A Comparative Business Site-Location Feasibility Analysis using Geographic Information Systems and the Gravity Model

Jesse K. Pearson
Department of Resource Analysis, Saint Mary's University of Minnesota, Minneapolis, MN 55408.

Abstract

A site-location analysis was conducted using Geographic Information Systems (GIS) to locate where future Kowalski's Markets could be located in the seven county Twin Cities metropolitan area. Kowalski's Markets will here after be referred to as Kowalski's. Kowalski's provided their preferred demographic information from their two most successful store locations of which the potential market analysis was based on. A comparative analysis was completed between the final GIS analysis and the gravity model, where both site location analysis techniques were combined to provide results of the areas that have the highest market potential as well as site selection recommendations. There were four recommended site locations and four future potential areas to build new Kowalski's stores and expand their market.

Introduction

Kowalski's was founded in 1983 by renovating a Red Owl Country Store and differentiating itself from the discounters by adding value to the shopping experience. Kowalski's specializes in providing fresh bakery products, specialty meats, and organic and natural foods. To expand Kowalski's market area the traditional path has been to buy underperforming grocery stores and covert them into Kowalski's Markets, but in the past 7 years they have also built a couple of stores.

The goal of the site-location analysis is to provide Kowalski's with 4 or 5 possible site locations to develop future stores. Site selection is a crucial part in expanding any business as it is a very costly process if the store fails. One of the most important decisions a retailer can make is where to locate a retail outlet; a retail store can prosper or fail solely based on its location (Mendes and Themido, 2004). Kowalski's experienced a store closing when they had to close their Lakeville store in January of 2006 due to lower than predicted disposable income data which resulted in projected sales figures that were half of expectations. Kowalski's inferred the Lakeville store had poor visibility from the main roadway, which is an important factor in the retail industry.

In the first part of the analysis, GIS will be used to determine new potential store locations based on key demographic information from two of the most successful Kowalski's stores. The second part of the analysis will be to use the gravity model to determine the domestic market potential for each town

selected. The final GIS analysis will then be compared to the gravity model, which has been a proven model in the location analysis industry.

Classical Procedures Verses GIS

Classical location analysis procedures are still employed by many retailing companies for site-location. Only recently in the past 10 years has GIS had a major impact on the business industry. Of the classical procedures, the most frequently applied techniques for retailers are some form of the analog method, statistical modeling, and gravitational models (Rogers, 2004). Cleary, there is not one best method for site selection, and as such, this has paved the way for GIS, which includes many of these classical procedures embedded in GIS software.

   The use of GIS has enabled organizations to move away from a gut feeling to having factual information relating to a location (Bennison and Hernandez, 2000). It is no surprise then that companies who deploy GIS will have a competitive advantage over companies that do not.

Background

Kowalski's operates in the Twin Cities metropolitan area which is comprised of seven counties: Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington. For the purposes of this research, the seven county Twin Cities metropolitan area has been designated as the study area for this project.

Methods

Software Requirements

The software used to perform the tasks in the site-location analysis were ArcGIS 9.1 (ArcInfo), Spatial Analyst, and Microsoft Excel.

Data Acquisition and Manipulation

Background data for the seven county Twin Cities metropolitan area included census tracts, county boundaries, lakes, highways, streets, and shopping centers. The data was obtained from the Metro GIS DataFinder (www.datafinder.org). Customized demographic information from two of the most successful Kowalski's stores in 2006 was received from Kowalksi's Markets. The Twin Cities Census information was obtained from the 2000 U.S. census taken from the Metro GIS DataFinder in DBASE IV format.

   Kowalski's eight store locations were obtained from their website (www.kowalskis.com) where the address, city and zipcode were entered into an Excel spreadsheet. The same process was completed for the 21 Lunds and Byerly's competitor locations found at (www.lundsandbyerlys.com). After the store locations were recorded, the spreadsheet was converted into a database file in DBASE IV format. The database file was geocoded using the address attribute based off of the streets data as the reference layer. The final result was a point shapefile representing the precise geographic location of each store for Kowalski's, Lunds and Byerly's.

   Each entry in the shapefile contained information regarding the name, address, city and zip code of each store (Table 1). The Kowalski's and competitor's stores were then exported out into separate shapefiles (Figure 1).

Table 1. Example records of geocoded store shapefile.

