Market Penetration Research Analysis

Jemal Gulum | Data Analyst

Purpose of Analysis

The objective of this study is motivated by the request of a business client currently invested within the furniture retail market**.** With no definitive objective given to the client's goal of expanding within the specified retail market, the presented analysis reflects both an underlining assumption of either purchasing an existing retail furniture store or potentially opening a new location. The research is geared in presenting value in revealing best options to either cases.

Brief Description to the Analysis

Store Correlation Study

The analytic module allows one to assign a drive time trade area for each location in a sample store file based on the surrounding population density and urban definition in order to perform a consumer segmentation analysis. The results are then compared to a user defined potential store location that takes the same trade area definition mentioned. Its segmentation profile mix is correlated to all those most like it in the existing store network. The greater in similarity to the new location's consumer segmentation profile mix in relation to the existing network's locations, the higher the correlation rating is likely to be. Thus, the higher correlation rating, the more accurate the projection of potential sales is for the potential store.

Tools Used for the Analysis

Alteryx: Data Blending and Advance Analytics Engine (© 2014 Alteryx, Inc.)

Esri's GIS (Geographic Information Systems) Mapping Software (© 2014 Esri)

Research Procedure on Model & ETL Process

The module for this study stems from two input files. One of the two files contain data of current existing furniture retail stores network: address locations and monthly sales estimates. From this input of location data of the existing stores, we can create geographic coordinates for a geography-based analysis. This is also done for the given potential store location, being the second input file, so we may merge that potential site with the existing retail network.

Adjunct Model Integration: Potential 'Hot Spot' Distribution to Polygon Mapping.

In regards to acquiring potential locations, a twenty minute drive time radius for forty three AFHS (Ashley Home Furniture Stores) within the states of New York, New Jersey, Connecticut and Pennsylvania were calculated. The process began by viewing only census blocks not included in any of the twenty stores' radii. The model consisted of eight variables:

- Market Penetration (Census block received the market penetration of the closest store)
- Customer Count (Amount of AFHS customers located within each census block)
- Distance to closet AFHS store (Closer a polygon was to a store, the less 'points' were given)
- Distance to closets Bob's Furniture OR Raymour & Flanagan store (Closer the polygon to a store, the more 'points' given)
- Sales (Census blocks received the total sales of the closest store)
    - Scaled as followed:
        - 25-39 minutes away = .75 sales
        - 30-40 minutes away = .5 sales
        - 40-60 minutes away = .25 sales
        - > 60 minutes away = .1 sales
- 5 Year Projected Growth
- Furniture Demand
- Aggregate Furniture Expenditure

After each census block got a value associated with it, the best blocks in each market were expanded upon, using the same model and process as mentioned above. An important note to consider is that the below suggestions may not exactly coincide with the best blocks in the overall model.

