## Case Study

A financial services company markets several financial service products and wants to identify new segments within its existing customer base for a cross selling strategy. Through the application of data mining/CRM software<sup>1</sup> the company can easily identify those customers who have already purchased various of the company's products. As Figure 13a shows, the software produces a Venn diagram showing the number of customers who have purchased Account Type A, Type B and Type C. 83,048 people have account A only and their names and addresses are quickly displayed (these are obviously 'scrambled' for this exercise). This is a potential segment for a cross selling campaign concerning another of the company's products, say Account Type B.

Figure 13a



The company could target these customers immediately, with a promotional offer for account B. But this would undervalue customer and transactional data as an asset. In addition, the company would also want the highest return on marketing investment. As we have already mentioned, it is increasingly important to satisfy the strategic criterion for segmentation concerned with financial returns.

If the company mailed all 83,048 at say, £1 per piece, the spend would be £83,048 uniformly across the target, missing the opportunity to 'gravitate' spend towards more profitable groups. So instead the company could use the data mining/CRM software to interrogate existing customers who have both A and B accounts. Data mining can identify what makes these customers different from others and what makes them more or less likely to take both products.

<sup>&</sup>lt;sup>1</sup> This case has been adapted by the author from one provided by SmartFocus. The data mining/CRM software concerned is their VIPER product.

Taking the overlap area in the middle of the top of the Venn diagram the data mining/CRM software could be used to overlay geodemographic profiles for these customers (Figure 13b)



Figure 13b MOSAIC Profiles of customers who have purchased both A and B

The results show that the existing A&B account holding customers come mainly from areas classified as Blue collar owners, High income families, Suburban semi's, and Low rise council. This information could be vital to the identification of target segments for the cross-sell campaign. It can also be valuable for other campaigns, especially acquisition targeting, where customer and transactional data on existing customers can not always be used. Figure 13c extends this analysis and demonstrates how the same data mining/CRM software can fuse the transactional data with Experian's MSAIC geodemographic system and MapInfo's geographical information system (GIS) to show 'hot spots' geographically of where these potentially best target segments might be found. As we have seen, MOSAIC uses postcode and the data mining/CRM software has spatial analysis capabilities through a dynamic link to Mapinfo.

Figure 13c GIS Data Fusion



To further hone the characteristics of this 'best prospect' segment, the data mining/CRM software can be used to overlay other customer characteristics onto the map in order to redraw and filter this target segment further. Here, the first map has been filtered using Income over £35,000, Marital Status = Married, and Age Bands in 40 to 60. These are the characteristics that the same data mining/CRM software identified as being the ones possessed by the 'best' current customers of both Account Type A and B, according to their RFM profile. The data mining/CRM software extracts the names and addresses of customers with these same characteristics who currently have only purchased Account Type A as representing the best prospect segment for the cross selling campaign for Type B. This is done by merely selecting the 'hot spot zones' from the second map in Figure 13c and dragging them out and names and addresses are produced almost instantly, providing a contact list that satisfies the accessible criterion for segmentation. This target segment would presumably have a higher propensity to purchase both A and B products. Although the segment is composed of those who have currently only purchased A, it contains those who possess the characteristics of the best customers who have purchased both products.

However, the fullest benefit from existing customer data comes from looking at all of the attributes together. The easiest way to achieve this is via CHAID, which in this case is an integral component of the Data Mining/CRM software being used. Here, various customer and transactional attributes have been investigated to see which best explain what characterises of customers who have both A & B. A 'tree' structure represents different 'hot' and 'cold' 'branches' through the data. Each branch represents a different level of importance in explaining who the A & B customers are. Each attribute is assessed and the most important or 'significant' forms the first split. Taking the entire customer base in this instance, 26.44% of all customers have both A & B accounts.

Figure 13d



By following the 'hottest branch' the company can understand which characterises possessed by those customers who have purchased both A & B Account Types. Figure 13d shows these to be: Married and Male. For this group of customers the percentage with both A & B accounts rises to 65.11% compared with 26.44% of entire un-segmented base.

Further branches of the CHAID tree miht cascade down to even more segments based on whichever variables prove to be significant. Space prevents showing further stages here, but assume the analysis produced 60 target segments. Each of these would have significant and different characteristics. Targeting could be done on a 'test' basis in which a sample from each might be targeted and those with better response rates could then be targeted with the full 'roll-out' campaign. Also, each could be targeted with different treatments, according to whatever gender, age, marital status or geodemographic characteristic might underpin the 'creative'.

CHAID can use any variable within the dataset. The traditional segmentation variables could be totally irrelevant if those that prove to be more significant turn out to be star sign, biorhythm or whatever! In this way, the old rules have been overturned in favour of almost any combination of whatever personal details the organisation has on its databases.

It is interesting to compare this with the traditional use of market reports in which entire countries' consumption profiles might be based on samples of just 1,000 and profiled in terms of gender, age and some measure of social class (such as occupation, in UK). By contrast, the new data driven segmentation and targeting approaches might be based on details of several million customers, each with hundreds of transaction records over several years and which can be further overlaid by Geodemographics. Geodemographics, itself, would be based on the national census which, in the case of UK represents 60million people and the analysis of up to 80 census variables plus financial and other variables used to cluster neighbourhood groups. This is further overlaid with personalised details of 15 million individuals and the hundreds of product/service interests they have marked on lifestyle surveys.