

Nazareno Philippi Lehmkuhl

**Metodologia de baixo custo para  
implantação de estratégias de CRM em  
ambiente supermercadista**

Dissertação apresentada ao  
Programa de Pós-Graduação em  
Engenharia de Produção da  
Universidade Federal de Santa Catarina  
como requisito parcial para obtenção  
do grau de Mestre em  
Engenharia de Produção

Orientador: Prof<sup>a</sup>. Elizabeth Sueli Specialski, Dr<sup>a</sup>  
Florianópolis  
2003

Nazareno Philippi Lehmkuhl

**Metodologia de baixo custo para  
implantação de estratégias de CRM em  
ambiente supermercadista**

Esta dissertação foi julgada e aprovada para a  
obtenção do grau de **Mestre em Engenharia de Produção**  
no **Programa de Pós Graduação em Engenharia de  
Produção** da Universidade Federal de Santa Catarina

Florianópolis, 19 de dezembro de 2003

---

Prof. Edson Pacheco Paladini  
Ph.D.

Coordenador do Programa

Banca Examinadora

---

Prof<sup>a</sup> Elizabeth Sueli Specialski, Dr<sup>a</sup>  
*Universidade Federal de Santa Catarina*  
**Orientadora**

---

Prof. José Leomar Todesco, Dr.  
*Universidade Federal de Santa Catarina*

---

Prof. João Bosco Motta Alves, Dr.  
*Universidade Federal de Santa Catarina*

À minha esposa Maria Amélia,  
amiga, companheira e confidente  
de todas as horas

À minha filha Sofia, o  
farol que ilumina minha vida.

## I. AGRADECIMENTOS

A **Universidade Federal de Santa Catarina**, minha segunda casa durante os cursos de graduação e pós-graduação.

Ao **Departamento de Engenharia de Produção**, mantenedor do Programa de Pós-Graduação em Engenharia de Produção e Sistemas.

A minha orientadora **Elizabeth Sueli Specialski**, por sua dedicação ao ensino e por sua paciência e perseverança na condução de seus orientandos.

A professora **Maria Marta Leite**, por sua dedicação no meu acompanhamento durante todo o período de pós-graduação e na condução do grupo de estudos em CRM, formado por alunos de pós-graduação e graduação desta Universidade.

Aos professores **Olga Cardoso, Jane Costa, Aran Morales, José L. Todesco, Maria Marta Leite, Elizabeth S. Specialski e Hugo C. Hoeschel**, por sua dedicação na condução das disciplinas ministradas durante o curso.

Aos diretores e funcionários da empresa **Autocom Informática Sistemas e Automação Ltda.**, em cujas atividades adentrei no fantástico mundo do varejo supermercadista, por seu apoio nos momentos iniciais de minha pós-graduação.

Aos diretores dos **Supermercados Imperatriz Ltda, Srs. Tiago Lohn, João Batista Lohn, Marcos Lohn, Vidal Lohn Fº, José Renato Lohn e Júlio César Lohn**, e ao Gerente do Depto. de Informática da rede, **Sr. Vilmar Olavo de Souza**, pelo apoio incondicional para a conclusão desta empreitada e pela cessão dos dados utilizados para o desenvolvimento do presente estudo.

Aos meus pais, **Lauro Leonardo Lehmkuhl e Ivone Philippi Lehmkuhl**, que me encaminharam para a vida com seus exemplos de dedicação, trabalho e honestidade, e aos meus irmãos **Luiz**,

**Lauro, Paulo, Guido, M<sup>a</sup> Aparecida, Iara e Terezinha**, pelo apoio e exemplos de vida.

A **Luciane, João Péricles, Renato, Terezinha e Newton**, pelo apoio constante para a conclusão desta etapa.

A minha esposa, **Maria Amélia Ramlow Lehmkuhl** e a minha filha **Sofia Ramlow Lehmkuhl**, por seu apoio e, principalmente, por sua paciência demonstrados durante este período.

Aos meus colegas do grupo de estudos em CRM, **João, Magali, Elaine, Osinaldo, Cristiano e Coradini**, pelas boas horas de estudos que passamos juntos.

Aos meus colegas do Depto. de Informática do Super Imperatriz, **Elzenilton, Klaus, Márcio, Valdir e Rômulo**, pelo apoio prestado.

A **Deus**, Inteligência suprema, causa primária de todas as coisas.

E a todos que, direta ou indiretamente, me auxiliaram e contribuíram para a realização deste estudo.

*“Toda empresa, qualquer que seja seu grau de descentralização, agora tem a possibilidade de consolidar as informações sobre seus clientes e obter um quadro muito mais fiel de quem são seus clientes, que produtos e serviços eles compram e como gostam de ser atendidos.”*

Patrícia B. Seybold

## II. RESUMO

As empresas varejistas, imersas em um ambiente de concorrência agressiva, necessitam de processos e ferramentas gerenciais que auxiliem na captação e, principalmente, na manutenção dos seus clientes.

A Gestão de Relacionamentos com Clientes (CRM, em sua sigla em inglês) traz uma série de ferramentas que permitem à empresa administrar e manter relacionamentos de longo prazo com seus clientes. Entretanto, os pacotes de software que implementam as ferramentas de CRM têm, muitas vezes, preços proibitivos para as empresas de pequeno e médio porte.

O presente trabalho tem por objetivo demonstrar que uma organização varejista pode, a partir de dados já existentes de suas transações correntes, descobrir informações valiosas a respeito de seus clientes, tais como seus hábitos de consumo.

Para realizar esta demonstração, foi realizado um estudo de caso apresentando formas de se conhecer o perfil dos clientes a partir dos dados gerados pelos sistemas de frente de caixa.

Este estudo, levado a efeito em uma rede de supermercados de médio porte do estado de Santa Catarina, apresentou como resultados, diferentes meios de se conhecer os padrões de compra de clientes já previamente cadastrados, com base nos dados de venda.

Palavras-chave: CRM, fidelização, supermercados, varejo, mineração de dados.

### III. ABSTRACT

Into an aggressive competition view, the retail companies needs process and management tools that helps to get and to hold their customers.

The Customers Relationship Management (CRM) has the tools and is process allow to manage and to keep long term relationship with their customers. Although that packages implement CRM tools, they have a prohibitive cost.

This work intends to prove that a retail company can, with its pre-existent transaction data, finds valid customer's informations and knowledgement like their purchase profile

A case study was performed, pointing out different ways to get knowledge from real data, provided by the Point-of-Sale's data in a middle-size supermarket company.

The case study was performed into a middle-sized supermarket company, at state of Santa Catarina, Brazil, and shows differents approachs to get knowledge of consumer patterns, based on data generated by comercial transactions beetwen consumer and the company.

Keywords: CRM, loyalty, supermarkets, retail, data mining.

## IV. SUMÁRIO

I. Agradecimentos .....	IV
II. Resumo.....	VII
III. Abstract.....	VIII
IV. Sumário .....	IX
V. Lista de Figuras e Gráficos.....	XII
VI. Lista de Quadros .....	XIII
VII. Lista de Abreviaturas / Glossário .....	XIV
1. Introdução .....	1
1.1. O tema central.....	1
1.2. Definição do problema.....	4
1.3. Objetivos.....	4
1.3.1. Objetivo Geral:.....	4
1.3.2. Objetivos Específicos:.....	4
1.4. Delimitação do trabalho.....	5
1.5. Estrutura do trabalho:.....	5
2. O varejo no Brasil.....	7
2.1. Definições.....	7
2.2. Histórico do varejo.....	9
2.3. Canais de distribuição .....	11
2.4. Classificações e formatos varejistas.....	13
2.4.1. Classificação conforme a propriedade .....	13
2.4.2. Classificação conforme o tipo .....	14
2.4.3. Varejo sem loja .....	15
2.5. O varejo alimentício no Brasil .....	17
2.6. O varejo não-alimentício.....	20
2.7. O varejo de serviços.....	22
2.8. O setor supermercadista .....	23
2.8.1. Histórico do setor .....	23
2.8.2. Definições .....	25
2.8.3. Situação atual .....	27
2.8.3.1. O Ranking da Associação Brasileira de Supermercados – ABRAS .....	29
2.8.3.2. Crescimento e estrutura das lojas .....	31
2.8.3.3. Recursos Humanos .....	33
2.8.3.4. Automação e produtividade pessoal.....	33
2.8.3.5. Meios de pagamento .....	35
2.8.4. Massificação dos clientes .....	36
3. Marketing de Relacionamento e CRM .....	38
3.1. A necessidade da retenção do cliente.....	38
3.2. O que é o Marketing de Relacionamentos .....	39
3.3. Customer Relationship Management (CRM).....	43
3.4. Pressupostos do marketing de relacionamentos e do CRM.....	46
3.4.1. Conhecimento do cliente.....	47
3.4.2. Personalização dos produtos e das ofertas .....	50
3.4.3. Acesso ao cliente.....	52

3.4.4. Atendimento.....	56
3.4.5. A retenção dos clientes.....	58
3.5. CRM e a tecnologia da informação.....	64
3.5.1. Bancos de dados de marketing (database marketing).....	65
3.5.2. Data Warehouse.....	67
3.5.3. Data Mining.....	69
3.5.4. Raciocínio Baseado em Casos.....	72
3.6. Conclusões iniciais.....	74
4. A implantação do CRM.....	76
4.1. Utilização eficiente dos dados de venda.....	77
4.2. A identificação do cliente.....	81
4.3. Um modelo genérico de implantação de CRM.....	84
4.3.1. DI_Data.....	87
4.3.2. DI_Hora.....	88
4.3.3. DI_Produtos.....	88
4.3.4. DI_Clientes.....	88
4.3.5. DI_Loja.....	88
4.3.6. F_Vendas.....	89
4.4. O passo alternativo: conhecer o cliente.....	94
5. Estudo de Caso.....	97
5.1. A empresa.....	97
5.2. Características tecnológicas.....	98
5.2.1. Frente de caixa.....	98
5.2.2. Servidores.....	99
5.2.3. Terminais de consulta.....	100
5.2.3. Equipamentos de interconexão de redes.....	101
5.2.4. Sistemas gerenciadores de bancos de dados.....	102
5.3. Dados a serem utilizados.....	102
5.3.1. Dimensão “DI_Dia”.....	102
5.3.2. Dimensão “DI_Hora”.....	103
5.3.3. Dimensão “DI_Loja”.....	103
5.3.4. Dimensão “DI_Clientes”.....	103
5.3.5. Dimensão “DI_Produtos”.....	104
5.3.6. Dados de vendas.....	105
5.4. Montagem do <i>data mart</i> .....	105
5.4.1. Dimensão DI_Dia.....	106
5.4.2. Dimensão DI_Clientes.....	107
5.4.3. Dimensão DI_Produtos.....	107
5.4.4. Tabela de Fatos F_Vendas.....	109
5.5. Análise dos dados.....	111
5.6. Análise estatística.....	112
5.7. Análise por mineração de dados ( <i>data mining</i> ).....	121
5.7.1. Classificação por árvore de decisão.....	121
5.7.2. Agrupamentos.....	131
5.7.3. Regras de associação.....	136
6. Conclusão e recomendações.....	141
6.1. Conclusão.....	141
6.2. Recomendações.....	142

7. Referências bibliográficas .....	144
8. Anexos .....	148

## V. LISTA DE FIGURAS E GRÁFICOS

Figura 2.1. Estrutura dos canais de distribuição de bens de consumo.....	12
Figura 2.2. Os cinco fluxos do varejo.....	12
Figura 2.3. Evolução do número de lojas no setor de auto-serviço .....	32
Figura 2.4. Participação dos formatos de lojas (%) .....	33
Figura 2.5. Nº de lojas automatizadas (frente de caixa) .....	34
Figura 2.6. Formas de Pagamento .....	35
Figura 4.1. Exemplo de modelo dimensional.....	78
Figura 4.2. Modelo lógico do <i>data Mart</i> básico para implantação de CRM.....	85
Gráfico 5.1. Gráfico de dispersão entre as variáveis “ <i>valor total</i> ” e “ <i>mercearia</i> ”.	
.....	116
Gráfico 5.2. Gráfico de dispersão entre as variáveis “ <i>Mercearia</i> ” e “ <i>Limpeza</i> ”	117
Gráfico 5.3. Gráfico de dispersão entre as variáveis “ <i>Laticínios</i> ” e “ <i>Frios</i> ”. ....	117
Gráfico 5.4: Gráfico de dispersão entre as variáveis “ <i>Açougue</i> ” e “ <i>Frios</i> ”. ....	119
Gráfico 5.5. Gráfico de dispersão entre as variáveis “ <i>Recência</i> ” e “ <i>Bebidas</i> ”. ....	121
Figura 5.1. Árvore de decisão para a variável “ <i>Açougue</i> ”.....	123
Figura 5.2. Rede de dependência do primeiro estudo de classificação por árvore de decisão .....	127
Figura 5.3. Árvore de decisão para a variável “ <i>Açougue</i> ” no segundo estudo. .....	128
Figura 5.4. Rede de dependências do segundo estudo de classificação por árvore de decisão. ....	130

## VI. LISTA DE QUADROS

Quadro 2.1. Varejo Alimentício – formato de lojas .....	18
Quadro 2.2. Critérios e tipos de varejo de serviços .....	23
Quadro 2.3. Síntese do setor de auto-serviço no ano de 2001 .....	27
Quadro 2.4. Comparativo entre o setor de Auto-Serviço e o segmento de Supermercados. Síntese dos resultados do ano 2001 .....	28
Quadro 2.5. Ranking simplificado – ABRAS/2001 .....	30
Quadro 2.6. Faturamento médio por funcionário (base ano 2001) .....	34
Quadro 4.1. Atributos da entidade DI_Dia .....	86
Quadro 4.2. Atributos da entidade DI_Hora.....	86
Quadro 4.3. Atributos da entidade DI_Podutos .....	86
Quadro 4.4. Atributos da entidade DI_Lojas.....	86
Quadro 4.5. Atributos da entidade DI_Clientes .....	86
Quadro 4.6. Atributos da entidade F_Vendas.....	87
Quadro 5.1. Análise dos dados das compras dos clientes selecionados .....	113
Quadro 5.2. Correlações “fortes” entre os atributos .....	115
Quadro 5.3. Correlações “fracas” entre os atributos.....	118
Quadro 5.4. Correlações “muito fracas” entre os atributos.....	120
Quadro 5.5. Análise da classificação por árvore de decisão, primeiro estudo. ....	125
Quadro 5.6. Inter-relacionamento entre as variáveis do primeiro estudo de classificação por árvore de decisão.....	127
Quadro 5.7. Análise da classificação por árvore de decisão, primeiro estudo. ....	129
Quadro 5.8. Inter-relacionamento entre as variáveis do segundo estudo de classificação por árvore de decisão.....	130
Quadro 5.9. Resultado do primeiro estudo de agrupamentos .....	132
Quadro 5.10. Variáveis em destaque no primeiro estudo de agrupamentos..	133
Quadro 5.11. Percentual de cada variável, por grupo, no primeiro estudo de agrupamentos.....	133
Quadro 5.12. Resultado do segundo estudo de agrupamentos .....	134
Quadro 5.13. Variáveis em destaque no segundo estudo de agrupamentos .	135
Quadro 5.14. Percentual de cada variável, por grupo, no segundo estudo de agrupamentos.....	135
Quadro 5.15. Valores dos quartis das variáveis, primeiro estudo de regras de associação .....	136
Quadro 5.16. Regras geradas no primeiro estudo com Regras de Associação .....	137
Quadro 5.17. Regras geradas no segundo estudo com Regras de Associação .....	139
Quadro 5.18. Regras geradas no segundo estudo com Regras de Associação e $\delta = 0,02$ .....	140

## VII. LISTA DE ABREVIATURAS / GLOSSÁRIO

ABRAS	Associação Brasileira de Supermercados – Entidade que congrega os supermercados em nível nacional.
Check-out	Balcão munido de máquinas registradoras, impressoras fiscais e/ou equipamentos de informática utilizados para o registro dos itens da compra e de seu respectivo pagamento, normalmente situados na saída das lojas.
CPF	Cadastro de Pessoa Física. Também conhecido como CIC (cadastro individual do contribuinte)
CNPJ	Cadastro Nacional de Pessoa Jurídica. Antigo CGC (Cadastro Geral do Contribuinte)
CRM	Customer Relationship Management. Sigla em inglês para “Gestão do Relacionamento com Clientes”.
DM	Data Mart. Porção de um data warehouse que congrega dados departamentais
DW	Data Warehouse, ou Armazém de dados. Conjunto de dados baseado em assuntos, integrado, não volátil e variável em relação ao tempo, de apoio às decisões gerenciais
ECF	Emissor de Cupom Fiscal. Impressora inteligente que registra todos os itens de uma compra, calculando automaticamente os impostos a serem pagos pelo estabelecimento, armazenando as informações fiscais em uma memória não volátil.
EDI	Electronic Data Interchange. Sigla em inglês para Troca Eletrônica de Dados, mecanismo pelo qual as empresas trocam informações entre si através de redes de comunicação.
IA	Inteligência Artificial
ICMS	Imposto sobre Circulação de Mercadorias. Tarifa de âmbito estadual, embutido no preço final das mercadorias comercializadas.
ISS	Imposto Sobre Serviços. Junto com o ICMS, é calculado automaticamente pelo ECF no momento do lançamento de valores referentes à serviços.
LOG	Denominação comum aos arquivos que contém informações sobre determinadas operações efetuadas por sistemas computacionais. No contexto deste estudo,
MBA	Market Basket Analysis. Sigla em inglês para Análise da Cesta de Compras, técnica que permite prever os produtos que um determinado cliente irá adquirir baseado nas compras anteriores.

OLAP	On-Line Analytical Processing. Sigla em inglês para Processamento Analítico on-line, técnica que permite, pela manipulação interativa de dados, descobrir informações e relacionamentos desconhecidos. Necessita sempre da intervenção humana.
RBC	Raciocínio Baseado em Casos. Técnica de IA que possibilita aprender através de situações que aconteceram no passado, ou seja, é o aprendizado por experiência.
RG	Registro Geral. Número impresso na Carteira de Identidade (CI).
SGBD	Sistema Gerenciador de Bancos de Dados.
SPC	Serviço de Proteção ao Crédito. Serviço de informações sobre a situação creditícia de pessoas físicas ou jurídicas. Normalmente vinculados à entidades que congregam empresas de varejo.
SERASA	Serviço de informações de crédito, semelhante ao SPC, com a diferença de estar vinculado ao sistema bancário nacional.
TEF	Transferência Eletrônica de Fundos. Mecanismo que permite o pagamento on-line de compras através de cartões de crédito e/ou débito.
TI	Tecnologia da Informação.
WEKA	Waikato Environment for Knowledge Analysis. Pacote de programas de mineração de dados desenvolvidos na universidade de Waikato, Nova Zelândia.
SWEDA	Marca comercial de Emissor de Cupom Fiscal.
URANO	Marca comercial de Emissor de Cupom Fiscal.
SCO/UNIX	Sistema operacional desenvolvido pela empresa Santa Cruz Operation.
LINUX	Sistema Operacional gratuito, baseado no UNIX e desenvolvido originalmente pelo finlandês Linus Torvalds.
WINDOWS	Nome genérico para uma família de sistemas operacionais desenvolvidos pela empresa Microsoft Co.

# 1. INTRODUÇÃO

## 1.1. O tema central

O setor varejista é hoje uma das principais forças econômicas de nosso país, com uma extensa rede capilarizada por todos os recantos do território nacional. Segundo PARENTE (2000), o setor é responsável por mais de 10% do PIB brasileiro, com um volume de vendas superior a R\$ 100 bilhões, sendo os produtos comercializados por cerca de um milhão de estabelecimentos.

Dentre todos os tipos diferentes de estabelecimento de varejo, o de maior destaque é, sem dúvidas, o chamado “varejo alimentar de auto-serviço”, ou seja, aquelas lojas em que o próprio cliente serve-se das mercadorias. Este tipo de estabelecimento representa 86,1% de todas as vendas do setor de varejo alimentar no Brasil, enquanto que o varejo alimentar tradicional, representado pelas mercearias, padarias, empórios etc, representam apenas 13,9% deste faturamento (PARENTE, 2000).

Para ser considerado um supermercado, conforme definição da Associação Brasileira de Supermercados (ABRAS), é necessário que a loja opere em regime de auto-serviço e apresente dois ou mais pontos de saída (*check-out*), onde os produtos serão registrados e onde haverá a cobrança dos mesmos. Quando o estabelecimento possui apenas um ponto de saída (mesmo sendo auto-serviço), ele é classificado como minimercado ou loja de conveniência.

Segundo o ranking elaborado pela ABRAS, as cinco maiores redes detêm 39,1% de todo o faturamento do setor, deixando o restante (60,9%) para ser disputado pelas demais redes (SUPERHIPER, 2002). Assim, as redes de porte médio enfrentam uma dupla concorrência. De um lado, as grandes redes com

seu marketing agressivo e seu poder de negociação com a indústria; de outro, as lojas de varejo tradicional, que atendem à demanda imediata dos consumidores devido, em grande parte, pela proximidade e pelo atendimento diferenciado.

Esta concorrência, aliada às incertezas e turbulências econômicas que se abateram sobre o Brasil após a crise cambial de 1998/1999, torna difícil a vida dos supermercados de porte médio. A necessidade premente de manter ou aumentar o faturamento tornou-se vital para estas empresas. Toda nova técnica administrativa ou financeira que resulte em melhores resultados pode e deve ser implementada. Técnicas como a qualidade total, gerenciamento de rede de suprimento, gerenciamento de categorias, dentre outras, já fazem parte do rol de ferramentas a disposição dos administradores. Dentre estas novas técnicas, uma se destaca: a Gestão de Relacionamento com Clientes (Customer Relationship Management – CRM).

O CRM é uma nova forma de gestão empresarial, onde o foco de atuação da empresa passa a ser o cliente. É formado por um conjunto de conceitos e técnicas que permitem aumentar o grau de satisfação do cliente durante todo o seu período de relacionamento com a empresa.

Ao contrário do que alguns fornecedores de sistemas apregoam, CRM é muito mais do que a simples implantação de um pacote de programas de computador. É toda uma mudança de mentalidade gerencial; é a implantação de serviços de atendimento personalizado; é a disponibilização de todos os dados dos clientes para todos os pontos de contato que o cliente possa ter com

a empresa; é ofertar ao cliente exatamente o que ele deseja para satisfazer suas necessidades.

No caso dos supermercados de médio porte, este tipo de mudança de paradigmas pode ser dificultada por outros fatores, tais como a capacidade de investimento. Em um setor que se caracteriza por baixas taxas líquidas de lucro (entre 1,5% e 3%), os investimentos devem ser realizados com muito critério e visando, sempre, formas de aumentar o faturamento a curto e médio prazo. Além disso, tem forte necessidade de gerar fluxo de clientes nas lojas para incrementar as vendas

Assim, a implantação de estratégias de CRM, que são caracterizadas por mudanças estruturais e de alto custo, acaba sendo prejudicada por esta dificuldade em se convencer os dirigentes de que é uma estratégia viável, mas com resultados a médio e longo prazo.

Entretanto, pode-se gerar informações a partir de dados que a própria empresa já possui e que são utilizados de forma incompleta, como por exemplo:

- Dados de venda sub aproveitados, pois são utilizados, geralmente, apenas para processos de controle de estoque e financeiro;
- Cadastro de clientes para controle de recebimento de cheques e para controle de limites no caso da utilização de cheques pré-datados, além de emissão de notas fiscais;
- Pouco ou nenhum conhecimento dos hábitos dos consumidores. Normalmente, este conhecimento é do gerente da loja, especialmente se for uma loja pequena, de comércio de vizinhança;

Como se pode observar, consegue-se manipular estas informações de maneira a se ter um conhecimento de hábitos de consumo a partir de dados que já existem dentro da empresa, sem a necessidade de se fazer uma implantação completa de estratégias de CRM, tais como os programas de fidelização com distribuição de prêmios.

## **1.2. Definição do problema**

Tendo em vista o exposto, pode-se definir o problema a ser trabalhado neste trabalho: *“Como obter informações a respeito dos hábitos de consumo dos clientes sem a necessidade de se implantar uma estratégia completa de CRM?”*

## **1.3. Objetivos**

Para a correta resolução do problema proposto, traçou-se os seguintes objetivos:

### **1.3.1. Objetivo Geral:**

Desenvolver uma metodologia de baixo custo que permita, a partir do processamento das vendas e do cruzamento dos dados com o cadastro de clientes, gerar informações para ações de marketing direcionado, ou seja, uma versão inicial de um sistema de CRM.

### **1.3.2. Objetivos Específicos:**

- Propor uma metodologia simples para implantação das ações necessárias para atingir o objetivo proposto;

- Projetar um pequeno *data mart* que armazenará as informações de venda, possibilitando o cruzamento das informações;
- Projetar os fluxos de informação necessários para atualizar e alimentar as dimensões e tabelas de fato do *data mart*;
- Propor soluções e ferramentas de baixo custo para a implantação do projeto;

#### **1.4. Delimitação do trabalho**

O estudo será realizado tendo em vista estabelecimentos de varejo de auto-serviço (supermercados) de médio porte, ou seja, aquelas redes listadas no ranking da ABRAS a partir da vigésima posição.

Para o correto desenvolvimento deste trabalho, será necessário que a empresa tenha, de alguma forma, a identificação de pelo menos uma parte da clientela no momento da compra, para que o cruzamento de informações possa ser realizado.

Finalmente, a empresa deverá ter um cadastro mínimo de clientes, com informações básicas sobre estes e que permita, de alguma forma, identificar este cliente no momento da compra.

#### **1.5. Estrutura do trabalho:**

No capítulo 1 haverá uma pequena introdução com as informações acima, além de um posicionamento genérico do projeto.

No capítulo 2 serão abordadas as características gerais do setor varejista, suas particularidades e sua estruturação no Brasil; abordando principalmente o

setor supermercadista em maior profundidade, com suas características principais.

No capítulo 3 serão tratadas as definições básicas sobre marketing de relacionamento e CRM, além de um amplo estudo das ferramentas de TI (*data warehouse* e *data mining*) disponíveis para uso do CRM.

No capítulo 4 serão abordadas as formas de resolução dos objetivos geral e específico, com a discussão da proposta de trabalho.

No capítulo 5 será apresentado um estudo de caso, com a aplicação prática das propostas do capítulo 4.

No capítulo 6 serão apresentadas as conclusões e recomendações a serem desenvolvidas em trabalhos futuros.

Por fim, no capítulo 7 temos as referências bibliográficas utilizadas para o desenvolvimento deste projeto.

## **2. O VAREJO NO BRASIL**

Não é possível discorrer sobre técnicas de gerenciamento de clientes em um ambiente de varejo de massa, sem as devidas definições e posicionamento da situação atual do mesmo no Brasil. O objetivo deste capítulo é trazer estas definições, fazendo ao mesmo tempo, uma revisão bibliográfica sobre o setor de varejo, em especial o supermercadista.

### **2.1. Definições**

Segundo FERREIRA (2003), varejo é a “atividade comercial de venda de produtos ou serviços feita diretamente ao consumidor final; comércio de mercadorias em pequenas quantidades; local em que se pratica esse comércio”.

PARENTE (2000, p. 22) diz que:

“Varejo consiste em todas as atividades que englobam o processo de venda de produtos e serviços para atender a uma necessidade pessoal do consumidor final. O varejista é qualquer instituição cuja atividade principal consiste no varejo, isto é, na venda de produtos e serviços para o consumidor final”.

Segundo LEVY e WEITZ (2000, p.27 apud DELUCA, 2003):

“O varejo é um conjunto de atividades de negócio que adiciona valor a produtos e serviços vendidos a consumidores para uso pessoal e familiar. As pessoas freqüentemente consideram o varejo somente como a venda de produtos em lojas. Entretanto, o varejo evoluiu também à venda de serviços: estadia de uma noite em um hotel de estrada, um

exame médico, um corte de cabelo, o aluguel de uma fita de vídeo ou uma pizza entregue em casa”.

Como se pode verificar nestas primeiras definições, o varejo é caracterizado tanto pela venda de produtos quanto de serviços. Não é comum considerar um prestador de serviços, como um barbeiro ou um médico, como varejista. No entanto, a partir do momento em que o estabelecimento ou o profissional atende diretamente ao consumidor final, é considerado varejista.

Conforme RICHERT (1954 *apud* DELUCA 2003, p.36), o varejo é o processo que envolve a compra em quantidades elevadas de um atacadista e outros fornecedores e as vendas em quantidades menores ou mesmo unitárias, ao consumidor final.

Já SPOHN (1977 *apud* DELUCA 2003, p.36) coloca que “varejo é uma atividade comercial responsável por providenciar mercadorias e serviços desejados pelos consumidores”.

STANTON (1984 *apud* DELUCA 2003, p.36) mostra que todo aquele que atende diretamente ao consumidor final pode ser considerado um varejista:

“Quando uma empresa vende ao consumidor final, ela está praticando uma atividade varejista. Muitas empresas estão ligadas ao varejo dos mais diversos tipos, como lojas, boutiques, padarias, restaurantes, supermercados, farmácias, drogarias, postos de gasolina etc”.

KOTLER (1998, p.493) complementa afirmando que: “Um **varejo** ou **loja de varejo** é qualquer empresa cujo volume de vendas decorre, principalmente, do fornecimento por unidades ou pequenos lotes” (grifos do autor). O autor afirma

que, mesmo que a empresa produza o bem ou serviço comercializado, se ela o vende para o consumidor final, é considerada uma empresa varejista.

PARENTE (2000, p. 22) contrapõe-se a esta idéia, dizendo que:

“Quando fabricantes e atacadistas vendem diretamente para o consumidor final, estão também desempenhando atividades de varejo, porém não são consideradas como varejo, pois essa não é a sua principal fonte de receita”.

Assim, a empresa de varejo pode ser caracterizada como aquela que vende produtos ou serviços para o consumidor final e que faz desta venda a sua principal fonte de receita.

## **2.2. Histórico do varejo**

As atividades de atacado e varejo já eram observadas na mais longínqua antiguidade (BADIN, 1997; KUTWAK, 2000). Os primeiros registros de transações comerciais surgiram na Mesopotâmia, juntamente com os primeiros registros da escrita. No antigo Egito, os escribas egípcios tinham a função de registrar e controlar todas as movimentações de cereais dos templos (TIME-LIFE, 1991, p.10-37). O código de Hamurabi, uma das primeiras compilações de leis que se tem notícias, já dedicava várias passagens à regulamentação do comércio (KUTWAK, 2000). Na antiga Roma, as lojas tornaram-se numerosas, formando verdadeiros centros comerciais (BADIN, 1997). Com a queda do império romano, verifica-se uma redução acentuada na estrutura varejista. Este cenário somente se altera a partir do século XIV, com o surgimento do movimento mercantilista, que deu origem às chamadas guildas, ou seja, a associação de empreendedores para o estabelecimento de seus negócios. O

varejo nas Américas sofreu forte influência dos colonizadores europeus. Tanto nos Estados Unidos como no Brasil, as relações comerciais eram determinadas pela metrópole (BADIN, 1997; KUTWAK, 2000).

Já no século XIX surgiram, nos Estados Unidos, as chamadas *General Stores*, ou armazéns gerais. Eram lojas que vendiam uma variada gama de produtos. Conforme BADIN (1997):

“Estas lojas comercializavam praticamente de tudo: produtos novos, produtos alimentícios, tecidos, pólvora, armas, dentre outros, tendo estas, importante papel social por atuarem como centro de reuniões do setor rural. Muitos compradores dirigiam-se aos estabelecimentos apenas para conversar e se informar. Todavia, estas lojas são consideradas como âncoras para atividades sociais da época”.

Ainda no mercado norte-americano, com o desenvolvimento das ferrovias, surgiram as primeiras empresas que realizavam vendas através de catálogos. Esta atividade iniciou-se em 1872 com a *Montgomery Ward*, seguida, em 1886, pela *Sears* (KUTWAK, 2000).

No início do século XX apareceram as primeiras concentrações comerciais nas cidades de maior porte, normalmente seguindo os trilhos dos chamados *street cars*, os populares bondes elétricos. Por esta época surgiram, também, as primeiras lojas de departamento, de vendas pelo correio e as primeiras cadeias de lojas (BADIN, 1997).

No ano de 1929 surgiu, ainda nos Estados Unidos, o conceito de auto-serviço, uma evolução das *general stores* que dispensava a intervenção de

vendedores. Este conceito chegou ao Brasil na década de 1950, quando o primeiro supermercado nacional foi inaugurado em São Paulo.

### **2.3. Canais de distribuição**

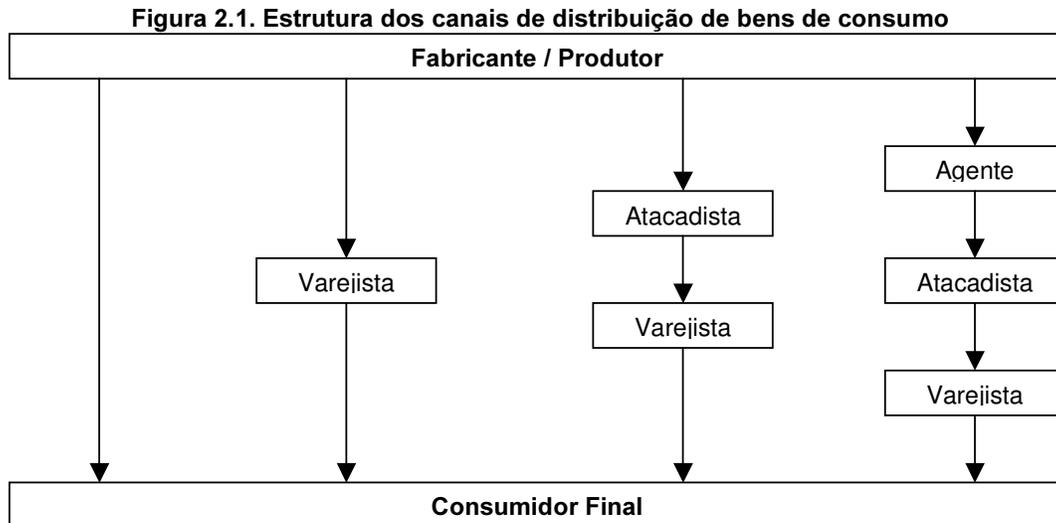
Tipicamente, entre o fornecedor de um produto ou serviço e o seu consumidor final, podem existir vários canais de distribuição. Em alguns casos, existe a venda direta do fabricante ou produtor ao consumidor final. Entre estes dois personagens, entretanto, podem existir um ou mais intermediários, cada qual agregando valor ao produto (PECK et al, 1999, P.36; PARENTE, 2000, p.22; RICHERS, 2000, p.252).

