



Análise multivariada para avaliação do comportamento de grupos supermercadistas brasileiros¹

Multivariate analysis for determine the behaviors of brazilian supermakets groups

Patrícia Prado Belfiore²

Luiz Paulo Lopes Fávero³

Claudio Felisoni de Angelo⁴

Resumo

O presente trabalho analisa o comportamento de grupos supermercadistas no Brasil no ano de 2004, por meio da aplicação de técnicas estatísticas multivariadas de análise fatorial, análise de conglomerados e análise discriminante. Para a realização deste estudo, utilizaram-se dados relativos ao faturamento, número de *check-outs*, área total das lojas, número de lojas, número total de funcionários, faturamento por *check-out*, faturamento por m², faturamento por funcionário, número de funcionários por *check-out*, número de funcionários por 100 m², número de *check-outs* por loja, m² por *check-out* e m² por loja, referentes aos 100 maiores grupos supermercadistas no Brasil, segundo a revista SuperHiper 2005. Os grupos estão classificados em um *ranking* pelo seu faturamento e o tratamento dos dados se deu por meio do software estatístico SPSS (*Statistical Package for Social Sciences*).

Palavras-chave: Varejo, análise fatorial, análise de conglomerados, análise discriminante.

Abstract

The present paper focuses in determining behaviors in supermarket groups, in the last year in Brazil, using multivariate analyses as factor analysis, cluster analysis and discriminant analysis. In order to study this, it was collected data such as revenues, *check-out* numbers, total area of stores, number of stores, total number of employees, revenues per *check-out*, revenues per area, revenues per employees, number of employees per *check-out*, number of employees per 100 m², *check-out* numbers per store, m² per *check-out* and m² per store, of the 100 largest supermarket groups in Brazil, according to the magazine SuperHiper 2005. The groups are ranked by their total revenues and the data was dealt with through the statistics software SPSS (*Statistical Package for Social Sciences*).

Keywords: Retail, factor analysis, cluster analysis, discriminant analysis.

¹ Recebido em 21 de setembro de 2005, aprovado em 11 de fevereiro de 2006.

² Doutora em Engenharia de Produção pela Escola Politécnica da Universidade de São Paulo. Professora do Departamento de Engenharia de Produção da Faculdade de Engenharia Industrial. E-mail: patricia.belfiore@poli.usp.br

³ Professor Doutor do Departamento de Contabilidade da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo. E-mail: lpfavero@usp.br

⁴ Professor Titular do Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo. E-mail: cfa@usp.br



Introdução

Na economia brasileira atual, de forma semelhante ao que acontece em alguns países desenvolvidos, o setor terciário e de serviços apresenta-se como o de maior importância, representando cerca de 50% do produto interno bruto. Entre os vários serviços ofertados por esse setor, destaca-se o comércio, mais especificamente, o varejo, uma atividade cuja relevância não decorre apenas de sua enorme expressão econômica, mas também por ser o elo final da cadeia de abastecimento. Ademais, por manter um contato direto com os consumidores, serve como um indicador das tendências do mercado de consumo e da distribuição de bens em geral.

Durante muitas décadas, os aspectos operacionais relacionados à administração varejista foram negligenciados e as vastas e rentáveis possibilidades oferecidas pelo mercado financeiro permitiam que as empresas do setor compensassem com grande facilidade seus problemas operacionais (BLECHER, 2001, p. 32). Por outro lado, a estabilidade econômica alterou significativamente a forma de atuação de grupos varejistas e reestruturou o mercado de consumo. A nova realidade econômica, aliada ao avanço tecnológico, à globalização e ao surgimento de novos formatos, alterou significativamente a forma de o varejo operar, abrindo espaço para questionamentos e aprimoramento de seu gerenciamento.

As transformações descritas estimularam o aumento de tamanho das empresas supermercadistas (SESSO FILHO, 2001, p. 22), porém o modo como as empresas estão agrupadas em função de determinados indicadores ainda desperta dúvida. Dessa forma, o desenvolvimento deste trabalho utiliza os dados do *ranking* da Revista SuperHiper 2005 para analisar o comportamento das 100 maiores redes supermercadistas, classificadas por seu faturamento no ano de 2004.

Primeiramente, aplica-se análise fatorial para identificação de fatores ou associações entre as variáveis observacionais analisadas (faturamento, número de *check-outs*, área total das lojas, número de lojas, número total de funcionários, faturamento por *check-out*, faturamento por m², faturamento por funcionário, número de funcionários por *check-out*, número de funcionários por 100 m², número de *check-outs* por loja, m² por *check-out* e m² por loja), o que facilita a

interpretação dos dados. Com o intuito de tornar possível a determinação de comportamentos semelhantes de atuação em grupos supermercadistas no Brasil, em função dos fatores obtidos, utiliza-se o método multivariado de análise de conglomerados. A análise discriminante tem como objetivo determinar as variáveis que discriminam os grupos, validar a análise de *cluster* e confirmar os resultados da análise fatorial. O tratamento dos dados se deu por meio do software estatístico SPSS (*Statistical Package for Social Sciences*).

Embora essa reflexão possa ser aplicada a todos os segmentos que compõem os setores produtivos e prestadores de serviços no Brasil, não há dúvidas de que o ramo composto pelos supermercados mereça atenção especial (ANGELO, SILVEIRA e SIQUEIRA, 2001, p. 3). Atualmente, grande parte da comercialização de bens no Brasil realiza-se por meio desse tipo de comércio e os supermercados vêm expandindo seus ramos de atuação e ampliando o sortimento dos produtos oferecidos. Suas atividades, outrora restritas à comercialização de produtos alimentares e de higiene e limpeza, estão sendo expandidas para outros segmentos, como os relacionados a bens duráveis, semiduráveis e têxteis, entre outros.

A estrutura do artigo está descrita a seguir. O item 2 mostra a evolução e o panorama do setor supermercadista no Brasil. O item 3 descreve o método empregado no estudo e a revisão teórica dos conceitos. Já o item 4 apresenta e analisa os resultados das técnicas multivariadas implementadas. Finalmente, no item 5, estão as conclusões e futuras pesquisas.

Evolução e panorama do setor supermercadista no Brasil

Com a política de redução drástica do ritmo inflacionário, introduzida em 1994 com o Plano Real, os supermercados estão sendo expostos a desafios cada vez maiores no sentido de atingirem a excelência em operações e produtividade (MARQUES e NORONHA, 2001, p. 177). Antes de 1994, as empresas compensavam seus problemas operacionais com os ganhos obtidos por meio da aplicação financeira, que, quando acompanhava a inflação, podia atingir patamares de até 80% ao mês. Com a estabilidade dos patamares inflacionários, os grupos supermercadistas deixaram de lucrar com as aplicações financeiras. A necessidade de ofertar produtos e serviços de melhor qualidade,

diferenciados e a preços coerentes e competitivos, obrigou os supermercadistas a se preocuparem com as operações e com a oferta. Assim, a satisfação do consumidor passa a ser extremamente importante para o bom desempenho dos grupos supermercadistas (MARQUES e NORONHA, 2001, p. 177).

Em praticamente cinco décadas de implantação no Brasil, o supermercado tornou-se um elemento indissociável do modo de vida e das preocupações de qualquer consumidor. Os principais supermercados foram instalados no país a partir dos anos 60 e, na década seguinte, houve sedimentação das grandes cadeias, inclusive com a abertura dos primeiros hipermercados (MARQUES e NORONHA, 2001, p. 177-178).

A década de 80 foi caracterizada por instabilidades políticas e econômicas no país, enquanto os anos 90 foram marcados por uma enorme competição no setor, inclusive com o aparecimento de grandes aquisições e expansões realizadas por alguns grupos supermercadistas (ROJO, 1998, p. 10-20). Na transição dos anos 90 para o novo século, esse processo foi muito intenso,

tanto pela incorporação de várias redes nacionais pelas estrangeiras, como também pela aquisição das redes de pequeno e médio portes pelas redes nacionais maiores, resultando em uma mudança estrutural do mercado, com a presença de maiores concentrações nessa área de varejo (GHISI, CAMARGO e MARTINELLI, 2003, p. 219). Como empresas varejistas estrangeiras, podemos citar o Wal-Mart (norte-americano), o Carrefour (francês), o Grupo Jerônimo Martins (português), o Grupo Royal Ahold (holandês) e o Sonae (português). O ano de 2002 foi de acomodação das aquisições antes realizadas pelas empresas do setor (SUPERHIPER, 2003, p. 66-67).

