

UNIVERSIDADE DE SÃO PAULO
Escola de Comunicações e Artes
Departamento de Relações Públicas, Propaganda e Turismo
“Especialização em Pesquisa de Mercado Aplicada em Comunicações”

LARISSA SOUZA CATALÁ

**A ESCOLHA DOS *INPUTS* EM UM
PROJETO DE SEGMENTAÇÃO**
Uma aplicação em análise de agrupamentos
de eleitores brasileiros

São Paulo, 2011

UNIVERSIDADE DE SÃO PAULO
Escola de Comunicações e Artes
Departamento de Relações Públicas, Propaganda e Turismo
“Especialização em Pesquisa de Mercado Aplicada em Comunicações”.

LARISSA SOUZA CATALÁ

**A ESCOLHA DOS *INPUTS* EM UM
PROJETO DE SEGMENTAÇÃO**
Uma aplicação em análise de agrupamentos
de eleitores brasileiros

Monografia apresentada ao Departamento de Relações Públicas, Propaganda e Turismo da Escola de Comunicações e Artes da Universidade de São Paulo, em cumprimento parcial às exigências do Curso de Pós-Graduação-Especialização, para obtenção do título de especialista em “Pesquisa de Mercado Aplicada em Comunicações”, sob orientação do Prof. Paulo Cidade.

Banca Examinadora

Dedico este trabalho aos profissionais de pesquisa, para que sigam sempre questionando as técnicas e métodos utilizados.

Dedico também à minha família, amigos e namorado pelo apoio e carinho constantes.

Em especial, dedico ao meu pai, quem mais me incentiva na busca contínua por conhecimento.

Agradeço à USP e à IPSOS pelo curso oferecido e aos professores pelo profissionalismo e dedicação.

Ao IBOPE por ter disponibilizado o material que serviu de referência para o desenvolvimento deste trabalho e a seus funcionários que de alguma forma colaboraram com ideias, orientações ou incentivos durante a realização deste estudo.

Ao professor orientador Paulo Cidade pelo conhecimento compartilhado.

RESUMO

Os avanços computacionais das últimas décadas têm permitido o constante aperfeiçoamento das análises estatísticas em estudos de segmentação. Devido ao interesse e à utilidade dessas técnicas em diferentes áreas, bem como a necessidade de atuar de forma crítica frente aos diversos métodos disponíveis, o presente trabalho se propõe tratar da escolha dos *inputs* na Análise de Agrupamentos, abordando o uso da Análise Fatorial para reduzir o conjunto original de variáveis. Considerando este propósito, são apresentadas as etapas da Análise de Agrupamentos, uma breve descrição da Análise Fatorial e algumas considerações sobre seu uso na definição dos *inputs* da Análise de Agrupamentos, e, finalmente, são avaliados os resultados da comparação de duas segmentações: a primeira construída com uso da Análise Fatorial, e a segunda construída a partir das próprias variáveis originais, obtidas da mesma pesquisa. Os dados utilizados como base para a comparação foram obtidos da pesquisa que originou o estudo “*Cenário eleitoral 2010: mudanças e continuidades no comportamento eleitoral do Brasil pós-Lula*” realizada pelo IBOPE Inteligência¹ em março de 2010.

Palavras-chave: Segmentação, Análise de Agrupamentos, *Inputs*, Análise Fatorial, Variáveis Originais, Eleitores.

¹ IBOPE Inteligência é uma empresa do Grupo IBOPE especializada em pesquisas e informações nas áreas de consumo, marca e opinião pública (www.ibope.com). As análises, discussões e conclusões apresentadas no presente trabalho são de responsabilidade exclusiva da autora, não refletindo necessariamente o entendimento e a posição da empresa sobre os temas aqui tratados.

ABSTRACT

The computational advances in recent decades have allowed the constant improvement of the statistical analysis in segmentation studies. Due to the interest and usefulness of these techniques in different areas, as well as the need to act in a critical way forward to the various methods available, The present work aims to deal with the choice of inputs in Cluster Analysis, addressing the use of Factor Analysis to reduce the original set of variables. Considering this purpose, are presented the stages of Cluster Analysis, a brief description of Factor Analysis and some thoughts about your application to define the inputs of Cluster Analysis, and, finally, are evaluated the results by comparing two segmentations: the first constructed using factor analysis, and the second built on the original variables, obtained from the same survey. The data used as basis for comparison were obtained from the research that led to the study "*Electoral Scenario: Changes and Continuities in the post-Lula Brazilian Electoral Behavior*" conducted by IBOPE Inteligência² in March 2010.

Key words: Segmentation, Cluster Analysis, Inputs, Factor Analysis, Original Variables, Voters.

² IBOPE Inteligência is a IBOPE Group company specializing in research and information in areas of consumer, brand and public opinion (www.ibope.com). The analysis, discussion and conclusions presented in this work are the responsibility exclusive of the author, not necessarily reflecting the understanding and the company's position on issues presented here.

RESUMEN

Los avances tecnológicos en las últimas décadas han permitido la mejora constante de los análisis estadísticos en los estudios de segmentación. Debido al interés y la utilidad de estas técnicas en diferentes áreas, así como la necesidad de actuar de una manera crítica frente los distintos métodos disponibles, este documento propone tratar la elección de los *inputs* en el Análisis de Conglomerados, dirigiéndose a la utilización del Análisis Factorial para reducir el conjunto original de variables. Teniendo en cuenta este objetivo, se presentan las etapas del Análisis de Conglomerados, una breve descripción del Análisis de Factores y algunas reflexiones acerca de su aplicación para definir los *inputs* del Análisis de Conglomerados, y, finalmente, se evalúan los resultados de la comparación de dos segmentaciones: la primera construida con el uso del Análisis Factorial, y la segunda, construida con las variables originales, obtenidas a partir de la misma investigación. Los datos utilizados como base para la comparación fueron obtenidos de la investigación que generó el estudio "Escenario Electoral 2010: cambios y continuidades en el comportamiento electoral en Brasil post-Lula", realizado por IBOPE Inteligência³ en marzo del 2010.

Palabras clave: Segmentación, Análisis de Conglomerados, *Inputs*, Análisis Factorial, Variables Originales, Electores.

³ IBOPE Inteligência es una empresa del Grupo IBOPE especializada en investigación e información en las áreas de consumo, marca y opinión pública (www.ibope.com). El análisis, discusión y conclusiones presentados en este trabajo son responsabilidad exclusiva del autor y no necesariamente reflejan la comprensión y la posición de la empresa en los temas tratados.

SUMÁRIO

| | | |
|-----------|--|-----------|
| 1 | INTRODUÇÃO | 2 |
| 2 | OBJETIVO | 3 |
| 3 | ANÁLISE DE AGRUPAMENTOS | 4 |
| 3.1 | DEFINIÇÃO DO PROBLEMA DE PESQUISA E SELEÇÃO DAS VARIÁVEIS DE AGRUPAMENTO | 5 |
| 3.2 | ESCOLHA DA MEDIDA DE SIMILARIDADE | 6 |
| 3.3 | SELEÇÃO DO ALGORITMO DE AGRUPAMENTO | 8 |
| 3.3.1 | MÉTODOS HIERÁRQUICOS | 8 |
| 3.3.2 | MÉTODOS NÃO-HIERÁRQUICOS | 9 |
| 3.3.3 | COMBINAÇÃO DE AMBOS OS MÉTODOS | 11 |
| 3.4 | NÚMERO DE AGRUPAMENTOS A SEREM FORMADOS | 11 |
| 3.5 | INTERPRETAÇÃO DOS AGRUPAMENTOS | 12 |
| 3.6 | VALIDAÇÃO DO PROCESSO DE AGRUPAMENTO | 13 |
| 4 | ANÁLISE FATORIAL | 15 |
| 4.1 | CONSIDERAÇÕES SOBRE O USO DA ANÁLISE FATORIAL EM ANÁLISE DE AGRUPAMENTOS | 16 |
| 5 | A PESQUISA CENÁRIO ELEITORAL 2010: MUDANÇAS E CONTINUIDADES NO COMPORTAMENTO ELEITORAL DO BRASIL PÓS-LULA | 18 |
| 5.1 | PESQUISA QUANTITATIVA | 19 |
| 5.2 | SEGMENTAÇÃO ATITUDINAL PROPOSTA | 20 |
| 5.2.1 | CARACTERIZAÇÃO DOS GRUPOS | 21 |
| 5.2.2 | RELAÇÃO DOS GRUPOS COM A INTENÇÃO DE VOTO PARA PRESIDENTE NA ELEIÇÃO DE 2010 | 23 |
| 6 | METODOLOGIA | 25 |
| 7 | APLICAÇÕES E RESULTADOS | 27 |
| 8 | CONCLUSÕES | 32 |
| 9 | ANEXOS | 34 |
| 10 | REFERÊNCIAS BIBLIOGRÁFICAS | 38 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 – Distribuição da amostra de eleitores por sexo, escolaridade, faixa etária e região do país..... | 19 |
| Tabela 2 – Intenção de voto no 1º turno das eleições presidenciais de 2010..... | 24 |
| Tabela 3 – Concordância entre as Segmentações com base na Análise Fatorial e com base nas Variáveis Originais | 27 |
| Tabela 4 – Intenção de voto no 1º turno das eleições presidenciais de 2010..... | 30 |
| Tabela 5 – Principais diferenças sociodemográficas dos grupos atitudinais | 34 |
| Tabela 6 – “Se pudesse iria embora do Brasil” | 34 |
| Tabela 7 – “Eu costumo conversar sobre a política e o governo com outras pessoas” | 35 |
| Tabela 8 – “Seria bom que com o próximo presidente o Brasil continuasse no rumo que vai” | 35 |
| Tabela 9 – “Costumo acompanhar as notícias nos diversos meios de comunicação como TV, radio, jornal, revistas etc. para ser uma pessoa com maior consciência social e política” | 36 |
| Tabela 10 – "O casamento entre homossexuais deveria ser permitido" | 36 |
| Tabela 11 – "O meu voto em um determinado candidato pode mudar muito a minha vida, tanto para melhor, quanto para pior" | 36 |
| Tabela 12 – “A gente vê que hoje em dia no Brasil os recursos do governo chegam para os mais pobres”..... | 37 |
| Tabela 13 – “Me incomoda votar para presidente numa pessoa que está concorrendo numa eleição pela primeira vez” | 37 |

LISTA DE FIGURAS

| | |
|---|---|
| Figura 1 - Etapas da análise de agrupamentos..... | 5 |
|---|---|

1 INTRODUÇÃO

O grande avanço computacional ocorrido principalmente nas duas últimas décadas tem propiciado o constante aperfeiçoamento das análises de dados psicológicos, sociológicos e atitudinais, além dos comportamentais e demográficos. Essas informações são amplamente utilizadas em estudos de segmentação.

