



**Karine de Almeida Karam**

**Regressão Logística: Um modelo de  
Risco de Cancelamento de Clientes**

**Dissertação de Mestrado**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Administração de Empresas do Departamento de Administração da PUC-Rio.

Orientador: Prof. Jorge Ferreira da Silva

Rio de Janeiro

Março de 2006

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



**Karine de Almeida Karan**

**Regressão Logística: Um modelo de  
Risco de Cancelamento de Clientes**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Administração de Empresas da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

**Prof. Jorge Ferreira da Silva**

Orientador

Departamento de Administração – PUC-Rio

**Prof. Marie Agnes Chauvel**

Departamento de Administração - PUC-Rio

**Prof. Leticia Moreira Casotti**

UFRJ

**Prof. João Pontes Nogueira**

Vice-Decano de Pós-Graduação do CCS

Rio de Janeiro, 06 de março de 2006

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, da autora e do orientador.

**Karine de Almeida Karam**

Graduou-se em Estatística pela UERJ.

Ficha Catalográfica

Karam, Karine de Almeida

Regressão logística: um modelo de risco de cancelamento de clientes / Karine de Almeida Karam ; orientador: Jorge Ferreira da Silva. – Rio de Janeiro : PUC, Departamento de Administração, 2006.

98 f. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Administração.

Inclui referências bibliográficas.

1. Administração – Teses. 2. Retenção. 3. Lealdade. 4. Satisfação. 5. Marketing de relacionamento. 6. Regressão logística. 7. Segmentação. I. Silva, Jorge Ferreira da. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Administração. III. Título.

CDD: 658

Para minha mãe, Maria José e minha irmã, Mônica pelo apoio e confiança.

Para Maurício, que me acompanhou neste ano de inteira dedicação.

Para os amigos, Flávia, Elisa e Sandro, pela inspiração.

Para Fontinha, pelos conselhos no momento certo.

Para Duda, pela orientação profissional.

## Agradecimentos

Ao meu orientador Professor Jorge Ferreira da Silva pelo estímulo e parceria na realização deste trabalho e pelo aprendizado durante todo o curso. Agradeço, principalmente, pelas dicas preciosas e inteligentes nos momentos mais difíceis da construção deste projeto.

À PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

A Sonia da Biblioteca da PUC, pela incansável dedicação nas pesquisas deste tema.

Aos meus colegas da PUC-Rio.

Aos professores que participaram da Comissão examinadora.

A todos os professores e funcionários do Departamento pelos ensinamentos e pela ajuda.

A todos os amigos e familiares que de uma forma ou de outra me estimularam ou ajudaram-me.

## Resumo

Karam, Karine de Almeida; da Silva, Jorge Ferreira; (Orientador). **Regressão Logística: Um Modelo de Risco de Cancelamento de Clientes.** Rio de Janeiro, 2006. 98p. Dissertação de Mestrado – Departamento de Administração de Empresas, Pontifícia Universidade Católica do Rio de Janeiro.

O tema central deste projeto é a retenção de clientes como estratégia competitiva para aumentar a lucratividade da empresa. O objetivo é desenvolver um modelo estatístico que relacione variáveis transacionais, demográficas e dados sobre o histórico de eventos com a probabilidade de cancelamento dos clientes assinantes de jornal e definir o perfil dos clientes com maior risco de desligamento. Em uma primeira etapa, este estudo fornece uma revisão teórica sobre lealdade, satisfação e marketing de relacionamento, a fim de buscar uma relação com a retenção de clientes. Em seguida, a revisão de literatura levantou as variáveis mais usadas na segmentação de clientes tais como: variáveis transacionais, geográficas, demográficas, psicográficas e comportamentais para definir o perfil dos clientes que cancelam e dos que não cancelam sua assinatura. Depois de construir um modelo teórico, a regressão logística foi utilizada como técnica estatística para desenvolver um modelo de previsão de cancelamento. Os resultados foram analisados com o auxílio do programa estatístico SPSS e conclui-se que o perfil do cliente que cancela a assinatura do jornal é o jovem de até 30 anos; com baixo nível sócio-demográfico; morador da baixada, subúrbio e outros estados que não o Rio de Janeiro; que tenha adquirido sua assinatura através do canal telemarketing ativo; com a assinatura da modalidade anual e forma de pagamento em boleto ou débito em conta corrente; clientes que adquiriram sua assinatura mais recentemente; que compreendem menos de 3 produtos da empresa e que não tenham feito reclamações através da central de atendimento. O modelo final de previsão de cancelamento contou com 11 variáveis e a tabela de classificação mostrou uma taxa de acerto geral de 75,3%. A última etapa apresenta algumas conclusões, implicações e sugestões para pesquisas futuras.

## Palavras-chave

Retenção, lealdade, satisfação, marketing de relacionamento, regressão logística, segmentação e cancelamento.

## Abstract

Karam, Karine de Almeida; da Silva, Jorge Ferreira; (Advisor). **Logistic Regression: A Model to Measure Signature's Cancellation Risk**. Rio de Janeiro, 2006. 98p. M.Sc. Dissertation – Departamento de Administração de Empresas, Pontifícia Universidade Católica do Rio de Janeiro.

The core subject of this project is the customers' retention as a competitive strategy to increase the company's profitability. The goal is to develop a statistical model that links transactional and demographic variables and customer's history data with the subscribers' churn of a certain publication. In the first part, this study provides a revision on loyalty, satisfaction and relationship marketing theory in order to find a relation with customers' retention. After that, the literature revision raised the most used variables for the segmentation of customers, such as: transactional, geographic, demographic, psychological and behavior variables to define the profile of the customer who churns and the profile of that one who doesn't. After constructing a theoretical model, the logistic regression was used as a statistical technique to develop a model of cancellation forecasting. The results has been analyzed with the aid of statistical program SPSS and conclude that the profile of the customer who cancels the subscription of the publication is young up to 30 years old; with low social-demographic level; living at Baixada, Suburb, and other states than Rio De Janeiro; that bought the subscription through the outbound telemarketing sales channel; with one year subscription and payment through invoice or direct debit in current account; customers who has bought its signature more recently; that do not buy less than 3 other products of the company and that have not made complaints through the customer service. The final model of churn forecasting uses 11 variables and the classification table showed an accuracy of 75,3%. The last part presents some conclusions, implications and suggestions for future research.

## Keywords

Retention, loyalty, satisfaction, relationship marketing, logistic regression, segmentation and churn.

## Sumário

1. Introdução	10
1.1. Objetivos do Estudo	13
1.2. Relevância do Estudo	14
1.3. Delimitação do Estudo	15
1.4. A Indústria Jornalística	15
2. Revisão de Literatura	19
2.1. O Conceito de Retenção	19
2.1.1. Retenção e MKt de Relacionamento	21
2.1.2. Retenção e Lealdade	25
2.1.3. Retenção e Satisfação	28
2.1.4. A Relação entre Retenção, Marketing de Relacionamento, Lealdade e Satisfação	30
2.2. Segmentação do Mercado Consumidor	32
2.2.1. Variáveis Geográficas	36
2.2.2. Variáveis Demográficas	37
2.2.3. Variáveis Psicográficas	39
2.2.4. Variáveis Comportamentais	41
2.3. Variáveis Transacionais	43
3.1. Tipo de Pesquisa	47
3.2. Universo e Amostra	48
3.3. Coleta dos Dados	49
3.4. Limitações do Método	49
3.5. Tratamento dos Dados	50
3.5.1. Problema e Objetivo	50
3.5.2. O Método de Regressão Logística	51
3.5.3. Coleta e Análise dos Dados	60
3.5.4. Definição das Variáveis	60
3.5.5. Descrição da Base de Dados	67
4. Análise dos Resultados	69
4.1. Construção do Modelo de Regressão Logística	69
4.2. Avaliação do Ajuste Geral	71
4.3. Precisão da Estimativa	73
4.4. Validação do Modelo ( <i>holdout sample</i> )	75
4.5. Interpretação dos Resultados	76
5. Conclusões e Recomendações	84
5.1. Conclusões	84
5.2. Recomendações.	85
5.3. Recomendação para futuras pesquisas	87
6. Referências Bibliográficas	89
Anexo A – Estatísticas Descritivas	93
Anexo B – Distribuição Geográfica	97

## Lista de Gráficos

Gráfico 1: Circulação Média Mensal / Ano – Mercado do Rio de Janeiro	16
Gráfico 2: Circulação Média Mensal / Ano – Mercado do Rio de Janeiro	17
Gráfico 3: VPL do Cliente x Risco de Desligamento	20
Gráfico 4: Curva Logística	54
Gráfico 5: Total de Assinaturas x Score de Risco de Cancelamento.	77
Gráfico 6: Score de Risco de Cancelamento x Gênero	78
Gráfico 7: Score de Risco de Cancelamento x Idade	78
Gráfico 8: Score de Risco de Cancelamento x Tipo De Assinatura	79
Gráfico 9: Score de Risco de Cancelamento x Região	79
Gráfico 10: Score de Risco de Cancelamento x Região	80
Gráfico 11: Score de Risco de Cancelamento x Participação em Ações de Fidelização	81
Gráfico 12: Score de Risco De Cancelamento x Canal de Venda	81
Gráfico 13: Score de Risco de Cancelamento x qtde de Produtos Agregados Comprados	82
Gráfico 14: Score de Risco de Cancelamento x Tempo de Permanência	82
Gráfico 15: Score de Risco de Cancelamento x SD &W	83
Gráfico 16: Score de Risco de Cancelamento x Indicador de Reclamação	83

# 1

## Introdução

As empresas estão constantemente aprimorando formas de maximizar seu lucro. Para atingir este objetivo, elas podem atuar na linha das receitas e das despesas. Além da redução de custos, as empresas podem aumentar preços, vender para um número maior de clientes, vender mais para os clientes atuais, vender para clientes que permaneçam ativos por mais tempo ou podem perder menos clientes.

Como a redução de custos não pode ser realizada indefinidamente, os clientes não absorvem aumentos de preços constantes e vender é uma atividade muito cara, reter clientes passa a ser tão ou mais importante do que atrair novos consumidores.

Do ponto de vista estratégico, a atração de novos clientes deveria ser considerada um passo intermediário, uma vez que o fundamental é desenvolver relacionamentos estreitos com os clientes atuais, objetivando sua lealdade (Berry, 2002). Os benefícios dessa lealdade geralmente se refletem nos motivos pelos quais um competidor é mais lucrativo do que outro (Reichheld, 1993). É imprescindível que todos os colaboradores entendam o valor dos clientes ao longo do tempo, almejando a perda zero de clientes (Reichheld & Sasser Jr., 1990).

Em um mercado cada vez mais competitivo, as empresas devem adotar estratégias para obter vantagem competitiva sustentável (Porter, 1989; Hunt *et al.*, 2002). Assim, o marketing de relacionamento, conjugado à estratégia de retenção de clientes, torna-se uma alternativa atraente. Entendendo o conceito de marketing de relacionamento como atrair, manter e aumentar os relacionamentos com clientes (Berry, 2002) e, quando necessário, terminá-los (Grönroos, 2000). Há profunda mudança de paradigma implícita no conceito – da visão que aborda o marketing transacional, focado na venda como um fim em si mesmo, para o marketing relacional, focado em relacionamentos de longo prazo com os clientes (Dwyer *et al.*, 1987; Morgan & Hunt, 1994; Gummesson, 1998).

Dependendo do setor de atuação, as empresas podem aumentar seus lucros de 25% a 85%, reduzindo em apenas 5% a perda de clientes (Reichheld & Sasser Jr., 1990). Segundo Asbrand (1997) e Knowles (1997), a retenção pode custar 4 a 8 vezes menos do que a aquisição de clientes.

A taxa de retenção de clientes é vista por Rust & Zahorik (1993) como o componente mais importante para que a empresa consolide sua participação no mercado, sendo direcionada pela satisfação do consumidor. Embora a satisfação não seja obrigatoriamente revertida em lealdade, as empresas devem persegui-la e reconhecer que ela é alcançada pela entrega consistente de valor superior ao cliente (Reichheld, 1993; Reichheld *et al.*, 2000).

As atividades de relacionamento com o cliente já estão sendo incorporadas ao negócio das empresas, pois são uma prática de marketing cujo foco é a retenção de clientes, em que o grande desafio é reconhecê-los, mostrando-lhes o quanto a empresa os estima por terem lhe conferido a preferência (Vavra & Pruden, 1995).

Para reconhecer esses clientes é preciso um database marketing bem estruturado que armazene o histórico de transações dos consumidores e os dados de endereço, mantendo a integridade das informações para monitorar continuamente o comportamento dos clientes. Segundo Stone e Shaw (1987, p.13), o database marketing é uma abordagem interativa para a comunicação de marketing que usa os meios de comunicação para expandir sua audiência, para estimular a demanda desta audiência e para estar próximo dela, além de manter a memória dos clientes no banco de dados para melhorar os contatos futuros.

Nash (1994) afirma que a espinha dorsal de muitos programas de relacionamento é o sistema de database marketing.

Os programas de relacionamento têm por objetivo aumentar o *life time value* dos clientes e, segundo Vavra (1994), por quanto mais tempo se mantém um cliente, mais lucrativo ele poderá ser para o negócio, e por quanto mais tempo o cliente comprar de uma mesma organização, mais dependente poderá tornar-se de seus produtos ou serviços e estará menos suscetível a ofertas com preços mais baixos da concorrência.

Programas voltados à retenção de clientes devem analisar sua viabilidade econômico-financeira. Portanto, torna-se relevante conhecer o perfil do cliente de

alto risco e identificar quais clientes têm maior probabilidade de serem perdidos para permitir que a empresa priorize suas ações de retenção.

O objetivo deste estudo é utilizar a base de dados de clientes de um jornal para desenvolver um modelo de regressão logística que determine o risco de cancelamento de clientes que possuem uma assinatura de jornal, bem como traçar o perfil dos clientes com maior risco de desligamento para propor ações de retenção que aumentem a lucratividade da carteira de assinantes desta empresa.

Segundo Hair Jr. *et al.* (1998), o erro mais crítico no desenvolvimento de modelos com base teórica é a omissão de uma ou mais variáveis independentes, um problema conhecido como erro de especificação.

Hair Jr. *et al.* (1998) argumentam ainda que o diagrama de definição das variáveis deve ser definido em termos de construtos, para então buscar métricas que irão mensurar esses construtos.

A estrutura desta fundamentação contempla a elaboração de um modelo teórico antes de definir as variáveis que devem compor o modelo matemático de previsão de cancelamento.

Em uma primeira etapa, faz-se importante abordar os diversos conceitos que se relacionam com a retenção de clientes. Em seguida, é preciso definir as variáveis que influenciam o comportamento de compra do consumidor e conseqüentemente ajudam a descrever o grupo que cancela e o grupo que não cancela. Segundo Kotler (1998), os fatores culturais, sociais, pessoais e psicológicos influenciam o comportamento de compra e auxiliam na segmentação. O perfil de um segmento é classificado em termos de atitude, comportamento, demografia e psicografia.

O modelo teórico contará com três grupos de variáveis. O primeiro contempla as variáveis que avaliam a relação cliente-empresa com a retenção, são elas: marketing de relacionamento, lealdade e satisfação. O segundo é o grupo das variáveis relativas às características do cliente e que ajudam na segmentação do mercado consumidor, tais como: geográficas, demográficas, psicográficas e comportamentais. Por fim, contará com as variáveis transacionais que descrevem a interação operacional da empresa com o cliente.

O banco de dados da empresa analisada não possui todas as variáveis do modelo teórico, mas é importante estudá-las. Uma limitação deste estudo é que

nem todas as variáveis levantadas no modelo teórico foram aplicadas no modelo matemático.

## 1.1

### Objetivo

Diante da busca constante para maximizar o lucro das empresas, faz-se necessária uma gestão mais inteligente dos recursos para otimizar as ações de maior impacto. As empresas estão focadas em medir o retorno de cada projeto para investirem tempo e dinheiro naqueles que maximizem os resultados da organização.

Dados os altos custos de aquisição, uma forma de obter lucro é aumentar o tempo de permanência dos clientes atuais, com isso a retenção dos clientes antigos passa a ser tão ou mais importante do que atrair novos clientes.

Pesquisas têm sido realizadas para identificar os clientes com risco de cancelamento e atuar sobre eles com ações de retenção para reverter ou impedir a perda destes clientes.

Faz-se necessário medir o risco de perda de cada cliente, calculando a probabilidade de cancelamento da assinatura, identificar os fatores que explicam o perfil do cliente que cancela e classificá-los em dois grupos: o dos que cancelam e o dos que não cancelam.

Desta forma, este estudo desenvolve um modelo que relaciona variáveis transacionais, demográficas e dados sobre o histórico de eventos com a probabilidade de cancelamento do cliente.

Este estudo se propõe a responder à seguinte pergunta: é possível definir o perfil do cliente que cancela uma assinatura de jornal?

A técnica utilizada para indicar o risco de cancelamento dos assinantes deste jornal é a regressão logística binária que determina a probabilidade de uma assinatura pertencer ao grupo dos que cancelam ou ao grupo dos que não cancelam.

Ao identificar o perfil dos clientes com risco de cancelamento será possível atuar sobre eles através de ações de retenção, que revertam a perda desses clientes.

Além do objetivo principal que é definir o perfil do cliente que cancela a assinatura do jornal, apresentam-se alguns objetivos secundários:

1. revisão de literatura para construir o entendimento dos conceitos de retenção e levantar as variáveis necessárias para a segmentação dos clientes;
2. medir o risco de cancelamento de cada cliente, que nada mais é do que a probabilidade de cancelamento da assinatura;
3. identificar as variáveis que explicam o perfil do cliente que cancela;
4. classificar os clientes em dois grupos: o que cancela e o que não cancela.

## 1.2

### Relevância

Os resultados deste estudo podem contribuir para a tomada de decisão em marketing. Conhecendo os riscos de cancelamento da base de clientes, as empresas podem direcionar ações de retenção orientadas ao público com alta propensão de desligamento em vez de aplicar suas ações para grupos de clientes cujos riscos de saída são baixos. Assim, a organização pode otimizar seu investimento em marketing.

Quando uma empresa aumenta sua taxa de retenção de clientes, poderá reduzir despesas de marketing e vendas e obter lucratividade pelo incremento do tempo de permanência dos seus clientes atuais. Segundo Griffin (1996) a redução nos custos pode surgir das seguintes maneiras:

- a. custos de marketing, pois evitam-se os elevados custos para atrair uma grande quantidade de novos clientes;
- b. custos de transação mais baixos, através de negociações de contratos e processamento de ordens;
- c. despesas geradas pela rotatividade dos consumidores.

Esta análise também poderá contribuir para a aquisição de consumidores à medida que descreve um perfil para o segmento de clientes de maior lucratividade. Assim, este perfil de clientes poderá ser “clonado” através da aquisição de clientes com o mesmo perfil dos clientes de maior retenção.

### 1.3

#### Delimitação

Um assinante de jornal pode ser pessoa física (PF) ou pessoa jurídica (PJ). Uma característica deste último é que um único cliente recebe mais de um exemplar de jornal para cada assinatura. Além disso, a decisão de cancelar a assinatura em uma empresa (PJ) muitas vezes não é do usuário, mas sim de um setor ou departamento da empresa. Para simplificar a análise, este estudo está restrito aos clientes do tipo pessoa física.

O jornal estudado tem prioritariamente clientes das classes AB, logo o resultado deste projeto não pode ser extrapolado para consumidores de todas as classes sociais. E também não pode ser considerado para qualquer segmento de mercado já que o objeto de estudo é o assinante de jornal.

### 1.4

#### A Indústria Jornalística

Até o início dos anos 80 o segmento de jornais no Rio de Janeiro era caracterizado de um lado por jornais "policialescos" (sensacionalistas e de qualidade gráfica inferior) e por outro lado pelos considerados *Quality Papers* (O Globo e Jornal do Brasil). Inexistia uma visão de gestão voltada para resultados.

Em 1994, o Plano Real veio incrementar o poder de consumo das classes populares e em 1998 a Infoglobo que tinha apenas O Globo como seu título, lançou o jornal Extra no segmento popular. Em 2001, o Jornal do Brasil, que enfrenta graves problemas financeiros desde o início da década de 80, foi vendido para o empresário Nelson Tanure e apesar disso não demonstrou sinais de recuperação. O que significa dizer que continua perdendo clientes e manteve seu alto nível de endividamento.

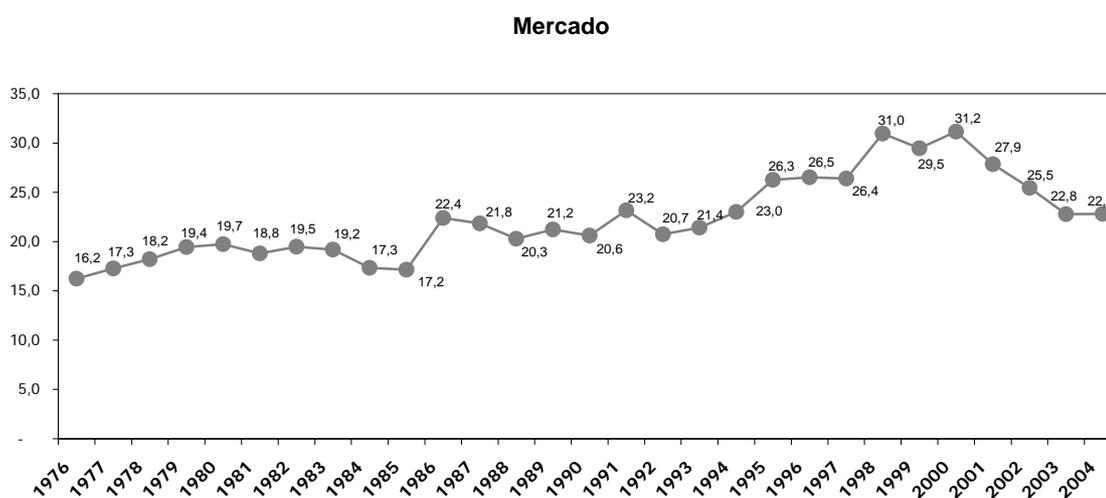
A circulação de jornais diários no Brasil expandiu-se em 0,8% em 2004 (IVC – Instituto Verificador de Circulação, 2005). É o primeiro desempenho positivo após três anos consecutivos de declínio. Houve uma retração de 7,2% em 2003, de 9,1% em 2002 e de 2,7% em 2001. Esses recuos interromperam uma linha ascendente iniciada em 1996. Em outras palavras, após quatro anos de crescimento contínuo, a circulação dos jornais brasileiros registrou três quedas

seguidas (IVC, 2005). O resultado de 2004 faz prever um novo ciclo de crescimento, uma vez que a economia brasileira apresentou expansão e já se comprovou a existência de forte relação entre a circulação de jornais no país e o desempenho do Produto Interno Bruto (ANJ – Associação Nacional de Jornais, 2005).

Dentre todos os mercados do Brasil, o mercado do Rio de Janeiro é o que conta com a competição mais acirrada. O Rio de Janeiro é o segundo mercado de maior penetração de leitura com 58% da população leitora de jornal, enquanto Porto Alegre é o primeiro com 77%, em terceiro está Brasília com 49%, em quarto Curitiba com 45% e em quinto está São Paulo com 43% de penetração de leitura (Pesquisa Marplan, 2005).

O mercado carioca, que cresceu a partir do final da década de 80, caiu nos últimos 3 anos e conseguiu ficar estável no ano de 2004 (ANJ, 2005).

