

Integração de Conhecimento de Especialistas na Tipificação de Lojas de Retalho

Armando B. Mendes, Dep. de Matemática, Universidade dos Açores. R. da Mãe de Deus,
9501-801 PONTA DELGADA (e-mail: amendes@notes.uac.pt)

Margarida G.M.S. Cardoso, Dep. de Métodos Quantitativos, ISCTE
Av. das Forças Armadas 1649-026 LISBOA (e-mail: margarida.cardoso@iscte.pt)

Trabalho em Curso;

Área Temática: Análise classificatória;

Área de Aplicação: *Marketing*;

Palavras-chave: *marketing*; localização de lojas de retalho; áreas de influência; análise classificatória; *multidimensional scaling*; regressão

Resumo

Motivação

O sector do retalho em Portugal está a passar por uma reestruturação, tendo-se observado nos últimos anos a perda da liderança em termos de volume de vendas por parte das grandes superfícies em favor das pequenas e médias superfícies. Já em 1996 os supermercados foram os únicos a registar um crescimento simultaneamente no número de lojas e no volume de vendas e conseqüentemente a aumentar a quota de mercado de 28 para 34% no universo Nielsen.

Em 97 os supermercados atingiram a liderança e consolidaram a sua estratégia de expansão, principalmente no que se refere a cadeias como o *Pingo Doce*, *Intermarché*, *Dia* e *Lidl*. Segundo os dados mais recentes, em 2000, as vendas em supermercados eram já largamente superiores às vendas em hipermercados: 44% contra apenas 35% do total das vendas alimentares em Portugal (Distribuição Hoje, Dezembro de 2002).

Perante este cenário alguns grupos de retalho Portugueses estão a investir fortemente em lojas de pequena dimensão, que ofereçam uma gama de produtos limitada, mas surjam como uma alternativa aos grandes espaços principalmente pela comodidade da proximidade e da rapidez de serviço.

No entanto, a localização de lojas de pequenas dimensões é especialmente crítica, uma vez que os investimentos apresentam poucas economias de escala e uma localização e/ou gestão errada pode conduzir ao encerramento ou trespasse da loja (McGoldrick, 2000 e Salvaneschi, 1996)

Este trabalho situa-se no âmbito de um estudo de localização de lojas de retalho alimentar de pequena e média dimensão, pretendendo tipificar um conjunto de lojas existentes. Os referidos agrupamentos são úteis não apenas para se poder avaliar o desempenho relativo das localizações e gestão das lojas mas também e, por utilização de métodos de previsão por analogia, para avaliação de novas localizações potenciais (Lilien. e Rangaswamy, 2002, Clarke, *et al.*, 2001 e Ghosh e McLaferty, 1987).

Recolha de Informação

O problema da localização é complexo por ser necessário considerar inúmeras variáveis na avaliação de localizações de lojas (Mendes e Themido, 2003, McGoldrick, 2000, Salvaneschi, 1996, Themido *et al.*, 1998 e Bowlby *et al.*, 2002). Os retalhistas cedo se aperceberam da importância da localização, mas tentar perceber todos os aspectos do desempenho de lojas, potenciais localizações e comportamentos do consumidor obriga a enormes quantidades de informação de vários tipos como geográfica, demográfica, socioeconómica e referente a dinâmicas de competição.

Na tentativa para obter um grande número de variáveis que cobrisse todos os aspectos da avaliação de localizações reuniram-se, neste estudo, dados de:

- inquéritos realizados todos os dias de duas semana consecutivas, a todas as lojas existentes da cadeia (cerca de 20), abrangendo um total de 3.766 inquiridos. Estes inquéritos permitiram avaliar a opinião dos clientes quanto a variáveis como a configuração da loja, acessibilidades, configuração da localização, caracterização do cliente, caracterização da relação cliente-loja (motivação, meios de deslocação à loja, escolhas e preferências) e ainda identificação da concorrência;
- várias variáveis de localização referidas a uma lista de itens preenchida por observação *in loco*;
- base geográfica Nacional do INE com informação demográfica do censo de 2001. Estas variáveis são utilizadas na avaliação de áreas de influência com utilização de Sistemas de Informação Geográficos e uma base de dados com a localização de mais de 800 lojas de retalho alimentar em Portugal.