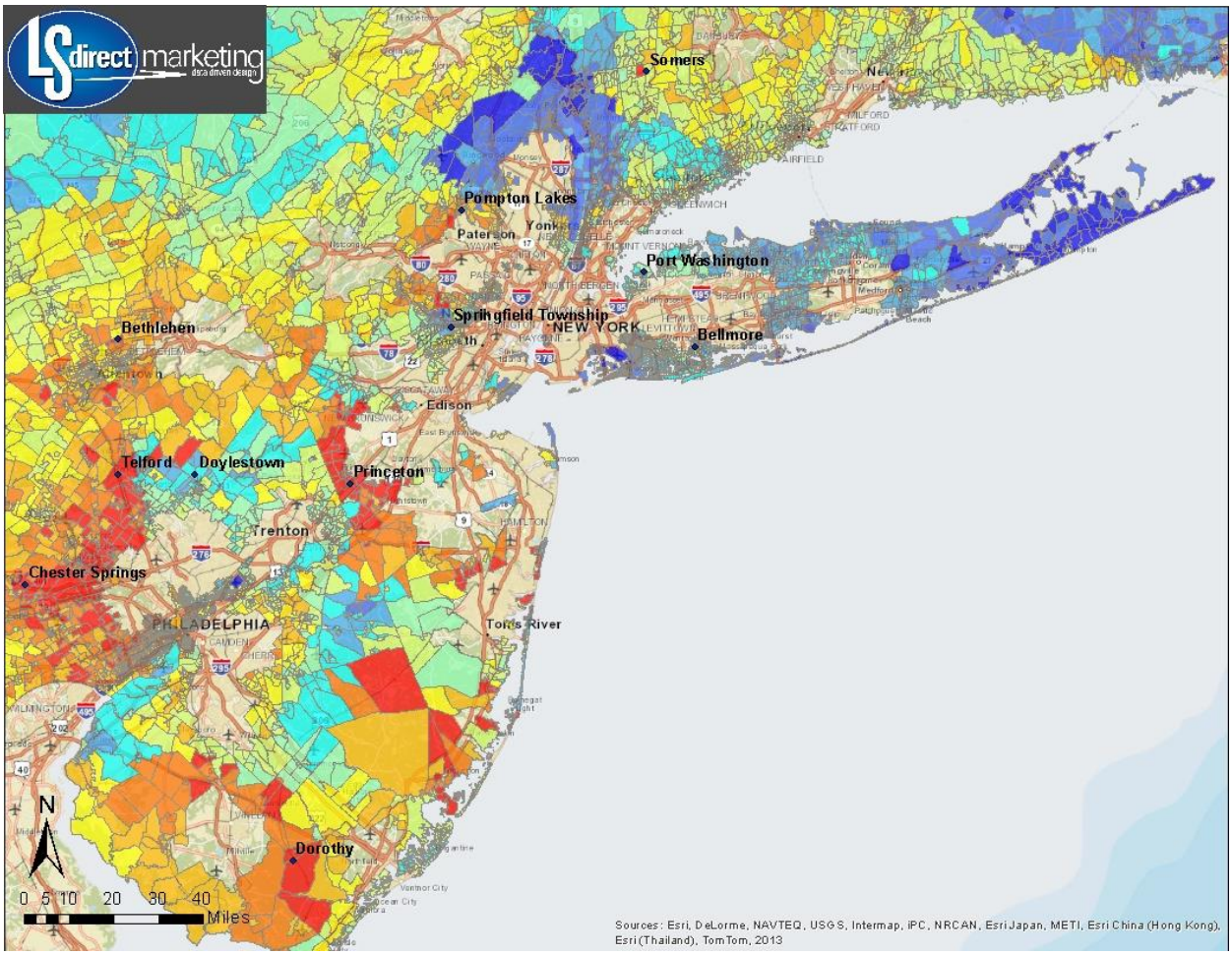

Fig 1 - Heat map of weighted census blocks, with red having the most "points", and blue being the worst.

This is due to the breakdown by market. Without this restriction, most blocks would end up in the Philadelphia area. After 'best' area determination, the store correlation model was used between exiting stores within the network and that new location. We are able to append total households and population to each trade area within a 7mile radius. We defined a trade area classification and drive time in order to establish an 'urbanicity' level for each retail store location within the network as well as for the potential new store location. In utilizing demographic analyses for building a correlation model, we retrieve drive time trade area for each mosaic group household distributions within the residential category by population.

