Nigel Magson

Nigel is the MD and co-founder of Talking Numbers with Target Direct Marketing. The company has developed a broad range of award winning marketing, statistical and information solutions, which are applied on behalf of organisations across many market sectors. Nigel speaks regularly at industry-recognised conferences and is a published author and editor of fundraising, database and analysis articles and books.

Keywords: marketing, metrics, lifetime value, segmentation, IT

Metrics: The kaleidoscope of value measurement at customer, campaign, function or enterprise level

Nigel Magson

Received (in revised form): 1 June 2001

Abstract

Many different organisations have built databases which house marketing information about individual customers or organisations. Obtaining and understanding the value that these individuals have represented and continue to represent is key to ensuring that marketing activity is effectively directed. Many organisations still see the database as supporting direct marketing activity only, and fail to use its potential to link back to their business-planning processes. So while measures are calculated at a macro level, these are rarely converted to the end customer base.

This paper is based on practical experience in building and producing value analysis and metrics within different sectors, and aims to reveal a workable approach to creating different metrics.

Introduction

This paper explores a number of areas:

- defining and creating value metrics: the definition of different measures
- building measures: the practical applications and some of the pitfalls in building measures
- an IT methodology of building information suites that will allow metric creation for different applications within the organisation.

This is done in the context of organisations covering different market sectors, such as FMCG, holiday, retailing, utility, charitable and financial services sectors.

Metric definition

It is rare that a single metric alone is used when describing marketing activity. The days of response rates being quoted as an evaluation of a campaign, without the consideration of ROI (return on investment), appear long gone. While a combination of metrics appear to be used in most industries, they can vary at different organisational levels. Metrics have been developed to describe a particular aspect of the business and the description carries a particular perspective — for example,

Nigel Magson
Talking Numbers Data based
Solutions Ltd.,
Lorraine House,
Market Place,
Cirencester GL7 2NW, UK
Tel: +44 (0)1285 644220;
E-mail:
nmagson@talkingnumbers.com

recruitment metrics viewed by the new business function can ignore calculations about ongoing worth.

The adoption of CRM (customer relationship management) within many organisations would suggest that these organisations are now better placed to have an overview of how different metrics integrate and can be used to describe their business activities. Stone and Mason consider the organisational issues of metric development as part of the use of customer relationship management in a major UK insurance company.¹

When considering developing an existing metric or adding a new metric, the following questions can be posed:

- will/does the business accept it?
- is it practical to calculate and maintain?
- will it be worthwhile and comparable over time?
- at what organisational level is the metric current?
- does it fit with other metrics in use?

Given the combination of metrics that may be used, it is also worth considering:

- do these metrics reflect and describe the key factors that affect the business?
- are there gaps in the business which are not described?
- are the metrics common parlance across the business or for just one area?

It may be that a common measure in use within an industry can be developed to give more value. For instance, in the holiday sector, occupancy rate is a key measure of hotel performance. An organisation can extend and develop this measure to allow performance to be gauged by defined seasonality, by day of week and weekend bookings, and by rack and discounted rates. This would allow an organisation to identify quickly areas of strong and poor performance by outlet, and then target marketing activity to redress the shortfall.

The measure could be further extended to allow comparative performance either against other group hotels or outlets or industry published metrics. However, high occupancy alone does not necessarily relate to high profitability, so an associated measure linking to yield is necessary. McMahon-Beattie and Palmer note that the concept of yield management originated in the airline industry to describe yield per available seat mile.²

‘Simply put, YM (yield management) is the process of allocating the right type of capacity or inventory unit to the right kind of customer at the right price so as to maximise revenue or yield.’³

To extend a concept such as yield management to service industries depends on the interpretation of key metrics through a range of descriptive statistics which can be applied to the basic measure, such as means, modes, medians etc. McMahon-Beattie and Palmer⁴ explore the

price variation of different hotels using normalised standard deviations in terms of different targeted segments to show that greater segmentation is possible at higher price levels. It should be noted that means must be understood in the context of the spread of data and the number of standard deviations within the dataset. The use of *t*-tests will help to uncover whether there is a significant difference between two means.

