Data Mining Applications for Multi-Channel Marketing

By

Matthew Johnson

B.Com., The University of British Columbia, 1999

A Thesis Submitted in Partial Fulfillment Of
The Requirements for the Degree Of

Master of Science (Business Administration)

In

The Faculty Of Graduate Studies

(Department of Marketing; School of Commerce)

FACULTY OF COMMERCE AND BUSINESS ADMINISTRATION

We accept this thesis as conforming
to the required standard

The University of British Columbia

August 2002

Department of Marketing

The University of British Columbia
Vancouver, Canada

Date 29 August 2002

Abstract

The management of multi-channel marketing is one of the critical issues facing marketing practitioners today. The emergence of e-commerce has presented new opportunities to communicate with and serve customers. As a result, managers face uncertainties as to what is the best way to incorporate an e-commerce channel into their existing marketing strategies.

Working in conjunction with a national retail chain (ABC)*, this thesis analyzes a multi-channel management situation using data mining techniques. The retailer has seven stores in major Canadian cities. In addition, ABC operates a telephone and mail order store and recently launched an online store at www.abc.ca.

Preliminary analysis focused on ABC cataloging practices. Weekly sales data was analyzed using linear regression models to determine if a promotional impact could be seen in the weeks following catalog mailings. While there was no statistical evidence to support a promotional impact in the retail store sales, there was a promotional impact was found in the mail order sales. While this finding was interesting, the financial impact was modest.

In order to identify opportunities with high sales and contribution potential, the focus of the analysis then shifted to customer's channel shopping patterns. A longitudinal study of customers shopping indicated that as customers increased/decreased the number of channels used, their spending increased/decreased. Financial analysis confirmed that converting single-channel customers to multi-channel would have high payoffs.

To determine which customers would be most easily converted to multi-channel, logistic regression was used. Customers that were most likely to convert to multi-channel displayed four characteristics:

- lived more than 50km from an ABC store,
- made more frequent purchases prior to becoming multi-channel
- spent larger amounts of money prior to becoming multi-channel
- have been a ABC customer for a longer period

The central implication is that marketing efforts should be targeted at "store shoppers more likely to go multi-channel". A campaign that achieves a 10 percent conversion to multi-channel shopping has the potential to produce a significant increase in sales.

[*] ABC is used here as a fictitious identifier of the retail chain.

Table of Contents

List of Tables

List of Figures

Introduction

Practitioners have come to recognize that use of e-commerce provides both opportunities and problems. On the one hand e-commerce provides an additional method of reaching out to potential customers. On the other it presents an additional area to be managed and integrated into the marketing strategy. Will the addition of an e-commerce channel provide new customers? Will the new channel better serve current customers? Will customers expect to order from one channel and take delivery via a second? Will all customers use all channels or will different channels serve different customer segments? Does the addition of e-commerce provide the opportunity to increase efficiency, or to increase market penetration, or to focus on high potential customers? Then there are questions regarding data and analysis. What customer data can be gathered to guide multi-channel planning and what analytical tools can be used to interpret this data?

This thesis looks at the multi-channel management situation faced by ABC[1], a retailer with thirty years experience in clothing and accessory products. In addition to retail stores in major cities across Canada ABC uses catalogues to support telephone and mail order sales, and as of 2001, ABC products are available on the ***abc.ca*** website.

The purpose of this thesis is to explore the use of data mining techniques to support multi-channel marketing strategy. In brief longitudinal analysis of ABC customer spending indicated that as individual customers moved from single to multi-channel shopping (and vise versa) their spending levels indicated a marked

[1] ABC is used here as a fictitious identifier of the retail chain.

increase (and decrease). Predictive modeling techniques were used to identify customers with a high potential to become multi-channel shoppers.

In addition more general implications were identified in two areas. First, the problem definition stage of this work provided a reminder of the strategic importance of directing data mining efforts to areas of high profit potential. Second, this research suggests the importance of thinking of multi-channel marketing as more than providing different customers with different alternatives, but rather as an opportunity to develop stronger customer relationships and greater share of spending.

This thesis is organized as follows. The first section discusses the importance of multi-channel marketing in today's marketplace. The second section provides an overview of data mining tools as they apply to marketing. The third section introduces the multi-channel setting of interest and the preliminary analysis that shaped the problem definition phase of the thesis. The fourth section describes the predictive modeling used to identify likely multi-channel users, and the final section presents a summary of the modeling results, as well as more general implications for future applications of data mining in multi-channel marketing settings.

Part I – Multi-Channel Marketing

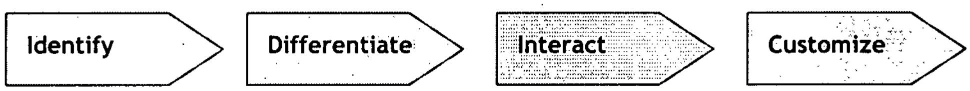
McKinsey & Company, an international consulting firm, makes the distinction between multiple channel and multi-channel strategies[2]. According to McKinsey, a multiple channel strategy involves offering customers multiple points of access. In other words some customers access the firm in one way while other customers use other ways. A multi-channel strategy, on the other hand, actively promotes the synergies that can be achieved by having individual customers using more than one channel. In other words a customer might use the Internet to gather preliminary product information, use a retail outlet to evaluate look and feel, and use a telephone call-line for installation advice.

Why the recent growth in interest in multi-channel marketing? Sears has offered been offering integrated mail order and retail stores for years. In business-to-business settings there is nothing new in the integrated use of catalogues, sales representatives, and trade shows. The "new" aspects of multi-channel integration have to do with technology. The Internet provides new ways of sending and receiving information from customers. Data mining tools makes it possible to track and evaluate customers through the storage and analysis of extensive data sets. And finally, customer relationship management software (CRM) facilitates the coordination of data and communications across a company's multiple customer touch points.

[2] McKinsey Marketing Report, pp3

A prominent proponent of the potential of these new technologies is the 1-to-1 consulting firm of Peppers and Rogers[3]. The 1-to-1 philosophy argues that because of new technologies it is possible to tailor different offerings based on the interests of individual customers. As shown in Figure 1.1, they suggest the IDIC process to achieve this result.

Figure 1.1 – The 1-to-1 Process



The implication is that the offering presented may differ by customer, and that part of this offering may be the channels used to inform, transact, service and interact. Clearly a high level of multi-channel integration is needed for this philosophy of marketing to be effective.

At a more pragmatic level the argument for multi-channel strategies is that the number of customers interested in multi-channel access is growing, and that these customers spend more money. McKinsey research in 2000 suggested that in two or three years, 50% of customers in the US, will be using multiple channels to make purchases. In addition, these multi-channel customers will purchase up to 50% more than single channel customers do[4].

Given the level of interest in multi-channel strategies, what are the difficulties? According to Tom Siebel, founder and CEO of Siebel Systems, "nobody knows

[3] Peppers & Rogers Website, http://www.1to1.com
[4] McKinsey Marketing Report, pp3

how to manage multiple distribution channels well."[5] This is a rather dramatic statement, especially from a man whose company "provides the industry's most comprehensive family of multi-channel eBusiness applications and services."[6]

It seems possible that Siebel's remarks are a comment on the underlying difficulty that marketers have in understanding their customers. Understanding customers' preferences and motivations is difficult when the only contact a company has with their customers is face-to-face in a retail store. With the emergence of e-commerce and growth in telephone and mail order sales, it is ever more difficult to understand customer motivations. The implication is that greater and greater efforts need to be focused on the collection and analysis of customer data, and hence the increasing interest in data mining.

Multi-Channel Marketing at ABC

ABC is a major Canadian retail chain, serving customers from locations in Vancouver, Calgary, Edmonton, Winnipeg, Toronto, Ottawa, and Halifax. Instead of maximizing profits, maximizing customer satisfaction is the paramount goal for ABC management. Providing top-quality merchandise at a reasonable price, and being a leader in environmentally friendly activities are at the core of the ABC approach.

Multi-channel marketing is major part of the ABC approach. In addition to the retail stores, customers have the option of purchasing goods via telephone and

---

[5] Siebel, Tom

[6] Siebel Systems Website, http://www.siebel.com

mail order, or online at www.abc.ca. While the vast majority of current sales come from the retail stores, ABC has a long tradition of catalog based mail order sales. In fact, original ABC customers had only one purchase option, via mail order. In 2001, ABC expanded its multi-channel offering to include an Internet website.

Another special characteristic of the ABC situation is an extensive customer-purchases database. Every ABC customer has a distinct loyalty club number recorded with each purchase. Thanks to an advanced point-of-sale system and relational database, ABC has data on what each customer has purchased, as well as where and when each purchase was made. ABC also has a record of customer names and addresses to facilitate communications with customers. Historically these communications have been limited to catalogs and newsletters, but the opportunity to add marketing communications to select segments of the loyalty club is obvious.

