

# Enhancing Customer Loyalty with Market Basket Analysis Using Innovative Methods: A Python Implementation Approach

**Ahmad M. A. Zamil<sup>a</sup>, Ahmad Al Adwan<sup>b</sup>, T. G. Vasista<sup>c</sup>,** <sup>a</sup>Associate Professor, Department of Marketing, College of Business Administration, Prince Sattam bin Abdul Aziz University, 165 Al-Kharj, 11942, Saudi Arabia, <sup>b</sup>Associate Professor and Head of Marketing Department, Business School, Al-Ahliyya Amman University, Amman, Jordan, <sup>c</sup>Senior Consultant, Vasista Consulting and Performing Services OPC Pvt. Ltd. Rajamahendravaram, India, Email: <sup>a</sup>[am.zamil@psau.edu.sa](mailto:am.zamil@psau.edu.sa), <sup>b</sup>[a.aladwan@ammanu.edu.jo](mailto:a.aladwan@ammanu.edu.jo), <sup>c</sup>[tgkvasista@gmail.com](mailto:tgkvasista@gmail.com)

Currently, customer options have increased to purchasing grocery items either in brick and mortar based or online stores. The challenge in retail management is to encourage customers to revisit stores and purchase items. Today, companies are using loyalty programs as a technique to enlarge their customer base so that both consumers and companies can reap benefits. This paper focuses on how to enhance customer loyalty at retail stores using state of the art information technology and programming applications such as Data analytics and Python. By applying the Apriori algorithm libraries of Python programming language in Anaconda Navigator, it is possible to data mine the association rules. Eighteen such association rules have been extracted from a secondary data set file available through various Internet sources such as Kaggle.com and Github, to learn customers' purchasing behaviour at retail stores in order to work on customer relationship management strategy.

**Key words:** *E-Marketing, Enhancing Customer Loyalty, Market Basket Analysis, Python Implementation of MBA, Application of Apriori Algorithm*

## Introduction

Customer options have increased to purchasing grocery items either physically or online through virtual stores. The challenge in retail management is to let customers revisit stores for purchasing higher quantity or other related items. Companies today are using loyalty programs as a technique to grow their customer base so that both customers and companies can reap benefits. Customer satisfaction alone is no longer enough; customer loyalty and retention are becoming more important for the growth of profit and to influence customers to purchase and encourage repurchase. According to Singh and Khan (2012), on average online retailers lose 25% of their customers every year. It can be understood that a small increase in customer retention can increase profits by more than 25%. Customer loyalty is evidently perceived to have a strong positive relationship between satisfaction loyalty and profitability based on research conducted by Helgesen (2006). A business model for the retail sector represents value creation and addition for its customers in markets. Business model innovations help to build sustainable competitive advantage in the market place (Sorescu, et. al., 2011). AlSudairi and Vasista (2012) propose an innovative theoretical framework for an e-enterprise as a business mode that essentially maintains that case study dealings in business management have to be considered. They designed a three-level business model that encapsulates aspects relevant to appropriate abstract levels covering strand, taxonomy as template and instance level in e-business model for 21<sup>st</sup> century compatibility business dealing. In cinch, the four aspects of e-enterprise model are e-value, e-people, e-knowledge and e-action; when customised for the present scenario of retail management. Thus, the scope and definition of business model in general (Sorescu et. al., 2011) and the retail business model in specific (European Commission B-1049 Brussels report) are in sync with the above proposed framework idea of AlSudairi and Vasista (2012) as well as the CRASP methodology suggested by AlSudairi and Vasista (2014), in which e-people include 'Customers' and 'Providers as neutral entities. Responding to customers reflects the e-value. 'Adopting' the strategy/technique reflects e-knowledge. 'Sensing' the data through sensors as well as 'Actuating' the information as a response to what providers offer through actuators reflects e-action. Groceries are related to food retail management and are part of the three basic needs for humans: food, accommodation and clothing. As mentioned in Reinartz, Wiegand and Imschloss (2019), retail management is becoming a central function to all economies. It bridges the needs of consumers and the offerings of producers. The retailing value chain includes brand manufacturers, institutional retailers and consumers whose sole income derives from conducting retail activities. Within the retail value chain, retail trading is shifted from physical store front to Internet based online store format as seen from contemporary online stores such as Amazon, Burberry, Flipcart, Safeway, Target and Walmart. Both internal and external technological innovation in retail stores impact the retail industry. Examples include but are not limited to targeted location-based applications and personalised smart phones, laptops and smart phone point-of-sale,

laptop point-of-sale personal shopping assistants and radio frequency recognition technology etc. (Prepletany, 2013). Internet of Thing (IoT) facilitates implementing new business models and strategies by viewing it as a platform to improve workflow efficiency and increasing employee engagement. IoT gathers huge data through number connected devices so that it can become an important commodity for making better decisions in retail businesses. Smaller devices are generally equipped with RFID chips and sensors to gather electronic data with decreased costs (Sawatzke, 2019).

### ***Market Basket Analysis as an E-Marketing & E-Retail Management Strategy***

Retailers face the challenge of operating vast amounts of goods using tens of thousands of people. They have to make decisions based on price, promotion, space, location and varied collection of items for each product holding stock. The decision for each Stock Keeping Unit (SKU) has an impact not only on sale but also on the sale of other related items. Category management approach tries to tackle this challenge by breaking down this management problem into sub-problems. Each of these categories is practically considered independent of the rest, by grouping highly interrelated products. Category management helps retailers control their strategic business units or division (Musalem, Aburto and Bosch, 2018). Market Basket Analysis as an exploratory scientific method of calculation strengthens the association of pairs of products from individual retailers. This kind of data analysis gathered at the point of sale also helps in understanding consumer behaviour (Bhasin, 2013).

In grocery retail management, Market Basket Analysis (MBA) can inform a super-market pricing strategy. MBA helps in developing strategies related to profitable advertising and promotion targets and revenue increase etc. (Musungwini, Zhoum Gumbo and Mzikamwi, 2014). When combined with scanning technologies for point-of-sale (POS), those retailers collect tremendous amounts of personally identifiable POS transaction data. When analysed, this data helps in understanding behavioural information about prime customers' purchasing patterns across categories, targeting segment-specific cross-selling and up-selling initiatives (Boztug and Retutterer, 2008).