Neste estudo utilizou-se, ainda, informação relativa à área de influência das lojas. Tradicionalmente, as áreas de influência são delimitadas com *buffers* ou circunferências, com um raio adequado e calibrado utilizando resultados a inquéritos a clientes (Birkin *et al.*, 1996 e McMullin, 2000) ou delimitando coroas circulares baseadas em ‘tempos de viagem’ na deslocação à loja e em algoritmos de caminho mais curto (Salvaneschi, 1996, Cowen, et al. 2000 e Boots, 2002). A informação sobre áreas de influência provém, no caso presente, da aplicação de um método que as define a partir de diagramas de Voronoi multiplicativos ponderados. Ao contrário dos restantes, este método permite, simultaneamente, incorporar a atractividade da loja e a presença de concorrência nas proximidades (Gonçalves e Mendes, 2002).

Integração do conhecimento de especialistas por validação a posteriori

Apesar da abundância de dados, o número de lojas com informação disponível é muito reduzido - cerca de duas dezenas - o que dificulta o processo de agrupamento das mesmas lojas. Neste contexto, a utilização de conhecimento de especialistas revelou-se essencial, quer no processo de selecção das variáveis base de agrupamento, quer na validação dos resultados obtidos.

Numa primeira abordagem deste problema, o conhecimento de especialistas foi integrado na fase de apreciação de resultados de sucessivos agrupamentos hierárquicos (Cardoso e Mendes, 2002). Após muitas tentativas ao longo de vários meses obteve-se uma tipificação que se revelou consensual baseada na utilização de apenas duas variáveis de agrupamento: as vendas anuais das lojas e a percentagem de inquiridos que provinha de casa e voltava a casa após as compras na loja.

Na Figura 1 apresentam-se os agrupamentos obtidos incluindo a denominação para os grupos de lojas baseada na caracterização dos agrupamentos. Nos três anos de vendas disponíveis a variação dentro dos grupos tem apresentado alguma coerência, como pode ser observado da figura.

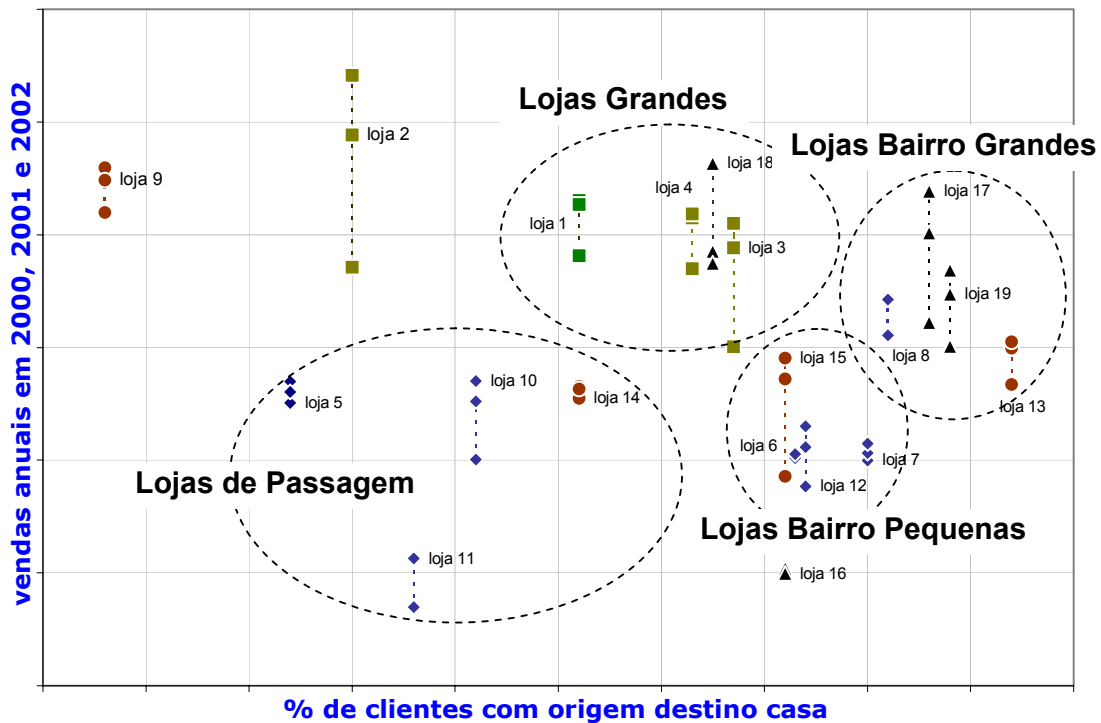


FIGURA 1 TIPIFICAÇÃO DAS LOJAS OBTIDA POR VALIDAÇÃO À POSTERIORI.