**GRÁFICO 1: Circulação Média Mensal/ ano – Mercado do Rio de Janeiro**



Fonte: IVC – Instituto Verificador de Circulação

O Jornal do Brasil, que teve 26% de *market share* na década de 70, diminuiu sua participação de mercado para 10% em 2004 e vem perdendo 2% (*cagr*<sup>1</sup>) ao ano desde 1976 (IVC, 2005).

<sup>1</sup> *Compound Annual Growth Rate* é a taxa de crescimento anual média de um período em relação a outro.

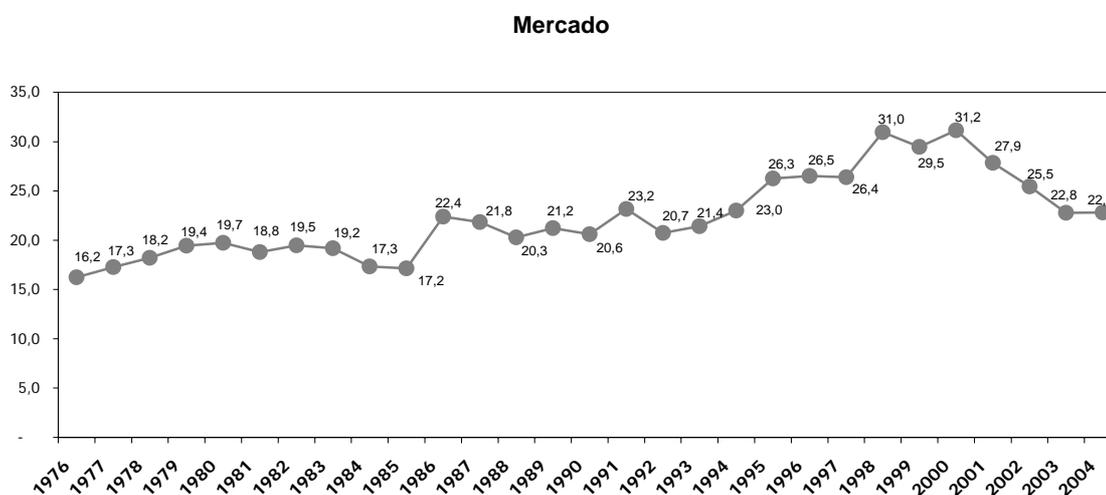
$$Cagr = \left( \sqrt[Qtdedeanos]{ValorFinal \div ValorInicial} \right) - 1$$

O Dia que cresceu muito em circulação entre 1993 e 1999 pelo aumento do poder de consumo das classes BC proporcionado pelo Plano Real, caiu 12,3% (cagr) ao ano desde 1998 (IVC, 2005).

O jornal Extra que foi lançado em 1998 aumentou em 7% sua circulação média em 2004 depois de 3 anos consecutivos de queda.

O jornal O Globo que cresceu 4,2% (cagr) ao ano até 1995, perdeu 6% (cagr) de sua circulação entre 1995 e 2003. Em 2004, o jornal se recuperou crescendo 3,9% em relação a 2003 (IVC, 2005).

**GRÁFICO 2: Circulação Média Mensal/ ano – Mercado do Rio de Janeiro**



Fonte: IVC – Instituto Verificador de Circulação

Em agosto de 2005, dois novos jornais foram lançados no mercado carioca. O jornal Meia Hora – um popular que começou circulando de segunda até sexta em formato tablóide e é atualmente o jornal mais barato do mercado. E o jornal Q! – um vespertino com circulação de segunda a sexta, em formato tablóide para as classes AB, vendido apenas em pontos de grande fluxo de pessoas, como o Metrô.

Cinco meses depois, o jornal Meia Hora está vendendo aproximadamente 83.357 exemplares na média dos dias úteis e lançou uma edição aos sábados e domingos. Ao contrário de todos os jornais, a venda do jornal Meia Hora nos finais de semana é menor que a venda nos dias úteis, aproximadamente 43.215 exemplares na média. O jornal Q! suspendeu a circulação da sua versão impressa e passou a publicar seu conteúdo apenas na Internet (IVC, 2006).

A Infoglobo, empresa que edita os jornais O Globo e Extra, é líder no Rio de Janeiro com 71% do market share carioca. O jornal O Globo é o principal jornal quality paper do Rio com 36% de market share de circulação, 81% das suas vendas são através de assinaturas (IVC, 2005). Um jornal como este sempre será alvo para o crescimento dos demais. A melhor estratégia para a manutenção da circulação do Globo é a retenção dos seus clientes atuais. Além disso, segundo Reichheld (1996), a retenção de clientes tem uma relação positiva com os lucros.

## 2

### Revisão de Literatura

#### 2.1

##### O Conceito de Retenção

Page *et al.* (1996) afirmam que as empresas podem crescer de três maneiras: atraindo novos consumidores, perdendo menos clientes e fazendo mais negócios com seus consumidores; as duas últimas parecem ser a forma mais eficiente de usar os recursos da empresa.

Para Berry e Parasuraman (1991), o marketing para atrair novos consumidores é apenas uma etapa intermediária no processo de marketing. O marketing não se exaure na compra. Segundo Vavra (1993), o marketing deve mudar a mentalidade de completar uma venda para a de iniciar um relacionamento; de fechar um negócio para construir lealdade.

Como vários estudos indicam, reter consumidores sai mais barato e é competitivamente mais vantajoso do que atrair novos clientes (Berry, 1995; Parvatiyar e Sheth, 2000; Koshas, 1997; Aaker, 1998). Uma estratégia defensiva voltada à retenção de clientes pode resultar em melhor desempenho empresarial, principalmente, enfocando aspectos como a melhoria da qualidade dos serviços e o devido tratamento das reclamações (Barcellos, 2000).

Para Vavra e Pruden (1995), a retenção de clientes é a chave para o sucesso no mercado atual. Kaplan e Norton (1997, p.73) comentam que assegurar a retenção dos clientes atuais é uma forma ideal de manter ou aumentar a participação de mercado em segmentos específicos.

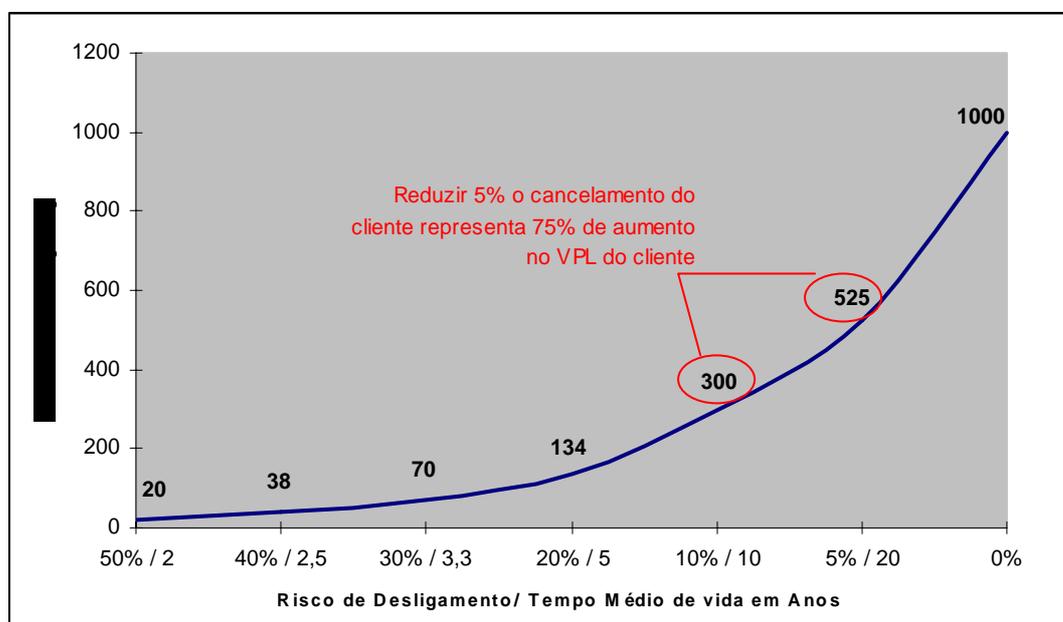
Nash (1994), Frigstad (1995) e Kotler e Armstrong (1995) afirmam que o custo para conservar um cliente atual satisfeito é menor do que a despesa para conquistar um novo. Vavra (1993) demonstra que a aquisição de novos clientes pode custar até 5 vezes mais do que o custo necessário para satisfazer e reter clientes já existentes. O custo de retenção de um cliente antigo é menor do que o custo de aquisição de um cliente novo (Alfor e Sherrell, 1996).

Quanto maior o tempo durante o qual uma empresa retém seu cliente, maior é o lucro gerado por ele (Reichheld, 1996). As empresas geralmente perdem entre 15% e 20% de seus clientes por ano, uma redução nestas taxas pode representar um crescimento significativo nos lucros (Reichheld, 1990). Reichheld e Sasser *apud* Berry e Parasuraman (1995, p.159) chegam à conclusão de que as empresas podem aumentar os lucros de 25% a 85%, reduzindo as fugas de clientes em apenas 5%.

Zeithaml, Berry e Parasuraman (1996, p.34) acrescentam que a retenção de clientes traz impacto financeiro para as organizações, dizendo que a longevidade do relacionamento com o cliente influencia favoravelmente a lucratividade.

Reichheld e Sasser (1990) construíram uma curva de cancelamento de clientes em uma empresa de cartões de crédito e mostraram que uma redução no cancelamento de 20% para 10%, ou um aumento de 5 para 10 anos no período de utilização do serviço, provoca um aumento do valor do cliente de 134 para 300 dólares (Gráfico 3).

**GRÁFICO 3: VPL do cliente x Risco de Desligamento**



Fonte: Reichheld, Frederick F., Sasser Jr, W.Eart. Zero Defections: Quality Comes to Service. Harvard Business Review, v.68, n.5, p.109, Sep/Oct. 1990.

Provavelmente as empresas não conseguem eliminar todos os cancelamentos – e não devem tentar – embora, elas possam e devam reduzi-los.

Um estudo feito por Oliver (1999) em 14 ramos de negócios diferentes mostrou que ao aumentar em 5% a retenção dos consumidores gera-se um incremento de 25% a 95% no valor presente líquido (VPL).

Além da redução dos custos operacionais, o relacionamento de longo prazo pode permitir em alguns casos que as empresas cobrem mais por seus produtos e serviços.

Segundo Trubik e Smith (2000) de todos os cancelamentos de clientes, 35% são gerados por fatores externos incontroláveis diante da perspectiva de satisfação do cliente. Os outros 65% são causados por fatores internos controláveis, relacionados ao tratamento oferecido pelas empresas aos seus consumidores.

Segundo Vavra (1993), as razões mais comuns que levam o cliente a abandonar uma empresa são:

- a. insatisfação com o produto, a entrega, a instalação, os serviços ou os preços;
- b. reclamações ignoradas, minimizadas ou mal solucionadas;
- c. descontentamento com mudanças de preço, políticas ou vendedores;
- d. insatisfação com o tratamento ou falta de cortesia;
- e. novos funcionários ou novas políticas da empresa;
- f. aceitação de uma oferta do concorrente.

Apesar disso, Reinartz e Kumar (2000) obtiveram os seguintes resultados de um estudo focado em compra por catálogos:

- a. não existe associação positiva entre relacionamentos de longo prazo e lucro para a empresa;
- b. a lucratividade pode ser gerada igualmente por clientes de longo prazo e clientes recentes;
- c. os custos de servir não são menores para os clientes de longo prazo;
- d. clientes de longo prazo não pagam mais caro pelos produtos ou serviços.

### 2.1.1

#### **Retenção e Marketing de Relacionamento**

O marketing moderno não se limita à transação apenas, refere-se à batalha pela fidelidade do cliente, Mckenna (1993).

Kotler (2000) argumenta que as empresas devem avaliar a relação custo/benefício entre o esforço de fidelização via investimento no relacionamento e a conseqüente melhoria de seus níveis de receita. Ele classifica os níveis de investimento no relacionamento em cinco diferentes tipos de abordagem de marketing:

- a. Marketing Básico: o vendedor simplesmente vende o produto;
- b. Marketing Reativo: o vendedor vende o produto e encoraja o cliente a contatar em caso de dúvidas, comentários e queixas;
- c. Marketing Responsável: o vendedor contata o cliente após a venda e verifica o atendimento de expectativas, investiga decepções e pede sugestões de melhoria. As informações obtidas auxiliam a empresa a melhorar seu desempenho;
- d. Marketing Pró-ativo: o vendedor entra em contato com o cliente de forma regular para falar de melhorias na utilização dos produtos ou de novos produtos;
- e. Marketing de Parceria: a empresa trabalha continuamente com o cliente buscando formas de melhorar seu desempenho.

Cada vez mais as empresas têm se preocupado em crescer de forma sustentável e lucrativa ao longo do tempo. Há muito, as ações de médio e longo prazo passaram a fazer parte da agenda dos executivos das grandes organizações. Segundo Day (2000), a habilidade de uma empresa em criar e manter relacionamentos com os seus mais valiosos clientes é uma base duradoura para uma vantagem competitiva.

Diante deste cenário, houve profunda mudança no conceito que aborda o marketing transacional, focado na venda como um fim em si mesmo, para o marketing relacional, focado em relacionamentos de longo prazo com os clientes (Dwyer *et al.*, 1987; Morgan & Hunt, 1994; Gummesson, 1998).

Além disso, as profundas mudanças que ocorrem no comportamento do consumidor fazem aumentar a exigência dos clientes e influenciam transformações na administração de marketing. Essa nova forma obriga as organizações a passarem por transformações fundamentais de crenças e valores (McKenna, 1999:I) e exigem das organizações mudanças na sua forma de atuação,

direcionando sua estratégia para o foco no cliente com o objetivo de conseguir acompanhar essa evolução.

O foco está na necessidade de um relacionamento de longo prazo em detrimento das transações com objetivos de curto prazo, ou seja, é preciso buscar a fidelização de clientes (Ribeiro, 1999). O marketing após a venda é tão importante quanto aquele marketing que leva à venda (Light, 1994).

Para manter contato com o cliente após a venda, Vavra (1993) propõe a retenção dos clientes que utiliza várias táticas de pós-marketing. Esta seria uma oportunidade de transformar um comprador qualquer em um comprador leal (Light, 1994).

A interação que o comprador estabelece com o fornecedor, embora seja comercial, constitui um relacionamento e isso requer visão de longo prazo e respeito mútuo, afirma Gummerson (1998). O relacionamento e a busca de fidelização são formas de alcançar maior competitividade para as empresas e maior satisfação para os clientes (Berry, 1995).

Peppers e Rogers (2000) definem marketing de relacionamento através do termo marketing *one to one*, estabelecendo-o como uma estratégia voltada ao entendimento e à antecipação das necessidades dos clientes atuais e potenciais de uma empresa, sempre com o auxílio da tecnologia da informação. Existem 4 etapas para atingir esse resultado: identificar, diferenciar, interagir e personalizar. Neste sentido, é importante identificar e diferenciar clientes com alto risco de cancelamento antes de comprometer parcelas importantes dos investimentos de marketing no relacionamento com eles.

Para Sheth e Parvatiyar (2000) o marketing de relacionamento é um processo em curso de engajamento em atividades cooperativas e colaborativas e programas com consumidores finais para criar e intensificar valor econômico e redução de custos mútuos.

Existem outras vantagens para as empresas além dos resultados financeiros, segundo Kahn (2000) o marketing de relacionamento é baseado na idéia de estabelecer uma relação de aprendizado com os consumidores procurando aumentar sua satisfação e minimizar qualquer problema.

Wulf, Schröder e Iacobucci (2001) propõem um modelo onde algumas táticas de marketing de relacionamento como correspondência direta, tratamento preferencial, comunicação interpessoal e recompensas tangíveis levariam o

consumidor a perceber um alto nível de investimento no relacionamento por parte da empresa, conduzindo a um relacionamento de qualidade com satisfação, confiança e comprometimento, que resultaria em um comportamento de lealdade.

Para a construção de relacionamentos com o cliente, as empresas devem entender as mudanças no comportamento do consumidor ao longo do tempo e precisam aprender em cada interação, personalizando os contatos e fortalecendo o vínculo entre o cliente e a empresa (Newell, 2000). O acesso à informação sobre o consumidor é um aspecto fundamental no relacionamento com o cliente (Wyner, 1999), apesar disso, este relacionamento não pode estar baseado apenas no database marketing que segundo Barnes (1994) é utilizado geralmente em sentido único e iniciado pela empresa. Muitas vezes o consumidor não percebe um relacionamento mesmo sendo impactado por correspondências da empresa freqüentemente.

Barnes (1994) propõe a existência do relacionamento com o consumidor em 3 diferentes formas:

- *Customer staff relationships*: a relação de troca aproxima-se de um relacionamento pessoal ou social. Acontece quando há um contato próximo ou em empresas pequenas nas quais se mantém contato quase sempre com os mesmos funcionários.

- *Customer firm relationships*: busca a proximidade dos consumidores com a empresa e não com os funcionários. O database marketing e o telemarketing podem ser utilizados, embora haja o risco de serem considerados invasivos e ofensivos pelos clientes.

- *Customer brand relationships*: ocorre nos casos em que o contato direto com a empresa e seus funcionários não é possível ou desejável, por exemplo, nos produtos vendidos nos supermercados. Os consumidores podem estabelecer um relacionamento muito maior com a empresa que vende os produtos e seus funcionários do que com a própria empresa que os produz.

Para Barnes (1994) o mais importante nos três casos é quando e sob que circunstâncias o consumidor percebe que existe um relacionamento. Seu estudo identifica alguns aspectos da interação que podem conduzir a um relacionamento mais próximo de acordo com a análise de 40 grupos de foco com consumidores de empresas de serviços: frequência de contato, complexidade da interação, negociação sempre com a mesma pessoa, contato personalizado e risco percebido.

Além destes, destacam-se: confiança, comprometimento, comunicação, desejo de manutenção de promessas, conforto, benefícios derivados do relacionamento, valores compartilhados/compreensão, empatia, ética, competência, familiaridade, cooperação, disposição para resolver conflitos, tratamento preferencial e sinceridade.

Alguns aspectos negativos também podem ser observados. Para Grayson e Ambler (1999) o relacionamento pode causar um aumento exagerado nas expectativas dos clientes e fazê-los acreditar que a empresa está se aproveitando da confiança construída entre as partes para agir de forma oportunista.

Como a administração de transações está sendo substituída pela administração de relacionamentos (McIntyre e Peck, 1998), as empresas têm empregado programas de retenção ou de lealdade (Kotler e Armstrong, 1995) e utilizam o marketing de relacionamento como ferramenta para se aproximarem dos seus consumidores. Apesar disso, nem sempre esta estratégia é aplicada a todas as empresas e a todos os clientes.

Os pré-requisitos para a construção do marketing de relacionamento referem-se à identificação dos grupos de clientes mais lucrativos e dos fatores e grau de lealdade dos grupos de clientes-chave, considerando atitudes e comportamentos (Lemos, Popadi e Berndt, 2000).

### **2.1.2**

#### **Retenção e Lealdade**

A prática de promoção de descontos tem desgastado a lealdade do consumidor à marca (Ribeiro *et al.*, 1999). Segundo Aaker (2000), é necessário desenvolver vantagens competitivas mais sustentáveis que sejam baseadas em competições fora da área de preços.

Para fugir da guerra de preços, muitas empresas têm tentado estimular a lealdade do consumidor para que fiquem menos sensíveis às variações de preço. Existem diversas definições para lealdade, entre elas a intenção do cliente de comprar novamente (Bateson e Hoffman, 2001).

Reichheld (1996) define que a lealdade de um cliente está na permanência dele com uma empresa durante muitos anos enquanto Gerpott, Rams e Schindler (2001) entendem lealdade como a atitude favorável de um cliente perante uma

empresa e seus serviços. Reichheld (1996, p.3) propõe ainda uma gestão baseada na lealdade. Segundo o autor, o lucro, obviamente, é indispensável, o que, junto com a lealdade, compõe a verdadeira essência de qualquer instituição comercial duradoura e bem sucedida.

Segundo Lacoeyilhe (1997) as primeiras definições de lealdade baseavam-se na observação do comportamento do consumidor fiel quando este comprava regularmente a mesma marca. Esta é uma abordagem comportamental ou behaviorista que considera a lealdade uma resposta passiva baseada na análise do comportamento de compra. Hart *et al.* (1999) afirmam ainda que a falta de opção, o hábito e a conveniência são fatores que podem levar à lealdade comportamental.

A abordagem atitudinal permite considerar o aspecto ativo da fidelização de clientes. O conceito de comprometimento à marca traduz uma tendência a resistir à mudança. As escalas atitudinais são usadas para medir a intenção do consumidor de adotar um determinado comportamento. Para Lee e Lee (1999) a visão atitudinal demonstra resistência a um produto superior.

Existe ainda uma terceira abordagem considerada por Neal (1999), Baldinger e Rubinson (1996) que busca enfatizar aspectos tanto comportamentais quando atitudinais quando se fala de lealdade. Sob este ponto de vista, Baldinger e Rubison (1996) acreditam que consumidores comportamentalmente fiéis a uma marca são esperados a considerar essa marca atitudinalmente mais favorável do que aquelas que eles nunca compraram ou compraram com baixa frequência.

A partir de todas essas abordagens, Lacoeyilhe (1997) considera leal aquele consumidor que possui uma atitude efetiva de compra e não só a intenção de comprar. “O sucesso em longo prazo de uma marca é baseado não no número de consumidores que a compram apenas uma vez, mas no número de consumidores que se tornam compradores repetidos” (Jacoby e Chestnut, 1978).

Quando relacionada diretamente à compra, a lealdade do cliente pode ser o principal ativo de uma empresa, porém esta lealdade não é construída apenas pelo oferecimento do produto certo, no momento correto e pela comunicação adequada, mas também pela administração do relacionamento com o cliente (Stone, Shaw, 1987). A lealdade ocorre de forma crescente, desenvolvendo-se até chegar a um comprometimento profundo com um produto ou serviço (Oliver, 1996).

Segundo Dowling e Uncles (1997) um programa de lealdade pode ser útil para uma empresa quando neutraliza o programa dos competidores, aumenta a disponibilidade de produtos e serviços ou incrementam a percepção de valor do produto ou serviço.

Os programas de lealdade não são destinados a todos os clientes, afinal “os clientes não são iguais, pois 80% do total de recompras de bens e serviços provêm de 20% dos clientes atuais, os projetos de fidelidade concentram esforços específicos nos 20% mais valiosos de seus bancos de dados” (Bob Stone, 1989, p.3).

A obsessão pela lealdade do consumidor não é consequência de um interesse intensificado pela filantropia. Ela provém da consciência de que essa lealdade está diretamente relacionada aos resultados financeiros dos negócios (McIntyre e Peck, 1998, p. 98). Embora, os programas de lealdade não sejam adequados para todos os negócios (Hochman, 1992).

Alguns autores acreditam que a lealdade traz aumento dos lucros. As empresas sabem intuitivamente que quando a lealdade do consumidor aumenta, os lucros crescem também (Reichheld, 1993). Berry (1995) também acredita que os consumidores fiéis geram mais receita durante mais tempo. Reichheld e Sasser (1990) complementam dizendo que a perda de poucos clientes pode trazer grande impacto nos lucros.

Para Griffin (1996) os clientes leais apresentam quatro comportamentos específicos: compram regularmente, compram produtos e serviços oferecidos pela venda cruzada, recomendam o produto a outros consumidores e são protegidos do assédio de outros competidores. Rapp (1996, p.25) faz um contraponto ao afirmar que recompensar os melhores clientes só gera fidelidade à marca quando seu concorrente não está oferecendo uma recompensa igual ou melhor.

