

Using a Geographic Information System (GIS) to Visualize and Analyze Spatial Location in a Retail Environment

Robert J. Marros

Department of Resource Analysis, Saint Mary's University of Minnesota, Winona, MN 55987;

Keywords: Geographic Information Systems (GIS), Environmental Systems Research Institute (ESRI), target market, geocoding, demographics, US Census Bureau

Abstract

This paper illustrates how a Geographic Information System (GIS) is used to visualize the spatial distribution of Starbucks Coffee locations in Chicagoland. The location of Starbucks Coffee was compared to the data available by census tract from the US Census Bureau. In particular, the income by census tract and location were considered. By visualizing and analyzing the spatial locations it is possible to determine areas that are underserved and determine areas that seem to have a predominant demographic that Starbucks Corporation prefers.

Introduction

In order for a company to be successful, it must determine what its target market is and locate their store where they are going to be most visible and accessible to that target market. A GIS can be a very useful tool to help identify locations that are suitable, as well as markets that are underserved.

This project came about while I was looking at all of the Starbucks locations I would pass as I went to work or drove around Chicagoland. It is amazing to see how a corporation can grow so much that its identity seems to blend in with the landscape and the culture of a neighborhood. This paper will show the disparity of the distribution of over 300 Starbucks locations in the Chicagoland area.

Starbucks Corporation is the parent of Starbucks Coffee locations. Starbucks sells coffee. They have also greatly expanded their product line to capitalize on other consumers that fall into their demographic. Starbucks not only now sells a myriad of hot coffee beverages, they also sell frozen drinks such as frappuccino, which is a blended ice beverage, as well as tea, hot and cold and small food items with

some locations even selling sandwiches. Starbucks along with the food and beverages that they offer have also begun to offer other services such as T-Mobile Hotspot which provides high-speed internet access and they forge special deals with music artists to sell their CDs.

The location of a Starbucks Store is a carefully planned out execution of a marketing plan that is intended to provide Starbucks Corporation with a maximum return on their investment. The business of coffee in Chicagoland is very competitive with many different competitors emerging everyday to take a piece of the market. To be successful, Starbucks spends a lot of money ensuring they are targeting the right demographic in exactly the correct location in a neighborhood. For example, selecting a corner location or a location where there is a heavy amount of foot traffic.

The importance of location to Starbucks can be summed up in a quote from the National Real Estate Investor Magazine, "As it expanded in Chicago during the 1990s, Starbucks stuck to a proven strategy: building outlets where there were plenty of BMWs and young professionals. Its green and white logos

spread through the affluent neighborhoods along Lake Michigan and the inner loop. Then, the company followed the new money into gentrifying areas and into the suburban strip malls, where middle-income residents lived” (Luxenburg, 2005). Mapping the spatial locations of Starbucks will mirror what is described in this quote.

Background

Starbucks Corporation was incorporated in 1971 in Seattle, WA. Currently, the company has almost 7,000 locations in the United States and has stores in 34 countries with a location total of almost 9,500 stores employing more than 90,000 partners (employees). The reach of Starbucks is truly that of a ‘global’ corporation. Starbucks adds hundreds of new stores all over the world every single year.

Starbucks has also been recognized in several magazines including “Top Ten Most Admired Companies in America” by Fortune Magazine and “100 Best Companies to Work For” also by Fortune Magazine in 2005.

Software

For this project, a number of software packages were used. ESRI’s (Environmental Systems Research Institute) ArcInfo 9 and its extensions including Maplex, Spatial Analyst, 3D Analyst and StreetMap USA were used to complete data analysis, reporting, graphics creation, and geocoding. Microsoft Access and Notepad were used to store, manage, and manipulate address location data for Starbucks locations. Microsoft Excel was used to extrapolate income data from census tables. Lastly, Adobe Photoshop CS2 was utilized to create and modify some of the graphics that are included in this paper.

Preparation of Data

In the initial stages of this project I knew there were a number of data sources that would be necessary to obtain. In order to facilitate this process, I made a list of basic data that would be required and potential sources of data. The details of the data requirements and steps to organize and format are described in the following sections.

