INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES
www.hrmars.com
ISSN: 2222-6990

Market Basket Analysis for Sales Transaction in Shopping Stores

Mohd Noor Azam Nafi, Azni Sharlina Zakaria, Nur Izzati Mohamad Arif, Siti Nurhafizah Mohd Shafie, Nasuhar Ab. Aziz, Omar Kairan

To Link this Article: http://dx.doi.org/10.6007/IJARBSS/v13-i2/15439 DOI:10.6007/IJARBSS/v13-i2/15439

Received: *09* December *2022, **Revised:*** *12* January 2023*, **Accepted:*** *28* January 2023

Published Online: *07* February 2023

In-Text Citation: (Nafi et al., 2023)

To Cite this Article: Nafi, M. N. A., Zakaria, A. S., Arif, N. I. M., Shafie, S. N. M., Aziz, N. A., & Kairan, O. (2023). Market Basket Analysis for Sales Transaction in Shopping Stores. *International Journal of Academic Research in Business and Social Sciences*, *13*(2), 981 – 993.

Vol. 13, No. 2, 2023, Pg. 981 – 993

http://hrmars.com/index.php/pages/detail/IJARBSS | **JOURNAL HOMEPAGE**

INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES
www.hrmars.com
ISSN: 2222-6990

Market Basket Analysis for Sales Transaction in Shopping Stores

Mohd Noor Azam Nafi[1], Azni Sharlina Zakaria[2], Nur Izzati Mohamad Arif[3], Siti Nurhafizah Mohd Shafie[4], Nasuhar Ab. Aziz[5], Omar Kairan[6]
[1,4,5,6]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Kelantan, Kota Bharu Campus, 15150 Kota Bharu, Kelantan, Malaysia, [2]Department of Statistics, Malaysia Block C6, Complex C, Federal Government Administrative Centre, 62514, Putrajaya, Malaysia, [3]M35, Ground Floor, Jalan Melur 1, Taman Melur, Jalan Wang Tepus, 06000 Jitra, Kedah, Malaysia
Corresponding Author's Email: sitinurhafizah@uitm.edu.my

Abstract
Market Basket Analysis (MBA) system is a widely used technique among marketers, especially for undirected data mining analysis. MBA is also known as product association analysis and the outcome of this analysis is called association rules. The outcome can be used to schedule marketing or advertising strategies and design catalogs for different shop layouts. Discovering the pattern from the customer's buying habits in the shopping stores was collected in their buying transaction. This study aims to compare the item purchased by the respondents between Store A and Store B and to find out the most potential products that customers have bought along with a specific category of products. Convenience non-probability sampling was involved with structured questionnaires of items in store was collected to analyze data. Association analysis was used by analyzing the result from support, confidence, and lift. The findings showed that there are 13 interesting rules of association revealed in this study. Moreover, the result also found that most products that were purchased together are tissues, condiments, instant food, cooking oil, meat, biscuits, dry goods, beverages, and cleaning products.
Keywords: Association Rule, Market Basket Analysis, Sales Transaction

Introduction

Nowadays, Market Basket Analysis (MBA) system is a widely used technique among the marketers. MBA also known as product association analysis and the outcome of this analysis is called association rules. In a supermarket, the manager may want to know more about customers' purchasing habits. MBA from Association Rule Mining can be conducted on the store's retail customer transaction data to answer this question. The outcome can be used to schedule marketing or advertising strategies and design for catalogs different shop layouts. Items that are often bought together can be put in close proximity in one of the strategies to further encourage the joint sale of such items. MBA can assist distributors schedule which

items to sell at reduced rates (Raorane et al., 2012). In retailing, most purchases are bought on desire. MBA provides clues as to what a customer might have purchased if the idea came to them. Therefore, MBA can be used to decide where goods are located and promoted in a store. Based on this assumption of independence, empirical researchers typically estimate incidence decisions for household purchases for each product category separately whether or not a household will buy ketchup during a visit to the store is modeled independently of whether or not it will purchase other products in the store (Chib et al., 2002).
Data mining is performed using a special tool that performs data operations defined based on the analysis model. Data mining is the extraction of important or interesting information or patterns from large database data that were unknown but potentially useful. Catering to the tastes of customers can be a challenge in a multi-ethnic nation such as Malaysia, with vast demographic and cultural diversity. The retailers enable to understand customers buying behavior when visiting the store by analyzing both the commercial and social aspects (Nandy, 2018).
Due to implementation of Movement Control Order (MCO) the sales from retail industry decline to curb the Covid-19 virus in this country. Thus, the association hopes that the government will come up with an additional stimulus package that is able to quickly restore the economy and business. Currently, there is no specific strategy for small or medium retail shops that focus on reference of buying behavior from their customers. The technique of MBA is needed in order to survive in business environment that can benefit a business increase better understanding into their customer's purchasing behavior (Kaur and Kang, 2016). The hidden pattern of the customer buying behavior also could be help the retailer in correct decision making.
Supermarket and hypermarket also grew rapidly across the country during the 2000s in Malaysia (The Star Online, 2019). With the current economic situation that has not changed so far, closing under-performing stores is no surprise to them. The closure is not supposed to be worsened than this year, unless a downturn occurred next year (The Star Online, 2019). This can be said that if Malaysian customers turn to mini markets instead of searching for deals in hypermarkets and supermarkets. These issues can be said many distributors face the issue of placing the products in the supermarket that convenience the customers. The decisions are also determined as to which item to stock more, cross-sell and up-sell and shop shelf arrangement. The objective of the study was to estimate the customers buying pattern, possibly the goods in the superstore within the categories that customer would like to pick and best possible combinatory of the products or services which are frequently bought by the customers. Other than that, this study also wants to improve the effectiveness of marketing and sales tactics using the sales transaction from the supermarket. Lastly, MBA wants to give the information to the supermarket to arrange the products in such a way to increase customer satisfaction and profit of the company.

