado não se restringem ao âmbito econômico, englobando o aspecto social, por identificar a parcela da produção referente a cada um dos envolvidos no processo (HALLER; STOLOWY, 1998). Conforme o *Accounting Standards Steering Committee* (ASSC) (1975, p. 49), o “valor adicionado é a riqueza criada por conta própria pela entidade juntamente com os funcionários. Essa demonstração evidencia como o valor adicionado foi utilizado para pagar aqueles que contribuíram para a sua criação”.

Bao e Bao (1998) corroboram que o valor adicionado mede a contribuição de uma empresa para a sociedade, pois evidencia como os benefícios oriundos dos esforços das empresas foram partilhados entre as partes interessadas. Para Siqueira (2007) e Follmann, Casagrande e Paiva (2009), a importância social da demonstração está centrada na distribuição do valor adicionado, pois nesse enfoque a empresa evidencia a quem está destinando a sua riqueza.

Conforme De Luca (1998) e Santos (2007), nos aspectos internos inclui-se, a folha de pagamento sobre o valor adicionado, o valor adicionado por empregado, o valor adicionado em relação às vendas e à contribuição do lucro das operações para o valor adicionado. Enquanto os aspectos externos elencados pelos autores propiciam uma análise sobre a participação da empresa no cenário econômico, a contribuição da sociedade ponderada pelo pagamento de impostos, a remuneração da força de trabalho e a contribuição da empresa para a riqueza nacional. De acordo com o exposto, a DVA apresenta um conjunto de informações que auxiliam na análise e avaliação do desempenho econômico e social, devido ao enfoque fornecido para a geração de riqueza e sua posterior distribuição (LANA et al., 2011).

Nesse sentido, tem-se a Responsabilidade Social Corporativa, que é influenciada pela forma como as empresas alinham seus valores e comportamentos, a partir do esperado pelos envolvidos e interessados. Para tanto, é necessário satisfazer os interesses deles, o que apresentará reflexos no desempenho social (AHMED; ISLAM; HASAN, 2012). A partir dessa perspectiva, Kahreh et al. (2014) afirmam que a RSC é importante, principalmente, pela competitividade que há entre as empresas.

Conforme Erhemjamts, Li e Venkateswaran (2013), empresas bem-sucedidas podem gastar mais no que tange aos aspectos ambientais que as demais, contudo, a RSC ajuda no



sucesso financeiro das organizações. Da mesma forma, grandes empresas são capazes de implementar atividades de Responsabilidade Social Corporativa, estas que possibilitam ações e resultados para com a sociedade (YOUN; HUA; LEE, 2015). Contudo, as Pequenas e Médias Empresas (PME) se diferem devido ao tipo de desempenho (financeiro, social, ambiental, competitivo) (MADUENO et al., 2016).

Diante do exposto, percebe-se a relevância da DVA como uma ferramenta, um instrumento de mensuração utilizado para apresentar um conjunto de informações aos seus interessados, sendo esta uma vertente do Balanço Social. Nesse sentido, a mineração dos dados poderá auxiliar as organizações a identificar e utilizar banco de dados que contenha as informações necessárias para a gestão e tomada de decisões contábeis.

2.2 Análise Fatorial como ferramenta de mineração de dados (*Data Mining*)

Conforme Dias (2002), a mineração de dados, é uma parte do processo de descoberta de conhecimento em banco de dados, pois os resultados obtidos a partir da mineração de dados podem ser utilizados no processamento de pedidos de informação, gerenciamento de informação, tomada de decisão e controle de processos.

Por meio da mineração de dados é que surgiu a ideia de construir programas, que filtrem automaticamente os bancos de dados em busca de padrões ou regularidades. Não há mágica de como aprender a lidar com a mineração de dados, e sim técnicas simples e práticas que podem extrair informações úteis a partir de dados brutos (WITTEN; FRANK, 2005). De acordo com Hand (2007), a mineração de dados é ainda, a descoberta de estruturas valiosas, inesperadas, interessantes em grandes conjuntos de dados.

Para tanto, tem-se a análise fatorial de dados, como uma ferramenta de *data mining*, ou seja, uma técnica que permite agrupar, com base na matriz de correlações, um conjunto grande de variáveis em um conjunto menor de variáveis não observáveis. Alguns trabalhos denominam esse conjunto menor por variáveis latentes, fatores latentes, ou simplesmente fatores. Dessa forma, este método de análise busca o agrupamento de informações correlacionadas, diferindo, por exemplo, de técnicas com análise de *clusters* que apresenta por objetivo o agrupamento de observações.

De acordo com Johnson e Wichern (1988), o surgimento do método é difuso, entretanto, sabe-se que os primeiros estudos referentes ao tema foram de Karl Pearson e de Charles



Spearman. O advento desse conhecimento ocorreu com os cientistas de psicometria, em que os pesquisadores conseguiram perceber diferentes características entre alguns indivíduos, principalmente, no que se refere às ponderações que cada variável receberia na formação de um escore único de descrição da inteligência dos indivíduos investigados.

Seja $\underline{X} \sim * (\underline{\mu}, \Sigma)$, um vetor de informações observadas, com qualquer distribuição de probabilidade. O modelo Fatorial considera a existência de uma variável aleatória latente de informações não observadas, F_1, F_2, \dots, F_m , (para $m \leq p$) sendo estas chamadas de Fatores Comuns e p fontes de variações, o modelo ainda considera $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ chamados de erros ou fatores específicos. Desta forma a Análise Fatorial é dada pelo sistema (Equação 1):

$$\begin{cases} X_1 - \mu_1 = l_{11}F_1 + l_{12}F_2 + \dots + l_{1m}F_m + \varepsilon_1 \\ X_2 - \mu_2 = l_{21}F_1 + l_{22}F_2 + \dots + l_{2m}F_m + \varepsilon_2 \\ \vdots \\ X_p - \mu_p = l_{p1}F_1 + l_{p2}F_2 + \dots + l_{pm}F_m + \varepsilon_p \end{cases} \quad (1)$$

Onde:

μ_i – média da i -ésima variável;

ε_i – erro ou fator específico da i -ésima variável;

F_j – j -ésimo fator comum;

l_{ij} – carregamento ou carga fatorial da i -ésima variável no j -ésimo fator.

