1

Data Mining for Shopping Centres – Customer Knowledge-Management Framework

Charles Dennis, David Marsland and Tony Cockett
Brunel Business School, Brunel University
Uxbridge, Middlesex, UB8 3PH, UK
Phone: +44 (0) 1895 265242   Fax: +44 (0) 1895 203149
Email: charles.dennis@brunel.ac.uk

KEY WORDS: Shopping centers; Customer knowledge management; Customer profiling.

ABSTRACT

Shopping centers are an important part of the UK economy and have been the subject of considerable research. Relying on complex interdependencies between shoppers, retailers and owners, shopping centers are ideal for knowledge management study. Nevertheless, although retailers have been in the forefront of data mining, little has been written on Customer Knowledge Management for shopping centers. In this chapter, the authors aim to demonstrate the possibilities and draw attention to the possible implications of improving customer satisfaction. Aspects of customer knowledge management for shopping centers are considered using analogies drawn from an exploratory questionnaire survey. The objectives of a Customer Knowledge Management system could include increasing rental incomes and bringing new life back into shopping centers and towns.

INTRODUCTION

Shopping centers are an interesting topic for knowledge management – relying on interdependency between owner, retailers and shoppers. Why are shopping centers important? Firstly, planned shopping centers comprise a substantial part of the UK economy, employing over three-quarters of a million people and playing a 'key role in the investments of pension funds' (Davies *et al.,* 1993; OXIRM, 1999). Shopping centers are therefore important not just to customers, but also employees and indeed to many others because of the investments of their pensions. Secondly, retail and shopping centers form the heart of UK towns and create a focus for the community. Shoppers tend to follow the provision of attractive shopping areas. Improving shopper satisfaction can lead to changes in population, expenditure, residence patterns and bring new life to run-down areas (Dennis *et al.,* forthcoming). The findings of the research could be applicable to traditional high streets and towns as they are to purpose-built shopping malls – if there is in place some form of central administration such as Town Center Managers. This chapter considers the possibilities for shopping centers to make their offer more attractive using techniques of data mining and customer knowledge management.

DATA MINING AND CUSTOMER KNOWLEDGE MANAGEMENT IN THE RETAIL CONTEXT

Data mining has been defined as:

> 'The process of exploration and analysis, by automatic or semi-automatic means, of large quantities of data in order to discover meaningful patterns and rules.'
>
> (Berry and Linoff, 1997)

Berry and Linoff (2000) list six data mining activities: (1) Classification; (2) Estimation; (3) Prediction; (4) Affinity grouping or association rules; (5) Clustering; and (6) Description and visualization. Retail studies have included many other techniques (e.g. Sequence-based analysis; Fuzzy logic; Neural networks; Fractal-based algorithms - Rao, 2000; Rensselaer Polytechnic Institute, 1999). Nevertheless, Berry's and Linoff's six categories serve our purposes here.

Data mining has many uses, but the aspect of most concern here is what is usually known as 'Customer Relationship Management' (CRM). Good CRM means (1) Presenting a single image of the organisation; (2) Understanding who customers are and their likes and dislikes; (3) Anticipating customer needs and addressing them proactively; and (4) Recognizing when customers are dissatisfied and taking corrective action (Berry and Linoff, 2000).

Some UK retailers recognize the potential of data mining in discovering customer knowledge. For example, Halfords and Sainsbury's use Brann Viper software; Tesco and John Lewis Dunn Humby (Computer Weekly 16 January and 29 May 1997). Most, though, jealously guard their customer knowledge capital. The authors argue that dissemination of this knowledge to a shopping center owner could result in meeting shopper requirements better.

Since the mid-1980s, there has been an increasing recognition that 'knowledge is a fundamental factor behind an enterprise's success' (Wiij, 1994) – a statement that applies in the retail industry as in others. This chapter will consider shopping center customer knowledge management from Wiij's (1998) third, broadest focus: 'all knowledge activities affecting success … using knowledge assets to realize their value'. The specific concern is with customer knowledge management - the management and exploitation of customer knowledge. There are two aspects of this knowledge: (1) Knowledge about customers; and (2) Knowledge possessed by customers (Rowley, 2001). The empirical study reported here concerns the first aspect, but we will conclude with a recommendation for further research on the second.

Richards *et al.* (1998) argued that the marketing success of an enterprise is founded on 'a continuous dialogue with users, leading to a real understanding … the more mundane the category [shopping centers?], the more dependent on knowledge'. In the early 1980s, data warehousing transformed operational data into knowledge for decision-making. As retail IT systems company NCR put it: 'For retailers the key … is to establish data warehouses to improve and manage customer relationships' (Teresko, 1999).

Data mining can use programming methods to identify patterns among data objects – for example between products in a shopping basket. The well-known early example is the 'diapers-beer' link on Friday evenings spotted by Wal-Mart in the US. By placing the two side by side, more fathers took home extra beer when they went to buy the diapers after work. Woolworths (UK) have installed a system costing UK£2 million, claimed to have boosted sales in women's toiletries alone by more than UK£5 million per year (Bird, 1996). The authors contend that incorporating data mining and customer database aspects within a framework of Knowledge Management can help increase knowledge value.