Part II – Data Mining Explained

Data mining has a different definition depending on whom you ask. Berry & Linoff, authors of *Data Mining Techniques for Marketing, Sales, and Customer Support*[7] define data mining as "exploration and analysis, by automatic or semiautomatic means, of large quantities of data in order to discover meaningful patterns and rules." A similar definition is offered by the SAS Institute[8], "the process of data selection, exploration and building models using vast data stores to uncover previously unknown patterns." (The SAS Institute was the leading producer of data mining software in 2001, in terms of market share) The common theme is that data mining involves finding patterns in extremely large data sets.

The amount of data that companies collect is absolutely phenomenal. A few years ago, IBM began what it called the "Terabyte Club" for customers that had amassed one terabyte of data in their IBM data warehouses. One terabyte is equal to approximately one thousand gigabytes or one million megabytes of data. In 2002, IBM reported that there are over 200 customers currently in the "Terabyte Club" and according to Jim P. Kelly, VP of Marketing for IBM's Data Management organization, "...we could be seeing data warehouses packed with petabytes (approximately 1000 terabytes) in the near future."[9]

The following quote is from the 1995 Conference in Knowledge Discovery in Databases (KDD) held in Montreal, "It is estimated that the amount of

---

7 Berry & Linoff, pp 5
8 SAS Institute Website, http://www.sas.com
9 IBM Website, http://www.ibm.com

information in the world doubles every 20 months. What are we supposed to do with this flood of raw data? Clearly little of it will ever be seen by human eyes."[10] Seven years and more then four doubling periods later, marketers are still learning to cope with the vast amount of data they receive on their products and the people that buy them.

While the software to handle these large data sets is relatively new, the statistical techniques that are used in data mining are largely not new at all. These tools are of three general types: OLAP, predictive modeling and clustering. The following provides a general description of each of these tools, along with references that can provide additional detail.

OLAP

Data mining is concerned with finding trends in the data that are difficult to see with simple summary statistics. Often these trends appear when the data is presented in a different form. Many people would agree that graphs are often a more insightful way of presenting data than using a data table. Seeing the data in different formats can allow analysts and managers to gain new insights into their business problems. The desire to query large databases and visualize data on many different dimensions has given rise to OLAP (On-Line Analytical Processing). According to the OLAP Council, OLAP "enables analysts, managers and executives to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information."[11] An

[10] SPSS White Paper, Data Mining: An Introduction, http://www.spss.com
[11] OLAP Council Website, http://www.olapcouncil.org

OLAP server is basically a database that is optimized to query multidimensional data. The OLAP server receives its information directly from a company's data warehouse and stores the data in what is called a "cube". OLAP cubes often have dozens of dimensions depending on management's informational needs.

The major strength of an OLAP cube is the ability to quickly find answers to multidimensional questions. For example, retailers could use an OLAP cube to determine the sales of a particular product in each of their ten stores for the past twelve months. Queries of this type are generally referred to as "slicing and dicing" the data, with the number of queries only limited by the manager's imagination. Managers can use OLAP cubes to perform queries almost instantly without the help of the IS department.

Berry & Linoff identify a number of strengths and weaknesses of OLAP. Some of these strengths and weaknesses are shown below.[12]

OLAP Strengths:

- powerful visualization tool
- provides fast, interactive response times
- good for analyzing time series
- useful to find some clusters and outliers
- many vendors offer OLAP products

OLAP Weaknesses:

- setting up the OLAP cubes can be difficult
- OLAP cubes can become quickly out of date

[12] Berry and Linoff, pp 407-9

The initial time and effort to create the OLAP cubes can be cumbersome and expensive, depending on the size of the company's database, which is a major reason why all companies do not use OLAP. Another weakness of OLAP cubes is that they need to be updated often. Since the OLAP server is usually separate from the data warehouse, they are not always up to date. However the OLAP server can be set to refresh itself on a continuous basis.

There is some debate among data miners as to the role of OLAP in data mining. Berry & Linoff argue that OLAP is not really data mining. They see OLAP as a tool to visualize the data instead of exploiting it. OLAP vendors such as Hyperion (www.hyperion.com) argue that their OLAP solutions can do much more than that. Regardless of which point of view you believe OLAP systems can provide companies with significant cost savings and highlight opportunities for growth. One of Hyperion's success stories came while working with Sears, Roebuck and Co. Sears claims that Hyperion's OLAP solution saved them $10 million (USD) in its first year of operation. Hyperion was able to gather data from Sears' transaction processing system and store it in a way that was conducive to analysis. Queries that took days to run on the old mainframe system took minutes on the new OLAP server. This allowed business analysts to spend more of their time analyzing the data instead of collecting it.[13]

Predictive Modeling

Predictive models use the mass of historical data about a customer to help estimate what he/she is most likely to do in the future. The basic assumption that is implicit in this approach is that "the past will predict the future". As long as this assumption is valid, the predictive model will be useful. A predictive model could be developed to identify characteristics of people that have done something of interest (shopped online, made a donation, etc.), and then used to predict which people are most likely to engage in this activity in the future. For example, a model of current online shoppers might indicate that they were more likely to have children, own three cars, and vote NDP. The implication would be that efforts to attract more online shoppers would be more effective if focused on customers with these characteristics.

The following example is a classic application of predictive modeling to a marketing problem. The UBC Development Office is responsible for soliciting and collecting donations from UBC alumni. Response rates from their mail and telephone solicitations had been quite low for many years the UBC Development Office felt that changes needed to be made in order to increase the response rate. UBC researchers[14] developed a predictive model to identify alumni that would be most likely to donate if contacted. An alumnus' previous solicitation history and demographic variables were key components of this model.

One of the most frequently used tools for predictive modeling is regression. The goal of regression models is to use a handful of predictor variables to find a line that most accurately fits the data. Regression equations often take the form of $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, where y is the target variable x_1 and x_2 are the

[13] Hyperion Website, http://www.hyperion.com

[14] Professor Dan Putler did this research with the assistance of the author.

predictor variables and β_0, β_1, and β_2 are the parameter estimates. For more information about regression models, please see Berry and Linoff, Chapter 7.

Another common modeling technique involves the use of a decision tree. According to Berry and Linoff, decision trees are rule-based tools used for classification and prediction.[15] The output of a decision tree model is reasonably easy to explain to managers, since the output is a tree diagram. However the inner workings of a decision tree is rather complex. Berry and Linoff dedicate an entire chapter to the discussion of decision trees, which is an ideal place to gather more information on this topic.

A third type of predictive model is the neural network model. Originally constructed as a means of mimicking a human neuron in artificial intelligence research, neural network models are extremely difficult to understand. However, they can be very good predictive models under the right circumstances. According to Berry and Linoff, these circumstances are:[16]

- the input data is well understood,
- the variable that you would like to predict is well understood,
- experience is available (ie. a large data set is available to train the model)

Please refer to Berry and Linoff, Chapter 13, for a more in-depth explanation of neural network models.

One of the biggest issues that data miners face is managerial support for their models. Managers often have limited statistical knowledge and need to be convinced of the usefulness of the model before they will allow it to be used in

[15] Berry and Linoff, Chapter 12

practice. One of the best ways to do this is to show how well a model predicts on a data set that was not used to create the model. Using a completely new data set eliminates many worries that the model is only mimicking the data that was used to create the model. Another way to gain managerial acceptance is to use models that are easiest to explain to others. Decision tree models and regression models are reasonably easy to explain, while neural network models are extremely difficult.

Clustering

Data miners and marketers generally use clustering techniques to segment customers. Each customer is described using previous purchase history and other demographic variables that the company has collected. Clustering techniques sift through this data and determine which customers are similar to each other. The result is a clustering solution that defines a number of groups. Within each group, the customers are reasonably homogenous, while different from customers in other groups.

The biggest difficulty associated with clustering is determining the number of clusters needed to provide the best solution. While there are many techniques to help data miners determine the optimal number of clusters, the central concerns are inter vs. intra cluster homogeneity and face validity.

---

[16] Berry and Linoff, pp 290

Berry and Linoff devote a chapter to discussing clustering and go into more detail about the types of algorithms that can be used.[17] It is an excellent source of information about the marketing applications of clustering.

Data Mining Software

A recent search on www.google.com for "data mining software" revealed 735,000 hits. This suggests that there are many data mining software vendors. While the SAS Institute and Hyperion are the most popular data mining and OLAP packages, respectively, many other firms also produce data mining software. A partial list of products and vendors is shown below, for a more complete list visit http://www.kdnuggets.com/software/suites.html.

- Enterprise Miner by the SAS Institute
- Essbase by Hyperion
- Clementine by SPSS
- IBM Intelligent Miner
- Oracle Enterprise Mining Suite
- Affinum Model by Unica Corporation
- Microsoft SQL Server 2000 (actually a relational database, but can construct decision trees and supports clustering)
- Insightful Miner by Insightful Corporation
- Crystal Analysis Professional by Crystal Decisions
- Statistica by StatSoft

Marketing Applications of Data Mining

For the most part, marketers use data mining to learn about customers. Data mining techniques can help segment customers into homogenous groups and

[17] Berry and Linoff, Chapter 10

predict which customers are most likely to respond to an offer or a new product. One of the most common marketing applications of data mining is direct marketing. According to the Direct Marketing Association (DMA) [18], direct marketing can be defined as "any direct communication to a consumer or business recipient that is designed to generate a response in the form of an order, a request for further information, and/or a visit to a store or other place of business for purchase of specific product(s) or service(s)."