Segundo PARENTE (2000, p.22), “O varejista faz parte dos sistemas de distribuição entre o produtor e o consumidor, desempenhando um papel de intermediário, funcionando como um elo de ligação entre o nível de consumo e o nível do atacado ou da produção”.

Entretanto, qualquer que seja o número de intermediários entre o produtor e o consumidor final, cabe sempre ao varejista a posição de contato direto com este último, como podemos verificar pela figura 2.1.

Conforme o esquema apresentado, apesar de o varejista ocupar um lugar de destaque na cadeia de distribuição, ele é fortemente afetado pelo comportamento dos outros participantes da mesma cadeia. Assim, existe a necessidade premente da criação de parcerias e alianças, além de um trabalho de equipe de todos os personagens envolvidos, principalmente com as técnicas de EDI (*eletronic data interchange* – troca eletrônica de dados), Reposição Contínua e Gerenciamento de Categorias.

PECK et al (1999) mostra esta necessidade através da teoria dos seis mercados, onde o varejista tem a necessidade da criação de relacionamentos de parceria com seus clientes, enquanto cria alianças com seus fornecedores, tudo dentro de um fluxo contínuo de informações entre os operadores do canal de distribuição.



Fonte: Adaptado de PECK et Al (1999, p.34) e PARENTE (2000, p.22)

Este fluxo de informações faz parte de uma cadeia maior, composto de cinco fluxos, conforme ilustrado na figura 2.2:

**Figura 2.2. Os cinco fluxos do varejo**

Fluxo Físico	Fluxo de Posse	Fluxo de Informação	Fluxo de Pagamento	Fluxo de Promoção
FORNECEDORES (fabricantes ou atacadistas)				
↓	↓	↓↑	↑	↓
VAREJISTAS				
↓	↓	↓↑	↑	↓
CONSUMIDOR				

Fonte: Adaptado de PARENTE (2000, p.23)

O **Fluxo Físico** indica a movimentação física do produto ou serviço, sempre partindo do produtor ou atacadista (níveis acima do varejista na estrutura de canais) até chegar ao consumidor final.

O **Fluxo de Posse** acompanha o fluxo físico e indica como a posse do produto muda, conforme o mesmo caminha na estrutura.

O **Fluxo de Informações** deve ser, sempre que possível, bidirecional, pois as informações coletadas pelo varejista são de extrema importância para o produtor ou atacadista, da mesma forma que estes, tendo uma visão macro do cenário de vendas, pode prover o varejista de dados relevantes para o incremento dos negócios com os seus clientes.

O **Fluxo de Pagamento** é o inverso do fluxo de posse, pois a mesma muda mediante o pagamento de valores em troca do produto. Assim, o consumidor paga ao varejista, que paga ao fornecedor.

O **Fluxo de Promoção** é o fluxo de comunicação persuasiva na forma de propaganda, venda pessoal e promoções, partindo dos fornecedores para os varejistas e consumidores, e do varejista para os consumidores.

## **2.4. Classificações e formatos varejistas**

As instituições varejistas podem ser classificadas de várias formas, com os autores apresentando diferentes metodologias.

### **2.4.1. Classificação conforme a propriedade**

Segundo STANTON (1984, apud DELUCA, 2003, p. 36), os varejistas podem ser classificados conforme o número de pontos de venda: independentes (uma única loja), rede (várias filiais) e associações varejistas

(união de varejistas independentes para uma negociação mais efetiva com os fornecedores).

Além destes, PARENTE (2000, pp.25-30) ainda acrescenta dois outros tipos: as franquias (utilização de uma marca registrada e de sua infraestrutura através do pagamento de *royalties*) e os chamados “departamentos alugados” (departamentos de uma loja que são operados e gerenciados por uma outra empresa).

#### **2.4.2. Classificação conforme o tipo**

Segundo KOTLER (1998, p.494-495), as instituições varejistas podem ser classificadas da seguinte forma:

- **Lojas de Conveniências:** Normalmente instaladas junto a postos de gasolina, trabalham com um *mix* limitado e de alta rotatividade, em horário estendido, praticando preços acima da concorrência;
- **Lojas de Departamentos:** Lojas com *mix* normalmente não-alimentício (brinquedos, roupas, eletrodomésticos etc), com estrutura departamentalizada e cada departamento responsável por suas próprias metas de venda;
- **Lojas de Especialidade:** São lojas especializadas em uma linha de produtos. São as livrarias e lojas de artigos esportivos, dentre outras;
- **Supermercados:** Lojas caracterizadas pelo auto-serviço, com *mix* principalmente orientado para produtos alimentícios, higiene e limpeza. Operam com grandes volumes de vendas com reduzida margem de lucro;

- **Lojas de Descontos:** Trabalham com grandes volumes de venda e margem reduzida, vendendo mercadorias padronizadas;
- **Varejos de Liquidação:** Compram produtos abaixo do preço de atacados e repassam estes descontos para os consumidores. Normalmente os produtos oferecidos são do tipo ponta-de-estoque;
- **Superlojas:** Lojas com área acima dos 11.000 m<sup>2</sup> e que atendem a todas as necessidades dos clientes, como um supermercado, mas oferecendo outros serviços, como farmácias e bancas de revistas.

Além dos tipos descritos por Kotler, ainda existem:

- **Hipermercados:** São lojas que, além de oferecer produtos alimentícios, oferecem grande sortimento de produtos não alimentícios, tais como eletrodomésticos, roupas, artigos de cama, mesa e banho, dentre outros;
- **Shopping-Centers:** Estruturas que comportam várias lojas de vários ramos distintos. Normalmente possuem as chamadas lojas-âncora, que são aquelas que geram fluxo de clientes, auxiliando as lojas menores. Oferecem ainda outros benefícios aos clientes, tais como estacionamento, segurança, ambiente agradável e climatizado, as chamadas “praças de alimentação”, cinemas e áreas de lazer, entre outros.

### 2.4.3. Varejo sem loja

Além das atividades varejistas que são baseadas em um local físico ou loja, existem casos em que o comércio é realizado sem este suporte físico.

PARENTE (2000, p.35) diz, a esse respeito:

“A maior parte das vendas varejistas ainda é realizada por lojas. Entretanto, pela enorme expansão das compras virtuais, o varejo sem loja vem crescendo mais rapidamente que o varejo com loja. O varejo sem loja utiliza-se de uma estratégia de marketing que não está baseada na loja tradicional”.

O autor supra citado relaciona as seguintes formas de varejo sem loja:

- **Marketing Direto:** é mais conhecido por sua forma de divulgação, que é realizada através de catálogos. O consumidor faz o pedido diretamente ao fornecedor (via correio, telefone ou internet, por exemplo) ou através de um representante ou vendedor;
- **Venda Direta:** é um sistema direto de venda ao consumidor que envolve explicações pessoais e demonstrações dos produtos, na casa ou escritório do cliente;
- **Máquinas automáticas de venda (*Vending Machines*):** Sistema em que as vendas são realizadas através de máquinas automáticas, sem a intervenção de qualquer operador que não o cliente. A compra é realizada normalmente através de fichas, moedas, cartões magnéticos etc. A única exigência por parte do fornecedor é garantir o abastecimento regular de produtos nessas máquinas;
- **Varejo Virtual:** Uma das mais novas modalidades de varejo utiliza, a infra-estrutura da rede mundial de computadores (*internet*) para oferecer produtos aos usuários da rede através de *home-pages*, portais, lojas virtuais etc. Requer uma estrutura de logística bem estruturada, pois possibilita que qualquer usuário de qualquer ponto

do planeta adquira os produtos oferecidos. Como meios de pagamento, utiliza-se normalmente o cartão de crédito ou o pagamento através de boleto bancário gerado na hora e impresso pelo próprio usuário. Segundo CARVALHO (2003, p.82) esta modalidade de varejo está em franca expansão no Brasil, onde a maior receita deste ramo provém da venda de livros e CDs. Como exemplo, o autor mostra que, no ano de 2002, o setor de livros teve um crescimento de vendas da ordem de 8% nas lojas físicas, contra um incremento de 20% nas vendas *on-line*.

Além destas modalidades, DELUCA (2003, p.39), citando STANTON (1984) ainda destaca a venda por reembolso postal (como um exemplo do marketing direto), a venda de porta em porta (exemplo de venda direta) e a venda por telefone (também um exemplo de marketing direto).

## **2.5. O varejo alimentício no Brasil**

Dentre todas as modalidades de varejo estudadas, a mais desenvolvida em nosso país é com certeza o varejo alimentício. Nesta categoria estão concentrados os principais formatos de lojas de atendimento ao público.

De acordo com um estudo coordenado pela ABRAS (Associação Brasileira de Supermercados), o formato das lojas empregadas neste tipo de varejo varia de 20 a mais de 10.000 m<sup>2</sup> de área de venda, com ou sem a presença dos chamados *check-outs*, ou seja, balcões munidos de máquinas registradoras, impressoras fiscais e/ou equipamentos de informática utilizados para o registro dos itens da compra e de seu respectivo pagamento. O quadro 2.1 trás um resumo do referido estudo:

Quadro 2.1. Varejo Alimentício – formato de lojas

Formatos de loja	Área de Vendas/m <sup>2</sup>	Nº Médio de Itens	Nº de <i>check outs</i>	Seções
Bares	20-50	300	*	Mercearia, lanches e bebidas
Mercearias	20-50	500	*	Mercearia, frios, laticínios e bazar
Padarias	50-100	1.000	*	Padaria, Mercearia, frios, laticínios e lanches
Minimercado	50-100	1.000	1	Mercearia, frios, laticínios e bazar
Loja de conveniências	50-250	1.000	1-2	Mercearia, frios, laticínios, lanches e bazar
Supermercado compacto	300-700	4.000	2-6	Mercearia, hortifruiti, carnes, aves, frios, laticínios, bazar
Supermercado convencional	700-2.500	9.000	7-20	Mercearia, hortifruiti, carnes, aves, frios, laticínios, bazar, peixaria
Superloja	3.000-5.000	14.000	25-36	Mercearia, hortifruiti, carnes, aves, frios, laticínios, bazar, peixaria, padaria, têxtil, eletrônicos
Hipermercado	7.000-16.000	45.000	55-90	Mercearia, hortifruiti, carnes, aves, frios, laticínios, bazar, peixaria, padaria, têxtil, eletrônicos
Clube atacadista	5.000-12.000	5.000	25-35	Mercearia, hortifruiti, carnes, aves, frios, laticínios, bazar, têxtil, eletrônicos

Fonte: Adaptado de estudo coordenado pela ABRAS. BRITO (1998 apud PARENTE 2000, p30)

Os assim denominados **Bares** são especializados na distribuição de bebidas alcoólicas e não alcoólicas, tanto para consumo local quanto fora dele. Geralmente fornecem também um sortimento limitado de lanches e uma linha restrita de mercearia, concentrada em produtos de consumo imediato, tais como cigarros, balas, chocolates etc. Segundo o censo da consultoria ACNielsen (ACNIELSEN, 2000), existiam cerca de 381 mil estabelecimentos deste formato no Brasil no ano 2000.

As **mercearias**, também denominadas de armazéns, são lojas pequenas que oferecem uma linha básica de produtos das linhas de mercearia, frios, laticínios, higiene e limpeza. Normalmente estão localizadas em regiões de baixo potencial de consumo, onde não é justificada a existência de

estabelecimentos maiores. Juntamente com as padarias, somam cerca de 300 mil estabelecimentos, detendo cerca de 14% da distribuição de produtos alimentares industrializados (PARENTE, 2000, p. 31).

As **padarias**, originalmente responsáveis pela fabricação de pães e bolos, estão em processo acelerado de modernização, com o objetivo de atender um mercado cada vez mais exigente. Melhor atendimento, *mix* variado de produtos e de lanches, eliminação da venda de bebidas alcoólicas são alguns exemplos destas mudanças. Juntamente com as mercearias, são responsáveis pela distribuição de considerável parcela de produtos alimentícios.

Os **minimercados**, de forte expressão nos bairros de classe baixa e de periferia, já apresentam um formato de auto-serviço, apesar de possuírem apenas um, ou no máximo dois, *check-outs*. Apresentam uma linha de produtos semelhantes ao de mercearias.

As **Lojas de Conveniência** adotam um modelo importado dos Estados Unidos na década de 80. Localizadas normalmente junto a postos de combustíveis, apresentam um *mix* de produtos relativamente diversificado. Por operarem em horários ampliados (normalmente no mesmo horário do posto a que estão vinculados), possuem um preço acima da média do mercado. Operam no formato de auto-serviço, com no máximo dois *check-outs*.

Os **supermercados** são caracterizados pelo sistema de auto-serviço, com os produtos dispostos de maneira acessível ao consumidor, que utiliza carrinhos ou cestas durante a estada na loja. O pagamento é realizado através dos chamados *check-outs*, ou seja, balcões providos de máquinas que registram as compras. O *mix* de cada loja depende do porte da mesma, onde

os gêneros alimentícios, de higiene e limpeza, sempre terão o papel de destaque. Entretanto, em lojas maiores é possível encontrar produtos não-alimentícios, como têxteis, vestuário e eletro-eletrônicos.

Os **Clubes atacadistas** são grandes lojas que vendem tanto no atacado quanto no varejo. São caracterizadas por grandes instalações, despojadas, com limitada gama de serviços e preços baixos. Seus clientes são na maioria dos casos institucionais, ou seja, micro e pequenos comerciantes que se utilizam desta estrutura para se abastecerem e atender ao consumidor final.

## 2.6. O varejo não-alimentício

Segundo PARENTE (2000, p.33), o segmento não-alimentício do varejo brasileiro não apresenta o grau de desenvolvimento ou maturidade de sua contra-partida alimentar.

Os principais modelos de varejo não-alimentício existentes no Brasil são descritos a seguir:

- **Lojas Especializadas:** Este tipo de loja domina o varejo não-alimentício. Sua característica principal é a de concentrar suas vendas em uma determinada linha de produtos, tais como livros, discos, calçados, confecções etc. Em geral apresentam um bom sortimento na linha escolhida e atendem um público específico. Existe uma grande variedade de tipos e formatos, além de um porte diversificado que depende do tipo de produto oferecido. Podem ser desde grandes lojas de materiais de construção, com centenas de funcionários até pequenas lojas de perfumes ou confecções, com área reduzida e pouquíssimos funcionários. Estas

lojas podem ser independentes ou fazer parte de uma rede, própria ou franqueada;

- **Lojas de Departamentos:** São grandes lojas, com área de venda superior a 4.000 m<sup>2</sup>, que apresentam uma ampla gama de produtos e serviços estruturados em departamentos distintos. Do ponto de vista organizacional, cada departamento é gerido como uma unidade de negócios autônoma, ou seja, seus gestores tomam as decisões de compra, venda, promoções, metas de venda etc. Isso faz com que estes gestores se especializem nestes segmentos. Tradicionalmente, as lojas de departamentos se baseiam na oferta de produtos da linha “dura” ou “*hard*” (eletrodomésticos, móveis, brinquedos, ferramentas) e da linha “mole” ou “*soft*” (têxteis, confecções, calçados, cama, mesa e banho);
- **Minilojas de Departamentos ou Magazines:** Modelos compactos das grandes lojas de departamento, com maior ênfase na linha “dura”. Com área de venda bem menor (de até 1.000 m<sup>2</sup>), a ambientação departamental não é tão evidente, assim como a gestão, que é mais centralizada. Muitas vezes este tipo de loja é confundida com uma loja especializada, principalmente se concentrar o foco em uma linha específica de departamentos. Operam neste formato as duas maiores redes varejistas não-alimentícias do Brasil: Casas Bahia e Ponto Frio.

## 2.7. O varejo de serviços

Conforme as definições listadas no início deste capítulo, o conceito de varejo abrange também a prestação de serviços. Convém lembrar que pode ser considerado varejista qualquer estabelecimento que forneça produtos ou serviços ao consumidor final.

No setor varejista, existem estabelecimentos em que a prestação de serviços responde a quase totalidade da receita, como nos salões de beleza, academias de ginástica e clínicas médicas, dentre outros.

Existem ainda aqueles em que a prestação de serviços divide o faturamento com a venda de produtos. Como exemplo, uma loja de produtos para animais pode, além de comercializar alimentos e rações, oferecer serviços de banho, tosa e consultas veterinárias.

Segundo PARENTE (2000, p.34), existem quatro características básicas que diferenciam um produto de um serviço:

- **Intangibilidade:** Ao contrário dos produtos, os serviços são menos perceptíveis aos cinco sentidos, ou seja, são menos palpáveis. Por exemplo, academias de ginástica, consultas médicas;
- **Perecibilidade:** Muitos serviços são perecíveis, não podendo ser armazenados. Por exemplo, viagens aéreas;
- **Inseparabilidade:** Ao contrário dos produtos, que primeiro são fabricados para depois serem consumidos, não se pode fazer a separação para os serviços. Por exemplo, o corte de um cabelo;

- **Heterogeneidade:** Podem existir variações do serviço prestado dependendo do dia e/ou da hora. Por exemplo, uma clínica médica com vários especialistas.

O quadro 2.2 apresenta alguns critérios para a classificação dos varejistas de serviços:

**Quadro 2.2. Critérios e tipos de varejo de serviços**

Critérios	Tipos	Exemplos
Grau de Tangibilidade	Produtos Alugados	Locadores de roupas de festas, de veículos
	Produtos de que o consumidor mantém a posse	Oficinas de carros, assistências técnicas de eletrodomésticos
	Sem produtos	Escolas de idiomas, academias de ginástica
Grau de Competência	Alta Especialidade	Clínicas médicas
	Baixa Especialidade	Reparos em sapatos, borracharias
Intensidade de mão de Obra	Baseado em pessoas	Salões de beleza
	Baseado em equipamentos	Lavanderias
Grau de contato com o cliente	Contato alto	Lanchonetes, restaurantes, hotéis
	Contato baixo	Cinemas
Objetivos	Objetivo de lucro	Bancos
	Não-lucrativas	Bibliotecas, postos de saúde

Fonte: PARENTE (2000, p.35), adaptado de BERMAN & EVANS (1998, p. 630)

## 2.8. O setor supermercadista

Dentre todas as formas de varejo existentes, com certeza o setor supermercadista é um dos mais complexos, tanto no planejamento quanto na operacionalização e logística. Tendo em vista a importância desta forma de varejo neste estudo, será feita uma análise deste setor e o impacto do mesmo sobre a economia nacional, além de discorrer sobre as dificuldades enfrentadas para melhorar os resultados apresentados pelo setor.

### 2.8.1. Histórico do setor

Conforme visto anteriormente, as chamadas *general stores* eram grandes lojas que vendiam praticamente de tudo, de gêneros de primeira necessidade

até armas e munições, passando por artigos têxteis, ferragens e outros materiais. Essas lojas possuíam o chamado atendimento tradicional, ou seja, cada cliente era atendido por um funcionário da loja, que se encarregava de apanhar os produtos e embalá-los.

Durante a chamada Grande Depressão, a crise econômica americana derivada da quebra da bolsa de Nova Iorque no ano de 1929, as empresas passaram a ter a necessidade vital de melhorar seus resultados, pois o custo da mão-de-obra era alto e os clientes tinham poucos recursos para dispendir em compras, mesmo as de primeira necessidade. Neste cenário, surgiu então o conceito de auto-serviço, ou seja, o cliente seria o responsável por escolher, apanhar e levar os produtos necessários até a saída da loja, onde então seria realizada a cobrança, através de máquinas registradoras colocadas sobre balcões, os chamados *check-outs*. Segundo KASPER (1981 apud DELUCA, 2003), o conceito surgiu da necessidade das empresas em reduzir seus custos com a mão-de-obra. Desta forma, poderiam oferecer produtos a preços mais baixos que a concorrência.

Em pouco tempo este novo formato de varejo ganhou força, sendo implantado também em outros países. No Brasil, este conceito foi implantado no início da década de 1950, com a inauguração das *Lojas Americanas*, no Rio de Janeiro. Essa loja entretanto atuava no setor de varejo não-alimentício.

O primeiro supermercado, ou seja, a primeira loja de auto-serviço de varejo alimentício, foi a Sirva-se, inaugurada no ano de 1953, na cidade de São Paulo (SM, 2003, p.13).

A partir desta loja pioneira, o setor teve um crescimento vertiginoso, respondendo já por uma média de 6% do PIB nacional nos últimos anos (MERCADOR, 2002, p.76). Sua atividade foi regulamentada pela Lei 7.208, de 13/11/68 (DELUCA, 2003, p.41). O setor apresenta, hoje, um perfil extremamente competitivo, operando sempre com pequenas margens de lucro.

### **2.8.2. Definições**

Antes da definição sobre o que é um supermercado, é necessário que se faça uma clara distinção entre uma empresa de varejo tradicional e uma de auto-serviço.

Segundo FERREIRA (2003), auto-serviço é o “Sistema em que o próprio consumidor ou usuário se serve (em restaurante, supermercado, posto de gasolina, etc.) e paga em caixa ou utiliza mecanismo operado por moeda, ficha, cartão eletrônico, etc”.

Ao contrário, as empresas de varejo tradicional são aquelas em que é necessária a presença de um vendedor ou balconista para selecionar, apanhar e embalar os produtos adquiridos (ACNIELSEN, 2000).

Segundo FERREIRA (2003), supermercado é definido como: “loja de auto-serviço, onde em ampla área se expõe à venda grande variedade de mercadorias, particularmente gêneros alimentícios, bebidas, artigos de limpeza doméstica e perfumaria popular”.

Para PARENTE (2000, p. 32):

“Os supermercados caracterizam-se pelo sistema de auto-serviço, *check-outs* (caixas registradoras sobre balcão na saída da loja) e

produtos dispostos de maneira acessível, que permitem aos fregueses 'auto-servirem-se', utilizando cestas ou carrinhos”.

KOTLER (2000 apud BORGES, 2001) define que:

“Supermercado é o auto-serviço que desenvolve operações relativamente grandes, de baixo custo, baixa margem e alto volume, projetado para atender a todas as necessidades de alimentação, higiene e limpeza doméstica”.

Segundo a consultoria ACNielsen, para ser considerada uma empresa de auto-serviço,

“...é necessário que, além de terem sido classificadas como 'alimentares', tenham como característica fundamental a presença do *check-out*, ou seja, balcão na saída da loja, com caixa registradora, terminal de ponto de venda (PDV), máquina de calcular, máquina de somar e conferência de compras. Além disso, deverá ter carrinhos ou cestas à disposição dos fregueses. A maioria dos produtos, nesses estabelecimentos, é disposta de maneira acessível, permitindo aos fregueses se auto-servirem”. (ACNIELSEN, 2000).

Assim, as características principais de um supermercado são: a possibilidade do próprio cliente pesquisar, escolher e selecionar os produtos, somente tendo contato obrigatório com os funcionários do estabelecimento no momento do pagamento; e a presença dos chamados *check-outs*, onde este pagamento é realizado, assim como a embalagem dos produtos para transporte.

### 2.8.3. Situação atual

O setor de auto-serviço é responsável hoje por cerca de 6,2% do total do Produto Interno Bruto nacional, com um faturamento da ordem de 72,5 bilhões de reais no ano de 2001 (SUPERHIPER, 2002, p. 12). Grande empregador, fechou o ano de 2001 com quase 720 mil empregos diretos gerados, em um total de 69.396 lojas. O quadro 2.3 mostra uma síntese da evolução do setor nos últimos anos (n.d. – dados não disponíveis):

**Quadro 2.3. Síntese do setor de auto-serviço no ano de 2001**

	1996	1997	1998	1999	2000	2001
Número de Lojas	43.763	47.787	51.502	55.313	61.259	69.396
Faturamento Anual (em R\$ bilhões nominais)	46,8	50,4	55,5	60,1	67,6	72,5
Participação % do Faturamento Sobre o PIB	6,2	6,02	6,1	6,0	6,2	6,2
Número de Empregos diretos	625.000	655.000	666.752	670.086	701.622	710.743
Área de Vendas (em m <sup>2</sup> milhões)	n.d.	12	12,7	13,1	14,3	15,3
Nº <i>Check-outs</i>	n.d.	123.170	125.867	135.914	143.705	156.022

Fonte: SUPERHIPER (2002, p.14)

Analisando os dados acima, é possível verificar que o setor passou por uma acentuada expansão a partir do ano de 1996. O faturamento cresceu cerca de 55%, o número de lojas cresceu aproximadamente 58,5%, com um aumento da área de vendas da ordem de 27,5% (comparado a 1997), enquanto que o incremento total de empregos diretos gerados foi de apenas 13,7%. Houve um grande aumento de produtividade no setor, quando o faturamento por funcionário saltou de R\$ 74.880,00 (em 1996) para R\$ 102.005,93 (em 2001), com um incremento de 73,4%.

Para efeitos de ranking, a Associação Brasileira de Supermercados (ABRAS) considera que, para ser definida como supermercado, a loja de auto-serviço deve apresentar um mínimo de 02 (dois) *check-outs*. As lojas com apenas 01 (um) *check-out* são classificadas como minimercados ou lojas de conveniência, dependendo do *mix* de produtos, da localização e do horário de funcionamento. Assim, o quadro 2.4 mostra uma relação entre os totais do setor de auto-serviço e os referentes aos supermercados.

**Quadro 2.4. Comparativo entre o setor de Auto-Serviço e o segmento de Supermercados. Síntese dos resultados do ano 2001**

	Outros Estab. Auto Serviço	Supermercados	Auto-Serviço Total
Faturamento (em R\$ bi)	6,2	66,3	72,5
Nº de lojas	39.793	29.603	69.396
Nº Check-Outs	39.793	116.229	156.022
Nº Funcionários	167.179	543.564	710.743
Área de Vendas em m2 (milhões)	3,8	11,5	15,3

Fonte: SUPERHIPER (2002, p. 16)

Na análise desses dados é possível provar que o segmento de supermercados apresenta um resultado muito superior ao total do segmento:

- Enquanto o faturamento por loja com um único *check-out* foi da ordem de R\$ 155.806,30 no ano, o faturamento por loja do segmento de supermercados foi de R\$ 2.239.637,87 no ano - um faturamento mais de 14 vezes superior;
- O faturamento por funcionário, nos supermercados, ficou na média de R\$ 133.378,95. Já nos demais estabelecimentos, esta média ficou em R\$ 37.086,00. Ou seja, um funcionário de supermercado gera um faturamento 3,5 vezes superior ao funcionário de um minimercado;
- O faturamento por área de venda mostra que enquanto as pequenas empresas geram R\$ 1.631,57 por m<sup>2</sup>, um supermercado fatura uma média de R\$ 4.738,56 pela mesma unidade de área.

#### **2.8.3.1. O Ranking da Associação Brasileira de Supermercados – ABRAS**

A Associação Brasileira de Supermercados (ABRAS) elabora anualmente um ranking das empresas associadas. Este estudo tem por objetivo traçar um retrato do setor em termos de faturamento e produtividade. Para efeitos de ranking, foram consideradas apenas as empresas de auto-serviço que apresentam pelo menos 02 (dois) *check-outs*. Como base para o presente trabalho, será utilizado o ranking relativo ao ano de 2001, divulgado no mês de maio de 2002 pela revista Superhiper, órgão de divulgação oficial da entidade.

O ranking é baseado no faturamento bruto conseguido pelas empresas no período de 01 de janeiro a 31 de dezembro do ano em estudo. Além dessa classificação oficial, é possível realizar novas classificações a partir dos chamados *indicadores de desempenho*, dentre os quais se destacam: o faturamento por funcionário, o faturamento por metro quadrado e o faturamento por *check-out*.

Na adoção de algum destes indicadores como base para o ranking, ocorrerão alterações substanciais nas posições do mesmo. Como um exemplo, tendo como base o faturamento por funcionário, o líder do ranking cai para a oitava posição, enquanto que a liderança passaria a ser ocupada pelo 36º colocado do ranking oficial.

Para efeitos de estudos e comparações, dadas as grandes diferenças de porte entre as empresas pesquisadas, divide-se a relação em vários grupos: 5 maiores, 20 maiores, 50 maiores, 300 maiores etc. O quadro 2.5 mostra uma relação simplificada do ranking, listando apenas o grupo chamado de “50 maiores”. A análise deste quadro já oferece uma boa dimensão da diversidade do setor, principalmente no que tange ao porte das empresas. O faturamento

dos cinco primeiros colocados do ranking (28,35 bilhões de reais) é 45% superior à soma do faturamento das outras 450 empresas que compõem o ranking (19,54 bilhões de reais).

**Quadro 2.5. Ranking simplificado – ABRAS/2001**

	Empresa	Sede	Fat. Bruto (em milhões de R\$)	Check-outs	Área de Vendas	Num. Lojas	Num. Func.
1	Cia. Brasileira Distribuição	SP	9.857,53	6.891	866.280	443	52.060
2	Carrefour Com. Ind. Ltda.	SP	9.236,68	7.240	1.015.485	227	45.424
3	Sonae Distribuição do Brasil S/A	RS	3.411,42	3.677	445.483	168	21.028
4	Bompreço S/A Sup. Nordeste	PE	3.222,58	2.532	312.032	110	19.344
5	Casas Sendas Com. Ind. S/A	RJ	2.622,10	1.910	207.728	84	14.417
6	Wal-Mart Brasil S/A	SP	1.482,90	777	191.751	22	6.828
7	Jerônimo Martins Dist. Brasil Ltda.	SP	1.043,64	963	97.273	62	6.984
8	Cia. Zaffari Com. Ind.	RS	838,62	799	96.417	22	6.850
9	G. Barbosa Com. Ltda.	SE	730,67	690	61.762	35	5.764
10	COOP Coop. De Consumo	SP	639,46	338	47.331	17	3.188
11	Irmãos Bretas, Filhos Cia. Ltda	MG	601,37	546	66.955	34	4.209
12	A. Angeloni & Cia. Ltda.	SC	542,87	394	49.268	17	4.468
13	Líder Sup. e Mag. Ltda.	PA	431,36	303	36.000	11	3.286
14	Sonda Sup. Exp. Imp. Ltda.	SP	429,63	345	31.967	9	2.335
15	DMA Distribuidora S/A	MG	417,50	471	45.589	42	3.370
16	Sup. Mundial Ltda	RJ	412,59	378	24.985	15	2.812
17	Condor Supercenter Ltda.	PR	340,71	324	40.757	18	2.715
18	D'Ávo Supermercados Ltda.	SP	285,13	296	37.200	7	1.991
19	Y. Yamada S/A Com. Ind.	PA	270,62	190	21.673	12	2.050
20	Sup. Bahamas Ltda.	MG	237,47	159	18.123	17	1.404
21	Emp. Bahiana de Alim. Ltda.	BA	224,47	1.160	116.200	418	4.358
22	Carvalho e Fernandes Ltda.	PI	205,30	242	22.842	27	1.445
23	Sup. Gimenes Ltda	SP	204,19	225	21.834	24	1.142
24	Sup. Nordeste Ltda	RN	203,57	119	13.072	6	1.030
25	Savegnago Sup. Ltda.	SP	193,14	156	15.047	13	1.012
26	Sup. Bergamini Ltda.	SP	178,94	177	16.000	3	1.350
27	Sup. Irmãos Lopes Ltda.	SP	171,84	184	19.610	11	1.260
28	Unisuper Distr. S/A	ES	166,97	200	18.500	19	1.350
29	Nazaré Coml. Alim. Mag. Ltda	PA	166,32	75	8.300	3	1.050
30	Coml. Nazaré S/A	ES	154,50	165	14.910	8	1.037
31	Giassi & Cia. Ltda.	SC	145,33	166	21.106	8	1.482
32	Império da Banha Ltda.	RJ	144,47	128	9.950	12	1.100
33	Formosa Sup. Mag. Ltda.	PA	140,43	89	12.000	2	946
34	Dias Pastorinho S/A Com. Ind.	SP	131,18	144	16.420	9	1.086
35	Enxuto Coml. Ltda.	SP	128,96	141	11.686	7	792
36	Viabrazil Com. Ind. Ltda.	MG	126,94	58	7.641	2	541
37	Sup. Lusitana Ltda.	MA	124,07	163	18.510	16	976
38	Supermercados Imperatriz Ltda.	SC	123,95	141	15.362	19	1.111
39	Emp. Pernambucana de Alim.	PE	122,17	128	11.840	12	671
40	Coml. Delta Ponto Certo Ltda.	SP	121,61	100	8.550	6	998
41	Coml. Unida de Cereais Ltda.	RS	115,23	169	17.176	20	1.334
42	Sup. Superpão Ltda.	PR	111,76	162	17.600	11	1.073
43	Futurama Sup. Ltda.	SP	109,10	98	11.700	8	880
44	Santa Cruz Imp. Com. Alim. Ltda.	GO	108,98	90	11.000	4	745
45	Sup. Alto da Posse Ltda.	RJ	108,65	135	8.759	9	1.350
46	Jad Zogheib & Cia. Ltda.	SP	106,08	52	4.800	3	589
47	Coop. Caf. Citric. SP	SP	100,86	100	8.947	7	755
48	Casa Avenida Com. Imp. Ltda.	SP	100,16	100	15.710	15	822
49	Covabra Coml. Ltda.	SP	98,85	121	8.200	7	510
50	Coop. Consumo Cooperica	SP	91,08	77	5.650	3	495

Fonte: Adaptado de SUPERHIPER (2002, p.38-40)

Estas cinco empresas detêm uma participação equivalente a 39,1% do setor, deixando os outros 60,9% para serem disputados pelas demais empresas.

É importante observar que apenas a primeira colocada (Companhia Brasileira de Distribuição – CBD), proprietária das bandeiras Extra, Pão de Açúcar, Barateiro e Eletro, é uma empresa com controle acionário em mãos brasileiras. As próximas quatro empresas são controladas por grupos estrangeiros que aportaram no Brasil a partir da década de 80 e, principalmente, a partir da estabilização econômica advinda do Plano Real: O Carrefour (bandeiras Carrefour, RDC, Stoc, Champion e Dia Brasil) é uma empresa de capital francês; o grupo Sonae (bandeiras Big, Cândia, Mercadorama, Nacional e Maxxi) é de origem portuguesa; o grupo Bompreço (bandeiras Bompreço e Balaio) é controlado pela Ahold, empresa de capital holandês.

