LogForum

> Scientific Journal of Logistics <

2020, 16 (4), 535-547

http://doi.org/10.17270/J.LOG.2020.512

http://www.logforum.net p-ISSN 1895-2038 e-ISSN 1734-459X

ORIGINAL PAPER

INVENTORY MANAGEMENT AND LOGISTICS OPTIMIZATION: A DATA MINING PRACTICAL APPROACH

Rafael Granillo-Macías

Autonomous University of Hidalgo State, Hidalgo, **México**

ABSTRACT. **Background:** In the current economic scenarios, characterized by high competitiveness and disruption in supply chains, the latent need to optimize costs and customer service has been promoted, placing inventories as a critical area with high potential to implement improvements in companies. Appropriate inventory management leads to positive effects on logistics performance indices. In economic terms, about 15% of logistics costs are attributed to warehousing operations. With a practical approach, using a case study in a company in the food sector, this article proposes an inventory classification method with qualitative and quantitative variables, using data mining techniques, categorizing the materials using variables such as picking frequency, consumption rates and qualitative characteristics regarding their handling in the warehouse. The proposed model also integrates the classification of materials with techniques for locating facilities, to support decision-making on inventory management and storage operations.

Methods: This article uses a method based on the Partitioning Around Medoids algorithm that includes, in an innovative way, the application of a strategy for the location of the optimal picking point based on the cluster classification considering the qualitative and quantitative factors that represent the most significant impact or priority for inventory management in the company.

Results: The results obtained with this model, improve the routes of distributed materials based on the identification of their characteristics such as the frequency of collection and handling of materials, allowing to reorganize and increase the storage capacity of the different SKUs, passing from a classification by families to a cluster classification. Furthermore, the results support decision-making on storage capacity, allowing the space required by the materials that make up the different clusters to be identified.

Conclusions: This article provides an approach to improving decision-making for inventory management, showing a proposal for a warehouse distribution design with data mining techniques, which use indicators and key attributes for operational performance for a case study in a company. The use of data mining techniques such as PAM clustering makes it possible to group the inventory into different clusters considering both qualitative and quantitative factors. The clustering proposal with PAM offers a more realistic approach to the problem of inventory management, where factors as diverse as time and capacities must be considered, to the types and handling that must be had with the materials inside the warehouse.

Key words: cluster, Partitioning Around Medoids, facility location, supply chain.

INTRODUCTION

The environment of high global competition and the increase in demand for products and services by customers have driven the search for options to improve efficiency in companies' operations. For many organizations, internal logistics and warehousing are considered areas with high potential for the implementation of improvements, which can have a favorable impact on the operational efficiency of the business [Anđelković, Radosavljević 2018, Grosse, Glock 2015]. Warehousing operations are a key factor for the operational success of an organization, since they promote and support the fulfillment of the set of requirements and expectations of customers, throughout the supply chain. Proper inventory

Citation: *Granillo-Macías R., 2020. Inventory management and logistics optimization: A Data Mining practical approach. LogForum 16 (4), 535-547, http://doi.org/10.17270/J.LOG.2020.512*

Received: 24.06.2020, Accepted: 06.08.2020, on-line: 18.09.2020.

management has an important effect on logistics performance indices, mainly in those organizations that seek to reduce their costs and improve efficiency in their product preparation and delivery processes [Zhang et al. 2019]. In economic terms, warehouses represent about 15% of total logistics costs in developed countries [Guo et al. 2016].

A warehouse is an intermediary facility between suppliers and customers, whose utility is to dampen demand considering time and cost variables, seeking to reduce the gap between the production and consumption of goods [Aqlan 2017]. Operations carried out in a warehouse are generally divided into reception, storage, order picking, sorting, and shipping [Zhang et al. 2019, Çelik, Süral 2019], with order picking operations consuming the most time and work. Order picking is the most expensive operation, constituting around 55% of a warehouse's operating costs [Bottani et al. 2019, Çelik, Süral 2019, Grosse, Glock 2015], this being the main reason for which the preparation of orders is considered as an area of study opportunity to achieve the improvement of productivity in the company. In logistics, order picking refers to the process of selecting a set of Stock Keeping Units (SKUs), retrieving them from their various storage locations, and transporting them for review, packaging, and shipping for fulfillment of customer orders. internal or external. Order picking activities are also the most time consuming, in the case of manual operations, caused by heavy work and repetitive activities carried out in the warehouse. In manual operations where the labor force intervenes, the cost of harvesting is mainly related to the time used to transport products within the warehouse. According to studies carried out by [Bottani et al. 2019], the collection cost represents approximately 50% of the total order picking time.

Transportation time is related to the distance that must be traveled to collect the items requested in a customer's order. Therefore, minimizing the distance and the collection time is an essential objective to achieve efficiency and competitiveness in the warehouse [Faia Pinto, Nagano 2019]. In planning for order picking, several decisions need to be made at the tactical and operational levels [Çelik, Süral 2019, Bottani et al. 2019]. Decisions at the tactical level include: (1) the allocation of products to storage areas, which describes the rules for determining the assignments of SKUs to storage locations, and (2) the zoning of storage areas. collection, which is a means through which policies are decided on how to divide the order picking area into zones and determine the locations of the order picking areas [van Gils et al. 2018]. Operation-al decisions are influenced by: (1) order batch processing, based on rules that define the mix of customer orders in a single selection round, and (2) routing policies, which define the sequence of storage locations that must be visited to collect all the SKUs necessary for the formation of an order. Current trends in supply chain management promote the optimization of inventories, as a support to storage operations, through the use of different technologies [Aqlan 2017]. Technological tools based on data analysis, such as Warehouse Management Systems (WMS), are becoming increasingly sophisticated. WMSs are designed to provide efficient information for making decisions about storage, inventory, and SKU movements. In the Big Data era, analytical techniques such as data mining and business intelligence are being used in inventory management to provide accurate and up-to-date information to make better decisions [Xindong Wu et al. 2014]. Specifically, data mining analysis techniques are considered as the fundamental basis for Big Data.

