DATA ANALYTICS FOR RETAIL LAYOUT DECISIONS

Özlem Akçay Kasapoğlu,

Istanbul University, Faculty of Business
Operations Management Department, Istanbul, Turkey. Email: ozlemak@istanbul.edu.tr

ABSTRACT
There is hidden information locked up in the huge data of the companies' databases. The unidentified information is potentially important for the companies' success. The relations among the products of the retailers can be efficiently extracted from large retail databases which effects the sales and profitability. The operations managers can alter both with overall arrangements of the store and the allocation of the space to various products with this hidden information gathered by the data mining analysis. In this paper association rules are applied to the data set and rule sets are gathered to achieve a better layout and shelf arrangement for the retailer. Layout has strategic importance to a firm. Placement of the power items which have a very high exposure rate, maximize profitability per square foot of the floor space.

Key words: Layout, Data mining, Erp, Retail

INTRODUCTION
Retail layouts are based on the idea that sale and profitability vary directly with customer exposure product. The most retail managers try to expose as many products as possible. Studies show that the greater rate of exposure, the greater sales and the higher return on investment. The operations managers can alter both with overall arrangements of the store and the allocation of the space to various products within arrangements. In Retail/service layout the problem is to expose customer to high-margin items, allocating shelf space and responding to customer behavior. (Heizer, Render, 2008, p:351) In this study data mining tools are applied to the customer purchase data set sets of a furniture and decoration retailer. Association analysis is done to uncover interesting cross-sells and related products. The implemented technique used to find best combination of the products that should be exposed together in the shelves to increase the sales and higher return on investment. Those decisions effects the layout of the retail or which has strategic importance to a firm. Proper layout enables higher utilization of space, equipment and people, improved flow of information, materials, or people, improved employee morale and safer working conditions, improved customer/client interaction and flexibility.

In recent years, the field of data mining has seen an explosion of interest from both academia and industry. Increasing volume of data, increasing awareness of inadequacy of human brain to process data and increasing affordability of machine learning are reasons of growing popularity of data mining. Data mining is a set of automated techniques used to extract buried or previously unknown pieces of information from large databases, using different criteria, which makes it
possible to discover patterns and relationships. This new derived information can be utilized in the areas such as decision support, prediction, forecasting and estimation to make important business decisions, which can help in giving a particular business the competitive edge. The extraction of hidden predictive information from large databases is a powerful tool with great potential to help organizations to define the information market needs of tomorrow. Data mining tools predict future trends and behaviors, allowing businesses to make knowledge-driven decisions that will affect the company, both short term and long term. Data mining tools search databases for hidden patterns, finding predictive information that experts may miss because it was outside their expectations. At present the techniques are becoming more commonly used but have not been applied adequately in the store layout. Store layout problem is motivated by applications known as market basket analysis to find relationships between items purchased by customers. Mirajkar et al. (2016)

LITERATURE REVIEW
Yu et al. (2015) worked on retail design and the visually impaired: A needs assessment. This research explores the lived experience of consumers with visual impairment to better understand their everyday shopping challenges, gaps in retail design, and opportunities for improved service. Three focus group interviews were conducted with 17 informants. Data were audio-recorded, transcribed, and analyzed using QSR NVivo 8. Findings highlighted many challenges: mobility and getting into the store; judging quality; distinguishing colour; reading labels, store signage, and receipts; negotiating store lighting, merchandise layout, and fitting rooms; and interacting with sales associates. This paper identifies visually impaired shoppers’ need for universal retail design, discusses implications, and provides recommendations to retailers and product, graphic, and interior designers.

Vrechopoulos et al.(2004), studied virtual store layout: an experimental comparison in the context of grocery retail. Interface design and the interaction between customer and computer are factors critical to business effectiveness over the Web. A key aspect of conventional retailing is store layout; an analogous aspect in virtual retail is virtual store layout. This paper reports on an experimental investigation into the use of three different layouts in online grocery retailing: freeform, grid, and racetrack. These three most common conventional retailing layout types were transformed into virtual layouts for computer-mediated interfaces. Subjects in Greece and the UK participated in a laboratory experiment: they were given a planned shopping task with money to spend, and performed their shopping through a virtual store with layout as the manipulated variable. The results show that layout significantly affects online consumer behavior, but that predictions generated from the literature of conventional retailing about differences in the outcome of layouts do not generally hold in a virtual setting. Some of the findings can be explained by reference to research in Human Computer Interaction (HCI).

