

# 10 Common Mistakes in Site Selection

WHITE PAPER:  
RETAIL INDUSTRY

Devon Wolfe • Managing Director • Strategy and Analytics Services for the Americas  
for Pitney Bowes MapInfo



## 10 Common Mistakes in Site Selection

2

### OVERVIEW

RETAILERS, RESTAURANTS AND OTHER BRICKS-AND-MORTAR COMPANIES OFTEN CREATE ESTIMATES OF THE SALES POTENTIAL FOR THEIR NEW LOCATIONS. HOWEVER, THE METHODS, DEGREE OF COMPLEXITY AND ACCURACY CAN VARY WIDELY, FROM BASIC "SEAT OF THE PANTS" ESTIMATES TO MORE ROBUST SPATIAL INTERACTION MODELS. WITH THE ADDED VALUE OF LOCATION INTELLIGENCE SOLUTIONS, SUCH AS THOSE OFFERED BY PITNEY BOWES MAPINFO®, MANY COMPANIES ARE SEEKING TO MAKE THEIR FORECASTING PROCESS MORE SCIENTIFIC AND ACCURATE, HIRING RESEARCHERS OR CONSULTANTS TO BUILD SITE MODELS.

THROUGH MANY YEARS OF ASSISTING COMPANIES IN THEIR SITE MODELLING PRACTICES, PITNEY BOWES MAPINFO RETAIL EXPERTS HAVE OBSERVED CERTAIN PRACTICES THAT SEEM TO PLAGUE BUSINESSES AGAIN AND AGAIN. SO HERE, IN NO PARTICULAR ORDER, ARE WHAT WE DEEM TO BE THE 10 MOST COMMON MISTAKES IN SITE MODELLING.

## 1. Assuming you can build a single model for all situations

Retailers today locate in a variety of situations: neighbourhood centres, out-of-town shopping centres, retail parks and traditional high streets, etc. Yet many retailers try to construct a “one-size fits all” model, or add a “situation weighting” to increase or decrease a centres attractiveness. This simply doesn’t work. The key to good retail modelling is to understand which situations have strong implications on performance, and then to develop sub-models that address why those situations are so important. An across-the-board “retail park weighting” may seem logical, but it makes more sense to understand how being in such a scheme affects the shape of the catchment, and then incorporate those findings into your model. In fact, a completely separate model will likely be warranted for a retail park location, since factors that impact the traffic flow within the park (positioning, co-tenancies, etc.) often have a high impact on an individual store’s performance.

## 2. Failing to understand the limitations of a model

Like all other social science models, site models attempt to explain highly complex and changing environments. Many factors that are not easily measurable (e.g., operations) affect store performance, while other factors (e.g., visibility ratings) can only be measured in an imperfect manner. It is important to note that retail models cannot directly model situations that aren’t present in a database of stores that already exist. A retail concept with an entirely suburban real estate strategy will be hard pressed to accurately model the first entry into the inner city. Likewise, any chain doing business in London will be sorely disappointed with a forecast from a model based on non-London sites. It is naive to assume that the model will be able to adjust for atypical situations.

## 3. Failing to verify store locations

While commercial geocoders become more accurate with every release, it is not uncommon to find that – due to factors such as address errors – the geocoder returns a location that is 1/4 mile, 1/2 mile or even several miles from the true location. Yet, many companies continue to make multi-million pound decisions from points entered into a geocoder without any additional verification. Of course, verification may be time-consuming, but simply asking your store employees to return a short note with a physical description of their location (e.g., the northeast corner of Oxford Street and Regent Street) and

comparing it to the geocode can make an enormous difference in the results you get from your model. And if you haven’t verified before you calibrate a model, you are likely to make the same mistakes again, only this time they will be more difficult to detect.

## 4. Using inadequate or undersized samples

In the process of building most site location models, it is critical to analyse existing customers. But not all customer files or store databases are equally suited to the process of model building. Site modellers for banks or membership warehouse clubs may have incredibly detailed customer data at their disposal because they track in detail every transaction from every customer. But many retailers have no data at all or they have partial data – a limited-period customer survey, a database of in-store returns, comment cards, mailing lists, etc. Some of these databases may in fact be representative of the overall base of customers, while others may be significantly biased.

To get an accurate and representative customer profile, you must use customer databases that are adequate in size and free from bias. If these aren’t available, a carefully designed survey instrument can be deployed to capture the data that is needed. Building a model from a biased database can be incredibly misleading and can lead to costly commercial property mistakes.

## 5. Using inappropriate variables

Many analysts are amazed to find some of the “nuggets” in their data – like a furniture retailer that learns that the number of households with five cars or more is highly correlated with sales. “This is significant!” they might say. “Let’s start a marketing campaign!” But a closer look may reveal that only 1.3% of EMEA households display such a characteristic. What does this mean? Remember that statistical techniques like correlation measure association, not causation. The relationship observed by the furniture dealer was probably a coincidence of some observations with a higher-than-average percentage of households with five or more cars having high sales levels as well. Variables that describe a very small portion of the population are much more likely to produce false correlations than variables that are well represented in the population. Further, variables need to pass a common sense check. If it doesn’t make sense that pet ownership and Rolls Royce ownership should be highly related, it’s probably not a good idea to use that relationship in a model.