Data mining is the process of extracting information from a data set, its main strengths are combining statistical models and autonomous learning, offering versatility to treat different types of data. Choi et al. [2018], Arora and Chana [2014], Tsai et al. [2015] point out that the areas focused on Big Data analysis in data mining include: (1) grouping techniques, (2) distributed and parallel processing, and (3) processing multimedia. Grouping techniques divide a data set into different groups. The grouping process consists of assigning a large number of data points to a smaller number of groups, so that the data points in the same group share the same properties, while the data of other groups are different. Grouping consists of classifying the input data based on certain values or attributes

[Aqlan 2017]. Clustering is used in different areas, including artificial intelligence, marketing, scientific analysis, and engineering [Xindong Wu et al. 2014]. The analysis through cluster grouping allows inventories to group SKUs according to certain characteristics. Specifically, this study presents an approach to inventory management, using cluster grouping based on variables related to the frequency of collection, storage, and warehouse returns. Using data mining techniques based on the Partitioning Around Medoids (PAM) algorithm, a case study is analyzed in a company in the food sector. The objective is to propose a method for inventory management, using clustering techniques and optimal picking point, categorizing materials through variables such as collection frequency, consumption rates, and qualitative characteristics regarding their handling in the warehouse.

The hypothesis was that using cluster grouping techniques it is possible to manage inventory, distribution, and storage logistics in a company, including both qualitative and quantitative variables. This article begins by identifying recent contributions to inventory management, order picking, and picking. Subsequently, the proposed approach based on data mining techniques is presented. Using data collected from a company in the food sector, the proposal is analyzed through a case study. Afterward, the results obtained are discussed, highlighting the variables that have the greatest contribution to storage logistics operations. Finally, the conclusions and future works for this study are presented.

LITERATURE REVIEW

The planning and control of materials and products, which support production functions, maintenance activities, and customer service, is carried out and coordinated through inventory management. The latent need to optimize costs and customer service has placed inventories as a fundamental area for improvement in companies, due to the high level of cost that they can reach in an organization. According to the extensive literature review carried out by Gu et al. [2007], the problems for inventory management are classified according to storage (reception, storage, order preparation, and shipping). Among the traditional and most widely used techniques by organizations for inventory management is ABC analysis. Class-based storage according to the ABC demand curve, divides stored items considering policies such as inventory turnover or cost [Guo et al. 2016]. Grosse and Glock [2015] develop an analytical model that helps predict performance on certain elements of the order picking system. With a quantitative approach, Jemelka et al. [2017] present a variant of the ABC analysis for the determination of inventories using a recursive model that considers the rates of return of materials and the redistribution of the sections for the location of SKUs within from a warehouse.

To distinguish themselves from the competition, reducing the time of preparation of orders, van Gils et al. [2017] propose the forecast of the workload in a warehouse con-text with emphasis on collection areas. Exploring the performance of different tools for order picking, de Vries et al. [2016] suggests that the human factor plays an important role for order picking within the warehouse. In the research by van Gils et al. [2018] statistically analyze and test the relationships between storage, order processing, batching, zoning, and routing by a full factorial Analysis of Variance (ANOVA). These authors conclude that significant benefits can be achieved in inventory management, while simultaneously considering storage, order processing, zone selection, and routing policies. Zhang et al. [2019] present the concept of Demand Correlation Pattern (DCP), to describe the correlation between SKU's, based on which they show a model to address the Storage Location Assignment Problem (SLAP). With the DCP proposal, Zhang et al. [2019] conclude that the class-based storage strategy, which divides SKUs into several classes and as-signs each class to a storage area, is one of the strategy of the inventory management most commonly adopted in practice.

Designing a multi-criteria inventory classification approach, Lolli et al. [2014] presents a method based on a hybrid model that combines the K-means algorithm and

Analytic Hierarchy Process (AHP). Analyzing multi-zone storage systems within the configuration of a WMS, Yuan et al. [2018] explore decisions on zones stowage, to determine the best distribution of products that arrive through multiple storage areas. In the research of Anđelković and Radosavljević [2018] use cluster analysis to identify which are the inventory management processes that can achieve the most significant benefits for the implementation of the WMS, as a result, these authors point out that the order processing operations are the most appropriate to implement information technologies based on the WMS. Examining the order picking problem (OPP), Çelik and Süral [2019] makes use of the properties of graph theory, integrating a model based on a heuristic approach, to determine the route that minimizes the times of transfer required for storage operations. A proposal on the minimization of the transfer times for collection is presented by Matthews and Visagie [2019], to reach an adequate arrangement for the collection operations of SKUs in a warehouse. Faia Pinto and Nagano [2019] propose a computational tool called GA-OPS, which is formulated based on two genetic algorithms, to minimize the number of picking trips, meeting the requirements requested in the different production orders.