Em alguns casos, a lealdade do consumidor se dá por falta de opção. Koskas (1997) afirma que quanto mais competitivo for um mercado e maiores as opções de escolha, menor será a fidelidade dos clientes.

Johnson (1998) classifica os programas de lealdade em:

- a. *Rewards* (programas de recompensa): bonifica o cliente com pontos pelas compras e estes pontos podem ser trocados por recompensas não relacionadas com a marca do produto comprado;

- b. *Rebate*: oferece um vale resgatável no momento da realização de novas compras;
- c. *Appreciation* (programa de milhagem): recompensa com um produto ou serviço da própria empresa em troca de pontos acumulados;
- d. *Partnership* (programa de parceria): recompensa as compras acumuladas dos clientes com produtos ou serviços de uma empresa parceira;
- e. *Affinity* (programa de afinidade): o cliente recebe benefícios especiais, comunicação diferenciada e reconhecimento como cliente especial.

Para Reichheld (1996), existem três tipos de clientes leais. O primeiro tipo é composto por clientes previsíveis e leais, independente da empresa, eles preferem relacionamentos estáveis e duradouros. O segundo possui clientes mais rentáveis do que outros, já que gastam mais, pagam em dia e demandam menos serviços. O terceiro grupo é formado por consumidores que vêem mais valor nos produtos ou serviços de uma empresa em relação a outras, pois estes produtos vão ao encontro das necessidades desses clientes.

### 2.1.3

#### **Retenção e Satisfação**

As empresas competitivas devem se dedicar a oferecer satisfação para os seus clientes, entendendo o nível de expectativa deles. Conhecer estas expectativas é importante para que a empresa possa gerenciar a relevância e o impacto que cada mudança trará na satisfação dos clientes.

Kotler (1996) define satisfação como sendo o sentimento resultante da comparação do desempenho de um produto ou serviço em relação às suas expectativas. Bote (1996) afirma ainda que a satisfação é uma extensão pela qual as expectativas dos clientes, em relação a um produto ou serviço, são atendidas ou não através do valor percebido.

Os clientes experimentam vários níveis de satisfação ou desapontamento após cada experiência de serviço à medida que suas expectativas são atingidas ou ultrapassadas (Lovelock e Wright, 2001).

Yin (1990) conceitua satisfação como o julgamento atitudinal que se realiza após a compra ou após uma série de interações do consumidor com um produto.

Oliver (1999) acrescenta que satisfação é um preenchimento prazeroso, ou seja, o consumidor sente que o consumo preenche alguma necessidade, desejo ou objetivo de forma prazerosa, gerando um desempenho favorável comparado a um padrão existente.

Melo *apud* Oliver (2002) define cinco maneiras diferentes de satisfação: a) contentamento (com o seu sentimento primário de aceitação ou tolerância); b) prazer (um reforço positivo que envolve a evocação ou a melhoria de uma experiência positiva e um sentimento de felicidade); c) alívio ( um reforço negativo que ocorre quando um estado de aversão é removido); d) novidade (expectativa relacionada ao não esperado que forma um sentimento primário de interesse ou excitação); e) surpresa (um sentimento primário de deleito que ocorre quando o produto excede às expectativas).

Fournier (1999) nota que a maioria das pesquisas sobre satisfação estão baseadas na visão da transação. Vários observadores têm criticado a função de marketing por tratar satisfação como uma avaliação estática originada através de um evento isolado, destacando que padrões de comparação são dinâmicos e se alteram com a experiência de consumo (Iacobucci, Gryson e Ostrom, 1994).

A satisfação do cliente é o resultado fundamental do processo de marketing (Bateson e Hoffman, 2001). A satisfação do consumidor vem sendo encarada como um fator crucial para o sucesso dos mais variados tipos de organizações, especialmente por influenciar diretamente a lealdade à marca, repetição de compras, comunicação boca a boca positiva, lucratividade e participação de mercado (Oliver, 1997). Entendendo essas dimensões, as organizações conseguirão influenciá-las e controlá-las.

Apesar disso, pesquisas realizadas em diversas empresas mostram que entre 60% e 80% dos clientes que cancelaram estavam satisfeitos ou muito satisfeitos (Reichheld, 1996). Há indicações de que apenas a satisfação dos clientes não é suficiente para garantir a recompra (Zeithaml, Parasuraman e Grewal, 1996).

Por outro lado, Cardozo (1965, p.244) cita que a satisfação do consumidor em relação a um produto leva à repetição de compra, aceitação de outros produtos similares, além do boca a boca favorável.

Se as empresas tentarem aumentar a satisfação de seus clientes indefinidamente, reduzindo preços ou melhorando serviços, poderão obter lucros menores (Kotler, 2000).

Para Rapp (1996, p.25), “a satisfação dos clientes é uma necessidade ao se fazer negócios. É tanto uma necessidade para seus concorrentes como para você mesmo. Assim, é correto supor que satisfazer os clientes pode evitar que você perca negócios; mas não garante vantagem competitiva que leve ao aumento dos lucros”.

#### **2.1.4**

### **A Relação entre Retenção, Marketing de Relacionamento, Lealdade e Satisfação**

As forças da globalização, a velocidade com que as mudanças tecnológicas se apresentam, somadas às mudanças, cada vez mais rápidas, das necessidades dos clientes, formam o pano de fundo deste novo ambiente competitivo (Kotler, 2000).

No marketing, as últimas tendências apontam para a busca de relações duradouras e de lealdade com os clientes, como estratégia de grande parte das organizações. Vários autores têm se preocupado com esta questão e com frequência recomendam que as empresas procurem uma estratégia de mercado focada no relacionamento de longo prazo.

Para Kotler (1998), as empresas devem modificar suas estratégias para se adaptar ao novo ambiente competitivo. Devem trocar a gestão de produtos pela gestão de clientes.

Além de repetirem a compra, clientes satisfeitos falam favoravelmente da empresa (Berry e Parasuraman, 1995), são menos sensíveis a preço (Bateson e Hoffman, 2001), são propensos a permanecerem leais por mais tempo a despeito de ofertas competitivas atraentes (Lovelok e Wright, 2001) e custam menos para serem mantidos (Kotler, 2000).

A satisfação sozinha não é suficiente para garantir a retenção de clientes (Reichheld, 1996), ela precisa ser convertida em relacionamentos e lealdade (Day, 2001, p.31), mas contribui para a criação de um vínculo emocional cujo resultado é um alto grau de lealdade (Kotler, 2000). As empresas devem construir um relacionamento mais intenso que culmine na retenção de seus clientes (Kurtz e Clow, 1998).

Para obter vantagem competitiva diante dos concorrentes, empresas tentam reter seus clientes através do aumento de sua satisfação e de sua lealdade (Gerpott, *et al.*, 2001). Além disso, muitos clientes em potencial já foram conquistados e as empresas com poucos clientes certamente vão tentar atrair os de outras empresas (Furlong, 1994, p.4).

Kurtz e Clow (1998) sugerem que uma alta satisfação se reverterá em alta lealdade. Mas Bateson e Hoffman (2001) alertam que, em mercados de alta competição, pequenas mudanças na satisfação dos clientes podem ser associadas a grandes mudanças na lealdade.

Kotler (2000) afirma que “a chave para a retenção de clientes é a satisfação”. Para Oliver (1999) a satisfação dos clientes não é suficiente para gerar lealdade, pois é apenas um determinante da lealdade, mas não é o único.

Zeithaml e Bitner (1996) sugerem que a retenção é uma das conseqüências da lealdade.

Para Berry e Parasuraman (1995) a retenção de clientes é uma conseqüência desejável do marketing de relacionamento.

Gerpott *et al.* (2001) sugerem que os conceitos de satisfação, lealdade e retenção estejam ligados em dois estágios: a satisfação influencia a lealdade e a lealdade influencia a retenção de clientes.

Para Kotler e Armstrong (1995) e Jones e Sasser (1995), a melhor maneira de manter clientes é oferecer altos níveis de satisfação e valor que resultem em lealdade.

Rapp (1996) e Reichheld (1993) compartilham a opinião de que a satisfação do cliente não é um substituto para a retenção. Apesar de ser intuitivo que um aumento na satisfação do cliente terá como conseqüência um aumento na retenção e nos lucros, os fatos são controversos. Dos clientes que se desligam de uma empresa, entre 65% e 85% deles estavam satisfeitos ou muito satisfeitos com seus fornecedores. Para Reichheld (1993) os sistemas de medida de satisfação não foram desenvolvidos para oferecer um entendimento sobre quantos clientes permanecem leais à empresa e por quanto tempo.

Independente das opiniões contraditórias sobre a relação entre lealdade e lucro, muitas empresas utilizam o marketing de relacionamento como ferramenta para aumentar a lealdade dos consumidores e garantir a retenção. O marketing de relações é essencial ao desenvolvimento de liderança, fidelidade do consumidor e

rápida aceitação de novos produtos e serviços no mercado. A criação de relações sólidas e duradouras é uma tarefa árdua, de difícil manutenção. Mas acredita-se que, em um mundo onde o cliente tem tantas opções, mesmo em segmentos limitados, uma relação pessoal é a única forma de manter a fidelidade do cliente (McKenna, 1999:105). Segundo Wehner (2000) as empresas estão se conscientizando de que a base de clientes é o ativo mais importante e precioso.

Diante da acirrada concorrência, as empresas têm o desafio de reter seus clientes ativos uma vez que encontrar novos clientes está cada vez mais raro. Surge neste cenário, o marketing de relacionamento como uma alternativa ao marketing de massa. O objetivo é desenvolver uma relação individual com os clientes e desenvolver com eles um relacionamento duradouro.

Com o marketing de relacionamento, forma-se uma rede de marketing composta pela empresa e os *stakeholders* (consumidores, funcionários, fornecedores, distribuidores, varejista, agências de propaganda, cientistas universitários e outros), cuja principal tarefa seria criar forte lealdade dos consumidores (Kotler, 1998:59).

Alguns autores afirmam que níveis maiores de satisfação do cliente levam a níveis maiores de lealdade do cliente (Anderson e Sullivan, 1993; LaBarbera e Mazursky, 1983; Oliver e Swan, 1989; Yi, 1991). Através da lealdade, alguns autores dizem que a satisfação ajuda a assegurar rendimentos futuros (Fornell, 1992; Rust *et al.*, 1994), reduzir custos de transações futuras (Reichheld e Sasser, 1990), diminuir a elasticidade de preços (Anderson, 1996) e minimizar a probabilidade de os clientes trocarem de fornecedor (Anderson e Sullivan, 1993).

## 2.2

### **Segmentação do Mercado Consumidor**

O consumidor é influenciado por vários fatores que vão determinar seu poder de compra (Rodrigues e Jupi, 2004) e que ajudam na classificação do mercado.

Para Kotler, o perfil de cada conglomerado pode ser classificado em termos de atitude, comportamento, demografia e psicografia. Já em outra visão, Kotler e Armstrong (1993) destacam que os fatores que podem influenciar o

comportamento do consumidor são as motivações, a personalidade e as percepções.

O mercado é composto por uma grande diversidade de clientes produtos e necessidades. O papel do marketing é determinar quais segmentos oferecem maiores chances para a empresa alcançar seus objetivos. O agrupamento em segmentos se dá através de diversos fatores: geográficos, demográficos, psicográficos e comportamentais (Guagliardi *et al.*, 2002).

Rodrigues e Jupi afirmam que os consumidores são influenciados ainda por outras variáveis muito importantes em seu comportamento, tais como: classes sociais, variáveis sociais, variáveis econômicas e variáveis culturais. Cobra (1992) defende os fatores culturais, sociais e pessoais como aqueles que influenciam o comportamento de compra dos clientes.

Castro (2004) destaca que os fatores situacionais têm o poder de influenciar não só o comportamento no ato da compra, mas também a maneira como ele se relaciona com as marcas e os produtos.

O campo do comportamento do consumidor surgiu no final da década de 60 através da influência de escritores como Engel, Blackweel e Miniard (2000) que visavam a identificar as variáveis que rodeiam o consumidor.

Para Karsaklian (2000), por mais que se busque conhecer o consumidor, é difícil saber quais comportamentos ele vai ter em relação a fatos que ocorram durante a vida.

Blackweel (2000) defende que o que leva as pessoas a decidirem por um determinado produto vem da sua personalidade, do seu estilo de vida e das forças sociais. Kotler complementa ainda que um indivíduo é influenciado pelos vários pequenos grupos com os quais interage.

Os grupos de referência e a família têm forte apelo no momento da decisão de compra, apesar do comportamento do consumidor ser explicado acima de tudo pelas preferências e pelas intenções individuais. Às vezes, o comportamento é mais afetado por pressões do ambiente social do que por atitudes individuais (Engel, Blackweel e Miniard, 2000, p. 255).

Czinkota (2001, p.138) afirma que o comportamento do consumidor é uma atividade mental e física que resulta em decisões e ações de pagar, comprar e usar produtos tanto para bens como serviços. Defende ainda que entender o consumidor na sua percepção individual, em suas decisões de compra, desde as

questões culturais até as sociais, significa entender até que ponto estas variáveis interferem nas decisões do consumidor.

Para Kotler e Armstrong (1993), a cultura é de grande importância no comportamento de uma pessoa e esse comportamento pode ser aprendido, como no caso das crianças que aprendem seus valores ao crescerem em uma determinada sociedade. A família, a escola e a igreja transmitem padrões básicos de comportamento.

Segundo Kotler, dois grupos amplos de variáveis são usados para segmentar mercados consumidores. Alguns pesquisadores tentam formar segmentos utilizando as características do consumidor e outros examinam as respostas dos consumidores.

Aqueles que tentam formar segmentos pelas respostas dos consumidores examinam benefícios percebidos, ocasiões de uso ou marcas. Já os pesquisadores, que como Kotler utilizam as características para segmentar o mercado, analisam as variáveis geográficas, demográficas, psicográficas e comportamentais.

<b>Variáveis Geográficas</b>	
Região	Pacífico; montanhas; noroeste central; sudoeste central; nordeste central; sudeste central; atlântico sul; atlântico médio; nova Inglaterra.
População da Cidade	Abaixo de 4.999; 5.000 – 19.999; 20.000 – 49.999; 50.000 - 99.999; 100.000 - 249.999; 250.000 - 499.999; 500.000 – 999.999; 1.000.000 - 3.999.999; 4.000.000 ou mais
Concentração	Urbana; suburbana; rural
Clima	Norte; sul
<b>Variáveis Demográficas</b>	
Faixa Etária	Abaixo de 6 anos de idade; 6 - 11; 12 - 19; 20 - 34; 35 – 49; 50 – 64; 65 ou mais
Tamanho da Família	1 a 2 pessoas; 3 a 4; 5 ou mais
Ciclo de Vida da Família	Jovem solteiro(a); casal jovem sem filhos; casal jovem com filho abaixo de 6 anos; casal jovem com filho de 6 anos ou mais; casal de meia-idade com filhos; casal de meia-idade sem filhos menores; solteiro(a) de meia-idade; outros
Sexo	Masculino; feminino
Renda (anual)	Abaixo de \$9.999; \$10.000 – \$14.999; \$15.000 - \$19.999;

	\$20.000 – \$29.999; \$30.000 - \$49.999; \$50.000 - \$99.999 - \$100.000 ou mais
Ocupação	Profissionais liberais e técnicos; gerentes, funcionários públicos e proprietários; funcionários de escritório e vendedores; artesãos e encarregados; operários; fazendeiros; aposentados; estudantes; donas de casa; desempregados
Formação Educacional	Básico completo; básico incompleto; médio completo; médio incompleto; superior completo; superior incompleto; pós-graduado
Religião	Católica; protestante; judaica; muçulmana, hindu e outras
Raça	Branca; negra; asiática
Geração	Pós-guerra; geração X
Nacionalidade	Norte-americana; sul-americana; britânica; francesa; alemã; italiana; japonesa
Classe Social	Miseráveis; pobres; classe operária, classe média; classe média-alta; classe alta; classe alta-alta
<b>Variáveis Psicográficas</b>	
Estilo de Vida	Objetivos; extravagantes; intelectuais
Personalidade	Compulsiva; gregária; autoritária; ambiciosa
<b>Variáveis Comportamentais</b>	
Ocasões	Normais; especiais
Benefícios	Qualidade; serviço; economia; rapidez
Status de Usuário	Pequeno; médio; grande
Status de Lealdade	Nenhum; médio; forte; completo
Estágio de aptidão de compra	Inconsciente, consciente; informado; interessado; desejoso; disposto a comprar
Atitude relativa ao produto	Entusiástica; positiva; indiferente; negativa; hostil

Fonte: Kotler, Philip. *Administração de Marketing. Atlas, 5ª edição, p.232, 1998.*

Essas bases de segmentação também foram pesquisadas em outras obras (Frank, Massy e Wind, 1972; Schewe e Smith, 1982; Grisi, 1986; Las Casas, 1987; Beane e Ennis, 1987; Gunter e Furnham, 1992; Cobra, 1993; Raaij e Verhallen, 1994; Weinstein, 1995; Belch e Belch, 1995; Hooley e Saunders, 1996; Kotler, 1997; Kotler e Armstrong, 1998; Marchioni, 1998; Carvalho, 2001; Chisnall, 1995; Engel, Blackwell e Miniard, 2000; Richers, 2000; Mittal e Newman, 2001; Simpson, 2001). Embora não se tenha encontrado consenso, a

grande maioria destaca quatro tipos de segmentação: geográfica, sócio-demográfica, psicográfica e comportamental.

Os requisitos para que se possa segmentar um mercado, segundo Guagliardi (1984), Sheth, Mittal e Newman (2001) e vários outros estudiosos, são:

a) Mensurabilidade/ Identificabilidade – que o tamanho e o poder de compra dos segmentos possam ser medidos, cada segmento possa ser identificável, de modo que o profissional de marketing possa saber quem é o cliente e quais são suas necessidades, desejos e recursos;

b) Acessibilidade – que os segmentos possam ser alcançados e servidos sem desperdício de recursos. Isto significa que deve haver a possibilidade de discriminar os segmentos pelos hábitos de mídia, pela demografia, etc. de modo que o profissional de marketing determine a forma de anunciar e distribuir o produto;

c) Substancialidade – que os segmentos sejam grandes e/ou suficientemente lucrativos para desenvolver esforços de marketing.

Aranha (1996) afirma que na base do sucesso em marketing de relacionamento está a identificação de um grande número de segmentos específicos e homogêneos de clientes, correntes ou *prospects* (clientes em potencial), sobre os quais atuar. A atenção às particularidades de cada grupo definido permite oferecer os meios exatos para a satisfação precisa de suas necessidades. Uma segmentação competente resulta de um processo longo e complexo.

## 2.2.1

### Variáveis Geográficas

A segmentação geográfica propõe dividir o mercado em unidades geográficas diferentes, como países, estados, regiões, cidades ou bairros (Kotler, 1998). Além de a empresa poder decidir operar em diferentes regiões geográficas, o comportamento de um consumidor pode variar de acordo com a geografia.

Tem crescido o uso de sistemas de informação geográficos na tomada de decisão de marketing. Esses sistemas foram desenvolvidos no início da década de 60 para empresas mineradoras e petrolíferas (Aranha, 1996) e foram recentemente descobertos por empresas industriais e comerciais.

É compreensível o crescimento destas empresas produtoras de softwares de mapeamento, pois eles permitem associar informações de um banco de dados a um mapa e, por meio deles, identificar padrões, analisar tendências, quantificar variáveis, analisar mercados e alocar recursos.

As variáveis de localização são importantes para a gestão eficiente de uma empresa, seja pela análise do potencial de mercado ou pela segmentação da base de clientes atuais ou potenciais. Isto contribui para o planejamento e a projeção de respostas a campanhas ou projeções de tendências.

Na literatura (Tomanari, 2003), encontram-se algumas variáveis geográficas tais como: região, estados, municípios, cidade, metragem, densidade, clima, limites políticos, área comercial, localização, transporte e acesso, topografia, tráfego e centro de compras.

O comportamento de um cliente pode variar profundamente em termos de geografia e isso pode levar uma empresa a adotar estratégias diferentes de acordo com a região onde se encontram seus clientes e seus *prospects*.

## 2.2.2

### Variáveis Demográficas

Segundo Kotler, as variáveis demográficas como idade, tamanho da família, ciclo de vida da família, sexo, renda, ocupação, formação educacional, religião, raça, geração, nacionalidade ou classe social, são as bases mais comuns para se distinguir grupos de consumidores.

Existe uma dificuldade natural em captar e mensurar todas as variáveis que sejam importantes para se realizar uma segmentação de mercado adequada. Kotler afirma que as variáveis demográficas são mais fáceis de serem mensuradas que a maioria de outras variáveis, além de acreditar que os desejos, preferências e taxas de uso dos consumidores estão frequentemente associados às variáveis demográficas.

Em relação ao caráter cultural, a literatura destaca a influência dos valores, normas e expectativas sobre os desempenhos sociais da população em geral e de subgrupos com características sócio-demográficas específicas em termos de sexo, idade, ocupação e papéis sociais (Argyle, 1984, 1994; Bresnahan, Sherman & Lee, 2002; Caballo, 1993; A. Del Prette. & Z. Del Prette, 1999, 2001; Fray &

Hector, 1987; Furnham, 1979, 1993; Trower, Bryant & Argyle, 1978), que podem constituir subculturas com seus padrões próprios de conduta aceitáveis. Dentre esses aspectos, sexo e idade têm sido os mais críticos, levando a estudos sobre padrões de desenvolvimento e de gênero que devem ser necessariamente diferenciados quando se buscam normas de referência. As diferenças na competência social de homens e mulheres ocorrem desde a infância (LaFrenière & cols., 2002). As variáveis sócio-demográficas afetam o desempenho social, porém o peso desta influência pode ser diferente sob contextos culturais diversos.

Para Kotler e Armstrong (1993), renda, ocupação, educação e riqueza são variáveis essenciais para a classificação de uma classe social. A classe social mede o poder de compra de um consumidor e pode explicar a intensidade do relacionamento cliente-empresa.

Bennet e Kassarian (1975) afirmam que uma classe social é um grande número de indivíduos ou famílias aproximadamente iguais uns aos outros e diferenciados dos estranhos em termos de status social. Este status social é definido a partir da combinação de diversas variáveis determinando, assim, a classe social.

As variáveis escolhidas por cada sociedade dependem dos valores mais importantes para essa sociedade (Tomanari, 2003). Segundo Bennet e Kassarian (1975), nos Estados Unidos, a ascensão através da estrutura de classe se dá, normalmente e principalmente, através da educação.

No Brasil, a classificação social é definida principalmente pelos itens de posse da família e pelo grau de instrução do chefe da família, uma vez que os respondentes brasileiros não costumam oferecer dados de renda familiar.

A idade é uma variável demográfica bastante usada na segmentação de mercado, ainda mais que o acelerado crescimento da população idosa no mundo tem suscitado grande interesse em todos os aspectos do processo de envelhecimento, particularmente para os cientistas sociais (Nussbaum *et al.*, 1997). Tal interesse cobre um amplo espectro de fenômenos, entre os quais podem ser citados: qualidade de vida, saúde pública, motivação, problemas de adaptação, aposentadoria e lazer (Neugarten, 1996; Furchtgott e Fuchtgott, 1999).

Há evidência de que, com o avanço da idade, acentua-se a tendência a preferir a TV a materiais impressos (Goodman, 1992; Owsn, 1996) e outras fontes de informação (Hilt, 1997; Scales, 1996). Isto se manifesta nos índices de

audiência em diferentes países obtidos ao longo de décadas de pesquisa. As pessoas idosas são o segmento da população que mais assiste TV diariamente (Grajczyk e Zoelner, 1998). Essas pessoas em idade avançada tendem a consumir um amplo leque de conteúdos, a manifestar uma grande afinidade com a mídia e uma menor propensão a prescindir dela (Comstock, 1989).