De acordo com a Revista SuperHiper (2005), os 20 maiores grupos supermercadistas presentes no Brasil apresentam um faturamento bruto de 49,954 bilhões de reais, o que equivale, em valores, a 3% do PIB nacional (crescimento de 4,8%). A tabela 1 apresenta a evolução do faturamento desses grupos nos últimos 10 anos. Os valores apresentados estão em moeda corrente, ou seja, foram traduzidos para a base atual pela variação do IPCA.

Tabela 1. Evolução do faturamento bruto (R\$) dos 20 maiores grupos supermercadistas brasileiros no período de 1995-2004

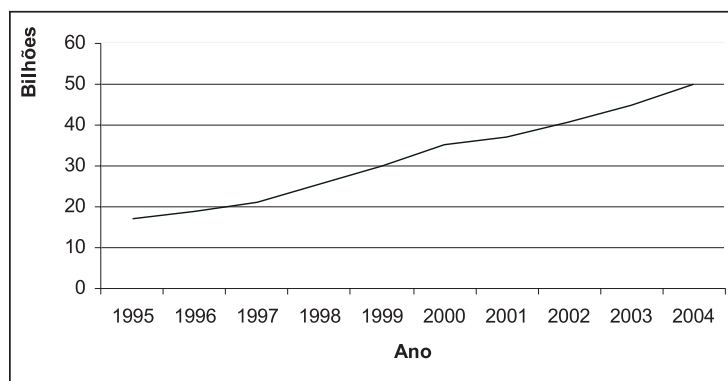
Ano	Faturamento bruto dos 20 maiores grupos supermercadistas (R\$)
1995	17.099
1996	18.929
1997	20.966
1998	25.539
1999	29.987
2000	35.135
2001	37.055
2002	40.609
2003	44.758
2004	49.954

Fonte: Revista SuperHiper, 2005.



A evolução histórica do faturamento bruto dos 20 maiores grupos supermercadistas presentes no Brasil é apresentada no Gráfico 1.

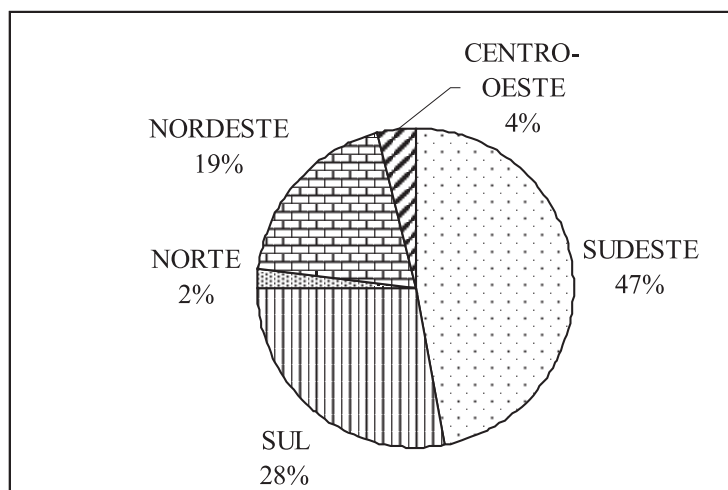
Gráfico 1. Evolução temporal do faturamento dos 20 maiores grupos supermercadistas no Brasil



A amostra dos grupos supermercadistas coletada para este estudo compõe um faturamento bruto em torno de R\$ 65 bilhões em 2004: os dez maiores grupos concentram 68% do faturamento total da amostra e os 20 maiores representam 77% da amostra. Além da concentração do faturamento bruto em poucas

empresas, ocorre também uma concentração do número de empresas supermercadistas, dentro do grupo das 100 maiores, que possuem sede nas regiões sudeste e sul do Brasil. No gráfico 2, são apresentados os percentuais do número de empresas, pertencentes ao grupo das 100 maiores, com sede nas cinco regiões geográficas do país.

Gráfico 2. Percentual do número de empresas com sede em cada uma das cinco regiões geográficas do País

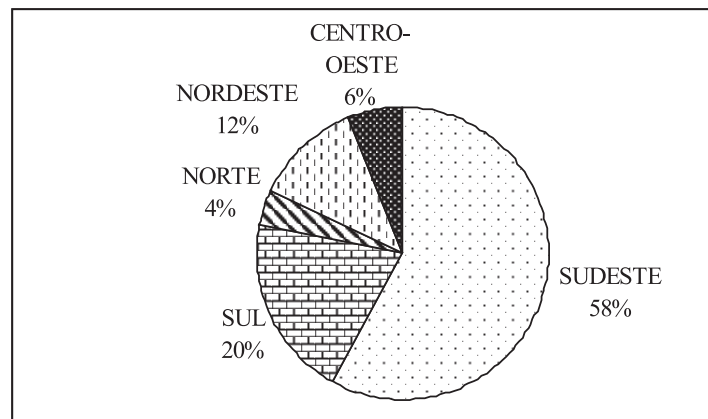




Além disso, a concentração é ainda maior para os estados do sudeste, quando o cálculo deixa de ser efetuado pelo número de empresas e passa a ser feito em relação ao faturamento dessas empresas,

segmentadas por região em que se localiza a sede. O gráfico 3 mostra os percentuais de faturamento das empresas pertencentes ao grupo das 100 maiores, em função da região geográfica em que se localiza a sede.

Gráfico 3. Percentual do faturamento com sede em cada uma das cinco regiões geográficas do País



Dessa forma, faz-se necessário um estudo mais profundo sobre a formação de eventuais *clusters*, em função de indicadores secundários, tornando possível identificar características comuns de cada agrupamento

de empresa no ano de 2004.

O Quadro 1 apresenta o *ranking* dos 100 maiores supermercadistas brasileiros no ano de 2004, de acordo com a Revista SuperHiper 2005.

Quadro 1. *Ranking* dos supermercadistas brasileiros no ano de 2004 em ordem decrescente de faturamento

1. Cia. Brasileira de Distribuição	15. Yamada S/A Com. e Ind.	29. Formosa Sup. e Mag. Ltda.
2. Carrefour Comércio e Indústria Ltda.	16. Empresa Baiana Alimentos (Ebal)	30. Nazaré Comercial de Alim. e Magazine Ltda.
3. Wal-Mart Brasil Ltda.	17. Savegnago Superm. Ltda.	31. Petrobras Distribuidora S/A
4. Sonae Distribuição Brasil S/A	18. D'avó Superm. Ltda.	32. Império da Banha Auto-Serviço Ltda.
5. Cia. Zaffari Com. e Indústria	19. Carvalho e Fernandes Ltda.	33. Superm. Superpão Ltda.
6. DMA Distribuidora S/A - EPA	20. Superm. Bahamas Ltda.	34. Comercial Unida de Cereais Ltda.
7. G Barbosa Comercial Ltda.	21. Giassi & Cia. Ltda.	35. Santa Cruz Imp. e Com. de Alimentos Ltda.
8. Irmãos Bretas, Filhos e Cia	22. Superm. Gimenes Ltda.	36. Comercial Delta Ponto Certo Ltda.
9. Coop – Cooperat. de Consumo	23. Superm. Nordeste Ltda.	37. Dias Pastorinho S/A Com. e Ind.
10. A. Angeloni & Cia. Ltda.	24. Unisuper Distribuidora S/A	38. Supermaia Superm. Ltda.
11. Irmãos Muffato e Cia. Ltda.	25. Superm. Modelo Ltda.	39. Covabra Comercial Ltda.
12. Condor Super Center Ltda.	26. Futurama Superm. Ltda.	40. Coop. dos Cafeic. e Citricult. de São Paulo
13. Sonda Sup. Exp. e Imp. Ltda.	27. Superm. Irmãos Lopes Ltda.	41. Jad Zogueib e Cia. Ltda.
14. Líder Sup. e Magazine Ltda.	28. AM PM Comestíveis Ltda.	42. Casa Avenida Comércio e Importação Ltda.





