

WHITE
PAPER

THE IMPACT OF BAD DATA ON BUSINESS PERFORMANCE

Bill Kandohla
Business Consultant

... poor-quality customer data leads to significant costs such as higher customer turnover, excessive expenses from customer contact processes like mail-outs, and missed sales opportunities.

But companies are now discovering that data quality has a significant impact on their most strategic business initiatives, not only sales and marketing. Other back-office functions like budgeting, manufacturing and distribution are also affected.

- Gartner Research

THE IMPACT OF BAD DATA ON BUSINESS PERFORMANCE

1. Introduction

Twenty years ago a database was something rare. Today, every organisation has multiple 'islands' of data. These are usually managed by standard database software and spreadsheets, although larger organisations have invested in powerful CRM and ERP systems to bring all customer-facing information together. Despite the promises of huge data systems, most organisations still run on stand-alone databases, enjoying the many performance improvements that have emerged since the first spreadsheet product was launched back in the 1980s.

These islands of data help managers to get their jobs done; and save their organisations the huge costs of custom applications. Continuing software improvements have given managers access to capable, off-the-shelf tools which help them to function daily with the data at their disposal. That is why most businesses today have a marketing database to track and capture prospects, a customer database for the accounts department, a sales contact system, a customer services database, and many more.

Any system, including a database, will inevitably develop errors. No matter who collects the data, manages it or uses it, different databases will always go their own ways. For the business, the costs of such 'bad' data start to add up. They can include lost marketing and sales opportunities and an eroded brand image. This is why today's critical issue with data is neither storage nor availability, but quality. Once the quality of current data is addressed managers can go on to change other parts of the data capture and management process to engineer bad data out of the system in the future. In this white paper we will discuss data quality and outline some of the things managers need to take into account when improving it.

2. So what's the problem?

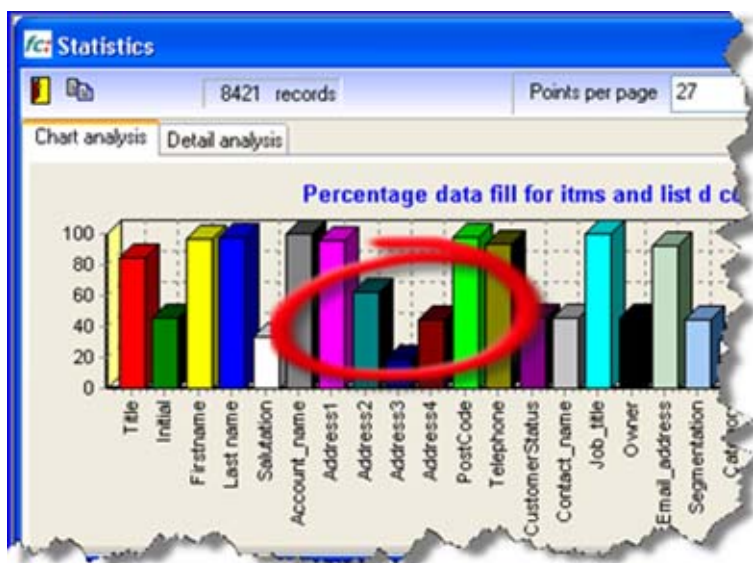
Take a database, any database. Create one if you like. Then let someone else work on it. Within a month you will have a number of errors. These will range from ones that upset customers but keep the cash flowing, like spelling their names wrong, to ones that impede customer service, like getting the postcode wrong so your delivery vans cannot get through.

The most common errors are missing data, spelling and punctuation mistakes and inconsistencies, the right data in the wrong field, partial data, and repeated data. **Our experience suggests that the error rate in most commercial B2B databases is around 30%.** That's millions of records in the UK alone with incorrect data of one kind or another. Even if we allow a percentage for 'churn' as businesses change address and personnel, or employees change job titles and contact numbers, there's still a huge amount of avoidable clutter in every database. Almost certainly yours.

3. Data is always missing.

You want a database to store all the data you need - but it almost certainly doesn't. Every time we examine a database, there's missing data. Take some simple examples. The 'title' field in a database is typically only 30% filled. The 'job title' field is usually just 50% complete, and of the remaining 50% a significant number are strange or invented. Many entries confuse job title with job function, and in some prospecting databases job title/function confusion is widespread and many contacts are listed with 'unclassified' as the job title. We also find many empty 'first name' fields or just an initial. But for us the biggest single surprise is always the number of missing phone numbers - up to 30% of records.

THE IMPACT OF BAD DATA ON BUSINESS PERFORMANCE



SEEING WHAT'S THERE.

No database is ever complete. The problem is to quantify its weaknesses. A data utility can show managers where the gaps are and enable them to drill-down for extra detail.