What sets direct marketing apart from other types of a marketing communications is that the offer is directed to a specific person or address. Often, the offer is based on little more than a name and an address, but frequently the direct marketer has descriptive information about potential customers. The more information a direct marketer can compile on potential customers, the more tailored the offer can be. This is where data mining techniques can be useful. Data mining techniques can be used to build predictive models or segment customers and hence help target the direct marketing offer to people who are more likely to respond.

---

[18] The DMA is a US based international association of approximately 4700 businesses that are involved in direct marketing, database marketing and interactive marketing. http://www.the-dma.org

Part III – Discussion of Preliminary Research

ABC Cataloging Practices

The initial focus of this research was the evaluation of ABC cataloging practices. Preliminary analysis led to the conclusion that the focus of attention should be shifted from catalog shoppers to multi-channel shoppers. The following section describes the preliminary analysis and the motivation for shifting attention to multi-channel shoppers.

At the outset of this research, ABC cataloging practice involved mailing two general catalogs per year, one in March that covered the full range of summer products and another in August that covered the full range of winter products. In order to be included on the mailing list, ABC customers must have made a purchase in the previous twenty months from any store, including mail order and online.[19] In 2001, this amounted to approximately 700,000 catalogs mailed out every six months. ABC catalogs provide in-depth product information as well as colorful product pictures. Scattered around the catalog are pictures of ABC products "in-action" at various spots on the planet. Together, the information and the often exhilarating photos provide a package that customers find very enjoyable.

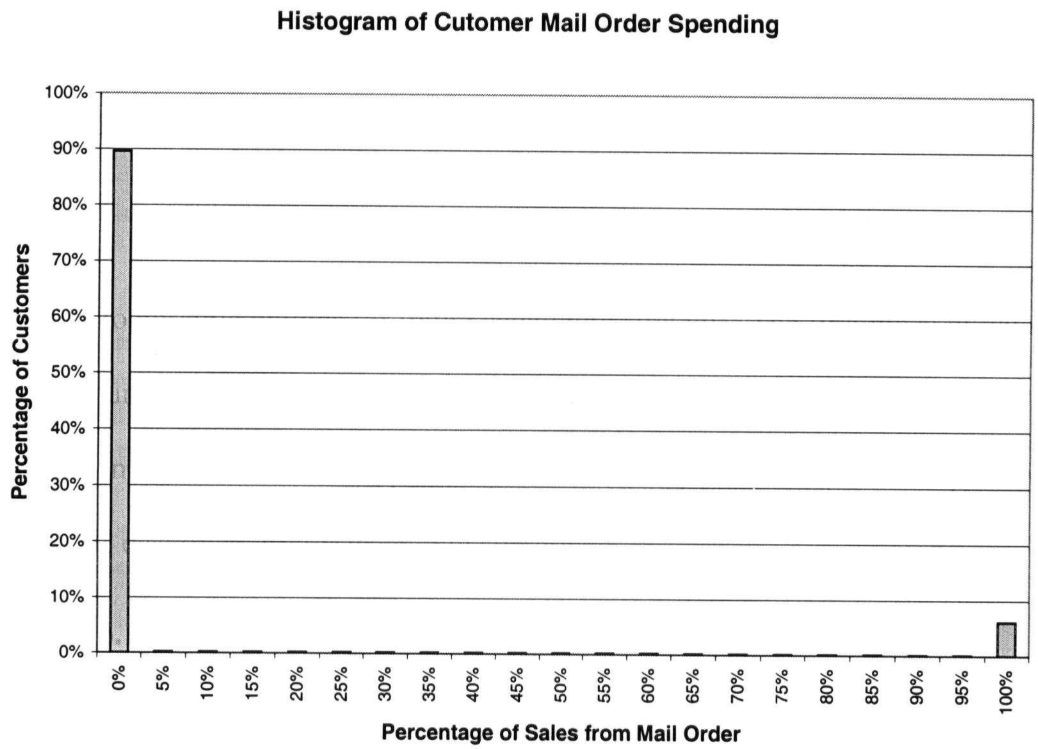
Catalog Impact

The focus of the preliminary research was to determine what impact the catalog had on both store and mail order sales. In addition, this analysis would provide

[19] ABC defines a store to be any of its seven retail locations, its mail order operations and its e-commerce website.

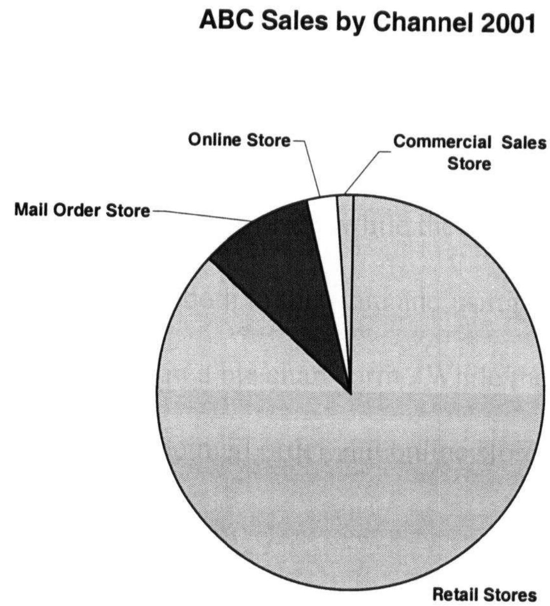
clues as to possible options for future cataloging strategy. The first piece of data that was analyzed involved determining how many customers shopped at stores and how many customers shopped via mail order. Figure 3.1 below clearly shows that ABC customers prefer to shop in stores.

Figure 3.1 – Mail Order vs. Retail Store Spending in 1999



Almost 90% of ABC's 1999 sales came from customers that shopped exclusively in the retail stores. Approximately 6% of these sales came from customers that shopped exclusively via mail order, while the remaining sales came from customers that shopped both in the store and using mail order. Figure 3.2 shows more recent sales data in a pie chart form. While there has been a small amount of migration towards the mail order and online stores, the retail stores still lead the way in sales.

Figure 3.2 – Sales by Channel in 2001



The next step in the analysis was to explore sales trends for the retail stores and mail order. Figures 3.3 and 3.4 show ABC's weekly sales for each sales channel. The cyclical nature of ABC's sales is not very surprising, since many retailers experience a busy summer and Christmas season with a lull in sales during the first few months of the year. Upon closer inspection, there appears to be an increase in sales for the weeks following a catalog mailing, especially in the mail order store.