Data Collection/Assembly

Data for this project was collected from a number of sources. Census tract data was downloaded from ESRI’s website (www.esri.com) and provided by the United States Census Bureau. Income and housing data was compiled by and downloaded from Northeastern Illinois Planning Commission’s (NIPC) website (www.nipc.org). Starbucks location data was collected from their website. Lastly, base mapping data was collected from ESRI’s Data and Maps 2004 and NIPC; this data included basic data such as counties, zip codes, cities and community areas (Figure 1). All data was in ESRI’s shapefile format and when compiled, loaded into a Geodatabase for manageability. Additional information such as detailed census tract data was in Microsoft Excel format and was also converted to a .DBF and imported into a separate Geodatabase. A combination of two data sets from the aforementioned sources were used to produce a Composite Geocoding Service in order to accurately locate the over 300 Starbucks locations, which will be discussed in detail in the following section.

Geocoding

The most important aspect of compiling all of the Starbucks location data was to ensure that all locations were correctly identified.

This included making sure those locations had the proper information to be used for geocoding. In order to collect the data, I used the Starbucks website (www.starbucks.com). This was the most accurate and complete resource to find the location data of Starbucks locations. The process included searching for all of the Starbucks locations in Illinois, then copying each location's information from HTML into Microsoft Notepad in order to manipulate the data and format it for use in a GIS. After all 300-plus locations were copied into Microsoft Notepad; I imported the location data into Microsoft Access so I could more efficiently manipulate the data and use it in my geocoding service. Many of the addresses were not sufficient for geocoding and needed to be edited or modified.

Geocoding the location data was very important to the success of my analysis.

“Geocoding is the process of assigning a location, usually in the form of coordinate values (points), to an address by comparing the descriptive location elements in the address to those present in the reference material” (ESRI, help file, 2005).

In order to geocode the location data, I utilized three sources of data in a Composite Geocoding Service in ESRI's ArcGIS. The Composite Geocoding Service was constructed by creating three separate Address Locators. The Address Locators consisted of “US Streets with Zone” locator types and were file-based locators. To achieve the best results, it was imperative that I created a geocoding service that would not only look at primary street names, but also a service that would take into account alternative street names. Two sources of data were used in my Composite Geocoding Service, the US Census Bureau TIGER (Topologically Integrated Geographic