Hübner and Kuhn (2012) worked on retail category management, they did state-of-the-art review of quantitative research and software applications in assortment and shelf space management. Retail requires efficient decision support to manage increasing product proliferation and various consumer choice effects with limited shelf space. Our goal is to identify, describe and compare decision support systems for category planning. This research analyzes quantitative models and software applications in assortment and shelf space management and contributes to a more integrated modeling approach. There are difficulties commonly involved in the use of commercial software and the implementation and transfer of
scientific models. Scientific decision models either focus on space-dependent demand or substitution effects, whereas software applications use simplistic rules of thumb. They show that retail assortment planning models neglect space-elastic demand and largely also ignore constraints of limited shelf space. Shelf space management streams on the other hand, mostly omit substitution effects between products when products are delisted or out-of-stock, which is the focus of consumer choice models in assortment planning. Also, the problem sizes of the models are often not relevant for realistic category sizes. Addressing these issues, this paper provides a state-of-the-art overview and research framework for integrated assortment and shelf space planning.

Wood and Reynolds (2012) worked on leveraging locational insights within retail store development? Assessing the use of location planners’ knowledge in retail marketing. The retail geography literature has long recognised the importance of spatial and catchment analysis to inform decision-making relating to store development. However, less attention has been directed to store development “in practice” and, more specifically, how location research and geographical knowledge is leveraged across the wider retail business – in particular informing the marketing function. Through the use of a semi-structured interview and focus group methodology involving approximately 40 location planning, property and marketing analysts, they find that while some larger retailers have established close links between store development and marketing functions in the exchange of catchment, customer, competitor and loyalty card data to inform local marketing, product ranging, promotional mailings and post-opening store performance reviews, this tends to be the exception rather than the rule. They suggest there is a need for location planners to develop their intra-organisational legitimacy to engender a culture of knowledge-sharing and challenge the departmentalised, silo cultures that exist within some retailers in order to better leverage geographical insights and assist in the realisation of appropriate customer propositions and marketing strategies.

Tripathi and Dave (2013) made a store format choice and relationship quality in apparel retail: A study of young and early-middle aged shoppers in New Delhi region. Shopping format choice has been an interesting and important area of research in the academic literature from long time. However, research in this area is presently at a nascent stage in India. The present study compares three different retail formats (Discount stores, Exclusive stores and Multi-brand outlets) based on the shopper's perception of relationship quality (RQ) of these stores using multivariate techniques. For addressing, the objective of this study a sample of 313 shoppers is used. Results suggest that “conflict due to store” and “combined overall RQ due to the store and its employees” influence the store format choice. The RQ levels among the apparel store format are highly competitive.

Griffith (2005), made an examination of the influences of store layout in online retailing. Store layout is a critical factor driving consumer elaboration and response in retailing. While considerable attention has focused on store layout in relation to brick and mortar retailing, store layout has been virtually ignored in the study of online retailing, a critical limitation in the literature given its performance implications. In this study, information processing theory and aspects of the technology acceptance model (TAM) are used to theorize how two types of online store layout (tree and tunnel web site structures) influence consumer elaboration and response. Results of an experiment indicate that consumers exposed to a tree structured online store layout perceived the store to be easier to use, experienced greater elaboration of product related
information, had higher levels of product and brand recall, greater purchase intentions and a
more positive attitude toward the retailer than those exposed to a tunnel structure online store
layout. Implications for academics and practitioners are discussed.