Segundo Myers (1996), “Segmentar um mercado consiste em dividi-lo em grupos de consumidores, cujas reações ao esforço de marketing da empresa sejam similares para os que pertencem ao mesmo grupo e distintas dos demais grupos”. Essa definição destaca o uso da segmentação na formulação de estratégias específicas de mercado.

A necessidade de analisar o comportamento e a atitude dos indivíduos não se restringe apenas ao âmbito do mercado de serviços e produtos. Pesquisas de opinião pública frequentemente fazem uso da análise de segmentação com o objetivo de identificar grupos de indivíduos semelhantes quanto à reflexão e atuação sobre temas relevantes. Dessa forma, é possível criar políticas públicas, desenvolver programas e até mesmo campanhas eleitorais voltadas às necessidades específicas de cada grupo, otimizando assim os esforços despendidos e os resultados obtidos.

Devido ao grande interesse pelo assunto e a difusão de técnicas de segmentação em setores como indústria, governo, empresas de serviços e centros de pesquisa acadêmicos, muitos livros e artigos têm sido publicados para tratar dos aspectos teóricos e matemáticos dessas ferramentas.

Somado a isso, a possibilidade de se trabalhar com grande quantidade de dados em um período relativamente curto de tempo tem permitido ao pesquisador se dedicar a uma avaliação mais crítica dos modelos teóricos empregados.

Dessa forma, como não há um consenso sobre o melhor método para segmentar um grupo de indivíduos ou objetos, este trabalho pretende explorar uma etapa importante da análise de agrupamentos, a escolha dos *inputs* a serem utilizados no modelo.

2 OBJETIVO

Esse trabalho se propõe descrever as etapas da análise de agrupamentos, comumente utilizada para segmentar indivíduos ou objetos, e comparar empiricamente os resultados de duas análises de agrupamentos: a primeira construída a partir de *inputs* gerados pelo uso da técnica Análise Fatorial (usualmente utilizada para agrupar variáveis), cujos resultados já foram publicados, e a segunda construída a partir das próprias variáveis originais, obtidas da mesma pesquisa, cujo desenvolvimento foi realizado exclusivamente para este trabalho. Além disso, será abordada também a importância do planejamento de pesquisa em um projeto de segmentação.

Para atingir os objetivos propostos, foram utilizados os dados levantados pela pesquisa que originou o estudo “*Cenário eleitoral 2010: mudanças e continuidades no comportamento eleitoral do Brasil pós-Lula*” realizada pelo IBOPE Inteligência em 2010, cujos objetivos eram analisar o comportamento eleitoral dos brasileiros após duas décadas ininterruptas de eleições presidenciais e desenvolver e testar uma segmentação atitudinal dos eleitores com base no conhecimento acumulado e da literatura sobre comportamento eleitoral (Nunes et al., 2010).

A comparação proposta pretende apresentar um novo olhar para o problema da segmentação dos eleitores brasileiros, enriquecendo assim as análises já publicadas, e contribuir para que pesquisadores em geral, estatísticos e não estatísticos, reflitam sobre aspectos relevantes da análise de segmentação.

3 ANÁLISE DE AGRUPAMENTOS

Segundo Frei (2006), a análise de agrupamentos é o nome dado a um conjunto de procedimentos que busca reunir indivíduos ou objetos em grupos homogêneos, segundo algum critério pré-determinado. Os agrupamentos resultantes desse processo devem apresentar alta homogeneidade interna (dentro dos grupos) e alta heterogeneidade externa (entre os grupos).

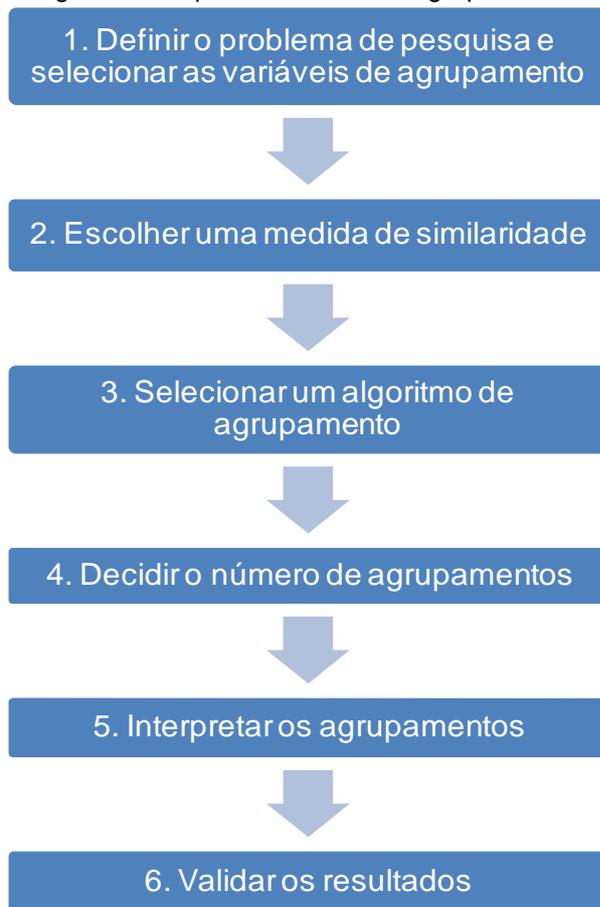
Dessa forma, dada uma amostra de n indivíduos, cada um deles medido segundo p variáveis, procura-se um esquema de classificação que agrupe os indivíduos em k grupos (Bussab, 1990 apud Frei, 2006). Os indivíduos são mensurados nas diversas variáveis de interesse fornecendo uma matriz de dados de n indivíduos por p variáveis, a qual será manuseada através de algoritmos para a obtenção dos grupos homogêneos.

A análise de agrupamentos é comumente utilizada em projetos de segmentação devido à sua utilidade para reduzir dados através da redução da informação de uma população ou amostra para a informação sobre subgrupos (segmentos) específicos e menores, com perda mínima de informação.

Os resultados obtidos são, então, utilizados para compreender melhor o público de interesse o que permite a elaboração de estratégias específicas de marketing e comunicação capazes de intervir no comportamento, opinião e atitude da população investigada.

A análise de agrupamentos pode ser realizada com o uso de diferentes métodos em diferentes etapas do processo. No entanto, independente dos métodos utilizados, ela deverá ser realizada de acordo com as seguintes etapas:

Figura 1 - Etapas da análise de agrupamentos



Fonte: Adaptada de HAIR et al., 2005.

As seções seguintes descrevem cada uma das etapas que compõem a análise de agrupamentos.

3.1 DEFINIÇÃO DO PROBLEMA E SELEÇÃO DAS VARIÁVEIS DE AGRUPAMENTO

Para que uma análise de segmentação se revele útil, o pesquisador deve definir claramente o objetivo do projeto, seja ele exploratório ou confirmatório, dispor de um desenho amostral e tamanho de amostra adequados e dedicar atenção especial à elaboração do questionário, pois será através deste instrumento que serão mensuradas as variáveis utilizadas para caracterizar os indivíduos a serem agrupados.

A solução de um problema de segmentação é totalmente dependente das variáveis usadas como base para a medida de similaridade. De acordo com Milligan (1980, apud Hair et al., 2005), o acréscimo ou a eliminação de variáveis relevantes

pode ter um impacto substancial sobre a solução resultante. Como a técnica de análise de agrupamentos não distingue variáveis relevantes e irrelevantes, ela apenas refletirá a estrutura inerente dos dados expressa pelas variáveis consideradas na análise pelo pesquisador.

A seleção das variáveis deve considerar questões teóricas, conceituais e práticas. Pesquisas anteriores, o conhecimento acumulado e até mesmo as hipóteses que serão testadas no estudo podem contribuir no processo de identificação das variáveis relevantes para a análise. Essa etapa do processo requer o julgamento e a intuição do pesquisador, como sugere Malhotra (2001).

Um aspecto importante que deve ser examinado é a presença de multicolinearidade⁴ nos dados. O pesquisador deve examinar as variáveis utilizadas na análise de agrupamentos e, caso encontre multicolinearidade substancial, deve reduzir as variáveis a números iguais em cada dimensão ou usar uma medida de similaridade que compense essa correlação (Hair et al., 2005).

Os resultados da análise também devem ser examinados e aquelas variáveis que não são distintas entre os segmentos obtidos devem ser eliminadas. Esse procedimento permite que as técnicas de aglomeração definam maximamente segmentos com base apenas nas variáveis que exibem diferenças entre os grupos.

Considerando a relevância da etapa de seleção de *inputs* para um projeto de segmentação realizado com base na análise de agrupamentos, e sendo este o tema central do presente trabalho, na seção 4 mais adiante, será abordado o uso da técnica de análise fatorial na geração de variáveis de agrupamento e os possíveis inconvenientes do uso desta técnica em uma análise de agrupamentos.

3.2 ESCOLHA DA MEDIDA DE SIMILARIDADE

A medida de similaridade determina quão semelhantes, ou quão diferentes, são os objetos que estão sendo agrupados. Os métodos mais utilizados são: a correlação, a distância e a associação.

Medidas de correlação entre indivíduos ou objetos investigam a correspondência de padrões ao longo das variáveis e, dessa forma, expressam muito mais os padrões do que as magnitudes. Na maioria dos projetos de

⁴A multicolinearidade se faz presente quando existe algum nível de inter-relação entre as variáveis estudadas.

segmentação deseja-se comparar os indivíduos em termos da magnitude de suas características e, por isso, esse tipo de medida é raramente utilizado.

A similaridade calculada pela distância entre pares de objetos ou indivíduos em um espaço n-dimensional é a abordagem mais comum. Os agrupamentos formados com base na distância entre observações apresentam valores mais similares no conjunto de variáveis, ainda que internamente os objetos ou indivíduos possam ter padrões diferentes (Hair et al., 2005).

Diversas medidas de distância estão disponíveis. A mais usada é a *distância euclidiana* representada pela raiz quadrada da soma dos quadrados das diferenças dos valores para cada variável. Geralmente trabalha-se com a *distância euclidiana quadrada ou absoluta*, expressa da mesma forma que a anterior sem calcular a raiz quadrada, o que torna o tempo de computação mais rápido, conforme descrito por Hair et al. (2005).