## 10 Common Mistakes in Site Selection

4

### 6. Overfitting

While everyone knows there are many complex factors that influence sales performance, many people may not recognise that overfitting a model can cause more harm than good. For example, let's say a retailer has 70 stores and wishes to build a store-level regression model to forecast sales. The retailer does a regression analysis and builds a model with 22 variables. The model forecasts at +/- 7% accuracy. Sounds good, doesn't it? The problem is that the retailer doesn't realise the dangers of overfitting statistical models. Statisticians use a concept called parsimony to describe their aims in building a model. While it's a somewhat obscure word, all it means is this: the best modellers aim to maximise their predictive power with the minimum number of variables. Related to this is the fact that the fewer the observations in the data set, the fewer variables you can use before those variables start to explain relationships that aren't really there. In practice, this means that the 22-variable model may seem to do a good job of explaining the data from which it was drawn, but it will do poorly in new situations.

### 7. Inadequately measuring competition

Let's say you have five competitors in a five-mile ring around your site. Does that mean you shouldn't establish a site there? Not necessarily. Many models in the location industry look only at radius counts of competition without considering how store performance is affected by the positioning of competition in the catchment. A competitor who is better positioned in a large portion of the catchment is far more damaging than one who is situated on the periphery and affects only a small part of the catchment. Similarly, some retail concepts benefit from nearby competition; others don't. A model that hasn't considered issues like these hasn't adequately addressed competition.

### 8. Using the wrong base for the customer profile

Segmentation systems can be powerful tools in helping you understand your customer base. But should those tools be misused, customer profiling can be misleading. One common mistake retailers make is appending a segmentation code to a list of customers and then profiling with the national average as the base. In fact, the national

average is not an appropriate base for a brick-and-mortar retailer, since it makes sense that only people within a reasonable distance of a store are eligible to shop there. If you limit the base for the profile to only those people eligible to shop at the stores, you'll get a much more accurate customer profile.

### 9. Failing to recognise non-linear relationships

Every retailer knows that there is a one-to-one relationship between adding square footage to a store and increased sales – right? Not so. In fact, most retailers experience the opposite to be true. Similarly, relationships such as how sales decrease over distance (distance decay) tend to be non-linear as well. Despite that, countless models use a linear equation for square footage or other factors that are clearly not linear relationships. Using curve estimation techniques or non-linear regression can help you to fit the model more accurately to the data.

### 10. Using inexperienced modellers

Some modellers say they can build a database of store information and do an adequate model using accepted statistical techniques – without the benefit of having training in spatial concepts, retailing concepts, catchment patterns or markets. Don't believe it. While a modeller certainly doesn't need to visit every store and market in the model database, he or she should evaluate the retail landscape, site considerations and catchment patterns or, at the very least, review the model with someone with firsthand knowledge of these factors.

Building reliable and actionable site models is an entirely achievable goal, but it is one that requires care and diligence throughout the modelling process. We at Pitney Bowes MapInfo hope that – armed with the knowledge of common site modelling mistakes, you'll know what to watch out for, whether you are building your own model or hiring someone else to build it for you.

---

DEVON WOLFE IS MANAGING DIRECTOR OF STRATEGY AND ANALYTICS SERVICES FOR THE AMERICAS FOR PITNEY BOWES MAPINFO.

IN THIS CAPACITY, DEVON DIRECTS ALL ASPECTS OF THE NORTH AMERICAN STRATEGY AND ANALYTICS CONSULTING PRACTICE, ENCOMPASSING OVER 100 EMPLOYEES SERVING THE RETAIL, FINANCIAL, AND PUBLIC SECTOR VERTICALS. DEVON'S PRIMARY PROFESSIONAL PRACTICE HAS BEEN TO WORK WITH CLIENTS TO DESIGN CUTTING-EDGE RESEARCH TO ASSIST IN SITE AND MARKET STRATEGY.

EUROPE/UNITED KINGDOM

Minton Place  
Victoria Street, Windsor  
Berkshire SL4 1EG UK  
main: +44 (0)1753 848200  
fax: +44 (0)1753 621140  
europe@mapinfo.com  
www.mapinfo.co.uk

GERMANY

Kelsterbacher Str. 23  
65479, Raunheim, Germany  
main: +49 (0) 6142 203 400  
fax: +49 (0) 6142 203 444  
germany@mapinfo.com  
www.mapinfo.de

SPAIN

C/Ribera del Loira, 42  
Campo de las Naciones  
28042 Madrid, Spain  
main: +34 91 418 5083  
fax: +34 91 555 9957  
spain@mapinfo.com  
www.mapinfo.es

ITALY

Via di Vigna Murata,  
40 |00143, Roma, Italy  
main: +39 (0) 65 483 2224  
fax: +39 (0) 65 483 4035  
italy@mapinfo.com  
www.mapinfo.it