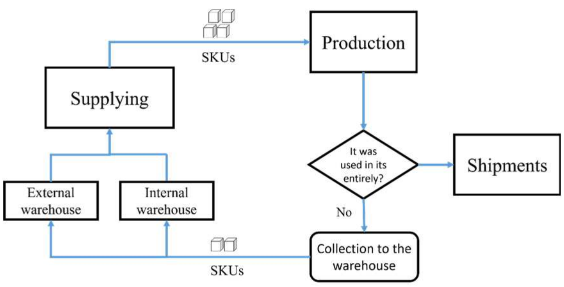
Djatna and Hadi [2017] through a multi-objective mathematical model, they present the order preparation problem in the warehouse of a beverage company with a drive-in rack system. Djatna and Hadi [2017] concludes that the integration between order picking (warehousing allocation, routing, batch processing, zoning and warehouse design) and other aspects (queue, operational and material handling aspects), is a current challenge for the order picking research area. Combining cluster analysis and simulated annealing algorithm to search for optimal classification in a warehouse, the authors Liu et al. [2016] presents a methodology that builds hierarchies of similar inventory groups, and then applies a simulated annealing algorithm to optimize inventory classifications on different hierarchy levels. Kusrini [2015] present a study that supports the process to determine the minimum stock and profit margin, using a model that groups the SKUs into categories of "fast movement" and "slow movement", using the grouping method from k-means.

Using data mining techniques, Aqlan [2017] categorizes inventories with cluster grouping, based on variables of collection frequency, time in storage, price, and sensitivity of products to transport. [Hong 2019] analyzing the elements involved in inventory management, he proposes the variables of flow time, work in process, and throughput in terms of pick probability. Considering that the classification and categorization of inventories require using multiple criteria to control different functions of inventory management, in the approach of Aktepe et al. [2018] an algorithm called functional-normal-and small (FNS) is analyzed, the FNS combines ABC analysis with variables of handling frequency, lead time, contract manufacturing process and specialty, which are used as input criteria for this model. In general, most of the literature on inventory management is focused on aspects related to the optimization of time, transfer distances, and use of resources, these being a reference to quantify improvements in warehousing operations in a company. Inventory analysis, proposed by various authors based on categories, is a predominant element that contributes to the improvement of storage operations. Specifically, the collection and preparation of orders are variables that are recognized as important factors that need to be considered to achieve effectiveness in inventory management.

Regarding classification methods, the k-means method is the most frequently used data mining technique, however, in this research, the cluster method is proposed through the PAM algorithm using quantitative variables such as (1) collection frequency, (2) average quantity per order, (3) daily rate of consumption, (4) daily rate of returns, (5) the average amount of returns per order and (6) frequency of return and related qualitative variables in the form of storage.

METHODOLOGY

The proposed approach to inventory management in this research is based on cluster grouping, to identify common elements among the different SKU´s that are located in a warehouse. Each cluster is integrated considering variables with particular characteristics, making them different from the other clusters formed. For this research, an international company in the food sector was used as a case study, which focuses on the production of food and beverages, specifically analyzing the problems of a business unit located in Mexico. This business unit with operations in Mexico, specializes in the preparation and bottling of beverages. The inventory management of the said company is carried out through internal warehouses (inside the facilities) and external warehouses (located outside the facilities), in a make to order environment. Figure 1 shows the logistics carried out in the warehouse, with the production department as an internal customer.



Source: own work

Fig. 1. Logistics in warehouse

The company has the particularity of SKU's return operations, due to the certain formulation and mixing processes, so the materials in a production order are not consumed in their entirety and are returned from the production area to the warehouse. High inventory levels, storage capacity, lack of storage spaces, and frequent variations in demand are some of the problems present in the management of company inventories. With this background, the proposed methodology has as a first step the identification of SKUs found in the warehouse. This stage includes selecting the methods and tools to ensure the identification, location, and location of the different materials and products used by the company in its supply chain, present throughout the different production processes. By locating barcodes on the various materials and using reports from the Computer System for Resource Planning (SAP), 203 SKUs were identified that entered the warehouse, during a production period of one year. The materials are determined by families, which are: chemicals, powders, liquid concentrates, lids, pallets, packaging, labels, PET lids, and cardboard. Once the different SKUs that make up the warehouse have been described, the factors that present the most significant impact or priority for inventory management were subsequently selected. In this case, the following were chosen for exits from the warehouse: daily consumption rate (DRC) (1), average quantity per order (AQO) (2) and pick frequency (PF) (3); Regarding the factors with the greatest impact on returns, the company considered: daily rate of return (DRR) (4), the average number of returns per order (ARO) (5) and frequency of return (RF) (6). The PF determines the frequency with which the materials are required by the production department [Aqlan 2017], while the RF indicates the frequency with which these are returned from the production department to the warehouse.

