Volume 2, Issue 2, 2026, Pages 105–114 *Journal of Digital Business and International Marketing*
| Regular Article | Open Access

Geo-Referenced Customer Segmentation Using K-Means and SOM for Spatially Informed Marketing Decisions

Walid Assaf
Faculty of Law and Political and Administrative Sciences, Lebanese University, Museum, Beirut, Lebanon.
assaf@ul.edu.lb

Zhu Ying
Faculty of Arts and Humanities, Sun Yat-sen University, Guangdong Province, Guangzhou, Haizhu, China.
zhuying@qq.com

Article Info
Journal of Digital Business and International Marketing
https://www.ansispublications.com/journals/jdbim/jdbim.html

https://doi.org/10.64026/JDBIM/2026011

Received 10 January 2026.
Revised from 16 February 2026.
Accepted 26 February 2026.
Available online 05 April 2026.
Published by Ansis Publications.

Corresponding author(s):
Zhu Ying, Faculty of Arts and Humanities, Sun Yat-sen University, Guangdong Province, Guangzhou, Haizhu, China.
Email: zhuying@qq.com

Abstract – This research presents a customer segmentation model as a geo-referenced geo-latitude and longitude-based framework, which combines transactional behavior with latitude longitude coordinates through the k-means approximation and Self-Organizing Maps (SOM). Our pipeline methodology includes a data preprocessing phase, elbow-based cluster exploration, SOM configuration and visualization in the multilayer of the information to support spatially aware marketing decision making. The results suggest that k-means clustering models based on elbow-optimization procedure produce limited segregation, which does not provide evident geographical or behavioral delimitation of the customer population. On the other hand, a 4 by 4 hexagonal SOM trained with a Gaussian neighborhood function and evaluated with topological error values produces 16 topology preserving clusters that distinguish finished goods, spare part and repair customers on both homogeneous and non-homogeneous regions.

Keywords – Self-Organizing Map, K-Means Clustering, Geo-Referenced Data, Customer Segmentation, Spatial Marketing Analytics, U-Matrix Visualization, Quantization Error, Topological Error.

I. INTRODUCTION

Traditionally, the segmentation process has mostly depended on demographic factors such as age, income and geographical location. These methods as far as they helped to gain useful insights cannot be deemed effective in capturing the changing nature of consumer behavior in the modern data-rich world. The current marketing trends have been changed, and there is a strong tendency to shift to the application of sophisticated analytics and machine learning (ML) algorithms. These methods have changed the client segmenting and precision of targeted marketing greatly [1].

Customer segmentation is a technique that is widely applicable in contemporary marketing to assist in better identifying consumer groups. The approach also focuses on attributes of behavior, which is better than the conventional classification systems. In the past, demographic and psychographic based segmentation patterns resulted in generalized marketing approaches which did not consider the individual preferences and behaviors. The scope of big data analytics is growing as the number of digital footprints is increasing. Social media, purchasing history and browser information have become essential in profiling customers.

The segmentation that is based on data allows building more accurate clients profiles, which in turn allows creating very specific marketing interventions. It is widely acknowledged that the strategic value of big data analytics in segmentation creates significant commercial value. A detailed model is introduced by Dahiya et al. [2] that explains how big-data analytics can be used to achieve the competitive advantage of organizations by converting raw data into strategic information. As stressed by Stone et al. [3], artificial (AI) can be used in an innovative way in marketing, especially in the optimization of decision-making, anticipation of the needs of consumers, and facilitating hyper-personal marketing communication. All these studies note the significance of applying AI-driven insights to the whole marketing process, starting with collecting data to the interaction. The key role of AI, ML, predictive analytics, and deep learning algorithms in segmentation advancement is emphasized by contemporary research.

Ahang et al. [4] provides an extensive analysis of data-mining based segmentation algorithms including clustering algorithms, classification algorithms, and combination techniques that are applied to identifying salient consumer segments. The scholars believe that behavioral data can be increased with the help of AI and ML, which will help to understand its clients better and target them more accurately. Kotras [5] explores how artificial intelligence can be used to personalize marketing, as automated and predictive algorithms could generate personalized experience in large scale.

To visualize a high-dimensional feature space, which is defined by the multivariate input data, on a two-dimensional discretized representation (the SOM grid), Park [6] uses a Self-Organizing Map (SOM) an unsupervised competitive-learning neural network. SOM algorithm is one of the most influential algorithms of unsupervised learning, and the task of SOM algorithm is to compress a high-dimensional set of data into a manageable two-dimensional representation. An example of a dataset is a set of measurements of p measurable variables. SOM systematically structures similar observations in groups and visualizes them in the map. This is referred to as Kohonen maps or Kohonen networks that was revised by Melssen, Wehrens and Buydens [7].

SOM training is competitive in nature compared to the traditional neural networks which are trained on the principles of error-correction. The biological models that inspired Kohonen were brain models, specifically, brain models, and Alan Turing breakthrough theories of morphogenesis, of which morphogenesis [8]. SOMs will also come in handy to analyze and map the complex data environments and encourage improved knowledge of the numerous patterns and relations of multidimensional information. Similar to most of the artificial neural-network architectures, SOM have two distinct stages: training and mapping. The training stage involves making use of a set of input data (representing the input space) to generate a reduced-dimensional form of the original data (describing the map space). Afterwards, during mapping, the extra input data is classified by the map that has been developed.