43. Imp.e Exp. de Cereais S/A	62. Irmãos Boa Ltda.	81. Sup. Araújo Imp. e Exp. Ltda.
44. Com. Oswaldo Cruz Ltda.	63. Master ATS Superm. Ltda.	82. Supermercado Rex Ltda.
45. Andorinha Superm. Ltda.	64. Algodoeira Sertaneja Ltda.	83. A. Daher & Cia. Ltda.
46. Araújo Hipermercados S/A	65. Superm. Andreazza Ltda.	84. Unicompra Superm. Ltda.
47. Santa Terezinha Distrib. de Prod. Industr. Ltda.	66. Asun Comércio de Gêneros Alimentícios Ltda.	85. Lopes Supermercado Ltda.
48. Econ Distribuição S/A	67. Maglioni Ribeiro e Cia. Ltda.	86. Superm. Yamaughi Ltda.
49. Barcelos e Cia. Ltda.	68. Superm. Rossi Monza Ltda.	87. Padaria Moderna Ltda.
50. Superm. da Família Ltda.	69. Diplomata Ind. e Comercial	88. Cooperativa de Consumo de Inúbia Paulista
51. Supermercado Beal Ltda.	70. Irmãos Gonçalves Comércio e Indústria Ltda.	89. Coop. C. E. Usiminas Ltda.
52. Bonanza Supermercado Ltda.	71. Supermercado Moreira Ltda.	90. Distribuidor de Atacado e Varejo Amigão Ltda.
53. Libraga e Cia. Ltda.	72. Superm. Baklizi Ltda.	91. J. Martins Sup. Planalto Ltda.
54. Big Trans Comércio Ltda.	73. Rede Superm.Passarelli Ltda.	92. Catricala & Cia. Ltda.
55. Superm. Guanabara S/A	74. Comercial S. Scrochio Ltda.	93. Tischler & Cia.
56. Supermercado Archer S/A	75. Supermercado Xande Ltda.	94. Organização Verdemar Ltda.
57. Com. e Distrib. Sales Ltda.	76. Comércio e Indústria Breithaupt S/A.	95. Coop. Agroindustrial Lar
58. Comercial Zaffari Ltda.	77. Superm. Gonçalves Ltda.	96. W. M. Tannous Ltda.
59. Cofesa Com. Fer. Santos S/A	78. Alexandre Batista Correa & Cia. Ltda.	97. Germans Distribuição de Cosméticos Ltda.
60. J. Melo Imp. e Exp. Ltda.	79. Luiz Tonin e Cia. Ltda.	98. Vilela Ribeiro & Filhos Ltda.
61. Coop. Consumo Cooperca	80. Atakarejo Distrib. de Alim. e Bebidas Ltda.	99. Althoff Supermercado Ltda.
		100. Cerealista Maranhão Ltda.

Fonte: Revista SuperHiper, 2005.

Método empregado no estudo e revisão teórica dos conceitos

O presente trabalho implementa técnicas de análise fatorial, análise de agrupamentos e análise discriminante em grupos supermercadistas brasileiros. Para o desenvolvimento deste trabalho, utilizaram-se os dados secundários provenientes do *ranking* dos 100 maiores grupos supermercadistas em 2004 atuantes no Brasil

da Revista SuperHiper, 2005. Foram coletados dados referentes ao faturamento bruto, número de *check-outs*, área total de vendas, número de lojas, número de funcionários, faturamento por *check-out*, faturamento por m², faturamento por funcionário, número de funcionários por *check-out*, número de funcionários por 100 m², número de *check-outs* por loja e m² por *check-out*, do ano de 2004. As variáveis coletadas encontram-se na Tabela 2.

Tabela 2. Variáveis utilizadas no estudo

Código	Variável
Fatur	Faturamento bruto
Check	Número total de check-outs
M ²	Área total de vendas
Lojas	Número de lojas
Func	Número de funcionários
Faturcheck	Faturamento / check-out





Código	Variável
Fatum ²	Faturamento / M ²
Faturfunc	Faturamento / funcionário
Funccheck	Funcionários / check-out
Func100m ²	Funcionários / 100 M ²
Checkloja	Check-outs / lojas
M ² check	M ² / check-out
M ² loja	M ² / loja

Fonte: Revista SuperHiper, 2005.

Análise fatorial

Dessa forma, foi elaborado, primeiramente, o método de análise fatorial para a determinação de fatores que compõem as variáveis apresentadas. Segundo HAIR, ANDERSON, TATHAM e BLACK (1998, p. 89), um fator representa uma combinação linear de variáveis originais. Segundo AAKER, KUMAR e DAY (1998, p. 582), a análise fatorial serve para a combinação de variáveis que criam novos fatores. No âmbito da análise fatorial, o que se pretende é a identificação de possíveis associações entre as variáveis observacionais, de modo que se defina a existência de um fator comum entre elas. Assim, pode-se dizer que a análise fatorial, ou análise de fator comum tem como objetivo a identificação de fatores ou *constructos* subjacentes às variáveis observacionais, o que, sem dúvida, contribui para facilitar sobremaneira a interpretação dos dados (RODRIGUES, 2002, p. 76).

Segundo JOHNSON e WICKERN (1992, p. 396), na análise do fator comum, as variáveis são agrupadas em função de suas correlações. Isso significa que variáveis que compõem um determinado fator devem ser altamente correlacionadas entre si e fracamente correlacionadas com as variáveis que entram na composição do outro fator.

O modelo de análise fatorial estima os fatores e as variâncias, de modo que as covariâncias ou as correlações previstas estejam o mais próximo possível dos valores observados. Para tal, os métodos de estimação ou extração mais usados são o das componentes principais e o da máxima verossimilhança. Neste trabalho será usado o método das componentes principais.

Antes de iniciar a análise fatorial, deve-se explorar individualmente cada variável em termos de *outliers* e do enviesamento da distribuição. Embora a normalidade não seja um pressuposto necessário de análise fatorial, tanto as distribuições muito enviesadas como os *outliers* podem distorcer os resultados, uma vez que alteram as estimativas das médias e dos desvios-padrão, modificando as estimativas das covariâncias e das correlações. Deve-se igualmente analisar o número de não-respostas, pois, quando são elevadas (correspondem pelo menos a 20% dados) e não são aleatórias, podem distorcer os resultados da análise fatorial.

No item análise fatorial, será apresentado um roteiro básico que, segundo PESTANA e GAGEIRO (2000, p. 389 a 427), é fundamental para avaliar-se a adequação da análise.

Análise de agrupamentos

A variável estatística de agrupamento pode ser definida como um conjunto de variáveis ou características que representam objetos a serem agrupados e é utilizada para calcular a similaridade entre os objetos. O método de análise de agrupamentos reúne um grupo de técnicas multivariadas cuja finalidade primária é agregar objetos com base nas características que eles possuem. A análise de agrupamentos classifica objetos (indivíduos, empresas, produtos ou mesmo comportamentos), de modo que, no agrupamento, cada objeto é muito semelhante aos outros em relação a algum critério de seleção predeterminado. Os agrupamentos resultantes de objetos devem então exibir





elevada homogeneidade interna dentro dos agrupamentos e elevada heterogeneidade externa entre agrupamentos. Dessa forma, se a classificação for bem sucedida, os objetos dentro dos agrupamentos estarão próximos entre si, e os objetos de diferentes agrupamentos estarão distantes entre si (HAIR, ANDERSON, TATHAM e BLACK, 1998).

A análise de agrupamentos tem sido chamada de análise de *clusters* (grupos), análise de conglomerados, análise Q, construção de tipologia, análise de classificação e taxonomia numérica. Essa variedade de nomes decorre em parte do uso de métodos de agrupamentos nas mais diversas áreas, como psicologia, biologia, sociologia, economia, engenharia e administração. Em todos os casos, o pesquisador está procurando uma estrutura natural entre as observações com base em um perfil multivariado (HAIR, ANDERSON, TATHAM e BLACK, 1998).

Assim, foi aplicado o método multivariado de análise de agrupamentos. A análise de agrupamentos visa ao agrupamento não mais de variáveis conforme ocorre na análise fatorial, mas de indivíduos, ou unidades observacionais, segundo determinados critérios de distância entre os respectivos vetores de dados. Como afirmam JOHNSON e WICHERN (1992, p. 573), o método corresponde também a uma importante técnica exploratória que busca identificar uma estrutura de agrupamentos com o intuito de avaliar a dimensionalidade dos dados, identificar *outliers* e fornecer interessantes hipóteses acerca de associações.

Neste estudo, a análise de agrupamentos tem como objetivo principal determinar em qual grupo se insere cada uma das redes supermercadistas, permitindo avaliar algumas características semelhantes de atuação, em função de aspectos organizacionais.