Integração do conhecimento de especialistas por validação a priori

Neste trabalho pretende-se abordar o problema da integração da opinião dos especialistas na selecção das variáveis base de agrupamento das lojas, procurando formalizar a sua contribuição.

Para tal solicitou-se a dois especialistas, profundamente conhecedores das lojas, o preenchimento de um questionário onde se comparam pares de lojas segundo uma escala ordinal de 1 a 9. A comparação é genérica tendo, no entanto, sido realçado os aspectos da localização e caracterização da loja e da gestão da mesma.

A matriz simétrica de dissemelhanças obtida da forma anterior pode ser utilizada de dois modos distintos: directamente para obter agrupamentos de lojas, ou utilizando uma análise MDS (*multidimensional scaling*) não métrica e posterior extracção de variáveis relevantes para a quantificação das dissemelhanças.

Ainda que a primeira via seja possível não foi considerada neste trabalho tendo em conta o objectivo de previsão de vendas para novas lojas, que exige a utilização de variáveis para afectar novas localizações a uma das tipificações pré-definidas.

Assim, realizou-se uma análise MDS não métrica, usando o algoritmo *ALSCAL*, um algoritmo de Takane, Young and Leeuw (Cox e Cox, 1994). Desta análise resultou uma solução com quatro dimensões, principais responsáveis pelas dissemelhanças entre as lojas, solução à qual que corresponde a um valor de *Stress* de 0,078 e *RSQ* de 96%.

Posteriormente foram realizadas regressões lineares de diversas variáveis disponíveis, sobre as dimensões extraídas. Deste modo foi possível identificar seis variáveis conducentes a regressões estatisticamente significativas (ao nível de significância 0,05).

Na Figura 2 representam-se as variáveis identificadas no espaço das dimensões MDS utilizando os coeficientes da regressão padronizados. Nesta figura é possível observar a elevada dispersão das variáveis no espaço perceptual o que é desejável já que se pretende que as variáveis identificadas representem fielmente as dimensões MDS.