Store Name	Address	City	Zip Code
Kowalski's	1261 Grand Ave.	St. Paul	55105
Kowalski's	5327 Lyndale Ave. S.	Minne-apolis	55419
Byerly's	7171 France Ave. S.	Edina	55435

Figure 1. Twin Cities Metropolitan Area with geocoded Kowalski's and competitor's stores.

In order to perform any analysis, the economic, social, and housing demographic census data, which was in DBASE IV format, had to be converted into geodemographic data through a join operation. Each of the three demographic datasets were joined to the Twin Cities metropolitan area census shapefile through a common attribute and then exported as a new shapefile.

The resulting polygon shapefile contained geodemographic census information including median family income, percent employed, median rent and population among many others (Figure 2).

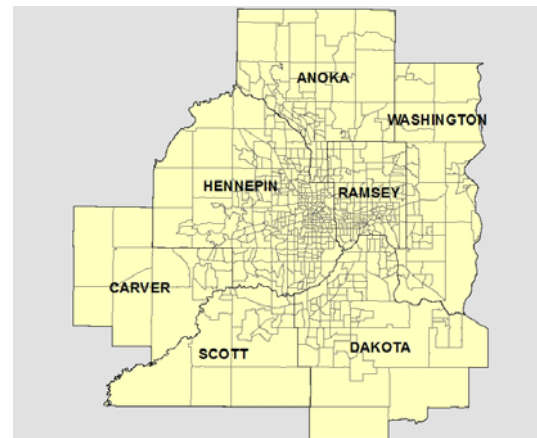A Thiessen (Voronoi) polygon was created to define individual areas of

Figure 2. Geodemographic Census Data by Census Tract for the Seven County Twin Cities Metropolitan Area.

influence around each store, or each store's trade area. The Thiessen polygon tool was used in this process. First, the store point layer was converted into a triangulated irregular network (TIN), which was then generated into perpendicular bisectors for each triangle edge. The locations at which the bisectors intersect determine the locations of the Thiessen polygon vertices (ESRI, 2005).

Next, Spatial Analyst was used to create a series of grid data sets. From the geodemographic census data layer, six polygons were created. These were then converted into grids representing median age, median family income, median household income, percent of population with bachelor's degree or higher, median mortgage and median housing values. These grids were created for the purpose of input data sets

to be used in a raster calculation to determine the geodemographic target market and to locate new Kowalski store locations. The input grid datasets can be viewed in Appendix A.

Lastly, the gravity model was used to determine the domestic market potential for each town. The following formula is given by:

$$Mx = \sum (P * a) / D$$

Where:

Mx = domestic market potential of town X
P = population of each region/city
a = mean per capita income for the region/city
D = distance between town X and the median centers of regional/city populations

Examining the gravity model market potential may be viewed as the gravitational attraction in terms of consumer purchasing power (Hammond and McCullagh, 1978). The distance variable was calculated between each town's major shopping centers by the actual drive time distance from the streets layer to get a more accurate measurement.

Analysis

The first phase of the analysis was to determine the key demographic information that Kowalski's looks for in a new store location. The first four criteria in Table 2 for the analysis are based from Kowalski's two most profitable stores. Criteria five and six were added to create a more specific potential market area. Criteria five and six were determined by taking the census

tracks that fell within one-mile of the two most profitable Kowalski's stores. This was achieved by intersecting a one-mile buffer of the two most profitable Kowalski's stores with the geodemographic census track data layer. Kowalski's determined the one-mile radius was where most of their retail food dollars originated. Additionally, the 2000 census information these criterion were based on have not been factored for inflation.

Table 2. Key Demographic Information to be used for Criteria in the Analysis.

To be a potential market area a census tract must have:
1. Median Household Income Greater than $45,000
2. Median Family Income Greater than $60,000
3. Bachelor's Degree or Higher Greater than 40%
4. Median Age Between 25-46 Years Old
5. Median Mortgage Greater than $1,050
6. Median House Value Greater than $139,000

The first calculation created a polygon layer and determined the potential markets for new stores by census tract, which is shown in Figure 3. This potential market area was created with the six aforementioned criteria using the SQL map calculator with the SQL syntax commands shown in Table 3.

To obtain a more detailed suitability analysis, the seven county Twin Cities metropolitan area census tract polygon layer was queried out by each criterion into six separate polygons. The selected query was then exported into a new polygon layer. This process was performed for all six criteria. Using Spatial Analyst, the six polygon layers were then converted into grids.

4

These grids now consisted of the desirable areas in the seven county Twin Cities metropolitan area for each set of criterion based on their geodemographic information.

