11-2020

Capitalising product associations in a supermarket retail environment

Michelle L. F. CHEONG
Singapore Management University, michcheong@smu.edu.sg

Yong Qing CHIA
Singapore Management University, yqchia.2017@mitb.smu.edu.sg

2020 Annual Conference Proceedings

Decision Sciences in the Age of Connectivity

DECISION SCIENCES INSTITUTE

ISBN: 978-0-578-62648-2
ISSN: 2471-884X

51st Annual Conference

November 21st - 23rd 2020

Table of Contents

DECISION SCIENCES INSTITUTE

Capitalising Product Associations in a Supermarket Retail Environment

Michelle Cheong Lee Fong
Singapore Management University
Email: michcheong@smu.edu.sg

Chia Yong Qing
Singapore Management University
Email: yqchia.2017@mitb.smu.edu.sg

ABSTRACT

This paper explores methods to capitalise on retail companies' transactional databases, to mine meaningful product associations, and to design product placement strategies as a means to drive sales. We implemented three in-store initiatives based on our hypotheses – placing products with high associations together will induce an increase in sales of consequent; introducing an antecedent that is new to store will bring about a similar impact on sales of consequent based on established product association rules uncovered from other stores. Sales tracking over twelve weeks revealed that there were improvements in sales of consequents across all three initiatives performed in-store.

KEYWORDS: Association Mining, Market Basket Analysis, Shopper Marketing

INTRODUCTION

Market surveys revealed that the global online shopping market size is expected to hit 4 trillion by 2020 (Clement, 2020). Amidst heightened competition from e-commerce players, it is a challenge for traditional retailers to retain customer's patronage and market share. Increasingly, brick-and-mortar companies are turning to shopper marketing for innovative practices to market to their customers in-stores. (Kopalle, et al., 2009). A key factor exacerbating this trend is the availability and use of data by retailers to understand customer's profiles, their motivations for purchase, as well as the optimal channels to reach out to them.

This paper attempts to discover new ways to leverage consumers' purchasing patterns to trigger successful purchases at the point of sales. To do so, we partnered with a corporate sponsor, NTUC FairPrice, which is the largest retailer in Singapore with more than 50% of the market share in the physical retail segment. Despite being the market leader in its domain, FairPrice's supermarket business is not spared from disruption by online contenders such as Lazada's RedMart and Amazon Prime Now. With over 100,000 customer touch points daily, it is of FairPrice's interest to reinvent ways to tap on its rich transaction database to generate additional sales and value for its customers.

LITERATURE REVIEW

Shopper Marketing

Shopper marketing is defined as the act of planning and execution of marketing activities that influence a consumer along his shopping process (Shankar, 2011). There have been continuous developments in the field of shopper marketing due to a shift in environmental factors that results in changes in shoppers' behaviours (Shankar, Inman, Mantrala, Kelley, & Rizley, 2011). One such shift is the increase in e-commerce adoption which has altered shoppers' search and browsing habits. While research has shown that most consumers prefer to browse for products online, more than half still prefer being in-store when making the final purchase decision (Briggs, 2016). Therefore, there is a potential upside for retailers to focus on marketing at the point of purchase (Löfgren, 2005).

In this paper, we referenced literature on innovations in merchandising which is a form of shopper marketing in the retail environment. While there has been extensive research on shoppers' navigation behaviours to create a positive shopper-centric environment, there is limited study on the use of consumers' purchasing patterns to design placement strategies that capitalise on synergies between products to trigger sales.

Association Rule Mining

We reviewed literature that explored association rule mining as a technique to identify the relationship between items in a large set of data. Particularly, we were interested in Market Basket Analysis, an application of association rules mining, to analyse customers' shopping patterns and formulate relationships amongst the various items placed in shoppers' basket (Gupta & Mamtora, 2014). The relationship between two product items can be represented as an association rule whereby X implies Y such that

X ==> Y (X is an antecedent and Y is a consequent)

Apriori is the most widely applied algorithm in the field of association rule mining where rules are discovered as the algorithm is passed repeatedly over a dataset and product pairs that are deemed to be infrequent are deleted (Yabing, 2013). Past literature studied different benchmarks to select rules which are useful and of interest to the user. Both statistical methods such as Support, Confidence, and Lift, as well as subjective methods, have been discussed.

