Koc. J. Sci. Eng., 1(2): (2018) 21-27

Kocaeli University

Kocaeli Journal of Science and Engineering
(KOJOSE)

http://dergipark.gov.tr/kojose

Placement of Products and Improvement of Order Picking Process through Association Analysis: A Case Study in Pharmaceutical Warehouse

Fatma DİNÇER [1,*], Alpaslan FIĞLALI[1]

[1] *Department of Industrial Engineering, Kocaeli University, Kocaeli, 41310, Turkey*

Article Info

Research paper

Received : *28/05/2018*
Accepted : *12/06/2018*

Keywords

Apriori Algorithm
Association Analysis
Data Mining
Order Picking

Abstract

While taking strategic decisions for the future in today's business world, which has a constantly changing and dynamic structure, several implications can be drawn with the information obtained from databases and the correct processing and analysis of this information. With regard to these implications, various decisions can be taken to improve the processes. Businesses can execute their processes more efficiently and they are able to make better decisions on the future by processing the data that will benefit by using the data mining techniques. In this study, the past orders of a pharmaceutical company were analyzed and an association analysis of the products in these orders was performed and a methodological framework has been presented based on the Apriori Algorithm results to ensure that the products are placed at the optimum level in the warehouse. Therefore, this study contributes to the improvement of order picking process.

1. Introduction

Warehouses and storage processes have become an integral part of production systems today. Being flexible and responsive to ever changing customer needs is essential in the dynamic business world. In order for the products to reach end users, raw materials, semi-finished products and products must be transported from one place to another. The products are kept in certain warehouses till customers demand them. Warehouses are not only the places where raw materials, semi-finished products and products are kept but also business units where goods are accepted, placed on shelves, orders are picked, replenishment, barcoding, packaging and shipment operations are carried out. When the processes from goods acceptance to shipment in warehouses are examined in detail, these may vary depending on the sector, the customer and the product, but it can also be said that basically many activities are common. In order to efficiently manage warehouse operations, it is very important that raw materials, semi-finished products and products are stored in the optimum way and can be prepared as soon as they are demanded by the customers. Delivering orders in a timely and accurate manner is an important criterion both for an efficient operation and for the image of the company. In the study by Van den Berg and Zijm, where the basic warehouse processes are addressed in four steps, it was determined that order picking costs accounted for more than 60% of the total cost of these four basic storage activities as seen in Fig. 1 [1]. When this data is evaluated, it can be said that order picking is the most costly process in warehouse operations.

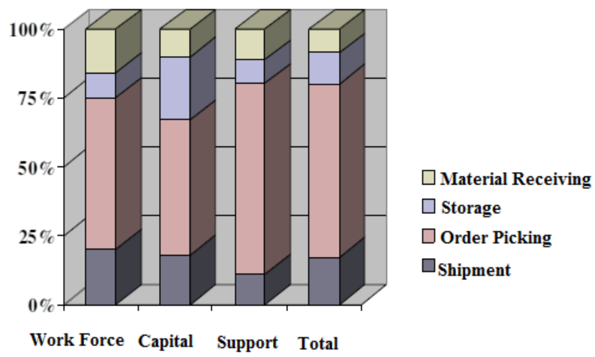

Figure 1. Storage costs according to activities.

[*] Corresponding Author: fatmatacar88@gmail.com

In this study, the products in a drug company's warehouse were placed according to the results of "Association Analysis", one of the most widely used methods in Data Mining, and the order picking process was improved. The possibility of products to be included in the orders was determined through "Apriori Algorithm" using past data stacks the layout of the products in the warehouse was rearranged in the direction of the data obtained from the analysis. Random orders were selected for products and the distances the products covered were calculated when the order picking process was performed according to the current layout. The measurements of the same orders were made again for the improved layout in the direction of the data obtained after the association analysis; the results were compared and the improvement rates were shared.

This study consists of five sections. The first section is the "Introduction" section and the second section is the "Literature" section. In the third section, "Association Analysis" and Apriori Algorithm which are data mining applications are explained and in the fourth section, data sets are analyzed. In the final section includes the results.