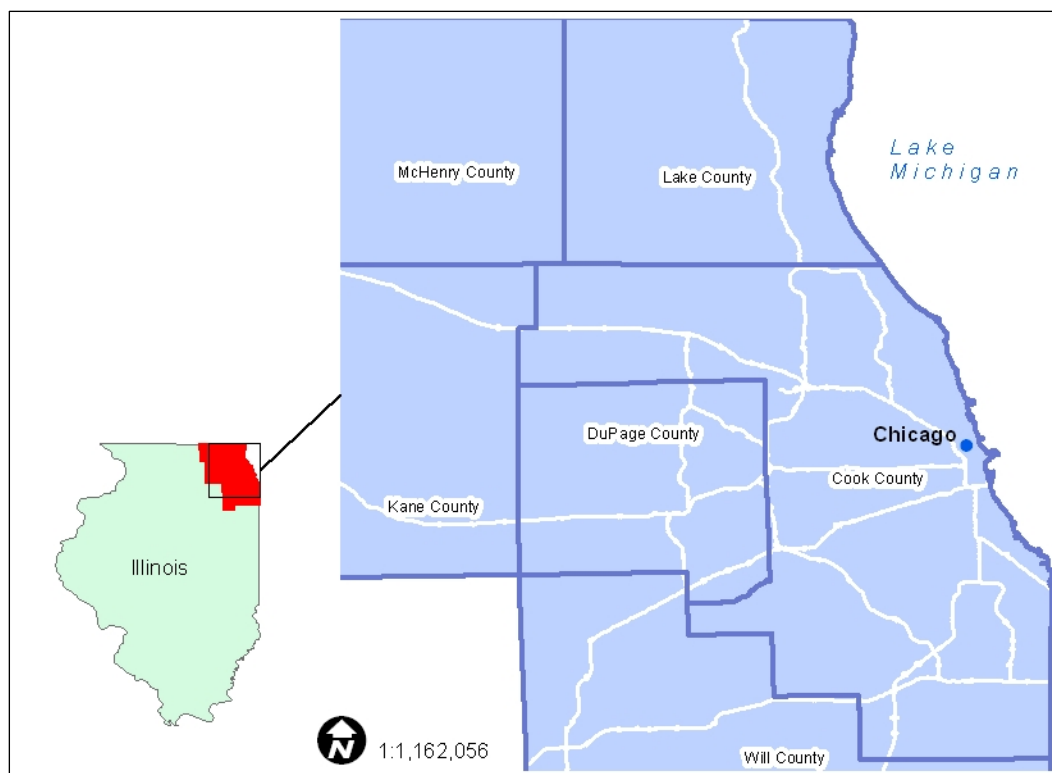


Figure 1: Chicagoland Overview Map

Encoding and Referencing system) Files as well as ESRI's StreetMap USA streets shapefile. While the TIGER data provided a good start for my geocoding, it did not have all of the information I needed to achieve the results I wanted. The ESRI Street Map USA data was essentially Census Bureau street line data enhanced with data from GDT, Inc (Geographic Data Technology, Inc. recently merged with Tele Atlas NV and is a digital map and geographic content provider); this provided a more comprehensive dataset with alternate names and other data enhancements. I determined that creating a Composite Geocoding Service would be most beneficial to ensure complete coverage of reference data. The StreetMap USA dataset provided the second and third source of data for my Composite Geocoding Service. This dataset provided the enhanced street name information as well as providing another source of primary and alternative street names.

Through trial and error, I arrived at the solution of creating the Composite Geocoding Service. The solution of the Composite Geocoding Service was needed because of the varying location address information from the Starbucks website as well as street names that varied from the TIGER dataset and StreetMap USA dataset. For example, the TIGER dataset would have a street attributed as Illinois Route 59 while the StreetMap USA dataset would have the same street attributed as Main Street. There were also many discrepancies within the datasets used for geocoding; for example, many state routes would lack the route number.

While this was the most accurate way to automatically geocode my data, I was still left with manual geocoding. I was able to automatically geocode about 75% of my data leaving about 70 locations that had to be "interactively" placed. I would interactively place locations by going