Segundo Romesburg (1984, apud Frei, 2006), quando as variáveis apresentam escalas de mensuração muito diferentes, a solução por agrupamento será influenciada pelas unidades de medida, pois as variáveis com maior dispersão terão maior peso no valor da similaridade final. Assim, para evitar que isso ocorra, Frei (2006) sugere que a padronização das variáveis de agrupamento seja realizada, a fim de permitir que as variáveis contribuam de forma mais igualitária para a similaridade entre os objetos.

Por outro lado, para evitar que a padronização reduza também as diferenças entre os agrupamentos em variáveis que melhor discriminam os aglomerados, é conveniente excluir da análise as observações atípicas (*outliers*), sempre quando consideradas não representativas, de acordo com H. Charles Romsburg (1990, apud Malhotra, 2001).

A *distância de Mahalanobis* é um tipo de distância euclidiana que incorpora automaticamente um procedimento de padronização. Sua utilização é recomendada quando as variáveis são correlacionadas (positiva ou negativamente), pois ela ajusta intercorrelações e pondera todas as variáveis igualmente. Mais detalhes em Hair et al. (2005).

Outra possibilidade é a padronização dos dados por respondentes. Scaninger, C.M. e W.C. Bass (1986, apud Hair et al., 2005) sugerem que este procedimento seja utilizado para eliminar efeitos de estilo de respostas, que são os padrões sistemáticos de respostas a um conjunto de questões. Isso porque, na

maioria dos projetos de segmentação e principalmente em análise de dados atitudinais, o que se deseja obter é a importância relativa de uma variável sobre a outra.

As medidas de similaridade por associação são usadas para comparar objetos cujas características são medidas apenas em termos não-métricos (nominal ou ordinal). No entanto, muitos programas de computador têm suporte limitado para medidas de associação. Mais detalhes em Hair et al. (2005).

O uso de diferentes medidas de similaridade pode resultar em diferentes soluções de agrupamento. Por isso, recomenda-se considerar na análise diferentes medidas de similaridade e comparar os resultados com padrões teóricos ou conhecidos.

3.3 SELEÇÃO DO ALGORITMO DE AGRUPAMENTO

Uma vez selecionadas as variáveis e obtida a matriz de similaridade, o passo seguinte é eleger um algoritmo de agrupamento ou conjunto de regras mais apropriado ao problema investigado.

Os algoritmos mais usados se dividem em duas classes: hierárquica e não-hierárquica. A primeira se caracteriza pelo estabelecimento de uma hierarquia ou estrutura em forma de árvore, enquanto que a segunda designa objetos a aglomerados conforme o número k de grupos estabelecidos previamente.

3.3.1 MÉTODOS HIERÁRQUICOS

Os métodos hierárquicos podem ser aglomerativos ou divisivos. Os métodos aglomerativos começam com cada indivíduo ou objeto formando um aglomerado. Em passos seguintes, são formados aglomerados maiores com o agrupamento de aglomerados formados em estágios anteriores. Os métodos divisivos procedem em direção oposta aos métodos aglomerativos, começando com um grande agregado formado por todos os indivíduos ou objetos que, em sucessivas partições, dão origem a aglomerados menores.

Os métodos hierárquicos aglomerativos são os mais empregados na prática e estão disponíveis nos softwares mais comumente utilizados (MALHOTRA, 2001). Os principais métodos são: ligação individual, ligação completa, ligação média,

método de Ward e método centróide. A diferença entre eles está no cálculo da distância entre os agrupamentos.

O método da ligação individual, também conhecido como regra do vizinho mais próximo, baseia-se na distância mínima, onde a distância entre dois aglomerados é a distância entre seus dois pontos (objetos ou indivíduos) mais próximos. O método da ligação completa, também conhecido como regra do vizinho mais afastado, baseia-se na distância máxima, ou seja, a distância entre aglomerados é calculada entre seus dois pontos mais afastados. O método da ligação média, por sua vez, funciona de maneira análoga, onde a distância entre aglomerados é definida como a média das distâncias entre todos os pares de objetos ou indivíduos, onde cada membro de um par provém de cada um dos aglomerados. O fato de este método considerar informações sobre todos os pares de distância, e, por isso, ser menos suscetível a valores extremos, faz com que ele seja o preferido entre os métodos de ligação (MALHOTRA, 2001).

O método de Ward é um método de variância que procura gerar agrupamentos minimizando a variância dentro dos grupos. Esse procedimento tende a construir agregados com aproximadamente o mesmo número de observações. Mais detalhes em Hair et al. (2005).

No método centróide, a distância entre dois agrupamentos é a distância entre seus centróides (médias para todas as variáveis). Cada vez que se agrupam objetos, calcula-se um novo centróide.

Entre os métodos hierárquicos, os da ligação média e de Ward têm-se mostrado superiores aos demais, de acordo com Richard A. Johnson e Dean A. Wichern (1998, apud MALHOTRA, 2001).

3.3.2 MÉTODOS NÃO-HIERÁRQUICOS

Os métodos não-hierárquicos, normalmente chamados de k-médias, buscam a melhor solução de acordo com o número k de agrupamentos previamente especificado. O uso deste método requer a seleção de sementes - que são os centros iniciais dos agrupamentos cuja distância a cada objeto será medida -, a especificação de uma distância de referência e a escolha de um método de designação de objetos aos aglomerados.

De acordo com Hair et al. (2005), as abordagens mais comuns para designar indivíduos ou objetos a um dos aglomerados são a referência sequencial, referência paralela e a otimização.

O método da referência sequencial começa pela seleção de uma semente de aglomeração, e todos os objetos a menos de uma distância pré-especificada desta semente são agrupados. Quando todos os objetos dentro da distância são incluídos, uma segunda semente é selecionada e todos os objetos dentro da distância pré-especificada são agrupados. Em seguida, uma terceira semente é selecionada e o processo segue como descrito anteriormente. Um objeto aglomerado a uma semente não é mais considerado para aglomerações subsequentes.

O método da referência paralela opera de maneira semelhante ao da referência sequencial, com a diferença de que se escolhem simultaneamente várias sementes de aglomeração e os objetos dentro da distância pré-estabelecida são agrupados às sementes mais próximas.

O terceiro método, conhecido como procedimento de otimização, difere dos dois anteriores pelo fato de permitir a redesignação de objetos aos agrupamentos, pois, se no decorrer do processo de aglomeração um dos objetos se torna mais próximo de outro agrupamento que não aquele ao qual ele está associado no momento, ele é transferido para o agrupamento mais próximo (mais semelhante).

Segundo Hair et al. (2005), o principal problema enfrentado ao utilizar um procedimento de agrupamento não-hierárquico é a seleção das sementes. Alguns programas escolhem os primeiros k casos sem valores faltantes como centros iniciais de aglomeração, outros utilizam uma seleção aleatória de pontos sementes. Essas duas abordagens podem apresentar diferentes resultados para cada conjunto de pontos sementes selecionados. Por outro lado, há programas que oferecem como alternativa a possibilidade de que o próprio pesquisador especifique as sementes iniciais, o que deve ser feito segundo alguma base prática, objetiva ou teórica.

3.3.3 COMBINAÇÃO DE AMBOS OS MÉTODOS

Os métodos hierárquicos são mais suscetíveis aos impactos de observações atípicas e menos tratáveis para analisar amostras muito grandes. Além disso, combinações iniciais indesejáveis construídas com esses métodos podem persistir na análise e conduzir a resultados artificiais.

Os métodos não hierárquicos, por sua vez, embora possibilitem a alteração de pertinência dos objetos aos grupos e sejam menos suscetíveis a observações atípicas, têm como principais desvantagens a necessidade de especificação do número de agrupamentos e a arbitrariedade da escolha dos pontos sementes.

Dessa forma, para minimizar os principais problemas destes métodos, Milligan (1980, apud HAIR et al., 2005) e B. Everitt (1993, apud MALHOTRA, 2001) sugerem utilizar conjuntamente os métodos hierárquicos e não-hierárquicos a fim de obter os benefícios de ambos. Primeiramente, uma técnica hierárquica pode identificar observações atípicas, estabelecer o número de agrupamentos e caracterizar os centros de aglomeração. Em seguida, excluídas as observações atípicas, as demais observações podem então ser agrupadas por um método não-hierárquico onde o número de grupos e as sementes iniciais serão aqueles resultantes do processo hierárquico.

3.4 NÚMERO DE AGRUPAMENTOS A SEREM FORMADOS

Não existe um procedimento padrão e objetivo para resolver o problema de determinação do número final de agrupamentos a serem formados. Geralmente, essa decisão exige julgamento por parte do pesquisador.

O pesquisador deve ter em conta que, quando o número de agrupamentos diminui, a homogeneidade dentro dos grupos necessariamente diminui. Assim, Hair et al. (2005) sugerem que o pesquisador busque um equilíbrio entre definir a estrutura mais básica (menos agrupamentos) e, ainda assim, conseguir o nível necessário de similaridade dentro dos agrupamentos. Além disso, os tamanhos relativos dos aglomerados devem ser significativos para permitir maior precisão na “leitura” de informações referentes aos grupos.

A maioria dos softwares oferece uma opção para a decisão do número de grupos. Um procedimento relativamente simples, de acordo com Frei (2006) consiste

em examinar alguma medida de similaridade ou distância entre os grupos a cada passo sucessivo. A solução final é dada quando a medida escolhida apresenta um salto abrupto entre dois passos consecutivos. Quando isso ocorre, o pesquisador deve selecionar a solução anterior sob o argumento de que a combinação seguinte provocaria uma queda substancial de similaridade. Em um procedimento hierárquico, a análise visual de um dendograma⁵ pode auxiliar nessa tarefa (FREI, 2006).

Existem outros procedimentos mais complexos, baseados em regras ou testes estatísticos, como, por exemplo, o Critério de Agrupamento Cúbico (CCC, disponível no software estatístico SAS). No entanto, Hair et al. (2005) consideram estes testes excessivamente complexos diante da pouca melhoria que oferecem quando comparados a procedimentos mais simples.