Figure 1: Holiday company

Which level depends on data availability

At what level can the calculation be made?

One of the next decisions that will need to be made is at what level it is possible to create the calculation — market, business, segment or individual level. Figure 1 shows an example of a holiday company where different levels will require different metrics. This is illustrative of different levels that can be generated for an organisation.

The availability of data may be channel dependent. For instance, in the holiday sector, while direct sales and details are well recorded, sales via the trade (travel agents) are sometimes only partially available through post-holiday questionnaire data. Alternatively the type of data available may be different, eg transactional, attitudinal, lifestyle and demographic. Koslowsky explores the use of different data types through cluster analysis in creating attitudinal and behavioural segmentation models.⁵

Macro and micro metrics

Generally, creating macro and micro metrics at a business or segment level is dependent on the reporting and recording of information through source systems, such as general sales levels, perhaps at store level. The ability to 'push' information up and down these levels will be a function of the accounting and recording of information and the level of systems integration.

In the retail sector, credit card or loyalty card sales will allow transactional histories to be built, while many cash sales will be missed. So calculating total value at an individual level will not be possible, whereas an accurate accounting calculation can be made at higher levels, eg store. Table 1 and Figures 2 and 3 show a number of metrics generated at store level. Given the difficulty in some instances of driving the metric from the market level through to the individual level, metric calculations can only be at the individual level through sampling and research of key customer groups.

Table 1: Retail metrics at store level

	York	Bristol	Basingstoke	All branches
Mean purchases (number)	1.98	3.49	2.44	2.45
Mean lifetime value	95.72	218.72	104.48	156.26

Figure 2: Average value to date by branch

Figure 3: Distance between customers and city centres in miles

Linking market level and individual level data

Example macro to micro metrics — market and individual volumetric segmentation

A simple example below reveals how market level volumes and share of specific brands can be related to individual consumers and their 'share of wallet'. These metric calculations were developed for an alcoholic drinks manufacturer from an individual consumer-level database, which allows individual levels to be categorised initially by the overall amount of a particular drink they consume per annum. Initially each consumer is assigned to a segment, which is defined by their overall annual drink consumption for a particular category, eg cognac, whisky, rum, champagne.

- Light 1–2 bottles:
- Medium 3–6 bottles:
- Heavy 7–12 bottles:
- Superheavy 13–52 bottles:
- Superheavy plus 52+ bottles:

Each consumer is then categorised within each category as:

- loyal (ie consuming a high percentage of the client's brands)
- repertoire (ie consuming a mid percentage of the client's brands)
- competitor (ie consuming a low percentage of the client's brands).

Segments are then describable as 'heavy loyal' or 'heavy competitors'. A value is attributable to each consumer and each segment based on the net profit of each nine-litre case consumed. This information is built from transactional-level information loaded to a drinking habits table within the database system. The derived scores are held so that changes in behaviour can be understood, and change scores generated. An association can then be made between the marketing activity received by an individual and the change in their behaviour. Significance testing (sales

lift analysis) will be used to see whether this is different from a control group receiving no activity.

The profiles and motivations behind each segment were recently explored using the Claritas household-level PRIZM segmentation system. Different communication programmes have now been adopted for specific segments.

It is worth noting that in examples such as these there is an in-built bias to the data, as the emphasis is on consumers who are likely to be consuming the client's brands. However, this bias can be allowed for by relating these figures to available market-level consumption statistics. (Note that the issue of self-selecting populations is prevalent within internal segmentation and historic campaign audience selection.) Dupin proposes that using set theory for sampling provides a more accurate way of predicting rollout than conventional testing methodologies.⁶ The information allows the client to understand their overall share of the market by category when compared to market-level figures, and also to explore regional penetration differences.