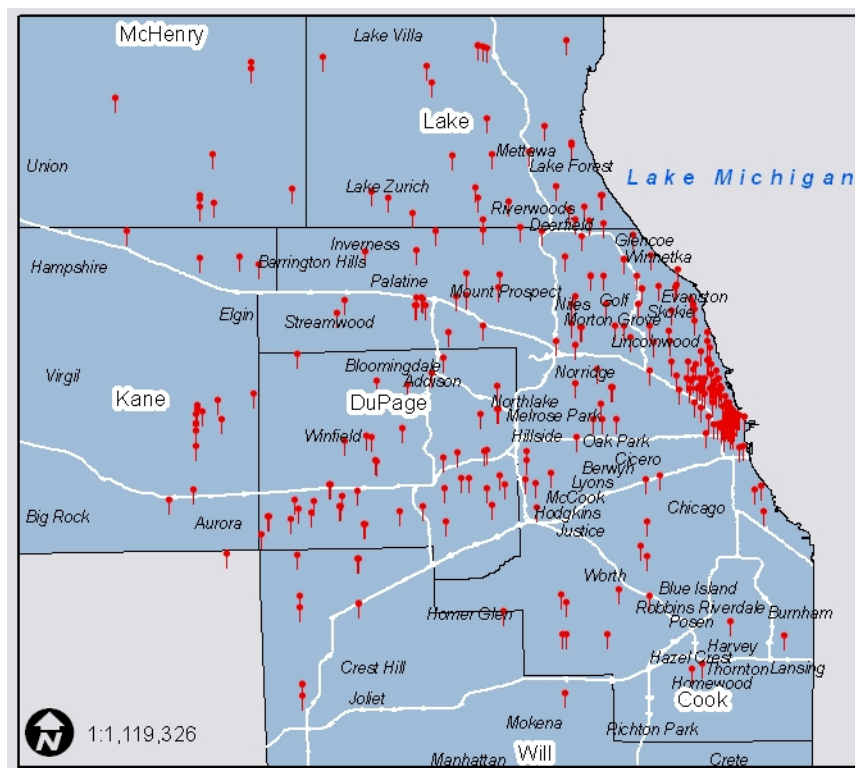


Figure 2: Geocoded Starbucks Locations

through each location and manipulating the address to at least find a place to start from. The data I had used for my street information often did not have full address range information, or in fact the street name was not correct. I utilized a number of sources to identify the correct location for each of the Starbucks that I had to interactively locate. To locate the most appropriate location, I used a combination of Starbucks website mapping, which utilizes Microsoft MapPoint technology, and could not always find the correct locations as well as Google Maps (www.maps.google.com). Google Maps provided another source of information regarding locations of Starbucks. With Google Maps, I could enter address information and retrieve a map of that location so I could more accurately geocode Starbucks. Between all of these sources, I could be certain that I found the correct location of Starbucks. My own knowledge of Starbucks and the geography

of Chicagoland was also useful in determining the correct location. To further support my geocoding effectiveness, I randomly selected locations that had been automatically placed and verified their placement (Figure 2).

I used default settings provided by ESRI's geocoding services and changed the spelling sensitivity to 60% to match a few more locations that would not automatically get matched. There were often spelling errors in the data procured from the Starbucks website, or in the data I used as a source for geocoding.

Methods

After the Starbucks locations were geocoded, I was able to begin analyzing data related to these locations. The most important data and the data in which my project is based, was the census tract data. I was most interested in the Median

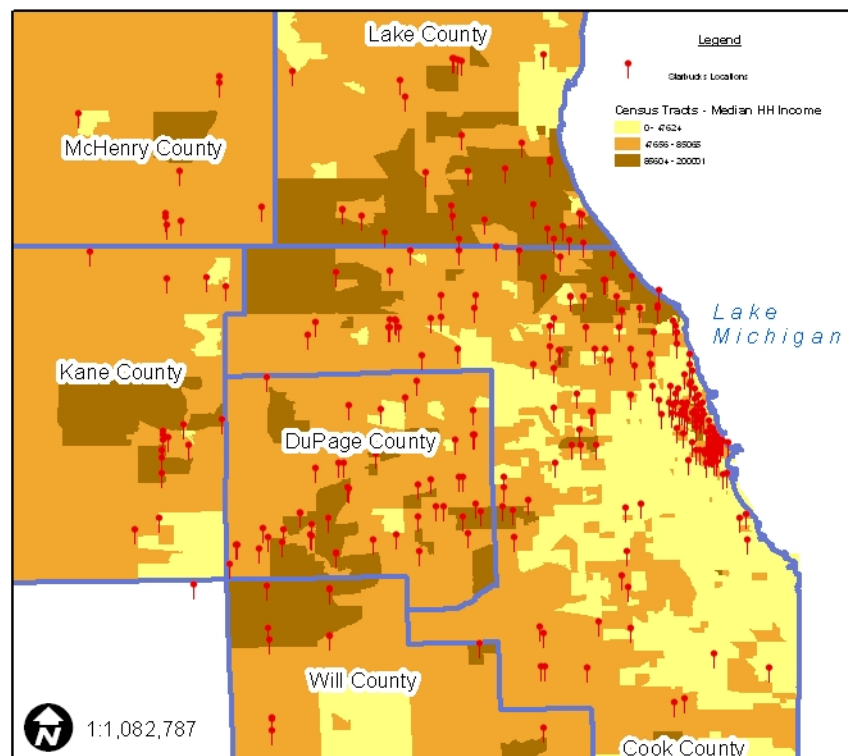
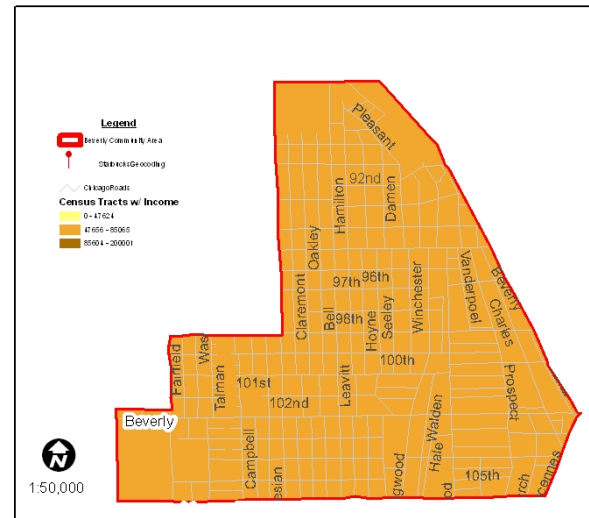


