



Manuela Alencar da Cruz Dantas

**Um Modelo de Risco de Cancelamento de Clientes de
Telefonia Fixa – A Aplicação da Regressão Logística
para Retenção de Clientes**

Dissertação de Mestrado

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

Orientador: Prof. Jorge Ferreira da Silva

Rio de Janeiro
Maio de 2009



Manuela Alencar da Cruz Dantas

**Um Modelo de Risco de Cancelamento de Clientes de
Telefonia Fixa – A Aplicação da Regressão Logística
para Retenção 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. Marcos Cohen
Departamento de Administração – PUC-Rio

Prof. Luis Antonio da Rocha Dib
UFRJ

Prof. Nizar Messari
Vice-Decano de Pós-Graduação do CCS

Rio de Janeiro, 05 de maio de 2009

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

Manuela Alencar da Cruz Dantas

Graduada em Administração de Empresas e Pós-graduada em Administração Financeira pela UECE.

Ficha Catalográfica

Dantas, Manuela Alencar da Cruz

Um modelo de risco de cancelamento de clientes de telefonia fixa: a aplicação da regressão logística para retenção de clientes / Manuela Alencar da Cruz Dantas ; orientador: Jorge Ferreira da Silva. – 2009.

128 f. : il. (col.) ; 30 cm

Dissertação (Mestrado em Administração)– Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.

Inclui bibliografia

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

CDD: 658

Aos meus pais, exemplos eternos para
minha vida de amor e dedicação.
Aos meus amores, Ricardo e Luana,
pela paciência e permanente estímulo.

Agradecimentos

A Deus, por sempre me guiar e iluminar meus caminhos para mais esta conquista.

Ao meu orientador Prof. Dr. Jorge Ferreira da Silva pelas valiosas contribuições, paciência e apoio, essenciais ao sucesso deste trabalho.

Aos professores e equipe do IAG pela atenção, aprendizado e ajuda.

Aos meus colegas da turma de mestrado PUC-Rio; foi um período difícil, mas muito divertido e produtivo.

Aos professores da comissão examinadora.

A todos amigos e familiares que direta ou indiretamente me ajudaram e me incentivaram em mais uma vitória.

Dedico um agradecimento especial a meus pais, irmãos, madrinha, sogros e cunhados pelo amor e apoio incondicional às minhas escolhas e atividades e por toda tranquilidade proporcionada nos momentos difíceis.

Por fim, aos meus amores Ricardo e Luana, pois eles são a minha força, a minha vontade, a base da minha vida.

Resumo

Dantas, Manuela Alencar da Cruz; Silva, Jorge Ferreira da. **Um modelo de risco de cancelamento de clientes de telefonia fixa – a aplicação da regressão logística para retenção de clientes**. Rio de Janeiro, 2009. 128p. Dissertação de Mestrado – Departamento de Administração de Empresas, Pontifícia Universidade Católica do Rio de Janeiro.

O atual ambiente organizacional está marcado por uma alta competitividade, elevada turbulência e por mudanças rápidas e descontínuas no macro ambiente das empresas. O cenário exige foco no cliente e estratégias voltadas para manutenção de um relacionamento profícuo para ambas as partes (cliente e empresa), com visão de longo prazo. Esta dinâmica de mercado é o foco principal desta pesquisa, que está centrada na retenção de clientes, como estratégia competitiva para aumento de valor para empresa. O objetivo, então, é desenvolver uma ferramenta que auxilie na definição do perfil de clientes mais propensos a romper o relacionamento (*churn*), permitindo que a empresa antecipe-se a este acontecimento, tornando-se mais pró-ativa, e assim, sendo mais eficiente nos seus processos de negócio. Com este propósito foi feita uma revisão bibliográfica sobre comportamento do consumidor, marketing de relacionamento, retenção de clientes, segmentação de mercado e *churn* em telecomunicações. Depois do embasamento teórico e direcionamento oferecido pelas obras consultadas, foi realizada a aplicação prática de uma ferramenta estatística na base de dados disponível de uma empresa de telecomunicações, utilizando regressão logística binária para o modelo de propensão de cancelamento. O resultado alcançou uma taxa de acerto geral de 79,2%, composto por 13 variáveis que podem facilitar a identificação do perfil de clientes com maior probabilidade de cancelamento, indicando ainda, que apenas 8% da base apresentam mais de 80% de chance de serem cancelados. Estes resultados demonstram a possibilidade e importância da identificação de clientes propensos ao *churn*, assim como daqueles que a empresa deve atrair e reter para ser capaz de desenvolver práticas eficientes e lucrativas.