In this example, the first level analysis shows that several fields are under-populated. Most of them are important, especially for a company which uses email marketing or wishes to address its customers correctly.

Missing data can usually be found. Junk data can be identified and corrected. It is not rocket science. And if the organisation is too busy, service providers such as Dun & Bradstreet, Harte Hanks and IDG can help. The question for managers is: *If you don't have critical data like the right name, an email address or a phone number, why keep the data?* Many managers never ask this question because they never analyse the database. Others keep redundant data because they are assessed on activity: a larger database impresses their directors.

4. Data is always wrong.

It is obvious that an internal database - one that holds employee details for example - will become out of date as soon as the information is captured. The question for managers with this sort of database is: *How up to date do you want to keep it?* In an ideal world, every employee's manager should tell you when a staff member leaves, but that is never going to happen. The alternative is to contact managers regularly and update the records. This is resource-intensive and not always practical.

The issue is this: *If it is this hard to keep an in-house database correct, how much harder is it to keep a customer database correct?* Your employees can let you know if you spell their names wrong; your customers are more likely to move their accounts to companies that can at least get their names right.

5. Data is always in disguise.

Bad data comes in several forms:

- *Duplicates*. The same record, twice or more. Easy to spot and correct, even using a standard spreadsheet.
- *Duplicates Level 1*. Same name/different address, or same address/different name. Is the same person? Has the person moved? Some people may have two addresses because they work from two or more sites. They may use a home address to subscribe and a work address for other contacts. The only way to clean up these issues is to contact the person and check. A costly approach.
- *Duplicates Level 2*. Variations within the same field. So here we may have A. Smith, or A.Smith, or A Smith, or Alan Smith or Allen Smith, all at the same address. Once again calling is the best way to sort this out.

THE IMPACT OF BAD DATA ON BUSINESS PERFORMANCE

- *Duplicates Level 3*. Inconsistent company names. For example, BT, B T, B.T., B. T., British Telecommunications, British Telecoms PLC, BT Limited and so on.
- *Invalid or obsolete data*. This includes 'gone away' contacts. Every database will have names that have moved on, and the older the database the more there will be. The typical failure rate can range from 10% for a new list to 50% for an older list. In business, these errors are usually detected by the sales team when they try to call lists provided by the marketing department! However, old data is inevitable in any database. The question for managers is: *What are you going to do about it?*
- *Bad addresses*. These are usually basic failures in data entry, such as putting the town in the wrong column. Postcodes make some of these issues less important.

6. Counting the cost - cash.

So what? Well, how much do you earn per customer per year? **How much more could you earn if you could market to more 'clean' records in your marketing, customer and prospecting databases?** Most businesses generate 80% of their cashflow from a solid core (around 20%) of their customers. Losing even 1% of this loyal base through bad data can potentially reduce cashflow by around 2-3%. It's the **quality** of your data that matters, not the **quantity** as many managers still believe.

7. Counting the cost - reputation.

In addition to the cash opportunities lost by bad data, managers also need to consider the impact of junk on brand image and customer relationships. Customer campaigns based on bad data quickly run out of potential. Provided that today's customers remain loyal the impact may not be significant. But as and when they fall away (and they will) cashflow will suffer. Every business needs to recruit new customers to replace outgoing ones, but if it cannot even get their names or addresses right how successful will it be?

It takes years to build a good reputation, and months to throw it away. For some reason, managers hardly ever consider the impact of data quality on a brand. Some years ago we came across an instance where a mailshot had been sent with the salutation *Dear Blah Blah Blah*. It was a simple error - forgetting to merge the salutation field in the letter. It was laughed-off at the time but we can only imagine the impact this had on the brand.

8. High impact, low priority.

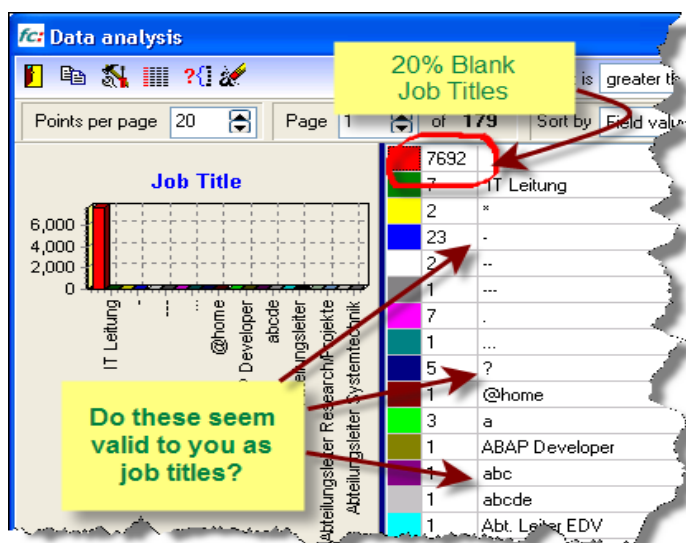
Many managers believe that it is cheaper to address data issues as they emerge than to manage data proactively. This is based on a false calculation. True, spending time to maintain data adds to overheads - but dirty data loses profits! This is why neglecting data is a false economy. Unless a database is actively maintained, bad data will accumulate until it reaches critical mass. It is at that tipping point that a business can lose real money.