We propose an analysis of a geo-referenced customer segmentation model based on integrating behavioral and spatial variables, k-means and SOM to simultaneously harness both, measure clustering quality and produce interpretable visual decision-support artefacts to allocate marketing resources and plan marketing channels in geographically dispersed customer populations. Our study has been structured as follows in the remaining sections: Section II provides a background study of SOM algorithm and its notations. Section III describes the dataset, clustering methods, and evaluation approach we employed in the study. Section IV provides a detailed analysis of findings, which integrates the employment of k-means clustering algorithm, and Python module K-Means. Section V concludes the study confirming that traditional k-means clustering provides limited comprehension of complex spatial behavioral heterogeneity of geo-based customer data.

II. SOM ALGORITHM AND NOTATION

A Self-Organizing Map (SOM) is composed of neurons that are organized in a two-dimensional grid. The quantity of neurons may be limited to dozens of a few thousands. The neurons are modeled by an n-dimensional weight vector, which is denoted by a single value: $m[m_1, \dots, m_n]$, where n is the dimensionality of the input vectors. The neurons are formed as relationships between themselves and other neurons which form the topology of the map. A hexagonal or a rectangular neighborhood (**Fig. 1** and **2**a-e respectively) is regularly used.

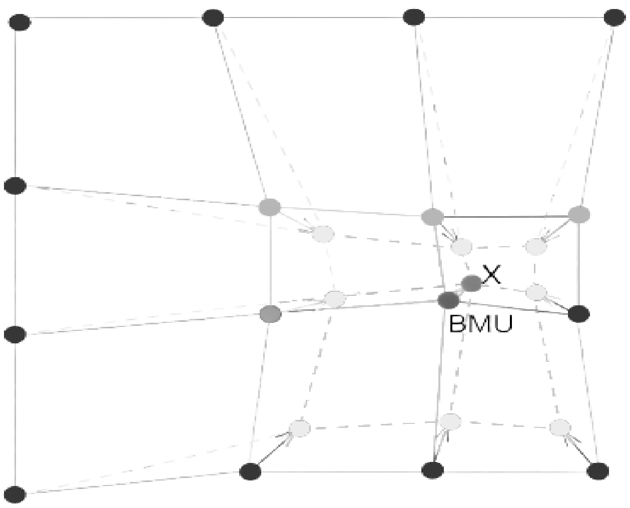

Fig 1. BMU and its Adjacent Neighbors Adapt to the Input Sample Represented by X.

The SOM approach normally facilitates a large number of repetitions until the node's reference vector of a 2D net are able to represent the nearest input patterns to the nodes (vector quantization). Finally, each sample of the dataset is projected to a network node through vector projections. SOM algorithm determines a non-linear mapping of the input field to the network topology, i.e. SOM output grid, during this optimization process. The mapping preserves the topological relationships between patterns of input wherever possible thus allowing the description of similarities and dissimilarities in the output grid. The foundational incremental SOM approach can then be simply using Eq. (1) and (2).

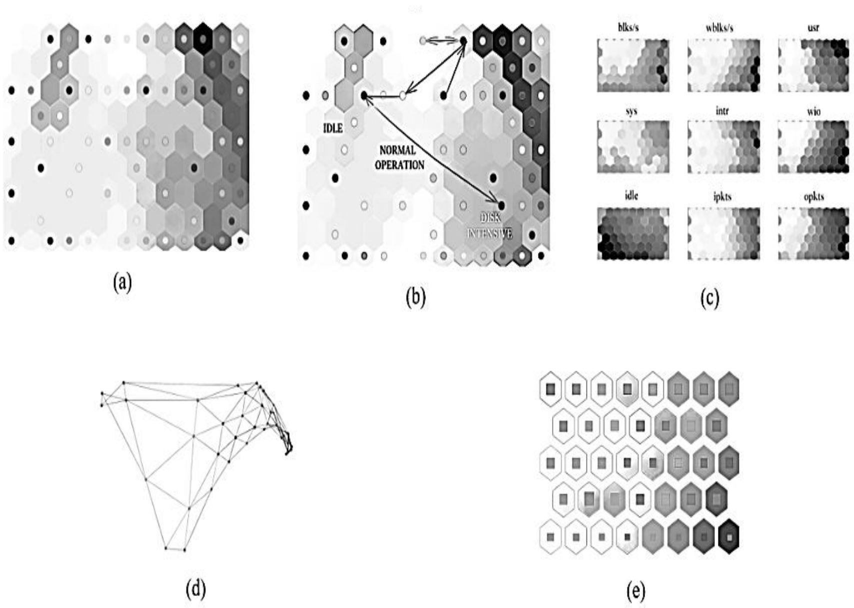

Fig 2. SOM Topology and Visualization of Hexagonal Grids.

$$\mathcal{X} = \{\mathrm{x_j}: \mathrm{j} = 1,2, \dots, \mathrm{m}\} \subset \mathbb{R}^{\mathrm{n}} \tag{1}$$

where:

$$x_j = [x_{j1}, x_{j2} \dots x_{jp}]^T \epsilon\ \mathbb{R}^n \tag{2}$$

Given a set of patterns (training patterns) in a set, denoted as $\mathcal{X}$ containing m patterns with n dimensions (variables). In the input domain of the network $\mathcal{R}^n$, each node of the network is specified and in the output domain of the SOM a typical unit grid is used. Input data domain and the map grid represent node i, and this node is represented by the reference (p-dimensional) vector m_i and the location (p-dimensional) vector r_i, computed using Eq. (3) and Eq. (4), respectively.