Table 3. Grid query used in ArcGIS to identify potential market areas.

"Econ_tract.IHMED" >= 45000 AND
"Econ_tract.IFMED" >= 60000 AND
"Social_tract.GRADCOL" >= 40.00
AND "Housing_tract.HVMED" >=
139000 AND "Housing_tract.MTMED"
>= 1050 AND
"Metro_Area.MED_AGE" >= 25 AND
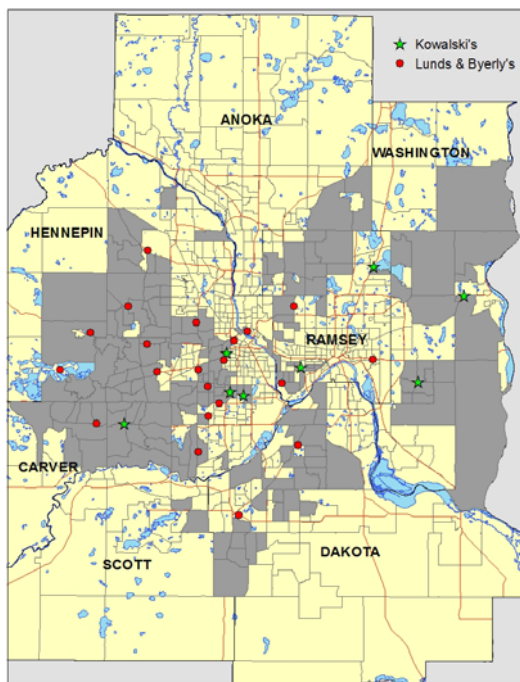"Metro_Area.MED_AGE" <= 46

Figure 3. Kowalski's Potential Market Areas by Census Tract in the seven county Twin Cities metro area with current Kowalski's and competitor store locations. The dark grey areas in the polygon layer represent the selected census tracks from the criteria.

When comparing multiple grids of data with different geodemographic information, they need to be reclassified into similar values. This was decided so that all the geodemographic information carried a similar value and was done using the Spatial Analyst "Reclassify" tool. All of the criteria except median age had a ranking value of one to nine. Median age had to be represented with a value of one for people between the range of 25-46 years old and a value of zero for null values.

To determine more specific potential market areas, the six grids were added together using the Spatial Analyst "Raster Calculator." This process created a suitability raster; this raster calculation identified the most suitable census tracts around the Twin Cities metropolitan area. Figure 4 displays the census tracts identified by the raster calculation which gives a more specified suitability analysis of where to further evaluate new Kowalski's stores.
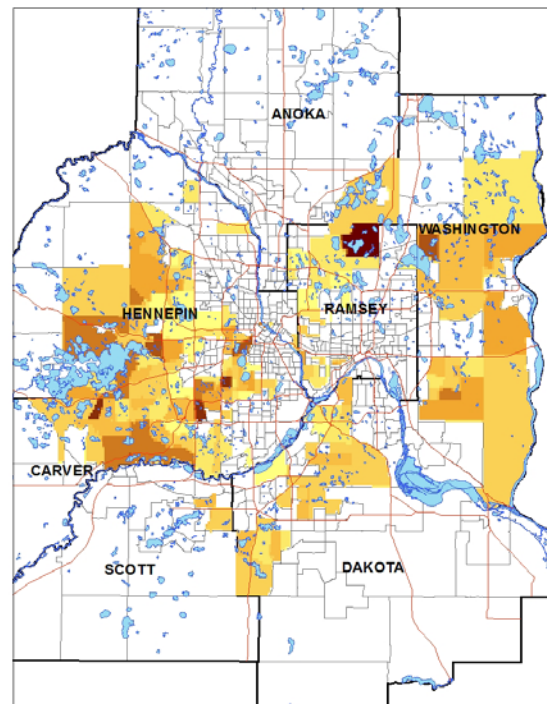

Figure 4. Raster calculation of Kowalski's Potential Market Areas by Census Tract in the seven county Twin Cities metro area. The darker areas represent a higher market potential area.

