

B05.1

Geographic Information Enabling Business Improvement

Keith Widdop, Business Metrics Analyst, Thomas Cook

Why this topic?

Thomas Cook has experience of merging customer and transaction data with market and geodemographic data. This has given it a very powerful understanding of customers, markets, store performance and competitor impact. Thomas Cook has brought together Customer Insight, Network Analysis, Market Research, Marketing Analysis, Loyalty Programme, Database and Geodemographic capability in an acknowledged Centre of Excellence available to all business units.

This is enabling Thomas Cook to widen the application of GI beyond the traditional areas of Marketing and store location analysis into optimising business efficiency.

"As a business - Thomas Cook UK Limited - our mission is to perfect the personal leisure experience. Underpinning this mission statement is a clear vision to build a close proactive relationship with our customers. Using the tools mentioned in this article, we are able to understand the needs and demands of our varied customer base and deliver appropriate operational initiatives. Even in the limited time that we have been proactively using the output of the analyses, we have seen a demonstrable improvement in our business. The analysis that we have worked with to date crosses all areas within our business: Product Offering, Foreign Exchange, Operations and Customer Delivery to name a few - and the possibilities are endless!" (Carla Stent Director of Operations, Thomas Cook UK)

Business improvement – Start with the customer in focus

Many retailers have the challenge of selling products and services through a variety of distribution channels to a geographically dispersed and demographically diverse customer base. It should not be surprising, therefore, that GI has an important role to play in business improvement.

Many of the big retailer brands have their own site location analysis tools and in-house marketing analysis, customer insight or GIS departments. Others will contract these functions out either to GIS suppliers or to marketing agencies. Typically, these retailers will have extensive data from POS and may also benefit from a loyalty card data. Smaller companies, however, may have little or no data and not know where to start.

This paper touches on some of the applications of GI in Retail. It will then go into some greater detail on the application of GI to business improvement. Whilst many companies use GI and GIS for site location and marketing purposes, how many are leveraging their investment in GI data and tools in the area of business improvement? It is suggested that there are far wider opportunities for business improvement, indeed, to confine GIS to marketing and site location analysis is to sub-optimize on your return on investment in GI. Given the need to focus more on profit post September 11, smart companies are changing the way they use GI from being almost exclusively the tool of specialist staff analysis to becoming a strategic tool to target competitive advantage.

The paper illustrates how Customer Insight including GI has been used to put the customer at the heart of the business at Thomas Cook UK.

Exploiting commercial opportunities on your customer base

It's all about getting closer to the customer. Who are your customers? What sort of people are they – age, affluence, family composition and lifestyle. Where do they live in relation to your store and competing stores? What are the most appropriate media to target them?

Geodemographic data can add huge richness to an organisation's understanding of its customers, yet many organisations perceive the cost of acquiring this data and the tools to visualise and manipulate it to be unaffordable. Where this is the case good value can still be obtained by starting with the new census statistics from ONS described as the most comprehensive survey of the UK population (see <http://www.statistics.gov.uk/census2001/default.asp> for England and Wales and <http://www.scrol.gov.uk/scrol/public/index.html> for Scotland). Indeed, in the context of Business Improvement, Thomas Cook is already starting to derive value from the ONS website in a number of areas.

Getting Value from Census 2001

The screenshot shows the 'Neighbourhood Statistics' section of the UK National Statistics website. The page is dated 20 July 2001. It includes a navigation menu on the left with options like 'Neighbourhood Statistics', 'Statistics by subject', and 'Interactive map'. The main content area is titled 'Welcome to Neighbourhood Statistics' and provides instructions on how to use the site to view, compare, or download statistics for local areas. It features a search box for 'Summary statistics for your area' and two buttons: 'GO' and 'Find detailed statistics by subject'. Below this, there is a section for 'Find detailed statistics by area name'. The right-hand sidebar contains 'Latest News' (New statistics: New 2001 Census data), 'Related links' (Census 2001, Small area population estimates project, Nomis - online labour market statistics), and 'External links' (renewal.net, Neighbourhood Renewal Unit, Social Exclusion Unit, CLIP).