Building measures: The practical applications and some of the pitfalls

This section explores some of the issues involved in forming metric calculations at an individual level for certain types of measures — in particular lifetime value calculations. There are a number of considerations when working with data, and this section picks up specific issues and obstacles that can be met when working with data elements. These issues are explored in more depth in relation to creating donor lifetime values.⁷

Customer-level metrics

Figure 4 is a non-exhaustive list of metrics that could be generated at a customer level. The examples could apply in this instance to the not-for-profit or mail-order sectors, although a change in nomenclature would make these measures applicable for different sectors.

In order to be able to understand the customer's behaviour and the effect of marketing over a period of time, these fields are recalculated periodically, and then change fields generated. By using 'snapshotting' techniques and creating suites of metrics held against each customer, it is possible to trade trends in movements and changes, relating these to marketing material sent. This helps to move away from the concept of single campaign evaluation to consider the cumulative effect of activity. Standard statistical measures should be used to analyse the change and determine its significance.

'Pushing down' information from segment to individual level

Information that is typically held at a campaign or segment level, perhaps outside of a main database, will need to be attributed to the individual. Often this type of information may be costs incurred in the marketing activity that need to be spread across the individuals within the segment.

However, in the case of campaign costs, this can become quite complex

Measure	Definition of measures to be created for each unique customer
Customer URN	Unique record number for each customer
Customer gender	Identified sex of customer
Customer type/segment	Types defined by organisation roles/existing segmentation
RFV	Recency frequency value calculation built from transactional history over defined period of time, held as separate elements
Response ratio	No of responses: no of communications $\pm$ weighting constant
Total gross income	Sum of cash (including VAT)
Total cash income	Sum of all cash transactions
Total DD/other income	Sum of all other payment methods
Total net income	Total gross income: (recruitment cost $\pm$ warm costs) $\div$ VAT
Total warm net income	Total gross income (excluding first transaction) $\div$ warm costs
Annual gross income	Sum of gross income by year
Annual net income	Sum of net income by year
Upgrade to regular subscription/purchase club date	Date of first regular subscription/purchase club payment
Highest transaction ever	Largest identified transaction
Lowest transaction	Smallest identified transaction
Highest five transactions ever	Record of five highest transactions
Modal transaction	Most common transaction
Median transaction	50th percentile
Mean transaction	Sum of transactions divided by number of transactions
Transaction type	Method of transaction
Transaction code	SKU or categorised code for each transaction
Date of last transaction	Most recent transaction
Date of first transaction	First recorded transaction
Amount of first transaction	Amount of first transaction
Amount of last transaction	Amount of last transaction
Recruitment method	Media/creative/offer used to recruit
Recruitment topic	Recruitment creative or proposition code
Cost of recruitment	Campaign cost divided by number of new responders
Smoothed cost of recruitment	Campaign cost as above smoothed to take mean recruitment cost per annum
Total warm communication cost	Sum of all warm communication costs
Total number of warm communications	Sum of all warm communications ever sent
Total number of transactional responses	Count of all response transactions ever sent
Total number of non-transactional responses	Count of all non-transaction responses ever received
Total warm communications cost	No of warm communications $\times$ pack constant $\pm$ (no of newsletters $\times$ newsletter constant) $\pm$ total fulfilment cost
Total number of newsletters received	Count of all newsletters sent
Total fulfilment cost	Total number of responses $\times$ fulfilment constant

DD = Direct debit

RFV = Recency frequency value

SKU = Stock-keeping unit

URN = Unique record number

VAT = Value-added tax

Figure 4: Example customer-level metrics for charity or catalogue retailer

if all the individuals who have received the campaign are not held on the underlying segment dataset — typically as media spills to segments that were not targeted, such as above-the-line media or even mailings or doordrops which are passed out of the area. There is a danger of attributing too much cost to a single individual and therefore adversely affecting their ‘value’. A realistic and common-sense approach has to be adopted.