2. Literature Review

One of the processes that have the highest priority among the costs that must be met in warehouse management is the order picking [1]. Efficient warehouse processes and the efficient use of warehouse are among the biggest goals of the enterprises. The literature review revealed that generally the improvements made in the warehouse with the back to back shelf system are striking in the association analysis of the warehouses. The association analysis seems to be used in general to determine the tendency of customers to products. Liao and Chen proposed the product maps they obtained using the Apriori algorithm as a new product development source [2]. Yang and Lai compared the performance of decisions in product promotion based on the information obtained from online shopping behaviors [3]. Ay and Çil formed association rules in a retail chain operation [4]. They proposed the premises order in line with the data obtained from the association rules formed in the study. They worked on the use of data in layout planning. A study by Kılınç presented a method for the association analysis. The rules generated by the Apriori algorithm were eliminated and applied in the production and goods acceptance quality data in an electronics firm [5]. Koç et al. carried out research to examine the use of social networking by students of department computer teaching. Survey questions were asked to students to determine the extent to which social networks affected them and the association analysis was applied to the data obtained [6]. A study by Yurtay et al. aimed to determine why computer engineering students choose technical elective courses [7]. Doğan et al. conducted an association analysis on a data set related to the customers of an insurance company in their study. At the end of the study, it is reported that the types of insurance that can be purchased together were determined and information that can guide the activities for marketing were obtained [8]. Doğrul et al. utilized association analysis rules to analyze the data on traffic accidents in their work. In this study, the places and times where the accidents often occurred were determined and it was stated that in such cases the accidents can be prevented by increasing the measures [9]. K. Kaur and Kang noted that the Apriori algorithm works on static data stacks that do not take the time into account and suggested a new algorithm for it. The algorithms they suggested take into account not only static properties but also time-varying properties [10]. Yener et al. used association analysis and genetic algorithm methods to solve the order stack problem. Due to the fact that there are orders covering similar products within the stacks formed by the combination of multiple orders, they tried to save time and distance lost during the order picking [11]. When we look at the literature, there are association analysis studies to determine the tendency of customers to buy in marketing and to make layout arrangements in this direction. No study was found in the literature review in which the association analysis was applied directly according to the order rate of the products and the products in the drug warehouse were placed with the Apriori Algorithm.

3. Method

3.1. Association Rules

Due to the continuous and rapid development of computer systems nowadays, data can be stored digitally for many years. With too much increase in the data stacks and the need to analyze these stacks and produce meaningful results from the data, the concept of data mining emerged. Innovative businesses throughout the world use data mining to assess customers' needs and complaints, reorganize their products, or reduce their losses to a minimum [12]. The purpose of data mining is to analyze the data in the forms of stacks and to reveal the relations and tendencies between these data with various algorithms and to interpret the data obtained as a result of these data to produce meaningful results for the future [13]. One of the methods that allows the generation of these meaningful results is "Association Analysis". Association rules analyze a business's retrospective sales and orders so that the relationship between these orders is revealed and make forecasts for the future based on those results. This method was developed by Agrawal and Srikant at the IBM Almaden Research Center in 1994. There have been many studies carried out on this subject in recent years. Through the obtained associations, consumer-oriented activities such as shelf product designs, customer personal preferences, promotional arrangements can be done in a more orderly fashion [14]. Association rule algorithms are used in many areas such as economics, health, storage and banking. They are often used in areas where customer-based databases are available; they are also preferred to reveal a variety of relationships in communication systems where there is a lot of data [15].

Association rules were mathematically expressed by Agrawal and Srikant as follows [16]:

- $I = \{I_1, I_2, \ldots I_m\}$ represents an array-object set
- $T = \{ t_1, t_2, \ldots t_n \}$ represents the operations (order) in the database. The value for each t_k is 0 or 1. If $t_k = 0$, it means I_k was not purchased; if $t_k = 1$, it means I_k was purchased. There is a separate record in the database for each operation. The t_k value corresponding to each I_k in X is $t_k = 1$.

This association is expressed with its rule in the following way:

- $X \Rightarrow I_J$ is a sub-set of X, I. I_j is any element in the I and this element is covered by X. In order to say that $X \rightarrow I_J$, rule is appropriate for T, it is necessary to mention a certain level of confidence. That is, how many of all the X's obtained for T cover I_k should be expressed in terms of % c. In this case, the association rule with $0 \leq c \leq 1$ confidence level can be expressed as follows:
- $X \Rightarrow I_J \quad X \rightarrow I_J \,|\, c$

Confidence level refers to the power of a defined rule. The rule also has the concept of support level. The support level is a concept that indicates how much of the operations in T provide X. Revealing the association rules through analyzing the data in a database is the determination of rules with greater confidence and support levels than the smallest support level and the smallest confidence level that the user would give. The object sets that provide the smallest support level are called the large object set, while others are called the small object set [16].