$$m_i = [m_{i1}, m_{i2} \dots, m_{in}]^T \in \mathbb{R}^n \tag{3}$$

$$r_i = [r_{i1}, r_{i2} \dots r_{ip}]^T \epsilon\ \mathbb{N}^p \tag{4}$$

Every reference vector, m_i, needs to be set up and specified in the data input field prior to the commencement of the learning process [9]. In addition, the SOM output field, herein referred to as the SOM coordinate space is determined by the lattice structure; it could be a rectangle, hexagon, and so forth. A Best Matching Unit (BMU) is found to a particular input pattern when the network is trained with a sequence of input patterns, denoted by x_j and reference vectors, denoted by m_i, which can be found by the assessment of the Euclidean distance. BMU is related to node c, as illustrated in in Eq. (5) and in the input data space, node c is represented by the reference vector c within the system of the form of m_c.

$$c = \underset{i}{\operatorname{argmin}}\{d(x_j, m_i\} \tag{5}$$

In which $d(x_j, m_i)$ signifies the Euclidean distance in between two vectors in the input data field (n bands). The network goes into the training phase of the input trend x_j by reaching to m_c and other benchmarking vectors which are in the close proximity of x_j. When the desired number of epochs is attained or when some other termination rule is attained, the training process is then terminated. Another implementation of the SOM approach is the batch processing algorithm, which like the traditional incremental SOM approach, is iterative.

Here, units are updated once the whole set of training has been introduced on the map and the cumulative effects of all patterns of input is incorporated [10]. The experimental methodology used in the research study employed the batch algorithm to train since it is significantly fast to compute when compared to sequential SOM technique and the performance achieved is often similar or better than that of the sequential technique. Fort, Cottrell, and Letremy [11] states that the batch method is much better than the classic SOM algorithm in the stability of asymptotic values of the reference vectors m_i and also in convergence problems.

III. METHODOLOGY

Dataset Description and Preprocessing

We utilized a geo-referenced transactional database, which tracks customer purchasing patterns as well as spatial co-ordinates given in latitudes and longitudes. These data consist of categorical variables, which define the category of products (finished goods, spares and repair services) and quantitative variables which measure the frequency of purchasing, intensity of interaction and distributional factors (see **Table 1**). Combining behavioral and spatial features, the dataset will allow the simultaneous market segmentation and geographic clustering which is essential to spatially informed decision-support systems.

Table 1. Overview of Dataset Characteristics

Category of Attributes	Description	Type	Purpose in Analysis
Product Category	Finished goods, spare parts, repair	Categorical	Cluster interpretation
Purchase Behavior	Frequency and interaction measures	Numerical	Pattern discovery
Latitude/Longitude	Geographic coordinates	Numerical	Spatial clustering

A critical preprocessing pipeline was done before the implementation of the model to ensure statistical reliability and analytical consistency. The preprocessing step entailed the systematic process of missing data, checking the ranges of the coordinates to remove spatial artifacts and normalization of numerical attributes to reduce the disproportionate influence of the high magnitude variables on distance base clustering. Similarity involving normalization Standardized normalization procedures were used to obtain feature scaling, so that spatial and behavioral features make a similar contribution to similarity calculations.

Fig. 3 shows the pipeline or consecutive steps used in the preprocessing pipeline and converting raw transaction records to an analytically processed dataset that can be used in unsupervised learning. This methodological design is used to guarantee that clustering results are true representations of inherent structural relationships, as opposed to results produced by measurement scales and non-observation.

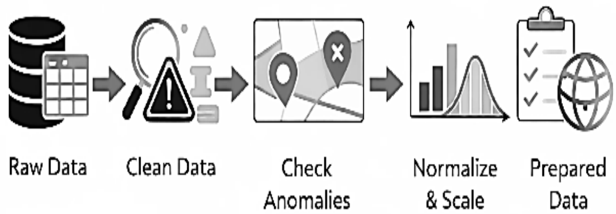

Fig 3. Geo Referenced Customer Data Preprocessing Workflow to Prepare Data to Be Clustered.

Clustering Methods and Model Development

Two unsupervised learning methods complementary to each other were used to test customer segmentation patterns: partition-based k-means clustering algorithm and topology preserving SOM neural network. The two-model approach has made possible comparative analysis between traditional variance-reduction clustering and nonlinear mapping of manifolds. The k-means algorithm divides observations into homogeneous clusters by minimizing within-cluster Euclidean distance [12]. Since k-means needs the a priori number of clusters, the elbow optimization strategy was used to assess the candidate cluster count on the basis of variance-reduction behavior. This optimization process, which is conceptually associated with the modeling pipeline in **Fig. 4**, empirically justifies the extent of segmentation before SOM analysis.

Table 2. Parameters of a Clustering and SOM Training Model

Model	Parameter	Value	Rationale
K-Means	Cluster search range	Elbow-evaluated	Identify the best segmentation.
SOM	Grid size	4 × 4	Exposure and elucidation of balance
	Iterations	1000	Make sure there is stability in convergence
	Sigma	1.0	Neighborhood of control influence
	Learning rate	0.5 with decay	Go permanent in adaptive learning.
	Training mode	Random	Improve generalization

After k-means exploration, SOM was applied in order to enjoy nonlinear topological associations in the geo-referenced behavioral space. The grid size of 4×4 hexagonal neuron grid was determined based on experimental trials to create a good balance between representational resolution and interpretation. The network was optimized by 1,000 iterations with a Gaussian neighborhood function with $2 = 1.0$ and an initial training learning rate of 0.5 which decreased linearly. Random sampling training mode was used to enhance the convergence stability between measurements which are spatially dispersed. These configuration options have been summarized in **Table 2** and are correlated with the analysis sequence as in **Fig. 4**.