Recruitment cost substitution

Perhaps one of the most important metrics needed by an organisation is the cost of recruiting a new customer. In the charity sector this has been

well explored,⁸ and also in the financial services sector.⁹ In the holiday sector, one measure for this is described as 'a cost per booking', in mail order as 'cost per response'. For the marketer, there is generally a large difference in the cost of getting repeat business and the cost of getting new customers.

The costs and perhaps the immediate income generated by a recruitment activity will be held at a campaign or segment level. The issue comes when trying to move these elements down to an individual level in order to develop individual-level lifetime calculations. This is because most recruitment campaigns will 'catch' both existing and prospective customers.

Recruitment is generally defined as the first known transaction with the organisation, and an assumption is made that this can only occur once. In order to evaluate the effectiveness of cold recruitment, and if allocating cold costs to a dataset, then responses received from existing customers should not carry the same cost per response as those from new customers.

Subsequent response to recruitment activity can be handled by a process called 'warm substitution', which is where a mean warm cost of response for that type of campaign is allocated in a relevant time frame. The main reason for warm substitution is so that a fair comparison can be made between two customers, for future activity, when one responded to an unsuccessful recruitment campaign. Warm substitution follows from the accounting used to calculate a recruitment or response cost (ie the overall campaign costs divided by the number of responses) rather than the variable communication spread over all recipients of an activity (eg house mailing), irrespective of whether they respond or not. An example is given below.

$$\text{Recruitment cost} \sim \frac{\text{Campaign costs } \pounds 10,000}{\text{Number of responses } 1,000} \sim \pounds 10$$

$$\text{Communications cost} \sim \frac{\text{Campaign costs } \pounds 10,000}{\text{Number of recipients } 20,000} \sim \pounds 0.50$$

Generally a recruitment cost would be written back to all new customers, but where an existing customer had responded to a recruitment appeal, a 'warm substitution' cost would be used. This would allow a representative cost to be attributed without penalising a customer for responding to acquisition activity (as this invariably incurs a higher cost).

Lifting up information from individual to segment level

Sometimes the converse of the above is necessary in creating a meaningful dataset.

In the retail sector, product coding is typically driven through the use of SKU numbers (stock-keeping units), with thousands of product lines. Individual metrics on specific product lines are typically generated through the underlying retail system. For analysis and generation of higher-level metrics, the SKU numbers are categorised into meaningful 'departments' or 'product groups' so that there is a need for sets of look-up tables which recode individual SKUs to a new category.

A further common example of lifting up information occurs with appeal or marketing activity codes. For instance, individual newspaper title codes may be too spread to be meaningful, and therefore categorisation to broadsheet or tabloid, daily or weekend will give broader segments to allow analysis to be tackled.

Treatment of cost and income over time

An accountant's view of money over time would use NPVs of historic transactions, and discounted cash flows for future revenue streams. When considering lengthy historic transactional files, applying some form of NPV weighting will increase the historic worth of the file.

Marketers may be more comfortable with the raw unweighted view of currency, as this may tally back to sales and campaign reports from the period. When trying to forecast future revenue streams a similar process can take place, calculations being based on current or discounted margins.

If dealing with multiple currencies in creating global calculations, conversion to a single common currency can be complex. Assuming this is the pound or the US dollar then transaction conversions could be made in a number of ways. This may be by using the currency rate on the day of the transaction, or alternatively on the date when the goods' costs were charged. From a comparative point of view using an exchange rate average by year is simpler, but may hide historic fluctuations in currency rates.

Indexing is a technique which allows comparisons to be made by comparing to a common base. Figure 5 shows life stage segments of multiple rather than one-off purchasers indexed over time against mean purchase values for a retailer. The indexing helps to reveal meaningful differences between the different life stages. Further information was uncovered when affluence and life stage were combined in the segment description.