Sipahi and Enginoglu (2015) worked on retail planning studies, they did an application oriented
at consumers’ perception of the quality of retail environment. The retail sector has a
continuously developing and changing structure. This change affects the economical, social,
cultural and geographical structure and at the same time it is affected from all of them. One of
the main targets of retail planning politics is creating an environment which will increase quality
of the shopping environment for consumers and will let every class of people to reach retail
possibilities easily with equal conditions. The aim of the study is determining the factors that is
affecting Kuşadası’s shopping environment in terms of consumer and consumers’ perception
of the quality of retail environment. Consumers’ perception of the quality of retail environment in
center of Kuşadası have been evaluated by four characteristics (characteristics of store, physical
environment, security and cleaning, pedestrian access). The findings of the study is providing
information to consumers, retailers and local governents which regulate the retail environment.

Pizzi and Scarpi (2016) studied the effect of shelf layout on satisfaction and perceived
assortment size, they did an empirical assessment. This research compares the effect of different
shelf visual layouts on decision satisfaction and perceptions of the retailer assortment, with a
focus on how (e.g. vertically vs. horizontally) to display products rather than on how many
products to display. The combined evidence from three experimental studies shows that
displaying assortment breadth horizontally and depth vertically led participants to perceive a
larger selection and to be more satisfied. Furthermore, linear displays increase decision
satisfaction but decrease perceived assortment size and preference strength for the top brand,
whereas visual layouts with the preferred brand in the central position increase decision
satisfaction, make the assortment appear larger, and enhance preference strength for the top
brand. The findings provide guidelines about how to organize products on the shelves and
suggest that – even when actual assortment size is constant – consumer responses to the
assortment depend from how the retailer displays the products.

Choi et al. (2005) worked on a prioritization of association rules in data mining with a multiple
criteria decision approach. Data mining techniques, extracting patterns from large databases are
the processes that focus on the automatic exploration and analysis of large quantities of raw data
in order to discover meaningful patterns and rules. In the process of applying the methods, most
of the managers who are engaging the business encounter a multitude of rules resulted from the
data mining technique. In view of multi-faceted characteristics of such rules, in general, the rules
are featured by multiple conflicting criteria that are directly related with the business values,
such as, e.g. expected monetary value or incremental monetary value. In their paper, they present
a method for rule prioritization, taking into account the business values which are comprised of
objective metric or managers’ subjective judgments. The proposed methodology is an attempt to
make synergy with decision analysis techniques for solving problems in the domain of data
mining. They believe that this approach would be particularly useful for the business managers
who are suffering from rule quality or quantity problems, conflicts between extracted rules, and
difficulties of building a consensus in case several managers are involved for the rule selection.
Videla-Cavieres and Ríos (2014) worked on extending market basket analysis with graph mining techniques. A common problem for many companies, like retail stores, is to find sets of products that are sold together. The only source of information available is the history of sales transactional data. Common techniques of market basket analysis fail when processing huge amounts of scattered data, finding meaningless relationships. They developed a novel approach for market basket analysis based on graph mining techniques, able to process millions of scattered transactions. They demonstrate the effectiveness of our approach in a wholesale supermarket chain and a retail supermarket chain, processing around 238,000,000 and 128,000,000 transactions respectively compared to classical approach.

Le et al. (2013) studied association rule hiding in risk management for retail supply chain collaboration. Association rule hiding is an efficient solution that helps enterprises avoid the risk caused by sensitive knowledge leakage when sharing data in their collaborations. This study examines how data sharing has the potential to create risk for enterprises in retail supply chain collaboration and proposes a new algorithm to remove sensitive knowledge from the released database based on the intersection lattice of frequent itemsets. The proposed algorithm specifies the victim item such that the modification of this item causes the least impact on frequent itemsets and the non-sensitive association rule. In the experiment described in this paper, this algorithm is used in risk avoidance for a retailer sharing data in retail supply chain collaboration. The results indicate that our approach is applicable in a real context and outperforms previous mechanisms.