5

A Thiessen (Voronoi) polygon analysis was conducted from each of the Kowalski's and competitor's stores to determine the individual areas of influence around each store, or each store's trade area. This model assumes that customers choose a desired store considering a trade-off between proximity and store attractiveness. The proximity comparison in Figure 5 shows there are definite locations where current store trade areas have minimal coverage or overlap each other's trade areas.
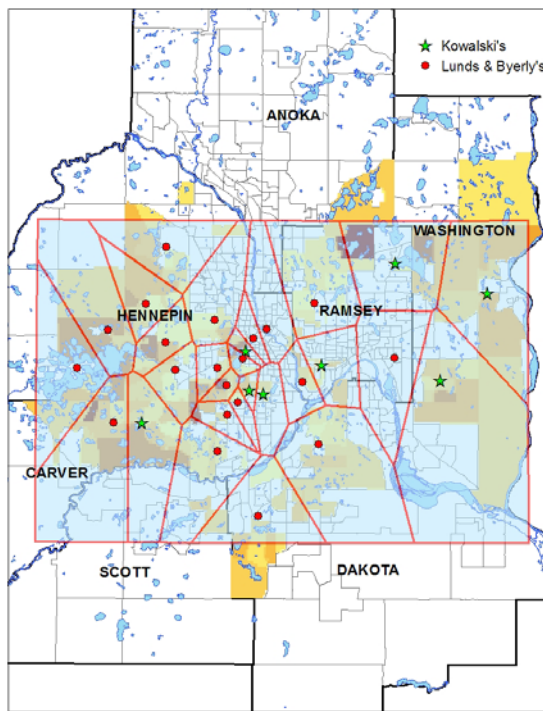

Figure 5. Thiessen (Voronoi) polygon proximity comparison of each Kowalksi's and competitor's stores.

The Thiessen (Voronoi) polygon analysis is a very easily applied model that can be valid in densely populated areas without geographical barriers in addition to relatively homogeneous demographic factors (Mendes and Themido, 2004).

After completion of the final raster calculation which gave the potential market areas and the Thiessen (Voronoi) polygon analysis which produced the competitor and current Kowalski store trade areas, the two methods were compared and analyzed from Figure 5 to find store trade areas with minimal coverage that lie within the potential market area. The results found there were eight locations that met all suitable criteria in the GIS site-location analysis that Kowalski's evaluates for building a successful store. Further analysis by Kowalski's at a finer level could interpret the results more closely. Kowalski's determined major success factors for their current stores on a micro level were high traffic counts in front of the store, appropriate store size for the area and good visibility from the main road. This paper does not go into a micro level analysis.

When the GIS analysis was completed, one of the factors to determine was that a new store had to be greater than three miles away of straight-line distance from a current store or competitor's store. Three miles was determined by Kowalski's as the trade area of an upscale food store. This was performed with the help of the Thiessen (Voronoi) polygon tool. This showed where the current store trade areas had minimal coverage, or overlapped each other's trade areas, and showed the weak penetration areas of a store.

Once this information was obtained, the distance was determined using the buffer tool on the current Kowalski's and competitor store's layers. This process found the potential store locations outside the three mile radius factor but still inside the potential market area.

There were four sites that had the strongest results from the GIS analysis and are ranked in order of suitability:

Shoreview/North Oaks, Minnetonka, SW Edina and Chaska. There were four other areas that had weaker analysis results but could be future possibilities since they each fit the model as well. These sites, also ranked in order of suitability are, Eagan/Mendota Heights, Maplewood/Oakdale, Lakeville, and Brooklyn Park (Figure 6).
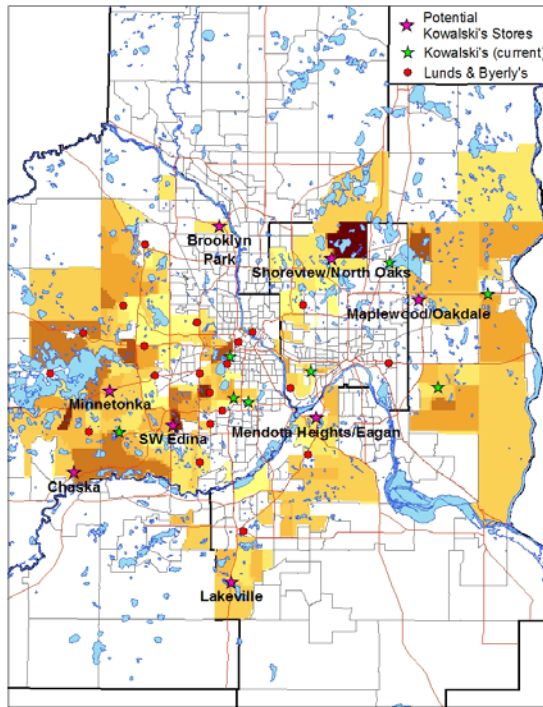

Figure 6. Kowalski's potential store locations in the seven county Twin Cities metro area.