### 2.2.3

#### **Variáveis Psicográficas**

Para Kotler, na segmentação psicográfica os compradores são divididos em grupos diferentes baseados no estilo de vida e na personalidade das pessoas. Pessoas de um mesmo grupo geográfico ou demográfico podem exibir perfis psicográficos muito diferentes.

Existem diversos fatores psicológicos que influenciam na decisão de compra e conseqüentemente servem de base para uma segmentação psicográfica. Além de estilo de vida e personalidade, outros autores citam atitude; atividades, interesses e opiniões; valores e percepções.

Sant'Anna (1989) afirma que para alguém tomar a decisão de comprar alguma coisa é preciso que na sua mente se desenvolvam, sucessivamente, os seguintes estados: a existência de uma necessidade, a consciência dessa necessidade, o conhecimento do objeto que pode satisfazê-la, o desejo de satisfazê-la e a decisão por determinado produto ou marca que, a seu ver, melhor satisfará o desejo.

Para analisar os fatores psicológicos de decisão de compra é preciso analisar ainda as teorias da percepção e da motivação.

A percepção está no meio do caminho entre um fluxo contínuo de sentir e pensar e aquilo que percebemos é muitas vezes um produto daquilo que queremos perceber, podendo não corresponder à realidade existente fora de nós (Meys e Reynolds, 1972). Pikunas (1981) completa que o processo de aprendizado de um indivíduo, para conhecer o mundo e a si próprio, começa com a percepção. Ele ainda afirma que “perceber é um processo em que o insumo sensorial é unificado e codificado de acordo com um número crescente de tendências operativas de grupamento e da experiência passada” (Pikunas, 1981, p.100). Meys e Reynolds (1972) afirmam que as mensagens publicitárias são percebidas seletivamente e

que, apesar das várias tentativas de bombardear a mente do consumidor através da mídia, nem todas as informações passam através do filtro perceptivo. Portanto, na elaboração de uma mensagem e na escolha do meio pelo qual ela será passada, também devemos levar em conta dois fatores nos quais está baseada a percepção: fatores de estímulo - características do objeto físico, como tamanho, cor e textura (muitas dessas características determinam o quanto de atenção será produzido); fatores do indivíduo - características próprias do indivíduo, seus processos sensoriais básicos, suas experiências passadas, suas expectativas, motivações e estados afetivos.

A motivação, segundo Cobra (1992), pode ser definida como uma força que move um indivíduo a optar por caminhos de satisfação e necessidade. Ele também divide a motivação da seguinte forma: a) motivos primários: são considerados como sendo as necessidades biológicas que levam as pessoas à sobrevivência; b) motivos secundários: são apontados como necessidades de filiação, aprovação, status, exploração, segurança e realização; c) motivos internos e incentivos: a motivação interna é relacionada a estímulos como a fome, que é oriunda do organismo, enquanto que a motivação, à base de incentivos, vem do desejo ao objeto externo; d) fatores motivadores inconscientes e conscientes: nem sempre os motivos de desejo são conscientes. Quando uma necessidade começa a tornar-se perceptível, estamos entrando no estado da motivação. Festinger (1975) também aponta a dissonância cognitiva como um dos fatores motivacionais. Ele define dissonância como incoerência e cognição como conhecimento. Segundo Meyers e Reynolds (1972, p.95), a “dissonância cognitiva mostra a formação de processos sistemáticos na mente para mudar impressões iniciais e assim reduzir algumas discrepâncias mais significativas”. Um indivíduo em dissonância é aquele que toma atitude adversa àquilo que ele ou a sociedade consideram como correto. Um exemplo citado por Festinger (1975) é o de uma pessoa que adquire um carro de valor muito elevado. Neste caso, o indivíduo terá que procurar motivos que justifiquem esta compra sejam eles a satisfação pessoal, status ou outros.

Existem duas teorias fundamentais que também explicam os fatores da motivação: a teoria behaviorista e a escola cognitiva. Para Lazzaroto e Rossi (1991), um dos conceitos básicos do behaviorismo é o condicionamento. De acordo com a escola behaviorista, o comportamento pode ser reflexo/respondente ou operante. O comportamento reflexo tem como causa um estímulo. A

conjugação estímulo/resposta forma o reflexo. Lazzaroto e Rossi (1991) ainda afirmam que, segundo a escola behaviorista, o comportamento está diretamente vinculado a um processo de condicionamento. As pessoas podem ser mecanicamente condicionadas e conduzidas por um mecanismo de prêmio-castigo.

A escola cognitiva, segundo Cobra (1992), é uma antítese da behaviorista. Isto porque ela enfatiza que o comportamento é dirigido a metas e propósitos e a motivação é vista como uma tendência a mover as pessoas às metas desejadas, enquanto que o behaviorismo defende uma ligação direta entre estímulo e resposta, sem qualquer fator interveniente. A escola cognitiva enfatiza que, ocorrendo o estímulo, uma pessoa busca informações para delinear uma resposta viável.

O aprendizado, para Meyers e Reynolds (1972), pode ocorrer através de informações, sugestões, pensamentos, necessidades e envolve também a formação de associações entre objetos e estímulos. Quando as associações se desenvolvem a partir de repetições, podem levar à formação de hábitos. Cobra (1992) ressalta que a importância prática da teoria do aprendizado é que ela pode estimular a demanda de um produto através da associação. Meyers e Reynolds (1972) também afirmam que a associação produzida por uma marca pode determinar a escolha de uma delas. Segundo estes autores, uma pesquisa realizada com 465 donas-de-casa da área de Chicago constatou indiscutível relação entre as preferências afirmadas, a intenção de compra e a compra real das marcas.

Através de ações e aprendizados, as pessoas adquirem suas crenças e atitudes que acabam por influenciar o comportamento do consumidor (Cobra, 1992). Se alguma dessas crenças estiver errada e inibir a compra, o fabricante do produto precisará fazer uma campanha para corrigir essas crenças.

#### **2.2.4**

#### **Variáveis Comportamentais**

Segundo Kotler (1998), muitas empresas acreditam que as variáveis comportamentais são os melhores pontos de partida para o desenvolvimento de segmentos de mercado. Para o autor, variáveis como ocasiões, benefícios, status de usuário, taxa de uso, status de lealdade, estágios de aptidão de compra e atitude

são variáveis importantes para dividir o mercado em grupos tomando como base o conhecimento do consumidor, sua atitude e o uso e a resposta para um produto.

A segmentação comportamental divide o mercado de acordo com características de conhecimento, lealdade e consumo de produtos. Indivíduos com o mesmo hábito de consumo ou mesmo grau de lealdade, ou mesmo conhecimento com relação a uma categoria de produto/serviço formam um segmento (Tomanari, 2003).

As variáveis comportamentais podem servir de base para uma segmentação de mercado bastante eficiente. Clientes com comportamentos similares tendem a apresentar características iguais, explicadas pelas variáveis:

- a) Ocasão de compra (normal, especial)
- b) Status de uso (não usuário, ex-usuário, usuário potencial, usuário pela primeira vez, usuário regular)
- c) Grau de uso (*light user*, *médium user*, *heavy user*)
- d) Modo de utilização
- e) Razão de compra (presente, reposição, festa)
- f) Frequência de compra
- g) Local de compra
- h) Estágio de atenção para a compra (desconhece, atento/conhece, interessado, desejoso, com intenção de compra)
- i) Atitude com relação ao produto (entusiasta, positiva, indiferente, negativa, hostil).

A segmentação comportamental pode ser realizada através dos resultados dos estudos de hábitos e atitudes. Segundo Galvão (2002), os hábitos e as atitudes são instrumentos de coleta de dados demográficos além das necessidades e motivos de consumo do produto, além da quantidade consumida. Com essa informação, consegue-se separar os *heavy users* e constatar sua participação no universo de consumidores. Concentrar os esforços de marketing nos *heavy users* é uma forma de segmentação embora não a denominem desta forma.

Para Gama (2002), pode-se encontrar no mercado de consumo variados comportamentos e tipos de consumidores. O autor destaca cinco grupos de consumidores e como se comportam frente aos seus direitos no mercado, são eles

os alienados, os tolerantes, os responsáveis pacíficos, os responsáveis exigentes e os renitentes.

Cada grupo apresenta diferentes níveis de consciência do que é ser consumidor e age de forma diferente na sua relação com a empresa.

- Alienados: é o maior grupo do mercado, representa a maior parte da população, são aqueles que aceitam sem filtro o que acontece nas relações de consumo, acreditando que tudo é sorte ou azar do destino;
- Tolerantes: têm consciência dos seus direitos, mas se desvalorizam diante das dificuldades, porque acreditam que não adianta reclamar;
- Responsáveis Pacíficos: também têm consciência dos seus direitos, mas acreditam que reclamar é falta de tempo e de dinheiro;
- Responsáveis Exigentes: querem ser respeitados e por isso reclamam e não aceitam desculpas sem fundamento, pois têm plena consciência de seus direitos;
- Renitentes: têm consciência dos seus direitos e reagem contra os danos que sofrem, ainda dão real valor ao seu dinheiro.

Para Howard (1989), o comportamento de compra dos consumidores é uma base fraca para a segmentação. Segundo o autor, mesmo a atitude de compra estando intimamente relacionada ao comportamento, tanto o comportamento de compra como a atitude podem mudar e, com isso, parecem não ser confiáveis. Além disso, o comportamento de compra de um cliente é frequentemente influenciado por fatores além da atitude, como preço e disponibilidade do produto, o que torna o comportamento de compra uma base fraca para segmentar os consumidores.

Schultz (2002) defende uma segmentação comportamental dinâmica que se ajuste às mudanças de vida dos consumidores.

## 2.3

### Variáveis Transacionais

As transações são de grande importância para a geração de valor de uma organização. Entende-se por transação qualquer interação que o cliente tenha com

uma empresa que pode ser classificada em termos de canal de contato, forma de pagamento, tipo de produto adquirido e tempo de relacionamento, que explicarão o valor financeiro que cada consumidor gera ao longo do tempo.

Os clientes não são iguais e expressam essa diversidade através de transações diferentes, que geram custos maiores ou menores para a empresa e impactam de maneira distinta na geração de caixa de uma companhia.

Um cliente pode gerar mais valor do que outro de acordo com a forma que interage com a empresa. Por isso, diversos autores (Zajac e Olsen, 1993; Hunt e Morgan, 1995, 1996; Srisvastava *et al.* 1998; Gosh e John, 1999) têm se preocupado com o impacto das atividades e estratégias de marketing na geração e apropriação de valor para a empresa e seus acionistas.

Adicionalmente, Srivastava (1999, p.168) observa que, se o marketing contribui para a renovação e o crescimento dos negócios e para a obtenção e a retenção de consumidores, ele precisa resultar em fluxos de caixa melhorados – um requisito central para a geração de valor ao acionista.

O processo de obtenção e retenção de consumidores e o impacto na geração de caixa ocorrem por meio da transação, considerada o elemento focal do contato entre dois agentes econômicos (Dwyer *et al.*, 1987).

A capacidade de apropriação do valor transacional deve ser função de uma série de variáveis, como o tipo de relacionamento entre as partes, o modo como elas estruturam o ambiente da transação e a capacidade individual de criação de valor (Dwyer *et al.*, 1987).

Torna-se relevante a compreensão das variáveis que influenciam a capacidade que um cliente tem de gerar valor para a organização. Williamson (1991) defende que as firmas sejam estruturas de governança formadas com o objetivo de economizar em custos de transação gerados pelo mercado.

A criação de valor para o consumidor é alcançada pela obtenção de posições favoráveis como resultado da gestão e da alavancagem de ativos baseados no mercado (Srivastava *et al.*, 1998), ou seja, o objetivo da obtenção e da retenção de consumidores é a criação de valor financeiro para a empresa (Souza, 2002).

Segundo Srivastava *et al.* (1998), a alavancagem destes ativos pela empresa pode ocorrer com o propósito de: a) diminuir os custos através do conhecimento dos canais e dos consumidores ou pela construção de relacionamentos; b) obter de prêmios em preços, já que o patrimônio da marca conduz à percepção de maior

valor agregado; c) gerar de barreiras competitivas pela construção de lealdade e criação de custos de mudança (Gosh e John, 1999); d) criar um limite competitivo; e) possibilitar de extensão da marca e de categorias de produtos.

Uma alternativa para a economia de custos de transação é o arranjo dos canais de marketing. Esses arranjos são estruturas que, de acordo com a sua administração e controle, podem caracterizar-se pela verticalização ou terceirização (Souza, 2002). A teoria do agente (Brickley *et al.*, 1991; Krueger, 1991; Bergen *et al.*, 1992; Lafontaine, 1992; Gupta e Romano, 1998; Combs e Ketchen Jr., 1999) estuda os fenômenos de custos gerados quando uma empresa titular concede a um agente o direito de representá-lo na administração dos recursos e capacidades da firma.

Embora conteste-se que o único objetivo de uma empresa seja a maximização do lucro (Ouchi, 1979), não se pode negar que se este não for o objetivo principal, é, pelo menos, o indicador fundamental de desempenho utilizado pelo mercado financeiro para avaliar a qualidade da administração de uma empresa.

Em função da estrutura de custos transacionais de uma empresa, Williamson (1979) defende que existe um elenco de oportunidades de redução desses custos.

Os bancos, por exemplo, acompanham o rastro de informações deixadas por um cliente. Utilizando sistemas que processam bilhões de pequenas pistas, como emissão de cheques, investimento em fundos, saldo em poupança, além de variáveis sócio-demográficas, como ocupação, renda, presença de filhos na família, e seu histórico de compras e de pagamentos, conseguindo assim identificar clientes com alta probabilidade de adquirir um produto de previdência privada (Stephan, 2000). Ao aumentarem a taxa de acertos em encontrar clientes mais propensos a comprar determinado produto, reduzem seus custos de aquisição que podem variar de acordo com o canal de venda, a forma de pagamento, a oferta de vendas, promoção e o tipo de produto.

Um dos usos mais avançados de dados transacionais é o gerenciamento de risco de crédito por portadores de cartões e tomadores de empréstimo. Os grandes bancos vêm aperfeiçoando seus modelos de *score* de risco desde o final dos anos 80. Utilizando variáveis transacionais como frequência de compra, data da última compra, valor e tipo de produto, histórico de pagamentos e manutenção de saldos negativos, os bancos conseguem antecipar problemas de crédito futuros (Souza,

2002). Uma outra variável transacional muito importante é o tempo de permanência. Para Stephan & Grion (2003), os novos clientes têm uma probabilidade muito maior de abandonar uma empresa do que um cliente antigo. Isto pode ocorrer porque um cliente quando decide comprar de uma empresa pela primeira vez está iniciando uma transição em sua vida. Essas pessoas podem ter necessidades e expectativas diferentes dos clientes mais antigos. Desta forma, acompanhar sua evolução no primeiro ano de forma diferenciada é um esforço que pode trazer resultados excepcionais.

O uso de informações transacionais dos clientes para guiar as decisões de marketing deve ser adotado por todas as empresas para redirecionar seus esforços e obter maior efetividade nas ações estratégicas da empresa com o objetivo de aumentar a rentabilidade dos clientes, alavancando os fluxos de caixa da organização, através da redução dos custos de transação ou pelo aumento do faturamento de cada cliente.

## 3

### Metodologia

#### 3.1

##### Tipo de Pesquisa

O principal objetivo de uma pesquisa é contribuir para o conhecimento, endereçando algumas das muitas perguntas não respondidas. Para fazê-lo de forma satisfatória e para avaliar se houve contribuição significativa para a coleção do conhecimento, um pesquisador deve cumprir o método científico (Remenyi *et al.*, 1995).

O método científico é um conjunto de regras informal, mas rigoroso, que foi desenvolvido para garantir a integridade, a confiabilidade e a reprodutibilidade do trabalho (Remenyi *et al.*, 1995).

A escolha da metodologia está intimamente ligada à forma de observação do objeto de estudo. Segundo Remenyi *et al.* (1995), a essência do conhecimento científico é que ele é derivado das observações realizadas no mundo. Para o autor, tudo o que não pode ser observado, direta ou indiretamente, através de seus efeitos ou conseqüências, está fora do domínio da ciência.

De acordo com Remenyi (2002), a pesquisa científica pode ser classificada em dois grupos distintos: o teórico e o empírico. Para Creswell (1994), os trabalhos podem ser considerados teóricos quando não se fundamentam em dados empíricos produzidos ou analisados diretamente pelo pesquisador e/ou se propõem a discutir conceitos teóricos e/ou propor uma nova teoria em administração. Já os trabalhos empíricos utilizam dados coletados no campo ou em bases já existentes, e que utilizarão esses dados para um experimento ou hipótese. O empírico pode ser subdividido pelo tipo de pesquisa qualitativa e quantitativa, sendo que as duas perspectivas se propõe a testar uma teoria ou hipótese.

Seguindo o critério do autor, este trabalho pode ser classificado como empírico, com aplicação da metodologia quantitativa. Esta pesquisa é empírica porque se propõe a determinar a influência das variáveis demográficas, dos dados

do histórico de eventos e de transação dos clientes sobre o risco de cancelamento, utilizando uma base de dados já existente para definir o perfil do cliente que cancela sua assinatura de jornal. Quantitativa, porque a partir desta base de dados foi aplicado um modelo de regressão logística binário para prever o risco de desligamento dos clientes assinantes de jornais.

Este estudo se propõe a responder à seguinte pergunta: é possível definir o perfil do cliente que cancela uma assinatura de jornal?

Remenyi (2002) afirma ainda que na escolha da metodologia, a revisão da literatura deve ser analisada junto com alguns fatores. Esta revisão de literatura deve revelar um modelo de pesquisa e uma metodologia adequada e que já tenha sido aplicada antes a esse tipo de questão. Essa opção é consistente com a estrutura adotada nesta fundamentação, que serve de base para a elaboração de um modelo teórico antes de definir as variáveis que devem compor o modelo matemático de previsão de cancelamento.

## 3.2

### Universo e Amostra

O universo da pesquisa é composto pelos clientes, ativos e inativos, registrados no banco de dados da empresa editora de jornal. A base de dados possui todas as assinaturas que estejam ativas, incluindo as vendidas desde o momento em que a empresa lançou o serviço de assinaturas em abril de 1976. Para as assinaturas inativas existe registro das assinaturas canceladas desde julho de 1995.

Os consumidores de assinaturas podem ser do tipo pessoa física e pessoa jurídica, mas a amostra foi selecionada considerando apenas a população de clientes do tipo pessoa física.

Foi definida uma janela de análise. Foram analisadas assinaturas canceladas entre junho de 2004 e maio de 2005 e assinaturas ativas em maio de 2005. De um total de 230.858 assinaturas ativas e 105.524 assinaturas canceladas extraiu-se duas amostras aleatórias simples, uma de treinamento com 35.549 casos e outra amostra para validação com 4.796 casos.

### 3.3

#### Coleta dos Dados

A coleta dos dados foi realizada da seguinte maneira:

- a. pesquisa bibliográfica em livros, dissertações, *papers* e periódicos, para conceituar marketing de relacionamento, satisfação, lealdade e retenção de clientes, variáveis transacionais, segmentação de mercado, variáveis geográficas, demográficas, psicográficas e comportamentais, além de métodos estatísticos para tratamento e modelagem dos dados do histórico de eventos dos assinantes. Além de levantar as variáveis do modelo teórico de previsão de cancelamento;
- b. pesquisa documental na base de dados fornecida pela empresa editora de jornais observando três grupos de variáveis: histórico de eventos, dados transacionais e variáveis demográficas. A unidade de medida do tempo empregada é o mês que é uma variável contínua e a unidade de análise é o assinante pessoa física.

### 3.4

#### Limitações do Método

Uma grande limitação deste estudo refere-se à coleta dos dados. As variáveis selecionadas estavam pré-definidas no banco de dados. Com isso, variáveis importantes dos clientes podem ter sido desconsideradas.

Outra restrição é o pequeno número de trabalhos publicados sobre a utilização de modelos de regressão logística para prever o risco de cancelamento de clientes.

O modelo desenvolvido na dissertação é único para a empresa analisada, portanto, este modelo não pode ser generalizado para outras empresas deste ou de outro setor.

### 3.5

#### Tratamento dos Dados

Foi escolhida uma amostra aleatória simples com 40.345 assinaturas de pessoas físicas de uma população de 1.765.221 assinaturas. Seguindo a proporção populacional, esta amostra é composta por 69% de assinantes ativos e 31% de inativos. Desta amostra de 40.345 assinaturas foi extraída uma nova amostra aleatória com 4.796 assinaturas, 35.549 assinaturas foram utilizadas para o desenvolvimento e aplicação do modelo e 4.796 assinaturas utilizadas para a validação do modelo que identifica o risco de cancelamento.

A técnica utilizada para indicar o risco de cancelamento dos assinantes deste jornal é a regressão logística binária que determina a probabilidade de uma assinatura pertencer ao grupo dos que cancelam ou ao grupo dos que não cancelam.

Tanto a variável dependente quanto as variáveis independentes são categóricas.

O tratamento estatístico dos dados e o desenvolvimento do modelo foram realizados com o apoio do SPSS 13.0.

#### 3.5.1

##### Problema e Objetivo

Cada vez as empresas estão mais preocupadas em otimizar seu investimento em marketing, direcionando ações diferenciadas para públicos específicos. A segmentação passou a ser uma prática comum para se conhecer os clientes e implementar estratégias mais adequadas e, conseqüentemente, mais eficazes.

Diante disso, identificar o público mais propenso ao cancelamento torna-se necessário para se conhecer o tamanho do risco de perda na carteira de clientes de uma empresa e também para desenvolver ações de retenção específicas para este público.

Este estudo propõe um modelo de regressão logística que diferencia o grupo de clientes que cancela do grupo que não cancela. Sendo possível assim, conhecer o perfil dos clientes canceladores e desenvolver ações de retenção que aumentem

o tempo de permanência destes clientes. Diante disso, o objetivo é definir o perfil do cliente que cancela a assinatura do jornal.

### 3.5.2

#### O Método de Regressão Logística

Na literatura, existem duas técnicas com a capacidade de separar dois grupos ou alocar um novo elemento em um desses grupos. Esta é uma situação enfrentada por muitos pesquisadores. Ambas as técnicas, Análise Discriminante e Regressão Logística, se enquadram na classe de métodos estatísticos multivariados de dependência, pois relacionam um conjunto de variáveis independentes com uma variável dependente categórica (Sharma, 1996; Hair *et al.*, 1998; Morgan e Griego, 1998).

As técnicas de discriminação buscam uma função ou conjunto de funções que discrimine os grupos definidos pela variável categórica, visando a minimizar erros de classificação. Em contexto no qual o conjunto de variáveis independentes possui comportamento probabilístico de normalidade multivariada, a análise discriminante é adequada, porque minimiza os erros de classificação (Sharma, 1996; Hair *et al.*, 1998). Portanto a suposição de normalidade multivariada é necessária para que os resultados da análise discriminante sejam satisfatórios.

