

## O MARKETING E A ANÁLISE DE DADOS PARA A TOMADA DE DECISÕES

FERNANDO AUGUSTO DE SÁ NEVES DOS SANTOS\*

MARIA MANUELA CARIA FIGUEIRA DE SÁ NEVES\*\*

### Resumo

As decisões de marketing que têm por objectivo definir os melhores planos estratégicos para abordar o mercado, escolher a melhor campanha publicitária, seleccionar o segmento e o tipo de produto a oferecer, têm de resultar de uma análise técnica e profissional da informação ou dos dados disponíveis. As técnicas estatísticas multivariadas conjuntamente com as diferentes soluções de *software* existentes tornam hoje esta tarefa em algo fácil. Necessitamos apenas de optar pela técnica mais adequada ao nosso propósito, trabalhar uma amostra para obter resultados e posteriormente validar esses resultados com novas amostras.

Muitos entendidos ainda encaram com dúvidas e temem a análise numérica de dados como forma de analisar os resultados obtidos com as suas acções e opções. Receiam a confrontação entre o que desejavam e os resultados alcançados pelas empresas. Não encaram com confiança a análise de dados, por receio de não gostarem dos resultados ou por não dominarem muitas das novas técnicas de análise de dados. Assim, desvalorizam-na embora cada vez mais esta possa ser encarada como uma base fundamental para a tomada de decisões. A utilização das novas técnicas é fundamental para comprovar que foram alcançados os resultados pretendidos com as acções e estratégias delineadas pelas empresas.

O desenvolvimento de *software* e a multiplicidade de soluções que existem hoje tornam cada vez mais fácil a análise de problemas, cada vez mais complexos se considerarmos o número crescente de informação, de dados e de variáveis que envolvem cada questão organizacional e de mercado. Quando se trata de analisar

---

\* Professor Adjunto do Departamento de Gestão da Escola Superior de Tecnologia e Gestão – Instituto Politécnico da Guarda.

\*\* Professora Adjunta do Departamento de Matemática, Secção de Estatística da Escola Superior de Tecnologia e Gestão – Instituto Politécnico da Guarda.

modelos com um número cada vez mais elevado de variáveis esta tarefa tem vindo a ser facilitada. Para empresários, estudantes, etc., a informatização facilitou a análise de resultados das empresas, dos efeitos da publicidade ou acções de publicidade, nas vendas, na empatia com marcas, na imagem de instituições, etc..

Conhecer os consumidores, as suas características, quais os mais rentáveis, que padrões de comportamento assumem e que preferências apresentam, é fundamental para melhor as empresas se adaptarem a essas exigências. A tecnologia de hoje permite-nos recorrer a muitas ferramentas e técnicas para a análise de dados e facilita a interpretação de sinais de comportamento, de preferências, de hábitos, de desejos, etc., que os consumidores apresentam, não inteligíveis numa análise directa dos valores de negócio obtido.

A segmentação de consumidores, de empresas, de marcas ou produtos permitem-nos compreender melhor o mercado, quer ao nível de comportamentos, quer ao nível de organização e distribuição. Segmentar significa encontrar agrupamentos de indivíduos, objectos, etc., que partilhem, associem ou sejam entendidos como tendo algumas características comuns. Com a segmentação pretende-se encontrar diferentes grupos com características homogéneas. Definir agrupamentos tem por finalidade eliminar ou diminuir a arbitrariedade entre grupos de consumidores, produtos ou objectos para estabelecer as melhores políticas da empresa. Pretende conseguir a definição de metas, objectivos, políticas de acção ou publicidade adequadas a esses produtos, marcas, objectos ou consumidores.

Devemos identificar características próprias para cada segmento: demográficas, psicográficas ou de comportamento entre consumidores, imagem, características associadas ao produto ou serviço, ao armazenamento, à embalagem, à utilização dos produtos, ao hábito de compra. Valor, notoriedade, imagem, atributos de marca. Apesar de os resultados não constituírem uma clara premonição de comportamentos, uma segmentação simples pode diferenciar um conjunto de indivíduos, produtos ou objectos em grupos básicos que permitam construir parâmetros das estratégias a implementar para conseguir o tratamento, comportamentos mais adequados a cada grupo. Estes grupos podem depois, através de instrumentos mais sofisticados ser novamente desagregados ou segmentados para encontrar, identificar mais particularidades que permitam uma melhor e mais correcta compreensão do seu comportamento ou necessidades.

O número de instrumentos e metodologias tem aumentado em número significativo e vem responder de forma cada vez mais eficiente aos interesses dos empresários.

