consumers select largely on the basis of function and convenience, but subject to the brand satisfying socially accepted standards of appropriate corporate conduct?

As a comparatively light user of social media, I could not help wondering how pervasive the effect of social media is on brand reputation. What proportion of the adult population are infrequent or non-users of social media? Of those who are regular users, how many are disposed to engage in conversations about brands?

Enabling opponents to self-organise

Or does social media's impact on brand reputation depend not so much on its reach as on it being a medium whereby opinion can rapidly permeate the awareness of an elite of highly influential communicators? Or is it because it is such an effective mobilisation tool, enabling opponents to self-organise rapidly and effectively?

For advocates struggling to influence sceptical colleagues, this will prove a highly motivating text with no shortage of detailed advice and plenty of relevant examples.

Richard Webber

Data smart: Using data science to transform information into insight

Journal of Direct, Data and Digital Marketing Practice (2014) **15**, 354–355.
doi:10.1057/ddmp.2014.33

Real value for self-help

What do you expect when you read a business or reference book? Ideally, the knowledge it contains should transfer positively into your working life as soon as possible. After all, why bother if there is no hope of that happening? (Although the continued growth of the self-help publishing sector suggests that hope tends to triumph over experience.) When that knowledge transfer actually happens, the book in question becomes one you treasure, recommend and constantly return to.

Highest possible recommendation

That is why I have to give John W. Foreman's new book (Wiley, US, ISBN: 978-1118661468) the highest possible recommendation because of my direct experience of what it contains transferring into my professional life. In April, the IDM ran its Data Discovery Workshop to provide a group of students an insight into how data is used in the commercial world and encourage these mathematicians, psychologists, engineers and the like to consider a career in the industry.

Bayes and a bag of words

A drinks reception was thrown for them on the evening of the second day. Earlier, I had some time to kill and so I continued to read my review copy of *Data Smart*. In Chapter 3, called 'Naïve Bayes and the incredible lightness of being an idiot', I learned about techniques for classifying texts (such as tweets) to understand whether they are positive or negative. Based

around natural language processing, it provides a step-by-step example of applying the ‘bag of words’ approach, Bayes rules and additive smoothing to train an Excel function to identify positive and negative references to a brand. As with everything in this book, Foreman is clear and persuasive, to the extent that after an hour I was convinced that I too could run some textual analysis of the type described.

Conversation made possible

I did not have to wait long to find out whether I had learned anything of value. That evening, I was introduced to one of the IDM’s students doing a PhD in computational linguistics, focusing on how images are tagged online. Armed with what I had read in Foreman, I was able to talk about MAP rules, artificial intelligence and the value these have for businesses trying to understand their status in social media. Without this book, I would not have been able to have that conversation.

Valuable to SMEs and large businesses alike

So it goes for the whole of *Data Smart*. If you have enough knowledge of Excel to create pivot tables, you will gain the sort of practical skills needed to create insight that you assume only super-users or those highly trained statisticians in the customer insight team possess. As a result, this is a highly valuable publication for any small- or medium-sized business that lacks those resources (or anybody working in a large enterprise who wants to get to grips with the impenetrable jargon being used around data-driven insight.)

K-means and Manhattan distances

Have you ever wondered how your customer clusters have been created, or what regression is really about? There are chapters here on both and after reading them you are likely to find yourself having conversations about K-means and Manhattan distances with a level of confidence that previously seemed unlikely. Throughout, the book provides everything you need to know to run customer problems through Excel, including lines of code, screen grabs and even links to sample data sets online, yet always in accessible and entertaining language. By the end, it has opened the door to the new world of data analytics using R, a next-generation data mining application.

Data science or classical analytics?

Foreman’s credentials are as good as they get, having been an analyst at blue-chip organizations such as Coca-Cola, Royal Caribbean and Intercontinental Hotels, before moving to his current role as chief data scientist for MailChimp.com. If there is one point on which I quibble with him, it is only in his use of the term data science. Throughout, the examples of projects discussed fit into existing customer insight and analytical practice, typically the need to test hypotheses and find evidence for decisions. Only in the latter stages does the underlying data move into more complex realms where patterns are investigated ahead of the theory as to why they exist, which is typical of data science and big data sets.

However, this is a minor point and does nothing to limit the appeal and value of this book. If you want to become better informed or even get hands on with customer insight, this book will undoubtedly help you to live up to its title and become data smart.

David Reed FIDM