The following equations were used to calculate each of these factors :

$$DRC = \frac{Q_n^p}{t} \tag{1}$$

$$AQO = \frac{Q_n^p}{f_n^p} \tag{2}$$

$$PF = \frac{DRC}{AOQ} \tag{3}$$

In the case of returns, the equations are:

$$DRR = \frac{B_n^p}{t} \tag{4}$$

$$ARO = \frac{B_n^p}{r_n^p} \tag{5}$$

$$RF = \frac{DRR}{ARO} \tag{6}$$

where:

Q_n^p: demand quantity Q of SKU p in period n
B_n^p: quantity B of SKU p in period n returning to the warehouse
t: period in days

f_n^p: SKU request frequency p during period n

r_n^p : frequency of returns of SKU p during period n

Due to the characteristics of some materials, which require occupying positions in "rack" and "floor" locations, the company added this qualitative factor for inventory management, also including the unit of measure factor in which the materials are accounted (kilograms, pieces, or gallons). A location on the floor indicates that the material does not require special storage conditions, while a location on racks is that location where materials such as chemicals, powdered ingredients, and concentrates are stored, which requires a storage system using racks and in some instances under controlled temperature conditions, to preserve the materials they protect in optimal conditions. After calculating these factors, the next step was classification using the cluster grouping technique. In a cluster analysis, a set of data, in this case, the SKUs, are grouped by similarity in the input variables, for this study six quantitative factors (continuous) and two categories, to identify groups that are internally as homogeneous as possible but differ from each other as much as possible. Producing a reasonable grouping and classified in a more similar series is one of the main advantages of cluster analysis [Akay and Yüksel 2018]. The clustering methods are also designed for applications where the data varies over time.

The clustering algorithm used in this study was Partitioning Around Medoids (PAM). The PAM algorithm minimizes the sum of the differences of each observation for its medoid. Since in the operations of the company some SKUs record atypical consumption and storage under consignment (customer property), the PAM algorithm was used using k-medoids. A medoid refers to the element of a cluster whose average distance (difference) between it and all other items in the same cluster is the shortest possible. Using medoids instead of centroids makes the PAM method more robust, being less affected by outliers or noise, compared to algorithms like k-means [Akay and Yüksel 2018, Kaufman and Rousseeuw 2005]. The PAM algorithm is developed with the following steps: (1) select k random observations as initial medoids, (2) calculate the distance matrix between all observations, (3) assign each observation to its closest medoid, (4) to each cluster created, check if selecting another observation as medoid reduces the distance of the cluster and (5) check if at least one medoid has changed, otherwise, the process ends. For the dataset, the Gower distance metric was used, which is not possible with other algorithms, for example, k-means, which only allows Euclidean or Manhattan distances. Gower distance metric is a powerful method proposed by Gower [1971] and extended by Kaufman and Rousseeuw [2005], applied to databases with continuous, ordinal or categorical variables at the same time. Gower distance is based on Gower´s General Similarity Coefficient Sij, comparing two cases i and j, defined as:

$$\boldsymbol{S_{ij}} = \frac{\sum_k^n \boldsymbol{w_{ijk} S_{ijk}}}{\sum_k^n \boldsymbol{w_{ijk}}} \qquad (7)$$

where:

S_{ijk} : indicates the contribution provided by the k-th variable.

w_{ijk} : it is usually 1 or 0 depending on whether the comparison is valid for the k-th variable.

Through the cluster grouping, the SKU's classification is obtained according to the proposed factors. Finally, with the classification obtained through PAM, the inventory analysis is performed, proposing a redistribution of materials in the warehouse. In this case, the optimal picking point was identified as an additional strategy for inventory management. The optimal picking point location strategy based on the redistribution of materials was a suggestion from the company to determine the location of the collection point that minimizes the transfer distances. Through the facility location problem, points a^1, … a^m ∈ R^2 was minimized where a represents a location within the warehouse. Using the Euclidean distances between points, the calculation for the location of the picking point was performed with equations (8) and (9).

$$\boldsymbol{d_2^2(x, a^i) := (x_1 - a_1^i)^2 + (x_2 - a_2^i)^2} \qquad (8)$$

for all $(x_1, x_2) \in R^2$

and $a^i := (a_1^i, a_2^i), i = 1, \dots, m$

$$\sum_{i=1}^{m} v_i \cdot d_2^2(x, a^i) = \sum_{i=1}^{m} v_i \cdot \left((x_1 - a_1^i)^2 + (x_2 - a_2^i)^2\right) \quad (9)$$

where $v_1, \dots, v_m \in R$ are weights assigned to the materials, according to the cluster classification obtained in the previous step.

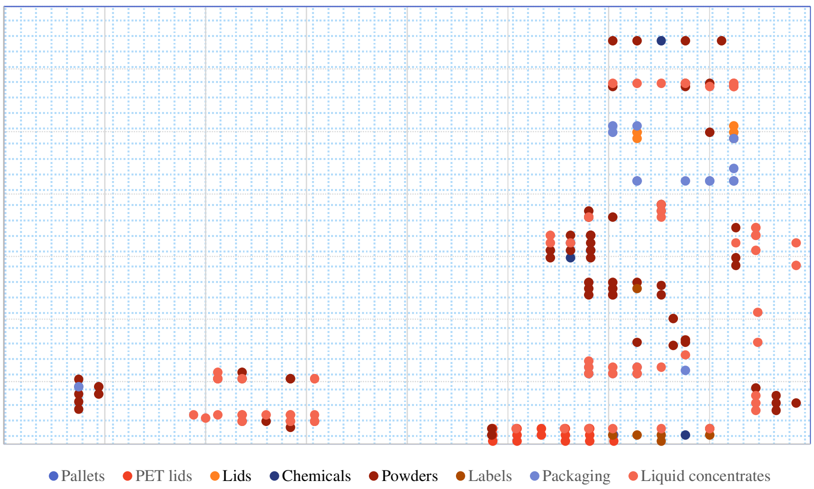
RESULTS

The data on the SKUs were obtained through the SAP system and processed in spreadsheets, identifying the inputs and outputs during a production period of one year. With the information from the SAP system, the frequencies and quantities of materials required by the production area were also calculated, in addition to the volume occupied and its location of each SKU. Table 1 shows the identification, location, quantity, and volume occupied of the SKUs present in the warehouse.

