



LUDMILA DE MELO SOUZA

**ANÁLISE DAS VARIÁVEIS QUE EXPLICAM O NÍVEL DE SIMILARIDADE DOS
RELATÓRIOS DA ADMINISTRAÇÃO DE EMPRESAS NEGOCIADAS NO
MERCADO ACIONÁRIO BRASILEIRO**

Orientador: Professor Dr. César Augusto Tibúrcio Silva

BRASÍLIA

2010

Professor Doutor José Geraldo de Sousa Júnior
Reitor da Universidade de Brasília

Professora Doutora Denise Bomtempo Birche de Carvalho
Decana de Pesquisa e Pós-Graduação

Professor Doutor Tomás de Aquino Guimarães
Diretor da Faculdade de Economia, Administração e Contabilidade.

Professor Mestre Elivânio Geraldo de Andrade
Chefe do Departamento de Ciências Contábeis e Atuariais

Professor Doutor César Augusto Tibúrcio Silva
Coordenador-geral do Programa Multiinstitucional e Inter-regional de Pós-Graduação
em Ciências Contábeis da UnB, UFPB e UFRN

LUDMILA DE MELO SOUZA

**ANÁLISE DAS VARIÁVEIS QUE EXPLICAM O NÍVEL DE SIMILARIDADE DOS
RELATÓRIOS DA ADMINISTRAÇÃO DE EMPRESAS NEGOCIADAS NO
MERCADO ACIONÁRIO BRASILEIRO**

Dissertação apresentada como requisito parcial à obtenção do título de mestre em Ciências Contábeis do Programa Multiinstitucional e Inter-regional de Pós-Graduação em Ciências Contábeis da Universidade de Brasília, Universidade Federal da Paraíba e Universidade Federal do Rio Grande do Norte.

Linha de pesquisa: Contabilidade e Mercado Financeiro

Grupo de pesquisa: Informação Contábil e Mercado de Capitais

Orientador: Prof. Dr. César Augusto Tibúrcio Silva.

BRASÍLIA

2010

Souza, Ludmila de Melo.
Análise das variáveis que explicam o Nível de Similaridade dos
Relatórios da Administração de empresas negociadas no mercado
acionário brasileiro/ Ludmila de Melo Souza – Brasília, 2010. 99 f.

Orientador: Prof. Dr. César Augusto Tibúrcio Silva.

Dissertação (mestrado) – Universidade de Brasília, Universidade
Federal da Paraíba e Universidade Federal do Rio Grande do
Norte, Programa Multiinstitucional e Inter-Regional de Pós-
Graduação em Ciências Contábeis, Brasília, 2010.

Inclui bibliografia.

1. Relatórios da Administração. 2. Similaridade. 3. Divulgação de
Informações. I. Silva, César Augusto Tibúrcio. II. Título: Análise das
variáveis que explicam o Nível de Similaridade dos Relatórios da
Administração de empresas negociadas no mercado acionário brasileiro.

LUDMILA DE MELO SOUZA

**ANÁLISE DAS VARIÁVEIS QUE EXPLICAM O NÍVEL DE SIMILARIDADE DOS
RELATÓRIOS DA ADMINISTRAÇÃO DE EMPRESAS NEGOCIADAS NO
MERCADO ACIONÁRIO BRASILEIRO**

Dissertação apresentada como requisito à obtenção do título de mestre em Ciências Contábeis do Programa Multiinstitucional e Inter-regional de Pós-Graduação em Ciências Contábeis da Universidade de Brasília, Universidade Federal da Paraíba e Universidade Federal do Rio Grande do Norte.

Linha de pesquisa: Contabilidade e Mercado Financeiro

Grupo de pesquisa: Informação Contábil e Mercado de Capitais

Professor Doutor César Augusto Tibúrcio Silva

Programa Multiinstitucional e Inter-regional de Pós-Graduação em Ciências Contábeis da UnB/UFPB/ UFRN (Orientador)

Professora Doutora Marisa Bräscher Basílio Medeiros

Faculdade de Ciência da Informação da Universidade de Brasília (membro externo)

Professor Doutor Jorge Katsumi Niyama

Programa Multiinstitucional e Inter-regional de Pós-Graduação em Ciências Contábeis da UnB/UFPB/ UFRN (membro interno)

Brasília, 30 de julho de 2010

À Deus, “inteligência suprema, causa primeira de todas as coisas”

Aos amores da minha vida, Vera e Bruno, base de tudo que eu construí. A vocês meu
amor incondicional!

Ao vovô Divino (*in memoriam*). Na certeza de estar feliz com a minha conquista.

AGRADECIMENTOS

Agradeço a Deus pela minha existência e por realizar coisas tão maravilhosas em minha vida.

À minha família, que sempre será base de tudo.

À vovó Júlia, pelas orações, amor e dedicação à família.

À minha mãe pelo amor incondicional. Você, sem dúvida, é a inspiração para tudo que eu faço na minha vida. Eu amo você!

Ao meu pai, que mesmo a distância, torce pelo meu sucesso.

Ao meu irmão, Bruno, pelo carinho e amizade.

Ao Danniell Muniz, pelo amor, companheirismo e paciência. E por tornar todos os momentos vividos ao seu lado mais que especiais! Amo você!

Agradeço especialmente ao Prof. Dr. César Augusto Tibúrcio Silva, amigo e orientador, por ter acreditado em mim (até naqueles momentos em que eu mesma não acreditava!) e por ter sido meu orientador no Mestrado e em dois Programas de Iniciação Científica. O seu exemplo de Docente foi fundamental para minha escolha pela Vida Acadêmica!

Ao ex-coordenador geral do PMIRPGCC, Prof. Dr. Jorge Katsumi Niyama, pelo exemplo de dedicação e seriedade na condução desse reconhecido Programa.

A todos os funcionários da Faculdade de Economia, Administração e Contabilidade (FACE). Em especial, a Vera Lúcia Cavalcanti (minha mãe na UnB), pelo carinho e atenção com que me tratou nesses anos.

Aos professores do curso que participaram da minha formação, Prof. Dr. Jorge Katsumi Niyama, Prof. Dr. Marcelo Quirino, Prof. Dr. Paulo Roberto Barbosa Lustosa, Prof. Dra. Fátima de Souza Freire, Prof. Dr. José Dionísio G. da Silva, Prof. Dr. César Augusto Tibúrcio Silva, Prof. PhD. Otávio Ribeiro de Medeiros e Prof. Dr. Edilson Paulo. Suas lições ficarão guardadas para sempre em minha memória.

À Prof. Dra. Marisa Bräscher Basílio Medeiros e ao Prof. Dr. Jorge Katsumi Niyama pela participação na banca de avaliação dessa dissertação e pelas valiosas contribuições.

Aos amigos da Secretaria da Pós-graduação: Renato, Rodolfo e Aline, em especial pelo carinho e atenção constantemente demonstrados.

Aos professores do Departamento de Ciências Contábeis e Atuariais (CCA) por todos os ensinamentos e conhecimentos transmitidos ainda na graduação. Em especial, agradeço ao Prof. Mestre Cláudio Moreira Santana pela disposição e atenção.

Aos colegas e amigos da 17^o Turma do Mestrado: Abimael, Antônio Daniel, Bruno, Domingos, Ednei, Eric, Eurípedes, Evandro, Juliana e Meg, com os quais tive a honra de aprimorar e compartilhar conhecimentos e estabelecer relações de amizade.

Aos amigos: Brunna Hísla, Camila Machado, Danielle Nunes, Denise Rochael, Diones Rocha, Humberto Cruzeiro e José Lúcio Tozetti, por me “adotarem” como mais uma integrante da 15^o Turma do Mestrado.

Ao amigo José Lúcio Tozetti pela troca de ensinamentos e amizade sincera. A sua companhia contribuiu para que os momentos vividos na UnB fossem ainda mais agradáveis.

À amiga Meg pelas várias manifestações de apoio e amizade! E ao Sérgio, ao Pedro e ao “pequeno” André, que a “dividiram” comigo e me receberam em casa, com muita simpatia.

À Isabel Sales, cuja energia e amizade foram imprescindíveis nos momentos finais dessa pesquisa – as suas mensagens me desejando ânimo foram essenciais para a conclusão desse trabalho.

À amiga Fernanda, com quem tive a oportunidade de engrandecer o meu conhecimento e criar laços de amizade.

À querida Nídia, pelo carinho, profissionalismo e dedicação com que me acompanhou.

Aos grandes amigos Carol, Mara, Kaká, Quel, Enio, Fiel, Paulinho e Rapha.

Aos participantes do Curso de Extensão Pesquisa em Contabilidade VI, pelas contribuições.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela concessão da bolsa de estudos fundamental para a realização da pesquisa e conclusão do mestrado.

A todos aqueles que indesculpavelmente não foram citados, porém de suma importância para conclusão dessa pesquisa.

MUITO OBRIGADA!

“...e a terceira forma de felicidade é viver a vida como a de um pesquisador”

(Aristóteles)

RESUMO

Os Relatórios da Administração (RA) são utilizados como uma ferramenta de comunicação das empresas para transmitir uma mensagem ou imagem para seus acionistas. Considerando a importância de que esses relatórios apresentem, cada vez mais, informações compreensíveis, relevantes, confiáveis e comparáveis - atributos da informação contábil - este estudo propôs uma investigação quanto a sua estrutura, a fim de verificar se há repetição nas informações que a administração divulga ao mercado de um exercício social para outro; tendo em vista que não foram encontrados estudos semelhantes no Brasil e na área de Contabilidade. O objetivo do estudo foi identificar quais as variáveis explicam a similaridade entre os RA divulgados em anos subsequentes pelas empresas negociadas no mercado acionário brasileiro. Utilizando-se de um *software* de análise de conteúdo, foram analisados os relatórios de 311 empresas listadas na bolsa brasileira, publicados juntamente com as Demonstrações Contábeis encerradas em 31/12/2008 e 31/12/2009. Em seguida, foi mensurada a similaridade entre esses RA. Dessa maneira, construiu-se um indicador que representasse o nível de similaridade dos relatórios, bem como verificou quais variáveis relacionadas à complexidade linguística e organizacional, desempenho, risco, tamanho e custos políticos das empresas ajudariam explicar esse indicador. Os resultados encontrados demonstraram que mais de 70% das empresas brasileiras constantes da amostra não apresentaram alterações significativas na estrutura do RA de 2009, em relação ao de 2008. Os resultados obtidos permitem concluir que grande parte das empresas analisadas realizou apenas modificações pontuais como: data, nomes e valores. Adicionalmente, verificou-se que o tamanho da empresa influencia positivamente nas alterações das informações contidas nos relatórios. Assim, quanto maior o porte da empresa, menor a similaridade dos seus RA em anos subsequentes. De maneira geral, o fato de outras variáveis independentes não terem se mostrado explicativas nos modelos obtidos, pode sugerir que a manutenção de estruturas similares de relatórios é uma prática comum nas empresas brasileiras estudadas, independentemente das suas características. Além disso, as empresas podem estar, também, perdendo uma grande oportunidade de divulgar suas realizações ao mercado.

Palavras Chaves: 1. Relatórios da Administração. 2. Similaridade. 3. Divulgação de Informações

ABSTRACT

Management Reports (MR) are used as companies' communication tool for broadcasting a message or image to its shareholders. In regard to the importance that these reports increasingly show comprehensible, relevant and comparable data – which are attributes of accounting information – this study presents an investigation of the structure of Management Reports with the purpose of verifying if there is a repetition of data, when comparing different fiscal years; given that no similar studies were found in Brazil and in the accounting field. The aim of the study was to identify which variables explain the similarity between the MR that were published by companies listed in the Brazilian stock market in subsequent years. With the help of a statistics software, there were analyzed the annual reports of 311 (three hundred and eleven) companies listed at the stock market, which were published between 12/31/2008 and 12/31/2009. Afterwards, it was measured the similarities between these MR. Thus, an indicator that represents the level of similarity between the reports was built, and it was, as well, verified if the indicator could be explained by variables related to linguistic complexity and organizational performance, risk, size and political costs of the companies. The results signalized that over 70% of the sample of Brazilian firms showed no significant changes in the structure of their MR in 2009, when compared to 2008. Therefore it was concluded that most of these companies had only specific modifications at their reports, such as: date, names and values. Additionally, it was found that firm size positively influences the change of information contained in the reports. Thus, the larger the company size, the smaller the similarity of their MR in subsequent years. In general, the fact that other independent variables were not shown as explanatory at the obtained model, may suggest that the maintenance of similar reports' structures is a usual practice at the studied Brazilian companies, regardless of their characteristics. In addition, companies may be also missing a great opportunity to report their achievements to the market.

Key-words: Management Reports. Similarity. Disclosure.

LISTA DE ABREVIATURAS E SIGLAS

CPC	Comitê de Pronunciamentos Contábeis
CVM	Comissão de Valores Mobiliários
RA	Relatório(s) da Administração
PL	Patrimônio Líquido

LISTA DE QUADROS

QUADRO 1.1 – EXEMPLOS DE RA DIVULGADOS EM 2007 E 2008	18
QUADRO 3.1 – HIPÓTESES DO TESTE DE MÉDIA	39
QUADRO 3.2 – VARIÁVEIS INDEPENDENTES	42
QUADRO 4.1 – RESUMOS DOS RESULTADOS E VERIFICAÇÃO DAS HIPÓTESES DA PESQUISA	59

LISTA DE ILUSTRAÇÕES

ILUSTRAÇÃO 3.1 – EXEMPLIFICAÇÃO DA CRIAÇÃO DOS INDICADORES	40
---	-----------

LISTA DE TABELAS

TABELA 4.1 – ESTATÍSTICAS DESCRITIVAS (GRAUS 100 E 60)	48
TABELA 4.2 – VARIÁVEIS QUE EXPLICAM O INDICADOR DE SIMILARIDADE 100, TOTAL	50
TABELA 4.3 – VARIÁVEIS QUE EXPLICAM O INDICADOR DE SIMILARIDADE 60, TOTAL	53
TABELA 4.4 – VARIÁVEIS QUE EXPLICAM O INDICADOR DE SIMILARIDADE 100, TERCIL (1º)	54
TABELA 4.5 – VARIÁVEIS QUE EXPLICAM O INDICADOR DE SIMILARIDADE 100, TERCIL (3º)	56
TABELA 4.6 – VARIÁVEIS QUE EXPLICAM O INDICADOR DE SIMILARIDADE 60, TERCIL (3º)	57

SUMÁRIO

1 INTRODUÇÃO.....	15
1.1 PROBLEMA E RELEVÂNCIA	17
1.2 OBJETIVOS DA PESQUISA.....	18
1.3 HIPÓTESES DA PESQUISA.....	19
1.4 DELIMITAÇÃO DA PESQUISA	20
1.5 A PESQUISA.....	20
1.6 ESTRUTURA DO TRABALHO.....	21
2 REFERENCIAL TEÓRICO.....	22
2.1 RELATÓRIOS DA ADMINISTRAÇÃO.....	22
2.1.1 PESQUISAS EMPÍRICAS NO ÂMBITO INTERNACIONAL	24
2.1.2 PESQUISAS EMPÍRICAS NO BRASIL	27
2.2 OFUSCAÇÕES LINGÜÍSTICAS	29
2.3 SIMILARIDADE	32
3 PROCEDIMENTOS DE PESQUISA	36
3.1 COLETA DE DADOS	36
3.2 POPULAÇÃO E AMOSTRA.....	36
3.3 OBTENÇÃO DAS ESTATÍSTICAS DE SIMILARIDADE.....	37
3.4 CRITÉRIOS UTILIZADOS PARA OBTENÇÃO DOS INDICADORES DE SIMILARIDADE.....	39
3.5 INSTRUMENTOS ESTATÍSTICOS PARA A APURAÇÃO DOS RESULTADOS.....	41
3.6 RESULTADOS ESPERADOS	46
4 RESULTADOS	47
4.1 ESTATÍSTICAS DESCRITIVAS	47
4.2 RELAÇÕES ENTRE AS VARIÁVEIS INDEPENDENTES E INDICADORES DE SIMILARIDADE, TOTAL.....	49
4.2.1 GRAU DE SIMILARIDADE 100	49
4.2.2 GRAU DE SIMILARIDADE 60	52

4.3 RELAÇÕES ENTRE AS VARIÁVEIS INDEPENDENTES E INDICADORES DE SIMILARIDADE, TERCIL	53
4.3.1 PRIMEIRO TERCIL – GRAU DE SIMILARIDADE 100	53
4.3.2 TERCEIRO TERCIL – GRAU DE SIMILARIDADE 100	56
4.3.3 PRIMEIRO TERCIL – GRAU DE SIMILARIDADE 60	57
4.3.4 TERCEIRO TERCIL – GRAU DE SIMILARIDADE 60	57
4.4 ANÁLISE DAS PREMISSAS DO MODELO CLÁSSICO DE REGRESSÃO LINEAR.....	58
4.5 SÍNTESE DAS HIPÓTESES DA PESQUISA E RESULTADOS ENCONTRADOS	59
5 CONSIDERAÇÕES FINAIS	62
REFERÊNCIAS.....	66
LISTA DE APÊNDICES	70
LISTA DE ANEXOS	91

1 INTRODUÇÃO

A informação contábil está relacionada ao objetivo da contabilidade de mensurar o conjunto de eventos econômicos e comunicar os resultados encontrados às partes interessadas (YAMAMOTO; SALOTTI, 2006). Nesse contexto, a Estrutura Conceitual Básica da Contabilidade constante do CPC (2008, p. 10) estabelece que as Demonstrações Contábeis devam apresentar atributos qualitativos básicos para que sejam úteis aos usuários, quais sejam: compreensibilidade, relevância, confiabilidade e comparabilidade.

No entanto, nem todas as informações contábeis são proporcionadas por demonstrações contábeis tradicionais; existem outras formas alternativas de divulgação dessas informações, como por exemplo, por meio dos Relatórios da Administração (RA). Segundo Yamamoto e Salotti (2006, p. 7), “a aceitação de outras fontes relevantes de informações contábeis é uma *ampliação da idéia* de que todas as informações importantes devem estar refletidas no corpo das demonstrações contábeis tradicionais” (grifo nosso).

Os RA têm sido utilizados como uma ferramenta de comunicação das empresas para transmitir uma mensagem ou imagem para seus acionistas (HERREMA; RYANS, 1995). Nesse sentido, esses relatórios necessitam de se expandir para atender as mudanças nas necessidades de informações requeridas pelo mercado e fornecer dados que reforçam a transparência e a *accountability*. Assim, em todo mundo, a comunicação narrativa em RA é vista como elemento crucial para alcançar a mudança desejada na qualidade dos relatórios das empresas (BEATTIE; MCINNES; FEARNLEY, 2004).

Llewellyn (1998) defende que a narração se apresenta como um dos modos de explicar o pensamento e de convencer; e é tão legítimo quanto aqueles respaldados por raciocínios quantitativos. Para Bryan (1997), os relatórios narrativos fornecem informações

aos investidores que não podem ser facilmente incluídas nas Demonstrações Contábeis propriamente ditas, como por exemplo, dados sobre as operações futuras da empresa. Apesar dessa suposta legitimidade dos textos narrativos, estes estão sendo subutilizados pelas Ciências Sociais (LLEWELLYN, 1998). Isso se deve principalmente a idéia difundida pelos teóricos de que, apesar da utilidade e relevância para os usuários da contabilidade, esses relatórios contêm viés. Dye (2001) afirma que qualquer entidade que tenha a possibilidade de realizar uma divulgação apresentará informações que lhe sejam favoráveis e omitirá as que lhe são desfavoráveis.

Ressalta-se, nesse ponto, que não há obrigatoriedade de realização de auditoria independente nas informações contidas nos RA, contudo, eventualmente eles podem ser revisados (SPRINGER, 1992). Iudícibus, Martins e Gelbcke (2007) comentam que muitos desses relatórios são utilizados como uma forma de ostentar adjetivos de autopromoção aos próprios administradores e que, em alguns extremos, contêm interpretações favoráveis sob a situação patrimonial e financeira de determinada entidade **não condizentes** com a realidade apresentada nas Demonstrações Contábeis. Portanto, “por ser este relatório uma evidenciação narrativa não-auditada, tornam-se necessários estudos que confirmem a transparência das informações prestadas pela administração” (SILVA; RODRIGUES; ABREU, 2007, p.75).

Rutherford (2003) aponta dois aspectos problemáticos em relação à regulação e normatização das narrativas contábeis, quais sejam:

a) *A dificuldade de estabelecer normas com determinada precisão.* Se forem criadas normas prescritivas incompletas, poderão existir brechas e as empresas poderão, assim, omitir informações materialmente relevantes; caso contrário, se forem criadas normas muito específicas, poderão surgir relatórios prolixos com informações desnecessárias; e

b) *A dificuldade de fiscalização por parte dos reguladores.* Em alguns casos, muitas informações divulgadas nas narrativas não podem ter sua veracidade comprovada sem o acesso tempestivo às circunstâncias ocorridas, dificultando o trabalho da autoridade competente e conseqüentemente aumentando os custos de inspeção.

1.1 Problema e Relevância

Considerando a necessidade de que os RA apresentem informações compreensíveis, relevantes, confiáveis e comparáveis (atributos qualitativos da informação contábil) este estudo propõe uma investigação quanto à apresentação e estrutura desses documentos, a fim de verificar se há repetição nas informações que a administração divulga de um exercício social para outro ademais, não foram encontrados estudos semelhantes no Brasil e na área de Contabilidade.

Acredita-se que as empresas possuem uma tendência a reproduzir determinada estrutura de relatório ao longo dos anos. Essa prática pode trazer alguns prejuízos, como por exemplo, reduzir a capacidade do investidor de tomar decisões com base nas narrativas contábeis e/ou induzi-lo ao erro. Adicionalmente, essa prática representa uma oportunidade perdida, pela empresa, de divulgar suas realizações, além do fato de sugerir certo descuido por parte da administração em relação à divulgação de informações ao mercado.

O Quadro 1.1 apresenta trechos dos RA da empresa Sergen S.A divulgados em 2007 e 2008. Note-se que a Companhia não só manteve a mesma estrutura de relatório como também o mesmo texto; a única alteração observada foi a data do exercício. Vale ressaltar que até o mesmo erro de português (“monetária”) permaneceu.

Relatório da Administração 2007	Relatório da Administração 2008
<p>Senhores Acionistas A Direção da SERGEN - Serviços Gerais de Engenharia S/A de acordo com a legislação em vigor, submete à apreciação de V.Sas. as Demonstrações Financeiras relativas ao Exercício Social findo em 31/12/2007</p> <p>DESEMPENHO OPERACIONAL Considerando a prática dos atrasos de pagamento dos Governos, a Diretoria da SERGEN Serviços Gerais de Engenharia S/A decidiu reduzir as suas atividades para iniciar novas obras, dirigindo seus esforços no sentido de agilizar os diversos processos de cobranças judiciais com correção monetária e juros dos serviços realizados e que não foram pagos de acordo com os respectivos contratos.”</p>	<p>Senhores Acionistas A Direção da SERGEN - Serviços Gerais de Engenharia S/A de acordo com a legislação em vigor, submete à apreciação de V.Sas. as Demonstrações Financeiras relativas ao Exercício Social findo em 31/12/2008</p> <p>DESEMPENHO OPERACIONAL Considerando a prática dos atrasos de pagamento dos Governos, a Diretoria da SERGEN Serviços Gerais de Engenharia S/A decidiu reduzir as suas atividades para iniciar novas obras, dirigindo seus esforços no sentido de agilizar os diversos processos de cobranças judiciais com correção monetária e juros dos serviços realizados e que não foram pagos de acordo com os respectivos contratos.”</p>

Fonte: Relatório da Administração da Sergen S.A.

Quadro 1.1 - Exemplo de RA divulgado em 2007 e 2008

Com base no que foi mencionado anteriormente, tem-se o seguinte problema de pesquisa: **Qual a relação entre as características das empresas negociadas no mercado acionário brasileiro e nível de similaridade entre os Relatórios da Administração divulgados em períodos subsequentes?**

1.2 Objetivos da Pesquisa

Considerando o problema de pesquisa descrito, o objetivo do estudo é identificar qual a relação entre as características das empresas negociadas no mercado acionário brasileiro e o nível de similaridade entre os RA divulgados em períodos subsequentes.