A análise passa então por determinar o alvo ou interesse de investigadores e empresários, o quê, quando, onde e como e depois escolher a melhor metodologia para efectuar as medidas.

Esta análise passa então primeiro pelo “*Data Mining*”, ou seja, pela exploração simples dos dados. Este processo usa uma variedade de instrumentos ou ferramentas estatísticas para trabalhar os dados e encontrar padrões e relações que permitam estabelecer projecções válidas acerca do objecto do estudo. Existem várias técnicas que os investigadores podem utilizar para atingir os seus objectivos, que vão desde a simples análise das frequências de compra a análises de crescente complexidade. A análise dos clientes de maior valor (MVC’s), análise de compras por recenticidade-frequência-valor monetário (RFM), análise de perfis (affinity), árvores de decisão, análise factorial, análise de agrupamentos, regressão, análise discriminante e redes neuronais.

A segmentação de consumidores de acordo com os seus comportamentos, as preferências, os hábitos, os gostos, os locais ou canais de distribuição que habitualmente visitam ou preferem, as condições que exigem aos produtos, às embalagens, a imagem que associam com a marca e o valor que atribuem a esta, passa muitas vezes pela simplicidade da análise (Hair *et al.*, 1998).

O senso comum e a experiência são uma ajuda, contudo é necessário possuir dados e bases mais sustentáveis para planear as acções e fundamentar as opções de políticas de produto, marca, preço, etc., para chegar aos diferentes consumidores ou mais especificamente a determinados segmentos. A mais básica consiste em colocar questões simples a que a base de dados possa responder, ex.: quem compra, idade, sexo, reside, etc.. Outra dessas técnicas consiste na metodologia do RFM (Recently-Frequency-Monetary value), ou seja uma análise dos clientes com base nos montantes de despesa, nas últimas compras ou a frequência com que visitam a loja e com base no número de compras e de acordo com um período de tempo definido em consonância para todas as categorizações. Este método não permite a validação, mas tem sido utilizado com sucesso durante muitos anos. Apesar das limitações este método permite considerar outras variáveis que ajudam a definir o segmento ou mercado alvo, combinando os resultados com variáveis como a idade, rendimento, tempo de relação com o cliente, estilo de vida, etc.. A análise da relação entre produtos adquiridos (affinity) tendo em conta o peso relativo de cada um no negócio permite encontrar estratégias que alimentam ou aproveitem a associação das escolhas dos consumidores pela afinidade que estabelecem entre determinado tipo de produtos. A complexidade que ganha o tipo de relações com o cliente e a informação necessária de tratar, tem sido e será cada vez mais facilitada, pelo número e sofisticação das ferramentas disponíveis e para as quais as técnicas a seguir descritas ajudam a encontrar respostas.

## ÁRVORES DE DECISÃO

A análise de mercados ou de consumidores obriga a considerar grande número de dados, pelo que os analistas recorrem muitas vezes a técnicas de decisão com base nas tecnologias de árvores de decisão. Esta técnica procura encontrar os melhores atributos que produzam um nível de resposta desejada, classificando e agrupando os dados. Esta aprendizagem automática facilitada pelas recentes evoluções de programas informáticos, facilitam a leitura do output e são uma opção a considerar porque não exigem o domínio de outras técnicas mais complexas. As árvores de decisão traduzem-se em regras proposicionais através de outras variáveis intermédias que permitem classificar e agrupar os indivíduos ou objectos de forma homogénea, considerando simultaneamente múltiplas variáveis.

## ANÁLISE FACTORIAL E DE COMPONENTES PRINCIPAIS

A análise factorial pretende descobrir a estrutura de número elevado de variáveis interrelacionadas de modo a permitir definir uma escala de medida de factores que controlam as variáveis iniciais (Maroco, 2003). Serve para analisar as inter-relações entre um grande número de variáveis e explicar as relações subjacentes entre elas, através de um menor número de variáveis (Reis, 2001). Permite identificar variáveis não correlacionadas, padrões ou factores inerentes a um amplo número de variáveis, representando estes uma combinação linear das variáveis originais, cuja medida de quantidade de informação explicada por cada factor é a sua variância (Aaker, Kumar e Day, 2001).

## ANÁLISE DE CLUSTERS

“A análise de clusters é um procedimento multivariado para detectar grupos homogéneos nos dados, podendo os grupos ser constituídos por variáveis ou casos.”

PESTANA E GAGEIRO, 2000, P.429

A análise de *clusters* pretende organizar um conjunto de casos em grupos homogéneos, de tal modo que os indivíduos pertencentes a um grupo são o mais

semelhante possível entre si e diferenciados dos restantes. (Reis, 2000). Esta análise procura classificar um conjunto de objectos (indivíduos, produtos, etc.) em grupos ou categorias usando os valores observados das variáveis, sem que seja necessário definir critérios que classificam os dados que integram determinado grupo (Aaker, Kumar e Day, 2001).