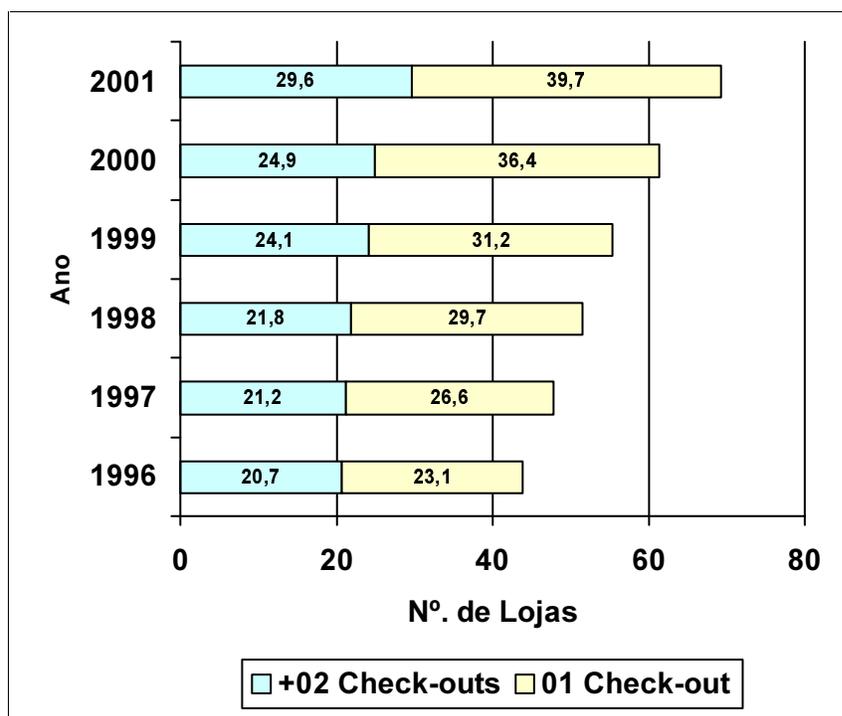
#### **2.8.3.2. Crescimento e estrutura das lojas**

A figura 2.3 mostra o crescimento em número de lojas no setor de auto-serviços. Na comparação entre o crescimento dos dois segmentos nos últimos anos, nota-se que o número de lojas com apenas 01 (um) *check-out* teve um crescimento de mais de 58% entre os anos de 1996 e 2001. Enquanto isso, o setor supermercadista apresentou um crescimento de cerca de 43% no mesmo período, com um maior aumento entre os anos de 2000 e 2001, quando demonstrou um crescimento de cerca de 18,9% em apenas um ano.

Estes aumentos podem ser explicados, segundo a ABRAS (SUPERHIPER, 2002, p. 16), com a transformação de lojas que atendiam pelo método tradicional para o método de auto-serviço com um único *check-out* passando,

posteriormente, para o segmento de supermercados com a adição de mais *check-outs*.

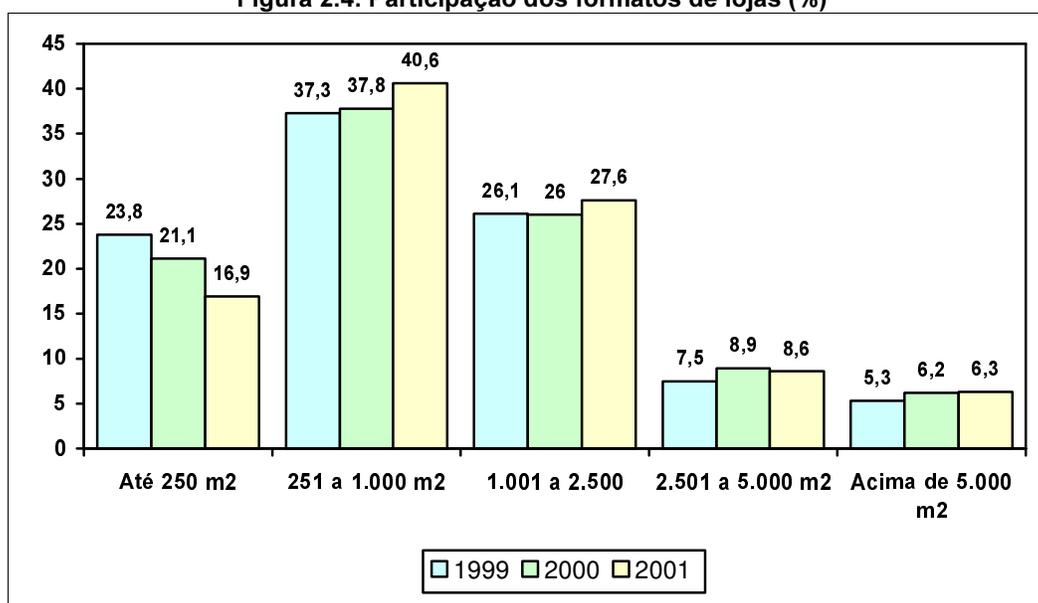
**Figura 2.3. Evolução do número de lojas no setor de auto-serviço**



Fonte: adaptado de SUPERHIPER (2002, p.15)

Com relação ao tamanho físico das lojas, existe uma forte concentração na faixa de até 2.500 m<sup>2</sup>, conforme mostra a figura 2.4, que apresenta uma amostragem de lojas por área nos anos de 1999 a 2001. Analisando o gráfico em questão, destaca-se a redução no número de lojas na primeira faixa (até 250 m<sup>2</sup>), com conseqüente aumento na quantidade das mesmas nas segunda e terceira faixas (entre 251 e 2.500 m<sup>2</sup>). Segundo a ABRAS, "O crescimento da participação das lojas com área entre 251 e 2.500 m<sup>2</sup> confirma a tendência de expansão das lojas de vizinhança. Já o formato hipermercado se encontra estável" (SUPERHIPER, 2002, p.31).

Figura 2.4. Participação dos formatos de lojas (%)



Fonte: adaptado de SUPERHIPER (2002, p. 31)

#### 2.8.3.3. Recursos Humanos

O setor supermercadista é um dos agentes com maior capacidade de geração de empregos na economia nacional. O número de empregos diretos gerados pelo setor chegou a 710,7 no ano de 2001, com um crescimento de 1,3% em relação ao ano 2000, quando gerou 701,6 empregos diretos (SUPERHIPER, 2002, p. 12).

Entretanto, o aumento no número de empregos não acompanhou o crescimento no número de lojas no mesmo período (13,37%). Isto indica que o setor está crescendo baseado no aumento da produtividade.

#### 2.8.3.4. Automação e produtividade pessoal

O setor investiu muito nos últimos anos na automação de seus processos. Hoje o índice de automação entre as 500 maiores empresas do setor chega a impressionantes 99,5% (SUPERHIPER, 2002, p.22). Este é um dos motivos apontados pela ABRAS para a discrepância entre o crescimento no número de lojas e o aumento no número de empregos diretos no setor. Analisando as

empresas listadas no ranking da ABRAS, nota-se que quanto mais próximo ao topo, maior a produtividade por funcionário, conforme o quadro 2.6:

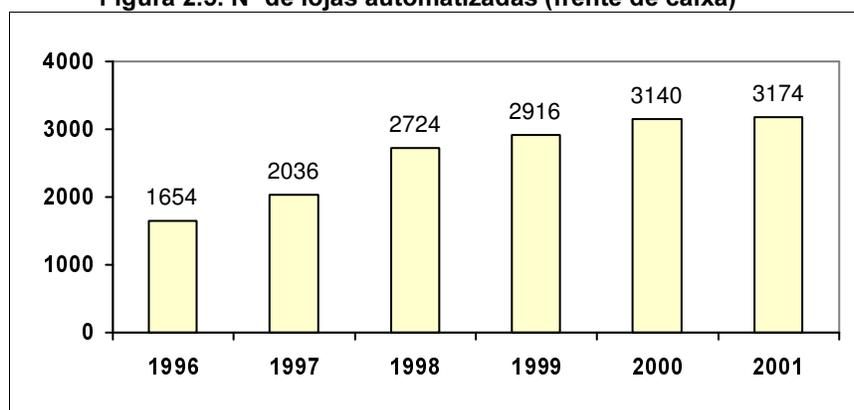
**Quadro 2.6. Faturamento médio por funcionário (base ano 2001)**

Posição	Faturamento Médio
1º à 5º	186.181
6º à 10º	159.900
11º à 20º	138.592
21º à 50º	127.039
51º à 100º	110.501
101º à 200º	109.190
201º à 300º	116.985

Fonte: adaptado de SUPERHIPER (2002, p.22)

Como o crescimento no índice de automação entre os anos de 2000 e 2001 foi de apenas 1,1% (conforme figura 2.5), é simples chegar à conclusão de que o crescimento registrado é apenas orgânico. Isso indica que o próximo passo de qualificação do setor será o de investir fortemente no gerenciamento dos dados adquiridos nos processos de venda (SUPERHIPER, 2002, p.22), sendo este um dos pilares do presente estudo.

**Figura 2.5. Nº de lojas automatizadas (frente de caixa)**



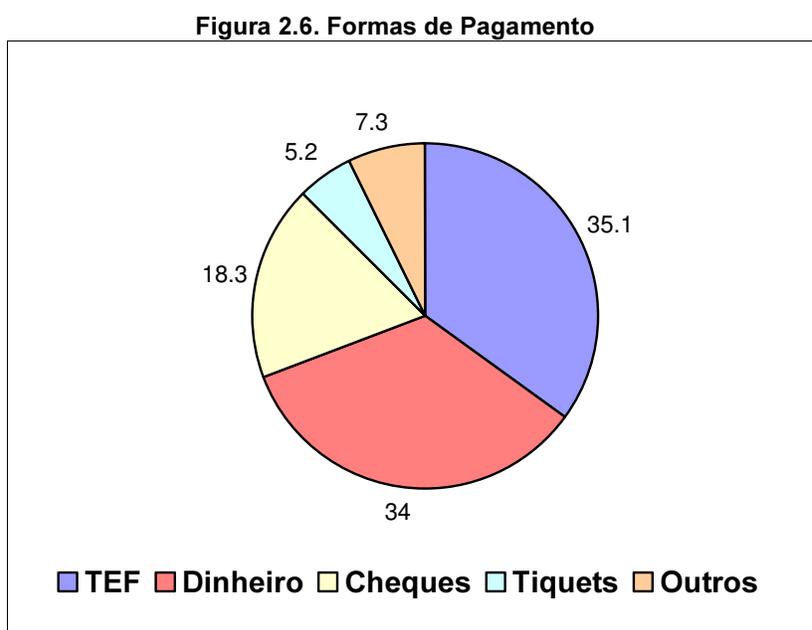
Fonte: adaptado de SUPERHIPER (2002, p.22)

Além disso, quanto maior o porte da empresa, maior a necessidade de interconectar os equipamentos em redes locais e de longa distância, com o objetivo de facilitar o fluxo de informações em tempo real, além de permitir pagamentos através de meios eletrônicos, tais como cartões de crédito e

débito, além de consultas on-line à administradoras de crédito, como SPC e SERASA.

#### 2.8.3.5. Meios de pagamento

A figura 2.6 mostra as principais formas de pagamento utilizadas pelos clientes de supermercados e suas participações na movimentação financeira total.



Fonte: SUPERHIPER (2002, p. 33)

Conforme a figura, os pagamentos realizados através da chamada Transferência Eletrônica de Fundos (TEF) são os mais numerosos. Entretanto, ao considerar que esta categoria pode ser subdividida em outras três subcategorias (cartão de crédito tradicional, cartão de crédito da empresa e cartão de débito), podemos considerar que o meio de pagamento mais utilizado ainda é o dinheiro em espécie.

Segundo a ABRAS (SUPERHIPER, 2002 p.33), existe uma tendência para o aumento de pagamentos por meio eletrônico, ao mesmo tempo em que

diminui a participação dos pagamentos com cheques, tanto à vista quanto pré-datados.

Mesmo respondendo por apenas 18,3% dos pagamentos, os cheques são considerados como uma das principais fontes de prejuízos para o setor. Além daqueles sem a devida provisão de fundos, os cheques roubados ou fraudados constituem uma parcela importante nos prejuízos advindos de problemas relacionados com o pagamento das compras por parte dos clientes. Este fato obriga as empresas a manterem um cadastro com os clientes emitentes de cheques, além de setores de cobrança e controle de crédito para evitar perdas com cheques devolvidos.

#### **2.8.4. Massificação dos clientes**

Com o desenvolvimento do setor e o aumento nos volumes de transações diárias, o cliente passou a ser encarado como um ser imaterial, cuja única função era a de ir à loja e adquirir os produtos. Já não mais eram conhecidos pelo primeiro nome, como nos antigos armazéns de varejo tradicional, nem seus hábitos de compra eram conhecidos. Os clientes passaram a fazer parte de uma massa, onde eram enquadrados em agrupamentos conforme algumas características geográficas, sociais, demográficas etc.

Neste estágio a tecnologia, a redução de custos, a briga com fornecedores e novos sistemas de gestão passaram a ser o principal foco do negócio. Tudo com o objetivo de melhorar a eficiência do setor. Durante muito tempo acreditou-se que o conceito do auto-serviço era a impessoalidade (SM, 2003, p.24).

Com o acirramento da concorrência (e o próprio ranking oficial da ABRAS é uma mostra disso), o setor está mudando este conceito. Já não basta mais apenas conquistar novos clientes. A chave está em reter os clientes, impedindo que os mesmos migrem para a concorrência, através de um processo denominado de **fidelização**, pois o cliente fiel sempre trará um alto retorno a empresa, tendo em vista que irá realizar uma quantidade maior de transações com a mesma do que o cliente eventual.

A adoção de programas de fidelização de clientes, uma realidade para os primeiros colocados no ranking, tornou-se uma tendência para as redes de porte médio, que fazem concorrência direta aos minimercados e pequenos armazéns e padarias, ainda sob a ótica do varejo tradicional e clientela reduzida, porém cativa.

Estes programas são, normalmente, baseados na obtenção de pontos que, posteriormente, podem ser trocados por produtos e/ou serviços, à semelhança dos programas de milhagem das companhias aéreas. Entretanto, o volume de dados a serem manipulados é muito grande. Uma rede de porte médio pode gerar, em apenas um único dia, mais de cem megabytes de dados de venda.

Para fazer frente a esta demanda, a tecnologia da informação (TI) dispõe de várias ferramentas, que serão tratadas no próximo capítulo, juntamente com a revisão teórica do marketing de relacionamento e da gestão de relacionamentos com clientes (Customer Relationship Management – CRM).

### **3. MARKETING DE RELACIONAMENTO E CRM**

Este capítulo traz uma pequena introdução ao marketing de relacionamento e ao CRM (*Customer Relationship Management* – Gestão de Relacionamento com Clientes). Suas principais definições estão colocadas em ordem cronológica, como uma introdução a essa fascinante área de estudo. O desenvolvimento do marketing de relacionamento depende de alguns pressupostos básicos, quais sejam o conhecimento do cliente, a personalização das ofertas e/ou produtos, o correto acesso ao mesmo e, por fim, o atendimento diferenciado. Para conhecer o cliente é necessário lançar mão de algumas técnicas e ferramentas específicas, de maneira a formar um quadro completo de seu perfil e potencialidades. A oferta de produtos e/ou serviços personalizados ou que possam ser montados em um pacote pelo próprio cliente, aumenta o valor do seu relacionamento com a empresa. Para alcançar o cliente, meios especiais de acesso personalizado mostram-se mais eficazes do que a mídia de massa. Por último o atendimento diferenciado, como forma de deixar o cliente à vontade, além de toda a estrutura necessária para o suporte dos contatos futuros do mesmo com a empresa.

#### **3.1. A necessidade da retenção do cliente**

As empresas atuam hoje em um ambiente altamente competitivo. A concorrência se desdobra para conquistar uma base de clientes que permita sua sobrevivência. As inovações adicionadas aos produtos trazem uma vantagem inicial, mas com o tempo essas inovações são imitadas pela concorrência. A globalização e a internet só pioram este quadro, pois a concorrência está apenas a um “click” de distância. Neste cenário, as

empresas estão buscando formas de se destacar em relação à concorrência. Os estrategistas de marketing estão aos poucos deixando de lado a idéia de que o importante é o produto.

Modelos como os chamado 4P's<sup>1</sup>, proposto por McCarthy (apud PECK, 1999) no início da década de 1960, já não são suficientes. Estes modelos foram desenvolvidos a partir de estudos sobre o mercado americano de bens de consumo durante os anos de pós-guerra (décadas de 1950 e 1960). Neste ambiente, por causa do grande aumento na demanda de consumo, as empresas tinham pouca razão ou incentivo, para considerar os relacionamentos com os clientes além daqueles necessários para efetivar uma única transação comercial (PECK, 1999). Desta forma, as estratégias de marketing tradicionais sempre foram mais direcionadas para *atrair* do que para *manter* os clientes.

Já no início da década de 1980, uma abordagem alternativa para a teoria (e prática) do marketing tradicional começou a ser delineada: o Marketing de Relacionamentos, onde o foco deixou de ser o produto/serviço e passou a ser o cliente.

### **3.2. O que é o Marketing de Relacionamentos**

O termo foi inicialmente utilizado na literatura especializada do setor de serviços e tratava do desenvolvimento e cultivo de relacionamentos lucrativos e mutuamente benéficos, de longa duração, entre uma organização e um grupo definido de clientes (PECK, 1999).

---

<sup>1</sup> Modelo que propunha o chamado "marketing mix", composto de *Product* (produto), *Price* (preço), *Place* (local de distribuição ou de localização) e *Promotion* (promoção, propaganda).

É nesse sentido que surge uma das primeiras definições de Marketing de Relacionamentos, atribuída a BERRY (1983): “É atrair, manter e - em organizações de multiserviços - ampliar relações com clientes”.

Com o passar do tempo, a definição de marketing de relacionamentos foi sendo ampliada.

PAYNE (1991) define que:

“Marketing de Relacionamentos trata sobre como maximizar o valor do cliente para a empresa, criando, construindo e prolongando relacionamentos de clientes com o fim de vender mais, fazer venda cruzada e manter clientes no longo prazo”.

MCKENNA (1992) traz uma visão mais abrangente em termos organizacionais:

“... é construir e sustentar infra-estrutura dos relacionamentos de clientes. É a integração dos clientes à empresa, desde o projeto, desenvolvimento, processos industriais e de vendas... Todos os empregados precisam estar no negócio de construir relacionamentos de clientes”.

VAVRA (1993) mostra de uma visão mais focada no cliente:

“É o processo de proporcionar satisfação contínua e reforço aos indivíduos ou organizações que são clientes atuais ou que já foram clientes. Os clientes devem ser identificados, reconhecidos, comunicados, auditados em relação à satisfação e respondidos. O objetivo do pós-marketing é construir relacionamentos duradouros com todos os clientes”

GRÖNROOS (1996), traz a seguinte definição:

“O Marketing de relacionamento é identificar, estabelecer, manter e ampliar relações lucrativas com clientes e outros interessados, de forma que os objetivos de todas as partes envolvidas sejam encontrados; e isto é feito por troca mútua e cumprimento de promessas”.

STONE e WOODCOCK (1998) trazem outra definição:

“É o uso de uma ampla gama de técnicas e processos de marketing, vendas, comunicação e cuidado com o cliente para identificar seus clientes de forma individual e nominal; criar um relacionamento entre sua empresa e esses clientes - um relacionamento que se prolonga por muitas transações; administrar esse relacionamento para o benefício dos seus clientes e da sua empresa.”.

GORDON (1999) diz que Marketing de Relacionamentos:

“... é o processo contínuo de identificação e criação de novos valores com clientes individuais e o compartilhamento de seus benefícios durante uma vida toda de parceria. Isso envolve a compreensão, a concentração e a administração de uma contínua colaboração entre fornecedores e clientes selecionados, para a criação e o compartilhamento de valores mútuos”.

Analisando estas definições, é possível extrair algumas informações:

- O foco deixou de ser o produto/serviço ofertado, passando a ser direcionado para o cliente. Isso não significa dar tudo para o cliente. Significa identificar suas necessidades e atendê-las (STONE & WOODCOCK, 1998);

- A ênfase passa a ser dada no sentido de manter (e reter) os clientes já existentes, pois é muito mais barato manter um cliente do que captar um novo. Uma empresa gasta entre cinco e seis vezes mais para conquistar um cliente do que para manter um já conquistado (ROSEMBERG & CZEPIEL, 1984; BUCHANAN & GILLIES, 1990; STONE et al, 2001);
- Para se conseguir um relacionamento perfeito, é necessário *conhecer* o cliente, ou seja, conhecer o seu padrão de compras e suas necessidades básicas. Uma empresa só será capaz de entender e atender as necessidades de seus clientes se os colocar em primeiro lugar (STONE & WOODCOCK, 1998);
- O relacionamento deve ser mútuo, ou seja, produtivo, agradável e rentável para ambas as partes (empresa e cliente);
- Deve-se criar novos valores durante o relacionamento; O relacionamento pode, e deve, ser mantido durante todo o chamado “ciclo de vida” do cliente, ou seja, aquele período em que ele tem condições comerciais de manter um relacionamento comercial com a empresa (GORDON, 1999);
- Existe a necessidade de se envolver toda a empresa na estratégia de relacionamento, não apenas os setores envolvidos diretamente com os clientes (GORDON, 1999);
- As empresas devem ter a capacidade de antecipar necessidades e desejos dos clientes, buscar permanentemente a sua fidelidade e utilizar largamente estratégias de personalização (MARINHO, 2003).

Desta forma, o planejamento estratégico das empresas deve, a partir de agora, levar em consideração todos esses aspectos, com o objetivo de se criar estratégias de longo prazo para se reter os clientes em carteira. Estas estratégias deverão envolver necessariamente todos os funcionários da empresa, com o objetivo de encantar o cliente a cada nova transação realizada.

### **3.3. Customer Relationship Management (CRM)**

O conceito de CRM surgiu praticamente junto com o de Marketing de Relacionamento. Algumas das definições do que seja a Gestão de Relacionamento com Clientes (CRM, em sua sigla em inglês) estão relacionadas à seguir.

STONE et al (2000) diz que CRM é:

“a utilização de uma ampla gama de abordagens de marketing, vendas, comunicação, serviço e atendimento aos clientes para identificar os clientes individuais de uma empresa, criar um relacionamento duradouro e vantajoso entre a empresa e os clientes e gerenciar esse relacionamento para beneficiar os clientes e a própria empresa”.

Para SWIFT (2001):

*“Gerência de Relacionamento com Clientes é uma abordagem empresarial destinada a entender e influenciar o comportamento dos clientes, por meio de comunicações significativas para melhorar as compras, a retenção, a lealdade e a lucratividade deles (...) CRM é um processo interativo que transforma informações sobre os clientes em relacionamentos positivos com os mesmos. (...) CRM deve ser integrado*

em tudo que a empresa faz, todos com os quais ela trabalha (até mesmo fornecedores) e todos com os quais ela transaciona.” (grifos do autor).

Através desta definição, SWIFT mostra que o conceito de marketing de relacionamento necessita do apoio de uma ferramenta poderosa para “transformar informações sobre os clientes em relacionamentos positivos com os mesmos”.

Já para BROW (2001):

“O Customer Relationship Management-CRM (...) é uma estratégia de negócios que visa entender, antecipar e administrar as necessidades dos clientes atuais e potenciais de uma organização. É uma jornada de estratégias, processos, mudanças organizacionais e técnicas pelas quais a empresa deseja administrar melhor o seu empreendimento acerca do comportamento dos clientes. Acarreta em adquirir e distribuir conhecimento sobre os clientes e usar essas informações por meio de vários pontos de contato para equilibrar rendimentos e lucros com o máximo de satisfação dos clientes”.

Nesta definição, o autor mostra também a necessidade de uma ferramenta capaz de “distribuir conhecimento sobre clientes” e de utilizar este conhecimento nos diversos setores da empresa.

Em sua definição, BRETZKE (2003) mostra que ferramenta é esta:

“Customer Relationship Management ou Gerenciamento do Relacionamento, como o próprio nome indica é a integração entre o marketing e a tecnologia da Informação para prover a empresa de meios mais eficazes e integrados para atender, reconhecer e cuidar do cliente,

em tempo real e transformar estes dados em informações que disseminadas pela organização permitem que o cliente seja 'conhecido' e cuidado por todos”.

Mais adiante, a mesma autora nos diz:

“O CRM é a combinação da filosofia do marketing de Relacionamento, que ensina a importância de cultivar os clientes e estabelecer com os mesmos um relacionamento estável e duradouro através do uso intensivo da informação, aliado a tecnologia da informação que prove os recursos de informática e telecomunicações integrados de uma forma singular”.

Assim, os autores mostram que a Tecnologia da Informação (TI) é uma ferramenta básica para se aplicar os conceitos de marketing de relacionamento.

BRETZKE (2000) traz a seguinte definição de TI:

“A Tecnologia da Informação (TI) deve ser entendida como a coleta, processamento, armazenamento, apresentação e transmissão da informação, em toda as suas formas. Inclui computadores, telecomunicações, redes de informações, internet, superinforvias, *on line data bases*, *softwares*, bem como fax, telefonia móvel, televisão à cabo e outras formas de comunicação pessoal e de massa”. (grifos da autora).

STONE et al (2000) justificam o uso da tecnologia da seguinte forma:

“... as novas tecnologias forneceram as ferramentas de ataque ideais, que lhes permitem escolher os clientes mais lucrativos ou os já existentes, oferecendo a esses clientes níveis de serviços sem

precedentes e que, há algum tempo, só seriam possíveis por meio de grandes investimentos físicos”.

Entretanto, apesar da TI ser uma ferramenta básica em estratégias de CRM, vale ressaltar que CRM *não é* um programa específico ou um sistema de informática, pura e simplesmente, como algumas empresas tentam vender. Como visto nas definições acima, CRM é, antes de tudo, um conjunto de técnicas de marketing que se valem do poder proporcionado pelas ferramentas de TI.

A seguir serão mostrados os pressupostos básicos do marketing de relacionamentos e como a TI aparece como poderosa ferramenta para implementação das técnicas de CRM.

### **3.4. Pressupostos do marketing de relacionamentos e do CRM**

Para atingir as metas propostas, o planejamento estratégico de uma empresa deve se basear em alguns pressupostos básicos. Esses pressupostos incluem:

- Conhecer o cliente;
- Fornecer produtos e/ou serviços personalizados para o cliente, afim de que ele possa satisfazer suas necessidades;
- Promover a comunicação individual com os clientes;
- Atender de maneira eficaz e personalizada cada cliente;
- Reter os melhores clientes.

A seguir estes tópicos serão analisados.

### 3.4.1. Conhecimento do cliente

Conhecer o cliente é o fundamental de todo processo de relacionamento. É através do conhecimento do perfil e das necessidades que a empresa terá a capacidade de criar serviços e/ou produtos feitos sob medida para atender as necessidades de cada um de seus clientes, criando assim um alto valor para o relacionamento.

SEYBOLD & MARSHAK (2000) dizem a respeito das estratégias competitivas:

“Qual a fórmula vencedora? Ela começa enfocando a base atual de clientes, tentando descobrir o que eles desejam e de que precisam e como sua vida pode ser facilitada. Depois, você pode expandir seus esforços, atraindo clientes em potencial. Com os clientes perto de você, fechar a venda e consolidar um relacionamento rentável e duradouro é tão fácil quanto estalar os dedos, porque já ficou fácil para o cliente fazer negócios com você!”.

STONE & WOODCOCK (1998) afirmam que:

“as informações acerca de seus clientes e respectivas necessidades devem ser levadas em conta durante o planejamento e devem estar disponíveis em um formato utilizável em pontos de contato específicos com o cliente. (...) O uso dos dados dessa maneira, tanto para o planejamento como para fins de comunicação, é importante para a ‘implementação da estratégia’ e para tornar seus planos mais visíveis para o cliente”.

Os profissionais de marketing tradicional, acostumados a trabalhar sempre com segmentos de mercado, agora se deparam com uma outra situação. Seja considerando os clientes demográfica, psicográfica ou comportamentalmente, ou até mesmo pelo estilo de vida, está cada vez mais difícil classificar os clientes em segmentos específicos (GORDON, 1999).

GORDON (1999) ainda enfatiza:

“Se a única categorização significativa é o comportamento real do comprador - o que as pessoas ou empresas compram, em vez de orientações subjacentes a esse comportamento - então *não existem mais segmentos de mercado, apenas clientes individuais*. (...) Se os segmentos de mercado não existem mais, os clientes precisam ser atendidos como indivíduos. As empresas podem dar aos clientes individuais, ou grupos lógicos de clientes, o valor que cada um deseja, utilizando a tecnologia adequadamente e por meio de toda a cadeia de valor”.

GORDON (1999) utiliza uma frase muito feliz neste contexto: “Segmentos não gastam dinheiro; clientes sim”.

Radicalizando este conceito, PEPPERS & ROGERS (2003) propõe o chamado *marketing one-to-one* (marketing um-para-um), da seguinte forma:

“*One-to-One* quer dizer relacionar-se com clientes de forma individual, um por vez. Assim, conhecendo cada um de nossos clientes, podemos identificar suas necessidades e trata-los de forma personalizada. *One-to-one*, em outras palavras, significa **‘tratar clientes diferentes de forma diferente’**. Focado no cliente individual, o marketing *one-to-one* é

baseado na idéia de uma empresa conhecer os seus clientes. Através de interações com estes consumidores, a empresa pode aprender como ele ou ela gostariam de serem tratados. A empresa estará apta, então, a tratar estes consumidores de forma diferente do que trata a outros diferentes”. (grifos do autor).

Assim, o conhecimento total do cliente é um fator chave nas estratégias de marketing de relacionamento. Conhecer os clientes é relativamente simples para pequenos estabelecimentos, como padarias e mercearias de bairro, ou para aqueles que tem uma clientela selecionada, como uma loja de alta costura. Para os grandes estabelecimentos, como os supermercados e grandes lojas de departamentos, esse “contato íntimo” é extremamente difícil. É muito mais fácil ser surpreendido por seu alfaiate, com a oferta gratuita da confecção de uma camisa do seu gosto, por exemplo, do que por um supermercado de venda em massa, que não sabe seu nome, não iniciará uma conversa e não fará muito por você (GORDON, 1999). Desta forma, as empresas menores possuem, em primeira instância, uma grande vantagem sobre as maiores. Entretanto, a tecnologia permite que essas empresas coloquem mais informação em sua memória organizacional sobre cada cliente do que as pequenas empresas conseguem fazer (GORDON, 1999).

Com os custos de armazenamento de dados em queda, está cada vez mais fácil para as empresas alimentar e manter sistemas de informações a respeito de seus clientes. Cada vez mais a tecnologia da informação (TI) é a ferramenta básica para tratar dos relacionamentos. GORDON (1999) diz:

“Atualmente, existe o potencial para os negociantes que vendem em massa preverem as necessidades dos clientes individuais. Quando abrem suas portas pela manhã, eles podem prever quem entrará. E podem dispor de um processo para atender cada um desses clientes conforme eles desejem individualmente”.

Portanto, o investimento em TI é fundamental para as empresas que desejam conhecer os seus clientes. SEYBOLD & MARSHAK (2000) diz, a esse respeito:

“Podemos concentrar todos os investimentos efetuados em TI em uma única estratégia vencedora: *facilitar os negócios do cliente com você!* (...) Com esse foco, você pode racionalizar seus investimentos em TI e evitar a duplicação de esforços. (...) No lugar de investir em novos sistemas de recursos humanos, manufatura ou finanças, você precisa investir pelo menos 50% de seu orçamento de TI com o objetivo de facilitar os negócios do cliente com você”. (grifos das autoras)

A tecnologia da informação é, na realidade, um conjunto de ferramentas, sistemas e técnicas que permitem aos estrategistas uma visão mais ampla do mercado sendo abastecido, e mostra as tendências e movimentações dos mesmos.

Haverá posteriormente, neste mesmo capítulo, uma discussão mais aprofundada sobre a Tecnologia da Informação (TI) e suas ferramentas.

### **3.4.2. Personalização dos produtos e das ofertas**

Muitas empresas ainda utilizam a técnica de “segmentar” o mercado. Segmentar significa reunir clientes com necessidades parecidas (mas não

idênticas) em grupos (os segmentos) com o objetivo de garantir que os clientes recebam as mercadorias e/ou serviços a fim de atender, ao menos em parte, suas necessidades. Mas nem sempre os clientes conseguem os produto ou serviços necessários para resolver totalmente um problema.

Até recentemente, as exigências de clientes individuais, embora relevantes para eles, eram amplamente ignoradas, pois a maioria dos compradores simplesmente não podia ser economicamente classificada, rastreada, prevista, administrada e atendida de modo exclusivo. Atualmente, com a diminuição dos custos de TI, as novas tecnologias e o advento do Marketing de Relacionamento, isso é possível (GORDON, 1999).

Hoje, os clientes individuais podem e devem ser considerados em todos os aspectos do negócio, incluindo o processo que conduz a novos esquemas de produtos e serviços. Essa atitude reconhece que os clientes não são iguais.

Assim, a forma histórica de produção - concepção, pesquisa de mercado, desenvolvimento, produção e lançamento - não dará os resultados esperados. Se, no entanto, a empresa envolver seus clientes prioritários, em tempo real, na resolução de seus problemas, os produtos/serviços resultantes serão o resultado de um processo que cria os valores que os clientes querem para cada componente do produto/serviço. Ele não compõe um pacote de vantagens tangíveis e intangíveis que a empresa reúne porque acha que é isso que os clientes querem comprar. Em vez disso, compreende uma agregação de vantagens individuais em cuja escolha ou planejamento os clientes tomaram parte. O “produto” resultante dessa colaboração pode ser exclusivo ou altamente adaptável às exigências do cliente (GORDON, 1999).

Além da personalização dos produtos/serviços, as empresas também podem personalizar suas ofertas. Isto está se tornando comum, principalmente por aquelas empresas que já tem uma experiência adquirida através de sistemas de mala-direta ou telemarketing. Envolve uma boa utilização dos dados já disponíveis, fornecidos pelos clientes de maneira direta (como através de formulários) ou indireta (através da determinação de seu perfil de compras). Pode melhorar significativamente os índices de resposta e de conversão da clientela da empresa, além de reduzir os custos de comunicação. As ofertas feitas a cada cliente são escolhidas dentre muitas opções selecionadas com antecedência. Essas ofertas levam em consideração os perfis dos clientes, sendo apresentada a cada um deles por meio de um recurso de personalização (STONE et al, 2001).