Table 1. Inventory and characteristics

Identification (family)	Number of SKUs	Location		Volume (in m^3)
		Internal	Internal/External	
Chemicals	6	x		12.32
Powders	65		x	519.29
Liquid concentrates	67		x	317.53
Lids	5		x	1.97
Pallets	1		x	1,169
Packaging	15	x		2.64
Labels	24	x		373.93
PET lids	9		x	0.53
Cardboard	11		x	641.71

Source: own work

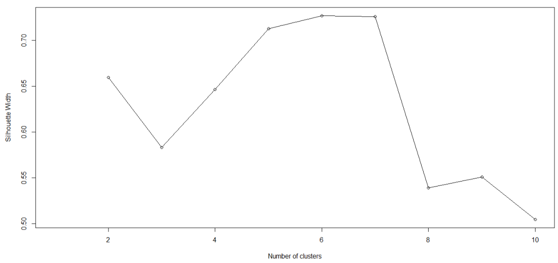


Source: own work

Fig. 2. Locations of the SKUs

Internal / external warehouse locations were considered those in which an SKU can be found, either in an internal warehouse, or in an external warehouse, mainly due to capacity and space constraints. Figure 2 presents the current locations of the SKUs in the internal warehouse.

Using the PAM clustering method, the SKUs were grouped based on the factors of DRC, AQO, PF, DRR, ARO, and RF, including two qualitative factors related to location (racks and floor) and the management unit (kilograms, pieces or gallons). Using the R studio software, modeling was performed with PAM, determining the number of clusters. By analyzing the similarity in the dataset and implementing the Gower distance, a similarity matrix was created. Establishing a similarity criterion allows the similarity of the elements to be related to each other; therefore, the proximity of an element is determined employing a similarity measure.



Source: own work

Fig. 3. Silhouette analysis

Subsequently, using the similarity matrix, with silhouette analysis, the size of the cluster was defined. With six clusters an average silhouette width of 0.73 was achieved, this being the highest result for the number of clusters (Figure 3).

Examining the context of each cluster, the solution with six clusters was the one that best adjusted to the diversity of data and requirements for warehouse administration. With the PAM clustering approach, the SKUs of the warehouse were grouped into six clusters as shown in Table 2.

Table 2. Quantitative factors

	Cluster					
	1	**2**	**3**	**4**	**5**	**6**
DRC	15.21	2.44	133.76	65 502.20	7 168.41	93 90.80
AQO	107.54	15.11	668.61	233 365.76	39 902.60	64 417.42
PF	0.14	0.15	0.18	0.28	0.25	0.09
DRR	1.72	0.2	17.6	12 428.9	828.2	1769.6
ARO	35.57	3.50	417.23	176 160.45	16 162.93	31 891.31
RF	0.04	0.05	0.04	0.08	0.05	0.04
# SKU´s	117	11	21	14	24	16
%	58	5	10	7	12	8
Volume (m^3)	616	33.65	841.08	23.64	1392	132.09
Average volume (m^3)	5.26	3.05	40.0	1.68	58.0	8.0

Source: own work

The SKUs were distributed as follows: 58% in Cluster 1, 5% in Cluster 2, 10% in Cluster 3, 7% in Cluster 4, 12% in Cluster 5 and 8% in Cluster 6. The factors with high values in PF and RF were clusters 4 and 5, which correspond to the families of labels, covers, and packaging materials. Cluster 6 includes families of label and packaging materials with the lowest PF. The DRC, AQO and DRR values are considerably higher in clusters 4, 5 and 6 because it corresponds to SKUs such as labels, can lids and packaging, which are used in large quantities during production. Cluster 5 includes, in addition to labels and packaging, the pallet family, so the volume occupied is greater than in other clusters. Without considering the family of pallets in cluster 5, cluster 3 is the one with the highest volume occupied in the warehouse. Cluster 1 is the one that contains the highest number of SKUs with 117, represented by materials from the families of chemicals, powders, and liquid concentrates.

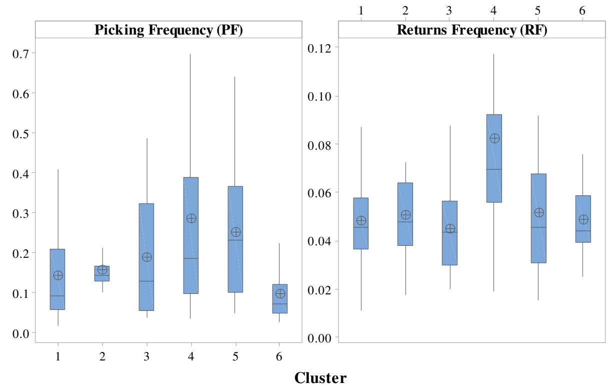
For qualitative factors, with the PAM grouping method, the results shown in Table 3 were obtained.