Process Flow of Market Basket Analysis
The process flow for the both descriptive and inferential statistics carried out by using SAS Enterprise Miner (SAS E-Miner). There are three groups of transaction data analyzed which are Store A, Store B and the combination of Store A and Store B called Retail Transaction in Figure 1. Descriptive statistics was analyze using Multiplot node that are used to construct bar chart for Store A, Store B and Retail Transaction. Then, Association node are used to identify the highest category of product purchased by the respondents and the rule for Store A, Store B, and Retail Transaction as shown in Figure 2.

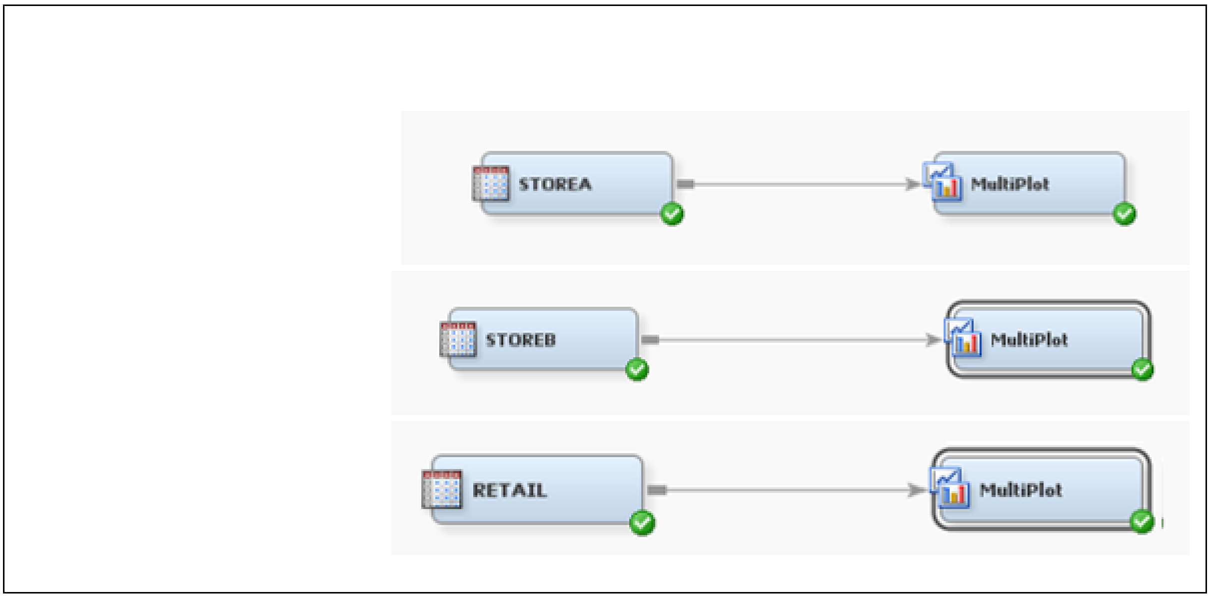

Figure 1. Multiplot for Store A, B and Retail Transaction

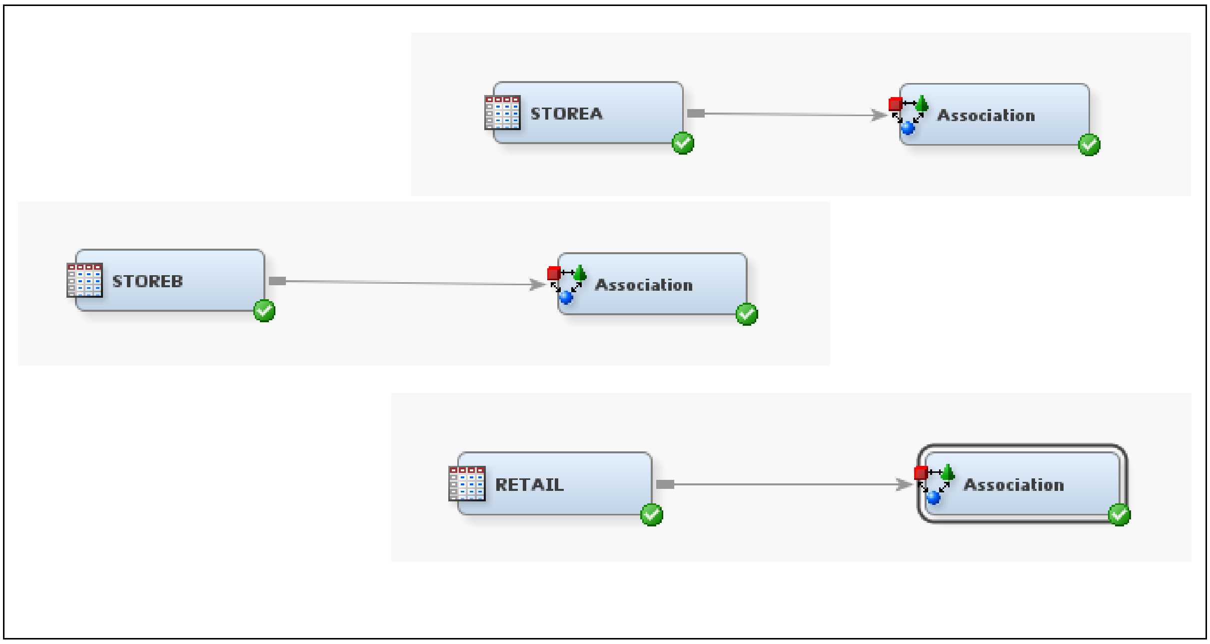

Figure 2. Association node for Store A, B and Retail Transaction

Data and Variables Description

The data was collected through with direct questionnaire to the customers within period whose visited the supermarket in Kota Bharu. Non-probability sampling which is convenience sampling has been applied in order to make comfortable accessible to the researcher. Subjects are chosen simply because they are easy to recruit. This technique is considered easiest, cheapest and least time consuming. There are two supermarkets chosen in this study which are Store A and Store B respectively. Based on the structured questionnaire provided, the list of items was selected in Table 1 that indicate the category of groceries considered in this study. There are 21 categories of groceries are involved.

Vol. 13, No. 2, 2023, E-ISSN: 2222-6990 © 2023 HRMARS

Table 1
List of Groceries Category

NO	CATEGORY	NO	CATEGORY	NO	CATEGORY
1.	Beverages	8.	Frozen Foods	15.	Personal Care
2.	Bread	9.	Meat	16.	Condiments
3.	Biscuits	10.	Vegetables	17.	Snacks
4.	Canned	11.	Fruits	18.	Cereal
5.	Cooking Oil	12.	Tissues	19.	Rice
6.	Dairy	13.	Eggs	20.	Instant Food
7.	Dry Goods	14.	Cleaning Product	21.	Instant Drink

Undirected Data Mining Method

The undirected data mining technique was used to identify the customers buying pattern, possibly the goods in the superstore within the categories. The purpose of the undirected data mining technique was association rule of market basket analysis contains numbers of items in the supermarket. However, for this analysis there is no specific target variables. The variable used is transaction data to identify the associations of the variables.

Association Rule