Chen et al. (2006) worked on a data mining approach for retail knowledge discovery with consideration of the effect of shelf-space adjacency on sales. Recent marketing research has suggested that in-store environmental stimuli, such as shelf-space allocation and product display, has a great influence upon consumer buying behavior and may induce substantial demand. Prior work in this area, however, has not considered the effect of spatial relationships, such as the shelf-space adjacencies of distinct items, on unit sales. This paper, motivated in great part by the prominent beer and diapers example, uses data mining techniques to discover the implicit, yet meaningful, relationship between the relative spatial distance of displayed products and the items’ unit sales in a retailer's store. The purpose of the developed mining scheme is to identify and classify the effects of such relationships. The managerial implications of the discovered knowledge are crucial to the retailer's strategic formation in merchandising goods. This paper proposes a novel representation scheme and develops a robust algorithm based on association analysis. To show its efficiency and effectiveness, an intensive experimental study using self-defined simulation data was conducted. The authors believe that this is the first academically researched attempt at exploring this emerging area of the merchandising problem using data mining.

Minami and Dawson (2008) worked on the CRM process in retail and service sector firms in Japan, they worked on loyalty development and financial return. Research on customer relationship management (CRM) in general has focused on the effects of customer satisfaction with CRM, customer retention and profit management, and the effects of CRM technique on performance. Conceptually, however, a sequence of effects of CRM is expected, from CRM implementation to financial performance, but this sequence has not been explored. Whilst several definitions of CRM have been proposed, this article defines CRM as relationship-development
programmes based on IT. CRM is regarded as the integration of relationship technology (i.e. data consolidating and data mining) with loyalty schemes. Survey research was conducted in Japan in the retail and service industries to test three hypotheses: (1) a firm's relationship orientation has a positive effect on CRM implementation (data warehousing, data mining, using customer data for decision making); (2) CRM implementation has a positive effect on return on equity; and (3) CRM implementation has an indirect effect on return on equity, mediated by customisation. Using a structural equation model the first hypothesis was supported, but the third hypothesis was only partially supported. In these analyses a direct effect of CRM implementation on return on equity (ROE) was supported; however, a negative impact of customisation on ROE was found.

Miguéis et al. (2012) studied customers with data mining for lifestyle segmentation. A good relationship between companies and customers is a crucial factor of competitiveness. Market segmentation is a key issue for companies to develop and maintain loyal relationships with customers as well as to promote the increase of company sales. This paper proposes a method for market segmentation in retailing based on customers’ lifestyle, supported by information extracted from a large transactional database. A set of typical shopping baskets are mined from the database, using a variable clustering algorithm, and these are used to infer customers’ lifestyle. Customers are assigned to a lifestyle segment based on their purchases history. This study is done in collaboration with an European retailing company.

Cil (2012) studied consumption universes based supermarket layout through association rule mining and multidimensional scaling. The success of retail business is influenced by its fast response and its ability in understanding consumers’ behaviors. Analysis of transaction data is the key for taking advantage of these new opportunities, which enables supermarkets to understand and predict customer behavior, has become a crucial technique for effective decision-making and strategy formation. They propose a methodological framework for the use of the knowledge discovery process and its visualization to improve store layout. This study examines the layout strategy in relation to supermarket retail stores and assists managers in developing better layout for supermarkets. They use the buying association measure to create a category correlation matrix and they apply the multidimensional scale technique to display the set of products in the store space. This is a new approach to supermarket layout from industrial categories to consumption universes that is consumer-oriented store layout approach through a data mining approach. This framework is useful for both academia and retail industry. For industry professionals, it may be used to guide development of successful layout. Retailers can utilize the proposed model to dynamically improve their in-store conversion rate. As the empirical study, a practical application proceeded for Migros Turk, a leading Turkish retailing company.

Mirajkar et al., (2016) studied on data mining based store layout architecture for supermarket. The mentioned system is designed to find the most frequent combinations of items. It is based on developing an efficient algorithm that outperforms the best available frequent pattern algorithms on a number of typical data sets. This will help in marketing and sales. The technique can be used to uncover interesting cross-sells and related products. The algorithms from association mining have been implemented and then best combination method is utilized to find more interesting results. The analyst then can perform the data mining and extraction and finally conclude the result and make appropriate decision. Market basket analysis is an important
component of analytical system in retail organizations to determine the placement of goods, designing sales promotions for different segments of customers to improve customer satisfaction and hence the profit of the supermarket. These issues for a leading supermarket are addressed here using frequent itemset mining. The frequent itemsets are mined from the market basket database using the efficient Apriori algorithm and then the association rules are generated.