Análise discriminante

A análise discriminante (*Discriminant Analysis – DA*) é aplicada para analisar a relação entre uma única variável dependente qualitativa ou categórica (grupos) e um conjunto de variáveis independentes quantitativas ou métricas.

Esse método tem como objetivo determinar as variáveis que diferenciam ou discriminam os grupos, permitindo, assim, a identificação de grupos similares e a classificação de novos casos, em que a inserção é feita no grupo a que o caso tem maior probabilidade

de pertencer. Assim, é possível obter uma função discriminadora por meio de combinações lineares de variáveis independentes, a partir das quais seja possível classificar os elementos em cada uma das categorias da variável dependente. Dessa forma, produzem-se combinações lineares das variáveis independentes que melhor discriminam os grupos estabelecidos pela variável dependente e, assim, são definidas as regras de classificação dos elementos em cada grupo. A análise discriminante valida a análise de *cluster* e confirma os resultados da análise fatorial (PESTANA e GAGEIRO, 2000).

A discriminação é conseguida estabelecendo-se os pesos da variável dependente para cada variável independente, com o objetivo de maximizar a variância entre os grupos em relação à variância dentro dos grupos. No item análise discriminante, serão apresentados as hipóteses e os resultados.

Apresentação e análise dos resultados

Análise fatorial

Matriz de correlações

Para poder aplicar-se o modelo de análise fatorial, deve haver correlação entre as variáveis. Se essas correlações forem pequenas, é pouco provável que essas variáveis partilhem fatores comuns. Assim, um valor de correlação pode ser considerado aceitável se estiver acima de 0,4. A Tabela 3 apresenta o resultado da matriz de correlação das variáveis.

Tabela 3. Matriz de correlação das variáveis

	FATUR	CHECK	M ²	LOJAS	FUNC	FATURCHECK K	FATURM ²	FATURFUN C	FUNC CHECK	FUNC 100M ²	CHECKLOJ A	M ² CHECK	M ² LOJA
FATUR	1.000	.993	.995	.583	.993	.180	.017	.310	-.084	-.151	.074	.205	.119
CHECK	.993	1.000	.992	.651	.991	.118	-.020	.275	-.130	-.165	.047	.159	.091
M ²	.995	.992	1.000	.592	.993	.155	-.017	.294	-.101	-.175	.069	.222	.123
LOJAS	.583	.651	.592	1.000	.601	-.230	-.165	-.034	-.319	-.098	-.183	-.223	-.138
FUNC	.993	.991	.993	.601	1.000	.168	.000	.276	-.071	-.148	.078	.207	.128
FATUR CHECK	.180	.118	.155	-.230	.168	1.000	.788	.597	.609	.368	.500	.273	.480
FATUR M ²	.017	-.020	-.017	-.165	.000	.788	1.000	.450	.505	.681	.243	-.330	.092
FATUR FUNC	.310	.275	.294	-.034	.276	.597	.450	1.000	-.227	-.284	.375	.180	.340
FUNC CHECK	-.084	-.130	-.101	-.319	-.071	.609	.505	-.227	1.000	.752	.246	.150	.237
FUNC 100M ²	-.151	-.165	-.175	-.098	-.148	.368	.681	-.284	.752	1.000	-.026	-.479	-.154
CHECK LOJA	.074	.047	.069	-.183	.078	.500	.243	.375	.246	-.026	1.000	.334	.933
M ² CHECK	.205	.159	.222	-.223	.207	.273	-.330	.180	.150	-.479	.334	1.000	.557
M ² LOJA	.119	.091	.123	-.138	.128	.480	.092	.340	.237	-.154	.933	.557	1.00

Pela análise da Tabela 3, percebe-se que todas as variáveis apresentam correlação linear entre si. Dessa forma, prossegue-se a verificação de adequação da análise fatorial.

KMO e teste de esfericidade de Bartlett

O segundo passo é o teste KMO (Kaiser-Meyer-Olkin) e o teste de esfericidade de Bartlett. Este último pode ser usado para testar a hipótese de a matriz das correlações ser a matriz identidade com determinante igual a 1. Esse teste requer que os dados provenham de uma população normal multivariada. Caso o nível de significância do teste seja inferior a 5% para uma

significância definida de 5%, deve-se rejeitar a hipótese de a matriz das correlações entre as variáveis ser a identidade, mostrando, portanto, que existe correlação entre as variáveis. Caso tal não se verificasse, dever-se-ia reconsiderar a utilização desse modelo fatorial.

O teste KMO, que varia entre 0 e 1, testa a adequação da amostra quanto ao grau de correlação parcial entre as variáveis, que deve ser pequeno. O KMO perto de 1 indica coeficientes de correlação parciais pequenos, enquanto valores próximos de zero indicam que a análise fatorial pode não ser adequada, pois existe uma correlação fraca entre as variáveis. A análise dos valores do KMO está apresentada na Tabela 4.

Tabela 4. Teste KMO (Kaiser-Meyer-Olkin)

KMO	Análise fatorial
1 – 0,9	Muito boa
0,8 – 0,9	Boa
0,7 – 0,8	Média
0,6 – 0,7	Razoável
0,5 – 0,6	Má
< 0,5	Inaceitável

Os resultados do teste KMO e do teste de esfericidade de Bartlett estão a seguir.

Tabela 5. Teste KMO e teste de Bartlett

KMO		.640
Teste de esfericidade de Bartlett	Approx. chi-square	2681.00
	df	78
	Sig.	.000

Nesse caso, o KMO é 0,64, valor considerado razoável. O teste de Bartlett tem associado um nível de significância de 0,000, valor inferior a 0,05, o que leva à rejeição da hipótese de a matriz de correlações na população ser a identidade, mostrando, portanto, que existe correlação entre as variáveis.

Matriz anti-imagem

A matriz anti-imagem é uma medida da adequação

amostral de cada variável para uso da análise fatorial, em que pequenos valores na diagonal levam a considerar a eliminação da variável.

As medidas da adequação da amostra (MAS) são os valores da diagonal principal da matriz, e, quanto maiores, melhor será a análise fatorial. Os valores fora da diagonal representam o simétrico da matriz de correlações e deverão ser pequenos para a aplicação do modelo das componentes principais. A Tabela 6 mostra os resultados da matriz anti-imagem.

Tabela 6. Matriz anti-imagem

	FATUR	CHECK	M ²	LOJAS	FUNC	FATURCHECK	FATURM ²	FATURFUNC	FUNC CHECK	FUNC 100M ²	CHECKLOJA	M ² CHECK	M ² LOJA
FATUR	.820(a)	-.618	-.313	.423	-.139	-.158	.138	-.031	.122	-.174	-.023	-.036	.072
CHECK	-.618	.768(a)	-.243	-.606	-.260	.249	-.169	-.122	-.203	.196	-.032	.080	-.009
M ²	-.313	-.243	.891 (a)	.126	-.387	-.009	.048	-.026	.155	-.185	.030	-.216	-.016
LOJAS	.423	-.606	.126	.541(a)	.032	-.474	.406	.275	.422	-.313	.254	.279	-.231
FUNC	-.139	-.260	-.387	.032	.911 (a)	-.093	-.024	.200	-.085	.176	.007	.136	-.028
FATUR CHECK	-.158	.249	-.009	-.474	-.093	.497(a)	-.786	-.599	-.748	.465	.072	-.357	-.122
FATUR M ²	.138	-.169	.048	.406	-.024	-.786	.509(a)	.027	.510	-.645	-.038	.406	.082
FATUR FUNC	-.031	-.122	-.026	.275	.200	-.599	.027	.605(a)	.584	.030	-.113	.064	.091
FUNC CHECK	.122	-.203	.155	.422	-.085	-.748	.510	.584	.420(a)	-.756	-.176	-.203	.174
FUNC 100M ²	-.174	.196	-.185	-.313	.176	.465	-.645	.030	-.756	.505(a)	.157	.309	-.163
CHECK LOJA	-.023	-.032	.030	.254	.007	.072	-.038	-.113	-.176	.157	.551(a)	.455	-.952
M ² CHECK	-.036	.080	-.216	.279	.136	-.357	.406	.064	-.203	.309	.455	.523(a)	-.460
M ² LOJA	.072	-.009	-.016	-.231	-.028	-.122	.082	.091	.174	-.163	-.952	-.460	.572 (a)

Como pode ser observado na Tabela 6, todas as variáveis apresentam valores elevados na diagonal, o que leva à permanência de todas as variáveis na análise.