### **Research Objective & Hypotheses**

The research objective is to describing how market basket analysis technique of retail management strategy can use these innovative technologies to help e-retailing companies in enhancing customer loyalty.

$H_0$  = Enhancing customer loyalty at global chain supermarkets is possible with market basket analysis and the use of state-of-the-art information technologies like IoT, Cloud Computing as well as Data Analytics tools like Python programming.

## **Research Methodology**

As cited in Ragab and Arisha (2018), research can be defined as the “systematic investigation into and study of materials and sources in order to establish facts and reach new conclusions” (Oxford Dictionary, 2010). The impact of management studies depends on the rigour and appropriateness of the research methods chosen (Scandura and Williams, 2000). Within the framework of an investigation of trends into management research methodology Scandura and Williams also mention that research strategy can refer to computer simulation (through computer programming) while statistical construction validity could refer to data analytics (p. 1249). The settings of organisational studies, research design and analysis have important implications for knowledge acquisition over time (p. 1248). Research method is the actual data collection and analysis method described and provides justification regarding why particular a research method was chosen. In this paper the case study research method has been chosen as it combines the grounded theory research method with secondary data analysis for understanding and acquiring knowledge. Secondary data is an abundant resource for researchers as the Internet has allowed vast amounts of data to become available. Further quantitative research is concerned with testing hypothesis derived from theory. It uses statistics and other mathematical sciences to report research findings. In order to realise theoretical concepts at the ground level, an effort is made to investigate a phenomenon and to make the context more clearly understandable within a real-life context of retail management solutions. Furthermore, we aim to build Python programming and accessing its libraries to show results. This study is conducted in order to extract association rules using market basket analysis. The data used in the study consist of the sales data of a supermarket received from Kaggle.com.

## **Enhancing Loyalty with Market Basket Analysis**

Numerous effective loyalty programs rely on refined data analytics to understand various factors that influence consumer behaviour. This allows building loyalty programs that add value by allowing advanced segmentation and product personalisation features. To this end, the adoption of Market Basket Analysis technique allows brands to strike a mark on customer satisfaction and in turn drive strategies that improve their loyalty programs. Although MBA represents different combinations of items that regularly occur together in transactions, its context can easily be extended to determine consumer behaviour over longer time periods and at different events. Loyalty is an extension of the brand. Therefore, branding is also becoming key (Neupane, 2015) to the formation and winning of loyalty programs. Advanced analytical techniques in retail marketing such as MBA, would allow understanding today’s Internet connected and digitally empowered customers (Banerjee, 2018).

## The Concept of Market Basket Analysis

Market Basket Analysis (MBA) focuses on customer carts/baskets to monitor and analyse the re-occurring patterns of items purchased together. Sales can be easily tracked on the product and item hierarchy of item classification and on different customer segments in multiple stores (Chen, et. al., 2005). MBA is one of the data mining techniques called frequent set purchased pattern where it focuses on the discovery of purchase patterns by extracting associations from the transaction data of the store. It defines the item set that is purchased together and helps in the designing the promotional campaigns based on the purchase of products and thus can be improved. It also helps in analysing and predicting consumer behaviour (Annie and Kumar, 2012). Furthermore, it is also known as affinity analysis and associations rule mining. Various algorithms such as Apriori etc. are available for MBA performance (Kaur and Kang, 2016). Retailers can make use of MBA insights in a number of ways:

- (i) Cross-Selling: Grouping products frequently purchased together by the customer in the store's product placement.
- (ii) Up-Selling: Encouraging the customer to purchase more quantity of the items he or she is interested in purchasing.
- (iii) Recommendation: Recommending related products that are frequently purchased together.
- (iv) Marketing promotions: Targeting the customer-marketing campaign and encouraging the purchase of related products for items they have been recently purchasing (Grosvenor, 2019).

As mentioned in Kaggle.com, data for market basket analysis is collected from point-of-sale terminals as transaction data or invoices. This data includes the list of products purchased, the unit price and the quantity of each item; usually the data set must be large and consist of 32 million records of 50k unique items.

As outlined in byGhumare (2019), Association Rule Mining technique is used to identify frequently occurring item combinations. It counts the frequency of items that are occurring together through a large collection of items or actions. The aim is to find associations between items that are taking place together. The rules appear in the form of IF {antecedent} THEN {Consequent} format i.e. IF {item1, item2} THEN {item3}.

A sample mini case on market basket analysis with metric computations of support, confidence, lift and conviction is discussed below:

**Table 1:** Sample in Transaction data at retail stores for Market Basket Analysis

Trans_id	Items bought
1	Jam, Perfume, Chips, Soap
2	Beer, Bread, Jam, Chips
3	Potato, Onion
4	Bread, Chocolate, Coffee, Jam
5	Beer, Coffee, Perfume, Chips, Soap
6	Bread, Chocolate, Jam, Soap
7	Onion, Perfume
8	Bread, Chocolate, Chips
9	Chocolate, Perfume, Soap
10	Beer, Bread, Jam, Chips

In a retail shop, the above items are bought in 10 different transactions during one hour which are presented in tabular format as shown in Table 1.

The metrics of market basket analysis are computed for the above tabular data as follows (Poojari, 2018); (Ghumare, 2019):

(i) Support: The popularity of the item or item set. It determines the proportion of transactions containing that item or item set.

Support (item set) = Number of transactions involving item set/Total number of transactions

Looking at Table 1, we can compute

Support for Soap =  $3/10 * 100 = 30\%$

Support for Perfume =  $4/10 * 100 = 40\%$

(ii) Confidence: The likelihood or trustworthiness of association rule, describing how often the association rule was valid.

Confidence of  $(X \rightarrow Y) = \text{Support}(X \cup Y) / \text{Support}(X)$

Confidence of  $(\text{Soap} \rightarrow \text{Perfume}) = \text{Support}(\text{Soap} \cup \text{Perfume}) / \text{Support}(\text{Soap})$

Confidence of  $(\text{Soap} \rightarrow \text{Perfume}) = 30/40 * 100 = 75\%$  i.e.