The main focus of this chapter concerns the opportunities for data mining and customer knowledge management for shopping centers. Data mining normally refers to large quantities of data, so our survey of 287 respondents must be near the smaller end of the scope. Nevertheless, the dataset has been useful in illustrating the utility of aspects of data mining and customer knowledge management that may be suitable for larger-scale use. Further, the exercise has demonstrated that a full data warehouse is not essential. Rather, effective data mining techniques can be applied to a smaller sample drawn from a large database. Another aspect for discussion (not the main focus of

this chapter, though) concerns the possibilities of extending customer knowledge management to the sharing of information between shopping center managers and potentially competing retailers. The case for such sharing is not clear-cut. Howard (1995) pointed out that shopping center landlord/tenant relationships are characterized by bargaining and outright conflict. Most UK shopping centers are not customer-orientated (according to Howard). This statement may be arguable, but we concur to the extent that more customer knowledge could help shopping centers to make their offers more attractive. Howard is on safer ground in pointing out that more marketing success could be achieved by a utilizing a partnership approach for collecting, sharing and using information. Howard (1997) cited a store manager at the successful Lakeside (UK) center as claiming that CSC (the owner of Lakeside) is different and has a more open relationship between retailers and center management. The information-sharing approach (Howard implied) has contributed substantially to CSC's success.

Some retailers, notably the UK market-leaders Tesco (supermarkets) and Boots (drugstores) have exploited customer knowledge by means of loyalty schemes. Such schemes have been successful for retailers but are unlikely to pay for themselves by increased loyalty (Field 1997). Rather, the benefits arise from their function of facilitating the flow of information and rewards between suppliers and consumers (Worthington 1999), i.e. as part of a customer knowledge management system. Some UK towns and shopping centers have experimented with loyalty schemes but as far as the authors are aware, the potential knowledge benefits have not been fully explored. In the Cobham (small town in Surrey, UK) and Lakeside (regional out-of-town shopping center in Essex, UK) schemes data from customer receipts had to be entered by hand. For the town or center management, the method provided access to customer transaction information, without needing the explicit agreement of individual retailers to data sharing. There is, of course, a privacy issue concerning the use of customer data in this way. The shopping centers may well have taken the view that the transaction data belonged to the individual shoppers – who gave written consent for the data use when they requested the loyalty card. Certain large retailer tenant(s), though, are understood to have considered that they owned the data concerning their shoppers' transactions, and to have objected to the use of that data by the shopping center landlord's the loyalty scheme. Some schemes including Lakeside and Cobham have been dropped under the burden of paperwork or lack of support from retailers (Hallsworth, 2000). Lakeside replaced the loyalty card with an 'affinity' credit card - the administrative load was transferred to banks, but customer data were lost to Lakeside. Nevertheless, the authors contend that loyalty schemes can be successful. The essential aspect is to design them from the start for customer knowledge management.

In the interests of providing a preliminary illustration, this chapter reports exploratory mall interview surveys at UK shopping centers. In a full-scale application, data mining for customer knowledge management would be applied to a customer database, but such a dataset was not available to the researchers. As an alternative, data mining techniques such as cluster analysis and predictive modeling have been applied to the findings of a questionnaire survey. The standard SPSS program has been used for the analysis, being less expensive and more applicable to this scale of project than would be a custom data warehouse. The authors have explored the differences in behaviour between shoppers and drawn attention to differences between exemplar segments as to which attributes are critical in shopping center choice.

EXPLORATORY STUDY
The results are from a survey of 287 respondents at six shopping centers varying in size from small in-town sub regional to large regional out-of-town. A 'regional' center is defined as having a gross retail area of greater than 50000 m^2 and a 'sub-regional' one 20000-50000 m^2 (based on Guy, 1994a, b; Marjenen, 1993; Reynolds, 1993). The objective was to determine which specific attributes of shopping centers were most associated with spend for various subgroups of shoppers.

If it can be demonstrated that customer knowledge management can enhance the attractiveness of shopping centers and lead to increased store sales, there will be an incentive for retailers to 'buy in' to the idea of sharing customer data.

The study evaluated shoppers' comparative ratings of two shopping centers, one of them being the center where the interview took place. The alternative center was the one where they shopped most (or next most after the interview center) for non-food shopping. The questionnaire instrument was based on the 'attributes of image' elements employed by McGoldrick and Thompson (1992a; b) together with additional constructs derived from analyses of preliminary unstructured interviews. Respondents stated their perceptions of the 'importance' of each of 38 attributes, including those identified by Guy, (1994a; b) as figuring in consumers' choices of shopping destination, for example, 'Quality of stores', 'Cleanliness' and 'Availability of rest rooms', following a similar procedure to that used by Hackett and Foxall (1994). Each attribute was also 'rated' for both the center studied and the alternative center. Respondents estimated perceived travel distance and time to both centers and supplied details such as age, location of residence and occupation of the main earner in the household. Examination of the characteristics of the sample indicated the distribution of socio-economic groups, age and sex reasonably representative of that anticipated at UK shopping centers. The number classified in the higher socio-economic groups of managerial, administrative, professional, supervisory or clerical (ABC1 on the UK JICTAR scale) was 59%. This compared, for example, with a figure of 63% for the Lakeside (UK) out-of-town regional center (owner's proprietary survey of 2000 respondents over two years) and 55% for the Treaty Centre, in-town, sub-regional (Hounslow, UK – from the center 'Education pack' citing 'street surveys'). The proportion in the younger age groups 16 to 44 years was 65% in our sample compared with 73% at Lakeside and 67% at the Treaty Centre. Our sample was 69% females compared with 60% at Lakeside and 59% at the Treaty Centre.