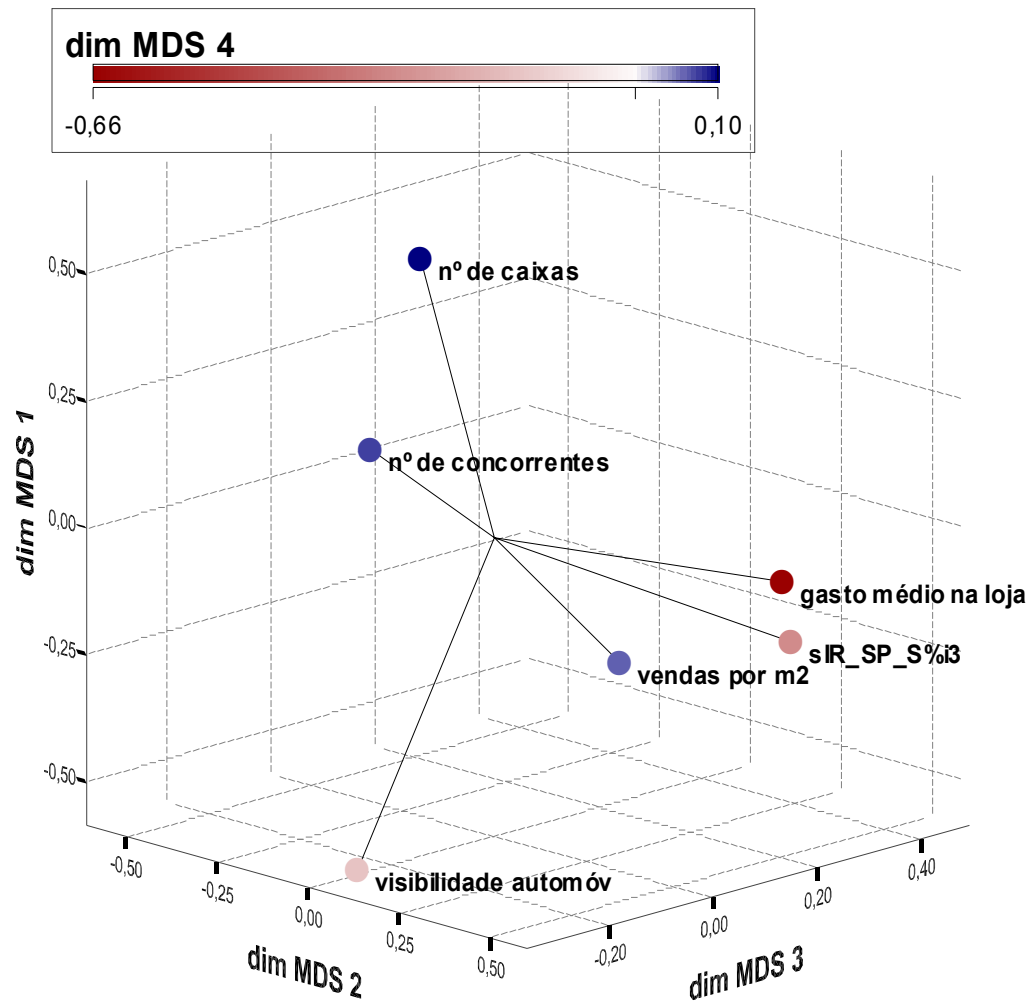


FIGURA 2 VARIÁVEIS ESTATISTICAMENTE SIGNIFICAS NA REGRESSÃO COM AS DIMENSÕES MDS.

Estas variáveis foram consideradas relevantes para a tipificação das lojas e foram posteriormente utilizadas em análises de agrupamento usando métodos hierárquicos e ainda árvores de regressão CART (Breiman, *et al.* 1984), tendo como variável alvo vendas associadas às lojas. Dessa análise resultou um novo agrupamento das lojas que se apresenta na Figura 3. Os grupos de lojas apresentados foram obtidos pelo método de Ward com distância Euclidiana quadrada.

Estes *clusters* são bem definidos não tendo sido identificada qualquer loja atípica ao contrário do que se verificou na análise *a posteriori* onde três lojas foram consideradas atípicas e excluídas.

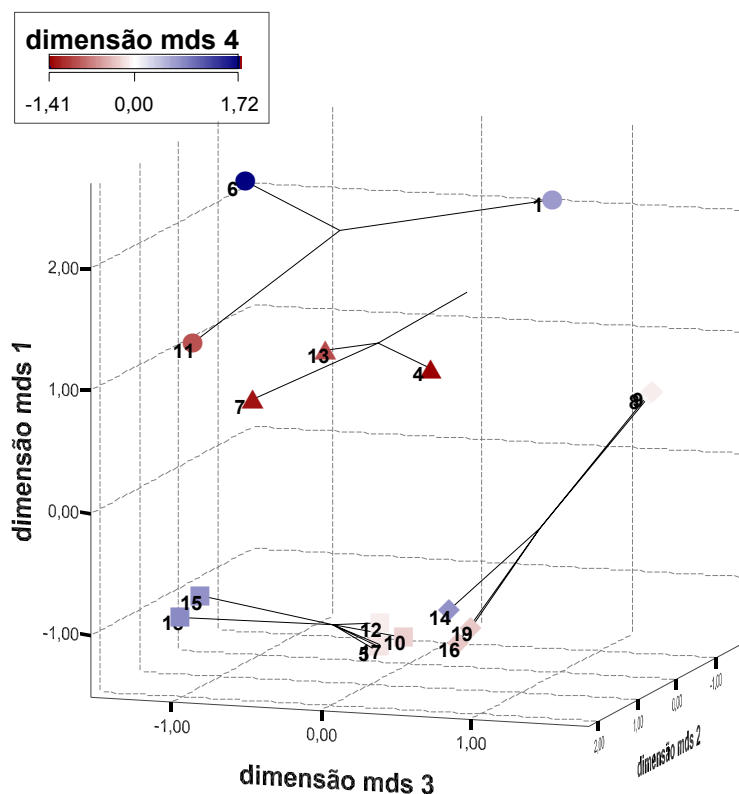


FIGURA 3 AGRUPAMENTOS DE WARD COM AS VARIÁVEIS IDENTIFICADAS POR REGRESSÃO.

Conclusões:

Para comparação das duas tipificações obtidas utiliza-se a fracção de variância de vendas anuais intra-grupos calculada excluindo as lojas atípicas. No Quadro 1 apresentam-se os valores referentes à fracção da variância dentro dos grupos sobre o total (*i.e.* considerando apenas um agrupamento) para as vendas anuais nos três anos disponíveis e para um total de 16 lojas.