Comunalidades

A proporção de variância de cada variável explicada pelas componentes principais retidas designa-se por comunalidade. As comunalidades exibem o valor inicial

e, após, a extração do número desejado de fatores. As comunalidades iniciais são iguais a 1 e, após a extração, variam entre 0 e 1, sendo 0 quando os fatores comuns não explicam nenhuma variância da variável e 1 quando explicam toda sua variância.

A Tabela 7 apresenta os resultados das comunalidades de cada variável e mostra que todas as variáveis têm uma forte relação com os fatores retidos, por terem comunalidades elevadas de acordo com a coluna Extração.

Tabela 7. Comunalidades

	Inicial	Extração
FATUR	1.000	.981
CHECK	1.000	.992
M ²	1.000	.982
LOJAS	1.000	.644
FUNC	1.000	.984
FATURCHECK	1.000	.915
FATURM ²	1.000	.983
FATURFUNC	1.000	.958
FUNCHECK	1.000	.962
FUNC100M ²	1.000	.963
CHECKLOJA	1.000	.734
M ² CHECK	1.000	.781
M ² LOJA	1.000	.849

Método de extração: análise dos componentes principais.

Fatores retidos e variância total explicada

Pelo critério de Kaiser, escolhe-se o número de fatores a reter, em função do número de valores próprios maiores que 1. Os valores próprios são ordenados por tamanho. No método de extração de componentes principais, a soma dos valores próprios iguala o número de variáveis.

Na coluna Total da Tabela 8, os valores próprios ordenam-se por tamanho.



Tabela 8. Variância total explicada

Componentes	Valores Próprios Iniciais			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% de Variância	% Cumulativa	Total	% de Variância	% Cumulativa	Total	% de Variância	% Cumulativa
1	4.686	36.044	36.044	4.686	36.044	36.044	4.452	34.249	34.249
2	3.433	26.406	62.449	3.433	26.406	62.449	2.661	20.472	54.721
3	2.338	17.983	80.432	2.338	17.983	80.432	2.611	20.083	74.804
4	1.271	9.775	90.207	1.271	9.775	90.207	2.002	15.404	90.207
5	.819	6.298	96.505						
6	.323	2.481	98.986						
7	.045	.346	99.333						
8	.040	.311	99.643						
9	.023	.179	99.822						
10	.009	.072	99.894						
11	.007	.050	99.945						
12	.005	.036	99.980						
13	.003	.020	100.000						

Método de extração: análise dos componentes principais.

Na situação inicial, uma vez que se utilizou o método de extração dos componentes principais, a soma dos valores próprios iguala o número de variáveis, que, neste caso, são 13. Como existem quatro valores próprios maiores que 1, podem reter 4 fatores pelo critério de Kaiser.

Matriz de componentes antes e após rotação

Conforme dito anteriormente, o modelo de análise fatorial estima os fatores *loadings* e as variâncias, de modo que as covariâncias ou as correlações previstas pelo modelo estejam o mais próximo possível dos valores observados, em geral pelo método de extração das componentes principais ou pelo método de estimação da máxima verossimilhança. A matriz de componentes inicial apresenta os coeficientes ou pesos (*loadings*) que correlacionam as variáveis com os fatores antes da rotação.

Na matriz de componentes, a soma (em coluna) dos quadrados dos *loadings* das variáveis para cada fator é o valor próprio dos componentes. Caso não haja nenhum *loading* elevado em ambos os fatores, é necessário o uso de rotações.

Os métodos de rotação podem ser ortogonais ou

oblíquos. O primeiro produz fatores que não estão correlacionados entre si chamados fatores ortogonais, os quais são interpretados a partir de seus *loadings*. Na rotação oblíqua, os fatores estão correlacionados e, para a interpretação da solução, torna-se necessário considerar simultaneamente a matriz das correlações e os *loadings*. As rotações Varimax, Quartmax e Equamax são exemplos de rotações ortogonais, enquanto o Direct Oblimin e o Promax são exemplos de rotações oblíquas. O Varimax, por exemplo, minimiza o número de variáveis com elevados *loadings* num fator. Em geral, consideram-se significativos os *loadings* maiores ou iguais a 0,5 por serem pelo menos responsáveis por 25% da variância.

A matriz de componentes após rotação ortogonal tem como objetivo extremar os valores dos *loadings*, de modo que cada variável se associe a apenas a um fator. Quanto menos iterações forem necessárias, mais os dados se são adequados ao modelo.

Na análise fatorial, as variáveis com baixos *loadings* devem ser eliminadas, de forma que sejam utilizadas apenas as variáveis com elevados *loadings*.

A matriz de componentes apresentada na tabela 9 mostra os coeficientes ou pesos (*loadings*) que correlacionam as variáveis com os fatores antes da rotação.



Tabela 9. Matriz de componentes

	Componentes			
	1	2	3	4
M ²	.975	-.065	.140	.088
FATUR	.973	-.043	.163	.072
CHECK	.972	-.106	.176	.068
FUNC	.972	-.050	.163	.106
LOJAS	.607	-.431	.297	-.006
FATURCHECK	.220	.913	.122	-.136
FATURM ²	-.015	.736	.552	-.368
CHECKLOJA	.197	.703	-.446	.045
FUNCHECK	-.162	.696	.371	.561
M ² LOJA	.262	.648	-.575	.172
FUNC100M ²	-.282	.445	.795	.230
M ² CHECK	.299	.253	-.690	.389
FATURFUNC	.423	.436	-.225	-.734

Método de extração: análise dos componentes principais.
4 componentes extraídos.

A soma (em coluna) dos quadrados dos *loadings* das variáveis para cada fator é o valor próprio dos componentes, conforme apresentado na tabela anterior.

Para o fator 1: $0,975^2 + 0,973^2 + \dots + 0,423^2 = 4,686$

Para o fator 2: $(-0,065)^2 + (-0,043)^2 + \dots + 0,436^2 = 3,433$

Para o fator 3: $0,140^2 + 0,163^2 + \dots + (-0,225)^2 = 2,338$

Para o fator 4: $0,088^2 + 0,072^2 + \dots + (-0,734)^2 = 1,271$

A Tabela 10 apresenta a matriz de componentes após rotação.

Tabela 10. Matriz de componentes rotacionados

	Componentes			
	1	2	3	4
CHECK	.990	.061	-.051	.075
FUNC	.983	.115	-.009	.072
M ²	.980	.117	-.042	.078
FATUR	.979	.107	-.019	.104
LOJAS	.714	-.317	-.153	-.098



	Componentes			
	1	2	3	4
M ² LOJA	.029	.893	.057	.220
M ² CHECK	.122	.840	-.177	-.170
CHECKLOJA	-.025	.771	.131	.348
FUNC100M ²	-.105	-.293	.930	.042
FUNCHECK	-.092	.296	.926	-.099
FATURFUNC	.184	.232	-.240	.901
FATURM ²	-.034	-.115	.666	.725
FATURCHECK	.084	.426	.556	.646

Método de extração: análise dos componentes principais.

Método de rotação: varimax com normalização Kaiser.

Rotações convergidas em 5 iterações.

De acordo com a Tabela 10, o fator 1 representa as variáveis número de *check-outs*, número de funcionários, área total das lojas, faturamento e número total de lojas. O fator 2 representa as variáveis m² por loja, m² por *check-outs* e *check-outs* por loja. O fator 3 engloba as variáveis funcionários por 100 m² e funcionários por *check-outs*. Finalmente, o fator 4 engloba as variáveis faturamento por funcionário, faturamento por m² e faturamento por *check-outs*.

Análise de conglomerados

A partir dos resultados obtidos na análise fatorial, aplicou-se a técnica de análise de conglomerados. Foram considerados, portanto, quatro variáveis (fatores) e 100 grupos supermercadistas. Cinco *clusters* foram formados. A composição de cada agrupamento e a análise dos resultados está detalhada a seguir.

Cluster 1– CBD (Cia. Brasileira de Distribuição) e Carrefour Comércio e Indústria Ltda.

Apesar de o grupo CBD apresentar valores superiores aos do Carrefour em termos de faturamento bruto, número de *check-outs*, área de vendas, número de lojas e número de funcionários (fator 1), as empresas foram agrupadas no mesmo *cluster* devido às características semelhantes de eficiência (fator 2, que corresponde às variáveis faturamento por *check-out*,

faturamento por m² e faturamento por funcionário), além da semelhança em relação aos fatores 3 e 4. As empresas apresentam indicadores de faturamento por funcionário bem acima da média, em especial o Carrefour, ocupando a terceira posição.