75% of the time when a customer bought soap, perfume was also bought.

(iii) Lift: The ratio of confidence and expected confidence

Lift  $(X \rightarrow Y) = \text{Confidence}(X, Y) / \text{Support}(Y)$

Lift  $(\text{Soap} \rightarrow \text{Perfume}) = \text{Confidence}(\text{Soap}, \text{Perfume}) / \text{Support}(\text{Perfume})$

Lift  $(\text{Soap} \rightarrow \text{Perfume}) = 75/40 = 1.875 > 1$

Soap and Perfume purchase has a strong relationship and its joint presence is definitely going to boost overall sales

(vi) Conviction ( $X \rightarrow Y$ ) =  $(1 - \text{Support}(Y)) / (1 - \text{Confidence}(X \rightarrow Y))$

Conviction (Soap  $\rightarrow$  Perfume) =  $(1 - \text{Support}(\text{Perfume})) / (1 - \text{Confidence}(X \rightarrow Y))$

Conviction (Soap  $\rightarrow$  Perfume) =  $(1 - 0.4) / (1 - 0.75) = 0.6 / 0.75 = 4/5 = 0.8$

Conviction (Soap  $\rightarrow$  Perfume) = 80%

### **Use of Internet of Things (Iot) As an Innovation Technology in the Retail Industry**

Industry 4.0 is the latest buzzword or term that is being used for current trends in automation and data exchange in manufacturing and related supporting services-based technologies. Embracing this change in terms of Retail 4.0 is also observed in the global retailing industry based on current practices of retail management. Retail concepts have influenced and transformed the processes of retail management for the basic concepts of retail management, store space, logistics and services extended by retailers (Malar, 2019). IoT as an innovative web service platform enables services as a part of the Software-as-a-Service (SaaS) model to bridge the gap between the physical and virtual world (Shankara, et. al., 2015). Retail 4.0 uses IoT and Data Analytics technology in retail stores to attract more customers. It uses Radio Frequency Identification Devices (RFID) and other related sensor technologies and consumer devices such as mobile phones etc. for mining data and to predict analytics. IoT includes devices such as laptops, smart phones, mobile apps, sensors and actuators that are connected to a network for uniquely identifying, generating, collecting, transferring and actuating data. In Retail 4.0, marketing intelligence is obtained by collecting and analysing user data to convert it to obtain semantic and useful information on user performance and purchase patterns for marketing purposes. This technology will have the capability to address the dynamics of retailing in terms of managing dynamic pricing depending on the supply and demand of the products in the retail outlet. Retail 4.0 uses smart shopping carts, automated beacon checkouts, real-time customer tracking and analytics to understand consumer behaviour and buying patterns. This kind of analysis leads to product retail intelligence to frame better strategic marketing of consumer products and services.

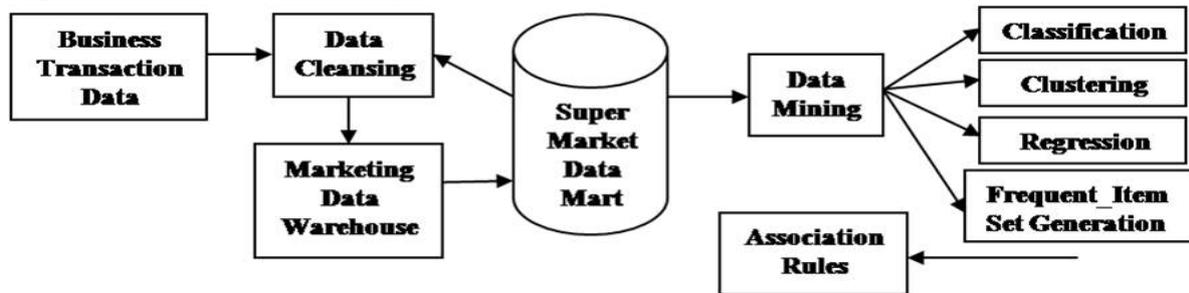
### **IoT and Business Data Analytics Framework for Retail Management**

The framework at a higher abstraction level consists of five layers:

(i) Connecting physical devices (ii) network devices (iii) middleware (iv) database and (v) application. The middle layer's role is to manage resources, event and data management and recovery through adopting service-oriented architecture (SOA) for end-users and applications. The database layer has two components: (i) the distributed server i.e.

shared database server and (ii) a structured database. Structured data is delivered to a cloud server where various data mining and machine learning algorithms can be used for information and knowledge extraction (Saqlain, et. al., 2019).

**Figure 1.** Business Data Analytics Framework



Basically, Internet of Things (IoT) connects huge numbers of devices that can sense data, communicate and potentially actuate using devices such as sensors, cameras, beacons, wi-Fi networks and actuators. IoT can help gather valuable data that can be analysed and used to optimise supply chain, staff deployment, storage and finally to respond with actuations (IoTUK, 2017). Actuators enable digital data, converting it into action. They alter a physical entity and respond to users in a useful manner. In the warehouse safety management domain, examples include warehouse door locks and fire alarms (ARTS, 2017).

According to Zaslavsky, Perera and Geogakopoulos (2013) Sensing-as-a-Service model of cloud computing has an important role to play in the IoT paradigm. It provides all features of a service model through cloud computing. In general, all these models can be called XaaS, where X can be almost anything to represent as a service (Vasista and AlSudairi, 2012).

For example, Wal-Mart generates a quantity of 2.5 bytes of customer transaction data every hour. Furthermore, if Wal-Mart operates RFID at an item level, it is expected to generate 7 terabytes of transaction data each day. In retail business, if this transaction data is analysed for customer movements in store or online, product searching such as consumer behaviour and consumer preferences can be better understood (Zaslavsky, Perera and Georgakopoulos, 2013). In this context, two management methodologies that are relevant to the present context are: (i) Customer-Respond-Adapt-Sense-Provider (CRASP) methodology (AlSudairi and Vasista, 2014).