Further questions concerned typical perceived monthly spend at each of the two centers. As McGoldrick and Thompson (1992b) pointed out, much of the variation in shoppers' expenditure relates to factors such as income or socio-economic groups, rather than travel distance or attributes of the shopping center. Following this approach, the main dependent variable was the 'individual relative spend'. A value of 100 indicated all expenditure at the center studied and none at the alternative center. A value of 50 indicated half of the expenditure at each center. The same approach was used to scale perceived travel distance and time producing the variables 'individual relative travel distance' and 'time'.

The view of 'attractiveness' taken by the authors is that any product (such as a shopping center) 'can be seen as a bundle of expected costs and rewards' which East (1997, page 131) found was 'upheld by research'. East drew support from Westbrook's (1980) finding that an overall measure of retail satisfaction correlated well with a simple addition of the satisfactions. In the authors' procedure, the measures of satisfaction and dissatisfaction have been taken from the respondents' ratings of the shopping center compared to their main alternative center (on 5-point semantic differential-type scales). These satisfactions for the individual attributes were weighted, firstly by the 'importance' of the attribute to the respondent (also on a 5-point scale) and secondly by the degree of association with the stated relative spend. Once weighted, satisfactions were added, giving an overall 'Attractiveness' measured value. The next stage was to combine the attractiveness measurements with the relative travel time or distance variables, to derive (statistically significant) models of individuals' relative spend. More detailed derivations of the attributes and models have been reported elsewhere (Dennis *et al.,* 1999; 2000a).

Attribute evaluations have been considered as interval rather than ordinal data (following the approach of Oppewal and Timmermans, 1999). Ordinary least squares regression analysis has been

used to investigate associations between shopping center attributes and shoppers' spend at the center studied compared to a competing center. For example, 'Cleanliness' was the attribute most associated with the spending of female shoppers, $R^2 = 0.075$. Individual regressions were performed for each variable; multiple regression was less appropriate on account of multicollinearity (Dennis *et al.,* 1999). Attribute ratings have been summed and combined with travel distance to allow comparisons between the subgroups of the fit of each model. R^2 values were between 0.09 and 0.40 - i.e. 'modest'.

In the analysis of the results we have firstly used conventional demographics to group shoppers, eliciting the most significant shopping center attributes separately, for example, for females and males. A further stage concerned the identification of attributes for various motivation clusters. Retail data mining schemes have aimed to identify subgroups that share similar shopping motivations. Researchers (Boedeker 1995; Boedeker and Marjenen, 1993; Jarrett, 1996) have identified shopping center motivation clusters. Targeted marketing mixes satisfy these more appropriately, increasing satisfaction, sales and profits. These researchers identified two subgroups (among others) that could be described as 'shopping' and 'service' motivations. It is hypothesized that members of these two groups can be identified as individuals for marketing communications purposes. Those primarily motivated to shop by attributes such as quality of the stores and selection of merchandise can be contrasted with those more interested in service and experience aspects such as the availability of good rest rooms and cleanliness. Accordingly, our study has also included a cluster analysis approach aimed at identifying the attributes critical for shoppers motivated by the importance of 'shops' *vs.* 'service'.

RESULTS
Table I lists the 'top six' attributes associated with individual relative spend for the subgroups. This table is designed to be read horizontally with comparative groups (e.g. females *vs.* males) side-by-side. The R^2 columns indicate the coefficients of determination of the specific attributes from linear regression with relative spend. Thus, these R^2 values are used here as a parameter to indicate the strength of the association between the attributes for the particular groups and shopper spending. Below the 'females *vs.* males' comparison follows a comparison of higher *vs.* lower socio-economic groups. 'ABC1' refers to the (UK, JICTAR) classifications of managerial, administrative, professional, supervisory and clerical. 'C2DE' categories include manual workers, senior citizens and unwaged. Comparisons of higher *vs.* lower income and age then travel by auto *vs.* public transport follow. The final comparison is of the shopper clusters that we have termed 'service importance' *vs.* 'shops importance' motivation.

[Take in Table I]

Conventional Demographics
Females *vs.* males: The attributes significant for females were clearly different to those for males, with 'cleanliness' top for females, significantly different with respect to the association with spend compared to males. Only one of the 'top six' attributes for females ('Nice place to spend time') was significant for males. Conversely, three out of the 'top six' attributes were significantly more associated with spend for males than for females ('Lighting', 'Sheltered access' and 'No undesirable characters'). Space limitations preclude a full discussion, but for females, two separate factors have been elicited (maximum likelihood extraction and varimax rotation - Kinnear and Gray, 1997). We have named these factors **shopping** (including, for example 'Selection of merchandise') and **experience** (exemplified by 'Friendly atmosphere'). On the other hand, for males the concerns were with the **center** ('Lighting' and 'Sheltered access' – the factor analysis did not produce separate factors for males). The interviewers reported that many males were in the center mainly to accompany females. Our interpretation of these results is that females, who were

enjoying the trip, were naturally concerned with 'shopping' and 'experience'. Conversely, males who were simply 'there' were more evaluative of the 'center'.

Upper *vs.* lower socio-economic groups: For managerial, administrative, professional, supervisory and clerical (ABC1s), 'Lighting' and 'Access by road' were significantly more associated. For manual workers, senior citizens and unwaged (groups C2DE), 'Good for children' and 'Quality of the stores' were among the most significant. The differences are to some extent understandable in the light of our observation that upmarket shoppers are more likely to travel by auto, whereas those from the lower socio-economic groups are more likely to bring children on shopping trips.