The first step in using your own data in spatial analysis is to 'geocode' the data i.e. to append **Grid References** (Easting and Northing), which enable you to display the locations on maps and to compute distance and drive time from stores. It may be that postcodes need to be reformatted prior to this step.

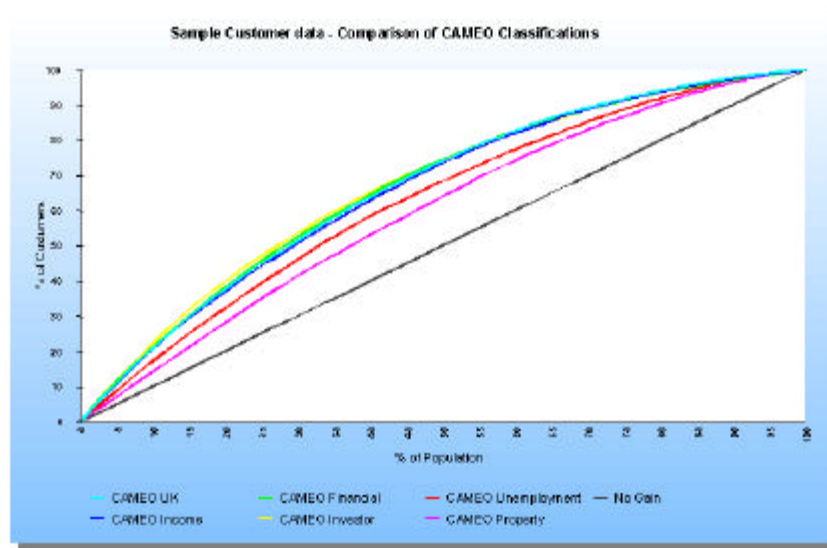
The next step is to profile your customers. Customer profiling is the analysis of a subset of your customer data based on a geodemographic classification used to determine the distribution of customers between the different groups or categories of that classification. A key benefit of profiling is that you can then identify which types of your customers tend to behave in a certain way, e.g. responding to promotions, purchasing particular products or services from your range, loyalty, credit risk etc. This type of analysis can therefore prove highly useful in identifying the most and least profitable customers and markets. This information can then be used to target direct mail, leaflet drops, siting of new branches, evaluation of existing branch performance and to optimize flying programmes.

There are a number of ways of interpreting profiles. The first and simplest is an **Absolute by Count Profile** . This presents a numeric count of the number of customers from your dataset, which fall into each of the classification categories. This provides an insight into the overall makeup of the profiled database.

The second profile is **Penetration by Count** , which gives an enhanced picture of your data because it takes into account the fact that different categories within each classification occur in the population in differing proportions. Even though a particular category may appear to be very significant in absolute terms, this may not be the case, particularly as different categories occur in different proportions in the total population. The resulting high count may therefore be deceiving. Similarly, a low count might be disguising the fact that you are reaching a high proportion of that category.

The profiles provide invaluable information for interpreting different subsets of a customer data set but it is important to select the most appropriate profile for the business question and product under consideration. This can be visualised using a Gains Chart. In simple terms the highest arc of a Gains Chart shows the best solution for segmenting customers and targeting new prospects.

Understanding the most appropriate profile to use



Business Intelligence (BI) tools can reveal customer insights that would otherwise remain hidden in your data. Different GI variables can be built into your BI application. This is not only the starting point for much spatial analysis but it also raises awareness of GI in the eyes of the management team.

The next step involves using this information to your advantage. This could include the following aspects:

Customer acquisition

"Lookalikes" for Direct Marketing who reflect the profile of your most profitable customers or those who typically buy a particular product can be identified and located using GIS.

Door drops can be planned based on the most relevant demographic categories ranking the postcodes by the degree of fit between the household types within the area and the total customer profile.