Em notação matricial o modelo pode ser descrito como (Equação 2):

$$\underline{X} - \underline{\mu} = L\underline{F} + \underline{\varepsilon} \quad (2)$$

Onde: \underline{X} e $\underline{\mu}$ são respectivamente o vetor de informações observadas e o vetor de médias das informações (Equação 3).

$$L = \begin{bmatrix} l_{11} & l_{12} & \dots & l_{1m} \\ l_{21} & l_{22} & \dots & l_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p1} & l_{p2} & \dots & l_{pm} \end{bmatrix} \underline{F} = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix} \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix} \quad (3)$$



Assume-se ainda que a Esperança dos fatores é nula, $E(\underline{F}) = 0$, e a variância dos fatores é igual a identidade, $V(\underline{F}) = E(\underline{F}\underline{F}^T) = I$. Para o erro, considera-se a situação em que a média do erro é nula, $E(\underline{\varepsilon}) = 0$, e a variância é dada por (Equação 4):

$$V(\underline{\varepsilon}) = E(\underline{\varepsilon}\underline{\varepsilon}^T) = \Psi = \begin{bmatrix} \Psi_1 & 0 & \cdots & 0 \\ 0 & \Psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Psi_p \end{bmatrix} \quad (4)$$

Ainda cabe considerar que \underline{F} e $\underline{\varepsilon}$ são independentes, ou seja, $cov(\underline{F}, \underline{\varepsilon}) = 0$. Considerando essas pressuposições o modelo descrito acima, é o modelo Fatorial Ortogonal, pois considera que os fatores são independentes entre si (JOHNSON; WICHERN, 1988).

A Técnica das Componentes Principais é o principal método utilizado para a estimação dos fatores, pois utiliza o conceito de autovalores e autovetores. A estimação dos fatores também pode ser feita por meio do método da máxima verossimilhança, entretanto, esse método pressupõe normalidade das variáveis, e o primeiro não necessita dessa condição.

A estimação dos fatores por meio das Componentes Principais é dada com base nos autovalores e autovetores associados à matriz de covariâncias das variáveis. Sejam os pares de autovalores e autovetores $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$, onde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$. Considerando que o número de fatores $m < p$ a matriz de pesos ou carregamentos é estimada por $\hat{L} = \hat{C}\hat{D}_\lambda^{\frac{1}{2}}$ onde (Equação 5):

$$\hat{C} = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1p} \\ e_{21} & e_{22} & \cdots & e_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ e_{p1} & e_{p2} & \cdots & e_{pp} \end{bmatrix} \hat{D}_\lambda^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_p} \end{bmatrix} \quad (5)$$

Para estimar a matriz de covariâncias é necessária a centralização dos dados em relação à média, ou à padronização dos dados. No primeiro caso, subtrai-se da observação a média obtida, e no segundo caso, subtrai-se a média e divide-se pelo desvio padrão. Quando adotada a padronização a matriz de covariâncias S obtida, é igual a matriz R que é a matriz de correlação. A matriz de covariâncias é necessária para a obtenção da variância específica ou não explicada, $\hat{\Psi} = S - \hat{L}\hat{L}^T$.



Cada variável possui uma associação para com os fatores do modelo, esta associação é dada pelos valores de l_{ij} . Com base nesses valores é possível determinar as comunalidades das variáveis para cada um dos fatores, seu cálculo é dado por $\hat{h}_i^2 = \hat{l}_{i1}^2 + \hat{l}_{i2}^2 + \dots + \hat{l}_{ip}^2$. As comunalidades descrevem o quanto da variância da variável é explicada pelos fatores selecionados.

Muitas são as discussões referentes à seleção do número de fatores utilizados na análise. O principal critério estabelecido e o mais utilizado é apresentado por Hair Jr., Anderson, Tatham e Black (2005) e Fávero, Belfiore, Silva e Chan (2009). Este é o critério de Kaiser, que estabelece como fatores importantes, apenas aqueles que possuem autovalores associados maiores do que 1.

Outros conceitos permeiam a análise fatorial, tais como a rotação dos dados para o aumento do poder explicativo das informações. Também a matriz dos resíduos que descreve o quanto as variáveis ainda não foram descritas pelo modelo. Entretanto, como este trabalho pretende estabelecer a possibilidade da análise fatorial como ferramenta de *data mining*, os conceitos já apresentados são fundamentais para esse entendimento da aplicação.

A mineração de dados por meio da análise fatorial já foi utilizada em alguns trabalhos, como Bezerra e Corrar (2006). O modelo permitido pelo método é baseado no percentual da variância explicada, quanto maior o número de fatores utilizados maior também deve ser o incremento no percentual de variância explicada. A seleção dos indicadores pode ser estabelecida conforme um ponto de corte preestabelecido pelos pesquisadores, considerando o número de fatores utilizados e a comunalidade de cada variável para os fatores.

Favorece a utilização da análise fatorial, frente a outros métodos, por exemplo, as componentes principais, o fato de ela medir tanto a variância explicada por fatores e por variáveis, enquanto o segundo método padroniza apenas a variância das componentes.

3 PROCEDIMENTOS METODOLÓGICOS

O presente estudo delinea-se como uma pesquisa descritiva, documental e quantitativa. Para a delimitação da população a ser investigada no estudo, utilizou-se por critério de seleção o setor econômico de maior representatividade, segundo a classificação estabelecida pela B3. Dessa forma, dentre um conjunto de nove setores que englobam 350 empresas não financeiras



com ações negociadas na Bolsa de Valores, verifica-se que o setor de consumo cíclico apresenta a maior representatividade com um total de 68 empresas.