### **3.4.3. Acesso ao cliente**

Um dos principais ganhos permitidos pelo marketing de relacionamentos é o melhor acesso ao cliente. Uma agência de turismo, por exemplo, pode enviar prospectos e folhetos de viagens a seus clientes ofertando pacotes específicos, de acordo com as últimas viagens de cada um. Isso é conseguido pela manutenção de um sistema de apoio (database marketing, data warehouse etc) e ferramentas específicas de pesquisa (como o data mining).

O chamado marketing de massa está sendo, aos poucos, substituído por formas mais direcionadas de acesso aos clientes, ou seja, a transmissão pública ou abrangente (ou “de um para muitos”) dá lugar a uma transmissão pública, porém estreita, atacando segmentos específicos (ou “de um para alguns”). Entretanto, em alguns casos já está se tornando uma comunicação

pessoal, tornando-se uma conversa, se a mesma for em tempo real e interativa (GORDON, 1999).

Note-se que esse último tipo de comunicação pode ser bi-direcional, ou seja, a iniciativa da comunicação pode partir tanto da empresa como do cliente. Os chamados *call-center* (centros de atendimento), *contact-center* (centros de contato) e a internet são as tecnologias de ponta mais utilizadas nestes casos.

Os *call-centers* surgiram da necessidade de fazer algum tipo de abordagem aos clientes por meio telefônico, dando início à era do telemarketing. Esta modalidade tanto pode ser passiva ou ativa.

O telemarketing passivo é aquele em que uma estrutura de atendimento é montada para a recepção de pedidos, reclamações, sugestões etc. No Brasil tornaram-se comuns a partir da entrada em vigor do Código de Defesa do Consumidor, que passou a exigir das empresas uma estrutura de atendimento aos clientes. Normalmente são acessados por números de acesso gratuito (0800) ou de tarifa reduzida (0300), de abrangência nacional.

Já o telemarketing ativo é aquele em que a estrutura é montada como apoio às estratégias mercadológicas da empresa. A iniciativa da ligação parte da empresa, normalmente direcionada a clientes (ou potenciais clientes) selecionados, escolhidos por uma série de fatores. Na sua forma mais rústica, os contatos são escolhidos por segmento de mercado. Assim, para exemplificar, uma construtora, ao oferecer salas comerciais em construção, pode executar uma operação de telemarketing ativo para todos os médicos e dentistas da região, a fim de que montem seus consultórios no empreendimento.

Após de ter iniciado apenas com a utilização da telefonia como meio de contato (quando eram chamados apenas de *call-center*), o setor de comunicação com o cliente melhorou e passou a fazer uso de outras mídias, como a internet, para ter acesso ao cliente. Com essa evolução, este segmento passou a denominar-se *contact-center* (centro de contato), ou seja, passou a ser o setor da empresa que centraliza todas as formas de contato com o cliente.

Outra forma de acesso muito utilizada hoje em dia é o correio eletrônico, ou e-mail. Novamente, tanto pode ser receptivo quanto ativo. Um supermercado, por exemplo, pode deixar a disposição de seus clientes vários endereços eletrônicos de contato, inclusive de seu *Ombudsman* (ou Ouvidor). Por outro lado, utiliza este mesmo canal de comunicação para enviar, aos seus clientes, boletins com ofertas e outras informações.

Em ambas as formas de comunicação ativa, é preciso destacar que o seu uso deve ser feito com muito cuidado. Poucas pessoas toleram serem abordadas em casa, altas horas da noite, por uma ligação oferecendo um *test-drive* em um novo modelo de carro. Da mesma forma, a comunicação por e-mail deve seguir padrões éticos mínimos. Os usuários de internet não gostam, de forma alguma, de receber uma enxurrada de ofertas em suas caixas postais eletrônicas. A forma mais ética seria a de apenas enviar mensagens para aqueles clientes (ou potenciais) que concordassem em utilizar este meio como forma de recepção de informações. Ainda assim, com a possibilidade de interromper o envio no momento em que assim desejar. Atualmente existem ofertas de bancos de endereços eletrônicos, alguns com milhões de endereços,

por preços muito baixos. Uma empresa séria deve evitar a todo custo o uso deste tipo de informação.

Outro meio de comunicação que pode ser muito eficaz, é o velho e conhecido sistema de mala direta. Neste sistema, a empresa envia, através dos correios, folhetos, prospectos e cartas comerciais, para clientes ou potenciais compradores. Em sua forma mais massificada, enviar a mesma correspondência padronizada para um grande segmento social, trás um retorno muito baixo, da ordem de apenas 3%.

Em todas as formas de acesso, a tecnologia permite que a comunicação seja muito mais eficaz. Utilizando as técnicas ou produtos descritos no item de conhecimento do cliente, a empresa pode criar ofertas personalizadas, baseada nas características de cada pessoa que receberá o contato.

Um exemplo pode ser dado com uma rede de supermercados. Ao analisar os dados de compra de seus clientes, é determinado o perfil de cada um deles. A partir daí, pode-se tomar iniciativas de marketing direcionado para cada cliente. Uma iniciativa pode ser a realização de eventos com palestras de enólogos e degustações de vinhos nacionais e importados. Para estes eventos, serão emitidos convites para aqueles clientes que já são consumidores da bebida ou que possuem um perfil de compra que os habilita a adquirir bebidas mais refinadas. O retorno advindo destes eventos irá superar, em muito, os gastos com os mesmos.

Ainda, é muito importante enfatizar que o cliente deve ser acessado pela forma que melhor lhe convier. Alguns irão preferir o e-mail, enquanto que outros não abrirão mão de serem abordados por pessoas, mesmo que seja do

outro lado de um telefone. Além disso, é muito importante que todos os contatos do cliente com a empresa, independente de quem foi a iniciativa, devem ser registrados e disponibilizados para todos os setores que porventura possam ter contato com o mesmo. É extremamente irritante para o cliente ter que explicar várias vezes, para pessoas diferentes, em departamentos diferentes, o mesmo problema. Um banco de chamadas, centralizado e acessível por todos os níveis da empresa, resolve este dilema.

GORDON (1999) diz a respeito do assunto:

“Na era do marketing de relacionamento, a concorrência brigará primeiro pelo acesso ao cliente. Aqueles que o conseguirem se beneficiarão de um maior alcance. Os que não tiverem êxito, se tornarão fornecedores de outros ou desaparecerão. As empresas que hoje identificam os melhores clientes e perseguem seus negócios com paixão serão mais capazes de ganhar acesso e colocar em vigor os processos para se tornarem ainda mais importantes.”

#### **3.4.4. Atendimento**

Atender bem não significa apenas dizer “bom dia” ou “boa noite”. Não é apenas dar um sorriso para o cliente ao final da transação. Isto não é diferencial. É obrigação.

CYмбаUM & MÁLACCO (1999) colocam que:

“... O atendimento agrega valor segundo a percepção do cliente, estimulando sua fidelidade por estabelecer um padrão a que ele se habitua e, até mesmo, criando uma identificação pessoal com o atendente. Melhora a comunicação, fazendo com que o cliente se sinta

mais à vontade para expressar suas opiniões sobre os serviços oferecidos”.

O atendimento, principalmente no ambiente varejista, deve ser primoroso, cordial, envolvente. Diante da falta de diferenciação generalizada, as lojas precisam oferecer experiências de compras ainda melhores, capacitando seus funcionários para atenderem os consumidores de maneira rápida, eficiente e personalizada (MARINHO, 2003). Entretanto, isso não se consegue apenas com algumas palestras motivacionais.

A necessidade de preparar os funcionários da empresa para a nova realidade do marketing de relacionamentos deasfiou vários autores a se empenharem em mostrar que as mudanças na gestão dos recursos humanos são necessidades prementes.

PECK et al (1999) mostra uma visão das áreas de atuação (ou “mercados”) do marketing de relacionamentos. Num total de seis, com o “mercado de clientes” no centro dos outros cinco, um grande destaque é dado ao chamado “mercado interno”, onde é tratado tudo o que é necessário para “vender” a empresa para seus próprios funcionários, ou seja, fazer com que eles assumam verdadeiramente o seu papel como co-autores nas realizações da empresa. Neste aspecto, uma política de treinamento constante e progressivo, acessível a todos os funcionários, é um pré-requisito para alcançar uma excelência no atendimento.

Entretanto, “atender” não significa apenas “vender” algo. Um dos pontos mais delicados dos setores de atendimento é justamente aquele que deve tratar com as reclamações provenientes dos clientes. O correto atendimento

das demandas dos clientes é um poderoso instrumento para o marketing de relacionamento.

Uma forma de se conseguir isso é através da figura do *Ombudsman* (GIANGRANDE 1999). É importante ressaltar que o *Ombudsman* deve ter autonomia dentro da organização, com acesso direto à cúpula dirigente, com o objetivo de sanar rapidamente as reclamações dos clientes.

Outra alternativa é a criação de comitês de clientes (HOLOVKO, 1999), que se define como uma reunião entre clientes e funcionários graduados da empresa com o objetivo de discutir e analisar as reclamações, sugestões e/ou elogios recebidos pela empresa. Além disso, o comitê de clientes pode, de uma forma pró-ativa, se antecipar e detectar possíveis focos de descontentamento dentro da organização.

A correta abordagem estatística das reclamações pode detectar problemas sistêmicos dentro da empresa, ou seja, aqueles problemas que ações gerenciais corretivas eliminam facilmente. DESOUZA (1992) mostra que o correto tratamento das reclamações é um passo importante dentro de uma estratégia de retenção de clientes.

#### **3.4.5. A retenção dos clientes**

A necessidade da manutenção, ou retenção, da carteira de clientes de uma empresa é muitas vezes menosprezada no momento da definição do planejamento estratégico (DESOUZA, 1992; BUCHAN & GILLIES, 1990).

Como parte de uma estratégia competitiva, a retenção de clientes é vital para as empresas atuais, até por uma questão econômica. Uma empresa gasta entre cinco e seis vezes mais para conquistar um cliente do que para manter

um já conquistado (ROSEMBERG & CZEPIEL, 1984; BUCHANAN & GILLIES, 1990; STONE et al 2001).

Além disso, estudos demonstram que um pequeno incremento na taxa de retenção pode proporcionar um impacto expressivo na lucratividade das empresas. Um estudo realizado pela empresa de consultoria Bain & Co. mostra que um incremento de apenas 5% na taxa de retenção irá impactar num aumento entre 25% e 125% na lucratividade, dependendo do tipo de mercado que a empresa atua (REICHHELD & SASSER, 1990; DESOUZA, 1992; ROSEMBERG & CZEPIEL, 1984; BUCHANAN & GILLIES, 1990; STONE et al 2001; PECK et al, 1999; REICHHELD, 1994).

Quanto mais a retenção é eficiente, menores os custos de marketing (DESOUZA, 1992; BUCHANAN & GILLIES, 1990; MAISTER, 1989; PECK et al, 1999), além do aumento da lucratividade. Sob este aspecto, BUCHANAN & GILLIES (1990) e PECK et al (1999) enumeram seis razões pelas quais um cliente retido é mais lucrativo:

- Clientes regulares fazem pedidos freqüentes, consistentes e com pouca variação, acarretando um menor custo de atendimento;
- Clientes retidos tendem a comprar mais;
- Clientes satisfeitos podem, ocasionalmente, pegar preços mais elevados;
- A retenção dos clientes torna mais difícil para a concorrência penetrar no mesmo mercado ou aumentar sua penetração no mesmo;
- Clientes satisfeitos geralmente indicam novos clientes praticamente sem custos (DESOUZA, 1992; PECK et al, 1999);

- O custo de aquisição de novos clientes pode ser substancial. Uma alta taxa de retenção implica que poucos clientes necessitam ser conquistados, a um custo mais baixo.

Com relação ao penúltimo item, DESOUZA (1992) cita como exemplo a aquisição de um complexo sistema de telecomunicações. Por ser um investimento de alto risco, a tendência do comprador será confiar mais em colegas ou amigos do que simplesmente nas informações prestadas pelo fornecedor ou por sua propaganda.

De um outro ponto de vista, a perda de um cliente pode trazer prejuízos substanciais, pois o custo da perda será o valor que aquela pessoa poderia desembolsar enquanto estivesse envolvida em um relacionamento de negócios com a empresa durante um ciclo de vida.

Para se conseguir a retenção dos clientes, alguns passos básicos são necessários.

Primeiro é necessário criar mecanismos de medição da taxa de retenção da empresa, pois parte-se do pressuposto que “o que não é medido não pode ser gerenciado” (DESOUZA, 1992; REICHHELD, 1994). Além da simples contagem de “mantidos” x “perdidos”, deve-se tomar o cuidado de se atribuir pesos diferenciados para cada tipo de cliente (o ideal seria um peso específico para cada cliente), a fim de se determinar corretamente, qual a taxa de retenção da empresa.

Outra ação importante é ouvir o relato dos antigos clientes, pois para entender porque os clientes deixam de negociar com sua empresa, devemos sempre lhes perguntar, seja diretamente ou através de serviços especializados.

A informação assim conseguida tenderá a ser mais acurada do que aquela conseguida por pesquisas de opinião e/ou satisfação. Ao se fazer este tipo de questionamento, as empresa podem fazer descobertas interessantes. THOMAS (2001) diz que, no ramo de seguros, a maioria dos clientes cita o preço como fator de desistência. Entretanto, ao fazer uma entrevista com 300 ex-clientes de uma determinada empresa, descobriu que apenas 20 haviam realmente desistido do negócio baseado apenas no preço. O restante já havia passado por outras experiências desagradáveis, sendo que o preço havia sido apenas a chamada “gota d’água”.

O correto gerenciamento das reclamações dos clientes também auxilia na retenção. É natural considerar as reclamações como enfadonhas e irritantes. Entretanto, os dados oriundos das reclamações podem ser uma mina de ouro para o analista que desejar identificar problemas que causam as defecções dos consumidores (DESOUZA, 1992; THOMAS, 2001). As pesquisas mostram, no entanto, que a grande maioria dos clientes com problemas não reclama. THOMAS (2001) diz que de cada 100 (cem) clientes, apenas 04 (quatro) reclamam. Ouvir e atuar sobre essas queixas pode ajudar a manter não apenas aqueles clientes que reclamam, mas principalmente aqueles não o fazem. THOMAS (2001) ainda diz que devemos considerar cada cliente que reclama como sendo um porta-voz de outros 25.

Existe, é claro, a necessidade de se fazer um uso estatístico dos dados das reclamações. Elementos individuais devem ser assinalados para identificar padrões nos dados, bem como elementos que estão fora dos padrões normais. Sem métodos estatísticos, tentativas de se implantar um processo desse tipo

será em vão. A maioria dos problemas resultam de fatores sistêmicos, como falhas de projeto, supervisão pobre e/ou maquinário fora das especificações. A análise das reclamações pode revelar que um modelo particular ou fabricante aponta para um número desproporcional de queixas. Isso indica que o problema é sistêmico e pode ser eliminado por ações gerenciais (DESOUZA, 1992). Um correto tratamento estatístico deverá ser utilizado para medir as reduções das reclamações baseados em sua *causa*, não ao seu *volume* (JOHNSTON, 2001).

O correto tratamento das reclamações pode fazer com que o cliente, às vezes a um passo da defecção, seja recuperado, fazendo com que o mesmo passe a se sentir satisfeito. JOHNSTON (2001), mostra que a maior parte dos clientes altamente satisfeitos passou por momentos desconfortáveis no relacionamento, mas perceberam que a empresa realmente se esforçou em eliminar os problemas. Um gerenciamento eficiente das reclamações, com sua consequente correção, alavanca sobremaneira a formação da lealdade do cliente.

Por último, a empresa deve se esforçar para identificar as chamadas *barreiras de troca*, ou seja, aqueles outros produtos ou serviços agregados que fazem com que seus clientes relutem em deixar de negociar com a empresa.

Essas barreiras podem ser tangíveis, tais como um serviço de suporte ou orientação ao usuário, ou intangíveis, como a criação de grupos de discussão de usuários de um mesmo produto.

Segundo SEYBOLD & MARSHAK (2000), apesar da noção de fomentar comunidades não ser nova no mundo dos negócios, a internet possibilitou que

peças com os mesmos interesses se reunissem em torno de listas de discussão, mesmo separadas por enormes distâncias. Desta forma, a empresa que possibilitar essa reunião de interesses entre seus clientes estará construindo uma barreira de troca.

Uma outra forma de criar barreiras de troca é o atendimento. Sempre que possível, o atendimento personalizado deve ser buscado, como uma forma de melhor atender o cliente. O cliente, por sua vez, prefere fazer negócios com empresas que conheçam suas necessidades e que buscam resolver os seus problemas. Quando uma empresa consegue este nível de atendimento, a barreira de troca já estará criada, fazendo com que este cliente se torne refratário às tentativas da concorrência em captá-lo.

Às vezes é necessário que as empresas “olhem para fora” de seus nichos de mercado, procurando por práticas de excelência em outros ramos de atividade. Se a empresa conseguir adaptar com sucesso uma dessas experiências alienígenas, estará criando uma barreira surpreendente para a concorrência. Como um exemplo, uma determinada ótica começou a atender ao público infantil em uma de suas lojas. Uma das dificuldades encontradas foi justamente como atender este tipo de público, pois a criança é naturalmente dispersiva. A solução encontrada foi contratar pedagogas para atender este público específico. Ou seja, inseriu um profissional estranho ao meio para o atendimento de um público específico.

Manter os clientes é, sem sombra de dúvidas, muito mais eficaz e econômico do que simplesmente atraí-los para a esfera de uma empresa. O

cliente retido traz mais lucro e custa menos para a empresa. Daí a necessidade de se criar mecanismos de retenção.

### **3.5. CRM e a tecnologia da informação**

Segundo BRETZKE (2000), o CRM auxilia a implementação de técnicas de marketing de relacionamento de diversas maneiras, entre as quais:

- Tornar os esforços de marketing mais eficazes através da identificação dos clientes mais importantes, apresentando a eles a oferta do produto/serviço mais adequado, no momento correto;
- Utilizar as ferramentas da tecnologia da informação (TI) para administrar a grande quantidade de dados de que a empresa necessita para interagir com seus clientes da maneira mais personalizada possível;
- Manter o diálogo com o cliente a partir da identificação das interações do mesmo com a empresa, que podem revelar alterações no comportamento de compras e permitem à empresa antecipar mudanças futuras;
- Facilitar o desenvolvimento de novos produtos, a partir do conhecimento de quem compra um produto, como ele satisfaz o consumidor e quais mudanças enriqueceriam o desempenho do produto.

Como visto anteriormente, o CRM pode ser definido como a junção das técnicas de marketing de relacionamento com as ferramentas colocadas à disposição pela tecnologia da informação (TI).

As principais ferramentas da tecnologia da informação, para o contexto do CRM, são:

- Banco de Dados de Marketing (database marketing);
- Data Warehouse;
- Data Mining;
- Raciocínio Baseado em Casos;

Segue a análise de cada uma dessas ferramentas de TI que permitem à empresa conhecer e interagir com os seus clientes.

### **3.5.1. Bancos de dados de marketing (database marketing)**

Para uma empresa conhecer os seus clientes, é necessário armazenar os dados produzidos por eles e sobre eles. O gerenciamento do relacionamento com os clientes, possivelmente em vários locais diferentes, seria impossível sem a tecnologia da informação (TI), particularmente nas áreas de telecomunicações e gerenciamento de bancos de dados.

No caso específico do seu uso em marketing ele é chamado de “database marketing”. BRETZKE (2000) diz que: “é um termo amplamente usado em marketing, mas costuma ter diferentes conceituações, dependendo de quem o está definindo”. Segundo ela, o National Center for Database Marketing tem empregado a seguinte definição:

“Database marketing é o gerenciamento de um sistema dinâmico de base de dados inteligíveis, atualizados, com informações relevantes sobre clientes atuais e potenciais, para:

- Identificar os clientes atuais e potenciais mais propensos a responder às ações de marketing;

- Desenvolver um relacionamento de alta qualidade de longo prazo com compras repetidas;
- Através do desenvolvimento de modelos preditivos, capacitar o envio das mensagens desejadas no tempo certo, no formato certo e para as pessoas certas;
- Enviar mensagens que encantarão o cliente;
- Construir lealdade;
- Fazer com que as despesas de marketing tenham a melhor relação custo-resultado;
- Aumentar o lucro.” (BRETZKE, 2000)

Segundo STONE & WOODCOCK (1998) e STONE et al (2001), as principais características de um banco de dados de marketing são as seguintes:

- Cada cliente efetivo ou potencial é identificado em um registro do banco de dados;
- Cada registro de cliente contém informações de identificação, necessidades e características, comunicações de campanha, reações do cliente às campanhas realizadas, transações realizadas no passado etc;
- A informação deverá estar disponível durante cada comunicação com o cliente;
- O banco de dados é usado para registrar as reações do cliente às iniciativas da empresa;

- A informação também deverá estar disponível para os estrategistas de marketing, para permitir que se decida sobre que mercados-alvo são adequados para cada produto ou serviço e qual marketing-mix é adequado para cada produto em cada mercado alvo;
- Usar o banco de dados visando assegurar que a abordagem do cliente seja coordenada e consistente. Esse passo é vital no marketing de relacionamento;
- O banco de dados eventualmente substitui a maior parte da pesquisa de mercado. As campanhas de marketing são concebidas para que a reação dos clientes a ela forneça as informações desejadas.

Utilizado de forma adequada, o database marketing pode ajudar as empresas a dirigirem sua atenção para onde estará direcionado o marketing dos próximos anos: cultivando e retendo os clientes atuais e usando o conhecimento e a informação para buscar novos consumidores mais rentáveis (BRETZKE, 2000).

Para que o database marketing seja efetivo como sustentação da estratégia de Marketing de Relacionamento, deve ser totalmente integrado à operação da empresa. A grande vantagem é a atualização constante de um grande número de informações obtidas pelas diversas áreas que possuem relacionamento com os clientes efetivos e/ou potenciais.

### **3.5.2. Data Warehouse**

A tecnologia de Data Warehouse (DW) foi descrita originalmente por INMON (1997), que diz: “Um Data Warehouse é um conjunto de dados *baseado em*

*assuntos, integrado, não volátil e variável em relação ao tempo, de apoio às decisões gerenciais”.*

Desta forma, os bancos de dados de marketing estão sendo estruturados cada vez mais como Data Warehouses, pois utilizam conceitos semelhantes.

Numa primeira vista, os conceitos de database marketing e DW são muito semelhantes. A maior diferença está na forma como os dados são estruturados. O DW é um repositório de dados que contém informações oriundas de diversas fontes de dados (vários departamentos e/ou sistemas diferentes dentro da mesma empresa), armazenadas de forma consolidada, enquanto que os bancos de dados normais podem conter dados redundantes e não consolidados, além de estarem descritos de formas diferenciadas em cada um deles.

É possível ter um mesmo dado, que represente a mesma grandeza, sendo tratado por sistemas diversos e em formatos diferentes. Quando for necessária uma decisão estratégica que contemple um conjunto de dados deste tipo, juntamente com um processo de análise que os compare, releve, analise tendências e projeções ou outro processo analítico deste tipo, teremos problemas em unir estas versões diversas de uma mesma grandeza (JAMIL, 2001). O DW resolve este tipo de problema através da consolidação destes dados.

GORDON (1999) diz que “os DW dos clientes podem ser definidos como enormes depósitos de informações sobre os clientes, a partir de fontes internas da empresa e de fontes do cliente ou de terceiros, como o governo, empresas de cartão de crédito e de pesquisa de mercado. Os dados podem incluir

comportamentos, preferências, informações sobre o estilo de vida, dados transacionais e dados sobre comunicações com a empresa antes, durante e depois da venda, além de incluir informações sobre a lucratividade, satisfação, retenção, fidelidade e referências do cliente”.

A maior vantagem desse recurso é permitir que as empresas elaborem sistemas de marketing e de vendas que possam ser otimizados para lidar com transações e com a estrutura organizacional específica da empresa. Em conjunto com sistemas de exploração de dados (data mining - próximo tópico), são elementos de apoio ao desenvolvimento de estratégias e operações mais competitivas (STONE et al, 2001).

### **3.5.3. Data Mining**

É a aplicação de técnicas estatísticas e de inteligência artificial em grandes quantidades de dados, para descobrir relações e padrões entre os dados. Também pode ser dito que é o processo de construir modelos baseados em um conjunto de dados, que permita fazer previsões, controlar ou melhorar algum processo (MORALES et al, 2001).

O seu objetivo é descobrir informações e conhecimentos, através do reconhecimento de novos fatos e relações existentes em uma base de dados. A idéia chave é tornar disponível para o gerenciamento as informações críticas que podem ser usadas para futuros processamentos analíticos e tomadas de decisões. Tecnologias de mineração de dados são caracterizadas pela intensiva computação em grandes volumes de dados.

Os processos de data mining podem ser utilizados para:

- **Previsão:** envolve a utilização de algumas variáveis (atributos da base de dados) para prever valores desconhecidos ou futuros de outras variáveis de interesse;
- **Descrição:** procura por padrões que descrevam os dados e que sejam interpretáveis por seres humanos.

As ferramentas de data mining podem ainda atuar como um dos dois modelos a seguir:

- **Modelo de Verificação:** O modelo de verificação toma uma hipótese do usuário e testa a sua validade contra os dados. A ênfase está com o usuário, que é responsável pela formulação das hipóteses e realização das consultas nos dados para afirmar ou negar a hipótese. Uma aplicação prática é a seleção de clientes para mala direta.

Exemplificando melhor, em uma divisão de marketing com orçamento limitado para uma campanha de mala direta com vistas ao lançamento de um novo produto, é importante identificar o segmento da população que provavelmente compraria o produto. O usuário formula uma hipótese para identificar clientes potenciais e as características que estes dividem. Dados históricos sobre compras e informações demográficas podem ser consultadas para revelar compras comparáveis e as características divididas pelas mesmas, que então podem ser usados como alvo para uma campanha de mala direta. Todas as operações podem ser refinadas de modo que as hipóteses reduzam o conjunto retornado cada vez, até que um limite determinado seja alcançado.

- **Modelo de Descoberta:** O modelo de descoberta difere na sua ênfase, considerando que são descobertas automaticamente informações importantes escondidas nos dados. Os dados são vasculhados na procura de padrões freqüentes, tendências e generalizações, sem intervenção ou ajuda do usuário. As ferramentas de descoberta ou mineração de dados tentam revelar um grande número de fatos sobre os mesmos no menor tempo possível.

Um exemplo deste modelo é o repositório de dados de um banco, que é analisado para a descoberta de vários grupos de clientes para uma campanha de mala direta. Os dados são procurados sem nenhuma hipótese em mente a não ser que o sistema deve agrupar os clientes de acordo com as características comuns encontradas.

Consultas tradicionais em bancos de dados contrastam com mineração de dados, simplesmente porque estão limitados a questões simples tais como "quais foram as vendas de suco de laranja em janeiro de 1995 em São Paulo?". A análise multidimensional, geralmente chamada de On-Line Analytical Processing (OLAP), habilita os usuários a fazer consultas muito mais complexas, tais como a comparação das vendas programadas com as vendas reais em uma determinada região nos anos de interesse. A mineração de dados, por outro lado, através do uso de algoritmos específicos ou mecanismos de busca, tenta descobrir padrões discerníveis e tendências nos dados e inferir regras para os mesmos. Com estas regras ou funções, o usuário está habilitado a suportar, revisar e examinar decisões em alguma área comercial ou científica relacionada (MORALES et al, 2001). Desta forma, a mineração de

dados é uma ferramenta extremamente poderosa para ser utilizada no marketing de relacionamento.

#### **3.5.4. Raciocínio Baseado em Casos**

Raciocínio Baseado em Casos (RBC) é uma técnica de IA que se baseia na premissa de que seres humanos utilizam um raciocínio analógico ou experimental para aprender e resolver problemas complexos. Sistemas RBC utilizam a filosofia de aprender através de situações que aconteceram no passado (aprendizado por experiência). A literatura apresenta dois pressupostos para o emprego do raciocínio baseado em analogia: problemas similares têm soluções similares e os tipos de problemas se repetem (COSTA, 1999).

Segundo WATSON (apud COSTA, 1999), as principais características do RBC são:

- Sistemas RBC não requerem um modelo explícito do domínio. A recuperação do conhecimento é feita através da recuperação de casos armazenados;
- As dificuldades de implementação são reduzidas. A tarefa é encontrar os aspectos relevantes que descrevem um caso, ao invés da criação de um modelo explícito;
- Sistemas RBC podem ser combinados com técnicas de banco de dados para gerenciar grandes volumes de informação;
- A manutenção do conhecimento é simplificada pela capacidade que os sistemas RBC possuem de aprender novas informações na forma de casos.

RBC é uma ferramenta bastante atual, cujas principais aplicações são os sistemas de diagnósticos médicos e de *help-desk* (atendimento a clientes), por sua capacidade de manter um histórico dos casos anteriormente atendidos, aprendendo com os mesmos e propondo soluções baseadas na similaridade destes com um novo problema apresentado.

Em ambientes de marketing, pode ser utilizado em sistemas de avaliação (utilizados para determinar valores para variáveis, comparando-as com o valor conhecido de um caso similar) ou de apoio à decisão (pois trabalham com problemas complexos oferecendo suporte para a recuperação de informações relevantes em problemas similares) (COSTA, 1999). Além disso, pode servir para atender um problema específico de um cliente, baseado na forma de atendimento (e soluções propostas) de problemas semelhantes de outros clientes. O site [amazon.com](http://amazon.com), a grande livraria virtual, é um caso de utilização da técnica de RBC. Quando se começa a fazer negócios com a livraria, começam a acontecer as surpresas. Sempre que o cliente volta ao site, é recebido com boas vindas pelo sistema, que mostra o seu nome e sugere vários novos livros semelhantes àqueles já adquiridos, ou até de outras áreas, mas que abrangem assuntos correlatos aos já adquiridos. Por trás desse serviço aparentemente simples existe toda uma base de conhecimento adquirido. Este sistema, todo fundamentado nesta técnica, é quem dá as dicas de novos livros, baseado nas suas próprias compras e nas compras realizadas por pessoas com interesses semelhantes.

### 3.6. Conclusões iniciais

Competição é a palavra chave no mundo dos negócios atualmente. As empresas atuam em ambientes cada vez mais competitivos, onde os resultados devem ser alcançados com estratégias corretas. Erros de estratégia de atuação custam, às vezes, a sobrevivência da empresa.

É neste cenário competitivo que o marketing de relacionamentos e o CRM vem auxiliar as empresas a conquistar e, muito mais importante, *reter* os clientes conquistados. As estratégias e ferramentas, fornecidas pelo CRM e quando bem utilizadas, permitem às empresas conhecer melhor os seus clientes, auxiliando na oferta de produtos e serviços com alto valor para os consumidores.

Conhecer o cliente, desde seu perfil de consumo, passando por suas formas preferidas de acesso, até a melhor forma de ofertar os produtos, garante à empresa um alto valor estratégico e competitivo. Os clientes, quando alcançam os mais altos degraus da escada de relacionamento, passam a atuar como verdadeiros advogados da empresa, tornando-se altamente refratários às investidas da concorrência (GORDON, 1999; PECK et al, 1999). E a empresa que alcança este degrau de lealdade nos seus clientes tornará muito difícil a vida de seus concorrentes. A correta análise e aplicação dos pressupostos aqui colocados irão possibilitar à empresa, com toda a certeza, uma grande vantagem competitiva.

No próximo capítulo será mostrado como se pode aplicar os pressupostos acima descritos em um ambiente altamente competitivo: o setor supermercadista. Além disso, será mostrado que a introdução destas técnicas

poderão ser realizadas com um custo extremamente baixo em comparação com a implantação de um projeto de CRM levado a efeito por consultorias especializadas.

## 4. A IMPLANTAÇÃO DO CRM

Como visto no capítulo anterior, a fidelização dos clientes hoje é uma necessidade, muito mais do que um mero modismo. Entretanto, quando se fala em fidelização de clientes em um ambiente de varejo, principalmente o varejo de massa, vê-se que as dificuldades em se alcançar este objetivo são muito maiores do que para uma empresa de serviços, por exemplo.

Uma simples análise da quantidade de clientes atendidos e do volume de transações diárias efetuadas por uma rede de supermercados de porte médio mostrará que, sem uma infraestrutura de tecnologia de informação (TI), é praticamente impossível a implantação de políticas de marketing de relacionamento por estas empresas.

Entretanto os custos para se implantar um sistema de fidelização ainda são muito elevados. Os gastos com consultoria, equipamentos e sistemas de CRM por vezes inviabilizam a adoção destas ferramentas por parte de redes de supermercados de porte médio, ainda mais se considerarmos que o setor trabalha com margens líquidas de lucro entre 1% e 2%, e é deste lucro que sai a maior parte dos investimentos em novas tecnologias e serviços.

O objetivo primário deste estudo é justamente mostrar que se consegue implantar um conjunto de ferramentas e ações de marketing com um baixo custo, aproveitando toda a infraestrutura de TI e as estruturas de dados existentes na empresa. Além disso, a utilização de sistemas abertos permite uma redução substancial nos custos de implantação.

Hoje é perfeitamente possível se implementar um *data warehouse* baseado em ferramentas *freeware*. Como um exemplo, um sistema operacional como o

Linux ou FreeBSD, em conjunto com o SGBD MySQL já permite a implantação de um DW com custo zero em licenças de software. Para os estudos de *data mining*, existem ferramentas de baixo ou nenhum custo, dentre as quais destaca-se o pacote WEKA, desenvolvido pela Universidade de Waikato, na Nova Zelândia.

#### **4.1. Utilização eficiente dos dados de venda**

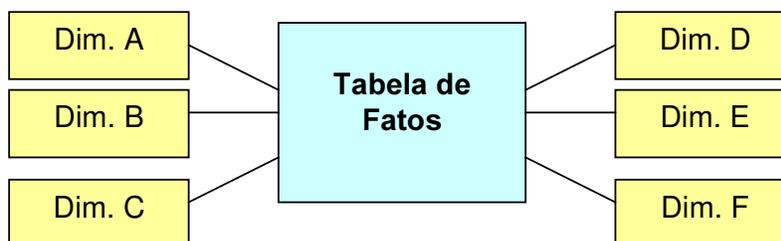
Por força do convênio ICMS 85, de 28/09/2001, todos os estabelecimentos varejistas com faturamento anual acima de R\$ 120.000,00 são obrigados a registrar todas as transações de venda através dos chamados Emissores de Cupom Fiscal (ECF), impressoras inteligentes que registram, a cada item vendido, quanto deve ser pago de imposto (ICMS ou ISS, dependendo do caso) pelo estabelecimento.