Figure 3.3 – Weekly Store Sales in 2000

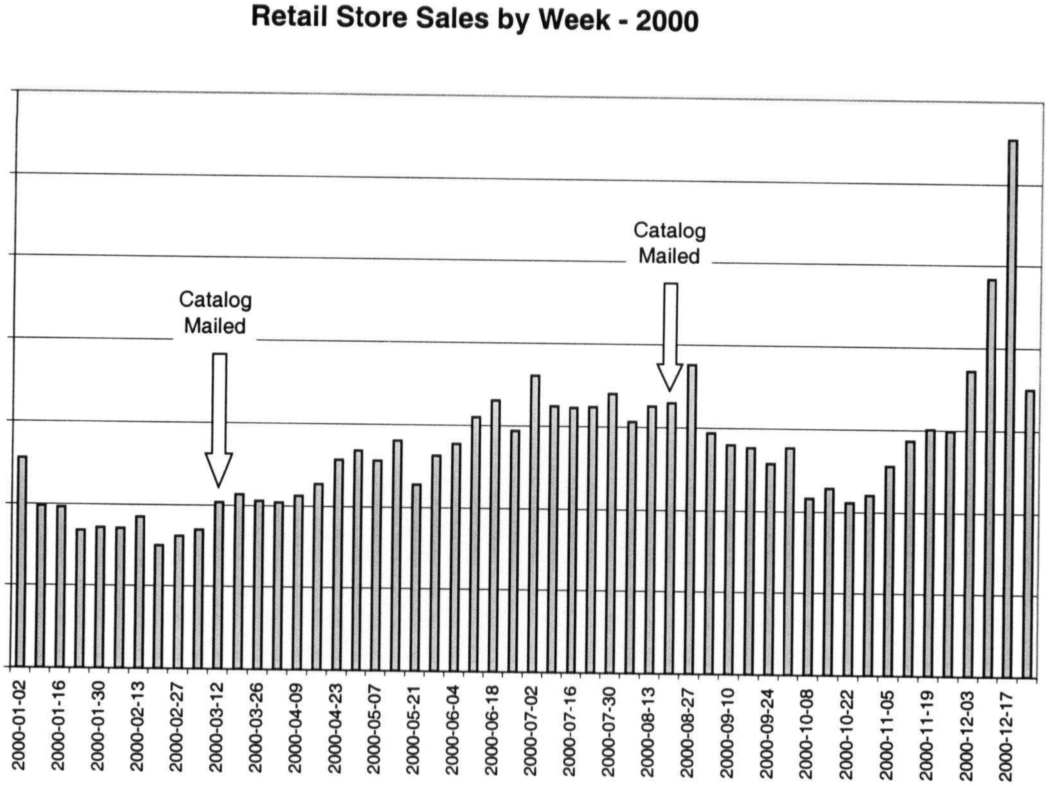
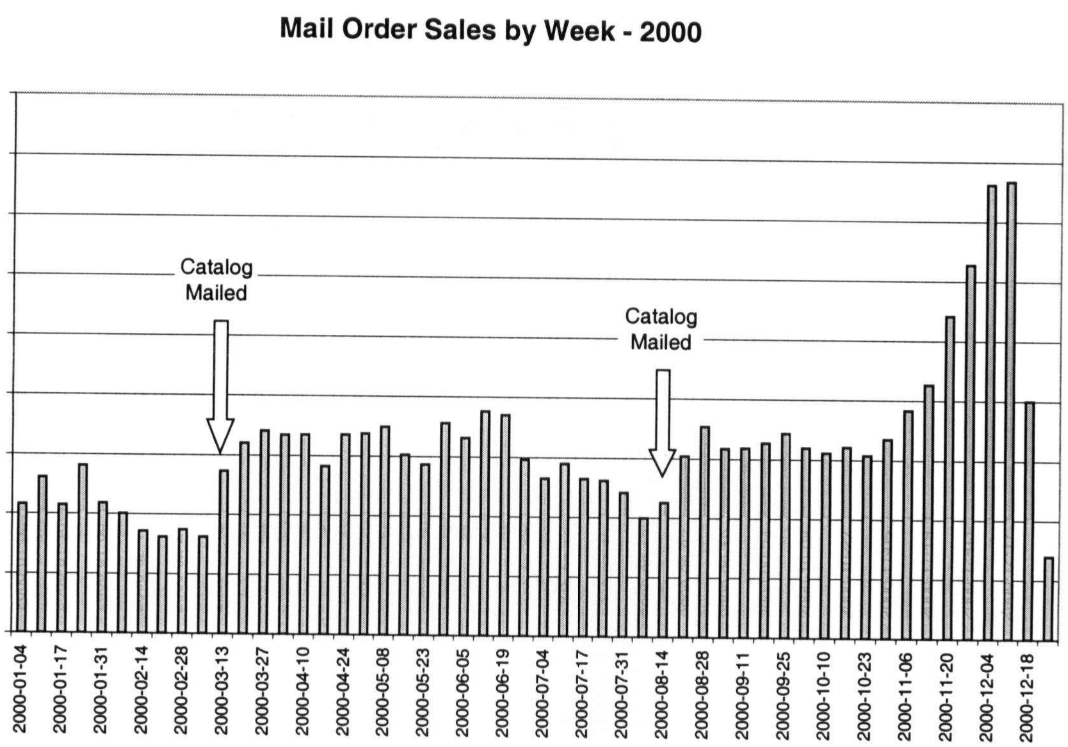


Figure 3.4 – Weekly Mail Order Sales in 2000



If it could be shown that a catalog creates a promotional impact (ie. a spike in sales) then the timing and frequency of catalog mailing is extremely important to the overall cataloging strategy. A promotional impact may provide support for publishing more catalogs per year, as the costs of producing and mailing the catalog would be offset by the gains in contribution from the increased sales. However, if no promotional impact is found then perhaps customers use the catalog more as a reference document. This would imply that the mail-out date is not very important and maybe one large catalog per year would be less costly and as effective.

To determine if the catalog created a promotional impact, two linear regression models were constructed. One model predicted the weekly sales via retail stores while the other predicted weekly sales via mail order (48 months of data were used in each model). The seasonal nature of the sales data complicated the building of the model. It was very difficult to determine if a jump in sales for a particular week was the result of a catalog drop or a seasonal increase. The seasonality needed to be accounted for in the model. Other research suggested that using harmonic variables [sine (x), cosine (x)] could solve these seasonality problems. The use of harmonic variables in regression models is rather rare in the marketing literature, but fortunately it is quite common in the agricultural economics literature. An article by Doran and Quilkey from the *American Journal of Agriculture Economics* provided the following model[20].

---

[20] Doran, H. E., & J.J. Quilkey

$$y_t = \sum_{k=1}^{6} \{\alpha_k \cos \lambda_k t + \beta_k \sin \lambda_k t\} + u_t \text{ where } \lambda_k = 2\pi k / 12 \text{ and } \xi(u_t) = 0$$

In this model, t is the number of months that observations have been taken and α_k and β_k are OLS predictors. The λ_k parameter alters the periodicity of the sine and cosine waves to mimic the seasonality of the data.

After testing the model with the harmonic variables, it was found that three sets of harmonic variables (k= 1, 2, 3) did the best job of controlling the seasonality. In addition to the harmonic variables, the completed model included a time trend (retail$t.trend) and dummy variables for the following weeks:

- four weeks prior to Christmas (retail$xmas4min, retail$xmas3min, retail$xmas2min & retail$xmas1min)
- four weeks following Christmas (retail$xmas1, retail$xmas2, retail$xmas3, retail$xmas4)
- Christmas week (retail$xmas)
- Spring catalog mailing week (retail$sdrop)
- four weeks following Spring catalog mailing (retail$sdrop1, retail$sdrop2, retail$sdrop3, retail$sdrop4)
- Fall catalog mailing week (retail$fdrop)
- four weeks following Fall catalog mailing (retail$fdrop1, retail$fdrop2, retail$fdrop3, retail$fdrop4)
- weeks that included a statutory holiday (stathol$stathol)

The regression output for the two models is shown in Table 3.1 and 3.2.

Table 3.1 – Regression Output for Retail Store Sales Model

```
****   REGRESSION FOR WEEKLY RETAIL STORE SALES    ****

> harmreg3(retail, lambda, stathol)
$SSE
[1] 1.329389

$SUMM

Call:
lm(formula = log(retail$total) ~ retail$xmas4min + retail$xmas3min +
    retail$xmas2min + retail$xmas1min + retail$xmas + retail$xmas1 +
    retail$xmas2 + retail$xmas3 + retail$xmas4 + retail$sdrop +
    retail$sdrop1 + retail$sdrop2 + retail$sdrop3 + retail$sdrop4 +
    retail$fdrop + retail$fdrop1 + retail$fdrop2 + retail$fdrop3 +
    retail$fdrop4 + stathol$stathol + retail$t.trend + cos(2 *
    pi * retail$t.trend/52) + sin(2 * pi * retail$t.trend/52) +
    cos(6 * pi * retail$t.trend/52) + sin(6 * pi * retail$t.trend/52) +
    cos(4 * pi * retail$t.trend/52) + sin(4 * pi * retail$t.trend/52))

Residuals:
     Min       1Q   Median       3Q      Max
-0.27561 -0.04824 -0.00141  0.04693  0.22377

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                 14.4405515  0.0152389 947.614  < 2e-16 ***
retail$xmas4min              0.0533069  0.0536474   0.994 0.321726
retail$xmas3min              0.2016056  0.0567237   3.554 0.000484 ***
retail$xmas2min              0.4279044  0.0595483   7.186 1.71e-11 ***
retail$xmas1min              0.7613053  0.0616202  12.355  < 2e-16 ***
retail$xmas                  0.6183235  0.0615545  10.045  < 2e-16 ***
retail$xmas1                 0.4246492  0.0661957   6.415 1.21e-09 ***
retail$xmas2                 0.0602200  0.0601516   1.001 0.318105
retail$xmas3                -0.0022946  0.0584635  -0.039 0.968735
retail$xmas4                -0.0115396  0.0550668  -0.210 0.834251
retail$sdrop                 0.0465330  0.0502661   0.926 0.355824
retail$sdrop1                0.0592677  0.0507318   1.168 0.244249
retail$sdrop2                0.0786919  0.0514053   1.531 0.127571
retail$sdrop3                0.0159455  0.0516355   0.309 0.757823
retail$sdrop4               -0.0658244  0.0494449  -1.331 0.184784
retail$fdrop                 0.0001819  0.0497244   0.004 0.997085
retail$fdrop1                0.0778693  0.0502847   1.549 0.123242
retail$fdrop2                0.1748713  0.0503965   3.470 0.000652 ***
retail$fdrop3                0.0273771  0.0505212   0.542 0.588562
retail$fdrop4               -0.0452565  0.0492052  -0.920 0.358934
stathol$stathol             -0.0383876  0.0170844  -2.247 0.025859 *
retail$t.trend               0.0016842  0.0001024  16.449  < 2e-16 ***
cos(2*pi*retail$t.trend/52) -0.2069302  0.0152389 -13.579  < 2e-16 ***
sin(2*pi*retail$t.trend/52) -0.1520761  0.0119933 -12.680  < 2e-16 ***
cos(6*pi*retail$t.trend/52)  0.0505282  0.0145062   3.483 0.000622 ***
sin(6*pi*retail$t.trend/52) -0.0658870  0.0113781  -5.791 3.07e-08 ***
cos(4*pi*retail$t.trend/52)  0.0757762  0.0163021   4.648 6.45e-06 ***
sin(4*pi*retail$t.trend/52) -0.0705190  0.0106191  -6.641 3.57e-10 ***
---
Signif. codes:  0  '***'  0.001  '**'  0.01  '*'  0.05  '.'  0.1  ' '  1

Residual standard error: 0.08594 on 180 degrees of freedom
Multiple R-Squared: 0.929,      Adjusted R-squared: 0.9184
F-statistic: 87.29 on 27 and 180 degrees of freedom,    p-value:    0
```