Selecionada a população do estudo, delimita-se a amostra da presente pesquisa. A delimitação de uma amostra faz-se necessária, pois nem todas as empresas apresentaram as informações necessárias para o cálculo dos indicadores sociais. Assim sendo, a amostra a ser investigada composta por 43 empresas do setor de Consumo Cíclico da B3. Essa amostra apresenta uma representatividade de 63,23% em relação à população estipulada. O setor de consumo cíclico representa 20,87% da economia brasileira, em uma análise dos nove setores cujas empresas possuem ações negociadas na B3.

Vale ressaltar que a escolha desse setor se pauta na sua relevância para o mercado de capitais, tendo em vista que a B3 criou o Índice Bovespa de Consumo, também conhecido como (ICON), este que mede o comportamento das ações de empresas mais representativas tanto do setor de consumo cíclico como do setor de consumo não cíclico.

A coleta dos dados para análise do desempenho social das empresas ocorreu por meio de uma pesquisa documental, utilizando-se dados secundários, que se encontram disponíveis no sítio da B3. Dessa forma, os indicadores sociais das empresas foram calculados com base nas informações obtidas pelos relatórios contábeis: o Balanço Patrimonial, a Demonstração de Resultado e a Demonstração do Valor Adicionado.

Os indicadores selecionados evidenciam o desempenho social das empresas, pautado em uma perspectiva da Demonstração do Valor Adicionado. Assim, tem-se um conjunto de seis indicadores: i) Gastos com Pessoal (GP), ii) Gastos com Tributos (GT), iii) Quociente do Valor Adicionado (QVA), iv) Taxa de Variação do Valor Adicionado Bruto (TVAB), v) Taxa de Variação Total do Valor Adicionado Bruto (TVT) e vi) Potencial do Ativo em Gerar Riqueza (PAGR). A utilização desses indicadores está evidenciada em Morley (1979), De Luca (1998), Santos (1999), Riahi-Belkaoui (1999), Cunha (2002), Aswegen et al. (2005) e Santos (2007).

Os relatórios analisados referem-se às demonstrações contábeis consolidadas do período de 2008 a 2011, pois para o cálculo de alguns indicadores sociais foram necessárias informações dos demonstrativos referentes ao período de 2008. Perante os dados documentais realizou-se uma análise transversal dos dados, utilizando por base de análise o período de 2009, 2010 e 2011.



Depois de efetivada a coleta, a etapa seguinte consiste na análise dos dados. Nessa etapa, aplicou-se a ferramenta de mineração de dados (*data mining*), utilizando-se a técnica de análise fatorial, conforme descrito anteriormente.

4 ANÁLISE E INTERPRETAÇÃO DOS DADOS

Com a importância que os aspectos sociais das empresas vêm apresentando, este estudo busca de forma geral analisar as informações contidas na DVA, de forma a selecionar os indicadores que possuem maior poder informacional. Primeiramente, verificou-se se os dados seguem a distribuição normal de probabilidade, e se os pressupostos de adequação da Amostra KMO e o teste de Esfericidade de Bartlett são satisfeitos, conforme descrito na Tabela 1.

Tabela 1 – Verificação da normalidade e testes de KMO e Esfericidade de Bartlett

		GP	GT	QVA	TVAB	TVT	PAGR
Estatística de teste	2009	0,442	0,327	0,126	0,197	0,512	0,536
	2010	0,461	0,530	0,372	0,411	0,449	0,536
	2011	0,095	0,148	0,162	0,174	0,218	0,537
Sig	2009	0,000	0,000	0,082	0,000	0,000	0,000
	2010	0,000	0,000	0,000	0,000	0,000	0,000
	2011	0,200	0,019	0,006	0,002	0,000	0,000
					2009	2010	2011
Medida Kaiser-Meyer-Olkin de adequação de amostragem.					0,47	0,56	0,44
Teste de esfericidade de Bartlett	Aprox. Qui-quadrado				116,93	362,82	26,96
	df				15,00	15,00	15,00
	Sig.				0,00	0,00	0,03

Fonte: Resultados da pesquisa.

Os resultados apresentados na Tabela 1, evidenciam que as variáveis são, na maior parte dos anos observados, pertencentes a uma distribuição normal de probabilidade. Os casos dos indicadores GP 2011 e QVA 2009 foram considerados normais, mesmo que suas significâncias auferem valores superiores a 0,05, haja vista que o desempenho histórico desses indicadores descreve a normalidade dos dados.

Para os pressupostos associados à Análise Fatorial, destacados por Hair Jr. et al. (2005), Fávero et al. (2009), a esfericidade de Bartlett foi satisfeita nos três anos analisados. O Teste de KMO foi satisfatório apenas para o ano de 2010, entretanto, assumiu-se como satisfeita a pressuposição, visto que os valores auferidos foram muito próximos ao valor crítico de 0,5.



Analisou-se também o nível informacional que cada fator possui sobre o poder explicativo que cada variável possui por meio das comunalidades. Os resultados referentes às análises das variâncias de cada um dos fatores estão apresentados na Tabela 2.

Tabela 2 – Autovalores, Variância Observada e Variância Acumulada dos Fatores

Fatores	Autovalores			Variância Observada			Variância Acumulada		
	2009	2010	2011	2009	2010	2011	2009	2010	2011
1	2,107	2,789	1,728	35,123	46,475	28,798	35,123	46,475	28,798
2	1,544	1,992	1,103	25,731	33,206	18,387	60,854	79,681	47,185
3	1,019	0,996	1,038	16,983	16,592	17,306	77,837	96,273	64,491
4	0,958	0,183	1,008	15,965	3,044	16,795	93,801	99,318	81,287
5	0,323	0,039	0,813	5,377	0,642	13,550	99,178	99,959	94,837
6	0,049	0,002	0,310	0,822	0,041	5,163	100,000	100,000	100,000

Fonte: Resultados da pesquisa.