Table 3. Qualitative factors

		Location		Unit of measure		
		Floor	Rack	Kilograms	Pieces	Gallons
Cluster	1		x	x		
	2		x			x
	3	x		x		
	4	x			x	
	5	x			x	
	6		x		x	

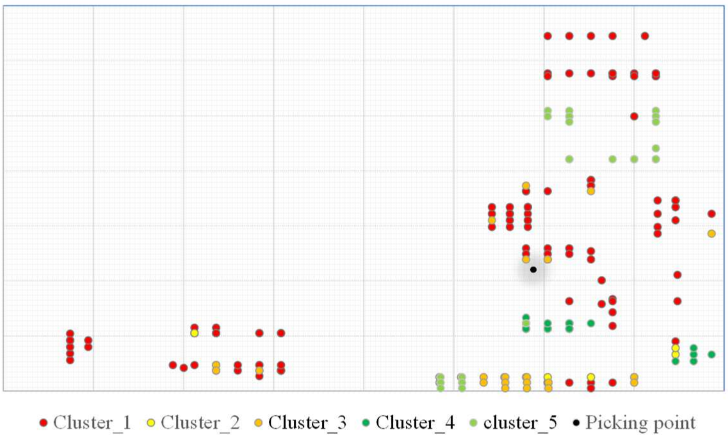
Source: own work

For example, for Cluster 2, the SKUs that form a group, belonging to families of liquid and chemical concentrates, that are in racks and identified as gallon material handling units. Also, these qualitative factors shown through the different clusters, allow the company to relate the characteristics of the materials and movement requirements with the systems and capacities of the equipment for the handling of materials. Based on the features that make up each of the grouped groups, for the location of the optimal collection point, and applying the Pareto principle, a weight is assigned to each of the factors. The weights give relevance to those factors that affect with a higher score towards the achievement of the different performance indicators, these weights are applied through equation (9) taking values of vi to reduce the variable d. In this case, the picking frequency and return frequency factors were determined by the company, as the main factors that favor efficiency in order processing. Eliminating outliers, as shown in Figure 4, clusters 4 and 5 were the main ones to consider for the PF and RF factors.



Source: own work

Fig. 4. Main factors



Source: own work

Fig. 5. Proposed warehouse layout

The company ranked clusters 4 and 5 with the highest weights, followed by clusters 2 and 3, and with the lowest weight to cluster 1. Regarding cluster 6, the company considered sending these materials to an external warehouse, such as the effect of low PF levels, a situation that favors the capacity of the internal warehouse. Applying equations (8) and (9), using the Facility Location Optimizer module of MATLAB software, the location of the optimal picking point was calculated. Graphically the solution is shown in Figure 5.

With this layout proposal, the calculated picking point was based on the weights of the clusters, with relevance in the picking frequency and warehouse returns, giving priority to these factors. This location is the point that minimizes the transfer distance for the preparation of the different SKUs, a situation that supports the improvement of logistics performance within the warehouse. With this distribution, the development of materials favors clusters 4 and 5, combining a policy of cluster collection and picking point, it is also possible to determine the volume required for the operation. The proposed picking point minimizes the routes of materials from clusters 4 and 5, with an average volume capacity of 60 m3. The configuration in Figure 5 also shows the distribution of the clusters throughout the warehouse, assisting in making decisions about planning the necessary space, depending on the volume occupied by each SKU that makes up the clusters. Compared to the current distribution of the company, this distribution proposal based on a cluster allows a reorganization of the SKUs increasing the storage capacity by 8%, in addition to avoiding the dispersed distribution of materials, going from a classification by families to a cluster classification.

CONCLUSIONS

The warehouse is an important component in the supply chain, due to reasons that include, among others, the fluctuations of demand and value-added service to the customer. Space, time, and costs are pillars for measuring storage efficiency. By optimizing inventory management, costs and time are minimized. With this proposal, it was possible to implement a methodology for the optimal identification and location of materials, without the need for expensive information systems, with a focus on the characteristics and factors that affect order preparation operations. The use of data mining techniques such as PAM clustering makes it possible to group the inventory into different clusters considering both qualitative and quantitative factors.

This article demonstrated how the variables of daily consumption rate, average quantity per order, picking frequency, the daily rate of return, the average number of returns per order, and frequency of return could be integrated into a distribution design, also including attributes related to the handling of materials within the warehouse.

The clustering proposal with PAM offers a more realistic approach to the problem of inventory management, where factors as diverse as time and capacities must be considered, to the types and handling that must be had with the materials inside the warehouse. The traditional approach to storage design issues for inventory management ignores the dynamic nature of customer demand. With this proposal, decision-makers in the company can analyze the dynamic environment of orders using factors such as picking frequency and return frequency. Also, with this analysis, the characteristics and qualities of the inventories can be periodically reviewed and the locations of the SKUs can be modified to benefit the improvement in supply logistics.

The inventory supply and administration process, through PAM, allows adapting the material selection environment, increasing the collection speed, and reducing the distance traveled. Furthermore, the results support decision-making on storage capacity, allowing the space required by the materials that make up the different clusters to be identified.