Cluster 2 – Wal-Mart Brasil Ltda. e Sonae Distribuição Brasil S/A

Cluster 3– Todas as empresas que não se situaram nos demais grupos.

Da mesma forma que o primeiro *cluster*, as empresas Wal-Mart e Sonae apresentaram características semelhantes em relação aos fatores 2, 3 e 4, o que explica a formação deste *cluster*.

Cluster 4 Coop – Cooperativa de Consumo, A Angeloni & Cia. Ltda., Líder Superm. E Magazine Ltda., Supermercado Nordeste Ltda., Formosa Supermercado e Magazine Ltda., Nazaré Com. Alimentos e Mag. Ltda., Jad Zogheib & Cia. Ltda., Andorinha Supermercado Ltda., Algodoeira Sertaneja Ltda., Supermercado Yamaughi Ltda., Coop. Consumo de Inúbia Paulista, Organização Verdemar Ltda.

Os supermercados deste grupo apresentaram indicadores de *performance* (em especial faturamento por *check-out*), bem acima da média, o que explica a



formação do *cluster*. Os melhores resultados deste indicador foram obtidos pelos supermercados Yamaughi Ltda. e Jad Zogheib & Cia. Ltda., respectivamente. Com relação ao indicador faturamento por m², os melhores resultados são das empresas Yamaughi Ltda., Jad Zogheib & Cia. Ltda. e Nordestão, respectivamente. Já em relação ao indicador faturamento por funcionário, os melhores resultados deste grupo foram obtidos dos supermercados Coop. de Consumo de Inúbia Paulista e Coop – Cooperativa de Consumo, respectivamente.

Embora esteja na posição 41, a rede Jad Zogheib & Cia. Ltda. demonstra, há três anos, o melhor resultado em faturamento por metro quadrado (R\$ 26,2 mil) das 50 maiores. Segundo o gerente comercial da rede, esse desempenho é fruto de uma sinergia constante, da integração entre produtos com valor agregado e serviço ao cliente. Ele atribui o resultado a um longo trabalho de interação social que resulta no reconhecimento do serviço como principal diferencial da rede. A rede busca atender às necessidades dos clientes (SUPERHIPER, 2005).

Apesar de ter somente três lojas, a rede Yamaughi (SP), que ocupa a posição 86 no *ranking*, obteve a melhor *performance* de faturamento por metro quadrado e faturamento por *check-out* das cem maiores empresas. O segredo, segundo a empresa, é simples: suas três lojas estão no mesmo bairro paulistano, a tradicional Mooca. A filosofia da empresa é “crescer junto com o bairro”, o que vem fazendo há 80 anos. Assim, muitos clientes compram nas lojas Yamaughi como faziam seus avós, o que facilita um ajuste fino entre sua expectativa e o suprimento delas pelo supermercadista. Mas os diretores assinalam ainda um outro fator de sucesso, a boa parceria com os fornecedores, o que permite esse nível de atendimento (SUPERHIPER, 2005).

Cluster 5 – Empresa Baiana de Alimentos, AM PM Comestíveis Ltda. e Petrobras Distribuidora S/A

Todas as empresas agrupadas neste *cluster* apresentaram indicadores de *performance* (principalmente faturamento por *check-out*) bastante inferior à média. A Empresa Baiana de Alimentos está classificada em décimo sexto lugar no *ranking* e possui 423 lojas, valores próximos apenas aos do grupo CBD e Carrefour, que estão classificados em primeiro e segundo lugares, respectivamente, no *ranking* de faturamento. A Empresa Baiana de Alimentos apresenta um faturamento por *check-out* de R\$ 444.386, enquanto a média das 100 maiores empresas é de R\$ 1.342.353. A AM PM Comestíveis Ltda. está classificada em 28º lugar no *ranking*, possui 466 lojas e apresenta um faturamento por *check-out* de R\$ 368.479. Já a Petrobras Distribuidora S/A está classificada em 31º lugar no *ranking*, possui 692 lojas e apresenta um faturamento por *check-out* de R\$ 254.912.

De fato, algumas redes regionais estão conseguindo subir rapidamente no *ranking*, apostando também na construção de novas lojas. Mudanças orgânicas importantes dentro da atividade supermercadista vêm ocorrendo, mas, apesar de projetadas desde o início da década, só encontram espaço para efetivar-se com o cenário econômico propício que o país vive (www.abrasnet.com.br).

Análise discriminante

Este item apresenta os resultados da análise discriminante, com base no roteiro de PESTANA e GAGEIRO (2000). A Tabela 11 mostra, para cada grupo, as médias, os desvios-padrão e o número de elementos de cada variável correspondente às 100 empresas.

Tabela 11. Estatísticas dos grupos

GRUPO		Média	Desvio-padrão	Valid N	
				Unweighted	Weighted
1	FATOR 1	62.103.818	129.410.603	2	2.000
	FATOR 2	.1291840	.28133771	2	2.000
	FATOR 3	-.1479367	.49386504	2	2.000
	FATOR 4	.6769809	.39465615	2	2.000



GRUPO		Média	Desvio-padrão	Valid N	
				Unweighted	Weighted
2	FATOR 1	23.931.520	.40995029	2	2.000
	FATOR 2	12.856.660	.72293648	2	2.000
	FATOR 3	-.2986527	.33316393	2	2.000
	FATOR 4	-.0930943	.05549979	2	2.000
3	FATOR 1	-.2144276	.14568711	81	81.000
	FATOR 2	-.0551405	.82442318	81	81.000
	FATOR 3	-.1835367	.72861023	81	81.000
	FATOR 4	-.1090391	.93480396	81	81.000
4	FATOR 1	.8269327	.22858032	3	3.000
	FATOR 2	-2.411.342	.37997884	3	3.000
	FATOR 3	-.8140854	.44107840	3	3.000
	FATOR 4	-.9422565	.60768021	3	3.000
5	FATOR 1	-.1932691	.22130069	12	12.000
	FATOR 2	.7392254	121.164.286	12	12.000
	FATOR 3	15.168.259	145.346.398	12	12.000
	FATOR 4	.8742637	114.659.035	12	12.000
Total	FATOR 1	.0000000	100.000.000	100	100.000
	FATOR 2	.0000000	100.000.000	100	100.000
	FATOR 3	.0000000	100.000.000	100	100.000
	FATOR 4	.0000000	100.000.000	100	100.000

Hipóteses da análise discriminante

O software estatístico SPSS apresenta, por meio do quadro *Log Determinants*, as dispersões existentes nos grupos. No grupo 3, observa-se a maior dispersão (25,4), enquanto o grupo 2 (Wal-Mart e Sonae) apresenta a menor dispersão (10,7). O teste Box's M verifica se as diferentes dispersões observadas são ou não estatisticamente significativas, e é muito sensível a afastamentos da normalidade (PESTANA e GAJEIRO, 2000). Um vez que este teste tem associado um nível de significância de 0,06, valor superior a 0,05, pode-se concluir pela não-significância das diferenças observadas, ou seja, pela igualdade das dispersões entre os grupos.

A verificação da normalidade recorre aos comandos *Save Unstandardized Residuals* do *General Linear Model*, que cria as variáveis referentes aos resíduos de cada variável original, sobre as quais se aplica o teste K-S.

Dessa forma, é possível afirmar que todas as variáveis independentes são normais a 0,05.

Pela análise das correlações entre as variáveis, conclui-se que não há problemas de multicolinearidade entre as variáveis independentes.

Assim, verificadas as hipóteses, pode-se prosseguir com a análise discriminante.

Resultados da análise discriminante

A estatística de Wilk's Lambda oferece informação sobre as diferenças entre os grupos, para cada variável individualmente. Obtém-se pela razão da variação dentro dos grupos (variação não explicada) sobre a variação total. Varia de 0 e 1, em que os pequenos valores indicam grandes diferenças entre os grupos, enquanto os valores elevados indicam não haver diferenças entre os mesmos.



Como esse teste não considera as correlações entre as variáveis explicativas, a interpretação do nível de significância associado a cada variável não obriga a mantê-la (sig. < 0,05) ou removê-la (sig. > 0,05) do modelo. A não-rejeição da hipótese de igualdade da

média de uma variável nos grupos aumenta a probabilidade de ser classificada incorretamente em outro grupo (PESTANA e GAJEIRO, 2000).