Cross-sell

An example of these insights might be cross-sell, where customers who buy one product may also buy one or more other products. This is visualised in a database query tool by a Venn Diagram.



In the above diagram if product 1 and 2 are basic products from your range (e.g. in the case of the travel industry say summer and winter holidays) and product 3 is an added value product (e.g. holiday car hire) you would be interested in the different characteristics of Group 4, who buy all three products. These customers are very valuable to your business. Analysis might show¹ that these customers, typically affluent homeowners, are three times more likely to hire a car than other demographic segments to the same destination and seven times more likely to hire a car than all segments travelling to all European destinations. Insights such as this can be used to assist productivity improvements and to target increased “share of wallet”.

Up sell

Most retail organisations face the problem of relatively high labour turnover. Young inexperienced staff may lack the life skills, confidence and product knowledge to recommend with confidence the most appropriate product, which meets their customer’s needs. Younger staff will often under-sell to a customer based on their own lifestyle and buying power. GI, even at a very basic level, can give inexperienced staff a good understanding of their local market. This may be as simple as “people from this part of town are older and more affluent and typically prefer a higher quality product; people from this part of town, however, are aspiring younger families on tight budgets, to whom value is more important”

Churn

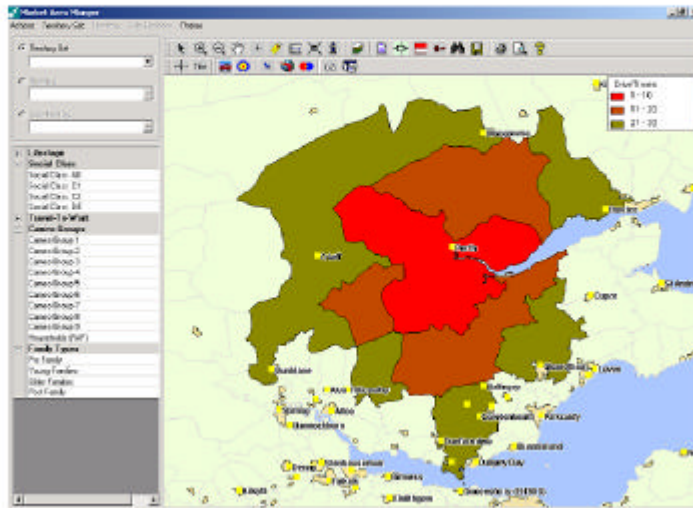
It can be important to understand churn (customer defections). Customers churn for many reasons – price or price perception, poor service, lack of convenience, product quality, competitor promotional activity, loyalty schemes, defection to new channels etc. GIS can visualize problem areas and GI can help identify problems contributing to churn.

¹ Fictitious illustration to protect commercial sensitivity

Drive Time Analysis

It is possible to analyse the distribution of customers from a store or retail centre, according to Distance or Drive Time bands. This is useful in establishing primary catchment areas around outlets, enabling you to identify how far the majority of your customers travel to reach you. The drive time “decay” of different types of centre can be very different.

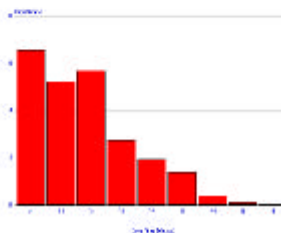
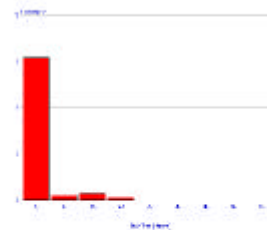
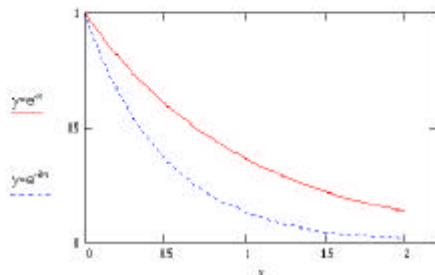
Drive Time Analysis helps you understand the size and profile of your catchment area



This enables you to target local marketing campaigns using the most appropriate media. To target a marketing campaign to a finer level it is possible to identify individual postcodes that share the characteristics of your customer database. Customers who have given mailing permission from these postcodes can then be mailed from your existing customer database or individuals selected from prospect lists in these postcodes.