Palavras-chave

Retenção de clientes; marketing de relacionamento; regressão logística; segmentação e *churn*.

Abstract

Dantas, Manuela Alencar da Cruz; Silva, Jorge Ferreira da (Advisor). **A model to measure customers' cancellation risk in telecommunications – the application of logistic regression for customer retention.** Rio de Janeiro, 2009. 128p. MSc. Dissertation – Departamento de Administração de Empresas, Pontifícia Universidade Católica do Rio de Janeiro.

The current organizational environment is marked by a high competitiveness, high turbulence and rapid and discontinuous changes in companies' macro environment. Scenario requires focus on customers and strategies geared towards maintaining a fruitful relationship for both parties (customer and company), with long-term vision. This dynamics of this market is the main focus of this research, which focuses on customer retention as a competitive strategy to increase value for companies. The objective then is to develop a tool that assists in defining the profile of customers most likely to break the relationship (churn), allowing the company forward to this event, becoming more pro-active, and thus being more efficient in their business processes. In this regard was a comprehensive literature review on consumer behavior, relationship marketing, customer retention, market segmentation and churn in telecommunications. After the theoretical basis and guidance offered by works consulted, the practice was performed by a statistical tool based on data available from a telecommunications company, using binary logistic regression model for the propensity of cancellation. The result achieved an overall success rate of 79,2%, composed of 13 variables that may facilitate identification of the profile of customers most likely to cancellation, further indicating that only 8% of the database have more than 80% chance to be canceled. These results demonstrate the possibility and importance of the identification of customers likely to churn, as well as those that the company must attract and retain to be able to develop efficient and profitable practice.

Keywords

Customers retention; relationship marketing; logistic regression; segmentation and churn.

Sumário

1. O Problema	17
1.1. Introdução	17
1.2. Objetivos	18
1.3. Relevância do Estudo	19
1.4. Delimitação do Estudo	20
1.5. Organização do Trabalho	20
2. Referencial Teórico	22
2.1. Comportamento do Consumidor	22
2.1.1. Expectativas dos Clientes Acerca dos Serviços	23
2.1.2. Satisfação e Valor para o Cliente	25
2.2. Relacionamento com Clientes	29
2.2.1. Conceito de Marketing de Relacionamento	29
2.2.2. Benefícios da Construção de Relacionamentos	30
2.2.3. Valor do Ciclo de Vida de um Cliente	32
2.3. Retenção de Clientes	34
2.3.1. Foco na Retenção de Clientes	34
2.3.2. Estratégias de Retenção	36
2.3.3. Obtendo Informações sobre o Cliente	41
2.4. Segmentação do Mercado Consumidor	44
2.4.1. Variáveis Geográficas	47
2.4.2. Variáveis Demográficas	47
2.4.3. Variáveis Psicográficas	50
2.4.4. Variáveis Comportamentais	51
2.5. <i>Churn</i> na Telefonia Fixa	52
2.5.1. O Setor de Telefonia	53
2.5.2. O Problema do <i>Churn</i>	58
2.5.3. Gerenciamento do <i>Churn</i>	60

3. Metodologia	62
3.1. Tipo de Pesquisa	62
3.2. Universo e Amostra	64
3.3. Coleta de Dados	64
3.4. Dados Disponíveis e suas Limitações	65
3.5. Tratamento de Dados	65
3.5.1. <i>Oversampling</i>	65
3.5.2. Descrição das Variáveis	66
3.5.3. Validação, Exploração, Limpeza da Base de Dados e Transformações	71
3.5.4. Seleção das Variáveis	73
3.5.5. Base Seleccionada	74
3.5.6. O Modelo: Regressão Logística Binária	74
4. Análise dos Resultados	80
4.1. Construção do Modelo de Regressão Logística	80
4.2. Avaliação do Ajuste Geral do Modelo	83
4.3. Precisão da Estimativa	84
4.4. Validação do Modelo (<i>Holdout Sample</i>)	86
4.5. Interpretação dos Resultados	86
5. Conclusões e Recomendações	105
5.1. Conclusões	105
5.2. Recomendações e Trabalhos Futuros	108
6. Referências Bibliográficas	109
7. Apêndice	116