(ii) Transform-Analyse-Measure-Predict-Act (TAMPA) as a closed loop information methodology (Vasista, 2007) and an extension of the MPA methodology referred to by Berson, Smith and Thearling (2000).

IoT becomes smarter and adds value when connected to valuable technologies such as data mining, which involves discovering new, interesting and potentially useful patterns from large data sets and applying relevant algorithms to extract hidden trends and patterns of information. Before deriving the knowledge, the data available as an integrated data at data source is called a data warehouse. The required data is extracted and available in the form of data mart. Data cleansing is completed as part of data pre-processing before applying data mining techniques and algorithms. Visualisation in the form of graphs and charts adds more value when combined with data mining (Chen, Deng, et. al., 2015).

### **Implementing and Discussing Beliefs in Python towards Market Basket Analysis**

The following points are discussed with reference to the Python coding in Screen shots presented in Appendix-I (see end of document):

- (i) Implementation in Python language involves declaring library files and packages.
- (ii) The Apriori library requires the data set to be in the form of lists or list of lists. Each data transaction in the data set can appear in the form of an inner list within the outer list. Currently, data is available in the form of panda's data frame and FOR loop is used to convert this data to a list of lists.
- (iii) Apriori algorithm works with Python 3.6 and higher. It is a machine-learning algorithm used to uncover hidden structures in categorical data and gain insights into the structured relationship between the various subjects involved.
- (iv) The Apriori algorithm in Python is a class that requires value-based parameters to work with (Poojari, 2018). The first parameter is the list of lists from which association rules are extracted. The second, third and fourth parameters are the min\_support, min\_confidence and min\_lift values to be specified as threshold values in order to filter rules with more than or equal to the threshold values specified. Another parameter called min\_length parameter specifies the minimum number of items in the rules. For example, min\_length = 2 indicates that at least two products are required by the rules.
- (v) The association rules found and extracted by Apriori class appear in LIST FORM.
- (vi) The current processing of transaction file extracted 18 rules. This is outlined by the print (Len (association results)).
- (vii) According to Relation Record, ground beef is purchased with spaghetti as a basic frozen vegetable.
- (viii) Lift > 1, means the sale of items in the association rule has a strong positive relationship.



## Conclusion

The concept of smart retail goes beyond the application of modern technology to the retail process. IoT helps physical store retailers in gaining insights into consumer data. IoT application can synthesise data from video surveillance cameras, mobile devices and websites on social media; it predicts customer behaviour. Retail businesses can enhance procurement planning and can be made more effective with the use of store shelf sensors, RFID tags, beacons, video monitoring and digital price tags. Retailers can use IoT based automated point of sale systems to reduce waiting time in the check-out queue lines when customers leave the store and stop customers from leaving without making any purchases. Retail stores can conduct better predictive equipment maintenance. IoT improves RoI, enhances customer satisfaction, loyalty and customer experience as well as brand loyalty (Gaikwad, 2018). Furthermore, data mining methods and algorithms will enhance organisational performance and gain a competitive advantage. Market Basket Analysis (MBA) used by retailers determines the optimal locations to promote products. MBA acts as a powerful tool for the implementation of cross-selling and up-selling strategies (Hemalatha, 2012).

## REFERENCES

- AlSudairi, M. A. T. and Vasista, T. G. K. (2012). “Design of strategic business model for electronic enterprise in digital society. *International Journal of Digital Society*, Vol. 3, Iss. 3 and 4, pp 690 – 697.
- Alsudairy, M. A. T. & Vasista, T. G. K. (2014). “CRASP—a strategic methodology perspective for sustainable value chain management,” In *Proceedings of the 23rd IBIMA Conference*, Valencia, Spain, May 13-15.
- Annie, L. C. and Kumar, A. D. (2012). “Market basket analysis for a Supermarket based on Frequent Itemset Mining,” *International Journal of Computer Science Issues*, Vol. 9, Iss. 5, No. 3, pp. 257-264.
- ARTS (2017). *The Internet of Things in Retail*. ARTS white paper, [Available at] [https://cdn.nrf.com/sites/default/files/2018-10/ARTS\\_IoTOperations\\_TechnicalReport\\_20170208\\_FINAL.pdf](https://cdn.nrf.com/sites/default/files/2018-10/ARTS_IoTOperations_TechnicalReport_20170208_FINAL.pdf) [Retrieved on 06-06-2020].
- Banerjee, K. (2018). *Enhancing Customer Loyalty using Market basket analysis, Big Data and Analytics Customer experience, customer sentiment analysis, Enterprise Analytics, market basket analysis*, April 2.
- Berson, A., Smith, J. S. and Kurt, T. *Building Data mining applications for CRM*, First Edition, USA: McGraw-Hill Professional, 2000, 510-610.
- Bhasin, R. 2019. *Shopper Behaviour Exploration and Market Basket Analysis using Apache Spark*. [Available at] <https://medium.com/analytics-vidhya/shopper-behavior-exploration-and-market-basket-analysis-using-spark-650656d6a0e1> [Retrieved on 03-07-2020].
- Boztug, Y and Reutterer, T (2008) a combined approach for segment-specific market basket analysis. *European Journal of Operational Research*, 187, 1. pp. 294-312.
- Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A., & Rong, X. (2015). Data mining for the internet of things: Literature review and challenges. *International Journal of Distributed Sensor Networks*, 2015, Article ID: 431047.
- Chen, Y.-L., K. Tang, R.-J. Shen and Y.-H. Hu. 2005. Market basket analysis in a multiple store environment. *Decision Support Systems* 40 (2):339–54. doi:10.1016/j.dss.2004.04.009.