MBA is widely used by large retailers to discover associations between items. It works by searching for combinations of items that often occur in transactions together. To put it another way, it enables retailers to define association between the items that the customer buy. MBA is also a popular data mining method. Association Rules are commonly used to analyze retail basket or transaction data and are designed to define strong rules found in transaction data using interesting measurements based on the concept of strong rules. The transaction contains list of items. Normally, a single customer purchase is a transaction, and the things that they bought is the items. The statement of the form (Itemset X) => (Itemset Y) is an association rule. The purpose of the analysis is to identify the strength of all the association rules among a set of items. The strength of the association can be measured by the support and confidence of the rule.

Support (S) is the percentage of transactions containing all items in an itemset as in (1). The more frequently the itemset occurs, the higher the support. Therefore, rules with high support are preferred. The support for the rule X => Y is the probability that the two item sets occur together. The support of the rule X => Y is predicted as follows. Note that there is symmetric support. That is, the support of the rule X => Y is the same as the support of the rule Y => X (Linoff & Berry, 2011).

$$S = \frac{\sum(Ta + Tc)}{\sum(T)} \tag{1}$$

Where, $\sum(Ta + Tc)$ is the number of transactions that contains antecedent and consequent, and $\sum(T)$ = the number of transaction

Confidence (C) is the percentage that a transaction containing the items on the left side of the rule will also contain the item on the right side. The higher the confidence, the higher the probability of buying item on the right side. In another word, the higher the return rate can be expected for a given rule. The confidence of an association rule X => Y is the conditional probability of the transaction containing item set Y given that it contains item set X (Linoff & Berry, 2011). The formula of confidence is:

$$C = \frac{\sum(Ta + Tc)}{\sum(Ta)} \tag{2}$$

Where, $\sum(Ta + Tc)$ is the number of transactions that contains antecedent and consequent, and $\sum(Ta)$ = the number of transactions that contains antecedent

It is difficult to interpret the implications (= >) in association rules. High confidence and support do not mean cause and effect. The rule does not have to be interesting. There may not even be a correlation between the two products. The word confidence is not related to the use of statistics. Therefore, there is no repeated sampling interpretation.