O modelo logístico é mais comumente utilizado porque a análise discriminante impõe às variáveis independentes condições como: serem normalmente distribuídas e terem suas matrizes de variância-covariância iguais entre os dois grupos de classificação. Outro ponto que é motivo de crítica quanto ao método da análise discriminante é que o resultado da expressão discriminante fornece um *score* que possui pouca interpretação intuitiva (Castro, 2003). Segundo Ohlson (1980, p.112), este *score* é basicamente um dispositivo (discriminatório) de classificação ordinal, não tendo embutido nenhum aspecto probabilístico. Além de não depender da exigência de normalidade das variáveis independentes e da igualdade de matrizes de covariância, a regressão logística é semelhante a uma regressão múltipla, pois possui o poder de incorporar efeitos não lineares (Hair, 1998). A possibilidade de interpretação direta dos coeficientes como medidas de associação é uma das grandes vantagens da regressão logística;

a interpretação destes coeficientes pode ser estendida para qualquer problema prático (Paula, 1999)

Hair (1998) afirma que existem algumas razões pelas quais a regressão logística representa uma alternativa atraente à análise discriminante sempre que a variável dependente tiver somente duas categorias. Em primeiro lugar, a regressão logística é menos afetada pelas desigualdades variância/covariância dentre os grupos, um pressuposto básico da análise discriminante. Em segundo lugar, a regressão logística pode cuidar facilmente de variáveis categóricas independentes, enquanto que na análise discriminante o uso de variáveis *dummy* criou problemas com as igualdades variância/covariância. Finalmente, a regressão logística é similar à regressão múltipla em termos de sua interpretação e nas medidas de diagnóstico direcionadas a casos disponíveis para o exame de resíduos.

Segundo Hosmer e Lemeshow (1989), a técnica de regressão logística tornou-se um método padrão de análise de regressão para variáveis medidas de forma dicotômica, especialmente nas áreas de ciência da saúde. Muitas situações em análise de dados envolvem prever o valor de uma variável de resultado categórico. Essas situações incluem aplicações na medicina prevendo o estado de saúde de um paciente, na pesquisa de marketing prevendo se uma pessoa irá comprar o produto ou em escolas prevendo o êxito de um aluno. A regressão logística é uma técnica que pode ser muito útil nessas situações ou em muitas outras. Mesmo quando a resposta de interesse não é originalmente do tipo binário, alguns pesquisadores têm dicotomizado a resposta de modo que a probabilidade de sucesso possa ser modelada através da regressão logística (Paula, 1999). O mesmo modelo pode ser utilizado com enfoque discriminatório, conforme descrevem Krzanowski (1988) e McLachlan (1992). Esses autores defendem o modelo logístico de discriminação como um método utilizado de forma mais abrangente, pois não faz suposições quanto à forma funcional das variáveis independentes, e o número de parâmetros envolvidos no processo de estimação provavelmente será menor.

Ao contrário da análise discriminante, a regressão logística exige um número menor de pressupostos que são menos rígidos, são eles: as variáveis independentes devem ser intervalo, taxa ou dicotômicas; todos os previsores relevantes foram incluídos, nenhum predictor irrelevante foi incluído e a forma do relacionamento é linear; o valor esperado por termo de erro é zero; não há

autocorrelação; não há correlação entre o erro e as variáveis independentes; há uma ausência de multicolinearidade perfeita entre as variáveis independentes (SPSS, 2003).

Além disso, a regressão logística funciona quase tão bem quanto a análise discriminante mesmo quando os pressupostos desta são satisfeitos. Comparando as duas técnicas, Krzanowski (1988) diz que é consenso geral que a discriminação logística deve ser preferida, quando as distribuições são claramente não-normais. A mesma afirmação é sustentada por Press e Wilson (1978). Hair *et al.* (1998) apontam uma lista de motivos que levariam o pesquisador a optar pela regressão logística: não é necessário supor normalidade multivariada; é uma técnica mais genérica e mais robusta, pois sua aplicação é apropriada em muitas situações e é uma técnica similar à regressão linear múltipla.

Outra vantagem da regressão logística é sua abordagem probabilística, já que essa regressão estima a chance de ocorrer um certo evento a partir de uma série de variáveis independentes ou explanatórias. Segundo Hower e Lemeshow (1989), o objetivo da regressão logística é achar o melhor relacionamento entre a variável resposta (de saída ou dependente) e um conjunto de variáveis explanatórias ou preditivas, sendo o modelo final aquele que apresentar o melhor ajuste matemático e for naturalmente razoável de se explicar. A regressão logística é projetada para utilizar uma combinação de variáveis previsoras contínuas e categóricas para prever uma variável de resultado categórico ou dependente.

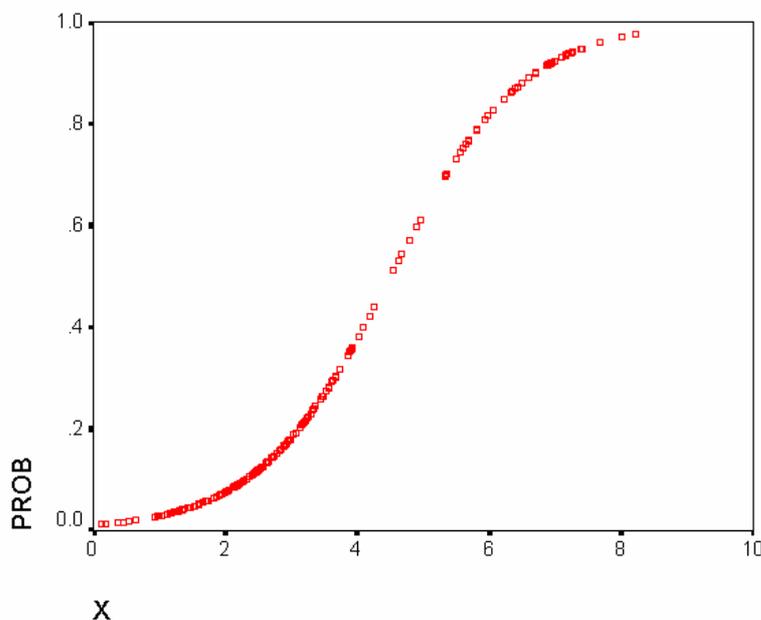
Hair (1999) defende que a regressão múltipla não pode ser usada para prever os valores de uma variável dependente dicotômica, pois, quando tentamos prever os valores de uma variável codificada, digamos, 0 ou 1, podemos considerar os valores estimados como sendo probabilidades, isto é, a probabilidade de obter um valor previsto de 1. Numa regressão múltipla com um ajuste de linha reta aos dados, este é freqüentemente o caso em que valores inferiores a 0 ou superiores a 1 são previstos. Ou seja, os modelos de regressão linear não são apropriados para se analisar a relação de diversos fatores e a probabilidade de ocorrência de um evento. Isto ocorre, porque o valor esperado da variável resposta, dado um conjunto de variáveis explicativas, pode assumir valores que vão de menos infinito até mais infinito e isto viola as leis de probabilidade. Além disso, um dos pressupostos da regressão é a homogeneidade da variância. Entretanto, para uma

variável dicotômica, o desvio médio e padrão são relacionados porque o desvio padrão está, onde  $p$  é a média da variável. Já que existe um relacionamento funcional entre o desvio padrão e a média, a homogeneidade de variância dentre os valores da variável dependente não pode ser satisfeita.

A regressão logística foi desenvolvida nos anos 60 como uma solução para este problema. Segundo Hair (1999), a regressão logística é uma técnica robusta e bem adequada quando há violação do pressuposto de igualdade das matrizes de variância/covariância dentre os grupos.

Ao predizer o valor de uma variável numa escala de 0 a 1, faz sentido ajustar uma curva em forma de S aos dados.

**GRÁFICO 4: Curva Logística**



A função logística está limitada em 0 e 1, de modo que previsões impossíveis não podem ocorrer. Há efetivamente uma família inteira de funções em forma de S, sendo o *probit* uma variante bem conhecida. Devido a várias considerações, a maior parte dos estatísticos concorda que a logística é um modelo opcional para a regressão com uma medida dependente dicotômica (SPSS, 2003).

O primeiro passo é verificar se os dados foram coletados em escala contínua ou discreta. Para situações tempo-discreto, Allison, P. (1984) sugere o uso do modelo *logit* e do método *maximum likelihood* para estimação dos parâmetros do

modelo. Se todas as variáveis explicativas forem categóricas, a estimação do modelo *logit* pode ser feita por método log-linear.

Para Moore (1994), a análise *logit* é um método que determina quais variáveis independentes devem ser incluídas no modelo para se prever adequadamente a variável dependente categórica. Como fator limitador para uso da análise *logit*, as variáveis dependentes e independentes devem ser categóricas, obrigatoriamente. Ott e Markewich *apud* Moore (1994) afirmam que a regressão logística é o método mais apropriado de análise quando as variáveis independentes são contínuas.

A regressão logística reescreve o modelo clássico de regressão linear de modo a confirmar o valor da variável resposta para a faixa de 0 a 1, ao mesmo tempo em que as variáveis independentes possam variar continuamente. Isto é obtido pela equação abaixo, também conhecida como função logística.

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

que pode ser linearizada pela transformação:

$$g(x) = \ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \beta_1 X$$

A regressão logística binária é a regressão aplicada a uma variável dependente dicotômica, onde a variável dependente não representa os valores de dados brutos, mas representa a probabilidade do evento estudado ocorrer. A equação geral para regressão logística é:

$$\ln(Odds) = \alpha + B_1 X_1 + B_2 X_2 + \dots + B_k X_k$$

onde os termos da direita são os termos padrão para as variáveis independentes e o intercepto numa equação de regressão. Entretanto, do lado esquerdo está o log natural da probabilidade e a quantidade  $\ln(Odds)$  é chamada de *logit* e pode variar

de menos até mais infinito. Portanto retirando o problema de predição para fora dos limites da variável dependente. As probabilidades são relacionadas por:

$$Odds = \frac{prob}{1 - prob}$$

Na regressão logística há um relacionamento linear com as variáveis independentes, mas é linear nas probabilidades de log e não nas probabilidades originais. Como o objeto de estudo é a probabilidade de ocorrência de um evento, a equação logística pode ser transformada numa equação na probabilidade (Hair, 1999). Assume então esta forma:

$$prob(event) = \frac{1}{1 + e^{-(\alpha + B_1 X_1 + B_2 X_2 + \dots + B_k X_k)}}$$

Diferente da regressão linear clássica, os erros desse modelo não seguem uma distribuição normal, mas sim a de Bernoulli. Assim, enquanto na regressão linear o método usado para estimar os coeficientes  $\beta_0, \dots, \beta_n$  é o método dos mínimos quadrados, na regressão logística usa-se o método da máxima verossimilhança. Este último método encontra os valores dos parâmetros  $\beta_0, \dots, \beta_n$ , que maximizem a probabilidade de se obter o conjunto observado de dados (Hosmer e Lemeshow, 1989).

Ao utilizarmos regressão logística, a principal suposição é a de que o logaritmo da razão entre as probabilidades de ocorrência e não ocorrência do evento é linear. Por essa razão, ao interpretar os coeficientes da regressão logística, interpreta-se  $e$  e não  $\ln$ . Contudo, quando se utiliza o modelo logístico do ponto de vista de discriminação entre grupos, não há grande interesse na interpretação dos coeficientes (Garson, 2000).

A regressão logística é estimada de forma bem semelhante à regressão múltipla pelo fato de que um modelo de base é primeiro estimado visando a fornecer um padrão para comparação. Na regressão múltipla, a média é utilizada para estabelecer o modelo base e calcular a soma total dos quadrados. Na regressão logística, o mesmo processo é utilizado, com a média usada no modelo

estimado não para estabelecer a soma dos quadrados, mas para estabelecer o valor de probabilidade log. Neste modelo, as correlações parciais para cada variável podem ser estabelecidas e a variável mais discriminante pode ser escolhida de acordo com o critério de seleção (Hair, 1998).

Para definir o ponto de corte é necessário conhecer a probabilidade *a priori* de um assinante cancelar a sua assinatura. Com isso, o ponto de corte para este estudo deve ser o valor que minimize os erros de classificação dos assinantes (erros Tipo I e Tipo II). O maior problema enfrentado pelos profissionais envolvidos é a obtenção do valor de corte. A grande questão é como obter um valor de corte confiável a ponto de evitar perdas para a empresa, tanto pela não classificação de risco para clientes que podem cancelar e, conseqüentemente, não serão impactados pelas ações de marketing, quanto pela atribuição de risco para clientes sem propensão ao cancelamento e que consomem parcelas importantes dos investimentos de marketing utilizados na retenção de clientes. Analisada a base de clientes dos últimos 12 meses percebe-se que a proporção entre ativos e cancelados é de 69% de clientes ativos. Conseqüentemente, define-se o ponto de corte com sendo 0,69.

O modelo de regressão logística minimiza o número de variáveis para que o modelo resultante seja mais facilmente generalizado e mais estável numericamente, dado que quanto mais variáveis são incluídas no modelo, mais ele se torna dependente dos dados. O uso da técnica *stepwise* na regressão logística é o processo de inclusão ou exclusão de variáveis do modelo, baseado em critérios tais com a estatística G e o teste Wald.

Existem os métodos *enter*, *backward* e *forward*. O método *enter* incorpora ao modelo todas as variáveis e deve ser utilizado principalmente quando se tem certeza de que todas as variáveis são necessárias para se estimar os parâmetros do modelo. O método *backward* caracteriza-se por incorporar todas as variáveis e após percorrer várias etapas, uma variável por vez pode ser eliminada. Se em uma determinada etapa não houver eliminação de alguma variável, o processo é então interrompido e as variáveis restantes definem o modelo final. Numa dada etapa, temos um determinado modelo que denominamos modelo completo da etapa e são investigadas as contribuições individuais das variáveis a esse modelo. A variável de pior desempenho é eliminada comparando-se o modelo completo com o modelo reduzido, pela retirada de tal variável (Charnet *et al*, 2000).

O método *forward* caracteriza-se por considerar a variável de maior coeficiente de correlação amostral observado com a variável resposta. A cada etapa, uma variável pode vir a ser incorporada. Se em uma etapa não houver uma inclusão, o processo é interrompido e as variáveis selecionadas até esta etapa definem o modelo final (Charnet *et al.*, 2000).

Em uma determinada etapa chega-se a um modelo definido como modelo reduzido. O modelo reduzido de cada etapa é comparado ao modelo em que uma nova variável é acrescentada. Existindo um modelo de melhor desempenho, a correspondente variável é incorporada ao elenco de variáveis já escolhidas. Enquanto em uma etapa do procedimento *backward* comparam-se vários modelos reduzidos com um único modelo completo devido ao objetivo de eliminar uma variável, em uma etapa do procedimento *forward* as comparações são feitas entre vários modelos completos e um único modelo reduzido, graças ao objetivo de incorporar uma variável (Charnet *et al.*, 2000).

Os três métodos diferem pela forma que utilizam para selecionar uma variável. Os métodos direcionados a passos utilizam a estatística Wald, a mudança na probabilidade ou a estatística condicional como método de escolha para a eliminação ou inclusão de variáveis. (SPSS, 2003). O método escolhido foi o *forward conditional* que utiliza a estatística condicional para incluir variáveis ao modelo é menos intensivo computacionalmente.

Para a obtenção do modelo final de regressão logística, após estimar os coeficientes da equação de regressão, é necessário verificar se cada variável é significativamente relacionada com a variável resposta do modelo. Isto é realizado geralmente através da formulação de testes de hipóteses estatísticas, que avaliam o modelo com a variável e sem a variável (Hower e Lemeshow, 1989).

Há dois testes estatísticos para a significância do modelo final. Primeiro, um teste *chi-quadrado* para mudança no valor  $-2LL$  do modelo base é comparável com o teste F geral em regressão múltipla. Além disso, a medida Hosmer e Lemeshow de ajuste geral tem um teste estatístico que indica se houve ou não diferença estatisticamente significativa entre as classificações observadas e previstas. Estas duas medidas, em combinação, fornecem suporte para que se aceite o modelo de regressão logística como significativo. Estes testes asseguram a evidência de significância estatística das variáveis, devendo se considerar outros

relevantes fatores como a importância da variável em relação ao evento modelado e a influência conjunta de outras importantes variáveis (Hair, 1998).

Segundo Hair (1998), o ajuste geral do modelo pode ser avaliado utilizando-se algumas medidas como o  $-2LL$ . Se no modelo avaliado houver um decréscimo no valor  $-2LL$  comparado ao modelo base, existe melhora no modelo, pois valores menores da medida  $-2LL$  indicam o melhor ajuste do modelo. A seguir, as medidas de adequação de ajuste comparam as probabilidades estimadas com as probabilidades observadas sendo que os valores mais altos indicam um ajuste melhor. Existem ainda três medidas comparáveis ao  $R^2$  no modelo de regressão múltipla. A medida  $R^2$  Cox e Snell opera da mesma forma, com valores mais altos indicando maior ajuste do modelo. Entretanto, esta medida está limitada pelo fato de que não consegue alcançar o valor máximo de 1, de modo que Nagelkerke propôs uma modificação que tem o alcance de 0 para 1. A terceira medida é a medida  $R^2$  "pseudo" com base na melhoria do valor  $-2LL$ . O pseudo  $R^2$  é calculado como:

$$R_{\log it}^2 = \frac{2LL_{base} - (-2LL_{modelo})}{-2LL_{base}}$$

A medida final do ajuste do modelo é o valor Hosmer e Lemeshow, que mede a correspondência dos valores efetivos e previstos da variável dependente. Neste caso, o melhor ajuste do modelo é indicado por uma diferença menor na classificação observada e prevista. Um bom ajuste de modelo é indicado por um valor *chi-quadrado* não significativo.

A medida Hosmer e Lemeshow ainda mostra a não-significância, indicando a ausência de diferença na distribuição de valores dependentes efetivos e previstos.

Finalmente, as matrizes de classificação, idênticas em natureza às utilizadas na análise discriminante, mostram se as taxas de acerto são altas ou baixas para os casos corretamente classificados no modelo.

A validação do modelo de regressão logística pode ser obtida através do mesmo método utilizado na análise discriminante: criação de amostras de treinamento e validação (Hair, 1998). Se as taxas de acertos da amostra de treinamento e da amostra de validação forem similares, pode-se concluir que o

modelo tem suporte empírico no mesmo nível para explicar as variáveis dependentes.

### 3.5.3

#### Coleta e Análise dos Dados

A amostra inicial contava com 36.000 assinaturas. Após a análise da base de dados, foram excluídos os assinantes com as seguintes características:

- a. assinaturas com forma de pagamento “não informado”;
- b. assinaturas com tipo de assinatura “em branco”;
- c. assinantes com idade inferior a 18 anos e superior a 90 anos.

Depois desse tratamento, a amostra ficou com 35.549 assinaturas, sendo 24.492 ativas e 11.057 canceladas. A base final para estudo contém as variáveis: status (ativa ou inativa), tipo de assinatura, fonte de venda, forma de pagamento, região, tempo de permanência, LTV, SD&W, gênero, faixa etária, indicativo de participação em ações de fidelização, publicação de anúncio nos últimos 12 meses, histórico de reclamação, compra de produto agregado, quantidade de produtos agregados comprados.

No banco de dados da empresa editora de periódicos a proporção é de 31% de inativos e 69% de ativos.

### 3.5.4

#### Definição das Variáveis

Em uma primeira etapa foram consideradas as seguintes variáveis:

<b>Variáveis Transacionais</b>		
<b>Descrição</b>	<b>Definição</b>	<b>Níveis</b>
Status do cliente	Indica se o cliente continua assinando o jornal ou se cancelou o serviço.	Ativo e Inativo
Tipo de Assinatura	Período pelo qual o cliente contratou a assinatura	Anual; boleto mensal; débito automático; semestral e trimestral.
Fonte de Venda	Representa o canal de vendas pelo qual o cliente adquiriu uma assinatura.	Agente; antigo (na migração do sistema os dados foram convertidos como antigo); Internet; outros; sem fonte; terceiros; telemarketing ativo; telemarketing receptivo.
Forma de pagamento	Meio de pagamento utilizado pelo cliente	Boleto bancário; cartão de crédito no débito automático; cartão de crédito parcelado; débito em conta corrente automático; débito em conta corrente parcelado; não informada.
Tempo de permanência	Quantidade de meses que o cliente ficou ativo no caso dos cancelados; ou que ainda está ativo, no caso dos atuais clientes	De 0 até 3 meses; de 3 até 6 meses; de 6 meses até 1 ano; de 1 ano até 2 anos; de 2 anos até 3 anos; mais de 3 anos.
Quantidade de Produto Agregado	Indica o volume de produtos comprados pelos clientes	Até de 3 produtos; entre 4 e 7 produtos; entre 8 e 10 produtos e mais de 10 produtos
LTV	Score de segmentação dos clientes por <i>life time value</i> que é igual ao tempo de permanência atual em meses+ tempo de vida futuro em meses (estimado por análise de sobrevivência) x margem (R\$) da assinatura mensal do cliente.	Platina; Diamante; Ouro; Prata; Bronze; Lata; Indefinido.

Indicador de compra de produto agregado	Indica se o cliente já comprou ou não um produto agregado (livros, guias, CDs e outros colecionáveis)	Sim ou Não
Indicador de compra de anúncio	Indica se o cliente já comprou um anúncio nos classificados ou não	Sim ou Não
Indicador de Reclamação	Indica se o cliente já fez uma reclamação através do <i>call center</i> ou pela internet	Sim ou Não

### Variáveis Geográficas

Descrição	Definição	Níveis
Região	Bairro ou cidade onde o assinante recebe o jornal agrupado por similaridade nos custos de distribuição	Baixada, subúrbio, Barra, Zona Norte, Zona Sul, Brasília, interior do RJ, Niterói, outros estados, São Paulo/Belo Horizonte

### Variáveis Demográficas

Descrição	Definição	Níveis
Gênero	Indica o sexo do assinante; alguns assinantes não têm o campo preenchido	Feminino; masculino e indefinido
Faixa Etária	Agrupamento que indica a idade do assinante titular	menos de 20 anos; de 20 a 25 anos; de 26 a 30 anos; de 31 a 35 anos; de 36 a 40 anos; de 41 a 50 anos; de 51 a 60 anos; de 61 a 80 anos; mais de 80 anos; sem preenchimento
SD&W	Score de classificação sócio-demográfica utilizando variáveis do senso do IBGE para os setores censitários do Rio de Janeiro	Solidez; segurança; equilíbrio em conquista; estabilidade; sobrevivência; privação; penúria; subsistência e indefinido.

### Variáveis de Fidelização

Descrição	Definição	Níveis
Indicador de participação em ações de fidelização	Indica se alguma vez o cliente participou das ações de relacionamento tais como vantagens, descontos ou participação em eventos exclusivos.	Sim ou Não

A SD&W é uma empresa de consultoria que desenvolve um modelo que utiliza variáveis sócio- demográficas do censo para classificar a população em 9 segmentos distintos de acordo com as características destes indivíduos.

O objetivo deste modelo é desenvolver indicadores (ou scores) capazes de traduzir informações tais como classe de renda, instrução, população, moradia, infra-estrutura.

Inicialmente as seguintes variáveis do censo 2000 foram utilizadas: número de domicílios, população, renda domiciliar, grau de instrução do chefe de família, tipo de moradia (casa, apartamento/ favela ou não), condição de moradia (próprio, alugado, cedido), número de moradores no domicílio, faixa etária da população, condição na família (chefe, cônjuge, filhos), existência de empregados domésticos no domicílio, infra-estrutura (abastecimento de água, instalações sanitárias, coleta de lixo).

Através de análise fatorial foram criados fatores (ou indicadores) capazes de agregar informações correlacionadas, de maneira a não se perder o conteúdo das mesmas. Sendo assim, cada fator responde por uma parcela da variabilidade dos dados. O ideal é se obter o menor número de fatores com um maior percentual de variabilidade explicada.