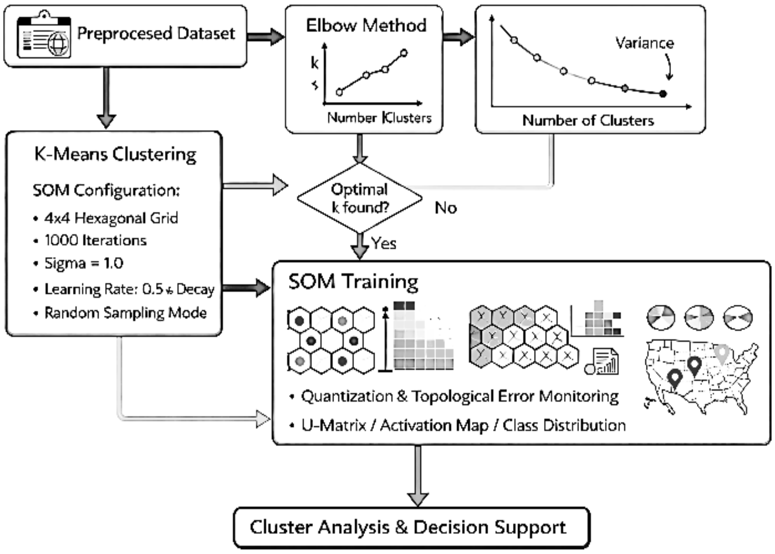

Fig 4. Topological Preserving Clustering Som Topology Analytical Modelling Pipeline Based on K-Means.

Measures of Evaluation and Presentation

The topological and quantization error were the two complementary performance measures applied in the quality of clustering and preserving SOM topology. The mean Euclidean distance between the input vectors and the nearest best-matching units (BMUs) is estimated using quantization error therefore providing an objective estimate of representational resolution [13]. The surrogate of the continuity of neighborhoods in the learned manifold is the topological error that is the ratio of cases when the BMUs are not adjacent neurons.

The constant observation of these measurements during the training process (which is repeated over time as shown in the **Fig.5**) will help to ensure that the convergence is stable and the mappings obtained will be reliable.

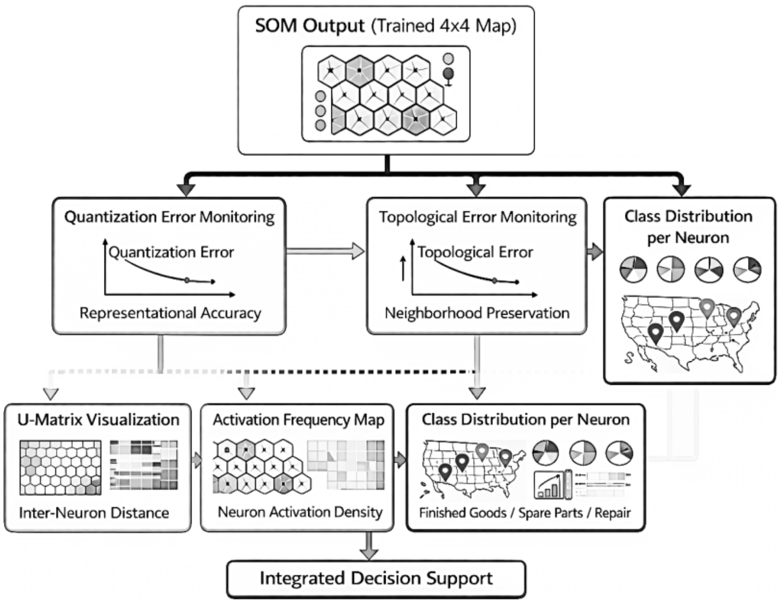

Fig 5. SOM Assessment System Connecting Errors Measures, Topology Mapping and Geographic Interpretation.

Besides the numerical approach to evaluation, visual analytics were an essential part of the interpretative scheme. A combination of complementary visualization methods was implemented, which includes U-Matrix distance mapping to explain the structures of inter-neuron similarity, maps of activation frequencies to locate dominant response loci, class-distribution pie charts per neuron to query the structure of categorical composition, and geographic scatterplots of clusters of SOM to combine topological data with spatial distributions in the real world. The architecture of integrated visualization process (deciding) is described in **Fig. 5**, whereby the SOM outputs are converted into consumable decision-support formats.

By incorporating quantitative error monitoring and the multilayer visualization, Section III provides an analytical base on which the pattern of geo-referenced customer segmentation is interpreted. These methodological processes directly endorse the empirical results provided in Section IV.

IV. RESULTS AND ANALYSIS

Using k-means clustering algorithm, we grouped data values into a specified cluster count (k) whereby every data point will belong to a cluster that has a nearest centroid (i.e. mean). The algorithm starts by picking k initial centroids at random and then attaches the points to the nearest centroid and then redefines each centroid as an arithmetic mean of all points in the member of the corresponding cluster. The elbow method aids in the determination of the optimum value of k; that is, one simply plots the within-cluster variance versus the cluster count and use the point where the variance reduction levels off. **Fig. 6** shows that the elbow criterion finds optimal cardinality of clusters of $k = 7$ in the current dataset.

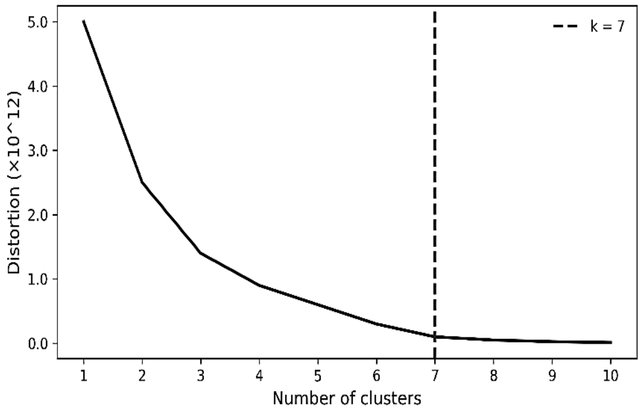

Fig 6. The Elbow Method.