The support criterion indicates how often an association is in the data and confidence value means by what probability a product is likely to be purchased along with another product. Each rule is expressed by a value of support and confidence [16]:

- $A \Rightarrow B$ [support = 3%, confidence = 70%]

The 3% support value for the association rule means that 3% of the entire shopping have A and B products sold together. A confidence level of 70% indicates that 70% of customers who bought product A also bought product B in the same shopping.

Various algorithms have been developed to make such relations and derive rules. The Apriori algorithm is one of them.

3.2. Apriori Algorithm

The Apriori algorithm, a simple and well-known algorithm for deriving association rules from data sets, was developed in 1994 by Agrawal et al. has been the most widely applied algorithm for deriving association rules in data mining [13]. With Apriori algorithm, it is ensured that association rules with support and confidence over a specified threshold value are derived [17].

It has a recursive nature and the database is scanned many times to identify frequently repeated data sets in the databases. In order for the association rules in large data sets to be analyzed quickly and correctly, Agrawal, Srikant, et al. developed the Apriori algorithm. With Apriori algorithm, association rules data mining is performed on real data. The Apriori algorithm provides an effective solution by eliminating some candidate object clusters without counting the support values of them [16].

Agrawal and Srikant explain the details of the operation of the algorithm as follows [16]:

- For the detection of large object clusters during the first scan of the data, all objects are counted.
- If the next scan is regarded as the kth scan, it consists of two steps;
- Apriori-gene function is used to construct C_k candidate object clusters with L_{k-1} object clusters obtained in the (k-1)th scan,
- The database is then scanned and the support of candidates in C_k is counted.
- For a quick count, candidates constituting C_k in a given I process must be well known.

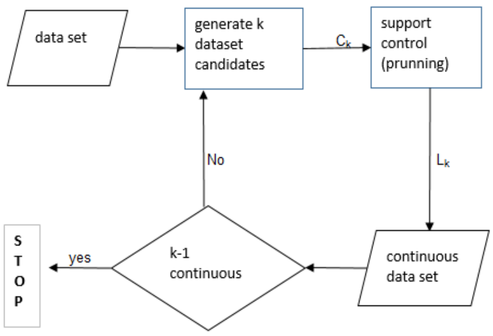

Figure 2. Apriori algorithm flow.

The apriori algorithm is often used to determine inter-product relationships in markets and to determine the future trends of customers' future decisions. Apriori algorithm is used on many institutional resources in the world, and it is possible to find many examples in Turkey as well. The Apriori algorithm can also be used to identify associations to optimize layouts in warehouses.

4. Results and Discussion

4.1. Work Study

This study aims to analyze the past sales data of a pharmaceutical company and to place the products in the warehouse optimally according to the associations that will arise as a result of this analysis. As a result of this optimum placement, it is aimed to reduce distances while order picking is being carried out. The greatest factor affecting the order picking period is the distance between the products in the order. As the operators carrying out the order picking collect the products in the order list, the more distance between the products, the more they will be walking in the warehouse. For this reason, it is important to focus on placing the largest selling products near the exit, as well as the possibility of the products being ordered together in an order. If we place products that are likely to be ordered together in the same order close to each other, the distance operator will walk while performing the order picking process will be minimized. Sales order data was received from the SAP system and these orders were analyzed on an item basis. While doing this process, the data taken from the system was primarily analyzed on the order and product code basis on excel since it covers a significant amount of data. The "Association Analysis" method was utilized to calculate the likelihood of products being ordered together in the same order. As a result of this analysis, the likelihood of products being ordered in the same order was determined, and the products that were most likely to be ordered together were placed in close proximity to each other.