By combining the optimization of the picking point, a considerable benefit was achieved for the company, not only in streamlining the order preparation process but also in reducing the costs related to inventory management. By minimizing the transfer distances and using the identification of the materials, it is possible to fulfill orders faster and with high levels of satisfaction for different customers. Warehouse design decisions are another element that should be considered in inventory management, as it affects various aspects related to performance, including material handling, space cost, and capacity. As additional proposals, this study could be extended to the optimization of the warehouse design considering other factors, such as the routing for the collection, the delivery dates of the orders, the definition of picking areas, and the policies for the storage of materials.

ACKNOWLEDGMENTS AND FUNDING SOURCE DECLARATION

This research was supported by PRODEP [grant number UAEH-EXB-152] and CONACYT.

REFERENCES

Akay Ö., Yüksel G., 2018. Clustering the mixed panel dataset using Gower's distance and k-prototypes algorithms, Communications in Statistics - Simulation and Computation, 47(10), 3031–3041. http://doi.org/10.1080/03610918.2017.1367806.

Aktepe A., Ersoz S., Turker A., Barisci N., Dalgic A, 2018. An Inventory classification approach combining expert systems, clustering, and fuzzy logic with the Abc method, And An Application, South African Journal of Industrial Engineering, http://doi.org/10.7166/29-1-1784.

Anđelković A., Radosavljević M, 2018. Improving the order-picking process through the implementation of warehouse management system, Strategic Management, 23(2), 3–10, http://doi.org/10.5937/StraMan1801003A.

Aqlan F, 2017. Dynamic clustering of inventtory parts to enhance warehouse management, European J. of Industrial Engineering, 11(4), 469, http://doi.org/10.1504/EJIE.2017.086184.

Arora S., Chana I, 2014. A survey of clustering techniques for big data analysis. 2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence), 59–65, http://doi.org/10.1109/CONFLUENCE.2014.6949256.

Bottani E., Volpi A., Montanari R., 2019. Design and optimization of order picking systems: An integrated procedure and two case studies, Computers & Industrial Engineering, 137, http://doi.org/10.1016/j.cie.2019.106035.

Çelik M., Süral H, 2019. Order picking in parallel-aisle warehouses with multiple blocks: complexity and a graph theory-based heuristic, International Journal of Production Research, 57(3), 888–906, http://doi.org/10.1080/00207543.2018.1489154.

Choi T.-M., Wallace S.W., Wang Y, 2018. Big Data Analytics in Operations Management, Production and Operations Management, 27(10), 1868–1883, http://doi.org/10.1111/poms.12838.

de Vries J., de Koster R., Stam D, 2016. Exploring the role of picker personality in predicting picking performance with pick by voice, pick to light and RF-terminal picking, International Journal of Production Research, 54(8), 2260–2274. http://doi.org/10.1080/00207543.2015.1064184

Djatna T., Hadi M. Z, 2017. Implementation of an ant colony approach to solve multi objective order picking problem in beverage warehousing with drive-in rack system, 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 137–142, http://doi.org/10.1109/ICACSIS.2017.8355024

Faia Pinto A.R., Nagano M.S, 2019. An approach for the solution to order batching and sequencing in picking systems, Production Engineering-Research and Development, 13(3–4), 325–341, http://doi.org/10.1007/s11740-019-00904-4

Gower J.C, 1971. A General Coefficient of Similarity and Some of Its Properties. Biometrics, 27(4), 857, http://doi.org/10.2307/2528823

Grosse E.H., Glock C.H, 2015. The effect of worker learning on manual order picking processes, International Journal of Production Economics, 170, 882–890, http://doi.org/10.1016/j.ijpe.2014.12.018.

Gu J., Goetschalckx M., McGinnis L.F, 2007. Research on warehouse operation: A comprehensive review, European Journal of Operational Research, 177(1), 1–21, http://doi.org/10.1016/j.ejor.2006.02.025.

Guo X., Yu Y., De Koster R.B.M, 2016. Impact of required storage space on storage policy performance in a unit load warehouse, International Journal of

Production Research, 54(8), 2405–2418, http://doi.org/10.1080/00207543.2015.1083624.

Hong S, 2019. A performance evaluation of bucket brigade order picking systems: Analytical and simulation approaches, Computers & Industrial Engineering, 135, 120–131, http://doi.org/10.1016/j.cie.2019.05.037.

Jemelka M., Chramcov B., Kriz P., Bata T, 2017. ABC analyses with recursive method for warehouse, 2017 4th International Conference on Control, Decision and Information Technologies (CoDIT), 0960–0963, http://doi.org/10.1109/CoDIT.2017.8102722.

Kaufman L., Rousseeuw P, 2005. Introduction. In Finding Groups in Data: An Introduction to Cluster Analysis, (pp. 1–67), John Wiley, Hoboken, http://doi.org/10.1002/9780470316801.ch1.

Kusrini K, 2015. Grouping of Retail Items by Using K-Means Clustering, Procedia Computer Science, 72, 495–502, http://doi.org/10.1016/j.procs.2015.12.131.

Liu J., Liao X., Zhao W., Yang N, 2016. A classification approach based on the outranking model for multiple criteria ABC analysis, Omega, 61, 19–34, http://doi.org/10.1016/j.omega.2015.07.004.

Lolli F., Ishizaka A., Gamberini R, 2014, New AHP-based approaches for multi-criteria inventory classification, International Journal of Production Economics, 156, 62–74. http://doi.org/10.1016/j.ijpe.2014.05.015.

Matthews J., Visagie S.E, 2019 SKU arrangement on a unidirectional picking line. International Transactions in Operational Research, 26(1), 100–130, http://doi.org/10.1111/itor.12550.

