Big Data Analytics: A GIS approach on Market Segmentation

Yohn Jairo Parra, PhD Student[1], Sudha Yerramilli[2], PhD, Richard Aló[*] PhD
[1]Jackson State University, USA, *sudha.yerramilli@jsums.edu*, faculty mentor
[2]Jackson State University, USA, j00758975@students.jsums.edu

Abstract– This research applies an automated Geospatial tool that integrates Big Data Analytics to segment the markets based on geographic, demographic, psychographic and behavioural profiles of the region. Traditional market segmentation processes involve categorization of potential buyers into groups based on common characteristics of static facts such as age, gender, income, and geography. With the availability of data on consumer spending behaviours and psychographic profiles, Big Data Analytics has changed/influenced the approach of segmenting the markets by integrating this dynamic dimension to the process. ESRI's tapestry and consumer spending data (available at a fine geographic level (Block groups)) had identifies potential market segments by generating hot spots of customer profiles through spatial analytical models. Organizations/companies to gain insight into customer hotspots; hence identify communications to the targeted customer segments.

Keywords-- Big Data Analytics, Block Groups, Consumer Spending, Hot Spots, Market Segmentation, Spatial Analysis

I. INTRODUCTION

Toyota and Nissan, two major mass car manufacturers, continuously seek to increase market share by exploring spatial dynamics. Market research, at Nissan is sponsoring numerous GIS projects in four areas: market segmentation, distribution network, network operation and marketing [1].

All these areas are focused on strategies involving highly customized customer segmentation. GIS location-based solutions have been continuously evolving in understanding the spatial data and analyze the fundamental characteristics of the target market from a geo-demographic point of view.

Market segmentation is instrumental in understanding target customers and is defined as the categorization of potential buyers into groups based on common characteristics such as age, gender, income, and geography or other attributes relating to purchase or consumption behavior [2].

Marketing explains customer diversity and incorporates wide range of geo demographic, business and market potential data. This helps organizations/companies to gain insight into their customer hotspots and thus identify and locate their businesses in untapped/undeserved areas.

Customer's insights define the criteria for a profitable market and is possible to accurately evaluate potential in they trade area, analyzes market characteristics, and determines the success or failure of a site, such as new locations, new competitors, and site performance, as well as measure the impact of a new store and others. Big data analytics is crucial to understand and uncover hidden patterns in market trends and studying the ESRI tapestry data information allows categories new data that conduct filter and convert those elements into geographic layers of information. The utility of the conventional market segmentation, which is now evolving into Big Data Analytics (BDA), can be improved by incorporating the Big Data Tools (BDT) that support the vast amount of data like Hadoop, Spark and others. This will ensure that accurate results are obtained quickly, leading to better and more informed strategies since visual presentation of data and information enhances communication text and tables.

II. MARKET SEGMENTATION

Segmenting a market allow marketing and sales companies to focus on the subset of prospects that are "most likely" to purchase. The marketing concept calls for understanding customers and satisfying their needs by treating them different.

Market segmentation is defined by McDonald and Wilson as a process for determining the value propositions to meet the needs of the customer groups (segments), communicating these value propositions to all those people in the organization and playing an appropriate part in delivering these value propositions [3].

The importance to have a target audience from a spatial perspective is evident, since good segmentation contribute to a full understanding of the market, the ability to predict behavior with great precision depends how companies advance detecting and exploiting new market opportunities [4].

Traditional marketing used to work with small data sets with limited analytic platforms and implementation capacity. Big data analytics on market segmentation primarily describes data sets that are so large, unstructured and complex that require advance and unique technologies to store, manage, analyse, and visualize [5].

High performance firms use more advanced (BDA) Big data analytics in the web 3.0 era or prescriptive era, BDA impart companies with methods to reinforce products consumers are more willing to buy. Therefore, understanding the marketing environment and consumer's expectations is challenging because they change rapidly.

Value prepositions refer to the decision- making process of

---

[*] Richard.a.alo@jsums.edu, faculty mentor, Jackson State University

deciding what the offering to the customer is to be and what value the customer will receive. Alternative approaches considered have included size of purchase, customer characteristics, product attributes, benefits sought, service quality, buying behavior and, more recently, propensity to switch suppliers.

There are two phases to market segmentation. These are developing segments and prioritizing and selecting segments.

Developing segments covers the essential steps like customers, competitors as well as products and services to develop a segmented structure for a market. In prioritizing and selecting segments the process requires selecting which of the concluding segments fix in the business.

The purpose for segmenting a market into Big Data Analytics is to allow specific segments like psychographics segment provide the means of target groups by which psychographic aspects are isolated and reach particular behaviors that can take the data and use Big Data to find new micro markets.

III. BACKGROUND OF THE PROBLEM

Toyota had always striven for "every customer is the only customer." It further meant delivering cars at an affordable price with perfect timing. This ambitious goal had seemed nearly elusive after the Second World War, since most people in japan could not afford a car even at cost. In addition, the country's labor productivity was only one-eighth of that United States. In essence, Toyota was challenge to cut cost dramatically, but without the scale of economies. It needed an entirely new source of economies to satisfy customers with variety, quality, and timeliness, all at reasonable price.

The Toyota Production System evolved as Toyota's answer to this challenge, and served as a common frame of reference among all its employees. Toyota Company needs to find the right model to satisfy customers with all the variables that are surrounded and apply the best strategy to identify communicate to the targeted customer segments.