4.2. Modelling

The data from the last 3 years taken from SAP have been reviewed separately and made available for data processing in SPSS Clementine. In order to run the Apriori algorithm in SPSS Clementine, the data must be tabulated. Pivot tables were created to determine which products were ordered in which orders. The data obtained from these pivot tables was converted to tables. In the table format obtained in Excel, order numbers were used in the rows and product codes in the columns, and the data was made available for processing in the SPSS Clementine. The table has a section from the dataset converted to table format. It is not possible to list the entire dataset here. For this reason, a section has been presented in order to understand the data set.

Table 1. A section of the dataset to be processed in SPSS Clementine.

Order	A14000730	A14000984	A14000525	A14000533	A14000547	A14000697	A14002398	A14002699	…	A14003138	A14003151
1192543	0	0	1	0	0	0	1	1	…	1	0
1192545	1	0	0	0	0	0	0	0	…	0	0
1192550	0	0	0	0	0	0	0	0	…	0	0
…	…	…	…	…	…	…	…	…	…	…	…
1192551	0	0	0	0	0	0	0	1	…	0	0
1193587	0	0	0	0	0	1	0	0	…	0	0
1192588	1	0	0	0	0	0	0	0	…	0	0
1192589	0	1	0	0	0	0	0	0	…	0	0
1192590	1	0	0	0	0	0	0	0	…	0	0

In this study, Apriori algorithm, one of the algorithms used in association analysis, was utilized. The installation of the model in the SPSS Clementine module is shown in Fig. 3.

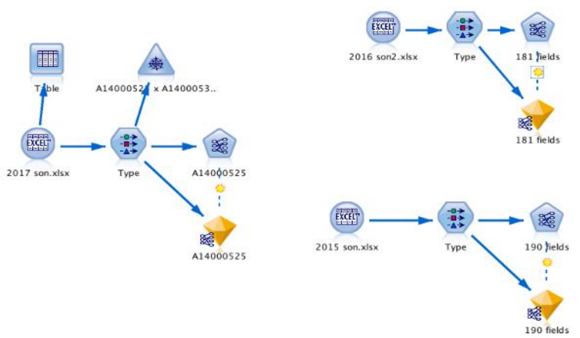

Figure 3. Installation of the model in SPSS.

Apriori algorithm was applied in SPSS by evaluating the data of the last 3 years drawn from SAP separately. There were 180 different products and the ratios of each of these products in total sales and the percentage of products in the order were determined and the average of these two values was found. According to these values, rates of 1.5% and above were taken and 27 different product codes meeting this value were identified. Studies were carried out on these identified products. As a result of the examination of the data, the support value was determined as 2% and the confidence value as 10%.

When we examine the analysis conducted for the product coded A14000523; the possibility of customers who order this product to order products A14000525, A14002766, A14000984, A14000533 etc. in the same order is shown in Fig. 4. As a result of this analysis, the probability of the product coded A14000525 is being ordered in the order in which the product coded A14000523 ordered is 43.275%.

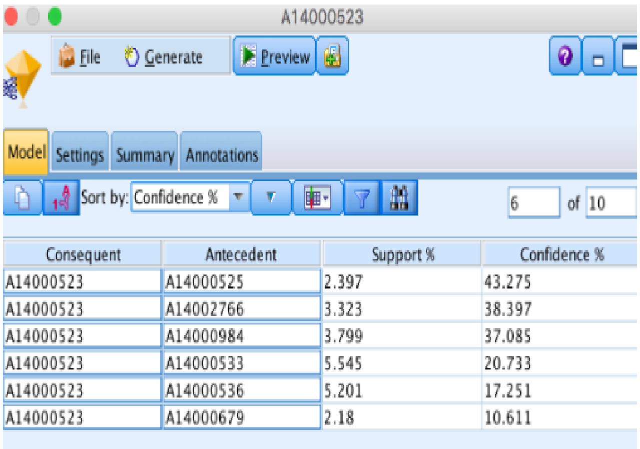

Figure 4. Association analysis according to product code.

These operations were made separately for the 27 different products previously identified. Association analysis revealed that there were associations in products other than the previously identified 27 different products.

In the current lay out, the products are arranged as shown in Fig. 5. In this layout each color represents a different product.