Podem ser utilizados métodos hierárquicos, que obrigam ao cálculo de uma matriz de semelhança/distâncias ou os não hierárquicos que se aplicam directamente sobre os dados originais e que partem de uma repartição inicial dos indivíduos por um número de grupos definidos pelo investigador. Um exemplo de método não hierárquico é o *K-means*, que consiste na transferência de um indivíduo para o cluster cujo centróide se encontra a menor distância (Hair *et al.*, 1998; Reis, 2001; Pérez, 2001).

Os métodos hierárquicos podem ser aglomerativos ou divisivos. No primeiro cada objecto ou observação parte como sendo um cluster e nos passos subsequentes os dois objectos mais próximos vão-se agregando num só cluster. O divisivo é um processo inverso, em que se parte de um só grupo que inclui todos os indivíduos e através de divisões sucessivas e sistemáticas as observações mais afastadas vão sendo retiradas e constituem-se clusters mais pequenos (Hair *et al.*, 1998; Reis, 2001).

## ANÁLISE DE CORRESPONDÊNCIAS

Constituem-se como técnicas exploratórias que pretendem encontrar e representar as relações existentes entre variáveis num espaço multidimensional, reduzindo a dimensão de variáveis com dados qualitativos para obter um pequeno número de factores. Esta análise gera um mapa de percepção no qual atributos e marcas são posicionados. Ajudam a seleccionar características a oferecer em produtos, na distribuição, etc..

A análise de correspondências múltiplas extrai relações entre categorias e define similaridades ou dissemelhanças entre elas, caso se verifique que correspondem, permitirá o seu agrupamento.

A análise de homogeneidade é uma das técnicas de análise de correspondências, que visa encontrar as relações de interdependência, se estas existirem, entre múltiplos indicadores, categorias de variáveis nominais ou outras tratadas como tal, num espaço multidimensional. As categorias de cada variável tendem a registar projecções distantes e a proximidade entre as categorias de variáveis distintas representam objectos com perfis semelhantes (Aaker, Kumar e Day, 2001; Carvalho, 2000; Pestana e Gageiro, 2000). Procede à transformação das variáveis de entrada

atribuindo-lhes uma quantificação, quer para as categorias, quer para os objectos (Carvalho, 2001).

A quantificação de uma categoria é representativa dos objectos que ela contém, ou seja, constitui a média dos *scores* dos objectos que se enquadram nessa categoria e os *scores* dos objectos são proporcionais à média das quantificações das categorias de que fazem parte. Assim, as categorias efectuem a partição dos objectos por subgrupos distanciados uns de outros, mas localizando de forma próxima os que partilhem as mesmas categorias. Desta forma através da análise de homogeneidade (HOMALS) vamos encontrar um conjunto de características e categorias de variáveis que permitem definir subgrupos homogéneos (Carvalho, 2001).

### POSICIONAMENTO MULTIDIMENSIONAL (MDS)

Com o recurso à metodologia para o posicionamento num espaço, o posicionamento multidimensional, (*Multidimensional Scaling* – MDS), posiciona os objectos ou indivíduos num espaço para podermos obter informação sobre a proximidade entre os objectos. O propósito desta técnica é transformar os julgamentos semelhantes ou preferências dos indivíduos em distâncias representadas num espaço multidimensional (Hair *et al.*, 1998).

O posicionamento multidimensional permite efectuar uma avaliação visual rápida das diferenças entre indivíduos a partir de uma matriz de distâncias (Reis, Moreira, 1993). Os modelos MDS sobre dados de preferências e escolhas, podem ser construídos com base na consideração de segmentos latentes, substituindo os dados iniciais por informação agregada que represente os segmentos que partilham as coordenadas de representação nos eixos perceptuais. As coordenadas dos segmentos são dadas pelas médias das variáveis como forma de pesar as escolhas (Cardoso, 2001). Podemos desta forma e como é do nosso interesse, determinar ou identificar se existem diferenças entre culturas, realçando algumas das características que permitem efectuar a comparação com diferentes países (Mackay, Fersenmaier, 2000).

## ANÁLISE DISCRIMINANTE

Esta técnica é utilizada para classificar indivíduos em grupos estatisticamente distintos, com base em características conhecidas (Reis, 2001). Tem como objectivo entender as diferenças entre grupos e prever a classe ou grupo a que pertence determinado objecto ou indivíduo com base em várias variáveis independentes (Hair *et al.*, 1998).