Combining data from traditional market segmentation and using big data tools like GIS and Hadoop, companies can look at the complete picture and get involved in the analysis of new variables.

IV. METHODOLOGY

Toyota Motor Corporation (TMC) is a widely-recognized car manufacturer around the world. Toyota has come up with a variety of models to tend to the needs of Jackson Mississippi like Corolla, Camry and Hilux just to name a few. This study focused in building customer profile that integrates Big Data Analytics to segment the markets of Toyota Company based on Geographic, demographic, psychographic and behavioral profiles of the region. Without a map showing the customer profile using Big Data tools, Cluster Analysis and Hot Spot Analysis cannot determine the hidden patterns.

The customer profiling for Toyota Car Industry was built at

a ZIP code level from the Tapestry data. Of the 12 life mode groups, L2 category (Upscale Avenues) with segment codes 11 (Pacific Heights), matches the Toyota customer profiles that suits their lifestyle and behavioral patterns. These locations of target customers are derived from the tapestry data for the Jackson, Mississippi region as shown in figure 1.

figure 1: Tapestry Data for the Jackson Area

The versatility and predictive power of tapestry segmentation allow users to integrate their own data into tapestry segmentation to identify their best market segments. Tapestry data derived the life groups by combining cluster analysis and traditional statistical methods to develop geodemographic data. The customer profile data was obtained from the 12-lifemode summary groups and summarize in the psychographic segmentation [6].

Block Groups data is organized in smaller numbers of markets that represent 207.721 variables and had been used to analyze certain consumption pattern.

V. RESULTS AND ANALYSIS

The random spatial patterns, to a naked eye, display some degree of clustering to statistically derive patterns of hot spots or cold spots, the study use statistical methods like Getis-Ord Gi* to quantify spatial patterns [7]. Figure 2 and Figure 3 respectively show the cluster and hot spots of Targeted Customers.

These statistically derived hotspots can provide where the clustering occurs and the processes behind them. Once we segment the market based on this clustering, the statistically significant information can be used in variety of business improvement strategies.

The Hot Spot Analysis tool calculates the Getis-Ord Gi* statistic for each feature in a dataset. The resultant Z score tells you where features with either high or low values cluster spatially [6].

Figure 2: A map showing Zip Code Clusters

The G_i^* statistic is a z score

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n}$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2}$$
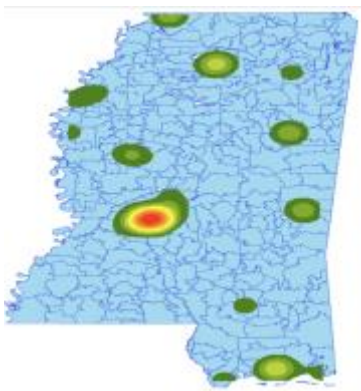
so, no further calculations are required.

Figure 3: A map showing Hot spots

This tool works by looking at each feature within the context of neighboring features. A feature with a high value is interesting, but may not be a statistically significant hot spot. To be a statistically significant hot spot, a feature will have a high value and be surrounded by other features with high values as well. The general model will take the form shown below.

$$Gi^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j} \right)^2 \right]}{n-1}}}$$

Where x_j is the attribute value for feature j , $w_{i,j}$

the spatial weight between feature i and j n i equal to the total number of features and:

VI. CONCLUSIONS

This study has demonstrated how Big Data Analytics combined with Geographic Information Systems (GIS) can be used to market segments and identifies potential targeted customers by generating hot spots of customer profiles. [8] The following were achieved:

i. a map showing locations of target customers are derived from the tapestry data for the Jackson, Mississippi region.
ii. A map show the cluster and hot spots of Targeted Customers by Toyota Company
iii. A statistical model to quantify spatial patterns

The integral value and role of Cluster Analysis into the market segmentation that integrate Big Data cannot be overestimated. This study has highlighted the importance of business enterprises to apply market segmentation using geospatial techniques to incorporated data and analytics into their operations and visualize their assets graphically.

REFERENCES

[1] Grimshaw, D. J. (2000). Bringing Geographical Information Systems Into Business. Canada: John Wiley and Sons.
[2] Civic Technologies. (2009). Using Market Segmentation for Better Customer Service and More Effective Strategic Planning. CivicTechnoligies.com.
[3] Dunbar, Malcolm McDonald (2012) Definition of Market segmentation.
[4] DeSarbo, W., Grisaffe, D.: Combinatorial optimization approaches to constrained market segmentation: An application to industrial market segmentation. Marketing Letters 9(2), 115–134 (1998

[5] Zhenning Xu, G. L. (2015). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*

[6] Keechoo Choi, W . J. (2000). Development of a transit network from a street map database with spatial analysis and dynamic segmentation. *Transportation Research Part C: Emerging Technologies* .

[7] Easterfield, C. (2013, March 11). *Mapping the GIS Market: A Market Research Model from Cambashi.* Retrieved July 18, 2015, from DirectionsMag: http://www.directionsmag.com/entry/mappi ng-the-gis-market-a-market-research-model- from-cambashi/31501

[8] ESRI. (2014). *Detailed Data for Profiling Customers or Constituents* . Retrieved August 7, 2015, from ESRI: http://www.esri.com/news/arcuser/0104/tape stry.html
.