Higher *vs.* lower income groups: 'Lively or exciting' and 'Covered shopping' were significantly more associated for the lower income respondents. The authors speculate that lower income (and lower socio-economic group) shoppers might tend to live nearby, patronizing as alternatives small, unexciting local centers. Therefore, they might tend to appreciate the benefits of lively covered shopping centers more than do the more upmarket customers who may take these benefits for granted.

Older *vs.* younger shoppers: 'Eating and drinking' was in the 'top six' for the older shoppers who we expect might shop at a slower pace than younger ones and take more refreshment breaks.

Shoppers travelling by auto *vs.* public transport: 'General layout', 'Choice of major stores' and 'Eating and drinking' were in the 'top six' for shoppers travelling by auto but not significant for public transport. Four of the 'top six' were significantly more associated for 'public transport': 'Shoppers nice people', 'Availability of seats', 'Big shopping center' and 'Value for money'. The authors consider that most of these attribute differences are related to differences in spending power. For example, shoppers travelling by public transport are more likely to appreciate (free) seats, compared to the more affluent auto travelers who choose to relax in a restaurant, bar or café.

Cluster analysis: shoppers motivated by the 'importance' of 'Shops' *vs.* 'Service'
An alternative to the conventional demographics approach was the search for clusters of buyers who shared needs or wants for particular benefits. A cluster analysis (SPSS 'K-means', minimizing the squared distances of each value from its cluster center, (Kinnear and Gray, 1997) based on 'importance' scores has identified distinct subgroups of shoppers classified by 'importance motivation'. The main attributes that distinguished the clusters (with the average 'importance' scores on the 1 to 5 scale, where 1 was 'no relevance' and 5 was 'extremely important') are listed in Table II. These segments were described as 'Shops importance motivation' (abbreviated to 'Shops'), Table II (a) and 'Service importance motivation' ('Service'), Table II (b). The two 'importance motivation' clusters were strikingly different in attributes significantly related to relative spend (the final section of Table I). As hypothesized, 'Quality of the stores' and 'Selection of merchandise' were both in the 'top six' for the 'Shops' group, with 'quality of the stores' being significantly more associated with spend for the 'Shops' group than for 'Service'. For the 'Service' shoppers, 'Availability of good rest rooms' and 'Cleanliness' were among the most significant. Not so expected, 'Other shoppers nice people' and 'Lively or exciting' were significantly more associated with spend for 'Shops' than 'Service' shoppers. One possible interpretation might be that consumers motivated by 'shops' are evaluating not just the tangible merchandise but also the shopping experience. Our term 'shops' encompasses not just the physical environment of the shops but also the wider systemic shopping environment.

Compared to 'Shops', the 'Service' shoppers were slightly higher socio-economic group (63% ABC1s *vs.* 59%), income (60% UK£20000 per year + *vs.* 53%) and age (42% 45 + *vs.* 33%) than the 'Shops' group. They predominantly traveled by car (90% *vs.* 52% - see Figure 1).

[Take in Table II]
[Take in Figure 1]

MODELS OF RELATIVE SPEND
The regression models for the various groups are reported in Table III. These are introduced in the same order as in the reporting of the critical attributes for these groups in the 'Results' section above and in Table I. The models describe the relationships observed between relative spend for the groups *vs.* the attractiveness of the centers and the distance shoppers travel. The first column is the group for which the model applies (the numbers of respondents in each group were indicated in Table I). The second column is the constant from the regression equation, representing the amount of relative spend not associated with the variations in attractiveness and distance. The third and fourth columns are the regression coefficients for attractiveness and distance respectively. The fifth column is the coefficient of determination, R^2 of the regression equation, and the sixth the degree of significance (p-value). These two columns indicate modest correlations. All of the models would be normally be described as 'significant'. All except model numbers 6 (lower income) and 10 (travel by public transport) would actually be considered 'highly significant'. The final column is simply the identification number allocated to each model to facilitate discussion.

For example, for the 'Shops' group:
(11) **Spend = 19.4 + 0.70 X Attractiveness - 0.21 X Distance**
Whereas for the 'Service' shoppers:
(12) **Spend = 39.6 + 0.54 X Attractiveness - 0.28 X Distance**

These models mean that we can be confident (at normal test levels) that an increase in the attractiveness of a center would result in an increase in spend at that center. For example, for the 'Shops' group (11), the increase in spend for a given improvement in attractiveness would be greater than for the 'Service' group (12). By going back to the weighting that each attribute carried in the attractiveness model, it is possible to predict by how much spend would be likely to increase for any given improvement in any attribute. The models also mean that spend was inversely related to the distance that shoppers traveled to the center.

[Take in Table III]

In the 'Exploratory study' section above the procedure for calculating respondents' weighted satisfactions was outlined. The satisfactions for all attributes were summed to give each respondent's total satisfaction score for the center studied. The average of the respondents' satisfaction scores represented a measured attractiveness score for each center, the 'Brunel Attractiveness Index'. This index has been described more fully elsewhere (Dennis *et al.,* 2000b). Stated briefly, the Brunel Attractiveness Index is an empirically derived measure of shoppers' evaluations of the attractiveness of shopping centers.

The utility of the models has been investigated by examining the relationship between the empirically measured attractiveness and the estimated sales turnover. Figure 2 illustrates the relationship between our measured attractiveness, the Brunel Attractiveness Index and the estimated sales turnover for the six centers. The sales value scale has been changed by an arithmetical factor in order to disguise commercially sensitive data. The sales turnover values are necessarily estimates based on the questionnaire responses plus footfall data of unknown reliability supplied by the center managements. The estimates, though, were made before the models were designed – and were not used in the development of the index. From an inspection of Figure 2, it would appear that the

modeling procedure has been effective in measuring attractiveness in a manner relevant to sales turnover.