- European Commission B-1049 Brussels Report (2014). Consumers as Innovation Drivers: Six perspectives on Retail Innovation – Expert Group on Retail Sector Innovation, Luxembourg: Publications Office of the European Union, pp. 11-15.
- Gaikwad, G. (2018). Retail IoT: Applications, challenges, and solutions. [Available at] <https://www.iotforall.com/retail-iot-applications-challenges-solutions/> [Retrieved on 07-06-2020].
- Ghumare, R. (2019). Market Basket Analysis using Association Rule Mining. [Available at] <https://medium.com/@ghumare64/market-basket-analysis-using-association-rule-mining-2f87e4064ee7> [Retrieved on 03-07-2020].
- Grosvenor, M. (2019). Market Basket Analysis 101: Anticipating customer behaviour. [Available at] <https://smartbridge.com/market-basket-analysis-101/> [Retrieved on 03-06-2020].
- Helgesen, Oy. (2006). Are loyal customers profitable? Customer satisfaction, customer (action) loyalty and customer profitability at the individual level. *Journal of Marketing Management*, 2006, 22, 245-266.
- Hemalatha, M. (2012). Market basket analysis – A data mining application in Indian retailing. *International Journal of Business Information Systems*, vol. 10, iss. 1, pp. 109-120.
- IoTUK (2017). Internet of Things in Retail, How can SME vendors encourage retailers to adopt the latest technologies?2017 <https://iotuk.org.uk/wp-content/uploads/2017/03/IoT-in-Retail.pdf>.
- Kaur, M., and S. Kang. 2016. Market basket analysis: Identify the changing trends of market data using association rule mining. *Procedia Computer Science* 85:78–85. doi:10.1016/j.procs.2016.05. 180.
- Malar M. S. (2019). Industry 4.0 transforming global retail industry. *Indian Journal of Computer Science*, 4(1), 19–25. 10.17010/ijcs/2019/v4/i1/142414.
- Musalem, A. Aburto, L. and Bosch, M. (2018). Market Basket Analysis insights to support category management. *European Journal of Marketing*, vol. 52, iss. 7/8.108-117.
- Musungwini, S., Zhou, T.G., Rebanowako, T.G., Gumbo, R. & Mzikamwi, and T. (2014) The Relationship between 4Ps and Market Basket Analysis: A Case Study of Grocery Retail Shops in Gweru Zimbabwe. *International Journal of Scientific and Technology Research*.



- Neupane, R. (2015). The Effects of brand image on customer satisfaction and loyalty intention in retail supermarket chains in UK. *International Journal of Social Sciences and Management*. Vol. 2, Iss-1, pp. 9-26.
- Oxford Dictionary. (Ed.) (2010). Third Edition, Oxford University Press
- Poojari, D. (2018). Apriori Algorithm in Python (Recommendation Engine). [Available at] <https://medium.com/@deepak.r.poojari/apriori-algorithm-in-python-recommendation-engine-5ba89bd1a6da> [Retrieved on 03-07-2020]
- Prepletany, D., (2013). The impact of digital technologies on innovations in retail business models. Master's thesis, Aalborg University, Aalborg, Denmark.
- Ragab, M A F and Arisha Amr (2018). Research Methodology in Business: A Starter's Guide. *Management and Organisational Studies*, Vol. 5, No. 1, pp. 1-23.
- Reinartz, W., Wiegand, N. and Imschloss, M., 2019. The impact of digital transformation on the retailing value chain. *International Journal of Research in Marketing*, 36(3), pp. 350-366
- Saqlain, M., Piao, M., Shim, Y., & Lee, J. Y. 2019. Framework of an IoT-based Industrial Data Management for Smart Manufacturing. *Journal of Sensor and Actuator Networks*, 8 (2) 25.
- Sawatzke, D. (2019). IoT and RFID: What's the connection? *RFID Journal*. [Available at] <https://www.rfidjournal.com/iot-and-rfid-whats-the-connection> [Retrieved on 01-06-2020]
- Scandura, T. A. and Williams, E. A. (2000). Research Methodology in Management: Current Practices, Trends and Implications for Future Research. *Academy of Management Journal*, Vol. 43, No. 6, pp. 1248-1264.
- Shankara, P., Mahanta, P., Arora, E. and Srinivasamurthy, G., 2015, October. Impact of Internet of Things in the Retail Industry. In *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"* (pp. 61-65).
- Singh, R and Khan, I.A. (2012). "An Approach to Increase Customer Retention and Loyalty in B2C World," *International Journal of Scientific and Research Publications*, vol. 2 no. 6, pp. 1-5, 2012.
- Sorescu, A., Frambach, R. T., Singh, J., Rangaswamy, A., & Bridges, C. (2011). Innovations in retail business models. *Journal of Retailing*, 87, pp. 3-16.



Vasista, T G K and AlSudairi, M A T (2014). CRASP – A strategic methodology perspective for sustainable value chain management. 23<sup>rd</sup> IBIMA Conference, May 13-14, Valencia, Spain.

Vasista, T. G. K. (2007). Wise CRM engine. Synergy-The Journal of Marketing, vol. 5, iss. 1, pp. 123–127.

Vasista, T. G. K. and AlSudairi, M. A. T (2012). Web Portal Integration Architecture Approaches. International Journal on Web Service Computing, vol. 3, no. 3, pp. 39-47.

Yüksel Akay Ünvan (2020): Market basket analysis with association rules, Communications in Statistics - Theory and Methods, DOI: 10.1080/03610926.2020.1716255.

Zaslavsky, A., Perera, C., and Georgakopoulos, D. (2013). Sensing as a service and big data. <https://arxiv.org/ftp/arxiv/papers/1301/1301.0159.pdf>