The Python module KMeans was used to perform the K-means clustering, and the quantity of clusters was set to $k = 7$ as it has been established earlier. Consequently, 7 clusters together with their centrals are depicted in **Fig. 7**.

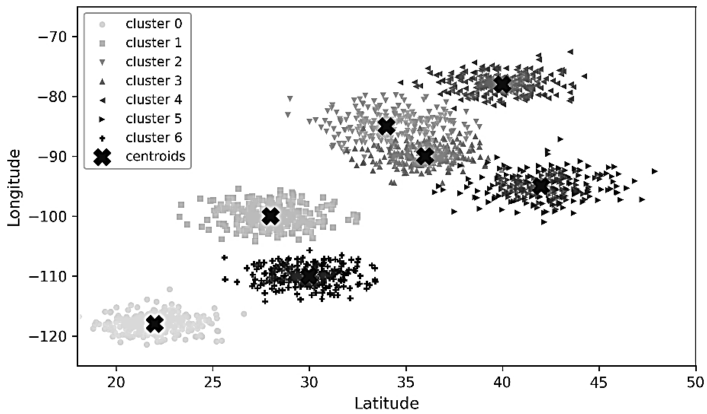

Fig 7. K-Means Clustering.

Geo-referenced data values help in cluster plotting according to geographic coordinates. The deployment of the k -means on the dataset failed to provide meaningful clusters and neither did it provide distinct boundaries or patterns among the clusters. Traditional approaches to clustering are commonly employed in marketing analytics to undertake segmentation exercises. The support capabilities are distributed based on the number of fixed clusters that have already been established by the marketing managers. To this end, a faulty analysis of the cluster number can create misinterpretation and therefore affect the tactical decision-making process.

The self-organizing maps (SOMs) were realized using Python Minisom package that provides two available training modalities, namely, train batch and train random. The train random option causes the model to pick randomly the data points throughout the training process whereas train batch picks the data set in the sequence it was set in. The train random strategy was used in this work.

The SOM algorithm works based on unsupervised learning and hence does not consider class labels during the training process. The product category, which is the target variable, was only used in the subsequent visualization of clusters in the SOM and not training itself. A number of different map dimensions were tested to see the optimal size of SOM. A 4 4 grid was finally chosen consisting of 16 neurons and it was trained over 1000 cycles. This arrangement provides a large feature map to support class expansion, and also brings out class overlap. A Gaussian neighborhood function was used in the model. All the other Minisom parameters were set as follows: neighborhood width (sigma) = 1.0, learn rate = 0.5 and a decreasing learning rate with each new iteration.

A SOM that is used as an accurate representation model should conserve the inherent topology and neighborhood relationship available in the input data [14]. In this regard, two standard quality measurements, namely, quantization error and topological error, were employed to evaluate the model performance.

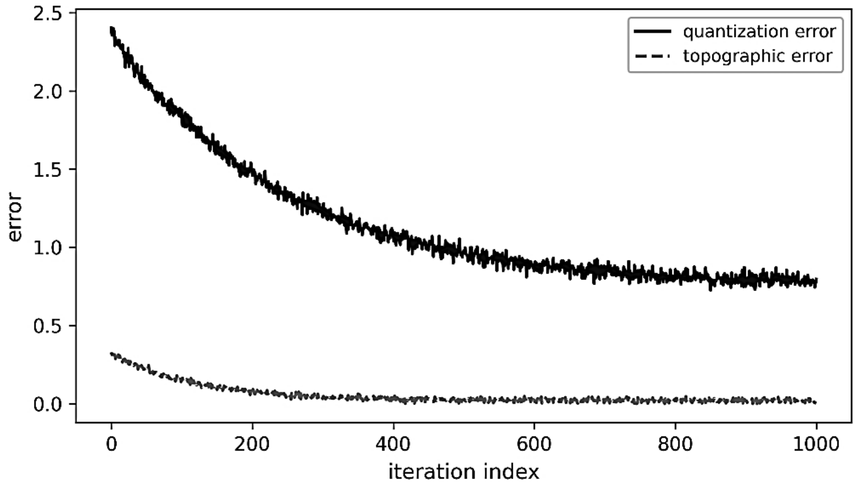

Fig 8. SOM with Quantization and Topography Error After 1000 Cycles.

Quantization error is used to estimate the granularity of the mapping and is an error metric of the difference between input vectors and the closest matching Best Matching Units (BMUs). Topological error or topographic error is a measure of how much the arrangement of neighboring input vectors is maintained; it is the fraction of data vectors with the first and second BMUs non-adjacent in the map [15]. The two metrics were calculated at each of the training cycles. Quantization error attained approximately 0.78 after 1000 cycles. **Fig.8** illustrates the changes in both the quantization and topological error throughout the entire training, which provides information on the training dynamics.

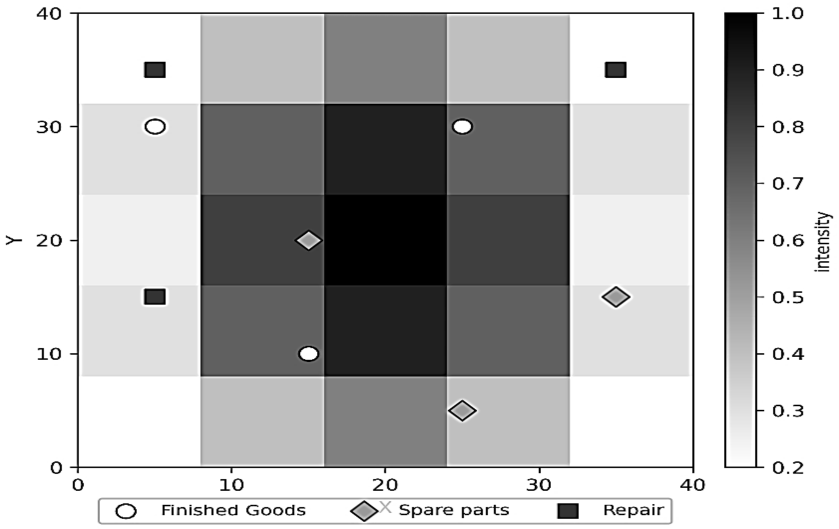

Fig 9. SOM Representing Distance with Marked Markers at the End of the 1000 Cycles.