Traditionally, high Support and Confidence values are used to evaluate the certainty of mined association rules. Support is defined as the percentage of transaction records that contain both products; for a pair of products X and Y, how often they are bought together in a transaction. Confidence measures the strength and accuracy of the rule, that is, how often do consumers who buy X also buy Y (Kotsiantis & Kanellopoulos, 2006). Lift takes this consideration one step further by measuring the ratio of Support divided by the probability that the antecedent and consequent occur together if the two products are independent. A lift value greater than 1 shows that the antecedent and consequent appear more frequently together than expected. Therefore, the larger the lift, the greater the association between the two products.

While rules with high Support shows that a product pair is often bought together, such pairings are generally well known and do not necessarily generate new insights (Coenen, Goulbourne, & Leng, 2004). Rules with low Support and high Confidence, on the other hand, are more likely to be counter-intuitive and interesting for business considerations. Beyond statistical measures, literature works such as (Geng & Hamilton, 2006) also suggested the use of a combination of objective, subjective and semantic measures to validate the association rules. Specifically, the user's domain knowledge is a good source of qualifier to rank the discovered rules according to various interestingness measures such as unexpectedness and conformity (Liu, Hsu, Chen, & Ma, 2000).

Though there have been extensive efforts to apply Market Basket Analysis to understand which products are commonly bought in combination with one another in a physical retail setting, most research works fell short on the practical application of the mined association rules for store merchandising purposes. Some applications include improving in-store layout (Surjandari & Seruni, 2010) and product category management (Griva, Bardaki, Pramatari, & Papakiriakopoulos, 2018) with limited focus on product placement. This paper attempts to bridge the gap in the study of product placement at the stock-keeping unit (SKU) level based on established association pairings and to evaluate the improvements before and after actual implementations. Specifically, we would like to investigate the impact of sales due to new product placements, and the addition of a new antecedent on the sales of its associated consequent in a retail store based on rules mined from other stores.

DATA SOURCE AND PREPARATION

Transactional data from receipts generated at four retail stores across the period of 1st Jun 2018 to 31st Aug 2018 was provided by the corporate sponsor. The four stores, selected based on their customer profile mix, were aligned to FairPrice's target consumer group and agreed upon during the initial discussions with the sponsor. The extracted dataset contains 9,729,581 rows of transactions. <TRAN_NO>, which reflects the transaction ID of a receipt, was repeated across the different stores and therefore not unique. To create a unique transaction ID, a new column <Unique_Trans_ID> was created by concatenating <Business_Date>, <Store>, <Reg_No>, <Tran_No>. Columns that are not relevant to our research have been streamlined, and the remaining dataset contains 17 columns as shown in Table 1. The pre-processed data file was then used to perform Market Basket Analysis using the association node function in SAS Enterprise Miner.

Table 1: Metadata Table Of 17 Data Fields

COLUMN	SAMPLE DATA	DESCRIPTION
UNIQUE_TRAN_ID	6/1/2018 12:00:00 AM31315368	Unique ID created by concatenating <Business_Date>, <Store>, <Reg_No>, <Tran_No>.
BUSINESS_DATE	6/1/2018 12:00:00 AM	Data and time of transaction in MM/DD/YYYY in 24-hour clock format.
STORE	313	Unique store code.
REG_NO	1	POS register counter number in store.
TRAN_NO	5368	Transaction number at POS counter.
LINE_NO	5	Sequence of scanned item in a transaction.
SKU	209280	Stock Keeping Unit (SKU) number, varying from 6-8 digits.
EAN	20183388610	International Article Number, varying from 11 to 13 digits.
ARTICLE_NAME	VALLEYCHEF CHK FRANKS 340G	Description of item.
DEPT	21	Unique department code ranging from 1-9.
DEPT_NAME	FROZEN FOOD	Description of department.
CLASS	21047	Unique ID of product class.
CLASS_NAME	FROZEN PROCESSED FD	Description of product class.