Para tanto, foram estabelecidos os seguintes Objetivos Específicos:

- a) Mensurar o nível de similaridade entre os RA divulgados em 2008 e 2009 das empresas negociadas no mercado acionário brasileiro;

- b) Elaborar um indicador que represente o nível de similaridade entre os RA divulgados no período analisado; e
- c) Identificar quais variáveis explicam esse nível de similaridade.

1.3 Hipóteses da Pesquisa

Segundo Martins e Theóphilo (2009, p. 31), um dos conceitos de hipótese de pesquisa é “proposição que pode ser colocada à prova – ser testada – para determinar a sua validade” ou “proposição de suposta explicação do fenômeno que está sendo investigado”. Dessa maneira, as hipóteses desse estudo são:

- I. As empresas brasileiras tendem a reproduzir as estruturas de seus RA ao longo do tempo;
- II. Não há modificações significativas no que tange à quantidade de informações divulgadas nos RA das empresas brasileiras, para o período estudado;
- III. Em função do seu tamanho, as empresas tendem a apresentar RA **mais ou menos** similares em períodos subsequentes;
- IV. Empresas que possuem maiores custos políticos¹ tendem a apresentar RA **menos** similares;
- V. Empresas que apresentam melhores indicadores de desempenho tendem a apresentar RA **menos** similares;
- VI. Empresas que apresentam estruturas organizacionais mais complexas tendem a apresentar RA **menos** similares;

¹ Custos decorrentes de contratos e *lobbies*, provenientes da regulação do Estado para as atividades da empresa (WATTS; ZIMMERMAM, 1976). Para Ribeiro et al (2009) a Hipótese de Custos Políticos presume que grandes empresas inseridas em setores sensíveis a questões políticas, tais como, setor farmacêutico, florestal, de utilidade públicas e de petróleo e gás estão sujeitos à custos potencialmente elevados, relacionado à evidenciação de seus ganhos também elevados.

- VII. Empresas com maiores indicadores de concentração acionária tendem a apresentar RA **mais** similares; e
- VIII. Empresas que representam maiores riscos de investimento tendem a divulgar RA **mais** similares.

1.4 Delimitação da Pesquisa

Os objetos de estudo são os RA divulgados em 2008 e 2009 juntamente com as Demonstrações Contábeis Societárias Anuais das empresas listadas na BMF&BOVESPA. Portanto, os resultados dessa pesquisa ficam restritos às empresas presentes na amostra.

1.5 A Pesquisa

Para Le Coadic (2004, p. 17), “toda ciência é uma atividade social determinada por condições históricas e socioeconômicas.” Nesse sentido, Marconi e Lakatos (2009, p. 4) afirmam que “a pesquisa social é um processo que utiliza metodologia científica, por meio da qual se pode obter novos conhecimentos no campo da realidade social”.

Portanto, a abordagem empregada nesse estudo é a hipotético-dedutiva, pois nele se utiliza como base a formulação de hipóteses para tentar explicar a questão evidenciada no problema de pesquisa (WALLIMAN, 2005).

Adicionalmente, trata-se de uma pesquisa interdisciplinar, pois estuda um fenômeno investigado em outros campos da ciência além da Contabilidade, tais como a Ciência da Informação, Ciência da Computação e Linguística.

1.6 Estrutura do Trabalho

Esta pesquisa está estruturada em cinco capítulos que abrangem os seguintes conteúdos:

Capítulo um: descreve a justificativa e o problema de pesquisa, os objetivos, as delimitações do estudo, os aspectos metodológicos e a estrutura do trabalho;

Capítulo dois: trata da revisão de literatura, que contempla: (i) a obrigatoriedade da publicação do RA e as pesquisas empíricas relacionadas ao tema, tanto no âmbito internacional como no nacional; (ii) a ofuscação linguística; e (iii) a similaridade.

Capítulo três: descreve os procedimentos da pesquisa, a coleta dos dados; a população e os critérios de seleção da amostra para realização do estudo; a operacionalização das variáveis, a identificação dos indicadores de similaridade e os instrumentos estatísticos utilizados para a apuração dos resultados;

Capítulo quatro: apresenta a análise do resultado encontrado na pesquisa; e

Capítulo cinco: apresenta as considerações finais, as limitações da pesquisa e as sugestões estudos futuros.

2 REFERENCIAL TEÓRICO

Nesta seção é apresentado um conjunto de teorias e estudos os quais norteiam o desenvolvimento dessa pesquisa. Primeiramente, serão abordados os requisitos legais para a elaboração e divulgação dos RA, suas características e os estudos empíricos já realizados. Em seguida, será tratado sobre a ofuscação linguística e por último sobre a similaridade e as algumas pesquisas encontradas sobre o assunto.

2.1 Relatórios da Administração

A Lei 6.404/76 e suas modificações estabelecem que as sociedades por ações de capital aberto devam divulgar o RA, as Demonstrações Contábeis, Notas Explicativas e o Parecer dos Auditores Independentes. Esses documentos representam o conjunto de informações que os gestores fornecem aos diversos usuários da contabilidade. Especificamente no artigo 133 da referida Lei, o legislador estabelece que o objetivo do RA é evidenciar os negócios sociais bem como, os principais fatos administrativos ocorridos no exercício findo.

A Comissão de Valores Mobiliários (CVM), no Parecer de Orientação 15/87 reproduz os temas cuja divulgação é exigida pela Lei societária em relatório anual: a) aquisição de debêntures de sua própria emissão (art. 55, § 2º); b) política de reinvestimento de lucros e distribuição de dividendos constantes de acordo de acionistas (art. 118, § 5º); c) negócios sociais e principais fatos administrativos ocorridos no exercício (art. 133, inciso I); e d) relação dos investimentos em sociedades coligadas e/ou controladas evidenciando as modificações ocorridas durante o exercício (art. 243).

No entanto, além de reproduzir as obrigatoriedades já estabelecidas pela Lei, a CVM também recomenda outras informações que podem constar nesses relatórios: a) descrição dos negócios, produtos e serviços; b) comentários sobre a conjuntura econômica geral; c) recursos humanos; d) investimentos; e) pesquisa e desenvolvimento; f) novos produtos e serviços; g) proteção ao meio-ambiente; h) reformulações administrativas; i) investimentos em controladas e coligadas: indicação dos investimentos efetuados e objetivos pretendidos com as inversões; j) direitos dos acionistas e dados de mercado; e k) perspectivas e planos para o exercício em curso e os futuros (CVM, 1987).

É importante notar que alguns dos itens recomendados pela CVM induzem a similaridade (descrição de negócios, produtos e serviços, por exemplo), enquanto que outros motivam mudanças nas informações divulgadas ao longo dos anos (como comentários sobre a conjuntura econômica geral). De certa forma, é possível afirmar que as recomendações não permitem que um RA seja completamente diferente ao longo do tempo.

O Comitê de Pronunciamentos Contábeis (CPC) também se pronuncia a respeito das informações que devem estar contidas nos RA. Segundo o CPC, esses relatórios devem trazer informações que descrevam e expliquem as características principais de desempenho e da posição financeira e patrimonial da entidade e as principais incertezas às quais ela está sujeita. E sugere, ainda, que se incluam análise contendo os principais fatores que determinam o desempenho, as fontes de financiamento da entidade e a respectiva relação pretendida entre passivo e patrimônio líquido e os recursos da entidade não reconhecidos nas Demonstrações Contábeis de acordo com os outros pronunciamentos emitidos pelo referido Comitê. (CPC 26, 2009).

Portanto, tanto a CVM quanto o CPC destacam a importância das empresas produzirem RA visando fornecer informações completas, tempestivas, fidedignas e

transparentes e que demonstrem a real situação da empresa, haja vista a dinâmica e a evolução do mercado.

A literatura existente sobre relatórios anuais adota uma variedade de abordagens para a análise das narrativas contábeis. Para Beattie, McInnes, Fearnley (2004), o estudo dos RA se divide em duas grandes metodologias: a subjetiva e a semi-objetiva. Um exemplo de metodologia subjetiva seria a “Avaliação dos Analistas”. Dentre as metodologias classificadas como semi-objetivas, os autores fazem uma divisão entre aquelas que se utilizam de índice de evidenciação (utilizando variáveis binárias, pesos, categorias hierárquicas) e aquelas que se pautam pela análise textual (análise de conteúdo, legibilidade, análise linguística).

2.1.1 Pesquisas Empíricas no Âmbito Internacional

Uma série de estudos sobre os RA foram desenvolvidos no âmbito internacional. Martson e Shrivies (1991) realizaram uma revisão sobre a utilização dos índices de evidenciação na pesquisa contábil, particularmente nos relatórios anuais narrativos. Segundo os autores, o índice calculado para uma determinada empresa pode representar uma medida do grau de divulgação, mas isso não necessariamente significa qualidade da divulgação. Isso ocorre devido à subjetividade da ponderação desses índices por partes dos vários pesquisadores. Nisso, os autores concluíram que embora a construção de índices inevitavelmente envolva subjetividade, essa metodologia ainda permanece como um instrumento valioso de pesquisa quando o tema é a divulgação de informações por parte das empresas.

Jones e Shoemaker (1994) revisaram trinta e dois estudos de legibilidade, vinte seis deles relacionados à utilização da Legibilidade *Flesch* na mensuração da facilidade de leitura dos RA. Os pesquisadores verificaram que existe consenso geral de que os RA são, na maioria

das vezes, classificados como “difíceis” ou “muito difíceis” de ler. No entanto, os referidos autores reforçam a idéia de que apesar da sua objetividade e confiabilidade, existem muitos problemas associados à aplicação da Metodologia *Flesch* em textos narrativos da Contabilidade. Primeiro porque a técnica foi desenvolvida para a aplicação em textos infantis e, segundo porque ela leva em consideração as palavras e as frases, e não o texto como um todo e, por último, por não considerar os interesses dos leitores.

Abrahamson e Amir (1996) realizaram um estudo com o objetivo de verificar se as informações contidas na seção mensagem aos acionistas dos RA estavam dentre as variáveis analisadas pelos investidores quando da análise dos números contábeis. Utilizando como ferramenta um *software* de análise de conteúdo, os autores identificaram que as informações contidas na seção considerada são o ponto de partida para uma análise da qualidade do lucro e a investigação sobre gerenciamento de resultados.

Courtis (1998) realizou um estudo com a intenção de investigar a presença de variabilidade na seção carta do presidente contida no RA. O autor calculou o Coeficiente de Variação (CV) considerando três estatísticas *Flesch* de 120 RA de empresas de capital aberto de Hong Kong. O pesquisador verificou que as empresas cujos RA apresentaram baixa legibilidade (RA de “difícil” ou “muito difícil” leitura) divulgam más notícias. Além disso, informações sobre as referidas empresas são recorrentes na imprensa.

Sydserrff and Weetman (1999) desenvolveram um índice textual apresentando um novo método de avaliação dos RA. Os autores justificaram que esse índice captura um conjunto mais rico de características do texto que não podem ser abrangidas pelo método de legibilidade. Verificaram, também, que o modelo geral estabelecido pela linguística pode ser adaptado à situação específica de uma narrativa em Contabilidade. Isso sugere que o índice textual é uma potencial ferramenta para análise de narrativas contábeis e para a realização testes de associação.

Boo e Simnett (2002) realizaram uma pesquisa com o objetivo de verificar se as informações contidas nos RA das companhias que estão enfrentando dificuldade financeira são esclarecedoras quanto à viabilidade futura dessas entidades. Foram identificados RA de 140 empresas australianas que apresentaram perdas significativas. Os relatórios foram classificados como otimistas, pessimistas, mistos e neutros. Os autores constataram que as empresas que estão em dificuldades financeiras e não evidenciam essa situação em seus relatórios, estão mais propensas a falir do que àquelas que estão em dificuldades e apresentam esse fato nos RA, na maioria das vezes de forma otimista.

Yuthas, Rogers e Dillard (2002) analisaram RA utilizando-se das normas de *Habermas*. Estas normas que pressupõem que a comunicação deve ser compreensível, confiável, sincera e legítima. Para isso, os autores selecionaram uma amostra de empresas que obtiveram lucro ou prejuízos não previstos no primeiro trimestre de 2001. Nesse estudo, foi verificado que àquelas empresas que utilizam os RA para veicularem informações confiáveis e fidedignas, evidenciando a sua verdadeira posição financeira, antecipam ganhos ou perdas.

A pesquisa de Beattie, Mcinnes, Fearnley (2004) contribuiu com o estudo dos RA em duas perspectivas: (i) a primeira contribuição é a introdução na literatura acadêmica de uma estrutura de quatro dimensões para a análise de conteúdo das narrativas contábeis e a utilização de uma ferramenta computacional para o desenvolvimento dessas dimensões; e (ii) a segunda relaciona-se à exploração do conceito de qualidade e sua difícil tarefa de mensuração. Dessa maneira, a pesquisa de Beattie, Mcinnes e Fearnley (2004) foi uma tentativa preliminar de identificar alguns atributos de qualidade da evidenciação contábil, sugerindo algumas *proxies* para mensuração dessa qualidade: tamanho do relatório, complexidade do texto e propagação da divulgação.

2.1.2 Pesquisas Empíricas no Brasil

No Brasil, um dos primeiros estudos em Contabilidade abordando os RA foi o de Rodrigues (2005). O referido estudo analisou os RA relativos aos anos de 2001 a 2003 e tinha como objetivo verificar a relação entre o conteúdo existente nos relatórios e o desempenho financeiro das companhias. A autora concluiu que os gestores apresentam o cenário econômico de forma pessimista quando obtém desempenho inferior ao esperado, atribuindo culpa ao ambiente econômico. Constatou também que a administração tende a ser otimista ao apresentar suas reformas, e nesse caso atribui a si o mérito pelo resultado alcançado.

Gallon e Beuren (2006) analisaram a relação entre o grau de evidenciação nos RA e o Nível de Governança Corporativa em que as empresas se encontram. A análise das informações de cinquenta oito companhias abertas teve como finalidade identificar a presença de itens obrigatórios de evidenciação ou recomendados, e os itens não obrigatórios e nem recomendados. O estudo conclui que há relação entre o grau de evidenciação de informações narrativas e o nível de governança das empresas. Portanto, confirmou-se o pressuposto de que as empresas apresentam melhoria na prestação de informação ao mercado em função da sua participação em um dos níveis diferenciados de governança.

Silva, Rodrigues e Abreu (2007) utilizaram para análise os RA divulgados no ano de 2002 e verificaram a relação entre os referidos relatórios e o resultado financeiro das companhias. As principais conclusões revelaram que os relatórios considerados como otimistas apresentaram maior volume de frases sobre reforma administrativa; já os relatórios classificados como pessimistas centraram sua atenção na conjuntura econômica. Constatou-se que empresas que apresentaram crescimento no ativo possuem relatórios mais extensos. Por outro lado, identificou-se que o aumento no lucro provoca uma redução no tamanho do relatório.

Beuren, Boff e Hein (2008) averiguaram se as empresas familiares de capital aberto evidenciaram nos RA, no período de 1997 a 2006, as informações relacionadas aos recursos humanos e questões relativas à proteção ao meio ambiente, em atendimento ao recomendado no Parecer de Orientação nº 15/87 da CVM. Os autores concluíram que todas as empresas pesquisadas alteraram a evidenciação das informações nos relatórios relativos ao período analisado, com destaque para o período de 1999 a 2001, no qual mais alterações foram realizadas. No entanto, verificou-se que nem todas as empresas estudadas adaptaram o RA com vistas a evidenciar as informações inerentes aos recursos humanos e à proteção do meio ambiente conforme recomendado no mencionado Parecer.

Cunha e Silva (2009) analisaram as Notas Explicativas e os RA das empresas brasileiras listadas na Bovespa. O objetivo era evidenciar o gerenciamento de impressões utilizando-se da Fórmula *Flesch* para a verificação do grau de facilidade de leitura. Os resultados encontrados sugerem que há fracas evidências de gerenciamentos de impressões nas narrativas estudadas, porém existe relação entre o porte da empresa, o crescimento contínuo e a melhoria na facilidade de leitura de seus textos.

Scotá (2009) avaliou a viabilidade da previsão do risco de descontinuidade das empresas brasileiras por meio de um modelo que utiliza apenas as informações das seções narrativas dos RA. Foram selecionados relatórios do exercício anterior à sua concordata, ou falência relativa das empresas. Utilizando-se dos modelos de Smith e Taffler, o autor concluiu que os modelos são consistentes e previram a falência das empresas em mais de 90% dos casos.

Tessarollo e Pagliarussi (2009) analisaram os informativos contábeis narrativos sob a ótica da Teoria da Sinalização, baseando-se no *Self-Serving Attributions*, um tipo de raciocínio causal que permite ao escritor assumir o crédito para as boas notícias e evitar a culpa pelas más. Foram analisados os RA das 49 empresas. A seleção dos relatórios levou em

consideração os anos considerados “bons” e “ruins” para o mercado de capitais brasileiro. Os resultados demonstraram que o raciocínio baseado no *Self-serving* é uma justificativa para o desempenho organizacional. Assim, os autores concluíram que as companhias que apresentaram desempenho positivo foram mais prolixas. Dessa maneira, os autores contataram que (i) as empresas que apresentaram um desempenho inferior, se sentem obrigadas a dar mais explicações quando se trata de um ano bom para o mercado; e (ii) que não importa quão ruim seja o desempenho, o tom dos relatórios é sempre otimista. Vale ressaltar que os resultados obtidos nessa pesquisa corroboram com os encontrados por Silva, Fernandes e Abreu (2007).

Luz e Pagliarussi (2009) investigaram o possível relacionamento entre atribuições causais (*self-serving*) nas seções narrativas dos RA e o Índice de Governança Corporativa das empresas listadas na Bovespa. Os resultados mostraram que essas atribuições causais foram encontradas nos relatórios anuais independentemente da participação das empresas nos Níveis de Governança Corporativa.

2.2 Ofuscações Linguísticas²

Halliday (1976, 1985) propôs uma teoria que tenta explicar a organização interna da Língua. Para esse autor, a língua oferece três funções: a (i) *ideacional*, que trata da expressão do conteúdo, tanto no âmbito exterior (sociedade) quanto no interior (mente) e se utiliza de métodos como a transitividade verbal para ser medida; (ii) a *interpessoal* que é empregada para a interação do usuário na sociedade e utiliza-se dos modos, pessoas e tempos verbais para detectar intenções; e a (iii) *textual* que se preocupa com a organização do texto, sua coesão e sistema temático.

² Utilização de técnicas da linguística com o intuito de ofuscar ou desviar a atenção de determinados tipos de informações nos relatórios contábeis narrativos.

O estudo das Narrativas em Linguística foi desenvolvido inicialmente por Beaugrande e Dressler em 1981. Na teoria de comunicação narrativa eles definiram um conjunto de sete princípios que determinam a efetividade de uma comunicação, quais sejam: coesão, coerência, aceitabilidade, intencionalidade, informatividade, situacionalidade, intertextualidade (BEAUGRANDE; DRESSLER, 1981).

Reeves et al (2005) afirmam que os diferentes documentos narrativos publicados pelas empresas estão relacionados às diferentes intenções que cada uma possui e que essas intenções estão correlacionadas com às escolhas linguísticas dos agentes.

Segundo Rutherford (2003), as narrativas contábeis se apresentam como um desafio ainda maior para os reguladores do que os próprios números contábeis. Isso se deve tanto à dificuldade dos fiscalizadores do mercado em detectar a ofuscação e a manipulação linguística nesses tipos de documento, quanto ao obstáculo de se exigir que os relatórios sejam auditados por peritos independentes, tais como professores universitários da área de Letras, nos casos de falta de clareza das informações.

Assim, a imprecisão ou até mesmo lacuna nas normas relativas à clareza das informações contidas nos RA, bem como a impossibilidade de acesso às diversas circunstâncias descritas nesses documentos, têm conduzido os pesquisadores a desenvolverem técnicas de análise dos textos narrativos (RUTHERFORD, 2003).

Amernic e Crag (2004) chamam a atenção para a importância e o poder das palavras. Os autores destacam que todas as pessoas pensam, falam e escrevem por meio de palavras, dessa maneira, para os diretores das empresas isso não é diferente. Por meio de uma análise retórica, é possível então compreender a ideologia dos gestores (ARMENIC; CRAG, 2004).

Zimmerman (1998) assevera que os agentes têm preferências por uma ampla variedade de coisas que não somente bens e serviços; mas também por coisas intangíveis como prestígio, amor e respeito. Relacionado a isso, Jönsson (1998) afirma que os gerentes

trabalham com as palavras, e muito bem, reforça o autor. E ainda defende que a relevância da informação contábil está ligada à habilidade que ela tem de transmitir o trabalho dos gestores.

A Hipótese dos Custos Políticos defende que empresas de grande porte inseridas em setores regulados (tais como utilidade pública, petróleo e gás, telecomunicações) estão sujeitas a custos políticos potencialmente elevados, devido ao seus altos lucros. E, à medida que essas empresas ficam sujeitas a uma potencial transferência de recursos, os gestores tendem a adotar procedimentos que reduzam essa transferência, como por exemplo, a redução das *expectativas e variabilidade dos lucros* e a ofuscação linguística de informações (WATTS; ZIMMERMAM, 1978; RUTHERFORD, 2003; RIBEIRO ET AL, 2009).

Nesse tocante, alguns estudos foram realizados para verificar a utilização de técnicas de linguística com o intuito de ofuscação de informações nos relatórios contábeis narrativos. Kong (2001) investigou como identificar ofuscação e manipulação linguística nos RA. Usando a Teoria de Halliday, o autor analisou cinco seções dos relatórios: três sobre *marketing* organizacional, uma sobre as operações da companhia e uma sobre evidenciação sócio-ambiental. Os resultados mostraram que a seção de *marketing* organizacional é construída de forma sofisticada e que há um apelo para a intertextualidade, principalmente no que tange às relações com as operações da companhia.

Rutherford (2005) examinou se as empresas realizam um mau uso da complexidade textual a fim de ofuscar as informações operacionais e financeiras e assim prejudicar uma comunicação eficaz e as melhores práticas de governança. Para isso, o autor analisou os textos narrativos publicados pelas empresas baseando-se em algumas *proxys* já consideradas em estudos anteriores, como por exemplo, o número de caracteres, sentenças, parágrafos, palavras e legibilidade. Adicionalmente, o autor coletou os dados referentes às características das empresas e ao seu desempenho financeiro como porte, margem, retorno sobre o patrimônio líquido, retorno sobre o investimento, aumento ou diminuição da receita, aumento

ou diminuição da lucratividade, risco, número de segmentos explorados, regulação do setor, entre outras. O autor não encontrou relação significativa entre a utilização da complexidade textual e as características das empresas ou seu mau desempenho.

Moffit e Burns (2009) analisaram duzentas e duas empresas a fim de verificar a relação entre escolhas linguísticas e ofuscação da informação. Para isso, eles dividiram as empresas selecionadas na amostra em dois grupos denominados como fraudulentos e não fraudulentos com base em apontamentos da *Securities Exchange Commission (SEC)*. Utilizando-se de *softwares* que capturavam características do texto analisado, os autores desenvolveram vinte e cinco critérios linguísticos que poderiam sugerir ofuscação. Os resultados da pesquisa demonstraram que as empresas classificadas como fraudulentas colocaram em suas narrações palavras mais complexas e apelaram para aquelas com o sentido de “realização” e “causa”. Portanto, os autores concluíram que as empresas consideradas fraudulentas se utilizam da manipulação linguística para ofuscar seu baixo desempenho e alardear seus pontos positivos.

No Brasil, Ribeiro et al (2009) analisaram a existência de relação entre custos políticos, escolhas linguísticas e gerenciamento de imagem observando as seções narrativas dos RA da empresa Aracruz Celulose S.A, sob a perspectiva da gramática *Halliday*. Os resultados sugerem que a empresa priorizou enfatizar nos relatórios sua identidade institucional ao reforçar suas atitudes ao invés de informar suas realizações e seus planos.

2.3 Similaridade

Similaridade é um conceito complexo e amplamente discutido na linguística, na filosofia e na ciência da informação (HATZIVASSILOGLOU; KLAVANS; ESKIN, 1999)

Para Hatzivassiloglou; Klavans e Eskin (1999, p. 204), duas unidades de texto são similares quando compartilham o mesmo enfoque sobre um conceito, ator, objeto ou ação comum.