SOM does not impose any manual algorithms of data points allocation to clusters and the limits of the clusters, as opposed to the traditional clustering algorithms. In this regard, it is necessary to visualize the SOM output in order to support users in the analysis process. Numerous visualization methods have been designed, each of which has mostly been created to represent the underlying cluster structure of the data [16]. In order to envisage the training performance, one may use the U-Matrix, and distance mapping (see **Fig.9**) in which a false color encoding scheme is used and each neuron is presented as a grid of cells, whose color representing the Euclidean distance of the weight-vector to the nearest neighbors.

Components with the same patterns are placed closely near each other. The representation of maps shows the neuron density using markers and colors that identify its scale. The black areas are more compact groups and the proximity of the nodes to each other and light colors show a distinction of a very long distance between them. The clusters involve the ones categorized as black (done regions), gray (spare regions), and white (regions). The SOM has detailed each cluster of the map in topological regions. Several of these clusters have been discovered in different areas of the SOM field, and other clusters have been discovered encompassing specific areas, therefore proving the effectiveness of the SOM. According to the U-Matrix, the data of several clusters occupy the areas with the high density of points.

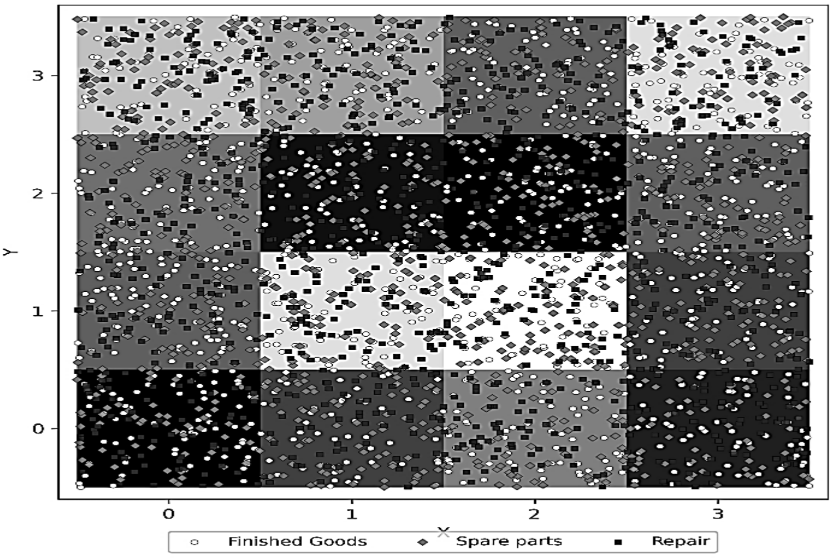

Fig 10. SOM Representing the Distances and Frequencies of Responses After 1000 Cycles.

Fig. 10 depicts the clusters of colors, representing the frequency of the activities in the neurons on the map. SOM mapping clearly highlights the data patterns as well as shows that consumers can be divided into sixteen distinct groups. On the lower part of the map, there are some green clusters, as well as red one, which overlap in some places, which means that they are similar. The clusters are also well defined towards the midpoint of the map. Clusters related to repair are easily recognized in the upper part and the two dominating groups on the right have the same feature.

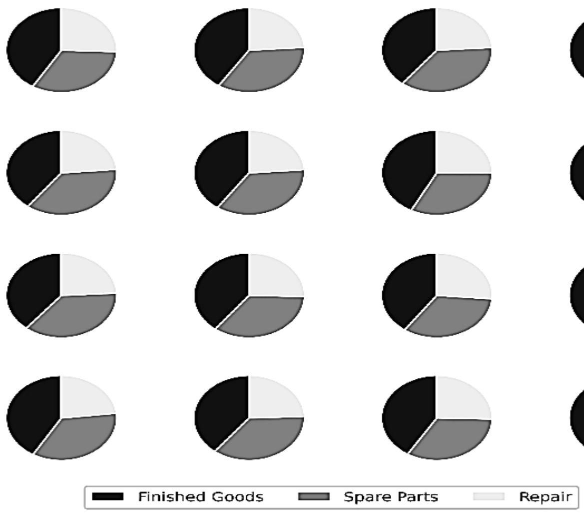

Fig 11. Class Distributions of SOM (4 × 4) At the End of 1000 Clusters.

Fig. 11 illustrates the count of samples in every cluster of a given neuron and is presented as a pie chart of the neuron. The class pie chart shows that there are four overall groups of customers that are associated with the completed items, three overall groups of consumers who purchase spare parts and three overall groups of customers that repair. The other six clusters include heterogeneous grouping of clients with different tendencies.

SOM clusters are represented as seen through their geographic distribution with the help of the geographic coordinates of the points of data noted at the stage of training. **Fig. 12** is a scatterplot demonstrating that the 16 SOM clusters are distributed spatially based on their coordinate attributes. The two-dimensional coordinate values of every centroid of the clusters are also shown in the plot. **Fig.13** illustrates superimposition of the geographic location of both clusters on a map of the United States and their centroid. The map shows a number of populated areas and especially close to the East Coast, which extends all the way in Boston up to New York and even the Southern states, which stretches to North Carolina up to Florida. The clusters are also disproportionately concentrated in the Eastern regions such as Michigan, Indiana, Ohio, and

Illinois. This means that future marketing channel allocation decisions can be based on a study of cluster centers on the geospatial clustering map.