Figure 3: Starbucks Locations and Census Tract Income

Household Income data. According to American FactFinder Website, (<http://factfinder.census.gov>) “The median divides the income distribution into two equal parts: one-half of the cases falling below the median income and one-half above the median.” The information that I used was collected from the 2000 US Census and compiled by NIPC. Median Household Income data can be found in table SF-3 of the 2000 Census. SF-3 is the Summary File that includes data from the 2000 census long form that covers social, economic, and housing characteristics. The income data was in a table that had been loaded in my Geodatabase and I joined it to the matching census tract field in the census tract feature class. To determine what type of analysis I needed to complete, I developed a list of questions that could help determine what areas of Chicagoland were of importance and prominence to Starbuck’s Corporation.

Initial Analysis

Using GIS data, I was able to visualize the locations of Starbucks in Chicagoland. The simple action of looking at this data on the same map, leads to certain conclusions. It is apparent that in Chicagoland, there is a large part of Cook County that falls into the lower end of the three income groups that are used in this project. These same areas have the



NIPC, I chose two neighborhoods that are similar from a demographic standpoint. The two community areas analyzed were the Beverly Community Area (Figure 4) on Chicago's Southwest side and the Lincoln Park Community Area (Figure 5) on the North side of Chicago. I was interested in these areas because of the somewhat similar demographics that are characteristic of these neighborhoods. For each neighborhood, I selected the neighborhood from the Chicago Community Area data, and then selected all of the census tracts that fell within each. For the Beverly neighborhood, there were seven census tracts. The Lincoln Park neighborhood consisted of 20 census tracts. While Lincoln Park has almost three times as many census tracts than Beverly, both neighborhoods are about 3.2 sq. miles in area. After all of the census tracts in each neighborhood were selected, I was able to determine the average household income for each neighborhood. Beverly had an average median household income of \$69,000, while

Lincoln Park had an average median household income that was a bit higher at \$81,000. While Beverly has an average median household income that is \$12,000 lower than that of Lincoln Park, Beverly still has a higher average median household income than what was determined to be Starbucks's average target market household income of \$62,000. The income data is important because looking at this factor alone would lead you to believe that this is a prime market for a Starbucks's location. However, there actually are no Starbucks's within the Beverly neighborhood. In fact, the closest Starbucks's location to the Beverly neighborhood is almost 4 miles away, while the relatively similar Lincoln Park neighborhood is home to 11 Starbucks's locations.

Density Analysis

I utilized the Spatial Analyst Extension to visually represent the density of the