Tsai C.-W., Lai C.-F., Chao H.-C., Vasilakos, A. V, 2015, Big data analytics: a survey. Journal of Big Data, 2(1), 21, http://doi.org/10.1186/s40537-015-0030-3.

van Gils T., Ramaekers K., Braekers K., Depaire B., Caris A, 2018. Increasing order picking efficiency by integrating storage, batching, zone picking, and routing policy decisions, International Journal of Production Economics, 197, 243–261, http://doi.org/10.1016/j.ijpe.2017.11.021.

van Gils T., Ramaekers K., Caris A., Cools M, 2017. The use of time series forecasting in zone order picking systems to predict order pickers' workload, International Journal of Production Research, 55(21), 6380–6393, http://doi.org/10.1080/00207543.2016.1216659.

van Gils T., Ramaekers K., Caris A., de Koster R.B.M., 2018. Designing efficient order picking systems by combining planning problems: State of the art classification and review, EUROPEAN JOURNAL OF OPERATIONAL RESEARCH, 267(1), 1–15, http://doi.org/10.1016/j.ejor.2017.09.002.

Xindong Wu, Xingquan Zhu, Gong-Qing Wu, Wei Ding, 2014. Data mining with big data, IEEE Transactions on Knowledge and Data Engineering, 26(1), 97–107, http://doi.org/10.1109/TKDE.2013.109.

Yuan R., Cezik T., Graves S.C, 2018. Stowage decisions in multi-zone storage systems, International Journal of Production Research, 333–343, http://doi.org/10.1080/00207543.2017.1398428.

Zhang R.-Q., Wang M., Pan X, 2019. New model of the storage location assignment problem considering demand correlation pattern, COMPUTERS & INDUSTRIAL ENGINEERING, 129, 210–219, http://doi.org/10.1016/j.cie.2019.01.027.

ZARZĄDZANIE ZAPASEM ORAZ OPTYMALIZACJA LOGISTYKI - PODEJŚCIE OPARTE NA EKSPLORACJI DANYCH

STRESZCZENIE. **Wstęp:** W obecnych warunkach ekonomicznych, charakteryzujących się wysoką konkurencyjności i nieprzewidywalności działalności w obrębie łańcucha dostaw, istotne jest dążenia do optymalizacji kosztów i poprawy poziomu obsługi klientów, poprzez prawidłowe zarządzanie zapasem, jako czynnikiem kluczowym. Właściwe zarządzanie zapasami prowadzi do pozytywnych wpływów na wyniki logistyczne. W ujęciu ekonomicznym, około 15% kosztów logistycznych jest związane z operacjami magazynowymi. Poprzez zastosowanie studium przypadku z branży spożywczej, w pracy proponowana jest metoda klasyfikacji zapasów z zmiennymi jakościowymi i ilościowymi, przy zastosowaniu technik eksploracji danych, kategoryzując materiały przy użyciu zmiennych takich jak częstotliwość pobrań, poziom konsumpcji, jak i charakterystyki jakościowe związane z operacjami magazynowymi. Proponowany model łączy klasyfikację materiałową z technikami lokalizacyjnymi w celu ułatwienia procesu decyzyjnego w obszarze zarządzania zapasem oraz operacji magazynowych.
Metody: Zastosowana metoda opiera się na algorytmie Partitioning Around Medoids, który w innowacyjny sposób, stosuje strategię lokalizacji optymalnego punktu poboru w oparciu o klasyfikację klastrową, uwzględniając jakościowe jak i ilościowe czynniki, mające duży wpływ na określanie priorytetów w zarządzaniu zapasem w przedsiębiorstwie.
Wyniki: Uzyskane wyniki poprawiają marszruty dystrybuowanych materiałów w oparciu o identyfikację ich charakterystyk takich jak częstotliwość pobrań i handlingu, pozwalając na reorganizację i wzrost pojemności magazynowej różnych indeksów materiałowych, przechodząc z klasyfikacji na podstawie rodzin do klasyfikacji opartej na klusterze. Dodatkowo, wyniki wspomagają proces decyzyjny związany ze zdolnościami magazynowymi, umożliwiając identyfikację na najniższym poziomie miejsca magazynowego.
Wnioski: Praca prezentuje podejście do poprawy procesu decyzyjnego w zarządzaniu zapasem poprzez propozycję projektu magazynu w oparciu o techniki eksploracji danych, które stosują mierniki i wskaźniki działań operacyjnych. Zastosowanie technik eksploracji danych takich jak klustrowanie PAM umożliwia grupowanie zapasów przy uwzględnieniu różnych czynników jakościowych i ilościowych. Proponowana metod PAM umożliwia bardziej realistyczne podejście do problemów zarządzania zapasem, gdzie muszą być uwzględnione tak różne czynniki jak czas czy zdolności oraz typu operacji magazynowych.

Słowa kluczowe: Cluster, Partitioning Around Medoids, lokalizacja zasobów, łańcuch dostaw

Rafael Granillo-Macías ORCID ID: https://orcid.org/0000-0002-1015-667X
Autonomous University of Hidalgo State, UAEH
Campus Sahagun
Department of Industrial Engineering
Corredor Industrial, Ciudad Sahagun, Hidalgo, **México**
e-mail: rafaelgm@uaeh.edu.mx