Figure 5: The differentiation of life stage on purchase value

It has already been noted that the length of time covered by a dataset, and in driving comparisons, will have an impact on financial comparisons. Length of time on a dataset will also start to allow calculations of typical 'lifetime' or repurchase pattern frequencies and other useful metrics. Clearly, ensuring data integrity is important.

Selecting the best technique and other issues

The practicalities of building the information at the defined levels will no doubt reveal situations where information and assumptions will need to be made in conjunction with the organisation's marketing and information departments.

Since large historical datasets can be a function of several systems over time, they almost invariably contain 'imperfections'. Experience has shown that once data are audited and examined in depth there will be a need to edit or adapt the dataset. Otherwise for a combination of reasons — such as an absence of data or recording multiple transactions to a sink or bucket code — biased and inaccurate results will be produced. For instance, it is typically expected that some cost substitution may be needed with some campaigns.

Once built, metrics can be combined and descriptive statistics used to reveal insights into the marketing approach. Figure 6 shows a mail responsiveness score cross-tabulated against net value for a charity. It reveals that the most responsive donors are not necessarily those with the highest net value. The categorical nature of marketing data lends itself to techniques such as CHAID. Madgison explores the use of more advanced scoring to determine key variables that influence response for the Amoco Motor Club.¹⁰ He reviews different statistical techniques and their evolution for marketing purposes.¹¹

There is a wide variety of software tools that allow more complex statistical analysis, such as SPSS or SAS.

Figure 6: Number of donors by response ratio, including average net LTV for each RR band

An approach to building information suites that will allow different applications within the organisation

The first section has dealt with some of the key questions to ask to help define the measures needed; the second has dealt with issues to be aware of when creating these measures; this section deals with the concept of an 'analysis information suite'. There are different approaches to the IT creation of value metrics. RDBMS (relational database management systems) are not necessarily best placed to work with analysis measures, as normalised design principles mean that data items should not be replicated. These systems are typically optimised for large transactional systems and speed of querying and selections, normally using a form of SQL. The emergence of OLAP (online analytical processing) tools has allowed greater manipulation and the ability to mine database structures, although these tools are limited by the fact that summaries and aggregation of data have to be predefined. They can nevertheless be powerful tools for regular reporting and measurement rather than investigative analysis.

An alternative practical IT delivery of a receptacle to hold and work with measures could be described as a 'datamart'.¹² While few organisations have all their systems integrated to one central 'warehouse', many have the capability to develop departmental or functional-level systems with limited transferring of key data between departments. Duffill explores the development of a datamart for metric creation for the charity Shelter.¹³ Micheaux and Gayet discuss the content of such systems, and the types of data that can be used to populate them.¹⁴

The generation of a range of meaningful marketing metrics often needs a combination of elements from different systems. If metrics are needed on a regular basis, then an interim system is needed to collect and hold the raw data, as well as perform the calculations needed. This is a form of datamart and can be turned into an information suite that can be worked on with different toolsets. Figure 7 shows a possible information flow into a datamart system. Figure 8 shows that the information suite could contain a number of data types, which could be calculated from the combined pool of data imported to the system.

This approach to creating a 'customer datamart' has a number of benefits:

- predictable response times for queries and reports
- no effect on operational systems as it is offline
- data can be 'scrubbed' and made easier to understand
- other information can be easily incorporated, such as communication costs, analysis scores and external data (eg PRIZM)
- a group of people can work with the same agreed version of the truth
- more flexible and easier to change than a RDBMS
- the creation and therefore knowledge of its functions and capabilities are understood.

The potential downsides of this approach are:

Core database information eg Socio-demographics, transactions and communications, customer satisfaction	Analysis scores eg RFV, LTV	Historic scores not held on database Market-level info — volumetrics, brand usage
	Segmentation and marketing codes	
	Additional information eg PRIZM, attitudes	

Figure 8: Types of information held on datamart

- need for regular data refreshes to keep the system up to date, which require scheduling and load code
- data cycle; the frequency of refresh can be reviewed after a period of live operation
- possibility of having systems ‘out of sync’, although the analysis and reporting server can be configured so that it automatically takes a refresh from other systems at predetermined times
- it is an additional system to maintain and resource; this process is best managed by a central IT function.