**LAYOUT**

Facility layout is to develop an economical layout which will meet the requirements of: product design and volume (product strategy), process equipment and capacity (process strategy), quality of work life (human resource strategy), building and site constraints (location strategy). Types of layout are; fixed-position layout, process-oriented layout, work cells, focused work center, office layout, retail layout, warehouse layout, product-oriented layout, assembly-line factory. The objectives of layout strategy that will develop an economical layout are considered to meet the requirements of: product design and volume (product strategy), process equipment and capacity (process strategy), quality of work life (human resource strategy) and building and site constraints (location strategy). Layout has strategic importance to a firm. Proper layout enables: Higher utilization of space, equipment and people, improved flow of information, materials, or people, improved employee morale and safer working conditions, improved customer/client interaction and flexibility. Requirements of a good layout are: an understanding of capacity and space requirements, selection of appropriate material handling equipment, decisions regarding environment and aesthetics, identification and understanding of the requirements for information flow, identification of the cost of moving between the various work areas. Some Rules of Thumb in Retail Layouts are: Locating high-draw items around the periphery of the store, using prominent locations such as the first or last aisle for high-impulse and high margin items, removing crossover aisles that allow customers the opportunity to move between aisles, Distributing what are known in the trade as “power items” (items that may dominate a shopping trip) to both sides of an aisle, and disperse them to increase the viewing of other items, using end aisle locations because they have a very high exposure rate. (Heizer, Render, 2008, p:351)

Once the overall layout of a retail store has been decided products need to be arranged for sale. Many considerations go to this arrangement, however the main objective of retail layout is to maximize profitability per square foot of floor space. **Figure 1: An example on Retail Layout - Grid Design**

An additional and somewhat conversional issue in retail layout is called slotting. Slotting fees are fees manufacturers pay to get goods on shelf in a retail store or supermarket chain. The competition for shelf space is advanced by POS systems and scanner technology which improve supply chain management and inventory control. Many small firms question the legality and ethics of slotting fees, claiming the fees stifle new products, limit their ability to expand and cost consumers money. (Heizer, Render, 2008, p:351)
DATA MINING
Data mining is a technology for processing large volumes of data. It has also opened up exciting opportunities for exploring and analyzing new types of data and for analyzing old types of data in new ways. It has attracted a great deal of attention in the information industry and in society as a whole in recent years, due to the wide availability of huge amounts of data and the imminent need for turning such data into useful information and knowledge.

Data mining is the process of discovering useful information in databases, data warehouses, or other large data repositories. (Tan, Steinbach, Kumar, 2006) It is an interdisciplinary field, drawing from areas such as statistics, machine learning, data visualization, information retrieval, and high performance computing. Other contributing areas include neural networks, pattern recognition, and many application fields, such as business, economics, and bioinformatics. (Han, Kamber, 2006)

Data mining techniques can be used to support a wide range of business intelligence applications such as customer profiling, targeted marketing, workflow management, store layout, production control and fraud detection.

There are many data mining tasks. Some of the common ones are supervised learning, unsupervised learning, association rule mining, and sequential pattern mining. (Liu, 2008) In this study association rules analysis is used for the application.

ASSOCIATION RULES ANALYSIS
Frequent pattern mining searches for recurring relationships in a given data set. It leads to the discovery of interesting associations and correlations between itemsets in transactional and relational databases. With massive amounts of data continuously being collected and stored, many industries are becoming interested in mining such patterns from their databases. The discovery of interesting correlation relationships among huge amounts of business transaction records can help in many business decision-making processes, such as catalog design, cross-marketing, layout and customer shopping behavior analysis. (Han, Kamber, p.228)

The typical example of association rules is market basket analysis. Besides market basket data, association analysis is also applicable to other application domains such as production management, layout, bioinformatics, medical diagnosis, web mining and etc.