Os ECFs operam acoplados a computadores, sob as ordens de programas específicos, que enviam as transações para os ECFs e, ao mesmo tempo, gravam essas informações (além de muitas outras) em arquivos ou bancos de dados. As informações gravadas são utilizadas basicamente para controles de estoque (informações de venda) e financeiros (informações sobre os pagamentos efetuados).

Como se vê, existe uma grande quantidade de informações gravadas nestes equipamentos que podem ser melhor aproveitadas. Com algumas pequenas alterações nos programas existentes, é possível obter a identificação do cliente que está realizando uma determinada transação. A partir de sua identificação, é possível cruzar as informações do cliente com as informações a respeito das compras efetuadas pelo mesmo.

Para realizar este cruzamento e transformar estas informações em conhecimento, a TI trabalha com os chamados *data warehouses* (DW), ou armazéns de dados. O DW é uma forma de modelar informações em um SGBD relacional de maneira a facilitar este cruzamento de informações. Dentre os vários modelos de DW, um dos mais comuns é o modelo *dimensional*, em que os dados são distribuídos entre dimensões e as chamadas *tabelas de fatos*. Em um DW dimensional típico, temos várias dimensões ligadas a tabelas de fatos. Este arranjo é denominado de *esquema estrela*, pois as tabelas de dimensões são representadas em volta das tabelas de fatos, como as pontas de uma estrela, conforme a figura 4.1. Uma das principais características dos DW é o fato de que ele tem uma forte característica temporal, ou seja, ele armazena os fatos em função do tempo.

**Figura 4.1. Exemplo de modelo dimensional**



Os DW são compostos normalmente de vários *data marts* (DM), ou seja, pequenos "DW departamentais". Cada DM é composto de uma tabela de fatos e suas respectivas tabelas dimensionais.

Com as informações vindas dos programas de ECF, é possível montar um pequeno (mas poderoso) DM de vendas, cujas dimensões principais seriam as de tempo (data e hora), de clientes, de produtos e de lojas (considerando uma empresa com várias filiais). Na tabela de fatos, seriam armazenadas as vendas

em si, com a quantidade e os valores de custo e venda de cada item. Com este DM é possível responder a várias perguntas, tais como:

- Quais produtos um determinado cliente x adquiriu? Quando? Em que loja? Qual o lucro gerado?
- Os clientes preferem a loja próxima à sua residência? Ou preferem outras, mas bem sortidas?
- Em que loja os clientes geram mais lucro? E prejuízos?
- Em que períodos os clientes adquirem produtos com maior valor agregado?

Dependendo da granularidade dos dados armazenados na tabela de fatos, é possível que o espaço ocupado por este arranjo seja muito grande, demandando recursos especiais de software e hardware para comportar a massa de dados.

A partir do momento em que os dados estão armazenados no DM, é possível utilizá-los para alimentar sistemas de mineração de dados. Estes sistemas vasculham grandes quantidades de informações a procura de padrões que estejam ocultos, ou pelo menos não perceptíveis numa primeira análise.

As ferramentas de *data mining* podem ser utilizadas, por exemplo, para selecionar clientes para uma determinada campanha de marketing a partir do seu perfil de compras. Pode, ainda, agrupá-los tendo por base seu perfil. Ou então determinar a existência de padrões de compra não visíveis a olho nu. Uma outra aplicação destas técnicas é a chamada “Análise de Cesta de Compras” (*MBA – Market Basket Analysis*), ou seja, a análise das compras do

cliente para comparação com as compras de clientes com perfil semelhante, para que se possa oferecer produtos que o mesmo normalmente não utiliza mas que possa a vir utilizar. Um dos grandes usuários de sistemas de *data mining* neste setor é a rede norte-americana Wal-Mart, que possui uma base de dados de mais de 100 Terabytes (1 Tb =  $2^{40}$  bytes) (SWIFT, 2001).

A partir dessas ferramentas é possível criar, então, um mecanismo de ações de marketing dirigido, fazendo com que os clientes recebam ofertas personalizadas. Sabe-se que este tipo de campanha é muito mais eficaz do que as campanhas de marketing. Uma mala direta ofertando um produto novo, dirigido à um pequeno número de clientes com perfil compatível é muito mais eficiente do que uma campanha dirigida a todos os clientes indistintamente. Enquanto uma campanha baseada em mala direta tradicional consegue entre 1 e 3% de resposta positiva uma campanha dirigida, focando os clientes certos, pode trazer respostas positivas bem acima de 60%.

Eis alguns exemplos de ações que poderiam ser criadas:

- A empresa planeja, no início do inverno, realizar uma feira especial para a comercialização de vinhos. Pode-se fazer uma mala direta para os clientes que são habituais consumidores de vinho convidando para o evento;
- Em parceria com um fornecedor de produtos para higiene de animais domésticos, promover um evento para demonstração do novo produto. Os convidados seriam os clientes que habitualmente adquirem rações para seus animais de estimação;

- Enviar um cartão de desconto para aquisição de iogurtes para os clientes que consomem leite, mas não outros laticínios;
- Utilizar os dados de compras anteriores, em conjunto com o padrão de compras de clientes com perfil semelhante, para sugerir novos produtos ao cliente.

Note-se que essas ações podem facilmente ser implementadas a partir do DM proposto acima.

#### **4.2. A identificação do cliente**

Um dos principais problemas desta proposta está justamente na identificação do cliente. Normalmente utiliza-se a implantação de programas de relacionamento, que incluem a necessidade de se identificar o cliente no momento da compra mediante a apresentação de um cartão. Em troca, são oferecidas vantagens específicas para o mesmo, tais como descontos, troca de pontos por produtos ou serviços etc.

A utilização destes programas desde o início de uma implementação de técnicas de CRM pode ser muito oneroso, pois envolve a implantação de uma grande infraestrutura apenas para administrar e operacionalizar este método de identificação. A questão que se coloca é a validade de um gasto tão alto apenas para se iniciar um programa deste tipo, sem uma perspectiva de retorno de curto prazo.

Em um supermercado comum, o cliente somente é identificado quando realiza o pagamento em cheque. Neste momento, é comum a obtenção de dados como CPF e RG, principalmente quando o estabelecimento realiza

consultas aos serviços de proteção ao crédito. Nas outras formas de pagamento, não existe a identificação.

Para resolver este problema, utiliza-se dos chamados *cartões de relacionamento*. Com o uso deste cartão, o cliente acumula pontos ou bonificações que podem ser trocados por produtos ou serviços. Este mecanismo de troca é necessário para estimular o cliente a se identificar a cada compra.

Entretanto, a simples implantação de um programa de recompensas não pode ser considerado um programa de fidelização, pois ele, apesar de ser necessário, permite única e exclusivamente a identificação do cliente. Um programa mais completo de fidelização utiliza as informações geradas por estes cartões de relacionamento para a implantação de uma política de marketing direcionada, o que acaba criando as chamadas “barreiras de troca”, ou seja, faz com que o cliente tenha uma grande resistência à mudança para um concorrente.

A implantação de um cartão de relacionamentos pode ser realizada de duas formas: criar uma campanha institucional, conclamando os clientes a se cadastrar no programa em troca de recompensas; ou partir de um cadastro inicial de clientes previamente cadastrados.

É muito comum as empresas manterem cadastros de clientes para controle de créditos, seja de prestações ou para controle de cheques pré-datados. Se a empresa já possui um cadastro desse tipo, é natural que o mesmo seja aproveitado para iniciar o programa. Além disso, essa abordagem permite começar com um projeto piloto, selecionando clientes já com um

relacionamento estabelecido e praticar, com esses clientes, algumas das técnicas de reconhecimento de perfil e ações de marketing como as propostas anteriormente.

Em qualquer dos dois casos, não é possível convencer os clientes a se identificarem sem oferecer algum tipo de recompensa em troca. Como já falado anteriormente, é um passo necessário e, de certa maneira, justo. É o que o varejista oferece em troca da identificação da compra.

Em termos de privacidade, é interessante notar que a imensa maioria dos clientes não tem noção de como expõe seus hábitos e suas preferências quando utiliza um cartão de relacionamento, principalmente aqueles que oferecem vantagens em vários estabelecimentos comerciais distintos, tais como o cartão Smiles, da Varig, ou o Smart Club, de uma associação de empresas, incluindo gigantes como as Lojas Americanas. A partir do cruzamento das informações coletadas em vários estabelecimentos, é possível determinar, com alto grau de exatidão, o comportamento de consumo de um determinado cliente.

Isso, entretanto, não é exclusivo dos cartões de relacionamento. Os bancos e administradoras de cartões de crédito também conseguem fazer este tipo de cruzamento de informações (pois sabem exatamente quanto cada cliente gastou, quando e em quais estabelecimentos). No entanto, essas instituições não têm por objetivo fazer com que o cliente troque um fornecedor de produtos ou serviços por outro concorrente.

### 4.3. Um modelo genérico de implantação de CRM

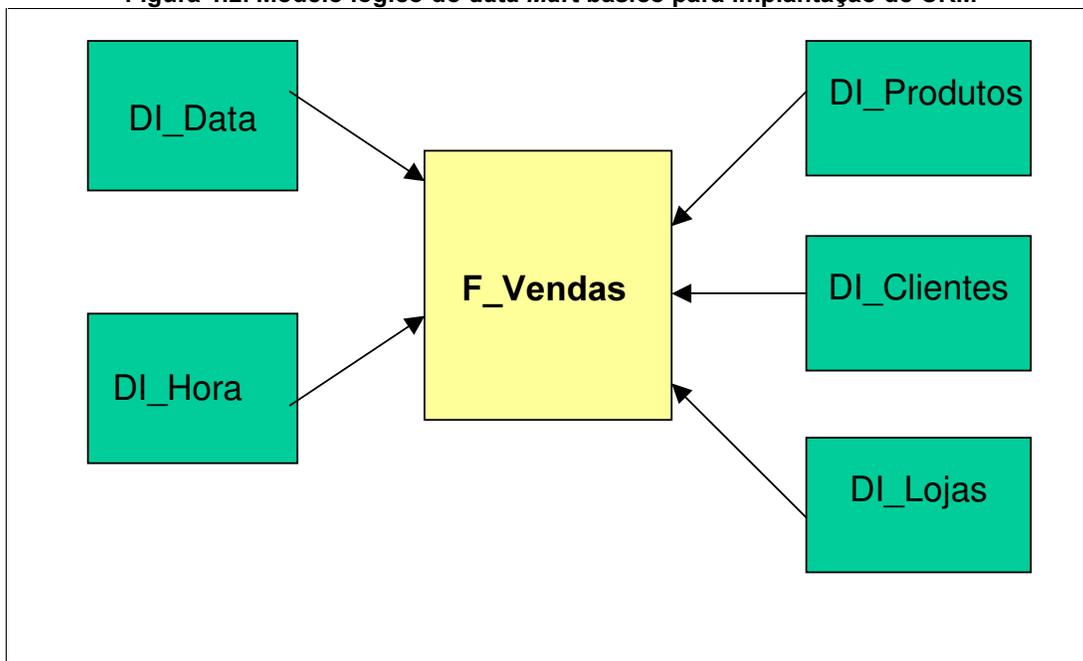
Para iniciar os procedimentos de implantação de uma abordagem de CRM (considerando que as definições de foco da empresa já estão resolvidas), é proposta a seguinte metodologia:

- a) Criar, a partir dos cadastros de produto e de clientes (cheques), acrescidos das informações de venda dos caixas (LOGs), um DM contendo as informações referentes às transações com cada um desses clientes;
- b) A partir deste DM selecionar os clientes através de algum critério prático. Pode-se selecionar, por exemplo, os clientes cadastrados com um número de transações acima da média (frequência, compras ou lucratividade) e utilizá-los como *database marketing* de testes;
- c) Planejar e executar ações de marketing visando a associação dos clientes selecionados ao programa;
- d) Implementar uma política de ações de marketing tendo como público-alvo os clientes associados;
- e) Mensurar o retorno das ações-piloto, fazendo os ajustes necessários no programa;
- f) Definir a forma de recompensa que será adotada pela empresa, baseada em algum parâmetro mensurável (frequência, valor das compras etc). Definir, também, como estes pontos serão convertidos em recompensas (troca por produtos e/ou serviços, qual a relação pontos X custo que seja sustentável etc);
- g) Finalmente, expandir o programa para o público em geral;

Como visto, o primeiro passo é a definição do DM que irá armazenar as informações referentes às vendas, com o cruzamento de produtos, lojas, clientes e o componente temporal (data e hora).

O modelo lógico deste DM é o mostrado na figura 4.2, a seguir:

**Figura 4.2. Modelo lógico do *data Mart* básico para implantação de CRM**



Do lado direito temos três dimensões representando os dados corporativos: Produtos, Clientes e Lojas. Já do lado esquerdo temos as dimensões temporais, divididas em duas: Data e Hora. Ao centro, a tabela de fatos de venda, que irá efetivamente fazer o cruzamento de todas as informações que são necessárias para o desenvolvimento de nossas atividades de marketing. Os quadros 4.1 a 4.6 mostram os atributos de cada uma dessas entidades:

**Quadro 4.1. Atributos da entidade DI\_Dia**

Entidade	Atributos
DI_Data	Ch_Data (*) Dia Mês Ano Dia_Semana Data Estação Temporada Feriado Vespera_Feriado

**Quadro 4.2. Atributos da entidade DI\_Hora**

Entidade	Atributos
DI_Hora	Ch_Hora (*) Hora Período

**Quadro 4.3. Atributos da entidade DI\_Podutos**

Entidade	Atributos
DI_Produtos	Ch_Produto (*) CodProduto Produto Secao Depto Grupo

**Quadro 4.4. Atributos da entidade DI\_Lojas**

Entidade	Atributos
DI_Lojas	Ch_Loja (*) Num_Loja Endereço Bairro Cidade Estado Gerente Regional Area NumCaixas NumFuncionarios

**Quadro 4.5. Atributos da entidade DI\_Clientes**

Entidade	Atributos
DI_Clientes	Ch_Cliente (*) CodCliente Nome Endereço Bairro Cidade Estado FaixaRenda

**Quadro 4.6. Atributos da entidade F\_Vendas**

Entidade	Atributos
F_Vendas	Ch_Dia (*) Ch_Hora (*) Ch_Produto (*) Ch_Cliente (*) Ch_Loja (*) NumEcf (*) NumCF (*) Qtd ValorVenda Lucratividade

Os atributos marcados com (\*) fazem parte das chaves primárias de cada uma das entidades.

Note-se que os atributos das tabelas de dimensão possuem vários campos que permitem uma série de consultas detalhadas, de forma clara e objetiva. Note-se também que são tabelas desnormalizadas, ou seja, não existem tabelas relacionadas a elas que não seja a tabela de fatos. Para tanto, os códigos (como os de cidade ou bairro) foram substituídos por seus valores reais.

A seguir, uma breve explanação de cada uma das entidades.

#### **4.3.1. DI\_Data**

Dimensão temporal, contém informações sobre as diversas datas que podem ser utilizadas pelo DM. Além da data real (data), temos a mesma decomposta em suas unidades básicas. Este tipo de decomposição é recomendável para que se possa fazer análises entre períodos distintos de vendas. Os atributos de estação, temporada, feriado e véspera servem ao mesmo propósito. O atributo temporada pode conter valores indicativos da época (verão, carnaval, quaresma/páscoa, dia dos namorados, dia das mães, dia dos pais, festas juninas, férias, dia das crianças, natal, volta às aulas etc).

#### **4.3.2. DI\_Hora**

Segunda dimensão temporal, contém informações sobre as diversas horas do dia e a que período ela pertence (madrugada, manhã, tarde e noite). Novamente, os atributos foram projetados tendo em vista a facilidade de tratamento em consultas por período ou cubos OLAP.

#### **4.3.3. DI\_Produtos**

Dimensão corporativa, contém informações a respeito de cada produto que conste no mix da empresa. Esta é, talvez, a dimensão que irá ter a maior quantidade de variantes, pois dependerá do ramo de atividade principal da empresa. No caso específico, estamos considerando que a empresa alvo é uma rede de supermercados. Os atributos Seção, Depto e Grupo são os indicadores de nível das diversas categorias. Como um exemplo, um desodorante será da Seção “HIGIENE”, do Depto “PERFUMARIA” e do grupo “COLONIAS E DESODORANTES”.

#### **4.3.4. DI\_Clientes**

Esta é, num ambiente de CRM, a dimensão mais importante. É aqui que estarão cadastrados os clientes e seus respectivos dados principais utilizáveis para análise. Como estamos desenvolvendo um modelo mais genérico, optamos pelos dados principais, acrescidos do atributo FaixaRenda para indicar o rendimento mensal do cliente.

#### **4.3.5. DI\_Loja**

A última dimensão é utilizada para armazenar as lojas e suas diversas variáveis de análise, tanto geográficas quanto administrativas. Ressaltamos

aqui os atributos NumCaixas, NumFuncionarios e Area. Esses atributos são muito utilizados para determinar taxas de desempenho, tais como faturamento por m<sup>2</sup>, por ECF e por funcionário. Essas taxas de produtividade são adotadas a nível nacional para avaliação de desempenho das unidades varejistas. Além disso, existem os atributos de Regional e Classe, que são utilizados para comparação de desempenho entre as lojas de uma mesma regional ou de classes iguais.

#### **4.3.6. F\_Vendas**

A maior tabela do DM, quanto à quantidade de registros. Além das informações das chaves de dia, hora, produto, cliente e loja, fazem parte da chave primária o número do ecf (NumECF) e o número do cupom fiscal (NumCF). Pode-se considerar estes atributos como falsas dimensões, válidas conforme os modelos dimensionais propostos por KIMBALL (1996), KIMBALL et al (1998) e INMON (1997), mas que levam a granularidade ao seu nível máximo. Assim, o modelo poderá determinar exatamente o que, quando e onde cada cliente levou. Como medidas dos fatos, temos os três atributos básicos: a quantidade de produto levada (Qtd), o preço pago por eles (ValorVenda) e a lucratividade alcançada (Lucratividade).

Para se chegar a este *Data Mart*, os dados tendem a provir de várias fontes distintas. As dimensões temporais (DI\_Data e DI\_Hora) podem ser criadas através de um sistema de controle de dimensões, o que oportuniza permitir a definição de faixas de datas para as estações do ano e para as temporadas de venda. As dimensões de produtos (DI\_Produtos) e clientes (DI\_Clientes) são alimentadas pelas informações existentes no sistema de retaguarda da

empresa. A dimensão de lojas (DI\_Lojas) normalmente é totalmente editada pelo sistema de controle.

De posse das dimensões alimentadas e validadas, resta a parte principal do processamento. As vendas são colhidas diretamente dos arquivos de venda (Logs de transações) dos ECFs. Para cada item vendido, é encontrada a sua correspondência nas dimensões. Este processamento é realizado através de tabelas auxiliares que funcionam como o *back room* do DM. Nestas tabelas é que conseguimos cruzar as informações de venda com os clientes.

Algumas transformações são necessárias para se chegar aos dados finais. Como exemplo podemos citar os produtos. Cada produto tem um código interno, que representa aquele produto no mix. Entretanto, este produto pode ter um ou mais códigos de barra (chamados de *códigos EAN*) associados a ele. Como os logs de transações possuem apenas o código EAN, é necessário cruzar esta informação em uma tabela auxiliar, onde co-existem os dois códigos. Além disso, esta tabela de produtos auxiliar também vai armazenar o preço de custo da mercadoria, para poder determinar a lucratividade de cada compra efetuada pelo cliente.

Definido o DM que irá armazenar as informações, falta selecionar os clientes, dentre os já cadastrados, para o teste piloto. Esta seleção pode ser realizada de duas formas diferentes:

- A partir do cadastro do sistema de retaguarda;
- A partir dos arquivos de vendas, onde estão cadastrados todas as operações com cheques realizados durante o dia.

Se os dados forem retirados do sistema de frente de caixa, deve-se ainda cruzar as informações com o cadastro da retaguarda, pois somente são armazenadas informações como CPF/CNPJ do emitente.

Independente da fonte é possível selecionar os clientes por três critérios diferentes: pela frequência, pelo valor das compras ou pela lucratividade gerada pelo cliente.

Após a definição do filtro de seleção, recomenda-se ainda selecionar apenas os clientes que estejam acima da média do filtro escolhido. Pode-se ainda escolher apenas os clientes de uma região delimitada, para poder-se testar todas as funcionalidades do programa de forma controlada.

No quarto passo, após a definição do conjunto piloto de clientes, deve se definir as estratégias de marketing para fazer com que os mesmos associem-se ao programa de fidelização em implantação. Uma possível forma seria a emissão de um bônus com direito a um desconto na compra de um determinado produto em troca da atualização cadastral e aceitação da participação no programa. Ou então, trocar por um brinde de valor simbólico.

O contato com os clientes poderia ser via mala direta (com o risco de uma baixa de retorno) ou telemarketing ativo, com as operadoras explicando como funcionaria o programa. Uma programação mista de contato talvez seja a melhor solução: primeiro, o contato via telemarketing, com as explicações necessárias, perguntando por fim se o cliente gostaria de receber o material de inscrição. Este material seria então remetido via correio, somente para aqueles clientes que demonstrassem interesse em se associar.

Quando o quarto passo for ativado, o programa de frente de caixa já deverá estar habilitado a identificar o cliente durante a compra mediante a apresentação de um cartão de identificação. Além desta opção, pode-se configurar o caixa para executar uma série de funções extras, com o intuito de executar ações de marketing instantâneas (ver discussão adiante).

É interessante fixar, neste passo, um tempo limite para que os clientes se associem ao programa. Algo em torno de 30 a 45 dias (contando desde o início dos contatos via telemarketing) pode ser considerado razoável.

Tendo os clientes se associado e começado a se identificar no momento da compra, chegamos ao quinto passo, que é o desenvolvimento de ações-piloto de marketing com o intuito de testar a receptividade deste grupo de controle.

Nesta etapa iniciamos o uso mais intensivo do DM, extraído do mesmo diversos grupos de clientes com determinadas características e aplicando ações de marketing. Essas ações podem ser desde promoções até ações de venda cruzada (*cross selling*). Estas ações podem, inclusive, envolver algum tipo de parceria com fornecedores, pois em troca de descontos maiores a empresa pode repassar informações de consumo de seus clientes.

Há duas formas clássicas de se aplicar estas ações. A primeira utilizando malas diretas ofertando, por exemplo, cupons de desconto para determinadas mercadorias. A segunda forma seria, por exemplo, utilizar as redes de comunicação de dados para realizar ações instantâneas.

Nos dias atuais, é extremamente comum que todas as lojas sejam conectadas, via redes de longa distância, aos computadores da matriz. Isto é válido inclusive para os computadores que operam os ECFs, pois os mesmos

devem ter a capacidade de realizar transações de TEF (transferência eletrônica de fundos) e de consultas diversas, como por exemplo consultas aos serviços de proteção ao crédito (como SERASA e SPC). Esta infraestrutura pode ser aproveitada para a realização de ações instantâneas, bastando que o programa envie ao computador central o código do cliente e receba, do mesmo, instruções de ação. Este procedimento pode ser realizado de forma rápida e eficiente através de um protocolo assíncrono de alto nível entre o programa do caixa e o computador que armazene o banco de dados de ações de marketing. Um exemplo de ação instantânea poderia ser a de ofertar um brinde simbólico ao cliente na semana de seu aniversário.

Este tipo de protocolo pode ser facilmente especificado e implementado, tendo por base uma rede de comunicações já em uso, principalmente se a rede é baseada no protocolo TCP/IP, que já é, hoje, um padrão de interconexão de estabelecimentos. Obviamente, o sistema de frente de caixa deverá estar preparado para operar este protocolo de alto nível.

O sexto passo deve ser executado em concomitância com o quinto, pois é neste passo que serão medidas as respostas às ações de marketing executadas no passo anterior. O objetivo desta medição é determinar a correção ou não tanto da ação quanto do agrupamento de clientes-alvo selecionados.

Com os resultados, podemos afinar e calibrar o sistema de maneira a obter o maior retorno possível para as ações planejadas.

O último ponto deve ser definido com muito critério, pois um programa com baixas recompensas pode não atrair o cliente ao mesmo, ao passo que um

programa com recompensas dispendiosas podem atrair em um primeiro momento, mas desestimular o cliente quando o mesmo perceber que será muito difícil alcançar os pontos necessários. Tome-se, como exemplo, um programa que atribua 3 pontos a cada R\$ 10,00 em compras. Se o cliente deseja trocar os pontos por um produto caro, por exemplo uma garrafa de whisky importado, terá que acumular 1.350 pontos, o equivalente a R\$ 4.500,00 reais em compras. Sabendo-se que este produto tem um preço de venda de R\$ 88,00 (oitenta e oito reais), podemos considerar que o mesmo equivale a aproximadamente 1,95% do total dispendido em compras. Pode-se dizer que este percentual é descontado da compra de cada produto para ser creditado em favor do cliente para trocas futuras. Se o percentual fosse de, por exemplo, 1.5%, o cliente deverá dispendir R\$ 5.866,66 para conquistar o direito ao mesmo prêmio (considerando-se inalterada a relação 3 pontos para R\$ 10,00 de compra). Como se pode notar, esta definição é mais comercial/administrativa do que técnica.

Esta metodologia permite que um projeto piloto de implantação de CRM em ambiente de varejo de massa possa ser implementado num período que varia entre 180 e 260 dias. Ao final deste período, pode-se começar a implantação definitiva, com uma campanha maciça de marketing conclamando os clientes a se associarem ao programa.

#### **4.4. O passo alternativo: conhecer o cliente**

Este modelo genérico pode ser utilizado para uma implementação de CRM em qualquer empresa. Isso envolve, no entanto, uma grande quantia de tempo e, principalmente, recursos financeiros para o seu completo desenvolvimento.

Um passo alternativo pode ser realizado internamente na empresa sem que o modelo como um todo seja implementado: trata-se de utilizar e manipular os dados de venda e as informações básicas de retaguarda de modo a traçar um perfil dos clientes e de seus perfis de compra.

Este passo pode ser realizado com uma grande economia de recursos, necessitando apenas de recursos humanos com conhecimentos básicos sobre bancos de dados, *data warehouses*, mineração de dados e estatística básica. Além das pessoas envolvidas, são necessários apenas um equipamento com boa capacidade de processamento, de um bom Sistema Gerenciador de Bancos de Dados (SGBD) e de programas estatísticos e de mineração de dados.

Atualmente, com o advento dos chamados softwares livres (também conhecidos como *freewares*), existem excelentes programas que podem ser utilizados para este fim, a um custo mínimo. Todos os programas necessários para o desenvolvimento deste passo podem ser adquiridos sem custos.

Mesmo que se torne necessário a utilização de softwares que demandem o pagamento de licenças, o custo das mesmas torna-se irrisório em comparação aos preços praticados pelos desenvolvedores de programas de CRM.

Este passo alternativo compõe-se das seguintes tarefas:

- a) Definir e construir o *data mart* proposto;
- b) Alimentar o referido *data mart* com todos os dados necessários aos estudos;
- c) Recuperar as informações pertinentes às compras de cada cliente;
- d) Realizar análise estatística sobre estes dados;

- e) Realizar estudos com técnicas de mineração de dados sobre os mesmos;
- f) Propor ações de marketing a partir do conhecimento adquirido com este passo e realimentar o *data mart* com os novos dados de venda, reprocessando novamente as informações a partir da tarefa **(b)** acima.

O estudo de caso apresentado no próximo capítulo mostra como implantar este passo e como se pode realizar os estudos a cerca do comportamento dos clientes.

## **5. ESTUDO DE CASO**

Para avaliar a metodologia proposta, torna-se necessário realizar um estudo de caso, com o objetivo primordial de validar ou não esta metodologia. O presente capítulo traz um estudo limitado, realizado junto a uma rede de supermercados de porte médio.

### **5.1. A empresa**

Para o desenvolvimento da metodologia, foi escolhida uma rede de supermercados de porte médio, com atuação apenas no estado de Santa Catarina.

Fundada em 14 de agosto de 1974 na cidade de Santo Amaro da Imperatriz (SC), possui 20 (vinte) lojas, com maior concentração de unidades na Região Metropolitana de Florianópolis (15) e dois centros de distribuição, um para cargas secas e bebidas, situado junto à administração central, e outro para produtos que exigem refrigeração ou congelamento. Conta atualmente com cerca de 1.500 colaboradores divididos entre os setores de operações (lojas), apoio (centros de distribuição) e administrativo/financeiro (administração central).

Com um faturamento anual de aproximadamente R\$ 150 milhões, situa-se na 38ª posição no ranking da ABRAS do ano de 2001. A administração da empresa ainda é familiar, com a diretoria dividida entre seis irmãos, filhos do fundador da empresa. Entretanto, já projetando uma futura profissionalização da administração, a Diretoria Operacional e de Recursos Humanos é ocupada por um executivo alheio à família.

Com um mix de aproximadamente 25.000 produtos, a maioria das lojas pode ser considerada loja de bairro, com menos de dez check-outs por unidade e área de vendas igual ou menor que 800 m<sup>2</sup>. Entretanto, as unidades mais recentes já foram construídas em instalações mais amplas, com mais de 1.500 m<sup>2</sup> de área de vendas e até 20 check-outs, com o objetivo de atender a uma parcela maior de clientes e com maior valor agregado nos produtos oferecidos. Possui duas unidades nos maiores shoppings da Região Metropolitana de Florianópolis, ambas atuando em regime de 24 horas. Para atender a todos os clientes, possui 174 check-outs em todas as unidades.

Considerando apenas os clientes que pagam com cheques, que farão parte deste estudo, o fluxo é de aproximadamente 798 clientes por dia no período, com uma média de R\$ 86,71 por compra.

## **5.2. Características tecnológicas**

Para fazer frente às necessidades de atendimento, a rede possui uma grande quantidade de equipamentos de informática, divididos entre *check-outs*, servidores, terminais de consulta e equipamentos de interconexão de redes, tanto locais quanto de longa distância.

### **5.2.1. Frente de caixa**

Os equipamentos de frente de caixa (*check-outs*) são aqueles utilizados para o registro dos produtos e das formas de pagamento de cada compra. Cada *check-out* é composto por um computador, monitor de vídeo, *scanner* (equipamento que faz a leitura e decodificação dos códigos de barras dos produtos), teclado reduzido (mais apropriado à operação do que os teclados

comuns) e os chamados *PIN/PAD*, pequenos terminais utilizados para ler as tarjas dos cartões magnéticos e as senhas pessoais dos clientes, quando necessário. A parte principal de cada *check-out* é representada pelas impressoras fiscais, equipamentos inteligentes que imprimem os cupons de venda, ao mesmo tempo em que registra os valores das vendas e dos impostos a serem recolhidos pelo estabelecimento. A rede em estudo possui três modelos de impressoras fiscais, das marcas SWEDA (IF-7000 II), URANO (2EFC) e UNISYS (Calypso).

Nos *check-outs* equipados com impressoras SWEDA e URANO, que correspondem a 170 caixas, é utilizado um programa que foi desenvolvido por uma empresa local e que roda sob versões do sistema operacional Microsoft Windows 95 ou superior. Todas as informações pertinentes às vendas são gravadas em arquivos de formato texto, denominado *log de transações*, sendo um arquivo para cada dia (ou data fiscal). Estes arquivos são de fácil acesso e manipulação e foram utilizados na montagem do *data mart* de vendas que é a base de nosso estudo. Os outros 04 (quatro) caixas, da marca UNISYS, utilizam programas proprietários do fabricante, que não forneceu as especificações dos arquivos de venda. Desta forma, e considerando o baixo impacto representado por estes equipamentos no total das transações, resolveu-se não considerar as vendas registradas por eles.

### **5.2.2. Servidores**

Cada uma das lojas possui o seu próprio servidor. São computadores rodando uma versão de UNIX (SCO ou Conectiva Linux) e banco de dados ZIM. Estes servidores rodam o sistema de retaguarda de cada unidade,

responsável pelo controle de estoque, comercialização, gestão de produtos e preços etc. Esses servidores atuam, além disso, como servidores de arquivos, armazenando cópias dos *logs* de transações dos caixas, que são utilizados como fontes de dados para o processamento das vendas de cada loja.

Na administração central são utilizados basicamente 03 (três) servidores distintos: o servidor central (responsável pelo sistema de retaguarda da Central Administrativa e Depósito Central); o servidor de TEF (responsável pela interconexão da empresa e as administradoras de cartão) e o servidor de internet/intranet (responsável pelo suporte a correio eletrônico, transferência de arquivos e publicação de páginas na *web*).

### **5.2.3. Terminais de consulta**

Existem dois tipos de terminais de consulta nas lojas. Um, na saída do depósito, é utilizado pelos reposidores, conferentes e promotores de venda para conferir os preços dos produtos e seus cadastrados antes que os mesmos sejam encaminhados para as gôndolas. Por este motivo, este tipo de terminal recebe o apelido de “mata-burro” e roda um ou mais módulos do sistema de retaguarda.

Já na área de vendas existe outro tipo de terminal, mais simples, que simplesmente informa o preço unitário a partir da leitura do código de barras de um produto. É utilizado pelos clientes para dirimir dúvidas em relação ao preço de cada produto. É conhecido como “São Tomé” e roda, normalmente, um programa específico.

### 5.2.3. Equipamentos de interconexão de redes

Cada loja possui uma rede local ligando todos os seus computadores. Em todas elas se utilizam redes com topologia em estrela, com o nó central sendo ocupado por um *hub* ou *switch*. O meio físico utilizado é composto por cabos UTP categoria 5, na maior parte dos casos, ou cabos coaxiais em casos específicos. O protocolo utilizado é o TCP/IP, o que permite a utilização do gerenciador de arquivos SAMBA nos servidores. O SAMBA é um *freeware* (software livre) que faz com que um servidor rodando UNIX ou LINUX “apareça” na rede como uma máquina rodando MS Windows, permitindo a gravação de arquivos diretamente no servidor a partir das estações e dos caixas.

Para a interligação das lojas e Centros de Distribuição é utilizada uma rede de longa distância por comutação de pacotes com tecnologia *frame relay*, o que permite uma comunicação estável, direta e de alta velocidade entre as unidades e os servidores centrais. Além da rede *frame relay*, é utilizada uma rede com protocolo X.25 para a conexão com administradoras de cartão (de crédito e/ou débitos) e serviços de proteção ao crédito (SERASA/SPC). Em cada unidade é necessário um roteador, com o objetivo de prover a interconexão da rede interna (local) com a rede de longa distância. Na administração é necessário um roteador com dupla função, ou seja, que acesse tanto a rede *frame relay* quanto a rede X.25.