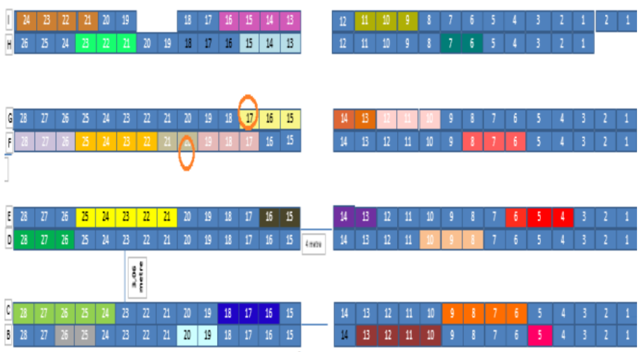

Figure 5. Current layout of products in warehouse.

Table 2 shows the walking distances for an order in the warehouse.

Table 2. Order picking by current layout.

S.N	Stages of Work	Distance (m)	Products on Order
1	Taking and controlling the order	****	
2	Controlling the location of products with RF handheld terminal	****	
3	Picker pick up	7.2	
4	Taking empty pallet	4.3	
5	Going to the product location -1 (F1028)	43.1	
	Picking the products on order	****	
	Going to the product location -2 (G1034)	7.5	1-Product A14000536 (300 Pcs)
	Picking the products on order	****	
6	Bringing the products to the closing area	35.8	2-Product A14000587 (120 Pcs)
7	Barcode reading process of products	****	
8	Labeling	****	
Total		**97.9**	

Then the products were ranked according to the support ratio and the first 4 products were identified. The association rules for some of these 27 products are shown in Fig. 5. The last 2, 3 and 4 digits of the product codes were taken in the visual.

Figure 6. Quadruple association analysis of products

Table 3 shows the walking distances measured for the corresponding order at the proposed layout.

Table 3. Order picking by layout after association analysis

S.N	Stages of Work	Distance (m)	Products on Order
1	Taking and controlling the order	****	
2	Controlling the location of products with RF handheld terminal	****	
3	Picker pick up	7.2	
4	Taking empty pallet	4.3	
5	Going to the product location -1 (G1012)	15.8	
	Picking the products on order	****	
	Going to the product location -2 (H1020)	6.3	1-Product A14000536 (300 Pcs)
	Picking the products on order	****	
6	Bringing the products to the closing area	28.3	2-Product A14000587 (120 Pcs)
7	Barcode reading process of products	****	
8	Labeling	****	
Total		**61.9**	

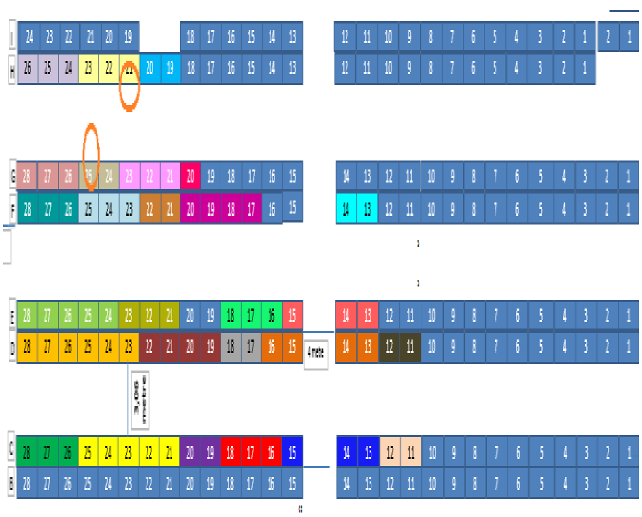

Figure 7. Proposed layout for the products

The layout of the products according to the association analysis is as shown in Fig. 6.

For 10 different orders, the distance covered during the order picking was measured primarily according to the current layout according to the layout created by the association analysis result. Then, the distance covered in the warehouse for these 10 different orders was measured. Table 4 shows the distance covered and the improvement rates according to the current layout and the improved layout. In current situation, it is necessary to walk more distance while collecting related orders. In proposed situation, the distance traveled when collecting the related order is reduced. The order picking processes has been improved.