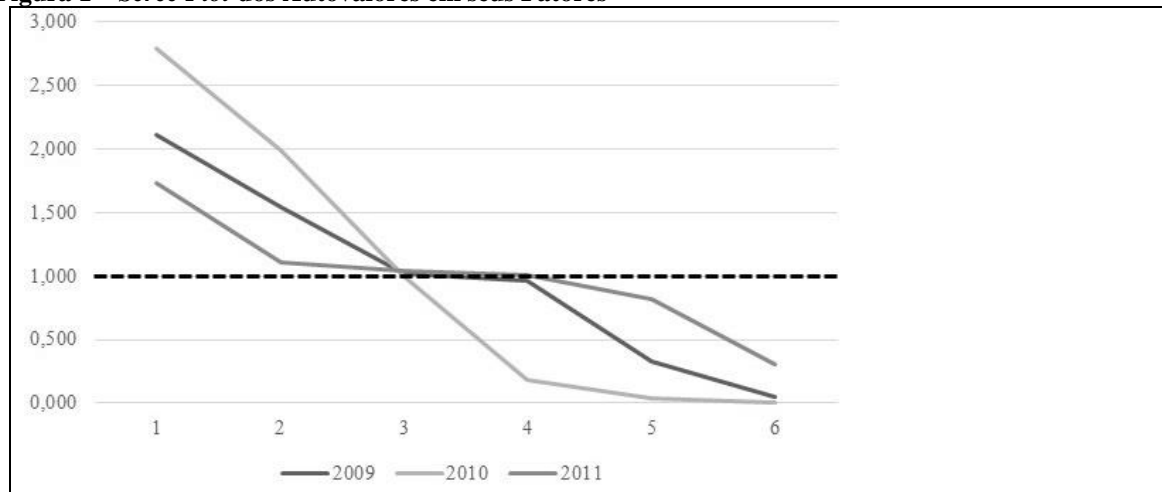
De acordo com os resultados apresentados o ano de 2010 apresentou maior conformidade com os modelos clássicos de Análise Fatorial. Foi esse o ano que auferiu maior percentual de variância sobre o primeiro fator analisado.

O ano de 2011 apresenta-se como o ano em que os resultados foram mais atípicos, visto que o percentual de variância explicada pelos fatores está distribuído com maior uniformidade. Fato percebido, pois, a variância acumulada de 2011 ultrapassa a faixa dos 50% apenas no terceiro fator que é metade do total de fatores analisados.

Para 2009, a análise está em um meio termo do ano de 2010 e 2011, deve-se considerar 2009 um ano atípico, visto que a DVA estava em fase de implementação. A Figura 1 contribui para a análise do peso informacional dos fatores analisados, ela apresenta o *Scree Plot* dos três anos analisados.



Figura 1 – Scree Plot dos Autovalores em seus Fatores



Fonte: Resultados da pesquisa.

O *Scree Plot*, como está associado ao valor da variância dos fatores, confirma a concentração de informação maior no ano de 2010, em comparação para com os demais anos. Verifica-se a barra tracejada representa a linha de corte pelo critério de Kaiser para definição do número de fatores. Poderia ser estabelecido para cada ano os seguintes fatores, considerando o critério de significativa perda de informação dos fatores, 2009 estabeleceria quatro fatores, 2010 estabeleceria três fatores, 2011 com cinco fatores.

Verificado o número de fatores adotados em cada um dos anos analisados, a seleção dos indicadores pode ser feita por meio da análise das comunalidades. A comunalidade de uma variável, segundo Johnson e Wichern (1988), representa o percentual de variância explicada pelos fatores considerados. A Tabela 3 descreve a comunalidade cumulativa das variáveis para com os fatores.

Tabela 3 – Comunalidades das Variáveis por Fator de Análise

	1 Fator	2 Fatores	3 Fatores	4 Fatores	5 Fatores	6 Fatores
Análise 2009						
GP	0,817*	0,970	0,970	0,972	0,975	1,000
GT	0,818*	0,965	0,970	0,970	0,976	1,000
QVA	0,362	0,824	0,832	0,832	1,000	1,000
TVAB	0,030	0,185	0,428	0,977	1,000	1,000
TVT	0,080	0,699	0,762	0,878	1,000	1,000
PAGR	0,000	0,009	0,709	0,999	1,000	1,000
Análise 2010						
GP	0,016	0,997	0,999	0,999	0,999	1,000
GT	0,008	0,997	0,998	0,999	0,999	1,000



QVA	0,888*	0,894	0,895	0,995	1,000	1,000
TVAB	0,906*	0,910	0,910	0,991	1,000	1,000
TVT	0,967*	0,973	0,974	0,975	1,000	1,000
PAGR	0,004	0,009	1,000	1,000	1,000	1,000
Análise 2011						
GP	0,117	0,534	0,549	0,620	0,999	1,000
GT	0,062	0,082	0,780	0,833	0,992	1,000
QVA	0,016	0,656	0,709	0,809	0,987	1,000
TVAB	0,683*	0,683	0,834	0,834	0,872	1,000
TVT	0,817*	0,841	0,842	0,842	0,845	1,000
PAGR	0,032	0,035	0,155	0,939	0,995	1,000

Fonte: Resultados da pesquisa.

Esta etapa da análise permite um comparativo importante para a condução da pesquisa. É possível perceber pela Tabela 3, que o número de fatores importantes para a análise dos dados estabelecido com base na Figura 1, busca possibilitar um nível aceitável de comunalidades de cada variável. Ou seja, seguindo a análise fatorial de forma padrão não seria possível a mineração dos dados, visto que as comunalidades estão em alto poder explicativo.

Pelas Componentes Principais, como método de estimação dos fatores é possível estabelecer como critério a seleção da variável cuja comunalidade é superior a 50% no primeiro fator, visto que o fator 1 é o que possui maior peso informacional.