Lista de figuras

Figura 1 – Zona de tolerância	24
Figura 2 – Relação entre satisfação e fidelidade de clientes em indústrias competitivas	25
Figura 3 – As conseqüências do valor superior para os clientes e empresas	28
Figura 4 – Metas com relação a clientes no marketing de relacionamento: obter, satisfazer, reter e ampliar clientes, Serviços de Informação	30
Figura 5 – Ciclo virtuoso de lealdade para criação de vantagem no relacionamento	36
Figura 6 – Níveis de estratégias de retenção	37
Figura 7 – Regiões das operadoras fixa	42
Figura 8 – Formato da Curva Logística	76

Lista de tabelas

Tabela 1 – Diferenças conceituais entre satisfação e valor	27
Tabela 2 – Estratégias de Retenção – Laços Financeiros	38
Tabela 3 – Estratégias de Retenção – Laços Sociais	39
Tabela 4 – Exemplos de Informações sobre o cliente, em banco de dados de algumas empresas	42
Tabela 5 – Variáveis de segmentação e dimensões	45
Tabela 6 – Estágios no ciclo de vida da família	49
Tabela 7 – Concessionárias e Empresas Espelho	55
Tabela 8 – Distribuição de acessos fixos em serviço entre concessionárias e autorizadas.	57
Tabela 9 – Terminais em serviços de telefonia fixa das três principais empresas	59
Tabela 10 – Descrição da variável alvo ou dependente	66
Tabela 11 – Descrição das variáveis independentes	67
Tabela 12 – Estatísticas mínimo, máximo e desvio padrão das variáveis numéricas	72
Tabela 13 – Tabela de classificação – proporção de <i>churn</i>	80
Tabela 14 – Passos do método <i>Stepwise Forward</i>	81
Tabela 15 – Cox & Snell R^2 , Nagelkerke R^2 e -2LL do modelo	83
Tabela 16 – Teste estatístico Hosmer e Lemeshow	84
Tabela 17 – Matriz de classificação das amostras de treinamento e de validação	85
Tabela 18 – Coeficiente da variável quantidade de fixos	89
Tabela 19 – Coeficientes da variável segmento	90
Tabela 20 – Coeficientes da variável grupo de produtos/serviços	91
Tabela 21 – Coeficientes da variável tempo do último fixo	92
Tabela 22 – Coeficiente da variável intenção anterior de cancelamento (nos últimos 6 meses)	94

Tabela 23 – Coeficientes da variável UF	95
Tabela 24 – Coeficiente da variável disponibilidade de banda larga da empresa	96
Tabela 25 – Coeficientes da variável idade do cliente	97
Tabela 26 – Coeficientes da variável quantidade de meses de faturamento de ligação para celular	98
Tabela 27 – Coeficiente da variável atuação do concorrente 1	99
Tabela 28 – Coeficiente da variável utilização de internet discada	100
Tabela 29 – Coeficientes da variável tempo da primeira banda larga	101
Tabela 30 – Coeficientes da variável tempo do primeiro móvel	102
Tabela 31 – Tabela resumo das variáveis mais relevantes para o modelo	102

Lista de gráficos

Gráfico 1 – Penetração de telefonia nos domicílios Brasileiros	53
Gráfico 2 – <i>Market share</i> de terminais fixos em serviço	58
Gráfico 3 – Distribuição da base por <i>score</i> de risco de cancelamento	87
Gráfico 4 – Distribuição da amostra de <i>churn</i> por quantidade de terminais fixos na residência do cliente	88
Gráfico 5 – Distribuição da amostra de <i>churn</i> por segmento do cliente	89
Gráfico 6 – Distribuição da amostra de <i>churn</i> por grupo de produtos/serviços	91
Gráfico 7 – Distribuição da amostra de <i>churn</i> por tempo do último fixo	92
Gráfico 8 – Distribuição da amostra de <i>churn</i> por intenção anterior de cancelamento (nos últimos 6 meses)	93
Gráfico 9 – Distribuição da amostra de <i>churn</i> por UF	94
Gráfico 10 – Distribuição da amostra de <i>churn</i> por disponibilidade de banda larga	95
Gráfico 11 – Distribuição da amostra de <i>churn</i> por idade do cliente	96
Gráfico 12 – Distribuição da amostra de <i>churn</i> por quantidade de meses de faturamento de ligação para celular	97
Gráfico 13 – Distribuição da amostra de <i>churn</i> por atuação do concorrente 1	98
Gráfico 14 – Distribuição da amostra de <i>churn</i> por uso ou não de internet discada	99
Gráfico 15 – Distribuição da amostra de <i>churn</i> por tempo da primeira banda larga	100
Gráfico 16 – Distribuição da amostra de <i>churn</i> por tempo do primeiro móvel	101