A definição do número de agrupamentos também deve considerar questões de ordem prática. Pode-se considerar, por exemplo, que trabalhar e fazer comunicação em um cenário com três a seis grupos seja mais viável. O pesquisador então pode gerar soluções de segmentação para cada um desses números de grupos e selecionar a melhor alternativa depois de avaliar todas elas. Lembrando sempre que aspectos teóricos e conceituais relacionados ao problema de pesquisa podem e devem ser levados em conta nesta etapa do processo (MALHOTRA, 2001).

3.5 INTERPRETAÇÃO DOS AGRUPAMENTOS

A etapa de interpretação dos agrupamentos formados consiste em entender suas características através do estudo minucioso das variáveis utilizadas para construí-los, e de outras variáveis adicionais importantes. A partir dessa análise é possível criar um nome ou rótulo que descreva adequadamente a natureza dos segmentos obtidos.

A identificação do perfil dos agrupamentos permite avaliar a correspondência dos grupos formados com aqueles propostos por alguma teoria ou experiência prática, além de permitir o melhor direcionamento dos esforços de marketing.

⁵Dendograma é uma representação gráfica (em árvore) dos resultados de um procedimento hierárquico no qual cada objeto é colocado em um eixo e o outro eixo representa os passos no procedimento hierárquico. Começando com cada objeto representado como um agrupamento separado, o dendograma mostra como os agrupamentos são combinados em cada passo do procedimento até que todos estejam contidos em um único agrupamento (Hair et al., 2005).

Em um projeto de segmentação atitudinal é possível, por exemplo, interpretar os segmentos utilizando a bateria de frases para as quais os entrevistados tiveram que declarar seu grau de concordância ou importância e, adicionalmente, utilizar variáveis demográficas e de consumo para complementar a descrição do perfil dos grupos.

Quando a análise de agrupamentos é realizada com base nos dados iniciais, sem padronização ou qualquer outra transformação, uma avaliação dos centróides dos aglomerados pode auxiliar no processo de interpretação dos segmentos (HAIR et al., 2005). Caso contrário, o pesquisador pode comparar os escores médios dos agrupamentos em relação às variáveis investigadas, para tentar explicar diferenças relevantes entre os grupos.

Outra ferramenta que pode ser utilizada nesta fase do projeto é a análise discriminante. Segundo Malhotra (2001), essa técnica estatística multivariada consiste na obtenção de funções discriminantes, ou combinações lineares das variáveis independentes que melhor discriminam entre as categorias da variável dependente (grupos).

Neste contexto, a análise discriminante é utilizada para verificar se existem diferenças significativas entre os segmentos formados pela análise de agrupamentos, onde a variável dependente categórica será aquela que identifica previamente o grupo a que pertencem as observações, e as variáveis independentes serão as características demográficas, psicográficas, de consumo, de utilização de meios de comunicação, etc. Além disso, essa técnica permite determinar as variáveis que mais contribuem para as diferenças entre os grupos e avaliar a precisão da classificação.

3.6 VALIDAÇÃO DOS RESULTADOS

Devido à natureza subjetiva da verificação da qualidade dos resultados da análise de agrupamentos, o pesquisador deve tomar muito cuidado ao avaliar a estabilidade e generalidade da solução proposta.

Apesar de não haver métodos que garantam a validade e significância práticas das soluções de segmentação, diversas abordagens foram propostas para fornecer alguma base para o pesquisador verificar a qualidade dos resultados obtidos. Malhotra (2001) sugere ao pesquisador:

- Fazer análises de agrupamentos sobre os mesmos dados utilizando diferentes medidas de distância e algoritmos de agrupamento e comparar os resultados para verificar a estabilidade da solução.
- Dividir aleatoriamente a base de dados em duas partes. Fazer a segmentação separadamente sobre cada parte e comparar os centróides dos agrupamentos formados nas duas subamostras.
- Fazer a segmentação com base em um conjunto reduzido de variáveis e comparar os resultados com aqueles obtidos pela segmentação baseada em todo o conjunto de variáveis.
- Para métodos de agrupamentos não-hierárquicos, como a solução pode depender da ordem dos casos na base de dados, é recomendável realizar múltiplas análises utilizando diferentes ordens de casos até que a solução se estabilize.

Pela descrição das etapas envolvidas na análise de agrupamentos, relacionadas anteriormente, pode-se notar a importância de se aplicar a técnica de maneira consciente para se obter resultados confiáveis e úteis. Para isso, o pesquisador deve dispor de tempo, cuidado e muita atenção.

4 ANÁLISE FATORIAL

Análise fatorial é um nome dado a um conjunto de técnicas estatísticas multivariadas utilizadas basicamente para reduzir e sumarizar variáveis através da análise da relação de interdependência existente entre elas, mantendo a maior parte da informação contida nos dados originais (MALHOTRA, 2001).

A análise fatorial pode ser comparada à análise de regressão múltipla, pelo fato de cada variável ser expressa como uma combinação linear de fatores comuns subjacentes. Esses fatores, por sua vez, podem ser expressos como combinações lineares das variáveis observáveis, o que permite que eles sejam interpretados e descritos.

A diferença entre as várias técnicas de análise fatorial está na maneira como elas selecionam os pesos que serão associados às variáveis originais na construção dos fatores. A análise de componentes principais, principal técnica de análise fatorial, constrói fatores ortogonais, ou seja, não correlacionados entre si de forma linear, e as combinações lineares que os definem são formadas de tal maneira que cada escore fatorial contribui com uma proporção decrescente da variância das variáveis originais (GREEN e TULL, 1975).

Essa técnica é comumente utilizada em pesquisa de mercado e opinião para identificar fatores que expliquem as correlações entre um conjunto de variáveis, selecionar um subconjunto representativo de variáveis para uma análise posterior, e, freqüentemente, para substituir o conjunto original de variáveis correlacionadas pelo conjunto, menor, de variáveis não-correlacionadas (escores fatoriais) em uma análise multivariada subsequente, como por exemplo, em uma análise de agrupamentos.

Assim como na análise de agrupamentos, o resultado obtido com a análise fatorial é totalmente dependente do conjunto de variáveis submetidas à análise. O pesquisador implicitamente especifica as dimensões potenciais identificáveis por meio do caráter e da natureza das variáveis selecionadas para análise. Por isso, essa técnica torna-se mais eficiente quando dimensões conceitualmente definidas podem ser representadas pelos fatores determinados.

A fim de evitar que os fatores originados pela análise fatorial sejam excessivamente “aderentes” à amostra, ou seja, pouco generalizáveis, a amostra

selecionada deve ter um tamanho compatível com o número de variáveis investigadas.

Para mais detalhes, ver Green e Tull (1975).

4.1 CONSIDERAÇÕES SOBRE O USO DA ANÁLISE FATORIAL EM ANÁLISE DE AGRUPAMENTOS

Em estudos de segmentação atitudinal é comum investigar o grau de concordância ou importância associado a um grande número de atributos ou afirmações relacionadas ao tema de interesse.

Em situações como essa, o pesquisador tem a opção de: considerar o conjunto completo de variáveis ou trabalhar com um conjunto reduzido de variáveis que expressem as idéias centrais da investigação, utilizando para isso alguma técnica de redução de dados como, por exemplo, a análise fatorial.

De acordo com Mayers (1996), a maioria dos pesquisadores prefere a segunda opção. Em uma situação onde são investigadas, por exemplo, de 100 a 125 atributos ou afirmações, o uso da análise fatorial pode resultar na construção de 20 a 30 fatores que juntos capturam a maior parte do significado de todas as declarações combinadas. Quando isso acontece, o pesquisador então gera os escores fatoriais associados a cada fator para cada entrevistado e os utiliza como variáveis de agrupamento para a análise de segmentação. Essa abordagem tem a vantagem de tornar a aplicação de um algoritmo de agrupamento computacionalmente mais rápida, além de reduzir a complexidade de interpretação dos resultados da segmentação.

No entanto, Mayers (1996) aponta alguns inconvenientes no uso dessa técnica. São eles:

1. Quando se combinam diversas variáveis em um único escore fatorial uma grande quantidade de informação pode ser perdida. Fatores raramente explicam mais de 60-75% da variância total do conjunto completo de variáveis e, assim, o restante não estará disponível para influenciar na formação dos aglomerados.

2. Alguns atributos ou afirmações não serão significativamente representadas em nenhum dos fatores e, portanto, não impactarão a formação dos aglomerados baseados em escores fatoriais. Entretanto, qualquer um desses

atributos ou afirmações relativamente independentes poderia ser o único que melhor distinguiria os grupos atitudinais, mas, no entanto, é misturado a outros que não discriminam tão bem na construção dos escores fatoriais. Isso pode ocorrer porque escores fatoriais não têm, necessariamente, uma relação com a definição dos agrupamentos.

3. Escores fatoriais, da maneira como são usualmente definidos, são totalmente independentes uns dos outros. Isso tende a espalhar os indivíduos no espaço n -dimensional de tal maneira que, computacionalmente, é mais difícil encontrar aglomerados de pontos em áreas de alta densidade. É importante lembrar, entretanto, que tampouco é desejável que as variáveis originais sejam fortemente correlacionadas entre si, o que torna necessária a eliminação a priori de variáveis redundantes.

Um estudo empírico conduzido por Yeung e Ruzzo (2001), cujo objetivo era estudar a eficácia da técnica Componentes Principais na captura da estrutura dos agrupamentos, mostrou que pode existir algum outro conjunto de m componentes principais, que não necessariamente as primeiras, as quais explicam a maior parte da variabilidade dos dados, que proporcionam maior qualidade nos resultados de agrupamento quando comparado à utilização das primeiras componentes principais.

Além disso, Yeung e Ruzzo (2001) concluíram que a análise de agrupamentos gerada com componentes principais só aumenta a qualidade da segmentação quando o número correto de componentes ou quando o conjunto correto de componentes principais é escolhido, o que na prática é difícil de garantir já que não existe uma metodologia satisfatória para determinar o número adequado de componentes ou um conjunto informativo de componentes principais sem depender de critérios externos ao conjunto de dados.

Apesar das questões aqui levantadas, não é uma escolha simples optar entre o uso de escores fatoriais ou de variáveis originais em uma análise de agrupamentos. A seguir, será apresentada a pesquisa "*Cenário eleitoral 2010: Mudanças e continuidades no comportamento eleitoral do Brasil pós-Lula*" e os resultados da segmentação atitudinal realizada com base em escores fatoriais, cujos resultados serão comparados empiricamente aos resultados obtidos com a segmentação com base nas variáveis originais.