Neste caso particular é fácil concluir pela superioridade do método *a posteriori* de integração do conhecimento de especialistas na tipificação de lojas de retalho de pequena dimensão, ainda que se verifique uma redução da percentagem de variância das vendas explicada pelo agrupamento à medida que mais dados de vendas vão sendo disponibilizados.

Pode-se ainda concluir pela maior facilidade em aceitar \ rejeitar os agrupamentos após a sua constituição do que a utilização de grande número de comparações binárias. Esta foi considerada difícil pelos especialistas por ser necessário manter uma visão global das restantes comparações (foi referido que por várias vezes foi necessário rever classificações já atribuídas por comparação com novas classificações).

QUADRO 1 FRACÇÃO DA VARIÂNCIA DAS VENDAS ANUAIS EXPLICADA PELOS AGRUPAMENTOS

ano	análise a posteriori	análise a priori
2000	86,07%	41,05%
2001	78,92%	34,52%
2002	62,47%	40,67%

Bibliografia

- Birkin, Mark; Clarke, Graham; Clarke, Martin e Wilson, Alan (1996) *Intelligent GIS: Location decisions and strategic planning*. GeoInformation International: Cambridge.
- Boots, Barry (2002) Using local statistics for boundary characterization. Em: Boots, Barry; Okabe, Atsuyuki e Thomas, Richard (Eds.) *Modelling Geographical Systems: Statistical and computational applications*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Bowlby, S.; Breheny, M. e Foot, D. (2002) Store location: Problems and methods 2: Expanding into new geographical areas, *Retail & Distribution Management* (Nov/Dec) pp. 41-46.
- Breiman, L., Friedman, J., Olshen, R., Stone, C. (1984). *Classification and Regression Trees*, Wadsworth, Inc., California.
- Cardoso, Margarida G.M.S. e Mendes, Armando B. (2002) Segmentação de clientes de lojas de pequena dimensão. Em: Carvalho, Lucília; Brilhante, Fátima e Rosado, Fernando (Eds.) *Novos Rumos em Estatística*. Actas do Congresso Anual da Sociedade Portuguesa de Estatística 9ª ed. Sociedade Portuguesa de Estatística, Ponta Delgada, pp. 157-170.
- Clarke, Ian; Mackness, William; Ball, Barbara e Horita, Masahide (2001) The devil is in the detail: Visualising analogical thought in retail location decision-making. Em: *Recent Advances in Retailing & Services Science*. European Institute of Retailing & Services Science Conference 8ª ed..
- Cox, Trevor F. e Cox, Michael A.A. (1994) *Multidimensional Scaling*. Chapman & Hall.
- Cowen, David J.; Jensen, John R.; Shirley, W. Lynn; Zhou, Yingming e Remington, Kevin (2000) Commercial real estate GIS site evaluation models: Interfaces to ArcView GIS. Em: *Proceedings*. Annual ESRI International User Conference 20ª ed. ESRI online Library.
- Ghosh, Avijit e McLaferty, Sara L. (1987) *Location Strategies for Retail and Service Firms*. Lexington Books: Lexington, Massachusetts.
- Gonçalves, Alexandre B. e Mendes, Armando B. (2002) Caracterização de áreas de influência de lojas de retalho alimentar de pequena dimensão com base em diagramas de Voronoi ponderados. Em: *ESIG'2002*. Encontro de Utilizadores de Informação Geográfica 7ª ed. USIG, Lisboa.
- Lilien, Gary L. e Rangaswamy, Arvind (2002) *Marketing Engineering: Computer-assisted marketing analysis and planning*. 2ª ed., Prentice Hall.
- McGoldrick, P. (2000) *Retail Marketing*. 2ª ed., McGraw-Hill: Maidenhead.
- McMullin, Shaun K. (2000) Where are your customers raster based modeling for customer prospecting. Em: *Proceedings*. Annual ESRI International User Conference 20ª ed. ESRI online Library.
- Mendes, Armando B. e Themido, Isabel Hall (2003) Multi outlet retail site location assessment: A state of the art, *International Transactions in Operations Research* Sumetido para publicação.
- Salvaneschi, Luigi (1996) *Location, Location, Location: How to select the best site for your business*. Psi Successful Business Library, Psi Research - Oasis Press.
- Themido, Isabel Hall; Quintino, António e Leitão, José (1998) Modelling the retail sales of gasoline in a Portuguese metropolitan area, *International Transactions in Operations Research* 5 (2) pp. 89-102.