Table 3.2– Regression Output for Mail Order Model

```
****   REGRESSION FOR WEEKLY MAIL ORDER SALES   ****

> catreg(catsales, lambda, catshrld, stathol)
$SUMM

Call:
lm(formula =log(catsales$total) ~ catsales$xmas4min + catsales$xmas3min +
  catsales$xmas2min + catsales$xmas1min + catsales$xmas + catsales$xmas1
  + catsales$xmas2 + catsales$xmas3 + catsales$xmas4 + catsales$sdrop +
  catsales$sdrop1 + catsales$sdrop2 + catsales$sdrop3 + catsales$sdrop4 +
  catsales$fdrop + catsales$fdrop1 + catsales$fdrop2 + catsales$fdrop3 +
  catsales$fdrop4 + retail$t.trend + stathol$stathol +
  cos(4*pi*catshrld[,trend]/52) + sin(4*pi*catshrld[,trend]/52) +
  cos(2*pi*catshrld[,trend]/52) + sin(2*pi*catshrld[,trend]/52) +
  cos(6*pi*catshrld[,trend]/52) + sin(6*pi*catshrld[,trend]/52))

Residuals:
      Min        1Q    Median        3Q       Max
-0.391302 -0.082537 -0.005423  0.077622  0.464496

Coefficients:
                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                    12.3934909  0.0249303 497.126  < 2e-16 ***
catsales$xmas4min               0.0596637  0.0877655   0.680 0.497499
catsales$xmas3min               0.2721864  0.0927982   2.933 0.003792 **
catsales$xmas2min               0.4159814  0.0974191   4.270 3.16e-05 ***
catsales$xmas1min               0.0952144  0.1008087   0.945 0.346177
catsales$xmas                  -0.7270726  0.1007012  -7.220 1.41e-11 ***
catsales$xmas1                 -0.9242371  0.1082942  -8.535 5.77e-15 ***
catsales$xmas2                 -0.3397291  0.0984061  -3.452 0.000693 ***
catsales$xmas3                 -0.2054453  0.0956445  -2.148 0.033050 *
catsales$xmas4                 -0.3565238  0.0900876  -3.958 0.000109 ***
catsales$sdrop                  0.0322693  0.0822337   0.392 0.695219
catsales$sdrop1                 0.3340419  0.0829957   4.025 8.39e-05 ***
catsales$sdrop2                 0.3992087  0.0840975   4.747 4.20e-06 ***
catsales$sdrop3                 0.4320443  0.0844741   5.115 8.00e-07 ***
catsales$sdrop4                 0.3124395  0.0808903   3.863 0.000156 ***
catsales$fdrop                 -0.0876570  0.0813476  -1.078 0.282672
catsales$fdrop1                 0.0630454  0.0822642   0.766 0.444456
catsales$fdrop2                 0.2835257  0.0824472   3.439 0.000726 ***
catsales$fdrop3                 0.1138964  0.0826511   1.378 0.169902
catsales$fdrop4                 0.0561835  0.0804982   0.698 0.486110
retail$t.trend                  0.0012550  0.0001675   7.492 2.94e-12 ***
stathol$stathol                -0.1087878  0.0279495  -3.892 0.000140 ***
cos(4*pi*catshrld[,trend]/52)   0.1459401  0.0266697   5.472 1.47e-07 ***
sin(4*pi*catshrld[,trend]/52)  -0.2933686  0.0173726 -16.887  < 2e-16 ***
cos(2*pi*catshrld[,trend]/52)   0.0521726  0.0249304   2.093 0.037775 *
sin(2*pi*catshrld[,trend]/52)  -0.1881585  0.0196206  -9.590  < 2e-16 ***
cos(6*pi*catshrld[,trend]/52)   0.1390270  0.0237317   5.858 2.18e-08 ***
sin(6*pi*catshrld[,trend]/52)   0.0059694  0.0186142   0.321 0.748816
---
Signif. codes:  0  '***'  0.001  '**'  0.01  '*'  0.05  '.'  0.1  ' '  1

Residual standard error: 0.1406 on 180 degrees of freedom
Multiple R-Squared: 0.8726,     Adjusted R-squared: 0.8535
F-statistic: 45.65 on 27 and 180 degrees of freedom,    p-value:    0
```

Summary of Regression Analysis

a) Summary Retail Store Sales Model (Table 3.1)

- Overall Fit: Adjusted R-squared of 0.92
- Significant Variables:
    - All six seasonality variables (sine and cosine variables)
    - Retail time trend variable (indicates a positive trend in sales)
    - Five Christmas variables (greatest sales during the period three weeks before Christmas to one week after)
    - Statutory holiday variable (low sales in weeks when stores are closed for one day)
    - One catalog mailing variable (increased sales two weeks after the fall mailing)

b) Summary Mail Order Store Sales Model (Table 3.2)

- Overall Fit: Adjusted R-squared of 0.85
- Significant Variables:
    - All six seasonality variables (sine and cosine variables)
    - Retail time trend variable (positive trend in mail order sales)
    - Christmas variables (greatest sales 2-3 weeks prior to Christmas, sharp decline in sales during Christmas week and the four weeks following)
    - Stat holiday variable (low sales in weeks with one day closed)
    - Catalog mailing variables (increased sales in the three weeks following spring catalog launch in the second week following the fall catalog launch)

c) Implications

In brief, neither the spring catalog nor the fall catalog creates a promotional impact in the retail store sales, except for a mysterious increase in sales two weeks following the fall launch. This increase in sales is likely due to last minute "back-to-school" shopping rather than a catalog launch. Looking at the mail order sales model, it is quite apparent that there is a statistically significant promotional impact for the spring catalog mailing. This promotional effect is not as strong for the fall launch mailing.

Justifying an increase in the number of catalogs published per year based on the previous evidence does not appear to be warranted. While there was evidence of a promotional impact for mail order sales, the size of the mail order market is too small to support the costs of an extra catalog.

Since there was limited promotional impact on store sales, the next question was "can we reduce the number of catalogs sent without impacting sales?" In order to do this catalogs must be delivered only to customers that were most likely to purchase. The costs associated with producing and mailing a catalog, and the cost savings that might be achieved can be seen Table 3.3 below.

Table 3.3 – Cost Savings from reducing the number of catalogs mailed[21]

Catalog	Catalogs Mailed	Cost per Catalog Mailed
Spring/Summer	700,000	$2.89
Fall/Winter	700,000	$2.28

Reduction in Mailed Catalogs	Cost Savings (Spring/Summer)	Cost Savings (Fall/Winter)	Total Savings
50,000	144,513	113,883	258,396
75,000	216,770	170,824	387,594
100,000	289,027	227,765	516,792

Table 3.3 suggests that reducing the number of catalogs mailed by 75,000 per mailing or about 10% would save ABC about $390,000 per year. Preliminary

[21] Costs were multiplied by a constant to maintain confidentiality.

plans were formulated for a predictive model to determine which customers would be potential targets, but it became apparent that this might not be the best solution.

The concern over this particular model was that it might be an efficient solution, but not an effective solution. In other words, a model focused on increasing the efficiency of current cataloging practice was not likely to have as much impact as efforts aimed at a strategic change in practice.

In order to find a new effective strategic solution, it was clear that further analysis of customer data was needed.

Customer Analysis

The first step in the customer analysis phase of this project was to define a suitable sample of ABC customers to analyze. Since there are more than 1.7 million customers in the loyalty club, some sampling decisions were necessary. After consulting with ABC management, a decision was made to track every customer that had purchased from mail order and/or online at least one time during the five year period between 1997-2001. The online store at www.abc.ca opened in March 2001, so only the final nine months of data included online sales.

A SQL query was developed to identify these mail order/online shoppers. Approximately 200,000 customers were identified. Using this "hit-list", each customer's entire purchase history for the 1997-2001 period was extracted from the ABC database. It was somewhat complicated to design the query to extract the data since it was important to retrieve every purchase that these customers made, whether it was from a retail store, the mail order store, or the online store. In the end, more than 4 million records were extracted. Each record contained the customer number of the purchaser, the store where the sale was made, the date of the purchase, the cost of the product and information regarding the department that the product came from. ABC divides all of its products into one of approximately thirty departments. Information was also gathered about the individual customers that were part of the "hit-list." Each customer's distinct loyalty club number as well as their postal code, province and the date they joined the loyalty club were collected. While ABC also collects addresses, phone numbers and other sensitive information, this information was not used in order to protect customer anonymity.

Once the data was extracted from ABC's database and imported into SAS, the analysis could begin. Classifying each customer based on which channels they purchased from quickly paid dividends. The five-year window was cut into three periods. The first period included 1997 and 1998 data, while the second period included 1999 and 2000 data. A two-year period was chosen due to the fact that the inter-purchase time for ABC customers is generally longer than it is for most other retailers. 2001 data was placed in the third period due to the fact that the online store opened in that year.