### **Authors**

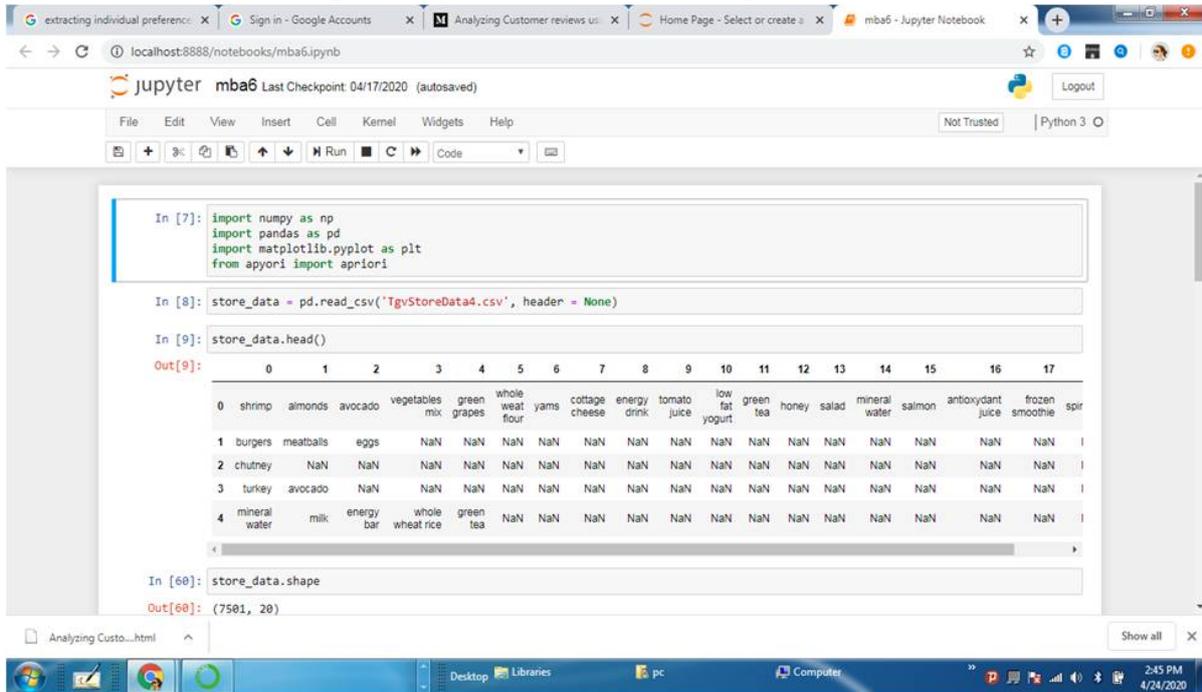
Dr. Ahmad M. A. Zamil is currently working as an Associate Professor in Prince Sattam Bin Abdulaziz University, AlKharj, KSA. He has 24 years of progressive experience in teaching and research in marketing specialisation in Jordan and Saudi Arabia.

Dr. Ahmad Al-Adwan is currently working as an Associate Professor and Head in Al-Ahliyya Amman University, Jordan.

Dr. T. G. Vasista is currently working as senior consultant at Vasista Consulting and Performing Services OPC Pvt. Ltd., Rajamahendravaram, India. He has a total of 27 years of versatile professional experience in the domain of construction project management, e-business, e-governance in academics, research and the IT industry.

## APPENDIX-I

### Screenshot 1: Python Implementation of MBA – Reading transaction data file



The screenshot shows a Jupyter Notebook interface with the following code and output:

```
In [7]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from apyori import apriori

In [8]: store_data = pd.read_csv('TgvStoreData4.csv', header = None)

In [9]: store_data.head()

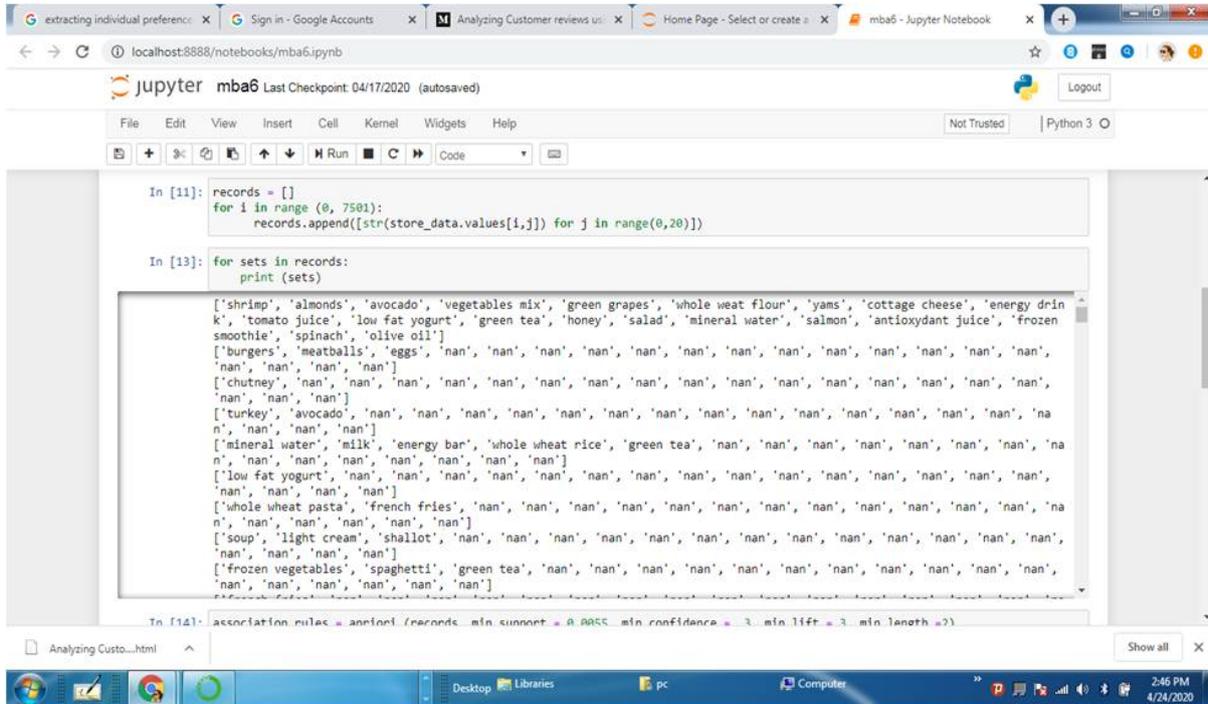
Out[9]:
```

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
0	shrimp	almonds	avocado	vegetables mix	green grapes	whole wheat flour	yams	cottage cheese	energy drink	tomato juice	low fat yogurt	green tea	honey	salad	mineral water	salmon	antioxydant juice	frozen smoothie	spir
1	burgers	meatballs	eggs	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	chutney	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	turkey	avocado	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	mineral water	milk	energy bar	whole wheat rice	green tea	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