Drive time “decay”

- Negative exponential equations $y = e^{-x}$
- Varies by type of store and type of retail centre
 - shopping malls
 - city centres
 - market towns
 - remote rural centres
 - suburbs

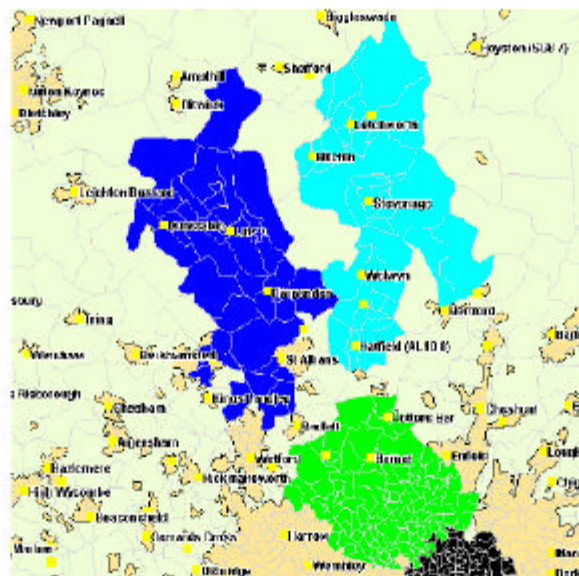


Retail Network

Most commercial organizations use GI to understand the impact of managing a property portfolio locating a new store, relocating an existing store or closing a store.

Many retailers and financial services institutions have store overlaps as the result of mergers and acquisitions, which may be hard to continue to justify. A GIS can help you understand the overlap of store catchment areas and reveal opportunities for consolidation or differentiation. Having identified the area of overlap it is relatively simple to compute the sales that would naturally migrate to the remaining outlet as opposed to being lost to competitors using a desktop local area gravity model.

Understanding catchment areas and network gaps



Fictitious representation of the catchment areas of 3 stores

Notwithstanding these overlaps, there may still be gaps in your distribution network. Mapping the primary and secondary catchment areas of surrounding stores can identify major gaps. The value of any gap can then be quantified by building a local market model based on Demand Estimates for your products at small area level. These are built based on market research, Family Expenditure Survey, TGI data, census data, your own customer data etc. Flows of expenditure may then be computed by a spatial interaction or gravity model. It is particularly important here to take commuter data into account in the case of conurbations and the South East; this is done based on Government Work Place Statistics².

² See <http://www.national-statistics.gov.uk/CCI/nscl.asp?ID=6680>

Gravity model

Spatial interaction of demand and supply

How do gravity models work?



- Demand
- Supply
- Interaction from residential zone i to supply zone j (service-related travel)
- Interaction from residential zone i to work zone j (work-related travel).
- Attractiveness of retail centre
- Distance or drive time between zone i and j
- "Micro" locational factors

Traditional retailers have reacted to the competitive threat from dotcom pureplays by creating their own new media channels through digital and broadcast television and the Internet. It is important to understand the demographic and geographic profile of customers using different channels to manage your property portfolio and your customer proposition in the context of multi- rather than single-channel. GI can then be used to understand areas of higher or lower new media penetration.

Performance Management

Spatial data can add significant value to managing performance against potential. Any performance variable e.g. total sales, product sales, performance ratios, stock control etc. can be analysed in the new breed of market analysis tools which give the user far greater control over the GIS than older black box solutions. Stores can then be benchmarked against their own peer group.