Table 1: Metadata Table Of 17 Data Fields		
COLUMN	SAMPLE DATA	DESCRIPTION
SUBCAT1	2104701	Unique ID of subcategory 1.
SUBCAT1_NAME	FROZEN MEAT PROCESSED FOOD	Description of subcategory 1.
SUBCAT2	210470101	Unique ID of subcategory 2.
SUBCAT2_NAME	FROZEN SAUSAGE/ FRAN	Description of subcategory 2.

THEORETICAL DEVELOPMENT

The research methodology was structured to uncover useful and interesting association rules and to design and test methods to capitalise on this information for sales-driving purposes. An overview of the approach is listed down below,

I. Market Basket Analysis of the pre-processed dataset was performed. A list of meaningful association rules was shortlisted based on Confidence and Lift value above a certain cutoff value. The list was then further streamlined based on the author’s domain knowledge and assessment of the interestingness of the rules.

II. Research hypotheses were raised to study the impact of the presence of an antecedent on the sales performance of the consequent. These hypotheses were derived from the selected association mining rules discovered. In-store initiatives were then designed and implemented to measure the sales of selected consequent in the presence and absence of an antecedent, taking into account any macro and seasonal trends affecting the overall product category sales in the pilot store.

III. The proposed initiatives were implemented in a pilot store over 12 weeks. Sales tracking was performed, and the results and insights were evaluated and communicated to the corporate sponsors.

Selection Of Association Rules

In this paper, we were primarily concerned with rules which have low Support and relatively high Confidence and/or Lift values. We set the cutoff value for Support to be <0.1 and Confidence and/or Lift to be >10. This suggests that even though the two products were not purchased frequently, but whenever they were, it was very likely for them to be purchased together. Products that were on bulk purchase promotion, for example, buy 2 at a discounted price, were likely to have high Support, therefore they were not considered. The selected rules were also chosen to conform to the author’s domain knowledge in the retail business to remain feasible for in-store applications. The three selected pairs of association rules are shown below,

Glad Cling Wrap 200ft ==> Glad Aluminum Foil 75sqft 1s
Support – 0.02
Confidence – 7.41
Lift – 37.07

Kim Poh Dish/wash liquid 5kg ==> Budget Kitchen Towel 6roll 60s
Support – 0.02
Confidence – 10.34
Lift – 11.05

Cirio Chopped Tomatoes 400g ==> Beef Minced kg
Support - 0.05
Confidence – 13.87
Lift – 11.86

DESIGN AND IMPLEMENTATION OF EXPERIMENT

Based on the criteria for the selection of rules, three promotional mechanisms were designed and implemented at a pilot store over a period of 12 weeks. The hypotheses tested are as follows,

- Hypothesis 1 – Placing products with high association together generates sales for consequents where X (antecedent) ==> Y (consequent).

- Hypothesis 2 – Introducing an antecedent (new product to store) will bring about an increase in sales of consequent based on established product associations from another store.

Table 2: In-store Initiatives Implemented In Test Store

S/N	ANTECEDENT	CONSEQUENT	HYPOTHESIS	IN-STORE INITIATIVES
1	Glad Cling Wrap 200ft	Glad Aluminum Foil 75sqft 1S	Hypothesis 1 – Placing products with high association together generates sales for consequents where antecedent ==> consequent.	

2	Kim Poh Dish/wash liquid 5kg	Budget Kitchen Towel 6Roll 60s	Hypothesis 1 – Placing products with high association together generates sales for consequents where antecedent ==> consequent.	
3	Cirio Chopped Tomatoes 400G	Minced Beef KG	Hypothesis 2 – Introducing a new antecedent to store will bring about an increase in sales of consequent based on established product associations derived from another store.	

Sales tracking of the above initiatives in FY2018/19 was performed over the duration of 12 weeks before implementation (Period 1) and 12 weeks after implementation (Period 2). We also compared the sales results during the same periods from 2017 to 2018 (denoted as FY2017/18). Such a comparison will take into account any sales improvement or decline due to Business As Usual (BAU) so that we can isolate any sales improvements of the consequents solely due to the experiments. The results of sales tracking and business insights derived are presented in the next section of the paper.