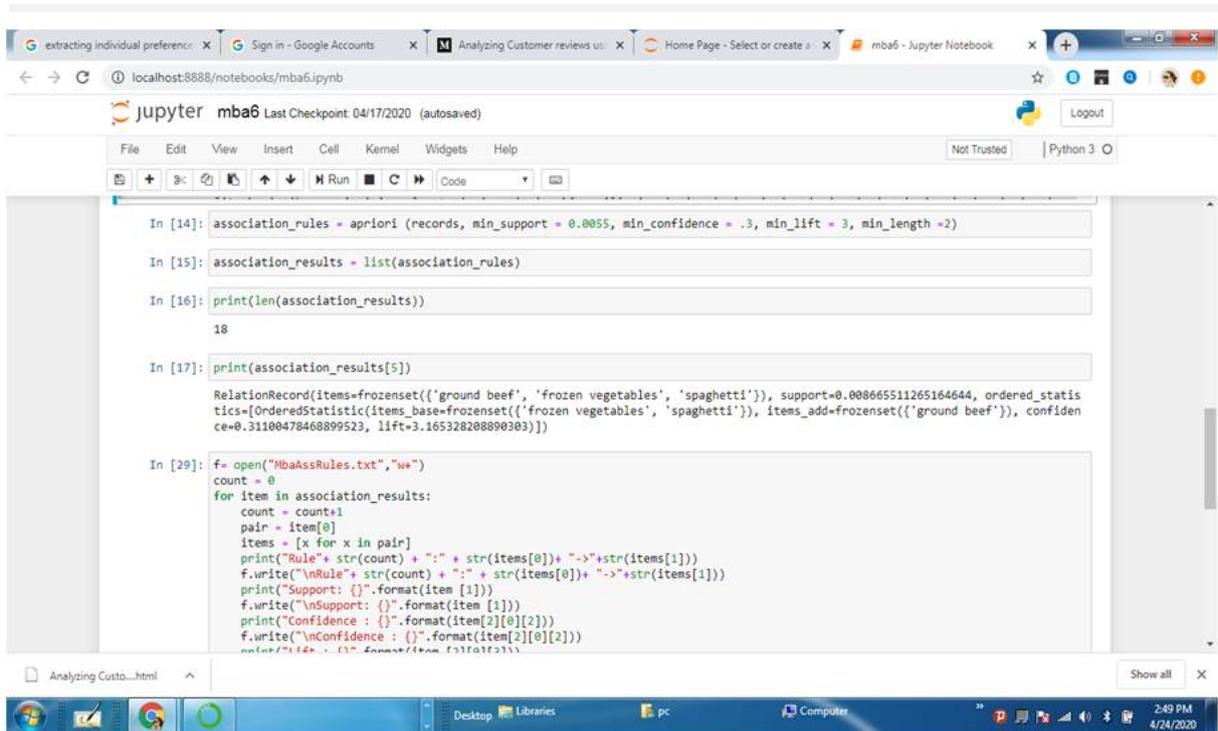
```
In [60]: store_data.shape

Out[60]: (7501, 20)
```

### Screenshot 2: Python Implementation of MBA – Listing Item sets



### Screenshot 3: Python Implementation of MBA – Extracting Association rules by applying Apriori algorithm



```
In [14]: association_rules = apriori(records, min_support = 0.0055, min_confidence = .3, min_lift = 3, min_length = 2)

In [15]: association_results = list(association_rules)

In [16]: print(len(association_results))

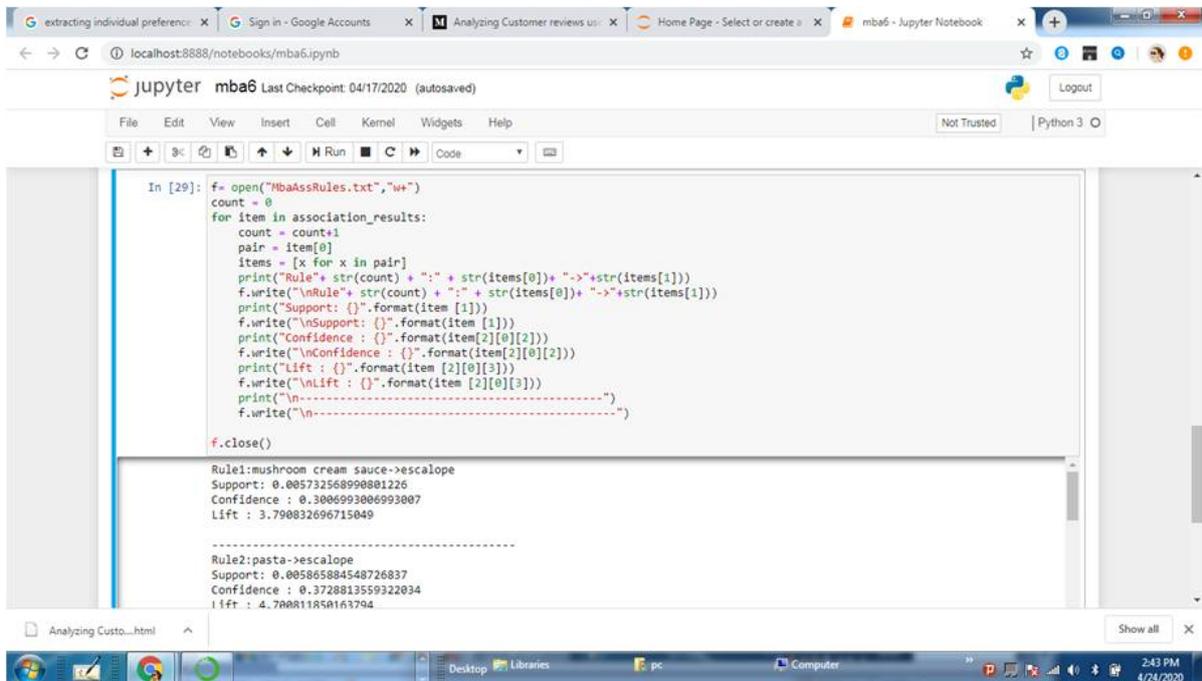
18

In [17]: print(association_results[5])

RelationRecord(items=frozenset({'ground beef', 'frozen vegetables', 'spaghetti'}), support=0.008665511265164644, ordered_statistics=[OrderedStatistic(items_base=frozenset({'frozen vegetables', 'spaghetti'}), items_add=frozenset({'ground beef'}), confidence=0.31180478468899523, lift=3.165328208890303)])

In [29]: f= open("MbaAssRules.txt","w+")
count = 0
for item in association_results:
    count = count+1
    pair = item[0]
    items = [x for x in pair]
    print("Rule"+ str(count) + ":" + str(items[0])+"->" +str(items[1]))
    f.write("\nRule"+ str(count) + ":" + str(items[0])+"->" +str(items[1]))
    print("Support: {}".format(item [1]))
    f.write("\nSupport: {}".format(item [1]))
    print("Confidence : {}".format(item[2][0][2]))
    f.write("\nConfidence : {}".format(item[2][0][2]))
    print("Lift : {}".format(item [2][0][3]))
    f.write("\nLift : {}".format(item [2][0][3]))
```