A análise discriminante envolve a derivada de uma combinação linear de duas ou mais variáveis independentes, para encontrar a separação máxima entre grupos. Permite conhecer as variáveis mais importantes para a discriminação dos grupos, classificar novos casos de acordo com a segmentação já conhecida, identificar grupos similares, validar a análise de *clusters* e confirmar os resultados da análise de componentes principais (Pestana e Gageiro, 2000). Procede-se a uma análise discriminante para verificar da percentagem de indivíduos que ficaram correctamente classificados em grupos obtidos através de outras técnicas.

## MODELOS DE EQUAÇÕES ESTRUTURAIS

A análise utilizada para testar os dados com uma determinada hipótese inicial designa-se de confirmatória (Pestana e Gageiro, 2000). Pode um investigador com base em estudos anteriores querer definir o número de dimensões e estipular a distribuição das variáveis segundo essas dimensões (Reis, 2001). Entre os modelos de análise confirmatória, salientam-se os Modelos de Equações Estruturais (SEM) que estimam as relações de dependência múltiplas e inter-relacionadas entre indicadores observáveis e variáveis não observadas designadas por latentes e verificam as relações teóricas de um modelo (Salgueiro e Reis, 2000; Byrne, 2001).

Os modelos de equações estruturais dão-nos segurança para inferir dos resultados ao contrário de outros métodos que são apenas descritivos na análise que fazem dos dados. Procede-se à avaliação da dimensionalidade de escalas de medida de determinado facto, valores, etc., através da análise factorial confirmatória (Arbuckle e Wothke, 1999).

O trabalho a desenvolver para testar os modelos de equações estruturais podem ser de 3 tipos, estritamente confirmatório, modelos alternativos e a gerar novos modelos. No primeiro caso a tarefa consiste em testar o nível de ajustamento entre o modelo e os dados da amostra. No segundo o investigador propõe modelos alternativos com base em teorias fundamentadas até encontrar aquele que melhor representa os dados da amostra. O último cenário do trabalho a desenvolver representa a situação em

que o investigador depois de rejeitar o modelo teórico devido ao seu fraco ajustamento, procede à exploração do modelo introduzindo alterações ao modelo inicial. O método mais comum consiste no terceiro tipo, de gerar novos modelos (Byrne, 2001).

## REGRESSÃO

A análise de regressão é um modelo estatístico utilizado para relacionar duas ou mais variáveis, prever o comportamento de uma variável, dependente, através de outras variáveis, independentes. O objectivo desta análise consiste em descrever, prever e controlar a variável de interesse com base nas variáveis independentes (Aacker, Kumar e Day, 2001; Reis e Moreira, 1993). A regressão permite ainda identificar variáveis independentes que melhor definem e explicam as dimensões resultantes de outras técnicas estatísticas (Pestana e Gageiro, 2000).

## REDES NEURONAIS

Tem uma abordagem diferente das restantes técnicas porque em vez de partirem de um modelo ou conceptualizar um problema matemático, apreendem cada problema apenas com a informação existente.

Como o próprio nome indica, estas técnicas constituem-se como modelos que pressupõem uma idêntica complexidade da associação de células que constituem o cérebro. O fim de uma rede neuronal é estabelecer um conjunto de redes que permitam prognosticar um grande conjunto de dados simultaneamente. Pretende-se obter a predição ou “profecia” para aplicação a outros casos ou obter novos conhecimentos. As redes “aprendem” através de exemplos e de repetições, tal como as pessoas. Quando o modelo detecta variáveis independentes responde com determinado output, prognóstico ou resultado. A rede pressente quando surgem erros e faz a aprendizagem para encontrar o padrão mais adequado entre os *inputs* e *outputs* (Hair *et al.*, 1998).

Como exemplo, o histórico da informação permite conhecer ou caracterizar um indivíduo, mesmo que para o momento não seja importante e essa informação permite à rede neuronal classificá-lo de determinada forma em função das suas escolhas, constituindo o processo de aprendizagem. Outros indivíduos com as mesmas características terão por isso as mesmas preferências e podem ser objecto de apelos a produtos e serviços que interessaram aos primeiros. Quando o resultado destas acções não está conforme a primeira predição, os novos resultados vão alimentar a aprendizagem da rede e refazer a classificação ou prognóstico.



