ABSTRACT

Many corporations and not-for-profits have old data in their archives which can yield significant insight into marketing decisions they face today. This data can be instrumental in estimating customer lifetime value (LTV). Knowing your customers’ LTV can unleash the full power of database marketing. SAS can be used to provide LTV estimates for customers/donors. Lifetime value can be calculated for various segments of customer/donor base allowing for comparison of various acquisition and retention programs.

INTRODUCTION

This paper describes how child sponsorship works at Compassion International, explains briefly the theory behind lifetime value analysis, describes the value and use of lifetime value to marketing, and finally, shows how to use SAS to read a file of transactions and calculate estimates of a donor’s lifetime value.

The following application is useful to either not-for-profit or for-profit organizations seeking to develop a relationship with their donors or customers. It will provide a way to know from actual historic data which acquisition methods work better.

CHILD SPONSORSHIP AT COMPASSION

Compassion International is a child development organization working with partner churches and schools in 21 countries to provide education, medical care, and Christian training to needy children. The method of assistance is to link each child to a donor (usually called a sponsor) who provides $24 a month to Compassion for the child. The sponsor is also encouraged to pray for, exchange letters with, and even visit the child. This sponsorship can be a relationship of very long standing. Some of our sponsors have been assisting children for decades.

The interesting question for Compassion’s marketing department is ‘which marketing activities attract the most loyal sponsors?’

LIFETIME VALUE ANALYSIS

The proliferation of computers and disk space has given rise to a new type of marketing: database marketing. This new discipline focuses on tracking the history of interactions between company and consumer. This history is made up of the stream of communications going to the customer and the sequence of purchases coming from the customer.

A major assumption in marketing is that there is a cause and effect relationship between advertising and sales. The marketing database tracks every stimuli and every result for extended periods of time. With this history it is possible to calculate the value of each customer to the company. It is also possible to discover what marketing and retention activities influence lifetime values.

A key premise of database marketing is that it is easier to keep a current customer than to find a new one. It is easier to win greater loyalty from existing customers than to break through the initial obstacles faced to acquire new customers. Database marketing can tell the marketer whether or not a loyalty program is working or not.

But all of this depends of having a good sense of the lifetime value of a customer. Lifetime value in this context is the dollar amount that the customer will give over the course of their relationship with the organization. A baseline must be calculated and variations from that baseline form the basis of comparing various marketing strategies.

As an example of lifetime value calculation, consider donors to Compassion’s child sponsorship program. To find the average lifetime value (or 1 year value in this example) of a donor from a particular media source, identify all the sponsorships attributed that source... say 100 from magazine advertisements. How much did the sponsors of those children give during the 12 months following their initial gift?... say $250,000. Take this total and divide by 100 ($250,000/100=$250). The result is the average one year value of sponsors acquired from magazine ads.

Dividing by 100 rather than by the number of sponsors still remaining after 12 months means that the one year value incorporates the impact of attrition. Attrition, size of gift, and frequency of gift are all included in the one year value.

Another key point in database marketing is that the unit of analysis is not the individual customer but rather groups of customers who share relevant characteristics. Predicting the behavior of a single customer is very difficult but predicting the behavior of a group of 1000 customers when you know certain things about them is much more reliable.

ESTIMATING LIFETIME VALUE

Historical Data

Hanging around in most company’s computer tape libraries are tapes of transaction histories containing customer name, ID number, date, amount paid, accounting codes and other codes. SAS can read these files no matter what form they may be in. The following code was used to read my organization’s files when they were flat files on magnetic tape.
Posters

Later this data was summarized and made available through a data warehouse. Using SAS/ACCESS and the following code the same information in a slightly summarized form was read.

**Figure 2 Read Data Using ODBC**

```
libname cidata 'd:\cidata\odbc';
proc sql stimer-feedback noprint number flow=78;
   connect to odbc as mycon (dsn="CC Data Warehouse" uid='~DW'
log);
   execute tuse ccdwl by mycon;
create table raw84 as
   select CONSTITU as connum format=8.,
          CALENARIO as year format=4,<
          CALENDAR as mon format=2.,
          FUND_ID as fund format=6.,
          CON_MONTH as amount format=10.2,
          CON_MONO as count format=8.
   from connection to mycon
   (select constituent_number,
              calendaryear,
              calendar_month,
              fund_id,
              con_month_fund_item_total,
              con_month_fund_item_count
   from cc_con_month_givng_by_fund_1984
   where not con_month_fund_item_total = 0);
```

These two sets of code show how to read data from the 1984 file. Twelve files (1984 - 1995) with hundreds of thousands of records in each were read and stored in SAS datasets in preparation for selection and analysis.