Lift is the probability that all items in a rule will occur together divided by the product of the probability that the items on the left and right side will occur as if they were not associated (equation 3). The lift of the rule X => Y is the confidence of the rule divided by the expected confidence, assuming the item sets independent. The lift can be viewed as a general association measure between the two sets of items. Values higher than 1 show positive correlation, values equal to 1 show zero correlation, and values less than 1 show negative correlation. Lift is also a symmetric. The rule of the association can also be interpreted by using Statistics Line Plot and Rule Matrix. The statistics line plot graphs the lift, expected confidence, confidence and support for each of the rules by rule index (Linoff & Berry,2011).

$$L = \frac{S}{S(A) * S(B)} \tag{3}$$

Where, S is Support and $S(A) * S(B)$ is Support A multiply by Support B.

Results and Discussion

Figure 3 shows that there are 1,000 transactions collected from Store A and the total is 6,002 of products. There are 21 types of products that were considered in this study. The highest category purchased by the respondents were Vegetables, Bread and Cleaning Products. The least category purchased by the respondents were Cereal, Rice, and Tissues. Meanwhile, Figure 4 indicate that there are 1,000 transactions collected from Store B and the total is 5,787 of products. There are 21 types of products that were considered too. The highest category of products purchased by the respondents were Bread, Cleaning Products, and Vegetables. The least category purchased by the respondents were Cereal and Rice.

Therefore, by combining all the 2,000 transactions collected from Store A and B, the total is 11,789 of products. Based on Figure 5, the highest time for the data collected are on the Evening followed by Morning and Afternoon. The highest category purchased by the respondents were Bread, Cleaning Products and Vegetables. The least category purchased by the respondents were Cereal, Rice, and Tissues.