Table 3: Period Of Experiment

	FY2017/2018	FY2018/2019
Period 1 (12 Weeks Before)	06 September 2017 ~ 28 November 2017	06 September 2018 ~ 28 November 2018
Launch Week	29 November 2017 ~ 05 December 2017	29 November 2018 ~ 05 December 2018
Period 2 (12 Weeks After)	06 December 2017 ~ 27 February 2018	06 December 2018 ~ 27 February 2019

RESULTS

Evaluation of Hypothesis 1

Experiment 1 - Glad Cling Wrap 200ft ==> Glad Aluminum Foil 75sqft 1s

Figure 1: Sales Of Consequent (Aluminum Foil)

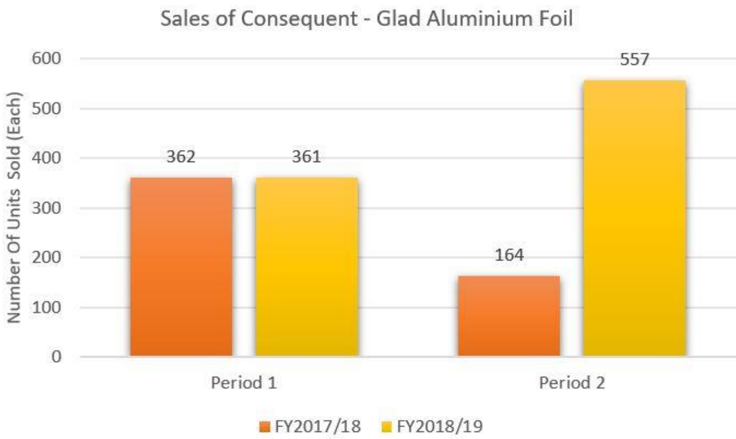


Figure 2: Category Sales For Party Needs

In Figure 1, the sales of consequent Glad Aluminum Foil 75sqft 1s decreased by 1 unit in Period 1 when we compared between FY2017/18 and FY2018/19, accounting for Business As Usual (BAU) negative growth from one year to the next. This BAU growth can be used to offset the expected BAU growth in Period 2 so that we can isolate the sales growth solely due to the implementation. In Period 2, we achieved an increase of 393 units when we compared between FY2017/18 and FY2018/19. With the BAU decline offset of 1 unit, we will achieve a net increase of 394 units. To put the comparison in context with Figure 2, we computed the BAU growth in sales dollar ($) for the Party Needs product category, where the consequent belongs to. The BAU sales dollar increased by $6070 in Period 1 when we compared between FY2017/18 and FY2018/19. In Period 2, we achieved only an increase of $4265 when we compared between FY2017/18 and FY2018/19, which signified that this product category performed worse in Period 2 than in Period 1, unable to achieve the expected BAU sales dollar increase. Such a comparison further confirmed that while the party needs product category performed worse in Period 2 of FY2018/19, the consequent performed better than expected BAU, defying the expected downward trend in its product category.

Experiment 2 - Kim Poh Dish/wash liquid 5kg ==> Budget Kitchen Towel 6roll 60s

Figure 3: Sales Of Consequent (Kitchen Towel)