9. Is your data really necessary?

Bad data sometimes gets that way because it is never used. We go further: we believe that a lot of it need never be collected in the first place. Think of the web sites that demand a full set of contact details before releasing a free document or adding you to a newsletter. What is the point? This is why many data collection forms are filled with rubbish. People will enter junk to get where they want to be.

THE IMPACT OF BAD DATA ON BUSINESS PERFORMANCE

Amazingly, some of this junk data is actually used! Mailers are sent to 'x' at an address that's 'x', 'x', 'x', 'x'. Why? Because once rubbish has been collected it is cheaper in the short term to mail it than to clean it.



ASSESSING THE DAMAGE.

This drill-down shows how bad the job title field has become. This database has not been quality checked even though it has been used to contact customers.

The hardest part of managing data is to accept that it can be wrong – and then doing something about it. Identifying problems is not a big issue. Corrections are easy to make. But are you ready to admit that your data could be wrong?

10. Six steps to better data.

- *Review it as it is collected.* It costs ten times more to clean data after it has been stored. This is why so many organisations accumulate bad data and carry the costs of mailing to the wrong (or non-existent) person or address. Simple housekeeping should be in place to validate data during collection to achieve quality gains and cost savings.
- *Analyse the database.* What is in it? Is it what you need? How reliable is it? Are there any errors like duplicates? What is your strategy for maintaining it?
- *Measure the challenge.* How can you manage an issue if you cannot put a number on it? Database problems have to be analysed to help managers measure the scale and importance of the cleaning task. Managers need to 'see' where data is missing, inconsistent, corrupt or duplicated so they can decide what needs to be done and choose the most appropriate response. Some form of analysis tool that's faster and easier to use than a standard spreadsheet may be helpful, especially for managers who do not regularly touch data but need to keep information under control when they need it.
- *Get it right for the future.* Once data problems are highlighted, they can be engineered-out for the future by changing collection and entry processes. Any errors in existing data will have to be tackled as resources permit, although something beyond a spreadsheet may be required.
- *Develop a plan.* You know that a problem almost certainly exists in your data! Do you have a plan to overcome it? Every manager needs a strategy for measuring and managing data quality problems, as well as the tools to make the necessary improvements.
- *Get started!* The market is rich in tools that will help you to address any data issue. From desktop utilities to enterprise applications, they need to be evaluated to ensure a perfect fit for your budget and needs. Service providers such as Dun & Bradstreet, Harte Hanks and IDG are also there to help. The days of struggling with a rigid spreadsheet can soon be over.

THE IMPACT OF BAD DATA ON BUSINESS PERFORMANCE

11. Conclusion.

Data management is not about collecting and storing huge amounts of data to meet a target or impress the board. **It is about storing relevant, productive data and removing unwanted and unnecessary junk.** It is better to have a relationship with 100 known customers than to store the details of 10,000 unknown ones. There is always a cost per record - the more you have the more it costs. Wasteful data costs even more, so why spend your tight budget on it?

Data managers must take decisions. For example, if you mainly market using emails, why retain records with no email address? As we put it to our clients: *Is all that data really necessary?*

You need a quality strategy, not a quantity strategy. Data problems arise because the raw material has no visible value, even though the future of the organisation may depend upon it. Data management is passed down the line and no manager assumes responsibility for its quality. This has to change if businesses are to grow and protect their brands. That's why more and more businesses are now appointing managers with responsibility for all the data within an organisation.

Although an accountable manager, powerful data software and managed collection structures are essential to maintaining data quality, managers at the sharp end also need desktop tools to help them work with the data they've got, not the data they would like to have. **Taking control of 'real world' databases on the desktop is not an option** if managers are to merge, clean, de-dupe and analyse it and deploy it more effectively to reach-out to prospects and customers.

Before the web, prospects would soak up all the information a company sent. Today, they are scanning the internet to check-out your competitors. Customers and prospects are better educated, better informed and more determined to do a better deal. But reaching them and persuading them is much harder. **If up to 30% of your database is wrong at the start, you're already a third of the way behind your competitors.** This is why data quality is vital, and why every manager needs to take it a lot more seriously.