Lista de quadros

Quadro 1 – Distribuição dos terminais por Estado	116
Quadro 2 – Distribuição dos terminais por gênero	116
Quadro 3 – Distribuição dos terminais por grupo de produtos	117
Quadro 4 – Distribuição dos terminais por disponibilidade de banda larga	117
Quadro 5 – Distribuição dos terminais por atuação da concorrência 1	117
Quadro 6 – Distribuição dos terminais por atuação da concorrência 2	117
Quadro 7 – Distribuição dos terminais por retenção (intenção de <i>churn</i>)	117
Quadro 8 – Distribuição dos terminais por valor para empresa	117
Quadro 9 – Distribuição dos terminais por quantidade de fixo RES	118
Quadro 10 – Distribuição dos terminais por quantidade de fixo NRES	118
Quadro 11 – Distribuição dos terminais por quantidade de móveis pré	118
Quadro 12 – Distribuição dos terminais por quantidade de móveis pós	118
Quadro 13 – Distribuição dos terminais por quantidade de banda larga RES	119
Quadro 14 – Distribuição dos terminais por quantidade de banda larga NRES	119
Quadro 15 – Distribuição dos terminais por quantidade de fixos	119
Quadro 16 – Distribuição dos terminais por quantidade de banda larga	119
Quadro 17 – Distribuição dos terminais por quantidade de insapoios	120
Quadro 18 – Estatísticas de frequência de tempo dos terminais fixos	120
Quadro 19 – Estatísticas de frequência de tempo dos terminais móveis	121
Quadro 20 – Estatísticas de frequência de tempo de banda larga	121
Quadro 21 – Estatísticas de frequência de tempo de relacionamento	121
Quadro 22 – Correlação variáveis numéricas	122
Quadro 23 – Correlação variáveis numéricas 2	123
Quadro 24 – Distribuição da base de <i>churners</i> por Estado	124
Quadro 25 – Distribuição da base de <i>churners</i> por gênero	124

Quadro 26 – Distribuição da base de <i>churners</i> por grupo de produtos	124
Quadro 27 – Distribuição da base de <i>churners</i> por disponibilidade de banda larga	124
Quadro 28 – Distribuição da base de <i>churners</i> – concorrência 1	125
Quadro 29 – Distribuição da base de <i>churners</i> – concorrência 2	125
Quadro 30 – Distribuição da base de <i>churners</i> – intenção de cancelamento	125
Quadro 31 – Distribuição da base de <i>churners</i> por valor para empresa	125
Quadro 32 – Distribuição da base de <i>churners</i> por terminais RES	125
Quadro 33 – Distribuição da base de <i>churners</i> por terminais NRES	125
Quadro 34 – Distribuição da base de <i>churners</i> por móveis pré	126
Quadro 35 – Distribuição da base de <i>churners</i> por móveis pós	126
Quadro 36 – Distribuição da base de <i>churners</i> por banda larga RES	126
Quadro 37 – Distribuição da base de <i>churners</i> por banda larga NRES	126
Quadro 38 – Distribuição da base de <i>churners</i> por terminais fixos	126
Quadro 39 – Distribuição da base de <i>churners</i> por banda larga	126
Quadro 40 – Distribuição da base de <i>churners</i> por <i>insapoios</i>	127
Quadro 41 – Distribuição da base de <i>churners</i> por idade	127
Quadro 42 – Distribuição da base de <i>churners</i> por tempo do último fixo	127
Quadro 43 – Distribuição da base de <i>churners</i> por tempo do primeiro móvel	127
Quadro 44 – Distribuição da base de <i>churners</i> por tempo da primeira banda larga	127
Quadro 45 – Distribuição da base de <i>churners</i> por tempo de relacionamento	128
Quadro 46 – Distribuição da base de <i>churners</i> por internet discada	128
Quadro 47 – Distribuição da base de <i>churners</i> por receita de longa distância	128
Quadro 48 – Distribuição da base de <i>churners</i> por receita total	128
Quadro 49 – Distribuição da base de <i>churners</i> por receita de ligação fixo – móvel	128