**Screenshot 4:** Python Implementation of MBA – printing and writing association rules onto a file



```
In [29]: f= open("MbaAssRules.txt","w+")
count = 0
for item in association_results:
    count = count+1
    pair = item[0]
    items = [x for x in pair]
    print("Rule"+ str(count) + ":" + str(items[0])+"->" +str(items[1]))
    f.write("\nRule"+ str(count) + ":" + str(items[0])+"->" +str(items[1]))
    print("Support: {}".format(item [1]))
    f.write("\nSupport: {}".format(item [1]))
    print("Confidence : {}".format(item[2][0][2]))
    f.write("\nConfidence : {}".format(item[2][0][2]))
    print("Lift : {}".format(item [2][0][3]))
    f.write("\nLift : {}".format(item [2][0][3]))
    print("\n-----")
    f.write("\n-----")

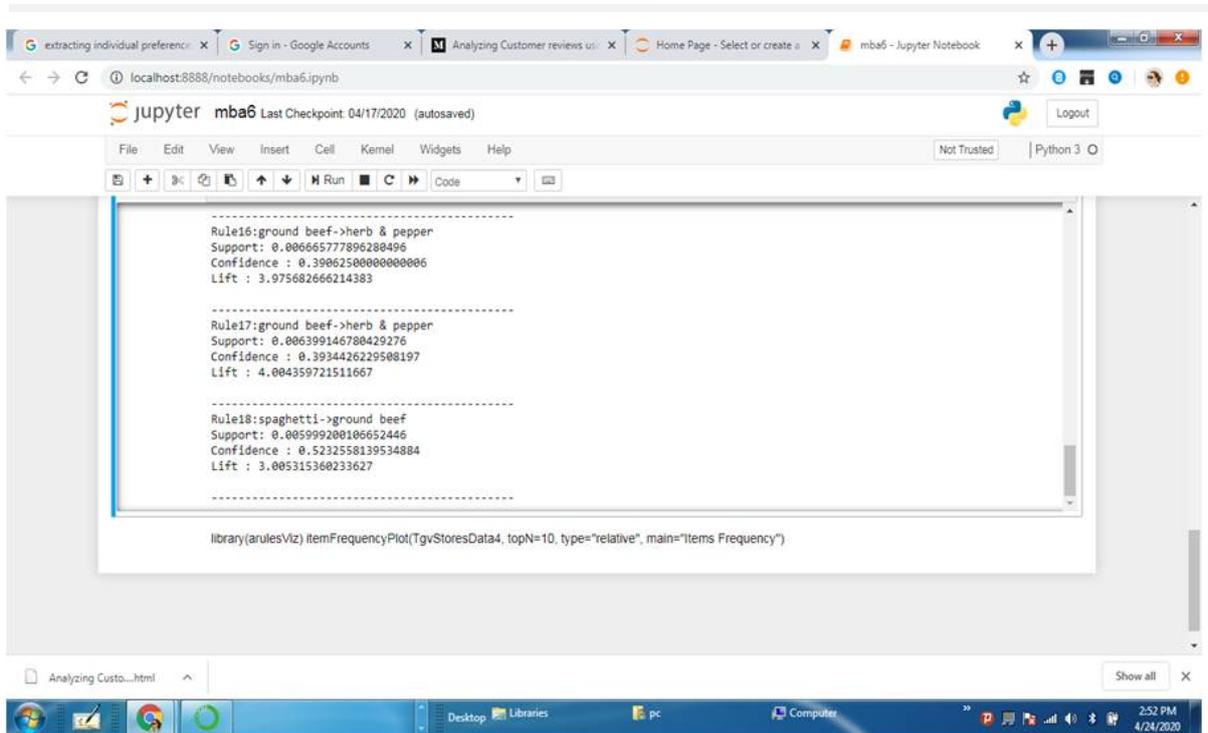
f.close()

Rule1:mushroom cream sauce->escalope
Support: 0.005732568990801226
Confidence : 0.3006093006093007
Lift : 3.790832696715049

-----

Rule2:pasta->escalope
Support: 0.005865884548726837
Confidence : 0.3728813559322034
Lift : 4.700811850163794
```

**Screenshot 5:** Python Implementation of MBA – 18 Association rules extracted



The screenshot displays a Jupyter Notebook environment within a web browser. The notebook contains three association rules, each with its support, confidence, and lift values. The rules are as follows:

Rule	Association	Support	Confidence	Lift
Rule16	ground beef->herb & pepper	0.006665777896280496	0.39062500000000006	3.975682666214383
Rule17	ground beef->herb & pepper	0.006399146780429276	0.3934426229508197	4.004359721511667
Rule18	spaghetti->ground beef	0.005999200106652446	0.5232558139534884	3.005315360233627

Below the rules, the following code is visible:

```
library(arulesViz) itemFrequencyPlot(TgvStoresData4, topN=10, type="relative", main="Items Frequency")
```