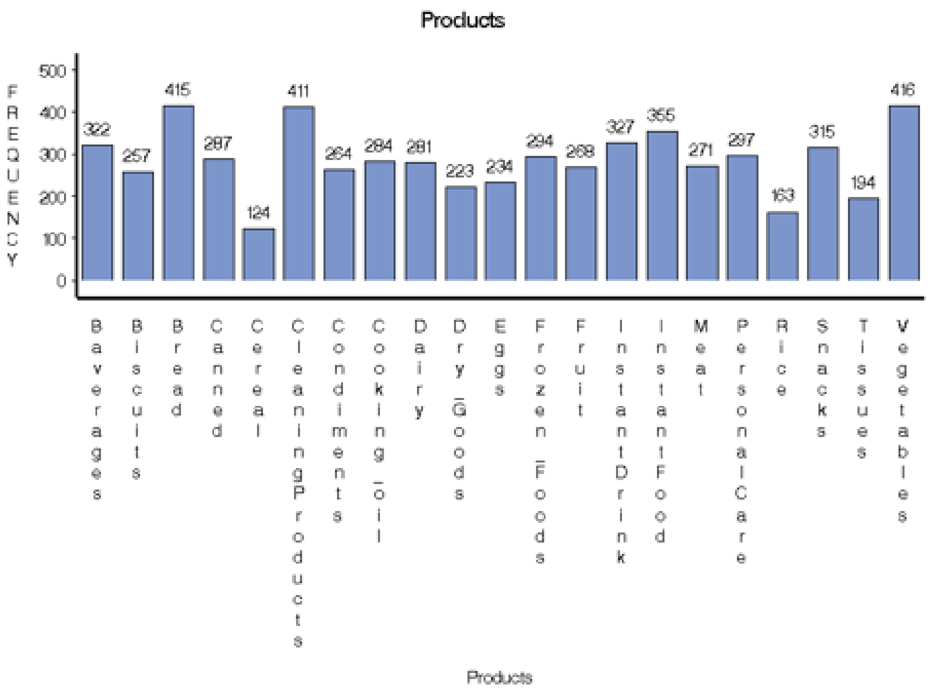

Figure 3. Bar Graph of Quantity of Products Purchased for Store A

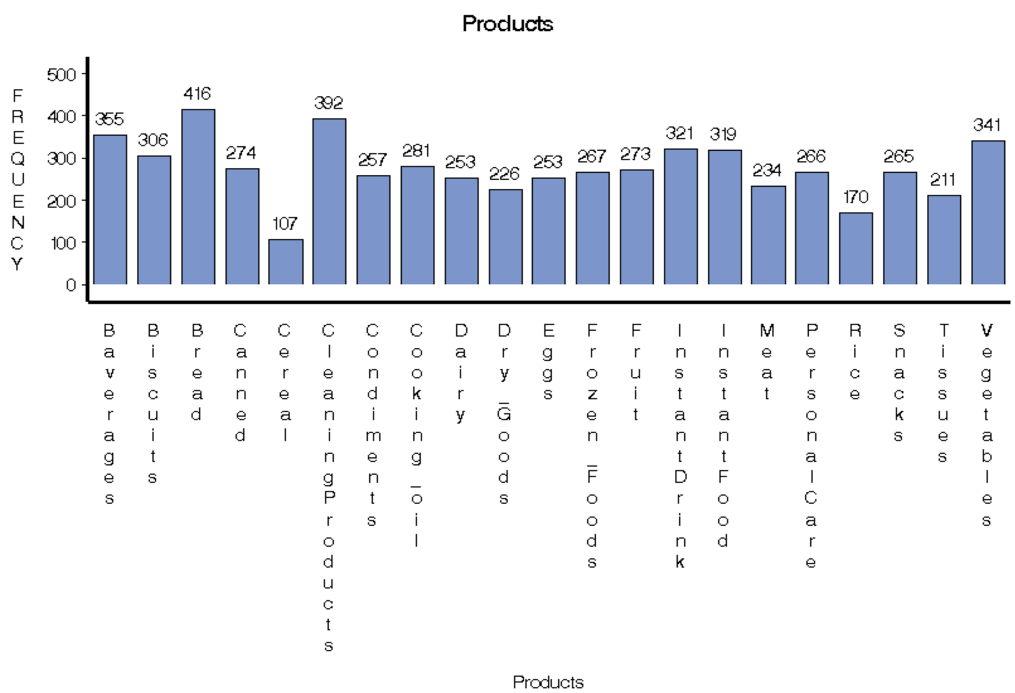

Figure 4. Bar Graph of Quantity of Products Purchased for Store B

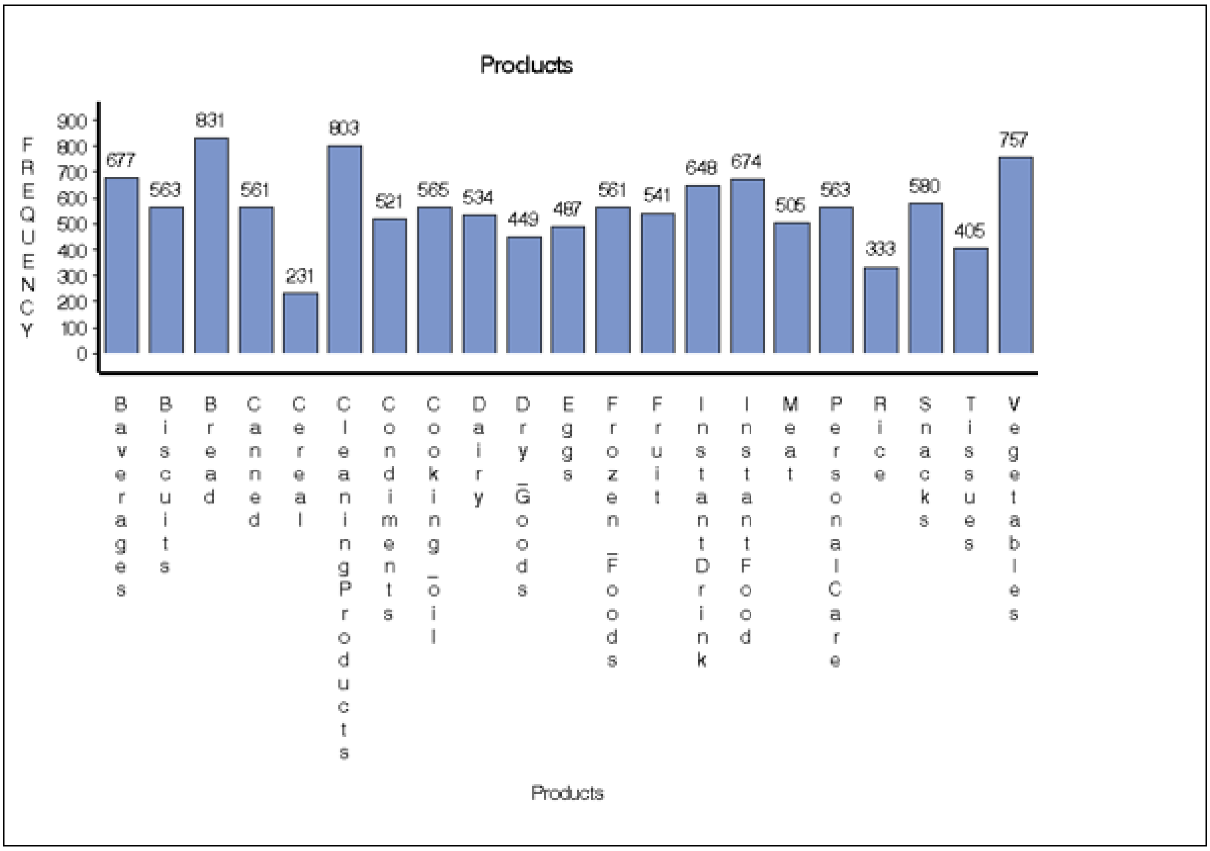

Figure 5. Bar Graph of Quantity of Products Purchased for Retail Transaction

There is a total of 1,000 transaction from Store A were collected and qualified for the analysis and there were 6,002 transactions formed. In the analysis, the quantity of each item that the customers bought is usually not considered. Data processing was done to find pairs of item group which were sold together. Market basket analysis has been applied on each season dataset, the discovered rules have high number, whereby selection of the rules should be considered in order to ensure the result actually shows a strong relationship and has attractiveness to retail company. Since if A => B is equal to B => A, the similar rule was removed and the highest value of confidence was selected. Table 2 shows the rule considered in this study. All of these rules have lift value more than one which indicate all rules generated has strong dependency bond.

Table 2
Rule Description of Store A

No	Antecedent	Consequent	Confidence (%)	Support (%)	Lift
1	Tissues, Condiments	Instant Food, Cooking oil	48.00	2.40	4.62
2	Tissues, Dry Goods	Dairy, Cooking oil	38.89	2.10	4.52
3	Meat, Biscuits	Dry Goods, Beverages	41.51	2.20	4.28
4	Tissues, Eggs	Cooking oil, Biscuits	37.04	2.00	4.16
5	Tissues, Condiments	Cooking oil, Cleaning Products	52.00	2.60	4.09
6	Tissues, Dry Goods	Dairy, Canned	38.89	2.10	4.09
7	Tissues, Canned	Dry Goods, Dairy	31.82	2.10	4.08
8	Dry Goods, Dairy	Frozen Foods, Beverages	41.03	2.20	4.02
9	Tissues, Eggs	Cooking oil, Beverages	44.44	2.40	3.97
10	Eggs, Biscuits	Dry Goods, Cooking oil	35.94	2.30	3.91
11	Eggs, Biscuits	Tissues, Cooking oil	31.25	2.00	3.91
12	Tissues, Canned	Frozen Foods, Dry Goods	31.82	2.10	3.88
13	Dairy, Cooking oil	Tissues, Canned	25.58	2.20	3.88

There are 13 rules considered for Store A. The minimum confidence is 25.58% and the maximum confidence is 52%. While, the highest support is 2.60% and the lowest is 2.00%. For the Rule 1 the rule are Tissues, Condiments => Instant Food, Cooking oil can be said that support is 2.40% of customers purchased Tissues, Condiments, Instant Food, and Cooking Oil. For the confidence 48% of the customers that bought Tissues and Condiments