[Take in Figure 2]

The models are useful in estimating changes in spending that could result from improving aspects of a shopping center. For the high spending 'Service' shoppers (model 12 in Table III), a 25% improvement in the ratings for cleanliness and rest rooms could be associated with an increase in spend for those shoppers of 10%, equivalent to an increase in the total center sales turnover of over 3%. One measure of the validity of the subgroups is the improvement in 'fit' of the models. 'Service' *vs.* 'Shops' had the best fit, with R^2 increased to an average of 0.195 for the two subgroups. Apart from 'Income' (average R^2 0.175), the models from the other pairs of groups did not improve the fit above the overall level of 0.16. 'Service' *vs.* 'Shops' discriminated well between high and low customer spend, with the 'Service' segment's average stated monthly spend UK£82, compared with the overall average of UK£65.

DISCUSSION AND CONCLUSIONS
Information from a customer database can be used to identify needs of different groups of customers. This knowledge can help shopping centers to improve marketing communications and customer satisfaction. Cluster analysis has identified a group of customers that shopping centers and retailers will want to target: high-spending 'Service' shoppers. How can they be identified, given the high costs of a data warehouse? Firstly, this experimental study has demonstrated that a full data-mining system is not essential. Analysis, identification of target segments and assessment of cost-effectiveness can be carried out on a small sample, with only simple processing needed on a complete database. As in this experiment, the SPSS program can be used - saving the costs of custom software. For future, larger-scale projects, though, the authors recommend the use of a multi-agent system. Such systems can handle text alongside quantitative data and furnish individual shoppers with a 'personal agent'. This represents customized marketing segmentation – a software 'personal shopper' for every participating consumer. So far in this chapter we have considered the knowledge **about** customers aspect of customer knowledge management. The personal agent system could address the knowledge **possessed by** customers aspect. It could be argued that such a system might not work in the UK cultural context. For the customers, though, this would be a small step from the well-established loyalty card. The customer might only be aware of the difference when presenting a 'smart card' to obtain benefits or information. Customers having a personal agent could receive communications specifically targeted to their needs and wants. There are a number of ways that this could be achieved, but one of the simplest would be for customers to present their card for reading at an information kiosk in order to receive personalized vouchers and information sheets.

In the 'Models of relative spend' section above, it was pointed out that a (probably achievable) improvement of 25% in the ratings for 'cleanliness' and 'rest rooms' could be associated with an increase in spending by the service shoppers of 10%. The 10% increase for this group would add 3% to the total center sales turnover. A regional shopping center would gain tens of millions of dollars sales, with retailers seeing a seven-figure increase in gross profits. In the medium term, rental incomes follow sales: shopping center owners could expect US$2 million in increased rents.

Customer knowledge management systems could be based on data sourced from loyalty schemes. Worthington (1999) reviewed the typology of local loyalty cards in the UK. Integrated chips (e.g. Nottingham), and magnetic stripe payment (Hereford; Lakeside) or non-payment (Chester; Meadowhall) are applicable and cost-effective for cities and regional shopping centers. The main distinguishing feature of the higher-spending 'Service' shopper cluster was the preponderance of

auto as the means of travel - 90% of the group (illustrated in Figure 1). Therefore, for smaller centers, a scheme could be based on parking. For in-town centers that charge for parking, our solution is the '**parking lot membership scheme**'. Shoppers would buy a '*carnet*' of tickets at a discount and fill in a detailed 'lifestyle' questionnaire including the information needed for the database. Parking lot schemes are already in use in Australia (Worthington and Hallsworth, 1999). For centers that offer free parking, the suggestion is to recruit shoppers at a kiosk in the parking lot, offering incentives such as a prize draw.

The results presented in this chapter have demonstrated what can be achieved using some of the typical data mining activities applied to a simple dataset of survey data. In terms of Berry's and Linoff's six activities, we have (1) 'classified' using, for example, standard socio-economic groupings, evaluating critical attributes for those segments; have (2) 'estimated' potential increases in sales arising from changes to these critical attributes using (3) 'predictive' modeling. We did not use (4) 'association rules' in the usual basket analysis context. Rather our 'affinity grouping' was achieved using (5) 'cluster' analysis - the most effective classification technique of our modeling exercise. Finally, we contend that our analysis and modeling process has assisted the (6) 'description and visualization' of shopper behavior.

In terms of Berry's and Linoff's four components of CRM, we have outlined an effective procedure for measuring the (1) 'image' of a shopping center. Evaluating the image of the different customer groups has led us to (2) a greater 'understanding of who the customers are and their likes and dislikes'. Although the methodology does not (3) 'anticipate' customer needs, the survey approach does at least allow needs and wants to be 'identified and addressed proactively'. Similarly, using survey data in the database has identified a number of instances of (4) 'customer dissatisfaction', leading to recommendations for 'corrective action'. CRM is normally implemented by a system of personalized communications (e.g. welcome letter; satisfaction questionnaire; special offers and so on). The details are beyond the scope of this chapter, but an applicable strategy should be facilitated by the installation of a simple data mining and customer knowledge management system. Shopping center managers could obtain a similar level of data to ours from (for example) a membership questionnaire, and could use a similar analysis process to that described here. Such activities would have to comply with data protection principles, but in the UK at least, many shoppers are willing to part with personal and transaction information in exchange for benefits – the principle behind the success of the Tesco and Boots loyalty cards.