5 A PESQUISA CENÁRIO ELEITORAL 2010: MUDANÇAS E CONTINUIDADES NO COMPORTAMENTO ELEITORAL DO BRASIL PÓS-LULA

A eleição presidencial de 2010 foi um marco na história política do Brasil. Isso porque, entre outras razões, o país completou com ela duas décadas ininterruptas de eleições diretas e encerrou o governo do presidente Lula, o qual, em seu segundo mandato, havia disputado as cinco últimas eleições presidenciais brasileiras.

Somado a esses fatores, as mudanças ocorridas na composição etária da população, distribuição de renda, consumo, acesso à informação, escolaridade, a participação da mulher no mercado de trabalho, entre outros, tem tornado ainda mais complexa a compreensão do comportamento eleitoral dos brasileiros.

A pesquisa *Cenário Eleitoral 2010* foi desenvolvida pelo IBOPE Inteligência e seu principal objetivo foi avaliar o comportamento dos eleitores brasileiros na eleição presidencial de 2010, considerando os avanços políticos, sociais e econômicos do país nas últimas décadas.

Os resultados alcançados forneceram importantes subsídios para o planejamento de campanhas eleitorais bem como para a realização de políticas públicas e iniciativas de comunicação por parte de organizações e instituições políticas no ano de 2010 e nos próximos anos.

Para atingir os objetivos propostos, o estudo foi estruturado em três etapas: *desk research*⁶, pesquisa qualitativa e pesquisa quantitativa. A *desk research* foi realizada com o propósito de analisar as principais mudanças ocorridas na sociedade. A pesquisa qualitativa teve o objetivo de aprofundar a investigação sobre o comportamento eleitoral brasileiro e fornecer subsídios para sustentar os resultados obtidos pela etapa quantitativa. E a pesquisa quantitativa foi estruturada para que um estudo de segmentação atitudinal dos eleitores brasileiros fosse desenvolvido, entre outras análises.

Uma vez que a pesquisa quantitativa desse trabalho é o foco e principal recurso do presente estudo, maiores detalhes sobre seu planejamento e resultados obtidos com a segmentação atitudinal serão descritos na próxima seção.

⁶*Desk Research*, também conhecida como pesquisa de dados secundários, consiste no levantamento de informações disponíveis em diferentes fontes, como: instituições governamentais, na internet, revistas especializadas, jornais, relatórios de empresas, publicações de institutos de pesquisa, etc. (www.multifocus.com.br)

5.1 PESQUISA QUANTITATIVA

Foram realizadas 2002 entrevistas pessoais com eleitores brasileiros através de um questionário estruturado em Março de 2010. A amostra foi selecionada em três estágios: no primeiro estágio foi feita a seleção dos municípios com PPT (Probabilidade Proporcional ao Tamanho) sistemático, tomando como base a população de 16 anos ou mais de cada um deles; no segundo estágio foi feita a seleção dos setores censitários, também com PPT sistemático, considerando como medida de tamanho a população com 16 anos ou mais residente nos setores; e, finalmente, no terceiro estágio foi selecionado em cada setor um número fixo de eleitores segundo cotas das variáveis sexo, grupos de idade, instrução e ramo de atividade⁷.

A tabela abaixo mostra o perfil dos entrevistados segundo a distribuição de sexo, escolaridade, faixa etária e região do país para que pudessem representar o eleitorado brasileiro:

Tabela 1 – Distribuição da amostra de eleitores por sexo, escolaridade, faixa etária e região do país

| SEXO | % | ESCOLARIDADE | % |
|-------------|----------|---------------------|----------|
| Masculino | 48 | Até a 4ª série | 29 |
| Feminino | 52 | Da 5ª a 8ª série | 23 |
| | | Ensino Médio | 34 |
| | | Universidade | 14 |

| FAIXA ETÁRIA | % | REGIÃO DO PAÍS | % |
|---------------------|----------|-----------------------|----------|
| 16 a 24 anos | 21 | Norte | 7 |
| 25 a 29 anos | 13 | Nordeste | 27 |
| 30 a 39 anos | 22 | Sudeste | 44 |
| 40 a 49 anos | 19 | Sul | 15 |
| 50 anos ou mais | 26 | Centro-Oeste | 7 |

Fonte: IBOPE Inteligência

O questionário aplicado abordou temas relacionados à política em geral, ao contexto eleitoral de 2010, avaliação do presidente e do país sobre diversas áreas, fonte de informações utilizadas pelos eleitores, além das baterias de frases que

⁷Fontes de dados para elaboração da amostra: Censo 2000, PNAD 2008 e TSE 2008.

trataram de temas relativos à moral, religião, economia, valores pessoais, justiça, ética, entre outros.

5.2 SEGMENTAÇÃO ATITUDINAL PROPOSTA

A segmentação proposta pelo estudo *Cenário Eleitoral 2010* se baseou no Modelo Michigan, que representa a tradição psicossociológica de explicação do comportamento eleitoral. De acordo com Nunes et al., (2010), essa perspectiva considera que

... as atitudes são consolidadas pelos indivíduos em todo o seu processo de socialização e, independentemente dos contextos históricos específicos, tornam-se estáveis e influenciam as avaliações e comportamentos que os indivíduos têm sobre os diversos aspectos da sua existência, incluindo o aspecto político.

Assim, com o objetivo de agrupar os indivíduos em grupos com atitudes e comportamentos similares em relação às várias dimensões investigadas (moral, religiosa, econômica, profissional, política, eleitoral, social, etc.), foram incluídas no questionário da pesquisa um total de 89 frases, com um número aproximado de frases por dimensão, as quais os repondentes tinham que avaliar seu nível de concordância segundo uma escala de 5 pontos (1 – Discorda Totalmente e 5 – Concorda Totalmente).

A maioria das frases foram selecionadas de pesquisas sobre temas políticos e sociais como a World Values Survey e a Latinobarometro e outras pesquisas brasileiras sobre cultura e comportamento político; outras foram obtidas de pesquisas de mercado e opinião pública realizadas pelo próprio IBOPE Inteligência e de comentários espontâneos dos participantes dos grupos focais da primeira etapa do estudo.

Para obter os segmentos ou grupos homogêneos de eleitores segundo as dimensões abordadas, foram realizadas análises fatoriais⁸ com o conjunto das 89 variáveis (frases), e com os escores fatoriais obtidos foram desenvolvidas análises de agrupamentos com a técnica *k-means* (não-hierárquico). Para alcançar

⁸Método de Rotação: Varimaxwith Kaiser Normalization

resultados satisfatórios, modelos de segmentação com diferentes número de escores fatoriais foram avaliados (foram testadas soluções com até 20 fatores).

Uma das soluções obtidas resultou em cinco fatores principais, dos quais um era composto por 20 frases com alto nível de concordância por parte dos entrevistados. Essas afirmações representavam atitudes comuns à população brasileira de maneira geral, como nacionalismo, importância da família, religião e trabalho, valor da democracia e igualdade social, otimismo em relação ao futuro, etc. Considerando a generalidade dessas afirmações e a baixa relevância desses temas na explicação das escolhas políticas e do comportamento eleitoral, optou-se por excluí-las da análise de segmentação.

Com as 69 afirmações restantes foram realizadas novas análises fatoriais e, considerando o percentual acumulado de variabilidade dos dados e, principalmente, a interpretabilidade dos fatores, decidiu-se utilizar para a análise de segmentação final uma solução com 4 fatores, que juntos explicaram apenas 25% da variabilidade dos dados. Com a análise de agrupamentos realizada a partir dos escores fatoriais dessa solução, foram construídos 4 grupos atitudinais de eleitores brasileiros, sendo que cada um deles representava um fator específico.

5.2.1 CARACTERIZAÇÃO DOS GRUPOS

De acordo com o estudo *Cenário Eleitoral 2010*, as principais diferenças sociodemográficas⁹ e atitudinais observadas entre os grupos permitiu rotulá-los e caracterizá-los da seguinte forma, conforme seus perfis:

Grupo 1 – “Apegados” (19%):

- Maior presença de homens
- Eleitores de municípios menores da região Sul do país
- Mais adeptos das religiões Evangélica e Protestante
- Temerosos, inseguros, ameaçados
- Traços de autoritarismo como uma reação à insegurança

⁹As principais diferenças sociodemográficas entre os quatro grupos podem ser observadas na tabela 5, em anexo.

- Tradicionalistas
- Moralmente conservadores
- Decepcionados com a política
- O grupo é a referência mais importante, preocupados com sua reputação na sociedade;
- Pouco críticos em relação à mídia.

Grupo 2 – “Meu governo, minha vida” (24%):

- Geralmente mais velhos
- Menor nível de escolaridade e renda familiar reduzida
- Maioria são casados ou possuem uma relação estável
- Concentrados em cidades pequenas e médias no interior do Nordeste brasileiro
- Maior parcela de beneficiários do programa Bolsa-Família
- Admitem e parecem aprovar uma relação de dependência em relação ao Governo
- Moralmente conservadores
- Satisfeitos com sua própria vida
- Demonstram satisfação com as conquistas sociais e econômicas do país
- Orgulhosos do Brasil e da nacionalidade
- Possuem sentimento de eficácia política e valorizam a escolha eleitoral
- Otimistas sobre seu futuro e sobre o futuro do país.

Grupo 3 – “Livres e Antenados” (21%):

- Eleitores mais jovens
- Maior nível de escolaridade
- Pertencentes a domicílios com maior renda familiar
- Residentes nas principais cidades do Sudeste e Sul do país
- Grande parte solteiros e sem filhos
- Declaram não possuir religião ou ateus

- Forte participação em programas sociais que não o Bolsa-Família. (provavelmente programas de distribuição de remédios com baixo custo ou gratuita, programas de financiamento habitacional, programas de financiamento do ensino superior, etc.)
- Modernos, tem alguns traços de valores pós-materialistas
- Defendem as liberdades individuais e são avessos ao autoritarismo
- Valorizam o prazer, hedonistas
- Abertos a novas experiências
- Gostam de mudanças
- Antenados("hype", "cool"), porém críticos
- Parecem ser mais envolvidos, participativos, em vários aspectos e temas

Grupo 4 – “Cada cabeça, uma sentença” (36%):

- O maior grupo atitudinal
- Normalmente vivem no interior do país
- Católicos
- Não-beneficiários dos programas sociais do Governo Federal
- Egoístas, individualistas
- Não se comprometem
- Revelam pragmatismo; orientam-se por ganhos tangíveis e por seus próprios interesses
- São liberais em termos morais
- Demonstram possuir um autoritarismo circunstancial, se isso os favorece
- Pouco participativos
- E imunes ao discurso nacionalista

5.2.2 RELAÇÃO DOS GRUPOS COM A INTENÇÃO DE VOTO PARA PRESIDENTE NA ELEIÇÃO DE 2010

A análise da intenção de voto para presidente, por grupo atitudinal, mostrou que a maioria dos eleitores do Grupo 1 – “Apegados” pretendia eleger o candidato da oposição José Serra, enquanto que no Grupo 2 - “Meu governo, minha vida”

maior parte do eleitorado declarou intenção de votar na candidata Dilma Rousseff, apoiada pelo presidente Lula, demonstrando seu desejo de continuidade do governo, o qual havia sido bem avaliado por este grupo (conforme Tabela 13, em anexo).