Create combinations of metrics

Forming a central repository of data out of which to build the required metrics can have its own complexities in design and the interfaces it must make to other systems. However, this approach, if performed consistently, will create a reliable resource for marketing planning.

Summary

This paper has outlined an approach to developing metrics at different levels within an organisation. The paper takes the view that a combination of metrics will be required for describing the performance of marketing within the organisation. Experience would indicate that the metrics that will become common parlance are those which are understandable to the key user group, easy to calculate and maintain, and give an accessible and meaningful view of the business area.

Organisationally the acceptance of the metric will depend on key ownership of the definition and concept at specific organisational levels. Many organisations could benefit from reviewing and assessing the metrics that are used and conducting a 'gap analysis' of business areas that are poorly described.

References

1. Stone, M. and Mason, V. (2000) 'CRM in intermediate financial services markets', *Journal of Financial Services Marketing*, Vol. 5, No. 1, pp. 61–72.
2. McMahon-Beattie, U. and Palmer, A. (2000) 'One for all or all for one: A comparison of everyday low pricing and yield management strategies in the hotel industry', *Journal of Targeting, Measurement and Analysis for Marketing*, Vol. 8, No. 3, pp. 249–258.
3. Kimes, S. (1997) 'Yield management: An overview', in Yeoman, I. and Ingold, A. (Eds) *Yield Management, Strategies for the Service Industries*, Cassell, London, pp. 3–11.
4. McMahon-Beattie and Palmer, ref. 2 above.
5. Koslowsky, S. (2000) 'The use of behaviours and attitudes in market segmentation', *International Journal of Customer Relationship Management*, Vol. 3, No. 2, pp. 119–129.
6. Dupin, D. (2000) 'Direct response multi-list market testing', *Journal of Interactive Marketing*, Vol. 2, No. 2, pp. 120–128.
7. Magson, N. (1999) 'Donors: How much do they give in a lifetime?', *Journal of Nonprofit and Voluntary Sector Marketing*, Vol. 4, No. 1, pp. 11–26.
8. Aldrich, T. M. (2000) 'How much are new donors worth? Making donor recruitment investment decisions based on lifetime value analysis', *Journal of Nonprofit and Voluntary Sector Marketing*, Vol. 5, No. 1, pp. 81–90.
9. Bradley, T. (1999) 'RecruitmentPlanner — A methodology for optimising recruitment spend on higher value customers', *Journal of Financial Services Marketing*, Vol. 3, No. 3, pp. 271–281.
10. Madgison, J. (1988) 'Improved statistical techniques for response modeling', *Journal of Direct Marketing*, Vol. 2, No. 4.
11. Madgison, J. (1990) 'CHAID.LOGIT and log-linear modeling', Datapro Reports on Marketing Information Systems, McGraw-Hill.
12. Brooks, P. (1997) 'March of the data marts', Database management system (DBMS), March.
13. Duffill, M. (2001) 'Shelter donor datamart', *Journal of Nonprofit and Voluntary Sector Marketing*, Vol. 6, No. 4 (in production).
14. Micheaux, A. and Gayet, A. (2001) 'Turning a marketing database into a relationship marketing database', *Journal of Interactive Marketing*, Vol. 2, No. 4, pp. 327–346.

Further reading

Huff, D. (1954) *How to Lie with Statistics*, Pelican Books.

Koslowsky, S. (1999) 'Reducing your risk: What's happening in retail database marketing', *Direct Marketing* (US).

Miller, B. and van Wyk, J. (2000) 'Life-cycle-based donor performance analysis', *Journal of Nonprofit and Voluntary Sector Marketing*, Vol. 5, No. 1, pp. 73–81.

Shaw, R. (1998) *Improving Marketing Effectiveness*, Economist Books.