Adding real sales transaction information would enrich the possibilities – (although this raises the possibility of conflict with retailers over ownership of the data). Shoppers might be grouped according to spend on fashion/designer styles rather than bargains? A Knowledge Management network between retailers and the center would be a further stage - allowing wider access to graphs, patterns and associations in the data. There is a parallel in systems that multiple grocery retailers operate with suppliers. It is understood that supermarkets such as Tesco allow suppliers direct, real time access to individual store sales and stock data via the Internet. In this model, a retailer and (potentially competing) suppliers share data in a knowledge network managed by the retailer. Bearing in mind the relationships of conflict rather than co-operation (mentioned in the earlier part of this chapter) that, according to Howard (1997), dominate shopping center landlord/tenant relationships, co-operation in a Knowledge Management network might seem unlikely. Nevertheless, Howard did identify one UK shopping center owner (CSC) that was the exception. On the basis of the limited empirical results reported here, little further analysis of this issue is possible, but we contend that further research into retailer/shopping center networks could be worthwhile for the more enlightened centers and retailers.

Dennis *et al.* (2000b) have argued that the most successful shopping centers are those where 'active marketing' and 'pro-active management' feature. Bennett and Gabriel (1999) contended that market orientation is central to the rapid introduction of Knowledge Management in UK companies, pre-supposing and spreading customer information. Change-friendly enterprises are more likely to have extensive Knowledge Management systems than others are. The authors predict that a rapid uptake of Knowledge Management is likely for the most successful, marketing orientated shopping centers. There are substantial benefits to be gained from the customer knowledge database.

Finally, the authors accept that there have been many limitations in this small exploratory study. A true data mining system would be expected to work on a much larger dataset. The benefits predicted from a customer knowledge management system are purely speculative at this stage. Therefore, a more extensive pilot and research program is recommended. This could take the form of (1) a further questionnaire survey with more respondents and shopping centers; (2) A pilot scheme based on exchange of customer information for parking discount benefits (at a paid-for parking lot); and (3) A pilot 'personal agent' trial based on a smart card. This trial could run on shopper data at a single shopping center gathered by, for example, a parking lot membership scheme as outlined above. If this pilot were to achieve no more success than confirming the effects of cleanliness and rest rooms found in our exploratory survey (which was carried out at nominal cost), the center could expect a medium-term increase in rental income alone of US$2m. In our view, there is a clear case for the cost-effectiveness of further research in this area.

REFERENCES
Bennett R and **Gabriel H** (1999) 'Organizational factors and knowledge management within large marketing departments: an empirical study', *Journal of Knowledge Management,* 3 (3): 212-225.
Berry M J A and **Linoff G S** (1997) *Data Mining Techniques for Marketing, Sales and Customer Support,* Wiley.
Berry M J A and **Linoff G S** (2000) *Mastering Data Mining,* Wiley.
Boedeker M (1995) 'New type and traditional shoppers: a comparison of two major consumer groups', *International Journal of Retail and Distribution Management,* 23(3): 17-26.
Boedeker M and **Marjanen H** (1993) 'Choice orientation types and their shopping trips: an empirical study of shopping trips to the city center *vs.* to an edge-of-town retail park', *7ᵗʰ International Conference on Research in the Distributive Trades,* Stirling, University of Stirling, Institute for Retail Studies.
Bird J (1996) 'Data in knowledge out', *Management Today,* September: 70-72.
Davies R, Howard E and **Reynolds J** (1993) *The Shopping Centre Industry 1993: Thirty Years of Growth,* Oxford, Oxford Institute of Retail Management/British Council of Shopping Centres.
Dennis C E, Marsland D and **Cockett W A** (1999) 'Why do people shop where they do?' *Recent Advances in Retailing and Services Science, 6th International Conference*, Organizer: The European Institute of Retailing and Services Studies, Eindhoven, The Netherlands. Location: Puerto Rico, July 7-10.
Dennis C E, Marsland D and **Cockett W A** (2000a), 'Objects of desire: attraction and distance in shopping Centre choice', *International Journal of New Product Development and Innovation Management,* (2) 2: 43-60.
Dennis C E, Marsland D and **Cockett W A** (2000b), 'Can a shopping Centre be a brand?', *Proceedings of the 11ᵗʰ International EARCD Conference on Retail Innovation,* (CD-ROM) ESADE, Barcelona, July 13-14.
Dennis C E, Marsland D and **Cockett W** (forthcoming) 'Central Place Practice: shopping Centre attractiveness measures, hinterland boundaries and the UK retail hierarchy', *Journal of Retailing and Consumer Services*.
East R (1997) *Consumer Behaviour,* Hemel Hempstead, Prentice Hall.
Field C (1997) 'Data goes to market', *Computer Weekly,* 16 January: 44-45.

Guy C M (1994a) *The Retail Development Process: Location, Property and Planning,* London, Routledge.

Guy C M (1994b) 'Whatever happened to regional shopping centres', *Geography* 79 (4) 293-312.

Hackett P M W and **Foxall G R** (1994) 'A factor analytic study of consumers' location specific values: a traditional high street and a modern shopping mall', *Journal of Marketing Management,* 10: 163-178.

Hallsworth A (2000) 'Britain's loyalty cards - an unmanageable retail innovation?' *International Journal of New Product Development and Innovation Management,* 2 (2): 133-144.

Howard E (1995) *Partnerships in Shopping Centres,* Oxford, Oxford Institute of Retail Management.

Howard E (1997) 'The management of shopping centres: conflict or collaboration?' *The International Review of Retail, Distribution and Consumer Research* 7 (3): 261-286.