Utilizando as 11 variáveis do censo, foram determinados 5 indicadores cuja nomenclatura está diretamente relacionada às características das variáveis que o compõem. São eles:

<b>Nome</b>	<b>Componentes principais do indicador</b>
Padrão de Vida (1)	<ul style="list-style-type: none"> <li>• Domicílios com renda superior a 15 S.M.</li> <li>• Chefe do Domicílio c/ superior completo</li> <li>• Empregados no Domicílio</li> <li>• Domicílios do tipo Apartamento</li> </ul>
Composição Familiar Reduzida e Madura (2)	<ul style="list-style-type: none"> <li>• Domicílios com até 2 moradores</li> <li>• Domicílios alugados</li> <li>• Moradores com 60 anos ou mais</li> <li>• Ausência de filhos/ enteados</li> </ul>
Ausência de infra-estrutura de desenvolvimento. (3)	<ul style="list-style-type: none"> <li>• Ausência de Rede Geral ou Fossa Séptica</li> <li>• Uso de vala (para esgoto)</li> <li>• Lixo queimado</li> <li>• Ausência de coleta de lixo</li> </ul>
Adensamento Populacional (4)	<ul style="list-style-type: none"> <li>• Número de domicílios</li> <li>• População Total</li> </ul>
Favelização (5)	<ul style="list-style-type: none"> <li>• Proporção de domicílios do tipo Casa em aglomerado subnormal (favela)</li> </ul>

Cada um destes indicadores foi calculado para a menor unidade de área disponível - o setor censitário (fonte: IBGE/2000). Em seguida, a técnica estatística utilizada foi a análise de clusters, que consiste na formação de grupos com maior homogeneidade interna e maior heterogeneidade entre si, segundo um determinado grupo de variáveis. Neste caso, as variáveis utilizadas são os indicadores.

Desta análise de cluster surgiram os grupos:

- **SOLIDEZ (4% da População):** Grupo que apresenta solidez seja ela conquistada através do tempo ou herdada.
  - É o mais Alto Padrão de Vida: Classe A (65%), chefe de família com nível superior, presença de empregados domésticos.
  - Vivem em apartamento próprio, com total infra-estrutura de desenvolvimento.
  - Família de tamanho tradicional, mas não há destaque para presença de filhos / crianças ou adolescentes
  - Somente Rio de Janeiro e Niterói.
  - Alguns bairros onde predomina: no RJ: parte da Urca, São Conrado, Joá, Leblon, Lagoa, parte da Barra da Tijuca. Em Niterói: parte de Icaraí, Santa Rosa, Ingá.
- **SEGURANÇA (10%):** Grupo baseado na segurança obtida através do desenvolvimento profissional.
  - Alto Padrão de Vida: Classe A/B (41% na Classe A), chefe de família com nível superior, alguma presença de empregados domésticos.
  - Vivem em apartamento, alguns alugados.
  - Composição familiar madura e reduzida, destaque para domicílios com até 2 moradores (cerca de metade dos domicílios).
  - Somente Rio de Janeiro e Niterói, com destaque para Niterói.

- Alguns bairros onde predomina: no RJ: Copacabana, Tijuca, Botafogo, Maracanã, parte do Grajaú. Em Niterói: parte de Icaraí, Santa Rosa, Ingá e Cubango.
- EQUILÍBRIO EM CONQUISTA (3%): Grupo de famílias em processo de conquista de equilíbrio financeiro, com filhos ainda pequenos/jovens.
  - Alto Padrão de Vida: Classe A/B (36% na Classe A), boa proporção de chefes de família com nível superior, alguma presença de empregados domésticos.
  - Vivem em casa (70%) própria. Em alguns casos, não contam com total infra-estrutura urbana de desenvolvimento (não têm rede geral de esgoto - usam fossa séptica; usam poço/ nascente para abastecimento de água)
  - Presença de filhos pequenos e/ou adolescentes.
  - Família de tamanho tradicional, em alguns casos com até mais de 4 moradores.
  - Rio de Janeiro e Niterói, sendo que cerca de 25% dos setores está em Niterói.
  - Alguns bairros no RJ: parte da Barra da Tijuca, parte do Itanhangá, parte de Jacarepaguá. Em Niterói: Itaipú, Jacaré, parte da Lagoa de Piratininga, parte de São Francisco.
- ESTABILIDADE (4%): Grupo de família madura e reduzida, com estabilidade calcada no emprego.
  - Bom Padrão de Vida: Classe A/B (55%), boa proporção de chefes de família com nível superior.
  - Vivem em apartamento próprio ou alugado (maior proporção de alugados entre os segmentos)
  - Moram sozinhos ou no máximo com mais uma pessoa (dois terços do segmento possuem até 2 moradores). Idade mais madura.
  - Rio de Janeiro e Niterói

- Alguns bairros no RJ: centro, parte da Ilha do Governador, Santa Tereza, Glória, parte de Copacabana, Flamengo. Em Niterói: Centro.
- **SOBREVIVÊNCIA (49%):** Grupo de sobrevivência relacionada ao emprego e escolaridade intermediária.
  - Padrão de Vida intermediário: Classe B/C (58%), chefe de família com nível de escolaridade intermediário.
  - Vivem em casa (a maioria) ou apartamento próprio, com boa infraestrutura urbana.
  - Família de tamanho tradicional.
  - É o grupo mais freqüente em praticamente todos os municípios estudados.
  - Nova Iguaçu, Caxias. Alguns bairros no RJ: a maior parte da Zona Norte: Méier, Penha, São Cristóvão, Realengo, Marechal Hermes, Olaria. Em Niterói : Barreto, parte de Fonseca, Santa Bárbara.
- **SUBSISTÊNCIA (15%):** Grupo onde a subsistência está vinculada à proximidade do desenvolvimento urbano, mas sem usufruir completamente da infra-estrutura.
  - Baixo Padrão de Vida: Classe D/E (57%), baixo nível de escolaridade
  - Vivem em casa (a maioria) própria, com fraca infra-estrutura de desenvolvimento, embora não seja favela.
  - Família mais numerosa, com presença de filhos /jovens.
  - Nova Iguaçu, Caxias e Rio de Janeiro.
  - Alguns bairros no RJ: Vargem Grande, Guaratiba, parte de Campo Grande, parte de Bangu, Santa Cruz. Em Niterói: Ititoca, Maceió.
- **PRIVAÇÃO (11%):** Grupo com baixíssimo padrão de vida, porém com uso-fruto de uma infra-estrutura existente.
  - Baixíssimo Padrão de Vida: Classe D/E (63%), baixo nível de escolaridade do chefe de família

- Vivem em casa (a maioria) própria, em favela, porém desfrutam de razoável infra-estrutura urbana.
  - Família numerosa, com presença de filhos /jovens.
  - Rio de Janeiro (principalmente), Caxias e Niterói.
  - Alguns bairros no RJ: parte da Rocinha, Complexo do Alemão, Vidigal, parte do Catumbi. Em Niterói: Morro do Estado.
- PENÚRIA (4%): Ausência de padrão de vida e de acesso à infraestrutura de desenvolvimento.
    - Baixíssimo Padrão de Vida: Classe D/E (68%), baixo nível de escolaridade do chefe de família.
    - Vivem em casa (a maioria) própria, em favela, em precárias condições de infra-estrutura.
    - Família numerosa, com presença de filhos /jovens.
    - Rio de Janeiro (principalmente), Niterói e Caxias.
    - Alguns bairros no RJ: parte do Catumbi (Rua Itapiru), Tijuca (próximo Rua São Miguel), parte da Rocinha.

A empresa desenvolve uma série de ações de fidelização para os assinantes do jornal. Essas ações estão relacionadas a vantagens em estabelecimentos conveniados, por exemplo, a rede de farmácias Drogasmil e o Mc Donald's, além da participação em eventos exclusivos como a pré-estréia de filmes ou peças de teatro. Esses clientes são selecionados pela empresa e recebem uma carta para comunicar o benefício oferecido.

### 3.5.5

#### Descrição da Base de Dados

Após serem selecionadas as 14 variáveis que poderiam explicar o cancelamento de uma assinatura de jornal, estatísticas descritivas foram obtidas para avaliar a consistência da base de dados.

Não foram identificados padrões inusitados, *outliers*, problemas de dados ausentes e incoerências. Com isso, foi aplicado o modelo de regressão logística objetivando investigar a relação dessas variáveis com a probabilidade de cancelamento de uma assinatura.

Quanto ao tipo de assinatura, a base é composta por 88% das assinaturas em débito automático, 11% por assinaturas da modalidade anual e 1% em outras modalidades. Metade paga no cartão de crédito, 25% no débito em conta corrente e 25% através de boleto bancário.

Quase metade das assinaturas foram adquiridas através do telemarketing ativo – interno (37%) e terceirizado (7%) – e 18% através do telemarketing receptivo.

Dois terços dos assinantes estão localizados na Zona Sul, Barra, Niterói e Zona Norte do Rio de Janeiro, a distribuição geográfica dos assinantes ativos e dos clientes que cancelam está no Anexo B deste estudo. Sendo 52% da base classificada nos segmentos prioritários da SD&W solidez, segurança, equilíbrio em conquista e estabilidade. Dos clientes, 22% possuem LTV platina, diamante ou ouro, sendo que 36% da base é cliente há mais de 3 anos.

Da base analisada, 51% pertencem ao sexo masculino e 68% dos assinantes possuem mais de 40 anos de idade.

Participaram de uma ação de fidelização 73% dos clientes e 62% fizeram alguma reclamação sobre o produto ou os serviços da empresa.

Dos clientes assinantes, 35% compraram um outro produto da empresa. Destes, 86% compraram um produto agregado, 9% compraram um anúncio de publicidade e 5% compraram ambos.

As tabelas com as análises descritivas da base podem ser encontradas no Anexo A.

## 4

### Análise dos Resultados

#### 4.1

#### Construção do Modelo de Regressão Logística

No SPSS 13.0, foi aplicado o modelo de regressão logística binário, método *stepwise forward*, para definir o modelo final que minimiza o número de variáveis e maximiza a precisão do modelo.

Um ponto importante é a definição do ponto de corte. Analisado o universo de assinaturas, nota-se que 69% das assinaturas existentes no banco de dados são ativas. Como a amostra respeita a proporção populacional, adotar esta taxa como ponto de corte parece ser o mais adequado. Quando não se conhece a proporção populacional, costuma-se usar o ponto de corte 0,5 que define probabilidades iguais para os dois grupos.

O resultado do modelo inicial apresenta a tabela de classificação considerando o modelo com apenas uma constante, ou seja, se arbitrariamente todas as assinaturas fossem consideradas canceladas, a taxa de acerto seria de 31%. O modelo de regressão logística que irá estimar o risco de cancelamento de clientes precisa ser mais assertivo na classificação dos clientes.

Classification Table<sup>a,b</sup>

Observed		Predicted		
		STATUS		Percentage Correct
Step 0	STATUS	CANCELAD	ATIVO	
	CANCELAD	11057	0	100.0
	ATIVO	24492	0	.0
	Overall Percentage			31.1

a. Constant is included in the model.

b. The cut value is .690

A primeira variável a ser incluída no modelo será aquela que tiver a estatística de pontuação mais alta, estatística Wald, no caso a variável tempo de permanência é selecionada a compor o modelo. Em segundo lugar, a variável

forma de pagamento é incorporada ao modelo. E em seguida, o indicador de reclamação. Essas 3 variáveis contribuem com 92,1% do poder explanatório do modelo.

Na tabela abaixo, verificamos que a análise direcionada a passos utilizando a estatística Wald consumiu 11 passos até se obter o modelo final. Observando-se as significâncias estatísticas do modelo, constatamos que o coeficiente é significativo a cada passo.

**Omnibus Tests of Model Coefficients**

		Chi-square	df	Sig.
Step 1	Step	8848.068	5	.000
	Block	8848.068	5	.000
	Model	8848.068	5	.000
Step 2	Step	1499.479	2	.000
	Block	10347.547	7	.000
	Model	10347.547	7	.000
Step 3	Step	297.834	1	.000
	Block	10645.381	8	.000
	Model	10645.381	8	.000
Step 4	Step	272.349	3	.000
	Block	10917.730	11	.000
	Model	10917.730	11	.000
Step 5	Step	217.242	6	.000
	Block	11134.972	17	.000
	Model	11134.972	17	.000
Step 6	Step	229.119	9	.000
	Block	11364.091	26	.000
	Model	11364.091	26	.000
Step 7	Step	165.895	8	.000
	Block	11529.987	34	.000
	Model	11529.987	34	.000
Step 8	Step	148.315	6	.000
	Block	11678.302	40	.000
	Model	11678.302	40	.000
Step 9	Step	125.520	12	.000
	Block	11803.821	52	.000
	Model	11803.821	52	.000
Step 10	Step	49.970	1	.000
	Block	11853.791	53	.000
	Model	11853.791	53	.000
Step 11	Step	26.362	4	.000
	Block	11880.154	57	.000
	Model	11880.154	57	.000

Após diversas interações, o modelo final selecionou 11 das 14 variáveis incluídas inicialmente no modelo. Excluiu as variáveis: gênero, indicador de compra de produto agregado e indicador de compra de um anúncio de publicidade. As variáveis que resultaram do modelo final são: “tempo de permanência”, “forma de pagamento”, “indicador de reclamação”, “tipo de assinatura”, “fonte de venda”, “faixa etária”, “SD&W”, “LTV”, “região”, “indicador de participação em ações de fidelização” e “quantidade de produtos agregados comprados”.

## 4.2

### Avaliação do Ajuste Geral

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	35227.478 <sup>a</sup>	.220	.310
2	33727.998 <sup>a</sup>	.253	.355
3	33430.165 <sup>a</sup>	.259	.364
4	33157.816 <sup>b</sup>	.264	.372
5	32940.573 <sup>c</sup>	.269	.378
6	32711.454 <sup>c</sup>	.274	.385
7	32545.559 <sup>c</sup>	.277	.390
8	32397.244 <sup>c</sup>	.280	.394
9	32271.724 <sup>d</sup>	.283	.398
10	32221.754 <sup>d</sup>	.284	.399
11	32195.392 <sup>d</sup>	.284	.400

- a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.
- b. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.
- c. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.
- d. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

A cada passo, quando uma nova variável é incluída no modelo, a estatística de probabilidade – 2log diminui indicando uma melhora no modelo. Em contrapartida, as medidas pseudo  $R^2$  aumentam à medida que previsoires são adicionados. O pseudo  $R^2$  de Nagelkerke no último passo aumentou em 30% o poder de explicação do modelo obtido no passo 1.

**Hosmer and Lemeshow Test**

Step	Chi-square	df	Sig.
1	.000	4	1.000
2	139.350	7	.000
3	135.151	7	.000
4	187.903	7	.000
5	129.408	8	.000
6	134.869	8	.000
7	134.234	8	.000
8	120.247	8	.000
9	144.911	8	.000
10	142.347	8	.000
11	135.804	8	.000

A medida Hosmer e Lemeshow de ajuste geral tem um teste estatístico que indica que não houve diferença estatisticamente significativa entre as classificações observadas e previstas para todos os modelos com duas ou mais variáveis.

O valor Hosmer e Lemeshow mede a correspondência dos valores efetivos e previstos da variável dependente. Neste caso, o melhor ajuste do modelo é indicado por uma diferença menor na classificação observada e prevista. Um bom ajuste de modelo é indicado por um valor *chi-quadrado* não significativo (Hair, 1998).

No modelo do último passo, todas as medidas de ajuste melhoraram. O valor -2LL diminuiu para 32.195. Os valores  $R^2$  variam de 0,310 para 0,400, indicando melhoria no modelo de 11 variáveis, embora seja um valor distante dos valores  $R^2$  geralmente encontrados em regressão múltipla. A medida Hosmer e Lemeshow indica a ausência de diferença significativa na distribuição de valores dependentes efetivos e previstos.

Essas medidas combinadas sugerem a aceitação do modelo do último passo como um modelo significativo de regressão logística.

### 4.3

#### Precisão da Estimativa

Pode haver problemas na utilização de métodos de regressão logística direcionados a passos quando o objetivo da análise é a precisão da estimativa. Os algoritmos direcionados a passos buscam um subconjunto de variáveis que maximize a probabilidade, mas isto não é o mesmo que maximizar a precisão da estimativa (SPSS, 2003).

As matrizes de classificação, idênticas em natureza às utilizadas na análise discriminante (Hair, 1998), mostram taxas de acerto extremamente altas de casos corretamente classificados para o modelo de 11 variáveis. A taxa de acerto geral é de 75,3%, além disso, as taxas de acerto de grupos individuais são consistentemente altas e não indicam um problema na previsão de qualquer um dos dois grupos. Apesar de altas, as taxas de acerto do grupo que cancela é maior que a taxa do grupo que não cancela, 77,4% contra 74,3%.

O modelo inicial que considerava apenas a constante tinha uma taxa geral de acerto de 31,1%. O modelo completo com 11 variáveis aumenta 2,5 vezes a taxa de acerto na previsão.

A partir do passo 9, a melhora no  $R^2$  é pequena e a taxa de acerto geral do modelo não se altera. Isto indica que as variáveis 10 e 11 poderiam ser descartadas do modelo final porque quanto menos variáveis um modelo tiver, menor o tempo de processamento. No entanto, as variáveis “indicador de participação em ações de fidelização” e “quantidade de produtos agregados comprados” foram mantidas neste modelo não para aumentar a precisão, mas para ajudar na definição do perfil dos clientes que cancelam a assinatura do jornal.

Classification Table<sup>a</sup>

Observed	STATUS	CANCELED ATIVO	Predicted		Percentage Correct
			STATUS		
			CANCELED	ATIVO	
Step 1	STATUS	CANCELED ATIVO	8519 7140	2538 17352	77.0 70.8
	Overall Percentage				72.8
Step 2	STATUS	CANCELED ATIVO	8528 6633	2529 17859	77.1 72.9
	Overall Percentage				74.2
Step 3	STATUS	CANCELED ATIVO	8464 6477	2593 18015	76.5 73.6
	Overall Percentage				74.5
Step 4	STATUS	CANCELED ATIVO	8462 6407	2595 18085	76.5 73.8
	Overall Percentage				74.7
Step 5	STATUS	CANCELED ATIVO	8634 6590	2423 17902	78.1 73.1
	Overall Percentage				74.6
Step 6	STATUS	CANCELED ATIVO	8509 6371	2548 18121	77.0 74.0
	Overall Percentage				74.9
Step 7	STATUS	CANCELED ATIVO	8560 6438	2497 18054	77.4 73.7
	Overall Percentage				74.9
Step 8	STATUS	CANCELED ATIVO	8553 6347	2504 18145	77.4 74.1
	Overall Percentage				75.1
Step 9	STATUS	CANCELED ATIVO	8549 6286	2508 18206	77.3 74.3
	Overall Percentage				75.3
Step 10	STATUS	CANCELED ATIVO	8558 6275	2499 18217	77.4 74.4
	Overall Percentage				75.3
Step 11	STATUS	CANCELED ATIVO	8554 6291	2503 18201	77.4 74.3
	Overall Percentage				75.3

a. The cut value is .690

#### 4.4

#### Validação do Modelo (*holdout sample*)

Para conseguir a eficiência classificatória do modelo, a amostra foi separada em duas partes: uma utilizada para estimação do modelo, e outra para testar a eficiência da classificação – *holdout sample* (Hair *et al.*, 1998). A amostra utilizada para estimação, também chamada de amostra de treinamento, contou com 35.549 assinantes. O processo de escolha foi realizado no software SPSS através da geração de números aleatórios.

A validação do modelo de regressão logística é obtida através da aplicação do modelo na amostra de validação (Hair, 1998). As taxas de acerto na amostra de validação são quase idênticas às taxas de acerto da amostra de treinamento. Isto leva à conclusão de que o modelo de regressão logística possui forte suporte empírico tanto na amostra de validação quanto na de treinamento.

#### Amostra de Validação – Tabela de Classificação

Classification Table<sup>a</sup>

Observed			Predicted		
			STATUS		Percentage Correct
			CANCELAD	ATIVO	
Step 11	STATUS	CANCELAD	1162	365	76.1
		ATIVO	809	2460	75.3
	Overall Percentage				75.5

a. The cut value is .690

#### Amostra de Treinamento – Tabela de Classificação

Classification Table<sup>a</sup>

Observed			Predicted		
			STATUS		Percentage Correct
			CANCELAD	ATIVO	
Step 11	STATUS	CANCELAD	8554	2503	77.4
		ATIVO	6291	18201	74.3
	Overall Percentage				75.3

a. The cut value is .690

## 4.5

### Interpretação dos Resultados

O modelo de regressão logística selecionou as 11 variáveis que melhor explicam o cancelamento de uma assinatura de jornal. Gerado o modelo, atribui-se uma probabilidade de cancelamento a cada indivíduo da base em estudo. Todos os indivíduos foram classificados em dois grupos: o grupo que cancela e o grupo que não cancela.

O cliente pode se desligar da empresa motivado por diversos fatores. Alguns desses fatores podem ser descritos por meio de variáveis e utilizados para estabelecer uma relação entre cancelamento e a ocorrência ou não de situações relacionadas aos fatores em questão (Barros, 2002).

O modelo de regressão logística, utilizado neste trabalho, permite identificar o risco de cancelamento de clientes a partir de variáveis transacionais, demográficas e do histórico de eventos armazenados no banco de dados.

Ao desenvolver o modelo de regressão logística sobre a base de dados, foi possível determinar um modelo que contempla as seguintes variáveis: “tempo de permanência”, “forma de pagamento”, “indicador de reclamação”, “tipo de assinatura”, “fonte de venda”, “faixa etária”, “SD&W”, “LTV”, “região”, “indicador de participação em ações de fidelização” e “quantidade de produtos agregados comprados”.

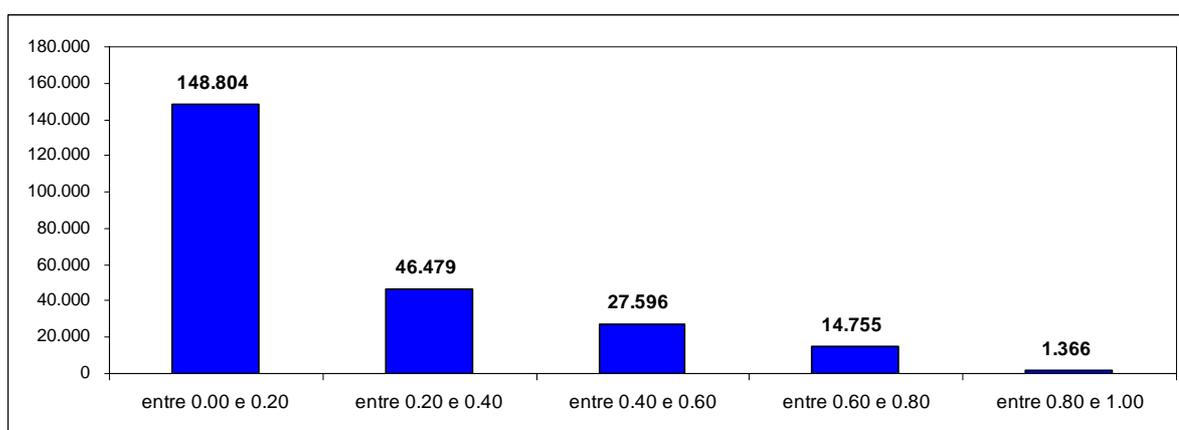
Essas variáveis aplicadas ao modelo de regressão logística geraram um *score* de probabilidade de cancelamento para cada um dos clientes da base de dados conforme o quadro a seguir.