Dessa maneira, é possível descrever o conjunto de indicadores selecionados em cada ano como o seguinte:

2009 – Gasto com Pessoal (GP), Gasto com Tributos (GT);

2010 – Quociente de Valor Adicionado (QVA), Taxa de Variação do Valor Adicionado Bruto (TVAB), taxa de Variação Total do Valor Adicionado Bruto (TVT);

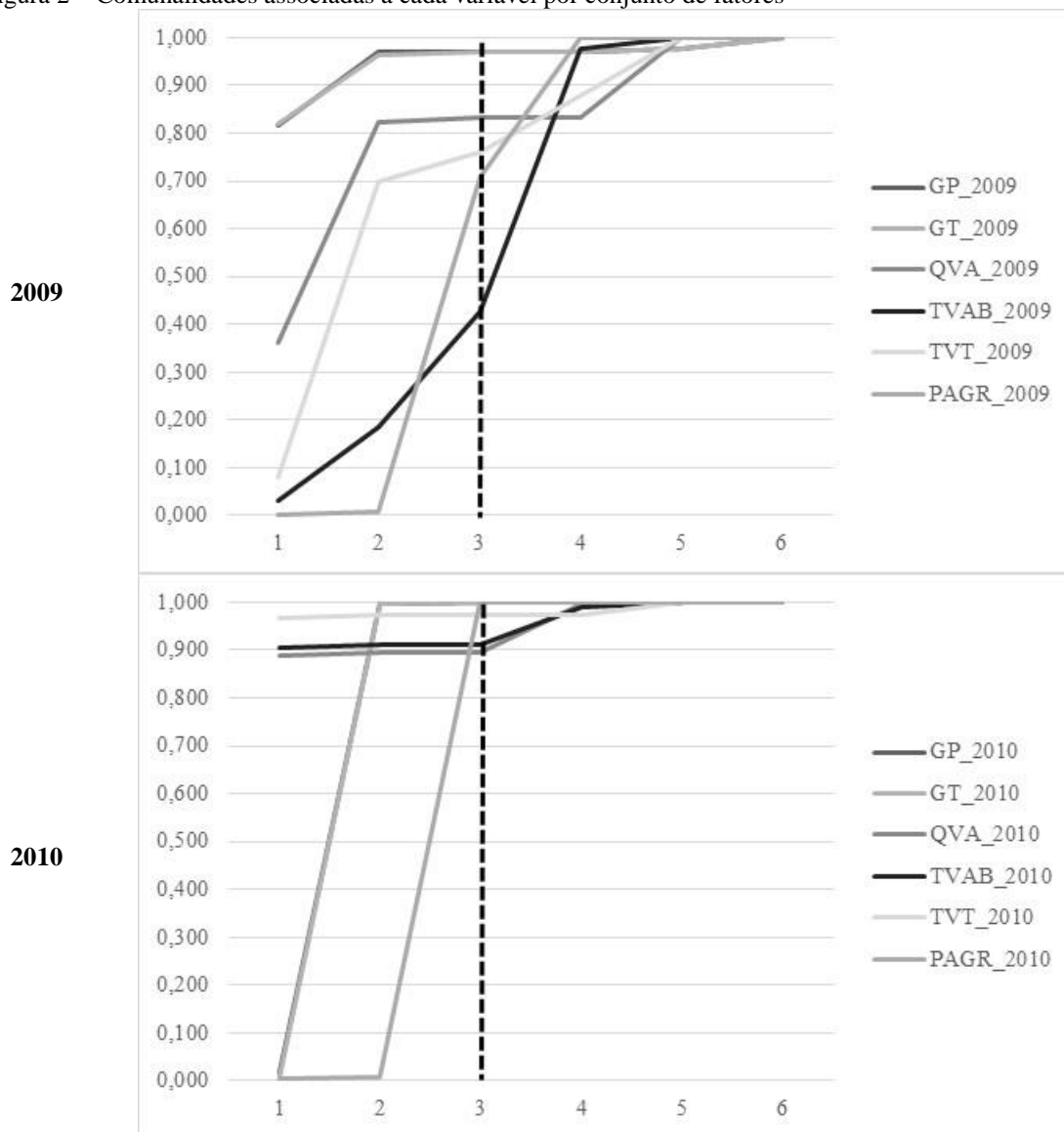
2011 – Taxa de Variação do Valor Adicionado Bruto (TVAB), Taxa de Variação Total do Valor Adicionado Bruto (TVT).

Conseguiu-se nos resultados a conformidade dos anos de 2010 e 2011. O indicador Potencial do Ativo em Gerar Riqueza (PAGR) não foi selecionado como informação importante dentre o conjunto de informações analisadas. Cabe ressaltar que a mineração de dados por meio da análise fatorial considera como principal critério a variância das informações em conjunto, com baixa covariância.

A Figura 2 evidencia a evolução das comunalidades para o conjunto de fatores considerados, possibilitando a percepção da perda de poder de mineração dos dados conforme o avanço do número de fatores admitidos pelo modelo.