#### 5.2.4. Sistemas gerenciadores de bancos de dados

Para fazer frente às necessidades de manipulação de dados provenientes das atividades comerciais da empresa, são utilizados dois sistemas gerenciadores de bancos de dados (SGBD):

- Nas lojas e no sistema de retaguarda da administração central é utilizado o SGBD ZIM, desenvolvido pela empresa canadense Zim Technology Inc.;
- Para sistemas de consulta *ad-hoc* e para os sistemas de gerenciamento (*business management*), utiliza-se o SGBD MS SQL Server, versão 2000, que possui várias ferramentas de auxílio a criação e manutenção de *Data Warehouses*, além de ferramentas de *Data Mining*.

### 5.3. Dados a serem utilizados

Para a montagem do DM de vendas, tornou-se necessário obter dados de várias fontes distintas dos sistemas da empresa. Cada uma das tabelas de dimensão do modelo proposto é montada e/ou atualizada de forma diferente, inclusive utilizando-se de mais de uma fonte de dados. A seguir será explanado como estas tabelas foram montadas e de que fontes os dados foram retirados:

#### 5.3.1. Dimensão “DI\_Dia”

Esta dimensão é gerada automaticamente por uma rotina do programa de atualização do DM. Para cada dia foi gerado um registro na tabela, com os atributos *ch\_dia*, *data*, *dia*, *mês*, *ano* e *dia\_da\_semana* sendo atribuídos diretamente pela rotina. Os atributos *feriado*, *vespera\_feriado* e *temporada*

foram atribuídos manualmente, tendo em vista as diferentes formas de se determinar os feriados móveis e o início e término das temporadas.

### **5.3.2. Dimensão “DI\_Hora”**

Esta dimensão, por ser de pequeno porte, foi totalmente preenchida de forma manual, com os valores dos atributos *ch\_hora* e *hora* sendo coincidentes.

### **5.3.3. Dimensão “DI\_Loja”**

Esta dimensão também foi preenchida de forma manual, tendo em vista que as informações contidas nela são de baixíssima atualização e a quantidade de registros também é baixa. Todos os dados foram conseguidos através de entrevistas aos diversos departamentos da empresa.

Os dados cadastrais (número, endereço, cidade, bairro, cep etc) das lojas foram repassados pelo Departamento Comercial; os dados sobre as áreas de venda foram repassados pelo Departamento Operacional; o número de funcionários foi repassado pelo Departamento de Recursos Humanos e o número de caixas foi repassado pelos Departamentos de Informática e de Contabilidade.

### **5.3.4. Dimensão “DI\_Clientes”**

Esta dimensão é formada por dados oriundos do cadastro de clientes do sistema de retaguarda, armazenados em duas tabelas relacionadas, chamadas ESM300 (dados principais) e ESM302 (complemento do cadastro).

Os dados são recuperados inicialmente em um arquivo texto localizado no servidor central, através de comandos apropriados emitidos no ambiente

interativo do sistema gerenciador ZIM. Após os dados terem sido gravados no arquivo texto, o mesmo é transferido para o servidor do DM para ser processado pelo programa de manutenção do mesmo.

#### **5.3.5. Dimensão “DI\_Produtos**

Os dados referentes aos produtos podem ser divididos em duas categorias: uma se refere aos produtos propriamente ditos; a outra se refere ao cadastro de famílias, como são chamadas as subdivisões sob as quais os produtos são agrupados. Como visto no capítulo 2, estas subdivisões auxiliam no correto gerenciamento do mix de produtos da empresa. No caso da empresa em estudo existem três subdivisões, denominadas **Seção**, **Departamento** e **Item**. Apesar do nome, a última subdivisão refere-se a um grupo de produtos, não à um item específico. Como um exemplo, uma determinada marca de ração para cães estará cadastrada na seção *Mercearia*, no departamento *Alimentos para Animais* e nos itens (grupo) *Rações e Sementes*.

Os dados referentes ao cadastro de famílias residem em uma tabela denominada ESMXXX no sistema de retaguarda, enquanto que os dados sobre os produtos em si ficam armazenados nas tabelas ESM100 e ESM975.

A recuperação dessas informações é realizada em duas etapas. Primeiro, o cadastro de famílias é recuperado através de comandos para o sistema ZIM e gravado em um arquivo texto. Em seguida, os dados referentes aos produtos são extraídos pelo mesmo processo. Ao final, restarão os dados gravados em dois arquivos textos separados e residentes no servidor principal do sistema de retaguarda. A forma como esses arquivos são combinados será explanada no item 5.4.

### **5.3.6. Dados de vendas**

Os dados oriundos das vendas são, com certeza, os dados mais importantes deste estudo. São, também, os de maior volume e de maior complexidade para serem extraídos.

Durante o processo de venda, todos os dados relativos aos produtos, pagamentos e totalizações são armazenados em arquivos texto, com uma cópia no caixa e outra no servidor da loja. Para cada dia de vendas, é gravado um arquivo diferente por caixa. O conjunto de arquivos dos caixas de um determinado dia contém todas as informações financeiras, administrativas e de vendas necessárias para gerenciar o estoque da loja e a movimentação financeira dos caixas.

Para se obter estes arquivos, tornou-se necessário a implantação de uma rotina de busca desses arquivos nas lojas e o seu posterior processamento para, em conjunto com os dados oriundos de outras fontes, formarem o DM de vendas. A rotina de busca consiste em criar “pacotes” compactados com todos os arquivos que formam o movimento de um determinado dia e transmiti-los para o equipamento responsável pelo seu processamento.

## **5.4. Montagem do *data mart***

A partir do momento em que foram determinadas as fontes de cada uma das informações que irão compor o DM, foi necessário determinar como este seria armazenado e como os dados brutos seriam processados para formar o mesmo.

O sistema gerenciador de bancos de dados (SGBD) escolhido foi o Microsoft SQL Server 2000 Personal Edition, que foi escolhido por três motivos básicos

- a) O mesmo já era utilizado pelo cliente;
- b) O SGBD possui ferramentas bem desenvolvidas de mineração de dados e de OLAP (*on-line analytical process* – processo de análise *on line*);
- c) Ter capacidade de operar com grandes conjuntos de dados com agilidade.

Como o objetivo do estudo é demonstrar que se pode realizar análises das compras com um baixo custo, poder-se-ia ter escolhido um SGBD gratuito, como o MySQL, o Interbase ou o Firebird (sucessor do Interbase). Mesmo sendo uma versão paga, ainda assim pode-se considerar que uma análise baseada no SGBD escolhido terá um custo muito baixo, se comparados aos custos mínimos de implantações de pacotes de CRM por parte das empresas especializadas.

Definida a questão da escolha do SGBD, desenvolveu-se um pequeno programa em Visual Basic (versão 6.0) para a alimentação do DM com os dados. Para cada fonte de dados criou-se uma rotina especial, pois os dados chegam de várias fontes e nos mais diversos formatos, com única semelhança de serem arquivos texto.

#### **5.4.1. Dimensão DI\_Dia**

Esta dimensão é gerada automaticamente pelo programa, conforme já explanado. É necessária apenas a informação da faixa de datas a serem criadas. O programa cria os registros, um por dia, com as informações básicas

a respeito de cada data (dia, mês, ano, data completa, dia da semana) além da chave primária da tabela, criada automaticamente e que será utilizada na tabela de fatos.

#### 5.4.2. Dimensão **DI\_Clientes**

Esta dimensão é criada de forma indireta. Como primeiro passo, os dados de clientes são extraídos do sistema de retaguarda. Nesta extração, todos os clientes são gravados em um arquivo texto, que é copiado para o computador responsável por sua manipulação. O segundo passo consiste em carregar este arquivo em uma tabela temporária (chamada "**Cientes1**"), gerando um campo *hash* para cada registro. No caso do banco de dados utilizado, um campo *hash* é um campo que irá armazenar o resultado de uma operação que terá como operadores os demais campos do registro.

Após a carga da tabela **Cientes1**, são executados mais dois passos: Primeiro, são inseridos os clientes que pertencem à tabela **Cientes1** e que não pertençam à dimensão **DI\_Clientes**. Esses clientes são aqueles que foram incluídos no sistema de retaguarda mas que ainda não haviam migrado para o DM. Após a inclusão, é executado o segundo passo, que é a atualização dos clientes já existentes. Os clientes que sofreram algum tipo de alteração nos dados terão o campo *hash* diferente nas duas tabelas. Desta forma, basta alterar os dados em **DI\_Clientes** a partir dos dados de **Cientes1**.

#### 5.4.3. Dimensão **DI\_Produtos**

A exemplo da dimensão **DI\_Clientes**, esta dimensão também é criada de forma indireta, através da junção de duas tabelas auxiliares, a de **Produtos** e a

de **Famílias**. Estas tabelas são, inicialmente, geradas pelo sistema de retaguarda e carregadas, uma por vez, nas tabelas auxiliares.

A tabela de **Famílias** traz toda a divisão hierárquica dos produtos no cadastro, conforme já explanado no item 5.3.5; ela é composta da descrição dos níveis e de um código numérico de seis dígitos representando cada folha da árvore. Este código também aparece no cadastro de produtos e será utilizado como chave para a junção das duas tabelas.

A tabela de **Produtos** contém as outras informações a respeito de cada um dos produtos comercializados. A principal característica desta tabela é a de possuir dois códigos para cada produto: o código interno, que identifica univocamente o produto no cadastro da retaguarda, e o código de barras (nos padrões EAN-13, EAN-8 e/ou UPC-12), que é o utilizado pelo sistema de frente de caixa e que é gravado nos arquivos de transação. A existência de dois códigos para cada produto deve-se ao fato de que um mesmo produto pode ter mais de um código de barras. Como exemplo, uma determinada marca de óleo de soja, com as mesmas características de embalagem, volume e composição, produzido por duas fabricas diferentes, pode ter um código de barras diferente para cada unidade fabril.

Para a formação da dimensão de produtos, o código de barras será suprimido, pois toda a referencia será feita através do código interno do produto, mesmo na tabela de fatos.

O processo de carga da dimensão de produtos segue os mesmos passos do descrito para a dimensão de clientes: carga nas tabelas auxiliares (**Famílias** e **Produtos**), junção das duas tabelas e inserção na dimensão dos produtos

incluídos e a posterior modificação, na mesma dimensão, dos produtos alterados.

#### **5.4.4. Tabela de Fatos F\_Vendas**

Esta tabela demanda um pouco mais de trabalho para ser carregada do que as dimensões. Em primeiro lugar, é necessário buscar todos os arquivos de transação das lojas e armazená-los no equipamento responsável por seu processamento, conforme relatado no item 5.3.6.

Após os arquivos estarem todos a disposição, o processo de montagem da tabela de fatos é disparado. Em primeiro lugar, todos os índices da tabela de fatos são derrubados, para melhorar o desempenho da inserção de blocos de informações. Na seqüência, um arquivo de transações é aberto e processado, registro a registro. Conforme se pode observar pelos formatos de registros destes arquivos (listados no Anexo), cada operação é identificada por um código específico. Para os propósitos do presente estudo, são necessários apenas os registros com códigos 100 (venda de item), 130 (cancelamento de item), 200 (cancelamento de venda) e 370 (registro de pagamento com cheque). Todos os registros do tipo 100 são decompostos e gravados em uma tabela chamada **TRX100**. Entre os dados gravados estão: a data, hora, número do caixa, número do cupom fiscal e os dados sobre o produto comercializado (código de barras, quantidade, preço unitário e preço total) . Quando um registro 130 é encontrado, o mesmo é decomposto para se obter as especificações do item de venda a ser cancelado. Com essas informações, é emitido um comando de exclusão na tabela **TRX100** para eliminar o registro específico. O processamento do registro 200 é semelhante, com a diferença

que todos os itens que fazem parte de um determinado cupom fiscal são eliminados de uma só vez. Se o pagamento de uma compra específico foi realizado com cheque, o registro 370 irá trazer, entre outras informações, o CPF/CNPJ do cliente e o número do cupom fiscal objeto do pagamento. De posse destes dados, o campo **CodCliente** da tabela **TRX100** é então preenchido com o código interno daquele cliente.

Quando o arquivo texto é processado em sua totalidade, os dados armazenados na tabela **TRX100** poderão, então, serem transferidas para a tabela de fatos. Entretanto, nesta tabela os produtos estão identificados pelo seu código de barras e não pelo código interno, que é o utilizado pela dimensão de produtos. Este problema é contornado fazendo-se a junção da tabela **TRX100** com a tabela auxiliar **Produtos** (ver item 5.4.3) através do código de barras. Essa junção vai fornecer, assim, o código interno do produto. Feita a junção, basta incluir os registros da tabela **TRX100** na tabela de fatos, através de um único comando SQL de inserção. Este comando faz a junção de todas as tabelas necessárias para fornecer os dados apropriados para cada registro. Como essa inserção é realizada ao final de cada arquivo texto processado, o tempo gasto é otimizado em função da pequena quantidade de registros manipulados de cada vez.

Após o processamento do arquivo de transações, a tabela **TRX100** é então esvaziada e um novo arquivo é aberto e processado. Estas etapas são repetidas até que todos os arquivos a disposição sejam manipulados. Ao final de todo o processo, os índices da tabela de fatos são então reconstruídos, para aumentar o desempenho das pesquisas a serem realizadas.

## 5.5. Análise dos dados

Para se mostrar o poder do modelo para a análise das informações de vendas individuais, fez-se uma análise dos dados utilizando-se duas ferramentas básicas: a análise estatística e a mineração de dados.

Ambas as análises tomaram por base um subconjunto do *data mart* completo. Como o *data mart* possui informações sobre todas as vendas realizadas no período do estudo, ele inclui aquelas vendas em que não houve uma identificação efetiva do cliente. Lembrando que o objetivo deste estudo é analisar as vendas dos clientes identificados, foi necessário retirar um subconjunto apenas com as vendas em que houve o pagamento com cheque e, portanto, houve a identificação efetiva do cliente. Esta primeira divisão reduziu o tamanho da tabela de fatos de 38.753.204 registros para apenas 4.630.897, mostrando que, dos 152.699 clientes cadastrados, apenas 56.616 (correspondente a 37,07%) efetuaram pelo menos uma compra com cheques no período de outubro/2002 a maio/2003.

Após esta primeira divisão, optou-se por uma segunda subdivisão. Isto se deveu ao fato de que muitos clientes identificados efetuaram poucas compras. Assim, como um ponto de corte, foram selecionados apenas os clientes que realizaram um mínimo de 08 (oito) compras no período, com média de uma compra por mês no período de estudo. Esta segunda divisão reduziu a tabela de fatos para apenas 2.712.954 registros e mostrou que apenas 6.481 clientes preencheram o pré-requisito de 08 (oito) compras no período de estudo. Estes clientes adquiriram 14.496 produtos diferentes de um *mix* de 25.155 produtos cadastrados. Esta discrepância entre a quantidade de produtos cadastrados e

os efetivamente comercializados deve-se ao fato de que a tabela de produtos possui muitos registros de artigos que já estão fora de linha (não mais comercializados), além dos chamados produtos sazonais, ou seja, produtos que somente são comercializados em determinadas épocas do ano, como os ovos de páscoa por exemplo.

Para facilitar a análise, foi criada uma tabela (**ComprasCli**) contendo vários fatos agregados de cada um destes 6.481 clientes. A primeira informação é a identificação do cliente na tabela **DI\_Clientes**, necessário para uma correta identificação do mesmo. Uma das mais importantes informações é o valor total de suas compras. Outra informação relevante é a sua frequência, ou seja, quantas compras distintas foram realizadas no período de estudos. Um terceiro atributo é a recência, ou seja, à quantos dias (tomando por base a data de 01/06/2003) o cliente não comparece à uma loja (quanto menor o valor, mais recentemente o cliente efetuou uma compra). Os outros dez atributos referem-se ao total das compras realizada por seção (mercearia, bebidas, açougue, bazar, frios, higiene, hortifrutigranjeiros, laticínios, padaria e limpeza).

Criadas as condições básicas, foram realizados dois estudos, um estatístico e outro por mineração de dados, descritos nos dois tópicos a seguir.

## 5.6. Análise estatística

A análise estatística foi realizada com o auxílio de uma planilha eletrônica, por suas facilidades de manipulação numérica.

Todos os dados da tabela **ComprasCli** foram transferidos para a planilha e procedeu-se à análise básica dos dados e seus atributos. O quadro 5.1 traz as medidas básicas de cada um dos atributos.

**Quadro 5.1. Análise dos dados das compras dos clientes selecionados**

Atributo	Mínimo	Máximo	Moda	Mediana	Média	Desvio Padrão
Freqüência (vezes)	8	121	8	12	15,12	9,19
Recência (dias)	1	221	1	12	22,31	27,72
Valor Total (R\$)	73,83	21.574,12	1.098,47	1.017,58	1.354,01	1.168,57
Mercearia (R\$)	0	6.874,26	175,61	374,60	512,52	473,48
Bebidas (R\$)	0	20.684,33	0	110,81	175,72	345,69
Açougue (R\$)	0	6.376,86	0	142,90	228,72	297,89
Bazar (R\$)	0	1.007,92	0	21,42	40,47	58,28
Frios (R\$)	0	2.309,88	0	51,87	77,63	91,97
Higiene (R\$)	0	2.135,77	0	67,18	105,16	123,12
Hortifruti (R\$)	0	2.094,59	0	58,26	92,92	117,53
Laticínios (R\$)	0	1.729,66	0	84,07	118,59	121,70
Padaria (R\$)	0	655,71	0	16,14	27,05	35,55
Limpeza (R\$)	0	1.963,40	0	92,23	140,66	152,86

Fonte: Dados de venda, período outubro/2002 a maio/2003

Algumas explicações são necessárias para o correto entendimento deste quadro:

- Como foram escolhidos apenas os clientes com um mínimo de 08 (oito) compras no período, este é, obviamente, o valor mínimo do atributo *freqüência*;
- Os valores máximos apresentam valores muito acima da média dos atributos. Isto decorre dos poucos registros que mostram comportamentos fora do padrão normal de compras, ou seja, fora do intervalo  $\bar{X} - 3s$  até  $\bar{X} + 3s$  (onde  $\bar{X}$  é a média e  $s$  representa o desvio padrão). Estatisticamente considera-se que 99,7% das observações estejam, dentro deste intervalo (DORIA F<sup>o</sup>, 2001). O caso do valor total do setor de bebidas é típico, considerando que o intervalo superior corresponde à R\$ 1.212,79 (valor correspondente a  $\bar{X} + 3s$ ), muito inferior ao máximo apresentado pelo atributo, que é de R\$ 20.684,33;

- É comum a ocorrência de valor zero nos totais por departamento, pois caracterizam aqueles clientes que deixaram de levar algum produto de um determinado departamento;
- Também é comum o fato da moda (valor com maior incidência no conjunto) de algumas variáveis possuírem o valor zero, pelo mesmo motivo;
- É possível verificar que o valor médio de venda por departamento está mais próximo do limite inferior do que do limite superior. Isto caracteriza distribuições assimétricas, à esquerda (média > moda).

Um dado que chama a atenção é o valor máximo do setor de bebidas. Ao se fazer uma análise mais detalhada das compras do cliente que gerou este valor, verificou-se que o mesmo estava concentrando suas compras apenas em refrigerantes, chegando a adquirir mais de 800 unidades em uma única compra, o que estava caracterizando uma venda por atacado, completamente fora dos padrões e normas da empresa.

Para verificar um possível inter-relacionamento entre os atributos da tabela **ComprasCli**, foi realizado um estudo de correlação linear. Para este estudo, foi tomado um determinado atributo e calculado o chamado *Coefficiente de Correlação* ( $r$ ) entre os mesmos, sendo que o valor de  $r$  pode variar entre  $-1$  e  $+1$ . Se a correlação é positiva, as variáveis são diretamente proporcionais. Se a correlação é negativa, as variáveis são inversamente proporcionais.

Segundo (CRESPO, 2002), para que exista uma correlação significativa entre as duas variáveis, é necessário que:  $0,6 \leq |r| \leq 1$ . Se o valor de  $r$  variar entre  $0,3$  e  $0,6$  ( $0,3 \leq |r| < 0,6$ ) existe uma correlação relativamente fraca.

Entretanto, se o valor de  $r$  variar entre 0 (zero) e 0,3 ( $0 \leq |r| < 0,3$ ) a correlação é muito fraca e, praticamente, nada podemos concluir sobre a relação entre as variáveis em estudo.

Tendo isso em mente, realizou-se o estudo que apontou as relações mais fortes entre os atributos analisados, como mostrado no quadro 5.2.

Como se pode observar, o maior grau de inter-relacionamento entre os atributos está entre o valor total das compras e o valor gasto com produtos do setor de mercearia ( $r=0,89211$ ). Como o coeficiente de correlação é positivo, isso significa que quanto maior o valor total, maior o valor gasto com produtos de mercearia.

**Quadro 5.2. Correlações “fortes” entre os atributos**

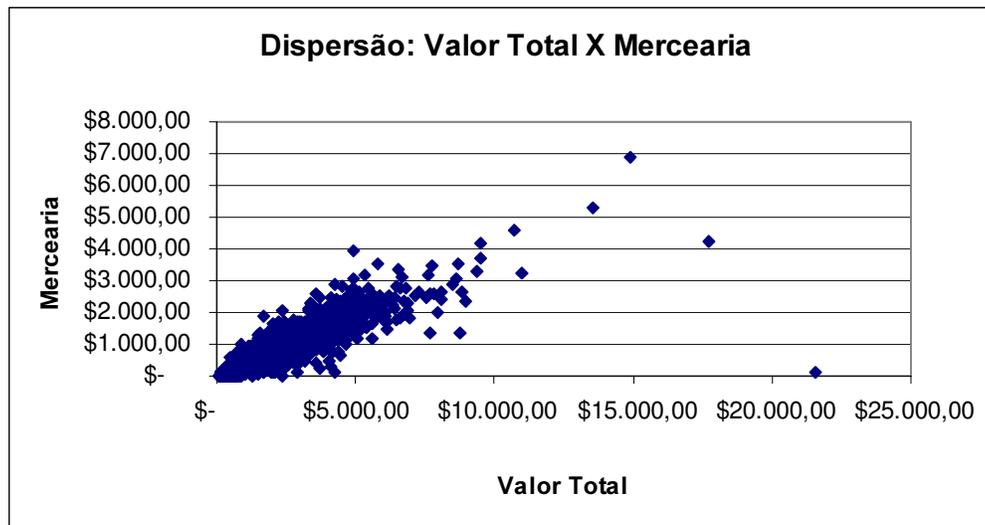
<b>Correlações</b>	<b>r</b>
Valor Total x Mercearia	0,89211
Mercearia x Limpeza	0,80488
Valor Total x Limpeza	0,80064
Valor Total x Laticínios	0,77624
Mercearia x Laticínios	0,74476
Valor Total x Frio	0,72303
Frios x Laticínios	0,68977
Valor Total x Higiene	0,68426
Higiene x Limpeza	0,67512
Laticínios x Limpeza	0,67330
Valor Total x Hortifruti	0,66961
Mercearia x Higiene	0,66784
Mercearia x Frios	0,66547
Valor Total x Açougue	0,66495
HortiFruti x Laticínios	0,64208

Fonte: Dados de venda, período outubro/2002 a maio/2003

Este tipo de relacionamento pode ser observado visualmente através do chamado *gráfico de dispersão*, onde são marcados os valores correspondentes às duas variáveis (uma em cada eixo), para todas as observações da mesma. Quanto mais a figura (normalmente uma elipse) se parecer com uma reta

inclinada, maior a relação entre as variáveis. No caso específico em estudo (valor total X mercearia), o gráfico 5.1 mostra esta relação de interdependência.

**Gráfico 5.1. Gráfico de dispersão entre as variáveis “valor total” e “mercearia”.**

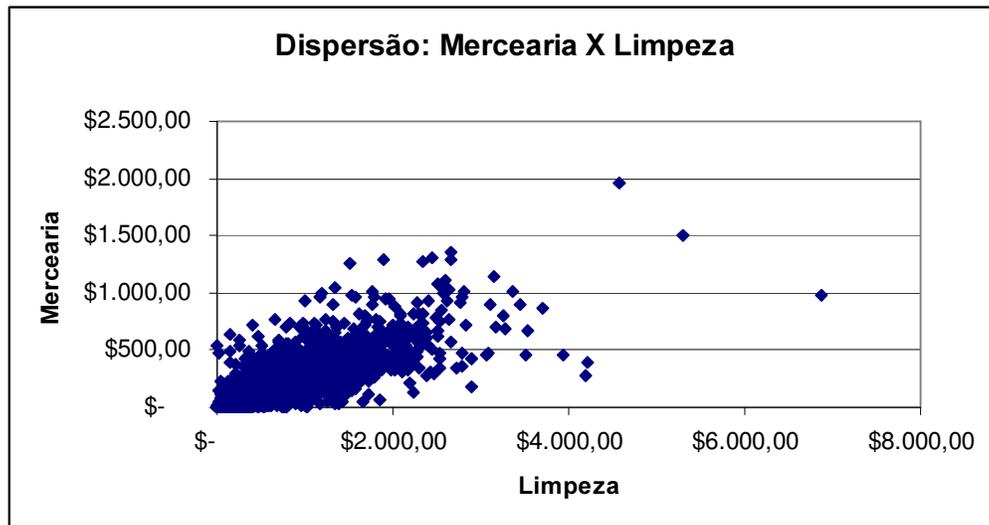


Fonte: Dados de venda, período outubro/2002 a maio/2003

A forte correlação entre o valor total gasto e os vários departamentos já era esperada. O que chama a atenção neste estudo é a forte correlação entre departamentos como, por exemplo, entre o valor gasto com produtos de mercearia e de limpeza ( $r = 0,80488$ ), ou entre o valor despendido nos departamentos de frios e laticínios ( $r = 0,68977$ ), como mostram os gráficos 5.2 e 5.3 a seguir.

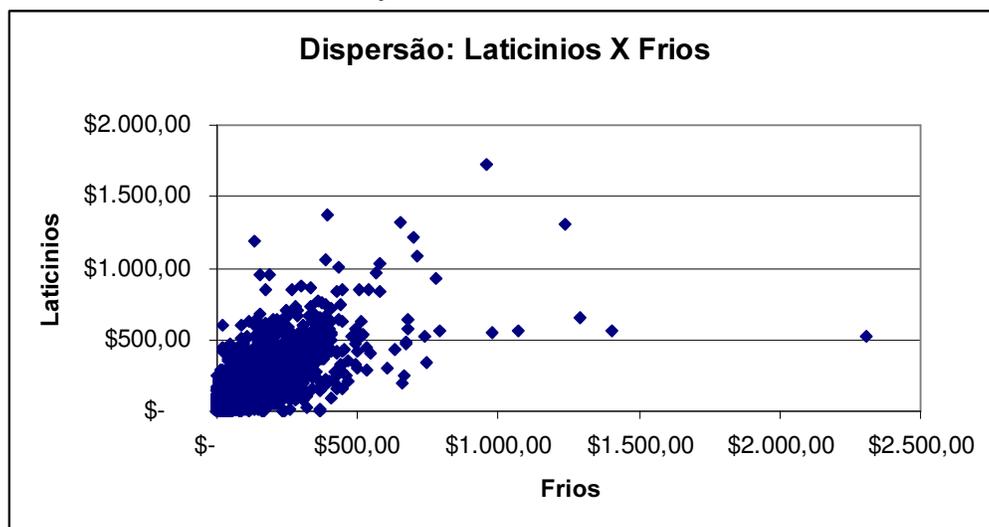
Como observado no gráfico 5.3, apesar da correlação entre as compras realizadas nos departamentos de laticínios e no de frios ser considerada estatisticamente “forte”, a elipse se apresenta mais aberta do que na primeira. Entretanto, ainda mostra um comportamento retilíneo, com uma tendência de crescimento semelhante para ambas as variáveis.

Gráfico 5.2. Gráfico de dispersão entre as variáveis “Mercearia” e “Limpeza”



Fonte: Dados de venda, período outubro/2002 a maio/2003

Gráfico 5.3. Gráfico de dispersão entre as variáveis “Laticínios” e “Frios”.



Fonte: Dados de venda, período outubro/2002 a maio/2003

Uma das correlações que mais chamou a atenção foi a que mostra uma forte tendência de compra conjunta entre os departamentos de hortifruti e o de laticínios ( $r = 0,64208$ ), teoricamente departamentos sem muita afinidade entre si.

**Quadro 5.3. Correlações “fracas” entre os atributos**

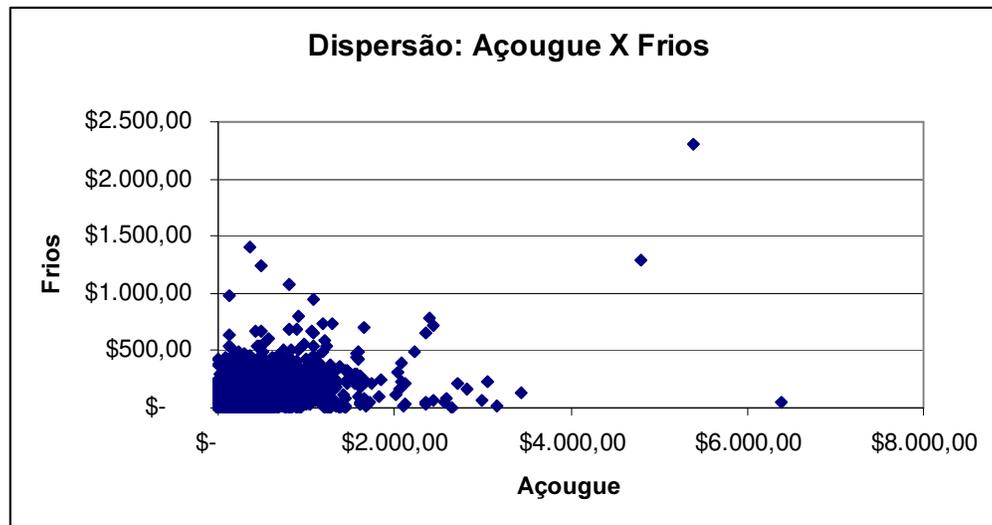
<b>Correlações</b>	<b>r</b>
Mercearia x Hortifruti	0,59979
Valor Total x Bebidas	0,58170
Valor Total x Freqüência	0,56964
Valor Total x Bazar	0,56556
Frios x Limpeza	0,56296
Higiene x Laticínios	0,56234
Freqüência x Laticínios	0,55425
Hortifruti x Limpeza	0,54895
Bazar x Limpeza	0,53880
Mercearia x Açougue	0,53297
Freqüência x Hortifruti	0,52534
Frios x Hortifruti	0,52307
Mercearia x Bazar	0,51065
Freqüência x Mercearia	0,49598
Frios x Higiene	0,49293
Freqüência x Frios	0,48015
Bazar x Higiene	0,47076
Bazar x Laticínios	0,45536
Açougue x Frios	0,44199
Freqüência x Padaria	0,43359
Freqüência x Limpeza	0,43315
Higiene x Hortifruti	0,42705
Bazar x Frios	0,41854
Valor Total x Padaria	0,41533
Açougue x Limpeza	0,40954
Bazar x Hortifruti	0,39449
Açougue x Laticínios	0,38531
Mercearia x Padaria	0,38005
Açougue x Hortifruti	0,37931
Freqüência x Açougue	0,37121
Frios x Padaria	0,36820
Laticínios x Padaria	0,36729
Hortifruti x Padaria	0,35960
Bebidas x Frios	0,31561
Açougue x Higiene	0,31208
Freqüência x Higiene	0,30920
Bebidas x Laticínios	0,30488
Mercearia x Bebidas	0,30301

Fonte: Dados de venda, período outubro/2002 a maio/2003

Além dessas correlações ditas “fortes”, foram detectadas outras que podem ser consideradas “fracas” ( $0,3 \leq |r| < 0,6$ ), conforme mostradas no quadro 5.3. Estes casos indicam que existe um certo grau de inter-relacionamento entre as variáveis, mas que não permite que sejam tiradas muitas conclusões sobre elas. Estas correlações apresentam gráficos de dispersão bem mais

espalhados, como mostrado no gráfico 5.4, que mostra a correlação entre os departamentos de açougue e de frios ( $r=0,44199$ ).

**Gráfico 5.4: Gráfico de dispersão entre as variáveis “Açougue” e “Frios”.**



Fonte: Dados de venda, período outubro/2002 a maio/2003

Como se pode observar, a figura mostra-se achatada, com um formato que lembra quase um retângulo ao invés de uma elipse, concentrada próximo à origem, com alguns pontos bem distintos, que representam os valores com menor incidência e mais distantes das médias de ambas as variáveis.

Por fim, aparecem aquelas correlações em que o valor do coeficiente  $r$  situa-se entre 0 (zero) e 0,3. Estas correlações estão listadas no quadro 5.4.

O que chama a atenção neste quadro é a presença maciça da variável que representa a recência. Isto significa que o fato do cliente ir várias ou poucas vezes às lojas não indica uma tendência do mesmo adquirir preferencialmente produtos deste ou daquele departamento. Além disso, não existe relação entre a recência e o valor total, nem entre a recência e a frequência ao estabelecimento.