Para Choi, Hasting e Moore (2006), a similaridade métrica estima a probabilidade de dois segmentos descreverem o mesmo assunto. Segundo os autores, são dois os métodos de estimação da similaridade: o Método de Coesão Lexical desenvolvido por *Halliday* (1976) que detecta partes do texto com vocabulários semelhantes; e o Método da Repetição de Palavras, que detecta a sequência de palavras apresentadas.

De maneira semelhante, Uzuner, Davis e Katz (2004b) reforçam que para se avaliar a similaridade entre dois documentos é preciso analisar tanto o conteúdo quanto a expressão presente nos documentos. Os autores definem conteúdo como os fatos relatados no documento, enquanto que o termo expressão está ligado às preferências linguísticas dos autores enquanto apresentam esses fatos (conteúdos).

Para Frawley (1992 *apud* Uzuner, Davis e Katz, 2004b) existem duas formas de se avaliar a similaridade: pela igualdade entre os textos ou pela diferença. Assim, existem dois tipos de similaridade: a representacional (ou sintática) e a comportamental (ou semântica). A *similaridade sintática* relaciona o texto como uma sequência de caracteres formando uma estrutura textual complexa. Dessa maneira, esse tipo de similaridade pode ser definido em termos de forma, propriedades ou características de representação; e sua análise é realizada por meio do exame de blocos, de sentenças e de classes de textos. Já para o outro tipo, a similaridade *semântica*, dois textos são similares quando a semelhança está em seu significado; vale ressaltar que é mais difícil de capturar semelhanças desse tipo (WALENSTEIN ET AL, 2007).

Uzuner, Davis e Katz (2004a) afirmam que existe uma variedade de circunstâncias nas quais é útil determinar se o conteúdo de dois textos é similar, incluindo da detecção de plágio

e infrações contra direitos autorais à seleção e organização de documentos que retornaram de uma busca na internet.

Não obstante a sua importância, o estudo da similaridade textual tem sido esquecido pela comunidade científica (UZUNER; DAVIS; KATZ, 2004b). Nesse sentido, algumas pesquisas estão sendo desenvolvidas na área de Ciência da Informação e Programação com o objetivo de desenvolver metodologias de análise de similaridade entre dois documentos.

Hatzivassiloglou, Klavans, Eskin (1999) apresentaram uma metodologia de similaridade métrica que combina informações de vários indicadores linguísticos com a finalidade de medir a distância semântica entre pares de unidades textuais. Os autores verificaram que a técnica desenvolvida por eles supera as técnicas utilizadas até então, que é a busca de palavras comuns e posição das palavras no texto. Esse método tem sido empregado com sucesso em um sistema de computação.

Uzuner, Randall e Katz (2004a) revisaram metodologias de similaridade e verificaram que a maioria dessas metodologias utiliza-se da comparação de palavras e frases em dois textos. Os autores, argumentando que somente essas técnicas não eram suficientes para capturar a similaridade textual, propuseram a utilização de quatro classes de recursos para a análise da similaridade:

- 1) *Características de Aparência*: número médio de frases no texto, número médio de palavras, desvio padrão da quantidade de palavras, desvio padrão da quantidade de frases, entre outras;
- 2) *Características de Discurso*: frequência de palavras padrões, frequência de frases padrões, entre outras;
- 3) *Características Sintáticas*: frequência de frases declarativas, interrogativas e imperativas, frequência de frases na voz ativa e passiva, entre outras; e

- 4) *Características Semânticas*: frequência de negações (não, nunca), frequência de incerteza (talvez, provavelmente), entre outras.

Uzuner, Randall e Katz (2004b) realizaram um experimento como o objetivo de verificar a habilidade das pessoas em reconhecer conteúdos e expressões similares de documentos usando suas próprias experiências e conhecimento. Os pesquisadores constataram que similaridade sintática pode ser detectada mais facilmente.

Walenstein et al (2007) apresentaram modelos de mensuração de similaridade realizados por programas de computação. Esses *softwares* podem ser utilizados tanto para o tipo sintático quanto para o tipo semântico de similaridade. Os modelos de mensuração do tipo sintático são, por exemplo:

- ✓ similaridade textual: determina o quanto uma sequência de caracteres é semelhante a outra; e
- ✓ similaridade métrica: procura objetos que sejam semelhantes ou diferentes.

Um dos modelos de mensuração por tipo semântico é o modelo de entradas e saídas, que verifica a relação do produto adicionado e excluído, gerando em cada campo o mesmo resultado (WALENSTEIN ET AL, 2007).

3 PROCEDIMENTOS DE PESQUISA

Nesta seção são detalhados os procedimentos adotados para o alcance do objetivo proposto para a pesquisa. Inicialmente, são descritos a coleta e a fonte dos dados, os critérios para a seleção da amostra, a operacionalização dos indicadores de similaridade e o instrumento estatístico utilizado.

3.1 Coleta de Dados

Os RA das empresas brasileiras listadas na Bolsa de Valores e de Mercadorias e Futuros de São Paulo, referentes ao período de 2008 e 2009 foram coletados, um a um, no sítio da CVM³.

3.2 População e Amostra

O universo da pesquisa compreende as empresas brasileiras listadas na bolsa brasileira que em fevereiro de 2010 totalizavam 544.

A seleção da amostra levou em consideração dois fatores: (i) a disponibilidade dos RA, em formato *doc*, na página da CVM, uma vez que o *software* utilizado para cálculo da similaridade trabalha apenas com textos nesse formato; e (ii) a disponibilidade de dados econômico-financeiros no banco de dados da Economática®. Considerando-se esses fatores, a amostra final do estudo foi de 311 empresas, o que corresponde a 57% do universo de estudo (Apêndice A).

³ <http://www.cvm.gov.br/>

3.3 Obtenção das Estatísticas de Similaridade

Após a coleta, os RA foram submetidos a uma análise de similaridade textual. A metodologia utilizada para essa análise foi a de avaliação da similaridade por meio da detecção da diferença entre dois textos (conforme sugere FRAWLEY, 1992 APUD UZUNER, DAVIS E KATZ, 2004b). Para isso, foi selecionado um *software* de análise de conteúdo. O programa escolhido foi o *ExamDiff Pro*, que pode ser utilizado para análise de textos em língua portuguesa.

Essa ferramenta permite a comparação de arquivos de texto, destacando as diferenças até o nível de linhas, palavras e caracteres. Nesse sentido, para fins desse estudo, utilizou-se a similaridade sintática.

O programa realiza, também, correspondência difusa de forma que se permita alinhar as partes semelhantes do texto e detectar, automaticamente, as diferenças existentes entre os arquivos de textos comparados.

Após a comparação dos arquivos de textos, o *ExamDiff Pro* oferece três estatísticas de diferenças, apresentadas a seguir:

- ✓ Diferenças Totais entre os textos
- ✓ Diferenças de Linhas (ou Sentenças)
- ✓ Diferenças Intra-Linhas (ou Palavras)

A Estatística “Diferenças Totais” é uma ponderação realizada pelo programa das “Diferenças de Linhas” e “Intra-Linhas”. Já as estatísticas de Diferenças de Linhas e Intra-Linha são o resultado do somatório do número de sentenças ou palavras, conforme o caso, adicionadas (LADIC/ ILADIC), deletadas (LDEL/ ILDEL) e modificadas (LMOD/ILMOD).

Assim, tem-se a seguinte equação: Diferenças de Linhas = Número de Linhas Adicionadas + Número de Linhas Deletadas + Número de Linhas Modificadas. De maneira semelhante é obtido o cálculo de Diferenças Intra-Linhas.

Para o cálculo das estatísticas de similaridade, o *software* apresenta diversas opções de configurações, como por exemplo, a de desconsiderar linhas e espaços em branco e a opção de se aumentar o grau de detalhamento da análise (estatísticas aos níveis de parágrafos, sentenças, palavras e caracteres). Vale ressaltar que o *software* utilizado desconsidera, para fins de diagnóstico, gráficos, tabelas e figuras e, essa é uma de suas limitações. Além disso, a investigação da similaridade textual só é realizada para documentos em formato *doc*. A ferramenta oferece, ainda, a opção de se escolher o grau de similaridade esperado para dois documentos, variando de “0” completamente dissimilar, a “100” completamente similar. Quanto mais próximo de 100, menor a correspondência difusa considerada. O grau padrão recomendado pelo programa é de 60.

Assim, os RA foram importados para o *software* a fim de que fossem comparados. Essa comparação foi realizada empresa por empresa, ou seja, foram comparados os relatórios de 2008 e de 2009 de cada empresa presente na amostra. Essa comparação foi realizada para quatro graus de similaridade: 100, 90, 60 e 50⁴. Faz-se necessário destacar que para fins de comparação, o nível de detalhamento escolhido foi o de sentenças e palavras, bem como desconsiderados os espaços em branco.

Adicionalmente, foram obtidos o número de caracteres, palavras, sentenças e o valor da Legibilidade *Flesch*, de cada relatório presente na amostra, por meio do *Microsoft Word*.

⁴ A ausência de explicações no tutorial do programa sobre a razão de se considerar o grau de similaridade 60 como o padrão conduziu a análise de diferentes graus de similaridade.

3.4 Critérios Utilizados para a Obtenção dos Indicadores de Similaridade

Antes da criação dos indicadores de similaridade, foi necessária a realização de um Teste t (teste de média) para as três estatísticas de similaridade fornecidas pelo programa, ou seja, Diferenças Totais, Diferenças de Linhas e Diferenças Intra-Linhas.

A realização do teste t teve como finalidade identificar se as diferenças entre os relatórios de 2008 e 2009 são significantes para os quatro graus de similaridade considerados (100, 90, 60 e 50). A expectativa inicial, quando da aplicação do Teste t, era de que não fossem constatadas diferenças significantes para esses graus, o que tornaria mais simples a criação dos indicadores de similaridade.

Para a interpretação do teste de média, é necessária a verificação da hipótese nula (H_0) e da hipótese alternativa (H_1). Assim, foi realizado o teste de hipótese bi-caudal para verificar se a média encontrada na amostra de um grau de similaridade é igual ou diferente da média amostral encontrada em outro grau. O intervalo de confiança foi de 95%. O quadro 3.1 mostra as hipóteses a serem testadas:

Hipótese	Descrição	Hipótese	Descrição
H_{0a}	$\mu 100 = \mu 90$	H_{1a}	$\mu 100 \neq \mu 90$
H_{0b}	$\mu 100 = \mu 60$	H_{1b}	$\mu 100 \neq \mu 60$
H_{0c}	$\mu 100 = \mu 50$	H_{1c}	$\mu 100 \neq \mu 50$
H_{0d}	$\mu 90 = \mu 60$	H_{1d}	$\mu 90 \neq \mu 60$
H_{0e}	$\mu 90 = \mu 50$	H_{1e}	$\mu 90 \neq \mu 50$
H_{0f}	$\mu 60 = \mu 50$	H_{1f}	$\mu 60 \neq \mu 50$

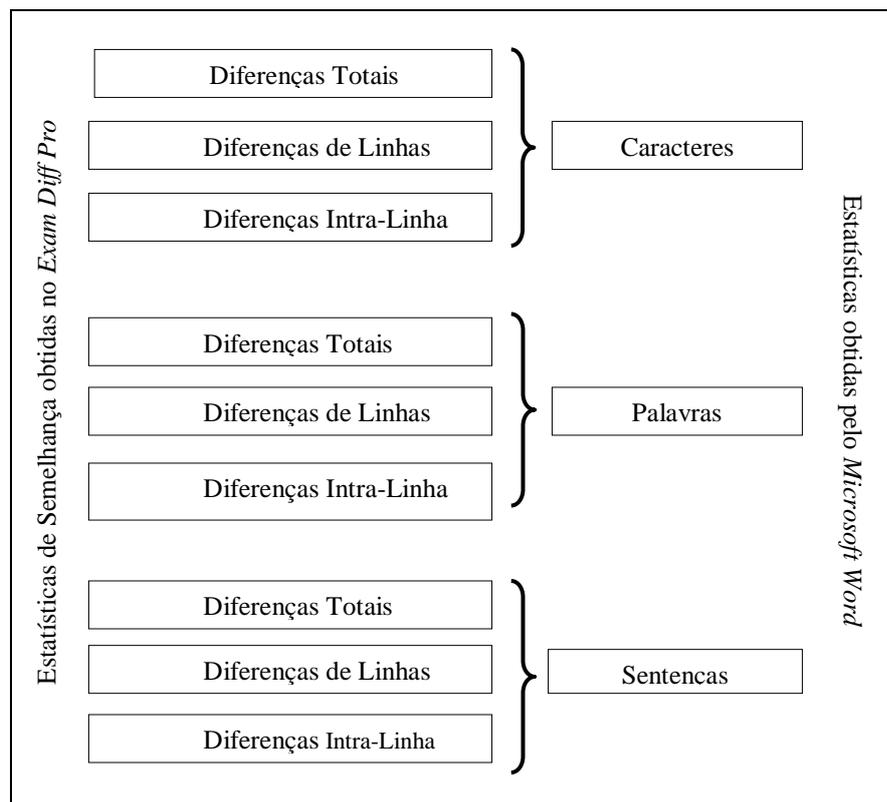
Fonte: elaboração própria.

Quadro 3.1 - Hipóteses do Teste de Média

Após a realização do teste de média verificou-se que as diferenças entre os graus de similaridade não eram significantes para as estatísticas Diferenças Totais e Diferenças de Linhas. Por conseguinte, a hipótese nula não poderia ser rejeitada. Esse resultado confirma que para a criação de indicadores, o grau de similaridade escolhido (100, 90, 60, 50) não

influenciaria o seu poder explicativo. Entretanto, para o indicador Diferenças Intra-Linhas, verificou-se uma única diferença significativa, entre os graus 100 e 60 (*p-value* 0,02). (Apêndice B)

Fundamentado no resultado do Teste t foram criados indicadores cruzando-se as variáveis obtidas no *software*, para os graus 100 e 60, com aquelas disponibilizadas pelo programa de texto. Esse procedimento se mostrou necessário devido ao problema de escala. A ilustração 3.1 exemplifica a forma que os indicadores foram criados:



Fonte: Elaboração Própria

Ilustração 3.1 - Exemplificação da Criação dos Indicadores

Tendo em vista que o procedimento efetuado, mencionado no parágrafo anterior resultou em quantidade razoável de variáveis (nove para cada grau de similaridade), realizou-se o cálculo da matriz de correlação, com o objetivo de verificar as inter-relações entre os indicadores e assim reduzi-los a **um único Indicador de Similaridade considerando o grau**

100 e a um único Indicador de Similaridade considerando o grau 60. Desse modo, foram calculados os coeficientes de correlação de *Pearson* (paramétrica) e de *Spearman* (não-paramétrica) e, em seguida realizado o teste de significância considerando 95% de confiança. Segundo Hair *et al* (2009, p. 150), o coeficiente de correlação indica “a força da associação entre quaisquer duas variáveis métricas. O sinal mais ou menos indica a direção dessa relação. O valor pode variar de -1 a +1, onde +1 indica uma perfeita relação positiva e, -1 uma perfeita relação negativa.”.

Com a Correlação de *Pearson*, verificou-se que os nove indicadores de *cada grau estudado* possuem correlação positiva entre si de no mínimo 0,60, significativa a 5% (*p-value* 0,00). Para a Correlação de *Spearman*, apurou-se que os indicadores apresentavam um coeficiente de correlação positivo entre si de no mínimo 0,90; também significativa a 5% (*p-value* 0,00). (Apêndice C)

Considerando as altas correlações verificadas entre os indicadores criados, foi escolhido apenas um indicador de cada grau para ser utilizado para fins de elaboração do modelo de regressão. No entanto, como no teste de média verificou-se que para a estatística “Diferenças Intra-Linhas” há diferenças significantes entre os dois graus (100 e 60) fornecidos, optou-se por utilizá-la como numerador do Indicador de Similaridade.

3.5 Instrumentos Estatísticos para a Apuração dos Resultados

A fim de verificar as hipóteses da pesquisa relacionadas à inclusão de maiores quantidades de informações nos RA e a manutenção de suas estruturas, de um período para o outro, procedeu-se com a análise das estatísticas descritivas das seguintes informações: Linhas Adicionadas (LADIC), Linhas Deletadas (LDEL), Linhas Modificadas (LMOD),

Intra-Linhas Adicionadas (ILADIC), Intra-Linhas Deletadas (ILDEL) e Intra-Linhas Modificadas (ILMOD).

Em seguida, para averiguar a influência das características: (i) complexidade linguística; (ii) complexidade organizacional; (iii) desempenho financeiro; (iv) risco; (v) custo político; e (vi) tamanho da empresa no Indicador de Similaridade dos RA e testar as demais hipóteses elencadas nessa pesquisa, usou-se a técnica de Regressão Múltipla (conforme utilizado por RUTHERFORD, 2003). Nesse sentido, foram utilizadas variáveis independentes constantes no quadro 3.2.

Variável	Proxy de Mensuração	Nome	Definição	Tipo	Origem dos Dados
Complexidade Organizacional	Setor de Atuação	SETOR	Conjunto de variáveis <i>dummies</i> que determinam setor de atuação da empresa Utilidade Pública, Financeiro e Outros, Bens Industriais, Materiais Básicos, Construção e Transportes, Materiais Básicos, Consumo Cíclico, Consumo não Cíclico, Tecnologia da Informação, Telecomunicações, Petróleo e Gás, Não Classificado	<i>Dummy</i>	BMF&BOVESPA
	Nível de Governança Corporativa	GC	Conjunto de variáveis <i>dummies</i> que determinam se a empresa participa de algum dos níveis de governança da BMF&BOVESPA; 0 = Não e 1 = Sim	<i>Dummy</i>	BMF&BOVESPA
	Imprensa	IP	Quantidade de notícias publicadas na imprensa especializada entre o período de 01 de janeiro de 2008 a 31 de dezembro de 2009	<i>Métrica</i>	<i>Factiva</i> ⁵
	Negociação em Bolsa Estrangeira	ADR	Conjunto de variáveis <i>dummies</i> que determinam se a empresa possui valores mobiliários negociados em bolsas norte-americanas (<i>American Depositary Receipts</i>)	<i>Dummy</i>	NYSE ⁶
	Concentração Acionária	CONCENT	Porcentagem de participação do 1º maior acionista no capital da empresa	<i>Métrica</i>	<i>Economática</i>
	Auditoria	AUDIT	Conjunto de variáveis <i>dummies</i> que determinam se a empresa é auditada por uma <i>Big Four</i> ou não em 2008; 0 = não e 1 = sim <i>Big Four: Ernest&Young, PriceWaterHouse Coopers, KPMG e Deloitte</i>	<i>Dummy</i>	<i>Economática</i>

⁵ Banco de Dados pertence ao grupo *Down Jones* (www.factiva.com)

⁶ Bolsa de Valores de Nova Iorque (www.nyse.com)

Variável	Proxy de Mensuração	Nome	Definição	Tipo	Origem dos Dados
Desempenho Financeiro	Variação do Patrimônio Líquido Contábil	DΔPL	Conjunto de variáveis <i>dummies</i> que determinam se o Patrimônio Líquido em 2009 aumentou ou diminuiu em relação ao do ano de 2008; 0 = diminuição e 1 = aumento	<i>Dummy</i>	<i>Econômica</i>
	Variação da Lucratividade	DΔLL	Conjunto de variáveis <i>dummies</i> que determinam se o Lucro Líquido em 2009 aumentou ou diminuiu em relação ao do ano de 2008; 0 = diminuição e 1 = aumento	<i>Dummy</i>	<i>Econômica</i>
	Variação da Receita	DΔREC	Conjunto de variáveis <i>dummies</i> que determinam se a Receita Operacional Líquida em 2009 aumentou ou diminuiu em relação a do ano de 2008; 0 = diminuição e 1 = aumento	<i>Dummy</i>	<i>Econômica</i>
	Liquidez de Mercado	LIQMERC	Relação entre o volume de negociação em 2008 e o patrimônio líquido em 2008	<i>Métrica</i>	<i>Econômica</i>
	Independência Financeira	IF	Relação entre o Patrimônio Líquido e o Ativo Total em 31/12/2008	<i>Métrica</i>	<i>Econômica</i>
	Liquidez Corrente	LC	Relação entre o Ativo Circulante e o Passivo Circulante em 31/12/2008	<i>Métrica</i>	<i>Econômica</i>
Complexidade Linguística	Variação do Tamanho do RA	ΔRA	Diferença entre o tamanho do RA (em caracteres) do ano de 2009 e o tamanho do RA do ano de 2008	<i>Métrica</i>	<i>Microsoft Word</i>
	Variação do Tamanho do RA	DΔRA	Conjunto de variáveis <i>dummies</i> que determinam se o tamanho (em caracteres) do RA de 2009 aumentou ou diminuiu em relação ao ano de 2008; 0 = diminuição e 1 = aumento	<i>Dummy</i>	<i>Microsoft Word</i>
	Legibilidade	LEG	Facilidade de Leitura ⁷	<i>Métrica</i>	<i>Microsoft Word</i>
Risco	Beta	BETA	Medida estatística que reflete o risco de um ativo em relação ao risco da carteira de mercado (ASSAF NETO, 2006)	<i>Métrica</i>	<i>Econômica</i>
Custos Políticos	Setor Regulado	SREG	Conjunto de variáveis <i>dummies</i> que determinam se a empresa pertence a um setor com forte regulamentação por parte do Governo ou não; 0 = não e 1 = sim Setores Regulados: utilidade pública, financeiros e outros; petróleo e gás e telecomunicações	<i>Dummy</i>	<i>BMF & BOVESPA</i>
Tamanho da Empresa	Porte	PORTE	Logaritmo Natural do Ativo Total (em milhares de reais) em 31/12/2008	<i>Métrica</i>	<i>Econômica</i>

Fonte: Elaboração própria
Quadro 3.2 – Variáveis Independentes

⁷ A Legibilidade está relacionada à Facilidade de Leitura de um texto, considerando variáveis como comprimento de frases (quantidade média de palavras) e de palavras (quantidade média de sílabas ou letras). Nesse sentido, a Legibilidade Flesch classifica a facilidade de leitura de um texto numa escala de 0 a 100 (difícil a muito fácil), considerando a quantidade de sílabas por palavras e a quantidade de palavras por sentença (SMITH; TAFLER, 1992).

A análise da influência das variáveis independentes no Indicador de Similaridade foi realizada por total e por grupo.

Primeiramente aplicou-se a técnica de regressão múltipla para todas as empresas selecionadas na amostra para os graus 100 e 60. Em seguida, o indicador de cada grau foi dividido em ordem crescente de similaridade e segregado em *tercis* de alta, média e baixa similaridade.⁸

Posteriormente, comparou-se o indicador agregado dos *tercis* de valores altos e baixos indicadores de similaridade com as respectivas características de complexidade linguística e organizacional, desempenho financeiro, risco, custo político e tamanho das empresas pertencentes a cada *tercil*.

Nesse sentido, o modelos da regressão, por total, para os graus 100 e 60 de similaridade assume a forma mostrada na expressão 1:

$$\begin{aligned}
 IS_i = & \beta_1 + \beta_2 \cdot Setor_i + \beta_3 \cdot GC_i + \beta_4 \cdot Porte_i + \beta_5 \cdot LC_i + \beta_6 \cdot IP_i + \\
 & \beta_7 \cdot ADR_i + \beta_8 \cdot Concent_i + \beta_9 \cdot Audit_i + \beta_{10} \cdot Leg_i + \beta_{11} \cdot D\Delta RA_i + \\
 & \beta_{12} \cdot D\Delta PL_i + \beta_{13} \cdot D\Delta LL_i + \beta_{14} \cdot D\Delta Re c_i + \beta_{15} \cdot LiqMerc_i + \\
 & \beta_{16} \cdot Beta_i + \beta_{17} \cdot IF_i + \varepsilon
 \end{aligned} \quad (1)$$

Onde:

IS= Indicador de Similaridade da empresa i

B_n = Coeficiente da Variável n, para a empresa i

ε = termo de erro da regressão

E os modelos por grupo para os graus 100 e 60 de similaridade assume a forma mostrada na expressão 2:

$$\begin{aligned}
 IS_{i,j} = & \beta_1 + \beta_2 \cdot Setor_{i,j} + \beta_3 \cdot GC_{i,j} + \beta_4 \cdot Porte_{i,j} + \beta_5 \cdot LC_{i,j} + \beta_6 \cdot IP_{i,j} + \\
 & \beta_7 \cdot ADR_{i,j} + \beta_8 \cdot Concent_{i,j} + \beta_9 \cdot Audit_{i,j} + \beta_{10} \cdot Leg_{i,j} + \beta_{11} \cdot D\Delta RA_{i,j} + \\
 & \beta_{12} \cdot D\Delta PL_{i,j} + \beta_{13} \cdot D\Delta LL_{i,j} + \beta_{14} \cdot D\Delta Re c_{i,j} + \beta_{15} \cdot LiqMerc_{i,j} + \\
 & \beta_{16} \cdot Beta_{i,j} + \beta_{17} \cdot IF + \varepsilon
 \end{aligned} \quad (2)$$

⁸ Ressalta-se, mais uma vez, que nesse trabalho a Similaridade é medida pela Diferença (FRAWLEY, 1992 APUD UZUNER, DAVIS E KATZ, 2004b)

Onde:

IS = Indicador de Similaridade da empresa i, no grupo (tercil) j;

B_n = Coeficiente da Variável n, para a empresa i, no grupo (tercil) j;

ε = termo de erro da regressão

Para a seleção das variáveis pertencentes aos modelos foi adotado o método *stepwise* a um nível de 95% de confiança, o qual seleciona as variáveis para a inclusão no modelo de acordo com o melhor preditor da variável dependente, utilizando-se para isso valores de probabilidade F entre 0,05 e 0,1. As variáveis preditoras adicionais são selecionadas conforme o poder explicativo incremental que oferecem (HAIR ET AL, 2009).