Support and confidence rates are measures that show the strength of an association rule. Support determines how often a rule is applicable to a given data set, while confidence determines how frequently items in Y appear in transactions that contain X. The inference made by an association rule suggests a strong co-occurrence relationship between items in the antecedent and consequent of the rule. (Tan, Steinbach, Kumar, p.330)

In this study customer data set from a global furniture and decoration retailer is used. Data has 83233 different purchase records for a particular period. Dataset consists of “transaction, item description and item category” variables. Item categories transposed and coded as F and T according to customers’ buying behavior. If a customer buys the item it is coded as “T”. (Table 1) 2449 items and 250,000 records used for the analysis.
Table 1 - Dataset

The built model can be seen at Table 2.

Table 2 - Model
Association rules analysis is applied to the dataset and the rulesets can be seen at Table 3.

<table>
<thead>
<tr>
<th>Antecedent</th>
<th>Consequent</th>
<th>Support (%)</th>
<th>Confidence(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Living room seating</td>
<td>Home decoration</td>
<td>32,863</td>
<td>68,941</td>
</tr>
<tr>
<td>Living room seating</td>
<td>Lighting accessories</td>
<td>30,917</td>
<td>59,63</td>
</tr>
<tr>
<td>Sofa&amp; Armchair</td>
<td>Pillow</td>
<td>28,105</td>
<td>58,389</td>
</tr>
<tr>
<td>Wardrobe</td>
<td>Bedroom textiles</td>
<td>26,483</td>
<td>50,529</td>
</tr>
<tr>
<td>Wardrobe</td>
<td>Rug</td>
<td>25,825</td>
<td>40,447</td>
</tr>
<tr>
<td>Children’s bed</td>
<td>Children’s textiles</td>
<td>25,611</td>
<td>35,122</td>
</tr>
<tr>
<td>Children’s furniture</td>
<td>Toys</td>
<td>24,146</td>
<td>32,854</td>
</tr>
<tr>
<td>Kitchen tools</td>
<td>Table Set</td>
<td>42,059</td>
<td>29,75</td>
</tr>
</tbody>
</table>
Table 3- Association Rules

The rule sets in the table 3 could be interpreted like that: % 68.941 of the customers who purchased living room seating also purchased Home decoration. This pattern could be seen in %32.863 of the records. %59.63 of the customers who purchased lighting accessories also bought living room seating. This pattern could be seen in %30.917 of the records. % 58.389 of the customers who purchased Sofa & Armchair also purchased pillow with a support rate of %28.10 and the rest of the rule sets could be seen in table 4. Retail data mining can help identify customer buying behaviors, discover customer shopping patterns and trends, improve the quality of customer service, design more effective store layout, good transportation and distribution policies, and reduce the cost of business.

CONCLUSION

In this study customer data set from a global furniture and decoration retailer is used. One month purchase records is analyzed by data mining tools. Association rules helped to find meaningful rule sets that lead the way to make better layout decisions. This might help the managers for both impulse and overal arrangement of the store and the allocation of the space on the space of various products. Those rule sets will help to position to the high impulse and high margin items to the hot points and close to each other. Application of different promotion strategies are possible. Allocation of attraction products on both sides of a row, and dispersion of them to increase the exposition of adjacent products, using the end of a row as a place for exposition, elimination of the aisles that allow the customers to go from one row to other without going through them completely could be some alternatives. Considering ambient conditions like background characteristics such as lighting, sound, smell, and temperature, spatial layout and functionality, which involve customer circulation path planning, signs, symbols could be important. Since the objective is to maximize the net benefit per m² of shelves. If the brick and mortar retail sales are directly related to the exposition of products to the customer, the objective will consist of exposing as many high impulse and high margin products as possible to the customers in the available space. This has to take into account the rule sets driven out of the customer purchase data for making better layout and shelf arrangements of the products.

REFERENCES