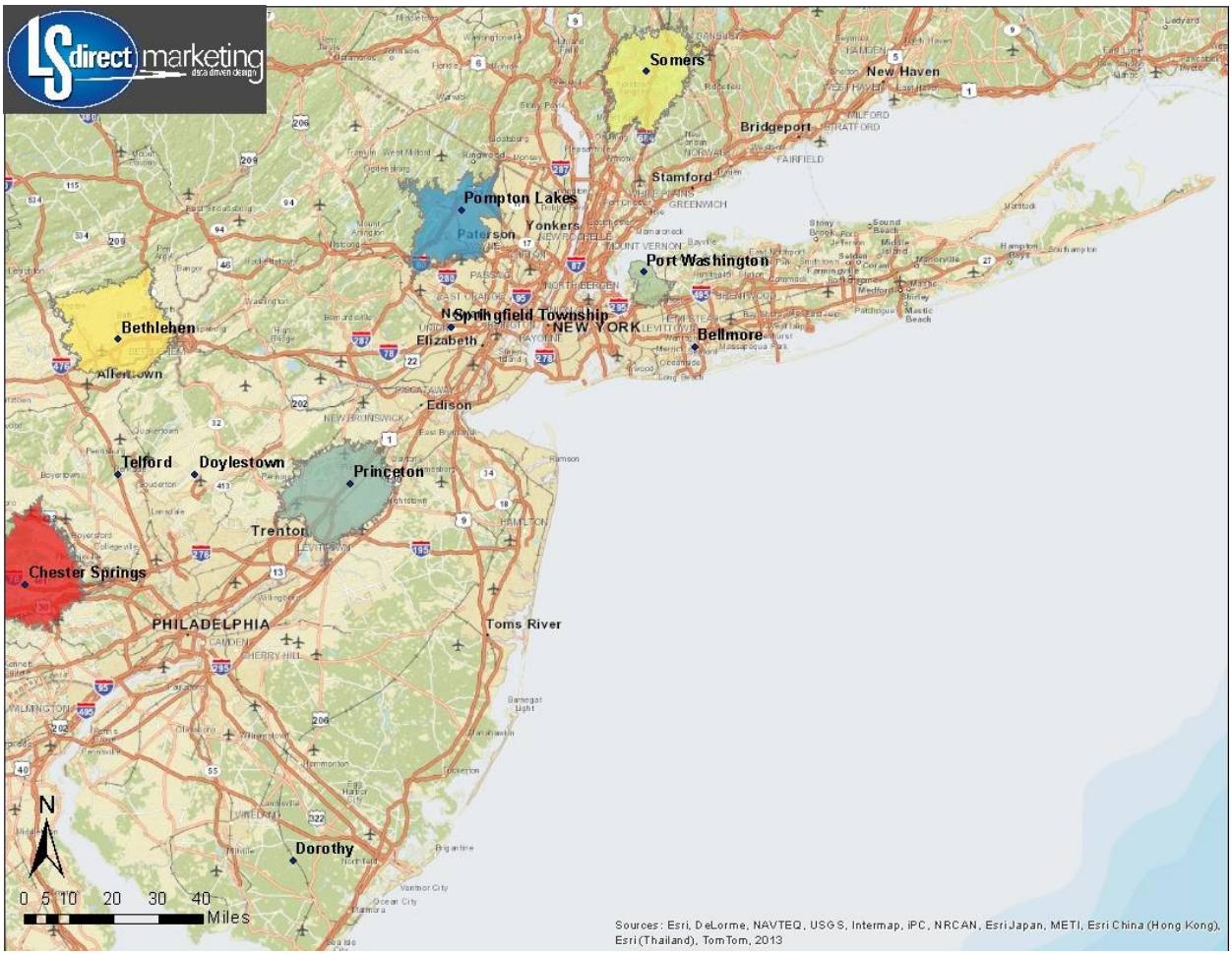

Fig 2 – The "best" areas to open a new store in each market, as determined by Fig 1.

(Other similar models were used in order to reveal the same distribution based on both Male and Female Adult Population. This was done to reflect either a positive or negative reinforcement to a recent study done by the University of Missouri proposing both gender and income differences to furniture purchasing behavior. [1])

After running several data manipulations to both input files, we ran a Spearman correlation for each existing location to prospect location. This data investigation method, being the core principle to the store correlation analysis, reveals the relationship between the variables created which, in effect, was done by transposing the value pairs that details the mosaic group breakdown by trade area. The Spearman's rank correlation assess how well an arbitrary monotonic function, an ever increasing or decreasing function, could describe the relationship between two variables, without making any other assumption about the particular nature (frequency distribution) of the relationship between the variables. [2]

In simply interpreting the Spearman's correlation, for the purpose of the main objective, the sign of the correlation indicates the direction of association between the two variables created within the model. One can interpret that the closer the Spearman correlation is to the positive value of 1 for each existing location in relation to the potential location, involving variables being represented by the segmentation profile mix to trade area by drive-time distance, the greater the probability that each potential store location will receive similar monthly sales numbers. In other words, the higher correlation rating, the more accurate the projection of potential sales is for the potential store.

- **Group and Segment Description**: Top Percentage of Mosaic Groups Exposed [3]

  Group A: *Power Elite* are considered the wealthiest households in the US, living in the most exclusive neighborhoods, and enjoying all that life has to offer.

  Group B: *Flourishing Families* are affluent, middle-aged families and couples earning prosperous incomes and living very comfortable as well as having an active lifestyle.

  Group C: *Booming with Confidence* reveal prosperous, established couples in their peak earning years living in suburban homes.

  Group D: *Suburban style* are middle-aged, ethnically-mixed suburban families and couples earning upscale incomes.

  Group E: *Thriving Boomers* are viewed as upper-middle-class baby boomer-age couples living comfortable lifestyles settled in town and exurban homes.

  Group H: *Middle-Class Melting Pot* stand as mid-scale, middle-aged (Baby Boomers) and established couples in suburban and fringe homes.