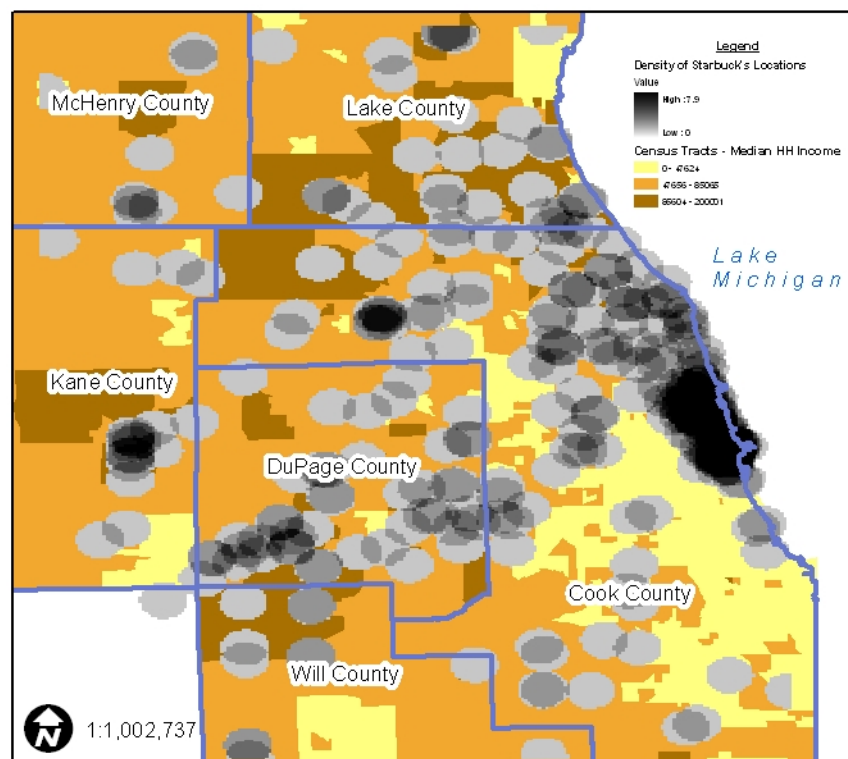


Figure 6: Starbuck's Location Density

locations of Starbucks (Figure 6). This figure reveals that there are Starbucks throughout Chicagoland, with the greatest density of Starbucks in the “Loop” or downtown business district. In the Loop, there are not many residents, but there are a large number of people that travel in and out of the Loop area on a daily basis. In fact, there is one census tract in the middle of the Loop that has 11 Starbucks in an area that is less than ½ square mile in size. The density of Starbucks locations is greatest near downtown Chicago, to the north of Chicago and in the western suburbs. This is also the same area that has the highest median household incomes. South of the Loop, there is a large area that has very few Starbucks. This also corresponds to the area in Chicagoland that has the lowest median income.

Visual Analysis

Using ESRI’s ArcScene and 3D Analyst, I was able to take the Starbucks’ geocoded locations and determine the number of

location’s in each census tract. This was an interesting analysis because I could then use a visual representation to see census tracts colored thematically, and the corresponding numbers of Starbucks that are located within each one. I then extruded the polygons based on the number of Starbucks’ locations. The result is a 3D graphic (Figure 7) representing census tracts and the number of Starbucks’s. This can be a helpful analysis when there are such a great number of census tracts that we are dealing with. By looking at the graphic, we see that there are many census tracts on the north and west areas of Chicagoland that are taller, or have higher extruded polygons (more Starbucks present). This analysis allows us to quickly visualize the census tracts and the number of Starbucks’ locations that are within their boundaries.

This type of density analysis shows the power of combining multiple sets of data and its spatial location in a GIS.