Além disso, para confirmar se as premissas do modelo clássico de regressão linear foram mantidas e assim proceder com a análise adequada dos resultados, foram realizados os seguintes testes de robustez para os modelos calculados, disponíveis no *software* estatístico (SPSS): (i) teste de normalidade, que verifica a normalidade dos erros; (ii) o teste de *Durbin-Watson* (DW) que averigua a existência de autocorrelação de primeira ordem entre os termos de erros do modelo; (iii) a análise do Fator de Inflação da Variância (VIF⁹), que avalia a existência de multicolinearidade entre as variáveis independentes.

3.6 Resultados Esperados

Como resultado das Regressões Múltiplas, espera-se encontrar relação positiva ($\beta > 0$) entre o indicador de similaridade e as variáveis relacionadas à complexidade linguística, ao desempenho financeiro, ao risco, aos custos políticos e ao tamanho da empresa. Se esse resultado for verificado, demonstrar-se-á a preocupação dos administradores em evidenciar as realizações da administração, visto será evidenciada uma tentativa dos gestores de justificar para os acionistas e demais interessados, os aumentos nos níveis de atividade.

⁹ Sigla para o termo em inglês *Variance Inflation Factor*.

Para as variáveis referentes à complexidade organizacional, acredita-se que a participação da empresa nos níveis de governança corporativa, a negociação de valores mobiliários no mercado estrangeiro e o fato de as empresas serem auditadas por uma das quatro maiores empresas de auditoria do mundo afetaria consideravelmente os relatórios, provocando um aumento de informações disponíveis aos investidores dessas empresas e, dessa maneira, aumentando as diferenças entre os relatórios de um ano para o outro. Portanto, para essas variáveis espera-se encontrar também relação positiva ($\beta > 0$) com o Indicador de Similaridade (ver, por exemplo, WATTS; ZIMMERMAN, 1976; COURTIS, 1998; FERNANDEZ, 2008; RUTHERFORD, 2003).

No entanto, acredita-se que quanto maior a concentração acionária e o risco de investimento que as empresas oferecem, menor as diferenças entre os RA divulgados de um período para o outro ($\beta < 0$). A concentração do controle acionário nas mãos de um pequeno grupo de acionistas pode ocasionar uma menor agregação de informações aos relatórios, visto que não há interesse dos acionistas majoritários e, ainda, demanda do mercado por determinados tipos de informação da empresa, uma vez que a maior parte dos interessados participa da gestão.

Acredita-se, também, que quanto maior o risco atribuído a empresa, menos similares deveriam ser os seus relatórios, visto que esse fato diminuiria a assimetria informacional e, por conseguinte a incerteza para o investidor.

4 RESULTADOS

Nesta seção são apresentadas as estatísticas descritivas, bem como os resultados das regressões para cada Indicador de Similaridade (100 e 60) com as características de complexidade linguística e organizacional, de desempenho financeiro, de risco, de custos políticos e de tamanho das empresas.

A maneira com que essas características explicam o Indicador de Similaridade dos RA é verificada pelo sinal e valor dos coeficientes beta (β) e pelo valor do coeficiente de determinação do modelo, R^2 ajustado. A confirmação dessa influência é apurada pela análise da significância estatística das variáveis independentes individualmente, pela estatística *t - Student* (sig. teste t), e em conjunto por meio do teste F (sig. teste F). As tabelas com os resultados das regressões estão apresentadas nessa seção, juntamente as respectivas análises e o Apêndice E contem os resultados dos testes de normalidade dos erros.

4.1 Estatísticas Descritivas

A tabela 4.1 apresenta as estatísticas descritivas do número de Linhas (sentenças) e Intra-Linhas (palavras) adicionadas, deletadas e modificadas quando comparados os RA de 2008 e 2009 para os graus de similaridade 60 e 100, respectivamente.

Tabela 4.1 – Estatísticas Descritivas

		N	Mínimo	Máximo	Média	Desvio-Padrão	Assimetria	Curtose
Grau 60	LADIC	311	0	465	7,60	2,09	8,79	92,72
	ILADIC	311	0	1508	197,21	12,28	2,28	7,16
	LDEL	311	0	81	3,43	0,61	4,44	22,27
	ILDEL	311	0	1644	182,95	12,18	2,77	10,98
	LMUD	311	0	3876	364,05	23,20	3,13	18,47
	ILMUD	311	0	7868	905,98	58,22	2,69	10,42
Grau 100	LADIC	311	0	465	7,65	2,09	8,79	92,69
	ILADIC	311	0	1521	207,22	12,82	2,28	6,88
	LDEL	311	0	81	3,43	0,61	4,44	22,27
	ILDEL	311	0	1609	193,50	12,67	2,61	9,29
	LMUD	311	0	3876	366,13	23,16	3,14	18,57
	ILMUD	311	0	7650	922,58	59,25	2,60	9,36

Onde:

LADIC = Linhas Adicionadas; LDEL = Linhas Deletadas;

LMUD = Linhas Mudadas ILADIC = Intra-Linhas Adicionadas;

ILDEL = Intra-Linhas Deletadas; ILMUD = Intra-Linhas Modificadas

Ainda analisando a Tabela 4.1, observa-se que o número médio de linhas modificadas é significativamente maior do que o número médio de linhas adicionadas (o número de mudanças nas linhas é aproximadamente 52 vezes maior que a adição de novas linhas). É importante frisar que a estatística de modificações é calculada quando a estrutura das frases permanece a mesma e as alterações ocorridas estão relacionadas às palavras pontuais. Os relatórios da Sergen S.A e Abyara Planejamentos Imobiliários S.A exemplificam essas constatações (ver Anexos A, B, C e D). No RA da Companhia Sergen S.A (ver Anexos A e B) nenhuma linha do texto foi adicionada ou deletada e as alterações realizadas no relatório de 2009, quando comparado ao de 2008, trata-se apenas de alterações dentro da linha, ou seja, nas palavras. Nesses casos, a análise de conteúdo detecta apenas a estatística de linhas modificadas, que no caso específico da companhia em questão é igual a três. De maneira contrária, no relatório da Abyara Planejamentos Imobiliários S.A (ver Anexos C e D) ocorreram alterações significativas na estrutura do RA, dessa forma, o programa detectou 74 linhas a mais no relatório de 2008 quando comparado ao de 2009.

Adicionalmente, ao se analisar a Tabela de Frequências de Sentenças Adicionadas e Deletadas para o grau 60 (Apêndice D), identificou-se que aproximadamente 70% das empresas estudadas não adicionaram nem deletaram nenhuma sentença ao RA de 2009 em relação ao de 2008. Resultados semelhantes foram encontrados quando os mesmos dados foram verificados para o grau 100 (Apêndice D).

Dessa forma, os resultados corroboram com as hipóteses dessa pesquisa de que grande parte das empresas tende a manter um padrão de RA, apesar das alterações ocorridas no cenário macroeconômico nos últimos anos.

Além disso, os relatórios permanecem com a mesma quantidade de informação, apresentando apenas alterações nas datas e valores. Contudo, entende-se que informações relacionadas ao desempenho, posição financeira e patrimonial e principais incertezas às quais uma companhia está sujeita deveriam ocasionar adição na quantidade de informações divulgadas ao mercado em anos subsequentes.

4.2 Relação entre as Variáveis Independentes e o Indicador de Similaridade, total

4.2.1 Grau de Similaridade 100

No modelo de regressão calculado para verificar se as características de complexidade organizacional e linguística, de desempenho financeiro, de risco, de custos políticos e de tamanho da empresa explicavam o Índice de Similaridade grau 100, somente as variáveis: (i) variação do tamanho do RA; (ii) porte; legibilidade, variação do PL contábil, governança corporativa, setor de telecomunicações e setor de materiais básicos retornaram resultados significativos (sig. do teste $t < 0,05$). Quando se tem um coeficiente significante

estatisticamente, implica dizer que a variável independente em questão pode ajudar a explicar o comportamento da variável dependente, no caso deste estudo, o Indicador de Similaridade.

Em relação ao teste de existência do modelo, obteve-se um coeficiente F no valor de 10,33, significativa a 5%. Nesse sentido, se aceita a hipótese da existência da regressão. Além disso, o modelo apresentou um poder explicativo (R^2 ajustado= 0,175) de 17,5%. A tabela 4.2 demonstra os resultados.

Tabela 4.2 – Variáveis que explicam o Indicador de Similaridade 100, por total

Variável	Variáveis Preditoras	β		Teste		VIF	R^2 Ajustado	ϵ	Teste		
		β	Padr.	t	Sig.				F	Sig.	
	Constante	6,20		2,57	0,01		0,18	5,41	1,96	10,33	0,03
Complexidade Linguística	DARA	3,24	0,27	5,17	0,00	1,03	0,18	5,41	1,96	10,33	0,03
	LEGIB	0,32	0,12	2,21	0,03	1,17					
Desempenho Financeiro	DAPL	-0,14	-0,20	-3,76	0,00	1,04	0,18	5,41	1,96	10,33	0,03
Complexidade Organizacional	GC	3,19	0,18	3,45	0,00	1,04	0,18	5,41	1,96	10,33	0,03
	STELEC	1,70	0,14	2,53	0,01	1,18					
	SMATBAS	4,20	0,13	2,32	0,02	1,08					
Tamanho da Empresa	PORTE	1,60	0,12	2,17	0,03	1,08	0,18	5,41	1,96	10,33	0,03

Onde:

DARA = Dummy da Variação do Tamanho do RA; LEGIB = Legibilidade2008; DAPL = Dummy Variação do PL Contábil; GC = Governança Corporativa;

STELEC = Setor de Telecomunicações; SMATBAS = Setor de Materiais Básicos; PORTE = Porte da Empresa

Os valores críticos de DW para 100 ou mais observações a 95% de confiança são: dL= 1,59 e dU=1,76

Para a variável Legibilidade encontrou-se um $\beta > 0$, conforme o esperado. Esse resultado demonstra que a facilidade de leitura está positivamente relacionada com o aumento das diferenças nos relatórios narrativos de um ano para o outro e, portanto, com a diminuição da similaridade.

A variável variação do PL apresentou um $\beta < 0$. Isso significa que quando o Patrimônio Líquido aumenta, verifica-se também um aumento na similaridade nos RA de um ano para o outro. Isso demonstra que, possivelmente, os gestores não se preocuparam em detalhar as informações de desempenho ao mercado. Esse resultado corrobora com os disponíveis na literatura de que quando a empresa apresenta um desempenho igual ou superior, não há

preocupação dos gestores em explicar o desempenho ao mercado. No entanto, nos casos em que ocorrem desempenhos inferiores, há um aumento de informações disponibilizadas aos investidores, evidenciando, provavelmente, uma tentativa dos gestores em justificar o desempenho apresentado.

De acordo com o esperado, a variável governança corporativa se mostrou significativa e com um $\beta > 0$. Esse resultado evidencia que quando uma empresa possui melhores práticas de governança corporativa devido ao fato de pertencer a algum dos níveis de governança da Bovespa, o nível de similaridade dos RA diminui. Esse resultado demonstra uma maior preocupação dos gestores em evidenciar a boa administração, atendendo os preceitos de transparência e proteção ao acionista.

De maneira semelhante, o porte também se apresentou significativo e com o $\beta > 0$. Isso comprova que quanto maior o porte da empresa, menor a similaridade dos seus RA divulgados em anos subsequentes. Os resultados encontrados corroboram com a hipótese de custos políticos no sentido de que, as empresas de maior porte por terem maior visibilidade no mercado e na sociedade, se preocupam em prestar maior quantidade de informações ao mercado visando justificar as suas atividades e ofuscar informações a cerca do seu desempenho.

Foi possível verificar também que o setor de materiais básicos tende a apresentar relatórios narrativos menos similares, ou seja, esse setor agrega mais informação em seus RA de um período para o outro ($\beta > 0$). Destaca-se que ramos de atividades como papel e celulose e defensivos e fertilizantes fazem parte desse setor, segundo a classificação da BMF&BOVESPA. Nesse sentido, o resultado encontrado pode ser justificado, já que empresas desses ramos de atividades devem possuir uma maior preocupação em quantificar e evidenciar informações relativas às questões ambientais.

Semelhantemente, o setor de telecomunicações também se mostrou significativo e com $\beta > 0$. Essa informação sugere que as empresas desse setor tendem a apresentar relatórios menos similares. É importante frisar que esse setor é considerado um setor regulado e alvo de constantes fiscalizações por parte do governo. Por conseguinte, o fato de apresentar relatórios com maior agregação de informação pode ser fruto de uma demanda da sociedade, do mercado e do governo de maneira geral ou até mesmo uma tentativa de ofuscação linguística, conforme sugere a literatura.

De maneira geral, o fato de outras variáveis independentes não terem se mostrado explicativas para esse grau de similaridade pode sugerir que a manutenção de estruturas similares de relatórios é uma prática comum nas empresas brasileiras estudadas na amostra.

4.2.2 Grau de Similaridade 60

No segundo modelo de regressão calculado, somente cinco variáveis retornaram como significativas (sig. do teste $t < 0,05$): (i) variação do tamanho do RA; (ii) porte; (iii) legibilidade; (iv) variação do PL contábil; e (v) setor de materiais básicos. Além disso, o modelo apresentou um poder explicativo de 12,5% (R^2 Ajustado = 0,125).

No que tange ao teste de existência da regressão, o coeficiente F apresentou o valor de 9,813 significativo a 5%; dessa maneira aceita-se a hipótese da existência do modelo, conforme demonstrado na tabela 4.3.

Tabela 4.3 - Variáveis que explicam o Indicador de Similaridade 60, por total

Variável	Variáveis Preditoras	β	β Padr.	Teste t	Sig.	VIF	R^2 Ajustado	ϵ	DW	Teste F	Sig.
	Constante	5,11		1,94	0,05		0,13	6,09	1,81	9,81	0,03
Complexidade Linguística	DARA	2,76	0,21	3,92	0,00	1,02	0,13	6,09	1,81	9,81	0,03
	LEGIB	-0,11	-0,14	-2,64	0,01	1,01					
Desempenho Financeiro	DAPL	2,15	0,14	2,60	0,01	1,06	0,13	6,09	1,81	9,81	0,03
Complexidade Operacional	SMATBAS	4,07	0,21	3,93	0,00	1,03	0,13	6,09	1,81	9,81	0,03
Tamanho da Empresa	PORTE	0,34	0,12	2,21	0,03	1,04	0,13	6,09	1,81	9,81	0,03

Onde:

DARA = Dummy da Variação do RA; LEGIB = Legibilidade2008; DAPL= Dummy da Variação do PL

SMATBAS = Setor de Materiais Básicos; PORTE = Porte da Empresa

Os valores críticos de DW para 100 ou mais observações a 95% de confiança são: dL= 1,59 e dU=1,76

Analisando-se a Tabela 4.3, observa-se que essas mesmas variáveis se mostraram significantes no modelo de regressão anterior considerando grau 100. No entanto, as variáveis legibilidade e variação do PL contábil apresentaram sinal inverso ao obtido anteriormente. Assim, para o grau de Similaridade 60, a variável legibilidade apresentou um $\beta < 0$; e a variável variação do PL apresentou o coeficiente $\beta > 0$.

Portanto, nesse caso, um aumento na facilidade de leitura dos relatórios provoca um aumento na similaridade dos RA. E, um aumento no PL provoca uma diminuição da similaridade dos RA em períodos subsequentes. Destaca-se que apesar dos resultados serem contraditórios com os encontrados para o grau 100 e até mesmo os constatados pela teoria, esse grau de similaridade é o padrão estabelecido pelo programa.

4.3 Relação entre as Variáveis Independentes e Indicador de Similaridade, tercil

4.3.1 Primeiro Tercil - Grau de Similaridade 100

Os resultados do modelo de regressão para o grupo com os menores valores de indicadores de similaridade retornou como significantes (sig. do teste $t < 0,05$) as variáveis variação do lucro líquido, beta, auditoria, variação do tamanho do RA e setor de construção e transporte.

Em relação ao teste de existência da regressão, obteve-se um coeficiente F no valor de 9,478, significativa a 5%. Nesse sentido, se aceita a hipótese da existência da regressão. O poder explicativo do modelo foi de 64,8% (R^2 ajustado = 0,648). A tabela 4.4 demonstra os resultados encontrados.

Tabela 4.4 - Variáveis que explicam o Indicador de Similaridade 100, por tercil (primeiro tercil)

Variável	Variáveis Preditoras	β	β Padr.	Teste t	Sig.	VIF	R^2 Ajustado	ϵ	DW	Teste F	Sig.
	Constante	4,89		8,91	0,00		0,65	0,85	1,21	9,48	0,00
Complexidade Linguística	DARA	1,01	0,29	2,14	0,05	1,29	0,65	0,85	1,21	9,48	0,00
Desempenho Financeiro	DALL	1,21	0,42	2,99	0,01	1,39	0,65	0,85	1,21	9,48	0,00
Complexidade Operacional	AUDIT	-1,27	-0,37	-2,67	0,02	1,21	0,65	0,85	1,21	9,48	0,00
	SCONST	-2,49	-0,77	-5,28	0,00	1,24					
Risco	BETA	1,44	0,34	2,50	0,02	1,24	0,65	0,85	1,21	9,48	0,00

Onde:

DARA = Dummy da Variação do RA;

DALL = Dummy da Variação do Lucro Líquido; SCONST = Setor de Construção e Transporte;

AUDIT = Auditoria;

BETA = Beta

Esse modelo testou a influência da complexidade linguística e organizacional, desempenho financeiro, custo político, risco e tamanho para aquelas empresas que divulgaram RA com menores diferenças de um período para o outro, ou seja, cujos relatórios apresentam maior similaridade.

Para esse teste, a variável de desempenho variação do lucro líquido se mostrou significativa e com um $\beta > 0$, conforme o esperado. Esse resultado evidencia que para as empresas que apresentaram RA mais similares de um ano para o outro, o fato de aumentar o

lucro líquido provoca um aumento na quantidade de informações disponíveis nos RA dessas empresas.

A variável auditoria apresentou um $\beta < 0$. Isso demonstra que as empresas que apresentaram os RA **mais similares** de um período para o outro, em sua maioria, **não** foram auditadas pelas maiores firmas de auditoria do mundo. Portanto, o fato de uma empresa não ser auditada por uma das *big four* contribui para que elas apresentem RA mais similares; ressalta-se nesse ponto que apesar de os RA não serem objeto de auditoria independente, a literatura destaca que eventualmente a empresa de auditoria pode revisá-lo.

O risco também se mostrou significativo nesse modelo de regressão, apresentando $\beta > 0$. Depreende-se, portanto que as empresas que apresentam RA mais similares se mostram como maiores risco de investimentos. Como o preço de qualquer ativo é formado com base nas diversas informações publicamente disponível aos investidores, conforme defende a literatura, em casos em que a informação é reduzida, a assimetria da informação faz com que se aumente o risco e a incerteza de quem está aplicando recursos na empresa.

Outro fator apontado pelo modelo de regressão, é que o setor de construção e transporte influencia negativamente ($\beta < 0$) o nível de similaridade dos RA. Isso significa que as empresas que apresentaram RA mais similares, em sua maioria, não pertencem aos referidos setores. Deduz-se que existe uma preocupação por parte dos gestores das empresas desses setores em divulgar as recentes realizações, além de justificar os elevados resultados alcançados em virtude dos altos índices de crescimento ocorrido nos últimos anos, com o reaquecimento da construção civil e do transporte.

4.3.2 Terceiro Tercil - Grau de Similaridade 100

Os resultados do modelo de regressão para o grupo com os maiores valores de indicadores de similaridade retornou como significativo (sig. do teste $t < 0,05$) somente a variável imprensa.

Obteve-se um coeficiente F no valor de 26,281, significativo a 5%; confirmando a hipótese de existência da regressão. O poder explicativo do modelo foi de 45,7% (R^2 ajustado= 0,457). Os resultados estão demonstrados na tabela 4.5.

Tabela 4.5 - Variáveis que explicam o Indicador de Similaridade 100, por tercil (terceiro tercil)

Variável	Variáveis Preditoras	β	β Padr.	Teste t	Sig.	VIF	R^2 Ajustado	ϵ	DW	Teste F	Sig.
	Constante	11,73		25,54	0,00		0,46	2,37	0,14	26,28	0,00
Complexidade Organizacional	IP	0,00	0,69	5,13	0,00	1,00	0,46	2,37	0,14	26,28	0,00

Onde:

IP = Quantidade de notícias publicadas pela imprensa especializada

Para esse teste, a imprensa apresentou um $\beta > 0$, conforme o esperado. Esse resultado evidenciou que para as empresas que apresentaram RA menos similares, quanto maior o número de notícias veiculadas pela imprensa especializada, maior a quantidade de informações incrementais nos relatórios de um período para o outro.

A imprensa pode influenciar o mercado acionário por duas razões: a primeira é que ela consegue atingir grande número de pessoas e a segunda relaciona-se à capacidade da imprensa em formar opiniões. Portanto, o fato de a quantidade de notícias veiculadas pela mídia influenciar a publicação de RA **menos similares** pode ser justificado pela maior necessidade da administração em confirmar e/ou tentar ofuscar as informações noticiadas pela imprensa.

4.3.3 Primeiro Tercil – Grau de Similaridade 60

Para o modelo de regressão para os indicadores que apresentaram menores valores de indicadores de similaridade (primeiro tercil) referente ao grau 60, nenhuma variável se mostrou significativa.

4.3.4 Terceiro Tercil - Grau de Similaridade 60

Em relação às empresas que apresentaram maiores valores de indicador de similaridade (terceiro tercil) relacionado ao grau 60, a única variável significativa (sig. do teste $t < 0,05$) foi a Variação do PL Contábil.

O modelo apresentou um poder explicativo de 19,1% (R^2 ajustado= 0,191) um coeficiente F de 7,3932 significativa a 5%, confirmando a hipótese de existência do modelo calculado, conforme pode ser verificado pela tabela 4.6.

Tabela 4.6 - Variáveis que explicam o Indicador de Similaridade 60, por tercil (terceiro tercil)

Variável	Variáveis Preditoras	β	β Padr.	Teste t	Sig.	VIF	R^2 Ajustado	ϵ	DW	Teste F	Sig.
	Constante	15,93		12,21	0,00		0,19	2,92	0,12	7,39	0,01
Desempenho Financeiro	D Δ PL	-3,91	-0,47	-2,72	0,01	1,00	0,19	2,92	0,12	7,39	0,01

Onde:

D Δ PL= Variação do PL Contábil

A variação do PL contábil influenciou negativamente ($\beta < 0$) no valor do indicador de similaridade. Portanto, para aquelas empresas que apresentaram relatórios **menos similares**, o aumento incremental da quantidade de informações no RA de 2009 em relação ao de 2008 está relacionado à diminuição do valor do PL contábil. O que se pode concluir é que há uma

tentativa por parte dos gestores em justificar a diminuição do PL de um ano para o outro com vistas a justificar uma possível piora no desempenho, conforme defende a literatura.