also bought Instant Food and Cooking Oil. The second rule, when buying Tissues and Dry Goods there are 38.89% confidence that the customers will buy Dairy and Cooking Oil and there are 2.10% support that the customers purchased Tissues, Dry Goods, Dairy, and Cooking Oil together. The third rule (Meat, Biscuits => Dry Goods, Beverages), there are 41.51% confidence that the customers who bought Meat and Biscuits they will also purchase Dry Goods and Beverages and with the support explained there are 2.20% customers bought Meat, Biscuits, Dry Goods, and Beverages. The fourth rule Tissues, Eggs => Cooking Oil, Biscuits explained that the customers always buy the four items together with 2% support and there are 37.04% confidence that if the customers who bought Tissues and Eggs also bought Cooking Oil and Biscuits.

The outcome reveals that there is a strong correlation on products such as tissues, condiments, cooking oil, instant food, dry goods, dairy, and eggs. Therefore, the store can considers placing these products together to ease the customer and give them idea what products to purchase together with their visit to the supermarket.

Besides that, at Store B involved total of 1,000 transactions were observed and formed 5,787 transactions of products. The result shows that there is a strong relationship and has attractiveness to retail company. The similar rule was removed and the highest value of confidence was selected. Table 3 shows the rule considered in this study. All these rules have lift value more than one which indicate all rules generated has strong dependency bond.

Table 3
Rule Description of Store B

No	Antecedent	Consequent	Confidence (%)	Support (%)	Lift
1	Tissues, Condiments	Personal Care, Dry Goods	45.45	2.02	6.71
2	Tissues, Eggs	Dry Goods, Dairy	47.76	3.24	6.22
3	Tissues, Frozen Foods	Meat, Biscuits	37.10	2.33	5.92
4	Instant Food, Cereal	Instant Drink, Biscuits	61.90	2.63	5.89
5	Tissues, Condiments	Instant Food, Dry Goods	47.73	2.12	5.83
6	Meat, Biscuits	Fruit, Dry Goods	41.94	2.63	5.76
7	Tissues, Condiments	Instant Drink, Dry Goods	47.73	2.12	5.76
8	Dairy, Cereal	Instant Drink, Biscuits	60.00	2.12	5.71
9	Meat, Biscuits	Frozen Foods, Dry Goods	50.00	3.13	5.68
10	Instant Food, Cereal	Instant Drink, Dairy	52.38	2.22	5.63
11	Vegetables, Tissues	Meat, Dairy	41.43	2.93	5.61
12	Tissues, Condiments	Personal Care, Cooking oil	50.00	2.22	5.56
13	Meat, Dairy	Tissues, Frozen Foods	34.25	2.53	5.46