On the Final Analysis

- **New Locations**

Using both geo-location and transactional data from several other furniture retail networks a top correlation value comparison was done to further reveal the strength of both models. In revealing some potential stores, mined from the map in Fig 2, we note that the Projected Sales Totals are only considering additional sales on top of current expected **** sales. In other words, taking into account a cannibalization feed of potential sales that the potential store locations would normally generate for the existing locations are not being included.

AFHS Location	CNJ Correlated	Correlation	Proj. Sales Low	Proj. Sales High
Springfield Twp, NJ	Linden, NJ	0.84	$ 8.1 Mil	$ 14.2 Mil
Bellmore, NY	Freehold, NJ	0.81	$ 4.5 Mil	$ 7.9 Mil
Pompton Lakes, NJ	Freehold, NJ	0.91	$ 3.3 Mil	$ 5.8 Mil
Princeton, NJ	Eatontown, NJ	0.74	$ 3.8 Mil	$ 6.7 Mil

- **Existing Locations**

Further analysis was done using the same store correlation model as before for existing store locations within the Long Island network i.e. Brooklyn, Carle Place, Long Island, etc. Taking into consideration only Long Island stores with the most correlated demographic segmentation profile mix, projected sales were calculated. Calculations were based on the presumption that the Home Store furniture group would achieve the same amount of aggregate furniture penetration within Long Island as they did in their current store locations.

Location	CNJ Correlated	Correlation	Proj. Sales Low	Proj. Sales High
Carle Place, NY	Freehold, NJ	0.74	$ 9.2 Mil	$ 14.3 Mil
Patchogue, NY	Freehold, NJ	0.83	$ 5.1 Mil	$ 6.2 Mil
Elmhurst, NY	North Brunswick, NJ	0.76	$ 6.0 Mil	$ 13.2 Mil
Farmingdale, NY	Freehold, NJ	0.81	$ 6.3 Mil	$ 11.5 Mil
Flatbush, NY	Freehold, NJ	0.77	$ 5.4 Mil	$ 10.5 Mil
Bayridge, NY	North Brunswick, NJ	0.73	$ 4.6 Mil	$ 11.4 Mil

Stores showing the highest potential of sales are Carle Place and Elmhurst. It is essential to note for future performance modeling of existing locations the spread between high and low can result in becoming very wide in scale. This is especially accounted for higher urbanicity level trading areas where slight upticks in market penetration can equate to much higher sales volume. In demonstrating a straightforward example, a densely considered area where Total Expected Furniture Demand (2010 US Census data) is $150M, for each percent increase to market share there multiplies a $1.5M of additional sales. Due to this sensitivity to market share an increase of variance makes potential sales estimates less predictable than those of lesser density in trading areas.

- **Competitor Locations**

A similar analysis was also done for existing competing locations. For each competitor location an append was done for closest drivetime to an AFHS location, giving distances and a point estimate of potential sales at a 5% market share. We've considered a radius of a 20 minute drivetime as well as taking into account a projected growth over the next 5 years. We also note that within the sample data, we have seen the percentage of total sales that come from within a 20 minute range from 90% of the more urban locations down to 55% for the more rural locations. On the aggregate, for the metropolitan area, a 72% of sales coming from the 20 minute drivetime is observed, giving a safe estimate to say, for non-city areas, an extra 30-40% when projecting total sales.

Company	Distance (Miles)	AFHS Location	Potential Sales (approx 5% mkt share)
Raymour & Flanigan	5.59	New Rochelle, NY	$ 18,467,895
Raymour & Flanigan	13.84	Cherry Hill, NJ	$ 12,692,606
Raymour & Flanigan	9.82	Cherry Hill, NJ	$ 11,173,504
Bob's Furniture	6.76	New Rochelle, NY	$ 11,139,890
Raymour & Flanigan	10.44	Brooklyn (86th-Bayridge), NY	$ 9,136,994
Raymour & Flanigan	12.49	Fairless Hills, PA (Oxford Valley)	$ 7,174,947
Raymour & Flanigan	12.59	New Rochelle, NY	$ 6,652,001
Raymour & Flanigan	51.35	Danbury, CT	$ 6,225,177
Bob's Furniture	8.53	Carle Place, NY	$ 5,971,888
Raymour & Flanigan	13.61	Elmhurst, NY	$ 5,901,497
Raymour & Flanigan	29.99	Danbury, CT	$ 5,638,773
Bob's Furniture	29.84	Danbury, CT	$ 5,630,104
Raymour & Flanigan	30.2	Danbury, CT	$ 5,567,359

Fig 2 – Comparison of Competitor locations to AFHS Stores within the Metro Area.