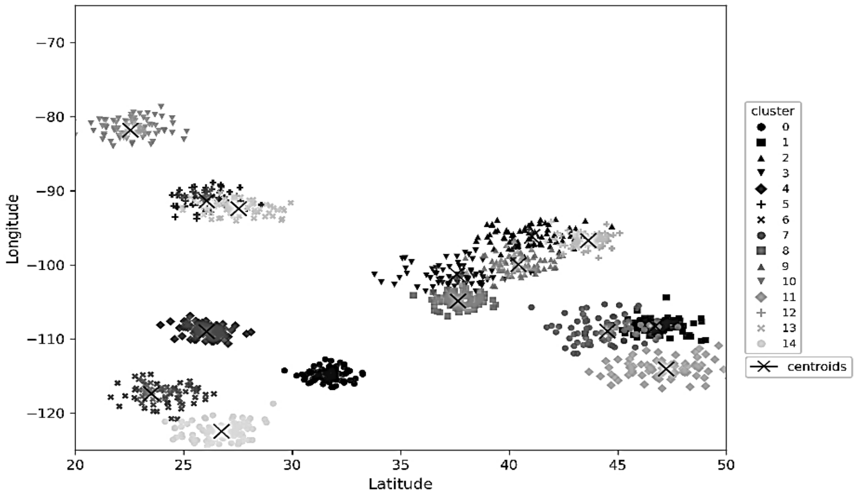

Fig 12. SOM Clustering Scatterplot of 1000 Cycles.

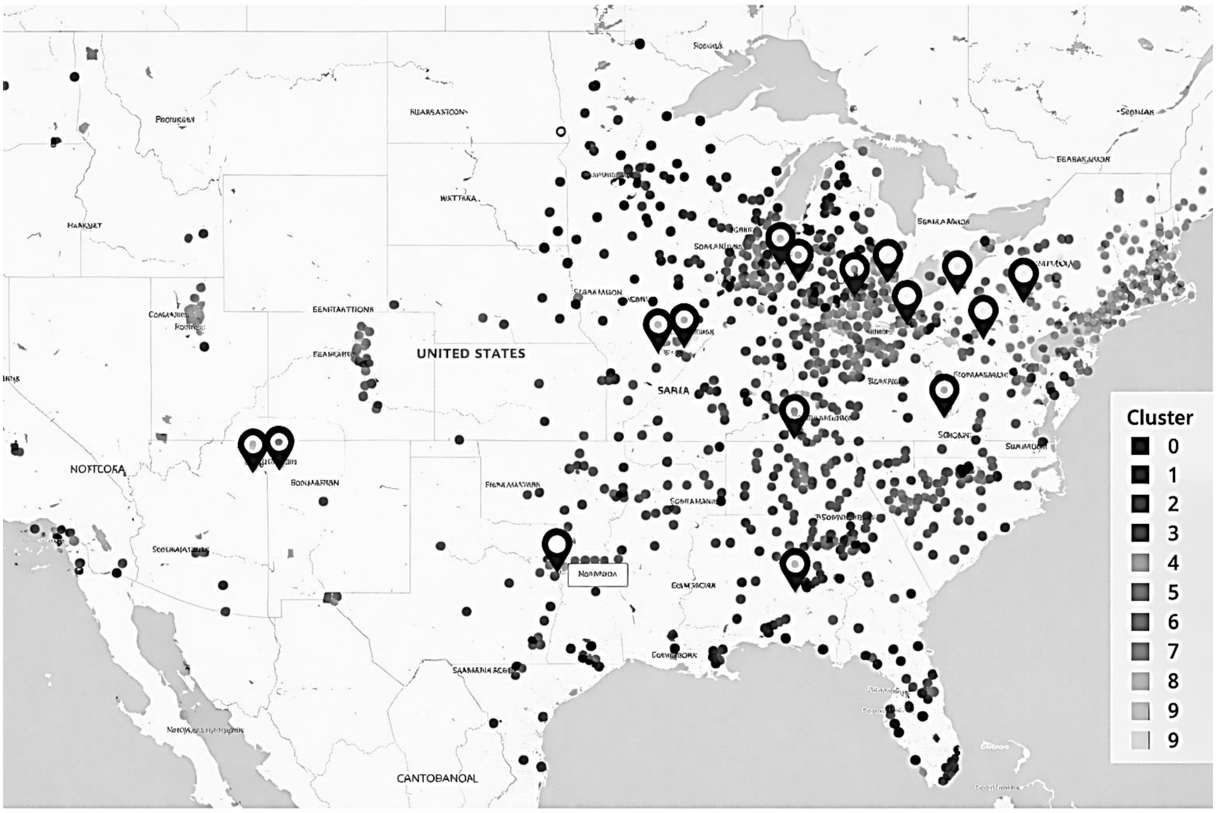

Fig 13. SOM Map of Geo-Clustering With 1000 Cycles.

V. CONCLUSION

Our study confirms that the traditional k-means clustering, with or without the accompanying elbow approach, provides limited understanding of the complex spatial-behavioral heterogeneity of georeferenced customer data. However, a topology-preserving representation, provided by a SOM set up with quantization, topological error evaluation, has been found to provide a representation of distinct customer clusters in distinct map segments, as well as to provide an identification of overlapping and mixed purchase behaviors. The U-matrix, activation-frequency maps, and neuron-level class pie charts provide the necessary information to answer the questions on the structure, purity and interrelationship between finished goods, spare parts and repair services. These patterns are then visualized by geo scatterplots of SOM clusters and centroids thus overlaying them onto spatially explicit marketing opportunities and allowing more accurate channel positioning and region targeted strategies.