Line managers and functional heads often have preconceptions, which are based on anecdote, rather than fact. This may be because the necessary information to challenge these preconceptions has either never been available or has never been openly shared across the organisation. GI has a valuable role to play in educating managers and in changing perceptions. This particularly applies in the context of managing performance and can turn conventional wisdom on its head. It may be that a store which has been acknowledged to be performing well is in fact coasting well below its true market potential, whereas another store which has been considered a persistent under-performer may in fact be performing particularly well in a difficult market.

It's how you tell 'em that counts

People absorb more than 80 percent of what they learn through the sense of sight. This means if you show people something visual, they are far more likely to remember it. Simple persuasive graphics produced from a GIS have far more impact than detailed cross-tabs and experience shows that senior managers in particular relate well to heat maps. A picture is worth a thousand words but, nevertheless, the words that accompany it are all important. Being able to develop GI analysis to answer a business question, to summarise the findings succinctly in terms that management can understand and to draw conclusions and to make recommendations are what really counts.

New ways of working

To maximise your return on investment from GI it is essential that it be made accessible across the enterprise, rather than keeping it within the GIS department.

Thinking and working with spatial data will be new to many areas of the business and this may at first seem an unwelcome development to functions that have previously worked in "silos". GI enabled analysis may cause you to rethink many processes and will approach a problem from a different perspective than purely financial analysis.

GIS can produce some very powerful and persuasive graphics to support a business case for change, however the challenge of gaining Executive buy-in should not be underestimated. This is an important step that should not be overlooked; it takes time, but it will help in the long run in breaking down barriers.

The challenge is to integrate spatial data into the core business functions

- Consolidating GI with diverse data sources into a Single View data warehouse ("one version of the truth")
- Distributing spatial data across the organisation
- Offering new spatial-based insights to business functions that have previously operated in silos e.g. HR, Logistics, Product Development, Operations, Customer Relations

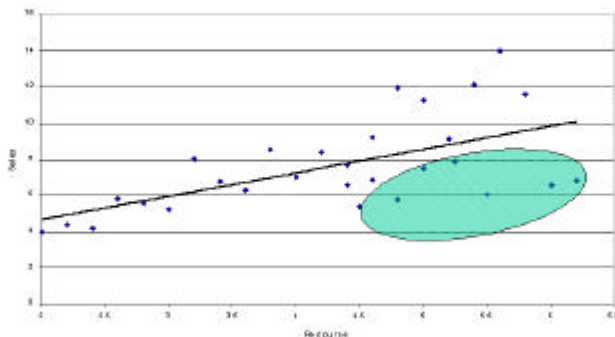
Some Business Improvement Techniques using GI

What follows is based on both personal experience and inspired by Fotheringham, Brunson and Charlton's excellent book "Geographically Weighted Regression"³.

In analysing performance it is assumed that the relationship between variables is constant over space. Experience suggests that this is not always the case and this may be a fundamental weakness in the approach taken by many analysts, who tend to look at global relationships.

The spatial component of the data may be useful in explaining variances in performance. This is important in business improvement as a business case, which ignores this aspect of variance will under or overstate the benefits arising from the change.

There is a need to understand
spatial variation in relationships



³ Fotheringham S., Brunson C. and Charlton M. 2000, Geographically Weighted Regression (The analysis of spatially varying relationships) ISBN 0-471-49616-2

In the above chart global regression analysis would indicate a linear relationship between the x and y variables, although the amount of variance explained by the model (R^2) is not particularly high. There is however one cluster of stores, which perform well below the benchmark and it would be important to understand the factors contributing to this. In this example excluding the cluster in question dramatically increases the amount of variance explained by the model.

Methodology

The methodology utilised is the CRISP DM, cross industry standard process - an industry- and tool-neutral Data Mining process model. Analysis has been performed in SPSS based on GI supplied by Eurodirect and GMAP.

The approach is to merge all customer, sales, competitor and market data into a single data warehouse. This can then be interrogated by the Customer Insight Team to address a business question. This is then merged with resource utilisation data and analysed within a statistical package. Geographically weighted regression models are then built to understand the cost saving and / or opportunity cost from resource reallocation. A limited trial is then put in place to test the hypothesis, which is then evaluated, refined and rolled out to the whole business.