4.4 Análise das Premissas do Modelo Clássico de Regressão Linear

A fim de verificar se as premissas relacionadas ao modelo clássico de regressão linear foram mantidas, realizou-se os testes de robustez.

Pela estatística *Durbin-Watson (DW)*, foi aceita a hipótese de que não há evidências significantes de autocorrelação de resíduos de primeira ordem para os modelos em que estão presentes todas as empresas da amostra. No entanto, para os modelos em que a amostra foi dividida em grupo (tercis), verificou-se autocorrelação positiva dos resíduos. Os modelos com presença de autocorrelação são tidos como regressões espúrias, isto é, a estimação da variável dependente é prejudicada pela ausência de variáveis ou pela formulação inadequada do modelo¹⁰. Portanto, esses modelos foram analisados, mas não serão usados para fins de conclusão dessa pesquisa.

A estatística de colinearidade fornecida pelo **Fator de Inflação da Variância (VIF)** apresenta valores próximos a 1,0 para todos os modelos analisados, sugerindo que a correlação entre as variáveis independentes é baixa¹¹.

Em relação à hipótese de normalidade dos erros, foram realizados dois testes diagnósticos: (i) o **Gráfico de Probabilidade Normal** que compara a distribuição cumulativa de valores de dados reais com a distribuição cumulativa de uma distribuição normal; e (ii) o **teste de Kolmogorov-Sminorv**, que calcula o nível de significância para as diferenças em relação a uma distribuição normal. A aplicação do teste de normalidade foi efetivada para

¹⁰ Segundo Brooks (2008), em casos onde existe autocorrelação positiva dos resíduos, os valores dos erros obtidos serão subavaliados em relação aos verdadeiros valores. Nesse sentido, aumenta a probabilidade de se incorrer em erro tipo I, ou seja, tendência a rejeitar a hipótese nula quando ela é verdadeira.

¹¹ Segundo Hair et al (2009), se o VIF é igual a 1,0 significa que a tolerância é 1,0 e assim não há multicolinearidade, e o erro padrão não é afetado.

ambos os graus na análise total, utilizando-se um nível de 95% de confiança, obtendo-se que a distribuição dos dados é normal (Apêndice E). Pela análise gráfica de probabilidade normal dos resíduos, verificou-se que os resíduos se comportam de maneira normal, porém sem elevações¹² (Apêndice F).

No entanto, para a análise por tercil, o resultado da aplicação do teste de *Kolmogorov-Sminov* demonstrou que algumas variáveis não apresentaram distribuição normal. No entanto, isso não invalida os resultados visto que a amostra utilizada nesses modelos é suficientemente grande e pode-se recorrer ao Teorema do Limite Central.¹³

4.5 Síntese das Hipóteses de Pesquisa e Resultados Encontrados

Nesta seção será apresentada uma síntese dos resultados encontrados, bem como a verificação das hipóteses de pesquisa levantadas, conforme demonstra o quadro 4.1.

Com base nas estatísticas de sentenças e palavras adicionadas, deletadas e modificadas constatou-se que em mais de 70% das empresas brasileiras reproduziram a estrutura do RA em anos subsequentes. Além disso, os relatórios dessas empresas permaneceram com a mesma quantidade de informação, apresentando apenas alterações nas datas e valores (Hipóteses I e II).

Hipótese da Pesquisa	Descrição	Aceita	Rejeitada	Parcialmente Confirmada
I	As empresas brasileiras tendem a reproduzir as estruturas de seus RA ao longo do tempo	X		

¹² Segundo Hair et al (2009, p. 83), “A distribuição normal forma um reta diagonal e os dados do gráfico são comparados com a diagonal. Se uma distribuição é normal, a linha que representa a distribuição real dos dados segue muito próxima a diagonal.”

¹³ Brooks (2008) afirma que para amostras suficientemente grandes, violações das premissas de normalidade não provocam grandes consequências. Utilizando-se o Teorema do Limite Central, as estatísticas do teste assintoticamente atingirão as distribuições adequadas, mesmo na ausência de normalidade.

Hipótese da Pesquisa	Descrição	Aceita	Rejeitada	Parcialmente Confirmada
II	Não há modificações significativas no que tange a quantidade de informações divulgadas nos RA das empresas brasileiras para o período estudado	X		
III	Em função do seu tamanho, as empresas tendem a apresentar RA mais ou menos similares em períodos subsequentes	X		
IV	Empresas pertencentes a setores regulados tendem a apresentar RA menos Similares			X
V	Empresas que apresentam melhores indicadores de desempenho tendem a apresentar RA menos Similares		X	
VI	Empresas que apresentam estruturas organizacionais mais complexas tendem a apresentar RA menos Similares			X
VII	Empresas com maiores indicadores de concentração acionária tendem a apresentar RA mais Similares		X	
VIII	Empresas que representam maiores riscos de investimentos tendem a divulgar RA mais similares		X	

Fonte: Elaboração própria

Quadro 4.1 – Resumo dos Resultados e verificação das Hipóteses da Pesquisa

Para os modelos de regressão por total analisados, o Indicador de Similaridade foi influenciado pelo porte da empresa. Nesse sentido, o tamanho da empresa está diretamente relacionado à divulgação de RA menos similares (Hipótese III).

No que tange à Hipótese IV, que pressupõe que empresas pertencentes a setores regulados tendem a apresentar narrativas contábeis menos similares, essa hipótese foi parcialmente confirmada. Isso porque somente um dos modelos de regressão analisados apresentou como significativo um dos setores considerados regulados, o de telecomunicações. Além disso, a variável setor regulado, que abrange todos os setores regulados no modelo, não se apresentou como significativa em nenhum dos testes realizados. Vale salientar que, apesar de não ser eminentemente regulado, o setor de materiais básicos se mostrou como significativo no que se refere à divulgação de RA menos similares.

A Hipótese V versava que empresas com melhores indicadores de desempenho financeiro apresentavam RA com menos similaridade. No entanto, essa hipótese foi rejeitada.

Inicialmente porque somente um dos indicadores de desempenho se mostrou relevante (*dummy* da variação do PL contábil) e depois porque esse mesmo indicador apresentou diferentes sinais de coeficiente β nos modelos analisados.

No que se refere à Hipótese VI de que empresas com estruturas organizacionais mais complexas apresentariam RA com menores Índices de Similaridade, essa foi parcialmente confirmada, em virtude de somente a Governança Corporativa variável ter se mostrado significativa, e em apenas um dos modelos analisados.

A Hipótese VII, que prevê que empresas com maiores concentrações acionárias tendem a apresentar RA com maior similaridade devido ao fato de o controle acionário estar nas mãos de poucos acionistas não pode ser confirmada porque em nenhum dos modelos investigados a variável concentração se mostrou significativa.

E, por fim, a hipótese de que as empresas que representam maiores riscos de investimentos tendem a divulgar RA mais similares também não pode ser confirmada, embora essa variável tenha retornado como significativa no modelo de regressão por *tercil* (Hipótese VIII).

5 CONSIDERAÇÕES FINAIS

Os RA são utilizados como uma ferramenta de comunicação das empresas para transmitir uma mensagem ou imagem para seus acionistas. Considerando a importância de que esses relatórios apresentem, cada vez mais, informações compreensíveis, relevantes, confiáveis e comparáveis, este estudo propôs uma investigação quanto à estrutura, a fim de verificar se há repetição nas informações que a administração divulga de um exercício social para outro; tendo em vista que não foram encontrados estudos semelhantes na área de Contabilidade no Brasil. O objetivo do estudo foi, ainda, identificar quais as variáveis explicam a similaridade entre os RA divulgados em anos subsequentes pelas empresas negociadas no mercado acionário brasileiro.

Utilizando-se o *Exam Diff Pro*, foram analisados trezentos e onze Relatórios de Administração de empresas listadas na bolsa brasileira, publicados juntamente com as Demonstrações Contábeis encerradas em 31/12/2008 e 31/12/2009. Feito isso, foi mensurada a similaridade entre os RA. Dessa maneira, construiu-se um indicador que representasse o nível de similaridade desses relatórios, bem como se verificou quais variáveis relacionadas à complexidade linguística e organizacional, desempenho financeiro, risco, tamanho e custos políticos das empresas poderiam ajudariam explicar esse indicador.

Os resultados encontrados demonstraram que mais de 70% das empresas brasileiras constantes da amostra não apresentaram alterações significativas na estrutura do Relatório da Administração de 2009, em relação ao de 2008. Constatou-se, portanto, que grande parte das empresas realizou apenas modificações pontuais como: data, nomes e valores. Esse resultado confirma as seguintes hipóteses levantadas nesse estudo: (i) a de que as empresas brasileiras reproduzem determinadas estruturas em seus RA em anos subsequentes; e (ii) a hipótese de que não houve alterações significativas no que tange à quantidade de informações inseridas

nos relatórios referentes os anos de 2008 e 2009. Ao confirmar essa hipótese, esse estudo evidencia que as empresas brasileiras podem estar deixando de contribuir com os investidores em suas decisões, bem como os induzindo a possíveis erros. Além disso, as empresas podem estar, também, perdendo uma grande oportunidade de divulgar suas realizações ao mercado.

Adicionalmente, verificou-se que o tamanho da empresa influencia positivamente na modificação das informações contidas nos relatórios (iii). Assim, quanto maior o porte da empresa, menor a similaridade dos seus RA em anos subsequentes. Assim, o que se pode constatar é que existe uma preocupação dos administradores das grandes empresas em apresentar sua gestão, por meio de alterações e/ou de agregação de informações em suas narrativas. Além disso, o resultado verificado vai ao encontro do que está descrito na Literatura de que as empresas de maior porte, por apresentarem custos políticos elevados e grande visibilidade no mercado, apresentam relatórios com vistas a justificar o maior nível de suas atividades.

Vale salientar que três hipóteses levantadas não puderam ser comprovadas nesse estudo e, nesse sentido, foram rejeitadas: (v) a de que empresas que apresentam melhores indicadores de desempenho tendem a apresentar relatórios menos similares; (vii) a de que empresas com maiores índices de concentração acionária tendem a apresentar relatórios mais similares; (viii) e a de que empresas que representam maior risco de investimento apresentam RA mais similares. Em relação à primeira, apesar de o resultado referente à variação do PL ter se mostrado significativo nos dois modelos, o comportamento dessa variável se mostrou contraditório nas duas situações. Ademais, não foi possível comprovar nesse estudo, por meio dos testes realizados, se as empresas pertencentes aos setores considerados regulados (iv) e àquelas com estruturas organizacionais mais complexas (vi) tendem a apresentar relatórios menos similares para o período analisado.

De maneira geral, o fato de outras variáveis independentes não terem se mostrado explicativas nos modelos obtidos, pode sugerir que a manutenção de estruturas similares de relatórios é uma prática comum nas empresas contidas na amostra, independente das suas características, conforme afirma Rutherford (2003).

A falta de confirmação de apenas parte das hipóteses de pesquisa levantadas pode estar relacionada ao período abrangido no estudo. O curto período analisado é um fator limitador da pesquisa. Entretanto, essa restrição se fez necessária, visto que a análise de períodos adicionais implicaria em um aumento exponencial dos indicadores do estudo.

Outro aspecto limitante da pesquisa está relacionado à impossibilidade de se controlar as variáveis macro-econômicas que afetam as informações contidas nas narrativas contábeis, tais como crise financeira e particularidades do setor de atuação das empresas. Somado a isso, a dificuldade em se ponderar de maneira objetiva o momento de transformação que a Contabilidade está vivenciando com o advento das mudanças nas normas e a busca internacional pela convergência.

E ainda, as restrições inerentes ao tipo de mensuração de similaridade utilizada. A similaridade sintática relaciona dois textos analisando uma sequência de palavras, portanto, mede somente as diferenças existentes entre as palavras apresentadas e a suas quantidades, não fazendo ponderações se essas diferenças apresentam o mesmo valor semântico.

Nesse sentido, para futuras pesquisas sugere-se aumentar o período analisando, a fim de se obter resultados mais abrangentes. Além disso, seria interessante verificar o nível de similaridade dos relatórios anuais de empresas de outros países. Acredita-se que fatores como cultura, regime político e nível de amadurecimento do mercado de capitais podem estar relacionados à prática de manutenção de estruturas semelhantes de relatórios em períodos subsequentes.

É importante destacar que o método utilizado pode ser aplicado para avaliar similaridade em outros instrumentos de divulgação ao mercado, como por exemplo, os prospectos de oferta pública, laudos de avaliação de empresas, fatos relevantes e até mesmo formulários de referência.

REFERÊNCIAS

- ABRAHAMSON, E.; AMIR, E. The information content of the President's Letter to Shareholders. **Journal of Business Finance Accounting**, v.23, n. 8, 1996
- ARMENIC, Joel H.; CRAIG, Russell J. 9/11 in Service of Corporate Rhetoric: Southwest Airlines 2001 Letters to Shareholders. **Journal of Communications Inquiry**, v. 28, 2004
- ASSAF NETO, Alexandre. **Finanças Corporativas e Valor**. 2ed. São Paulo: Atlas, 2006
- BEATTIE, Vivian; MCINNES, Bill; FEARNLEY Stella. A methodology for analyzing and evaluating narratives in annual reports: a comprehensive descriptive profile and metrics for disclosure quality attributes. **The Accounting Forum**, v. 28, p. 205-236, 2004
- BEAUGRANDE, R.; DRESSLER, W. U. **Introduction to text linguistics**. London: Longman, 1981
- BEUREN, Ilse; BOFF, Marines L.; HEIN, Nelson. Informações Recomendadas Pelo Parecer de Orientação N° 15/87 da CVM nos RA de Empresas Familiares de Capital Aberto. 8° In: CONGRESSO USP DE CONTROLADORIA E CONTABILIDADE. 8, 2008, São Paulo, **Anais...** São Paulo: EAC/FEA/USP. 2008. Disponível em <http://www.congressoeac.locaweb.com.br/new/artigos82008/69.pdf> Acesso 09 mar. 2010
- BOO, El'Fred; SIMNETT, Roger. The Information Content of Management's Prospective Comments in Financially Distressed Companies: A Note. **ABACUS**, v. 38 n.2, p. 280-295, 2002.
- BRASIL. Lei das Sociedades por Ações. Lei 6.404, de 15.12.1976. Dispõe sobre as Sociedades por Ações. Publicada no Diário Oficial da União em 17.12.77. Disponível em <www.planalto.gov.br/>. Acesso em: 07 mar. 2010.
- BROOKS, Chris. **Introductory Econometrics for Finance**. Cambridge: Cambridge University Press, 2008.
- BRYAN, S. H. Incremental information content of required disclosures contained in management discussion and analysis. **The Accounting Review**, v. 72 n. 2, p. 285–301, 1997
- CHOI, Freddy Y. Y.; HASTING, Peter Wiemer; MOORE, Johanna. Latent Semantic Analysis for Text Segmentation. **Computational Linguistics**, v. 32, n. 1, p. 5-12, 2006
- COMISSÃO DE VALORES MOBILIÁRIOS (CVM). Parecer de Orientação CVM N° 15, de 28 de dezembro de 1987. Disponível em <www.cvm.gov.br/>. Acesso em: 07 mar. 2010.
- COMITÊ DE PRONUNCIAMENTOS CONTÁBEIS (CPC). Pronunciamento 26. Disponível em <www.cpc.org.br/>. Acesso em: 07 mar. 2010.
- COURTIS, J. K. Annual report readability variability: Tests of the obfuscation hypothesis. **Accounting Auditing and Accountability Journal**, v. 11, n. 4, p. 459–471, 1998.

CUNHA, Rafael Koifman C.; SILVA, César Augusto T.; Análise da Facilidade de Leitura das Demonstrações Contábeis das Empresas Brasileiras: Uma Investigação do Gerenciamento de Impressões nas Narrativas Contábeis. 9º In: CONGRESSO USP DE CONTROLADORIA E CONTABILIDADE. 9, 2009, São Paulo, **Anais...** São Paulo: EAC/FEA/USP. 2009. Disponível em <http://www.congressousp.fipecafi.org/artigos92009/132.pdf> Acesso 09 mar. 2010

DYE, Ronald. An evaluation of “essays on disclosure” and the disclosure literature in Accounting. **Journal of Accounting and Economics**, n. 32, p. 181-135, 2001

FERNANDEZ, Marcelo Otávio L. **Prêmio de Controle no Brasil**. 2008. 148 f. Programa de Pós-Graduação e Pesquisa em Administração e Economia. Ibmec. 2008

GALLON, Alessandra V.; BEUREN, Ilse M. Análise da Relação entre Evidenciação nos RA e o Nível de Governança das Empresas na Bovespa. 30º In: ENCONTRO NACIONAL DA ANPAD. 30, 2006, Salvador, **Anais...** Salvador: ANPAD. 2006.

HAIR, Jr., J. F.; ANDERSON, R.E.; TATHAM, R. L.; BLACK, W. C.. **Análise multivariada de dados**. 6 ed.. São Paulo: Bookman, 2009.

HALLIDAY, M. A. K. **System and function in language**. London: Oxford University Press, 1976

_____. **A Introduction of Funcional Grammar**. London: Edwars Arnold, 1985

HATZIVASSILOGLOU, Vasileios; KLAVANS, Judith; ESKIN, Eleazar. Detecting Text Similarity over Short Passages: Exploring Linguistic Feature Combinations via Machine Learning. **The College of Information Sciences and Technology**, v. 37, 1999. Disponível on line em <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.9798> Acesso em 10 mar. 2010

HERREMAN, I. RYANS, J. Jr. The case for better measurement and reporting of marketing performance. **Business Horizons**, v. 38, n. 5, p. 51-60, 1995

IUDÍCIBUS, Sérgio; MARTINS, Eliseu; GELBCKE, Ernesto R. **Manual de Contabilidade das Sociedades por Ações**. 7 ed.. São Paulo: Atlas, 2007

JONES, M. J., SHOEMAKER, P. A. Accounting narratives: A review of empirical studies of content and readability. **Journal of Accounting Literature**, v. 13, p. 142–184, 1994.

JÖNSON, S. Related Management Accounting Research to Managerial Work! **Accounting, Organizations and Society**, v.23, n. 4, p. 411-434, 1998

KONG, Kenneth C. C. Marketing of Belief: Intertextual Construction of Network marketers' identities. **Discussion&Society**, v.12, 2001

LE COADIC, Yves-François. **A Ciência da Informação**. Brasília: Editora Briquet de Lemos, 2004

LLEWELLYN, Sue. Narratives in accounting and management research. **Journal of Accounting, Auditing and Accountability**, v. 12, n. 2, p. 220-236, 1999

LUZ, Antonio Thadeu M.; PAGLIARUSSI, Marcelo S. An analysis of the influence of Corporate Governance principles on the use of self-serving attributions in corporate Annual Reports. 3º, 2009. In: CONGRESSO IAAER-ANPCONT. 3, 2009, São Paulo, **Anais...** São Paulo: ANPCONT. 2009. Disponível em <http://www.anpcont.com.br/site/docs/congressoIII/02/268.pdf> Acesso em 09 mar. 2010

MARCONI, Marina Andrade; LAKATOS, Eva Maria. **Técnicas de Pesquisa**. 7ed. São Paulo: Atlas, 2009

MARSTON, C. L., SHRIVES, P. J. The use of disclosure indices in accounting research: A review article. **British Accounting Review**, v. 23, n. 3, p. 195–210, 1991

MARTINS, Gilberto Andrade.; THEÓPHILO, Carlos Renato. **Metodologia da Investigação Científica para as Ciências Sociais** Aplicadas. 2 ed. São Paulo: Atlas, 2009

MOFFIT, Kevin; BURNS, Mary B. What Does That Mean? Investigating Obfuscation and Readability Cues as Indicators of Deception in Fraudulent Financial Reports. **Proceeding of the Fifteenth Americas Conference on Information Systems**, 2009.

REEVES, Terrie C.; FORD, Eric W.; DUNCAN, W. Jack; GINTER, Peter M. Communication clarity in strategic management data sources. **Strategic Organization**, v. 3, n. 3, p. 243-278, 2005

RIBEIRO, Jusimeire F., PAGLIARUSSI, Marcelo S.; SILVA, Alfredo R. L.; SILVA JÚNIOR, Annor. Escolhas Linguísticas, Custos Políticos e Gerenciamento da Imagem Cooperativa: O Caso da Aracruz Celulose S.A.. 3º, 2009. In: CONGRESSO IAAER-ANPCONT. 3, 2009, São Paulo, **Anais...** São Paulo: ANPCONT. 2009. Disponível em <http://www.anpcont.com.br/site/docs/congressoIII> Acesso em 09 mar. 2010

RODRIGUES, Fernanda F. **Análise das Variáveis que Influenciam as Informações Divulgadas nos RA das Companhias Abertas Brasileiras: um estudo empírico nos anos de 2001 a 2003**. 2005. 118f. Dissertação (mestrado em Ciências Contábeis) – Programa Multiinstitucional e Inter-Regional de Pós- Graduação em Ciências Contábeis, Universidade de Brasília, Brasília, 2005.

RUTHERFORD, Brian A.. Obfuscation, Textual Complexity and The Role of Regulated Narrative Accounting Disclosure in Corporate Governance. **Journal of Management and Governance**, v. 7, p. 187-203, 2003

SCAPENS, R.W.; ROBERTS, J. Accounting and control: a case study of resistance to Accounting change. **Management Accounting Research**, vol. 4, n. 1, p. 1-32, 1993

SILVA, César Augusto T., RODRIGUES, Fernanda F., ABREU, Robson L. Análise dos RA das Companhias Abertas Brasileiras: um Estudo do Exercício Social de 2002. **Revista de Administração Contemporânea**, v. 11, n. 2, p. 71-92, 2007

SCOTÁ, Renato. O uso do conteúdo descritivo dos RA para inferência da continuidade das empresas. 3º, 2009. In: CONGRESSO IAAER-ANPCONT. 3, 2009, São Paulo, **Anais...** São Paulo: ANPCONT. 2009. Disponível em <http://www.anpcont.com.br/site/docs/congressoIII/02/190.pdf> Acesso em 09 mar. 2010

SMITH, Malcolm; TAFFLER, Richard. Readability and Understandability: Different Measures of Textual Complexity of Accounting Narrative. **Accounting, Auditing and Accountability Journal**, v. 5, n. 4, p. 84-98, 1992

SPRINGER, Lisa. Enhancing the Annual Report: Investor Relations and the MD&A. **The Journal of Bank Accounting & Auditing**, v. 5, n. 2, p.27-29, 1992

SYDSERFF, R., WEETMAN, P. A texture index for evaluating accounting narratives: An alternative to readability formulas. **Accounting, Auditing and Accountability Journal**, v.12, n. 4, p. 459–488, 1999.

TESSAROLLO, Izabella F.; PAGLIARUSSI, Marcelo S. The Justification of Organizational Performance in Annual Report Narratives. . 3º, 2009. In: CONGRESSO IAAER-ANPCONT. 3, 2009, São Paulo, **Anais...** São Paulo: ANPCONT. 2009. Disponível em <http://www.anpcont.com.br/site/docs/congressoIII/02/270.pdf> Acesso em 09 mar. 2010

UZUNER, Olzem; DAVIS, Randall; KATZ, Boris. Recognizing Text Similarity. MIT **Computer Science and Artificial Intelligence Laboratory**. 2004a

UZUNER, Olzem; DAVIS, Randall; KATZ, Boris. Using Empirical Methods for Evaluation Expression and Content Similarity. **Hawaii International Conference on System Sciences**. 2004

ZIMMERMAM, Jerold L. **Accounting for Decision Making and Control**. Boston: Irwin, 1998

WALENSTEIN, A., EL-RAMLY, M., CORDY, J. R., EVANS, W. S., MAHDAVI, K., PIZKA, M., RAMALINGAM, G., VON GUDENBERG, J. W. Similarity in programs. **Dagstuhl Seminar Proceedings**. 2007 Disponível em <http://www4.informatik.tu-muenchen.de/~pizka/mp06g.pdf> Acesso 10 mar. 2010

WALLIMAN, Nicholas. **Your Research Project – A step by step guide for the first-time researcher**. 2ed. London: Sage Publications, 2005

WATTS, Ross L.; ZIMMERMAM, Jerold L. **Positive Accounting Theory**. New Jersey: Prentice-Hall, 1976

YAMAMOTO, Marina Mitiyo; SALOTTI, Bruno Meirelles. **Informação Contábil: Estudo sobre a sua Divulgação no Mercado de Capitais**. São Paulo: Atlas, 2006

YUTHAS, Kristi et al. Communicative Action and Corporate Annual Reports. **Journal of Business Ethics**, v. 41, p. 141, 2002.