There are 13 rules considered for Store B. The minimum confidence is 34.25% and the maximum confidence is 60%. Highest support is 3.24% and the lowest is 2.02%. For the Rule 1 the rule are Tissues, Condiments => Personal Care, Dry Goods can be said that support is 2.02% of customers purchased Tissues, Condiments, Personal Care and Dry Goods. For the confidence 45.45% of the customers that bought Tissues and Condiments also bought Personal Care and Dry Goods. The second rule, when buying Tissues and Eggs there are 47.76% confidence that the customers will buy Dry Goods and Dairy and there are 3.24% support that the customers purchased Tissues, Condiments, Dry Goods, and Dairy together. The third rule (Tissues, Frozen Foods => Meat, Biscuits), there are 37.10% confidence that the customers who bought Tissues and Frozen Foods they will also purchase Meat and Biscuits and with the support explained there are 2.33% customers bought Tissues, Frozen Foods, Meat, and Biscuits. The fourth rule Instant Food, Cereal=> Instant Drink, Biscuits explained that the customers always buy the four items together with 2.63% support and there are 61.90% confidence that if the customers who bought Instant Food and Cereal also bought Instant Drink and Biscuits.

As of the observations recorded indicates that there is a strong correlation on products such as dairy, cereal, instant drink, dry goods, meat, biscuits, frozen foods, and instant food. Therefore, the supermarket can considers placing these products together to ease the customer and give them idea what products to purchase together with their visit to the supermarket.

Therefore, there were 11,789 transactions formed based on the combination of two stores. The quantity of each item that the customers bought is usually not considered. Whether the customer buys a dozen of chili sauce or one bottle of tomato sauce would be considered as the same set of condiments and recorded as one for condiments. When the researchers record the observations, not all customers included since only customers that purchase more than one item are considered as data. Transactions of single item are not used for the analysis. Data processing was done to find pairs of item group which were sold together.

Table 3
Rule Description of Store A and B Simplify

No.	Antecedent	Consequent	Confidence (%)	Support (%)	Lift (%)
1	Tissues, Eggs	Dry Goods, Dairy	42.15	2.56	5.44
2	Meat, Biscuits	Fruit, Dry Goods	35.65	2.06	5.14
3	Meat, Biscuits	Frozen Foods, Dry Goods	42.61	2.46	5.01
4	Tissues, Condiments	Instant Food, Cooking oil	50.00	2.36	4.92
5	Eggs, Condiments	Personal Care, Dry Goods	32.06	2.11	4.83
6	Meat, Biscuits	Dry Goods, Dairy	37.39	2.16	4.83
7	Tissues, Condiments	Dry Goods, Cleaning Products	43.62	2.06	4.69
8	Tissues, Dry Goods	Dairy, Cooking oil	40.63	2.61	4.67
9	Meat, Biscuits	Dry Goods, Beverages	45.22	2.61	4.64
10	Tissues, Frozen Foods	Dry Goods, Dairy	35.66	2.31	4.61
11	Tissues, Dry Goods	Eggs, Dairy	39.84	2.56	4.58
12	Tissues, Eggs	Dairy, Cooking oil	39.67	2.41	4.56
13	Tissues, Eggs	Dry Goods, Biscuits	36.36	2.21	4.55

There are 13 rules considered for Retail transaction which are the combination of Store A and Store B. The minimum confidence is 32.06% and the maximum confidence is 50.00%. Highest support is 2.61% and the lowest is 2.06%. For the first rule Tissues, Eggs => Dry Goods, Dairy can be concluded that when buying Dry Goods and Dairy there are 42.15% confidence that the customers will buy Tissues and Eggs and there are 2.56% of support that the customers will purchase the items together. The second rule, when buying Fruits and Dry Goods there are 35.65% confidence that the customers will buy Meat and Biscuits and with 2.06% support that the customers bought the items together. The third rule, there are 42.61% confidence and 2.46% support explained if the customers buy Frozen Foods and Dry Goods, they will also purchase Meat and Beverages. The fourth rule Tissues, Condiments => Instant Food, Cooking Oil explained that the customers always buy the four items together with 2.36% support and customers who bought Tissues and Condiments there are 50.00% confidence that the customer also bought Instant Food and Cooking Oil. From the observations recorded it shows that there is a strong correlation on products such as tissues, condiments, instant food, cooking oil, meat, biscuits, dry goods, beverages, and cleaning products. Therefore, the supermarket can considers placing these products together to ease the customer and give them idea what products to purchase together with their visit to the supermarket.