Jarratt D E (1996) 'A shopper typology for retail strategy development', *The International Review of Retail Distribution and Consumer Research,* 6 (2): 196-215.

Kinnear P R and **Gray C D** (1997) *SPSS for Windows*, Hove, Psychology.

McGoldrick P J and **Thompson M G** (1992a) 'The role of image in the attraction of the out of town Centre', *The International Review of Retail, Distribution and Consumer Research* 2 (1): 81-98.

McGoldrick P J and **Thompson M G** (1992b) *Regional Shopping Centres: Out-of-Town versus In-Town,* Aldershot, Avebury.

Marjanen H (1993) *Store Location Analysis and the Mystery of Consumer Spatial Behaviour: competition between downtown shopping areas and out-of-town shopping centres as a special case,* Turku (Finland), Turku School of Economics and Business Administration.

Oppewal H and **Timmermans H** (1999) 'Modeling consumer perception of public space in shopping centers', *Environment and Behaviour,* 31 (1): 45.

OXIRM (1999) *The Shopping Centre Industry: its Importance to the UK Economy,* Oxford, Oxford Institute of Retail Management/MCB.

Rao B P (2000) 'Improving retail effectiveness through technology: a survey of analytical tools for physical and on-line retailers', *Technology in Society,* 22: 111-122.

Rensselaer Polytechnic Institute (1999) *Data mining,* http://www.rpi.edu/~arunmk/dml.html.

Reynolds J (1993) 'The proliferation of the planned shopping Centre' in Bromley D F and Thomas C J (eds.), *Retail Change Contemporary Issues,* London, UCL.

Richards I, **Foster D** and **Morgan R** (1998) 'Brand Knowledge Management: growing brand equity, *Journal of Knowledge Management,* 2 (1): 47-54.

Rowley J (2001) 'Eight questions for customer knowledge management in e-business', *Proceedings of the e-Business Workshop,* IBM/de Montfort University.

Teresko J (1999) 'Information rich, knowledge poor?' *Industry Week,* 248 (3): 19.

Westbrook R A (1980) 'Intrapersonal affective influences upon consumer satisfaction', *Journal of Consumer Research,* 7 (June): 49-54.

Wiij K M (1994) *Knowledge Management: The Central Management Focus for Intelligent-Acting Organisations,* Arlington, USA, Schema.

Wiij K M (1998) 'Knowledge Management: an introduction and perspective', *Journal of Knowledge Management,* 1 (1): 6-14.

Worthington S (1999) 'A classic example of a misnomer: the loyalty card', *Journal of Targeting, Measurement and Analysis for Marketing,* 8 (3): 222-234.

Worthington S and **Hallsworth A** (1999) 'Cards in context - the comparative development of local loyalty schemes', *International Journal of Retail and Distribution Management,* 27 (10): 420-428.

Table I The 'top six' significant attributes for each segment, ranked in order of the coefficient of determination, R^2, associated with individual relative spend.

	R^2		R^2
FEMALES **(199 respondents: UK£68 per month)**		*MALES* **(88 respondents: UK£58 per month)**	
Cleanliness *	**0.075**	General layout	0.104
Nice place to spend time	0.063	Nice place to spend time	0.086
Availability good rest rooms	0.056	**Lighting ***	**0.085**
Friendly atmosphere	0.053	**Sheltered access ***	**0.081**
Selection of merchandise	0.051	Helpfulness of staff	0.069
Eating and drinking	0.048	**No undesirable characters ***	**0.067**
ABC1 **(168: UK£73)**		*C2DE* **(113: UK£53)**	
Nice place to spend time	0.156	Nice place to spend time	0.049
Lighting *	0.118	Cleanliness	0.044
Access by road *	**0.113**	Good for children	0.043
Friendly atmosphere	0.101	Quality of stores	0.038
General layout	0.101	General layout	0.037
Cleanliness	0.092	Availability good rest rooms	0.036
INCOME UK£20000 + **(101: UK£89)**		*INCOME UP TO UK£20000* **(81: UK£59)**	
Nice place to spend time	0.077	**Lively or exciting ***	**0.110**
General layout	0.069	General layout	0.095
Cleanliness	0.062	**Covered shopping ***	**0.093**
Availability good rest rooms	0.046	Cleanliness	0.088
Selection of merchandise	0.045	Selection of merchandise	0.084
Quality of the stores	0.043	Nice place to spend time	0.074
AGE UP TO 44 YEARS **(186: UK£65)**		*AGE 45 YEARS +* **(100: UK£65)**	
General layout	0.070	Nice place to spend time	0.074
Availability good rest rooms	0.069	Cleanliness	0.058
Selection of merchandise	0.039	General layout	0.053
Nice place to spend time	0.038	Availability good rest rooms	0.046
Lighting	0.035	Friendly atmosphere	0.042
Value for money	0.034	Eating and drinking	0.042
TRAVEL BY AUTO **(149: UK£81)**		*PUBLIC TRANSPORT* **(57: UK£60)**	
Nice place to spend time	0.079	Selection of merchandise	0.155
Covered shopping	0.072	Quality of the stores	0.131
General layout	0.069	**Shoppers nice people ***	**0.110**
Selection of merchandise	0.044	**Availability of seats ***	**0.080**
Choice of major stores	0.039	**Big shopping center ***	**0.080**
Eating and drinking	0.038	**Value for money ***	**0.076**
SERVICE IMPORTANCE **(74: UK£82)**		*SHOPS IMPORTANCE* **(213: UK£59)**	
General layout	0.104	Nice place to spend time	0.080
Relative travel distance	0.099	**Shoppers nice people ***	**0.067**
Cleanliness	0.078	**Quality of the stores ***	**0.065**
Availability good rest rooms	0.069	Friendly atmosphere	0.057
Nice place to spend time	0.059	**Lively or exciting ***	**0.056**
Good for children	0.057	Selection of merchandise	0.052

All listed attributes were significantly associated with individual relative spend at $p = 0.05$.