Table 4. Improvement rates of the distances walked

	Total Walking Distance While Order Picking (m)		
Order	**Current Layout**	**Improved Layout**	**Improvement Rate**
1	160.9	121.5	24%
2	151.5	115.6	24%
3	97.9	61.9	36%
4	187.9	78.6	58%
5	151.5	50.1	67%
6	143.6	90.5	37%
7	122.2	111.0	9%
8	149.6	131.4	12%
9	196.9	156.4	21%
10	118.0	55.7	53%

5. Conclusions

Order picking is one of the processes that cause high costs in warehouse processes. It is important for the efficiency of the processes that the operators cover minimum levels of distances when orders are picked. Placing products that are likely to be in the same order close to each other ensures that the distance covered during the order picking operation is reduced to the minimum level; increasing the order picking efficiency. When Apriori algorithm is applied on the datasets that the company has, it has been observed that some product groups are in the same order. Association analysis ratios were determined for product groups. In accordance with these association rules, the existing layout of the products was changed and the products were replaced according to the results obtained from the association analysis results. It was observed after this new layout that the distance the operator covered while picking orders was reduced at certain rates and that the order picking process was improved. This study is an important and fruitful study showing the effect of association analysis on the location of products. The work to be done on this subject can be extended.

References

[1] Van den Berg J. P., Zijm W. H. M., 1999. Models for warehouse management: Classification and examples. Int. J. Production Economics, **519**, 519–528.

[2] Chen M. C., Lin C. P.,2007. A Data Mining Approach to Product Assortment and Shelf Space Allocation. Expert Systems with Applications, **32**, 976-986.

[3] Yang T. C., Lai H., 2006. Comparison of Product Bundling Strategies on Different Online Shopping Behaviors. Electronic Commerce Research and Applications, **5**, 295-304.

[4] Ay D., Çil İ., 2008. Use of Association Rules in Layout Planning AT Migros Türk A.Ş. Journal of Industrial Engineering, **21**, 14-29.

[5] Kılınç Y., 2009. Mining Association Rules For Quality Related Data In An Electronics Company. Master's Thesis, Middle East Technical University, Industrial Engineering, 1-23.

[6] Koç M., Karabatak M., 2011. Investigation of the Effect of Social Networks on Students Using Data Mining. 5th International Computer & Instructional Technologies Symposium, Fırat University, Elazığ-Turkey, 22-24 September.

[7] Güngör E., Yalçın N., Yurtay N., 2013. Technical Elective Course Selection Analysis with Apriori Algorithm. National Distance Education and Technologies Symposium, Selçuk University, Konya,1-3 November, 114-119.

[8] Doğan B., Erol B., Buldu A., 2014. Usage of Association Rules for Customer Relationship Management in the Insurance Sector. Marmara Journal of Science, **3**, 105-114.

[9] Doğrul G., Akay D., Kurt M., 2015. Analysis of Traffic Accidents by Association Rules. Gazi Journal of Engineering Sciences, **1**, 265-284.

[10] Kaur M., Kang S., 2016. Market Basket Analysis: Identify The Changing Trends of Market Data Using Association Rule Mining. International Conference on Computational Modeling and Security, Procedia Computer Science, **85**, 78-85.

[11] Yener F., Yazgan H., Cömert S., Kır S., Kaya Y., 2016. Solution of Order Batching Problem with Association Rules and Genetic Algorithm: A Case Study in Pharmacy Warehouse. Journal of Transportation and Logistics, **1**, 130-142.

[12] Edelstein H. A., Edelstein H. C., 1997. Building, Using, and Managing the Data Warehouse. In: Data Warehousing Institute, 1st edn., Prentice Hall PTR, Englewood, Cliffs.

[13] Han J., Kamber M., 2006. Data Mining: Concepts and Techniques, 2rd edn., Morgan Kaufmann Publishers.

[14] Karaibrahimoğlu A., 2014. Analyzing Breast Cancer Data Using Association Rule Mining. Phd Thesis, Selçuk University, Graduate School of Natural Sciences, 1-126.

[15] Kotsiantis S., Kanellopoulos D., 2006. Association Rules Mining: A Recent Overwiev. GESTS Int'l Transactions on Computer Science and Engineering, **78**, 71-82.

[16] Agrawal R., Srikant R., 1994. Fast Algorithms for Mining Association Rules. Proceedings of the 20th VLDB Conferance, Santiago, Chile, 487-499.

[17] Bothorel G., Serrurier M., Hurter C., 2011. Using Visual Data Mining Tools to Explore a Set of Association Rules. IHM '11 23rd French Speaking Conference on Human-Computer Interaction, Sophia Antipolis, France, October 24-27.