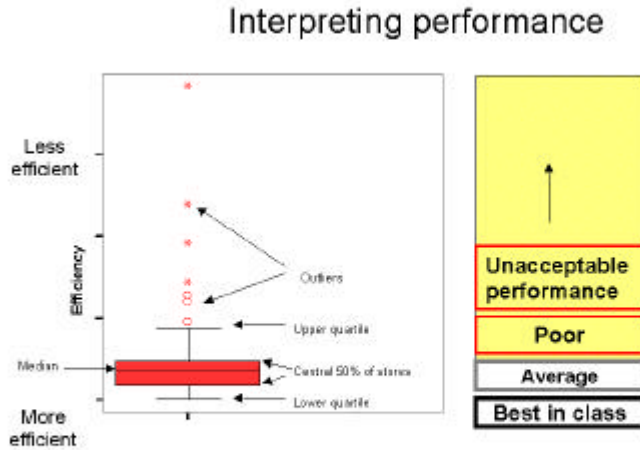
The CRISP process



Note: CRISP = cross industry standard process, an industry- and tool-neutral Data Mining process model.

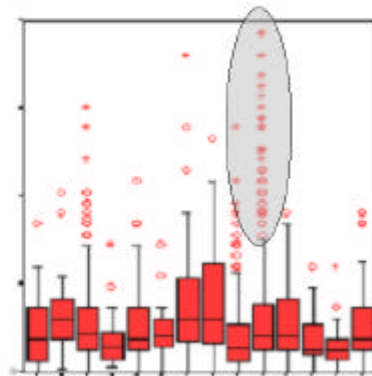
Boxplots⁴

A Boxplot is a useful technique, which is particularly useful for contrasting performance of different groups of stores trading in different geographic markets.



The following box plot illustrates how GI can be used to understand spatial variations in efficiency. The outliers relative to their peer group require particular focus whilst the upper (least efficient) quartile in a number of markets is a significant opportunity.

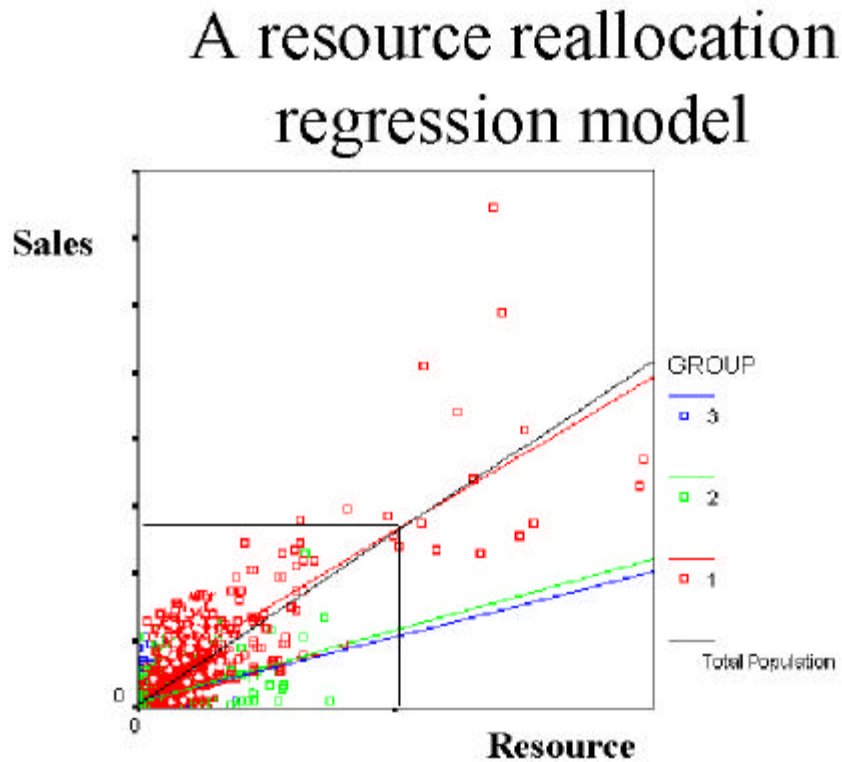
Using Box Plots to understand efficiencies