A Tabela 12 mostra que existem diferenças significativas nas médias de cada variável nos grupos.

Tabela 12. Teste de igualdade de médias

	Wilk's Lambda	F	df1	df2	Sig.
FATOR 1	.042	538.218	4	95	.000
FATOR 2	.721	9.175	4	95	.000
FATOR 3	.671	11.633	4	95	.000
FATOR 4	.861	3.825	4	95	.006

A matriz de estrutura evidencia a contribuição de cada variável para a função discriminante, realçando com um asterisco as variáveis mais importantes. Quanto maiores forem os coeficientes, em valor absoluto, mais a função discriminante detém informação contida nessas variáveis.

Caso duas variáveis tenham correlação 1, incluir

ambas não fornece mais informação do que uma só. Como todas as variáveis têm baixa correlação entre si, consideram-se seus coeficientes individuais na matriz estrutura e todas as funções discriminantes. Pela Tabela 13, verifica-se que o fator 1 faz parte da primeira função; o fator 4, da segunda função; os fatores 2 e 3, da terceira função discriminante.

Tabela 13. Matriz de estrutura

	Function		
	1	2	3
FATOR 1	.911(*)	-.370	-.173
FATOR 4	.001	.940(*)	-.163
FATOR 2	-.016	-.183	.891(*)
FATOR 3	.010	-.554	-.663(*)

* Largest absolute correlation between each variable and any discriminant function

Assim sendo, a Tabela 14 apresenta as estatísticas e os testes que permitem avaliar a importância das funções discriminantes. Os valores próprios (*eigenvalues*) são a razão

da variação entre os grupos pela variação dentro dos mesmos. Quanto mais afastado de 1, maior será a variação entre os grupos explicada pela função discriminante.

Tabela 14. Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	24.563(a)	77.4	77.4	.983
2	6.516(a)	20.5	97.9	.776
3	.669(a)	2.1	100.0	.633

a) First 4 canonical discriminant functions were used in the analysis.

Pela análise da Tabela 14, pode-se observar que a primeira função contribui com 77,4% para o total da variância entre os grupos, sendo efetivamente a que tem o maior poder de separação. A segunda função explica 20,5% da variância intergrupar, enquanto a terceira função explica apenas 2,1%.

Como complemento dessas estatísticas, procede-se ao teste de Wilk's Lambda, conforme mostra a Tabela 15, para determinar-se o número de funções a serem retidas. Assim, testa-se também a hipótese das médias das duas funções discriminantes serem iguais nos três grupos, a qual é rejeitada (sig. = 0,000).

Tabela 14. Wilk's Lambda

Test of Function(s)	Wilk's Lambda	Chi-square	df	Sig.
1 through 3	.007	466.898	24	.000
2 through 3	.200	150.257	15	.000
3	.504	63.976	8	.000

A estatística F é utilizada para descrever os grupos mais parecidos e testar a igualdade das médias (centróides) dos grupos. Pode ser entendida como uma medida de distância entre cada par de grupos (PESTANA e GAJEIRO, 2000). A Tabela 16 permite afirmar que o par de grupos 1 e 3 tem o maior F

(532.783), indicando que esses grupos são os que mais se diferenciam. Os centróides desses grupos são significativamente diferentes (sig. = 0,000). Já os grupos 2 e 5 são os mais próximos (F = 28.160), porém seus centróides também são significativamente diferentes (sig. = 0,000).

Tabela 16. Comparações entre os grupos

Step	Average Linkage (Between Groups)						
			1	2	3	4	5
1	1	F		330.851	1.829.336	789.657	1.596.153
		Sig.		.000	.000	.000	.000
	2	F	330.851		301.334	66.838	260.386
		Sig.	.000		.000	.000	.000
	3	F	1.829.336	301.334		71.230	.106
		Sig.	.000	.000		.000	.745
	4	F	789.657	66.838	71.230		56.718
		Sig.	.000	.000	.000		.000
	5	F	1.596.153	260.386	.106	56.718	
		Sig.	.000	.000	.745	.000	
2	1	F		183.419	1.019.916	431.879	933.676
		Sig.		.000	.000	.000	.000
	2	F	183.419		169.436	35.479	167.059
		Sig.	.000		.000	.000	.000
	3	F	1.019.916	169.436		44.737	23.362
		Sig.	.000	.000		.000	.000
	4	F	431.879	35.479	44.737		54.279
		Sig.	.000	.000	.000		.000
	5	F	933.676	167.059	23.362	54.279	
		Sig.	.000	.000	.000	.000	
3	1	F		122.416	673.413	286.793	619.686
		Sig.		.000	.000	.000	.000
	2	F	122.416		112.457	30.717	110.343
		Sig.	.000		.000	.000	.000
	3	F	673.413	112.457		39.624	23.876
		Sig.	.000	.000		.000	.000
	4	F	286.793	30.717	39.624		54.219
		Sig.	.000	.000	.000		.000
	5	F	619.686	110.343	23.876	54.219	
		Sig.	.000	.000	.000	.000	



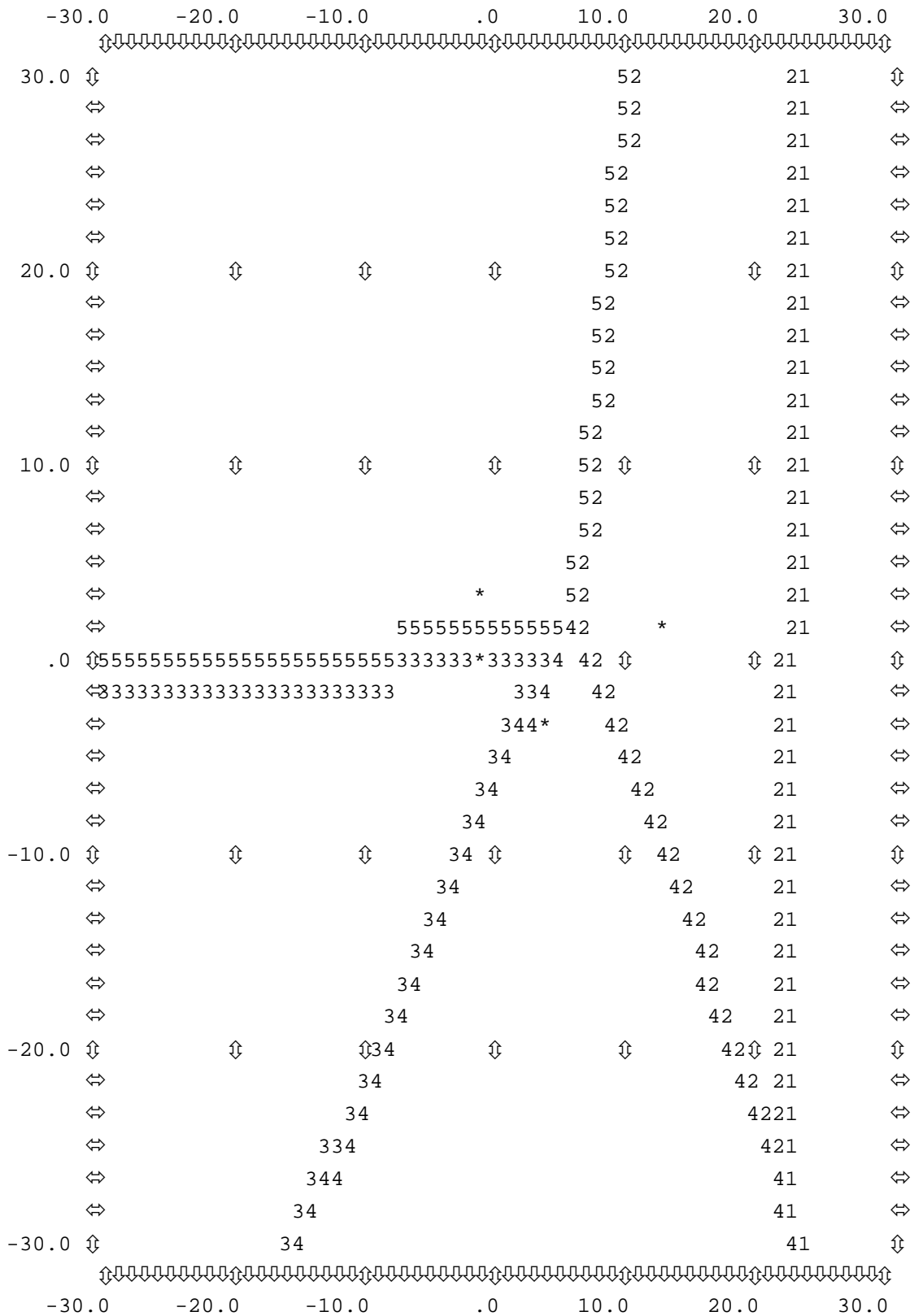
Step	Average Linkage (Between Groups)						
			1	2	3	4	5
4	1	F		96.695	532.783	237.494	476.315
		Sig.		.000	.000	.000	.000
	2	F	96.695		89.068	28.160	82.670
		Sig.	.000		.000	.000	.000
	3	F	532.783	89.068		29.901	28.473
		Sig.	.000	.000		.000	.000
	4	F	237.494	28.160	29.901		45.143
		Sig.	.000	.000	.000		.000
	5	F	476.315	82.670	28.473	45.143	
		Sig.	.000	.000	.000	.000	