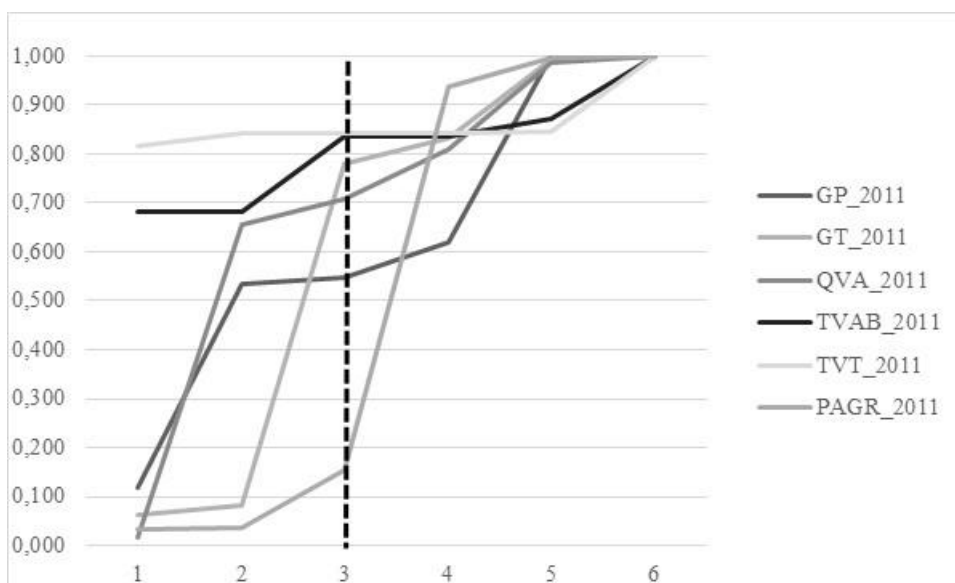


Figura 2 – Comunalidades associadas a cada variável por conjunto de fatores





2011



Fonte: Resultados da pesquisa.

Conforme os resultados apresentados e com o conjunto de fatores estabelecidos, verifica-se que durante um período agruparam-se as variáveis de gastos com pessoal e com tributos.

Por outro lado, pela análise dos períodos posteriores, estabeleceu-se um grupo de fatores com os indicadores: Quociente de Valor Adicionado (QVA), Taxa de Variação do Valor Adicionado Bruto (TVAB), taxa de Variação Total do Valor Adicionado Bruto (TVT).

O QVA é denominado como a produtividade das vendas, pois indicará para um determinado período, quanto cada unidade monetária vendida se transformou em riqueza para a empresa, representando a eficiência econômica de uma empresa (ASWEGEN; STEYN; HAMMAN, 2005; SANTOS, 2007). OTVAB demonstra o acréscimo (decréscimo) no nível de atividade econômica de uma empresa, evidenciando o desempenho de um período para o outro (CARVALHO, 2007; SANTOS, 2007). A TVT analisa a taxa de incremento (ou redução do valor adicionado), bem como sua relação com as vendas (CARVALHO, 2007).

Por outro lado, o fato de o Potencial do Ativo em Gerar Riqueza (PAGR) não ter sido selecionado como informação importante dentre o conjunto de informações analisadas pode estar associado ao fato de ser uma variável com informação financeira, de contribuição dos ativos das empresas, buscando demonstrar a sua eficiência na utilização enquanto geração de valor.

Dessa forma, os resultados obtidos indicam que a análise fatorial pode ser utilizada como forma de mineração de dados, para avaliar o âmbito social das empresas. As informações



referentes à DVA são importantes como um todo, considerando que nela está descrita o impacto social da empresa no processo de geração de riqueza. Entretanto o processo de mineração de dados possibilita uma análise mais aprofundada e mais rica dos indicadores formados a partir dela.

5 CONSIDERAÇÕES FINAIS

O objetivo deste estudo buscou identificar as contribuições da análise fatorial, como ferramenta de mineração de dados, para a análise do desempenho social das empresas do setor de consumo cíclico. Para tanto, realizou-se uma análise das empresas listadas na B3 no período de 2009 a 2011, sendo a amostra do estudo composta por 43 empresas.

Os resultados da pesquisa apresentaram que a Análise Fatorial pode ser utilizada de forma coerente como mineração de dados. Foram analisados três anos, 2009, 2010 e 2011, e os resultados foram distintos para cada um dos períodos. O ano de 2009 foi a composição dos resultados de 2010 e de 2011, e para 2010, os resultados apontam para maior concentração da informação nos primeiros fatores e para 2011 ocorreu uma maior suavização na distribuição da variância.

Os resultados ainda apontaram que o critério de Kaiser, que recomenda utilizar apenas fatores com autovalor superior a 1, não deve ser utilizado para a mineração dos dados, pois as comunalidades tendem a ser altas para todas as variáveis. Para melhor mineração foi adotado apenas o fator com autovalor máximo, visto que este possui maior peso informacional.

Os indicadores selecionados para o ano de 2009 foram gastos com pessoal e tributos, possuindo maior variabilidade, por causa desse fato que os indicadores foram selecionados como possuindo maior informação. Deve ser levado em consideração que 2009 foi o segundo ano de divulgação da DVA, ou seja, ainda estava sendo implementada.

Para o ano de 2010, os indicadores selecionados foram os de Quociente de Valor Adicionado, Taxa de Variação do Valor Adicionado e Taxa de Variação Total do Valor Adicionado Bruto, todas as informações diretamente relacionadas à geração de riqueza por parte das empresas.

O ano de 2011 só não selecionou o indicador de Quociente de valor Adicionado, entretanto o enfoque dos indicadores foi também voltado à geração de riqueza das empresas. O



indicador do potencial de geração de riqueza das empresas por parte dos ativos está associado ao fato de ser uma informação mais financeira.

Dessa forma, conclui-se que as informações predominantes para os anos de 2010 e 2011 foram referentes à geração de valor por parte das empresas, esses dois períodos podem ser considerados mais consistentes, visto que o processo de elaboração da DVA está mais institucionalizado e reflete um processo mais estabelecido.

Por sua vez, a análise fatorial possibilitou a seleção das informações relevantes com base na matriz de variância dos dados. Esta técnica como ferramenta de mineração de dados deve ser mais discutida, visto a necessidade de um critério para a seleção do número de fatores utilizados pela análise, pois conforme os resultados apontaram o critério de Kaiser não auxilia na redução de dimensão das informações.

O presente trabalho buscou contribuir para despertar no meio acadêmico o interesse em selecionar os indicadores de desempenho social das empresas, mediante aplicação da análise fatorial como ferramenta de *data mining*. Da mesma forma, o estudo possibilitará o avanço das discussões sobre o desempenho social e as ferramentas de *data mining*, visto ser um tema com potencial a ser explorado e discutido.

