

Bookbinders Case-Study

About 50,000 new titles, including new editions are published in the US each year, giving rise to a \$20 billion book publishing industry (in 1994). About 10% of the books are sold through mail order.

Book retailing in the 1970s was characterized by growth of chain bookstore operations in concert with the development of shopping malls. Traffic in bookstores in the 1980s was enhanced by the spread of discounting. In 1990s, the superstore concept of book retailing was responsible for the double-digit growth of the book industry. Generally situated near large shopping centers, superstores maintain large inventories of anywhere from 30,000 to 80,000 titles. Superstores are putting intense competitive pressure on book clubs, mail-order firms and retail outlets. Recently, online superstores such as Amazon have emerged, carrying 1-2.5million titles and further intensifying the pressure on book clubs and mail order firms. In response to these pressures, book clubs are starting to look at alternative business models that will make them more responsive to their customer's preferences.

Historically, book clubs offered their readers continuity and negative option programs that were based on an extended contractual relationship between the club and its subscribers. In a continuity program, popular in such genres as children's book, a reader signs up for an offer of several books for a few dollars each (plus shipping and handling on each book) and agrees to receive a shipment of one or two books each month thereafter. In a negative option program, subscribers get to choose which and how many additional books they will receive, but the default option is that the club's selection will be delivered to them each month. The club informs them of the monthly selection, and they must mark "no" on their order forms if they do not want to receive it. Some firms are now beginning to offer books on a positive-option basis, but only to selected segments of their customer lists that they deem receptive to specific offers.

Book clubs are also beginning to use database marketing techniques to work smarter rather than expand the coverage of their mailings. According to Doubleday present Marcus Wilhelm, "The database is the key to what we're doing... We have to understand what our customers want and be more flexible. I doubt book clubs can survive if they offer the same 16 offers, the same fulfillment to everybody"¹. Doubleday uses modeling techniques to look at more than 80 variables, including geography and the type of books customer's purchase, and selects three to five variables that are most influential predictors.

The BBB Club, Bookbinders Book Club, was established in 1986 for the purpose of selling specialty books through direct marketing. BBBC is strictly a distributor and does not publish any of the books it sells. In anticipation of using database marketing, BBBC made a strategic decision right from the start to build and maintain a detailed database about its members containing all relevant information about them. Readers fill out an insert and return it to BBBC which then enters the data into the database. The company currently has a database of 500,000 readers and sends out a mailing about once a month. ‘

BBBC is exploring whether to use predictive modeling approaches to improve the efficacy of its direct mail program. For a recent mailing, the company selected 20,000 customers in Pennsylvania, New York, and Ohio from its database and included with their regular mailing a specially produced brochure for the book *The Art History of Florence*. This resulted in a 9.03% response rate (1806 orders) for the purchase of the book. BBBC then developed a database to calibrate a response model to identify the factors that influences these purchases.

¹ 1: DM News, May 23, 1994.

For this case analysis², we will use a subset of the database available to BBBC. It consists of data for 400 customers who purchased the book and 1200 customers who did not, thereby over-representing the response group. The dependent variable for the analysis is Choice – purchase or no purchase of the book. BBBC also selected several independent variables that it thought might explain the observed choice behavior. The descriptions of the variables are:

Choice: Whether the customer purchased the The Art History of Florence. 1 corresponds to a purchase and 0 corresponds to a nonpurchase.

Gender: 0 = Female and 1 = Male.

Amount_purchased: Total money spent on BBBC books.

Frequency: Total number of purchases in the chosen period (used as a proxy for frequency.)

Last_purchase (recency of purchase): Months since last purchase.

First_purchase: Months since first purchase.

P_Child: Number of children's books purchased.

P_Youth: Number of youth books purchased.

P_Cook: Number of cookbooks purchased.

P_DIY: Number of do-it-yourself books purchased.

P_Art: Number of art books purchased.

BBBC is evaluating three different modeling methods: linear regression, a logit model, and support vector machines. While they want to isolate the factors that most influenced customers to buy the book, they also want to develop a highly accurate model. Note that linear regression is not the appropriate, so you may get poor results. Comment on this and why it is inappropriate.

- *Summarize the results of your analysis for the three models. The training, testing, and prediction data can be found on Blackboard.*
- *Interpret the results of the models. In particular, for models the influential covariates and their coefficients, provide insights.*
- *BBBC is considering a similar mail campaign in the Midwest where it has data for 50,000 customers. Such mailings typically promote several books. The allocated cost of the mailing is \$0.65/addressee (including postage) for the art book, and the book costs \$15 to purchase and mail. The company allocates overhead to each book at 45% of cost. The selling price of the book is \$31.95. Based on the model, which customers should Bookbinders target? How much more profit would you expect the company to generate using these models as compare to sending the mail offer to the entire list.*
- *Please also summarize the advantages and disadvantages of the three models, as you experienced in the modeling exercise. Should the company develop expertise in either (or all) of these methods to develop in-house capability to evaluate its direct mail campaigns.*
- *How would you simplify and automate your recommended method(s) for future modeling efforts at the company?*

² This case and data were developed by Professors Nissan Levin and Jacob Zahavi at Tel Aviv University. The materials have been adapted here and further discussed in the book Marketing Engineering.