LISTA DE APÊNDICES

APÊNDICE A – EMPRESAS PRESENTES NA AMOSTRA	71
APÊNDICE B – RESULTADO DO TESTE T	78
APÊNDICE C – MATRIZES DE CORRELAÇÃO	79
APÊNDICE D – TABELAS DE FREQUÊNCIAS DAS ESTATÍSTICAS DE SIMILARIDADE	83
APÊNDICE E – RESULTADOS DOS TESTES DE NORMALIDADE	84
APÊNDICE F – GRÁFICOS DE PROBABILIDADE NORMAL	89

APENDICE A – Empresas Presentes na Amostra

Nome	Setor	Governança
521 PARTICIPACOES S.A.	Utilidade Pública	-
ABYARA PLANEJAMENTO IMOBILIARIO S.A.	Construção e Transporte	-
ACOS VILLARES S.A.	Materiais Básicos	-
AES ELPA S.A.	Bens Industriais	-
AES SUL DISTRIB GAUCHA DE ENERGIA S.A.	Bens Industriais	-
AES TIETE S.A.	Bens Industriais	-
AETATIS SECURITIZADORA S.A.	Financeiro e Outros	-
AFLUENTE GERAÇÃO DE ENERGIA ELÉTRICA S.A.	Utilidade Pública	-
AGRA EMPREENDIMENTOS IMOBILIARIOS S.A.	Construção e Transporte	-
ALFA HOLDINGS S.A.	Financeiro e Outros	-
ALIANSCÉ SHOPPING CENTERS S.A.	Financeiro e Outros	NM
ALL AMERICA LATINA LOGISTICA S.A.	Construção e Transporte	N2
ALTERE SECURITIZADORA S.A.	Financeiro e Outros	-
AMERICAN BANKNOTE S.A.	Bens Industriais	NM
AMIL PARTICIPACOES S.A.	Consumo Não Cíclico	NM
AMPLA ENERGIA E SERVICOS S.A.	Utilidade Pública	-
AMPLA INVESTIMENTOS E SERVICOS S.A.	Utilidade Pública	-
ANHANGUERA EDUCACIONAL PARTICIPACOES S.A	Consumo Cíclico	N2
B2W - COMPANHIA GLOBAL DO VAREJO	Consumo Cíclico	NM
BAHEMA S.A.	Financeiro e Outros	-
BANESTES S.A. - BCO EST ESPIRITO SANTO	Financeiro e Outros	-
BCO ABC BRASIL S.A.	Financeiro e Outros	N2
BCO ALFA DE INVESTIMENTO S.A.	Financeiro e Outros	-
BCO AMAZONIA S.A.	Financeiro e Outros	-
BCO BRADESCO S.A.	Financeiro e Outros	N1
BCO BRASIL S.A.	Financeiro e Outros	NM
BCO CRUZEIRO DO SUL S.A.	Financeiro e Outros	N1
BCO DAYCOVAL S.A.	Consumo Não Cíclico	NM
BCO ESTADO DO PARA S.A.	Financeiro e Outros	-
BCO INDUSVAL S.A.	Bens Industriais	NM
BCO MERCANTIL DO BRASIL S.A.	Construção e Transporte	-
BCO NORDESTE DO BRASIL S.A.	Consumo Cíclico	-
BCO PANAMERICANO S.A.	Financeiro e Outros	N1
BCO PINE S.A.	Consumo Cíclico	N1
BCO SOFISA S.A.	Financeiro e Outros	N2
BEMATECH S.A.	Tecnologia da Informação	NM
BHG S.A. - BRAZIL HOSPITALITY GROUP	Consumo Cíclico	NM
BOMBRIIL S.A.	Consumo Não Cíclico	-
BR MALLS PARTICIPACOES S.A.	Financeiro e Outros	NM
BRASIL BROKERS PARTICIPACOES S.A.	Construção e Transporte	NM
BRASIL ECODIESEL IND COM BIO.OL.VEG.S.A.	Petróleo, Gás e Combustível	NM
BRASIL TELECOM S.A.	Telecomunicações	-
BRASKEM S.A.	Materiais Básicos	N1
BRASMOTOR S.A.	Consumo Cíclico	-
BRAZILIAN FINANCE E REAL ESTATE S.A.	Financeiro e Outros	-
BRB BCO DE BRASILIA S.A.	Financeiro e Outros	-
BRF - BRASIL FOODS S.A.	Consumo Não Cíclico	NM
BROOKFIELD INCORPORAÇÕES S.A.	Construção e Transporte	NM
CAMARGO CORREA DESENV. IMOBILIARIO S.A.	Construção e Transporte	NM
CAMBUCI S.A.	Consumo Cíclico	-

Nome	Setor	Governança
CELULOSE IRANI S.A.	Materiais Básicos	-
CEMEPE INVESTIMENTOS S.A.	Financeiro e Outros	-
CENTRAIS ELET BRAS S.A. – ELETROBRAS	Utilidade Pública	N1
CENTRAIS ELET DE SANTA CATARINA S.A.	Utilidade Pública	N2
CENTRAIS ELET DO PARA S.A. – CELPA	Utilidade Pública	-
CENTRAIS ELET MATOGROSSENSSES S.A.- CEMAT	Utilidade Pública	-
CESP - CIA ENERGETICA DE SAO PAULO	Financeiro e Outros	-
CETIP S.A. - BALCÃO ORGANIZADO DE ATIVOS E DERIV.	Financeiro e Outros	NM
CIA BANDEIRANTES DE ARMAZENS GERAIS	Construção e Transporte	-
CIA BRASILEIRA DE DISTRIBUICAO	Consumo Não Cíclico	N1
CIA CACIQUE DE CAFE SOLUVEL	Consumo Não Cíclico	-
CIA CATARINENSE DE AGUAS E SANEAM.-CASAN	Utilidade Pública	-
CIA CELG DE PARTICIPACOES – CELGPAR	Utilidade Pública	-
CIA CONCESSOES RODOVIARIAS	Construção e Transporte	NM
CIA DISTRIB DE GAS DO RIO DE JANEIRO-CEG	Utilidade Pública	-
CIA ELETRICIDADE EST. DA BAHIA – COELBA	Utilidade Pública	-
CIA ENERGETICA DE BRASILIA	Utilidade Pública	-
CIA ENERGETICA DE MINAS GERAIS – CEMIG	Financeiro e Outros	N1
CIA ENERGETICA DE PERNAMBUCO – CELPE	Utilidade Pública	-
CIA ENERGETICA DO CEARA – COELCE	Consumo Não Cíclico	-
CIA ENERGETICA DO MARANHAO – CEMAR	Utilidade Pública	-
CIA ENERGETICA DO RIO GDE NORTE – COSERN	Construção e Transporte	-
CIA ESTADUAL DE DISTRIB ENER ELET-CEEE-D	Utilidade Pública	-
CIA ESTADUAL GER.TRANS.ENER.ELET-CEEE-GT	Utilidade Pública	-
CIA FERRO LIGAS DA BAHIA – FERBASA	Materiais Básicos	-
CIA FIACAO TECIDOS CEDRO CACHOEIRA	Consumo Cíclico	N1
CIA GAS DE SAO PAULO – COMGAS	Utilidade Pública	-
CIA HABITASUL DE PARTICIPACOES	Financeiro e Outros	-
CIA HERING	Consumo Cíclico	NM
CIA IGUACU DE CAFE SOLUVEL	Consumo Não Cíclico	-
CIA INDUSTRIAL CATAGUASES	Consumo Cíclico	-
CIA INDUSTRIAL SCHLOSSER S.A.	Consumo Cíclico	-
CIA MELHORAMENTOS DE SAO PAULO	Materiais Básicos	-
CIA PARANAENSE DE ENERGIA – COPEL	Utilidade Pública	N1
CIA PARTICIPACOES ALIANCA DA BAHIA	Financeiro e Outros	-
CIA PROVIDENCIA INDUSTRIA E COMERCIO	Materiais Básicos	NM
CIA SANEAMENTO BASICO EST SAO PAULO	Utilidade Pública	NM
CIA SANEAMENTO DE MINAS GERAIS-COPASA MG	Utilidade Pública	NM
CIA SANEAMENTO DO PARANA – SANEPAR	Utilidade Pública	-
CIA SIDERURGICA NACIONAL	Materiais Básicos	-
CIA TECIDOS SANTANENSE	Telecomunicações	-
CIELO S.A.	Financeiro e Outros	NM
CIMS S.A.	Financeiro e Outros	-
CLARION S.A. AGROINDUSTRIAL	Consumo Não Cíclico	-
COARI PARTICIPACOES S.A.	Financeiro e Outros	-
COBRASMA S.A.	Bens Industriais	-
CONFAB INDUSTRIAL S.A.	Materiais Básicos	N1
CONSERVAS ODERICH S.A.	Consumo Não Cíclico	-
CONSORCIO ALFA DE ADMINISTRACAO S.A.	Financeiro e Outros	-
CONSTRUTORA BETER S.A.	Construção e Transporte	-

Nome	Sector	Governança
CONSTRUTORA LIX DA CUNHA S.A.	Construção e Transporte	-
CONSTRUTORA SULTEPA S.A.	Construção e Transporte	-
CONSTRUTORA TENDA S.A.	Construção e Transporte	-
CONTAX PARTICIPACOES S.A.	Bens Industriais	-
CPFL ENERGIA S.A.	Utilidade Pública	NM
CPFL GERACAO DE ENERGIA S.A.	Utilidade Pública	-
CR2 EMPREENDIMENTOS IMOBILIARIOS S.A.	Construção e Transporte	NM
CREMER S.A.	Consumo Não Cíclico	NM
CTEEP - CIA TRANSMISSÃO ENERGIA ELÉTRICA PAULISTA	Utilidade Pública	N1
DIAGNOSTICOS DA AMERICA S.A.	Consumo Não Cíclico	NM
DIMED S.A. DISTRIBUIDORA DE MEDICAMENTOS	Consumo Não Cíclico	-
DINAMICA ENERGIA S.A.	Financeiro e Outros	-
DIRECIONAL ENGENHARIA S.A.	Construção e Transporte	NM
DIXIE TOGA S.A.	Materiais Básicos	-
DOHLER S.A.	Consumo Cíclico	-
DROGASIL S.A.	Consumo Não Cíclico	NM
DUKE ENERGY INT. GER. PARANAPANEMA S.A.	Utilidade Pública	-
DURATEX S.A.	Materiais Básicos	NM
ECORODOVIAS INFRAESTRUTURA E LOGÍSTICA S.A.	Não Classificado	-
EDP - ENERGIAS DO BRASIL S.A.	Utilidade Pública	NM
ELECTRO ACO ALTONA S.A.	Bens Industriais	-
ELEKEIROZ S.A.	Materiais Básicos	-
ELEKTRO - ELETRICIDADE E SERVICOS S.A.	Utilidade Pública	-
ELETROBRÁS PARTICIPAÇÕES S.A. – ELETROPAR	Utilidade Pública	-
ELETROPAULO METROP. ELET. SAO PAULO S.A.	Utilidade Pública	N2
ELUMA S.A. INDUSTRIA E COMERCIO	Materiais Básicos	-
EMAE - EMPRESA METROP.AGUAS ENERGIA S.A.	Utilidade Pública	-
EMBRAER-EMPRESA BRAS DE AERONAUTICA S.A.	Bens Industriais	NM
EMPRESA ENERG MATO GROS.SUL S.A.-ENERSUL	Utilidade Pública	-
EMPRESA NAC COM REDITO PART S.A.ENCORPAR	Consumo Cíclico	-
ENERGISA S.A.	Utilidade Pública	-
EQUATORIAL ENERGIA S.A.	Utilidade Pública	NM
ESPIRITO SANTO CENTR.ELETR. S.A.-ESCELSA	Utilidade Pública	-
ESTACIO PARTICIPACOES S.A.	Consumo Cíclico	NM
ETERNIT S.A.	Construção e Transporte	NM
EUCATEX S.A. INDUSTRIA E COMERCIO	Materiais Básicos	-
EXCELSIOR ALIMENTOS S.A.	Consumo Não Cíclico	-
EZ TEC EMPREEND. E PARTICIPACOES S.A.	Construção e Transporte	NM
FABRICA TECIDOS CARLOS RENAUX S.A.	Consumo Cíclico	-
FERROVIA CENTRO-ATLANTICA S.A.	Construção e Transporte	-
FERTILIZANTES FOSFATADOS S.A.- FOSFERTIL	Materiais Básicos	-
FERTILIZANTES HERINGER S.A.	Materiais Básicos	NM
FIBAM COMPANHIA INDUSTRIAL	Materiais Básicos	-
FIBRIA CELULOSE S.A.	Materiais Básicos	N1
FINANCEIRA ALFA S.A.- CRED FINANC E INVS	Financeiro e Outros	-
FINANSINOS S.A.- CREDITO FINANC E INVEST	Financeiro e Outros	-
FLEURY S.A.	Consumo Não Cíclico	NM
FRAS-LE S.A.	Bens Industriais	N1
GAFISA S.A.	Construção e Transporte	NM
GENERAL SHOPPING BRASIL S.A.	Financeiro e Outros	NM

Nome	Setor	Governança
GERDAU S.A.	Materiais Básicos	N1
GLOBEX UTILIDADES S.A.	Consumo Cíclico	-
GOL LINHAS AEREAS INTELIGENTES S.A.	Construção e Transporte	N2
GP INVESTMENTS, LTD.	Financeiro e Outros	-
GPC PARTICIPACOES S.A.	Materiais Básicos	-
GRAZZIOTIN S.A.	Consumo Cíclico	-
GRENDENE S.A.	Consumo Cíclico	NM
GRUCAI PARTICIPACOES S.A.	Construção e Transporte	-
GUARARAPES CONFECÇOES S.A.	Consumo Cíclico	-
GVT (HOLDING) S.A.	Telecomunicações	NM
HAGA S.A. INDUSTRIA E COMERCIO	Construção e Transporte	-
HELBOR EMPREENDIMENTOS S.A.	Construção e Transporte	NM
HOPÍ HARI S.A.	Consumo Cíclico	-
HOTEIS OTHON S.A.	Consumo Cíclico	-
HYPERMARCAS S.A.	Consumo Não Cíclico	NM
IDEIASNET S.A.	Tecnologia da Informação	NM
IGUATEMI EMPRESA DE SHOPPING CENTERS S.A.	Financeiro e Outros	NM
INDUSTRIAS J B DUARTE S.A.	Financeiro e Outros	-
INDUSTRIAS ROMI S.A.	Bens Industriais	NM
INPAR S.A.	Construção e Transporte	NM
INVESTIMENTOS BEMGE S.A.	Financeiro e Outros	-
IOCHPE MAXION S.A.	Bens Industriais	NM
ITAITINGA PARTICIPACOES S.A.	Financeiro e Outros	-
ITAU UNIBANCO HOLDING S.A.	Financeiro e Outros	N1
ITAUSA INVESTIMENTOS ITAU S.A.	Financeiro e Outros	N1
ITAUTEC S.A. - GRUPO ITAUTEC	Tecnologia da Informação	-
JBS S.A.	Consumo Não Cíclico	NM
JHSF PARTICIPACOES S.A.	Construção e Transporte	-
KARSTEN S.A.	Consumo Não Cíclico	-
KEPLER WEBER S.A.	Bens Industriais	-
KLABIN S.A.	Materiais Básicos	N1
KROTON EDUCACIONAL S.A.	Consumo Cíclico	N2
LA FONTE TELECOM S.A.	Telecomunicações	-
LARK S.A. MAQUINAS E EQUIPAMENTOS	Bens Industriais	-
LIGHT S.A.	Utilidade Pública	NM
LLX LOGISTICA S.A.	Construção e Transporte	NM
LOCALIZA RENT A CAR S.A.	Consumo Cíclico	NM
LOG-IN LOGISTICA INTERMODAL S.A.	Construção e Transporte	NM
LOJAS AMERICANAS S.A.	Consumo Cíclico	-
LOJAS HERING S.A.	Consumo Cíclico	-
LOJAS RENNER S.A.	Consumo Cíclico	NM
LPS BRASIL - CONSULTORIA DE IMOVEIS S.A.	Construção e Transporte	NM
M G POLIESTER S.A.	Materiais Básicos	-
MAGNESITA REFRACTORIOS S.A.	Materiais Básicos	NM
MAHLE-METAL LEVE S.A.	Bens Industriais	-
MANGELS INDUSTRIAL S.A.	Materiais Básicos	N1
MARCOPOLO S.A.	Bens Industriais	N2
MARFRIG ALIMENTOS S/A	Consumo Não Cíclico	NM
MARISA S.A.	Consumo Cíclico	NM
MARISOL S.A.	Consumo Cíclico	-

Nome	Setor	Governança
MEDIAL SAUDE S.A.	Consumo Não Cíclico	NM
MENDES JUNIOR ENGENHARIA S.A.	Construção e Transporte	NM
MERCANTIL BRASIL FINANC S.A. C.F.I.	Financeiro e Outros	-
METALFRIO SOLUTIONS S.A.	Bens Industriais	NM
METALGRAFICA IGUACU S.A.	Materiais Básicos	-
METALURGICA DUQUE S.A.	Materiais Básicos	-
METALURGICA GERDAU S.A.	Materiais Básicos	N1
METALURGICA RIOSULENSE S.A.	Bens Industriais	-
METISA METALURGICA TIMBOENSE S.A.	Bens Industriais	-
MILLENNIUM INORGANIC CHEMICALS BR S.A.	Materiais Básicos	-
MINERVA S.A.	Consumo Não Cíclico	NM
MINUPAR PARTICIPACOES S.A.	Consumo Não Cíclico	-
MMX MINERACAO E METALICOS S.A.	Materiais Básicos	NM
MONTEIRO ARANHA S.A.	Financeiro e Outros	-
MPX ENERGIA S.A.	Utilidade Pública	NM
MRV ENGENHARIA E PARTICIPACOES S.A.	Construção e Transporte	NM
MULTIPLAN - EMPREEND IMOBILIARIOS S.A.	Financeiro e Outros	N2
NADIR FIGUEIREDO IND E COM S.A.	Consumo Cíclico	-
NATURA COSMETICOS S.A.	Consumo Não Cíclico	NM
NEOENERGIA S.A.	Utilidade Pública	-
NET SERVICOS DE COMUNICACAO S.A.	Consumo Cíclico	N2
NORDON INDUSTRIAS METALURGICAS S.A.	Bens Industriais	-
NUTRIPLANT INDUSTRIA E COMERCIO S.A.	Materiais Básicos	-
OBRASCON HUARTE LAIN BRASIL S.A.	Construção e Transporte	-
ODONTOPREV S.A.	Consumo Não Cíclico	NM
OGX PETROLEO E GAS PARTICIPACOES S.A.	Petróleo, Gás e Combustível	NM
OSX BRASIL S.A.	Petróleo, Gás e Combustível	NM
PANATLANTICA S.A.	Materiais Básicos	-
PARANA BCO S.A.	Financeiro e Outros	N1
PARANAPANEMA S.A.	Materiais Básicos	N1
PDG REALTY S.A. EMPREEND E PARTICIPACOES	Construção e Transporte	NM
PETROLEO BRASILEIRO S.A. PETROBRAS	Petróleo, Gás e Combustível	-
PETROPAR S.A.	Materiais Básicos	-
PETTENATI S.A. INDUSTRIA TEXTIL	Consumo Cíclico	-
PLASCAR PARTICIPACOES INDUSTRIAIS S.A.	Bens Industriais	-
POLPAR S.A.	Financeiro e Outros	-
PORTO SEGURO S.A.	Financeiro e Outros	NM
PORTOBELLO S.A.	Construção e Transporte	NM
POSITIVO INFORMATICA S.A.	Tecnologia da Informação	NM
PRÓ METALURGIA S.A.	Consumo Cíclico	-
PROFARMA DISTRIB PROD FARMACEUTICOS S.A.	Consumo Não Cíclico	NM
PRONOR PETROQUIMICA S.A.	Materiais Básicos	-
QUATTOR PETROQUIMICA S.A.	Materiais Básicos	-
RANDON S.A. IMPLEMENTOS E PARTICIPACOES	Bens Industriais	N1
RASIP AGRO PASTORIL S.A.	Consumo Não Cíclico	-
RECRUSUL S.A.	Bens Industriais	-
REDE ENERGIA S.A.	Utilidade Pública	-
REDECARD S.A.	Financeiro e Outros	NM
REFINARIA DE PETROLEOS MANGUINHOS S.A.	Petróleo, Gás e Combustível	-
RENAR MACAS S.A.	Consumo Não Cíclico	NM

Nome	Setor	Governança
RENOVA ENERGIA S.A.	Não Classificado	-
RIMET EMPREEND INDS E COMERCIAIS S.A.	Materiais Básicos	-
RIO GRANDE ENERGIA S.A.	Utilidade Pública	-
RODOBENS NEGOCIOS IMOBILIARIOS S.A.	Construção e Transporte	NM
ROSSI RESIDENCIAL S.A.	Construção e Transporte	NM
SANSUY S.A. INDUSTRIA DE PLASTICOS	Materiais Básicos	-
SANTOS BRASIL PARTICIPACOES S.A.	Construção e Transporte	N2
SANTOS BRASIL S.A.	Construção e Transporte	-
SAO CARLOS EMPREEND E PARTICIPACOES S.A.	Financeiro e Outros	NM
SAO PAULO ALPARGATAS S.A.	Consumo Cíclico	N1
SAO PAULO TURISMO S.A.	Consumo Cíclico	-
SARAIVA S.A. LIVREIROS EDITORES	Consumo Cíclico	N2
SAUIPE S.A.	Consumo Cíclico	-
SCHULZ S.A.	Bens Industriais	N2
SEB - SISTEMA EDUCACIONAL BRASILEIRO S.A	Consumo Não Cíclico	-
SERGEN SERVICOS GERAIS DE ENG S.A.	Construção e Transporte	-
SIDERURGICA J. L. ALIPERTI S.A.	Materiais Básicos	-
SLC AGRICOLA S.A.	Consumo Não Cíclico	NM
SONDOTECNICA ENGENHARIA SOLOS S.A.	Construção e Transporte	-
SPRINGER S.A.	Consumo Cíclico	NM
SUL AMERICA S.A.	Financeiro e Outros	N2
SUZANO HOLDING S.A.	Materiais Básicos	-
SUZANO PAPEL E CELULOSE S.A.	Materiais Básicos	N1
TAM S.A.	Construção e Transporte	N2
TEC TOY S.A.	Consumo Cíclico	-
TECBLU TECELAGEM BLUMENAU S.A.	Consumo Cíclico	-
TECNISA S.A.	Construção e Transporte	NM
TECNOSOLO ENGENHARIA S.A.	Construção e Transporte	-
TEGMA GESTAO LOGISTICA S.A.	Construção e Transporte	NM
TEKA-TECELAGEM KUEHNRIK S.A.	Consumo Cíclico	-
TELE NORTE CELULAR PARTICIPACOES S.A.	Telecomunicações	-
TELE NORTE LESTE PARTICIPACOES S.A.	Telecomunicações	-
TELEC BRASILEIRAS S.A. TELEBRAS	Financeiro e Outros	-
TELEC DE SAO PAULO S.A. – TELESP	Telecomunicações	-
TELEMAR NORTE LESTE S.A.	Telecomunicações	-
TEMPO PARTICIPACOES S.A.	Consumo Não Cíclico	NM
TEXTIL RENAUXVIEW S.A.	Consumo Cíclico	-
TIM PARTICIPACOES S.A.	Telecomunicações	-
TIVIT TERC. DE PROC., SERV. E TEC. S.A.	Tecnologia da Informação	NM
TOTVS S.A.	Tecnologia da Informação	NM
TPI - TRIUNFO PARTICIP. E INVEST. S.A.	Construção e Transporte	NM
TRACTEBEL ENERGIA S.A.	Utilidade Pública	NM
TRANSMISSORA ALIANÇA DE ENERGIA ELÉTRICA S.A.	Utilidade Pública	N2
TREVISA INVESTIMENTOS S.A.	Construção e Transporte	-
TRISUL S.A.	Construção e Transporte	NM
TUPY S.A.	Bens Industriais	-
ULTRAPAR PARTICIPACOES S.A.	Financeiro e Outros	N1
UNIPAR UNIAO DE IND PETROQ S.A.	Materiais Básicos	N1
UNIVERSO ONLINE S.A.	Tecnologia da Informação	N2
USINAS SID DE MINAS GERAIS S.A.-USIMINAS	Materiais Básicos	N1

Continuação do Apêndice A -6/6

Nome	Setor	Governança
VALE S.A.	Materiais Básicos	N1
VBC ENERGIA S.A.	Utilidade Pública	-
VICUNHA TEXTIL S.A.	Consumo Cíclico	-
VIVO PARTICIPACOES S.A.	Telecomunicações	-
VULCABRAS S.A.	Consumo Cíclico	-
WEG S.A.	Bens Industriais	NM
WEMBLEY SOCIEDADE ANONIMA	Consumo Cíclico	-
WETZEL S.A.	Bens Industriais	-
WHIRLPOOL S.A.	Consumo Cíclico	-
WLM - INDUSTRIA E COMERCIO S.A.	Bens Industriais	-
YARA BRASIL FERTILIZANTES S.A.	Materiais Básicos	-

Fonte: Elaboração Própria

Quadro A1: Empresas pertencentes à Amostra Final: N1 = Nível 1; N2 = Nível 2; NM = Novo Mercado

APÊNDICE B – Resultado do Teste t (teste de média)

Tabela B1: Resultado do Teste t

Pares de Amostra	Média	Desvio-Padrão	T	df	Sig. (2-tailed)
DifTotais100 - DifTotais90	0,04	0,59	1,26	311	0,21
DifLinhas100 - DifLinhas90	1,10	18,42	1,06	311	0,29
DifPalavras100 - DifPalav90	16,94	172,99	1,73	311	0,08
DifTotais100 - DifTotais60	0,08	1,84	0,77	311	0,44
DifLinhas100 - DifLinhas60	2,11	45,03	0,83	311	0,41
DifPalavras100 - DifPalav60	37,20	279,26	2,35	311	0,02
DifTotais100 - DifTotais50	-0,05	2,68	-0,34	311	0,74
DifLinhas100 - DifLinha50	-2,37	43,07	-0,97	311	0,33
DifPalavras100 - DifPalav50	22,45	303,21	1,31	311	0,19
DifTotais90 - DifTotais60	0,04	1,74	0,39	311	0,70
DifLinhas90 - DifLinhas60	1,01	41,12	0,43	311	0,66
DifPalav90 - DifPalav60	20,25	197,38	1,81	311	0,07
DifLinhas90 - DifLinha50	-3,47	38,83	-1,58	311	0,12
DifPalav90 - DifPalav50	5,46	233,19	0,41	311	0,68
DifTotais90 - DifTotais50	-0,09	2,62	-0,63	311	0,53
DifTotais60 - DifTotais50	-0,13	2,70	-0,86	311	0,39
DifLinhas60 - DifLinha50	-4,48	42,72	-1,85	311	0,07
DifPalav60 - DifPalav50	-14,83	211,86	-1,23	311	0,22

Onde:

DifTotais = Diferenças Totais; DifLinhas = Diferenças de Linhas; DifPalavras = Diferenças Intra-Linhas, para os graus de Similaridade 100, 90, 60 e 50.