The number of respondents and the average monthly spend for each subgroup is indicated in parenthesis.

* Segments significantly different at $p = 0.05$ with respect to the association with spend of these attributes (combination of Monte Carlo and t-test, Dennis *et al.,* 1999a).

Table II(a) Shops importance motivation cluster

	Final cluster center 'Importance' scores
Variety of the stores	3.49
Quality of the stores *	**3.41**
Covered shopping	3.30
Access by public transport **	**3.14**

Table II(b) Service importance motivation cluster

	Final cluster center 'Importance' scores
Parking facilities **	**4.47**
Access by auto **	**4.29**
Cleanliness **	**4.22**
Availability of good rest rooms **	**4.01**
Value for money **	**3.99**
Helpfulness of the staff **	**3.96**

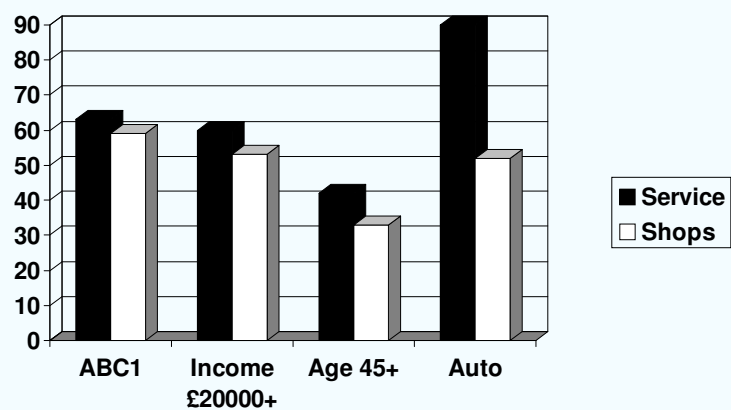**Differences between clusters 'Importance' scores significant at: * p = 0.05   ** p = 0.001.**
'Importance' scores are on the 1 to 5 scale, where 1 is 'no relevance' and 5 is 'extremely important'.
Only attributes above the scale mid-point (3.00) are listed, and each attribute is listed once only, in the cluster where most dominant.

Table III Models for shopper segments

	Constant	Attractiveness Coefficient	Distance Coefficient.	R^2	Significance p	Model number
Females	**28.3**	**0.63**	**-0.24**	**0.19**	**<0.0001**	**1**
Males	**21.1**	**0.49**	**0**	**0.09**	**<0.01**	**2**
ABC1	**19.0**	**0.72**	**-0.19**	**0.20**	**<0.01**	**3**
C2DE	**34.4**	**0.50**	**-0.24**	**0.13**	**<0.01**	**4**
Income UK£20000+	**28.6**	**0.62**	**-0.24**	**0.17**	**<0.01**	**5**
Income up to UK£19000	**27.0**	**0.58**	**-0.19**	**0.18**	**<0.05**	**6**
Age up to 44	**29.3**	**0.58**	**-0.23**	**0.16**	**<0.0001**	**7**
Age 45 +	**18.0**	**0.61**	**0**	**0.14**	**0.0001**	**8**
Auto	**32.8**	**0.53**	**-0.20**	**0.15**	**<0.01**	**9**
Public transport	**31.8**	**0.58**	**-0.22**	**0.19**	**<0.05**	**10**
'Shops motivation'	**19.4**	**0.70**	**-0.21**	**0.17**	**0.0001**	**11**
'Service motivation'	**39.6**	**0.54**	**-0.28**	**0.22**	**<0.01**	**12**
All respondents	**26.0**	**0.62**	**-0.20**	**0.16**	**0.0001**	**13**

Figure 1 Characteristics of the 'Service' and 'Shops' clusters

$$y = 0.0014x^2 + 1.11x$$
$$R^2 = 0.99$$
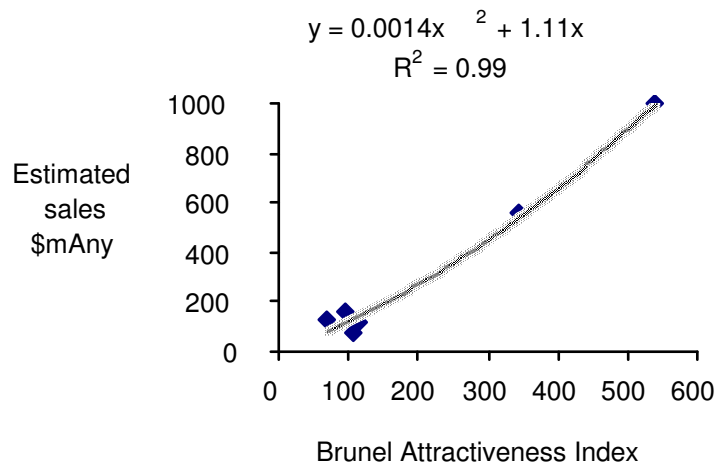
Estimated sales $mAny

Brunel Attractiveness Index

Figure 2 Estimated sales of shopping centres *vs.* the Brunel Attractiveness Index - polynomial plot forced through the origin