For periods 1 and 2, each customer was given a code to signify which channels they purchased from in that period:

1. retail stores only (multiple store locations allowed)
2. mail order store only
3. both retail and mail order stores
4. did not purchase during this period
5. was not a customer in this period

Customers were placed into groups according to their channel purchasing habits and the average sales for each group of these groups was calculated. Table 3.4 below shows the results of this analysis.

Table 3.4 – Analysis of Customers based on their channel usage.[22]

	Period 1 (1997-98)		Period 2 (1999-2000)	
Group	Members	Indexed Average Sales for Period	Members	Indexed Average Sales for Period
Retail	40,430	159.02	48,289	171.25
Mail Order	40,872	100.00	47,784	105.08
Both	51,661	254.80	55,615	262.93
No Purchase	13,453	0.00	31,019	0.00
Not a Member	55,357	0.00	19,066	0.00

It is quite easy to see that in both periods, customers that purchased in both the retail and the mail order stores spent much more than customers that shopped in only one channel. This was a very strong finding and deserved further attention. The quest for the multi-channel shopper had begun.

[22] Average sales for mail order in Period 1 was used as the index base.

The Quest for the Multi-Channel Shopper

In an attempt to find more conclusive evidence to support the hypothesis that multi-channel shoppers were a lucrative segment of the ABC customers, the groups from Table 3.4 were split based on their behavior across Periods 1 and 2. The resulting table is shown as Table 3.5 on the next page.

Look carefully at the group of customers that purchased in the retail store in Period 1 (S) and in both channels in Period 2 (S&O). Their purchases increased by over 60% from Period 1 to Period 2. The same trend is evident when looking at the customers that purchased from the mail order store in Period 1 (O) and in both channels in Period 2 (S&O). This time the percentage increase in sales is 87%. As a customer moves from single-channel shopping to multi-channel shopping, their spending increases significantly.

Another interesting trend that is apparent in Table 3.5 is that if a customer moves from a multi-channel shopper to a single-channel shopper, their spending decreases. Customers that purchase from both channels in Period 1 (S&O) but from only the retail store (S) or only the mail order store (O) in Period 2 spent significantly less (30% and 41%, respectively) in Period 2 than they did in Period 1.

For <u>every</u> group in which, customers chose to change their behavior and shop from multiple channels, average sales increase. Conversely, when customers of a multi-channel group in the first period shopped in only a single channel, average sales decrease. The importance of this analysis is that it shows that there are <u>not</u>

single channel and multi-channel shoppers. Rather, individual shoppers move from single channel to multi-channel (and vice versa). The implication being that marketing activities might be used to increase the shift from single channel to multi-channel.

Table 3.5 – Time Based Behavior Analysis of ABC Customers[23]

Period 1	Period 2	Group Members Number	Indexed Av Group Member Spending 1	Indexed Av Group Member Spending 2
S	S	13,513	261.94	304.54
	O	5,883	113.88	101.36
	S&O	18,317	224.61	362.46
	NP	2,717	109.24	0
O	S	5,024	100.00	113.15
	O	12,887	127.23	182.72
	S&O	6,096	171.68	321.55
	NP	16,865	102.05	0
S&O	S	20,674	347.02	245.84
	O	6,032	290.92	170.26
	S&O	16,638	417.31	434.55
	NP	8,317	213.80	0
NP	S	2,068	0	129.12
	O	5,161	0	114.31
	S&O	3,602	0	251.75
	NP	2,622	0	0
NM	S	7,010	0	150.50
	O	17,821	0	121.47
	S&O	10,962	0	264.17
	NP	498	0	0

Legend

S	Retail Store Only
O	Mail Order Only
S&O	Both Retail and Mail Order Store
NP	No Purchases
NM	Not A Member

[23] Average group customer spending for customers switching from Mail Order Only in Period 1 to Store Only in Period 2 was used as the index base.

In Period 3, there was a natural experiment that can be used to discover if the multi-channel trend continues. ABC opened its online store in March 2001 so nine months of online store data can be analyzed. Adding a new channel to the mix will provide a strong test to the hypothesis that as a customer moves from single channel purchasing to multi-channel purchasing their sales increase. Table 3.6 on the next page shows the results of this analysis.

After looking at the results in Table 3.6, there is little doubt that this trend is real. In every case where a group of customers increased the number of channels shopped, their average purchases increased. Also, every time a group of customers choose to shop in fewer channels their purchases declined.

Many companies can produce tables similar to Table 3.4, which shows that multi-channel customers are comparative "big spenders". What is difficult to discern from this snapshot is the nature of causality. Specifically, do big-spenders cause multi-channel behavior or does multi-channel behavior cause customers to be big-spenders. The data presented above appears to support the latter argument. By following groups of people over several periods of time, we can see that ABC customers migrate back and forth from multi-channel to single channel and vice versa. We also know that sales rise when customers shop in more channels and fall when they shop in fewer channels. In any case, we can see that the multi-channel segment of ABC customers is not totally comprised of loyal ABC customers who always buy from multiple channels.

Table 3.6 – The Multi-Channel Trend continues in Period 3[24]

Period 2	Period 3	Group Members Number	Indexed Av Group Member Spending 2 — 2 Year Av	Indexed Av Group Member Spending 2 — 1 Year Av	Indexed Av Group Member Spending 3 — 1 Year Av
	S	13,507	507.00	253.50	274.40
	O	5,459	220.79	110.39	141.29
	W	1,991	224.71	112.36	115.23
S	S&O	13,311	539.45	269.73	504.38
	S&W	3,777	491.84	245.92	429.64
	O&W	337	263.60	131.80	280.03
	S&O&W	702	584.18	292.09	710.15
	NP	9,432	209.97	104.99	0
	S	5,263	207.36	103.68	176.28
	O	11,032	353.59	176.79	224.54
	W	1,101	236.02	118.01	142.65
O	S&O	3,166	354.02	177.01	415.99
	S&W	339	234.43	117.21	345.38
	O&W	696	421.06	210.53	381.32
	S&O&W	276	432.15	216.07	649.36
	NP	25,911	200.00	100.00	0
	S	18,841	674.96	337.48	261.59
	O	6,906	600.56	300.28	223.23
	W	788	486.88	243.44	120.65
S&O	S&O	9,992	873.20	436.60	538.94
	S&W	1,100	725.16	362.58	425.53
	O&W	635	691.12	345.56	378.70
	S&O&W	1,058	1005.42	502.71	724.32
	NP	16,295	420.66	210.33	0
	S	2,866	0	0	143.85
	O	3,914	0	0	153.78
	W	1,126	0	0	117.41
NP	S&O	1,814	0	0	368.15
	S&W	548	0	0	285.00
	O&W	186	0	0	346.62
	S&O&W	89	0	0	481.91
	NP	20,653	0	0	0
	S	0	0	0	0
	O	7,571	0	0	165.67
	W	5,118	0	0	129.49
NM	S&O	3,689	0	0	388.65
	S&W	1,798	0	0	342.45
	O&W	578	0	0	331.68
	S&O&W	312	0	0	578.82
	NP	0	0	0	0

---

[24] Average yearly spending for customers that shopped in the Mail Order Store in Period 2 and did not purchase in Period 3 was used as the index base.

Table 3.6 (cont.) – The Multi-Channel Trend continues in Period 3

Legend

S	Retail Store Only	S&W	Both Retail and Online Stores
O	Mail Order Only	O&W	Both Mail Order and Online Stores
W	Online Store Only	S&O&W	All Three Channels
S&O	Both Retail and Mail Order Stores	NM	New Member in Period 3
S&W	Both Retail and Online Stores	NP	No Purchases

It is reasonable to think that one of the reasons why multi-channel customers spend more at ABC is that ABC is doing a better job of meeting their needs as a consumer of outdoor equipment. ABC has done very little to promote multi-channel purchasing, but these customers have been smart enough to realize that there are advantages to purchasing through different channels. What might happen if ABC decided to actively target potential multi-channel customers from the vast majority of customers that shop exclusively through the retail stores?

Multi-Channel Market Potential

Estimating some measure of market potential is generally an important factor in determining which opportunities are worth pursuing. In the discussion in Section 3 on alternative cataloging strategies, an opportunity appeared that might save ABC almost $400,000 by reducing the number of catalogs it sent to customers. While a $400,000 cost savings is not insignificant, the goal here is to look for opportunities to increase sales and contribution.

Consider the following four opportunities. ABC could:

- attempt to increase the number of new customers attracted to mail order.
- attempt to increase the number of new customers attracted to online shopping.
- attempt to increase the number of new customers to the retail stores.
- attempt to convert store-only customers to multi-channel shoppers.

a) Increasing the number of new customers attracted to mail order

New Mail Order customers in 2001:	7,571
Average spending per customer:	$328.33*
Total Spending for this group:	$2,485,770

Suppose a marketing initiative increased the number of new mail order customers by 10%

The impact of this campaign would be: **$248,577**

b) Increasing the number of new customers attracted to online shopping

New Online customers in 2001:	5,118
Average spending per customer:	$256.64*
Total Spending for this group:	$1,313,467

Suppose a marketing initiative increased the number of new online customers by 10%

The impact of this campaign would be: **$131,347**

c) Increasing the number of new customers attracted to retail stores

New Retail Customers in 2000 (approx.):	139,000
Average spending per customer:[25]	$289.23*
Total Spending for this group:	$40,203,071

Suppose a marketing initiative increased the number of new mail order customers by 10%

The impact of this campaign would be: **$4,020,307**

---

* These numbers have been multiplied by the same constant used on page 24 to maintain confidentiality.