⁴ A boxplot, or box and whisker diagram, provides a simple graphical summary of a set of data. It shows a measure of central location (the median), two measures of dispersion (the range and inter-quartile range), the skewness (from the orientation of the median relative to the quartiles) and potential outliers (marked individually). Boxplots are especially useful when comparing two or more sets of data.

Regression Model

In the following chart it can be seen that successful repositioning of resource from Areas 2 and 3 would have a significant return on investment.



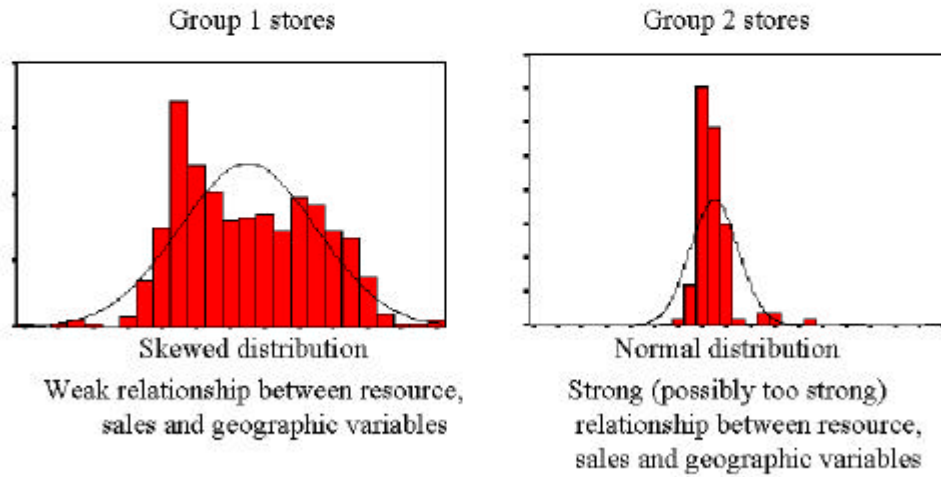
It may be appropriate to weight values in the analysis as differences in parameters may arise depending on the size of the areal units. The question here is – are there real differences in the parameters themselves, or do the differences result from differences in the size of areal units? An example of this would be to weight postal areas by population or by target households, so that variances in highly populated areas are more significant, and variances in less populated areas are less significant. Having computed performance data you can test whether the differences from the predicted model are statistically significant.

It may seem that these techniques require a high degree of statistical understanding and specialist software, however, as with GIS, this is not really the case. One can in fact get quite a good understanding of your data using a technique known as Spearman Rank Correlation, which would rank the performance variable relative to other variables such as standard region, type of retail centre, size of store etc and then look for correlations and anomalies between the rankings.

Using the output to drive change

Having built the regression model and visualised the output you then have the basis for building a business improvement model – a “what-if” model” – to compute the result of the change in performance. Some of the modelling output (as in the following chart) can be particularly persuasive in building the case for change with the senior management team:

Regression output



Note

The Geodemographic and GIS illustrations contained in this paper are based on Cameo provided by EuroDirect and MicroVision, which is a suite of software modules with a core component of mapping and reporting for quick and easy reporting of market metrics supplied by GMAP. It is acknowledged that a number of other suppliers have similar classifications and tools.

Reference

Fotheringham S., Brunson C. and Charlton M. 2000, Geographically Weighted Regression (The analysis of spatially varying relationships) ISBN 0-471-49616-2

The views expressed in this paper are personal views and not necessarily the views of Thomas Cook UK. The contents of this paper may not be duplicated or further distributed without permission of the author.