Por outro lado, a intenção de voto para os candidatos José Serra e Dilma Rousseff mostrou-se mais equilibrada entre os eleitores do Grupo 4 - “Cada cabeça, uma sentença”, o qual apresenta maior percentual de respondentes que pretendiam anular o voto ou votar em branco.

O Grupo 3 – “Livres e antenados”, por sua vez, apresentou uma tendência similar àquela observada para o total dos eleitores entrevistados.

Esses resultados podem ser conferidos na Tabela 2 abaixo:

Tabela 2 – Intenção de voto no 1º turno das eleições presidenciais de 2010

| | Grupos | | | | |
|--------------------------------|------------|-----------------|-----------------------------------|------------------------------|--|
| | TOTAL | 1 “Apegados” | 2 “Meu governo, minha vida” | 3 “Livres e Antenados” | 4 “Cada cabeça, uma sentença” |
| Dilma Rousseff (PT) | 35% | 28% | 42% | 34% | 36% |
| José Serra (PSDB) | 38% | 46% | 36% | 41% | 33% |
| Marina Silva (PV) | 7% | 8% | 5% | 9% | 5% |
| Branco/ Nulo | 11% | 9% | 8% | 10% | 14% |
| Não sabe/ Não respondeu | 9% | 9% | 9% | 6% | 12% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

A análise desses resultados mostrou que no contexto eleitoral de 2010, apenas os aspectos atitudinais abordados na pesquisa não foram suficientes para explicar a intenção de voto nos principais candidatos.

6 METODOLOGIA

Para verificar o efeito do uso da técnica análise fatorial na segmentação dos eleitores brasileiros no estudo *Cenário Eleitoral 2010* foi realizada uma análise de agrupamentos em SPSS¹⁰ considerando como *inputs* do processo as 69 variáveis originais (padronizadas por respondente, para evitar vício de respostas) ao invés dos 4 escores fatoriais mencionados na seção 5.2. Para isso, foram utilizadas as mesmas especificações utilizadas na segmentação com base na análise fatorial, apresentada por aquele estudo.

O algoritmo de agrupamento utilizado foi o k-means (não-hierárquico) com a distância euclidiana como medida de similaridade e número de grupos igual a 4. A abordagem utilizada para designar indivíduos aos aglomerados foi a otimização e as sementes iniciais de agrupamento correspondente as primeiras observações da base de dados.

O tratamento para casos faltantes em ambas as segmentações foi o mesmo: exclusão de casos quando não havia nenhuma resposta válida para as variáveis de agrupamento. No entanto, a Análise Fatorial aplicada na primeira segmentação foi realizada definindo que os dados faltantes fossem substituídos pela média das demais observações, o que fez com que todos os indivíduos da base apresentassem valores válidos para cada escore fatorial e a segmentação fosse processada com 2.002 casos válidos. A segunda segmentação, por sua vez, como foi executada com as variáveis originais, contou com 1.985 casos válidos, já que 17 indivíduos se negaram a responder às baterias de frases atitudinais.

Obtida a segmentação com as variáveis originais, construiu-se uma tabela cruzada com as classificações obtidas pelas duas segmentações, com e sem o uso da análise fatorial, para avaliar a concordância entre as duas abordagens. Para isso, foram utilizados a estatística Kappa¹¹ e o teste de McNemar¹².

¹⁰ SPSS (Statistical Package for the Social Sciences) é um software estatístico utilizado por profissionais de diversas áreas.

¹¹ A estatística Kappa é um coeficiente que mede a concordância entre duas ou mais medidas subtraindo da proporção global de concordância observada a proporção global de concordância esperada ao acaso (Nunes *et al.*, 2004).

¹² Teste de McNemar é um teste de concordância que testa a hipótese nula de que as discordâncias observadas são fruto do acaso (<http://leg.ufpr.br/~silvia/CE055/node83.html>).

A análise discriminante também foi utilizada a fim de verificar se as variáveis que melhor discriminavam os grupos coincidiam entre as duas abordagens. Além disso, comparou-se também o percentual de classificação correta fornecido pelos modelos construídos com as variáveis que apresentaram significância estatística para discriminar os grupos.

A caracterização atitudinal dos grupos foi feita com a comparação do valor médio das respostas de cada atributo (frases atitudinais) dentro dos grupos em relação à média de respostas no total da amostra, além da análise de outras perguntas do questionário. Já a identificação do perfil sociodemográfico se deu com a análise da distribuição das variáveis sociodemográficas em cada um dos grupos, também comparada com o perfil geral dos respondentes.

Finalmente, analisaram-se as características atitudinais e sociodemográficas dos grupos formados pela segmentação com base nas variáveis originais e compararam-se os perfis dos grupos obtidos pelas duas abordagens a fim de avaliar as possíveis diferenças reveladas por elas. Também se comparou a relação dos grupos, gerados pelas duas propostas, com a intenção de voto para presidente na eleição de 2010.

7 APLICAÇÕES E RESULTADOS

A segmentação com base nas 69 variáveis originais realizada com 1.985 casos válidos resultou em 4 grupos de eleitores que podem ser relacionados diretamente aos grupos obtidos com a segmentação com base na Análise Fatorial, devido à alta concordância entre as duas classificações, como pode ser observado na tabela abaixo e nos resultados obtidos com a estatística Kappa ($k=0,684$, classificada como substancial¹³) e o Teste de MC Nemar (p -valor=0,000, resultando na rejeição da hipótese H_0 : não há concordância entre as classificações).

Tabela 3 – Concordância entre as Segmentações com base na Análise Fatorial e com base nas Variáveis Originais

| | | Grupos da 2ª Segmentação (com Variáveis Originais) | | | | Total |
|--|-----------------------------|---|-----------------------------|-----------------------------|-----------------------------|-------|
| | | 2 | 3 | 1 | 4 | |
| Grupos da 1ª Segmentação (com Análise Fatorial) | “Apegados” | 217 (60,4%) ^a | 92 | 42 | 22 | 373 |
| | “Meu governo, minha vida” | 21 | 354 (76,0%) ^a | 84 | 28 | 487 |
| | “Livres e antenados” | 80 | 12 | 300 (69,1%) ^a | 21 | 413 |
| | “Cada cabeça, uma sentença” | 41 | 8 | 8 | 655 (90,2%) ^a | 712 |
| | Total | 359 | 466 | 434 | 726 | 1985 |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

^a estes percentuais foram calculados sobre os eleitores classificados de acordo com a 2ª segmentação (com variáveis originais)

A comparação entre os resultados das Análises Discriminantes processadas para ambas as classificações, utilizando o mesmo conjunto de variáveis (as 69 variáveis utilizadas nas análises de segmentação), mostrou que 2/3¹⁴ das variáveis identificadas como as que mais discriminam os grupos classificados pela 2ª segmentação também foram apontadas como relevantes para segmentar os eleitores classificados de acordo com a 1ª segmentação. O percentual de

¹³ Para avaliar o valor de Kappa, utilizou-se a classificação proposta por Landis & Koch (*apud* Nunes et al., (2004)): concordância quase perfeita (0,81-1,00), substancial (0,61-0,80), moderada (0,41-0,60), sofrível (0,21-0,40) e leve (0,00-0,20).

¹⁴ Das 21 variáveis identificadas como sendo as mais discriminantes para a classificação obtida com a segmentação com base nas Variáveis Originais, 14 também foram identificadas como tal na Análise Discriminante realizada para a classificação obtida com a segmentação com base na Análise Fatorial.

classificação correta fornecido pelos modelos resultantes das Análises Discriminantes foi de 83,3% para a classificação obtida com a 1ª segmentação e de 85% para a classificação obtida com a 2ª segmentação.

A análise do perfil sociodemográfico e atitudinal dos grupos, em relação às dimensões investigadas, resultantes da 2ª segmentação, não demonstrou nenhuma contradição com respeito às características definidas para os grupos equivalentes da 1ª segmentação. Entretanto, quando se comparam os resultados obtidos com ambas as segmentações, observa-se que alguns aspectos se tornaram mais evidentes e outros menos nos grupos da segmentação com base nas Variáveis Originais, os quais são destacados a seguir:

Grupo 2 (18%) – Equivalente ao “Apegados”:

- ❖ Maior representação na região Sudeste e maior concentração em municípios com mais de 100.000 habitantes.
- ❖ Maior descontentamento com o país.
- ❖ Menos engajados.
- ❖ Menor desejo de continuidade.
- ❖ Menos informados.

As tabelas 5, 6, 7, 8 e 9, em anexo, apresentam alguns resultados que evidenciam as conclusões acima. Com exceção da tabela 5, as demais foram construídas com base nas variáveis de agrupamento.

Este grupo foi o que apresentou maior variação nas respostas relativas às dimensões atitudinais, isso porque, comparado com os demais, ele foi o que teve menor percentual de membros do grupo equivalente na 1ª segmentação, apenas 60,4%.

No entanto, as diferenças observadas foram de encontro com o perfil dos “Apegados” definido na segmentação com base na Análise Fatorial. Pode-se notar que algumas características definidas para este grupo foram reforçadas nos resultados obtidos com a 2ª segmentação.

Grupo 3 (23%) – Equivalente ao “Meu governo, minha vida”:

- ❖ Maioria do sexo masculino e maior proporção de evangélicos/protestantes.
- ❖ Mais contrários à igualdade dos direitos dos homossexuais.

Os resultados das tabelas 5 e 10, em anexo, evidenciam estas diferenças. A variável da tabela 10 foi utilizada para segmentar os eleitores.

Grupo 1 (22%) – Equivalente ao “Livres e antenados”:

- ❖ Maior proporção de mulheres e eleitores com filhos.
- ❖ Maior consciência do poder do voto.
- ❖ Maior percepção das avanços sociais e econômicos do país.