O Gráfico 4 mostra o mapa territorial, que é uma versão das duas funções discriminantes, onde estão representadas linhas de separação, ou fronteiras, entre os grupos. As fronteiras com números marcam as regiões em que cada grupo é classificado. A leitura do mapa territorial mostra a existência de fronteiras bem

delimitadas para os sete grupos, e os asteriscos representam os centróides de cada grupo. Como se pode observar, os centróides estão bem afastados uns dos outros, indicando que existem diferenças significativas entre os grupos supermercadistas, já comprovadas pelos testes F.



Gráfico 4. Mapa territorial com centróides para cada grupo



Nota: O eixo y representa a função discriminante 2 e o eixo x, a função discriminante 1.



Por fim, deve-se analisar a matriz de classificação, apresentada na Tabela 17.

Tabela 17. Matriz de classificação

		GRUPO	Predicted Group Membership					Total
			1	2	3	4	5	
Original	Count	1	2	0	0	0	0	2
		2	0	2	0	0	0	2
		3	0	0	79	0	2	81
		4	0	0	0	3	0	3
		5	0	0	1	0	11	12
		%	1	100.0	.0	.0	.0	.0
	2	.0	100.0	.0	.0	.0	100.0	
	3	.0	.0	97.5	.0	2.5	100.0	
	4	.0	.0	.0	100.0	.0	100.0	
	5	.0	.0	8.3	.0	91.7	100.0	
Cross-validated(a)	Count	1	2	0	0	0	0	2
		2	0	2	0	0	0	2
		3	0	0	78	0	3	81
		4	0	0	0	3	0	3
		5	0	0	1	0	11	12
		%	1	100.0	.0	.0	.0	.0
	2	.0	100.0	.0	.0	.0	100.0	
	3	.0	.0	96.3	.0	3.7	100.0	
	4	.0	.0	.0	100.0	.0	100.0	
	5	.0	.0	8.3	.0	91.7	100.0	

a) Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b) 97,0% of original grouped cases correctly classified.

c) 96,0% of cross-validated grouped cases correctly classified.

Pela Tabela 17, é possível afirmar que o percentual de 96% de acertos sugere que a classificação foi adequada. Esse resultado é bastante superior ao critério do acaso (50%).

Conclusões e perspectivas

O presente trabalho analisou o comportamento dos 100 maiores grupos supermercadistas brasileiros,

por meio da aplicação de técnicas de análise fatorial, análise de conglomerados e análise discriminante. Os resultados da análise discriminante validaram a análise de conglomerados, confirmando a importância das variáveis fator 1 (faturamento, número de *check-outs*, área total de vendas, número de lojas e número de funcionários) e fator 4 (faturamento por *check-out*, faturamento por m² e faturamento por funcionário)



para a discriminação dos grupos.

Primeiramente, é importante ressaltar o poder explicativo das variáveis analisadas sobre o comportamento da atuação de grupos supermercadistas no Brasil. É claro que não se pode ignorar o fato de que muitas dessas variáveis representam um resultado da própria operação organizacional, já que um faturamento alto pode propiciar um maior aumento total de área de lojas, número de funcionários, expansão de novas lojas e aumento do número de *checkouts*, porém seu incremento ao longo do tempo oferece subsídios a esta análise.

Além disso, o aumento da concentração regional e organizacional ao longo dos últimos anos demonstra a reestruturação desse mercado, onde muitas empresas têm expandido as suas participações por meio de aquisições de grupos menores e, dessa forma, a sobrevivência das menores redes passa necessariamente pelo aumento de desempenho diante das gigantes do setor, tanto em gestão de custos, como em termos de vendas e diferenciação (SESSO FILHO, 2001, p. 42-43).

Por fim, percebe-se uma nova fase do setor supermercadista no Brasil, com um incremento exacerbado de competitividade, que faz com que os *players* ofertantes de produtos e serviços apresentem diferenciação no nível de serviços, quer seja pelo gerenciamento de estoques, quer seja pela automação ou pela melhora dos indicadores de atendimento e prazo de entrega. Independentemente de qual seja o principal atributo de diferenciação, é importante ressaltar que os aspectos organizacionais analisados neste trabalho representam o sucesso da organização no médio prazo e, portanto, devem ser modelados e avaliados para que sejam determinados padrões e tendências de comportamento.

Sugestões para novos estudos podem relacionar-se com a continuidade da análise, incluindo novas variáveis no modelo e a segmentação por regiões ou por estados. Ao invés da variável faturamento, poderia ser utilizada a variável lucro operacional, uma vez que já são deduzidos os custos e despesas operacionais, permitindo a análise de novos desempenhos. Porém, como a informação referente ao lucro operacional de muitas empresas não estava disponível, optou-se pela variável faturamento. Pela inclusão de novas variáveis e pela segmentação por regiões ou estados, podem ser

fornecidas respostas diferentes de comportamento e desempenho sobre a formação de *clusters* de empresas, permitindo o alcance de resultados mais profundos sobre a atuação varejista no setor de supermercados no Brasil.

Referências bibliográficas

- AAKER, D.; KUMAR, V.; DAY, G. Marketing research, 6 ed.. New York: John Wiley & Sons Inc., 1998.
- ANGELO, C. F.; SILVEIRA, J. A. G.; SIQUEIRA, J. P. L. O comportamento de consumo nos supermercados. *Pesquisa PROVAR*– Programa de Administração de Varejo da FIA, 2001.
- BLECHER, N. Quer preço ou quer marca? *Revista Exame*, ano 35, n. 12, p. 32, 2001.
- GHISI, F. A.; CAMARGO, S. V.; MARTINELLI, D. P. Redes de compras de supermercados de pequeno e médio porte: um estudo multicaso no interior do Estado de São Paulo. *Varejo Competitivo*, São Paulo, v. 7, p. 217-240, 2003.
- HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. *Multivariate data analysis*. 5. ed. New Jersey: Prentice Hall, 1998.
- JOHNSON, R.; WICHERN, D. *Applied multivariate statistical analysis*. 3 ed. New Jersey: Prentice Hall, 1992.
- MARQUES, E. F.; NORONHA, A. B. Avaliação e dimensionamento da percepção de excelência de serviços no setor supermercadista: uma abordagem quantitativa. *Varejo Competitivo*, São Paulo, v. 5, p. 173-197, 2001.
- PESTANA, M. H.; GAJEIRO, J. N. *Análise de Dados para Ciências Sociais*. 2. ed. Lisboa: Edições Sílabo, 2000.
- RODRIGUES, M. C. P. Potencial de desenvolvimento dos municípios fluminenses: uma metodologia alternativa ao IQM, com base na análise fatorial exploratória e na análise de *clusters*. *Caderno de Pesquisas em Administração*, São Paulo, v. 9, n. 1, p. 75-89, jan - mar 2002.
- ROJO, J. G. *Supermercados no Brasil: qualidade, marketing de serviços, comportamento do consumidor*. São Paulo: Atlas, 1998.
- SESSO FILHO, U. A. Crescimento e desempenho de redes de supermercado na década de 90. *Varejo Competitivo*, São Paulo, v. 6, p. 21-44, 2001.
- SUPERHIPER. Revista da Associação Brasileira de Supermercados. São Paulo, ano 29, n. 333, maio 2003.
- SUPERHIPER. Revista da Associação Brasileira de Supermercados. São Paulo, ano 31, n. 353, maio 2005.
- www.abrasnet.com.br. Acesso em 24/5/2005.