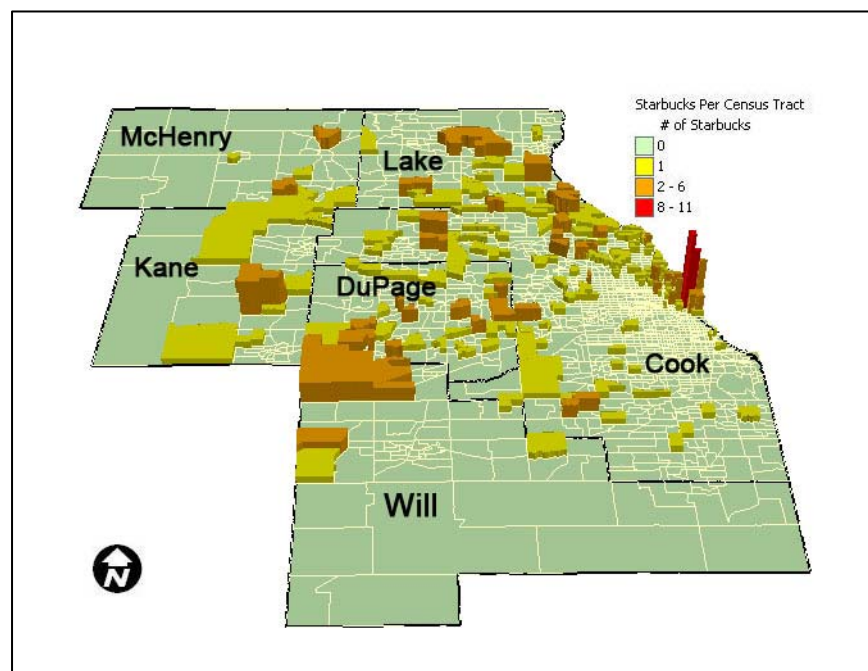


Figure 7: Starbucks Locations by Census Tract - 3D Graphic

Discussion

This project shows how GIS can be used to visually represent and analyze data. While I was able to effectively determine areas in Chicagoland that have Starbucks Coffee locations, I only looked at a small portion of the information that must be gathered before a company would make a logical decision about where to position a retail location. There are many other datasets that could be used to analyze a market. Very specific demographic data is available to more precisely determine the suitability of a market or a demographic. Selecting the correct data is paramount in employing a GIS to make important business decisions. Also, understanding the question you are attempting to answer and the data that you are using are critically important.

Many factors determine where to locate a business like Starbucks. This project does not intend to show locations that should be served by Starbucks. Chicagoland is a competitive market with many different obstacles that must be overcome to run a successful business. Driving throughout Chicagoland, it is hard not to notice over 300 Starbucks locations, let alone understand the spatial relationship that exists between them.

Results

When you look at the spatial distribution of Starbucks locations in Chicagoland, there seems to be Starbucks almost everywhere. There are certainly some places in the six-county area that can be considered rural and do not have the customer base to support something like Starbucks. However, it is apparent when we visualize the spatial location's of Starbucks in a GIS that on the South Side of Chicago there is a significant lack of Starbucks locations. This is also the same area of Chicago that has the lowest median incomes. It is logical that a

company would want to position itself in markets where it will be able to be most profitable. However, there are still many locations such as the Beverly area that would most likely be able to profitably support a Starbucks location.

There is also a higher density of Starbucks locations in the "Loop" in downtown Chicago. These locations cater primarily to business people who are downtown for work. These locations are unlike other markets served by Starbucks as they are in locations that are quite far from dense populations. However, their business relies on people walking to and from work. Another distinction in the locations of Starbucks is the locations that are in shopping centers or store locations such as those that are inside the store Target. These locations will primarily serve customers that are patronizing the host store as opposed to someone coming purely for Starbucks.

While it appears that there are a few census tracts on the South Side of Chicago that actually do have a Starbucks location, this is also misleading; these Starbucks locations actually exist through a partnership with basketball legend Earvin "Magic" Johnson's Development Corporation which works with corporation's such as Starbucks to develop locations within the inner-city (Luxenburg, 2005). Of the five Starbucks locations on the South Side of Chicago, three of them are actually a product of the partnership between Starbucks and the Johnson Development Corporation. The name of the partnership is Urban Coffee Opportunities (UCO). Two of these three locations are actually located near the University of Chicago campus. By locating Starbucks in this location, UCO will also be supported by college students that are a large portion of Starbucks target market.

Conclusion

In conclusion, I have determined that while data used in this project can give us an idea of the demographic that Starbucks is interested in; there are many other factors that must be analyzed in order to reach a definitive conclusion about the location of Starbucks's Coffee.

We have also determined that by visualizing the spatial location of Starbucks's Coffee that there is a correlation to be made between income and the carefully selected location of a Starbucks. This is not surprising when you look at the price of coffee at a Starbucks's; to be successful, they must position themselves where people have the disposable income to spend.

Acknowledgements

I would like to thank Dr. McConville for his guidance and support through the development of this project. I would also like to thank the instructors of the Saint Mary's University of Minnesota's Department of Resource Analysis who taught me the skills needed to develop my career in GIS.

Lastly, I would like to thank my classmates who truly made my experience learning GIS at SMU exciting and memorable.

References

- ESRI ArcGIS 9 Help file. 2005.
Environmental Systems Research Institute
Inc. (ESRI) Redlands, CA
Starbucks's Corporation Website. 2005.
12 May 2005, www.starbucks.com
Luxenburg, Stan. "Urban Magic." *National Real Estate Investor*. 1 May 2005
Northeastern Illinois Planning Commission
Website. 2005. 19 May 2005,
<www.nipc.org>

- American FactFinder Website. 2005. 14
June 2005, <<http://factfinder.census.gov>>
Google Maps Website. 2005. 25 May 2005,
<www.maps.google.com>