Conclusions

The result shown most generated rules in the set of data has similarity with others data. However, there are also found unique rules for each data set. There was discovered that there are four categories that have been identified which are store, time, status and products play the important attribute in this study. Then, based on the comparison of the items purchased by the respondents between Store A and Store B indicated that the most purchased items for both stores are similar which are vegetables, bread and cleaning products. There are 13 rules for each store and for the combination of store called retail transaction can lead for retail manager in order to plan the promotional offers, cross-selling strategy and place the products that will be purchased together near to each other which may be helpful to increase the sales and generate more profit. Hence, the result also obtained that the most products that were purchased together are tissues, condiments, instant food, cooking oil, meat, biscuits, dry goods, beverages, and cleaning products.

References

Abdulsalam, S. O., Adewole, K. S., Akintola, A. G., & Hambali, M. A. (2014). Data mining in market basket transaction: An association rule mining approach. *International Journal of Applied Information Systems (IJAIS)*, *7*(10), 15-20.

Andrews, R. L., & Currim, I. S. (2002). Identifying segments with identical choice behaviors across product categories: An intercategory logit mixture model. *International Journal of Research in Marketing*, *19*(1), 65-79.

Annie, L. C. M., & Kumar, A. D. (2012). Market basket analysis for a supermarket based on frequent itemset mining. *International Journal of Computer Science Issues* (IJCSI), 9(5), 257. https://doi.org/10.1016/S0167-8116(02)00048-4.

Ayu, S. K., Surjandari, I., & Zulkarnain, Z. (2018, July). Mining association rules in seasonal transaction data. In *2018 5th International Conference on Information Science and Control Engineering (ICISCE)* (pp. 321-325). IEEE. https://doi.org/10.1016/10.1109/ICISCE.2018.00074

Cheung, D. W. L., Fu, A. W. C., & Han, J. (1994). Knowledge discovery in databases: A rule-based attribute-oriented approach. *In International Symposium on Methodologies for Intelligent Systems* (pp. 164-173). Springer, Berlin, Heidelberg. Available at https://link.springer.com/chapter/10.1007/3-540-58495-1_17

Chib, S., Seetharaman, P. B., and Strijnev, A. (2002), "Analysis of multi-category purchase incidence decisions using IRI market basket data", *Advances in Econometrics* (*Advances in Econometrics, Vol. 16*), Emerald Group Publishing Limited, Bingley, pp. 57-92. https://doi.org/10.1016/S0731-9053(02)16004-X

Zadeh, H. A., Schiller, S., & Duffy, K. (2016). Teaching Analytics: A Demonstration of Association Discovery with SAS Enterprise Miner. Available at https://web.archive.org/web/20220803134744id_/https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1264&context=amcis2016

INN, T. (2017). No room for large hypermarkets now. [online] The Star Online. Available at: https://www.thestar.com.my/news/nation/2017/04/30/no-room-for-large-hypermarkets-now-tesco-ceo-they-are-no-longer-feasible

Kaur, M., & Kang, S. (2016). Market Basket Analysis: Identify the changing trends of market data using association rule mining. Procedia computer science, 85, 78-85.

Khattak, A. M., Khan, A. M., Lee, S., & Lee, Y. K. (2010). Analyzing association rule mining and clustering on sales day data with xlminer and weka. *International Journal of Database Theory and Application*, *3*(1), 13-22.

Kim, H. K., Kim, J. K., & Chen, Q. Y. (2012). A product network analysis for extending the market basket analysis. *Expert Systems with Applications*, *39*(8), 7403-7410. https://doi.org /10.1016/j.eswa.2012.01.066

Kurniawan, F., Umayah, B., Hammad, J., Nugroho, S. M. S., & Hariadi, M. (2018). Market Basket Analysis to identify customer behaviours by way of transaction data. *Knowledge Engineering and Data Science*, *1*(1), 20. http://doi.org/10.17977/um018v1i12018p20-25

Linoff, G. S., & Berry, M. J. (2011). *Data mining techniques: for marketing, sales, and customer relationship management*. John Wiley & Sons.

Raorane, A. A., Kulkarni, R. V., & Jitkar, B. D. (2012). Association rule–extracting knowledge using market basket analysis. *Research Journal of Recent Sciences ISSN*, *2277*, 2502.

Setiawan, A., Budhi, G. S., Setiabudi, D. H., & Djunaidy, R. (2017). Data mining applications for sales information system using market basket analysis on stationery company. In *2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT)* (pp. 337-340). IEEE. https://doi.org/10.1109/IC SIIT.2017.39

The Star Online. (2019). Supermarkets and hypermarkets in Malaysia going through consolidation. [online] The Star Online. Available at: https://www.thestar.com.my/ business/business-news/2019/09/17/supermarkets-and-hypermarkets-in-malaysia-going-through-consolidation

Videla-Cavieres, I. F., & Rios, S. A. (2014). Extending market basket analysis with graph mining techniques: A real case. *Expert Systems with Applications*, *41*(4), 1928-1936. https://doi.org/10.1016/j.eswa.2013.08.088