



Driving Profit Through Quality Marketing Data

> Articles ([back to magazine](#))

Marketing data (MD) often fails to achieve the planned or desired results. So non-marketers see its use to generate profit with some scepticism... Marketing data quality must underpin the analysis and support any subsequent marketing actions. Let's explore how this can be achieved.

Poor data quality has a negative impact for two reasons:

- 1) The analysis and its subsequent use are incorrect because the data has too many errors to give the correct statistical result.
- 2) Any MD set, used for analysis or action, is usually already a subset as one or more data variables has been used to pre-select it. The data would be drawn from several different databases all with their own idiosyncrasies. If there are errors in the variables used to make a selection (which there often are), any subsequent activity or customer interactions are based upon an incomplete (smaller) set of records/customers.

The aim of this article is to illustrate a simple approach (used prior to marketing analysis) to address this crucial Data quality issue. This could save you time and money in the long run, and if done periodically enables you to avoid the replication of the data quality before any marketing analysis, and improving the profitability of your marketing efforts.

You can overcome many data quality issues, if you know what you are up against! – before waiting for an interaction with your Customer to expose data quality issues, risking the potential negative impact on development activity.

Business Data Coverage

The first step is to: *Confirm that the extracted MD set meets business expectation.*

The way to do this is to undertake a set of "yearly" analysis and compare this against what is expected from the business MI for the same period, i.e. the correct Business Data Coverage. As a minimum check the variables suggested below per year (calendar year or business see example table 1 below:

Table 1

Year 1

Data Field Title	From MD Set	Business MI	Variance	Percentage Variance
In force 1st Jan to 31st Dec				
Direct				
Number of Customers				
Number of Sales				
Value of Sales				
Number of different types of Products sold 1st Jan to 31st Dec				



Martin Hickley

[email](#)

[web](#)

Full list of articles for August 2006

- [Turn data into information; only then can it become actionable knowledge that equals profit](#)
- [Data and Strategic Marketing Decision-making \(webinar replay\)](#)
- [Driving Profit Through Quality Marketing Data](#)
- [Dirty data reduces Marketing's return on investment](#)
- [Predictive data triggers a new era](#)
- [Is data dated or dateless?](#)
- [Accountability: your duty to turn data into profit](#)
- [Data mining for marketing profit](#)
- [Data aggregation and consistency for best return on investment](#)
- [Marketing data defined: How can you increase profitability?](#)

Number of renewals
 Through Third Party
 Number of Sales Third Party
 Value of Sales per Third Party

It is important to include industry specific indicators, for example, Net Receivables for Credit Cards; Number or Annual Services for Automotive, number and value of Claims for Insurance, etc.

Note: If The Business Data Coverage variance is less than 5%, that is exceptionally good. Our experience is that for most companies at least one field will be more that 30% different, and we have experienced up to a 90% variance where there is data missing. If the variance is more than 5% the MD should be re-extracted.

Valid Data Population Analysis

Step 2:

The next step is to assess the correctness of each individual field. This can be achieved by building a descending hierarchy of fields, as per the examples in the table 2 below, to undertake the validation analysis:

Table 2

MD Field Name	Validation Rules	Number which Pass	Percentage Correct
Total Number of Records in MD	NA		
Customer ID	8 Characters Numeric		
Title	Comparison with Titles Tables		
Initials	Alpha only		
Surname	Alpha only; First Character upper case, Subsequent lower case		
Postcode	PAF valid		
First line of Address	PAF valid		
Second line of Address	PAF valid		
Third line of Address	PAF valid		
Fourth line of Address	PAF valid		
Date of Birth	DD/MM/YYYY		
Gender	M or F		
Telephone	Valid std code and 9 or 8 Numeric		
Email address	Alpha and numeric characters, with one @ and (.net .org . com, .co.uk, etc)		
Account ID	10 Characters, numeric only		
Product Start Date	DD/MM/YYYY. Do Frequency checks		
Product End Date	DD/MM/YYYY Do Frequency checks		
Sales Value	Numeric only, two decimal places		
Product ID	10 Character, First 4 Alpha, 6 Numeric		
Marketing Consent	"Y" or "N"		

The examples used above have been simplified but illustrate the types of validation rules that can be built from the front-end system data capture rules, internal system rules and common sense; they don't have to be embedded into any system.

Note: This analysis normally takes several iterations, because the data has been collected over several years, system formats have changed and the validation rules are the current ones people remember. Often the results are unexpected and can shake the heart of your core operational databases, for example the idea that all customer id's or account id's are in the correct format. In our experience this is rarely true. Similarly inception dates can be a minefield; this is an area where system loads/migrations have contributed to the errors, and e.g. all dates being set to the date of load.

Finally some of the valid field populations can be surprisingly low, e.g. valid Telephone number 40%, valid email address 5%.

Often the original extracts will have to be redone after this analysis due to the lessons learned and the Business Data Coverage will be much closer.

Drop Down Reports

Whilst data quality at field level is useful, to really understand your business data quality you need to undertake this analysis at a record level.

Step 3

Take the previous table 2 and now run it at record level, from the top data field, then the top and second data field and so on.

Table 3:

Field Number	MD Field Name	Records that Pass Validation Rules for:	Number which Pass	Percentage Correct
	Total Number of Records in MD	NA		
1	Customer ID	1		
2	Title	1 & 2		
3	Initials	1, 2 & 3		
4	Surname	1 to 4		
5	Postcode	1 to 5		
6	First line of Address	1 to 6		
7	Second line of Address	1 to 7		
8	Third line of Address	1 to 8		
9	Fourth line of Address	1 to 9		
10	Date of Birth	1 to 10		
11	Gender	1 to 11		
12	Telephone	1 to 12		
13	Email address	1 to 13		
14	Account ID	1 to 14		
15	Product Start Date	1 to 15		
16	Product End Date	1 to 16		
17	Sales Value	1 to 17		
18	Product ID	1 to 18		
19	Marketing Consent	1 to 19		

Often for field 19 in the hierarchy, the percentage is between 60% to 85%, although in an intermediated business 40% is not unusual.

The 3 steps of data quality analysis

The 3 steps of data quality analysis will give you a good insight to your MD quality.

The analysis clearly identifies three sets of data:

- 1) MD fit for use,
- 2) MD which needs some work,
- 3) MD so bad you need to archive it.

It will identify both where you have problems and usually the necessary remedial action – both IT lead, and changes to front-end systems.

These steps will give you a fully auditable, correct MD to be used in analysis. This can be used, with confidence to gain the insight to inform and plan activity to drive profit. If done of a partial set this will allow you to do the analysis based upon correct data and widen it to the full customer base.

Finally you can undertake customer development activity with confidence, knowing that each customer interaction will not risk exposing your Customers to the reality of your data quality.

By Martin Hickley,
Customer Essential
Email: Martin.Hickley@customeressential.com
Web: <http://www.customeressential.com>

About Customer Essential

Customer Essential provides experienced and practical consulting to drive more profitable customer management activity.

:: theMarketingLeaders is a trademark and its respective community and publications are copyright Bipedal Ltd. :: All rights reserved.

:: [Bipedal](#) :: [Use](#)