Notice that all transactions with a non-zero amount were read. This is because refunds and adjustments happen which can change the actual value of donations given.

**Selection of Analysis Group**

The data above represented several hundred thousand sponsors and millions of donation transactions. Taking a subset speeds things up at the data processing and analysis stage. To begin with, all the sponsors who began sponsoring a child in 1989 were subsetted. Additional information needed for this group was the marketing channels through which the sponsors were acquired, i.e., TV, magazine ads, concerts, volunteers, etc.

The following output from SAS shows that we attracted 19,356 new sponsors in 1989. Their sources are distributed as shown in the output of a PROC FREQ.

**Table 1 Distribution of Donor Sources**

<table>
<thead>
<tr>
<th>MEDIA</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Frequency</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional</td>
<td>13</td>
<td>0.1</td>
<td>13</td>
<td>0.1</td>
</tr>
<tr>
<td>Alumni Insert</td>
<td>2032</td>
<td>10.5</td>
<td>2045</td>
<td>10.6</td>
</tr>
<tr>
<td>Audio Visual</td>
<td>241</td>
<td>1.2</td>
<td>242</td>
<td>1.2</td>
</tr>
<tr>
<td>Carnival</td>
<td>2479</td>
<td>12.8</td>
<td>4765</td>
<td>24.6</td>
</tr>
<tr>
<td>Convention</td>
<td>13</td>
<td>0.1</td>
<td>4778</td>
<td>24.7</td>
</tr>
<tr>
<td>Card Deck</td>
<td>449</td>
<td>2.3</td>
<td>5227</td>
<td>27.0</td>
</tr>
<tr>
<td>Direct Mail</td>
<td>58</td>
<td>0.3</td>
<td>5285</td>
<td>27.3</td>
</tr>
<tr>
<td>Festival</td>
<td>265</td>
<td>1.4</td>
<td>5550</td>
<td>28.7</td>
</tr>
<tr>
<td>General</td>
<td>594</td>
<td>2.8</td>
<td>7799</td>
<td>40.2</td>
</tr>
<tr>
<td>Magazine</td>
<td>1407</td>
<td>7.3</td>
<td>9195</td>
<td>47.5</td>
</tr>
<tr>
<td>Music</td>
<td>1072</td>
<td>5.5</td>
<td>10268</td>
<td>53.0</td>
</tr>
<tr>
<td>Referral</td>
<td>94</td>
<td>0.5</td>
<td>10275</td>
<td>53.3</td>
</tr>
<tr>
<td>Resource</td>
<td>43</td>
<td>0.2</td>
<td>10.708</td>
<td>58.3</td>
</tr>
<tr>
<td>Speaker</td>
<td>125</td>
<td>0.9</td>
<td>11.313</td>
<td>61.0</td>
</tr>
<tr>
<td>Television</td>
<td>527</td>
<td>2.8</td>
<td>19.248</td>
<td>94.3</td>
</tr>
<tr>
<td>Other</td>
<td>64</td>
<td>0.3</td>
<td>7789</td>
<td>40.2</td>
</tr>
<tr>
<td>TV</td>
<td>19,356</td>
<td>100.0</td>
<td>19,356</td>
<td>100.0</td>
</tr>
</tbody>
</table>

TV was the biggest source for new sponsors with 6,365. Some sources brought in only a handful of sponsors.

**Data Preparation**

The next step is to collect all the historical giving transactions of the sponsors who began in 1989. The transactions resided on several tapes and numbered in the millions.

Plowing through all the giving transactions proved to be too time consuming so transaction records were excluded based on their sponsor identification number. Records with unneeded sponsor identification numbers were excluded before trying to sort and merge giving data from the 19,356 sponsors in my target group.