<b>Faixa Probabilidade</b>	<b>Ativo</b>	<b>Cancelado</b>	<b>Total Geral</b>	<b>%canc.</b>
<b>entre 0.00 e 0.20</b>	15,249	1,601	16,850	9%
<b>entre 0.20 e 0.40</b>	4,763	1,656	6,419	18%
<b>entre 0.40 e 0.60</b>	2,828	2,635	5,463	32%
<b>entre 0.60 e 0.80</b>	1,512	3,848	5,360	54%
<b>entre 0.80 e 1.00</b>	140	1,317	1,457	61%

Nota-se que o percentual de assinaturas canceladas cresce à medida que o *score* de probabilidade de cancelamento aumenta. Isto é coerente com a tabela de classificação que acerta em 75,3% dos casos.

Aplicado o *score* de probabilidade na base total de assinantes pode-se concluir que a maior parte dos clientes ativos apresenta baixa probabilidade de cancelar a assinatura do jornal, 62% da base de clientes tem até 20% de chance de cancelar a assinatura do jornal.

**GRÁFICO 5: TOTAL DE ASSINATURAS X SCORE DE RISCO DE CANCELAMENTO.**

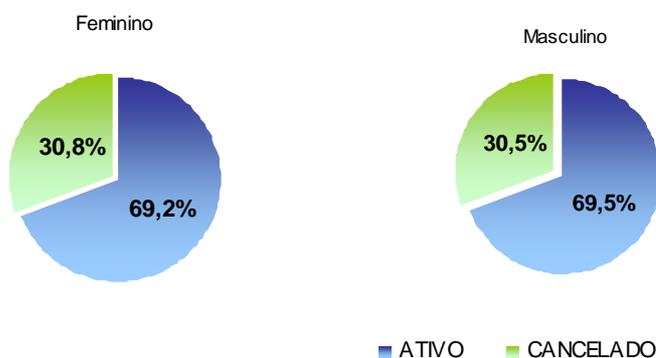


Na análise de regressão logística é possível identificar qual a mudança nas chances de um evento ocorrer dada a presença de um fator (variável categórica) ou alteração em uma variável contínua.

É possível descrever ainda o perfil dos clientes que cancelam, o que pode auxiliar na definição de ações profiláticas que ajudem a reduzir a perda de clientes.

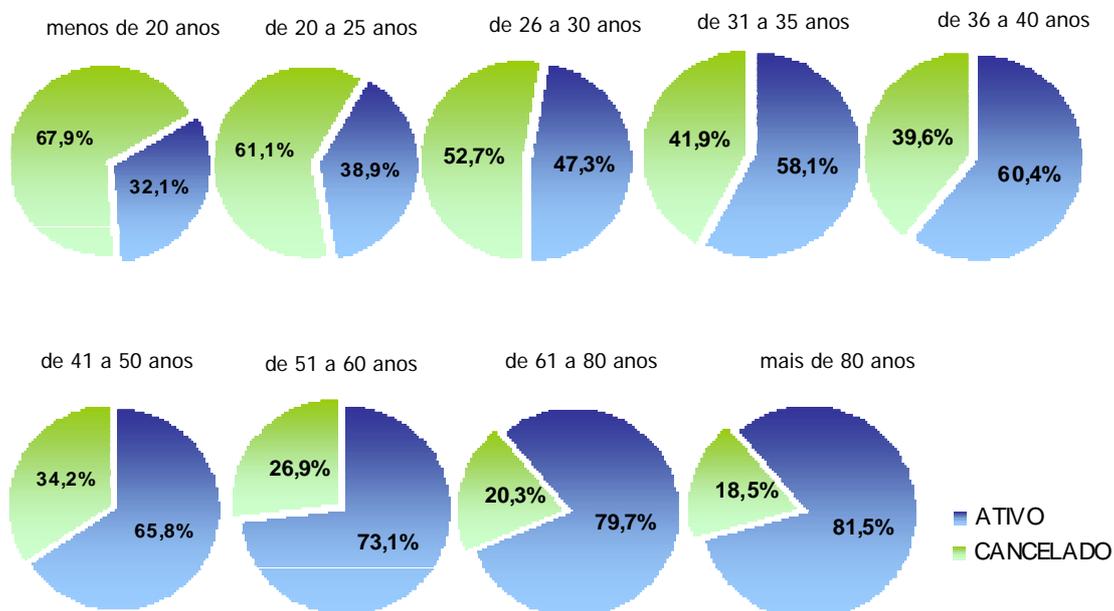
O gênero não discrimina o cancelamento. Não existe diferença no cancelamento entre homens e mulheres.

GRÁFICO 6: SCORE DE RISCO DE CANCELAMENTO x GÊNERO



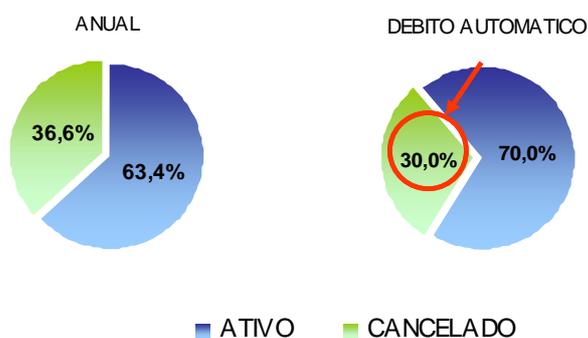
Os mais jovens cancelam mais que os clientes mais velhos. Assinantes com menos de 20 anos têm 2 vezes mais chances de cancelar do que um cliente com mais de 50 anos. O *score* de risco de cancelamento vai se reduzindo com o aumento da faixa etária dos clientes.

GRÁFICO 7: SCORE DE RISCO DE CANCELAMENTO x IDADE



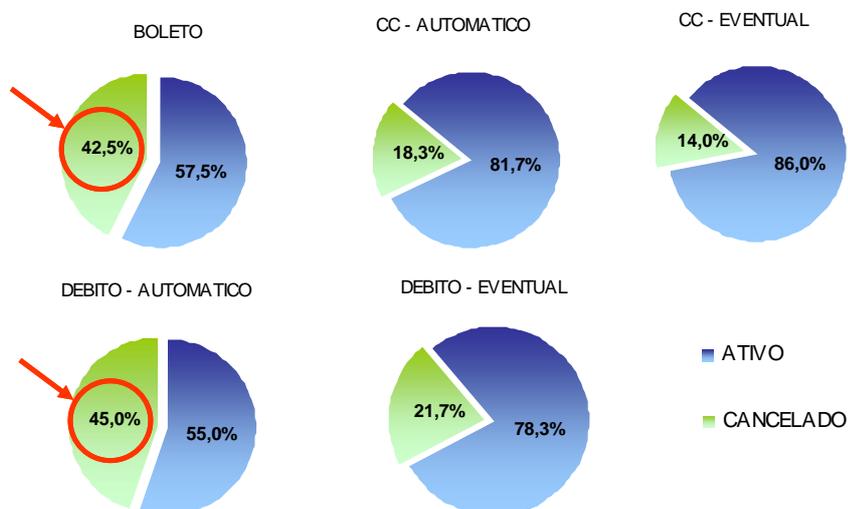
As assinaturas do tipo anual apresentam taxas de cancelamento mais elevadas que as assinaturas da modalidade débito automático. Uma assinatura anual tem 18% mais chance de cancelar do que uma assinatura débito automático.

**GRÁFICO 8: SCORE DE RISCO DE CANCELAMENTO x TIPO DE ASSINATURA**



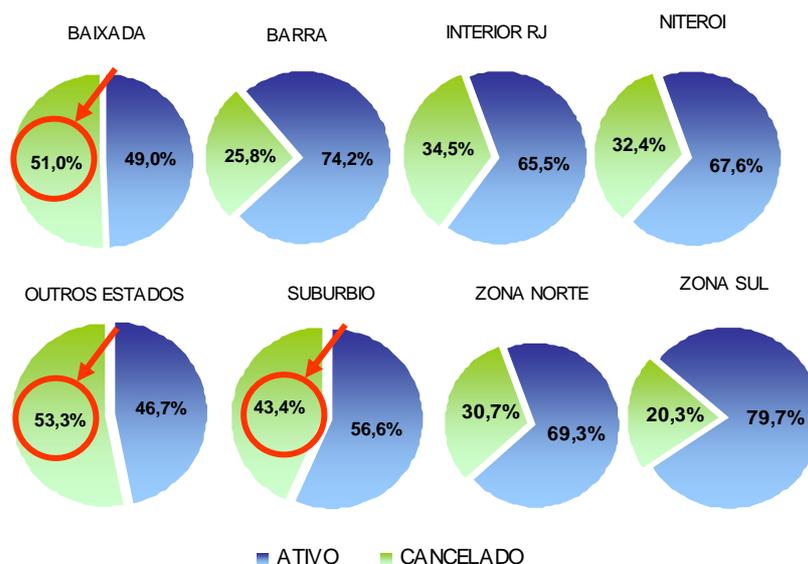
Forma de pagamento é a segunda variável mais importante do modelo e explica o cancelamento. Uma assinatura paga no débito em conta corrente ou boleto bancário tem 1,5 vezes mais chance de cancelar do que uma assinatura paga no cartão de crédito. As assinaturas pagas no cartão de crédito apresentam as menores chances de cancelamento.

**GRÁFICO 9: SCORE DE RISCO DE CANCELAMENTO x FORMA DE PAGAMENTO**



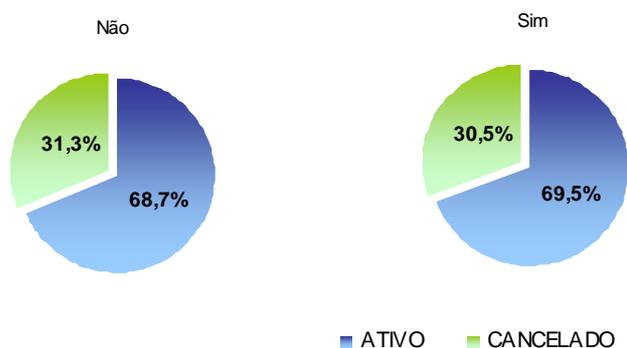
A Zona Sul é a região que menos incrementa o *score* de risco de cancelamento de uma assinatura, enquanto as assinaturas fora do estado aumentam o *score* de risco em 2,6 vezes. As regiões que apresentam o maior risco de cancelamento são Baixada, subúrbio e outros estados.

**GRÁFICO 10: SCORE DE RISCO DE CANCELAMENTO x REGIÃO**



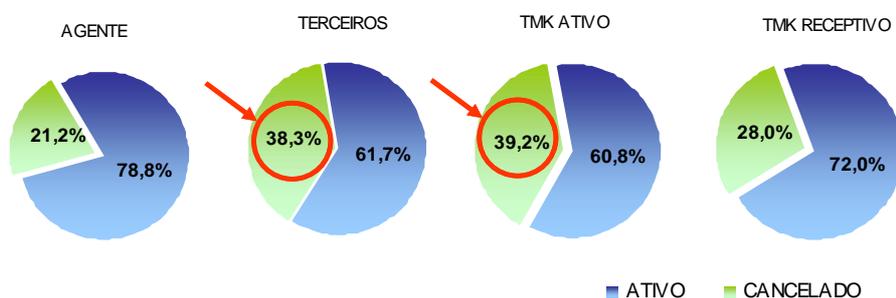
Os clientes que foram impactados por ações de fidelização não apresentam diferença no *score* de cancelamento comparado aos assinantes que nunca foram impactados por estas ações. Talvez seja preciso redefinir as ações de marketing de relacionamento para que contribuam de forma mais efetiva para a fidelização de clientes.

**GRÁFICO 11: SCORE DE RISCO DE CANCELAMENTO x PARTICIPAÇÃO EM AÇÕES DE FIDELIZAÇÃO**



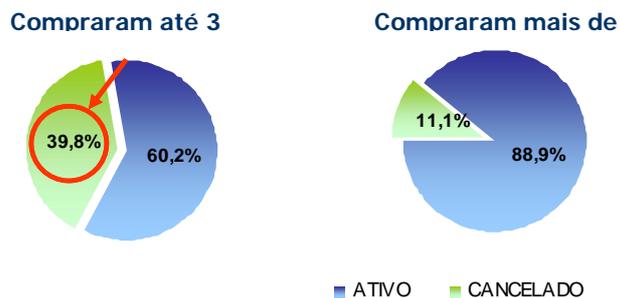
Dos clientes classificados no grupo dos que cancelam, 55% adquiriram sua assinatura através do telemarketing ativo, no grupo dos que não cancelam 38% fizeram a assinatura pelo canal telemarketing ativo.

**GRÁFICO 12: SCORE DE RISCO DE CANCELAMENTO x CANAL DE VENDA**



Assinantes que compram menos de 3 produtos agregados como livros, CDs, DVDs e guias apresentam um *score* de risco de cancelamento 3,6 vezes maior do que o de clientes que compram estes produtos da empresa. Dos clientes que compram mais de 3 produtos agregados, 11% cancelam a sua assinatura, enquanto a taxa de cancelamento da base é de 31%. Além disso, quanto mais produtos um cliente compra, menor o *score* de cancelamento.

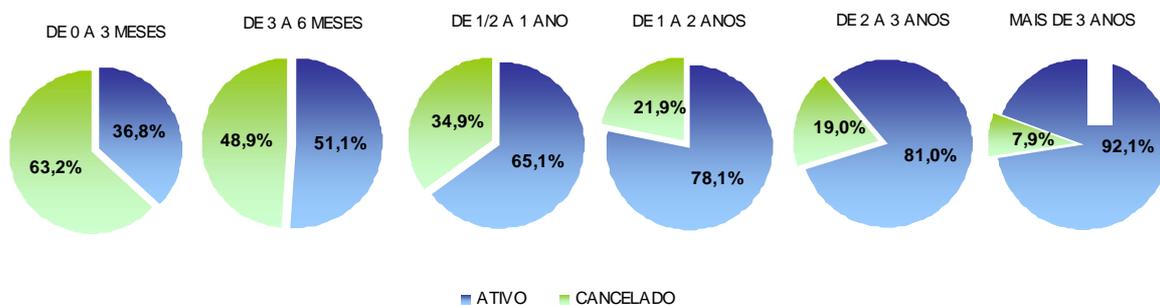
**GRÁFICO 13: SCORE DE RISCO DE CANCELAMENTO x QTDE DE PRODUTOS AGREGADOS COMPRADOS**



Apesar disso, clientes que fizeram um anúncio de publicidade apresentam taxas de cancelamento similares aos que não anunciaram.

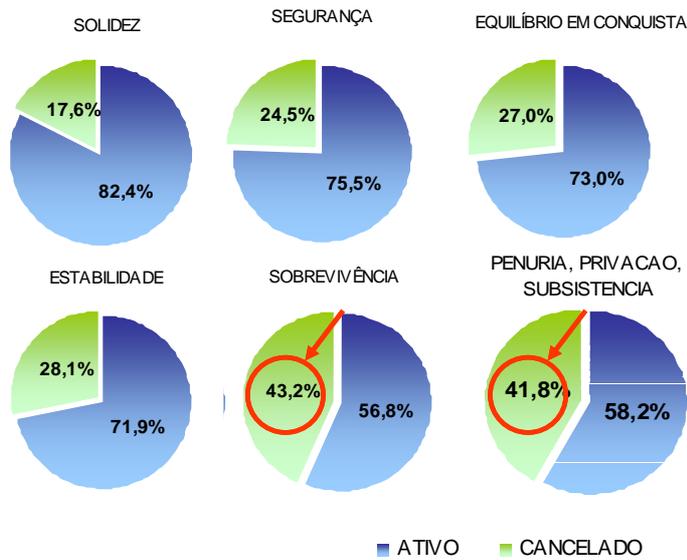
O risco de cancelamento é decrescente à medida que o tempo de permanência do cliente na carteira aumenta. Dos clientes que estão na carteira no máximo há 3 meses, 63% cancelam a sua assinatura. A taxa de cancelamento entre os clientes com mais de 3 anos na base é 8%.

**GRÁFICO 14: SCORE DE RISCO DE CANCELAMENTO x TEMPO DE PERMANÊNCIA**



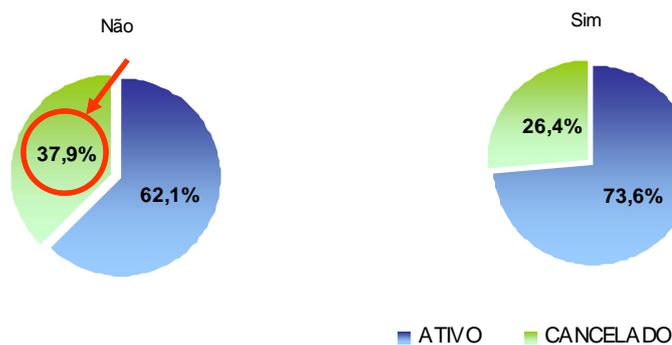
Quanto menor o nível sócio demográfico do indivíduo, maior o *score* de risco de cancelamento. Os menores níveis de risco de desligamento estão entre os clientes dos segmentos SD&W “solidez”, “segurança”, “equilíbrio em conquista” e “estabilidade”. Os assinantes dos segmentos sócio-demográficos “sobrevivência”, “privação”, “penúria” e “subsistência” têm risco de cancelamento 1,8 vez maior do que o risco dos demais segmentos.

**GRÁFICO 15: SCORE DE RISCO DE CANCELAMENTO x SD&W**



Os clientes que não fizeram reclamações apresentam taxas de cancelamento maiores que os clientes que entraram em contato com a empresa para reclamar. Clientes que não reclamam apresentam um risco de cancelamento 40% maior do que os clientes que contataram a empresa para apresentar a sua reclamação.

**GRÁFICO 16: SCORE DE RISCO DE CANCELAMENTO x INDICADOR DE RECLAMAÇÃO**



## 5

### Conclusões e Recomendações

#### 5.1

##### Conclusões

O objetivo deste estudo foi utilizar a base de dados de clientes de uma empresa para desenvolver um modelo de regressão logística que determine o risco de cancelamento de clientes que possuem uma assinatura de jornal, bem como traçar o perfil dos clientes com maior risco de desligamento para propor ações de retenção que aumentem a lucratividade da carteira de assinantes desta empresa.

É possível afirmar que o modelo de regressão logística pode ser eficiente para a determinação do risco de cancelamento de clientes a partir de variáveis sócio-demográficas, transacionais e do histórico de eventos.

O modelo foi aplicado a uma amostra de dados com 35.549 clientes e as variáveis compostas no modelo final foram: “tempo de permanência”, “forma de pagamento”, “indicador de reclamação”, “tipo de assinatura”, “fonte de venda”, “faixa etária”, “SD&W”, “LTV”, “região”, “indicador de participação em ações de fidelização” e “quantidade de produtos agregados comprados”.

A tabela de classificação mostrou que a taxa de acerto geral do modelo de regressão logística é de 75,3% e que as taxas de acerto dos grupos individuais são altas e indicam uma consistência na previsão de qualquer um dos dois grupos. O grupo que cancela apresentou taxa de acerto de 77,4% enquanto o grupo que não cancela tem taxa de acerto de 74,3%.

O pseudo  $R^2$  de Nagelkerke apresentou um poder de explicação de 0,4 e a medida Hosmer e Lemeshow de ajuste geral através de um teste estatístico indica que não houve diferença estatisticamente significativa entre as classificações observadas e previstas para o modelo final. Além disso, o valor de -2LL aumentou a cada passo.

A combinação dessas medidas de avaliação do ajuste e da precisão do modelo indica a aceitação deste como um modelo de regressão logística significativa. Além disso, o modelo foi aplicado a uma base de validação diferente

da amostra utilizada na fase de estimação dos coeficientes do modelo. As taxas de acertos na amostra de validação são quase idênticas às taxas de acerto da amostra de treinamento. Isto indica que o modelo de regressão logística possui suporte empírico tanto na amostra de validação quanto na de treinamento.

Assim, de acordo com as variáveis assumidas no modelo de regressão logística, conclui-se que o perfil do assinante com maior risco de cancelamento da sua assinatura é: o jovem de até 30 anos; dos segmentos sócio-demográfico de menor nível (sobrevivência, penúria, privação e subsistência); morador da Baixada, subúrbio ou de outros estados; que tenha adquirido sua assinatura através do telemarketing ativo; cuja assinatura seja da modalidade anual, com forma de pagamento boleto ou débito em conta corrente; que tenham comprado sua primeira assinatura recentemente; sem compra de produto agregado ao jornal e que nunca tenham feito reclamações.

Os resultados obtidos indicam que o modelo é bastante eficiente em distinguir assinantes que possivelmente cancelariam sua assinatura dos que não cancelariam. Assim, ao implementar uma ação de marketing com o objetivo de reter os clientes, as informações fornecidas pelo modelo permitirão que se atinja o grupo com maior risco de cancelamento com mais eficiência do que sem a utilização do modelo.

Uma seleção aleatória para uma ação de retenção traria 31% dos clientes canceladores de sua assinatura para serem impactados pela ação. Estariam sendo atingidos 69% de clientes cujos riscos de cancelamento são pequenos, ou seja, há um desperdício de investimento nesta ação. A utilização do modelo permite que se alcance 75% de clientes potenciais canceladores. Com isso, pode haver maior eficácia na ação de marketing, à medida que se diminui a dispersão do investimento e aumenta-se a probabilidade de o público alvo ser atingido.

## 5.2

### Recomendações

Os resultados obtidos indicam algumas ações de marketing e de negócio que podem contribuir para aumentar a rentabilidade da empresa através da retenção de clientes ou da implantação de ações que melhorem a qualidade da venda, o que irá se reverter em clientes que permaneçam mais tempo na carteira.

A empresa poderia estimular a venda de assinaturas do tipo débito automático que apresentam taxas de risco de cancelamento 19% menores do que as assinaturas anuais. Poderia mudar o foco da operação de vendas praticando um comissionamento maior para assinaturas em débito automático. As assinaturas pagas em boleto e débito em conta corrente poderiam ter um preço maior já que têm uma probabilidade de cancelamento mais de 2 vezes maior que o cartão de crédito.

Quanto aos canais de vendas, pode-se estimular a venda de assinaturas deste jornal através do telemarketing receptivo que possui score de risco de cancelamento 28% menor que o risco do telemarketing ativo. Deve-se investigar ainda os motivos de cancelamento no canal telemarketing ativo, talvez a venda esteja sendo “forçada” para um público sem perfil de compra e isso gera mais cancelamentos futuramente.

Os assinantes das regiões “Baixada”, “subúrbio” e “outros estados” têm as maiores taxas de cancelamento entre todas as regiões, portanto devem ser implementadas ações de retenção que aumentem o tempo de permanência dos clientes dessas áreas. Seria necessário investigar através de pesquisa de mercado ou teste quais ações seriam mais eficientes para o público dessas regiões, tais como brindes, programa de recompensas ou vantagens em outros serviços. É recomendável que se tenha cautela para vender novas assinaturas para estas áreas.

O público mais jovem que apresenta taxas mais elevadas de cancelamento, também precisa ser impactado por ações de fidelização. Estes programas podem envolver o cliente mais jovem com a empresa e estabelecer uma relação afetiva com a marca até que se desenvolva o hábito pela leitura e pela assinatura de jornal.

Como a taxa de cancelamento dos clientes diminui com o aumento do tempo na carteira, a empresa pode implantar ações que aumentem o tempo de permanência do cliente. Ou então, estimular o uso do produto aos novos clientes. O cancelamento no início do ciclo de vida pode estar ocorrendo porque o cliente desconhece o produto e conseqüentemente não percebe o seu valor, interrompendo o relacionamento com a empresa.

Os maiores riscos de cancelamento estão entre os clientes com menor nível sócio-demográfico. A empresa deve vender menos para estes segmentos ou

vender assinaturas de modalidades mais baratas como as assinaturas de final de semana.