Como limitações da pesquisa, elenca-se a seleção da amostra, composta apenas por empresas de um único setor. Além disso, verificou-se na literatura um número reduzido de estudos relacionados à análise dos indicadores de desempenho social com a mineração de dados, o que dificultou a escolha das variáveis que possibilitassem responder, de forma mais precisa, à problemática deste estudo.

Desse modo, pesquisas futuras podem corrigir essas limitações, pois com uma nova amostra e novos indicadores, será possível utilizar este estudo para fins de comparabilidade dos resultados. Além disso, sugere-se a aplicação de outras técnicas de mineração de dados, para a mesma amostra, no intuito de verificar a eficiência desta ferramenta.

REFERÊNCIAS

AHMED, Sarwar Uddin; ISLAM, Zahidul Md; HASAN, Ikramul. Corporate Social Responsibility and financial performance linkage: Evidence from the banking sector of Bangladesh. **Journal of Organizational Management**, v. 1, n. 1, p. 14-21, 2012.



ACCOUNTING STANDARDS STEERING COMMITTEE. **The corporate report: a discussion paper published for comment.** London: Accounting Standards Steering Committee of the Institute of Chartered Accountants in England and Wales, 1975.

ALLEN, Alexis M.; PELOZA, John. Someone to watch over me: The integration of privacy and Corporate Social Responsibility. **Business Horizons**, v. 58, n. 6, p. 635-642, 2015.

AMANI, F. A.; FADLALLA, A. M. Data mining applications in accounting: A review of the literature and organizing framework. **International Journal of Accounting Information Systems**, v. 24, p. 32-58, 2017.

ANGELIA, Dessy; SURYANINGSIH, Rosita. The Effect of Environmental Performance and Corporate Social Responsibility Disclosure Towards Financial Performance (Case Study to Manufacture, Infrastructure, And Service Companies That Listed At Indonesia Stock Exchange). **Procedia-Social and Behavioral Sciences**, v. 211, p. 348-355, 2015.

ASWEGEN, N. V.; STEYN, B. W.; HAMMAN, W. D. Trends in the distribution of added value of listed industrial companies – 1990 to 2002. **South African Journal of Business Management**, v. 36, n. 2, p. 85-94, 2005.

BAO, B.-H.; BAO, D.-H. Usefulness of value added and abnormal economic earnings: an empirical examination. **Journal of Business Finance & Accounting**, v. 25, n.1-2, p. 251-264, 1998.

BEZZERA, F. A.; CORRAR, L. Utilização da análise fatorial na identificação dos principais indicadores para avaliação do desempenho financeiro: uma aplicação nas empresas de seguros. **Revista Contabilidade & Finanças**, v. 42, p. 50-62, 2006.

BLANCO, E. R. R.; ORTEGA, A. C.; VANTI, A. A.; JOHANN, S. Imágenes organizacionales en empresas brasileñas: detección y análisis con técnicas de minería de datos. **Revista de Ciências da Administração**, v. 15, n. 37, p. 105-120, 2013.

CĂRBUREANU, M. The Annual Inflation Rate Analysis Using Data Mining. **Techniques. Economic Insights – Trends and Challenges**, v. 1, n. 4, p. 121-128, 2012.

CARVALHO, F. R. **Demonstração do valor adicionado e informações sociais:** contribuição de evidenciação para as empresas na região de Uberaba. 2007. Dissertação (Mestrado) - Pontifícia Universidade Católica de São Paulo, São Paulo, SP, 2007.

COMITÊ DE PRONUNCIAMENTOS CONTÁBEIS. *Pronunciamento técnico CPC 09. Demonstração do Valor Adicionado*, 2008. Disponível em: <<http://www.cpc.org.br/CPC>>. Acesso em: 24 nov.2014.

COSENZA, J. P. A eficácia informativa da demonstração do valor adicionado. **Revista Contabilidade & Finanças**, Edição comemorativa, p. 7-29, 2003.



CUADRADO-BALLESTEROS, B.; RODRÍGUEZ-ARIZA, L.; GARCÍA-SANCHEZ, I.-M.; MARTÍNEZ-FERRERO, J. The mediating effect of ethical codes on the link between family firms and their social performance. **Long Range Planning**, p. 1-10, 2016.

CUNHA, J. V. A. **Demonstração contábil do valor adicionado (DVA):** um instrumento de mensuração da distribuição da riqueza das empresas para os funcionários. 2002. Dissertação (Mestrado) - Universidade de São Paulo, São Paulo, SP, 2002.

DE LUCA, M. M. M. **Demonstração do Valor Adicionado:** do cálculo da riqueza criada pela empresa. São Paulo: Atlas, 1998.

DIAS, M. M. Parâmetros na escolha de técnicas e ferramentas de mineração de dados. **Acta Scientiarum**, v. 24, n. 6, p. 1715-1725, 2002.

DUCASSY, I.; MONTANDRAU, S. Corporate social performance, ownership structure, and corporate governance in France. **Research in International Business and Finance**, v. 34, n. 1, p. 383-396, 2015.

DOČEKALOVÁ, M. P.; KOČMANOVÁ, A. Composite indicator for measuring corporate sustainability. **Ecological Indicators**, v. 61, n. 1, p. 612-623, 2016.

ERHEMAMTS, Otgontsetseg; LI, Qian; VENKATESWARAN, Anand. Corporate social responsibility and its impact on firms' investment policy, organizational structure, and performance. **Journal of Business Ethics**, v. 118, n. 2, p. 395-412, 2013.

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

GOEBEL, M.; GRUENWALD, L. A survey of data mining and knowledge discovery software tools. **ACM SIGKDD Explorations Newsletter**, v. 1, n. 1, p. 20-33, 1999.