Gravity Model

The gravity model is used on both ends of the retail industry. It is used for decision-making related to total shoppers and sales flow. Sales flow is best explained as the calculation of potential sales originating from an area. The gravity model is also used to provide insight related to shopping centers and department stores. More significantly, however, it is extensively used by convenience goods retailers, such as supermarkets and fuel retailers, for whom distance is a crucial factor (Rogers, 2004).

In this second phase of analysis, the gravity model was used to verify the domestic market potential for each town selected from the GIS analysis. The purpose of the gravity model in this analysis is to obtain another perspective of how attractive one area is to another. Using the formula previously mentioned, the domestic market potential was obtained by multiplying the population of the city by its mean per capita income and dividing by the distance to an adjacent city. The resulting figures for all the adjacent cities were added together to achieve the market potential for a city. The gravity model found that Minnetonka and Edina had the highest market potential. This can be attributed to the higher per capita income in both of those cities. The following table shows the relative difference in market potential between one city and another in its regional setting (Table 4).

Table 4. Results of the Gravity Model illustrating the domestic market potential for the prospective Kowalski's store locations.

City (Rank)	Market Potential
Minnetonka (1)	2,248,332,396
Edina (2)	1,988,535,798
Brooklyn Park (3)	1,426,958,249
Shoreview (4)	1,258,585,598
Maplewood (5)	1,137,074,533
Eagan (6)	974,224,704
Lakeville (7)	590,315,938
Chaska (8)	204,793,526

Comparative Analysis of GIS and the Gravity Model Results

The results from the gravity model showed that there was a slight variation between a cities' overall market

potential and the GIS analysis results with the exception of Chaska being at the bottom of the list. The distance from nearby cities and lower per capita income were significant factors explaining why Chaska had the lowest market potential between the potential store locations. One obvious disadvantage the gravity model has is it does not have the superb visual analysis GIS has. In its defense, the gravity model is not meant to be a graphics heavy model or to evaluate in-depth to the census tract or block group. Rather, it provides general trade areas for a city. Though the gravity model is still widely used across the industry, it is a crude technique that does not consider demographics factors. Mendes and Themido (2004) state gravity models have little sensitivity to demographic variations or market segmentation and therefore should not be used for specialized retailing.

Even though the gravity model is a crude technique it did, however, shed validation on the GIS results. The gravity model verified the strongest results from the GIS analysis and shed validation on areas with poor market potential that would not have shown in the GIS analysis results. The gravity model provides an advantage to show the domestic market potential quantitatively, even though the results may be more generalized than the GIS results.

Most retailers cannot rely solely on the standard gravity model to make million dollar store location decisions. The gravity model can be used to enhance decision-making by combining it with GIS. This is why more businesses are turning to the geographic advantage of GIS. The GIS analysis for Kowalski's only illustrated a few main

geographic techniques, but there are a plethora of tools available such as the Network or Business Analyst to obtain an even more precise store location. In addition, only six criteria were used in this site-location analysis using customized demographic information. If actual customer information was used the potential market areas would likely be more precise.

Results

The final GIS results in Figure 7 encapsulate all the geographic techniques used in this analysis. Figure 7 conveys potential market areas by census track ranking from high to low priority areas. It also illustrates each store's trade area using the Thiessen (Voronoi) polygon proximity comparison, and the three-mile buffer around current Kowalski's and competitor stores. The GIS analysis results suggest eight site locations would be suitable to build new Kowalski's stores based off these initial results.

Now that the gravity model's results have verified the market potential for each of the cities, the final suitability ranking was determined by a combination of the GIS and gravity model results. After taking into consideration both of the site location analysis results, the following are the final site recommendations for Kowalski's (Table 5).

One of the biggest notable changes was that Chaska dropped out of the recommended site locations and into the future possibilities area, which are considered as areas that have potential but did not have the highest analysis results. This was undoubtedly due to having the lowest market potential for the gravity model results. Minnetonka

became the number one recommended site location because of a combination of the strongest market potential from the gravity model results and the GIS analysis.

Table 5. Final recommended site locations and future possibilities based off the feasibility analysis for Kowalski's.