[25] This number was calculated from a sample of ABC customers that had purchased from the retail store only.

d) Convert store-only customers to multi-channel shoppers

Customers purchased in 2001 (approx.)	500,000
Percentage Retail Store-only (approx.)	85%
Retail Store only Customers in 2001	425,000
Conservative Increase in Sales per converted Multi-Channel Customer	$281.17*

Suppose a marketing initiative could convert 10% of retail store customers to multi-channel

The impact of this campaign would be: **$11,949,833**

While it should not be assumed that achieving a 10% increase in each of these options is equally easy (or equally costly), the relative magnitudes of the outcomes is a strong argument for making the multi-channel conversion initiative a top priority.

Part IV – Modeling Multi-Channel Shopping

The preliminary analysis indicates that a key question is "which retail store only customers can be converted to multi-channel customers?" This is a classic marketing application of a predictive model. For this particular case, a logistic regression model is a useful approach, because it will attach a probability of conversion to each customer. SAS Enterprise Miner was used to develop this model.

A data set of 35,225 customers was created in SAS from customers that had purchased from the retail store only in Period 1 (1997-98). 17,225 of these customers had converted to multi-channel customers in Period 2 (1999-2000), while the remaining customers continued to purchase exclusively from the retail stores only. This data set was "oversampled" with multi-channel customers. Oversampling is very common in data mining, because often the target group (in this case, multi-channel customers) is very small in comparison to the non-target group (single channel customers). By oversampling the data set, the model receives enough data on the target group to make a prediction. The goal of the model was to predict which of the customers converted to multi-channel and which stayed as single-channel customers.

Data Available

Data was collected for all customers who had made one or more purchases via a non-store channel. For each of these customers, a complete history of purchases over a 5-year period was compiled. Data on product description, spending amount, channel, and date was collected for each purchase, while for each customer, a postal code and the loyalty join date was collected as well. In addition, similar data was collected for a sample of customers that have purchased at least once in the 5-year period, but shopped exclusively in the retail stores.

This data was structured to form the following variables that were available as predictor variables in this model:

1. Total $ Purchases in Period 1
2. Number of Purchases in Period 1
3. $ Purchases in Product Segment 1
4. $ Purchases in Product Segment 2
5. $ Purchases in Product Segment 3
6. $ Purchases in Product Segment 4
7. $ Purchases in Product Segment 5
8. $ Purchases in Product Segment 6
9. $ Purchases in Product Segment 7
10. $ Purchases in Product Segment 8
11. $ Purchases in Product Segment 9
12. How many months since the customer joined loyalty club
13. The distance to the closest retail store
14. The distance to the closest retail store less than 25 km (dummy variable)
15. The distance to the closest retail store between 25-50 km (dummy variable)

While each product that ABC sells come from one of about 30 product categories, for the purposes of this model, product categories were amalgamated into six product segments. Table 4.1 below describes each product segment.

Table 4.1 – ABC Product Segments

Product Categories	
1	Clothing
2	Accessories
3	Housewares
4	Appliances
5	Sporting Goods
6	Automotive

A potentially important variable in this data set is distance to the closest ABC retail store. Computing this distance was a challenging process requiring patience and spherical trigonometry. The first step was to collect postal code information for ABC customers. Fortunately, ABC captures this information. The second step involved converting these postal codes into enumeration areas (EA). This postal code to EA conversion is necessary since Statistics Canada publishes a latitude and longitude for the centroid of each EA. Using the latitude and longitude of each customer's EA, together with the latitude and longitude of the nearest store, it was possible to calculate the distance (as the crow flies) that a customer lives from a store.[26] The final step involved trying to find a formula that would do this calculation. A search on www.google.ca led to a US Census Bureau website (http://www.census.gov/cgi-bin/geo/gisfaq?Q5.1) that explained the pros and cons of different formulas. The Haversine Formula was judged to be the best as it is not vulnerable to round-off errors when the two points are close together.[27] The Haversine Formula is shown on the next page.

[26] A better measure of distance would be driving distance to the nearest ABC retail store. Unfortunately, a complete set of Canadian street network maps was not available at this time.

[27] Sinnott, R.W.

$$
\begin{aligned}
&dlon = lon2 - lon1 \\
&dlat = lat2 - lat1 \\
&a = [\sin(dlat/2)]^2 + \cos(lat1) * \cos(lat2) * [\sin(dlon/2)]^2 \\
&c = 2 * \arcsin\{\min[1, sqrt(a)]\} \\
&d = R * c, \text{ where } R = \text{radius of Earth}
\end{aligned}
$$

Using the Haversine formula, a distance from each customer to each of the 7 ABC retail store locations was calculated and the shortest distance was used as the distance to closest store variable.

Preliminary modeling

The data was loaded into SAS Enterprise Miner and the model building process began. After building preliminary models using logistic regression, decision trees and neural networks, the logistic regression model was chosen for the final model.

Results of the Logistic Regression Model

A stepwise logistic regression model was run using the 15 variables shown above. Stepwise regression uses an iterative approach to adding variables to the model. The computer runs the model with every variable and then decides which variable does the best job of predicting. This variable is then added to the model. The model is re-run with the first variable and each of the remaining variables and decides which new variable is the best predictor. The computer runs the model over and over until there are no more variables that are statistically significant (95% confidence level). Table 4.2 shows the variables that became part of the logistic regression model.

Table 4.2 – Predictor Variables from Logistic Regression Model

Variable Name	Parameter Estimate	T-Score
Intercept	0.362	8.14
Distance Less Than 25km	-1.983	-48.66
Distance Between 25-50 km	-1.228	-19.65
Distance to Closest Store	0.00122	13.01
Months Since Joining ABC	0.00151	5.51
Segment 1 Purchases (Clothing)	0.000932	8.28
Segment 6 Purchases (Automotive)	-0.000593	-2.38
Number of Purchases in Period 1	0.025467	4.51
Purchases in Period 1	0.000418	6.03

To see how the model affects different customers, Table 4.3 shows two customers and how they scored on the model.

Table 4.3 – Comparison of Jane's and Tom's score on the logistic regression model

Variable Name	Parameter Estimate	Jane	Jane's Score	Tom	Tom's Score
Intercept	0.362	1	0.362	1	0.36
Distance Less Than 25km	-1.983	No	0.000	Yes	-1.98
Distance Between 25-50 km	-1.228	No	0.000	No	0.000
Distance to Closest Store	0.00122	141	0.17211	9	0.0110
Months Since Joining ABC	0.00151	102	0.15376	36	0.0543
Segment 1 Purchases (Clothing)	0.000932	159	0.148115	0	0.000
Segment 6 Purchases (Automotive)	-0.000593	20	-0.011858	0	0.000
Number of Purchases in Period 1	0.025467	9	0.229203	2	0.0509
Purchases in Period 1	0.000418	1351	0.564469	21	0.00873
		Total	1.618	Total	-1.496
		Probability	0.834	Probability	0.183

As you can see, Jane has a far higher probability of converting to a multi-channel shopper than Tom does. The main reason for the difference in probabilities is the fact that Jane lives much further away from an ABC store than Tom does. Another reason is that Jane spent $1351 at ABC during 9 shopping trips while Tom spent only $21 during his 2 visits.

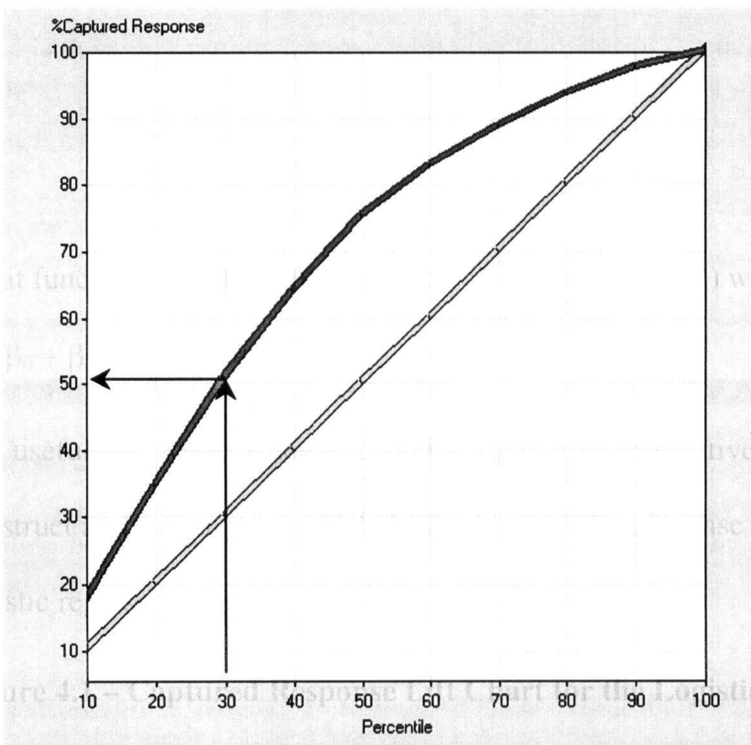
Notice that the "Total" for Jane and Tom is different than their "Probability". The Total is the sum of all the scores, while the Probability is calculated using the

Logit function. The logit function is: $P = e^y \div (1+e^y)$ where

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \varepsilon$

The useful way to tell whether a model is providing predictive power is to construct a lift chart. Figure 4.1 shows the captured response lift chart for the logistic regression model.