CRediT Author Statement
The authors confirm contribution to the paper as follows:
Conceptualization: Walid Assaf and Zhu Ying; **Methodology:** Walid Assaf and Zhu Ying; **Writing- Original Draft Preparation:** Walid Assaf; **Visualization:** Zhu Ying; **Investigation:** Walid Assaf and Zhu Ying; **Supervision:** Walid Assaf and Zhu Ying; **Validation:** Walid Assaf and Zhu Ying; **Writing- Reviewing and Editing:** Walid Assaf and Zhu Ying. All authors reviewed the results and approved the final version of the manuscript.

Data Availability
No data was used to support this study.

Conflicts of Interests
The authors declare no conflict of interest.

Funding
No funding agency is associated with this research.

Competing Interests
There are no competing interests.

References
[1]. C. Marcus, “A practical yet meaningful approach to customer segmentation,” Journal of Consumer Marketing, vol. 15, no. 5, pp. 494–504, Oct. 1998, doi: 10.1108/07363769810235974.
[2]. R. Dahiya, S. Le, J. K. Ring, and K. Watson, “Big data analytics and competitive advantage: the strategic role of firm-specific knowledge,” Journal of Strategy and Management, vol. 15, no. 2, pp. 175–193, Feb. 2021, doi: 10.1108/jsma-08-2020-0203.
[3]. M. Stone et al., “Artificial intelligence (AI) in strategic marketing decision-making: a research agenda,” The Bottom Line, vol. 33, no. 2, pp. 183–200, Apr. 2020, doi: 10.1108/bl-03-2020-0022.
[4]. F. Ahang, A. Imani, M. Abbasi, H. Ghaffari, and M. Mehdi, “Customer segmentation to identify key customers based on RFM model by using data mining techniques,” Shilap Revista De Lepidopterología, Mar. 2022, doi: 10.22105/riej.2021.291738.1229.
[5]. B. Kotras, “Mass personalization: Predictive marketing algorithms and the reshaping of consumer knowledge,” Big Data & Society, vol. 7, no. 2, Jul. 2020, doi: 10.1177/2053951720951581.
[6]. Dong-Chul Park, “Centroid neural network for unsupervised competitive learning,” IEEE Transactions on Neural Networks, vol. 11, no. 2, pp. 520–528, Mar. 2000, doi: 10.1109/72.839021.
[7]. W. Melssen, R. Wehrens, and L. Buydens, “Supervised Kohonen networks for classification problems,” Chemometrics and Intelligent Laboratory Systems, vol. 83, no. 2, pp. 99–113, Sep. 2006, doi: 10.1016/j.chemolab.2006.02.003.
[8]. H. Greif, A. Kubiak, and P. Stacewicz, “Turing’s Biological Philosophy: Morphogenesis, Mechanisms and Organicism,” Philosophies, vol. 8, no. 1, p. 8, Jan. 2023, doi: 10.3390/philosophies8010008.
[9]. F.-P. An, “Human Action Recognition Algorithm Based on Adaptive Initialization of Deep Learning Model Parameters and Support Vector Machine,” IEEE Access, vol. 6, pp. 59405–59421, 2018, doi: 10.1109/access.2018.2874022.
[10]. Baraldi and F. Parmiggiani, “A neural network for unsupervised categorization of multivalued input patterns: an application to satellite imaee clustering,” IEEE Transactions on Geoscience and Remote Sensing, vol. 33, no. 2, pp. 305–316, Mar. 1995, doi: 10.1109/tgrs.1995.8746011.
[11]. J.-C. Fort, M. Cottrell, and P. Letremy, “Stochastic on-line algorithm versus batch algorithm for quantization and self-organizing maps,” Neural Networks for Signal Processing XI: Proceedings of the 2001 IEEE Signal Processing Society Workshop (IEEE Cat. No.01TH8584), pp. 43–52, doi: 10.1109/nnsp.2001.943109.
[12]. S. M. Miraftabzadeh, C. G. Colombo, M. Longo, and F. Foiadelli, “K-Means and Alternative Clustering Methods in Modern Power Systems,” IEEE Access, vol. 11, pp. 119596–119633, 2023, doi: 10.1109/access.2023.3327640.
[13]. Y. Bernard and B. Girau, “Fast Parallel Search of Best Matching Units in Self-organizing Maps,” Advances in Self-Organizing Maps, Learning Vector Quantization, Clustering and Data Visualization, pp. 11–20, 2022, doi: 10.1007/978-3-031-15444-7_2.
[14]. P. Rousset, C. Guinot, and B. Maillet, “Understanding and reducing variability of SOM neighbourhood structure,” Neural Networks, vol. 19, no. 6–7, pp. 838–846, Jul. 2006, doi: 10.1016/j.neunet.2006.05.017.
[15]. D. Rajashekar, A. N. Zincir-Heywood, and M. I. Heywood, “Smart Phone User Behaviour Characterization Based on Autoencoders and Self Organizing Maps,” 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), pp. 319–326, Dec. 2016, doi: 10.1109/icdmw.2016.0052.
[16]. E. Ventocilla and M. Riveiro, “A comparative user study of visualization techniques for cluster analysis of multidimensional data sets,” Information Visualization, vol. 19, no. 4, pp. 318–338, Jul. 2020, doi: 10.1177/1473871620922166.

ISSN: 3104-4115