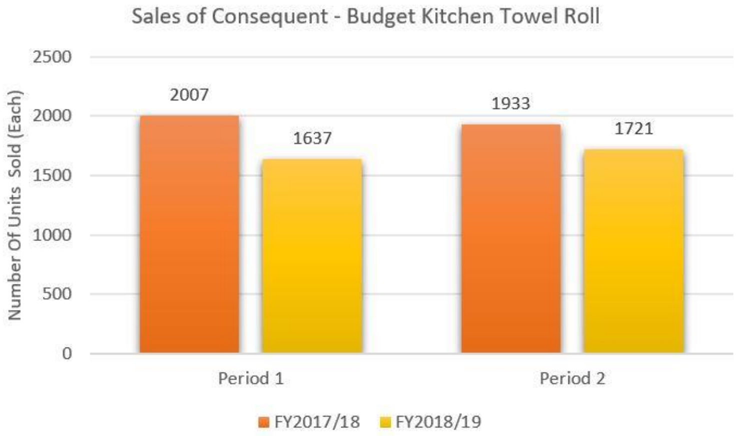
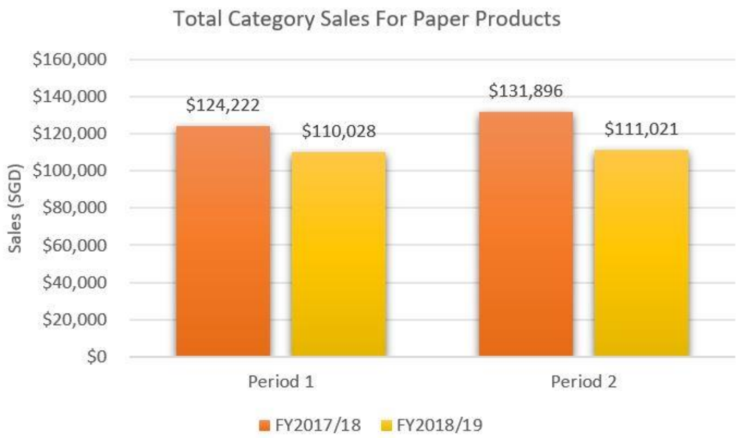


Figure 4: Category Sales For Paper Products



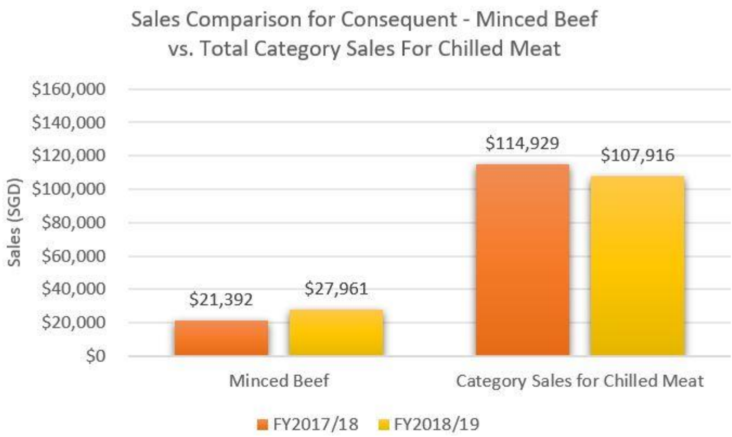
Evaluation of the sales of consequent Budget Kitchen Towel revealed similar positive results after implementation. In Figure 3, the BAU decline in the sales of the consequent in Period 1 between FY2017/18 and FY2018/19 was 370 units. In Period 2, the decline was reduced to 212 units. Offsetting the BAU decline, it was shown that the sales decline was reduced to a lower number due to the implementation, and thus achieved a net positive gain of 158 units. In Figure 4, an investigation of the Paper Products category which the consequent belongs to, showed an overall decline of $14,194 in sales dollar in Period 1 when we compared FY2017/18 with FY2018/19, representing the BAU sales dollar decline. In Period 2, the decline was increased to $20,875 between FY2017/18 and FY2018/19. Despite a higher sales dollar decline in the product category in Period 2, the consequent performed better than the expected BAU decline, slowing the rapid downward trend.

The results from both Experiments 1 and 2 provided insights to hypothesis 1, which suggested that product (antecedent) with a high association is likely to have a positive impact on the sales of the consequent when both products are placed next to each other in the sales display, despite an expected decline in sales and growth in their respective product categories.

Evaluation of Hypothesis 2

<u>Experiment 3 - Cirio Chopped Tomatoes 400g ==> Beef Minced kg</u>

Figure 5: Sales for Consequent (Minced Beef) vs. Total Category Sales For Chilled Meat



The antecedent Cirio Chopped Tomatoes 400g was a new brand of product introduced in the pilot store from the launch week and thereafter. Sales of the consequent Minced Beef was monitored in Period 2 against the same period last year without the presence of the antecedent. In comparison to last year, sales of consequent went up by $6,569, from $21,392 to $27,961 (30.7%) despite a decline in the overall category sales of Swiss Butchery concessionary category by $7013, from $114,929 to $107,916 (-6.1%).