## CONCLUSÕES

A evolução das técnicas estatísticas e sobretudo das soluções de *software* vêm facilitando cada vez mais o trabalho dos analistas. Não basta conhecer as técnicas estatísticas para que se possam tirar conclusões acertadas sobre um mercado de produtos, comportamentos de consumidores ou de produtos, mas são hoje em dia instrumentos fundamentais para decidir. A experiência e o conhecimento adequado de produtos do grupo de consumidores constituem a chave que falta para tirar as ilações mais correctas e definir as estratégias em função dos *outputs* obtidos com a análise técnica. Porém devemos apenas ressaltar que qualquer que seja a técnica de análise de dados utilizada ela tem de ser conduzida de forma profissional para garantir que os resultados representam a realidade em estudo.

A realocação dos meios para melhor tirar proveito das condições do mercado só pode ser possível se estivermos na posse de informação correcta e válida. Algo que só a análise técnica de dados nos permite conseguir ou garantir através do estudo exploratório e da validação de resultados. Para conseguir resultados para a organização devemos ajustar os investimentos, as práticas e as políticas em função de factos comprovados.

## BIBLIOGRAFIA

Aaker, David; Kumar, V.; Day, George (2001); *Pesquisa de Marketing*; Editora Atlas; São Paulo.

Arbuckle, James; Wothke, Werner (1999); *AMOS 4.0 User's Guide*; SmallWaters Corporation; Chicago.

Byrne, Barbara (2001); “*Structural Equation Modeling With AMOS: Basics Concepts, Applications, and Programming*”; Lawrence Erlbaum Associates; New Jersey.

Cardoso, Margarida (2001); “Modelos de Segmentos Latentes: Aplicações em Marketing” in *Métodos Quantitativos 2* ed. Manuel Alberto Ferreira, Rui Menezes e Margarida Cardoso; Ed. Sílabo, Lisboa.

Carvalho, Helena (2001); “Análise de Homogeneidade (HOMALS) – Quantificação Óptima e Múltipla de Dados Qualitativos” in *Métodos Quantitativos 2* ed. Manuel Alberto Ferreira, Rui Menezes, Margarida Cardoso; Ed. Sílabo, Lisboa.

Carvalho, Helena (2000); “Homogeneidade e Correspondências Múltiplas: Comparação de Dois métodos de Análise” in *Métodos Quantitativos I* ed. Elizabeth Reis, Manuel Alberto Ferreira; Ed. Sílabo, Lisboa.

Pérez, César (2001); “Técnicas Estadísticas con SPSS”; Pearson Educación, SA; Madrid.

Chatfield, C.; Collins, A. (1992); *Introduction to Multivariate Analysis*; Chapman & Hall, New York.

Hair, Joseph et al. (1998) ; *Multivariate Data Analysis*; Fifth edition, Prentice-Hall; New Jersey.

Jobson, J. D. (1992); “*Applied Multivariate Data Analysis vol. II Categorical and Multivariate Methods*”; Springer-Verlag; New York.

Jolliffe, I. T. (1986); *Principal Component Analysis*; Springer-Verlag; New York.

MacKay, Kelly J.; Fesenmaier, Daniel R. (2000); “An Exploration of Cross-Cultural Destination Image Assessment”; *Journal of Travel Research*; vol.38; Iss.4; May; pp.417-423.

Maroco, João (2003); *Análise Estatística com utilização do SPSS*; Edições Sílabo; Lisboa.

Pérez, César (2001); *Técnicas Estadísticas con SPSS*; Pearson Educación, SA; Madrid.

Pestana, Maria H.; Gageiro, João N. (2000); “*Análise de Dados para Ciências Sociais: A complementaridade do SPSS*”; 2ª edição; Edições Sílabo; Lisboa.

Reis, Elizabeth (2000); “A Análise de Clusters e as Aplicações às Ciências Empresariais: uma visão crítica da teoria dos grupos estratégicos” in *Métodos Quantitativos I* ed. Elizabeth Reis e Manuel A Ferreira; Edições Sílabo; Lisboa.

Reis, Elizabeth (2001); *Estatística Multivariada Aplicada*; 2ª edição; Edições Sílabo; Lisboa.

Reis, Elizabeth (1998); *Estatística Descritiva*; Edições Sílabo; Lisboa.

Reis, Elizabeth; et al. (1999); *Estatística Aplicada*; Edições Sílabo; Lisboa

Reis, Elizabeth; Moreira, Raúl (1993); *Pesquisa de Mercados*; Edições Sílabo; Lisboa.

Salgueiro, Maria F.; Reis, Elizabeth (2000); “Socialização dos Jovens no Trabalho: Aplicação de um Modelo de Equações Estruturais” in *Métodos Quantitativos I* ed. Elizabeth Reis e Manuel A Ferreira; Edições Sílabo; Lisboa.

Tacq, Jacques (1997); “*Multivariate Analysis Techniques in Social Science Research*”; Sage; London.