As tabelas 5, 11 e 12, em anexo, apresentam alguns resultados que evidenciam as constatações acima. Vale mencionar que as tabelas 11 e 12 foram construídas com base nas variáveis de agrupamento.

Grupo 4 (37%) – Equivalente ao “Cada cabeça, uma sentença”:

Este grupo não apresentou nenhuma diferença relevante em seu perfil quando comparado com o grupo equivalente da 1ª segmentação, o que é justificado pela presença de 90,2% dos eleitores classificados no grupo “Cada cabeça, uma sentença” da 1ª segmentação.

É importante notar que, apesar das variações percentuais dos níveis de concordância entre os grupos equivalentes das duas segmentações, os padrões de respostas e perfis sociodemográficos, de maneira geral, se mantiveram. Por exemplo, o Grupo “Apegados” na 1ª segmentação foi caracterizado como tendo uma maior representação na região Sul, 25% contra a média dos eleitores que é de 15%, enquanto que o percentual de residentes na região Sudeste foi de 41%, contra 44% da média geral. Com a 2ª segmentação, apesar de a região Sudeste ter tido maior representação no grupo equivalente ao “Apegados”, 57% contra a média de 44%, e a região Sul menor representação, 18% contra 15% da média, o padrão “maioria reside na região Sudeste, seguida das regiões Sul e Nordeste” permaneceu, conforme tabela 5, em anexo. Isso também ocorre com as características atitudinais.

A comparação da intenção de voto para presidente, por grupo atitudinal, entre as duas propostas de segmentação, mostra que a principal diferença está no equilíbrio entre o percentual de eleitores que pretendiam votar nos candidatos José Serra (38%) e Dilma Rousseff (39%) no grupo 3 da segmentação com base nas variáveis originais, correspondente ao grupo “Meu governo, minha vida” da 1ª segmentação, conforme tabela abaixo:

Tabela 4 – Intenção de voto no 1º turno das eleições presidenciais de 2010

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|------------------------------------|--|-----|------------|-----|-----|---|-----|------------|-----|-----|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Dilma Rousseff | 35% | 28% | 42% | 34% | 36% | 35% | 26% | 39% | 36% | 38% |
| José Serra | 38% | 46% | 36% | 41% | 33% | 38% | 47% | 38% | 40% | 33% |
| Marina Silva | 7% | 8% | 5% | 9% | 5% | 7% | 5% | 6% | 11% | 5% |
| Branco/ Nulo | 11% | 9% | 8% | 10% | 14% | 11% | 13% | 8% | 8% | 14% |
| Não sabe/ Não respondeu | 9% | 9% | 9% | 6% | 12% | 9% | 9% | 9% | 6% | 11% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

A diferença apontada pode ser justificada pelo maior percentual de eleitores neste grupo que declararam se sentir incomodados em “votar para presidente numa pessoa que está concorrendo numa eleição pela primeira vez” quando comparado com o mesmo percentual de eleitores do grupo “Meu Governo, minha vida” da 1ª segmentação, como mostra a tabela 13 em anexo.

Além disso, a proximidade da intenção de voto dos candidatos José Serra e Dilma Rousseff nestes grupos pode ter sido resultado do período em que foi realizada a pesquisa, março de 2010, onde a candidata Dilma ainda não era amplamente conhecida pelos eleitores, principalmente aqueles com menor grau de escolaridade, característica marcante no grupo “Meu governo, minha vida” da 1ª segmentação e mais expressivo no grupo equivalente da 2ª segmentação (conforme tabela 5, em anexo).

Os demais grupos da 2ª segmentação seguiram com a mesma tendência observada na 1ª segmentação quanto à intenção de voto: “Apegados”, com a maioria (mais de 45%) favorável ao candidato da oposição, José Serra; “Livres e antenados”, com uma distribuição similar àquela observada para o total dos eleitores entrevistados; e “Cada cabeça, uma sentença”, com o menor percentual de eleitores

com intenção de votar no candidato José Serra, quando comparado com os demais grupos.

A análise da relação entre os grupos atitudinais da segmentação com base nas variáveis originais e a intenção de voto para presidente em 2010 também sugere que apenas os aspectos atitudinais abordados na pesquisa não foram suficientes para explicar a intenção de voto nos principais candidatos naquele contexto eleitoral.

8 CONCLUSÕES

A análise comparativa dos resultados obtidos com as duas segmentações de eleitores, com e sem o uso da Análise Fatorial, considerando as mesmas variáveis e técnica de agrupamento, mostrou que os grupos formados por ambas as propostas possuem perfis atitudinais e sociodemográficos equivalentes.

As diferenças detectadas entre as duas classificações, descritas na seção anterior, ressaltaram alguns aspectos dos grupos identificados e minimizaram outros, porém, não foram suficientes para contradizer as características definidas na 1ª segmentação, com base na Análise Fatorial, apresentada no estudo *Cenário Eleitoral 2010*.

Ainda assim, o resultado empírico aqui apresentado não é suficiente para afirmar a superioridade de uma abordagem em detrimento da outra. Uma investigação teórica mais profunda e a realização de outros estudos comparativos, inclusive utilizando outros métodos de agrupamento (medidas de similaridade, algoritmo de aglomeração, etc.), são necessárias para um maior amadurecimento dos ganhos e perdas que o uso da Análise Fatorial pode ocasionar na análise de segmentação.

Na prática, as principais desvantagens em utilizar os escores fatoriais resultantes da Análise Fatorial como *inputs* em um projeto de segmentação é o tempo adicional que deve ser despendido para selecionar um número parcimonioso, representativo e interpretável de fatores a utilizar. Além disso, ainda que um conjunto de fatores que responda por um percentual baixo de variabilidade dos dados proporcione resultados semelhantes aos obtidos com o uso de variáveis originais, como visto na comparação realizada no presente estudo, o uso dessa informação pode gerar menor confiança quanto aos resultados alcançados por parte do pesquisador e daqueles que farão uso das análises.

A análise prévia da correlação existente entre as variáveis selecionadas para agrupar casos ou indivíduos, assim como o uso da análise fatorial e discriminante, além de outras técnicas, podem ser úteis na tarefa de identificar aquelas variáveis que são relevantes para discriminar os grupos e descrever com maior precisão a estrutura dos dados.

Cabe lembrar que as decisões tomadas em cada etapa de um projeto de segmentação – contextualização do problema, elaboração do questionário, seleção

8. CONCLUSÕES

da amostra, aplicação da técnica de agrupamento - podem impactar a consistência e utilidade dos resultados finais. Por isso, o pesquisador deve ter consciência de cada escolha feita, atuar com criticidade, utilizar diferentes métodos, realizar diferentes provas, comparar os resultados e, sempre que possível, validá-los com base em estudos anteriores e teorias conhecidas.

9 ANEXOS

Tabela 5 – Principais diferenças sociodemográficas dos grupos atitudinais

| Principais características sociodemográficas | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|---|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Gênero - Masculino | 48% | 52% | 47% | 47% | 48% | 48% | 51% | 52% | 41% | 49% |
| Idade – 16 a 29 anos | 34% | 29% | 25% | 43% | 36% | 33% | 31% | 22% | 43% | 36% |
| Idade – 50 anos ou mais | 26% | 29% | 31% | 17% | 26% | 26% | 24% | 37% | 16% | 25% |
| Escolaridade - até Primária Completa | 29% | 32% | 39% | 12% | 29% | 28% | 28% | 44% | 11% | 29% |
| Escolaridade – Superior em diante | 14% | 13% | 10% | 24% | 12% | 14% | 13% | 8% | 25% | 13% |
| Região do país - Nordeste | 27% | 18% | 42% | 17% | 28% | 27% | 14% | 38% | 22% | 28% |
| Região do país - Sudeste | 44% | 41% | 30% | 63% | 44% | 44% | 57% | 29% | 51% | 45% |
| Região do país - Sul | 15% | 25% | 8% | 12% | 15% | 15% | 18% | 14% | 13% | 15% |
| Capital | 27% | 30% | 28% | 34% | 22% | 27% | 28% | 29% | 36% | 20% |
| Periferia | 13% | 9% | 9% | 21% | 14% | 13% | 11% | 9% | 17% | 15% |
| Interior | 59% | 61% | 63% | 46% | 64% | 59% | 61% | 62% | 47% | 64% |
| Municípios com até 20.000 habitantes | 16% | 21% | 23% | 5% | 16% | 17% | 17% | 24% | 8% | 17% |
| Municípios com 20.000 a 100.000 habitantes | 26% | 19% | 33% | 21% | 29% | 26% | 17% | 30% | 24% | 29% |
| Municípios com mais de 100.000 habitantes | 57% | 60% | 45% | 74% | 55% | 58% | 66% | 46% | 68% | 55% |
| Renda familiar > 5 salários mínimos | 12% | 14% | 9% | 19% | 10% | 13% | 13% | 8% | 20% | 11% |
| Renda familiar até 2 salários mínimos | 47% | 44% | 61% | 37% | 45% | 47% | 42% | 61% | 40% | 45% |
| Beneficiários do Bolsa-família | 16% | 17% | 24% | 9% | 15% | 16% | 14% | 23% | 13% | 14% |
| Beneficiários de outros programas sociais do Governo Federal | 5% | 6% | 5% | 10% | 2% | 6% | 6% | 6% | 9% | 3% |
| Não-beneficiários de programas sociais do Governo Federal | 95% | 94% | 94% | 94% | 98% | 95% | 94% | 94% | 94% | 96% |
| Aproximadamente 50% ou menos da renda familiar é recebida do Estado | 23% | 23% | 30% | 16% | 17% | 20% | 21% | 29% | 19% | 15% |
| Estado civil - solteiro | 31% | 28% | 23% | 38% | 33% | 31% | 33% | 19% | 37% | 33% |
| Estado civil - casado | 59% | 60% | 64% | 53% | 59% | 59% | 57% | 67% | 52% | 58% |
| Com filhos em idade escolar (7 a 22 anos) | 32% | 36% | 36% | 31% | 28% | 32% | 36% | 35% | 32% | 27% |
| Sem filhos | 40% | 37% | 32% | 46% | 42% | 39% | 39% | 32% | 42% | 43% |
| Religião - Católica | 66% | 61% | 67% | 63% | 70% | 66% | 64% | 62% | 67% | 69% |
| Religião - Evangélico/ Protestante | 21% | 28% | 24% | 16% | 18% | 21% | 26% | 30% | 13% | 18% |
| Ateu/ Sem religião | 9% | 7% | 7% | 13% | 9% | 9% | 8% | 6% | 12% | 10% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