**Quadro 5.4. Correlações “muito fracas” entre os atributos**

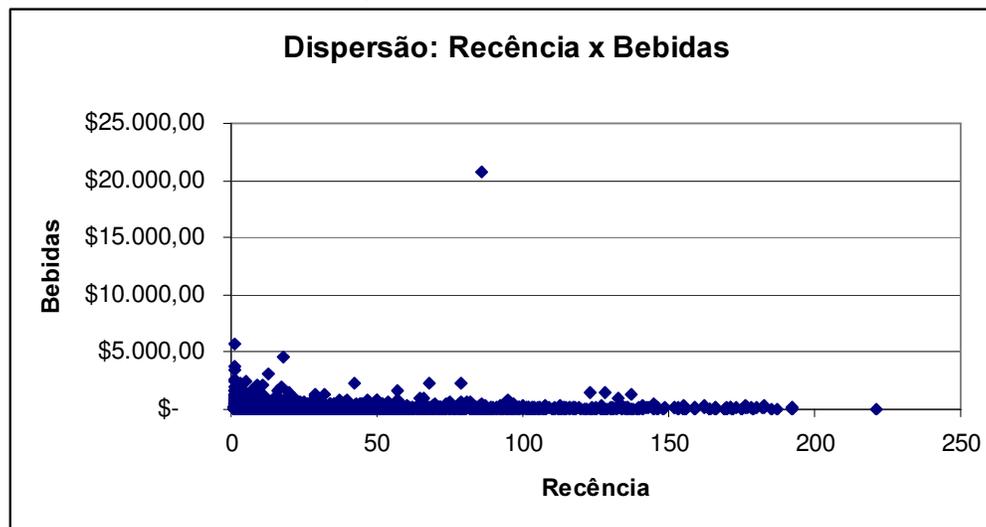
<b>Correlação</b>	<b>r</b>
Frequência x Bazar	0,29901
Bebidas x Limpeza	0,29728
Frequência x Bebidas	0,29281
Bazar x Padaria	0,29053
Padaria X Limpeza	0,28806
Açougue x Bazar	0,28535
Higiene x Padaria	0,25647
Bebidas x Hortifruti	0,25604
Bebidas x Higiene	0,24803
Açougue x Padaria	0,24216
Bebidas x Bazar	0,24166
Bebidas x Açougue	0,23303
Bebidas x Padaria	0,18011
Recência x Bebidas	-0,05727
Recência x Bazar	-0,08461
Recência x Açougue	-0,11567
Recência x Higiene	-0,13560
Recência x Padaria	-0,14804
Recência x Hortifruti	-0,15126
Recência x Frios	-0,17502
Valor Total x Recência	-0,18004
Recência x Limpeza	-0,18163
Recência x Laticínios	-0,19349
Recência x Mercearia	-0,22017
Frequência x Recência	-0,26060

Fonte: Dados de venda, período outubro/2002 a maio/2003

Como uma forma de demonstrar visualmente esta falta de correlação, foi criado o gráfico de dispersão entre a variável “recência” e o valor gasto no departamento de “bebidas”, que é mostrado no gráfico 5.5 abaixo.

Verifica-se facilmente que a falta de correlação entre as variáveis, com uma tendência muito forte dos valores de se aproximarem do eixo das abscissas (X).

Gráfico 5.5. Gráfico de dispersão entre as variáveis “Recência” e “Bebidas”.



Fonte: Dados de venda, período outubro/2002 a maio/2003

## 5.7. Análise por mineração de dados (*data mining*)

A análise estatística já fornece muitas informações valiosas, tais como a análise de correlação estudada anteriormente. Além dessa análise, utiliza-se também a chamada **mineração de dados**, ou seja, a análise dos dados disponíveis à procura de padrões não detectáveis através da análise estatística. Para o presente estudo, realizou-se a mineração através de três técnicas distintas: **Classificação por árvore de decisão**, **agrupamentos (*clustering*)** e **regras de associação**. A seguir, serão mostradas as conclusões deste estudo para cada uma das técnicas utilizadas.

### 5.7.1. Classificação por árvore de decisão

Esta técnica permite montar, a partir dos dados registrados, uma árvore que irá prever, com um determinado grau de probabilidade, o valor de uma das variáveis, baseado nos valores das demais variáveis em estudo. Como um exemplo, podemos utilizar esta árvore para prever, com um determinado grau

de certeza, se um determinado cliente será classificado em uma determinada faixa de consumo de produtos de mercearia, a partir dos valores apresentados nas compras de outros setores.

A tabela **ComprasCli**, já utilizada para a análise estatística, também foi utilizada como fonte de informações para esta análise classificatória. Para manipular os dados foi utilizado o algoritmo sistema Microsoft Analysis Manager, que faz parte do pacote do SGBD SQL Server 2000 Personal Edition, já utilizado para armazenamento do *data mart* de vendas.

Em uma primeira análise, foram inseridos todos as variáveis que compõem a tabela, com exceção da coluna que identificava o cliente. Como resultado, o sistema montou uma árvore classificatória para cada uma das demais variáveis.

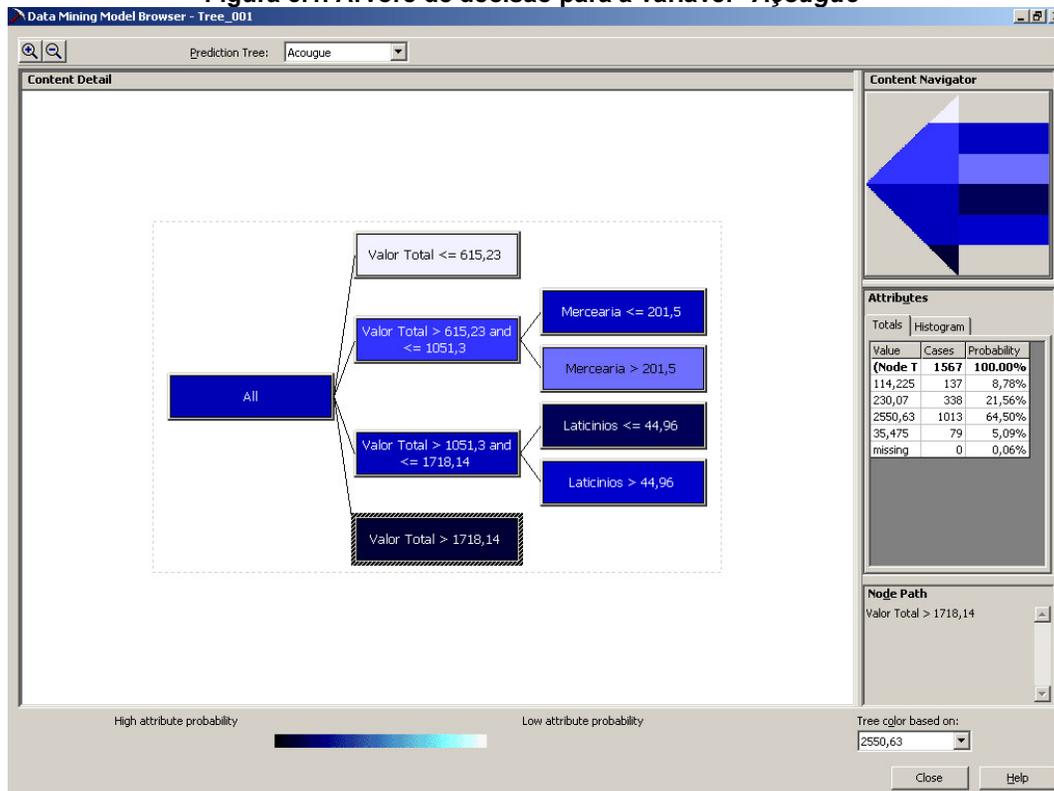
Como a maioria das variáveis representa os valores gastos em cada um dos departamentos, elas podem ser consideradas variáveis contínuas. No entanto, para que o algoritmo utilizado funcionasse a contento, foi necessário indicar ao programa que as variáveis deveriam ser automaticamente categorizadas, ou seja, o programa deveria separar os dados contínuos em categorias ou classes. Para tanto, o programa separa automaticamente os dados em quatro grandes grupos, determinados pelos três quartis de cada variável. Para as variáveis **Frequência** e **Recência**, que já são discretas, não houve a necessidade de se categorizar os dados.

No caso mais detalhado, o sistema montou uma árvore com quatro níveis, quando no caso menos detalhado a árvore possuía apenas a raiz. Neste último caso, o que ocorreu é que a árvore não pode ser montada, pois os dados

fornecidos para montar e “treinar” a árvore não conseguiram demonstrar algum tipo de relação possível entre as variáveis de forma a “prever” a variável-alvo.

A figura 5.1 mostra uma tela típica do sistema, mostrando a árvore montada para a análise da variável “Açougue”.

**Figura 5.1. Árvore de decisão para a variável “Açougue”**



Fonte: Tela do programa Microsoft Analysis Manager

No canto superior direito temos uma visualização gráfica da árvore como um todo. Nota-se que a árvore possui dois níveis. No primeiro nível existem quatro ramificações, sendo que duas delas sofrem ainda mais duas ramificações cada. No painel central estas ramificações estão mais evidentes, mostrando toda a árvore. No painel central direito tem-se uma tabela com informações estatísticas de cada um dos nodos da árvore. No caso, o nodo selecionado é o último de cima para baixo. A coluna **value** traz um valor numérico, utilizado pelo algoritmo para criar as categorias em que os dados são divididos

(conforme já explanado anteriormente). A coluna **Cases** indica quantos casos do conjunto de dados utilizado foram direcionados para aquele nodo, enquanto que a coluna **Probability** mostra qual a probabilidade de que um novo registro seja classificado para aquele nodo dependendo do valor de outras variáveis.

No caso do nodo em questão, bastará que o novo registro tenha a variável **ValorTotal** > 1.718,14 para que exista uma probabilidade de 64,5% de que o valor das compras no setor de açougue esteja no último quartil (representado pelo valor 2.550,63, que é um valor no intervalo entre o terceiro quartil – Q3 e o valor máximo daquela variável). Esta informação está detalhada no painel inferior direito (**Node Path**), que mostra quais as escolhas que devem ser feitas nos dados para se localizar os casos que representam aquele nodo.

Assim, quando analisamos o terceiro nodo do segundo nível, temos a informação que ele representa os casos em que a variável **ValorTotal** apresenta valores entre 1.051,30 e 1.718,14 e que a variável **Laticínios** apresenta um valor inferior a 44,96. Neste caso, a probabilidade de que as compras no setor de açougue estejam no último quartil é de 49,66%.

Como se pode observar, existe uma relação entre as variáveis **Açougue** (que é o alvo desta análise em particular) e as variáveis **ValorTotal**, **Laticínios** e **Mercearia**. Ao se comparar estes resultados com a análise de correlação, verifica-se que estas variáveis já estavam correlacionadas. No entanto, apenas a correlação entre as variáveis **Açougue** e **ValorTotal** pode-se considerar “forte” ( $r=0,66495$ ), enquanto que entre **Açougue** e **Mercearia** ( $r=0,53297$ ) e **Laticínios** ( $r=0,38531$ ) a correlação é considerada “fraca”. Vale notar, além disso, que existem outras correlações mais fortes entre as variáveis que não

foram levadas em consideração pelo algoritmo de classificação na montagem da árvore.

**Quadro 5.5. Análise da classificação por árvore de decisão, primeiro estudo.**

Variável	Nº de níveis	Maior Probabilidade	Variáveis de que depende	Seleção da maior probabilidade
Açougue	2	64,50	ValorTotal, Mercearia, Limpeza	Valor Total > 1718,14
Bazar	3	62,93	ValorTotal, Limpeza	Limpeza > 193,62 e Valor Total > 1718,14
Bebidas	3	68,96	ValorTotal, Mercearia, Laticínios, Frios	Laticínios > 44,96 e Frios > 104,04 e Valor Total > 1718,14
Freqüência	2	13,25	ValorTotal	Valor Total > 615,23 e Valor Total <= 1718,14
Frios	3	74,78	ValorTotal, Laticínios	Laticínios > 148,06 e Valor Total > 1718,14
Higiene	3	72,67	ValorTotal, Limpeza	Limpeza > 193,62 e Valor Total > 1718,14
Hortifruti	2	69,14	ValorTotal, Laticínios, Açougue	Laticínios > 148,06 e Valor Total > 1718,14
Laticínios	3	84,02	ValorTotal, Frios, Mercearia, Limpeza	Frios > 104,04 and Valor Total > 1718,14
Limpeza	3	88,29	ValorTotal, Mercearia, Higiene	Higiene > 147,09 e Mercearia > 688,34 e Valor Total > 1718,14
Mercearia	3	58,70	ValorTotal, Limpeza, Laticínios	Limpeza > 42,24 e Limpeza <= 193,62 e Valor Total > 1718,14
Padaria	2	51,26	Frios, Hortifruti, ValorTotal	Frios > 104,04 e Valor Total > 1718,14
Recência	1	-	-	-
ValorTotal	4	97,67	Mercearia, Bebidas, Açougue, Limpeza, Laticínios, Hortifruti, Padaria	Limpeza > 193,62 e Bebidas > 206,3 e Mercearia > 688,34

Fonte: Dados de venda período outubro/2002 a maio/2003

Considerando que nosso estudo deseja identificar os caminhos em que os clientes tenham o maior consumo em todos os departamentos (ou seja, estariam no último quartil), chegamos às conclusões listadas no quadro 5.5 a seguir. Cópias das telas com as árvores completas estão listadas no Anexo 2.

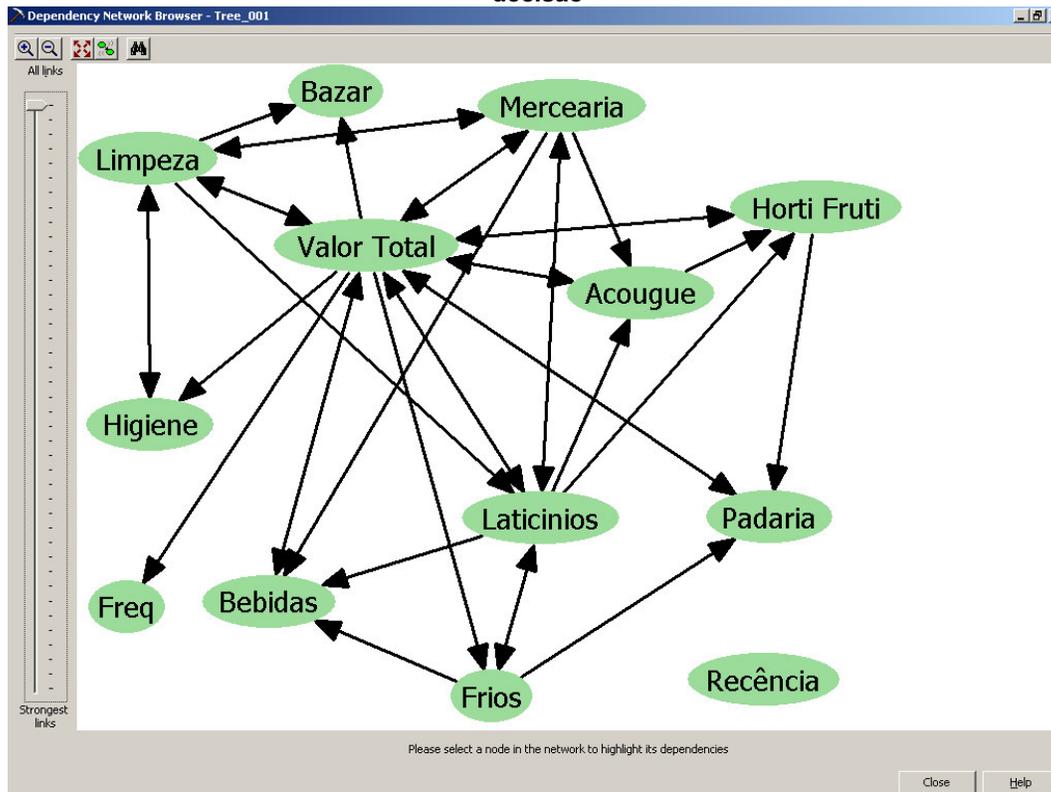
Ao se analisar o quadro 5.5 observa-se que a variável **Recência** é a única em que não se pode montar a árvore de decisão, pois os dados não estão correlacionados. Isto já era esperado, pois uma verificação ao quadro 5.4

mostra que esta variável possui alguns dos menores índices de correlação com as demais variáveis.

Este comportamento torna-se mais evidente ainda quando se analisa a chamada “Rede de Dependência”, gráfico derivado do processo de criação da árvore de decisão e que mostra como as variáveis se inter-relacionam. A rede é composta de esferas conectadas, representando as variáveis em estudo, onde a direção da seta indica a relação de dependência entre as variáveis. Na figura 5.2, pode-se observar que a variável **Recência** é a única que não possui nenhum tipo de relacionamento com as outras variáveis. Outra variável que chama a atenção nesta figura é a **Frequência** (representada pela esfera **Freq**), que possui somente uma relação, ainda assim de dependência, com a variável **ValorTotal** que, por sua vez, é a variável que mais influencia (e também sofre influência) as demais. O quadro 5.6 mostra como estas variáveis se inter-relacionam.

Um outro estudo de mineração foi realizado, apenas com as variáveis que representam os valores despendidos nos vários departamentos. Assim, as variáveis **ValorTotal**, **Frequência** e **Recência** não foram utilizadas.

Figura 5.2. Rede de dependência do primeiro estudo de classificação por árvore de decisão



Fonte: Tela do programa Microsoft Analysis Manager

Quadro 5.6. Inter-relacionamento entre as variáveis do primeiro estudo de classificação por árvore de decisão

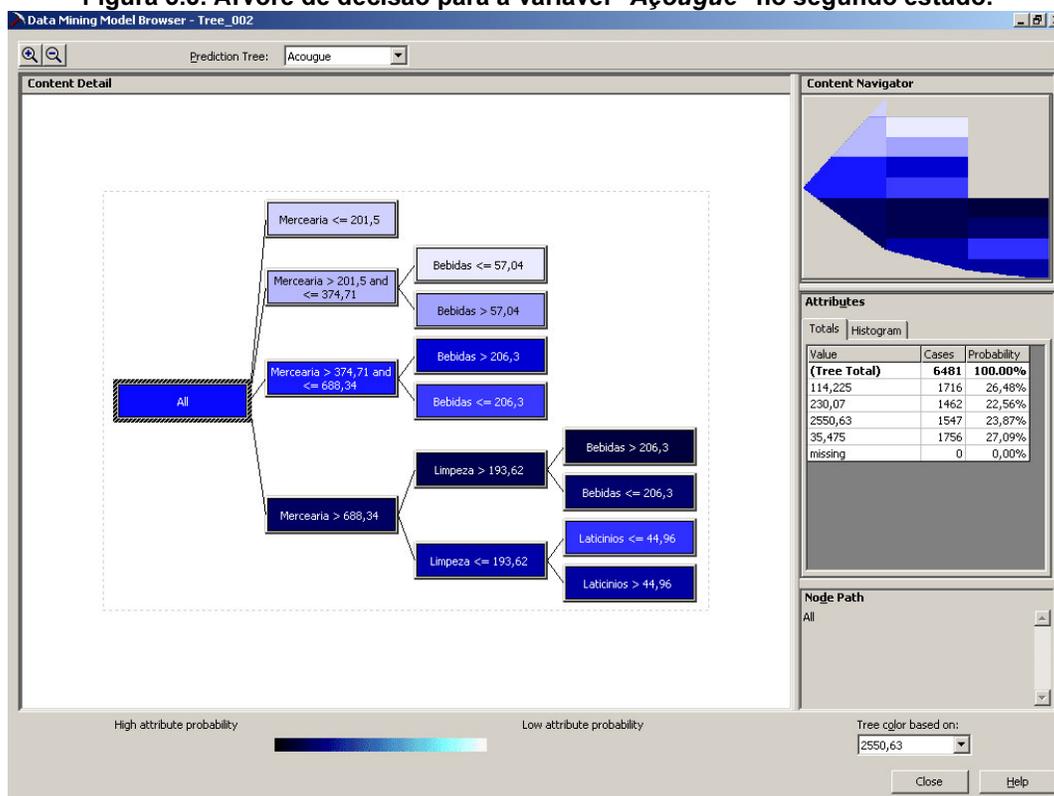
Variável	Qual variável influencia	É influenciada por
Açougue	Hortifruiti, ValorTotal	Laticínios, mercearia, ValorTotal
Bazar	-	Limpeza, ValorTotal
Bebidas	ValorTotal	Frios, Laticínios, Mercearia, ValorTotal
Freqüência	-	ValorTotal
Frios	Bebidas, Laticínios, Padaria	Laticínios, ValorTotal
Higiene	Limpeza	Limpeza, ValorTotal
Hortifruiti	Padaria, ValorTotal	Açougue, Laticínios, ValorTotal
Laticínios	Açougue, Bebidas, Frios, Hortifruiti, Mercearia, ValorTotal	Frios, Limpeza, Mercearia, ValorTotal
Limpeza	Bazar, Higiene, Laticínios, Mercearia, ValorTotal	Higiene, Mercearia, ValorTotal
Mercearia	Açougue, Bebidas, Laticínios, Limpeza, ValorTotal	Laticínios, Limpeza, ValorTotal
Padaria	ValorTotal	Frios, Hortifruiti, ValorTotal
Recência	-	-
ValorTotal	Todas, exceto recência	Açougue, Bebidas, Hortifruiti, Laticínios, Limpeza, Mercearia, Padaria

Fonte: Dados de venda, período outubro/2002 a maio/2003

Este segundo estudo tem por objetivo verificar como as compras efetuadas em um determinado departamento influenciam na decisão de compra de produtos de outros departamentos.

Como esperado, as árvores montadas para cada departamento foram bem distintas, conforme mostra a figura 5.3 que mostra a árvore para a variável **Açougue**. Uma rápida comparação com a figura 5.1 já mostra claramente a diferença.

**Figura 5.3. Árvore de decisão para a variável “Açougue” no segundo estudo.**



Fonte: Tela do programa Microsoft Analysis Manager

O quadro 5.7, a seguir, mostra as conclusões das análises deste segundo estudo. Como no primeiro, o objetivo é reconhecer quais os caminhos para se chegar ao valor máximo de vendas de cada departamento. Como se pode observar em comparação ao quadro 5.5, existe um maior grau de

interdependência entre as variáveis do estudo pois a variável **ValorTotal** é prevalente no estudo anterior. Vê-se, inclusive, que existe um aumento no número de níveis das árvores geradas.

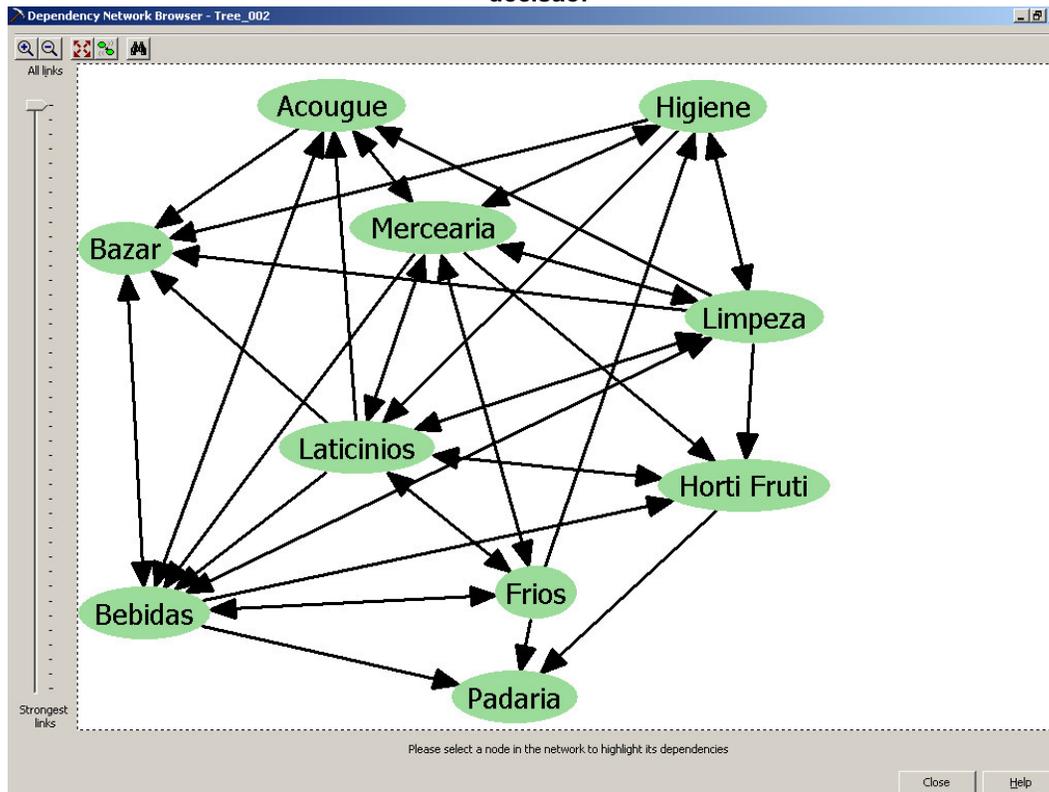
**Quadro 5.7. Análise da classificação por árvore de decisão, primeiro estudo.**

Variável	Nº de níveis	Maior Probabilidade	Variáveis de que depende	Seleção da maior probabilidade
Açougue	3	70,78	Mercearia, Limpeza, Bebidas, Laticínios	Limpeza > 193,62 e Bebidas > 206,3 e Mercearia > 688,34
Bazar	4	68,27	Limpeza, Higiene, Bebidas, Laticínios, Açougue	Limpeza > 193,62 e Laticínios > 148,06 e Higiene > 147,09
Bebidas	4	68,34	Mercearia, Laticínios, Frios, Limpeza, Bazar, Açougue	Laticínios > 148,06 e Frios > 104,04 e Mercearia > 688,34
Frios	3	79,97	Laticínios, Mercearia, Bebidas	Laticínios > 148,06 e Bebidas > 206,3 e Mercearia > 688,34
Higiene	3	78,87	Limpeza, Mercearia, Frios	Limpeza > 193,62 e Frios > 104,04 e Mercearia > 688,34
Hortifruti	3	75,12	Laticínios, Mercearia, Bebidas, Limpeza	Limpeza > 193,62 e Laticínios > 148,06 e Bebidas > 206,3
Laticínios	4	87,62	Frios, Hortifruti, Mercearia, Limpeza	Limpeza > 193,62 e Frios > 104,04 e Mercearia > 688,34
Limpeza	3	89,85	Mercearia, Higiene, Laticínios, Bebidas	Higiene > 147,09 e Mercearia > 688,34 e Valor Total > 1718,14
Mercearia	4	96,79	Limpeza, Laticínios, Açougue, Higiene, Frios	Limpeza > 193,62 e Laticínios > 148,06 e Higiene > 147,09 e Açougue > 302,64
Padaria	3	53,41	Frios, Hortifruti, Bebidas	Hortifruti > 111,94 e Frios > 104,04 e Bebidas > 57,04

Fonte: Dados de venda período outubro/2002 a maio/2003

A rede de dependências desta segunda análise (figura 5.4) mostra este novo grau de inter-relacionamento entre as variáveis. A rede está resumida no quadro 5.8, a seguir. Como se pode observar, a ausência da variável **ValorTotal** fez com que as outras variáveis mostrassem a sua força na determinação do padrão de compra.

**Figura 5.4. Rede de dependências do segundo estudo de classificação por árvore de decisão.**



Fonte: Tela do programa Microsoft Analysis Manager

**Quadro 5.8. Inter-relacionamento entre as variáveis do segundo estudo de classificação por árvore de decisão**

Variável	Qual variável influencia	É influenciada por
Açougue	Bazar, Bebidas, Mercearia	Bebidas, Mercearia, Laticínios, Limpeza
Bazar	Bebidas	Açougue, Bebidas, Higiene, Laticínios, Limpeza
Bebidas	Açougue, Bazar, Frios, Hortifruiti, Limpeza, Padaria	Açougue, Bazar, Frios, Limpeza, Laticínios, Mercearia
Frios	Bebidas, Higiene, Laticínios, Mercearia, Padaria	Bebidas, Laticínios, Mercearia
Higiene	Bazar, Laticínios, Limpeza, Mercearia	Frios, Limpeza, Mercearia
Hortifruiti	Laticínios, Padaria	Bebidas, Laticínios, Limpeza, Mercearia
Laticínios	Açougue, Bazar, Bebidas, Frios, Hortifruiti, Limpeza, Mercearia	Frios, Limpeza, Higiene, Hortifruiti, Mercearia
Limpeza	Açougue, Bazar, Bebidas, Higiene, Hortifruiti, Laticínios, Mercearia	Bebidas, Higiene, Laticínios, Mercearia
Mercearia	Açougue, Bebidas, Frios, Higiene, Laticínios, Limpeza	Açougue, Frios, Higiene, Laticínios, Limpeza
Padaria	-	Bebidas, Frios, Hortifruiti

Fonte: Dados de venda, período outubro/2002 a maio/2003

### 5.7.2. Agrupamentos

Outra forma de se analisar os dados apresentados é separar os registros em grupos com características em comum. Para tanto, a maioria das ferramentas de mineração de dados provêm algoritmos de criação destes agrupamentos (ou *clusters*).

Estes algoritmos utilizam técnicas de refinamentos sucessivos para agrupar os casos em “vizinhanças” que exibam características similares e previsíveis. Estas “vizinhanças” são formadas, normalmente, através da verificação da distância de cada caso até o “ponto central” (centróide) de cada grupo, que é calculado pelo algoritmo conforme as técnicas utilizadas pelo mesmo (algoritmos diferentes tendem a ter diferentes formas de cálculos da distância e do centróide). Os casos então são agrupados conforme esta distância até a centróide mais próxima, o que determina em que grupo cada caso será colocado.

Alguns fatores influenciam de forma decisiva a forma como os dados são agrupados. O principal fator é a quantidade de grupos que se deseja formar. Um pequeno número de grupos pode não mostrar detalhes importantes, pois tendem a ocorrer grandes distâncias entre os casos e o centróide. Por outro lado, um grande número de grupos tende a dispersar, em dois ou mais grupos, casos com características minimamente diferenciadas. Como não existe regra geral que especifique quantos grupos devem ser formados, somente a experimentação poderá mostrar qual o número ideal de agrupamentos.

Para o estudo em curso, e a exemplo da análise de classificação por árvore de decisão, foi utilizado o programa Microsoft Analysis Manager, manipulando

a mesma tabela das análises anteriores (**ComprasCli**), que nada mais é do que um resumo do que cada cliente gastou em cada departamento no período em estudo (outubro de 2002 a maio de 2003).

Uma primeira análise foi realizada utilizando-se todas as variáveis contidas na tabela, à exceção da variável que informa o código do cliente. Após algumas experimentações, concluiu-se que o ideal é se formar 06 (seis) grupos distintos. O resultado dos agrupamentos realizados está mostrado no quadro 5.9, colocado em forma de matriz, onde cada coluna representa um agrupamento e os valores contidos são as médias dos valores das variáveis correspondentes, dentro de cada grupo.

**Quadro 5.9. Resultado do primeiro estudo de agrupamentos**

	<b>G1</b>	<b>G2</b>	<b>G3</b>	<b>G4</b>	<b>G5</b>	<b>G6</b>
<b>Açougue</b>	135,68	78,51	181,61	376,87	266,63	793,36
<b>Bazar</b>	18,20	13,52	32,46	64,31	55,96	148,02
<b>Bebidas</b>	109,40	56,65	142,78	287,63	179,78	661,59
<b>Freqüência</b>	12,04	9,67	17,00	21,97	12,20	30,23
<b>Frios</b>	43,96	24,52	81,17	140,78	60,70	265,59
<b>Higiene</b>	52,99	27,39	89,11	200,27	128,22	328,13
<b>Hortifruti</b>	54,95	28,19	91,88	171,56	74,15	317,07
<b>Laticínios</b>	69,19	39,06	121,98	217,22	92,41	392,93
<b>Limpeza</b>	71,67	32,67	122,95	279,55	148,73	464,88
<b>Mercearia</b>	288,74	147,60	489,86	969,58	494,25	1605,77
<b>Padaria</b>	16,00	13,06	37,50	40,77	16,65	789,99
<b>Recência</b>	20,38	34,29	11,31	9,40	38,63	13,82
<b>ValorTotal</b>	774,86	427,12	1.230,87	2.416,55	1.383,69	4.452,92
<b>Qtd. Casos</b>	<b>1.515</b>	<b>1.394</b>	<b>1.178</b>	<b>1.044</b>	<b>970</b>	<b>380</b>

Fonte: Dados de vendas, período outubro/2002 a maio/2003

Como se pode observar, cada grupo possui diferentes quantidades de casos. Chama a atenção o grupo **G6** pelo fato de ele conter os clientes com as maiores médias (com exceção da variável **Recência**). Este grupo contém, com certeza, os melhores clientes da rede. Ao contrário, o grupo **G2** possui as menores médias e o maior valor médio de **Recência**, ou seja, são aqueles clientes eventuais, que não fazem visitas regulares às lojas e com baixo valor médio da compra como um todo.

Os demais grupos destacam uma ou mais variável em relação às outras. Como um exemplo, no grupo **G1** a variável **Bebidas** tem mais destaque (dentro do grupo) do que nos demais. Este destaque pode ser também pelo baixo peso de determinada variável no grupo, como é o caso da variável **Padaria** no grupo **G5**. O quadro 5.10 mostra estes destaques, marcados como positivo (+), para as variáveis de maior peso, ou negativo (-) para aquelas de menor peso:

**Quadro 5.10. Variáveis em destaque no primeiro estudo de agrupamentos**

	<b>G1</b>	<b>G2</b>	<b>G3</b>	<b>G4</b>	<b>G5</b>	<b>G6</b>
<b>Açougue</b>			-		+	
<b>Bazar</b>	-				+	
<b>Bebidas</b>			-			+
<b>Frios</b>			+		-	
<b>Higiene</b>		-			+	
<b>Hortifruti</b>			+		-	
<b>Laticínios</b>			+		-	
<b>Limpeza</b>		-		+		
<b>Mercearia</b>				+		-
<b>Padaria</b>		+			-	

Fonte: Dados de venda, período outubro/2002 a maio/2003

O quadro 5.11 mostra os percentuais de cada variável dentro de cada grupo:

**Quadro 5.11. Percentual de cada variável, por grupo, no primeiro estudo de agrupamentos**

	<b>G1</b>	<b>G2</b>	<b>G3</b>	<b>G4</b>	<b>G5</b>	<b>G6</b>
<b>Açougue</b>	15,76%	17,02%	13,05%	13,71%	17,57%	15,69%
<b>Bazar</b>	2,11%	2,93%	2,33%	2,34%	3,69%	2,93%
<b>Bebidas</b>	12,71%	12,28%	10,26%	10,46%	11,85%	13,08%
<b>Frios</b>	5,11%	5,32%	5,83%	5,12%	4,00%	5,25%
<b>Higiene</b>	6,16%	5,94%	6,40%	7,29%	8,45%	6,49%
<b>Hortifruti</b>	6,38%	6,11%	6,60%	6,24%	4,89%	6,27%
<b>Laticínios</b>	8,04%	8,47%	8,77%	7,90%	6,09%	7,77%
<b>Limpeza</b>	8,33%	7,08%	8,84	10,17%	9,80%	9,19%
<b>Mercearia</b>	33,54%	32,01%	35,21%	35,28%	32,57%	31,76%
<b>Padaria</b>	1,86%	2,83%	2,70	1,48%	1,10%	1,56%

Fonte: Dados de venda, período outubro/2002 a maio/2003

Como se pode observar, apesar de ser o grupo com maiores médias, **G6** apresenta o menor peso da variável **Mercearia**. Isto demonstra que este grupo, além das maiores médias, possui uma maior homogeneidade nas compras de

todos os departamentos. Já o grupo **G5** pode ser considerado o grupo dos extremos, pois possui 03 (três) variáveis com destaque positivo e outras 04 (quatro) com destaque negativo.