Os clientes que reclamam cancelam menos que aqueles que têm o perfil de não reivindicar. Aqueles que não reclamam interrompem o relacionamento com a empresa sem dizer a ela o motivo da sua insatisfação. Será importante realizar uma pesquisa de mercado com os clientes que cancelaram sua assinatura para entender os motivos de cancelamento da assinatura ou disponibilizar no *call center* uma marcação no sistema para que os operadores indiquem os motivos que levaram ao cancelamento no momento que o cliente telefona.

A empresa precisa rever as ações de fidelização que estão sendo utilizadas atualmente, pois não estão atingindo o objetivo de reter clientes. Os clientes que participaram de ações de fidelização apresentam *score* de cancelamento quase idêntico ao dos assinantes que não participaram. Estas ações não estão se revertendo em menos cancelamento para a empresa.

### 5.3

#### **Recomendação para Futuras Pesquisas**

As empresas que têm por objetivo entender o comportamento de seus clientes devem montar uma base de dados que registre o relacionamento dos produtos com seus consumidores e também o perfil dos clientes e *prospects*.

Algumas variáveis não foram consideradas neste estudo porque não estavam disponíveis na base de dados da empresa. Novos estudos utilizando essas variáveis poderiam definir melhor o perfil dos clientes com risco de cancelamento. As variáveis seriam:

Transacionais	Motivo de cancelamento
	Oferta na compra da assinatura
	Histórico de inadimplência
	Natureza das reclamações
	Promoção na compra
	Oferta recebida na compra
	Valor pago real
Geográfica	Setor censitário
Demográficas	Grau de instrução
	Profissão
	Estado Civil
	Renda
Psicográficas	Estilo de Vida
	Personalidade
Comportamentais	Motivos das ligações para o call center
	Nível de lealdade à marca
	Satisfação com produto e serviço
	Nível de afinidade com o meio jornal
	Hábito de leitura de jornal na família

Um modelo similar a este pode ser aplicado para um determinado período do ciclo de vida dos clientes. Por exemplo, os clientes no início do ciclo de vida. Como 50% dos cancelamentos ocorrem entre os clientes que possuem até 6 meses na base, pode-se desenvolver um modelo de regressão logística apenas para esses clientes.

Este estudo pode ser aplicado a qualquer empresa cuja interação com os clientes seja contínua, tais como: telefonia móvel e fixa, TV por assinatura, internet banda larga, cartão de crédito, planos de saúde, seguradoras, revistas. Cada uma dessas empresas precisa definir que variáveis podem influenciar o cancelamento antes de modelar o risco de desligamento de seus clientes.

## Referências Bibliográficas

AAKER, D.A. Brand Equity – Gerenciando o Valor a Marca. São Paulo: Negócio Editora, 1998.

AAKER, D.A.; KUMAR, V. & DAY, G.S. **Pesquisa de marketing**. São Paulo: Atlas, 2001.

ARAÚJO, P.F. de. RELACIONAMENTO COM O CLIENTE: UM ESTUDO EM SHOPPING CENTERS DO RIO DE JANEIRO. 01/09/1999. 1v. 164p. Mestrado. UNIVERSIDADE FEDERAL DO RIO DE JANEIRO --ADMINISTRAÇÃO. Orientador(es): REBECCA ARKADER. Biblioteca Depositária: COPPEAD/UFRJ1.

BALDINGER, A.L. & RUBINSON, J. Brand Loyalty: The link Between attitude and behavior. In: **Journal of Advertising Research**. Nov/Dez., p. 22-34, 1996.

BARROS, J.E.G. Análise de Sobrevivência: Modelo de Risco de Desligamento de Clientes. 01/03/2002, 1v. 123p. Mestrado. UNIVERSIDADE FEDERAL DO RIO DE JANEIRO – ADMINISTRAÇÃO. Orientador(es): CELSO FUNCIA LEMME. Biblioteca Depositária: COPPEAD/UFRJ.

BERRY, L.L. & PARASURAMAN, A. **Servicos de Marketing**. 3ª Ed. São Paulo: Editora Maltese, 1992.

BERRY, L.L. & PARASURAMAN, A. Building a new academic field – thw case of services marketing. **Journal of Retailing**, Vol. 69 (1), 13-18, Spring 1993.

COATES, D.; FINLAY, P. & WILSON, J. Validation in marketing models. **Journal of the Market Research Society**. V.33, n.2, p.83-90, April 1991.

CRAVENS, D.W. & CRAVENS, K.S. Horizontal Alliances for Relationship Marketing. In: SHETH, J.N.; PARVATIYAR, A. **Handbook of Relationship Marketing**. Sage Publications, 200.

CUNHA, A.G. da. Potencial de retenção dos usuários de ônibus de uma empresa de transporte. 01/04/2003. 1v. 212p. Mestrado. UNIVESIDADE FEDERAL DO RIO DE JANEIRO – ENGENHARIA DE TRANSPORTES. Orientador(es): MILENA BODMER. Biblioteca Depositária: Biblioteca do Centro de Tecnologia da UFRJ.

Dowling, G.R., UNCLES, Mark. Do customer loyalty programs really work? **Sloan Management Review**, v.38, n.4, 9.71-82, Summer 1997.

ECHEVESTE, S.S. Lealdade do Consumidor: um estudo sobre retenção de consumidores do cartão Visa/Banrisul. 01/09/1999. 1v. 139p. Mestrado. UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL – ADMINISTRAÇÃO. Orientador(es): Luiz Antônio Siongo. Biblioteca Depositária: Gladys W. Amaral.

FADER, P. & SCHMITTLEIN, D. Excess Behavioral Loyalty for High-Share Brands: Deviations from the Dirichlet Model for Repeat Purchasing. In: **Journal of Marketing Research**. Vol. 30, p. 478-493, November/1993.

FILHO, R.C.P. ESTRATÉGIA DE LEALDADE: UM ESTUDO DE CASO. 01/09/1999. 1v. 157p. Mestrado. UNIVERSIDADE FEDERAL DO RIO DE JANEIRO – ADMINISTRAÇÃO. Orientador(es): CESAR GONÇALVES NETO. Biblioteca Depositária: COPPEAD/UFRJ.

GROFF, R. Mensurando de Relacionamento: uma aplicação da escala de Wilson & Vlosky no mercado de bens de capital. 01/07/2001. 1v. 160p. Profissionalizante. UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL – ADMINISTRAÇÃO. Orientador(es): Fernando Bins Luce. Biblioteca Depositária: Administração.

HAIR, J.F. Et al. **Multivariate data Analysis with Readings**. 4 ed. New Jersey: Prentice Hall, 1995.

HOCHMAN, K. **Customer loyalty programs**. In: NASH, Edward L. (Org.) The direct marketing handbook. 2.ed. New York: McGraw-Hill, 1992. p. 781-799.

IMBER, J. & TOFFLER, Betsy-Ann. **Dictionary of marketing terms**. 2.ed. New York: Barron's Educational Series, 1994. 579p.

JOHNSON, K. Choosing the right program. In: **Direct Marketing**. P. 36-45, June/1998.

JONES, T.; SASSER JR. & EARL, W. Why satisfied customer defect. **Harvard Business Review**, v. 73, n. 6, 9. 88-99, Nov./Dec. 1995.

KOCKKANNY, M.M. Relacionamento e Lealdade: um estudo em companhias aéreas nacionais. 2003. Mestrado. UNIVERSIDADE FEDERAL DO PARANÁ – ADMINISTRAÇÃO. Orientador(es): Renato Zancan Marchetti. Biblioteca Depositária: Administração.

KOTLER, P. **Administração de Marketing**. 10. Ed. São Paulo: Prentice Hall, 2000.

KOTLER, P. & ARMSTRONG, G. **Princípios de marketing**. 7.ed. Rio de Janeiro: Prentice-Hall, 1995. Cap. 18.

KURTZ, D.L. & CLOW, K.E. Services marketing. New York: John Wiley & Sons, 1998.

LOVERLOCK, C.H. & WRIGHT, L. **Services: marketing e gestão**. Rio de Janeiro: Saraiva, 2001.

MACEDO, M.G.C. de. O Uso de Smart Cards Para Retenção e Fidelização de Clientes: O Caso Smart Club. 01/03/2001. 2v. 106p. Mestrado. Universidade Federal do Rio De Janeiro – Administração. Orientador(es): Rebecca Arkader. Biblioteca Depositária: COPPEAD/UFRJ.

MARTINHO, M.A.S. Fatores Decisivos para a Retenção de Alunos em escolas de Educação Infantil. 01/06/2003. 8v. 138p. Mestrado. Universidade Presbiteriana Mackenzie – Administração de Empresas. Orientador(es): Eliane Pereira Zamith Brito. Biblioteca Depositária: Biblioteca Central – George Alexander.

McKENNA, R. **Marketing de relacionamento: estratégias bem-sucedidas para a era do cliente**. Rio de Janeiro: Campus, 1993. 254p.

MELO, A.S.C. de. O marketing de relacionamento como fator de influência na lealdade do cliente: um estudo de caso no mercado de assinatura de revistas. 01/10/2002. 1v. 172p. Mestrado. Universidade de São Paulo – Administração. Orientador(es): Geraldo Luciano Toledo. Biblioteca Depositária: Biblioteca da FEA-USP.

NASH, E.L **Database marketing**: uma ferramenta atual e decisiva do marketing. São Paulo: Makron, 1994. 270p.

OLIVEIRA, B.E.M. de. Conquista e Retenção de Clientes no Setor Varejista de Eletro-Eletrônicos. 01/10/2002. 1v. 197p. Mestrado. Universidade Federal da Paraíba / João Pessoa – Administração. Orientador(es): Sônia Trigueiro de Almeida. Biblioteca Depositária: Setorial e Central da UFPB.

PEPPERS, D. **Os segredos dos programas de fidelidade bem sucedidos**. Disponível em <http://www.1tol.com.br> Acesso em: Mar. 2001.

PINTO, Solange da C. Marketing de Relacionamento e Retenção de Clientes: um estudo de caso na indústria jornalística. 01/11/1997. 1v. 128p. Mestrado. Universidade Federal de Minas Gerais – Administração. Orientador(es): Luiz Antônio Antunes Teixeira. Biblioteca Depositária: Biblioteca da FACE/UFGM/Bibl. Universitária/UFGM/Bibl./UFLA.

PORTER, M.E. & MILLAR, V.E. How information gives you competitive advantage. **Harvard Business Review**, v. 63, n. 4, p. 149-160, July/Aug. 1985.

REICHHELD, F. Loyalty-Based Management. In: **Harvard Business Review**. P. 64-73, March/April, 1993.

REICHHELD, F.F. **The loyalty effect: the hidden force behind growth, profits, and lasting value**. Boston: Harvard Business Scholl Press, 1996.

REMENYI, D.S.J. **Doing Research in Business and Management: An Introduction in Process and Method**. Boston: Sage Publications Ltda, 1998.

REZENDE, A.S. O impacto da satisfação e lealdade na retenção de clientes no serviço de telefonia celular – um modelo integrativo. 01/05/2003. 1v. 1p. Mestrado. Pontífca Universidade Católica do Rio de Janeiro – Administração de Empresas. Orientador(es): Jorge Ferreira da Silva. Biblioteca Depositária: Biblioteca Central.

ROBERTS, J.H. Marketing approaches to forecasting problems. **Journal of Forecasting**, v. 17, n. 3/4, p. 169-174, June/July 1998.

SANTOS, C.P. dos. Impacto do Gerenciamento de Reclamações na Confiança e Lealdade do Consumidor, no Contexto de Trocas Relacionadas de Serviços: Construção e Teste de Um Modelo Teórico. 01/07/2001. 1v. 253p. Doutorado. Universidade Federal do Rio Grande do Sul – Administração. Orientador(es): Carlos Alberto Vargas Rossi. Biblioteca Depositária: Administração.

SINCICH, T. **Business statistics by example**. 5. Ed. New Jersey: Prentice Hall, 1996.

**SPSS Advanced Statistics 6.1**. Chicago: SPSS Inc., 1994. Cap. 1, 10, 12.

**SPSS for Windows Release 8.0.0**. Chicago: SPSS Inc., 1997. CD Rom.

TOMANARI, S.A.A. Segmentação de Mercado com Enfoque em Valores e Estilo de Vida (Segmentação Psicográfica) – Um Estudo Exploratório. 01/02/2003.

Mestrado. Universidade De São Paulo – Ciências Da Comunicação.  
Orientador(es): Mitsuru Higuchi Yanaze.

TURCHAN, M.P. & MATEUS, P. The Value of Relationships. **Journal of Business Strategy**, v.22, n.6, p.29-32, November/December 2001.

VAVRA, T.G. **Marketing de relacionamento: Aftermarketing**. São Paulo: Atlas, 1993. 323p.

ZEITHAML, V.A. Consumer Perceptions of Price, Quality na Value: A Means-End model and Synthesis of Evidence. In: **Journal of Marketing**. Vol. 52, p. 2-22, July/1998.

## Anexo A

### Estatísticas Descritivas

As estatísticas descritivas da amostra de 35.549 assinaturas são as seguintes de acordo com os quadros abaixo:

**QUADRO 01 – Distribuição dos assinantes por status da assinatura**

Status	Frequência	%
Ativo	24492	69%
Cancelado	11057	31%
Total	35549	100%

**QUADRO 02 – Distribuição dos assinantes por tipo de assinatura**

Tipo de Assinatura	Frequência	%
Anual	4008	11%
Débito Automático	31279	88%
Semestral	67	0%
Trimestral	195	1%
Total geral	35549	100%

**QUADRO 03 – Distribuição dos assinantes por fonte de vendas**

Fonte de Vendas	Frequência	%
Agente	1169	3%
Antigos	4798	13%
Outros	13	0%
Sem Fonte	7834	22%
Terceiros	2342	7%
Tmk Ativo	12984	37%
Tmk Receptivo	6409	18%
Total geral	35549	100%

**QUADRO 04 – Distribuição dos assinantes por forma de pagamento**

Forma de Pagamento	Frequência	%
Boleto	8783	25%
Cartão de Crédito	18010	51%
Débito em Conta	8756	25%
Total geral	35549	100%

**QUADRO 05 – Distribuição dos assinantes por tempo de permanência**

Tempo de Permanência	Frequência	%
Entre 0 e 3 meses	9089	26%
Entre 3 e 6 meses	2976	8%
De 6 meses a 1 ano	3594	10%
Entre 1 e 2 anos	4375	12%
Entre 2 e 3 anos	2794	8%
Mais de 3 anos	12721	36%
Total geral	35549	100%

**QUADRO 06 – Distribuição dos assinantes por região**

Região	Frequência	%
Baixada	1296	4%
Barra	4930	14%
Brasília	307	1%
Interior do RJ	3509	10%
Niterói	2957	8%
Outros Estados	193	1%
São Paulo/ BH	401	1%
Subúrbio	5082	14%
Zona Norte	6871	19%
Zona Sul	10003	28%
Total geral	35549	100%

**QUADRO 07 – Distribuição dos assinantes por gênero**

Gênero	Total	%
F	14499	41%
M	18147	51%
I	2903	8%
Total geral	35549	100%

**QUADRO 08 – Distribuição dos assinantes por indicador de compra de produtos****agregados**

Indicador de compra de produto agregado	Total	%
Sim	11189	31%
Não	24360	69%
Total geral	35549	100%

**QUADRO 09 – Distribuição dos assinantes por quantidade de produtos agregados**

Qtde de produtos agregados	Total	%
0 produtos	24360	69%
até 3 produtos	7500	21%
entre 4 e 7 produtos	1903	5%
entre 8 e 10 produtos	1024	3%
mais de 10 produtos	762	2%
Total geral	35549	100%

**QUADRO 10 – Distribuição dos assinantes por LTV (Life Time Value)**

LTV	Total	%
Platina	214	1%
Diamante	2591	7%
Ouro	4858	14%
Prata	925	3%
Bronze	22879	64%
Lata	10	0%
Indefinido	4072	11%
Total geral	35549	100%

**QUADRO 11 – Distribuição dos assinantes por faixa etária**

Faixa Etária	Total	%
Menos de 20 anos	84	0%
De 20 a 25 anos	899	3%
De 26 a 30 anos	1876	5%
De 31 a 35 anos	2691	8%
De 36 a 40 anos	3019	8%
De 41 a 50 anos	7902	22%
De 51 a 60 anos	7638	21%
De 61 a 80 anos	7828	22%
Mais de 80 anos	1139	3%
Sem preenchimento	2473	7%
Total geral	35549	100%

**QUADRO 12 – Distribuição dos assinantes por SD&W**

SD&W	Total	%
Solidez	5157	15%
Segurança	9027	25%
Equilíbrio em conquista	1734	5%
Estabilidade	2398	7%
Sobrevivência	5365	15%
Privação	109	0%
Penúria	33	0%
Subsistência	475	1%
Indefinido	11251	32%
Total geral	35549	100%

**QUADRO 13– Distribuição dos assinantes por indicador de compra de anúncios de classificados**

Indicador de compra de anúncio	Total	%
Sim	1689	5%
Não	33860	95%
Total geral	35549	100%

**QUADRO 14– Distribuição dos assinantes por indicador de reclamação**

Indicador de reclamação	Total	%
Sim	22091	62%
Não	13458	38%
Total geral	35549	100%

**QUADRO 15– Distribuição dos assinantes por indicador de participação em ações de fidelização**

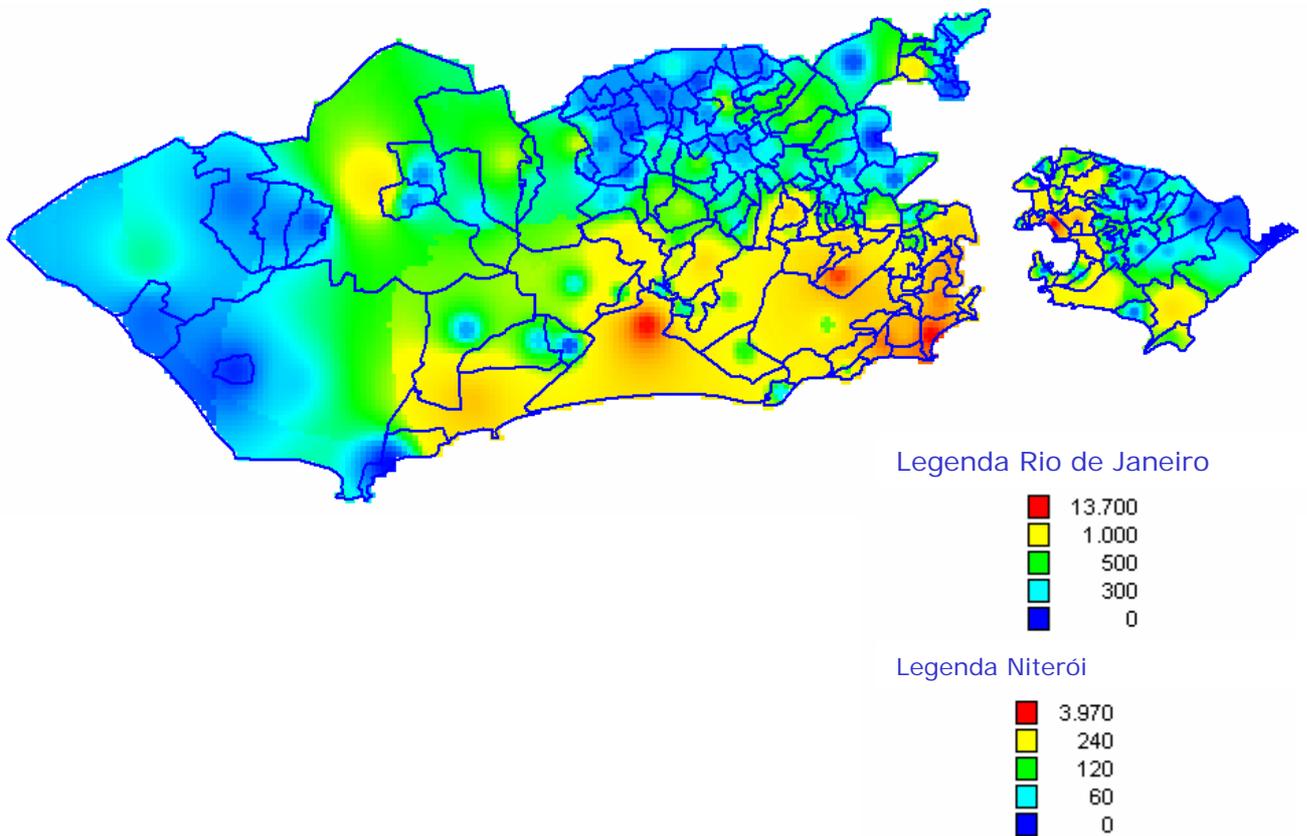
Participação em ações de fidelização	Total	%
Sim	25942	73%
Não	9607	27%
Total geral	35549	100%

## Anexo B

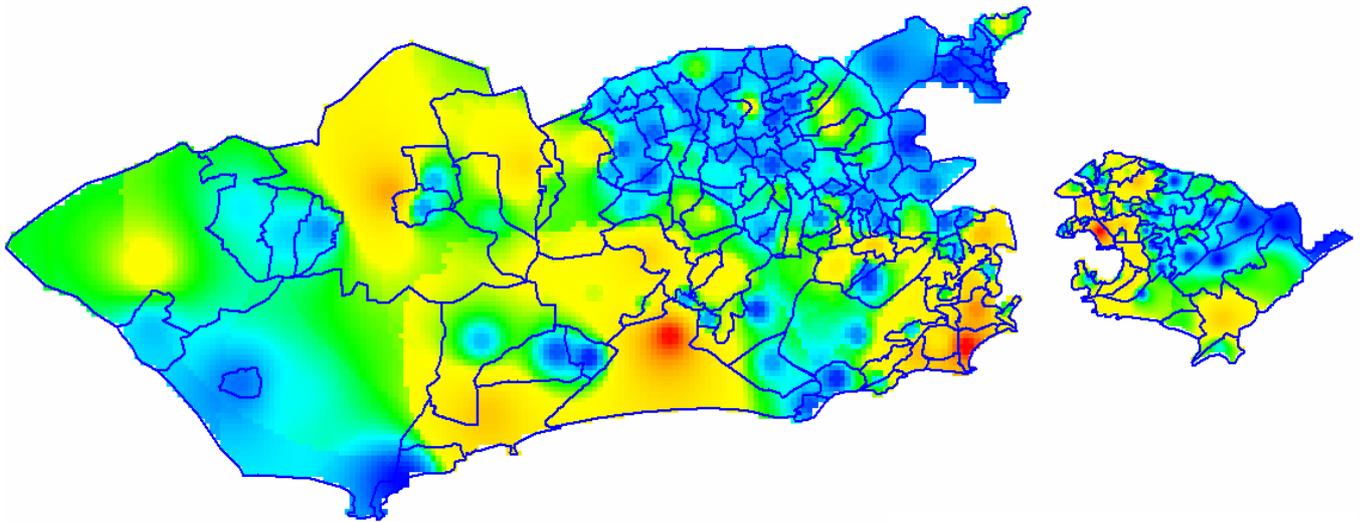
### Distribuição Geográfica

A carteira de assinantes e os cancelamentos seguem a seguinte distribuição geográfica:

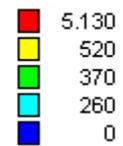
**MAPA 01 – Distribuição geográfica dos assinantes**



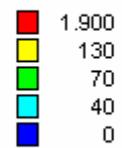
MAPA 02 – Distribuição geográfica dos assinantes cancelados



Legenda Rio de Janeiro



Legenda Niterói



# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)