GUERREIRO, R. **Modelo conceitual de sistema de informação de gestão econômica:** uma contribuição a teoria da comunicação da contábil. 1989. Tese (Doutorado) - Universidade de São Paulo, São Paulo, SP, 1989.

HAIR JR., J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. **Análise multivariada de dados.** 5. ed. Porto Alegre: Bookman, 2005.

HALLER, A.; STOLOWY, H. Value added in financial accounting a comparative study between Germany and France. **Advances in International Accounting**, v. 11, n. 1, p. 23-51, 1998.

HAND, D. J. Principles of data mining. **Drug safety**, v. 30, n. 7, p. 621-622, 2007.

ISHIBASHI, K.; IWASAKIA, T.; OTOMASAA, S.; YADAA, K. Model Selection for Financial Statement Analysis: Variable Selection with Data Mining Technique. **Procedia Computer Science**, v. 96, p. 1681-1690, 2016.



JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 2nd ed. New Jersey: Prentice-Hall, 1988.

KADŁUBEK, Marta. The Essence of Corporate Social Responsibility and the Performance of Selected Company. **Procedia-Social and Behavioral Sciences**, v. 213, p. 509-515, 2015.

KAHREH, Mohammad Safari; BABANIAB, Asghar; TIVEC, Mohammad; MIRMEHDID, Seyed Mehdi. An examination to effects of Gender Differences on the Corporate Social Responsibility (CSR). **Procedia-Social and Behavioral Sciences**, v. 109, p. 664-668, 2014.

LANA, C. A. M.; BERNARDO, D. C. R.; NAZARETH, L. G. C.; MENDONÇA, F. M. Um estudo das ações para divulgar e consolidar o balanço social no Brasil. **Revista Gestão Organizacional**, v. 4, n. 2, p. 311-329, 2011.

MADORRAN, Cristina; GARCIA, Teresa. Corporate Social Responsibility and Financial Performance: the Spanish Case. **RAE - Revista de Administração de Empresas**, v. 56, n. 1, p. 20-28, 2016.

MADRUGA, Sérgio Rossi. **Estágio de Maturidade da Responsabilidade Social Corporativa e o Desempenho Econômico-Financeiro**: estudo em empresas Brasileiras. 184 f. 2014. Tese (Doutor em Ciências) - Programa de Pós-Graduação em Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, 2014.

MADUEÑO, Jesús Herrera; JORGEA, Manuel Larrán; CONESAB, Isabel Martínez; MARTÍNEZ-MARTÍNEZA, Domingo. Relationship between Corporate Social Responsibility and competitive performance in Spanish SMEs: Empirical evidence from a stakeholders' perspective. **BRQ Business Research Quarterly**, v. 19, 55-72, 2015.

MEEK, G. K.; GRAY, S. J. The value added statement: an innovation for U.S. companies? **Accounting Horizons**: p. 73-81, 1988.

MORLEY, M. F. The Value Added Statement in Britain. **THE Accounting Review**, v. 3, 1979.

NANDI, K. C. Performance measures: an application of value added statement. **The IUP Journal of Operations Management**, v. 10, n. 3, p. 39-62, 2011.

RIAHI-BELKAOUI, A. Value Added reporting and research: state of the art. In: RIAHI-BELKAOUI, A. **The story behind value added reporting**. Westport: Quorum Books: pp. 1-16, 1999.

RODRIGUEZ-FERNANDEZ, Mercedes. Social Responsibility and Financial Performance: The role of good Corporate Governance. **BRQ Business Research Quarterly**, 2015, 19, 137-151.



SANTOS, A. **Demonstração contábil do valor adicionado- DVA**: um instrumento para medição da geração e distribuição de riqueza das empresas. Tese (Livre Docência) - Universidade de São Paulo, São Paulo, SP, 1999.

SANTOS, A. **Demonstração do Valor Adicionado**: como elaborar e analisar a DVA. São Paulo: Atlas, 2007.

SANTOS, M. I. C.; SILVA, M. S. Utilização da demonstração do valor adicionado – DVA como ferramenta de medição da riqueza no setor de telefonia no Brasil. **Pensar Contábil**, v. 11, n. 46, p. 39-45, 2009.

SILVA, C. V. S.; RALHA, C. G. Detecção de cartéis em licitações públicas com agentes de mineração de dados. **Revista Eletrônica de Sistemas de Informação**, v. 10, n. 1, p. 1-19, 2011.

SIQUEIRA, J. R. M. DVA: vantagens e limitações de seu uso como instrumento de avaliação da estratégia social corporativa. **Revista Del Instituto Internacional de Costos**, v. 1, p. 116-140, 2007.

STAINBANK, L. J. The value added statement: does it add any value? **Meditari Accountancy Research**, v. 17, n. 2, p. 137-149, 2009.

SUOJANEN, W. W. Accounting theory and the large corporation. **The Accounting Review**, p. 391-398, 1954.

TINOCO, J. E. P. **Balanco social**: uma abordagem da transparência e da responsabilidade pública das organizações. São Paulo: Atlas, 2001.

VAN STADEN, C. J. **The value added statement**: bastion of social reporting or dinosaur of financial reporting? Palmerston North, NZ: School of Accountancy, Massey University, 2000.

WITTEN, I. H.; FRANK, E. **Data Mining**: practical machine learning tools and techniques. Morgan Kaufmann, 2005.

YOUN, Hyewon; HUA, Nan; LEE, Seoki. Does size matter? Corporate Social Responsibility and firm performance in the restaurant industry. **International Journal of Hospitality Management**, v. 51, p. 127-134, 2015.

YUEN, K. F.; THAI, V. V.; WONG, Y. D. The effect of continuous improvement capacity on the relationship between of corporate social performance and business performance in maritime transport in Singapore. **Transportation Research Part E: Logistics and Transportation Review**, v. 95, n. 1, p. 62-75, 2016.

ZHOU, Z. H. Three perspectives of data mining. **Artificial Intelligence Journal**. v. 143, n. 1, p. 139-146, 2003.