No segundo estudo, apenas as variáveis referentes às compras foram utilizadas, ficando de fora as variáveis **ValorTotal**, **Freqüência** e **Recência**. Para efeitos de comparação, este estudo também faz a divisão em 06 (seis) grupos distintos. O quadro 5.12 aponta o resultado deste estudo. Como no estudo anterior, é mostrado o valor médio de compra por variável dentro de cada agrupamento:

**Quadro 5.12. Resultado do segundo estudo de agrupamentos**

	<b>G1</b>	<b>G2</b>	<b>G3</b>	<b>G4</b>	<b>G5</b>	<b>G6</b>
<b>Açougue</b>	214,37	114,05	89,34	380,10	231,57	835,06
<b>Bazar</b>	38,70	15,43	9,44	71,66	43,26	159,91
<b>Bebidas</b>	157,25	138,72	56,97	304,66	96,23	713,22
<b>Frios</b>	86,10	47,86	24,38	139,47	45,62	278,02
<b>Higiene</b>	132,12	60,13	25,52	198,13	57,91	342,00
<b>Hortifruti</b>	96,84	62,08	28,16	171,96	53,08	330,98
<b>Laticínios</b>	132,88	77,11	39,54	215,06	63,44	404,81
<b>Limpeza</b>	165,12	84,71	32,41	279,37	63,42	490,86
<b>Mercearia</b>	594,29	322,17	154,96	949,12	285,74	1.674,55
<b>Padaria</b>	25,47	12,63	12,30	44,71	32,18	80,77
<b>Qtd. Casos</b>	<b>1.379</b>	<b>1.369</b>	<b>1.332</b>	<b>1.089</b>	<b>991</b>	<b>321</b>

Fonte: Dados de vendas, período outubro/2002 a maio/2003

A exemplo do primeiro estudo, o grupo **G6** concentra os casos com maiores médias de compra. Entretanto, o grupo com menores médias passou a ser o **G3**, ao invés do **G2**, como no estudo anterior.

Vale ressaltar que estes grupos são diferentes daqueles formados no primeiro estudo, como indica a quantidade de casos em cada um dos agrupamentos. Isto demonstra como a inclusão ou exclusão de variáveis pode alterar completamente os resultados de cada análise, fazendo com que alguns casos migrem de um agrupamento para outro.

O quadro 5.13 mostra os destaques positivos e negativos deste novo estudo, a exemplo do anterior:

**Quadro 5.13. Variáveis em destaque no segundo estudo de agrupamentos**

	G1	G2	G3	G4	G5	G6
<b>Açougue</b>		-			+	
<b>Bazar</b>		-			+	
<b>Bebidas</b>	-	+				
<b>Frios</b>	+				-	+
<b>Higiene</b>	+		-			
<b>Hortifruti</b>		+			-	
<b>Laticínios</b>			+		-	
<b>Limpeza</b>				+	-	
<b>Mercearia</b>	+				-	
<b>Padaria</b>		-			+	

Fonte: Dados de venda, período outubro/2002 a maio/2003

Como podemos observar, o quadro apresenta-se distinto daquele apresentado no primeiro estudo. Dois pontos chamam a atenção para o grupo **G5**. Primeiro, ele mantém a característica de ser o grupo dos extremos, com 03 (três) destaques positivos e 05 (cinco) destaques negativos. Segundo, a reviravolta da variável **Padaria**, que passou de destaque negativo (no primeiro estudo conforme quadro 5.10) para destaque positivo dentro do grupo.

O quadro 5.14 mostra os percentuais de cada variável dentro de cada agrupamento, base para a análise do quadro 5.13:

**Quadro 5.14. Percentual de cada variável, por grupo, no segundo estudo de agrupamentos**

	G1	G2	G3	G4	G5	G6
<b>Açougue</b>	13,05%	12,20%	18,89%	13,80%	23,81%	15,73%
<b>Bazar</b>	2,36%	1,65%	2,00%	2,60%	4,45%	3,01%
<b>Bebidas</b>	9,57%	14,84%	12,04%	11,06%	9,90%	13,43%
<b>Frios</b>	5,24%	5,12%	5,15%	5,06%	4,69%	5,24%
<b>Higiene</b>	8,04%	6,43%	5,40%	7,19%	5,95%	6,44%
<b>Hortifruti</b>	5,89%	6,64%	5,95%	6,24%	5,46%	6,23%
<b>Laticínios</b>	8,09%	8,25%	8,36%	7,81%	6,52%	7,62%
<b>Limpeza</b>	10,05%	9,06%	6,85%	10,14%	6,52%	9,24%
<b>Mercearia</b>	36,17%	34,46%	32,76%	34,46%	29,38%	31,53%
<b>Padaria</b>	1,55%	1,35%	2,60%	1,62%	3,31%	1,52%

Fonte: Dados de venda, período outubro/2002 a maio/2003

### 5.7.3. Regras de associação

O terceiro método de análise por mineração de dados é denominado “regras de associação”. Este método analisa os dados e verifica as possíveis associações entre as variáveis em estudo, baseado em um suporte estatístico.

Para a realização do estudo, utilizou-se o programa Waikato Environment for Knowledge Analysis (WEKA), desenvolvido pela Universidade de Waikato, Nova Zelândia. Este sistema, totalmente desenvolvido em JAVA e de uso gratuito, congrega vários algoritmos de mineração de dados, inclusive de classificação por árvore de decisão e de agrupamentos.

Por uma imposição do algoritmo utilizado, todas as variáveis precisaram ser divididas em classes, pois o mesmo só trabalha com variáveis discretas. Para seguir o mesmo padrão dos estudos anteriores, em que os dados eram automaticamente divididos em quartis pelo programa, foi realizada a mesma divisão. O primeiro passo, então, foi a definição dos limites das classes baseados nos quartis, conforme o quadro 5.15:

**Quadro 5.15. Valores dos quartis das variáveis, primeiro estudo de regras de associação**

	Min	Q1	Q2	Q3	Max
<b>Açougue</b>	0,00	65,17	142,90	293,28	6.376,86
<b>Bazar</b>	0,00	7,05	21,42	51,93	1.007,92
<b>Bebidas</b>	0,00	53,05	110,81	209,36	20.684,33
<b>Frios</b>	0,00	25,17	51,87	97,20	2.309,88
<b>Higiene</b>	0,00	27,15	67,18	136,31	2.135,77
<b>Hortifruti</b>	0,00	24,50	58,26	116,19	2.094,59
<b>Laticínios</b>	0,00	44,93	84,07	148,06	1.729,66
<b>Limpeza</b>	0,00	40,43	92,23	186,37	1.963,40
<b>Mercearia</b>	0,00	200,84	374,60	660,11	6874,26
<b>Padaria</b>	0,00	6,63	16,14	34,32	655,71
<b>Freqüência</b>	8	9	12	18	121
<b>Recência</b>	1	5	12	29	221
<b>ValorTotal</b>	73,83	613,30	1.017,58	1.681,93	21.574,12

Fonte: Dados de venda, período outubro/2002 a maio/2003

A partir deste quadro, criou-se uma segunda tabela de dados, denominada

**CompraCliClasse**, onde os registros contém um número inteiro, entre 1 e 4,

indicando à qual quartil o valor original pertencia (primeiro, segundo, terceiro ou quarto). A partir desta tabela foi gerado o arquivo texto que serviu de entrada para o algoritmo utilizado.

O algoritmo foi configurado para gerar um máximo de 100 regras, com um suporte mínimo de 70% de confiança. Com isto, foram geradas as regras listadas no quadro 5.16, todas no formato “x → y” (x implica em y):

**Quadro 5.16. Regras geradas no primeiro estudo com Regras de Associação**

Regra (x)	Casos	Implica (y)	Casos	Confiança
Limpeza=4 Mercearia=4	1192	ValorTotal=4	1094	0,92
Laticínios=4 Mercearia=4	1085	ValorTotal=4	993	0,92
ValorTotal=4 Limpeza=4	1214	Mercearia=4	1094	0,90
ValorTotal=4 Laticínios=4	1136	Mercearia=4	993	0,87
ValorTotal=4 Mercearia=4	1323	Limpeza=4	1094	0,83
ValorTotal=4 Mercearia=4	1621	Mercearia=4	1323	0,82
Mercearia=4	1621	ValorTotal=4	1323	0,82
ValorTotal=1	1620	Mercearia=1	1222	0,75
Mercearia=1	1620	ValorTotal=1	1222	0,75
ValorTotal=4 Mercearia=4	1323	Laticínios=4	993	0,75
ValorTotal=4 Limpeza=4	1621	Limpeza=4	1214	0,75
Limpeza=4	1621	ValorTotal=4	1214	0,75
Limpeza=4	1621	Mercearia=4	1192	0,74
Mercearia=4	1621	Limpeza=4	1192	0,74
ValorTotal=4	1621	Laticínios=4	1136	0,70
Laticínios=4	1621	ValorTotal=4	1136	0,70

Fonte: Dados de venda, período outubro/2002 a maio/2003

Neste quadro, a primeira coluna mostra as regras geradas, a segunda coluna mostra quantos casos satisfazem a regra, a terceira coluna é a consequência da regra, a quarta coluna traz o número de casos em que esta consequência satisfaz as exigências da regra e a quinta coluna traz o percentual de confiança, obtido pela divisão entre a quarta e a segunda coluna. Esta última coluna é que mostra quão forte uma regra pode ser. Assim, a primeira regra do quadro pode ser lida como: “SE Limpeza = 4 e Mercearia = 4,

ENTÃO ValorTotal = 4”, com uma confiança de 92%. Uma outra característica que emerge são as regras de dupla implicação, ou seja, “x  $\leftrightarrow$  y”, como nos casos das regras (com 82% de confiança):

“ValorTotal = 4  $\rightarrow$  Mercearia = 4” e

“Mercearia = 4  $\rightarrow$  ValorTotal = 4”

Estas regras indicam que grandes valores na variável **Mercearia** implicam em altos valores na variável **ValorTotal** e vice-versa.

É interessante notar também que as regras mostram padrões nos dois extremos, tanto para grandes valores de compra como para os pequenos valores, como mostram as regras (com 75% de confiança):

“ValorTotal = 1  $\rightarrow$  Mercearia = 1” e

“Mercearia = 1  $\rightarrow$  ValorTotal = 1”

Estas regras demonstram que pequenos valores na variável **Mercearia** implicam em pequenos valores na variável **ValorTotal**.

Algumas outras conclusões podem ser tiradas em uma análise rápida do quadro 5.16, tais como a prevalência da força da variável **ValorTotal** em determinar o comportamento das demais variáveis. Além disso, como visto no estudo com árvores de decisão, as variáveis **Laticínios**, **Limpeza** e **Mercearia** também possuem muita força na determinação do comportamento das demais. Tanto que apenas estas quatro variáveis foram utilizadas para gerar as regras acima, mesmo que todas as outras também tenham participado do estudo.

Da mesma forma que nas análises anteriores, um segundo estudo foi levado a efeito eliminando as variáveis **ValorTotal**, **Frequência** e **Recência**, com o objetivo de determinar como um departamento influi em outro. Este

estudo foi executado com a mesma configuração do anterior, isto é, com um suporte mínimo de 70%.

Ao se executar o estudo, no entanto, apenas duas regras foram alcançadas, como mostra o quadro 5.17:

**Quadro 5.17. Regras geradas no segundo estudo com Regras de Associação**

Regra (x)	Casos	Implica (y)	Casos	Confiança
Limpeza = 4	1621	Mercearia = 4	1192	0,74
Mercearia = 4	1621	Limpeza = 4	1192	0,74

Fonte: Dados de venda, período outubro/2002 a maio/2003

Como se pode observar, estas duas regras mostram a força das duas variáveis descritas: **Mercearia** e **Limpeza**, mostrando claramente que os clientes costumam comprar produtos dos dois departamentos. Isto já foi demonstrado, inclusive, na análise estatística (quadro 5.2) onde as mesmas variáveis aparecem com alto índice de correlação ( $r=0,80488$ ).

Para aumentar a chance de aparecimento de outras regras, foi ajustado um parâmetro (**delta**), que permite um ajuste mais fino na forma como o algoritmo encontra as regras, mas que acaba por fazê-lo mais lento. O algoritmo parte da análise de regras que tenham 100% de confiança. Após cada iteração, este percentual de confiança é diminuído do valor do parâmetro **delta** e nova iteração ocorre à procura de regras que satisfaçam o novo percentual. O valor padrão deste parâmetro é de 0,05, equivalente a 5%. Com o valor ajustado para 0,02 (equivalente a 2%), o sistema conseguiu gerar as regras mostradas no quadro 5.18, sendo que algumas possuem maior índice de confiança que a execução original (com **delta** assumindo o valor padrão de 0,05).

**Quadro 5.18. Regras geradas no segundo estudo com Regras de Associação e delta = 0,02**

<b>Regra (x)</b>	<b>Casos</b>	<b>Implica (y)</b>	<b>Casos</b>	<b>Confiança</b>
Laticínios=4 Limpeza=4	1025	Mercearia=4	891	0,87
Frios=4 Limpeza=4	933	Mercearia=4	809	0,87
Higiene=4 Mercearia=4	1035	Limpeza=4	894	0,86
Higiene=4 Limpeza=4	1065	Mercearia=4	894	0,84
Frios=4 Mercearia=4	1004	Laticínios=4	828	0,82
Laticínios=4 Mercearia=4	1085	Limpeza=4	891	0,82
Frios=4 Mercearia=4	1004	Limpeza=4	809	0,81
Frios=4 Laticínios=4	1065	Mercearia=4	828	0,78
Laticínios=4 Mercearia=4	1085	Frios=4	828	0,76
Limpeza=4 Mercearia=4	1192	Higiene=4	894	0,75
Limpeza=4 Mercearia=4	1192	Laticínios=4	891	0,75
Limpeza=4	1621	Mercearia=4	1192	0,74
Mercearia=4	1621	Limpeza=4	1192	0,74

Fonte: Dados de venda, período outubro/2002 a maio/2003

Analisando-se este quadro, pode-se verificar que a variável mais forte passa a ser **Mercearia**, exercendo uma forte influência nas demais. Está presente em todas as regras listadas no quadro 5.18, seja como regra, seja como implicante. Outras variáveis aparecem, neste estudo, com força: **Limpeza**, **Higiene**, **Frios** e **Laticínios**. Isto demonstra uma forte influência destes departamentos nas compras dos clientes analisados.

## 6. CONCLUSÃO E RECOMENDAÇÕES

### 6.1. Conclusão

O setor supermercadista é, hoje, um dos mais concorridos da economia brasileira. Toda e qualquer nova tecnologia que venha a auxiliar as empresas do setor a melhorarem seu faturamento e, por conseqüência, sua margem de lucro será bem-vinda.

Assim, a implantação de estratégias de fidelização de clientes se mostra vital para a capacidade de sobrevivência destas empresas. No entanto, o alto custo de implantação destas estratégias levam as empresas a postergarem sua implementação efetiva.

O problema inicial deste trabalho foi exatamente este: *“Como obter informações a respeito dos hábitos de consumo dos clientes sem a necessidade de se implantar uma estratégia completa de CRM?”* Após o seu desenvolvimento, pode-se dizer que é plenamente possível.

Pode-se afirmar, com segurança, que os objetivos traçados para a resolução deste problema foram plenamente alcançados. Tanto o objetivo geral quanto os objetivos específicos foram desenvolvidos à exaustão, de maneira a mostrar capacidade de se obter informações preciosas onde antes havia apenas dados brutos:

- O projeto do *data mart* responsável por armazenar os dados foi realizado, conforme mostrado nos capítulos 4 e 5;
- O fluxo de informações necessário para o processamento dos dados e seu posterior armazenamento no *data mart* foi desenvolvido no capítulo 5;

- A proposição de soluções e ferramentas de baixo custo para a implantação do projeto, além das ferramentas de análise, foram realizadas nos capítulos 4 e 5.

Além disso, a academia exige que projetos como este devam atender a alguns preceitos básicos para serem considerados válidos: a contribuição, a relevância e a viabilidade.

No que diz respeito à contribuição, este trabalho mostra uma metodologia simples e economicamente viável para descobrir, a partir de dados sub utilizados pelas empresas, informações a respeito de hábitos de consumo dos clientes atendidos por ela.

Quanto à relevância, mostra que estes dados podem se tornar armas poderosas para enfrentar os desafios enfrentados diariamente pelo setor. O baixo custo da metodologia apresentada faz com que ela possa ser utilizada por uma variada gama de empresas, a partir de pressupostos básicos.

E quanto à viabilidade, mostra que estas informações podem ser recuperadas a um baixo custo, tanto de implementação quanto de processamento, tornando-se uma vantagem competitiva em um setor que se caracteriza pela forte concorrência.

## **6.2. Recomendações**

Após a conclusão do trabalho, algumas recomendações são necessárias, tanto para a empresa em que o trabalho foi desenvolvido quanto para futuros acréscimos a ele:

- Deve-se ter consciência que o cadastro de clientes é um dos bens mais preciosos para a empresa que deseja implantar estratégias de

fidelização de clientes e CRM. Assim, é necessário que o mesmo esteja o mais atualizado e completo possível. É necessário que este trabalho seja constante, permitindo que as análises sócio-econômicas e psicográficas possam ser realizadas a contento;

- Ao *data mart* proposto, pode-se agregar outro com as informações de pagamento realizados pelos clientes, tornando possível um cruzamento de informações que permitam traçar um perfil mais detalhado do consumidor;
- Implementar formas de atualização automática, tanto do *data mart* quanto das tabelas utilizadas como base para as análises estatísticas e de mineração de dados;
- Desenvolver estratégias de consulta *ad hoc* ao *data mart* e às ferramentas de mineração de dados, para que os próprios usuários tenham acesso às informações sem a necessidade de auxílio do pessoal de informática da empresa.

Esses foram alguns dos aspectos que se destacaram durante a execução do presente estudo.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- ACNIELSEN, **Censo ACNielsen 2000**. Disponível em <http://www.acnielsen.com.br/varejo/indice.htm>. Acesso em 22/03/2003.
- BADIN, Neiva T. **Avaliação da Produtividade de Supermercados e seu Benchmarking**. 1997. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, UFSC, Florianópolis. Disponível em: <http://www.eps.ufsc.br/disserta98/neiva/cap2.htm>. Acesso em 21/03/2003.
- BERMAN, Barry; EVANS, Joel R. **Retail Management: a strategic approach**. Prentice Hall. Upper Saddle River. 1998. *apud* PARENTE, Juracy. Varejo no Brasil. Ed. Atlas. São Paulo. 2000.
- BERQUÓ, E., SOUZA, J. M., GOTLIEB, S. L. D. **Bioestatística**. Ed. Pedagógica e Universitária Ltda (EPU). São Paulo. 1981.
- BERRY, L.B., Relationship Marketing. In: BERRY, L.L., SHOSTACK, G.L., UPAH, G.D. **Emerging Perspectives in Services Marketing**. AMA. Chicago. 1983
- BERSON, A., SMITH, S., THEARLING, K. **Building Data Mining Applications for CRM**. McGraw-Hill. New York. 2000.
- BORGES, Admir R. **Marketing de Varejo: as estratégias adotadas pelos supermercados de vizinhança para conquistar e fidelizar clientes**. 2001. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, UFSC, Florianópolis.
- BRETZKE, Mirian. Marketing de Relacionamento e Competição em Tempo Real. Ed. Atlas. São Paulo. 2000.
- BRETZKE, Mirian. **O conceito de CRM viabilizando o Marketing de Relacionamento para competir em tempo real**. Disponível em <http://www.bretzke-marketing.com.br/artigos/crm1.htm>. Acessado em 29/04/2003
- BROWN, Stanley A. CRM – Uma ferramenta estratégica para o mundo e-business. Ed. Makron Books. São Paulo. 2001.
- BUCHANAN, R.W.T, GILLIES, C.S, **Value Managed Relationships: The Key to Customer Retention and Profitability**. European Management Journal, 1990. In: PAYNE, A., CHRISTOPHER, M., CLARK, M., PECK, H. Relationship Marketing for Competitive Advantage, pg. 247-253. Butterworth-Heinemann, Oxford, 1995.
- CARVALHO, Denise. **O Negócio da Internet que deu Certo**. Revista Veja. São Paulo, 19 mar. 2003. pp. 82-83
- COSTA, M.T.C., **Uma Arquitetura Baseada em Agentes para Suporte ao Ensino à Distância**. Tese. Programa de Pós-Graduação em Engenharia de Produção-UFSC. 1999.

- CRESPO, Antonio A., **Estatística Fácil**. 18ª edição. Ed. Saraiva. São Paulo. 2002.
- CYмбаUM, J. J., MÁLACCO, R.L.S. **Qualidade de Atendimento no Varejo: um Desafio de Recursos Humanos**. In: ANGELO, C.F., GIANGRANDE, V. Marketing de Relacionamento no Varejo. Ed. Atlas. São Paulo. 1999.
- DELUCA, Marcelo. Uma análise dos Supermercados da Grande Florianópolis. Ed. Insular. Florianópolis. 2003.
- DESOUZA, G., **Designing a Customer Retention Plan**. The Journal of Bussines Strategy, 1992. In: PAYNE, A., CHRISTOPHER, M., CLARK, M., PECK, H. Relationship Marketing for Competitive Advantage, pg. 293-299. Butterworth-Heinemann, Oxford, 1995.
- DORIA Fº, Ulysses. **Introdução à Bioestatística**. Negócio Editora. São Paulo. 2001.
- FERREIRA, Aurélio B. H. Disponível em: <http://www.uol.com.br/aurelio>. Acesso em 15/03/2003.
- GIANGRANDE, Vera. **Marketing de Relacionamento e o Papel do Ombudsman**. 1999. In: Angelo, C., GIANGRANDE, V., Marketing de Relacionamento no Varejo. Pg. 15-26, Ed. Atlas. São Paulo. 1999.
- GORDON, I.H. **Marketing de Relacionamentos**. Ed. Futura. São Paulo, 1999.
- GRÖNROOS, C. **Relationship Marketing Logic**. Asia-Australia Marketing Journal, Vol. 4, 1996.
- HOLOVKO, B.I., **Comitês de Clientes**. 1999. In: Angelo, C., GIANGRANDE, V., Marketing de Relacionamento no Varejo. Pg. 119-129, Ed. Atlas. São Paulo. 1999.
- INMON, W. H. **Como Construir um Data Warehouse**. Ed. Campus. São Paulo. 1997.
- JAMIL, G.L. **Repensando a TI na Empresa Moderna**. Ed. Axcel Books. São Paulo. 2001.
- JOHNSTON, R., **Linking Complaint Management to Profit**. International Journal of Service Industry Management. Vol. 12, Nº 02. 2001.
- KASPER, J. F. de Pireno. **Estratégias de marketing em supermercados**. Rigel. Porto Alegre. 1981.
- KIMBALL, R., REEVES, L, ROSS, M., THORNTHWAITE, W. **The Data Warehouse Lifecycle Toolkit**., John Wiley & Sons. New York. 1998.
- KIMBALL, Ralph. **Data Warehouse Toolkit**. Makron Books. São Paulo. 1996.
- KOTLER, Philip. Administração de Marketing: Análise, planejamento, implementação e controle. 5ª Edição. Ed. Atlas. São Paulo. 1998.
- KUTWAK, Luiz. **Shopping Center**. 29/06/2001. Disponível em: <http://www.estacio.br/scripts/direito/cadernojuridico/artigos.asp?codigo=82>. Acesso em 21/03/2003.

- LEVY e WEITZ. **Administração de Varejo**. Ed. Atlas. São Paulo. 2000.
- MAISTER, D., **Marketing to Existing Clients**. Journal of Management Consultancy, 1989. In: PAYNE, A., CHRISTOPHER, M., CLARK, M., PECK, H. Relationship Marketing for Competitive Advantage, pg. 253-263. Butterworth-Heinemann, Oxford, 1995.
- McKENNA, R. **Marketing de Relacionamentos**. Estratégias Bem-Sucedidas para a Era do Cliente. Ed. Campus. Rio de Janeiro. 1992.
- MERCADOR. **Anuário Brasileiro de Supermercados**. Mercador S/A. São Paulo. 2002.
- MORALES, A.B.T, RODRIGUES, A.M., GONÇALVES, A.L. **Data Mining**. Disponível em: <http://www.stela.ufsc.br/disciplinas/datamining/kdd/pagdm-introducao-dm.htm>. Acessado em 25/09/2001.
- PARENTE, Juracy. **Varejo no Brasil**. Ed. Atlas. São Paulo. 2000.
- PAYNE, A., CHRISTOPHER, M., BALLANTYNE, D. Relationship Marketing: bringing quality, customer service and marketing together. Butterworth-Heinemann. Oxford. 1991.
- PECK, H., PAYNE, A., CHRISTOPHER, M., CLARK, M. **Relationship Marketing - Strategy and Implementation**. Butterworth-Heinemann, Oxford, 1999.
- PEPPERS, Don, ROGERS, Martha. **CRM series – marketing 1 to 1**. 2ª edição. Peppers & Rogers Group. São Paulo. 2003.
- REICHHELD, F.F., **Loyalty and the Renaissance of Marketing. Marketing Management**, 1994. In: PAYNE, A., CHRISTOPHER, M., CLARK, M., PECK, H. Relationship Marketing for Competitive Advantage, pg. 232-247. Butterworth-Heinemann, Oxford, 1995.
- REICHHELD, F.F., SASSER, W.E, **Zero Defections: Quality Comes to Services**. Harvard Business Review, 1990
- RICHERS, Raimar. **Marketing - Uma visão brasileira**. Negócio Editora. São Paulo. 2000.
- ROSEMBERG, L.J., CZEPIEL, J.A., **A Marketing Approach for Customer Retention**. Journal of Consumer Marketing, 1984. In: PAYNE, A., CHRISTOPHER, M., CLARK, M., PECK, H. Relationship Marketing for Competitive Advantage, pg. 269-277. Butterworth-Heinemann, Oxford, 1995.
- SEYBOLD, P., MARSHAK, T.R. **Cientes.com**. Ed. Makron Books. São Paulo, 2000.
- SM. **Supermercado Moderno**. Informa Publicações Especializadas. São Paulo. Janeiro/2003.
- STONE, M., WOODCOCK, N. **Marketing de Relacionamento**. Ed. Littera Mundi. São Paulo. 1998.
- STONE, M., WOODCOCK, N., MACHTYNGER, L. **CRM - Marketing de Relacionamento com os Clientes**. Ed. Futura. São Paulo. 2001.

SUPERHIPER. **Bom ano apesar da instabilidade.** Revista Super Hiper, Nº 320. São Paulo. Maio 2002.

SWIFT, Ronald. CRM – O revolucionário marketing de relacionamento com o cliente. Ed. Campus. São Paulo. 2001.

THOMAS, L., **Agents Must Commit to Customer Retention.** Disponível em: [www.iianny.org/fyiretention.htm](http://www.iianny.org/fyiretention.htm). Acessado em 26/09/2001.

TIME-LIFE. **A Era dos Reis Divinos: 3000-1500 a.C.**, Abril Livros. Rio de Janeiro. 1991.

VAVRA, T.G. Marketing de Relacionamento: Aftermarketing. Ed. Atlas. São Paulo. 1993.

## 8. ANEXOS

LAYOUT DE TABELAS E ARQUIVOS

<b>Arquivo:</b>	Cxddmmaa.ppp					
<b>Descrição:</b>	Log de Transações do PDV					
<b>Transação:</b>	TR 100 - Venda de Item					
<b>Cliente:</b>	Genérico					
<b>Módulo:</b>	<input type="checkbox"/> Monitor <input checked="" type="checkbox"/> Caixa					
<b>Formato:</b>	MS-DOS					
<b>Tamanho:</b>	152 bytes					
	Campo	Tipo	De	Até	Tamanho	
					Int	Dec
C	Data da abertura fiscal	N	001	006	006	
A	Data do sistema ou operação	N	007	012	006	
B	Hora do sistema ou operação	N	013	018	006	
Ç	Número do PDV ou caixa	N	019	021	003	
.	Tipo de registro (transação)	N	022	024	003	
	Número de seqüência ou cupom	N	025	030	006	
C	Código do supervisor liberador	N	031	036	006	
O	Código do operador liberador	N	037	042	006	
M	Ind. de operação executada off-line	N	043	043	001	
U	Número da loja	N	044	047	004	
M	Livre ( 'XXX' )	C	048	050	003	
	Código do produto (EAN-13)	N	051	063	013	
	(P1) Preço unitário de tabela	N	064	072	007	002
	(P2) Preço de tabela menos desconto	N	073	081	007	002
	(Q) Quantidade vendida do produto	N	082	088	004	003
	Preço Total (QxP1 ou QxP2, se dif. 0)	N	089	102	012	002
	Código da alíquota de ICMS	N	103	104	002	
	Código da seção (família)	N	105	107	003	
	% redução da base de cálculo	N	108	111	002	002
	Livre (`.´)	C	112	144	033	
	Sequencial do registro	C	145	150	006	
	CR+LF (fixo)	C	151	152	002	

LAYOUT DE TABELAS E ARQUIVOS

<b>Arquivo:</b>	Cxddmmaa.ppp					
<b>Descrição:</b>	Log de Transações do PDV					
<b>Transação:</b>	TR 130 - Cancelamento de Item de Venda					
<b>Cliente:</b>	Genérico					
<b>Módulo:</b>	<input type="checkbox"/> Monitor <input checked="" type="checkbox"/> Caixa					
<b>Formato:</b>	MS-DOS					
<b>Tamanho:</b>	152 bytes					
	Campo	Tipo	De	Até	Tamanho	
					Int	Dec
C	Data da abertura fiscal	N	001	006	006	
A	Data do sistema ou operação	N	007	012	006	
B	Hora do sistema ou operação	N	013	018	006	
Ç	Número do PDV ou caixa	N	019	021	003	
.	Tipo de registro (transação)	N	022	024	003	
	Número de seqüência ou cupom	N	025	030	006	
C	Código do supervisor liberador	N	031	036	006	
O	Código do operador liberador	N	037	042	006	
M	Ind. de operação executada off-line	N	043	043	001	
U	Número da loja	N	044	047	004	
M	Livre ( 'XXX' )	C	048	050	003	
	Código do produto (EAN-13)	N	051	063	013	
	(P1) Preço unitário de tabela	N	064	072	007	002
	(P2) Preço de tabela menos desconto	N	073	081	007	002
	(Q) Quantidade vendida do produto	N	082	088	004	003
	Preço Total (QxP1 ou QxP2, se dif. 0)	N	089	102	012	002
	Código da alíquota de ICMS	N	103	104	002	
	Código da seção (família)	N	105	107	003	
	% redução da base de cálculo	N	108	111	002	002
	Livre (`.´)	C	112	144	033	
	Sequencial do registro	C	145	150	006	
	CR+LF (fixo)	C	151	152	002	

LAYOUT DE TABELAS E ARQUIVOS

<b>Arquivo:</b>	Cxddmmaa.ppp					
<b>Descrição:</b>	Log de Transações do PDV					
<b>Transação:</b>	TR 200 - Cancelamento da Venda					
<b>Cliente:</b>	Genérico					
<b>Módulo:</b>	<input type="checkbox"/> Monitor <input checked="" type="checkbox"/> Caixa					
<b>Formato:</b>	MS-DOS					
<b>Tamanho:</b>	152 bytes					
	Campo	Tipo	De	Até	Tamanho	
					Int	Dec
C	Data da abertura fiscal	N	001	006	006	
A	Data do sistema ou operação	N	007	012	006	
B	Hora do sistema ou operação	N	013	018	006	
Ç	Número do PDV ou caixa	N	019	021	003	
.	Tipo de registro (transação)	N	022	024	003	
	Número de seqüência ou cupom	N	025	030	006	
C	Código do supervisor liberador	N	031	036	006	
O	Código do operador liberador	N	037	042	006	
M	Ind. de operação executada off-line	N	043	043	001	
U	Número da loja	N	044	047	004	
M	Livre ( 'XXX' )	C	048	050	003	
	Total da venda cancelada	N	051	064	012	002
	Livre ( '.' )	C	065	144	080	
	Sequencial do registro	C	145	150	006	
	CR+LF (fixo)	C	151	152	002	

LAYOUT DE TABELAS E ARQUIVOS

<b>Arquivo:</b>	Cxddmmaa.ppp					
<b>Descrição:</b>	Log de Transações do PDV					
<b>Transação:</b>	TR 370 - Venda com cheques					
<b>Cliente:</b>	Genérico					
<b>Módulo:</b>	<input type="checkbox"/> Monitor <input checked="" type="checkbox"/> Caixa					
<b>Formato:</b>	MS-DOS					
<b>Tamanho:</b>	152 bytes					
	Campo	Tipo	De	Até	Tamanho	
					Int	Dec
C	Data da abertura fiscal	N	001	006	006	
A	Data do sistema ou operação	N	007	012	006	
B	Hora do sistema ou operação	N	013	018	006	
Ç	Número do PDV ou caixa	N	019	021	003	
.	Tipo de registro (transação)	N	022	024	003	
	Número de seqüência ou cupom	N	025	030	006	
C	Código do supervisor liberador	N	031	036	006	
O	Código do operador liberador	N	037	042	006	
M	Ind. de operação executada off-line	N	043	043	001	
U	Número da loja	N	044	047	004	
M	Livre ( 'XXX' )	C	048	050	003	
	Fixo ('00')	N	051	052	002	
	Tipo Pessoa ( F / J )	C	053	053	001	
	Banco	N	054	056	003	
	Agencia	N	057	061	005	
	Conta	N	062	071	010	
	Cheque	N	072	078	007	
	Fixo ('00000')	N	079	083	005	
	Valor	N	084	094	011	002
	Data	N	095	102	008	
	Condição de Recebimento	N	103	104	002	
	Autorizador	N	105	110	006	
	Código Cliente	N	111	125	015	
	Sequencial do registro	N	126	131	006	
	Fixo ('.')	C	132	150	019	
	CR+LF (fixo)	C	151	152	002	