**Figure 3 Summarize Giving Data**

```
data totgive;
   set svdata.givings9 svdata.giving90
          svdata.giving91 svdata.giving92
          svdata.giving93 svdata.giving94
          svdata.giving95;
   if Iconnum ge 180000 and connum le 418000 or
      (connum ge 921000 and connum le 923000);
   proc sort data=totgive by connum;
   proc sort data=cdata class59 by connum;
   data temp89;
   merge cdata.class89{in=inclass}
      totgive{iID='intemp'};
   if inclas$ and intemp;
   months=mdy(month,15,year);+1;
   keep connum media strtdate amount;
   proc sort data=temp89; by connum;
   data cdata.class59; by connum;
   data temp89; merge cdata.class59{in=inclass}
      temp89{iID='intemp'};
   by connum;
   if inclas$ and intemp;
   months=mdy(month,15,year);+1;
   keep connum media strtdate amount;
   proc sort data=temp89; by connum;
   proc means noprint;
   by connum months;
   md media strtdate; var amount;
   output out=cdata.f89 mean(amount)=amount;
run;
```
The PROC MEANS statement above collapses the information about all the constituents into a summary by sponsor and by month of sponsorship. The analysis variables (media and startdate) are carried along for use in future procedures.

Analysis
The goal in this LTV analysis is to identify which acquisition method produces more loyal sponsors? The following SAS code distills the previously created dataset into a table which sheds light on that question.

Figure 4 Analyze Data

```
libname cidata 'd:\cidaea\sugi';

proc format;
  value $fmt 'AC' = 'Additional';
  'AT' = 'Album Insert';
  'AV' = 'Audio Visual';
  'Bk' = 'Book Insert';
  'CGV' = 'Convention';
  'CARD' = 'Card Deck';
  'DM' = 'Direct Mail';
  'FN' = 'Festival';
  'MAG' = 'Magazine';
  'NP' = 'Newspaper';
  'OTH' = 'Other';
  'PA' = 'Print Ad';
  'PF' = 'Public Info';
  'RA' = 'Radio';
  'RCT' = 'Receipt';
  'REF' = 'Referral';
  'RP' = 'Resource';
  'SPK' = 'Speaker';
  'TV' = 'Television';
  'UN' = 'Unknown';
  'VOL' = 'Volunteer';
run;

%************ get denominators */
proc sort data=cidata.f89;
  by month;
run;

/* get one record per constituent */
set cidata.f89;
by month;
if first.month then delete;
keep media months amount cum count;

format media $stmt.;
count = 13;
when ('AC') count= 13;
when ('AT') count= 2012;
when ('AV') count= 241;
when ('CG') count= 2479;
when ('DVA') count= 13;
when ('CRD') count= 449;
```

Part of the output of the above code is shown in Table 2. This output shows that sponsors with a 'Album Insert' source have an average one year value (12 month value) of $252. The average two year value is $470 and the average five year value is $1020. This is higher than the five year values of Audio Visuals ($962) and Concerts ($879).

### Table 2 LTV by Media Sources

<table>
<thead>
<tr>
<th>Media Source</th>
<th>Months of Sponsorship</th>
<th>Starting Sponsors</th>
<th>Cumulative Total</th>
<th>Long Term Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Album Insert</td>
<td>12</td>
<td>2,032</td>
<td>$515,326</td>
<td>$252</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>2,032</td>
<td>$954,680</td>
<td>$470</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>2,032</td>
<td>$1,350,242</td>
<td>$669</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>2,032</td>
<td>$2,730,829</td>
<td>$952</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>2,032</td>
<td>$2,073,276</td>
<td>$1,020</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Media Source</th>
<th>Months of Sponsorship</th>
<th>Starting Sponsors</th>
<th>Cumulative Total</th>
<th>Long Term Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio Visual</td>
<td>12</td>
<td>241</td>
<td>$58,089</td>
<td>$241</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>241</td>
<td>$107,986</td>
<td>$448</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>241</td>
<td>$155,210</td>
<td>$644</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>241</td>
<td>$2,009,674</td>
<td>$812</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>241</td>
<td>$2,071,790</td>
<td>$982</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Media Source</th>
<th>Months of Sponsorship</th>
<th>Starting Sponsors</th>
<th>Cumulative Total</th>
<th>Long Term Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concert</td>
<td>12</td>
<td>2,479</td>
<td>$563,771</td>
<td>$219</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>2,479</td>
<td>$1,004,598</td>
<td>$405</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>2,479</td>
<td>$1,409,319</td>
<td>$569</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>2,479</td>
<td>$1,893,326</td>
<td>$727</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>2,479</td>
<td>$2,178,449</td>
<td>$879</td>
</tr>
</tbody>
</table>
A few more simple SAS statements yield the five year value by media source in descending order.