The encouraging results from experiment 3 reinforced hypothesis 2 which states that an introduction of a new product (antecedent) may have a positive impact on the sales of the corresponding consequent based on established association rules from other stores, where traditionally, association rule implementations were often based on rules discovered within the same store. All in all, association rules are potential sources of information for designing in-store displays for synergy across different products, and to drive sales, which is especially important for retailers amidst declining growth.

CONCLUSION

This paper attempts to explore new methods using Market Basket Analysis to capitalise on an established supermarket retailer's database of transactional receipts as a proxy for consumers' purchase patterns. We sought to address two research hypotheses – whether the placement of antecedent near its corresponding consequent would bring about an increase in the sales of the consequent, and whether the introduction of a new product (antecedent) to store based on established association rules from other stores would have a similar positive impact on the sales of the consequent. Three association rules with low Support and high Confidence and/or Lift values were selected, and in-store initiatives were designed and implemented in a pilot store over a period of 12 weeks. From the results, we observed an improvement in sales of consequents after implementation across all three initiatives, when compared to the 12 weeks before implementation, as well as when compared to the same period last year and, despite an overall poorer performance of their corresponding product categories.

ACKNOWLEDGMENT

The authors would like to thank NTUC FairPrice for the data used for the analysis presented in this paper.

REFERENCES

Briggs, F. (2016, July 6). *Shoppers Prefer To Make Final Purchase Decisions In-Store, Despite Researching Online, Study Shows*. Retrieved from Forbes: https://www.forbes.com/sites/fionabriggs/2016/07/06/shoppers-prefer-to-make-final-purchase-decisions-in-store-despite-researching-online-study-shows/#611637a05456

Clement, J. (2020, Aug 27). *Retail e-commerce sales worldwide from 2014 to 2023*. Retrieved from Statista : https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/

Coenen, F., Goulbourne, G., & Leng, P. (2004). Tree Structures for Mining Association Rules. *Data Mining and Knowledge Discovery*, (pp. 25-51).

Geng, L., & Hamilton, H. J. (2006). Interestingness Measures for Data Mining: A Survey. *ACM Computing Surveys.* Association for Computing Machinery.

Griva, A., Bardaki, C., Pramatari, K., & Papakiriakopoulos, D. (2018). Retail business analytics: Customer visit segmentation using market basket data. *Expert Systems with Applications*, (pp. 1-16).

Gupta, S., & Mamtora, R. (2014). A Survey on Association Rule Mining in Market Basket Analysis. *International Journal of Information and Computation Technology* (pp. 409-414). International Research Publications House.

Kopalle, P., Biswas, D., Chintagunta, P. K., Fan, J., Pauwels, K., Ratchford, B. T., & Sills, J. A. (2009). Retailer Pricing and Competitive Effects. *Journal of Retailing*, (pp. 56-70).

Kotsiantis, S., & Kanellopoulos, D. (2006). Association Rules Mining: A Recent Overview. *GESTS International Transactions on Computer Science and Engineering*, (pp. 71-82).

Liu, B., Hsu, W., Chen, S., & Ma, Y. (2000). Analyzing the subjective interestingness of association rules. *IEEE Intelligent Systems and their Applications* (pp. 47-55). IEEE.

Löfgren, M. (2005). Winning at the First and Second Moments of Truth: An Exploratory Study. ” *Managing Service Quality*, (pp. 102-115).

Shankar, V. (2011). *Shopper Marketing.* Marketing Science Institute.

Shankar, V., Inman, J. J., Mantrala, M., Kelley, E., & Rizley, R. (2011). Innovations in Shopper Marketing: Current Insights and. *Journal of Retailing*, (pp. 29-42).

Surjandari, I., & Seruni, A. C. (2010). Design of Product Placement Layout in Retail Shop. *Theory of Computing Systems*, (pp. 43-47).

Yabing, J. (2013). Research of an Improved Apriori Algorithm in Data Mining Association Rules. *International Journal of Computer and Communication Engineering.*