Description and Documentation for the MultiChannel Gift Company Dataset
Version 1.0

By
Richard J. Courtheoux
President
Marketing Analysis Applications, Inc.
rcourtheoux@maa-online.com

A multichannel company with sales of several hundred million dollars per year has provided the Direct Marketing Educational Foundation with a dataset that can be used for classroom teaching. The company has a network of retail stores, a well-established traditional catalog channel and a website. Its brand is very well known nationally and it has a strong positive reputation along with very good long term customer loyalty. The core of the company's business consists of food products which are often purchased as gifts during the Christmas season. Their marketing challenges include:

➢ Customer relationship management, including contact management, across multiple channels.

➢ A very competitive marketplace for Christmas season gifts

➢ Generating business during the non-Christmas months of the year

The DMEF dataset consists of 4 files that have been extracted from the company's marketing database and that can be linked based on customer ID numbers. The files are:

➢ 100,051 customer records that have a customer ID, buying activity summarized by channel for 8 seasons, channel summary Recency/Frequency/Monetary information prior to the 8 seasons, gift recipients counts by season, catalog and e-mail circulation counts by season for 6 seasons, first order characteristics, overlay demographic and psychographic data, estimated distance to a retail store and the sectional center code (i.e., the first 3 digits of the customer's ZIPCode).

➢ 3,389,329 marketing contact records. These include both catalog and e-mail contacts with data fields containing the customer ID, the contact date and the type of contact. The records cover the years 2005 – 2007.

➢ 241,366 order or store trip records. Each of these records has a customer ID, an order number, date, channel and payment method. The records cover the years 2001 - 2007.

➢ 618,661 line item records. These records have a customer ID, an order number, date, price, a gift flag and (if applicable) the ID of the gift recipient. These records also cover the period from 2001 - 2007.

More detailed information about the records is provided below.

These files also provide the opportunity to illustrate for students the process of going from marketing database files that are maintained at a detailed, transactional level to analytical files containing many derived or summarized variables.

The company which provided the raw data for this dataset wishes to remain anonymous. In order to respect that wish some adjustments to the raw data have been made to make it less likely that the company can be successfully identified based on the dataset; also different sampling rates based on customer characteristics were applied to select customers for this teaching sample. If someone were to make a lucky guess about the identity of the company, the basic business ratios in this teaching dataset would not accurately match the real numbers which characterize this company's business. The use of these adjustment and sampling procedures should not reduce the utility of the dataset as a teaching resource for student analytical assignments.

Detailed Data Description

All 4 files are provided in CSV format. The first field in all 4 files is a customer ID which can be used to link the files together; all 4 files are sorted by customer ID.

Customer File: DMEFExtractSummaryV01.csv

Cust_ID is the customer ID. The file is sorted on this variable in ascending order.

The purchasing behavior variables have a naming scheme composed of three components:

- ➤ Channel – options are Ret, Int and Cat
- ➤ Season – these usually begin with either F (fall) or S (spring) and are followed by the last 2 digits of the year (e.g., '07' for the year 2007). Summaries for the period prior to 2004 have the designation Pre04.
- ➤ Measure – options are Dollars, GDollars, NGDollars, Orders, Trips and Lines. Retail sales are not allocated to gift or non-gift so they are just "Dollars" whereas the Internet and catalog sales are reported separately for gift and non-gift. In the Pre04 variables there is also a Recency measure which is the recency of last purchase measured in number of seasons prior to 2004 (e.g., a value 2 indicates that the most recent purchase prior to 2004 was in Spring 2003); a value of 99 indicates that the customer had not purchased in the channel prior to 2004.

So, for example, the variable IntF06GDollars has a customer's Fall 2006 Internet gift sales dollar total.

Variables FirstYYMM, FirstChannel, and FirstDollar give characteristics of the first purchase found in the purchase data.

AcqDate is the date (years and months only) that the company originally added the customer to its database. In some cases the company originally acquired the customer prior to the earliest retained transactional data (i.e., prior to 2001).

StoreDist gives the estimated distance from the customer's address to the nearest company store in miles. SCF_Code gives the first three digits of the customer's ZIPCode.

Demographic and psychographic variables were obtained through an overlay with a data vendor company. Not all customer records could be matched in the vendor company's database, so many of these fields are blank.

- ➤ 'Y' and 'N' codes are used to indicate whether a customer has an interest in a particular topic. These codes pertain to variables Travel, CurrAff, CurrEv, Wines, FineArts,

Exercise, SelfHelp, Collect, Needle, Sewing, DogOwner, CarOwner, Cooking, Pets, Fashion, Camping, Hunting and Boating.

➢ 'Y' and 'N' codes are also used to indicate presence of children of various age ranges in a household. These variables are Child0_2, Child3_5, Child6_11, Child12_16 and Child17_18.

➢ 'Y' and 'N' codes for variable Email indicate whether or not the company has permission to send e-mails to the customer.

➢ The HomeValue variable simply gives the estimated value of a customer's home in dollars.

➢ Specific coding schemes are used for the AgeCode, IncCode, HomeCode, Dwelling, LengthRes and OccupCd variables. Reference tables for these codes are in separate tabs of the spreadsheet DMEF Demo Codes Reference which is provided along with this document.

Contact File: DMEFExtractContactsV01.csv

The variables in this file are:

➢ Cust_ID is the customer ID. The file is sorted on this variable in ascending order.

➢ ContactDate is the date (YYYYMMDD) the contact was made

➢ ContactType indicates whether the contact was a catalog ('C') or e-mail ('E')

Order File: DMEFExtractOrdersV01.csv

The variables in this file are:

➢ Cust_ID is the customer ID. The file is sorted on this variable in ascending order.

➢ OrderNum is the order number. It is used to link order and line item records.

➢ OrderDate is the date (YYYYMMDD) that the order was placed

➢ OrderMethod indicates the method (channel) used for purchasing. Possible values are 'ST' (store), 'I' (Internet), 'P' (phone) and 'M' (mail). Phone and mail are the two methods by which catalog channel orders are placed.

➢ PaymentType gives the form of payment used. Possible values are 'BC' (bankcard), 'CA' (cash), 'CK' (check), 'GC' (gift certificate), 'HA' (house account), 'NV' and 'PC'. 'NV' and 'PC' codes were not documented by the company. Virtually all payments are bankcard, cash or check.

Line Item File: DMEFExtractLinesV01.csv

The variables in this file are:

➢ Cust_ID is the customer ID. The file is sorted on this variable in ascending order.

➢ OrderNum is the order number. It is used to link order and line item records.

➢ OrderDate is the date (YYYYMMDD) that the order was placed.

➢ LineDollars is the selling price of the line item in dollars.

- Gift indicates whether the item was given as a gift. For Internet and catalog orders this field is either 'Y' (yes) or 'N' (no – self purchase). For retail channel purchases this field is blank as the company systems do not record whether or not the item is a gift.

- RecipNum is the ID number of a gift recipient. It can be used to see if a recipient receives multiple gifts over time.