Tabela 6 – “Se pudesse iria embora do Brasil”

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|-------------------------|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Concorda | 28% | 23% | 12% | 24% | 44% | 28% | 34% | 6% | 21% | 44% |
| (4) | 15% | 7% | 6% | 11% | 27% | 15% | 12% | 3% | 9% | 27% |
| (5) Concorda totalmente | 13% | 16% | 6% | 13% | 17% | 13% | 22% | 3% | 12% | 17% |
| Discorda | 57% | 63% | 81% | 65% | 32% | 57% | 47% | 89% | 70% | 33% |
| (2) | 14% | 9% | 11% | 16% | 18% | 14% | 9% | 10% | 17% | 17% |
| (1) Discorda totalmente | 42% | 54% | 70% | 49% | 14% | 42% | 38% | 79% | 53% | 15% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

Tabela 7 – “Eu costumo conversar sobre a política e o governo com outras pessoas”

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|-------------------------|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Concorda | 46% | 40% | 46% | 42% | 53% | 46% | 27% | 50% | 44% | 55% |
| (4) | 25% | 18% | 25% | 18% | 32% | 25% | 14% | 26% | 19% | 33% |
| (5) Concorda totalmente | 22% | 22% | 21% | 24% | 21% | 22% | 13% | 24% | 25% | 23% |
| Discorda | 30% | 38% | 37% | 36% | 18% | 30% | 50% | 33% | 34% | 17% |
| (2) | 12% | 13% | 11% | 14% | 12% | 12% | 14% | 11% | 14% | 11% |
| (1) Discorda totalmente | 18% | 25% | 25% | 22% | 6% | 18% | 36% | 22% | 19% | 6% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

Tabela 8 – “Seria bom que com o próximo presidente o Brasil continuasse no rumo que vai”

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|-------------------------|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Concorda | 64% | 58% | 86% | 57% | 56% | 64% | 47% | 83% | 70% | 57% |
| (4) | 29% | 23% | 28% | 27% | 34% | 29% | 23% | 24% | 32% | 33% |
| (5) Concorda totalmente | 22% | 35% | 58% | 29% | 22% | 35% | 24% | 59% | 38% | 24% |
| Discorda | 14% | 18% | 4% | 21% | 14% | 14% | 27% | 4% | 12% | 14% |
| (2) | 7% | 10% | 2% | 9% | 9% | 7% | 12% | 2% | 6% | 9% |
| (1) Discorda totalmente | 6% | 8% | 2% | 12% | 4% | 6% | 14% | 2% | 6% | 5% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

Tabela 9 – “Costumo acompanhar as notícias nos diversos meios de comunicação como TV, radio, jornal, revistas etc. para ser uma pessoa com maior consciência social e política”

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|-------------------------|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Concorda | 68% | 74% | 75% | 71% | 59% | 68% | 59% | 77% | 80% | 60% |
| (4) | 29% | 21% | 29% | 26% | 35% | 29% | 21% | 28% | 27% | 35% |
| (5) Concorda totalmente | 39% | 53% | 45% | 45% | 24% | 39% | 38% | 49% | 52% | 25% |
| Discorda | 10% | 8% | 8% | 12% | 11% | 10% | 15% | 7% | 7% | 10% |
| (2) | 6% | 5% | 6% | 6% | 6% | 6% | 9% | 4% | 4% | 6% |
| (1) Discorda totalmente | 4% | 3% | 2% | 5% | 5% | 4% | 6% | 3% | 3% | 5% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

Tabela 10 – "O casamento entre homossexuais deveria ser permitido"

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|-------------------------|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Concorda | 37% | 19% | 16% | 63% | 45% | 37% | 24% | 3% | 66% | 48% |
| (4) | 15% | 7% | 8% | 16% | 24% | 15% | 10% | 1% | 17% | 25% |
| (5) Concorda totalmente | 22% | 12% | 9% | 47% | 21% | 22% | 13% | 2% | 48% | 23% |
| Discorda | 46% | 68% | 72% | 20% | 33% | 46% | 58% | 87% | 17% | 31% |
| (2) | 8% | 7% | 6% | 5% | 12% | 8% | 10% | 5% | 5% | 11% |
| (1) Discorda totalmente | 38% | 61% | 66% | 15% | 21% | 38% | 49% | 82% | 13% | 20% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

Tabela 11 – "O meu voto em um determinado candidato pode mudar muito a minha vida, tanto para melhor, quanto para pior"

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|-------------------------|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Concorda | 63% | 54% | 82% | 65% | 53% | 63% | 46% | 78% | 73% | 55% |
| (4) | 27% | 22% | 25% | 24% | 33% | 27% | 23% | 23% | 24% | 34% |
| (5) Concorda totalmente | 36% | 32% | 57% | 41% | 19% | 36% | 23% | 55% | 50% | 20% |
| Discorda | 13% | 20% | 5% | 11% | 15% | 13% | 23% | 7% | 8% | 14% |
| (2) | 7% | 5% | 3% | 6% | 10% | 7% | 8% | 3% | 4% | 9% |
| (1) Discorda totalmente | 6% | 15% | 1% | 5% | 5% | 6% | 15% | 3% | 4% | 4% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

Tabela 12 – “A gente vê que hoje em dia no Brasil os recursos do governo chegam para os mais pobres”

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|-------------------------|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Concorda | 55% | 44% | 79% | 43% | 52% | 55% | 39% | 73% | 54% | 53% |
| (4) | 32% | 25% | 37% | 29% | 35% | 32% | 23% | 34% | 33% | 34% |
| (5) Concorda totalmente | 23% | 19% | 42% | 14% | 18% | 23% | 15% | 39% | 21% | 19% |
| Discorda | 19% | 26% | 7% | 31% | 16% | 19% | 33% | 8% | 22% | 17% |
| (2) | 12% | 15% | 5% | 17% | 12% | 12% | 17% | 6% | 13% | 13% |
| (1) Discorda totalmente | 7% | 11% | 2% | 14% | 4% | 7% | 15% | 2% | 9% | 4% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

Tabela 13 – “Me incomoda votar para presidente numa pessoa que está concorrendo numa eleição pela primeira vez”

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|-------------------------|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Concorda | 34% | 40% | 23% | 18% | 46% | 34% | 31% | 27% | 20% | 47% |
| (4) | 20% | 17% | 14% | 11% | 31% | 20% | 17% | 13% | 11% | 31% |
| (5) Concorda totalmente | 14% | 23% | 10% | 7% | 15% | 14% | 14% | 14% | 8% | 16% |
| Discorda | 40% | 38% | 50% | 63% | 20% | 40% | 46% | 46% | 62% | 19% |
| (2) | 17% | 12% | 22% | 22% | 13% | 17% | 14% | 19% | 24% | 12% |
| (1) Discorda totalmente | 23% | 26% | 29% | 41% | 7% | 23% | 32% | 27% | 38% | 7% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

Tabela 14 – Avaliação do presidente Lula

| | Grupos (Segmentação com Análise Fatorial) | | | | | Grupos (Segmentação com Variáveis Originais) | | | | |
|-----------------|--|------------|------------|------------|------------|---|------------|------------|------------|------------|
| | TOTAL | 1 | 2 | 3 | 4 | TOTAL | 2 | 3 | 1 | 4 |
| Positiva | 75% | 73% | 87% | 67% | 73% | 75% | 66% | 85% | 73% | 74% |
| Ótima | 28% | 23% | 40% | 25% | 25% | 28% | 20% | 39% | 29% | 25% |
| Boa | 47% | 50% | 47% | 42% | 47% | 47% | 46% | 46% | 44% | 49% |
| Regular | 20% | 23% | 10% | 27% | 20% | 20% | 27% | 13% | 24% | 18% |
| Negativa | 5% | 4% | 2% | 6% | 7% | 5% | 6% | 2% | 3% | 8% |
| Ruim | 3% | 2% | 1% | 3% | 4% | 3% | 4% | 1% | 1% | 4% |
| Péssima | 2% | 2% | 1% | 3% | 3% | 2% | 3% | 1% | 2% | 3% |

Fonte: Elaborada pela autora a partir dos dados da pesquisa realizada pelo IBOPE Inteligência.

10 REFERÊNCIAS BIBLIOGRÁFICAS

FREI, Fernando. **Introdução à análise de agrupamentos: Teoria e prática**. 1ª edição. São Paulo: Editora UNESP, 2006.

GREEN, Paul E.; TULL, Donald S. **Research for marketing decisions**. 3ed. Englewood Cliffs: Prentice-Hall Inc. 1975

HAIR, Joseph F.; ANDERSON, Rolph E.; TATHAM, Ronaldo L.; BLACK William C. **Análise Multivariada de Dados**. Porto Alegre: Bookman, 2005.

MALHOTRA, Naresh K. **Pesquisa de Marketing: Uma Orientação Aplicada**. 3ª edição. Porto Alegre: Bookman, 2001.

MYERS, James H. **Segmentation and positioning for strategic marketing decisions**. Chicago: American Marketing Association, 1996.

NUNES, Jacira et al . **Confiabilidade e validade das declarações de óbitos por câncer de útero no município de Belém, Pará, Brasil**. Cad. Saúde Pública, Rio de Janeiro, v. 20, n. 5, Oct. 2004. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-311X2004000500020&lng=en&nrm=iso>. Acesso em 21/11/2011.

NUNES, M. C., Resende, J. R., Cervellini, S., Giani, M. **2010 Electoral Scenario: Changes and Continuities in the post-Lula Brazilian Electoral Behavior**. In: WAPOR 63rd ANNUAL CONFERENCE, Chicago, 2010.

RICHARDSON, Roberto J. **Pesquisa Social: métodos e técnicas**. 3ª edição. São Paulo: Atlas, 2011.

Yeung, J. Y.; Ruzzo, W. L. **Principal Component Analysis for clustering gene expression data, 2001**. Disponível em <<http://bioinformatics.oxfordjournals.org/content/17/9/763.full.pdf+html>>. Acesso em 4/7/2011.

SITES

Desk Research. Multifocus. Disponível em <<http://www.multifocus.com.br>>. Acesso em 21/11/2011.

Teste de McNemar. Notas de aula. Disponível em <<http://leg.ufpr.br/~silvia/CE055/node83.html>>. Acesso em 25/11/2011.