APÊNDICE C – Matrizes de Correlação

Tabela C1: Resultados para a Correlação de *Pearson* (grau 100)

	DT/ Carac	DL/ Carac	DIL/ Carac	DT/ Palav	DL/ Palav	DIL/ Palav	DT/ Sentença	DL/ Sentença	DIL/ Sentença
DT/Caracteres	1,00	0,13	0,05	0,61	0,13	0,06	0,91	0,04	-0,10
		0,02	0,36	0,00	0,02	0,30	0,00	0,45	0,08
DL/Caracteres	0,13	1,00	0,76	0,09	0,89	0,56	0,04	0,88	0,45
	0,02		0,00	0,10	0,00	0,00	0,43	0,00	0,00
DIL/Caracteres	0,05	0,76	1,00	0,03	0,67	0,70	0,06	0,74	0,88
	0,36	0,00		0,66	0,00	0,00	0,28	0,00	0,00
DT/Palavras	0,61	0,09	0,03	1,00	0,45	0,60	0,59	0,06	-0,05
	0,00	0,10	0,66		0,00	0,00	0,00	0,30	0,42
DL/Palavras	0,13	0,89	0,67	0,45	1,00	0,82	0,08	0,78	0,40
	0,02	0,00	0,00	0,00		0,00	0,18	0,00	0,00
DIL/Palavras	0,06	0,56	0,70	0,60	0,82	1,00	0,10	0,55	0,62
	0,30	0,00	0,00	0,00	0,00		0,08	0,00	0,00
DT/Sentença	0,91	0,04	0,06	0,59	0,08	0,10	1,00	0,04	0,02
	0,00	0,43	0,28	0,00	0,18	0,08		0,43	0,69
DL/Sentença	0,04	0,88	0,74	0,06	0,78	0,55	0,04	1,00	0,61
	0,45	0,00	0,00	0,30	0,00	0,00	0,43		0,00
DIL/Sentença	-0,10	0,45	0,88	-0,05	0,40	0,62	0,02	0,61	1,00
	0,08	0,00	0,00	0,42	0,00	0,00	0,69	0,00	

Onde:

¹ DT/Caracteres = Diferenças Totais/Caracteres; DL/Caracteres = Diferenças de Linhas/Caracteres; DIL/Caracteres = Diferenças Intra-Linha/Caracteres

² DT/Palavras = Diferenças Totais/Palavras; DL/Palavras = Diferenças de Linhas/Palavras; DIL/Palavras = Diferenças Intra-Linha/Palavras

³ DT/Sentenças = Diferenças Totais/Sentenças; DL/Sentenças = Diferença de Linhas/Sentenças; DIL/Sentenças = Diferença Intra-Linha/Sentenças

Tabela C2: Resultados para a Correlação de Spearman (grau 100)

	DT/ Carac	DL/ Carac	DIL/ Carac	DT/ alav	DL/ Palav	DIL/ Palav	DT/ Sentença	DL/ Sentença	DIL/ Sentença
DT/Carac	1,00	-0,09	-0,01	0,99	-0,08	0,01	0,95	-0,24	-0,14
	.	0,11	0,89	0,00	0,17	0,86	0,00	0,00	0,01
DL/Carac	-0,09	1,00	0,19	-0,09	0,99	0,19	-0,13	0,93	0,12
	0,11	.	0,00	0,12	0,00	0,00	0,03	0,00	0,03
DIL/Carac	-0,01	0,19	1,00	-0,02	0,18	0,99	0,02	0,22	0,92
	0,89	0,00	.	0,79	0,00	0,00	0,78	0,00	0,00
DT/Palav	0,99	-0,09	-0,02	1,00	-0,07	0,02	0,95	-0,23	-0,14
	0,00	0,12	0,79	.	0,24	0,76	0,00	0,00	0,01
DL/Palav	-0,08	0,99	0,18	-0,07	1,00	0,19	-0,11	0,93	0,12
	0,17	0,00	0,00	0,24	.	0,00	0,05	0,00	0,03
DIL/Palav	0,01	0,19	0,99	0,02	0,19	1,00	0,04	0,23	0,92
	0,86	0,00	0,00	0,76	0,00	.	0,45	0,00	0,00
DT/Sentença	0,95	-0,13	0,02	0,95	-0,11	0,04	1,00	-0,21	-0,04
	0,00	0,03	0,78	0,00	0,05	0,45	.	0,00	0,50
DL/Sentença	-0,24	0,93	0,22	-0,23	0,93	0,23	-0,21	1,00	0,26
	0,00	0,00	0,00	0,00	0,00	0,00	0,00	.	0,00
DIL/Sentença	-0,14	0,12	0,92	-0,14	0,12	0,92	-0,04	0,26	1,00
	0,01	0,03	0,00	0,01	0,03	0,00	0,50	0,00	.

Onde:

¹ DT/Carac = Diferenças Totais/Caracteres; DL/Carac = Diferenças de Linhas/Caracteres; DIL/Carac = Diferenças Intra-Linha/Caracteres

² DT/Palav = Diferenças Totais/Palavras; DL/Palav = Diferenças de Linhas/Palavras; DIL/Palav = Diferenças Intra-Linha/Palavras

³ DT/Sentenças = Diferenças Totais/Sentenças; DL/Sentenças = Diferença de Linhas/Sentenças; DIL/Sentenças = Diferença Intra-Linha/Sentenças

Tabela C3: Resultados para a Correlação de *Pearson* (grau 60)

	DT/ Carac	DL/ Carac	DIL/ Carac	DT/ Palav	DL/ Palav	DIL/ Palav	DT/ Sentença	DL/ Sentença	DIL/ Sentença
DT/Carac	1,00	0,14	0,09	0,61	0,13	0,09	0,91	0,05	-0,04
		0,02	0,11	0,00	0,02	0,13	0,00	0,40	0,51
DL/Carac	0,14	1,00	0,76	0,10	0,89	0,57	0,05	0,88	0,48
	0,02		0,00	0,09	0,00	0,00	0,40	0,00	0,00
DIL/Carac	0,09	0,76	1,00	0,05	0,66	0,72	0,11	0,77	0,90
	0,11	0,00		0,38	0,00	0,00	0,06	0,00	0,00
DT/Palav	0,61	0,10	0,05	1,00	0,45	0,60	0,60	0,06	-0,01
	0,00	0,09	0,38		0,00	0,00	0,00	0,27	0,88
DL/Palav	0,13	0,89	0,66	0,45	1,00	0,82	0,08	0,78	0,42
	0,02	0,00	0,00	0,00		0,00	0,17	0,00	0,00
DIL/Palav	0,09	0,57	0,72	0,60	0,82	1,00	0,13	0,59	0,65
	0,13	0,00	0,00	0,00	0,00		0,02	0,00	0,00
DT/Sentença	0,91	0,05	0,11	0,60	0,08	0,13	1,00	0,05	0,08
	0,00	0,40	0,06	0,00	0,17	0,02		0,38	0,16
DL/Sentença	0,05	0,88	0,77	0,06	0,78	0,59	0,05	1,00	0,66
	0,40	0,00	0,00	0,27	0,00	0,00	0,38		0,00
DIL/Sentença	-0,04	0,48	0,90	-0,01	0,42	0,65	0,08	0,66	1,00
	0,51	0,00	0,00	0,88	0,00	0,00	0,16	0,00	

Onde:

¹ DT/Carac = Diferenças Totais/Caracteres; DL/Carac = Diferenças de Linhas/Caracteres; DIL/Carac = Diferenças Intra-Linha/Caracteres

² DT/Palav = Diferenças Totais/Palavras; DL/Palav = Diferenças de Linhas/Palavras; DIL/Palav = Diferenças Intra-Linha/Palavras

³ DT/Sentenças = Diferenças Totais/Sentenças; DL/Sentenças = Diferença de Linhas/Sentenças; DIL/Sentenças = Diferença Intra-Linha/Sentenças

Tabela C4: Resultados para a Correlação de Spearman (grau 60)

	DT/ Carac	DL/ Carac	DIL/ Carac	DT/ Palav	DL/ Palav	DIL/ Palav	DT/ Sentença	DL/ Sentença	DIL/ Sentença
DT/Carac	1,00	-0,07	-0,01	0,99	-0,06	0,00	0,95	-0,22	-0,15
	.	0,21	0,88	0,00	0,30	0,97	0,00	0,00	0,01
DL/Carac	-0,07	1,00	0,25	-0,07	0,99	0,25	-0,11	0,94	0,18
	0,21	.	0,00	0,22	0,00	0,00	0,06	0,00	0,00
DIL/Carac	-0,01	0,25	1,00	-0,02	0,23	0,98	0,01	0,27	0,92
	0,88	0,00	.	0,77	0,00	0,00	0,92	0,00	0,00
DT/Palav	0,99	-0,07	-0,02	1,00	-0,05	0,01	0,95	-0,21	-0,15
	0,00	0,22	0,77	.	0,39	0,88	0,00	0,00	0,01
DL/Palav	-0,06	0,99	0,23	-0,05	1,00	0,25	-0,09	0,93	0,17
	0,30	0,00	0,00	0,39	.	0,00	0,11	0,00	0,00
DIL/Palav	0,00	0,25	0,98	0,01	0,25	1,00	0,03	0,27	0,92
	0,97	0,00	0,00	0,88	0,00	.	0,65	0,00	0,00
DT/Sentença	0,95	-0,11	0,01	0,95	-0,09	0,03	1,00	-0,19	-0,05
	0,00	0,06	0,92	0,00	0,11	0,65	.	0,00	0,35
DL/Sentença	-0,22	0,94	0,27	-0,21	0,93	0,27	-0,19	1,00	0,30
	0,00	0,00	0,00	0,00	0,00	0,00	0,00	.	0,00
DIL/Sentença	-0,15	0,18	0,92	-0,15	0,17	0,92	-0,05	0,30	1,00
	0,01	0,00	0,00	0,01	0,00	0,00	0,35	0,00	.

Onde:

¹ DT/Carac = Diferenças Totais/Caracteres; DL/Carac = Diferenças de Linhas/Caracteres;

DIL/Carac = Diferenças Intra-Linha/Caracteres

² DT/Palav = Diferenças Totais/Palavras; DL/Palav = Diferenças de Linhas/Palavras;

DIL/Palav = Diferenças Intra-Linha/Palavras

³ DT/Sentenças = Diferenças Totais/Sentenças; DL/Sentenças = Diferença de Linhas/Sentenças;

DIL/Sentenças = Diferença Intra-Linha/Sentenças

APÊNDICE D – Tabela de Frequências das Estatísticas de Linhas Adicionadas e Deletadas

Tabela D1 – Frequências de Linhas Adicionadas e Deletadas

Grau 60								Grau 100							
LADIC				LDEL				LADIC				LDEL			
	F	%	%Acum		F	%	%Acum		F	%	%Acum		F	%	%Acum
0	230	74,0	74,0	0	223	71,7	71,7	0	229	73,6	73,6	0	224	72,0	72,0
1	28	9,0	83,0	1	32	10,3	82,0	1	28	9,0	82,6	1	31	10,0	82,0
2	2	0,6	83,6	2	2	0,6	82,6	2	2	0,6	83,3	2	2	0,6	82,6
3	6	1,9	85,5	3	9	2,9	85,5	3	6	1,9	85,2	3	9	2,9	85,5
5	5	1,6	87,1	5	6	1,9	87,5	5	5	1,6	86,8	5	6	1,9	87,5
6	3	1,0	88,1	6	2	0,6	88,1	6	3	1,0	87,8	6	2	0,6	88,1
7	3	1,0	89,1	7	4	1,3	89,4	7	3	1,0	88,7	7	4	1,3	89,4
8	1	0,3	89,4	8	1	0,3	89,7	8	1	0,3	89,1	8	1	0,3	89,7
9	2	0,6	90,0	9	2	0,6	90,4	9	2	0,6	89,7	9	2	0,6	90,4
11	3	1,0	91,0	10	2	0,6	91,0	11	3	1,0	90,7	10	2	0,6	91,0
14	1	0,3	91,3	11	1	0,3	91,3	13	1	0,3	91,0	11	1	0,3	91,3
17	1	0,3	91,6	12	1	0,3	91,6	14	1	0,3	91,3	12	1	0,3	91,6
18	2	0,6	92,3	13	2	0,6	92,3	17	1	0,3	91,6	13	2	0,6	92,3
19	1	0,3	92,6	14	1	0,3	92,6	18	2	0,6	92,3	14	1	0,3	92,6
22	1	0,3	92,9	15	3	1,0	93,6	19	1	0,3	92,6	15	3	1,0	93,6
25	2	0,6	93,6	17	1	0,3	93,9	22	1	0,3	92,9	17	1	0,3	93,9
27	1	0,3	93,9	19	1	0,3	94,2	25	2	0,6	93,6	19	1	0,3	94,2
30	1	0,3	94,2	21	1	0,3	94,5	27	1	0,3	93,9	21	1	0,3	94,5
35	2	0,6	94,9	24	2	0,6	95,2	30	1	0,3	94,2	24	2	0,6	95,2
41	1	0,3	95,2	27	1	0,3	95,5	35	2	0,6	94,9	27	1	0,3	95,5
43	1	0,3	95,5	29	1	0,3	95,8	41	1	0,3	95,2	29	1	0,3	95,8
48	1	0,3	95,8	32	2	0,6	96,5	43	1	0,3	95,5	32	2	0,6	96,5
50	2	0,6	96,5	34	2	0,6	97,1	48	1	0,3	95,8	34	2	0,6	97,1
51	1	0,3	96,8	39	2	0,6	97,7	50	2	0,6	96,5	39	2	0,6	97,7
52	1	0,3	97,1	42	1	0,3	98,1	51	1	0,3	96,8	42	1	0,3	98,1
55	1	0,3	97,4	43	1	0,3	98,4	52	1	0,3	97,1	43	1	0,3	98,4
71	1	0,3	97,7	51	2	0,6	99,0	55	1	0,3	97,4	51	2	0,6	99,0
85	1	0,3	98,1	67	1	0,3	99,4	71	1	0,3	97,7	67	1	0,3	99,4
117	1	0,3	98,4	74	1	0,3	99,7	85	1	0,3	98,1	74	1	0,3	99,7
147	1	0,3	98,7	81	1	0,3	100,0	117	1	0,3	98,4	81	1	0,3	100,0
154	1	0,3	99,0	Tot al	31 1	100, 0		147	1	0,3	98,7	Tota l	31 1	100, 0	
171	1	0,3	99,4					154	1	0,3	99,0				
307	1	0,3	99,7					171	1	0,3	99,4				
465	1	0,3	100,0					307	1	0,3	99,7				
		100,0						465	1	0,3	100,0				

Onde: LADIC = Linhas Adicionadas; LDEL = Linhas Deletadas

APÊNDICE E – Resultados do Testes de Normalidade

Tabela E1 – Resultado do Teste de *Kolmogorov-Smirnov*, por total 60

	DILSent	SUtilPubl	SFINANC	SBENSIND	SCONST	SMATBAS
N	311,00	311,00	311,00	311,00	311,00	311,00
Kolmogorov-Smirnov Z	2,86	9,08	8,98	9,40	9,03	9,14
Asymp. Sig. (2-tailed)	0,00	0,00	0,00	0,00	0,00	0,00
	SConsCic	SCONSNC	STI	STELEC	SPETR	SNC
N	311,00	311,00	311,00	311,00	311,00	311,00
Kolmogorov-Smirnov Z	9,08	9,35	9,48	9,52	9,43	9,27
Asymp. Sig. (2-tailed)	0,00	0,00	0,00	0,00	0,00	0,00
	SREG	Governança	PORTE	LC	IMPRESA	ADR
N	311,00	311,00	311,00	311,00	311,00	311,00
Kolmogorov-Smirnov Z	7,37	6,55	1,69	7,98	7,41	9,45
Asymp. Sig. (2-tailed)	0,00	0,00	0,01	0,00	0,00	0,00
	CONCENT	AUDIT	LEGIB	CARACT	DummyVariarRelat	DummyVariarPL
N	308,00	311,00	311,00	311,00	311,00	311,00
Kolmogorov-Smirnov Z	1,78	7,22	1,39	3,23	6,41	8,27
Asymp. Sig. (2-tailed)	0,00	0,00	0,04	0,00	0,00	0,00
	DummyVarLL	LIQMERC	INDEPENDENCIA			
N	311,00	311,00	311,00			
Kolmogorov-Smirnov Z	7,05	6,95	8,87			
Asymp. Sig. (2-tailed)	0,00	0,00	0,00			

Onde:

DILSENT = Diferenças Intra-Linha/Sentença; SUP = Setor de Utilidade Pública; SFINANC = Setor Financeiro e Outros
SBENSIND = Setor de Bens Industriais; SCONST = Setor de Construção e Transportes; SMATBAS = Setor de Materiais Básicos

SREG = Setor Regulado; LC = Liquidez Corrente; IMPRESA = Imprensa; ADR = American Depositary Receipts;
CONCENT = Concentração; AUDIT = Auditoria; LEGIB = Legibilidade; PORTE = Porte; SConsCic = Setor de Consumo Cíclico

SConsNC = Setor de Consumo Não Cíclico; STI = Setor de Tecnologia da Informação; SPETR = Setor de Petróleo e Gás;
SNC = Setor Não Classificado; GC = Governança Corporativa; CARACT = Caracteres; DVR = Dummy Variação do RA;
DVPL = Dummy Variação do PL; DVLL = Dummy Variação do LL; LIQMERC = Liquidez de Mercado; IF = Independência Financeira

Tabela E2 – Resultado do Teste de Kolmogorov-Smirnov, por total 100

	DILSENT	SUtilPubl	SFINANC	SBENSIND	SCONST	SMATBAS
N	311,00	311,00	311,00	311,00	311,00	311,00
Kolmogorov-Smirnov Z	2,40	9,08	8,98	9,40	9,03	9,14
Asymp. Sig. (2-tailed)	0,00	0,00	0,00	0,00	0,00	0,00
	SConsCic	SCONSNC	STI	STELEC	SPETR	SNC
N	311,00	311,00	311,00	311,00	311,00	311,00
Kolmogorov-Smirnov Z	9,08	9,35	9,48	9,52	9,43	9,27
Asymp. Sig. (2-tailed)	0,00	0,00	0,00	0,00	0,00	0,00
	SREG	Governança	PORTE	LC	IMPRESA	ADR
N	311,00	311,00	311,00	311,00	311,00	311,00
Kolmogorov-Smirnov Z	7,37	6,55	1,69	7,98	7,41	9,45
Asymp. Sig. (2-tailed)	0,00	0,00	0,01	0,00	0,00	0,00
	CONCENT	AUDIT	LEGIB	CARACT	DummyVa rRelat	DummyVarP L
N	308,00	311,00	311,00	311,00	311,00	311,00
Kolmogorov-Smirnov Z	1,78	7,22	1,39	3,23	6,41	8,27
Asymp. Sig. (2-tailed)	0,00	0,00	0,04	0,00	0,00	0,00
	DummyVarLL	LIQMERC	INDFINANC			
N	311,00	311,00	311,00			
Kolmogorov-Smirnov Z	7,05	6,95	8,87			
Asymp. Sig. (2-tailed)	0,00	0,00	0,00			

Onde:

DILSENT = Diferenças Intra-Linha/Sentença; SUP = Setor de Utilidade Pública; SFINANC = Setor Financeiro e Outros
SBENSIND = Setor de Bens Industriais; SCONST = Setor de Construção e Transportes; SMATBAS = Setor de Materiais Básicos

SREG = Setor Regulado; LC = Liquidez Corrente; IMPRESA = Imprensa; ADR = American Depositary Receipts;
CONCENT = Concentração; AUDIT = Auditoria; LEGIB = Legibilidade; PORTE = Porte; SConsCic = Setor de Consumo Cíclico

SConsNC = Setor de Consumo Não Cíclico; STI = Setor de Tecnologia da Informação; SPETR = Setor de Petróleo e Gás;
SNC = Setor Não Classificado; GC = Governança Corporativa; CARACT = Caracteres; DVR = Dummy Variação do RA;
DVPL = Dummy Variação do PL; DVLL = Dummy Variação do LL; LIQMERC = Liquidez de Mercado; IF = Independência Financeira

Tabela E3 – Resultado do Teste de Kolmogorov-Smirnov para primeiro tercil 100

	DIL100S	SUP	SFINANC	SBENSIND	SCONST	SMATBAS
N	104,00	104,00	104,00	104,00	104,00	104,00
Kolmogorov-Smirnov Z	1,13	5,42	5,12	5,44	5,21	5,39
Asymp. Sig. (2-tailed)	0,15	0,00	0,00	0,00	0,00	0,00
	SCONSC	SCONNC	STI	STELEC	SPETR	SNC
N	104,00	104,00	104,00	104,00	104,00	104,00
Kolmogorov-Smirnov Z	4,99	5,39	5,47	5,51	5,40	
Asymp. Sig. (2-tailed)	0,00	0,00	0,00	0,00	0,00	
	SREG	GC	PORTE	LC	IMPrensa	ADR
N	104,00	104,00	104,00	104,00	104,00	104,00
Kolmogorov-Smirnov Z	4,38	4,33	1,49	3,57	4,18	5,51
Asymp. Sig. (2-tailed)	0,00	0,00	0,02	0,00	0,00	0,00
	CONCENT	AUDIT	LEGIB	DVRELAT	VPL	DVLL
N	102,00	104,00	104,00	104,00	104,00	104,00
Kolmogorov-Smirnov Z	1,00	3,72	0,68	4,08	4,47	4,08
Asymp. Sig. (2-tailed)	0,27	0,00	0,74	0,00	0,00	0,00
	DVREC	LIQMERC	BETA	IF		
N	68,00	104,00	35,00	104,00		
Kolmogorov-Smirnov Z	2,99	4,05	0,51	5,10		
Asymp. Sig. (2-tailed)	0,00	0,00	0,96	0,00		