Figure 4.1 – Captured Response Lift Chart for the Logistic Regression Model



The captured response lift chart measures the gains in efficiency by using the predictive model. For example, if you randomly chose to look at 30% of the customers in the data set to see how many were multi-channel shoppers in Period 2, you would find that you had about 30% of the entire list of multi-channel shoppers. This is to be expected as you chose randomly from the entire data set. What the lift chart shows is that if you used the regression model to rank each customer by their probability of being a multi-channel shopper and then chose to look at the top 30% of that list you would have about 50% of the multi-channel shoppers. This gain in efficiency is called "lift". In this case the lift at 30% is about 1.7 (51% ÷ 30% = 1.7)

Another way to evaluate the model is to look at how many correct predictions were made by the model. The easiest way to see this is by producing a table showing the predicted behavior vs. the actual behavior, as is seen below in Table 4.4.

Table 4.4 – Cross-tab Showing Correct and Incorrect Predictions

		Model Predicted Behavior					
		Single		Multi			
Actual Behavior	Single	8,223	39%	2,420	11%	10,643	50%
	Multi	2,930	14%	7,562	36%	10,492	50%
		11,153	53%	9,982	47%	21,135	

Table 4.4 shows that the model correctly predicts a customer's behavior in Period 2 (1999/2000) 75% of the time (39% + 36%). This is much better than the 50% probability of picking a multi-channel customer from the data set randomly.

To double check that the model is providing useful predictions the model was used to predict behavior in Period 3 (2001). This is a good test of the model as this is an entirely different data set than the one used to create the model. In addition, using a more recent data set would determine how robust the model was. Often models need to be revised over time so that they maintain their predictive power. The results of this test are shown in Table 4.5 below.

Table 4.5 – Cross-tab Showing Results of Model Test

		Model Predicted Behavior					
		Single		Multi			
Actual Behavior	Single	15,422	41%	4,578	12%	20,000	54%
	Multi	7,364	20%	9,867	27%	17,231	46%
		22,786	61%	14,445	39%	37,231	

When applied to the new data set, the model did not predict quite as well as it did on the previous data; 68% of the customers were correctly identified as single channel or multi-channel customers. Looking at Table 4.5, it is clear that the model was under-predicting multi-channel customers. It is predicting that a customer will be a single channel shopper while in reality that customer was actually a multi-channel shopper. This problem may suggest that the model has lost predictive power and needs to be revised with new data, perhaps due to a change in ABC customer's buying habits in 2001.

To determine whether or not the model has lost predictive power, a new model was created, using Period 2 data to predict behavior in Period 3. The results of this model are shown in Table 4.6.

Table 4.6 – Cross-tab Showing Results of New Model

		Model Predicted Behavior					
		Single		Multi			
Actual Behavior	Single	8,937	40%	3,075	14%	12,012	54%
	Multi	3,744	17%	6,583	29%	10,327	46%
		12,681	57%	9,658	43%	22,339	

Looking at Table 4.6, you can see that the new model predicted correctly 69% of the time. This is very similar to the first model that predicted 68% correctly on the same year's data. This provides evidence for the scenario that something has changed ABC customer's buying habits. In fact there was a fundamental change in 2001, the introduction of the online store in March.

The addition of a new channel for ABC shoppers to purchase from has increased the total number of multi-channel shoppers in a way that that model cannot predict at this time. Unfortunately, 2001 is the last full year of data that could be extracted from ABC's database. In early 2003, it would be interesting to gather the data from 2001 and 2002 and create a new logistic regression model that may be better at determining multi-channel shoppers for 2003.

The findings of the logistic regression model indicate that customers who are most likely to be converted to multi-channel customers:

- live further from ABC
- have been a member of the loyalty club for a longer period
- make more frequent retail store purchase trips
- spend larger amounts of money at ABC (especially in the clothing segment)

Part V – Summary

The primary goal of this thesis was to provide analysis to help ABC enhance its multi-channel services. The preliminary analysis focused on cataloging practices and found that the vast majority (90%) of ABC customers purchased from the retail store exclusively. Weekly sales data from both the retail store and the mail order store was analyzed to determine whether or not the mailing of a catalog produced a promotional effect. Two regression models were produced, using harmonic variables (sine & cosine) to control the seasonality in the data. In each of the models, the harmonic variables and a linear time trend were key predictors of weekly sales as were dummy variables signifying the Christmas buying season and statutory holidays. For the retail store sales, it appeared unlikely that there was an increase in sales in the weeks following a catalog mailing, however there was evidence of a promotional impact after a catalog mailing for mail order sales.

Next, financial analysis was performed to discover the cost savings gained by reducing the number of catalogs sent to customers. This analysis showed that reducing the mailing list by 10% would provide cost savings of approximately $400,000.

Then an attempt was made to discover ways to increase revenues and contribution. Customer channel-purchase data was gathered and analyzed. The results of this analysis indicated that customers that had purchased from multiple channels spent more than those customers that purchased from only one channel. To cast light on the causality behind these findings, longitudinal analysis of these customers was completed. The results of this analysis showed that in <u>every</u> case when a customer moved from single-channel behavior to multi-channel behavior, spending increased. Similarly, when a customer moved from a multi-channel behavior to single-channel behavior, spending decreased.

Market potential analysis was done next to determine whether potential payoffs of converting store-only customers to multi-channel customers. This was compared to several other courses of action. The multi-channel opportunity indicated greater potential than all other alternatives.

With the course of action identified, a logistic regression model was chosen to predict which customers were most likely to switch from single-channel to multi-channel shoppers. A major predictor variable for this model involved the distance that a customer lived from a ABC store (Distance was calculated by converting postal codes to EAs and using StatsCan information on EA latitude and longitude).

The final model indicates that customers that live more than 50km from an ABC store are more likely to be converted to multi-channel, as are customers that make frequent trips to ABC and spend large amounts of money. Customers that have been with ABC for a long period of time are also more likely to convert to multi-channel customers.

An overarching purpose of this thesis was to learn how data mining techniques can support multi-channel marketing initiatives. One of the biggest lessons learned was that data mining techniques can easily be focused on the wrong task. In other words, data mining techniques can make any solution more efficient, regardless of whether the solution is effective or not. It is important, as both a marketer and a data miner to know the difference between an efficient solution and an effective solution. The difference between the two can be very subtle, and require clear thinking to recognize which solutions are efficient and which solutions are improvements to inferior practices and which are focused on a superior practice.

From a multi-channel perspective, an important lesson that can be learned is that a company cannot provide multiple channels for their customers and expect the customers to find them. Retailers need to actively promote their alternative channels through every different type of media available. Customers must be reminded that purchasing from alternative channels can be just as useful as purchasing from a retail store. In some cases, using an alternative channel can be much more convenient.

Multi-channel retailing is becoming more and more important. In the coming years, more people are going to use multiple channels to purchase their goods and services. Customers will become increasingly familiar with purchasing items from mail order or online stores. While a few items may still require a visit to a retail store, many others can be purchased more quickly, easily and cheaply via alternative channels. I predict that retailers that are proactive in multi-channel marketing initiatives will become the retail leaders of the future.

Bibliography

Berry, Michael J.A., and Gordon Linoff. *Data Mining Techniques For Marketing Sales and Customer Support.* J. Wiley & Sons Publishing Co., 1997

Direct Marketing Association Website, http://www.the-dma.org

Doran, H.E., and J.J. Quilkey. "Harmonic Analysis of Seasonal Data: Some Important Properties". *American Journal of Agricultural Economics.* 54(1972):646-651.

Hyperion Website, http://www.hyperion.com

IBM Website, http://www.ibm.com

Kdnuggetts (Knowledge Discovery) Website, http://www.kdnuggets.com/software/suites.html

McKinsey Marketing, "Multi-Channel Marketing: Making Bricks and Clicks Stick" 2000
http:// marketing.mckinsey.com/solutions/McK-Multi-Channel.pdf

OLAP Council Website, http://www.olapcouncil.org

Peppers & Rogers Website, http://www.1to1.com

SAS Institute Website, http://www.sas.com

Siebel Systems Website, http://www.siebel.com

Siebel, Tom. "Internet Explosion Has Sparked A Fundamental Change" *Financial Times.*
June 7, 2000

Sinnott, R.W. "Virtues of the Haversine." *Sky and Telescope* 68(1984):159.

SPSS White Paper, Data Mining: An Introduction, http://www.spss.com