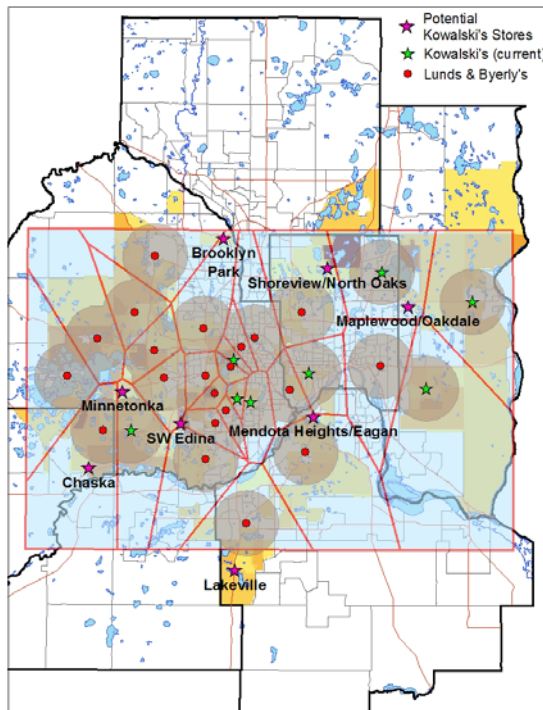| Final Recommended Site Locations by Rank: |
| --- |
| 1. Minnetonka |
| 2. Shoreview/North Oaks |
| 3. SW Edina |
| 4. Maplewood/Oakdale |
| Future Possibilities of Site Locations by Rank: |
| 5. Eagan/Mendota Heights |
| 6. Brooklyn Park |
| 7. Chaska |
| 8. Lakeville |

Figure 7. Kowalski's Potential Store Locations in the seven county Twin Cities metro area with the Thiessen (Voronoi) polygon analysis and a three-mile buffer around current Kowalski's and competitor stores.

From the GIS results, the Thiessen (Voronoi) polygon analysis illustrated Minnetonka did not fall within a current Kowalski's or competitor store's trade area. This makes Minnetonka a prime area for a new Kowalski's store. Also, the Maplewood/Oakdale site became a recommended site location not only because of strong market potential results, but because Kowalski's already has a dominant presence in the Twin Cities eastern metro area and would compliment the other three stores in the area.

Conclusions

This site-location analysis provides Kowalski's the recommended site locations at the census track level. In addition, it provides a traditional and technological process to identify areas that have the highest market potential as well as future possibilities based off of the criteria defined in this analysis. The analysis was based off of Kowalski's selected customer demographic attributes that were used to determine the potential market areas in the seven county Twin Cities metro area. Future analysis on a micro level would involve investigating the census tracks that were recommended and find a few optimal locations, that at a minimum, had excellent visibility from the main road, and peak traffic counts in front of the future store location.

In this study, there were three main analysis techniques that were used. The raster calculator derived the potential market areas in the seven county Twin Cities metro area. The Thiessen (Voronoi) polygon and buffer analysis determined each store's trade area, and the gravity model showed the domestic market potential for each city. For this site-location analysis to be

successful, it was very important to have multiple analysis techniques and to take each one into consideration.

As shown in this study, GIS technology is a very powerful and essential tool in making million-dollar store location decisions. With the help of GIS, Kowalski's could have avoided a store closing in Lakeville due to being a poor location. Although GIS is being used by many retailers to locate new stores, there is still a need to have industry knowledge to make the final decisions.

With the help of this study, Kowalski's has gained the geographic advantage through the power of GIS and the gravity model. Kowalski's now possesses the knowledge of what initial cities and potential market areas to target for new stores from this analysis. With this significant geographic and competitive advantage Kowalski's will have a decreased chance of a new store closing in the future and will see their market share increase with successful store openings.

References

Bennison, D. and T. Hernandez. 2000. The art and science or retail location decisions. *International Journal of Retail & Distribution Management*, Vol. 28 No. 8, pp. 357-367. Retrieved February 2, 2007, EBSCO*host* database.

Environmental Science Research Institute. 2005. ArcMap 9.1 [computer software]. Redlands, CA: ESRI.

Hammond, R. and P. McCullagh. 1978. *Quantitative Techniques in Geography,* (2nd Edition), Oxford University Press., pp. 303-307.

Mendes, A. B. and I.H. Themido. 2004. Multi-outlet retail site location assessment. *International Transactions in Operational Research.* 11, pp. 1-18. Retrieved February 2, 2007, EBSCO*host* database.

Rogers, D. 2004, January 26. A status report on location analysis. *DSNRetailing Today*, 14, Retrieved February 2, 2007, EBSCO*host* database.