Table 3 Five year LTV by Media Source

<table>
<thead>
<tr>
<th>OBS</th>
<th>MEDIA</th>
<th>AMOUNT</th>
<th>COUNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Leaflet</td>
<td>$1,898</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Additional</td>
<td>$1,425</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>Convention</td>
<td>$1,237</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>Magazine</td>
<td>$1,130</td>
<td>2,174</td>
</tr>
<tr>
<td>5</td>
<td>Speaker</td>
<td>$1,057</td>
<td>175</td>
</tr>
<tr>
<td>6</td>
<td>Radio</td>
<td>$1,043</td>
<td>1,972</td>
</tr>
<tr>
<td>7</td>
<td>Other</td>
<td>$1,041</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>Direct Mail</td>
<td>$1,036</td>
<td>58</td>
</tr>
<tr>
<td>9</td>
<td>Referral</td>
<td>$1,036</td>
<td>549</td>
</tr>
<tr>
<td>10</td>
<td>Card Deck</td>
<td>$1,031</td>
<td>449</td>
</tr>
<tr>
<td>11</td>
<td>Unknown</td>
<td>$1,025</td>
<td>545</td>
</tr>
<tr>
<td>12</td>
<td>Album Insert</td>
<td>$1,020</td>
<td>2,312</td>
</tr>
<tr>
<td>13</td>
<td>Volunteer</td>
<td>$1,014</td>
<td>563</td>
</tr>
<tr>
<td>14</td>
<td>Print Ad</td>
<td>$999</td>
<td>1,407</td>
</tr>
<tr>
<td>15</td>
<td>Audio Visual</td>
<td>$992</td>
<td>261</td>
</tr>
<tr>
<td>16</td>
<td>Receipt</td>
<td>$988</td>
<td>50</td>
</tr>
<tr>
<td>17</td>
<td>Festival</td>
<td>$963</td>
<td>245</td>
</tr>
<tr>
<td>18</td>
<td>Resource</td>
<td>$895</td>
<td>433</td>
</tr>
<tr>
<td>19</td>
<td>Concert</td>
<td>$879</td>
<td>2,479</td>
</tr>
<tr>
<td>20</td>
<td>Television</td>
<td>$760</td>
<td>6,365</td>
</tr>
</tbody>
</table>

From the above report it is easy to see that among the major media sources, the most loyal sponsors are acquired through magazines with a five year value of $1,898. TV was at the other end with a five year value of $760. We speculate that the TV sponsor's decision to sponsor a child was a little more emotion based and the magazine sponsor's was more rational. It could also be that people who read magazines make more thoughtful, long-lived decisions than those who watch TV.

With this information the marketing director is better able to determine how much he should invest to acquire new sponsors through various channels.

OTHER ISSUES

There are several other issues which need to be addressed to arrive at an accurate lifetime value. First, since income from a donor takes place over months and years, an adjustment for net present value must be made. Since we often use this information to compare the relative LTV of various sources we do not need exact figures but a simple ranking and magnitude of the differences.

Second, the researcher never gets to see the actual lifetime value of a donor until they die! Hence, projections must be made to estimate what they will give.

Third, when comparing dollars from different periods of time, inflation becomes a problem. Amounts from transaction files are nominal dollars not real dollars. This is very important when comparing donors acquired in different years. $300 raised during 1955 is not equal to $300 raised in 1995. One way to approach this problem is to deflate all dollar figures to a standard dollar value (1995 dollars) at the beginning of the analysis.

Finally, this paper only considers the income stream from a donor. The cost of acquiring a donor and the cost of maintaining the donor must be applied to arrive at a net lifetime value for various donor types.

CONCLUSION

This paper has shown one way to derive estimates of donor lifetime value from old accounting transaction files. Comparison of lifetime value by acquisition sources and retention programs can provide valuable information for marketing directors and donor satisfaction directors.

ACKNOWLEDGMENTS

Terry Spence provided significant editorial help for this article.

SAS and SAS/ACCESS are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

Thayer Allison
Compassion International
PO Box 7000
Colorado Springs, CO 80933
719-594-9900x126
tallison@ci.org