Onde:

DILSENT = Diferenças Intra-Linha/Sentença; SUP = Setor de Utilidade Pública; SFINANC = Setor Financeiro e Outros
SBENSIND = Setor de Bens Industriais; SCONST = Setor de Construção e Transportes; SMATBAS = Setor de Materiais Básicos

SREG = Setor Regulado; LC = Liquidez Corrente; IMPrensa = Imprensa; ADR = American Depositary Receipts;
CONCENT = Concentração; AUDIT = Auditoria; LEGIB = Legibilidade; PORTE = Porte; SConsCic = Setor de Consumo Cíclico

SConsNC = Setor de Consumo Não Cíclico; STI = Setor de Tecnologia da Informação; SPETR = Setor de Petróleo e Gás;
SNC = Setor Não Classificado; GC = Governança Corporativa; CARACT = Caracteres; DVR = Dummy Variação do RA;
DVPL = Dummy Variação do PL; DVLL = Dummy Variação do LL; LIQMERC = Liquidez de Mercado; IF = Independência Financeira

Tabela E4 – Resultado do Teste de *Kolmogorov-Smirnov* para terceiro tercil 100

	DIL100S	SUP	SFINANC	SBENSIND	SCONST	SMATBAS
N	103,00	103,00	103,00	103,00	103,00	103,00
Kolmogorov-Smirnov Z	2,81	5,32	5,17	5,39	5,01	5,05
Asymp. Sig. (2-tailed)	0,00	0,00	0,00	0,00	0,00	0,00
	SCONSC	SCONNC	STI	STELEC	SPETR	SNC
N	103,00	103,00	103,00	103,00	103,00	103,00
Kolmogorov-Smirnov Z	5,48	5,36	5,48	5,44	5,47	
Asymp. Sig. (2-tailed)	0,00	0,00	0,00	0,00	0,00	
	SREG	GC	PORTE	LC	IMPRESA	ADR
N	103,00	103,00	103,00	103,00	103,00	103,00
Kolmogorov-Smirnov Z	4,39	3,73	0,88	4,90	3,78	5,32
Asymp. Sig. (2-tailed)	0,00	0,00	0,42	0,00	0,00	0,00
	CONCEN	AUDIT	LEGIB	DVRELAT	VPL	DVLL
N	103,00	103,00	103,00	103,00	103,00	103,00
Kolmogorov-Smirnov Z	1,36	4,39	1,02	4,63	5,17	4,14
Asymp. Sig. (2-tailed)	0,05	0,00	0,25	0,00	0,00	0,00
	DVREC	LIQUMER	BETA	IF		
N	69,00	103,00	45,00	103,00		
Kolmogorov-Smirnov Z	3,10	3,90	1,18	3,19		
Asymp. Sig. (2-tailed)	0,00	0,00	0,12	0,00		

Onde:

DILSENT = Diferenças Intra-Linha/Sentença; SUP = Setor de Utilidade Pública; SFINANC = Setor Financeiro e Outros; SBENSIND = Setor de Bens Industriais; SCONST = Setor de Construção e Transportes; SMATBAS = Setor de Materiais Básicos

SREG = Setor Regulado; LC = Liquidez Corrente; IMPRESA = Imprensa; ADR = American Depositary Receipts; CONCEN = Concentração; AUDIT = Auditoria; LEGIB = Legibilidade; PORTE = Porte; SConsCic = Setor de Consumo Cíclico

SConsNC = Setor de Consumo Não Cíclico; STI = Setor de Tecnologia da Informação; SPETR = Setor de Petróleo e Gás; SNC = Setor Não Classificado; GC = Governança Corporativa; CARACT = Caracteres; DVR = Dummy Variação do RA; DVPL = Dummy Variação do PL; DVLL = Dummy Variação do LL; LIQMERC = Liquidez de Mercado; IF = Independência Financeira

Tabela E5 – Resultado do Teste de Kolmogorov-Smirnov para terceiro tercil 60

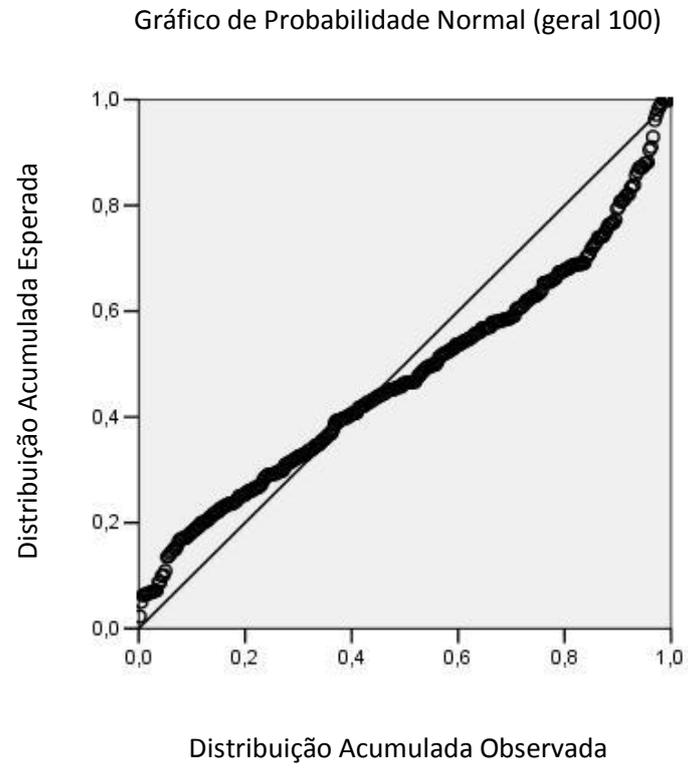
	<i>DIL60</i>	<i>SUP</i>	<i>SFINANC</i>	<i>SBENSIND</i>	<i>SCONST</i>	<i>SMATBAS</i>
N	104,00	104,00	104,00	104,00	104,00	104,00
Kolmogorov-Smirnov Z	3,18	5,35	5,21	5,49	5,08	5,04
Asymp. Sig. (2-tailed)	0,00	0,00	0,00	0,00	0,00	0,00
	<i>SCONSC</i>	<i>SCONNC</i>	<i>STI</i>	<i>STELEC</i>	<i>SPETR</i>	<i>SNC</i>
N	104,00	104,00	104,00	104,00	104,00	104,00
Kolmogorov-Smirnov Z	5,44	5,42	5,51	5,50	5,50	
Asymp. Sig. (2-tailed)	0,00	0,00	0,00	0,00	0,00	
	<i>SREG</i>	<i>GC</i>	<i>PORTE</i>	<i>LC</i>	<i>IMPREN</i>	<i>ADR</i>
					<i>SA</i>	
N	104,00	104,00	104,00	104,00	104,00	104,00
Kolmogorov-Smirnov Z	4,38	3,57	0,78	3,48	4,22	5,42
Asymp. Sig. (2-tailed)	0,00	0,00	0,58	0,00	0,00	0,00
	<i>CONCENT</i>	<i>AUDIT</i>	<i>LEGIB</i>	<i>DVRELAT</i>	<i>VPL</i>	<i>DVLL</i>
N	104,00	104,00	104,00	104,00	104,00	104,00
Kolmogorov-Smirnov Z	1,26	4,13	1,09	4,81	4,95	4,08
Asymp. Sig. (2-tailed)	0,08	0,00	0,19	0,00	0,00	0,00
	<i>DVREC</i>	<i>LIQUMER</i>	<i>BETA</i>	<i>IF</i>		
		<i>C</i>				
N	66,00	104,00	46,00	104,00		
Kolmogorov-Smirnov Z	2,89	4,09	0,95	2,38		
Asymp. Sig. (2-tailed)	0,00	0,00	0,33	0,00		

Onde:

DILSENT = Diferenças Intra-Linha/Sentença; *SUP* = Setor de Utilidade Pública; *SFINANC* = Setor Financeiro e Outros; *SBENSIND* = Setor de Bens Industriais; *SCONST* = Setor de Construção e Transportes; *SMATBAS* = Setor de Materiais Básicos

SREG = Setor Regulado; *LC* = Liquidez Corrente; *IMPRESA* = Imprensa; *ADR* = American Depositary Receipts; *CONCENT* = Concentração; *AUDIT* = Auditoria; *LEGIB* = Legibilidade; *PORTE* = Porte; *SConsCic* = Setor de Consumo Cíclico

SConsNC = Setor de Consumo Não Cíclico; *STI* = Setor de Tecnologia da Informação; *SPETR* = Setor de Petróleo e Gás; *SNC* = Setor Não Classificado; *GC* = Governança Corporativa; *CARACT* = Caracteres; *DVR* = Dummy Variação do RA; *DVPL* = Dummy Variação do PL; *DVLL* = Dummy Variação do LL; *LIQMERC* = Liquidez de Mercado; *IF* = Independência Financeira

APÊNDICE F – Gráficos de Probabilidade Normal**Figura 1: Gráfico da Probabilidade Normal (por total 100)**

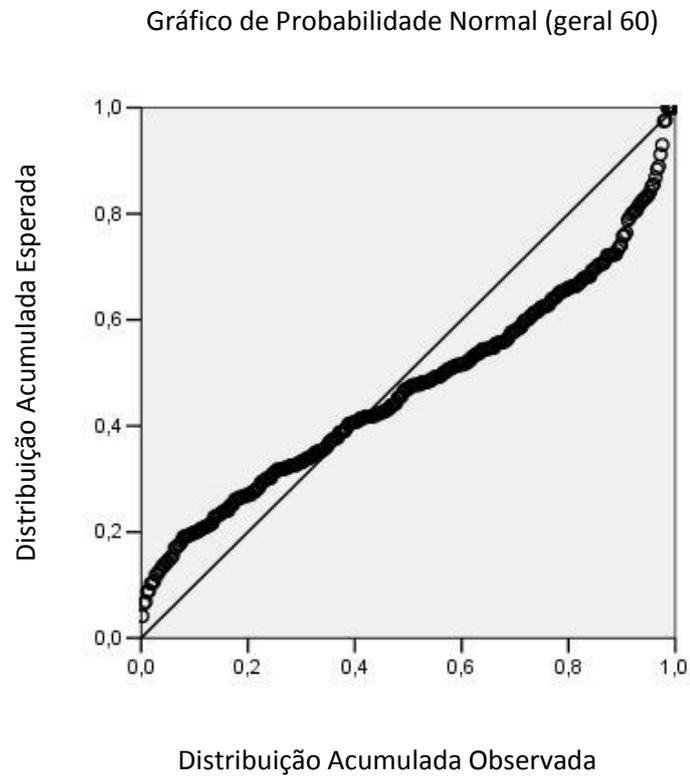


Figura 1: Gráfico da Probabilidade Normal (por total 60)

LISTA DE ANEXOS

ANEXO A – RA DA COMPANHIA SEGEN S.A. EM 31/12/2008	92
ANEXO B – RA DA COMPANHIA SEGEN S.A. EM 31/12/2009	93
ANEXO C – RA DA COMPANHIA ABYARA PLANEJ. IMOBILIÁRIOS S.A. EM 31/12/2009	94
ANEXO D – RA DA COMPANHIA ABYARA PLANEJ. IMOBILIÁRIOS S.A. EM 31/12/2008	95

ANEXO A – Relatório da administração Companhia Sergen S.A em 31.12.2008**RELATÓRIO DA ADMINISTRAÇÃO**

Senhores Acionistas

A Direção da SERGEN - Serviços Gerais de Engenharia S/A de acordo com a legislação em vigor, submete à apreciação de V.Sas. as Demonstrações Financeiras relativas ao Exercício Social findo em 31/12/**2008**

DESEMPENHO OPERACIONAL

Considerando a prática dos atrasos de pagamento dos Governos, a Diretoria da SERGEN Serviços Gerais de Engenharia S/A decidiu reduzir as suas atividades para iniciar novas obras, dirigindo seus esforços no sentido de agilizar os diversos processos de cobranças judiciais com correção **monetária** e juros dos serviços realizados e que não foram pagos de acordo com os respectivos contratos .

MERCADO DE CAPITAIS

As ações da SERGEN são negociadas na Bolsa de Valores de São Paulo.

PERSPECTIVAS

Considerando as decisões favoráveis à SERGEN em diversos processos em andamento, esperamos o recebimento efetivo dos valores devidos.

Relacionamento com Auditores

Cabe informar, que A Moreira & Associados Auditores , em conformidade com o estabelecido na Legislação societária vigente, não prestou nenhum serviço além de auditoria externa .

.

AGRADECIMENTOS

Agradecemos aos funcionários, clientes, fornecedores e bancos, o valioso apoio recebido.

Rio de Janeiro, 28 de Março de **2009**

A ADMINISTRAÇÃO

ANEXO B – Relatório da Administração Companhia Sergen S.A em 31.12.2009

RELATÓRIO DA ADMINISTRAÇÃO

Senhores Acionistas

A Direção da SERGEN - Serviços Gerais de Engenharia S/A de acordo com a legislação em vigor, submete à apreciação de V.Sas. as Demonstrações Financeiras relativas ao Exercício Social findo em 31/12/2009

DESEMPENHO OPERACIONAL

Considerando a prática dos atrasos de pagamento dos Governos, a Diretoria da SERGEN Serviços Gerais de Engenharia S/A decidiu reduzir as suas atividades para iniciar novas obras, dirigindo seus esforços no sentido de agilizar os diversos processos de cobranças judiciais com correção monetária e juros dos serviços realizados e que não foram pagos de acordo com os respectivos contratos .

MERCADO DE CAPITAIS

As ações da SERGEN são negociadas na Bolsa de Valores de São Paulo.

PERSPECTIVAS

Considerando as decisões favoráveis à SERGEN em diversos processos em andamento, esperamos o recebimento efetivo dos valores devidos.

Relacionamento com Auditores

Cabe informar, que A Moreira & Associados Auditores , em conformidade com o estabelecido na Legislação societária vigente, não prestou nenhum serviço além de auditoria externa .

.

AGRADECIMENTOS

Agradecemos aos funcionários, clientes, fornecedores e bancos, o valioso apoio recebido.

Rio de Janeiro, 28 de Março de 2010

A ADMINISTRAÇÃO

ANEXO C – Relatório da Administração da Companhia Abyara Planejamentos Imobiliários em 31.12. 2009

Atendendo aos dispositivos legais, estatutários e à regulamentação do mercado de valores mobiliários, submetemos à apreciação de V. Sas. as demonstrações financeiras devidamente auditadas, relativas aos exercícios findos em 31 de dezembro de 2009 e de 2008, da Abyara Planejamento Imobiliário S.A..

Relatório da Administração

Em 2009 a Abyara se associou a Agra e a Klabin Segall para formar uma das maiores companhias do setor imobiliário, a Agre.

A Agre é resultado da combinação das melhores práticas individuais das três companhias fortalecida pela diversificação geográfica da complementaridade dos landbanks.

A companhia já nasceu entre as empresas líderes e mais líquidas do setor, com patrimônio líquido de R\$ 1,8 bi, valor de mercado de R\$ 2,3bi, 33.000 clientes ativos e um landbank de R\$ 18,3bi em VGV potencial distribuído em 16 estados e suficiente para assegurar os lançamentos dos próximos anos.

A partir do segundo semestre a velocidade de lançamentos e vendas foi retomada atingindo R\$1,2bi em lançamentos e 1,6bi em vendas no ano, com velocidade de vendas (VSO) de 47% em 2009.

Relacionamento com Auditores

De acordo com a instrução CVM nº 381, de 14 de janeiro de 2003, informamos que a Companhia não contraiu outros serviços junto ao auditor independente responsável pelo exame das demonstrações financeiras que não sejam relativos a trabalhos de auditoria.

Companhia está vinculada à arbitragem na Câmara de Arbitragem do Mercado, conforme cláusula compromissória constante do seu Estatuto Social.

ANEXO D – Relatório da Administração da Companhia Abyara Planejamentos

Imobiliários em 31.12. 2008

ABYARA PLANEJAMENTO IMOBILIÁRIO S.A. E CONTROLADAS RELATÓRIO DA ADMINISTRAÇÃO

Arbitragem. A Companhia está vinculada à arbitragem na Câmara de Arbitragem do Mercado, conforme cláusula compromissória constante do seu Estatuto Social.

Atendendo as disposições legais e estatutárias, a Administração da Abyara Planejamento Imobiliário S.A. (“Companhia”) apresenta o relatório de Administração e Demonstrações Financeiras completas (com o respectivo parecer), referentes ao exercícios findos em 2008 e 2007.

Mensagem da Administração

O ano de 2008 foi desafiante e decisivo para o futuro da Abyara. A trajetória de crescimento no ano esperada nos primeiros meses de 2008 foi bruscamente interrompida pela redução da demanda e mais especificamente pela crise de liquidez e crédito vivida pela Companhia. Por esta razão, tivemos que nos reestruturar rapidamente, revisando nossa meta de lançamentos, diminuindo nossa estrutura e conseqüentemente nossas despesas, vendendo ativos para gerar liquidez e buscando soluções de capitalização para a empresa.

A partir do segundo semestre, tomamos ações no sentido de fortalecer a estrutura de capital da Abyara. Anunciamos a venda de 51% do negócio de corretagem por R\$250 milhões, dos quais boa parte foi utilizada para pagar dívidas bancárias. Iniciamos um processo de negociação com nossos bancos credores, visando a redução expressiva do endividamento da Abyara através da dação de ativos em pagamento e um alongamento de prazos excepcional. Em paralelo, revisamos nossa estrutura administrativa e reduzimos nosso quadro de pessoal. Este movimento capacitou a empresa para a operação de mudança de controle realizada em 18 de fevereiro de 2009.

Estes desafios trouxeram um grande amadurecimento e experiência para a Companhia, o que ficou comprovado pelos resultados positivos que alcançamos, em quatro principais frentes:

- I. **Consolidação de nossa posição diversificada.** A estratégia de nos tornarmos incorporadores plenos permitiu solidificar nosso posicionamento como incorporadores, o que está comprovado por nossa diversificação regional e de segmento de renda. Ao final de 2008, estávamos presentes em 12 Estados, com produtos voltados para todos os segmentos de renda. Nosso *landbank*, que atualmente é de R\$ 3,2 bilhões, é igualmente diversificado geograficamente e voltado para o segmento de média e baixa renda, proporcionando à Companhia a manutenção de sua estratégia de crescimento no mercado brasileiro de incorporação imobiliária. Este crescimento acontece agora num novo contexto mercadológico, que recomenda uma estratégia comercial mais conservadora.

1. incorporação imobiliária. Este crescimento acontece agora num novo contexto mercadológico, que recomenda uma estratégia comercial mais conservadora.
2. **Melhora da estrutura de capital.** Com a venda de parte do negócio de corretagem no terceiro trimestre, a Companhia reduziu seu endividamento em aproximadamente R\$145 milhões e teve um incremento de R\$128 milhões em seu patrimônio líquido. No último trimestre do ano, iniciamos um processo de renegociação do passivo bancário da Abyara com nove instituições credoras, objetivando a redução do endividamento e o alongamento da dívida remanescente para patamares compatíveis com o ciclo econômico-financeiro de nossa atividade.
3. **Redução de despesas e venda de ativos não estratégicos.** Também reduzimos nossa estrutura administrativa, até então dimensionada para um crescimento do mercado maior do que o apontado pelo novo contexto de desaceleração econômica verificada a partir do terceiro trimestre do ano. Reduzimos nosso quadro de colaboradores e outras despesas relacionadas. Também motivados pela crise no cenário econômico em geral e no mercado imobiliário em particular, vendemos alguns terrenos de nosso *landbank*, aqueles considerados de menor relevância para a estratégia da Companhia. Estas vendas proporcionaram um maior conforto de liquidez para a Abyara ao longo do segundo semestre.
4. **Bons resultados de vendas.** Ao longo de 2008, lançamos 16 empreendimentos, dos quais 8 possuem participação 100% Abyara. A despeito do impacto no crédito e na confiança dos consumidores causados pela desaceleração econômica intensificada a partir do quarto trimestre, vendemos R\$1,04 bilhão ao longo do ano de 2008, superando o total lançado, de R\$ 875 milhões, demonstrando nossa força e conhecimento de vendas. Ao final do ano de 2008, tínhamos um total acumulado de R\$1,2 bilhão em contas a receber e R\$609 milhões em estoque de unidades a vender. Acreditamos na força de nossa inteligência imobiliária, acumulada ao longo dos 14 anos em que a Abyara cresceu e se consolidou como uma das maiores marcas do mercado imobiliário brasileiro. Esta inteligência tem se mostrado decisiva em nossa atividade de incorporação.

Estas conquistas só foram possíveis graças ao comprometimento de nossa equipe, que vem trabalhando incansavelmente na obtenção de resultados positivos para a Companhia.

Em 18 de Fevereiro de 2009, o controle da Abyara foi vendido. Os antigos acionistas controladores venderam a totalidade de suas ações para a IPU Participações, que tem como acionistas a Kove Participações, com 70% e a Agra Empreendimentos Imobiliários, com 30%. Os novos controladores comprometeram-se a realizar um

bancário da Companhia. A operação coloca a Abyara como um *player* em posição privilegiada no mercado brasileiro, devido à viabilização da melhora da estrutura de capital da Companhia e da expertise trazida por Agra e Veremonte, acionista controladora da Kove Participações.

A Abyara tem hoje um *landbank* perfeitamente adaptado ao novo momento e uma capacidade estratégica ampliada, que permitirão um forte desenvolvimento dos negócios da Companhia, sob uma nova cultura de valorização de nossa equipe e de disciplina financeira.

Cenário Econômico

O ano de 2008 assistiu à brusca interrupção do mais robusto período de crescimento da economia mundial no pós-guerra, que havia começado em 2003. O pânico que se seguiu à quebra do Lehman Brothers, em setembro, agravou subitamente a crise que se desenrolava lentamente no sistema financeiro desde meados de 2007. Dada a globalização financeira, o problema se espalhou de forma rápida por todo o mundo. A tese de que os mercados emergentes ficariam descolados da crise não se confirmou e todos foram atingidos, em maior ou menor grau, pelo vendaval originado nos Estados Unidos.

No Brasil, a crise se instalou instantaneamente via três canais: crédito, comércio exterior e confiança. De uma hora para outra, as empresas brasileiras se viram diante de uma contração violenta do crédito, incluindo a suspensão de operações já aprovadas. O bom ritmo de crescimento da economia perdeu fôlego. Até o terceiro trimestre de 2008, a expansão ocorria a uma velocidade de 6,8%. No quarto trimestre, houve uma brusca desaceleração, que fez o PIB tombar 3,6% em relação ao mesmo período de 2007.

Felizmente, fruto do amadurecimento das transformações institucionais dos últimos 25 anos, o Brasil está mais preparado para enfrentar esta crise do que em qualquer outra das últimas décadas. Democracia, autonomia operacional do Banco Central, estabilidade macroeconômica, câmbio flutuante, inflação sob controle, endividamento público sustentável e nível confortável de reservas internacionais são conquistas desse período, que nos têm permitido navegar bem as águas turbulentas deste momento.

Graças a este amadurecimento, o Brasil tem tudo para sair bem da crise. A percepção de risco do país tem melhorado ao longo da tormenta, relativamente à média dos países emergentes. Assim, o ano de 2009 tenderá a ser um período duro de ajuste à situação internacional, o que nos legará uma estagnação da atividade econômica.

Do lado positivo, deveremos ver inflação em queda, bom nível de solvência do setor público e juros em forte declínio. É o oposto do que ocorreu em outras crises. Assim, a menos de um agravamento da situação nos Estados Unidos, que sinalize um ambiente de depressão (hipótese pouco provável), a economia brasileira deve emitir sinais de recuperação a partir do segundo trimestre de 2009, entrando 2010 em trajetória de crescimento de 3% a 4%, ainda abaixo do potencial, mas melhor do que estarão as economias desenvolvidas e muitos países emergentes. Não será pouco.